2d. Grundlagen
Optimalitätsbedingungen

Optimierung SoSe 2020

Dr. Alexey Agaltsov

Plan

- Allgemeine Optimalitätsbedingungen
- Optimalitätsbedingungen für konvexe Programme
- Starke Konvexität

Optimalitätsbedingungen

$$\text{Minimiere} \quad f(x)$$
$$U \subseteq \mathbb{R}^n \text{ offen, } f \in C^1(U)$$

- Wie kann man die optimalen Lösungen erkennen?
- Notwendige Optimalitätsbedingungen (NOB):

  Sei x_* ein lokales/globales Minimum, so gilt …

- Hinreichende Optimalitätsbedingungen (HOB):

  Erfülle $x_* \in U$ …, so ist x_* ein lokales/globales Minimum

Minimiere $f(x)$

Bemerkung: Maximierungsprobleme

- Der Einfachheit halber werden wir die Theorie nur für die Minimierungsprobleme entwickeln
- Notwendige Modifikationen für Maximierungsprobleme sind geradlinig
- In Beispielen wird eventuell Maximierung von Funktionen auftreten

Erinnerung: Taylor-Formel

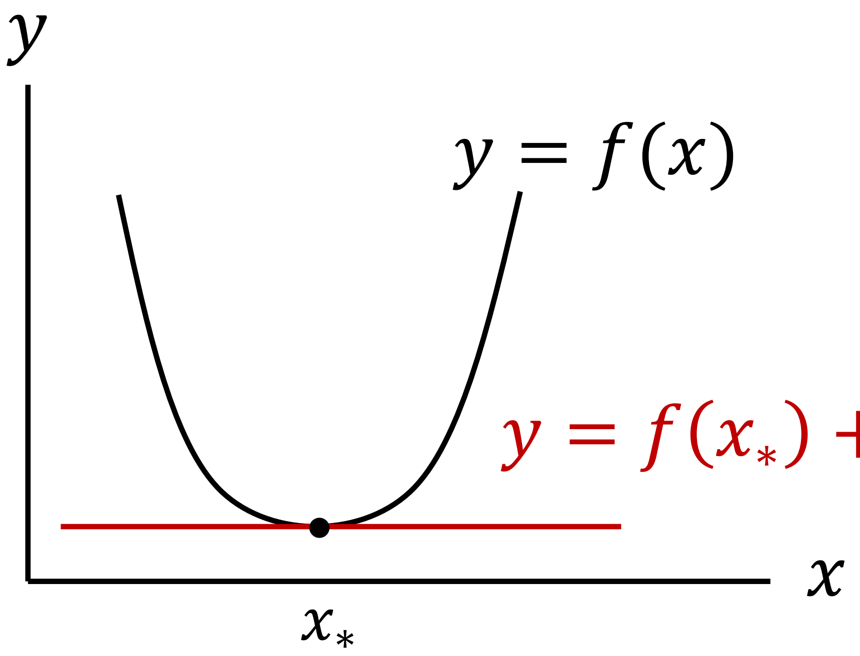
- Seien $f \in C^1(U)$, U offen und konvex, $x, x+p \in U$. Dann:

  $$\exists \xi \in [0,1]:\ \ f(x+p) = f(x) + \underbrace{\nabla f(x+\xi p)^T p}_{\text{Restglied nach Lagrange}}$$

- Ist weiterhin $f \in C^2(U)$, so gilt:

1. $$\exists \eta \in [0,1]:\ \ f(x+p) = f(x) + \nabla f(x)^T p + \overbrace{p^T \nabla^2 f(x+\eta p) p}$$
2. $$\nabla f(x+p) = \nabla f(x) + \underbrace{\int_0^1 \nabla^2 f(x+tp) p\, dt}_{\text{Restglied in Integralform}}$$

Wann ist x_* ein lokales Minimum von $f \in C^1$?



$$y = f(x_*) + \underbrace{\nabla f(x_*)^T}_{0}(x - x_*)$$

Tangentialebene bei x_* ist horizontal

- x_* heißt stationärer Punkt von f falls $\nabla f(x_*) = 0$
- Lemma. Sei $U \subseteq \mathbb{R}^n$ offen, $f \in C^1(U)$. Ist $x_* \in U$ ein lokales Minimum so gilt $\nabla f(x_*) = 0$

Beweis

$$\exists \varepsilon > 0: \overbrace{B_\varepsilon(x_*)}^{\{x \in \mathbb{R}^n : \|x - x_*\|_2 < \varepsilon\}} \subseteq U, f(x_*) \leq f(x) \; \forall x \in B_\varepsilon(x_*)$$

Seien $d \in \underbrace{S_1(0)}_{\{x \in \mathbb{R}^n : \|x\|_2 = 1\}}, t \in (0, \varepsilon)$

