

Ausgleichsprobleme:

①

1. Fall:

Gegeben: - m Datenpunkte (x_i, y_i) mit $x_i, y_i \in \mathbb{R}$, $1 \leq i \leq m$
- n Funktionen $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$

Gesucht: Lineare Kombination dieser Funktionen, so dass die Abstände dieser Gesamtfunktion zu den Datenpunkten minimal wird:

$$\min \sum_{k=1}^m \left| \underset{\substack{\uparrow \\ \text{bekannt}}}{y_k} - \sum_{i=1}^n \underset{\substack{\uparrow \\ \text{bekannt}}}{a_i} f_i(x_k) \right|^2$$

Die Koeffizienten a_i sind gesucht.

Diesen Ausdruck kann man umformen zu

$$\min \sum_{k=1}^m \left| y_k - r_k^T a \right|^2, \text{ wobei}$$

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$$

$$r_k = \begin{bmatrix} f_1(x_k) \\ \vdots \\ f_n(x_k) \end{bmatrix} \in \mathbb{R}^n \text{ gilt.}$$

Insgesamt kann man damit noch weiter umformen zu

$$\min \left\| \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} - \begin{bmatrix} f_1(x_1) & \dots & f_n(x_1) \\ \vdots & & \vdots \\ f_1(x_m) & \dots & f_n(x_m) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \right\|_2^2$$
$$= \begin{bmatrix} y_1 - r_1^T a \\ \vdots \\ y_m - r_m^T a \end{bmatrix} \in \mathbb{R}^m$$

Wir erhalten damit ein Problem der Form

$$\min \| b - Ma \|_2^2 \quad \text{mit} \quad b \in \mathbb{R}^m, M \in \mathbb{R}^{m \times n}, a \in \mathbb{R}^n$$

2. Fall (mehrdimensional)

②

Gegeben: - Jeweils m Datenpunkte $x_i \in \mathbb{R}^q$, $y_i \in \mathbb{R}^p$

- n bekannte Funktionen $f_i: \mathbb{R}^q \rightarrow \mathbb{R}^p$

Gesucht: Wieder eine Linearkombination dieser Funktionen wie in Fall 1:

$$\min \sum_{k=1}^m \left\| \underset{\substack{\uparrow \\ \text{bekannt}}}{y_k} - \sum_{i=1}^n a_i \underset{\substack{\uparrow \\ \text{bekannt}}}{f_i(x_k)} \right\|_2^2 \quad \text{Wieder sind die Koeffizienten } a_i \text{ gesucht.}$$

Diesen Ausdruck kann man umschreiben zu

$$\min \sum_{k=1}^m \| y_k - R_k a \|_2^2 \quad \text{mit} \quad a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$$

$$R_k = \left[f_1(x_k) \mid \dots \mid \underset{\substack{\uparrow \\ \text{Spaltenvektor}}}{f_n(x_k)} \right] \in \mathbb{R}^{n \times p}$$

Damit kann man weiter umformen zu

$$\min \left\| \underbrace{\begin{bmatrix} \text{Vektor in } \mathbb{R}^p \\ y_1 \\ \vdots \\ y_m \end{bmatrix}}_{\substack{\uparrow \\ \text{bekannt}}} - \underbrace{\begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix}}_{\substack{\uparrow \\ \text{bekannt}}} a \right\|_2^2$$

$$= \left\| \begin{bmatrix} y_1 - R_1 a \\ \vdots \\ y_m - R_m a \end{bmatrix} \right\|_2^2 \in \mathbb{R}^{pm}$$



Wieder erhalten wir ein Problem in der Form

$$\min \| b - M a \|_2^2, \quad \text{hier ist } b \in \mathbb{R}^{p \cdot m}$$

$$M \in \mathbb{R}^{p \cdot m \times n}$$

$$a \in \mathbb{R}^n$$

Wie soll man nun ein Problem der Form

(3)

$$\min \|b - Ma\|_2^2 \quad \text{lösen?}$$

$$(b - Ma, b - Ma)_2$$

quadratische Funktion,
nach oben geöffnet; also Minimum
im Scheitel, was man
durch Ableiten findet

