Klausur zur Bachelorprüfung
Optimierungstheorie

vom 07.09.2018

Aufgabe 1: Zu einer Matrix $A \in \mathbb{R}^{m\times n}$, $m, n \in \mathbb{N}$, und $b \in \mathbb{R}^m$ sei $M = \{x \in \mathbb{R}^n : Ax \leq b\}$ gegeben. Zeigen Sie, dass $x \in M$ genau dann eine Ecke von M ist, wenn

$$V := \operatorname{span}\{a_{i*}^\top : i \in I(x)\} = \mathbb{R}^n$$

gilt, wobei a_{i*} die i-te Zeile der Matrix A bezeichnet, und $I(x) = \{i \in \{1,\dots,m\} : (Ax)_i = b_i\}$ die aktiven Indizes an der Stelle x angibt.

Aufgabe 2: Lösen Sie mit Hilfe des Simplex-Verfahrens das lineare Optimierungsproblem

$$\operatorname*{Min}_{x\in M} \; -x_1 - x_2 - x_3$$

auf

$$M = \left\{x \in \mathbb{R}^3_{\geq 0} : x_1 + 2x_2 \leq 2,\; x_3 \geq 1,\; -x_1 - 2x_2 + 2x_3 \leq 2\right\}\,.$$

Aufgabe 3: Gegeben ist

$$\text{(P)} \qquad \operatorname*{Min}_{x\in M} \; f(x) = \mathrm{e}^{(x_1^2 + (x_2-1)^2)}$$

auf

$$M = \left\{x \in \mathbb{R}^2 : |x_2| - x_1 \leq 0\,,\; x_2^2 + x_1 \leq 1\right\}\,.$$

(a) Zeigen Sie, dass es sich um ein konvexes Optimierungsproblem handelt.
(b) Begründen Sie, dass (P) eine Lösung besitzt.
(c) Berechnen Sie eine Lösung zu (P) unter Verwendung der Dualitätstheorie und dem Sattelpunktsatz.

Aufgabe 4: Das Optimierungsproblem

$$\text{(P)} \qquad \operatorname*{Min}_{x\in M} \; f(x) = -x_2^3 + x_3^2 \quad \text{auf} \quad M = \left\{x \in \mathbb{R}^3 : (2 - x_1)(x_2^2 + x_3^2) = x_1^3\,,\; x_2 \leq 1\right\}$$

ist lösbar. Berechnen Sie eine Minimalstelle $\hat{x} \in M$.

Aufgabe 5: Es seien $p \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$ mit $\operatorname{Rg}(A) = m$ gegeben. Gesucht ist eine Lösung $x \in \mathbb{R}^n$ des Gleichungssystems $Ax = b$ mit kleinstem euklidischen Abstand zu p.
(a) Formulieren Sie das Problem durch ein äquivalentes quadratisches Optimierungsproblem.
(b) Begründen Sie, dass es zu jedem $p \in \mathbb{R}^n$ mindestens eine Lösung des Optimierungsproblems gibt.
(c) Zeigen Sie, dass für eine Lösung $\hat{x}$ die Differenz $\hat{x} - p \in \{A^\top v : v \in \mathbb{R}^m\}$ im Bild von $A^\top$ liegt.

Aufgabe 1:

Lösung: "$\Rightarrow$:" Angenommen es ist $V \neq \mathbb{R}^n$. Da V Unterraum ist, gibt es ein $z \in \mathbb{R}^n$ mit $z \neq 0$ und $z \perp V$, d.h. $(Az)_i = 0$ für $i \in I(x)$. Betrachten wir die Konvexkombinationen

$$x = \frac{1}{2}(x + \varepsilon z) + \frac{1}{2}(x - \varepsilon z)\,.$$

Dann ist $(A(x \pm \varepsilon z))_i = b_i$ für $i \in I(x)$. Außerdem lässt sich, wegen $(Ax)_i < b_i$ für $i \notin I(x)$, $\varepsilon > 0$ hinreichend klein wählen, sodass $(A(x \pm \varepsilon z))_i \leq b_i$ gilt. Also gilt $x \pm \varepsilon z \in M$. Somit ist x keine Ecke von M.

"$\Leftarrow$" Ist $V = \mathbb{R}^n$ und $x = \lambda z^1 + (1-\lambda) z^2$ eine Darstellung von $x \in M$ mit $z^j \in M$, $j = 1, 2$, und $\lambda \in (0,1)$. Für $i \in I(x)$ gilt

$$b_i = (Ax)_i = \lambda (Az^1)_i + (1-\lambda)(Az^2)_i \leq b_i\,.$$

Also ist $(Az^1)_i = (Az^2)_i = b_i$ für jedes $i \in I(x)$. Es folgt $(A(z^1 - z^2))_i = 0$, und wir erhalten $z^1 - z^2 \in V^\perp = \{0\}$ bzw. $z^1 = z^2$. Somit ist x Ecke von M.

