

Mathematik II – SoSe 2020

Dr. Jana Bielagk

Inhaltsverzeichnis

1	Vektorräume	1
1.1	Vorüberlegung: Was ist ein Körper?	1
1.2	Vektorräume	3
1.3	Norm und Skalarprodukt	5
1.4	Lineare Unabhängigkeit und Erzeugendensystem	7
1.5	Basis, Dimension, Koordinaten	9
1.6	Orthogonale Projektion	11
1.7	Gram-Schmidt'sches Orthonormalisierungsverfahren	12
2	Lineare Abbildungen und Matrizen	14
2.1	Lineare Abbildungen	14
2.2	Matrizen	18
2.2.1	Rechnen mit Matrizen	19
2.3	Matrizen und lineare Abbildungen	22
3	Lineare Gleichungssysteme	25
3.1	Einführung in LGS – Lösbarkeit und Lösungsmenge	25
3.2	Methoden zur Lösung von LGS	27
3.2.1	Gauß-(Jordan-)Verfahren	29
4	Differentialrechnung für Funktionen mehrerer reeller Variablen	33
4.1	Differenzierbarkeit von Funktionen in mehreren Variablen	33
4.2	Ableitungen höherer Ordnung in mehreren Variablen	35
4.3	Vektorfelder	36
5	Determinanten	37
5.1	Definition, Berechnung und Eigenschaften	37
5.2	Anwendung: Lösung von LGS mit der Cramer'schen Regel	40
5.3	Anwendung: Berechnung der Inversen einer Matrix	40
5.4	Anwendung im Spezialfall 2D / 3D: Fläche / Volumen berechnen	41
5.5	Anwendung: Polynominterpolation — Vandermonde-Determinante	42
6	Eigenwertprobleme	44
6.1	Definition, Berechnung, Eigenschaften	44
6.1.1	Invertierbarkeit	46
6.2	Definitheit symmetrischer Matrizen	46
6.2.1	Anwendung: Extremwertaufgaben	48
6.3	Diagonalisierbarkeit von Matrizen	53
6.3.1	Vielfachheiten	53
6.3.2	Diagonalisierung	55

6.4	Anwendung: Matrixfunktionen	56
7	Differentialgleichungen	60
7.1	Einführung	60
7.2	Elementare Lösungsmethoden	62
7.2.1	Getrennte Variablen	62
7.2.2	Lineare DGL 1. Ordnung	63
7.2.3	Homogene DGL	64
7.3	Existenz und Eindeutigkeit von Lösungen	65
7.4	Exakte Differentialgleichungen	67
8	Lineare Differentialgleichungen höherer Ordnung und Systeme von linearen Differentialgleichungen	69
8.1	Systeme von linearen DGLen	69
8.1.1	Homogene DGL-Systeme mit konstanten Koeffizienten – Lösungsmethode	69
8.1.2	Homogene DGL-Systeme – Charakterisierung der Lösungsmenge	71
8.1.3	Homogene DGL-Systeme mit nicht-konstanten Koeffizienten	72
8.1.4	Inhomogene DGL-Systeme mit konstanten Koeffizienten	73
8.2	Lineare DGLen höherer Ordnung	73
8.2.1	Homogene lineare DGLen höherer Ordnung mit konstanten Koeffizienten	73
8.2.2	Inhomogene lineare DGLen höherer Ordnung mit konstanten Koeffizienten	75
9	Flächen- und Raumintegrale	77

1 Vektorräume

Vektoren sind in der Schule Elemente des \mathbb{R}^2 oder \mathbb{R}^3 – sie werden wahlweise als Zahlentupel oder geometrisch als ein Pfeil dargestellt. Von dieser Vorstellung wollen wir uns hier lösen und einen allgemeinen Begriff eines Vektorraumes über einem Körper einführen. Neben dem reellen Vektorraum \mathbb{R}^n werden wir auch die komplexen Zahlen als Vektorraum auffassen und Vektorräume von Funktionen kennen lernen.

Beispiel 1.1. Betrachten wir die Menge der reellen Polynome vom Grad höchstens 2, d.h.

$$\mathcal{P} := \{ p: \mathbb{R} \rightarrow \mathbb{R} \mid p(x) = ax^2 + bx + c, a, b, c \in \mathbb{R} \}.$$

Addieren wir Elemente aus \mathcal{P} , erhalten wir wieder ein Polynom vom Grad höchstens 2:

$$(a_1x^2 + b_1x + c_1) + (a_2x^2 + b_2x + c_2) = (a_1 + a_2)x^2 + (b_1 + b_2)x + (c_1 + c_2).$$

Die Multiplikation führt jedoch aus \mathcal{P} heraus: Ist $p(x) = x^2$ und $q(x) = 3x$, dann ist $(p \cdot q)(x) = 3x^3$, das heißt $p \cdot q \notin \mathcal{P}$. Die skalare Multiplikation, d.h. die Multiplikation eines Polynoms $p \in \mathcal{P}$ mit einer Zahl $\lambda \in \mathbb{R}$, liefert jedoch auf jeden Fall wieder ein Element aus \mathcal{P} :

$$\lambda \cdot p(x) = \lambda \cdot (ax^2 + bx + c) = (\lambda \cdot a)x^2 + (\lambda \cdot b)x + (\lambda \cdot c).$$

Im Folgenden werden wir erst sagen, was wir unter einem Körper verstehen – dabei können wir \mathbb{R} im Hinterkopf behalten als Beispiel für einen Körper. Dann werden wir den Begriff eines Vektorraumes über einem Körper einführen, welcher die in dem Beispiel gesehenen Eigenschaften hat: Vektoren kann man addieren oder mit einem Element des Körpers (einem Skalar) multiplizieren und das Ergebnis soll jeweils wieder ein Element des Vektorraumes sein.

1.1 Vorüberlegung: Was ist ein Körper?

Definition 1.2. Ein Körper ist eine Menge \mathbb{K} mit zwei Verknüpfungen $+$ und \cdot mit folgenden Eigenschaften:

1. $(\mathbb{K}, +)$ ist eine Abel'sche Gruppe, d.h. es existiert ein Element $0 \in \mathbb{K}$, so dass gilt:

- (i) $x + y = y + x$ für alle $x, y \in \mathbb{K}$ (Kommutativgesetz),
- (ii) $(x + y) + z = x + (y + z)$ für alle $x, y, z \in \mathbb{K}$ (Assoziativgesetz),
- (iii) $x + 0 = x$ für alle $x \in \mathbb{K}$ (neutrales Element),
- (iv) zu jedem $x \in \mathbb{K}$ existiert $-x \in \mathbb{K}$ so dass $x + (-x) = 0$ (inverses Element).

2. $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine Abel'sche Gruppe (mit neutralem Element 1), d.h.

- (i) $x \cdot y = y \cdot x$ für alle $x, y \in \mathbb{K} \setminus \{0\}$ (Kommutativgesetz),
- (ii) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ für alle $x, y, z \in \mathbb{K} \setminus \{0\}$ (Assoziativgesetz),
- (iii) $x \cdot 1 = x$ für alle $x \in \mathbb{K} \setminus \{0\}$ (neutrales Element),
- (iv) zu jedem $x \in \mathbb{K} \setminus \{0\}$ existiert $x^{-1} \in \mathbb{K} \setminus \{0\}$ so dass $x \cdot x^{-1} = 1$ (inverses Element).

3. Es gelten die Distributivgesetze, d.h. für alle $x, y, z \in \mathbb{K}$ gilt

$$\begin{aligned}x \cdot (y + z) &= x \cdot y + x \cdot z, \\(x + y) \cdot z &= x \cdot z + y \cdot z.\end{aligned}$$

Beispiel 1.3. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ sind Körper. $(\mathbb{Z}, +, \cdot)$ ist hingegen kein Körper, da das Multiplikativ Inverse einer ganzen Zahl außer ± 1 (z.B. $\frac{1}{3}$ für die 3) nicht mehr in der Menge der ganzen Zahlen liegt.

Beispiel 1.4. Gegeben sei die Menge $\mathbb{K} = \{0, 1, 2, 3, 4\}$. Mit den in folgenden Tabellen festgelegten Operationen Addition und Multiplikation (Rechnen mit Rest modulo 5) ist \mathbb{K} ein Körper:

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Aufgabe. Sei $\text{Abb}(\mathbb{R}; \mathbb{R})$ die Menge der Abbildungen von \mathbb{R} nach \mathbb{R} . Für zwei Funktionen $f, g: \mathbb{R} \rightarrow \mathbb{R}$ sei $(f + g)(x) := f(x) + g(x)$ und $(f \circ g)(x) := f(g(x))$.

- a) Überprüfen Sie, dass $(\text{Abb}(\mathbb{R}; \mathbb{R}), +)$ eine Abel'sche Gruppe ist. Wie lautet das neutrale Element "0"? Wie lautet das Inverse einer Funktion f bezüglich der Addition?
- b) Ist $(\text{Abb} \setminus \{0\}, \circ)$ eine Abel'sche Gruppe mit dem Nullelement "0" aus a)?

Lösung. a) Kommutativität und Assoziativität werden von der Addition auf \mathbb{R} geerbt. So gilt z.B. für $f, g \in \text{Abb}(\mathbb{R}; \mathbb{R})$ und $x \in \mathbb{R}$ stets

$$(f + g)(x) \stackrel{\text{Def.}}{=} f(x) + g(x) \stackrel{\text{KG in } \mathbb{R}}{=} g(x) + f(x) \stackrel{\text{Def.}}{=} (g + f)(x).$$

Das neutrale Element der Addition ist die konstante Nullfunktion – schreiben wir diese als

$$n: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto 0.$$

Dann gilt für beliebige Abbildungen f und beliebige $x \in \mathbb{R}$

$$(f + n)(x) = f(x) + n(x) = f(x) + 0 = f(x) = 0 + f(x) = (n + f)(x).$$

Das Inverse einer Funktion $f \in \text{Abb}(\mathbb{R}; \mathbb{R})$ ist gegeben durch

$$-f: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto (-1) \cdot f(x).$$

Dann gilt nämlich für alle $x \in \mathbb{R}$

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) + (-1) \cdot f(x) = 0.$$

- b) Die Antwort ist "Nein". Zunächst einmal ist das neutrale Element der Multiplikation in diesem Fall die Identität

$$i: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x,$$

denn dann gilt für jede Abbildung $f \in \text{Abb}(\mathbb{R}; \mathbb{R})$ und für alle $x \in \mathbb{R}$ stets

$$(f \circ i)(x) = f(i(x)) = f(x), \quad \text{also} \quad f \circ i = f.$$

Wenn $g := f^{-1}$ das Inverse einer Abbildung f bezüglich \circ sein soll, muss für alle $x \in \mathbb{R}$

$$(f \circ g)(x) = f(g(x)) \stackrel{!}{=} i(x) = x$$

gelten. Das heißt, g müsste die zu f inverse Abbildung (dem üblichen Sprachgebrauch nach) sein. Allerdings ist nicht jede Abbildung invertierbar – das trifft nur auf bijektive Abbildungen zu. Folglich ist $(\text{Abb} \setminus \{0\}, \circ)$ keine Abel'sche Gruppe. Insbesondere ist $(\text{Abb}, +, \circ)$ kein Körper.

1.2 Vektorräume

Während auf einem Körper die Addition und Multiplikation von Elementen des Körpers definiert ist, wollen nun für einen Vektorraum die Addition von Vektoren untereinander und die Multiplikation eines Vektors mit einem Skalar (d.h. einem Element des Körpers) definieren. Um beide Operationen für den Anfang auseinander halten zu können, führen wir für die Vektorraumoperationen neue Symbole ein. Sobald wir sicher im Unterschied zwischen beiden sind, werden wir wieder zu den bekannten Symbolen zurückkehren.

Definition 1.5 (Vektorraum). Sei $(\mathbb{K}, +, \cdot)$ ein Körper. Eine nichtleere Menge V , auf der eine Vektoraddition $\oplus: V \times V \rightarrow V$ und eine Skalarmultiplikation $\odot: \mathbb{K} \times V \rightarrow V$ definiert sind, heißt \mathbb{K} -Vektorraum (V, \oplus, \odot) , falls gilt:

1. (V, \oplus) bildet eine Abel'sche 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:

$$\begin{aligned} a \odot (v \oplus w) &= a \odot v + a \odot w, \\ (a + b) \odot v &= a \odot v + b \odot v. \end{aligned}$$

Beispiel 1.6. Sei $(\mathbb{R}, +, \cdot)$ der Körper der reellen Zahlen. Sei $V = \mathbb{R}^3$ mit folgenden Operationen:

$$v \oplus w = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \oplus \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} := \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix}, \quad a \odot v = a \odot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} := \begin{pmatrix} a \cdot v_1 \\ a \cdot v_2 \\ a \cdot v_3 \end{pmatrix},$$

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 Abel'sche 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

$$-v = \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix}, \quad \text{da } v \oplus (-v) = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \oplus \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix} = \begin{pmatrix} v_1 - v_1 \\ v_2 - v_2 \\ v_3 - v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0.$$

- 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:

$$a \odot (v \oplus w) = a \cdot \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix} = \begin{pmatrix} a \cdot (v_1 + w_1) \\ a \cdot (v_2 + w_2) \\ a \cdot (v_3 + w_3) \end{pmatrix} = \begin{pmatrix} a \cdot v_1 + a \cdot w_1 \\ a \cdot v_2 + a \cdot w_2 \\ a \cdot v_3 + a \cdot w_3 \end{pmatrix} = a \odot v \oplus a \odot w,$$

$$(a + b) \odot v = (a + b) \odot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} (a + b) \cdot v_1 \\ (a + b) \cdot v_2 \\ (a + b) \cdot v_3 \end{pmatrix} = \begin{pmatrix} a \cdot v_1 + b \cdot v_1 \\ a \cdot v_2 + b \cdot v_2 \\ a \cdot v_3 + b \cdot v_3 \end{pmatrix} = a \odot v \oplus b \odot v.$$

Wir können also grundsätzlich Vektoren addieren oder mit Elementen aus dem zu Grunde liegenden Körper \mathbb{K} multiplizieren. Eine Multiplikation von Vektoren ist also im Allgemeinen nicht definiert, ebenso wie die Division eines Vektors durch einen anderen.

Wenn wir im Folgenden von einem reellen Vektorraum sprechen, verstehen wir darunter einen \mathbb{R} -Vektorraum (d.h. $\mathbb{K} = \mathbb{R}$). Analog ist ein komplexer Vektorraum ein \mathbb{C} -Vektorraum.

Beispiel 1.7 (Gegenbeispiel). Sei $\mathbb{K} = \mathbb{R}$. Auf der Teilmenge $V := \{v \in \mathbb{R}^3 \mid v_1 = 1\}$ definieren wir die Addition \oplus und Multiplikation mit Skalaren \odot wie in Beispiel 1.6. Warum ist (V, \oplus, \odot) kein \mathbb{R} -Vektorraum?

Es sind

$$v = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \in V, w = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in V, \text{ aber } z := v \oplus w = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} \notin V.$$

Somit ist (V, \oplus) keine Gruppe, da die Summe zweier Vektoren aus V nicht mehr in V liegen muss. Folglich handelt es sich auch nicht um einen Vektorraum.

Ab jetzt werden wir wieder ausschließlich die Notation $+$ und \cdot benutzen, wissend dass es sich ggf. um unterschiedliche Operationen handelt.

Definition 1.8. 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 1.9. 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

$$v + w \in U \quad \text{für alle } v, w \in U \text{ und}$$

$$a \cdot v \in U \quad \text{für alle } v \in U, a \in \mathbb{K}.$$

Beispiel 1.10. Die Menge

$$\mathcal{E} := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid z = 0 \right\}$$

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 $a \in \mathbb{R}$

$$a \cdot v = \begin{pmatrix} a \cdot v_1 \\ a \cdot v_2 \\ a \cdot 0 \end{pmatrix} \in \mathcal{E}.$$

Folgende Menge ist jedoch kein Untervektorraum des \mathbb{R}^3 , da die Abgeschlossenheit bezüglich der Addition nicht gegeben ist:

$$\mathcal{E}' := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid z = 5 \right\}.$$

Es ist nämlich

$$\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} \in \mathcal{E}', \quad \begin{pmatrix} 3 \\ 3 \\ 5 \end{pmatrix} \in \mathcal{E}', \quad \text{aber} \quad \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 10 \end{pmatrix} \notin \mathcal{E}'.$$

1.3 Norm und Skalarprodukt

Im Folgenden wollen wir Längen von Vektoren bestimmen und feststellen, wann Vektoren senkrecht aufeinander stehen. Dafür führen wir die Begriffe *Norm* und *Skalarprodukt* ein.

Definition 1.11. Sei $(V, +, \cdot)$ ein reeller oder komplexer Vektorraum, d.h. $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Eine nicht-negative reellwertige Funktion $\|\cdot\| : V \rightarrow [0, \infty)$ heißt Norm, falls sie folgende Eigenschaften besitzt:

1. $\|v\| \geq 0$ für alle $v \in V$ und $\|v\| = 0 \iff v = 0$. (Positive 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).

Beispiel 1.12. Beispiele für Normen auf dem \mathbb{R} -Vektorraum $V = \mathbb{R}^n$ sind (mit $v = (v_1, \dots, v_n)$):

$$\|v\|_\infty := \max\{|v_1|, \dots, |v_n|\}, \quad \|v\|_2 := \sqrt{v_1^2 + \dots + v_n^2}, \quad \|v\|_1 := |v_1| + \dots + |v_n|.$$

Die eben eingeführte 2-Norm $\|\cdot\|_2$ wird auch Euklidische Norm genannt. Bevor wir uns überzeugen können, dass diese tatsächlich eine Norm ist, brauchen wir noch den Begriff des Skalarproduktes.

Definition 1.13 (Skalarprodukt). Gegeben sei ein \mathbb{R} -Vektorraum V . Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ heißt Skalarprodukt auf V , falls gilt:

1. $\langle v, w \rangle = \langle w, v \rangle$ für alle $v, w \in V$ (Symmetrie);
2. $\langle av, w \rangle = a \langle v, w \rangle$ für alle $v, w \in V$ und $a \in \mathbb{R}$ (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 $v \neq 0$ impliziert $\langle v, v \rangle > 0$ (positive Definitheit).

Beispiel 1.14. Ist $V = \mathbb{R}^n$, dann ist $\langle v, w \rangle := \sum_{k=1}^n v_k w_k = v_1 w_1 + \dots + v_n w_n$ ein Skalarprodukt.

Definition 1.15. Sei V ein \mathbb{C} -Vektorraum. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ heißt Skalarprodukt auf V , falls $\langle v, w \rangle = \overline{\langle w, v \rangle}$ für alle $v, w \in V$ gilt (d.h. $\langle \cdot, \cdot \rangle$ ist hermitesch) und zudem Eigenschaften 2-4 aus Definition 1.13.

Definition 1.16. Ein \mathbb{R} -Vektorraum mit einem Skalarprodukt heißt euklidischer Vektorraum; ein \mathbb{C} -Vektorraum mit Skalarprodukt heißt unitärer Vektorraum.

Satz 1.17. 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. Überprüfen wir die drei Eigenschaften einer Norm [im Fall eines euklidischen VR](#):

1. $\|v\| = \sqrt{\langle v, v \rangle} \geq 0$ für alle $v \in V$, da $\langle v, v \rangle \geq 0$ ist. Zudem folgt aus $\langle v, v \rangle = 0 \Rightarrow v = 0$, dass $\|v\| = 0$ genau dann erfüllt ist, wenn $v = 0$ ist.

2. Seien $a \in \mathbb{R}$ und $v \in V$. Dann ist mit den Eigenschaften des Skalarproduktes

$$\|a \cdot v\| = \sqrt{\langle av, av \rangle} = \sqrt{a \langle v, av \rangle} = \sqrt{a \langle av, v \rangle} = \sqrt{a^2 \langle v, v \rangle} = |a| \cdot \sqrt{\langle v, v \rangle} = |a| \|v\|.$$

3. Für die Überprüfung der Dreiecksungleichung seien $v, w \in V$. Es gilt

$$\|v + w\|^2 = \langle v+w, v+w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \stackrel{(*)}{\leq} \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$$

Durch Wurzelziehen erhält man die Aussage. Die Ungleichung (*) folgt aus folgendem Resultat:

□

Satz 1.18 (Cauchy-Schwarz'sche Ungleichung). Sei V ein \mathbb{K} -Vektorraum ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) 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 $a \in \mathbb{K}$ existiert, so dass $w = a \cdot v$ ist.

Beispiel 1.19. Seien $v, w \in \mathbb{R}^n$ und sei $\langle v, w \rangle = v_1 w_1 + \dots, v_n w_n$. Dann lautet die Cauchy-Schwarz'sche Ungleichung:

$$|\langle v, w \rangle| = \left| \sum_{k=1}^n v_k w_k \right| \leq \sqrt{\sum_{k=1}^n v_k^2} \cdot \sqrt{\sum_{k=1}^n w_k^2},$$

bzw. äquivalent dazu (durch Quadrieren)

$$\left(\sum_{k=1}^n v_k w_k \right)^2 \leq \left(\sum_{k=1}^n v_k^2 \right) \cdot \left(\sum_{k=1}^n w_k^2 \right).$$

Beispiel 1.20. Sei $V = \mathbb{R}^n$ mit der übliche Vektoraddition und skalaren Multiplikation und sei $\mathbb{K} = \mathbb{R}$. Sei $\|v\| := \sqrt{v_1^2 + \dots + v_n^2}$ die sog. Euklidische Norm. Diese ist in der Tat eine Norm:

1. Die Summe von Quadraten ist grundsätzlich nichtnegativ, so dass $\|v\| \geq 0$ gilt. Ist $0 = (0, \dots, 0)$ der n -dimensionale Nullvektor, dann ist $\|0\| = 0$. Andererseits gilt für $v \neq 0$, dass mindestens eine Koordinate $v_i \neq 0$ ist, damit also $v_i^2 > 0$ und folglich

$$\|v\| \geq \sqrt{v_i^2} = |v_i| > 0.$$

2. Seien $v = (v_1, \dots, v_n)$ und $a \in \mathbb{R}$. Dann ist

$$\|av\| = \sqrt{(av_1)^2 + \dots + (av_n)^2} = \sqrt{a^2 \cdot (v_1^2 + \dots + v_n^2)} = |a| \cdot \|v\|.$$

3. Seien $v = (v_1, \dots, v_n), w = (w_1, \dots, w_n) \in V$. Dann gilt (mit der 1. Binomischen Formel):

$$\begin{aligned} \|v + w\|^2 &= (v_1 + w_1)^2 + \dots + (v_n + w_n)^2 \\ &= (v_1^2 + \dots + v_n^2) + (w_1^2 + \dots + w_n^2) + 2(v_1 w_1 + \dots + v_n w_n) \\ &\leq \|v\|^2 + \|w\|^2 + 2|\langle v, w \rangle| \\ &\leq \|v\|^2 + \|w\|^2 + 2\|v\| \cdot \|w\| \\ &= (\|v\| + \|w\|)^2, \end{aligned}$$

wobei bei der letzten Ungleichung die Cauchy-Schwarz-Ungleichung benutzt wurde. Zieht man auf beiden Seiten die Wurzel, folgt daraus die behauptete Dreiecksungleichung.

Mit Hilfe des Skalarproduktes können wir den Begriff der Orthogonalität einführen:

Definition 1.21. 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 $v \parallel w$.

Beispiel 1.22. Im \mathbb{R}^2 mit dem Skalarprodukt $\langle v, w \rangle = v_1 w_1 + v_2 w_2$ 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 1.23. Ein Sonderfall in einem jeden Vektorraum ist der Nullvektor $\mathbf{0} \in V$, denn einerseits gilt $\langle \mathbf{0}, v \rangle = 0$ für jeden Vektor $v \in V$, d.h. $\mathbf{0}$ ist orthogonal zu jedem anderen Vektor. Zugleich gilt $\mathbf{0} = 0 \cdot v$ für jeden Vektor $v \in V$, d.h. $\mathbf{0}$ ist auch parallel zu jedem Vektor. Der Nullvektor $\mathbf{0}$ ist der einzige Vektor, der beide Eigenschaften hat.

Definition 1.24. Ist $U \subset V$ ein Untervektorraum eines \mathbb{K} -Vektorraumes V , so nennt man

$$U^\perp := \{v \in V \mid v \perp u, \quad \forall u \in U\}$$

den zu U orthogonalen Untervektorraum (oder das orthogonale Komplement).

Mit Worten kann man das orthogonale Komplement eines Untervektorraumes $U \subset V$ als die Menge der Vektoren des gesamten Vektorraumes V beschreiben, die zu allen Vektoren aus U orthogonal sind.

Beispiel 1.25. Seien $V = \mathbb{R}^3$ und $U = \{u = (u_1, u_2, u_3) \in V \mid u_2 = u_3 = 0\}$. Sei $v = (v_1, v_2, v_3) \in V$ beliebig. Es gilt $\langle u, v \rangle = u_1 v_1 = 0$ für alle $u \in U$ genau dann, wenn $v_1 = 0$ ist. Damit ist also

$$U^\perp = \{v = (v_1, v_2, v_3) \in V \mid v_1 = 0\}.$$

Es lässt sich leicht nachprüfen, dass das orthogonale Komplement des orthogonalen Komplements des Unterraumes wieder der Unterraum selbst ist, d.h. $(U^\perp)^\perp = U$.

Wir werden sehen, dass die Untersuchung von Orthogonalität nützlich ist für den Nachweis von linearer Unabhängigkeit – diesen Begriff sehen wir uns als nächstes an.

1.4 Lineare Unabhängigkeit und Erzeugendensystem

Definition 1.26. Sei V ein \mathbb{K} -Vektorraum. Für $r \in \mathbb{N}$ seien $v_1, \dots, v_r \in V$ Vektoren in V und seien zudem $a_1, \dots, a_r \in \mathbb{K}$ Skalare. Die endliche Summe

$$\sum_{k=1}^r a_k v_k = a_1 v_1 + \dots + a_r v_r$$

heißt dann Linearkombination der Vektoren v_1 bis v_r . Die Menge aller endlichen Linearkombinationen aus Vektoren einer Menge heißt lineare Hülle oder Span, d.h. ist $M \subset V$ eine (evtl. unendliche) Teilmenge von V , so ist

$$\mathcal{L}(M) := \left\{ \sum_{k=1}^r a_k v_k \mid r \in \mathbb{N}, \{v_1, \dots, v_r\} \subset M, \{a_1, \dots, a_r\} \subset \mathbb{K} \right\}.$$

Ist $M = \{v_1, \dots, v_n\} \subset V$ eine endliche Menge, dann ist

$$\mathcal{L}(M) = \left\{ \sum_{k=1}^n a_k v_k \mid v_1, \dots, v_n \in M, a_1, \dots, a_n \in \mathbb{K} \right\}.$$

Satz 1.27. Ist $U = \{v_1, \dots, v_n\} \subset V$, dann ist der Span $\mathcal{L}(U)$ ein Unterraum von V . Er ist zugleich der kleinste Unterraum von V , der U enthält.

Definition 1.28. Eine Teilmenge $U \subset V$ heißt Erzeugendensystem von V , wenn $\mathcal{L}(U) = V$ ist.

Beispiel 1.29. Folgende Mengen sind Erzeugendensysteme von \mathbb{R}^3 :

$$\begin{aligned} \mathcal{M}_1 &:= \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \\ \mathcal{M}_2 &:= \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \right\} = \mathcal{M}_1 \cup \left\{ \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \right\}, \\ \mathcal{M}_3 &:= \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix} \right\}, \\ \mathcal{M}_4 &:= \left\{ \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \right\}. \end{aligned}$$

Definition 1.30. Eine Menge $U \subset V$ heißt linear unabhängig, wenn für jede Linearkombination von Vektoren $v_1, \dots, v_r \in U$ gilt:

$$\sum_{k=1}^r a_k v_k = a_1 v_1 + \dots + a_r v_r = 0 \implies a_1 = \dots = a_r = 0.$$

Die Menge heißt linear abhängig, wenn sie nicht linear unabhängig ist.

Lemma 1.31. Eine Menge $U \subset V$ ist genau dann linear unabhängig, wenn kein Vektor $v \in U$ als Linearkombination der anderen Vektoren dargestellt werden kann.

Beispiel 1.32. Die Vektoren aus \mathcal{M}_1 aus Beispiel 1.29 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:

$$\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Satz 1.33. Sei V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Sei $U \subset V \setminus \{0\}$ eine Teilmenge von V mit paarweise orthogonalen (und von Null verschiedenen) Vektoren, d.h. es gilt $\langle v, w \rangle = 0$ für alle $v, w \in U$ mit $v \neq w$. Dann ist die Menge U linear unabhängig.

Beispiel 1.34. Die Vektoren der Menge \mathcal{M}_1 aus Beispiel 1.29 sind in der Tat paarweise orthogonal. Es gilt zum Beispiel

$$\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 0 = 0.$$

Die Vektoren der Menge \mathcal{M}_2 sind hingegen nicht alle paarweise orthogonal. Es gilt zum Beispiel

$$\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \right\rangle = 1 \cdot 1 + 0 \cdot 3 + 0 \cdot 5 = 1 \neq 0.$$

Allerdings kann aus fehlender Orthogonalität noch nicht geschlossen werden, dass die Menge nicht linear unabhängig sei. So ist beispielsweise \mathcal{M}_4 linear unabhängig. Allerdings sind die Vektoren nicht paarweise orthogonal zueinander:

$$\left\langle \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix} \right\rangle = 1 \cdot 0 + 4 \cdot (-1) + 6 \cdot 3 = 14 \neq 0.$$

1.5 Basis, Dimension, Koordinaten

Definition 1.35. Eine Menge $U \subset V$ heißt Basis von V , wenn U linear unabhängig und ein Erzeugendensystem von V ist.

Beispiel 1.36. \mathcal{M}_1 aus Beispiel 1.29 ist eine Basis des \mathbb{R}^3 . Weitere Basen sind \mathcal{M}_3 und \mathcal{M}_4 . 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.

Wir werden uns später noch mit dem systematischen Lösen von linearen Gleichungssystemen befassen, was die Untersuchung von \mathcal{M}_4 erleichtert.

Satz 1.37. Sind $B = \{b_1, \dots, b_m\}$ und $C = \{c_1, \dots, c_n\}$ zwei verschiedene Basen des Vektorraumes V , so gilt $m = n$, d.h. die Anzahl der Basiselemente ist gleich.

