Prof. Dr. Volker Kaibel
Dr. Dirk Oliver Theis
Otto-von-Guericke U MD
SS 2009

Lineare Optimierung
Übungsblatt 11

Abgabetermin: 24.06.2009
25 Punkte = 100%
Fragen/Kommentare/Errata: `http://dirkolivertheis.wordpress.com/2009/06/11/linopt09-blatt-11`

Aufgabe 1 Barrier-Methode (6 Punkte)

Seien $X = [0,1]^2 \subseteq \mathbb{R}^2$, $\eta_0 = 1$, $x^{(0)} = (0.2, 0.6)$, und $c = (5,2)$.

(a) Bestimmen Sie eine Barrier-Funktion (siehe Kor. 5.26) $f\colon \operatorname{int} X \to \mathbb{R}$ und für jedes $x \in \operatorname{int} X$ die Hesse-Matrix von f in x, sowie für die Funktionen f_η den Newton-Schritt in x.

(b) Berechnen Sie gemäß der Short-Step-Barrier Methode $x^{(1)}, \dots, x^{(4)}$, wobei $\eta_k = 1.8\eta_{k-1}$ (anstatt der in der Algorithmenbeschreibung angegebenen Formel $\eta_k = (1+1/(8\sqrt{\vartheta_f}))\eta_{k-1}$ gesetzt werden soll.

Bemerkung. Mit dem Faktor 1.8 (statt $1/(8\sqrt{\vartheta_f}) \approx 1.176$) garantiert die Theorie nicht mehr, dass man $\operatorname{int} X$ nicht verlässt oder sich zu weit vom zentralen Pfad entfernt; bei den ersten vier Schritten geht aber im obigen Fall alles gut, und der Fortschritt ist größer. Es ist interessant, ein kleines Programm zu schreiben, mit dem man den Algorithmus mit Faktor 1.176 über mehr Iterationen laufen lässt. Dann kann man auch ein wenig mit den anderen Parametern (wie $x^{(0)}$ und η_0) experimentieren.

Lösung:

(a). Gemäß Korollar 5.26 ist

$$f\colon (x,y) \mapsto -\ln(x) - \ln(y) - \ln(1-x) - \ln(1-y)$$

eine Barrierfunktion. Es gilt

$$\operatorname{grad}_{(x,y)} f = \begin{pmatrix} 1/(1-x) - 1/x \\ 1/(1-y) - 1/y \end{pmatrix} \quad \text{und} \quad H_{(x,y)} f = \begin{pmatrix} \frac{2x^2-2x+1}{x^2(1-x)^2} & 0 \\ 0 & \frac{2y^2-2y+1}{y^2(1-y)^2} \end{pmatrix}.$$

also erhält man

$$\operatorname{grad}_{(x,y)} f_\eta = \begin{pmatrix} 5\eta + 1/(1-x) - 1/x \\ 2\eta + 1/(1-y) - 1/y \end{pmatrix} \quad \text{und} \quad H_{(x,y)} f_\eta = H_{(x,y)} f.$$

Man rechnet weiter

$$\operatorname{new}_x f = -\begin{pmatrix} \frac{x^2(1-x)^2}{2x^2-2x+1} & 0 \\ 0 & \frac{y^2(1-y)^2}{2y^2-2y+1} \end{pmatrix} \begin{pmatrix} 5\eta + 1/(1-x) - 1/x \\ 2\eta + 1/(1-y) - 1/y \end{pmatrix}$$

$$= \begin{pmatrix} -\dfrac{5\eta x^2(1-x)^2}{2x^2-2x+1} - \dfrac{x^2(1-x)}{2x^2-2x+1} + \dfrac{x(1-x)^2}{2x^2-2x+1} \\ -\dfrac{2\eta y^2(1-y)^2}{2y^2-2y+1} - \dfrac{y^2(1-y)}{2y^2-2y+1} + \dfrac{y(1-y)^2}{2y^2-2y+1} \end{pmatrix}$$

(b). (Keine Musterlösung für (b).)

Aufgabe 2 Kodierungslängen von Lösungen von Gleichungssystemen (5 Punkte)

Für alle Matrizen (und Vektoren) X sei $|X|_\infty$ das Maximum der Beträge der Einträge: $|X|_\infty := \max_{k,l} |X_{k,l}|$.

