R. Herzog, G. Müller
Universität Heidelberg

Übung 10 (Lösung)

Ausgabedatum: 16. Dezember 2024
Abgabedatum: 7. Januar 2025

Hausaufgabe 10.1 (Schranke an den Abstand zum globalen Minimierer) 4 + 1 = 5 Punkte

Es sei $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ eine eigentliche, μ-stark konvexe Funktion.

(a) Zeigen Sie, dass falls x^* ein globaler Minimierer von f über $\mathbb{R}^n$ ist, dann ist

$$\|x - x^*\|^2 \le \frac{2}{\mu} |f(x) - f(x^*)|.$$

(b) Geben Sie ein Gegenbeispiel an, das zeigt, dass Punkt (a) i. A. nicht gilt, wenn x^* kein globaler Minimierer von f ist.

Lösung.

(a) Für μ-stark konvexes f gilt

$$f(\alpha x + (1-\alpha) x^*) + \frac{\mu}{2} \alpha (1-\alpha) \|x - x^*\|^2 \leqslant \alpha f(x) + (1-\alpha) f(x^*) \tag{$*$}$$

für alle $x \in \mathbb{R}^n$ und $\alpha \in [0, 1]$. Da x^* ein globaler Minimierer ist, gilt außerdem

$$f(x^*) \le f(\alpha x + (1-\alpha) x^*) \quad \forall x \in \mathbb{R}^n \text{ und } \alpha \in [0, 1].$$

Damit erhalten wir aus Gleichung ($*$), dass

$$\frac{\mu}{2} \alpha (1-\alpha) \|x - x^*\|^2 \le \alpha f(x) - \alpha f(x^*) \quad \forall x \in \mathbb{R}^n \text{ und } \alpha \in [0, 1].$$

(2 Punkte)

Für $\alpha \neq 0$ können wir α kürzen und erhalten daher

$$\frac{\mu}{2}(1-\alpha)\|x-x^*\|^2 \leq f(x) - f(x^*) \quad \forall x \in C \text{ und } \alpha \in (0,1].$$

Für $\alpha \longrightarrow 0$ folgt nun

$$\frac{\mu}{2}\|x-x^*\|^2 \leq f(x) - f(x^*) \quad \forall x \in C,$$

und wegen $f(x) - f(x^*) \geq 0$ gilt damit auch die gesuchte Abschätzung

$$\|x-x^*\|^2 \leq \frac{2}{\mu}|f(x) - f(x^*)|.$$

(2 Punkte)

(b) Wir betrachten die Funktion $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ mit $f(x) = x^2$, welche 2-stark konvex auf ganz $\mathbb{R}$ ist. Wählen wir $x_1 = a$ und $x_2 = -a$ mit $a > 0$, dann gilt $|f(x_1) - f(x_2)| = 0$ und $\|x_1 - x_2\|^2 > 0$. Damit kann es keine Konstante μ geben, sodass die obige Ungleichung gilt. (1 Punkt)

Hausaufgabe 10.2 (Beispiele von Projektionsaufgaben) 5 + 1 = 6 Punkte

(a) Bestimmen Sie eine explizite Darstellung der orthogonalen Projektion auf einen affinen Unterraum $A \subseteq \mathbb{R}^n$.

(b) Zeigen Sie, dass die Konvexität der Zielmenge $C \subseteq \mathbb{R}^n$ in der Projektionsaufgabe (Beispiel 15.1) entscheidend für die Wohldefiniertheit der Aufgabe ist. Geben Sie dazu eine nichtleere, kompakte und nichtkonvexe Zielmenge C sowie einen Punkt $p \in \mathbb{R}^n$ an, bei dem die gesamte Menge C Lösung der Projektionsaufgabe (15.1) für den Punkt p ist.

Lösung.

(a) Die Optimierungsaufgabe zur Bestimmung der Projektion eines Punkts $p \in \mathbb{R}^n$ auf die Menge A ist gegeben durch

$$\text{Minimiere} \quad \|x-p\| + \delta_A(x) \quad \text{über } x \in \mathbb{R}^n$$

bzw. die nach Beispiel 15.1 äquivalente Form

$$\text{Minimiere} \quad \frac{1}{2}\|x-p\|^2 + \delta_A(x) \quad \text{über } x \in \mathbb{R}^n.$$

Da A ein affiner Unterraum ist, hat er die Struktur $A = x_0 + U$ für einen linearen Unterraum U. Der Unterraum U hat eine Dimension $k \le n$ und eine entsprechende Basis, es existiert also eine Vollrangmatrix $B \in \mathbb{R}^{n\times k}$, so dass $A = \{x_0 + By \mid y \in \mathbb{R}^k\}$.

