Humboldt-Universität zu Berlin
Institut für Mathematik
C. Wendl, A. Fauck,
L. Kotan, L. Upmeier zu Belzen

Analysis III
WiSe 2019–20

Übungsblatt 3: Musterlösung zur Aufgabe 3.3

Aufgabe 3.3 (3 + 3 + 6 + 3 Punkte)
Wir betrachten die sogenannte *Gradientenflussgleichung*

$$\dot{\mathbf{x}}(t) = \nabla f(\mathbf{x}(t)) \tag{1}$$

für eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ der Klasse C^2, die im Punkt $0 \in \mathbb{R}^n$ ein lokales Maximum annimmt. Wir bezeichnen mit $\mathbf{A} := Hf(0) \in \mathbb{R}^{n\times n}$ die Hesse-Matrix von f im kritischen Punkt $0 \in \mathbb{R}^n$, und nennen den kritischen Punkt **nicht entartet**, falls $\mathbf{A}$ invertierbar ist. Beweisen Sie:

a) Es gibt eine stetige matrixwertige Funktion $\mathbf{B} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ mit $\mathbf{B}(0) = 0$, so dass $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{B}(\mathbf{x})\mathbf{x}$ für alle $\mathbf{x} \in \mathbb{R}^n$ gilt.

Lösung:
Da f von der Klasse C^2 ist und ein lokales Maximum in $0 \in \mathbb{R}^n$ annimmt, muss 0 auch ein kritischer Punkt von f sein, also $\nabla f(0) = 0$, und die Hesse-Matrix $\mathbf{A} = Hf(0)$ ist außerdem negativ-semidefinit. Die gewünschte Formel folgt im Wesentlichen aus dem Mittelwertsatz in Integralform, der eine Konsequenz vom Hauptsatz der Differential- und Integralrechnung ist. Der Gradient $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ ist eine C^1-Funktion, und ihre Jacobi-Matrix in jedem Punkt $\mathbf{x} \in \mathbb{R}^n$ ist die Hesse-Matrix von f,

$$D(\nabla f)(\mathbf{x}) = Hf(\mathbf{x}),$$

also gilt insb. $D(\nabla f)(0) = \mathbf{A}$. Für ein gegebenes $\mathbf{x} \in \mathbb{R}^n$ ist die Funktion $[0, 1] \to \mathbb{R}^n : s \mapsto \nabla f(s\mathbf{x})$ stetig differenzierbar, also folgt aus dem Hauptsatz,

$$\begin{aligned}
\nabla f(\mathbf{x}) = \nabla f(\mathbf{x}) - \nabla f(0) &= \int_0^1 \frac{d}{ds}\nabla f(s\mathbf{x})\,ds \\
&= \int_0^1 D(\nabla f)(s\mathbf{x})\frac{d}{ds}(s\mathbf{x})\,ds = \int_0^1 D(\nabla f)(s\mathbf{x})\mathbf{x}\,ds = \left(\int_0^1 D(\nabla f)(s\mathbf{x})\,ds\right)\mathbf{x} \\
&= \mathbf{A}\mathbf{x} + \left(\int_0^1 \left[D(\nabla f)(s\mathbf{x}) - D(\nabla f)(0)\right]\,ds\right)\mathbf{x}.
\end{aligned}$$

Wir definieren $\mathbf{B}(\mathbf{x}) \in \mathbb{R}^{n\times n}$ als das Integral in Klammern in der letzten Zeile. Das ist nach dem Satz über parameterabhängige Integrale (mit $\mathbf{x} \in \mathbb{R}^n$ in diesem Fall als Parameter) eine stetige Funktion von $\mathbf{x}$, und sie erfüllt $\mathbf{B}(0) = 0$.