Sei A eine invertierbare $(n \times n)$-Matrix mit ganzzahligen Einträgen, $b \in \mathbb{Z}^n$ und $x \in \mathbb{R}^n$ mit $Ax = b$. Zeigen Sie mit Hilfe der Cramer'schen Regel und der Hadamard'schen Ungleichung: Die Einträge von x sind rational, und man kann ganze Zahlen q und $p_1, \dots, p_n$ wählen mit $x_l = \frac{p_l}{q}$, so dass gilt

$$|q| \leq n^{n/2} |A|_\infty^n \quad \text{und} \quad |p_l| \leq n^{n/2} |A|_\infty^{n-1} |b|_\infty.$$

Schätzen Sie die Kodierungslänge von x ab durch diejenigen von A und b ab.

Lösung:

Es sei $A^{\setminus(k,l)}$ die Matrix, die man erhält, wenn man aus A die k-te Zeile und die l-te Spalte löscht, und $A^{(l),b}$ die Matrix, die man erhält, wenn man in A die l-te Spalte durch b ersetzt. Nach der Cramer'schen Regel ist ja

$$x_l = \frac{\sum_{k=1}^n (-1)^{k+l} \det A^{\setminus(k,l)} b_k}{\det A} = \frac{\det A^{(l),b}}{\det A}.$$

Wir definieren also $q := \det A$ und $p_l := \det A^{(l),b}$. Da die Determinante ein Polynom in den Einträgen ist sind q und die p_k ganze Zahlen.

Für die Ungleichung über die Koeffizienten muss man die Determinanten abschätzen. Die Hadamard'sche Ungleichung liefert

$$|\det B| \leq \prod_{l=1}^n |B_{*,l}|_2 \leq \prod_{l=1}^n \sqrt{n}\, |B_{*,l}|_\infty \leq n^{n/2} |B|_\infty^n.$$

Damit folgt sofort die Abschätzung für $|q|$. Für p_l hat man

$$|p_l| = |\det A^{(l),b}| \leq |b| \prod_{j \neq l} |A_{*,j}|_2 \leq n^{n/2} |A|_\infty^{n-1} |b|_\infty.$$

Für den letzten Teil der Aufgabe beobachten wir, dass die Kodierungslänge einer Zahl ungefähr deren Logarithmus ist, und erhalten als obere Schranke

$$n(\tfrac{n}{2} \log n + n \log |A|_\infty + \log |b|_\infty) = \tfrac{n^2}{2} \log n + n^2 \log |A|_\infty + n \log |b|_\infty.$$

Die letzten beiden Terme sind aber gerade die Kodierungslängen von A bzw. b.

Aufgabe 3 Seiten von konvexen Mengen (8 Punkte)

Sei K eine abgeschlossene konvexe Menge.

- Eine Teilmenge $X \subseteq K$ heisst *extremal in K,* wenn jedes Geradensegment $[x, y] := \{(1-t)x + ty \mid t \in [0,1]\}$, dessen Endecken x, y in K liegen und dessen Schnitt mit X nicht nur aus höchstens einem der Punkte x,y besteht, ganz in X liegt. Formaler bedeutet das, dass für alle $x, y \in K$ gilt: $[x, y] \cap X \notin \{\emptyset, \{x\}, \{y\}\} \Longrightarrow [x, y] \subseteq X$.
- Eine Teilmenge $F \subseteq K$ heisst *exponiert in K,* wenn es einen Halbraum H gibt, der K enthält, und für den der Schnitt seiner begrenzende Hyperebene mit K gerade F ist. Exponierte Teilmengen konvexer Mengen nennt man auch *Seiten.*

Wenn klar ist, auf welche konvexe Obermenge man sich bezieht, lässt man das „in K" oft weg. Zeigen Sie:

(a) Jede extremale Teilmenge von K ist konvex.

(b) Jede exponierte Teilmenge von K ist auch extremal.

(c) Es gibt extremale echte Teilmengen von konvexen Körpern, die nicht exponiert sind. (Geben sie einen konvexen Körper K und eine extremale echte Teilmenge X von K an, so dass X in K nicht exponiert ist.)

(d) Jede extremale Teilmenge X' einer extremalen Teilmenge X von K ist eine extremale Teilmenge von K.

(e) Es gibt exponierte Teilmengen von exponierten Teilmengen von konvexen Körpern K, die nicht exponiert in K sind. (Geben sie einen konvexen Körper K, eine exponierte (echte) Teilmenge X von K und eine exponierte (echte) Teilmenge X' von X an, so dass X' in K nicht exponiert ist.)

Lösung:

(a). Seien $x, y \in X$. Aufgrund der Definition von extremal ist damit sofort $[x, y] \subseteq X$. Damit ist die Konvexität gezeigt.

(b). Es sei F eine Seite von K und H ein Halbraum wie in der Definition, und ∂H sein Rand. Hat man $x, y \in K$ so dass $[x, y] \cap F$ weder leer noch $\{x\}$ noch $\{y\}$ ist, so gilt zunächst $[x, y] \subseteq \partial H$. Da auch $[x, y] \subseteq K$ gilt, erhält man $[x, y] \subseteq F$ wie gefordert.

