Technische Universität Darmstadt
Fachbereich Mathematik
Prof. Dr. S. Ulbrich

Wintersemester 2010/2011

Blatt 4

Übungen zur Vorlesung Innere-Punkte-Verfahren der konvexen Optimierung: Lösungsvorschlag

G10.

Sei (x, u, y) primal-dual zulässig. Dann gilt wegen $Qx + c = A^T u + y$ und $Ax = b$

$$q(x) - d(x, u, y) = c^T x + x^T Qx - b^T u = x^T A^T u + x^T y - b^T u = x^T y \geq 0.$$

Gilt $q(x) - d(x, u, y) = 0$, dann ist wie eben gezeigt $x^T y = 0$. Also gilt

$$Ax = b, \quad Qx + c - A^T u - y = 0, \quad x \geq 0, \quad y \geq 0, \quad x^T y = 0. \tag{*}$$

Dies sind die KKT-Bedingungen für (CQP) und diese sind wegen der Konvexität von (CQP) hinreichend für die primale Optimalität von x.

(*) sind auch die KKT-Bedingungen für (DQP) (wir schreiben (DQP) als Minimierungsproblem mit der negativen Zielfunktion (DQPm)), denn die KKT-Bedingungen für (DQPm) sind mit Multiplikatoren v, z

$$\begin{aligned} Qx + c - A^T u - y &= 0, \quad y \geq 0, \\ Qx - Q^T v &= 0, \\ -b + A^T v &= 0, \\ 0 + v - z &= 0, \quad z \geq 0, \\ y^T z &= 0, \end{aligned}$$

und mit der Wahl $v = z := x$ ist dies genau (*). Also ist (x, u, y) auch ein KKT-Punkt von (DQPm) und die KKT-Bedingungen sind hinreichend für das konvexe Problem (DQPm). Daher ist (x, u, y) optimal für (DQP).

Sei schließlich $f^* = q(\bar{x})$ der Optimalwert von (CQP). Dann gelten mit geeigneten $(\bar{x}, \bar{u}, \bar{y})$ die KKT-Bedinungen. $(\bar{x}, \bar{u}, \bar{y})$ ist also wie eben gezeigt primal-dual optimal mit $q(\bar{x}) = d(\bar{x}, \bar{u}, \bar{y})$. Also gilt wegen der primal-dualen Optimalität

$$q(x) \geq q(\bar{x}) = d(\bar{x}, \bar{u}, \bar{y}) \geq d(x, u, y),$$

also

$$0 \leq q(x) - q(\bar{x}) \leq q(x) - d(x, u, y) = x^T y.$$

G11.

zu a): Setze $G(x, u, y, \mu) := F_\mu(x, u, y)$. Dann gilt

$$G(x(\mu), u(\mu), y(\mu), \mu) = 0.$$

Nach Vorlesung ist der Satz über implizite Funktionen anwendbar und liefert die Differenzierbarkeit von $\mu \mapsto (x(\mu), u(\mu), y(\mu))$, wobei

$$G'_{(x,u,y)}(x(\mu), u(\mu), y(\mu)) \begin{pmatrix} x'(\mu) \\ u'(\mu) \\ y'(\mu) \end{pmatrix} + G'_\mu(x(\mu), u(\mu), y(\mu)) = 0.$$

Berechnen der Ableitungen liefert

$$\begin{pmatrix} 0 & -A^T & -I \\ A & 0 & 0 \\ Y(\mu) & 0 & X(\mu) \end{pmatrix} \begin{pmatrix} x'(\mu) \\ u'(\mu) \\ y'(\mu) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix}.$$

zu b): Multiplikation der letzten Zeile mit $X(\mu)^{-1/2}Y(\mu)^{-1/2}$ liefert

$$D^{-1}x'(\mu) + Dy'(\mu) = X(\mu)^{-1/2}Y(\mu)^{-1/2}e = \frac{1}{\sqrt{\mu}}e,$$

wobei wir $x_i(\mu)y_i(\mu) = \mu$ im letzten Schritt genutzt haben.

