O. Schnürer, L. S. Krapp
Sommersemester 2020

Universität Konstanz
Fachbereich Mathematik und Statistik

Übungen zur Vorlesung Analysis II
Blatt 11

Abgabe von: Musterstudent
Tutor(in): Mein Lieblingstutor

1	2	3	4	Σ
5	6	4	9	24

Allgemeiner Hinweis: Für die Bearbeitung dieses Übungsblatts werden alle Resultate bis einschließlich Bemerkung 6.36 vorausgesetzt. Freiwillige Zusatzaufgaben sind mit einem * gekennzeichnet. Alle Aussagen sind stets zu beweisen.

Definition. Seien E, F Banachräume, sei $\Omega \subset E$ offen und sei $f\colon \Omega \to F$. Dann heißt f in $x \in \Omega$ *Gâteaux-differenzierbar*, falls $A \in L(E, F)$ existiert, sodass für alle $h \in E$

$$\lim_{t \to 0} \frac{1}{t}\left(f(x+th) - f(x) - A\langle th\rangle\right) = 0$$

gilt. (Hierbei nimmt t Werte in $\mathbb{R}$ an.) Wir nennen A die *Gâteaux-Ableitung* von f in x und bezeichnen diese mit $f'_G(x)$. Ist f in jedem Punkt $x \in \Omega$ Gâteaux-differenzierbar, so heißt f *Gâteaux-differenzierbar* und die Abbildung $x \mapsto f'_G(x)$ heißt *Gâteaux-Ableitung* von f.
Die Funktion f heißt in $x \in \Omega$ *Fréchet-differenzierbar*, falls $B \in L(E, F)$ existiert, sodass

$$\lim_{h \to 0} \frac{1}{\|h\|}\left(f(x+h) - f(x) - B\langle h\rangle\right) = 0$$

gilt. (Hierbei nimmt h Werte in E an.) Wir nennen B die *Fréchet-Ableitung* von f in x und bezeichnen diese mit $f'_F(x)$. Analog zu oben definieren wir *Fréchet-Differenzierbarkeit* und die *Fréchet-Ableitung*.
Die Funktion f heißt *stetig Gâteaux-differenzierbar* bzw. *stetig Fréchet-differenzierbar*, falls ihre Gâteaux- bzw. Fréchet-Ableitung stetig ist.

Aufgabe 11.1 (Gâteaux- und Fréchet-Ableitung) **[1 + 1 + 2 + 1* Punkte]**
Seien E, F Banachräume, sei $\Omega \subset E$ offen und sei $f\colon \Omega \to F$.

(i) Zeigen Sie, dass f genau dann Fréchet-differenzierbar ist, wenn es differenzierbar ist. Folgern Sie, dass in diesem Fall $f'_F = Df$ gilt.

(ii) Sei f Fréchet-differenzierbar. Zeigen Sie, dass f Gâteaux-differenzierbar ist und $f'_G = f'_F$ erfüllt.

(iii) Sei f stetig Gâteaux-differenzierbar. Zeigen Sie, dass f auch stetig Fréchet-differenzierbar ist und $f'_F = f'_G$ erfüllt.

(Hinweis: Zeigen Sie zunächst, dass f die Gleichung

$$\begin{aligned} f(x+h) - f(x) &= \int_0^1 \frac{\partial}{\partial t} f(x+th)\, dt = \int_0^1 f'_G(x+th)\langle h\rangle\, dt \\ &= f'_G(x)\langle h\rangle + \int_0^1 \left(f'_G(x+th)\langle h\rangle - f'_G(x)\langle h\rangle\right) dt \end{aligned}$$

erfüllt.)

(iv)* Erläutern Sie in Worten den Unterschied zwischen Gâteaux- und Fréchet-Differenzierbarkeit und finden Sie eine Funktion, die Gâteaux-, aber nicht Fréchet-differenzierbar ist.

Lösung:

(i) Sei f Fréchet-differenzierbar. Dann gilt für alle $x_0 \in \Omega$ (durch die Substitution $h = x - x_0$):

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'_F(x_0)\langle x - x_0 \rangle}{\|x - x_0\|} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - f'_F(x_0)\langle h \rangle}{\|h\|} = 0.$$

Insbesondere ist f in x_0 differenzierbar mit $Df(x_0) = f'_F(x_0)$.

Für die Rückrichtung können wir analog argumentieren, indem wir $x = x_0 + h$ substituieren.

(ii) Sei $x_0 \in \Omega$. Wir zeigen, dass f in x_0 Gâteaux-differenzierar ist und $f'_G(x_0) = f'_F(x_0)$ erfüllt. Der Beweis ist klar für $h = 0$. Sei also $h \in E \setminus \{0\}$. Dann gilt $th \to 0$ für $t \to 0$ (mit $t \in \mathbb{R}$). Nach Definition der Fréchet-Ableitung erhalten wir

$$\lim_{t \downarrow 0} \frac{f(x_0 + th) - f(x_0) - f'_F(x_0)\langle th \rangle}{t} = \|h\| \lim_{t \downarrow 0} \frac{f(x_0 + th) - f(x_0) - f'_F(x_0)\langle th \rangle}{\|th\|} = 0,$$

sowie

$$\lim_{t \uparrow 0} \frac{f(x_0 + th) - f(x_0) - f'_F(x_0)\langle th \rangle}{t} = -\|h\| \lim_{t \uparrow 0} \frac{f(x_0 + th) - f(x_0) - f'_F(x_0)\langle th \rangle}{\|th\|} = 0.$$

Dies zeigt die Behauptung.