b) Ist der kritische Punkt $0 \in \mathbb{R}^n$ nicht entartet, dann gilt

$$\langle \mathbf{x}, \mathbf{A}\mathbf{x}\rangle \le -\lambda\|\mathbf{x}\|^2 \quad \text{für alle } \mathbf{x} \in \mathbb{R}^n,$$

mit einer Konstante $\lambda > 0$. Finden Sie den optimalen Wert für diese Konstante.
Hinweis: λ sollte etwas mit den Eigenwerten von $\mathbf{A}$ *zu tun haben.*

Lösung:
Wie schon bemerkt muss $\mathbf{A}$ negativ-semidefinit sein, was heißt, alle Eigenwerte von $\mathbf{A}$ sind nicht positiv. Aber wenn der kritische Punkt auch nicht entartet ist, dann muss

$\mathbf{A}$ invertierbar sein, also darf kein Eigenwert 0 sein. Wir bezeichnen die Eigenwerte mit $\lambda_1, \ldots, \lambda_n < 0$, und definieren

$$\lambda := \min\{-\lambda_1, \ldots, -\lambda_n\} > 0,$$

also gilt $\lambda_j \leq -\lambda$ für alle $j = 1, \ldots, n$. Zu den Eigenwerten gibt es eine orthonormale Basis von Eigenvektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, mit $\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$ für $j = 1, \ldots, n$. Ein gegebener Vektor $\mathbf{x} \in \mathbb{R}^n$ kann dann als lineare Kombination $\mathbf{x} = \sum_{j=1}^n x_i \mathbf{v}_i$ geschrieben werden, und es gilt

$$\langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle = \sum_{i,j} \langle x_i \mathbf{v}_i, \mathbf{A} x_j \mathbf{v}_j \rangle = \sum_{i,j} x_i x_j \lambda_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_j x_j^2 \lambda_j \leq -\lambda \sum_j x_j^2 = -\lambda \|\mathbf{x}\|^2.$$

c) Sei $\mathbf{x} : (t_-, t_+) \to \mathbb{R}^n$ eine maximale Lösung zur Gleichung (1) mit $0 \in (t_-, t_+)$. Ist der kritische Punkt $0 \in \mathbb{R}^n$ nicht entartet, dann gibt es für jedes $\epsilon \in (0, \lambda)$ eine Umgebung $\mathcal{U}_\epsilon \subset \mathbb{R}^n$ von 0 mit der folgenden Eigenschaft: liegt $\mathbf{x}(0)$ in $\mathcal{U}_\epsilon$, dann gilt auch $\mathbf{x}([0, t_+)) \subset \mathcal{U}_\epsilon$, $t_+ = \infty$, $\lim_{t\to\infty} \mathbf{x}(t) = 0$ und

$$\|\mathbf{x}(t)\| \leq C e^{-(\lambda - \epsilon)t} \quad \text{für alle} \quad t \geq 0 \tag{2}$$

mit einer Konstante $C > 0$.
Hinweis: Wenn $\mathbf{B}(\mathbf{x}(t))$ klein ist, dann ist $z(t) := \|\mathbf{x}(t)\|^2$ eine Unterfunktion für die Differentialgleichung $\dot{y}(t) = -2(\lambda - \epsilon)y(t)$. (Warum?)

Lösung:
Die Funktion $\mathbf{B} : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ aus Teilaufgabe (a) ist stetig und erfüllt $\mathbf{B}(0) = 0$, also existiert für jedes $\epsilon > 0$ ein $\delta > 0$, so dass

$$\|\mathbf{x}\| < \delta \quad \Rightarrow \quad \|\mathbf{B}(\mathbf{x})\| < \epsilon.$$