Dieser Satz rechtfertigt folgende Definition:

Definition 1.38. Besitzt ein Vektorraum V eine Basis $B = \{b_1, \dots, b_n\}$ 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 unendlich dimensional und man schreibt $\dim(V) = \infty$.

Beispiel 1.39. 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 ein 2-dimensionaler \mathbb{R} -Vektorraum mit der Basis $\{1, i\}$.
- Der Vektorraum $(C([0, 1]; \mathbb{R}), +, \cdot)$ der stetigen Funktionen von $[0, 1]$ nach \mathbb{R} mit den Operationen

$$\begin{aligned} (f + g)(x) &:= f(x) + g(x), & x \in [0, 1], \\ (a \cdot f)(x) &:= a \cdot f(x), & x \in [0, 1] \end{aligned}$$

ist ein unendlich-dimensionaler Vektorraum.

Satz 1.40. Sei $B = (b_1, \dots, b_n)$ eine geordnete Basis des \mathbb{K} -Vektorraumes V . Dann gibt es für jeden Vektor $v \in V$ genau ein Tupel $(a_1, \dots, a_n) \in \mathbb{K}^n$ mit der Eigenschaft dass

$$v = \sum_{k=1}^n a_k b_k.$$

Die Zahlen a_k heißen Koordinaten von v bezüglich der Basis B .

Bemerkung 1.41. Nach vorherigem Satz ist die Abbildung $\kappa: V \rightarrow \mathbb{K}^n$, die jedem Vektor $v \in V$ seine Koordinaten $(a_1, \dots, a_n) \in \mathbb{K}^n$ zuordnet, eine Bijektion.

Beispiel 1.42. Betrachten wir den Vektor $v := (1, 3, 5) \in \mathbb{R}^3$ aus Beispiel 1.29. Bezüglich der Basis \mathcal{M}_1 hat v die Koordinaten $(1, 3, 5)$. Bezüglich der Basis \mathcal{M}_3 hingegen hat v die Koordinaten $(\frac{1}{2}, \frac{3}{4}, \frac{5}{7})$.

Das Berechnen der Koordinaten ist nicht immer einfach. Es gibt jedoch spezielle Basen, für die die Koordinaten ohne Lösen eines Gleichungssystems berechnet werden können.

Definition 1.43. Sei V ein euklidischer oder unitärer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Sei weiterhin $B \subset V \setminus \{0\}$ eine Teilmenge paarweise orthogonaler Vektoren. Dann nennt man B Orthogonalsystem. Gilt zudem $\|v\| = \sqrt{\langle v, v \rangle} = 1$ für alle Vektoren $v \in B$, so heißt B Orthonormalsystem. Ist B eine Basis von V , so spricht man von einer Orthogonal- bzw. Orthonormalbasis (kurz ONB).

Beispiel 1.44. \mathcal{M}_1 aus 1.29 ist eine ONB (man nennt sie auch die kanonische Basis des \mathbb{R}^3), \mathcal{M}_3 ist hingegen nur eine Orthogonalbasis, da die Vektoren nicht Einheitslänge haben. \mathcal{M}_4 ist zwar eine Basis, jedoch weder eine Orthonormal- noch eine Orthogonalbasis.

Handelt es sich bei der Basis um eine ONB, so lassen sich die Koordinaten eines Vektors bezüglich der Basis wie folgt berechnen:

Satz 1.45. Sei $B = \{b_1, \dots, b_n\}$ eine ONB des Vektorraumes V und sei $v \in V$ beliebig. Dann gilt

$$v = \sum_{k=1}^n \langle v, b_k \rangle b_k.$$

Beispiel 1.46. Betrachten wir wieder \mathcal{M}_1 aus 1.29, wobei wir die Basisvektoren der Reihe nach als b_1, b_2 und b_3 bezeichnen. Sei zudem der Vektor $v = \begin{pmatrix} 7 \\ 3 \\ 9 \end{pmatrix}$ gegeben. Dann gilt in der Tat bezüglich des kanonischen Skalarproduktes:

$$\begin{aligned} \sum_{k=1}^n \langle v, b_k \rangle b_k &= \left\langle \begin{pmatrix} 7 \\ 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} 7 \\ 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} 7 \\ 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= 7 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 9 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 7 \\ 3 \\ 9 \end{pmatrix} = v. \end{aligned}$$

1.6 Orthogonale Projektion

Wir wissen bereits, dass man mit Hilfe von Normen Längen von Vektoren definieren kann. Ebenso kann man damit auch den Abstand zwischen einem Punkt im Raum (gegeben durch einen Vektor) und einem Untervektorraum definieren.

Definition 1.47. Sei V ein (eukl. oder unitärer) Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Sei $U \subset V$ eine (nichtleere) Teilmenge von V . Zu jedem Vektor $v \in V$ definieren wir den Abstand von v zu U als

$$d(v, U) := \inf \{ \|v - u\| \mid u \in U \}.$$

Satz 1.48. Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Sei $U \subset V$ ein Untervektorraum von V . Zu jedem Vektor $v \in V$ existiert genau ein Vektor $v^* \in U$, so dass der Abstand $d(v, U)$ minimal wird, d.h. $d(v, U) = \min \{ \|v - u\| \mid u \in U \} = \|v - v^*\|$. Die Abbildung

$$P_U: V \rightarrow U, \quad P_U(v) = v^*$$

nennt man orthogonale Projektion von V auf U .

Für die orthogonale Projektion gilt:

- $P_U(v) \in U$ für alle $v \in V$;
- $P_U(v) = v$ genau dann, wenn $v \in U$ ist;
- $P_U \circ P_U = P_U$, d.h. für alle $v \in V$ ist $P_U(P_U(v)) = P_U(v)$;
- für alle $v \in V$ ist $v - P_U(v) \in U^\perp$ und $\|v - P_U(v)\| = d(v, U)$, d.h. $P_U(v)$ ist der Punkt aus U , dessen Abstand zu v minimal ist unter allen Punkten aus U ;
- $P_U(v) = 0$ genau dann, wenn $v \in U^\perp$ ist ($\|v - u\| = \sqrt{\langle v - u, v - u \rangle}$).

Ist $B = \{b_1, \dots, b_m\}$ eine ONB des Untervektorraumes $U \subset V$, dann gilt für alle $u \in U$ die Darstellung

$$u = \sum_{k=1}^m \langle u, b_k \rangle b_k.$$

Speziell für die orthogonale Projektion gilt:

Satz 1.49. Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Sei $U \subset V$ ein Untervektorraum mit ONB $B = \{b_1, \dots, b_m\}$. Dann gilt für alle $v \in V$:

$$P_U(v) = \sum_{k=1}^m \langle v, b_k \rangle b_k.$$

Beispiel 1.50. Sei $V = \mathbb{R}^3$ und $U = \{v = (v_1, v_2, v_3) \in V \mid v_3 = 0\}$. Eine ONB von U ist

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Beispielhaft sehen wir uns den Vektor $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in V$ an. Die orthogonale Projektion von v auf den

Untervektorraum U ist gegeben durch

$$P_U(v) = \sum_{k=1}^2 \langle v, b_k \rangle b_k = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

1.7 Gram-Schmidt'sches Orthonormalisierungsverfahren

Kennt man eine beliebige Basis $B = \{b_1, \dots, b_n\}$ eines Vektorraumes V , so ist es möglich, mit folgendem Algorithmus eine ONB zu konstruieren:

1. Schritt: Setze $e_1 := \frac{b_1}{\|b_1\|}$. (Damit ist $\|e_1\| = 1$.)

2. Schritt (1): Setze $e_2^* := b_2 - \langle e_1, b_2 \rangle e_1$. (Damit ist $\langle e_2^*, e_1 \rangle = 0$.)

2. Schritt (2): Setze $e_2 := \frac{e_2^*}{\|e_2^*\|}$. (Damit ist $\|e_2\| = 1$.)

3. Schritt (1): Setze $e_3^* := b_3 - \langle e_1, b_3 \rangle e_1 - \langle e_2, b_3 \rangle e_2$. (Damit ist e_3^* orthogonal zu e_1, e_2 .)

3. Schritt (2): Setze $e_3 := \frac{e_3^*}{\|e_3^*\|}$. (Damit ist $\|e_3\| = 1$.)

⋮

n. Schritt (1): Setze $e_n^* := b_n - \sum_{k=1}^{n-1} \langle e_k, b_n \rangle e_k$. (Damit ist e_n^* orthogonal zu e_1, \dots, e_{n-1} .)

n. Schritt (2): Setze $e_n := \frac{e_n^*}{\|e_n^*\|}$. (Damit ist $\|e_n\| = 1$.)

Beispiel 1.51. Wir haben gesehen, dass \mathcal{M}_3 aus Beispiel 1.29 eine Basis des \mathbb{R}^3 ist. Wir wenden nun das obige Verfahren an, um daraus eine ONB zu erhalten.

1. Schritt:

$$e_1 := \frac{b_1}{\|b_1\|} = \frac{1}{\left\| \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\|} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2^2}} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

2. Schritt (1):

$$e_2^* := b_2 - \langle e_1, b_2 \rangle e_1 = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} - 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}$$

2. Schritt (2):

$$e_2 := \frac{e_2^*}{\|e_2^*\|} = \frac{1}{\sqrt{4^2}} \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

3. Schritt (1):

$$e_3^* := b_3 - \langle e_1, b_3 \rangle e_1 - \langle e_2, b_3 \rangle e_2 = \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix}$$

3. Schritt (2):

$$e_3 := \frac{e_3^*}{\|e_3^*\|} = \frac{1}{\sqrt{7^2}} \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Das Verfahren fassen wir in folgendem Satz zusammen:

Satz 1.52. Seien b_1, \dots, b_n linear unabhängige Vektoren des (eukl. oder unitären) Vektorraumes V . Dann bilden die Vektoren

$$e_i := \frac{b_i - \sum_{k=1}^{i-1} \langle b_i, e_k \rangle e_k}{\left\| b_i - \sum_{k=1}^{i-1} \langle b_i, e_k \rangle e_k \right\|}, \quad i = 1, \dots, n$$

ein Orthonormalsystem, wobei $\mathcal{L}(b_1, \dots, b_n) = \mathcal{L}(e_1, \dots, e_n)$ ist.

Ist $\{b_1, \dots, b_n\}$ bereits eine Basis, so ist $\{e_1, \dots, e_n\}$ eine ONB von V .

2 Lineare Abbildungen und Matrizen

2.1 Lineare Abbildungen

Im Folgenden sei \mathbb{K} ein Körper und V und W seien \mathbb{K} -Vektorräume.

Definition 2.1. Eine Abbildung $f: V \rightarrow W$ heißt lineare Abbildung, falls gilt:

1. $f(x + y) = f(x) + f(y)$ für alle $x, y \in V$,
2. $f(\lambda x) = \lambda f(x)$ für alle $x \in V$ und $\lambda \in \mathbb{K}$.

Alternativ hat man folgendes Kriterium:

Lemma 2.2. Eine Abbildung $f: V \rightarrow W$ ist genau dann linear, wenn

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y), \quad \text{für alle } x, y \in V \text{ und } \lambda, \mu \in \mathbb{K}.$$

Beispiel 2.3. Sehen wir uns Beispiele für lineare und nicht lineare Funktionen an.

i) Die Abbildung

$$\begin{cases} f: \mathbb{R} \rightarrow \mathbb{R}, \\ x \mapsto f(x) = ax + b \end{cases} \quad (2.1)$$

mit $a, b \in \mathbb{R}$ ist genau dann eine lineare Abbildung, wenn $b = 0$ ist. (Andernfalls heißt f affin.)
Um das zu sehen, reicht es, das erste Kriterium zu betrachten:

$$f(x + y) = a(x + y) + b = ax + ay + b, \quad (2.2)$$

$$f(x) + f(y) = (ax + b) + (ay + b) = ax + ay + 2b. \quad (2.3)$$

ii) Sei 0_W der Nullvektor des Vektorraumes W . Die Abbildung

$$\begin{cases} f: V \rightarrow W, \\ x \mapsto f(x) = 0_W \end{cases}$$

ist linear, denn für beliebige $x, y \in V$ und $\lambda, \mu \in \mathbb{K}$ ist

$$\begin{aligned} f(\lambda x + \mu y) &= 0_W, \\ \lambda f(x) + \mu f(y) &= \lambda \cdot 0_W + \mu \cdot 0_W = 0_W. \end{aligned}$$

iii) Folgende Funktion ist ebenfalls linear:

$$\begin{cases} f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x + 3y \\ ex + \pi y \end{pmatrix}. \end{cases}$$

iv) Folgende Funktion ist nicht linear:

$$\begin{cases} f: \mathbb{R} \rightarrow \mathbb{R}, \\ x \mapsto f(x) = x^2 \end{cases}$$

In der Tat ist $f(x + y) = (x + y)^2 = x^2 + 2xy + y^2 \neq x^2 + y^2 = f(x) + f(y)$ für $x, y \in \mathbb{R} \setminus \{0\}$.

Definition 2.4. Sei $f: V \rightarrow W$ eine lineare Abbildung und sei 0_W der Nullvektor in W . Dann ist

$$\text{im}(f) := f(V) = \{f(v) \mid v \in V\} \subset W$$

das Bild (englisch image) von f und

$$\text{ker}(f) := f^{-1}(0_W) = \{v \in V \mid f(v) = 0_W\}$$

der Kern (englisch kernel) von f .

Beispiel 2.5. Sei

$$U = \left\{ u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \in \mathbb{R}^3 \mid u_3 = 0 \right\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}.$$

Seien weiterhin

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Die Menge $\{e_1, e_2\}$ ist eine ONB von U , daher ist die orthogonale Projektion von $v = (v_1, v_2, v_3) \in V = \mathbb{R}^3$ gegeben durch

$$P_U(v) = \sum_{k=1}^2 \langle v, e_k \rangle e_k = \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix}.$$

Zunächst einmal ist $P_U: V \rightarrow U$ eine lineare Abbildung. Seien dazu $x = (x_1, x_2, x_3), y = (y_1, y_2, y_3) \in V$ und $\lambda, \mu \in \mathbb{R}$. Dann ist

$$P_U(\lambda x + \mu y) = \begin{pmatrix} \lambda x_1 + \mu y_1 \\ \lambda x_2 + \mu y_2 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix} = \lambda P_U(x) + \mu P_U(y).$$

Berechnen wir Bild und Kern der orthogonalen Projektion P_U :

$$\text{im}(P_U) = \{u \in U \mid \exists v \in V: u = P_U(v)\} = U,$$

$$\text{ker}(P_U) = \left\{ v \in V \mid P_U(v) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3 \mid v_1 = v_2 = 0 \right\} = \left\{ a \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mid a \in \mathbb{R} \right\}.$$

Lemma 2.6. Sei $f: V \rightarrow W$ eine lineare Abbildung zwischen zwei \mathbb{K} -Vektorräumen V und W . Dann gilt:

- i) $\text{ker}(f) \subseteq V$ ist ein Untervektorraum von V ,
- ii) $\text{im}(f) \subseteq W$ ist ein Untervektorraum von W .

Beweis.

- i) Seien $x, y \in \text{ker}(f)$ und $\lambda, \mu \in \mathbb{K}$. Dann ist

$$f(\lambda x + \mu y) \stackrel{f \text{ linear}}{=} \lambda f(x) + \mu f(y) = \lambda \cdot 0_W + \mu \cdot 0_W = 0_W,$$

d.h. in der Tat ist $\lambda x + \mu y \in \text{ker}(f)$.

ii) Seien $w_1, w_2 \in \text{im}(f)$, d.h. es existieren $v_1, v_2 \in V$ mit $f(v_1) = w_1$ und $f(v_2) = w_2$. Seien zudem $\lambda, \mu \in \mathbb{K}$. Dann ist

$$f(\lambda v_1 + \mu v_2) = \lambda f(v_1) + \mu f(v_2) = \lambda w_1 + \mu w_2,$$

d.h. $\lambda w_1 + \mu w_2 \in \text{im}(f)$.

□

Es bezeichne jetzt wieder $\dim(V)$ die Dimension eines Vektorraumes V .

Definition 2.7. Seien V und W endlich-dimensionale Vektorräume und $f: V \rightarrow W$ eine lineare Abbildung. Der Rang von f ist definiert als die Dimension des Bildes von f :

$$\text{rg}(f) := \dim(\text{im}(f)).$$

Eine kurze Erinnerung: Eine Funktion $f: A \rightarrow B$ heißt

- injektiv, falls aus $f(x) = f(y)$ folgt, dass $x = y$ sein muss,
- surjektiv, falls zu jedem $b \in B$ (mind.) ein $a \in A$ existiert mit $f(a) = b$,
- bijektiv, falls f sowohl injektiv als auch surjektiv ist.

Satz 2.8. Sei $f: V \rightarrow W$ eine lineare Abbildung zwischen zwei endlich-dimensionalen Vektorräumen. Dann gilt:

i) f ist surjektiv genau dann, wenn $\text{rg}(f) = \dim(W)$ ist.

ii) f ist injektiv genau dann, wenn $\dim(\ker(f)) = 0$ ist.

Allgemein gilt die Formel

$$\dim(V) = \dim(\ker(f)) + \text{rg}(f).$$

Daraus folgt insbesondere für lineare Funktionen $f: V \rightarrow W$ mit $\dim(V) = \dim(W)$:

$$f \text{ ist injektiv} \iff f \text{ ist surjektiv} \iff f \text{ ist bijektiv.}$$

Beispiel 2.9. Betrachten wir nochmal die orthogonale Projektion aus Beispiel 2.5. Es ist $\text{im}(P_U) = U$, also gilt zwangsläufig $\text{rg}(P_U) = \dim(\text{im}(P_U)) = \dim(U)$, d.h. P_U ist surjektiv als Funktion von V nach U . Der Kern von P_U ist

$$\ker(P_U) = \left\{ \begin{pmatrix} 0 \\ 0 \\ v_3 \end{pmatrix} \mid v_3 \in \mathbb{R} \right\},$$

also ist $\dim(\ker(P_U)) = 1$. Die orthogonale Projektion ist also nicht injektiv. Dies können wir auch wie folgt sehen:

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, y = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \text{ aber } P_U(x) = P_U(y) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Definition 2.10. Ist eine lineare Abbildung $f: V \rightarrow W$ bijektiv, so sagt man, f sei ein Isomorphismus. Die Vektorräume V und W heißen dann zueinander isomorph.

Satz 2.11. Sei V ein beliebiger n -dimensionaler \mathbb{K} -Vektorraum. Dann ist V isomorph zum Vektorraum \mathbb{K}^n .

Beweis. Dazu wählen wir eine Basis $B = \{b_1, \dots, b_n\} \subset V$ von V . Jeder Vektor $v \in V$ besitzt dann eine eindeutige Darstellung der Form

$$v = \sum_{k=1}^n a_k b_k.$$

Damit sind die Koordinaten von v bezüglich B gegeben durch $(a_1, \dots, a_n) \in \mathbb{K}^n$. Die Koordinatenabbildung $\kappa: V \rightarrow \mathbb{K}^n$ mit $\kappa(v) = (a_1, \dots, a_n)$ ist, wie man leicht nachprüfen kann, linear. Gleichzeitig ist κ eine Bijektion (Bemerkung 1.41). Damit ist die Koordinatenabbildung κ ein Isomorphismus. \square

Zuletzt bemerken wir noch, dass eine lineare Abbildung durch endlich viele Werte bereits eindeutig bestimmt ist:

Satz 2.12. Sei V ein (n -dimensionaler) Vektorraum mit Basis $B = \{b_1, \dots, b_n\}$. Seien $w_1, \dots, w_n \in W$. Dann gibt es genau eine lineare Funktion $f: V \rightarrow W$ mit der Eigenschaft, dass $f(b_k) = w_k$ ist für $k = 1, \dots, n$.

Beweis. Ist $v = \sum_{k=1}^n a_k b_k$, d.h. sind die Koordinaten von v bezüglich B gegeben durch (a_1, \dots, a_n) , dann ist

$$f(v) = f\left(\sum_{k=1}^n a_k b_k\right) \stackrel{(*)}{=} \sum_{k=1}^n a_k f(b_k) = \sum_{k=1}^n a_k w_k,$$

wobei $(*)$ gelten muss, da die Funktion linear sein soll. Da die Koordinaten (a_1, \dots, a_n) eines jeden Vektors v aus V bezüglich einer gegebenen Basis eindeutig sind, ist somit der Funktionswert des Vektors eindeutig durch $\sum_{k=1}^n a_k w_k$ bestimmt. Man kann zeigen, dass diese Abbildung tatsächlich linear ist. (Siehe Satz 7.15 im Skript Lineare Algebra I) \square

Beispiel 2.13. Seien $B = (b_1, b_2, b_3)$ die (geordnete) Standard-Basis von $V = \mathbb{R}^3$ und $C = (c_1, c_2)$ die (geordnete) Standard-Basis von $W = \mathbb{R}^2$, d.h.

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad c_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad c_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Wir wollen eine lineare Abbildung $f: V \rightarrow W$ konstruieren, wobei gelten soll:

$$f(b_1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad f(b_2) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \quad f(b_3) = \begin{pmatrix} 7 \\ 6 \end{pmatrix}.$$

Damit müssen folgende Bedingungen erfüllt sein:

$$\begin{aligned} f(b_1) &= 1 \cdot c_1 + 2 \cdot c_2, \\ f(b_2) &= 3 \cdot c_1 + 5 \cdot c_2, \\ f(b_3) &= 7 \cdot c_1 + 6 \cdot c_2. \end{aligned}$$

Die lineare Abbildung f ist also durch das Zahlenschema

$$A = \begin{pmatrix} 1 & 3 & 7 \\ 2 & 5 & 6 \end{pmatrix}$$

vollständig definiert. Ein solches Zahlenschema nennt man eine Matrix.

2.2 Matrizen

Seien V und W endlich-dim. \mathbb{K} -Vektorräume mit geordneten Basen $B = (b_1, \dots, b_n)$ (von V) und $C = (c_1, \dots, c_m)$ (von W). Damit ist also insbesondere $\dim(V) = n$ und $\dim(W) = m$. Ist $f: V \rightarrow W$ eine lineare Abbildung, dann gibt es eindeutig bestimmte Skalare $\alpha_{k,j}$ ($k = 1, \dots, m$ und $j = 1, \dots, n$) so dass

$$f(b_j) = \sum_{k=1}^m \alpha_{k,j} c_k, \quad j \in \{1, \dots, n\}.$$

Wir ordnen diese in einem rechteckigen Schema wie folgt an:

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

Man spricht in diesem Fall von einer $m \times n$ -Matrix. Die Menge der $m \times n$ -Matrizen mit Einträgen aus dem Körper \mathbb{K} wird mit $\mathbb{K}^{m \times n}$ notiert. Wenn der Körper bekannt ist, schreibt man auch $\text{Mat}(m, n)$.

Definition 2.14. A heißt die der linearen Abbildung $f: V \rightarrow W$ zugeordnete $m \times n$ -Matrix bezüglich der Basen B und C .

$$(\alpha_{k,1} \quad \cdots \quad \alpha_{k,n})$$

heißt k -ter Zeilenvektor von A ,

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

heißt j -ter Spaltenvektor von A .

Beispiel 2.15.

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

Es ist $\alpha_{1,2} = 2$ und $\alpha_{2,1} = 3$. (Merksatz für die Reihenfolge: „Zeile zuerst, Spalte später“)

Definition 2.16.

- (i) Eine Matrix $A \in \mathbb{K}^{m \times n}$ heißt quadratisch, falls $m = n$ ist.
- (ii) Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt symmetrisch, falls $\alpha_{k,j} = \alpha_{j,k}$ ist für alle Indizes k, j .
- (iii) Eine Diagonalmatrix $A \in \mathbb{K}^{n \times n}$ ist von der Form

$$A = \begin{pmatrix} \alpha_{1,1} & & 0 \\ & \ddots & \\ 0 & & \alpha_{n,n} \end{pmatrix} =: \text{diag}(\alpha_{1,1}, \dots, \alpha_{n,n}),$$

d.h. $\alpha_{k,j} = 0$ für alle $k \neq j$.

Beispiel 2.17. Eine $(n \times n)$ -Diagonalmatrix, in deren (Haupt-)Diagonale alle Einträge 1 sind, heißt Einheitsmatrix, d.h.

$$E_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathbb{K}^{n \times n}$$

2.2.1 Rechnen mit Matrizen

Zunächst einmal definieren wir die Addition von Matrizen und die skalare Multiplikation. Dazu 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, $A, B \in \mathbb{K}^{m \times n}$. Sei zudem $c \in \mathbb{K}$. Dann gilt:

Addition: $A + B := C = (\gamma_{k,j})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}}$ mit $\gamma_{k,j} := (\alpha_{k,j} + \beta_{k,j})$, d.h.

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

Skalare Multiplikation: $cA := (c\alpha_{k,j})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}}$, d.h.

$$cA = c \cdot \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} = \begin{pmatrix} c\alpha_{1,1} & \cdots & c\alpha_{1,n} \\ \vdots & & \vdots \\ c\alpha_{m,1} & \cdots & c\alpha_{m,n} \end{pmatrix}$$

Beispiel 2.18. Seien

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

Dann gilt:

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

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

Satz 2.19. Sei \mathbb{K} ein Körper und seien $m, n \in \mathbb{N}$. Die Menge der Matrizen $\mathbb{K}^{m \times n}$ mit den Operationen Addition und skalarer Multiplikation ist ein $m \cdot n$ -dimensionaler \mathbb{K} -Vektorraum.

Insbesondere gelten damit die üblichen Rechengesetze, z.B. $A + B = B + A$ (Kommutativität) und $c(A + B) = cA + cB$ (Distributivgesetz).

Das neutrale Element der Addition ist die Nullmatrix und ist $A = (\alpha_{k,j})$, dann ist $-A = (-\alpha_{k,j})$ die bezüglich der Addition inverse Matrix zu A , d.h. $A + (-A) = 0$ (Nullmatrix).

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

Definition 2.20. Seien $A = (\alpha_{k,j})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m \times n}$ und $B = (\beta_{k,j})_{\substack{1 \leq k \leq n \\ 1 \leq j \leq r}} \in \mathbb{K}^{n \times r}$. Dann ist $A \cdot B = C = (\gamma_{k,j})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq r}} \in \mathbb{K}^{m \times r}$ gegeben durch

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

d.h.

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

Beispiel 2.21. Im Allgemeinen gilt $A \cdot B \neq B \cdot A$:

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

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

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

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

Es gelten folgende Rechengesetze:

Satz 2.22. Seien $A \in \mathbb{K}^{m \times n}$, $B, B' \in \mathbb{K}^{n \times r}$ und $C \in \mathbb{K}^{r \times s}$ beliebige Matrizen und sei $E_d \in \mathbb{K}^{d \times d}$ die $d \times d$ -Einheitsmatrix für beliebiges $d \in \mathbb{N}$. Dann gilt:

(i) $E_m \cdot A = A \cdot E_n = A$,

(ii) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$,

(iii) $A \cdot (B + B') = A \cdot B + A \cdot B'$ und $(B + B') \cdot C = B \cdot C + B' \cdot C$.

Der erste Teil des Satzes besagt, dass die Einheitsmatrix das neutrale Element der Multiplikation ist, sofern wir uns auf quadratische Matrizen beschränken. (Andernfalls gibt es ein links- und ein rechts-neutrales Element, da sich die Dimensionen unterscheiden!) Aus diesem Grund ist es nur für quadratische Matrizen sinnvoll, von einer (multiplikativ) inversen Matrix zu sprechen:

Definition 2.23. Sei $A \in \mathbb{K}^{n \times n}$. Eine Matrix $B \in \mathbb{K}^{n \times n}$ heißt inverse Matrix von A , falls $A \cdot B = B \cdot A = E_n$ gilt. Wir bezeichnen Sie mit $B = A^{-1}$.

Bevor wir uns darüber Gedanken machen können, ob eine Matrix eine Inverse besitzt und wie man diese berechnet, brauchen wir noch etwas mehr Handwerkszeug.

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

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

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

Um zwischen beiden Notationen wechseln zu können, führen wir folgenden Begriff ein:

Definition 2.24. Sei $A = (\alpha_{k,j})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m \times n}$. Dann ist die zu A transponierte Matrix $A^T = (\alpha_{j,k})_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}} \in \mathbb{K}^{n \times m}$, 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 2.25.

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

Es gelten folgende Rechenregeln:

Satz 2.26. Seien $A, B \in \mathbb{K}^{m \times n}$ und $c \in \mathbb{K}$.

- (i) $(A + B)^T = A^T + B^T$,
- (ii) $(cA)^T = cA^T$,
- (iii) $(A^T)^T = A$.

Definition 2.27. Sei $A \in \mathbb{R}^{n \times n}$.

- (i) A heißt symmetrisch, falls $A = A^T$ ist.
- (ii) A heißt orthogonal, falls $A \cdot A^T = E_n$ ist.

Beispiel 2.28. Die Matrix $A = \begin{pmatrix} 2 & -5 \\ -5 & 3 \end{pmatrix}$ ist symmetrisch, die Matrix $B = \begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix}$ ist orthogonal, denn

$$A^T = \begin{pmatrix} 2 & -5 \\ -5 & 3 \end{pmatrix} = A \quad \text{und} \quad BB^T = \frac{1}{5^2} \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2.$$

Lemma 2.29. Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist genau dann orthogonal, wenn ihre Spalten eine ONB des \mathbb{R}^n bilden bezüglich des kanonischen Skalarproduktes.

Komplexwertige Matrizen

Für komplexwertige Matrizen gibt es Begriffe, die analog sind zu denen von symmetrischen und orthogonalen reellwertigen Matrizen.

Definition 2.30. Sei $A = (\alpha_{k,j})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}} \in \mathbb{C}^{m \times n}$.

- (i) $A^* = (\alpha_{j,k}^*)_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}}$ mit $\alpha_{k,j}^* := \overline{\alpha_{k,j}}$ heißt die zu A adjungierte Matrix.
- (ii) Ist $A \in \mathbb{C}^{n \times n}$ und gilt $A = A^*$, so heißt die Matrix A hermitesch.
- (iii) Ist $A \in \mathbb{C}^{n \times n}$ und gilt $A \cdot A^* = E_n$, so heißt A unitär.

2.3 Matrizen und lineare Abbildungen

Satz 2.31. Sei $A \in \mathbb{K}^{m \times n}$. Die Abbildung $f_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ mit $f_A(x) = A \cdot x$ (für $x \in \mathbb{K}^n$) ist linear.

