

**Aufgabe S. 2.1** (2 + 2 + 3 + 3)

Gegeben sei das Optimierungsproblem

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

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  $f: M \rightarrow \mathbb{R}$  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

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

und damit:

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

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:

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

$M$  ist aufgrund der Stetigkeit der Restriktionsfunktion als untere Niveaumenge abgeschlossen<sup>1</sup> bzw. falls man anstatt  $g$  komponentenweise  $g_i$  als Restriktion betrachtet, so ist  $M$  als Schnitt abgeschlossener Mengen ebenfalls abgeschlossen:

$$M = \bigcap_{i \in I} \{w \in \mathbb{R}^m : g_i(w) \leq 0\}.$$

Aus der Abgeschlossenheit der Menge  $M$  folgt, dass die Funktion  $f$  koerziv ist, denn nach Definition 1.2.37 müssen nur alle Folgen  $(x^\nu) \subseteq M$  mit  $\|x^\nu\| \rightarrow \infty$  betrachtet werden und für die gilt mittels 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.  $\square$

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  $\mathbb{R}^m$  konvexe Funktionen sind, dann ist

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

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

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

womit nach 2.5.10  $f$  insbesondere konvex und  $P$  ein konvexes Optimierungsproblem ist.  $\square$

---

<sup>1</sup>Dies wurde in der Übung (o. B.) kurz angesprochen; deshalb ist ein kurzer Beweis hinten angehängt

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

*Beweis:* Die Lagrange Funktion zum restringierten, konvexen Optimierungsproblem  $P$  lautet:

$$\begin{aligned} L(w, \lambda) &= f(w) + \sum_{i \in I} \lambda_i g_i(w) = \frac{1}{2} \|w\|_2^2 + \lambda^T (e - Aw) \\ &= \frac{1}{2} \sum_{i=1}^m w_i^2 + \sum_{i=1}^d \lambda_i (1 - a_i \cdot w) \end{aligned}$$

mit  $a_i$  i-ter Zeilenvektor von  $A$  und  $\lambda = (\lambda_1, \dots, \lambda_d)^T$ . Das Dualproblem lässt sich dadurch folgendermaßen umformulieren:

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

mit der zulässigen Menge

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

Dabei ist

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

Zusammengefasst also:

$$\begin{aligned} D : \quad & \max_{(w, \lambda) \in \mathbb{R}^m \times \mathbb{R}^d} \frac{1}{2} \sum_{i=1}^m w_i^2 + \sum_{i=1}^d \lambda_i (1 - a_i \cdot w) \\ \text{s.t.} \quad & \begin{pmatrix} w_1 - \sum_{i=1, \dots, d} \lambda_i A_{i1} \\ \vdots \\ w_m - \sum_{i=1, \dots, d} \lambda_i A_{im} \end{pmatrix} = 0 \end{aligned}$$

□

- d) Formulieren Sie, ausgehend von Aufgabenteil b), ein duales Optimierungsproblem, das nur von der Dualvariabel  $\lambda$  abhängt.

*Beweis:* Aus der Forderung aus c)

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

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

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

Somit lassen sich die  $w_i$  in  $D$  ersetzen und das Dualproblem reduziert sich zu

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

□

### Aufgabe S. 2.2 (3 + 4 + 5)

Gegeben sei das Optimierungsproblem

$$P : \min_{x \in \mathbb{R}^2} x_1^2 + 2x_2^2 + x_1 \text{ s.t. } x_1 + x_2 \leq \alpha$$

für einen Parameter  $\alpha \in \mathbb{R}$ .

- a) Zeigen Sie, ohne  $P$  explizit zu lösen, dass  $P$  einen **eindeutigen** globalen Minimalpunkt besitzt.

*Beweis:*  $M$  ist trivialerweise nicht-leer, da die Menge nur durch eine lineare Funktion

$$g(x) = x_1 + x_2 - \alpha \leq 0$$

beschränkt ist. Außerdem ist  $M$  aufgrund der Stetigkeit (analog zu 1. a) der Restriktionsfunktion als untere Niveaumenge von  $g$  abgeschlossen<sup>2</sup>. Da lineare Funktionen insbesondere konvex sind, ist nach Übung 4.4 die Menge  $M$  konvex. Als Polynom ist  $f \in C^2(\mathbb{R}^2)$  und Ableiten der Zielfunktion ergibt:

$$\nabla f(x) = \begin{pmatrix} 2x_1 + 1 \\ 4x_2 \end{pmatrix} \text{ und damit } D^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \succ 0.$$

Nach 2.5.10 ist  $f$  somit gleichmäßig konvex. Die eindeutige Lösbarkeit folgt nun mit Satz 2.3.3 c).  $\square$

- b) Lösen Sie  $P$  mit Hilfe der KKT-Bedingungen  
(*Hinweis: die Lösung ist abhängig von  $\alpha$* ).

