Lineare Algebra und Analytische Geometrie I & II

Dr. Jana Bielagk
Humboldt-Universität zu Berlin

Wintersemester 2022/23 & Sommersemester 2023

Inhaltsverzeichnis

Einleitung

Warum die komplizierte Schreibweise? Einerseits ist Sprache selten eindeutig zu verstehen. Andererseits gehen bei langen Texten Informationen unter und die üblichen stilistischen Anforderungen an Texte widersprechen den Anforderungen der Mathematik. (Es wird immer wieder etwas geschlussfolgert und Bedingungen werden überprüft - häufige Wiederholungen und langweiliger Stil sind vorprogrammiert!)

Warum die große Allgemeinheit? Ein wesentlicher Aspekt der Mathematik besteht darin, nicht nur ein spezielles Problem lösen zu können, sondern ganze Klassen von Problemen. Statt also Vektoraddition erst im $\mathbb{R}^2$, dann im $\mathbb{R}^3$ und dann im $\mathbb{R}^4$ einzuführen, wird gleich die Addition im $\mathbb{R}^n$ für beliebiges $n \in \mathbb{N}$ eingeführt.

Wie viel muss man auswendig lernen? Die Bezeichnungen von Objekten folgen gewissen Regeln, z.B. werden in der Regel Matrizen mit Großbuchstaben bezeichnet, Winkel mit kleinen griechischen Buchstaben und Funktionen mit kleinen oder großen Buchstaben - je nach Situation. Es ist sinnvoll, sich an solche Konventionen zu halten. Abgesehen davon sind die Bezeichnungen für mathematische Objekte in Sätzen und Definitionen beliebig austauschbar. Gelernt werden sollte also in der Regel nicht „Zeichen-für-Zeichen“, sondern optimalerweise der Inhalt von Sätzen und Definitionen. Das geht am besten durch das vielfache Anwenden und nur als Ergänzung sind Karteikarten o.ä. Lernsysteme sinnvoll.

Was erwartet mich bei LAAG? Die analytische Geometrie beschäftigt sich mit der rechnerischen Beschreibung geometrischer Objekte und der Lösung geometrischer Probleme mit algebraischen Mitteln. Die lineare Algebra befasst sich in erster Linie mit der Lösung von linearen Gleichungen bzw. linearen Gleichungssystemen. Wir werden im Laufe des Kurses sehen, dass beide Gebiete Überschneidungen haben und sich wunderbar ergänzen.

Kapitel 1

Grundlagen der Logik und Mengenlehre

Ziele

- □ korrekte logische Schlussfolgerungen formulieren und falsche Schlussfolgerungen erkennen
- □ die grundlegenden mathematischen Beweismethoden (insb. Kontraposition, Widerspruchsbeweis) kennen und an einfachen Beispielen anwenden
- □ Aussagen mit und ohne Quantoren korrekt negieren
- □ mit Wahrheitstafeln grundlegende Regeln der Aussagenlogik begründen
- □ die verschiedenen Schreibweisen von Mengen kennen und interpretieren
- □ Mengen verknüpfen und Verknüpfungen von Mengen korrekt interpretieren
- □ Äquivalenz- von anderen Umformungen unterscheiden

1.1 Grundlagen der Logik

Die Mathematik ist dem Namen nach die *Kunst des Lernens*. Zwar gibt es keine allgemeingültige Definition der Mathematik selbst, jedoch ist ein fundamentaler Aspekt der Mathematik die Verwendung logischer Argumente um den Wahrheitsgehalt von Aussagen zu überprüfen. Das Beweisen ist somit ein elementarer Bestandteil aller Bereiche der Mathematik.

Beispiel 1.1. *Sind folgende Deduktionen korrekt?*

(i) *Elefanten sind Säugetiere. Dumbo ist ein Elefant. Also ist Dumbo ein Säugetier.*

(ii) *Dozenten stehen an der Tafel. Thorsten steht an der Tafel. Also ist Thorsten Dozent.*

(iii) *Berlin liegt in Deutschland, Deutschland liegt in Europa, also liegt Berlin in Europa.*

(iv) *Einige Pflanzen sind Fleischfresser. Einige Fleischfresser sind Katzen. Also sind einige Pflanzen Katzen.*

Um richtige und falsche Schlussweisen besser unterscheiden und selbst korrekte Beweise führen zu können, befassen wir uns zunächst mit elementarer Logik.

1.1.1 Aussagen und Logische Verknüpfungen

Definition 1.2

Unter einer Aussage verstehen wir die Beschreibung eines Sachverhaltes mit eindeutigem Wahrheitswert (wahr oder falsch).

Beispiel 1.3.

- *„3 ist eine Primzahl“ und „$1 = 1$“ sind wahre Aussagen.*
- *„4 ist eine Primzahl“ und „$1 = 2$“ sind falsche Aussagen.*
- *„Was bedeutet das? “ ist zwar ein sinnvoller Satz, jedoch keine Aussage. Ebenso ist „$1 + 3$“ keine Aussage. „$x - 3 = 4$“ ist ebenfalls keine Aussage (sondern eine* Aussageform*), „wenn $x = 7$ ist, dann ist $x - 3 = 4$“ hingegen schon.*

Definition 1.4

Ist A eine Aussage, so bezeichnet $\neg A$ (sprich „nicht A“) ihre Negation. Die Negation $\neg A$ ist genau dann wahr, wenn A falsch ist.

Man kann verschiedene Aussagen durch sogenannte *Junktoren* verknüpfen.

Definition 1.5

Konjunktion Die Konjunktion $A \wedge B$ (sprich „A und B“) zweier Aussagen A und B ist genau dann wahr, wenn beide Aussagen wahr sind.

Disjunktion Die Disjunktion $A \vee B$ (sprich „A oder B“) zweier Aussagen A und B ist genau dann falsch, wenn beide Aussagen falsch sind.

Hervorzuheben ist hierbei, dass es sich bei $A \vee B$ nicht um das ausschließende „entweder A oder B“ handelt, d.h. es dürfen durchaus sowohl A als auch B wahr sein.

Beispiel 1.6.

Die Sonne scheint	***und***	*die Kinder spielen im Garten.*
A	$\wedge$	B
Die Sonne scheint	***oder***	*die Kinder spielen im Garten.*
A	$\vee$	B

Definition 1.7

Implikation Die Aussage $A \Rightarrow B$ ist genau dann falsch, wenn A wahr und B falsch ist.

Äquivalenz Zwei Aussagen A und B sind genau dann logisch äquivalent ($A \iff B$), wenn sie entweder beide wahr oder beide falsch sind.

Hat man eine wahre Aussage $A \Rightarrow B$ vorliegen, so sagt man, A sei *hinreichend* für B bzw. B sei *notwendig* für A. Gilt $A \iff B$, so sagt man daher auch, A sei *hinreichend und notwendig* für B.

Abhängig davon, ob Aussagen A und B wahr (w) oder falsch (f) sind, kann man anhand der Definitionen den Wahrheitsgehalt der verknüpften Aussagen bestimmen:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \iff B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Satz 1.8: De Morgan'sche Regeln der Aussagenlogik

Seien A und B beliebige Aussagen. Dann gelten folgende Regeln:

$$\begin{aligned} \neg(A \wedge B) &\iff \neg A \vee \neg B, \\ \neg(A \vee B) &\iff \neg A \wedge \neg B. \end{aligned}$$

Bevor wir uns ansehen, wie eine solche Aussage bewiesen werden kann, wollen wir uns anhand eines Beispiels die Aussage dieser Regeln verdeutlichen.

Beispiel 1.9. *Betrachten wir zwei Aussagen und die daraus entstehenden zusammengesetzten Aussagen:*

$$\begin{aligned} A &= \text{„Die S-Bahn kommt pünktlich."} \\ B &= \text{„Der Bus kommt pünktlich."} \\ A \wedge B &= \text{„Die S-Bahn und der Bus sind pünktlich."} \\ \neg(A \wedge B) &= \text{„S-Bahn und Bus sind nicht beide pünktlich."} \\ \neg A &= \text{„Die S-Bahn kommt nicht pünktlich."} \\ \neg B &= \text{„Der Bus kommt nicht pünktlich."} \\ \neg A \vee \neg B &= \text{„Die S-Bahn oder der Bus oder beide kommen nicht pünktlich."} \end{aligned}$$

Sowohl $\neg(A \wedge B)$ als auch $\neg A \vee \neg B$ sagen also aus, dass mindestens eines der beiden Verkehrsmittel unpünktlich kommt.

Beweis von Satz 1.8.
Wir zeigen an dieser Stelle nur die erste der beiden Aussagen. Die Beweismethode ist für die zweite Aussage identisch.
Um zu zeigen, dass zwei Aussagen identisch sind, müssen wir mit Hilfe von Wahrheitstafeln alle möglichen Fälle untersuchen.

A	B	$A \wedge B$	$\neg(A \wedge B)$	$\neg A$	$\neg B$	$\neg A \vee \neg B$
w	w	w	f	f	f	f
w	f	f	w	f	w	w
f	w	f	w	w	f	w
f	f	f	w	w	w	w

Da in allen Fällen $\neg(A \wedge B)$ und $\neg A \vee \neg B$ die gleichen Wahrheitswerte haben, sind beide Aussagen äquivalent. Man könnte dies auch durch eine weitere Spalte mit $\neg(A \wedge B) \iff \neg A \vee \neg B$, in der alle Wahrheitswerte *w* sind, hervorheben. $\square$

Lemma 1.10

Es gelten folgende Rechenregeln für Konjunktionen und Disjunktionen und beliebige Aussagen A, B und C:

$$\begin{aligned} A \wedge B &\iff B \wedge A & A \vee B &\iff B \vee A \\ A \wedge (B \wedge C) &\iff (A \wedge B) \wedge C & A \vee (B \vee C) &\iff (A \vee B) \vee C \\ A \wedge (B \vee C) &\iff (A \wedge B) \vee (A \wedge C) & A \vee (B \wedge C) &\iff (A \vee B) \wedge (A \vee C) \end{aligned}$$

1.1.2 Grundlegende Beweistechniken

Satz 1.11

Seien A, B und C beliebige Aussagen. Dann gelten folgende Regeln:

1. $A \iff B$ ist äquivalent zu $(A \Rightarrow B) \wedge (B \Rightarrow A)$;
2. $A \Rightarrow B$ ist äquivalent zu $\neg B \Rightarrow \neg A$ (Kontraposition);
3. $A \Rightarrow B$ ist äquivalent zu $\neg(A \wedge \neg B)$ (Widerspruchsbeweis);
4. $A \wedge (A \Rightarrow B)$ impliziert B (Abtrennregel);
5. $(A \Rightarrow B) \wedge (B \Rightarrow C)$ impliziert $A \Rightarrow C$ (Kettenregel).

Die Aussagen dieses Satzes sind die Grundpfeiler mathematischer Beweise. Es gibt noch weitere Beweismethoden (vollständige Induktion, Beweis durch Fallunterscheidung, Diagonalverfahren, Schubfachprinzip, ...), die jedoch besser an entsprechender Stelle dieser oder anderer Vorlesungen vorgestellt werden.

Bemerkung 1.12. *Insbesondere ist im Gegensatz zur Kontraposition $A \Rightarrow B$ nicht äquivalent zu $B \Rightarrow A$. Ist eine Aussage vom Typ „aus A folgt B“ zu zeigen, d.h. eine Aussage vom Typ $A \Rightarrow B$, so nennt man A* Prämisse *(oder* Voraussetzung*) und B* Konklusion *(oder* Schlussfolgerung*). Bei Beweisaufgaben ist es sinnvoll, zunächst Voraussetzungen und Schlussfolgerungen zu identifizieren um Fehlschlüsse zu vermeiden.*

Beweis von Satz 1.11.
Beweis von Aussage 1:

A	B	$A \iff B$	$A \Rightarrow B$	$B \Rightarrow A$	$(A \Rightarrow B) \wedge (B \Rightarrow A)$
w	w	w	w	w	w
w	f	f	f	w	f
f	w	f	w	f	f
f	f	w	w	w	w

Beweis von Aussage 2:

A	B	$A \Rightarrow B$	$\neg A$	$\neg B$	$\neg B \Rightarrow \neg A$
w	w	w	f	f	w
w	f	f	f	w	f
f	w	w	w	f	w
f	f	w	w	w	w

Beweis von Aussage 3:

A	B	$A \Rightarrow B$	$\neg B$	$A \wedge \neg B$	$\neg(A \wedge \neg B)$
w	w	w	f	f	w
w	f	f	w	w	f
f	w	w	f	f	w
f	f	w	w	f	w

Beweis von Aussage 4:
In diesem Fall ist keine Äquivalenz zwischen Aussagen zu zeigen, sondern es muss nachgewiesen werden, dass die Aussage $A \wedge (A \Rightarrow B) \Rightarrow B$ immer wahr ist.

A	B	$A \Rightarrow B$	$A \wedge (A \Rightarrow B)$	$A \wedge (A \Rightarrow B) \Rightarrow B$
w	w	w	w	w
w	f	f	f	w
f	w	w	f	w
f	f	w	f	w

Beweis von Aussage 5:

A	B	C	$A \Rightarrow B$	$B \Rightarrow C$	$(A \Rightarrow B) \wedge$ $(B \Rightarrow C)$	$A \Rightarrow C$	$((A \Rightarrow B) \wedge (B \Rightarrow C))$ $\Rightarrow (A \Rightarrow C)$
w	w	w	w	w	w	w	w
w	f	w	f	w	f	w	w
f	w	w	w	w	w	w	w
f	f	w	w	w	w	w	w
w	w	f	w	f	f	f	w
w	f	f	f	w	f	f	w
f	w	f	w	f	f	w	w
f	f	f	w	w	w	w	w

□

Bemerkung 1.13.
Betrachten wir die Aussageformen A „$x+1$ ist eine gerade Zahl“ und B „x ist eine ungerade Zahl“.

1. *Die Äquivalenz von Aussagen A und B kann gezeigt werden, indem sowohl die „Hinrichtung“ $A \Rightarrow B$ als auch die „Rückrichtung“ $A \Leftarrow B$ (bzw. $B \Rightarrow A$) bewiesen werden. In diesem Fall zeigt man also:*
   - *$A \Rightarrow B$: „Wenn $x+1$ gerade ist, dann ist x ungerade.“*
   - *$B \Rightarrow A$: „Wenn x ungerade ist, dann ist $x+1$ gerade.“*
2. *Äquivalent zu $A \Rightarrow B$ ist die Aussage $\neg B \Rightarrow \neg A$, d.h. „Wenn x nicht ungerade ist, dann ist $x+1$ nicht gerade.“*
3. *Ebenfalls äquivalent zu $A \Rightarrow B$ ist die Aussage $\neg(A \wedge \neg B)$, d.h. „Es gilt nicht gleichzeitig, dass $x+1$ gerade ist und x nicht ungerade ist.“*
4. *Die vierte Aussage sagt uns: Wenn Aussage A wahr ist und aus Aussage A Aussage B folgt, dann ist Aussage B wahr. Lässt man die Überprüfung der Wahrheit von A weg, so kann man alles Mögliche schlussfolgern. Beispielsweise ist die Aussage „wenn 2 ungerade ist, dann ist 3 gerade“ wahr. Sinnvoll angewandt wird eine Implikation $A \Rightarrow B$ also nur dann, wenn wir mit einer wahren Aussage starten, z.B. „4 ist gerade und der Vorgänger einer geraden Zahl ist ungerade, folglich ist 3 ungerade.“*

5. *Die fünfte Aussage erlaubt es uns, eine Implikation durch Aneinanderreihung mehrerer bekannter Implikationen zu zeigen. Zudem können wir diese Methode nutzen um die Äquivalenz mehrerer Aussagen zu zeigen. Zum Beispiel folgt aus $A \Rightarrow B \Rightarrow C \Rightarrow A$, dass die Aussagen A, B und C äquivalent sind. Der Unterschied zwischen Aussagen (iii) und (iv) im Einführungsbeispiel 1.1, die beide scheinbar diese Form haben, werden wir im nächsten Abschnitt erarbeiten.*

Folgerung 1.14: Die Beweismethoden Kontraposition & Widerspruchsbeweis

Wenn $A \Rightarrow B$ gezeigt werden soll, d.h. aus der Gültigkeit von A folgt die Gültigkeit von B, so kann man an Stelle eines direkten Beweisen eine der folgenden Beweisstrategien aus Satz 1.11 wählen:

Kontraposition: Man nimmt an, dass $\neg B$ wahr ist und zeigt, dass dies die Gültigkeit von $\neg A$ impliziert.

Widerspruchsbeweis: Mann nimmt an, dass A und $\neg B$ wahr sind und kommt durch logische Argumente zu einem Widerspruch.

Beispiel 1.15. *Als Beispiel wollen wir zeigen, dass aus $x = y$ folgt, dass $x^2 = y^2$ ist.*

Indirekter Beweis: *Nehmen wir an, dass $x^2 \neq y^2$ ist. Dann ist $0 \neq x^2 - y^2 = (x-y)(x+y)$. Da ein Produkt genau dann Null ist, wenn einer der Faktoren Null ist, impliziert das, dass $x - y \neq 0$ und $x + y \neq 0$ ist, d.h. insbesondere ist $x \neq y$.* □

Widerspruchsbeweis: *Nehmen wir an, dass $x = y$ und $x^2 \neq y^2$ gelten. Dann ist $0 \neq x^2 - y^2 = (x-y)(x+y) = 0 \cdot (x+y) = 0$, d.h. $0 \neq 0$. Dies ist ein Widerspruch. Folglich kann die Annahme nicht stimmen.* □

1.2 Mengen

> Unter einer 'Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die 'Elemente' von M genannt werden) zu einem Ganzen.
>
> Georg Cantor (1895)

Beispiel 1.16 (Beispiele für Mengen)**.**

- $\{0, 1\}$ *ist die endliche Menge, die nur die Zahlen 0 und 1 enthält;*
- $\{\square, ⬠, ⬡\}$ *ist eine Menge, die ein Vier-, ein Fünf- und ein Sechseck enthält;*
- $\mathbb{N} = \{1, 2, 3, \ldots\}$ *ist die Menge der natürlichen Zahlen;*
- $\mathbb{R} = \{x \mid x \text{ reelle Zahl}\}$ *bezeichnet die Menge der reellen Zahlen.*

Ob Elemente zu einer Menge gehören oder nicht[1], schreiben wir wie folgt verkürzt:

[1]Eine grundlegende Forderung an Mengen ist in diesem Sinne, dass feststellbar sein muss, ob Elemente zu einer Menge gehören oder nicht. Aus diesem Grund spricht man auch nicht von *der Menge aller Mengen*, da es sonst zu Widersprüchen kommen kann. In diesem Zusammenhang spricht man stattdessen von *Klassen*. Andernfalls müsste man sich mit dem Paradoxon befassen, welche Menge als die Menge aller Mengen, die sich nicht nicht selbst enthalten, beschrieben werden kann. Etwas anschaulicher: Definieren wir den Barbier als denjenigen, der genau all die Menschen rasiert, die sich nicht selbst rasieren. Und wer rasiert den Barbier?

Notation

Ist x Element der Menge A, so schreiben wir $x \in A$; sonst schreiben wir $x \notin A$.

Aus den Beispielen sehen wir, dass es unterschiedliche Arten gibt, Mengen zu notieren.

1. Eine endliche Menge wird üblicherweise durch Aufzählung ihrer Elemente dargestellt, z.B. $A = \{1, 2, 3\}$.
2. Auch unendliche Mengen können durch Aufzählung notiert werden, sofern sie durch logische Fortsetzung eindeutig identifiziert werden können:

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

ist die Menge der ganzen Zahlen.
3. Ist eine Aufzählung der Elemente einer Menge nicht möglich oder erscheint dies als zu kompliziert, so kann man sie durch Eigenschaften ihrer Elemente charakterisieren, zum Beispiel ist das abgeschlossene Intervall von 0 bis 1

$$[0, 1] = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}.$$

Bei allen Notationen mit Mengenklammern (geschweiften Klammern, d.h. $\{\dots\}$) spielt die Reihenfolge der Elemente für die Menge keine Rolle. Insbesondere bezeichnen $\{0, 1\}$ und $\{1, 0\}$ die gleiche Menge. Es ist eine übliche Konvention, dass jedes Element einer Menge nur einmal aufgezählt wird. Wird eine Menge wie in Punkt 1 (neu) definiert, so schreibt man häufig

$$A := \{1, 2, 3\}$$

um hervorzuheben, dass es nicht einfach um die Gleichheit zweier Mengen geht, sondern dass die links stehende durch den Ausdruck auf der rechten Seite definiert wird.

Beispiel 1.17 (Weitere Beispiele für Mengen).

- $\mathbb{Q}$ *bezeichnet die Menge der rationalen Zahlen;*
- $(0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$ *ist das offene Intervall von 0 bis 1;*
- $\emptyset$ *bzw.* $\{\}$ *ist die* leere Menge*;*
- $2\mathbb{Z} = \{z \in \mathbb{Z} \mid z \text{ ist durch 2 teilbar}\}$ *ist die Menge der geraden Zahlen.*

Definition 1.18: Teilmenge

Eine Menge A heißt *Teilmenge* einer Menge B, falls aus $x \in A$ auch $x \in B$ folgt.

Notation

- $A \subset B$ oder $A \subsetneq B$ für echte Teilmengen, d.h. A ist eine Teilmenge von B und es existiert mindestens ein $x \in B$ mit $x \notin A$;
- $A \subset B$ oder $A \subseteq B$ für einfache Teilmenge, d.h. A ist eine Teilmenge von B, möglicherweise ist sogar $A = B$.

Kurzschreibweise der Definition:

$$A \subseteq B \quad \Longleftrightarrow \quad (x \in A \Rightarrow x \in B)$$

Beispiel 1.19 (Beispiele für Teilmengen und solche, die es nicht sind)**.**

- $\{1, 2, 3\} \subseteq \mathbb{N}$;
- $-1 \notin \mathbb{N}$, *also ist* $\{-1, 1\} \not\subseteq \mathbb{N}$;

Satz 1.20

Für jede Menge M gilt sowohl $M \subseteq M$ als auch $\emptyset \subseteq M$.

Beweis. Sei M eine beliebige Menge. Beweisen wir zunächst die erste Aussage durch direkte Anwendung der Definition.

$$x \in M \quad \Longrightarrow \quad x \in M$$

ist gültig für jede Menge M. Daraus folgt, dass $M \subseteq M$ ist.
Die zweite Aussage beweisen wir durch Annahme des Gegenteils und Herleitung eines Widerspruchs: Nehmen wir an, es existiere eine Menge M mit der Eigenschaft dass $\emptyset \not\subseteq M$. Dann gilt (mit den Regeln zur Negation einer Aussage), dass ein $x \in \emptyset$ existiert mit $x \notin M$. Da die leere Menge allerdings keine Elemente enthält, kann sie insbesondere kein Element mit dieser Eigenschaft enthalten, d.h. wir haben einen Widerspruch. Es kann also keine Menge M geben mit der Eigenschaft $\emptyset \not\subseteq M$. Im Umkehrschluss heißt das also, dass jede Menge M die leere Menge enthält, was gerade die zu behauptende Aussage ist. □

Beispiel 1.21. *Ob ein Objekt Element oder Teilmenge einer Menge ist, ist unter Umständen vom Kontext abhängig. Betrachten wir dazu folgende Menge von Mengen, welche man daher auch als* Mengensystem *bezeichnet:*

$$M := \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\},$$

d.h. die Elemente *von M sind die natürlichen, die ganzen, die rationalen und die reellen Zahlen. Man schreibt damit*

$$\mathbb{N} \in M, \quad \textit{aber} \quad \mathbb{N} \subseteq \mathbb{Z}.$$

Definition 1.22: Potenzmenge

Die *Potenzmenge* $\mathcal{P}(M)$ einer Menge M ist die Menge der Teilmengen von M, d.h.

$$\mathcal{P}(M) := \{A \mid A \subseteq M\}.$$

Bemerkung 1.23. *Man kann zeigen, dass wenn M genau n Elemente enthält, ihre Potenzmenge $\mathcal{P}(M)$ genau 2^n Elemente enthält. (Beispiel 1.41) Aus diesem Grund wird in der Literatur auch die Notation 2^M für die Potenzmenge verwendet.*

Definition 1.24: Gleichheit von Mengen

Zwei Mengen A und B heißen *gleich*, in Zeichen $A = B$, wenn A Teilmenge von B und B Teilmenge von A ist.

Kurzschreibweise der Definition:

$$A = B \quad \Longleftrightarrow \quad ((A \subseteq B) \wedge (B \subseteq A)) \quad \Longleftrightarrow \quad ((x \in A) \iff (x \in B))$$

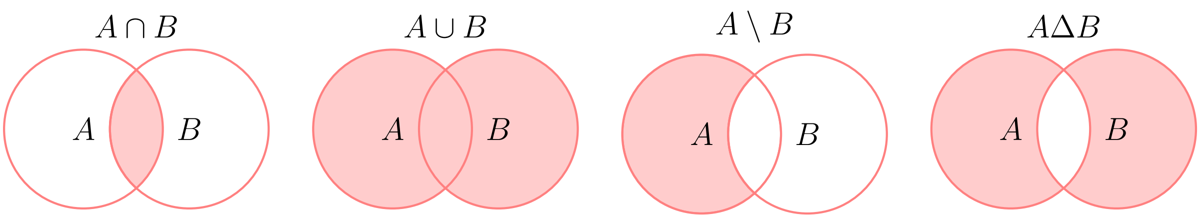
Beispiel 1.25. *Wir haben bereits behauptet, dass die Mengen $A = \{0, 1\}$ und $B = \{1, 0\}$ gleich sind. Überzeugen wir uns, dass dies auch nach der Definition der Mengengleichheit der Fall ist:*

- *Einerseits kann man über die gegenseitige Inklusion argumentieren: Für alle $x \in A$ (nämlich 0 und 1) gilt auch $x \in B$, d.h. $A \subseteq B$. Mit dem gleichen Argument gilt auch die Umkehrung.*
- *Andererseits kann man über die Elemente von A und B und die Äquivalenzaussage argumentieren. Einerseits gilt $x \in A \Rightarrow x \in B$, wie wir schnell für die einzigen beiden Elemente, 0 und 1, überprüfen. Andererseits gilt auch die Umkehrung: $x \in B \Rightarrow x \in A$. Daraus folgt die behauptete Äquivalenz.*

Definition 1.26: Mengenoperationen

1. Sind A und B Mengen, so bezeichnet man die Menge der Elemente, die sowohl in A als auch in B enthalten sind, als *Durchschnitt* von A und B.
$$A \cap B = \{x \mid (x \in A) \wedge (x \in B)\}$$
2. Sind A und B Mengen, so bezeichnet man die Menge der Elemente, die entweder in A oder in B oder in beiden enthalten sind, als *Vereinigung* von A und B.
$$A \cup B = \{x \mid (x \in A) \vee (x \in B)\}$$
3. Sind A und B Mengen, so bezeichnet man die Menge der Elemente, die in A, aber nicht in B enthalten sind, als *Differenz* von A und B.
$$A \setminus B = \{x \mid (x \in A) \wedge (x \notin B)\} = \{x \in A \mid x \notin B\}$$
4. Sind A und B Mengen, so bezeichnet man die Menge der Elemente, die entweder in A oder in B enthalten sind, als *symmetrische Differenz* von A und B.
$$A \Delta B = \{x \mid (x \in A \setminus B) \vee (x \in B \setminus A)\}$$

Diese Definitionen kann man sich mit Hilfe von **Venn-Diagrammen** verdeutlichen:



Beispiel 1.27. *Seien* $A = \{1,2,3\}$ *und* $B = \{3,4\}$. *Dann gilt:*

$$\begin{aligned} A \cap B &= \{1,2,3\} \cap \{3,4\} = \{3\}, \\ A \cup B &= \{1,2,3\} \cup \{3,4\} = \{1,2,3,4\}, \\ A \setminus B &= \{1,2,3\} \setminus \{3,4\} = \{1,2\}, \\ A \Delta B &= \{1,2,3\} \Delta \{3,4\} = \{1,2,4\}. \end{aligned}$$

Satz 1.28: Rechengesetze für Durchschnitte und Vereinigungen von Mengen

Für beliebige Mengen A, B und C gelten die *Kommutativgesetze*

$$A \cap B = B \cap A \qquad \text{und} \qquad A \cup B = B \cup A,$$

die *Assoziativgesetze*

$$A \cap (B \cap C) = (A \cap B) \cap C \quad \text{und} \quad A \cup (B \cup C) = (A \cup B) \cup C$$

und die Distributivgesetze

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{und} \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Beweis. Beginnen wir mit dem Kommutativgesetz für Schnittmengen:
Es ist $x \in A \cap B$ genau dann, wenn $x \in A$ und $x \in B$ ist, d.h.

$$x \in A \cap B \iff (x \in A) \wedge (x \in B). \tag{1.2.1}$$

Da die Konjunktion ($\wedge$) zweier Aussagen symmetrisch bezüglich der Aussagen ist, d.h. da $A \wedge B$ äquivalent ist zur Aussage $B \wedge A$, gilt also

$$(x \in A) \wedge (x \in B) \iff (x \in B) \wedge (x \in A). \tag{1.2.2}$$

Setzen wir wieder die Definition der Schnittmenge ein, so erhalten wir

$$(x \in B) \wedge (x \in A) \iff x \in B \cap A. \tag{1.2.3}$$

Aus der Kette von Äquivalenzen in (1.2.1), (1.2.2) und (1.2.3) folgt $A \cap B = B \cap A$.

Etwas kompakter zeigen wir noch das erste der beiden Distributivgesetze, wobei wir das entsprechende Distributivgesetz für die logischen Junktoren $\wedge$ und $\vee$ benutzen:

$$\begin{aligned} x \in A \cap (B \cup C) &\iff (x \in A) \wedge (x \in (B \cup C)) \\ &\iff (x \in A) \wedge ((x \in B) \vee (x \in C)) \\ &\iff ((x \in A) \wedge (x \in B)) \vee ((x \in A) \wedge (x \in C)) \\ &\iff (x \in A \cap B) \vee (x \in A \cap C) \\ &\iff x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

Die übrigen Aussagen zeigt man analog. $\square$

Auf Grund der Kommutativ- und Assoziativgesetze für Vereinigungen und Schnitte von Mengen können wir Vereinigungen und Schnittmengen von mehr als zwei Mengen definieren ohne auf die Reihenfolge oder Klammern achten zu müssen.

Für Vereinigungen und Schnittmengen von mehr als zwei Mengen gibt es eine geeignete Notation:

$$\bigcup_{i\in I} A_i = \bigcup \{\, A_i \mid\, i \in I\} \qquad \text{bzw.} \qquad \bigcap_{i\in I} A_i = \bigcap \{\, A_i \mid\, i \in I\}\,,$$

wobei I eine beliebige Indexmenge sein kann. Man nennt i in diesem Fall den *Index* der Menge A_i.

Beispiel 1.29 (Vereinigungen und Schnittmengen).

- *endliche Vereinigung:* $\bigcup_{i=1}^{3} A_i = A_1 \cup A_2 \cup A_3$*;*
- *endlicher Schnitt:* $\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n$ *(mit* $n \in \mathbb{N}$*);*
- *(abzählbar) unendliche Vereinigung:* $\bigcup_{n\in\mathbb{N}}\{n\} = \mathbb{N}$*;*
- *(abzählbar) unendlicher Schnitt:* $\bigcap_{n\in\mathbb{N}} \left[0, \frac{1}{n}\right] = \{0\}$*;*
- *(überabzählbar) unendliche Vereinigung:* $\bigcup_{x\in[0,1]}\{x\} = \{\, x \mid\, x \in [0,1]\} = [0,1]$.

Während bei den bisherigen Mengen die Reihenfolge der Aufzählung keine Rolle gespielt hat und auch wiederholt aufgezählte Elemente die Menge nicht verändern, gibt es auch die Möglichkeit, Mengen nach ihrer Sortierung zu unterscheiden.

Während $\{a, b\}$ die Menge ist, die die Elemente a und b enthält, verstehen wir unter (a, b) das *geordnete Paar* der Elemente a und b. Während Elemente einer Menge nach Cantors Definition *wohlunterschieden* sein sollen, dürfen die Komponenten eines geordneten Paars durchaus gleich sein.

Definition 1.30: kartesisches Produkt

Sind A und B zwei Mengen, so heißt die Menge

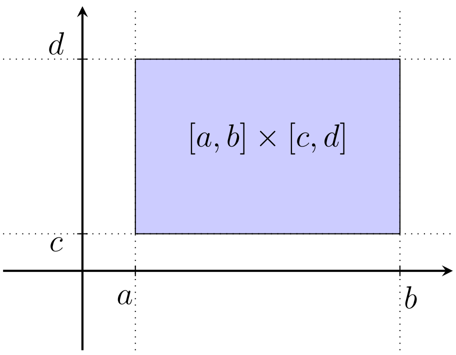
$$A \times B := \{\, (a, b) \mid\, a \in A, b \in B\}$$

das *kartesische Produkt* der Mengen A und B.

Beispiel 1.31 (kartesisches Produkt zweier Intervalle).

Seien $[a, b]$ *und* $[c, d]$ *zwei Intervalle von reellen Zahlen, d.h.* $[a, b], [c, d] \subset \mathbb{R}$*. Dann ist das kartesische Produkt beider Intervalle ein Rechteck:*

$$[a, b] \times [c, d] = \{\, (x, y) \mid\, x \in [a, b], y \in [c, d]\}$$



Bemerkung 1.32. *Das Beispiel zeigt, dass das kartesische Produkt zweier Mengen im Allgemeinen nicht kommutativ ist, d.h. es gilt* $A \times B \neq B \times A$*. So beschreibt* $[1,5] \times [1,3]$ *nicht das gleiche Rechteck wie* $[1,3] \times [1,5]$.

Verallgemeinern wir dieses Konzept gleich auf mehr als zwei Komponenten:

Man definiert Tripel, Quadrupel oder allgemeine n-*Tupel* wie folgt:

$$A_1 \times \ldots \times A_n := \{ (a_1, \ldots, a_n) \mid a_1 \in A_1, \ldots, a_n \in A_n \} .$$

Damit können wir die übliche Schreibweise $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$, $\mathbb{R}^3 := \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ oder den allgemeinen Fall für $n \in \mathbb{N}$,

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n-\text{mal}},$$

einführen.

Achtung!

Je nach Art der Klammer können unterschiedliche Objekte gemeint sein:

- $\{a, b, c, \ldots\}$ kennzeichnet eine unsortierte Menge von Objekten;
- $(a, b, c, \ldots)$ kennzeichnet ein Tupel, d.h. eine geordnete Menge – auch: Vektor oder Zahlenfolge;
- (a, b) oder $]a, b[$ kennzeichnet das offene Intervall zwischen zwei Zahlen a, b, wohingegen $[a, b]$ das abgeschlossene Intervall kennzeichnet – Achtung! Ob mit (a, b) ein offenes Intervall oder ein Tupel gemeint ist, erkennt man nur aus dem Zusammenhang heraus.
- In der Regel drückt $(\cdot)$ beim Rechnen die Reihenfolge von Operationen aus, z.B. in $(1+2) \cdot 3$; bei mehreren Klammern werden bisweilen auch eckige und geschweifte Klammern benutzt um die Lesbarkeit zu verbessern, z.B. $[(1+2) \cdot 3]^2 = 81$.

Notation

Für n-Tupel (bzw. Vektoren) gibt mehrere Schreibweisen:

Tupel: $\mathbf{a} = (a_1, a_2, \ldots, a_n)$

Zeilenvektor: $\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$

Spaltenvektor: $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$

1.3 Aussagen mit Quantoren

Neben den bisher genannten Aussagen, die meist ein einzelnes Objekt betrafen, gibt es auch Aussagen, die eine ganze Klasse von Objekten betreffen. Einige Beispiele:

- „Nachts sind alle Katzen grau.“
- „Es gibt natürliche Zahlen a, b und c, die die Gleichung $a^2 + b^2 = c^2$ erfüllen.“
- „Es gibt Vierecke, deren Seiten alle gleich lang sind.“
- „Es gibt keine lineare Gleichung, die genau zwei Lösungen hat.“

Wir führen für die Quantoren „für alle“ und „es gibt“ Schreibweisen ein, die zwei Vorteile haben. Zum einen ist es so möglich, Sätze durch Zeichenfolgen stark verkürzt darzustellen. Zum anderen ist es auf diese Art leichter, die Negation von Aussagen systematisch zu lernen.

$\forall x \in X : A(x)$	bedeutet	„für alle x in (der Menge) X ist $A(x)$ wahr“
$\exists x \in X : A(x)$	bedeutet	„es existiert ein x in X, für das $A(x)$ wahr ist“
$\nexists x \in X : A(x)$	bedeutet	„es existiert kein x in X, für das $A(x)$ wahr ist“
$\exists! x \in X : A(x)$	bedeutet	„es existiert genau ein x in X, für das $A(x)$ wahr ist“

Bemerkung 1.33. *Das Symbol $\forall$ nennt man* Allquantor*, das Symbol $\exists$ nennt man* Existenzquantor.

Beispiel 1.34. *Die Aussage „Zu jeder reellen Zahl x existiert eine reelle Zahl y, so dass $x + y = 0$ ist, lässt sich wie folgt übersetzen:*

$$\forall x \in \mathbb{R}\, \exists y \in \mathbb{R} : x + y = 0.$$

Die Reihenfolge der Quantoren ist dabei entscheidend, was wir auch an folgendem Beispiel sehen.[2]

Beispiel 1.35. *Die Mengen M und S sollen alle Menschen und Obstsorten bezeichnen. Sei $A(m, s)$ die Aussage, dass Mensch $m \in M$ die Obstsorte $s \in S$ gern isst. Dann sind folgende Aussagen grundverschieden:*

- $\forall m \in M\, \exists s \in S : A(m, s)$ *bedeutet „für alle Menschen existiert (mindestens) eine Obstsorte, die dieser Mensch gerne isst“.*
- $\exists s \in S\, \forall m \in M : A(m, s)$ *bedeutet „es existiert eine Obstsorte, die alle Menschen gern essen.“*

Kommen wir nun zur Negation von Aussagen.

Merkregel 1.36: Negation von Aussagen mit Quantoren

Die Verneinung einer Existenzaussage ist eine Allaussage und die Verneinung einer Allaussage ist eine Existenzaussage.

$$\begin{aligned} \neg(\exists x \in M : A(x)) \quad &\Longleftrightarrow \quad \forall x \in M : \neg A(x) \\ \neg(\forall x \in M : A(x)) \quad &\Longleftrightarrow \quad \exists x \in M : \neg A(x) \end{aligned}$$

[2] Die Reihenfolge kann nur dann ignoriert werden, wenn mehrere Quantoren der gleichen Art hintereinander stehen, z.B. ist $\forall x \forall y$ gleichbedeutend mit $\forall y \forall x$ und wird somit auch zu $\forall x, y$ abgekürzt.

Beispiel 1.37. *Sei P die Menge der Primzahlen. Für natürliche Zahlen $a, b \in \mathbb{N}$ schreiben wir $a \mid b$, falls a durch b teilbar ist, d.h. b ist ein Vielfaches von a.*

Die Negation der Aussage „alle Primzahlen sind nicht durch 2 teilbar" ist „es gibt (mindestens) eine Primzahl, die durch 2 teilbar ist". In Symbolen können wir dies wie folgt ausdrücken:

$$\neg\,(\forall p \in P \colon 2 \nmid p) \iff (\exists p \in P \colon 2 \mid p)\,.$$

An einem Beispiel wollen wir zeigen, wie man systematisch eine beliebige Aussage, die Quantoren enthält, negieren kann. Wir betrachten dazu folgende Aussage: „Alle Rosen sind verwelkt oder teuer."

Mathematisierung: Sei $x \in X$ eine freie Variable, die für jede beliebige Rose stehen kann. Für ein festes x formulieren wir zwei Aussagen:

$$\begin{aligned} A(x) &= \text{„Die Rose ist verwelkt."} \\ B(x) &= \text{„Die Rose ist teuer."} \end{aligned}$$

Kritische Betrachtung: Bei der Mathematisierung nehmen wir an, dass wir wissen, auf welche Gesamtheit an Rosen (X) die Aussage zutreffen soll (in einem bestimmten Blumenladen, am Valentinstag in jedem Blumenladen Deutschlands oder gar in jedem Blumenladen überall in der Welt). Zudem müssen wir ein gemeinsames Verständnis dafür haben, wann eine Rose als *verwelkt* oder *teuer* gilt.

Formulierund der Aussage: $\forall x \in X \colon (A(x) \vee B(x))$

Negation der Aussage: $\exists x \in X \colon \neg(A(x) \vee B(x))$

Vereinfachung der negierten Aussage: Mit Hilfe der De Morgan'schen Regeln kann man die negierte Aussage umformulieren:

$$\exists x \in X \colon (\neg A(x) \wedge \neg B(x))$$

Rückübersetzung: „Es gibt (mindestens) eine Rose, die weder verwelkt noch teuer ist."

Folgerung 1.38

Um eine Allaussage zu widerlegen, genügt es, ein einziges Gegenbeispiel zu finden.

Beispiel 1.39. *Die Behauptung, alle Primzahlen seien ungerade, lässt sich durch Angeben der Primzahl 2 widerlegen. Um als Beweis zu gelten, muss dabei gezeigt (oder wenigstens erwähnt) werden, dass 2 eine gerade Zahl ist und dass es eine Primzahl ist.*

1.4 Vollständige Induktion

Eine wichtige Beweismethode, die üblicherweise in der Analysis bei der Diskussion der Zahlenbereiche eingeführt wird, ist der Beweis durch vollständige Induktion. Wir verzichten hier auf eine Begründung, warum dieses Beweisprinzip funktioniert und verweisen stattdessen auf die Analysis oder die Literatur, zum Beispiel [Dal].

Beweise über vollständige Induktion werden immer dann geführt, wenn eine Aussage $A(n)$ für alle natürlichen Zahlen $n \in \mathbb{N}$ oder alle natürlichen Zahlen ab einem Startwert $n_0 \in \mathbb{N}$ gezeigt werden sollen.

Vollständige Induktion

Ist eine Aussage $A(n)$ für alle natürlichen Zahlen $n \geq n_0$ zu zeigen, so genügt es, folgende Schritte auszuführen:

Induktionsanfang: Nachweis, dass $A(n_0)$ wahr ist.

Induktionsvoraussetzung: Wir nehmen an, dass eine natürliche Zahl $n \geq n_0$ existiert, für die $A(n)$ eine wahre Aussage ist.

Induktionsbehauptung: Wir behaupten, dass $A(n+1)$ wahr ist.

Induktionsschluss: Beweis, dass aus $A(n)$ wirklich $A(n+1)$ folgt.

Beweisende Feststellung, dass damit $A(n)$ für alle natürlichen Zahlen $n \geq n_0$ bewiesen wurde.

Bemerkung 1.40.

- *Induktionsvoraussetzung, -behauptung und -schluss werden auch bisweilen zum* ***Induktionsschritt*** *zusammengefasst. Wenn bereits gesagt wurde, dass der Beweis mit vollständiger Induktion über eine Variable $n \geq n_0$ geführt wird, so reicht als Beweisende das übliche Zeichen $\square$.*
- *Die Aussage kann auch für nicht-natürliche Zahlen beweisen werden – wichtig ist nur, dass die Aussage für abzählbar viele aufeinanderfolgende Zahlen gezeigt wird. So könnte auch eine Aussage $A(n)$ für alle $n \geq -1$ gezeigt werden.*

Beispiel 1.41 (Beweis von Bemerkung 1.23)**.** *Sei M eine n-elementige Menge mit $n \in \mathbb{N}$. Dann hat die Potenzmenge $\mathcal{P}(M)$ genau 2^n Elemente; wir schreiben das kurz als $|\mathcal{P}(M)| = 2^n$.*

Diese Aussage beweisen wir mittels vollständiger Induktion:

Induktionsanfang: *Ist $n = 1$, so ist M von der Form $M = \{m\}$ und $\mathcal{P}(M) = \{\emptyset, M\}$, d.h. in der Tat ist $|\mathcal{P}(M)| = 2 = 2^n$.*

Induktionsvoraussetzung: *Angenommen es existiert ein beliebiges (aber festes) $n \in \mathbb{N}$ derart, dass $|M| = n$ und $|\mathcal{P}(M)| = 2^n$ ist. Es sei $M := \{m_1, \dots, m_n\}$.*

Induktionsbehauptung: *Wir wollen zeigen, dass für $M' := \{m_1, \dots, m_{n+1}\}$ mit $m_{n+1} \notin M$ dann $|\mathcal{P}(M')| = 2^{n+1}$ gilt.*

Induktionsschluss: *Sei $\mathcal{P}(M) := \{U_1, \dots, U_{2^n}\}$. Wegen $M \subseteq M'$ ist $\mathcal{P}(M) \subseteq \mathcal{P}(M')$. Zusätzlich von den Mengen aus $\mathcal{P}(M)$ enthält $\mathcal{P}(M')$ auch diejenigen Mengen, die entstehen, wenn man die Elemente von $\mathcal{P}(M)$ mit $\{m_{n+1}\}$ vereint, d.h.*

$$\mathcal{P}(M') = \{U_1, \dots, U_{2^n}, U_1 \cup \{m_{n+1}\}, \dots, U_{2^n} \cup \{m_{n+1}\}\}.$$

Somit hat in der Tat $\mathcal{P}(M')$ genau $2^n + 2^n = 2^{n+1}$ Elemente. $\square$

Man könnte sogar $n = 0$ zulassen, d.h. eine 0-elementige Menge. In diesem Fall ist der Induktionsanfang $|\mathcal{P}(M)| = |\{\emptyset\}| = 1 = 2^0 = 2^n$ und der Rest des Beweises ist unverändert.

1.5 Gleichungen, Lösungsmengen und Äquivalenzumformungen

Eine Gleichung hat die allgemeine Struktur

$$\text{Term}_1 = \text{Term}_2,$$

wobei in beiden Termen Variablen (Unbekannte) vorkommen können.

Beispiel 1.42.

- *Die Gleichung* $2 = 3$ *enthält keine Variablen; es handelt sich um eine falsche Aussage.*
- *Die Gleichung* $2 + 5 = 7$ *enthält ebenfalls keine Variablen; es handelt sich um eine wahre Aussage.*
- *Die Gleichung* $2x + 3 = 7$ *enthält die Variable* x*; durch Einsetzen von* $x = 2$ *erhält man eine wahre und durch Einsetzen von jedem beliebigen anderen Wert eine falsche Aussage.*
- *Die Gleichung* $y = 2x+3$ *enthält die Variablen* x *und* y*; alle Paare von Punkten* (x, y)*, die diese Gleichung zu einer wahren Aussage machen, liegen auf einer Geraden.*

Definition 1.43

Treten in einer Gleichung $T_1 = T_2$ die n Variablen $x_1, \ldots, x_n$ auf, so verstehen wir unter der *Lösungsmenge der Gleichung* die Menge aller n-Tupel $(x_1, \ldots, x_n)$, für die $T_1 = T_2$ eine wahre Aussage ergibt.

Beispiel 1.44. *Die Gleichung* $2x + 3 = 7$ *hat die Lösungsmenge* $L = \{2\}$*. Die Gleichung* $2x = 4$ *hat ebenfalls die Lösungsmenge* $L = \{2\}$*.*

Definition 1.45

Zwei Gleichungen heißen *äquivalent*, wenn sie die gleiche Lösungsmenge besitzen.

Eine Gleichung *impliziert* eine andere, wenn ihre Lösungsmenge eine Teilmenge der Lösungsmenge der anderen Gleichung ist.

Notation

Sind zwei Gleichungen äquivalent, so kann man dies durch einen Äquivalenzpfeil $\Longleftrightarrow$ darstellen.

Impliziert eine Gleichung eine andere, so verwendet man den Folgepfeil $\Longrightarrow$.

Beispiel 1.46. *Es ist*

$$2x + 3 = 7 \quad \Longleftrightarrow \quad 2x = 4 \quad \Longleftrightarrow \quad x = 2.$$

Andererseits hat man die Implikation

$$x = 2 \quad \Longrightarrow \quad x^2 = 2x,$$

da die linke Gleichung die Lösungsmenge $L_1 = \{2\}$ *hat, die rechte Gleichung hingegen die Lösungsmenge* $L_2 = \{0, 2\}$.

Wenn man auf beiden Seiten einer Gleichung die selbe Termumformung durchführt und dadurch eine äquivalente Gleichung entsteht, so spricht man von einer Äquivalenzumformung. Beispiele für Äquivalenzumformungen in $\mathbb{R}$ sind Addition der gleichen Zahl auf beiden Seiten oder Multiplikation beider Seiten mit der selben von Null verschiedenen Zahl.

1.6 Aufgaben

Aufgabe 1.1. *Bilden Sie analog zur Vorlesung die Aussagen $A \wedge B$, $A \vee B$, $A \Rightarrow B$, $A \iff B$ und $\neg A$ zu den Aussagen*

$$\begin{aligned} A&: \text{„Es herrschen frostige Temperaturen.“} \\ B&: \text{„Ich ziehe Handschuhe an.“} \end{aligned}$$

Aufgabe 1.2. *Diskutieren Sie, ob folgender Satz eine Aussage ist in dem Sinne, dass ihm ein Wahrheitsgehalt zugeschrieben werden kann:*

„Dieser Satz ist falsch.“

Aufgabe 1.3. *Betrachten Sie die folgenden Verknüpfungen von Aussagen:*

- *„Fritz bekam Bauchschmerzen und nahm Medizin.“*
- *„Fritz nahm Medizin und bekam Bauchschmerzen.“*

Diskutieren Sie anhand dieser Sätze die Kommutativität der Konjunktion in der Mathematik und in der Alltagssprache.

Aufgabe 1.4. *Ein Börsenspekulant ist sich sicher: Wenn der Goldpreis steigt, dann steigt auch der Ölpreis. Als nun der Ölpreis zu steigen beginnt, investiert er sein ganzes Vermögen in Gold. Ist das aus Sicht der Logik eine sinnvolle Entscheidung?*

Aufgabe 1.5. *Beweise mit Wahrheitstafeln*

a) *Zeigen Sie, dass für Aussagen A und B stets $A \Rightarrow A \vee B$ gilt.*
b) *Zeigen Sie die folgende in Lemma 1.10 postulierte Rechenregel für Aussagen A, B und C:*
$$A \wedge (B \vee C) \iff (A \wedge B) \vee (A \wedge C)$$

Aufgabe 1.6. *Negieren Sie folgende Aussagen ohne einfach „nicht“ einzufügen.*

a) *„Jede ganze Zahl ist durch 2 teilbar.“*
b) *„Jeden Tag scheint die Sonne oder es regnet.“*
c) *„Es gibt eine Stadt, in der alle Menschen Hüte tragen.“*

Aufgabe 1.7. *Ein Schwimmbad hat folgende Regeln, die den Zugang beschränken:*

- *Personen ohne Eintrittskarte dürfen das Schwimmbad nicht betreten.*
- *Wer betrunken ist, darf das Schwimmbad nicht betreten.*

Zur Formalisierung seien folgende Aussagen gegeben:

$$\begin{aligned} A &= \text{„Die Person darf die Schwimmhalle betreten.“} \\ B &= \text{„Die Person hat eine Eintrittskarte.“} \\ C &= \text{„Die Person ist betrunken.“} \end{aligned}$$

a) Notieren Sie die erste Regel formal. Welche Bedingung ist hinreichend und welche ist notwendig für die jeweils andere?

b) Welche Schlussfolgerung kann man daraus ziehen, dass eine Person eine Eintrittskarte besitzt?

c) Welche Schlussfolgerung kann man daraus ziehen, dass eine Person die Schwimmhalle betreten hat?

Aufgabe 1.8. *Seien m und n ganze Zahlen. Zeigen Sie einmal durch Kontraposition und einmal durch einen Widerspruchsbeweis, dass folgende Behauptung wahr ist:*

Gilt $m \cdot n \neq 0$, so folgt, dass $m \neq 0$ und $n \neq 0$ ist.

Aufgabe 1.9. *Entscheiden Sie wieder, welche der Aussagen hinreichend und welche notwendig für die jeweils andere sind.*

a) „Rex ist ein Schäferhund“ und „Rex ist ein Hund“

b) „Felix hat seine Geige dabei“ und „Felix spielt gerade auf seiner Geige“

c) „$x = 2$“ und „$x^2 = 4$“

Aufgabe 1.10. *Seien* $M_1 = \{x \in \mathbb{R} \mid |x| \leq 2\}$ *und* $M_2 = \{x \in \mathbb{R} \mid |x| < 2\}$.

a) Geben Sie Mengen M_1 und M_2 in einer anderen (kürzeren) Schreibweise an.

b) Notieren Sie, welche Menge die andere enthält.

c) Betrachten Sie die Aussageformen

$$A\colon \text{„}x \in M_1\text{“} \quad \text{und} \quad B\colon \text{„}x \in M_2\text{“}.$$

Notieren Sie, welche der Aussageformen die andere impliziert und formulieren Sie damit, welche notwendig und welche hinreichend ist.

Aufgabe 1.11. *Gegeben sind folgende Mengen:*

$$\emptyset, \quad \{1\}, \quad \{n \in \mathbb{N} \mid n \leq 5\}, \quad \mathbb{Z}, \quad [0,1], \quad \{z \in \mathbb{Z} \mid |z-1| \leq 1\}, \quad (0,1).$$

Vereinfachen Sie, wenn möglich, die Darstellung der Mengen. Bilden Sie alle möglichen Inklusionsketten mit der jeweils größtmöglichen Anzahl von Elementen der Kette.

Aufgabe 1.12. *Beschreiben Sie folgende Mengen durch die korrekte mathematische Symbolik:*

a) das offene Intervall von 3 bis 5;

b) die Menge, die nur die Zahlen 0 und 1 enthält;

c) die Menge aller geraden natürlichen Zahlen;

d) die Menge der ganzen Zahlen, deren Abstand zur 7 maximal 3 beträgt;

e) die Menge aller natürlichen Zahlen, die echt größer sind als 4.

Aufgabe 1.13. *Skizzieren Sie folgende Mengen:*

$$\begin{aligned}
A &= \{0,1\} \times \{1,2,3\}, \\
B &= \mathbb{N} \times \mathbb{N}, \\
C &= [0,1] \times [0,1] \times [0,1], \\
D &= \left\{ (x,y,z) \in \mathbb{R}^3 \,\middle|\, |x| \leq 1, |y| \leq 2, z = 3 \right\}, \\
E &= \left\{ (x,y) \in \mathbb{R}^2 \,\middle|\, \frac{1}{2}x + 1 \leq y \leq 3x - 4 \right\}.
\end{aligned}$$

Aufgabe 1.14. *Zeigen Sie, dass folgende Rechenregeln für Mengen A, B, C und Ω gelten:*

a) $A \subseteq A \cup B$,

b) $A \cap (B \cap C) = (A \cap B) \cap C$,

c) Seien $A, B \subseteq \Omega$. Dann gilt $A \cap B = \emptyset \Longrightarrow A \subseteq \Omega \setminus B$.

Aufgabe 1.15. *Wo versteckt sich der Fehler in folgendem „Beweis"?*

Behauptung: $1 = 2$

Beweis: *Es gilt $x = x$. Durch Quadrieren beider Seiten erhalten wir $x^2 = x^2$.*

Subtrahieren wir auf beiden Seiten x^2, so erhalten wir

$$x^2 - x^2 = x^2 - x^2.$$

Termumformungen auf beiden Seiten führen zu

$$x(x - x) = (x + x)(x - x).$$

Nun können wir auf beiden Seiten durch $x - x$ dividieren und erhalten

$$x = x + x.$$

Setzen wir speziell $x = 1$ ein, so erhalten wir die behauptete Gleichung $1 = 2$.

Aufgabe 1.16. *Irgendetwas stimmt hier nicht. Fügen Sie jeweils zwischen zwei aufeinanderfolgenden Zeilen die richtige logische Implikation $\Leftarrow$, $\Rightarrow$ oder $\Leftrightarrow$ ein.*

a)

$$\begin{array}{rcll}
\sqrt{2+x} & = & -4 & |()^2 \\
2 + x & = & 16 & |-2 \\
x & = & 14 &
\end{array}$$

b)

$$\begin{array}{rcll}
x & = & 2 & |()^2 \\
x^2 & = & 4 & |\sqrt{\ } \\
\sqrt{x^2} & = & 2 & \\
|x| & = & 2 & \\
(x = 2) & \vee & (x = -2) &
\end{array}$$

c)

$$\begin{array}{rcl}
5 \cdot (6 \cdot x + 1) & = & 6 \cdot (5 \cdot x + 1) \\
30 \cdot x + 5 & = & 30 \cdot x + 6 \\
5 & = & 6
\end{array}$$

Kapitel 2

Matrizen

Ziele

- ☐ Matrizen addieren und mit Skalaren und Matrizen multiplizieren
- ☐ Vektoren als spezielle Matrizen auffassen und damit rechnen
- ☐ Matrizen invertieren, falls möglich
- ☐ offene Fragen als Motivation der weiteren Kapitel

2.1 Definition und Rechenregeln

Definition 2.1

Ein rechteckiges Schema mit m Zeilen und n Spalten der Form

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} = (\alpha_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}}$$

nennen wir $m \times n$-Matrix mit Einträgen $\alpha_{k,j}$ für $k \in \{1, \ldots, m\}$ und $j \in \{1, \ldots, n\}$.

Notation

Die Menge der $m \times n$-Matrizen mit Einträgen aus der Menge $\mathbb{K}$ wird mit $\mathbb{K}^{m\times n}$ notiert. Im Moment beschränken wir uns auf $\mathbb{K} = \mathbb{R}$, d.h. die Matrizen aus $\mathbb{R}^{m\times n}$.

Ist aus dem Kontext bereits klar, in welcher Menge die Einträge liegen sollen (beispielsweise in unserem Fall gerade $\mathbb{R}$), so schreibt man auch $\mathrm{Mat}(m, n)$.

Beispiel 2.2.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix} \in \mathbb{R}^{3\times 2}$$

Es ist $\alpha_{1,2} = 2$ *und* $\alpha_{2,1} = 3$.

Um nicht nur über einzelne Elemente der Matrix, sondern über ganze Zeilen oder Spalten sprechen zu können, führen wir folgende Begriffe ein:

Definition 2.3

Ist

$$A = (\alpha_{k,j})_{\substack{1 \leq k \leq m, \\ 1 \leq j \leq n}}$$

so heißt

$$\begin{pmatrix} \alpha_{k,1} & \cdots & \alpha_{k,n} \end{pmatrix}$$

k-ter Zeilenvektor von A und

$$\begin{pmatrix} \alpha_{1,j} \\ \vdots \\ \alpha_{m,j} \end{pmatrix}$$

j-ter Spaltenvektor von A.

Einige Bezeichnungen für Matrizen sind intuitiv leicht zu verstehen – wir wollen sie der Vollständigkeit halber jedoch formal einführen.

Definition 2.4

1. Zwei Matrizen $A = (\alpha_{k,j})$ und $B = (\beta_{k,j})$ sind genau dann gleich, wenn sie die gleiche Anzahl Zeilen und Spalten haben und alle Einträge gleich sind, d.h. $\alpha_{k,j} = \beta_{k,j}$ für alle k, j.
2. Eine Matrix $A \in \mathrm{Mat}(m, n)$ heißt *Nullmatrix*, falls alle ihre Einträge 0 sind.
3. Eine Matrix $A \in \mathrm{Mat}(m, n)$ heißt *quadratisch*, falls sie gleich viele Zeilen und Spalten hat, d.h. falls $m = n$ ist. In diesem Fall nennt man die Elemente $\alpha_{k,k}$ *Diagonalelemente*; sie bilden zusammen die *Hauptdiagonale* der Matrix.
4. Eine quadratische Matrix $D \in \mathrm{Mat}(n, n)$ heißt *Diagonalmatrix*, falls sie nur auf der Hauptdiagonale von Null verschiedene Einträge hat, d.h.

$$D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

5. Eine quadratische Matrix $E_n \in \mathrm{Mat}(n, n)$ heißt *Einheitsmatrix*, falls E_n eine Diagonalmatrix ist, auf deren Hauptdiagonale überall der Eintrag 1 steht, d.h.

$$E_n = \mathrm{diag}(1, \ldots, 1) = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

Definieren wir nun die Addition von Matrizen und die skalare Multiplikation.

Definition 2.5: Addition von Matrizen

Seien $A = (\alpha_{k,j})_{\substack{1\leq k\leq m\\1\leq j\leq n}}$ und $B = (\beta_{k,j})_{\substack{1\leq k\leq m\\1\leq j\leq n}}$ Matrizen, $A, B \in \mathbb{R}^{m\times n}$. Dann ist $A + B := C = (\gamma_{k,j})_{\substack{1\leq k\leq m\\1\leq j\leq n}}$ mit $\gamma_{k,j} := (\alpha_{k,j} + \beta_{k,j})$, d.h.

$$A+B = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} + \begin{pmatrix} \beta_{1,1} & \cdots & \beta_{1,n} \\ \vdots & & \vdots \\ \beta_{m,1} & \cdots & \beta_{m,n} \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}+\beta_{1,1} & \cdots & \alpha_{1,n}+\beta_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1}+\beta_{m,1} & \cdots & \alpha_{m,n}+\beta_{m,n} \end{pmatrix}$$

Achtung! Zwei Matrizen können nur genau dann addiert werden, wenn sie gleich viele Zeilen und gleich viele Spalten haben.

Beispiel 2.6. *Seien*

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 0 \\ 2 & -1 \\ 1 & -4 \end{pmatrix}, \quad \textit{und} \quad C = \begin{pmatrix} 1 & 4 & 5 \\ 2 & -1 & 7 \\ 6 & 2 & 0 \end{pmatrix}.$$

Dann gilt:

$$A + B = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 2 & -1 \\ 1 & -4 \end{pmatrix} = \begin{pmatrix} 1+3 & 2+0 \\ 3+2 & 5-1 \\ 7+1 & 6-4 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 5 & 4 \\ 8 & 2 \end{pmatrix},$$

die Summen $A + C$ und $B + C$ sind hingegen nicht definiert.

Definition 2.7: Multiplikation einer Matrix mit einem Skalar

Sei $A \in \mathbb{R}^{m\times n}$, $A = (\alpha_{k,j})_{\substack{1\leq k\leq m\\1\leq j\leq n}}$, und sei $\lambda \in \mathbb{R}$. Dann ist $\lambda A := (\lambda\alpha_{k,j})_{\substack{1\leq k\leq m\\1\leq j\leq n}}$, d.h.

$$\lambda A = \lambda \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda\alpha_{1,1} & \cdots & \lambda\alpha_{1,n} \\ \vdots & & \vdots \\ \lambda\alpha_{m,1} & \cdots & \lambda\alpha_{m,n} \end{pmatrix}$$

Beispiel 2.8. *Seien*

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix} \quad \textit{und} \quad \lambda = 2.$$

Dann ist

$$\lambda A = 2 \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix} = \begin{pmatrix} 2\cdot 1 & 2\cdot 2 \\ 2\cdot 3 & 2\cdot 5 \\ 2\cdot 7 & 2\cdot 6 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 10 \\ 14 & 12 \end{pmatrix}.$$

Folgende Rechengesetze können wir beweisen, indem wir die entsprechenden Rechengesetze in $\mathbb{R}$ ausnutzen.

Satz 2.9

Die Matrixaddition und Multiplikation mit einem Skalar erfüllen das Kommutativ-, das Assoziativ- und das Distributivgesetz, d.h. für $A, B, C \in \mathbb{R}^{m\times n}$ und $\lambda \in \mathbb{R}$ gilt:

$$\begin{aligned} A + B &= B + A, \\ (A + B) + C &= A + (B + C), \\ \lambda(A + B) &= \lambda A + \lambda B. \end{aligned}$$

Beweis. Beispielhaft wollen wir die erste Aussage beweisen:

$$\begin{aligned} A + B &= (\alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} + (\beta_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} \stackrel{*}{=} (\underbrace{\alpha_{k,j} + \beta_{k,j}}_{\text{Addition in } \mathbb{R}})_{\substack{1\le k\le m\\1\le j\le n}} = (\beta_{k,j} + \alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} \\ &\stackrel{*}{=} (\beta_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} + (\alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} = B + A, \end{aligned}$$

wobei wir bei $*$ jeweils die Definition der Summe zweier Matrizen benutzt haben. □

Neben skalarer Multiplikation und Addition von Matrizen ist die **Multiplikation von Matrizen** wie folgt definiert:

Definition 2.10: Matrixmultiplikation

Seien $A = (\alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} \in \mathbb{R}^{m\times n}$ und $B = (\beta_{k,j})_{\substack{1\le k\le n\\1\le j\le r}} \in \mathbb{R}^{n\times r}$. Dann ist $A \cdot B = C = (\gamma_{k,j})_{\substack{1\le k\le m\\1\le j\le r}} \in \mathbb{R}^{m\times r}$ gegeben durch

$$\gamma_{k,j} = \sum_{i=1}^{n} \alpha_{k,i}\beta_{i,j}, \quad k \in \{1, \dots, m\},\ j \in \{1, \dots, r\},$$

d.h.

$$C = \begin{pmatrix} (\alpha_{1,1}\beta_{1,1} + \ldots + \alpha_{1,n}\beta_{n,1}) & \cdots & (\alpha_{1,1}\beta_{1,r} + \ldots + \alpha_{1,n}\beta_{n,r}) \\ (\alpha_{2,1}\beta_{1,1} + \ldots + \alpha_{2,n}\beta_{n,1}) & \cdots & (\alpha_{2,1}\beta_{1,r} + \ldots + \alpha_{2,n}\beta_{n,r}) \\ \vdots & & \vdots \\ (\alpha_{m,1}\beta_{1,1} + \ldots + \alpha_{m,n}\beta_{n,1}) & \cdots & (\alpha_{m,1}\beta_{1,r} + \ldots + \alpha_{m,n}\beta_{n,r}) \end{pmatrix}$$

Beispiel 2.11. *Im Allgemeinen gilt $A \cdot B \neq B \cdot A$:*

$$\begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -4 & 3 \\ 3 & -9 & 7 \\ 7 & -4 & 5 \end{pmatrix}, \textit{ aber}$$

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 6 \\ 5 & -3 \end{pmatrix}.$$

Die Multiplikation von Matrizen kann man mit Hilfe des *Falk-Schemas* verdeutlichen:

$$\begin{array}{cc|ccc} & & 1 & 2 & -1 \\ & & 0 & -3 & 2 \\ \hline 1 & 2 & 1 & -4 & 3 \\ 3 & 5 & 3 & -9 & 7 \\ 7 & 6 & 7 & -4 & 5 \end{array} \quad \text{bzw.} \quad \begin{array}{ccc|cc} & & & 1 & 2 \\ & & & 3 & 5 \\ & & & 7 & 6 \\ \hline 1 & 2 & -1 & 0 & 6 \\ 0 & -3 & 2 & 5 & -3 \end{array}$$

Ein n-dimensionaler Vektor kann als Zeilen- oder Spaltenvektor aufgefasst werden:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^{n\times 1} \quad \text{oder} \quad \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} = (v_1, \dots, v_n) \in \mathbb{R}^{1\times n}.$$

Sofern nicht anders gesagt, verstehen wir ab jetzt unter einem Vektor grundsätzlich einen Spaltenvektor.

Ein Spezialfall der Nullmatrix ist der Nullvektor $\left(\begin{smallmatrix} 0 \\ \vdots \\ 0 \end{smallmatrix}\right)$, der auch als $0 \in \mathbb{R}^n$ notiert werden kann. Durch die Angabe des Raumes ist eine Verwechslung mit der Zahl $0 \in \mathbb{R}$ ausgeschlossen. Der Nullvektor wird uns bei Gleichungssystemen noch des Öfteren begegnen.

Wenn wir einen Vektor als Zeilen- statt Spaltenvektor auffassen wollen, so nennt man diese Uminterpretation *Transponieren*. Dies ist nicht nur für Vektoren, sondern auch für Matrizen möglich:

Definition 2.12: Transponierte Matrix

Sei $A = (\alpha_{k,j})_{\substack{1\le k\le m \\ 1\le j\le n}} \in \mathbb{R}^{m\times n}$. Dann ist die zu A *transponierte Matrix*

$$A^T = (\alpha_{j,k})_{\substack{1\le j\le n \\ 1\le k\le m}} \in \mathbb{R}^{n\times m},$$

d.h.

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}, \quad A^T = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{m,1} \\ \vdots & & \vdots \\ \alpha_{1,n} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

Beispiel 2.13.

$$\textit{Zu } A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix} \textit{ ist } A^T = \begin{pmatrix} 1 & 3 & 7 \\ 2 & 5 & 6 \end{pmatrix}.$$

Definition 2.14: Symmetrische Matrix

Eine quadratische Matrix $A = (\alpha_{k,j})_{\substack{1\le k\le n \\ 1\le j\le n}} \in \mathbb{R}^{n\times n}$ heißt *symmetrisch*, falls $A^T = A$ ist, d.h. falls $\alpha_{k,j} = \alpha_{j,k}$ ist für $j, k \in \{1, \dots, n\}$.

Beispiel 2.15. *Matrix A aus Beispiel 2.13 ist schon deshalb nicht symmetrisch, weil sie nicht quadratisch ist. Betrachten wir*

$$B = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \quad \text{und} \quad C = \begin{pmatrix} -1 & 4 \\ 4 & 2 \end{pmatrix}.$$

Es ist

$$B^T = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \neq B \quad \text{und} \quad C^T = \begin{pmatrix} -1 & 4 \\ 4 & 2 \end{pmatrix} = C,$$

d.h. B ist nicht symmetrisch, C hingegen schon.

Satz 2.16: Rechenregeln für das Transponieren

Seien $A, B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times r}$ und $\lambda \in \mathbb{R}$. Dann gelten folgende Rechenregeln:

1. $(A + B)^T = A^T + B^T$,
2. $(\lambda A)^T = \lambda A^T$,
3. $(A^T)^T = A$,
4. $(A \cdot C)^T = C^T \cdot A^T$.

Beweis. Seien $A = (\alpha_{k,j})$, $B = (\beta_{k,j})$ und $C = (\gamma_{j,l})$ mit $k \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$ und $l \in \{1, \dots, r\}$.

1. $(A + B)^T = (\alpha_{k,j} + \beta_{k,j})^T =: (c_{k,j})^T = (c_{j,k}) = (\alpha_{j,k} + \beta_{j,k}) = A^T + B^T$
2. $(\lambda A)^T = (\lambda\alpha_{k,j})^T = (\lambda\alpha_{j,k}) = \lambda A^T$
3. $(A^T)^T = ((\alpha_{k,j})^T)^T = (\alpha_{j,k})^T = (\alpha_{k,j}) = A$
4. $(A \cdot C)^T = (\sum_i \alpha_{k,i}\gamma_{i,l})^T =: (c_{k,l})^T = (c_{l,k}) = (\sum_i \alpha_{i,k}\gamma_{l,i}) = (\sum_i \gamma_{l,i}\alpha_{i,k}) = C^T \cdot A^T$

□

2.2 Rechenoperationen für Vektoren

In diesem Abschnitt wollen wir die Rechenoperationen, die wir für Matrizen kennen gelernt haben, auf Vektoren anwenden und das Ergebnis für den Spezialfall von Vektoren im $\mathbb{R}^2$ interpretieren.

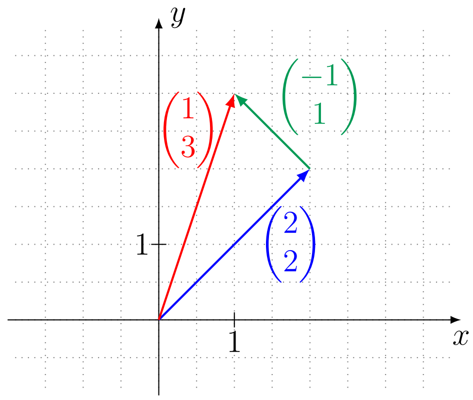
Addition von Vektoren

Allgemein: Zwei Vektoren v und w kann man genau dann addieren, wenn sie aus dem gleichen Raum sind, z.B. $v, w \in \mathbb{R}^n$. In diesem Fall gilt

$$v + w = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}.$$

Beispiel im $\mathbb{R}^2$:

$$\begin{pmatrix}2\\2\end{pmatrix} + \begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix}1\\3\end{pmatrix}$$



Interpretation: Interpretiert man die Vektoren als Verschiebung, so entspricht die Addition von Vektoren der Nacheinanderausführung zweier Verschiebungen.

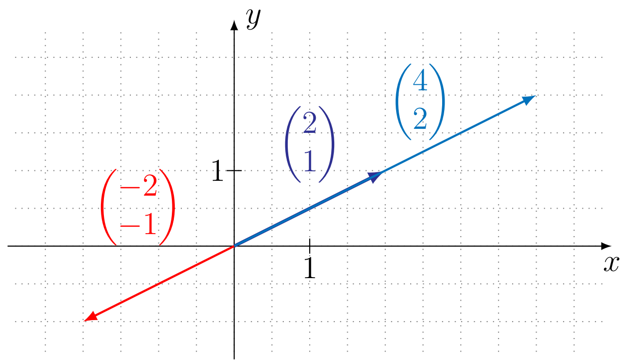
Multiplikation mit einem Skalar

Allgemein: Ist $\lambda \in \mathbb{R}$ und $v \in \mathbb{R}^n$, so ist

$$\lambda\, v = \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

Beispiel im $\mathbb{R}^2$:

$$2\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}4\\2\end{pmatrix} \quad \text{und} \quad (-1)\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}-2\\-1\end{pmatrix}.$$



Interpretation: Multiplikation mit einem Skalar $|\lambda| > 1$ ($|\lambda| < 1$) entspricht einer Streckung (Stauchung) um den Faktor $|\lambda|$ und ist $\lambda < 0$, so wird die Richtung des Vektors um 180 Grad gedreht.

Multiplikation von Vektoren

Allgemein: Zwei Vektoren v und w kann man genau dann (im Sinne der Multiplikation von Matrizen) multiplizieren, wenn der zweite Vektor transponiert ist oder wenn beide Vektoren aus dem gleichen Raum sind und der erste Vektor transponiert ist. Es gilt dann

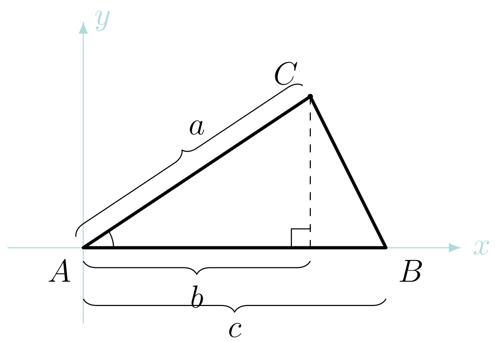
$$vw^T = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \begin{pmatrix} w_1 & \cdots & w_n \end{pmatrix} = \begin{pmatrix} v_1w_1 & \cdots & v_1w_n \\ \vdots & & \vdots \\ v_mw_1 & \cdots & v_mw_n \end{pmatrix},$$

$$v^Tw = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \sum_{k=1}^{n} v_k w_k.$$

Das erste Produkt hat für uns in diesem Kurs keine weitere Bedeutung;das zweite Produkt ist das *Skalarprodukt* beider Vektoren. Diesen Begriff werden wir später noch genauer kennen lernen.

Beispiel im $\mathbb{R}^2$:

$$\begin{pmatrix} 3 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 3 \cdot 4 + 2 \cdot 0 = 12.$$



Nach den Sätzen zu Winkeln am rechtwinkligen Dreieck ist

$$\cos\theta = \frac{b}{a},$$

also ist

$$b \cdot c = a \cdot c \cdot \cos\theta.$$

Man kann zeigen, dass dies gerade das oben definierte Skalarprodukt ist.

Interpretation: Im $\mathbb{R}^2$ ist $v^Tw = |v||w|\cos(\sphericalangle(v,w))$, wobei $|v|$ und $|w|$ die Länge der Vektoren v und w bezeichnet.

2.3 Inverse Matrizen

Betrachten wir die Produkte der Matrizen

$$A = \begin{pmatrix} 5 & 1 \\ 4 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & -1 \\ -4 & 5 \end{pmatrix}:$$

Es ist

$$A \cdot B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad B \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Wir nennen B die *inverse Matrix* zu A.

Definition 2.17: Inverse

Sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix. Existiert eine Matrix $B \in \mathbb{R}^{n\times n}$, so dass

$$A \cdot B = B \cdot A = E_n$$

erfüllt ist, wobei $E_n \in \mathbb{R}^{n\times n}$ die Einheitsmatrix sei, so heißt A *invertierbar*. Man schreibt $B = A^{-1}$ und nennt A^{-1} die *Inverse* zu A.

Bemerkung 2.18. *Warum werden nur quadratische Matrizen in der Definition zugelassen? In Beispiel 2.11 hatten wir Matrizen A und B betrachtet, für die $A \cdot B \in \mathbb{R}^{3\times 3}$, aber $B \cdot A \in \mathbb{R}^{2\times 2}$ ist. Bei nicht quadratischen Matrizen ergibt die Multiplikation von links und von rechts mit derselben Matrix zwei Matrizen unterschiedlicher Größe – sofern die Produkte überhaupt definiert sind.*

Bevor wir zeigen, dass die Inverse einer Matrix, sofern sie existiert, tatsächlich eindeutig bestimmt ist, müssen wir ein paar Rechenregeln für die Multiplikation von Matrizen festhalten.

Satz 2.19

Die Multiplikation von Matrizen erfüllt das Assoziativ- und das Distributivgesetz, jedoch in der Regel nicht das Kommutativgesetz. Mit anderen Worten, für Matrizen A, B und C, für die die folgenden Summen und Produkte gebildet werden können, gilt:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C), \tag{2.3.1}$$

$$(A + B) \cdot C = A \cdot C + B \cdot C, \tag{2.3.2}$$

$$A \cdot (B + C) = A \cdot B + A \cdot C. \tag{2.3.3}$$

Insbesondere gelten die o.g. Rechenregeln also für Matrizen $A, B, C \in \mathbb{R}^{n\times n}$.

Beweis. Alle drei Aussagen werden bewiesen, indem man sich auf die Ebene der Matrixeinträge begibt und die Rechenregeln in $\mathbb{R}$ verwendet. Exemplarisch werden wir Gleichung (2.3.1) für drei quadratische Matrizen $A, B, C \in \mathbb{R}^{n\times n}$ zeigen.

Seien $A = (\alpha_{i,j})_{(i,j)\in\mathcal{N}}$, $B = (\beta_{j,k})_{(j,k)\in\mathcal{N}}$ und $C = (\gamma_{k,l})_{(k,l)\in\mathcal{N}}$ mit $\mathcal{N} := \{1, \ldots, n\}^2$. Dann ist gemäß der Definitionen von Produkten von Matrizen

$$A \cdot B = \left(\sum_{j=1}^{n} \alpha_{i,j}\beta_{j,k}\right)_{(i,k)\in\mathcal{N}} =: (\delta_{i,k})_{(i,k)\in\mathcal{N}}$$

$$(A \cdot B) \cdot C = \left(\sum_{k=1}^{n} \delta_{i,k}\gamma_{k,l}\right)_{(i,l)\in\mathcal{N}}$$

$$B \cdot C = \left(\sum_{k=1}^{n} \beta_{j,k}\gamma_{k,l}\right)_{(j,l)\in\mathcal{N}} =: (\eta_{j,l})_{(j,l)\in\mathcal{N}}$$

$$A \cdot (B \cdot C) = \left(\sum_{j=1}^{n} \alpha_{i,j} \eta_{j,l} \right)_{(i,l) \in \mathcal{N}}$$

Auf Ebene der Einträge der Matrizen können wir Kommutativ-, Assoziativ- und Distributivgesetze (kurz KG, AG und DG) der reellen Zahlen anwenden und erhalten

$$\begin{aligned}
(A \cdot B) \cdot C &= \left(\sum_{k=1}^{n} \delta_{i,k} \gamma_{k,l} \right)_{(i,l) \in \mathcal{N}} && (\delta_{i,k} \text{ durch Summe ersetzen}) \\
&= \left(\sum_{k=1}^{n} \left(\sum_{j=1}^{n} \alpha_{i,j} \beta_{j,k} \right) \gamma_{k,l} \right)_{(i,l) \in \mathcal{N}} && (DG) \\
&= \left(\sum_{k=1}^{n} \left(\sum_{j=1}^{n} \alpha_{i,j} \beta_{j,k} \gamma_{k,l} \right) \right)_{(i,l) \in \mathcal{N}} && (KG) \\
&= \left(\sum_{j=1}^{n} \left(\sum_{k=1}^{n} \alpha_{i,j} \beta_{j,k} \gamma_{k,l} \right) \right)_{(i,l) \in \mathcal{N}} && (AG) \\
&= \left(\sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{i,j} \left(\beta_{j,k} \gamma_{k,l} \right) \right)_{(i,l) \in \mathcal{N}} && (DG) \\
&= \left(\sum_{j=1}^{n} \alpha_{i,j} \left(\sum_{k=1}^{n} \beta_{j,k} \gamma_{k,l} \right) \right)_{(i,l) \in \mathcal{N}} && (\text{Summe durch } \eta_{j,l} \text{ ersetzen}) \\
&= \left(\sum_{j=1}^{n} \alpha_{i,j} \eta_{j,l} \right)_{(i,l) \in \mathcal{N}} \\
&= A \cdot (B \cdot C).
\end{aligned}$$

□

Satz 2.20

Die Inverse zu einer quadratischen Matrix ist, sofern sie existiert, eindeutig bestimmt.

Beweis. Seien B und C zwei zu $A \in \mathbb{R}^{n \times n}$ inverse Matrizen. Dann gilt

$$B = B \cdot E_n = B \cdot (A \cdot C) \stackrel{(*)}{=} (B \cdot A) \cdot C = E_n \cdot C = C,$$

wobei wir in $(*)$ die Assoziativität der Multiplikation von Matrizen aus Satz 2.19 benutzt haben. □

Im Folgenden wollen wir überlegen, wie wir die Inverse einer quadratischen Matrix, sollte sie existieren, berechnen können.

Beispiel 2.21. *Gegeben sei die Matrix*

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}.$$

Sei

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \textit{mit } a, b, c, d \in \mathbb{R}$$

ein Kandidat für die Inverse zu A. Dann gilt:

$$A \cdot B = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a + c & 2b + d \\ 5a + 3c & 5b + 3d \end{pmatrix} \overset{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Wir müssen also die folgenden zwei linearen Gleichungssysteme (LGS) lösen:

$$\begin{aligned} 2a + c &= 1 \\ 5a + 3c &= 0 \end{aligned} \qquad \textit{und} \qquad \begin{aligned} 2b + d &= 0 \\ 5b + 3d &= 1 \end{aligned}$$

Im ersten LGS erhält man durch Umstellen der ersten Gleichung $c = 1 - 2a$ und Einsetzen in die zweite Gleichung liefert

$$5a + 3 \cdot (1 - 2a) = 3 - a \overset{!}{=} 0.$$

Daraus folgt $a = 3$, also $c = -5$.

Im zweiten LGS erhält man analog zunächst durch Umstellen $d = -2b$, Einsetzen in die zweite Gleichung liefert die Bedingung $-b = 1$, woraus $b = -1$ und damit $d = 2$ folgt. Insgesamt ist also die gesuchte Inverse

$$A^{-1} = B = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}.$$

Man kann prüfen, dass in der Tat $A \cdot B = B \cdot A = E_2$ ist.

Es ist klar, dass die Berechnung inverser Matrizen durch Lösen von LGS zeitintensiv ist, insbesondere, wenn man nicht weiß, ob eine Matrix überhaupt invertierbar ist. Im Laufe der Vorlesung werden wir also (unter anderem) folgende Fragen beantworten:

- Können wir vorhersagen, ob eine Matrix invertierbar ist, ohne ein LGS lösen zu müssen?
- Können wir vorhersagen, ob ein LGS lösbar ist, ohne es bereits zu lösen?
- Wie lässt sich ein LGS systematisch lösen? Wie kann die Lösungsmenge aussehen?

2.4 Aufgaben

Aufgabe 2.1. *Gegeben sind folgende Matrizen:*

$$A = \begin{pmatrix} 2 & 5 \\ 3 & 6 \\ 4 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 3 \\ 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

Berechnen Sie $-2A + 3B$ *und* B^T.

Aufgabe 2.2. *Zeigen Sie am Beispiel der folgenden Matrizen, dass die Multiplikation von Matrizen nicht kommutativ ist, d.h. dass im Allgemeinen* $A \cdot B \neq B \cdot A$ *ist.*

$$A = \begin{pmatrix} -6 & 9 & -3 \\ -4 & 6 & -2 \\ -2 & 3 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{pmatrix}.$$

Aufgabe 2.3. *Seien* $A, B, A \cdot B \in \mathbb{R}^{n \times n}$ *invertierbare Matrizen. Zeigen Sie, dass gilt:*

a) $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$,

b) $(A^{-1})^{-1} = A$,

c) $(A^T)^{-1} = (A^{-1})^T$.

Aufgabe 2.4. *Für reelle Zahlen* $a, b, c, d \in \mathbb{R}$, *welche* $a \cdot d \neq b \cdot c$ *erfüllen, seien*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{und} \quad B = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Überprüfen Sie, dass $A \cdot B = B \cdot A = E_2$ *ist, d.h. dass* B *die Inverse zu* A *ist.*

Aufgabe 2.5. *Seien*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 4 & 3 \\ 2 & 2 \end{pmatrix}.$$

a) Berechnen Sie A^{-1} *und* B^{-1}.

b) Berechnen Sie $A \cdot B$.

c) Berechnen Sie $(A \cdot B)^{-1}$.

d) Berechnen Sie das Produkt $B^{-1} \cdot A^{-1}$ *und vergleichen Sie Ihr Ergebnis mit dem aus Aufgabenteil c).*

Aufgabe 2.6. *Sei* $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. *Finden Sie eine Matrix* B *derart, dass* $A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ *ist.*

Wie viele solche Matrizen gibt es?

Aufgabe 2.7. *Seien* $a, b, c, d, \lambda \in \mathbb{R}$ *beliebig.*

a) Berechnen Sie $A_k \cdot B$ *für*

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

b) Für welche Matrizen $D_k \in \mathbb{R}^{2 \times 2}$ *(*$k \in \{1, 2\}$*) ist* $D_k \cdot B = C_k$ *mit* B *wie eben und*

$$C_1 = \begin{pmatrix} \lambda a & \lambda b \\ c & d \end{pmatrix} \quad \text{bzw.} \quad C_2 = \begin{pmatrix} a & b \\ a - c & b - d \end{pmatrix}?$$

Kapitel 3

Algebraische Strukturen: Gruppen, Ringe und Körper

Ziele

- ☐ mit den Grundbegriffen zu Abbildungen (inkl. Verknüpfungen) umgehen können
- ☐ mit den Begriffen Gruppe, Ring, Körper umgehen können und jeweils Beispiele und Gegenbeispiele, sowie wichtige Eigenschaften kennen
- ☐ Beweise unter Verwendung der Definitionen dieser Begriffe durchführen
- ☐ die Begriffe Äquivalenzrelationen und -klassen verstehen und Beispiele dafür kennen
- ☐ eine Familie von endlichen Körpern kennen und ihre Konstruktion nachvollziehen können

3.1 Abbildung und Verknüpfung

Definition 3.1: Abbildung

Sind X und Y beliebige Mengen, so versteht man unter einer *Abbildung* von X nach Y eine Vorschrift f, die jedem Element $x \in X$ eindeutig ein $f(x) \in Y$ zuordnet.

Notation

Eine Abbildung f von X nach Y gibt man wie folgt an:

$$f\colon X \to Y, \qquad x \mapsto f(x)$$

Beispiel 3.2 (Abbildungen).

a) $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) := 3x + 5$

b) $f\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$, $\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \mapsto \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$ *(Addition von Vektoren im* $\mathbb{R}^2$*)*

c) $f\colon \mathbb{R}^3 \to \mathbb{R}^2$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) := \begin{pmatrix} 3x + 2y + z \\ -2x + y + 4z \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ -2 & 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

Bemerkung 3.3. *Sei* $M = \{1, 2, \ldots, 10\}$. *Die Zuordnungsvorschrift*

$$x \mapsto f(x) := \{m \in M \mid m \text{ teilbar durch } x\}$$

kann auch als Abbildung aufgefasst werden, obwohl beispielsweise $f(2) = \{2, 4, 6, 8, 10\}$ *nicht ein einziges Element, sondern eine ganze Menge von Elementen von* M *enthält. Dazu wird* f *als Abbildung von* M *nach* $\mathcal{P}(M)$ *definiert (*$f\colon M \to \mathcal{P}(M)$*), wobei* $\mathcal{P}(M)$ *die Potenzmenge von* M *ist.*

Definition 3.4

Ist $f\colon X \to Y$ eine Abbildung, so nennt man X *Definitionsmenge* oder *Definitionsbereich* und Y *Zielmenge* oder *Wertevorrat* von f.

Ist $A \subseteq X$ eine beliebige Teilmenge von X, so ist

$$f(A) := \{\, f(x) \mid x \in A\}$$

das *Bild von* A *unter* f. Für $A = X$ bezeichnen wir

$$\operatorname{Bild}(f) := \{\, f(x) \mid x \in X\} = f(X)$$

als *Bild*, *Bildmenge* oder *Bildbereich* von f. Ist $B \subseteq Y$, so heißt

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}$$

das *Urbild von* B *unter* f.

Bemerkung 3.5. *An Stelle der Bezeichnungen* $\operatorname{Bild}(f)$ *oder* $f(X)$ *ist auch die Notation* $\operatorname{Im}(f)$ *für das Bild von* f *üblich, da diese Menge im Englischen* image *genannt wird. Die Abkürzung* Im *wird jedoch auch für den in Kürze auftretenden* imaginären Anteil *benutzt, was zu Missverständnissen führen kann.*

Bemerkung 3.6. *Gibt es zwei Werte* $x_1, x_2 \in X$, *so dass* $f(x_1) = f(x_2)$ *ist, so wird dieser Funktionswert formal doppelt im Bild von* f *auftauchen. Um dies zu verhindern, kann man das Bild auch folgendermaßen beschreiben:*

$$\operatorname{Bild}(f) = \{y \in Y \mid \exists x \in X\colon f(x) = y\}\,.$$

Beispiel 3.7. *Betrachten wir die Abbildung* $f\colon \mathbb{R} \to \mathbb{R}$ *mit* $f(x) = x^2$. *Dann ist*

- $\mathbb{R}$ *Definitions- und auch Zielmenge von* f;
- $f(\mathbb{R}) = \mathbb{R}_{\geq 0} = [0, \infty)$ *das Bild von* f;
- $f([1, 2]) = [1, 4]$ *das Bild des abgeschlossenen Intervalls* $[1, 2]$ *unter* f;
- $f^{-1}([1, 4]) = [-2, -1] \cup [1, 2]$ *das Urbild des Intervalls* $[1, 4]$ *unter* f.

Definition 3.8: injektiv, surjektiv, bijektiv

Seien X und Y beliebige Mengen und sei $f\colon X \to Y$ eine Abbildung. f heißt

1. *injektiv*, falls aus $f(x) = f(x')$ folgt, dass $x = x'$ ist;
2. *surjektiv*, falls zu jedem $y \in Y$ mindestens ein $x \in X$ existiert mit $f(x) = y$;
3. *bijektiv*, falls f injektiv und surjektiv ist.

Beispiel 3.9. *Analysieren wir diese Eigenschaften anhand der Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$.*

- *f ist nicht injektiv, da $f(-2) = 4 = f(2)$ ist; die Injektivität erreicht man für diese Funktion durch Einschränken der Definitionsmenge auf $\mathbb{R}_{\geq 0}$.*
- *f ist nicht surjektiv, da beispielsweise $-1 \notin f(\mathbb{R})$ ist; die Surjektivität erreicht man für diese Funktion durch Einschränken der Zielmenge auf das tatsächliche Bild, d.h. auf $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$.*
- *f ist somit insbesondere nicht bijektiv. Die Einschränkung $f|_{\mathbb{R}_{\geq 0}}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ hingegen ist bijektiv.*

Wir sehen, dass nur dann sinnvoll entschieden werden kann, ob eine Funktion injektiv, surjektiv oder gar bijektiv ist, wenn Definitions- und Zielmenge angegeben werden.

Definition 3.10: Verknüpfung

Eine *Verknüpfung* auf einer Menge X ist eine Abbildung, die zwei Elementen der Menge ein neues Element der Menge zuordnet, d.h.

$$*\colon X \times X \to X, \qquad (x, y) \mapsto x * y$$

Neben b) aus Beispiel 3.2 sind auch folgende Abbildungen Verknüpfungen:

Beispiel 3.11 (Verknüpfungen).

a) $+\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *mit* $(x, y) \mapsto x + y$ *(Addition natürlicher Zahlen)*

b) $\cdot\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *mit* $(x, y) \mapsto x \cdot y$ *(Multiplikation reeller Zahlen)*

c) $*\colon \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ *mit* $(x, y) \mapsto \frac{1}{2}(x + y)$ *(Durchschnitt zweier rationaler Zahlen)*

Was sind hingegen Abbildungen, die keine Verknüpfungen sind? Beispiele a) und c) aus Beispiel 3.2 sind keine Verknüpfungen, da die Abbildungen nur ein Argument besitzen. Ebenso ist

$$-\colon \mathbb{N} \times \mathbb{N} \to \mathbb{Z} \quad \text{mit} \quad (x, y) \mapsto x - y$$

keine Verknüpfung, da das Ergebnis der Differenz zweier natürlicher Zahlen keine natürliche Zahl sein muss.

3.2 Gruppen

Definition 3.12: (Halb-)Gruppe

Eine Halbgruppe $(G,*)$ ist eine nichtleere Menge G zusammen mit einer assoziativen Verknüpfung

$$*\colon G\times G\to G,$$

d.h. es gilt

(AG) $(x*y)*z=x*(y*z)$ für alle $x,y,z\in G$.

Gelten zudem die Bedingungen

(N) es gibt ein neutrales Element $e\in G$ mit $e*x=x*e=x$ für alle $x\in G$,

(I) jedes Element besitzt ein Inverses, d.h. zu jedem $x\in G$ existiert $x'\in G$ mit $x'*x=x*x'=e$,

so nennen wir $(G,*)$ eine *Gruppe*. Ist die Verknüpfung zudem kommutativ, d.h. gilt

(KG) $x*y=y*x$ für alle $x,y\in G$,

so heißt die Gruppe *kommutativ* oder *abelsch*.

Beispiel 3.13.
Betrachten wir die Mengen und Verknüpfungen aus Beispiel 3.11:

a) $(\mathbb{N},+)$ *ist keine Gruppe, da (N) (und somit auch (I)) verletzt ist; es ist eine Halbgruppe.*

b) $(\mathbb{R}_{>0},\cdot)$ *ist eine abelsche Gruppe; das neutrale Element ist die* 1 *und das Inverse zu* $x\in\mathbb{R}_{>0}$ *ist* $x^{-1}=\frac{1}{x}\in\mathbb{R}_{>0}$*. Allerdings ist* $(\mathbb{R},\cdot)$ *nur eine Halbgruppe, da* $0\in\mathbb{R}$ *kein inverses Element besitzt.*

c) $(\mathbb{Q},*)$ *ist keine Halbgruppe und somit auch keine Gruppe, da (AG) verletzt ist.*

Notation

Bei Gruppen mit der Addition als Verknüpfung bezeichnet man das neutrale Element üblicherweise als 0 und das zu x inverse Element als $-x$. Bei Gruppen mit der Multiplikation als Verknüpfung bezeichnet man das neutrale Element häufig als 1 und das zu x inverse Element als x^{-1}.

Bemerkung 3.14. *Da Halbgruppen und Gruppen nicht notwendigerweise kommutativ sind, beinhalten die Eigenschaften (N) und (I) eigentlich jeweils zwei Aussagen:*

(N1) es gibt ein linksneutrales Element $e\in G$ *mit* $e*x=x$ *für alle* $x\in G$*;*

(N2) ein linksneutrales Element ist zugleich rechtsneutral, d.h. es gilt $x*e=x$ *für alle* $x\in G$*;*

(I1) jedes Element $x \in G$ besitzt ein linksinverses Element $x' \in G$ so dass $x' * x = e$ ist;

(I2) ein zu $x \in G$ linksinverses Element x' ist zugleich rechtsinvers, d.h. $x * x' = e$.

Die Definition ist jedoch bewusst so gewählt, da wir gleich sehen werden, dass linksneutrale bzw. -inverse Elemente immer auch gleichzeitig rechtsneutral bzw. -invers sind. Zudem können wir von dem *neutralen Element bzw.* dem *zu $x \in G$ inversen Element sprechen, da es nie mehr als ein solches geben kann.*

Satz 3.15: Eindeutigkeit des neutralen Elements und der inversen Elemente

In jeder Gruppe gibt es genau ein neutrales Element und jedes Element einer Gruppe hat genau ein inverses Element. Zudem ist jedes linksneutrale Element gleichzeitig rechtsneutral und jedes linksinverse Element ist gleichzeitig auch rechtsinvers.

Beweis. Der Beweis besteht aus mehreren Schritten.

Jedes Linksinverse ist auch Rechtsinvers. Es gelte $x' * x = e$, d.h. x' ist ein Linksinverses von x. Sei weiterhin x'' ein Linksinverses von x', d.h. $x'' * x' = e$. Dann gilt

$$\begin{aligned} x * x' &= e * (x * x') = (x'' * x') * (x * x') = x'' * ((x' * x) * x') = x'' * (e * x') \\ &= x'' * x' = e, \end{aligned}$$

d.h. x' ist in der Tat ein Rechtsinverses von x.

Jedes linskneutrale Element ist auch rechtsneutral. Es gelte $e * x = x$ und $x' * x = x * x' = e$. Dann ist e rechtsneutral, denn

$$x * e = x * (x' * x) = (x * x') * x = e * x = x.$$

Es gibt nur ein neutrales Element. Es seien sowohl e_1 als auch e_2 neutrale Elemente:

$$e_1 * x = x * e_1 = x, \quad \forall x \in G, \tag{3.2.1}$$
$$e_2 * x = x * e_2 = x, \quad \forall x \in G. \tag{3.2.2}$$

Setzen wir e_2 in (3.2.1) und e_1 in (3.2.2) ein, so erhalten wir

$$e_1 * e_2 = e_2 \quad \text{und} \quad e_1 * e_2 = e_1,$$

also ist $e_1 = e_1 * e_2 = e_2$.

Für jedes $x \in G$ gibt es nur ein zu x inverses Element. Seien y und z zu x inverse Elemente. Dann gilt

$$y = y * e = y * (x * z) = (y * x) * z = e * z = z.$$

□

Bemerkung 3.16. *Obwohl eine Gruppe im Allgemeinen nicht kommutativ sein muss, ist – das zeigt das gerade bewiesene Resultat – die Verknüpfung eines Elementes mit seinem Inversen oder mit dem neutralen Element immer kommutativ.*

Folgendes Lemma erleichtert die Suche nach einem neutralen Element, da ein Kandidat nicht mit allen Elementen der Gruppe verknüpft werden muss um nachzuweisen, dass er tatsächlich das neutrale Element ist.

Lemma 3.17

Sei $(G,*)$ eine Gruppe. Gilt für ein $x \in G$ die Gleichheit $y * x = x$, dann ist y das neutrale Element der Gruppe.

Beweis. Sei e das neutrale Element der Gruppe $(G,*)$ und für Elemente $x, y \in G$ gelte $y * x = x$. Sei weiterhin $x' \in G$ das zu x inverse Element. Dann gilt

$$y = y * e = y * (x * x') = (y * x) * x' = x * x' = e,$$

d.h. y stimmt mit dem (eindeutig bestimmten) neutralen Element der Gruppe überein. $\square$

Beispiel 3.18 (Rechnen mit Rest). *Herr Itô setzt sich um 17:20 Uhr Ortszeit in Helsinki in einen Flieger und landet exakt $9\frac{1}{2}$ Stunden später (um 8:50 Uhr Ortszeit) in Tokyo. Unterwegs fiel ihm etwas ein, dass er seinem Kollegen in Helsinki schnellstmöglich mitteilen möchte. Doch wie spät ist es bei seiner Landung in Helsinki?*

Rechnen wir kurz nach: Nach einer halben Stunde ist es 17:50 Uhr und weitere 9 Stunden später ist es also 2:50 Uhr. Herr Itô wird seinen Kollegen in Finnland also lieber erst ein paar Stunden später anrufen.

Was hat dieses Beispiel mit Gruppentheorie zu tun? Im Allgemeinen rechnen wir $17+9 = 26$. In unserem Beispiel ist die Zeit allerdings nicht 26:50 Uhr, sondern 2:50 Uhr. Addieren wir zwei Stundenangaben und ist das Ergebnis keine Zahl zwischen 0 und 24, so müssen wir so oft 24 abziehen, bis wir in diesem Intervall landen. Mit anderen Worten: Wir berechnen den Rest bei Division durch 24. Beim Rechnen mit Minuten verfahren wir analog mit Division durch 60 (ggf. mit Änderung der zugehörigen Stundenzahl).

Das einfachste Beispiel für Rechnen mit Rest ist innerhalb der Menge $\{0,1\}$. Das Ergebnis können wir in einer Additions- und einer Multiplikationstabelle festhalten:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Anhand der Tabellen erkennen wir, dass $(\{0,1\},+)$ eine abelsche Gruppe ist (mit neutralem Element 0); $(\{0,1\},\cdot)$ ist hingegen keine Gruppe, da zwar ein neutrales Element existiert (1), das Element 0 jedoch kein Inverses besitzt.

Bei Abbildungen zwischen Gruppen heben wir diejenigen Abbildungen hervor, die in folgendem Sinne mit den Verknüpfungen innerhalb der Gruppen kompatibel sind.

Definition 3.19: Homomorphismus

Seien $(G,*)$ und $(H,\circ)$ Gruppen. Eine Abbildung $\varphi\colon G \to H$ heißt *Homomorphismus* (oder Gruppenhomomorphismus), falls

$$\varphi(x * y) = \varphi(x) \circ \varphi(y), \quad \forall x, y \in G.$$

Ein Homomorphismus heißt *Isomorphismus*, wenn er bijektiv ist.

Beispiel 3.20. *Die* lineare *Abbildung* $f\colon \mathbb{R} \to \mathbb{R}$ *mit* $f(x) = 3x$ *ist ein Gruppenhomomorphismus zwischen den Gruppen* $(\mathbb{R}, +)$ *und* $(\mathbb{R}, +)$*, denn*

$$f(x + y) = 3(x + y) = 3x + 3y = f(x) + f(y).$$

Die affine *Abbildung* $g\colon \mathbb{R} \to \mathbb{R}$ *mit* $g(x) = 3x+2$ *hingegen ist zwischen den selben Gruppen kein Homomorphismus:*

$$g(x + y) = 3(x + y) + 2 \neq (3x + 2) + (3y + 2) = g(x) + g(y).$$

Beispiel 3.21. *Die Abbildung* $\varphi\colon \mathbb{R} \to \mathbb{R}_{>0}$ *mit* $\varphi(x) = e^x$ *ist ein Gruppenhomomorphismus zwischen* $(\mathbb{R}, +)$ *und* $(\mathbb{R}_{>0}, \cdot)$*, denn*

$$\varphi(x + y) = e^{x+y} = e^x \cdot e^y = \varphi(x) \cdot \varphi(y), \qquad \forall x, y \in \mathbb{R}.$$

Folgender Satz illustriert, warum Homomorphismen dem Wort nach als *strukturerhaltend* bezeichnet werden:

Satz 3.22

Sei $\varphi\colon (G, *) \to (H, \circ)$ ein Gruppenhomomorphismus und seien e_G und e_H die neutralen Elemente der Gruppen $(G, *)$ und $(H, \circ)$. Dann gilt

$$\varphi(e_G) = e_H.$$

Beweis. Wir wissen bereits aus Satz 3.15, dass das neutrale Element einer Gruppe eindeutig bestimmt ist. Wir zeigen also, dass $\varphi(e_G)$ ein neutrales Element von $(H, \circ)$ ist und auf Grund der Eindeutigkeit folgt die Behauptung. Sei $x \in G$ beliebig. Dann gilt:

$$\varphi(x) = \varphi(x * e_G) \overset{\text{Hom.}}{=} \varphi(x) \circ \varphi(e_G).$$

Mit Lemma 3.17 folgt, dass $\varphi(e_G)$ das neutrale Element von $(H, \circ)$ ist. □

Wir sehen also, dass ein Gruppenhomomorphismus immer das neutrale Element der einen Gruppe auf dem neutralen Element der anderen Gruppe abbildet. Falls der Homomorphismus nicht injektiv ist, kann es jedoch sein, dass noch mehr Gruppenelemente auf dem neutralen Element abgebildet werden. All diese Elemente fassen wir unter folgendem Begriff zusammen:

Definition 3.23: Kern

Ist $\varphi\colon (G, *) \to (H, \circ)$ ein Gruppenhomomorphismus, so nennt man

$$\ker \varphi := \{x \in G \mid \varphi(x) = e_H\}$$

den *Kern* von φ.

Alternative Abkürzungen für den Kern im deutschsprachigen Raum sind auch $\operatorname{Kern} \varphi$ oder $\operatorname{Ker} \varphi$. Die kleingeschriebene Abkürzung leitet sich von dem englischen Begriff *kernel* ab.

Beispiel 3.24.

Der Kern des Gruppenhomomorphismus $\varphi\colon \mathbb{R} \to \mathbb{R}_{>0}$ mit $\varphi(x) = e^x$ (Beispiel 3.21) ist

$$\ker \varphi = \{x \in \mathbb{R} \mid \varphi(x) = 1\} = \{x \in \mathbb{R} \mid e^x = 1\} = \{0\}.$$

Wie bereits das Rechnen mit Rest zeigt, können wir mehr als eine Verknüpfung auf einer Menge haben – beispielsweise sowohl die Addition als auch die Multiplikation. Das führt uns zu den Begriffen *Ring* und *Körper*.

3.3 Ringe

Definition 3.25: Ringe

Seien R eine Menge und $+$ und $\cdot$ zwei Verknüpfungen auf R. Das Tripel $(R, +, \cdot)$ heißt *Ring*, falls gilt:

(R1) $(R, +)$ ist eine abelsche Gruppe, dessen neutrales Element wir 0 nennen;

(R2) $(R, \cdot)$ ist eine Halbgruppe;

(R3) Es gelten die Distributivgesetze, d.h. für $a, b, c \in R$ gilt

$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \text{und} \quad a \cdot (b + c) = a \cdot b + a \cdot c.$$

Ein Ring $(R, +, \cdot)$ heißt *kommutativ*, wenn die Multiplikation $\cdot$ kommutativ ist, d.h. wenn $a \cdot b = b \cdot a$ ist für alle $a, b \in R$.

Ein Element $1 \in R$ heißt *Einselement*, wenn es ein neutrales Element bezüglich der Multiplikation ist, d.h. wenn für alle $a \in R$ gilt, dass $1 \cdot a = a \cdot 1 = a$ ist.

Beispiel 3.26. $(\mathbb{Z}, +, \cdot)$ *ist ein kommutativer Ring mit Einselement.*

Im Folgenden wollen wir das Beispiel 3.18 wieder aufgreifen.

Beispiel 3.27. *Sei $\mathbb{F}_2 = \{0, 1\}$ mit den Operationen Addition und Multiplikation, die wie in Beispiel 3.18 definiert sind, d.h.*

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

$(\mathbb{F}_2, +, \cdot)$ *ist ein kommutativer Ring mit Einselement:*

- *Das neutrale Element der Addition ist 0, denn $0 + 0 = 0$ und $0 + 1 = 1 + 0 = 1$.*
- *Das additiv Inverse (d.h. das Negative) zu 0 ist 0 und das zu 1 ist 1.*
- *Die Addition ist kommutativ, wie man anhand der Symmetrie der Tabelle erkennt.*
- *Die Addition ist assoziativ: Man kann alle acht Kombinationen von Summen dreier Elemente von $\mathbb{F}_2$ einsetzen um zu bestätigen, dass tatsächlich $(a + b) + c = a + (b + c)$ ist für beliebige $a, b, c \in \mathbb{F}_2$.*

- *Das neutrale Element der Multiplikation ist 1, denn* $1 \cdot 0 = 0 \cdot 1 = 0$ *und* $1 \cdot 1 = 1$.
- *Die Multiplikation ist kommutativ, wie man anhand der Symmetrie der Tabelle erkennt.*
- *Die Multiplikation ist assoziativ: Man kann alle acht Kombinationen von Produkten dreier Elemente von* $\mathbb{F}_2$ *einsetzen um zu bestätigen, dass tatsächlich* $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ *ist für beliebige* $a, b, c \in \mathbb{F}_2$.
- *Die Distributivgesetze gelten: Hier kann man wieder alle acht Kombinationen aus Zahlen* $(a, b, c) \in \mathbb{F}_2^3$ *bilden und überprüfen, dass in der Tat* $(a + b) \cdot c = a \cdot c + b \cdot c$ *und* $a \cdot (b + c) = a \cdot b + a \cdot c$ *gelten.*

Definition 3.28: Nullteiler

Sei $(R, +, \cdot)$ ein kommutativer Ring und sei $a \in R$. Falls $b \in R \setminus \{0\}$ existiert, so dass $a \cdot b = 0$ ist, so heißt a *Nullteiler*. Der Ring R heißt *nullteilerfrei*, falls $0 \in R$ der einzige Nullteiler in R ist.

Lemma 3.29

In einem nullteilerfreien Ring $(R, +, \cdot)$ kann gekürzt werden, d.h. für $a, b, c \in R$ und $a \neq 0$ gilt:

$$a \cdot b = a \cdot c \quad \Rightarrow \quad b = c.$$

Beweis. Aus $a \cdot b = a \cdot c$ folgt mit dem Distributivgesetz $0 = a \cdot b - a \cdot c = a \cdot (b - c)$. Wegen $a \neq 0$ und der Nullteilerfreiheit folgt $b - c = 0$, also $b = c$. □

3.4 Körper

3.4.1 Definition und Eigenschaften

Neben den Bedeutungen, die das Wort *Körper* in der Biologie (als Gestalt eines Lebewesens) und der Geometrie (als dreidimensionale geometrische Figur) hat, gibt es eine Bedeutung im Gebiet der Algebra und um diese geht es an dieser Stelle.

Definition 3.30: Körper

Seien $\mathbb{K}$ eine Menge und $+$ und $\cdot$ zwei Verknüpfungen auf $\mathbb{K}$. Das Tripel $(\mathbb{K}, +, \cdot)$ heißt *Körper*, falls gilt:

(K1) $(\mathbb{K}, +)$ ist eine abelsche Gruppe, dessen neutrales Element wir 0 nennen;

(K2) $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe;

(K3) Es gelten die Distributivgesetze, d.h. für $a, b, c \in \mathbb{K}$ gilt

$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \text{und} \quad a \cdot (b + c) = a \cdot b + a \cdot c.$$

Beispiel 3.31.

a) $(\mathbb{R}, +, \cdot)$ *ist ein Körper.*

b) $(\mathbb{R}^{n\times n}, +, \cdot)$ *ist kein Körper, denn die Multiplikation von Matrizen ist nicht kommutativ.*

c) $(\mathbb{Z}, +, \cdot)$ *ist kein Körper, da die multiplikativ inversen Elemente in* $\mathbb{Q}$*, aber nicht immer in* $\mathbb{Z}$ *liegen.*

Beispiel 3.32 (Der endliche Körper $\mathbb{F}_2$[1])**.** *Betrachten wir nochmals* $\mathbb{F}_2 = \{0, 1\}$ *mit der Addition und Multiplikation aus Beispiel 3.27. Es handelt sich dabei sogar um einen Körper, denn das multiplikativ Inverse zu* 1 *ist* 1.

Unter Verwendung der Körpereigenschaften können wir die uns vertrauten Rechenregeln beweisen, zum Beispiel das folgende Resultat, welches wir im Anschluss benötigen werden.

Lemma 3.33

Sei $(\mathbb{K}, +, \cdot)$ ein Körper mit neutralem Element 0 bezüglich der Addition. Dann gilt

$$0 \cdot a = 0, \qquad \forall a \in \mathbb{K}.$$

Beweis. Nutzen wir einerseits, dass 0 das neutrale Element ist und andererseits, dass das Distributivgesetz gilt, so erhalten wir die Gleichungskette

$$0 \cdot a \overset{\text{(N)}}{=} (0 + 0) \cdot a \overset{\text{(DG)}}{=} 0 \cdot a + 0 \cdot a.$$

Somit ist $0 \cdot a$ ein neutrales Element für $0 \cdot a$. Aus der Eindeutigkeit des neutralen Elements folgt, dass $0 = 0 \cdot a$ ist. □

In den reellen Zahlen kennen wir die Regel, dass ein Produkt zweier Zahlen genau dann Null ist, wenn (mindestens) einer der Faktoren Null ist. Diese Eigenschaft haben alle Körper, jedoch nicht alle Ringe.

Satz 3.34: Körper sind nullteilerfrei.

Sei $(\mathbb{K}, +, \cdot)$ ein Körper. Gilt $a \cdot b = 0$ für zwei Elemente $a, b \in \mathbb{K}$, so ist mindestens eines davon 0.

Beweis. Da $\mathbb{K}$ ein Körper ist, gibt es zu $a \in \mathbb{K} \setminus \{0\}$ ein multiplikativ Inverses a^{-1}. Damit gilt dann

$$b \overset{\text{(N)}}{=} 1 \cdot b \overset{\text{(I)}}{=} (a^{-1} \cdot a) \cdot b \overset{\text{(AG)}}{=} a^{-1} \cdot (a \cdot b) \overset{\text{Vor.}}{=} a^{-1} \cdot 0 \overset{\text{Lemma 3.33}}{=} 0.$$

□

Wir haben bei der Definition des Körpers einerseits verlangt, dass $\mathbb{K}$ mit der Addition eine abelsche Gruppe ist, andererseits soll dies bezüglich der Multiplikation nur für die Menge $\mathbb{K} \setminus \{0\}$ gelten. Es stellt sich also die Frage, ob es auch Körper gibt, die diese Einschränkung nicht haben.

[1]Die Abkürzung des hier eingeführten Körpers geht auf die englische Bezeichnung für Körper zurück: *field*.

Satz 3.35

Sei $(\mathbb{K}, +, \cdot)$ ein Körper mit neutralem Element 0 bezüglich der Addition und 1 bezüglich der Multiplikation. Zudem sei $(\mathbb{K}, \cdot)$ eine abelsche Gruppe. Dann enthält $\mathbb{K}$ nur genau ein Element, d.h. $\mathbb{K} = \{0\}$.

Beweis. Wenn $(\mathbb{K}, \cdot)$ eine Gruppe ist, existiert ein zu 0 (multiplikativ) inverses Element – nennen wir es 0^{-1}. Dieses Element erfüllt die Bedingung $0^{-1} \cdot 0 = 1$.

Seien nun $a, b \in \mathbb{K}$ beliebige Elemente des Körpers. Nach Lemma 3.33 gilt $0 \cdot a = 0 = 0 \cdot b$, d.h. $0 \cdot a = 0 \cdot b$. Multiplikation (von links) mit 0^{-1} ergibt

$$0^{-1} \cdot (0 \cdot a) = 0^{-1} \cdot (0 \cdot b) \overset{\text{(AG)}}{\iff} (0^{-1} \cdot 0) \cdot a = (0^{-1} \cdot 0) \cdot b \overset{\text{Vor.}}{\iff} 1 \cdot a = 1 \cdot b \overset{\text{(N)}}{\iff} a = b.$$

Folglich besteht $\mathbb{K}$ nur aus einem einzigen Element – dem neutralen Element der Addition, welches in diesem Fall auch neutrales Element der Multiplikation ist:

$$0 + 0 = 0 \quad \text{und} \quad 0 \cdot 0 = 0.$$

□

In Kapitel 4 werden wir noch einen weiteren Körper neben $(\mathbb{R}, +, \cdot)$ und $(\mathbb{F}_2, +, \cdot)$ kennen lernen – den Körper der komplexen Zahlen. Bevor wir dieses Kapitel angehen, wollen wir jedoch endliche Körper ansehen, die analog zum $\mathbb{F}_2$ konstruiert werden. Dafür benötigen wir Äquivalenzrelationen und -klassen.

3.4.2 Exkurs: Relationen und Äquivalenzrelationen

Für zwei gegebene reelle Zahlen können wir immer sagen, welche der beiden größer ist oder ob beide gleich groß sind. Zum Beispiel ist

$$-1 < 0, \qquad \frac{1}{3} = \frac{2}{6}, \qquad 1 \leq 2 \quad \text{und} \quad \pi > 3.$$

Für Vektoren ist dies hingegen keineswegs klar: Ist $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ kleiner oder größer als $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ oder gar gleich (groß)? Wollen wir Beziehungen zwischen mathematischen Objekten darstellen, so tun wir das mit Hilfe von *Relationen*.

Definition 3.36: Relationen

Seien X und Y nichtleere Mengen. Eine Teilmenge $\mathcal{R}$ des kartesischen Produkts $X \times Y$ heißt *Relation in* $X \times Y$.

Gilt $\mathcal{R} \subseteq X \times X$, so heißt $\mathcal{R}$ *Relation in* X.

Beispiel 3.37.