Beweis. Seien $c \in \mathbb{K}$,

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Dann ist

$$f_A(cx) = A \cdot (cx) = \begin{pmatrix} \alpha_{1,1}cx_1 + \dots + \alpha_{1,n}cx_n \\ \vdots \\ \alpha_{m,1}cx_1 + \dots + \alpha_{m,n}cx_n \end{pmatrix} = c \begin{pmatrix} \alpha_{1,1}x_1 + \dots + \alpha_{1,n}x_n \\ \vdots \\ \alpha_{m,1}x_1 + \dots + \alpha_{m,n}x_n \end{pmatrix} = cf_A(x).$$

Des Weiteren ist (wie in Satz 2.22 zu den Rechengesetzen behauptet)

$$\begin{aligned} f_A(x+y) &= A \cdot (x+y) = \begin{pmatrix} \alpha_{1,1}(x_1+y_1) + \dots + \alpha_{1,n}(x_n+y_n) \\ \vdots \\ \alpha_{m,1}(x_1+y_1) + \dots + \alpha_{m,n}(x_n+y_n) \end{pmatrix} \\ &= \begin{pmatrix} (\alpha_{1,1}x_1 + \dots + \alpha_{1,n}x_n) + (\alpha_{1,1}y_1 + \dots + \alpha_{1,n}y_n) \\ \vdots \\ (\alpha_{m,1}x_1 + \dots + \alpha_{m,n}x_n) + (\alpha_{m,1}y_1 + \dots + \alpha_{m,n}y_n) \end{pmatrix} \\ &= A \cdot x + A \cdot y = f_A(x) + f_A(y). \end{aligned}$$

Folglich ist f_A in der Tat linear. □

Die Verknüpfung linearer Abbildung lässt sich übersetzen in die Multiplikation der zugehörigen Matrizen:

Satz 2.32. Seien $A \in \mathbb{K}^{m \times n}$ und $B \in \mathbb{K}^{n \times r}$. Den Matrizen zugeordnet seien die linearen Abbildungen $f_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ und $f_B: \mathbb{K}^r \rightarrow \mathbb{K}^n$. Dann ist die Verkettung $f_A \circ f_B: \mathbb{K}^r \rightarrow \mathbb{K}^m$ gegeben durch $f_A \circ f_B = f_{A \cdot B}$, d.h. für $x \in \mathbb{K}^r$ ist

$$(f_A \circ f_B)(x) = f_A(f_B(x)) = f_A(B \cdot x) = A \cdot (B \cdot x) = (A \cdot B) \cdot x.$$

Beispiel 2.33. Seien $f_B: \mathbb{R} \rightarrow \mathbb{R}^2$ und $f_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ wie folgt gegeben:

$$\begin{aligned} f_B(x) &= \begin{pmatrix} 2x \\ 3x \end{pmatrix}, \quad f_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+y \\ 3y \end{pmatrix}, \quad \text{d.h.} \\ B &= \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}. \end{aligned}$$

Einerseits ist

$$(f_A \circ f_B)(x) = f_A(f_B(x)) = f_A \begin{pmatrix} 2x \\ 3x \end{pmatrix} = \begin{pmatrix} 2 \cdot 2x + 3x \\ 3 \cdot 3x \end{pmatrix} = \begin{pmatrix} 7x \\ 9x \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \end{pmatrix} \cdot x.$$

Andererseits ist in der Tat

$$A \cdot B = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4+3 \\ 0+9 \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \end{pmatrix}.$$

Wir werden den Zusammenhang zwischen linearen Abbildungen und Matrizen im Folgenden benutzen, um Begriffe, die wir bereits für lineare Abbildungen eingeführt haben, auf Matrizen zu übertragen.

Definition 2.34. Die maximale Anzahl linear unabhängiger Spaltenvektoren einer Matrix A heißt Spaltenrang von A und wir schreiben $\text{rg}_S(A)$. Die maximale Anzahl linear unabhängiger Zeilenvektoren einer Matrix A heißt Zeilenrang von A und wir schreiben $\text{rg}_Z(A)$.

Satz 2.35. Seien V, W endlich dimensionale \mathbb{K} -Vektorräume mit geordneten Basen $B = (b_1, \dots, b_n)$ (von V) und $C = (c_1, \dots, c_m)$ (von W). Sei $f: V \rightarrow W$ eine lineare Abbildung mit zugehöriger Matrix $A = (\alpha_{k,j})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m \times n}$, d.h. $f(b_j) = \sum_{k=1}^m \alpha_{k,j} c_k$. Dann gilt:

$$\text{rg}(f) = \text{rg}_S(A) = \text{rg}_Z(A).$$

Definition 2.36. Der Rang einer Matrix A ist definiert als $\text{rg}(A) := \text{rg}_S(A) = \text{rg}_Z(A)$. Eine quadratische Matrix $A \in \mathbb{K}^{n \times n}$ heißt regulär, falls sie maximalen Rang hat, d.h., falls $\text{rg}(A) = n$ ist. Andernfalls heißt A singular.

Beispiel 2.37. Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 7 & 6 \end{pmatrix}$$

hat zwei linear unabhängige Spaltenvektoren, damit gilt $\text{rg}_S(A) = 2$. Der Rang der Matrix A ist somit $\text{rg}(A) = 2$. Wir können auch prüfen, dass tatsächlich der dritte Zeilenvektor von den anderen beiden abhängig ist, d.h. dass auch $\text{rg}_Z(A) = 2$ ist: Das lineare Gleichungssystem

$$\begin{aligned} 1 \cdot a + 3 \cdot b &= 7 \\ 2 \cdot a + 5 \cdot b &= 6 \end{aligned}$$

hat die Lösung $(a, b) = (-17, 8)$, d.h.

$$(-17) \cdot \begin{pmatrix} 1 & 2 \end{pmatrix} + 8 \cdot \begin{pmatrix} 3 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 6 \end{pmatrix}.$$

Satz 2.38. Eine Matrix $A \in \mathbb{K}^{n \times n}$ ist genau dann regulär, wenn die zugehörige lineare Abbildung f_A bijektiv ist, d.h. wenn $f_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ ein Isomorphismus ist. In diesem Fall besitzt A eine Inverse A^{-1} und es gilt $f_{A^{-1}} = (f_A)^{-1}$.

Beispiel 2.39. Eine Diagonalmatrix

$$A = \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

ist genau dann regulär, wenn auf der Diagonalen keine Null steht, d.h. $\lambda_1, \dots, \lambda_n \in \mathbb{K} \setminus \{0\}$. In diesem Fall ist die Inverse von A gegeben durch

$$A^{-1} = \text{diag}(\lambda_1^{-1}, \dots, \lambda_n^{-1}) = \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix}.$$

Beispiel 2.40. Sei $f_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch die Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R}.$$

Falls $ad - bc \neq 0$ ist, ist A invertierbar (dies werden wir später noch herleiten) und die Inverse lautet

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

Überzeugen wir uns zunächst, dass die Inverse korrekt ist:

$$A \cdot A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} ad - bc & -ab + ab \\ cd - cd & -bc + ad \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2.$$

Betrachten wir nun die zugehörigen Funktionen:

$$f_A(x, y) = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix},$$

$$f_{A^{-1}}(x, y) = \frac{1}{ad - bc} \begin{pmatrix} dx - by \\ -cx + ay \end{pmatrix}.$$

Verknüpfen wir beide Funktionen, so gilt in der Tat

$$\begin{aligned} (f_A \circ f_{A^{-1}})(x, y) &= f_A \left(\frac{1}{ad - bc} \begin{pmatrix} dx - by \\ -cx + ay \end{pmatrix} \right) \\ &= \frac{1}{ad - bc} \cdot f_A \begin{pmatrix} dx - by \\ -cx + ay \end{pmatrix} \\ &= \frac{1}{ad - bc} \cdot \begin{pmatrix} a(dx - by) + b(-cx + ay) \\ c(dx - by) + d(-cx + ay) \end{pmatrix} \\ &= \frac{1}{ad - bc} \cdot \begin{pmatrix} (ad - bc)x \\ (ad - bc)y \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

Für zwei invertierbare Matrizen gilt folgende Rechenregel:

Satz 2.41. Seien $A, B \in \mathbb{K}^{n \times n}$ regulär. Dann ist auch $A \cdot B$ regulär und es gilt

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

Beweis. Es gilt

$$(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot ((B \cdot B^{-1}) \cdot A^{-1}) = A \cdot (E_n \cdot A^{-1}) = A \cdot A^{-1} = E_n.$$

□

Bemerkung 2.42. Analog zu obiger Formel für die Inverse des Produktes zweier Matrizen gilt für die Transponierte von $A \in \mathbb{K}^{m \times n}$ und $B \in \mathbb{K}^{n \times r}$ die Formel

$$(A \cdot B)^T = B^T \cdot A^T.$$

3 Lineare Gleichungssysteme

3.1 Einführung in LGS – Lösbarkeit und Lösungsmenge

Ein lineares Gleichungssystem (LGS) ist ein System der Form

$$\begin{aligned}\alpha_{1,1}x_1 + \dots + \alpha_{1,n}x_n &= b_1 \\ &\vdots \\ \alpha_{m,1}x_1 + \dots + \alpha_{m,n}x_n &= b_m.\end{aligned}$$

In diesem Fall ist es ein System aus m Gleichungen mit n Unbekannten. Mit den Matrizen und Vektoren

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

lässt sich das LGS wie folgt übersetzen:

$$A \cdot x = b. \tag{3.1}$$

Ist $b = 0 \in \mathbb{K}^m$, so heißt das LGS (3.1) *homogen*, andernfalls heißt das LGS *inhomogen*. Das LGS (3.1) kann auch als *erweiterte Koeffizientenmatrix* dargestellt werden:

$$(A|b) := \left(\begin{array}{ccc|c} \alpha_{1,1} & \cdots & \alpha_{1,n} & b_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & b_m \end{array} \right) \tag{3.2}$$

Sei

$$L(A, b) := \{x \in \mathbb{K}^n \mid A \cdot x = b\}$$

die Menge aller Lösungen des LGS $A \cdot x = b$. Überlegen wir zunächst, wie solch eine Lösungsmenge beschaffen ist.

Satz 3.1.

- (i) Ist f_A die zur Matrix $A \in \mathbb{K}^{m \times n}$ gehörende lineare Abbildung, d.h. $f_A(x) = A \cdot x$, so ist die Lösungsmenge von $A \cdot x = 0$ ein Untervektorraum von \mathbb{K}^n und es gilt

$$L(A, 0) = \ker(f_A).$$

- (ii) Gilt $y \in L(A, b)$, d.h. $A \cdot y = b$, dann ist die Lösungsmenge des inhomogenen LGS gegeben durch

$$L(A, b) = \{x + y \mid x \in \mathbb{K}^n, A \cdot x = 0\} = y + L(A, 0).$$

- (iii) Das LGS $A \cdot x = b$ ist genau dann lösbar, wenn $b \in \text{im}(f_A)$ liegt. Das ist genau dann der Fall, wenn der Rang von A gleich dem Rang der erweiterten Koeffizientenmatrix ist, d.h. $\text{rg}(A) = \text{rg}(A|b)$.

- (iv) Es gilt $\dim(\text{im}(f_A)) = \text{rg}(A) = n - \dim(\ker(f_A))$.

Beweis von (ii). Seien $y \in L(A, b)$ und $x \in L(A, 0)$, d.h. $A \cdot y = b$ und $A \cdot x = 0$ (Nullvektor). Dann ist

$$A \cdot (x + y) = A \cdot x + A \cdot y = b + 0 = b \Rightarrow x + y \in L(A, b).$$

Daraus folgt zunächst einmal $y + L(A, 0) \subseteq L(A, b)$. Sei nun $z \in L(A, b)$, also $A \cdot z = b$. Dann gilt

$$A \cdot (z - y) = A \cdot z - A \cdot y = b - b = 0 \Rightarrow z - y \in L(A, 0).$$

Folglich ist $z = y + (z - y) \in y + L(A, 0)$, also $L(A, b) \subseteq y + L(A, 0)$. □

Bemerkung 3.2.

- Punkt (ii) kann man auch so formulieren: Kennt man eine beliebige (man sagt auch eine partikuläre) Lösung des inhomogenen LGS, so erhält man sämtliche Lösungen des inhomogenen LGS indem man die Lösungen des homogenen LGS zu dieser partikulären Lösung addiert.
- Zu der ersten Aussage von (iii): $b \in \text{im}(f_A)$ genau dann, wenn ein Vektor x existiert, für den $f_A(x) = A \cdot x = b$ gilt und das ist dann gerade ein Element der Lösungsmenge des inhomogenen LGS, d.h. $x \in L(A, b)$.
- Um den zweiten Teil von Aussage (iii) zu zeigen, verwendet man die Eigenschaft, dass die Spalten einer Matrix A eine Basis des Bildes der linearen Abbildung $f_A(x) = A \cdot x$ bilden

Beispiel 3.3. Betrachten wir das LGS mit der Matrix A aus Beispiel 2.37:

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

Wir haben bereits gesehen, dass $\text{rg}(A) = 2$ ist, d.h. der Lösungsraum des homogenen LGS $A \cdot x = 0 \in \mathbb{R}^3$ (mit $x \in \mathbb{R}^2$, also $n = 2$) ist 0-dimensional:

$$2 = \text{rg}(A) = 2 - \dim(\ker(f_A)) \Rightarrow \dim(\ker(f_A)) = 0.$$

Dies können wir rechnerisch bestätigen: $A \cdot x = 0$ lässt sich übersetzen als LGS

$$1x_1 + 2x_2 = 0$$

$$3x_1 + 5x_2 = 0$$

$$7x_1 + 6x_2 = 0,$$

welches als einzige Lösung $(x_1, x_2) = (0, 0)$ hat. Sei nun $b = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$. Dann hat das LGS $A \cdot x = b$ die

Lösung $(x_1, x_2) = (1, -1)$. Nach obigem Satz ist das tatsächlich die einzige Lösung des inhomogenen LGS.

Überprüfen wir noch den Rang der erweiterten Matrix $(A|b)$:

$$(A|b) = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 5 & -2 \\ 7 & 6 & 1 \end{pmatrix}.$$

Wir haben bereits festgestellt, dass sich der dritte Spaltenvektor als Linearkombination der ersten beiden darstellen lässt:

$$1 \cdot \begin{pmatrix} 1 \\ 3 \\ 7 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}.$$

Folglich sind nicht alle drei Vektoren linear unabhängig, der Spaltenrang von $(A|b)$ ist also (höchstens und gleichzeitig genau) zwei. Da $\text{rg}(A) = \text{rg}_S(A) = 2$ ist, gilt also $\text{rg}(A) = \text{rg}(A|b)$ wie vorhergesagt.

Aus dem Satz 3.1 folgt sofort:

Korollar 3.4. Ist $A \in \mathbb{K}^{n \times n}$ eine reguläre Matrix (d.h. $\text{rg}(A) = n$), dann ist das LGS $A \cdot x = b$ für jeden beliebigen Vektor $b \in \mathbb{K}^n$ eindeutig lösbar. Der Lösungsvektor ist $x = A^{-1} \cdot b$.

Beweis. Ist $A \in \mathbb{K}^{n \times n}$ regulär, dann ist $\text{rg}(A) = n$. Nach Aussage (iv) des Satzes ist damit

$$\text{rg}(A) = n = n - \dim(\ker(f_A)) \Rightarrow \dim(\ker(f_A)) = 0.$$

Nach Teil (i) ist die Dimension der Lösungsmenge des homogenen LGS $\dim(L(A, 0))$ Null, d.h. $L(A, 0)$ besteht nur aus dem Nullvektor. Die Lösungsmenge $L(A, b)$ kann also maximal aus einem Element bestehen nach (ii). Da A regulär ist, existiert A^{-1} und da $x = A^{-1}b$ das inhomogene LGS löst, ist die gewünschte Aussage bewiesen. \square

Beispiel 3.5. Betrachten wir folgendes (inhomogenes) LGS:

$$\begin{aligned} 3x_1 + 5x_2 &= 4 \\ 1x_1 + 2x_2 &= -2. \end{aligned}$$

Dieses Problem lässt sich in Matrixform schreiben als $A \cdot x = b$, wobei

$$A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ -2 \end{pmatrix}.$$

Die Matrix hat zwei voneinander linear unabhängige Spaltenvektoren und damit vollen Rang ($\text{rg}(A) = 2$). A ist damit regulär, also invertierbar. In Beispiel 2.40 haben wir die Inverse einer regulären 2×2 -Matrix kennen gelernt. Wenden wir die Formel auf Matrix A an, so erhalten wir

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

Die einzige Lösung des LGS ist damit

$$x = A^{-1} \cdot b = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 18 \\ -10 \end{pmatrix}.$$

Wie erkennt man möglichst einfach und systematisch den Rang einer Matrix und bestimmt Lösungen von linearen Gleichungssystemen? Im nächsten Abschnitt befassen wir uns mit verschiedenen Lösungsmethoden.

3.2 Methoden zur Lösung von LGS

Die folgenden Methoden sind gerade bei kleinen übersichtlichen LGS sinnvoll:

Einsetzungsverfahren. Man stellt eine Gleichung nach einer Variablen um und ersetzt die Variable in den übrigen Gleichungen durch die gefundene Darstellung:

$$\begin{cases} 3x - y = 3 & \Rightarrow y = 3x - 3 & \Rightarrow y = \frac{33}{8} - 3 = \frac{9}{8} \\ 2x + 2y = 5 & \Rightarrow 2x + 2(3x - 3) = 5 & \Leftrightarrow x = \frac{11}{8} \end{cases}$$

Gleichsetzungsverfahren. Man löst zwei Gleichungen nach der selben Variablen auf und setzt die entstehenden Ausdrücke gleich:

$$\begin{cases} 3x - y = 3 & \iff y = 3x - 3 \\ 2x + 2y = 5 & \iff y = \frac{5}{2} - x \end{cases}$$

Gleichsetzen liefert $3x - 3 = \frac{5}{2} - x$, also $x = \frac{11}{8}$. Einsetzen in jede beliebige der Gleichungen ergibt dann $y = \frac{9}{8}$.

Eliminationsverfahren. Ein Vielfaches einer Gleichung wird zu einem Vielfachen einer anderen Gleichung addiert, damit eine der Variablen herausfällt:

$$3x - y = 3 \tag{3.3}$$

$$2x + 2y = 5 \tag{3.4}$$

$2 \cdot (3.3) + (3.4)$ ergibt $6x + 2x = 6 + 5$, also wieder $x = \frac{11}{8}$. Einsetzen in eine beliebige Gleichung liefert $y = \frac{9}{8}$.

Wir haben gesehen, dass wir durch Vertauschungen, geschicktes Multiplizieren und Addieren die Lösung bekommen. Dies lässt sich auch auf die Matrix-Schreibweise übertragen, indem man elementare Zeilen- und Spaltenumformungen durchführt.

- (S1) Vertauschung zweier Spalten
- (S2) Multiplikation einer Spalte mit einem Skalar $\lambda \neq 0$
- (S3) Addition einer Spalte zu einer anderen Spalte
- (Z1) Vertauschung zweier Zeilen
- (Z2) Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0$
- (Z3) Addition einer Zeile zu einer anderen Zeile

Es ist wichtig zu bemerken, dass diese (elementaren) Umformungen den Rang einer Matrix nicht ändern.

Für die erweiterte Koeffizientenmatrix $(A|b)$ gilt zudem, dass sich die Lösungsmenge nach der Durchführung von elementaren Zeilenumformungen nicht ändert.

Die Idee hinter dem im Folgenden vorgestellten Verfahren ist es, diese elementaren Zeilen- und Spaltenumformungen an der erweiterten Koeffizientenmatrix $(A|b)$ durchzuführen bis die Matrix in der sog. *Zeilenstufenform* vorliegt. $(A|b)$ wird dann auch obere Dreiecksmatrix genannt.

Definition 3.6. Die erweiterte Koeffizientenmatrix $(A|b)$ mit $A = (\alpha_{k,j})$ ($1 \leq k \leq m, 1 \leq j \leq n$) und $b = (b_k)$ ($1 \leq k \leq m$) liegt in Zeilenstufenform vor, falls gilt (mit der Notation $\alpha_{k,n+1} := b_k$):

- Sei $\alpha_{k,j} \neq 0$ der erste von Null verschiedene Eintrag der k -ten Zeile, genannt *Pivotelement* oder *führende Koeffizient*. Dann ist der führende Koeffizient der $k + 1$. Zeile rechts davon. Insbesondere ist $\alpha_{k+1,j} = 0$.
- Zeilen, die nur Nullen enthalten, stehen ganz unten in der Matrix.

Die Matrix liegt in der erweiterten/normierten Zeilenstufenform vor, falls

- Das Pivotelement einer jeden Zeile ist 1, sofern die Zeile nicht nur aus Nullen besteht.
- Alle Einträge in einer Pivotspalte außer dem Pivotelement sind Null.

Beispiel 3.7. Betrachten wir die Matrizen

$$A = \begin{pmatrix} 3 & 7 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

A liegt in Zeilenstufenform vor, B in der erweiterten Zeilenstufenform.

3.2.1 Gauß-(Jordan-)Verfahren

Das Gauß-Verfahren überführt eine Matrix in die Zeilenstufenform; das Gauß-Jordan-Verfahren überführt sie in die erweiterte Zeilenstufenform.

Beispiel 3.8. Betrachten wir das homogene lineare Gleichungssystem mit Matrix

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 4 & -1 & 3 \\ 6 & 3 & 2 \end{pmatrix}, \quad \text{also } (A|0) = \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 4 & -1 & 3 & 0 \\ 6 & 3 & 2 & 0 \end{array} \right).$$

Transformieren wir die Matrix $(A|b)$ schrittweise in die Zeilenstufenform und dann noch in die erweiterte Zeilenstufenform. Dazu benutzen wir die Kurznotation Z_k für die k -te Zeile.

$$\begin{aligned} Z_2 \rightsquigarrow 4 \cdot Z_1 - Z_2 & \quad \text{ergibt} \quad \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 5 & 0 \\ 6 & 3 & 2 & 0 \end{array} \right) \\ Z_3 \rightsquigarrow 6 \cdot Z_1 - Z_3 & \quad \text{ergibt} \quad \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & -3 & 10 & 0 \end{array} \right) \\ Z_3 \rightsquigarrow 3 \cdot Z_2 + 1 \cdot Z_3 & \quad \text{ergibt} \quad \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 25 & 0 \end{array} \right) \text{ — das ist die Zeilenstufenform} \\ Z_3 \rightsquigarrow \frac{1}{25} \cdot Z_3 & \quad \text{ergibt} \quad \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \\ Z_2 \rightsquigarrow Z_2 - 5 \cdot Z_3 & \quad \text{ergibt} \quad \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \\ Z_1 \rightsquigarrow Z_1 - 2 \cdot Z_3 & \quad \text{ergibt} \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \text{ — das ist die erweiterte Zeilenstufenform} \end{aligned}$$

Das neue LGS hat offensichtlich als einzige Lösung $x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Damit haben wir insbesondere gezeigt, dass \mathcal{M}_4 aus Beispiel 1.29 in der Tat eine Basis des \mathbb{R}^3 ist.

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 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 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 Schritt 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 freien Variablen angegeben.

Bemerkung 3.9.

- (i) Man kann auch mehrere LGS mit verschiedenen rechten Seiten $b, c, \dots \in \mathbb{R}^m$ gleichzeitig lösen. Dazu bildet man in Schritt 1 die erweiterte Koeffizientenmatrix $(A|bc \dots)$ und führt die Schritte wie beschrieben durch.
- (ii) Will man die Inverse zu $A \in \mathbb{K}^{n \times n}$ berechnen, so bildet man die erweiterte Koeffizientenmatrix $(A|E_n)$ für die n -dimensionale Einheitsmatrix.

Beispiel 3.10. Berechnen wir die Inverse zu $A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ aus Beispiel 3.5.

Die erweiterte Koeffizientenmatrix lautet

$$(A|E_2) = \left(\begin{array}{cc|cc} 3 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array} \right).$$

Elementare Zeilenumformungen ergeben

$$\left(\begin{array}{cc|cc} 3 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 3 & 5 & 1 & 0 \\ 0 & -1 & 1 & -3 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 5/3 & 1/3 & 0 \\ 0 & 1 & -1 & 3 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & 2 & -5 \\ 0 & 1 & -1 & 3 \end{array} \right)$$

Am Ende kann man in der Tat die Inverse ablesen: $A^{-1} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$.

Beispiel 3.11 (Keine Lösung). *Betrachten wir das LGS*

$$\begin{cases} x_1 - 2x_2 & = 5 \\ -3x_1 + 6x_2 & = 8. \end{cases}$$

Die Umformung der erweiterten Koeffizientenmatrix liefert

$$(A|b) = \left(\begin{array}{cc|c} 1 & -2 & 5 \\ -3 & 6 & 8 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & -2 & 5 \\ 0 & 0 & 23 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & -2 & 5 \\ 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 1 \end{array} \right)$$

Die letzte Spalte, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, ist also Pivotspalte. Somit ist das LGS nicht lösbar. In der Tat ist nach obigen Umformung das ursprüngliche LGS äquivalent zu

$$\begin{cases} x_1 - 2x_2 & = 0 \\ 0 & = 1, \end{cases}$$

welches erkennbar keine Lösung besitzt.

Beispiel 3.12 (Mehr als eine Lösung – Darstellung in Parameterform). *Betrachten wir das LGS mit folgender erweiterter Koeffizientenmatrix:*

$$(A|b) = \left(\begin{array}{cccc|c} 0 & 0 & 1 & 3 & 2 \\ 1 & 2 & 1 & 4 & 3 \\ 1 & 2 & 2 & 7 & 5 \end{array} \right).$$

Elementare Zeilenumformungen ergeben:

$$\left(\begin{array}{cccc|c} 0 & 0 & 1 & 3 & 2 \\ 1 & 2 & 1 & 4 & 3 \\ 1 & 2 & 2 & 7 & 5 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 2 & 7 & 5 \\ 1 & 2 & 1 & 4 & 3 \\ 0 & 0 & 1 & 3 & 2 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 2 & 7 & 5 \\ 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 3 & 2 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 2 & 7 & 5 \\ 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Das (zu dem ursprünglichen LGS äquivalente) LGS lautet damit

$$\begin{cases} x_1 + 2x_2 + 2x_3 + 7x_4 & = 5 \\ x_3 + 3x_4 & = 2 \\ 0 & = 0. \end{cases}$$

Die Variablen x_2 und x_4 sind frei, da die 2. und 4. Spalte keine Pivotspalte sind. Aus der zweiten Gleichung ergibt sich die Bedingung

$$x_3 = 2 - 3x_4,$$

eingesetzt in die 1. Zeile ergibt sich

$$x_1 = 5 - 2x_2 - 2(2 - 3x_4) - 7x_4 = 1 - 2x_2 - x_4.$$

Die Lösungsmenge des LGS ist somit

$$\begin{aligned} L(A, b) &= \left\{ x \in \mathbb{R}^4 \mid x_1 = 1 - 2x_2 - x_4, x_3 = 2 - 3x_4 \right\} \\ &= \left\{ \begin{pmatrix} 1 - 2x_2 - x_4 \\ x_2 \\ 2 - 3x_4 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 \mid x_2, x_4 \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -3 \\ 1 \end{pmatrix} \mid x_2, x_4 \in \mathbb{R} \right\} \end{aligned}$$

Interpretieren wir dieses Ergebnis in Hinblick auf Abschnitt 3.1:

- $A \in \mathbb{R}^{3 \times 4}$, d.h. $f_A: \mathbb{R}^4 \rightarrow \mathbb{R}^3$. Damit ist $m = 3$ und $n = 4$.
- $\text{rg}(A) = \text{rg}(A|b) = 2$, somit ist $\dim(\ker(f_A)) = n - \text{rg}(A) = 4 - 2 = 2$, d.h. die Lösungsmenge des homogenen LGS ist 2-dimensional.
- Die Lösungsmenge des inhomogenen LGS ist somit $y + \ker(f_A)$, wenn y eine Lösung von $A \cdot x = b$ ist.

- Nach obiger Rechnung ist $y = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}$ die partikuläre Lösung des inhomogenen LGS und

$$\ker(f_A) = \left\{ x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -3 \\ 1 \end{pmatrix} \mid x_2, x_4 \in \mathbb{R} \right\}.$$

4 Differentialrechnung für Funktionen mehrerer reeller Variablen

4.1 Differenzierbarkeit von Funktionen in mehreren Variablen

Definition 4.1. Seien $n \in \mathbb{N}$, $M \subset \mathbb{R}^n$, $a = (a_1, \dots, a_n)^T \in M$ und $f: M \rightarrow \mathbb{R}$ gegeben. Wenn für ein $j \in \{1, \dots, n\}$ der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - f(a)}{h} =: g \in \mathbb{R}$$

existiert, so heißt f partiell differenzierbar in Richtung von x_j (oder in der j -ten Komponente) und g heißt partielle Ableitung von f in Richtung von x_j .

Notation: $g = \frac{\partial f}{\partial x_j}(a) = \partial_{x_j} f(a) = f_{x_j}(a)$ oder auch verkürzt $\partial_j f(a)$.

Definition 4.2. Ist f in allen Punkten $a \in M$ partiell differenzierbar in Richtung x_j , so heißt f partiell differenzierbar in Richtung x_j . Ist f in allen Punkten $a \in M$ partiell differenzierbar in alle Richtungen x_1, \dots, x_n , so heißt f partiell differenzierbar.

Beispiel 4.3. Sei $f(c, x) := cx^2$. Dann ist

$$\begin{aligned} \frac{\partial f}{\partial c}(c, x) &= \lim_{h \rightarrow 0} \frac{f(c+h, x) - f(c, x)}{h} = \lim_{h \rightarrow 0} \frac{(c+h)x^2 - cx^2}{h} = x^2, \\ \frac{\partial f}{\partial x}(c, x) &= \lim_{h \rightarrow 0} \frac{f(c, x+h) - f(c, x)}{h} = \lim_{h \rightarrow 0} \frac{c(x+h)^2 - cx^2}{h} = \lim_{h \rightarrow 0} (2cx + ch) = 2cx. \end{aligned}$$

Man berechnet die partielle Ableitung, indem man alle Komponenten, nach denen man nicht ableitet, als Parameter behandelt.

Beispiel 4.4. Sei $M = \mathbb{R}^2$ und für $x = (x_1, x_2)^T \in M$ sei $f(x_1, x_2) := 3x_1^2 - x_2^5 + x_1x_2 + 7$. Dann sind die partiellen Ableitungen $\partial_1 f(x_1, x_2) = 6x_1 + x_2$ und $\partial_2 f(x_1, x_2) = -5x_2^4 + x_1$.