$\exists \xi \in [0,1]$
Taylor-Formel

$$\underbrace{f(x_* + td) - f(x_*)}_{\geq 0} = t\nabla f(x_* + \xi t d)^T d$$

$$\nabla f(x_* + \xi t d)^T d \geq 0$$

$t \to 0$
$f \in C^1(U)$

$$\nabla f(x_*)^T d \geq 0$$

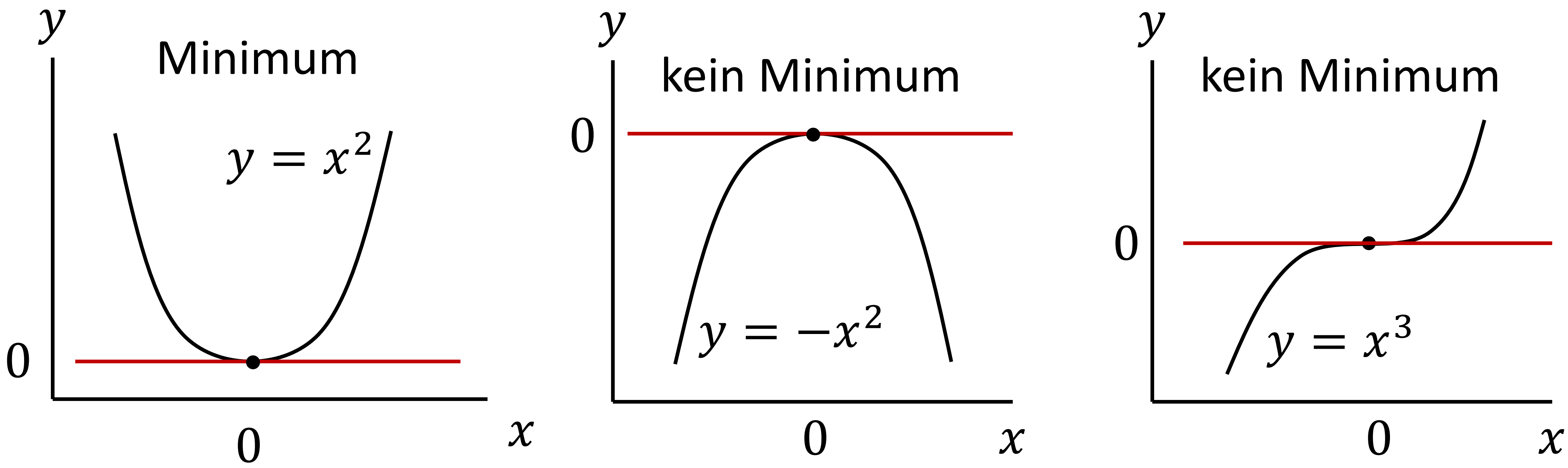
$d = -\nabla f(x_*)/\|\nabla f(x_*)\|_2$

$$-\|\nabla f(x_*)\|_2 = 0$$

$\square$

Wann ist x_* ein lokales Minimum von $f \in C^2$?

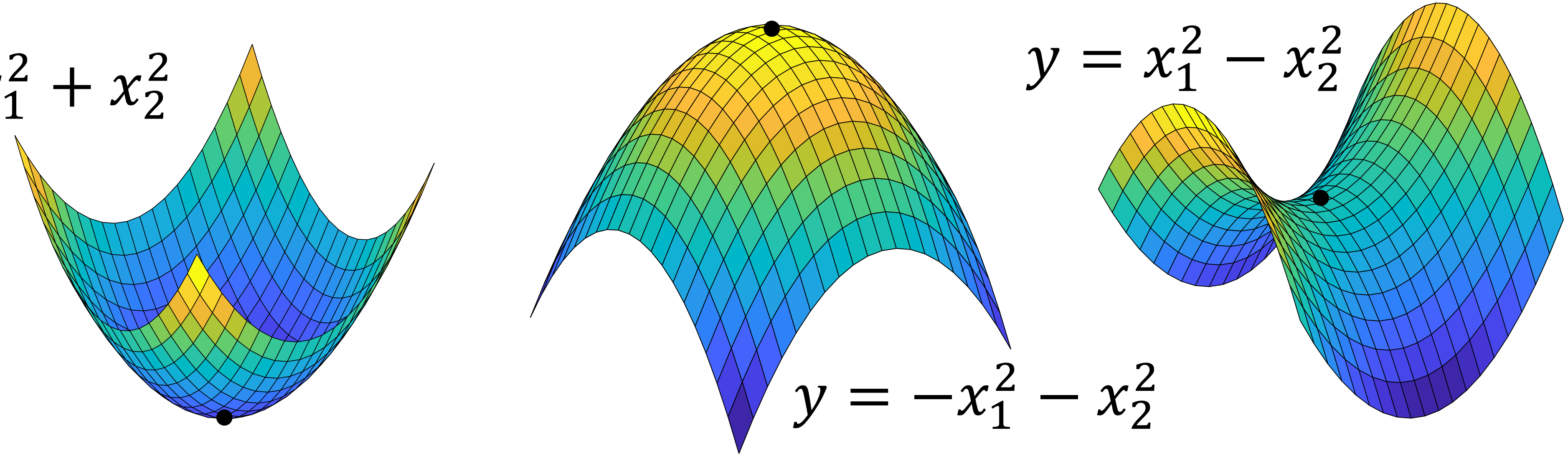
Ist $\nabla f(x_*) = 0$, so kann x_* kein lokales Minimum sein



$f''(0) > 0$ $\quad$ $f''(0) < 0$ $\quad$ $f''(0) = 0$

Für $f \in C^2(U)$ kann man diese Fälle unterscheiden !

