Karlsruher Institut für Technologie

Fakultät für Mathematik
Institut für Angewandte und Numerische Mathematik
PD. Dr. Frank Hettlich
Leonie Fink

Karlsruhe, den 6. Juni 2025

Optimierungstheorie
Lösung zum 6. Übungsblatt – Sommersemester 2025

Aufgabe 16: Es sei $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $e = (1,\dots,1)^\top \in \mathbb{R}^m$ und es gelte $b \notin \{Ax:\ x \in \mathbb{R}^n\}$. Die diskrete Tschebyscheffsche Approximationsaufgabe lautet:

$$\text{(T)}\ \ \text{Minimiere } \|Ax-b\|_\infty \text{ unter } x \in \mathbb{R}^n.$$

(a) Beweisen Sie: $\widehat{x} \in \mathbb{R}^n$ ist eine Lösung von (T) mit $\|A\widehat{x}-b\|_\infty = \widehat{x}_0$ genau dann, wenn $(\widehat{x}_0,\widehat{x}) \in \mathbb{R}\times\mathbb{R}^n$ eine Lösung des Problems (P) ist:

$$\text{(P)}\quad \text{Maximiere } -x_0 \text{ unter } Ax - x_0 e \le b,\ -Ax - x_0 e \le -b \text{ und } -x_0 \le 0.$$

(b) Formulieren Sie das duale Problem (D) zu (P).

(c) Zeigen Sie: Die Probleme (P) und (D) sind lösbar.

(d) Es sei $(\widehat{x}_0,\widehat{x}) \in \mathbb{R}\times\mathbb{R}^n$ eine Lösung von (P) und $\widehat{u} \in \mathbb{R}^{2m+1}$ eine Lösung von (D). Beweisen Sie, dass folgendes gilt:

$$\begin{aligned} \widehat{u}_j = 0 \quad &\text{oder} \quad (A\widehat{x})_j - b_j = \widehat{x}_0 \ \text{ für } j=1,\dots,m, \text{ sowie} \\ \widehat{u}_{j+m} = 0 \quad &\text{oder} \quad (A\widehat{x})_j - b_j = -\widehat{x}_0 \ \text{ für } j=1,\dots,m. \end{aligned}$$

(e) Ist $\widehat{x}$ eine Lösung von (T), dann gibt es ein $\widehat{y} \in \mathbb{R}^m \setminus \{0\}$ mit $\sum_{j=1}^m \widehat{y}_j \operatorname{sign}((A\widehat{x}-b)_j)\, a_j = 0$, wobei $a_j = (a_{j1},\dots,a_{jn})^\top$ die (transponierte) j-te Zeile von A bezeichne.

Lösung: Zu (a): Es gelten die folgenden Äquivalenzen:

$$\begin{aligned} & \widehat{x} \in \mathbb{R}^n \text{ löst (T) mit } \|A\widehat{x}-b\|_\infty = \widehat{x}_0 \text{ und } \widehat{x}_0 \ge 0 \text{ minimal.} \\ \iff\ & \text{Für } \widehat{x} \in \mathbb{R}^n \text{ gilt } \max_{i=1,\dots,m} |(A\widehat{x}-b)_i| = \widehat{x}_0 \text{ und } \widehat{x}_0 \ge 0 \text{ minimal.} \\ \iff\ & \text{Für } \widehat{x} \in \mathbb{R}^n \text{ gilt } |(A\widehat{x}-b)_i| \le \widehat{x}_0\ (i=1,\dots,m) \text{ und } \widehat{x}_0 \ge 0 \text{ minimal.} \\ \iff\ & \text{Für } \widehat{x} \in \mathbb{R}^n \text{ gilt } (A\widehat{x}-b)_i \le \widehat{x}_0\ (i=1,\dots,m) \text{ und } -(A\widehat{x}-b)_i \le \widehat{x}_0\ (i=1,\dots,m) \\ & \text{und } \widehat{x}_0 \ge 0 \text{ minimal.} \\ \iff\ & \text{Für } \widehat{x} \in \mathbb{R}^n \text{ gilt } A\widehat{x} - \widehat{x}_0 e \le b \text{ und } -A\widehat{x} - \widehat{x}_0 e \le -b \\ & \text{und } -\widehat{x}_0 \le 0 \text{ maximal.} \\ \iff\ & (\widehat{x}_0,\widehat{x}) \in \mathbb{R}\times\mathbb{R}^n \text{ löst das Problem (P).} \end{aligned}$$

