PROSEMINAR:
4. FORTSETZUNG ZU KONVEXEN FUNKTIONEN

Wiederholung

Definition 3.2. Sei $X \subseteq \mathbb{R}^n$ eine konvexe Menge. Eine Funktion $f\colon X \to \mathbb{R}$ heißt
(a) *konvex* (auf X), wenn für alle $x, y \in X$ und alle $\lambda \in (0,1)$ gilt:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y);$$

(b) *strikt konvex* (auf X), wenn für alle $x, y \in X$ mit $x \neq y$ und alle $\lambda \in (0,1)$ gilt:

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y);$$

(c) *gleichmäßig konvex* (auf X), wenn es ein $\mu > 0$ gibt mit

$$f(\lambda x + (1-\lambda)y) + \mu\lambda(1-\lambda)\left\|x-y\right\|^2 \leq \lambda f(x) + (1-\lambda) f(y).$$

für alle $x, y \in X$ und alle $\lambda \in (0,1)$. (Man bezeichnet f dann auch als *gleichmäßig konvex mit Modulus* μ.)

Satz 3.5. *Seien $X \subseteq \mathbb{R}^n$ eine offene und konvexe Menge und $f\colon X \to \mathbb{R}$ stetig differenzierbar. Dann gelten:*
(a) f konvex (auf X) $\Longleftrightarrow \forall x, y \in X : f(x) - f(y) \geq \nabla f(y)^\top (x-y)$.
(b) f strikt konvex (auf X) $\Longleftrightarrow \forall x, y \in X, x \neq y : f(x) - f(y) > \nabla f(y)^\top (x-y)$.
(c) f gleichmäßig konvex (auf X) $\Longleftrightarrow \exists \mu > 0 : f(x) - f(y) \geq \nabla f(y)^\top (x-y) + \mu\left\|x-y\right\|^2 \ \forall x, y \in X$

Definition 3.6. Sei $X \subseteq \mathbb{R}^n$ eine gegebene Menge. Eine Funktion $F\colon X \to \mathbb{R}^n$ heißt
(a) *monoton* (auf X), wenn

$$(x-y)^\top (F(x) - F(y)) \geq 0$$

für alle $x, y \in X$ gilt;
(b) *strikt monoton* (auf X), wenn

$$(x-y)^\top (F(x) - F(y)) > 0$$

für alle $x, y \in X$ mit $x \neq y$ gilt;
(c) *gleichmäßig monoton* (auf X), wenn es ein $\mu > 0$ gibt mit

$$(x-y)^\top (F(x) - F(y)) \geq \mu\left\|x-y\right\|^2$$

für alle $x, y \in X$. (Man bezeichnet F dann auch als *gleichmäßig monoton mit Modulus* μ.)

Satz 3.7. *Sei $X \subseteq \mathbb{R}^n$ eine offene und konvexe Menge und $f\colon X \to \mathbb{R}$ stetig differenzierbar. Dann gelten:*
(a) f konvex $\Longleftrightarrow \nabla f$ monoton
(b) f strikt konvex $\Longleftrightarrow \nabla f$ strikt monoton
(c) f gleichmäßig konvex $\Longleftrightarrow \nabla f$ gleichmäßig monoton.

Satz A.3 (Mittelwertsatz in Integralform)**.** *Sei $F\colon \mathbb{R}^n \to \mathbb{R}^m$ stetig differenzierbar, sowie $x, y \in \mathbb{R}^n$ gegeben. Dann gilt*

$$F(x) = F(y) + \int_0^1 F'(y + \tau(x-y))(x-y)\mathrm{d}\tau.$$

Hier nutzen wir $F(x) = \nabla f(x)$, d.h. es gilt

$$\nabla f(x) = \nabla f(y) + \int_0^1 \nabla^2 f(y + \tau(x-y))(x-y)\mathrm{d}\tau. \tag{1}$$

Fortsetzung

Wir zeigen zuerst ein Resultat über zweimal stetig differenzierbare (strikt, gleichmäßig) konvexe Funktionen unter Zuhilfenahme von Satz 3.7:

Satz 3.8. *Sei $X \subseteq \mathbb{R}^n$ eine offene und konvexe Menge und $f\colon X \to \mathbb{R}^n$ zweimal stetig differenzierbar. Dann gilt*
(a) f konvex (auf X) $\iff$ $\nabla^2 f(x)$ positiv semidefinit für alle $x \in X$
(b) $\nabla^2 f(x)$ positiv definit für alle $x \in X$ $\implies$ f strikt konvex (auf X)
(c) f gleichmäßig konvex (auf X) $\iff$ $\nabla^2 f(x)$ gleichmäßig positiv definit auf X, d.h., wenn es ein $\mu > 0$ gibt mit

$$d^\top \nabla^2 f(x) d \geq \mu \left\|d\right\|^2 \tag{2}$$

für alle $x \in X$ und für alle $d \in \mathbb{R}^n$.

Beweis.

(c) "$\Rightarrow$": Sei f gleichmäßig konvex. Wegen Satz 3.7 *(c)* ist ∇f gleichmäßig monoton, d.h. es existiert ein $\mu > 0$ so, dass für alle $x, y \in X$ gilt

$$(x-y)^\top (\nabla f(x) - \nabla f(y)) \geq \mu \left\|x - y\right\|^2. \tag{3}$$

Da ∇f stetig differenzierbar ist folgt mit einer geeigneten Konstanten $\mu > 0$ also

$$\begin{aligned} d^\top \nabla^2 f(x) d &= d^\top \lim_{t\to 0} \frac{\nabla f(x+td) - \nabla f(x)}{t} \\ &= \lim_{t\to 0} \frac{td^\top(\nabla f(x+td) - \nabla f(x))}{t^2} \\ &\overset{(3)}{\geq} \lim_{t\to 0} \frac{1}{t^2} \mu \left\|td\right\|^2 = \lim_{t\to 0} \mu \left\|d\right\|^2 \\ &= \mu \left\|d\right\|^2 \end{aligned}$$

für alle $x \in X$ und alle $d \in \mathbb{R}^n$, d.h., $\nabla^2 f(x)$ ist gleichmäßig positiv definit (auf X).

"$\Leftarrow$": Angenommen es gilt (2), Aus dem Mittelwertsatz in der Integralform A.3 und der Monotonie des Integrals ergibt sich

$$\begin{aligned} (x-y)^\top (\nabla f(x) - \nabla f(y)) &\overset{(1)}{=} \int_0^1 (x-y)^\top \nabla^2 f(y + \tau(x-y))(x-y)\mathrm{d}\tau \\ &\overset{(2)}{\geq} \mu \int_0^1 \left\|x-y\right\|^2 \mathrm{d}\tau \\ &= \mu \left\|x-y\right\|^2, \end{aligned} \tag{4}$$

d.h., nach Definition 3.6 (c) ∇f ist gleichmäßig monoton auf X. Wegen Satz 3.7 *(c)* ist f selbst daher gleichmäßig konvex auf X.

(a) Folgt aus (c) wenn man $\mu = 0$ setzt.
(b) Sei $\nabla^2 f(z)$ positiv definit für alle $z \in X$. Dann ist $\theta(\tau) := (x-y)^\top \nabla^2 f(y+\tau(x-y))(x-y) > 0$ für alle $\tau \in [0,1]$ und alle $x, y \in X$ mit $x \neq y$. Folglich ist

$$(x-y)^\top (\nabla f(x) - \nabla f(y)) \overset{(1)}{=} \int_0^1 \theta(\tau) \mathrm{d}\tau > 0 \tag{5}$$

für alle $x, y \in X$ mit $x \neq y$, vgl. (4). Also ist ∇f nach Definition 3.6 (b) strikt monoton und somit f selbst strikt konvex aufgrund des Satzes 3.7 *(b)*.