- *Die Gleichheitsrelation in* $\mathbb{N}$ *lässt sich beschreiben als*

$$\mathcal{R} = \{(1,1), (2,2), (3,3), \ldots\} = \{(n,n) \mid n \in \mathbb{N}\}.$$

- *$<$, $\leq$, $=$, $\geq$ und $>$ definieren Relationen auf $\mathbb{R}$.*
- *Parallelität definiert eine Relation in der Menge aller Geraden in einer Ebene.*

Ist $\mathcal{R} \subset X \times Y$ eine Relation, und ist das Paar $(x, y) \in \mathcal{R}$, so sagt man, x *steht in Relation zu* y.

Notation

Gilt für $(x, y) \in X \times Y$, dass x in Relation steht zu y, d.h. $(x, y) \in \mathcal{R}$, so schreibt man auch $x \sim y$.

Definition 3.38: Äquivalenzrelation

Eine Relation $\sim$ in einer nichtleeren Menge X heißt *Äquivalenzrelation*, falls für beliebige $x, y, z \in X$ folgende Eigenschaften erfüllt sind:

Reflexivität: $x \sim x$,

Symmetrie: $x \sim y$ impliziert $y \sim x$,

Transitivität: $x \sim y$ und $y \sim z$ impliziert $x \sim z$.

Bemerkung 3.39. *Die Eigenschaft der Symmetrie kann auch alternativ wie folgt definiert werden:*

$$x \sim y \iff y \sim x, \quad \forall x, y \in X.$$

Diese Definition ist inhaltlich äquivalent zu der oben gegebenen Definition.

Beispiel 3.40.

- *Die Relation $<$ auf $\mathbb{R}$ ist zwar transitiv, aber weder reflexiv noch symmetrisch.*
- *Die Relation $\leq$ auf $\mathbb{R}$ ist reflexiv und transitiv, jedoch nicht symmetrisch.*
- *Die Relation $=$ auf $\mathbb{R}$ ist reflexiv, symmetrisch und transitiv – es handelt sich also um eine Äquivalenzrelation.*

Beispiel 3.41. *Sei X die Menge der in der Mensa zur Verfügung stehenden Essen. Definieren wir als Relation*

$$(x, y) \in \mathcal{R} :\iff \text{ich mag } x \text{ genau so gerne wie oder lieber als } y.$$

Diese Relation ist üblicherweise reflexiv, jedoch nicht symmetrisch. Nehmen wir an, ich mag Pasta lieber als Reis und Reis lieber als Kartoffeln, so bedeutet Transitivität, dass ich Pasta lieber mag als Kartoffeln.

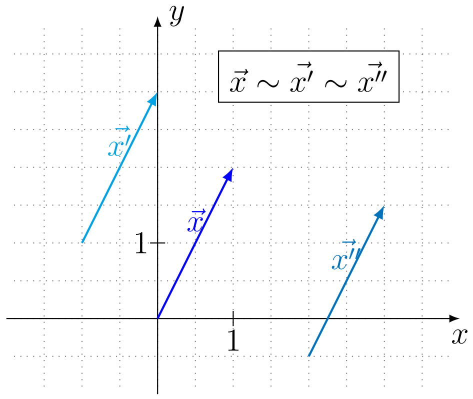
Die bisher beschriebene Relation nennt man Präferenzrelation, *wenn sie zudem die Eigenschaft der* Vollständigkeit *hat, d.h. wenn gilt:*

$$\forall x, y \in X : \ (x, y) \in \mathcal{R} \lor (y, x) \in \mathcal{R}.$$

Mit anderen Worten: Eine Relation ist vollständig auf einer Menge, wenn man für je zwei Elemente der Menge immer eines der beiden in Relation zum anderen steht.

Beispiel 3.42. *Interpretieren wir Vektoren $x \in \mathbb{R}^2$ als Pfeile mit einem Anfangspunkt, einer Richtung und einer Länge. Auf der Menge dieser Pfeile definieren wir folgende Äquivalenzrelation:*

$$\vec{x} \sim \vec{y} \;:\Longleftrightarrow\; \vec{x} \text{ und } \vec{y} \text{ haben die gleiche Richtung und Länge.}$$



Auf Grund der Transitivität sind somit alle Pfeile einer vorgegebenen Richtung und Länge äquivalent zu dem Pfeil, der im Koordinatenursprung startet und die vorgegebene Richtung und Länge hat. Somit können wir Vektoren $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ als Repräsentanten einer ganzen Klasse *äquivalenter Pfeile interpretieren. Alle zueinander äquivalenten Pfeile haben gemeinsam, dass die Differenz zwischen End- und Anfangspunkt gerade den Vektor x ergeben.*

Den Ansatz, alle (in Bezug auf eine fixierte Äquivalenzrelation) zueinander äquivalenten Elemente einer Menge zu einer *Äquivalenzklasse* zusammenzufassen, wollen nun formalisieren:

Definition 3.43: Äquivalenzklassen

Sei $\sim$ eine Äquivalenzrelation in einer nichtleeren Menge X. Die *Äquivalenzklasse* eines Elements $x_0 \in X$ ist definiert als

$$[x_0] := \{x \in X \mid x \sim x_0\}.$$

Soll die Relation hervorgehoben werden, so schreibt man auch $[x_0]_\sim$.

Elemente einer Äquivalenzklasse werden *Vertreter* oder *Repräsentanten* der Äquivalenzklasse genannt.

Die Menge aller Äquivalenzklassen einer Äquivalenzrelation $\sim$ in einer Menge X,

$$X/\sim\; := \{[x]_\sim \mid x \in X\},$$

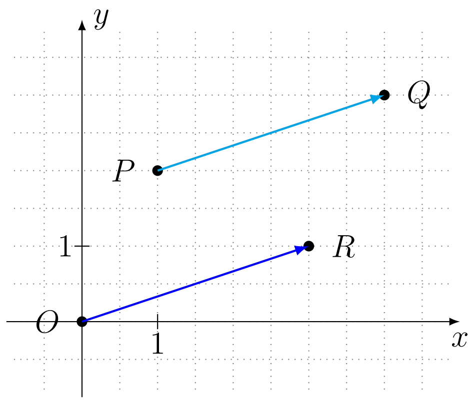
wird als *Faktormenge* oder *Quotientenmenge* bezeichnet.

Die Menge aller Pfeile gleicher Länge und Richtung bilden eine Äquivalenzklasse, welche durch einen Vektor des $\mathbb{R}^n$ repräsentiert wird. Die Menge der Vektoren können wir also als Quotientenmenge der in Beispiel 3.42 eingeführten Relation auf der Menge der Pfeile verstehen.

Beispiel 3.44 (Fortsetzung von Beispiel 3.42)**.** *Seien*

$$O = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad P = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad Q = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \quad \text{und} \quad R = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Punkte im Raum $\mathbb{R}^2$. Dann ist $\overrightarrow{PQ} \sim \overrightarrow{OR}$, d.h. $\overrightarrow{PQ} \in [\overrightarrow{OR}]$.



Alle Pfeile dieser Äquivalenzklasse werden durch den Vektor

$$v = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

repräsentiert. Der Pfeil $\overrightarrow{OR}$ *heißt auch* Ortspfeil *oder* Ortsvektor *zum Vektor* v. *Unter allen zueinander äquivalenten Pfeilen wird so der im Koordinatenursprung startende Pfeil hervorgehoben.*

Satz 3.45

Sei $\sim$ eine Äquivalenzrelation in einer nichtleeren Menge X. Dann gilt:

1. Keine der durch $\sim$ erzeugten Äquivalenzklassen ist leer, d.h.
$$[x] \neq \emptyset, \quad \forall x \in X.$$
2. Sind zwei Äquivalenzklassen nicht gleich, so sind sie disjunkt, d.h.
$$[x] \neq [y] \implies [x] \cap [y] = \emptyset.$$
3. X ist die Vereinigung aller Äquivalenzklassen bezüglich $\sim$, d.h.
$$X = \bigcup_{x \in X} [x].$$

Beweis.

1. Als Äquivalenzrelation ist $\sim$ insbesondere reflexiv. Somit ist für jedes $x \in X$ insbesondere $x \in [x]$, so dass jede Äquivalenzklasse mindestens ein Element enthält.
2. Zeigen wir die Aussage durch Kontraposition. Wir nehmen an, es sei $[x] \cap [y] \neq \emptyset$, d.h. es existiert ein $z \in [x] \cap [y]$. Dann gilt insbesondere $z \sim x$ und $z \sim y$. Wegen Symmetrie und Transitivität von $\sim$ gilt somit auch $x \sim y$. Mit Symmetrie und Transitivität haben wir also für alle $x' \in [x]$ und $y' \in [y]$:
$$\begin{array}{lllllll} x' \sim x & \text{und} & x \sim y & \implies & x' \sim y & \implies & x' \in [y], \\ y' \sim y & \text{und} & x \sim y & \implies & y' \sim x & \implies & y' \in [x], \end{array}$$
d.h. $[x] \subseteq [y]$ und $[y] \subseteq [x]$. Zusammenfassend haben wir also $[x] = [y]$.
3. Es ist $x \in [x]$ für jedes $x \in X$ und man hat immer $X = \bigcup_{x \in X} x$, also folgt $X \subseteq \bigcup_{x \in X} [x]$. Andererseits ist per Definition $[x] \subseteq X$ für jedes $x \in X$ und somit ist auch $\bigcup_{x \in X} [x] \subseteq X$. Insgesamt folgt somit die behauptete Gleichheit der Mengen.

$\square$

3.4.3 Endliche Körper

Wir hatten bereits $(\mathbb{Z}, +, \cdot)$ als Ring kennen gelernt. Er ist als Teilmenge des Körpers $(\mathbb{R}, +, \cdot)$ nullteilerfrei (Vgl. Satz 3.34), d.h. sind $a, b \in \mathbb{Z}^* := \mathbb{Z}\setminus\{0\}$, so ist auch $a \cdot b \in \mathbb{Z}^*$.

Wir wollen Restklassenringe definieren. Dafür müssen wir zunächst geeignete Notationen einführen und nochmals das Rechnen mit Rest formalisieren.

Notation

Für $n \in \mathbb{N}$ seien

$$\begin{aligned} n\mathbb{Z} &:= \{n \cdot k \mid \; k \in \mathbb{Z}\}\,, \\ \mathbb{Z}_n &:= \{0, 1, \ldots, n-1\}, \\ r + n\mathbb{Z} &:= \{r + n \cdot k \mid \; k \in \mathbb{Z}\}\,, \quad r \in \mathbb{Z}_n. \end{aligned}$$

Für ein $n \in \mathbb{N}$ führen wir nun folgende Äquivalenzrelation auf $\mathbb{Z}$ ein:

$$a \sim_n b \quad :\Longleftrightarrow \quad |b - a| \in n\mathbb{Z}.$$

Man könnte diese Anforderung auch so umformulieren, dass $|b - a|$ durch n teilbar sei, d.h. es existiert eine Zahl $k \in \mathbb{Z}$ derart, dass $|b - a| = n \cdot k$ ist. Es ist eine Übungsaufgabe, die Eigenschaften der Äquivalenzrelation zu überprüfen.

Beispiel 3.46. *Für* $n = 6$ *ist* $3 \sim_6 -9$, *denn* $|3 - (-9)| = 12 = 6 \cdot 2 \in 6\mathbb{Z}$.

Ebenso ist $6 \sim_6 0$, *denn* $|6 - 0| = 6 \in 6\mathbb{Z}$.

Es ist jedoch $8 \not\sim_6 5$, *denn* $|8 - 5| = 3 \notin 6\mathbb{Z}$.

Lemma 3.47

Mit obiger Notation gilt für festes $n \in \mathbb{N}$ und jedes $a \in \mathbb{Z}_n$ bezüglich der Äquivalenzrelation $\sim_n$:

$$[a] = a + n\mathbb{Z} \tag{3.4.1}$$

und insbesondere

$$[0] = n\mathbb{Z}. \tag{3.4.2}$$

Beweis. Sei $b \in a + n\mathbb{Z} = \{a + k \cdot n \mid \; k \in \mathbb{Z}\}$, d.h. $b = a + k_b \cdot n$ für ein $k_b \in \mathbb{Z}$. Dann ist $|b - a| = |k_b| \cdot n \in n\mathbb{Z}$, d.h. nach Definition ist $a \sim_n b$ und somit $b \in [a]$.

Ist andererseits $b \in [a]$, d.h. $a \sim_n b$, also $|b - a| \in n\mathbb{Z}$, so existiert $k \in \mathbb{Z}$ derart, dass $|b - a| = |k| \cdot n$, also ist $b \in a + n\mathbb{Z}$.

Damit haben wir (3.4.1) gezeigt. Gleichung (3.4.2) folgt aus (3.4.1) mit $a = 0$. □

Die Quotientenmenge bezüglich dieser Äquivalenzrelation notieren wir damit wie folgt:

Notation

Für $n \in \mathbb{N}$ ist

$$\mathbb{Z}/n\mathbb{Z} := \mathbb{Z}/\sim_n = \{[a] \mid a \in \mathbb{Z}\}$$

die Menge der *Restklassen modulo* n.

Um leichter mit Resten rechnen zu können, definieren wir $\varrho_n \colon \mathbb{Z} \to \mathbb{Z}_n$ via

$$r = \varrho_n(a) \quad :\iff \quad a \sim_n r \quad \iff \quad a \in [r]. \tag{3.4.3}$$

Satz 3.45 sagt uns, dass jede ganze Zahl in genau einer Restklasse liegt, so dass diese Funktion wohldefiniert ist.

Beispiel 3.48. *Für $n = 6$ ist $\varrho_6(-9) = 3$, denn 3 ist die einzige Zahl aus $\mathbb{Z}_n$, zu der -9 bezüglich $\sim_6$ äquivalent ist.*

Zudem ist $\varrho_6(3) = 3$, denn $3 \in \mathbb{Z}_6$ und $3 \sim_6 3$.

Lemma 3.49

Die in (3.4.3) definierte Abbildung $\varrho_n \colon \mathbb{Z} \to \mathbb{Z}_n$ hat folgende Eigenschaften:

a) $\varrho_n(a) = a \iff a \in \mathbb{Z}_n$

b) $\varrho_n(a) = \varrho_n(b) \iff a - b \in n\mathbb{Z}$

c) $\varrho_n(a+b) = \varrho_n(\varrho_n(a) + \varrho_n(b))$ und $\varrho_n(a \cdot b) = \varrho_n(\varrho_n(a) \cdot \varrho_n(b)), \quad a, b \in \mathbb{Z}$.

Beweis.

a) Wenn $\varrho_n(a) = a$ ist, dann ist nach Definition $a \sim_n a$, was für alle $a \in \mathbb{Z}_n$ der Fall ist, da $\sim_n$ eine Äquivalenzrelation ist. Andersherum folgt aus $a \in \mathbb{Z}_n = \{0, \ldots, n-1\}$, dass $0 \leq a < n$ ist und somit $\varrho_n(a) = a$ ist.

b) $\varrho_n(a) = \varrho_n(b)$ ist genau dann der Fall, wenn $r \in \mathbb{Z}_n$, sowie $k_a, k_b \in \mathbb{Z}$ derart existieren, dass $a = k_a \cdot n + r$ und $b = k_b \cdot n + r$ ist. Das impliziert jedoch $a - b = (k_a - k_b) \cdot n$, also $a - b \in n\mathbb{Z}$. Ist andersherum $a - b \in n\mathbb{Z}$, d.h. $a - b = k \cdot n$ für ein $k \in \mathbb{Z}$, so ist folglich für $b = k_b \cdot n + r$ nun $a = b + k \cdot n = (k_b \cdot n + r) + k \cdot n = (k_b + k) \cdot n + r$, also $\varrho_n(a) = r = \varrho_n(b)$.

c) Seien $r_a := \varrho_n(a)$ und $r_b := \varrho_n(b)$. Dann existieren $k_a, k_b \in \mathbb{Z}$ derart, dass

$$\begin{aligned} a &= k_a \cdot n + r_a \quad \text{und} \\ b &= k_b \cdot n + r_b, \end{aligned}$$

d.h. insbesondere ist $r_a = \varrho_n(a)$ und $r_b = \varrho_n(b)$. Folglich ist

$$\begin{aligned} a + b &= (k_a + k_b) \cdot n + (r_a + r_b), \\ a \cdot b &= (k_a \cdot k_b \cdot n + k_a \cdot r_b + k_b \cdot r_a) \cdot n + (r_a \cdot r_b). \end{aligned}$$

Somit sind

- $(a+b)-(r_a+r_b) \in n\mathbb{Z}$, also nach b) $\varrho_n(a+b) = \varrho_n(r_a+r_b) = \varrho_n(\varrho_n(a)+\varrho_n(b))$;
- $(a \cdot b) - (r_a \cdot r_b) \in n\mathbb{Z}$, also nach b) ebenfalls $\varrho_n(a \cdot b) = \varrho_n(r_a \cdot r_b)$, woraus die Behauptung folgt.

□

Notation

Gilt $\varrho_n(a) = \varrho_n(b)$ für zwei Zahlen $a, b \in \mathbb{Z}$, so heißt a *kongruent* b *modulo* n und man schreibt auch $a \equiv b \pmod{n}$.

Satz 3.50: Restklassenring

Sei $n \in \mathbb{N}$. Für $a, b \in \mathbb{Z}_n$ seien Summe und Produkt wie folgt definiert:

$$\begin{aligned} a +_n b &:= \varrho_n(a+b), \\ a \cdot_n b &:= \varrho_n(a \cdot b). \end{aligned}$$

Damit ist $(\mathbb{Z}_n, +_n, \cdot_n)$ ein kommutativer Ring mit Einselement.

Beispiel 3.51. *Für* $n = 6$ *ist* $(-5) +_6 (-4) = \varrho_6(-9) = 3$ *und* $2 \cdot_6 3 = \varrho_6(6) = 0$.

Beweis. Zunächst ist $(\mathbb{Z}_n, +_n)$ eine abelsche Gruppe mit neutralem Element $0 \in \mathbb{Z}_n$:

- Die Verknüpfung $+_n$ ist in $\mathbb{Z}_n$ abgeschlossen, denn zu $a, b \in \mathbb{Z}_n$ ist $a +_n b = \varrho_n(a+b) \in \mathbb{Z}_n$, da ϱ_n nur Werte in $\mathbb{Z}_n$ annimmt.
- $+_n$ ist assoziativ: Sind $a, b, c \in \mathbb{Z}_n$, so ist (unter Verwendung von a) und c) von Lemma 3.49)

$$\begin{aligned} a +_n (b +_n c) &= \varrho_n(a) +_n \varrho_n(b+c) = \varrho_n(\varrho_n(a) + \varrho_n(b+c)) = \varrho_n(a + (b+c)) \\ &= \varrho_n((a+b)+c) = \varrho_n(\varrho_n(a+b) + \varrho_n(c)) = \varrho_n(a+b) +_n \varrho_n(c) \\ &= (a +_n b) +_n c. \end{aligned}$$

- $+_n$ ist kommutativ: Für $a, b \in \mathbb{Z}_n$ ist $a +_n b = \varrho_n(a+b) = \varrho_n(b+a) = b +_n a$.
- Das neutrale Element ist $0 \in \mathbb{Z}_n$, denn für alle $a \in \mathbb{Z}_n$ ist $a +_n 0 = \varrho_n(a+0) = \varrho_n(a) = a$.
- Jedes Element besitzt ein Inverses: Ist $a \in \mathbb{Z}_n$, so ist $n - a \in \mathbb{Z}_n$ das zugehörige Inverse, denn $a +_n (n-a) = \varrho_n(a+n-a) = \varrho_n(n) = 0$.

Des Weiteren ist $(\mathbb{Z}_n, \cdot_n)$ eine kommutative Halbgruppe mit Einselement $1 \in \mathbb{Z}_n$. (Aufgabe 3.5) Letztlich gelten auch die Distributivgesetze, von denen wir wegen der Kommutativität beider Verknüpfungen nur eines zeigen müssen: Für $a, b, c \in \mathbb{Z}_n$ gilt

$$\begin{aligned} a \cdot_n (b +_n c) &= \varrho_n(a) \cdot_n \varrho_n(b+c) = \varrho_n(\varrho_n(a) \cdot \varrho_n(b+c)) = \varrho_n(a \cdot (b+c)) \\ &= \varrho_n(a \cdot b + a \cdot c) = \varrho_n(\varrho_n(a \cdot b) + \varrho_n(a \cdot c)) = \varrho_n(a \cdot_n b + a \cdot_n c) \\ &= (a \cdot_n b) +_n (a \cdot_n c). \end{aligned}$$

□

Satz 3.52

Der Ring $(\mathbb{Z}_n, +_n, \cdot_n)$ $(n \in \mathbb{N})$ ist genau dann ein nullteilerfrei, wenn n eine Primzahl oder 1 ist. Nur in diesem Fall ist es ein Körper.

Beweis. $(\mathbb{Z}_1, +_1, \cdot_1)$ ist ein Körper, denn es ist $\mathbb{Z}_1 = \{0\}$ mit den Operationen $0 +_1 0 = 0$ und $0 \cdot_1 0 = 0$ der Körper aus Satz 3.35.

$(\mathbb{Z}_2, +_2, \cdot_2)$ ist nullteilerfrei und es ist ein Körper, es ist nämlich gerade $\mathbb{F}_2$ aus den Beispielen 3.27 und 3.32.

Sei $n \geq 3$. Ist n keine Primzahl, so existieren $a, b \in \mathbb{Z}_n \setminus \{0\}$ derart, dass $n = a \cdot b$, d.h.

$$a \cdot_n b = \varrho_n(n) = 0.$$

Damit ist gezeigt, dass $\mathbb{Z}_n$ nicht nullteilerfrei ist. Aus Satz 3.34 folgt, dass dieser Ring somit kein Körper sein kann.

Sei nun n eine Primzahl. Angenommen, $\mathbb{Z}_n$ wäre nicht nullteilerfrei, d.h. es existieren $a, b \in \mathbb{Z}_n \setminus \{0\}$ (also insbesondere $a, b < n$) derart, dass $a \cdot_n b = 0$. Dann wäre $\varrho_n(a \cdot b) = 0$, d.h. $a \cdot b$ ist durch n teilbar. Da n eine Primzahl ist, muss somit bereits a oder b durch n teilbar sein (oder beide). Das widerspricht jedoch $a, b < n$. Folglich ist unsere Annahme widerlegt, d.h. $\mathbb{Z}_n$ ist nullteilerfrei.

Es bleibt zu zeigen, dass für eine Primzahl $n \in \mathbb{N}$ der nullteilerfreie Ring $(\mathbb{Z}_n, +_n, \cdot_n)$ sogar ein Körper ist. Wir haben bereits gezeigt, dass der Ring kommutativ ist und das Einselement $1 \in \mathbb{Z}_n$ enthält. Es bleibt also nur noch zu zeigen, dass jedes Element $a \in \mathbb{Z}_n^* := \mathbb{Z}_n \setminus \{0\}$ ein multiplikativ inverses Element $b \in \mathbb{Z}_n^*$ enthält, so dass $a \cdot_n b = 1$ gilt.

Betrachten wir dazu die Menge $M := \{a \cdot_n z \mid z \in \mathbb{Z}_n^*\}$. Da $a \neq 0$ ist und $\mathbb{Z}_n$ nullteilerfrei ist, ist $M \subseteq \mathbb{Z}_n^*$. Wir wollen nun zeigen, dass sogar Gleichheit der Mengen gilt. Da die Mengen endlich sind (mit $|\mathbb{Z}_n^*| = n - 1$), genügt es dazu, zu zeigen, dass alle Elemente der Form $a \cdot_n z$ für $z \in \mathbb{Z}_n^*$ verschieden sind. Seien dazu $z, z' \in \mathbb{Z}_n^*$ und angenommen es gilt $a \cdot_n z = a \cdot_n z'$. Da $\mathbb{Z}_n$ nullteilerfrei ist, können wir die Kürzungsregel aus Lemma 3.29 verwenden, d.h. es folgt $z = z'$. Da $M = \mathbb{Z}_n^*$ ist, gibt es also ein $z \in \mathbb{Z}_n*$, so dass $a \cdot_n z = 1$ ist, d.h. es existiert ein zu a inverses Element in $\mathbb{Z}_n^*$. □

Bemerkung 3.53. *Wir haben uns nur sehr kurz mit endlichen Körpern beschäftigt. Wer mehr wissen möchte, findet dies u.a. in [Kur08].*

3.5 Aufgaben

Aufgabe 3.1. *Sei* $M = \{1, 2, 3\}$.

a) *Notieren Sie alle Bijektionen von* M *nach* M. *Verwenden Sie folgende Notation:*

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix},$$

falls $1 \mapsto a$, $2 \mapsto b$ *und* $3 \mapsto c$. *Die Menge aller solcher Bijektionen bezeichnen wir mit* S_3.

b) *Wir definieren die Verknüpfung* $\circ \colon S_3 \times S_3 \to S_3$ *als*

$$(\sigma \circ \tau)(m) := \sigma(\tau(m)), \quad m \in M.$$

Stellen Sie die Verknüpfungstafel auf:

	τ	$\cdots$
σ	$\sigma \circ \tau$	
$\vdots$		

c) *Überprüfen Sie, ob es sich bei* $(S_3, \circ)$ *um eine Gruppe handelt. Ist die Verknüpfung kommutativ?*

Aufgabe 3.2. *Entscheiden Sie, welche der Eigenschaften (AG), (KG), (N), (I) von Gruppen folgende Paare aus Menge* M *und Verknüpfung* $*$ *besitzen:*

a) $M = \mathbb{Q}$, $x * y := x + y$,

b) $M = \mathbb{Z}$, $x * y := x \cdot y$,

c) $M = \mathbb{R}^{m \times n}$, $A * B := A + B$,

Aufgabe 3.3. *Zeigen Sie, dass die Menge aller Abbildungen* $\mathrm{Abb}(\mathbb{R}; \mathbb{R}) := \{f \colon \mathbb{R} \to \mathbb{R}\}$ *mit der punktweisen Addition*

$$(f + g)(x) := f(x) + g(x)$$

eine abelsche Gruppe ist.

Aufgabe 3.4. *Zeigen Sie:* $(\mathbb{Z}, +, \cdot)$ *ist ein kommutativer Ring mit Einselement. Ist es ein Körper?*

Aufgabe 3.5. *Vervollständigen Sie den Beweis von Satz 3.50, d.h. zeigen Sie, dass* $(\mathbb{Z}_n, \cdot_n)$ *eine kommutative Halbgruppe mit Einselement ist.*

Aufgabe 3.6. *Sei* $\sim$ *die auf* $(\mathbb{Z} \times \mathbb{Z} \setminus \{0\})$ *wie folgt definierte Relation:*

$$(a, b) \sim (c, d) :\iff a \cdot d = c \cdot b.$$

Zeigen Sie, dass dies eine Äquivalenzrelation ist.

Aufgabe 3.7. *Für die Äquivalenzrelation aus der vorherigen Aufgabe ist eine Äquivalenzklasse wie folgt definiert:*

$$\frac{a}{b} := \{(m, n) \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\} \mid (m, n) \sim (a, b)\}.$$

Für diese Äquivalenzklassen definieren wir eine Addition

$$\frac{a}{b} \oplus \frac{c}{d} := \frac{a \cdot d + c \cdot b}{b \cdot d},$$

wobei $+$ *und* $\cdot$ *für die übliche Addition und Multiplikation ganzer Zahlen stehen.*

a) *Bestimmen Sie* $\frac{1}{3} \oplus \frac{1}{2}$ *und* $\frac{2}{6} \oplus \frac{3}{6}$.

b) *Zeigen Sie, dass die Addition* $\oplus$ *auf den Äquivalenzklassen* repräsentantenunabhängig *ist, d.h. dass für alle* $(a_1, b_1), (a_2, b_2), (c_1, d_1), (c_2, d_2) \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ *aus* $(a_1, b_1) \sim (a_2, b_2)$ *und* $(c_1, d_1) \sim (c_2, d_2)$ *folgt, dass* $(a_1 \cdot d_1 + c_1 \cdot b_1, b_1 \cdot d_1) \sim (a_2 \cdot d_2 + c_2 \cdot b_2, b_2 \cdot d_2)$ *ist. Mit anderen Worten, es gilt:*

$$\left[\frac{a_1}{b_1} = \frac{a_2}{b_2}\right] \wedge \left[\frac{c_1}{d_1} = \frac{c_2}{d_2}\right] \Rightarrow \left[\frac{a_1}{b_1} \oplus \frac{c_1}{d_1} = \frac{a_2}{b_2} \oplus \frac{c_2}{d_2}\right]$$

Kapitel 4

Der Fundamentalsatz der Algebra

Ziele

- ☐ mit komplexen Zahlen in der kartesischen Darstellung rechnen können und dazu die zugehörigen Begriffe kennen
- ☐ Eigenschaften komplexer Zahlen beweisen
- ☐ sämtliche Nullstellen von Polynomen in $\mathbb{R}$ oder $\mathbb{C}$ berechnen, Polynome so in lineare oder quadratische Faktoren zerlegen
- ☐ Hintergrundwissen zu den Beweisen zur Anzahl der Nullstellen von Polynomen und zu der Faktorisierung besitzen

4.1 Der Körper der komplexen Zahlen

Wieso benötigen wir die komplexen Zahlen?

- Für $m, n \in \mathbb{N}$ ist $m + n \in \mathbb{N}$, aber möglicherweise ist $m - n \in \mathbb{Z} \setminus \mathbb{N}$
- Für $x, y \in \mathbb{Z}$ sind $x + y, x - y, x \cdot y \in \mathbb{Z}$, aber möglicherweise ist $\frac{x}{y} \in \mathbb{Q} \setminus \mathbb{Z}$
- Für $x, y \in \mathbb{Q}$ sind $x + y, x - y, x \cdot y, \frac{x}{y}, x^2 \in \mathbb{Q}$ (abgesehen von der Division durch 0), aber möglicherweise ist $\sqrt{|x|} \in \mathbb{R} \setminus \mathbb{Q}$

Die komplexen Zahlen ermöglichen es, $\sqrt{x}$ auch für negative reelle Zahlen x zu berechnen. Beispielsweise hat die Gleichung $x^2 = -1$, welche innerhalb von $\mathbb{R}$ keine Lösung besaß, im Raum der komplexen Zahlen zwei Lösungen.

Für komplexe Zahlen gibt es verschiedene Darstellungsmöglichkeiten – wir führen zunächst die kartesische Darstellung[1] ein.

[1]Das kartesische Koordinatensystem ist benannt nach dem französischen Mathematiker René Descartes (1596-1650). Die Achsen stehen in einem rechten Winkel aufeinander, die Position eines Punktes wird anhand von Parallelen zu den Koordinatenachsen angegeben. Anders funktioniert das Polarkoordinatensystem, bei dem Punkte durch ihren Abstand zu einem Festen Punkt (Pol) und den Winkel zu einer festen Richtung angegeben werden.

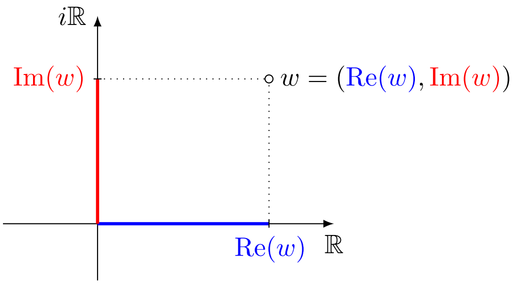
Definition 4.1: kartesische Darstellung

Die *imaginäre Einheit* i ist definiert als eine Lösung der Gleichung $x^2 = -1$. Wir schreiben dies als $i = \sqrt{-1}$.

Die Menge der komplexen Zahlen ist $\mathbb{C} := \{a + ib \mid a, b \in \mathbb{R}\}$. Für $w = a + ib \in \mathbb{C}$ nennt man $a = \operatorname{Re}(w) \in \mathbb{R}$ den *Realteil* von w und $b = \operatorname{Im}(w) \in \mathbb{R}$ den *Imaginärteil* von w. Die Darstellung in der Form $w = \operatorname{Re}(w) + i \operatorname{Im}(w)$ heißt kartesische Darstellung der komplexen Zahl w.

Die Notation variiert in der Literatur: Für $\operatorname{Re}(w)$ schreibt man auch $\Re(w)$, für $\operatorname{Im}(w)$ schreibt man auch $\Im(w)$.

Komplexe Zahlen kann man auch als Tupel schreiben mit $w = (\operatorname{Re}(w), \operatorname{Im}(w)) = (x, y) \in \mathbb{R}^2$. So lassen sich komplexe Zahlen in der *Gauß'schen Zahlenebene* (auch *komplexe Zahlenebene* genannt) darstellen.



Die rellen Zahlen kann man als Teilmenge der komplexen Zahlen auffassen:

$$\mathbb{R} = \{w \in \mathbb{C} \mid \operatorname{Im}(w) = 0\}.$$

Auf $\mathbb{C}$ kann man Addition und Multiplikation wie folgt definieren:

Definition 4.2

Für $w = a + ib \in \mathbb{C}$ und $z = c + id \in \mathbb{C}$ ist

$$\begin{aligned} w + z &:= (a + ib) + (c + id) = (a + c) + i \cdot (b + d), \\ w \cdot z &:= (a + ib) \cdot (c + id) = ac + iad + ibc + i^2 bd = (ac - bd) + i(bc + ad). \end{aligned}$$

Man prüft leicht nach, dass für zwei reelle Zahlen (d.h. für $b = d = 0$) die übliche Addition und Multiplikation entsteht.

Beispiel 4.3. *Seien $w = 1 + i$ und $z = 2 - i$. Dann sind nach obiger Notation $a = 1, b = 1, c = 2$ und $d = -1$. Summe und Produkt von w und z sind:*

$$\begin{aligned} w + z &= (1 + i) + (2 - i) = (1 + 2) + i \cdot (1 - 1) = 3, \\ w \cdot z &= (1 + i) \cdot (2 - i) = 2 - i + 2i - i^2 = 2 + i - (-1) = 3 + i. \end{aligned}$$

Ist $w = a + ib$, dann kann man $w^{-1} = \frac{1}{w}$ bestimmen, indem man ausnutzt, dass $w \cdot \frac{1}{w} = 1$ ist. Sei also $w^{-1} := c + id$. Nach obiger Rechenregel ist dann

$$w \cdot \frac{1}{w} = (ac - bd) + i(bc + ad) \stackrel{!}{=} 1 + i \cdot 0.$$

Die reellen Zahlen c und d müssen also ein lineares Gleichungssystem erfüllen:

$$\begin{aligned} ac - bd &= 1 \\ bc + ad &= 0. \end{aligned}$$

Löst man dies, erhält man folgendes Resultat:

Lemma 4.4

Für $w \in \mathbb{C} \setminus \{0\}$ mit $w = a + ib$ (mit $a, b \in \mathbb{R}$) ist

$$\frac{1}{w} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i.$$

Alternativ zur Herleitung über das Gleichungssystem kann man zur Berechnung von $\frac{1}{a+ib}$ auch die 3. Binomische Formel geschickt anwenden:

$$\frac{1}{a + ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2}.$$

Beispiel 4.5. *Sei $w = 3 + 4i$. Dann ist $\frac{1}{w} = \frac{3}{25} - \frac{4}{25} i$. Dass dies korrekt ist, können wir durch Multiplikation überprüfen:*

$$w \cdot \frac{1}{w} = (3 + 4i) \cdot \left(\frac{3}{25} - \frac{4}{25} i \right) = \frac{9}{25} - \frac{12}{25} i + \frac{12}{25} i - \frac{16}{25} i^2 = \frac{9}{25} + \frac{16}{25} = 1.$$

Satz 4.6

Die Menge der komplexen Zahlen bildet zusammen mit den eingeführten Verknüpfungen Addition und Multiplikation den Körper $(\mathbb{C}, +, \cdot)$.

Beweis. Einige Aspekte wurden gerade gezeigt. In Gänze ist dies der Inhalt einer Übungsaufgabe. □

Definition 4.7

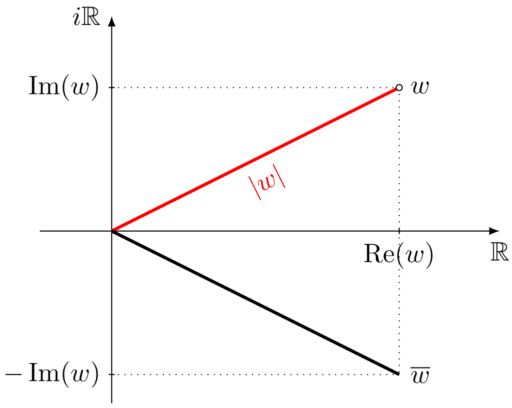
Für $w \in \mathbb{C}$ mit $w = a + ib$ ist der Betrag von w definiert als $|w| := \sqrt{a^2 + b^2}$ und die *konjugiert komplexe Zahl* zu w ist $\overline{w} := a - ib$.

Bemerkung 4.8. *Nicht nur die Addition und Multiplikation in $\mathbb{C}$ ergeben, angewendet auf reelle Zahlen, die aus $\mathbb{R}$ bekannten Verknüpfungen Addition und Multiplikation. Auch der Betrag einer Zahl $w \in \mathbb{R} \subset \mathbb{C}$ ist der bereits bekannte Betrag einer reellen Zahl: Ist $w = a + 0i \in \mathbb{R}$, so ist*

$$\begin{aligned} |w| = \sqrt{a^2 + 0^2} &= \begin{cases} a & \text{, falls } a \geq 0 \text{ ist,} \\ -a & \text{, falls } a < 0 \text{ ist} \end{cases} \\ &= |a|. \end{aligned}$$

Bemerkung 4.9. *Wie kann man die beiden neuen Begriffe in der Gauß'schen Zahlenebene veranschaulichen?*

Einerseits erhält man die zu $w \in \mathbb{C}$ *konjugiert komplexe Zahl* $\overline{w}$ *durch Spiegelung an der reellen Achse. Andererseits ist der Betrag von* w *gerade der Abstand vom Koordinatenursprung zum Punkt* $(\mathrm{Re}(w), \mathrm{Im}(w))$ *in der Gauß'schen Zahlenebene.*



Satz 4.10

Seien $w, z \in \mathbb{C}$. Dann gilt:

(i) $w \in \mathbb{R} \subset \mathbb{C} \iff \mathrm{Im}(w) = 0 \iff w = \bar{w}$,

(ii) $w + \overline{w} = 2\,\mathrm{Re}(w)$,

(iii) $|w| = \sqrt{w \cdot \overline{w}}$ und damit $|w|^2 = w \cdot \overline{w} = \mathrm{Re}(w)^2 + \mathrm{Im}(w)^2$,

(iv) $|\,\mathrm{Re}(w)| \leq |w|$ und $|\,\mathrm{Im}(w)| \leq |w|$,

(v) $\overline{w + z} = \overline{w} + \overline{z}$,

(vi) $\overline{w \cdot z} = \overline{w} \cdot \overline{z}$,

(vii) $|w \cdot z| = |w| \cdot |z|$,

(viii) $|w + z| \leq |w| + |z|$ (Dreiecksungleichung).

Beweis.

(i) Mit $\mathbb{R} = \{w \in \mathbb{C} \,|\, \mathrm{Im}(w) = 0\}$ folgt die erste Äquivalenz. Sei nun $w = a + ib$ mit $a, b \in \mathbb{R}$. Es gilt $w = \overline{w} \iff a + ib = a - ib \iff b = 0 \iff \mathrm{Im}(w) = 0$, da nach Definition ist $b = \mathrm{Im}(w)$ ist. Somit ist auch die zweite Äquivalenz gezeigt.

(ii) Sei $w = a + ib$ mit $a = \mathrm{Re}(w)$ und $b = \mathrm{Im}(w)$. Dann ist

$$w + \overline{w} = (a + ib) + (a - ib) = (a + a) + i(b - b) = 2a = 2\,\mathrm{Re}(w).$$

(iii) Der Betrag einer komplexen Zahl $w = a + ib$ ist gemäß Definition $|w| = \sqrt{a^2 + b^2}$. Somit ist

$$\sqrt{w \cdot \overline{w}} = \sqrt{(a + ib) \cdot (a - ib)} = \sqrt{(a^2 - b \cdot (-b) + i(ba + a \cdot (-b))} = \sqrt{a^2 + b^2} = |w|.$$

Durch Quadrieren folgt hieraus

$$|w|^2 = w \cdot \overline{w} = (a + ib) \cdot (a - ib) = a^2 + b^2 = \operatorname{Re}(w)^2 + \operatorname{Im}(w)^2.$$

(iv) Das Quadrat einer reellen Zahl ist immer nichtnegativ. Somit ist für $w = a + ib$ mit $a, b \in \mathbb{R}$

$$\begin{aligned} |w| &= \sqrt{a^2 + b^2} \geq \sqrt{a^2} = |a| = |\operatorname{Re}(w)| \quad \text{und} \\ |w| &= \sqrt{a^2 + b^2} \geq \sqrt{b^2} = |b| = |\operatorname{Im}(w)|. \end{aligned}$$

Hierbei benutzen wir, dass die Wurzelfunktion $x \mapsto \sqrt{x}$ $(x \in \mathbb{R})$ monoton wachsend ist. Dies ist entweder aus der Schule bekannt oder wird in Analysis I gezeigt.

(viii) Man rechnet mit dem Quadrat der linken Seite und zieht am Ende die Wurzel auf beiden Seiten der (Un-)gleichungskette

$$\begin{aligned} |w + z|^2 &= (w + z) \cdot \overline{w + z} \\ &= (w + z) \cdot (\overline{w} + \overline{z}) \\ &= w\overline{w} + z\overline{z} + w\overline{z} + z\overline{w} \\ &= |w|^2 + |z|^2 + w\overline{z} + \overline{\overline{z}} \cdot \overline{w} \qquad (\text{wegen } \overline{\overline{z}} = z \text{ und (iii)}) \\ &= |w|^2 + |z|^2 + 2\operatorname{Re}(w\overline{z}) \qquad (\text{wegen (ii) und (vi)}) \\ &\leq |w|^2 + |z|^2 + 2|w\overline{z}| \qquad (\text{wegen (iv)}) \\ &= |w|^2 + |z|^2 + 2|w||z| \qquad (\text{wegen (vii) und } |z| = |\overline{z}|) \\ &= (|w| + |z|)^2. \end{aligned}$$

Die nicht bewiesenen Aussagen sind als Übung zu zeigen. □

Kommen wir noch einmal zurück zum Ausgangsproblem: Wir wollen die Wurzel aus einer negativen Zahl ziehen können. Bisher haben wir nur $\sqrt{-1} = i$ gesetzt. Dies wollen wir nun auf andere negative Zahlen verallgemeinern:

Lemma 4.11

Sei $c \in \mathbb{R}_{>0}$ eine positive reelle Zahl. Dann hat die Gleichung $x^2 = -c$ in $\mathbb{C}$ genau zwei Lösungen, nämlich $x_1 = i\sqrt{c}$ und $x_2 = -i\sqrt{c}$.

Beweis. Nehmen wir an, $a + ib \in \mathbb{C}$ sei eine Lösung der Gleichung, d.h.

$$(a + ib)^2 = (a^2 - b^2) + 2abi \overset{!}{=} -c \in \mathbb{R} \tag{4.1.1}$$

Es ist genau dann $(a + ib)^2 \in \mathbb{R}$, wenn $\operatorname{Im}((a + ib)^2) = 2ab = 0$ ist, d.h. wenn $a = 0$ oder $b = 0$ ist. Wäre $b = 0$, so würde aus (4.1.1) die Bedingung $a^2 = -c < 0$ folgen und eine solche Zahl $a \in \mathbb{R}$ gibt es nicht. Folglich muss $a = 0$ sein, d.h. aus (4.1.1) folgt $-b^2 = -c \iff b = \pm\sqrt{c}$. Wir erhalten also gerade die Kandidaten x_1 und x_2. Überprüfen wir, dass beide tatsächlich Lösungen von $x^2 = -c$ sind:

$$\begin{aligned} (i\sqrt{c})^2 &= (0 + i\sqrt{c}) \cdot (0 + i\sqrt{c}) = -\sqrt{c}^2 + 0i = -c, \\ (-i\sqrt{c})^2 &= (0 - i\sqrt{c}) \cdot (0 - i\sqrt{c}) = -((-\sqrt{c})^2) + 0i = -c. \end{aligned}$$

□

Beispiel 4.12. *Die Gleichung $x^2 = -4$ hat in $\mathbb{C}$ zwei Lösungen, denn wenden wir obiges Lemma mit $c = 4$ an, so erhalten wir $(-2i)^2 = (2i)^2 = -4$.*

4.2 Lösung quadratischer Gleichungen in $\mathbb{R}$ und in $\mathbb{C}$

Quadratische Gleichungen in einer Variablen haben die allgemeine Form

$$ax^2 + bx + c = 0, \qquad a \in \mathbb{R} \setminus \{0\},\ b, c \in \mathbb{R}. \tag{4.2.1}$$

Durch quadratische Ergänzung lässt sich die bekannte Lösungsformel herleiten:

$$\begin{aligned} & & ax^2 + bx + c &= 0 \\ &\iff & x^2 + \frac{b}{a}x + \frac{c}{a} &= 0 \\ &\iff & \left[x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2\right] - \left(\frac{b}{2a}\right)^2 + \frac{c}{a} &= 0 \\ &\iff & \left[x + \frac{b}{2a}\right]^2 &= \frac{b^2 - 4ac}{4a^2} \end{aligned} \tag{4.2.2}$$

Damit gilt nach dem Radizieren (Wurzelziehen) und Subtrahieren von $\frac{b}{2a}$ auf beiden Seiten, dass die allgemeine Lösung der quadratischen Gleichung (4.2.1) die Form

$$x = -\frac{b}{2a} \pm \sqrt{\frac{b^2 - 4ac}{4a^2}} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{4.2.3}$$

hat, *sofern die auftretende Wurzel definiert ist.*

Bemerkung 4.13. *Alternativ kann jede quadratische Gleichung in die Normalform*

$$x^2 + px + q = 0, \qquad p, q \in \mathbb{R},$$

gebracht werden. Die Lösung hat dann die Form

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}.$$

Bemerkung 4.14. *Ganz ohne Rechenaufwand lassen sich die Lösungen einer quadratischen Gleichung finden, wenn diese in Produktform*

$$(x - x_1)(x - x_2) = 0, \qquad x_1, x_2 \in \mathbb{R} \tag{4.2.4}$$

vorliegt. Mit der Nullteilerfreiheit von $\mathbb{R}$ lässt sich die Lösungsmenge von (4.2.4) *sofort als $L = \{x_1, x_2\}$ ablesen.*

Mit obigen Rechnungen und unseren Kenntnissen über die komplexen Zahlen folgt direkt die folgende Aussage zu quadratischen Gleichungen in $\mathbb{C}$:

Satz 4.15

In $\mathbb{C}$ hat die quadratische Gleichung (4.2.1) eine Lösung, nämlich $x = -\frac{b}{2a}$, falls $b^2 = 4ac$ ist und andernfalls zwei Lösungen, welche gegeben sind durch (4.2.3)

Beweis. Die Aussagen folgen direkt aus (4.2.3) mit den Kenntnissen über komplexe Zahlen. □

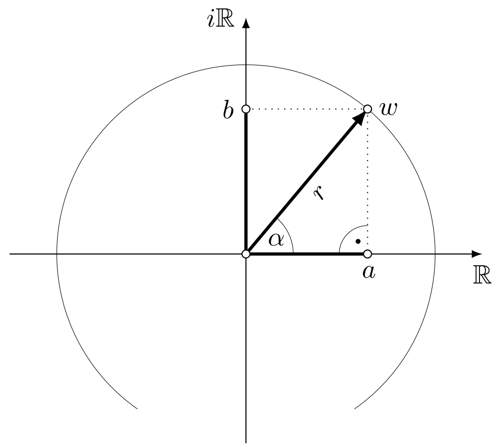
4.3 Polarkoordinaten komplexer Zahlen

Definition 4.16: Polarkoordinaten

Jede komplexe Zahl $w \in \mathbb{C}$ lässt sich darstellen als $w = r \cdot (\cos\alpha + i \sin\alpha)$ mit $r \in [0, \infty)$ und $\alpha \in (-\pi, \pi]$. Die Zahlen r und α heißen *Polarkoordinaten* von w. Dabei ist $r = |w|$ der *Betrag* von w und $\alpha = \arg(w)$ nennt man *Argument* von w.

Bemerkung 4.17. *Die Darstellung in Polarkoordinaten ist eindeutig, sofern man nur die Zahlen in* $\mathbb{C} \setminus \{0\}$ *zulässt. In diesem Fall erhält man eine Bijektion zwischen* $\mathbb{C} \setminus \{0\}$ *und* $(0, \infty) \times (-\pi, \pi]$ *via* $(a + ib) \mapsto (|a + ib|, \arg(a + ib))$.

Die Berechnung des Arguments α erfordert Kenntnisse der trigonometrischen Funktionen, speziell des Arkustangens. Aus diesem Grund werden wir in dieser Veranstaltung zwar den Betrag, nicht jedoch das Argument einer komplexen Zahl berechnen. Das Argument dient als Winkel lediglich der Veranschaulichung in der Gauß'schen Zahlenebene.



$$w = a + ib = r(\cos\alpha + i\sin\alpha)$$
$$|w| = \sqrt{a^2 + b^2} = r$$
$$a = \mathrm{Re}(w), \quad b = \mathrm{Im}(w)$$
$$\alpha = \arg(w)$$

Für natürliche Zahlen $n \in \mathbb{N}$ ist die n-te Potenz einer komplexen Zahl $w \in \mathbb{C}$ induktiv definiert:

$$w^1 := w \quad \text{und} \quad w^{n+1} := w^n \cdot w.$$

Ebenso definiert man analog zu den Potenzen reeller Zahlen

$$w^0 := 1 \quad \text{und} \quad w^{-n} := \left(\frac{1}{w}\right)^n = \frac{1}{w^n}.$$

Bemerkung 4.18. *Die Veranschaulichung von Produkten und Potenzen komplexer Zahlen anhand der Polarkoordinaten erfolgt ausführlich in Analysis I. Ebenso wird in Analysis genauer auf die Problematik des Radizierens (d.h. des Wurzelziehens) in* $\mathbb{C}$ *eingegangen. Da wir in dieser Veranstaltung beides nicht benötigen, soll an dieser Stelle der Verweis auf die Parallelveranstaltung genügen.*

Mit Hilfe der eingeführten Potenzen kommen ab sofort komplexe Zahlen als Kandidaten für Nullstellen beliebiger Polynome in Frage.

4.4 Der Polynomring

Definition 4.19: Nullpolynom und Grad eines Polynoms

Seien $(\mathbb{K}, +, \cdot)$ ein Körper und $p(x)$ ein Polynom über $\mathbb{K}$, d.h. $p(x)$ habe die Form

$$p(x) = \sum_{k=0}^{n} a_k x^k \qquad \text{mit } a_0, a_1, \ldots, a_n \in \mathbb{K} \text{ für ein } n \in \mathbb{N}_0.$$

1. Gilt $a_k = 0$ für alle $k \in \mathbb{N}_0$, so heißt p *Nullpolynom* und man schreibt $p = 0$.
2. Der *Grad* von p ist definiert als

$$\deg(p) := \begin{cases} -\infty, & \text{falls } p = 0 \text{ ist,} \\ \max\{k \in \mathbb{N}_0 \mid a_k \neq 0\}, & \text{sonst.} \end{cases}$$

Beispiel 4.20. *Das Polynom $p(x) = x^2 + x$ ist ein Polynom vom Grad 2, d.h. $\deg(p) = 2$. Insbesondere ist p also nicht das Nullpolynom. Betrachtet man hingegen die zugehörige Polynomfunktion über dem Körper $\mathbb{F}_2$, $p\colon \mathbb{F}_2 \to \mathbb{F}_2$ mit $x \mapsto p(x)$, so ist*

$$p(0) = 0^2 + 0 = 0 \quad \textit{und} \quad p(1) = 1^2 + 1 = 0,$$

d.h. es handelt sich um die Nullfunktion. Dieses Beispiel illustriert, dass über endlichen Körpern eine Unterscheidung zwischen Nullpolynom und Nullfunktion nötig ist.

Während Polynomfunktionen so wie allgemeine Abbildungen addiert und multipliziert werden, definieren wir diese Verknüpfungen auf Polynomen folgendermaßen:

Definition 4.21: Polynomring

Die Menge der Polynome über einem Körper $(\mathbb{K}, +, \cdot)$ bezeichnen wir mit $\mathbb{K}[x]$. Sind $p, q \in \mathbb{K}[x]$ mit den Darstellungen

$$p(x) = \sum_{k=0}^{n} a_k x^k \quad \text{und} \quad q(x) = \sum_{k=0}^{m} b_k x^k,$$

so definieren wir die Summe und das Produkt der beiden Polynome als

$$(p \oplus q)(x) := \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) x^k \quad \text{und}$$

$$(p \odot q)(x) := \sum_{k=0}^{m+n} c_k x^k, \quad \text{wobei } c_k = \sum_{i+j=k} a_i \cdot b_j \text{ist.}$$

Falls $m < n$ ist, so setzen wir $b_{m+1} = \ldots = b_n = 0$ und analog falls $m > n$ ist.

Satz 4.22

Mit den eingeführten Verknüpfungen ist $(\mathbb{K}[x], \oplus, \odot)$ ein kommutativer Ring mit Einselement, genannt der *Polynomring* über $\mathbb{K}$.

Beweis. Zum Beweis müssen wieder die Ringeigenschaften, sowie das Kommutativgesetz der Multiplikation überprüft werden. Dies ist eine Übungsaufgabe. □

Bemerkung 4.23. *Es gibt an dieser Stelle wieder zwei verschiedene Additionen und Multiplikationen: $+$ und $\cdot$ im Körper $\mathbb{K}$ und $\oplus$ und $\odot$ im Polynomring $\mathbb{K}[x]$. Ab jetzt verwenden wir für beide die Symbole $+$ bzw. $\cdot$ um die Schreibarbeit nicht unnötig zu verkomplizieren. Zudem werden wir ab jetzt die Verknüpfungen eines Körpers, sofern nicht für die Aufgabe nötig, unterschlagen und einfach $\mathbb{K}$ anstelle von $(\mathbb{K}, +, \cdot)$ schreiben wenn wir den Körper meinen.*

4.5 Nullstellen von Polynomen

Lemma 4.24

Sei $\mathbb{K}$ ein Körper. Sind $p, q \in \mathbb{K}[x]$, so gilt

$$\deg(p \cdot q) = \deg(p) + \deg(q).$$

Beweis. Seien $p(x) = \sum_{k=0}^{n} a_k x^k$ und $q(x) = \sum_{k=0}^{m} b_k x^k$ mit $a_n \neq 0$ und $b_m \neq 0$, d.h. $\deg(p) = n$ und $\deg(q) = m$. Hieraus folgt $a_n b_m \neq 0$ und somit ist $\deg(p \cdot q) = n + m$, da der größtmögliche von Null verschiedene Koeffizient nach Definition des Produktes zweier Polynome gerade $c_{m+n} = a_n b_m$ ist.

Betrachten wir noch den Fall, dass mindestens eines der beiden Polynome das Nullpolynom ist. Ist beispielsweise $p = 0$, so ist auch $p \cdot q = 0$ und somit $\deg(p \cdot q) = -\infty = -\infty + \deg(q) = \deg(p) + \deg(q)$ für beliebige Werte $\deg(q) \in \mathbb{N}_0 \cup \{-\infty\}$. □

Satz 4.25: Analogon zur Division mit Rest für Polynome

Seien $\mathbb{K}$ ein Körper und $p, \widetilde{p} \in \mathbb{K}[x]$. Gilt $\widetilde{p} \neq 0$, so existieren eindeutig bestimmte Polynome $q, r \in \mathbb{K}[x]$, so dass

$$p = q \cdot \widetilde{p} + r \quad \text{und } \deg(r) < \deg(\widetilde{p}) \text{ ist.}$$

Beweis. Der Beweis erfolgt in zwei Schritten.

Eindeutigkeit: Seien $q_1, q_2, r_1, r_2 \in \mathbb{K}[x]$ derart, dass

$$\begin{aligned} p &= q_1 \cdot \widetilde{p} + r_1 \qquad &&\text{mit} \quad \deg(r_1) < \deg(\widetilde{p}) \quad \text{und} \\ p &= q_2 \cdot \widetilde{p} + r_2 \qquad &&\text{mit} \quad \deg(r_2) < \deg(\widetilde{p}). \end{aligned}$$

Durch Subtraktion beider Gleichungen voneinander folgt

$$0 = (q_1 - q_2) \cdot \widetilde{p} + (r_1 - r_2), \qquad \text{d.h.} \quad (q_1 - q_2) \cdot \widetilde{p} = r_2 - r_1. \tag{4.5.1}$$

Nach Lemma 4.24 ist $\deg((q_1 - q_2) \cdot \widetilde{p}) = \deg(q_1 - q_2) + \deg(\widetilde{p})$. Andererseits haben zwei gleiche Polynome auch gleichen Grad, d.h. es gilt auch $\deg((q_1 - q_2) \cdot \widetilde{p}) = \deg(r_2 - r_1)$. Aus der Voraussetzung $\widetilde{p} \neq 0$ folgt $\deg(\widetilde{p}) \geq 0$.

Angenommen $q_1 \neq q_2$, dann ist $\deg(q_1 - q_2) \geq 0$ und somit

$$\deg(r_2 - r_1) = \deg(q_1 - q_2) + \deg(\widetilde{p}) \geq \deg(\widetilde{p}),$$

was im Widerspruch zu $\max\{\deg(r_1), \deg(r_2)\} < \deg(\widetilde{p})$ steht. Folglich muss $q_1 = q_2$ sein und aus (4.5.1) folgt $r_2 - r_1 = 0$, d.h. $r_1 = r_2$.

Existenz: Die Existenz der Polynome $q, r \in \mathbb{K}[x]$ mit den gewünschten Eigenschaften weisen wir nach, indem wir sie schrittweise konstruieren. Das benutzte Verfahren ist die *Polynomdivision*.

1. Zunächst schreibt man p und $\widetilde{p}$ mit der Größe nach absteigend sortierten Potenzen, d.h.

$$p(x) = a_n x^n + \ldots + a_1 x + a_0 \quad \text{und} \quad \widetilde{p}(x) = b_m x^m + \ldots + b_1 x + b_0,$$

wobei $a_n, b_m \neq 0$ seien. Ist $n < m$, so setzt man $q = 0$ und $r = p$, so dass

$$p = 0 \cdot \widetilde{p} + p \quad \text{und} \quad n = \deg(p) < \deg(\widetilde{p}) = m \text{ ist.}$$

Andernfalls geht man zum nächsten Schritt.

2. Man setzt $q_1 := \frac{a_n}{b_m} x^{n-m}$ und definiert $p_1 := p - q_1 \cdot \widetilde{p}$. Nach Konstruktion ist nun $\deg(p_1) < \deg(p)$.
3. Falls $\deg(p_1) < m$ ist, so setzt man $q := q_1$ und $r := p_1$ und ist fertig. Andernfalls wiederholt man Schritt 2 mit p_1 statt p und erhält somit den nächsten Term von q, nämlich q_2, und man setzt

$$p_2 := p_1 - q_2 \cdot \widetilde{p} \quad \text{mit } \deg(p_2) < \deg(p_1).$$

4. Da die Grade der Polynome p_j bei jedem Schritt um mindestens 1 abnehmen, existiert ein $k \leq n - m$ (sofern nicht schon in Schritt 1 Schluss war) derart, dass

$$p_k := p_{k-1} - q_k \cdot \widetilde{p} \quad \text{die Bedingung } \deg(p_k) < \deg(\widetilde{p}) \text{ erfüllt}$$

und an dieser Stelle bricht der Algorithmus ab. Durch Einsetzen aller Zwischenergebnisse erhält man

$$p = q_1 \cdot \widetilde{p} + p_1 = (q_1 + q_2) \cdot \widetilde{p} + p_2 = \ldots = (q_1 + \ldots + q_k) \cdot \widetilde{p} + p_k.$$

Die gesuchten Polynome sind somit

$$q := q_1 + \ldots + q_k \quad \text{und} \quad r := p_k.$$

□

Beispiel 4.26 (Beispiel Polynomdivision).

$$\begin{array}{l}
(x^3 - 2x^2 - 5x + 6) : (x - 1) = x^2 - x - 6 \\
-\underline{(x^3 - 1x^2)} \\
\quad - x^2 - 5x \\
\quad - \underline{(-x^2 + x)} \\
\qquad - 6x + 6 \\
\qquad - \underline{(-6x + 6)} \\
\qquad\qquad 0
\end{array}$$

Satz 4.27

Sei $p \in \mathbb{K}[x]$ ein Polynom n-ten Grades, wobei $n \in \mathbb{N}$ sei. Ist $c \in \mathbb{K}$ Nullstelle von p, d.h. gilt $p(c) = 0$, so existiert ein Polynom $q \in \mathbb{K}[x]$ mit $\deg(q) = n - 1$ so dass $p(x) = q(x)(x - c)$ ist.

Bemerkung 4.28.

1. *Den Term* $(x - c)$ *in obigem Satz nennt man* linearen Faktor *oder* Linearfaktor.
2. *Der Satz ist nur für* $n \in \mathbb{N}$ *formuliert, da ein Polynom 0. Grades konstant ist (und nicht das Nullpolynom) und somit keine Nullstelle besitzt.*

Beweis. Wenden wir Satz 4.25 mit $\widetilde{p}(x) = (x - c)$, so erhalten wir die Aussage, dass Polynome $q, r \in \mathbb{K}[x]$ existieren, so dass Folgendes gilt:

$$p(x) = q(x) \cdot (x - c) + r(x), \qquad \text{wobei} \qquad \deg(r) < \deg(\widetilde{p}) = 1 \text{ ist.} \tag{4.5.2}$$

Somit ist r ein konstantes Polynom – es bleibt also nur zu zeigen, dass diese Konstante 0 ist.

Setzen wir die Nullstelle c des Polynoms p in Gleichung (4.5.2) ein, so erhalten wir

$$0 = p(c) \overset{(4.5.2)}{=} q(c) \cdot (c - c) + r(c) \quad \Longleftrightarrow \quad r(c) = 0,$$

d.h. es ist in der Tat $r = 0$. Die Aussage zur Dimension folgt aus Lemma 4.24: Nennen wir $\widetilde{p}(x) := x - c$, so ist

$$\deg(p) = \deg(q) + \deg(\widetilde{p}) \implies \deg(q) = \deg(p) - \deg(\widetilde{p}) = n - 1.$$

□

Beispiel 4.29. *Das Polynom* $p \in \mathbb{R}[x]$ *mit* $p(x) = x^2 - 4$ *hat in* $\mathbb{R}$ *die Nullstellen* $c_1 = 2$ *und* $c_2 = -2$ *und lässt sich in der Tat schreiben als* $p(x) = q_1(x) \cdot q_2(x)$ *mit* $q_1(x) = (x - c_1) = (x - 2)$ *und* $q_2(x) = (x - c_2) = (x + 2)$.

Kennt man eine Nullstelle von p, d.h. ein $c \in \mathbb{C}$ mit $p(c) = 0$, so kann man das Polynom q aus Satz 4.27 durch Polynomdivision berechnen.

Satz 4.30

Für ein $n \in \mathbb{N}_{\geq 2}$ sei $p \in \mathbb{K}[x]$ ein Polynom n-ten Grades mit Nullstelle $c \in \mathbb{K}$ und Darstellung $p(x) = q(x)(x - c)$. Zudem sei $\widetilde{c} \in \mathbb{K}$ mit $\widetilde{c} \neq c$ eine von c verschiedene Nullstelle von p. Dann ist $q(\widetilde{c}) = 0$.

Beweis. Durch zweifaches Anwenden von Satz 4.27 wissen wir, dass Polynome $q, \widetilde{q} \in \mathbb{K}[x]$ existieren mit $\deg(q) = n - 1$ und $\deg(\widetilde{q}) = n - 2$, so dass

$$p(x) = q(x)(x - c) = \widetilde{q}(x)(x - c)(x - \widetilde{c})$$

ist. Folglich ist $q(x) = \widetilde{q}(x)(x - \widetilde{c})$. Durch Einsetzen von $\widetilde{c}$ sehen wir, dass in der Tat

$$q(\widetilde{c}) = \widetilde{q}(\widetilde{c})(\widetilde{c} - \widetilde{c}) = 0$$

ist. □

Als nächstes wollen wir überlegen, wie viele Nullstellen ein Polynom n-ten Grades haben kann. Eine obere Schranke können wir mit Hilfe aller bisherigen Aussagen bereits festlegen:

Korollar 4.31

Sei $p \in \mathbb{K}[x] \setminus \{0\}$ für einen beliebigen Körper $\mathbb{K}$. Sei $N \in \mathbb{N}_0$ die Anzahl der Nullstellen von p in $\mathbb{K}$. Dann ist $N \leq \deg(p)$.

Beweis. Diese Aussage beweisen wir mittels vollständiger Induktion über den $\deg(p)$:

Induktionsanfang $\deg(p) = 0$**:** Ein konstantes (von 0 verschiedenes) Polynom hat keine Nullstellen – die Behauptung ist also wahr.

Induktionsannahme: Die Behauptung sei für alle Polynome $q \in \mathbb{K}[x]$ mit $\deg(q) \leq n$ bewiesen.

Induktionsschritt: Sei nun $p \in \mathbb{K}[x]$ mit $\deg(p) = n+1$. Entweder p hat keine Nullstelle in $\mathbb{K}$ – in diesem Fall ist die Behauptung wahr – oder p besitzt eine Nullstelle $c \in \mathbb{K}$. In diesem Fall existiert nach Satz 4.27 ein Polynom $q \in \mathbb{K}[x]$ mit $p(x) = (x - c) \cdot q(x)$ und $\deg(q) = \deg(p) - 1 = n$. Nach Satz 4.30 sind alle Nullstellen von p (außer eventuell c) zugleich Nullstellen von q und nach Induktionsannahme gibt es davon höchstens n Stück. Insgesamt kann p (mit c zusammen) also höchstens $n + 1$ Nullstellen besitzen.

$\square$

Ein Polynom 2. Grades kann in zwei Linearfaktoren zerlegt werden, d.h.

$$p(x) = a \cdot (x - c_1) \cdot (x - c_2)$$

für Konstanten $a, c_1, c_2 \in \mathbb{K}$, obwohl es unter Umständen nur eine Nullstelle hat.

Das Polynom $p(x) = x^2 - 2x + 1$ hat beispielsweise nur eine Nullstelle, nämlich $c = 1$. Das können wir auch daran sehen, dass wir die zugehörige Polynomfunktion auch in der Form $p(x) = (x - 1)^2$ schreiben können. Um solche Fälle zu beschreiben, führen wir folgenden Begriff ein:

Definition 4.32

Sei $p \in \mathbb{K}[x]$ mit $\deg(p) \geq 0$, d.h. $p \neq 0$. Sei $c \in \mathbb{K}$ eine Nullstelle von p. Dann heißt die Zahl

$$\max\{n \in \mathbb{N} \mid \exists q \in \mathbb{K}[x] : \; p(x) = (x - c)^n \cdot q(x)\}$$

Vielfachheit der Nullstelle c.

Damit ist $c = 1$ eine 2-fache Nullstelle des Polynoms $p(x) = x^2 - 2x + 1$.

Selbst wenn wir Nullstellen entsprechend ihrer Vielfachheiten zählen, hängt die Anzahl dennoch vom gewählten Körper $\mathbb{K}$ ab.

Beispiel 4.33. *Betrachten wir das Polynom* $p(x) = x^2 + 1$.

- *In $\mathbb{R}$ besitzt p keine Nullstelle.*
- *In $\mathbb{F}_2$ ist $p(1) = 0$ und $p(0) = 1$, d.h. p besitzt eine Nullstelle.*
- *In $\mathbb{C}$ besitzt p zwei Nullstellen, nämlich i und $-i$.*

Der folgende Satz, der als *Fundamentalsatz der Algebra* oder auch *Satz von der Linearfaktorzerlegung* bekannt ist, liefert die genau Anzahl von Nullstellen im Körper der komplexen Zahlen. Es gibt diverse Beweise dieses Satzes, die jedoch auf Wissen aus anderen Bereichen der Mathematik aufbauen (siehe zum Beispiel das Buch *Analysis 1* [Kö04]). Aus diesem Grund werden wir hier auf einen Beweis verzichten.

Satz 4.34: Fundamentalsatz der Algebra

1. Jedes Polynom $p \in \mathbb{C}[x]$ vom Grad $\deg(p) \geq 1$ besitzt (mindestens) eine Nullstelle in $\mathbb{C}$.
2. Jedes Polynom n-ten Grades $p \in \mathbb{C}[x]$ lässt sich in n lineare Faktoren zerlegen, d.h. es existieren $a, c_1, \ldots, c_n \in \mathbb{C}$ so dass
$$p(x) = a \cdot \prod_{k=1}^{n} (x - c_k).$$

Der Fundamentalsatz gilt *nicht* für reelle Polynome. Um die Nullstellen von Polynomen $p \in \mathbb{R}[x]$ zu verstehen, müssen wir einen Umweg über (echt) komplexe Nullstellen von Polynomen mit reellen Koeffizienten machen.

Lemma 4.35

Für ein $n \in \mathbb{N}$ sei p ein Polynom n-ten Grades mit reellen Koeffizienten, d.h. $p(x) = \sum_{k=0}^{n} a_k x^k$ mit $a_0, a_1, \ldots, a_n \in \mathbb{R}$. Sei $c \in \mathbb{C}$ mit $p(c) = 0$. Dann ist auch $p(\overline{c}) = 0$.

Beweis. Ist $c \in \mathbb{R}$, dann ist $\overline{c} = c$ und somit gilt sofort $p(c) = 0 \iff p(\overline{c}) = 0$.

Sei $p(x) = \sum_{k=0}^{n} a_k x^k$. Es gelte $p(c) = 0$. Da die Koeffizienten reell sind, gilt $a_k = \overline{a_k}$ für $k \in \{0, 1, \ldots, n\}$. Mit Aussagen (vi) und (v) von Satz 4.10 folgern wir

$$p(\overline{c}) = \sum_{k=0}^{n} a_k \cdot \overline{c}^k = \sum_{k=0}^{n} \overline{a_k} \cdot \overline{c}^k \overset{(vi)}{=} \sum_{k=0}^{n} \overline{a_k \cdot c^k} \overset{(v)}{=} \overline{p(c)} = \overline{0} = 0.$$

□

Beispiel 4.36. *Das Polynom $p(x) := x^2 + 1$ hat zwei komplexe Nullstellen, nämlich $c_1 = i$ und $c_2 = -i = \overline{c_1}$.*

Satz 4.37

Jedes Polynom n-ten Grades $p \in \mathbb{R}[x]$ lässt sich in höchstens n lineare oder quadratische Faktoren zerlegen.

Beispiel 4.38. *$p(x) = (x^2 + 1)(x - 5)$ besitzt die Nullstellen 5, i und $-i$ in $\mathbb{C}$. In $\mathbb{R}$ lässt sich dieses Polynom allerdings nicht weiter zerlegen.*

Zum Beweis benötigen wir einen kleinen Hilfssatz:

Lemma 4.39

Sei $c \in \mathbb{C} \setminus \mathbb{R}$ eine echt komplexe Nullstelle von $p \in \mathbb{R}[x]$ und sei $g(x) := (x - c) \cdot (x - \overline{c})$. Dann gilt:

(i) $g \in \mathbb{R}[x]$;

(ii) es existiert ein Polynom $q \in \mathbb{R}[x]$ mit $p = q \cdot g$.

Mit anderen Worten: Je zwei Faktoren von p, die zu komplex konjugierten Nullstellen gehören, ergeben zusammen einen quadratischen Faktor von p mit reellen Koeffizienten.

Beweis von Lemma 4.39. Zeigen wir zunächst (i):

Ist $c \in \mathbb{C} \setminus \mathbb{R}$ eine echt komplexe Nullstelle des Polynoms p, dann ist nach Lemma 4.35 auch $\overline{c}$ eine Nullstelle von p, d.h. p hat die Darstellung $p(x) = q(x)(x - c)(x - \overline{c})$. Nach Satz 4.10 ist

$$g(x) = (x - c)(x - \overline{c}) = x^2 - (c + \overline{c})x + c \cdot \overline{c} = x^2 - 2\,\mathrm{Re}(c)x + |c|^2,$$

wobei $1, \mathrm{Re}(c), |c|^2 \in \mathbb{R}$ sind. Damit ist also $g \in \mathbb{R}[x]$.

Um (ii) zu zeigen, wenden wir wieder Satz 4.25 an. Dieser sagt uns, dass Polynome $q, r \in \mathbb{R}[x]$ existieren, so dass

$$p(x) = q(x) \cdot g(x) + r(x) \quad \text{ist mit} \quad \deg(r) < \deg(g) = 2.$$

Andererseits wissen wir, dass ein Polynom $r \in \mathbb{R}[x] \setminus \{0\}$ mit $\deg(r) \leq 1$ (nach Korollar 4.31) höchstens eine Nullstelle haben kann. Es gilt jedoch nach Konstruktion

$$r(c) = p(c) - q(c) \cdot g(c) = 0 \quad \text{und} \quad r(\overline{c}) = p(\overline{c}) - q(\overline{c}) \cdot g(\overline{c}) = 0,$$

wobei $c \neq \overline{c}$ ist. Folglich muss $r = 0$ sein, was (ii) beweist. □

Beweis von Satz 4.37. Die Nullstellen eines Polynoms mit reellen Koeffizienten sind entweder reell oder echt komplex. Durch den Fundamentalsatz wissen wir, dass p, wenn wir alle Nullstellen entsprechend ihren Vielfachheiten ggf. mehrfach zählen, genau $n = \deg(p)$ Nullstellen besitzt. Bezeichnen wir diese mit

$$c_1, \ldots, c_k \in \mathbb{R} \quad \text{und} \quad c_{k+1}, \ldots, c_n \in \mathbb{C} \setminus \mathbb{R}.$$

Nach Satz 4.35 ist $n - k$ (d.h. die Anzahl echt komplexer Nullstellen) gerade und nach Lemma 4.39 können jeweils zwei komplex konjugierte Nullstellen multiplizieren um reelle Polynome $g_1, \ldots, g_m \in \mathbb{R}[x]$ (mit $m = \frac{n-k}{2}$) zu erhalten. Ersetzen wir die Linearfaktoren der echt komplexen Nullstellen durch die reellen quadratischen Polynome, so erhalten wir direkt als Folge des Fundamentalsatzes die Darstellung

$$p(x) = a \cdot \prod_{j=1}^{k} (x - c_j) \cdot \prod_{j=1}^{m} g_j(x),$$

wobei $a \in \mathbb{R}$ ist, da $p(x) \in \mathbb{R}[x]$ ist. □

Zusammenfassung der Ergebnisse für $\mathbb{R}[x]$

Sei $p(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{R}[x]$ ein Polynom n-ten Grades.

- Ist $c \in \mathbb{R}$ eine Nullstelle von p, so gibt es ein reelles Polynom q vom Grad $n-1$, so dass $p(x) = q(x) \cdot (x - c)$ ist. Die übrigen Nullstellen von p sind ebenfalls Nullstellen von q.
- Das Polynom q berechnet man durch Polynomdivision.
- Ist $c \in \mathbb{C} \setminus \mathbb{R}$ eine Nullstelle von p, so ist $\bar{c}$ eine weitere Nullstelle von p und mit der Festlegung $g(x) := (x - c) \cdot (x - \bar{c})$ existiert ein reelles Polynom q vom Grad $n - 2$, so dass $p(x) = q(x) \cdot g(x)$ ist; q berechnet man wieder durch Polynomdivision.
- Insgesamt hat p bei Beachtung der Vielfachheiten genau n (möglicherweise komplexe) Nullstellen.

Zum Abschluss des Kapitels noch eine Bemerkung:

Bemerkung 4.40. *Die komplexen Zahlen ermöglichen uns, die Anzahl der zu erwartenden reellen oder komplexen Nullstellen besser voraussagen zu können. Zudem hilft die Polynomdivision, den Grad des Polynoms zu verringern, wenn man bereits eine Nullstelle erraten hat. So lange wir jedoch keine Nullstelle erraten können, helfen nur numerische oder graphische Lösungen. Bei reellen Funktionen bestehen letztere darin, den Schnittpunkt des Graphen einer Funktion mit der x-Achse (zumindest näherungsweise) zu bestimmen.*

4.6 Aufgaben

Aufgabe 4.1. *Berechnen Sie Realteil, Imaginärteil, die komplex Konjugierten und die Beträge der folgenden komplexen Zahlen:*

$$z_1 = \frac{2}{3}i + \frac{1}{2} - \frac{4}{3}i, \qquad z_2 = (1+2i)(2-i), \qquad z_3 = (1+2i)^2 - \frac{2-i}{i^3-1}.$$

Aufgabe 4.2. *Sind* $(\mathbb{K}_1, +, \cdot)$ *und* $(\mathbb{K}_2, \oplus, \odot)$ *Körper, so heißt die Abbildung* $\varphi\colon \mathbb{K}_1 \to \mathbb{K}_2$ *Körperhomomorphismus, falls für alle* $x, y \in \mathbb{K}_1$ *gilt:*

$$\varphi(x+y) = \varphi(x) \oplus \varphi(y) \qquad \textit{und}$$
$$\varphi(x \cdot y) = \varphi(x) \odot \varphi(y).$$

Auf der Menge $\mathbb{K} := \mathbb{R} \times \mathbb{R}$ *definieren wir eine Addition und Multiplikation wie folgt:*

$$(a,b) \oplus (c,d) := (a+c, b+d)$$
$$(a,b) \odot (c,d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d)$$

für beliebige $(a,b), (c,d) \in \mathbb{K}$*, wobei* $+$*,* $-$ *und* $\cdot$ *die bekannten Operationen auf* $\mathbb{R}$ *sind.*

Finden Sie Homomorphismen

a) $\varphi\colon (\mathbb{C}, +, \cdot) \to (\mathbb{R}^2, \oplus, \odot)$*,*

b) $\psi\colon (\mathbb{R}^2, \oplus, \odot) \to (\mathbb{C}, +, \cdot)$*,*

so dass $\varphi \circ \psi = \mathrm{id}_{\mathbb{R}^2}$ *und* $\psi \circ \varphi = \mathrm{id}_{\mathbb{C}}$ *ist.*

Zusatzinformation: *Mit den gewünschten Eigenschaften sind* φ *und* ψ *bijektiv, man spricht also von Isomorphismen. Man sagt,* $\mathbb{C}$ *sei zu* $\mathbb{R}^2$ isomorph.

Aufgabe 4.3. *Skizzieren Sie in der Gauß'schen Zahlenebene folgende Mengen:*

a) $A = \{z \in \mathbb{C} \mid \operatorname{Re}(z) \geq 0\}$*,*

b) $B = \{z \in \mathbb{C} \mid z = \overline{z}\}$*,*

c) $C = \{z \in \mathbb{C} \mid |z| \leq 3\}$*,*

d) $D = \{z \in \mathbb{C} \mid z \cdot \overline{z} = 4\}$*,*

e) $E = \{z \in \mathbb{C} \mid |z - w_1| > |z - w_2|\}$ *für feste Zahlen* $w_1, w_2 \in \mathbb{C}$*.*

Aufgabe 4.4. *Beweisen Sie, dass für* $z, w \in \mathbb{C}$ *das folgende Parallelogrammgesetz gilt:*

$$|z+w|^2 + |z-w|^2 = 2(|z|^2 + |w|^2).$$

Aufgabe 4.5. *Berechnen Sie mittels Polynomdivision:*

a) $(3x^3 - 2x^2 - 4x + 1) : (x+1)$*,*

b) $(x^3 + x^2y - 3xy^2 + y^3) : (x - y)$*,*

c) $(x^3 - 6x^2 + 9x - 4) : (x - 1)$*.*

Aufgabe 4.6. *Faktorisieren Sie die folgenden Polynome jeweils in $\mathbb{C}$ und in $\mathbb{R}$:*

$$p(x) = x^4 + 2x^3 + 5x^2 + 8x + 4,$$
$$q(x) = x^3 - 1.$$

Aufgabe 4.7. *Gegeben sei das Polynom $p(x) = x^3 + 2x$.*

a) *Bestimmen Sie alle Nullstellen von $p \in \mathbb{C}[x]$.*

b) *Bestimmen Sie alle Nullstellen von $p \in \mathbb{R}[x]$.*

c) *Bestimmen Sie alle Nullstellen von $p \in \mathbb{F}_3[x]$, wobei $\mathbb{F}_3 := \mathbb{Z}/3\mathbb{Z}$ sei.*

Geben Sie jeweils die Faktorisierung von p in Polynome kleinsten möglichen Grades an.

Aufgabe 4.8. *Sei $(\mathbb{K}[x], \oplus, \odot)$ die Menge der Polynome mit der Addition und Multiplikation aus Definition 4.21. Zeigen Sie, dass die Assoziativgesetze für $\oplus$ und $\odot$ gelten.*

Bemerkung: *Das ist ein Teil der Behauptung von Satz 4.22, welcher in der Vorlesung nicht bewiesen wurde.*

Kapitel 5

Lineare Gleichungssysteme

Ziele

- ☐ mehrere Lösungsmethoden zum Lösen linearer Gleichungssysteme kennen und anwenden können
- ☐ Lösungsmengen von Gleichungssystemen berechnen und mathematisch korrekt aufschreiben
- ☐ den Gauß-Algorithmus verstehen und anwenden können

5.1 Lineare Gleichungen

5.1.1 Lineare Gleichungen mit einer Variablen

Eine lineare Gleichung in einer Variablen ist von der allgemeinen Form

$$ax = b, \tag{5.1.1}$$

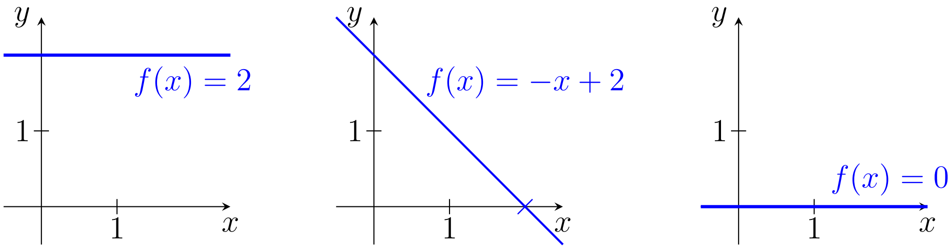
wobei $a, b \in \mathbb{R}$ gegebene Konstanten sind und $x \in \mathbb{R}$ die gesuchte Variable ist.

Untersuchen wir, welche Möglichkeiten es für die Lösungsmenge dieser Gleichung gibt, d.h. für die Menge $L := \{x \in \mathbb{R} \mid ax = b\}$.

Beispiel 5.1.

- $0x = 2$ *hat keine Lösung;*
- $-x = -2$ *hat eine Lösung, nämlich* $x = 2$*;*
- $0x = 0$ *hat unendlich viele Lösungen, nämlich alle reellen Zahlen.*

Die Lösung(en) kann man als Nullstelle(n) des Graphen der Funktion $f(x) = ax - b$ veranschaulichen:



Bemerkung 5.2. *Bevor wir eine Gleichung lösen, müssen wir festlegen, innerhalb welcher Menge wir die Lösungen einer Gleichung suchen. Beispielsweise hat $2x = 3$ eine Lösung in $\mathbb{R}$ und ebenso in $\mathbb{Q}$, jedoch nicht in $\mathbb{Z}$, da die reelle (bzw. rationale) Lösung $x = \frac{3}{2}$ keine ganze Zahl ist. Sofern nichts anderes gesagt wird, suchen wir in dieser Veranstaltung reelle Lösungen von Gleichungen.*

Analysieren wir systematisch, wie die Lösungsmenge von (5.1.1) aussehen kann:

1. Fall $a \neq 0$**:** Man darf durch $a \neq 0$ teilen und erhält die eindeutige Lösung $x = \frac{b}{a}$.

2. Fall $a = 0$**:** Die Gleichung $0 = b$ ist wahr für $b = 0$ und falsch für $b \neq 0$. Entsprechend gilt die Fallunterscheidung:

2.1. Fall $a = 0, b = 0$**:** Die Lösungsmenge ist $\mathbb{R}$.

2.2. Fall $a = 0, b \neq 0$**:** Die Lösungsmenge ist $\emptyset$.

Fassen wir das gefundene Ergebnis zusammen:

Lemma 5.3

Die lineare Gleichung $ax = b$ mit $a, b \in \mathbb{R}$ hat in $\mathbb{R}$ entweder keine Lösung ($L = \emptyset$), genau eine Lösung ($L = \{\frac{b}{a}\}$) oder unendlich viele Lösungen ($L = \mathbb{R}$).

Bemerkung 5.4. *Die Fallunterscheidung ist eine nützliche Beweismethode. Wichtig ist dabei, dass tatsächlich alle Fälle untersucht werden. Eine Fallunterscheidung ist insbesondere dann nötig, wenn eine Variable oder ein Parameter den Wert Null annehmen kann und im weiteren Verlauf durch die Variable oder den Parameter geteilt wird.*

5.1.2 Lineare Gleichungen mit mehreren Variablen

Betrachten wir die lineare Gleichung

$$ax + by = c, \tag{5.1.2}$$

wobei x und y die gesuchten reellwertigen Variablen (d.h. die Unbekannten) und $a, b, c \in \mathbb{R}$ gegebene Konstanten sind. Lösungen solcher Gleichungen sind Paare $(x, y) \in \mathbb{R} \times \mathbb{R}$, für die (5.1.2) eine wahre Aussage wird.

Wenn $b = 0$ ist, so ist die Gleichung (abgesehen von der Bezeichnung der Koeffizienten) gerade die im vorherigen Abschnitt betrachtete lineare Gleichung mit einer Variablen. Daher nehmen wir jetzt an, dass $b \neq 0$ sei. Dann gilt

$$ax + by = c \quad \iff \quad y = \frac{c - ax}{b} \quad \iff \quad y = -\frac{a}{b}x + \frac{c}{b}.$$

Fassen wir die Ergebnisse aus dem vorherigen Abschnitt und obiger Rechnung zusammen, so erhalten wir mit einer Fallunterscheidung alle möglichen Lösungsmengen von (5.1.2):

Fall $a = b = c = 0$**:** Die Gleichung ist immer wahr, d.h. $L = \mathbb{R} \times \mathbb{R}$.

Fall $a = b = 0$, $c \neq 0$**:** Die Gleichung ist immer falsch, d.h. $L = \emptyset$.

Fall $b = 0$, $a \neq 0$**:** Die Lösungsmenge ist gegeben durch die Gerade

$$L = \left\{ (x, y) \in \mathbb{R} \times \mathbb{R} \,\middle|\, x = \frac{c}{a} \right\} = \left\{ \frac{c}{a} \right\} \times \mathbb{R}.$$

Fall $b \neq 0$**:** Die Lösungsmenge ist gegeben durch die Gerade

$$L = \left\{ (x, y) \in \mathbb{R} \times \mathbb{R} \,\middle|\, y = -\frac{a}{b}x + \frac{c}{b} \right\}.$$

Analog kann man als Übung die möglichen Lösungsmengen für eine lineare Gleichung mit n Variablen,

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b, \qquad a_1, \ldots, a_n, b \in \mathbb{R},$$

untersuchen, wobei $n > 2$ eine beliebige natürliche Zahl sein darf. Lösungen schreiben wir in diesem Fall verkürzt als $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ auf.

5.2 Lineare Gleichungssysteme

Ein lineares Gleichungssystem (LGS) ist ein System der Form

$$\begin{aligned} \alpha_{1,1} x_1 + \ldots + \alpha_{1,n} x_n &= b_1 \\ &\vdots \\ \alpha_{m,1} x_1 + \ldots + \alpha_{m,n} x_n &= b_m. \end{aligned}$$

In diesem Fall ist es ein System aus m Gleichungen mit n Unbekannten. Eine Lösung eines solchen LGS ist ein n-Tupel $(x_1, \ldots, x_n)$, welches alle Gleichungen *simultan* zu wahren Aussagen macht.

Verkürzte Schreibweise von LGS

Betrachten wir die folgenden Gleichungssysteme:

$$\begin{aligned} 3x + y + 2z &= 3 \\ x + 2y + z &= 2 \\ 2y + 3z &= -4 \end{aligned} \qquad\qquad \begin{aligned} 3a + b + 2c &= 3 \\ a + 2b + c &= 2 \\ 2b + 3c &= -4 \end{aligned}$$

Diese Systeme unterscheiden sich offensichtlich nur in der Benennung der Variablen. Sie sind durch die gleichen Rechenschritte lösbar und haben die gleiche Lösungsmenge. Aus Sicht der Mathematik sind sie also identisch. Daher können wir eine verkürzte Schreibweise für LGS einführen, die auf die Benennung der Variablen verzichtet.

Notation

Betrachten wir ein LGS der Form

$$\begin{aligned} \alpha_{1,1}x_1 + \ldots + \alpha_{1,n}x_n &= b_1 \\ \alpha_{2,1}x_1 + \ldots + \alpha_{2,n}x_n &= b_2 \\ &\vdots \\ \alpha_{m,1}x_1 + \ldots + \alpha_{m,n}x_n &= b_m. \end{aligned}$$

Die zugehörige Schreibweise als *erweiterte Koeffizientenmatrix* ist

$$\left[\begin{array}{ccc|c} \alpha_{1,1} & \cdots & \alpha_{1,n} & b_1 \\ \alpha_{2,1} & \cdots & \alpha_{2,n} & b_2 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & b_m \end{array}\right],$$

was wiederum eine Abkürzung des Problems ist, den Vektor $x \in \mathbb{R}^n$ zu bestimmen, der die Gleichung

$$A \cdot x = b$$

löst, wobei

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \cdots & \alpha_{2,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad \text{ist.}$$

Beispiel 5.5. *Das LGS*

$$\begin{aligned} 3x + y + 2z &= 3 \\ x + 2y + z &= 2 \\ 2y + 3z &= -4 \end{aligned}$$

hat folgende Darstellung als erweiterte Koeffizientenmatrix*:*

$$\left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ 1 & 2 & 1 & 2 \\ 0 & 2 & 3 & -4 \end{array}\right]$$

5.2.1 Lösungsmethoden

Beispiel 5.6. *Das lineare Gleichungssystem*

$$\begin{aligned} 3x &= 6 \\ 4y &= 2 \end{aligned}$$

ist leicht lösbar, da beide Gleichungen unabhängig voneinander sind, d.h. sie enthalten je nur eine Variable. Die Lösungsmenge ist $L = \{(2, \frac{1}{2})\}$.

Im Folgenden werden wir verschiedene Lösungsmethoden kennen lernen, mit denen lineare Gleichungssysteme gelöst werden können.

Betrachten wir das LGS

$$2x + y = 7 \tag{5.2.1}$$
$$3x - y = 8 \tag{5.2.2}$$

Es gibt verschiedene Lösungsmethoden, die wir mit Hilfe dieses LGS illustrieren wollen:

Einsetzungsverfahren: Wir lösen eine Gleichung nach einer der Variablen auf und ersetzen diese Variable in der anderen Gleichung durch den entstandenen Ausdruck.

In diesem Fall können wir beispielsweise (5.2.1) nach y umstellen und erhalten $y = 7 - 2x$. Einsetzen in (5.2.2) ergibt

$$3x - (7 - 2x) = 8 \quad \Longleftrightarrow \quad 5x = 15 \quad \Longleftrightarrow \quad x = 3.$$

Den gefundenen Wert können wir in den Ausdruck für y einsetzen und erhalten $y = 7 - 2 \cdot 3 = 1$. Die Lösungsmenge des LGS ist somit $L = \{(3, 1)\}$.

Gleichsetzungsverfahren: Zwei Gleichungen werden nach der selben Variablen aufgelöst und die so entstehenden Ausdrücke gleichgesetzt.

In unserem Beispiel bietet es sich an, beide Gleichungen nach y aufzulösen, d.h. man erhält

$$y = 7 - 2x \qquad \text{und} \qquad y = 3x - 8. \tag{5.2.3}$$

Setzen wir beide Ausdrücke gleich, so erhalten wir

$$7 - 2x = 3x - 8 \quad \Longleftrightarrow \quad 15 = 5x \quad \Longleftrightarrow \quad x = 3.$$

Einsetzen des Wertes in eine der Gleichungen aus (5.2.3) liefert

$$y = 7 - 2 \cdot 3 = 1 \qquad \text{bzw.} \qquad y = 3 \cdot 3 - 8 = 1.$$

Eliminationsverfahren: Durch Addition eines geeigneten Vielfachen einer Gleichung zur anderen soll diese so vereinfacht werden, dass eine Variable eliminiert wird.

In unserem Beispiel können wir die (5.2.2) zu (5.2.1) hinzuaddieren und erhalten

$$2x + 3x + y - y = 7 + 8 \quad \Longleftrightarrow \quad 5x = 15 \quad \Longleftrightarrow \quad x = 3.$$

Durch Einsetzen in eine der beiden Ursprungsgleichungen erhält man wie bei den vorherigen Methoden wieder $y = 1$.

Hat man mehr als 2 Gleichungen und 2 Variablen, so kann die geschickte Kombination der genannten Methoden zum Ziel führen.

Beispiel 5.7. *Betrachten wir das folgende LGS:*

$$\begin{array}{rl} 2x + y \phantom{{}+z} = 7 & (I) \\ 3x - y + z = 8 & (II) \\ 2y - z = 5 & (III) \end{array}$$

Löst man (I) nach y auf, so erhält man

$$y = 7 - 2x.$$

Einsetzen in (III) ergibt

$$2(7-2x) - z = 5 \qquad \Longleftrightarrow \qquad z = 9 - 4x.$$

Ersetzen wir y und z in (II) durch die gefundenen Terme, so erhalten wir

$$3x - (7-2x) + (9-4x) = 8 \qquad \Longleftrightarrow \qquad x = 6.$$

Ersetzen wir x in (I) und (III) (bzw. in den umgeformten Gleichungen) durch den gefundenen Wert, so erhalten wir

$$y = 7 - 2 \cdot 6 = -5 \quad \textit{und} \quad z = 9 - 4 \cdot 6 = -15,$$

d.h. das LGS hat die eindeutige Lösung $L = \{(6, -5, -15)\}$.

Bisher haben wir nur Gleichungssysteme mit eindeutiger Lösung gefunden. Wie schon für lineare Gleichungen gilt jedoch auch für LGS, dass die Lösungsmenge aus keiner, einer oder unendlich vielen Lösungen bestehen kann. Sehen wir uns dazu ein Beispiel an.

Beispiel 5.8. *Betrachten wir das unterbestimmte LGS aus 2 Gleichungen und 3 Variablen:*

$$\begin{aligned} 2x + y \phantom{{}+z} &= 7 \qquad (I) \\ 3x - y + z &= 8 \qquad (II) \end{aligned}$$

Einerseits können wir beide Gleichungen addieren und erhalten somit die Bedingung $5x + z = 15$, bzw. $z = 15 - 5x$. Lösen wir zudem Gleichung (I) nach y auf, so erhalten wir als zweite Bedingung $y = 7 - 2x$. Insgesamt ist unsere Lösungsmenge somit

$$L = \left\{ (x,y,z) \in \mathbb{R}^3 \,\middle|\, y = 7 - 2x, z = 15 - 5x \right\}.$$

Dies können wir auch in folgender Weise schreiben:

$$L = \{ (x, 7-2x, 15-5x) \mid x \in \mathbb{R} \} \qquad \text{☹ } \textit{Unübersichtlich!}$$

Da die Schreibweise von Tupeln nicht nur bei Kommazahlen unvorteilhaft ist, sondern bei längeren Termen die Lesbarkeit erschwert, nutzen wir ab jetzt grundsätzlich die Schreibweise als Spaltenvektoren:

$$L = \left\{ \begin{pmatrix} x \\ 7-2x \\ 15-5x \end{pmatrix} \,\middle|\, x \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 0 \\ 7 \\ 15 \end{pmatrix} + x \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix} \,\middle|\, x \in \mathbb{R} \right\}.$$

Die letzte Darstellung zeigt uns, dass die Lösungsmenge eine Gerade im 3-dimensionalen Raum ist. (Genaueres dazu folgt in Abschnitt 6.6.4.)

Abbildung 5.1 zeigt die Position der Ebenen, die durch (I) und (II) gegeben sind. Die Gerade, in der sich beide Ebenen schneiden, stellt die Lösungsmenge des Gleichungssystems dar.

Die vorgestellten Methoden erfordern ein gewisses Maß an Kreativität bei der Anwendung, da wir gerade bei großen LGS nach jedem Schritt entscheiden müssen, mit welcher Methode wir welche Variablen berechnen oder eliminieren wollen. Im Folgenden wollen wir uns eine Methode aneignen, die man auch einem Computer beibringen kann und die bei jedem beliebigen LGS anwendbar ist. Die drei Methoden dieses Abschnittes sind vor allem dann vorteilhaft, wenn ein kleines und leicht überschaubares LGS von einem Menschen mit möglichst geringem Aufwand gelöst werden soll.

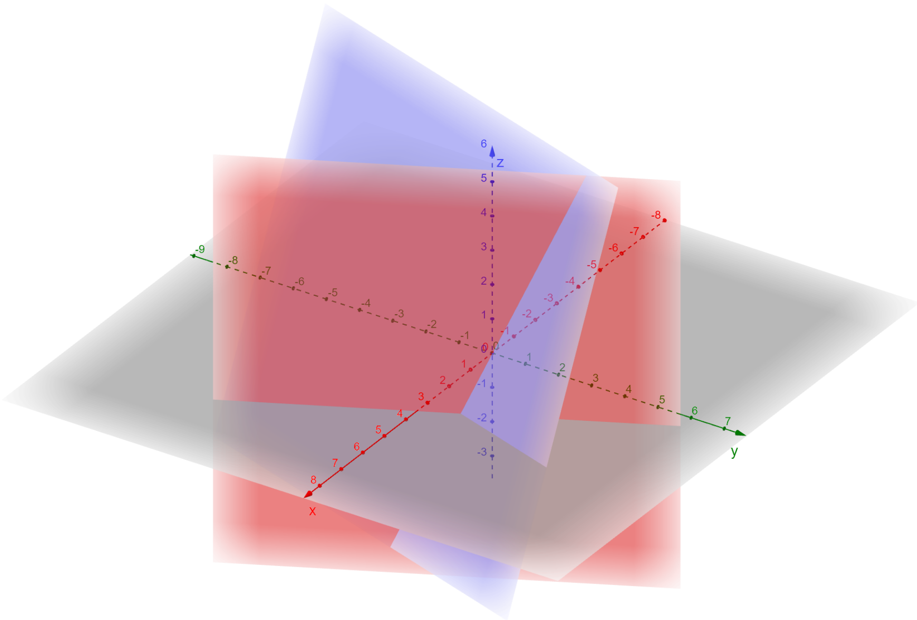


Abbildung 5.1: Ebene (I) ist rot, Ebene (II) blau und die x-y-Ebene ist grau dargestellt.

5.2.2 Gauß(-Jordan)-Algorithmus

Motivation der Methode

Die Idee hinter dem Gauß- (bzw. Gauß-Jordan-)Verfahren zu verstehen, sehen wir uns am besten mehrere Gleichungssysteme an, die leicht lösbar sind.

Beispiel 5.9. *Betrachten wir zunächst folgendes LGS:*

$$\begin{aligned} x \qquad\qquad &= 3 \\ y \qquad &= 1 \\ z &= -6 \end{aligned}$$

Da alle drei Gleichungen unabhängig voneinander sind, lösen wir jede für sich und erhalten so die Lösungsmenge $L = \left\{ \begin{pmatrix} 3 \\ 1 \\ -6 \end{pmatrix} \right\}$.

Beispiel 5.10. *Betrachten wir nun ein etwas komplizierteres LGS:*

$$\begin{aligned} x - 2y + z &= 3 \\ y + z &= 1 \\ z &= -6 \end{aligned}$$

Wir lösen die Gleichungen in diesem Fall von unten nach oben:

- *Aus der dritten Gleichung lesen wir sofort ab:* $z = -6$.
- *Setzen wir den Wert für* z *in die zweite Gleichung ein, so erhalten wir*

$$y - 6 = 1 \quad \Longleftrightarrow \quad y = 7.$$

- *Setzen wir die gefundenen Werte für y und z in die erste Gleichung ein, so erhalten wir*
$$x - 2 \cdot 7 - 6 = 3 \qquad \Longleftrightarrow \qquad x = 23.$$

Somit ist die Lösungsmenge $L = \left\{ \begin{pmatrix} 23 \\ 7 \\ -6 \end{pmatrix} \right\}$.

Das Ziel des Gauß-Verfahrens ist es, ein LGS in die Form von Beispiel 5.10 zu überführen, wohingegen das Gauß-Jordan-Verfahren die Form von 5.9 anstrebt. Im folgenden Beispiel sehen wir uns einmal an, wie diese Formen bei einem unterbestimmten LGS aussehen:

Beispiel 5.11. *Betrachten wir das folgende unterbestimmte LGS:*

$$\begin{aligned} x - 2y + z &= 3 \\ y + z &= 1 \end{aligned}$$

Da es zwei Gleichungen und drei Variablen gibt, können wir eine Variable als freie Variable *festlegen und die Werte der übrigen Variablen werden in Abhängigkeit von der freien Variablen festgelegt. In diesem Beispiel bietet es sich an, z als freie Variable zu fixieren.*

Wir lösen das LGS nun wieder „von unten nach oben“:

- *Da z nicht bestimmt wird, lösen wir die zweite Gleichung nach y auf und erhalten*
$$y = 1 - z.$$
- *Ersetzen wir y durch $1 - z$ in der ersten Gleichung, so erhalten wir*
$$x - 2(1 - z) + z = 3 \qquad \Longleftrightarrow \qquad x = 5 - 3z.$$

Somit lautet die Lösungsmenge des LGS

$$L = \left\{ \begin{pmatrix} 5 - 3z \\ 1 - z \\ z \end{pmatrix} \;\middle|\; z \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} \;\middle|\; z \in \mathbb{R} \right\}.$$

Alternativ erhalten wir durch zweifache Addition der zweiten Gleichung zur ersten Gleichung das LGS in der Form analog zu Beispiel 5.9:

$$\begin{aligned} x \qquad + 3z &= 5 \\ y + z &= 1 \end{aligned}$$

Der Vorteil dieser Umformung ist, dass die Gleichungen für x und y (fast) sofort ablesbar sind.

Äquivalenzumformungen von LGS

Analog zur Äquivalenz von Gleichungen definieren wir die Äquivalenz von Gleichungssystemen:

Definition 5.12

Zwei Gleichungssysteme heißen *äquivalent*, wenn sie die gleiche Lösungsmenge besitzen.

Der letzte Schritt hin zum Gauß- bzw. Gauß-Jordan-Algorithmus ist die Überlegung, welche Operationen ein LGS in ein dazu äquivalentes LGS überführen.

- Vertauschen von Gleichungen ändert die Lösungsmenge nicht.
- Multiplikation einer Zeile mit einem von Null verschiedenen Faktor ändert die Lösungsmenge nicht, da dies eine Äquivalenzumformung für Gleichungen ist:

  $$a_1x_1 + \ldots + a_nx_n = b \quad \Longleftrightarrow \quad ca_1x_1 + \ldots + ca_nx_n = cb \quad \text{für beliebige } c \in \mathbb{R} \setminus \{0\}.$$

- Das Addieren einer Gleichung zu einer anderen ändert die Lösungsmenge nicht. Dies haben wir bereits beim Eliminationsverfahren benutzt.

Bemerkung 5.13. *Schreibt man das LGS in der Form $A \cdot x = b$, so entsprechen die genannten Operationen der Multiplikation beider Seiten mit sogenannten* Elementarmatrizen *(Vgl. Aufgabe 2.7). Wir werden später zeigen, dass die Elementarmatrizen alle invertierbar sind. Schon jetzt können wir zeigen, dass die Multiplikation mit einer invertierbaren Matrix die Lösungsmenge nicht ändert:*

Ist der x Lösung der Gleichung $A \cdot x = b$ mit $A \in \mathbb{R}^{m\times n}$ (und damit des zugehörigen LGS), so gilt für jede invertierbare Matrix $B \in \mathbb{R}^{m\times m}$ $BAx = Bb$. Ist andererseits x Lösung von $BAx = Bb$, so gilt

$$Ax = (B^{-1}B)Ax = B^{-1}(BAx) = B^{-1}(Bb) = (B^{-1}B)b = b,$$

d.h. x ist auch Lösung von $Ax = b$.

Sobald wir nachweisen, dass die Elementarmatrizen invertierbar sind, wissen wir also, dass die Multiplikation beider Seiten eines LGS der Form $Ax = b$ die Lösungsmenge nicht ändert und damit der Gauß-Jordan-Algorithmus tatsächlich die Lösungsmenge nicht ändert.

In Bezug auf die zu einem LGS gehörende erweiterte Koeffizientenmatrix können wir somit folgende *elementaren Zeilenoperationen* festhalten, die die Lösungsmenge nicht verändern:

Elementare Zeilenoperationen

1. Vertauschen von Zeilen
2. Multiplizieren einer Zeile mit einer von Null verschiedenen Zahl
3. Addieren einer Zeile zu einer anderen Zeile

Gauß- und Gauß-Jordan-Algorithmus

Mit Hilfe der gerade benannten elementaren Zeilenoperationen soll eine erweiterte Koeffizientenmatrix in eine Zeilenstufenform (dann handelt es sich um den Gauß-Algorithmus) oder in die normierte Zeilenstufenform (dann handelt es sich um den Gauß-Jordan-Algorithmus) gebracht werden.

Definieren wir zunächst genau, was wir unter diesen Darstellungsformen verstehen.

Definition 5.14: Zeilenstufenform einer Matrix

Eine Matrix liegt in Zeilenstufenform vor, falls sie folgende Eigenschaften hat:

1. Zeilen, die nur Nullen enthalten, stehen unter den übrigen Zeilen.
2. Der *führende* (d.h. der am weitesten links stehende von Null verschiedene) Koeffizient einer Zeile befindet sich rechts von dem führenden Koeffizienten in der Zeile darüber. Der führende Koeffizient einer Zeile heißt auch *Pivotelement*; die Spalte, in der ein Pivotelement vorkommt, heißt auch *Pivotspalte*.

Insbesondere sind alle Einträge in einer Spalte unterhalb eines Pivotelementes Null.

Anhand eines Beispiels wollen wir sehen, wie wir eine Matrix in diese Form bringen können:

Beispiel 5.15. *Die erweiterte Koeffizientenmatrix*

$$\left[\begin{array}{ccc|c} \color{blue}{3} & 1 & 2 & 3 \\ \color{red}{1} & 2 & 4 & 6 \\ 0 & 5 & 10 & 15 \end{array}\right]$$

liegt nicht in Zeilenstufenform vor, da das Pivotelement der zweiten Zeile (die rot markierte 1*) nicht rechts von dem Pivotelement der ersten Zeile (die blau markierte* 3*) ist. Um dies zu beheben, können wir die zweite Zeile mit* (-3) *multiplizieren und zur ersten Zeile addieren. Somit erhalten wir die neue Matrix*

$$\left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ \color{blue}{0} & \color{blue}{-5} & \color{blue}{-10} & \color{blue}{-15} \\ 0 & \color{red}{5} & 10 & 15 \end{array}\right].$$

Nun ist das Pivotelement der dritten Zeile (in rot) noch unter dem der zweiten Zeile. Wir ersetzen jetzt die dritte Zeile durch die Summe der zweiten und dritten und erhalten somit

$$\left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ 0 & -5 & -10 & -15 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Hiermit haben wir die Zeilenstufenform erreicht. Übersetzen wir die erweiterte Koeffizientenmatrix wieder zurück in ein LGS, so erhalten wir:

$$\begin{aligned} 3x + \ \ y + \ \ 2z \quad &= 3 \\ -\,5y - 10z \quad &= -15\,. \\ 0 &= 0 \end{aligned}$$

Die dritte Gleichung können wir weglassen, da sie immer wahr ist, und die zweite kürzen, so dass wir das vereinfachte LGS

$$\begin{aligned} 3x + y + 2z &= 3 \\ y + 2z &= 3 \end{aligned}$$

erhalten. Subtrahieren wir die zweite von der ersten Gleichung, so können wir das LGS nochmals vereinfachen zu

$$\begin{aligned} 3x \qquad &= 0 \\ y + 2z &= 3 \end{aligned}$$

und bekommen somit die Lösungsmenge

$$L = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x = 0, y = 3 - 2z \right\} = \left\{ \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \;\middle|\; z \in \mathbb{R} \right\}.$$

Wir sehen dass nach dem erfolgreichen Umformen in die Zeilenstufenform noch etliche Rechenschritte nötig waren um die Lösungsmenge zu bestimmen. Der Gauß-Jordan-Algorithmus geht daher weiter und fordert die Umformung in die normierte Zeilenstufenform:

Definition 5.16: Normierte Zeilenstufenform einer Matrix

Eine Matrix liegt in normierter Zeilenstufenform vor, falls sie in Zeilenstufenform vorliegt und zudem folgende Eigenschaften hat:

1. Das Pivotelement in jeder Zeile, die nicht nur Nullen enthält, ist 1.
2. Jedes Pivotelement ist der einzige von 0 verschiedene Eintrag in der Pivotspalte.

Beispiel 5.17 (Fortsetzung von Beispiel 5.15)**.** *Die in Zeilenstufenform vorliegende Matrix*

$$\left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ 0 & -5 & -10 & -15 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

wollen wir nun in die normierte Zeilenstufenform bringen. Dazu teilen wir zunächst die zweite Zeile durch (-5) und erhalten somit

$$\left[\begin{array}{ccc|c} 3 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Anstatt nun die erste Zeile durch 3 zu teilen, subtrahieren wir zuerst die zweite Zeile, da dies den Rechenaufwand verringern wird, und erhalten somit

$$\left[\begin{array}{ccc|c} 3 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Jetzt müssen wir noch das Pivotelement der ersten Zeile zu 1 machen, indem wir diese Zeile durch 3 teilen:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Dies ist die normierte Zeilenstufenform. Das zugehörige LGS lautet

$$\begin{aligned} x \qquad\qquad &= 0 \\ y + 2z &= 3 \\ 0 &= 0 \end{aligned}$$

und die Lösungsmenge kann, wie in Beispiel 5.15 gesehen, (fast sofort) abgelesen werden.

Während wir Menschen die Möglichkeit haben, die Zeilenstufenform (auch Treppenform genannt) zu *sehen* und aus möglichen Rechenoperationen diejenige zu wählen, welche uns am einfachsten durchführbar erscheint, müssen wir einem Computer eine feste Reihenfolge an Operationen vorgeben – einen Algorithmus. Es gibt keine festgelegte Schrittfolge, daher wollen wir an dieser Stelle einen möglichen Algorithmus vorstellen, wie er auf der englischen Wikipedia-Seite gefunden werden kann. Es handelt sich um *Pseudocode*, d.h. der Algorithmus ist in keiner speziellen Programmiersprache verfasst und er vermischt mathematische Ausdrücke und natürliche Sprache.

```
h := 1     /* Initialsierung der Pivotzeile */
k := 1     /* Initialsierung der Pivotspalte */
so lange h ≤ m und k ≤ n
    /* Finde das k-te Pivotelement: */
    i_max := argmax (i = h ... m, abs(A[i, k]))
    falls A[i_max, k] = 0
        /* Kein Pivotelement in dieser Spalte, gehe zur nächsten Spalte */
        k := k+1
    sonst
        vertausche Zeilen(h, i_max)
        /* Für alle Zeilen unterhalb des Pivotelements: */
        für i = h + 1 ... m:
            f := A[i, k] / A[h, k]
            /* Fülle Pivotspalte unterhalb des Pivotelements mit Nullen: */
            A[i, k] := 0
            /* Für die restlichen Elemente der Zeile: */
            für j = k + 1 ... n:
                A[i, j] := A[i, j] - A[h, j] * f
        /* Erhöhe Pivotzeile und -spalte */
        h := h+1
        k := k+1
```

Bemerkung 5.18 (Hinweise zur Interpretation des Pseudocode).

- *Kommentare stehen zwischen* `/* ... */`
- `h := 1` *ordnet der Variablen h den Wert 1 zu*
- `A[h,k]` *steht für den Eintrag der h-te Zeile (englisch* row*) und k-ten Spalte (englisch* column*) der (Koeffizienten-)Matrix*
- `abs(...)` *steht für den Betrag (english* absolute value*)*
- `.../...` *steht für die Division zweier Zahlen,* `...*...` *für Multiplikation*

Das Aufschreiben der Lösungsmenge

Wir haben bereits mehrere Lösungsmengen aufgeschrieben, wollen diesen Prozess an dieser Stelle jedoch noch einmal systematisch anschauen. Dazu sehen wir uns jeweils ein Beispiel an, welches bereits in Zeilenstufenform vorliegt.

Beispiel 5.19 (Keine Lösung)**.** *Betrachten wir das folgende LGS (in Zeilenstufenform):*

$$\begin{aligned} 2x + y \quad &= 3 \\ 0 &= 1 \end{aligned} \qquad \textit{bzw.} \qquad \left[\begin{array}{cc|c} 2 & 1 & 3 \\ 0 & 0 & 1 \end{array}\right]$$

In der Schreibweise als LGS erkennt man, dass die zweite Gleichung immer eine falsche Aussage ist. In der erweiterten Koeffizientenmatrix sieht man dies daran, dass das die Ergebnisspalte eine Pivotspalte ist, d.h. der erste von Null verschiedene Wert der 2. Zeile steht in der Ergebnisspalte. In diesem Fall ist die Lösungsmenge

$$L = \emptyset.$$

Beispiel 5.20 (Eine Lösung)**.** *Betrachten wir das folgende LGS (in Zeilenstufenform):*

$$\begin{aligned} x + \;\; y &= 3 \\ 2y &= 2 \end{aligned} \qquad \textit{bzw.} \qquad \left[\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 2 & 2 \end{array}\right]$$

Das Gleichungssystem lösen wir „von unten nach oben" und finden so die eindeutige Lösung, nämlich $y = 1$ und (durch Einsetzen in die erste Gleichung) $x = 2$. Die Lösungsmenge schreiben wir wir folgt auf:

$$L = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}.$$

Immer dann, wenn links von der Ergebnisspalte gleich viele Zeilen und Spalten sind und alle auftretenden Spalten Pivotspalten sind, hat das zur erweiterten Koeffizientenmatrix gehörende LGS genau eine Lösung.

Beispiel 5.21 (Unendlich viele Lösungen)**.** *Betrachten wir das folgende LGS (in Zeilenstufenform):*

$$\begin{aligned} x_1 + 2x_2 \qquad\quad + x_4 &= 4 \\ x_3 - x_4 &= 1 \end{aligned} \qquad \textit{bzw.} \qquad \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 4 \\ 0 & 0 & 1 & -1 & 1 \end{array}\right]$$

Man kann die Lösungsmenge wie folgt aufschreiben:

$$L = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \;\middle|\; x_1 + 2x_2 + x_4 = 4,\; x_3 - x_4 = 1 \right\}.$$

Diese Schreibweise ist allerdings nicht sehr aussagekräftig. Wir führen also folgende Begriffe ein:

Die Variablen, die den Pivotspalten entsprechen, heißen führende Variablen*, alle anderen Variablen heißen* freie Variablen*. In dem Beispiel sind x_1 und x_3 führende Variablen, x_2 und x_4 freie Variablen.*

Nun stellen wir beide Gleichungen nach der führenden Variablen um. Indem wir wieder „von unten nach oben“ arbeiten, können wir dafür sorgen, dass die führenden Variablen mit Hilfe der freien Variablen ausgedrückt werden:

$$x_3 = 1 + x_4 \qquad \textit{und}$$
$$x_1 = 4 - 2x_2 - x_4.$$

Fügen wir nun noch hinzu, dass x_2 und x_4 beliebig sind, so erhalten wir folgendes LGS:

$$\begin{aligned} -2x_2 - x_4 + 4 &= x_1 \\ x_2 \qquad\quad &= x_2 \\ x_4 + 1 &= x_3 \\ x_4 \quad &= x_4 \end{aligned}$$

Damit können wir die Koeffizienten für unsere Lösungsmenge ablesen:

$$L = \left\{ \begin{pmatrix} -2a - b + 4 \\ a \\ 1 + b \\ b \end{pmatrix} \middle| \; a, b \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 4 \\ 0 \\ 1 \\ 0 \end{pmatrix} + a \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \middle| \; a, b \in \mathbb{R} \right\}.$$

Zusammenfassung

Der Gauß-Jordan-Algorithmus zum Lösen eines LGS kann wie folgt zusammengefasst werden:

1. Man bildet die erweiterte Koeffizientenmatrix $(A|b)$.
2. Durch Vertauschen von Zeilen sorgt man dafür, dass das die Pivotelemente niemals links von den Pivotelementen der darüber liegenden Zeilen sind. Insbesondere sind dann alle Null-Zeilenvektoren ganz unten.
3. Man fixiert das erste Pivotelement, welches direkt unter einem anderen Pivotelement steht. Nennen wir die Elemente $\alpha_{k,j}$ und $\alpha_{k+1,j}$. Dann ersetzt man Zeile Z_{k+1} durch $Z_{k+1} - \frac{\alpha_{k+1,j}}{\alpha_{k,j}} Z_k$. Diesen Schritt wiederholt man ggf. mit den folgenden Zeilen bis unter $\alpha_{k,j}$ nur noch Nullen stehen.
4. Man wiederholt Schritte 2 und 3 so lange, bis die Matrix in Zeilenstufenform da steht.
5. Man geht alle Zeilen durch und sollte das Pivotelement $(\alpha_{k,j})$ nicht 1 sein, ersetzt man die Zeile durch $\frac{1}{\alpha_{k,j}} Z_k$.
6. Statt von oben nach unten geht man die Matrix von unten nach oben durch und ändert alle Spalten analog zu Schritt 3 derart, dass über und unter Pivotelementen nur Nullen stehen.
7. Zur Auswertung gibt es drei mögliche Fälle:

   1. Fall: Ist die Ergebnisspalte (ganz rechts) eine Pivotspalte, d.h. ist der erste von Null verschiedene Zeileneintrag in dieser Spalte, so ist das LGS nicht lösbar.