Es bezeichne $F: \mathbb{R}^n \rightarrow \mathbb{R}$ die Abbildung

$$F(a) = (b - Ma, b - Ma)_2$$

Produktregel

$$\begin{aligned} DF(a)h &= 2(b - Ma, -Mh)_2 \\ &= -2(M^T b - M^T M a, h)_2 \end{aligned}$$

$$\Rightarrow \nabla F(a) = 2(M^T M a - M^T b) \stackrel{!}{=} 0$$

$$\Rightarrow M^T M a = M^T b$$

($M^T M$ ist regulär $\Leftrightarrow$
 M hat volle Spaltenrang)

$$\text{wenn } M^T M \text{ regulär} \Rightarrow a = (M^T M)^{-1} M^T b$$

Damit haben wir die gewünschte Linearkombination gefunden,
die Koeffizienten stehen im Vektor a .

Was kann man jetzt damit machen?

Wenn man z.B. ein Polynom $p: \mathbb{R} \rightarrow \mathbb{R}$ vom Grad ≤ 5 sucht, welches
minimalen Abstand zu den Datenpunkten hat, kann man

$$f_1 \equiv 1, f_2(x) = x, f_3(x) = x^2, \dots, f_{s+1}(x) = x^s \quad \text{als}$$

Funktionen nehmen, da man als Ergebnis eine Linearkombination
dieser Funktionen und damit eben genau so ein Polynom bekommt.

Bei mehrdim. Datenpunkten muss man nur beachten, dass

$$f_i: \mathbb{R}^q \rightarrow \mathbb{R}^p \text{ geht und dementsprechend passende}$$

$$\text{Funktionen wählt. z.B. } f_1: \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{bmatrix} x^2 \\ y^2 \\ 0 \end{bmatrix} \quad (\text{rein fiktives Beispiel!})$$

😊 Das hat Spaß gemacht

Aufgabe 5

$$f(x) = x_1^2 + x_2^2$$

$$g_1(x) = -x_2$$

$$g_2(x) = x_1^3 - x_2$$

$$g_3(x) = x_1^3(x_2 - x_1^3)$$

$$\nabla g_1(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\nabla g_2(x) = \begin{pmatrix} 3x_1^2 \\ -1 \end{pmatrix}$$

$$\nabla g_3(x) = \begin{pmatrix} 3x_1^2x_2 - 6x_1^5 \\ x_1^3 \end{pmatrix}$$

gibt es bei uns gar keine

Es gilt $T_c(\bar{x}, g, h) = \{ d \in \mathbb{R}^2 \mid \nabla g_i(\bar{x})^T d \leq 0 \text{ f\"ur } i \in A_{\bar{x}}, \nabla h_j(\bar{x})^T d = 0 \forall j \}$

und wir sollen den linearisierten Tangentialkegel in $\bar{x} = (0, 0)$ bestimmen.

$$\nabla g_1(0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \nabla g_2(0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \nabla g_3(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \leftarrow \text{ist im Skalarprodukt mit allen Vektoren}$$