(c). Betrachte den konvexen Körper

$$K := \{(x, y) \in \mathbb{R}^2 \mid -\pi/2 \leq x \leq \pi/2,\ -1 \leq y \leq 1,\ y \leq \cos x\ \forall x \leq 0\}.$$

Der Punkt $(0, 1)$ ist extremal aber nicht exponiert: Es gibt nur eine Gerade, die den Punkt enthält und K nicht durchschneidet, und die enthält das ganze Geradensegment $[-1, 0] \times \{1\}$.

(d). Es seien $x, y \in K$ mit $[x, y] \cap X' \neq \{\emptyset, \{x\}, \{y\}\}$. Wegen $[x, y] \cap X \supset [x, y] \cap X'$ gilt auch $[x, y] \cap X \neq \{\emptyset, \{x\}, \{y\}\}$, und somit $[x, y] \subseteq X$. Da aber X' exponiert in X ist, folgt sogar $[x, y] \subseteq X'$, wie gefordert.

(e). Mit dem gleichen konvexen Körper wie in (c), betrachte die exponierte Menge $[-1, 0] \times \{1\}$. Offenbar ist der Punkt $(0, 1)$ exponiert in diesem Geradensegment, aber, wie wir in (c) überlegt haben, ist der Punkt nicht exponiert in K.

Aufgabe 4 Seiten von Poyedern (6 Punkte)

Es seien U, V endliche Mengen in $\mathbb{R}^n$, sowie $P = \operatorname{conv} V + \operatorname{ccone} U$ ein Polyeder mit einer Seite F definiert von der Ungleichung $a^\top x \leq \beta$. Zeigen Sie:

(a) Mit $U_F := \{u \in U \mid a^\top u = 0\}$ und $V_F := \{v \in V \mid a^\top v = \beta\}$ gilt $F = \operatorname{conv} V_F + \operatorname{ccone} U_F$.

(b) Ist F nicht leer, so gilt

(b1) $\operatorname{char} F = \operatorname{char} P \cap H^=(a, 0)$ (= die von $a^\top x \leq 0$ definierte Seite von $\operatorname{char} P$)

(b2) $\operatorname{lineal} F = \operatorname{lineal} P$

(Das ist Satz 6.17 der Vorlesung.)

Lösung:

(a). Wegen $F \supset V_F$ ist die Inklusion $F \supset \operatorname{conv} V_F + \operatorname{ccone} U_F$ ist klar. Sei $x \in F$. Aufgrund der Bedingung an P gibt es $v_1, \ldots, v_r$ und $u_1, \ldots, u_s$ sowie $\lambda_1, \ldots, \lambda_r \geq 0$ und $\mu_1, \ldots, \mu_s \geq 0$ mit $\sum_j \lambda = 1$ und $x = \sum_j \lambda_j v_j + \sum_j \mu_j u_j$. Wir zeigen, dass die v_j in V_F und die u_j in U_F liegen. Es gilt

$$\beta = a^\top x = \sum_j \lambda_j a^\top v_j + \sum_j \mu_j a^\top u_j.$$

Da jedes der $a^\top v_j$ höchstens β ist, und jedes der $a^\top u_j \leq 0$, die gewichtete Summe aber gleich β, kann keine der Ungleichungen $a^\top v_j \leq \beta$ oder $a^\top u_j \leq 0$ strikt sein. Also sind die $v_j \in V_F$ und die $u_j \in U_F$.

(b1). Nach Satz 6.6 gilt $\operatorname{char} F = \operatorname{ccone} U_F$, und es bleibt zu zeigen, dass $\operatorname{ccone} U_F = \{x \in \operatorname{ccone} U \mid a^\top x = 0\}$. Dies ist aber ein Spezialfall von (a) mit $V := \{0\}$.

(b2). Offensichtlich gilt $\operatorname{lineal} F \subseteq \operatorname{lineal} P$. Mit Satz 6.11 (oder durch eine sehr kurze Überlegung) folgt aus $y \in \operatorname{lineal} P$, dass $a^\top y = 0$ für *jede* lineare Ungleichung $a^\top x \leq \beta$ gilt, die zulässig für P ist, insbesondere auch für die gegebene. Somit ist für jedes $\lambda \in \mathbb{R}$ der Punkt $x + \lambda y$ in P und erfült $a^\top(x + \lambda y) = \beta$, ist also in F. Damit ist $y \in \operatorname{lineal} F$ gezeigt.