2c. Grundlagen
Konvexe Funktionen

Optimierung SoSe 2020

Dr. Alexey Agaltsov

Plan

- Äquivalente Definitionen der Konvexität
- Konvexitätserhaltende Operationen
- Glatte konvexe Funktionen
- Minima konvexer Funktionen

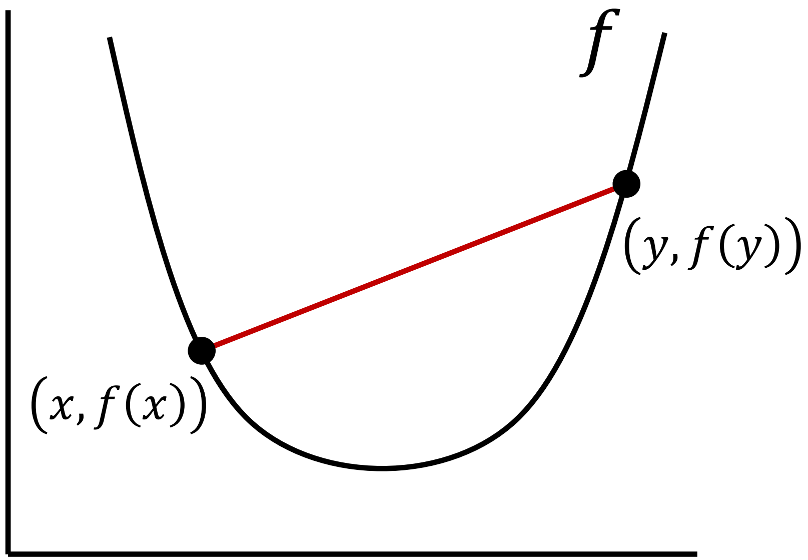
Konvexe Funktionen

- Sei $K \subseteq \mathbb{R}^n$ konvex und $f: K \to \mathbb{R}$. Wir bezeichnen f als konvex falls:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$
$$\forall x, y \in K, \forall \alpha \in (0,1)$$

- Jede Verbindungsstrecke zweier Punkte liegt oberhalb des Graphen

- Gilt die Ungleichung für $x \neq y$ mit $<$, so heißt f strikt konvex



$(y, f(y))$

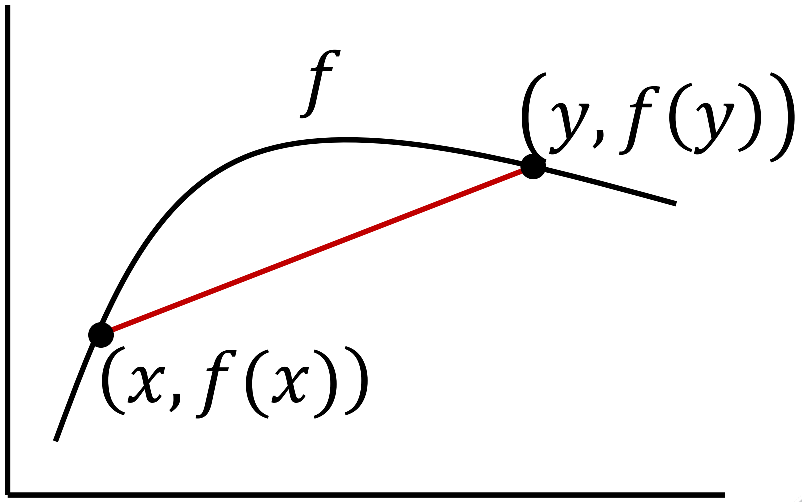
Konkave Funktionen

- Sei $K \subseteq \mathbb{R}^n$ eine konvexe Menge. $f: K \to \mathbb{R}$ heißt konkav falls:

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha) f(y)$$
$$\forall x, y \in K \quad \forall \alpha \in (0,1)$$

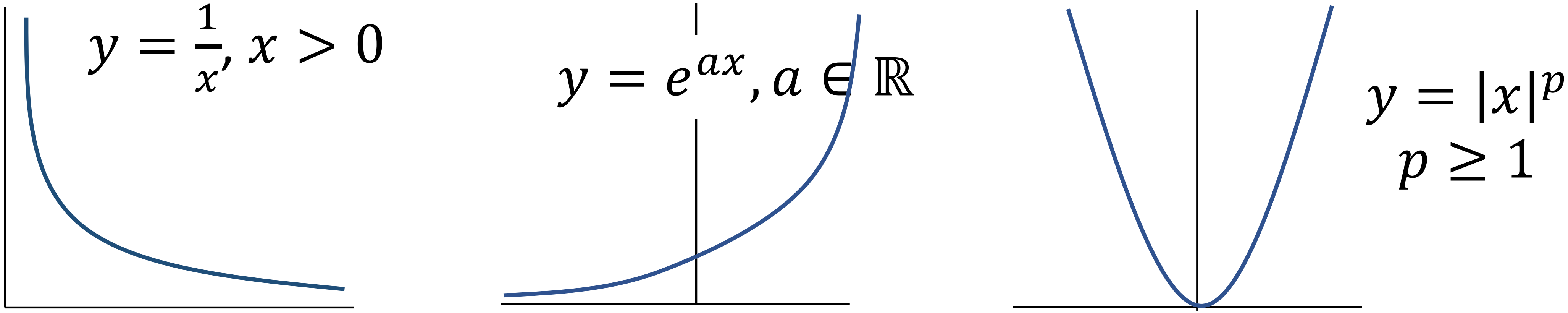
- Jede Sekantenstrecke liegt unterhalb des Graphen

