

Nichtlineare Optimierung - 2. Hausaufgabe

Claudia Wohlgemuth	<i>366323</i>
Thorsten Lucke	<i>363089</i>
Felix Thoma	<i>358638</i>

Tutor: Mathieu Rosière, Di 8-10 Uhr

14. Mai 2017

2.1	2.2	2.3	2.4	Σ

Anmerkungen:

Aufgabe 2.1

Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ konvex und $x \in \mathbb{R}^n$. Für $h \in \mathbb{R}^n$ bezeichne $f'(x, h)$ die Richtungsableitung von f im Punkt x in Richtung h und weiter definieren wir

$$g_h : \mathbb{R} \rightarrow \mathbb{R}, \quad g_h(t) = f(x + th).$$

- (i) ¹ Sei $h \in \mathbb{R}^n$ beliebig, aber fest gewählt. Mit der Konvexität von f folgt auch die Konvexität von g_h ; betrachte dazu für $t_1, t_2 \in \mathbb{R}$, $\lambda \in [0, 1]$

$$\begin{aligned} g_h(\lambda t_1 + (1 - \lambda)t_2) &= f((\lambda + (1 - \lambda))x + (\lambda t_1 + (1 - \lambda)t_2)h) \\ &= f(\lambda(x + t_1 h) + (1 - \lambda)(x + t_2 h)) \\ &\leq \lambda f(x + t_1 h) + (1 - \lambda)f(x + t_2 h) = \lambda g_h(t_1) + (1 - \lambda)g_h(t_2). \end{aligned}$$

Dann gilt für $t, t_1, t_2 \in \mathbb{R}$ mit $t_1 < t < t_2$ wie aus Analysis I bekannt

$$\frac{g(t) - g(t_1)}{t - t_1} \leq \frac{g(t_2) - g(t_1)}{t_2 - t_1} \leq \frac{g(t_2) - g(t)}{t_2 - t}.$$

Diese Ungleichung zeigt, dass der Differenzenquotient monoton wachsend ist (fixiere dazu jeweils t_1, t_2 beziehungsweise t). Damit ist insbesondere $\frac{g(t) - g(0)}{t}$ für $t \searrow 0$ monoton fallend und durch $\frac{g(0) - g(-1)}{-1}$ nach unten beschränkt. Somit existiert

$$\lim_{t \rightarrow 0} \frac{g(t) - g(0)}{t} = \lim_{t \rightarrow 0} \frac{f(x + th) - f(x)}{t} = f'(x, h).$$

- (ii) Sei $h \in \mathbb{R}^n$ und $\lambda > 0$, dann gilt mit $s = \lambda t$

$$\begin{aligned} f'(x, \lambda h) &= \lim_{t \rightarrow 0} \frac{f(x + t\lambda h) - f(x)}{t} = \lambda \lim_{t \rightarrow 0} \frac{f(x + t\lambda h) - f(x)}{\lambda t} \\ &= \lambda \lim_{s \rightarrow 0} \frac{f(x + sh) - f(x)}{s} = \lambda f'(x, h). \end{aligned}$$

- (iii) Seien $h_1, h_2 \in \mathbb{R}^n$. Mit (ii) gilt dann

$$\begin{aligned} f'(x, h_1 + h_2) &= \lim_{t \rightarrow 0} \frac{f(x + t(h_1 + h_2)) - f(x)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f((\frac{1}{2} + \frac{1}{2})x + \frac{1}{2}2th_1 + \frac{1}{2}2th_2) - (\frac{1}{2} + \frac{1}{2})f(x)}{t} \\ &\leq \lim_{t \rightarrow 0} \frac{\frac{1}{2}f(x + 2th_1) - \frac{1}{2}f(x) + \frac{1}{2}f(x + 2th_2) - \frac{1}{2}f(x)}{t} \\ &= \frac{1}{2}f'(x, 2h_1) + \frac{1}{2}f'(x, 2h_2) \\ &= f'(x, h_1) + f'(x, h_2). \end{aligned}$$

