Optimierung Klausurzettel

muss handschriftlich sein!

Begriffe / Wiederholung

- kompakt = beschränkt und abgeschlossen
- regulär = invertierbar ($\det A \neq 0$)
- Eine symmetrische Matrix $H \in \mathbb{R}^{n \times n}$ heißt positiv definit, falls $y^T H y > 0 \forall y \in \mathbb{R}^n$.
  - äquivalent: det(`H[:i][:i]`) > 0 (Hauptminoren)
- Mittelwertsatz: $\forall x, y \in \mathbb{R}^n \exists \xi \in \{x + t(y - x) \mid t \in (0,1)\} : f(x) - f(y) = \nabla f(\xi)(x - y)$
- Moore-Penrose-Inverse: $A^+ := \left(A^T A\right)^{-1} A^T$, minimiert $\left\|Ax - b^\delta\right\|_2$

Grundlagen

- allgemeines restringiertes Optimierungsproblem: min $f(x)$ u.d.N. $g(x) \leq 0$, $h(x) = 0$
- Lagrange-Funktion: $\mathcal{L}(x, \mu, \lambda) := f(x) + \sum_{i \in I_U} \mu_i g_{i(x)} + \sum_{j \in I_G} \lambda_j h_{j(x)}$
- Weierstrass: Jedes Programm mit nichtleerem, kompaktem zulässigem Bereich und stetiger Zielfunktion hat eine Lösung.
- Konvexe Menge $K \subset V$: $\forall x, y \in K : \{x + t(y - x) \mid t \in [0,1]\} \subset K$
- Sei $K \subset \mathbb{R}^n$ nichtleer, abgeschlossen und konvex. Dann existiert für alle $y \in \mathbb{R}^n$ eine eindeutige Lösung $P_K(y)$ von $\min\|y - x\|$ u.d.N. $x \in K$.
  - $P_K(y) =: x_0$ heißt beste Approximation oder metrische Projektion und ist eindeutig charakterisiert durch $\langle y - x_0, x - x_0\rangle \leq 0 \forall x \in K$.
  - Ist K ein linearer Unterraum, heißt P_K orthogonale Projektion, es gilt $\langle x, P_K(y)\rangle = \langle P_K(x), y\rangle$ und $\langle y - P_K(y), z\rangle = 0 \forall z \in K$.
- Epigraph: $\operatorname{epi}(f) := \{(x, \alpha) \in \mathcal{F} \times \mathbb{R} \mid f(x) \leq \alpha\}$
- konvexe Funktion: $f(tx + (1-t)y) \leq t f(x) + (1-t) f(y) \forall x, y \in \mathcal{F}, t \in [0,1]$. Äquivalent: $\operatorname{epi}(f)$ ist konvex.
  - $\alpha f, f + g, \sup_{i \in \mathcal{J}} f_i$ sind auch konvex ($\alpha > 0$)
  - jedes lokale Minimum x mit $\nabla f(x) = 0$ ist auch globales Minimum von f

Abstiegsverfahren

- Finde Suchrichtung d_k mit $d_k^T g_k < 0$ und Schrittweite $s_k > 0$ mit $f(x_k + s_k d_k) < f(x_k)$. Setze $x_{k+1} = x_k + s_k d_k$
  - Liniensuche: $s_k = \arg\min_{s_k \in \mathbb{R}^+} f(x_k + s_k d_k)$
- Gradientenverfahren: $d_k = -g_k$
- Newtonverfahren: $d_k = -((Hf)(x_k))^{-1} g_k$
- Gauß-Newton-Verfahren: $d_k = -\left[DF(x_k)^T DF(x_k)\right]^{-1} g_k$
- Quasi-Newton-Verfahren: $d_k = -B_k^{-1} \nabla f(x_k) = -H_k \nabla f(x_k)$
  - $y_k := \nabla f(x_{k+1}) - \nabla f(x_k)$, $s_k := x_{k+1} - x_k$, $\gamma_k := \frac{1}{y_k^T s_k}$
  - DFP-Update-Formel: $B_{k+1}^{\text{DFP}} := \left(I - \gamma_k y_k s_k^T\right) B_k^{\text{DFP}} \left(I - \gamma_k s_k y_k^T\right) + \gamma_k y_k y_k^T$
  - BFGS-Update-Formel: $H_{k+1}^{\text{BFGS}} := \left(I - \gamma_k s_k y_k^T\right) H_k^{\text{BFGS}} \left(I - \gamma_k y_k s_k^T\right) + \gamma_k s_k s_k^T$

