Übung 13 (Lösung)

Ausgabedatum: 20. Januar 2025
Abgabedatum: 27. Januar 2025

Hausaufgabe 13.1 (Richtungsableitungen von Normen) 2 + 2 + 2 = 6 Punkte

Verifizieren Sie Beispiel 16.13, also die folgenden Aussagen für $d \in \mathbb{R}^n$:

(a) Für $f(x) = \|x\|_1 = \sum_{i=1}^n |x_i|$ ist

$$f'(x;d) = \sum_{\substack{i=1\\x_i>0}}^n d_i - \sum_{\substack{i=1\\x_i<0}}^n d_i + \sum_{\substack{i=1\\x_i=0}}^n |d_i|.$$

(b) Für $f(x) = \|x\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ ist

$$f'(x;d) = \begin{cases} \frac{x^\top d}{\|x\|_2}, & \text{falls } x \neq 0, \\ \|d\|_2, & \text{falls } x = 0. \end{cases}$$

(c) Für $f(x) = \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$ und $x \neq 0$ ist

$$f'(x;d) = \max\left\{(\operatorname{sgn} x_i)\, d_i \,\middle|\, i = 1,\ldots,n,\ |x_i| = \|x\|_\infty\right\}$$

sowie

$$f'(0;d) = \|d\|_\infty.$$

Lösung.

Am einfachsten folgen die Aussagen mit der max formula

$$f'(x;d) = \sup\left\{s^\top d \,\middle|\, s \in \partial f(x)\right\} \in \mathbb{R} \cup \{\infty\}.$$

aus Satz 16.18 und aus der Charakterisierung der Subdifferentiale der gegebenen Normen von Blatt 11 bzw. Beispiel 16.5 aus dem Skript. Da in den Fällen der Normen $\operatorname{dom} f = \mathbb{R}^n$ ist, wissen wir auch, dass das Supremum ein Maximum ist.

(a) Wir wissen, dass $s \in \partial \|x\|_1$ genau dann, wenn

$$s_i \in \begin{cases} \{-1\} & \text{falls } x_i < 0, \\ [-1, 1] & \text{falls } x_i = 0, \\ \{1\} & \text{falls } x_i > 0 \end{cases}$$

für alle $i = 1, \dots, n$. Es ist also für $s \in \partial f(x)$

$$s^\top d = \sum_{\substack{i=1 \\ x_i > 0}}^n d_i - \sum_{\substack{i=1 \\ x_i < 0}}^n d_i + \sum_{\substack{i=1 \\ x_i = 0}}^n s_i d_i$$

was für $s_i = \operatorname{sgn}(d_i)$ für i mit $x_i = 0$ über das Subdifferential sein Maximum annimmt. (2 Punkte)

(b) Wir wissen, dass

$$\partial \|x\|_2 = \begin{cases} \{\frac{x}{\|x\|_2}\}, & \text{falls } x \neq 0, \\ \{s \in \mathbb{R}^n \mid \|s\|_2 \leqslant 1\}, & \text{falls } x = 0. \end{cases}$$

Überall außerhalb der Null ist die 2-Norm differenzierbar und das Subdifferential einelementig, das Maximum ist also sofort ablesbar und es gilt

$$s^\top d = \left(\frac{x}{\|x\|_2} \right)^\top d = \left(\frac{x^\top d}{\|x\|_2} \right)$$

wie behauptet. Für $x = 0$ und $s \in \partial f(0) = \{s \in \mathbb{R}^n \mid \|s\|_2 \leq 1\}$ ist nach der Cauchy-Schwarzschen Ungleichung

$$s^\top d \leq \|s\|_2 \|d\|_2 = \|d\|_2$$

und für $s = \frac{d}{\|d\|_2}$ wird der Wert mit Gleichheit angenommen, es handelt sich also um einen Maximierer, weshalb

$$f'(0; d) = \|d\|_2$$

(2 Punkte)

(c) Wir wissen, dass für $x \neq 0$ das Subdifferential die Form

$$\partial f(x) = \left\{ s \in \mathbb{R}^n \;\middle|\; \begin{array}{r} \|s\|_1 = 1,\ s_i \geqslant 0 \text{ für diejenigen } i \text{ mit } x_i = \|x\|_\infty, \\ s_i \leqslant 0 \text{ für diejenigen } i \text{ mit } -x_i = \|x\|_\infty, \\ \text{und } s_i = 0 \text{ für diejenigen } i \text{ mit } |x_i| < \|x\|_\infty \end{array} \right\}$$

hat, und dass

$$\partial \|0\|_\infty = \{s \in \mathbb{R}^n \mid \|s\|_1 \leqslant 1\}.$$

Für $x = 0$ und $s \in \partial f(0)$ ist

$$s^\mathsf{T} d = \sum_{i=1}^n s_i d_i \leq \sum_{i=1}^n |s_i||d_i| \leq \|s\|_1 \|d\|_\infty \leq \|d\|_\infty$$

und Gleichheit wird für jedes $s \in \partial f(0)$ mit $\operatorname{sgn}(s_i) = \operatorname{sgn}(d_i)$ und $s_i = 0$ wenn $d_i < \|d\|_\infty$ angenommen.

