

$$\text{Setze } x_2 = x_2^* \Rightarrow \|x_1^* - \hat{x}\| \leq \|x_1 - \hat{x}\| \quad \text{für alle } x_1 \in C_1,$$

$$\text{Setze } x_1 = x_1^* \Rightarrow \|x_2^* - \hat{x}\| \leq \|x_2 - \hat{x}\| \quad \text{für alle } x_2 \in C_2,$$

d. h., es gilt (**).

Aus dem [Projektionssatz 12.2](#) folgt daher

$$\underbrace{(x_1^* - \hat{x})}_{=-a}^\top (x_1 - x_1^*) \geq 0 \quad \text{für alle } x_1 \in C_1.$$

$$\Rightarrow a^\top x_1 \leq a^\top x_1^* = a^\top \hat{x} + a^\top (x_1^* - \hat{x}) = \beta - \|a\|^2 < \beta \quad \text{für alle } x_1 \in C_1.$$

Analog: $a^\top x_2 \geq \beta + \|a\|^2 > \beta$ für alle $x_2 \in C_2$. □

§ 13 Das Subdifferential und die Richtungsableitung konvexer Funktionen

Literatur: Geiger, Kanzow, 2002, Kapitel 6.3, Alt, 2004, Kapitel 2.4–2.8

Ziel: Verallgemeinerung der Ableitung für nicht-glatte konvexe Funktionen

In diesem Abschnitt sei $C \subset \mathbb{R}^n$ konvex und $f : C \rightarrow \mathbb{R}$ eine konvexe Funktion.

Definition 13.1 (Subdifferential).

Es sei $x_0 \in C$.

(a) Ein Vektor $s \in \mathbb{R}^n$ heißt ein **Subgradient**³² von f in x_0 , wenn gilt:

$$f(x) \geq f(x_0) + s^\top (x - x_0) \quad \text{für alle } x \in C. \quad (13.1)$$

Man sagt: Die rechte Seite in (13.1) ist eine **lineare Minorante**³³ (mit Gradient s), die f in x_0 **stützt**, kurz: eine **lineare Stützfunktion**.³⁴

(b) Die Menge $\partial f(x_0)$ aller Subgradienten im Punkt x_0 heißt das **Subdifferential** von f in x_0 .³⁵

Beachte: (13.1) verallgemeinert die Ungleichung (11.2), die für konvexe *diffbare* Funktionen gilt. **Es ist hinreichend für die Bestimmung des Subdifferentials, die Ungleichung (13.1) in einer Umgebung $U(x_0) \cap C$ zu prüfen.**

Beachte: Das Subdifferential hängt vom Definitionsgebiet C der Funktion ab!³⁶

Beispiel 13.2 (Beispiele zum Subdifferential, siehe [Übung 11, Hausaufgabe 25](#)).

Wir betrachten verschiedene Normen auf \mathbb{R}^n .

³²Ein besserer Begriff wäre eigentlich **Subableitung**.

³³In [Übung 11, Aufgabe 46](#) wird die Menge aller linearen Minoranten an einem Beispiel betrachtet.

³⁴Eine andere Anschauung: $(s, -1)$ ist Normalenvektor einer Hyperebene durch den Punkt $(x_0, f(x_0))$, die den (Epi-)Graphen in diesem Punkt stützt.

³⁵**Beachte:** Das Subdifferential hängt vom Definitionsbereich der Funktion ab.

In [Übung 11, Aufgabe 48](#) wird gezeigt, dass das Subdifferential einer konvexer Funktionen ein monotoner Operator ist. (Es ist sogar unter bestimmten Voraussetzungen maximal monoton.) Subdifferentials einiger Normen werden in [Übung 11, Hausaufgabe 25](#) berechnet. In [Übung 11, Aufgabe 47](#) wird eine Kettenregel für das Subdifferential bewiesen.

³⁶Bei erweiterten konvexen Funktionen fordert man dann natürlich die Ungleichung (13.1) auf ganz \mathbb{R}^n .