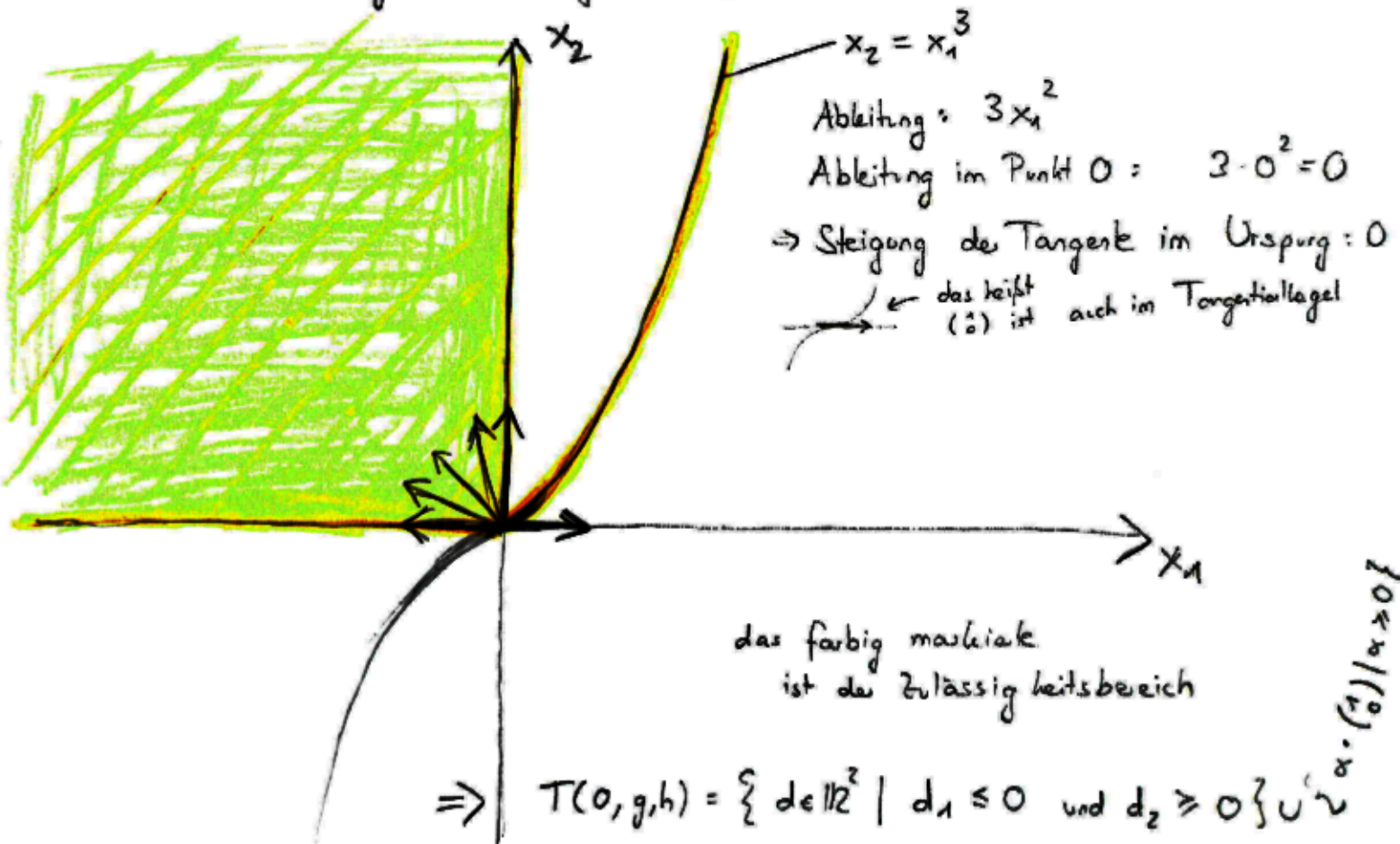
sowie 0.

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \leq 0$$

$$\Leftrightarrow -d_2 \leq 0 \Leftrightarrow d_2 \geq 0$$

$$\Rightarrow T_c(0, g, h) = \{ d \in \mathbb{R}^2 \mid d_2 \geq 0 \}$$

F\"ur die Bestimmung des Tangentialkegels schauen wir uns den Zul.bereich an:



duale Kegel:
 $0 \in \mathbb{R}^n$

Wenn K ein Kegel ist, ist der duale Kegel so definiert:

$$\text{dual}(K) = \{ d \in \mathbb{R}^n \mid (d, x)_2 \geq 0 \quad \forall x \in K \}$$