zu c): Es gilt folgt unmittelbar

$$\|D^{-1}x'(\mu) + Dy'(\mu)\|_2^2 = \|\frac{1}{\sqrt{\mu}}e\|_2^2 = \frac{1}{\mu}\|e\|_2^2 = \frac{n}{\mu}$$

Andererseits haben wir

$$\begin{aligned} \|D^{-1}x'(\mu) + Dy'(\mu)\|_2^2 &= (D^{-1}x'(\mu) + Dy'(\mu))^T(D^{-1}x'(\mu) + Dy'(\mu)) \\ &= \|D^{-1}x'(\mu)\|_2^2 + \|Dy'(\mu)\|_2^2 + 2(D^{-1}x'(\mu))^T(Dy'(\mu)) \\ &= \|D^{-1}x'(\mu)\|_2^2 + \|Dy'(\mu)\|_2^2 + 2x'(\mu)^T y'(\mu). \end{aligned}$$

Weiter gilt $x'(\mu)^T y'(\mu) = 0$, denn die zweite Zeile liefert $Ax'(\mu) = 0$ und somit liefert $x'(\mu)^T \cdot$ 1. Zeile

$$0 = -x'(\mu)^T A u'(\mu) - x'(\mu)^T y'(\mu) = 0 - x'(\mu)^T y'(\mu).$$

Damit:

$$\|D^{-1}x'(\mu) + Dy'(\mu)\|_2^2 = \|D^{-1}x'(\mu)\|_2^2 + \|Dy'(\mu)\|_2^2 = \frac{n}{\mu}.$$

zu d): Offensichtlich gilt

$$\frac{y_i(\mu)}{x_i(\mu)}x_i'(\mu)^2 \le \sum_{i=1}^n \frac{y_i(\mu)}{x_i(\mu)}x_i'(\mu)^2 = \|D^{-1}x'(\mu)\|_2^2 \le \frac{n}{\mu}$$

und somit

$$x_i'(\mu)^2 \le \frac{x_i(\mu)}{y_i(\mu)}\frac{n}{\mu} = \frac{x_i(\mu)^2}{x_i(\mu)y_i(\mu)}\frac{n}{\mu} = \frac{x_i(\mu)^2}{\mu^2}n.$$

Wegen $x_i(\mu) \le C_1\mu$ für $i \in N$ haben wir

$$x_i'(\mu)^2 \le \frac{C_1^2\mu^2}{\mu^2}n = C_1^2 n.$$

Für $i \in B$ wissen wir lediglich $x_i(\mu) \le C$ und erhalten

$$x_i'(\mu)^2 \le \frac{C^2}{\mu^2}n.$$

Die Abschätzungen für $y_i'(\mu)$ verlaufen völlig analog mit Verwendung des Terms $\|Dy'(\mu)\|_2^2$.

zu e): Sei nun $|N| = n - m$ und $(A^T, (e_i)_{i\in N}$ nichtsingulr. es gilt

$$Ax'(\mu) = A_B x_B'(\mu) + A_N x_N'(\mu) = 0$$

mit Aufspaltung in Teilmatrizen bzw. -vektoren für die Indexmengen B, N. Umsortieren der Zeilen von $(A^T, (e_i)_{i\in N}$ liefert die Matrix $\left(\begin{smallmatrix} A_B & 0 \\ A_N & I \end{smallmatrix}\right)$ und somit ist der Block A_B invertierbar. Dies ergibt

$$x_B'(\mu) = -A_B^{-1} A_N x_N'(\mu)$$

und somit wegen $|x_i'(\mu)| \leq C_1\sqrt{n}$, $i \in N$,

$$\|x_B'(\mu)\| \leq \|A_B^{-1}A_N\|\|x_N'(\mu)\| \leq \|A_B^{-1}A_N\|\sqrt{n}C_1\sqrt{n}.$$

zu f): Es gilt für alle $0 < \mu < \nu$

$$\|x(\nu) - x(\mu)\| = \left\|\int_\mu^\nu x'(\mu)\,dx\right\| \leq \int_\mu^\nu \|x'(\mu)\|\,dx \leq |\nu - \mu|\kappa C_1 n.$$

Damit läst sich $x(\mu)$ stetig in $\mu = 0$ fortsetzen.