§ 13. SUBDIFFERENTIAL UND RICHTUNGSABL. KONVEXER FUNKTIONEN 87

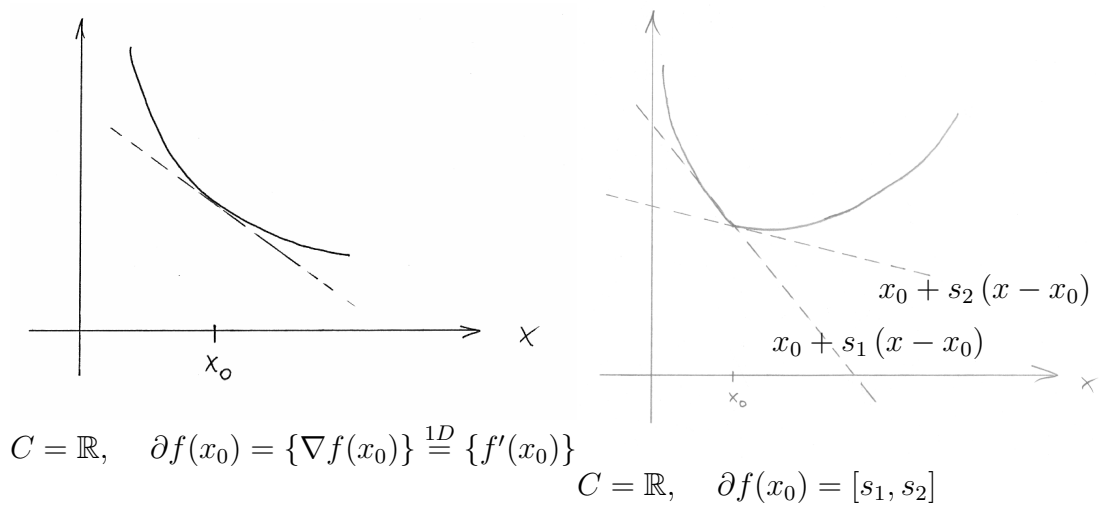
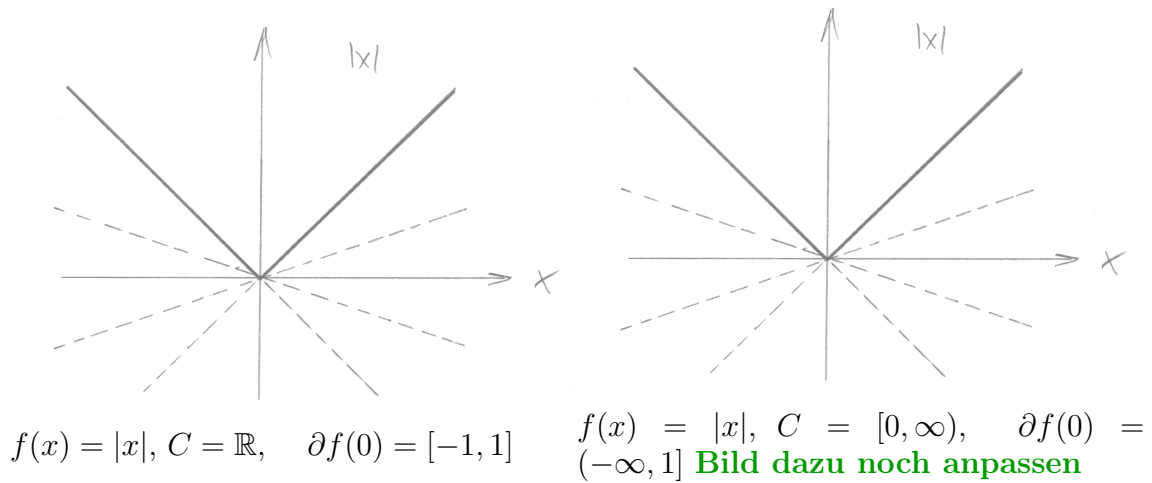


ABBILDUNG 13.1. Das Subdifferential besteht aus den *Steigungen* aller linearen Minoranten an f im Punkt x_0 .

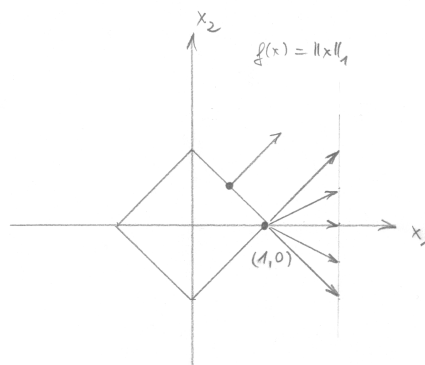


ABBILDUNG 13.2. Das Subdifferential der 1-Norm $\|x\|_1 = |x_1| + |x_2|$

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

$$s \in \partial f(x) \iff \text{für alle } i = 1, \dots, n \text{ gilt } s_i \in \begin{cases} \{-1\} & \text{falls } x_i < 0, \\ [-1, 1] & \text{falls } x_i = 0, \\ \{1\} & \text{falls } x_i > 0. \end{cases}$$

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

$$\partial f(x) = \begin{cases} \left\{ \frac{x}{\|x\|_2} \right\}, & \text{falls } x \neq 0, \\ \{s \in \mathbb{R}^n : \|s\|_2 \leq 1\}, & \text{falls } x = 0. \end{cases}$$