Löst man die Nebenbedingung aus dem Optimierungsproblem mit dieser Beziehung explizit auf, dann erhält man die Aufgabe

$$\text{Minimiere} \quad f(y) \coloneqq \frac{1}{2}\|By + x_0 - p\|^2 \quad \text{über } y \in \mathbb{R}^k.$$

(1 Punkt)

Untersucht man die Stationaritätsbedingungen erster Ordnung aus Satz 3.1 für die Funktion

$$\begin{aligned} f(y) &= \frac{1}{2}\|By + x_0 - p\|^2 \\ &= \frac{1}{2}(By + x_0 - p)^\top (By + x_0 - p) \\ &= \frac{1}{2} y^\top B^\top B y + (B^\top(x_0 - p))^\top y + \frac{1}{2}(x_0 - p)^\top (x_0 - p), \end{aligned}$$

(1 Punkt)

dann ergibt sich sofort die Bedingung

$$f'(y) = B^\top B y + B^\top (x_0 - p) = 0.$$

(1 Punkt)

Da B vollen Spaltenrang hat, ist

$$y^\top B^\top B y = \|By\|^2 > 0 \quad \text{für alle } y \in \mathbb{R}^k \setminus \{0\},$$

und $B^\top B$ ist damit positiv definit und somit insbesondere invertierbar, was uns den einzigen stationären Punkt $y^* \coloneqq (B^\top B)^{-1} B^\top (p - x_0)$ mit dem dazugehörigen Punkt $x^* \coloneqq x_0 + By^*$ liefert. (1 Punkt)

Dass x^* tatsächlich die Lösung des Problems ist, sieht man direkt daran, dass nach Beispiel 15.1 eine eindeutige Lösung der Aufgabe existiert (die damit auch ein stationärer Punkt ist). Alternativ sieht man, dass die hinreichenden Bedingungen zweiter Ordnung für $f''(y) = B^\top B$, an y^*, wie oben gezeigt, erfüllt sind. (1 Punkt)

Ein alternativer Beweis nutzt ausschließlich die Variationsungleichung Gleichung (15.4) aus Satz 15.3, denn dadurch, dass A ein affiner Unterraum ist, wird die Variationsungleichung

$$(x^* - p)^\top (x - x^*) \ge 0 \quad \text{für alle } x \in A,$$

zu einer Variationsgleichung der Form

$$(x^* - p)^\top z = 0 \quad \text{für alle } z \in U,$$

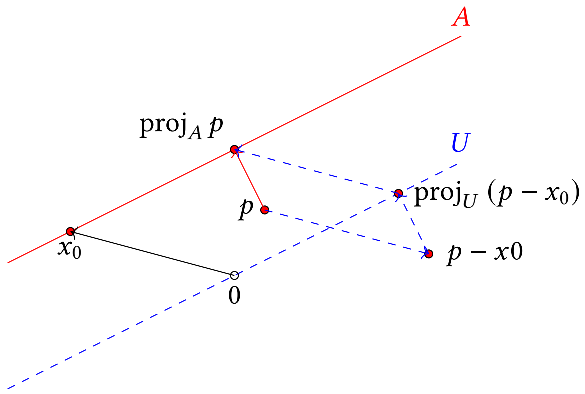

Abbildung 0.1: Beziehung von Projektion eines Punkts p auf einen affinen Unterraum A und Projektion eines um einen beliebigen Aufpunkt $x_0 \in A$ verschobenen Punkts auf den Richtungsraum ($\operatorname{proj}_A p = x_0 + \operatorname{proj}_U (p - x_0)$)

von der wir wissen, dass sie eine eindeutige Lösung besitzt. Ersetzt man nun wieder die Variable x^* durch die Darstellung $x_0 + By^*$ für ein $y^* \in \mathbb{R}^k$ und z durch By, dann ergibt sich die Variationsgleichung

$$(B^\top B y^* + B^\top (x_0 - p))^\top y = 0 \quad \text{für alle } y \in \mathbb{R}^k.$$