- f ist konkav $\Leftrightarrow -f$ ist konvex

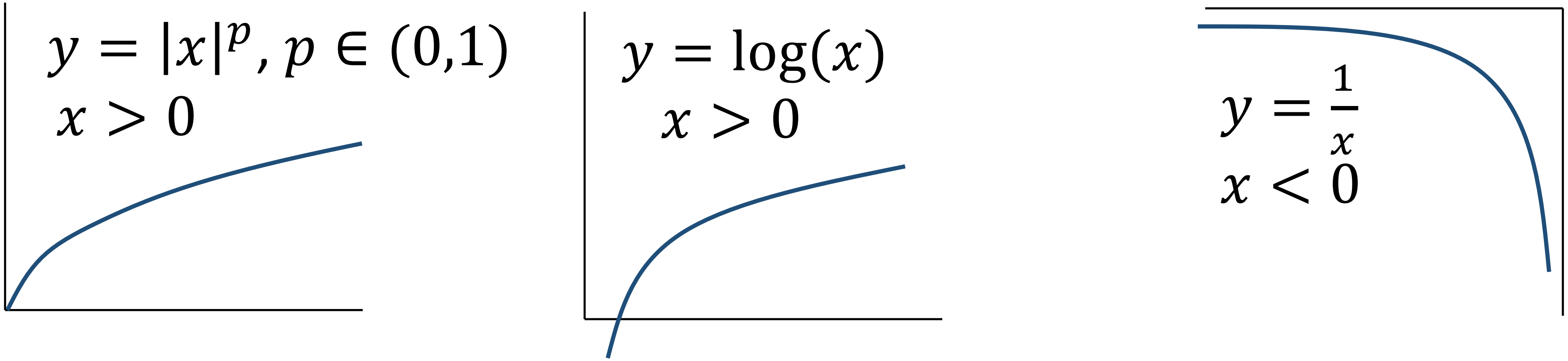


Beispiele in 1D

konvexe Funktionen



konkave Funktionen



Beispiel: Normen

$$f(x) = \|x\|$$

beliebige Norm auf $\mathbb{R}^n$

Seien $x, y \in \mathbb{R}^n$, $\alpha \in (0,1)$

$$\|\alpha x + (1-\alpha)y\| \leq \|\alpha x\| + \|(1-\alpha)y\| \qquad \text{(Dreiecksungleichung)}$$

$$= \alpha\|x\| + (1-\alpha)\|y\| \qquad \text{(Homogenität)}$$

Erweiterte Konvexe Funktionen

- $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ heißt konvex falls:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$$
$$\forall x, y \in \mathbb{R}^n \quad \forall \alpha \in (0,1)$$

- Lemma. $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ ist konvex gdw. $\mathbf{dom}(f)$ und $f|_{\mathbf{dom}(f)}$ konvex sind

$$\mathbf{dom}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}$$

Beweis: $\Rightarrow$

- Sei $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ konvex $\Longrightarrow$ $\mathbf{dom}(f)$ ist konvex

$\longrightarrow$ $f|_{\mathbf{dom}(f)}$ ist konvex

Seien $x, y \in \mathbf{dom}(f), \alpha \in (0,1)$

$$f(\alpha x + (1-\alpha)y) \leq \alpha \underbrace{f(x)}_{<\infty} + (1-\alpha)\underbrace{f(y)}_{<\infty} < \infty$$

$$\alpha x + (1-\alpha)y \in \mathbf{dom}(f)$$

Beweis: $\Leftarrow$

- $\mathbf{dom}(f)$ und $f|_{\mathbf{dom}(f)}$ sind konvex $\Longrightarrow f$ ist konvex