Für $x \neq 0$ und $s \in \partial f(x)$ ist

$$\begin{aligned} s^\mathsf{T} d &= \sum_{i, |x_i| = \|x\|_\infty} s_i d_i \\ &\leq \max\left\{(\operatorname{sgn} x_i)\, d_i \,\middle|\, i = 1, \ldots, n,\ |x_i| = \|x\|_\infty\right\} \end{aligned}$$

und Gleichheit wird erreicht für s komplett Null außer $s_i = \operatorname{sgn}(x_i)$ für ein $i \in \operatorname{arg\,max}\left\{(\operatorname{sgn} x_i)\, d_i \,\middle|\, i = 1, \ldots, n,\ |x_i| = \|x\|_\infty\right\}$ (2 Punkte)

Hausaufgabe 13.2 (Kettenregel für die Richtungsableitungen) 2 + 2 = 4 Punkte

(a) Geben Sie ein Beispiel für Funktionen $f, g\colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$, $x_0 \in \operatorname{dom} g$ und $d \in \mathbb{R}$ an, das zeigt, dass i. A. keine Kettenregel für die Richtungsableitung gilt.

(b) Wir können die Definition 16.12 der Richtungsableitung direkt auf vektorwertige Funktionen $f\colon R^n \to \mathbb{R}^m$ übertragen, indem wir fordern, dass der Grenzwert existieren muss, aber in jeder Komponente endlich ist.

Es seien $f\colon \mathbb{R}^m \to \mathbb{R}^l$, $g\colon \mathbb{R}^n \to \mathbb{R}^m$ gegeben, so dass

(i) die Richtungsableitung von g an x_0 in Richtung d existiert (und endlich ist),

(ii) die Richtungsableitung von f an $g(x_0)$ in Richtung $g'(x_0; d)$ existiert (und endlich ist) und

(iii) die Funktion f in einer Umgebung von $g(x_0)$ Lipschitz-stetig ist.

Zeigen Sie, dass in diesem Fall die Richtungsableitung von $f \circ g$ an x_0 in Richtung d existiert, und dass

$$f \circ g'(x_0; d) = f'(g(x_0); g'(x_0, d)) \qquad \text{für alle } d \in \mathbb{R}^n. \tag{0.1}$$

Lösung.

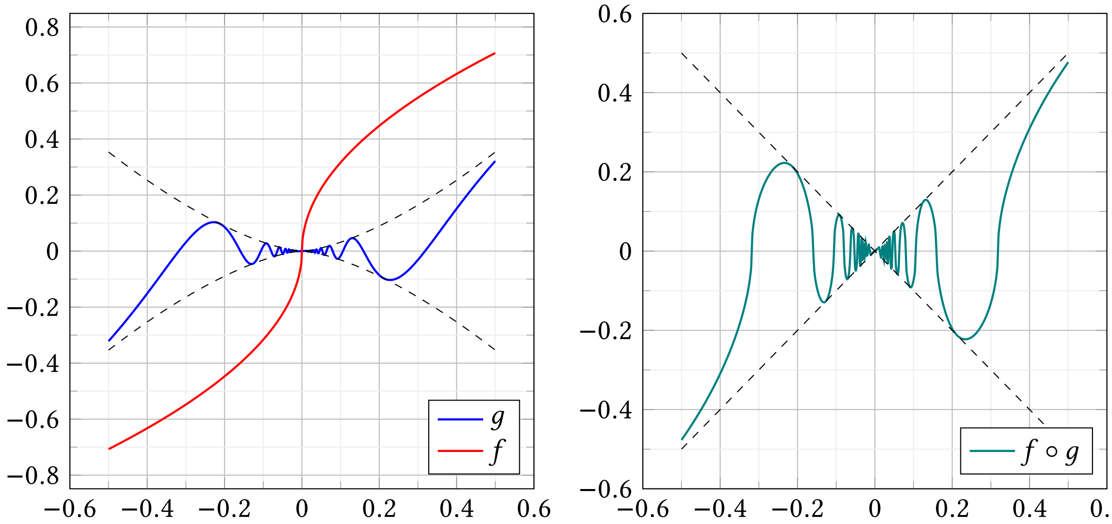

Abbildung 0.1: Darstellung des charakteristischen Verhaltens von f, g und $f \circ g$. Statt g ist auf Grund von Darstellungsschwierigkeiten die Funktion $x^{1.25} \sin(1/x)$ geplottet, die etwas langsamer wächst.

