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

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:

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

M ist aufgrund der Stetigkeit der Restriktionsfunktion als untere Niveaumenge abgeschlosse[1] bzw. falls man anstatt g komponentenweise g_i als Restriktion betrachtet, so ist M als Schnitt abgeschlossener Mengen ebenfalls abgeschlossen:

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

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 $\|x^\nu\| \to \infty$ 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 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 nach 2.5.10 f insbesondere konvex und P ein konvexes Optimierungsproblem ist. $\square$

[1] Dies wurde in der Übung (o. B.) kurz angesprochen; deshalb ist ein kurzer Beweis hinten angehängt

c) Stellen Sei 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 $\lambda = (\lambda_1, \ldots, \lambda_d)^T$. 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

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

Zusammengefasst also:

$$D: \quad \max_{(w,\lambda)\in\mathbb{R}^m\times\mathbb{R}^d} \frac{1}{2}\sum_{i=1}^m w_i^2 + \sum_{i=1}^d \lambda_i (1 - a_i \cdot w) \text{ s.t. } \lambda \geq 0,$$

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

□

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

Beweis: Aus der Forderung aus c)

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

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

$$P_{red}: \quad \max_{\lambda \in \mathbb{R}^d} \sum_{i=1,\dots,m} \left(\sum_{j=1,\dots,d} \lambda_j A_{ji} \right)^2 + \sum_{i=1,\dots,d} \lambda_i \left(1 - \left(\sum_{k=1,\dots,m} A_{ik} \sum_{j=1,\dots,d} \lambda_j A_{jk} \right) \right) \text{ s.t. } \lambda \geq 0$$

$$\iff \max_{\lambda \in \mathbb{R}^d} \|\lambda^T A\|_2^2 + \lambda^T \left(e - AA^T \lambda \right) \text{ s.t. } \lambda \geq 0$$

$$\iff \max_{\lambda \in \mathbb{R}^d} \|\lambda^T A\|_2^2 + \lambda^T e - \left(\lambda^T A \right) \left(\lambda^T A \right)^T \text{ s.t. } \lambda \geq 0$$

$\square$

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

Gegeben sei das Optimierungsproblem

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

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 S. 2.1 a) der Restriktionsfunktion als untere Niveaumenge von g abgeschlossen[2]. 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 α*).

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 λ

und μ falls folgendes System von Gleichungen und Ungleichungen erfüllt ist:

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

[2] vgl. S. 2.1 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 λ, μ).

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:

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

Für die Komplementärbedinungen 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-Bedninungen 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:

$$\frac{-1-\lambda}{2} + \frac{-\lambda}{4} - \alpha = 0.$$

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

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

$\square$

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

Beweis: Setzt man die obige Lösung für $\alpha \leq -0.5$ nun in die Zielfunktion ein, so ergibt sich folgende von α abhängige Funktion für den Optimalwert von P:

$$\begin{aligned} v(\alpha) &= \left(\frac{2}{3}\alpha - \frac{1}{6}\right)^2 + 2 \cdot \left(\frac{1}{3}\alpha + \frac{1}{6}\right)^2 + \frac{2}{3}\alpha - \frac{1}{6} \\ &= \frac{2}{3}\alpha^2 + \frac{2}{3}\alpha - \frac{1}{12} \\ &= \frac{2}{3}\left(\alpha^2 + \alpha\right) - \frac{1}{12} \\ &= \frac{2}{3}\left(\alpha + \frac{1}{2}\right)^2 - \frac{1}{4}. \end{aligned}$$

v ist somit für $\alpha \leq -0.5$ als Parabel insbesondere konvex, denn $v''(\alpha) = \frac{4}{3} > 0$. Für $\alpha > -0.5$ ist v konstant und damit auch konvex. v ist außerdem stetig, da

$$\lim_{\alpha \to -0.5_-} v(\alpha) = \lim_{\alpha \to -0.5_-} \frac{2}{3}\left(\alpha + \frac{1}{2}\right)^2 - \frac{1}{4} = -0.25 = (-0.5)^2 - 0.5 = \lim_{\alpha \to -0.5_+} v(\alpha).$$

v ist als Komposition stetig differenzierbarer Funktionen differenzierbar und es gilt

$$v'(\alpha) = \begin{cases} \frac{4}{3}\alpha + \frac{2}{3}, & \alpha \leq -0.5 \\ 0, & \alpha > -0.5 \end{cases},$$

also

$$\lim_{\alpha \to -0.5_-} v'(\alpha) = \lim_{\alpha \to -0.5_-} \frac{4}{3}\alpha + \frac{2}{3} = 0 = \lim_{\alpha \to -0.5_+} v'(\alpha),$$

womit v stetig differenzierbar ist. Nun ist v' wegen

$$v''(\alpha) = \frac{4}{3} > 0 \text{ für } \alpha \leq -0.5 \text{ und } v''(\alpha) = 0 \text{ für } \alpha > -0.5$$

auf der konvexen Menge $\mathbb{R}$ monoton und v nach Satz 2.5.16 konvex. □

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

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

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

Beweis:

□

```
%%
% Nadine Kostorz (1972005)
% Martin Belica  (1775706)

function x_lsg = kelley_opt(linear_model, g, jacobi_g, eps)

% lineares Modell
lp     = linear_model;
result = gurobi(lp);

% Anfangsparameter
j          = 0;
x          = result.x;
val        = g(x);
[g_max, k] = max(val);

while g_max > eps
    jacobi_val = jacobi_g(x);

    b       = -val(k) - dot(jacobi_val(k, :),-x);
    fmatrix = full(lp.A);
    fmatrix(end+1:end+size(jacobi_val(k, :), 1), :) = jacobi_val(k, :);
    j = j+1;

    % Model-Update
    lp.A            = sparse(fmatrix);
    lp.rhs(end+1)   = b;
    lp.sense(end+1) = ['<'];

    result = gurobi(lp);

    x = result.x;
    val = g(result.x);
    [g_max, k] = max(val);
end

x_lsg = x;

end
```

b) Testen Sie Ihren Algorithmus an dem Optimierungsproblem für $\epsilon := 10^{-7}$

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

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

Beweis:

□

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Globale Optimierung - Sonderuebung II
%
% Aufgabe S.2.3
% b): Testen Sie Ihren Algorithmus an dem Optimierungsproblem
%
% P: \min_{x \in \R^2} x_2 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] $$
%          fuer $\epsilon \coloneqq 10^{-7}$.
%%
% Nadine Kostorz (1972005)
% Martin Belica  (1775706)

epsilon = 10^(-7);

g   = @(x) [x(1).^2+x(2).^2-1 exp(x(1))-x(2)];
ddg = @(x) [2*x(1) 2*x(2); exp(x(1)) -1];

clear linear_model;
linear_model.A          = sparse([1 -1]);
linear_model.obj        = [0 1];
linear_model.modelsense = 'min';
linear_model.rhs        = [0];
linear_model.sense      = ['<'];
linear_model.lb         = [-1; -1];
linear_model.ub         = [-0.5; 0.9];

x_lsg = kelley_opt(linear_model, g, ddg, epsilon)
```

Anhang

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

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

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

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

$$\lim_{\nu \to \infty} f(x^\nu) =: f(x),$$

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

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

Aufgrund der Stetigkeit der Funktion f folgt dann:

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

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