- (iv) Als Gegenbeispiel untersuchen wir die euklidische Norm $\|\cdot\| : \mathbb{R}^2 \rightarrow \mathbb{R}$ und ihre Richtungsableitung in Null. Dann ist

$$f'(0, e_1) = 1 \quad \text{und} \quad f'(0, e_2) = 1 \quad \text{sowie} \quad f'(0, e_1 + e_2) = \sqrt{2}.$$

¹vgl. http://www.mathematik.uni-dortmund.de/~tdohnal/TEACH/Seminar_AnalIII_SS2013/Strickmann_Konvexe_Fkt.pdf; S.4

Aufgabe 2.2

Gegeben sei $z \in \mathbb{R}^n$. Wir untersuchen zu $\lambda \geq 0$

$$f_\lambda: \mathbb{R}^n \rightarrow \mathbb{R}, \quad f_\lambda(x) = \frac{1}{2} \|x - z\|_2^2 + \lambda \|x\|_1 = \frac{1}{2} \sum_{k=1}^n |x_k - z_k|^2 + \lambda \sum_{k=1}^n |x_k|.$$

- (i) Wir zeigen, dass für jedes $\lambda \geq 0$ ein eindeutiges $x_\lambda = \arg \min_{x \in \mathbb{R}^n} f_\lambda(x)$ gibt. Sei also $\lambda \geq 0$ beliebig. Wir zeigen, dass f_λ koerziv und streng konvex ist. Mit

$$f_\lambda(x) = \frac{1}{2} \|x - z\|_2^2 + \lambda \|x\|_1 \geq \frac{1}{2} \|x - z\|_2^2 \geq \|x\|_2^2 - \|x\|_2 \|z\|_2 + \|z\|_2^2 \geq \|x\|_2 (\|x\|_2 - \|z\|_2)$$

folgt die Koerzitivität von f_λ .

Für den Nachweis der Konvexität seien $x, y \in \mathbb{R}^n$, $x \neq y$ und $t \in (0, 1)$. Mit der Parallelogrammgleichung gilt

$$\begin{aligned} f_\lambda(tx + (1-t)y) &= \frac{1}{2} \|tx + (1-t)y - z\|_2^2 + \lambda \|tx + (1-t)y\|_1 \\ &\leq \frac{1}{2} \|tx + (1-t)y - tz - (1-t)z\|_2^2 + \lambda(t\|x\|_1 + (1-t)\|y\|_1) \\ &= \frac{1}{2} \|t(x-z) + (1-t)(y-z)\|_2^2 + \lambda(t\|x\|_1 + (1-t)\|y\|_1) \\ &= \|t(x-z)\|_2^2 + \|(1-t)(y-z)\|_2^2 \\ &\quad - \frac{1}{2} \|t(x-z) - (1-t)(y-z)\|_2^2 + \lambda(t\|x\|_1 + (1-t)\|y\|_1) \\ &\leq \|t(x-z)\|_2^2 + \|(1-t)(y-z)\|_2^2 + \lambda(t\|x\|_1 + (1-t)\|y\|_1) \\ &= t^2 \|x-z\|_2^2 + \lambda t\|x\|_1 + (1-t)^2 \|(y-z)\|_2^2 + \lambda(1-t)\|y\|_1 \\ &< \frac{t}{2} \|x-z\|_2^2 + \lambda t\|x\|_1 + \frac{(1-t)}{2} \|(y-z)\|_2^2 + \lambda(1-t)\|y\|_1 \\ &= tf_\lambda(x) + (1-t)f_\lambda(y). \end{aligned}$$

Insgesamt ist f_λ also streng konvex und somit insbesondere stetig sowie koerziv und besitzt damit einen eindeutigen Minimierer x_λ .

(ii)

Aufgabe 2.3

Sei $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear und $b \in \mathbb{R}^m$. Wir betrachten die Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f(x) = \frac{1}{2} \|Ax - b\|^2 = \frac{1}{2} \langle Ax - b, Ax - b \rangle = \frac{1}{2} (\langle A^T A x, x \rangle - 2 \langle A^T x, b \rangle + \langle b, b \rangle).$$