Definition 4.5. Seien $n \in \mathbb{N}$, $M \subset \mathbb{R}^n$, $a \in M$ und $f: M \rightarrow \mathbb{R}$. Falls eine lineare Abbildung $L: \mathbb{R}^n \rightarrow \mathbb{R}$ existiert, so dass

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - L(h)}{|h|} = 0$$

gilt, dann heißt f (total) differenzierbar in a und $df(a) := L$ Ableitung (oder Differential) von f in a . Ist f in allen Punkten $a \in M$ differenzierbar, so heißt f differenzierbar auf M .

Bemerkung 4.6.

- (i) Ist $h = (h_1, \dots, h_n)^T \in \mathbb{R}^n$, dann ist $|h| := \|h\|_2 = \sqrt{h_1^2 + \dots + h_n^2}$. Das ist die euklidische Norm des \mathbb{R}^n .
- (ii) Das Differential df (in der Literatur auch mit Df bezeichnet) hat zwei Argumente – die Stelle an der f abgeleitet wird ($a \in M$) und einen weiteren Vektor ($h \in \mathbb{R}^n$, nahe Null).
- (iii) Für den Spezialfall $n = 1$ gilt $df(a)(h) = f'(a) \cdot h$ und damit $f'(a) = df(a)(1)$.
- (iv) Das Differential ist, sofern es existiert, eindeutig bestimmt.
- (v) Ist f in einem Punkt a differenzierbar, so ist f in diesem Punkt auch stetig.

Wir hatten bereits festgestellt, dass für eine Matrix $A \in \mathbb{R}^{m \times n}$ die Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, welche durch $x \mapsto f(x) = A \cdot x$ gegeben ist, linear ist. Andersherum können wir jeder linearen Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Matrix $A \in \mathbb{R}^{m \times n}$ zuordnen, so dass $f(x) = A \cdot x$ für alle $x \in \mathbb{R}^n$. Im Spezialfall $m = 1$ ist dann für das kanonische Skalarprodukt $\langle \cdot, \cdot \rangle$ des \mathbb{R}^n :

$$f(x) = A \cdot x = \langle A^T, x \rangle.$$

Da die Ableitung eine lineare Abbildung ist, können wir sie durch das Skalarprodukt mit einem festen Vektor darstellen.

Definition 4.7. Für $M \subset \mathbb{R}^n$ sei $f: M \rightarrow \mathbb{R}$ in $a \in M$ differenzierbar. Dann heißt der Vektor $v \in \mathbb{R}^n$, welcher die Bedingung

$$df(a)(h) = \langle v, h \rangle, \quad \forall h \in \mathbb{R}^n$$

erfüllt, Gradient von f in a . Wir schreiben $v =: \text{grad } f(a) =: \nabla f(a)$.

Satz 4.8. Seien $n \in \mathbb{N}$, $M \subset \mathbb{R}^n$ und $f: M \rightarrow \mathbb{R}$.

1. Wenn f differenzierbar ist, dann ist f auch partiell differenzierbar und für alle $a \in M$ ist $\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)^T$.
2. Wenn f partiell differenzierbar ist und alle partiellen Ableitungen stetig sind, dann ist f (total) differenzierbar.

Bemerkung 4.9. Ist eine Funktion f differenzierbar, so kann man das Differenzial also mit Hilfe des Gradienten ausdrücken:

$$df(a)(h) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(a) \cdot h_k = \langle \nabla f(a), h \rangle, \quad h = (h_1, \dots, h_n)^T.$$

Bemerkung 4.10. Ein kleiner Vorgriff: Für Differentialgleichungen hat das totale Differential die Darstellung $df = \sum_{k=1}^n \frac{\partial f}{\partial x_k} dx_k$.

Definition 4.11. Seien $n \in \mathbb{N}$, $M \subset \mathbb{R}^n$, $a \in M$, $b \in \mathbb{R}^n$ mit $|b| = 1$ und $f: M \rightarrow \mathbb{R}$. f besitzt in a eine Ableitung in Richtung b , falls

$$\nabla_b f(a) := \lim_{h \rightarrow 0} \frac{f(a + hb) - f(a)}{h} \quad (\text{hier } h \in \mathbb{R} \setminus \{0\})$$

existiert.

Die partielle Ableitung ist ein Spezialfall der Richtungsableitung. So ist z.B. $\frac{\partial f}{\partial x_i}(a) = \nabla_{e_i} f(a)$, wenn e_i der Vektor ist, der überall den Eintrag 0 hat außer in der i -ten Komponente, wo eine 1 steht.

Bemerkung 4.12. Der Gradient ∇f steht immer senkrecht auf die Niveaulinien

$$N_f(c) = \{x \in M \mid f(x) = c\} = f^{-1}(\{c\}), \quad c \in \mathbb{R},$$

und zeigt in Richtung des steilsten Anstiegs von f , d.h. ist $b \in \mathbb{R}^n$ mit $|b| = 1$ und $a \in M$ wie bisher, dann gilt $\nabla_b f(a) \leq |\nabla f(a)|$ und Gleichheit gilt g.d.w. $b = \frac{\nabla f(a)}{|\nabla f(a)|}$. (Achtung! $\nabla_b f(a) \in \mathbb{R}$, aber $\nabla f(a) \in \mathbb{R}^n$.)

Ist f in a differenzierbar, dann ist für b wie eben $df(a)(b) = \nabla_b f(a) = \langle \nabla f(a), b \rangle$.

Beispiel 4.13. Sei $M = \mathbb{R}^2$ und für $x = (x_1, x_2)^T \in M$ sei $f(x_1, x_2) := 2x_1 - x_2 + x_1x_2 + 5$. Wir berechnen: $\partial_1 f(x_1, x_2) = 2 + x_2$ und $\partial_2 f(x_1, x_2) = -1 + x_1$. Sei jetzt $b := \frac{1}{5}(3, 4)^T$. Dann ist $|b| = \frac{1}{5}\sqrt{3^2 + 4^2} = 1$. Die Richtungsableitung ist

$$\begin{aligned} \nabla_b f(x) &= \lim_{h \rightarrow 0} \frac{2(x_1 + \frac{3}{5}h) - (x_2 + \frac{4}{5}h) + (x_1 + \frac{3}{5}h)(x_2 + \frac{4}{5}h) + 5 - f(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{2 \cdot \frac{3}{5}h - \frac{4}{5}h + x_1 \cdot \frac{4}{5}h + x_2 \cdot \frac{3}{5}h + \frac{3}{5}h \frac{4}{5}h}{h} \\ &= \lim_{h \rightarrow 0} \left(\frac{6}{5} - \frac{4}{5} + x_1 \cdot \frac{4}{5} + x_2 \cdot \frac{3}{5} + \frac{12}{25}h \right) \\ &= \frac{2}{5} + \frac{4}{5} \cdot x_1 + \frac{3}{5} \cdot x_2. \end{aligned}$$

Andererseits ist

$$\begin{aligned} \langle \nabla f(x), b \rangle &= \frac{\partial f}{\partial x_1}(x) \cdot \frac{3}{5} + \frac{\partial f}{\partial x_2}(x) \cdot \frac{4}{5} \\ &= (2 + x_2) \cdot \frac{3}{5} + (-1 + x_1) \cdot \frac{4}{5} \\ &= \frac{6}{5} + \frac{3}{5} \cdot x_2 - \frac{4}{5} + \frac{4}{5} \cdot x_1 = \nabla_b f(x) \end{aligned}$$

für beliebiges $x \in M$ wie behauptet.

Der Gradient an der Stelle $(0, 0)^T$ ist $\nabla f(0, 0) = (2, -1)^T$. Für eine beliebige Richtung ist die Richtungsableitung $\langle \nabla f(0, 0), b \rangle = (2+0)b_1 + (-1+0)b_2 = 2b_1 - b_2$. In Richtung des (normierten) Gradienten erhalten wir also $\langle \nabla f(0, 0), \frac{\nabla f(0,0)}{|\nabla f(0,0)|} \rangle = \frac{5}{\sqrt{5}} = \sqrt{5}$.

Bemerkung 4.14. Ist eine Funktion f differenzierbar, so ist sie auch stetig. Ist die Funktion jedoch nur partiell differenzierbar, so braucht sie nicht stetig zu sein. Ein Gegenbeispiel ist $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) := \begin{cases} \frac{xy}{(x^2+y^2)^2}, & \text{falls } (x, y) \neq (0, 0), \\ 0, & \text{falls } (x, y) = (0, 0). \end{cases}$$

Im Punkt $(0, 0)^T$ (und auch in dessen Umgebung) ist die Funktion partiell differenzierbar, sie ist jedoch nicht stetig in diesem Punkt. [→ Übungsblatt]

Satz 4.15 (Kettenregel – mehrdimensional). Seien $I \subset \mathbb{R}$ und $M \subset \mathbb{R}^n$ für $n \in \mathbb{N}$ (offene) Mengen. Sei $g = (g_1, \dots, g_n)^T: I \rightarrow M$ mit Koordinatenfunktionen g_k ($k = 1, \dots, n$), die differenzierbar sind in $t_0 \in I$. Sei weiterhin $f: M \rightarrow \mathbb{R}$ differenzierbar in $g(t_0) \in M$. Dann ist die Komposition $f \circ g: I \rightarrow \mathbb{R}$ mit $(f \circ g)(t) = f(g(t))$ für $t \in I$ differenzierbar in t_0 mit Ableitung

$$(f \circ g)'(t_0) = \langle \nabla f(g(t_0)), g'(t_0) \rangle = \sum_{k=1}^n \partial_k f(g(t_0)) \cdot g'_k(t_0).$$

4.2 Ableitungen höherer Ordnung in mehreren Variablen

Definition 4.16. Seien $n \in \mathbb{N}$, $M \subset \mathbb{R}^n$ und $f: M \rightarrow \mathbb{R}$ differenzierbar. Seien alle partiellen Ableitungen $g_k := \frac{\partial f}{\partial x_k}: M \rightarrow \mathbb{R}$ ($k = 1, \dots, n$) ebenfalls differenzierbar. Dann sind die zweiten Ableitungen von f gegeben durch

$$\frac{\partial^2 f}{\partial x_k \partial x_l} := \frac{\partial g_l}{\partial x_k} \quad \text{für } k, l \in \{1, \dots, n\}.$$

In diesem Fall ist die Hesse-Matrix von f :

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} = \begin{pmatrix} \nabla \frac{\partial f}{\partial x_1}^T \\ \nabla \frac{\partial f}{\partial x_2}^T \\ \vdots \\ \nabla \frac{\partial f}{\partial x_n}^T \end{pmatrix}.$$

Beispiel 4.17. Sei $M = \mathbb{R}^2$ und für $x = (x_1, x_2)^T \in M$ sei $f(x) := 2x_1 - x_2 + x_1x_2 + 5$. Wir berechnen: $\partial_1 f(x_1, x_2) = 2 + x_2$ und $\partial_2 f(x_1, x_2) = -1 + x_1$. Der Gradient von $\partial_1 f(x_1, x_2) = 2 + x_2$ ist $(0, 1)^T$, der Gradient von $\partial_2 f(x_1, x_2) = -1 + x_1$ ist $(1, 0)^T$. Damit ist insgesamt

$$H_f(x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Definition 4.18. Eine Funktion $f: M \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt k -fach stetig differenzierbar, falls alle partiellen Ableitungen von f der Ordnung $\leq k$ existieren und stetig sind.

Notation: $f \in C^k(M; \mathbb{R}^m)$.

Satz 4.19 (Satz von Schwarz). Ist $f \in C^k(M; \mathbb{R}^m)$ für $M \subset \mathbb{R}^n$ und $k, m, n \in \mathbb{N}$, dann sind alle partiellen Ableitungen der Ordnung $\leq k$ unabhängig von der Reihenfolge des Differenzierens.

Im Spezialfall $k = 2$ bedeutet das: Existieren die partiellen Ableitungen erster und zweiter Ordnung und sind sie stetig, dann ist $\frac{\partial^2 f}{\partial x_k \partial x_l} = \frac{\partial^2 f}{\partial x_l \partial x_k}$ für $k, l \in \{1, \dots, n\}$.

4.3 Vektorfelder

Definition 4.20 (Vektorfeld). Für $n \in \mathbb{N}$ sei $M \subset \mathbb{R}^n$. Eine Abbildung $f: M \rightarrow \mathbb{R}^n$ heißt Vektorfeld.

Beispiel 4.21. Ist $f: M \subset \mathbb{R}^n \rightarrow \mathbb{R}$ partiell differenzierbar, dann wird durch

$$\nabla f: M \rightarrow \mathbb{R}^n, \quad x = (x_1, \dots, x_n)^T \mapsto \nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)^T(x)$$

ein Vektorfeld definiert – genannt Gradientenfeld.

Sind sowohl Definitions- als auch Wertebereich einer Funktion mehrdimensional, tritt an die Stelle des Gradienten die sogenannte Jacobi-Matrix. Konkret heißt das für $f = (f_1, \dots, f_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $n, m \in \mathbb{N}$:

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} = \begin{pmatrix} (\nabla f_1)^T \\ (\nabla f_2)^T \\ \vdots \\ (\nabla f_m)^T \end{pmatrix}.$$

Für den Spezialfall $m = 1$ ist die Jacobi-Matrix also die Transponierte des Gradienten.

Definition 4.22 (exaktes Vektorfeld). Ein Vektorfeld $f: M \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt exaktes Vektorfeld, wenn es eine Funktion $g: M \rightarrow \mathbb{R}$ gibt, so dass $f = \nabla g$.

Satz 4.23. Sei $M \subset \mathbb{R}^n$ für ein $n \in \mathbb{N}$. Ein Vektorfeld $f: M \rightarrow \mathbb{R}^n$ ist genau dann exakt, wenn für alle $k, l \in \{1, \dots, n\}$ (mit $k \neq l$) gilt: $\frac{\partial f_k}{\partial x_l} = \frac{\partial f_l}{\partial x_k}$.

Etwas leichter zu merken ist vielleicht die äquivalente Aussage:

Ein Vektorfeld f ist genau dann exakt, wenn die Jacobi-Matrix Df symmetrisch ist.

5 Determinanten

Determinanten sind (1-dimensionale) Kennzahlen von quadratischen Matrizen, die sich als sehr nützlich erweisen werden — u.a. bei der Untersuchung der Invertierbarkeit einer Matrix und (im nächsten Kapitel) bei der Berechnung von Eigenwerten von Matrizen und sie taucht bei der Transformationsregel der Integration auf.

5.1 Definition, Berechnung und Eigenschaften

Sei im Folgenden

$$A = (\alpha_{k,j})_{\substack{1 \leq k \leq n \\ 1 \leq j \leq n}} \in \mathbb{K}^{n \times n}, \quad A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_n \end{pmatrix},$$

d.h. $\boldsymbol{\alpha}_k$ sei die k -te Zeile von A .

Definition 5.1 (Untermatrix). *Seien $I \subset \{1, \dots, n\}$ und $J \subset \{1, \dots, n\}$ Teilmengen der natürlichen Zahlen $\{1, \dots, n\}$. Dann ist die Untermatrix*

$$A_{IJ} := (\alpha_{i,j})_{\substack{i \in \{1, \dots, n\} \setminus I \\ j \in \{1, \dots, n\} \setminus J}} \in \mathbb{K}^{(n-|I|) \times (n-|J|)}$$

die Matrix mit $n - |I|$ Zeilen und $n - |J|$ Spalten, die entsteht, wenn man alle Zeilen mit den Indizes $i \in I$ und alle Spalten mit den Indizes $j \in J$ aus A streicht. Insbesondere ist

$$A_{k,j} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{k-1,1} & \cdots & \alpha_{k-1,j-1} & \alpha_{k-1,j+1} & \cdots & \alpha_{k-1,n} \\ \alpha_{k+1,1} & \cdots & \alpha_{k+1,j-1} & \alpha_{k+1,j+1} & \cdots & \alpha_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{pmatrix}.$$

Definition 5.2 (Determinante). *Eine Funktion $\det: \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ ($n \in \mathbb{N}$) heißt Determinante, falls sie folgende Eigenschaften erfüllt:*

- (i) *det ist linear bezüglich jeder Zeile der Matrix, d.h. für $k \in \{1, \dots, n\}$ und $\mu, \lambda \in \mathbb{K}$ und $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{K}^{1 \times n}$ gilt*

$$\det \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{k-1} \\ \lambda \boldsymbol{\alpha} + \mu \boldsymbol{\beta} \\ \boldsymbol{\alpha}_{k+1} \\ \vdots \\ \boldsymbol{\alpha}_n \end{pmatrix} = \lambda \det \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{k-1} \\ \boldsymbol{\alpha} \\ \boldsymbol{\alpha}_{k+1} \\ \vdots \\ \boldsymbol{\alpha}_n \end{pmatrix} + \mu \det \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{k-1} \\ \boldsymbol{\beta} \\ \boldsymbol{\alpha}_{k+1} \\ \vdots \\ \boldsymbol{\alpha}_n \end{pmatrix}.$$

- (ii) *det ist alternierend, d.h. besitzt A zwei identische Zeilen, so ist $\det(A) = 0$.*

- (iii) *det ist normiert, d.h. $\det(E_n) = 1$.*

Bemerkung 5.3.

- 1.) Es lässt sich zeigen, dass für jedes $n \in \mathbb{N}$ genau eine Abbildung mit diesen Eigenschaften existiert.
- 2.) Alternativ könnte man die Determinante über die Spalten der Matrix definieren, d.h. \det ist linear bezüglich der Spalten der Matrix und besitzt eine Matrix A zwei identische Spalten, so ist $\det(A) = 0$.
- 3.) Man schreibt auch $|A|$ statt $\det(A)$ für die Determinante einer Matrix, sofern keine Verwechslungsgefahr mit dem Betrag besteht. Anders als der Betrag oder eine Norm kann die Determinante jedoch negative Werte annehmen!
- 4.) Eine wichtige Folgerung aus obigen Eigenschaften ist die Determinanten einer 1-dimensionalen Matrix $A = (\alpha_{1,1}) \in \mathbb{K}^{1 \times 1}$: $\det(A) = \alpha_{1,1}$.

Es gibt mehrere Möglichkeiten, die Determinante zu berechnen. Eine davon bietet der Entwicklungssatz von Laplace, der die Determinante einer $n \times n$ -Matrix mit Hilfe von Determinanten von $(n-1) \times (n-1)$ -dimensionalen Untermatrizen definiert.

Satz 5.4 (Entwicklungssatz von Laplace). Sei $A = (\alpha_{k,j}) \in \mathbb{K}^{n \times n}$. Dann gilt einerseits die Entwicklungsformel nach der k -ten Zeile von A

$$\det(A) = \sum_{j=1}^n \alpha_{k,j} (-1)^{k+j} \det(A_{k,j})$$

und andererseits die Entwicklungsformel nach der j -ten Spalte von A

$$\det(A) = \sum_{k=1}^n \alpha_{k,j} (-1)^{k+j} \det(A_{k,j}),$$

wobei jeweils der festgehaltene Index $j, k \in \{1, \dots, n\}$ beliebig ist.

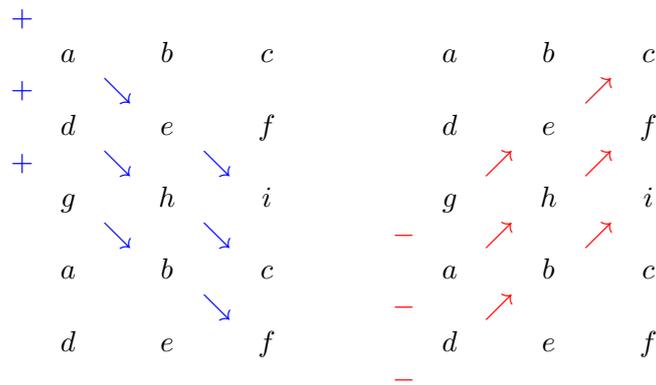
Beispiel 5.5 (2×2 -Matrix). Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Führen wir eine Entwicklung nach der 1. Zeile der Matrix durch:

$$\begin{aligned} \det(A) &= \alpha_{1,1} \cdot (-1)^{1+1} \cdot \det(A_{1,1}) + \alpha_{1,2} \cdot (-1)^{1+2} \cdot \det(A_{1,2}) \\ &= a \cdot 1 \cdot \det(d) + b \cdot (-1) \cdot \det(c) \\ &= ad - bc. \end{aligned}$$

Beispiel 5.6 (3×3 -Matrix). Sei $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Dann ist

$$\begin{aligned} \det(A) &= a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix} \\ &= a(ei - hf) - b(di - gf) + c(dh - ge) \\ &= aei + dhc + gbf - gec - hfa - idb. \end{aligned}$$

Diese Form lässt sich mit der sog. Regel von Sarrus merken.



Folgende Eigenschaften der Determinante können deren Berechnung erleichtern und geben gleichzeitig erste Hinweise auf die Nützlichkeit der Determinante:

Satz 5.7. Seien $A, B \in \mathbb{K}^{n \times n}$ quadratische Matrizen. Elementare Zeilen-/Spaltenumformungen ändern die Determinante wie folgt:

- (i) Geht B durch Vertauschen zweier Zeilen oder Spalten aus A hervor, so ist $\det(B) = -\det(A)$.
- (ii) Geht B durch Multiplikation einer Zeile oder Spalte von A mit $\lambda \in \mathbb{K}$ hervor, so gilt $\det(B) = \lambda \det(A)$. Insbesondere ist damit $\det(\lambda \cdot A) = \lambda^n \cdot \det(A)$ für $\lambda \in \mathbb{K}$.
- (iii) Geht B aus A durch Addition des λ -fachen einer Zeile/Spalte zu einer anderen Zeile/Spalte hervor, so ist $\det(B) = \det(A)$.

Weiterhin gilt:

- (iv) $\det(A^T) = (\det(A))^T$, d.h. Transponieren ändert den Wert der Determinante nicht.
- (v) $\det(A) \cdot \det(B) = \det(A \cdot B)$.
- (vi) $\det(A) \neq 0$ genau dann, wenn A invertierbar ist.
- (vii) Falls A invertierbar ist, so gilt $\det(A^{-1}) = (\det(A))^{-1}$.

Beispiel 5.8. Berechnen wir die Determinante der Matrix, deren Spalten aus \mathcal{M}_4 aus Beispiel 1.29 bestehen. Nach der Regel von Sarrus ist

$$\det \begin{pmatrix} 1 & 0 & 2 \\ 4 & -1 & 3 \\ 6 & 3 & 2 \end{pmatrix} = -2 + 24 + 0 - (-12) - 9 - 0 = 25.$$

Da die Determinante nicht Null ist, ist die Matrix invertierbar. Eine invertierbare Matrix ist regulär; sie hat also in diesem Fall Rang 3. Folglich sind alle Spaltenvektoren linear unabhängig, was wiederum beweist, dass \mathcal{M}_4 in der Tat eine Basis des \mathbb{R}^3 ist.

Betrachten wir noch einen besonders einfachen Sonderfall einer Determinante. Durch wiederholte Entwicklung nach der ersten Spalte kann man folgende Aussage beweisen:

Satz 5.9. Die Matrix $A \in \mathbb{K}^{n \times n}$ liege in der Zeilenstufenform vor, d.h. $\alpha_{k,j} = 0$ für $k > j$. Dann ist

$$\det(A) = \prod_{k=1}^n \alpha_{k,k}.$$

5.2 Anwendung: Lösung von LGS mit der Cramer'schen Regel

Wir wissen bereits, dass das LGS $Ax = b$ für $A \in \mathbb{K}^{n \times n}$ und $b \in \mathbb{K}^n$ genau dann **genau** eine Lösung hat, wenn der Rang der erweiterten Matrix gleich dem der Matrix ist **und der Lösungsraum des homogenen LGS Dimension 0 hat**, d.h. falls $\text{rg}(A|b) = \text{rg}(A) = n$ ist, also wenn A regulär ist, d.h. wenn $\det(A) \neq 0$ ist.

Satz 5.10 (Cramer'sche Regel). Sei $A \in \mathbb{K}^{n \times n}$ eine reguläre quadratische Matrix und sei $b \in \mathbb{K}^n$. Sei weiterhin $A_k(b)$ die Matrix, die entsteht, wenn man die k -te Spalte von A durch den Vektor b ersetzt. Die Lösung des LGS $Ax = b$ ist dann durch den Vektor $x = (x_1, \dots, x_n)^T \in \mathbb{K}^{n \times 1}$ gegeben mit

$$x_k = \frac{\det(A_k(b))}{\det(A)}, \quad k \in \{1, \dots, n\}.$$

Beispiel 5.11. Sei wieder $A = \begin{pmatrix} 1 & 0 & 2 \\ 4 & -1 & 3 \\ 6 & 3 & 2 \end{pmatrix}$. Sei zudem $b = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Dann ist $\text{rg}(A_k(b)) = 2$ für $k \in \{1, 2, 3\}$, da immer eine Spalte der Matrix Null ist. Folglich ist in allen diesen Fällen $\det(A_k(b)) = 0$. Wir haben bereits $\det(A) = 25$ berechnet. Somit ist die einzige Lösung von $Ax = b$ gegeben durch

$$x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Dies ist nicht überraschend, da die Spalten von A linear unabhängig sind.

Beispiel 5.12. Betrachten wir das LGS

$$\begin{cases} x_1 + 3x_2 = 2 \\ -x_1 + 2x_2 = 1 \end{cases}$$

Es lässt sich übersetzen in $Ax = b$ mit

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

Die Determinanten für die Cramer'sche Regel sind

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} = 5, \\ \det(A_1(b)) &= \det \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} = 1, \\ \det(A_2(b)) &= \det \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = 3. \end{aligned}$$

Damit ist die Lösung des LGS $x = \left(\frac{1}{5}, \frac{3}{5}\right)^T$.

5.3 Anwendung: Berechnung der Inversen einer Matrix

Bereits bei dem Entwicklungssatz von Laplace (Satz 5.4) sind uns Koeffizienten der Form

$$\gamma_{k,j} := (-1)^{k+j} \det(A_{k,j}), \quad k, j \in \{1, \dots, n\} \quad (5.1)$$

für eine Matrix $A = (\alpha_{k,j})$ begegnet.

Definition 5.13. Sei $A = (\alpha_{k,j}) \in \mathbb{K}^{n \times n}$ und sei $C = (\gamma_{k,j}) \in \mathbb{K}^{n \times n}$ mit Einträgen wie in (5.1) definiert. Die Einträge $\gamma_{k,j}$ heißen Kofaktoren von A , die Matrix C die Kofaktorenmatrix von A und deren Transponierte, C^T , wird Adjunkte von A genannt.

Satz 5.14. Sei $A \in \mathbb{K}^{n \times n}$ regulär und sei C^T die Adjunkte von A . Dann gilt

$$A^{-1} = \frac{1}{\det(A)} C^T.$$

Beispiel 5.15 (2×2 -Matrix). Sei A wie in Beispiel 5.5, d.h. $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist nach obigem Satz und mit der bereits in Beispiel 5.5 berechneten 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 5.16. Sei

$$A = \begin{pmatrix} 1 & 0 & -2 \\ -1 & 1 & 2 \\ 0 & 4 & -1 \end{pmatrix}.$$

Zunächst einmal ist A invertierbar, denn

$$\det(A) = \begin{vmatrix} 1 & 0 & -2 \\ -1 & 1 & 2 \\ 0 & 4 & -1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 2 \\ 4 & -1 \end{vmatrix} - 2 \cdot \begin{vmatrix} -1 & 1 \\ 0 & 4 \end{vmatrix} = -9 + 8 = -1 \neq 0.$$

Berechnen wir nun die Kofaktoren:

$$\begin{aligned} \gamma_{1,1} &= + \begin{vmatrix} 1 & 2 \\ 4 & -1 \end{vmatrix} = -9 & \gamma_{1,2} &= - \begin{vmatrix} -1 & 2 \\ 0 & -1 \end{vmatrix} = -1 & \gamma_{1,3} &= + \begin{vmatrix} -1 & 1 \\ 0 & 4 \end{vmatrix} = -4 \\ \gamma_{2,1} &= - \begin{vmatrix} 0 & -2 \\ 4 & -1 \end{vmatrix} = -8 & \gamma_{2,2} &= + \begin{vmatrix} 1 & -2 \\ 0 & -1 \end{vmatrix} = -1 & \gamma_{2,3} &= - \begin{vmatrix} 1 & 0 \\ 0 & 4 \end{vmatrix} = -4 \\ \gamma_{3,1} &= + \begin{vmatrix} 0 & -2 \\ 1 & 2 \end{vmatrix} = 2 & \gamma_{3,2} &= - \begin{vmatrix} 1 & -2 \\ -1 & 2 \end{vmatrix} = 0 & \gamma_{3,3} &= + \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1 \end{aligned}$$

Damit ist

$$A^{-1} = - \begin{pmatrix} -9 & -1 & -4 \\ -8 & -1 & -4 \\ 2 & 0 & 1 \end{pmatrix}^T = \begin{pmatrix} 9 & 8 & -2 \\ 1 & 1 & 0 \\ 4 & 4 & -1 \end{pmatrix}.$$

5.4 Anwendung im Spezialfall 2D / 3D: Fläche / Volumen berechnen

Satz 5.17.

- (i) Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine lineare Abbildung mit zugehöriger Matrix A , d.h. $f(x) = Ax$. Ist $M \subset \mathbb{R}^2$ eine Menge mit Flächeninhalt $|M|$, dann ist der Flächeninhalt von $f(M)$ gegeben durch $|\det(A)| \cdot |M|$.
- (ii) Sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ eine lineare Abbildung mit zugehöriger Matrix A , d.h. $f(x) = Ax$. Ist $M \subset \mathbb{R}^3$ eine Menge mit Volumen $|M|$, dann ist das Volumen von $f(M)$ gegeben durch $|\det(A)| \cdot |M|$.