(c) $f(x) = \|x\|_\infty = \max_{i=1, \dots, n} |x_i|$. Dann gilt für $x \neq 0$

$$\partial f(x) = \{s \in \mathbb{R}^n : \|s\|_1 \leq 1, \text{sgn } s_i = \text{sgn } x_i \text{ für } i \text{ mit } |x_i| = \|x\|_\infty \\ \text{und } s_i = 0 \text{ für } i \text{ mit } |x_i| < \|x\|_\infty\}$$

sowie

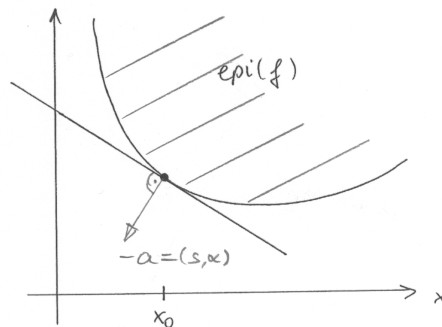
$$\partial f(0) = \{s \in \mathbb{R}^n : \|s\|_1 \leq 1\}.$$

Satz 13.3 (Existenz des Subdifferentials).

Es sei $x_0 \in \text{int}(C)$. Dann ist $\partial f(x_0) \neq \emptyset$.

Beweis: Betrachte den Epigraphen

$$\text{epi}(f) = \{(x, \alpha) \in C \times \mathbb{R} : \alpha \geq f(x)\}.$$



Nach [Satz 11.11](#) ist $\text{epi}(f) \subset \mathbb{R}^n \times \mathbb{R}$ konvex. Es ist $(x_0, f(x_0)) \notin \text{int}(\text{epi}(f))$, sonst gäbe es einen Punkt $(x_0, f(x_0) - \varepsilon) \in \text{epi}(f)$, Widerspruch zur Definition von $\text{epi}(f)$.

Nach [Satz 12.5](#) [Trennung von Punkt und konvexer Menge] existiert $a = -(s, \alpha) \in \mathbb{R}^n \times \mathbb{R}$, $(s, \alpha) \neq 0$ mit

$$\begin{pmatrix} s \\ \alpha \end{pmatrix}^\top \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} s \\ \alpha \end{pmatrix}^\top \begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix} \quad \text{für alle } (x, y) \in \text{epi}(f). \quad (*)$$

Behauptung: Es gilt $\alpha \leq 0$, denn:

Fall (i): Ist $f(x_0) = 0$, wähle $x = x_0$, $y = 1$

$$\stackrel{(*)}{\Rightarrow} s^\top x_0 + \alpha \leq s^\top x_0 + 0 \quad \Rightarrow \quad \alpha \leq 0.$$

Fall (ii): Ist $f(x_0) > 0$, wähle $x = x_0, y = 2f(x_0)$

$$\stackrel{(*)}{\Rightarrow} s^\top x_0 + 2\alpha f(x_0) \leq s^\top x_0 + \alpha f(x_0) \Rightarrow \alpha \leq 0.$$

Fall (iii): ist $f(x_0) < 0$, wähle $x = x_0, y = 0$

$$\stackrel{(*)}{\Rightarrow} s^\top x_0 + 0 \leq s^\top x_0 + \alpha f(x_0) \Rightarrow \alpha \leq 0.$$

Behauptung: Es gilt sogar $\alpha < 0$, denn:

$$x_0 \in \text{int}(C) \Rightarrow x_0 + \varepsilon s \in C \text{ für ein } \varepsilon > 0.$$

Setze in (*) $x = x_0 + \varepsilon s$ und $y = f(x)$.

$$\begin{aligned} \Rightarrow s^\top x + \alpha y &= s^\top x_0 + \varepsilon \|s\|^2 + \alpha f(x_0 + \varepsilon s) \stackrel{(*)}{\leq} s^\top x_0 + \alpha f(x_0) \\ \Rightarrow \varepsilon \|s\|^2 &\leq \alpha [f(x_0) - f(x_0 + \varepsilon s)]. \end{aligned}$$

Aus $\alpha = 0$ würde $s = 0$ folgen, Widerspruch zu $(s, \alpha) \neq 0$.

Wir können O. B. d. A. $\alpha = -1$ annehmen, dann ergibt (*) mit $y = f(x)$

$$\begin{aligned} s^\top x - f(x) &\leq s^\top x_0 - f(x_0) \text{ für alle } x \in C \\ \Rightarrow f(x) &\geq f(x_0) + s^\top(x - x_0) \text{ für alle } x \in C, \end{aligned}$$

d. h., $s \in \partial f(x_0)$. □

Beispiel 13.4 (Leeres Subdifferential).

Am Rand der Menge C kann das Subdifferential leer sein. Betrachte z. B. die konvexe (unstetige) Funktion

$$f(x) := \begin{cases} 1 & x = 0 \\ x & x > 0 \end{cases}$$

auf $C = [0, 1]$.