- (i) Dann ist die Ableitung in $x \in \mathbb{R}^n$, wie aus Analysis II bekannt, gegeben durch

$$f'(x) = A^T A x - A^T b$$

gegeben. Damit ergibt sich der Gradient

$$\nabla f(x) = f'(x)^T = x^T A^T A - b^T A^T.$$

- (ii) Ist $\bar{x} \in \mathbb{R}^n$ ein Minimierer von f , dann ist insbesondere die notwendige Optimalitätsbedingung erfüllt, d.h. es gilt $f'(\bar{x}) = 0$ und somit $A^T A \bar{x} = A^T b$. Wie aus Numerik bekannt, ist $A^T A$ eine symmetrische, positiv semi-definite Matrix. Wegen $f''(x) = A^T A$ für alle $x \in \mathbb{R}^n$ ist f also konvex. Damit ist die Erfüllung der notwendigen Optimalitätsbedingung bereits hinreichend für Minimalität in einem Punkt.
- (iii) Nach (i) genügt es zu zeigen, dass $A^T b \in \text{im } A^T A$ ist. Wir zeigen sogar $\text{im } A^T = \text{im } A^T A$. Dabei ist die Inklusion $\text{im } A^T \supset \text{im } A^T A$ klar. Für $x \in \mathbb{R}^n$ gilt weiterhin

$$\langle A^T A x, x \rangle = \langle A x, A x \rangle$$

d.h. $x \in \ker A^T A$ genau dann, wenn $x \in \ker A$ ist, kurz $\ker A^T A = \ker A$. Damit ist

$$\text{rank } A^T A = \text{rank } A = \text{rank } A^T.$$

Daher muss bereits $\text{im } A^T = \text{im } A^T A$ gelten, d.h. es gibt eine Lösung $\bar{x} \in \mathbb{R}^n$ von $A^T A x = A^T b$.

- (iv) Nach obiger Beobachtung ist $n = \text{rank } A = \text{rank } A^T A$ und somit $A^T A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ebenfalls injektiv und daher sogar bijektiv. Folglich erfüllt nur $x = (A^T A)^{-1} A^T b$ die notwendige Optimalitätsbedingung und ist nach (ii) einziger Minimierer.

Aufgabe 2.4

Gegeben sei eine stetig differenzierbare Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ und ein Punkt $\bar{x} \in \mathbb{R}^n$. Für $d \in \mathbb{R}^n$ definieren wir

$$g_d: \mathbb{R} \rightarrow \mathbb{R}, \quad g_d(t) = f(\bar{x} + td).$$

- (i) Sei f konvex und besitze für jedes $d \in \mathbb{R}^n$ die Funktion g_d in $\bar{x} \in \mathbb{R}$ ein lokales Minimum bei $t = 0$. Mit f ist auch g_d für $d \in \mathbb{R}^n$ differenzierbar und mit der notwendigen Optimalitätsbedingung folgt

$$0 = g'_d(0) = \lim_{t \rightarrow 0} \frac{g_d(t) - g_d(0)}{t} = \lim_{t \rightarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t} = f'(\bar{x}, d).$$

Da f stetig differenzierbar ist, so ist die Ableitung in einem Punkt durch die partiellen Ableitungen gegeben. Insgesamt ist also $f'(\bar{x}) = 0$. Zusammen mit der Konvexität folgt bereits, dass \bar{x} ein Minimierer von f ist.

- (ii) Wir betrachten die stetig differenzierbare Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = (y - x^2)(y - 2x^2).$$

Für $d = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \setminus \{0\}$ ist

$$g_d(t) = (ty - t^2x^2)(ty - 2t^2x^2).$$

Für den Fall $x = 0$ und $y \neq 0$ vereinfacht sich die Funktion sogar zu

$$g_d(t) = t^2y^2 > 0 = g_d(0) \text{ für } t \neq 0$$

und falls $y = 0$ und $x \neq 0$ zu

$$g_d(t) = 2t^4x^4 > 0 = g_d(0) \text{ für } t \neq 0.$$

Andernfalls ist für $|t| < \frac{2|y|}{x^2}$ auch

$$g_d(t) > 0 = g_d(0) \text{ mit } t \neq 0.$$

Insgesamt ist $t = 0$ für jedes $d \in \mathbb{R}^2 \setminus \{0\}$ ein lokaler Minimierer von g_d . Trotzdem ist

$$f(0, 0) = 0 > -1 = f(1, 1).$$