   2. Fall: Sind die zu A gehörigen Spalten alle Pivotspalten, so ist das LGS eindeutig lösbar. Die Lösung ist in der letzten Spalte ablesbar.

3. Fall: Sind zu A gehörige Spalten keine Pivotspalten, so nennt man die zugehörigen Variablen *frei* und die Werte der nicht freien ("führenden") Variablen (die zu den Pivotspalten gehören) werden in Abhängigkeit von den den freien Variablen angegeben.

5.3 Invertieren von Matrizen

Erinnern wir uns an Beispiel 2.21:

Beispiel 5.22. *Gegeben ist die Matrix*

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}.$$

Ist

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ein Kandidat für die Inverse zu A, dann gilt:

$$A \cdot B = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a+c & 2b+d \\ 5a+3c & 5b+3d \end{pmatrix} \overset{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Wir mussten also die folgenden zwei linearen Gleichungssysteme (LGS) lösen:

$$\begin{aligned} 2a + c &= 1 \\ 5a + 3c &= 0 \end{aligned} \qquad \textit{und} \qquad \begin{aligned} 2b + d &= 0 \\ 5b + 3d &= 1 \end{aligned}$$

Als Lösung erhielten wir

$$A^{-1} = B = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}.$$

In Matrixschreibweise heißt das Problem:

$$\left[\begin{array}{cc|c} 2 & 1 & 1 \\ 5 & 3 & 0 \end{array}\right] \qquad \textit{und} \qquad \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 5 & 3 & 1 \end{array}\right].$$

Da beide linken Seiten identisch sind – das ist jeweils Matrix A – schreiben wir beide Systeme zusammen in eine erweiterte Koeffizientenmatrix:

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array}\right]$$

Die üblichen Zeilenumformungen ergeben:

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & -1 & 5 & -2 \end{array}\right] \rightsquigarrow \left[\begin{array}{cc|cc} 2 & 0 & 6 & -2 \\ 0 & -1 & 5 & -2 \end{array}\right] \rightsquigarrow \left[\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -5 & 2 \end{array}\right]$$

Die Matrix, die sich aus den Ergebnisspalten zusammensetzt, ist gerade die gesuchte Inverse:

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}.$$

Motiviert durch obiges Beispiel postulieren wir folgenden Satz, dessen Beweis allerdings erst später erfolgen kann, da unser Wissen über Matrizen dafür noch nicht ausreicht.

Satz 5.23: Invertieren von Matrizen mit dem Gauß-Jordan-Algorithmus

Eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann invertierbar, wenn die erweiterte Koeffizientenmatrix $(A|E_n)$ mit Hilfe von elementaren Zeilenumformungen in die Form $(E_n|B)$ gebracht werden kann. In diesem Fall ist $B = A^{-1}$.

Wir werden im Laufe der Vorlesung noch weitere Möglichkeiten kennen lernen, die Invertierbarkeit einer Matrix vorherzusagen und ggf. die Inverse zu berechnen. Sehen wir uns jedoch zunächst an, wozu die Inverse einer Matrix nützlich ist.

Beispiel 5.24. *Betrachten wir folgendes LGS:*

$$\begin{aligned} 2x + y &= 2 \\ 5x + 3y &= 1 \end{aligned}, \qquad \text{d.h.} \qquad \left[\begin{array}{cc|c} 2 & 1 & 2 \\ 5 & 3 & 1 \end{array}\right]$$

In Matrixschreibweise haben wir das LGS

$$A \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Multiplikation beider Seiten mit der Inversen von A, die wir gerade berechnet haben, ergibt auf der linken Seite

$$A^{-1} \cdot \left(A \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) = (A^{-1} \cdot A) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

und auf der rechte Seite

$$A^{-1} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ -8 \end{pmatrix}.$$

Gleichsetzen ergibt die Lösungsmenge

$$L = \left\{ \begin{pmatrix} 5 \\ -8 \end{pmatrix} \right\}.$$

Fassen wir den gefundenen Zusammenhang zwischen dem Lösen von LGS und der Invertierbarkeit von Matrizen zusammen:

Invertierbare Matrix und LGS

Ein reellwertiges LGS mit n Gleichungen und n Variablen lässt sich immer in der Form $A \cdot x = b$ schreiben, wobei $A \in \mathbb{R}^{n\times n}$ und $x, b \in \mathbb{R}^n$ sind.

Falls A invertierbar ist, so ist die eindeutige Lösung des LGS gegeben durch

$$x = A^{-1} \cdot b.$$

Allgemeiner als in obigem Beispiel kann man überprüfen, dass $x^* := A^{-1} \cdot b$ wirklich immer eine Lösung von $A \cdot x = b$ ist:

$$A \cdot x^* = A \cdot (A^{-1} \cdot b) = (A \cdot A^{-1}) \cdot b = E_n \cdot b = b,$$

d.h. x^* löst tatsächlich das LGS. Die Eindeutigkeit der Lösung folgt daraus, dass die Multiplikation mit einer Matrix ($\neq 0$) eine Äquivalenzumformung ist.

5.4 Aufgaben

Aufgabe 5.1. *Zeigen Sie anhand des LGS*

$$\begin{aligned} x \quad &= 1 \\ y &= 2 \end{aligned},$$

dass folgende Operationen (angewandt auf die erweiterte Koeffizientenmatrix) die Lösungsmengen von Gleichungssystemen möglicherweise ändern:

a) *Vertauschen von Spalten,*

b) *Multiplizieren einer Spalte mit einer von Null verschiedenen Zahl,*

c) *Addieren einer Spalte zu einer anderen.*

Aufgabe 5.2. *Nehmen wir an, nach Anwendung des Gauß-Verfahrens erhalten Sie die linearen Gleichungssysteme*

a)

$$\begin{aligned} x_1 - 3x_3 \qquad\quad + x_5 &= 0 \\ x_3 - 2x_4 + x_5 &= 7 \end{aligned}$$

b)

$$\begin{aligned} x_1 \qquad\qquad &= 4 \\ x_2 + x_4 &= 1 \end{aligned}$$

Bringen Sie diese ggf. auf die normierte Zeilenstufenform und geben Sie ihre Lösungsmenge an. Beachten Sie, dass die Lösungsmenge eine Teilmenge des $\mathbb{R}^5$ (in Teil a) bzw. $\mathbb{R}^4$ (in Teil b) ist.

Aufgabe 5.3. *Um eventuelle Rechenfehler schneller zu bemerken, gibt es folgenden Trick:*

Gegeben sei ein LGS in Form der erweiterten Koeffizientenmatrix

$$(A|b) = \left[\begin{array}{ccc|c} \alpha_{1,1} & \cdots & \alpha_{1,n} & b_1 \\ \alpha_{2,1} & \cdots & \alpha_{2,n} & b_2 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & b_m \end{array}\right].$$

Nun fügen wir nach einem Doppelstrich eine weitere Spalte hinzu und notieren in der Spalte die Summen aller in einer Zeile stehenden Werte:

$$(A|b) = \left[\begin{array}{ccc|c||c} \alpha_{1,1} & \cdots & \alpha_{1,n} & b_1 & \alpha_{1,1} + \ldots + \alpha_{1,n} + b_1 \\ \alpha_{2,1} & \cdots & \alpha_{2,n} & b_2 & \alpha_{2,1} + \ldots + \alpha_{2,n} + b_2 \\ \vdots & & \vdots & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & b_m & \alpha_{m,1} + \ldots + \alpha_{m,n} + b_m \end{array}\right].$$

Ab jetzt machen wir alle Zeilenoperationen auch für die neue Spalte und überprüfen nach jedem Schritt, ob in der Kontrollspalte immer noch die Summe aller links vom Doppelstrich stehenden Zahlen steht.

Wenden Sie diese Methode auf folgendes LGS an:

$$\begin{aligned} 2x + 4y + 6z &= 12 \\ x + 3y + 7z &= 16 \\ 3x + 3y - 2z &= -9 \end{aligned} \quad .$$

Aufgabe 5.4. *Wenden Sie das Gauß-Jordan-Verfahren auf das folgende LGS an und bestimmen Sie seine Lösungsmenge.*

$$\begin{aligned} x_1 + 2x_2 + 2x_3 + x_4 &= 1 \\ x_2 + 3x_4 &= 0 \\ x_1 + 3x_3 &= 2 \\ 2x_1 + 3x_2 + 5x_3 + 4x_4 &= 3 \\ x_1 + x_2 + 3x_3 + 3x_4 &= 2 \end{aligned}$$

Aufgabe 5.5. *Unter einem $m \times n$-System verstehen wir ein LGS mit m Gleichungen und n Variablen.*

a) Bilden Sie ein 2×1-System mit unendlich vielen Lösungen.

b) Bilden Sie ein 2×2-System mit keiner Lösung.

c) Bilden Sie ein 5×2-System mit genau einer Lösung.

Aufgabe 5.6. *Berechnen Sie die Inversen der Matrizen*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \textit{und} \quad B = \begin{pmatrix} 4 & 3 \\ 2 & 2 \end{pmatrix}$$

mit Hilfe des Gauß-Algorithmus.

Kapitel 6

Vektorräume und ihr Zusammenhang zu LGS

Ziele

- □ den Begriff Vektorraum kennen und überprüfen können, ob es sich bei Beispielen um Vektorräume handelt
- □ das Konzept der linearen Unabhängigkeit mit der algebraischen und geometrischen Vorstellung kennen
- □ Vektoren in verschiedenen Vektorräumen auf lineare Unabhängigkeit überprüfen, den Rang einer Matrix bestimmen
- □ die Beziehungen zwischen den Begriffen *lineare Unabhängigkeit*, *Erzeugendensystem* und *Basis* kennen und auf Beispiele anwenden, Beweise dazu
- □ über die fehlende Eindeutigkeit von Basen in Form von diversen Sätzen reflektieren und Koordinaten als basisabhängig erkennen
- □ die Dimension eines Vektorraumes geometrisch und algebraisch verstehen
- □ Lösungsmengen von LGS als affine Räume erkennen und Beziehungen zwischen der Dimension des Lösungsraum und dem Rang der zum LGS gehörenden Matrix erkennen, Beweis dazu

6.1 Vektorräume

Wir haben Elemente des $\mathbb{R}^n$ ($n \in \mathbb{N}$) bereits *Vektoren* genannt ohne jedoch zu klären, was das sein soll. In diesem Kapitel werden wir sehen, dass auch andere Objekte so genannt werden können sobald sie Elemente eines *Vektorraumes* sind. Diese Struktur baut auf dem bekannten Begriff einer *Gruppe* (Definition 3.12) und eines *Körpers* (Definition 3.30) auf.

Definition 6.1: Vektorraum

Sei $(\mathbb{K}, +, \cdot)$ ein Körper. Eine nichtleere Menge V, auf der eine Vektoraddition $\oplus\colon V \times V \to V$ und eine Skalarmultiplikation $\odot\colon \mathbb{K} \times V \to V$ definiert sind, heißt $\mathbb{K}$-Vektorraum $(V, \oplus, \odot)$, falls gilt:

1. $(V, \oplus)$ bildet eine abelsche Gruppe;
2. $1 \odot v = v$ für alle $v \in V$ mit $1 \in \mathbb{K}$;
3. $a \odot (b \odot v) = (a \cdot b) \odot v$ für alle $a, b \in \mathbb{K}$ und $v \in V$;
4. Es gelten die Distributivgesetze, d.h. für $a, b \in \mathbb{K}$ und $v, w \in V$ gilt:

$$a \odot (v \oplus w) = (a \odot v) \oplus (a \odot w),$$
$$(a + b) \odot v = (a \odot v) \oplus (b \odot v).$$

Beispiel 6.2. *Sei $(\mathbb{R}, +, \cdot)$ der Körper der reellen Zahlen. Sei $V = \mathbb{R}^3$ mit folgenden Operationen:*

$$v \oplus w = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \oplus \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} := \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix}, \qquad a \odot v = a \odot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} := \begin{pmatrix} a \cdot v_1 \\ a \cdot v_2 \\ a \cdot v_3 \end{pmatrix},$$

wobei $a \in \mathbb{R}$ und $v \in V = \mathbb{R}^3$. Überprüfen wir die geforderten Eigenschaften (mit $u, v, w \in V$ und $a, b \in \mathbb{R}$):

- *$(\mathbb{R}^3, \oplus)$ ist eine abelsche Gruppe:*

$$v \oplus w = w \oplus v, \quad u \oplus (v \oplus w) = (u \oplus v) \oplus w, \quad v \oplus \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = v,$$

  d.h. die Null (das neutrale Element) ist in diesem Fall der Nullvektor, und das Inverse Element ist

$$-v = \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix}, \quad \text{da } v \oplus (-v) = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \oplus \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix} = \begin{pmatrix} v_1 - v_1 \\ v_2 - v_2 \\ v_3 - v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0.$$

  Kommutativität und Assoziativität sind von den Operationen für Matrizen bereits bekannt.

- *In der Tat ist $1 \odot v = v$, denn*

$$1 \odot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 1 \cdot v_1 \\ 1 \cdot v_2 \\ 1 \cdot v_3 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

- *Multiplikation in $\mathbb{R}$ und skalare Multiplikation sind verträglich:*

$$a \odot (b \odot v) = a \odot \left(b \odot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right) = a \odot \begin{pmatrix} b \cdot v_1 \\ b \cdot v_2 \\ b \cdot v_3 \end{pmatrix} = \begin{pmatrix} a \cdot b \cdot v_1 \\ a \cdot b \cdot v_2 \\ a \cdot b \cdot v_3 \end{pmatrix} = (a \cdot b) \odot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = (a \cdot b) \odot v.$$

- *Es gelten die Distributivgesetze:*

$$a \odot (v \oplus w) = a \odot \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix} = \begin{pmatrix} a \cdot (v_1 + w_1) \\ a \cdot (v_2 + w_2) \\ a \cdot (v_3 + w_3) \end{pmatrix} = \begin{pmatrix} a \cdot v_1 + a \cdot w_1 \\ a \cdot v_2 + a \cdot w_2 \\ a \cdot v_3 + a \cdot w_3 \end{pmatrix} = a \odot v \oplus a \odot w,$$

$$(a + b) \odot v = (a + b) \odot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} (a + b) \cdot v_1 \\ (a + b) \cdot v_2 \\ (a + b) \cdot v_3 \end{pmatrix} = \begin{pmatrix} a \cdot v_1 + b \cdot v_1 \\ a \cdot v_2 + b \cdot v_2 \\ a \cdot v_3 + b \cdot v_3 \end{pmatrix} = a \odot v \oplus b \odot v.$$

Wir können also grundsätzlich Vektoren addieren oder mit Elementen aus dem zu Grunde liegenden Körper $\mathbb{K}$ multiplizieren. Eine Multiplikation von Vektoren hingegen ist im Allgemeinen nicht definiert, ebenso wie die Division eines Vektors durch einen anderen.

Die folgenden Eigenschaften sind für den Vektorraum $(\mathbb{R}^n, +, \cdot)$ bekannt. Für allgemeine Vektorräume müssen wir sie jedoch überprüfen.

Satz 6.3

Sei $(V, \oplus, \odot)$ ein $\mathbb{K}$-Vektorraum, wobei 0 und 1 die neutralen Elemente bezüglich der Addition und Multiplikation in $\mathbb{K}$ seien und $e \in V$ das neutrale Element bezüglich der Vektoraddition in V. Weiterhin bezeichnen wir dass additiv Inverse zu $v \in V$ mit $-v \in V$. Dann gilt:

1. $0 \odot v = e$ für alle $v \in V$;
2. $a \odot e = e$ für alle $a \in \mathbb{K}$;
3. $(-a) \odot v = -(a \odot v)$ für alle $a \in \mathbb{K}$ und $v \in V$;
4. Seien $a \in \mathbb{K}$ und $v \in V$. Ist $a \odot v = e$, dann ist $a = 0$ oder $v = e$.

Beweis.

1. Sei $v \in V$ beliebig. Aus $0 \odot v = (0 + 0) \odot v \overset{\text{(DG)}}{=} (0 \odot v) \oplus (0 \odot v)$ folgt, dass $0 \odot v$ das neutrale Element in $(V, \oplus)$ ist, d.h. $0 \odot v = e$.
2. Sei $a \in \mathbb{K}$ beliebig. Man hat analog zu 1.
$$a \odot e = a \odot (e \oplus e) \overset{\text{(DG)}}{=} (a \odot e) \oplus (a \odot e),$$
d.h. $a \odot e$ ist wieder das neutrale Element in $(V, \oplus)$.
3. Seien $a \in \mathbb{K}$ und $v \in V$ beliebig. Dann gilt
$$e \overset{1.}{=} 0 \odot v = (a + (-a)) \odot v \overset{\text{(DG)}}{=} (a \odot v) + ((-a) \odot v),$$
d.h. $(-a) \odot v$ ist das additiv Inverse zu $a \odot v$; in Zeichen: $(-a) \odot v = -(a \odot v)$.
4. Seien $a \in \mathbb{K}$ und $v \in V$. Es gelte $a \odot v = e$ und $a \neq 0$. Dann ist mit Eigenschaften 2 und 3 des Vektorraumes
$$v = 1 \odot v = (a^{-1} \cdot a) \odot v = a^{-1} \odot (a \odot v) \overset{\text{Vor.}}{=} a^{-1} \odot e \overset{2.}{=} e.$$

$\square$

Beispiel 6.4 (Polynomfunktionen). *Ist $(\mathbb{K},+,\cdot)$ ein Körper und sind für $n \in \mathbb{N}_0$ Koeffizienten $a_0, a_1, \ldots, a_n \in \mathbb{K}$, $a_n \neq 0$, gegeben, so wissen wir bereits, dass ein Ausdruck der Form*

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = \sum_{k=0}^{n} a_k x^k$$

Polynom n-ten Grades über $\mathbb{K}$ *genannt wird.*

Eine Abbildung $p \colon \mathbb{R} \to \mathbb{R}$ der Form

$$p(x) = \sum_{k=0}^{n} a_k x^k$$

mit Koeffizienten $a_0, a_1, \ldots, a_n \in \mathbb{R}$ und $a_n \neq 0$ heißt (reelle) Polynomfunktion n-ten Grades.

Sei $\mathcal{P}_n$ die Menge aller reellen Polynomfunktionen vom Grad höchstens n, so bildet diese Menge mit der üblichen Addition von Polynomen und der skalaren Multiplikation einen $\mathbb{R}$-Vektorraum.

6.2 Untervektorräume

Definition 6.5: Untervektorraum

Sei $(V,+,\cdot)$ ein $\mathbb{K}$-Vektorraum und sei $U \subset V$ eine nichtleere Teilmenge von V. U heißt Untervektorraum von V, falls $(U,+,\cdot)$ selbst ein $\mathbb{K}$-Vektorraum ist.

Satz 6.6

Sei $(V,+,\cdot)$ ein $\mathbb{K}$-Vektorraum und sei $U \subset V$ eine nichtleere Teilmenge von V. $(U,+,\cdot)$ ist genau dann ein Untervektorraum von V, wenn

$$\begin{aligned} v + w \in U \quad &\text{für alle } v, w \in U \text{ und} \\ a \cdot v \in U \quad &\text{für alle } v \in U, a \in \mathbb{K} \text{ gilt.} \end{aligned}$$

Bemerkung 6.7. *Man nennt die beiden Eigenschaften auch* Abgeschlossenheit der Addition und skalaren Multiplikation *innerhalb des Vektorraumes $(U,+,\cdot)$. Das bedeutet, dass man durch Addition von Vektoren oder durch skalare Multiplikation keine Vektoren außerhalb des Vektorraumes erreicht.*

Beweis. Zeigen wir zunächst, dass die genannten Eigenschaften ein hinreichendes Kriterium für Unterräume sind:

Gelten KG, AG und DG in $(V,+,\cdot)$, so gelten sie auch innerhalb jeder Teilmenge von V, also insbesondere in $(U,+,\cdot)$. Ebenso gilt

$$1 \cdot v = v,\ \forall v \in V \implies 1 \cdot v = v,\ \forall v \in U.$$

Die Abgeschlossenheit der Addition und skalaren Multiplikation sind nach Voraussetzung gegeben, d.h. die Addition und skalare Multiplikation sind Abbildungen $+\colon U \times U \to U$ und $\cdot\colon \mathbb{K} \times U \to U$.

Es bleibt zu zeigen, dass $(U,+)$ die Eigenschaften (N) und (I) besitzt. Sei dafür e das neutrale Element der Gruppe $(V,+)$. Einerseits enthält jeder Körper die 0 und andererseits ist $U \neq \emptyset$ per Annahme. Mit $0 \in \mathbb{K}$, $v \in U$ und der Abgeschlossenheit bezüglich der skalaren Multiplikation folgt $0 \cdot v = e \in U$. Das additiv Inverse zu einem beliebigen Vektor $v \in U$ ist $-v = (-1) \cdot v \in V$. Da $-1 \in \mathbb{K}$ ist (mit den Körpereigenschaften) und die skalare Multiplikation abgeschlossen ist, gilt also sogar $-v \in U$.

Überprüfen wir nun, dass die Eigenschaften notwendig sind für Untervektorräume. Sei also $(U,+,\cdot)$ ein Untervektorraum von $(V,+,\cdot)$, d.h. $U \subseteq V$ und $(U,+,\cdot)$ ist selbst ein Vektorraum. Dann ist $+\colon U \times U \to U$ eine Verknüpfung und $\cdot\colon \mathbb{K} \times U \to U$ eine Abbildung, deren Ergebnisse jeweils (wie gefordert) in U liegen. Es ist also nichts weiter zu zeigen. $\square$

Beispiel 6.8. *Die Ebene*

$$\mathcal{E} := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; z = 0 \right\}$$

ist ein Untervektorraum des reellen Vektorraumes $V = \mathbb{R}^3$ mit der üblichen Vektoraddition und skalaren Multiplikation. Seien hierzu $v, w \in \mathcal{E}$, d.h. es muss gelten $v_3 = w_3 = 0$. Dann ist

$$v + w = \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ 0 \end{pmatrix} \in \mathcal{E}.$$

Ebenso gilt für $v \in \mathcal{E}$ und $a \in \mathbb{R}$

$$a \cdot v = \begin{pmatrix} a \cdot v_1 \\ a \cdot v_2 \\ a \cdot 0 \end{pmatrix} \in \mathcal{E}.$$

Die folgende im Raum verschobene Ebene ist jedoch kein Untervektorraum des $\mathbb{R}^3$, da die Additivität nicht gegeben ist:

$$\mathcal{E}' := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; z = 5 \right\}.$$

Es ist nämlich

$$\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} \in \mathcal{E}', \quad \begin{pmatrix} 3 \\ 3 \\ 5 \end{pmatrix} \in \mathcal{E}', \quad \textit{aber} \quad \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 10 \end{pmatrix} \notin \mathcal{E}'.$$

6.3 Erzeugendensysteme und linear unabhängige Mengen

6.3.1 Lineare Hülle und Erzeugendensystem

In Abschnitt 6.2 haben wir Untervektorräume eingeführt und ein Kriterium kennen gelernt, mit dessen Hilfe wir untersuchen können, ob eine gegebene Teilmenge eines Vektorraumes selbst einen Vektorraum bildet.

In diesem Abschnitt geht es um die Frage, wie der kleinste Untervektorraum $(U, +, \cdot)$ eines Vektorraumes $(V, +, \cdot)$ aussieht, der gegebene Vektoren $v_1, \ldots, v_n \in V$ enthält.

Definition 6.9: Linearkombination

Sei $(V, +, \cdot)$ ein $\mathbb{K}$-Vektorraum. Für $n \in \mathbb{N}$ seien $v_1, \ldots, v_n \in V$ Vektoren in V und seien zudem $a_1, \ldots, a_n \in \mathbb{K}$ Skalare. Die endliche Summe

$$\sum_{k=1}^{n} a_k v_k = a_1 v_1 + \ldots + a_n v_n$$

heißt dann *Linearkombination* der Vektoren v_1 bis v_n.

Beispiel 6.10. *Es sind* $3 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 0 \cdot \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$ *zwei verschiedene Linearkombinationen der Vektoren* $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ *und* $\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$*, die den gleichen Vektor ergeben.*

Dank der Untervektorraum-Kriterien aus Vorlesung (Satz 6.6) und Übung (Aufgabe 6.4) ist schnell klar, dass für Vektoren $v_1, \ldots, v_n \in U$ eines Untervektorraumes $(U, +, \cdot)$ von $(V, +, \cdot)$ auch jede Linearkombination dieser n Vektoren in U liegt. Daher fassen wir alle Vektoren, die sich als Linearkombination von fest gewählten Vektoren darstellen lassen, zu einer neuen Menge zusammen.

Definition 6.11: lineare Hülle

Die Menge aller endlichen Linearkombinationen aus Vektoren einer Menge heißt *lineare Hülle* oder *Span*, d.h. ist $M \subseteq V$ eine (evtl. unendliche) Teilmenge von V, so ist

$$\mathcal{L}(M) := \left\{ \sum_{k=1}^{n} a_k v_k \;\middle|\; n \in \mathbb{N}, v_1, \ldots, v_n \in M, a_1, \ldots, a_n \in \mathbb{K} \right\}.$$

Ist $M = \{v_1, \ldots, v_m\} \subseteq V$ eine endliche Menge, dann ist

$$\mathcal{L}(M) = \left\{ \sum_{k=1}^{m} a_k v_k \;\middle|\; v_1, \ldots, v_m \in M, a_1, \ldots, a_m \in \mathbb{K} \right\},$$

d.h. die Anzahl an Summanden einer jeden Linearkombination ist gleich. Dies ist jedoch keine wirkliche Einschränkung, da die Koeffizienten $a_k \in \mathbb{K}$ auch Null sein dürfen.

Satz 6.12

Sei $(V, +, \cdot)$ ein $\mathbb{K}$-Vektorraum. Ist $M := \{v_1, \ldots, v_n\} \subseteq V$ mit $n \in \mathbb{N}$, dann ist der Span $\mathcal{L}(M)$ mit den Operationen $+$ und $\cdot$ aus $(V, +, \cdot)$ ein Unterraum von V. Er ist zugleich der kleinste Unterraum von V, der M enthält, d.h. falls $(U, +, \cdot)$ ein UVR von $(V, +, \cdot)$ ist mit $M \subseteq U$, dann gilt $\mathcal{L}(M) \subseteq U$.

Beispiel 6.13. *Für* $M = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} \right\}$ *ist, wie wir gesehen haben,* $\begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} \in \mathcal{L}(M)$. *Es ist*

$$\mathcal{L}(M) = \left\{ a \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \;\middle|\; a \in \mathbb{R} \right\} =: \mathcal{A},$$

denn einerseits ist $\mathcal{A} \subseteq \mathcal{L}(M)$ *und andererseits gilt: Ist*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathcal{L}(M), \quad \text{d.h.} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = a \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + b \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} \quad (a, b \in \mathbb{R}),$$

so ist gleichzeitig

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (a + 2b) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathcal{A},$$

d.h. es ist $\mathcal{L}(M) \subseteq \mathcal{A}$, *woraus* $\mathcal{L}(M) = \mathcal{A}$ *folgt.*

Bemerkung 6.14. *Da die Operationen Addition* $(+)$ *und Multiplikation mit Skalaren* $(\cdot)$ *ab jetzt in der Regel immer durch die gleichen Symbole dargestellt werden, vernachlässigen wir beide ab sofort, d.h. sprechen vom Vektorraum* V *als Kurzform für* $(V, +, \cdot)$.

Beweis. Zeigen wir zunächst mit Hilfe des UVR-Kriteriums (Satz 6.6), dass $\mathcal{L}(M)$ tatsächlich ein UVR von V ist.

Zunächst einmal ist $M \neq \emptyset$, denn aus $n \in \mathbb{N}$ in der Formulierung des Satzes folgt, dass M mindestens einen Vektor $v_1 \in V$ enthält. Somit ist auch $\mathcal{L}(M) \neq \emptyset$, da insbesondere $M \subseteq \mathcal{L}(M)$ ist. Zudem sind Summen und Vielfache von Vektoren aus V wieder in V, so dass $\mathcal{L}(M) \subseteq V$ ist. Überprüfen wir nun, ob Addition und Multiplikation mit Skalaren in $\mathcal{L}(M)$ abgeschlossen sind:

- Seien $u_1, u_2 \in \mathcal{L}(M)$, d.h.

$$\begin{aligned} u_1 &= a_1 v_1 + \ldots + a_n v_n, \quad && \text{mit } a_1, \ldots, a_n \in \mathbb{K}, \\ u_2 &= b_1 v_1 + \ldots + b_n v_n, \quad && \text{mit } b_1, \ldots, b_n \in \mathbb{K}. \end{aligned}$$

  Dann ist auch

$$\begin{aligned} u_1 + u_2 &= (a_1 v_1 + \ldots + a_n v_n) + (b_1 v_1 + \ldots + b_n v_n) \\ &= (a_1 + b_1) v_1 + \ldots + (a_n + b_n) v_n \quad \in \mathcal{L}(M). \end{aligned}$$

- Seien $\lambda \in \mathbb{K}$ und $u \in \mathcal{L}(M)$ mit

$$u = a_1 v_1 + \ldots + a_n v_n, \qquad \text{mit } a_1, \ldots, a_n \in \mathbb{K}.$$

  Dann ist ebenfalls

$$\lambda u = \lambda(a_1 v_1 + \ldots + a_n v_n) = (\lambda a_1) v_1 + \ldots + (\lambda a_n) v_n \in \mathcal{L}(M).$$

Somit ist $\mathcal{L}(M)$ tatsächlich ein UVR von V.

Es bleibt zu zeigen, dass $\mathcal{L}(M)$ der kleinste UVR von V ist, der M enthält. Um dies zu zeigen, nehmen wir an, es sei U ein UVR von V, der M enthält, d.h. $M \subseteq U \subseteq V$. Wir zeigen nun, dass $\mathcal{L}(M) \subseteq U$ ist:

Mit $M = \{v_1, \ldots, v_n\} \subseteq U$ ist nach dem UVR-Kriterium aus der Übung auch $a_1 v_1 + \ldots + a_n v_n \in U$ für beliebige $a_1, \ldots, a_n \in \mathbb{K}$, d.h. alle Linearkombinationen der Vektoren aus M liegen in U. Definitionsgemäß ist also ganz $\mathcal{L}(M) \subseteq U$. □

Definition 6.15

Eine Teilmenge $M \subseteq V$ heißt *Erzeugendensystem von V*, wenn $\mathcal{L}(M) = V$ ist.

Beispiel 6.16. *Folgende Mengen sind Erzeugendensysteme von $\mathbb{R}^3$:*

$$\mathcal{M}_1 := \left\{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \right\},$$

$$\mathcal{M}_2 := \left\{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}1\\3\\5\end{pmatrix} \right\} = \mathcal{M}_1 \cup \left\{ \begin{pmatrix}1\\3\\5\end{pmatrix} \right\},$$

$$\mathcal{M}_3 := \left\{ \begin{pmatrix}1\\4\\-3\end{pmatrix}, \begin{pmatrix}0\\-1\\3\end{pmatrix}, \begin{pmatrix}1\\2\\2\end{pmatrix} \right\}.$$

Menge $\mathcal{M}_1$: *Sei $v = \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix} \in \mathbb{R}^3$ ein beliebiger Vektor des $\mathbb{R}^3$. Es ist $v \in \mathcal{L}(\mathcal{M}_1)$, falls Koeffizienten $a_1, a_2, a_3 \in \mathbb{R}$ existieren, so dass*

$$\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix} = a_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + a_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + a_3 \begin{pmatrix}0\\0\\1\end{pmatrix} \quad \textit{ist.}$$

Dieses Gleichungssystem hat die eindeutige Lösung $(a_1, a_2, a_3) = (v_1, v_2, v_3)$, d.h. $\mathbb{R}^3 = \mathcal{L}(\mathcal{M}_1)$.

Menge $\mathcal{M}_2$: *Wenn sich $v \in \mathbb{R}^3$ als Linearkombination der Vektoren aus $\mathcal{M}_1$ darstellen lässt, so ist dies auch mit Vektoren aus $\mathcal{M}_2$ möglich, so dass $\mathbb{R}^3 = \mathcal{L}(\mathcal{M}_2)$ gilt. Beispielsweise multipliziert man den zusätzlich Vektor immer mit dem Faktor 0 und behält ansonsten die Darstellung bei. Diese Darstellung ist jedoch nicht die einzige Möglichkeit. Am offensichtlichsten ist das für den vierten Vektor von $\mathcal{M}_2$:*

$$\begin{aligned}\begin{pmatrix}1\\3\\5\end{pmatrix} &= 1\cdot\begin{pmatrix}1\\0\\0\end{pmatrix}+3\cdot\begin{pmatrix}0\\1\\0\end{pmatrix}+5\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}+0\cdot\begin{pmatrix}1\\3\\5\end{pmatrix}\\ &= 0\cdot\begin{pmatrix}1\\0\\0\end{pmatrix}+0\cdot\begin{pmatrix}0\\1\\0\end{pmatrix}+0\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}+1\cdot\begin{pmatrix}1\\3\\5\end{pmatrix}.\end{aligned}$$

Menge $\mathcal{M}_3$: *Für die Menge $\mathcal{M}_3$ müssen wir etwas mehr rechnen. Das LGS*

$$\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}=a_1\begin{pmatrix}1\\4\\-3\end{pmatrix}+a_2\begin{pmatrix}0\\-1\\3\end{pmatrix}+a_3\begin{pmatrix}1\\2\\2\end{pmatrix}$$

hat die folgende Darstellung als erweiterte Koeffizientenmatrix:

$$\left[\begin{array}{ccc|c}1&0&1&v_1\\4&-1&2&v_2\\-3&3&2&v_3\end{array}\right]$$

Bezeichnen wir die quadratische Matrix links von der Ergebnisspalte mit A, so können wir A mit Hilfe des Gauss-Jordan-Algorithmus invertieren und erhalten

$$A^{-1}=\begin{pmatrix}-8&3&1\\-14&5&2\\9&-3&-1\end{pmatrix}.$$

Somit ist die Lösung des obigen LGS gegeben durch

$$L=\{A^{-1}\cdot v\}=\left\{\begin{pmatrix}-8v_1+3v_2+v_3\\-14v_1+5v_2+2v_3\\9v_1-3v_2-v_3\end{pmatrix}\right\}.$$

Da die Lösungsmenge für jeden Vektor $v\in\mathbb{R}^3$ nicht-leer ist, ist $\mathbb{R}^3=\mathcal{L}(\mathcal{M}_3)$.

Wie findet man heraus, ob ein Erzeugendensystem minimal ist, d.h. ob ein Vektor weggelassen werden kann ohne dass die lineare Hülle kleiner wird?

6.3.2 Lineare Unabhängigkeit

Definition 6.17

Eine Menge $M\subseteq V$ heißt *linear unabhängig*, wenn für jede Linearkombination von endlich vielen Vektoren $v_1,\ldots,v_n\in M$ gilt:

$$\sum_{k=1}^{n}a_kv_k=a_1v_1+\ldots a_nv_n=e \quad\Longrightarrow\quad a_1=\ldots=a_n=0,$$

wobei $e\in V$ der Nullvektor und $0\in\mathbb{K}$ die Null des Körpers ist.
Die Menge heißt *linear abhängig*, wenn sie nicht linear unabhängig ist.

Beispiel 6.18. *Betrachten wir zwei verschiedene einelementige Mengen:*

- *Die Menge* $M := \{v\} \subseteq \mathbb{R}^3$ *mit* $v := \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$ *ist linear unabhängig, denn*

$$a \cdot \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff a = 0.$$

- *Die Menge* $\widetilde{M} := \{e\} \subseteq \mathbb{R}^3$ *mit* $e := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ *ist linear abhängig, denn*

$$a \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \forall a \in \mathbb{R}.$$

Lemma 6.19

Ist $M_1 \subseteq V$ linear abhängig und ist $M_2 \subseteq V$ beliebig, so ist auch $M_1 \cup M_2$ linear abhängig.

Beweis. Übungsaufgabe 6.8. □

Bemerkung 6.20. *Sollen wir die leere Menge* $\emptyset \subseteq V$ *als linear abhängig oder als linear unabhängig definieren? Nehmen wir an,* $\emptyset \subseteq \mathbb{R}^3$ *wäre linear abhängig. Nach obigem Lemma ist die Menge* $M := M \cup \emptyset$ *mit* M *aus Beispiel 6.18 dann auch linear abhängig. Das widerspricht jedoch dem bereits erbrachten Nachweis, dass* M *linear unabhängig ist. Um solche Widersprüche zu vermeiden, muss die leere Menge als linear unabhängig definiert werden.*

Eine alternative Charakterisierung linear unabhängiger Mengen bietet folgender Satz:

Satz 6.21

Eine Menge $M \subseteq V$ ist genau dann linear unabhängig, wenn kein Vektor $v \in M$ als Linearkombination der anderen Vektoren dargestellt werden kann.

Bemerkung 6.22. *Die* leere Summe *definiert man als*

$$\sum_{j \in J} a_j := e \text{ (Nullvektor)} \quad \text{für} \quad J = \emptyset.$$

Damit lässt sich obiges Lemma auch für einelementige Mengen anwenden.

Beweis von Satz 6.21. Wir beweisen beide Implikationen separat.

"$\Longrightarrow$": Nehmen wir zunächst an, $M \subseteq V$ sei linear unabhängig und dennoch existiere ein Vektor $v \in M$, der als Linearkombination von endlich vielen Vektoren aus $M \setminus \{v\}$ dargestellt werden kann, d.h. wenn wir nur alle Vektoren mit von Null verschiedenen Koeffizienten notieren, so gilt:

$$\exists a_1, \ldots, a_r \in \mathbb{K} \setminus \{0\}, \exists v_1, \ldots, v_n \in M \setminus \{v\} : \sum_{k=1}^{r} a_k v_k = v.$$

Stellen wir diese Gleichung um, so erhalten wir

$$(-1) \cdot v + \sum_{k=1}^{r} a_k v_k = e,$$

was im Widerspruch steht zur Annahme, M sei linear unabhängig. Wir haben also per Widerspruchsbeweis gezeigt, dass aus linearer Unabhängigkeit folgt, dass kein Vektor als Linearkombination der anderen dargestellt werden kann.

"$\Longleftarrow$": Nehmen wir nun an, $M \subseteq V$ sei linear abhängig. Nach Definition existieren dann Vektoren $v_1, \ldots, v_n \in M$ und zugehörige Koeffizienten $a_1, \ldots, a_n \in \mathbb{K}$, so dass

$$\sum_{k=1}^{n} a_k v_k = e \tag{6.3.1}$$

ist, wobei nicht alle Koeffizienten Null sind. Sei o.B.d.A.[1] $a_1 \neq 0$. Dann lässt sich (6.3.1) nach v_1 umstellen:

$$v_1 = -\frac{1}{a_1} \left(\sum_{k=2}^{n} a_k v_k \right) = \sum_{k=2}^{n} \left(-\frac{a_k}{a_1} \right) v_k.$$

v_1 besitzt also in der Tat eine Darstellung als Linearkombination der übrigen Vektoren. Wir haben also mit Hilfe einer Kontraposition gezeigt, dass wenn kein Vektor als Linearkombination der anderen dargestellt werden kann, die Menge linear unabhängig ist.

Somit sind beide Implikationen der Äquivalenzaussage gezeigt. □

Beispiel 6.23. *Die Vektoren aus $\mathcal{M}_1$ aus Beispiel 6.16 sind linear unabhängig: Seien dazu $a, b, c \in \mathbb{R}$. Dann gilt:*

$$a \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff a = b = c = 0.$$

Die Menge $\mathcal{M}_2$ ist hingegen linear abhängig. Zum Beispiel ist

$$(-1) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-3) \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + (-5) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Dies sieht man ebenfalls, da der letzte Vektor als Linearkombination der ersten drei dargestellt werden kann:

$$\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

[1] o.B.d.A. = ohne Beschränkung der Allgemeinheit

Die Begriffe lineare Unabhängigkeit und lineare Hülle zusammen ermöglichen uns, folgendes Resultat zu formulieren:

Satz 6.24

Für eine beliebige Teilmenge M eines $\mathbb{K}$-Vektorraumes V sind folgende Aussagen äquivalent:

(1) M ist linear unabhängig;

(2) Jeder Vektor $v \in \mathcal{L}(M)$ lässt sich in eindeutiger Weise als Linearkombination von Vektoren aus M darstellen.

Beweis. Wir beweisen beide Implikation mit Hilfe der Kontraposition.

(1) $\Longrightarrow$ (2): Angenommen $v \in \mathcal{L}(M)$ habe zwei verschiedene Darstellungsmöglichkeiten als Linearkombination von Vektoren aus M:

$$v = \sum_{k \in I} a_k v_k = \sum_{k \in J} b_k v_k.$$

Da nach Annahme beide Darstellungen verschieden sind, gibt es mindestens einen Index $k^* \in I \cup J$ so dass $a_{k^*} \neq b_{k^*}$ ist. Somit ist

$$\sum_{k \in I \cup J} (a_k - b_k) v_k = e$$

eine Darstellung des Nullvektors, bei der nicht alle Koeffizienten Null sind. Folglich ist M linear abhängig.

(2) $\Longrightarrow$ (1): Sei M linear abhängig, d.h. es existieren Vektoren $v_1, \ldots, v_n \in M$ und Skalare $a_1, \ldots, a_n \in \mathbb{K}$ (nicht alle Null) derart, dass $\sum_{k=1}^n a_k v_k = e$ ist. Angenommen ein Vektor $v \in \mathcal{L}(M)$ habe die Darstellung $v = \sum_{k=1}^m \widetilde{a}_k \widetilde{v}_k$ mit $\widetilde{a}_1, \ldots, \widetilde{a}_m \in \mathbb{K}$ und $\widetilde{v}_1, \ldots, \widetilde{v}_m \in M$. Dann ist

$$v = \left(\sum_{k=1}^n a_k v_k \right) + \left(\sum_{k=1}^m \widetilde{a}_k \widetilde{v}_k \right)$$

eine weitere Darstellung von v als Linearkombination von Vektoren aus M, womit gezeigt ist, dass die Darstellung nicht eindeutig ist.

$\square$

6.3.3 Rang einer Matrix

Definition 6.25

Die maximale Anzahl linear unabhängiger Spaltenvektoren einer Matrix A heißt *Spaltenrang* von A und wir schreiben $\mathrm{rg}_S(A)$. Die maximale Anzahl linear unabhängiger Zeilenvektoren einer Matrix A heißt *Zeilenrang* von A und wir schreiben $\mathrm{rg}_Z(A)$.

Beispiel 6.26. *Betrachten wir die Matrizen*

$$A := \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 7 \end{pmatrix} \quad \text{und} \quad B := \begin{pmatrix} 1 & 2 & 4 \\ -3 & -6 & -12 \\ 2 & 4 & 8 \end{pmatrix}.$$

Es ist leicht zu überprüfen, dass sowohl alle Zeilen als auch alle Spalten von A linear unabhängig voneinander sind, d.h. es gilt $\mathrm{rg}_S(A) = \mathrm{rg}_Z(A) = 3$.

Bei der Matrix B sieht man, dass jeder Spaltenvektor ein Vielfaches des ersten Spaltenvektors ist und jeder Zeilenvektor ein Vielfaches des ersten Zeilenvektors ist. Somit gibt es keine zwei linear unabhängigen Zeilen oder Spalten – es ist also $\mathrm{rg}_S(B) = \mathrm{rg}_Z(B) = 1$.

Es ist keineswegs eine Besonderheit der Matrizen aus dem Beispiel, dass Spalten- und Zeilenrang übereinstimmen.

Satz 6.27: Rang einer Matrix

Sei $A \in \mathbb{K}^{m\times n}$ eine beliebige Matrix. Dann gilt

$$\mathrm{rg}_S(A) = \mathrm{rg}_Z(A),$$

d.h. Spalten- und Zeilenrang der Matrix stimmen überein. Die somit eindeutig bestimmte Zahl $\mathrm{rg}(A) := \mathrm{rg}_S(A)$ nennen wir den *Rang* der Matrix A.

Dies lässt sich anschaulich mit Hilfe von Matrizen in der (normierten) Zeilenstufenform (aus Abschnitt 5.2.2) illustrieren.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & * & \dots & * \\ 0 & 0 & 1 & 0 & 0 & * & \dots & * \\ 0 & 0 & 0 & 1 & 0 & * & \dots & * \\ 0 & 0 & 0 & 0 & 1 & * & \dots & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix} \underset{\rightsquigarrow}{\text{Spalte 2 nach rechts verschieben}} \begin{pmatrix} 1 & 0 & 0 & 0 & * & \dots & * \\ 0 & 1 & 0 & 0 & * & \dots & * \\ 0 & 0 & 1 & 0 & * & \dots & * \\ 0 & 0 & 0 & 1 & * & \dots & * \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Der Satz ist also klar, sobald eine Matrix in der normierten Zeilenstufenform vorliegt. Es bleibt somit noch die Frage, ob durch die elementaren Zeilenumformungen, mit deren Hilfe wir eine Matrix in diese Form überführt haben, der Rang der Matrix verändert wird.

Erinnern wir uns zu diesem Zweck daran, was elementare Zeilenumformungen sind:

Elementare Zeilenoperationen

(Z1) Vertauschen von zwei Zeilen

(Z2) Multiplizieren einer Zeile mit einer von Null verschiedenen Zahl

(Z3) Addieren einer Zeile zu einer anderen Zeile

Analog können wir auch elementare Spaltenumformungen für Matrizen definieren:

Elementare Spaltenoperationen

(S1) Vertauschen von zwei Spalten

(S2) Multiplizieren einer Spalte mit einer von Null verschiedenen Zahl

(S3) Addieren einer Spalte zu einer anderen Spalte

Während wir elementare Spaltenumformungen nicht auf die erweiterte Koeffizientenmatrix anwenden durften, da sich die Lösungsmenge sonst ändern würde, können wir sowohl elementare Zeilen- als auch Spaltenumformungen auf eine Matrix anwenden ohne dass sich deren Rang ändert.

Satz 6.28

Elementare Zeilen- und Spaltenumformungen ändern weder den Zeilenrang einer Matrix noch deren Spaltenrang.

Beweis für (S1) und (Z1). Da die Reihenfolge der Vektoren bei der Untersuchung von deren linearer Unabhängigkeit keine Rolle spielt, ist klar, dass (Z1) den Zeilenrang und (S1) den Spaltenrang unverändert lässt.

Zeigen wir nun, dass (Z1) den Spaltenrang nicht ändert. Dazu sei $A = (\alpha_{k,j}) \in \mathbb{K}^{m\times n}$. Sei zudem $B = (\beta_{k,j}) \in \mathbb{K}^{m\times n}$ die Matrix, die aus A entsteht, wenn die ersten beiden Zeilen vertauscht werden. Insbesondere gilt damit:

- $\beta_{k,j} = \alpha_{k,j}$ für $k \in \{3, \ldots, m\}$ und $j \in \{1, \ldots, n\}$,
- $\beta_{1,j} = \alpha_{2,j}$ und $\beta_{2,j} = \alpha_{1,j}$ für $j \in \{1, \ldots, n\}$.

Wir müssen uns nun vergewissern, dass wenn maximal $l := rg_S(A)$ Spalten von A linear unabhängig voneinander sind, dies auch auf B zutrifft, d.h. es gibt keine $l+1$ Spalten von B, die linear unabhängig sind und gleichzeitig ist es möglich, l Spalten von B auszuwählen, die linear unabhängig sind.

Bezeichnen wir mit $A_j := (\alpha_{k,j})_{k\in\{1,\ldots,m\}}$ den j-ten Spaltenvektor von A und analog mit B_j den j-ten Spaltenvektor von B. Sei zudem $e \in \mathbb{K}^m$ der Nullvektor im Vektorraum $\mathbb{K}^m$.

Nehmen wir o.B.d.A. an, dass es möglich ist, aus den ersten l Spalten von A eine linear unabhängige Menge zu bilden. Dann gilt für Koeffizienten $\lambda_1, \ldots, \lambda_l \in \mathbb{K}$:

$$
\begin{aligned}
\sum_{j=1}^{l} \lambda_j B_j = e &\iff \sum_{j=1}^{l} \lambda_j \beta_{k,j} = 0, \quad \forall k \in \{1, \dots, m\} \\
&\iff \begin{cases} \sum_{j=1}^{l} \lambda_j \beta_{1,j} = 0, \\ \sum_{j=1}^{l} \lambda_j \beta_{2,j} = 0, \\ \sum_{j=1}^{l} \lambda_j \beta_{k,j} = 0, \quad \forall k \in \{3, \dots, m\} \end{cases} \\
&\iff \begin{cases} \sum_{j=1}^{l} \lambda_j \alpha_{2,j} = 0, \\ \sum_{j=1}^{l} \lambda_j \alpha_{1,j} = 0, \\ \sum_{j=1}^{l} \lambda_j \alpha_{k,j} = 0, \quad \forall k \in \{3, \dots, m\} \end{cases} \\
&\iff \sum_{j=1}^{l} \lambda_j \alpha_{k,j} = 0, \quad \forall k \in \{1, \dots, m\} \\
&\iff \sum_{j=1}^{l} \lambda_j A_j = e \\
&\overset{\text{Ann.}}{\iff} \lambda_1 = \ldots = \lambda_l = 0.
\end{aligned}
$$

Folglich sind auch die ersten l Spalten von B linear unabhängig. Mit der gleichen Rechnung sieht man auch, dass durch Hinzunahme eines beliebigen Spaltenvektors $B_j = A_j$ mit $j > l$ (dies ist nur möglich, wenn $l < n$ ist) die Menge aus diesen $l + 1$ Vektoren linear abhängig ist. Somit gilt $\mathrm{rg}_S(B) = l = \mathrm{rg}_S(A)$.

Die Aussage, dass die elementare Spaltenumformung (S1) den Zeilenrang einer Matrix nicht ändert, lässt sich analog zeigen oder indem man den folgenden Zusammenhang zwischen Zeilen- und Spaltenrang einer Matrix und ihrer Transponierten verwendet:

$$\mathrm{rg}_S(A) = \mathrm{rg}_Z(A^T) \quad \text{und} \quad \mathrm{rg}_Z(A) = \mathrm{rg}_S(A^T).$$

Diese Gleichheiten gelten, da die Zeilen von A gerade die Spalten von A^T sind und die Spalten von A gerade die Zeilen von A^T. □

Die übrigen Aussagen zeigt man analog – einige davon als Übungsaufgabe.

6.4 Basis und Dimension

Definition 6.29

Sei $(V, +, \cdot)$ ein Vektorraum. Eine Menge $U \subseteq V$ heißt *Basis von V*, wenn U linear unabhängig und ein Erzeugendensystem von V ist.

Beispiel 6.30. *$\mathcal{M}_1$ und $\mathcal{M}_3$ aus Beispiel 6.16 sind Basen des $\mathbb{R}^3$. Die Menge $\mathcal{M}_2$ ist hingegen keine Basis des $\mathbb{R}^3$, da sie zwar ein Erzeugendensystem ist, jedoch die Vektoren nicht alle linear unabhängig sind.*

Beispiel 6.31. *Sei $V = \{e\}$ der Nullvektorraum. Dann ist $B = \emptyset$ eine Basis von V, denn einerseits hatten wir in Bemerkung 6.20 die leere Menge als linear unabhängig definiert und andererseits in Bemerkung 6.22 die leere Summe als Nullvektor festgelegt.*

Es wird sich im Folgenden als nützlich erweisen, folgende äquivalente Charakterisierungen für Basen zur Verfügung zu haben.

Satz 6.32

Sei $V \neq \{e\}$ ein $\mathbb{K}$-Vektorraum und sei $B = \{v_1, \ldots, v_n\} \subseteq V$ eine endliche Menge von Vektoren aus V. Dann sind folgende Aussagen äquivalent:

(1) B ist eine Basis, d.h. ein linear unabhängiges Erzeugendensystem von V.

(2) B ist ein „unverkürzbares" Erzeugendensystem von V, d.h. für ein beliebiges $r \in \{1, \ldots, n\}$ ist
$$\{v_1, \ldots, v_{r-1}, v_{r+1}, v_n\}$$
kein Erzeugendensystem von V mehr.

(3) Zu jedem $v \in V$ gibt es eindeutig bestimmte $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ mit
$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n,$$
d.h. B ist ein Erzeugendensystem mit einer Eindeutigkeitseigenschaft.

(4) B ist „unverlängerbar" linear unabhängig, d.h. B ist linear unabhängig und für jedes $v \in V$ ist $B \cup \{v\}$ linear abhängig.

Bemerkung 6.33 (Anwendung des Satzes auf $B = \mathcal{M}_1$ aus Beispiel 6.16)**.**
Wir haben bereits überprüft, dass $\mathcal{M}_1$ eine Basis des $\mathbb{R}^3$ ist. Nach Eigenschaft (2) kann keiner der drei Vektoren weggelassen werden ohne dass die Eigenschaft eines Erzeugendensystems verlorengeht. Nach Eigenschaft (3) gibt es für jeden Vektor des $\mathbb{R}^3$ genau eine Darstellung als Linearkombination der Vektoren aus $\mathcal{M}_1$. Zuletzt gilt nach Eigenschaft (4), dass die Menge $\mathcal{M}_2$, da sie $\mathcal{M}_1$ als echte Teilmenge enthält, nicht linear unabhängig sein kann.

Bemerkung 6.34. *Um zwischen Vektoren und Skalaren (d.h. Körperelementen) unterscheiden zu können, bezeichnen wir Vektoren mit lateinischen und Körperelemente mit griechischen Buchstaben.*

Beweis von Satz 6.32. Wir zeigen die Äquivalenz dieser vier Aussagen mit Hilfe eines Ringschlusses.

(1) $\Longrightarrow$ (2): Angenommen B ist ein verkürzbares Erzeugendensystem von V. O.B.d.A. sei $r = 1$, d.h. $\{v_2, \ldots, v_n\}$ ist ein Erzeugendensystem von V. Das bedeutet, $V = \mathcal{L}(\{v_2, \ldots, v_n\})$. Insbesondere ist damit $v_1 \in \mathcal{L}(\{v_2, \ldots, v_n\})$, d.h. es existieren $\lambda_2, \ldots, \lambda_n \in \mathbb{K}$ so dass $v_1 = \lambda_2 v_2 + \ldots + \lambda_n v_n$ ist. Nach Satz 6.21 ist B linear abhängig. Damit ist die Kontraposition der Aussage gezeigt.

(2) $\Longrightarrow$ (3): Sei B ein Erzeugendensystem, welches die Eindeutigkeitseigenschaft verletzt. Dann existiert ein $v \in V$ mit
$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n = \mu_1 v_1 + \ldots + \mu_n v_n$$
und o.B.d.A. $\lambda_1 \neq \mu_1$.

Subtrahieren wir beide Darstellungen voneinander, so erhalten wir

$$\begin{aligned} e &= (\lambda_1 - \mu_1)v_1 + \ldots + (\lambda_n - \mu_n)v_n \\ \iff \quad v_1 &= \frac{\lambda_2 - \mu_2}{\mu_1 - \lambda_1}v_2 + \ldots + \frac{\lambda_n - \mu_n}{\mu_1 - \lambda_1}v_n, \end{aligned}$$

Somit ist $B \setminus \{v_1\}$ immer noch ein Erzeugendensystem von V, denn setzen wir $\beta_k := \frac{\lambda_k - \mu_k}{\mu_1 - \lambda_1}$ für $k \in \{2, \ldots, n\}$, so hat können wir jede Linearkombination der Vektoren aus B mit Hilfe der Vektoren aus $B \setminus \{v_1\}$ darstellen:

$$\sum_{k=1}^{n} \alpha_k v_k = \alpha_1 \underbrace{\sum_{k=2}^{n} \beta_k v_k}_{=v_1} + \sum_{k=2}^{n} \alpha_k v_k = \sum_{k=2}^{n} (\alpha_k + \alpha_1 \beta_k) v_k.$$

Folglich ist B verkürzbar, was die Kontraposition der Aussage beweist.

(3) $\implies$ (4): Es gelte (3). Ist B ein Erzeugendensystem, so ist $\mathcal{L}(B) = V$. Gilt die Eindeutigkeitseigenschaft, so folgt mit Satz 6.24, dass B linear unabhängig ist. Es bleibt zu zeigen, dass B *unverlängerbar* linear unabhängig ist. Sei also $v \in V$ beliebig. Da B ein Erzeugendensystem ist, existieren $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ derart, dass

$$v = \sum_{k=1}^{n} \lambda_k v_k \quad \text{ist.}$$

Da wir v als Linearkombination der Vektoren aus B darstellen können, ist nach Satz 6.21 die Menge $\{v_1, \ldots, v_n, v\} = B \cup \{v\}$ linear abhängig.

(4) $\implies$ (1): Sei B unverlängerbar linear unabhängig. Für den Nachweis, dass B eine Basis ist, muss nur noch gezeigt werden, dass B ein Erzeugendensystem von V ist. Sei $v \in V$ beliebig. Nach Voraussetzung ist $B \cup \{v\}$ linear abhängig, d.h. es existieren $\lambda_1, \ldots, \lambda_n, \lambda \in \mathbb{K}$ derart, dass

$$\lambda_1 v_1 + \ldots + \lambda_n v_n + \lambda v = e \quad \text{ist,}$$

wobei nicht alle Vorfaktoren $\lambda_1, \ldots, \lambda_n, \lambda$ Null sind. Da B linear unabhängig ist, muss insb. $\lambda \neq 0$ sein. Folglich können wir die Gleichung nach v umstellen:

$$v = \sum_{k=1}^{n} -\frac{\lambda_k}{\lambda} v_k.$$

Somit kann v als Linearkombination der Vektoren aus B dargestellt werden. Da $v \in V$ beliebig war, ist B ein Erzeugendensystem von V.

□

Definition 6.35

Sei $(V, +, \cdot)$ ein $\mathbb{K}$-Vektorraum, für den eine endliche Menge $M = \{v_1, \ldots, v_n\} \subseteq V$ derart existiert, dass $\mathcal{L}(M) = V$ ist. Dann nennt man V *endlich erzeugt*.

Beispiel 6.36. *Der Vektorraum* $(\mathbb{R}^n, +, \cdot)$ *ist für beliebiges* $n \in \mathbb{N}$ *ein endlich erzeugter Vektorraum. Für jedes* $n \in \mathbb{N}$ *kann man leicht eine Basis angeben, nämlich*

$$B := \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^n.$$

Wir nennen diese Basis auch Standardbasis *oder* kanonische Basis *des* $\mathbb{R}^n$*. Die jeweils* n *Basisvektoren, die zusammengeschrieben die Einheitsmatrix* E_n *ergeben, nennen wir passend dazu* Einheitsvektoren *und bezeichnen sie mit* $e_1, \dots, e_n$*. Somit können wir die Basis auch folgendermaßen schreiben:*

$$B = \{ e_k \in \mathbb{R}^n \mid k \in \{1, \dots, n\} \}, \text{ wobei } e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k\text{-te Stelle}.$$

Satz 6.32 bezieht sich ausschließlich auf endlich erzeugte Vektorräume. Dennoch kann man aus dem Beweis von (4) $\Longrightarrow$ (1) auch etwas über nicht endlich erzeugte Vektorräume schlussfolgern:

Korollar 6.37

Sei V ein $\mathbb{K}$-Vektorraum, der nicht endlich erzeugt ist. Dann existiert eine unendliche linear unabhängige Familie von Vektoren aus V.

Beweis. Gegeben sei eine endliche linear unabhängige Menge $M := \{v_1, \dots, v_n\}$. Wäre für jeden Vektor $v \in V \setminus M$ die Menge $M \cup \{v\}$ linear abhängig, so könnten wir wie im Beweis (4) $\Longrightarrow$ (1) jeden Vektor aus $V \setminus M$ als Linearkombination aus Vektoren aus M darstellen. Damit wäre M allerdings ein Erzeugendensystem von V, was der Annahme widerspricht, dass V nicht endlich erzeugt ist. □

Beispiel 6.38. *In einer Hausaufgabe (Aufgabe 3 von Blatt 8) wurde bereits gezeigt, dass die reellen Polynomfunktionen vom Grad* $\leq n$ *einen Vektorraum bilden. Betrachten wir nun die Menge aller reellen Polynomfunktionen*

$$\mathcal{P} := \left\{ p\colon \mathbb{R} \to \mathbb{R} \;\middle|\; \exists n \in \mathbb{N}_0, \exists a_0, \dots, a_n \in \mathbb{R} \colon p(x) = \sum_{k=0}^{n} a_k x^k, \forall x \in \mathbb{R} \right\}.$$

Den Beweis für $\mathcal{P}_n$ *kann man direkt auch auf* $\mathcal{P}$ *übertragen, d.h.* $\mathcal{P}$ *mit der Addition von Funktionen und der Multiplikation von Polynomen mit Skalaren aus* $\mathbb{R}$ *bildet einen* $\mathbb{R}$*-Vektorraum.*

Ein kanonisches Erzeugendensystem für $\mathcal{P}$ ist die Menge aller Monome,

$$M := \{p \in \mathcal{P} \mid \exists n \in \mathbb{N}_0, : p(x) = x^n, \forall x \in \mathbb{R}\}.$$

Diese Menge können wir kurz schreiben als $M = \{1, x, x^2, x^3, \dots\}$. Durch direktes Einsetzen in die Definition erhält man sofort, dass $\mathcal{L}(M) = \mathcal{P}$ ist, d.h. M ist in der Tat ein nicht endliches Erzeugendensystem von $\mathcal{P}$. Die Menge M ist auch, wie wir gleich zeigen werden, linear unabhängig, d.h. es ist eine Basis von $\mathcal{P}$. Insbesondere ist es eine unendliche Familie von Vektoren aus $\mathcal{P}$, die linear unabhängig ist – deren Existenz war die Aussage von Korollar 6.37.

Satz 6.39

Sei $\mathcal{P}$ der Vektorraum der rellen Polynomfunktionen. Die Menge der Monome $M = \{1, x, x^2, x^3, \dots\} \subseteq \mathcal{P}$ ist linear unabhängig.

Beweis. Betrachten wir eine beliebige endliche Linearkombination von Monomen, d.h. seien $n \in \mathbb{N}_0$ und $a_0, a_1, \dots, a_n \in \mathbb{R}$.[2] Wir wollen zeigen, dass folgende Implikation gilt:

$$p(x) := \sum_{k=0}^{n} a_k x^k = 0, \forall x \in \mathbb{R} \quad \Longrightarrow \quad a_0 = a_1 = \dots = a_n = 0.$$

Nehmen wir an, mindestens ein Koeffizient von p sei von Null verschieden – o.B.d.A. sei $a_n \neq 0$. Das zur Polynomfunktion p gehörende Polynom in $\mathbb{R}[x]$ hat somit Grad n. Nach Korollar 4.31 hat p also in $\mathbb{R}$ höchstens n Nullstellen. Dies widerspricht der Voraussetzung, dass die Polynomfunktion p unendlich viele Nullstellen hat (nämlich alle reellen Zahlen). Folglich müssen alle Koeffizienten Null sein. $\square$

Wir haben nun bereits einiges über Basen gelernt. Unter anderem haben wir gesehen, dass ein Vektorraum mehr als eine Basis besitzen kann. Allerdings ist eine wichtige Frage bisher noch nicht beantwortet: Gibt es möglicherweise Vektorräume, die keine Basis besitzen?

Basisexistenzsatz 6.40

Jeder Vektorraum besitzt eine Basis.

Man muss bei diesem Satz zwei Fälle unterscheiden – die endlich erzeugten und nicht nicht endlich erzeugten Vektorräume. Für letztere geht der Beweis über den Anspruch dieses Kurses hinaus.[3] Für endlich erzeugte Vektorräume können wir den Satz in folgender Form jedoch beweisen:

[2] Es müssen nicht alle Koeffizienten vorkommen – zur vereinfachten Notation fügen wir die ausgelassenen Koeffizienten jedoch hinzu und setzen sie Null.

[3] Der Beweis findet sich z.B. ab Seite 42 in dem Buch *Lineare Algebra* von S. Bosch [Bos14].

Basisauswahlsatz 6.41

Aus jedem endlichen Erzeugendensystem eines Vektorraumes kann man eine Basis auswählen. Insbesondere hat jeder endlich erzeugte Vektorraum eine endliche Basis.

Beweis. Sei $M := \{v_1, \dots, v_n\} \subseteq V$ ein Erzeugendensystem von V, d.h. $V = \mathcal{L}(M)$. Ist M linear unabhängig, so ist es bereits eine Basis und es gibt nichts zu zeigen. Ist M nicht linear unabhängig, so können wir mindestens einen Vektor aus M als Linearkombination der anderen darstellen. Lassen wir diesen weg und überprüfen, ob die Menge aus $n-1$ Vektoren nun linear unabhängig ist. Dies führen wir so lange durch, bis das Erzeugendensystem unverkürzbar und somit linear unabhängig ist. Da am Anfang nur endlich viele Vektoren zur Auswahl stehen, führt dieses Verfahren nach endlich vielen Schritten zu einer Basis. □

Als letztes wichtiges Resultat zu Basen wollen wir uns mit der Anzahl der Basiselemente beschäftigen. Bisher haben wir mehrere Basen des $\mathbb{R}^3$ kennen gelernt, die alle drei Elemente enthielten. In der Tat gibt es keine Basis des $\mathbb{R}^3$, die echt weniger oder mehr als drei Vektoren enthält. Allgemein können wir dieses Resultat so formulieren:

Satz 6.42

Sei V ein $\mathbb{K}$-Vektorraum. Sind $B = \{b_1, \dots, b_m\}$ und $C = \{c_1, \dots, c_n\}$ zwei verschiedene Basen von V, so gilt $m = n$, d.h. die Anzahl der Basiselemente ist gleich.

Um diesen Satz zu beweisen, brauchen wir folgenden Satz:

Austauschsatz 6.43

Sei V ein $\mathbb{K}$-Vektorraum mit endlicher Basis $B = \{v_1, \dots, v_n\}$. Sei weiterhin $M = \{w_1, \dots, w_m\} \subseteq V$ eine linear unabhängige Teilmenge von V. Dann ist $m \leq n$ und es gibt Indizes $k_1, \dots, k_m \in \{1, \dots, n\}$ derart, dass man nach Austausch der Vektoren $v_{k_1}, \dots, v_{k_m}$ gegen $w_1, \dots, w_m$ wieder eine Basis von V erhält. Nummeriert man die Vektoren geeignet um, so bedeutet das, dass

$$\widetilde{B} := \{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$$

eine Basis von V ist.

Bemerkung 6.44. *Ein wichtiger Teil der Aussage des Austauschsatzes ist die Ungleichung $m \leq n$. Sie sagt aus, dass es nicht mehr linear unabhängige Vektoren als Basisvektoren geben kann. Dies ist der Schlüssel für den Beweis von Satz 6.42.*

Beweis von Satz 6.42. Wenden wir den Austauschsatz auf die Basis B und die linear unabhängige Menge C an, so erhalten wir die Aussage $n \leq m$. Wenden wir den Austauschsatz auf die Basis C und die linear unabhängige Menge B an, so erhalten wir die Aussage $m \leq n$. Zusammen ergibt sich also $m = n$. □

Beweis des Austauschsatzes 6.43. Wir beweisen die Aussage mit Hilfe vollständiger Induktion über $m \in \mathbb{N}_0$.

Induktionsanfang: Für $m = 0$ ist nichts zu zeigen.[4]

Induktionsannahme: Sei $m \geq 1$. Die Aussage sei für $m - 1$ bereits bewiesen.

Induktionsschritt: Aus Lemma 6.19 folgt, dass eine (echte) Teilmenge einer linear unabhängigen Menge selbst linear unabhängig ist. Somit ist insbesondere

$$M' := \{w_1, \ldots, w_{m-1}\}$$

linear unabhängig. Nach Induktionsannahme gilt, dass $m - 1 \leq n$ ist und dass (ggf. nach geeigneter Nummerierung)

$$B_{m-1} := \{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\}$$

eine Basis von V ist. Angenommen, es wäre $m - 1 = n$. Dann wäre allerdings M' bereits eine Basis von V. Dann könnte die Menge $M \supsetneq M'$ nicht linear unabhängig sein (nach Eigenschaft (4) aus Satz 6.32, der „Unverlängerbarkeit“). Da $m - 1 \leq n$ und $m - 1 \neq n$ gilt, folgt, da $m, n \in \mathbb{N}_0$ sind, die behauptete Ungleichung $m \leq n$. Es bleibt noch zu zeigen, dass der m-te Vektor w_m einen der noch verbliebenen Basisvektoren $v_m, \ldots, v_n$ ersetzen kann. Wir schreiben zunächst w_m als Linearkombination der Basisvektoren aus B_{m-1}:

$$w_m = \lambda_1 w_1 + \ldots + \lambda_{m-1} w_{m-1} + \lambda_m v_m + \ldots + \lambda_n v_n. \tag{6.4.1}$$

Wäre $\lambda_m = \ldots = \lambda_n = 0$, so hätten wir eine Darstellung von w_m als Linearkombination der übrigen Vektoren aus M, was der linearen Unabhängigkeit von M widerspricht. Folglich existiert ein Index $k^* \in \{m, \ldots, n\}$ derart, dass $\lambda_{k^*} \neq 0$ ist. Nummerieren wir die Vektoren der Basis ggf. um, so ist $\lambda_m \neq 0$.

Weisen wir nun nach, dass $\widetilde{B}$ tatsächlich eine Basis von V ist.

- *$\widetilde{B}$ ist ein Erzeugendensystem:* Sei $v \in V$ beliebig. Da B_{m-1} eine Basis von V ist, hat v die eindeutige Darstellung

$$v = \sum_{k=1}^{m-1} \mu_k w_k + \sum_{k=m}^{n} \mu_k v_k \tag{6.4.2}$$

mit $\mu_1, \ldots, \mu_n \in \mathbb{K}$. Andererseits folgt wegen $\lambda_m \neq 0$ aus (6.4.1)

$$v_m = \frac{1}{\lambda_m}\left(w_m - \sum_{k=1}^{m-1} \lambda_k w_k - \sum_{k=m+1}^{n} \lambda_k v_k\right)$$

Ersetzen wir v_m in (6.4.2) durch diesen Ausdruck, so erhalten wir

$$\begin{aligned} v &= \sum_{k=1}^{m-1} \mu_k w_k + \frac{\mu_m}{\lambda_m}\left(w_m - \sum_{k=1}^{m-1} \lambda_k w_k - \sum_{k=m+1}^{n} \lambda_k v_k\right) + \sum_{k=m+1}^{n} \mu_k v_k \\ &= \sum_{k=1}^{m-1} \left(\mu_k - \frac{\mu_m \lambda_k}{\lambda_m}\right) w_k + \frac{\mu_m}{\lambda_m} w_m + \sum_{k=m+1}^{n} \left(\mu_k - \frac{\mu_m \lambda_k}{\lambda_m}\right) v_k, \end{aligned}$$

d.h. in der Tat ist $v \in \mathcal{L}(\widetilde{B})$.

[4] Dieser Fall ist tatsächlich der sinnvolle Anfang der Induktion, denn die leere Menge ist linear unabhängig und somit ist $m = 0$ tatsächlich die kleinste Anzahl an Elementen einer linear unabhängigen Teilmenge eines Vektorraumes.

- $\widetilde{B}$ ist linear unabhängig: Für $\mu_1, \ldots, \mu_n \in \mathbb{K}$ gelte

$$\sum_{k=1}^{m} \mu_k w_k + \sum_{k=m+1}^{n} \mu_k v_k = e. \tag{6.4.3}$$

Ersetzen wir w_m durch den Ausdruck in (6.4.1), so erhalten wir

$$\sum_{k=1}^{m-1} \mu_k w_k + \mu_m \left(\sum_{k=1}^{m-1} \lambda_k w_k + \sum_{k=m}^{n} \lambda_k v_k \right) + \sum_{k=m+1}^{n} \mu_k v_k = e.$$

Dies ist äquivalent zu

$$\sum_{k=1}^{m-1} (\mu_k + \mu_m \lambda_k) w_k + \mu_m \lambda_m v_m + \sum_{k=m+1}^{n} (\mu_k + \mu_m \lambda_k) v_k = e.$$

Diese Linearkombination der Vektoren aus B_{m-1} kann wegen ihrer linearen Unabhängigkeit nur dann Null ergeben, wenn alle Koeffizienten Null sind, d.h. wenn gilt:

$$\mu_k + \mu_m \lambda_k = 0, \quad \forall k \in \{1, \ldots, m-1\} \cup \{m+1, \ldots, n\} \tag{6.4.4}$$

und $\mu_m \lambda_m = 0$. Da $\lambda_m \neq 0$ ist, muss $\mu_m = 0$ sein. Damit wird aus (6.4.4)

$$\mu_1 = \ldots = \mu_{m-1} = \mu_{m+1} = \ldots = \mu_n = 0,$$

d.h. aus (6.4.3) folgt, dass alle Koeffizienten Null sind.

□

Korollar 6.45

Sei V ein $\mathbb{K}$-Vektorraum. Besitzt V eine endliche Basis, so ist jede Basis von V endlich.

Beweis. Der Beweis ist eine Übungsaufgabe. □

Insbesondere folgt aus diesem Korollar, dass jeder Vektorraum entweder nur endliche oder nur unendliche Basen besitzt. Diese Feststellung rechtfertigt zusammen mit Satz 6.42 die folgende Definition:

Definition 6.46

Besitzt ein Vektorraum V eine Basis $B = \{b_1, \ldots, b_n\}$ aus n Elementen, so nennt man n die *Dimension* des Vektorraumes und schreibt $\dim(V) = n$. Gibt es keine Basis aus endlich vielen Vektoren, so heißt V *unendlichdimensional* und man schreibt $\dim(V) = \infty$.

Beispiel 6.47. *Einige Beispiele für Vektorräume, die uns schon begegnet sind:*

- *Der Vektorraum* $(\mathbb{R}^n, +, \cdot)$ *ist* n*-dimensional (mit* $n \in \mathbb{N}$*).*
- *Der Vektorraum* $(\mathbb{C}, +, \cdot)$ *ist isomorph zu* $(\mathbb{R}^2, \oplus, \odot)$*, welcher ein 2-dimensionaler* $\mathbb{R}$*-Vektorraum ist.* $\mathbb{C}$ *selbst ist ebenfalls ein 2-dimensionaler* $\mathbb{R}$*-Vektorraum mit* $\{1, i\}$ *als Basis.*
- *Der Vektorraum* $\mathcal{P}$ *als UVR von* $\mathrm{Abb}(\mathbb{R}; \mathbb{R})$ *mit den Operationen*

$$(f+g)(x) := f(x) + g(x), \quad x \in \mathbb{R},$$
$$(a \cdot f)(x) := a \cdot f(x), \quad x \in \mathbb{R}$$

ist ein unendlichdimensionaler Vektorraum mit den Monomen als Basis (Satz 6.39).

6.5 Koordinaten

Satz 6.48

Sei $B = (b_1, \ldots, b_n)$ eine geordnete Basis des $\mathbb{K}$-Vektorraumes V. Dann gibt es für jeden Vektor $v \in V$ genau ein Tupel $(\alpha_1, \ldots, \alpha_n)^T \in \mathbb{K}^n$ mit der Eigenschaft dass

$$v = \sum_{k=1}^{n} \alpha_k b_k.$$

Die Zahlen α_k heißen *Koordinaten von* v bezüglich der Basis B.

Wir schreiben dafür kurz $\kappa_{\mathcal{B}}(v) = (\alpha_1, \ldots, \alpha_n)^T$.

Beweis. Diese Aussage ist gerade Teil (3) von Satz 6.32. □

Bemerkung 6.49. *Betrachten wir die Abbildung* $\kappa_{\mathcal{B}} \colon V \to \mathbb{K}^n$*, die jedem Vektor* $v \in V$ *seine Koordinaten* $(\alpha_1, \ldots, \alpha_n)^T \in \mathbb{K}^n$ *(bezüglich einer geordneten Basis* $\mathcal{B}$ *von* V*) zuordnet. Es ist leicht ersichtlich, dass* $\kappa_{\mathcal{B}}^{-1} \colon \mathbb{K}^n \to V$ *mit* $(\alpha_1, \ldots, \alpha_n)^T \mapsto \sum_{k=1}^{n} \alpha_k b_k$ *die Umkehrabbildung zu* $\kappa_{\mathcal{B}}$ *ist. Folglich ist* $\kappa_{\mathcal{B}}$ *eine bijektive Abbildung. Wir werden später nochmal auf diese Abbildung zurückkommen.*

Beispiel 6.50. *Betrachten wir folgende geordneten Basen*

$$\mathcal{B}_1 := \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \quad \text{und} \quad \mathcal{B}_3 := \left(\begin{pmatrix} 1 \\ 4 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \right)$$

des $\mathbb{R}^3$*, welche den ungeordneten Basen* $\mathcal{M}_1$ *und* $\mathcal{M}_3$ *aus Beispiel 6.16 entsprechen.*[5]

Betrachten wir den Vektor $v := \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \in \mathbb{R}^3$*. Bezüglich der Basis* $\mathcal{B}_1$ *hat* v *die Koordinaten* $\kappa_{\mathcal{B}_1}(v) = (1, 3, 5)^T$*, denn*

$$1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = v.$$

[5] Die veränderte Bezeichnung soll darauf hinweisen, dass dies geordnete Mengen von Vektoren sind.

Bezüglich der Basis $\mathcal{B}_3$ hingegen hat v die Koordinaten $\kappa_{\mathcal{B}_3}(v) = (6, 11, -5)^T$, denn

$$6 \cdot \begin{pmatrix} 1 \\ 4 \\ -3 \end{pmatrix} + 11 \cdot \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix} + (-5) \cdot \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = v.$$

Die Koordinaten findet man durch Lösen des entsprechenden LGS.

Im vorherigen Beispiel ist $\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ gleichzeitig ein Vektor und es sind die Koordinaten des Vektors bezüglich der Standardbasis $\mathcal{B}_1$ des $\mathbb{R}^3$. Sehen wir uns daher noch ein anderes Beispiel an, in dem $\mathbb{K}^n \neq V$ ist:

Beispiel 6.51. *Es sind*

$$\mathcal{B} := (1, x, x^2) \quad \textit{und} \quad \mathcal{C} := (1 - x, 1 + x, x^2)$$

zwei geordnete Basen des $\mathbb{R}$-Vektorraumes $\mathcal{P}_2$ der reellen Polynomfunktionen von höchstens Grad 2. Sei $p(x) = a + bx + cx^2 \in \mathcal{P}_2$ (mit $a, b, c \in \mathbb{R}$) ein beliebiges Element von $\mathcal{P}_2$. Dann liest man die Koordinaten bezüglich $\mathcal{B}$ direkt ab:

$$\kappa_{\mathcal{B}}(p) = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Für die Koordinaten bezüglich $\mathcal{C}$ beachtet man, dass

$$\begin{aligned} \frac{1}{2} \cdot (1 - x) + \frac{1}{2} \cdot (1 + x) + 0 \cdot x^2 &= 1, \\ \left(-\frac{1}{2}\right) \cdot (1 - x) + \frac{1}{2} \cdot (1 + x) + 0 \cdot x^2 &= x, \\ 0 \cdot (1 - x) + 0 \cdot (1 + x) + 1 \cdot x^2 &= x^2 \end{aligned}$$

ist. Damit ist

$$\begin{aligned} a + bx + cx^2 &= a \cdot \left[\frac{1}{2} \cdot (1 - x) + \frac{1}{2} \cdot (1 + x)\right] + b \cdot \left[\left(-\frac{1}{2}\right) \cdot (1 - x) + \frac{1}{2} \cdot (1 + x)\right] + c \cdot x^2 \\ &= \frac{1}{2}(a - b) \cdot (1 - x) + \frac{1}{2}(a + b)(1 + x) + c \cdot x^2, \end{aligned}$$

d.h.

$$\kappa_{\mathcal{C}}(p) = \begin{pmatrix} \frac{1}{2}(a - b) \\ \frac{1}{2}(a + b) \\ c \end{pmatrix}.$$

6.6 LGS und Unterverktorräume

6.6.1 Fragestellung

Betrachten wir nochmals lineare Gleichungssysteme, d.h. Systeme der Form

$$\begin{aligned} \alpha_{1,1}x_1 + \ldots + \alpha_{1,n}x_n &= b_1 \\ \alpha_{2,1}x_1 + \ldots + \alpha_{2,n}x_n &= b_2 \\ &\vdots \\ \alpha_{m,1}x_1 + \ldots + \alpha_{m,n}x_n &= b_m, \end{aligned}$$

wobei alle Koeffizienten $\alpha_{k,j}$ und β_k Elemente eines Körpers $\mathbb{K}$ seien und die Unbekannten $x_1, \ldots, x_n$ ebenfalls in der Regel innerhalb desselben Körpers liegen sollen. Betrachten wir für diesen Abschnitt ausschließlich den Körper $\mathbb{K} = \mathbb{R}$.

Die zum LGS gehörige Schreibweise als *erweiterte Koeffizientenmatrix* ist $(A|b)$, d.h.

$$\left[\begin{array}{ccc|c} \alpha_{1,1} & \cdots & \alpha_{1,n} & b_1 \\ \alpha_{2,1} & \cdots & \alpha_{2,n} & b_2 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & b_m \end{array}\right],$$

was wiederum eine Abkürzung des Problems ist, den Vektor x zu bestimmen, der die Gleichung

$$A \cdot x = b$$

löst, wobei

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \cdots & \alpha_{2,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{R}^{m \times n}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m \quad \text{ist.}$$

Wir wissen bereits, dass ein reelles LGS entweder keine, eine oder unendlich viele Lösungen im $\mathbb{R}^n$ besitzt. Im Folgenden wollen einerseits untersuchen, ob Lösungsmengen von LGS Vektorräume sind und andererseits können wir mit Hilfe der neuen Begriffe die Untersuchung von LGS auf Lösbarkeit vereinfachen.

6.6.2 Lösbarkeit von LGS – einführende Beispiele

Im Folgenden betrachten wir immer LGS vom Typ $A \cdot x = b$ und nennen A die Koeffizientenmatrix des LGS und $(A|b)$ wie bisher die erweiterte Koeffizientenmatrix.

Beispiel 6.52. *Seien*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \textit{und} \quad \widetilde{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

In Aufgabe 2.5 hatten wir bereits deren Inverse berechnet:

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

Die Lösungen der Gleichungssysteme $A \cdot x = b$ und $A \cdot x = \widetilde{b}$ lassen sich nun einfach berechnen:

$$L = \left\{A^{-1} \cdot b\right\} = \left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\} \quad \text{und} \quad \widetilde{L} = \left\{A^{-1} \cdot \widetilde{b}\right\} = \left\{\begin{pmatrix} -1 \\ 4 \end{pmatrix}\right\}.$$

Einerseits sehen wir an dieser Stelle exemplarisch, dass für jeden Vektor $b \in \mathbb{R}^2$ das LGS $A \cdot x = b$ mit der gegebenen Matrix A genau eine Lösung besitzt, nämlich $x = A^{-1} \cdot b$.

Andererseits ist zwar L ein Untervektorraum des $\mathbb{R}^2$, $\widetilde{L}$ hingegen nicht.

Beispiel 6.53. *Seien nun*

$$A = \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{und dazu} \quad b_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad b_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

A ist in diesem Fall nicht invertierbar, denn man kann A nicht mit Hilfe elementarer Zeilenumformungen in die Einheitsmatrix überführen. Somit müssen wir die Gleichungssysteme direkt lösen.

$$\begin{aligned} 2x_1 - x_2 &= 0 \\ 0x_1 + 0x_2 &= 0 \end{aligned}$$

hat die freie Variable x_2, da die zweite Spalte keine Pivotspalte des LGS in Zeilenstufenform ist. Damit hat das LGS mit b_1 als rechter Seite einen Vektorraum als Lösungsmenge, nämlich[6]

$$\begin{aligned} L_1 &= \left\{ x \in \mathbb{R}^2 \;\middle|\; x_1 = \frac{1}{2}x_2 \right\} = \left\{ x_2 \cdot \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} \;\middle|\; x_2 \in \mathbb{R} \right\} = \left\{ a \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} \;\middle|\; a \in \mathbb{R} \right\} \\ &= \mathcal{L}\left(\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}\right) = \mathcal{L}\begin{pmatrix} 1 \\ 2 \end{pmatrix}. \end{aligned}$$

Betrachten wir nun das LGS $A \cdot x = b_2$:

$$\begin{aligned} 2x_1 - x_2 &= 2 \\ 0x_1 + 0x_2 &= 3 \end{aligned}$$

ist nicht lösbar, d.h. $L_2 = \emptyset$.

Zu guter Letzt betrachten wir das LGS $A \cdot x = b_3$:

$$\begin{aligned} 2x_1 - x_2 &= 1 \\ 0x_1 + 0x_2 &= 0 \end{aligned}$$

ist lösbar. Die Lösungsmenge

$$L_3 = \left\{ x \in \mathbb{R}^2 \;\middle|\; 2x_1 - x_2 = 1 \right\} = \left\{ \begin{pmatrix} \frac{1}{2} + \frac{1}{2}x_2 \\ x_2 \end{pmatrix} \;\middle|\; x_2 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} + a \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} \;\middle|\; a \in \mathbb{R} \right\}$$

ist in diesem Fall jedoch kein Vektorraum.

[6]Zur besseren Lesbarkeit lassen wir bei linearen Hüllen von einem einzigen Vektor die zwei äußeren Klammern ab jetzt weg.

6.6.3 Vorarbeit: Summen von Mengen

Wir haben gerade gesehen, dass Lösungsmengen nicht immer Vektorräume sind, jedoch fast die Struktur einer linearen Hülle besitzen. Diesen Gedanken wollen wir genauer mit Hilfe geeigneter mathematischer Begriffe in Worte fassen:

Definition 6.54

Seien A und B beliebige Teilmengen eines Vektorraumes V. Dann ist $A+B$ definiert als

$$\begin{aligned} A + B :&= \{a + b \mid a \in A, b \in B\} && (6.6.1)\\ &= \{v \in V \mid \exists a \in A, b \in B\colon a + b = v\}\,. && (6.6.2) \end{aligned}$$

Bemerkung 6.55. *Warum ist es sinnvoll, neben der Darstellung* (6.6.1) *noch die zusätzliche Darstellung* (6.6.2) *anzugeben? Sehen wir uns dazu folgendes einfaches Beispiel an: Zu* $A = B = \{0, 1\}$ *ist*

$$A + B = \{0 + 0, 0 + 1, 1 + 0, 1 + 1\} = \{0, 1, 1, 2\},$$

d.h. ein Element der Menge wurde doppelt gezählt. Dies wird bei der zweiten Darstellung vermieden:

$$A + B = \{0, 1, 2\},$$

denn $0 = 0 + 0$*,* $1 = 0 + 1 = 1 + 0$ *und* $2 = 1 + 1$.

Die Definition können wir auf mehr als nur zwei Mengen anwenden. Durch Anwendung auf Untervektorräume erhalten wir einen neuen UVR:

Satz 6.56: Summe von Untervektorräumen

Sei V ein $\mathbb{K}$-Vektorraum mit Untervektorräumen $U_1, \ldots, U_n$. Dann ist

$$U_1 + \ldots + U_n := \{v \in V \mid \exists v_1 \in U_1, \ldots, v_n \in U_n : v = v_1 + \ldots + v_n\}$$

ein UVR von V.

Beweis. Bezeichne e den Nullvektor in V. Da $U_1, \ldots, U_n$ Untervektorräume von V sind, ist insbesondere $e \in U_k$ für alle $k \in \{1, \ldots, n\}$. Damit ist insbesondere $e = e + \ldots + e \in U_1 + \ldots + U_n$, d.h. $U_1 + \ldots + U_n \neq \emptyset$. Zudem ist $U_1 + \ldots + U_n$ nach Definition eine Teilmenge von V.

Überprüfen wir nun das UVR-Kriterium, d.h. die Abgeschlossenheit bezüglich Addition und Multiplikation mit Skalaren:

- Seien $u, v \in U_1 + \ldots + U_n$ mit

  $$u = u_1 + \ldots + u_n \quad \text{und} \quad v = v_1 + \ldots + v_n,$$

  wobei $u_k, v_k \in U_k$ seien für $k \in \{1, \ldots, n\}$. Somit ist

  $$u + v = (u_1 + v_1) + \ldots + (u_n + v_n).$$

Da U_1 bis U_n selbst das UVR-Kriterium erfüllen, ist

$$u_k + v_k \in U_k, \quad \forall k \in \{1, \dots, n\}.$$

Somit ist insgesamt auch $u + v \in U_1 + \ldots + U_n$.

- Sei $v \in U_1 + \ldots + U_n$ mit der gleichen Darstellung wie eben und sei $\lambda \in \mathbb{K}$. Dann ist

$$\lambda v = \lambda(v_1 + \ldots + v_n) = \lambda v_1 + \ldots + \lambda v_n \in U_1 + \ldots + U_n,$$

da jeder einzelne Summand auf Grund des UVR-Kriteriums in einem der Unterräume liegt.

□

Bemerkung 6.57. *Wir wissen bereits, dass für beliebige Teilmengen $M_1, M_2 \subseteq V$ eines $\mathbb{K}$-Vektorraumes V die linearen Hüllen $\mathcal{L}(M_1)$ und $\mathcal{L}(M_2)$ Untervektorräume von V sind. Aus dem Satz können wir nun unmittelbar folgern, dass auch $\mathcal{L}(M_1) + \mathcal{L}(M_2)$ ein UVR von V ist.*

Die Summe von Mengen ist ein Begriff, der uns hilft, die folgende Definition zu verstehen:

Definition 6.58: affiner Unterraum

Sei V ein $\mathbb{K}$-Vektorraum und sei $W \subseteq V$ eine Teilmenge von V. W heißt *affiner Unterraum* von V, falls ein $v \in V$ und ein UVR $U \subseteq V$ existieren, so dass

$$W = v + U = \{w \in V \mid \exists u \in U \colon w = v + u\} \quad \text{ist.}$$

Bemerkung 6.59. *Während die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(x) = ax$ (für $a \in \mathbb{R}$) als* lineare *Abbildung bezeichnet wird, heißt die etwas allgemeinere Funktion $g\colon \mathbb{R} \to \mathbb{R}$ mit $g(x) = ax + b$ ($a, b \in \mathbb{R}$)* affine *Abbildung. Analog dazu gibt es passend zur Bezeichnung von $v + U$ als* affinen *Unterraum die Bezeichnung* linearen *Unterraum für den UVR U.*

Auf den ersten Blick scheint bei festem UVR U für jede Wahl von $v \in V$ ein anderer affiner Unterraum zu entstehen. Dies ist jedoch nicht unbedingt der Fall, wie wir uns zunächst anhand eines Beispiels verdeutlichen wollen.

Beispiel 6.60. *Für $V = \mathbb{R}^3$ seien*

$$v = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}, \; v' = \begin{pmatrix} 0 \\ 5 \\ -1 \end{pmatrix} \;\; \text{und} \;\; U = \left\{ a \cdot \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \;\middle|\; a \in \mathbb{R} \right\} = \mathcal{L}\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}.$$

Wir haben einerseits $v' \in v + U$ (mit $a = -1$) und andererseits $v \in v' + U$ (mit $a = 1$).

Somit können wir zeigen, dass $v + U = v' + U$ ist:

$$\begin{aligned}
w \in v' + U &\iff \exists a \in \mathbb{R} \colon w = \begin{pmatrix} 0 \\ 5 \\ -1 \end{pmatrix} + a \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \\
&\iff \exists a \in \mathbb{R} \colon w = \left[\begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right] + a \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \\
&\iff \exists a \in \mathbb{R} \colon w = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + (a-1) \cdot \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \\
&\iff \exists b \in \mathbb{R} \colon w = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \\
&\iff w \in v + U.
\end{aligned}$$

Allgemein gilt folgender Satz:

Satz 6.61

Seien V ein $\mathbb{K}$-Vektorraum, $v \in V$ ein beliebiger Vektor und $U \subseteq V$ ein UVR von V. Sei $v + U$ der für diese Wahl entstehende affine Unterraum. Dann gilt:

1. Für jedes $w \in v + U$ ist $v + U = w + U$.
2. Gilt $v + U = v' + U'$ für $v' \in V$ und einen UVR $U' \subseteq V$, so gilt $U = U'$ und $v' - v \in U$.

Beweis.

1. Wir zeigen die Mengengleichheit, indem wir beide Inklusionen nachweisen. Sei dazu $w \in v + U$ beliebig mit der Darstellung $w = v + u_w$ für ein $u_w \in U$.

   $\underline{v + U \subseteq w + U}$:

   $$\begin{aligned}
   w' \in v + U &\Longrightarrow w' = v + u' \ \text{ für ein } u' \in U \\
   &\Longrightarrow w' = (w - u_w) + u' \\
   &\Longrightarrow w' = w + (u' - u_w) \quad \text{mit } u', u_w \in U, \text{ also } u' - u_w \in U \\
   &\Longrightarrow w' \in w + U.
   \end{aligned}$$

   $\underline{w + U \subseteq v + U}$: Analog zur vorherigen Inklusion gilt

   $$\begin{aligned}
   w'' \in w + U &\iff \exists u'' \in U : w'' = w + u'' \\
   &\Longrightarrow w'' = (v + u_w) + u'' = v + (u_w + u'') \in v + U.
   \end{aligned}$$

2. Es gelte $v + U = v' + U'$. Da $v' \in v' + U'$ ist, folgt hieraus $v' \in v + U$, d.h. es existiert ein $u \in U$ mit $v' = v + u$. Folglich ist $v' - v = u \in U$ (wie behauptet).

Für die Behauptung zu den Untervektorräumen schreiben wir die Differenz zweier Teilmengen eines VR (nicht zu verwechseln mit der allgemeinen Differenzmenge $A \setminus B$) analog zu deren Summe, d.h.

$$A - B := \{a - b \mid a \in A, b \in B\}$$

Somit ist

$$\begin{aligned}(v + U) - (v + U) &= \{w \in v + U \mid \exists u_1, u_2 \in U \colon w = (v + u_1) - (v + u_2) = u_1 - u_2\} \\ &= U \quad \text{und} \\ (v' + U') - (v' + U') &= \{w \in v' + U' \mid \exists u_1', u_2' \in U' \colon w = (v' + u_1') - (v' + u_2') = u_1' - u_2'\} \\ &= U'\end{aligned}$$

Aus $v + U = v' + U'$ folgt somit $U = U'$.

□

Der zweite Teil des Satzes liefert ein notwendiges Kriterium für die Gleichheit zweier affiner Unterräume.

Beispiel 6.62. *Seien*

$$W_1 = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + \mathcal{L}\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \quad \textit{und} \quad W_2 = \begin{pmatrix} 0 \\ 7 \\ 4 \end{pmatrix} + \mathcal{L}\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}.$$

Es ist $W_1 \neq W_2$, denn

$$\begin{pmatrix} 0 \\ 7 \\ 4 \end{pmatrix} - \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix} \notin \mathcal{L}\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}.$$

Auch bei affinen Unterräumen sprechen wir von Dimensionen.

Definition 6.63

Sei V ein $\mathbb{K}$-Vektorraum mit affinem Unterraum $W := v + U$. Die Dimension von W ist definiert als die Dimension des zugehörigen (linearen) Untervektorraumes U, d.h.

$$\dim(W) := \dim(U).$$

Bemerkung 6.64. *Die leere Menge $\emptyset \subseteq V$ wird auch als affiner Unterraum aufgefasst; ihre Dimension wird als -1 festgelegt. Wir werden später sehen, warum das sinnvoll ist.*

6.6.4 Geraden und Ebenen als affine Unterräume des $\mathbb{R}^3$

Bevor wir zu den eher abstrakten Lösungsmengen linearer Gleichungssysteme zurückkommen, wollen wir zunächst ein paar sehr anschauliche Beispiele für affine Unterräume des $\mathbb{R}^3$ betrachten – Geraden und Ebenen.

Geraden und Ebenen sind elementare Objekte der Geometrie – sie auf elementarem Niveau zu definieren ist hingegen nicht einfach. Nehmen wir folgende Sätze, die man so wörtlich im Internet finden kann:

- „Die Gerade in der Mathematik ist ein Strich ohne Kurven. Die Gerade hat keinen Anfang und kein Ende.“ (Was ist dann eine Kurve? Wie ist ein Strich definiert?)
- „Eine Gerade ist eine geometrische Figur die unendlich lang, unendlich dünn und in beide Richtungen unbegrenzt ist.“ (Hier fehlt die entscheidende Eigenschaft, dass Geraden gerade sind.)
- „Die kürzeste Verbindung zweier Punkte ist gerade und wird als Strecke bezeichnet. Eine gerade, unendlich lange, unendlich dünne und in beide Richtungen unbegrenzte Linie nennt man eine Gerade.“ (aus Wikipedia)

Demgegenüber können wir mit dem gerade erworbenen Wissen zu affinen Unterräumen Geraden und Ebenen wunderbar exakt definieren:

Definition 6.65

Im n-dimensionalen $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ heißt ein eindimensionaler affiner Unterraum *(affine) Gerade* und ein zweidimensionaler affiner Unterraum *(affine) Ebene*.

Somit finden wir speziell für den $\mathbb{R}^3$ die als *Parameterform* bekannte Darstellung von Geraden und Ebenen:

- Jede (affine) Gerade im $\mathbb{R}^3$ hat eine Darstellung als

$$G = v + \mathcal{L}(u) = \{v + a \cdot u \mid a \in \mathbb{R}\}$$

für feste Vektoren $v, u \in \mathbb{R}^3$, wobei $\{u\}$ linear unabhängig, also u nicht der Nullvektor sei.

- Jede (affine) Ebene im $\mathbb{R}^3$ hat eine Darstellung als

$$E = v + \mathcal{L}(u) + \mathcal{L}(\widetilde{u}) = \{v + a \cdot u + b \cdot \widetilde{u} \mid a, b \in \mathbb{R}\}$$

für feste Vektoren $v, u, \widetilde{u} \in \mathbb{R}^3$, wobei $\{u, \widetilde{u}\}$ linear unabhängig sei, also u und $\widetilde{u}$ nicht kollinear sein dürfen (d.h. $\widetilde{u} \notin \mathcal{L}(u)$).

Man nennt v Stützvektor und u und $\widetilde{u}$ Richtungsvektoren der Gerade bzw. Ebene. Nach Satz 6.61 ist eine solche Darstellung nicht eindeutig.

Beispiel 6.66. *Die beiden affinen Unterräume aus Beispiel 6.62,*

$$W_1 = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + \mathcal{L}\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \quad \textit{und} \quad W_2 = \begin{pmatrix} 0 \\ 7 \\ 4 \end{pmatrix} + \mathcal{L}\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix},$$

sind somit (parallele) Geraden. Die folgende Menge ist hingegen eine Ebene, da die beide Richtungsvektoren linear unabhängig sind:

$$W_3 = \left\{ \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} + a \cdot \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}$$

6.6.5 Lösungsmengen von LGS – affine Unterräume

Kommen wir nun zurück zu der Ausgangsfrage dieses Abschnitts – wie sehen die Lösungsmengen linearer Gleichungssysteme aus und können wir anhand der Koeffizientenmatrix bereits etwas über die Struktur der Lösungsmenge voraussagen? Wir beschränken und hier wieder auf endlich-dimensionale reelle Vektorräume und bezeichnen den Nullvektor des $\mathbb{R}^m$ mit $0_m \in \mathbb{R}^m$.

In Abschnitt 6.6.2 haben wir bereits einige Lösungsmengen linearer Gleichungssysteme gesehen und alle sind affine Unterräume (einschließlich der leeren Menge $\emptyset$).

Um systematisch die verschiedenen Möglichkeiten für Lösungsmengen untersuchen zu können, benötigen wir folgenden Begriff:

Definition 6.67: homogenes LGS

Seien $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$.

Ein LGS der Form $A \cdot x = b$ heißt

- *homogenes LGS*, falls $b = 0_m$ der Nullvektor ist;
- *inhomogenes LGS*, falls b nicht der Nullvektor ist.

Notation

Wenn wir die Lösungsmenge zu einem LGS der Form $A \cdot x = b$ notieren, so schreiben wir statt L auch $L(A, b)$ um die Lösungsmengen für verschiedene Matrizen und Vektoren vergleichen zu können.

Wir werden zunächst ein Kriterium finden, welches uns sagt, ob ein LGS eine Lösung hat, und anschließend die Struktur der Lösungsmenge genauer untersuchen.

Existenz einer Lösung

Satz 6.68

Ein LGS der Form $A \cdot x = b$ mit $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$ ist genau dann lösbar, wenn $\mathrm{rg}(A) = \mathrm{rg}(A|b)$ ist, wobei $(A|b)$ die erweiterte Koeffizientenmatrix ist.

Beweis. Bezeichnen wir die Spaltenvektoren von A wieder mit $A_1, \ldots, A_n$.

Sei $\mathrm{rg}(A) = l$. Das heißt, es existieren l Spaltenvektoren von A, die linear unabhängig sind. O.B.d.A. seien das die ersten l Spalten, $A_1, \ldots, A_l$.

- Ist $\mathrm{rg}(A|b) = l$, so ist insbesondere $\{A_1, \ldots, A_l, b\}$ linear abhängig und man kann b als Linearkombination dieser l Spaltenvektoren schreiben:

$$b = \sum_{k=1}^{l} \lambda_k A_k.$$

Damit ist

$$\sum_{k=1}^{l} \lambda_k A_k + \sum_{k=l+1}^{n} 0 \cdot A_k = \begin{pmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_n \\ | & | & & | \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_l \\ 0 \\ \vdots \\ 0 \end{pmatrix} = b.$$

Folglich ist

$$x = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_l \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$$

eine Lösung von $A \cdot x = b$.

- Angenommen $x = (x_k)_{k \in \{1,\ldots,n\}} \in \mathbb{R}^n$ ist eine Lösung von $A \cdot x = b$, d.h.

$$A \cdot x = \begin{pmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_n \\ | & | & & | \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{k=1}^{n} x_k A_k = b$$

Somit können wir b als Linearkombination der Spalten von A darstellen, d.h. $\{A_1, \ldots, A_n, b\}$ ist linear abhängig. Folglich ist $\mathcal{L}(\{A_1, \ldots, A_n\}) = \mathcal{L}(\{A_1, \ldots, A_n, b\})$. Nach Aufgabe 6.17 ist somit

$$\mathrm{rg}(A|b) = \dim(\mathcal{L}(\{A_1, \ldots, A_n, b\})) = \dim(\mathcal{L}(\{A_1, \ldots, A_n\})) = \mathrm{rg}(A).$$

□

Homogene LGS

Betrachten wir das homogene LGS $A \cdot x = 0_m$ für $A \in \mathbb{R}^{m \times n}$, so können wir eine Lösung sofort erkennen: den Nullvektor $x = 0_n$.

Satz 6.69

Seien $x_1, \ldots, x_r \in \mathbb{R}^n$ Lösungen des homogenen LGS $A \cdot x = 0_m$. Dann ist auch jede Linearkombination dieser Vektoren Lösung des LGS, d.h. $\mathcal{L}(\{x_1, \ldots, x_r\}) \subseteq L(A, 0_m)$.

Beweis. Nach Voraussetzung ist $A \cdot x_1 = \ldots = A \cdot x_r = 0_m$. Nach Satz 2.19 ist für beliebige $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$

$$A \cdot \left(\sum_{k=1}^{r} \lambda_k x_k \right) = \sum_{k=1}^{r} A \cdot (\lambda_k x_k) = \sum_{k=1}^{r} \lambda_k (A \cdot x_k) = \sum_{k=1}^{r} \lambda_k \cdot 0_m = 0_m.$$

Folglich ist auch $\sum_{k=1}^{r} \lambda_k x_k$ eine Lösung des LGS und die Menge der Linearkombinationen ist gerade die lineare Hülle. □

Satz 6.70

Die Lösungsmenge eines homogenen LGS $A \cdot x = 0_m$ für $A \in \mathbb{R}^{m \times n}$ ist ein Untervektorraum des $\mathbb{R}^n$.

Beweis. In Aufgabe 6.16 wurde bereits gezeigt, dass der Kern des VR-Homomorphismus $\varphi\colon V \to W$ ein UVR von V ist. Dieses Ergebnis wenden wir auf $\varphi(x) = A \cdot x$ (mit $V = \mathbb{R}^n$, $W = \mathbb{R}^m$) an. □

Inhomogene LGS

Satz 6.71

Seien $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Sei $x^* \in L(A, b)$ eine beliebige Lösung des inhomogenen LGS $A \cdot x = b$. Dann ist

$$L(A, b) = x^* + L(A, 0_m).$$

Beweis. Um die Gleichheit der Mengen $L(A, b)$ und $x^* + L(A, 0_m)$ zu zeigen, zeigen wir beide Inklusionen.

$L(A, b) \supseteq x^* + L(A, 0_m)$**:** Sei $x \in x^* + L(A, 0_m)$, d.h. es existiert $x_h \in L(A, 0_m)$ mit $x = x^* + x_h$. Dann ist insbesondere $A \cdot x_h = 0_m$ und somit

$$A \cdot x = A \cdot (x^* + x_h) = A \cdot x^* + A \cdot x_h = A \cdot x^* = b,$$

d.h. es ist in der Tat $x \in L(A, b)$.

$L(A, b) \subseteq x^* + L(A, 0_m)$**:** Sei $x \in L(A, b)$, d.h. $A \cdot x = b$. Damit gilt

$$A \cdot (x - x^*) = A \cdot x - A \cdot x^* = b - b = 0_m,$$

d.h. $x - x^* \in L(A, 0_m)$. Das bedeutet, es existiert ein $x_h \in L(A, 0_m)$ mit $x - x^* = x_h$, also $x = x^* + x_h$. Folglich ist in der Tat $x \in x^* + L(A, 0_m)$.

□

Korollar 6.72

Die Lösungsmenge eines inhomogenen LGS der Form $A \cdot x = b$ mit $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$ ist ein affiner Unterraum des $\mathbb{R}^n$.

Beweis. Ist $L(A,b)=\emptyset$, so ist nichts zu zeigen, da festgelegt wurde, dass die leere Menge ein affiner Unterraum ist. Falls $L(A,b)\neq\emptyset$ ist, existiert (mindestens) eine Lösung des inhomogenen LGS (x^*) und da laut Satz 6.70 $L(A,0_m)$ ein Vektorraum ist, folgt durch Anwendung von Satz 6.71, dass $L(A,b)$ ein affiner Unterraum ist. □

Können wir die Dimension von $L(A,0_m)$ und somit auch die von $L(A,b)$ vorhersagen?

Satz 6.73

Ist $A\in\mathbb{R}^{m\times n}$, so ist $\dim(L(A,0_m))=n-\operatorname{rg}(A)$.

Beweis. Sei $(A|0_m)$ die erweiterte Koeffizientenmatrix des homogenen LGS $A\cdot x=0_m$. Wenden wir den Gauß-Algorithmus an, so wissen wir dank Satz 6.28, dass die elementaren Zeilenumformungen weder die Lösungsmenge noch den Rang von $(A|0_m)$ ändern. Nennen wir $\widetilde{A}$ die so entstandene Matrix und $(\widetilde{A}|0_m)$ die zugehörige erweiterte Koeffizientenmatrix. $\widetilde{A}\in\mathbb{R}^{m\times n}$ hat $\operatorname{rg}(A)=\operatorname{rg}(\widetilde{A})$ Zeilenvektoren, die nicht der Nullvektor sind, und ebenso viele Pivotspalten. Die zu den übrigen $l:=n-\operatorname{rg}(A)$ Spalten gehörenden Variablen sind die freien Variablen — dies seien o.B.d.A. $x_1,\dots,x_l$.[7] Die Lösungsmenge hat dann folgende Form:

$$L(A,0_m)=L(\widetilde{A},0_m)=\left\{ x_1\begin{pmatrix}1\\0\\0\\\vdots\\0\\*\end{pmatrix}+x_2\begin{pmatrix}0\\1\\0\\\vdots\\0\\*\end{pmatrix}+\dots+x_l\begin{pmatrix}0\\0\\\vdots\\0\\1\\*\end{pmatrix}\;\middle|\; x_1,\dots,x_l\in\mathbb{R}\right\},$$

wobei $*$ für die fehlenden $n-l$ Einträge der Vektoren stehen, die an dieser Stelle unwichtig sind. Entscheidend ist, dass $L(A,0_m)$ die lineare Hülle von l linear unabhängigen Vektoren ist, d.h. $\dim(L(A,0_m))=l=n-\operatorname{rg}(A)$. □

6.7 Nützliche Anwendungen des Ranges von Matrizen

Der folgende Satz liefert nicht nur ein nützliches Resultat, sondern wir wollen gleich mehrere Beweismöglichkeiten zeigen, um uns an die bereits bekannten Resultate zu erinnern.

Satz 6.74

Sei $A\in\mathbb{R}^{n\times n}$ eine Matrix und sei $\mathcal{B}$ die Menge, die aus den Spaltenvektoren von A besteht. Dann gilt:

$\mathcal{B}$ ist genau dann eine Basis des $\mathbb{R}^n$, wenn $\operatorname{rg}(A)=n$ ist.

[7] Wenn das nicht die ersten l Variablen sind, kann dies durch Umbenennung der Variablen erreicht werden.

Beweis 1. Der Rang einer Matrix ist gleich dem Spaltenrang, d.h. $rg(A) = n$ genau dann, wenn n Spalten von A linear unabhängig sind und es keine $n+1$ linear unabhängigen Spalten von A gibt.

Sei nun $\mathcal{B}$ eine Basis des $\mathbb{R}^n$. Somit sind alle Vektoren aus $\mathcal{B}$ – also gerade alle n Spaltenvektoren von A – linear unabhängig. Andererseits ist $A \in \mathbb{R}^{n\times n}$ und hat somit höchstens n linear unabhängige Spaltenvektoren. Folglich ist in der Tat $\mathrm{rg}(A) = n$.

Für die Rückrichtung sei $\mathrm{rg}(A) = n$, d.h. alle n Spaltenvektoren von A und somit alle Vektoren von $\mathcal{B}$ sind linear unabhängig. Da $\dim(\mathbb{R}^n) = n$ ist, können wir eine beliebige Basis des $\mathbb{R}^n$ nehmen und alle n Vektoren durch die von $\mathcal{B}$ ersetzen und nach dem Austauschsatz erhalten wir als Ergebnis eine Basis des $\mathbb{R}^n$ – das ist in diesem Fall also $\mathcal{B}$ selbst. □

Beweis 2. Sei $v \in \mathbb{R}^n$ ein beliebiger Vektor.

Sei $\mathcal{B}$ eine Basis des $\mathbb{R}^n$. Nach Satz 6.32 hat v eine eindeutige Darstellung als Linearkombination der Vektoren aus $\mathcal{B}$. Die Koeffizienten sind die Lösungsmenge des LGS $A \cdot x = v$. Da die Koeffizienten eindeutig bestimmt sind, hat das LGS genau eine Lösung. Nach Satz 6.71 enthält $L(A, 0_n)$ nur den Nullvektor, da $L(A, v)$ sonst ein affiner Unterraum von mindestens Dimension 1 wäre. Nach Satz 6.73 ist somit

$$0 = \dim(L(A, 0_n)) = n - \mathrm{rg}(A),$$

d.h. $\mathrm{rg}(A) = n$.

Sei andererseits $\mathrm{rg}(A) = n$. Da $A \in \mathbb{R}^{n\times n}$ ist und der Rang gleich der Anzahl Zeilen und gleich der Anzahl Spalten der Matrix ist, ist auch

$$n = \mathrm{rg}(A) = \mathrm{rg}(A|v),$$

d.h. nach Satz 6.68 existiert eine Lösung von $A \cdot x = v$. Zudem gilt wieder mit Satz 6.73

$$\dim(L(A, 0_n)) = n - \mathrm{rg}(A) = n - n = 0,$$

d.h. $L(A, v)$ enthält entweder genau eine oder keine Lösung – nach vorheriger Überlegung also genau eine. Da $v \in \mathbb{R}^n$ beliebig war, können wir mit Eigenschaft (3) von Satz 6.32 schlussfolgern, dass $\mathcal{B}$ eine Basis des $\mathbb{R}^n$ ist. □

Satz 6.75

Sei $A \in \mathbb{R}^{n\times n}$ eine beliebige quadratische Matrix. A ist genau dann invertierbar, wenn $\mathrm{rg}(A) = n$ ist, d.h. wenn A den maximalen Rang besitzt.

Beweis. $A \in \mathbb{R}^{n\times n}$ ist genau dann invertierbar, wenn die Matrix durch elementare Zeilenumformungen in die n-dimensionale Einheitsmatrix E_n überführt werden kann. Da elementare Zeilenumformungen den Rang einer Matrix nicht ändern, gilt also

$$A \text{ invertierbar} \iff \mathrm{rg}(A) = \mathrm{rg}(E_n) = n.$$

□

6.8 Aufgaben

Aufgabe 6.1. *Es sei* $(V, +, \cdot)$ *ein* $\mathbb{K}$*-Vektorraum und* $u, v, w \in V$*, wobei*

$$u = a \cdot v + b \cdot w \qquad \textit{mit} \qquad a, b \in \mathbb{K}, \quad a \neq 0$$

gelte. Stellen Sie diese Gleichung nach v *um, genauer: Zeigen Sie, dass gilt:*

$$u = a \cdot v + b \cdot w \qquad \Longleftrightarrow \qquad v = \frac{1}{a} \cdot u - \frac{b}{a} \cdot w.$$

Welche der Vektorraumaxiome aus Definition 6.1 haben Sie bei Ihren Umformungen benutzt?

Aufgabe 6.2. *Sei* $(V, +, \cdot)$ *ein* $\mathbb{K}$*-Vektorraum. Ist er ein Untervektorraum (UVR) von sich selbst? Wenn* $e \in V$ *das neutrale Element von* $(V, +)$ *ist, ist dann auch* $(\{e\}, +, \cdot)$ *ein UVR von* $(V, +, \cdot)$*?*

Aufgabe 6.3. *Sei* $(V, +, \cdot)$ *ein* $\mathbb{K}$*-Vektorraum und sei* $U \subseteq V$ *eine nichtleere Teilmenge von* V*. Sei* $e \in V$ *das neutrale Element der Gruppe* $(V, +)$*, d.h. der Nullvektor. Zeigen Sie: Wenn* $e \notin U$ *ist, dann ist* $(U, +, \cdot)$ *kein UVR.*

Aufgabe 6.4. *Sei* $U \subseteq V$ *eine nichtleere Teilmenge von* V *und sei* $(V, +, \cdot)$ *ein* $\mathbb{K}$*-Vektorraum. Zeigen Sie, dass* $(U, +, \cdot)$ *genau dann ein UVR von* $(V, +, \cdot)$ *ist, wenn gilt*

$$a \cdot v + b \cdot w \in U, \quad \forall a, b \in \mathbb{K}, v, w \in U.$$

Aufgabe 6.5. *Betrachten wir den Vektorraum* $(\mathbb{R}^3, +, \cdot)$ *mit der üblichen Addition und der Multiplikation mit Skalaren aus* $\mathbb{R}$ *und mit dem Nullvektor (d.h. Koordinatenursprung)* $e = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Unter Geraden *bzw.* Ebenen *im* $\mathbb{R}^3$ *verstehen wir Teilmengen* $G \subseteq \mathbb{R}^3$ *bzw.* $E \subseteq \mathbb{R}^3$ *der Form*

$$\begin{aligned} G &= \{u + a \cdot v \mid a \in \mathbb{R}\} && \textit{mit } u \in \mathbb{R}^3, v \in \mathbb{R}^3 \setminus \{e\}, \\ E &= \{u + a \cdot v + b \cdot w \mid a, b \in \mathbb{R}\} && \textit{mit } u \in \mathbb{R}^3, v, w \in \mathbb{R}^3 \setminus \{e\}, w \notin \mathcal{L}(\{v\}). \end{aligned}$$

Zeigen Sie: Geraden und Ebenen sind genau dann Untervektorräume des $\mathbb{R}^3$*, wenn sie den Koordinatenursprung enthalten.*

Aufgabe 6.6. *Überprüfen Sie, ob das Polynom* $p(x) = 3x^2 - x + 4$ *in der linearen Hülle von* $M := \{1 - x, 1 + 2x, 1 + x^2\}$ *liegt.*

Aufgabe 6.7. *Geben Sie die linearen Hüllen der folgenden Mengen an.*

a) $M_1 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$*,*

b) $M_2 = \{1, x, x^2, x^3\} \subseteq \mathbb{R}[x]$*,*

c) $M_3 = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 12 \\ 24 \\ 36 \end{pmatrix}, \begin{pmatrix} -1/3 \\ -2/3 \\ -1 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$*.*

Aufgabe 6.8. *Zeigen Sie die Aussage von Lemma 6.19: Ist $M_1 \subseteq V$ linear abhängig und ist $M_2 \subseteq V$ beliebig, so ist auch $M_1 \cup M_2$ linear abhängig.*

Aufgabe 6.9. *Seien $v_1, \ldots, v_n$ linear unabhängige Vektoren in einem Vektorraum $(V, +, \cdot)$. Sei $\lambda \neq 0$. Zeigen Sie, dass auch $\lambda v_1, v_2, \ldots, v_n$ linear unabhängig sind.*

Aufgabe 6.10. *Sei V ein Vektorraum und $M \subseteq V$. Zeigen oder widerlegen Sie folgende Aussagen:*

a) $\mathcal{L}(\mathcal{L}(M)) = \mathcal{L}(M)$.

b) Ist M linear unabhängig, dann ist auch $\mathcal{L}(M)$ linear unabhängig.

