Universität zu Köln
Mathematisches Institut
Prof. Dr. F. Vallentin

Einführung in die Mathematik des Operations Research

Sommersemester 2017

— Lösungsskizze Blatt 7 —

Aufgabe 7.1 (10 Punkte)

a) Zeigen Sie, dass die folgende Menge von univariaten Polynomen konvex ist:

$$P = \{(p_0, p_1, \ldots, p_d) \in \mathbb{R}^{d+1} : p_0 + p_1 x + \cdots + p_d x^d \geq 0 \text{ für alle } x \in \mathbb{R}\}.$$

b) Bestimmen Sie einen inneren Punkt von P.

c) Bestimmen Sie die konvexen Hüllen der folgenden Mengen:

$$A = \{(x, 1/x) \in \mathbb{R}^2 : x \in \mathbb{R}_{>0}\} \quad \text{und} \quad B = \{(x, \sin x) \in \mathbb{R}^2 : x \in \mathbb{R}\}.$$

Lösung:

a) Seien $p, q \in P$, und sei $f(x) = p_0 + p_1 x + \cdots + p_d x^d$ und entsprechend $g(x)$ für q. Dann sind $f(x)$ und $g(x)$ nicht-negativ für alle $x \in \mathbb{R}$, also auch $(1-\alpha)f(x) + \alpha g(x)$ für $\alpha \in [0, 1]$.

b) Falls d grade ist, betrachte beispielsweise p mit $p_0 = 1 = p_d$ und $p_i = 0$ sonst.

Es ist $(1-\varepsilon) - \varepsilon x - \cdots - (1-\varepsilon)x^d \geq 0$ für grades d und $\varepsilon > 0$ ausreichend klein, also ist p ein innerer Punkt.

Falls d ungerade ist, gibt es keine inneren Punkte. Hierzu beachte, dass der führende Koeffizient p_d in einem Punkt aus P Null sein muss (warum?), aber in jeder Umgebung gibt es Punkte, deren letzte Koordinate nicht Null ist.

c) Bei der Menge A erhalten wir den Epigraph der Funktion $f(x) = 1/x$ mit $x > 0$, also $\operatorname{conv} A = \{(x, \beta) : x > 0, 1/x \leq \beta\}$. Hierzu sei $y = (y_1, y_2) \in \operatorname{epi} f$, also $y_2 \geq 1/y_1$. Betrachte die Grade $g = \{y + \lambda(1/2, -1/2)^{\mathsf{T}} : \lambda \in \mathbb{R}\}$, dann schneidet g den Graph von f zwei mal (wieso?). Die beiden Schnittpunkte haben dann y auf ihrer Verbindungsstrecke, also ist $\operatorname{epi} f \subseteq \operatorname{conv} A$. Da $\operatorname{epi} f$ konvex ist (!), ist dann $\operatorname{epi} f = \operatorname{conv} A$.

Es ist $\operatorname{conv} B = \{(x, y) : x \in \mathbb{R}, y \in [-1, 1]\}$. Hierzu beobachten wir zunächst, dass $V = \{(x, 1) : x \in \mathbb{R}\}$ und $W = \{(x, -1) : x \in \mathbb{R}\}$ in $\operatorname{conv} B$ enthalten sein müssen. Dies ist der Fall, da es für jedes $x \in \mathbb{R}$ Werte $x_0, x_1 \in \mathbb{R}$ gibt mit $x_0 \leq x < x_1$ und $(x_0, 1), (x_1, 1) \in B$, analog für W.

Nun ist $\operatorname{conv} B \supseteq \operatorname{conv}(V \cup W) = \{(x, y) : x \in \mathbb{R}, y \in [-1, 1]\} \supseteq B$, und da $\operatorname{conv} B$ die kleinste konvexe Menge ist, die B enthält, sind wir fertig.

Aufgabe 7.2 (10 Punkte) Welche der folgenden Funktionen $f \colon \mathbb{R}^n \to \mathbb{R}$ sind konvex?

a) $f(x) = \max\{g(x), h(x)\}$ für $g, h \colon \mathbb{R}^n \to \mathbb{R}$ konvex.

b) $f(x) = x^{\mathsf{T}} \begin{pmatrix} -6/5 & 12/5 \\ 12/5 & 1/5 \end{pmatrix} x.$

c) $f(x) = x^\mathsf{T} \begin{pmatrix} 4/3 & 0 & 2/3 \\ 0 & 2/3 & -2/3 \\ 2/3 & -2/3 & 1 \end{pmatrix} x.$

d) $f(x) = e^{x^\mathsf{T} x}$.

Lösung:

a) Wir beweisen die Konvexität von f über die Jensen Ungleichung.

$$\begin{aligned} f((1-\alpha)x + \alpha y) &= \max\{g((1-\alpha)x+\alpha y), h((1-\alpha)x+\alpha y)\} \\ &\overset{g,h \text{ konvex}}{\leq} \max\{(1-\alpha)g(x)+\alpha g(y), (1-\alpha)h(x)+\alpha h(y)\} \\ &\overset{g(z)\leq\max\{g(z),h(z)\}}{\leq} \max\{(1-\alpha)g(x), (1-\alpha)h(x)\} + \max\{\alpha g(y), \alpha h(y)\} \\ &= (1-\alpha)\max\{g(x),h(x)\} + \alpha\max\{g(y),h(y)\} \\ &= (1-\alpha)f(x) + \alpha f(y) \end{aligned}$$

b) Bei Funktionen der Art $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = x^T Bx$, mit B symmetrisch, entspricht $2B$ genau der Hessematrix der Funktion f. $2B$ ist positiv semidefinit genau dann wenn B positiv semidefinit ist. Es bleibt also zu prüfen, ob $B = \begin{pmatrix} -6/5 & 12/5 \\ 12/5 & 1/5 \end{pmatrix}$ positiv semidefinit ist. Da $Tr(B) = \lambda_1 + \lambda_2 = -\frac{6}{5} + \frac{1}{5} = -1$ ist, ist mindestens einer der Eigenwerte < 0 und B somit nicht positiv semidefinit, womit auch f nicht konvex ist.

