Zusammenfassung zu Optimierung I

Sara Adams

9. Juli 2004

Diese Zusammenfassung basiert auf der Vorlesung
Optimierung I
gehalten im Sommersemester 2004
von Prof. Martin Buhmann (Ph.D.)
an der Justus-Liebig Universität Gießen

Inhaltsverzeichnis

Teil I
Lineare Optimierung

$$f(x) = c^T x \to \max \text{ oder } \min, \quad x \in \mathbb{R}^n, c \in \mathbb{R}^n \text{ Koeffizientenvektor mit linearen NB}$$

1 Grundlegendes

1.1 Normalform

$$c^T x \to \min \qquad x, c \in \mathbb{R}^n, \quad x \geq 0 \text{ (komponentenweise)}$$
$$\text{Nebenbedingungen} \qquad Ax = b \qquad A = (a^{(1)} \dots a^{(n)}) \in \mathbb{R}^{m \times n}$$
$$M = \{x \in \mathbb{R}^n : \, x \geq 0, Ax = b\} \text{ zulässiger Bereich}$$

- $d^T \to \max \qquad \Rightarrow \; c = -d$
- $y_i < 0 \quad \Rightarrow \; y_i = x_1 - x_2, x_1, x_2 \geq 0$
- NB mit "$\geq$" durch Multiplikation mit -1 auf "$\leq$"
- NB $a^T x \leq b$ durch Schlupfvariable auf "=" transformieren: $y \geq 0, \; a^T x + y = b$

1.2 Ecken

$$x \in M \text{ Ecke } \Leftrightarrow \nexists x^{(1)} \neq x^{(2)} \in M, \; \lambda \in (0,1) : \; x = \lambda x^{(1)} + (1-\lambda) x^{(2)}$$

1.3 Sätze

- $x \in M$ Ecke $\Leftrightarrow$ die Spaltenvektoren in der Menge $B(x) = \{a^{(i)} : \; i \in I(x)\}, \; I(x) = \{i : x_i > 0\}$ sind linear abhängig
- Hat das lineare Optimierungsproblen eine Lösung, so gibt es eine Ecke $x^* \in M$, die Lösung ist.

1.4 Lexikographische Ordnung

$u \in \mathbb{R}^n$ heißt lexikographisch positiv ($u \vec{>} 0$), wenn $u \neq 0$ und die erste nicht verschwindende Komponente positiv ist.
u heißt lexikographisch größer als v ($u \vec{>} v$ für $u, v \in \mathbb{R}^n$), falls $u - v \vec{>} 0$.

2 Die Simplexmethode

Es wird zwischen Basisvariablen (BV) und Nichtbasisvariablen (NBV) unterschieden.

Phase I: Auffinden einer Startecke

Ist das Optimierungsproblem gegeben als

$$c^T x \to \min, \quad x \geq 0, \quad Ax \leq b, \quad b \geq 0,$$

so ist eine Startecke sofort gegeben:

$$\tilde{c}^T \tilde{x} \to \min$$
$$\tilde{x} = (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m})^T \in \mathbb{R}^{n+m}, \quad \tilde{c} = (c_1, \dots, c_n, 0, \dots, 0)^T \in \mathbb{R}^{n+m}$$
$$\tilde{A}\tilde{x} = b, \qquad \tilde{A} = (AI_m)$$

äquivalent zum ursprünglichen Problem mit Startecke $x^0 = (0, \dots, 0, b_1, \dots, b_m)^T \in \mathbb{R}^n$. Hier sind $x_1, \dots, x_n$ NBV, $x_{n+1}, \dots, x_{n+m}$ BV.

Ist $b < 0$ oder sind einige NB als "=" gegeben, so ist die Startecke nicht offensichtlich:
Hilfsproblem:

$$\hat{c}^T \hat{x} \to \min, \quad \hat{x} \geq 0, \quad \hat{c} = (0, \dots, 0, 1, \dots, 1)^T \in \mathbb{R}^{n+m}, \; \hat{x} = (x, v)^T$$
$$\hat{A}\hat{x} = b, \qquad \hat{A} = (AI_m)$$

Löse dieses Problem mit dem Simplexverfahren, Lösung $(x^*, v^*)^T \in \mathbb{R}^{n+m}$. Ist $v^* = 0$, so ist x^* eine Startecke des ursprünglichen Problems. Ist $v^* \neq 0$, so ist das ursprüngliche System nicht lösbar.