Wann ist x_* ein lokales Minimum von $f \in C^2$?



$y = x_1^2 + x_2^2$ $\qquad$ $y = -x_1^2 - x_2^2$ $\qquad$ $y = x_1^2 - x_2^2$

$\nabla^2 f(x_*) \succ 0$ $\qquad$ $\nabla^2 f(x_*) \prec 0$ $\qquad$ $\nabla^2 f(x_*)$ indefinit

Lemma. Sei $U \subseteq \mathbb{R}^n$ offen, $f \in C^2(U)$, $x_* \in U$ ein Punkt mit $\nabla f(x_*) = 0$

1. Sei x_* ein lokales Minimum, so gilt $\nabla^2 f(x_*) \succcurlyeq 0$
2. Gelte $\nabla^2 f(x_*) \succ 0$, so ist x_* ein striktes lokales Minimum

Beweis: x_* ein lokales Min. $\Longrightarrow \nabla^2 f(x_*) \succcurlyeq 0$

$$\exists \varepsilon > 0\colon\ B_\varepsilon(x_*) \subseteq U, f(x_*) \leq f(x)\ \forall x \in B_\varepsilon(x_*)$$

Seien $d \in S_1(0)$, $t \in (0, \varepsilon)$

$\exists \eta \in [0,1]$
Taylor-Formel

$$\underbrace{f(x_* + td) - f(x_*)}_{\geq 0} = t\underbrace{\nabla f(x_*)^T}_{0} d + \frac{1}{2}t^2 d^T \nabla^2 f(x_* + \eta t d)d$$

$$d^T \nabla^2 f(x_* + \xi t d)d \geq 0$$

$t \to 0$
$f \in C^2(U)$

$$d^T \nabla^2 f(x_*)d \geq 0$$

Also $\nabla^2 f(x_*) \succcurlyeq 0$

Beweis: zu 2

Behauptung. $\nabla f(x_*) = 0, \nabla^2 f(x_*) \succ 0 \Longrightarrow$ ist x_* ein striktes lokales Minimum

$$\exists \varepsilon > 0\colon \quad \nabla^2 f(x) \succ 0 \quad \forall x \in B_\varepsilon(x_*)$$

Sei $d \in B_\varepsilon(0) \backslash \{0\}$:

$\exists \eta \in [0,1]$

Taylor-Formel

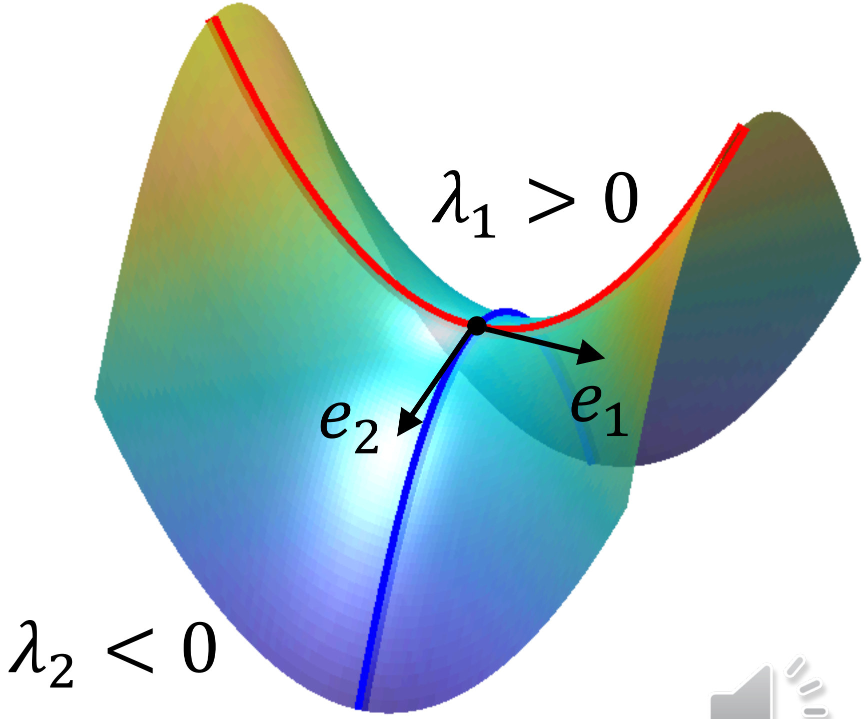
$$f(x_* + d) = f(x_*) + \underbrace{\nabla f(x_*)^T}_{0} d + \underbrace{d^T \nabla^2 f(x_* + \eta d) d}_{> 0}$$