(iii) Wir beweisen zunächst den Hinweis. Sei $x \in \Omega$ und sei $r > 0$ so, dass $B_r(x) \subset \Omega$ gilt. Sei $h \in E \setminus \{0\}$ mit $\|h\| < r$. Dann gilt $x + th \in \Omega$ für $t \in [-1, 1]$. Wir betrachten die Funktion $g(t) = f(x + th)$ definiert auf $[0, 1]$. Weil f Gâteaux-differenzierbar ist, ist g differenzierbar:

$$\begin{aligned}
\frac{\partial}{\partial t} f(x + th) &= g'(t) \\
&= \lim_{s \to 0} \frac{g(t + s) - g(t)}{s} \\
&= \lim_{s \to 0} \frac{f(x + th + sh) - f(x + th)}{s} \\
&= f'_G(x + th)\langle h \rangle.
\end{aligned}$$

Weil f'_G nach Voraussetzung stetig ist, ist g' als Verkettung stetiger Funktionen stetig. Wir können also den Hauptsatz anwenden:

$$\begin{aligned}
f(x + h) - f(x) &= g(1) - g(0) \\
&= \int_0^1 g'(t)\, dt \\
&= \int_0^1 \frac{\partial}{\partial t} f(x + th)\, dt \\
&= \int_0^1 f'_G(x + th)\langle h \rangle\, dt \\
&= \int_0^1 (f'_G(x + th)\langle h \rangle + f'_G(x)\langle h \rangle - f'_G(x)\langle h \rangle)\, dt \\
&= f'_G(x)\langle h \rangle + \int\limits_0^1 (f'_G(x + th)\langle h \rangle - f'_G(x)\langle h \rangle)\, dt.
\end{aligned}$$

Dies zeigt die Gültigkeit des Hinweises.

Wir zeigen nun, dass f Fréchet-differenzierbar ist und $f'_F = f'_G$ erfüllt. Hieraus folgt bereits, dass f'_F stetig ist.

Sei $x_0 \in \Omega$. Dann gilt für h wie oben:

$$f(x_0 + h) - f(x_0) = f'_G(x_0)\langle h\rangle + \int_0^1 (f'_G(x_0 + th)\langle h\rangle - f'_G(x)\langle h\rangle)\, dt.$$

Wir erhalten

$$\begin{aligned}\|f(x_0 + h) - f(x_0) - f'_G(x_0)\langle h\rangle\| &= \left\|\int_0^1 (f'_G(x_0 + th)\langle h\rangle - f'_G(x)\langle h\rangle)\, dt\right\| \\ &\le \sup_{0\le t\le 1} \|f'_G(x + th)\langle h\rangle - f'_G(x)\langle h\rangle\|_F \\ &\le \|h\| \sup_{0\le t\le 1} \|f'_G(x + th) - f'_G(x)\|_{L(E,F)}.\end{aligned}$$

Aufgrund der Stetigkeit von f'_G gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass $\|f'_G(y) - f'_G(x)\|_{L(E,F)} < \varepsilon$, falls $\|y - x\| < \delta$. Für $\|h\| < \delta$ folgt

$$\sup_{0\le t\le 1} \|f'_G(x_0 + th) - f'_G(x_0)\|_{L(E,F)} < \varepsilon.$$

Das impliziert die Konvergenz

$$\frac{1}{\|h\|} \|f(x_0 + h) - f(x_0) - f'_G(x_0)\langle h\rangle\| \le \sup_{0\le t\le 1} \|f'_G(x_0 + th) - f'_G(x_0)\|_{L(E,F)} \to 0$$

mit $h \to 0$. Dies zeigt unsere Behauptung.

(iv) Die Definitionen sehen recht ähnlich aus, sind aber verschieden. Tatsächlich ist die Gâteaux-Ableitung in Richtung h also $f'(x)\langle h\rangle$ gerade die Ableitung von $t \mapsto f(x + th)$. Die Abbildung f ist also gerade dann in $x \in \Omega$ Gâteaux-differenzierbar, wenn f in x in alle Richtungen differenzierbar ist. Zusätzlich müssen diese Richtungsableitungen stetig linear von der Richtung abhängen. Bei der Fréchet-Ableitung steht die lineare Approximierbarkeit im Vordergrund: Eine Funktion ist Fréchet-differenzierbar in einem Punkt, wenn sie dort in erster Ordnung durch eine affin lineare Funktion approximiert wird. Diese Näherung hängt nur von $\|h\|$ ab. Im Fall der Gâteaux-Ableitung überprüfen wir diese Approximierbarkeit nur entlang von Geraden durch den Punkt und die Güte der Approximation darf von der Richtung der Geraden abhängen.
Mit obiger Erklärung wird klar, warum die Funktion in Beispiel A.5 im Ursprung Gâteaux-, aber nicht Fréchet-differenzierbar ist. (Siehe hierzu auch Bemerkung A.16 (iv).)

Aufgabe 11.2 (Bilineare Abbildungen) **[2 + 2 + 0* + 2* Punkte]**
Seien E_1, E_2, F Banachräume und sei $A \in L(E_1, E_2; F)$.

(a) Zeigen Sie, dass die folgenden Aussagen äquivalent sind:

(i) A ist stetig.

(ii) A ist in 0 stetig.

(iii) Die Operatornorm $\|A\|$ ist endlich.

(iv) Es gibt ein $c \geq 0$, sodass die Abschätzung

$$\|A\langle x, y\rangle\| \leq c\,\|x\|\,\|y\|$$

für alle $(x, y) \in E_1 \times E_2$ gilt.

(b) Zeigen Sie, dass $(L(E_1, E_2; F), \|\cdot\|)$ ein Banachraum ist.

