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

G1. Strikt zulässige Punkte

zu a): $S^\circ \subset \mathring{S}$:

Da $h : \mathbb{R}^n \to \mathbb{R}^l$ stetig ist, ist S° offensichtlich offen. Zudem gilt $S^\circ \subset S$, also insgesamt $S^\circ \subset \mathring{S}$.

$S^\circ \supset \mathring{S}$:

Nach Voraussetzung existiert ein $y \in S^\circ$, es gilt also $h_i(y) < 0$, $1 \le i \le l$.

Sei $x \in \mathring{S}$ beliebig. Wir zeigen $x \in S^\circ$. Da $\mathring{S}$ offen ist, existiert eine offene Kugel $B_\varepsilon(x) \subset \mathring{S}$. Daher gilt für $\delta > 0$ klein genug auch $z = y + (1+\delta)(x-y) \in B_\varepsilon(x) \subset \mathring{S}$, also $h(z) \le 0$.

Nun ist

$$x = \frac{1}{1+\delta} z + \left(1 - \frac{1}{1+\delta}\right) y$$

und daher wegen h_i konvex mit $h_i(y) < 0$, $h_i(z) \le 0$

$$h_i(x) \le \frac{1}{1+\delta} h_i(z) + \left(1 - \frac{1}{1+\delta}\right) h_i(y) < 0.$$

Dies zeigt $x \in S^\circ$ und wegen $x \in \mathring{S}$ beliebig $S^\circ \supset \mathring{S}$.

G2.

zu a):

$$\min \varphi(x;\mu) := \frac{x_1^2}{2} + x_2 - \mu(\ln(x_1) + \ln(x_2)) \quad \text{u.d. NB} \quad x_1 > 0,\ \ x_2 > 0. \qquad (\mathrm{BP}_\mu)$$

$$\nabla\varphi(x;\mu) = \begin{pmatrix} x_1 \\ 1 \end{pmatrix} - \mu \begin{pmatrix} 1/x_1 \\ 1/x_2 \end{pmatrix},$$
$$\nabla^2\varphi(x;\mu) = \begin{pmatrix} 1 + \frac{\mu}{x_1^2} & 0 \\ 0 & 1 + \frac{\mu}{x_2^2} \end{pmatrix}.$$

zu b): $\nabla^2\varphi(x;\mu)$ is positiv definit auf $]0,\infty[^2$. Also ist $\varphi(\cdot;\mu)$ streng konvex und daher höchstens ein lokales (=globales) Minimum $x(\mu)$. Jedes lokale Minimum x^* von $\varphi(\cdot;\mu)$ erfüllt

$$\nabla\varphi(x^*;\mu) = \begin{pmatrix} x_1^* \\ 1 \end{pmatrix} - \mu \begin{pmatrix} 1/x_1^* \\ 1/x_2^* \end{pmatrix} = 0,$$

also $x(\mu) = x^* = (\sqrt{\mu}, \mu)^T = (\tau, \tau^2)$ mit $\tau = \sqrt{\mu}$.

Skizze: Der zentrale Pfad ist der rechte Ast einer Normalparabel).

zu c): Es gilt $\lim_{\mu \searrow 0} x(\mu) = (0,0)^T$. Offensichtlich ist $\bar{x} = (0,0)^T$ Lösung von (QP), da die Zielfunktion von (QP) auf dem zulässigen Bereich durch 0 nach unten beschränkt ist und $\bar{x} = (0,0)^T$ den Zielfunktionswert 0 liefert.

zu d): Dann gilt

$$\varphi(x;\mu) = \frac{x_1^2}{2} + x_2 - \mu(m\ln(x_1) + \ln(x_2)),$$
$$\nabla\varphi(x;\mu) = \begin{pmatrix} x_1 \\ 1 \end{pmatrix} - \mu \begin{pmatrix} 1/x_1 \\ m/x_2 \end{pmatrix}.$$

Also liefert

$$\nabla\varphi(x^*;\mu) = \begin{pmatrix} x_1^* \\ 1 \end{pmatrix} - \mu \begin{pmatrix} 1/x_1^* \\ m/x_2^* \end{pmatrix} = 0$$

$x(\mu) = x^* = (\sqrt{\mu}, m\mu)^T = (\tau, m\tau^2)$ mit $\tau = \sqrt{\mu}$. Für grosse m verläuft der zentrale Pfad also immer näher entlang der x_2-Achse.

G3. Konvexe QPs als Second Order Cone Problem

Setze $q(x) := \frac{1}{2}x^TQx + c^Tx + \gamma$. Mit $\bar{x} = -Q^{-1}c$ gilt $\nabla q(\bar{x}) = Q\bar{x} + c = 0$ und daher liefert Taylorentwicklung in $\bar{x}$

$$q(x) := \frac{1}{2}x^TQx + c^Tx + \gamma = q(\bar{x}) + \nabla q(\bar{x})^T(x-\bar{x}) + \frac{1}{2}(x-\bar{x})^TQ(x-\bar{x}) = q(\bar{x}) + \frac{1}{2}(x-\bar{x})^TQ(x-\bar{x}).$$

Nun existiert $R \in \mathbb{R}^{n,n}$ mit $Q = R^TR$. Also gilt

$$(x-\bar{x})^TQ(x-\bar{x}) = (x-\bar{x})^TR^TR(x-\bar{x}) = \|R(x-\bar{x})\|^2.$$

Dies ergibt $q(x) = q(\bar{x}) + \frac{1}{2}\|R(x-\bar{x})\|^2$ und somit

$$q(x) \le 0 \quad \Longleftrightarrow \quad \|R(x-\bar{x})\|^2 \le -2q(\bar{x}) \quad \Longleftrightarrow \quad \|R(x-\bar{x})\| \le \sqrt{-2q(\bar{x})}.$$

Die rechte Ungleichung ist eine Second Order Cone Nebenbedingung mit $C = R$, $a = -R\bar{x}$, $d = 0$, $e = \sqrt{-2q(\bar{x})}$.