Seien $x, y \in \mathbb{R}^n$ und $\alpha \in (0,1)$

Seien $x \notin \mathbf{dom}(f)$ oder $y \notin \mathbf{dom}(f)$

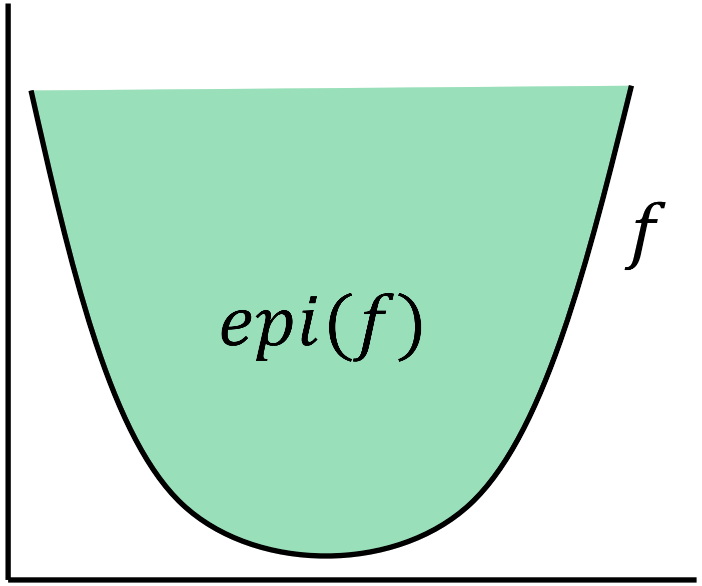
$$\infty = \alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha) y)$$

Seien $x, y \in \mathbf{dom}(f)$

$f|_{\mathbf{dom}(f)}$ ist konvex

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

Epigraph konvexer Funktion



$$\mathbf{epi}(f) = \{(x, z) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq z\}$$

- Lemma. f ist konvex $\Leftrightarrow$ $\mathbf{epi}(f)$ ist konvex

Beweis: $\Rightarrow$

- f ist konvex $\implies$ $\mathbf{epi}(f)$ ist konvex

Seien $(x, z), (y, w) \in \mathbf{epi}(f)$ und $\alpha \in (0,1)$

$$z \geq f(x), w \geq f(y)$$

$(1)\times\alpha + (2)\times(1-\alpha)$

$$\alpha z + (1-\alpha)w \geq \alpha f(x) + (1-\alpha) f(y)$$

f ist konvex

$$\geq f(\alpha x + (1-\alpha)y)$$

$$\Leftrightarrow (\alpha x + (1-\alpha)y, \alpha z + (1-\alpha)w) \in \mathbf{epi}(f)$$

Beweis: $\Leftarrow$

- $\mathbf{epi}(f)$ ist konvex $\Longrightarrow f$ ist konvex

Seien $x, y \in \mathbb{R}^n$, $\alpha \in (0,1)$

Sei $f(x) = \infty$ oder $f(y) = \infty$

$$\infty = \alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha) y)$$

Sei $f(x) < \infty$ und $f(y) < \infty$

$$\big(x, f(x)\big) \in \mathbf{epi}(f), (y, f(y)) \in \mathbf{epi}(f)$$

$\mathbf{epi}(f)$ konvex

$$\big(\alpha x + (1-\alpha) y, \alpha f(x) + (1-\alpha) f(y)\big) \in \mathbf{epi}(f)$$

$$\Leftrightarrow f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$

$\square$

Aufgabe. Einschränkungen auf Geraden

- Seien $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ und $x \in \mathbb{R}^n$, $d \in \mathbb{R}^n \setminus \{0\}$
- Setze $\varphi_{x,d}: \mathbb{R} \to \overline{\mathbb{R}}$, $\varphi_{x,d}(t) = f(x + td)$

Aufgabe. f ist konvex $\Leftrightarrow \varphi_{x,d}$ ist konvex $\forall x \in \mathbb{R}^n, d \in \mathbb{R}^n \setminus \{0\}$

Satz 2.11. Äquiv. Definitionen der Konvexität

Sei $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$. Die folgenden Aussagen sind äquivalent:

1. f ist konvex
2. $\mathbf{dom}\, f$ und $f|_{\mathbf{dom}(f)}$ sind konvex
3. $\mathbf{epi}(f)$ ist konvex
4. $\varphi_{x,d}$ ist konvex $\forall x \in \mathbb{R}^n, d \in \mathbb{R}^n \setminus \{0\}$, wobei:

$$\varphi_{x,d}\colon \mathbb{R} \to \overline{\mathbb{R}}, \quad \varphi_{x,d}(t) = f(x + td)$$

Plan

- Äquivalente Definitionen der Konvexität
- Konvexitätserhaltende Operationen
- Glatte konvexe Funktionen
- Minima konvexer Funktionen

Satz 2.12. Konvexitätserhaltende Operationen

1. $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ konvex $\Rightarrow \lambda f$ konvex $\forall \lambda > 0$
2. $f, g: \mathbb{R}^n \to \overline{\mathbb{R}}$ konvex $\Rightarrow f + g$ konvex
3. $f_\alpha: \mathbb{R}^n \to \overline{\mathbb{R}},\ \alpha \in I$ konvex $\Rightarrow \sup\limits_{\alpha \in I} f_\alpha$ konvex