(c)* Sei $B\colon E_1 \times E_2 \to F$ eine lineare Abbildung, sei $(x, y) \in E_1 \times E_2$ und sei $\lambda \in \mathbb{K}$. Vergewissern Sie sich, dass

$$A\langle \lambda x, \lambda y\rangle = \lambda^2 A\langle x, y\rangle \quad \text{und}$$
$$B\langle(\lambda x, \lambda y)\rangle = \lambda B\langle(x, y)\rangle$$

gelten.

(d)* Zeigen Sie, dass

$$\psi\colon L(E_1; L(E_2, F)) \to L(E_1, E_2; F), \quad \alpha \mapsto ((x, y) \mapsto (\alpha(x))(y))$$

einen normerhaltenden Vektorraumisomorphismus definiert.

Lösung:

(a) „(i) ⇒ (ii)“:

Stetigkeit in 0 ist nur ein Spezialfall von Stetigkeit. Somit ist diese Implikation klar.

„(ii) ⇒ (iii)“:

Nehme an, die Operatornorm ist nicht endlich. Dann gibt es eine Folge $((x_k, y_k))_k \subset E_1 \times E_2$ mit $x_k, y_k \neq 0$ für alle k, sodass

$$\frac{\|A\langle x_k, y_k\rangle\|}{\|x_k\|\|y_k\|} > k^2$$

gilt. Wir beachten die Stetigkeit von A in der Null und $\frac{x_k}{k\|x_k\|}, \frac{y_k}{k\|y_k\|} \to 0$ für $k \to \infty$ und erhalten einen Widerspruch:

$$1 < \left\| A\left\langle \frac{x_k}{k\|x_k\|}, \frac{y_k}{k\|y_k\|} \right\rangle \right\| \to 0.$$

„(iii) ⇒ (iv)“:

Sei $(x, y) \in E_1 \times E_2$. Gilt $x = 0$ oder $y = 0$, dann gilt $A\langle x, y\rangle = 0$. In diesem Fall ist (iv) trivialerweise erfüllt. Seien nun $x, y \neq 0$.

$$\frac{\|A\langle x, y\rangle\|}{\|x\|\|y\|} \leq \sup_{(a,b)\in E_1\times E_2, (a,b)\neq 0} \frac{\|A\langle a, b\rangle\|}{\|a\|\|b\|} = \|A\|.$$

Umordnen liefert (iv) mit $c = \|A\|$.

„(iv) ⇒ (i)“:

Es gilt für $x = (x^1, x^2)$, $y = (y^1, y^2) \in E_1 \times E_2$

$$\begin{aligned}
&\left\| A\left\langle x^1, x^2\right\rangle - A\left\langle y^1, y^2\right\rangle \right\| \\
&\leq \left\| A\left\langle x^1 - y^1, x^2\right\rangle \right\| + \left\| A\left\langle y^1, x^2 - y^2\right\rangle \right\| \\
&\leq c\left(\left\|x^1 - y^1\right\| \left\|x^2\right\| + \left\|y^1\right\| \left\|x^2 - y^2\right\| \right) \to 0.
\end{aligned}$$

für $(y^1, y^2) \to (x^1, x^2)$.

(b) Die Nichtnegativität der Operatornorm ist klar. Die Homogenität und Dreiecksungleichung sieht man durch ($\lambda \in \mathbb{R}$, $A, B \in L(E_1, E_2; F)$):

$$\begin{aligned} \|\lambda A + B\| &= \sup_{0 \neq x^i \in E_i} \frac{\|\lambda A \langle x^1, x^2 \rangle + B \langle x^1, x^2 \rangle\|}{\|x^1\|\|x^2\|} \\ &\leq \sup_{0 \neq x^i \in E_i} \left(|\lambda| \frac{\|A \langle x^1, x^2 \rangle\|}{\|x^1\|\|x^2\|} + \frac{\|B \langle x^1, x^2 \rangle\|}{\|x^1\|\|x^2\|} \right) \\ &\leq |\lambda| \cdot \sup_{0 \neq x^i \in E_i} \frac{\|A \langle x^1, x^2 \rangle\|}{\|x^1\|\|x^2\|} + \sup_{0 \neq x^i \in E_i} \frac{\|B \langle x^1, x^2 \rangle\|}{\|x^1\|\|x^2\|} \\ &= |\lambda| \cdot \|A\| + \|B\|. \end{aligned}$$

Zur Definitheit: Ist $A\langle x_0, y_0 \rangle \neq 0$ für ein $(x_0, y_0) \in E_1 \times E_2$, dann gilt $x_0, y_0 \neq 0$. Somit

$$\|A\| = \sup_{x,y \neq 0} \frac{\|A \langle x, y \rangle\|}{\|x\|\|y\|} \geq \frac{\|A \langle x_0, y_0 \rangle\|}{\|x_0\|\|y_0\|} > 0.$$

Zur Vollständigkeit: Sei $(A_n)_n \subset L(E_1, E_2; F)$ eine Cauchyfolge. Dann gibt es zu $\varepsilon > 0$ eine Zahl N_ε, sodass $\|A_n - A_m\| < \varepsilon$ für $n, m \geq N_\varepsilon$ gilt. Für festes $x \in E_1 \times E_2$ erhalten wir für $m, n \geq N_\varepsilon$

$$\|A_n \langle x \rangle - A_m \langle x \rangle \| < \varepsilon \cdot \|x^1\|\|x^2\|. \tag{1}$$

Also ist $(A_n\langle x \rangle)_n$ eine Cauchyfolge in F. Weil F vollständig ist, konvergiert $(A_n\langle x \rangle)$. Wir definieren $A\langle x \rangle := \lim_{n\to\infty} A_n\langle x \rangle$. Aufgrund von Grenzwertsätzen ist auch A multilinear. Aus (1) erhalten wir durch den Grenzübergang $n \to \infty$

$$\|(A - A_m) \langle x \rangle \| < \varepsilon \cdot \|x^1\|\|x^2\|.$$

Weil x beliebig war, also $\|A - A_m\| < \varepsilon$ für $m \geq N_\varepsilon$. Damit erhalten wir aus der umgekehrten Dreiecksungleichung die Beschränktheit von A in der Operatornorm, also $A \in L(E_1, E_2; F)$, und natürlich die Konvergenz von $(A_m)_m$ gegen A bezüglich der Operatornorm. Bemerkung: Es reicht aus, dass F ein Banachraum ist.

