Lösungsvorschlag zur Vorlesung
Nichtglatte Optimierung und Anwendungen

G13. Subdifferential des größten Eigenwerts

Sei $S_n = \{A \in \mathbb{R}^{n,n} \ : \ A = A^T\}$ und

$$\lambda_{max} : A \in S_n \mapsto \max_{\|v\|=1} v^T A v$$

der größte Eigenwert.

Unser Ziel ist der Nachweis, dass

$$\partial\lambda_{max}(A) = \{W \in S_n \ : \ W \succeq 0, \ \text{spur}(W) = 1, \ W \bullet A = \lambda_{max}(A)\}.$$

Bemerkung: Es gilt $\text{spur}(B^{-1}AB) = \text{spur}(A)$ und $A \bullet B = \text{spur}(AB^T)$.

Zeigen Sie:

a) Sei $W \in \partial\lambda_{max}(A)$. Dann muss gelten $W \bullet A = \lambda_{max}(A)$.
Tip: $\lambda_{max}(tA) = t\lambda_{max}(A)$ für alle $t \geq 0$.

Die Subgradientenungleichung lautet

$$\lambda_{max}(B) - \lambda_{max}(A) \geq W \bullet (B - A) \quad \forall\, B \in S_n.$$

Für $B = tA$, $t \geq 0$ ergibt sich

$$\lambda_{max}(tA) - \lambda_{max}(A) = (t-1)\lambda_{max}(A) \geq (t-1)W \bullet A \quad \forall\, t \geq 0.$$

Einsetzen von $t = 2$ und $t = 0$ liefert $\lambda_{max}(A) = W \bullet A$.

b) Sei $W \in \partial\lambda_{max}(A)$. Dann gilt $\text{spur}(W) = 1$. **Tip:** Betrachte $B = A + tI$.

Für $B = A + tI$ ergibt sich

$$\lambda_{max}(A + tI) - \lambda_{max}(A) = t \geq tW \bullet I = t\text{spur}(W) \quad \forall\, t.$$

Dies zeigt $\text{spur}(W) = 1$.

c) Sei $W \in \partial\lambda_{max}(A)$. Dann gilt $W \bullet B \leq \lambda_{max}(B)$ für alle $B \in S_n$. Folgern Sie, dass $W \succeq 0$ gelten muss.

Wegen a) liefert die Subgradientenungleichung

$$\lambda_{max}(B) \geq W \bullet B \quad \forall\, B \in S_n.$$

Wäre W nicht positiv definit, dann existiert $v \in \mathbb{R}^n$ mit $v^T W v = W \bullet (vv^T) < 0$. Die Wahl $B = -vv^T$ ergibt nun den Widerspruch

$$0 = \lambda_{max}(-vv^T) \geq -v^T W v > 0.$$

d) Sei umgekehrt $W \in S_n$ mit $W \succeq 0$, $\operatorname{spur}(W) = 1$, $W \bullet A = \lambda_{max}(A)$. Zeigen Sie dass dann gilt $W \bullet B \leq \lambda_{max}(B)$ für alle $B \in S_n$ und folgern Sie $W \in \partial\lambda_{max}(A)$. **Tip:** Nutzen Sie $W \bullet B = \operatorname{spur}(WB)$ Transformieren Sie B auf Diagonalform und nutzen Sie die Eigenschaften der Spur.

Die Subgradientenungleichung ist äquivalent zu

$$\lambda_{max}(B) \geq W \bullet B \quad \forall B \in S_n.$$

Sei nun Q orthogonal mit $Q^T BQ = D = \operatorname{diag}(\lambda_1, \lambda_n)$. und setze $Q^T WQ = (m_{ij})$ Dann gilt $m_{ii} \geq 0$, da $Q^T WQ \succeq 0$, und

$$\operatorname{spur}(W) = \operatorname{spur}(Q^T WQ) = \sum_i m_{ii} = 1.$$

Weiter gilt

$$\begin{aligned} W \bullet B &= \operatorname{spur}(WB) = \operatorname{spur}(Q^T WQQ^T BQ) = \operatorname{spur}(Q^T WQD) \\ &= \sum_i m_{ii}\lambda_i \leq \lambda_{max}(B) \sum_i m_{ii} = \lambda_{max}(B). \end{aligned}$$

G14. Optimalitätsbedingungen für nichtglatte restringierte Probleme

Sei $f : \mathbb{R}^n \to \mathbb{R}$ konvex und $C \subset \mathbb{R}^n$ nichtleer, abgeschlossen und konvex.

