R. Herzog, G. Müller
Universität Heidelberg

Übung 14 (Lösung)

Ausgabedatum: 27. Januar 2025
Abgabedatum: 3. Februar 2025

Hausaufgabe 14.1 (Eigenschaften von $\mathcal{F}_M(x)$, $\mathcal{K}_M(x)$ und deren Zusammenhang) 1 + 1 + 1 + 3 = 6 Punkte

Es sei $M \subseteq \mathbb{R}^n$ beliebig und $x \in M$. Zeigen Sie Satz 17.7 aus dem Skript, also die folgenden Aussagen.

(a) $\mathcal{F}_M(x)$ und $\mathcal{K}_M(x)$ sind spitze Kegel.

(b) $\mathcal{F}_M(x) \subseteq \mathcal{K}_M(x)$.

(c) $M \subseteq x + \mathcal{K}_M(x)$.

(d) Es sei $C \subseteq \mathbb{R}^n$ konvex. Dann ist $\mathcal{K}_C(x)$ für jedes $x \in c$ ein spitzer konvexer Kegel, und es gilt $\mathcal{F}_C(x) = \mathcal{K}_C(x)$.

Lösung.

(a) Für $x \in M$ ist $x + t\,0 = x \in M$ für jedes beliebige $t \in \mathbb{R}$ und somit $0 \in \mathcal{F}_M(x)$. (0.5 Punkte)

Für $x \in M$ ist außerdem $\beta(x - x) = \beta\,0 = 0$ für alle $\beta \in \mathbb{R}$ und somit $0 \in \mathcal{K}_M(x)$. (0.5 Punkte)

(b) Sei $d \in \mathcal{F}_M(x)$ gegeben. Nach Definition existiert ein $\varepsilon > 0$, so dass $x + \varepsilon d \in M$. Für $\beta \coloneqq 1/\varepsilon$ ist damit $d = \beta(x + \varepsilon d - x)$ und somit $d \in \mathcal{K}_M(x)$. (1 Punkt)

(c) Sei $y \in M$ beliebig gegeben. Damit ist natürlich $y - x \in \mathcal{K}_M(x)$ und somit

$$y = x + (y - x) \in x + \mathcal{K}_M(x).$$

(1 Punkt)

(d) Dass $\mathcal{F}_C(x)$ und $\mathcal{K}_C(x)$ spitz sind, haben wir bereits in Aussage (a) gezeigt.

Die Inklusion $\mathcal{F}_C(x) \subseteq \mathcal{K}_C(x)$ folgte in Aussage (b). Für die gegenteilige Inklusion sei nun C konvex und $d \in \mathcal{K}_C(x)$, d. h., $d = \beta\,(y - x)$ für ein $\beta > 0$ und $y \in C$. Mit $\varepsilon \coloneqq 1/\beta$ folgt $x + \varepsilon\, d = x + (y - x) = y \in C$. Da C konvex ist, gilt wegen x und $x + \varepsilon\, d$ in C, dass auch $x + t\, d$ zu C gehört für alle $t \in [0, \varepsilon]$. Damit stimmen die Kegel überein. (1 Punkt)

Wir zeigen die Konvexität anhand des Radialkegels. Seien dafür $\alpha \in [0, 1]$ sowie d und $\tilde{d}$ aus $\mathcal{K}_C(x)$, also,

$$d = \beta\,(y - x) \quad \text{und} \quad \tilde{d} = \tilde{\beta}\,(\tilde{y} - x)$$

für β, $\tilde{\beta} > 0$ und y, $\tilde{y} \in C$. Dann ist auf Grund der Konvexität von C auch

$$\alpha\, d + (1-\alpha)\,\tilde{d} = \alpha\beta\,(y - x) + (1-\alpha)\tilde{\beta}\,(\tilde{y} - x) = \underbrace{\left(\alpha\beta + (1-\alpha)\tilde{\beta}\right)}_{>0} \left(\underbrace{\frac{\alpha\beta y + (1-\alpha)\tilde{\beta}\tilde{y}}{\left(\alpha\beta + (1-\alpha)\tilde{\beta}\right)}}_{\in C} - x \right) \in \mathcal{K}_C(x).$$

(2 Punkte)

Hausaufgabe 14.2 (Eigenschaften des Normalenkegels) 1.5 + 1.5 = 3 Punkte

Es sei $M \subseteq \mathbb{R}^n$ eine beliebige Menge und $x \in M$. Beweisen Sie Lemma 17.11 aus dem Skript, also die folgenden Aussagen:

(a) Der Normalenkegel $\mathcal{N}_M(x)$ ist ein konvexer abgeschlossener Kegel.