Lineare Optimierung

- Problem: min $f(x)$ u.d.N. $A_U x \geq b_U$, $A_G x = b_G$
- KKT-Bedingungen:
  - $\nabla f(x_*) = A_G^T \lambda_* + A_U^T \mu_*$
  - $b_U - A_U x_* \leq 0$
  - $b_G - A_G x_* = 0$
  - $\mu_* \geq 0$

- $(A_U x_* - b_U)^T \mu_* = 0$
- Kegel N: $x \in N \implies \lambda x \in N \forall \lambda > 0$
- konvexer Kegel: $\operatorname{cone}(A) := \left\{\sum_j \lambda_j a_j \mid \lambda_j \geq 0, a_j \in A\right\}$
- Lemma von Farkas: Sei $a_i \in \mathbb{R}^n$ und $g \in \mathbb{R}^n$. Dann ist entweder $g \in \operatorname{cone}\{a_i\}$, oder $\exists d \in \mathbb{R}^n : a_i^T d \geq 0 \land g^T d < 0$
  - äquivalent: entweder $A\lambda = g, \lambda \geq 0$ lösbar, oder $A^T d \geq 0, d^T g < 0$ lösbar

Lagrange-Dualität

- Sattelpunkt (x_*, λ_*): $\mathcal{L}(x_*, \lambda) \leq \mathcal{L}(x_*, \lambda_*) \leq \mathcal{L}(x, \lambda_*) \forall x, \lambda$
- primales Problem: $\min_x \sup_\lambda \mathcal{L}(x, \lambda)$
- duales Problem: $\max_\lambda \inf_x \mathcal{L}(x, \lambda)$
- primales lineares Programm: $\min c_x^T x + c_y^T y$ u.d.N. $A_U^x x + A_U^y y \geq b_U$, $A_G^x + A_G^y = b_G$, $x \geq 0$
- duales lineares Programm: $\max \mu^T b_U + \lambda^T b_G$ u.d.N. $(A_U^x)^T \mu + (A_G^x)^T \lambda \leq c_x$, $(A_U^y)^T \mu + (A_G^y)^T \lambda = c_y$, $\mu \geq 0$
- Standardform (LPS): $\min c^T x$ u.d.N. $Ax = b$, $x \geq 0$
  - Standardform des primalen linearen Programms: $\min c_x^T x + c_y^T y_+ - c_y^T y_-$ u.d.N. $A_U^x x + A_U^y y_+ - A_U^y y_- - z = b_U$, $A_G^x x + A_G^y y_+ - A_G^y y_- = b_G$, $x, y_+, y_-, z \geq 0$

Lineare Programmierung

- Zulässigbereich von LPS: $M = \{x \in \mathbb{R}^n \mid x \geq 0, Ax = b\}$
- Polyeder: Schnittmenge endlich vieler abgeschlossenen Halbräumen. Falls beschränkt -> Polytop
- x ist Ecke, falls $x = (1-t)y + tz, t \in (0,1), y \in P, z \in P \implies x = y = z$
- Basislösung $x \in M$ zur Basis $\mathcal{B}$: $x_\mathcal{N} = 0$ und $A[:, \mathcal{B}]$ regulär
  - $x \in M$ ist Ecke $\iff x$ ist Basislösung

Simplex-Tableau Algorithmus

- Eingabe: $c, A, b, \mathcal{B}$
- $T = T_\mathcal{B}^{-1} T_0 = \begin{pmatrix} 1 & -c_\mathcal{B}^T B \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & c^T & 0 \\ 0 & A & b \end{pmatrix}$
- while not $\mu \geq 0$:
  - $\mu = T[0, : -1]$
  - $q = \arg\min \mu$
  - $w = T[:, q]$
  - if $w \leq 0$ error
  - k = $\arg\min_{i \in \{j \mid w_j > 0\}} \frac{T[i, -1]}{w_i}$
  - $T[k, :] = \frac{T[k, :]}{w_k}$
  - $T[i, :] = T[i, :] - T[i, q] \cdot T[k, :]$ mit $i \neq k$
  - $\mathcal{B}(k) = q$
- Ausgabe: $x_\mathcal{B} = T[1 :, -1]$, $c^T x = -T[0, -1]$
- Hilfsproblem: $\min c^T x + N e^T z$ u.d.N. $Ax + z = b$, $x, z \geq 0$