Aufgabe 6.11. *Bestimmen Sie den Rang der Matrizen*

$$A := \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix} \quad \textit{und} \quad B := \begin{pmatrix} 1 & 12 & -1/3 \\ 2 & 24 & -2/3 \\ 3 & 36 & -1 \end{pmatrix}.$$

Aufgabe 6.12. *Bestimmen Sie alle $a \in \mathbb{R}$, für die die folgenden drei Vektoren des $\mathbb{R}^3$ linear unabhängig sind.*

$$u = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ t \\ 1 \end{pmatrix}, \quad w = \begin{pmatrix} 1 \\ 2 \\ t^2 \end{pmatrix}.$$

Aufgabe 6.13. *Analog zu Homomorphismen zwischen Gruppen und Körpern können wir auch Homomorphismen zwischen Vektorräumen definieren:*

Definition 6.76: Vektorraumhomomorphismus

Seien $(V, +, \cdot)$ und $(W, \oplus, \odot)$ $\mathbb{K}$-Vektorräume. Eine Abbildung $\varphi\colon V \to W$ heißt Vektorraumhomomorphismus, *wenn für alle $a \in \mathbb{K}$ und $v, w \in V$ gilt:*

$$\varphi(v + w) = \varphi(v) \oplus \varphi(w) \quad \textit{und} \quad \varphi(a \cdot v) = a \odot \varphi(v).$$

φ heißt Isomorphismus, *wenn φ ein bijektiver Homomorphismus ist.*

Seien $V = W = \mathbb{R}^n$ mit der bekannten Addition und skalaren Multiplikation von Vektoren (bzw. Matrizen) und sei $\varphi\colon V \to W$ gegeben durch $\varphi(x) := A \cdot x + b$ für eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$. Unter welchen Bedingungen an A und b ist φ ein Vektorraumhomomorphismus?

Aufgabe 6.14. *Es gilt zudem folgende Definition:*

Definition 6.77: Kern eines Vektorraumhomomorphismus

Sei $\varphi\colon V \to W$ ein Vektorraumhomomorphismus zwischen den Vektorräumen $(V, +, \cdot)$ und $(W, \oplus, \odot)$. Sei e_W das neutrale Element der Gruppe $(W, \oplus)$, d.h. der Nullvektor von W. Der Kern von φ ist definiert als

$$\ker \varphi := \{v \in V \mid \varphi(v) = e_W\}.$$

Betrachten wir den VR-Homomorphismus $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ mit $\varphi(x) = A \cdot x$, wobei $A \in \mathbb{R}^{n\times n}$ invertierbar sei. Bestimmen Sie den Kern von φ.

Aufgabe 6.15. *Gegeben sei die Matrix* $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & -1 \\ 4 & 17 & 13 \end{pmatrix}$.

a) *Bestimmen Sie den Rang der Matrix.*

b) *Bestimmen Sie den Kern des VR-Homomorphismus $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^3$ mit $\varphi(x) = A \cdot x$.*

c) *Bestimmen Sie* $\dim(\ker(\varphi))$.

Aufgabe 6.16. *Seien V und W zwei $\mathbb{K}$-Vektorräume und $\varphi\colon V \to W$ ein Vektorraumhomomorphismus, d.h.*

$$\begin{aligned} \varphi(x+y) &= \varphi(x) + \varphi(y), \quad \forall x, y \in V, \\ \varphi(\lambda \cdot x) &= \lambda \cdot \varphi(x), \quad \forall \lambda \in \mathbb{K}, x \in V. \end{aligned}$$

a) *Zeigen Sie, dass* $\ker(\varphi)$ *ein Untervektorraum von V ist.*

b) *φ sei injektiv. Bestimmen Sie für diesen Fall* $\ker(\varphi)$ *und* $\dim(\ker(\varphi))$.

Aufgabe 6.17. *Sei $M := \{v_1, \ldots, v_n\} \subseteq \mathbb{R}^m$. Sei weiterhin $A \in \mathbb{R}^{m\times n}$ die Matrix, deren Spaltenvektoren genau die Vektoren aus M sind. Zeigen Sie:*

$$\dim(\mathcal{L}(M)) = \mathrm{rg}(A).$$

Aufgabe 6.18. *Seien V ein $\mathbb{K}$-Vektorraum und $M \subseteq V$. Unter welchen Bedingungen ist M eine Basis von $\mathcal{L}(M)$?*

Aufgabe 6.19. *Sei $\mathcal{P}_n$ der schon bekannte Vektorraum der reellen Polynomfunktion vom Grad $\leq n$. Bestimmen Sie* $\dim(\mathcal{P}_n)$.

Aufgabe 6.20. *Gegeben seien*

$$A = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & 1 \\ 1 & 2 & -5 \end{pmatrix} \quad \textit{und} \quad b = \begin{pmatrix} 3\beta + 4 \\ 4\beta - 1 \\ 4\beta + 7 \end{pmatrix} \quad \textit{mit} \quad \beta \in \mathbb{R}.$$

a) *Bestimmen Sie* $\mathrm{rg}(A)$ *und* $\mathrm{rg}(A|b)$ *in Abhängigkeit von $\beta \in \mathbb{R}$. Für welche Werte von β hat das LGS $A \cdot x = b$ (mindestens) eine Lösung?*

b) *Bestimmen Sie die Lösungsmenge von $A \cdot x = b$ in Abhängigkeit von $\beta \in \mathbb{R}$.*

c) *Bestimmen Sie die Lösungsmenge des zugehörigen homogenen LGS $A \cdot x = 0_3$.*

Aufgabe 6.21. *Begründen Sie folgende Aussagen:*

a) *Ist $B = \{b_1, \ldots, b_n\}$ eine Basis und $U = \{u_1, \ldots, u_n\}$ eine linear unabhängige Teilmenge von V, die ebenso viele Elemente enthält wie B, so ist U ebenfalls eine Basis von V.*

b) *Ist V ein n-dimensionaler Vektorraum und E ein ES von V, das aus n Vektoren besteht, so ist E eine Basis von V.*

Aufgabe 6.22. *Betrachten Sie im reellen Vektorraum $\mathbb{R}^3$ die Unterräume*

$$U_1 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 + x_2 + x_3 = 0 \right\} \quad \textit{und} \quad U_2 := \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; \lambda \in \mathbb{R} \right\}.$$

Bestimmen Sie die Dimensionen beider Unterräume und zeigen Sie, dass $U_1 + U_2 = \mathbb{R}^3$ gilt.

Aufgabe 6.23. *Sei $U_1 \subseteq V$ ein UVR eines endlich erzeugten $\mathbb{K}$-Vektorraumes V mit Nullvektor e.*

a) Begründen Sie, dass wenn B_1 eine Basis von U_1 ist, Sie diese zu einer Basis B von V ergänzen können.

b) Folgern Sie hieraus, dass $\dim(U_1) \leq \dim(V)$ ist.

Sei U_1 nun ein echter UVR von V, d.h. $\dim(U_1) < \dim(V)$. Sei $U_2 := \mathcal{L}(B \setminus B_1)$ der UVR, der von diesen zusätzlichen Basisvektoren aus a) erzeugt wird.

c) Zeigen Sie, dass $U_1 \cap U_2 = \{e\}$ ist.

d) Zeigen Sie, dass $U_1 + U_2 = V$ ist.

Kapitel 7

Lineare Abbildungen

Ziele

- □ lineare Abbildungen als VR-Homomorphismen kennen lernen
- □ die Besonderheiten von Kern und Bild bei linearen Abbildungen kennen einschließlich der Dimensionsformel, Beweis dazu
- □ mit Hilfe von Matrizen lineare Abbildungen definieren und deren Eigenschaften an Hand der Matrixeigenschaften erkennen
- □ jeder linearen Abbildung eine Matrix zuordnen können, Existenzbeweis
- □ Transformationsmatrix berechnen können

7.1 Der Vektorraum der linearen Abbildungen

Definition 7.1

Sei $\mathbb{K}$ ein Körper und seien V, W zwei $\mathbb{K}$-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt *linear*, falls

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y), \qquad \forall x, y \in V, \lambda, \mu \in \mathbb{K}$$

gilt, d.h. falls f ein Vektorraum-Homomorphismus ist.

Eine bijektive lineare Abbildung heißt *Isomorphismus*.

Sehen wir uns zwei Beispiele für lineare Abbildungen an, die auch im weiteren Verlauf des Kapitels eine Rolle spielen sollen:

Beispiel 7.2. *Sei $\mathcal{P}_n$ wie bisher der $(n+1)$-dimensionale Vektorraum der reellen Polynomfunktionen, deren Grad höchstens n betragen soll für eine natürliche Zahl $n \in \mathbb{N}$.*

Wir kennen bereits eine Basis von $\mathcal{P}_n$:

$$\mathcal{B} = (1, x, x^2, \ldots, x^n).$$

Ist $p(x) = \sum_{k=0}^{n} a_k x^k \in \mathcal{P}_n$, so sind die Koordinaten von p bezüglich der Basis $\mathcal{B}$ gegeben durch den Vektor $(a_0, a_1, \ldots, a_n)^T \in \mathbb{R}^{n+1}$. Sei $\kappa_{\mathcal{B}}\colon \mathcal{P}_n \to \mathbb{R}^{n+1}$ die in Bemerkung 6.49 *eingeführte Koordinatenabbildung, d.h.*

$$p(x) = \sum_{k=0}^{n} a_k x^k \quad \mapsto \quad \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Diese Abbildung ist nicht nur bijektiv, sondern auch linear, d.h. $\kappa_\mathcal{B}$ *ist ein Isomorphismus.*

Überprüfen wir die Linearität. Seien $p(x) = \sum_{k=0}^{n} a_k x^k$ *und* $q(x) = \sum_{k=0}^{n} b_k x^k$. *Seien weiterhin* $\lambda, \mu \in \mathbb{R}$. *Dann gilt*

$$\begin{aligned}
\kappa_\mathcal{B}(\lambda p + \mu q) &= \kappa_\mathcal{B}\left(\lambda \cdot \sum_{k=0}^{n} a_k x^k + \mu \cdot \sum_{k=0}^{n} b_k x^k\right) \\
&= \kappa_\mathcal{B}\left(\sum_{k=0}^{n} (\lambda a_k + \mu b_k) x^k\right) \\
&= \begin{pmatrix} \lambda a_0 + \mu b_0 \\ \lambda a_1 + \mu b_1 \\ \vdots \\ \lambda a_n + \mu b_n \end{pmatrix} \\
&= \lambda \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} + \mu \cdot \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} \\
&= \lambda \cdot \kappa_\mathcal{B}\left(\sum_{k=0}^{n} a_k x^k\right) + \mu \cdot \kappa_\mathcal{B}\left(\sum_{k=0}^{n} b_k x^k\right) \\
&= \lambda \cdot \kappa_\mathcal{B}(p) + \mu \cdot \kappa_\mathcal{B}(q),
\end{aligned}$$

d.h. $\kappa_\mathcal{B}$ *ist in der Tat ein Isomorphismus.*

Beispiel 7.3. *Sei*

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

Dann ist $f \colon \mathbb{R}^3 \to \mathbb{R}^3$ *mit* $f(x) = A \cdot x$ *eine lineare Abbildung – dies haben wir in Aufgabe 6.13 gesehen.*

Zur Erinnerung: Die Linearität basiert auf dem Distributivgesetz für die Multiplikation von Matrizen – Eigenschaft 3 von Satz 2.19.

Notation

Seien V, W zwei $\mathbb{K}$-Vektorräume. Wir haben folgende Mengen von Abbildungen:

$$\begin{aligned}\mathrm{Abb}(V;W) &:= \{f\colon V \to W\},\\ L(V;W) &:= \{f \in \mathrm{Abb}(V;W) \mid f \text{ ist linear}\},\\ \mathcal{P}_n &:= \left\{p \in \mathrm{Abb}(\mathbb{R};\mathbb{R}) \;\middle|\; p(x) = \sum_{k=0}^{n} a_k x^k\right\}.\end{aligned}$$

Satz 7.4

Seien V und W zwei $\mathbb{K}$-Vektorräume. Dann ist $\mathrm{Abb}(V;W)$ ein $\mathbb{K}$-Vektorraum mit der punktweisen Addition und der Multiplikation mit Skalaren, die wie folgt definiert sind. Für $f, g \in \mathrm{Abb}(V;W)$ und $\lambda \in \mathbb{K}$ definiert man $f+g$ und $\lambda \cdot f$ durch

$$\begin{aligned}(f+g)(x) &:= f(x) + g(x), \quad \forall x \in V,\\ (\lambda \cdot f)(x) &:= \lambda \cdot f(x), \quad \forall x \in V.\end{aligned}$$

Zudem ist $L(V;W)$ ein Untervektorraum von $\mathrm{Abb}(V;W)$.

Beweis. In Aufgabe 1 von Übungsblatt 6 wurde gezeigt, dass $(\mathrm{Abb}(\mathbb{R};\mathbb{R}), +, \cdot)$ ein $\mathbb{R}$-Vektorraum ist. Der Beweis lässt direkt auf den allgemeinen Raum $\mathrm{Abb}(V;W)$ übertragen.

Nach Definition ist $L(V;W) \subseteq \mathrm{Abb}(V;W)$. Wir können also das UVR-Kriterium anwenden.

- $L(V;W) \neq \emptyset$: Sei $f\colon V \to W$ mit $f(x) = e_W$ für alle $x \in V$. Dann gilt für alle $x, y \in V$ und $\lambda, \mu \in \mathbb{K}$

  $$\lambda f(x) + \mu f(y) = \lambda e_W + \mu e_W = e_W = f(\lambda x + \mu y),$$

  d.h. $f \in L(V;W)$. Diese Funktion ist zugleich der Nullvektor in $L(V;W)$.

- Seien $f, g \in L(V;W)$ und $a, b \in \mathbb{K}$. Dann gilt für alle $x, y \in V$ und $\lambda, \mu \in \mathbb{K}$

  $$\begin{aligned}(af+bg)(\lambda x + \mu y) &= (af)(\lambda x + \mu y) + (bg)(\lambda x + \mu y)\\ &= a \cdot f(\lambda x + \mu y) + b \cdot g(\lambda x + \mu y)\\ &= a \cdot (\lambda f(x) + \mu f(y)) + b \cdot (\lambda g(x) + \mu g(y)) \quad \text{da } f,\, g \text{ linear sind}\\ &= a\lambda f(x) + a\mu f(y) + b\lambda g(x) + b\mu g(y)\\ &= \lambda \cdot (af(x) + bg(x)) + \mu \cdot (af(y) + bg(y))\\ &= \lambda \cdot (af+bg)(x) + \mu \cdot (af+bg)(y),\end{aligned}$$

  d.h. $af + bg \in L(V;W)$.

□

7.2 Kern und Bild

Erinnern wir uns an die Definitionen von Kern und Bild:

Definition 7.5

Seien V und W $\mathbb{K}$-Vektorräume mit Nullvektoren e_V und e_W und sei $f \in L(V;W)$.

- Das *Bild* von f ist definiert als
$$\mathrm{Bild}(f) := f(V) = \{w \in W \mid \exists v \in V\colon f(v) = w\}.$$
- Der *Kern* von f ist definiert als
$$\ker(f) := \{v \in V \mid f(v) = e_W\}.$$

Beispiel 7.6 (Fortsetzung von Beispiel 7.2). *Sei $\kappa_\mathcal{B}\colon \mathcal{P}_n \to \mathbb{R}^{n+1}$ die Abbildung, die jedem reellen Polynom vom Grad $\leq n$ seine Koordinaten bezüglich der Basis $\mathcal{B} = (1, x, x^2, \ldots, x^n)$ zuordnet.*

- $\mathrm{Bild}(\kappa_\mathcal{B}) = \kappa_\mathcal{B}(\mathcal{P}_n) = \mathbb{R}^{n+1}$*, denn zu jedem* $a = (a_0, \ldots, a_n)^T \in \mathbb{R}^{n+1}$ *gibt es ein Polynom in* $\mathcal{P}_n$ *mit diesen Koordinaten, nämlich* $p(x) = \sum_{k=0}^n a_k x^k$.
- $\ker(\kappa_\mathcal{B}) = \{0\}$*, d.h. der Kern besteht ausschließlich aus dem Nullpolynom, da gilt:*
$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \iff a_0 = a_1 = \ldots = a_n = 0 \iff p(x) = \sum_{k=0}^n a_k x^k = 0.$$

Beispiel 7.7 (Fortsetzung von Beispiel 7.3). *Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ mit*
$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Wir berechnen wieder Bild und Kern:

- $(a,b,c)^T \in f(\mathbb{R}^3)$*, falls das Gleichungssystem* $A \cdot x = (a,b,c)^T$ *(mindestens) eine Lösung besitzt.*
$$\left[\begin{array}{ccc|c} 1 & 4 & 7 & a \\ 2 & 5 & 8 & b \\ 3 & 6 & 9 & c \end{array}\right] \overset{2Z_2 - Z_3 \to Z_3}{\rightsquigarrow} \left[\begin{array}{ccc|c} 1 & 4 & 7 & a \\ 2 & 5 & 8 & b \\ 1 & 4 & 7 & 2b-c \end{array}\right] \overset{Z_1 - Z_3 \to Z_3}{\rightsquigarrow} \left[\begin{array}{ccc|c} 1 & 4 & 7 & a \\ 2 & 5 & 8 & b \\ 0 & 0 & 0 & a-2b+c \end{array}\right]$$
$$\overset{2Z_1 - Z_2}{\rightsquigarrow} \left[\begin{array}{ccc|c} 1 & 4 & 7 & a \\ 0 & 3 & 6 & 2a-b \\ 0 & 0 & 0 & a-2b+c \end{array}\right]$$
Es ist $\mathrm{rg}(A) = 2 = \mathrm{rg}(A|(a,b,c)^T)$ *genau dann, wenn* $a - 2b + c = 0$ *ist. Somit ist*
$$\mathrm{Bild}(f) = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; a - 2b + c = 0 \right\} = \left\{ b \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + c \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; b, c \in \mathbb{R} \right\}.$$

- $x \in \ker(f)$, *falls* $A \cdot x = (0,0,0)^T$ *gilt. Mit den bereits gemachten Umformungen gilt also*

$$\left[\begin{array}{ccc|c} 1 & 4 & 7 & 0 \\ 2 & 5 & 8 & 0 \\ 3 & 6 & 9 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 4 & 7 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 4 & 7 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

d.h.

$$\begin{aligned} \ker(f) &= \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, a - c = 0, b + 2c = 0 \right\} \\ &= \left\{ c \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \,\middle|\, c \in \mathbb{R} \right\}. \end{aligned}$$

Während im ersten Beispiel das Bild ein $n+1$-dimensionaler und der Kern ein 0-dimensionaler $\mathbb{R}$-Vektorraum sind, ist das Bild im zweiten Beispiel ein 2-dimensionaler und der Kern ein 1-dimensionaler $\mathbb{R}$-Vektorraum.

Das Ziel ist es nun, für beliebige Abbildungen $f \in L(V;W)$ die Dimensionsformel

$$\dim(V) = \dim(\ker(f)) + \dim(\mathrm{Bild}(f))$$

zu zeigen (vorausgesetzt $\dim(V) < \infty$).

Satz 7.8

Für $\mathbb{K}$-Vektorräume V und W sei $f \in L(V;W)$. Dann gilt:

1. $\ker(f)$ und $\mathrm{Bild}(f)$ sind Untervektorräume von V bzw. W.
2. f ist genau dann surjektiv, wenn $\mathrm{Bild}(f) = W$ ist.
3. f ist genau dann injektiv, wenn $\ker(f) = \{e_V\}$ ist.

Beispiel 7.9. *Die Koordinatenabbildung* $\kappa_{\mathcal{B}} \colon \mathcal{P}_n \to \mathbb{R}^{n+1}$ *aus Beispiel 7.2 und 7.6 erfüllt die Bedingungen aus 2. und 3., d.h. es ist, wie wir bereits an anderer Stelle gezeigt hatten, eine bijektive lineare Abbildung, also ein Isomorphismus.*

Die Abbildung f *aus Beispiel 7.3 und 7.7 erfüllt jedoch keine der Bedingungen, sie ist also weder injektiv noch surjektiv.*

Beweis von Satz 7.8.

1. In Aufgabe 6.16 wurde gezeigt, dass der Kern ein UVR von V ist. Es bleibt die Aussage für das Bild zu zeigen. Wir wenden wieder das UVR-Kriterium an:

   - Da $f(e_V) = e_W$ ist[1], ist $e_W \in \mathrm{Bild}(f)$ und somit ist $\mathrm{Bild}(f) \neq \emptyset$.

[1]Erinnerung: In Satz 3.22 wurde gezeigt, dass ein Gruppenhomomorphismus $\varphi\colon (G,*) \to (H,\circ)$ die Eigenschaft $\varphi(e_G) = e_H$ besitzt und auch in Aufgabe 6.16 a haben wir für VR-Homomorphismen nochmals gezeigt, dass der Nullvektor des Ausgangsraumes (das neutrale Element der Addition) auf dem Nullvektor des Zielraumes abgebildet wird.

- Seien $w_1, w_2 \in \mathrm{Bild}(f)$ und $\lambda_1, \lambda_2 \in \mathbb{K}$. Dann existieren $v_1, v_2 \in V$ mit $f(v_1) = w_1$ und $f(v_2) = w_2$. Folglich ist

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 w_1 + \lambda_2 w_2 \in \mathrm{Bild}(f).$$

2. f ist der Definition gemäß genau dann surjektiv, wenn zu jedem $w \in W$ ein $v \in V$ existiert, so dass $f(v) = w$ ist. Das ist (gemäß der Definition des Bildes) genau dann der Fall, wenn $\mathrm{Bild}(f) = W$ ist. Damit ist die Behauptung gezeigt.
3. Sei f injektiv, d.h. aus $f(x) = f(y)$ folge $x = y$. Ist Dann $x \in \ker(f)$, so folgt aus $f(x) = e_W = f(e_V)$ sofort $x = e_V$, d.h. $\ker(f) = \{e_V\}$.

   Angenommen, f sei nicht injektiv, d.h. es gibt zwei Vektoren $x, y \in V$ mit $x \neq y$ und $f(x) = f(y)$. Dann folgt hieraus

$$e_W = f(x) - f(y) \overset{f \text{ linear}}{=} f(x - y),$$

   d.h. $x - y \in \ker(f)$ mit $x - y \neq e_V$. Damit ist die Kontraposition der Rückrichtung bewiesen.

□

Satz 7.10

Seien V und W zwei $\mathbb{K}$-Vektorräume und $f \in L(V;W)$. Es gelte $\dim(V) < \infty$. Sind $(v_1, \ldots, v_m)$ eine Basis von $\ker(f)$ und $(w_1, \ldots, w_n)$ eine Basis von $\mathrm{Bild}(f)$, sowie beliebige Urbildvektoren $u_1 \in f^{-1}(w_1), \ldots, u_n \in f^{-1}(w_n)$ gegeben, so ist

$$B := (u_1, \ldots, u_n, v_1, \ldots, v_m)$$

eine Basis von V. Insbesondere gilt die Dimensionsformel

$$\dim(V) = \dim(\ker(f)) + \dim(\mathrm{Bild}(f)). \tag{7.2.1}$$

Beweis. Zeigen wir zunächst, dass B ein Erzeugendensystem von V ist:

Sei $x \in V$ ein beliebiger Vektor. Dann existieren eindeutig bestimmte Koeffizienten $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ derart, dass

$$\mathrm{Bild}(f) \ni f(x) = \lambda_1 w_1 + \ldots + \lambda_n w_n$$

ist. Setzen wir

$$y := \lambda_1 u_1 + \ldots + \lambda_n u_n, \tag{7.2.2}$$

so ist, da f linear ist und dank der Definition der u_k,

$$f(y) = \lambda_1 f(u_1) + \ldots + \lambda_n f(u_n) = \lambda_1 w_1 + \ldots + \lambda_n w_n = f(x).$$

Damit ist $f(x - y) = f(x) - f(y) = e_W$, d.h. $x - y \in \ker(f)$. Nutzen wir nun die Basis des Kerns, so finden wir Koeffizienten $\mu_1, \ldots, \mu_m \in \mathbb{K}$ derart, dass

$$x - y = \mu_1 v_1 + \ldots + \mu_m v_m \tag{7.2.3}$$

ist. Addieren wir (7.2.3) und (7.2.2), so erhalten wir

$$x = \mu_1 v_1 + \ldots + \mu_m v_m + \lambda_1 u_1 + \ldots + \lambda_n u_n.$$

Da $x \in V$ beliebig war, haben wir $\mathcal{L}(B) = V$ gezeigt.

Es bleibt noch die lineare Unabhängigkeit von B zu zeigen. Ist

$$\mu_1 v_1 + \ldots + \mu_m v_m + \lambda_1 u_1 + \ldots + \lambda_n u_n = e_V, \tag{7.2.4}$$

so folgt, indem wir die lineare Abbildung f auf diese Gleichung anwenden,

$$\begin{aligned} & f(\mu_1 v_1 + \ldots + \mu_m v_m + \lambda_1 u_1 + \ldots + \lambda_n u_n) \\ &= \mu_1 f(v_1) + \ldots + \mu_m f(v_m) + \lambda_1 f(u_1) + \ldots + \lambda_n f(u_n) \\ &= \mu_1 \cdot e_W + \ldots + \mu_m \cdot e_W + \lambda_1 w_1 + \ldots + \lambda_n w_n \\ &= \lambda_1 w_1 + \ldots + \lambda_n w_n \\ &\overset{!}{=} f(e_V) = e_W. \end{aligned}$$

Da $(w_1, \ldots, w_n)$ eine Basis des Bildes und somit insbesondere linear unabhängig ist, gilt diese Gleichung also nur dann, wenn $\lambda_1 = \ldots = \lambda_n = 0$ ist. Setzen wir dies in (7.2.4) ein, so erhalten wir

$$\mu_1 v_1 + \ldots + \mu_m v_m = e_V.$$

Da $(v_1, \ldots, v_m)$ eine Basis des Kerns und somit ebenfalls linear unabhängig ist, folgt schließlich auch $\mu_1 = \ldots = \mu_m = 0$, was den Beweis vervollständigt, dass B eine linear unabhängige Teilmenge von V und, da es bereits als Erzeugendensystem erkannt wurde, eine Basis von V ist. Also ist $\dim(V) = m + n$. Da zudem $\dim(\ker(f)) = m$ und $\dim(\mathrm{Bild}(f)) = n$ ist, folgt die Dimensionsformel. □

Korollar 7.11

Seien V und W zwei $\mathbb{K}$-Vektorräume, wobei $\dim(V) < \infty$ sei. Falls ein Isomorphismus $f\colon V \to W$ existiert, dann gilt $\dim(V) = \dim(W)$.

Beispiel 7.12. *Da $\kappa_{\mathcal{B}}\colon \mathcal{P}_n \to \mathbb{R}^{n+1}$ aus Beispiel 7.2 und 7.6 ein Isomorphismus ist, ist* $\dim(\mathcal{P}_n) = \dim(\mathbb{R}^{n+1}) = n + 1$.

Beweis von Korollar 7.11.
Falls f ein Isomorphismus ist, d.h. eine bijektive lineare Abbildung, so folgt aus Satz 7.8

$$\dim(\ker(f)) + \dim(\mathrm{Bild}(f)) = \dim(\{e_V\}) + \dim(W) = 0 + \dim(W) = \dim(W).$$

Mit der Dimensionsformel (7.2.1) gilt zudem

$$\dim(\ker(f)) + \dim(\mathrm{Bild}(f)) = \dim(V).$$

Beide Aussagen zusammen ergeben die Behauptung. □

7.3 Lineare Abbildungen und Matrizen

Satz 7.13

Seien V und W endlich-dimensionale $\mathbb{K}$-Vektorräume. Seien zudem $(v_1, \dots, v_r) \subseteq V$ und $(w_1, \dots, w_r) \subseteq W$.

1. Sind die Vektoren $v_1, \dots, v_r$ linear unabhängig, so gibt es mindestens eine lineare Abbildung $f \in L(V; W)$ mit
$$f(v_k) = w_k, \quad \forall k \in \{1, \dots, r\}.$$
2. Ist $(v_1, \dots, v_r)$ eine Basis von V, so gibt es genau eine lineare Abbildung $f \in L(V; W)$ mit
$$f(v_k) = w_k, \quad \forall k \in \{1, \dots, r\}. \tag{7.3.1}$$
Es gilt zudem:
   (a) $\mathrm{Bild}(f) = \mathcal{L}(\{w_1, \dots, w_r\})$;
   (b) f ist genau dann injektiv, wenn $w_1, \dots, w_r$ linear unabhängig sind.

Beweis. Es ist in diesem Fall günstig, zunächst Teil 2 zu beweisen und damit dann Teil 1 zu zeigen.

2. Jeder Vektor $v \in V$ hat eine eindeutige Darstellung als Linearkombination der Basisvektoren, d.h. $v = \sum_{k=1}^{r} \alpha_k v_k$. Ist f linear, so ist der Funktionswert $f(v)$ eindeutig festgelegt als
$$f(v) = f\left(\sum_{k=1}^{r} \alpha_k v_k\right) = \sum_{k=1}^{r} \alpha_k f(v_k) = \sum_{k=1}^{r} \alpha_k w_k. \tag{7.3.2}$$
Folglich kann es höchstens eine lineare Abbildung mit der gewünschten Eigenschaft geben. Zeigen wir nun, dass durch (7.3.2) eine lineare Abbildung definiert wird, die die gewünschte Eigenschaft hat. Dann ist durch explizite Konstruktion die Existenz gezeigt.

Linearität: Seien $v, w \in V$ mit $v = \sum_{k=1}^{r} \beta_k v_k$ und $w = \sum_{k=1}^{r} \gamma_k v_k$ und seien $\lambda, \mu \in \mathbb{K}$. Dann ist
$$\begin{aligned} f(\lambda v + \mu w) &= f\left(\lambda \sum_{k=1}^{r} \beta_k v_k + \mu \sum_{k=1}^{r} \gamma_k v_k\right) = f\left(\sum_{k=1}^{r} (\lambda\beta_k + \mu\gamma_k) v_k\right) \\ &\overset{(7.3.2)}{=} \sum_{k=1}^{r} (\lambda\beta_k + \mu\gamma_k) w_k \\ &= \lambda \sum_{k=1}^{r} \beta_k w_k + \mu \sum_{k=1}^{r} \gamma_k w_k \\ &= \lambda f(v) + \mu f(w). \end{aligned}$$

(7.3.1) **ist erfüllt:** Da $v_k = \sum_{j=1}^r \delta_{k,j} v_j$ ist mit

$$\delta_{k,j} = \begin{cases} 1, & \text{falls } k = j \text{ ist,} \\ 0, & \text{sonst,} \end{cases}$$

gilt wie gefordert (durch Vertauschen der endlichen Summen)

$$f(v_k) \overset{f \text{ linear}}{=} \sum_{j=1}^r \delta_{k,j} f(v_j) \overset{\text{Def. } f}{=} \sum_{j=1}^r \delta_{k,j} w_j = w_k, \qquad \forall k \in \{1, \ldots, r\}.$$

Zeigen wir nun die beiden zusätzlichen Aussagen:

(a) Wir zeigen wieder, dass jede Menge in der anderen enthalten ist.

$\underline{\mathrm{Bild}(f) \subseteq \mathcal{L}(\{w_1, \ldots, w_r\})}$**:** Sei $w \in \mathrm{Bild}(f)$, d.h. zu w existiert ein $v \in V$ mit $f(v) = w$. Dieser Vektor v hat eine eindeutige Darstellung als Linearkombination der Basiselemente, d.h. $v = \sum_{k=1}^r \alpha_k v_k$. Dann gilt (unter Verwendung der Eigenschaften von f)

$$w = f(v) = f\left(\sum_{k=1}^r \alpha_k v_k\right) = \sum_{k=1}^r \alpha_k f(v_k) = \sum_{k=1}^r \alpha_k w_k \in \mathcal{L}(\{w_1, \ldots, w_r\}).$$

Da $w \in \mathrm{Bild}(f)$ beliebig war, folgt somit $\mathrm{Bild}(f) \subseteq \mathcal{L}(\{w_1, \ldots, w_r\})$.

$\underline{\mathcal{L}(\{w_1, \ldots, w_r\}) \subseteq \mathrm{Bild}(f)}$**:** Nach Eigenschaft (7.3.1) ist $\{w_1, \ldots, w_r\} \subseteq \mathrm{Bild}(f)$. Da $\mathrm{Bild}(f)$ ein Vektorraum ist und $\mathcal{L}(\{w_1, \ldots, w_r\})$ der kleinste VR, der diese Vektoren enthält, folgt die Behauptung.

(b) Nehmen wir an, $\{w_1, \ldots, w_r\}$ sei linear abhängig. Dann existieren $\alpha_1, \ldots, \alpha_r \in \mathbb{K}$, nicht alle Null, so dass $\sum_{k=1}^r \alpha_k w_k = e_W$ ist. Folglich ist

$$f(\sum_{k=1}^r \alpha_k v_k) = \sum_{k=1}^r \alpha_k f(v) = \sum_{k=1}^r \alpha_k w_k = e_W,$$

d.h. $\sum_{k=1}^r \alpha_k v_k \in \ker(f)$, obwohl dies nicht der Nullvektor in V ist, da $\{v_1, \ldots, v_r\}$ eine Basis von V ist. Nach Satz 7.8 ist f also nicht injektiv.
Sei andererseits $\{w_1, \ldots, w_r\}$ linear unabhängig. Für ein $v \in V$ gelte $f(v) = e_W$. Da v die Darstellung $v = \sum_{k=1}^r \alpha_k v_k$ besitzt, folgt sofort aufgrund der Linearität von f

$$e_W = f(v) = f\left(\sum_{k=1}^r \alpha_k v_k\right) = \sum_{k=1}^r \alpha_k f(v_k) = \sum_{k=1}^r \alpha_k w_k.$$

Mit der linearen Unabhängigkeit von $\{w_1, \ldots, w_r\}$ folgt, dass $\alpha_1 = \ldots = \alpha_r = 0$ ist, d.h. $v = e_V$. Somit ist $\ker(f) = \{e_V\}$. Wieder schlussfolgern wir mit Satz 7.8, dass f in diesem Fall injektiv sein muss.

1. Ist $\{v_1, \ldots, v_r\}$ nur eine linear unabhängige Teilmenge von V, so können wir sie mit Hilfe des Austauschsatzes zu einer Basis $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ von V ergänzen. Wählen wir Vektoren $w_{r+1}, \ldots, w_n \in W$ beliebig, so gibt es nach Teil 2 genau eine lineare Abbildung mit $f(v_k) = w_k$ für $k \in \{1, \ldots, n\}$. Da die Vektoren $w_{r+1}, \ldots, w_n \in W$ beliebig gewählt wurden, kann ohne deren Festlegung keine Eindeutigkeit der linearen Abbildung erzielt werden.

□

Erinnern wir uns daran, dass für $m, n \in \mathbb{N}$ und eine Matrix $A \in \mathbb{K}^{m \times n}$ die Abbildung $f\colon \mathbb{K}^n \to \mathbb{K}^m$ mit $f_A(x) := A \cdot x$ stets linear ist. Der folgende Satz gibt uns die Umkehrung:

Korollar 7.14

Zu jeder linearen Abbildung $f \in L(\mathbb{K}^n; \mathbb{K}^m)$ existiert genau eine Matrix $A \in \mathbb{K}^{m \times n}$, so dass gilt:

$$f(x) = A \cdot x, \quad \forall x \in \mathbb{K}^n.$$

Beweis. Wählen wir die Standardbasis des $\mathbb{K}^n$, $\mathcal{B} := (e_1, \ldots, e_n)$ mit den Einheitsvektoren e_k. Für $k \in \{1, \ldots, n\}$ sei $w_k := f(e_k)$. Nach Satz 7.13 gibt es genau eine lineare Abbildung mit der Eigenschaft $f(e_k) = w_k$ für alle $k \in \{1, \ldots, n\}$. Finden wir also eine Matrix A, für die gilt, dass $A \cdot e_k = w_k$ ist für $k \in \{1, \ldots, n\}$, so ist automatisch $f(x) = A \cdot x$ für alle $x \in \mathbb{K}^n$.

Sei A nun die Matrix, die entsteht, indem die Spaltenvektoren $f(e_1), \ldots, f(e_n) \in \mathbb{K}^m$ aneinandergereiht werden. Nach Konstruktion ist $A \in \mathbb{K}^{m \times n}$. Zudem gilt für $k \in \{1, \ldots, n\}$:

$$A \cdot e_k = \begin{pmatrix} | & | & & | & & | \\ f(e_1) & f(e_2) & \cdots & f(e_k) & \cdots & f(e_n) \\ | & | & & | & & | \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = f(e_k).$$

□

Dieses Korollar lässt sich verallgemeinern für $f \in L(V; W)$ statt $f \in L(\mathbb{K}^n; \mathbb{K}^m)$.

Satz 7.15

Gegeben seien zwei $\mathbb{K}$-Vektorräume V und W mit Basen $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_m)$. Sei $f\colon V \to W$ eine lineare Abbildung. Dann gibt es eine eindeutig bestimmte Matrix $A_f^{\mathcal{B},\mathcal{C}} = (\alpha_{k,j}) \in \mathbb{K}^{m \times n}$ so dass

$$f(v_j) = \sum_{k=1}^{m} \alpha_{k,j} w_k, \quad \forall j \in \{1, \ldots, n\}, \tag{7.3.3}$$

gilt. Die Abbildung $F\colon L(V; W) \to \mathbb{K}^{m \times n}$ mit $F(f) = A_f^{\mathcal{B},\mathcal{C}}$ ist ein Isomorphismus zwischen zwei $\mathbb{K}$-Vektorräumen.

Bemerkung 7.16. *In diesem Satz ist es nicht sinnvoll, $f(v_j) = w_k$ zu fordern, da die Indizes möglicherweise über verschiedenen Teilmengen von $\mathbb{N}$ laufen.*

Notation

- Man bezeichnet $A_f^{\mathcal{B},\mathcal{C}}$ als die darstellende Matrix von f bezüglich $\mathcal{B}$ und $\mathcal{C}$.
- Wenn $V = W$ ist und beide Basen identisch sind, so schreibt man statt $A_f^{\mathcal{B},\mathcal{B}}$ auch kürzer $A_f^{\mathcal{B}}$ für die darstellende Matrix von f bezüglich $\mathcal{B}$.
- Sind die Basen kanonisch gewählt (also die Standardbasen, sofern solche existieren) oder für einen längeren Abschnitt festgehalten, so kann man verkürzt auch A_f statt $A_f^{\mathcal{B},\mathcal{C}}$ schreiben.

Beweis. Da $\mathcal{C}$ eine Basis von W ist, hat jeder Vektor aus W eine eindeutige Darstellung als Linearkombination der Vektoren $\{w_1, \ldots, w_m\}$, also auch $f(v_1), \ldots, f(v_n) \in W$.

Zeigen wir nun, dass F ein Isomorphismus ist.

F **ist linear:** Seien $f, g \in L(V;W)$ mit zugehörigen Matrizen $A_f = (\alpha_{k,j}), A_g = (\beta_{k,j}) \in \mathbb{K}^{m\times n}$. Seien zudem $\lambda, \mu \in \mathbb{K}$. Dann gilt

$$(\lambda f+\mu g)(v_j) = \lambda f(v_j)+\mu g(v_j) = \lambda \sum_{k=1}^{m} \alpha_{k,j} w_k + \mu \sum_{k=1}^{m} \beta_{k,j} w_k = \sum_{k=1}^{m} (\lambda\alpha_{k,j}+\mu\beta_{k,j}) w_k,$$

d.h. $F(\lambda f + \mu g) = \lambda A_f + \mu A_g = \lambda F(f) + \mu F(g)$.

F **ist bijektiv:** Da $\mathcal{B}$ eine Basis von V ist, gibt es nach Teil 2 von Satz 7.13 zu jeder Matrix $A \in \mathbb{K}^{m\times n}$ genau eine lineare Abbildung $f_A \in L(V;W)$, die (7.3.3) erfüllt. Damit ergibt sich eine Umkehrabbildung zu F. Folglich ist F bijektiv.

□

7.4 Der Isomorphismus zwischen dem Raum der linearen Abbildungen und den zugehörigen Matrizen

Die Inhalte dieses Abschnitts wollen wir uns anhand zweier Beispiele verdeutlichen.

Beispiel 7.17. *Seien $f\colon \mathbb{R}^4 \to \mathbb{R}^3$ und $g\colon \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch*

$$f\begin{pmatrix} a\\ b\\ c\\ d \end{pmatrix} = \begin{pmatrix} a+2b+3c+4d\\ b+2c+3d\\ a+3b+5c+7d \end{pmatrix} \quad \text{bzw.} \quad g\begin{pmatrix} x\\ y\\ z \end{pmatrix} = \begin{pmatrix} y\\ z\\ x \end{pmatrix}.$$

Bezüglich der Standardbasen des $\mathbb{R}^3$ und $\mathbb{R}^4$ sind die zu f bzw. g gehörenden darstellenden Matrizen

$$A_f = \begin{pmatrix} 1 & 2 & 3 & 4\\ 0 & 1 & 2 & 3\\ 1 & 3 & 5 & 7 \end{pmatrix} \quad \text{bzw.} \quad A_g = \begin{pmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0 \end{pmatrix}.$$

Untersuchen wir die linearen Abbildung f und g auf Injektivität, Surjektivität und damit Bijektivität.

- *Wäre f eine lineare bijektive Abbildung, also ein Isomorphismus, dann müssten nach Korollar 7.11 Ausgangs- und Zielraum gleiche Dimension haben. In unserem Fall ist jedoch* $\dim(\mathbb{R}^4) = 4 \neq 3 = \dim(\mathbb{R}^3)$*, d.h. f kann kein Isomorphismus sein. Nach Aufgabe 7.3 ist* $\mathrm{Bild}(f) = \mathcal{L}(\mathcal{A}_f)$*, wenn $\mathcal{A}_f$ die Menge der Spaltenvektoren von A_f bezeichnet. Somit gilt* $\dim(\mathrm{Bild}(f)) = \mathrm{rg}(A_f)$*. Da die ersten zwei Zeilen von A_f linear unabhängig sind, deren Summe jedoch genau die dritte Zeile ergibt, ist* $\mathrm{rg}(A_f) = 2$ *und mit der Dimensionsformel folgt*

$$\dim(\ker(f)) = \dim(\mathbb{R}^4) - \dim(\mathrm{Bild}(f)) = 4 - \mathrm{rg}(A_f) = 2.$$

  Da weder $\dim(\mathrm{Bild}(f)) = \dim(\mathbb{R}^3)$ *noch* $\dim(\ker(f)) = \dim\{(0,0,0,0)^T\} = 0$ *ist, ist f weder surjektiv noch injektiv.*

- *Nach Aufgabe 7.6 ist g entweder bijektiv oder weder injektiv noch surjektiv. Da alle Zeilen/Spalten von A_g linear unabhängig sind, ist* $\dim(\mathrm{Bild}(g)) = \mathrm{rg}(A_g) = 3 = \dim(\mathbb{R}^3)$*. Da zudem nach Definition des Bildes* $\mathrm{Bild}(g) \subseteq \mathbb{R}^3$ *ist, folgt hieraus* $\mathrm{Bild}(g) = \mathbb{R}^3$*, d.h. g ist surjektiv, also sogar bijektiv. Insbesondere enthält der Kern von g nur den Nullvektor des $\mathbb{R}^3$ und g ist invertierbar. Letzteres gilt auch für die Matrix A_g:*

$$A_g^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

  Es ist sogar $A_{g^{-1}} = A_g^{-1}$, denn für die zu A_g^{-1} gehörende lineare Abbildung gilt:

$$h\begin{pmatrix} x \\ y \\ z \end{pmatrix} := \begin{pmatrix} z \\ x \\ y \end{pmatrix} \quad \textit{erfüllt} \quad (h \circ g)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = h\left(g\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = h\begin{pmatrix} y \\ z \\ x \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

  d.h. $h \circ g = \mathrm{id}_{\mathbb{R}^3}$ (und ebenso $g \circ h = \mathrm{id}_{\mathbb{R}^3}$).

Satz 7.18

Seien V und W zwei $\mathbb{K}$-Vektorräume mit Basen $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_m)$. Sei zudem $f \in L(V; W)$ mit darstellender Matrix $A_f = A_f^{\mathcal{B},\mathcal{C}} \in \mathbb{K}^{m \times n}$. Dann gilt:

1. $\dim(\mathrm{Bild}(f)) = \mathrm{rg}(A_f)$ und somit ist $\dim(\ker(f)) = \dim(V) - \mathrm{rg}(A_f)$;
2. f ist genau dann surjektiv, wenn $\mathrm{rg}(A_f) = \dim(W)$ ist und f ist genau dann injektiv, wenn $\mathrm{rg}(A_f) = \dim(V)$ ist;
3. f ist genau dann bijektiv, wenn $\mathrm{rg}(A_f) = \dim(V) = \dim(W)$ ist. In diesem Fall gilt $A_f^{-1} = A_{f^{-1}}$. Insbesondere ist somit die Inverse einer linearen Abbildung, falls sie existiert, selbst linear.

Bemerkung 7.19.

i) Aussage 1 impliziert, dass der Rang der darstellenden Matrix einer linearen Abbildung nicht von der Wahl der Basen abhängig ist.

ii) Aussage 2 kann man sich auch folgendermaßen merken: f ist genau dann surjektiv, wenn A_f den maximalen Zeilenrang hat und f ist genau dann injektiv, wenn A_f den maximalen Spaltenrang hat.

iii) Für den Spezialfall $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ wird in Aufgabe 7.3a) $\mathrm{Bild}(f) = \mathcal{L}(\mathcal{A}_f)$ *gezeigt, wobei $\mathcal{A}_f$ die Menge der Spaltenvektoren von A_f ist. Da* $\dim(\mathcal{L}(\mathcal{A}_f))$ *die Anzahl linear unabhängiger Spaltenvektoren aus $\mathcal{A}_f$ ist, was per Definition gerade* $\mathrm{rg}_S(A_f) = \mathrm{rg}(A_f)$ *ist, folgt die Aussage zu* $\dim(\mathrm{Bild}(f))$*. Für diesen Spezialfall ist zudem* $\ker(f) = L(A_f, 0_m)$*, so dass man die Formel für die Dimension des Kerns wahlweise über die Dimensionsformel* (7.2.1) *oder über* $\dim(L(A_f, 0_m))$ *aus Satz 6.73 herleiten kann.*

Bevor wir den Satz beweisen können, benötigen wir ein kleines Hilfsresultat:

Lemma 7.20

Sei $f \in L(V; W)$ für zwei $\mathbb{K}$-Vektorräume V und W. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V. Dann gilt $\mathrm{Bild}(f) = \mathcal{L}(f(\mathcal{B}))$.

Beweis. Sei einerseits $w \in \mathrm{Bild}(f)$, d.h. es existiert $v \in V$ mit $f(v) = w$. Mit der Darstellung $v = \sum_{k=1}^n \alpha_k v_k$ und der Linearität von f folgt

$$w = f(v) = f\left(\sum_{k=1}^n \alpha_k v_k\right) = \sum_{k=1}^n \alpha_k f(v_k) \in \mathcal{L}(f(\mathcal{B})).$$

Sei nun andererseits $w' \in \mathcal{L}(f(\mathcal{B}))$, d.h. $w' = \sum_{k=1}^n \beta_k f(v_k)$. Da f linear ist, folgt sofort

$$w' = \sum_{k=1}^n \beta_k f(v_k) = f\left(\sum_{k=1}^n \beta_k v_k\right) \in \mathrm{Bild}(f).$$

$\square$

Beweis von Satz 7.18.

1. Nach Lemma 7.20 ist $\mathcal{L}(f(\mathcal{B})) = \mathrm{Bild}(f)$. Es bleibt zu zeigen, dass $f(\mathcal{B})$ aus genau so vielen linear unabhängigen Vektoren besteht wie A_f linear unabhängige Spalten besitzt.

   $\mathcal{A}_f = \{A_1, \dots, A_n\} \subseteq \mathbb{K}^m$ sei die Menge der Spaltenvektoren von A_f. Schreiben wir analog zu Satz 7.15 $A_f = (\alpha_{k,j})_{k\in\{1,\dots,m\}, j\in\{1,\dots,n\}}$, so gilt

$$\sum_{j=1}^r \lambda_j A_j = 0_m \in \mathbb{K}^m \quad \Longleftrightarrow \quad \sum_{j=1}^r \lambda_j \alpha_{k,j} = 0,\ \forall k \in \{1, \dots, m\}. \tag{7.4.1}$$

   Sei $J \subseteq \{1, \dots, n\}$ derart, dass $\{A_j \mid j \in J\}$ linear abhängig ist. Dann existieren $\{\lambda_j \mid j \in J\} \subseteq \mathbb{K}$ (nicht alle 0), so dass $\sum_{j\in J} \lambda_j A_j = 0_m \in \mathbb{K}^m$ ist. Daraus folgt

$$\sum_{j\in J} \lambda_j f(v_j) \overset{(7.3.3)}{=} \sum_{j\in J} \lambda_j \sum_{k=1}^m \alpha_{k,j} w_k = \sum_{k=1}^m \left(\sum_{j\in J} \lambda_j \alpha_{k,j}\right) w_k \overset{(7.4.1)}{=} \sum_{k=1}^m 0 \cdot w_k = e_W,$$

   d.h. $\{f(v_j) \mid j \in J\}$ ist linear abhängig.

Sei nun $I \subseteq \{1, \dots, n\}$ so gewählt, dass $\{ f(v_i) \mid i \in I\}$ linear abhängig ist. Dann existiert eine Menge $\{ \lambda_i \mid i \in I\} \subseteq \mathbb{K}$ (nicht alle 0) derart, dass $\sum_{i\in I} \lambda_i f(v_i) = e_W$ ist. Daraus folgt

$$e_W = \sum_{i\in I} \lambda_i f(v_i) \overset{(7.3.3)}{=} \sum_{i\in I} \lambda_i \sum_{k=1}^{m} \alpha_{k,i} w_k = \sum_{k=1}^{m} \left(\sum_{i\in I} \lambda_i \alpha_{k,i} \right) w_k.$$

Da $\mathcal{C} = (w_1, \dots, w_m)$ eine Basis von W ist, ist dies eine Linearkombination linear unabhängiger Vektoren, die den Nullvektor ergibt. Folglich müssen alle Koeffizienten Null sein, d.h.

$$\sum_{i\in I} \lambda_i \alpha_{k,i} = 0, \quad \forall k \in \{1, \dots, m\}.$$

Mit (7.4.1) folgt hieraus, dass die Linearkombination der Spaltenvektoren $\sum_{i\in I} \lambda_i A_i$ den Nullvektor ergibt, obwohl nicht alle Koeffizienten Null sind. Folglich ist die Menge $\{ A_i \mid i \in I\}$ linear abhängig.

Wir haben also gezeigt: Beliebig gewählte Spaltenvektoren von A_f sind genau dann linear abhängig, wenn die zugehörigen Vektoren $f(v_j)$ linear abhängig sind. Die gleiche Äquivalenz gilt (per Kontraposition) für die lineare Unabhängigkeit, so dass hieraus folgt, dass $\mathcal{A}_f$ und $f(\mathcal{B})$ die gleiche Anzahl linear unabhängiger Vektoren enthält und somit $\mathrm{rg}(A_f) = \dim(\mathcal{L}(f(\mathcal{B}))) = \dim(\mathrm{Bild}(f))$ folgt.

Die Aussage über die Dimension des Kerns folgt aus der Dimensionsformel (7.2.1).

2. f ist genau dann surjektiv, wenn $\mathrm{Bild}(f) = W$ ist. Da $\mathrm{Bild}(f)$ ein UVR von W ist, ist $\mathrm{Bild}(f) = W$ äquivalent zu $\dim(\mathrm{Bild}(f)) = \dim(W)$, was nach 1. wiederum äquivalent ist zu $\mathrm{rg}(A_f) = \dim(W)$.

   Andererseits ist f genau dann injektiv, wenn $\ker(f) = \{e_V\}$ ist, was mit $\ker(f) \subseteq V$, der Dimensionsformel und 1. äquivalent ist zu $\dim(\ker(f)) = 0 = \dim(V) - \dim(\mathrm{Bild}(f)) = \dim(V) - \mathrm{rg}(A_f)$.

3. f ist genau dann bijektiv, wenn es surjektiv und injektiv ist und nach 2. ist das genau dann der Fall, wenn $\mathrm{rg}(A_f) = \dim(W)$ und $\mathrm{rg}(A_f) = \dim(V)$ ist. Kommen wir nun zur Inversen.

   - Sei f ein Isomorphismus. Wir wollen zunächst zeigen, dass auch die Inverse f^{-1} linear ist. Seien dazu $w, w' \in W$ mit $f^{-1}(w) =: v$ und $f^{-1}(w') =: v'$ und seien $\lambda, \mu \in \mathbb{K}$. Dann ist

$$\begin{aligned} f^{-1}(\lambda w + \mu w') &= f^{-1}(\lambda f(v) + \mu f(v')) \\ &\overset{f \text{ linear}}{=} f^{-1}(f(\lambda v + \mu v')) \\ &= \lambda v + \mu v' \\ &= \lambda f^{-1}(w) + \mu f^{-1}(w'). \end{aligned}$$

   - Betrachten wir nun die zu f^{-1} zugehörige Matrix. Falls wir eine Matrix A' finden, welche $f^{-1}(\mathcal{C}) = A' \cdot \mathcal{B}$ erfüllt, so ist $A' = A_{f^{-1}}$ nach (7.3.3).
   - A_f erfüllt die Bedingung $f(\mathcal{B}) = A_f \cdot \mathcal{C}$. Multiplikation mit der inversen Matrix A_f^{-1} ergibt $A_f^{-1} \cdot f(\mathcal{B}) = \mathcal{C}$. Da f linear ist, folgt hieraus $f(A_f^{-1} \cdot \mathcal{B}) = \mathcal{C}$. Anwendung von f^{-1} auf beiden Seiten ergibt schließlich $A_f^{-1} \cdot \mathcal{B} = f^{-1}(\mathcal{C})$, woraus mit dem ersten Punkt folgt, dass $A_f^{-1} = A_{f^{-1}}$ ist.

□

Definition 7.21

Seien V und W zwei $\mathbb{K}$-Vektorräume und sei $f \in L(V;W)$. Dann definieren wir den *Rang der linearen Abbildung* f als die Dimension des Bildes von f, d.h.

$$\mathrm{rg}(f) := \dim(\mathrm{Bild}(f)).$$

Bemerkung 7.22. *Zu fest bestimmten Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W haben wir im vorherigen Satz gesehen, dass $\mathrm{rg}(A_f^{\mathcal{B},\mathcal{C}}) = \dim(\mathrm{Bild}(f))$ ist (unabhängig von der Wahl der Basen), was erklärt, warum für die Dimension des Bildes ebenfalls die Bezeichnung* Rang *gewählt wurde.*

Beispiel 7.23. *Die Verknüpfung $g \circ f$ der Abbildungen aus Beispiel 7.17 ist nach Aufgabe 1 des 11. Übungszettels ebenfalls linear. Ihre Abbildungsvorschrift lautet*

$$(g \circ f)\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} b + 2c + 3d \\ a + 3b + 5c + 7d \\ a + 2b + 3c + 4d \end{pmatrix}.$$

Ihre darstellende Matrix bezüglich der Standardbasen lautet

$$A_{g \circ f} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 5 & 7 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

Andererseits ist

$$A_g \cdot A_f = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 5 & 7 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 5 & 7 \\ 1 & 2 & 3 & 4 \end{pmatrix} = A_{g \circ f}.$$

Es spielt also keine Rolle ob zuerst zwei lineare Abbildungen verknüpft und dann die zugehörige Matrix aufgestellt wird oder ob zuerst die Matrizen aufgestellt werden und diese anschließend multipliziert werden.

Satz 7.24

Seien U, V, W drei $\mathbb{K}$-Vektorräume mit Basen $\mathcal{B}_U = (u_1, \ldots, u_{n_1})$, $\mathcal{B}_V = (v_1, \ldots, v_{n_2})$ und $\mathcal{B}_W = (w_1, \ldots, w_{n_3})$. Seien $f \in L(U;V)$ und $g \in L(V;W)$ zwei lineare Abbildungen mit darstellenden Matrizen $A_f = A_f^{\mathcal{B}_U,\mathcal{B}_V} \in \mathbb{K}^{n_2 \times n_1}$ und $A_g = A_g^{\mathcal{B}_V,\mathcal{B}_W} \in \mathbb{K}^{n_3 \times n_2}$. Dann gilt:

1. $g \circ f \in L(U;W)$ und $A_{g \circ f} = A_g \cdot A_f \in \mathbb{K}^{n_3 \times n_1}$.
2. Ist $n_1 = n_2 = n_3 =: n$ und sind f und g bijektiv, so ist auch $g \circ f$ bijektiv und es gilt $A_{(g \circ f)^{-1}} = (A_{g \circ f})^{-1} = A_f^{-1} \cdot A_g^{-1} \in \mathbb{K}^{n \times n}$.

Den im Satz beschriebenen Zusammenhang zwischen linearen Abbildungen und den zugehörigen Matrizen kann man mit folgendem Diagramm verdeutlichen:

$$\begin{tikzcd}
U \arrow[rr, "g \circ f"] \arrow[dr, "f"] \arrow[ddd, "\kappa_{\mathcal{B}_U}"'] & & W \arrow[ddd, "\kappa_{\mathcal{B}_W}"] \\
& V \arrow[ur, "g"] \arrow[d, "\kappa_{\mathcal{B}_V}"'] & \\
& \mathbb{K}^{\dim(V)} \arrow[dr, "A_g"] & \\
\mathbb{K}^{\dim(U)} \arrow[ur, "A_f"] \arrow[rr, "A_g \cdot A_f"'] & & \mathbb{K}^{\dim(W)}
\end{tikzcd}$$

Beweis von Satz 7.24.

1. In Aufgabe 1 des 11. Übungszettels wurde bewiesen, dass die Verknüpfung linearer Abbildungen linear ist. Die darstellenden Matrizen $A_f = (\alpha_{i,j})_{i\in\{1,\dots,n_2\},j\in\{1,\dots,n_1\}}$ und $A_g = (\beta_{k,i})_{k\in\{1,\dots,n_3\},i\in\{1,\dots,n_2\}}$ erfüllen die Bedingungen

$$f(u_j) = \sum_{i=1}^{n_2} \alpha_{i,j} v_i \quad \text{und} \quad g(v_i) = \sum_{k=1}^{n_3} \beta_{k,i} w_k.$$

   Damit ist

$$\begin{aligned}
(g \circ f)(u_j) = g\left(\sum_{i=1}^{n_2} \alpha_{i,j} v_i\right) &= \sum_{i=1}^{n_2} \alpha_{i,j} g(v_i) \\
&= \sum_{i=1}^{n_2} \alpha_{i,j} \sum_{k=1}^{n_3} \beta_{k,i} w_k = \sum_{k=1}^{n_3} \left(\sum_{i=1}^{n_2} \beta_{k,i}\alpha_{i,j}\right) w_k =: \sum_{k=1}^{n_3} \gamma_{k,j} w_k
\end{aligned}$$

   mit $\gamma_{k,j} = \sum_{i=1}^{n_2} \beta_{k,i}\alpha_{i,j}$. Die Matrix $C = (\gamma_{k,j})_{k\in\{1,\dots,n_3\},j\in\{1,\dots,n_1\}} \in \mathbb{K}^{n_3\times n_1}$ mit $C = A_g \cdot A_f$ ist somit die darstellende Matrix von $g \circ f$ bezüglich $\mathcal{B}_U$ und $\mathcal{B}_W$, d.h. $A_{g\circ f} = A_g \cdot A_f$.

2. Sind f und g bijektiv, so gilt für alle $u \in U$ und $w \in W$

$$(g\circ f)(u) = g(f(u)) = w \iff f(u) = g^{-1}(w) \iff u = f^{-1}(g^{-1}(w)) = (f^{-1}\circ g^{-1})(w).$$

   Somit haben wir nicht nur gezeigt, dass $g \circ f$ invertierbar ist, sondern wir haben auch die Inverse als $(g \circ f)^{-1}$ identifiziert. Nach Teil 3 von Satz 7.18 gilt

$$A_{(g\circ f)^{-1}} = A_{g\circ f}^{-1} \overset{1.}{=} (A_g \cdot A_f)^{-1} \overset{(*)}{=} A_f^{-1} \cdot A_g^{-1},$$

   wobei wir $(*)$ in Aufgabe 2.3 gezeigt haben.

□

7.5 Die Koordinaten- und Basistransformation

Koordinaten von Vektoren sind abhängig von der gewählten Basis des jeweiligen Vektorraumes. Hat ein Vektorraum V die zwei geordneten Basen $\mathcal{B} := (b_1, \dots, b_n)$ und $\mathcal{C} := (c_1, \dots, c_n)$, so hat jeder Vektor $v \in V$ zwei verschiedene Koordinatenvektoren aus den folgenden Linearkombinationen:

$$v = \sum_{k=1}^{n} \beta_k b_k \quad \text{und} \quad v = \sum_{j=1}^{n} \gamma_j c_j.$$

Wir haben gesehen, dass die Koordinaten durch Lösen eines LGS gefunden werden. Kennt man bereits die Koordinaten von Vektoren bezüglich Basis $\mathcal{B}$, so stellt sich die Frage, ob es eine einfache Methode gibt, die Koordinaten bezüglich $\mathcal{C}$ zu finden, ohne für jeden Vektor erneut ein LGS lösen zu müssen.

Beispiel 7.25 (Fortsetzung von Beispiel 6.51)**.** *Der Vektorraum $\mathcal{P}_2$ hat die zwei geordneten Basen*

$$\mathcal{B} := (1, x, x^2) \quad \textit{und} \quad \mathcal{C} := (1-x, 1+x, x^2).$$

Die Koordinaten eines beliebigen Polynoms $p(x) = a + bx + cx^2$ bezüglich $\mathcal{B}$ konnten wir direkt ablesen:

$$\kappa_{\mathcal{B}}(p) = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Um die Koordinaten bezüglich $\mathcal{C}$ zu bestimmen, suchen wir Koeffizienten $\alpha, \beta, \gamma \in \mathbb{R}$ derart, dass

$$a + bx + cx^2 = \alpha(1-x) + \beta(1+x) + \gamma x^2 = (\alpha+\beta)\cdot 1 + (\beta-\alpha)\cdot x + \gamma \cdot x^2. \tag{7.5.1}$$

Koeffizientenvergleich führt auf das LGS

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ -1 & 1 & 0 & b \\ 0 & 0 & 1 & c \end{array}\right]$$

mit der eindeutigen Lösung

$$\kappa_{\mathcal{C}}(p) = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(a-b) \\ \frac{1}{2}(a+b) \\ c \end{pmatrix}.$$

Wir haben das Polynom $p(x) = a + bx + cx^2 \in \mathcal{P}_2$ bewusst allgemein gewählt, denn dank der Parameter können wir die Koordinaten jedes beliebigen Polynoms aus $\mathcal{P}_2$ sofort durch Einsetzen in die Formeln für $\kappa_{\mathcal{B}}(p)$ und $\kappa_{\mathcal{C}}(p)$ berechnen.

Bezeichnen wir die beiden auftretenden Matrizen mit

$$A_1 := \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \textit{und} \quad A_2 := \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Spalten von A_1 sind dabei gerade die Koordinaten der Vektoren aus $\mathcal{C}$ bezüglich $\mathcal{B}$ und die Spalten von A_2 sind die Koordinaten der Vektoren von $\mathcal{B}$ bezüglich $\mathcal{C}$, d.h. für alle $p \in \mathcal{B}$ und $q \in \mathcal{C}$ gilt

$$\kappa_{\mathcal{B}}(q) = A_1 \cdot \kappa_{\mathcal{C}}(q) \quad \textit{und} \quad \kappa_{\mathcal{C}}(p) = A_2 \cdot \kappa_{\mathcal{B}}(p).$$

Letztere hatten wir in Beispiel 6.51 bereits berechnet – hier führen wir sie nochmals zum Vergleich auf.

$$\begin{aligned}
\frac{1}{2} \cdot (1-x) + \frac{1}{2} \cdot (1+x) + 0 \cdot x^2 &= 1, \\
\left(-\frac{1}{2}\right) \cdot (1-x) + \frac{1}{2} \cdot (1+x) + 0 \cdot x^2 &= x, \\
0 \cdot (1-x) + 0 \cdot (1+x) + 1 \cdot x^2 &= x^2
\end{aligned}$$

Setzen wir $f := \mathrm{id}_{\mathcal{P}_2}$, so ist

$$A_1 = A_f^{\mathcal{C},\mathcal{B}} \quad \textit{und} \quad A_2 = A_f^{\mathcal{B},\mathcal{C}}.$$

Definieren wir Abbildungen $f_1, f_2 \colon \mathbb{R}^3 \to \mathbb{R}^3$ als $f_1(x) := A_1 \cdot x$ und $f_2(x) := A_2 \cdot x$, so erhalten wir das folgende kommutative Diagramm:

$$\begin{tikzcd}
\mathcal{P}_2 \arrow[r, "\mathrm{id}_{\mathcal{P}_2}"] \arrow[d, "\kappa_{\mathcal{B}}"'] & \mathcal{P}_2 \arrow[d, "\kappa_{\mathcal{C}}"] \\
\mathbb{R}^3 \arrow[r, "f_2 = (f_1)^{-1}"'] & \mathbb{R}^3
\end{tikzcd}$$

Definition 7.26

Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und seien $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_n)$ geordnete Basen von V. Dann heißt die darstellende Matrix $T \in \mathbb{K}^{n \times n}$ der Identität $\mathrm{id} \colon V \to V$ bezüglich $\mathcal{B}$ und $\mathcal{C}$, d.h. $T = A_{\mathrm{id}_V}^{\mathcal{B},\mathcal{C}}$, die *Transformations-* oder *Übergangsmatrix des Basiswechsels* von $\mathcal{B}$ nach $\mathcal{C}$.

Bemerkung 7.27. *Fasst man die geordneten Basen $\mathcal{B}$ und $\mathcal{C}$ als Matrizen auf (wie schon im Beweis von 3. von Satz 7.15), so gilt*

$$\mathcal{B} = (A_f^{\mathcal{B},\mathcal{C}})^T \mathcal{C} = ((A_f^{\mathcal{C},\mathcal{B}})^{-1})^T \mathcal{C} \quad \textit{und} \quad \mathcal{C} = (A_f^{\mathcal{C},\mathcal{B}})^T \mathcal{B} = ((A_f^{\mathcal{B},\mathcal{C}})^{-1})^T \mathcal{B},$$

d.h. die Basistransformationsmatrix, welche eine Basis auf einer anderen (desselben Raumes!) abbildet, ist die Transponierte der Inversen der Koordinatentransformationsmatrix, welche zu diesem Basiswechsel gehört.

In obigem Beispiel gilt für die Basistransformation:

$$\begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1-x \\ 1+x \\ x^2 \end{pmatrix} \quad \textit{und} \quad \begin{pmatrix} 1-x \\ 1+x \\ x^2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix}.$$

7.6 Aufgaben

Aufgabe 7.1. *Füllen Sie die Lücken in folgenden Sätzen.*

a) *Seien V und W zwei $\mathbb{K}$-Vektorräume. Dann bezeichnen wir mit $L(V;W)$*

____.

b) *Eine Abbildung $f\colon V \to W$ heißt Vektorraum-Homomorphismus, falls gilt:*

$$f(x+y) = f(x) + f(y), \qquad \forall\underline{\qquad\qquad\qquad}$$
$$\underline{\qquad\qquad\qquad\qquad}, \qquad \forall \lambda \in \mathbb{K}, \underline{\qquad\qquad\qquad}$$

Beide Bedingungen kann man zu einer zusammenfassen:

$$f(\lambda x + \mu y) = \underline{\qquad\qquad\qquad\qquad\qquad}, \qquad \forall\underline{\qquad\qquad\qquad\qquad\qquad}.$$

Ein Vektorraum-Homomorphismus wird synonym auch als ________________ Abbildung bezeichnet.

c) *Eine Abbildung $f\colon V \to W$ heißt ________________, falls gilt:*

$$\forall w \in W \;\exists\, v \in V\colon f(v) = w.$$

Etwas kürzer lässt sich das formulieren als $f(V) = \mathrm{Bild}(f) =$ ________, wobei das Bild definiert ist als

$$\mathrm{Bild}(f) = \underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}.$$

d) *Eine Abbildung $f\colon V \to W$ heißt ________________, falls für alle $x, y \in V$ gilt:*

$$f(x) = f(y) \Longrightarrow x = y.$$

Äquivalent hierzu ist die Kontraposition:

$$\underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}$$

e) *Eine Abbildung ist genau dann bijektiv, wenn sie ____________ und ____________ ist. Zudem ist eine Abbildung genau dann invertierbar, wenn sie ____________ ist. Eine bijektive lineare Abbildung bezeichnet man auch als ____________.*

f) *Bezeichnen wir die Nullvektoren der Räume V und W mit e_V und e_W. Der Kern einer linearen Abbildung ist dann definiert als*

$$\ker(f) = \underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}.$$

Eine lineare Abbildung ist genau dann injektiv, wenn $\ker(f) =$ ________ ist.

Aufgabe 7.2. *Gegeben seien*

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}.$$

Bestimmen Sie $A \cdot v_k$ für $k \in \{1,2,3,4\}$.

Aufgabe 7.3. *Sei $A \in \mathbb{R}^{m\times n}$. Sei $f\colon \mathbb{R}^n \to \mathbb{R}^m$ die lineare Abbildung mit $f(x) = A \cdot x$ und sei $\mathcal{A}$ die Menge der Spaltenvektoren von A.*

a) *Zeigen Sie, dass $\mathcal{L}(\mathcal{A}) = \mathrm{Bild}(f)$ ist.*

b) *Folgern Sie damit, dass folgende Aussagen äquivalent sind:*

   (1) $A \cdot x = b$ *ist lösbar;*

   (2) $b \in \mathcal{L}(\mathcal{A})$;

   (3) $b \in \mathrm{Bild}(f)$.

Aufgabe 7.4. *Sei $f\colon \mathcal{P}_n \to \mathcal{P}_n$ die Abbildung, die jeder reellen Polynomfunktion ihre Ableitung zuordnet, d.h. $f(p(x)) = p'(x)$.*

a) *Ist f linear?*

b) *Bestimmen Sie das Bild von f, d.h. $f(\mathcal{P}_n)$.*

c) *Bestimmen Sie den Kern von f.*

d) *Ist f surjektiv und/oder injektiv?*

Aufgabe 7.5. *Betrachten Sie die Abbildung $f\colon \mathbb{C} \to \mathbb{C}$ mit $f(z) = \overline{z}$.*

a) *Ist f linear auf dem $\mathbb{R}$-Vektorraum $\mathbb{C}$?*

b) *Ist f linear auf dem $\mathbb{C}$-Vektorraum $\mathbb{C}$?*

Aufgabe 7.6. *Seien V und W zwei $\mathbb{K}$-Vektorräume mit $\dim(V) = \dim(W) < \infty$. Sei $f\colon V \to W$ eine lineare Abbildung. Dann gilt:*

$$f \text{ ist injektiv} \iff f \text{ ist surjektiv} \iff f \text{ ist bijektiv.}$$

Aufgabe 7.7. *Gibt es lineare Abbildungen $f\colon \mathbb{R}^3 \to \mathbb{R}^4$,*

a) *die injektiv sind?* b) *die surjektiv sind?* c) *die bijektiv sind?*

Aufgabe 7.8. *Gegeben ist $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ mit*

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+4y+7z \\ 2x+5y+8z \\ 3x+6y+9z \end{pmatrix}.$$

Ist f bijektiv? Welche Möglichkeiten haben Sie, um das zu überprüfen?

Aufgabe 7.9. *Sei $\mathcal{B}$ die Standardbasis des $\mathbb{R}^2$ und seien*

$$A_f^{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{und} \quad \mathcal{B}' = \left(\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix} \right).$$

Bestimmen Sie $A_f^{\mathcal{B}'}$.

Aufgabe 7.10. *Sie haben in Aufgabe 7.4 die lineare Abbildung $f\colon \mathcal{P}_4 \to \mathcal{P}_4$ mit $f(p(x)) = p'(x)$ kennen gelernt. Seien $\mathcal{B} := (1, x, x^2, x^3, x^4)$ und $\mathcal{C} := (1, 1+x, 1+x+x^2, 1+x+x^2+x^3, 1+x+x^2+x^3+x^4)$ gegebene Basen von $\mathcal{P}_4$. Bestimmen Sie die darstellenden Matrizen $A_f^{\mathcal{B}}$ und $A_f^{\mathcal{B},\mathcal{C}}$ von f bezüglich der jeweiligen Basen (Vgl. Satz 7.15).*

Aufgabe 7.11. *Sei* $g\colon \mathcal{P}_4 \to \mathcal{P}_4$ *gegeben durch* $g(p(x)) := p(x) - p'(x)$.

a) *Sei* $\mathcal{B} := (1, x, x^2, x^3, x^4)$ *die Standardbasis von* $\mathcal{P}_4$. *Bestimmen Sie die darstellende Matrix* $A_g^{\mathcal{B}}$.

b) *Bestimmen Sie Bild und Kern der Abbildung* $h\colon \mathbb{R}^5 \to \mathbb{R}^5$ *mit* $h(x) := A_g^{\mathcal{B}} \cdot x$.

c) *Ist* g *bijektiv?*

Aufgabe 7.12. *Für* $k \in \{1, \dots, 8\}$ *seien* $f_k\colon \mathbb{R}^2 \to \mathbb{R}^2$ *die folgendermaßen definierten Abbildungen:*

$$f_1\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}, \qquad f_2\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix},$$

$$f_3\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\cos(\alpha) & -\sin(\alpha)\\\sin(\alpha) & \cos(\alpha)\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}, \quad f_4\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\cos(\alpha) & \sin(\alpha)\\-\sin(\alpha) & \cos(\alpha)\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix} \quad (\alpha \in \mathbb{R}),$$

$$f_5\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}1/2 & -\sqrt{3}/2\\\sqrt{3}/2 & 1/2\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}, \quad f_6\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}1/2 & \sqrt{3}/2\\-\sqrt{3}/2 & 1/2\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix},$$

$$f_7\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3 & 0\\0 & 2\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}, \qquad f_8\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}.$$

Geben Sie eine geometrische Deutung der angegebenen Abbildungen. Mit anderen Worten: Was bewirken diese?

Aufgabe 7.13. *Sei* $f\colon \mathbb{R}^3 \to \mathbb{R}^4$ *gegeben durch*

$$f\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x+y\\x+y\\x+2y+z\\x\end{pmatrix}.$$

Seien $\mathcal{B}$ *und* $\mathcal{C}$ *die Standardbasen des* $\mathbb{R}^3$ *und* $\mathbb{R}^4$ *und sei*

$$\mathcal{B}' = \left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right).$$

a) *Berechnen Sie die Transformationsmatrix* $T := A_{\mathrm{id}_{\mathbb{R}^3}}^{\mathcal{B},\mathcal{B}'}$ *und deren Inverse* $T^{-1} = A_{\mathrm{id}_{\mathbb{R}^3}}^{\mathcal{B}',\mathcal{B}}$.

b) *Berechnen Sie die darstellenden Matrizen* $A_f^{\mathcal{B},\mathcal{C}}$ *und* $A_f^{\mathcal{B}',\mathcal{C}}$.

c) *Finden Sie einen Zusammenhang zwischen den drei von Ihnen berechneten Matrizen.*

Aufgabe 7.14. *Aufgabe zur Vorbereitung von Aufgabe 7.15:*

a) Ist $\begin{pmatrix} 4 \\ 3 \\ 1 \\ -1 \end{pmatrix} \in \mathcal{L}\left(\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}\right\}\right)$ *?*

b) Liegt $\begin{pmatrix} 4 \\ 3 \\ 1 \\ -1 \end{pmatrix}$ *im Bild der linearen Abbildung* $f_A \colon \mathbb{R}^3 \to \mathbb{R}^4$ *mit* $f_A(x) = A \cdot x$*, wobei*

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 3 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \text{ ist?}$$

c) Hat das folgende LGS eine Lösung?

$$\begin{aligned} x - y + z &= 4 \\ 3y + z &= 3 \\ x \phantom{{}+3y} + z &= 1 \\ -x + y + z &= -1 \end{aligned}$$

Aufgabe 7.15. *Wir betrachten im Folgenden ein* $m \times n$*-LGS der Form*

$$\begin{aligned} \alpha_{1,1}x_1 + \ldots + \alpha_{1,n}x_n &= b_1 \\ &\vdots \\ \alpha_{m,1}x_1 + \ldots + \alpha_{m,n}x_n &= b_m, \end{aligned} \tag{7.6.1}$$

die Matrix und die Vektoren

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{R}^{m \times n}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m \tag{7.6.2}$$

und die lineare Abbildung

$$f \colon \mathbb{R}^n \to \mathbb{R}^m, \ x \mapsto A \cdot x. \tag{7.6.3}$$

Die Spalten von A *bezeichnen wir mit* $S_1, \ldots, S_n$.

Die unten gegebenen Begriffe und Aussagen beziehen sich jeweils auf LGS, Matrizen oder lineare Abbildungen. Ordnen Sie die Begriffe und Aussagen einer der drei Kategorien zu und sortieren Sie sie in einer Tabelle der folgenden Form derart, dass inhaltlich zusammengehörende Begriffe oder Aussagen in gleichen Zeilen stehen.

LGS (7.6.1)	*Matrix* (7.6.2)	*lineare Abb.* (7.6.3)
x *ist Lösung von* (7.6.1)	$A \cdot x = b$	$f(x) = b$
$\vdots$	$\vdots$	$\vdots$

Begriffe und Aussagen:

- *f ist surjektiv.*
- $\{x \mid A \cdot x = 0_m\}$
- *Eine Lösung von* (7.6.1) *existiert für alle $b \in \mathbb{R}^m$.*
- $\mathrm{Bild}(f)$
- $L(A, b)$
- $\ker(f)$
- *A hat den maximalen Zeilenrang.*
- *A hat den maximalen Spaltenrang.*
- $\mathcal{L}(\{S_1, \ldots, S_n\}) = \mathbb{R}^m$.
- *A^{-1} existiert.*
- $b \in \mathrm{Bild}(f)$.
- *Für jedes $b \in \mathbb{R}^m$ existiert genau eine Lösung x von* (7.6.1).
- *f ist injektiv.*
- $\mathcal{L}(\{S_1, \ldots, S_n\})$
- *die Lösungsmenge des zu* (7.6.1) *gehörigen homogenen Systems*
- $\mathrm{rg}(f) = m$.
- $\mathrm{rg}(f) = n$.
- *Die Spalten von A sind eine Basis des $\mathbb{R}^m$.*
- *f^{-1} existiert.*
- *Mindestens eine Lösung von* (7.6.1) *existiert.*
- *f ist bijektiv.*
- $\{x \mid A \cdot x = b\}$
- $\mathrm{rg}(A) = m = n$.
- $\dim(\mathrm{Bild}(f)) = m$.
- *A ist invertierbar.*
- *Die Spalten von A sind linear unabhängig.*
- *Menge aller $b \in \mathbb{R}^m$, für die* (7.6.1) *lösbar ist*
- $\{x \in \mathbb{R}^n \mid f(x) = b\}$
- $b \in \mathcal{L}(\{S_1, \ldots, S_n\})$.
- *Lösung von* (7.6.1) *ist, wenn sie existiert, eindeutig für jedes $b \in \mathbb{R}^m$.*
- $\mathrm{rg}(f) = m = n$.
- $\ker(f) = \{0_n\}$.

Kapitel 8

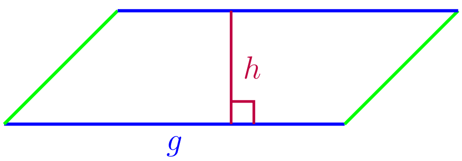
Determinanten

Ziele

- ☐ die Interpretation und Berechnung von Determinanten von 2×2- und 3×3-Matrizen kennen
- ☐ Determinanten von beliebig großen Matrizen berechnen
- ☐ mit Hilfe von Determinanten auf Invertierbarkeit einer Matrix schließen und die Inverse berechnen
- ☐ den abstrakten Begriff der Determinante und die davon abgeleiteten Eigenschaften und Rechenregeln kennen und anwenden können, Beweise dazu

8.1 Determinanten von 2×2- und 3×3-Matrizen

8.1.1 Der Flächeninhalt eines Parallelogramms



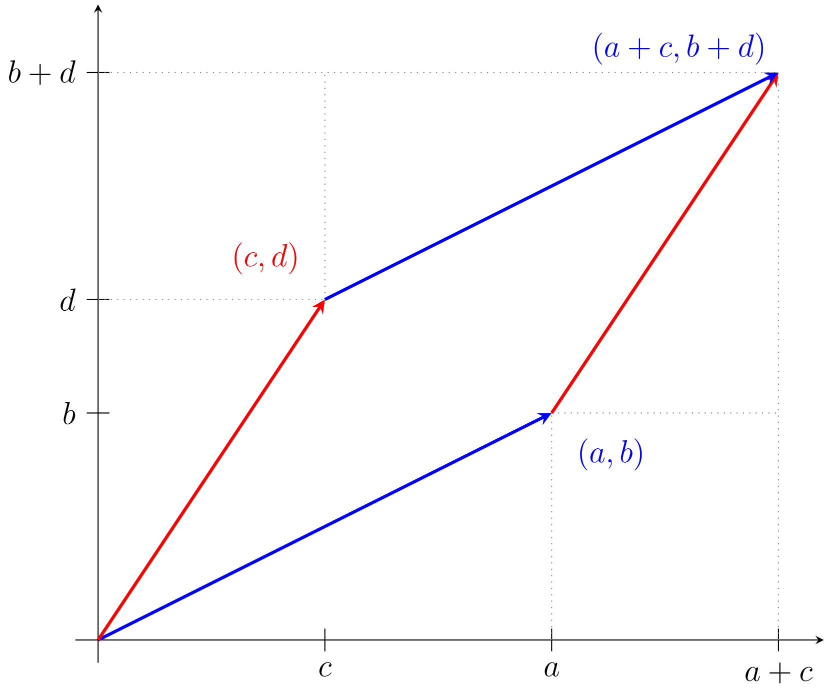
Der Flächeninhalt eines Parallelogramms mit einer Grundseite der Länge g und Höhe h ist gegeben durch $A = g \cdot h$.

Im Folgenden wollen wir den Flächeninhalt eines Parallelogramms bestimmen, dessen Seiten durch Vektoren angegeben werden. Dazu machen wir ein paar Vorüberlegungen:

- Je zwei gegenüberliegende Winkel im Parallelogramm sind gleich. Da die Innenwinkelsumme eines Vierecks stets $360°$ beträgt, ist einer der beiden auftretenden Winkel bei einem echten Parallelogramm(, dass also kein Rechteck ist), ein spitzer Winkel mit $< 90°$.
- Drehen und Verschieben eines Parallelogramms ändert seinen Flächeninhalt nicht. Somit können wir nach Drehen und Verschieben ein jedes Parallelogramm so positionieren, dass eine Ecke mit dem spitzen Winkel auf dem Koordinatenursprung liegt und das ganze Parallelogramm im 1. Quadranten liegt.

Nach diesen Vorüberlegungen wollen wir nun den Flächeninhalt eines von den Vektoren

$\begin{pmatrix} a \\ b \end{pmatrix}$ und $\begin{pmatrix} c \\ d \end{pmatrix}$ aufgespannten Parallelogramms berechnen.



Durch geschickte Wahl von rechtwinkligen Dreiecken und Rechtecken können wir den Flächeninhalt A_{Par} des Parallelogramms berechnen:

$$\begin{aligned} A_{\text{Par}} &= (a+c)\cdot(b+d) - \left(\frac{1}{2}ab + cb + \frac{1}{2}cd + \frac{1}{2}cd + bc + \frac{1}{2}ab\right) \\ &= (ab+ad+bc+cd) - (ab+cd+2bc) \\ &= ad - bc. \end{aligned}$$

Schreiben wir beide Vektoren zusammen in eine Matrix, so nennen wir

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc$$

die *Determinante* dieser Matrix. Da diese Zahl für jede 2×2-Matrix existiert, können wir die zugehörige Abbildung definieren.

Definition 8.1

Sei $\mathbb{K}$ ein Körper. Die Abbildung

$$\det\colon \mathbb{K}^{2\times 2} \to \mathbb{K}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc = \det(A)$$

heißt *Determinante*.

Die Determinante tauchte bereits in Aufgabe 2.4 auf. Das dort gefundene Ergebnis halten wir an dieser Stelle noch einmal fest:

Lemma 8.2

Sei $\mathbb{K}$ ein Körper und
$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2\times 2}.$$
Die Matrix A ist genau dann invertierbar, wenn $\det(A) \neq 0$ ist. In diesem Fall ist ihre Inverse gegeben durch
$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Kommen wir nochmals zurück zum Flächeninhalt des durch zwei Vektoren aufgespannten Parallelogramms und somit Determinanten reellwertiger 2×2-Matrizen.

Beispiel 8.3. *Der Flächeninhalt eines Quadrates mit Seitenlänge 1 ist 1 und genau das ist die Determinante der Einheitsmatrix E_2:*
$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

Beispiel 8.4. *Strecken wir eine der Seiten um den Faktor $\lambda > 0$, d.h. multiplizieren wir einen Vektor mit diesem Faktor, so wird auch der Flächeninhalt mit diesem Faktor multipliziert. Exemplarisch sehen wir das am ersten Vektor:*
$$\det \begin{pmatrix} \lambda a & \lambda b \\ c & d \end{pmatrix} = \lambda ad - \lambda bc = \lambda \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

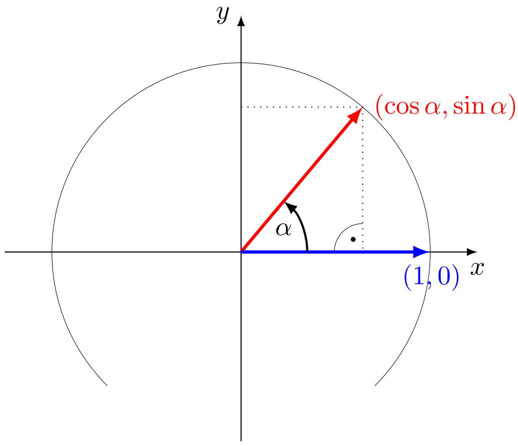
Beispiel 8.5 (Spiegelung an einer Koordinatenachse)**.** *Die Spiegelung an der x-Achse entspricht der Multiplikation der y-Koordinaten beider Vektoren mit -1. Während der Flächeninhalt eines gespiegelten Parallelogramms unverändert bleibt, wechselt die Determinante das Vorzeichen:*
$$\det \begin{pmatrix} a & -b \\ c & -d \end{pmatrix} = a \cdot (-d) - (-b) \cdot c = -(ad - bc) = -\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$
Gleiches gilt für die Spiegelung an der y-Achse:
$$\det \begin{pmatrix} -a & b \\ -c & d \end{pmatrix} = (-a) \cdot d - b \cdot (-c) = -(ad - bc) = -\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$
Die Determinante gibt also den orientierten Flächeninhalt[1] *des Parallelogramms an.*

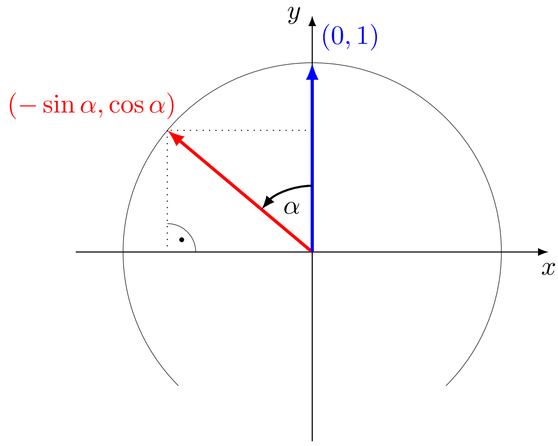
Beispiel 8.6 (Drehung um den Ursprung)**.** *Eine Drehung um den Winkel α ändert weder den Flächeninhalt des Parallelogramms noch die Determinante der zugehörigen Matrix. Wird die Drehung im mathematisch positiven Sinn (d.h. entgegen dem Uhrzeigersinn) durchgeführt, so kann eine Drehung durch die Multiplikation mit der Drehmatrix*
$$D_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

[1] Den Begriff der Orientierung werden wir später genauer ansehen. An dieser Stelle soll es genügen, festzuhalten, dass zwei Vektoren $x, y \in \mathbb{R}^2$ positiv orientiert sind, wenn x um einen Winkel $< 180°$ im mathematisch positiven Sinn gedreht werden muss um in die gleiche Richtung wie y zu zeigen.

beschrieben werden, was anhand der Wirkung auf die Standardbasis des $\mathbb{R}^2$ nachvollzogen werden kann:

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \qquad \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}.$$





Wenden wir nun die Drehmatrix auf eine beliebige Matrix an, so ändert dies nicht die Determinante und somit auch nicht den Flächeninhalt des von den Vektoren aufgespannten Parallelogramms:

$$\begin{aligned}
&\det\left(\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) \\
&= \det\begin{pmatrix} a\cos\alpha - c\sin\alpha & b\cos\alpha - d\sin\alpha \\ a\sin\alpha + c\cos\alpha & b\sin\alpha + d\cos\alpha \end{pmatrix} \\
&= (a\cos\alpha - c\sin\alpha)\cdot(b\sin\alpha + d\cos\alpha) - (b\cos\alpha - d\sin\alpha)\cdot(a\sin\alpha + c\cos\alpha) \\
&= ad\cos^2\alpha - bc\sin^2\alpha - bc\cos^2\alpha + ad\sin^2\alpha \\
&= (ad - bc)(\cos^2\alpha + \sin^2\alpha) \\
&= ad - bc \\
&= \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}.
\end{aligned}$$

Insbesondere folgt hiermit sofort (mit $a = d = 1$ und $b = c = 0$), dass $\det(D_\alpha) = 1$ ist.

8.1.2 Lösung eines 3×3-LGS

Betrachten wir ein LGS der Form $A \cdot x = b$, wobei $A \in \mathbb{K}^{3\times 3}$ und $b \in \mathbb{K}^3$ sei. Analog zu Lemma 8.2 können wir wieder überlegen, unter welchen Bedingungen die Matrix A invertierbar ist und das LGS somit eine eindeutige Lösung besitzt.

Da eine Matrix genau dann invertierbar ist, wenn sie vollen Rang besitzt, müssen wir also nur eine Bedingung dafür finden, dass der Rang der Matrix 3 ist.

Zunächst einmal muss in der ersten Spalte ein Eintrag von 0 verschieden sein, da die Matrix sonst höchstens Rang 2 haben kann. Nehmen wir also an, dass $a \neq 0$ ist. Dann gilt:

$$\begin{aligned}
\operatorname{rg}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= \operatorname{rg}\begin{pmatrix} a & b & c \\ 0 & ae-bd & af-cd \\ 0 & ah-bg & ai-cg \end{pmatrix} \\
&= \operatorname{rg}\begin{pmatrix} a & b & c \\ 0 & ae-bd & af-cd \\ 0 & 0 & (ai-cg)(ae-bd)-(af-cd)(ah-bg) \end{pmatrix}
\end{aligned}$$

Wie lautet der letzte Term ausmultipliziert?

$$\begin{aligned}
&(ai-cg)(ae-bd)-(af-cd)(ah-bg) \\
&= a^2ei - abdi - aceg + bcdg - a^2fh + abfg + acdh - bcdg \\
&= a(aei - bdi - ceg - afh + bfg + cdh) + b(cdg - cdg) \\
&= a(aei - bdi - ceg - afh + bfg + cdh)
\end{aligned}$$

Folglich ist (nachdem wir die Terme nach Vorzeichen sortiert haben)

$$\operatorname{rg}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = 3 \iff aei + bfg + cdh - afh - bdi - ceg \neq 0$$

Definition 8.7

Sei $\mathbb{K}$ ein Körper. Die Abbildung

$$\det\colon \mathbb{K}^{3\times 3} \to \mathbb{K}, \quad A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \mapsto aei + bfg + cdh - afh - bdi - ceg = \det(A)$$

heißt *Determinante*.

Analog zu Lemma 8.2 können wir das eben gefundene Invertierbarkeitskriterium festhalten:

Lemma 8.8

Sei $\mathbb{K}$ ein Körper und

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \in \mathbb{K}^{3\times 3}.$$

Die Matrix A ist genau dann invertierbar, wenn $\det(A) \neq 0$ ist.

Die Determinante einer 3×3-Matrix lässt sich mit der *Regel von Sarrus* (benannt nach dem französischen Mathematiker Pierre Frédéric Sarrus (1798-1861)) merken.

Regel von Sarrus

Man multipliziert die Einträge entlang der Diagonalen und addiert sie unter Beachtung des Vorzeichens:

$$
\begin{array}{cccccccccccccc}
+ & & & & & & & & & & & & & \\
 & a & & b & & c & \qquad & & a & & b & & c \\
+ & & \searrow & & & & & & & & & \nearrow & \\
 & d & & e & & f & & & d & & e & & f \\
+ & & \searrow & & \searrow & & & & & \nearrow & & \nearrow & \\
 & g & & h & & i & & & g & & h & & i \\
 & & \searrow & & \searrow & & & - & & \nearrow & & \nearrow & \\
 & a & & b & & c & & & a & & b & & c \\
 & & & & \searrow & & & - & & \nearrow & & & \\
 & d & & e & & f & & & d & & e & & f \\
 & & & & & & & - & & & & &
\end{array}
$$

Damit erhält man

$$\det(A) = a \cdot e \cdot i + d \cdot h \cdot c + g \cdot b \cdot f - g \cdot e \cdot c - a \cdot h \cdot f - d \cdot b \cdot i.$$

Beispiel 8.9. *Betrachten wir nochmals die Matrix*

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

Wir wissen bereits, dass $\operatorname{rg}(A) = 2$ *ist und in der Tat ist die Determinante Null:*

$$\begin{aligned} \det(A) &= 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 3 \cdot 5 \cdot 7 - 6 \cdot 8 \cdot 1 - 9 \cdot 2 \cdot 4 \\ &= 45 + 84 + 96 - 105 - 48 - 72 = 0. \end{aligned}$$

Beispiel 8.10. *Sehen wir uns wieder die Determinante der Einheitsmatrix an:*

$$\det(E_3) = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1,$$

d.h. die Determinante ist gleich dem Volumen eines Würfels mit Seitenlänge 1. Im nächsten Abschnitt werden wir sehen, dass auch die Determinante einer 3×3*-Matrix eine geometrische Interpretation hat – analog zu der einer* 2×2*-Matrix.*

Beispiel 8.11. *Da elementare Zeilenumformungen den Rang einer Matrix nicht ändern, sollten sie auch nichts daran ändern, ob eine Matrix Determinante 0 hat oder nicht. Sehen wir uns das genauer an.*

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}, \quad B = \begin{pmatrix} 3a & 3b & 3c \\ d & e & f \\ g & h & i \end{pmatrix}, \quad C = \begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}, \quad D = \begin{pmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{pmatrix}.$$

Die Determinanten der Matrizen sind:

$$\begin{aligned}
\det(A) &= aei + cdh + bfg - ceg - afh - bdi,\\
\det(B) &= 3aei + 3cdh + 3bfg - 3ceg - 3afh - 3bdi\\
&= 3\det(A),\\
\det(C) &= bdi + ceg + afh - bfg - aei - cdh\\
&= -\det(A),\\
\det(D) &= (a+d)ei + (c+f)dh + (b+e)fg - (c+f)eg - (a+d)fh - (b+e)di\\
&= \det(A) + dei + dfh + efg - efg - dfh - dei\\
&= \det(A).
\end{aligned}$$

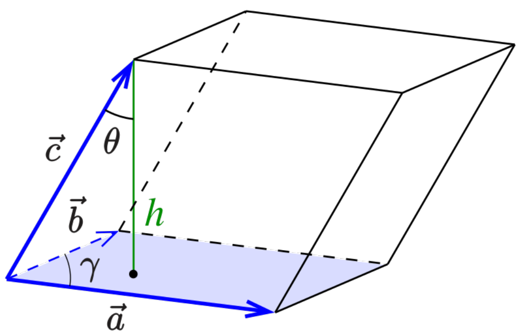
Welchen Einfluss haben also elementare Zeilenumformungen auf die Determinante?

Beobachtungen:

- *Multipliziert man eine Zeile einer Matrix mit einem (von 0 verschiedenen) Faktor, so wird auch die Determinante mit diesem Faktor multipliziert.*
- *Vertauscht man zwei Zeilen, so ändert sich das Vorzeichen der Determinante.*
- *Addiert man eine Zeile zu einer anderen hinzu, so ändert sich die Determinante der Matrix nicht.*

Bevor wir Determinanten beliebiger quadratischer Matrizen einführen, wollen wir noch kurz auf die geometrische Interpretation der Determinante einer 3×3-Matrix eingehen.

8.1.3 Das Volumen eines Spats



Ein *Spat*[a] (oder *Parallelepiped*) ist ein Körper, der von sechs paarweise kongruenten, in parallelen Ebenen liegenden Paralellogrammen begrenzt wird.

[a] Der Begriff *Spat* stammt von dem Mineral Calcit, welches auch als Kalkspat oder Doppelspat bezeichnet wird und dessen Kristalle diese Form aufweisen.

Das Volumen des Spats[2] berechnet sich gemäß der Formel $V = A \cdot h$, wobei A der Flächeninhalt der Grundfläche und h die Höhe des Spats ist.

Wenn ein Spat wie im Bild dargestellt von drei Vektoren im $\mathbb{R}^3$ aufgespannt wird, ist das (orientierte) Volumen des Spats gegeben durch die Determinante der Matrix, die aus den drei Vektoren gebildet wird. Details hierzu werden wir erst im nächsten Kurs betrachten können nach der Einführung des Skalarprodukts und des Begriffs Orthogonalität. In der Schule wird hier neben dem Skalarprodukt auch das sogenannte Spatprodukt verwendet.

[2] Bild von Ag2gaeh - Eigenes Werk, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=74888453

8.2 Definition und Eigenschaften von Determinanten

Die ersten Betrachtungen zu Determinanten stammen vom Ende des 16. Jahrhunderts. Die ursprüngliche Motivation war so wie in Abschnitt 8.1.2 das Herleiten einer Bedingung, die uns sagt, ob ein LGS eine eindeutige Lösung besitzt. Später folgten auch andere Anwendungen und verschiedene Definitionen. Die folgende Definition geht auf Karl Weierstraß (1815-1897) zurück.

Definition 8.12

Seien $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Eine Abbildung

$$\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}, \quad A \mapsto \det(A)$$

heißt *Determinante*, falls Folgendes gilt:

(D1) $\det$ ist linear in jeder Zeile, d.h. für jeden Index $k \in \{1, \ldots, n\}$ gilt:
Ist $z_k = \lambda z_k' + \mu z_k''$ für $\lambda, \mu \in \mathbb{K}$ und Zeilenvektoren $z_k', z_k'' \in \mathbb{K}^{1\times n}$, so ist

$$\det\begin{pmatrix} \vdots \\ z_k \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ \lambda z_k' \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ \mu z_k'' \\ \vdots \end{pmatrix} = \lambda \det\begin{pmatrix} \vdots \\ z_k' \\ \vdots \end{pmatrix} + \mu \det\begin{pmatrix} \vdots \\ z_k'' \\ \vdots \end{pmatrix},$$

wobei alle Zeilen außer der k-ten in allen Matrizen identisch sind.

(D2) $\det$ ist *alternierend*, d.h. hat A zwei gleiche Zeilen, so ist $\det(A) = 0$.

(D3) $\det$ ist *normiert*, d.h. $\det(E_n) = 1$.

Beispiel 8.13 (Ein Beispiel zur Linearität). *Die Linearität wollen wir anhand eines Beispiels verdeutlichen:*

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} + \det\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 6 \end{pmatrix} = 1\cdot 4\cdot 6 + 1\cdot 2\cdot 6 = 36$$
$$= 1\cdot 6\cdot 6 = \det\begin{pmatrix} 1 & 2 & 3 \\ 0 & 6 & 8 \\ 0 & 0 & 6 \end{pmatrix}$$

Satz 8.14

Die als Determinanten eingeführten Abbildungen aus Definitionen 8.1 und 8.7 erfüllen die Eigenschaften (D1), (D2) und (D3), d.h. es handelt sich auch gemäß Definition 8.12 um Determinanten.

Beweis. Wir haben bereits gesehen, dass die Determinanten aus den Definitionen 8.1 und 8.7 normiert sind. Zudem folgt aus zwei gleichen Zeilen, dass die Matrix keinen vollen Rang haben kann, also nicht invertierbar ist und somit nach Lemma 8.2 bzw.

Lemma 8.8 die Determinante der Matrix 0 ist. Die eingeführten Abbildungen haben also beide Eigenschaften (D2) und (D3). Eigenschaft (D1) haben wir oben einmal exemplarisch für 3×3-Matrizen nachgeprüft. Ein formaler Beweis soll hier nicht erfolgen. Ein Beweis für Determinanten von 2×2-Matrizen ist Inhalt von Aufgabe 8.6. □

Bemerkung 8.15. *Wir haben also gesehen, dass es für $n = 2$ und $n = 3$ Determinanten gibt, d.h. es gibt Abbildungen mit den Eigenschaften (D1), (D2) und (D3). Wir wissen jedoch weder ob es für $n > 3$ Determinanten gibt noch die Determinanten mit nur den gegebenen drei Eigenschaften eindeutig bestimmt sind, d.h. ob für festes $n \in \mathbb{N}$ höchstens eine Abbildung mit den Eigenschaften (D1), (D2) und (D3) existiert.*

Satz 8.16

Für einen Körper $\mathbb{K}$ und $n \in \mathbb{N}$ sei $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Determinante. Dann hat $\det$ folgende Eigenschaften:

(E1) Für jedes $\lambda \in \mathbb{K}$ und $A \in \mathbb{K}^{n\times n}$ ist $\det(\lambda \cdot A) = \lambda^n \cdot \det(A)$.

(E2) Ist eine Zeile von $A \in \mathbb{K}^{n\times n}$ der Nullvektor, so ist $\det(A) = 0$.

(E3) Seien $A, B \in \mathbb{K}^{n\times n}$. Ist B aus A durch Vertauschung zweier Zeilen entstanden, so ist $\det(B) = -\det(A)$.

(E4) Seien $\lambda \in \mathbb{K}$ und $A \in \mathbb{K}^{n\times n}$. Ist $B \in \mathbb{K}^{n\times n}$ die Matrix, die entsteht, wenn man das λ-Fache der i-ten Zeile von A zur j-ten Zeile von A hinzu addiert (mit $i \neq j$), so ist $\det(B) = \det(A)$.

(E5) Liegt $A \in \mathbb{K}^{n\times n}$ in Zeilenstufenform vor, d.h. gilt

$$A = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ 0 & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \alpha_{n,n} \end{pmatrix},$$

so ist $\det(A) = \alpha_{1,1} \cdot \ldots \cdot \alpha_{n,n}$.

(E6) Sei $n \geq 2$ und $A \in \mathbb{K}^{n\times n}$. Sei A von der Form

$$A = \begin{pmatrix} A_1 & C \\ O & A_2 \end{pmatrix},$$

wobei $A_1 \in \mathbb{K}^{n_1\times n_1}$ und $A_2 \in \mathbb{K}^{n_2\times n_2}$ mit $n_1 + n_2 = n$ sind, $C \in \mathbb{K}^{n_1\times n_2}$ beliebig ist und $O \in \mathbb{K}^{n_2\times n_1}$ die Nullmatrix ist. Dann ist $\det(A) = \det(A_1) \cdot \det(A_2)$.

(E7) Für $A \in \mathbb{K}^{n\times n}$ gilt $\det(A) = 0$ genau dann, wenn $\operatorname{rg}(A) < n$ ist.

(E8) Sind $A, B \in \mathbb{K}^{n\times n}$, so ist $\det(A \cdot B) = \det(A) \cdot \det(B)$.

(E9) Ist $A \in \mathbb{K}^{n\times n}$ invertierbar, so gilt $\det(A^{-1}) = (\det(A))^{-1}$.

Beweis.

1. Eine Matrix wird mit einem Faktor λ multipliziert indem man jede Zeile mit diesem Faktor multipliziert. Die n-fache Anwendung von Eigenschaft (D1) (mit $\mu = 0$) liefert also sofort (E1).
2. Wählt man $\lambda = \mu = 0$ in Eigenschaft (D1), so ergibt sich sofort (E2):

$$\det \begin{pmatrix} \vdots \\ 0 \cdot z_k \\ \vdots \end{pmatrix} = 0 \cdot \det \begin{pmatrix} \vdots \\ z_k \\ \vdots \end{pmatrix} = 0.$$

3. Nehmen wir an, dass B die Matrix ist, bei der Zeile i und Zeile j von A vertauscht wurden mit $i < j$. Bezeichnen wir mit $z_1, \ldots, z_n$ die Zeilen von A. Dann gilt

$$\begin{aligned}
\det(A) + \det(B) &= \det \begin{pmatrix} \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ z_j \\ \vdots \\ z_i \\ \vdots \end{pmatrix} \\
&\overset{(D2)}{=} \det \begin{pmatrix} \vdots \\ z_i \\ \vdots \\ z_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ z_j \\ \vdots \\ z_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ z_j \\ \vdots \\ z_j \\ \vdots \end{pmatrix} \\
&\overset{(D1)}{=} \det \begin{pmatrix} \vdots \\ z_i \\ \vdots \\ z_i + z_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ z_j \\ \vdots \\ z_i + z_j \\ \vdots \end{pmatrix} \overset{(D1)}{=} \det \begin{pmatrix} \vdots \\ z_i + z_j \\ \vdots \\ z_i + z_j \\ \vdots \end{pmatrix} \overset{(D2)}{=} 0.
\end{aligned}$$

4. Eigenschaft (E4) zu zeigen ist Teil von Aufgabe 8.7.
5. Sind alle Einträge der Hauptdiagonale von Null verschieden, d.h. $\alpha_{k,k} \neq 0$, $\forall k \in \{1, \ldots, n\}$, so gilt durch wiederholte Anwendung von (E4) gemäß dem Gauß-Algorithmus

$$\det(A) = \det \begin{pmatrix} \alpha_{1,1} & & \\ & \ddots & \\ & & \alpha_{n,n} \end{pmatrix} \overset{(D1)}{=} \alpha_{1,1} \cdot \ldots \cdot \alpha_{n,n} \cdot \det(E_n) \overset{(D3)}{=} \alpha_{1,1} \cdot \ldots \cdot \alpha_{n,n}.$$

   Falls nicht alle Diagonaleinträge von Null verschieden ist, so wählen wir den Index i derart, dass $\alpha_{i,i} = 0$ ist, jedoch $\alpha_{i+1,i+1}, \ldots, \alpha_{n,n} \neq 0$ sind. Dank der anfänglichen Treppenstruktur können wir wieder mit dem Gauß-Algorithmus erreichen, dass $\alpha_{i,i} = \alpha_{i,i+1} = \ldots = \alpha_{i,n} = 0$ ist, wobei sich dank (E4) die Determinante nicht verändert hat. Nun hat die Matrix eine Nullzeile, d.h. gemäß (E2) ist die Determinante Null, so dass auch $\det(A) = 0$ ist.

6. Durch die in (E3) und (E4) beschriebenen elementaren Zeilenumformungen kann man eine beliebige Matrix in die Zeilenstufenform bringen. In einem ersten Schritt wenden wir diese auf A_1 an, so dass dieser Block der Matrix zu einer Zeilenstufenform – nennen wir diese Matrix B_1 – wird. Dabei bleibt A_2 unverändert und aus C werde C'. Wurden k Zeilenvertauschungen vorgenommen, so gilt

$$\det(A_1) = (-1)^k \det(B_1). \tag{8.2.1}$$

Nun bringen wir den Block A_2 mit Hilfe von (E3) und (E4) in die Zeilenstufenform – nennen wir diese neue Matrix B_2. Dabei bleiben A_1 und C' unverändert und es gilt bei l Zeilenvertauschungen

$$\det(A_2) = (-1)^l \det(B_2). \tag{8.2.2}$$

Nun haben wir die neue Matrix

$$B = \begin{pmatrix} B_1 & C' \\ O & B_2 \end{pmatrix},$$

wobei B_1 und B_2 in Zeilenstufenform dastehen, so dass B insgesamt eine Matrix in Zeilenstufenform ist. In der Hauptdiagonale von B stehen gerade alle Einträge der Hauptdiagonalen von B_1 und B_2, so dass nach (E5)

$$\det(B) = \det(B_1) \cdot \det(B_2)$$

gilt. Andererseits ist

$$\det(A) = (-1)^{k+l} \det(B), \tag{8.2.3}$$

da B durch $k + l$ Zeilenvertauschungen und Addition von Zeilen entstanden ist. Aus (8.2.1), (8.2.2) und (8.2.3) folgt

$$\det(A) = \det(A_1) \cdot \det(A_2).$$

7. Eigenschaft (E7) zu zeigen ist Teil von Aufgabe 8.7.
8. Es gibt mehrere Möglichkeiten, diese Aussage zu zeigen. Wir werden zunächst einen anderen Satz über Determinanten beweisen und anschließend mit dessen Hilfe Eigenschaft (E8) zeigen.
9. Eigenschaft (E9) zu zeigen ist Teil von Aufgabe 8.7.

□

Wir haben gesehen, dass man aus den definierenden Eigenschaften (D1), (D2) und (D3) viele weitere Eigenschaften von Determinanten folgern kann. Zumindest für die Sonderfälle $n = 2$ und $n = 3$ haben wir auch schon Abbildungen gesehen, die Determinanten sind. Es bleibt die Frage, ob dies die einzigen Determinanten sind. Mit anderen Worten: Gibt es verschiedene Abbildungen, die die Bedingungen (D1), (D2) und (D3) erfüllen? Der folgende Satz beantwortet diese Frage.

Satz 8.17: Eindeutigkeit der Determinante

Seien $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gibt es höchstens eine Abbildung $\det \in \mathrm{Abb}(\mathbb{K}^{n\times n}; \mathbb{K})$, die die Bedingungen (D1), (D2) und (D3) aus Definition 8.12 erfüllt. Die Determinante wird also durch (D1), (D2) und (D3) eindeutig bestimmt.

Beweis. Seien $\det_1, \det_2 \in \mathrm{Abb}(\mathbb{K}^{n\times n}; \mathbb{K})$ zwei Abbildungen, die beide (D1), (D2) und (D3) erfüllen und sei $f\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ definiert über $f(A) := \det_1(A) - \det_2(A)$. Dann erfüllt f ebenfalls (D1) und (D2):

- f ist linear in jeder Zeile, d.h. für jeden Index $k \in \{1,\dots,n\}$ gilt:
  Ist $z_k = \lambda z_k' + \mu z_k''$ für $\lambda, \mu \in \mathbb{K}$ und Zeilenvektoren $z_k', z_k'' \in \mathbb{K}^{1\times n}$, so ist

$$\begin{aligned} f\begin{pmatrix} \vdots \\ z_k \\ \vdots \end{pmatrix} &= \det_1\begin{pmatrix} \vdots \\ z_k \\ \vdots \end{pmatrix} - \det_2\begin{pmatrix} \vdots \\ z_k \\ \vdots \end{pmatrix} \\ &\overset{(D1)}{=} \lambda\det_1\begin{pmatrix} \vdots \\ z_k' \\ \vdots \end{pmatrix} + \mu\det_1\begin{pmatrix} \vdots \\ z_k'' \\ \vdots \end{pmatrix} - \lambda\det_2\begin{pmatrix} \vdots \\ z_k' \\ \vdots \end{pmatrix} - \mu\det_2\begin{pmatrix} \vdots \\ z_k'' \\ \vdots \end{pmatrix} \\ &= \lambda(\det_1 - \det_2)\begin{pmatrix} \vdots \\ z_k' \\ \vdots \end{pmatrix} + \mu(\det_1 - \det_2)\begin{pmatrix} \vdots \\ z_k'' \\ \vdots \end{pmatrix} \\ &= \lambda f\begin{pmatrix} \vdots \\ z_k' \\ \vdots \end{pmatrix} + \mu f\begin{pmatrix} \vdots \\ z_k'' \\ \vdots \end{pmatrix}. \end{aligned}$$

- f ist alternierend: Hat A zwei gleiche Zeilen, so ist wegen (D2) $\det_1(A) = \det_2(A) = 0$, also ist $f(A) = \det_1(A) - \det_2(A) = 0$.

Zudem gilt durch Anwendung von (D3):

$$f(E_n) = \det_1(E_n) - \det_2(E_n) = 1 - 1 = 0.$$

Sei nun $A = (\alpha_{k,j})_{k,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$ eine beliebige Matrix mit Zeilenvektoren $z_1,\dots,z_n$ und seien $e_1,\dots,e_n$ die hier ausnahmsweise als Zeilenvektor interpretierten Einheitsvektoren. Dann gilt:

$$A = \begin{pmatrix} - & z_1 & - \\ - & z_2 & - \\ & \vdots & \\ - & z_n & - \end{pmatrix} \quad \text{und} \quad E_n = \begin{pmatrix} - & e_1 & - \\ - & e_2 & - \\ & \vdots & \\ - & e_n & - \end{pmatrix}.$$

Wir wollen zeigen, dass $f(A) = 0$ ist.