(c) Dies folgt direkt aus den Definitionen einer linearen bzw. einer bilinearen Abbildung.

(d) Wohldefiniertheit: Sei $\alpha \in L(E_1; L(E_2, F))$. Wir müssen zeigen, dass $\psi(\alpha)$ eine bilineare Abbildung von $E_1 \times E_2$ nach F ist. Hierfür zeigen wir Linearität in der ersten und zweiten Komponente separat: Seien $x, x' \in E_1$, $y, y' \in E_2$ und $\lambda \in \mathbb{K}$.

$$\psi(\alpha)(x + \lambda x', y) = (\alpha(x + \lambda x'))(y) = (\alpha(x) + \lambda\alpha(x'))(y) = \psi(\alpha)(x, y) + \lambda\psi(\alpha)(x', y).$$

$$\psi(\alpha)(x, y + \lambda y') = (\alpha(x))(y + \lambda y') = (\alpha(x))(y) + \lambda(\alpha(x))(y') = \psi(\alpha)(x, y) + \lambda\psi(\alpha)(x, y').$$

Stetigkeiten folgen, da jeweils Verknüpfungen von stetigen Funktionen vorliegen.

Normerhaltung: Sei $\alpha \in L(E_1; L(E_2, F))$.

$$\begin{aligned} \|\psi(\alpha)\| &= \sup_{(x,y) \in E_1 \times E_2, x,y \neq 0} \frac{\|\psi(\alpha) \langle x, y \rangle\|}{\|x\| \, \|y\|} \\ &= \sup_{(x,y) \in E_1 \times E_2, x,y \neq 0} \frac{\|(\alpha\langle x \rangle) \langle y \rangle\|}{\|x\| \, \|y\|} \\ &= \sup_{x \in E_1 \setminus \{0\}} \frac{1}{\|x\|} \sup_{y \in E_2 \setminus \{0\}} \frac{\|(\alpha\langle x \rangle) \langle y \rangle\|}{\|y\|} \\ &= \sup_{x \in E_1 \setminus \{0\}} \frac{1}{\|x\|} \|\alpha\langle x \rangle\| \\ &= \|\alpha\| . \end{aligned}$$

Linearität: Seien $\alpha, \beta \in L(E_1; L(E_2, F))$ und sei λ ein Skalar. Dann gilt für alle $(x, y) \in E_1 \times E_2$:

$$\begin{aligned}
\psi(\alpha + \lambda\beta)\langle x, y\rangle &= ((\alpha + \lambda\beta)(x))(y) \\
&= (\alpha(x) + \lambda\beta(x))(y) \\
&= (\alpha(x))y + \lambda(\beta(x))(y) \\
&= \psi(\alpha)\langle x, y\rangle + \lambda\psi(\beta)\langle x, y\rangle \\
&= (\psi(\alpha) + \lambda\psi(\beta))\langle x, y\rangle .
\end{aligned}$$

Injektivität: Sei $\alpha \in L(E_1; L(E_2, F))$ mit $\psi(\alpha) = 0$. Dann gilt für alle $(x, y) \in E_1 \times E_2$, dass $\alpha\langle x\rangle\langle y\rangle = 0$. Insbesondere gilt dann $\alpha(x) = 0$. Da x beliebig war, folgt $\alpha = 0$.

Surjektivität: Sei $A \in L(E_1, E_2; F)$. Wir definieren $\alpha_A \in L(E_1; L(E_2, F))$ durch $(\alpha_A(x))(y) = A\langle x, y\rangle$. Mit dieser Definition lässt sich leicht nachrechnen, dass α wohldefiniert ist und $\psi(\alpha_A) = A$ gilt. (Wir könnten φ durch $\varphi(A) = \alpha_A$ definieren und nun alle Schritte oben durchgehen, um nachzuweisen, dass φ wohldefiniert und linear ist. Daraus folgt dann $\varphi = \psi^{-1}$.)

Aufgabe 11.3 (Distanz zum Einheitsquadrat) **[1 + 3 Punkte]**

Sei $Q = \{(x, y) \in \mathbb{R}^2 \colon \max\{|x|, |y|\} = 1\}$ das Einheitsquadrat. Wir betrachten die Distanzfunktion

$$\Delta\colon \mathbb{R}^2 \to \mathbb{R},\ a \mapsto \inf_{\xi \in Q} \|a - \xi\| .$$