$$f(x) > f(x_*) \quad \forall x \in B_\varepsilon(x_*) \setminus \{x_*\}$$

Geometrische Interpretation in 2D

- $\nabla^2 f(x, y)$ ist die Matrix der zweiten Fundamentalform des Graphen

$z = f(x, y)$



Hauptkrümmungsrichtungen

- $\nabla^2 f(x, y) e_i = \lambda_i e_i$ für $i = 1{,}2$ mit $e_i \neq 0$

  λ_1, λ_2 sind Hauptkrümmungen

  e_1, e_2 sind Hauptkrümmungsrichtungen

Satz 2.19. Optimalitätsbedingungen

Sei U offen, $f \in C^1(U)$. Dann gelten die folgenden Aussagen:

- Ist $x_* \in U$ ein lokales Minimum von f, so gilt $\nabla f(x_*) = 0$
  Ist weiterhin $f \in C^2(U)$, so gilt $\nabla^2 f(x_*) \succcurlyeq 0$
- Ist $x_* \in U$ ein Punkt mit $\nabla f(x_*) = 0, \nabla^2 f(x_*) \succ 0$, so ist x_* ein striktes lokales Minimum von f

Notwendige Optimalitätsbedingungen (NOB)

Mithilfe der NOB kann man das globale Minimum von f finden:

- Bestimme alle Punkte die NOB erfüllen
- Wähle von denen den Punkt x_* mit minimalem Zielfunktionswert
- Nimmt f ihr Minimum an, so ist x_* ein Minimum von f

Beispiel: Nichtexistenz von Lösungen

- Vorsicht: Zuerst muss man die Existenz von Lösungen feststellen

Minimiere $f(x) = x^2 - x^4$

$$f'(x) = 2x - 4x^3$$
$$= 3x(1 - \sqrt{2}\,x)(1 + \sqrt{2}\,\mathrm{x})$$

$$f''(x) = 2 - 12x^2$$

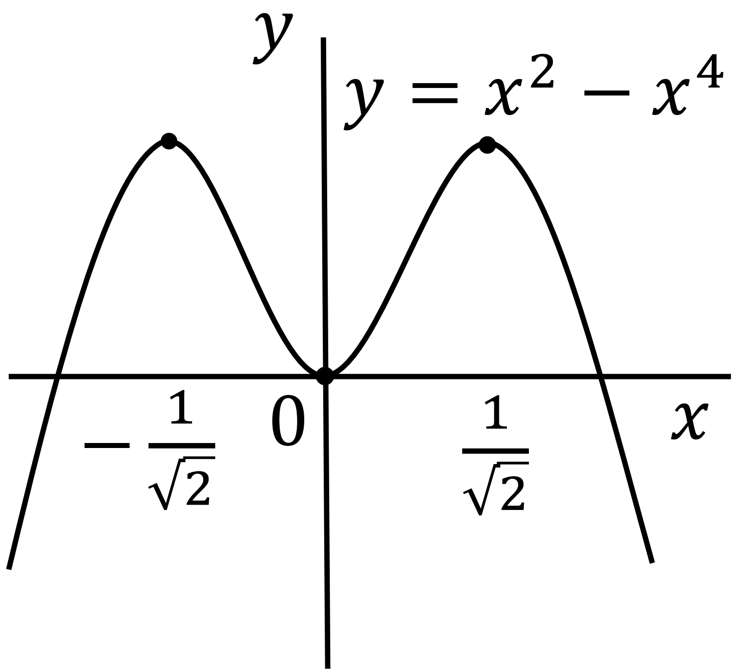
Beispiel: Nichtexistenz von Lösungen

$f'(x) = 0$ für $x \in \{0, \pm 1/\sqrt{2}\}$

$f''(0) = 2$ lokales Minimum

$f''(\pm 1/\sqrt{2}) = -4$ lokale Maxima

$f(x) \to -\infty$ für $x \to \pm\infty$



Plan

- Allgemeine Optimalitätsbedingungen
- Optimalitätsbedingungen für konvexe Programme
- Starke Konvexität