Beweis: zu 3

Seien $f_\alpha : \mathbb{R}^n \to \overline{\mathbb{R}},\ \alpha \in I$, konvex

$\xrightarrow{\text{Satz 2.11}}$

$\mathbf{epi}\ f_\alpha, \alpha \in I$ sind konvex

$\xrightarrow{\text{Lemma 2.6}}$

$\underbrace{\cap_\alpha \mathbf{epi}(f_\alpha)}_{\mathbf{epi}\left(\sup\limits_\alpha f_\alpha\right)}$ ist konvex

(Lemma 2.2)

$\xrightarrow{\text{Satz 2.11}}$

$\sup\limits_\alpha f_\alpha$ ist konvex

Beispiel 2.13: Maximaler Eigenwert

Sei $f(X)$ der maximale Eigenwert von $X \in \mathbb{S}^n$

$$\mathbb{S}^n \cong \mathbb{R}^m \text{ mit } m = \frac{n(n+1)}{2}$$

Wir erinnern daran, dass:

$$f(X) = \sup\{\underbrace{y^T X y}_{\text{linear in } X} : \|y\|_2 = 1\} \quad \textit{(Satz von Courant-Fischer)}$$

$$\xrightarrow{\text{Satz 2.12}}$$

f ist konvex

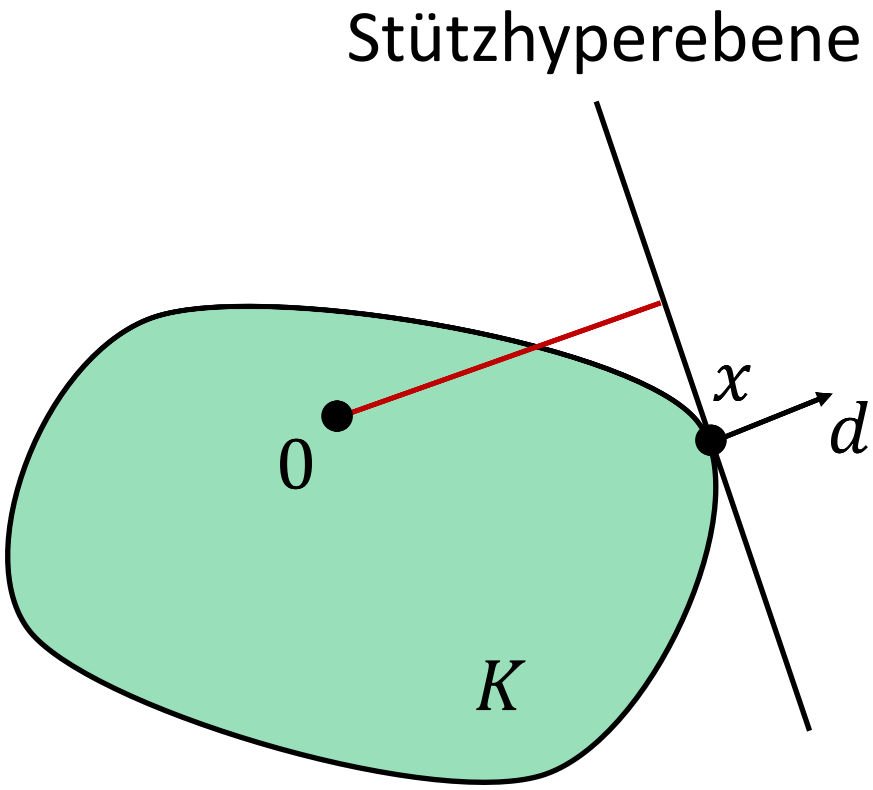
Beispiel 2.14: Stützfunktion

Sei $K \subseteq \mathbb{R}^n$, $K \neq \emptyset$ und $d \in \mathbb{R}^n$

$$S_K(d) := \sup_{x \in K} \underbrace{d^T x}_{\text{linear}}$$

Der Abstand (mit Vorzeichen) zwischen 0 und der Stützhyperebene in Richtung d für $\|d\|_2 = 1$

Nach Satz 2.12. ist $S_K : \mathbb{R}^n \to \mathbb{R}$ konvex



Aufgabe 2.15: Stützfunktion eines Ellipsoides

$$\mathcal{E} = \{x: (x - x_*)^T \Sigma^{-1} (x - x_*) \leq 1\}, \;\; \Sigma \in \mathbb{S}^n_{\succ}$$

$$S_{\mathcal{E}}(d) \quad \left[\begin{array}{l} \text{Maximiere } \; d^T x \;\; \text{über } x \in \mathbb{R}^n \\ \qquad\quad \text{u.d.N. } (x - x_*)^T \Sigma^{-1} (x - x_*) \leq 1 \end{array}\right.$$

Zeigen Sie, dass:

$$S_{\mathcal{E}}(d) = d^T x_* + \left\|\Sigma^{1/2} d\right\|_2$$

Tipp. Zeigen Sie, dass $x \in \mathcal{E}$ gdw. $x = x_* + \Sigma^{1/2} y$ mit $y \in \mathbb{R}^n$, $\|y\|_2 \leq 1$