(i) Zeichnen Sie die Niveaulinen von Δ, d. h. die Mengen

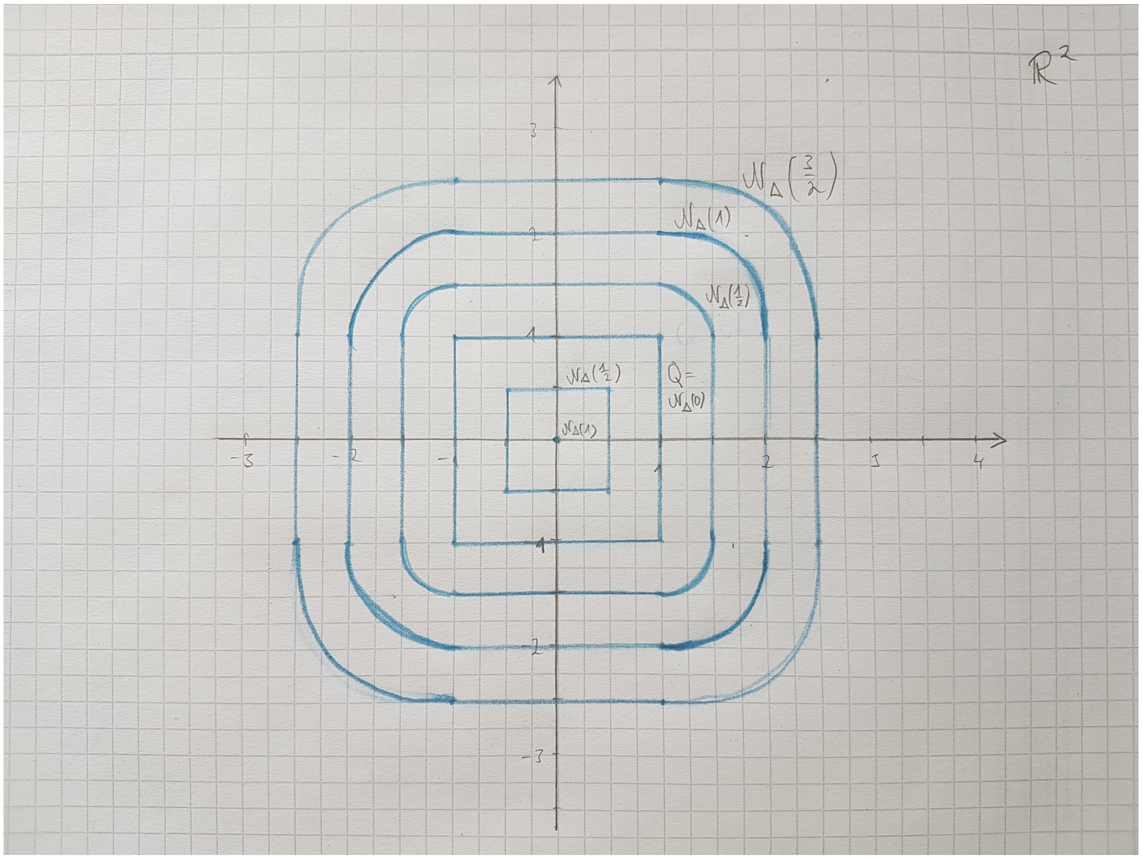
$$\mathcal{N}_\Delta(c) := \Delta^{-1}(\{c\}) \subset \mathbb{R}^2,$$

für $c = 0$, $c = \frac{1}{2}$, $c = 1$ und $c = \frac{3}{2}$.

(ii) Sei $f = \Delta|_{(0,\infty)^2}$. Prüfen Sie f auf Differenzierbarkeit und berechnen Sie die Ableitung von f in allen Punkten, in denen es differenzierbar ist.

Lösung:

(i)



(Die Rundungen sollen Viertelkreise sein, die zur nächstliegenden Ecke des Einheitsquadrat den jeweiligen Abstand c haben.)

(ii) *(In dieser Aufgabe geht es hauptsächlich um grafisches Verständnis, nicht um formelle Beweise, die bis ins kleinste Detail stimmen. Die Lösungen hier zeigen jedoch einen hohen Detailgrad.)*

Wir partitionieren die Menge $X = (0, \infty)^2$ in die folgenden Teilmengen:

$$\begin{aligned}
A_1 &:= \{(x, y) \in X : y < x < 1\}, \\
A_2 &:= \{(x, y) \in X : x < y < 1\}, \\
B_1 &:= \{(x, y) \in X : y < 1 < x\}, \\
B_2 &:= \{(x, y) \in X : x < 1 < y\}, \\
C &:= \{(x, y) \in X : 1 < x, 1 < y\}, \\
D &:= \{(x, y) \in X : x = y \leq 1\}, \\
E_1 &:= \{(x, y) \in X : y = 1, x > 1\}, \\
E_2 &:= \{(x, y) \in X : x = 1, y > 1\}, \\
F &:= (Q \cap X) \setminus \{(1, 1)\}.
\end{aligned}$$

Die Funktion ist in allen Punkten der Mengen A_1, A_2, B_1, B_2, C, E_1 und E_2 differenzierbar. Nur in den Mengen D und F ist f nicht differezierbar.

Im Folgenden berechnen wir die Ableitungen innerhalb der jeweiligen Mengen. Hierbei berechnen wir ∇f, da f' dann durch Definition 6.8 gegeben ist.