(b) Es gilt

$$\mathcal{N}_M(x) = \mathcal{K}_M(x)^\circ \coloneqq \{s \in \mathbb{R}^n \mid s^\mathsf{T} d \leqslant 0 \text{ für alle } d \in \mathcal{K}_M(x)\}.$$

Lösung.

(a) Dass es sich bei $\mathcal{N}_M(x)$ um einen Kegel handelt, ist klar, denn für $y \in M$, $s \in \mathcal{N}_M(x)$ und $\beta > 0$ ist

$$(\beta s)^\mathsf{T}(y - x) = \beta(s^\mathsf{T}(y - x)) \leq 0.$$

(0.5 Punkte)

Ebenso leicht folgt die Abgeschlossenheit, denn für $y \in M$ und eine Folge $s^{(k)}$ aus $\mathcal{N}_M(x)$ mit Grenzwert s ist

$$s^\top(y-x) = \lim_{k\to\infty} s^{(k)\top}(y-x) \le 0.$$

(0.5 Punkte)

Die Konvexität ist genauso offensichtlich, denn für $\alpha \in [0,1]$, $y \in M$ und $s, \tilde{s} \in \mathcal{N}_M(x)$ ist

$$(\alpha\, s + (1-\alpha)\,\tilde{s})^\top\,(y-x) = \alpha\, s^\top(y-x) + (1-\alpha)\,\tilde{s}^\top(y-x) \le 0.$$

(0.5 Punkte)

(b) Dass $\mathcal{K}_M(x)^\circ \subseteq \mathcal{N}_M(x)$ ist, ist ziemlich offensichtlich, denn für $y \in M$ ist natürlich immer $y - x \in \mathcal{K}_M(x)$. Für $s \in \mathcal{K}_C(x)^\circ$ ist also immer

$$s^\top(y-x) \le 0$$

und damit $s \in \mathcal{N}_C(x)$.

Dass außerdem $\mathcal{N}_M(x) \subseteq \mathcal{K}_M(x)^\circ$ sieht man daran, dass für $y \in M$, $\beta > 0$ und $s \in \mathcal{N}_C(x)$ immer

$$s^\top(\beta(y-x)) = \beta s^\top((y-x)) \le 0.$$

(1.5 Punkte)

Hausaufgabe 14.3 (Standortoptimierung) 1 + 6 + 2 + 2 = 11 Punkte

Der Standort einer Rettungswache soll geplant werden. Sie soll m Ortschaften versorgen. Maß für die Güte eines Standorts ist die gewichtete Summe der Abstände des Standorts zu den Ortschaften (je kleiner, desto besser). Dabei sind die Gewichte der Ortschaften proportional zu ihrer jeweiligen Bevölkerungszahl.

Das mathematische Modell hat die folgende Form: Es seien $a_1, \dots, a_m$ paarweise verschiedene Punkte des $\mathbb{R}^2$, welche die Lage der Ortschaften darstellen, positive Gewichte $w_1, \dots, w_m$ und $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x) \coloneqq \sum_{i=1}^m w_i \, \|x - a_i\|_2$ gegeben. Jeder Minimierer von f ist ein optimaler Standort.

(a) Geben Sie ein Beispiel an, das zeigt, dass die Aufgabe im Allgemeinen nicht eindeutig lösbar ist.

(b) Geben Sie hinreichende und notwendige Optimalitätsbedingungen für Minimierer von f an und zeigen Sie, dass ein Minimierer existiert.

(c) Der Standort darf nun ausschließlich im Kreisscheibensektor $K \coloneqq \{x \in \mathbb{R}^2 \mid x \geqslant 0,\ \|x\|_2 \leqslant 1\}$ liegen. Passen Sie die Optimalitätsbedingungen aus Teilaufgabe (b) an.

Hinweis: Sie dürfen den auftretenden Normalenkegel ohne technischen Nachweis angeben.

(d) Die Optimierungsaufgaben in Teilaufgaben (b) und (c) lassen sich mit einer mechanischen Konstruktion lösen. Geben Sie an, wie diese aussehen könnte.

Hinweis: Interpretieren Sie die Bedingungen aus Teilaufgaben (b) und (c) als Kräftegleichgewicht.

Lösung.