Aufgabe 2:

Lösung: Da nach Einführen von Schlupfvariablen für die Ungleichungsbedingungen noch keine Basislösung ablesbar ist, starten wir zunächst mit Phase 1 des Simplex-Verfahrens und stellen zu $\operatorname{Min} e^\top (b - Ax)$ auf $M_H = \{x \in \mathbb{R}^6_{\geq 0} : Ax \leq b\}$ das entsprechende Tableau auf:

$$\begin{array}{ccccccccc|c}
1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 2 \\
0 & 0 & \boxed{1} & 0 & -1 & 0 & 0 & 1 & 0 & 1 \\
-1 & -2 & 2 & 0 & 0 & 1 & 0 & 0 & 1 & 2 \\
\hline
-1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & \eta \\
\hline
0 & 0 & -3 & -1 & 1 & -1 & 0 & 0 & 0 & \gamma - 5
\end{array}$$

mit der Basislösung $z = (0,0,0,0,0,0,2,1,2)^\top$. Mit der Bland'schen Regel wählen wir das markierte Pivotelement und berechnen im ersten Schritt

$$\begin{array}{ccccccccc|c}
1 & 2 & 0 & \boxed{1} & 0 & 0 & 1 & 0 & 0 & 2 \\
0 & 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0 & 1 \\
-1 & -2 & 0 & 0 & 2 & 1 & 0 & -2 & 1 & 0 \\
\hline
-1 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & \eta + 1 \\
\hline
0 & 0 & 0 & -1 & -2 & -1 & 0 & 3 & 0 & \gamma - 2
\end{array}$$

Zwei weitere Gauß-Jordan-Umformungen liefern

$$\begin{array}{ccccccccc|c}
1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 2 \\
0 & 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0 & 1 \\
-1 & -2 & 0 & 0 & \boxed{2} & 1 & 0 & -2 & 1 & 0 \\
\hline
-1 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & \eta + 1 \\
\hline
1 & 2 & 0 & 0 & -2 & -1 & 1 & 3 & 0 & \gamma - 0
\end{array}$$

und

$$\begin{array}{ccccccccc|c}
1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 2 \\
-\frac{1}{2} & -1 & 1 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 1 \\
-\frac{1}{2} & -1 & 0 & 0 & 1 & \frac{1}{2} & 0 & -1 & \frac{1}{2} & 0 \\
\hline
-\frac{3}{2} & -2 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & \eta + 1 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & \gamma - 0
\end{array}$$

Also ist mit $\tilde{z} = (0,0,1,2,0,0)^\top$ eine Basislösung zum Optimierungsproblem (mit Schlupfvariablen) gegeben. Wir erhalten für die Phase 2 das Tableau

$$\begin{array}{cccccc|c} \boxed{1} & 2 & 0 & 1 & 0 & 0 & 2 \\ -\frac{1}{2} & -1 & 1 & 0 & 0 & \frac{1}{2} & 1 \\ -\frac{1}{2} & -1 & 0 & 0 & 1 & \frac{1}{2} & 0 \\ \hline -\frac{3}{2} & -2 & 0 & 0 & 0 & \frac{1}{2} & \eta+1 \end{array}$$

Mit dem gewählten Pivotelement ergibt sich

$$\begin{array}{cccccc|c} 1 & 2 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 2 \\ 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1 \\ \hline 0 & 1 & 0 & \frac{3}{2} & 0 & \frac{1}{2} & \eta+4 \end{array}$$

Aus dem letzten Tableau lesen wir die Lösung des Optimierungsproblems $\hat{x} = (2,0,2)^\top$ mit Zielfunktionswert $-\hat{x}_1 - \hat{x}_2 - \hat{x}_3 = -4$ ab.

Aufgabe 3:

Lösung: (a) Wir setzen $h_1(x) = |x_2| - x_1$ und $h_2(x) = x_2^2 + x_1 - 1$. Mit $\lambda \in [0,1]$ folgt für $x, y \in \mathbb{R}^2$

$$\begin{aligned} h_1(\lambda x + (1-\lambda)y) &= |\lambda x_2 + (1-\lambda)y_2| - (\lambda x_1 + (1-\lambda)y_1) \\ &\le \lambda(|x_2| - x_1) + (1-\lambda)(|y_2| - y_1) = \lambda h_1(x) + (1-\lambda)h_1(y)\,. \end{aligned}$$

Also ist h_1 konvex.

Weiter sehen wir aus

$$\nabla h_2(x) = \begin{pmatrix} 1 \\ 2x_2 \end{pmatrix} \qquad \text{und} \qquad \nabla^2 h_2(x) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}\,,$$

dass die Hessematrix positiv semidefinit ist. Also ist h_2 auch konvex.