□

Die Aussage des Satzes 3.8 *(b)* ist i.A. nur von hinreichendem Charakter, denn die Funktion$f(x) := x^4$ ist strikt konvex, aber $\nabla^2 f(0) = 0$ ist nur positiv semidefinit auf $X = [0,1] \subseteq \mathbb{R}$:

- Aus $f(x) = x^4$ folgt $\nabla f(x) = 4x^3$ und $\nabla^2 f(x) = 12x^2$. Also gilt $\nabla^2 f(0) = 0$ und f kann nicht positiv definit auf X sein.
- $f(x) = x^4$ ist strikt konvex genau dann, wenn $\nabla f(x)$ strikt monoton ist, d.h., wenn für alle $x, y \in X$ mit $x \neq y$ gilt

$$(x-y)(\nabla f(x) - \nabla f(y)) > 0.$$

O.B.d.A. können wir $x > y$ voraussetzen. Eingesetzt liefert dies

$$\underbrace{(x-y)}_{>0}(4x^3 - 4y^3) > 0$$

was wegen $x^3 > y^3$ erfüllt ist.

Dass Levelmengen von gleichmäßig konvexen Funktionen stets kompakt sind, zeigen wir im folgenden

Lemma 3.9. *Sei $f\colon \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, $x^0 \in \mathbb{R}^n$ beliebig, die Levelmenge*

$$\mathcal{L}(x^0) := \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\} \tag{6}$$

konvex und f gleichmäßig konvex auf $\mathcal{L}(x^0)$. Dann ist die Menge $\mathcal{L}(x^0)$ kompakt.

Beweis. Die Levelmenge $\mathcal{L}(x^0)$ ist nichtleer, da $x^0 \in \mathcal{L}(x^0)$. Die Funktion f ist gleichmäßig konvex auf $\mathcal{L}(x^0)$, d.h., es existiert ein geeignetes $\mu > 0$ mit $\lambda := \frac{1}{2}$ sodass für alle $x \in \mathcal{L}(x^0)$ und $y = x^0$ gilt (nach Einsetzen in die Definition 3.2 (c))

$$\begin{aligned} & f(\tfrac{1}{2}x + \tfrac{1}{2}x^0) + \mu\tfrac{1}{2} \cdot \tfrac{1}{2} \left\|x - x^0\right\|^2 \leq \tfrac{1}{2} f(x) + \tfrac{1}{2} f(x^0) \\ \Leftrightarrow \quad & f(\tfrac{1}{2}(x + x^0)) + \mu\tfrac{1}{4} \left\|x - x^0\right\|^2 \leq \tfrac{1}{2}\big(f(x) + f(x^0)\big) \end{aligned}$$

bzw. umgestellt

$$\begin{aligned} \tfrac{1}{4}\mu \left\|x - x^0\right\|^2 &\leq \tfrac{1}{2}(f(x) + f(x^0)) - f(\tfrac{1}{2}(x + x^0)) \\ &= \tfrac{1}{2}\underbrace{(f(x) - f(x^0))}_{\leq 0, \text{ da } x \in \mathcal{L}(x^0)} - \big(f(\tfrac{1}{2}(x + x^0)) - f(x^0)\big) \\ &\leq -\big(f(\tfrac{1}{2}(x + x^0)) - f(x^0)\big) \\ &\overset{3.5(a)}{\leq} -\tfrac{1}{2}\nabla f(x^0)^\top (x - x^0) \\ &\leq \tfrac{1}{2}\left\|\nabla f(x^0)^\top (x - x^0)\right\| \\ &\leq \tfrac{1}{2}\left\|\nabla f(x^0)\right\| \left\|(x - x^0)\right\| \end{aligned}$$

und daraus folgt

$$\left\|x - x^0\right\| \leq c := \frac{2\left\|\nabla f(x^0)\right\|}{\mu}$$

für alle $x \in \mathcal{L}(x^0)$. Also ist $\mathcal{L}(x^0)$ beschränkt. Aus Stetigkeitsgründen ist die Levelmenge $\mathcal{L}(x^0)$ aber auch abgeschlossen. Nach dem nach dem Satz von Heine-Borel ist $\mathcal{L}(x^0)$ also kompakt. □

Es ist hierbei egal ob man f als gleichmäßig konvex auf dem gesamten $\mathbb{R}^n$ oder auch nur auf einer konvexen Menge $X \subseteq \mathbb{R}^n$ voraussetzt: Es folgt automatisch, dass die Levelmenge $\mathcal{L}(x^0)$ konvex ist. Man könnte in diesem Fall also auf die explizit geforderte Voraussetzung der Konvexität von $\mathcal{L}(x^0)$ im Lemma 3.9 verzichten.