- A_1: Hier ist die Funktion durch $f(x,y) = 1 - x$ gegeben. Damit haben wir

$$\nabla f(x,y) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

- A_2: Hier ist die Funktion durch $f(x,y) = 1 - y$ gegeben. Damit haben wir

$$\nabla f(x,y) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

- B_1: Hier ist die Funktion durch $f(x,y) = x - 1$ gegeben. Damit haben wir

$$\nabla f(x,y) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

- B_2: Hier ist die Funktion durch $f(x,y) = y - 1$ gegeben. Damit haben wir

$$\nabla f(x,y) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

- C: Hier ist die Funktion durch

$$f(x,y) = \|(x,y) - (1,1)\| = \sqrt{(x-1)^2 + (y-1)^2}$$

gegeben. Damit haben wir

$$\nabla f(x,y) = \frac{1}{\sqrt{(x-1)^2 + (y-1)^2}} \begin{pmatrix} x-1 \\ y-1 \end{pmatrix}.$$

- D: Hier ist die Funktion durch $f(x,y) = 1 - x = 1 - y$ gegeben. Sie ist nicht differenzierbar, da sie nicht partiell differenzierbar ist: Sei $(x_0, y_0) \in D$. Dann ist $x_0 \in (0,1]$ und $y_0 = x_0$. Inbesondere erhalten wir für $x_0 \neq 1$

$$\lim_{t\downarrow 0} \frac{f(x_0 + t, x_0) - f(x_0, x_0)}{t} = \lim_{t\downarrow 0} \frac{(1 - (x_0 + t)) - (1 - x_0)}{t} = -1,$$

bzw. für $x_0 = 1$

$$\lim_{t\downarrow 0} \frac{f(1+t,1) - f(1,1)}{t} = \lim_{t\downarrow 0} \frac{t - 0}{t} = 1,$$

aber

$$\lim_{t\uparrow 0} \frac{f(x_0 + t, x_0) - f(x_0, x_0)}{t} = \lim_{t\uparrow 0} \frac{(1 - x_0) - (1 - x_0)}{t} = 0.$$

- E_1: Hier ist die Funktion durch $f(x,y) = f(x,1) = x - 1$ gegeben. Sie ist stetig partiell differenzierbar: Sei $x_0 > 1$. Wir haben

$$\frac{\partial f}{\partial x}(x_0, 1) = \lim_{t\to 0} \frac{f(x_0 + t, 1) - f(x_0, 1)}{t} = \lim_{t\to 0} \frac{(x_0 + t - 1) - (x_0 - 1)}{t} = 1$$

sowie

$$\begin{aligned}\lim_{t\downarrow 0} \frac{f(x_0, 1+t) - f(x_0, 1)}{t} &= \lim_{t\downarrow 0} \frac{\sqrt{(x_0-1)^2 + t^2} - (x_0 - 1)}{t} \\ &= \lim_{t\downarrow 0} \left(\sqrt{\frac{(x_0-1)^2}{t^2} + 1} - \sqrt{\frac{(x_0-1)^2}{t^2}} \right) \\ &= 0,\end{aligned}$$

und

$$\lim_{t\uparrow 0} \frac{f(x_0, 1+t) - f(x_0, 1)}{t} = \lim_{t\uparrow 0} \frac{(x_0 - 1) - (x_0 - 1)}{t} = 0,$$

also

$$\frac{\partial f}{\partial y}(x_0, 1) = 0.$$

Da E_1 im Rand von B_1 liegt und die partiellen Ableitungen übereinstimmen, ist die Funktion stetig partiell differenzierbar und somit auch differenzierbar. Der Gradient ist der gleiche wie in B_1.

- E_2: Hier ist die Funktion durch $f(x, y) = f(1, y) = y - 1$ gegeben. Analog wie oben in E_1 ist die Funktion differenzierbar mit gleichem Gradienten wie in B_2.
- F: Hier kann man ähnlich wie bei der eindimensionalen Betragsfunktion zeigen, dass die Funktion nicht partielle differenzierbar ist. Betrachte z.B. den Punkt $(1, 1/2)$. Dann gilt

$$\lim_{t\downarrow 0} \frac{f(1+t, 1/2) - f(1, 1/2)}{t} = \lim_{t\downarrow 0} \frac{|t| - 0}{t} = 1,$$

aber

$$\lim_{t\uparrow 0} \frac{f(1+t, 1/2) - f(1, 1/2)}{t} = \lim_{t\uparrow 0} \frac{|t| - 0}{t} = -1.$$

Aufgabe 11.4 (Numerische Suche nach einem Minimum) **[1 + 1* + 3 + 4* Punkte]**
Sei $n \in \mathbb{N}_{>0}$ und sei $u \in C^1(\mathbb{R}^n)$. Die Funktion u heißt *koerziv*, falls $u(x) \to \infty$ für $\|x\| \to \infty$ gilt. Sei $(\lambda_k)_k$ eine Folge in $[0, 1]$ und sei $x_0 \in \mathbb{R}^n$. Die Folge $(x_k)_k$ in $\mathbb{R}^n$ sei rekursiv durch

$$x_{k+1} := x_k - \lambda_{k+1} \nabla u(x_k)$$

für alle $k \in \mathbb{N}$ gegeben. Wir sagen, dass u die *Abstiegseigenschaft ab x_0 (mit Schrittfaktorfolge $(\lambda_k)_k$)* hat, falls es für alle $k \in \mathbb{N}$ die Ungleichung

$$u(x_{k+1}) \leq u(x_k) - \frac{\lambda_{k+1}}{3} \cdot \|\nabla u(x_k)\|^2$$

erfüllt.

(i) Sei $(\lambda_k)_k$ die konstante Folge mit $\lambda_k = \frac{1}{4}$ für alle $k \in \mathbb{N}$. Finden Sie eine koerzive Funktion $s \in C^1(\mathbb{R}^2)$, die für alle $x_0 \in \mathbb{R}^2$ die Abstiegseigenschaft ab x_0 hat.

(ii)* Sei $(\lambda_k)_k$ die konstante Folge mit $\lambda_k = \frac{1}{10}$ für alle $k \in \mathbb{N}$. Finden Sie eine koerzive Funktion $w \in C^1(\mathbb{R}^2)$, die für kein $x_0 \in \mathbb{R}^2 \setminus \{0\}$ die Abstiegseigenschaft ab x_0 hat.

