Langer Titel der Ausarbeitung

Name des Autors

17. Februar 2010

1 Optimalitätsbedingungen

Wir betrachten wieder die konvexe Optimierungsaufgabe

$$\text{Minimiere} \quad f(x) \quad \text{über } x \in C \tag{1.1}$$

mit konvexer Grundmenge $C \subset \mathbb{R}^n$. Die Zielfunktion $f : D \to \mathbb{R}$ sei konvex auf einer offenen konvexen Menge $D \supset C$.[1]

Satz 1.1 (Notwendige und hinreichende Optimalitätsbedingungen) *Sei $x^* \in C$. Die folgenden Aussagen sind äquivalent, vgl. auch [Alt(2004), Satz 3.3.11]:*

(a) x^* ist eine Lösung von (1.1).

(b) $\delta f(x^*; x - x^*) \geq 0$ für alle $x \in C$.

(c) $\delta f(x^*; d) \geq 0$ für alle $d \in K(C, x^*)$.

(d) $\delta f(x^*; d) \geq 0$ für alle $d \in \mathcal{T}_C(x^*)$.

(e) $0 \in \partial f(x^*) + \mathcal{N}_C(x^*)$.

Beachte: (e) bedeutet: Es existiert $s \in \partial f(x^*)$, sodass $-s \in \mathcal{N}_C(x^*)$ liegt. Also: C ist enthalten im Halbraum $\{z \in \mathbb{R}^n : s^\top z \geq s^\top x^*\}$ mit Normalenvektor s, siehe Abbildung 1.1.

Beweis: „(a) $\Rightarrow$ (b)“: Sei $x \in C$, also auch $x^* + t\,(x - x^*) \in C$ für $t \in [0, 1]$.

$$\begin{aligned}
& f(x^* + t\,(x - x^*)) \geq f(x^*) \quad \text{für alle } t \in [0,1] \\
\Rightarrow \quad & \frac{f(x^* + t\,(x - x^*)) - f(x^*)}{t} \geq 0 \quad \text{für alle } t \in (0,1] \\
\Rightarrow \quad & \delta f(x^*; x - x^*) \geq 0.
\end{aligned}$$

„(b) $\Rightarrow$ (a)“: Sei $x \in C$ beliebig.

$$\begin{aligned}
0 &\leq \delta f(x^*; x - x^*) \\
&\leq \frac{f(x^* + 1\,(x - x^*)) - f(x^*)}{1} \qquad \text{wegen der Monotonie des D-Quotienten} \\
&= f(x) - f(x^*),
\end{aligned}$$

[1] Dies sichert, dass ein Minimum x^* immer in $\operatorname{int}(D)$ liegt.

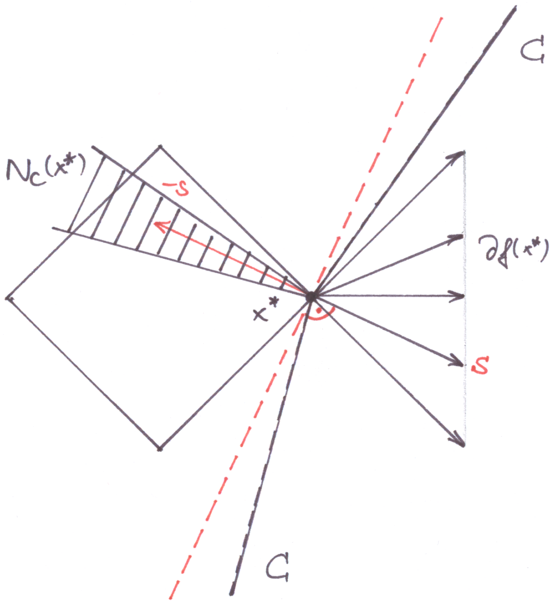

Abbildung 1.1: Illustration der Bedingung (e) für die Aufgabe min $\|x\|_1$ unter $x \in C$. Der Normalenkegel $\mathcal{N}_C(x^*)$ und das Subdifferential $\partial f(x^*)$ sind wieder verschoben in den Punkt x^* dargestellt.

d.h., x^* löst (1.1).

„(b) $\Rightarrow$ (c)“: Sei $d \in K(C, x^*)$, d.h., $d = \lambda(x - x^*)$ für ein $x \in C$ und $\lambda > 0$. Nach Voraussetzung gilt $\delta f(x^*; x - x^*) \geq 0$, und die positive Homogenität von $\delta f(x^*; \cdot)$ zeigt $\delta f(x^*; d) \geq 0$.