(a) Für zwei Punkte $a_1 \neq a_2$ und $w_1 = w_2 = 1$ ist jeder Punkt auf der Verbindungsgeraden von a_1 und a_2 optimal mit dem Optimalwert $\|a_1 - a_2\|_2$. (1 Punkt)

(b) Da $f(x)$ konvex ist, folgt aus Satz 18.1, dass x^* genau dann ein globaler Minimierer von f über R^2 ist, wenn $0 \in \partial f(x^*)$. (1 Punkt)

Für die Verkettung der Norm mit der affinen Verschiebung gilt die Kettenregel auch für das Subdifferential (in jedem Punkt außer der Null ist die Norm differenzierbar und in der Null ist das direkt sichtbar). Die Darstellung des Subdifferentials der 2-Norm kennen wir aus Beispiel 16.5. Damit erhalten wir, dass

$$\partial\|x - a_i\|_2 = \begin{cases} \left\{\frac{x-a_i}{\|x-a_i\|_2}\right\}, & x \neq a_i \\ \left\{s \in \mathbb{R}^2 \,\middle|\, \|s\|_2 \leq 1\right\}, & x = a_i. \end{cases}$$

(1 Punkt)

Da der eigentliche Definitionsbereich aller summierten Funktionen (der ganze $\mathbb{R}^n$) übereinstimmt, gilt auch die Summenregel für Subdifferentiale (Satz 16.9). Ein Punkt $x^* \in \mathbb{R}^n$ ist also genau dann Minimierer, wenn

$$0 \in \sum_{i=1}^m w_i \partial\|x - a_i\|_2$$

und damit genau dann, wenn

$$0 \in \begin{cases} \left\{\sum_{i=1}^m w_i \frac{x^*-a_i}{\|x^*-a_i\|}\right\} & \text{wenn } x^* \neq a_i \text{ für alle } i \\ \left\{\sum_{\substack{i=1 \\ i\neq j}}^m w_i \frac{x^*-a_i}{\|x^*-a_i\|}\right\} + w_j\{s \in \mathbb{R}^2 \mid \|s\| \leq 1\} & x^* = a_j. \end{cases}$$

Das ist äquivalent zu

$$
\begin{aligned}
\sum_{i=1}^m w_i \frac{x^*-a_i}{\|x^*-a_i\|} &= 0 \qquad \text{wenn } x^* \neq a_i \text{ für alle } i \text{, bzw.}\\
\left\| \sum_{\substack{i=1\\ i\neq j}}^m w_i \frac{x^*-a_i}{\|x^*-a_i\|} \right\| &\leq w_j \qquad \text{wenn } x^* = a_j.
\end{aligned}
$$

(2 Punkte)

Weiterhin ist die Funktion f stetig. Wir können die Existenz einer Lösung also aus dem Satz von Weierstraß folgern, wenn wir zeigen können, dass wir uns bei der Suche auf eine kompakte Menge einschränken können. Es gibt zwei naheliegende Wege das zu tun.

Ansatz 1 nutzt, dass wir uns von einem der Punkte a_i, o. B. d. A. sei das a_1, so weit entfernen können, dass allein der gewichtete Abstandsanteil zu a_1 den Funktionswert $f(a_1)$ übersteigt. Genauer ist für Punkte x außerhalb der Menge $\overline{B_{\frac{1}{w_1} f(a_1)}(a_1)}$

$$
f(x) = \sum_{i=1}^n w_i \|x-a_i\|_2 = w_1\|x-a_1\| + \sum_{i=2}^n w_i \|x-a_i\|_2 \geq f(a_1) + \sum_{i=2}^n w_i\|x-a_i\|_2 \geq f(a_1),
$$

daher muss der Minimierer in $\overline{B_{\frac{1}{w_1} f(a_1)}(a_1)}$ liegen, auf dem uns der Satz von Weierstraß die Existenz eines solchen liefert.

Ansatz 2 zeigt das interessantere Resultat, dass alle Minimierer in der kompakten Menge $\operatorname{conv}(a_1, \ldots, a_m)$ liegen müssen. **Beachte:** (Das gilt allgemein natürlich nur für positive Gewichte w_i. Für negativ gewichtete Probleme kann es sein, dass ein Punkt so "abstoßend" wirkt, dass ein Optimum außerhalb der konvexen Hülle liegt.)Genauer zeigen wir für eine beliebige abgeschlossene, nichtleere, konvexe Menge $C \subseteq \mathbb{R}^n$ und einen beliebigen Punkt $x \in \mathbb{R}^n$, dass

$$
\|x-c\|_2 \geq \|\operatorname{proj}_C(x) - c\| \quad \text{für alle } c \in C
$$

ist, wobei Gleichheit nur dann gilt, wenn $x = \operatorname{proj}_C(x) \in C$ war.