Optimierung über konvexe Mengen

Minimiere $f(x)$ über $x \in K$

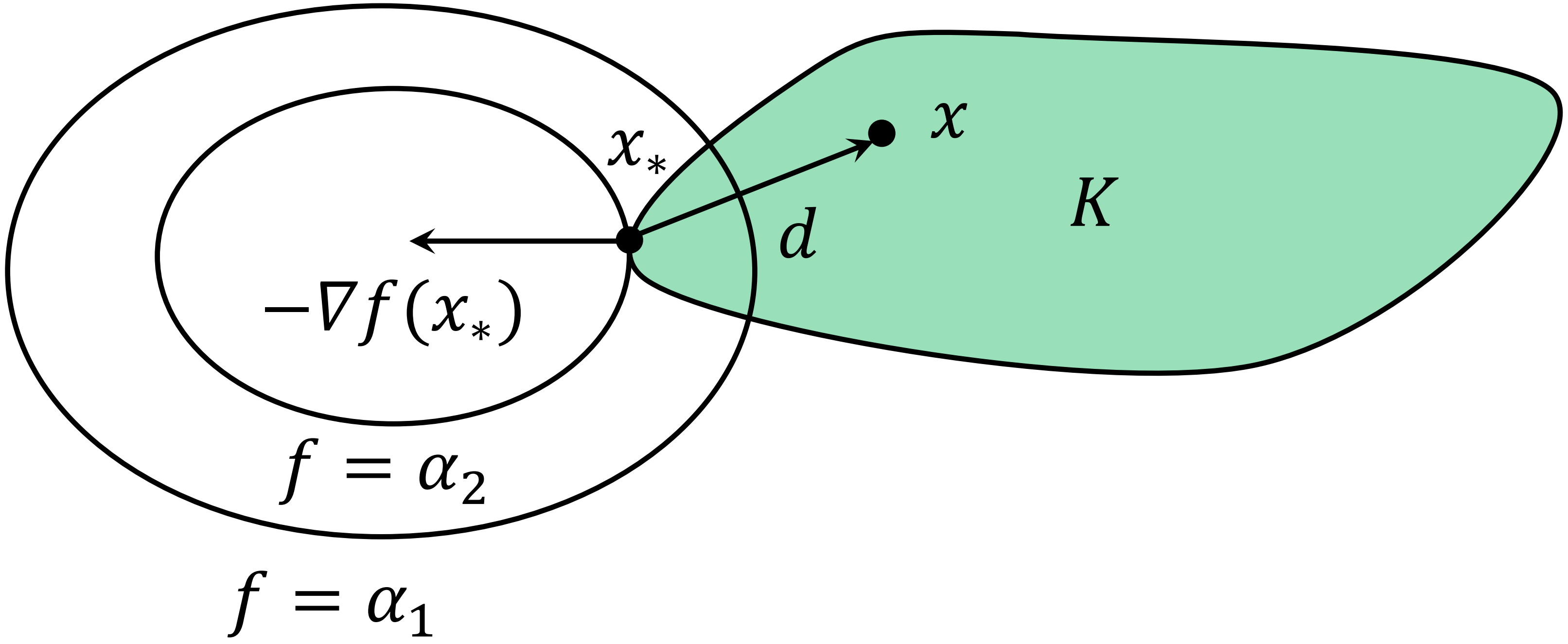
$f \in C^1(\mathbb{R}^n)$, $K \subseteq \mathbb{R}^n$ konvex

- Wir nehmen an, x_* ist eine optimale Lösung
- Ist x_* ein innerer Punkt von K, so gilt nach Satz 2.19:

$$\nabla f(x_*) = 0$$

- Was ist die Optimalitätsbedingung für ein allgemeines x?

Optimalitätsbedingung für konvexe Funktionen



$$-\nabla f(x_*)^T(x - x_*) \leq 0 \;\; \forall x \in K$$

Satz 2.20. Optimalitätsbedingungen für konvexe Funktionen

Sei $K \subseteq \mathbb{R}^n$ konvex und $f \in C^1(\mathbb{R}^n)$

- Ist $x_* \in K$ ein lokales Minimum von f auf K, so gilt:

$$-\nabla f(x_*)^T(x - x_*) \leq 0 \quad \forall x \in K$$

- Seien f konvex und $x_* \in K$ ein Punkt, der die obige Bedingung erfüllt
  Dann ist x_* ein globales Minimum von f auf K

Beweis: zu 1

Sei x_* ein lokales Minimum von f auf K

Seien $x \in K$ und $t \in (0,1)$

$\exists \xi \in [0,1]$

Taylor-Formel

$$\underbrace{f\big(x_* + t(x - x_*)\big) - f(x_*)}_{\geq 0 \text{ für } t \in (0,\varepsilon)} = t\nabla f\big(x_* + \xi t(x - x_*)\big)^T (x - x_*)$$

$$\nabla f\big(x_* + \xi t(x - x_*)\big)^T (x - x_*) \geq 0 \qquad t \in (0,\varepsilon)$$

lasse $t \to 0$

$f \in C^1(\mathbb{R}^n)$

$$\nabla f(x_*)^T (x - x_*) \geq 0$$

Beweis: zu 2

Sei f konvex und $x_* \in K$ ein Punkt, sodass:

$$-\nabla f(x_*)^T(x - x_*) \leq 0 \quad \forall x \in K$$

Satz 2.17

$$f(x) - f(x_*) \geq \overbrace{\nabla f(x_*)^T(x - x_*)}^{\geq 0}$$

$$f(x) \geq f(x_*)$$

x ist beliebig

x_* ist ein globales Minimum von f auf K

$\square$

Aufgabe 2.21. Quadratische Programme

$$\text{Minimiere} \quad f(x) = \frac{1}{2} x^T Q x - c^T x + r \quad \text{über } x \in \mathbb{R}^n$$

$$Q \in \mathbb{S}^n_{\succcurlyeq}, \mathrm{c} \in \mathbb{R}^n, r \in \mathbb{R}$$