Zu (b): Das Problem (P) lässt sich mit $x = (x_1,\dots,x_n)^\top$ formulieren als

$$\text{(P)}\qquad \text{Maximiere } (-1,0,\dots,0)\begin{pmatrix} x_0 \\ x \end{pmatrix} \text{ unter } \begin{pmatrix} -e & A \\ -e & -A \\ -1 & 0^\top \end{pmatrix}\begin{pmatrix} x_0 \\ x \end{pmatrix} \le \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}.$$

Daher ist das duale Problem zu (P) gegeben durch

$$\text{(D)} \qquad \text{Minimiere } (b^\top, -b^\top, 0) \begin{pmatrix} u_1 \\ \vdots \\ u_{2m+1} \end{pmatrix}$$

$$\text{unter } u \geq 0 \text{ und } \begin{pmatrix} -e^\top & -e^\top & -1 \\ A^\top & -A^\top & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_{2m+1} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Zu (c): (P) ist zulässig: Setze $x = 0$ und $x_0 = \|b\|_\infty$, dann gilt $\begin{pmatrix} -x_0 e \\ -x_0 e \\ -x_0 \end{pmatrix} \leq \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$.

Ebenso ist (D) zulässig. Setze dazu $u_1 = \ldots = u_{2m} = 0$ und $u_{2m+1} = 1$, dann gilt $u \geq 0$ und

$$\begin{pmatrix} -e^\top & -e^\top & -1 \\ A^\top & -A^\top & 0 \end{pmatrix} u = \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Mit dem schwachen Dualitätssatz folgt daher die Lösbarkeit beider Probleme und mit dem starken Dualitätssatz $\min(\text{P}) = \max(\text{D})$.

Zu (d): Es sei $(\widehat{x}_0, \widehat{x}) \in \mathbb{R} \times \mathbb{R}^n$ eine Lösung von (P) und $\widehat{u} \in \mathbb{R}^{2m+1}$ eine Lösung von (D). Dann gilt:

$$\widehat{u}^\top \begin{pmatrix} -e & A \\ -e & -A \\ -1 & 0^\top \end{pmatrix} \begin{pmatrix} \widehat{x}_0 \\ \widehat{x} \end{pmatrix} = (-1, 0, \ldots, 0) \begin{pmatrix} \widehat{x}_0 \\ \widehat{x} \end{pmatrix} = (b^\top, -b^\top, 0)\widehat{u},$$

$$\text{also} \qquad \left(\underbrace{\begin{pmatrix} b \\ -b \\ 0 \end{pmatrix} - \begin{pmatrix} -e & A \\ -e & -A \\ -1 & 0^\top \end{pmatrix} \begin{pmatrix} \widehat{x}_0 \\ \widehat{x} \end{pmatrix}}_{\geq 0} \right)^\top \underbrace{\widehat{u}}_{\geq 0} = 0.$$

Zeilenweise gelesen bedeutet dies

$$\begin{aligned} (b + \widehat{x}_0 e - A\widehat{x})_i = 0 \quad &\text{oder} \quad \widehat{u}_i = 0 \text{ für } j = 1, \ldots, m, \text{ sowie} \\ (-b + \widehat{x}_0 e + A\widehat{x})_i = 0 \quad &\text{oder} \quad \widehat{u}_{i+m} = 0 \text{ für } j = 1, \ldots, m, \end{aligned}$$

und wie man sieht, ist das gleichbedeutend mit der Behauptung.