Ende 19. V

Frage: Gibt es einen Zusammenhang mit der Richtungsableitung? (Satz 13.12)

Definition 13.5 ((Einseitige) Richtungsableitung).

Es sei $M \subset \mathbb{R}^n, g : M \rightarrow \mathbb{R}$ (hier nicht notwendig konvex), $x_0 \in M$ und $d \in \mathbb{R}^n$. Weiter gebe es ein $\delta > 0$, sodass $x_0 + td \in M$ liegt für alle $t \in (0, \delta)$. Dann heißt der Grenzwert (sofern er in \mathbb{R} existiert)

$$\delta g(x_0; d) := \lim_{t \searrow 0} \frac{g(x_0 + td) - g(x_0)}{t} \tag{13.2}$$

die **(einseitige) Richtungsableitung** der Funktion g im Punkt x_0 in Richtung d .³⁷

Wir führen die Abkürzung

$$q(t) := \frac{f(x_0 + td) - f(x_0)}{t}, \quad t > 0. \tag{13.3}$$

für den Differenzenquotienten unserer konvexen Funktion $f : C \rightarrow \mathbb{R}$ ein.

³⁷Die Richtungsableitung wird manchmal auch mit $g'(x_0; d)$ oder $Dg(x_0; d)$ bezeichnet.

Lemma 13.6 (Monotonie des Differenzenquotienten).

Es seien $x_0 \in C$ und $x_0 + d \in C$. Dann ist $q(t)$ monoton wachsend auf $(0, 1]$.

Beweis: Es sei $0 < t_1 < t_2 \leq 1$. Es gilt $x_0 + t_1 d \in C$, $x_0 + t_2 d \in C$ und

$$x_0 + t_1 d = \frac{t_1}{t_2}(x_0 + t_2 d) + \left(1 - \frac{t_1}{t_2}\right)x_0,$$

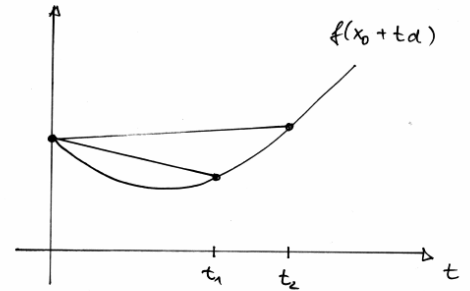
und C ist konvex, also ist $x_0 + t_1 d \in C$. Weiter gilt

$$f(x_0 + t_1 d) = f\left(\frac{t_1}{t_2}(x_0 + t_2 d) + \left(1 - \frac{t_1}{t_2}\right)x_0\right)$$

$$\leq \frac{t_1}{t_2}f(x_0 + t_2 d) + \left(1 - \frac{t_1}{t_2}\right)f(x_0)$$

$$\Rightarrow f(x_0 + t_1 d) - f(x_0) \leq \frac{t_1}{t_2}f(x_0 + t_2 d) - \frac{t_1}{t_2}f(x_0)$$

$$\Rightarrow q(t_1) \leq q(t_2).$$



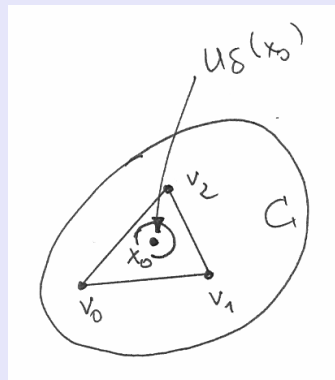
□

Um die Existenz von Richtungsableitungen sicherzustellen, zeigen wir in [Satz 13.8](#) zunächst einige weitere Eigenschaften konvexer Funktionen, das wir nun vorbereiten:

Lemma 13.7 (Zwischen eine Menge und einen inneren Punkt passt immer ein Simplex (wird benötigt in [Satz 13.8](#))).

Es sei $M \subset \mathbb{R}^n$ eine Menge und $x_0 \in \text{int}(M)$. Dann existieren affin unabhängige Vektoren $v_0, \dots, v_n \in M$, sodass deren konvexe Hülle $\Delta = \text{conv}(\{v_0, \dots, v_n\})$ folgende Eigenschaft hat:

$$x_0 \in \text{int}(\Delta) \subset M.$$



Beachte: Vektoren $v_0, \dots, v_k \in \mathbb{R}^n$ heißen **affin unabhängig**, wenn die Vektoren $v_1 - v_0, \dots, v_k - v_0$ linear unabhängig sind.

im Bild C durch M ersetzen

Beweis: Übung³⁸

□

Lemma 13.8 (Lokale Beschränktheit und lokale Lipschitz-Stetigkeit).