Für jedes $k \in \{1,\dots,n\}$ ist

$$z_k = \sum_{j=1}^{n} \alpha_{k,j} e_j.$$

Betrachten wir als ersten Schritt die Matrix

$$B_1 := \begin{pmatrix} z_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}\cdot e_1 + \alpha_{1,2}\cdot e_2 + \ldots + \alpha_{1,n}\cdot e_n \\ e_2 \\ \vdots \\ e_n \end{pmatrix}.$$

Da alle Zeilen bis auf die erste Zeile gleich sind, können wir die (D1) anwenden, d.h.

$$\begin{aligned}\det_1(B_1) &\overset{\text{(D1)}}{=} \alpha_{1,1}\cdot\det_1\begin{pmatrix}e_1\\e_2\\\vdots\\e_n\end{pmatrix} + \alpha_{1,2}\cdot\det_1\begin{pmatrix}e_2\\e_2\\\vdots\\e_n\end{pmatrix} + \alpha_{1,n}\cdot\det_1\begin{pmatrix}e_n\\e_2\\\vdots\\e_n\end{pmatrix}\\ &\overset{\text{(D2)}}{=} \alpha_{1,1}\cdot 1 + \alpha_{1,2}\cdot 0 + \ldots + \alpha_{n,n}\cdot 0\\ &= \alpha_{1,1}\end{aligned}$$

und analog $\det_2(B_1) = \alpha_{1,1}$, also insgesamt $f(B_1) = 0$. Im nächsten Schritt betrachten wir die Matrix

$$B_2 := \begin{pmatrix}z_1\\z_2\\e_3\\\vdots\\e_n\end{pmatrix} = \begin{pmatrix}z_1\\ \alpha_{2,1}\cdot e_1 + \alpha_{2,2}\cdot e_2 + \ldots + \alpha_{2,n}\cdot e_n\\ e_3\\\vdots\\e_n\end{pmatrix}$$

Wieder gilt nach Anwendung von (D1) und (D2) und für den ersten Vektor mit einer analogen Überlegung zu eben

$$\begin{aligned}\det_1(B_2) &= \alpha_{2,1}\cdot\det_1\begin{pmatrix}z_1\\e_1\\e_3\\\vdots\\e_n\end{pmatrix} + \alpha_{2,2}\cdot\det_1\begin{pmatrix}z_1\\e_2\\e_3\\\vdots\\e_n\end{pmatrix} + \ldots + \alpha_{2,n}\cdot\det_1\begin{pmatrix}z_1\\e_n\\e_3\\\vdots\\e_n\end{pmatrix}\\ &= \alpha_{2,1}\cdot\alpha_{1,2} + \alpha_{2,2}\cdot\det_1(B_1) + \alpha_{2,3}\cdot 0 + \ldots + \alpha_{n,n}\cdot 0\\ &= \alpha_{2,1}\cdot\alpha_{1,2} + \alpha_{2,2}\cdot\alpha_{1,1}\\ &= \det_2(B_2).\end{aligned}$$

Folglich ist wieder $f(B_2) = 0$. Führen wir dieses Verfahren fort, so ist $f(A) = f(B_n) = 0$, d.h. für jede beliebige Matrix $A \in \mathbb{K}^{n\times n}$ ist $f(A) = 0$. Nach Definition von f ist also $\det_1(A) = \det_2(A)$ für alle $A \in \mathbb{K}^{n\times n}$, was gerade der Behauptung des Satzes entspricht. □

Beispiel 8.18 (Determinante der Summe von Matrizen)**.** *Der Beweis des vorherigen Satzes lässt die Vermutung aufkommen, dass die Determinante der Summe zweier Matrizen gerade die Summe der Determinanten sei. Dies ist im Allgemeinen leider nicht der Fall, wie wir anhand eines einfachen Beispiels zeigen können:*

$$1 = \det\begin{pmatrix}1&0\\0&1\end{pmatrix} = \det\left(\begin{pmatrix}1&0\\0&0\end{pmatrix} + \begin{pmatrix}0&0\\0&1\end{pmatrix}\right) \neq \det\begin{pmatrix}1&0\\0&0\end{pmatrix} + \det\begin{pmatrix}0&0\\0&1\end{pmatrix} \overset{\text{(E2)}}{=} 0 + 0 = 0.$$

Beweis von Eigenschaft (E8) aus Satz 8.16. Seien $A, B \in \mathbb{K}^{n\times n}$. Wir wollen zeigen:

$$\det(A\cdot B) = \det(A)\cdot\det(B). \tag{8.2.4}$$

Nehmen wir zunächst an, es sei $\det(B) = 0$. Nach (E7) ist in diesem Fall $\operatorname{rg}(B) < n$ und mit der Formel aus Satz 6.73 folgt $\dim(L(B, 0_n)) = n - \operatorname{rg}(B) > 0$, d.h. das homogene LGS $B \cdot x = 0_n$ hat eine nichttriviale[3] Lösung $x^* \neq 0_n \in \mathbb{K}^n$. Dann ist auch $x^* \in L(A \cdot B, 0_n)$, denn mit dem Assoziativgesetz der Matrizenmultiplikation (Gleichung (2.3.1) in Satz 2.19) gilt

$$(A \cdot B) \cdot x^* = A \cdot (B \cdot x^*) = A \cdot 0_n = 0_n \in \mathbb{K}^n.$$

Da $L(A \cdot B, 0_n)$ mindestens eine nichttriviale Lösung enthält, ist $\dim(L(A \cdot B, 0_n)) > 0$. Durch nochmalige Anwendung von Satz 6.73 folgt

$$\mathrm{rg}(A \cdot B) = n - \dim(L(A \cdot B, 0_n)) < n.$$

Gemäß (E7) ist also auch $\det(A \cdot B) = 0$, was (8.2.4) für diesen Fall beweist.

Nehmen wir nun an, es sei $\det(B) \neq 0$. Für festes B definieren wir die Abbildung

$$f \colon \mathbb{K}^{n \times n} \to \mathbb{K}, \qquad f(A) := \frac{\det(A \cdot B)}{\det(B)}.$$

Wenn wir zeigen können, dass f die Bedingungen (D1), (D2) und (D3) erfüllt, so ist f eine Determinante. Da gemäß Satz 8.17 die Determinante eindeutig bestimmt ist, folgt sofort $f(A) = \det(A)$ und damit die gewünschte Formel (8.2.4).

f **erfüllt (D1):** Seien $A = (\alpha_{k,i})$ und $B = (\beta_{i,j})$. Dann ist $A \cdot B = C = (\gamma_{k,j})_{1 \leq k,j \leq n}$ gegeben durch

$$\gamma_{k,j} = \sum_{i=1}^{n} \alpha_{k,i}\beta_{i,j}, \quad k, j \in \{1, \ldots, n\}.$$

Zudem gelte für einen festen Index $k^* \in \{1, \ldots, n\}$

$$\alpha_{k^*,i} = \lambda \alpha'_{k^*,i} + \mu \alpha''_{k^*,i}, \qquad \forall i \in \{1, \ldots, n\}.$$

Bezeichnen wir mit A' und A'' die Matrizen, die aus A entstehen, wenn $\alpha_{k^*,i}$ durch $\alpha'_{k^*,i}$ bzw. durch $\alpha''_{k^*,i}$ ersetzt wird. Mit diesen Bezeichnungen wollen wir zeigen:

$$f(A) = \lambda f(A') + \mu f(A'').$$

Setzen wir nach und nach alle Voraussetzungen ein, so erhalten wir:

$$\begin{aligned}
f(A) = \frac{\det(A \cdot B)}{\det(B)} = \frac{\det(C)}{\det(B)} &= \frac{1}{\det(B)} \cdot \det \begin{pmatrix} \vdots \\ (\gamma_{k^*,j})_{j \in \{1,\ldots,n\}} \\ \vdots \end{pmatrix} \\
&= \frac{1}{\det(B)} \cdot \det \begin{pmatrix} \vdots \\ \left(\sum_{i=1}^{n} \alpha_{k^*,i}\beta_{i,j}\right)_{j \in \{1,\ldots,n\}} \\ \vdots \end{pmatrix} \\
&= \frac{1}{\det(B)} \cdot \det \begin{pmatrix} \vdots \\ \left(\sum_{i=1}^{n} (\lambda \alpha'_{k^*,i} + \mu \alpha''_{k^*,i})\beta_{i,j}\right)_{j \in \{1,\ldots,n\}} \\ \vdots \end{pmatrix} \\
&= \frac{1}{\det(B)} \cdot \det \begin{pmatrix} \vdots \\ \left(\lambda \sum_{i=1}^{n} \alpha'_{k^*,i}\beta_{i,j} + \mu \sum_{i=1}^{n} \alpha''_{k^*,i}\beta_{i,j}\right)_{j \in \{1,\ldots,n\}} \\ \vdots \end{pmatrix}
\end{aligned}$$

[3] Da jedes homogene LGS den Nullvektor als (nicht notwendigerweise einzige) Lösung besitzt, bezeichnet man den Nullvektor in diesem Kontext als die *triviale Lösung* des homogenen LGS.

Definieren wir nun die Matrizen C' und C'' als die Matrix C, in deren k^*-ter Zeile stattdessen Folgendes steht:

$$\gamma'_{k^*,j} := \sum_{i=1}^{n} \alpha'_{k^*,i}\beta_{i,j}, \qquad \gamma''_{k^*,j} := \sum_{i=1}^{n} \alpha''_{k^*,i}\beta_{i,j}.$$

Es ist gerade $C' = A' \cdot B$ und $C'' = A'' \cdot B$. Mit Eigenschaft (D1) der Determinante erhalten wir somit

$$\begin{aligned}
f(A) &= \frac{1}{\det(B)} \cdot \det \begin{pmatrix} \vdots \\ \left(\lambda\gamma'_{k^*,j} + \mu\gamma''_{k^*,j}\right)_{j\in\{1,\dots,n\}} \\ \vdots \end{pmatrix} \\
&= \frac{1}{\det(B)} \cdot [\lambda \det(C') + \mu \det(C'')] = \frac{1}{\det(B)} \cdot [\lambda \det(A' \cdot B) + \mu \det(A'' \cdot B)] \\
&= \lambda f(A') + \mu f(A'').
\end{aligned}$$

f **erfüllt (D2):** Seien wieder $A = (\alpha_{k,i})$, $B = (\beta_{i,j})$ und $A \cdot B = C = (\gamma_{k,j})_{1\le k,j\le n}$. Sind Zeile k und Zeile l von A identisch, so ist $\alpha_{k,i} = \alpha_{l,i}$ für alle $i \in \{1, \dots, n\}$. Folglich ist $\alpha_{k,i}\beta_{i,j} = \alpha_{l,i}\beta_{i,j}$ für alle $i, j \in \{1, \dots, n\}$. Summieren wir über $i \in \{1, \dots, n\}$, so erhalten wir:

$$\gamma_{k,j} = \gamma_{l,j}, \quad \forall j \in \{1, \dots, n\},$$

d.h. die k-te und j-te Zeile von $C = A\cdot B$ sind identisch. Mit (D2) folgt $\det(A\cdot B) = 0$, also ist $f(A) = 0$.

f **erfüllt (D3):** Es ist

$$f(E_n) = \frac{\det(E_n \cdot B)}{\det(B)} = \frac{\det(B)}{\det(B)} = 1.$$

□

8.3 Spaltenumformungen und Determinanten

In diesem Abschnitt wollen wir mit Hilfe der Produktformel nachweisen, dass alle Aussagen, die über Zeilenumformungen und deren Auswirkung auf die Determinante gemacht wurden, auch für die analogen Spaltenumformungen gültig sind.

Beispiel 8.19. *Seien*

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad \textit{und} \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dann gilt $\det(B) = -1$ *und somit nach Anwendung der Produktformel aus (E8):*

$$\det(B \cdot A) = \det \begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix} = \det(B) \cdot \det(A) = -\det(A),$$

$$\det(A \cdot B) = \det \begin{pmatrix} b & a & c \\ e & d & f \\ h & g & i \end{pmatrix} = \det(A) \cdot \det(B) = -\det(A).$$

Eine Multiplikation mit B von links bewirkt eine Vertauschung der ersten zwei Zeilen der Matrix und in der Tat ändert sich (wie in (E3) behauptet) nur das Vorzeichen der Determinante. Eine Multiplikation mit B von rechts bewirkt eine Vertauschung der ersten zwei Spalten der Matrix und wieder ändert sich nur das Vorzeichen.

Formulieren wir im nun analoge Aussagen zu (E2) bis (E5) aus Satz 8.16 für Spaltenumformungen:

Satz 8.20

Für einen Körper $\mathbb{K}$ und $n \in \mathbb{N}$ sei $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Determinante. Dann hat $\det$ folgende Eigenschaften:

(E2') Ist eine Spalte von $A \in \mathbb{K}^{n\times n}$ der Nullvektor, so ist $\det(A) = 0$.

(E3') Seien $A, B \in \mathbb{K}^{n\times n}$. Ist B aus A durch Vertauschung zweier Spalten entstanden, so ist $\det(B) = -\det(A)$.

(E4') Seien $\lambda \in \mathbb{K}$ und $A \in \mathbb{K}^{n\times n}$. Ist $A' \in \mathbb{K}^{n\times n}$ die Matrix, die entsteht, wenn man das λ-Fache der i-ten Spalte von A zur j-ten Spalte von A hinzu addiert (mit $i \neq j$), so ist $\det(A') = \det(A)$.

(E5') Liegt $A \in \mathbb{K}^{n\times n}$ in der unteren Zeilenstufenform vor, d.h. gilt $A = \begin{pmatrix} \alpha_{1,1} & & 0 \\ \vdots & \ddots & \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix}$, so ist $\det(A) = \alpha_{1,1} \cdot \ldots \cdot \alpha_{n,n}$.

Beweis.

(E2') Hat $A \in \mathbb{K}^{n\times n}$ eine Nullspalte, so ist $\operatorname{rg}(A) < n$. Nach Eigenschaft (E7) ist folglich $\det(A) = 0$.

(E3') Sei $E_n^{i,j}$ die Matrix, die aus der $n \times n$-Einheitsmatrix entstanden ist durch Vertauschen der Zeilen i und j. Dann ist zunächst einmal nach (E3) und (D3)

$$\det(E_n^{i,j}) = -\det(E_n) = -1.$$

Die Multiplikation von $A \in \mathbb{K}^{n\times n}$ von rechts mit $E_n^{i,j}$ bewirkt die Vertauschung der i-ten und j-ten Spalte von A und es gilt nach der Produktformel für Determinanten aus (E8):

$$\det(A \cdot E_n^{i,j}) = \det(A) \cdot \det(E_n^{i,j}) = -\det(A).$$

(E4') Seien $\lambda \in \mathbb{K}$; $i^*, j^* \in \{1, \ldots, n\}$ mit $i^* \neq j^*$ und $A = (\alpha_{k,i}), B = (\beta_{i,j}) \in \mathbb{K}^{n\times n}$ mit

$$\beta_{i,j} = \mathbb{1}_{\{i=j\}} + \lambda \mathbb{1}_{\{(i,j)=(i^*,j^*)\}},$$

d.h. B ist die Einheitsmatrix, bei der zusätzlich der Eintrag in der i^*-ten Zeile und j^*-ten Spalte auf λ gesetzt wurde. Es ist $\det(B) = 1$, was man beispielsweise mit

(D1) zeigen kann. Das Produkt $A \cdot B =: (\gamma_{k,j}) \in \mathbb{K}^{n\times n}$ hat die Einträge

$$\begin{aligned}\gamma_{k,j} &= \sum_{i=1}^{n} \alpha_{k,i}\beta_{i,j} = \sum_{i=1}^{n} \alpha_{k,i} \left(\mathbb{1}_{\{i=j\}} + \lambda \mathbb{1}_{\{(i,j)=(i^*,j^*)\}}\right) \\ &= \begin{cases} \alpha_{k,j}, & \text{falls } j \neq j^*, \\ \alpha_{k,j} + \lambda \alpha_{i^*,j^*}, & \text{falls } j = j^*. \end{cases}\end{aligned}$$

Es ist also $A \cdot B$ die Matrix, die aus A entsteht, wenn zur j^*-ten Spalte das λ-Fache der i^*-ten Spalte addiert wird. Mit der Produktformel aus (E8) schlussfolgern wir wieder

$$\det(A \cdot B) = \det(A) \cdot \det(B) = \det(A) \cdot 1 = \det(A).$$

(E5') Der Beweis von (E5) kann komplett übernommen werden, wobei an Stelle von (E2) und (E4) nun (E2') und (E4') verwendet werden.

□

Der Beweis von (E4') kann durch leichte Modifikation verwendet werden um zu zeigen, dass die Multiplikation einer Spalte mit einem Faktor $\lambda \in \mathbb{K}$ zur Vervielfachung der Determinante um diesen Faktor führt.

Lemma 8.21

Seien $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$, $\lambda \in \mathbb{K}$ und $j^* \in \{1, \ldots, n\}$. Entsteht A' aus $A \in \mathbb{K}^{n\times n}$ indem die j^*-te Spalte von A mit λ multipliziert wird, so ist $\det(A') = \lambda \cdot \det(A)$.

Beweis. Seien $A = (\alpha_{k,i}), B = (\beta_{i,j}) \in \mathbb{K}^{n\times n}$ mit

$$\beta_{i,j} = \mathbb{1}_{\{i=j\neq j^*\}} + \lambda \mathbb{1}_{\{i=j=j^*\}},$$

d.h. B ist die Einheitsmatrix, bei der der Eintrag in der j^*-ten Zeile auf λ gesetzt wurde. Mit (D1) ist $\det(B) = \lambda$. Das Produkt $A \cdot B =: (\gamma_{k,j}) \in \mathbb{K}^{n\times n}$ hat die Einträge

$$\begin{aligned}\gamma_{k,j} &= \sum_{i=1}^{n} \alpha_{k,i}\beta_{i,j} = \sum_{i=1}^{n} \alpha_{k,i} \left(\mathbb{1}_{\{i=j\neq j^*\}} + \lambda \mathbb{1}_{\{i=j=j^*\}}\right) \\ &= \begin{cases} \alpha_{k,j}, & \text{falls } j \neq j^*, \\ \lambda\alpha_{k,j}, & \text{falls } j = j^*. \end{cases}\end{aligned}$$

Es ist also $A \cdot B$ die Matrix, die aus A entsteht, wenn die j^*-te Spalte mit λ multipliziert wird. Mit der Produktformel aus (E8) schlussfolgern wir wieder

$$\det(A \cdot B) = \det(A) \cdot \det(B) = \det(A) \cdot \lambda = \lambda \det(A).$$

□

Anhand eines Beispiels wollen wir einerseits exemplarisch nachprüfen, dass das Transponieren die Determinante nicht ändert und andererseits illustrieren, wie wir das zeigen können.

Beispiel 8.22. *Betrachten wir die Matrizen*

$$A = \begin{pmatrix} 3 & 1 & 2 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \textit{und} \quad A^T = \begin{pmatrix} 3 & 0 & 1 \\ 1 & 4 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Es gilt $\det(A) = 12 - 8 = 4 = \det(A^T)$. *Im Folgenden wollen wir das mit Hilfe elementarer Zeilen- und Spaltenumformungen herausbekommen.*

$$\begin{aligned}
\det \begin{pmatrix} 3 & 1 & 2 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix} &\overset{\text{(E4),(D1)}}{=} -\det \begin{pmatrix} 3 & 1 & 2 \\ 0 & 4 & 0 \\ 0 & 1/3 & -1/3 \end{pmatrix} && \mid \textit{durch } \tfrac{1}{3}Z_1 + (-1)\cdot Z_3 \rightsquigarrow Z_3 \\
&\overset{\text{(D1)}}{=} -4 \cdot \frac{1}{3} \cdot \det \begin{pmatrix} 3 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} && \mid \textit{durch } 3Z_3 \rightsquigarrow Z_3 \textit{ und } \tfrac{1}{4}Z_2 \rightsquigarrow Z_2 \\
&\overset{\text{(E4),(D1)}}{=} \frac{4}{3} \det \begin{pmatrix} 3 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} && \mid \textit{durch } Z_2 + (-1)\cdot Z_3 \rightsquigarrow Z_3 \\
&\overset{\text{(E5)}}{=} \frac{4}{3} \cdot 3 = 4.
\end{aligned}$$

Wenden wir die analogen Spaltenumformungen auf A^T *an:*

$$\begin{aligned}
\det \begin{pmatrix} 3 & 0 & 1 \\ 1 & 4 & 0 \\ 2 & 0 & 1 \end{pmatrix} &\overset{\text{(E4'),Lemma}}{=} -\det \begin{pmatrix} 3 & 0 & 0 \\ 1 & 4 & 1/3 \\ 2 & 0 & -1/3 \end{pmatrix} && \mid \textit{durch } \tfrac{1}{3}S_1 + (-1)\cdot S_3 \rightsquigarrow S_3 \\
&\overset{\text{Lemma}}{=} -4 \cdot \frac{1}{3} \cdot \det \begin{pmatrix} 3 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & -1 \end{pmatrix} && \mid \textit{durch } 3S_3 \rightsquigarrow S_3 \textit{ und } \tfrac{1}{4}S_2 \rightsquigarrow S_2 \\
&\overset{\text{(E4'),Lemma}}{=} \frac{4}{3} \det \begin{pmatrix} 3 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} && \mid \textit{durch } S_2 + (-1)\cdot S_3 \rightsquigarrow S_3 \\
&\overset{\text{(E5')}}{=} \frac{4}{3} \cdot 3 = 4.
\end{aligned}$$

Korollar 8.23

Für $n \in \mathbb{N}$ und einen Körper $\mathbb{K}$ sei $A \in \mathbb{K}^{n\times n}$. Dann gilt

$$\det(A) = \det(A^T),$$

d.h. durch das Transponieren ändert sich die Determinante einer Matrix nicht.

Beweis. Wendet man den Gauß-Algorithmus auf A an um mit Hilfe elementarer Zeilenumformungen die Matrix auf die obere Zeilenstufenform B zu bringen, so führen die analogen elementaren Spaltenumformungen dazu, dass die transponierte Matrix A^T auf die untere Zeilenstufenform C gebracht wird, wobei $B^T = C$ gilt. In beiden Fällen kann die Determinante nun als Produkt der Diagonaleinträge berechnet werden, so dass $\det(B) = \det(C)$ gilt. Da (E2) und (E2') bzw. (E3) und (E3') die gleiche Wirkung auf die Determinante haben, gilt auch $\det(A) = \det(A^T)$. $\square$

Bemerkung 8.24. *Man kann auch zeigen, dass jede Determinante analog zu (D1) linear in den Spalten der Matrix ist. Den Nachweis ersparen wir uns an dieser Stelle.*

8.4 Entwicklungssatz von Laplace

Bisher können wir problemlos Determinanten von 2×2- und 3×3-Matrizen berechnen. Ziel dieses Abschnittes ist es, Determinanten von höherdimensionalen Matrizen auf die Determinanten von Matrizen niedrigerer Dimension zurückzuführen. Sehen wir uns dazu die bisher bekannten Dimensionen an:

$n = 1$: Gemäß (D3) ist $\det(1) = 1$ und laut (D1) ist folglich $\det(\lambda) = \lambda$ für jedes $\lambda \in \mathbb{K}$. Die Determinante einer 1×1-Matrix, d.h. einer Zahl, ist diese Zahl selbst.

$n = 2$: Es ist

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc = a \cdot \det(d) - c \cdot \det(b).$$

$n = 3$: Nach der Regel von Sarrus ist

$$\begin{aligned} \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= aei + dhc + gbf - gec - ahf - dbi \\ &= a(ei - hf) - d(bi - hc) + g(bf - ec) \\ &= a \cdot \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - d \cdot \det \begin{pmatrix} b & c \\ h & i \end{pmatrix} + g \cdot \det \begin{pmatrix} b & c \\ e & f \end{pmatrix}. \end{aligned}$$

Wir haben jeweils ein Element der ersten Spalte $\alpha_{k,1}$ mit der Determinante der Matrix multipliziert, die entsteht, wenn man die k-te Zeile und 1. Spalte streicht. Um dieses Verfahren zu verallgemeinern, benötigen wir folgende Definition:

Definition 8.25

Sei $A = (\alpha_{k,j})_{k,j \in \{1,\dots,n\}} \in \mathbb{K}^{n \times n}$. Für $k^*, j^* \in \{1, \dots, n\}$ definieren wir die Matrix $A_{k^*,j^*} \in \mathbb{K}^{(n-1)\times(n-1)}$ als

$$A_{k^*,j^*} = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,j^*-1} & \alpha_{1,j^*+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{k^*-1,1} & \cdots & \alpha_{k^*-1,j^*-1} & \alpha_{k^*-1,j^*+1} & \cdots & \alpha_{k^*-1,n} \\ \alpha_{k^*+1,1} & \cdots & \alpha_{k^*+1,j^*-1} & \alpha_{k^*+1,j^*+1} & \cdots & \alpha_{k^*+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,j^*-1} & \alpha_{n,j^*+1} & \cdots & \alpha_{n,n} \end{pmatrix},$$

d.h. A_{k^*,j^*} ist die Matrix, die aus A entsteht, wenn man die k^*-te Zeile und die j^*-te Spalte aus A streicht. Diese Matrix nennt man *Minor* oder *Untermatrix* von A.

Beispiel 8.26.

Zu $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$ *ist beispielsweise* $\alpha_{2,1} = 5$ *und* $A_{2,1} = \begin{pmatrix} 2 & 3 & 4 \\ 10 & 11 & 12 \\ 14 & 15 & 16 \end{pmatrix}$.

Satz 8.27: Entwicklungssatz von Laplace

Seien $\mathbb{K}$ ein Körper, $n \geq 2$ eine natürliche Zahl und $A = (\alpha_{k,j}) \in \mathbb{K}^{n \times n}$. Dann gilt für jedes $k \in \{1, \ldots, n\}$ folgende Formel für die *Entwicklung nach der k-ten Zeile*:

$$\det(A) = \sum_{j=1}^{n} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}).$$

Zudem gilt für jedes $j \in \{1, \ldots, n\}$ folgende Formel für die *Entwicklung nach der j-ten Spalte*:

$$\det(A) = \sum_{k=1}^{n} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}).$$

Bemerkung 8.28. *Die Entwicklungsformel hängt formal von der gewählten Spalte/Zeile ab. Da wir jedoch bereits wissen, dass die Determinante eindeutig bestimmt ist, ergebt jede beliebige Wahl der Spalte/Zeile die gleiche Abbildung, nämlich* die *Determinante.*

Notation

Bei der Anwendung des Entwicklungssatzes werden wir viele Determinanten schreiben müssen. Dafür gibt es eine verkürzte Schreibweise:

$$\det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \cdots & \alpha_{2,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} = \begin{vmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \cdots & \alpha_{2,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{vmatrix}$$

Einzig für 1×1-Matrizen ist diese Notation wegen der Verwechslungsgefahr mit dem Betrag nicht zu empfehlen, denn $\det(-5) = -5$, wohingegen der Betrag $|-5| = 5$ ist.

Beispiel 8.29. *Bevor wir den Satz beweisen, wollen wir anhand eines Beispiels illustrieren, dass die Formel tatsächlich eine Determinante liefert. Sehen wir uns speziell die Entwicklung nach der 1. Spalte an für Matrizen aus dem* $\mathbb{R}^{4 \times 4}$.

(D1) *Für* $\lambda, \mu \in \mathbb{R}$ *betrachten wir die Determinante von*

$$A = \begin{pmatrix} a & x_1 & x_2 & x_3 \\ \lambda u_0 + \mu v_0 & \lambda u_1 + \mu v_1 & \lambda u_2 + \mu v_2 & \lambda u_3 + \mu v_3 \\ b & y_1 & y_2 & y_3 \\ c & z_1 & z_2 & z_3 \end{pmatrix}.$$

Wir kürzen die 3*-dimensionalen Zeilenvektoren wie folgt ab:* $x = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$ *und analog* y*,* z*,* u *und* v*. Damit ist*

$$\det(A) = a \cdot \begin{vmatrix} \lambda u + \mu v \\ y \\ z \end{vmatrix} - (\lambda u_0 + \mu v_o) \cdot \begin{vmatrix} x \\ y \\ z \end{vmatrix} + b \cdot \begin{vmatrix} x \\ \lambda u + \mu v \\ z \end{vmatrix} - c \cdot \begin{vmatrix} x \\ \lambda u + \mu v \\ y \end{vmatrix}$$

Nehmen wir an, für 3×3*-Matrizen hätten wir bereits gezeigt, dass der Entwicklungssatz tatsächlich eine (bzw. die) Determinante liefert. Dann ist*

$$\begin{aligned} \det(A) &= a\lambda \begin{vmatrix} u \\ y \\ z \end{vmatrix} + a\mu \begin{vmatrix} v \\ y \\ z \end{vmatrix} - (\lambda u_0 + \mu v_o) \begin{vmatrix} x \\ y \\ z \end{vmatrix} + b\lambda \begin{vmatrix} x \\ u \\ z \end{vmatrix} + b\mu \begin{vmatrix} x \\ v \\ z \end{vmatrix} - c\lambda \begin{vmatrix} x \\ u \\ y \end{vmatrix} - c\mu \begin{vmatrix} x \\ v \\ y \end{vmatrix} \\ &= \lambda \cdot \left(a \begin{vmatrix} u \\ y \\ z \end{vmatrix} - u_0 \begin{vmatrix} x \\ y \\ z \end{vmatrix} + b \begin{vmatrix} x \\ u \\ z \end{vmatrix} - c \begin{vmatrix} x \\ u \\ y \end{vmatrix} \right) + \mu \cdot \left(a \begin{vmatrix} v \\ y \\ z \end{vmatrix} - v_0 \begin{vmatrix} x \\ y \\ z \end{vmatrix} + b \begin{vmatrix} x \\ v \\ z \end{vmatrix} - c \begin{vmatrix} x \\ v \\ y \end{vmatrix} \right) \\ &= \lambda \cdot \begin{vmatrix} a & x_1 & x_2 & x_3 \\ u_0 & u_1 & u_2 & u_3 \\ b & y_1 & y_2 & y_3 \\ c & z_1 & z_2 & z_3 \end{vmatrix} + \mu \cdot \begin{vmatrix} a & x_1 & x_2 & x_3 \\ v_0 & v_1 & v_2 & v_3 \\ b & y_1 & y_2 & y_3 \\ c & z_1 & z_2 & z_3 \end{vmatrix} . \end{aligned}$$

Damit ist die Linearität für diese Matrix gezeigt.

(D2) *Zur Illustration wählen wir mehrere Matrizen mit je zwei identischen Zeilen.*

$$A = \begin{pmatrix} a & 1 & 2 & 3 \\ a & 1 & 2 & 3 \\ b & 0 & 4 & 5 \\ c & 0 & 0 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} a & 1 & 2 & 3 \\ b & 0 & 4 & 5 \\ a & 1 & 2 & 3 \\ c & 0 & 0 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} a & 1 & 2 & 3 \\ b & 0 & 4 & 5 \\ c & 0 & 0 & 6 \\ a & 1 & 2 & 3 \end{pmatrix}.$$

Die Determinanten berechnen sich mit Hilfe der Entwicklung nach der 1. Spalte als

$$\det(A) = a \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} - a \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} + b \cdot \underbrace{\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 0 & 0 & 6 \end{vmatrix}}_{=0} - c \cdot \underbrace{\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 2 & 3 \end{vmatrix}}_{=0} = (a - a) \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} = 0,$$

$$\begin{aligned} \det(B) &= a \cdot \begin{vmatrix} 0 & 4 & 5 \\ 1 & 2 & 3 \\ 0 & 0 & 6 \end{vmatrix} - b \cdot \underbrace{\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 0 & 0 & 6 \end{vmatrix}}_{=0} + a \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} - c \cdot \underbrace{\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 2 & 3 \end{vmatrix}}_{=0} \\ &\overset{\text{(E3)}}{=} (-1) \cdot a \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} + a \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} = 0, \end{aligned}$$

$$\begin{aligned} \det(C) &= a \cdot \begin{vmatrix} 0 & 4 & 5 \\ 0 & 0 & 6 \\ 1 & 2 & 3 \end{vmatrix} - b \cdot \underbrace{\begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 1 & 2 & 3 \end{vmatrix}}_{=0} + c \cdot \underbrace{\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 2 & 3 \end{vmatrix}}_{=0} - a \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} \\ &\overset{\text{(E3)}}{=} (-1)^2 \cdot (-a) \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} + a \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} = 0. \end{aligned}$$

(D3) *Es ist*

$$\det\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 1 \cdot \det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \cdot 1 \cdot \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot 1 \cdot 1 \cdot \det(1) = 1.$$

Beweis von Satz 8.27. Da die Determinante einer Matrix eindeutig bestimmt ist, genügt es, Eigenschaften (D1), (D2) und (D3) für die im Entwicklungssatz definierten Abbildungen zu überprüfen. Wir beschränken uns auf die Entwicklung nach der j-ten Spalte für ein festes $j \in \{1, \dots, n\}$. (Die Entwicklung nach der Zeile folgt durch Transposition.) Wir führen den Beweis induktiv.

Induktionsanfang: Für $n = 2$ und $j = 1$ bzw. $j = 2$ besagt die Formel gerade, dass

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (-1)^{1+1} \cdot a \cdot \det(d) + (-1)^{2+1} \cdot c \cdot \det(b) = ad - bc \quad \text{bzw.}$$
$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (-1)^{1+2} \cdot b \cdot \det(c) + (-1)^{2+2} \cdot d \cdot \det(a) = ad - bc$$

ist. Das ist in beiden Fällen die uns aus Definition 8.1 bekannte Determinante.

Induktionsannahme: Nehmen wir nun an, für $n - 1 \geq 2$ sei die Aussage gezeigt.

Induktionsschritt: Wir wollen zeigen, dass die Formel für n ebenfalls korrekt ist. Dazu weisen wir die Eigenschaften (D1), (D2) und (D3) nach.

(D1) Für ein beliebiges (aber fest gewähltes) $i \in \{1, \dots, n\}$ gelte

$$\alpha_{i,j} = \lambda\beta_{i,j} + \mu\gamma_{i,j}, \quad \forall j \in \{1, \dots, n\}.$$

Wir wollen nun die Linearität in der i-ten Zeile nachweisen. Es ist

$$\begin{aligned} \det(A) &= \sum_{k=1}^{n} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}) \\ &= \sum_{k \neq i} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}) + (-1)^{i+j} \alpha_{i,j} \det(A_{i,j}) \\ &= \sum_{k \neq i} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}) + (-1)^{i+j} (\lambda\beta_{i,j} + \mu\gamma_{i,j}) \det(A_{i,j}) \end{aligned}$$

Wenn jeder einzelne Summand linear ist, ist auch die ganze Summe und damit die Determinante linear.

Nach Induktionsannahme ist die Determinante für $n - 1$ linear in jeder Zeile, was wir für die $(n-1) \times (n-1)$-Matrizen $A_{k,j}$ $(k \neq i)$ nutzen können. Nennen wir $B_{k,j}$ (bzw. $C_{k,j}$) die Matrizen, die aus $A_{k,j}$ entstehen, indem die ursprünglich i-te Zeile durch $(\beta_{i,l})_{l \neq j}$ (bzw. $(\gamma_{i,l})_{l \neq j}$) ersetzt wird. Dann ist nach Induktionsannahme

$$\det(A_{k,j}) = \lambda \det(B_{k,j}) + \mu \det(C_{k,j}).$$

Andererseits ist $A_{i,j}$ von der ursprünglich i-ten Zeile von A unabhängig, da diese gestrichen wurde. Die Linearität steht bereits in dem Faktor vor der Determinante.

Somit sind in der Tat alle Summanden linear, was die behauptete Linearität der Determinante in der i-ten Zeile beweist.

(D2) Hat A zwei gleiche Zeilen (Zeile l und Zeile i mit $l < i$), so gilt

$$\alpha_{l,j} = \alpha_{i,j}, \quad \forall j \in \{1, \dots, n\}. \tag{8.4.1}$$

Mit der Festlegung $K := \{1, \dots, n\} \setminus \{l, i\}$ ist

$$\begin{aligned}
\det(A) &= \sum_{k=1}^{n} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}) \\
&= \sum_{k \in K} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}) + (-1)^{l+j} \alpha_{l,j} \det(A_{l,j}) + (-1)^{i+j} \alpha_{i,j} \det(A_{i,j}) \\
&= \sum_{k \in K} (-1)^{k+j} \alpha_{k,j} \det(A_{k,j}) + \alpha_{l,j}((-1)^{l+j} \det(A_{l,j}) + (-1)^{i+j} \det(A_{i,j}))
\end{aligned}$$

Wir müssen nun eine Fallunterscheidung zu den Positionen der Indizes l und i machen.

$i = l + 1$**:** Da die beiden identischen Zeilen direkt untereinander liegen, ist $A_{l,j} = A_{i,j}$ und $(-1)^{i+j} = (-1)^{l+1+j} = (-1) \cdot (-1)^{l+j}$. Folglich ist

$$\alpha_{l,j}((-1)^{l+j} \det(A_{l,j}) + (-1)^{i+j} \det(A_{i,j})) = 0.$$

Zudem gilt für alle $k \in K$, dass die Minoren jeweils zwei identische Zeilen haben und damit deren Determinanten den Wert 0 haben. Insgesamt ist also $\det(A) = 0$.

$i = l + m$ **mit** $m \geq 2$**:** Alle Minoren $A_{k,j}$ mit $k \in K$ haben wieder zwei identische Zeilen, so dass ihre Determinanten verschwinden. Damit ist

$$\det(A) = (-1)^{l+j} \alpha_{l,j} \det(A_{l,j}) + (-1)^{i+j} \alpha_{i,j} \det(A_{i,j}).$$

Die beiden Minoren unterscheiden sich nur bezüglich der Reihenfolge ihrer Zeilen und durch $m - 1$ Zeilenvertauschungen kann der zweite Minor in die Form des ersten gebracht werden, d.h. dank (E3) und unter Beachtung von (8.4.1) ist

$$\begin{aligned}
\det(A) &\overset{\text{(E3)}}{=} (-1)^{l+j} \alpha_{l,j} \det(A_{l,j}) + (-1)^{m-1} \cdot (-1)^{i+j} \alpha_{i,j} \det(A_{l,j}) \\
&\overset{\text{(8.4.1)}}{=} (-1)^{l+j} \alpha_{l,j} \det(A_{l,j}) + (-1)^{m-1} \cdot (-1)^{l+m+j} \alpha_{l,j} \det(A_{l,j}) \\
&= \left[(-1)^{l+j} + (-1)^{2m-1}(-1)^{l+j}\right] \alpha_{l,j} \det(A_{l,j}) = 0.
\end{aligned}$$

(D3) Mit Hilfe der Induktionsannahme gilt direkt

$$\det(E_n) = (-1)^{j+j} \cdot \alpha_{j,j} \cdot \det(E_{n-1}) = \det(E_{n-1}) \overset{\text{Ind.-Ann.}}{=} 1.$$

□

Korollar 8.30: Existenz der Determinante

Für jedes $n \in \mathbb{N}$ und jeden Körper $\mathbb{K}$ existiert eine Determinante $\det\colon \mathbb{K}^{n \times n} \to \mathbb{K}$.

Beweis. Für $n = 1$ ist $\det(\lambda) := \lambda$ eine Determinante und für $n \geq 2$ liefert der Entwicklungssatz von Laplace eine Formel für die Determinante. □

8.5 Aufgaben

Aufgabe 8.1. *Entscheiden Sie, ob die folgenden Berechnungen für Determinanten korrekt sind. Korrigieren Sie die Rechnung gegebenenfalls.*

$$\det\begin{pmatrix}3&0&0\\0&3&0\\0&0&3\end{pmatrix} = 3\cdot\det\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = 3\cdot 1 = 3. \tag{8.5.1}$$

$$\det\begin{pmatrix}2&0&0\\0&1&0\\0&0&1\end{pmatrix} = \det\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + \det\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = 1+1 = 2 \tag{8.5.2}$$

$$\det\begin{pmatrix}1&0&0\\2&4&0\\7&0&1\end{pmatrix} = \det\begin{pmatrix}1&2&7\\0&4&0\\0&0&1\end{pmatrix} = 1\cdot 4\cdot 1 = 4 \tag{8.5.3}$$

Aufgabe 8.2. *Berechnen Sie die Determinanten der gegebenen Matrizen mit möglichst geringem Rechenaufwand oder durch eine geeignete Begründung.*

$$A = \begin{pmatrix}1&5&2\\-3&-1&-6\\2&10&4\end{pmatrix},\ B = \begin{pmatrix}1&-4&3\\0&7&5\\0&0&3\end{pmatrix},\ C = \begin{pmatrix}1&3&4&6\\2&7&-2&3\\0&0&3&3\\0&0&4&3\end{pmatrix}.$$

Aufgabe 8.3. *Für welche Werte $t \in \mathbb{R}$ werden die Determinanten der nachfolgenden Matrizen 0?*

$$A = \begin{pmatrix}1&2\\2&t\end{pmatrix}, \qquad B = \begin{pmatrix}2&t\\t&18\end{pmatrix}, \qquad C = \begin{pmatrix}3&t\\-t&12\end{pmatrix}.$$

Aufgabe 8.4. *Berechnen Sie die Determinante $\det(A\cdot B)$ für folgende Matrizen und vergleichen Sie Ihr Ergebnis mit $\det(A)\cdot\det(B)$.*

$$A = \begin{pmatrix}4&5&6\\0&2&3\\0&0&1\end{pmatrix}, \qquad B = \begin{pmatrix}2&7&8\\0&2&9\\0&0&3\end{pmatrix}$$

Aufgabe 8.5. *Berechnen Sie die Determinante $\det(A+B)$ für folgende Matrizen und vergleichen Sie Ihr Ergebnis mit $\det(A)+\det(B)$.*

$$A = \begin{pmatrix}1&1&0\\1&2&0\\0&0&2\end{pmatrix}, \qquad B = \begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}$$

Aufgabe 8.6. *Zeigen Sie, dass die Determinante einer 2×2-Matrix aus Definition 8.1 Eigenschaft (D1) aus Definition 8.12 erfüllt, d.h. dass für $\lambda, \mu \in \mathbb{K}$ und $a, b, c, d, a', b' \in \mathbb{K}$ gilt:*

$$\det\begin{pmatrix}\lambda a + \mu a' & \lambda b + \mu b'\\ c & d\end{pmatrix} = \lambda\cdot\det\begin{pmatrix}a&b\\c&d\end{pmatrix} + \mu\cdot\det\begin{pmatrix}a'&b'\\c&d\end{pmatrix}.$$

Aufgabe 8.7. *Zeigen Sie Eigenschaften (E4), (E7) und (E9) von Determinanten aus Satz 8.16.*

Aufgabe 8.8. *a) Berechnen Sie die Determinanten der folgenden Matrizen:*

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \end{pmatrix}, \qquad B = \begin{pmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{pmatrix}.$$

b) Bestimmen Sie die Lösungsmenge der folgenden Gleichung:

$$\det \begin{pmatrix} x & y & 1 \\ 4 & 4 & 1 \\ 8 & 6 & 1 \end{pmatrix} = 0.$$

Aufgabe 8.9. *Berechnen Sie die Determinanten der folgenden Matrizen unter Verwendung des Entwicklungssatzes 8.27.*

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 3 & 1 \\ 2 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 12 & -10 & 17 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 8 & -1 & 1 \end{pmatrix}.$$

Kapitel 9

Eigenwerte, -vektoren und -räume

Ziele

- ☐ die Definitionen von Eigenwert, Eigenvektor und Eigenraum kennen und anwenden können
- ☐ Eigenwerte, -vektoren und -räume berechnen können
- ☐ Determinanten beliebig großer Matrizen berechnen können
- ☐ die Eigenschaften von Determinanten kennen (auch aus dem 1. Semester)
- ☐ mit Hilfe von Determinanten LGS lösen und Matrizen invertieren können

9.1 Eigenwerte und Eigenvektoren

In diesem Kapitel werden wir uns mit linearen Abbildungen beschäftigen, deren Ursprungs- und Zielraum identisch sind. Solche Abbildungen heißen Endomorphismen. Zudem werden wir uns auf endlich-dimensionale Vektorräume beschränken.

Definition 9.1: Endomorphismus

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum. Eine lineare Abbildung $f \in L(V;V)$ bezeichnet man als *Endomorphismus*.

Beispiel 9.2. *Betrachten wir drei Beispiele:*

$$f\colon \mathbb{R}^2 \to \mathbb{R}^3, \quad f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+y \\ 3x-y \\ 4y \end{pmatrix},$$

$$g\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad g\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x^2 \\ 3y \end{pmatrix}, \qquad h\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad h\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+3y \\ x \end{pmatrix}.$$

f ist kein Endomorphismus, da Ausgangs- und Zielraum nicht identisch sind. g ist ebenfalls kein Endomorphismus, da die beiden Räume zwar identisch sind, die Abbildung selbst jedoch nicht linear ist. h erfüllt beide Bedingung, es ist also ein Endomorphismus.

Definition 9.3: Eigenwert und Eigenvektor einer linearen Abbildung

Für eine natürliche Zahl $n \in \mathbb{N}$ sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum mit Nullvektor e_V. Sei zudem $f \in L(V;V)$.

Ein Skalar $\lambda \in \mathbb{K}$ heißt *Eigenwert* (EW) von f, falls es einen Vektor $v \in V \setminus \{e_V\}$ gibt, so dass

$$f(v) = \lambda \cdot v \tag{9.1.1}$$

erfüllt ist. Ein Vektor $v \in V \setminus \{e_V\}$ heißt *Eigenvektor* (EV) von f, falls es einen Skalar $\lambda \in \mathbb{K}$ gibt, so dass (9.1.1) erfüllt ist. Man nennt v dann Eigenvektor zum Eigenwert λ.

Wir wissen, dass jede lineare Abbildung $f \in L(V;V)$ für einen n-dimensionalen $\mathbb{K}$-Vektorraum V eine darstellende Matrix $A \in \mathbb{K}^{n\times n}$ besitzt. Ist $V = \mathbb{K}^n$, so kann A (bzw. die zu Grunde liegende Basis) so gewählt werden, dass für alle $x \in \mathbb{K}^n$ gilt: $f(x) = A \cdot x$. Damit können wir Eigenwerte und -vektoren von Matrizen als die Eigenwerte und -vektoren der zugehörigen linearen Abbildung definieren. Alternativ kann man sie unabhängig von den linearen Abbildungen wie folgt definieren:

Definition 9.4: Eigenwert und Eigenvektor einer Matrix

Sei $V = \mathbb{K}^n$ für $n \in \mathbb{N}$ mit Nullvektor e_V. Sei zudem $A \in \mathbb{K}^{n\times n}$.

Ein Skalar $\lambda \in \mathbb{K}$ heißt *Eigenwert* (EW) von A, falls es einen Vektor $v \in V \setminus \{e_V\}$ gibt, so dass

$$A \cdot v = \lambda \cdot v \tag{9.1.2}$$

erfüllt ist. Ein Vektor $v \in V \setminus \{e_V\}$ heißt *Eigenvektor* (EV) von A, falls es einen Skalar $\lambda \in \mathbb{K}$ gibt, so dass (9.1.2) erfüllt ist. Man nennt v dann Eigenvektor zum Eigenwert λ.

Beispiel 9.5. *Betrachten wir die lineare Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, welche gegeben ist durch*

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2y \\ 2x \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Überprüfen wir die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

daraufhin, ob sie Eigenvektoren von f bzw. von $A = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ sind.

- *Überprüfen wir, ob es ein $\lambda \in \mathbb{R}$ gibt, für welches* (9.1.1) *erfüllt wird.*

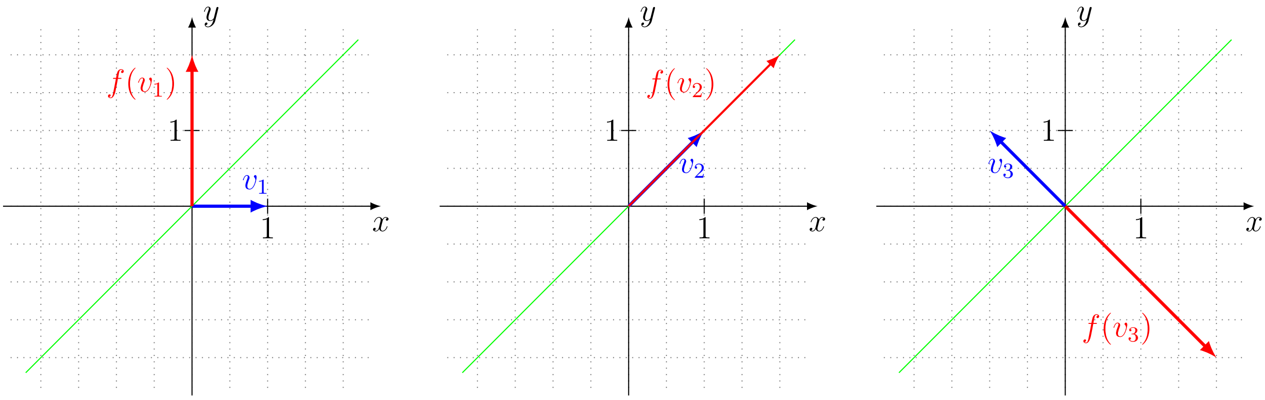
$$f(v_1) = f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

und da $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ kein Vielfaches von $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist, ist v_1 kein Eigenvektor von f bzw. von A.

- *Für den Vektor v_2 gilt*
$$f(v_2) = f\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2\\2\end{pmatrix} = 2\cdot v_2,$$
d.h. v_2 ist ein Eigenvektor von f (bzw. A) zum Eigenwert $\lambda = 2$.

- *Für den Vektor v_3 gilt*
$$f(v_3) = f\begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix}2\\-2\end{pmatrix} = (-2)\cdot v_3,$$
d.h. v_3 ist ein Eigenvektor von f (bzw. A) zum Eigenwert $\lambda' = -2$.

Erinnern wir uns daran, dass die Matrix A eine Spiegelung an der Winkelhalbierenden und eine Streckung um den Faktor 2 bewirkt.



Eigenvektoren haben die Eigenschaft, dass sie durch die lineare Abbildung (bzw. Matrix) höchstens gestreckt oder gestaucht werden. Identifizieren wir die Vektoren mit den Endpunkten der Ortsvektoren, so bedeutet das, dass der Eigenvektor und sein Bild auf einer Gerade, die durch den Ursprung geht, liegen.

Beispiel 9.6. *Betrachten wir zwei Extremfälle – die Identität und die Nullabbildung im $\mathbb{R}^2$:*

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2,\quad f\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\\y\end{pmatrix},$$
$$g\colon \mathbb{R}^2 \to \mathbb{R}^2,\quad g\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0&0\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}.$$

Der einzige Eigenwert von f ist $\lambda = 1$, denn für jeden Vektor $v \in \mathbb{R}^2 \setminus \{0\}$ gilt $f(v) = 1\cdot v$. Der einzige Eigenwert von g ist $\lambda = 0$, denn für jeden Vektor $v \in \mathbb{R}^2 \setminus \{0\}$ gilt $g(v) = 0\cdot v$. In beiden Fällen sind alle vom Nullvektor verschiedenen Vektoren Eigenvektoren zum jeweiligen Eigenwert.

Satz 9.7: Eigenvektoren zu verschiedenen Eigenwerten

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit Nullvektor e_V und $f \in L(V;V)$. Seien $v_1,\ldots,v_n$ ($n \in \mathbb{N}$) Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1,\ldots,\lambda_n$ von f. Dann ist $\{v_1,\ldots,v_n\} \subseteq V$ linear unabhängig.

Bemerkung 9.8. *Der Inhalt des Satzes lässt sich folgendermaßen kurz zusammenfassen:*

Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig.

Die Formulierung des Satzes ist zwar länger und auf den ersten Blick umständlicher, bietet jedoch den Vorteil, dass für den Beweis bereits alle Objekte benannt wurden.

Korollar 9.9

1. Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $f \in L(V;V)$. Dann besitzt f höchstens n Eigenwerte.

2. Sei $A \in \mathbb{K}^{n\times n}$. Dann besitzt A höchstens n Eigenwerte.

Beweis. Übungsaufgabe □

Beweis von Satz 9.7. Wir beweisen diesen Satz mittels vollständiger Induktion über $n \in \mathbb{N}$.

Induktionsanfang: Für $n = 1$ gilt in der Tat, dass $\{v_1\}$, d.h. die einelementige Menge, die nur einen Eigenvektor enthält, linear unabhängig ist, da dieser nach Voraussetzung nicht der Nullvektor ist.

Induktionsvoraussetzung: Für ein $m \geq 2$ seien die ersten $m-1$ Eigenvektoren, d.h. $\{v_1, \dots, v_{m-1}\}$, linear unabhängig.

Induktionsschritt: Wir wollen zeigen, dass $\{v_1, \dots, v_m\}$ linear unabhängig ist. Für Skalare $a_1, \dots, a_m \in \mathbb{K}$ gelte

$$\sum_{k=1}^{m} a_k v_k = a_1 v_1 + \ldots + a_m v_m = e_V. \tag{9.1.3}$$

Wir müssen zeigen, dass alle Skalare Null sind. Durch Anwendung der linearen Abbildung f auf (9.1.3) erhalten wir

$$\sum_{k=1}^{m} a_k f(v_k) \overset{f \text{ linear}}{=} f\left(\sum_{k=1}^{m} a_k v_k\right) \overset{(9.1.3)}{=} f(e_V) \overset{f \text{ linear}}{=} e_V,$$

Da für $k \in \{1, \dots, m\}$ jeweils v_k ein Eigenvektor zum Eigenwert λ_k ist, gilt $f(v_k) = \lambda_k \cdot v_k$, d.h wir erhalten

$$\sum_{k=1}^{m} a_k \lambda_k v_k = e_V. \tag{9.1.4}$$

Nun wenden wir einen Trick an: Wir wissen, dass das Vielfache des Nullvektors wieder der Nullvektor ist, d.h. insbesondere gilt $\lambda_m \cdot e_V = e_V$. Indem wir Gleichung (9.1.3) mit λ_m multiplizieren, erhalten wir also

$$\lambda_m \cdot \left(\sum_{k=1}^{m} a_k v_k\right) = e_V. \tag{9.1.5}$$

Subtrahieren wir nun (9.1.5) und (9.1.4) voneinander, so erhalten wir (nach Anwendung von KG und DG)

$$(\lambda_m - \lambda_1) \cdot a_1 v_1 + \ldots + (\lambda_m - \lambda_{m-1}) \cdot a_{m-1} v_{m-1} + (\lambda_m - \lambda_m) \cdot a_m v_m = e_V,$$

bzw.

$$(\lambda_m - \lambda_1) \cdot a_1 v_1 + \ldots + (\lambda_m - \lambda_{m-1}) \cdot a_{m-1} v_{m-1} = e_V.$$

Dies ist eine Linearkombination der nach Induktionvoraussetzung linear unabhängigen Vektoren $\{v_1, \ldots, v_{m-1}\}$. Da $\lambda_m \neq \lambda_k$ ist für $k \in \{1, \ldots, m-1\}$ müssen die übrigen Faktoren Null sein, d.h. $a_1 = \ldots = a_{m-1} = 0$. Setzen wir dies in (9.1.3) ein, so folgt, da $v_m \neq e_V$ ist, dass auch $a_m = 0$ ist, was den Beweis vervollständigt.

□

Im nächsten Abschnitt wollen wir untersuchen, wie genau die Menge der Eigenvektoren zu einem Eigenwert aussehen kann.

9.2 Eigenräume

Definition 9.10: Eigenraum

Für einen endlich-dimensionalen $\mathbb{K}$-Vektorraum V und einen Endomorphismus $f \in L(V;V)$ bezeichnen wir die Menge

$$\operatorname{Eig}(f, \lambda) := \{v \in V \mid f(v) = \lambda \cdot v\}$$

als *Eigenraum* von f zum Eigenwert $\lambda \in \mathbb{K}$.

Analog wird der Eigenraum einer Matrix zu einem Eigenwert definiert.

Bemerkung 9.11. *Der Nullvektor kann definitionsgemäß niemals ein Eigenvektor sein. Andererseits können wir uns schnell überzeugen, dass der Nullvektor in jedem Eigenraum enthalten ist:*

Wir wir erstmalig in Satz 3.22 festgestellt haben, bildet ein Gruppen-Homomorphismus das neutrale Element einer Gruppe immer auf dem neutralen Element der anderen Gruppe ab. Da eine lineare Abbildung gerade ein Vektorraum-Homomorphismus ist, gilt dieses Ergebnis auch für lineare Abbildungen. Zur Wiederholung geben wir den Beweis an dieser Stelle jedoch nochmals an:

Ist $f \in L(V;V)$, so gilt für jeden beliebigen Vektor $v \in V$

$$f(e_V) = f(0 \cdot v) \overset{f \text{ linear}}{=} 0 \cdot f(v) = e_V.$$

Somit gilt also insbesondere $f(e_V) = e_V = \lambda \cdot e_V$, d.h. in der Tat ist $e_V \in \operatorname{Eig}(f, \lambda)$ für jedes $\lambda \in \mathbb{K}$.

Satz 9.12: Eigenschaften des Eigenraumes

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit Nullvektor e_V und $f \in L(V;V)$ ein Endomorphismus.

1. $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert von f, wenn $\operatorname{Eig}(f,\lambda) \neq \{e_V\}$ ist.
2. $\operatorname{Eig}(f,\lambda) = \ker(f - \lambda \cdot \operatorname{id}|_V)$.
3. Für jedes $\lambda \in \mathbb{K}$ ist $\operatorname{Eig}(f,\lambda)$ ein Untervektorraum von V.

Beweis.

1. Ein Skalar $\lambda \in \mathbb{K}$ ist definitionsgemäß genau dann Eigenwert von f, wenn ein Vektor $v \in V \setminus \{e_V\}$ existiert, so dass $f(v) = \lambda \cdot v$ gilt. Diese Bedingung ist dazu äquivalent, dass ein Vektor $(V \setminus \{e_V\}) \cap \operatorname{Eig}(f,\lambda) \neq \emptyset$ ist, was wiederum äquivalent ist zur Behauptung.
2. Definitionsgemäß gilt:
$$\begin{aligned}\ker(f - \lambda \cdot \operatorname{id}|_V) &= \{v \in V \mid (f - \lambda \cdot \operatorname{id}|_V)(v) = e_V\} \qquad \text{(Def. Kern)}\\ &= \{v \in V \mid f(v) - \lambda \cdot v = e_V\}\\ &= \{v \in V \mid f(v) = \lambda \cdot v\}\\ &= \operatorname{Eig}(f,\lambda) \qquad \text{(Def. Eigenraum)}\end{aligned}$$
3. Diese Aussage folgt sofort aus 2., da der Kern einer linearen Abbildung, wie in Aufgabe 6.16 gezeigt wurde, stets ein Untervektorraum von V ist. Siehe auch Satz 6.70. Alternativ kann man zeigen, dass zu Vektoren $v, w \in \operatorname{Eig}(f,\lambda)$ und beliebigen Skalaren $a, b \in \mathbb{K}$ auch $av + bw \in \operatorname{Eig}(f,\lambda)$ ist, wobei man sowohl die definierende Eigenschaft des Eigenraumes als auch die Linearität von f benötigt. (Vgl. Präsenz- und Hausaufgabenblatt 1)

□

Korollar 9.13

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit Nullvektor e_V und $f \in L(V;V)$ ein Endomorphismus. Dann gilt jedes $\lambda \in \mathbb{K}$:

$$\dim \operatorname{Eig}(f;\lambda) = \dim(V) - \operatorname{rg}(f - \lambda \cdot \operatorname{id}|_V).$$

Insbesondere ist $\lambda \in \mathbb{K}$ genau dann ein Eigenwert von f, wenn $\operatorname{rg}(f - \lambda \cdot \operatorname{id}|_V) < \dim(V)$ ist.

Beweis. Diese Formel folgt sofort aus Eigenschaft 3 des vorherigen Satzes mit Hilfe der Dimensionsformel für lineare Abbildungen. Der Zusatz folgt wiederum aus Eigenschaft 1 des Satzes. □

Bemerkung 9.14. *Der Rang einer linearen Abbildung ist definiert als die Dimension ihres Bildes, d.h. für* $f \in L(V; W)$ *ist*

$$\mathrm{rg}(f) := \dim(\mathrm{Bild}(f)).$$

Hat $f \in L(\mathbb{R}^n; \mathbb{R}^m)$ *die darstellende Matrix* $A \in \mathbb{R}^{m \times n}$ *bezüglich der Standardbasis, d.h. gilt* $f(x) = A \cdot x$ *für alle* $x \in \mathbb{R}^n$*, so haben wir in Aufgabe 7.3 gezeigt, dass das Bild von* f *gleich der linearen Hülle der Spaltenvektoren von* A *ist. Folglich ist* $\mathrm{rg}(A) = \dim(\mathrm{Bild}(f))$*. Damit ist der Rang einer linearen Abbildung gleich dem Rang ihrer darstellenden Matrix.*

Das Korollar können wir auch für Matrizen statt Endomorphismen formulieren. Es lautet in diesem Fall wie folgt:

Korollar 9.15

Für ein $n \in \mathbb{N}$ und einen Körper $\mathbb{K}$ sei $A \in \mathbb{K}^{n \times n}$. Dann gilt jedes $\lambda \in \mathbb{K}$:

$$\dim \mathrm{Eig}(A; \lambda) = n - \mathrm{rg}(A - \lambda \cdot E_n).$$

Insbesondere ist $\lambda \in \mathbb{K}$ genau dann ein Eigenwert von A, wenn $\mathrm{rg}(A - \lambda \cdot E_n) < n$ ist.

Beispiel 9.16. *Bestimmen wir die Eigenräume* $\mathrm{Eig}(A; 0)$ *und* $\mathrm{Eig}(A; 1)$ *für die Matrix*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Da $\mathrm{rg}(A - 0 \cdot E_3) = \mathrm{rg}(A) = 3$ *ist, folgt aus Korollar 9.15 sofort* $\dim(Eig(A; 0)) = 3 - 3 = 0$*. Damit ist* $\mathrm{Eig}(A; 0) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$*. Auch ohne die Dimensionsformel können wir dies herausfinden:*

$$v \in \mathrm{Eig}(A; 0) \iff A \cdot v = 0 \cdot v = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Damit ist der Eigenraum gerade die Lösungsmenge des homogenen LGS, d.h.

$$\mathrm{Eig}(A; 0) = L(A, 0) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Folglich ist 0 *kein Eigenwert von* A*. Wiederholen wir dieses Vorgehen für* $\lambda = 1$*.*

$$\dim(\mathrm{Eig}(A; 1)) = 3 - \mathrm{rg}(A - 1 \cdot E_3) = 3 - \mathrm{rg} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 1 = 2.$$

In der Tat ist der Eigenraum 2-dimensional und somit 1 *ein Eigenwert von* A*:*

$$\begin{aligned}
\operatorname{Eig}(A;1) &= \left\{ v \in \mathbb{R}^3 \,\middle|\, A \cdot v = 1 \cdot v \right\} \\
&= \left\{ v \in \mathbb{R}^3 \,\middle|\, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right\} \\
&= \left\{ v \in \mathbb{R}^3 \,\middle|\, \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ v \in \mathbb{R}^3 \,\middle|\, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ v \in \mathbb{R}^3 \,\middle|\, v_2 = 0 \right\} \\
&= \left\{ a \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \,\middle|\, a, b \in \mathbb{R} \right\} \\
&= \mathcal{L}\left(\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right).
\end{aligned}$$

Den Eigenraum berechnen wir also genauso wie die Lösungsmenge eines homogenen LGS. Da wir nur quadratische Matrizen betrachten, können wir die Determinante der Matrix als Indikator nutzen, der uns sagt, ob das homogene LGS eine eindeutige Lösung besitzt (nämlich den Nullvektor) – in diesem Fall liegt kein Eigenwert vor und der Eigenraum besteht nur aus dem Nullvektor – oder ob die Lösungsmenge ein mindestens eindimensionaler Vektorraum ist, den wir dann als lineare Hülle schreiben können – in diesem Fall ist der Wert λ ein Eigenwert und der Eigenraum enthält neben dem Nullvektor alle zu λ gehörenden Eigenvektoren.

Bevor wir uns genauer ansehen, wie wir die Eigenwerte finden, wiederholen wir zunächst die Definition und die wichtigsten Eigenschaften einer Determinante, welche wir in Kapitel 8 eingeführt hatten.

9.3 Wiederholung: Determinanten

In Kapitel 8 haben wir in Definition 8.12 Determinanten quadratischer Matrizen eingeführt als eine Abbildung, die drei definierende Eigenschaften besitzt.

Wir haben gezeigt, dass für jedes $n \in \mathbb{N}$ eine Abbildung mit solchen Eigenschaften existiert und dass sie eindeutig bestimmt ist. Zur Berechnung für die Spezialfälle $n = 2$ und $n = 3$ gibt es Formeln, wohingegen für ein allgemeines $n \geq 2$ der Entwicklungssatz von Laplace (Satz 8.27) die Berechnung von Determinanten ermöglicht.

Sehen wir uns ein paar Beispiele für Determinanten an.

Beispiel 9.17. *Betrachten wir folgende Matrizen:*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & 2 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 6 & 8 \\ 0 & 0 & 6 \end{pmatrix}.$$

An der Matrix A wollen wir zunächst die Regel von Sarrus nochmal in Erinnerung rufen. Bei dieser summieren wir die Produkte der Einträge auf den Diagonalen von links oben nach rechts unten und subtrahieren davon die Produkte der Einträge auf den Diagonalen von links unten nach rechts oben.

$$\begin{aligned} \det(A) &= 1 \cdot 5 \cdot 9 + 4 \cdot 8 \cdot 3 + 7 \cdot 2 \cdot 6 - 7 \cdot 5 \cdot 3 - 8 \cdot 6 \cdot 1 - 9 \cdot 4 \cdot 2 \\ &= 45 + 96 + 84 - 105 - 48 - 72 \\ &= 0. \end{aligned}$$

Andererseits können wir durch Entwicklung nach der 1. Spalte die Determinante wie folgt berechnen:

$$\begin{aligned} \det(A) &= (-1)^{1+1} \cdot 1 \cdot \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} + (-1)^{2+1} \cdot 4 \cdot \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} + (-1)^{3+1} \cdot 7 \cdot \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} \\ &= \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} - 4 \cdot \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} + 7 \cdot \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} \\ &= (5 \cdot 9 - 8 \cdot 6) - 4 \cdot (2 \cdot 9 - 3 \cdot 8) + 7 \cdot (2 \cdot 6 - 3 \cdot 5) \\ &= -3 - 4 \cdot (-6) + 7 \cdot (-3) \\ &= 0. \end{aligned}$$

Dieses Beispiel illustriert Eigenschaft (E8) aus Satz 8.16, welche besagt, dass für eine Matrix $A \in \mathbb{K}^{n \times n}$ folgende Äquivalenz gilt:

$$\det(A) = 0 \quad \Longleftrightarrow \quad \operatorname{rg}(A) < n.$$

In der Tat gilt in unserem konkreten Fall $\operatorname{rg}(A) = 2 < 3$, denn[1]

$$\begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} = 2 \cdot \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} - \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}.$$

Berechnen wir Die Determinante von B zu Illustrationszwecken durch eine Entwicklung nach der 2. Zeile.

$$\begin{aligned} \det(B) &= (-1)^{2+1} \cdot 0 \cdot \det \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} + (-1)^{2+2} \cdot 4 \cdot \det \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} + (-1)^{2+3} \cdot 0 \cdot \det \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \\ &= 0 + 4 \cdot (3 \cdot 1 - 1 \cdot 2) + 0 \\ &= 4. \end{aligned}$$

Dies bestätigt nochmal die Berechnung aus Beispiel 8.22. Wenden wir uns zuletzt noch der Matrix C zu:

---

[1] Diese Rechnung beweist, dass die drei Spaltenvektoren von A nicht linear unabhängig sind, jeweils zwei von ihnen sind es jedoch. Damit ist der Spaltenrang und somit der Rang von A genau 2. Die Definition des Ranges kann in Definition 6.25 und Satz 6.27 nachgelesen werden.

Nach Eigenschaft (E5) aus Satz 8.16 ist die Determinante einer Matrix in Zeilenstufenform gegeben durch das Produkt der Einträge auf der Hauptdiagonale, d.h.

$$\det(C) = 1 \cdot 6 \cdot 6 = 36.$$

Überprüfen wir dies nochmal durch die Entwicklung nach der 1. Spalte von C:

$$\begin{aligned}\det(C) &= 1 \cdot \det\begin{pmatrix} 6 & 8 \\ 0 & 6 \end{pmatrix} - 0 \cdot \det\begin{pmatrix} 2 & 3 \\ 0 & 6 \end{pmatrix} + 0 \cdot \det\begin{pmatrix} 2 & 3 \\ 6 & 8 \end{pmatrix} \\ &= 6 \cdot 6 - 0 \cdot 8 \\ &= 36.\end{aligned}$$

Die Inhalte der wichtigsten Sätze zu Eigenschaften von Determinanten sind nachzulesen als Satz 8.16, Satz 8.20 und Korollar 8.23. Für den nächsten Abschnitt ist die wichtigste Eigenschaft, dass die Determinante ein Indikator dafür ist, ob eine Matrix invertierbar ist:

$$\text{Für } A \in \mathbb{K}^{n\times n} \text{ gilt:} \quad \det(A) = 0 \iff \operatorname{rg}(A) < n. \tag{E7}$$

9.4 Berechnung von Eigenwerten mit Hilfe von Determinanten

9.4.1 Das charakteristische Polynom einer Matrix

Da uns häufig der konkrete Rang einer Matrix nicht interessiert, sondern nur, ob der Rang maximal ist (und die Matrix somit invertierbar), führen wir folgende Bezeichnung ein, die wir im Anschluss gleich benutzen werden.

Definition 9.18: regulär, singulär

Für $n \in \mathbb{N}$ sei $A \in \mathbb{K}^{n\times n}$. Die Matrix A heißt *regulär*, falls A invertierbar ist, d.h. falls $\operatorname{rg}(A) = n$ ist. Andernfalls heißt A *singulär*.

Beispiel 9.19. *Die Matrix A aus Beispiel 9.17 ist singulär, die Matrizen B und C sind hingegen regulär. Dies folgt unmittelbar aus den berechneten Determinanten und Eigenschaft* (E7).

In diesem Abschnitt wenden wir uns wieder den Eigenwerten von Matrizen zu und der Frage, wie diese berechnet werden können.

Nehmen wir dazu an, $v \in \mathbb{K}^n$ sei ein Eigenvektor der Matrix $A \in \mathbb{K}^{n\times n}$ zum Eigenwert $\lambda \in \mathbb{K}$. Das heißt, dass v nicht der Nullvektor ist und die Gleichung

$$A \cdot v = \lambda \cdot v \tag{9.4.1}$$

erfüllt wird. Die Multiplikation des Vektors v mit dem Skalar λ lässt sich ebenfalls als Multiplikation mit einer Matrix umschreiben. Dazu multiplizieren wir v mit $\lambda \cdot E_n$, d.h. mit der Diagonalmatrix, auf deren Hauptdiagonale überall der Eintrag λ steht:

$$\lambda \cdot E_n \cdot v = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix} = \lambda \cdot v.$$

Gleichung (9.4.1) können wir somit als homogenes LGS umschreiben:

$$A \cdot v - \lambda \cdot v = A \cdot v - (\lambda \cdot E_n) \cdot v = (A - \lambda \cdot E_n) \cdot v = e_{\mathbb{R}^n}.$$

In der Notation, die wir in Abschnitt 6.6.5 eingeführt haben, können wir dies folgendermaßen umschreiben:

$$v \in L(A - \lambda \cdot E_n, 0).$$

Laut Satz 6.70 ist $L(A - \lambda \cdot E_n, 0)$ ein Untervektorraum des $\mathbb{R}^n$ und seine Dimension bestimmen wir mit Hilfe der Dimensionsformel aus Satz 6.73:

$$\dim(L(A - \lambda \cdot E_n, 0)) = n - \operatorname{rg}(A - \lambda \cdot E_n).$$

Da wir bereits angenommen hatten, dass v ein Eigenvektor von A zum Eigenwert λ sei, ist $\dim(L(A - \lambda \cdot E_n, 0)) \geq 1$ und somit $\operatorname{rg}(A - \lambda \cdot E_n) < n$. Wir haben also gezeigt, dass wenn λ ein Eigenwert der Matrix $A \in \mathbb{K}^{n \times n}$ ist, die Matrix $A - \lambda \cdot E_n$ singulär ist. Die Umkehrung dieser Aussage gilt ebenfalls, was wir in folgendem Satz gleich zusammenfassend feststellen.

Satz 9.20

Seien $A \in \mathbb{K}^{n \times n}$ und $\lambda \in \mathbb{K}$. Der Wert λ ist genau dann ein Eigenwert von A, wenn die Matrix $A - \lambda \cdot E_n$ singulär ist.

Beweis. Wir haben bereits die eine Richtung gezeigt, nämlich dass für einen Eigenwert λ von A die Matrix $A - \lambda \cdot E_n$ stets singulär ist. Zeigen wir nun die andere Richtung. Dazu nehmen wir an, dass $\operatorname{rg}(A - \lambda \cdot E_n) < n$ sei. Da damit aus der Dimensionsformel folgt, dass $\dim(L(A - \lambda \cdot E_n, 0)) \geq 1$ ist, gibt es (mindestens) einen Vektor $v \neq e_{\mathbb{R}^n}$ in der Lösungsmenge dieses homogenen LGS und dieser Vektor ist gerade ein Eigenvektor von A zum Eigenwert λ, denn es gilt

$$(A - \lambda \cdot E_n) \cdot v = e_{\mathbb{R}^n} \quad \Longleftrightarrow \quad A \cdot v = \lambda \cdot v.$$

□

Bemerkung 9.21. *Formal haben wir die Äquivalenz bewiesen indem wir beide Implikationen separat gezeigt haben. Da die Argumente jedoch in beiden Richtungen die gleichen waren, könnte man in diesem Fall den Beweis direkt als Aneinanderreihung von äquivalenten Aussagen aufschreiben.*

Die Eigenwerte sind also die Zahlen $\lambda \in \mathbb{K}$, für die $\det(A - \lambda \cdot E_n) = 0$ ist. Das folgt sofort aus dem Satz und der Eigenschaft (E7). Sehen wir uns diese Determinante einmal genauer an.

Beispiel 9.22. *Wir suchen die Eigenwerte der Matrizen B und C aus Beispiel 9.17.*

$$\begin{aligned}
\det(C - \lambda \cdot E_3) &= \det \begin{pmatrix} 1-\lambda & 2 & 3 \\ 0 & 6-\lambda & 8 \\ 0 & 0 & 6-\lambda \end{pmatrix} \\
&\overset{\text{(E5) aus Satz 8.16}}{=} (1-\lambda) \cdot (6-\lambda) \cdot (6-\lambda) \\
&= -\lambda^3 + 13\lambda^2 - 48\lambda + 36.
\end{aligned}$$

Die Eigenwerte von C sind also die Nullstellen dieses Polynoms, d.h. 1 und 6.

Wiederholen wir dieses Vorgehen für die leicht kompliziertere Matrix B:

$$\begin{aligned}
\det(B - \lambda \cdot E_3) &= \det \begin{pmatrix} 3-\lambda & 1 & 2 \\ 0 & 4-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{pmatrix} \\
&= (3-\lambda) \cdot \det \begin{pmatrix} 4-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix} - 0 + 1 \cdot \det \begin{pmatrix} 1 & 2 \\ 4-\lambda & 0 \end{pmatrix} \\
&= (3-\lambda) \cdot [(4-\lambda) \cdot (1-\lambda) - 0 \cdot 0] + [1 \cdot 0 - (4-\lambda) \cdot 2] \\
&= (3-\lambda) \cdot (4-\lambda) \cdot (1-\lambda) - 2 \cdot (4-\lambda) \\
&= (4-\lambda) \cdot [\lambda^2 - 4\lambda + 1] \\
&= -\lambda^3 + 8\lambda^2 - 17\lambda + 4
\end{aligned}$$

An der vorletzten Zeile erkennt man sofort eine Nullstelle dieses Polynoms, nämlich $\lambda_1 = 4$. Zudem hat das quadratische Polynom dieser Zeile die mit Hilfe der p-q-Formel leicht zu berechnenden Nullstellen $\lambda_2 = 2 + \sqrt{3}$ und $\lambda_3 = 2 - \sqrt{3}$. Diese drei Nullstellen sind die Eigenwerte der Matrix B.

Lemma 9.23

Sei $A \in \mathbb{K}^{n \times n}$. Dann ist $p_A(t) := \det(A - t \cdot E_n) \in \mathbb{K}[t]$ ein Polynom n-ten Grades.

Bemerkung 9.24. *Es erscheint ungewöhnlich, als Variable für das Polynom t statt x zu wählen. Da wir x jedoch häufig als Bezeichnung für einen Vektor verwenden, die Variable hier jedoch ein Körperelement sein soll, werde ich für p_A immer den Variablennamen t statt x benutzen.*

Beweis. Sei $A = (\alpha_{k,j})_{k,j \in \{1,\dots,n\}}$. Dann berechnet man durch wiederholte Anwendung des Entwicklungssatzes nach der jeweils ersten Spalte (oder Zeile):

$$\begin{aligned}
p_A(t) &= \det \begin{pmatrix} \alpha_{1,1} - t & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} - t & \cdots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} - t \end{pmatrix} \\
&= (\alpha_{1,1} - t) \cdot \det \begin{pmatrix} \alpha_{2,2} - t & \alpha_{2,3} & \cdots & \alpha_{2,n} \\ \alpha_{3,2} & \alpha_{3,3} - t & \cdots & \alpha_{3,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,2} & \alpha_{n,3} & \cdots & \alpha_{n,n} - t \end{pmatrix} + q_1(t) \\
&= (\alpha_{1,1} - t) \cdot (\alpha_{2,2} - t) \cdot \det \begin{pmatrix} \alpha_{3,3} - t & \cdots & \alpha_{3,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,3} & \cdots & \alpha_{n,n} - t \end{pmatrix} + q_2(t) + q_1(t) \\
&\;\vdots \\
&= (\alpha_{1,1} - t) \cdot (\alpha_{2,2} - t) \cdot \ldots \cdot (\alpha_{n,n} - t) + q(t).
\end{aligned}$$

Mit unseren Mitteln ist es etwas mühsam, nachzuweisen, dass q_1, q_2, ... und somit auch deren Summe, $q(t)$, Polynome sind und dass deren Grad höchstens $n-2$ beträgt. Mit Hilfe der *Leibniz*'schen Determinantenformel[2] folgt dies sofort ohne viel Mühe, daher überspringen wir dieses Detail ausnahmsweise.

Zudem ist $\widetilde{q}(t) := (\alpha_{1,1}-t)\cdot(\alpha_{2,2}-t)\cdot\ldots\cdot(\alpha_{n,n}-t)$ ein Polynom n-ten Grades, denn durch Ausmultiplizieren der Terme erhält man

$$\widetilde{q}(t) = (-1)^n t^n + (-1)^{n-1}(\alpha_{1,1}+\alpha_{2,2}+\ldots+\alpha_{n,n})t^{n-1} + r(t),$$

wobei $r(t)$ ein Polynom vom Grad $\leq n-2$ ist. Da weder q noch r Terme n-ten Grades enthält, folgt sofort, dass $\deg(p_A) = n$ ist wie behauptet. $\square$

Definition 9.25: charakteristisches Polynom

Sei $A \in \mathbb{K}^{n\times n}$. Die Abbildung $p_A\colon \mathbb{K} \to \mathbb{K}$ mit der Abbildungsvorschrift

$$p_A(t) := \det(A - t\cdot E_n)$$

heißt *charakteristisches Polynom* von A.

Mit dieser Bezeichnung können wir folgendes Resultat formulieren, welches die Berechnung der Eigenwerte einer Matrix ermöglicht. Wir werden das Korollar nicht beweisen, da es unmittelbar aus Satz 9.20 folgt.

Korollar 9.26

Die Eigenwerte einer Matrix $A \in \mathbb{K}^{n\times n}$ sind genau die Nullstellen des charakteristischen Polynoms der Matrix.

Beispiel 9.27. *Betrachten wir folgende drei Matrizen:*

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{pmatrix}.$$

Ihre Eigenwerte sind sehr einfach zu finden:

$$\begin{aligned} p_A(t) &= \det\begin{pmatrix} -t & 0 & 0 \\ 0 & -t & 0 \\ 0 & 0 & -t \end{pmatrix} \\ &= (-t)^3 \end{aligned}$$

besitzt die dreifache Nullstelle 0 *– dies ist somit der einzige Eigenwert von* A*.*

$$\begin{aligned} p_B(t) &= \det\begin{pmatrix} 1-t & 0 & 0 \\ 0 & 1-t & 0 \\ 0 & 0 & 1-t \end{pmatrix} \\ &= (1-t)^3 \end{aligned}$$

[2] Die Leibniz'sche Formel wird beispielsweise in Kapitel 7.4 des Buches *Lineare Algebra* von A. Beutelspacher [Beu14] präsentiert. Man benötigt zudem die Abschnitte 7.2 und 7.3 um 7.4 zu verstehen.

besitzt die dreifache Nullstelle 1 *– die Einheitsmatrix besitzt somit nur den Eigenwert* 1*.*

$$\begin{aligned} p_C(t) &= \det \begin{pmatrix} 1-t & 0 & 0 \\ 0 & 2-t & 0 \\ 0 & 0 & -5-t \end{pmatrix} \\ &= (1-t)\cdot(2-t)\cdot(-5-t) \end{aligned}$$

besitzt die drei Nullstellen 1*,* 2 *und* -5*, welche somit die Eigenwerte von* C *sind.*

9.4.2 Das charakteristische Polynom eines Endomorphismus

Während wir Eigenwerte und -vektoren noch sowohl für Matrizen als auch für lineare Abbildungen eingeführt hatten, können wir Determinanten nur für Matrizen berechnen. Zudem hatten wir die Eigenwerte und Eigenvektoren einer Matrix auf zweierlei Weise eingeführt – einerseits als die Eigenwerte und -vektoren der zugehörigen linearen Abbildung und andererseits über die Formel (9.1.4). Auch hier haben wir eine gewisse Asymmetrie so lange folgende Frage noch unbeantwortet ist:

Unterscheiden sich die Eigenwerte und -vektoren von verschiedenen darstellenden Matrizen bezüglich ein und derselben linearen Abbildung? Etwas formeller gefragt: Sei $f \in L(V;V)$ und seien B und C zwei (georndete) Basen des Vektorraumes V. Seien A_f^B und A_f^C die jeweiligen darstellenden Matrizen von f bezüglich dieser Basen. Sind dann die Eigenwerte und -vektoren beider Matrizen gleich?

Um diese Frage zu beantworten, wiederholen wir an dieser Stelle kurz die darstellende Matrix und die Basistransformationsmatrix.

Wiederholung: Basiswechsel und darstellende Matrix einer linearen Abbildung

Seien V ein n-dimensionaler und W ein m-dimensionaler $\mathbb{K}$-Vektorraum und seien $\mathcal{B} = (v_1, \dots, v_n) \subseteq V$ und $\mathcal{C} = (w_1, \dots, w_m) \subseteq W$ geordnete Basen von V bzw. W. Sei nun $f \in L(V;W)$ eine lineare Abbildung. Gemäß Satz 7.15 gibt es eine genau eine Matrix $A_f^{\mathcal{B},\mathcal{C}}$, die folgende Bedingung erfüllt:

$$f(v_j) = \sum_{k=1}^{m} \alpha_{k,j} w_k, \quad \forall j \in \{1, \dots, n\}. \tag{9.4.2}$$

Diese Matrix nennen wir die darstellende Matrix von f bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$.

In diesem Semester geht es vorrangig um Endomorphismen, d.h. $V = W$ und somit können beide Basen identisch gewählt werden. In diesem Fall schreiben wir für die darstellende Matrix von $f \in L(V;V)$ bezüglich $\mathcal{B}$ kurz $A_f^{\mathcal{B}}$ statt $A_f^{\mathcal{B},\mathcal{B}}$.

Beispiel 9.28. *Gegeben sei der Endomorphismus* $f \in L(\mathbb{R}^3;\mathbb{R}^3)$ *mit folgender Abbildungsvorschrift:*

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ -4x+7y \\ 3x+5y+3z \end{pmatrix}.$$

Wählen wir drei Basen des $\mathbb{R}^3$:

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right), \mathcal{C} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) \text{ und } \mathcal{D} = \left(\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \right).$$

Da $\mathcal{B}$ die Standardbasis ist, können wir $A_f^{\mathcal{B}}$ direkt anhand der Koeffizienten der Abbildungsvorschrift ablesen:

$$A_f^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 7 & 0 \\ 3 & 5 & 3 \end{pmatrix}.$$

Für die darstellende Matrix von f bezüglich $\mathcal{C}$ wenden wir Gleichung (9.4.2) an:

$$\begin{aligned}
f\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix} = 5 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-7) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\
f\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 3 \\ 8 \end{pmatrix} = (-2) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-5) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 8 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\
f\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ 3 \\ 11 \end{pmatrix} = (-2) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-8) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 11 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.
\end{aligned}$$

Die Koeffizienten des Gleichungssystems bilden, wenn wir sie spaltenweise in eine Matrix eintragen, die gesuchte Matrix

$$A_f^{\mathcal{C}} = \begin{pmatrix} 5 & -2 & -2 \\ -7 & -5 & -8 \\ 3 & 8 & 11 \end{pmatrix}.$$

Diese Rechnung können wir alternativ mit Hilfe von Matrizen ausdrücken. Dazu führen wir die zu den Basen gehörenden Matrizen

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad D = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

ein. Das Suchen der blauen Koeffizienten entspricht dann dem Lösen des LGS $C \cdot x = f(w_k)$, wobei $f(w_k)$ die Bilder der Basisvektoren von $\mathcal{C} = (w_1, w_2, w_3)$ darstellt. Da $\mathcal{C}$ eine Basis ist, hat C vollen Rang und ist somit invertierbar. Das LGS hat somit die eindeutige Lösung $C^{-1} \cdot f(w_k)$. Andererseits ist $f(w_k) = A_f^{\mathcal{B}} \cdot w_k$. Schreiben wir die Ergebnisvektoren aneinander, so erhalten wir die Gleichung

$$A_f^{\mathcal{C}} = C^{-1} \cdot A_f^{\mathcal{B}} \cdot C.$$

In der Tat ist

$$C^{-1} \cdot A_f^{\mathcal{B}} \cdot C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ -4 & 7 & 0 \\ 3 & 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -2 & -2 \\ -7 & -5 & -8 \\ 3 & 8 & 11 \end{pmatrix} = A_f^{\mathcal{C}}.$$

Analog berechnen wir die darstellende Matrix von f bezüglich $\mathcal{D}$:

$$A_f^{\mathcal{D}} = D^{-1} \cdot A_f^{\mathcal{B}} \cdot D = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ -4 & 7 & 0 \\ 3 & 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & -3 & 5 \\ 0 & 1 & 0 \\ 0 & 4 & 7 \end{pmatrix}.$$

Überprüfen wir nun andersherum, ob damit (9.4.2) *erfüllt ist:*

$$f\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -3 \end{pmatrix} = (-3) \cdot \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$f\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix} = (-3) \cdot \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 4 \cdot \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$f\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -7 \\ -5 \end{pmatrix} = 5 \cdot \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 7 \cdot \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

Der Vollständigkeit halber wollen wir nun sehen, wie man von $A_f^{\mathcal{C}}$ zu $A_f^{\mathcal{D}}$ gelangt. Es gelten folgende Gleichungen, die man durch Umstellen der vorherigen Gleichungen nach $A_f^{\mathcal{B}}$ erhält:

$$A_f^{\mathcal{B}} = C \cdot A_f^{\mathcal{C}} \cdot C^{-1} \quad \textit{und} \quad A_f^{\mathcal{B}} = D \cdot A_f^{\mathcal{D}} \cdot D^{-1}.$$

Somit gilt auch $C \cdot A_f^{\mathcal{C}} \cdot C^{-1} = D \cdot A_f^{\mathcal{D}} \cdot D^{-1}$, welches wir nach einer der darstellenden Matrizen umstellen können, zum Beispiel ist

$$A_f^{\mathcal{D}} = D^{-1} \cdot C \cdot A_f^{\mathcal{C}} \cdot C^{-1} \cdot D = (C^{-1} \cdot D)^{-1} \cdot A_f^{\mathcal{C}} \cdot (C^{-1} \cdot D).$$

Das charakteristische Polynom darstellender Matrizen zu verschiedenen Basen

Satz 9.29

Seien $\mathcal{B}$ und $\mathcal{C}$ zwei geordnete Basen eines n-dimensionalen $\mathbb{K}$-Vektorraumes V und sei $f \in L(V;V)$. Seien $A := A_f^{\mathcal{B}}$ und $A' := A_f^{\mathcal{C}}$ die darstellenden Matrizen von f bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$. Dann gilt $p_A(t) = p_{A'}(t)$ für alle $t \in \mathbb{K}$, d.h. die charakteristischen Polynome der darstellenden Matrizen stimmen überein.

Beweis. Sei T die Matrix, die den Basiswechsel von $\mathcal{B}$ nach $\mathcal{C}$ ermöglicht, so dass

$$A_f^{\mathcal{C}} = T^{-1} \cdot A_f^{\mathcal{B}} \cdot T$$

ist. Dann gilt unter Beachtung der Rechengesetze für Matrizen und für Determinanten:

$$\begin{aligned} p_{A'}(t) &= \det(A' - t \cdot E_n) \\ &= \det(T^{-1} \cdot A \cdot T - t \cdot E_n) \\ &= \det(T^{-1} \cdot A \cdot T - t \cdot T^{-1} \cdot E_n \cdot T) \\ &= \det(T^{-1} \cdot (A - t \cdot E_n) \cdot T) \\ &= \det(T^{-1}) \cdot \det(A - t \cdot E_n) \cdot \det(T) \\ &= (\det(T))^{-1} \cdot \det(T) \cdot \det(A - t \cdot E_n) \\ &= \det(A - t \cdot E_n) \\ &= p_A(t). \end{aligned}$$

□

Beispiel 9.30 (Fortsetzung von Beispiel 9.28). *Berechnen wir die charakteristischen Polynome von $A_f^{\mathcal{B}}$, $A_f^{\mathcal{C}}$ und $A_f^{\mathcal{D}}$.*

$$p_{A_f^{\mathcal{B}}}(t) = \det \begin{pmatrix} 1-t & 0 & 0 \\ -4 & 7-t & 0 \\ 3 & 5 & 3-t \end{pmatrix} = (1-t)(7-t)(3-t),$$

da die Matrix eine untere Dreiecksmatrix ist. (Das ist Eigenschaft (E5') aus Satz 8.20.)

Des Weiteren ist

$$\begin{aligned} p_{A_f^{\mathcal{C}}}(t) &= \det \begin{pmatrix} 5-t & -2 & -2 \\ -7 & -5-t & -8 \\ 3 & 8 & 11-t \end{pmatrix} \\ &= -(5-t)(5+t)(11-t) + 112 + 48 - 6 \cdot (5+t) + 64 \cdot (5-t) - 14 \cdot (11-t) \\ &= -t^3 + 11t^2 + 25t - 275 + 160 - 30 - 6t + 320 - 64t - 154 + 14t \\ &= -t^3 + 11t^2 - 31t + 21 \\ &= (1-t)(7-t)(3-t). \end{aligned}$$

Zum Abschluss berechnen noch das charakteristische Polynom von $A_f^{\mathcal{D}}$:

$$p_{A_f^{\mathcal{D}}}(t) = \det \begin{pmatrix} 3-t & -3 & 5 \\ 0 & 1-t & 0 \\ 0 & 4 & 7-t \end{pmatrix} = (3-t) \cdot \det \begin{pmatrix} 1-t & 0 \\ 4 & 7-t \end{pmatrix} = (3-t)(1-t)(7-t).$$

Da die Wahl der Basis keinen Einfluss auf das charakteristische Polynom der darstellenden Matrix eines Endomorphismus hat, ist es sinnvoll, das charakteristische Polynom eines Endomorphismus über die zugehörige darstellende Matrix zu definieren.

Definition 9.31

Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum mit einer geordneten Basis $\mathcal{B}$. Sei $f \in L(V;V)$ ein Endomorphismus und sei $A_f^{\mathcal{B}}$ die darstellende Matrix von f bezüglich $\mathcal{B}$. Dann ist das charakteristische Polynom von f definiert als

$$p_f(t) := p_{A_f^{\mathcal{B}}}(t) = \det(A_f^{\mathcal{B}} - t \cdot E_n).$$

Eigenwerte und Eigenvektoren beim Basiswechsel

Da die charakteristischen Polynome von darstellenden Matrizen zu verschiedenen Basen identisch sind, sind auch deren Nullstellen und somit die Eigenwerte identisch. Was ist jedoch mit den Eigenvektoren?

Nehmen wir dazu an, dass $\mathcal{B}$ und $\mathcal{C}$ zwei Basen eines Vektorraumes V sind. Reihen wir die Basisvektoren aneinander, so erhalten wir die Matrizen B und C. Zudem seien $A_f^{\mathcal{B}}$ und $A_f^{\mathcal{C}}$ die darstellenden Matrizen eines Endomorphismus $f \in L(V;V)$ bezüglich der jeweiligen Matrix.

Wenn $v \in V \setminus \{e_V\}$ ein Eigenvektor von $A_f^{\mathcal{C}}$ zum Eigenwert λ ist, dann gilt $A_f^{\mathcal{C}} \cdot v = \lambda \cdot v$. Zudem gilt nach der Formel aus Beispiel 9.28

$$A_f^{\mathcal{C}} = C^{-1} \cdot B \cdot A_f^{\mathcal{B}} \cdot B^{-1} \cdot C.$$

Damit gilt insgesamt

$$C^{-1} \cdot B \cdot A_f^{\mathcal{B}} \cdot B^{-1} \cdot C \cdot v = \lambda \cdot v.$$

Dies formen wir mit den Rechenregeln für Matrizen um und erhalten

$$A_f^{\mathcal{B}} \cdot \underbrace{B^{-1} \cdot C \cdot v}_{w} = \lambda \cdot \underbrace{B^{-1} \cdot C \cdot v}_{w},$$

d.h. $w := B^{-1} \cdot C \cdot v$ ist ein Eigenvektor von $A_f^{\mathcal{B}}$ zum Eigenwert λ.

Beispiel 9.32 (Fortsetzung von Beispiel 9.28)**.** *Die Eigenwerte von f sind, wie man an den gerade berechneten charakteristischen Polynomen ablesen kann, $\lambda_1 = 1$, $\lambda_2 = 3$ und $\lambda_3 = 7$. Berechnen wir exemplarisch die Eigenräume von $A_f^{\mathcal{B}}$ und $A_f^{\mathcal{C}}$ zum Eigenwert 7:*

$$\begin{aligned}
\mathrm{Eig}(A_f^{\mathcal{B}}; 7) &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} 1-7 & 0 & 0 \\ -4 & 7-7 & 0 \\ 3 & 5 & 3-7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} -6 & 0 & 0 \\ -4 & 0 & 0 \\ 3 & 5 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x = 0, 5y = 4z \right\} = \mathcal{L} \begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix}.
\end{aligned}$$

Andererseits ist

$$\begin{aligned}
\mathrm{Eig}(A_f^{\mathcal{C}}; 7) &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} 5-7 & -2 & -2 \\ -7 & -5-7 & -8 \\ 3 & 8 & 11-7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} -2 & -2 & -2 \\ -7 & -12 & -8 \\ 3 & 8 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}.
\end{aligned}$$

Lösen wir dieses LGS mit dem Gauß-Algorithmus:

$$\left[\begin{array}{ccc|c} -2 & -2 & -2 & 0 \\ -7 & -12 & -8 & 0 \\ 3 & 8 & 4 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ -7 & -12 & -8 & 0 \\ 3 & 8 & 4 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -5 & -1 & 0 \\ 0 & -5 & -1 & 0 \end{array}\right]$$
$$\rightsquigarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 1/5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & 4/5 & 0 \\ 0 & 1 & 1/5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Damit ist

$$\mathrm{Eig}(A_f^{\mathcal{C}}; 7) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{array}{rl} x & = -\frac{4}{5}z \\ y & = -\frac{1}{5}z \\ z & = z \end{array} \right\} = \mathcal{L} \begin{pmatrix} -4/5 \\ -1/5 \\ 1 \end{pmatrix} = \mathcal{L} \begin{pmatrix} -4 \\ -1 \\ 5 \end{pmatrix}.$$

Somit ist $w = \begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix}$ *ein Eigenvektor von* $A_f^{\mathcal{B}}$ *zum Eigenwert* 7 *und* $v = \begin{pmatrix} -4 \\ -1 \\ 5 \end{pmatrix}$ *ein Eigenvektor von* $A_f^{\mathcal{C}}$ *zum Eigenwert* 7*. Zudem gilt in der Tat*

$$B^{-1} \cdot C \cdot v = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ -1 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix} = w.$$

Die gleiche Beziehung lässt sich auch für die Eigenvektoren zu den Eigenwerten 1 und 3 bestätigen.

9.5 Die Cramer'sche Regel und die Berechnung der Inversen mit Hilfe von Determinanten

In diesem Abschnitt soll es nochmals um zwei bereits bekannte Problemstellungen gehen: das Lösen eines LGS und die Berechnung der Inversen einer Matrix (sofern dies möglich ist). In Kapitel 5 zum systematischen Lösen von linearen Gleichungssystemen haben wir den Gauß-Algorithmus kennen gelernt. Zudem haben wir in Abschnitt 6.6 Lösungsmengen als affine oder lineare Vektorräume kennen gelernt. Ist ein LGS in der Form $A \cdot x = b$ gegeben, wobei A invertierbar ist, so existiert genau eine Lösung des LGS, nämlich $x = A^{-1} \cdot b$. Daher ist die Untersuchung der Invertierbarkeit eng mit der Lösungstheorie verknüpft. Zum Berechnen der Inversen kennen wir bisher die in Abschnitt 5.3 vorgestellte Methode mit Hilfe des Gauß-Algorithmus.

Es gibt noch eine weitere Methode, die Inverse einer Matrix zu berechnen, welche auf der Berechnung von Determinanten beruht, und ebenso können wir die Lösungsmenge eines LGS mit Hilfe von Determinanten bestimmen, sofern die zu Grunde liegende Matrix quadratisch ist. Das sehen wir uns im Folgenden an.

9.5.1 Herleitung der Cramer'schen Regel

Betrachten wir ein LGS der Form

$$\begin{aligned} \alpha_{1,1}x_1 + \ldots + \alpha_{1,n}x_n &= b_1 \\ \alpha_{2,1}x_1 + \ldots + \alpha_{2,n}x_n &= b_2 \\ &\vdots \\ \alpha_{n,1}x_1 + \ldots + \alpha_{n,n}x_n &= b_n. \end{aligned} \tag{9.5.1}$$

Die zugehörige Schreibweise als *erweiterte Koeffizientenmatrix* ist

$$\left[\begin{array}{ccc|c} \alpha_{1,1} & \cdots & \alpha_{1,n} & b_1 \\ \alpha_{2,1} & \cdots & \alpha_{2,n} & b_2 \\ \vdots & & \vdots & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} & b_n \end{array}\right],$$

was wiederum eine Abkürzung des Problems ist, den Vektor $x \in \mathbb{K}^n$ zu bestimmen, der die Gleichung

$$A \cdot x = b$$

löst, wobei

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \cdots & \alpha_{2,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad \text{ist.}$$

Das Ziel ist es nun, eine Darstellung für x_k ($k \in \{1, \ldots, n\}$) zu bekommen.

Sehen wir uns nochmal das LGS an, so können wir $A \cdot x = b$ auch in folgender Form schreiben:

$$x_1 \cdot \begin{pmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \vdots \\ \alpha_{n,1} \end{pmatrix} + x_2 \cdot \begin{pmatrix} \alpha_{1,2} \\ \alpha_{2,2} \\ \vdots \\ \alpha_{n,2} \end{pmatrix} + \ldots + x_k \cdot \begin{pmatrix} \alpha_{1,k} \\ \alpha_{2,k} \\ \vdots \\ \alpha_{n,k} \end{pmatrix} + \ldots + x_n \cdot \begin{pmatrix} \alpha_{1,n} \\ \alpha_{2,n} \\ \vdots \\ \alpha_{n,n} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Subtrahieren wir den Vektor b auf beiden Seiten, so erhalten wir eine Linearkombination von Vektoren, die den Nullvektor darstellt:

$$x_1 \cdot \begin{pmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \vdots \\ \alpha_{n,1} \end{pmatrix} + x_2 \cdot \begin{pmatrix} \alpha_{1,2} \\ \alpha_{2,2} \\ \vdots \\ \alpha_{n,2} \end{pmatrix} + \ldots + 1 \cdot \begin{pmatrix} x_k\alpha_{1,k} - b_1 \\ x_k\alpha_{2,k} - b_2 \\ \vdots \\ x_k\alpha_{n,k} - b_n \end{pmatrix} + \ldots + x_n \cdot \begin{pmatrix} \alpha_{1,n} \\ \alpha_{2,n} \\ \vdots \\ \alpha_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Da mindestens einer der Koeffizienten nicht Null ist (nämlich 1), sind die Vektoren linear abhängig. Nach (E7) ist also die Matrix, die aus diesen Spaltenvektoren gebildet werden kann, singulär, d.h. ihre Determinante ist Null:

$$\det \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & x_k\alpha_{1,k} - b_1 & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & x_k\alpha_{2,k} - b_2 & \cdots & \alpha_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & x_k\alpha_{n,k} - b_n & \cdots & \alpha_{n,n} \end{pmatrix} = 0.$$

Da einerseits die Determinante nach Definition 8.12 linear in jeder Zeile ist und andererseits das Transponieren die Determinante nicht ändert, ist die Determinante auch linear in jeder Spalte, d.h. wir können wir diese Gleichung wie folgt umformen:

$$x_k \cdot \det \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,k} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,k} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,k} & \cdots & \alpha_{n,n} \end{pmatrix} - \det \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & b_1 & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & b_2 & \cdots & \alpha_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & b_n & \cdots & \alpha_{n,n} \end{pmatrix} = 0.$$

Der erste Summand ist das x_k-Fache von $\det(A)$. Sofern $\det(A) \neq 0$ ist, können wir diese Gleichung nach x_k umstellen und erhalten

$$x_k = \frac{\det \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & b_1 & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & b_2 & \cdots & \alpha_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & b_n & \cdots & \alpha_{n,n} \end{pmatrix}}{\det(A)},$$

wobei die k-te Spalte von A durch den Vektor b ersetzt wurde. $k \in \{1, \ldots, n\}$ war hierbei beliebig, so dass man auf diese Art den gesamten Lösungsvektor bekommt. Dieses Ergebnis ist bekannt unter dem Namen *Cramer'sche Regel*, benannt nach Gabriel Cramer (1704-1752), wurde jedoch bereits vorher von Gottfried Wilhelm Leibniz gefunden.

Satz 9.33: Cramer'sche Regel

Gegeben sei das LGS (9.5.1) mit der quadratischen regulären Koeffizientenmatrix A und einem beliebigen Ergebnisvektor b. Dieses LGS besitzt eine eindeutige Lösung $x = (x_1, \ldots, x_n)^t$, deren Komponenten gegeben sind durch die Formel

$$x_k = \frac{\det \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & b_1 & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & b_2 & \cdots & \alpha_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & b_n & \cdots & \alpha_{n,n} \end{pmatrix}}{\det(A)}, \qquad k \in \{1, \ldots, n\}.$$

Bemerkung 9.34. *Achtung! Diese Regel ist nur dann anwendbar, wenn die Koeffizientenmatrix A regulär ist, was insbesondere voraussetzt, dass sie quadratisch ist.*

Beispiel 9.35. *Wählen wir ein LGS, das leicht lösbar ist, so dass wir das Ergebnis gut überprüfen können:*

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ \qquad x_2 + x_3 + x_4 = -2 \\ \qquad\qquad x_3 + x_4 = 5 \\ \qquad\qquad\qquad x_4 = 4 \end{cases}$$

Dieses LGS können wir „von unten nach oben" lösen und erhalten die eindeutige Lösung

$$x = \begin{pmatrix} 3 \\ -7 \\ 1 \\ 4 \end{pmatrix}.$$

Schreiben wir das LGS in der Matrixschreibweise, so ist

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 1 \\ -2 \\ 5 \\ 4 \end{pmatrix}.$$

Da alle Spalten (und auch alle Zeilen) von A linear unabhängig sind, ist $\mathrm{rg}(A) = 4$ und somit ist A invertierbar und das LGS hat eine eindeutige Lösung, welche gegeben ist durch $x = A^{-1} \cdot b$. Anstatt die Inverse zu berechnen, was wir später noch tun werden, wenden wir an dieser Stelle die Cramer'sche Regel an. Dazu berechnen wir zunächst die Determinante von A. Da diese Matrix in Zeilenstufenform vorliegt, ist $\det(A) = 1$, was die kommende Rechnung vereinfacht. Somit müssen wir nur die Determinante aus dem Zähler berechnen.

Wir werden das Kürzel S1 *für die Entwicklung nach der ersten Spalte und* Z1 *für die Entwicklung nach der ersten Zeile verwenden (oder entsprechend andere Zahlen) und zwei senkrechte Striche für die Determinante benutzen.*

$$x_1 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ -2 & 1 & 1 & 1 \\ 5 & 0 & 1 & 1 \\ 4 & 0 & 0 & 1 \end{vmatrix} \overset{S2}{=} (-1) \cdot \begin{vmatrix} -2 & 1 & 1 \\ 5 & 1 & 1 \\ 4 & 0 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 1 & 1 \\ 5 & 1 & 1 \\ 4 & 0 & 1 \end{vmatrix} = -(-7) + (-4) = 3$$

$$x_2 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 0 & 5 & 1 & 1 \\ 0 & 4 & 0 & 1 \end{vmatrix} \overset{S1}{=} 1 \cdot \begin{vmatrix} -2 & 1 & 1 \\ 5 & 1 & 1 \\ 4 & 0 & 1 \end{vmatrix} = -7$$

$$x_3 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 4 & 1 \end{vmatrix} \overset{S1}{=} 1 \cdot \begin{vmatrix} 1 & -2 & 1 \\ 0 & 5 & 1 \\ 0 & 4 & 1 \end{vmatrix} = 1$$

$$x_4 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 4 \end{vmatrix} \overset{\textit{Diagonale}}{=} 1 \cdot 1 \cdot 1 \cdot 4 = 4.$$

Schreibt man diese vier Werte in einen Vektor, so erhält man in der Tat den eindeutigen Lösungsvektor x des LGS.

9.5.2 Invertieren von Matrizen

Erinnern wir uns zunächst daran, was wir bereits über das Invertieren wissen:

- Nur quadratische Matrizen können gleichzeitig vollen Zeilen- und Spaltenrang haben, d.h. das Suchen einer Inverse ist nur für quadratische Matrizen sinnvoll.
- Eine Matrix $A \in \mathbb{K}^{n \times n}$ ist genau dann invertierbar, wenn $\operatorname{rg}(A) = n$ ist und inzwischen wissen wir, dass das wiederum genau dann der Fall ist, wenn $\det(A) \neq 0$ ist.
- A heißt invertierbar, falls eine (ebenfalls quadratische) Matrix $B \in \mathbb{K}^{n \times n}$ existiert, so dass $A \cdot B = B \cdot A = E_n$ gilt. Diese Inverse bezeichnen wir mit $B =: A^{-1}$.
- Die Inverse zu A finden wir, indem wir elementare Zeilenumformungen auf die erweiterte Koeffizientenmatrix $[A|E_n]$ anwenden bis diese in der Form $[E_n|A^{-1}]$ vorliegt.

Beispiel 9.36. *Da die Matrix A aus Beispiel 9.35 vollen Rang hat (bzw. da ihre Determinante 1 und somit $\neq 0$ ist), ist A invertierbar. Sehen wir uns noch einmal den bisher bekannten Weg des Invertierens an:*

$$\left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1. \end{array}\right]$$

Somit ist

$$A^{-1} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Bereits bei dem Entwicklungssatz von Laplace (Satz 8.27) sind uns Koeffizienten der Form

$$\gamma_{k,j} := (-1)^{k+j} \det(A_{k,j}), \quad k, j \in \{1, \dots, n\} \tag{9.5.2}$$

für eine Matrix $A = (\alpha_{k,j})$ begegnet.

Definition 9.37

Sei $A = (\alpha_{k,j}) \in \mathbb{K}^{n\times n}$ und sei $C = (\gamma_{k,j}) \in \mathbb{K}^{n\times n}$ die Matrix mit den Einträgen (9.5.2). Die Einträge $\gamma_{k,j}$ heißen *Kofaktoren* von A, die Matrix C die *Kofaktorenmatrix* von A und deren Transponierte, C^T, wird *Adjunkte* oder *komplementäre Matrix* zu A genannt.

Satz 9.38

Sei $A \in \mathbb{K}^{n\times n}$ regulär und sei C^T die Adjunkte von A. Dann gilt $A \cdot C^T = C^T \cdot A = \det(A) \cdot E_n$ und somit

$$A^{-1} = \frac{1}{\det(A)} C^T.$$

Beispiel 9.39 (2 × 2-Matrix)**.** *Sei A eine allgemeine 2×2-Matrix d.h. $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ für $a, b, c, d \in \mathbb{K}$. Dann ist nach obigem Satz und mit der bereits bekannten Determinante*

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Beispiel 9.40 (Fortsetzung von Beispiel 9.36)**.** *Sei wieder*

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Es ist $\det(A) = 1$*, da* A *in Zeilenstufenform vorliegt und so nur die Diagonalelemente multipliziert werden müssen. Bestimmen wir nun die Kofaktoren von* A*:*

$$\gamma_{1,1} = (-1)^{1+1} \cdot \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1, \quad \gamma_{2,1} = (-1)^{2+1} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = -1, \quad \gamma_{3,1} = 0, \quad \gamma_{4,1} = 0,$$

$$\gamma_{1,2} = (-1)^{1+2} \cdot \begin{vmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 0, \quad \gamma_{2,2} = (-1)^{2+2} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1, \quad \gamma_{3,2} = -1, \quad \gamma_{4,2} = 0,$$

$$\gamma_{1,3} = (-1)^{1+3} \cdot \begin{vmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 0, \quad \gamma_{2,3} = (-1)^{2+3} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 0, \quad \gamma_{3,3} = 1, \quad \gamma_{4,3} = -1,$$

$$\gamma_{1,4} = (-1)^{1+4} \cdot \begin{vmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0, \quad \gamma_{2,4} = (-1)^{2+4} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0, \quad \gamma_{3,4} = 0, \quad \gamma_{4,4} = 1.$$

Dabei ist zu beachten, dass beispielsweise $A_{1,2}$ *entsteht, indem die 1. Zeile und 2. Spalte von* A *gestrichen wird – der zugehörige Kofaktor* $\gamma_{1,2}$ *steht durch das Transponieren innerhalb der Adjunkten jedoch in der 1. Spalte und 2. Zeile, weshalb die Kofaktoren bereits so angeordnet wurden. Die Adjunkte lautet in diesem Fall:*

$$C^T = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Da $\det(A) = 1$ *ist, ist also* $A^{-1} = \frac{1}{\det(A)} \cdot C^T = C^T$.

Beweis von Satz 9.38. Der Beweis erfolgt in Aufgabe 9.10. □

9.6 Aufgaben

Aufgabe 9.1. *Entscheiden Sie, ob die angegebenen Vektoren Eigenvektoren zur gegebenen Matrix sind und falls ja, bestimmen Sie den zugehörigen Eigenwert.*

a) $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, v_3 = \begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix};$

b) $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad w_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$

Aufgabe 9.2. *Können Sie aus der vorherigen Aufgabe Eigenwerte und -vektoren zu folgender linearer Abbildung nennen?*

$$f\colon \mathbb{R}^3 \to \mathbb{R}^3, \quad f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 4x + 5y + 6z \\ 7x + 8y + 9z \end{pmatrix}.$$

Aufgabe 9.3. *Sei $f \in L(\mathbb{R}^n; \mathbb{R}^n)$ eine lineare Abbildung und seien $v_1, v_2 \in \mathbb{R}^n$ zwei Eigenvektoren von f zum Eigenwert $\lambda \in \mathbb{R}$. Zeigen Sie, dass dann auch $v_1 + v_2$ ein Eigenvektor von f zum Eigenwert λ ist, sofern $v_1 + v_2$ nicht der Nullvektor ist.*

(i) *Schreiben Sie auf, was es bedeutet, dass v_1 ein Eigenvektor von f zum Eigenwert λ ist. Wiederholen Sie dies für v_2.*

(ii) *Wiederholen Sie Schritt (i) für $v_1 + v_2$.*

(iii) *Laut Aufgabenstellung soll f linear sein. Was bedeutet das genau?*

(iv) *Markieren Sie, welche der bisher notierten Aussagen vorausgesetzt werden dürfen und was gezeigt werden soll.*

(v) *Überlegen Sie, wie Sie die Voraussetzungen so kombinieren können, dass die Behauptung daraus folgt.*

(vi) *Schreiben einen sauberen Beweis auf, in dem Sie nur die notwendigen Bestandteile der bisherigen Überlegungen in einer sinnvollen Reihenfolge in vollständigen Sätzen verknüpfen.*

Aufgabe 9.4. *Sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix, deren Rang $m := \mathrm{rg}(A) < n$ ist.*

a) *Bestimmen Sie die Dimension des Kerns der Abbildung $f_A \in L(\mathbb{R}^n; \mathbb{R}^n)$ mit $f_A(x) = A\cdot x$. Begründen Sie Ihre Antwort.*

b) *Begründen Sie, warum ein Vektor $v \in \ker(f_A)$ existiert, der nicht der Nullvektor ist.*

c) *Zeigen Sie, dass der Vektor v aus b) ein Eigenvektor von A ist und geben Sie den zugehörigen Eigenwert an.*

Aufgabe 9.5. *Sei $f \in L(\mathbb{R}^n; \mathbb{R}^n)$ gegeben durch die Abbildungsvorschrift $f(x) = A \cdot x$, wobei $A \in \mathbb{R}^{n\times n}$ sei. Füllen Sie die Lücken so aus, dass lauter zueinander äquivalente Aussagen entstehen.*

a) ________ *ist ein Isomorphismus.*

b) ________ *ist eine invertierbare Matrix.*

c) $\mathrm{rg}(A)$ ________.

d) $\det(A)$ ________.

e) $\ker(f)$ ________.

f) $\mathrm{Bild}(f)$ ________.

g) *Die Spalten von A sind* ____.

h) *Es existiert eine Matrix $B \in \mathbb{R}^{n\times n}$, so dass $B \cdot A =$* ________ *ist.*

i) *Das LGS $A \cdot x = b$ ist* ____ *lösbar.*

j) $\lambda = 0$ *ist* __________ *Eigenwert von A.*

Aufgabe 9.6. *Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt* nilpotent, *wenn eine Zahl $k \in \mathbb{N}$ existiert, so dass $A^k = O$ die Nullmatrix des $\mathbb{K}^{n\times n}$ ist. Zeigen Sie, dass jede nilpotente Matrix singulär ist. Gehen Sie dazu folgende Schritte durch:*

a) *Nehmen Sie an, A sei regulär. Was wissen Sie dann über A?*

b) *Sei* $k^* := \min\left\{ k \in \mathbb{N} \,\middle|\, A^k = O \right\}$. *Was bedeutet das?*

c) *Zeigen Sie, dass für jede reguläre Matrix* A *aus* $A^k = O$ *bereits* $A^{k-1} = O$ *folgt und vervollständigen Sie den Beweis.*

Aufgabe 9.7. *Berechnen Sie zu folgenden Matrizen das charakteristische Polynom*

$$p_A(t) = \det(A - t \cdot E_n)$$

sowie dessen Nullstellen im angegebenen Körper.

$$A = \begin{pmatrix} 1 & -1 \\ -3 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -4 & 3 \\ 0 & 7 & 5 \\ 0 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 5 & 2 \\ 3 & -1 & -6 \\ 2 & 10 & 4 \end{pmatrix}$$

a) $p_A(t)$ *über* $\mathbb{K} = \mathbb{R}$

b) $p_B(t)$ *über* $\mathbb{K} = \mathbb{R}$

c) $p_C(t)$ *über* $\mathbb{K} = \mathbb{R}$

d) $p_C(t)$ *über* $\mathbb{K} = \mathbb{C}$

Aufgabe 9.8. *Berechnen Sie die gesuchten Eigenräume zu folgenden Matrizen:*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 \\ -3 & -1 \end{pmatrix}$$

a) $\mathrm{Eig}(A; 3)$

b) $\mathrm{Eig}(A; 2)$

c) $\mathrm{Eig}(B; 0)$

d) $\mathrm{Eig}(B; -2)$

Aufgabe 9.9. *Seien* $\lambda_1, \lambda_2 \in \mathbb{R}$ *zwei verschiedene Eigenwerte einer Matrix* $A \in \mathbb{R}^{5\times 5}$. *Bestimmen Sie*

$$\mathrm{Eig}(A; \lambda_1) \cap \mathrm{Eig}(A; \lambda_2).$$

Aufgabe 9.10. ***Beweis von Satz 9.38 – Erinnerung an Notation und Aussage:***

Für eine Matrix $A = (\alpha_{k,j})$ *seien die* Kofaktoren *von* A *definiert als*

$$\gamma_{k,j} := (-1)^{k+j} \det(A_{k,j}), \quad k, j \in \{1, \ldots, n\}. \tag{9.6.1}$$

Dann gilt:
Sei $A \in \mathbb{K}^{n\times n}$ *regulär und sei* $\mathrm{adj}(A) := C^T$ *die Adjunkte von* A, *d.h.* $C = (\gamma_{k,j}) \in \mathbb{K}^{n\times n}$. *Dann gilt*

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) \cdot E_n \tag{9.6.2}$$

und folglich ist

$$A^{-1} = \frac{1}{\det(A)} C^T. \tag{9.6.3}$$

Füllen Sie die Lücken im folgenden Beweis des Satzes:

a) *Sei wie üblich $A_{k,j}$ die $(n-1)\times(n-1)$-Matrix, die aus A durch Streichen der k-ten Zeile und j-ten Spalte entsteht und sei zudem*

$$A_{k,j}^{(1)} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \cdots & \alpha_{2,j-1} & \alpha_{2,j} & \alpha_{2,j+1} & \cdots & \alpha_{2,n} \\ \vdots & & \vdots & \vdots & & & \vdots \\ \alpha_{k-1,1} & \cdots & \alpha_{k-1,j-1} & \alpha_{k-1,j} & \alpha_{k-1,j+1} & \cdots & \alpha_{k-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \alpha_{k+1,1} & \cdots & \alpha_{k+1,j-1} & \alpha_{k+1,j} & \alpha_{k+1,j+1} & \cdots & \alpha_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots & \\ \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{pmatrix},$$

d.h. die ______. Zeile von A wird durch den Zeilenvektor $e_j^T = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}$ ersetzt, dessen ______. Eintrag 1 ist und alle anderen Einträge sind Nullen. Die Matrix $A_{k,j}^{(1)}$ hat ______ Zeilen und ______ Spalten.

b) *Für beliebiges $k \in \{1,\dots,n\}$ ist nach der Linearität der Determinante in jeder Zeile*

$$\sum_{j=1}^{n} \alpha_{k,j} \det(A_{i,j}^{(1)}) = \begin{cases} \underline{\hspace{2cm}}, & \textit{falls } k = i, \\ \underline{\hspace{2cm}}, & \textit{falls } k \neq i. \end{cases}$$

c) *Zur Illustration dieser Formel: Betrachten wir die Matrix*

$$A = \begin{pmatrix} 1 & 0 & -2 \\ -1 & 1 & 2 \\ 0 & 4 & -1 \end{pmatrix}.$$

Diese ist invertierbar, denn

$$\det(A) = \begin{vmatrix} 1 & 0 & -2 \\ -1 & 1 & 2 \\ 0 & 4 & -1 \end{vmatrix} = \underline{\hspace{1cm}} \cdot \begin{vmatrix} 1 & 2 \\ 4 & -1 \end{vmatrix} - 0 + \underline{\hspace{1cm}} \cdot \begin{vmatrix} -1 & 1 \\ 0 & 4 \end{vmatrix} = -9 + 8 = -1 \neq 0.$$

Wenden wir nun die Formel an:

$$\begin{aligned} \sum_{j=1}^{3} \alpha_{1,j} \det(A_{1,j}^{(1)}) &= 1 \cdot \det(A_{1,1}^{(1)}) + 0 \cdot \det(A_{1,2}^{(1)}) + (-2) \cdot \det(A_{1,3}^{(1)}) \\ &= 1 \begin{vmatrix} 1 & 0 & 0 \\ -1 & 1 & 2 \\ 0 & 4 & -1 \end{vmatrix} + 0 \begin{vmatrix} 0 & 1 & 0 \\ -1 & 1 & 2 \\ 0 & 4 & -1 \end{vmatrix} - 2 \begin{vmatrix} 0 & 0 & 1 \\ -1 & 1 & 2 \\ 0 & 4 & -1 \end{vmatrix} \\ &= \underline{\hspace{6cm}} \\ &= \underline{\hspace{1cm}} \end{aligned}$$

Andererseits ist

$$\begin{aligned}\sum_{j=1}^{3} \alpha_{1,j} \det(A_{2,j}^{(1)}) &= 1 \cdot \det(A_{2,1}^{(1)}) + 0 \cdot \det(A_{2,2}^{(1)}) + (-2) \cdot \det(A_{2,3}^{(1)}) \\ &= 1 \cdot \left| \qquad\qquad \right| + 0 \cdot \left| \qquad\qquad \right| + (-2) \cdot \left| \qquad\qquad \right| \\ &= \underline{\qquad\qquad\qquad\qquad\qquad\qquad} \\ &= \underline{\qquad}\end{aligned}$$

d) *Weiterhin gilt*

$$\det(A_{i,j}^{(1)}) = (-1)^{i+j} \det(A_{i,j}),$$

denn mit Hilfe elementarer Zeilenumformungen kann man $A_{i,j}^{(1)}$ *umformen in die Form*

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A_{i,j} & \\ 0 & & & \end{array}\right)$$

Dazu Addiert man zuerst Vielfache der ____. *Zeile zu den übrigen Zeilen um in der* ____. *Spalte Nullen zu erzeugen. Das ändert die Determinante nicht. Anschließend kann man durch* ______ *Vertauschungen benachbarter Zeilen und durch* ______ *Vertauschungen benachbarter Spalten die gewünschte Form erreichen. Jede solche Vertauschung bewirkt* ____ *und insgesamt gab es* ________ *Vertauschungen.*

Da

$$(-1)^{i+j} = (-1)^{(i-1)+(j-1)}$$

ist, folgt die behauptete Formel.

e) *Nun berechnen wir die Einträge von* $A \cdot \operatorname{adj}(A)$. *Dazu erinnern wir uns zunächst an die Definition des Produktes zweier Matrizen. Ist* $A = (\alpha_{k,j})$ *und* $B = (\beta_{j,l})$, *dann gilt für* $C = A \cdot B =: (\gamma_{k,l})$

$$\gamma_{k,l} = \sum_j \alpha_{k,j} \beta_{j,l}.$$

Nach Definition der Adjunkten ist $\operatorname{adj}(A) = (\gamma_{k,j})^T$. *Durch Verwendung der Formeln aus b) und d) ergibt sich, wenn* $A = (\alpha_{k,j})$ *und* $\operatorname{adj}(A) = (\beta_{j,l})$ *ist:*

$$\begin{aligned}\sum_{j=1}^{n} \alpha_{k,j} \beta_{j,i} &\overset{\text{Def. Adj.}}{=} \\ &\overset{\text{d)}}{=} \\ &\overset{\text{b)}}{=}\end{aligned}$$

woraus folgt, dass in der Tat $A \cdot \operatorname{adj}(A) = \det(A) \cdot E_n$ *ist. Durch Vertauschen der Rollen beider Matrizen folgt die zweite Gleichheit aus* (9.6.2).

f) *Die Formel für die Inverse von A, d.h.* (9.6.3)*, folgt aus* (9.6.2) *indem auf beiden Seiten der Gleichung durch* ______________ *dividiert wird. Hier wird nicht durch 0 geteilt, da* A _________________________ *ist. Zudem gilt: Die Inverse von A ist die Matrix, für die gilt:* _____________________________

Aufgabe 9.11. *Berechnen Sie, falls möglich, die Inversen der folgenden Matrizen:*

$$A = \begin{pmatrix} 1 & 5 & 2 \\ 3 & -1 & -6 \\ 2 & 10 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -4 & 3 \\ 0 & 7 & 5 \\ 0 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 & 4 & 6 \\ 2 & 7 & -2 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 4 & 3 \end{pmatrix}.$$

Kapitel 10

Diagonalisierbarkeit von Matrizen

Ziele

- □ diagonalisierbare Matrizen als Sonderfall von ähnlichen Matrizen auffassen
- □ überprüfen können, ob eine Matrix diagonalisierbar ist
- □ die Transformationsmatrizen bestimmen, mit deren Hilfe eine Matrix diagonalisiert wird
- □ die Diagonalgestalt als Sonderfall der Jordan'schen Normalform auffassen

10.1 Einführung: Wozu sind Diagonalmatrizen nützlich?

Es gibt diverse Anwendungen, die die Diagonalisierung von Matrizen erfordern. Einige davon erfordern zusätzliches Wissen aus der Analysis oder Inhalte, die erst später in diesem Semester behandelt werden. Daher sollen ein paar ausgewählte Beispiele zur Illustration genügen, auch wenn diese nicht die ganze Bandbreite der Möglichkeiten umfassen. Doch vorher erinnern wir nochmals kurz daran, was wir unter einer Diagonalmatrix verstehen:

Definition 10.1

Eine quadratische Matrix $D \in \mathrm{Mat}(n,n)$ heißt *Diagonalmatrix*, falls sie nur auf der Hauptdiagonale von Null verschiedene Einträge hat, d.h.

$$D = \mathrm{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Beispiel 10.2. *Betrachten wir zu Demonstrationszwecken die Matrix*

$$A = \mathrm{diag}(1,2,3,4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Man erkennt ohne jeden Rechenaufwand sofort, dass A regulär, also invertierbar, ist, da alle Spalten (bzw. Zeilen) von A linear unabhängig sind. Zudem ist die Determinante leicht zu berechnen:

$$\det(A) = 1 \cdot 2 \cdot 3 \cdot 4 = 24.$$

Die Eigenwerte der Diagonalmatrix sind gerade die Einträge auf der Diagonalen, d.h. A besitzt die Eigenwerte 1, 2, 3 und 4.