Es seien dafür $x \in \mathbb{R}^n$ und $c \in C$ beliebig gegeben. Dann ist auf Grund der Variationsungleichung für die Projektion aus Satz 15.3 auch

$$
\begin{aligned}
\|x-c\|^2 &= (x - \operatorname{proj}_C(x) + \operatorname{proj}_C(x) - c)^\top (x - \operatorname{proj}_C(x) + \operatorname{proj}_C(x) - c)\\
&= \underbrace{\|x-\operatorname{proj}_C(x)\|^2}_{\geq 0,\, >0 \text{ für } \operatorname{proj}_C(x)\neq c} + \|\operatorname{proj}_C(x) - c\|^2 + \underbrace{(x-\operatorname{proj}_C(x))^\top(\operatorname{proj}_C(x) - c)}_{\geq 0}\\
&\geq \|\operatorname{proj}_C(x) - c\|^2.
\end{aligned}
$$

Wir sehen also, dass der Abstand $\|x-a_i\|_2$ für jedes $x \in \mathbb{R}^n \setminus \operatorname{conv}(a_1, \dots, a_m)$ und a_i, $i = 1, \dots, m$ echt kleiner als der Abstand $\|\operatorname{proj}_{\operatorname{conv}(a_1,\dots,a_m)}(x) - a_i\|$ ist und damit auf Grund der positiven Gewichte auch $f(\operatorname{proj}_{\operatorname{conv}(a_1,\dots,a_m)}(x)) < f(x)$, womit kein Minimierer außerhalb der konvexen Hülle liegen kann. Da die konvexe Hülle nichtleer und kompakt ist, können wir den Satz von Weierstraß also anwenden, um die Existenz eines Minimierers in der konvexen Hülle zu erhalten. (2 Punkte)

(c) Wir fügen die zusätzlichen Nebenbedingungen zu unserem Problem in Form einer Indikatorfunktion hinzu und betrachten das Problem

$$\text{Minimiere } \underbrace{\sum_{i=1}^{n} w_i \|x - a_i\|}_{f(x)} + \delta_K(x) \text{ über } x \in \mathbb{R}^n$$

Nach den Optimalitätsbedingungen in Satz 18.1 und der Darstellung des Subdifferentials der Indikatorfunktion als Normalenkegel aus Lemma 17.12 ist x^* genau dann ein Minimierer, wenn

$$0 \in \partial f(x^*) + N_K(x^*).$$

(1 Punkt)

Das Subdifferential $\partial f(x^*)$ haben wir schon in Teilaufgabe (b) bestimmt, und der Normalenkegel $N_K(x^*)$ an den Kreisscheibensektor $\{x \in \mathbb{R}^2 \mid x \geqslant 0, \ \|x\|_2 \leqslant 1\}$ hängt von x^* ab. Dabei ist:

$$N_K(x) = \begin{cases} 0, & x \in \operatorname{int}(K) \\ \left\{\beta \frac{x}{\|x\|_2} \,\middle|\, \beta \geqslant 0\right\}, & \|x\|_2 = 1, x_1, x_2 \neq 0 \\ \{x \in \mathbb{R}^n \mid x_1 \geq 0, \ x_2 \leq 0\}, & x = (1, 0) \\ \{x \in \mathbb{R}^n \mid x_1 \leq 0, \ x_2 \geq 0\}, & x = (0, 1) \\ \{x \in \mathbb{R}^n \mid x_1 \leq 0, \ x_2 \leq 0\}, & x = (0, 0) \\ \{\beta(0, -1) \mid \beta \geq 0\}, & x_1 \in (0, 1), \ x_2 = 0 \\ \{\beta(-1, 0) \mid \beta \geq 0\}, & x_2 \in (0, 1), \ x_1 = 0 \end{cases}$$

(1 Punkt)

(d) In eine flache Platte bohrt man an den Stellen a_i Löcher. Von unten führt man Schnüre durch die Löcher und knotet alle Enden auf der Oberseite zusammen. An die freien Enden unterhalb der Platte knotet man Gewichte w_i. Die Kräfte, die auf den Knotenpunkt an der Stelle x auf der Oberfläche in der Ebene wirken ergeben sich genau als die Gewichtskräfte der Gewichte in Richtung des Lochs, durch das die Schnur des Gewichts gefädelt ist, gegeben durch

$$g\, w_i \frac{x - a_i}{\|x - a_i\|_2}.$$

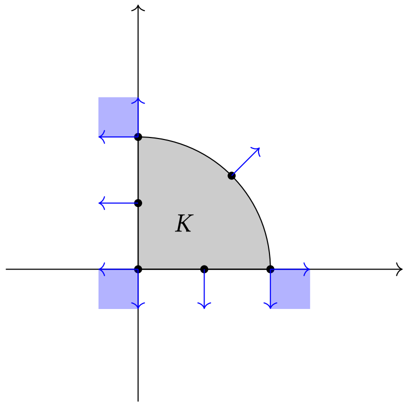

Abbildung 0.1: Kreisscheibensektor mit Normalenkegeln (blau) an verschiedenen Punkten