Es sei $x_0 \in \text{int}(C)$. Dann gibt es ein $\delta > 0$ mit $U_\delta(x_0) \subset C$, sodass gilt:

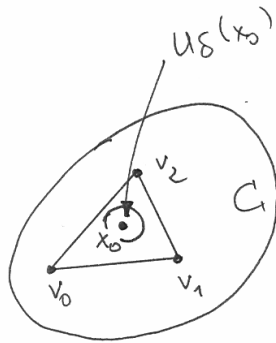
³⁸bewiesen in [Übung 10, Hausaufgabe 23](#)

- (a) Es existieren $m, M \in \mathbb{R}$ mit $f(x) \in [m, M]$ für alle $x \in U_\delta(x_0)$.
- (b) Es existiert $L(x_0) \geq 0$ mit
- $$|f(x) - f(y)| \leq L(x_0) \|x - y\| \quad \text{für alle } x, y \in U_{\delta/2}(x_0).$$

Beachte: Konvexe Funktionen sind im Inneren ihres Definitionsbereiches also **lokal beschränkt** und **lokal Lipschitz-stetig**.

Beweis: Es sei $x_0 \in \text{int}(C)$.

(a): Nach [Satz 13.7](#) existieren Vektoren $v_0, \dots, v_n \in C$ und $\delta > 0$, sodass mit $\Delta = \text{conv}(\{v_0, \dots, v_n\})$ gilt: $U_\delta(x_0) \subset \text{int}(\Delta) \subset C$ gilt.



Für $x \in U_\delta(x_0)$ gilt natürlich $x \in \Delta$, also

$$x = \sum_{i=0}^n \alpha_i v_i, \quad \text{mit } \alpha_i \geq 0 \text{ und } \sum_{i=0}^n \alpha_i = 1$$

$$\Rightarrow f(x) = f\left(\sum_{i=0}^n \alpha_i v_i\right) \stackrel{\text{Jensen}}{\leq} \sum_{i=0}^n \alpha_i f(v_i) \leq \underbrace{\max_{i=0, \dots, n} f(v_i)}_{=: M} \underbrace{\sum_{i=0}^n \alpha_i}_{=: 1} = M.$$

Die Jensensche Ungleichung wird in der Übung bewiesen.³⁹

Es sei nun $s \in \partial f(x_0)$ (Existenz: [Satz 13.3](#)). Wegen $f(x) \geq f(x_0) + s^\top(x - x_0)$ ist f auf $U_\delta(x_0)$ nach unten beschränkt durch $m := f(x_0) - \delta \|s\|$.

³⁹Beweis der Jensenschen Ungleichung in [Übung 9, Aufgabe 36](#)

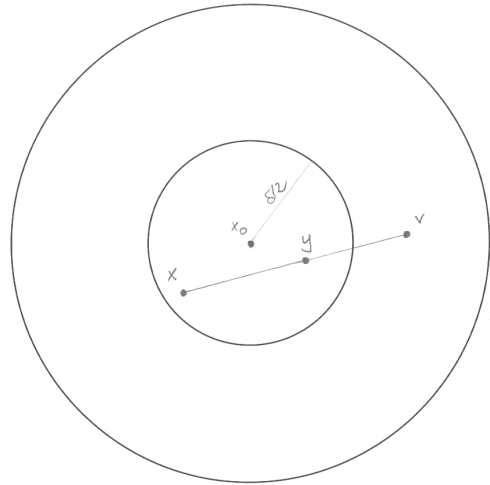
(b): Es seien $x, y \in U_{\delta/2}(x_0)$ und $x \neq y$.

Setze

$$v := y + \frac{\delta}{2} \frac{y - x}{\|y - x\|}.$$

Dann ist $v \in U_{\delta}(x_0)$, denn

$$\|v - x_0\| \leq \|v - y\| + \|y - x_0\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$



Setze $\alpha := \frac{2\|y-x\|}{\delta+2\|y-x\|} \in (0, 2/3)$.⁴⁰ Dann ist $y = \alpha v + (1 - \alpha)x$ (nachrechnen).

$$\begin{aligned} \Rightarrow f(y) - f(x) &= f(\alpha v + (1 - \alpha)x) - f(x) \\ &\leq \alpha f(v) + (1 - \alpha)f(x) - f(x) \quad \text{Konvexität von } f \\ &= \alpha(f(v) - f(x)) \\ &\leq \alpha(M - m) \quad \text{Teil (a)} \\ &\leq \frac{2(M - m)}{\delta} \|y - x\| \quad \text{denn } \alpha \leq \frac{2\|y - x\|}{\delta}. \end{aligned}$$

Durch Vertauschen von x und y erhalten wir die Behauptung mit $L(x_0) := \frac{2(M-m)}{\delta}$.
□