„(c) $\Rightarrow$ (b)“: Klar, weil $K(C, x^*)$ alle Richtungen $x - x^*$ für $x \in C$ enthält.

„(c) $\Rightarrow$ (d)“: Sei $d \in \mathcal{T}_C(x^*)$. Wegen $\mathcal{T}_C(x^*) = \overline{K(C, x^*)}$ gibt es eine Folge $\{d_k\} \subset K(C, x^*)$ mit $d_k \to d$. Aufgrund der Stetigkeit von $\delta f(x^*; \cdot)$ gilt

$$0 \leq \delta f(x^*; d_k) \to \delta f(x^*, d).$$

„(d) $\Rightarrow$ (c)“: Klar, weil $\mathcal{T}_C(x^*) = \overline{K(C, x^*)} \supset K(C, x^*)$ gilt.

„(d) $\Rightarrow$ (e)“: Wir zeigen „$\neg$ (e) $\Rightarrow$ $\neg$ (d)“ und nehmen an, dass $0 \notin \partial f(x^*) + \mathcal{N}_C(x^*) =: F$ gilt. Da $x^* \in \operatorname{int}(D)$ liegt, ist $\partial f(x^*)$ konvex und kompakt. Der Normalenkegel $\mathcal{N}_C(x^*)$ ist konvex und abgeschlossen, deshalb ist F konvex und abgeschlossen. Nach striktem Trennungssatz mit $C_1 = F$ [abgeschlossen] und $C_2 = \{0\}$ [kompakt] existiert $a \in \mathbb{R}^n$, $a \neq 0$ und $\beta \in \mathbb{R}$ mit

$$\begin{aligned} & a^\top s < \beta < 0 \quad \text{für alle } s \in F \\ \Rightarrow \quad & 0 > \beta \geq \sup_{s \in F} a^\top s = \max_{s_1 \in \partial f(x^*)} a^\top s_1 + \sup_{s_2 \in \mathcal{N}_C(x^*)} a^\top s_2. \end{aligned} \tag{$*$}$$

Behauptung: Es gilt $a \in \mathcal{T}_C(x^*)$ und $\sup_{s_2 \in \mathcal{N}_C(x^*)} a^\top s_2 = 0$.

$$0 \in \mathcal{N}_C(x^*) \quad \Rightarrow \quad \sup_{s_2 \in \mathcal{N}_C(x^*)} a^\top s_2 \geq a^\top 0 = 0$$

Wäre $a^\top s_2 > 0$ für ein $s_2 \in \mathcal{N}_C(x^*)$, so erreicht man durch Skalierung mit $\lambda > 0$ $\sup_{s_2 \in \mathcal{N}_C(x^*)} a^\top s_2 = \infty$ im Widerspruch zu ($*$). Also gilt

$$a^\top s_2 \leq 0 \quad \text{für alle } s_2 \in \mathcal{N}_C(x^*) \quad \text{und} \quad \sup_{s_2 \in \mathcal{N}_C(x^*)} a^\top s_2 = 0.$$

Das heißt, $a \in \mathcal{N}_C(x^*)^\circ = \mathcal{T}_C(x^*)^{\circ\circ} = \mathcal{T}_C(x^*)$, denn $\mathcal{T}_C(x^*)$ ist abgeschlossener und konvexer Kegel.

Aus ($*$) folgt nun

$$0 > \beta \geq \max_{s_1 \in \partial f(x^*)} a^\top s_1 + 0 = \delta f(x^*; a)$$

im Widerspruch zu Voraussetzung (d), da $a \in \mathcal{T}_C(x^*)$.

„(e) $\Rightarrow$ (d)“: Es existiert $s \in \partial f(x^*)$ mit $-s \in \mathcal{N}_C(x^*) = \mathcal{T}_C(x^*)^\circ$, d.h., $s^\top d \geq 0$ für alle $d \in \mathcal{T}_C(x^*)$.

Sei nun $d \in \mathcal{T}_C(x^*)$ beliebig.

$$\begin{aligned} \Rightarrow \quad \delta f(x^*; d) &= \max_{g \in \partial f(x^*)} g^\top d \\ &\geq s^\top d \geq 0, \qquad \text{da } s \in \partial f(x^*). \end{aligned}$$

□

2 Das Subgradientenverfahren

Literatur

[Alt(2004)] W. Alt. *Numerische Verfahren der konvexen, nichtglatten Optimierung.* Teubner, Stuttgart, 2004.