Zu (e): Es sei $\widehat{x}$ eine Lösung von (T), also $(\|A\widehat{x} - b\|_\infty, \widehat{x})$ eine von (P), und $\widehat{u} \geq 0$, $\widehat{u} \neq 0$ eine Lösung von (D). Dabei gilt $\widehat{u} \neq 0$, da sonst gelten würde:

$$\max(\text{D}) = 0 = \min(\text{P}), \text{ also } \|A\widehat{x} - b\|_\infty = 0, \text{ also } b \in \{Ax : x \in \mathbb{R}^n\}.$$

An dieser Stelle halten wir fest, dass die Ungleichungen $u_j > 0$ und $u_{j+m} > 0$ für $j \in \{1, \ldots, m\}$ nicht beide gleichzeitig gelten können, da dies Teil (d) widersprechen würde.
Da $\widehat{u} = (\widehat{u}_1, \ldots, \widehat{u}_{2m}, \widehat{u}_{2m+1})^\top \geq 0$ eine Łösung von (D) ist, gilt insbesondere (siehe Teil (b))

$$(*) \qquad A^\top \begin{pmatrix} \widehat{u}_1 \\ \vdots \\ \widehat{u}_m \end{pmatrix} - A^\top \begin{pmatrix} \widehat{u}_{m+1} \\ \vdots \\ \widehat{u}_{2m} \end{pmatrix} = \sum_{j=1}^m a_j (\widehat{u}_j - \widehat{u}_{j+m}) = 0.$$

Setzt man nun $\widehat{y}_j = \begin{cases} \widehat{u}_j, & \text{falls} \quad \widehat{u}_j > 0, \\ \widehat{u}_{j+m}, & \text{falls} \quad \widehat{u}_{j+m} > 0, \\ 0 & \text{sonst,} \end{cases}$ für $j = 1, \ldots, m$, dann ist $\widehat{y} = (\widehat{y}_1, \ldots, \widehat{y}_m)^\top$ das gesuchte $\widehat{y}$.

Wegen $\widehat{u} \neq 0$ gilt $\widehat{y} \neq 0$ und $\operatorname{sign}((A\widehat{x}-b)_j) = 1$, falls $\widehat{u}_j > 0$ für $j \in \{1,\dots,m\}$, sowie $\operatorname{sign}((A\widehat{x}-b)_j) = -1$, falls $\widehat{u}_{j+m} > 0$ für $j \in \{1,\dots,m\}$. Also erfüllt $\widehat{y}$ die angegebene Gleichung.

Aufgabe 17: Die Auszahlungsmatrix eines Spieles sei gegeben durch

$$A = \begin{pmatrix} 1 & -1 & -2 \\ -1 & 2 & 1 \\ 2 & -1 & 2 \end{pmatrix}.$$

Zeigen Sie, ohne die optimalen Strategien der Spieler zu bestimmen, dass das Spiel unfair ist, das heißt, dass folgendes gilt:

$$\min_{\substack{x\geq 0\\ e^\top x=1}} \max_{\substack{y\geq 0\\ e^\top y=1}} y^\top Ax \neq 0.$$

Hinweis: Nutzen Sie die aus der Vorlesung bekannte Identität $\max\limits_{\substack{y\geq 0\\ e^\top y=1}} y^\top Ax = \max\limits_{i=1,2,3} (Ax)_i$ und unterscheiden Sie die Fälle $x_2 \in [0, 1-\varepsilon]$ und $x_2 \in [1-\varepsilon, 1]$ mit geeignetem $\varepsilon \in (0,1)$.

Lösung: Es sei $\varepsilon \in (0,1)$, $x \geq 0$ und $e^\top x = 1$. Weiter bezeichne $x_0 = \max\limits_{i=1,2,3} (Ax)_i$. Dann gilt

$$3x_0 \geq (Ax)_1 + (Ax)_2 + (Ax)_3 = e^\top Ax = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix} x = 2x_1 + x_3.$$

Wegen $x_1 + x_2 + x_3 = 1$ folgt damit $3x_0 \geq 2x_1 + x_3 = 1 + x_1 - x_2 \geq 1 - x_2$, also $x_0 \geq \frac{1}{3}(1-x_2)$. Damit gilt mit dem im Hinweis zitierten Resultat $\max\limits_{\substack{y\geq 0\\ e^\top y=1}} y^\top Ax = \max\limits_{i=1,2,3} (Ax)_i$ das Folgende:

$$\min_x \max_{i=1,2,3} \left\{ (Ax)_i :\ x \geq 0,\ e^\top x = 1,\ x_2 \leq 1-\varepsilon \right\} \geq \frac{\varepsilon}{3}.$$