(iii) Sei $u \in C^1(\mathbb{R}^n)$ koerziv und sei $x_0 \in \mathbb{R}^n$. Wir nehmen an, dass u die Abstiegseigenschaft ab x_0 mit konstanter Schrittfaktorfolge $(\lambda_k)_k$ mit $\lambda_k = \frac{1}{4}$ besitzt. Zeigen Sie, dass eine Teilfolge der Folge $(x_k)_k$ gegen einen kritischen Punkt von u konvergiert.

(Hinweis: Zeigen Sie dazu zunächst, dass es für jedes $\varepsilon > 0$ und jedes $k_0 \in \mathbb{N}$ ein $k \geq k_0$ mit $\|\nabla u(x_k)\| < \varepsilon$ gibt.)

(iv)* Sei $v \in C^1(\mathbb{R}^n)$ koerziv und sei $x_0 \in \mathbb{R}^n$. Für alle $k \in \mathbb{N}$ wählen wir $m \in \mathbb{N}_{>0}$ minimal mit

$$v\left(x_k - \frac{\nabla v(x_k)}{m}\right) \leq v(x_k) - \frac{\|\nabla v(x_k)\|^2}{3m}$$

und setzen $\lambda_{k+1} = \frac{1}{m}$.

Zeigen Sie, dass die Folge $(\lambda_k)_k$ wohldefiniert ist und v die Abstiegseigenschaft ab x_0 hat. Folgern Sie, dass $(x_k)_k$ eine Teilfolge besitzt, die gegen einen kritischen Punkt von v konvergiert.

Lösung:

(i) Setze $s(x) = \|x\|^2$. Diese Funktion ist offensichtlich koerziv.

Sei $x_0 \in \mathbb{R}^2$. Wir haben $\nabla s(x) = 2x$ für alle $x \in \mathbb{R}^2$. Sei $k \in \mathbb{N}$. Dann gilt

$$\begin{aligned} s(x_{k+1}) &= s\left(x_k - \frac{2x_k}{4}\right) \\ &= s\left(\frac{x_k}{2}\right) \\ &= \frac{\|x_k\|^2}{4} \\ &\leq \frac{2}{3}\|x_k\|^2 \\ &= \|x_k\|^2 - \frac{4\|x_k\|^2}{12} \\ &= s(x_k) - \frac{\|\nabla s(x_k)\|^2}{12}. \end{aligned}$$

Dies zeigt die Behauptung.

(ii) Setze $w_c(x) = c\|x\|^2$ für ein $c > 0$. Diese Funktion ist offensichtlich koerziv.

Sei $x_0 \in \mathbb{R}^2$ mit $x_0 \neq 0$. Wir haben $\nabla w(x) = 2cx$ für alle $x \in \mathbb{R}^2$. Sei $k \in \mathbb{N}$. Dann gilt

$$\begin{aligned} w_c(x_{k+1}) &= w\left(x_k - \frac{2cx_k}{10}\right) \\ &= w_c\left(\frac{(5-c)x_k}{5}\right) \\ &= \frac{(5-c)^2 c\|x_k\|^2}{25} \end{aligned}$$

und

$$w_c(x_k) - \frac{1}{30}\|2cx_k\|^2 = \frac{15c - 2c^2}{15}\|x_k\|^2.$$

Wir benötigen also ein c mit

$$\frac{(5-c)^2 c}{25} > \frac{15c - 2c^2}{15}.$$

Diese Ungleichung ist für alle genügend großen $c > 0$ erfüllt.

(iii) Da die Folge $(u(x_k))_k$ monoton fällt, ist sie nach oben beschränkt. Wäre die Folge $(x_k)_k$ unbeschränkt so gäbe es eine Teilfolge $(y_j)_j$ mit $y_j = x_{k_j}$, sodass $\|y_j\| \to \infty$ für $j \to \infty$. Daraus würde folgen $u(y_j) \to \infty$ für $j \to \infty$, ein Widerspruch. Damit ist die Folge $(x_k)_k$ beschränkt.

Wir zeigen nun, dass die Folge $(u(x_k))_k$ nach unten beschränkt ist. Wäre sie es nicht, so gäbe es eine Teilfolge $(y_j)_j$ mit $y_j = x_{k_j}$, sodass $u(y_j) \to -\infty$ für $j \to \infty$. Nun ist die Folge $(y_j)_j$ als Teilfolge von $(x_k)_k$ auch beschränkt und besitzt somit eine konvergente Teilfolge $(z_\ell)_\ell$ mit $z_\ell = y_{j_\ell}$. Wegen der Stetigkeit von u erhalten wir

$$-\infty = \lim_{\ell\to\infty} u(z_\ell) = u\left(\lim_{\ell\to\infty} z_\ell\right),$$

ein Widerspruch.

Wir zeigen nun den Hinweis. Angenommen, es gibt ein $\varepsilon > 0$ und ein k_0, sodass für alle $\ell \in \mathbb{N}$ gilt: $\|\nabla u(x_{k_0+\ell})\| \geq \varepsilon$. Wir erhalten:

$$u(x_{k_0+1}) \leq u(x_{k_0}) - \frac{\|\nabla u(x_{k_0})\|}{12} \leq u(x_{k_0}) - \frac{\varepsilon}{12},$$

$$u(x_{k_0+2}) \leq u(x_{k_0+1}) - \frac{\|\nabla u(x_{k_0+1})\|}{12} \leq u(x_{k_0}) - \frac{2\varepsilon}{12},$$

und dies weiterführend induktiv

$$u(x_{k_0+\ell}) \leq u(x_{k_0}) - \frac{\ell\varepsilon}{12},$$

für alle $\ell \in \mathbb{N}$. Insbesondere ist die Folge $(u(x_k))_k$ nach unten unbeschränkt, ein Widerspruch.