Auch das Invertieren der Matrix ist einfach:

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1/4 \end{array}\right],$$

d.h.

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}.$$

Die Multiplikation der Matrix mit sich selber ist ebenfalls einfacher als bei beliebigen Matrizen:

$$A^2 = A \cdot A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 16 \end{pmatrix} = \operatorname{diag}(1^2, 2^2, 3^2, 4^2).$$

Es gilt analog $A^3 = \operatorname{diag}(1^3, 2^3, 3^3, 4^3)$ oder allgemeiner für eine beliebige Potenz $n \in \mathbb{N}$:

$$A^n = \operatorname{diag}(1^n, 2^n, 3^n, 4^n) = \begin{pmatrix} 1^n & 0 & 0 & 0 \\ 0 & 2^n & 0 & 0 \\ 0 & 0 & 3^n & 0 \\ 0 & 0 & 0 & 4^n \end{pmatrix}.$$

Beim Lösen mehrdimensionaler linearer Differentialgleichungen trifft man auf das Matrixexponential, d.h. das Analogon zur Exponentialfunktion für Matrizen und dieses ist für Diagonalmatrizen ebenfalls einfach berechenbar. Details hierzu führen an dieser Stelle jedoch zu weit.

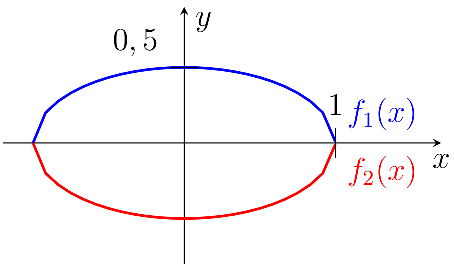
Beispiel 10.3. *Betrachten wir folgende Menge:*

$$M := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1 \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; x^2 + 4y^2 = 1 \right\}$$

Da x und y getrennt voneinander vorkommen, können wir die Bedingung an x und y nach y umstellen und so die Menge folgendermaßen umschreiben:

$$M = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; y = \pm\frac{1}{2}\sqrt{1 - x^2}, x \in [-1, 1] \right\}$$

Indem wir die Graphen der Funktionen $f_1(x) = \frac{1}{2}\sqrt{1 - x^2}$ und $f_2(x) = -\frac{1}{2}\sqrt{1 - x^2}$ über dem Intervall $[-1, 1]$ plotten, finden wir eine Darstellung der Menge M:



M ist also eine Ellipse um den Koordinatenursprung.

Liegt eine Ellipse nicht so günstig, ist es nicht so schnell möglich, sie als solche zu erkennen. Dieser Fragestellung werden wir später beim Thema Hauptachsentransformation begegnen.

Es ist also praktisch, wenn Matrizen in Diagonalform vorliegen. Da dies jedoch in der Regel nicht der Fall ist, behelfen wir uns damit, dass wir versuchen, einen Bezug zwischen der gegebenen und einer Diagonalmatrix herzustellen und diesen Bezug für das Rechnen auszunutzen. Was das genau bedeutet, sehen wir uns im nächsten Abschnitt an.

10.2 Ähnliche Matrizen

Definition 10.4

Zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ heißen *zueinander ähnlich*, wenn eine reguläre Matrix $S \in \mathbb{K}^{n\times n}$ existiert, so dass $B = S^{-1}AS$ ist.

Bemerkung 10.5. *Alternativ wird in der Literatur auch gefordert, dass eine reguläre Matrix $T \in \mathbb{K}^{n\times n}$ existieren soll, für die $B = TAT^{-1}$ gelte. Dann ist jedoch $B = (T^{-1})^{-1} \cdot A \cdot T^{-1}$, d.h. die Matrix $S = T^{-1}$ erfüllt die Bedingung aus unserer Definition. Beide Versionen sind also äquivalent zueinander. Um jedoch in den folgenden Abschnitten beim Rechnen ein einheitliches Vorgehen zu ermöglichen, arbeiten wir ausschließlich mit der gegebenen Definition.*

Beispiel 10.6. *In Beispiel 9.28 haben wir gesehen, dass darstellende Matrizen eines fest gewählten Endomorphismus $f \in L(\mathbb{K}^n; \mathbb{K}^n)$ zu unterschiedlichen Basen äquivalent zueinander sind, denn für zwei beliebige Basen $\mathcal{C}$ und $\mathcal{D}$ des $\mathbb{K}^n$ gilt (mit den in dem Beispiel eingeführten Matrizen)*

$$A_f^{\mathcal{D}} = (C^{-1} \cdot D)^{-1} \cdot A_f^{\mathcal{C}} \cdot (C^{-1} \cdot D).$$

In diesem Fall ist also $S = C^{-1} \cdot D$.

Angenommen, A ist ähnlich zu B – ist B dann auch ähnlich zu A? Und ist jede Matrix ähnlich zu sich selbst? Diese und andere Fragen beantwortet folgender Satz:

Satz 10.7

Sei $n \in \mathbb{N}$ beliebig und $\mathbb{K}$ ein Körper. Die Ähnlichkeit von Matrizen des $\mathbb{K}^{n\times n}$ stellt eine Äquivalenzrelation auf der Menge $\mathbb{K}^{n\times n}$ dar.

Beweis. Wir benutzen zum Beweis folgende Notation: $A \sim B$, falls A zu B ähnlich ist, d.h. falls eine reguläre Matrix S existiert, so dass $B = S^{-1}AS$ ist.

Reflexivität: Sei $A \in \mathbb{K}^{n\times n}$ beliebig. Dann gilt $A \sim A$, denn $A = (E_n)^{-1}AE_n$.

Symmetrie: Es gelte $A \sim B$, d.h. A sei ähnlich zu B. Aus $B = S^{-1}AS$ folgt $SBS^{-1} = A$, d.h. $A = (S^{-1})^{-1}BS^{-1}$. Somit ist auch $B \sim A$, d.h. B ist auch ähnlich zu A.

Transitivität: Für Matrizen A, B und $C \in \mathbb{K}^{n\times n}$ gelte $A \sim B$ und $B \sim C$, d.h. es existieren reguläre Matrizen $S, T \in \mathbb{K}^{n\times n}$, so dass gilt:

$$B = S^{-1}AS \quad \text{und} \quad C = T^{-1}BT.$$

Damit gilt insgesamt durch Einsetzen der ersten in die zweite Gleichung

$$C = T^{-1}S^{-1}AST = (ST)^{-1}A(ST),$$

d.h. in der Tat ist auch $A \sim C$.

□

Satz 10.8: Eigenschaften ähnlicher Matrizen

Seien A und B ähnliche Matrizen. Dann gilt:

1. $p_A \equiv p_B$, d.h. A und B haben das gleiche charakteristische Polynom und somit die gleichen Eigenwerte;
2. $\det(A) = \det(B)$; insbesondere ist A genau dann invertierbar, wenn B invertierbar ist;
3. $\mathrm{rg}(A) = \mathrm{rg}(B)$.

Beweis. Es gelte $B = S^{-1}AS$.

1. Die Gleichheit der charakteristischen Polynome folgt mit der gleichen Rechnung, die bereits im Beweis von Satz 9.29 benutzt wurde. Da die Eigenwerte die Nullstellen der charakteristischen Polynome sind, müssen diese auch übereinstimmen.
2. Da für alle $t \in \mathbb{K}$ die Gleichheit $\det(A - tE_n) = p_A(t) = p_B(t) = \det(B - tE_n)$ gilt, folgt für die konkrete Wahl $t = 0$, dass die Determinanten von A und B gleich sind. Da A genau dann invertierbar ist, wenn $\det(A) \neq 0$ ist, was wiederum genau dann der Fall ist, wenn $\det(B) \neq 0$ ist, folgt die zusätzliche Behauptung: Entweder A und B sind beide invertierbar oder keine von beiden Matrizen ist invertierbar.
3. A ist darstellende Matrix eines Endomorphismus $f \in (\mathbb{K}^n; \mathbb{K}^n)$, nämlich von $f(x) = A \cdot x$ bezüglich der Standardbasis. Dann ist jedoch B die darstellende Matrix von f bezüglich der Basis, die aus den Spaltenvektoren von S besteht (siehe Beispiel 9.28). Daher gilt
$$\mathrm{rg}(A) = \mathrm{rg}(f) = \mathrm{rg}(B).$$

□

Beispiel 10.9. *Mit Hilfe des Satzes 10.8 können in vielen Beispielen schnell erkennen, ob Matrizen ähnlich zueinander sein können. Betrachten wir zum Beispiel*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \textit{und} \quad B = \begin{pmatrix} 2 & 4 & 6 \\ 0 & 8 & 10 \\ 0 & 0 & 12 \end{pmatrix}.$$

Es ist zwar $B = 2A$, jedoch sind beide Matrizen nicht ähnlich, denn auf ihren jeweiligen Hauptdiagonalen können wir dank der Zeilenstufenform das charakteristische Polynom (und damit auch die Eigenwerte) ablesen:

$$p_A(t) = (1-t)(4-t)(6-t), \quad \textit{aber} \quad p_B(t) = (2-t)(8-t)(12-t).$$

Da die Polynome nicht gleich sind, können die Matrizen nach Satz 10.8 nicht ähnlich sein.

Beispiel 10.10. *Die in Satz 10.8 genannten Eigenschaften sind notwendig, aber nicht hinreichend für die Ähnlichkeit zweier Matrizen. Betrachten wir dazu folgende Matrizen:*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Überprüfen wir, dass die Matrizen die drei Eigenschaften besitzen:

1. *Da beide Matrizen in Zeilenstufenform vorliegen, kann man das charakteristische Polynom leicht bestimmen:*

   $$p_A(t) = \det \begin{pmatrix} 1-t & 0 \\ 0 & 1-t \end{pmatrix} = (1-t)^2 = \det \begin{pmatrix} 1-t & 1 \\ 0 & 1-t \end{pmatrix} = p_B(t).$$

   Damit haben auch beide Matrizen den doppelten Eigenwert 1.
2. $\det(A) = \det(B) = 1$.
3. $\mathrm{rg}(A) = \mathrm{rg}(B) = 2$, *da nach 2. beide Matrizen invertierbar sind.*

Dennoch sind die Matrizen nicht ähnlich, denn für jede invertierbare Matrix $S \in \mathbb{R}^{2\times 2}$ gilt:

$$S^{-1}AS = S^{-1}E_2S = S^{-1}S = E_2 \neq B.$$

Ein Sonderfall der ähnlichen Matrizen sind die Matrizen, die zu einer Diagonalmatrix ähnlich sind.

Definition 10.11

Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *diagonalisierbar*, falls eine invertierbare Matrix $S \in \mathbb{K}^{n\times n}$ existiert, so dass $S^{-1}AS$ eine Diagonalmatrix ist, d.h. es existieren $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, so dass gilt:

$$S^{-1}AS = \mathrm{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Beispiel 10.12. *Die Matrix* $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ *ist diagonalisierbar, denn setzen wir* $S = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ *und* $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$*, so gilt*

$$\begin{aligned} AS &= \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 2 & 6 \end{pmatrix} \quad \text{und} \\ SD &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 2 & 6 \end{pmatrix} \end{aligned}$$

und aus $AS = SD$ *folgt* $S^{-1}AS = D$*, d.h.* A *ist ähnlich zur Diagonalmatrix* D*.*

Das Beispiel zeigt zweierlei: Erstens müssen wir nicht notwendigerweise die Inverse S^{-1} berechnen um zu überprüfen ob für gegebene Matrizen $S^{-1}AS = D$ gilt. Zweitens ist es zwar leicht, bei gegebener Matrix S die Beziehung zu überprüfen – wir wissen jedoch noch nicht, wie wir auf die Matrix S kommen.

Beispiel 10.13 (Fortsetzung von Beispiel 10.12)**.** *Bezeichnen wir die Spalten von* S *mit* s_1 *und* s_2*, so gilt*

$$AS = A \cdot \begin{pmatrix} s_1 & s_2 \end{pmatrix} = \begin{pmatrix} As_1 & As_2 \end{pmatrix}.$$

Zudem gilt analog

$$SD = \begin{pmatrix} s_1 & s_2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 2s_1 & 3s_2 \end{pmatrix}.$$

Aus der Gleichheit folgt sofort, dass s_1 *und* s_2 *Eigenvektoren von* A *sind, denn es gilt beispielsweise* $As_1 = 2s_1$ *und gleichzeitig ist* s_1 *nicht der Nullvektor, da die Matrix* S *invertierbar ist.*

Inspiriert von diesem Beispiel werden wir im nächsten Abschnitt zeigen, dass die Matrix S, mit deren Hilfe eine Matrix A diagonalisiert werden kann, in der Tat aus linear unabhängigen Vektoren der Matrix A bestehen muss. Dazu müssen wir uns nochmals die Eigenwerte und die zugehörigen Eigenräume genauer ansehen.

10.3 Bedingungen für die Diagonalisierbarkeit von Matrizen

Definition 10.14

Sei $A \in \mathbb{K}^{n \times n}$ und sei $\lambda \in \mathbb{K}$ ein Eigenwert von A. Die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms p_A bezeichnen wir als *algebraische Vielfachheit* von λ. Die Dimension des Eigenraums $\dim(\operatorname{Eig}(A, \lambda))$ bezeichnen wir als *geometrische Vielfachheit* von λ.

Bemerkung 10.15. *Man kann auch sagen, die algebraische Vielfachheit eines Eigenwertes* λ *sei der größte ganzzahlige Exponent* $k \in \mathbb{N}$*, für den gilt:*

$$(t - \lambda)^k \mid p_A(t),$$

d.h. $(t - \lambda)^k$ *ist noch ein Teiler des charakteristischen Polynoms von* A*.*

Notation

Bezeichnen wir die Eigenwerte mit $\lambda_1, \dots, \lambda_m$, so notieren wir im Folgenden kurz deren algebraische Vielfachheiten als $\alpha_{\lambda_1}, \dots, \alpha_{\lambda_m}$ und deren geometrische Vielfachheiten als $\beta_{\lambda_1}, \dots, \beta_{\lambda_m}$.

Konkret berechnen kann man β_{λ_k} auch über die Formel

$$\beta_{\lambda_k} = \dim(\ker(A - \lambda_k E_n)) = n - \operatorname{rg}(A - \lambda_k E_n). \tag{10.3.1}$$

Beispiel 10.16. *Sei* $A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

Um zu entscheiden, ob A diagonalisierbar ist, bestimmen wir zuerst das charakteristische Polynom:

$$\begin{aligned} p_A(t) &= \det \begin{pmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{pmatrix} \\ &= t(2+t)(3-t) + 6 + 6 - (2(2+t) + 6t + 3(3-t)) \\ &= 6t + t^2 - t^3 + 12 - (4 + 2t + 6t + 9 - 3t) \\ &= -t^3 + t^2 + t - 1 \\ &= -(t-1)^2(t+1). \end{aligned}$$

Die Eigenwerte von A sind somit $\lambda_1 = 1$ mit algebraischer Vielfachheit $\alpha_1 = 2$ und $\lambda_2 = -1$ mit $\alpha_{-1} = 1$. Betrachten wir die Eigenräume, um die geometrischen Vielfachheiten zu bestimmen:

$\lambda_1 = 1:$ *Es ist*

$$\operatorname{rg}(A - 1E_3) = \operatorname{rg} \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} = 1,$$

da alle Zeilen der Matrix Vielfache voneinander sind. Somit ist

$$\beta_1 = n - \operatorname{rg}(A - \lambda_1 E_3) = 3 - \operatorname{rg}(A - 1E_3) = 3 - 1 = 2.$$

$\lambda_2 = -1$**:** *Es ist*

$$\operatorname{rg}(A + 1E_3) = \operatorname{rg} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} = 2,$$

denn einerseits sind mindestens zwei Zeilen linear unabhängig und andererseits ist $\det(A + 1E_n) = 0$ (nach Definition des charakteristischen Polynom), so dass die Matrix nicht vollen Rang 3 haben kann. Somit ist

$$\beta_{-1} = n - \operatorname{rg}(A - (-1)\cdot E_3) = 3 - \operatorname{rg}(A + E_3) = 3 - 2 = 1.$$

Da jeder Eigenraum $\mathrm{Eig}(A, \lambda)$ mindestens einen von Null verschiedenen Vektor enthält, ist die geometrische Vielfachheit immer mindestens 1. Im Umkehrschluss gilt in Anlehnung an Satz 9.20: Ist $\mathrm{rg}(A - \lambda E_n) = n$, so ist λ kein Eigenwert von $A \in \mathbb{K}^{n \times n}$ gewesen.

Die algebraische Vielfachheit eines Eigenwertes ist auch immer mindestens 1. Zudem ist die Summe der algebraischen Vielfachheiten genau dann n, wenn sich das charakteristische Polynom in Linearfaktoren zerlegen lässt.

Das folgende Ergebnis wird uns ermöglichen, ein Kriterium für Diagonalisierbarkeit zu finden und auch die Matrix S aufzustellen, mit deren Hilfe eine diagonalisierbare Matrix A diagonalisiert werden kann.

Satz 10.17

Seien $\lambda_1, \dots, \lambda_m$ alle paarweise verschiedenen Eigenwerte der Matrix $A \in \mathbb{K}^{n \times n}$ und seien $n_1 := \beta_{\lambda_1}, \dots, n_m := \beta_{\lambda_m}$ deren geometrische Vielfachheiten. Für jedes $i \in \{1, \dots, m\}$ sei zudem $\{v_1^{(i)}, \dots, v_{n_i}^{(i)}\}$ eine Basis des Eigenraumes $\mathrm{Eig}(A; \lambda_i)$. Dann ist $\{v_1^{(1)}, \dots, v_{n_1}^{(1)}, \dots, v_1^{(m)}, \dots, v_{n_m}^{(m)}\}$ eine linear unabhängige Menge.

Beweis. Seien $\mu_j^{(i)} \in \mathbb{K}$ ($i \in \{1, \dots, m\}$, $j \in \{1, \dots, n_i\}$) derart, dass die zugehörige Linearkombination der Eigenvektoren den Nullvektor ergibt, d.h.

$$\sum_{i=1}^{m} \sum_{j=1}^{n_i} \mu_j^{(i)} v_j^{(i)} = e_{\mathbb{K}^n}.$$

Da alle Eigenvektoren zu einem festen Eigenwert im zugehörigen Eigenraum liegen und der Eigenraum ein Vektorraum ist, gilt für jedes $i \in \{1, \dots, m\}$:

$$w_i := \sum_{j=1}^{n_i} \mu_j^{(i)} v_j^{(i)} \in \mathrm{Eig}(A; \lambda_i).$$

Die Vektoren $w_1, \dots, w_m$ sind entweder Eigenvektoren oder die Nullvektoren. Wären es tatsächlich Eigenvektoren, so hätten wir eine nicht-triviale Linearkombination aus Eingenvektoren zu paarweise verschiedenen Eigenwerten, welche den Nullvektor ergibt:

$$w_1 + \ldots + w_m = e_{\mathbb{K}^n}.$$

Nach Satz 9.7 sind Eigenvektoren zu paarweise verschiedenen Eigenwerten jedoch linear unabhängig. Daraus folgt, dass $w_1 = \ldots = w_m = e_{\mathbb{K}^n}$ ist. Da für jedes $i \in \{1, \dots, m\}$ die Menge $\{v_1^{(i)}, \dots, v_{n_i}^{(i)}\}$ eine Basis des Eigenraumes und als solche linear unabhängig ist, folgt $\mu_1^{(i)} = \ldots = \mu_{n_i}^{(i)} = 0$ für alle $i \in \{1, \dots, m\}$.

Damit ist bewiesen, dass die Menge $\{v_1^{(1)}, \dots, v_{n_1}^{(1)}, \dots, v_1^{(m)}, \dots, v_{n_m}^{(m)}\}$ linear unabhängig ist. □

Hieraus können wir folgern:

Korollar 10.18

Sei $A \in \mathbb{K}^{n\times n}$ und seien $\lambda_1, \ldots, \lambda_m$ alle paarweise verschiedenen Eigenwerte von A mit zugehörigen geometrischen Vielfachheiten $\beta_{\lambda_1}, \ldots, \beta_{\lambda_m}$. Dann gilt:

1. $\beta_{\lambda_1} + \ldots + \beta_{\lambda_m} \leq n$.
2. A ist genau dann diagonalisierbar, wenn $\beta_{\lambda_1} + \ldots + \beta_{\lambda_m} = n$ ist.

Beweis. 1. Nach Satz 10.17 gibt es $\beta_{\lambda_1} + \ldots + \beta_{\lambda_m}$ linear unabhängige Vektoren im $\mathbb{K}^n$ und da es (wie im Austauschsatz 6.33 bewiesen wurde) nicht mehr als n linear unabhängige Vektoren in einem n-dimensionalen Vektorraum geben kann, folgt sofort die gewünschte Ungleichung.

2. Wenn die Summe der geometrischen Vielfachheiten n beträgt, dann gibt es nach Satz 10.17 n linear unabhängige Eigenvektoren, die folglich eine Basis des $\mathbb{K}^n$ bilden. Sei v einer dieser Eigenvektoren mit zugehörigem Eigenwert λ. Dann gilt (nach Definition eines Eigenvektors) $A\cdot v = \lambda\cdot v$. Bilden wir aus den Eigenvektoren eine Matrix und nennen wir diese S, so gilt (Vgl. Beispiel 10.13)

$$AS = S \cdot \mathrm{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}_{\beta_{\lambda_1}\text{ Spalten}}, \cdots, \underbrace{\lambda_m, \cdots, \lambda_m}_{\beta_{\lambda_m}\text{ Spalten}}).$$

Hieraus folgt, dass $S^{-1}AS$ Diagonalgestalt hat, d.h. A ist diagonalisierbar.

Sei nun A diagonalisierbar, d.h. es existiert eine invertierbare Matrix $S \in \mathbb{K}^{n\times n}$, so dass $S^{-1}AS$ eine Diagonalmatrix ist. Dann gilt für jede Spalte s_i von S, dass ein $\mu_i \in \mathbb{K}$ existiert, so dass Folgendes gilt:

$$As_i = \mu_i s_i.$$

Da s_i nicht der Nullvektor sein kann, ist s_i ein Eigenvektor von A und da zudem S invertierbar ist, muss es insgesamt n linear unabhängige Eigenvektoren von A geben, welche gerade die Spalten von S sind. Da die Summe der Dimensionen der Eigenräume die maximale Anzahl linear unabhängiger Eigenvektoren von A angibt, folgt hieraus

$$\beta_{\lambda_1} + \ldots + \beta_{\lambda_m} = n.$$

□

Beispiel 10.19. *Nach diesem Satz ist die Matrix A aus Beispiel 10.16 diagonalisierbar, denn $\beta_1 + \beta_{-1} = 2 + 1 = 3 = n$. Überprüfen wir, dass wir mit Hilfe einer Basis aus Eigenvektoren in der Tat die Matrix diagonalisieren können:*

$\lambda_1 = 1:$ *Das LGS $Ax = 1x$ bzw. $(A - 1E_n)x = 0$ führt zu*

$$\left[\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ -3 & -3 & 3 & 0 \\ -2 & -2 & 2 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

d.h. $x_3 = x_1 + x_2$*. Eine Basis von* $\mathrm{Eig}(A, 1)$ *ist somit*

$$B_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

$\lambda_2 = -1$**:** *Das LGS* $Ax = -1x$ *bzw.* $(A + 1E_n)x = 0$ *führt zu*

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -3 & -1 & 3 & 0 \\ -2 & -2 & 4 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & -4 & 6 & 0 \\ 0 & -4 & 6 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & -4 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

d.h. $x_1 - x_2 + x_3 = 0$ *und* $2x_2 - 3x_3 = 0$*. Eine Lösung von diesem (vereinfachten) LGS ist* $x^T = \begin{pmatrix} 1 & 3 & 2 \end{pmatrix}$*. Eine Basis von* $\mathrm{Eig}(A, -1)$ *ist somit*

$$B_2 = \left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}.$$

Der Kandidat für die Matrix S *ist somit*

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}.$$

Die Inverse von S *ist*

$$S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Letztlich ist in der Tat

$$\begin{aligned} S^{-1} \cdot A \cdot S &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}. \end{aligned}$$

Kennt man die Eigenwerte einer Matrix, so lässt sich in einem Spezialfall bereits ohne Untersuchung der Eigenräume sagen, ob die Matrix diagonalisierbar ist:

Korollar 10.20

Hat eine Matrix $A \in \mathbb{K}^{n \times n}$ genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.

Beweis. Das Korollar wird in der Übung bewiesen. □

Bemerkung 10.21. *Das Korollar liefert nur ein hinreichendes, jedoch kein notwendiges Kriterium für Diagonalisierbarkeit. Für jedes $n \in \mathbb{N}$ ist die Einheitsmatrix E_n beispielsweise diagonalisierbar (mit $S = S^{-1} = E_n$), obwohl sie nur einen einzigen Eigenwert hat: $\lambda = 1$. Auch A aus Beispiel 10.16 ist diagonalisierbar, obwohl sie als 3×3-Matrix nur 2 paarweise verschiedene Eigenwerte besitzt.*

Eine strengere Obergrenze für β_k als n liefert folgender Satz:

Satz 10.22

Die geometrische Vielfachheit eines Eigenwertes ist immer kleiner oder gleich der algebraischen Vielfachheit des Eigenwertes, d.h. $\beta_\lambda \leq \alpha_\lambda$ für alle Eigenwerte λ.

Beweis. Sei $\lambda \in \mathbb{K}$ ein Eigenwert einer Matrix $A \in \mathbb{K}^{n \times n}$ und seien $\alpha := \alpha_\lambda$ und $\beta := \beta_\lambda$ die algebraische bzw. geometrische Vielfachheit des Eigenwertes λ.

Da $\beta = \dim(\mathrm{Eig}(A; \lambda))$ ist, existieren β linear unabhängige Eigenvektoren von A, nennen wir sie $v_1, \dots, v_\beta$. Da $\mathrm{Eig}(A; \lambda) \subseteq \mathbb{K}^n$ ist, ist $\beta \leq n$ und wir können die Eigenvektoren zu einer Basis des $\mathbb{K}^n$ ergänzen, d.h. es existieren $w_1, \dots, w_{n-\beta} \in \mathbb{K}^n$ derart, dass

$$\mathcal{B} := (v_1, \dots, v_\beta, w_1, \dots, w_{n-\beta})$$

eine Basis des $\mathbb{K}^n$ ist. Sei nun $f\colon \mathbb{K}^n \to \mathbb{K}^n$ gegeben durch $f(x) = A \cdot x$. Bezüglich der Basis $\mathcal{B}$ (in diesem Fall ist das nicht die Standardbasis!) hat f eine darstellende Matrix folgender Form:

$$A_f^{\mathcal{B}} = \begin{pmatrix} \lambda E_\beta & B \\ O & C \end{pmatrix},$$

wobei $B \in \mathbb{K}^{\beta \times (n-\beta)}$ und $C \in \mathbb{K}^{(n-\beta) \times (n-\beta)}$ beliebige Matrizen sind und $O \in \mathbb{K}^{(n-\beta) \times \beta}$ die Nullmatrix mit den entsprechenden Dimensionen ist. Diese Darstellung gilt, da für $k \in \{1, \dots, \beta\}$ die Gleichung $f(v_k) = \lambda v_k$ erfüllt ist. Da sowohl λE_β als auch C quadratische Matrizen sind, gilt für die Determinante von $A_f^{\mathcal{B}}$ nach (E6) aus Satz 8.16:

$$\det(A_f^{\mathcal{B}}) = \det(\lambda E_\beta) \cdot \det(C).$$

Für die Berechnung des charakteristischen Polynoms wenden wir die gleiche Rechenregel an und erhalten

$$p_{A_f^{\mathcal{B}}}(t) = p_{\lambda E_\beta}(t) \cdot p_C(t) = (\lambda - t)^\beta \cdot p_C(t),$$

d.h. $(\lambda - t)^\beta \mid p_{A_f^{\mathcal{B}}}(t)$. Da nach Satz 9.29 die charakteristischen Polynome darstellender Matrizen eines Endomorphismus zu verschiedenen Basen übereinstimmen, (da diese Matrizen ähnlich zueinander sind,) gilt $p_A(t) = p_{A_f^{\mathcal{B}}}(t)$, d.h. insbesondere gilt auch

$$(\lambda - t)^\beta \mid p_A(t).$$

Nach Bemerkung 10.15 ist die geometrische Vielfachheit α die größte natürliche Zahl, für die dies gilt, d.h. $\beta \leq \alpha$ wie behauptet. □

Mit dieser neuen Ungleichung können wir noch ein weiteres Diagonalisierbarkeitskriterium formulieren.

Korollar 10.23

Seien $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ sämtliche verschiedenen Eigenwerte von $A \in \mathbb{K}^{n\times n}$. Es gelte zudem, dass das charakteristische Polynom von A über $\mathbb{K}$ in Linearfaktoren zerfällt, d.h.

$$p_A(t) = (-1)^n (t-\lambda_1)^{\alpha_{\lambda_1}} \cdot \ldots \cdot (t-\lambda_m)^{\alpha_{\lambda_m}}.$$

Für $k \in \{1, \ldots, m\}$ sei $\beta_{\lambda_k} := \dim(\mathrm{Eig}(A, \lambda_k))$ die geometrische Vielfachheit des Eigenwertes λ_k. Dann gilt:

A ist genau dann diagonalisierbar, wenn algebraische und geometrische Vielfachheiten identisch sind, d.h. wenn $\alpha_{\lambda_k} = \beta_{\lambda_k}$ ist für alle $k \in \{1, \ldots, m\}$.

Beispiel 10.24. *Nach diesem Satz ist die Matrix A aus Beispiel 10.16 diagonalisierbar, denn $\alpha_1 = 2 = \beta_1$ und $\alpha_{-1} = 1 = \beta_{-1}$.*

Beweis von Korollar 10.23. Da jedes Polynom n-ten Grades über einem beliebigen Körper *höchstens* n Nullstellen besitzt, gilt immer

$$\sum_{k=1}^{m} \alpha_{\lambda_k} \leq n.$$

Die Voraussetzung, dass das charakteristische Polynom in Linearfaktoren zerfällt, impliziert, dass Gleichheit gilt, d.h.

$$\sum_{k=1}^{m} \alpha_{\lambda_k} = n. \tag{10.3.2}$$

"$\Longleftarrow$": Nehmen wir an, es gelte $\alpha_{\lambda_k} = \beta_{\lambda_k}$ für alle $k \in \{1, \ldots, m\}$. Dann folgt sofort

$$n \overset{(10.3.2)}{=} \sum_{k=1}^{m} \alpha_{\lambda_k} \overset{\text{Annahme}}{=} \sum_{k=1}^{m} \beta_{\lambda_k}.$$

Mit Teil 2 von Korollar 10.18 folgt hieraus, dass A diagonalisierbar ist.

"$\Longrightarrow$": Nehmen wir an, A sei diagonalisierbar. Dann gilt

$$\sum_{k=1}^{m} \beta_{\lambda_k} \overset{\text{Kor. 10.18}}{=} n \overset{(10.3.2)}{=} \sum_{k=1}^{m} \alpha_{\lambda_k}$$

Da nach Satz 10.22 kein Summand der linken Summe größer sein kann als sein entsprechender Summand der rechten Summe, müssen alle Summanden identisch sein, d.h. $\alpha_{\lambda_k} = \beta_{\lambda_k}$ für alle $k \in \{1, \ldots, m\}$.

$\square$

10.3.1 Die Rolle des Körpers bei der Diagonalisierung

Ob eine Matrix diagonalisierbar ist oder nicht, hängt möglicherweise von der Wahl des zugrunde liegenden Körpers ab. Wir wissen aus Abschnitt 4.5, dass jedes Polynom n-ten

Grades in $\mathbb{C}$ genau n Nullstellen besitzt – in $\mathbb{R}$ ist das jedoch nicht gegeben. Es kann also sein, dass eine Matrix $A \in \mathbb{R}^{n\times n}$ nicht im $\mathbb{R}^{n\times n}$ diagonalisierbar ist, da ihre Eigenwerte nicht reell, sondern (mindestens teilweise) komplex sind. Einen solchen Fall sehen wir uns in folgendem Beispiel an:

Beispiel 10.25. *In Beispiel 8.6 hatten wir die Drehmatrix betrachtet, die eine Drehung im $\mathbb{R}^2$ um den Koordinatenursprung in mathematisch positivem Sinn um den Winkel α bewirkt:*

$$D_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Anschaulich ist klar, dass diese Matrix nur dann reelle Eigenwerte und -vektoren besitzen kann, wenn α ein ganzzahliges Vielfaches von π ist und in der Tat ist D_α in diesen Fällen selbst bereits eine Diagonalmatrix. Berechnen wir das charakteristische Polynom von D_α, so ergibt sich für $t \in \mathbb{R}$:

$$\begin{aligned} p_{D_\alpha}(t) &= \begin{vmatrix} \cos\alpha - t & -\sin\alpha \\ \sin\alpha & \cos\alpha - t \end{vmatrix} \\ &= (\cos\alpha - t)^2 + (\sin\alpha)^2 \\ &= t^2 - 2\cos\alpha \cdot t + 1, \end{aligned}$$

da für alle $\alpha \in \mathbb{R}$ gilt: $\cos^2\alpha + \sin^2\alpha = 1$. Mit Hilfe dieser Beziehung können wir auch die Nullstellen des charakteristischen Polynoms, d.h. die Eigenwerte von D_α, bestimmen[1]*:*

$$\begin{aligned} t_{1/2} &= \cos\alpha \pm \sqrt{\cos^2\alpha - 1} \\ &= \cos\alpha \pm \sqrt{-\sin^2\alpha} \\ &= \cos\alpha \pm i\sin\alpha. \end{aligned}$$

Da die 2×2-Matrix genau 2 verschiedene Eigenwerte in $\mathbb{C}$ besitzt, ist sie gemäß Korollar 10.20 in $\mathbb{C}$ diagonalisierbar. In $\mathbb{R}$ ist sie hingegen nicht diagonalisierbar, da D_α nicht einmal einen einzigen reellen Eigenvektor besitzt, geschweige denn zwei linear unabhängige.

Um D_α zu diagonalisieren, bestimmen wir zunächst deren Eigenräume zu den Eigenwerten $\lambda_1 = \cos\alpha + i\sin\alpha$ und $\lambda_2 = \cos\alpha - i\sin\alpha$.

$$\begin{aligned} \mathrm{Eig}(D_\alpha;\lambda_1) &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 \;\middle|\; \begin{pmatrix} -i\sin\alpha & -\sin\alpha \\ \sin\alpha & -i\sin\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 \;\middle|\; \begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 \;\middle|\; -ix = y \right\} \\ &= \left\{ x \cdot \begin{pmatrix} 1 \\ -i \end{pmatrix} \;\middle|\; x \in \mathbb{C} \right\} \end{aligned}$$

[1] Ich benutze hier die übliche Schreibweise für die Nullstellen eines quadratischen Polynoms, welche mit der p-q-Formel berechnet werden.

und analog dazu:

$$\begin{aligned}
\operatorname{Eig}(D_\alpha;\lambda_2) &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 \;\middle|\; \begin{pmatrix} i\sin\alpha & -\sin\alpha \\ \sin\alpha & i\sin\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 \;\middle|\; \begin{pmatrix} i & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 \;\middle|\; ix = y \right\} \\
&= \left\{ x \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} \;\middle|\; x \in \mathbb{C} \right\}
\end{aligned}$$

Die geometrischen Vielfachheiten ergeben tatsächlich zusammen 2, was nochmal die Diagonalisierbarkeit von D_α in $\mathbb{C}$ bestätigt.

Wir bilden zum Diagonalisieren die Matrix, die aus linear unabhängigen Eigenvektoren von D_α besteht, und deren Inverse:

$$S = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \quad \textit{und} \quad S^{-1} = \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}.$$

In der Tat gilt (unter Verwendung von $\frac{1}{i} = -i$):

$$\begin{aligned}
S^{-1} D_\alpha S &= \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \\
&= \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix} \begin{pmatrix} \cos\alpha + i\sin\alpha & \cos\alpha - i\sin\alpha \\ \sin\alpha - i\cos\alpha & \sin\alpha + i\cos\alpha \end{pmatrix} \\
&= \frac{1}{2i} \begin{pmatrix} 2i\cos\alpha - 2\sin\alpha & 0 \\ 0 & 2i\cos\alpha + 2\sin\alpha \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.
\end{aligned}$$

Bemerkung 10.26. *Wir haben anhand dieses Beispiels gesehen, dass Matrizen, die über dem Körper der reellen Zahlen nicht diagonalisierbar sind,* möglicherweise *über dem Körper der komplexen Zahlen diagonalisiert werden können. Das kann jedoch nur dann passieren, wenn die Summe der algebraischen Vielfachheiten $\sum_k \alpha_{\lambda_k} < n$ ist, d.h. wenn das charakteristische Polynom in $\mathbb{R}[x]$ nicht in Linearfaktoren zerlegt werden kann. Matrix $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ aus Beispiel 10.10 hingegen ist nicht diagonalisierbar, denn wenn sie diagonalisierbar wäre, müsste sie ähnlich zu der Diagonalmatrix sein, in deren Diagonale gerade die Eigenwerte von B stehen, d.h. in diesem Fall zur Einheitsmatrix E_2. In Beispiel 10.10 hatten wir jedoch gezeigt, dass das nicht der Fall ist, d.h. B ist weder als reelle noch als komplexwertige Matrix diagonalisierbar.*

10.3.2 Zusammenfassung der Schritte zur Diagonalisierung

Fassen wir die wichtigsten Ergebnisse zur Diagonalisierung von Matrizen noch einmal zusammen:

- Ist $A \in \mathbb{K}^{n \times n}$ diagonalisierbar und hat A die Eigenwerte $\lambda_1, \ldots, \lambda_m$, so gilt $\beta_{\lambda_1} + \ldots + \beta_{\lambda_m} = n$, d.h. die Summe der Dimensionen der Eigenräume ist n.
- Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. (Satz 9.7)
- Jede Basis des Eigenraumes $\mathrm{Eig}(A, \lambda_k)$ besteht aus β_{λ_k} Eigenvektoren, die somit linear unabhängig voneinander sind. (Satz 10.17)
- Fixiert man für jeden Eigenraum eine Basis $\mathcal{B}_k$ und vereinigt diese zu $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_m$, so erhält man n linear unabhängige Eigenvektoren von A und somit ist $\mathcal{B}$ eine Basis des $\mathbb{K}^n$.

Aus diesem Wissen können wir eine Methode zur Diagonalisierung einer Matrix $A \in \mathbb{K}^{n \times n}$ formulieren:

1. Bestimmung von $p_A(t)$, dessen Nullstellen sind die Eigenwerte von A. Falls A n verschiedene Nullstellen besitzt, so ist A diagonalisierbar und man fährt mit Schritt 3 fort. Falls nicht, ist Schritt 2 nötig.
2. Mit der Formel $\beta_\lambda = n - \mathrm{rg}(A - \lambda E_n)$ bestimmt man die geometrischen Vielfachheiten der Eigenwerte. Ist deren Summe n, so ist A diagonalisierbar. In diesem Fall geht es weiter mit Schritt 3.
3. Man bestimmt alle Eigenräume (als $\mathrm{Eig}(A; \lambda) = L(A - \lambda E_n, 0)$) und bildet eine Basis $\mathcal{B}$ von $\mathbb{K}^n$, die ausschließlich aus Eigenvektoren von A besteht.
4. Die Matrix $S \in \mathbb{K}^{n \times n}$, die entsteht, wenn man alle Vektoren von $\mathcal{B}$ aneinanderreiht, diagonalisiert A, d.h. sind $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$ die Eigenwerte, gemäß ihrer algebraischen Vielfachheit mehrfach aufgezählt, so ist $\mathrm{diag}(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n) = S^{-1} \cdot A \cdot S$.

Bemerkung 10.27. *Man kann Schritt 2 auch zu Gunsten schon Schritt 3 auslassen, da man die geometrischen Vielfachheiten auch als Dimensionen der Eigenräume ablesen kann.*

10.4 Diagonalisierbarkeit von Endomorphismen

Während wir Eigenwerte, -vektoren und -räume für quadratische Matrizen und auch für Endomorphismen definiert hatten, haben wir bei der Diagonalisierung die Endomorphismen bisher (bis auf vereinzelte Gastauftritte in Beweisen) nicht betrachtet. Das wollen wir in diesem Abschnitt nachholen.

Bisher haben wir uns die Frage gestellt: Existiert zu einer quadratischen Matrix $A \in \mathbb{K}^{n \times n}$ eine reguläre Matrix $S \in \mathbb{K}^{n \times n}$, so dass $S^{-1}AS$ eine Diagonalmatrix ist? In diesem Fall heißt A diagonalisierbar.

Nun stellen wir uns die zugehörige Frage für Endomorphismen: Existiert zu $f \in L(V;V)$ eine Basis von V, bezüglich der die darstellende Matrix von f eine Diagonalmatrix ist?

Das folgende Beispiel zeigt, dass die so formulierte Problemstellung für Endomorphismen tatsächlich zu dem uns bereits bekannten Problem der Diagonalisierung von Matrizen führt.

Beispiel 10.28. *Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch*

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 + x_3 \\ -3x_1 - 2x_2 + 3x_3 \\ -2x_1 - 2x_2 + 3x_3 \end{pmatrix}$$

d.h. $f(x) = Ax$ für A aus Beispiel 10.16. Somit ist $A = A_f^{\mathcal{B}}$, wenn wir die Standardbasis des $\mathbb{R}^3$ mit $\mathcal{B}$ bezeichnen. Berechnen wir nun die darstellende Matrix von f bezüglich der neuen Basis

$$\mathcal{C} = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right).$$

Gehen wir dabei wie in Beispiel 9.28 vor:

$$\begin{aligned}
f\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \\
f\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \\
f\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} &= \begin{pmatrix} -1 \\ -3 \\ -2 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}.
\end{aligned}$$

Somit ist $A_f^{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Dies hätten wir auch durch Matrizenmultiplikation berechnen können. Dazu bezeichnen wir die Matrix mit den Vektoren aus $\mathcal{C}$ mit C und berechnen

$$A_f^{\mathcal{C}} = C^{-1}AC.$$

Diese Rechnung haben wir bereits in Beispiel 10.19 durchgeführt und es kommt in der Tat die Diagonalmatrix mit den Eigenwerten von A in der Diagonale heraus.

In der Literatur existieren verschiedene Definitionen dafür, wann ein Endomorphismus *diagonalisierbar* genannt wird. Im Folgenden stellen wir daher zwei Definitionen (angepasst an unsere Notation) vor, die jedoch zueinander äquivalent sind.

Definition 10.29: aus Liesen & Mehrmann (2015) [LM15]

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum. Ein Endomorphismus $f \in L(V;V)$ heißt *diagonalisierbar*, wenn es eine Basis $\mathcal{B}$ von V gibt, so dass $A_f^{\mathcal{B}}$ eine Diagonalmatrix ist.

Definition 10.30: aus Fischer (2011) [Fis17]

Sei V ein $\mathbb{K}$-Vektorraum mit $\dim(V) = n < \infty$ und $f\colon V \to V$ linear. f heißt *diagonalisierbar*, wenn V eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ besitzt, die aus Eigenvektoren von f besteht.

Überlegen wir uns kurz, dass tatsächlich beide Definitionen äquivalent sind:

Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine solche Basis von V, für die $A_f^{\mathcal{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ ist, so gilt $f(v_k) = \lambda_k v_k$ für alle $k \in \{1, \ldots, n\}$, d.h. alle Vektoren aus $\mathcal{B}$ sind (bekanntermaßen linear unabhängige) Eigenvektoren von f, die zusammen eine Basis von V bilden.

Andersherum folgt, wenn $\mathcal{B} = (v_1, \ldots, v_n)$ linear unabhängige Eigenvektoren von f sind, dass $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ existieren, so dass $f(v_k) = \lambda_k v_k$ für alle $k \in \{1, \ldots, n\}$ gilt, woraus $A_f^{\mathcal{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ folgt.

Um einen Endomorphismus auf Diagonalisierbarkeit zu prüfen, wählen wir zunächst eine beliebige Basis des zu Grunde liegenden Vektorraumes $\mathcal{B}$ und bestimmen die darstellende Matrix $A_f^{\mathcal{B}}$. Für diese wenden wir anschließend das bereits bekannte Verfahren zur Diagonalisierung von Matrizen an.

Beispiel 10.31. *Betrachten wir den Endomorphismus $f \in L(\mathbb{R}^{2\times 2}; \mathbb{R}^{2\times 2})$, der jeder Matrix $A \in \mathbb{R}^{2\times 2}$ ihre Transponierte $A^T \in \mathbb{R}^{2\times 2}$ zuordnet, d.h.*

$$f\begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \qquad a, b, c, d \in \mathbb{R}.$$

Überprüfen wir zunächst, dass diese Abbildung in der Tat linear ist:

$$\begin{aligned} f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix}\right) &= f\left(\begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}\right) = \begin{pmatrix} a+e & c+g \\ b+f & d+h \end{pmatrix} \\ &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} + \begin{pmatrix} e & g \\ f & h \end{pmatrix} = f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) + f\left(\begin{pmatrix} e & f \\ g & h \end{pmatrix}\right) \end{aligned}$$

oder kürzer: $(A+B)^T = A^T + B^T$. Zudem gilt für jeden Skalar $\mu \in \mathbb{R}$:

$$f\left(\mu \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = f\left(\begin{pmatrix} \mu a & \mu b \\ \mu c & \mu d \end{pmatrix}\right) = \begin{pmatrix} \mu a & \mu c \\ \mu b & \mu d \end{pmatrix} = \mu \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \mu \cdot f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right).$$

Da f linear ist, können wir eine darstellende Matrix bezüglich einer beliebigen Basis des $\mathbb{R}^{2\times 2}$ finden. Wählen wir die Standardbasis des $\mathbb{R}^{2\times 2}$:

$$\mathcal{B} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

Dann gilt

$$A_f^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Berechnen wir zunächst das charakteristische Polynom von $A := A_f^{\mathcal{B}}$*:*

$$\begin{aligned} p_A(t) = \det(A - tE_4) &= \begin{vmatrix} 1-t & 0 & 0 & 0 \\ 0 & -t & 1 & 0 \\ 0 & 1 & -t & 0 \\ 0 & 0 & 0 & 1-t \end{vmatrix} = (1-t) \cdot \begin{vmatrix} -t & 1 \\ 1 & -t \end{vmatrix} \cdot (1-t) \\ &= (1-t)^2 \cdot (t^2 - 1) = (1-t)^2 \cdot (t-1) \cdot (t+1) \end{aligned}$$

Somit hat A *die Eigenwerte* ± 1 *mit algebraischen Vielfachheiten* $\alpha_1 = 3$ *und* $\alpha_{-1} = 1$*. Die zugehörigen geometrischen Vielfachheiten sind*

$$\beta_1 = 4 - \operatorname{rg} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 4 - 1 = 3,$$

$$\beta_{-1} = 4 - \operatorname{rg} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = 4 - 3 = 1.$$

Mit Korollar 10.18 oder Korollar 10.23 kann man schließen, dass A *diagonalisierbar ist. Dazu berechnen wir nun die Eigenräume und schreiben sie als lineare Hüllen ihrer Basen:*

$$\begin{aligned} \operatorname{Eig}(A; 1) &= \left\{ x \in \mathbb{R}^4 \,\middle|\, (A - E_4)x = e_{\mathbb{R}^4} \right\} = \left\{ x \in \mathbb{R}^4 \,\middle|\, x_2 = x_3 \right\} \\ &= \mathcal{L}\left(\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right), \\ \operatorname{Eig}(A; -1) &= \left\{ x \in \mathbb{R}^4 \,\middle|\, (A + E_4)x = e_{\mathbb{R}^4} \right\} = \left\{ x \in \mathbb{R}^4 \,\middle|\, x_1 = x_4 = 0, x_2 = -x_3 \right\} \\ &= \mathcal{L}\left(\left\{ \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\} \right). \end{aligned}$$

Damit ist[2]

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \textit{und} \quad S^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1/2 & 1/2 & 0 \end{pmatrix}.$$

[2]Die Berechnung der Inversen kann mit Hilfe des Gauß-Algorithmus oder mit der Determinantenregel erfolgen. Wer die eine oder andere Methode noch einmal üben möchte, kann dieses Beispiel dafür nutzen.

Mit Hilfe dieser Matrizen können wir die Matrix $A = A_f^{\mathcal{B}}$ diagonalisieren:

$$\begin{aligned} S^{-1}AS &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1/2 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1/2 & -1/2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

Kommen wir nun zurück zum ursprünglichen Endomorphismus. Mit Hilfe der Basis aus Eigenvektoren von A bestimmen wir die zur Diagonalisierung von f geeigneten Basis:

$$\mathcal{C} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right).$$

Die vier Matrizen sind leicht als linear unabhängig erkennbar, da keine Matrix als Linearkombination der anderen dargestellt werden kann. Folglich ist dies tatsächlich eine Basis des Vektorraumes $\mathbb{R}^{2\times 2}$. Zudem ist $A_f^{\mathcal{C}} = S^{-1}A_f^{\mathcal{B}}S$ in der Tat diagonal, denn auch ohne die Matrix S lässt sich ausrechnen:

$$\begin{aligned} f\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) &= 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\ f\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) &= 1 \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ f\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) &= 1 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ f\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) &= (-1) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

Die Matrizen aus $\mathcal{C}$ sind also in der Tat Eigenvektoren *von f.*

10.5 Ausblick: Die Jordan'sche Normalform

In diesem Abschnitt verzichten wir auf die Beweise – für eine ausführliche Behandlung dieses Themas sei beispielsweise auf Abschnitt 16.2 in [LM15] oder 4.2.5-4.2.7 in [Fis17] verwiesen.

Wir haben in der Vorlesung und in den Übungen bereits mehrere nicht diagonalisierbare Matrizen gesehen. Die Jordan'sche Normalform dieser Matrizen ist gewissermaßen das Nächstbeste nach einer Diagonalmatrix, was für die gegebenen Matrizen machbar ist –

konkret bedeutet das, dass auch direkt über der Hauptdiagonalen Einsen stehen dürfen, überall sonst jedoch nur Nullen.

Sehen wir uns zwei Beispiele an, wie die Jordan'sche Normalform konkret aussehen kann:

Beispiel 10.32. *Die Matrix*

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$$

ist nicht diagonalisierbar, denn

- *das charakteristische Polynom*

$$p_A(t) = \begin{vmatrix} 3-t & 4 & 3 \\ -1 & -t & -1 \\ 1 & 2 & 3-t \end{vmatrix} = -(t-2)^3$$

  hat die 3*-fache Nullstelle* $\lambda = 2$*, d.h.* $\alpha_2 = 3$*;*

- *die geometrische Vielfachheit von* $\lambda = 2$ *ist*

$$\beta_2 = 3 - \mathrm{rg} \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} = 3 - 2 = 1.$$

Mit den Matrizen

$$S = \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \quad \text{und} \quad S^{-1} = \frac{1}{4} \begin{pmatrix} -1 & -3 & 0 \\ 2 & 2 & 0 \\ 0 & 4 & 4 \end{pmatrix}$$

erhalten wir die sogenannte Jordan'sche Normalform der Matrix A*:*

$$S^{-1}AS = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Beispiel 10.33. *Die Matrix*

$$A = \begin{pmatrix} 7 & 1 & -8 & -1 \\ 0 & 3 & 0 & 0 \\ 4 & 2 & -5 & -1 \\ 0 & -4 & 0 & -1 \end{pmatrix}$$

ist ebenfalls nicht diagonalisierbar, denn

- *das charakteristische Polynom*

$$\begin{aligned} p_A(t) &= \begin{vmatrix} 7-t & 1 & -8 & -1 \\ 0 & 3-t & 0 & 0 \\ 4 & 2 & -5-t & -1 \\ 0 & -4 & 0 & -1-t \end{vmatrix} = (3-t) \begin{vmatrix} 7-t & -8 & -1 \\ 4 & -5-t & -1 \\ 0 & 0 & -1-t \end{vmatrix} \\ &= (3-t)(-1-t) \begin{vmatrix} 7-t & -8 \\ 4 & -5-t \end{vmatrix} = (t-3)(t+1)(t^2-2t-3) \\ &= (t-3)^2(t+1)^2 \end{aligned}$$

  hat die doppelten Nullstellen 3 *und* -1*, d.h.* $\alpha_{-1} = \alpha_3 = 2$*;*

- *die geometrischen Vielfachheiten sind*

$$\beta_{-1} = 4 - \operatorname{rg}\begin{pmatrix} 8 & 1 & -8 & -1 \\ 0 & 4 & 0 & 0 \\ 4 & 2 & -4 & -1 \\ 0 & -4 & 0 & 0 \end{pmatrix} = 4 - 3 = 1 < \alpha_{-1},$$

$$\beta_{3} = 4 - \operatorname{rg}\begin{pmatrix} 4 & 1 & -8 & -1 \\ 0 & 0 & 0 & 0 \\ 4 & 2 & -8 & -1 \\ 0 & -4 & 0 & -4 \end{pmatrix} = 4 - 3 = 1 < \alpha_{3}.$$

Mit den Matrizen

$$S = \begin{pmatrix} -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad \textit{und} \quad S^{-1} = \begin{pmatrix} 1 & 1 & -2 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

erhalten wir die Jordan'sche Normalform von A:

$$S^{-1}AS = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Auf der Diagonalen stehen also nach wie vor die Eigenwerte der Matrix entsprechend ihren algebraischen Vielfachheiten und darüber können Einsen stehen. In der Tat ist die Jordan'sche Normalform eine Blockmatrix, deren Blöcke die Form

$$B = \begin{pmatrix} \lambda & * & 0 & \cdots & 0 \\ 0 & \lambda & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & * \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix},$$

wobei $*$ für eine 0 oder 1 stehen kann (je nachdem, wie viel größer die algebraische als die geometrische Vielfachheit von λ ist).

Nehmen wir an, für eine Matrix $A \in \mathbb{K}^{n\times n}$ gelte $B = S^{-1}AS$ und bezeichnen wir die Spalten von S mit $v_1, \dots, v_n$.

Steht über dem λ eine 0 (oder gar nichts), so ist der zugehörige Spaltenvektor von S ein Eigenvektor von A – in unserem Fall trifft das insbesondere auf v_1 zu.

Steht über dem zweiten λ eine 1, so folgt aus $AS = SB$, dass

$$Av_2 = v_1 + \lambda v_2, \quad \text{also} \quad (A - \lambda E_n)v_2 = v_1$$

sein muss. Da v_1 ein Eigenvektor ist, gilt folglich

$$(A - \lambda E_n)^2 v_2 = (A - \lambda E_n)v_1 = e_{\mathbb{R}^n}.$$

Steht über dem λ in der 3. Spalte ebenfalls eine 1, so gilt auch in diesem Fall

$$Av_3 = v_2 + \lambda v_3, \quad \text{also} \quad (A - \lambda E_n)v_3 = v_2.$$

Durch die vorherige Rechnung sehen wir also, dass

$$(A - \lambda E_n)^3 v_3 = (A - \lambda E_n)^2 v_2 = (A - \lambda E_n)v_1 = e_{\mathbb{R}^n}$$

ist.

Wie lange können wir diese Prozedur fortsetzen? Wenn das charakteristische Polynom in Linearfaktoren zerfällt, so wäre es wünschenswert, zu jedem Eigenwert λ genau α_λ Vektoren zu finden, da diese zusammen dann eine quadratische Matrix ergeben. Analog zum Eigenraum definieren wir somit den *Hauptraum*:

Definition 10.34

Eine Matrix $A \in \mathbb{K}^{n\times n}$ habe einen Eigenwert λ mit algebraischer Vielfachheit $\alpha_\lambda \in \mathbb{N}$. Dann ist der *Hauptraum* von A zum Eigenwert λ definiert als

$$\mathrm{Hau}(A;\lambda) := \left\{ v \in \mathbb{K}^n \,\middle|\, (A - \lambda E_n)^{\alpha_\lambda} v = e_{\mathbb{K}^n} \right\}.$$

Bemerkung 10.35. *Es gilt folgende Ungleichungskette für einen EW λ von $A \in \mathbb{K}^{n\times n}$:*

$$1 \le \beta_\lambda \le \alpha_\lambda = \dim(\mathrm{Hau}(A;\lambda)) \le n.$$

Die Dimension des Hauptraumes werden wir an anderer Stelle begründen.

Um mit oben beschriebener Strategie also genau α_λ Vektoren zu bekommen, müssen wir mit einem Vektor starten, der im Hauptraum liegt, für den jedoch

$$(A - \lambda E_n)^{\alpha_\lambda - 1} v \neq e_{\mathbb{R}^n}$$

gilt. Anhand der zwei Beispiele vom Beginn dieses Abschnittes wollen wir diese Strategie veranschaulichen. Auf einen formalen Beweis verzichten wir.

Beispiel 10.36 (Fortsetzung von Beispiel 10.32)**.** *Berechnen wir zunächst den Eigenraum zum einzigen Eigenwert $\lambda = 2$. Dazu zunächst eine Nebenrechnung:*

$$\left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ -1 & -2 & -1 & 0 \\ 1 & 2 & 1 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ -1 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Hiermit ist

$$\begin{aligned} \mathrm{Eig}(A;2) &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, x = z, y = -z \right\} = \mathcal{L}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}. \end{aligned}$$

Indem wir $A - 2E_3$ quadrieren, kommen wir dem Hauptraum um einen Schritt näher:

$$(A-2E_3)^2 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix},$$

d.h. wir bekommen den Hilfsraum

$$\begin{aligned} H := & \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ = & \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; y = -z \right\} = \mathcal{L}\left(\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\} \right). \end{aligned}$$

Da $(A - 2E_3)^3$ die Nullmatrix ist, ist der Hauptraum

$$\mathrm{Hau}(A;2) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \mathbb{R}^3.$$

Wie beschrieben wählen wir nun einen Vektor, der zwar im Hauptraum, jedoch nicht in H liegt, zum Beispiel $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in (\mathrm{Hau}(A;2) \setminus H)$. Dann ist

$$\begin{aligned} v_2 &= (A-2E_3)v_3 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \in (H \setminus \mathrm{Eig}(A;2)) \quad \textit{und} \\ v_1 &= (A-2E_3)^2 v_3 = (A-2E_3)v_2 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} \in \mathrm{Eig}(A;2). \end{aligned}$$

Die drei Vektoren v_1, v_2 und v_3 ergeben zusammen gerade die Matrix S, die A in die Jordan'sche Normalform transformiert.

Beispiel 10.37 (Fortsetzung von Beispiel 10.33). *Da A zwei Eigenwerte mit jeweils algebraischer Vielfachheit 2 besitzt, müssen wir beide Eigen- und Haupträume berechnen. Dazu quadrieren wir zunächst die Matrix $(A - \lambda E_4)$ für $\lambda \in \{-1, 3\}$:*

$$\begin{aligned} (A+E_4)^2 &= \begin{pmatrix} 8 & 1 & -8 & -1 \\ 0 & 4 & 0 & 0 \\ 4 & 2 & -4 & -1 \\ 0 & -4 & 0 & 0 \end{pmatrix} \begin{pmatrix} 8 & 1 & -8 & -1 \\ 0 & 4 & 0 & 0 \\ 4 & 2 & -4 & -1 \\ 0 & -4 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 32 & 0 & -32 & 0 \\ 0 & 16 & 0 & 0 \\ 16 & 8 & -16 & 0 \\ 0 & -16 & 0 & 0 \end{pmatrix}, \\ (A-3E_4)^2 &= \begin{pmatrix} 4 & 1 & -8 & -1 \\ 0 & 0 & 0 & 0 \\ 4 & 2 & -8 & -1 \\ 0 & -4 & 0 & -4 \end{pmatrix} \begin{pmatrix} 4 & 1 & -8 & -1 \\ 0 & 0 & 0 & 0 \\ 4 & 2 & -8 & -1 \\ 0 & -4 & 0 & -4 \end{pmatrix} = \begin{pmatrix} -16 & -8 & 32 & 8 \\ 0 & 0 & 0 & 0 \\ -16 & -8 & 32 & 8 \\ 0 & 16 & 0 & 16 \end{pmatrix} \end{aligned}$$

Lösen wir als Nebenrechnung die homogenen LGS:

$$\left[\begin{array}{cccc|c} 8 & 1 & -8 & -1 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 4 & 2 & -4 & -1 & 0 \\ 0 & -4 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 8 & 1 & -8 & -1 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 32 & 0 & -32 & 0 & 0 \\ 0 & 16 & 0 & 0 & 0 \\ 16 & 8 & -16 & 0 & 0 \\ 0 & -16 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 2 & 0 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 4 & 1 & -8 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 4 & 2 & -8 & -1 & 0 \\ 0 & -4 & 0 & -4 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 4 & 1 & -8 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} -16 & -8 & 32 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -16 & -8 & 32 & 8 & 0 \\ 0 & 16 & 0 & 16 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 2 & 1 & -4 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & 4 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right]$$

Somit sind die gesuchten Räume:

$$\begin{aligned}
\operatorname{Eig}(A;-1) &= \left\{ v \in \mathbb{R}^4 \,\middle|\, (A+E_4)v = e_{\mathbb{R}^4} \right\} = \left\{ v \in \mathbb{R}^4 \,\middle|\, v_1 = v_3, v_2 = v_4 = 0 \right\} \\
&= \mathcal{L} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\
\operatorname{Hau}(A;-1) &= \left\{ v \in \mathbb{R}^4 \,\middle|\, (A+E_4)^2 v = e_{\mathbb{R}^4} \right\} = \left\{ v \in \mathbb{R}^4 \,\middle|\, v_1 = v_3, v_2 = 0 \right\} \\
&= \mathcal{L} \left(\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right), \\
\operatorname{Eig}(A;3) &= \left\{ v \in \mathbb{R}^4 \,\middle|\, (A-3E_4)v = e_{\mathbb{R}^4} \right\} = \left\{ v \in \mathbb{R}^4 \,\middle|\, v_1 = 2v_3, v_2 = v_4 = 0 \right\} \\
&= \mathcal{L} \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\
\operatorname{Hau}(A;3) &= \left\{ v \in \mathbb{R}^4 \,\middle|\, (A-3E_4)^2 v = e_{\mathbb{R}^4} \right\} = \left\{ v \in \mathbb{R}^4 \,\middle|\, v_1 = 2v_3 + v_4, v_2 = -v_4 \right\} \\
&= \mathcal{L} \left(\left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\} \right)
\end{aligned}$$

Nennen wir unsere gesuchten Vektoren w_1, w_2, w_3, w_4, so haben wir:

$$w_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in (\mathrm{Hau}(A; -1) \setminus \mathrm{Eig}(A; -1)),$$

$$w_1 = (A + E_4)w_2 = \begin{pmatrix} 8 & 1 & -8 & -1 \\ 0 & 4 & 0 & 0 \\ 4 & 2 & -4 & -1 \\ 0 & -4 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \in \mathrm{Eig}(A; -1),$$

$$w_4 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \in (\mathrm{Hau}(A; 3) \setminus \mathrm{Eig}(A; 3)),$$

$$w_3 = (A - 3E_4)w_4 = \begin{pmatrix} 4 & 1 & -8 & -1 \\ 0 & 0 & 0 & 0 \\ 4 & 2 & -8 & -1 \\ 0 & -4 & 0 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \in \mathrm{Eig}(A; 3).$$

Damit ist

$$S = \begin{pmatrix} -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad \textit{und} \quad S^{-1} = \begin{pmatrix} 1 & 1 & -2 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

und wir erhalten die Jordan'sche Normalform

$$\begin{aligned} S^{-1}AS &= \begin{pmatrix} 1 & 1 & -2 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 7 & 1 & -8 & -1 \\ 0 & 3 & 0 & 0 \\ 4 & 2 & -5 & -1 \\ 0 & -4 & 0 & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 & 2 & 1 \\ 0 & -1 & 0 & -1 \\ 3 & 2 & -3 & 0 \\ 0 & -3 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}. \end{aligned}$$

Bemerkung 10.38. *An Stelle der Vektoren w_2 und w_4 hätte man jeden beliebigen Vektor des jeweiligen Hauptraumes wählen können, vorausgesetzt dieser ist nicht gleichzeitig im Eigenraum. Zudem entscheidet die Reihenfolge der Vektoren über das Aussehen der Jordan'schen Normalform. Hätten wir beispielsweise die Matrizen*

$$S = \begin{pmatrix} -2 & 1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & -1 & -1 \\ 0 & -1 & 0 & 1 \end{pmatrix} \quad \textit{und} \quad S^{-1} = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -2 & -1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

gewählt, so erhielten wir

$$S^{-1}AS = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

10.6 Aufgaben

Aufgabe 10.1. *Bestimmen Sie zu folgenden Matrizen algebraische und geometrische Vielfachheiten ihrer Eigenwerte.*

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 2 & -2 & 0 \end{pmatrix}$$

Dazu müssen Sie der Reihe nach Folgendes berechnen:

a) das charakteristische Polynom,

b) die Nullstellen des charakteristischen Polynoms ($\rightsquigarrow$ algebr. Vielfachheiten der EWe),

c) die Eigenräume zu den gefundenen Eigenwerten,

d) die Dimensionen der Eigenräume ($\rightsquigarrow$ geom. Vielfachheiten der EWe).

Aufgabe 10.2. *Beweisen Sie die Aussage von Korollar 10.20:*

Besitzt eine Matrix $A \in \mathbb{K}^{n\times n}$ (mit $n \in \mathbb{N}$) genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.

Aufgabe 10.3. *Entscheiden Sie mit* möglichst geringem Rechenaufwand*, ob die folgenden Matrizen diagonalisierbar sind:*

$$A = \begin{pmatrix} 5 & -1 & 2 \\ 0 & 2 & -4 \\ 0 & 0 & 10 \end{pmatrix}, \quad B = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 7 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 2 & 7 & 4 \\ 3 & 6 & 0 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Aufgabe 10.4. *Diagonalisieren Sie die Matrix*

$$B = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 2 & -2 & 0 \end{pmatrix}$$

aus Aufgabe 10.1.

Aufgabe 10.5. *Wir bezeichnen im Folgenden mit $\mathcal{P}_4$ den Vektorraum der reellen Polynome, deren Grad höchstens 4 beträgt. Sei $f\colon \mathcal{P}_4 \to \mathcal{P}_4$ gegeben durch die Vorschrift $f(p) := p'$.*

a) Sei $p(x) = \sum_{k=0}^{4} a_k x^k \in \mathcal{P}_4$. Bestimmen Sie $f(p)$.

b) Erinnerung: Wie zeigt man, dass f ein Endomorphismus ist?

c) Bestimmen Sie die darstellende Matrix von f bezüglich der Standardbasis $\mathcal{B} = (x^4, x^3, x^2, x, 1)$ des $\mathcal{P}_4$.

d) Überprüfen Sie, ob f diagonalisierbar ist.

Aufgabe 10.6. *Sei $A \in \mathbb{K}^{n\times n}$ regulär.*

a) Zeigen Sie (zur Wiederholung), dass 0 kein Eigenwert von A ist.

b) Zeigen Sie: Ist $\lambda \in \mathbb{K}$ ein Eigenwert von A, dann ist λ^{-1} ein Eigenwert von A^{-1}.

Aufgabe 10.7. *Gegeben ist folgende Matrix:*

$$A = \begin{pmatrix} 0 & -3 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

a) *Welchen Eigenwert λ hat A? Bestimmen Sie algebraische und geometrische Vielfachheit. Können Sie damit die Jordan'sche Normalform von A vorhersagen?*

b) *Berechnen Sie A^2 und A^3.*

c) *Finden Sie einen Vektor $v_3 \in \mathbb{R}^3$, für den gilt:*

$$A^3 v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \text{ aber } A^2 v_3 \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

d) *Berechnen Sie $v_2 := Av_3$ und $v_1 := Av_2$.*

e) *Überprüfen Sie, dass für die Matrix S mit den Spalten v_1, v_2, v_3 das Produkt $S^{-1}AS$ die Jordan'sche Normalform von A liefert.*

Kapitel 11

Euklidische und unitäre Vektorräume

Ziele

- □ die Begriffe Bilinearform, Skalarprodukt und Norm kennen und anwenden können auf Beispiele und in Beweisen
- □ mehrere verschiedene Skalarprodukte und Normen kennen
- □ die Cauchy-Schwarz'sche Ungleichung kennen
- □ den Bezug zwischen dem Standard-Skalarprodukt, der Euklidischen Norm und dem Winkelbegriff erkennen
- □ den Begriff der Isometrie kennen und auf Beispiele anwenden können
- □ Orthogonal- und Orthonormalsystem als Begriffe unterscheiden können
- □ überprüfen können, ob eine Menge eine ONB ist
- □ eine Basis mit Hilfe des Gram-Schmidt-Verfahrens in eine ONB umwandeln können

Ein wichtiger Begriff der Geometrie ist bisher in dieser Vorlesung noch nicht aufgetaucht: Orthogonalität. Wir haben Vektoren auf lineare Unabhängigkeit untersucht, jedoch niemals geschaut, ob sie senkrecht aufeinander stehen. Dieser Problematik werden wir uns in diesem Kapitel widmen und dabei über die anschauliche Vorstellung von senkrecht aufeinander stehenden Objekten hinausgehen.

Zudem wollen wir sehen wie wir insbesondere im $\mathbb{R}^n$, aber auch in allgemeinen Vektorräumen, einen Längen- und Abstandsbegriff bekommen.

11.1 Winkel, Längen und Abstände – die Begriffe Skalarprodukt und Norm

11.1.1 Bilinearformen

In Kapitel 7 haben wir lineare Abbildungen als Vektorraum-Homomorphismen kennen gelernt, d.h. f ist eine lineare Abbildung zwischen zwei $\mathbb{K}$-Vektorräumen V und W, in Zeichen $f \in L(V; W)$, wenn folgende Eigenschaften erfüllt sind:

$$\begin{aligned} f(v + v') &= f(v) + f(v'), && \forall v, v' \in V, \\ f(\lambda \cdot v) &= \lambda \cdot f(v), && \forall v \in V, \forall \lambda \in \mathbb{K}. \end{aligned}$$

Im Gegensatz zu linearen Abbildungen sind Bilinearformen Abbildungen, die immer zwei Argumente haben und die, wenn man eines von beiden festhält, in Bezug auf das übrige Argument linear sind. Zudem ist der Zielraum kein beliebiger Vektorraum, sondern der dem Ausgangsraum (bzw. den Ausgangsräumen) zu Grunde liegende Körper. Formalisieren wir diese Idee:

Definition 11.1

Seien V, W zwei $\mathbb{K}$-Vektorräume. Eine Abbildung $f\colon V \times W \to \mathbb{K}$ heißt *Bilinearform*, falls gilt:

$$\begin{aligned} f(v + v', w) &= f(v, w) + f(v', w), \quad f(\lambda v, w) = \lambda f(v, w), \quad \forall v, v' \in V, \forall w \in W, \forall \lambda \in \mathbb{K}, \\ f(v, w + w') &= f(v, w) + f(v, w'), \quad f(v, \lambda w) = \lambda f(v, w), \quad \forall v \in V, \forall w, w' \in W, \forall \lambda \in \mathbb{K}. \end{aligned}$$

Gilt zudem $V = W$, so heißt eine Bilinearform *symmetrisch*, falls gilt:

$$f(v, v') = f(v', v), \quad \forall v, v' \in V.$$

Beispiel 11.2. *Sei $I = [0, 1]$ das abgeschlossene Intervall von 0 bis 1 und sei $\mathcal{C}(I; \mathbb{R})$ die Menge der stetigen Funktionen auf diesem Intervall. Dann wird durch*

$$s\colon \mathcal{C}(I; \mathbb{R}) \times \mathcal{C}(I; \mathbb{R}) \to \mathbb{R}, \qquad s(f, g) := \int_0^1 f(x)g(x)\mathrm{d}x$$

eine symmetrische Bilinearform definiert. Weisen wir die notwendigen Eigenschaften nach.

(i) Seien $f, g, h \in \mathcal{C}(I; \mathbb{R})$. Dann gilt:

$$\begin{aligned} s(f + g, h) &= \int_0^1 (f + g)(x)h(x)\mathrm{d}x = \int_0^1 [f(x) + g(x)]h(x)\mathrm{d}x \\ &= \int_0^1 [f(x)h(x) + g(x)h(x)]\mathrm{d}x = \int_0^1 f(x)h(x)\mathrm{d}x + \int_0^1 g(x)h(x)\mathrm{d}x \\ &= s(f, h) + s(g, h). \end{aligned}$$

(ii) Seien $f, g \in \mathcal{C}(I; \mathbb{R})$ und $\lambda \in \mathbb{R}$. Dann gilt:

$$s(\lambda f, g) = \int_0^1 (\lambda f)(x)g(x)\mathrm{d}x = \int_0^1 \lambda f(x)g(x)\mathrm{d}x = \lambda \int_0^1 f(x)g(x)\mathrm{d}x = \lambda s(f, g).$$

(iii) Seien $f, g \in \mathcal{C}(I; \mathbb{R})$. Dann gilt:

$$s(f,g) = \int_0^1 f(x)g(x)\mathrm{d}x = \int_0^1 g(x)f(x)\mathrm{d}x = s(g,f).$$

(iv) Die Linearität bezüglich des zweiten Arguments kann man entweder direkt nachweisen oder mit Hilfe der bereits nachgewiesenen Eigenschaften. Hier wählen wir letzteren Zugang. Dazu seien wieder $f, g, h \in \mathcal{C}(I; \mathbb{R})$ und $\lambda \in \mathbb{R}$. Dann gilt:

$$\begin{aligned} s(f, g+h) &\overset{\text{(iii)}}{=} s(g+h, f) \overset{\text{(i)}}{=} s(g,f) + s(h,f) \overset{\text{(iii)}}{=} s(f,g) + s(f,h); \\ s(f, \lambda g) &\overset{\text{(iii)}}{=} s(\lambda g, f) \overset{\text{(ii)}}{=} \lambda s(g,f) \overset{\text{(iii)}}{=} \lambda s(f,g). \end{aligned}$$

Eine abschließende Bemerkung noch zu der Analysis im Hintergrund dieses Beispiels:

Da jede stetige Funktion auf dem kompakten Intervall I integrierbar ist und das Produkt stetiger Funktionen auf I selbst auch stetig, also integrierbar ist, haben wir mit s tatsächlich eine sinnvolle Abbildung definiert.

In den Übungen werden wir ein Beispiel für eine Bilinearform kennen lernen, bei der $V \neq W$ ist. Im weiteren Verlauf der Vorlesung beschränken wir uns jedoch auf den Sonderfall $V = W$.

Darstellende Matrizen von Bilinearformen im Spezialfall $V = W$

Da Bilinearformen nicht in beliebige Vektorräume, sondern in Körper abbilden, ist eine 1:1-Übernahme der Definition einer darstellenden Matrix von den linearen Abbildungen nicht sinnvoll. Überlegen wir uns daher zunächst, welche Matrix diese Rolle einnehmen sollte.

Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine (geordnete) Basis eines $\mathbb{K}$-Vektorraumes V und sei $f\colon V \times V \to \mathbb{K}$ eine Bilinearform. Können wir mit Hilfe der Bilder von Paaren von Basisvektoren bereits alle Werte von f berechnen? Sei dazu

$$\alpha_{k,j} := f(v_k, v_j), \quad k, j \in \{1, \ldots, n\}.$$

Seien nun $b, c \in V$ beliebige Vektoren mit ihren Koordinaten

$$\kappa(b) =: \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \ \kappa(c) =: \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \in \mathbb{K}^n$$

bezüglich $\mathcal{B}$, d.h. es gelte

$$b = \sum_{k=1}^n \beta_k v_k \quad \text{und} \quad c = \sum_{j=1}^n \gamma_j v_j.$$

Dann gilt

$$\begin{aligned} f(b,c) &= f\left(\sum_{k=1}^n \beta_k v_k, \sum_{j=1}^n \gamma_j v_j\right) = \sum_{k=1}^n \sum_{j=1}^n \beta_k \gamma_j f(v_k, v_j) \\ &= \sum_{k=1}^n \sum_{j=1}^n \beta_k \gamma_j \alpha_{k,j} \end{aligned}$$

Führen wir die Matrix $A := (\alpha_{k,j}) \in \mathbb{K}^{n\times n}$ ein, so gilt also

$$f(b,c) = \kappa(b)^T \cdot A \cdot \kappa(c).$$

Ist $V = \mathbb{K}^n$ und ist $\mathcal{B}$ die Standardbasis des $\mathbb{K}^n$, so folgt für jede Bilinearform $f\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ und für alle Vektoren $v, w \in \mathbb{K}^n$:

$$f(v,w) = v^T \cdot A \cdot w.$$

In Analogie zu den linearen Abbildungen bezeichnen wir die Matrix A als darstellende Matrix von f bezüglich der Basis $\mathcal{B}$.

Definition 11.3

Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine (geordnete) Basis eines $\mathbb{K}$-Vektorraumes V und sei $f\colon V \times V \to \mathbb{K}$ eine Bilinearform. Sei $A_f^{\mathcal{B}} := (\alpha_{k,j}) \in \mathbb{K}^{n\times n}$, die Matrix deren Einträge wie folgt berechnet werden:

$$\alpha_{k,j} := f(v_k, v_j), \quad k,j \in \{1,\ldots,n\}.$$

Dann bezeichnet man $A_f^{\mathcal{B}}$ als *darstellende Matrix* von f bezüglich der Basis $\mathcal{B}$.

Beispiel 11.4. *Sei $V = \mathbb{R}^2$ und sei $f\colon V \times V \to \mathbb{R}$ gegeben durch $f(x,y) := x_1y_1 + x_2y_2 = x^Ty$. Die Bilinearität von f werden wir in Kürze zeigen; jetzt wollen wir nur eine darstellende Matrix von f bestimmen. Wählen wir dazu die Standardbasis*

$$\mathcal{B} = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right),$$

so ist

$$f\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right) = 1, \qquad f\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right) = 0,$$
$$f\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right) = 0, \qquad f\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right) = 1,$$

also ist die darstellende Matrix von f bezüglich $\mathcal{B}$ die Einheitsmatrix $A_f^{\mathcal{B}} = E_2 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$.

Lemma 11.5

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum. Eine Bilinearform $f\colon V \times V \to \mathbb{K}$ ist genau dann symmetrisch, wenn für jede Basis $\mathcal{B}$ von V die darstellende Matrix $A_f^{\mathcal{B}}$ symmetrisch ist.

Beweis. Eine Matrix $A = (\alpha_{k,j}) \in \mathbb{K}^{n\times n}$ ist genau dann symmetrisch (d.h. $A = A^T$), wenn $\alpha_{k,j} = \alpha_{j,k}$ ist für alle $k,j \in \{1,\ldots,n\}$. Sei $\mathcal{B} = (v_1,\ldots,v_n)$ eine beliebige Basis von V.

Nach Definition der darstellenden Matrix gilt dann

$$\begin{aligned} A_f^{\mathcal{B}} = (A_f^{\mathcal{B}})^T &\iff \alpha_{k,j} = \alpha_{j,k},\ \forall k,j \in \{1,\dots,n\} \\ &\iff f(v_k,v_j) = f(v_j,v_k),\ \forall k,j \in \{1,\dots,n\} \\ &\overset{(*)}{\iff} f \text{ ist symmetrisch.} \end{aligned}$$

Die letzte Äquivalenz ist nicht ganz offensichtlich. Klar ist $\Longleftarrow$. Zeigen wir, dass auch $\Longrightarrow$ gilt. Nehmen wir also an, dass

$$f(v_k,v_j) = f(v_j,v_k),\ \forall k,j \in \{1,\dots,n\} \tag{11.1.1}$$

gelte. Seien wie im einleitenden Text $b, c \in V$ mit Koordinaten

$$\kappa(b) =: \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},\ \kappa(c) =: \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \in \mathbb{K}^n$$

bezüglich $\mathcal{B}$. Dann gilt in der Tat

$$\begin{aligned} f(b,c) &= f\left(\sum_{k=1}^n \beta_k v_k, \sum_{j=1}^n \gamma_j v_j\right) \\ &= \sum_{k=1}^n \sum_{j=1}^n \beta_k \gamma_j f(v_k,v_j) \overset{(11.1.1)}{=} \sum_{k=1}^n \sum_{j=1}^n \beta_k \gamma_j f(v_j,v_k) \\ &= f\left(\sum_{j=1}^n \gamma_j v_j, \sum_{k=1}^n \beta_k v_k\right) \\ &= f(c,b), \end{aligned}$$

d.h. f ist eine symmetrische Bilinearform. □

Bemerkung 11.6. *Das Lemma hat eine sehr praktische Seite: Indem wir eine symmetrische Matrix $A \in \mathbb{K}^{n\times n}$ nehmen, können wir eine symmetrische Bilinearform $f\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ konstruieren, nämlich $f(x,y) := x^T A y$.*

Definition 11.7

Sei V ein $\mathbb{R}$-Vektorraum und sei $f\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Man nennt f

1. *positiv definit*, falls $f(v,v) > 0$ ist für alle $v \in V \setminus \{e_V\}$;
2. *negativ definit*, falls $f(v,v) < 0$ ist für alle $v \in V \setminus \{e_V\}$;
3. *indefinit*, falls $v, w \in V \setminus \{e_V\}$ existieren, für die $f(v,v) > 0$ und $f(w,w) < 0$ gilt.

Bemerkung 11.8. *Warum haben wir uns hierbei auf $\mathbb{K} = \mathbb{R}$ beschränkt? Der Grund ist der, dass der Körper der komplexen Zahlen nicht angeordnet ist, d.h. es gibt keine "$<$"-Relation auf den komplexen Zahlen. Was eine Ordnungsrelation "$<$" erfüllen müsste und warum es keine solche auf $\mathbb{C}$ gibt, ist Inhalt der Analysis. Wir werden im weiteren Verlauf sehen, wie man dieses Problem umgehen kann und auf den komplexen Zahlen eine positiv definite* Sesquilinearform *definieren kann.*

Beispiel 11.9. *Für $a, b, d \in \mathbb{R}$ sei $A := \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ die symmetrische 2×2-Matrix mit diesen Einträgen. In Aufgabe 4 von Übungsblatt 5 wurde gezeigt, dass A diagonalisierbar ist, d.h. insbesondere existiert eine Basis $\mathcal{B} = (v, w)$ des $\mathbb{R}^2$, die aus Eigenvektoren von A besteht. Es existieren also $\lambda, \mu \in \mathbb{R}$, so dass gilt:*

$$Av = \lambda v \quad \text{und} \quad Aw = \mu w.$$

Nehmen wir an, $f\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) := x^T Ay$ sei positiv definit. Dann gilt insbesondere

$$0 < f(v, v) = v^T Av = v^T \lambda v = \lambda v^T v = \lambda(v_1^2 + v_2^2)$$

Das ist genau dann der Fall, wenn $\lambda > 0$ ist und durch die analoge Rechnung für w folgt ebenso $\mu > 0$, d.h. A hat zwei positive (nicht notwendigerweise verschiedene) Eigenwerte. Wir werden später noch sehen, dass sich diese Aussage auch auf höherdimensionale Matrizen übertragen lässt.

Diese Fragestellung hat eine Anwendung in der Optimierung von Funktionen in mehreren Variablen. Ein lokales Maximum/Minimum einer Funktion $f\colon \mathbb{R} \to \mathbb{R}$ findet man, indem man $f'(x) = 0$ nach x auflöst (das gibt uns die kritischen Punkte) und dann das Vorzeichen von $f''(x)$ in den kritischen Punkten bestimmt. Ist hingegen $f\colon \mathbb{R}^n \to \mathbb{R}$, so setzt man den Gradienten[1] gleich dem Nullvektor und bestimmt, ob die Hessematrix[2] positiv oder negativ definit ist, d.h. ob die zur Hessematrix gehörende Bilinearform positiv oder negativ definit ist.

Nachdem wir anhand des Beispiels gesehen haben, dass die Definitheit eine wichtige Eigenschaft von Bilinearformen ist, ist es nun nicht überraschend, dass wir für positiv definite symmetrische Bilinearformen einen eigenen Begriff haben – das Skalarprodukt.

11.1.2 Skalarprodukte

Definition 11.10: Skalarprodukt auf einem reellen Vektorraum

Gegeben sei ein $\mathbb{R}$-Vektorraum V. Eine Abbildung $\langle \cdot, \cdot \rangle\colon V \times V \to \mathbb{R}$ heißt *Skalarprodukt* auf V, falls gilt:

1. $\langle v, w \rangle = \langle w, v \rangle$ für alle $v, w \in V$ (Symmetrie);
2. $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$ für alle $v, w \in V$ und $\lambda \in \mathbb{R}$ (Homogenität);
3. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ für alle $u, v, w \in V$ (Additivität);
4. $\langle v, v \rangle \geq 0$ für alle $v \in V$ und $\langle v, v \rangle = 0 \iff v = e_V$ (positive Definitheit).

Bemerkung 11.11. *Eigenschaften 1-3 implizieren, dass ein Skalarprodukt auf einem $\mathbb{R}$-Vektorraum eine symmetrische Bilinearform ist, da auch die Linearität bezüglich des zweiten Eintrags gegeben ist:*

- *für beliebige $u, v, w \in V$ gilt:*

$$\langle u, v + w \rangle \overset{\text{Symmetrie}}{=} \langle v + w, u \rangle \overset{\text{Linearität}}{=} \langle v, u \rangle + \langle w, u \rangle \overset{\text{Symmetrie}}{=} \langle u, v \rangle + \langle u, w \rangle;$$

[1] Der Gradient ist der Vektor, der die partiellen Ableitungen enthält.

[2] Die Hessematrix ist die Matrix, deren Einträge gerade die zweiten partiellen Ableitungen sind.

- *für beliebige $v, w \in V$ und $\lambda \in \mathbb{R}$ gilt:*

$$\langle v, \lambda w\rangle \overset{\text{Symmetrie}}{=} \langle \lambda w, v\rangle \overset{\text{Homogenität}}{=} \lambda\langle w, v\rangle \overset{\text{Symmetrie}}{=} \lambda\langle v, w\rangle.$$

Zudem ist diese Definition positive Definitheit äquivalent zu der, die wir für Bilinearformen eingeführt hatten, denn einerseits folgt aus

$$\langle v, v\rangle \geq 0 \text{ für alle } v \in V \quad \text{und} \quad \langle v, v\rangle = 0 \iff v = e_V,$$

dass $\langle v, v\rangle > 0$ ist für alle $v \neq e_V$. Andererseits gilt für jedes Skalarprodukt $\langle e_V, e_V\rangle = 0$, da für jeden beliebigen Vektor $v \in V$ gilt:

$$e_V = 0 \cdot v \implies \langle e_V, e_V\rangle = \langle 0 \cdot v, 0 \cdot v\rangle \overset{\text{Homogenität}}{=} 0 \cdot \langle v, 0 \cdot v\rangle = 0.$$

Bemerkung 11.12. *Für das Skalarprodukt gibt es auch andere Schreibweisen in der Literatur, beispielsweise $v \cdot w$. In dieser Form ähneln Symmetrie, Homogenität und Additivität dem Kommutativ-, Assoziativ- und Distributivgesetz:*

$$v \cdot w = w \cdot v, \quad (\lambda v) \cdot w = \lambda(v \cdot w), \quad (u + v) \cdot w = u \cdot w + v \cdot w.$$

Da wir den Punkt bereits für die Multiplikation von Matrizen und die Multiplikation von Vektoren/Matrizen und Skalaren verwenden, wählen wir hier die eingeführte Notation um die Operationen voneinander zu unterscheiden.

Beispiel 11.13. *Ist $V = \mathbb{R}^n$, dann ist $\langle v, w\rangle := v^T w = \sum_{k=1}^n v_k w_k = v_1 w_1 + \ldots v_n w_n$ ein Skalarprodukt, genannt das* kanonische *oder* euklidische Skalarprodukt *des $\mathbb{R}^n$. Zum Beweis überprüfen wir die Eigenschaften eines Skalarproduktes:*

Symmetrie: *Da die Multiplikation in $\mathbb{R}$ kommutativ ist, gilt für beliebige $v, w \in \mathbb{R}^n$:*

$$\langle v, w\rangle = \sum_{k=1}^n v_k w_k \overset{\text{(KG)}}{=} \sum_{k=1}^n w_k v_k = \langle w, v\rangle.$$

Homogenität: *Diese Eigenschaft basiert auf dem Assoziativ- und Distributivgesetz in $\mathbb{R}$. Für $v, w \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt:*

$$\langle \lambda v, w\rangle = \sum_{k=1}^n (\lambda v_k) w_k \overset{\text{(AG)}}{=} \sum_{k=1}^n \lambda(v_k w_k) \overset{\text{(DG)}}{=} \lambda \sum_{k=1}^n v_k w_k = \lambda\langle v, w\rangle.$$

Additivität: *Seien $u, v, w \in \mathbb{R}^n$. Dann gilt dank Kommutativ- und Distributivgesetz in $\mathbb{R}$:*

$$\langle u+v, w\rangle = \sum_{k=1}^n (u_k+v_k) w_k \overset{\text{(DG)}}{=} \sum_{k=1}^n (u_k w_k + v_k w_k) \overset{\text{(KG)}}{=} \sum_{k=1}^n u_k w_k + \sum_{k=1}^n v_k w_k = \langle u, w\rangle + \langle v, w\rangle.$$

positive Definitheit: *Sei $v \in \mathbb{R}^n$ beliebig. Dann gilt*

$$\langle v, v\rangle = \sum_{k=1}^n v_k^2 \geq 0$$

da das Quadrieren in $\mathbb{R}$ immer nichtnegative Zahlen erzeugt. Zudem gilt

$$\begin{aligned}\langle v, v\rangle = 0 &\iff v_k^2 = 0, \forall k \in \{1, \ldots, n\} \iff v_k = 0, \forall k \in \{1, \ldots, n\} \\ &\iff v = e_{\mathbb{R}^n}.\end{aligned}$$

Bemerkung 11.14. *Es gibt ein Skalarprodukt $\langle \cdot, \cdot \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, welches aus der Schule bekannt sein dürfte:*

$$\langle v, w \rangle := \cos(\sphericalangle(v, w)) \cdot |v| \cdot |w| \,,$$

wobei $|v|$ die Länge des Vektors v bezeichnet. Wir werden uns dieses Skalarprodukt näher ansehen sobald wir den Begriff der Norm *eingeführt haben und somit über Längen von Vektoren sprechen können. Letztlich werden wir sehen, dass das kanonische Skalarprodukt mit dem hier vorgestellten übereinstimmt.*

Beispiel 11.15. *Über dem Körper $\mathbb{K} = \mathbb{R}$ sei $\mathcal{P}_2$ der Vektorraum der reellen Polynome, deren Grad höchstens 2 beträgt. Dann wird auf folgende Weise ein Skalarprodukt definiert:*

$$\langle \sum_{k=0}^{2} a_k x^k, \sum_{k=0}^{2} b_k x^k \rangle := 2a_2 b_2 + a_1 b_1 + a_1 b_0 + a_0 b_1 + 2a_0 b_0.$$

Die Eigenschaften des Skalarproduktes zu überprüfen ist eine Übungsaufgabe.

Kommen wir nun zu der Definition eines Skalarproduktes auf Vektorräumen über dem Körper der komplexen Zahlen.

Definition 11.16: Skalarprodukt auf einem komplexen Vektorraum

Gegeben sei ein $\mathbb{C}$-Vektorraum V. Eine Abbildung $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ heißt *Skalarprodukt* auf V, falls gilt:

1. $\langle v, w \rangle = \overline{\langle w, v \rangle}$ für alle $v, w \in V$ (Hermitezität);
2. $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$ für alle $v, w \in V$ und $\lambda \in \mathbb{C}$ (Homogenität);
3. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ für alle $u, v, w \in V$ (Additivität);
4. $\langle v, v \rangle \geq 0$ für alle $v \in V$ und $\langle v, v \rangle = 0 \iff v = e_V$ (positive Definitheit).

Bemerkung 11.17. *Die Adjektive zu den Eigenschaften lauten: $\langle \cdot, \cdot \rangle$ ist symmetrisch oder hermitesch, homogen, additiv (bzw. zusammengenommen linear) und positiv definit.*

Bemerkung 11.18. *Die positive Definitheit ist dank der Hermitezität sinnvoll, denn für alle $v \in V$ gilt:*

$$\langle v, v \rangle \overset{\textit{Herm.}}{=} \overline{\langle v, v \rangle},$$

was impliziert, dass $\langle v, v \rangle \in \mathbb{R}$ ist. Daher ist es für alle Vektoren aus V tatsächlich möglich, zu entscheiden ob $\langle v, v \rangle$ nichtnegativ ist.

Definition 11.19

Ein $\mathbb{R}$-Vektorraum mit einem Skalarprodukt heißt *euklidischer Vektorraum*; ein $\mathbb{C}$-Vektorraum mit Skalarprodukt heißt *unitärer Vektorraum*.

Lemma 11.20

Sei $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ ein Skalarprodukt auf einem $\mathbb{C}$-Vektorraum V. Dann gelten folgende Eigenschaften des Skalarproduktes:

1. Für alle $u, v, w \in V$ gilt $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$.

2. Für alle $v, w \in V$ und $\lambda \in \mathbb{C}$ gilt $\langle v, \lambda w\rangle = \overline{\lambda}\langle v, w\rangle$.

Beweis. Übungsaufgabe. □

Bemerkung 11.21. *Das Skalarprodukt über einem $\mathbb{C}$-Vektorraum ist somit linear im ersten Argument, aber nur semilinear (also wörtlich gesehen halb linear) im zweiten Argument. Daraus wurde die Bezeichnung für $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{C}$ als positiv definite hermitesche* Sesquilinearform *(also anderthalbfach lineare Form). Der Begriff* hermitesch *bzw.* Hermitezität *ist von dem Namen des französischen Mathematikers Charles Hermite (1822-1901) abgeleitet.*

Beispiel 11.22. *Das kanonische Skalarprodukt über dem $\mathbb{C}$-Vektorraum $\mathbb{C}^n$ ist definiert als*

$$\langle v, w\rangle := \sum_{k=1}^{n} v_k \overline{w_k}.$$

Die Eigenschaften des Skalarproduktes zu überprüfen ist eine Übungsaufgabe.

Wir hatten im Abschnitt zu Bilinearformen den Begriff der darstellenden Matrix einer Bilinearform eingeführt. Handelt es sich bei der Bilinearform (oder im komplexen Fall bei der Sesquilinearform) um ein Skalarprodukt, so ist die darstellende Matrix benannt nach dem dänischen Mathematiker Jørgen Pedersen Gram (1850-1916):

Definition 11.23

Sei V ein euklidischer oder unitärer Vektorraum mit Basis $\mathcal{B} = (v_1, \dots, v_n)$ und Skalarprodukt $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{K}$. Dann heißt die Matrix $A = (\alpha_{k,j}) \in \mathbb{K}^{n\times n}$ mit den Einträgen

$$\alpha_{k,j} = \langle v_k, v_j\rangle, \quad k, j \in \{1, \dots, n\},$$

Gram'sche Matrix des Skalarproduktes $\langle\cdot,\cdot\rangle$ bezüglich der Basis $\mathcal{B}$.

Beispiel 11.24. *Sei $V = \mathbb{C}^n$ und sei $\mathcal{B}$ die geordnete Standardbasis des $\mathbb{C}^n$, d.h.*

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right),$$

wobei wir die Basisvektoren wieder mit $e_1, \dots, e_n$ bezeichnen. Dann gilt:

$$\langle e_k, e_j\rangle = \begin{cases} 1, & k = j, \\ 0, & k \neq j. \end{cases}$$

Somit ist die Gram'sche Matrix des kanonischen Skalarproduktes des $\mathbb{C}^n$ bezüglich der Standardbasis $\mathcal{B}$ des $\mathbb{C}^n$ gerade die Einheitsmatrix.

Für Bilinearformen, also auch für Skalarprodukte über $\mathbb{R}$-Vektorräumen, haben wir bereits gezeigt, dass die darstellende Matrix, also auch die Gram'sche Matrix, symmetrisch ist. Da Skalarprodukte über $\mathbb{C}$-Vektorräumen nicht symmetrisch, sondern hermitesch sind, ist es naheliegend, dass die zugehörige Gram'sche Matrix im Allgemeinen nicht symmetrisch sein muss.

Lemma 11.25

Sei V ein unitärer Vektorraum mit dem Skalarprodukt $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{C}$. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und sei $A \in \mathbb{C}^{n\times n}$ die Gram'sche Matrix des Skalarproduktes bezüglich der Basis $\mathcal{B}$. Dann ist A *hermitesch*, d.h. es gilt $A^T = \overline{A}$.

Beweis. Ist das Skalarprodukt hermitesch, so gilt $\langle v, w\rangle = \overline{\langle w, v\rangle}$ für alle $v, w \in V$. Insbesondere gilt also für alle Basisvektoren und somit für die Einträge von A^T:

$$\alpha_{j,k} = \langle v_j, v_k\rangle = \overline{\langle v_k, v_j\rangle} = \overline{\alpha_{k,j}}.$$

□

11.1.3 Normen

Was ist die Länge eines Vektors $v \in \mathbb{R}^3$? Was ist die „Größe“ eines Vektors des $\mathcal{P}_3$, des Vektorraums der reellen Polynome, deren Grad höchstens 3 betragen darf? Während die erste Frage typisches Schulwissen ist, ist die Antwort auf die zweite Frage alles andere als klar. Fügen wir diesen beiden theoretischen Fragen noch eine praktische hinzu: Wo bin ich schneller, wenn ich in Adlershof starte? Am Hauptgebäude der HU oder an dem der FU? Spätestens bei dieser Frage lautet die Antwort: Je nachdem, welches Fortbewegungsmittel ich nutze.[3] Es ist also sinnvoll, mehr als nur genau einen Begriff für eine Länge/Größe zur Verfügung zu haben.

Im Folgenden führen wir den Begriff einer *Norm* ein. Das Ziel ist es, Vektoren (im allgemeinen Sinne, d.h. Elemente eines Vektorraumes) eine Größe zuordnen zu können.

Definition 11.26: Norm

Sei $(V, +, \cdot)$ ein reeller oder komplexer Vektorraum, d.h. $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Eine nichtnegative reellwertige Funktion $\|\cdot\| : V \to [0, \infty)$ heißt *Norm*, falls sie folgende Eigenschaften besitzt:

1. $\|v\| = 0$ impliziert $v = e_V$. (Definitheit)
2. $\|a \cdot v\| = |a| \cdot \|v\|$ für alle $a \in \mathbb{K}$ und $v \in V$. (absolute Homogenität)
3. $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$. (Dreiecksungleichung)

Bemerkung 11.27.

i) *Man bezeichnet die Eigenschaft, dass die Dreiecksungleichung erfüllt ist, auch als* Subadditivität.

[3] Antwort zu einer festen Zeit laut Google Maps: Zur FU kommt man mit ÖPNV in einer Stunde, mit dem Auto in unter einer halben Stunde. Zum Hauptgebäude der HU braucht man mit den öffentlichen Verkehrsmitteln ca. eine dreiviertel Stunde, mit dem Auto zwischen 35 und 45 Minuten.

ii) *Man kann Eigenschaft 1 durch positive Definitheit ersetzen:*

$$\|v\| \geq 0, \forall v \in V \quad \text{und} \quad \|v\| = 0 \iff v = e_V. \tag{11.1.2}$$

Aus der positiven Definitheit folgt sofort die Definitheit. Andersherum folgt $\|v\| \geq 0, \forall v \in V$ *bereits aus der Annahme, dass die Norm nur Werte in* $[0, \infty)$ *annimmt. Damit gilt dann zusammen mit der Definitheit, dass die Norm* (11.1.2) *erfüllt.*

iii) *Der Körper soll laut Definition* $\mathbb{R}$ *oder* $\mathbb{C}$ *sein. Der Grund dafür ist, dass wir für die* absolute Homogenität *den* Betrag *eines Körperelementes kennen müssen. Im Falle einer komplexen Zahl* $z = a + ib$ *mit Realteil* $\mathrm{Re}(z) = a$ *und Imaginärteil* $\mathrm{Im}(z) = b$ *ist* $|z| = \sqrt{a^2 + b^2}$. *Für reelle Zahlen* $x \in \mathbb{R}$ *ist der Betrag wie folgt definiert:*

$$|x| := \begin{cases} x, & x \geq 0 \\ -x, & x < 0. \end{cases}$$

Beispiel 11.28. *Auf dem Vektorraum* $V = \mathbb{R}$ *definiert die Betragsfunktion eine Norm. Zunächst einmal ist in der Tat* $|v| \geq 0$ *für alle* $v \in \mathbb{R}$. *Das zeigen wir mit Hilfe einer Fallunterscheidung:*

- *Ist* $v \geq 0$, *so ist* $|v| = v \geq 0$.
- *Ist* $v < 0$, *so ist* $|v| = -v > 0$. *(Das gilt, da* $\mathbb{R}$ *die Anordnungsaxiome*[4] *erfüllt.)*

Überprüfen wir nun die drei Eigenschaften einer Norm:

Definitheit: *Ist* $v \neq 0$, *so ist entweder* $v > 0$ *und somit* $|v| = v > 0$ *oder* $v < 0$ *und somit* $|v| = -v > 0$. *(Diese Fallunterscheidung gilt wieder dank der Anordnungsaxiome.) In beiden Fällen gilt also* $|v| \neq 0$. *Durch Kontraposition ist die Definitheit somit gezeigt.*

absolute Homogenität: *Für* $V = \mathbb{R}$ *lautet die zu zeigende Eigenschaft:*

$$|a \cdot v| = |a| \cdot |v|, \quad \forall a, v \in \mathbb{R}.$$

Dies kann mit Hilfe einer Fallunterscheidung in Bezug auf die Vorzeichen von a *und* v *und durch mehrfache Anwendung der Anordnungsaxiome (bzw. der daraus abgeleiteten Eigenschaften) gezeigt werden. In Kurzform notiert: Für jedes* $x \in \mathbb{R}$ *gilt* $|x| = |-x|$. *Wählen wir* $a_0, v_0 \geq 0$ *derart, dass* $a = \pm a_0$ *und* $v = \pm v_0$ *ist, so gilt*

$$|a \cdot v| = |(\pm a_0) \cdot (\pm v_0)| = |\pm a_0 \cdot v_0| = |a_0 \cdot v_0| = a_0 \cdot v_0 = |a_0| \cdot |v_0| = |a| \cdot |v|.$$

Dreiecksungleichung: *Auch hier werden wieder die Anordnungsaxiome benötigt. Seien also* $v, w \in \mathbb{R}$. *Zunächst einmal gilt* $|v| = \max\{v, -v\}$ *und somit ist immer* $-v \leq |v|$ *und* $v \leq |v|$ *(und analog für* w*). Machen wir eine Fallunterscheidung:*

- *Ist* $v + w < 0$, *so ist* $|v + w| = -(v + w) = (-v) + (-w) \leq |v| + |w|$, *wobei wir wieder das Anordnungsaxiom benutzt haben.*

[4] *Exkurs in die Analysis:* Ein Körper $\mathbb{K}$ heißt *angeordnet*, falls eine Relation $<$ existiert, für die gilt:

1. Für jedes $x \in \mathbb{K}$ gilt genau eine der drei Aussagen: $x > 0$, $x = 0$ oder $x < 0$.
2. Gilt für $x, y \in \mathbb{K}$ sowohl $x > 0$ als auch $y > 0$, so gilt auch $x + y > 0$ und $x \cdot y > 0$.

Für mehr Details siehe beispielsweise [For16], Kapitel 3.

- *Ist* $v + w \geq 0$*, so ist analog* $|v + w| = v + w \leq |v| + |w|$.
- *Beide Punkte zusammen ergeben* $|v + w| = \max\{-(v + w), v + w\} \leq |v| + |w|$.

Im Vektorraum $V = \mathbb{R}$ ist es schwer, Alternativen zum Betrag als Norm zu finden. Anders sieht es im $\mathbb{R}^n$ für $n > 1$ aus.

Beispiel 11.29. *Sehen wir uns drei Beispiele für Normen auf dem* $\mathbb{R}$*-Vektorraum* $V = \mathbb{R}^n$ *an. Dazu sei* $v = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^T \in V$.

$$\|v\|_\infty := \max\{|v_1|, \dots, |v_n|\}, \quad \|v\|_2 := \sqrt{v_1^2 + \dots v_n^2}, \quad \|v\|_1 := |v_1| + \dots + |v_n|.$$

Diese Normen heißen Maximumsnorm *oder* Unendlichnorm, 2-Norm *oder* euklidische Norm *und* 1-Norm. *Die verschiedenen Normen haben durchaus ihre Daseinsberechtigung. Dazu kann man sich überlegen, welche Norm am besten geeignet ist, die Weglängen auf folgender Karte zu beschreiben.*

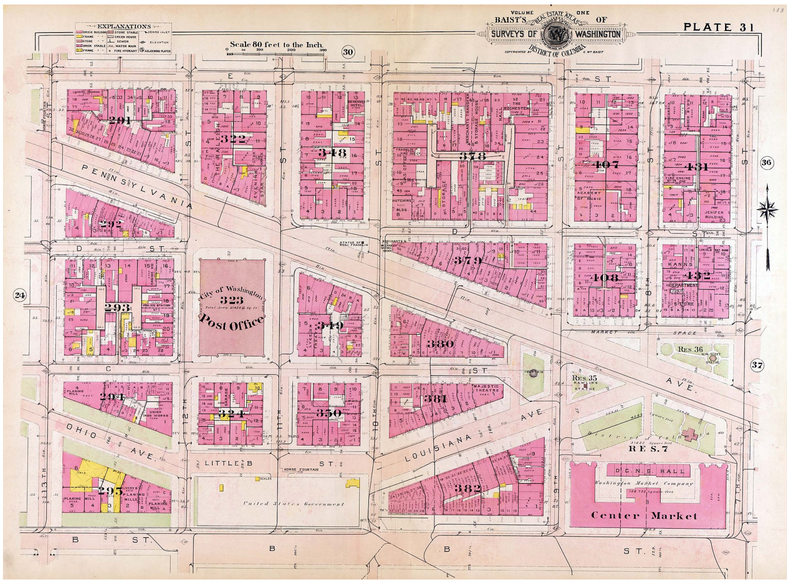


Abbildung 11.1: Bild: George William Baist [Public domain] – Quelle: `https://commons.wikimedia.org/wiki/File:1909_map_of_Downtown_Washington,_D.C..jpg`

Bemerkung 11.30 (Länge vs. Abstand)**.** *Ein Vektor* $v \in \mathbb{R}^n$ *kann als Punkt in einem Raum oder als Verschiebung in Richtung des Vektors um den Betrag* $|v|$ *interpretiert werden. Verwenden wir zur Unterscheidung ausnahmsweise* $\vec{v}$*, wenn wir die Verschiebung meinen, so ist* $|\vec{v}|$ *die* Länge *des Vektors, der die Verschiebung darstellt. Sind hingegen* $v, w \in \mathbb{R}^n$ *Punkte im Raum, so können wir nach dem* Abstand *zwischen den Punkten* v *und* w *fragen – in diesem Fall ist der Abstand* $|v - w|$*. Auf diese Art können wir in beliebigen Vektorräumen* V*, auf denen wir eine Norm* $\|\cdot\|$ *definiert haben, den Abstand zweier Vektoren* $v, w \in V$ *definieren als* $d(v, w) := \|v - w\|$*. Diese Abstandsfunktion* d *hat folgende Eigenschaften:*

1. $d(v, w) \geq 0$ *für alle* $v, w \in V$ *und* $d(v, w) = 0 \iff v = w$;
2. $d(v, w) = d(w, v)$ *für alle* $v, w \in V$;
3. $d(u, w) \leq d(u, v) + d(v, w)$ *für alle* $u, v, w \in V$.

Mit anderen Worten: Die Abstandsfunktion ist positiv definit und symmetrisch und sie erfüllt die Dreiecksungleichung. Erfüllt andersherum eine Abbildung $d\colon X \times X \to \mathbb{R}$ die drei genannten Eigenschaften, wobei X eine beliebige nichtleere Menge sei, so nennen wir d eine Metrik. *Aus jeder Norm können wir uns wie eben gesehen problemlos eine Metrik konstruieren; andererseits gibt es jedoch Räume, die mit einer Metrik versehen sind, auf denen keine Norm existiert. Da wir jedoch in euklidischen oder unitären Vektorräumen arbeiten, in denen es ein Skalarprodukt und folglich eine induzierte Norm gibt, ist der Abstandsbegriff natürlicherweise mit Hilfe der Norm als $d(v;w) := \|v-w\|$ gegeben.*

Die euklidische Norm hat einen offensichtlichen Bezug zum bereits eingeführten euklidischen Skalarprodukt: Für $v \in \mathbb{R}^n$ gilt

$$\|v\|_2 = \sqrt{\langle v, v\rangle}.$$

Bevor wir nachweisen, dass die euklidische Norm die Dreiecksungleichung erfüllt, wollen wir kurz festhalten, dass die gerade beobachtete Beziehung nicht nur in diesem Spezialfall eine Rolle spielt.

Definition 11.31

Sei V ein $\mathbb{K}$-Vektorraum ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) mit dem Skalarprodukt $\langle\cdot,\cdot\rangle$. Dann heißt die Abbildung $\|\cdot\| : V \to [0,\infty)$, welche definiert ist durch

$$\|v\| := \sqrt{\langle v, v\rangle}, \quad v \in V,$$

die *von dem Skalarprodukt induzierte Norm*.

Wenn wir zeigen können, dass jede von einem Skalarprodukt induzierte Norm tatsächlich die Normeigenschaften erfüllt, so ist damit gleichzeitig nachgewiesen, dass die euklidische Norm in der Tat eine Norm ist, da wir in Beispiel 11.13 bereits gezeigt hatten, dass das euklidische Skalarprodukt wirklich ein Skalarprodukt ist.

Zu diesem Beweis benötigen wir jedoch zuerst ein Hilfsresultat. Es geht zurück auf die Mathematiker Augustin-Louis Cauchy (1789-1857), Wiktor Jakowlewitsch Bunjakowski (1804-1889) und Hermann Amandus Schwarz (1843-1921), welche die Ungleichung in unterschiedlichen Räumen gezeigt hatten – Cauchy für Reihen, Bunjakowski und Schwarz hingegen für Funktionen.

Satz 11.32: Cauchy-Schwarz'sche Ungleichung

Sei V ein $\mathbb{R}$-Vektorraum mit dem Skalarprodukt $\langle\cdot,\cdot\rangle$ und der Norm $\|v\| := \sqrt{\langle v, v\rangle}$. Dann gilt für alle $v, w \in V$ die Ungleichung

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|,$$

wobei die Gleichheit genau dann gilt, wenn ein $a \in \mathbb{R}$ existiert, so dass $w = a \cdot v$ ist.