der duale Kegel zum lin. Tangentialkegel ist also bei uns

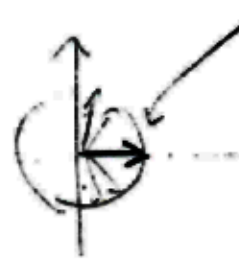
$$\{ d \in \mathbb{R}^2 \mid d_1 = 0 \wedge d_2 \geq 0 \}$$

der duale Kegel zum Tangentialkegel ist bei uns

$$\{ d \in \mathbb{R}^2 \mid d_1 = 0 \wedge d_2 \geq 0 \},$$

die dualen Kegel sind hier also identisch.

(Das kann man sich überlegen



mit diesem Vektor im Skalarprodukt ergeben die Vektoren aus dem schraffierten Bereich einen Wert ≥ 0 und das kann man sich für obige Kegel überlegen)

Aufgabe 6

Beispiel 1:

$$g_1(x) = -x_1^3 - x_2$$

$$g_2(x) = -x_2$$

$$\Rightarrow \nabla g_1(x) = \begin{pmatrix} -3x_1^2 \\ -1 \end{pmatrix}$$

$$\nabla g_2(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

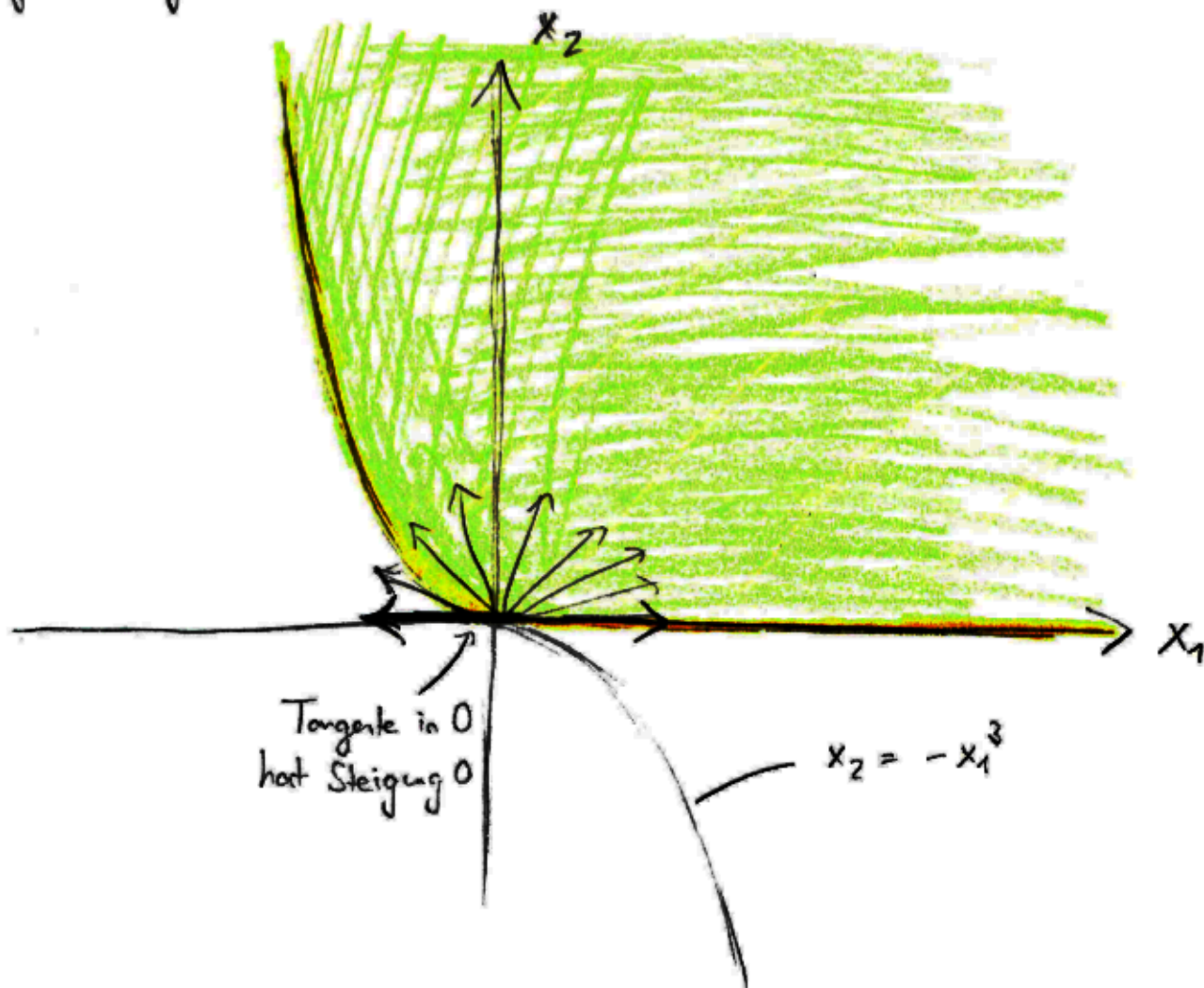
im Punkt $(0,0)$:

$$\nabla g_1(0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \nabla g_2(0) \leftarrow \text{nicht lin. unabh.} \Rightarrow \text{LICQ verletzt.}$$

Linearisierter Tangentialkegel in $(0,0)$:

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \stackrel{!}{\leq} 0 \Rightarrow T_e(0, g, h) = \{ d \in \mathbb{R}^2 \mid d_2 \geq 0 \}$$

Tangentialkegel? :



$$T(0, g, h) = \{ d \in \mathbb{R}^2 \mid d_2 \geq 0 \} = T_e(0, g, h) \checkmark$$

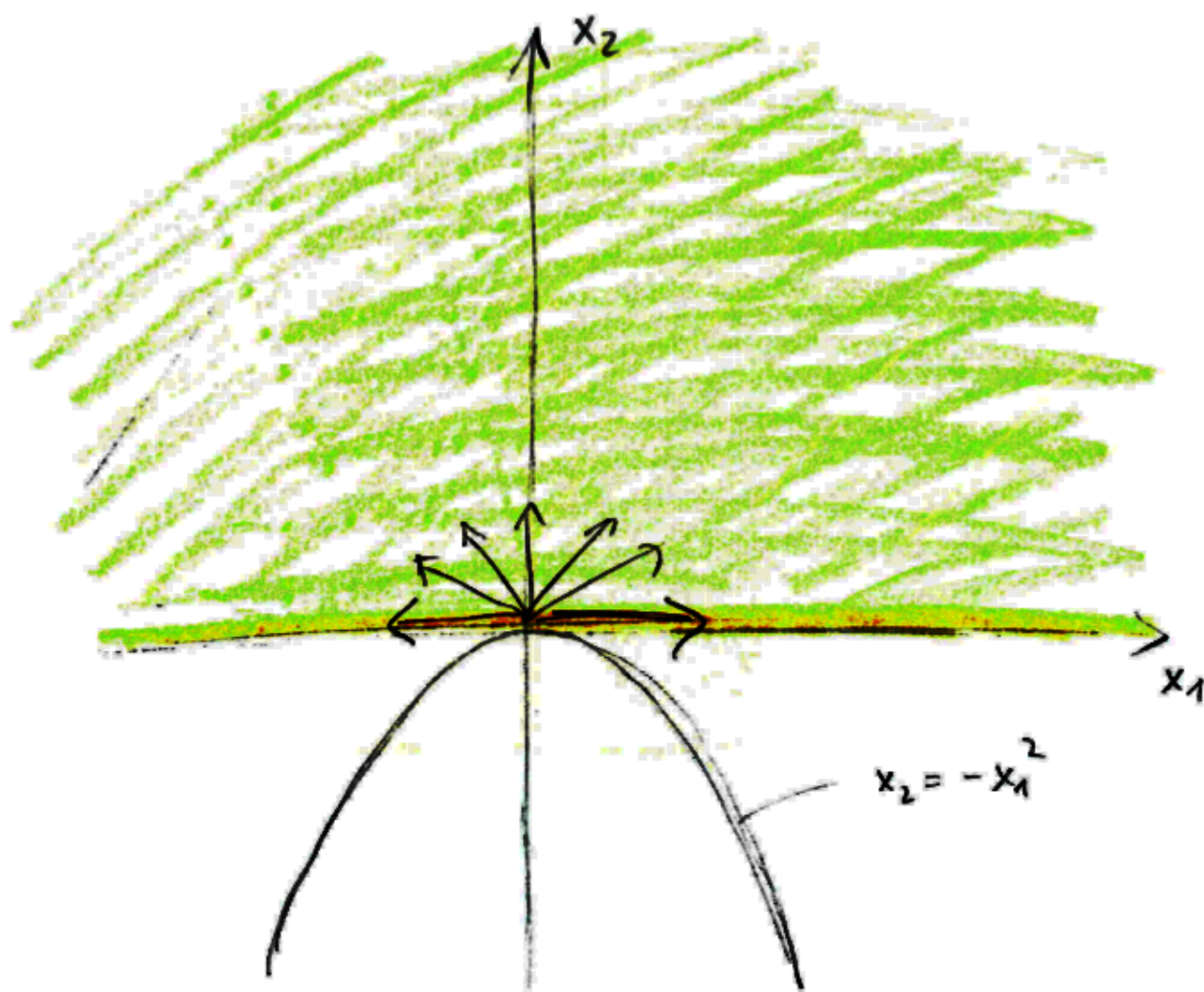
Beispiel 2: $g_1(x) = -x_2 - x_1^2$ $\nabla g_1(x) = \begin{pmatrix} -2x_1 \\ -1 \end{pmatrix}$
 $g_2(x) = -x_2$ $\nabla g_2(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