Die Optimalitätsbedingungen an glatten Punkten entsprechen also genau einer Kräftegleichheit im Knotenpunkt (der sich dann an der Lösung befindet). An den nichtglatten Punkten (den Bohrlöchern) entspricht die Optimalitätsbedingung genau der Bedingung, dass das Gewicht w_j in dem Loch a_j schwer genug ist, alle weiteren Kräfte dort auszugleichen. Hier sieht man schön, wie die Richtung, in die gezogen wird, degeneriert, wenn der Knoten auf dem Bohrloch angekommen ist. Ab da zählen nur noch Kraftbeträge, keine Richtungen mehr. Im Fall von Teilaufgabe (c) muss noch zusätzlich dafür gesorgt werden, dass der Kreisscheibensektor nicht verlassen wird (z. B. im Knoten noch eine Kugel einknoten und den Kreisscheibensektor mit Stangenbegrenzen, unter denen die Schnüre frei laufen können, nicht aber die Kugel drunter durch passt). Diese Ränder geben dann eine weitere Kraftkomponente an den Rändern dazu. Diese muss aus einer Konvexkombination der Normalenkräfte an den jeweils aktiven Nebenbedingungen entsprechen (die Ränder können nur nach innen drücken, nicht nach innen ziehen). Diese Konstruktion wird auch Varignon'scher Apparat genannt (2 Punkte)

Zusatzaufgabe 14.4 (Eindeutigkeit der Richtung des steilsten Abstiegs) 3 Bonuspunkte

Es sei $f\colon \mathbb{R}^n \to \mathbb{R}$ konvex und $x_0 \in \mathbb{R}^n$. Beweisen Sie Lemma 19.1 aus dem Skript, also die folgende Aussage:

Falls ein $d \in \mathbb{R}^n$ existiert, sodass $f'(x_0; d) < 0$ ist, dann ist die Lösung von

$$\begin{aligned} \text{Minimiere} \quad & f'(x_0; d) \quad \text{über } d \in \mathbb{R}^n \\ \text{unter} \quad & \|d\| \leqslant 1 \end{aligned} \qquad \text{(Gleichung (19.3))}$$

eindeutig.

Lösung.

Nehmen wir an, wir hätten zwei verschiedene globale Minimierer d_1 und d_2 von Gleichung (19.3). Aufgrund der Voraussetzung, dass es ein $d \in \mathbb{R}^n$ mit $f'(x_0; d) < 0$ gibt, gilt natürlich dann auch $f'(x_0; d_1) = f'(x_0; d_2) < 0$.

Wegen der positiven Homogenität der Richtungsableitung (Satz 16.16) wissen wir, dass $\|d_1\| = \|d_2\| = 1$ (sonst wäre z. B. $f'(x_0; d_1/\|d_1\|) < f'(x_0; d_1)$). (1 Punkt)

Andererseits folgt aus dem Hauptsatz der konvexen Optimierung 14.2, dass auch alle Konvexkombinationen $\alpha\, d_1 + (1-\alpha)\, d_2$ mit $\alpha \in [0,1]$ globale Minimierer von (19.3) sind. Jedoch gilt beispielsweise für $\overline{d} \coloneqq (d_1 + d_2)/2$:

$$\|\overline{d}\|^2 = \left\|\frac{d_1 + d_2}{2}\right\|^2 = \frac{1}{4}\|d_1 + d_2\|^2 = \frac{1}{2}\|d_1\|^2 + \frac{1}{2}\|d_2\|^2 - \frac{1}{4}\|d_1 - d_2\|^2 = 1 - \frac{1}{4}\|d_1 - d_2\|^2 < 1.$$

Hier wurde die Parallelogramm-Gleichung

$$2\,\|x\|^2 + 2\,\|y\|^2 = \|x+y\|^2 + \|x-y\|^2,$$

die in jedem Innenproduktraum gilt, verwendet. Es gilt also

$$f'(x_0; d_1) = f'(x_0; d_2) = f'(x_0; \overline{d}) < 0,$$

aber $\|\overline{d}\| < 1$. Wie oben bereits erwähnt, ergibt das aber mit der positiven Homogenität der Richtungsableitung einen Widerspruch zur Optimalität der Richtungen. (2 Punkte)

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