Genauso erhalten wir für $\tilde{h}(x) = x_1^2 + (x_2 - 1)^2$ mit der positiv definiten Hessematrix $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ eine konvexe Funktion $\tilde{h} : \mathbb{R}^2 \to \mathbb{R}$. Weiterhin ist die Exponentialfunktion mit $\tilde{g}(t) = \mathrm{e}^t$ konvex (da die 2.Abl. positiv ist) und streng monoton steigend auf $\mathbb{R}$. Somit ist auch $f : \mathbb{R}^2 \to \mathbb{R}$ eine konvexe Funktion, da mit der Konvexität von $\tilde{h}$ die Abschätzung

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \tilde{g}\left(\tilde{h}(\lambda x + (1-\lambda)y)\right) \\ &\le \tilde{g}\left(\lambda \tilde{h}(x) + (1-\lambda)\tilde{h}(y)\right) \\ &\le \lambda f(x) + (1-\lambda) f(y) \end{aligned}$$

für $x, y \in \mathbb{R}^2$ gilt.

(b) f ist als Verkettung stetiger Funktionen stetig und M ist beschränkt; denn aus $x_1 \ge |x_2| \ge 0$ und $x_1 \le 1 - x_2^2 \le 1$ ergibt sich $x_1 \in [0,1]$, und weiter liefert $x_2^2 \le 1 - x_1 \le 1$ die Beschränkung $|x_2| < 1$. Da darüberhinaus M abgeschlossen ist, ist M kompakt. Mit dem Satz von Weierstraß ergibt sich die Existenz von Lösungen $\hat{x} \in M$ zu (P).

(c) Zunächst folgt mit $x = (1/2, 0)^\top \in M$ aus $h_1(x) = -1/2 < 0$ und $h_2(x) = -1/2 < 0$, dass die Menge M die Slater-Bedingung erfüllt. Somit gibt es nach dem starken Dualitätssatz eine Lösung $\hat{u} \in \{u \in \mathbb{R}^2_{\ge 0} : F(u) > -\infty\}$ zum dualen Problem mit

$$F(u) = \inf_{x \in \mathbb{R}^2} \left(\mathrm{e}^{x_1^2 + (x_2-1)^2} + u_1(|x_2| - x_1) + u_2(x_2^2 + x_1 - 1) \right)\,.$$

Für das Paar $\hat{x}, \hat{u}$ gilt die Komplementaritätsbedingung $\hat{u}^\top h(\hat{x}) = 0$ und mit dem Sattelpunktsatz die Abschätzung

$$\begin{aligned} &\mathrm{e}^{\hat{x}_1^2+(\hat{x}_2-1)^2} + u_1(|\hat{x}_2| - \hat{x}_1) + u_2(\hat{x}_2^2 + \hat{x}_1 - 1) \\ &\leq \mathrm{e}^{\hat{x}_1^2+(\hat{x}_2-1)^2} + \hat{u}_1(|\hat{x}_2| - \hat{x}_1) + \hat{u}_2(\hat{x}_2^2 + \hat{x}_1 - 1) = \mathrm{e}^{\hat{x}_1^2+(\hat{x}_2-1)^2} \\ &\leq \mathrm{e}^{x_1^2+(x_2-1)^2} + \hat{u}_1(|x_2| - x_1) + \hat{u}_2(x_2^2 + x_1 - 1) \end{aligned}$$

für alle $x \in \mathbb{R}^2$, $u \in \mathbb{R}^2_{\geq 0}$.

Aufgabe 4:

Lösung: Mit $g(x) = (2 - x_1)(x_2^2 + x_3^2) - x_1^3$ und $h(x) = x_2 - 1$ betrachten wir zunächst die MFB Bedingung. Es gilt

$$\nabla g(x) = \begin{pmatrix} -3x_1^2 - x_2^2 - x_3^2 \\ 2(2 - x_1)x_2 \\ 2(2 - x_1)x_3 \end{pmatrix} \neq 0$$

genau dann, wenn $x \neq 0$ ist. Weiterhin gibt es $\tilde{z} \in \mathbb{R}^3$ mit $h(x) + h'(x)\tilde{z} = x_2 - 1 + \tilde{z}_2 < 0$ und $(\nabla g(x))^\top \tilde{z} = 0$. Somit ist die MFB für alle $x \in M$ mit $x \neq 0$ erfüllt.