Wegen [Satz 13.8](#) gilt für den Differenzenquotienten aus [\(13.3\)](#) einer konvexen Funktion $f : C \rightarrow \mathbb{R}$ in einem Punkt $x_0 \in \text{int}(C)$:

$$|q(t)| = \frac{1}{t} |f(x_0 + td) - f(x_0)| \leq \frac{1}{t} L(x_0) \|x_0 + td - x_0\| = L(x_0) \|d\|$$

für alle $t \in (0, \delta')$ mit einem geeigneten $\delta' > 0$.⁴¹ $q(t)$ ist also auf $(0, \delta')$ monoton wachsend ([Satz 13.6](#)) und beschränkt. Damit existiert der Grenzwert

$$\lim_{t \searrow 0} q(t) = \inf_{t \in (0, \delta')} q(t), \quad (13.4)$$

und wir haben gezeigt:

Satz 13.9 (Existenz der Richtungsableitung).

Es sei $x_0 \in \text{int}(C)$. Dann existiert die Richtungsableitung von f im Punkt x_0 in jede Richtung $d \in \mathbb{R}^n$:

$$\delta f(x_0; d) = \lim_{t \searrow 0} q(t) = \lim_{t \searrow 0} \frac{f(x_0 + td) - f(x_0)}{t}. \quad (13.5)$$

Merke: Konvexe Funktionen sind im Inneren ihres Definitionsbereiches richtungsdiffbar.

⁴⁰Es gilt $\alpha \approx 0$, falls $y \approx x$, und $\alpha \approx 2/3$, falls $\|y - x\| \approx \delta$.

⁴¹ δ' kann als Minimum des Kugelradius ε (sodass $B_\varepsilon(x_0) \subset C$) und von $\delta/2$ aus [Satz 13.8\(b\)](#) gewählt werden.

Beispiel 13.10 (Beispiele zur Richtungsableitung).

Wir betrachten verschiedene Normen auf \mathbb{R}^n .

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

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

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

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

(c) $f(x) = \|x\|_\infty = \max_{i=1, \dots, n} |x_i|$. Dann ist

$$\delta f(x; d) = \begin{cases} \max_{\substack{i=1, \dots, n \\ |x_i| = \|x\|_\infty}} (\operatorname{sgn} x_i) d_i, & x \neq 0 \\ \|d\|_\infty, & x = 0 \end{cases}$$

Lemma 13.11 (Eigenschaften der Richtungsableitung).

Es sei $x_0 \in \operatorname{int}(C)$. Die Richtungsableitung $\delta f(x_0; \cdot)$ aus (13.5) hat folgende Eigenschaften:

(a) positiv homogen:

$$\delta f(x_0; \lambda d) = \lambda \delta f(x_0; d) \quad \text{für alle } \lambda > 0 \text{ und } d \in \mathbb{R}^n$$

(b) konvex:

$$\delta f(x_0; \lambda d_1 + (1 - \lambda) d_2) \leq \lambda \delta f(x_0; d_1) + (1 - \lambda) \delta f(x_0; d_2)$$

für alle $\lambda \in [0, 1]$ und $d_1, d_2 \in \mathbb{R}^n$

(c) damit subadditiv:

$$\delta f(x_0; d_1 + d_2) \leq \delta f(x_0; d_1) + \delta f(x_0; d_2)$$

für alle $d_1, d_2 \in \mathbb{R}^n$

(d) beschränkt durch die lokale Lipschitz-Konstante aus Satz 13.8:

$$|\delta f(x_0; d)| \leq L(x_0) \|d\| \quad \text{für alle } d \in \mathbb{R}^n;$$

insbesondere gilt $\delta f(x_0; 0) = 0$

(e) Lipschitz-stetig bzgl. d :

$$|\delta f(x_0; d_1) - \delta f(x_0; d_2)| \leq L(x_0) \|d_1 - d_2\| \quad \text{für alle } d_1, d_2 \in \mathbb{R}^n.$$

Beweis:

(a) Sei $d \in \mathbb{R}^n$, betrachte

$$\delta f(x_0; \lambda d) = \lambda \lim_{t \searrow 0} \frac{f(x_0 + t \lambda d) - f(x_0)}{\lambda t} = \lambda \delta f(x_0; d).$$

- (b) Es seien $d_1, d_2 \in \mathbb{R}^n$ und $\lambda \in [0, 1]$ sowie $t > 0$ hinreichend klein, sodass $x_0 + t d_1$, $x_0 + t d_2$ und $x_0 + t(\lambda d_1 + (1 - \lambda) d_2)$ jeweils in C liegen. Betrachte

$$\begin{aligned} & f(x_0 + t(\lambda d_1 + (1 - \lambda) d_2)) - f(x_0) \\ &= f(\lambda(x_0 + t d_1) + (1 - \lambda)(x_0 + t d_2)) - f(x_0) \\ &\leq \lambda [f(x_0 + t d_1) - f(x_0)] + (1 - \lambda)[f(x_0 + t d_2) - f(x_0)] \end{aligned}$$

Division durch $t > 0$ und Grenzübergang liefert die Behauptung.