Beispiel 5.18. Sei $M = \{(x, y)^T \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ der Einheitskreis im \mathbb{R}^2 . Sei eine lineare Abbildung $f = f_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch die Matrix $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$. Dann ist

$$\begin{aligned} f(M) &= \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2 \mid A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} \text{ für } \begin{pmatrix} x \\ y \end{pmatrix} \in M \right\} \\ &= \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2 \mid ax = u, by = v, x^2 + y^2 \leq 1 \right\} \\ &= \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2 \mid \frac{u^2}{a^2} + \frac{v^2}{b^2} \leq 1 \right\}. \end{aligned}$$

$f(M)$ ist also eine Ellipse um den Mittelpunkt $(0, 0)^T$ mit Halbachsen der Längen a und b . Der Flächeninhalt des Einheitskreises ist bekanntlich π . Nach obigem Satz ist der Flächeninhalt der Ellipse

$$|f(A)| = |\det(A)| \cdot \pi = ab \cdot \pi.$$

5.5 Anwendung: Polynominterpolation — Vandermonde-Determinante

Gegeben seien drei Punkte

$$P_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \text{und} \quad P_3 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}.$$

Wir suchen eine quadratische Funktion $f(x) = ax^2 + bx + c$, so dass alle drei Punkte auf dem Graphen der Funktion liegen. Es soll also gelten:

$$\begin{cases} (-1)^2 \cdot a + (-1) \cdot b + c &= 1, \\ 1^2 \cdot a + 1 \cdot b + c &= 2, \\ 3^2 \cdot a + 3 \cdot b + c &= 6. \end{cases}$$

Um zu sehen, ob dies überhaupt möglich ist, berechnen wir die Determinante der Koeffizientenmatrix:

$$\begin{vmatrix} (-1)^2 & -1 & 1 \\ 1^2 & 1 & 1 \\ 3^2 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 9 & 3 & 1 \end{vmatrix} = 1 + 3 + (-9) - 9 - 3 - (-1) = -16$$

Das LGS ist also (eindeutig) lösbar, da die Matrix regulär, also invertierbar ist. In der Tat gilt dies immer für n gegebene Punkte und ein Polynom f vom Grad $n - 1$:

Die Koeffizientenmatrix für n Punkte mit Koordinaten (x_k, y_k) und ein Polynom $f(x) = a_0 + a_1x^1 + a_2x^2 + \dots + a_{n-1}x^{n-1}$ lautet

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix}.$$

Diese Matrix heißt *Vandermonde-Matrix* und ihre Determinante ist

$$\begin{aligned}
 \det(A) &= \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} && \text{Spalte } j - x_1 \text{ Mal Spalte } j - 1 \\
 &= \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_2x_1 & \dots & x_2^{n-1} - x_2^{n-2}x_1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n - x_1 & x_n^2 - x_nx_1 & \dots & x_n^{n-1} - x_n^{n-2}x_1 \end{vmatrix} && \text{Entwicklung nach 1. Zeile} \\
 &= 1 \cdot \begin{vmatrix} x_2 - x_1 & (x_2 - x_1)x_2 & \dots & (x_2 - x_1)x_2^{n-2} \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & (x_n - x_1)x_n & \dots & (x_n - x_1)x_n^{n-2} \end{vmatrix} && \text{zeilenweise ausklammern} \\
 &= (x_2 - x_1)(x_3 - x_1) \cdot \dots \cdot (x_n - x_1) \cdot \begin{vmatrix} 1 & x_2 & \dots & x_2^{n-2} \\ 1 & x_3 & \dots & x_3^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n^{n-2} \end{vmatrix} \\
 &= \dots \\
 &= \prod_{k>j} (x_k - x_j)
 \end{aligned}$$

Diese ist genau dann von Null verschieden, wenn alle x_k verschieden sind.

Schlussfolgerungen

- Wir können genau dann ein Polynom vom Grad $n - 1$ finden, so dass n gegebene Punkte auf dessen Graphen liegen, wenn die x -Koordinaten der Punkte alle verschieden sind.
- Die Determinante zeigt, dass die Polynome $P := \{1, x, x^2, x^3, \dots, x^n\}$ in der Tat linear unabhängig sind. Ist \mathcal{P}_n der Vektorraum der Polynome vom Grad $\leq n$, so ist P ein linear unabhängiges Erzeugendensystem von \mathcal{P}_n , also eine Basis. Folglich ist, wie wir bereits in den Übungen behauptet haben, $\dim(\mathcal{P}_n) = n + 1$.

6 Eigenwertprobleme

Im Folgenden werden wir den Nullvektor mit $\mathbf{0}$ bezeichnen, um ihn von der Zahl $0 \in \mathbb{R}$ unterscheiden zu können.

6.1 Definition, Berechnung, Eigenschaften

Beispiel 6.1. Sei $f = f_A$ mit $A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$ und seien $u = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Berechnen wir die Funktionswerte beider Vektoren, so erhalten wir

$$f(u) = A \cdot u = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \quad \text{und} \quad f(v) = A \cdot v = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3v.$$

Während u und $A \cdot u$ in unterschiedliche Richtungen zeigen, hat v die Eigenschaft, dass er in die gleiche Richtung zeigt wie sein Bild, $A \cdot v$. Vektoren mit dieser Eigenschaft werden wir im Folgenden näher betrachten.

Sei im Folgenden V ein n -dimensionaler \mathbb{K} -Vektorraum und sei $f: V \rightarrow V$ eine lineare Abbildung. Falls $V = \mathbb{K}^n$ ist, so sei $A \in \mathbb{K}^{n \times n}$ die zu $f = f_A$ gehörige Matrix, d.h. $f(v) = A \cdot v$ für $v \in \mathbb{K}^n$.

Definition 6.2. Ein Vektor $v \in V \setminus \{\mathbf{0}\}$ heißt Eigenvektor von f , falls $\lambda \in \mathbb{K}$ existiert, so dass

$$f(v) = \lambda v$$

gilt. Der Skalar λ heißt dann Eigenwert zum Eigenvektor v .

Ist $V = \mathbb{K}^n$ und $f = f_A$, so spricht man auch von Eigenvektoren und -werten der Matrix A .

Die Menge der Eigenvektoren zu einem fest gewählten Eigenwert ergänzt um den Nullvektor $\mathbf{0} \in V$ heißt Eigenraum und man schreibt

$$\begin{aligned} \text{Eig}(f, \lambda) &= \{v \in V \mid f(v) = \lambda v\} \quad \text{bzw.} \\ \text{Eig}(A, \lambda) &= \{v \in \mathbb{K}^n \mid A \cdot v = \lambda v\}. \end{aligned}$$

Das Eigenwertproblem besteht darin, alle Eigenwerte und -vektoren einer linearen Abbildung zu bestimmen.

Beispiel 6.3. Sei $f(x) = A \cdot x$ mit $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ und $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Dann ist

$$A \cdot x = A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 2x_2 \end{pmatrix}.$$

Damit $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ ein Eigenvektor ist, muss gelten:

$$A \cdot v = \begin{pmatrix} v_1 \\ 2v_2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \end{pmatrix} = \lambda v.$$

Ist $v_1 \neq 0$, dann muss $\lambda = 1$ sein und somit $2v_2 = v_2$, d.h. $v_2 = 0$.

Ist $v_1 = 0$, dann muss $\lambda = 2$ sein und v_2 darf nicht Null sein, ist ansonsten aber beliebig wählbar. A (und somit auch f) hat also die zwei Eigenwerte 1 und 2 und die zugehörigen Eigenvektoren sind von der Form $\begin{pmatrix} a \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ b \end{pmatrix}$ für beliebige $a, b \in \mathbb{R} \setminus \{0\}$. Somit ist

$$\text{Eig}(A, 1) = \left\{ a \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid a \in \mathbb{R} \right\} \quad \text{und} \quad \text{Eig}(A, 2) = \left\{ b \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid b \in \mathbb{R} \right\}.$$

Satz 6.4. Seien $V = \mathbb{K}^n$ und $A \in \mathbb{K}^{n \times n}$. Es gilt:

1. $\text{Eig}(A, \lambda)$ ist ein Untervektorraum von V .
2. $\text{Eig}(A, \lambda) = L(A - \lambda E_n, \mathbf{0})$.
3. Sind $\lambda, \mu \in \mathbb{K}$ zwei verschiedene Eigenwerte von A , so gilt $\text{Eig}(A, \lambda) \cap \text{Eig}(A, \mu) = \{\mathbf{0}\}$.
4. Sind v_1, \dots, v_m Eigenvektoren zu verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m$, so sind v_1, \dots, v_m linear unabhängig. Insbesondere gilt $m \leq n$.

Ein paar Bemerkungen zu der Interpretation der Aussagen des Satzes:

- Aussage 1 besagt, dass Summen von Eigenvektoren und Vielfache von Eigenvektoren selbst wieder Eigenvektoren zum gleichen Eigenwert sind.
- Aussage 2 ist für uns nützlich, da wir somit nur dann einen Eigenvektor finden können, wenn das homogene LGS $A \cdot x = \mathbf{0}$ mehr als nur den Nullvektor enthält. Dies werden wir gleich zum Bestimmen der Eigenwerte ausnutzen.
- Aussage 3 besagt, dass kein Eigenvektor gleichzeitig zu zwei verschiedenen Eigenwerten gehören kann.
- Aussage 4 folgt aus Aussage 3, denn je zwei Eigenräume haben nur den Nullvektor gemeinsam und somit sind ihre Basen disjunkt.

Definition 6.5. Ist $A \in \mathbb{K}^{n \times n}$ und E_n die Einheitsmatrix gleicher Dimensionen, so heißt

$$p_A(t) := \det(A - tE_n), \quad t \in \mathbb{K},$$

das charakteristische Polynom der Matrix A .

Satz 6.6. Die Matrix $A \in \mathbb{K}^{n \times n}$ hat genau dann den Eigenwert $\lambda \in \mathbb{K}$, wenn $\det(A - \lambda E_n) = 0$ ist, d.h. wenn λ eine Nullstelle des charakteristischen Polynoms von A ist.

Beispiel 6.7. Ist $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ wie in Beispiel 6.3, so lautet das charakteristische Polynom

$$p_A(t) = \det\left(\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} - t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} 1-t & 0 \\ 0 & 2-t \end{pmatrix} = (1-t)(2-t).$$

Die Nullstellen des Polynoms sind 1 und 2, welche wir in der Tat bereits als Eigenwerte von A identifiziert hatten.

Beispiel 6.8. Sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$ aus dem Einführungsbeispiel 6.1. Das charakteristische Polynom von A ,

$$p_A(t) = \det\begin{pmatrix} 1-t & 2 \\ 3 & -t \end{pmatrix} = (1-t) \cdot (-t) - 6 = t^2 - t - 6 = (t+2)(t-3),$$

hat die Nullstellen -2 und 3 — dies sind somit die Eigenwerte von A . Der Eigenraum $\text{Eig}(A, 3)$ ist $L(A - 3E_2, \mathbf{0})$, d.h. wir suchen Vektoren $x = (x_1, x_2)^T$, für die gilt:

$$(A - 3E_2) \cdot x = \begin{pmatrix} 1-3 & 2 \\ 3 & -3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2x_1 + 2x_2 \\ 3x_1 - 3x_2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Somit ist

$$\text{Eig}(A, 3) = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = x_2 \right\} = \mathcal{L}\left(\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}\right).$$

Analog berechnen wir den Eigenraum zum Eigenwert -2 als $L(A + 2E_2, \mathbf{0})$:

$$(A + 2E_2) \cdot x = \begin{pmatrix} 1+2 & 2 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + 2x_2 \\ 3x_1 + 2x_2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Somit ist

$$\text{Eig}(A, -2) = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid 3x_1 + 2x_2 = 0 \right\} \mathcal{L} \left(\left\{ \begin{pmatrix} 2 \\ -3 \end{pmatrix} \right\} \right).$$

Wie schon bei der Berechnung von Determinanten sind auch bei der Berechnung von Eigenwerten Dreiecksmatrizen, d.h. Matrizen in Zeilenstufenform, besonders einfach zu handhaben.

Lemma 6.9. Sei $A \in \mathbb{K}^{n \times n}$ eine Matrix in Zeilenstufenform. Dann sind die Eigenwerte der Matrix gerade die Einträge auf der Hauptdiagonalen.

Beispiel 6.10. Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$. Dann ist

$$p_A(t) = \det \begin{pmatrix} 1-t & 2 & 3 \\ 0 & 4-t & 5 \\ 0 & 0 & 6-t \end{pmatrix} = (1-t) \cdot (4-t) \cdot (6-t),$$

d.h. die Nullstellen des charakteristischen Polynoms sind 1, 4 und 6.

6.1.1 Invertierbarkeit

Die Matrix $A \in \mathbb{K}^{n \times n}$ habe den Eigenwert 0. Dann existiert ein Vektor $x \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ so dass $A \cdot x = 0x = \mathbf{0} \in \mathbb{K}^n$. Folglich ist die Matrix nicht invertierbar.

Sind jedoch alle Eigenwerte von Null verschieden, so existiert keine nicht-triviale Lösung von $A \cdot x = \mathbf{0} = 0 \cdot x$, d.h. die Matrix ist invertierbar.

Beide Erkenntnisse zusammen können wir wie folgt zusammenfassen:

Satz 6.11. Eine Matrix $A \in \mathbb{K}^{n \times n}$ ist genau dann invertierbar, wenn 0 kein Eigenwert von A ist.

6.2 Definitheit symmetrischer Matrizen

In diesem Abschnitt werden wir uns ausschließlich auf reelle symmetrische Matrizen beschränken, d.h. $A \in \mathbb{R}^{n \times n}$ mit $A = A^T$.

Satz 6.12. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann gilt:

1. Alle Eigenwerte von A sind reell.
2. Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander bezüglich des kanonischen Skalarproduktes.

Beispiel 6.13. Matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ aus Beispiel 6.7 hat die Eigenwerte $1, 2 \in \mathbb{R}$ und die Eigenräume sind in der Tat orthogonal zueinander:

$$\text{Eig}(A, 1) = \mathcal{L} \left(\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \right), \quad \text{Eig}(A, 2) = \mathcal{L} \left(\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \right).$$

Wir wissen bereits, dass es zu jeder linearen Abbildung $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$ eine eindeutig bestimmte Matrix $A \in \mathbb{K}^{n \times n}$ gibt, so dass $f(x) = A \cdot x$ ist für $x \in \mathbb{K}^n$.

Im Gegensatz dazu nennt man eine Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ eine *quadratische Form*, falls sie darstellbar ist als $f(x) = x^T \cdot A \cdot x$ für $x \in \mathbb{R}^n$ und eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$.

Beispiel 6.14. $f(x_1, x_2) = 3x_1^2 + 2x_2^2$ ist eine quadratische Form zur Matrix $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$:

$$x^T \cdot A \cdot x = (x_1 \quad x_2) \cdot \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \quad x_2) \cdot \begin{pmatrix} 3x_1 \\ 2x_2 \end{pmatrix} = 3x_1^2 + 2x_2^2.$$

Definition 6.15. Eine quadratische Form $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mit $f(x) = x^T Ax$ heißt

- positiv definit, wenn $f(x) > 0$ für alle $x \neq \mathbf{0}$ gilt;
- positiv semidefinit, wenn $f(x) \geq 0$ für alle $x \in \mathbb{R}^n$ gilt;
- negativ definit, wenn $f(x) < 0$ für alle $x \neq \mathbf{0}$ gilt;
- negativ semidefinit, wenn $f(x) \leq 0$ für alle $x \in \mathbb{R}^n$ gilt;
- indefinit, wenn $f(x)$ sowohl positive als auch negative Werte annimmt.

Andererseits heißt eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ positiv/negativ (semi-)definit oder indefinit, wenn die zugehörige quadratische Form $f(x) = x^T Ax$ diese Eigenschaft hat.

Beispiel 6.16. Die Matrix $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ aus dem vorherigen Beispiel ist positiv definit, da $f(x_1, x_2) = 3x_1^2 + 2x_2^2 > 0$ ist sofern $x \neq \mathbf{0}$ ist. Die Matrix $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ mit der zugehörigen quadratischen Form $f(x) = x^T Bx = x_1^2 - x_2^2$ ist hingegen indefinit, denn $f(1, 0) = 1 > 0$ und $f(0, 1) = -1 < 0$.

Die Definitheit einer Matrix erkennt man direkt an deren Eigenwerten:

Satz 6.17. Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Dann sind die Matrix und deren zugehörige quadratische Form $f(x) = x^T Ax$

- positiv definit genau dann, wenn alle Eigenwerte von A positiv sind;
- positiv semidefinit genau dann, wenn alle Eigenwerte von A nicht negativ sind;
- negativ definit genau dann, wenn alle Eigenwerte von A negativ sind;
- negativ semidefinit genau dann, wenn alle Eigenwerte von A nicht positiv sind;
- indefinit genau dann, wenn A sowohl positive als auch negative Eigenwerte hat.

Beispiel 6.18. Die Definitheit der quadratischen Form $f(x_1, x_2, x_3) = x_1^2 + 3x_2^2 + x_3^2 - 2x_1x_2 + 2\sqrt{2}x_2x_3$ ist nicht ganz so offensichtlich. Die zugehörige Matrix lautet

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 3 & \sqrt{2} \\ 0 & \sqrt{2} & 1 \end{pmatrix}.$$

Berechnen wir das charakteristische Polynom von A :

$$\begin{aligned} \det(A - tE_3) &= \begin{vmatrix} 1-t & -1 & 0 \\ -1 & 3-t & \sqrt{2} \\ 0 & \sqrt{2} & 1-t \end{vmatrix} \\ &= (1-t) \begin{vmatrix} 3-t & \sqrt{2} \\ \sqrt{2} & 1-t \end{vmatrix} + 1 \begin{vmatrix} -1 & \sqrt{2} \\ 0 & 1-t \end{vmatrix} \quad \text{durch Entw. nach 1. Zeile} \\ &= (1-t) \cdot (3-t)(1-t) - (1-t) \cdot 2 + 1 \cdot (-1)(1-t) \\ &= (1-t) \cdot (3-4t+t^2-2-1) \\ &= (1-t) \cdot (t^2-4t) \\ &= (1-t) \cdot (t-4) \cdot t, \end{aligned}$$

d.h. die Eigenwerte von A sind 1 , 4 und 0 . Somit ist A (und folglich auch f) positiv semidefinit.

Bevor wir zur Anwendung der Definitheit kommen, gibt es noch eine weitere Möglichkeit, die Definitheit einer Matrix zu bestimmen. Dazu betrachten wir Untermatrizen und deren Determinanten.

Sei dazu $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Bezeichnen wir mit $A_k \in \mathbb{R}^{k \times k}$ die Untermatrix von A , die entsteht, indem man die letzten $n-k$ Zeilen und Spalten von A streicht. Sei $\Delta_k := \det(A_k)$ die Determinante der entsprechenden Untermatrix, genannt *Unterdeterminante k -ter Ordnung*.

Satz 6.19. Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix mit zugehöriger quadratischer Form $f(x) = x^T A x$. Es gelten folgende Aussagen:

1. f ist genau dann positiv definit, wenn alle Unterdeterminanten von A positiv sind.
2. f ist genau dann negativ definit, wenn alle Unterdeterminanten von $-A$ positiv sind, d.h. wenn alle Unterdeterminanten ungerader Ordnung negativ und alle Unterdeterminanten gerader Ordnung positiv sind.
3. Ist $\det(A) \neq 0$ und gelten weder 1. noch 2., so ist f indefinit.
4. Gibt es gerade Unterdeterminanten von A , die negativ sind, so ist f indefinit.

Beispiel 6.20 (Fortsetzung von Beispiel 6.18). In Beispiel 6.18 ist

$$A_3 = A, \quad A_2 = \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix} \quad \text{und} \quad A_1 = (1).$$

und die zugehörigen Determinanten sind $\Delta_3 = 0$, $\Delta_2 = 2$ und $\Delta_1 = 1$.

6.2.1 Anwendung: Extremwertaufgaben

In diesem Abschnitt wollen wir Extremwertaufgaben mit und ohne Nebenbedingungen lösen. Zunächst geht es um die Frage, wie man lokale Extrema reellwertiger Funktionen in zwei Variablen findet. Formulieren wir zunächst einen Satz gänzlich ohne die Begriffe der linearen Algebra.

Satz 6.21. Seien $M \subset \mathbb{R}^2$ offen und $f \in C^2(M; \mathbb{R})$ (d.h. zweimal stetig differenzierbar). Sei $a = (a_1, a_2) \in M$ ein kritischer Punkt von f .

a) Falls

$$\left(\frac{\partial^2 f}{\partial x \partial y}(a) \right)^2 < \frac{\partial^2 f}{\partial x^2}(a) \cdot \frac{\partial^2 f}{\partial y^2}(a) \quad (6.1)$$

ist, so ist a ein eine lokale Extremalstelle von f .

b) Ist $\frac{\partial^2 f}{\partial x^2}(a) > 0$, dann ist a ein lokales Minimum von f .

c) Ist $\frac{\partial^2 f}{\partial x^2}(a) < 0$, dann ist a ein lokales Maximum von f .

d) Gilt " $>$ " an Stelle von " $<$ " in (6.1), so ist a ein Sattelpunkt von f .

Wie können wir dieses Ergebnis mit Hilfe der neuen Begriffe formulieren?

Die Hesse-Matrix einer Funktion $f \in C^2(M; \mathbb{R})$ ist gerade

$$H_f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

und da f nach Voraussetzung zweimal stetig differenzierbar ist, ist $H_f(x, y)$ nach dem Satz von Schwarz symmetrisch, d.h.

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$

Die Determinante der Hesse-Matrix in einem Punkt $a \in \mathbb{R}^2$ ist

$$\det H_f(a) = \frac{\partial^2 f}{\partial x^2}(a) \cdot \frac{\partial^2 f}{\partial y^2}(a) - \frac{\partial^2 f}{\partial y \partial x}(a) \cdot \frac{\partial^2 f}{\partial x \partial y}(a).$$

Auf Grund der Symmetrie bedeutet Bedingung (6.1), dass $\det(A) = \det(A_2) > 0$ ist für $A = H_f(a)$. Andererseits ist mit dieser Notation $A_1 = \frac{\partial^2 f}{\partial x^2}(a)$, d.h. b) ist erfüllt, falls $\det(A_1) > 0$ ist und c) ist erfüllt, falls $\det(A_1) < 0$ ist. Im ersten Fall ist A positiv definit, im zweiten Fall negativ definit. d) besagt, dass (mit Punkt 4 des obigen Satzes) A indefinit ist.

Fassen wir diese Beobachtungen in einem Satz zusammen, wobei wir gleich auf mehr als 2 Dimensionen verallgemeinern:

Satz 6.22 (Optimierung ohne Nebenbedingungen). Sei $f \in C^2(M; \mathbb{R})$ für eine offene Menge $M \subset \mathbb{R}^n$. Dann gilt:

1. Jede Extremstelle $a \in \mathbb{R}^n$ von f ist kritischer Punkt von f , d.h. $\nabla f(a) = \mathbf{0}$.
2. Ist a ein kritischer Punkt von f und ist $H_f(a)$ positiv definit, so besitzt f in a ein lokales Minimum.
3. Ist a ein kritischer Punkt von f und ist $H_f(a)$ negativ definit, so besitzt f in a ein lokales Maximum.
4. Ist a ein kritischer Punkt von f und ist $H_f(a)$ indefinit, so besitzt f in a einen Sattelpunkt.

Beispiel 6.23. Sei $f(x, y, z) = (x - 1)^2 + (y + 2)^2 + (z + 1)^2$. Es ist

$$\nabla f(x, y, z) = \begin{pmatrix} 2(x - 1) \\ 2(y + 2) \\ 2(z + 1) \end{pmatrix} \quad \text{und} \quad H_f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Alle kritischen Punkte von f müssen $\nabla f(a) = \mathbf{0}$ erfüllen. Der einzige kritische Punkt ist somit $a = (1, -2, -1)^T$. Die Hesse-Matrix ist konstant (d.h. unabhängig von x, y, z) und wir erhalten die Unterdeterminanten

$$\Delta_3 = \det(H_f(a)) = 8, \quad \Delta_2 = \det \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4 \quad \text{und} \quad \Delta_1 = \det(2) = 2.$$

Somit ist $H_f(a)$ positiv definit, d.h. f besitzt in $a = (1, -2, -1)^T$ ein lokales Minimum.

Extrema unter Nebenbedingungen

Betrachten wir nun das Problem des Optimierens unter Nebenbedingung, d.h. wir suchen lokale Extrema einer Funktion $f \in C^2(M; \mathbb{R})$ mit einer offenen Menge $M \subset \mathbb{R}^n$ unter der Nebenbedingung dass $g_1(x) = \dots = g_m(x) = 0$ ist, wobei $m \leq n$ sei. Der Kürze halber schreiben wir $g := (g_1, \dots, g_m)^T: M \rightarrow \mathbb{R}^m$.

Auflösungsmethode Falls $g(x_1, \dots, x_n) = 0$ eindeutig nach x_n auflösbar ist, d.h. falls eine (optimalerweise differenzierbare) Funktion $G: \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ existiert so dass

$$g(x_1, \dots, x_{n-1}, G(x_1, \dots, x_{n-1})) \equiv 0$$

ist, dann maximiert/minimiert man die neue Zielfunktion $f(x_1, \dots, x_{n-1}, G(x_1, \dots, x_{n-1}))$.

Lagrange-Methode Wir erzeugen mit Hilfe einer Hilfsvariablen $\lambda = (\lambda_1, \dots, \lambda_m)^T \in \mathbb{R}^m$ (genannt *Lagrange-Multiplikator(en)*) die *Lagrange-Funktion*

$$\mathcal{L}(\lambda; x) := f(x) - \langle \lambda, g(x) \rangle = f(x) - \sum_{k=1}^m \lambda_k g_k(x).$$

Alle Punkte $(\lambda; x)^T = (\lambda_1, \dots, \lambda_m; x_1, \dots, x_n)^T$, für die $\nabla \mathcal{L}(\lambda; x) = \mathbf{0}$ gilt, sind kritische Punkte. Zudem fügen wir alle Punkte hinzu, in denen $\nabla g_1(x), \dots, \nabla g_m(x)$ linear abhängig sind und $g(x) = \mathbf{0}$ ist. Anschließend überprüfen wir, ob es sich um lokale Minima oder Maxima handelt. Dafür gibt es zwei Möglichkeiten, die die Hesse-Matrix ausnutzen:

- (i) Man bestimmt die Hesse-Matrix von $\tilde{f}(x) := \mathcal{L}(\lambda; x)$ (λ aus kritischem Punkt einsetzen). Ist diese Hesse-Matrix positiv definit, liegt ein lokales Minimum vor; ist sie negativ definit, liegt ein lokales Maximum vor.
- (ii) Man bildet die sog. *geränderte Hesse-Matrix* $H_{\mathcal{L}}(\lambda; x) \in \mathbb{R}^{(n+m) \times (n+m)}$. Falls $\text{sgn}(\Delta_k) = (-1)^m$ ist für alle $k \in \{2m+1, \dots, m+n\}$, so liegt ein lokales Minimum vor; falls $\text{sgn}(\Delta_k) = (-1)^{k-m}$ ist für alle $k \in \{2m+1, \dots, m+n\}$, so liegt ein lokales Maximum vor.

Beispiel 6.24. Wir wollen die lokalen Extrema der Funktion $f(x, y, z) = x^2 + y^2 + z^2$ unter der Nebenbedingung $3x + 2y + z = 7$ finden.

Auflösungsmethode Es ist $3x + 2y + z = 7 \iff z = 7 - 3x - 2y$. Somit suchen wir lokale Extrema von

$$\begin{aligned} \tilde{f}(x, y) &:= f(x, y, 7 - 3x - 2y) = x^2 + y^2 + (7 - 3x - 2y)^2 \\ &= 10x^2 - 42x + 12xy - 28y + 5y^2 + 49. \end{aligned}$$

Die kritischen Punkte finden wir durch die Bedingung

$$\nabla \tilde{f}(x, y) = \begin{pmatrix} 20x - 42 + 12y \\ 12x - 28 + 10y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} 20 & 12 \\ 12 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 42 \\ 28 \end{pmatrix}.$$

Dieses Gleichungssystem lässt durch Kürzen mit 2 darstellen als

$$\begin{pmatrix} 10 & 6 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 21 \\ 14 \end{pmatrix}.$$

Die Inverse der Koeffizientenmatrix ist $A^{-1} = \frac{1}{14} \begin{pmatrix} 5 & -6 \\ -6 & 10 \end{pmatrix}$. Somit ist

$$\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1} \cdot b = \frac{1}{2} \begin{pmatrix} 5 & -6 \\ -6 & 10 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ 1 \end{pmatrix}.$$

Der einzige kritische Punkt ist (mit $z = 7 - 3x - 2y = 7 - \frac{9}{2} - 2 = \frac{1}{2}$) gegeben durch

$(x, y, z)^T = (\frac{3}{2}, 1, \frac{1}{2})^T$. Die Hesse-Matrix $H_f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ ist (unabhängig von

(x, y, z)) positiv definit, so dass wir nur Minima finden können. f hat also unter der genannten Nebenbedingung ein lokales Minimum in $(x, y, z) = (\frac{3}{2}, 1, \frac{1}{2})$. Die Hesse-Matrix $H_{\tilde{f}}$ ist übrigens ebenfalls positiv definit.

Lagrange-Methode Die Lagrange-Funktion lautet

$$\mathcal{L}(\lambda; x, y, z) = x^2 + y^2 + z^2 - \lambda(3x + 2y + z - 7).$$

Kritische Punkte finden wir wie folgt:

$$\nabla \mathcal{L}(\lambda; x, y, z) = \begin{pmatrix} 7 - (3x + 2y + z) \\ 2x - 3\lambda \\ 2y - 2\lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} \lambda \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{3}{2} \\ 1 \\ \frac{1}{2} \end{pmatrix}.$$

(i) Es ist $\tilde{f}(x, y, z) = \mathcal{L}(1; x, y, z) = x^2 + y^2 + z^2 - 3x - 2y - z + 7$. Die zugehörige

Hesse-Matrix $H_{\tilde{f}}(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ ist positiv definit. Wir haben also ein lokales

Minimum gefunden.

(ii) Die geränderte Hesse-Matrix ist

$$H_{\mathcal{L}}(\lambda; x, y, z) = \begin{pmatrix} 0 & -3 & -2 & -1 \\ -3 & 2 & 0 & 0 \\ -2 & 0 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix}$$

Mit $m = 1$ und $n = 2$ müssen wir also die Vorzeichen der Determinanten $\Delta_{2m+1} = \Delta_3$ und Δ_4 dieser Matrix bestimmen.

$$\Delta_3 = \begin{vmatrix} 0 & -3 & -2 \\ -3 & 2 & 0 \\ -2 & 0 & 2 \end{vmatrix} = -8 - 18 = -26 < 0,$$

$$\Delta_4 = \begin{vmatrix} 0 & -3 & -2 & -1 \\ -3 & 2 & 0 & 0 \\ -2 & 0 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{vmatrix} = \begin{vmatrix} -3 & -2 & -1 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \end{vmatrix} + 2 \begin{vmatrix} 0 & -3 & -2 \\ -3 & 2 & 0 \\ -2 & 0 & 2 \end{vmatrix}$$

$$= -4 + 2 \cdot (-8 - 18) = -56 < 0.$$

Es ist also in der Tat $\text{sgn}(\Delta_k) = (-1)^m = (-1)^1 = -1$ für $k \in \{3, 4\}$, d.h. es liegt ein lokales Minimum vor.