Satz 2.16. Minimum ist konvex

Sei $f: \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ konvex. Sei $K \subseteq \mathbb{R}^m$ konvex. Setze:

$$g(x) = \inf_{y \in K} f(x, y)$$

Dann gilt eine der folgenden Aussagen:

- $g(x) > -\infty \; \forall x \in \mathbb{R}^n$ und $g: \mathbb{R}^n \to \overline{\mathbb{R}}$ ist konvex
- $g(x) = -\infty \; \forall x \in \mathbb{R}^n$

Beweis

Seien $x_1, x_2 \in \mathbb{R}^n$ und $\alpha \in (0,1)$

Wir werden zeigen, dass:

$$g(\alpha x_1 + (1-\alpha)x_2) \leq \alpha g(x_1) + (1-\alpha)g(x_2)$$

$\forall y_1, y_2 \in K$

$$g(\alpha x_1 + (1-\alpha)x_2) = \inf_{y \in C} f(\alpha x_1 + (1-\alpha)x_2, y))$$

$$\leq f(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2)$$

f ist konvex

$$\leq \alpha f(x_1, y_1) + (1-\alpha)f(x_2, y_2)$$

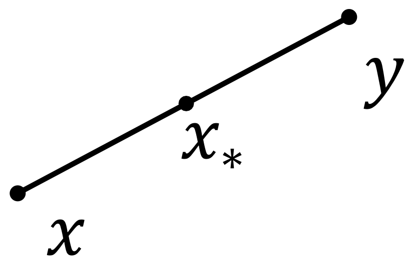
$\inf\limits_{y_1, y_2 \in K}$

Beweis

- $\exists x_*: g(x_*) > -\infty \implies g(x) > -\infty \; \forall x \in \mathbb{R}^n$

Sei $x \in \mathbb{R}^n$

Sei y so, dass $x_* = \frac{x+y}{2}$



$$-\infty < g(x_*) \leq \frac{g(x)+g(y)}{2}$$

Also ist $g(x) > -\infty$

$\square$

Beispiel: Abstand zur konvexen Menge

$$f(y) = \inf_{x \in K} \|x - y\| = \mathbf{dist}(y, K)$$

$K \subseteq \mathbb{R}^n$ ist nichtleer und konvex

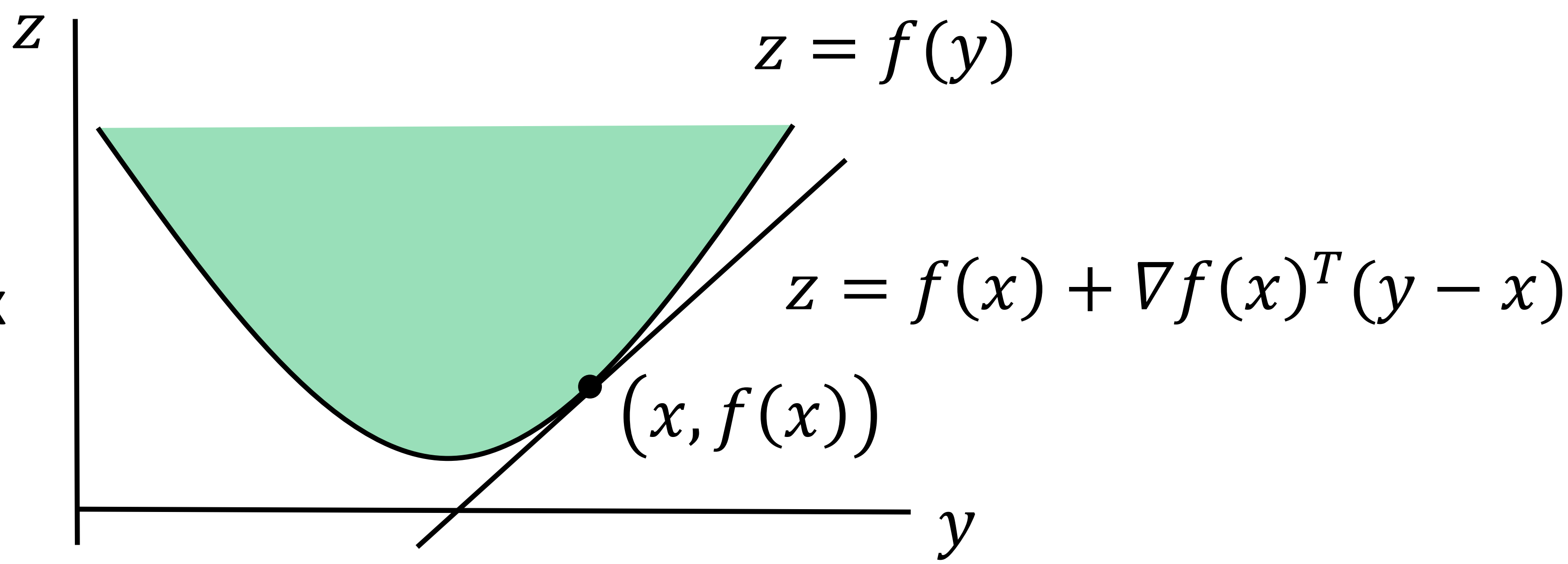
- $g(x, y) = \|x - y\|$ ist konvex auf $\mathbb{R}^n \times \mathbb{R}^n$
- $K \neq \emptyset \implies f(x) < \infty \;\; \forall x \in \mathbb{R}^n$
- Nach Satz 2.16 ist f konvex