Mit dem Hinweis erhalten wir eine Teilfolge $(y_j)_j$ von $(x_k)_k$, sodass

$$\nabla u(y_j) \to 0$$

für $j \to \infty$. Da die Folge $(x_k)_k$ beschränkt ist, ist auch $(y_j)_j$ beschränkt und hat somit eine konvergente Teilfolge. Für den Grenzwert y dieser Teilfolge gilt wegen der Stetigkeit von ∇u, dass $\nabla u(y) = 0$, also y ein kritischer Punkt von u ist.

(iv) Sei $u = v$. Wir können ähnlich wie in Aufgabenteil (iii) vorgehen. Lediglich die Wohldefiniertheit der Folge $(\lambda_k)_k$ und der Beweis des Hinweises sind zu zeigen. Die Abstiegseigenschaft folgt direkt aus der Wahl von m.

Wohldefiniertheit der Folge: Sei $x \in \mathbb{R}^n$ und sei $h > 0$. Dann gilt

$$\begin{aligned} u(x - h\nabla u(x_k)) &= u(x) - hDu(x)\langle \nabla u(x)\rangle + o(\|h\nabla u(x)\|) \\ &= u(x) - h\|\nabla u(x)\|^2 + o(\|h\nabla u(x)\|). \end{aligned}$$

Letzteres ist für genügend kleines h gleiner oder gleich

$$u(x) - \frac{h}{3}\|\nabla u(x)\|^2.$$

Dies impliziert die Existent eines geeigneten m. Nun ist dieses noch minimal zu wählen.

Wir zeigen nun den Hinweis: Sei $R > 0$ mit $(x_k)_k \subset B_R(0)$. Betrachte für $\varepsilon > 0$ die Menge

$$K_\varepsilon = \left\{x \in \overline{B_R(0)} \colon \|\nabla u(x)\| \geq \varepsilon\right\}.$$

Sei $k_0 \in \mathbb{N}$. Wir zeigen, dass es ein Folgeglied x_{k_ε} von $(x_k)_k$ mit $k_\varepsilon > k_0$ gibt, das nicht in K_ε liegt.

Angenommen die obige Aussage ist falsch. Dann gibt es ein $\varepsilon > 0$ und ein $k_0 \in \mathbb{N}$, sodass für alle $k \geq k_0$ gilt $x_k \in K_\varepsilon$.

K_ε ist kompakt. Zu $x \in K_\varepsilon$ gibt es ein $\delta_x > 0$ mit

$$\|\nabla u(x) - \nabla u(y)\| \leq \frac{1}{20}\|\nabla u(x)\|$$

für alle $y \in B_\delta(x)$. Seien nun $y, z \in B_{\delta_x}(x)$. Dann folgt

$$\begin{aligned}\left|\langle \nabla u(y), \nabla u(z)\rangle - \|\nabla u(y)\|^2\right| &= |\langle \nabla u(y), \nabla u(z) - \nabla u(y)\rangle| \\ &\leq \|\nabla u(y)\| \left(\|\nabla u(z) - \nabla u(x)\| + \|\nabla u(x) - \nabla u(y)\|\right) \\ &\leq \frac{1}{10} \|\nabla u(x)\| \|\nabla u(y)\| .\end{aligned}$$

Den Faktor $\|\nabla u(x)\|$ wollen wir weiter abschätzen.

$$\begin{aligned}\|\nabla u(x)\| &\leq \|\nabla u(x) - \nabla u(y)\| + \|\nabla u(y)\| \\ &\leq \frac{1}{10} \|\nabla u(x)\| + \|\nabla u(y)\| .\end{aligned}$$

Wir erhalten

$$\|\nabla u(x)\| \leq \frac{10}{9} \|\nabla u(y)\| .$$

Also folgt

$$\|\nabla u(y)\|^2 - \langle \nabla u(y), \nabla u(z)\rangle \leq \frac{1}{9} \|\nabla u(y)\|^2 .$$

Daraus folgt

$$\frac{1}{2} \|\nabla u(y)\|^2 \leq \frac{8}{9} \|\nabla u(y)\|^2 \leq \langle \nabla u(y), \nabla u(z)\rangle .$$

Nun überdecken endlich viele Bälle der Form $B_{\delta_x/2}(x)$ die abgeschlossene Kugel K_ε. Seien $\{x_1, \ldots, x_\ell\}$ die zugehörigen Elemente aus $\mathbb{R}^n$. Setze

$$\xi = \frac{1}{2} \min\{\delta_{x_1}, \ldots, \delta_{x_\ell}\}.$$

Für alle $x \in K_\varepsilon$ haben wir nun

$$u(x - t\nabla u(x)) \leq u(x) - \frac{1}{2} t \|\nabla u(x)\|^2$$

für alle $t \in [0, \xi / \|\nabla u(x)\|]$. Bemerke, dass $\xi / \|\nabla u(x)\| \leq \xi / \varepsilon$. Somit gilt

$$\frac{\xi}{2\varepsilon} \leq \frac{1}{m},$$

wobei m wie in der Definition der Aufgabe ist. Insbesondere erhalten wir

$$u(x_{k+1}) \leq u(x_k) - \frac{1}{3} \frac{\xi}{2\varepsilon} \varepsilon^2.$$

Wir erhalten nun durch eine Induktion wie in Aufgabenteil (iii) den gewünschten Widerspruch.

Abgabe: Bis **Freitag, 03. Juli 2020, 09:54 Uhr**, direkt an die Tutorin / den Tutor. Wir bitten die allgemeinen Hinweise zur Abgabe von Lösungen (siehe Homepage) zu beachten.