- (c) Es gilt

$$\delta f(x_0; d_1 + d_2) \stackrel{(a)}{=} 2 \delta f(x_0; \frac{1}{2}(d_1 + d_2)) \stackrel{(b)}{\leq} \delta f(x_0; d_1) + \delta f(x_0; d_2).$$

Ende 20. V

- (d) Aufgrund von [Satz 13.8\(b\)](#) gilt für hinreichend kleines $t > 0$:

$$|f(x_0 + t d) - f(x_0)| \leq L(x_0) t \|d\|.$$

Mit Division durch $t > 0$ und Grenzübergang erhalten wir die Behauptung.

- (e) Es gilt

$$\begin{aligned} \delta f(x_0; d_1) - \delta f(x_0; d_2) &= \delta f(x_0; d_1) - \delta f(x_0; d_1 + (d_2 - d_1)) \\ &\stackrel{(c)}{\geq} \delta f(x_0; d_1) - \delta f(x_0; d_1) - \delta f(x_0; d_2 - d_1) \\ &\stackrel{(d)}{\geq} -L(x_0) \|d_1 - d_2\| \end{aligned}$$

und analog

$$\begin{aligned} \delta f(x_0; d_1) - \delta f(x_0; d_2) &= \delta f(x_0; d_2 + (d_1 - d_2)) - \delta f(x_0; d_2) \\ &\stackrel{(c)}{\leq} \delta f(x_0; d_2) + \delta f(x_0; d_1 - d_2) - \delta f(x_0; d_2) \\ &\stackrel{(d)}{\leq} L(x_0) \|d_1 - d_2\| \end{aligned}$$

□

Satz 13.12 (Zusammenhang zwischen Subdifferential und Richtungsableitung).

Es sei $x_0 \in \text{int}(C)$. Dann gilt

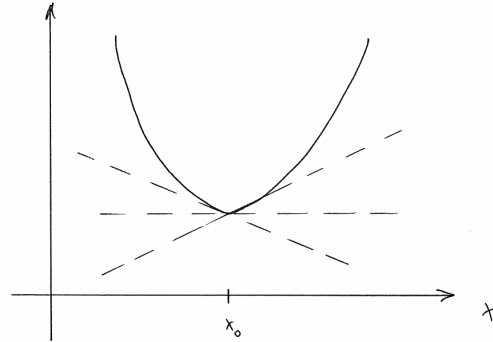
- (a)

$$\partial f(x_0) = \{s \in \mathbb{R}^n : s^\top d \leq \delta f(x_0; d) \text{ für alle } d \in \mathbb{R}^n\}.$$

- (b) Für $d \in \mathbb{R}^n$ gilt:⁴²

$$\delta f(x_0; d) = \max_{s \in \partial f(x_0)} s^\top d.$$

⁴²In [Übung 11, Aufgabe 47](#) wird die Darstellung (b) für den Beweis einer Kettenregel für das Subdifferential benutzt.



Beweis: (a) „ \subset “: Es sei $s \in \partial f(x_0)$ und $d \in \mathbb{R}^n$. Es sei $\delta > 0$ so, dass $x_0 + td \in C$ für $t \in (0, \delta)$.

$$\Rightarrow f(x_0 + td) \geq f(x_0) + ts^\top d \quad \text{für alle } t \in (0, \delta) \quad \text{Subgradientenungleichung (13.1)}$$

$$\Rightarrow \frac{f(x_0 + td) - f(x_0)}{t} \geq s^\top d \quad \text{für alle } t \in (0, \delta)$$

$$\Rightarrow \delta f(x_0; d) \geq s^\top d$$

„ \supset “: Es sei $s \in \mathbb{R}^n$ mit $s^\top d \leq \delta f(x_0; d)$ für alle $d \in \mathbb{R}^n$. Es sei $x \in C$ beliebig, setze $d := x - x_0$.

$$\Rightarrow s^\top d \leq \delta f(x_0; d) \leq \frac{f(x_0 + d) - f(x_0)}{1} \quad \text{wegen (13.4) mit } t = 1$$

$$\stackrel{t=1}{\Rightarrow} f(x) = f(x_0 + d) \geq f(x_0) + s^\top d = f(x_0) + s^\top (x - x_0),$$

d. h., $s \in \partial f(x_0)$.

(b): Es sei $d \in \mathbb{R}^n$. Nach Teil (a) gilt

$$\begin{aligned} \delta f(x_0; d) &\geq s^\top d \quad \text{für alle } s \in \partial f(x_0) \\ \Rightarrow \delta f(x_0; d) &\geq \sup_{s \in \partial f(x_0)} s^\top d. \end{aligned} \quad (*)$$

Ziel: Finde ein $s_d \in \partial f(x_0)$ mit $\delta f(x_0; d) = s_d^\top d$.