Bemerkung 6.25. *Lauten die Nebenbedingungen $g_k(x) = c_k$ ($k \in \{1, \dots, m\}$), so setzt man $\tilde{g}_k(x) := g(x) - c$ und erhält somit wieder die gewünschte Form der Nebenbedingung, $\tilde{g}_k(x) = 0$ für $k \in \{1, \dots, m\}$.*

Bemerkung 6.26. *Sind die Nebenbedingungen keine Gleichungen, sondern Ungleichungen, so gibt es eine Methode, die der von Lagrange ähnelt. Der einzige entscheidende Unterschied ist der, dass die (vormals Lagrange-)Multiplikatoren, die zu den Ungleichungen gehören, ein festes Vorzeichen haben müssen. Die Methode nennt sich in diesem Fall Kuhn-Tucker-Methode.*

6.3 Diagonalisierbarkeit von Matrizen

Es gibt Klassen von Matrizen, die gewisse Eigenschaften teilen — man nennt solche Matrizen ähnlich.

Definition 6.27. Zwei Matrizen $A, B \in \mathbb{K}^{n \times n}$ heißen ähnlich, falls eine invertierbare Matrix $U \in \mathbb{K}^{n \times n}$ existiert, so dass $B = U^{-1}AU$ ist.

Insbesondere ist in diesem Fall auch $A = UBU^{-1}$, d.h. wenn A ähnlich zu B ist, dann ist auch B ähnlich zu A .

Ein Sonderfall der ähnlichen Matrizen sind die Matrizen, die zu einer Diagonalmatrix ähnlich sind.

Definition 6.28. Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt diagonalisierbar, falls eine invertierbare Matrix $U \in \mathbb{K}^{n \times n}$ existiert, so dass $U^{-1}AU$ Diagonalgestalt hat, d.h. es existieren $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ so dass

$$U^{-1}AU = \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Folgender Satz gibt uns einen Hinweis darauf, wofür Diagonalisierbarkeit nützlich sein könnte:

Satz 6.29 (Eigenschaften ähnlicher Matrizen). Seien A und B ähnliche Matrizen. Dann gilt:

1. $\text{rg}(A) = \text{rg}(B)$;
2. $\det(A) = \det(B)$; insbesondere ist A genau dann invertierbar, wenn B invertierbar ist;
3. $p_A = p_B$, d.h. A und B haben das gleiche charakteristische Polynom und somit die gleichen Eigenwerte.

Korollar 6.30. Ist $A \in \mathbb{K}^{n \times n}$ diagonalisierbar und somit zu $B = \text{diag}(\lambda_1, \dots, \lambda_n)$ ähnlich, dann ist $\det(A) = \det(B) = \lambda_1 \cdot \dots \cdot \lambda_n$ und die Eigenwerte von A sind gegeben durch $\lambda_1, \dots, \lambda_n$.

Kennt man die Eigenwerte einer Matrix, so lässt sich möglicherweise bereits sagen, ob eine Matrix diagonalisierbar ist.

Satz 6.31. Hat $A \in \mathbb{K}^{n \times n}$ genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.

Allerdings ist die Einheitsmatrix E_n offensichtlich diagonalisierbar (mit $U = U^{-1} = E_n$), obwohl sie nur einen einzigen Eigenwert hat, nämlich 1. Um ein geeigneteres Kriterium für Diagonalisierbarkeit zu finden, müssen wir ein wenig ausholen.

6.3.1 Vielfachheiten

Da das charakteristische Polynom einer Matrix $A \in \mathbb{K}^{n \times n}$ ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) vom Grad $\leq n$ ist und somit in \mathbb{C} genau n und in \mathbb{R} höchstens n Nullstellen hat, wissen wir automatisch, dass A höchstens n verschiedene Eigenwerte (in \mathbb{C}) haben kann. Bei weniger als n Eigenwerten muss eine Nullstelle von p_A mehrfach vorkommen.

Definition 6.32. Die Vielfachheit eines Eigenwertes λ als Nullstelle des charakteristischen Polynoms bezeichnen wir als algebraische Vielfachheit von λ . Die Dimension des Eigenraums $\text{Eig}(A, \lambda)$ bezeichnen wir als geometrische Vielfachheit von λ .

Bezeichnen wir die Eigenwerte mit $\lambda_1, \dots, \lambda_m$, so notieren wir im Folgenden kurz deren algebraische Vielfachheiten als $\alpha_1, \dots, \alpha_m$ und deren geometrische Vielfachheiten als β_1, \dots, β_m . Da jeder Eigenraum $\text{Eig}(A, \lambda_k)$ mindestens einen von Null verschiedenen Vektor enthält, ist die geometrische Vielfachheit immer mindestens 1. Eine Obergrenze für β_k liefert folgender Satz:

Satz 6.33. Die geometrische Vielfachheit eines Eigenwertes ist kleiner oder gleich der algebraischen Vielfachheit, d.h. $\beta_k \leq \alpha_k$ für alle $k \in \{1, \dots, m\}$.

Konkret berechnen kann man β_k auch über die Formel

$$\beta_k = \dim(L((A - \lambda_k E_n), \mathbf{0})) = n - \text{rg}(A - \lambda_k E_n).$$

Nun sind wir so weit, ein hinreichendes und notwendiges Kriterium für die Diagonalisierbarkeit zu formulieren:

Satz 6.34. Seien $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ sämtliche verschiedenen Eigenwerte von $A \in \mathbb{K}^{n \times n}$. (Es ist insb. $n \geq m$.) Es sei $\beta_k := \dim(\text{Eig}(A, \lambda_k))$ für $k \in \{1, \dots, m\}$. A ist genau dann diagonalisierbar, wenn $\beta_1 + \dots + \beta_m = n$ ist. Das ist genau dann der Fall, wenn algebraische und geometrische Vielfachheiten identisch sind, d.h. $\alpha_k = \beta_k$ für alle $k \in \{1, \dots, m\}$.

Beispiel 6.35. 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:

$$\begin{aligned} p_A(t) &= \det \begin{pmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3 \end{pmatrix} \\ &= t(2+t)(3-t) + 6 + 6 - (2(2+t) + 6t + 3(3-t)) \\ &= 6t + t^2 - t^3 + 12 - (4 + 2t + 6t + 9 - 3t) \\ &= -t^3 + t^2 + t - 1 \\ &= -(t-1)^2(t+1). \end{aligned}$$

Die Eigenwerte von A sind somit $\lambda_1 = 1$ mit algebraischer Vielfachheit $\alpha_1 = 2$ und $\lambda_2 = -1$ mit $\alpha_2 = 1$. Betrachten wir die Eigenräume, um die geometrischen Vielfachheiten zu bestimmen:

$\lambda_1 = 1$: Es ist

$$\text{rg}(A - 1E_3) = \text{rg} \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} = 1,$$

da alle Zeilen der Matrix Vielfache voneinander sind. Somit ist

$$\beta_1 = n - \text{rg}(A - \lambda_1 E_3) = 3 - \text{rg}(A - 1E_3) = 3 - 1 = 2.$$

$\lambda_2 = -1$: Es ist

$$\text{rg}(A + 1E_3) = \text{rg} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} = 2,$$

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.

Nach obigem Satz ist die Matrix A also diagonalisierbar.

6.3.2 Diagonalisierung

Kommen wir nun zu der Frage, wie man eine gegebene diagonalisierbare Matrix diagonalisiert. Dazu machen wir eine Reihe von Überlegungen, wobei weiterhin $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ gelte.

- Ist $A \in \mathbb{K}^{n \times n}$ diagonalisierbar und hat A die Eigenwerte $\lambda_1, \dots, \lambda_m$, so gilt $\beta_1 + \dots + \beta_m = n$, d.h. die Summe der Dimensionen der Eigenräume ist n .
- Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. (Satz 6.4)
- Jede Basis des Eigenraumes $\text{Eig}(A, \lambda_k)$ besteht aus β_k Eigenvektoren, die somit linear unabhängig voneinander sind.
- Fixiert man für jeden Eigenraum eine Basis B_k und vereinigt diese zu $B = B_1 \cup \dots \cup B_m$, so erhält man n linear unabhängige Eigenvektoren von A und somit ist B eine Basis des \mathbb{K}^n .

Wir behaupten nun Folgendes:

Die Matrix $U \in \mathbb{K}^{n \times n}$, die entsteht, wenn man alle Vektoren von B aneinanderreihet, diagonalisiert A , d.h. sind $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$ die Eigenwerte, gemäß ihrer algebr. Vielfachheit mehrfach aufgezählt in der gleichen Reihenfolge, in der die Eigenvektoren in U stehen, so ist $\text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n) = U^{-1} \cdot A \cdot U$.

Überzeugen wir uns anhand von Beispiel 6.35, dass die Behauptung stimmt.

Beispiel 6.36 (Fortsetzung von Beispiel 6.35).

$\lambda_1 = 1$: 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 $\text{Eig}(A, 1)$ 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 $x^T = (1, 3, 2)$. Eine Basis von $\text{Eig}(A, -1)$ ist somit

$$B_2 = \left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}.$$

Der Kandidat für die Matrix U ist somit

$$U = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}.$$

Die Inverse von U ist

$$U^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Letztlich ist in der Tat

$$\begin{aligned} U^{-1} \cdot A \cdot U &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}. \end{aligned}$$

6.4 Anwendung: Matrixfunktionen

Beispiel 6.37. Betrachten wir nochmals die Matrix $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ aus Beispiel 6.14. Wir wollen A^n für beliebiges $n \in \mathbb{N}$ berechnen. Zunächst einmal ist

$$\begin{aligned} A^2 &= \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix}, \\ A^3 &= A \cdot A^2 = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 27 & 0 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 3^3 & 0 \\ 0 & 2^3 \end{pmatrix}. \end{aligned}$$

Da auch

$$\begin{pmatrix} 3^{n-1} & 0 \\ 0 & 2^{n-1} \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3^n & 0 \\ 0 & 2^n \end{pmatrix}$$

ist, gilt nach dem mathematischen Beweisprinzip der vollständigen Induktion, dass in der Tat $A^n = \begin{pmatrix} 3^n & 0 \\ 0 & 2^n \end{pmatrix}$ ist für beliebiges $n \in \mathbb{N}$.

Satz 6.38.

1. Sei $A = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^{n \times n}$ eine Diagonalmatrix. Dann ist für beliebige Potenzen $k \in \mathbb{N}$

$$A^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k) = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix}.$$

2. Sei $A \in \mathbb{K}^{n \times n}$ diagonalisierbar, d.h. $U^{-1}AU = \text{diag}(\lambda_1, \dots, \lambda_n) =: D$. Dann gilt $A = UDU^{-1}$ und somit

$$A^k = (UDU^{-1})^k = UD^kU^{-1} = U \cdot \text{diag}(\lambda_1^k, \dots, \lambda_n^k) \cdot U^{-1}$$

für beliebige Potenzen $k \in \mathbb{N}$.

Der Satz kann auch erweitert werden auf Exponenten $k \notin \mathbb{N}$, z.B. $k \in \mathbb{Z}$ oder $k \in \mathbb{Q}$.

Beispiel 6.39. Betrachten wir die Matrix $A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$ aus Beispiel 6.35. Wir hatten bereits die Diagonalisierbarkeit von A überprüft und die Matrizen D , U und U^{-1} berechnet als

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \quad U^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Berechnen wir das Quadrat von A :

$$A^2 = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Andererseits ist ebenfalls (wie im Satz behauptet)

$$UD^2U^{-1} = U \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (-1)^2 \end{pmatrix} \cdot U^{-1} = U \cdot U^{-1} = E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Mit Hilfe der Potenzen von Matrizen können wir die Exponentialfunktion für Matrizen definieren. Dafür müssen wir uns an die Taylor-Entwicklung der Exponentialfunktion erinnern:

$$\text{Ist } f(x) = e^x, \text{ so ist } Tf(x; 0) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Die Exponentialfunktion ist analytisch, d.h. sie stimmt mit ihrer Taylorreihe überein: $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

Definition 6.40. Sei $A \in \mathbb{K}^{n \times n}$. Dann ist das Matrixexponential definiert als

$$e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Es gelten folgende Rechenregeln:

- Ist $A = \text{diag}(\lambda_1, \dots, \lambda_n)$, so ist $e^A = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$.
- Ist $O \in \mathbb{K}^{n \times n}$ eine Nullmatrix und E_n die Einheitsmatrix gleicher Größe, so ist $e^O = E_n$.
- Sind $A, B \in \mathbb{K}^{n \times n}$, so ist $e^{A+B} = e^A \cdot e^B$.
- Ist A eine quadratische Matrix, so ist $e^{A^T} = (e^A)^T$.
- Für jede quadratische Matrix A ist e^A invertierbar mit der Inversen $(e^A)^{-1} = e^{-A}$.
- Ist $A = UDU^{-1}$, so ist $e^A = Ue^DU^{-1}$.

Beispiel 6.41. Ist $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, dann ist $e^A = \begin{pmatrix} e^3 & 0 \\ 0 & e^2 \end{pmatrix}$.

Ist $A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$ aus Beispiel 6.35 mit der zu ihr ähnlichen Matrix D , so ist

$$e^D = \begin{pmatrix} e^1 & 0 & 0 \\ 0 & e^1 & 0 \\ 0 & 0 & e^{-1} \end{pmatrix} = \begin{pmatrix} e & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & \frac{1}{e} \end{pmatrix}.$$

Da $A = UDU^{-1}$ ist, ist also

$$\begin{aligned} e^A &= Ue^DU^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} e & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & \frac{1}{e} \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} e & 0 & 1/e \\ 0 & e & 3/e \\ e & e & 2/e \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e+1/e & -e+1/e & e-1/e \\ -3e+3/e & -e+3/e & 3e-3/e \\ -2e+2/e & -2e+2/e & 4e-2/e \end{pmatrix}. \end{aligned}$$

Das Verfahren kann auch angewendet werden, wenn die Matrix mit komplexen Eigenwerten diagonalisierbar ist:

Beispiel 6.42. Sei $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Das charakteristische Polynom ist $p_A(t) = t^2 + 1$ und hat die Nullstellen $\pm i$, welche somit komplexe Eigenwerte von A sind. Berechnen wir die zugehörigen Eigenräume:

$$\text{Eig}(A, i) = \{x \in \mathbb{C}^2 \mid (A - iE_2)x = 0\} = \{x \in \mathbb{C}^2 \mid -ix_1 + x_2 = 0\} = \mathcal{L} \left\{ \begin{pmatrix} 1 \\ i \end{pmatrix} \right\},$$

$$\text{Eig}(A, -i) = \{x \in \mathbb{C}^2 \mid (A + iE_2)x = 0\} = \{x \in \mathbb{C}^2 \mid ix_1 + x_2 = 0\} = \mathcal{L} \left\{ \begin{pmatrix} 1 \\ -i \end{pmatrix} \right\}.$$

Somit ist $U = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ mit zugehöriger Inverser

$$U^{-1} = \frac{1}{-2i} \begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix} = \frac{i}{2} \begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Damit ist in der Tat

$$U^{-1}AU = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} =: D.$$

Das Matrixexponential von At ($t \in \mathbb{R}$) ist somit

$$\begin{aligned} e^{At} &= Ue^{Dt}U^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \cdot \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \\ &= \begin{pmatrix} e^{it} & e^{-it} \\ ie^{it} & -ie^{-it} \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e^{it} + e^{-it} & -i(e^{it} - e^{-it}) \\ i(e^{it} - e^{-it}) & e^{it} + e^{-it} \end{pmatrix} \\ &= \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}, \end{aligned}$$

da $\cos t = \frac{1}{2}(e^{it} + e^{-it})$ und $\sin t = \frac{1}{2i}(e^{it} - e^{-it})$ ist.

Unter Umständen ist die Exponentialfunktion leicht berechenbar, obwohl die Matrix nicht invertierbar ist.

Beispiel 6.43. Sei $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. Die Matrix hat den dreifachen Eigenwert 0 und der zugehörige Eigenraum ist, da A nicht regulär ist,

$$\text{Eig}(A, 0) = \left\{ x \in \mathbb{R}^3 \mid Ax = 0 \right\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mid a \in \mathbb{R} \right\}.$$

Es ist also $\alpha_1 = 3$, jedoch $\beta_1 = 1$. Folglich ist A nicht diagonalisierbar. Berechnen wir die ersten Potenzen von A :

$$A^2 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$A^3 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Somit ist $A^k = O \in \mathbb{R}^{3 \times 3}$ (die Nullmatrix) für $k \geq 2$ und das Matrixexponential ist

$$\begin{aligned} e^A &= A^0 + A + \frac{1}{2}A^2 + O \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Ist $t \in \mathbb{R}$ ein beliebiger Parameter, so kann man auf diese Art auch e^{At} berechnen:

$$\begin{aligned} e^{At} &= (At)^0 + At + \frac{1}{2}(At)^2 + O \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} t + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} t^2 \\ &= \begin{pmatrix} 1 & t & t+t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Definition 6.44. Eine quadratische Matrix $A \in \mathbb{K}^n$ heißt nilpotent, wenn ein $k \in \mathbb{N}$ existiert, so dass $A^k = O$ ist, d.h. wenn eine Ihrer Potenzen die Nullmatrix ergibt.

Die Matrix aus obigem Beispiel ist nilpotent, was die Berechnung des Matrixexponentials reduziert hat auf die Berechnung einer endlichen Summe von Matrizen. Folgender Satz sagt uns, wie nilpotente Matrizen zu erkennen sind:

Satz 6.45. Folgende Aussagen sind äquivalent für eine Matrix $A \in \mathbb{K}^{n \times n}$:

- (i) A ist nilpotent.
- (ii) A ist ähnlich zu einer strikten Dreiecksmatrix, d.h. einer Dreiecksmatrix auf deren Hauptdiagonale nur Nullen stehen.
- (iii) Das charakteristische Polynom von A ist $p_A(t) = \pm t^n$.

7 Differentialgleichungen

7.1 Einführung

Eine Differentialgleichung (kurz DGL) ist eine Gleichung, in der neben einer oder mehreren Variablen auch eine Funktion und deren Ableitungen auftreten.

Beispiel 7.1. Ist f implizit gegeben durch $F(x, y) = 0$, d.h. $F(x, f(x)) = 0$ für alle $x \in \mathbb{R}$, dann gilt nach der Kettenregel

$$0 = \partial_x F(x, f(x)) \cdot 1 + \partial_y F(x, f(x)) \cdot f'(x),$$

d.h. f erfüllt die DGL

$$f'(x) = \frac{-\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))},$$

sofern der Nenner nicht verschwindet.

Beispielsweise wird durch $F(x, y) = 0$ mit $F(x, y) = x^2 - y^2 - 1$ (implizit) eine Hyperbel definiert. Es ist $\frac{\partial F}{\partial y}(x, y) = -2y \neq 0$ für $y \neq 0$. Wenn wir diese Funktion lokal nach y auflösen können (also für $|x| \geq 1$ und $y \neq 0$), so dass für $f(x) = y(x)$ die Gleichung $F(x, y(x)) = 0$ gilt, dann ist die Ableitung

$$f'(x) = y'(x) = -\frac{2x}{-2y} = \frac{x}{y}.$$

Wir können überprüfen, dass dies korrekt ist, indem wir f ausrechnen:

$$x^2 - y^2 - 1 = 0 \Rightarrow y = \pm \sqrt{x^2 - 1}$$

Setzen wir beispielsweise $f(x) = \sqrt{x^2 - 1} = (x^2 - 1)^{1/2}$, so ist die Ableitung von f gegeben durch

$$f'(x) = \frac{1}{2}(x^2 - 1)^{-1/2} \cdot 2x = \frac{x}{\sqrt{x^2 - 1}} = \frac{x}{f(x)}, \quad \text{d.h.} \quad y' = \frac{x}{y}.$$

Die letzte Gleichung ist eine Differentialgleichung.

Beispiel 7.2 (Einfaches Populationsmodell). Betrachten wir die Population als von der Zeit t abhängige Funktion $n(t)$, so kann man vereinfacht annehmen, dass die Population proportional zum gegenwärtigen Bestand wächst, d.h. $n'(t) = \alpha n(t)$ für eine Konstante $\alpha > 0$.

Man unterscheidet zwischen gewöhnlichen und partiellen Differentialgleichungen. Letztere hängen von mehreren Variablen ab und es treten partielle Ableitungen nach den jeweiligen Variablen auf. Wir werden ausschließlich gewöhnliche Differentialgleichungen betrachten.

Definition 7.3. Seien $n, N \in \mathbb{N}$, $M \subset \mathbb{R} \times \mathbb{R}^{(n+1)N}$ und $F: M \rightarrow \mathbb{R}^N$. Eine Gleichung der Form

$$F(t, y, y', \dots, y^{(n)}) = 0 \tag{*}$$

mit Variablen $t \in \mathbb{R}$ und $y, y', \dots, y^{(n)} \in \mathbb{R}^N$ heißt N -dimensionale DGL n -ter Ordnung. Eine n -mal differenzierbare Funktion $\varphi: I \subset \mathbb{R} \rightarrow \mathbb{R}^N$ heißt Lösung der DGL $*$, falls

$$F(t, \varphi(t), \varphi'(t), \dots, \varphi^{(n)}(t)) = 0 \quad \text{für alle } t \in I.$$

Das Intervall I ist das zur Lösung φ gehörende Existenzintervall. Zudem gilt folgende Einteilung in Typen von Differentialgleichungen:

- Die DGL $*$ heißt linear, falls F affin linear ist in den Variablen $y, y', \dots, y^{(n)}$.

- Die DGL \star heißt autonom, falls F nicht von t abhängt.
- Die DGL \star heißt explizit, falls sie sich schreiben lässt als

$$0 = F(t, y, y', \dots, y^{(n)}) = y^{(n)} - f(t, y, y', \dots, y^{(n-1)})$$

mit $f: U \subset \mathbb{R} \times \mathbb{R}^{nN} \rightarrow \mathbb{R}^N$.

- Ist $N = 1$, so spricht man von einer skalaren DGL, sonst spricht man von einem System von Differentialgleichungen.
- Schreibt man $M = I \times X$ für $I \subset \mathbb{R}$ und $X \subset \mathbb{R}^{(n+1)N}$, so nennt man X auch Phasen- oder Zustandsraum.

Satz 7.4. Sei $f: U \subset \mathbb{R} \times \mathbb{R}^{nN} \rightarrow \mathbb{R}^N$ gegeben. Dann ist die (explizite) N -dimensionale DGL n -ter Ordnung

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)})$$

äquivalent zum nN -dimensionalen System von Differentialgleichungen 1. Ordnung

$$\begin{aligned} y'_0 &= y_1 \\ y'_1 &= y_2 \\ &\vdots \\ y'_{n-1} &= f(t, y_0, y_1, \dots, y_{n-1}) \end{aligned}$$

mit $y_0, y_1, \dots, y_{n-1}: \mathbb{R} \rightarrow \mathbb{R}^N$.

Beispiel 7.5. Die DGL 2. Ordnung $y'' + \alpha y = 0$ ist äquivalent zum 2-dimensionalen System 1. Ordnung

$$\begin{aligned} y'_0 &= y_1 \\ y'_1 &= -\alpha y_0. \end{aligned}$$

Daher werden wir uns in diesem Kapitel auf Systeme 1. Ordnung beschränken.

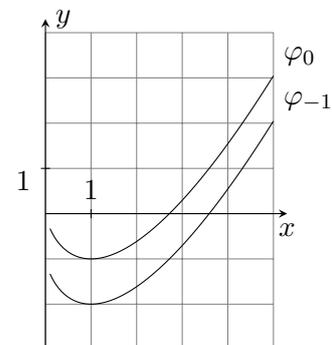
Definition 7.6 (AWP). Für $f: I \times X \subset \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ sei $y' = f(t, y)$ ein DGL-System 1. Ordnung. Seien $t_0 \in I$ und $y_0 \in X$. Dann heißt

$$y' = f(t, y), \quad y(t_0) = y_0$$

Anfangswertproblem (kurz AWP).

Beispiel 7.7.

Die DGL $y' = \ln t$ hat die Durch Integration leicht zu berechnende Lösung $\varphi_C(t) = t \ln(t) - t + C$ für $C \in \mathbb{R}$ und $t > 0$. Die Lösung ist also eine Schar von Funktionen mit von der Konstanten abhängigen Werten, z.B. $\varphi_C(e) = C$. Durch Festlegung eines Wertes (t_0, y_0) wählen wir eine Funktion aus dieser Lösungsschar aus.



7.2 Elementare Lösungsmethoden

Wir betrachten im Folgenden nur 1-dimensionale Differentialgleichungen 1. Ordnung.

7.2.1 Getrennte Variablen

Definition 7.8. Seien $I, J \subset \mathbb{R}$ und seien $g: I \rightarrow \mathbb{R}$ und $h: J \rightarrow \mathbb{R}$ stetig. Die Funktion $f: I \times J \rightarrow \mathbb{R}$ sei gegeben durch $f(t, y) = g(t)h(y)$. Dann heißt $y' = f(t, y) = g(t)h(y)$ DGL mit getrennten Variablen.

Ansatz:

Ist $h(y) \neq 0$ für alle $y \in J$, so folgt aus der Bedingung $y' = g(t)h(y)$ dass jede Lösung

$$\frac{\varphi'(t)}{h(\varphi(t))} = g(t)$$

erfüllt. Integrieren wir beide Seiten, so erhalten wir (mit Substitution $x = \varphi(s)$)

$$\int_{t_0}^t g(s) ds = \int_{t_0}^t \frac{1}{h(\varphi(s))} \varphi'(s) ds = \int_{\varphi(t_0)}^{\varphi(t)} \frac{1}{h(x)} dx$$

Die linke und rechte Seite obiger Gleichung hängen nur noch von t und $\varphi(t)$ ab. Ist die Stammfunktion von $\frac{1}{h}$ invertierbar, so kann man also nach φ umstellen.

Satz 7.9. Sei $(t_0, y_0) \in I \times J$ und $h(y) \neq 0$ für alle $y \in J$. Definiere die Funktionen $G(t) := \int_{t_0}^t g(s) ds$ und $H(y) := \int_{y_0}^y \frac{1}{h(x)} dx$. Ist $I' \subset I$ derart, dass $G(I') \subset H(J)$, dann besitzt das AWP

$$y' = g(t)h(y), \quad y(t_0) = y_0$$

genau eine Lösung in I' . Diese genügt der Beziehung $G(t) = H(\varphi(t))$.

Beispiel 7.10. In Beispiel 7.1 hatten wir für $F(t, y) = t^2 - y^2 - 1$ so die DGL $y' = \frac{t}{y}$ hergeleitet. Diese DGL können wir jetzt lösen:

1. Schreibe als DGL mit getrennten Variablen: $\frac{t}{y} = t \cdot \frac{1}{y}$, d.h. $g(t) = t$ und $h(y) = \frac{1}{y}$.
2. Bringe alles von y abhängige auf eine Seite, alles zeitabhängige auf die andere Seite:

$$y' = \frac{t}{y} \iff \frac{y'}{h(y)} = y \cdot y' = t.$$

3. Integration über $[t_0, t]$ auf beiden Seiten:

$$\begin{aligned} \int_{t_0}^t \frac{y'(s)}{h(y(s))} ds &= \int_{y(t_0)}^{y(t)} \frac{1}{h(x)} dx = \int_{y(t_0)}^{y(t)} x dx = \frac{1}{2} (y^2(t) - y^2(t_0)) \\ \int_{t_0}^t g(s) ds &= \int_{t_0}^t t dt = \frac{1}{2} (t^2 - t_0^2). \end{aligned}$$

4. Beides gleichsetzen (und $y(0) = y_0$):

$$\begin{aligned} \frac{1}{2} (y^2(t) - y^2(t_0)) &= \frac{1}{2} (t^2 - t_0^2) \\ \iff y^2(t) &= y_0^2 + (t^2 - t_0^2) \\ \iff \varphi(t) &= \pm \sqrt{t^2 + y_0^2 - t_0^2}, \quad t \in \mathbb{R} \setminus \left(-\sqrt{|y_0^2 - t_0^2|}, \sqrt{|y_0^2 - t_0^2|} \right) \end{aligned}$$

Mit dem Anfangswert $(t_0, y_0) = (1, 0)$ — dieser erfüllt in der Tat die Bedingung $F(t_0, y_0) = 0$ — erhalten wir $\varphi(t) = \pm \sqrt{t^2 - 1}$. Dies ist genau die Funktion, für die wir in Beispiel 7.1 die Differentialgleichung $y' = \frac{t}{y}$ bestimmt hatten.