Mit der Lagrange-Funktion

$$L(x, u, v) = -x_2^3 - x_3^2 + u(x_2 - 1) + v\left[(2 - x_1)(x_2^2 + x_3^2) - x_1^3\right]$$

ergibt sich für eine Extremalstelle die notwendge Bedingung

$$\nabla_x L(x, u, v) = \begin{pmatrix} 0 \\ -3x_2^2 \\ 2x_3 \end{pmatrix} + u \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + v \begin{pmatrix} -3x_1^2 - x_2^2 - x_3^2 \\ 2(2 - x_1)x_2 \\ 2(2 - x_1)x_3 \end{pmatrix} = 0\,.$$

Wir erhalten insgesamt sieben Bedingungen:
(1) $v(3x_1^2 + x_2^2 + x_3^2) = 0$
(2) $-3x_2^2 + u + 2(2 - x_1)x_2 v = 0$
(3) $2x_3 + 2(2 - x_1)x_3 v = 0$
(4) $(2 - x_1)(x_2^2 + x_3^2) - x_1^3 = 0$
(5) $x_2 \leq 1$
(6) $u(x_2 - 1) = 0$
(7) $u \geq 0$

Aus (1) folgt entweder $v = 0$ oder $x = 0$. Den Fall $x = 0$, bei dem die MFB verletzt ist, berücksichtigen wir später. Mit $v = 0$ erhalten wir aus (3) $x_3 = 0$ und aus (2) $u = 3x_2^2 \geq 0$. Mit (4) folgt $x_1 \neq 2$

Aufgabe 5:

Lösung: (a) Wir betrachten die Zielfunktion $f(x) = \|x-p\|^2$ und zerlegen $x = x^+ - x^-$ mit $x_i^+ = \max\{x_i, 0\}$ und $x_i^- = \max\{-x_i, 0\}$ für $i = 1, \dots, n$. Setzen wir $y = (x^+, x^-)^\top \in \mathbb{R}^{2n}$ so ergibt sich

$$\begin{aligned}\|x-p\|^2 &= (x,x) - (x,p) - (p,x) - (p,p) \\ &= y^\top \underbrace{\begin{pmatrix} I & -I \\ -I & I \end{pmatrix}}_{=:Q\in\mathbb{R}^{2n\times 2n}} y \underbrace{-2\begin{pmatrix} p \\ -p \end{pmatrix}}_{=:c\in\mathbb{R}^{2n}} y + \|p\|^2 .\end{aligned}$$

Insgesamt erhalten wir mit der Matrix Q und dem Vektor c das äquivalente quadratische Optimierungsproblem

$$\text{Min } y^\top Q y + c^\top y$$

auf

$$M = \left\{ y \in \mathbb{R}^{2n}_{\geq 0} : \big(A | -A\big) y = b \right\} .$$

(b) Mit

$$y^\top Q y + c^\top y = \|x-p\|^2 - \|p\|^2 > -\infty , \quad \text{für alle } y \in \mathbb{R}^{2n}_{\geq 0} ,$$

ist die Zielfunktion nach unten beschränkt. Außerdem ist M zulässig, da es zu $Ax = b$ wegen $\mathrm{Rg}(A) = m$ stets eine Lösung zu $Ax = b$ gibt. Mit dem allgemeinen Existenzsatz zu quadratischen Problemen folgt Existenz einer Lösung $\hat{y}$ bzw. $\hat{x}$ des Optimierungsproblems.

(c) Mit den Kuhn-Tucker Bedingungen für eine Lösung $\hat{y}$ zum quadratischen Problem gibt es $\hat{v} \in \mathbb{R}^m$ mit

$$\begin{aligned} &\text{(i)} \qquad 2Q\hat{y} + c + \begin{pmatrix} A^\top \\ -A^\top \end{pmatrix} \hat{v} \geq 0 \\ &\text{(ii)} \qquad \left(2Q\hat{y} + c + \begin{pmatrix} A^\top \\ -A^\top \end{pmatrix} \hat{v}\right)^\top \hat{y} = 0 . \end{aligned}$$

Mit $\hat{x}^+ = \hat{y}_j$ und $\hat{x}^- = \hat{y}_{j+m}$ ergibt sich aus der ersten Bedingung (i), dass $2(\hat{x}^+ - \hat{x}^- - p) + A^\top \hat{v} \geq 0$ und $-2(\hat{x}^+ - \hat{x}^- - p) - A^\top \hat{v} \geq 0$ gilt. Somit ist $\hat{x} - p = -\frac{1}{2} A^\top \hat{v} \in \{A^\top v : v \in \mathbb{R}^m\}$.

Bemerkung: Mit $b - Ap = A(\hat{x} - p) = -\frac{1}{2} A A^\top \hat{v}$ folgt $-\frac{1}{2}\hat{v} = (AA^\top)^{-1}(b - Ap)$, da die Inverse zu $AA^\top$ wegen $\mathrm{Rg}(A) = m$ existiert. Einsetzen liefert die Darstellung $\hat{x} = p - \frac{1}{2} A^\top \hat{v} = p + A^\dagger (b - Ap)$ mit der Pseudo-Inversen $A^\dagger = A^\top (AA^\top)^{-1}$ für die Lösung des Optimierungsproblems.