(a) Wir können beispielsweise die Funktionen

$$g(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases} \quad \text{und} \quad f(x) = \operatorname{sgn}(x)\sqrt{|x|}$$

am Punkt $x_0 = 0$ und in Richtung $d = 1$ betrachten.

Dann ist

$$g'(0; d) = \lim_{t \to 0^+} \frac{t^2 \sin\left(\frac{1}{t}\right)}{t} = 0$$
$$f'(0; g'(0; d)) = f'(0; 0) = 0$$

aber für die Nullfolge $t_k = \frac{1}{\frac{\pi}{2} + 2k\pi} > 0$ ist $\sin\left(\frac{1}{t_k}\right) = 1$ und damit g immer positiv und für die Nullfolge $\tilde{t}_k = \frac{1}{\frac{3\pi}{2} + 2k\pi} > 0$ ist $\sin\left(\frac{1}{\tilde{t}_k}\right) = -1$ und damit g immer negativ, und somit

$$\frac{f \circ g(t_k) - f \circ g(0)}{t_k} = \frac{f \circ g(t_k)}{t_k} = \frac{\operatorname{sgn}(g(t_k))\sqrt{|g(t_k)|}}{t_k} = \frac{|t_k|}{t_k}\sqrt{\left|\sin\left(\frac{1}{t_k}\right)\right|} = \operatorname{sgn}(t_k) = 1 \xrightarrow{k \to \infty} 1$$

aber

$$\frac{f\circ g(\tilde t_k) - f\circ g(0)}{\tilde t_k} = \frac{f\circ g(\tilde t_k)}{\tilde t_k} = \frac{\operatorname{sgn}(g(\tilde t_k))\sqrt{|g(\tilde t_k)|}}{\tilde t_k} = -\frac{|\tilde t_k|}{\tilde t_k}\sqrt{\left|\sin\left(\frac{1}{\tilde t_k}\right)\right|} = -\operatorname{sgn}(\tilde t_k) = -1 \xrightarrow{k\to\infty} -1$$

womit die Richtungsableitung nicht existiert. (2 Punkte)

(b) Es sei also f in $B_r(g(x_0))$ Lipschitz-stetig mit Konstante $L > 0$. Dann gilt

$$g(x_0 + td) = g(x_0) + tg'(x_0;d) + t\left(\frac{g(x+td) - g(x)}{t} - g'(x_0;d)\right) \xrightarrow{t\to 0^+} g(x_0)$$

und somit für genügend kleine t, so dass $g(x_0+td), g(x_0) + tg'(x_0;d) \in B_r(g(x_0))$, auch

$$\begin{aligned}
&\left\|\frac{f(g(x_0+td)) - f(g(x_0))}{t} - f'(g(x_0); g'(x_0;d))\right\| \\
\le{}& \left\|\frac{f(g(x_0+td)) - f(g(x_0) + tg'(x_0;d))}{t}\right\| \\
&+ \left\|\frac{f(g(x_0) + tg'(x_0;d)) - f(g(x_0))}{t} - f'(g(x_0); g'(x_0;d))\right\| \\
\le{}& L\left\|\frac{g(x_0+td) - g(x_0) - tg'(x_0;d)}{t}\right\| \\
&+ \left\|\frac{f(g(x_0) + tg'(x_0;d)) - f(g(x_0))}{t} - f'(g(x_0); g'(x_0;d))\right\| \xrightarrow{t\to 0^+} 0.
\end{aligned}$$

(2 Punkte)

Hausaufgabe 13.3 (Richtungsableitung und Subdifferential für max diffbarer Funktionen) 3 + 3 = 6 Punkte

Es seien $C \subseteq \mathbb{R}^n$ eine offene und konvexe Menge und $f_i\colon C \to \mathbb{R}$, $i = 1,\ldots,m$, konvexe und stetig differenzierbare Funktionen sowie für $x \in C$

$$f(x) \coloneqq \max_{i=1,\ldots,m} f_i(x)$$

und $A(x) \coloneqq \{i \in \{1,\ldots,m\} \mid f_i(x) = f(x)\}$. Zeigen Sie:

(a) Die Funktion f besitzt für alle Richtungen $d \in \mathbb{R}^n$ die Richtungsableitung

$$f'(x,d) = \max_{i\in A(x)} f_i'(x)d.$$

(b) Das Subdifferential der Funktion f ist

$$\partial f(x) = \operatorname{conv}\{f_i'(x)^\top \mid i \in A(x)\}.$$

Lösung.

