Nichtlineare Optimierung - 1. Hausaufgabe

Claudia Wohlgemuth *366323*
Thorsten Lucke *366515*
Felix Thoma *358638*

Tutor: Mathieu Rosière, Di 8-10 Uhr

23. April 2017

1.1	1.2	1.3	1.4	Σ

Anmerkungen:

Aufgabe 1.1

(i) Es sei $A\colon \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung. Wir werden zeigen, dass A genau dann positiv definit ist, wenn A stark positiv ist.

Beweis. Sei A zunächst stark positiv. Dann existiert ein $\alpha > 0$, sodass die Abschätzung

$$\langle Ax, x\rangle \geq \alpha \|x\|^2 \tag{1}$$

für alle $x \in \mathbb{R}^n$ erfüllt ist. Der positiven Definitheit von Normen wegen, folgt für alle $x \in \mathbb{R}^n \setminus \{0\}$ unmittelbar

$$\langle Ax, x\rangle \geq \alpha \|x\|^2 > 0.$$

Die Abbildung A ist also positiv definit.

Sei A nun positiv definit und $x \in \mathbb{R}^n$ beliebig gewählt. Ist $x = 0$, so gilt (1) für alle $\alpha > 0$. Sei daher x von Null verschieden. Dann gilt (1) genau dann, wenn

$$\frac{\langle Ax, x\rangle}{\|x\|^2} \geq \alpha$$

beziehungsweise

$$\langle Ay, y\rangle \geq \alpha$$

für alle $y \in \mathbb{R}^n$ mit $\|y\| = 1$ erfüllt ist. Da lineare Abbildungen – auf endlichen Vekorräumen operierend – beschränkt und stetig sind, ist auch die Abbildung

$$f\colon \mathbb{S}^1 \to \mathbb{R},\ y \mapsto \langle Ay, y\rangle$$

stetig. Da die Einheitssphäre $\mathbb{S}^1$ kompakt ist, folgt aus dem Satz von Weierstraß die Existenz eines Minimums $\alpha \in \mathbb{R}$ mit $f(y) \geq \alpha$ für alle $y \in \mathbb{S}^1$. Da A positiv definit ist, muss zudem $\alpha > 0$ gelten. Es folgt daher insgesamt

$$\langle Ax, x\rangle \geq \alpha \|x\|^2$$

für alle $x \in \mathbb{R}^n$ und damit die starke Positivität von A. □

(ii) Wir zeigen nun, dass (i) im Allgemeinen falsch ist. Dazu betrachten wir die Abbildung

$$A\colon \ell_2 \to \ell_2,\ (x_n)_{n\in\mathbb{N}} \mapsto \left(\frac{1}{n}x_n\right)_{n\in\mathbb{N}}.$$

Da für alle $n \in \mathbb{N}$ die Abschätzung $(x_n/n)^2 \leq (x_n)^2$ gilt, folgt wegen

$$\sum_{n=1}^{\infty} \left(\frac{1}{n}x_n\right)^2 \leq \sum_{n=1}^{\infty} (x_n)^2$$

die Wohldefiniertheit und Beschränktheit der Abbildung A aus dem Majorantenkriterium für Reihen.

Desweiteren folgt aus

$$0 \leq \langle A(x_n)_{n\in\mathbb{N}}, (x_n)_{n\in\mathbb{N}} \rangle = \sum_{n=1}^{\infty} \frac{1}{n} x_n^2$$

die positive Definitheit.

Seien nun $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \in \ell_2$ zwei Folgen und $\lambda \in \mathbb{R}$ beliebig gewählt. Dann ist A wegen

$$\begin{aligned} A(\lambda(x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}}) &= A((\lambda x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}}) \\ &= A((\lambda x_n + y_n)_{n\in\mathbb{N}}) \\ &= \left(\frac{1}{n}(\lambda x_n + y_n)\right)_{n\in\mathbb{N}} \\ &= \left(\lambda \frac{x_n}{n} + \frac{y_n}{n}\right)_{n\in\mathbb{N}} \\ &= \lambda \left(\frac{x_n}{n}\right)_{n\in\mathbb{N}} + \left(\frac{y_n}{n}\right)_{n\in\mathbb{N}} \\ &= \lambda A(x_n)_{n\in\mathbb{N}} + A(y_n)_{n\in\mathbb{N}} \end{aligned}$$

linear.