*Beweis:* Nach Definition ist 2.6.10 gilt:

Für ein  $C^1$ -Problem  $P$  heißt ein Punkt  $x \in \mathbb{R}^n$  KKT-Punkt mit Multiplikatoren  $\lambda$

und  $\mu$  falls folgendes System von Gleichungen und Ungleichungen erfüllt ist:

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0, \quad \bar{\lambda}_i g_i(\bar{x}) = 0, \quad \bar{\lambda}_i \geq 0, \quad g_i(\bar{x}) \leq 0, \quad h_j(\bar{x}) = 0$$

---

<sup>2</sup>vgl 1. a kurz bzw. kurzer Beweis im Anhang

für  $i \in I$  und  $j \in J$ . Nach Satz 2.6.12 gilt außerdem:

$P$  sei konvex und  $C^1$ , und  $x$  sei KKT-Punkt (mit Multiplikatoren  $\lambda, \mu$ ).

Dann ist  $x$  globaler Minimalpunkt von  $P$ .

Da  $M$  nur durch eine Ungleichung beschrieben wird, lässt sich die Lagrange-Funktion schreiben als:

$$L(x, \lambda) = x_1^2 + 2x_2^2 + x_1 + \lambda(x_1 + x_2 - \alpha).$$

Damit ergibt sich für die erste KKT-Bedingung:

$$\nabla_x L(x, \lambda) = \begin{pmatrix} 2x_1 + 1 + \lambda \\ 4x_2 + \lambda \end{pmatrix} = 0. \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{-1-\lambda}{2} \\ \frac{-\lambda}{2} \end{pmatrix}.$$

Für die Komplementärbedingungen betrachte man die beiden folgenden Fälle.

**1. Fall**  $I_0(x) = \emptyset \Rightarrow \lambda_1 = 0$ . Einsetzen von  $\lambda_1 = 0$  in die 1. Bedingung ergibt

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{-1}{2} \\ 0 \end{pmatrix}.$$

Nach Konstruktion werden die ersten 3 KKT-Bednningen erfüllt und da  $I_0(x) = \emptyset$  gilt, muss zudem

$$x_1 + x_2 - \alpha < 0$$

erfüllt sein. Einsetzen von  $x_1, x_2$  ergibt  $\frac{-1}{2} < \alpha$ .

**2. Fall**  $I_0(x) = \{1\} \Rightarrow x_1 + x_2 - \alpha = 0$ . Einsetzen von  $x_1, x_2$  ergibt:

$$\frac{-1-\lambda}{2} + \frac{-\lambda}{2} - \alpha = 0.$$

$$\Leftrightarrow \frac{-1}{2} - \frac{3}{4}\lambda = \alpha \Leftrightarrow \lambda = -\frac{4}{3}\alpha - \frac{2}{3}$$

Da zudem  $\lambda \geq 0$  gelten muss, ergibt sich  $\alpha \leq -\frac{1}{2}$ . Einsetzen von  $\lambda$  in die 1. Bedingung ergibt

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{-1}{2} + \frac{4}{6}\alpha + \frac{2}{6} \\ \frac{4}{12}\alpha + \frac{2}{12} \end{pmatrix} = \begin{pmatrix} \frac{2}{3}\alpha - \frac{1}{6} \\ \frac{1}{3}\alpha + \frac{1}{6} \end{pmatrix}.$$

d.h.  $\lambda = -\frac{4}{3}\alpha - \frac{2}{3}$  mit  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3}\alpha - \frac{1}{6} \\ \frac{1}{3}\alpha + \frac{1}{6} \end{pmatrix}$  erfüllt nach Konstruktion alle KKT-Bedingungen.

Unter Verwendung von Teilaufgabe a) und Satz 2.6.12 ist somit für  $\alpha \leq -\frac{1}{2}$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3}\alpha - \frac{1}{6} \\ \frac{1}{3}\alpha + \frac{1}{6} \end{pmatrix} \text{ und für } \alpha > -\frac{1}{2} \text{ ist } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -0.5 \\ 0 \end{pmatrix}$$

globaler Minimalpunkt.