Damit folgt sofort, dass

$$B^\top B y^* = B^\top (p - x_0)$$

und mit dem gleichen Invertierbarkeitsargument wie oben das Ergebnis

$$y^* = (B^\top B)^{-1} B^\top (p - x_0) \quad \text{bzw.} \quad x^* = x_0 + B(B^\top B)^{-1} B^\top (p - x_0).$$

Beachte: Setzt man $x_0 = 0$, dann erhält man genau die Projektion von p auf den Unterraum U (weil dann $A = U$ ist), die schon aus dem Plenum bekannt ist.

Anhand der Struktur

$$\operatorname{proj}_A p = x^* = x_0 + B(B^\top B)^{-1} B^\top (p - x_0) = x_0 + \operatorname{proj}_U (p - x_0)$$

kann man dann direkt ablesen, dass die Projektion von p auf A mit der um x_0 verschobenen Projektion von $p - x_0$ auf U übereinstimmt.

(b) Für jede Sphäre um den Mittelpunkt $\{x \in \mathbb{R}^n \mid \|x\| = r\}$ für $r > 0$ ist jeder Punkt der Menge Lösung der Projektionsaufgabe für den Ursprung. (1 Punkt)

Beachte: Es gibt natürlich auch konvexe Mengen, für die diese Eigenschaft gilt. Das sind genau die Punktmengen.

Hausaufgabe 10.3 (Dimension eines affinen Unterraums) 3 + 4 = 7 Punkte

Beweisen Sie Lemma 15.7 aus dem Skript, also die folgenden Aussagen für einen affinen Unterraum $A \subset \mathbb{R}^n$.

(a) A besitzt genau dann eine affine Basis $\{x_0, x_1, \ldots, x_k\}$ aus $k+1$ Elementen mit $k \geq 0$, wenn $\dim A = k$ ist.

(b) Ist $\{x_0, x_1, \ldots, x_k\}$ eine affine Basis von A, dann lässt sich jedes Element von A auf eindeutige Art und Weise aus $\{x_0, x_1, \ldots, x_k\}$ affinkombinieren. Genauer hat jedes $x \in A$ die Darstellung

$$x = \sum_{i=0}^{k} \alpha_i \, x_i$$

mit Koeffizienten $\alpha = (\alpha_0, \ldots, \alpha_k)^\top$, die sich aus der eindeutigen Lösung des linearen Gleichungssystems

$$\underbrace{\begin{bmatrix} 1 & \cdots & 1 \\ | & & | \\ x_0 & \cdots & x_k \\ | & & | \end{bmatrix}}_{=:B} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_k \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ | \\ x \\ | \end{pmatrix}}_{=:b} \tag{$*$}$$

ergeben. Die Matrix $B \in \mathbb{R}^{(n+1)\times(k+1)}$ hat Rang $k+1$. Daher ist $B^\top B$ regulär, und ($*$) kann äquivalent als

$$B^\top B \, \alpha = B^\top b$$

geschrieben werden.

Lösung.

(a) $\Rightarrow$: Es sei $\{x_0, x_1, \ldots, x_k\}$ eine affine Basis aus A von $k+1$ Elementen. Per Definition sind die Elemente $\{x_1 - x_0, \ldots, x_k - x_0\}$ Elemente von dem zu A gehörigen Unterraum $U = \{x - x_0 \mid x \in A\}$. Sie sind linear unabhängig und maximal und damit eine Basis von U, daher $\dim A = \dim U = k$. (1.5 Punkte)

$\Leftarrow$: Ist $\dim A = \dim U = k$ für den zu A gehörigen Unterraum U, so dass $A = U + x$ für jedes beliebige $x \in A$, dann existieren für alle nichttrivialen Fälle ($k \geqslant 1$) Basisvektoren $\{u_1, \ldots, u_k\}$ zu U. Wir erhalten aus der ursprünglichen Basis für jedes beliebige Element $x_0 \in A$ mit $\{x_0, u_1 + x_0, \ldots, u_k + x_0\}$ eine affine Basis. (1.5 Punkte)