Sei nun $\alpha > 0$ beliebig gewählt. Dann gibt es ein $k \in \mathbb{N}$ mit $\frac{1}{k} < \alpha$. Für den k-ten Einheitsbasisvektor $e_k = (\delta_{k,n})_{n\in\mathbb{N}}$ des ℓ_2 gilt dann

$$\langle Ae_k, e_k \rangle = \sum_{n=1}^{\infty} \frac{1}{n} \delta_{k,n}^2 = \frac{1}{k} < \alpha.$$

Somit kann ein derartiges α nicht existieren, das bedeutet, A ist nicht stark positiv.

Aufgabe 1.2

Es sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ eine Funktion.

(i) Sei f zusätzlich konvex und seien $(x, \alpha), (y, \beta) \in \operatorname{epi}(f)$ und $\lambda \in [0, 1]$. Dann folgt aus der Konvexität von f und der Definition des Epigraphen

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \leq \lambda\alpha + (1-\lambda)\beta,$$

das heißt, wegen $(\lambda x + (1-\lambda)y), \lambda\alpha + (1-\lambda)\beta) = \lambda(x, \alpha) + (1-\lambda)(y, \beta) \in \operatorname{epi}(f)$ ist $\operatorname{epi}(f)$ konvex.
Ist andererseits $\operatorname{epi}(f)$ konvex und seien $x, y \in \mathbb{R}^n$ und $\lambda \in [0, 1]$, so erhalten wir mit $(x, f(x)), (y, f(y)) \in \operatorname{epi}(f)$ unmittelbar $(\lambda x+(1-\lambda)y), \lambda f(x)+(1-\lambda)f(y)) \in \operatorname{epi}(f)$ und somit die Behauptung

$$f(\lambda x + (1-\lambda)y) \leq \lambda\alpha + (1-\lambda)\beta.$$

(ii) Diese Aussage ist leider falsch, wie sich leicht an

$$f\colon \mathbb{R} \to \mathbb{R}, x \mapsto \begin{cases} 0, & x \in \mathbb{R}\backslash 0, \\ -1, & x = 0, \end{cases}$$

einsehen lässt. Offensichtlich ist f nicht stetig. Dennoch ist

$$\operatorname{epi}(f) = \mathbb{R} \times [0, \infty) \cup \{0\} \times [-1, \infty)$$

abgeschlossen.

Wir zeigen stattdessen, dass f genau dann unterhalbstetig ist, wenn $\operatorname{epi}(f)$ abgeschlossen ist.
Sei dazu zunächst f unterhalbstetig, das heißt, für alle $x \in \mathbb{R}^n$ gelte

$$f(x) \leq \liminf_{t \to x} f(t).$$

Sei ferner $(x_n, \alpha_n)_{n\in\mathbb{N}} \subset \operatorname{epi}(f) \subset \mathbb{R}^{n+1}$ eine beliebige in $\mathbb{R}^{n+1}$ konvergente Folge. Wir setzen $x \coloneqq \lim_{n\to\infty} x_n \in \mathbb{R}^n$ und $\alpha \coloneqq \lim_{n\to\infty} \alpha_n$. Dann gilt wegen der Unterhalbstetigkeit von f

$$\alpha = \lim_{n\to\infty} \alpha_n \geq \liminf_{n\to\infty} f(x_n) \geq f(x),$$

also $(x, \alpha) \in \operatorname{epi}(f)$. Somit ist $\operatorname{epi}(f)$ abgeschlossen.
Sei andererseits $\operatorname{epi}(f)$ abgeschlossen und $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^n$ eine beliebige gegen $\mathbb{R}^n$ konvergente Folge. Dann ist wegen $(x_n, f(x_n)) \in \operatorname{epi}(f)$ für alle $n \in \mathbb{N}$ und der Abgeschlossenheit von $\operatorname{epi}(f)$ auch

$$(\lim_{n\to\infty} x_n, \liminf_{n\to\infty} f(x_n)) = (x, \liminf_{n\to\infty} f(x_n)) \in \operatorname{epi}(f)$$

und somit gilt

$$f(x) \leq \liminf_{n\to\infty} f(x_n).$$

(iii) Aus der Analysis-Übung vom 9. Dezember 2016 wissen wir, dass koerzive und unterhalbstetige Funktionen ein Minimum besitzen.