Wir definieren $\mathcal{U}_\epsilon \subset \mathbb{R}^n$ als die offene Kugel von Radius δ um 0. Wenn $\mathbf{x} : [0, t_+) \to \mathbb{R}^n$ die Gradientenflussgleichung (1) erfüllt, dann ist die Ableitung von $z(t) := \|\mathbf{x}(t)\|^2$ gegeben durch

$$\begin{aligned}
\dot{z}(t) &= \frac{d}{dt} \langle \mathbf{x}(t), \mathbf{x}(t) \rangle = 2 \langle \mathbf{x}(t), \dot{\mathbf{x}}(t) \rangle = 2 \langle \mathbf{x}(t), \nabla f(\mathbf{x}(t)) \rangle \\
&\overset{\text{(a)}}{=} 2 \left\langle \mathbf{x}(t), \mathbf{A}\mathbf{x}(t) \right\rangle + 2 \left\langle \mathbf{x}(t), \mathbf{B}(\mathbf{x}(t))\mathbf{x}(t) \right\rangle \overset{\text{(b)}}{\leq} -2\lambda \|\mathbf{x}(t)\|^2 + 2 \left\langle \mathbf{x}(t), \mathbf{B}(\mathbf{x}(t))\mathbf{x}(t) \right\rangle \\
&= -2\lambda z(t) + 2 \left\langle \mathbf{x}(t), \mathbf{B}(\mathbf{x}(t))\mathbf{x}(t) \right\rangle.
\end{aligned}$$

Liegt $\mathbf{x}(t)$ in $\mathcal{U}_\epsilon$, dann gilt $\|\mathbf{B}(\mathbf{x}(t))\| < \epsilon$, also $\langle \mathbf{y}, \mathbf{B}(\mathbf{x}(t))\mathbf{y} \rangle \leq \|\mathbf{y}\| \cdot \|\mathbf{B}(\mathbf{x}(t))\| \cdot \|\mathbf{y}\| < \epsilon \|\mathbf{y}\|^2$ für beliebige Vektoren $\mathbf{y} \in \mathbb{R}^n$, und es folgt

$$\dot{z}(t) \leq -2\lambda z(t) + 2\epsilon z(t) = -2(\lambda - \epsilon) z(t). \tag{3}$$

Insbesondere gilt $\dot{z}(t) \leq 0$.
Gilt $\mathbf{x}(0) \in \mathcal{U}_\epsilon$, dann können wir jetzt beweisen, dass $\mathbf{x}(t)$ für alle $t \in [0, t_+)$ noch in $\mathcal{U}_\epsilon$ bleibt: wenn dies nicht so ist, dann existiert das Infimum

$$t_1 := \inf \left\{ t \in [t_0, t_+) \;\middle|\; \mathbf{x}(t) \notin \mathcal{U}_\epsilon \right\},$$

und es gilt $t_1 > t_0$, da $\mathcal{U}_\epsilon$ offen und $\mathbf{x}$ stetig ist. Es müssen außerdem $\|\mathbf{x}(t_1)\| = \delta$ und $\|\mathbf{x}(t)\| < \delta$ für alle $t \in [t_0, t_1)$ gelten, aber dann ist $z(t_1)$ größer als $z(t_0)$, also muss $\dot{z}(t)$ für irgendein $t \in (t_0, t_1)$ positiv sein, was (3) widerspricht.

Als Nächstes behaupten wir: wenn $\mathbf{x} : [0, t_+) \to \mathcal{U}_\epsilon$ eine maximale Lösung ist, dann gilt $t_+ = \infty$. Wenn nicht, dann existiert ein $t_* \in (t_-, t_+)$, so dass für alle $t \in (t_*, t_+)$, $(t, \mathbf{x}(t))$ nicht in der kompakten Menge $[t_0, t_+] \times \overline{\mathcal{U}}_\epsilon$ liegt. Das wäre nur möglich, wenn $\mathbf{x}(t) \not\in \overline{\mathcal{U}}_\epsilon$ für alle t in der Nähe von t_+, aber wir hatten gerade eben bewiesen, dass das nicht stimmt. Also ist $\mathbf{x}(t)$ tatsächlich für alle $t \in [0, \infty)$ definiert und liegt in $\mathcal{U}_\epsilon$.
Als Letztes beweisen wir die Abschätzung (2) für alle $t \in [0, \infty)$. Sei $z_0 := \|\mathbf{x}(0)\|^2$; dann impliziert (3), dass $z(t) := \|x(t)\|^2$ auf $[0, \infty)$ eine Unterfunktion für das Anfangswertproblem