Plan

- Äquivalente Definitionen der Konvexität
- Konvexitätserhaltende Operationen
- Glatte konvexe Funktionen
- Minima konvexer Funktionen

Wann ist eine C^1-Funktion konvex?

$f \in C^1$ konvex



Der Graph von f liegt oberhalb Tangente

- Lemma. Sei $U \subseteq \mathbb{R}^n$ offen und konvex. Dann ist $f \in C^1(U)$ konvex gdw.

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \forall x, y \in U$$

Beweis: $\Rightarrow$

Sei $f \in C^1(U)$ konvex

Seien $x, y \in U, \alpha \in (0,1)$

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

$$\frac{f(x+\alpha(y-x))-f(x)}{\alpha} \leq f(y) - f(x)$$

$$\alpha \to 0$$

$$\nabla f(x)^T(y - x) \leq f(y) - f(x)$$

Beweis: $\Leftarrow$

Seien $x, y \in U, \alpha \in (0,1)$

$z := \alpha x + (1-\alpha)y$

$$(x-z)^T \nabla f(z) \leq f(x) - f(z)$$

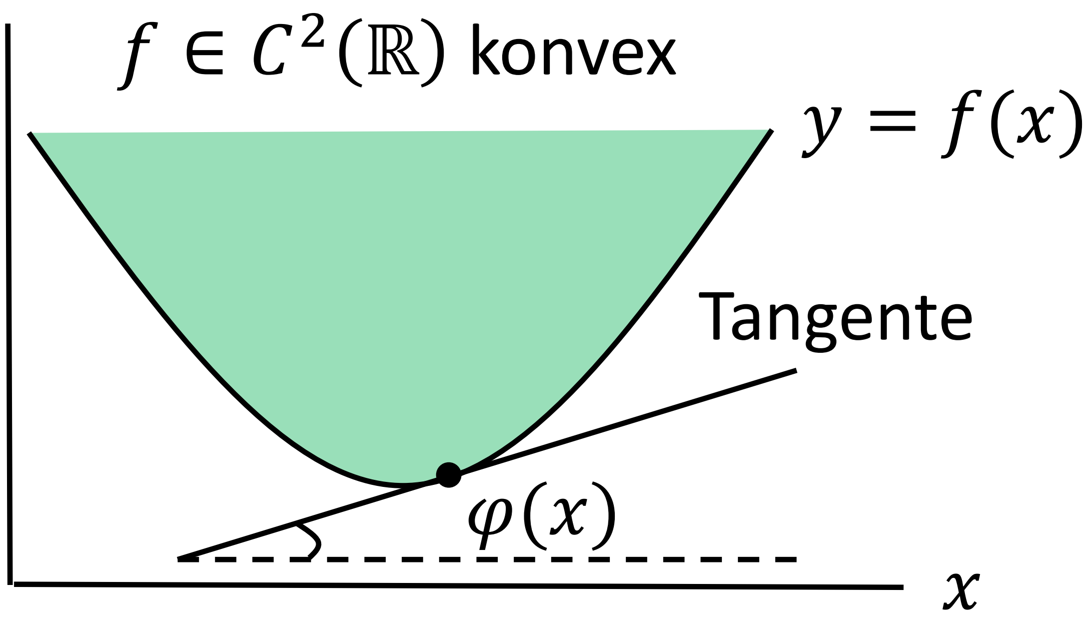
$$(y-z)^T \nabla f(z) \leq f(y) - f(z)$$

$(1)\times\alpha + (2)\times(1-\alpha)$

$$0 \leq \alpha f(x) + (1-\alpha) f(y) - f(z)$$

$\square$

Wann ist eine C^2-Funktion konvex?