Der Tangentialkegel von C in x ist definiert als

$$T_C(x) := \overline{\{s \in \mathbb{R}^n \ : \ \exists\, t > 0 : x + ts \in C\}}.$$

Wir betrachten das Optimierungsproblem

$$\min f(x) \quad \text{s.t.} \quad x \in C. \tag{P}$$

Wir definieren die Indikatorfunktion

$$\chi_C(x) = \begin{cases} 0, & x \in C \\ \infty, & x \notin C \end{cases}$$

Für $x \in C$ definieren wir analog wie bisher zu $\mathbb{R}$-wertigen Funktionen

$$\partial\chi_C(x) = \{g \in \mathbb{R}^n \ : \ \chi_C(y) - \chi_C(x) \geq g^T(y - x) \ \ \forall\, y \in C\}.$$

Zeigen Sie:

a) (P) is äquivalent zu

$$\min f(x) + \chi_C(x).$$

b) Sei $x \in C$. Dann gilt:

$$\partial\chi_C(x) = \{g \in \mathbb{R}^n \ : \ g^T s \leq 0 \quad \forall\, s \in T_C(x)\}.$$

$\subset$: Sei $g \in \partial\chi_C(x)$. Dann gilt

$$0 \geq g^T(y - x) \quad \forall\, y \in C.$$

Sei nun $s \in \mathbb{R}^n$ mit $x + ts \in C$ für ein $t > 0$. Dann folgt $0 \geq g^T(ts)$, also $g^T s \leq 0$. Aus Steigfkeitsgründen gilt dies auch für den Abschluss $T_C(x)$ dieser Richtungen.

$\subset$: Sei $g^T s \leq 0 \quad \forall\, s \in T_C(x)$. Für $y \notin C$ ist die Subgradientenungleichung trivial. Für $y \in C$ ist $s = y - x \in T_C(x)$ und daher gilt

$$g^T(y - x) \leq 0 = \chi_C(y) - \chi_C(x).$$

c) Ist $\bar{x} \in C$ optimale Lösung von (P), dann gilt:

$$\exists\, g \in \partial f(\bar{x}) \;:\; g^T s \geq 0 \;\; \forall\, s \in T_C(\bar{x}).$$

Bemerkung: Sie dürfen verwenden, dass in $\bar{x}$ die Summenregel für Subdifferentiale angewendet werden darf.

Es gilt

$$0 \in \partial(f + \chi_C)(\bar{x}) = \partial f(\bar{x}) + \partial \chi_C(\bar{x}).$$

Es gibt also $g \in \partial f(\bar{x})$ mit $-g \in \partial \chi_C(\bar{x})$. b) ergibt nun die Behauptung.

In der Nichtlinearen Optimierung lautet die Optimalitätsbedingung völlig analog:

$$\bar{x} \in C \quad \text{und} \quad \nabla f(\bar{x})^T s \geq 0 \;\; \forall\, s \in T_C(\bar{x}).$$

d) Sei nun $C := \{x \in \mathbb{R}^n \,:\, c_i(x) \leq 0, \;\; i = 1, \ldots, m\}$ mit stetig differenzierbaren konvexen Funktionen $c_i : \mathbb{R}^n \to \mathbb{R}$. Kann man unter einer Constraint Qualification aus c) analog zur Nichtlinearen Optimierung KKT-Bedingungen ableiten. Wie sehen sie aus?

Unter einer Constraint Qualification stimmen wie in der Nichtlinearen Optimierung der Tangentialkegel $T_C(\bar{x})$ und der Linearisierungskegel

$$T_L(\bar{x}, c) = \left\{ s \in \mathbb{R}^n \,:\, \nabla c_i(\bar{x})^T s \leq 0 \;\; \forall\, i \in \mathcal{A}(\bar{x}) \right\}$$

überein, wobei $\mathcal{A}(\bar{x})$ die aktive Indexmenge bezeichnet. Anwendung des Farkas Lemmas liefert nun die KKT-Bedingungen: Es gibt $g \in \partial f(\bar{x})$ und $\bar{\lambda} \in \mathbb{R}^m$ mit

$$\begin{aligned} c(\bar{x}) &\leq 0, \\ g + \nabla c(\bar{x}) \bar{\lambda} &= 0, \\ \bar{\lambda} \geq 0, \quad c(\bar{x})^T \bar{\lambda} &= 0. \end{aligned}$$