im Punkt $(0,0)$:

$$\nabla g_1(0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \nabla g_2(0) \Rightarrow \text{LICQ} \text{ verletzt}$$

Linearisierter Tangentialkegel wie im vorigen Bsp:

$$T_c(0, g, h) = \{ d \in \mathbb{R}^2 \mid d_2 \geq 0 \}$$



$$T(0, g, h) = \{ d \in \mathbb{R}^2 \mid d_2 \geq 0 \} = T_c(0, g, h) \checkmark$$

Aufgabenübersicht:

- ① $X_{\text{glob}}^* \subseteq X_{\text{lok}}^* \subseteq X_{\text{int}}^*$ für $f: \mathbb{R} \rightarrow \mathbb{R}$ stetig
- ② Rechenaufgabe, Bsp für ∞ -viele kritische Punkte
- ③ Niveaumenge ist für stetige Fkt. abgeschlossen
↳ wenn geschnitten mit D beschränkt nimmt f auf D Min. an
- ④ C^1 -Fkt. sind auf konvexen, kompakten Gebieten Lipschitz
- ① f konvex, D konvex $\Rightarrow$ Schnitt von Niveaumenge mit D ist konvex
 C^0 $\hookrightarrow D$ zw. abgeschl., f gleichmäßig konvex $\rightarrow$ genau 1 Minimum auf D
- ② Abschätzung von Taylor-Restglied
- ③ Kandidatenpunkte liegen dicht in X_{glob}^*
- ④ f konvex $\rightarrow$ unimodal $\Rightarrow$ quasikonvex für f stetig, $f: S \rightarrow \mathbb{R}$, S konvex
- ① + ② Rechenaufgaben, Beispiele
- ③ Bei Q-konvergenz der Urbilder konvergieren die Bilder in entgegengesetzter R-Geschw. für stetige Funktionen
- ① - strikt quasikonvexe Funktionen laufen ^{strik} monoton oder V-Konfig.
- durch Vergleich der beiden inneren Punkte, kann man aussagen in welchem Intervall x^* liegt
- ② Abschätzung, Rechenaufgabe, nicht so viele Konvergenz wegen Wurzel
- ③ $C\|x\| \leq \|Ax\| \leq C\|x\|$, C, c Max/Min Beträge von EW
- ① GSS konvergiert gegen $x^* \in X_{\text{glob}}^*$ für f konvex und $X_{\text{glob}}^* \cap [a, b] \neq \emptyset$
- ② $f \in C^1$, konvex oder strikt q.k.
- $X_{\text{glob}}^* \neq \emptyset$ oder
$$\lim_{x \rightarrow -\infty} f(x) = \inf_{x \in \mathbb{R}} f(x) \quad \text{für } f'(x^*) > 0$$