c) Analog zu b) betrachtet man die Hesse-Matrix

$$C = \begin{pmatrix} 4/3 & 0 & 2/3 \\ 0 & 2/3 & -2/3 \\ 2/3 & -2/3 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}^T + \frac{2}{3}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}^T.$$

Da Summen von positiv semidefiniten Matrizen wieder positiv semidefinit sind, ist damit auch C positiv semidefinit und f damit konvex. Alternativ kann man auch das charakteristische Polynom bestimmen und hieraus die Eigenwerte ausrechnen.

d) Die Hessematrix von $f(x) = e^{x^\mathsf{T} x}$ hat an i,j-ter Stelle die Einträge

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = 4x_i x_j e^{x^\mathsf{T} x}.$$

Sie lässt sich somit darstellen als $H(f)(a) = 4aa^T e^{a^\mathsf{T} a}$. Dies impliziert, dass sie positiv semidefinit ist, da $y^T 4e^{a^\mathsf{T} a} aa^T y = 4e^{a^\mathsf{T} a}(y^T a)^2 \geq 0$ für alle Vektoren y und a.

Aufgabe 7.3 (10 Punkte) Zeigen Sie: Sei $A \subseteq \mathbb{R}^n$ eine kompakte Menge, dann ist auch ihre konvexe Hülle conv A kompakt.
Hinweis: Satz von Carathéodory.

Lösung: Wir zeigen, dass jede Folge in conv A eine konvergente Teilfolge mit Grenzwert in conv A besitzt. Sei also $x_j \in \operatorname{conv} A$ für $j \in \mathbb{N}$. Nach dem Satz von Carathéodory können wir jedes x_j als Konvexkombination von $n+1$ Elementen in A beschreiben:

$$x_j = \sum_{i=1}^{n+1} \lambda_i^j a_i^j, \quad a_i^j \in A, \lambda_i^j \geq 0, \sum_{i=1}^{n+1} \lambda_i^j = 1.$$

Wir betrachten die entsprechenden Punkte $(\lambda_1^j, \dots, \lambda_{n+1}^j, a_1^j, \dots, a_{n+1}^j)$ in $\Delta^n \times A^{n+1}$, wobei

$$\Delta^n = \{(\lambda_1, \dots, \lambda_{n+1}) : \lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1\}$$

der Standardsimplex ist.
Da $\Delta^n \times A^{n+1}$ kompakt ist, was wir gleich zeigen werden, gibt es eine konvergente Teilfolge $(\lambda_1^{j_\ell}, \dots, \lambda_{n+1}^{j_\ell}, a_1^{j_\ell}, \dots, a_{n+1}^{j_\ell})$ dieser Folge mit Grenzwert $(\lambda_1, \dots, \lambda_{n+1}, a_1, \dots, a_{n+1})$ in $\Delta^n \times A^{n+1}$. Dann gilt $\lim_{\ell\to\infty} \lambda_i^{j_\ell} a_i^{j_\ell} = \lambda_i a_i$ und damit:

$$\lim_{\ell\to\infty} x_{j_\ell} = \lim_{\ell\to\infty} \sum_{i=1}^{n+1} \lambda_i^{j_\ell} a_i^{j_\ell} = \sum_{i=1}^{n+1} \lim_{\ell\to\infty} \lambda_i^{j_\ell} a_i^{j_\ell} = \sum_{i=1}^{n+1} \lambda_i a_i \in \text{conv } A.$$

Es bleibt zu zeigen, dass die Menge $\Delta^n \times A^{n+1}$ kompakt ist. Da A kompakt ist, ist auch A^{n+1} kompakt. Δ^n ist beschränkt, da für $(\lambda_1, \dots, \lambda_{n+1}) \in \Delta^n$ gilt:

$$\|(\lambda_1, \dots, \lambda_{n+1})\| \leq \sqrt{\sum_{i=1}^{n+1} \lambda_i^j} = 1.$$

Um zu zeigen, dass Δ^n abgeschlossen ist, betrachte eine konvergente Folge $(\lambda_1^j, \dots, \lambda_{n+1}^j)$ in Δ^n mit Grenzwert $(\lambda_1, \dots, \lambda_{n+1})$. Es gilt $\sum_{i=1}^{n+1} \lambda_i = \lim_{j\to\infty} \sum_{i=1}^{n+1} \lambda_i^j = 1$ und $\lambda_i = \lim_{j\to\infty} \lambda_i^j \geq 0$. Also ist $(\lambda_1, \dots, \lambda_{n+1})$ in Δ^n. Δ^n ist also kompakt, und damit auch $\Delta^n \times A^{n+1}$.

Alternative Lösung: Anstelle der Definition von Kompaktheit können wir auch verwenden, dass Bilder kompakter Mengen unter stetigen Abbildungen kompakt sind. Hierfür betrachten wir die Funktion $\phi : \Delta^n \times A^{n+1} \to \mathbb{R}^n$ definiert durch $\phi((\lambda_1, \dots, \lambda_{n+1}, a_1, \dots, a_{n+1})) = \sum_{i=1}^{n+1} \lambda_i a_i$. Nach dem Satz von Carathéodory ist das Bild dieser Abbildung $\operatorname{conv} A$. Da ϕ stetig ist (!) und $\Delta^n \times A^{n+1}$ kompakt (wie oben bewiesen), ist also $\operatorname{conv} A$ kompakt.