□

- c) Es sei  $v(\alpha)$  die Optimalwertfunktion von  $P$  abhängig vom Parameter  $\alpha$ . Schreiben Sie  $v$  explizit und zeigen Sie, dass  $v$  eine konvexe Funktion ist.

*Beweis:* Setzt man die obige Lösung nun in die Zielfunktion ein, so ergibt sich folgende von  $\alpha$  abhängige Funktion für den Optimalwert von P:

$$\begin{aligned} v(\alpha) &= \left(\frac{2}{3}\alpha - \frac{1}{6}\right)^2 + 2 \cdot \left(\frac{1}{3}\alpha + \frac{1}{6}\right)^2 + \frac{2}{3}\alpha - \frac{1}{6} \\ &= \frac{2}{3}\alpha^2 + \frac{2}{3}\alpha - \frac{1}{12} \\ &= \frac{2}{3}(\alpha^2 + \alpha) - \frac{1}{6} \\ &= \frac{2}{3}\left(\alpha + \frac{1}{2}\right)^2 - \frac{3}{8}, \end{aligned}$$

wobei  $\alpha \in M = \{x \in \mathbb{R} | x \leq 0.5\}$ . Als Parabel ist  $v(\alpha)$  insbesondere konvex, denn  $f''(x) = \frac{4}{3} > 0, \forall x \in M$ . □

### Aufgabe S. 2.3 (13 + 4)

- a) Implementieren Sie das Schnittebenenverfahren von Kelley (vgl. Algorithmus 2.1 im Skript), das eine Approximation eines globalen Minimalpunktes eines konvexen Optimierungsproblems

$$P : \min_{x \in \mathbb{R}^n} c^T x \text{ s.t. } g_i(x) \leq 0, \ i \in I, \ Ax \leq b$$

mit der nicht-leeren kompakten Mengen  $M^0 := \{x \in \mathbb{R}^n | Ax \leq b\}$  berechnet. Die in jeder Iteration auftretenden linearen Optimierungsprobleme sollen mit dem Solver Gurobi gelöst werden. Ihr Algorithmus erhält folgende Inputs:

- Das Gurobi-Modell zu Hilfsproblemen

$$LP^0 : \min_{x \in \mathbb{R}^n} c^T x \text{ s.t. } x \in M^0.$$

- Die Restriktionsfunktionen  $g_1(x), \dots, g_p(x)$ .
- Die Gradienten (bzw. Ableitungen) der Restriktionsfunktionen.
- Eine Toleranz  $\epsilon$ .

- b) Testen Sie Ihren Algorithmus an dem Optimierungsproblem

$$P : \min_{x \in \mathbb{R}^2} x_2 \text{ s.t.}$$

$$x_1^2 + x_2^2 \leq 1, \ e^{x_1} - x_2 \leq 0, \ x_1 - x_2 \leq 0, \ x_1 \in [-1, -\frac{1}{2}], \ x_2 \in [-1, 1]$$

für  $\epsilon := 10^{-7}$ .



## Anhang

**Theorem:** Die Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  sei stetig. Dann sind die Mengen

$$f_{\leq}^{\alpha} = \{x \in \mathbb{R}^n: f(x) \leq \alpha\}$$

für alle  $\alpha \in \mathbb{R}$  abgeschlossen.

*Beweis:* Sei  $(x^{\nu}) \subseteq \mathbb{R}^n$  eine Folge in  $f_{\leq}^{\alpha}$ . Weiterhin gelte

$$\lim_{\nu \rightarrow \infty} f(x^{\nu}) =: x.$$

Zu zeigen ist, dass  $x \in f_{\leq}^{\alpha}$ , woraus die Abgeschlossenheit folgt. Es gilt:

$$f(x^{\nu}) \leq \alpha \quad \forall \nu \in N.$$

Aufgrund der Stetigkeit der Funktion  $f$  folgt dann:

$$f(x) = f(\lim_{\nu \rightarrow \infty} x^{\nu}) = \lim_{\nu \rightarrow \infty} f(x^{\nu}) \leq \alpha,$$

also  $f(x) \leq \alpha$  bzw.  $x \in f_{\leq}^{\alpha}$ . □