Beweisen Sie die folgenden Aussagen:

- Ist f nach unten beschränkt, so gibt es eine optimale Lösung
- $x_* \in \mathbb{R}^n$ ist eine optimale Lösung $\Leftrightarrow Q x_* = c$

Plan

- Allgemeine Optimalitätsbedingungen
- Optimalitätsbedingungen für konvexe Programme
- Starke Konvexität

Starke Konvexität

Minimiere $f(x)$ über $x \in U$

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

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

- Wir nehmen an, es gibt eine eindeutige Lösung $x_* \in U$
- Also ist x_* durch $\nabla f(x_*) = 0$ gekennzeichnet
- Sei $x \in U$ ein Punkt mit $\|\nabla f(x)\|_2 < \varepsilon$

Wie kann man $\|x - x_*\|_2$ und $f(x) - f(x_*)$ durch ε abschätzen?

Starke Konvexität

Notation: $\forall X, Y \in \mathbb{S}^n$ *setzen wir* $X \succcurlyeq Y$ *falls* $(X - Y) \succcurlyeq 0$

- $f \in C^2(U)$ heißt stark konvex falls $\exists m > 0$, sodass:
$$\nabla^2 f(x) \succcurlyeq mI \quad \forall x \in U$$

- Gilt $\nabla^2 f(x_*) \succ 0$, so ist f stark konvex neben x_*

- Ist f stark konvex, so kann man die Optimalitätskriterien quantifizieren

Lemma 2.22. Starke Konvexität

Sei $U \subseteq \mathbb{R}^n$ offen und konvex, $f \in C^2(U)$. Wir nehmen an:

- f ist stark konvex mit Konstante $m > 0$
- $x_* \in U$ ist ein Minimum von f auf U und $f_* = f(x_*)$

Dann gelten die folgenden Abschätzungen:

1. $f(y) \geq f(x) + \nabla f(x)(y - x) + \frac{m}{2}\|y - x\|_2^2 \quad \forall x, y \in U$
2. $f(x) - f_* \leq \frac{1}{2m}\|\nabla f(x)\|_2^2 \quad \forall x \in U$
3. $\|x - x_*\|_2 \leq \frac{2}{m}\|\nabla f(x)\|_2 \quad \forall x \in U$

Beweis: zu 1

Behauptung: $f(y) \geq f(x) + \nabla f(x)(y-x) + \frac{m}{2}\|y-x\|_2^2 \;\; \forall x, y \in U$

Sei $f \in C^2(U)$, $\nabla^2 f \succcurlyeq mI$

Seien $x, y \in U$

$\exists z \in [x, y]$

Taylor-Formel

$$f(y) = f(x) + \nabla f(x)(y-x) + \underbrace{\tfrac{1}{2}(y-x)^T \nabla^2 f(z)(y-x)}_{\geq \frac{m}{2}\|y-x\|_2^2}$$

Beweis: zu 2

Behauptung: $f(x) - f_* \leq \frac{1}{2m} \|\nabla f(x)\|_2^2 \quad \forall x \in U$

$$f(y) \geq \underbrace{f(x) + \nabla f(x)(y - x) + \frac{m}{2}\|y - x\|_2^2}_{g(y)} \; \forall x, y \in U$$

minimiere $g(y)$

$$\nabla f(x) + m(y_* - x) = 0$$

$$y_* = x - \frac{1}{m}\nabla f(x)$$

$$f(y) \geq g(y_*) = f(x) - \frac{1}{2m}\|\nabla f(x)\|_2^2$$

$y := x_*$

$$f_* \geq f(x) - \frac{1}{2m}\|\nabla f(x)\|_2^2$$

Beweis: zu 3

Behauptung: $\|x - x_*\|_2 \leq \frac{2}{m}\|\nabla f(x)\|_2 \quad \forall x \in U$

$$f(x_*) \geq \overbrace{f(x)}^{\geq f(x_*)} + \underbrace{\nabla f(x)(x_* - x)}_{\geq -\|\nabla f(x)\|_2\|x - x_*\|_2 \text{ (Cauchy-Schwarz)}} + \frac{m}{2}\|x - x_*\|_2^2 \quad \forall x \in U \backslash \{x_*\}$$