$$\lim_{x \rightarrow +\infty} f(x) = \sup_{x \in \mathbb{R}} f(x) \quad \text{für } f'(x^*) < 0$$

EW monoton fallend von oben
- Als. der entweder V-Konfig. liefert oder Folgo mit $\otimes$
- ① Bsp für Optimierungsproblem, Formulierungsaufgabe, mit ∞ -viele Lsg
- ② $g \in C^2$, in x^* einf. Nst., $g''(x^*) \neq 0 \Rightarrow$ KO von Sekantenverfahren ist $\frac{1+\sqrt{5}}{2}$
- ① $f \in C^2$ $\uparrow$ 1. lok. Min ist $\nabla^2 f(x^*)$ pos. semi def.
- ② Daffert $\nabla f(x^*) = 0$, $\nabla^2 f(x^*)$ spd $\Rightarrow x^*$ stärkstes lok. Min.stelle
- ③ Übung zur Richtungsableitung (manche Fkt. haben überall Richtungsableitung sind aber nicht diff. bar)
- ④ $f: D \rightarrow \mathbb{R}$ konvex, D konvex. Alle Richtungsableit. existieren
- ⑤ Wenn Gradient ex. gilt $f'(x, d) = \nabla f(x)^T d$

- ① Rechenaufgabe, Bsp für Funktion, die kein ⁶⁶Min. in einem Punkt hat, auf verschiedene Geraden eingeschränkt aber schon
- ② Für Suchrichtungen, die fast senkrecht zur Gradientenrichtung sind, konv. Abstiegsverfahren evtl. nicht gegen Opt. Punkt, in dem Bsp der Aufgabe z.B.



bleibt immer Abstand

- ③ Bsp für wohldef., effiziente Schrittweitenregel
- für $\nabla f(x)^T d < 0 \Rightarrow$ Abstiegsrichtung
- ④ - Schrittweite berechnen für Quadr. Problem
- Q-lineare Konvergenz des Iterierten in A-Norm
 - Kond.-zahl ≈ 1
 - Vergleich mit $\|x_0 - x^*\|$
- ⑤ Matrix H bleibt inv., wenn man bel. Matrix aufaddiert mit $\|A^{-1}B\| < 1$ (Menge der inv. Matrizen ist offen)
- $H \mapsto H^{-1}$ ist stetige Abbildung

- ③ Rechenaufgabe für konkretes Bsp, Q-lineare Konvergenz

- ①+② Hilfsaussagen für ③

- ③ Satz von Dennis und Moré:
- $f \in C^2$, (H_k) Folge reg. Matrizen, $x^{k+1} = x^k - H_k^{-1} \nabla f(x^k)$ (Quasi-Newton), $x^k \rightarrow x^*$
- $\nabla^2 f(x^*)$ regulär.