(b) Da $x \in A$ und $\{x_0, x_1, \dots, x_k\}$ eine affine Basis ist, existieren eindeutige Koeffizienten $\alpha = (\alpha_1, \dots, \alpha_k)^\top$, so dass

$$x - x_0 = \sum_{i=1}^{k} \alpha_i (x_i - x_0) = \sum_{i=1}^{k} \alpha_i x_i - \sum_{i=1}^{k} \alpha_i x_0.$$

Nun setzen wir

$$\alpha_0 \coloneqq 1 - \sum_{i=1}^{k} \alpha_i$$

so dass

$$x = \sum_{i=0}^{k} \alpha_i x_i \quad \text{und} \quad \sum_{i=0}^{k} \alpha_i = 1.$$

Wir wissen also, dass das System $(*)$ für $x \in A$ eindeutig lösbar ist. (1 Punkt)

Da $\{x_0, x_1, \dots, x_k\}$ eine affine Basis ist, hat die Matrix $B \in \mathbb{R}^{(n+1)\times(k+1)}$ außerdem den Rang $k+1 \le n+1$, was man sofort sieht, wenn man die erste Spalte von den anderen Spalten abzieht. Damit ist $B^\top B$ regulär. (1 Punkt)

Dass die Lösung des Systems $(*)$ auch das System

$$B^\top B\, \alpha = B^\top b$$

löst ist offensichtlich, wenn man mit $B^\top$ durchmultipliziert. Sei nun irgendein α gegeben, dann ist (weil $(*)$ lösbar ist) $b \in \operatorname{Bild}(B)$ und damit auch

$$B\alpha - b \in \operatorname{Bild}(B)$$

und da $\operatorname{Bild}(B) = \ker(B^\top)^\perp$, ist $B^\top$ injektiv auf $\operatorname{Bild}(B)$, also ist eine Lösung des Systems

$$B^\top B\, \alpha = B^\top b$$

auch eine Lösung des Systems $(*)$. (2 Punkte)

Zusatzaufgabe 10.4 (Hauptsatz der linearen Optimierung aus Sicht der konvexen Optimierung) 5 Bonuspunkte

(a) Es seien $C \subset \mathbb{R}^n$ eine nichtleere, kompakte, konvexe Menge und $f\colon C \to \mathbb{R}$ eine stetige, konvexe Funktion. Zeigen Sie, dass die Funktion f ihr *Supremum* in einem Extremalpunkt von C annimmt.

(b) Nutzen Sie Aussage (a), um Satz 6.21 Aussage *(iii)* für beschränkte Polyeder zu beweisen, also die folgende Aussage:

Es sei P ein beschränktes Polyeder. Besitzt das Problem

$$\text{Minimiere} \quad c^\top x \quad \text{über } x \in P$$

eine Lösung, so ist auch ein Extremalpunkt (einer Ecke) von P eine Lösung.

Lösung.

(a) Der Satz von Weierstraß liefert sofort, dass die Funktion f ihr Supremum als Maximum annimmt. Es sei $x \in C$ ein dazugehöriger Maximierer. Auf Grund von **?? ??** wissen wir, dass Extremalpunkte $x_1, \dots, x_m$ von C und Konvexkoeffizienten $\alpha_1, \dots, \alpha_m \geqslant 0$, $\sum_{k=1}^m \alpha_k = 1$ existieren, sodass

$$x = \sum_{k=1}^m \alpha_k x_k,$$

(1 Punkt)

und auf Grund der Jensenschen Ungleichung (Hausaufgabe 9.5) ist

$$\sup_{y \in C} f(y) = f(x) = f\Big(\sum_{k=1}^m \alpha_k\, x_k\Big) \leqslant \sum_{k=1}^m \alpha_k\, f(x_k) \leqslant \sup_{y \in C} f(y)$$

und somit muss in das Supremum ebenfalls in jedem x_k mit $\alpha_k > 0$ und damit in mindestens einem Extremalpunkt angenommen werden. (2 Punkte)

(b) Beschränkte Polyeder sind konvex. Besitzt das Problem eine Lösung, so ist das Polyeder nichtleer und damit nichtleer, kompakt und konvex. Nach Aussage (a) nimmt die konvexe Funktion $x \mapsto -c^\top x$ ihr Maximum in einem Extremalpunkt an, welcher dann ein Minimierer von der Funktion $x \mapsto c^\top x$ ist. (2 Punkte)

Zusatzaufgabe 10.5 (Randlage von Maximierern konvexer Funktionen.) 3 Bonuspunkte

Es sei $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ konvex. Zeigen Sie: Wenn f den Wert

$$\sup_{x \in \operatorname{dom} f} f(x)$$

in einem Punkt in $\operatorname{rel\,int}(\operatorname{dom} f)$ annimmt, dann ist f konstant auf $\operatorname{dom} f$.