Winkel φ ist monotonsteigend

$$\frac{df}{dx} = \tan \varphi$$

$$\frac{d^2 f}{dx^2} = (1 + \tan(\varphi)^2)\ \frac{d\varphi}{dx}$$

$$\frac{d\varphi}{dx} \geq 0 \ \Leftrightarrow\ \frac{d^2 f}{dx^2} \geq 0$$

- Lemma. Sei $U \subseteq \mathbb{R}^n$ offen und konvex. Dann ist $f \in C^2(U)$ konvex gdw.

$$\nabla^2 f(x) \succcurlyeq 0 \quad \forall x \in U$$

Beweis: $\Rightarrow$

Sei $f \in C^2(U)$ konvex

Seien $x \in U, p \in \mathbb{R}^n, t > 0$

letztes Lemma

$$(tp)^T \nabla f(x) \leq f(x+tp) - f(x)$$

$$-(tp)^T \nabla f(x+tp) \leq f(x) - f(x+tp)$$

$$\frac{(1)+(2)}{t^2}$$

$$p^T \frac{\nabla f(x) - \nabla f(x+tp)}{t} \leq 0$$

$t \to 0$

$$-p^T \nabla^2 f(x) p \leq 0$$

Beweis: $\Leftarrow$

Sei $f \in C^2(U)$ mit $\nabla^2 f \succcurlyeq 0$ in U

Seien $x, y \in U$

Taylor-Formel

$\exists \xi \in [0,1]$ so dass:

$$f(y) - f(x) - \nabla f(x)^T(y-x) = \underbrace{\tfrac{1}{2}(y-x)^T \nabla^2 f(x + \xi(y-x))(y-x)}_{\geq 0 \text{ nach Voraussetzung}}$$

$$f(y) \geq f(x) + \nabla f(x)^T(y-x)$$

letztes Lemma

f konvex

$\square$

Satz 2.17. Glatte konvexe Funktionen

Sei $U \subseteq \mathbb{R}^n$ offen und konvex.

1. Sei $f \in C^1(U)$. Dann ist f konvex gdw.
$$f(y) \geq f(x) + \nabla f(x)^T(y-x) \quad \forall x, y \in U$$
2. Sei $f \in C^2(U)$. Dann ist f konvex gdw.
$$\nabla^2 f(x) \succcurlyeq 0 \quad \forall x \in U$$

Beispiel

$$f(x,y) = \frac{x^2}{y},\ x \in \mathbb{R}, y > 0$$

$$\nabla f(x,y) = \begin{bmatrix} \frac{2x}{y} \\ -\frac{x^2}{y^2} \end{bmatrix}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & 2\frac{x^2}{y^3} \end{bmatrix} = \frac{2}{y^3} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix}$$

$$= \frac{2}{y^3} \begin{bmatrix} y \\ x \end{bmatrix}^T \begin{bmatrix} y \\ x \end{bmatrix} \succcurlyeq 0 \quad \Longrightarrow f \text{ ist konvex}$$

Plan

- Äquivalente Definitionen der Konvexität
- Konvexitätserhaltende Operationen
- Glatte konvexe Funktionen
- Minima konvexer Funktionen

Lokales und globales Minimum

$$f: \mathbb{R}^n \to \overline{\mathbb{R}}$$

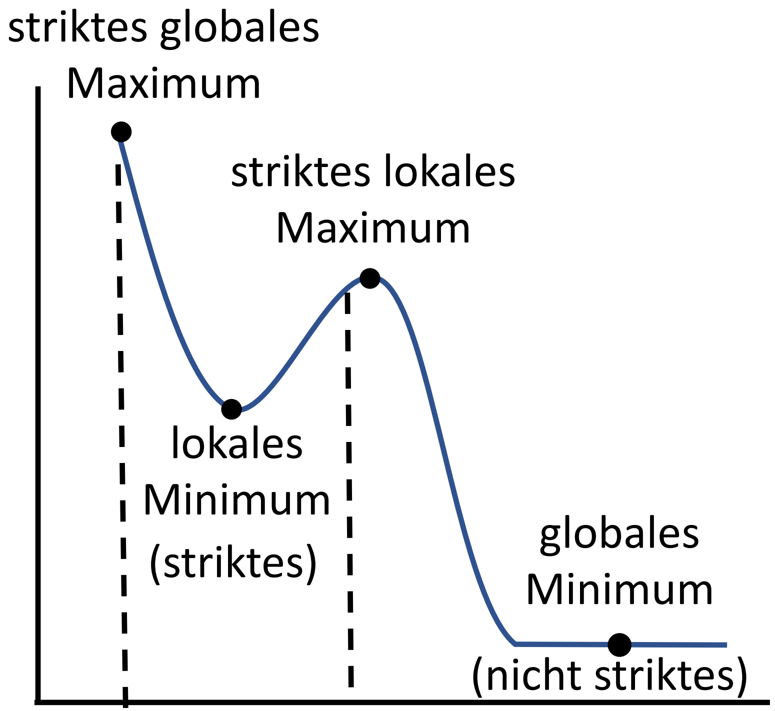
- $x_* \in \mathbb{R}^n$ heißt lokales Minimum von f falls:

$$\exists \varepsilon > 0: \quad f(x_*) \leq f(x) \quad \forall (x: |x - x_*| < \varepsilon)$$