7.2.2 Lineare DGL 1. Ordnung

Beispiel 7.11. Ist $y' = a(t) \cdot y$, d.h. eine lineare DGL 1. Ordnung, so berechnen wir analog die Lösungen für das AWP mit $y(t_0) = y_0$: (Umbenennung der Variablen von t in s , damit Variable und Integralgrenzen nicht identisch sind)

$$\frac{y'(t)}{y(t)} = a(t) \implies \int_{t_0}^t \frac{y'(s)}{y(s)} ds = \int_{t_0}^t a(s) ds \implies \ln(y(t)) - \ln(y_0) = \int_{t_0}^t a(s) ds$$

Mit den Rechenregeln für den Logarithmus ($\ln a - \ln b = \ln(\frac{a}{b})$) erhalten wir

$$\ln\left(\frac{y(t)}{y_0}\right) = \int_{t_0}^t a(s) ds \implies \boxed{y(t) = y_0 \cdot \exp\left(\int_{t_0}^t a(s) ds\right)} \quad (7.1)$$

Eine skalare lineare DGL 1. Ordnung ist von der Form

$$y' = a(t) \cdot y + b(t) \quad (7.2)$$

mit $t \in I \subset \mathbb{R}$ und $a, b: I \rightarrow \mathbb{R}$ stetig. Falls $b(t) \equiv 0$ ($\forall t \in I$), so heißt (7.2) *homogene lineare DGL 1. Ordnung*. Die Lösung der homogenen linearen DGL 1. Ordnung ist durch (7.1) gegeben. Die Lösung einer allgemeinen (nicht homogenen) linearen DGL 1. Ordnung berechnen wir mit folgendem Ansatz:

Ansatz – Variation der Konstanten: Sei ψ eine Lösung der homogenen linearen DGL $y' = a(t)y$, d.h.

$$\psi(t) = \psi(t_0) \cdot \exp\left(\int_{t_0}^t a(s) ds\right).$$

Dann ist für jede Konstante h auch $h \cdot \psi$ Lösung dieser DGL:

$$(h\psi)' = h \cdot \psi' = h \cdot a(t)\psi = a(t) \cdot (h\psi).$$

Nun ersetzen wir die Konstante h durch eine Funktion. (Wir "variieren die Konstante".) Sei $\boxed{\phi(t) := \psi(t)h(t)}$. Aus $\psi'(t) = a(t)\psi(t)$ folgt dann

$$\phi' = (\psi h)' = \psi' h + \psi h' = a\psi h + \psi h' = a\phi + \psi h'.$$

Damit ϕ Lösung von (7.2) ist, muss $\psi h' = b$ gelten, also $h' = \frac{b}{\psi}$. Setzen wir $\psi(t_0) = 1$, so erhalten wir durch Integration und Einsetzen von ψ :

$$h(t) = \int_{t_0}^t \frac{b(s)}{\psi(s)} ds + C = \int_{t_0}^t \psi(t_0)^{-1} \cdot \exp\left(-\int_{t_0}^s a(r) dr\right) b(s) ds + C.$$

Soll die Anfangsbedingung $y(t_0) = \phi(t_0) = y_0$ erfüllt sein, kann man $\boxed{\psi(t_0) = 1}$ setzen. Damit dann $\phi(t_0) = \psi(t_0)h(t_0)$ erfüllt ist, muss $C = y_0$ sein, d.h.

$$h(t) = y_0 + \int_{t_0}^t \exp\left(-\int_{t_0}^s a(r) dr\right) b(s) ds \quad \text{und damit}$$

$$\phi(t) = \exp\left(\int_{t_0}^t a(s) ds\right) \left(y_0 + \int_{t_0}^t \exp\left(-\int_{t_0}^s a(r) dr\right) b(s) ds\right).$$

Satz 7.12. Zu $t_0 \in I$ und $y_0 \in \mathbb{R}$ existiert genau eine Lösung $\phi: I \rightarrow \mathbb{R}$ des AWP

$$y'(t) = a(t)y(t) + b(t), \quad y(t_0) = y_0,$$

nämlich

$$\begin{aligned}\phi(t) &= \exp\left(\int_{t_0}^t a(s) ds\right) \left(y_0 + \int_{t_0}^t \exp\left(-\int_{t_0}^s a(r) dr\right) b(s) ds\right) \\ &= y_0 \exp\left(\int_{t_0}^t a(s) ds\right) + \int_{t_0}^t \exp\left(\int_s^t a(r) dr\right) b(s) ds.\end{aligned}$$

Beispiel 7.13. Gesucht ist die Lösung φ des AWP $y' = 2ty + t^3$ mit $y(0) = y_0$. Die homogene lineare DGL $y' = 2ty$ mit Anfangswert $y(0) = 1$ hat die Lösung

$$\psi(t) = 1 \cdot \exp\left(\int_0^t 2s ds\right) = e^{t^2}.$$

Zudem ist

$$h(t) = y_0 + \int_0^t e^{-s^2} s^3 ds.$$

Berechnen wir das Integral (mit Subst. $x = s^2$, $dx = 2s ds$ und danach part. Integration):

$$\begin{aligned}\int_0^t e^{-s^2} s^3 ds &= \int_0^{t^2} e^{-x} \cdot x \cdot \frac{1}{2} dx \\ &= \frac{1}{2} \int_0^{t^2} e^{-x} \cdot x dx = \frac{1}{2} [-e^{-x} x]_0^{t^2} + \frac{1}{2} \int_0^{t^2} e^{-x} dx \\ &= \frac{1}{2} [-e^{-x} x - e^{-x}]_0^{t^2} = \frac{1}{2} (1 - e^{-t^2} - t^2 e^{-t^2}).\end{aligned}$$

Damit ist insgesamt

$$\varphi(t) = \psi(t)h(t) = e^{t^2} \cdot \left(y_0 + \frac{1}{2} (1 - e^{-t^2} - t^2 e^{-t^2})\right) = \left(\frac{1}{2} + y_0\right) e^{t^2} - \frac{1}{2}(1 + t^2).$$

7.2.3 Homogene DGL

Definition 7.14 (homogene DGL). Seien $I \subset \mathbb{R} \setminus \{0\}$, $g: J \rightarrow \mathbb{R}$ und für $U := \{(t, y) \in I \times \mathbb{R} \mid \frac{y}{t} \in J\}$ sei $f: U \rightarrow \mathbb{R}$ mit $f(t, y) := g(\frac{y}{t})$ stetig. Dann heißt $y' = g(\frac{y}{t})$ homogene DGL.

Ansatz – Substitution: Sei φ eine Lösung der homogenen DGL $y' = g(\frac{y}{t})$ mit Anfangswert $y(t_0) = y_0$. Sei nun $z(t) := \frac{\varphi(t)}{t}$. Dann gilt

$$z' = \left(\frac{\varphi}{t}\right)' = \frac{\varphi' \cdot t - \varphi}{t^2} = \frac{\varphi'}{t} - \frac{\varphi}{t^2} = \frac{1}{t} \left(g\left(\frac{\varphi}{t}\right) - \frac{\varphi}{t}\right) = \frac{1}{t} (g(z) - z),$$

d.h. z ist Lösung der DGL $z' = \frac{1}{t} (g(z) - z)$ mit getrennten Variablen mit Anfangswert $z_0 = \frac{y_0}{t_0}$.

Satz 7.15. $\varphi: I \rightarrow \mathbb{R}$ ist genau dann eine Lösung von $y' = g(\frac{y}{t})$ mit Anfangswert $y(t_0) = y_0$, falls $z(t) = \frac{\varphi(t)}{t}$ Lösung der DGL $z' = \frac{1}{t} (g(z) - z)$ mit Anfangswert $z(t_0) = \frac{y_0}{t_0}$ ist.

Beispiel 7.16. Betrachte die DGL $y' = 1 + \frac{y}{t} + \frac{y^2}{t^2}$, d.h. nach obiger Notation ist $g(z) = 1 + z + z^2$. Setzen wir $z := \frac{y}{t}$, so gilt $z' = \frac{1}{t} (1 + z^2)$. Lösen wir diese DGL:

$$\begin{aligned}\frac{1}{t} &= \frac{z'}{1 + z^2} \text{ (mit } 1 + z^2 \neq 0 \text{ ergibt durch Integration)} \\ \ln\left(\frac{t}{t_0}\right) &= \int_{t_0}^t \frac{1}{s} ds = \int_{t_0}^t \frac{z'(s)}{1 + z^2(s)} ds = \int_{z(t_0)}^{z(t)} \frac{1}{1 + x^2} dx = \arctan(z(t)) - \arctan(z(t_0)).\end{aligned}$$

Umstellen nach $z(t)$ liefert

$$z(t) = \tan \left(\ln \left(\frac{t}{t_0} \right) + \arctan(z(t_0)) \right).$$

Mit Rücksubstitution von $z(t) = \frac{\varphi(t)}{t}$ und $z(t_0) = \frac{y_0}{t_0}$ erhält man die Lösung

$$\varphi(t) = t \cdot \tan \left(\ln \left(\frac{t}{t_0} \right) + \arctan \left(\frac{y_0}{t_0} \right) \right).$$

7.3 Existenz und Eindeutigkeit von Lösungen

Für diesen Abschnitt gelte wieder $n = 1$, d.h. wir betrachten nur (gewöhnliche) DGLen 1. Ordnung.

Satz 7.17 (Existenzsatz von Peano). Sei $f: I \times X \subset \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ stetig und sei $(t_0, y_0) \in I \times X$. Dann existiert ein $\alpha = \alpha(t_0, y_0) > 0$ so dass das AWP

$$y' = f(t, y), \quad y(t_0) = y_0 \tag{7.3}$$

auf dem Intervall $(t_0 - \alpha, t_0 + \alpha)$ (mindestens) eine Lösung besitzt.

Beispiel 7.18. Gegeben sei das AWP

$$y' = \sqrt{|y|}, \quad y(0) = 0.$$

Die Funktion $f(y) = \sqrt{|y|}$ ist auf ganz \mathbb{R} stetig. Dies ist eine DGL mit getrennten Variablen, wobei $g(t) \equiv 1$ und $h(y) = \sqrt{|y|}$ ist. Lösen wir diese:

1. Falls $y \neq 0$ ist, gilt

$$\begin{aligned} \int_{t_0}^t \frac{y'(s)}{\sqrt{|y(s)|}} ds &= \int_0^t 1 ds = t \\ \iff \int_{y(0)}^{y(t)} \frac{1}{\sqrt{|x|}} dx &= \int_0^{y(t)} |x|^{-1/2} dx = t \quad \text{mit Subst. } y(t) = x \\ \iff \left[2|x|^{1/2} \cdot \operatorname{sgn}(x) \right]_0^{y(t)} &= t \\ \iff \sqrt{|y(t)|} \cdot \operatorname{sgn}(y(t)) &= \frac{1}{2}t \\ \iff y(t) = \begin{cases} \frac{1}{4}t^2 & , t \geq 0, \\ -\frac{1}{4}t^2 & , t < 0. \end{cases} \end{aligned}$$

2. Die konstante Lösung $\varphi \equiv 0$ erfüllt auch $\varphi'(t) = 0 = \sqrt{|\varphi(t)|}$ und $\varphi(0) = 0$.

Eine DGL hat also nicht zwangsläufig eine eindeutige Lösung.

Beispiel 7.19. Gegeben sei das AWP

$$y' = y^2, \quad y(0) = y_0 > 0.$$

Die konstante Funktion $\varphi \equiv 0$ ist durch Wahl des Anfangswertes jetzt keine Lösung. Die (nicht-triviale) Lösung berechnen wir analog zum vorherigen Beispiel:

$$\begin{aligned} \int_0^t \frac{y'(s)}{y^2(s)} ds = t &\iff \int_{y(0)}^{y(t)} \frac{1}{x^2} dx = t \\ \iff -\frac{1}{y(t)} + \frac{1}{y_0} = t &\iff \frac{1}{y(t)} = \frac{1}{y_0} - t \iff y(t) = \frac{1}{\frac{1}{y_0} - t} \end{aligned}$$

Die Lösung $\varphi(t) := \frac{1}{\frac{1}{y_0} - t}$ existiert nur für $t < \frac{1}{y_0}$, da

$$\lim_{t \nearrow \frac{1}{y_0}} \varphi(t) = \lim_{t \nearrow \frac{1}{y_0}} \left(\frac{1}{\frac{1}{y_0} - t} \right) = +\infty$$

ist.

Definition 7.20 (Maximale Lösung). Eine Lösung $\varphi: I \rightarrow \mathbb{R}^N$ des AWP (7.3) heißt maximal, falls für jede andere Lösung $\psi: J \rightarrow \mathbb{R}^N$ gilt: $J \subset I$ und $\psi = \varphi|_J$.

Satz 7.21 (Picard-Lindelöf). Sei $f: I \times X \rightarrow \mathbb{R}^N$ stetig mit stetigen partiellen Ableitungen $\frac{\partial f}{\partial y_1}, \dots, \frac{\partial f}{\partial y_N}$. Dann besitzt das AWP (7.3) für jedes $(t_0, y_0) \in I \times X$ genau eine maximale Lösung $\varphi: J \rightarrow \mathbb{R}^N$. Das Intervall J ist offen.

Bemerkung 7.22 (Picard-Iteration). Die lokale Lösung des AWP (7.3) auf einem Intervall $I_\delta := (t_0 - \delta, t_0 + \delta)$ lässt sich iterativ bestimmen:

- $\varphi_0(t) \equiv y_0$ (Startwert);
- $\varphi_{n+1}(t) = y_0 + \int_{t_0}^t f(s, \varphi_n(s)) ds, t \in I_\delta$.

Beispiel 7.23. Betrachten wir das AWP $y' = 2ty, \quad y(0) = y_0$ mit $f(t, y) = 2ty$.

Startwert: $\varphi_0(t) = y_0$

1. Schritt: $\varphi_1(t) = y_0 + \int_0^t 2sy_0 ds = y_0 + 2 \cdot \frac{1}{2} y_0 t^2 = y_0(1 + t^2)$

2. Schritt: $\varphi_2(t) = y_0 + \int_0^t 2s \cdot y_0(1 + s^2) ds = y_0 \cdot \left(1 + \int_0^t 2(s + s^3) ds \right) = y_0 \cdot \left(1 + t^2 + \frac{1}{2} t^4 \right)$

Für $n \in \mathbb{N}$ führt dies auf $\varphi_n(t) = y_0 \cdot \left(\sum_{k=0}^n \frac{t^{2k}}{k!} \right)$. Für $n \rightarrow \infty$ erhalten wir die Lösung $\varphi(t) = y_0 \cdot \sum_{k=0}^{\infty} \frac{t^{2k}}{k!} = y_0 e^{t^2}$.

Mit den Methoden für DGLen mit getrennten Variablen können wir dies auch berechnen:

$$\int_0^t \frac{y'(s)}{y(s)} ds = \int_0^t 2s ds \implies \ln \left(\frac{y(t)}{y_0} \right) = t^2 \implies y(t) = y_0 e^{t^2}.$$

Bemerkung 7.24. Die Formulierung des Satzes von Picard-Lindelöf kann abgeschwächt werden. Statt stetiger Differenzierbarkeit der Funktion f genügt es, dass f lokal Lipschitz-stetig ist bezüglich y , d.h. zu jedem $(t_0, y_0) \in I \times X$ existieren $\delta, \varepsilon > 0$ so dass für $I' := (t_0 - \delta, t_0 + \delta) \subset I$ und $X' := (y_0 - \varepsilon, y_0 + \varepsilon) \subset X$ gilt: f ist auf $I' \times X'$ Lipschitz stetig, d.h. es existiert eine Konstante $L > 0$ so dass

$$|f(t, y) - f(t, y')| \leq L|y - y'|, \quad \forall t \in I', y, y' \in X'.$$

Die Funktion $f(t, y) = 2ty$ aus obigem Beispiel ist Lipschitz-stetig auf jedem abgeschlossenen Intervall $I = [a, b]$ mit $L = 2 \cdot \max\{|t| \mid t \in [a, b]\}$, denn

$$|f(t, y) - f(t, y')| = 2t|y - y'| \leq 2 \max\{|t| \mid t \in [a, b]\} |y - y'|.$$

Die Funktion $f(y) = \sqrt{|y|}$ hingegen ist nicht Lipschitz-stetig auf einem Intervall, welches den Punkt 0 enthält. Sie ist dort auch nicht stetig differenzierbar, denn $f'(y) = \text{sgn}(y) \cdot \frac{1}{2\sqrt{|y|}}$ ist unbeschränkt für $y \rightarrow 0$, also insb. nicht stetig.

7.4 Exakte Differentialgleichungen

Definition 7.25. Eine DGL vom Typ

$$\boxed{A(t,y) + B(t,y) y' = 0}, \quad t \in I \subset \mathbb{R}, y, y' \in X \subset \mathbb{R} \quad (7.4)$$

heißt exakt, falls eine Funktion $F: I \times X \rightarrow \mathbb{R}$ existiert, so dass für alle $(t, y) \in I \times X$ gilt:

$$\frac{\partial}{\partial t} F(t, y) = A(t, y), \quad \frac{\partial}{\partial y} F(t, y) = B(t, y).$$

Satz 7.26. Sind A und B stetig differenzierbar, so ist die DGL (7.4) exakt genau dann, wenn

$$\frac{\partial}{\partial y} A(t, y) = \frac{\partial}{\partial t} B(t, y)$$

Die DGL (7.4) ist also genau dann exakt, wenn das Vektorfeld

$$G(t, y) := \begin{pmatrix} A(t, y) \\ B(t, y) \end{pmatrix}$$

exakt ist. Das Potential ist dann durch F gegeben, d.h. $\nabla F(t, y) = G(t, y)$.

Satz 7.27. φ ist genau dann Lösung der DGL (7.4) mit Potential F , wenn $F(t, \varphi(t)) = c$ für eine Konstante $c \in \mathbb{R}$ gilt.

Beweis. φ erfüllt die Gleichung $F(t, \varphi(t)) = c$ genau dann, wenn $\frac{d}{dt} F(t, \varphi(t)) = 0$ ist. Mit der verallgemeinerten Kettenregel gilt andererseits

$$\frac{d}{dt} F(t, \varphi(t)) = \frac{\partial}{\partial t} F(t, \varphi(t)) + \frac{\partial}{\partial y} F(t, \varphi(t)) \varphi'(t) = A(t, \varphi(t)) + B(t, \varphi(t)) \varphi'(t).$$

Folglich erfüllt φ die exakte DGL (7.4) mit Potential F genau dann, wenn es die implizite Darstellung $F(t, y) = c$ besitzt. \square

Beispiel 7.28. Betrachten wir die DGL $t + 3y^2 \cdot y' = 0$, d.h. $A(t, y) = t$ und $B(t, y) = 3y^2$. Die Gleichung ist exakt:

$$\frac{\partial}{\partial y} A(t, y) = 0 = \frac{\partial}{\partial t} B(t, y).$$

Setzen wir $G(t, y) := \begin{pmatrix} t \\ 3y^2 \end{pmatrix}$. Das Potential der DGL, F , soll $\nabla F = G$ erfüllen. Wir berechnen es als $F(t, y) = \frac{1}{2}t^2 + y^3$. Die implizite Gleichung $F(t, y) = c$ hat die in diesem Fall leicht zu findende Lösung $\varphi(t) = \sqrt[3]{c - \frac{1}{2}t^2}$, welche nach obigem Satz auch die exakte DGL $t + 3y^2 \cdot y' = 0$ löst. Überprüfen wir dies kurz:

$$t + 3\varphi^2(t) \cdot \varphi'(t) = t + 3 \left(c - \frac{1}{2}t^2 \right)^{2/3} \cdot \frac{1}{3} \left(c - \frac{1}{2}t^2 \right)^{-2/3} \cdot (-t) = 0.$$

Falls eine Differentialgleichung von der Form $A(t, y) + B(t, y)y' = 0$ nicht exakt ist, so funktioniert möglicherweise folgender Ansatz:

Ansatz – Integrierender Faktor: Falls eine Funktion $\mu: I \times X \rightarrow \mathbb{R}$ existiert mit $\mu(t, y) \neq 0$ für alle $(t, y) \in I \times X$, dann ist

$$A(t, y) + B(t, y)y' = 0 \iff \mu(t, y)A(t, y) + \mu(t, y)B(t, y)y' = 0,$$

d.h. beide Differentialgleichungen besitzen die gleiche Lösung.

Ist die DGL $\mu(t, y)A(t, y) + \mu(t, y)B(t, y)y' = 0$ exakt, so nennt man μ *integrierenden Faktor*. Diesen findet man mit etwas Glück mit folgender Methode:

Definiere, sofern die jeweiligen Nenner nicht Null werden,

$$a(t, y) := \frac{\frac{\partial}{\partial y}A(t, y) - \frac{\partial}{\partial t}B(t, y)}{A(t, y)} \quad \text{und} \quad b(t, y) := \frac{\frac{\partial}{\partial y}A(t, y) - \frac{\partial}{\partial t}B(t, y)}{B(t, y)}.$$

- Ist $A(t, y) \neq 0$ und ist $a(t, y) = a(y)$, d.h. unabhängig von t , so setze $\mu(t, y) := \mu(y) := \exp\left(-\int a(y)dy\right)$.
- Ist $B(t, y) \neq 0$ und ist $b(t, y) = b(t)$, d.h. unabhängig von y , so setze $\mu(t, y) := \mu(t) := \exp\left(+\int b(t)dt\right)$.

Beispiel 7.29. Die DGL $t^2 + y - ty' = 0$ mit $A(t, y) = t^2 + y$ und $B(t, y) = -t$ ist nicht exakt, denn $\frac{\partial}{\partial y}A(t, y) = 1 \neq -1 = \frac{\partial}{\partial t}B(t, y)$. In diesem Fall ist $b(t, y) = \frac{2}{-t}$ unabhängig von y (wohldefiniert für $t \neq 0$) und damit lautet unser Kandidat für den integrierenden Faktor

$$\mu(t) := \exp\left(\int -\frac{2}{t}dt\right) = \exp(-2\ln t + C) = \exp(\ln(t^{-2}) + C) = \tilde{C} \cdot t^{-2}$$

für $\tilde{C} = e^C$. Ist μ ein integrierender Faktor, so ist auch $\tilde{C}\mu$ ein integrierender Faktor für beliebige Konstanten $\tilde{C} \neq 0$, d.h. wir können ohne Einschränkung $\tilde{C} = 1$ setzen. Damit erhalten wir die exakte DGL

$$\mu(t)A(t, y) + \mu(t)B(t, y)y' = \underbrace{1 + \frac{y}{t^2}}_{\tilde{A}(t, y)} + \underbrace{\left(-\frac{1}{t}\right)}_{\tilde{B}(t, y)}y' = 0.$$

Berechnen wir das Potential: Es soll $\partial_t F(t, y) = 1 + \frac{y}{t^2}$ gelten, also $F(t, y) = t - \frac{y}{t} + f(y)$. Zudem haben wir die Bedingung $\partial_y F(t, y) = -\frac{1}{t}$, welche mit dem vorherigen Ergebnis kombiniert zu $-\frac{1}{t} = \partial_y(t - \frac{y}{t} + f(y)) = -\frac{1}{t} + f'(y)$, also $f'(y) = 0$ führt. Damit ist also $F(t, y) = t - \frac{y}{t}$ (plus Konstante). Die Gleichung $F(t, y) = c$ wird also zu $t - \frac{y}{t} = c$ und hat die explizite Lösung $\varphi(t) = (t - c) \cdot t = t^2 - ct$. Überprüfen wir, dass φ tatsächlich die gewünschte DGL $t^2 + y - ty' = 0$ erfüllt: In der Tat ist

$$t^2 + (t^2 - ct) - t \cdot (2t - c) = 0.$$

8 Lineare Differentialgleichungen höherer Ordnung und Systeme von linearen Differentialgleichungen

In Mathe I hatten wir bereits gesehen, dass man jede gewöhnliche DGL höherer Ordnung in ein (dazu äquivalentes) System von DGLen erster Ordnung überführen kann. Nachdem wir uns in Mathe I ausschließlich mit DGLen erster Ordnung befasst hatten, wollen wir nun Lösungsverfahren für DGLen höherer Ordnung und Systeme linearer DGLen erster Ordnung betrachten.

8.1 Systeme von linearen DGLen

Definition 8.1. Sei $I \subset \mathbb{R}$ ein Intervall und seien

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} : I \rightarrow \mathbb{R}^{n \times n} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} : I \rightarrow \mathbb{R}^n$$

stetige Funktionen. Dann heißt

$$y' = A(t)y + b(t) \tag{8.1}$$

ein lineares Differentialgleichungssystem. Ist $b = 0$ der Nullvektor, so heißt (8.1) homogen; andernfalls heißt (8.1) inhomogen.

8.1.1 Homogene DGL-Systeme mit konstanten Koeffizienten – Lösungsmethode

Beispiel 8.2. Das System von linearen Differentialgleichungen

$$\begin{aligned} y_1' &= 2y_1 \\ y_2' &= 3y_2 \end{aligned}$$

lässt sich in Matrixschreibweise umschreiben als

$$y' = Ay \quad \text{bzw.} \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Da beide Gleichungen unabhängig voneinander sind, können wir sie separat lösen und erhalten die Lösungen

$$y_1(t) = c_1 e^{2t}, \quad y_2(t) = c_2 e^{3t},$$

wobei $c_1, c_2 \in \mathbb{R}$ beliebig wählbare Konstanten sind.

Im Allgemeinen hat eine 1-dim. DGL der Form $y' = ay$ (für eine Konstante $a \in \mathbb{R}$) mit Anfangswertbedingung $y(t_0) = y_0$ die eindeutige Lösung $y(t) = y_0 \exp(a(t - t_0))$, $t \in \mathbb{R}$.

Ob Differentialgleichungen unabhängig sind erkennt man auch daran, ob die zugehörige Matrix Diagonalf orm hat.

Hat man nicht das Glück, dass die DGLen unabhängig (man sagt auch *ungekoppelt*) sind, so kann man durch Diagonalisierung der Matrix, sofern dies möglich ist, das gekoppelte System von DGLen lösen:

Beispiel 8.3. Betrachten wir die gekoppelten DGLen

$$\begin{aligned} y_1' &= y_1 + y_2 \\ y_2' &= -2y_1 + 4y_2. \end{aligned}$$

Dieses homogen System von DGLen hat die zugehörige Matrix $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. Ihr charakteristisches Polynom lautet

$$p_A(t) = (1-t)(4-t) + 2 = t^2 - 5t + 6 = (t-2)(t-3),$$

somit hat A die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = 3$. Da beide Eigenwerte verschieden sind, ist A diagonalisierbar. Die Eigenräume sind

$$\text{Eig}(A, 2) = \{x \in \mathbb{R}^2 \mid (A - 2E_2)x = 0\} = \{x \in \mathbb{R}^2 \mid -x_1 + x_2 = 0\} = \left\{ a \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid a \in \mathbb{R} \right\},$$

$$\text{Eig}(A, 3) = \{x \in \mathbb{R}^2 \mid (A - 3E_2)x = 0\} = \{x \in \mathbb{R}^2 \mid -2x_1 + x_2 = 0\} = \left\{ a \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid a \in \mathbb{R} \right\}.$$

Somit diagonalisieren wir A mit Hilfe der Matrizen

$$U = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{und} \quad U^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Es ist in der Tat

$$U^{-1}AU = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -3 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} =: D.$$

Definieren wir nun eine neue vektorwertige Funktion $z = U^{-1}y$, so erfüllt diese die DGL

$$z' = U^{-1}y' = U^{-1}Ay = U^{-1}(UDU^{-1})y = DU^{-1}y = Dz.$$

Da D Diagonalform hat, ist dieses System von DGLen ungekoppelt und hat die allgemeine Lösung aus dem ersten Beispiel,

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} c_1 e^{2t} \\ c_2 e^{3t} \end{pmatrix}.$$

Durch die Rücksubstitution $y = Uz$ erhalten wir die Lösung der ursprünglichen DGL,

$$y = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} c_1 e^{2t} \\ c_2 e^{3t} \end{pmatrix} = \begin{pmatrix} c_1 e^{2t} + c_2 e^{3t} \\ c_1 e^{2t} + 2c_2 e^{3t} \end{pmatrix} = c_1 e^{2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{3t} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Die im Beispiel benutzte Methode ist allgemein anwendbar. Wir fassen sie in folgendem Satz zusammen:

Satz 8.4. Sei $A \in \mathbb{R}^{n \times n}$ eine diagonalisierbare Matrix mit Eigenwerten $\lambda_1, \dots, \lambda_n$ und der Matrix $U = (v_1 \ \dots \ v_n)$ aus Eigenvektoren von A , so dass $U^{-1}AU = \text{diag}(\lambda_1, \dots, \lambda_n)$ ist. Dann ist die allgemeine Lösung des DGL-Systems $y' = Ay$ gegeben durch

$$y = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \dots + c_n e^{\lambda_n t} v_n.$$

Bei gegebener Anfangswertbedingung $y(0) = y_0 \in \mathbb{R}^n$ berechnet man den Koeffizientenvektor $c = (c_1, \dots, c_n)^T$ als $c = U^{-1}y_0$.

Mit Hilfe des Matrixexponentials können wir die Lösung schreiben als

$$y(t) = e^{At}y(0).$$

In der Tat lässt sich der Satz somit sogar auf nicht diagonalisierbare Matrizen erweitern.

Satz 8.5. Sei $A \in \mathbb{R}^{n \times n}$. Dann ist die Lösung des AWP $y' = Ay$ mit $y(t_0) = y_0$ gegeben durch $y(t) = e^{A(t-t_0)}y(t_0)$, $y \in \mathbb{R}$.

Beispiel 8.6. Gegeben sei das System von Differentialgleichungen

$$\begin{aligned}y_1' &= y_2 + y_3 \\y_2' &= 2y_3 \\y_3' &= 0,\end{aligned}$$

so dass die zugehörige Matrix die nilpotente Matrix A aus Beispiel 6.43 ist. Das Matrixexponential hatten wir bereits berechnet als

$$e^{At} = \begin{pmatrix} 1 & t & t+t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{pmatrix}.$$

Ist die Anfangswertbedingung $y(0) = (a, b, c)^T$, so lautet die Lösung des Anfangswertproblems

$$y(t) = e^{At}y(0) = \begin{pmatrix} 1 & t & t+t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + bt + c(t+t^2) \\ b + 2ct \\ c \end{pmatrix}.$$

In der Tat ist $y(0) = (a, b, c)^T$ und es gilt

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} b + c + 2ct \\ 2c \\ 0 \end{pmatrix} = \begin{pmatrix} y_2 + y_3 \\ 2y_3 \\ 0 \end{pmatrix}.$$

8.1.2 Homogene DGL-Systeme – Charakterisierung der Lösungsmenge

Satz 8.7. Sei $I \subset \mathbb{R}$ ein Intervall und $A: I \rightarrow \mathbb{K}^{n \times n}$ eine stetige matrixwertige Funktion. Sei L_H die Menge aller Lösungen der homogenen linearen DGL $y' = A(t)y$. Dann ist L_H ein n -dimensionaler \mathbb{K} -Vektorraum. Ein k -Tupel von Lösungen $\varphi_1, \dots, \varphi_k \in L_H$ ist genau dann linear unabhängig, wenn ein $t_0 \in I$ existiert, so dass die Vektoren $\varphi_1(t_0), \dots, \varphi_k(t_0) \in \mathbb{K}^n$ linear unabhängig sind. In diesem Fall ist jedes k -Tupel $\varphi_1(t), \dots, \varphi_k(t) \in \mathbb{K}^n$ ($t \in I$ beliebig) linear unabhängig.