(a) Wir stellen fest, dass $f = g \circ h$ für

$$\begin{aligned} h\colon C \subseteq \quad & \mathbb{R}^n \to \mathbb{R}^m, & h(x) &= (f_1(x), \ldots, f_m(x))^\top \\ g\colon \quad & \mathbb{R}^m \to \mathbb{R}, & g(x) &= \max_{i=1,\ldots,m} x_i \end{aligned}$$

und wollen die Kettenregel für die Richtungsableitung anwenden.

Da g global Lipschitz-stetig mit Konstante $L = 1$ ist, können wir aus Aufgabe 3 folgern, dass die Kettenregel für die Richtungsableitung gilt. (1 Punkt)

Da die f_i differenzierbar sind, ist auch h differenzierbar und es gilt

$$h'(x) = (f_1'(x)^\top, \ldots, f_m'(x)^\top)^\top.$$

und somit

$$h'(x; d) = (f_1'(x)d, \ldots, f_m'(x)d)^\top.$$

(1 Punkt)

Es verbleibt zu zeigen, dass $g'(x; d) = \max_{i \in \tilde{A}(x)} d_i$ für jedes $x \in \mathbb{R}^m$ und die Menge $\tilde{A}(x) \coloneqq \{i \in \{1, \ldots, m\} \mid x_i = g(x)\}$. Für $i \in \tilde{I}(x) \coloneqq \{1, \ldots, m\} \setminus \tilde{A}(x)$ ist $x_i + t\, d_i < \max_{i=1,\ldots,m} x_i = g(x)$ für t ausreichend klein, also ist $\tilde{I}(x + td) = \tilde{I}(x)$ für t ausreichend klein und somit ebenfalls für t ausreichend klein

$$g(x + td) = \max_{i \in \tilde{A}(x)} x_i + t\, d_i = \max_{i=1,\ldots,m} (x_i) + t \max_{i \in \tilde{A}(x)} (d_i) = g(x) + t \max_{i \in \tilde{A}(x)} d_i.$$

(1 Punkt)

(b) Auf Grund von Aussage (a) ist $f_i'(x)^\top \in \partial f(x)$ für alle $i \in A(x)$ und da das Subdifferential konvex ist, ist auch $\operatorname{conv}\{(f_i'(x))^\top \mid i \in A(x)\} \subseteq \partial f(x)$. (1 Punkt)

Angenommen, es gäbe nun noch ein $s \in \partial f \setminus \operatorname{conv}\{(f_i'(x))^\top \mid i \in A(x)\}$. Die Menge $\{s\}$ ist offensichtlich kompakt. Die Menge $\operatorname{conv}\{(f_i'(x))^\top \mid i \in A(x)\}$ ist das Bild des (kompakten) Einheitssimplex $S \coloneqq \{\alpha \in \mathbb{R}^{|Ax|} \mid \alpha \geq 0,\ \sum_{i \in A(x)} \alpha_i = 1\}$ unter der stetigen Abbildung $\alpha \mapsto \sum_{i \in A(x)} \alpha_i f_i'(x)$ und somit ebenfalls kompakt. Wir können dann auf Grund von Satz 15.35 die kompakten, disjunkten und abgeschlossenen Mengen s und $\operatorname{conv}\{(f_i'(x))^\top \mid i \in A(x)\}$ strikt trennen. Es gibt also ein $a \in \mathbb{R}^n$, so dass

$$a^\top f_i'(x)^\top < a^\top s \quad \text{für alle } i \in A(x)$$

also

$$\max_{i \in A(x)} f_i'(x)a = f'(x;a) < s^\top a$$

im Widerspruch zu der max formula in Satz 16.18.

Bemerkung. Das sehr intuitive Resultat, dass die konvexe Hülle endlich vieler Punkte ein kompaktes Polyeder sein muss, wird in der Literatur unter anderem als "representation theorem for polytopes" bezeichnet und benötigt erstaunlich viel Maschinerie (Farkas Lemma, Satz von Caratheodory, Motzkin Eliminierung) um leicht zu folgen.

(2 Punkte)

Bitte reichen Sie Ihre Lösungen als ein PDF auf Mampf ein.