Entartete Ecken:
Falls $p' \neq p$ existiert mit

$$\min_{i:\, \alpha_{iq} > 0} \frac{x_i^0}{\alpha pq} = \frac{x_i^0}{\alpha_{pq}} = \frac{x_{p'}^0}{\alpha_{p'q}},$$

so wählt man p derart, dass folgender Vektor lexikographisch minimal wird:

$$u^p = \left(\frac{x_p^0}{\alpha_{pq}}, \frac{\alpha_{p1}}{\alpha_{pq}}, \dots, \frac{\alpha_{pm}}{\alpha_{pq}} \right)^T \in \mathbb{R}^{m+1}$$

Phase II: Minimierung

Simplexschritt:

1. Bestimmen des Pivotelements durch Wahl des Koeffizientens der Zielfunktion und der Pivotzeile.
2. Eintragen der neuen BV.
3. Bilden der Pivotzeile durch Division durch das Pivotelement
4. Bilden der neuen Pivotspalte
5. Bilden aller übrigen Elemente

Simplextableau:

BV	$x_k\ (k \notin I_0)$	
$x_i\ (i \in I_0)$	$(\alpha_{ik})_{i\in I_0, k\notin I_0}$	$x_i^0\ (i \in I_0)$
	$\gamma_k\ (k \notin I_0)$	Z

$x_i, i \in I_0$ sind die BV, $x_k, k \in \{1, ..., n\}\backslash I_0$ die NBV.

Simplexalgorithmus:

1. $\gamma_k \geq 0\ \forall k \notin I_0?\quad \Rightarrow$ fertig, Lösung $x_l = \begin{cases} x_l^0 & \text{falls } l \in I_0 \\ 0 & \text{falls } l \notin I_0 \end{cases}$
2. $\exists\, p:\ \alpha_{pq} > 0$ Wähle p, so dass $\frac{x_p^0}{\alpha_{pq}}$ minimal ist unter allen $\frac{x_i^0}{\alpha_{iq}}$ mit $\alpha_{iq} > 0$

$$p = \arg\min_{i:\,\alpha_{iq}>0} \frac{x_i^0}{\alpha_{iq}}$$

Transformation der $\alpha_{ik} \Rightarrow\ \alpha'_{ik}$ im neuen Simplextableau:

- $\alpha_{pq} \Rightarrow\ \alpha'_{qp} = -\frac{\alpha_{iq}}{\alpha_{pq}}$
- in der Pivotspalte:
  $\alpha_{iq} \Rightarrow\ \alpha'_{ip} = -\frac{\alpha_{iq}}{\alpha_{pq}}$
  x_p wird zu x_q
  γ_q wird zu $\gamma'_p = -\frac{\gamma_q}{alpha_{pq}}$
- in der Pivotzeile:
  $\alpha_{pk} \Rightarrow\ \alpha'_{qk} = \frac{\alpha_{pk}}{\alpha_{pq}}$
  $x_p^0 \Rightarrow\ x_q^1 = \frac{x_p^0}{\alpha_{pq}}$
- sonstige:
  alle $x_i, x_k,\ i$ bzw. k nicht p bzw. q bleiben unberührt
  $\alpha_{ik} \Rightarrow\ \alpha'_{ik} = \alpha_{ik} - \frac{\alpha_{iq}\alpha_{pk}}{\alpha_{pq}}$
  $x_i^0 \Rightarrow\ x_i^1 = x_i^0 - \frac{\alpha_{iq}x_p^0}{a_{pq}}$
  $\gamma_k \Rightarrow\ \gamma'_k = \gamma_k - \frac{\gamma_q\alpha_{pk}}{\alpha_{pq}}$