$$0 \geq -\|\nabla f(x)\|_2\|x - x_*\|_2 + \frac{m}{2}\|x - x_*\|_2^2$$

dividiere durch $\|x - x_*\|_2$

$$\|x - x_*\|_2 \leq \frac{2}{m}\|\nabla f(x)\|_2$$

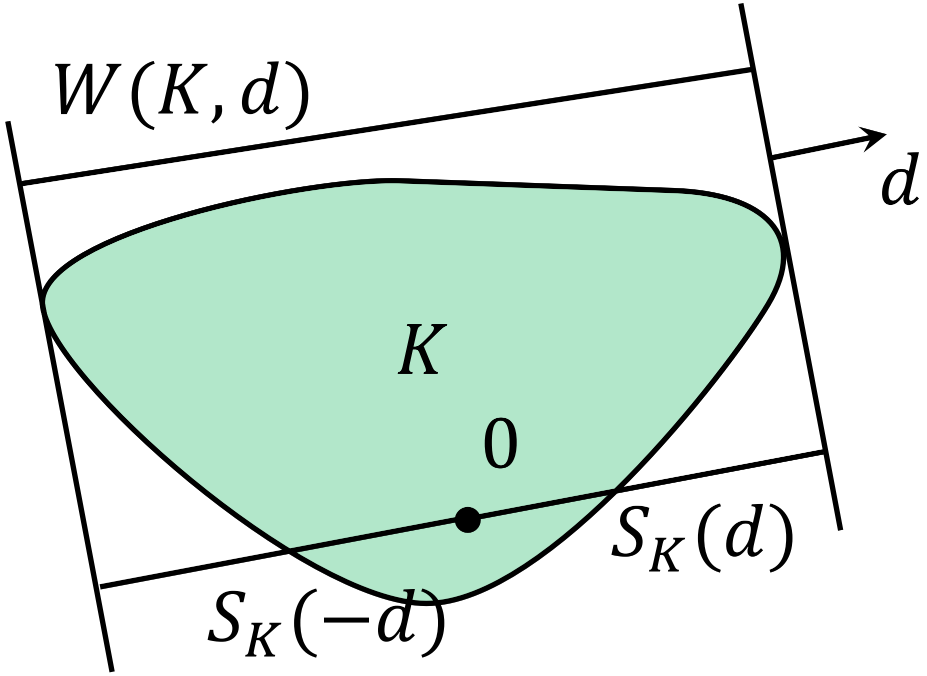
Geometrische Interpretation

- Sei $K \subseteq \mathbb{R}^n$ konvex und $d \in \overbrace{S_1^n(0)}^{\{x \in \mathbb{R}^n : \|x\|_2 = 1\}}$

- Die Breite von K in der Richtung d

$$W(K,d) = \underbrace{S_K(d) + S_K(-d)}_{\text{Stützfunktion}}$$

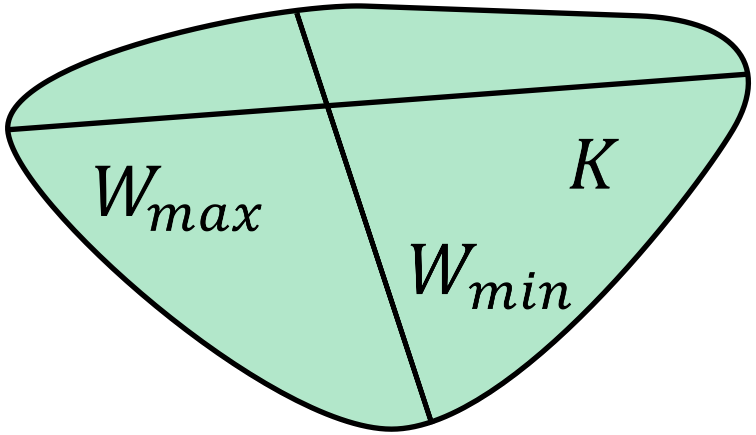
$$S_K(d) = \sup_{x \in K} d^T x$$



Maximale und minimale Breite

$$W_{max}(K) = \sup_{\|d\|_2=1} W(K,d)$$

$$W_{min}(K) = \inf_{\|d\|_2=1} W(K,d)$$



- Als die Kondition von K bezeichnet man

$$\kappa(K) := \frac{W_{max}^2(K)}{W_{min}^2(K)}$$

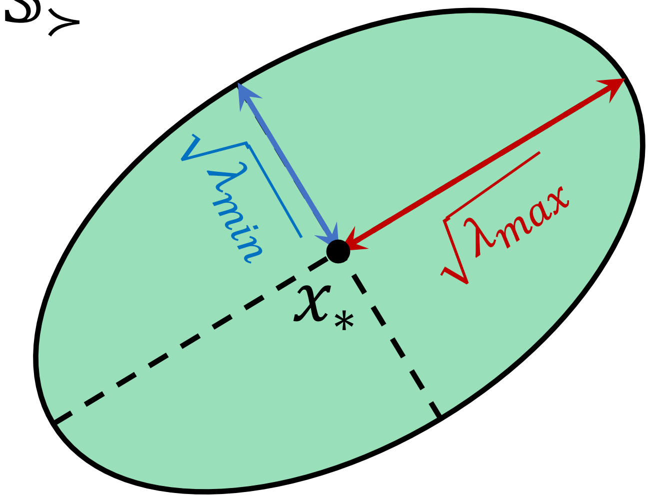
- Durch $\kappa(K)$ wird die Asymmetrie von K gemessen