Nichtlineare Programmierung

- Tangentialkegel: $d \in \mathbb{R}^n$ heißt tangential zu $M \subset \mathbb{R}^n$ im Punkt $x \in M$, falls $\exists (x_k)_{k \in \mathbb{N}}$ aus M und $(t_k)_{k \in \mathbb{N}}$ in $(0, \infty)$ sodass $x_k \implies x$, $t_k \implies 0$, $\frac{x_k - x}{t_k} \implies d$. $\mathcal{T}_M(x)$ ist die Menge aller d.
- linearisierter Tangentialkegel: Sei $\mathcal{A}(x) := \left\{i \in I_U : g_{i(x)} = 0\right\}$. Dann ist $\mathcal{T}^{\text{lin}}(x) := \{d \in \mathbb{R}^n : \nabla g_\mathcal{A}(x)^T d \leq 0, \nabla h(x)^T d = 0\} \supset \mathcal{T}_M(x)$
- Abadie CQ ist für $x_* \in M$ erfüllt, falls $\mathcal{T}^{\text{lin}}(x_*) = \mathcal{T}_M(x_*)$
- LICQ ist für $x_* \in M$ erfüllt, falls alle $v \in \{\nabla g_\mathcal{A}(x_*)\} \cup \left\{\nabla h_{I_G}(x_*)\right\}$ linear unabhängig
- KKT unter LICQ: Sei x_* lokales Minimum und erfülle LICQ. Dann $\exists \lambda_* \in \mathbb{R}^{m_G}, \mu_* \in \mathbb{R}^{m_U}$:

    - $\nabla_x \mathcal{L}(x_*, \mu_*, \lambda_*) = 0$
    - $g_{I_U}(x_*) \leq 0$
    - $h_{I_G}(x_*) = 0$
    - $\mu_{I_U,*} \geq 0$
    - $\sum_{i \in I_U} \mu_{i,*} g_{i(x_*)} = 0$
- konvexes Programm: min $f(x)$ u.d.N. $g_{I_U}(x) \leq 0$, $a_{I_G}^T x = b_{I_G}$ mit f, g_{I_U} konvex und stetig diffbar
- Slater-Bedingung für konvexes Programm ist erfüllt, falls $\exists \overline{x} \in M : g_{I_U}(\overline{x}) < 0$. Falls ja, sind die KKT-Bedingungen für globales Minimum hinreichend.

Quadratische Programmierung

- quadratisches Programm: min $\frac{1}{2} x^T Q x + c^T x$ u.d.N. $A_U x \leq b_U$, $A_G x = b_G$ mit Q symmetrisch
    - ersetze allgemeine Matrix Q mit $\frac{1}{2}(Q + Q^T)$ ohne Zielfunktion zu ändern
    - Q positiv semidefinit $\iff$ Zielfunktion konvex
- KKT-Bedingungen:
    - $Q x_* + A_U^T \mu_* + A_G^T \lambda_* = -c$
    - $A_U x_* \leq b_U$
    - $A_G x_* = b_G$
    - $\mu_* \geq 0$
    - $(b_U - A_U x_*)^T \mu_* = 0$

Strategie der aktiven Mengen

- Eingabe: Q positiv definit, c, A, b, $x_0 \in M$
- $\mathcal{A}_0 = \emptyset$
- while true:
    - Löse $\begin{pmatrix} Q & A_{U[\mathcal{A}_k,:]}^T & A_G^T \\ A_{U[\mathcal{A}_k,:]} & 0 & 0 \\ A_G & 0 & 0 \end{pmatrix} \begin{pmatrix} d_k \\ \mu_{k+1}[\mathcal{A}_k] \\ \lambda_{k+1} \end{pmatrix} = \begin{pmatrix} -\nabla f(x_k) \\ 0 \\ 0 \end{pmatrix}$
    - if $d_k = 0$:
        - if $\mu_{k+1}[\mathcal{A}_k] \geq 0$ return x_k
        - else:
            - $q = \arg\min_{\mathcal{A}_k} \mu$
            - $x_{k+1} = x_k$
            - $\mathcal{A}_{k+1} = \mathcal{A}_k \setminus \{q\}$
    - else:
        - $t_i^k := \frac{b_{u[i]} - A_{U[i,:]} x_k}{A_{U[i,:]} d_k}$ für $i \in \left\{ i \mid A_{U[i,:]} d_k > 0 \right\}$
        - if $\min t^k \geq 1$:
            - $x_{k+1} = x_k + d_k$
            - $\mathcal{A}_{k+1} = \mathcal{A}_k$
        - else:
            - $r = \arg\min t^k$
            - $x_{k+1} = x_k + t_r^k d_k$
            - $\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{r\}$