- $x_* \in \mathbb{R}^n$ heißt globales Minimum von f falls:

$$f(x_*) \leq f(x) \quad \forall x \in \mathbb{R}^n$$

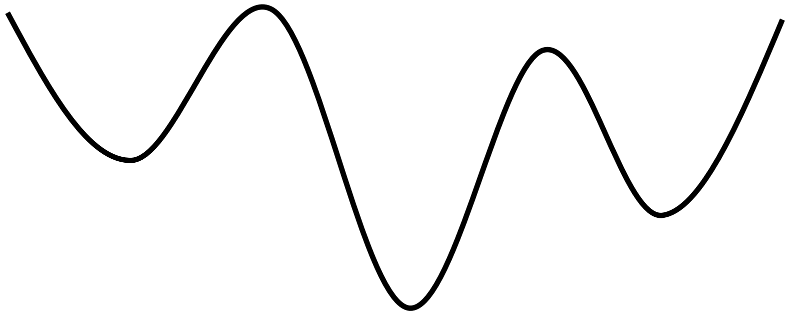
- " $<$ " für $x \neq x_* \Rightarrow$ striktes lokales/globales Minimum
- " $\leq$ " $\to$ " $\geq$ " lokales/globales Maximum



Lokale und globale Minima

$$\text{Minimiere } f(x) \text{ über } x \in \mathbb{R}^n$$
$$f: \mathbb{R}^n \to \overline{\mathbb{R}}$$

- Die lokalen Minima sind einfacher zu bestimmen als die globalen Minima
- Hat f viele lokale Minima, so ist das Problem typescherweise schwer zu lösen



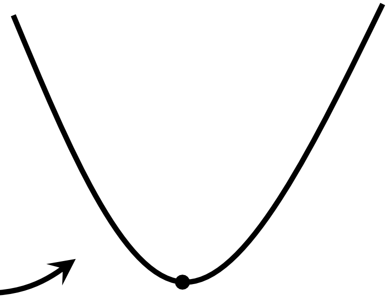
Satz 2.18. Minima konvexer Funktionen

Sei $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ konvex. Dann gilt:

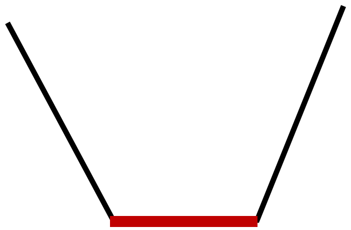
1. Ist x_* lokales Minimum, so ist x_* ein globales Minimum
2. $\operatorname{Argmin}(f)$ ist konvex

$$\operatorname{Argmin}(f) := \left\{x_*: f(x_*) = \inf_{x \in \mathbb{R}^n} f(x)\right\}$$

3. Ist f strikt konvex, so gibt es höchstens ein Minimum



x_*



$\operatorname{Argmin}(f)$

Beweis

Behauptung: f ist konvex $\Longrightarrow$ Jedes lokales Minimum ist global

Sei x_* ein lokales Minimum

$$\exists \varepsilon > 0:\ f(x_*) \leq f(x)\ \ \forall x \in \underbrace{B_\varepsilon(x_*)}_{\{x \in \mathbb{R}^n :\, \|x - x_*\|_2 \leq \varepsilon\}}$$

Angenommen $\exists y: f(y) < f(x_*)$

Sei $\alpha \in (0,1)$: $\alpha x_* + (1-\alpha)y \in B_\varepsilon(x_*)$

$$f(\alpha x_* + (1-\alpha)y) \leq \alpha f(x_*) + (1-\alpha) f(y)$$
$$< \alpha f(x_*) + (1-\alpha) f(x_*) = f(x_*)$$

Widerspruch!

Beweis

Behauptung: f ist konvex $\Longrightarrow \operatorname{Argmin}(f)$ ist konvex

Seien $x, y \in \operatorname{Argmin}(f), \alpha \in (0,1)$

$$f(\alpha x + (1-\alpha)y) \leq \alpha \underbrace{f(x)}_{\inf f} + (1-\alpha)\underbrace{f(y)}_{\inf f} = \inf f$$

$$\Longrightarrow \alpha x + (1-\alpha)y \in \operatorname{Argmin}(f)$$

Beweis

Behauptung: f ist strikt konvex $\Longrightarrow$ es gibt höchstens ein Minimum

Angenommen, $x, y \in \mathrm{Argmin}(f)$

$$x \neq y$$

$$f\left(\frac{x+y}{2}\right) < \frac{1}{2}\underbrace{f(x)}_{\inf f} + \frac{1}{2}\underbrace{f(y)}_{\inf f} = \inf f$$

Wiederspruch!

Zusammenfassung

- Äquivalente Definitionen der Konvexität
- Konvexitätserhaltende Operationen
- Glatte konvexe Funktionen
- Minima konvexer Funktionen

Nächstes Video

- 2d. Grundlagen: Optimalitätskriterien