Wir betrachten nun ein zulässiges x (d.h. $x \geq 0$, $x_1 + x_2 + x_3 = 1$) mit $x_2 \geq 1-\varepsilon$. Dann ist insbesondere $x_1 \in [0,\varepsilon]$ und $x_3 \in [0,\varepsilon]$, und wir erhalten

$$(Ax)_2 = -x_1 + 2x_2 + x_3 \geq -\varepsilon + 2(1-\varepsilon) + 0 = 2 - 3\varepsilon.$$

Damit ist insbesondere

$$\min_x \max_{i=1,2,3} \left\{ (Ax)_i :\ x \geq 0,\ e^\top x = 1,\ x_2 \geq 1-\varepsilon \right\} \geq 2 - 3\varepsilon.$$

Führt man beide Abschätzungen zusammen, so erhält man etwa für $\varepsilon = \frac{3}{5}$

$$\min_{\substack{x\geq 0\\ e^\top x=1}} \max_{\substack{y\geq 0\\ e^\top y=1}} y^\top Ax = \min_x \max_{i=1,2,3} \left\{ (Ax)_i :\ x \geq 0,\ e^\top x = 1 \right\} \geq \frac{1}{5}.$$

Aufgabe 18: Beweisen Sie die in Analogie zu Lemma 3.3 der Vorlesung stehende Aussage: $K \subseteq \mathbb{R}^n$ ist genau dann ein konvexer Kegel, wenn für alle $x_i \in K$ und alle $\lambda_i \geq 0$ ($i = 1,\dots,m$, $m \in \mathbb{N}$ beliebig) auch $\sum_{i=1}^m \lambda_i x_i \in K$ gilt.

Hinweis: Verwenden Sie vollständige Induktion.

Lösung: „$\Leftarrow$“: Ist $x_1 := x \in K$ und $\lambda_1 := \lambda \geq 0$, so gilt nach Voraussetzung auch $\lambda x = \lambda_1 x_1 \in K$. Folglich ist K ein Kegel.
Sind $x_1 := x$ und $x_2 := y \in K$ und ist $\lambda \in [0,1]$, so ist auch $1-\lambda \in [0,1]$. Mit $\lambda_1 := \lambda$ und $\lambda_2 := 1-\lambda$ folgt wiederum nach Voraussetzung

$$\lambda x + (1-\lambda)y = \lambda_1 x_1 + \lambda_2 x_2 \in K.$$

Daher ist K auch konvex, insgesamt K also ein konvexer Kegel.

„$\Rightarrow$“: Wir beweisen die Behauptung mittels vollständiger Induktion nach m.
Für $m = 1$ gilt: Ist $\lambda_1 \geq 0$, $x_1 \in K$, dann ist auch $\lambda_1 x_1 \in K$, da K ein Kegel ist.
Für $m = 2$ gilt: Sind $\lambda_1, \lambda_2 \geq 0$, $x_1, x_2 \in K$, wobei o.B.d.A. $\lambda_1 + \lambda_2 > 0$ angenommen werden kann, so ist

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2 \in K,$$

da K konvex ist, und damit auch

$$\lambda_1 x_1 + \lambda_2 x_2 = (\lambda_1 + \lambda_2) \cdot \underbrace{\left[\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2 \right]}_{\in K} \in K,$$

da K ein Kegel ist (oder nach Fall $m = 1$).
Die Aussage gelte nun für ein festes m und beliebige $x_i \in K$ und alle $\lambda_i \geq 0$, $i = 1, \ldots, m$.
Es seien jetzt $x_1, \ldots, x_{m+1} \in K$ und $\lambda_1, \ldots, \lambda_{m+1} \geq 0$. Nach Voraussetzung ist dann $\lambda_1 x_1 + \ldots + \lambda_m x_m \in K$. Damit gilt (siehe Fall $m = 2$)

$$\lambda_1 x_1 + \ldots + \lambda_{m+1} x_{m+1} = 1 \cdot \underbrace{(\lambda_1 x_1 + \ldots + \lambda_m x_m)}_{\in K} + \lambda_{m+1} \underbrace{x_{m+1}}_{\in K} \in K.$$