Lösung.

Es sei $x^* \in \operatorname{rel\,int}(\operatorname{dom} f)$ ein Maximierer von f über $\operatorname{dom} f$ und $y \in \operatorname{dom} f$ ein beliebiger weiterer Punkt, der o. B. d. A. nicht mit x^* übereinstimmt. Nach der Definition des relativen Inneren gibt es ein $\varepsilon > 0$, so dass der Punkt $z \coloneqq x^* - \varepsilon\,(y - x^*)$ noch in $\operatorname{rel\,int}(\operatorname{dom} f)$ liegt. Der Optimierer x^* liegt nun nach Konstruktion auf der (echten) Verbindungsstrecke von y und z, denn für $\alpha^* = 1 - \frac{1}{1+\varepsilon} \in (0,1)$ ist

$$\alpha^*\,y + (1-\alpha^*)\,z = \alpha^*\,y + (1-\alpha^*)(x^* - \varepsilon\,(y - x^*)) = (1-\alpha^*)(1+\varepsilon)\,x^* + (\alpha^* - \varepsilon\,(1-\alpha^*))\,y = x^*.$$

(2 Punkte)

Auf Grund der Konvexität von f ist nun

$$f(x^*) = f(\alpha^*\,y + (1-\alpha^*)\,z) \leqslant \alpha^* f(y) + (1-\alpha^*)\,f(z),$$

dabei ist die letzte Summe definiert, da y und z in $\operatorname{dom} f$ liegen. Da x^* ein Maximierer und $\alpha \in [0,1]$ ist, kann diese Ungleichung nur gelten, wenn $f(x^*) = f(y) = f(z)$. (1 Punkt)

Zusatzaufgabe 10.6 (Konvexität der Abstandsfunktion) 5 Bonuspunkte

Es sei $C \subseteq \mathbb{R}^n$ eine nichtleere, abgeschlossene und konvexe Menge. Zeigen Sie, dass der euklidische Abstand eines beliebigen Punktes $x \in \mathbb{R}^n$ zur Menge C, also

$$\begin{aligned} \operatorname{dist}_C &: \mathbb{R}^n \to \mathbb{R} \\ \operatorname{dist}_C(x) &\coloneqq \inf_{z \in C} \|x - z\|_2, \end{aligned}$$

eine konvexe Funktion ist.

Gilt die gleiche Aussage auch für nichtkonvexe Mengen C?

Lösung.

Lemma 15.2 sichert die Wohldefiniertheit der Distanzfunktion, der Wert $\operatorname{dist}(x)$ ist gerade der Optimalwert der Projektionsaufgabe (15.1) zum Punkt $x \in \mathbb{R}^n$. Für $x, y \in \mathbb{R}^n$ und $\lambda \in [0,1]$ ist nun

$$\begin{aligned} \operatorname{dist}_C(\lambda x + (1-\lambda)y) &= \inf_{z \in C} \|\lambda x + (1-\lambda)y - z\| \\ &\le \|\lambda x + (1-\lambda)y - \underbrace{(\lambda \operatorname{proj}_C(x) + (1-\lambda)\operatorname{proj}_C(y))}_{\in C,\text{ da } C \text{ konvex ist}}\| \\ &\le \lambda\|x - \operatorname{proj}_C(x)\| + (1-\lambda)\|y - \operatorname{proj}_C(y)\| \\ &= \lambda \operatorname{dist}_C(x) + (1-\lambda)\operatorname{dist}_C(y). \end{aligned}$$

(4 Punkte)

Als Beispiel, das zeigt, dass die gleiche Aussage für nichtkonvexe Mengen nicht gilt, kann man beispielsweise das gleiche Beispiel wie schon in Hausaufgabe 10.2 verwenden. Entlang jeder Linie durch den Urspruch ist die Abstandsfuntion zur 2-Sphäre wie ein W geformt und damit nicht konvex. (1 Punkt)

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