$$\dot{y}(t) = -2(\lambda - \epsilon) y(t), \qquad y(0) = z_0$$

ist. Die Lösung dieses Anfangswertproblems ist $y(t) = z_0 e^{-2(\lambda-\epsilon)t}$, also folgt aus dem Satz über Ober- und Unterfunktionen in der Vorlesung,

$$z(t) \le y(t) = z_0 e^{-2(\lambda-\epsilon)t},$$

und daher,

$$\|\mathbf{x}(t)\| = \sqrt{z(t)} \le \sqrt{z_0} e^{-(\lambda-\epsilon)t}$$

für alle $t \in [0, \infty)$.

d) Die Funktion $f(x, y) := -x^2 - y^4$ auf $\mathbb{R}^2$ hat ein lokales Maximum in $0 \in \mathbb{R}^2$, aber der kritische Punkt ist entartet. Finden Sie für diese Funktion eine Lösung $\mathbf{x} : (t_-, \infty) \to \mathbb{R}^2$ zur Gleichung (1), die bei $t \to \infty$ gegen 0 konvergiert aber keine Abschätzung der Form (2) erfüllt.

Lösung:
Dass $Hf(0)$ nicht invertierbar ist, liegt zum Teil daran, dass $\partial_y^2 f(0,0)$ verschwindet, was als Motivation dienen kann, Lösungen auf der Achse $\{x = 0\}$ zu untersuchen. Wir nehmen also als Ansatz, dass eine Lösung der Form $(x(t), y(t)) = (0, y(t))$ existiert. Der Gradient ist $\nabla f(x, y) = (-2x, -4y^3)$, also erfüllt eine Funktion dieser Art die Gradientenflussgleichung genau dann, wenn $y(t)$ die Differentialgleichung

$$\dot{y}(t) = -4[y(t)]^3$$

erfüllt. Diese Gleichung lässt sich durch Trennung der Variablen wie folgt lösen:

$$\begin{aligned} \frac{dy}{dt} = -4y^3 \quad &\Rightarrow \quad \frac{dy}{y^3} = -4\,dt \quad \Rightarrow \quad \int \frac{dy}{y^3} = \int -4\,dt \\ &\Rightarrow \quad -\frac{1}{2y^2} = -4t + C \quad \Rightarrow \quad 2y^2 = \frac{1}{4t - C} \quad \Rightarrow \quad y = \frac{1}{\sqrt{8t - 2C}}. \end{aligned}$$

Die Konstante $C \in \mathbb{R}$ kann durch eine Anfangsbedingung $y(0) = y_0 > 0$ bestimmt werden, und die eindeutige Lösung $\mathbf{x}(t) \in \mathbb{R}^2$ der Gradientenflussgleichung mit $\mathbf{x}(0) = (0, y_0)$ ist also gegeben durch

$$\mathbf{x}(t) = \left(0, \frac{1}{\sqrt{8t + \frac{1}{y_0^2}}}\right).$$

Diese Lösung konvergiert wie erwartet gegen $0 \in \mathbb{R}^2$ bei $t \to \infty$, aber die Konvergenz ist nicht exponentiell: wir behaupten, es gibt keine Konstanten $\lambda > 0$ und $C > 0$, so dass eine Abschätzung der Form

$$\|\mathbf{x}(t)\| \le Ce^{-\lambda t} \tag{4}$$

für alle $t \geq 0$ hinreichend groß erfüllt wird. Tatsächlich, wäre (4) für alle $t \geq 0$ hinreichend groß gültig, dann hätten wir auch

$$\frac{1}{8t + 1/y_0^2} \leq C^2 e^{-2\lambda t},$$

also müsste die Funktion $\frac{e^{2\lambda t}}{8t+1/y_0^2}$ bei $t \to \infty$ beschränkt sein. Aber die Regel von L'Hospital impliziert, dass eine Funktion dieser Form nie bei $t \to \infty$ beschränkt sein kann.