Nach Satz 13.11(c) ist $\delta f(x_0; \cdot)$ konvex auf \mathbb{R}^n . Nach Satz 13.3 existiert also ein Subgradient $s_d \in \mathbb{R}^n$ an $\delta f(x_0; \cdot)$ im Punkt d , d. h.,

$$\begin{aligned} \delta f(x_0; y) &\geq \delta f(x_0; d) + s_d^\top (y - d) \quad \text{für alle } y \in \mathbb{R}^n \quad (**) \\ y = 0 &\Rightarrow 0 \geq \delta f(x_0; d) - s_d^\top d \\ &\Rightarrow s_d^\top d \geq \delta f(x_0; d) \\ y = 2d &\Rightarrow \delta f(x_0; 2d) \stackrel{\text{Satz 13.11(a)}}{=} 2\delta f(x_0; d) \geq \delta f(x_0; d) + s_d^\top (2d - d) \\ &\Rightarrow \delta f(x_0; d) \geq s_d^\top d. \end{aligned}$$

Also gilt $s_d^\top d = \delta f(x_0; d)$. Noch zu zeigen ist $s_d \in \partial f(x_0)$. Aus (**) folgt

$$\delta f(x_0; y) \geq \delta f(x_0; d) + s_d^\top (y - d) = s_d^\top y \quad \text{für alle } y \in \mathbb{R}^n,$$

d. h., $s_d \in \partial f(x_0)$ nach Teil (a). Aus (*) folgt jetzt

$$\delta f(x_0; d) \geq \sup_{s \in \partial f(x_0)} s^\top d \geq s_d^\top d = \delta f(x_0; d),$$

und das Supremum wird durch s_d angenommen, was (b) beweist. \square

Folgerung 13.13 (Kompaktheit des Subdifferentials).

Es sei $x_0 \in \text{int}(C)$. Dann ist $\partial f(x_0) \neq \emptyset$, konvex und kompakt.

Beweis: $\partial f(x_0) \neq \emptyset$ wegen [Satz 13.3](#). Die Konvexität und Abgeschlossenheit von $\partial f(x_0)$ folgen direkt aus der [Satz 13.1](#). Für $s \in \partial f(x_0)$ gilt nach [Satz 13.12\(a\)](#)

$$s^\top d \leq \delta f(x_0; d) \leq L(x_0) \|d\| \quad \text{für alle } d \in \mathbb{R}^n$$

mit der lokalen Lipschitz-Konstante aus [Satz 13.11\(d\)](#). Mit $d = s$ folgt $\|s\| \leq L(x_0)$, d. h., $\partial f(x_0)$ ist beschränkt. \square

Satz 13.14 (Subdifferential einer konvexen *diffbaren* Funktion).

Es sei $x_0 \in \text{int}(C)$ und f diffbar in x_0 . Dann gilt

- (a) $\delta f(x_0; d) = \nabla f(x_0)^\top d$ für alle $d \in \mathbb{R}^n$
- (b) $\partial f(x_0) = \{\nabla f(x_0)\}$.

Beweis: Übung⁴³

\square

§ 14 Kegel

Literatur: [Geiger, Kanzow, 2002](#), Kapitel 2.2.1

Ziel: Beschreibung von Richtungen in der zulässigen Menge

Definition 14.1 (Kegel).

Eine Menge $K \subset \mathbb{R}^n$ heißt ein **Kegel** (*cone*), wenn $\lambda x \in K$ gilt für $x \in K$ und alle $\lambda > 0$. [Kurz: $\lambda K \subset K$ für alle $\lambda > 0$.] Der Kegel K heißt **spitz**, wenn $0 \in K$ ist, ansonsten **stumpf**.

Beachte: Mit jedem Punkt $x \in \mathbb{R}^n$ enthält ein Kegel bereits die ganze Halbgerade $\{\lambda x : \lambda > 0\}$.

Beispiel 14.2 (Kegel).

Beispiele für Kegel sind:

- (a) offene Halbgerade $\{\lambda a : \lambda > 0\}$ mit $a \in \mathbb{R}^n$, $a \neq 0$
- (b) abgeschlossene Halbgerade $\{\lambda a : \lambda \geq 0\}$ mit $a \in \mathbb{R}^n$, $a \neq 0$
- (c) offener Orthant $\{x \in \mathbb{R}^n : x > 0\}$
- (d) abgeschlossener Orthant $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x \geq 0\}$
- (e) der Lorentzkegel $K = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \|x\| \leq t\}$

⁴³Übung 11, Hausaufgabe 26