Aufgabe 1.3

Wir nutzen im Folgenden die zum Hinweis äquivalente Aussage: Ein normierter Raum X ist genau dann endlichdimensional, wenn beschränkte und abgeschlossene Mengen kompakt sind, dies ist insbesondere genau dann erfüllt, wenn die abgeschlossene Einheitskugel kompakt ist.

Sei $(X, \|\cdot\|)$ ein normierter Raum.

$\Rightarrow$ Sei X zusätzlich endlichdimensional und $a \in \mathbb{R}$ beliebig. Sei ferner f ein beliebiges stetiges und koerzives Funktional. Weil f koerziv ist, gibt es ein $r \in \mathbb{R}$, sodass für alle $x \in X$ mit $\|x\| > r$ bereits $f(x) > \alpha$ gilt. Somit ist $N(\alpha) \subset \overline{B(0,r)}$, d.h. $N(\alpha)$ ist beschränkt.
Sei nun $(x_n)_{n\in\mathbb{N}} \subset N(\alpha) \subset X$ eine gegen $x \in X$ konvergente Folge. Weil f stetig ist, gilt
$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \alpha = \alpha.$$
Somit ist auch $x \in N(\alpha)$ und die Niveaumenge abgeschlossen. Weil X endlichdimensional ist, folgt mit Heine-Borel die Kompaktheit von $N(\alpha)$.

$\Leftarrow$ Für jedes $\alpha \in \mathbb{R}$ und jedes stetige, koerzive Funktional f sei die Niveaumenge $N(\alpha)$ kompakt. Dann gilt dies insbesondere für $\|\cdot\| : X \to [0,\infty)$. Dabei ist $\|\cdot\|$ offensichtlich koerziv und stetig. Weiterhin ist $N(1) = \overline{B(0,1)}$ nach Voraussetzung kompakt. Dass die abgeschlossene Einheitskugel kompakt ist, ist äquivalent zur Endlichdimensionalität von X.

Aufgabe 1.4

(i) Es gilt nach Vorlesung

$$x^T Hx = \sum_{i=1}^{m} (x_1\xi_i + x_2)^2.$$

Sei nun $x \in \mathbb{R}^2$ beliebig. Wegen $(x_1\xi_i+x_2)^2 \geq 0$ ist $x^T Hx \geq 0$ und weil $\xi \mapsto x_1\xi+x_2$ nur eine Nullstelle besitzt, die ξ_i aber paarweise verschieden sind, gibt es ein ξ_i mit $(x_1\xi_i + x_2)^2 > 0$. Somit ist $x^T Hx > 0$ für jedes $x \in \mathbb{R}^2$ und H also positiv definit. Nun ist f zweimal stetig differenzierar und damit genau dann konvex, wenn $f'' = H$ positiv definit ist. Hiermit folgt die Behauptung.

(ii) Mit 1.1. wissen wir, dass H sogar stark positiv ist. Wir haben also mit Cauchy-Schwarz und der starken Positivität

$$\begin{aligned} f(x) &= \frac{1}{2}\langle Hx, x\rangle + \langle b, x\rangle + c \\ &\geq \frac{1}{2}\langle Hx, x\rangle + \|b\|_2\|x\|_2 + c \\ &\geq \frac{1}{2}\alpha\|x\|_2^2 + \|b\|_2\|x\|_2 + c =: g(\|x\|_2). \end{aligned}$$

Für quadratische, reelle Funktionen gilt bekanntlich

$$\infty = \lim_{\|x\|_2\to\infty} g(\|x\|_2) \leq \lim_{\|x\|_2\to\infty} f(x).$$

Somit ist f koerzitiv und zusammen mit der Stetigkeit folgt die Existenz eines Minimierers. Aufgrund der Konvexität von f ist dieser eindeutig.