Aufgabe S. 2.1

Gegeben sei das Optimierungsproblem

$$P: \quad \min_{w \in \mathbb{R}^m} \frac{1}{2}\|w\|_2^2 \text{ s.t. } Aw \geq e$$

mit einer (d, m)-Matrix A und

$$e := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^d.$$

Gehen Sie im Folgenden davon aus, dass ein $\tilde{w}$ existiert, für das gilt

$$A\tilde{w} \geq e.$$

a) Zeigen Sie, dass P lösbar ist.

Beweis: Nach Korollar 1.2.40 gilt:

Es sei M nicht-leer, und $f\colon M \to R$ sei stetig und koerziv auf M.

Dann ist S nicht-leer und kompakt.

Nach Aufgabenstellung ist mit $f(w) = \frac{1}{2}\|w\|_2^2$ und $g(w) = e - Aw$, sodass

$$M = \{w \in \mathbb{R}^m \colon g(w) \leq 0\}$$

und damit:

$$P: \quad \min_{w \in \mathbb{R}^m} f(w) \text{ s.t. } w \in M.$$

Nach Voraussetzung existiert ein $\tilde{w} \in M$, sodass M nicht-leer ist. Weiter ist die Funktion f stetig, da für $\|x - y\|_2 < \delta$ und $\epsilon := \frac{\delta^2}{2}$ mit der umgekehrten Dreiecksungleichung gilt:

$$|f(x) - f(y)| = \frac{1}{2}\left|\|x\|_2^2 - \|y\|_2^2\right| \leq \frac{1}{2}\|x - y\|_2^2 \leq \epsilon.$$

Behauptung: die Menge M ist abgeschlossen bzw. das Komplement

$$\mathbb{R}^m \setminus M = \{w \in R^m \colon Aw < e\}.$$

ist offen. Betrachte hierfür $w \in \mathbb{R}^m \setminus M$, dann gilt mit dem Radius

$$\epsilon := \min_{i=1,\dots,d} \frac{1}{2} |(Aw - e)_i| > 0,$$

dass alle

$$x \in B_\epsilon(w) = \{x \in \mathbb{R}^m \colon \|w - x\| < \epsilon\}$$

in M liegen, da für alle $i = 1, \dots, d$

$$|(e - Ax)_i| \geq \Big| |(e - Aw)_i| - |(Aw - Ax)_i| \Big| > \frac{2\epsilon - \epsilon}{2} = \frac{\epsilon}{2},$$

womit $Ax < e$ bzw. $x \in \mathbb{R}^m \setminus M$. Somit ist M abgeschlossen, woraus folgt, dass die Funktion f koerziv ist, denn nach Definition 1.2.37 müssen nur alle Folgen $(x^\nu) \subseteq M$ mit $\|x^\nu\| \to \infty$ betrachtet werden und für die gilt mit der Äquivalenz der Normen im $\mathbb{R}^m$:

$$f(x^\nu) = \frac{1}{2}\|x^\nu\|_2^2 \longrightarrow +\infty.$$

Damit sind alle Voraussetzungen von Korollar 1.2.40 erfüllt, S ist somit nicht-leer und kompakt und somit P lösbar. □

b) Zeigen Sie, dass P ein konvexes Optimierungsproblem ist.

Beweis: Nach Beispiel 2.1.9 gilt:
Falls f, g_i für $i \in I$ auf R^m konvexe Funktionen sind, dann ist

$$P: \quad \min f(x) \text{ s.t. } g_i(x) \leq 0, \; i \in I$$

ein konvexes Optimierungsproblem.

Beim gegebenen Problem ist $g_i(w) = -(Aw)_i + 1$ für alle $i \in \{1, \dots, d\}$ eine lineare und somit konvexe Funktion. Da $f \in C^2(\mathbb{R}^m)$ mit $f(w) = \frac{1}{2}\sum_{i=1,\dots,m} w_i^2$ folgt $\nabla_w f(w) = w$ und

$$D^2 f = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \succ 0,$$

womit f insbesondere konvex und damit P ein konvexes Optimierungsproblem ist. □

c) Stellen Sei das Wolfe-Dual D zu P auf.

Beweis: Die Lagrange Funktion zum (nach b)) restringierten, konvexen Optimierungsproblem P lautet:

$$L(w,\lambda) = f(w) + \sum_{i\in I} \lambda_i g_i(w) = \frac{1}{2}\|w\|_2^2 + \lambda^T(e - Aw)$$

mit $\lambda = (\lambda_1, \ldots, \lambda_d)^T$ heißt. Das Dualproblem lässt sich dadurch folgendermaßen umformulieren:

$$D: \quad \max_{w,\lambda} L(w,\lambda) \text{ s.t. } \nabla_w L(w,\lambda) = 0,\ \lambda \geq 0,$$

mit der zulässigen Menge

$$M_D = \left\{(w,\lambda) \in \mathbb{R}^m \times \mathbb{R}^d \colon \nabla_w L(w,\lambda) = 0, \lambda \geq 0\right\}.$$

Dabei ist

$$\begin{aligned}
\nabla_w L(w,\lambda) &= w - \nabla_w \left(\sum_{i=1,\ldots,d} \lambda_i \sum_{k=1,\ldots,m} A_{ik} w_k\right) \\
&= w - \nabla_w \left(\sum_{k=1,\ldots,m} \sum_{i=1,\ldots,d} \lambda_i A_{ik} w_k\right) \\
&= \begin{pmatrix} w_1 - \sum_{i=1,\ldots,d} \lambda_i A_{i1} \\ \vdots \\ w_m - \sum_{i=1,\ldots,d} \lambda_i A_{im} \end{pmatrix}
\end{aligned}$$

□

d) Formulieren Sie, ausgehend von Aufgabenteil b), ein duales Optimierungsproblem, das nur von der Dualvariabel λ abhängt.

Beweis: Aus der Forderung aus c):

$$\begin{pmatrix} w_1 - \sum_{i=1,\ldots,d} \lambda_i A_{i1} \\ \vdots \\ w_m - \sum_{i=1,\ldots,d} \lambda_i A_{im} \end{pmatrix} = 0$$

folgt $w_1 - \sum_{i=1,\ldots,d} \lambda_i A_{i1} = 0,\ \ldots,\ w_m - \sum_{i=1,\ldots,d} \lambda_i A_{im} = 0$ und damit ist

$$w_1 = \sum_{i=1,\ldots,d} \lambda_i A_{i1},\ \ldots,\ w_m = \sum_{i=1,\ldots,d} \lambda_i A_{im}.$$

Somit reduziert sich das Dualproblem zu

$$\begin{aligned} D_{w_{red}}: \quad & \max_{\lambda \in \mathbb{R}^d} \sum_{i=1,\ldots,m} \left(\sum_{j=1,\ldots,d} \lambda_j A_{ji} \right)^2 + \sum_{i=1,\ldots,d} \lambda_i \left(1 - \left(\sum_{k=1,\ldots,m} A_{ik} \sum_{j=1,\ldots,d} \lambda_j A_{jk} \right) \right) \\ & \iff \max_{\lambda \in \mathbb{R}^d} \|\lambda^T A\|_2^2 + \lambda^T \left(e - AA^T \lambda \right) \end{aligned}$$

$\square$