Ein homogenes (lineares) DGL-System $y' = A(t)y$ besitzt nach diesem Satz also unendlich viele Lösungen, die einen Vektorraum aufspannen. Insbesondere ist die Summe zweier Lösungen wieder eine Lösung – dies ist auch als *Superpositionsprinzip*¹ bekannt. Der Satz rechtfertigt also folgende Definition:

Definition 8.8. Unter einem Lösungs-Fundamentalsystem des DGL-Systems $y' = A(t)y$ versteht man eine Basis $(\varphi_1, \dots, \varphi_n)$ des Vektorraums seiner Lösungen.

Haben wir ein (geordnetes) Lösungs-Fundamentalsystem, können wir die (Spalten-)Vektoren $\varphi_k = (\varphi_{1,k}, \dots, \varphi_{n,k})^T$ in eine Matrix schreiben:

$$\Phi = \begin{pmatrix} \varphi_{1,1} & \cdots & \varphi_{1,n} \\ \vdots & & \vdots \\ \varphi_{n,1} & \cdots & \varphi_{n,n} \end{pmatrix}.$$

Jede Lösung kann man dann als Linearkombination der Vektoren des Lösungs-Fundamentalsystems schreiben:

$$\varphi = c_1\varphi_1 + \dots + c_n\varphi_n,$$

¹Superposition kommt aus dem Lateinischen und bedeutet Überlagerung. In der Mathematik besagt das Superpositionsprinzip, dass die Summe von Lösungen einer homogenen linearen Gleichung ebenfalls eine Lösung ist. In der Physik taucht der Begriff beispielweise bei der Überlagerung von Wellen auf.

d.h. für jeden beliebigen Vektor $c = (c_1, \dots, c_n)^T \in \mathbb{K}^n$ ist $\varphi = \Phi c$ eine Lösung von $y' = A(t)y$. Betrachten wir zunächst ein Beispiel eines homogenen DGL-Systems mit konstanten Koeffizienten bevor wir uns nicht-konstante Koeffizienten ansehen.

Beispiel 8.9. In Beispiel 8.6 hatten wir gesehen, dass die Lösung des DGL-Systems

$$\begin{aligned}y_1' &= y_2 + y_3 \\y_2' &= 2y_3 \\y_3' &= 0,\end{aligned}$$

für einen festen Startwert $y(0) = (a, b, c)^T$ gegeben ist durch

$$\varphi(t) = \begin{pmatrix} a + bt + c(t + t^2) \\ b + 2ct \\ c \end{pmatrix}.$$

Somit ist die Gesamtheit aller Lösungen des homogenen DGL-Systems (ohne Anfangswertbedingung) gegeben durch

$$L_H = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} t + t^2 \\ 2t \\ 1 \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}.$$

In der Tat erfüllen die drei ablesbaren Lösungen

$$\varphi_1(t) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \varphi_2(t) = \begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix}, \quad \varphi_3(t) = \begin{pmatrix} t + t^2 \\ 2t \\ 1 \end{pmatrix}$$

das DGL-System $y' = Ay$:

$$\begin{aligned}\varphi_1'(t) &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \varphi_{2,1}(t) + \varphi_{3,1}(t) \\ 2\varphi_{2,1}(t) \\ 0 \end{pmatrix} = A\varphi_1(t) \\ \varphi_2'(t) &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \varphi_{2,2}(t) + \varphi_{3,2}(t) \\ 2\varphi_{2,2}(t) \\ 0 \end{pmatrix} = A\varphi_2(t), \\ \varphi_3'(t) &= \begin{pmatrix} 1 + 2t \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \varphi_{2,3}(t) + \varphi_{3,3}(t) \\ 2\varphi_{2,3}(t) \\ 0 \end{pmatrix} = A\varphi_3(t).\end{aligned}$$

Die drei gefundenen Lösungen bilden auch in der Tat ein Lösungs-Fundamentalsystem, wie wir mit Hilfe der Determinante von Φ überprüfen können:

$$\det(\Phi) = \begin{vmatrix} 1 & t & t + t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0,$$

folglich sind die Spaltenvektoren von Φ (wie behauptet) linear unabhängig.

8.1.3 Homogene DGL-Systeme mit nicht-konstanten Koeffizienten

Betrachten wir DGL-Systeme der Form $y' = A(t)y$ für $A: I \subset \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$. Das Matrixexponential hilft uns nur in wenigen Sonderfällen:

Satz 8.10. Sei $A: \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ derart, dass $A(t)A'(t) = A'(t)A(t)$. Dann ist $e^{A(t)}$ ein Lösungs-Fundamentalsystem der DGL $y' = A'(t)y$.

Ist $B \in \mathbb{R}^{n \times n}$, so erfüllt $A(t) = Bt$ mit $A'(t) = B$ die Anforderungen des obigen Satzes.

8.1.4 Inhomogene DGL-Systeme mit konstanten Koeffizienten

Betrachten wir DGL-Systeme der Form $y' = Ay + b(t)$ für $A \in \mathbb{R}^{n \times n}$ und $b: I \subset \mathbb{R} \rightarrow \mathbb{R}^n$. Im letzten Semester haben wir zur Lösung inhomogener linearer DGLen die Methode *Variation der Konstanten* kennen gelernt. Diese Methode ist auch hier anwendbar und führt auf eine analoge Lösungsformel:

Satz 8.11. Sei $b: I \rightarrow \mathbb{R}^n$ stetig. Dann hat das inhomogene DGL-System $y' = Ay + b(t)$ zu jedem Anfangswert $(t_0, y_0) \in I \times \mathbb{R}^n$ genau eine Lösung, welche gegeben ist durch

$$\varphi(t) = e^{(t-t_0)A}y_0 + \int_{t_0}^t e^{(t-s)A}b(s)ds, \quad t \in I.$$

Beweis. Zunächst ist in der Tat der Anfangswert $\varphi(t_0) = y_0$. Überprüfen wir, dass die angegebene Lösung tatsächlich Lösung des DGL-Systems ist:

$$\begin{aligned} \varphi'(t) &= Ae^{(t-t_0)A}y_0 + \frac{d}{dt} \left[e^{tA} \int_{t_0}^t e^{-sA}b(s)ds \right] \\ &= Ae^{(t-t_0)A}y_0 + Ae^{tA} \int_{t_0}^t e^{-sA}b(s)ds + e^{tA}e^{-tA}b(t) \\ &= Ae^{(t-t_0)A}y_0 + A \int_{t_0}^t e^{(t-s)A}b(s)ds + b(t) \\ &= A\varphi(t) + b(t). \end{aligned}$$

Ist ψ eine weitere Lösung des AWP, so gilt $(\varphi - \psi)' = A(\varphi - \psi)$ und $(\varphi - \psi)(t_0) = y_0 - y_0 = 0$, d.h. $\varphi - \psi$ ist die eindeutige Lösung des AWP $y' = Ay$ mit Anfangswert $y(t_0) = 0$. Nach Satz 8.5 ist die einzige Lösung $\varphi - \psi \equiv 0$, d.h. das ursprünglich betrachtete inhomogene AWP besitzt genau eine Lösung. \square

8.2 Lineare DGLen höherer Ordnung

8.2.1 Homogene lineare DGLen höherer Ordnung mit konstanten Koeffizienten

Satz 8.12. Gegeben sei die homogene lineare DGL

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0.$$

Das charakteristische Polynom der DGL

$$p(x) := x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$$

habe die paarweise voneinander verschiedenen Nullstellen $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ mit algebraischen Vielfachheiten $\alpha_1, \dots, \alpha_m$. Dann bilden die Funktionen

$$\varphi_{k,\ell}(t) = t^\ell e^{\lambda_k t}, \quad k \in \{1, \dots, m\}, \ell \in \{0, \dots, \alpha_k - 1\}$$

ein Lösungs-Fundamentalsystem.

Beispiel 8.13. Gegeben sei die DGL $y'' + 4y' + 3y = 0$. Das zugehörige Polynom ist

$$p(x) = x^2 + 4x + 3$$

und hat die einfachen Nullstellen $\lambda_1 = -1$ und $\lambda_2 = -3$. Nach obigem Satz sollten also alle Lösungen der DGL von folgender Form sein:

$$\varphi(t) = c_1 e^{-t} + c_2 e^{-3t}, \quad c_1, c_2 \in \mathbb{R}.$$

In der Tat gilt

$$\begin{aligned} \varphi''(t) + 4\varphi'(t) + 3\varphi(t) &= (c_1 e^{-t} + 9c_2 e^{-3t}) + 4(-c_1 e^{-t} - 3c_2 e^{-3t}) + 3(c_1 e^{-t} + c_2 e^{-3t}) \\ &= 0. \end{aligned}$$

Formuliert man die DGL 2. Ordnung in ein System von linearen DGL 1. Ordnung um, so erhält man

$$\begin{aligned} y_1' &= -4y_1 - 3y_0 \\ y_0' &= y_1, \end{aligned}$$

d.h. $y' = Ay$ mit $A = \begin{pmatrix} -4 & -3 \\ 1 & 0 \end{pmatrix}$. Diese Matrix hat das charakteristische Polynom $p_A(t) = t^2 + 4t + 3$, d.h. das charakteristische Polynom der Matrix stimmt mit dem charakteristischen Polynom der DGL überein.

Sollten die Nullstellen des charakteristischen Polynoms der DGL komplex sein, so kann man mit Hilfe der Euler'schen Formel ($e^{ix} = \cos x + i \sin x$) dennoch ein reellwertiges Lösungsfundamentalsystem finden:

Beispiel 8.14. Betrachten wir die DGL $y'' + y = 0$. Ihr charakteristisches Polynom $p(x) = x^2 + 1$ hat die komplexen Nullstellen $\lambda_1 = i$ und $\lambda_2 = -i$. Die zugehörigen Lösungen lauten also $\varphi_1(t) = e^{it}$ und $\varphi_2(t) = e^{-it}$. Mit Hilfe der Euler'schen Formel können wir sie umschreiben in

$$\begin{aligned} \varphi_1(t) &= \cos(t) + i \sin(t), \\ \varphi_2(t) &= \cos(-t) + i \sin(-t) = \cos(t) - i \sin(t). \end{aligned}$$

Folglich lautet die allgemeine Lösung

$$\begin{aligned} \varphi(t) &= c_1(\cos(t) + i \sin(t)) + c_2(\cos(t) - i \sin(t)) \\ &= (c_1 + c_2) \cos(t) + (c_1 - c_2) i \sin(t). \end{aligned}$$

Hierbei können c_1 und c_2 beliebig gewählt werden, also können $c_1 - c_2$ und $c_1 + c_2$ beliebige reelle Werte annehmen. Wir finden somit das reelle Lösungsfundamentalsystem

$$\phi_1(t) = \cos(t), \quad \phi_2(t) = \sin(t). \tag{8.2}$$

Alternativer Lösungsweg:

Wir hätten die 2. Ordnung in ein 2-dim. DGL-System 1. Ordnung umwandeln können:

$$y'' = -y \iff \begin{cases} y_1' = -y_0 \\ y_0' = y_1. \end{cases} \iff y' = Ay \quad \text{mit} \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Diese hat die Diagonaldarstellung (Vgl. Beispiel 6.42)

$$U^{-1}AU = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} =: D$$

und somit das Matrixexponential

$$\begin{aligned} e^{At} &= Ue^{Dt}U^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e^{it} + e^{-it} & ie^{it} - ie^{-it} \\ -ie^{it} + ie^{-it} & e^{it} + e^{-it} \end{pmatrix} \\ &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \end{aligned}$$

Bei gegebenem Anfangswert $(y_1(0), y_2(0)) = (b, a)$ ist somit die Lösung von der DGL $y' = Ay$ gegeben durch

$$\begin{pmatrix} y_1(t) \\ y_0(t) \end{pmatrix} = e^{At} \begin{pmatrix} y_1(0) \\ y_0(0) \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} -a \sin t + b \cos t \\ a \cos t + b \sin t \end{pmatrix}$$

Übersetzen wir das System wieder zurück, so ist $\varphi(t) = y_0(t)$ und $\varphi'(t) = y_1(t)$, d.h. wie in (8.2) festgestellt hat jede beliebige Lösung der DGL $y'' = -y$ die Form

$$\varphi(t) = a \cos t + b \sin t.$$

Ist $\lambda = a + ib \in \mathbb{C}$ eine k -fache komplexe Nullstelle des charakteristischen Polynoms p , so erhält man für λ und $\bar{\lambda} = a - ib$ (welches ebenfalls eine Nullstelle von p ist) die $2k$ linear unabhängigen reellen Lösungen

$$\begin{aligned} &e^{at} \cos(bt), te^{at} \cos(bt), \dots, t^{k-1} e^{at} \cos(bt) \quad \text{und} \\ &e^{at} \sin(bt), te^{at} \sin(bt), \dots, t^{k-1} e^{at} \sin(bt). \end{aligned}$$

8.2.2 Inhomogene lineare DGLen höherer Ordnung mit konstanten Koeffizienten

Betrachten wir die DGL

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y + a_0 = b(t).$$

Ist φ_p eine partikuläre Lösung dieser DGL, so ist jede andere Lösung von der Form $\varphi = \varphi_p + \varphi_h$, wobei φ_h Lösung der zugehörigen homogenen DGL ist.

Partikuläre Lösungen findet man möglicherweise durch *Variation der Konstanten*:

Ist $\varphi_h(t) = \sum_{k=1}^n c_k \varphi_k(t)$ eine Lösung der homogenen DGL mit Fundamentallösungen $\varphi_1, \dots, \varphi_n$, so macht man den Ansatz

$$\varphi_p(t) = \sum_{k=1}^n c_k(t) \varphi_k(t).$$

Diese Funktion hat die Ableitung

$$\varphi_p'(t) = \sum_{k=1}^n [c_k'(t) \varphi_k(t) + c_k(t) \varphi_k'(t)].$$

Der Trick besteht nun darin, dass man $\sum_{k=1}^n c_k'(t) \varphi_k(t) = 0$ setzt. Das Ableiten wiederholt man nun bis zum Grad $n - 1$ und erhält somit Bedingungen, die man in die DGL einsetzen kann.

Beispiel 8.15. Betrachten wir das AWP

$$y'' + 4y' = \cos(2t), \quad y(0) = 0, \quad y'(0) = 1.$$

homogene DGL: Das charakteristische Polynom $p(x) = x^2 + 4x$ hat die Nullstellen 0 und -4 . Somit lautet die allgemeine homogene Lösung

$$\varphi_h(t) = c_1 e^{0t} + c_2 e^{-4t} = c_1 + c_2 e^{-4t}.$$

inhomogene DGL ohne Anfangswert: Wir machen den Ansatz $y(t) = c_1(t) + c_2(t)e^{-4t}$. Mit dem Trick $c_1'(t) + c_2'(t)e^{-4t} = 0$ ist

$$\begin{aligned} y'(t) &= c_1'(t) + c_2'(t)e^{-4t} - 4c_2(t)e^{-4t} = -4c_2(t)e^{-4t} \quad \text{und} \\ y''(t) &= -4c_2'(t)e^{-4t} + 16c_2e^{-4t} \end{aligned}$$

Damit ist

$$\begin{aligned} y''(t) + 4y'(t) &= -4c_2'(t)e^{-4t} + 16c_2e^{-4t} - 16c_2(t)e^{-4t} \\ &= -4c_2'(t)e^{-4t} \\ &\stackrel{!}{=} \cos(2t). \end{aligned}$$

Zusammen mit dem Trick haben wir also ein Gleichungssystem:

$$\begin{aligned} c_1'(t) &= -c_2'(t)e^{-4t} \\ c_2'(t) &= -\frac{1}{4}e^{4t} \cos(2t). \end{aligned}$$

Dies hat die eindeutige Lösung

$$c_1'(t) = \frac{1}{4} \cos(2t), \quad c_2'(t) = -\frac{1}{4} e^{4t} \cos(2t).$$

Integration ergibt (mit gewählter Integrationskonstante 0) die Funktionen

$$c_1(t) = \frac{1}{8} \sin(2t), \quad c_2(t) = -\frac{1}{40} (2 \cos(2t) + \sin(2t)) e^{4t}.$$

Die partikuläre Lösung ist somit

$$\begin{aligned} \varphi_p(t) &= \frac{1}{8} \sin(2t) - \frac{1}{40} (2 \cos(2t) + \sin(2t)) e^{4t} e^{-4t} \\ &= \frac{1}{10} \sin(2t) - \frac{1}{20} \cos(2t). \end{aligned}$$

Die allgemeine Lösung der inhomogenen DGL ohne Anfangswertbedingung ist somit

$$\varphi(t) = \varphi_p(t) + \varphi_h(t) = \frac{1}{10} \sin(2t) - \frac{1}{20} \cos(2t) + c_1 + c_2 e^{-4t}.$$

inhomogenes AWP: Setzen wir die Anfangswerte ein, so erhalten wir mit

$$\varphi'(t) = \frac{1}{5} \cos(2t) + \frac{1}{10} \sin(2t) - 4c_2 e^{-4t}$$

das Gleichungssystem

$$\begin{aligned} 0 &= 0 - \frac{1}{20} + c_1 + c_2 \\ 1 &= \frac{1}{5} + 0 - 4c_2 \quad \iff \quad c_2 = -\frac{1}{5}. \end{aligned}$$

Die Konstanten setzen wir also $c_1 = \frac{1}{4}$ und $c_2 = -\frac{1}{5}$.

Es gibt noch diverse weitere Lösungsmethoden, insb. für gewisse Klassen von Differentialgleichungen. Diese hier zu besprechen würde den Rahmen der Vorlesung sprengen.

9 Flächen- und Raumintegrale

Ziel dieses Abschnitts ist die Definition und Berechnung von Integralen der Form $\int_A f(x)dx$ für Mengen $A \subset \mathbb{R}^n$ und $f: A \rightarrow \mathbb{R}$.

Ist $A = \{x \in \mathbb{R} \mid a \leq x \leq b\} = [a, b]$, dann ist bekanntlich

$$\int_A f(x)dx = \int_a^b f(x)dx.$$

Definition 9.1. Sei $A \subset \mathbb{R}^n$ für $n \in \{2, 3\}$. Die Menge A heißt Normalbereich, falls

$n = 2$: $a_1, b_1 \in \mathbb{R}$ und stetige Funktionen $a_2, b_2: \mathbb{R} \rightarrow \mathbb{R}$ existieren, so dass

$$A = \{(x, y)^T \in \mathbb{R}^2 \mid a_1 \leq x \leq b_1, a_2(x) \leq y \leq b_2(x)\}.$$

$n = 3$: $a_1, b_1 \in \mathbb{R}$, stetige Funktionen $a_2, b_2: \mathbb{R} \rightarrow \mathbb{R}$ und $a_3, b_3: \mathbb{R}^2 \rightarrow \mathbb{R}$ existieren, so dass

$$A = \{(x, y, z)^T \in \mathbb{R}^3 \mid a_1 \leq x \leq b_1, a_2(x) \leq y \leq b_2(x), a_3(x, y) \leq z \leq b_3(x, y)\}.$$

Beispiel 9.2. Sei $A = \{(x, y)^T \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$. Dies ist ein Normalbereich, denn

$$A = \{(x, y)^T \in \mathbb{R}^2 \mid -1 \leq x \leq 1, -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}\}.$$

Satz 9.3. Ist $A \subset \mathbb{R}^n$ ($n \in \{2, 3\}$) ein Normalbereich und $f: A \rightarrow \mathbb{R}$ über A integrierbar, so berechnet man das Integral $\int_A f(x)dx$ wie folgt:

$n = 2$: Ist $A = \{(x, y)^T \in \mathbb{R}^2 \mid a_1 \leq x \leq b_1, a_2(x) \leq y \leq b_2(x)\}$, so ist

$$\int_A f(x, y)d(x, y) = \int_{a_1}^{b_1} \left(\int_{a_2(x)}^{b_2(x)} f(x, y)dy \right) dx.$$

$n = 3$: Ist $A = \{(x, y, z)^T \in \mathbb{R}^3 \mid a_1 \leq x \leq b_1, a_2(x) \leq y \leq b_2(x), a_3(x, y) \leq z \leq b_3(x, y)\}$, so ist

$$\int_A f(x, y, z)d(x, y, z) = \int_{a_1}^{b_1} \left(\int_{a_2(x)}^{b_2(x)} \left(\int_{a_3(x, y)}^{b_3(x, y)} f(x, y, z)dz \right) dy \right) dx.$$

Satz 9.4. Seien $A, B \subset \mathbb{R}^n$ mit $A \cap B = \emptyset$ und seien $f: A \cup B \rightarrow \mathbb{R}$ und $g: A \rightarrow \mathbb{R}$ gegeben. Zudem sei $\alpha \in \mathbb{R}$. Falls die auftretenden Integrale existieren, so gelten folgende Beziehungen:

i) $\int_A f(x) + g(x)dx = \int_A f(x)dx + \int_A g(x)dx,$

ii) $\int_A \alpha f(x)dx = \alpha \int_A f(x)dx,$

iii) $\int_{A \cup B} f(x)dx = \int_A f(x)dx + \int_B f(x)dx.$

Mit Hilfe von iii) lassen sich auch Integrale der Form $\int_A f(x)dx$ berechnen für Mengen, die kein Normalbereich sind, sich aber aus solchen zusammensetzen lassen.

Sehen wir uns ein Beispiel für die Integration einer Funktion über dem Einheitskreis an.

Beispiel 9.5. Sei $A \subset \mathbb{R}^2$ wie in Beispiel 9.2. Integrieren wir über A die konstante Funktion $f(x) \equiv 1$. Dann ist (mit Subst. $x = \sin u$, $dx = \cos u du$)

$$\begin{aligned} \int_A f(x, y) d(x, y) &= \int_{-1}^1 \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy \right) dx \\ &= \int_{-1}^1 \left(\sqrt{1-x^2} - (-\sqrt{1-x^2}) \right) dx \\ &= \int_{-1}^1 2\sqrt{1-x^2} dx \\ &= \int_{\arcsin(-1)}^{\arcsin(1)} 2\sqrt{1-(\sin u)^2} \cos u du \\ &= \int_{\arcsin(-1)}^{\arcsin(1)} 2(\cos u)^2 du \end{aligned}$$

Die Stammfunktion lautet

$$\int (\cos u)^2 du = \frac{1}{2} (u + \sin u \cos u) + C$$

und man erhält (mit $\arcsin(-x) = -\arcsin(x)$, $\cos(\arcsin x) = \sqrt{1-x^2}$ und $\arcsin(1) = \frac{\pi}{2}$)

$$\begin{aligned} \int_A f(x, y) d(x, y) &= [u + \sin u \cos u]_{\arcsin(-1)}^{\arcsin(1)} \\ &= \arcsin(1) - \arcsin(-1) + 1 \cdot \cos(\arcsin(1)) - (-1) \cdot \cos(\arcsin(-1)) \\ &= 2 \arcsin(1) + \cos(\arcsin(1)) + \cos(\arcsin(-1)) \\ &= 2 \arcsin(1) + \sqrt{1-1^2} + \sqrt{1-(-1)^2} \\ &= \pi. \end{aligned}$$

Dies ist der Flächeninhalt des Einheitskreises.

Lemma 9.6. Der Flächeninhalt einer Menge $A \subset \mathbb{R}^2$ ist $\int_A 1 d(x, y)$ und das Volumen einer Menge $B \subset \mathbb{R}^3$ ist $\int_B 1 d(x, y, z)$.

Definition 9.7 (Polarkoordinaten). Sei $g: [0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2$ gegeben durch

$$g(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \quad \text{mit} \quad Dg(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Ist $(x, y)^T = g(r, \varphi)$, dann heißen (r, φ) Polarkoordinaten des Punktes (x, y) .

Die Einschränkung $g: (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus ([0, \infty) \times \{0\})$ ist eine Bijektion.

Analog zur Substitutionsregel gibt es auch die Möglichkeit, mehrdimensionale Integrale zu transformieren, welche die Berechnung des Integrals unter Umständen erleichtert:

Satz 9.8 (Transformationsatz). Sei $A \subset \mathbb{R}^n$ eine offene Menge und sei $g: A \rightarrow g(A) \subset \mathbb{R}^n$ ein Diffeomorphismus, d.h. eine bijektive stetig differenzierbare Abbildung mit stetig differenzierbarer Umkehrabbildung. Ist f auf der Menge $g(A)$ integrierbar, dann gilt

$$\int_{g(A)} f(y) dy = \int_A f(g(x)) |\det(Dg(x))| dx,$$

wobei Dg die Jacobi-Matrix von g ist (siehe Abschnitt 4.3) und $\det(Dg)$ die Funktionaldeterminante von g , d.h. die Determinante von Dg .

Beispiel 9.9. Mit Hilfe von Polarkoordinaten kann man den (abgeschlossenen) Kreis mit Radius $R > 0$ um $(0, 0)^T$ auf folgende Weise darstellen:

$$A = \{ (x, y)^T \in \mathbb{R}^2 \mid x = r \cos \varphi, y = r \sin \varphi, 0 \leq \varphi \leq 2\pi, 0 \leq r \leq R \} \\ = \{ (r \cos \varphi, r \sin \varphi)^T \mid 0 \leq r \leq R, 0 \leq \varphi \leq 2\pi \}.$$

Berechnen wir analog zu Beispiel 9.5 den Flächeninhalt des Einheitskreises, d.h. wir setzen $R = 1$ und $f \equiv 1$. Weiter sei $g(r, \varphi) := (r \cos \varphi, r \sin \varphi)^T$ und damit

$$Dg(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}, \quad \text{also} \quad \det(Dg(r, \varphi)) = r(\cos^2 \varphi + \sin^2 \varphi) = r.$$

Mit der Transformationsregel ist folglich

$$\begin{aligned} \int_A f(x, y) d(x, y) &= \int_{g^{-1}(A)} f(r \cos \varphi, r \sin \varphi) \cdot rd(r, \varphi) \\ &= \int_0^1 \int_0^{2\pi} 1 \cdot r \, d\varphi dr \\ &= \int_0^1 2\pi r \, dr \\ &= 2\pi \left[\frac{1}{2} r^2 \right]_0^1 \\ &= \pi. \end{aligned}$$

Für die Integration über Teilmengen des \mathbb{R}^3 gibt es zwei Spezialfälle der Transformation, die wir uns ansehen wollen – der Übergang zu Zylinder- und zu Kugelkoordinaten.

Definition 9.10 (Zylinderkoordinaten). Sei $g: [0, \infty) \times [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3$ gegeben durch

$$g(r, \varphi, z) := \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \\ z \end{pmatrix} \quad \text{mit} \quad Dg(r, \varphi, z) = \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ist $(x, y, z)^T = g(r, \varphi, z)$, dann heißen (r, φ, z) Zylinderkoordinaten des Punktes (x, y, z) .

Definition 9.11 (Kugelkoordinaten). Sei $g: [0, \infty) \times [0, \pi) \times [0, 2\pi) \rightarrow \mathbb{R}^3$ gegeben durch

$$g(r, \theta, \varphi) := \begin{pmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{pmatrix} \quad \text{mit} \quad Dg(r, \theta, \varphi) = \begin{pmatrix} \cos \varphi \sin \theta & r \cos \varphi \cos \theta & -r \sin \varphi \sin \theta \\ \sin \varphi \sin \theta & r \sin \varphi \cos \theta & r \cos \varphi \sin \theta \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}.$$

Ist $(x, y, z)^T = g(r, \theta, \varphi)$, dann heißen (r, θ, φ) Kugelkoordinaten des Punktes (x, y, z) .

Fassen wir im Folgenden diese drei Spezialfälle des Transformationsatzes zusammen:

Satz 9.12. Seien $A_1 \subset \mathbb{R}^2$ und $A_2, A_3 \subset \mathbb{R}^3$ und $f_1: A_1 \rightarrow \mathbb{R}$ und $f_2: A_2 \rightarrow \mathbb{R}$, $f_3: A_3 \rightarrow \mathbb{R}$ integrierbar über ihrem Definitionsbereich. Dann gelten folgende Aussagen:

Polarkoordinaten: Hat A_1 die Darstellung

$$A_1 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} = g(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \mid \begin{pmatrix} r \\ \varphi \end{pmatrix} \in A_1^* \right\}$$

mit $\det(Dg(r, \varphi)) = r$, so gilt die Transformationsformel

$$\int_{A_1} f_1(x, y) d(x, y) = \int_{A_1^*} r \cdot f_1(r \cos \varphi, r \sin \varphi) d(r, \varphi).$$

Zylinderkoordinaten: Hat A_2 die Darstellung

$$A_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(r, \varphi, z) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix} \mid \begin{pmatrix} r \\ \varphi \\ z \end{pmatrix} \in B^* \right\}$$

mit $\det(Dg(r, \varphi, z)) = r$, so gilt die Transformationsformel

$$\int_{A_2} f_2(x, y, z) d(x, y, z) = \int_{B_2^*} r \cdot g(r \cos \varphi, r \sin \varphi, z) d(r, \varphi, z).$$

Kugelkoordinaten: Hat A_3 die Darstellung

$$A_3 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(r, \theta, \varphi) = \begin{pmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{pmatrix} \mid \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} \in A_3^* \right\}$$

mit $\det(Dg(r, \theta, \varphi)) = r^2 \sin \theta$, so gilt die Transformationsformel

$$\int_{A_3} f_3(x, y, z) d(x, y, z) = \int_{A_3^*} r^2 \sin \theta \cdot f_3(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) d(r, \theta, \varphi).$$

Die Funktionaldeterminante von Polar- und Zylinderkoordinatentransformation ist jeweils r , die der oben angegebenen Kugelkoordinatentransformation ist $r^2 \sin \theta$. Beide sind positiv auf den Definitionsbereichen der Transformationen (so wie oben angegeben), so dass kein Betrag nötig ist.

Beispiel 9.13. Sei $K_R(0)$ die offene Kugel mit Radius $R > 0$ mit Mittelpunkt $(0, 0, 0)$. Das Volumen der Kugel ist dann (mit Übergang zu Kugelkoordinaten)

$$\begin{aligned} \int_{K_R(0)} 1 d(x, y, z) &= \int_0^R \left(\int_0^\pi \left(\int_0^{2\pi} r^2 \sin \theta d\varphi \right) d\theta \right) dr \\ &= \int_0^R \left(\int_0^\pi (2\pi \cdot r^2 \sin \theta) d\theta \right) dr \\ &= \int_0^R \left(2\pi \cdot r^2 \int_0^\pi \sin \theta d\theta \right) dr \\ &= \int_0^R \left(2\pi \cdot r^2 \cdot \underbrace{[-\cos \pi + \cos 0]}_{=2} \right) dr \\ &= 4\pi \int_0^R r^2 dr \\ &= 4\pi \cdot \frac{1}{3} R^3 \\ &= \frac{4}{3} \pi R^3. \end{aligned}$$