Beweis. Seien $v, w \in V$. Ist $w = e_V$, so folgt die Aussage sofort: Einerseits ist

$$|\langle v, e_V\rangle| = |\langle v, 0 \cdot v\rangle| = |0 \cdot \langle v, v\rangle| = |0| = 0.$$

Andererseits ist

$$\|v\| \cdot \|e_V\| = \|v\| \cdot 0 = 0.$$

Nehmen wir nun an, dass $w \neq e_V$ sei. Für jedes $\lambda \in \mathbb{R}$ gilt dank der Bilinearität und positiven Definitheit des Skalarproduktes

$$0 \leq \langle v - \lambda w, v - \lambda w\rangle = \langle v, v\rangle - 2\lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle. \tag{11.1.3}$$

Zudem folgt aus $w \neq e_V$ mit der positiven Definitheit $\langle w, w\rangle \neq 0$, d.h. wir können konkret $\lambda := \frac{\langle v,w\rangle}{\langle w,w\rangle}$ setzen und erhalten somit

$$0 \leq \langle v, v\rangle - 2\frac{\langle v, w\rangle^2}{\langle w, w\rangle} + \frac{\langle v, w\rangle^2}{\langle w, w\rangle} \overset{\text{Def. Norm}}{=} \|v\|^2 - \frac{\langle v, w\rangle^2}{\|w\|^2}. \tag{11.1.4}$$

Durch Umstellen erhalten wir die Aussage

$$\langle v, w\rangle^2 \leq \|v\|^2 \cdot \|w\|^2 . \tag{11.1.5}$$

Durch Wurzelziehen auf beiden Seiten ergeben sich die Ungleichungen

$$\langle v, w\rangle \leq \|v\| \cdot \|w\| \quad \text{und} \quad -\langle v, w\rangle \leq \|v\| \cdot \|w\| . \tag{11.1.6}$$

Da $|\langle v, w\rangle| = \max\{-\langle v, w\rangle, \langle v, w\rangle\}$ ist, folgt hieraus die behauptete Ungleichung. Gleichheit gilt in (11.1.3) genau dann, wenn $v - \lambda w = e_V$ ist, d.h. genau dann, wenn $v = \lambda w$ ist. Die Ungleichungen in (11.1.4), (11.1.5) und (11.1.6) werden auch genau in diesem einen Fall zu Gleichungen, so dass auch die zusätzliche Behauptung bewiesen ist. □

Die Cauchy-Schwarz'sche Ungleichung gilt analog für $\mathbb{C}$-Vektorräume.

Beispiel 11.33. *Seien $v, w \in \mathbb{R}^n$ und sei $\langle v, w\rangle = v_1w_1 + \ldots, v_nw_n$. Dann lautet die Cauchy-Schwarz'sche Ungleichung:*

$$|\langle v, w\rangle| = |\sum_{k=1}^{n} v_k w_k| \leq \sqrt{\sum_{k=1}^{n} v_k^2} \cdot \sqrt{\sum_{k=1}^{n} w_k^2},$$

bzw. äquivalent dazu (durch Quadrieren)

$$\left(\sum_{k=1}^{n} v_k w_k\right)^2 \leq \left(\sum_{k=1}^{n} v_k^2\right) \cdot \left(\sum_{k=1}^{n} w_k^2\right).$$

Satz 11.34

Sei V ein euklidischer oder unitärer Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Dann wird durch $\|v\| := \sqrt{\langle v, v\rangle}$ eine Norm auf V definiert.

Beweis. Sei V ein $\mathbb{R}$-Vektorraum und $\langle\cdot,\cdot\rangle \colon V \times V \to \mathbb{R}$ ein Skalarprodukt. Wir überprüfen für diesen Fall, dass die induzierte Norm $\|v\| := \sqrt{\langle v, v\rangle}$ in der Tat die Normeigenschaften erfüllt.

Wertebereich und Definitheit: Sei $v \in V$. Dann gilt

$$\|v\| = \sqrt{\langle v, v\rangle} \overset{\text{pos. Definitheit des SKP}}{\geq} 0$$

und

$$\|v\| = 0 \iff \langle v, v\rangle = 0 \overset{\text{pos. Definitheit des SKP}}{\iff} v = e_V.$$

absolute Homogenität: Seien $v \in V$ und $a \in \mathbb{K}$. Dann gilt

$$\begin{aligned}\|a \cdot v\| = \sqrt{\langle a \cdot v, a \cdot v\rangle} &\overset{\text{Hom. des SKP}}{=} \sqrt{a\langle v, a \cdot v\rangle} \overset{\text{Symmetrie des SKP}}{=} \sqrt{a\langle a \cdot v, v\rangle} \\ &\overset{\text{Hom. des SKP}}{=} \sqrt{a^2\langle v, v\rangle} = \sqrt{a^2} \cdot \sqrt{\langle v, v\rangle} = |a| \cdot \|v\|.\end{aligned}$$

Dreiecksungleichung: Seien $v, w \in V$. Dann gilt

$$\begin{aligned}\|v + w\| &= \sqrt{\langle v + w, v + w\rangle} \\ &\overset{\text{Add. des SKP}}{=} \sqrt{\langle v, v + w\rangle + \langle w, v + w\rangle} \\ &\overset{\text{Symm. des SKP}}{=} \sqrt{\langle v + w, v\rangle + \langle v + w, w\rangle} \\ &\overset{\text{Add. des SKP}}{=} \sqrt{\langle v, v\rangle + \langle w, v\rangle + \langle w, v\rangle + \langle w, w\rangle} \\ &\overset{\text{Symm. des SKP}}{=} \sqrt{\|v\|^2 + 2\langle v, w\rangle + \|w\|^2} \\ &\overset{\text{Cauchy-Schwarz}}{\leq} \sqrt{\|v\|^2 + 2\,\|v\|\,\|w\| + \|w\|^2} \\ &\overset{\text{Bin. Formel}}{=} \sqrt{(\|v\| + \|w\|)^2} \\ &\overset{(*)}{=} \|v\| + \|w\|,\end{aligned}$$

wobei wir nach dem Radizieren in $(*)$ keinen Betrag schreiben müssen, da bereits gezeigt wurde, dass die induzierte Norm nichtnegativ ist.

Der Beweis für die induzierte Norm auf einem unitären Vektorraum funktioniert analog. □

Aus diesem Satz folgt insbesondere, dass die euklidische Norm in der Tat eine Norm ist. Wir hatten in Bemerkung 11.14 das aus der Schule bekannte Skalarprodukt

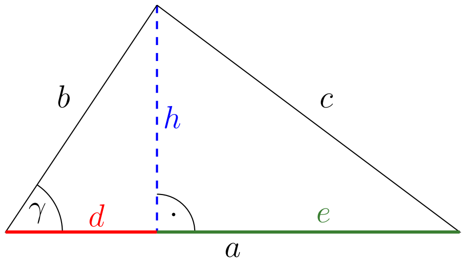
$$\langle v, w\rangle := \cos(\sphericalangle(v, w)) \cdot |v| \cdot |w|\,, \quad v, w \in \mathbb{R}^n, \tag{11.1.7}$$

angeführt. Wir wollen nun zeigen, dass dieses Skalarprodukt mit dem kanonischen Skalarprodukt übereinstimmt. Aus Gründen der Anschaulichkeit beschränken wir uns hierbei auf $V = \mathbb{R}^2$. Wir benötigen zunächst ein Resultat aus der Geometrie:

Lemma 11.35: Kosinussatz

Ein Dreieck habe Seiten der Längen a, b und c und der Winkel gegenüber der zuletzt genannten Seite sei γ. Dann gilt

$$c^2 = a^2 + b^2 - 2ab\cos(\gamma).$$



Beweis. Als Hilfsmittel benutzen wir nur den Satz des Pythagoras und die Formel für die Winkel am rechtwinkligen Dreieck. Die Bezeichnungen seien die aus der Skizze. Zunächst einmal gilt

$$c^2 = h^2 + e^2.$$

Für diese beiden haben wir wiederum die Darstellungen

$$\begin{aligned} h^2 &= b^2 - d^2 \quad \text{und} \\ e^2 &= (a-d)^2 = a^2 - 2ad + d^2. \end{aligned}$$

Durch Einsetzen in unsere erste Gleichung erhalten wir

$$c^2 = b^2 - d^2 + a^2 - 2ad + d^2 = a^2 + b^2 - 2ad.$$

Für den Winkel γ gilt, dass sein Kosinus gleich dem Quotienten aus Ankathete und Hypotenuse ist, d.h.

$$\cos(\gamma) = \frac{d}{b} \qquad \Longleftrightarrow \qquad d = b \cdot \cos(\gamma).$$

Insgesamt erhalten wir also

$$c^2 = a^2 + b^2 - 2ab \cdot \cos(\gamma).$$

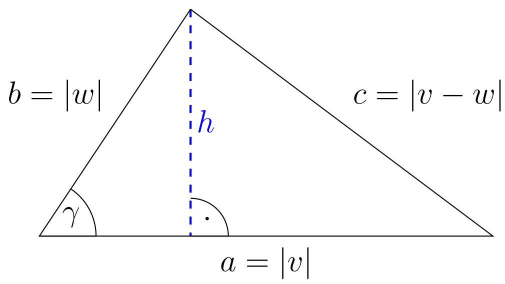
□

Mit Hilfe dieses Resultates können wir folgende Aussage beweisen:

Satz 11.36

Seien $v, w \in \mathbb{R}^2$ und sei $\gamma = \sphericalangle(v, w)$ der von beiden Vektoren eingeschlossene Winkel. Sei zudem $\langle \cdot, \cdot \rangle$ definiert wie in (11.1.7). Dann gilt $\langle v, w \rangle = v^T w$, d.h. das Skalarprodukt aus (11.1.7) und das kanonische Skalarprodukt auf $V = \mathbb{R}^2$ stimmen überein.

Beweis. Wir wollen den Kosinussatz auf folgendes Dreieck anwenden:



Nach Kosinussatz gilt

$$|v-w|^2 = |v|^2 + |w|^2 - 2\,|v|\,|w|\cos(\gamma),$$

bzw. durch Umstellen und Ersetzen des Skalarproduktes

$$\langle v, w\rangle = \cos(\gamma)\,|v|\,|w| = \frac{1}{2}\left(|v|^2 + |w|^2 - |v-w|^2\right).$$

Setzen wir für jeden Vektor auf der rechten Seite als dessen Länge die euklidische Norm ein, so erhalten wir hieraus

$$\begin{aligned}\langle v, w\rangle &= \frac{1}{2}\left[v_1^2 + v_2^2 + w_1^2 + w_2^2 - (v_1 - w_1)^2 - (v_2 - w_2)^2\right]\\ &= \frac{1}{2}\left[2v_1w_1 + 2v_2w_2\right]\\ &= v^T w.\end{aligned}$$

□

Insbesondere folgt hieraus, dass zwei Vektoren $v, w \in \mathbb{R}^2$ genau dann orthogonal zueinander sind, d.h. die durch sie erzeugten Geraden stehen senkrecht aufeinander, wenn $\langle v, w\rangle = 0$ ist, denn dann ist $\gamma = \arccos\left(\frac{\langle v,w\rangle}{|v||w|}\right)$ ein ungerades Vielfaches von π bzw. $90°$.

11.2 Orthonormalbasen

11.2.1 Orthogonalität und Parallelität

Mit Hilfe des Skalarproduktes können wir den Begriff der Orthogonalität einführen:

Definition 11.37

Sei V ein $\mathbb{K}$-Vektorraum ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) mit Skalarprodukt $\langle\cdot,\cdot\rangle$.

(i) Zwei Vektoren $v, w \in V$ heißen *orthogonal* (bezüglich $\langle\cdot,\cdot\rangle$), falls $\langle v, w\rangle = 0$ ist. Man schreibt dann $v \perp w$.

(ii) Zwei Vektoren $v, w \in V$ heißen *parallel* zueinander, falls ein $a \in \mathbb{K}$ existiert, so dass $v = a \cdot w$ ist. Man schreibt dann $v \| w$.

Beispiel 11.38. *Im $\mathbb{R}^2$ mit dem Skalarprodukt $\langle v, w\rangle = v^T w = v_1 w_1 + v_2 w_2$ gilt:*

- *Die Vektoren $v = \begin{pmatrix}1\\2\end{pmatrix}$ und $w = \begin{pmatrix}4\\8\end{pmatrix}$ sind parallel, denn $4 \cdot v = w$.*
- *Die Vektoren $v = \begin{pmatrix}1\\2\end{pmatrix}$ und $w = \begin{pmatrix}-2\\1\end{pmatrix}$ sind orthogonal zueinander, denn $\langle v, w\rangle = 1 \cdot (-2) + 2 \cdot 1 = 0$.*

Bemerkung 11.39. *In einer Übungsaufgabe wird gezeigt werden, dass der Nullvektor stets der einzige Vektor ist, der zu allen Vektoren sowohl parallel als auch orthogonal ist.*

Während lineare Unabhängigkeit für Mengen von Vektoren definiert ist, ist Orthogonalität eine Eigenschaft, die immer für Paare von Vektoren definiert ist. Der Kürze halber sagt man jedoch auch, eine *Menge* von Vektoren $\{v_1, \ldots, v_n\} \subseteq V$ sei *orthogonal*, wenn $v_k \perp v_j$ für alle Indizes $k \neq j$ gilt.

Zueinander orthogonale Vektoren sind, sofern wir den Nullvektor ausschließen, immer linear unabhängig; die Umkehrung dieser Aussage gilt jedoch nicht, wie man leicht an den Vektoren $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und dem kanonischen Skalarprodukt des $\mathbb{R}^2$ überprüfen kann.

Satz 11.40

Sei $M := \{v_1, \ldots, v_n\} \subseteq V \setminus \{e_V\}$ eine Teilmenge eines euklidischen oder unitären Vektorraumes V. Wenn die Vektoren aus M paarweise zueinander orthogonal sind, dann ist M linear unabhängig.

Beweis. Je nachdem, ob V ein euklidischer oder unitärer Vektorraum ist, sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Für $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gelte

$$\sum_{k=1}^{n} \lambda_k v_k = e_V.$$

Wählen wir einen beliebigen Index $j \in \{1, \ldots, n\}$ so gilt

$$\begin{aligned} 0 &= \langle e_V, v_j \rangle \\ &= \langle \sum_{k=1}^{n} \lambda_k v_k, v_j \rangle \\ &\overset{\text{Additivität}}{=} \langle \lambda_j v_j, v_j \rangle + \sum_{k \neq j} \langle \lambda_k v_k, v_j \rangle \\ &\overset{\text{Hom. in 1. Komp.}}{=} \lambda_j \langle v_j, v_j \rangle + \sum_{k \neq j} \lambda_k \langle v_k, v_j \rangle \\ &\overset{M \text{ orthogonal}}{=} \lambda_j \langle v_j, v_j \rangle. \end{aligned}$$

Da nach Annahme $v_j \neq e_V$ ist, folgt aus der positiven Definitheit des Skalarproduktes $\langle v_j, v_j \rangle > 0$. Folglich muss $\lambda_j = 0$ sein. Da der Index $j \in \{1, \ldots, n\}$ beliebig gewählt wurde, müssen auch alle anderen Koeffizienten Null sein, d.h. $\lambda_1 = \ldots = \lambda_n = 0$, was beweist, dass M linear unabhängig ist. □

11.2.2 Orthogonal- und Orthonormalsysteme

Betrachten wir die Einheitsvektoren des $\mathbb{R}^n$, so sind diese bezüglich des kanonischen Skalarproduktes paarweise orthogonal zueinander. Multiplizieren wir die Vektoren mit von Null verschiedenen Faktoren, so ändert sich das nicht. Beispielsweise ist

$$\langle \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -5 \end{pmatrix} \rangle = 0.$$

Die *Länge* der Vektoren hat also scheinbar keinen Einfluss darauf hat, ob Vektoren zueinander orthogonal sind. Überzeugen wir uns, dass dies kein Zufall war.

Lemma 11.41

Sei $M \subseteq V$ eine Teilmenge eines euklidischen oder unitären $\mathbb{K}$-Vektorraumes V mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Sind die Vektoren aus M paarweise orthogonal zueinander, dann sind auch beliebige Vielfache der Vektoren paarweise orthogonal zueinander.

Beweis. Seien $v, w \in M$ und seien $\lambda, \mu \in \mathbb{K}$. Sind die Vektoren aus M alle paarweise orthogonal zueinander, so gilt insbesondere $\langle v, w \rangle = 0$. Daraus folgt jedoch sofort

$$\langle \lambda v, \mu w \rangle = \begin{cases} \lambda\mu\langle v, w\rangle = 0, & \text{falls } \mathbb{K} = \mathbb{R} \text{ ist,} \\ \lambda\overline{\mu}\langle v, w\rangle = 0, & \text{falls } \mathbb{K} = \mathbb{C} \text{ ist.} \end{cases}$$

Folglich sind die Vielfachen der Vektoren aus M in der Tat paarweise orthogonal zueinander. $\square$

Da die Länge der Vektoren also keinen Einfluss auf die Orthogonalität hat, können wir sie auf eine einheitliche Länge stauchen oder strecken ohne die Orthogonalität zu verlieren.

Definition 11.42

Sei V ein euklidischer oder unitärer Vektorraum mit induzierter Norm $\|v\| := \sqrt{\langle v, v\rangle}$ für $v \in V$. Sei $M \subseteq V$. Man nennt M

1. *orthogonal*, falls $v \perp w$ ist für alle $v, w \in M$ mit $v \neq w$;
2. *orthonormal*, falls M orthogonal ist und zudem $\|v\| = 1$ für alle $v \in M$ gilt.

Falls zudem $e_V \notin M$, so heißt M *Orthogonal-* bzw. *Orthonormalsystem*. Ist M eine orthonormale Basis von V, so nennt man M auch *Orthonormalbasis* (kurz ONB).

Bemerkung 11.43 (Zusatzinformation). *In unendlich-dimensionalen Vektorräumen versteht man unter einer ONB nicht einfach eine Basis, deren Elemente paarweise orthogonal zueinander und normiert sind, sondern ein Orthonormalsystem, dessen lineare Hülle* dicht *in V liegt. Es gilt also für die ONB $\mathcal{B}$ nicht $\mathcal{L}(\mathcal{B}) = V$, sondern für jeden Vektor $v \in V$ und zu jedem $\varepsilon > 0$ existiert ein $w \in \mathcal{L}(\mathcal{B})$ mit $\|v - w\| < \varepsilon$.*

Mit diesen Bezeichnungen können wir die Aussage von Satz 11.40 wie folgt zusammenfassen: Orthogonalsysteme sind linear unabhängig. Da die Normierung nur eine zusätzliche Eigenschaft ist, folgt aus 11.40 sofort, dass auch Orthonormalsysteme linear unabhängig sind.

Beispiel 11.44. *Sei $V = \mathbb{R}^2$ und sei $M := \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.*

M ist bezüglich des kanonischen Skalarproduktes orthogonal:

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = 1 \cdot 1 + 1 \cdot (-1) = 0.$$

M ist jedoch nicht orthonormal, denn es gilt

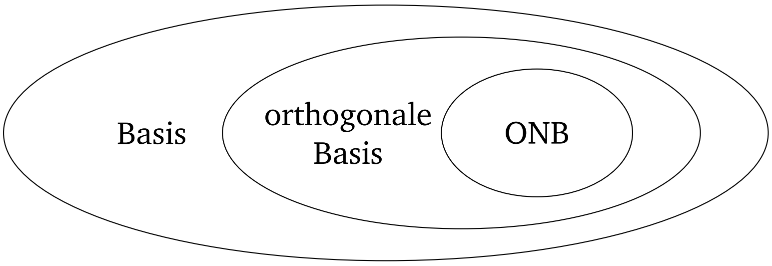
$$\left\|\begin{pmatrix}1\\1\end{pmatrix}\right\| = \left\|\begin{pmatrix}1\\-1\end{pmatrix}\right\| = \sqrt{1+1} = \sqrt{2} \neq 1.$$

Stauchen wir beide Vektoren um diesen Faktor, so erhalten wir die nun orthonormale Menge

$$M' := \left\{ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \right\}.$$

Da der Nullvektor nicht in M' enthalten ist, ist M' ein Orthonormalsystem. Zudem ist $\mathbb{R}^2 = \mathcal{L}(M')$, so dass M' sogar eine ONB des $\mathbb{R}^2$ ist.

Wir haben also bereits gesehen, dass es einfach ist, aus einem Orthogonal- ein Orthonormalsystem zu machen. Zudem ist jedes Orthogonalsystem linear unabhängig. Beschränken wir uns auf Basen, so können wir die Begriffe wie folgt in ein Venn-Diagramm eintragen:



11.2.3 Das Gram-Schmidt'sche Orthonormalisierungsverfahren

In diesem Abschnitt wollen wir aus beliebigen Basen Orthonormalbasen konstruieren. Bevor wir uns der Konstruktion zuwenden, sehen wir uns erstmal an, warum eine ONB tatsächlich nützlich ist.

Satz 11.45: Entwicklung nach einer ONB

Sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine orthogonale Basis eines euklidischen oder unitären Vektorraumes V. Dann hat jeder Vektor $v \in V$ folgende Darstellung:

$$v = \sum_{k=1}^{n} \frac{\langle v, b_k\rangle}{\langle b_k, b_k\rangle} b_k. \tag{11.2.1}$$

Ist $\mathcal{B}$ sogar eine ONB von V, so vereinfacht sich die Darstellung für $v \in V$ zu

$$v = \sum_{k=1}^{n} \langle v, b_k\rangle b_k. \tag{11.2.2}$$

Bemerkung 11.46. *Die Aussage, die in der Formel* (11.2.2) *steckt, kann man folgendermaßen beschreiben:*

Ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine ONB von V, so sind Koordinaten eines beliebigen Vektors $v \in V$ bezüglich $\mathcal{B}$ gerade durch das Skalarprodukt des Vektors mit den jeweiligen Basisvektoren *gegeben. Mit unserer Notation aus Bemerkung 6.49 und Abschnitt 7.5 heißt das:*

$$\kappa_{\mathcal{B}}(v) = \begin{pmatrix} \langle v, b_1 \rangle \\ \vdots \\ \langle v, b_n \rangle \end{pmatrix}.$$

Beweis. Aus (11.2.1) folgt sofort (11.2.2), da für eine ONB zusätzlich die Normierung $\|b_k\|^2 = \langle b_k, b_k \rangle = 1$ für alle $k \in \{1, \ldots, n\}$ gilt. Zeigen wir also (11.2.1):

Für jede Basis, also insbesondere für die orthogonale Basis $\mathcal{B}$ gilt nach Satz 6.22, dass es zu jedem Vektor $v \in V$ eindeutig bestimmte Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gibt, für die gilt:

$$v = \sum_{k=1}^{n} \lambda_k b_k.$$

Wählen wir einen beliebigen Index $j \in \{1, \ldots, n\}$, so folgt durch Bilden des Skalarproduktes mit dem Basisvektor b_j

$$\langle v, b_j \rangle = \langle \sum_{k=1}^{n} \lambda_k b_k, b_j \rangle \overset{\text{SKP linear in 1. Komp.}}{=} \sum_{k=1}^{n} \lambda_k \langle b_k, b_j \rangle \overset{\text{orth.}}{=} \lambda_j \langle b_j, b_j \rangle.$$

Stellen wir diese Gleichung nach λ_j um, so erhalten wir

$$\lambda_j = \frac{\langle v, b_j \rangle}{\langle b_j, b_j \rangle}.$$

□

Bemerkung 11.47. *Ist es von Bedeutung, ob in* (11.2.1) *und* (11.2.2) $\langle v, b_k \rangle$ *oder* $\langle b_k, v \rangle$ *steht? Im Falle eines euklidischen Vektorraumes nicht, da das Skalarprodukt symmetrisch ist. Sollte es sich jedoch um einen unitären Vektorraum handeln, so ist das Skalarprodukt nur in der ersten Komponente linear. Ein Vertauschen der Komponenten ist also nur zulässig, wenn danach komplex konjugiert wird.*

Um eine Idee zu bekommen, wie man eine ONB konstruieren kann, betrachten wir das kleinste sinnvolle Beispiel, den $\mathbb{R}^2$:

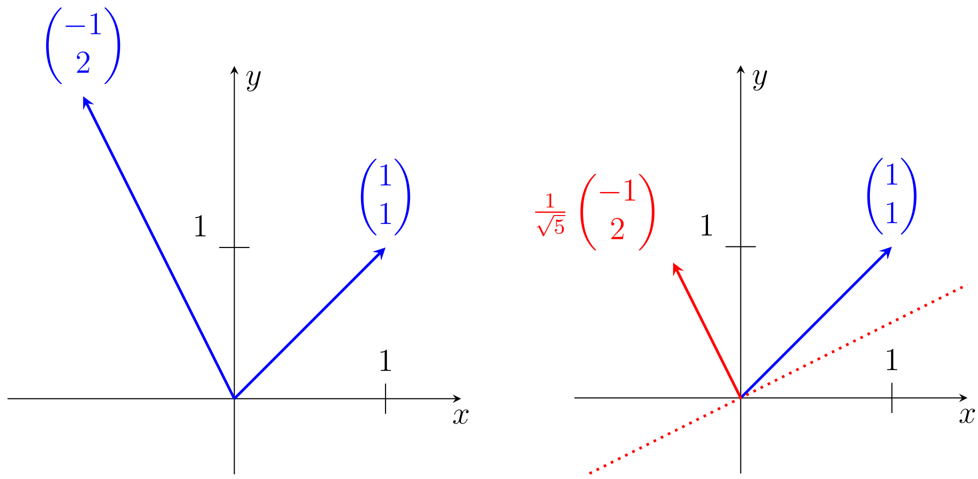
Beispiel 11.48. *Sei* $\mathcal{B} := \left(\begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$. *Diese Menge ist leicht als (geordnete) Basis des* $\mathbb{R}^2$ *zu erkennen und sie ist nicht orthogonal bezüglich des kanonischen Skalarproduktes, da*

$$\langle \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle = 1 \cdot (-1) + 1 \cdot 2 = 1 \neq 0$$

ist. Starten wir mit dem ersten Vektor, so können wir diesen zunächst auf die schon bekannte Art normieren:

$$\left\| \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5} \implies \left\| \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\| = 1.$$

Anschließend wollen wir gewissermaßen den zum ersten Vektor orthogonalen Anteil des zweiten Vektors finden. In der Skizze bedeutet das, dass wir einen Vektor suchen, der sich auf der roten gestrichelten Linie befindet.



Die Gerade, die durch diese gestrichelte Linie dargestellt wird, hat die Form

$$\left\{ \begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2 \,\middle|\, x = 2y \right\} = \mathcal{L}\begin{pmatrix}2\\1\end{pmatrix}.$$

Alle Vektoren aus dieser Menge, die zudem die auf 1 normierte Länge haben, sind $\pm\frac{1}{\sqrt{5}}\begin{pmatrix}2\\1\end{pmatrix}$*. Wählen wir einen dieser Vektoren aus, so erhalten wir beispielsweise*

$$\mathcal{B}' := \left(\frac{1}{\sqrt{5}}\begin{pmatrix}-1\\2\end{pmatrix}, \frac{1}{\sqrt{5}}\begin{pmatrix}2\\1\end{pmatrix} \right)$$

als ONB des $\mathbb{R}^2$*.*

Demonstrieren wir anhand dieses Beispiels nochmal die Aussage von Satz 11.45. Dazu sei $v = \begin{pmatrix}1\\1\end{pmatrix}$*. Nach Aussage des Satzes müsste für die Koordinaten von* v *bezüglich* $\mathcal{B}'$ *gelten:*

$$\kappa_{\mathcal{B}}(v) = \begin{pmatrix} \left\langle \begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{5}}\begin{pmatrix}-1\\2\end{pmatrix} \right\rangle \\ \left\langle \begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{5}}\begin{pmatrix}2\\1\end{pmatrix} \right\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{3}{\sqrt{5}} \end{pmatrix}.$$

In der Tat ist

$$\frac{1}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}}\begin{pmatrix}-1\\2\end{pmatrix} + \frac{3}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}}\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}(6-1)/5\\(2+3)/5\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}.$$

Aus einer Basis, die aus zwei Vektoren besteht, können wir mit etwas Nachdenken also eine ONB konstruieren. Allerdings können wir dieses Vorgehen noch nicht auf höhere Dimensionen übertragen ohne folgende Fragen beantwortet zu haben:

(I) Wie charakterisiert man die Menge aller Vektoren, die nicht nur zu einem Vektor orthogonal sind, sondern zu einem Orthogonal- oder Orthonormalsystem?

(II) Wie kann man einen Vektor auf eine Gerade (oder im Allgemeinen auf diesen Untervektorraum) projizieren?

Beantworten wir zunächst die erste Frage:

Definition 11.49

Sei V ein euklidischer oder unitärer Vektorraum und U ein (linearer oder affiner) Unterraum von V. Dann bezeichnet man

$$U^{\perp} := \{v \in V \mid \langle v, u \rangle = 0, \, \forall u \in U\}$$

als das *orthogonale Komplement* von U.

Beispiel 11.50. *Betrachten wir folgende Räume:*

$$U := \mathcal{L}\begin{pmatrix}1\\1\\1\end{pmatrix} = \left\{ \begin{pmatrix}a\\a\\a\end{pmatrix} \,\middle|\, a \in \mathbb{R} \right\}, \quad V := \mathcal{L}\left(\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}\right\}\right), \quad W := \mathbb{R}^3.$$

U, V und W sind 1-, 2- und 3-dimensionale $\mathbb{R}$-Vektorräume und es gilt die Inklusionskette $U \subset V \subset W$. Damit können wir U sowohl als UVR von V als auch als UVR von W auffassen.

Berechnen wir zunächst das orthogonale Komplement von $U \subset V$:

$$\begin{aligned}
U_V^{\perp} &= \left\{ v \in V \,\middle|\, \langle \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}, \begin{pmatrix}a\\a\\a\end{pmatrix} \rangle = 0, \forall a \in \mathbb{R} \right\} \\
&= \{v \in V \mid av_1 + av_2 + av_3 = 0, \; \forall a \in \mathbb{R}\} \\
&= \{v \in V \mid v_1 = -v_2 - v_3\} \\
&= \left\{v \in \mathbb{R}^3 \,\middle|\, v_1 = -v_2 - v_3, \, v_2 = v_3\right\} \\
&= \left\{v \in \mathbb{R}^3 \,\middle|\, v_1 = -2v_3, v_2 = v_3\right\} \\
&= \mathcal{L}\begin{pmatrix}-2\\1\\1\end{pmatrix} \subseteq V.
\end{aligned}$$

Andererseits gilt für U als UVR von W:

$$\begin{aligned}
U_W^{\perp} &= \left\{ w \in W \,\middle|\, \langle \begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}, \begin{pmatrix}a\\a\\a\end{pmatrix} \rangle = 0, \forall a \in \mathbb{R} \right\} \\
&= \left\{w \in \mathbb{R}^3 \,\middle|\, w_1 = -w_2 - w_3\right\} \\
&= \mathcal{L}\left(\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}\right\}\right) \subseteq W.
\end{aligned}$$

Bemerkung 11.51.

i) Gilt $\langle v, u \rangle = 0$ für alle $u \in U$, so schreibt man dafür auch kurz $v \perp U$.

ii) Man nennt zwei Unterräume $U_1, U_2 \subseteq V$ eines euklidischen oder unitären Vektorraumes orthogonal *zueinander, wenn jeder Vektor des einen Unterraumes zu jedem Vektor des anderen orthogonal ist. In diesem Fall schreibt man $U_1 \perp U_2$. Nach Definition ist $U^{\perp} \perp U$ für jeden Unterraum U von V.*

iii) Im Beispiel konnten wir folgendes beobachten:

$$U_V^\perp + U = V \text{ und } U_V^\perp \cap U = \{e_V\},$$
$$U_W^\perp + U = W \text{ und } U_W^\perp \cap U = \{e_W\}.$$

Lemma 11.52

Sei V ein euklidischer oder unitärer Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$ und sei $U \subseteq V$ ein UVR von V. Dann ist auch $U^\perp$ ein UVR von V.

Beweis. Wir wenden das UVR-Kriterium an. Zunächst einmal ist $U^\perp$ eine nichtleere Teilmenge von V, da der Nullvektor immer darin enthalten ist (siehe Bemerkung 11.39). Seien nun $v_1, v_2 \in U^\perp$, $a, b \in \mathbb{K}$ und $u \in U$. Da $v_1 \perp u$ und $v_2 \perp u$ ist, gilt

$$\langle av_1 + bv_2, u\rangle = a\langle v_1, u\rangle + b\langle v_2, u\rangle = a \cdot 0 + b \cdot 0 = 0,$$

d.h. $av_1 + bv_2 \in U^\perp$. Somit ist $U^\perp$ ein UVR von V. □

Bevor wir zu unserer Ausgangsfrage zurückkommen, erwähnen wir noch ein sehr nützliches Resultat, welches uns leicht ermöglicht, die Orthogonalität zweier (Unter-)Vektorräume zu überprüfen.

Lemma 11.53

Seien $\mathcal{B} := (v_1, \ldots, v_m)$ und $\mathcal{C} := (w_1, \ldots, w_n)$ Basen zweier Untervektorräume W und W' eines euklidischen oder unitären Vektorraumes V. Dann gilt folgende Äquivalenz:

$$W \perp W' \iff v_k \perp w_j, \forall k \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\}.$$

Beweis. Der Beweis ist eine Übungsaufgabe. □

Nachdem wir Frage (I) beantwortet haben, kommen wir nun zu Frage (II). Formulieren wir zunächst das Ziel:

Seien $\{v_1, \ldots, v_m\} \subseteq V$ ein Orthonormalsystem und $v \in V$ ein Vektor, für den die Menge $\{v_1, \ldots, v_m, v\}$ immer noch linear unabhängig ist. Wir suchen einen Vektor $w \in V$, so dass $\mathcal{L}(\{v_1, \ldots, v_m, v\}) = \mathcal{L}(\{v_1, \ldots, v_m, w\})$ ist und zudem $\{v_1, \ldots, v_m, w\}$ ein Orthonormalsystem ist. Es muss also $w \in \{v_1, \ldots, v_m\}^\perp$ und $\|w\| = 1$ sein und – und hier kommt der Zusammenhang mit dem ursprünglichen Vektor v – die lineare Hülle darf sich nicht verändern, d.h. es muss $w \in \mathcal{L}(\{v_1, \ldots, v_m, v\})$ gelten. Ignorieren wir zunächst die Normierung, so liefert das folgende Lemma die gewünschten Eigenschaften:

Lemma 11.54

Sei V ein euklidischer oder unitärer Vektorraum und sei $M := \{v_1, \ldots, v_m\}$ ein Orthonormalsystem in V. Sei $U := \mathcal{L}(M)$. Dann lässt sich jeder Vektor $v \in V$ auf eindeutige Weise als Summe $v = u + w$ mit $u \in U$ und $w \in U^\perp$ schreiben. Es ist

$$u = \sum_{k=1}^{m} \langle v, v_k \rangle v_k \quad \text{und} \quad w = v - \sum_{k=1}^{m} \langle v, v_k \rangle v_k. \tag{11.2.3}$$

Beweis. Wir müssen sowohl die Eindeutigkeit als auch die Existenz einer solchen Darstellung zeigen.

Eindeutigkeit: Angenommen es wäre $v = u+w = u'+w'$ mit $u, u' \in U$ und $w, w' \in U^\perp$. Da U und nach Lemma 11.52 auch $U^\perp$ Untervektorräume von V sind, ist $u-u' \in U$ und $w - w' \in U^\perp$. Damit gelten folgende Gleichungen:

$$(u - u') + (w - w') = e_V \quad \text{und} \quad \langle u - u', w - w' \rangle = 0.$$

Somit ist

$$\langle u - u', u - u' \rangle = \langle u - u', -(w - w') \rangle = -\langle u - u', w - w' \rangle = 0.$$

Da das Skalarprodukt (positiv) definit ist, folgt hiermit $u - u' = e_V$, also $u = u'$ und somit auch $w = w'$.

Existenz: Zu $v \in V$ seien u und w wie in (11.2.3) definiert. Diese erfüllen die geforderten Eigenschaften:

- $u + w = v$;
- $u = \sum_{k=1}^{m} \langle v, v_k \rangle v_k \in \mathcal{L}(M) = U$ nach Definition der linearen Hülle;
- $w \in U^\perp$, denn für alle $j \in \{1, \ldots, m\}$ gilt

$$\begin{aligned}
\langle w, v_j \rangle &= \langle v - \sum_{k=1}^{m} \langle v, v_k \rangle v_k, v_j \rangle \\
&= \langle v, v_j \rangle - \sum_{k=1}^{m} \langle v, v_k \rangle \underbrace{\langle v_k, v_j \rangle}_{=0,\text{ falls } k \neq j} \\
&= \langle v, v_j \rangle - \langle v, v_j \rangle \underbrace{\langle v_j, v_j \rangle}_{=1} \\
&= 0.
\end{aligned}$$

Die Orthogonalität von w zu allen Vektoren aus $\mathcal{L}(M)$ (und nicht nur den Basisvektoren) folgt mit Lemma 11.53.

□

Der in (11.2.3) definierte Vektor w erfüllt beide geforderten Eigenschaften: $w \in U^\perp$ und $w \in \mathcal{L}(\{v_1, \ldots, v_m, v\})$. Damit haben wir eine Methode gefunden, wie wir ein Orthonormalsystem um einen dazu orthogonalen Vektor erweitern können. Normieren wir diesen Vektor, so haben wir ein neues Orthonormalsystem. Damit können wir das Gram-Schmidt'sche Orthonormalisierungsverfahren formulieren:

Gram-Schmidt'sches Orthonormalisierungsverfahren

Sei $\mathcal{B} := (b_1, \ldots, b_n)$ eine Basis eines euklidischen oder unitären Vektorraumes V. Mit folgendem Algorithmus konstruiert man eine ONB $\mathcal{C} := (c_1, \ldots, c_n)$ von V:

1. $c_1 := \frac{b_1}{\|b_1\|}$.
2. Für $j \in \{2, \ldots, n\}$ setze

$$\widetilde{c}_j := b_j - \sum_{k=1}^{j-1} \langle b_j, c_k \rangle c_k \quad \text{und} \quad c_j := \frac{\widetilde{c}_j}{\|\widetilde{c}_j\|}.$$

Bevor wir das Verfahren zum Herleiten weiterer theoretischer Erkenntnisse verwenden, berechnen wir für ein konkretes Beispiel eine ONB:

Beispiel 11.55. *Betrachten wir die Basis*

$$\mathcal{B} := \left(\begin{pmatrix} 4 \\ 3i \end{pmatrix}, \begin{pmatrix} i \\ i \end{pmatrix} \right)$$

des unitären Vektorraumes $\mathbb{C}^2$ mit dem kanonischen Skalarprodukt $\langle v, w \rangle = v_1 \overline{w}_1 + v_2 \overline{w}_2$. Da

$$\langle \begin{pmatrix} 4 \\ 3i \end{pmatrix}, \begin{pmatrix} i \\ i \end{pmatrix} \rangle = 4 \cdot (-i) + 3i \cdot (-i) = -4i + 3 \neq 0$$

ist, handelt es sich nicht um ein Orthogonalsystem, also insbesondere nicht um eine ONB. Um eine solche zu konstruieren, normalisieren wir zunächst den ersten Basisvektor:

$$\left\| \begin{pmatrix} 4 \\ 3i \end{pmatrix} \right\| = \sqrt{\langle \begin{pmatrix} 4 \\ 3i \end{pmatrix}, \begin{pmatrix} 4 \\ 3i \end{pmatrix} \rangle} = \sqrt{4 \cdot 4 + 3i \cdot (-3i)} = \sqrt{25} = 5.$$

Wir setzen also $c_1 := \frac{1}{5} \begin{pmatrix} 4 \\ 3i \end{pmatrix}$. Gehen wir zum zweiten Schritt des Gram-Schmidt'schen Orthonormalisierungsverfahrens über.

$$\begin{aligned} \widetilde{c}_2 :&= b_2 - \langle b_2, c_1 \rangle c_1 = \begin{pmatrix} i \\ i \end{pmatrix} - \langle \begin{pmatrix} i \\ i \end{pmatrix}, \begin{pmatrix} 4/5 \\ 3i/5 \end{pmatrix} \rangle \cdot \begin{pmatrix} 4/5 \\ 3i/5 \end{pmatrix} \\ &= \begin{pmatrix} i \\ i \end{pmatrix} - \left(4i/5 - 3i^2/5\right) \cdot \begin{pmatrix} 4/5 \\ 3i/5 \end{pmatrix} \\ &= \begin{pmatrix} i \\ i \end{pmatrix} - \frac{1}{25} \begin{pmatrix} (4i+3) \cdot 4 \\ (4i+3) \cdot 3i \end{pmatrix} \\ &= \frac{1}{25} \begin{pmatrix} -12 + 9i \\ 12 + 16i \end{pmatrix} \end{aligned}$$

Berechnen wir wieder die Norm des Vektors:

$$\begin{aligned} \|\widetilde{c}_2\| &= \frac{1}{25} \sqrt{(-12 + 9i) \cdot (-12 - 9i) + (12 + 16i) \cdot (12 - 16i)} \\ &= \frac{1}{25} \sqrt{12^2 + 9^2 + 12^2 + 16^2} = \frac{1}{25} \sqrt{625} = 1. \end{aligned}$$

Der Vektor ist also bereits normiert, d.h. $c_2 = \widetilde{c}_2$. *Somit haben wir folgende ONB des unitären Vektorraumes* $\mathbb{C}^2$ *konstruiert:*

$$\mathcal{C} = \left(\frac{1}{5}\cdot\begin{pmatrix}4\\3i\end{pmatrix}, \frac{1}{25}\cdot\begin{pmatrix}-12+9i\\12+16i\end{pmatrix}\right).$$

Sehen wir uns am Beispiel des Vektors $v = \begin{pmatrix}1\\0\end{pmatrix}$ *an, dass Gleichung* (11.2.2) *tatsächlich erfüllt wird:*

$$\begin{aligned}
&\langle\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}4/5\\3i/5\end{pmatrix}\rangle\cdot\begin{pmatrix}4/5\\3i/5\end{pmatrix} + \langle\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}(-12+9i)/25\\(12+16i)/25\end{pmatrix}\rangle\cdot\begin{pmatrix}(-12+9i)/25\\(12+16i)/25\end{pmatrix}\\
&= \frac{4}{5}\cdot\begin{pmatrix}4/5\\3i/5\end{pmatrix} + \frac{1}{25^2}\cdot(-12-9i)\cdot\begin{pmatrix}-12+9i\\12+16i\end{pmatrix}\\
&= \frac{1}{25}\cdot\begin{pmatrix}16\\12i\end{pmatrix} + \frac{1}{25^2}\cdot\begin{pmatrix}12^2+9^2\\-12^2+9\cdot16-(9+16)\cdot12i\end{pmatrix}\\
&= \frac{1}{25}\cdot\begin{pmatrix}16\\12i\end{pmatrix} + \frac{1}{25}\cdot\begin{pmatrix}9\\-12i\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}.
\end{aligned}$$

Satz 11.56

Jeder endlich-dimensionale euklidische oder unitäre Vektorraum besitzt eine ONB.

Beweis. Seien $\mathcal{B}$ und $\mathcal{C}$ wie im Gram-Schmidt'schen Orthonormalisierungsverfahren beschrieben. Nach Konstruktion ist für jedes $j \in \{1,\ldots,n\}$ die Menge $\{c_1,\ldots,c_j\}$ eine ONB von $\mathcal{L}(\{b_1,\ldots,b_j\})$ und somit ist insbesondere $\mathcal{C}$ eine ONB von V. □

Zum Abschluss dieses Abschnittes wollen wir aus Teil iii) von Bemerkung 11.51 eine allgemeingültigen Aussage herleiten. Dazu führen wir die übliche Schreibweise für innere[5] direkte Summen ein:

Notation

Gilt für Untervektorräume $U_1, U_2 \subseteq V$ eines $\mathbb{K}$-Vektorraumes V

$$U_1 + U_2 = V \quad \text{und} \quad U_1 \cap U_2 = \{e_V\},$$

so schreiben wir auch $V = U_1 \oplus U_2$ und nennen V *innere direkte Summe* von U_1 und U_2 bzw. $U_1 \oplus U_2$ *direkte Zerlegung* von V.

Mit Hilfe der neu eingeführten Begriffe und Notationen können wir die Beobachtungen aus dem Beispiel als Satz formalisieren:

[5] Es gibt auch den Begriff einer *äußeren direkten Summe*, der jedoch in dieser Vorlesung keine Rolle spielen wird. Er steht im Zusammenhang zum Produkt von Vektorräumen und speziell in dem Fall, dass es sich um unendlich viele Vektorräume handelt – eine Situation, die uns nicht begegnen wird.

Satz 11.57: direkte Zerlegung durch das orthogonale Komplement

Sei V ein euklidischer oder unitärer endlich-dimensionaler Vektorraum. Sei zudem $U \subseteq V$ ein UVR von V. Dann gilt

$$U^{\perp} \oplus U = V.$$

Insbesondere gilt die Dimensionsformel

$$\dim(U^{\perp}) + \dim(U) = \dim(V).$$

Beweis von Satz 11.57. Der Satz enthält gleich mehrere Teilaussagen, die wir nacheinander beweisen, nämlich:

1. $U^{\perp} \cap U = \{e_V\}$;
2. $U^{\perp} + U = V$;
3. $\dim(U^{\perp}) + \dim(U) = \dim(V)$.

Zu 1.: Wäre $u \neq e_V \in U^{\perp} \cap U$, dann wäre insbesondere $u \perp u$, was laut Bemerkung 11.39 und der zugehörigen Übungsaufgabe der Serie 7 auf den Nullvektor, jedoch auf keinen anderen Vektor zutrifft.

Zu 2.: Da wir o.B.d.A. davon ausgehen können, dass U eine ONB besitzt, können wir Lemma 11.54 anwenden, woraus sofort die Behauptung folgt.

Zu 3.: In Aufgabe 4 des 10. Übungsblattes in LAAG I wurde folgende Aussage gezeigt:

Seien $U_1, U_2 \subseteq V$ Untervektorräume eines endlich-dimensionalen $\mathbb{K}$-Vektorraumes V. Dann ist $\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)$.

Wenden wir diese auf $U_1 = U$ und $U_2 = U^{\perp}$ an, so erhalten wir dank 1. und 2.

$$\dim(V) \overset{2.}{=} \dim(U + U^{\perp}) \overset{\text{ÜA, 1.}}{=} \dim(U) + \dim(U^{\perp}) - \dim(e_V) = \dim(U) + \dim(U^{\perp}).$$

□

Bemerkung 11.58. *Zwei Bemerkungen zu den Voraussetzungen des Satzes:*

i) Wenn man statt eines linearen einen affinen Unterraum $U \subseteq V$ nimmt, so ist $U^{\perp}$ immer noch ein linearer Unterraum (d.h. ein UVR) von V. Allerdings ist $U^{\perp} \cap U = \emptyset$. Als Beispiel kann man einen beliebigen Vektor $v \neq e_V$ wählen und $U := \{v\}$ setzen. Dann ist insbesondere $e_V \notin U$ und $v \notin U^{\perp}$, so dass $U^{\perp} \cap U = \emptyset$ ist.

ii) Für die direkte Zerlegung benötigen wir die Existenz einer ONB, welche dank des Gram-Schmidt'schen Orthogonalisierungsverfahrens für endlich-dimensionale euklidische oder unitäre Vektorräume gesichert ist. Für unendlich-dimensionale Vektorräume haben wir dies nicht und in der Tat gibt es Beispiele, für die gilt: $\dim(U) = \infty$, $U \subseteq V$ und $U \oplus U^{\perp} \subsetneq V$.

Es können auch Summen aus mehr als zwei Untervektorräumen gebildet werden. Was es anschaulich bedeutet, dass $V = U_1 \oplus \ldots \oplus U_n$ für Untervektorräume $U_1, \ldots, U_n \subseteq V$ ist, sagt folgender Satz.

Satz 11.59

Für eine natürliche Zahl $n \geq 2$ seien $U_1, \ldots, U_n \subseteq V$ Untervektorräume eines $\mathbb{K}$-Vektorraumes V. Sei $U := U_1 + \ldots + U_n$. Dann sind folgende Aussagen äquivalent:

(1) Die Summe ist direkt, d.h. für jedes $j \in \{1, \ldots, n\}$ ist

$$U_j \cap \left(\sum_{k \neq j} U_k \right) = \{e_V\}.$$

(2) Jeder Vektor $u \in U$ besitzt eine eindeutige Darstellung als $u = \sum_{k=1}^{n} u_k$ mit eindeutig bestimmten $u_k \in U_k$, $k \in \{1, \ldots, n\}$.

(3) Ist $\sum_{k=1}^{n} u_k = e_V$ für Vektoren $u_k \in U_k$, $k \in \{1, \ldots, n\}$, so folgt stets $u_k = e_V$ für alle $k \in \{1, \ldots, n\}$.

Beweis. Der Beweis dieses Satzes ist eine Übungsaufgabe. □

Beispiel 11.60. *Sei $A \in \mathbb{K}^{n \times n}$ eine diagonalisierbare Matrix mit Eigenwerten $\lambda_1, \ldots, \lambda_m$. Dann gilt die direkte Zerlegung*

$$\mathbb{K}^n = \mathrm{Eig}(A; \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(A; \lambda_m),$$

denn einerseits kann jeder Vektor als Linearkombination von Eigenvektoren dargestellt werden, da eine Basis aus Eigenvektoren existiert. Somit ist

$$\mathbb{K}^n = \mathrm{Eig}(A; \lambda_1) + \ldots + \mathrm{Eig}(A; \lambda_m).$$

Andererseits folgt aus

$$v \in \mathrm{Eig}(A; \lambda_j) \cap \left(\sum_{k \neq j} \mathrm{Eig}(A; \lambda_k) \right)$$

sofort $v = e_V$, da ein Eigenvektor eines Eigenwertes nicht als Linearkombination von Eigenvektoren der übrigen Eigenwerte geschrieben werden kann. Zur Erinnerung: In Satz 10.17 hatten wir bereits gezeigt, dass Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind. Damit ist die Summe direkt.

11.3 Orthogonale Projektion

11.3.1 Projektionen

Folgende Definition ist bereits in einer Übungsserie aufgetaucht:

Definition 11.61

Sei V ein $\mathbb{K}$-Vektorraum. Ein Endomorphismus $f \in L(V; V)$ heißt *Projektion*, falls er *idempotent* ist, d.h. falls $f \circ f = f$ ist.

Idempotenz ist eine nützliche Eigenschaft in der Praxis: Nehmen wir an, jemand möchte 50 Euro von seinem Konto abheben. Dann sollte der Kontostand nach dem Abheben um 50 Euro verringert werden. Benutzt die Bank die Operation

$$k \mapsto k - 50$$

für den Kontostand und bricht während des Abhebevorgangs die Verbindung ab, so darf die Operation nicht nochmal durchgeführt werden, da sonst im Zweifel der Kontostand um 100 Euro verringert wurde. Speichert man hingegen zu Beginn des Abhebevorgangs den Kontostand k^*, so kann man stattdessen die idempotente Operation

$$k \mapsto k^* - 50$$

(mit dem von k unabhängigen Wert $k^* - 50$) beliebig oft ausführen – es wird am Ende der korrekte Kontostand festgehalten werden.

Kommen wir nun zu den Eigenschaften von Projektionen. In der 5. Übungsserie wurde folgendes Resultat gezeigt:

Lemma 11.62

Jede Projektion $f \in L(V;V)$ ist diagonalisierbar und die einzigen Eigenwerte von f sind 0 und 1. Zudem gilt

$$\ker(f) = \operatorname{Eig}(f;0) \quad \text{und} \quad \operatorname{Bild}(f) = \operatorname{Eig}(f;1).$$

Nach Beispiel 11.60 gilt somit insbesondere

$$V = \ker(f) \oplus \operatorname{Bild}(f)$$

für jede Projektion $f \in L(V;V)$.

Beispiel 11.63. *Beispiele für Projektionen sind die Identität* $\mathrm{id}_V\colon V \to V$*, die jeden Vektor sich selbst zuordnet, d.h.*

$$\mathrm{id}_V(v) = v, \quad \forall v \in V.$$

Eine weitere Projektion ist die konstante Abbildung f_e*, die jeden Vektor* $v \in V$ *auf dem Nullvektor* $e_V \in V$ *abbildet, d.h.*

$$f_e(v) = e_V, \quad \forall v \in V.$$

Beispiel 11.64. *Sei* V *ein endlich-dimensionaler euklidischer Vektorraum und sei* $\mathcal{B} = (b_1, \ldots, b_m)$ *eine ONB eines Untervektorraumes* $U \subseteq V$ *von* V*. Dann definiert*

$$f\colon V \to V, \quad f(v) := \sum_{k=1}^{m} \langle v, b_k \rangle b_k$$

eine Projektion mit $\operatorname{Bild}(f) = U$ *und* $\ker(f) = U^\perp$*. Überprüfen wir dies:*

- *Die Abbildung* f *ist linear, denn für* $v, w \in V$ *und* $\lambda, \mu \in \mathbb{R}$ *ist*

$$\begin{aligned} f(\lambda v + \mu w) &= \sum_{k=1}^{m} \langle (\lambda v + \mu w), b_k \rangle b_k \overset{\text{SKP linear}}{=} \sum_{k=1}^{m} \left(\lambda \langle v, b_k \rangle + \mu \langle w, b_k \rangle \right) b_k \\ &= \lambda \sum_{k=1}^{m} \langle v, b_k \rangle b_k + \mu \sum_{k=1}^{m} \langle w, b_k \rangle b_k = \lambda f(v) + \mu f(w). \end{aligned}$$

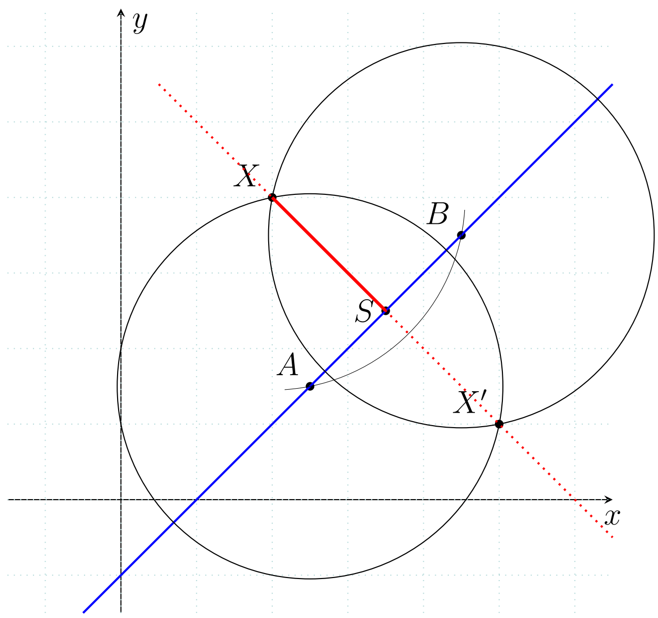
- *Es gilt* $\ker(f) = U^\perp$*, denn*
$$v \in \ker(f) \iff f(v) = e_V \overset{\mathcal{B}\text{ Basis}}{\iff} \langle v, b_k\rangle = 0, \forall k \in \{1,\dots,m\} \iff v \in U^\perp.$$
- *Da* $f(v) \in \mathcal{L}(\mathcal{B}) = U$ *ist, folgt* $\mathrm{Bild}(f) \subseteq U$*. Andererseits ist* $U \subseteq \mathrm{Bild}(f)$*, denn für jeden Vektor* $u = \sum_{j=1}^m \lambda_j b_j \in U$ *ist*
$$f(u) = \sum_{k=1}^m \langle \sum_{j=1}^m \lambda_j b_j, b_k\rangle b_k = \sum_{k=1}^m \sum_{j=1}^m \lambda_j \langle b_j, b_k\rangle b_k = \sum_{k=1}^m \lambda_k b_k = u.$$
- *Insbesondere folgt aus diesen beiden Berechnungen, dass* f *eine Projektion ist, denn*
$$(f\circ f)(v) = f(f(v)) = \begin{cases} f(v), & \text{für } v \in U = \mathrm{Eig}(f;1) = \mathrm{Bild}(f), \\ f(e_V) = e_V = f(v), & \text{für } v \in U^\perp = \mathrm{Eig}(f;0) = \ker(f), \end{cases}$$
d.h. $(f\circ f)(v) = f(v)$ *für alle* $v \in V$.

11.3.2 Der Spezialfall: Orthogonale Projektionen

Gegeben sei folgende Problemstellung:

Bestimmen Sie den Abstand des Punktes X von der Gerade G. Zur Verfügung haben Sie nur einen Zirkel und ein Lineal.

Eine mögliche Konstruktion ist im Folgenden beschrieben: In einem ersten Schritt konstruiert man eine Senkrechte auf die Gerade, die durch X geht (d.h. die Gerade, auf der X und X' liegen) indem man zuerst zwei Punkte A und B auf der Gerade konstruiert, die von X gleich weit entfernt sind, und anschließend mit dem Zirkel Kreise um A und B zieht, die jeweils durch X gehen und sich zudem in einem Weiteren Punkt, X', schneiden. In einem zweiten Schritt misst man den Abstand von dem Schnittpunkt der Senkrechten und der Geraden, S, zu X.



Im Hintergrund dieser Konstruktionstecken zwei Überlegungen:

1. Der Abstand von einem Punkt $x \in \mathbb{R}^2$ zu einer Geraden $G \subseteq \mathbb{R}^2$ ist definiert als
$$d(x, G) := \min\{\|x-y\| \mid y \in G\}.$$
2. Der Abstand ist gleich der Länge der *Lotstrecke* von dem Punkt X zu dem Punkt S, den man erreicht, wenn man von X senkrecht auf die Gerade projiziert.

In diesem Abschnitt werden wir diese Idee auf abstrakte Vektorräume übertragen.

Definition 11.65

Sei V ein euklidischer oder unitärer Vektorraum und sei $U \subseteq V$ ein endlich-dimensionaler Untervektorraum von V. Erfüllt $P_U \in L(V;V)$ die Eigenschaften

(1) $P_U(v) \in U$ für alle $v \in V$ und

(2) $P_U(v) - v \in U^\perp$ für alle $v \in V$,

so heißt P_U *Orthogonalprojektion* oder *orthogonale Projektion auf* U.

Beispiel 11.66. *Die lineare Abbildung f aus Beispiel 11.64 ist eine Orthogonalprojektion auf $U = \mathcal{L}(\mathcal{B})$, denn für alle $v \in V$ gilt $f(v) \in \mathcal{L}(\mathcal{B}) = U$ und nach Lemma 11.54 ist $v - f(v) \in U^\perp$ für alle $v \in V$. Da das orthogonale Komplement ein UVR ist, ist also auch $-(v - f(v)) = f(v) - v \in U^\perp$.*

Bemerkung 11.67.

i) *Nach Lemma 11.54 ist die Orthogonalprojektion eines Vektors $v \in V$ auf einen endlich-dimensionalen UVR $U \subseteq V$ eindeutig bestimmt. Bei bekannter ONB von U ist sie gegeben durch* (11.2.3) *bzw. durch die Abbildung aus Beispiel 11.64.*

ii) *Das Gram-Schmidt'sche Orthonormalisierungsverfahren basiert auf der sukzessiven Berechnung der Orthogonalprojektion eines gegebenen Basisvektors b_k auf den bisher berechneten Unterraum $\mathcal{L}(c_1, \ldots, c_{k-1})$ und anschließende Normierung dieses Vektors.*

Es gibt auch eine alternative Definition für die Orthogonalprojektion. Um beide Definitionen zunächst auseinanderhalten zu können, geben wir dieser zunächst ein anderes mathematisches Symbol.

Definition 11.68

Sei V ein euklidischer oder unitärer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und induzierter Norm $\|\cdot\|$. Sei $U \subseteq V$ ein endlich-dimensionaler Untervektorraum von V. Die *orthogonale Projektion* eines Vektors $v \in V$ auf U ist definiert als

$$\pi_U(v) := \arg\min \{ \|v - u\| \mid u \in U \} \tag{11.3.1}$$

Die zugehörige Abbildung bezeichnen wir mit $\pi_U : V \to V$.

Bemerkung 11.69. *Gleichung* (11.3.1) *sagt aus, dass $\pi_U(v)$ der Vektor $u \in U$ ist, für den der Abstand $\|v - u\|$ minimal wird. Insbesondere ist $\pi_U(v) = v$ für alle $v \in U$.*

Satz 11.70

Sei V ein euklidischer Vektorraum und sei $U \subseteq V$ ein endlich-dimensionaler Untervektorraum von V. Dann gilt $P_U(v) = \pi_U(v)$ für alle $v \in V$.

Beweis. Für einen festen Vektor $v \in V$ und jeden beliebigen Vektor $u \in U$ gilt

$$\|u-v\|^2 = \Bigg\|\underbrace{u - P_U(v)}_{\in U} + \underbrace{P_U(v) - v}_{\in U^\perp}\Bigg\|^2 \overset{(*)}{=} \|u-P_U(v)\|^2 + \|P_U(v)-v\|^2 \geq \|P_U(v)-v\|^2,$$

wobei wir in $(*)$ die in Aufgabe 11.5 bewiesene Version des Satzes des Pythagoras benutzt haben. Aus dieser Ungleichung folgt durch Radizieren sofort

$$P_U(v) = \arg\min\{\,\|u-v\| \mid u \in U\} = \pi_U(v).$$

□

Ab sofort bezeichnen wir die orthogonale Projektion auf einen Unterraum U mit P_U und können entscheiden, ob wir mit den Eigenschaften aus Definition 11.65 oder mit (11.3.1) aus Definition 11.68 arbeiten.

11.4 Orthogonale und unitäre Abbildungen und Matrizen, Isometrien

11.4.1 Orthogonale und unitäre Abbildungen

Definition 11.71

Seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei euklidische Vektorräume. Eine Abbildung $f\colon V \to W$ heißt *orthogonal*, wenn folgende Eigenschaft erfüllt ist:

$$\langle f(v), f(v')\rangle_W = \langle v, v'\rangle_V, \quad \forall v, v' \in V. \tag{11.4.1}$$

Bemerkung 11.72. *In* (11.4.1) *stecken zwei verschiedene Skalarprodukte. Wenn V und W zwei verschiedene Räume sind, beispielsweise $\mathbb{R}^m$ und $\mathbb{R}^n$ mit $m \neq n$, so besteht in der Regel keine Verwechslungsgefahr, um welches Skalarprodukt es sich jeweils handelt, so dass man kurz $\langle\cdot,\cdot\rangle$ (ohne Index) schreiben kann.*

Beispiel 11.73. *Sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch $f\begin{pmatrix} x \\ y\end{pmatrix} := \begin{pmatrix} y \\ x\end{pmatrix}$ und sei $\langle v, w\rangle := v^T w$ für $v, w \in \mathbb{R}^2$ das kanonische Skalarprodukt des $\mathbb{R}^2$. Dann gilt*

$$\langle f\begin{pmatrix} v_1 \\ v_2\end{pmatrix}, f\begin{pmatrix} w_1 \\ w_2\end{pmatrix}\rangle = \langle\begin{pmatrix} v_2 \\ v_1\end{pmatrix}, \begin{pmatrix} w_2 \\ w_1\end{pmatrix}\rangle = v_2w_2 + v_1w_1 = v_1w_1 + v_2w_2 = \langle\begin{pmatrix} v_1 \\ v_2\end{pmatrix}, \begin{pmatrix} w_1 \\ w_2\end{pmatrix}\rangle.$$

Folglich ist f orthogonal.

Für orthogonale Abbildungen $f\colon V \to W$ gilt

$$\langle v, w\rangle = 0 \Longrightarrow \langle f(v), f(w)\rangle = 0. \tag{11.4.2}$$

Allerdings impliziert (11.4.2) nicht, dass eine Abbildung orthogonal ist.

Beispiel 11.74. *Für ein $n \in \mathbb{N}$ und $a \in \mathbb{R}$ sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ gegeben durch $f(v) = a \cdot v$. Dann gilt für Vektoren $v, w \in \mathbb{R}^n$*

$$\langle v, w\rangle = 0 \implies \langle f(v), f(w)\rangle = \langle a \cdot v, a \cdot w\rangle \overset{f \text{ bilinear}}{=} a^2\langle v, w\rangle = 0.$$

Andererseits sieht man bereits an dieser Berechnung, dass f genau dann orthogonal ist, wenn $a^2 = 1$ ist, d.h. für $a \in \{\pm 1\}$.

Definition 11.75

Seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei unitäre Vektorräume. Eine Abbildung $f\colon V \to W$ heißt *unitär*, wenn folgende Eigenschaft erfüllt ist:

$$\langle f(v), f(v')\rangle_W = \langle v, v'\rangle_V, \quad \forall v, v' \in V.$$

Beispiel 11.76. *Sei $f\colon \mathbb{C}^2 \to \mathbb{C}^2$ gegeben durch*

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} iy \\ ix \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Bezogen auf das kanonische Skalarprodukt $\langle v, w\rangle := v^T\overline{w}$ ist dies eine unitäre Abbildung, denn für $x, y, v, w \in \mathbb{C}$ ist

$$\begin{aligned} \langle f\begin{pmatrix} x \\ y \end{pmatrix}, f\begin{pmatrix} v \\ w \end{pmatrix}\rangle &= \langle \begin{pmatrix} iy \\ ix \end{pmatrix}, \begin{pmatrix} iw \\ iv \end{pmatrix}\rangle \\ &= iy \cdot \overline{iw} + ix \cdot \overline{iv} \\ &= i \cdot y \cdot (-i) \cdot \overline{w} + i \cdot x \cdot (-i) \cdot \overline{v} \\ &= y \cdot \overline{w} + x \cdot \overline{v} \\ &= x \cdot \overline{v} + y \cdot \overline{w} \\ &= \langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} v \\ w \end{pmatrix}\rangle. \end{aligned}$$

In dieser Vorlesung befassen wir uns hauptsächlich mit linearen Abbildungen – auch orthogonale und unitäre Abbildungen gehören dazu, wie folgendes Lemma besagt.

Satz 11.77

Seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei euklidische (unitäre) Vektorräume. Ist $f\colon V \to W$ orthogonal (unitär), dann ist f eine injektive lineare Abbildung. Ist zudem $V = W$ mit $\dim(V) < \infty$, so ist f ein Isomorphismus.

Beweis. Der Beweis ist eine Übungsaufgabe. □

Beispiel 11.78. *Die orthogonale Abbildung f aus Beispiel 11.73 ist offensichtlich linear, denn*

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Sie ist auch bijektiv, denn die darstellende Matrix von f ist regulär. Somit ist f in der Tat ein Isomorphismus.

Es ist auf Grund der Bezeichnung naheliegend, zu vermuten, dass orthogonale Projektionen immer orthogonale Abbildungen im Sinne der obigen Definition seien. Das folgende Beispiel illustriert, dass das jedoch nicht der Fall ist. Die intuitive Erklärung ist die, dass orthogonale Abbildungen winkeltreu sind, wohingegen die orthogonale Projektion Winkel zwischen Vektoren in der Regel verändert.

Beispiel 11.79. *Sei $V = \mathbb{R}^3$ ausgestattet mit dem kanonischen Skalarprodukt und sei*

$$\mathcal{B} := \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

eine ONB von $U = \mathcal{L}(B)$.

Dann gilt mit der Formel aus Beispiel 11.64

$$P_U \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$
$$P_U \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

und somit ist

$$\langle P_U \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, P_U \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle = 0 \neq 1 = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \rangle.$$

Damit ist bewiesen, dass die Projektion P_U keine orthogonale Abbildung ist.

11.4.2 Isometrien

Orthogonale oder unitäre Abbildungen werden auch als *winkelerhaltend* oder *winkeltreu* bezeichnet. Analog dazu heißt eine Abbildung $f\colon V \to W$ für zwei euklidische oder unitäre Vektorräume V und W *abstandserhaltend* oder *längentreu*, wenn $\|f(v)\| = \|v\|$ ist für alle $v \in V$.

Den Begriff einer abstandserhaltenden Abbildung gibt es für allgemeine Mengen M.

Definition 11.80

Sei M eine beliebige Menge. Eine Abbildung $d\colon M \times M \to \mathbb{R}$ heißt *Metrik*, falls gilt:

1. $d(x,y) \geq 0$ für alle $x, y \in M$ und $d(x,y) = 0 \iff x = y$;
2. $d(x,y) = d(y,x)$ für alle $x, y \in M$;
3. $d(x,z) \leq d(x,y) + d(y,z)$ für alle $x, y, z \in M$.

Eine Menge, auf der eine Metrik definiert ist, heißt *metrischer Raum*.

Eine Abbildung zwischen zwei metrischen Räumen heißt dann *abstandserhaltend* oder *isometrisch*, wenn die Bilder zweier Elemente aus M den gleichen Abstand haben wie die Punkte selbst. Formal lautet die Definition folgendermaßen:

Definition 11.81

Seien (X, d_X) und (Y, d_Y) zwei metrische Räume. Eine Abbildung $f\colon X \to Y$ heißt *Isometrie*, falls

$$d(f(x_1), f(x_2))_Y = d(x_1, x_2)_X, \quad \forall x_1, x_2 \in X.$$

Da wir uns gerade auf euklidische oder unitäre Vektorräume beschränken, haben wir auf einem solchen Vektorraum $V, \langle \cdot, \cdot \rangle$ immer eine natürliche Metrik, nämlich

$$d(v, w) := \|v - w\| = \sqrt{\langle v - w, v - w \rangle}.$$

Für eine lineare Abbildung $f \in L(V; W)$ ist dann

$$d(f(v_1), f(v_2)) = \|f(v_1) - f(v_2)\| \overset{f \text{ linear}}{=} \|f(v_1 - v_2)\|, \quad \forall v_1, v_2 \in V,$$

d.h. f ist genau dann eine Isometrie, wenn $\|f(v)\| = \|v\|$ ist für alle $v \in V$.

Satz 11.82

Seien V und W euklidische (unitäre) Vektorräume und sei $f \in L(V; W)$. Die Abbildung f ist genau dann orthogonal (unitär), wenn f isometrisch ist, d.h. wenn Folgendes gilt:

$$\|f(v)\| = \|v\|, \quad \forall v \in V. \tag{11.4.3}$$

Beweis. Wir beweisen beide Implikationen separat für euklidische Vektorräume.

„$\Longrightarrow$“: Ist f orthogonal, dann gilt für alle $v \in V$

$$\|f(v)\| = \sqrt{\langle f(v), f(v) \rangle} \overset{f \text{ orthogonal}}{=} \sqrt{\langle v, v \rangle} = \|v\|.$$

„$\Longleftarrow$“: Ist $f \in L(V; W)$ isometrisch, dann gilt für $v, w \in V$

$$\begin{aligned}
\langle f(v) + f(w), f(v) + f(w) \rangle &\overset{f \text{ linear}}{=} \langle f(v + w), f(v + w) \rangle = \|f(v + w)\|^2 \\
&\overset{f \text{ isom.}}{=} \|v + w\|^2 = \langle v + w, v + w \rangle \\
&= \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle.
\end{aligned}$$

Andererseits gilt

$$\begin{aligned}
\langle f(v) + f(w), f(v) + f(w) \rangle &= \|f(v)\|^2 + \|f(w)\|^2 + 2\langle f(v), f(w) \rangle \\
&\overset{f \text{ isom.}}{=} \|v\|^2 + \|w\|^2 + 2\langle f(v), f(w) \rangle.
\end{aligned}$$

Gleichsetzen dieser Ausdrücke und Umstellen ergibt $\langle f(v), f(w) \rangle = \langle v, w \rangle$.

□

Satz 11.83

Seien V, W euklidische (unitäre) Vektorräume und sei $\mathcal{B} := \{b_1, \dots, b_n\}$ eine ONB von V. Eine lineare Abbildung $f \in L(V; W)$ ist genau dann orthogonal (unitär), wenn $\{f(b_1), \dots, f(b_n)\}$ ein Orthonormalsystem in W ist.

Beweis. Übungsaufgabe. □

11.4.3 Orthogonale und unitäre Endomorphismen und Matrizen

Während wir orthogonale Abbildungen für beliebige Kombinationen aus Ausgangs- und Zielräumen definiert haben (vorausgesetzt beide sind euklidisch oder beide sind unitär!), können nur quadratische Matrizen orthogonal oder unitär sein.

Definition 11.84

1. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *orthogonal*, wenn $A^T \cdot A = E_n$ ist.
2. Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt *unitär*, wenn $\overline{A^T} \cdot A = E_n$ ist.

Notation

Ist $A \in \mathbb{C}^{n\times n}$, so nennt man $A^H := \overline{A^T} = (\overline{A})^T$ auch die zugehörige *adjungierte Matrix*.

Bemerkung 11.85. *Es gibt noch diverse andere Notationen für die zu A adjungierte Matrix, z.B. A^*, $\mathrm{adj}(A)$, A^+, jedoch sind diese Notationen alle doppelt belegt – A^* wird in der Literatur auch für die komplex konjugierte Matrix $(\overline{A})$ benutzt, A^+ für die Pseudoinverse und $\mathrm{adj}(A)$ für die Adjunkte. Der Superindex H steht für* Hermite.

Beispiel 11.86. *Betrachten wir die Matrizen*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \textit{und} \quad B = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

$A \in \mathbb{R}^{2\times 2}$ ist orthogonal und $B \in \mathbb{C}^{2\times 2}$ ist unitär, denn

$$A^T A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$B^H B = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Lemma 11.87

Eine Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann orthogonal, wenn ihre Spalten ein Orthonormalsystem bezüglich des kanonischen Skalarproduktes bilden.

Beweis. Seien $v_1, \dots, v_n$ die Spalten von A. $A \in \mathbb{R}^{n\times n}$ ist nach Definition genau dann orthogonal, wenn $A^T A = E_n$ ist. Das bedeutet wiederum

$$\langle v_k, v_j\rangle = \begin{cases} 1, & k = j, \\ 0, & k \neq j, \end{cases}$$

d.h. die Menge der Spalten von A, $\{v_1, \dots, v_n\}$, bildet eine ONB des $\mathbb{R}^n$. □

Notation

Zum Arbeiten mit ONBs ist die Verwendung des *Kronecker-Delta* nützlich:

$$\delta_{i,j} := \begin{cases} 1 & \text{, falls } i = j, \\ 0 & \text{, falls } i \neq j. \end{cases}$$

Wie hängen orthogonale bzw. unitäre Matrizen und Endomorphismen zusammen? In Beispiel 11.86 haben wir gesehen, dass die darstellenden Matrizen der orthogonalen bzw. unitären Abbildungen aus Beispiel 11.73 und 11.76 selbst unitär sind. Diesen Ansatz wollen wir verallgemeinern.

Ist $A = A_f^{\mathcal{B}}$ die darstellende Matrix eines orthogonalen Endomorphismus $f \in L(\mathbb{R}^n; \mathbb{R}^n)$ bezüglich der Standardbasis $\mathcal{B} = (b_1, \dots, b_n)$ des $\mathbb{R}^n$, so ist $f(v) = A \cdot v$ für alle $v \in \mathbb{R}^n$ und somit ist

$$\langle v, w\rangle \overset{f \text{ orth.}}{=} \langle f(v), f(w)\rangle = \langle Av, Aw\rangle$$

Handelt es sich bei dem Skalarprodukt um das Standardskalarprodukt $\langle v, w\rangle = v^T w$, so folgt hieraus

$$v^T w = (Av)^T (Aw) = v^T (A^T A) w.$$

Soll diese Gleichung für alle $v, w \in \mathbb{R}^n$ erfüllt sein, so muss $A^T A = E_n$ sein, d.h. A ist orthogonal.

Satz 11.88

1. Eine Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann orthogonal, wenn sie die darstellende Matrix eines orthogonalen Endomorphismus $f \in L(\mathbb{R}^n; \mathbb{R}^n)$ bezüglich des kanonischen Skalarproduktes des $\mathbb{R}^n$ und bezüglich einer ONB des $\mathbb{R}^n$ ist.

2. Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann unitär, wenn sie die darstellende Matrix eines unitären Endomorphismus $f \in L(\mathbb{C}^n; \mathbb{C}^n)$ bezüglich des kanonischen Skalarproduktes des $\mathbb{C}^n$ und bezüglich einer ONB des $\mathbb{C}^n$ ist.

Beweis. Wir zeigen nur die Aussage für orthogonale Matrizen. Die Aussage für unitäre Matrizen wird analog bewiesen.

Wenn A orthogonal ist, dann gilt für $f\colon \mathbb{R}^n \to \mathbb{R}^n$ mit $f(x) = A \cdot x$, dass A die darstellende Matrix von f bezüglich A, bezogen auf die Standardbasis des $\mathbb{R}^n$, einer ONB, ist. In diesem Fall ist f orthogonal bezüglich des kanonischen Skalarproduktes, denn für alle $v, w \in \mathbb{R}^n$ gilt dann

$$\langle f(v), f(w)\rangle = \langle Av, Aw\rangle = (Av)^T (Aw) = v^T A^T A w = v^T E_n w = v^T w = \langle v, w\rangle.$$

Sei andererseits f orthogonal bezüglich des kanonischen Skalarproduktes und $A = A_f^{\mathcal{C}}$ für eine ONB $\mathcal{C}$ des $\mathbb{R}^n$. Sei C die Matrix, deren Spalten die Vektoren der ONB $\mathcal{C}$ sind. Dann ist C nach Lemma 11.87 orthogonal und $A_f^{\mathcal{B}} = CA_f^{\mathcal{C}}C^{-1}$ ist die darstellende Matrix von f bezüglich der kanonischen Basis, d.h. $f(x) = CAC^{-1}x$. Damit gilt wiederum

$$\begin{aligned}\langle v, w\rangle = \langle f(v), f(w)\rangle &= (CAC^{-1}v)^T \cdot CAC^{-1}w \\ &= v^T(C^{-1})^T A^T C^T CA(C^{-1})^T w \\ &= v^T(C^{-1})^T A^T E_n AC^{-1}w.\end{aligned}$$

Damit das für alle $v, w \in \mathbb{R}^n$ gilt, muss $(C^{-1})^T A^T AC^{-1} = E_n$ sein. Daraus folgt jedoch durch Multiplikation mit C^T von links und mit C von rechts:

$$A^T A = C^T C \overset{C \text{ orthogonal}}{=} E_n,$$

d.h. auch A ist orthogonal.

□

Bemerkung 11.89. *Man kann die Aussage des Satzes auch mit Hilfe von Lemma 11.87 zeigen. Die Idee ist, dass wenn $\mathcal{B} = (b_1, \ldots, b_n)$ die kanonische Basis des $\mathbb{R}^n$ ist, die Spalten von A gegeben sind durch die Vektoren $f(b_k)$ und f ist laut Satz 11.83 genau dann orthogonal, wenn diese Vektoren ein Orthonormalsystem, also eine ONB des $\mathbb{R}^n$ bilden.*

Lemma 11.90

Sei $A \in \mathbb{R}^{n\times n}$ orthogonal, d.h. es gelte $A^T A = E_n$. Dann gilt:

1. $\det A \in \{\pm 1\}$.
2. A ist invertierbar mit Inverser $A^{-1} = A^T$.

Beweis.

1. Nach der Produktregel für Determinanten (Eigenschaft (E8) in Satz 8.16) und Eigenschaft (D3) aus Definition 8.12 einer Determinanten ist

   $$1 = \det(E_n) = \det(A^T A) = \det(A^T) \cdot \det(A).$$

   Zudem gilt nach Korollar 8.23 $\det(A) = \det(A^T)$, d.h. es ist $1 = \det(A)^2$, woraus $\det(A) \in \{\pm 1\}$ folgt.

2. Aus $\det(A) \neq 0$ folgt sofort die Invertierbarkeit von A. Zudem folgt aus $A^T A = E_n$ durch Multiplikation mit A^{-1} von rechts, dass $A^T = A^{-1}$ ist.

□

Bemerkung 11.91. *Ist* $\det(A) = 1$*, so heißt die zugehörige Isometrie* orientierungserhaltend*, ist* $\det(A) = -1$*, so heißt sie* orientierungsvertauschend *oder* ungleichsinnig.

Lemma 11.92

Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $f\colon V \to V$ orthogonal. Dann kann f nur die Eigenwerte 1 und -1 haben.

Beweis. Sei v ein Eigenvektor von f zum Eigenwert λ. Dann ist

$$\langle v, v \rangle \overset{f \text{ orthogonal}}{=} \langle f(v), f(v) \rangle \overset{\text{EV}}{=} \langle \lambda v, \lambda v \rangle \overset{\text{SKP bilinear}}{=} \lambda^2 \langle v, v \rangle.$$

Da v als Eigenvektor nicht der Nullvektor sein kann, ist $\langle v, v \rangle > 0$ dank der positiven Definitheit des Skalarproduktes. Aus

$$(1 - \lambda^2)\langle v, v \rangle = 0$$

folgt damit $1 - \lambda^2 = 0$, d.h. $\lambda \in \{-1, 1\}$. □

Bemerkung 11.93. *In Beispiel 11.74 hatten wir bereits einen Spezialfall dieser Aussage gesehen, nämlich dass $f(v) = a \cdot v$ genau dann orthogonal ist, wenn $a \in \{\pm 1\}$ ist – und genau das ist dann auch der einzige auftretende Eigenwert.*

Kommen wir zum Analogon für unitäre Abbildungen.

Lemma 11.94

Sei $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum und $f\colon V \to V$ unitär. Ist $\lambda \in \mathbb{C}$ ein Eigenwert von f, so gilt $|\lambda| = 1$.

Beweis. Sei v ein Eigenvektor von f zum Eigenwert λ. Dann ist

$$\langle v, v \rangle \overset{f \text{ orthogonal}}{=} \langle f(v), f(v) \rangle \overset{\text{EV}}{=} \langle \lambda v, \lambda v \rangle \overset{\text{SKP sesquilinear}}{=} \lambda\overline{\lambda}\langle v, v \rangle = |\lambda|^2 \langle v, v \rangle.$$

Da v als Eigenvektor nicht der Nullvektor sein kann, ist $\langle v, v \rangle > 0$ dank der positiven Definitheit des Skalarproduktes. Aus

$$(1 - |\lambda|^2)\langle v, v \rangle = 0$$

folgt damit $1 - |\lambda|^2 = 0$, d.h. $|\lambda| = 1$. □

Nachdem wir die Eigenwerte angesehen haben, können wir nun auch Aussagen über die Eigenvektoren treffen.

Lemma 11.95

Eigenvektoren orthogonaler (unitärer) Abbildungen zu verschiedenen Eigenwerten sind orthogonal zueinander.

Beweis. Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und sei $f\colon V \to V$ eine orthogonale Abbildung. Damit hat f höchstens die zwei verschiedenen Eigenwerte $\lambda = 1$ und $\mu = -1$. Seien $v, w \in V$ Eigenvektoren zu diesen Eigenwerten. Dann gilt

$$\langle v, w\rangle \overset{f \text{ orthogonal}}{=} \langle f(v), f(w)\rangle \overset{\text{EV}}{=} \langle \lambda v, \mu w\rangle \overset{\text{SKP bilinear}}{=} \lambda\mu\langle v, w\rangle = -\langle v, w\rangle.$$

Hieraus folgt $\langle v, w\rangle = 0$.

Sei nun $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum und sei $f\colon V \to V$ eine unitäre Abbildung. Seien $\lambda, \mu \in \mathbb{C}$ zwei verschiedene Eigenwerte von f und seien $v, w \in V$ zugehörige Eigenvektoren. Dann gilt

$$\langle v, w\rangle \overset{f \text{ unitär}}{=} \langle f(v), f(w)\rangle \overset{\text{EV}}{=} \langle \lambda v, \mu w\rangle \overset{\text{SKP sesquilin.}}{=} \lambda\overline{\mu}\langle v, w\rangle.$$

Hieraus folgt entweder $\langle v, w\rangle = 0$ oder $\lambda\overline{\mu} = 1$, d.h. $\overline{\mu} = \lambda^{-1}$. Nehmen wir an, Letzteres wäre der Fall. Ist $\lambda = a + ib$ mit Betrag $|\lambda|^2 = a^2 + b^2 = 1$ und Inverser $\lambda^{-1} = \frac{a-ib}{a^2+b^2} = a - ib = \overline{\lambda}$, so folgt hiermit für den anderen Eigenwert μ:

$$\mu = \overline{\overline{\mu}} = \overline{\lambda^{-1}} = \overline{\overline{\lambda}} = \lambda,$$

was der Annahme $\mu \neq \lambda$ widerspricht. Somit muss für Eigenwerte $\mu \neq \lambda$ die Orthogonalität der zugehörigen Eigenvektoren gelten. □

Die Menge aller orthogonalen $n \times n$-Matrizen

Satz 11.96

Für ein $n \in \mathbb{N}$ sei

$$\mathcal{O}(n) := \left\{ A \in \mathbb{R}^{n\times n} \,\middle|\, A^T A = E_n \right\}$$

die Menge aller orthogonalen $n \times n$-Matrizen. Mit der Matrixmultiplikation $\cdot$ als Verknüpfung ist $(\mathcal{O}(n), \cdot)$ eine Gruppe.

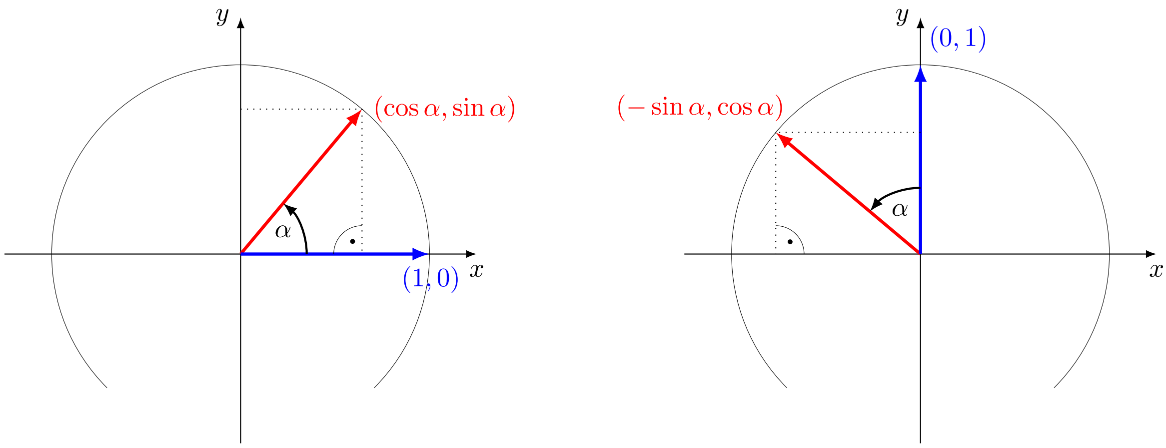
Beweis. Das wird in Aufgabe 11.11 gezeigt. □

Für den Spezialfall $n = 2$ können wir diese Gruppe genau charakterisieren. Zunächst einmal ist es intuitiv einleuchtend, dass Drehungen winkeltreu und somit orthogonale Abbildungen sind. Gleiches gilt für Spiegelungen. (Der Beweis dafür kommt noch.) Für den Fall $n = 2$ können wir sogar zeigen, dass Drehungen und Spiegelungen die einzigen orthogonalen Abbildungen aus $L(\mathbb{R}^2; \mathbb{R}^2)$ sind.

In Beispiel 8.6 und später nochmal in Beispiel 10.25 hatten wir uns die Drehung um den Winkel α um den Koordinatenursprung in mathematisch positivem Sinn angesehen. Die darstellende Matrix dieser Abbildung bezüglich der Standardbasis des $\mathbb{R}^2$ ist die Drehmatrix

$$D_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

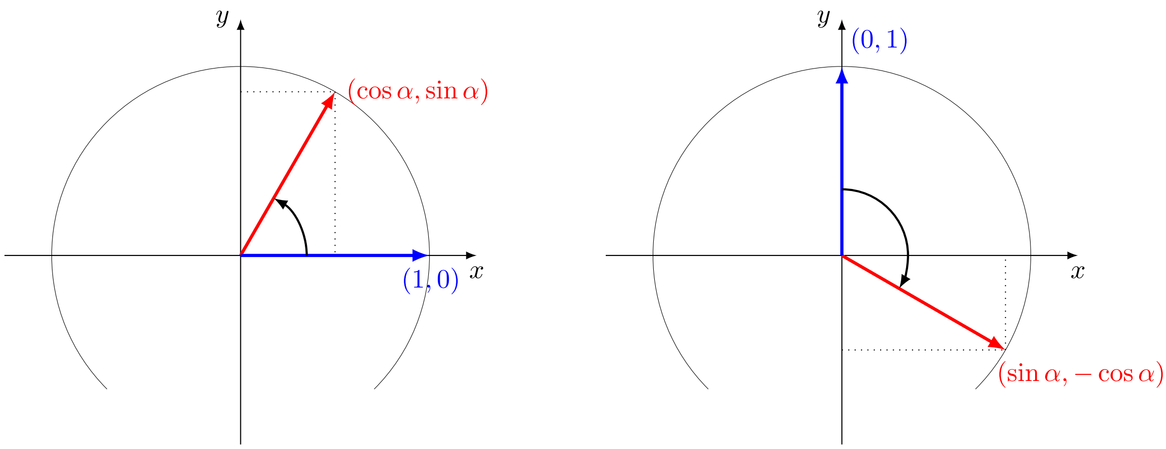
Ihre Wirkung haben wir mit Hilfe der Vektoren der Standardbasis graphisch dargestellt:



Andererseits sei

$$S_\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

Sehen wir uns zunächst wieder anhand der Vektoren der Standardbasis an, was S_α, bzw. die zugehörige lineare Abbildung, bewirkt.

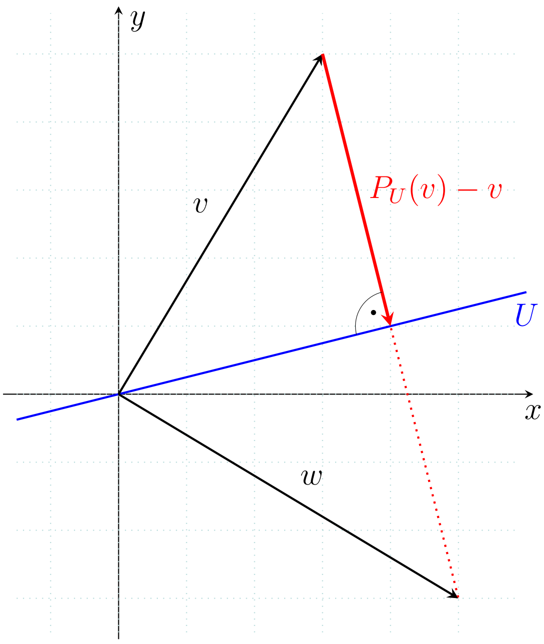


Anhand der Wirkung auf die Basisvektoren lässt sich vermuten, dass S_α eine Spiegelung an der Geraden

$$U = \mathcal{L}\begin{pmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \end{pmatrix}$$

bewirkt. Überprüfen wir diese Vermutung:

Sei $v \in \mathbb{R}^2$ beliebig. Sei $w \in \mathbb{R}^2$ der Vektor, der durch Spiegelung von v an der durch U gegebenen Geraden entsteht. Dann ist für $\gamma = \frac{\alpha}{2}$



$$
\begin{aligned}
w &= v + 2(P_U(v) - v) = 2P_U(v) - v \\
&= 2\langle v, \begin{pmatrix}\cos\gamma\\ \sin\gamma\end{pmatrix}\rangle \begin{pmatrix}\cos\gamma\\ \sin\gamma\end{pmatrix} - v \\
&= 2(v_1\cos\gamma + v_2\sin\gamma)\begin{pmatrix}\cos\gamma\\ \sin\gamma\end{pmatrix} - \begin{pmatrix}v_1\\ v_2\end{pmatrix} \\
&= \begin{pmatrix}2v_1\cos^2\gamma + 2v_2\cos\gamma\sin\gamma - v_1\\ 2v_1\cos\gamma\sin\gamma + 2v_2\sin^2\gamma - v_2\end{pmatrix} \\
&= \begin{pmatrix}2\cos^2\gamma - 1 & 2\cos\gamma\sin\gamma\\ 2\cos\gamma\sin\gamma & 2\sin^2\gamma - 1\end{pmatrix}\begin{pmatrix}v_1\\ v_2\end{pmatrix} \\
&\overset{(*)}{=} \begin{pmatrix}\cos(2\gamma) & \sin(2\gamma)\\ \sin(2\gamma) & -\cos(2\gamma)\end{pmatrix}\begin{pmatrix}v_1\\ v_2\end{pmatrix} \\
&= \begin{pmatrix}\cos\alpha & \sin\alpha\\ \sin\alpha & -\cos\alpha\end{pmatrix} v \\
&= S_\alpha v,
\end{aligned}
$$

wobei wir in $(*)$ die Additionstheoreme

$$
\begin{aligned}
&\cos(2\gamma) = 2\cos^2\gamma - 1 = \cos^2\gamma - \sin^2\gamma = 1 - \sin^2\gamma \qquad \text{und}\\
&\sin(2\gamma) = 2\sin\gamma\cos\gamma
\end{aligned}
$$

benutzt haben.

Die Rechnung bestätigt, dass S_α eine Spiegelung an der durch U gegebenen Geraden bewirkt. Mit diesen Vorüberlegungen können wir zeigen, dass die orthogonalen Matrizen im $\mathbb{R}^{2\times 2}$ in der Tat durch die Drehungen und Spiegelungen gegeben sind.

Lemma 11.97

Sei $\mathcal{O}(2) = \left\{ A \in \mathbb{R}^{2\times 2} \,\middle|\, A^T A = E_2 \right\}$. Dann gilt:

$$\mathcal{O}(2) = \{ D_\alpha \mid \alpha \in [0, 2\pi) \} \cup \{ S_\alpha \mid \alpha \in [0, 2\pi) \}\,.$$

Beweis. In zwei Übungsaufgaben wird gezeigt, dass

$$\mathcal{O}(2) = \mathcal{M} := \left\{ \begin{pmatrix} a & b\\ -b & a\end{pmatrix} \,\middle|\, a, b \in \mathbb{R}, a^2 + b^2 = 1 \right\} \cup \left\{ \begin{pmatrix} a & b\\ b & -a\end{pmatrix} \,\middle|\, a, b \in \mathbb{R}, a^2 + b^2 = 1 \right\}$$

ist. Einerseits erfüllen $a = \cos\alpha$ und $b = \sin\alpha$ für jedes $\alpha \in [0, 2\pi)$ die Bedingung $a^2 + b^2 = 1$, woraus folgt, dass die Dreh- und Spiegelungsmatrizen D_α und S_α in $\mathcal{M}$ liegen. Andererseits gibt es für jede Matrix $A \in \mathcal{M}$ einen Wert $\alpha \in [0, 2\pi)$, so dass A einer Dreh- oder Spiegelungsmatrix entspricht mit $a = \cos\alpha$ und $b = \sin\alpha$. □

11.5 Aufgaben

Aufgabe 11.1. *Sei V ein $\mathbb{R}$-Vektorraum. Zeigen Sie, dass*

$$b\colon V \times L(V; \mathbb{R}) \to \mathbb{R}, \qquad b(v, f) := f(v)$$

eine Bilinearform ist.

Aufgabe 11.2. *Skizzieren Sie die folgenden Mengen:*

a) $M_1 := \{x \in \mathbb{R}^2 \mid \|x\|_1 = 1\}$, *wobei* $\|x\|_1 = |x_1| + |x_2|$ *ist;*

b) $M_2 := \{x \in \mathbb{R}^2 \mid \|x\|_2 = 1\}$, *wobei* $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$ *ist;*

c) $M_\infty := \{x \in \mathbb{R}^2 \mid \|x\|_\infty = 1\}$, *wobei* $\|x\|_\infty = \max\{|x_1|, |x_2|\}$ *ist.*

Aufgabe 11.3. *Das kanonische Skalarprodukt über dem* $\mathbb{C}$*-Vektorraum* $\mathbb{C}^n$ *(mit* $n \in \mathbb{N}$*) ist definiert als*

$$\langle v, w \rangle := \sum_{k=1}^{n} v_k \overline{w_k}.$$

Überprüfen Sie, dass es sich dabei tatsächlich um ein Skalarprodukt handelt.

Aufgabe 11.4. *Seien* $v = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, w = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \in \mathbb{R}^2$.

a) *Skizzieren Sie das Parallelogramm, welches von den Vektoren* v *und* w *aufgespannt wird. Welche Vektoren repräsentieren die Diagonalen des Parallelogramms?*

b) *In einem Parallelogramm mit Seitenlängen* a *und* b *und Diagonalen mit Längen* e *und* f *gilt*

$$2(a^2 + b^2) = e^2 + f^2.$$

Übersetzen Sie diese Gleichung mit Hilfe der Vektoren und der euklidischen Norm als Länge eines Vektors.

c) *Überprüfen Sie die Gültigkeit der in b) hergeleiteten* Parallelogrammgleichung *für die konkret vorgegebenen Vektoren* v *und* w.

Aufgabe 11.5. *Beweisen Sie folgende abstrakte Version des Satzes des Pythagoras:*

Satz 11.98: Pythagoras

Sei V *ein euklidischer oder unitärer Vektorraum und seien* $v_1, \dots, v_n \subseteq V$ *paarweise orthogonale Vektoren. Dann gilt für die durch das Skalarprodukt induzierte Norm:*

$$\left\| \sum_{k=1}^{n} v_k \right\|^2 = \sum_{k=1}^{n} \|v_k\|^2 .$$

Aufgabe 11.6. *Konstruieren Sie mit Hilfe des Gram-Schmidt'schen Orthonormalisierungsverfahrens eine ONB von* $\mathcal{L}(M)$ *bezüglich des kanonischen Skalarproduktes, wobei* M *folgende Menge sei:*

$$M = \left(\begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 8 \\ 5 \\ -6 \end{pmatrix} \right).$$

Aufgabe 11.7. *Sei* $U \subseteq V$ *ein Untervektorraum eines euklidischen Vektorraumes. Zeigen Sie, dass für das orthogonale Komplement von* U *bezüglich* V *folgende Inklusion gilt:*

$$U \subseteq (U^\perp)^\perp.$$

Gilt die Aussage auch für den UVR eines unitären Vektorraumes?

Aufgabe 11.8. *In dieser Aufgabe werden ein paar Feinheiten zum Vektorraum der komplexen Zahlen untersucht.*

a) *Geben Sie eine ONB des unitären Vektorraumes $V = \mathbb{C}$ an, wobei das zugrunde liegende Skalarprodukt das kanonische Skalarprodukt $\langle v, w\rangle := v \cdot \overline{w}$ sei.*

b) *Zeigen Sie, dass $V = \mathbb{C}$ auch als $\mathbb{R}$-Vektorraum aufgefasst werden kann, indem Sie die Vektorraumeigenschaften überprüfen. Geben Sie eine Basis dieses Vektorraumes an.*

c) *Wird durch $\langle v, w\rangle := v \cdot w$ ein Skalarprodukt auf dem $\mathbb{R}$-Vektorraum $\widetilde{V} = \mathbb{R}$ definiert? Und auf dem $\mathbb{R}$-Vektorraum $V = \mathbb{C}$?*

d) *Der Körper der komplexen Zahlen wird auch als Menge $\mathbb{C} := \mathbb{R}^2$ mit folgenden Verknüpfungen definiert:*

$$(a, b) + (c, d) := (a + c, b + d), \quad (a, b) \cdot (c, d) := (ac - bd, bc + ad).$$

Finden Sie eine Basis des so definierten $\mathbb{R}$-Vektorraumes $\mathbb{C}$ und überprüfen Sie, dass

$$\langle (a, b), (c, d)\rangle := ac + bd$$

ein Skalarprodukt darauf definiert.

e) *Für $z \in \mathbb{C}$ sei $|z|$ der Betrag der komplexen Zahl z. Begründen Sie, dass der Betrag eine Norm auf dem Vektorraum $V = \mathbb{C}$ ist – sowohl für den $\mathbb{R}$- als auch für den $\mathbb{C}$-Vektorraum $\mathbb{C}$.*

f) *Kann man $\widetilde{V} = \mathbb{R}$ auch als $\mathbb{C}$-Vektorraum auffassen?*

Aufgabe 11.9. *Eine Ebene im $\mathbb{R}^3$ ist genau dann ein UVR des $\mathbb{R}^3$, wenn sie den Nullvektor enthält. In diesem Fall hat sie die Form*

$$\mathcal{E} = \{ a \cdot v + b \cdot w \mid a, b \in \mathbb{R}\} = \mathcal{L}(\{v, w\})$$

für zwei linear unabhängige Vektoren

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \textit{und} \quad w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

a) *Bestimmen Sie die Dimension von $\mathcal{E}^{\perp}$.*

b) *Das* Vektorprodukt *oder* Kreuzprodukt *zweier Vektoren $v, w \in \mathbb{R}^3$ ist definiert als*

$$v \times w := \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Zeigen Sie, dass $v \times w$ zu v und zu w orthogonal ist bezüglich des kanonischen Skalarproduktes des $\mathbb{R}^3$. Mit anderen Worten, zeigen Sie Folgendes:

$$v^T(v \times w) = 0 \quad \textit{und} \quad w^T(v \times w) = 0.$$

c) *Sei*

$$\mathcal{E}' := \left\{ u \in \mathbb{R}^3 \,\middle|\, \langle u, v \times w\rangle = 0 \right\}.$$

Begründen Sie, warum $\mathcal{E} = \mathcal{E}'$ ist.

Hintergrundinformation:

$\mathcal{E}$ nennt man eine Ebene in Parameterform, *$\mathcal{E}'$ ist eine Ebene in* Normalenform. *Ist $v \times w = (a, b, c)^T$, so ist*

$$\mathcal{E}' = \left\{ u \in \mathbb{R}^3 \;\middle|\; au_1 + bu_2 + cu_3 = 0 \right\}.$$

Aufgabe 11.10. *Sei V ein endlich-dimensionaler euklidischer Vektorraum und sei $\mathcal{B} = (b_1, \ldots, b_m)$ eine ONB eines Untervektorraumes $U \subseteq V$ von V. Zeigen Sie, dass die orthogonale Projektion*

$$f : V \to V, \quad f(v) := \sum_{k=1}^{m} \langle v, b_k \rangle b_k$$

eine selbstadjungierte Abbildung *definiert, d.h. überprüfen Sie, dass Folgendes gilt:*

$$\langle f(v), w \rangle = \langle v, f(w) \rangle, \quad \forall v, w \in V.$$

Aufgabe 11.11. *Sei $n \in \mathbb{N}$. Dazu sei*

$$\mathcal{O}(n) := \left\{ A \in \mathbb{R}^{n \times n} \;\middle|\; A^T A = E_n \right\}.$$

Zeigen Sie:

a) $E_n \in \mathcal{O}(n)$;

b) Alle Matrizen in $\mathcal{O}(n)$ sind invertierbar;

c) Ist $A \in \mathcal{O}(n)$, so ist auch $A^{-1} \in \mathcal{O}(n)$;

d) Sind $A, B \in \mathcal{O}(n)$, so ist auch $A \cdot B \in \mathcal{O}(n)$;

e) $\mathcal{O}(n)$ ist mit der Matrizenmultiplikation als Verknüpfung eine Gruppe.

Bemerkung: *Man nennt $\mathcal{O}(n)$ auch die* orthogonale Gruppe.

Kapitel 12

Affine Räume

Ziele

- □ die Feinheiten beim Arbeiten in affinen Räumen verstehen
- □ Beispiele für affine Räume kennen
- □ die Begriffe Basis, Dimension, Unterraum, Koordinaten und Abbildungen von linearen zu affinen Räumen übertragen
- □ Dimensionsformeln für affine Räume kennen
- □ Lagebeziehungen affiner Unterräume des $\mathbb{R}^n$ beschreiben können

12.1 Motivation

Aufbauend auf dem Begriff einer Gruppe haben wir bisher zwei wichtige Strukturbegriffe kennen gelernt – den *Körper*, der insbesondere die Gruppeneigenschaften bezüglich Addition und Multiplikation besitzt und den *Vektorraum* über einem Körper, in dem Vektoren addiert und mit Körperelementen multipliziert werden können. Wir werden jetzt einen weiteren Strukturbegriff kennen lernen.

Erinnern wir uns zunächst daran, dass es für Vektoren im $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ grundsätzlich zwei verschiedene Interpretations- bzw. Darstellungsmöglichkeiten gibt: Vektoren als Punkte im Raum und Vektoren als Äquivalenzklassen von Pfeilen. Beide haben unterschiedliche Eigenschaften:

Punkte

- Punkte haben eine feste Position relativ zum Koordinatenursprung;
- Punkte haben einen Abstand voneinander;
- Punkte kann man weder addieren noch strecken.

Äquivalenzklassen von Pfeilen

- Ortsvektor hat feste Position, ansonsten sind Vektoren frei im Raum;
- Pfeile haben keinen Abstand voneinander, aber dafür eine Länge;
- Vektoren kann man addieren und strecken.

Indem wir beide Vorstellungen miteinander verknüpfen, können wir beispielsweise affine Geraden und Ebenen, d.h. Geraden und Ebenen, die den Koordinatenursprung nicht enthalten, darstellen.

Beispiel 12.1. *Betrachten wir die Mengen*

$$G = \left\{ \begin{pmatrix}1\\1\\1\end{pmatrix} + a \cdot \begin{pmatrix}1\\0\\0\end{pmatrix} \;\middle|\; a \in \mathbb{R} \right\}, \quad E = \left\{ \begin{pmatrix}1\\1\\1\end{pmatrix} + a \cdot \begin{pmatrix}0\\1\\0\end{pmatrix} + b \cdot \begin{pmatrix}0\\0\\1\end{pmatrix} \;\middle|\; b \in \mathbb{R} \right\}.$$

G und E sind nach Definition 6.58 affine Unterräume des $\mathbb{R}^3$, ersteres ist eine Gerade, letzteres eine Ebene.

Dabei stellt $\begin{pmatrix}1 & 1 & 1\end{pmatrix}^T$ einen Punkt im Raum dar, der auf der Geraden bzw. in der Ebene liegt, wohingegen die übrigen Vektoren nicht als Punkte, sondern eher als Äquivalenzklassen von Pfeilen aufgefasst werden.

Wenn wir beide Objekte – Punkte und Pfeilklassen – als verschiedene Objekte auffassen, müssen wir formal einführen, was die Summe dieser Objekte ist.

12.2 Einführung des Begriffes

Basierend auf [Fil11] führen wir den Begriff eines *affinen Raumes* ein. Wir werden später sehen, dass dieser zu dem bereits eingeführten *affinen Unterraum* passt.

Definition 12.2

Sei V ein $\mathbb{K}$-Vektorraum. Eine nichtleere Menge A heißt *affiner Raum* über dem Vektorraum V, falls eine Verknüpfung $+ : A \times V \to A$ mit folgenden Eigenschaften existiert:

(A1) $P + e_V = P$ für alle $P \in A$.

(A2) Für zwei beliebige Punkte $P, Q \in A$ existiert genau ein Vektor $v \in V$, für den $Q = P + v$ gilt.

(A3) $P + (v + w) = (P + v) + w$ für alle $P \in A$ und $v, w \in V$.

Die Elemente von A werden als *Punkte* bezeichnet.

Ist $A = \emptyset$, so wird A ebenfalls als affiner Raum aufgefasst.

Bemerkung 12.3.

i) Soll der dem affinen Raum zu Grunde liegende Vektorraum genannt werden, so schreiben wir $A(V)$ statt A.

ii) Die Forderung, dass eine Verknüpfung $+ : A \times V \to A$ existieren soll, beinhaltet (analog zur Verknüpfung in Gruppen) die Bedingung, dass es zu jedem Paar aus einem Punkt $P \in A$ und einem Vektor $v \in V$ stets einen Punkt $Q \in A$ geben muss, für den $Q = P + v$ gilt.

iii) *Gilt $Q = P + v$, so schreiben wir auch $v = \overrightarrow{PQ}$. Eigenschaft (A1) können wir damit wie folgt umschreiben:*
$$\overrightarrow{PP} = e_V, \quad \forall P \in A.$$

iv) *Ab jetzt werden wir durch das Symbol $+$ drei verschiedene Paare von Objekten addieren können: zwei Vektoren, zwei Skalare (d.h. Elemente des Körpers) und einen Punkt und einen Vektor. In der Regel ist es nicht nötig, für die drei Fälle unterschiedliche Symbole zu benutzen, da entweder aus der Art der Objekte auf die Art des $+$ geschlossen werden kann oder die Wirkung der Operationen sich ohnehin nicht unterscheidet – Letzteres ist im kommenden Beispiel der Fall. Daher verzichten wir ab jetzt auf die farbliche Unterscheidung zwischen den $+$-Symbolen.*

Beispiel 12.4. *Sei $L(A, b)$ die Lösungsmenge des inhomogenen LGS der Form $A \cdot x = b$ mit $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ und $b \in \mathbb{R}^m$. Wir haben bereits in Satz 6.71 gesehen, dass folgende Beziehung gilt:*
$$L(A, b) = x^* + L(A, 0)$$
für ein beliebiges $x^* \in L(A, b)$. Zudem ist laut Satz 6.70 $L(A, 0)$ ein UVR des $\mathbb{R}^n$ und somit selbst ein Vektorraum. Damit ist $L(A, b)$ ein affiner Raum über dem Vektorraum $L(A, 0)$ mit der Addition von Vektoren des $\mathbb{R}^n$ als Verknüpfung, denn alle geforderten Eigenschaften werden erfüllt:

0. *$+\colon L(A, b) \times L(A, 0) \to L(A, b)$ ist in der Tat eine in $L(A, b)$ abgeschlossene Verknüpfung, denn für beliebige $x \in L(A, b)$ und $y \in L(A, 0)$ ist $Ax = b$ und $Ay = e_{\mathbb{R}^m}$, also ist*
$$A(x + y) = Ax + Ay = b + e_{\mathbb{R}^m} = b,$$
d.h. $x + y \in L(A, b)$.

1. *Für beliebige Punkte $x \in L(A, b)$ gilt $x + e_{\mathbb{R}^n} = x$, da die Addition die übliche Addition des $\mathbb{R}^n$ ist und der Nullvektor das neutrale Element der Addition von Vektoren ist. Folglich ist diese Bedingung für alle $x \in \mathbb{R}^n$, also insbesondere für alle $x \in L(A, b) \subseteq \mathbb{R}^n$ erfüllt.*

2. *Sind $x, x' \in L(A, b)$, d.h. gilt $Ax = Ax' = b$, so folgt $A(x - x') = Ax - Ax' = b - b = e_{\mathbb{R}^m}$, d.h. $x - x' \in L(A, 0)$. Damit ist*
$$x = x' + (x - x').$$
Sei andererseits $y \in L(A, 0)$ derart, dass
$$x = x' + y$$
ist. Durch Subtraktion von obiger Gleichung folgt hieraus jedoch $e_{\mathbb{R}^n} = (x - x') - y$, d.h. $y = x - x'$, was die geforderte Eindeutigkeit beweist.

3. *Seien $x \in L(A, b)$ und $y, y' \in L(A, 0)$. Dann gilt*
$$x + (y + y') = (x + y) + y',$$
da in der Gruppe $(\mathbb{R}^n, +)$ das Assoziativgesetz gilt und die Lösungsmengen des homogenen und des inhomogenen LGS Teilmengen des $\mathbb{R}^n$ sind.

Beispiel 12.5. *Sei $\mathcal{P}_n$ der $\mathbb{R}$-Vektorraum der reellen Polynom vom Grad $\leq n$. Sei*

$$*\colon \mathbb{R} \times \mathcal{P}_n \to \mathbb{R}, \quad (x,p) \mapsto p(x).$$

Dann ist $A = \mathbb{R}$ mit dieser Verknüpfung kein affiner Raum über dem Vektorraum $\mathcal{P}_n$, denn Eigenschaften (A1) bis (A3) sind verletzt. Beispielsweise gilt für das Nullpolynom $p \equiv 0$ und jedes $x \in \mathbb{R} \setminus \{0\}$

$$x * p = p(x) = 0 \neq x.$$

Beispiel 12.6. *Mit der Wahl $A = \mathbb{R}^n$ und $V = \mathbb{R}^n$ und der üblichen Addition n-dimensionaler Vektoren als Verknüpfung ist der $\mathbb{R}^n$ ein affiner Punktraum.*

Satz 12.7

Sei V ein $\mathbb{K}$-Vektorraum und sei A ein affiner Raum über V. Seien $P, Q, R, S \in A$ und $v, w \in V$. Dann gelten folgende Aussagen:

(i) $P + v = P + w \Longrightarrow v = w$.

(ii) $P + v = Q + v \Longrightarrow P = Q$.

(iii) $P + v = Q \Longrightarrow P = Q + (-v)$.

(iv) $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$.

(v) Für n Punkte $P_1, \ldots, P_n \in A$ (mit $n \in \mathbb{N}_{>1}$) gilt stets

$$\overrightarrow{P_1P_2} + \overrightarrow{P_2P_3} + \ldots + \overrightarrow{P_{n-1}P_n} = \sum_{k=1}^{n-1} \overrightarrow{P_kP_{k+1}} = \overrightarrow{P_1P_n}.$$

(vi) $\overrightarrow{PQ} + \overrightarrow{QP} = e_V$ und somit ist $\overrightarrow{PQ} = -\overrightarrow{QP}$.

(vii) $\overrightarrow{PQ} = \overrightarrow{RS} \Longrightarrow \overrightarrow{PR} = \overrightarrow{QS}$.

Beweis. Wir zeigen nur einige Aussagen, den Rest zu zeigen ist eine Übungsaufgabe.

(i) Da nach Eigenschaft (A2) genau ein Vektor $v \in V$ existiert mit $Q = P + v$, folgt aus $P + w = P + v =: Q$ sofort $v = w$.

(iv) Die Vektoren $\overrightarrow{PQ}, \overrightarrow{QR}, \overrightarrow{PR} \in V$ sind definitionsgemäß die eindeutig bestimmten Vektoren, für die folgende Gleichungen gelten:

$$Q = P + \overrightarrow{PQ}, \tag{I}$$

$$R = Q + \overrightarrow{QR}, \tag{II}$$

$$R = P + \overrightarrow{PR}. \tag{III}$$

Ersetzen wir Q in (II) durch dessen Darstellung aus (I), so erhalten wir

$$R = (P + \overrightarrow{PQ}) + \overrightarrow{QR} \overset{\text{(A3)}}{=} P + (\overrightarrow{PQ} + \overrightarrow{QR}).$$

Durch Vergleich mit (III) erhalten wir wegen der Eindeutigkeit des Vektors v, für den $R = P + v$ gilt, $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$.

(v) Diese Aussage folgt durch wiederholte Anwendung von (iv).

(vii) Es gelte $\overrightarrow{PQ} = \overrightarrow{RS}$. Dann folgt

$$\overrightarrow{PR} \overset{\text{(iv)}}{=} \overrightarrow{PQ} + \overrightarrow{QR} \overset{\text{Vor.}}{=} \overrightarrow{RS} + \overrightarrow{QR} \overset{\text{KG in } V}{=} \overrightarrow{QR} + \overrightarrow{RS} \overset{\text{(iv)}}{=} \overrightarrow{QS}.$$

□

Die Dimension eines affinen Raumes definieren wir über den zu Grunde liegenden Vektorraum:

Definition 12.8

Sei A ein nichtleerer affiner Raum über einem $\mathbb{K}$-Vektorraum V. Dann definieren wir die Dimension von A als

$$\dim(A) := \dim(V).$$

Ist $A = \emptyset$, so definieren wir $\dim(A) := -1$.

Beispiel 12.9 (Fortsetzung von Beispiel 12.4)**.** *Ist* $\dim(L(A, 0)) = n$*, so ist* $\dim(L(A, b)) = n$*. Insbesondere ist für* $A \in \mathbb{R}^{n\times n}$ *mit* $\operatorname{rg}(A) = n$ *für jeden beliebigen Vektor* $b \in \mathbb{R}^n$

$$\dim(L(A, b)) = \dim(L(A, 0)) = n - \operatorname{rg}(A) = 0.$$

Zugleich wissen wir, dass $L(A, b) = \{A^{-1}b\}$ *ist, d.h. ein einzelner Punkt im* $\mathbb{R}^3$ *hat die Dimension Null.*

Beispiel 12.10 (Fortsetzung von Beispiel 12.6)**.** *Der affine Punktraum* $A = \mathbb{R}^n$ *ist ein* n*-dimensionaler affiner Raum über dem Vektorraum* $V = \mathbb{R}^n$*.*

Definition 12.11

Ist A ein affiner Raum über einem euklidischen oder unitären Vektorraum V, so nennt man A (bzw. $(A, V, \langle\cdot,\cdot\rangle, +)$) einen *euklidischen* oder *unitären Punktraum*.

In der Regel wird in der Schule im euklidischen Punktraum $A = \mathbb{R}^n$ (üblicherweise $n \in \{2, 3\}$) über $V = \mathbb{R}^n$ mit dem kanonischen Skalarprodukt $\langle v, w\rangle = v^T w = \sum_{k=1}^n v_k w_k$ und der üblichen Addition von Vektoren des $\mathbb{R}^n$ gearbeitet.