Beispiel: Ellipsoid

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

Aufgabe 2.15

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



$$W(\mathcal{E}, d) = S_{\mathcal{E}}(d) + S_{\mathcal{E}}(-d) = 2\left\|\Sigma^{1/2} d\right\|_2$$

$W_{max}(\mathcal{E}) = 2\sqrt{\lambda_{max}}$ — max. Eigenwert von Σ

$W_{min}(\mathcal{E}) = 2\sqrt{\lambda_{min}}$ — min. Eigenwert von Σ

$$\kappa(\mathcal{E}) = \frac{W_{max}^2(\mathcal{E})}{W_{min}^2(\mathcal{E})} = \frac{\lambda_{max}}{\lambda_{min}} = \kappa(\Sigma)$$ Kondition von Σ

Satz 2.23. Kondition der Unterniveau-Mengen

Sei $f \in C^2(\mathbb{R}^n)$ und $S = \{x: f(x) \leq \alpha\}$ mit $\alpha \in \mathbb{R}$. Wir nehmen an:

- $\nabla^2 f \succcurlyeq mI$ auf S (bzw. $\nabla^2 f \preccurlyeq MI$ auf S)
- x_* ist ein globales Minimum von f und $f_* = f(x_*)$

Dann gilt:

$$W_{max}^2(S) \leq \frac{2}{m}(\alpha - f_*) \qquad (\text{ bzw. } W_{min}^2(S) \geq \frac{2}{M}(\alpha - f_*)\ )$$

Gelte gleichzeitig $mI \preccurlyeq \nabla^2 f \preccurlyeq MI$ auf S, so gilt:

$$\kappa(S) \leq \frac{M}{m}$$

Beweis

Behauptung: $\nabla^2 f \succcurlyeq mI$ auf $S \implies W_{max}^2(S) \le \frac{2}{m}(\alpha - f_*)$

$$f_* + \underbrace{\nabla f(x_*)^T}_{0}(x - x_*) + \frac{m}{2}\|x - x_*\|_2^2 \le f(x) \quad \forall x \in S \quad \textit{(Lemma 2.22)}$$

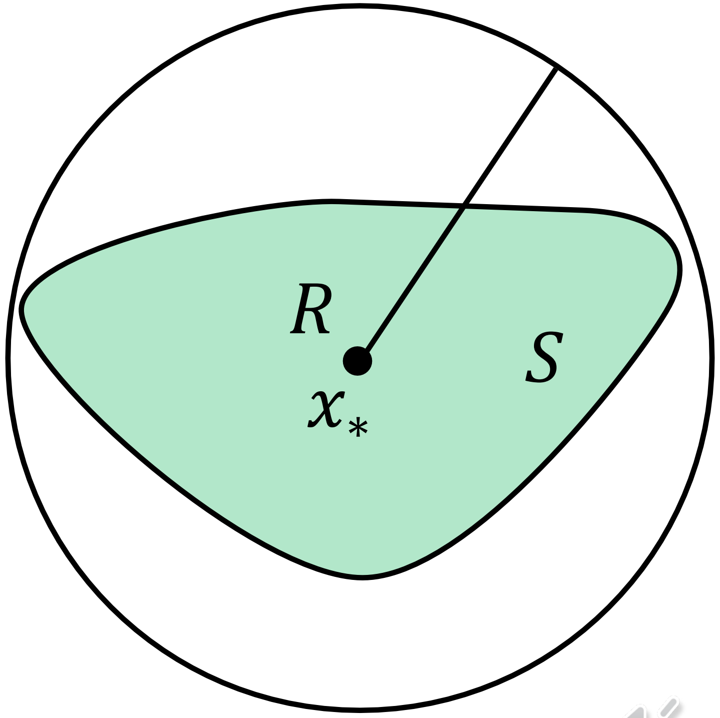
$$\underbrace{f(x) \le \alpha}_{x \in S}$$

$$f_* + \frac{m}{2}\|x - x_*\|_2^2 \le \alpha$$

$$\|x - x_*\|_2^2 \le \frac{2}{m}(\alpha - f_*)$$

$$x \in \overline{B_R}(x_*),\ R = \sqrt{\frac{2}{m}(\alpha - f_*)}$$

$$S \subseteq \overline{B_R}(x_*) \longrightarrow W_{max}(S) \le R$$



Beweis

Behauptung: $mI \preccurlyeq \nabla^2 f \preccurlyeq MI$ auf $S \Longrightarrow \kappa(S) \leq \frac{M}{m}$

$$W_{min}^2(S) \geq \frac{2}{M}(\alpha - f_*)$$

$$W_{max}^2(S) \leq \frac{2}{m}(\alpha - f_*)$$

$$\kappa(S) \coloneqq \frac{W_{max}^2(S)}{W_{min}^2(S)} \leq \frac{M}{m}$$

$\square$

Zusammenfassung

- Allgemeine Optimalitätsbedingungen
- Optimalitätsbedingungen für konvexe Programme
- Starke Konvexität

Nächstes Video

- 3a. Abstiegsverfahren: Definition und Schrittweitenwahl