$$x^k \rightarrow x^* \text{ superlinear, } \nabla f(x^*) = 0 \quad \Leftrightarrow$$

$$\|(\nabla^2 f(x^k) - H_k)(x^{k+1} - x^k)\| = o(\|x^{k+1} - x^k\|) \quad \Leftrightarrow$$

$$\|(\nabla^2 f(x^k) - H_k)(x^{k+1} - x^k)\| = o(\|x^{k+1} - x^k\|)$$

H_k müssen gegen $\nabla^2 f(x^*)$ z.B. konvergieren, bzw. $H_k - \nabla^2 f(x^k)$ muss, dann hat man superlin. Konvergenz

- ④ Bsp

- ① - Darstellung lin. Tangentialkegel über constraints
- Tangentialkegel ist im lin. Tangentialkegel enthalten

- ② Bsp zu dualen Problem, Dualitätssatz

- ③ $\nabla f(x)^T d < 0 \Leftrightarrow$ Es gibt z.B. Abstiegsrichtung ($\Leftarrow$ stimmt nur mit $\Leftarrow$)

- ⑤+⑥ Rechenaufgabe als Beispiele

- ① LICQ $\Rightarrow$ ACQ, SCQ $\Rightarrow$ MFCQ $\Rightarrow$ ACQ

- ② Bsp für Strafvorfahren

- ③ Rechenaufgabe für red. Gradientenmethode

- ④ Strafvorfahren konv. gegen opt. Lösung, wenn f, g, h stetig, $c_k \nearrow \infty$, inneres Lsg existiert ∇

- ⑤ Rechenaufgabe für Barriereverfahren

Algorithmen:

Diskretisierungsmethode ++++++ - viele Funktionsauswertungen (Verbesserungen möglich, dann ^{2. B.} Lipschitz-stetigkeit)
- alle Funktionen

Golden section search - Unterteilung des Intervalls im Goldenen Schnitt
- konv. für konvexes f (Üb. Aufgabe)

Bisektionsverfahren - halbiert Intervalle um NST. zu finden
- für stetige Fkt. ^{1. B.} f , mehrfache NST., die andere Verfahren Probleme machen
- konv. linear

Sehantenverfahren - sucht NST., legt Sehante durch 2 Heile, NST. dieser ist neue Heile
- Modifikation von Newton, Ableitung durch Diff.quotient ersetzt ^{zu Auslöschungseffekt}
- Superlineare Konvergenz, f muss stetig sein, genau 1 NST.

Regula Falsi - Kombination Bisektion/Sehante, solange f nicht stark konvex/konkav $\rightarrow$ Superlinear
sonst bleibt 1 Intervall mit NST. $\rightarrow$ sonst linear

Newton $f \in C^1, a$ init. NST. $\rightarrow$
- lokal quadratisch, man muss Startwert erreichen
- legt Tangente an $\rightarrow$ mehr mit NST. weiter, sucht NST.
- bei mehrfachen NST. Probleme mit Konvergenz

Steepest Abstieg - man läuft in Richtung des neg. Gradienten, konvergiert oft langsam
- konjugierte Gradienten ist Verbesserung
- f muss diffbar sein
- Schwachkonvergenz nicht belegt werden, s. Übungsaufgabe zu Sparrnfall

BFGS, DFP 2. Newtonverfahren, Approximation $\approx$ 2. Ableitung, Update Formeln Hessematrix

Zuwendig?

Strafverfahren - Ersetze restl. Problem durch ^{Folge} unrestl. Probleme, deren Lösungen gegen Opt. Lsg. konvergieren. Unrestl. Probleme lösen "Straflem" (0 bis zul. region, sonst $\neq 0$)
- bestraft Verletzung der Nebenbed.

Barriereverfahren (innere Punkte) - Einfall Straffunktion kommt dazu
- Heile werden getrennt zul. zu sein
- Barriere sorgt dafür das Verlassen des zul. Bereichs zu verhindern
einf. Problem kann dadurch schnell schwierig werden

SQP
- wende auf KKT-Bod Newton an
- Newtonl. Δx löst quadt. Min. Aufgabe
- quadratische Teilprobleme (bei unrestl. $\rightarrow$ Newton)
für Suchrichtung

findet „neu“ KKT-Punkte