12.3 Koordinatensysteme

In Definition 12.8 haben wir die Dimension eines affinen Raumes als die Dimension des zugehörigen Vektorraumes definiert. Zwei endlich-dimensionale Vektorräume haben genau dann die gleiche Dimension, wenn sie isomorph zueinander sind. (Dieses Ergebnis wird in der Übung noch genauer besprochen.) Da es innerhalb des affinen Punktraumes keine Addition oder skalare Multiplikation gibt, ist die Suche nach einem Isomorphismus nicht sinnvoll. Eine bijektive Abbildung gibt es jedoch.

Lemma 12.12

Sei A ein affiner Raum über einem Vektorraum V. Sei zudem $O \in A$ ein fixierter Punkt des affinen Raumes. Dann definiert die Abbildung

$$\varphi\colon A \to V, \qquad P \mapsto \overrightarrow{OP},$$

eine Bijektion zwischen A und V.

Beweis. Nach Eigenschaft (A2) ist $\overrightarrow{OP}$ der eindeutig bestimmte Vektor aus V, für den $P = O + \overrightarrow{OP}$ gilt. Damit ist die Abbildung wohldefiniert, da zu jedem Punkt $P \in A$ ein solcher Vektor existiert, eindeutig bestimmt ist und im korrekten Raum, nämlich V, liegt. Überprüfen wir die Bijektivität indem wir die Umkehrabbildung konkret angeben:

$$\psi\colon V \to A, \qquad v \mapsto O + v.$$

Diese Abbildung ist ebenfalls korrekt definiert und für alle $P \in A$ gilt

$$\psi(\varphi(P)) = O + \varphi(P) = O + \overrightarrow{OP} = P$$

und für $v \in V$ und $Q := O + v$ gilt nach Definition von $\overrightarrow{OQ}$:

$$\varphi(\psi(v)) = \varphi(O + v) = \varphi(Q) = \overrightarrow{OQ} = v.$$

Damit ist gezeigt, dass $\psi = \varphi^{-1}$ ist, d.h. φ ist invertierbar und somit bijektiv. $\square$

Die Bijektion ordnet, anschaulich gesprochen, jedem Punkt dessen Ortsvektor zu. Der Punkt $O \in A$ kann jedoch frei gewählt werden - für jede solche Wahl entsteht eine neue Bijektion.

In einem Vektorraum hat jeder Vektor eine eindeutige Darstellung als Linearkombination der Vektoren einer Basis. Die Koeffizienten dieser Darstellung sind gerade die *Koordinaten* des Vektors bezüglich der gegebenen Basis. Um in einem affinen Raum ebenfalls Koordinaten einführen zu können, müssen wir nach obiger Überlegung also nicht nur eine Basis des Vektorraumes, sondern auch einen Bezugspunkt O, den wir später *Koordinatenursprung* nennen werden, fixieren.

Satz 12.13

Sei $A \neq \emptyset$ ein affiner Raum über einem $\mathbb{K}$-Vektorraum V mit einer Verknüpfung $+$. Sei $O \in A$ ein beliebiger fixierter Punkt und $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V.

Dann besitzt jeder Punkt $P \in A$ eindeutig bestimmte Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, so dass P folgende Darstellung besitzt:

$$P = O + \sum_{k=1}^{n} \lambda_k b_k.$$

Beweis. Seien $O, P \in A$ fixiert. Dann existiert nach (A2) ein eindeutig bestimmter Vektor $v \in V$, für den $P = O + v$ gilt, und den wir mit $v = \overrightarrow{OP}$ notieren. Da $\mathcal{B}$ eine (geordnete) Basis von V ist, existieren eindeutig bestimmte Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, für die

$$\overrightarrow{OP} = v = \sum_{k=1}^{n} \lambda_k b_k$$

gilt, nämlich die Koordinaten von v bezüglich $\mathcal{B}$. Damit gilt sofort

$$P = O + \sum_{k=1}^{n} \lambda_k b_k.$$

$\square$

Definition 12.14

Sei $A \neq \emptyset$ ein affiner Raum über einem $\mathbb{K}$-Vektorraum V. Seien zudem $\mathcal{B} = (b_1, \ldots, b_n)$ eine (geordnete) Basis von V und $O \in A$ ein beliebiger Punkt des affinen Raumes. Dann nennen wir $K = (O; b_1, \ldots, b_n)$ *Koordinatensystem* des affinen Raumes A und O den *Koordinatenursprung* des Koordinatensystems.

Ist $\mathcal{B}$ eine ONB von V, so heißt K *kartesisches Koordinatensystem*.

Notation

Wir nennen die eindeutig bestimmten Skalare $\lambda_1, \ldots, \lambda_n$ aus Satz 12.13 die *Koordinaten* von P bezüglich K und schreiben dafür in Anlehnung an die Koordinaten von Vektoren

$$\kappa_K(P) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Beispiel 12.15 (Fortsetzung von Beispiel 12.6). *Das uns bekannte kartesische Koordinatensystem des $\mathbb{R}^3$ ist $K = (O; b_1, b_2, b_3)$ mit*

$$O = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ und } b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

In diesem konkreten Beispiel stimmt jeder Punkt selbst mit seinen Koordinaten bezüglich der Standardbasis und auch mit seinen affinen Koordinaten überein.

Beispiel 12.16. *Seien*

$$O = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 1 \\ -1/2 \end{pmatrix}, \text{ und } b_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2.$$

Dann ist $\mathcal{B} = (b_1, b_2)$ eine Basis des $\mathbb{R}$-Vektorraumes $V = \mathbb{R}^2$ und für $A = \mathbb{R}^2$ ist somit $K = (O; b_1, b_2)$ ein Koordinatensystem des affinen Raumes $A(V)$.

Gegeben seien nun die Punkte $P = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $Q = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ des affinen Raumes A. Berechnen wir deren Koordinaten:

Zu P: *Wir suchen $\lambda_1, \lambda_2 \in \mathbb{R}$, welche das LGS $P = O + \lambda_1 b_1 + \lambda_2 b_2$ erfüllen, d.h.*

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ -1/2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Durch Umstellen erhalten wir das LGS

$$\begin{cases} \lambda_1 + 2\lambda_2 = -\frac{1}{2} \\ -\frac{1}{2}\lambda_1 + \lambda_2 = 1 \end{cases}$$

welches eine eindeutige Lösung besitzt, nämlich $\lambda_1 = -\frac{5}{4}$ und $\lambda_2 = \frac{3}{8}$. Damit lauten die Koordinaten $\kappa_K(P) = \begin{pmatrix} -\frac{5}{4} \\ \frac{3}{8} \end{pmatrix}$.

Zu Q: *Analog zu P bestimmt man die Lösung (μ_1, μ_2) des LGS*

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ -1/2 \end{pmatrix} + \mu_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \Longleftrightarrow \quad \begin{cases} \mu_1 + 2\mu_2 = \frac{3}{2} \\ -\frac{1}{2}\mu_1 + \mu_2 = 0 \end{cases}$$

und erhält so die Koordinaten $\kappa_K(Q) = \begin{pmatrix} \frac{3}{4} \\ \frac{3}{8} \end{pmatrix}$.

Zu beliebigen Punkten: *Bestimmen wir zuletzt die Koordinaten eines beliebigen Punktes $X = \begin{pmatrix} x \\ y \end{pmatrix} \in A$. Das LGS lautet in diesem Fall*

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ -1/2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \Longleftrightarrow \quad \begin{cases} \lambda_1 + 2\lambda_2 = x - \frac{3}{2} \\ -\frac{1}{2}\lambda_1 + \lambda_2 = y - 1 \end{cases}$$

und dessen eindeutige Lösung gibt uns die Koordinaten

$$\kappa_K(X) = \begin{pmatrix} \frac{1}{4} + \frac{1}{2}x - y \\ -\frac{7}{8} + \frac{1}{4}x + \frac{1}{2}y \end{pmatrix}.$$

Dieser Ausdruck sagt uns, wie die Koordinaten eines beliebigen Punktes bezüglich des Standardkoordinatensystems

$$\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

in die Koordinaten bezüglich K überführt (also transformiert) werden. Wir haben also die Vorschrift für die Koordinatentransformation *bestimmt.*

12.4 Affine Unterräume

In Abschnitt 6.6.3 hatten wir zunächst Summen von beliebigen Mengen A und B als

$$\begin{aligned} A + B :&= \{a + b \mid a \in A, b \in B\} \\ &= \{v \in V \mid \exists a \in A, b \in B \colon a + b = v\}. \end{aligned}$$

definiert, wobei wir voraussetzen müssen, dass Elemente aus A und B addiert werden können, was beispielsweise auf Elemente eines Vektorraumes zutrifft. Anschließend haben wir einen affinen Unterraum W eines Vektorraumes V als Teilmenge von V definiert, welche in der Form

$$W = v + U = \{w \in V \mid \exists u \in U : w = v + u\}$$

geschrieben werden kann, wobei $v \in V$ ein beliebiger Vektor und $U \subseteq V$ ein UVR von V ist.

So lange die Menge A, die den affinen Raum darstellt, mit V übereinstimmt, hat diese Definition weiterhin ihre Gültigkeit. Sollte jedoch die Addition $+ : A \times V \to A$ *nicht* mit der Addition innerhalb des Vektorraumes V übereinstimmen, müssen wir mit einer neuen Definition eines affinen Unterraumes arbeiten.

Definition 12.17

Sei A ein affiner Raum über einem $\mathbb{K}$-Vektorraum V mit der Verknüpfung $+ : A \times V \to A$ und sei $P_0 \in A$ ein Punkt dieses Raumes. Ist $U \subseteq V$ ein Untervektorraum von V, so heißt

$$B := P_0 + U := \{P \in A \mid \exists v \in U : P = P_0 + v\} \tag{12.4.1}$$

affiner Unterraum von A.

Bemerkung 12.18. *Nach Eigenschaft (A2) eines affinen Unterraumes stimmt für den Spezialfall $U = V$ der zugehörige affine Unterraum B gerade mit dem affinen Raum A überein. Folglich gilt für den allgemeinen Fall $U \subseteq V$ insbesondere, dass jeder affine Unterraum tatsächlich einen Teilraum des affinen Raumes darstellt, d.h. $B \subseteq A$. Wir könnten* (12.4.1) *also durch*

$$B := \{P \mid \exists v \in U : P = P_0 + v\}$$

ersetzen, da automatisch $B \subseteq A$ gilt.

Lemma 12.19

Sei $B \subseteq A$ ein affiner Unterraum des affinen Raumes $A(V)$. Dann ist B selbst ein affiner Raum über einem UVR U von V.

Beweis. Nach Definition eines affinen Unterraumes existieren ein Punkt $P_0 \in A$ und ein UVR $U \subseteq V$ derart, dass

$$B = \{P \in A \mid \exists v \in U : P = P_0 + v\} \tag{12.4.2}$$

ist. Die Verknüpfung auf B ist die Verknüpfung, die auf A definiert ist, eingeschränkt auf B und U:

$$+ : B \times U \to B$$

Zunächst einmal ist $P + v$ für alle $P \in A$ und $v \in V$ definiert, und somit auch für alle $P \in B \subseteq A$ und $u \in U \subseteq V$. Zudem gilt nach (12.4.2) $P + u \in B$ für $P \in B$ und $u \in U$. Überprüfen wir nun Eigenschaften (A1) bis (A3) aus Definition 12.2:

Die Eigenschaften (A1) und (A3) werden von A auf $B \subseteq A$ und von V auf $U \subseteq V$ vererbt. Einzig Eigenschaft (A2) muss überprüft werden. Seien also $P, Q \in B$ mit

$$P = P_0 + v_P \quad \text{und} \quad Q = P_0 + v_Q,$$

wobei $v_P, v_Q \in U$ sind. Da es innerhalb von V nur genau einen Vektor gibt, der P und Q verbindet, kann es innerhalb von $U \subseteq V$ auch *höchstens einen* solchen Vektor geben. Es bleibt zu zeigen, dass dieser Vektor in der Tat in U liegt.

Bezeichnen wir die Vektoren wieder mit $\overrightarrow{P_0P} = v_P$ und $\overrightarrow{P_0Q} = v_Q$, so gilt mit den Eigenschaften aus Satz 12.7

$$\overrightarrow{PQ} \overset{\text{(iv)}}{=} \overrightarrow{PP_0} + \overrightarrow{P_0Q} \overset{\text{(vi)}}{=} -\overrightarrow{P_0P} + \overrightarrow{P_0Q}.$$

Folglich ist $Q = P + \overrightarrow{PQ} = P + (v_Q - v_P)$, wobei $v_Q - v_P \in U$ ist, da U ein UVR ist. □

Bereits in Kapitel 5 hatten wir überlegt, dass die Darstellung eines affinen Unterraumes als „Punkt plus Vektorraum" nicht eindeutig ist. Formulieren wir für affine Unterräume eines affinen Raumes das Analogon zu Satz 6.61:

Satz 12.20

Sei $A = (A, V, +)$ ein affiner Raum und seien $U, U' \subseteq V$ Untervektorräume von V. Dann gilt:

1. Für jedes $Q \in P_0 + U$ gilt $P_0 + U = Q + U$.
2. Gilt für zwei Punkte $P_0, P_1 \in A$

$$P_0 + U = P_1 + U',$$

so gilt $U = U'$ und $\overrightarrow{P_0P_1} \in U$.

Für Schnitte und Summen von Untervektorräumen haben wir Dimensionsformeln kennen gelernt. Sehen wir uns an, was sich ändert, wenn wir affine Unterräume betrachten.

Lemma 12.21

Seien U_1 und U_2 Untervektorräume eines $\mathbb{K}$-Vektorraumes V und seien $B_1(U_1)$ und $B_2(U_2)$ affine Unterräume eines affinen Raumes A. Ist $B_1 \cap B_2 \neq \emptyset$, so ist $B_1 \cap B_2$ ein affiner Unterraum von A mit zugehörigem Vektorraum $U_1 \cap U_2$ und es gilt

$$\dim(B_1 \cap B_2) = \dim(U_1 \cap U_2). \tag{12.4.3}$$

Beweis. Seien $B_1 = P_1 + U_1$ und $B_2 = P_2 + U_2$ affine Unterräume von $A(V)$. Ist $B_1 \cap B_2 \neq \emptyset$, so existiert ein Punkt $Q \in B_1 \cap B_2$. Nach Satz 12.20 ist

$$B_1 = Q + U_1 \quad \text{und} \quad B_2 = Q + U_2.$$

Damit ist die Schnittmenge $B_1 \cap B_2$ gegeben durch

$$B_1 \cap B_2 = \{P \in A \mid \exists u_1 \in U_1, u_2 \in U_2 \colon P = Q + u_1 = Q + u_2\}.$$

Da für jedes Paar von Punkten $P, Q \in A$ *genau ein* Vektor $v \in V$ existiert mit $P = Q + v$, müssen die beiden Vektoren $u_1 \in U_1$ und $u_2 \in U_2$ gleich sein, d.h.

$$B_1 \cap B_2 = \{P \in A \mid \exists u \in U_1 \cap U_2 \colon P = Q + u\} = Q + (U_1 \cap U_2).$$

(12.4.3) folgt sofort mit der Definition der Dimension eines affinen Raumes. □

Lemma 12.22

Seien U_1 und U_2 Untervektorräume eines $\mathbb{K}$-Vektorraumes V und seien $B_1(U_1)$ und $B_2(U_2)$ affine Unterräume eines affinen Raumes A. Zudem seien $P_1 \in B_1$ und $P_2 \in B_2$ beliebige Punkte. Dann gilt folgende Äquivalenz:

$$B_1 \cap B_2 \neq \emptyset \iff \overrightarrow{P_1P_2} \in U_1 + U_2.$$

Beweis.

$\Longrightarrow$: Ist $B_1 \cap B_2 \neq \emptyset$, so existiert ein Punkt $Q \in B_1 \cap B_2$ und somit liegen die Verbindungsvektoren von P_1 bzw. P_2 zu diesem Punkt in den jeweiligen Unterräumen, d.h.

$$\overrightarrow{P_1Q} \in U_1, \quad \overrightarrow{QP_2} \in U_2.$$

Somit gilt (nach Eigenschaft (iv) von Satz 12.7)

$$\overrightarrow{P_1P_2} = \overrightarrow{P_1Q} + \overrightarrow{QP_2} \in U_1 + U_2.$$

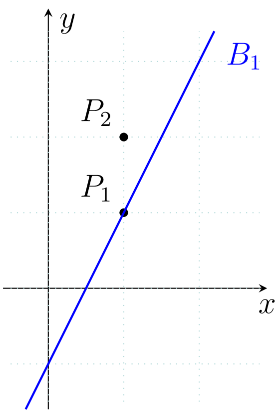
$\Longleftarrow$: Sei $\overrightarrow{P_1P_2} \in U_1 + U_2$, d.h. es existieren $u_1 \in U_1$ und $u_2 \in U_2$ derart, dass $\overrightarrow{P_1P_2} = u_1 + u_2$ ist. Setzen wir $Q := P_1 + u_1 \in B_1$, so gilt

$$\begin{aligned} Q = P_1 + u_1 &= P_1 + ((u_1 + u_2) - u_2) = P_1 + (\overrightarrow{P_1P_2} - u_2) \\ &= (P_1 + \overrightarrow{P_1P_2}) + (-u_2) = P_2 + (-u_2) \in B_2. \end{aligned}$$

Folglich ist $B_1 \cap B_2 \neq \emptyset$.

□

Beispiel 12.23.



Für zwei Parameter $v_1, v_2 \in \mathbb{R}$ seien

$$B_1 = B_1(U_1) = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + a \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} \;\middle|\; a \in \mathbb{R} \right\} \quad \textit{und}$$

$$B_2 = B_2(U_2) = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} + b \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \;\middle|\; b \in \mathbb{R} \right\}$$

zwei affine Unterräume des affinen Raumes $\mathbb{R}^2$. Es ist

$$\dim(B_1 \cap B_2) \leq \min\{\dim(B_1), \dim(B_2)\} = 1,$$

so dass $\dim(B_1 \cap B_2) \in \{-1, 0, 1\}$ ist, d.h. $B_1 \cap B_2$ kann a priori die leere Menge, ein einzelner Punkt oder eine affine Gerade sein.

Wir wollen die Schnittmenge von B_1 und B_2 in Abhängigkeit der Parameter untersuchen. Zunächst halten wir fest, dass es genau dann einen Punkt in $B_1 \cap B_2$ gibt, wenn es Zahlen $a, b \in \mathbb{R}$ gibt, so dass

$$\begin{pmatrix}1\\1\end{pmatrix} + a \cdot \begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix} + b \cdot \begin{pmatrix}v_1\\v_2\end{pmatrix} \tag{12.4.4}$$

gilt. Untersuchen wir nun, wann das der Fall ist, indem wir eine Fallunterscheidung in Bezug auf die Untervektorräume zu B_1 und B_2 durchführen..

$U_1 = U_2$**:** *Die affinen Unterräume B_1 und B_2 lassen sich folgendermaßen als lineare Hüllen schreiben:*

$$U_1 = \mathcal{L}\begin{pmatrix}1\\2\end{pmatrix} \quad \textit{und} \quad U_2 = \mathcal{L}\begin{pmatrix}v_1\\v_2\end{pmatrix}.$$

Diese sind genau dann gleich, wenn $v_2 = 2v_1 \neq 0$ ist. In diesem Fall wird aus (12.4.4)

$$a \cdot \begin{pmatrix}1\\2\end{pmatrix} + b \cdot \begin{pmatrix}-v_1\\-2v_1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix} \iff \left[\begin{array}{cc|c}1 & -v_1 & 0\\2 & -2v_1 & 1\end{array}\right].$$

Dieses LGS hat für keine mögliche Wahl von v_1 eine Lösung, d.h.

$$B_1 \cap B_2 = \emptyset.$$

In diesem Fall sind die beiden affinen Geraden parallel. Insbesondere ist Lemma 12.21 nicht anwendbar und die Formel des Lemmas gilt in der Tat nicht:

$$\dim(B_1 \cap B_2) = \dim(\emptyset) = -1 \neq \dim(U_1 \cap U_2) = \dim(U_1) = 1.$$

$U_1 \neq U_2$**:** *In diesem Fall muss $v_2 \neq 2v_1$ sein, so dass $U_1 \cap U_2 = \{e_{\mathbb{R}^2}\}$ ist. Aus* (12.4.4) *wird dann das eindeutig lösbare LGS*

$$\begin{pmatrix}1 & -v_1\\2 & -v_2\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}.$$

Die Inverse der Matrix ist

$$\frac{1}{2v_1 - v_2}\begin{pmatrix}-v_2 & v_1\\-2 & 1\end{pmatrix},$$

so dass die Lösung des LGS gegeben ist durch

$$\begin{pmatrix}a\\b\end{pmatrix} = \frac{1}{2v_1 - v_2}\begin{pmatrix}-v_2 & v_1\\-2 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{2v_1 - v_2}\begin{pmatrix}v_1\\1\end{pmatrix}.$$

In der Tat ist

$$\begin{pmatrix}1\\1\end{pmatrix} + \frac{v_1}{2v_1 - v_2}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix} + \frac{1}{2v_1 - v_2}\begin{pmatrix}v_1\\v_2\end{pmatrix} \in B_1 \cap B_2$$

der einzige gemeinsame Punkt der beiden affinen Geraden B_1 und B_2.

Die Schnittmengen affiner Unterräume des affinen Raumes $\mathbb{R}^3$ werden in den Übungen genauer untersucht.

Kommen wir nun zu Vereinigungen und Summen affiner Mengen und dem affinen Analogon zur linearen Hülle.

Definition 12.24

Sei $M \subseteq A(V)$ eine Teilmenge eines affinen Raumes A über einem $\mathbb{K}$-Vektorraum V. Der kleinste affine Unterraum von A, der M enthält, wird als *affine Hülle* von M bezeichnet und mit $\mathcal{A}(M)$ abgekürzt.

Bemerkung 12.25. *Sind $B_1(U_1)$ und $B_2(U_2)$ affine Unterräume eines affinen Raumes $A(V)$, so bezeichnen wir die affine Hülle $\mathcal{A}(B_1 \cup B_2)$ auch als* Verbindungsraum *von B_1 und B_2.*

Wie ein solcher Verbindungsraum zweier affiner Unterräume aussieht, sagt folgender Satz:

Satz 12.26

Seien $U_1, U_2 \subseteq V$ Untervektorräume eines $\mathbb{K}$-Vektorraumes und $B_1(U_1)$ und $B_2(U_2)$ affine Unterräume eines affinen Raumes $A(V)$. Seien $P_1 \in B_1$ und $P_2 \in B_2$, so dass

$$B_1 = P_1 + U_1 \quad \text{und} \quad B_2 = P_2 + U_2$$

gilt. Dann ist der Verbindungsraum von B_1 und B_2 gegeben durch

$$\mathcal{A}(B_1 \cup B_2) = P_1 + (\mathcal{L}(\overrightarrow{P_1P_2}) + U_1 + U_2).$$

Beweis. Sei U der $\mathcal{A}(B_1 \cup B_2)$ zu Grunde liegende Untervektorraum von V.

Nach Definition muss $B_1 \cup B_2 \subseteq \mathcal{A}(B_1 \cup B_2)$ sein, d.h. insbesondere ist

$$B_1 = P_1 + U_1 \subseteq \mathcal{A}(B_1 \cup B_2) \quad \text{und} \quad B_2 = P_2 + U_2 \subseteq \mathcal{A}(B_1 \cup B_2),$$

so dass $U_1, U_2 \subseteq U$ sein muss.

Zudem gilt insbesondere $P_1, P_2 \in \mathcal{A}(B_1 \cup B_2)$ und da $\mathcal{A}(B_1 \cup B_2)$ ein affiner Raum ist, muss der ihm zu Grunde liegende Untervektorraum U auch den Verbindungsvektor $\overrightarrow{P_1P_2}$ enthalten. Damit ist insbesondere

$$P_1 + \mathcal{L}(\overrightarrow{P_1P_2}) \subseteq \mathcal{A}(B_1 \cup B_2).$$

Erinnern wir uns daran, dass in Aufgabe 1 von Blatt 10 aus LAAG I für beliebige Teilmengen $M_1, M_2 \subseteq V$ eines Vektorraumes gezeigt wurde, dass

$$\mathcal{L}(M_1 \cup M_2) = \mathcal{L}(M_1) + \mathcal{L}(M_2)$$

ist, so erhalten wir

$$\mathcal{L}(\overrightarrow{P_1P_2}) + U_1 + U_2 \subseteq U$$

und somit

$$P_1 + (\mathcal{L}(\overrightarrow{P_1P_2}) + U_1 + U_2) \subseteq \mathcal{A}(B_1 \cup B_2).$$

Auf der linken Seite steht ein affiner Raum, der $B_1 \cup B_2$ enthält, so dass auf Grund der Minimalität der affinen Hülle sogar die Gleichheit beider Räume gilt. □

Mehrere Möglichkeiten, die Dimension eines Verbindungsraumes zu berechnen, liefert der folgende Satz.

Satz 12.27

Seien $U_1, U_2 \subseteq V$ Untervektorräume eines $\mathbb{K}$-Vektorraumes und $B_1 = P_1 + U_1$ und $B_2 = P_2 + U_2$ affine Unterräume eines affinen Raumes $A(V)$. Dann gilt:

1. Ist $B_1 \cap B_2 \neq \emptyset$, so ist
$$\begin{aligned}\dim(\mathcal{A}(B_1 \cup B_2)) &= \dim(U_1 + U_2) \\ &= \dim(B_1) + \dim(B_2) - \dim(U_1 \cap U_2) \\ &= \dim(B_1) + \dim(B_2) - \dim(B_1 \cap B_2)\end{aligned}$$

2. Ist $B_1 \cap B_2 = \emptyset$, so ist
$$\begin{aligned}\dim(\mathcal{A}(B_1 \cup B_2)) &= \dim(U_1 + U_2) + 1 \\ &= \dim(B_1) + \dim(B_2) - \dim(U_1 \cap U_2) + 1\end{aligned}$$

Beweis von 1. Ist $B_1 \cap B_2 \neq \emptyset$, so gilt nach Lemma 12.22 $\overrightarrow{P_1P_2} \in U_1 + U_2$. Damit wird aus der Formel für den Verbindungsraum aus Satz 12.26

$$\mathcal{A}(B_1 \cup B_2) = P_1 + (\mathcal{L}(\overrightarrow{P_1P_2}) + U_1 + U_2) = P_1 + (U_1 + U_2). \tag{12.4.5}$$

In Aufgabe 4 von Blatt 10 in LAAG I wurde die Dimensionsformel

$$\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2) \tag{12.4.6}$$

gezeigt, welche uns zusammen mit dem wegen $B_1 \cap B_2 \neq \emptyset$ anwendbaren Formel (12.4.3) aus Lemma 12.21 das Ergebnis liefert, nämlich

$$\begin{aligned}\dim(\mathcal{A}(B_1 \cup B_2)) &\overset{(12.4.5)}{=} \dim(U_1 + U_2) \\ &\overset{(12.4.6)}{=} \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2) \\ &\overset{(12.4.3)}{=} \dim(B_1) + \dim(B_2) - \dim(B_1 \cap B_2),\end{aligned}$$

wobei wir mehrfach verwendet haben, dass die Dimension eines affinen Raumes definiert ist als die Dimension des zu Grunde liegenden Vektorraumes. □

12.5 Begriffe zur Beschreibung von Lagebeziehungen affiner Unterräume

In Beispiel 12.23 hatten wir bereits die möglichen Lagebeziehungen einer Geraden zu einer anderen, von der nur ein Punkt feststeht, untersucht. Um ein gemeinsames Vokabular für solche Untersuchungen zu haben, sammeln wir in diesem Abschnitt die Begriffe, die für die Beschreibung von Lagebeziehungen notwendig sind. Zunächst einmal eine Erinnerung und Verallgemeinerung der Begriffe aus Abschnitt 6.6.4:

- Ein 0-dimensionaler affiner Unterraum des $\mathbb{R}^n$ heißt Punkt im $\mathbb{R}^n$.

- Ein 1-dimensionaler affiner Unterraum des $\mathbb{R}^n$ heißt Gerade im $\mathbb{R}^n$.
- Ein 2-dimensionaler affiner Unterraum des $\mathbb{R}^n$ heißt Ebene im $\mathbb{R}^n$.

Parallelität ist ein sehr anschaulicher Begriff, der jedoch auch abstrakt für beliebige affine Unterräume definiert werden kann.

Definition 12.28

1. Zwei affine Unterräume $B_1(U_1), B_2(U_2)$ eines affinen Raumes $A(V)$ heißen *parallel*, in Zeichen $B_1 \parallel B_2$, wenn $U_1 = U_2$ ist.
2. Zwei affine Unterräume $B_1(U_1), B_2(U_2)$ eines affinen Raumes $A(V)$ heißen *schwach parallel*, in Zeichen $B_1 \lhd B_2$, wenn $U_1 \subset U_2$ ist.

Insbesondere können mit dieser Bezeichnung eine Gerade und eine Ebene nie parallel, sondern höchstens schwach parallel sein, während zwei identische Ebenen parallel genannt werden. Zu weiteren üblichen Bezeichnungen kommen wir gleich. Vorher wollen wir jedoch die Frage beantworten, welche Lagebeziehungen zueinander parallele affine Unterräume (gleich welcher Dimension) haben können. Eine Antwort darauf liefert der folgende Satz:

Satz 12.29

Seien $B_1(U_1), B_2(U_2)$ zwei zueinander parallele affine Unterräume eines affinen Raumes $A(V)$. Dann gilt entweder $B_1 = B_2$ oder $B_1 \cap B_2 = \emptyset$.

Beweis. Sind $B_1(U_1)$ und $B_2(U_2)$ parallele affine Unterräume, so gilt nach Definition der Parallelität $U_1 = U_2$. Angenommen $B_1 \cap B_2 \neq \emptyset$, d.h. es existiert ein Punkt $P \in B_1 \cap B_2$. Nach Satz 12.20 gilt dann

$$B_1 = P + U_1 \overset{B_1 \parallel B_2}{=} P + U_2 = B_2.$$

$\square$

Um speziell über die Lagebeziehungen von Geraden und Ebenen sprechen zu können, führen wir zudem folgende Formulierungen ein, welche es ermöglichen, u.a. zwischen den im Satz genannten Möglichkeiten zu unterscheiden.

- Seien G eine affine Gerade und E eine affine Ebene.
  - Ist $G \subseteq E$, so sagen wir, die Gerade liegt in der Ebene.
  - Ist $G \cap E = \{P\}$ für einen Punkt P, so sagen wir, die Gerade schneidet die Ebene in einem Punkt.
  - Ist $G \lhd E$, so sagen wir, die Gerade ist (echt) parallel zur Ebene.
- Seien G_1 und G_2 zwei affine Geraden.
  - Ist $G_1 = G_2$, so sagen wir, die Geraden sind identisch.

- Ist $G_1 \parallel G_2$ und $G_1 \cap G_2 = \emptyset$, so sagen wir, die Geraden sind (echt) parallel zueinander.
- Ist $G_1 \cap G_2 = \{P\}$ für einen Punkt P, so sagen wir, die Geraden schneiden sich in einem Punkt.
- Ist $G_1 \cap G_2 = \emptyset$ und $G_1 \nparallel G_2$, so sagen wir, die Geraden sind zueinander windschief.

Die Begriffe für die Lagebeziehungen zwischen zwei Geraden sind auf zwei Ebenen übertragbar. Alle möglichen Lagebeziehungen einer Geraden und einer Ebene im $\mathbb{R}^3$ zueinander werden in der Präsenzübung besprochen; alle möglichen Lagebeziehungen zweier Ebenen zueinander zu charakterisieren ist eine Hausaufgabe.

12.6 Affine Abbildungen

Basierend auf einem Vektorraum haben wir einen affinen Raum als Punktraum eingeführt, dessen Punkte jeweils durch eindeutig bestimmte Vektoren verbunden sind und auf dem eine Verknüpfung existiert, die es uns ermöglicht, einen Vektor an einen Punkt anzuheften und so einen anderen Punkt zu erreichen.

Mit dieser Idee müssten folgende Abbildungen affin sein:

$$\begin{aligned} f\colon \mathbb{R} \to \mathbb{R}, \quad & f(x) = b + ax, \quad a, b \in \mathbb{R} \\ g\colon \mathbb{R}^n \to \mathbb{R}^m, \quad & g(x) = b + Ax, \quad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m. \end{aligned}$$

Führen wir zunächst den abstrakten Begriff ein und im Anschluss daran werden wir sehen, dass in beiden Fällen in der Tat affine Abbildungen vorliegen, wobei f als Spezialfall von g angesehen werden kann.

Definition 12.30

Seien $A_1(V_1)$ und $A_2(V_2)$ affine Räume über $\mathbb{K}$-Vektorräumen V_1 und V_2. Eine Abbildung $f\colon A_1 \to A_2$ heißt *affine Abbildung*, falls ein Punkt $P \in A_1$ existiert, so dass die Abbildung

$$f_P\colon V_1 \to V_2, \quad f_P(\overrightarrow{PQ}) := \overrightarrow{f(P)f(Q)}, \quad Q \in A_1,$$

linear ist.

Überlegen wir uns zunächst im allgemeinen Fall, wie man (unter Verwendung der Notation aus der Definition) zu einem beliebigen Vektor $v \in V_1$ dessen Bild $f_P(v)$ bestimmt.

i) Zu $v \in V_1$ liegt der Punkt $Q := P + v$ nach Definition der Verknüpfung $+\colon A_1 \times V_1 \to A_1$ in A_1.

ii) Bestimmt man zu den Punkten $P, Q \in A_1$ deren Bilder $f(P), f(Q) \in A_2$, so gibt es für diese beiden Punkte nach (A2) einen eindeutig bestimmten Verbindungsvektor $w := \overrightarrow{f(P)f(Q)}$ mit der Eigenschaft $f(Q) = f(P) + w$.

iii) Schlussendlich setzen wir $f_P(v) := w$.

Bemerkung 12.31. *Zur Berechnung der Verbindungsvektoren ist es wichtig, folgende Äquivalenz im Hinterkopf zu behalten:*

$$f_P(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)} \quad \Longleftrightarrow \quad f(Q) = f(P) + f_P(\overrightarrow{PQ}).$$

Nach unserer Vorüberlegung gibt es zu jeder beliebigen Abbildung $f\colon A_1 \to A_2$ eine solche Abbildung f_P, welche auf die angegebene Weise konstruiert wird. Für diese muss man letztlich überprüfen, ob sie linear ist, um nachzuweisen, dass f affin ist.

Bevor wir zeigen, dass f_P nicht von dem speziell gewählten Punkt P abhängt, sehen wir uns nochmal das Einführungsbeispiel an.

Beispiel 12.32. *Für $m, n \in \mathbb{N}$ sei $g\colon \mathbb{R}^n \to \mathbb{R}^m$ definiert vermöge $g(x) := b + Ax$, wobei $A \in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^m$ seien. Es gilt also $A_1 = V_1 = \mathbb{R}^n$ und $A_2 = V_2 = \mathbb{R}^m$.*

Um zu zeigen, dass g affin ist nach der obigen Definition bestimmen wir zunächst die Abbildung g_p für einen beliebig gewählten Punkt $p \in A_1$. Seien dazu $v \in V_1$ und $q := p + v$. Dann gilt

$$g(p) = b+Ap \quad \textit{und} \quad g(q) = b+Aq = b+A(p+v) = (b+Ap)+Av = g(p)+Av. \tag{12.6.1}$$

Somit setzen wir $g_p(v) := \overrightarrow{g(p)g(q)} = Av$. Die Abbildung $g_p\colon V_1 \to V_2$ mit $g_p(v) = Av$ ist bekanntlich linear und sie erfüllt nach Konstruktion die gewünschte Eigenschaft:

Wählen wir einen Punkt $p \in A_1$ aus und ist $q \in A_1$ mit zugehörigem Verbindungsvektor $v = \overrightarrow{pq}$, so gilt nach (12.6.1)

$$\overrightarrow{g(p)g(q)} = Av = g_p(\overrightarrow{pq}).$$

Folglich ist die angegebene Abbildung g in der Tat affin.

In diesem Beispiel war der Punkt p frei wählbar und f_p hing nicht von diesem Punkt ab. Das nächste Lemma besagt, dass dies immer der Fall ist.

Lemma 12.33

Die Definition einer affinen Abbildung ist unabhängig von der Wahl des Punktes P, d.h. für jeden Vektor $v \in V_1$ ist $f_P(v)$ unabhängig von P.

Beweis. Sei $Q \in A_1$ ein beliebig gewählter Punkt. Zu $v \in V_1$ sei $R := Q + v$, so dass $\overrightarrow{QR} = v$ ist. Da

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$$

ist, haben wir für v die alternative Darstellung

$$v = \overrightarrow{PR} - \overrightarrow{PQ}.$$

Die Definition von f_P und ihre Linearität geben uns $f(v)$, nämlich

$$f_P(v) = f_P(\overrightarrow{PR} - \overrightarrow{PQ}) = f_P(\overrightarrow{PR}) - f_P(\overrightarrow{PQ}) = \overrightarrow{f(P)f(R)} - \overrightarrow{f(P)f(Q)}.$$

Da wiederum

$$\overrightarrow{f(P)f(Q)} + \overrightarrow{f(Q)f(R)} = \overrightarrow{f(P)f(R)}$$

ist, folgt nun

$$f_P(v) = \overrightarrow{f(Q)f(R)},$$

d.h. $f_P(v)$ hängt nicht von P ab. □

12.7 Aufgaben

Aufgabe 12.1. *Sei $\mathcal{P}_n$ der $\mathbb{R}$-Vektorraum der reellen Polynom vom Grad $\leq n$. Sei*

$$*\colon \mathbb{R} \times \mathcal{P}_n \to \mathbb{R}, \quad (x,p) \mapsto p(x).$$

In der VL wurde in Beispiel 12.5 gezeigt, dass Eigenschaft (A1) der Definition eines affinen Raumes von $A = \mathbb{R}$ über dem VR $V = \mathcal{P}_n$ mit dieser Verknüpfung verletzt wird. Zeigen Sie, dass auch die Eigenschaften (A2) und (A3) (siehe Definition 12.2) verletzt werden.

Aufgabe 12.2. *Ziel dieser Aufgabe ist es, zu zeigen, dass Eigenschaft (A1) der Definition bereits durch die anderen Eigenschaften impliziert wird.*

Im Folgenden Sei A ein affiner Raum über einem Vektorraum V mit einer Verknüpfung $+\colon A \times V \to A$, welche Eigenschaften (A2) und (A3) der Definition erfüllt.

a) Sei $P \in A$. Warum existiert ein $v \in V$ derart, dass $P + v = P$ ist?

b) Es gelte $P = P + v$. Zeigen Sie, dass dann $P + e_V = P + v$ ist.

c) Warum folgt hieraus, dass $P + e_V = P$ ist?

Aufgabe 12.3. *Gegeben ist folgende Menge:*

$$A = \left\{ \begin{pmatrix} 7 \\ 5 \\ 3 \end{pmatrix} + s \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \;\middle|\; s, t \in \mathbb{R} \right\}.$$

Finden Sie einen geeigneten Vektorraum V und eine Verknüpfung $+\colon A \times V \to A$, so dass A über V bezüglich $+$ ein affiner Raum ist.

Es ist hier kein ausführlicher Nachweis gefordert, dass es sich um einen affinen Raum handelt.

Aufgabe 12.4. *Betrachten Sie die folgende Menge:*

$$M := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; y = 2x^2 - 12x + 16 \right\}.$$

Bestimmen Sie ein Koordinatensystem $K = (O_K; b_1, b_2)$, für welches die Koordinaten der Punkte aus M die Gleichung $\lambda_2 = \lambda_1^2$ erfüllen.

Hinweis: *Skizzieren Sie M im Standardkoordinatensystem. Wie müssen die Koordinatenachsen verschoben und gestaucht oder gestreckt werden, damit M zu einer Normalparabel im neuen Koordinatensystem wird?*

Aufgabe 12.5. *Gegeben sind*

$$P = \begin{pmatrix} -6 \\ -11 \end{pmatrix}, \; O_K = \begin{pmatrix} 7/2 \\ 1 \end{pmatrix}, \; b_1 = \begin{pmatrix} -1 \\ 4 \end{pmatrix}, \; b_2 = \begin{pmatrix} 3 \\ 8 \end{pmatrix} \in \mathbb{R}^2.$$

Bestimmen Sie die Koordinaten des Punktes P bezüglich $K = (O_K; b_1, b_2)$.

Aufgabe 12.6. *Eine affine Gerade G im $\mathbb{R}^3$ ist ein 1-dimensionaler affiner Unterraum von $A = V = \mathbb{R}^3$; eine affine Ebene E im $\mathbb{R}^3$ ist ein 2-dimensionaler affiner Unterraum. Diese seien gegeben durch*

$$G = P_G + U_G, \quad \textit{und} \quad E = P_E + U_E,$$

wobei U_G und U_E Untervektorräume des $\mathbb{R}^3$ sind.

a) *Welche Lagebeziehungen sind zwischen einer affinen Gerade $G \subseteq \mathbb{R}^3$ und einer affinen Ebene $E \subseteq \mathbb{R}^3$ möglich? Wie sehen in jedem dieser Fälle die Schnittmengen $G \cap E$ aus?*

b) *Finden Sie Kriterien (d.h. Eigenschaften von P_G, P_E, U_G und U_E), die Ihnen sagen, um welchen der Fälle es sich handelt.*

Aufgabe 12.7. *Bestimmen Sie die Schnittmenge $B_1 \cap B_2$ für*

$$B_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + a \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \;\middle|\; a, b \in \mathbb{R} \right\} \quad \textit{und}$$

$$B_2 = \left\{ \begin{pmatrix} 3 \\ 8 \\ 1 \end{pmatrix} + c \cdot \begin{pmatrix} -2 \\ 0 \\ -3 \end{pmatrix} \;\middle|\; c \in \mathbb{R} \right\}.$$

Aufgabe 12.8. *Wir betrachten den affinen Raum $A = \mathbb{R}^3$ mit zugehörigem Vektorraum $V = \mathbb{R}^3$ und der üblichen Vektoraddition als Verknüpfung. Sei $f \colon A \to A$ eine affine Abbildung.*

a) *Begründen Sie, warum f die Form $f(x) = b + g(x)$ besitzt für einen festen Punkt $b \in A$ und eine lineare Abbildung $g \in L(V; V)$.*

b) *Zeigen Sie, dass durch die vier Werte*

$$f\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}, \; f\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}, \; f\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix}, \; f\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}$$

eindeutig eine affine Abbildung definiert wird.

Hinweis: Verwenden Sie zunächst den Nullvektor und a) und erst im Anschluss die restlichen Vektoren.

c) *Bestimmen Sie eine Matrix $A \in \mathbb{R}^{3\times 3}$ und einen Vektor $b \in \mathbb{R}^3$, so dass für alle $x \in \mathbb{R}^3$ gilt:*

$$f(x) = Ax + b.$$

Kapitel 13

Quadratische Formen, Quadriken und die Hauptachsentransformation

Ziele

- □ die Begriffe quadratische Form und Quadrik kennen
- □ quadratische Formen mit Hilfe von Koordinatentransformation ohne gemischte Terme schreiben
- □ Quadriken mit Hilfe der Hauptachsentransformation in ihre Normalform überführen
- □ Besonderheiten beim Diagonalisieren symmetrischer Matrizen kennen

13.1 Quadratische Formen und Quadriken

In Abschnitt 11.1.1, genauer gesagt in Bemerkung 11.6, hatten wir festgestellt, dass für eine symmetrische Matrix $A \in \mathbb{K}^{n\times n}$ die Abbildung $f\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, welche gegeben ist durch $f(v,w) := v^T Aw$, eine Bilinearform darstellt.

Jetzt werden wir uns auf $\mathbb{K} = \mathbb{R}$ beschränken und an Stelle von zwei verschiedenen Vektoren v und w zweimal den gleichen Vektor einsetzen. Dann erhalten wir die Abbildung $g\colon \mathbb{R}^n \to \mathbb{R}$ mit $g(v) := v^T Av$. Anhand eines Beispiels wollen wir uns ansehen, warum wir einen Ausdruck dieser Form auch als *quadratische Form* bezeichnen.

Beispiel 13.1. *Sei* $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{R}^{2\times 2}$ *eine quadratische Matrix und sei und* $g\colon \mathbb{R}^2 \to \mathbb{R}$ *gegeben durch* $g(v) := v^T Av$. *Dann gilt*

$$\begin{aligned} g\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} ax + by \\ bx + cy \end{pmatrix} = x\cdot(ax+by) + y\cdot(bx+cy) \\ &= ax^2 + 2bxy + cy^2. \end{aligned}$$

Es entsteht ein quadratisches Polynom in zwei Variablen.

Definition 13.2: Quadratische Form

Für $n \in \mathbb{N}$ sei $A = (\alpha_{k,j}) \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Sei zudem $x = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}^T \in \mathbb{R}^n$. Eine Abbildung $Q_A\colon \mathbb{R}^n \to \mathbb{R}$ von der Form

$$Q_A(x_1, \dots, x_n) := x^T A x = \sum_{k,j=1}^{n} \alpha_{k,j} x_j x_k$$

heißt *quadratische Form* zur Matrix A.

Quadratische Formen sind also spezielle quadratische Polynome in n Variablen.

Bemerkung 13.3. *Allgemeine quadratische Funktionen dürfen neben den Termen einer quadratischen Form auch lineare und konstante Terme enthalten, d.h. eine allgemeine quadratische Funktion ist von der Form*

$$q\colon \mathbb{R}^n \to \mathbb{R}, \quad q(x) = x^T A x + b^T x + c,$$

wobei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix, $b \in \mathbb{R}^n$ ein Vektor und $c \in \mathbb{R}$ eine Zahl ist. Dieses Polynom können wir jedoch kompakter als quadratische Form schreiben, denn für

$$\widetilde{A} := \begin{pmatrix} A & \frac{1}{2}b \\ \frac{1}{2}b^T & c \end{pmatrix} \quad \textit{und} \quad \widetilde{x} := \begin{pmatrix} x \\ 1 \end{pmatrix}$$

gilt

$$\widetilde{x}^T \widetilde{A} \widetilde{x} = \begin{pmatrix} x^T & 1 \end{pmatrix} \begin{pmatrix} Ax + \frac{1}{2}b \\ \frac{1}{2}b^T x + c \end{pmatrix} = x^T A x + \frac{1}{2} x^T b + \frac{1}{2} b^T x + c = q(x)$$

d.h. q kann als quadratische Form aufgefasst werden, wenn die Matrix und der Vektor x entsprechend erweitert werden.

Uns interessieren gerade das Aussehen und die Eigenschaften der Menge der Nullstellen eines quadratischen Polynoms.

Definition 13.4: Quadrik

Eine *Quadrik* ist eine Teilmenge des $\mathbb{R}^n$ ($n \in \mathbb{N}$) von der Form

$$Q = \{x \in \mathbb{R}^n \mid q(x) = 0\},$$

wobei

$$q(x_1, \dots, x_n) = \sum_{k,j=1}^{n} \alpha_{k,j} x_k x_j + \sum_{i=1}^{n} \beta_i x_i + c$$

ist. Für eine symmetrische Matrix $A = (\alpha_{k,j}) \in \mathbb{R}^{n\times n}$, einen Vektor $b = (\beta_i) \in \mathbb{R}^n$ und ein Skalar $c \in \mathbb{R}$ ist somit

$$q(x) = x^T A x + b^T x + c.$$

Beispiel 13.5. *Für $n = 2$ und*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \textit{und} \quad c = -1$$

erhalten wir die Quadrik

$$Q = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; x^T A x + b^T x + c = 0 \right\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 = 1 \right\},$$

d.h. die Quadrik ist ein Einheitskreis.

Bemerkung 13.6. *Im Spezialfall $n = 2$ bezeichnen wir Quadriken auch als* Kegelschnitte *und im Spezialfall $n = 3$ als* quadratische Fläche. *Alternativ spricht man auch von* Kurven *oder* Flächen *(oder für ein allgemeines n von* Hyperflächen*)* zweiter Ordnung.

Das Ziel dieses Kapitels ist es, Quadriken gewissen Typen zuordnen zu können, die wiederum bestimmte Eigenschaften besitzen. Als Vorarbeit dafür müssen wir uns ansehen, wie symmetrische Matrizen diagonalisiert werden können und anschließend die affine Koordinatentransformation bemühen.

13.2 Symmetrische Matrizen

13.2.1 Diagonalisierbarkeit symmetrischer Matrizen

Für den Spezialfalls $n = 2$ hatten wir bereits in Aufgabe 4 von Übungsblatt 5 gesehen, dass symmetrische Matrizen des $\mathbb{R}^{2\times 2}$ immer diagonalisierbar sind, da sie ausschließlich reelle Eigenwerte besitzen und entweder bereits selbst diagonal sind oder zwei verschiedene reelle Eigenwerte haben müssen, was nach Korollar 10.20 hinreichend für die Diaognalisierbarkeit ist.

Diese Ergebnisse werden wir jetzt u.a. mit Hilfe der Kenntnisse über euklidische Vektorräume auf symmetrische Matrizen des $\mathbb{R}^{n\times n}$ für beliebige $n \in \mathbb{N}$ erweitern.

Satz 13.7: Spektralsatz

Für ein $n \in \mathbb{N}$ sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann gilt:

1. Alle Eigenwerte von A sind reell.
2. Eigenvektoren zu verschiedenen Eigenwerten von A sind orthogonal zueinander bezüglich des kanonischen Skalarproduktes des $\mathbb{R}^n$.
3. Es existiert eine orthogonale Matrix $S \in \mathcal{O}(n)$, welche A diagonalisiert, d.h. es gilt
$$S^T A S = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$
wenn $\lambda_1, \ldots, \lambda_n$ die entsprechend ihrer algebraischen Vielfachheit ggf. mehrfach gezählten Eigenwerte von A sind.

Bemerkung 13.8. *Eigenschaft 2 impliziert, dass wenn eine symmetrische Matrix diagonalisierbar ist, dies auch mit einer orthogonalen Matrix möglich sein muss. Dazu wählt man sich eine wegen der Diagonalisierbarkeit existierende Basis $\mathcal{B}$ aus Eigenvektoren von A. Diese ist zusammengesetzt aus Basen $\mathcal{B}_1, \ldots, \mathcal{B}_m$ der $m \leq n$ verschiedenen Eigenräume von A. Mit dem Gram-Schmidt'schen Orthonormalisierungsverfahren lassen sich aus $\mathcal{B}_1, \ldots, \mathcal{B}_m$ ONBs $\mathcal{C}_1, \ldots, \mathcal{C}_m$ der Eigenräume konstruieren, welche auf Grund von Eigenschaft 2 orthogonal zueinander sind. Folglich ist $\mathcal{C}_1 \cup \ldots \cup \mathcal{C}_m$ eine ONB des $\mathbb{R}^n$, welche aus Eigenvektoren von A besteht. Die zugehörige Matrix, deren Spalten aus diesen Basisvektoren besteht, ist somit orthogonal und hat die gewünschten Eigenschaften.*

Das Problem bei dieser Überlegung ist jedoch, dass wir die Diagonalisierbarkeit selbst erst zeigen müssen. Diese folgt noch nicht aus 1., da mehrfache reelle Eigenwerte dazu führen können, dass eine Matrix nicht diagonalisierbar ist, sofern die algebraische und geometrische Vielfachheit des Eigenwertes nicht übereinstimmen.

Beweis.

1. Sei $v \in \mathbb{C}^n$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$ von A. Das Ziel ist es, zu zeigen, dass $\lambda \in \mathbb{R}$ ist. Da A selbst reell ist, wissen wir, dass wir zu $\lambda \in \mathbb{R}$ einen Eigenvektor von A im $\mathbb{R}^n$ finden. So lange wir dies jedoch noch nicht gezeigt haben, verwenden wir das kanonische Skalarprodukt des $\mathbb{C}^n$, welches sich für reelle Vektoren jedoch nicht von dem kanonischen Skalarprodukt des $\mathbb{R}^n$ unterscheidet.

$$\begin{aligned}\overline{\lambda}\langle v, v\rangle &= \overline{\lambda} v^T \overline{v} = v^T \overline{\lambda v} = v^T \overline{Av} = v^T \overline{A}\,\overline{v} \overset{A \in \mathbb{R}^{n\times n}}{=} v^T A \overline{v} \overset{A^T = A}{=} v^T A^T \overline{v} \\ &= (Av)^T \overline{v} = (\lambda v)^T \overline{v} = \lambda \langle v, v\rangle.\end{aligned}$$

   Da v ein Eigenvektor ist, kann es nicht der Nullvektor sein, so dass wegen der positiven Definitheit des Skalarproduktes $\langle v, v\rangle > 0$ sein muss. Folglich muss $\overline{\lambda} = \lambda$ sein, was $\lambda \in \mathbb{R}$ impliziert.

2. Das ist eine Übungsaufgabe.

3. Wir führen eine vollständige Induktion über $n \in \mathbb{N}$ durch.

   Induktionsanfang: Für $n = 1$ ist nichts zu zeigen.

   Induktionsannahme: Für jede Matrix $A' \in \mathbb{R}^{(n-1)\times(n-1)}$ existiert eine orthogonale Matrix $S' \in \mathcal{O}(n-1)$, für welche $(S')^{-1}A'S' =: D'$ eine Diagonalmatrix ist. Da S' orthogonal ist, gilt $(S')^{-1} = (S')^T$, d.h. es gilt $S'D'(S')^T = A'$.

   Induktionsschritt: Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch, d.h. $A^T = A$. Da alle Eigenwerte von A reell sind, zerfällt das charakteristische Polynom p_A in $\mathbb{R}[x]$ in Linearfaktoren. Es gibt also mindestens einen Eigenwert $\lambda_1 \in \mathbb{R}$ von A und sei $v_1 \in \mathbb{R}^n$ ein zugehöriger Eigenvektor. Ohne Einschränkung sei dieser normiert, d.h. $\|v_1\| = 1$ für die durch das kanonische euklidische Skalarprodukt induzierte Norm $\|\cdot\|$.

   Mit Hilfe des Austauschsatzes 6.33 kann v_1 zu einer Basis $\mathcal{B} = (v_1, w_2, \ldots, w_n)$ des $\mathbb{R}^n$ ergänzt werden. Ohne Einschränkung[1] nehmen wir an, dass $\mathcal{B}$ bezüglich des kanonischen Skalarproduktes eine ONB ist.

[1] Sollte $\mathcal{B}$ keine ONB sein, können wir mit Hilfe des Gram-Schmidt'schen Orthonormalisierungsverfahrens eine konstruieren und arbeiten mit dieser weiter. Es ist an dieser Stelle wichtig, zu bemerken, dass wenn der erste Vektor einer Basis normiert ist, dieser durch das Gram-Schmidt-Verfahren nicht verändert wird.

Sei S_1 die Matrix, deren Spalten aus den Vektoren von $\mathcal{B}$ bestehen. Nach Lemma 11.87 ist $S_1 \in \mathcal{O}(n)$ orthogonal, da ihre Spalten nach Konstruktion eine ONB bezüglich des kanonischen Skalarproduktes bilden. Folglich ist $S_1^{-1} = S_1^T$. Für die Matrix $A_1 := S_1^{-1}AS_1$ gilt damit

$$A_1^T = (S_1^{-1}AS_1)^T = S_1^T A^T (S_1^{-1})^T = S_1^{-1} A^T (S_1^T)^T \overset{A^T=A}{=} S_1^{-1}AS_1 = A_1,$$

d.h. A_1 ist symmetrisch. Da die erste Spalte von S_1, v_1, Eigenvektor von A zum Eigenwert λ_1 ist, hat A_1 die Form

$$A_1 = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix},$$

wobei $A_2 \in \mathbb{R}^{n-1 \times n-1}$ symmetrisch ist, da A_1 symmetrisch ist.

Nach Induktionsannahme existiert eine Matrix $S_2 \in \mathcal{O}(n-1)$ derart, dass

$$S_2^{-1}A_2S_2 = S_2^T A_2 S_2 = \operatorname{diag}(\mu_2, \ldots, \mu_n)$$

ist mit den $n-2$ Eigenwerten $\mu_2, \ldots, \mu_n \in \mathbb{R}$ von A_2. Sei

$$S := S_1 \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix}.$$

Da beide Matrizen in der Gruppe $\mathcal{O}(n)$ sind, ist auch deren Produkt eine orthogonale Matrix. Die Inversen aller beteiligten Matrizen (also S, S_1 und S_2) sind also deren Transponierte. Damit können wir nachweisen, dass S die Matrix A diagonalisiert:

$$\begin{aligned}
S^{-1}AS &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix}^{-1} \cdot S_1^{-1}AS_1 \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix}^{T} \cdot A_1 \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2^T & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2^{-1}A_2S_2 & \\ 0 & & & \end{pmatrix}
\end{aligned}$$

$$= \mathrm{diag}(\lambda_1, \mu_2, \dots, \mu_n).$$

Da A somit ähnlich ist zu $\mathrm{diag}(\lambda_1, \mu_2, \dots, \mu_n)$, müssen $\mu_2, \dots, \mu_n$ auch Eigenwerte von A sein, d.h. es ist auf Grund der Orthogonalität von S

$$S^{-1}AS = S^T AS = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

□

Nachdem wir festgestellt haben, dass alle Eigenwerte symmetrischer Matrizen des $\mathbb{R}^{n\times n}$ reelle Zahlen sind, können wir diese daraufhin untersuchen, ob sie positive und/oder negative Eigenwerte besitzen. Das führt uns zu einer Anwendung des Wissens über symmetrische Matrizen.

13.2.2 Definitheit symmetrischer Matrizen

In Beispiel 11.9 hatten wir bereits gesehen, dass für eine symmetrische Matrix $A \in \mathbb{R}^{2\times 2}$ die symmetrische Bilinearform $f(x, y) := v^T Aw$ genau dann positiv definit ist, wenn A ausschließlich positive Eigenwerte besitzt. Mit dem obigen Resultat über die Eigenwerte symmetrischer Matrizen im Hinterkopf können wir diese Aussage auf beliebig große symmetrische Matrizen verallgemeinern. Vorher erinnern wir jedoch noch einmal an die Definition der Definitheit einer symmetrischen Bilinearform und ergänzen sie noch um dem Begriff der Semidefinitheit:

Man nennt eine symmetrische Bilinearform $f\colon V \times V \to \mathbb{R}$

1. *positiv definit*, falls $f(v, v) > 0$ ist für alle $v \in V \setminus \{e_V\}$;
2. *positiv semidefinit*, falls $f(v, v) \geq 0$ ist für alle $v \in V$;
3. *negativ definit*, falls $f(v, v) < 0$ ist für alle $v \in V \setminus \{e_V\}$;
4. *negativ semidefinit*, falls $f(v, v) \leq 0$ ist für alle $v \in V$;
5. *indefinit*, falls $v, w \in V \setminus \{e_V\}$ existieren, für die $f(v, v) > 0$ und $f(w, w) < 0$ gilt.

Bemerkung 13.9. *Die Definitheit einer Bilinearform sagt uns, ob die zugehörige quadratische Form $Q_A(v) := f(v, v) = v^T Av$ für alle Vektoren ein einheitliches Vorzeichen hat.*

Die Definitheit einer Matrix wird geknüpft an die Definitheit der zugehörigen Bilinearform.

Definition 13.10

Sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix und $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ gegeben durch $f(v, w) := v^T Aw$. A heißt genau dann *positiv (semi-)definit/ negativ (semi-)definit /indefinit*, wenn f diese Eigenschaft besitzt.

Die Definitheit können wir mit Hilfe der Eigenwerte der Matrix bestimmen.

Satz 13.11

Für $n \in \mathbb{N}$ sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Zudem sei $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ die Bilinearform, welche gegeben ist durch $f(v, w) := v^T Aw$. Dann ist f genau dann

1. positiv definit, wenn alle Eigenwerte von A positiv sind;
2. positiv semidefinit, wenn alle Eigenwerte von A nicht negativ sind;
3. negativ definit, wenn alle Eigenwerte von A negativ sind;
4. negativ semidefinit, wenn alle Eigenwerte von A nicht positiv sind;
5. indefinit, wenn A sowohl positive als auch negative Eigenwerte hat.

Beweis. Wir zeigen die Aussage für die positive Definitheit. Sei dazu $\lambda \in \mathbb{R}$ ein beliebiger Eigenwert von A mit zugehörigem Eigenvektor $v \in \mathbb{R}^n$. Bezeichnen wir mit $\langle x, y\rangle := x^T y$ das kanonische Skalarprodukt des $\mathbb{R}^n$, so gilt

$$f(v, v) = v^T Av = v^T \cdot \lambda v = \lambda\langle v, v\rangle > 0 \quad\iff\quad \lambda > 0, \tag{13.2.1}$$

da v als Eigenvektor nicht der Nullvektor sein kann und somit $\langle v, v\rangle > 0$ ist.

„$\Longrightarrow$“: Ist f positiv definit, so gilt $f(x, x) > 0$ für alle $x \in \mathbb{R}^n \setminus \{e_{\mathbb{R}^n}\}$, also insbesondere für alle Eigenvektoren von A. Nach (13.2.1) impliziert dies, dass alle Eigenwerte positiv sind.

„$\Longleftarrow$“: Sind alle Eigenwerte von A positiv, so impliziert (13.2.1), dass $f(v, v) > 0$ ist für alle Eigenvektoren von A. Da A als symmetrische reelle Matrix diagonalisierbar ist mit einer orthogonalen Matrix, gibt es eine ONB des $\mathbb{R}^n$, nennen wir sie $\mathcal{B} = (b_1, \ldots, b_n)$, welche aus Eigenvektoren von A besteht. Sei nun $v \in \mathbb{R}^n$ ein beliebiger Vektor mit (eindeutiger) Darstellung

$$v = \sum_{k=1}^{n} \mu_k b_k$$

bezüglich $\mathcal{B}$, wobei genau dann $\mu_1 = \ldots = \mu_n = 0$ ist, wenn $v = e_{\mathbb{R}^n}$ ist. Auf Grund der Bilinearität von f und deren Zusammenhang zu A gilt dann

$$\begin{aligned}
f(v, v) = f\left(\sum_{k=1}^{n} \mu_k b_k, \sum_{j=1}^{n} \mu_k b_k\right) &= \sum_{k,j=1}^{n} \mu_k\mu_j f(b_k, b_j) \\
&= \sum_{k,j=1}^{n} \mu_k\mu_j b_k^T Ab_j = \sum_{k,j=1}^{n} \mu_k\mu_j\lambda_j\langle b_k, b_j\rangle \overset{\mathcal{B}\text{ ONB}}{=} \sum_{k,j=1}^{n} \mu_k\mu_j\lambda_j\delta_{k,j} \\
&= \sum_{k=1}^{n} \mu_k^2 \underbrace{\lambda_k}_{>0 \text{ nach Vor.}} .
\end{aligned}$$

Für $v \neq e_{\mathbb{R}^n}$ ist mindestens ein Koeffizient $\mu_k \neq 0$ und somit ist $f(v, v) > 0$ für alle $v \neq e_{\mathbb{R}^n}$.

$\square$

Eine konkrete Anwendung findet die Untersuchung der Definitheit von symmetrischen Matrizen in der Analysis[2] bei mehrdimensionalen Optimierungsproblemen. Da die Analysis keine Voraussetzung für diese Lehrveranstaltung ist, werden wir auf die Details jedoch nicht eingehen.

13.3 Hauptachsentransformation

In diesem Abschnitt kommen wir zurück zu den Quadriken. Speziell wollen wir uns mit Kegelschnitten befassen, d.h. mit Mengen der Form

$$Q = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; q(x_1, x_2) := ax_1^2 + bx_1x_2 + cx_2^2 + dx_1 + ex_2 + f = 0 \right\}. \tag{13.3.1}$$

Mit Hilfe von Drehungen, Spiegelungen und Verschiebungen wollen wir den Typ der Quadrik bestimmen.

Normalformen von 2D-Quadriken

Ellipse: $\alpha^2 y_1^2 + \beta^2 y_2^2 = 1$

Hyperbel: $\alpha^2 y_1^2 - \beta^2 y_2^2 = 1$

parallele Geraden: $\alpha^2 y_1^2 = 1$

Punkt: $\alpha^2 y_1^2 + \beta^2 y_2^2 = 0$

sich schneidende Geraden: $\alpha^2 y_1^2 - \beta^2 y_2^2 = 0$

Gerade: $\alpha^2 y_1^2 = 0$

Parabel: $\alpha^2 y_1^2 - 2y_2 = 0$

Um ein geeignetes Koordinatensystem zu finden, bezüglich dessen eine gegebene Quadrik eine dieser Formen aufweist, gehen wir wie folgt vor:

1. Schritt: Wir schreiben (13.3.1) in Matrixschreibweise als

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} d \\ e \end{pmatrix} + f = 0.$$

2. Schritt: Die symmetrische Matrix kann mit Hilfe einer orthogonalen Matrix diagonalisiert werden, so dass für die neuen Koordinaten von Q folgende Gleichung gilt:

$$A(x_1')^2 + B(x_2')^2 + Cx_1' + Dx_2' + E = 0.$$

Dieser Schritt entspricht einer Drehung oder Spiegelung von Q.

[2] Analog zum hinreichenden Kriterium für Maxima und Minimal von zweimal stetig differenzierbaren Funktionen, welches nach dem Vorzeichen der zweiten Ableitung fragt, gilt für Funktionen $f\colon \mathbb{R}^n \to \mathbb{R}$, dass nach der Überprüfung des notwendigen Optimalitätskriteriums als hinreichendes Kriterium die Definitheit der Hesse-Matrix der Funktion überprüft werden muss, in welcher die zweiten partiellen Ableitungen der Funktion stehen.

3. Schritt: Mit Hilfe quadratischer Ergänzung wird ein geeigneter Koordinatenursprung gefunden, so dass nach einer erneuten Koordinatentransformation (in diesem Fall eine Verschiebung) die linearen Terme nach Möglichkeit verschwinden und die Normalform erreicht ist.

Sehen wir uns ein Beispiel dazu an:

Beispiel 13.12. *Betrachten wir die Quadrik, welche gegeben ist durch*

$$Q = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, q(x_1,x_2) := 3x_1^2 - 6x_1x_2 - 5x_2^2 + \frac{48}{\sqrt{10}}x_1 + \frac{64}{\sqrt{10}}x_2 - 19 = 0 \right\}.$$

Gehen wir die einzelnen Schritte durch.

1. *In Matrixform ist*

$$Q = \left\{ x \in \mathbb{R}^2 \,\middle|\, x^T \cdot \begin{pmatrix} 3 & -3 \\ -3 & -5 \end{pmatrix} \cdot x + x^T \cdot \frac{1}{\sqrt{10}} \begin{pmatrix} 48 \\ 64 \end{pmatrix} - 19 = 0 \right\}.$$

Nennen wir ab jetzt

$$A := \begin{pmatrix} 3 & -3 \\ -3 & -5 \end{pmatrix}, \quad b := \frac{1}{\sqrt{10}} \begin{pmatrix} 48 \\ 64 \end{pmatrix} \quad \textit{und} \quad c := -19.$$

2. *Um die Matrix A zu diagonalisieren, benötigen wir zunächst eine ONB aus Eigenvektoren, wofür wir wiederum die Eigenwerte bestimmen müssen.*

Bestimmung der Eigenwerte: *Das charakteristische Polynom von A lautet*

$$\begin{aligned} p_A(t) &= \det(A - tE_2) = (3-t)(-5-t) - 9 = t^2 + 2t - 24 \\ &= (t-4)(t+6) \end{aligned}$$

und dessen Nullstellen sind $\lambda_1 = 4$ und $\lambda_2 = -6$.

Bestimmung der Eigenräume: *Für $\lambda_1 = 4$ ist*

$$\begin{aligned} \operatorname{Eig}(A;4) &= \left\{ x \in \mathbb{R}^2 \,\middle|\, (A - 4E_2)x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ x \in \mathbb{R}^2 \,\middle|\, \begin{pmatrix} -1 & -3 \\ -3 & -9 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \mathcal{L}\begin{pmatrix} 3 \\ -1 \end{pmatrix}. \end{aligned}$$

Analog ist für $\lambda_2 = -6$

$$\operatorname{Eig}(A;-6) = \left\{ x \in \mathbb{R}^2 \,\middle|\, \begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \mathcal{L}\begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Normierung: *Wie in Satz 13.7 vorhergesagt sind die Eigenräume orthogonal zueinander. Wir müssen also nur aus jedem Eigenraum einen normierten Vektor wählen:*

$$\left\| \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\| = \sqrt{3^2 + (-1)^2} = \sqrt{10} = \left\| \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\|,$$

folglich wählen wir als ONB

$$\mathcal{B} = \left(\frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ -3 \end{pmatrix} \right).$$

Diagonalisierung: *Die orthogonale Matrix*

$$S := \frac{1}{\sqrt{10}} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} \quad \textit{mit} \quad S^{-1} = S^T = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}$$

diagonalisiert A, denn

$$\begin{aligned} S^{-1}AS &= \frac{1}{10} \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & -3 \\ -3 & -5 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} \\ &= \frac{1}{10} \begin{pmatrix} 12 & -4 \\ -6 & -18 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} \\ &= \frac{1}{10} \begin{pmatrix} 40 & 0 \\ 0 & -60 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -6 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} =: D. \end{aligned}$$

Lineare Koordinatentransformation: *Ersetzen wir x durch x' mit $x = Sx'$ bzw. $x' = S^T x$, so erhalten wir*

$$\begin{aligned} & x^T A x + x^T b + c = 0 \\ \iff & (Sx')^T A (Sx') + (Sx')^T b + c = 0 \\ \iff & (x')^T S^T A S x' + (x')^T S^T b + c = 0 \\ \iff & (x')^T D x' + (x')^T S^T b + c = 0 \\ \iff & 4(x_1')^2 - 6(x_2')^2 + \begin{pmatrix} x_1' & x_2' \end{pmatrix} \frac{1}{10} \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 48 \\ 64 \end{pmatrix} - 19 = 0 \\ \iff & 4(x_1')^2 - 6(x_2')^2 + 8x_1' + 24x_2' - 19 = 0. \end{aligned}$$

3. *Mit quadratischer Ergänzung erhalten wir*

$$\begin{aligned} & 4(x_1')^2 - 6(x_2')^2 + 8x_1' + 24x_2' - 19 \\ =\ & 4((x_1')^2 + 2x_1' + 1 - 1) - 6((x_2')^2 - 4x_2' + 4 - 4) - 19 \\ =\ & 4(x_1' + 1)^2 - 6(x_2' - 2)^2 + 1 = 0 \end{aligned}$$

Setzen wir $y_2 := x_1' + 1$ und $y_1 := x_2' - 2$, so erhalten wir die gewünschte Form:

$$6y_1^2 - 4y_2^2 = 1.$$

Es handelt sich bei der gegebenen Quadrik Q also um eine Hyperbel.

Bemerkung 13.13. *Die Wahl der ONB $\mathcal{B}$ lässt einen gewissen Spielraum, denn die Multiplikation eines oder beider Vektoren mit -1 ergibt eine neue ONB und welcher Vektor zuerst kommt ist ebenso frei wählbar. Daher ist auch S nicht eindeutig bestimmt, genauso wenig wie das Gesamtergebnis nach dem 2. Schritt. Erst die Normierung am Ende des 3. Schrittes gibt uns einen eindeutigen Typ.*

Auf einen Beweis der Vollständigkeit der Liste aller Normalformen und der gezeigten Methode und auf eine Verallgemeinerung auf beliebige Dimensionen verzichten wir.

13.4 Aufgaben

Aufgabe 13.1. *Gegeben sei die Matrizen*

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & -4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}.$$

Bestimmen Sie die zugehörigen quadratischen Formen $q_k(x) = x^T A_k x$.

Aufgabe 13.2. *Finden Sie je eine symmetrische Matrix* A*, so dass* $q(x) = x^T Ax$ *ist für* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

a) $q(x_1, x_2) = 7x_1^2 - 4x_1x_2 + 6x_2^2$

b) $q(x_1, x_2) = -3x_1^2 + 7x_1x_2$

Aufgabe 13.3. *Gegeben sei die Quadrik*

$$Q = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, x^2 - 2\sqrt{3}xy + 3y^2 - 4\sqrt{3}x - 4y + 24 = 0 \right\}.$$

a) Bestimmen Sie eine symmetrische Matrix $A \in \mathbb{R}^{2\times 2}$ *einen Vektor* $b \in \mathbb{R}^2$ *und eine Zahl* $c \in \mathbb{R}$*, so dass* Q *auf folgende Weise geschrieben werden kann:*

$$Q = \left\{ v \in \mathbb{R}^2 \,\middle|\, v^T Av + v^T b + c = 0 \right\}.$$

b) Bestimmen Sie die Eigenwerte und die zugehörigen Eigenräume von A*.*

c) Bestimmen Sie eine ONB des $\mathbb{R}^2$*, welche aus Eigenvektoren von* A *besteht.*

d) Diagonalisieren Sie A*, d.h. finden Sie* $S \in \mathcal{O}(2)$*, so dass* $S^T AS = D$ *eine Diagonalmatrix ist.*

e) Ersetzen Sie $\begin{pmatrix} x \\ y \end{pmatrix}$ *durch* $S\begin{pmatrix} v \\ w \end{pmatrix}$*. Welche Darstellung bekommt* Q *damit?*

f) Kann man mit einer quadratischen Ergänzung die linearen Terme verschwinden lassen?

Kapitel 14

Ergänzungen und Anwendungen

Ziele

- □ bekannte Begriffe und Methoden in neuen Problemen anwenden
- □ Anwendungsbereiche kennen lernen, die über LAAG hinausgehen

14.1 Determinante und Volumenänderung

Die Determinante war historisch gesehen zunächst ein Hilfsmittel für die Lösung linearer Gleichungssysteme – sie zeigt uns an, ob ein LGS aus n Gleichungen mit n Unbekannten eine eindeutige Lösung besitzt (d.h. ob die zugehörige Matrix invertierbar ist) und sie erlaubt uns, beispielsweise mit Hilfe der Cramer'schen Regel, die Lösung eines LGS zu bestimmen.

In diesem Abschnitt wollen wir einen anderen Aspekt der Determinante beleuchten. Dazu sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix und sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ die lineare Abbildung, welche gegeben ist durch die Vorschrift $f(x) = A \cdot x$. Ist $M \subseteq \mathbb{R}^n$ eine Menge, deren Volumen wir kennen, so ist die Frage, welches Volumen das Bild dieser Menge, $f(M)$, besitzt.

Beispiel 14.1. *Sei $M := [0,1]^3 \subseteq \mathbb{R}^3$ ein Einheitswürfel. Man kann M auch in folgender Form schreiben:*

$$M = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x_1, x_2, x_3 \in [0,1] \right\}.$$

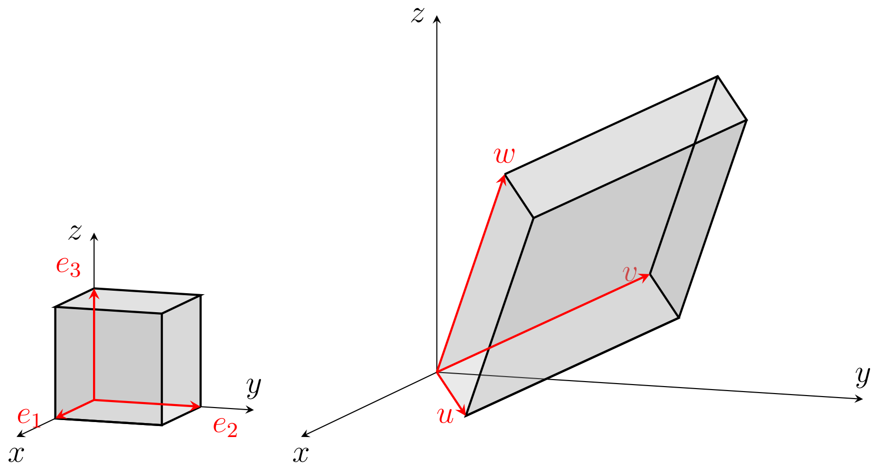
Seien $u, v, w \in \mathbb{R}^3$ drei Vektoren, deren Spalten die Matrix $A \in \mathbb{R}^{3\times 3}$ ergeben, d.h.

$$A = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

Das Volumen eines Würfels mit Kantenlänge 1 LE (Längeneinheit) ist bekanntlich 1 VE (Volumeneinheit). Sei nun $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch $f(x) = Ax$. Wir wollen das Volumen von $f(M)$ bestimmen. Dazu überlegen wir uns zunächst, wie $f(M)$ aussieht.

$$\begin{aligned} f(M) &= \left\{ y \in \mathbb{R}^3 \;\middle|\; \exists x \in M \colon f(x) = Ax = y \right\} \\ &= \left\{ y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \exists x \in M \colon x_1 u + x_2 v + x_3 w = y \right\} \\ &= \left\{ x_1 u + x_2 v + x_3 w \;\middle|\; x_1, x_2, x_3 \in [0,1] \right\}, \end{aligned}$$

d.h. $f(M)$ ist die Menge aller Linearkombinationen der Vektoren u, v und w, wobei die Koeffizienten nur aus $[0,1]$ stammen. Diese Menge ist gerade der durch die Vektoren u, v und w aufgespannte Spat (oder das Parallelepiped):



Das Volumen des Spats ist gegeben durch das Produkt aus dessen Grundfläche und Höhe. Die Grundfläche, d.h. das Parallelogramm, welches von u und v aufgespannt wird, hat den Flächeninhalt

$$|u| \cdot |v| \cdot |\sin \sphericalangle(u,v)| = |u \times v|$$

und die Höhe ist gegeben durch $|w|\,|\cos \sphericalangle(u \times v, w)|$, so dass das Volumen des Spats gegeben ist durch

$$\left| |u \times v| \cdot |w| \cdot \cos \sphericalangle(u \times v, w) \right| = |\langle u \times v, w \rangle| = |\det(A)|,$$

wobei die letzte Gleichheit auf Übungsblatt 9 gezeigt wurde.

Sonderfälle:

- *Sind die drei Vektoren u, v und w linear abhängig, so ist $\det(A) = 0$ und in der Tat entsteht eine Teilmenge einer Ebene, so dass $f(M)$ das Volumen 0 besitzt.*
- *Ist $A = \operatorname{diag}(a,b,c)$ für $a,b,c > 0$, so ist $\det(A) = a \cdot b \cdot c$. Die Menge $f(M)$ ist in diesem Fall ein Quader mit Seitenlängen a, b und c und dieser besitzt in der Tat das Volumen $a \cdot b \cdot c$.*

In dem Beispiel haben wir gesehen, dass das Volumen eines Spats gegeben ist durch den Betrag der Determinante der Matrix, deren Spaltenvektoren den Spat aufspannen. Mit dem gleichen Prinzip können wir die Volumina der Bilder von Mengen unter affinen Abbildungen bestimmen. Folgendes Resultat wird gern als Anwendung von Determinanten genannt.

Satz 14.2

Sei $M \subseteq \mathbb{R}^n$ eine Menge, deren Volumen $V := \mathrm{Vol}(M)$ betrage. Sei zudem $f\colon \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung vom affinen Raum $\mathbb{R}^n$ über dem Vektorraum $\mathbb{R}^n$ in sich selbst. Diese habe die Form $f(x) = A \cdot x + b$ für eine Matrix $A \in \mathbb{R}^{n \times n}$ und einen Vektor $b \in \mathbb{R}^n$. Dann gilt:

$$\mathrm{Vol}(f(M)) = |\det(A)| \cdot \mathrm{Vol}(M).$$

Bemerkung 14.3. *Um diesen Satz zu beweisen, müssten wir zunächst sagen, was das* Volumen *einer Teilmenge des $\mathbb{R}^n$ ist für beliebiges $n \in \mathbb{N}$. Für den Sonderfall $n = 3$ haben wir eine Vorstellung davon, was gemeint ist und für den Fall $n = 2$ können wir einfach sagen, das Volumen von $M \subseteq \mathbb{R}^2$ sei dessen Flächeninhalt. In höheren Dimensionen fehlt uns dafür die Vorstellung. Eine Möglichkeit ist es, zu fordern, dass ein n-dimensionaler Würfel des $\mathbb{R}^n$ mit Seitenlänge $a > 0$,*

$$[0,a]^n := \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \;\middle|\; x_k \in [0,a], k \in \{1,\dots n\} \right\},$$

das Volumen a^n haben soll und das Volumen jeder anderen Menge dadurch bestimmt wird, dass diese Menge durch beliebig kleine Würfel so genau wie möglich approximiert wird. Schon diese Formulierung signalisiert, dass es bei einer echten Definition der Begriffe der Analysis bedarf, weshalb wir in Ermangelung eines Volumenbegriffs auch auf den Beweis des Satzes verzichten.

Zeigen wir anhand einiger Beispiele, was wir in einfachen Fällen bereits aus diesem Satz ableiten können.

Beispiel 14.4. *Sei $M \subseteq \mathbb{R}^2$ eine beliebige Menge mit Flächeninhalt $\mathrm{Vol}(M) > 0$. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch eine Drehung und anschließende Verschiebung, d.h.*

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix}$$

für $\alpha \in [0, 2\pi)$ und $x', y' \in \mathbb{R}$. Dann gilt

$$\mathrm{Vol}(f(M)) = \det\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \mathrm{Vol}(M) = 1 \cdot \mathrm{Vol}(M) = \mathrm{Vol}(M).$$

Eine Drehung oder Spiegelung ändert genauso wenig am Flächeninhalt einer Menge des $\mathbb{R}^2$ wie eine Verschiebung der Menge – der Satz bestätigt unsere Erfahrung, da gerade orthogonale Matrizen immer die Determinante 1 besitzen.

Beispiel 14.5. *Sei*

$$M = \left\{ x \in \mathbb{R}^2 \;\middle|\; x^2 + y^2 \leq 1 \right\}$$

ein Kreis mit Radius 1. Dessen Flächeninhalt ist bekanntlich $\pi \cdot 1^2 = \pi$. Sei

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2 \quad \textit{gegeben durch} \quad f(x) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \cdot x$$

für $a, b > 0$. Dann ist der Flächeninhalt von der Ellipse $f(M)$, deren Halbachsen die Längen a und b besitzen, gegeben durch $\det(\mathrm{diag}(a,b)) \cdot \pi = ab\pi$.

14.2 Die Spektralzerlegung symmetrischer Matrizen

Satz 13.7 wird auch als *Spektralsatz* bezeichnet, da die Menge aller Eigenwerte einer Matrix als deren *Spektrum* bezeichnet wird.[1]

Aus diesem Satz folgt die sogenannte *Spektralzerlegung* einer symmetrischen Matrix:

Korollar 14.6

Sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix mit den reellen Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Sei $S \in \mathcal{O}(n)$ eine orthogonale Matrix, die A diagonalisiert, d.h.

$$S^T A S = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Sind $v_1, \dots, v_n$ die Spaltenvektoren von S, so gilt

$$A = \sum_{k=1}^{n} \lambda_k v_k v_k^T.$$

Beispiel 14.7. *Wählen wir ein einfach zu rechnendes Beispiel:*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Diese Matrix ist symmetrisch und die Einheitsmatrix ist eine orthogonale Matrix (da $E_3^T = E_3^{-1} = E_3$), welche A diagonalisiert. In der Tat gilt für die Spaltenvektoren e_1, e_2, e_3 von E_3

$$\begin{aligned} & 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \\ = & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} = A. \end{aligned}$$

Beweis. Aus Satz 13.7 folgt die Existenz einer solchen orthogonalen Matrix $S \in \mathcal{O}(n)$ und nach Lemma 11.87 ist $\{v_1, \dots, v_n\}$ eine ONB des $\mathbb{R}^n$ bezüglich des kanonischen Skalarproduktes. Da S orthogonal ist, ist $S^{-1} = S^T$, d.h. für $D := \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ gilt

$$\begin{aligned} A = SDS^T &= \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 v_1 & \cdots & \lambda_n v_n \end{pmatrix} \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} = \lambda_1 v_1 v_1^T + \ldots + \lambda_n v_n v_n^T. \end{aligned}$$

[1] Der Begriff *Spektrum* wird besonders jenseits von Endomorphismen auf einem endlich-dimensionalen Vektorraum verwendet. In diesem Fall gelten einige Subtilitäten, die bei Matrizen und Endomorphismen auf endlich-dimensionalen Vektorräumen ignoriert werden können. Beispielsweise gilt die Äquivalenz von Injektivität, Surjektivität und Bijektivität und somit Invertierbarkeit von Endomorphismen nur auf Grund der Dimensionsformel, welche nur für endliche Dimensionen gilt.

□

Bemerkung 14.8. *Die Matrizen* $v_k v_k^T$ *werden auch als* Projektionsmatrizen *bezeichnet, denn für einen beliebigen Vektor* $x \in \mathbb{R}^n$ *ist*

$$(v_k v_k^T)x = v_k(v_k^T x) = v_k \langle v_k, x\rangle = \langle v_k, x\rangle v_k,$$

d.h. durch Vergleich mit Beispiel 11.64 sehen wir, dass $x \mapsto (v_k v_k^T)x$ *die orthogonale Projektion des Vektors* x *auf den Unterraum, der von* v_k *aufgespannt wird, also auf* $\mathcal{L}(v_k)$*, darstellt.*

Der Satz kann dazu verwendet werden, eine symmetrische Matrix zu konstruieren, die vorgegebene Eigenwerte und -vektoren besitzt.

Beispiel 14.9. *Seien* $\lambda_1 = -3$ *und* $\lambda_2 = 2$ *Eigenwerte einer symmetrischen Matrix des* $\mathbb{R}^{2\times 2}$ *und seien*

$$v_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad \textit{und} \quad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

zugehörige Eigenvektoren. Da die Vektoren bereits orthogonal zueinander sind bezüglich des kanonischen Skalarproduktes, müssen wir diese nur noch normieren:

$$\|v_1\| = \|v_2\| = \sqrt{1^2 + 2^2} = \sqrt{5}.$$

Also setzen wir $b_1 := \frac{1}{\sqrt{5}} v_1$ *und* $b_2 := \frac{1}{\sqrt{5}} v_2$*. Damit erhalten wir*

$$\begin{aligned} A := \lambda_1 b_1 b_1^T + \lambda_2 b_2 b_2^T &= \frac{-3}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \end{pmatrix} + \frac{2}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \end{pmatrix} \\ &= \frac{-3}{5} \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} + \frac{2}{5} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} -10 & 10 \\ 10 & 5 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 2 & 1 \end{pmatrix}. \end{aligned}$$

Überprüfen wir, dass v_1 *und* v_2 *Eigenvektoren von* A *zu den Eigenwerten* λ_1 *und* λ_2 *sind:*

$$\begin{aligned} Av_1 &= \begin{pmatrix} -2 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \end{pmatrix} = -3 \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \lambda_1 v_1, \\ Av_2 &= \begin{pmatrix} -2 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \lambda_2 v_2. \end{aligned}$$

14.3 QR-Zerlegung

Kommen wir zurück zum Problem, lineare Gleichungssysteme zu lösen. Ist die Koeffizientenmatrix A des LGS $A \cdot x = b$ eine Matrix in Zeilenstufenform, so kann man das LGS „von unten nach oben“ lösen.

Beispiel 14.10. *Das LGS*

$$\begin{cases} x + 2y + 3z = a \\ \quad\;\; y + 4z = b \\ \qquad\quad\; z = c \end{cases}$$

hat die eindeutige Lösung

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a - 3c - 2(b - 4c) \\ b - 4c \\ c \end{pmatrix} = \begin{pmatrix} a - 2b + 5c \\ b - 4c \\ c \end{pmatrix}.$$

Nehmen wir an, für eine Koeffizientenmatrix $A \in \mathbb{R}^{m \times n}$ mit $m \geq n$ hätten wir eine Produktdarstellung als $A = QR$, wobei $Q \in \mathbb{R}^{m \times n}$ ist und $Q^T Q = E_n$ erfüllt und $R \in \mathbb{R}^{n \times n}$ eine (rechte) obere Dreiecksmatrix ist. Dann gilt für das LGS

$$Ax = b \iff Q^T Ax = Q^T b \overset{Q^T Q = E_n}{\iff} Rx = Q^T b,$$

d.h.

$$L(A; b) = L(R; Q^T b).$$

Anstatt ein LGS mit beliebiger Matrix zu lösen, müsste man bei bekannter QR-Zerlegung nur den Ergebnisvektor $Q^T b$ berechnen und dann das neue LGS $R \cdot x = Q^T b$ (von unten nach oben) lösen. Ob oder unter welchen Umständen dieser Weg sinnvoll ist, ist Inhalt anderer Vorlesungen – wir wollen in diesem Abschnitt nur zeigen, dass jede Matrix $A \in \mathbb{R}^{m \times n}$ mit $m \geq n$ und $\mathrm{rg}(A) = n$ eine solche QR-Zerlegung besitzt und wie diese mit Hilfe des Gram-Schmidt'schen Orthonormalisierungsverfahrens berechnet werden kann.

Satz 14.11

Sei $A \in \mathbb{R}^{m \times n}$ mit $m \geq n$ und $\mathrm{rg}(A) = n$. Dann existieren eine Matrix $Q \in \mathbb{R}^{m \times n}$ und eine Matrix $R \in \mathbb{R}^{n \times n}$, welche folgende Bedingungen erfüllen:

$$A = QR, \quad Q^T Q = E_n, \quad R = (\rho_{k,j}) \text{ mit } \rho_{k,j} = 0 \text{ für } k > j.$$

Bemerkung 14.12.

i) *Die Bedingung an die Einträge von R bedeutet, dass es sich um eine rechte obere Dreiecksmatrix handelt.*

ii) *Ist $A \in \mathbb{R}^{n \times n}$ quadratisch, so ist auch $Q \in \mathbb{R}^{n \times n}$ und somit bedeutet $Q^T Q = E_n$, dass Q eine orthogonale Matrix ist. Ist jedoch $A \in \mathbb{R}^{m \times n}$ mit $m > n$, so ist zwar $Q^T Q = E_n$ möglich und wir werden sehen, dass die Spalten von Q ein Orthonormalsystem bilden. Der Begriff einer orthogonalen Matrix ist jedoch für quadratische Matrizen mit dieser Eigenschaft vorbehalten.*

Der Beweis ist gleichzeitig eine Konstruktionsanleitung für Q und R. Es gibt noch andere Möglichkeiten zur Konstruktion – da es uns jedoch nicht um numerische Eigenschaften der Konstruktion geht, soll dies genügen.

Beweis. Sei $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt des $\mathbb{R}^m$ und $\|v\| := \sqrt{\langle v, v \rangle}$ die durch das Skalarprodukt induzierte Norm.

1. Seien $v_1, \ldots, v_n \in \mathbb{R}^m$ die Spaltenvektoren von $A \in \mathbb{R}^{m \times n}$. Ist wie gefordert $\operatorname{rg}(A) = n$, so bedeutet das, dass diese Vektoren linear unabhängig sind, so dass das Gram-Schmidt-Verfahren auf sie angewendet werden kann.

2. Seien $w_1, \ldots, w_n \in \mathbb{R}^m$ die zueinander orthogonalen und normierten Vektoren, die uns das Gram-Schmidt-Verfahren liefert. Mit anderen Worten,

$$\begin{aligned} w_1 &= \frac{v_1}{\|v_1\|}, \\ w_2 &= \frac{v_2 - \langle v_2, w_1 \rangle w_1}{\|v_2 - \langle v_2, w_1 \rangle w_1\|} \\ &\vdots \\ w_n &= \frac{v_n - \sum_{i=1}^{n-1} \langle v_n, w_i \rangle w_i}{\left\|v_n - \sum_{i=1}^{n-1} \langle v_n, w_i \rangle w_i\right\|}. \end{aligned}$$

3. Bezeichnen wir Q als die Matrix, deren Spalten aus den Vektoren $w_1, \ldots, w_n$ bestehen, so gilt $Q^T Q = E_n$, denn

$$Q^T Q = \begin{pmatrix} w_1^T \\ \vdots \\ w_n^T \end{pmatrix} \begin{pmatrix} w_1 & \cdots & w_n \end{pmatrix} = \begin{pmatrix} w_1^T w_1 & w_1^T w_2 & \cdots & w_1^T w_n \\ w_2^T w_1 & w_2^T w_2 & \cdots & w_2^T w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n^T w_1 & w_n^T w_2 & \cdots & w_n^T w_n \end{pmatrix} \overset{(*)}{=} E_n,$$

   wobei $(*)$ gilt, da $\{w_1, \ldots, w_n\}$ ein Orthonormalsystem ist, d.h. $w_k^T w_j = \langle w_k, w_j \rangle = \delta_{k,j}$ für $k, j \in \{1, \ldots, n\}$.

4. $R := Q^T A$ ist nach Konstruktion eine rechte obere Dreiecksmatrix, denn aus

$$Q^T A = \begin{pmatrix} w_1^T \\ \vdots \\ w_n^T \end{pmatrix} \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = (w_k^T v_j)_{k,j=1,\ldots,n} = (\langle w_k, v_j \rangle)_{k,j=1,\ldots,n}$$

   folgt, dass wenn $R = (\rho_{k,j})$ ist, die Matrixeinträge links unterhalb der Diagonale verschwinden, d.h. für $k > j$ ist $\rho_{k,j} = \langle w_k, v_j \rangle = 0$. Im Allgemeinen ist die Formel recht lang, daher zeigen wir dies nur für $\rho_{2,1} = \langle w_2, v_1 \rangle$. Zunächst einmal ist

$$\langle w_1, v_1 \rangle = \langle \frac{v_1}{\|v_1\|}, v_1 \rangle = \frac{\langle v_1, v_1 \rangle}{\|v_1\|} = \frac{\|v_1\|^2}{\|v_1\|} = \|v_1\|.$$

Mit den Formeln für w_1 und w_2 aus dem Gram-Schmidt-Verfahren ist damit

$$\begin{aligned}\rho_{2,1} &= \langle w_2, v_1\rangle = \langle \frac{v_2 - \langle v_2, w_1\rangle w_1}{\|v_2 - \langle v_2, w_1\rangle w_1\|}, v_1\rangle \\ &= \frac{\langle v_2 - \langle v_2, w_1\rangle w_1, v_1\rangle}{\|v_2 - \langle v_2, w_1\rangle w_1\|} \\ &= \frac{\langle v_2, v_1\rangle - \langle v_2, w_1\rangle\langle w_1, v_1\rangle}{\|v_2 - \langle v_2, w_1\rangle w_1\|} \\ &= \frac{1}{\|v_2 - \langle v_2, w_1\rangle w_1\|}\left(\langle v_2, v_1\rangle - \langle v_2, \frac{v_1}{\|v_1\|}\rangle\langle \frac{v_1}{\|v_1\|}, v_1\rangle\right) \\ &= \frac{1}{\|v_2 - \langle v_2, w_1\rangle w_1\|}\left(\langle v_2, v_1\rangle - \frac{1}{\|v_1\|^2}\langle v_2, v_1\rangle\langle v_1, v_1\rangle\right) \\ &= \frac{1}{\|v_2 - \langle v_2, w_1\rangle w_1\|}\left(\langle v_2, v_1\rangle - \frac{1}{\|v_1\|^2}\langle v_2, v_1\rangle\,\|v_1\|^2\right) \\ &= 0.\end{aligned}$$

Für beliebige Indizes $k > j$ folgt die Aussage analog.

□

Verdeutlichen wir die QR-Zerlegung anhand eines Beispiels.

Beispiel 14.13. *Sei*

$$A = \begin{pmatrix}0 & 1\\ 1 & 0\\ 1 & 1\end{pmatrix} \quad \textit{mit Spaltenvektoren} \quad v_1 = \begin{pmatrix}0\\1\\1\end{pmatrix} \quad \textit{und} \quad v_2 = \begin{pmatrix}1\\0\\1\end{pmatrix}.$$

Es ist offensichtlich $\operatorname{rg}(A) = 2$*, d.h. nach obigem Satz besitzt* A *eine QR-Zerlegung. Wenden wir das Gram-Schmidt-Verfahren auf* $\{v_1, v_2\}$ *an:*

$$w_1 = \frac{1}{\left\|\begin{pmatrix}0\\1\\1\end{pmatrix}\right\|}\begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}0\\1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix},$$

$$\begin{aligned}w_2 &= \frac{v_2 - \langle v_2, w_1\rangle w_1}{\|v_2 - \langle v_2, w_1\rangle w_1\|} \\ &= \frac{1}{\|v_2 - \langle v_2, w_1\rangle w_1\|}\left(\begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{1}{2}\langle\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}\rangle\begin{pmatrix}0\\1\\1\end{pmatrix}\right) \\ &= \frac{1}{\|v_2 - \langle v_2, w_1\rangle w_1\|}\left(\begin{pmatrix}1\\0\\1\end{pmatrix} - \begin{pmatrix}0\\1/2\\1/2\end{pmatrix}\right) \\ &= \frac{1}{\left\|\begin{pmatrix}1\\-1/2\\1/2\end{pmatrix}\right\|}\begin{pmatrix}1\\-1/2\\1/2\end{pmatrix} = \frac{1}{\sqrt{3/2}}\cdot\frac{1}{2}\cdot\begin{pmatrix}2\\-1\\1\end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix}2\\-1\\1\end{pmatrix}.\end{aligned}$$

Mit Hilfe dieser Vektoren bestimmen wir Q und R:

$$Q = \begin{pmatrix} 0 & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix},$$

$$R = Q^T A = \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{1/2} \\ 0 & \sqrt{3/2} \end{pmatrix}.$$

R hat gut erkennbar die gewünschte Form. Überprüfen wir noch, dass $Q^T Q = E_2$ ist:

$$Q^T Q = \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 0 & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Aus $R = Q^T A$ folgt durch Multiplikation mit Q von links sofort $A = QR$, d.h. wir haben die gewünschte Zerlegung gefunden.

Bemerkung 14.14. *Die QR-Zerlegung ist im Allgemeinen nicht eindeutig bestimmt. Beispielsweise kann man beide Matrizen mit -1 multiplizieren ohne die Eigenschaften zu verlieren. Ist jedoch $A \in \mathbb{R}^{n \times n}$ regulär und fordert man von R, dass die Diagonalelemente positiv sind, d.h. $\rho_{k,k} > 0$ für $k \in \{1, \ldots, n\}$, so erhalten wir eine eindeutig bestimmte QR-Zerlegung.*

Es gibt noch weitere Zerlegungen, die wir hier nicht behandeln werden. Beispielsweise ist die LU- bzw. LR-Zerlegung[2] einer Matrix A von der Form $A = PLU$, wobei L eine linke untere Dreiecksmatrix, U eine rechte obere Dreiecksmatrix und P eine Permutationsmatrix ist, die in jeder Zeile und Spalte genau eine 1 als Eintrag und sonst ausschließlich Nullen hat.

14.4 Der Satz von Cayley-Hamilton

In diesem Abschnitt befassen wir uns mit einem Ergebnis, welches auf den britischen Mathematiker Arthur Cayley (1821-1895) und den irischen Mathematiker und Physiker William Rowan Hamilton[3] (1805-1865) zurückgeht.

Satz 14.15: Satz von Cayley-Hamilton

Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und sei $f \in L(V;V)$ ein Endomorphismus. Dann gilt für das charakteristisch Polynom p_f von f

$$p_f(f) = 0,$$

wobei 0 die konstante Nullabbildung ist.

Aus diesem Satz folgt unmittelbar das bereits auf Übungsblatt 6 erwähnte folgende Resultat:

[2] LU steht für *left upper*, LR hingegen für *links rechts*.

[3] W. R. Hamilton ist die Entwicklung der *Quaternionen* zu verdanken, einem Zahlenbereich, der über reelle und komplexe Zahlen hinausgeht.

Korollar 14.16

Seien $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine Matrix (für $n \in \mathbb{N}$). Dann gilt für das charakteristische Polynom p_A von A

$$p_A(A) = O,$$

wobei $O \in \mathbb{K}^{n\times n}$ die Nullmatrix sei.

Auf Blatt 6 hatten wir einen sehr kurzen, aber leider nicht korrekten Beweis des Korollars gesehen. Jetzt wollen wir uns einen korrekten Beweis ansehen. Da wir den Beweis des Satzes in der Formulierung für Endomorphismen durchführen wollen, sehen wir uns vorher an einem Beispiel an, wie die Aussage in diesem Fall konkret aussieht.

Beispiel 14.17. *Sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch*

$$f(x) = Ax \quad \textit{für} \quad A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Das charakteristische Polynom eines Endomorphismus[4] *ist definiert als charakteristisches Polynom einer darstellenden Matrix von f. In unserem Fall ist $A = A_f^{\mathcal{B}}$ die darstellende Matrix von f bezüglich der kanonischen Basis $\mathcal{B}$ des $\mathbb{R}^2$, d.h.*

$$p_f(t) = p_A(t) = \det \begin{pmatrix} 1-t & 0 \\ 2 & 1-t \end{pmatrix} = t^2 - 2t + 1.$$

Damit gilt

$$\begin{aligned} p_f(f)(x) &= f^2(x) - 2f(x) + f^0(x) = (f(x))^2 - 2f(x) + \mathrm{id}_{\mathbb{R}^2}(x) \\ &= \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} 1-2+1 & 0-0+0 \\ 4-4+0 & 1-2+1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Ox, \end{aligned}$$

d.h. $p_f(f)$ ist in der Tat die lineare Abbildung, die jeden Vektor auf dem Nullvektor abbildet.

Das Beispiel zeigt also, dass wir zum Beweis des Satzes von Cayley-Hamilton zeigen müssen, dass $p_f(f)(x) = e_V$ ist für alle $x \in V$.

Beweis von Satz 14.15. Sei $f \in L(V;V)$ ein Endomorphismus und p_f das charakteristische Polynom von f. Für den Nullvektor e_V gilt $f(e_V) = e_V$, da f linear ist, und somit ist

[4] In Abschnitt 9.4.2 wurde gezeigt, dass die darstellenden Matrizen eines Endomorphismus bezüglich unterschiedlichen Basen identisch ist und daraufhin wurde das charakteristische Polynom eines Endomorphismus als das charakteristische Polynom der darstellenden Matrix bezüglich einer fest gewählten Basis definiert.

auch $f^m(e_V) = (f \circ \dots \circ f)(e_V) = e_V$. Zudem ist $\mathrm{id}_V(e_V) = e_V$, so dass für jedes Polynom $\widetilde{p} \in \mathbb{R}[t]$ gilt:

$$\widetilde{p}(f)(e_V) = e_V.$$

Dies gilt also insbesondere für das charakteristische Polynom, so dass wir ab jetzt $x \neq e_V$ annehmen und für diesen Vektor wollen wir $p_f(f)(x) = e_V$ zeigen.

Für ein festes $m \in \mathbb{N}_0$ betrachten wir

$$B_m := \{f^0(x), f^1(x), f^2(x), \dots, f^m(x)\} = \{x, f(x), f^2(x), \dots, f^m(x)\},$$

wobei $f^0 = \mathrm{id}_V$ ist. Für $m = 0$ ist $B_0 = \{x\}$ linear unabhängig wegen $x \neq e_V$. Für $m \geq n$ besteht B_m aus $m + 1 \geq n + 1$ Vektoren und ist wegen $\dim(V) = n$ linear abhängig. Folglich gibt es ein kleinstes $m^* \in \mathbb{N}$, für das B_{m^*} linear abhängig ist, d.h. B_{m^*-1} ist linear unabhängig und $B_{m^*} = M_{m^*-1} \cup \{f^{m^*}(x)\}$ ist linear abhängig. Schreiben wir ab jetzt m statt m^* für diesen ausgewählten Index.

Nach Konstruktion existieren (von x abhängige) Skalare $a_0, a_1, \dots, a_{m-1} \in \mathbb{K}$, so dass

$$f^m(x) = a_0 x + a_1 f(x) + \ldots + a_{m-1} f^{m-1}(x) \tag{14.4.1}$$

ist. Sei

$$q(t) := a_0 + a_1 t + \ldots + a_{m-1} t^{m-1} - t^m \in \mathbb{K}[t].$$

Dieses (von der Wahl von x abhängige) Polynom erfüllt $q(f)(x) = e_V$. Um zu zeigen, dass auch $p_f(f)(x) = e_V$ ist, betrachten wir den Untervektorraum

$$U := \mathcal{L}(B_{m-1}) = \mathcal{L}(\{x, f(x), f^2(x), \dots, f^{m-1}(x)\})$$

Es gilt

$$f(B_{m-1}) = \{f(x), f(f(x)), \dots, f(f^{m-1}(x))\} = \{f(x), f^2(x), \dots, f^m(x)\} \subseteq U,$$

da B_m linear abhängig ist und somit $f^m(x) \in \mathcal{L}(B_{m-1})$ ist. Folglich ist

$$f(U) = f(\mathcal{L}(B_{m-1})) \subseteq \mathcal{L}(U) = U, \tag{14.4.2}$$

da U selbst ein Untervektorraum ist.

Sei $g := f|_U \in L(U; U)$ die Einschränkung von f auf den UVR U. Da nach (14.4.2) $g(U) \subseteq U$ ist, ist g in der Tat ein Endomorphismus. Bezüglich der Basis $\mathcal{B}_{m-1}$ von U hat g die darstellende Matrix

$$\widetilde{A} := A_g^{\mathcal{B}_{m-1}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{m-2} \\ 0 & \cdots & 0 & 1 & a_{m-1} \end{pmatrix},$$

denn nach Definition von f und (14.4.1) ist

$$\begin{aligned} g(x) &= f(x) = 0 \cdot x + 1 \cdot f(x) + 0 \cdot f^2(x) + \ldots + 0 \cdot f^{m-1}(x), \\ g(f)(x) &= f^2(x) = 0 \cdot x + 0 \cdot f(x) + 1 \cdot f^2(x) + \ldots + 0 \cdot f^{m-1}(x), \\ &\vdots \\ g(f^{m-2})(x) &= f^{m-1}(x) = 0 \cdot x + 0 \cdot f(x) + 0 \cdot f^2(x) + \ldots + 1 \cdot f^{m-1}(x), \\ g(f^{m-1})(x) &= f^m(x) = a_0 \cdot x + a_1 \cdot f(x) + a_2 \cdot f^2(x) + \ldots + a_{m-1} \cdot f^{m-1}(x). \end{aligned}$$

Das charakteristische Polynom von g ist damit

$$p_g(t) = p_{\widetilde{A}}(t) = \det(\widetilde{A} - tE_m) = \begin{vmatrix} -t & 0 & \cdots & 0 & a_0 \\ 1 & -t & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -t & a_{m-2} \\ 0 & \cdots & 0 & 1 & a_{m-1} - t \end{vmatrix}$$

Addieren wir von unten nach oben gehend jeweils das t-Fache einer Zeile zur darüberstehenden Zeile, so ändert sich nach Eigenschaft (E4) aus Satz 8.16 die Determinante nicht, d.h.

$$p_g(t) = p_{\widetilde{A}}(t) = \begin{vmatrix} 0 & 0 & \cdots & 0 & a_0 + a_1 t + \ldots + a_{m-1}t^{m-1} - t^m \\ 1 & 0 & \cdots & 0 & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{m-2} + a_{m-1}t - t^2 \\ 0 & \cdots & 0 & 1 & a_{m-1} - t \end{vmatrix}$$

Entwicklung nach der 1. Zeile ergibt

$$p_g(t) = p_{\widetilde{A}}(t) = (-1)^{m+1} q(t) \cdot \det(E_{m-1}) = (-1)^{m+1} q(t).$$

Da $U = \mathcal{L}(B_{m-1})$ ein Untervektorraum von V ist, können wir $\mathcal{B}_{m-1}$ zu einer Basis B von V ergänzen. f hat dann bezüglich B eine Abbildungsmatrix folgender Blockgestalt:

$$A_f^{\mathcal{B}} = \begin{pmatrix} \widetilde{A} & C \\ O & D \end{pmatrix},$$

wobei $C \in \mathbb{K}^{m\times(n-m)}$ und $D \in \mathbb{K}^{(n-m)\times(n-m)}$ geeignete Matrizen sind und $O \in \mathbb{K}^{(n-m)\times m}$ die Nullmatrix ist. Nach Eigenschaft (E6) aus Satz 8.16, der Determinantenregel für Blockmatrizen, gilt

$$p_f = p_{A_f^{\mathcal{B}}} = p_{\widetilde{A}} \cdot p_D = (-1)^{m+1} p_D \cdot q.$$

Setzen wir $r(t) := (-1)^{m+1} p_D(t)$, so gilt $p_f = r \cdot q$ und damit $p_f(f) = r(f) \circ q(f)$. Damit folgt für unseren fixierten Vektor x

$$p_f(f)(x) = r(f)[q(f)(x)] = r(f)(e_V) \overset{r(f)\text{ linear}}{=} e_V.$$

Da $x \in V$ beliebig war, folgt $p_f(f)(v) = e_V$ für alle $v \in V$ und somit $p_f(f) = 0$. □

Aus diesem Satz folgt folgende Aussage:

Korollar 14.18

Wenn $A \in \mathbb{K}^{n\times n}$ das charakteristische Polynom $p_A(t) = \pm t^n$ hat, so ist A nilpotent, d.h. es existiert eine Zahl $k \in \mathbb{N}$, so dass $A^k = O$ (die Nullmatrix) ist.

Beweis. Nach Cayley-Hamilton folgt aus $p_A(t) = \pm t^n$ sofort $p_A(A) = \pm A^n = O$. □

14.5 Affine Koordinatentransformation an einem Beispiel

In diesem Abschnitt wollen wir an Beispiel 12.16 anknüpfen, in welchem wir die Koordinatentransformation zwischen dem Standard-Koordinatensystem und einem anderen Koordinatensystem angesehen haben. Anhand zweier Ebenen wollen wir die Koordinatentransformation zwischen zwei beliebigen Koordinatensystemen veranschaulichen.

Betrachten wir zwei affine Unterräume des affinen Raumes $A = \mathbb{R}^3$ über dem Vektorraum $V = \mathbb{R}^3$:

$$B_1 = \left\{ \begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} + a \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; a, b \in \mathbb{R} \right\} =: P_1 + U_1$$

$$B_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 7 \end{pmatrix} + c \cdot \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + d \cdot \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \;\middle|\; c, d \in \mathbb{R} \right\} =: P_2 + U_2.$$

Untersuchen wir zunächst die Lagebeziehungen dieser beiden Ebenen zueinander:

- $U_1 = U_2$, denn es gilt

$$\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

- $P_2 \in B_1$, denn es gilt

$$\begin{pmatrix} 1 \\ 0 \\ 7 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Aus beiden Punkten folgt, dass $B_1 = B_2$ ist, denn wegen $B_1 \cap B_2 \neq \emptyset$ gilt laut Dimensionsformel (12.4.3)

$$\dim(B_1 \cap B_2) = \dim(U_1 \cap U_2) = \dim(U_1) = 2,$$

d.h. der Schnitt beider Ebenen ist wieder eine Ebene und somit sind beide affinen Räume identisch. Nennen wir ab jetzt $B := B_1$.

Die verschiedenen Darstellung der Ebene liefern uns zwei verschiedene Koordinatensysteme:

$$K_1 = \left(\begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix}; \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$$

$$K_2 = \left(\begin{pmatrix} 1 \\ 0 \\ 7 \end{pmatrix}; \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \right).$$

Sehen wir uns die Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ an, welche den Koordinaten eines Punktes bezüglich des einen Koordinatensystems die Koordinaten des Punktes bezüglich des anderen Koordinatensystems zuordnet, d.h.

$$f(\kappa_{K_2}(x)) := \kappa_{K_1}(x), \quad x \in B.$$

Sei $x \in B$ mit $\kappa_{K_2}(x) = \begin{pmatrix} c \\ d \end{pmatrix}$, d.h.

$$x = \begin{pmatrix} 1 \\ 0 \\ 7 \end{pmatrix} + c \cdot \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + d \cdot \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$$

Auf Grund der bisherigen Überlegungen ist somit

$$\begin{aligned} x &= \left[\begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right] + c \cdot \left[2\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right] + d \cdot \left[1\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right] \\ &= \begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} + (-2 + 2c + d) \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + (-1 - c + d) \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Somit gilt

$$f\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} -2 + 2c + d \\ -1 - c + d \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} =: \begin{pmatrix} a \\ b \end{pmatrix}.$$

Die Koordinatenabbildung ist also eine affine Abbildung auf (in diesem konkreten Fall) dem affinen Raum $\mathbb{R}^2$. Sehen wir uns nun die umgekehrte Koordinatentransformation an:

$$\begin{aligned} & \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} \\ \iff & \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} a+2 \\ b+1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a+2 \\ b+1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 4/3 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}, \end{aligned}$$

d.h. $f^{-1}\colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $\kappa_{K_1}(x) \mapsto \kappa_{K_2}(x)$ ist die affine Abbildung mit der Vorschrift

$$f^{-1}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1/3 \\ 4/3 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}.$$

Sehen wir uns exemplarisch die Koordinaten der beiden Punkte $P_1, P_2 \in B$ an, welche jeweils bezüglich einem der Koordinatensysteme die sofort ablesbaren Koordinaten $\begin{pmatrix} 0 & 0 \end{pmatrix}^T$ haben.

Zu P_1: Es ist $\kappa_{K_1}(P_1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und somit laut Koordinatentransformation f^{-1}

$$\kappa_{K_2}\begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} = f^{-1}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 4/3 \end{pmatrix}.$$

In der Tat ist

$$\begin{pmatrix}3\\5\\8\end{pmatrix} = \begin{pmatrix}1\\0\\7\end{pmatrix} + \frac{1}{3}\begin{pmatrix}2\\3\\-1\end{pmatrix} + \frac{4}{3}\begin{pmatrix}1\\3\\1\end{pmatrix}.$$

Zu P_2: s ist $\kappa_{K_2}(P_2) = \begin{pmatrix}0\\0\end{pmatrix}$ und somit laut Koordinatentransformation f

$$\kappa_{K_1}\begin{pmatrix}1\\0\\7\end{pmatrix} = f\begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}-2\\-1\end{pmatrix}.$$

In der Tat ist

$$\begin{pmatrix}1\\0\\7\end{pmatrix} = \begin{pmatrix}3\\5\\8\end{pmatrix} - 2\begin{pmatrix}1\\2\\0\end{pmatrix} - \begin{pmatrix}0\\1\\1\end{pmatrix}.$$

Wir sehen also anhand des Beispiels, dass wir für affine Räume und affine Koordinaten keine neue Theorie erfinden müssen. Während wir in Vektorräumen die Koordinaten der Basisvektoren bezüglich der anderen Basis bestimmt haben um die allgemeine Koordinatentransformation zu bestimmen, müssen wir nun die Koordinaten der Basis und des Koordinatenursprungs bezüglich des anderen Koordinatensystems kennen um daraus die Abbildungsvorschrift der Koordinatentransformation herzuleiten.

Ein Nebenprodukt der obigen Rechnung soll als zusätzliche Information zu affinen Abbildungen noch erwähnt werden:

Ist $f\colon \mathbb{R}^n \to \mathbb{R}^n$ von der Form $f(x) = b + Ax$, d.h. eine affine Abbildung, so ist f genau dann invertierbar, wenn A invertierbar ist. Es ist nämlich in diesem Fall

$$y = b + A \cdot x \iff x = A^{-1} \cdot (y - b), \quad \text{d.h.} \quad f^{-1}(y) = A^{-1} \cdot (-b) + A^{-1} \cdot x.$$

Dieses Beispiel und die gerade gemachte Beobachtung stehen exemplarisch für eine ganze Reihe von Resultaten, die man in leicht modifizierter Form von Vektorräumen und linearen Abbildungen auf affine Räume und affine Abbildungen übertragen kann.

Literaturverzeichnis

[Beu14] BEUTELSPACHER, Albrecht: *Lineare Algebra : eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen*. 8., aktualisierte Aufl. Wiesbaden : Springer Spektrum, 2014 (Springer-Lehrbuch)

[Bos14] BOSCH, Siegfried: *Lineare Algebra von Siegfried Bosch*. 5., überarb. u. erweiterte Aufl. Berlin, Heidelberg : Springer Spektrum, 2014 (Springer-Lehrbuch)

[Dal] DALWIGK, Florian A.: *Vollständige Induktion : Beispiele und Aufgaben bis zum Umfallen*. Berlin : Springer Spektrum

[Fil11] FILLER, Andreas: *Elementare Lineare Algebra: Linearisieren und Koordinatisieren*. Springer-Verlag, 2011

[Fis17] FISCHER, Gerd: *Lernbuch Lineare Algebra und Analytische Geometrie : das Wichtigste ausführlich für das Lehramts- und Bachelorstudium*. 3., verbesserte und ergänzte Auflage. Wiesbaden : Springer Spektrum, 2017

[For16] FORSTER, Otto: *Analysis 1: Differential-und Integralrechnung einer Veränderlichen*. Springer-Verlag, 2016

[Kur08] KURZWEIL, Hans: *Endliche Körper: Verstehen, Rechnen, Anwenden*. 2., überarb. Aufl. Berlin ; Heidelberg : Springer-Verlag, 2008 (Springer-Lehrbuch)

[Kö04] KÖNIGSBERGER, Konrad: *Analysis. 1 : mit 250 Aufgaben samt ausgearbeiteten Lösungen*. 6., durchges. Aufl. Springer-Verlag, 2004

[LM15] LIESEN, Jörg ; MEHRMANN, Volker: *Lineare Algebra*. 2. Auflage. Springer Spektrum, 2015