3 Transportprobleme

gegeben: $a_1, ..., a_m, b_1, ..., b_n \geq 0:\ \sum_{i=1}^m a_i = \sum_{i=1}^n b_i$
gesucht: $x_{ik} \geq 0:\ Z = \sum_{i=1}^m \sum_{k=1}^n c_{ik}x_{ik} = \max!$

- **Ausgangstableau**: $\begin{array}{ccc|c} c_{11} & \dots & c_{1n} & a_1 \\ \vdots & & \vdots & \vdots \\ c_{m1} & \dots & c_{mn} & a_m \\ \hline b_1 & \dots & b_n & \end{array}$
- erste Basislösung: mit **Diagonalmethode**
  Sei x_{11} BV. Wird durch den Wert der letzten BN die

  1. Zeilensumme erreicht, so erhalten alle weiteren Variablen dieser Zeile den Wert 0 (NBV) und $x_{i+1,k}$ wird nächste BV.
  2. Spaltensumme erreicht, so erhalten alle weiteren Variablen dieser Spalte den Wert 0 (NBV) und $x_{i,k+1}$ wird nächste BV.
- **Turmzug**: Ein Turmzug zur NBV x_{ik} ist eine endliche Folge von Variablen (der Optimierungsaufgabe) mit erstem Element x_{ik} und ansonsten BV. Je zwei aufeinanderfolgende Elemente liegen abwechselnd in derselben Zeile oder Spalte.
- **Transporttabellen**: Jeder Zeile und Spalte im Tableau wird eine Variable u_i bzw. v_j zugeordnet, wobei gelten soll: $\begin{array}{ccc|c} & & & u_1 \\ & & & \vdots \\ & & & u_m \\ \hline v_1 & \dots & v_n & \end{array}$
  - ohne Einschränkung $u_1 = 0$
  - auf den Feldern der BV: $u_i + v_j = c_{ij}$
  - auf den Feldern der NBV: $c_{ij} - u_i - v_j$
- Werden in der Zielfunktion die BV mittels der NB des Transportproblems eliminiert, so ist $c_{ij} - u_i - v_j$ der Koeffizient der neuen BV $x_{ij}:\quad Z \to\ Z + (c_{ij} - u_i - v_j)p$, wenn $x_{ij} = p$
- **Austauschschritt**:
  1. Erzeugen des Hilfstableaus zur Bestimmung der NBV, die ausgetauscht werden soll (Transporttableau)
  2. Fertig (bei max ist man fertig, wenn $c_{ij} - u_i - v_j \leq 0$ bei den NBV)
  3. Bestimmen des maximalen p, das auf dem Turmzug zur NBV x_{ij} gewählt werden darf
  4. Eintragen der neuen Basislösung, GOTO 1.

  Bemerkung: bei ≤ 0 haben wir eine Lösung, bei < 0 **die** Lösung.

Teil II
Nichtlineare Optimierung

Aufgabenstellung:

$$f: \mathbb{R}^n \supset D \to \mathbb{R} \in \mathcal{C}^1(D) \text{ minimieren mit/ohne NB (lokale Minima)}$$
$$g_1(x) = 0\ \dots\ g_m(x) = 0,\ g_{m+1}(x) \leq 0,\ \dots\ g_n(x) \leq 0$$

- $D \ni x^*$ lokales Minimum von $f\ :\Leftrightarrow$
  $\exists \varepsilon > 0:\ f(x) \geq f(x^*)\ \forall x \in B_\varepsilon(x^*) = \{x:\ ||x - x^*|| \leq \varepsilon\}$
- x^* lokales Minimum von $f \in \mathcal{C}^1(D) \Rightarrow\ \nabla f(x^*) = 0$ (**stationärer Punkt**)

4 Minimieren ohne Nebenbedingungen

- $f \in \mathcal{C}^1(D),\ x^*$ lokales Minimum $\Rightarrow\ \nabla f(x^*) = 0$
- $f \in \mathcal{C}^2(D),\ x^*$ lokales Minimum $\Rightarrow\ \nabla^2 f(x^*)$hat keine negativen Eigenwerte
- $\nabla f(x^*) = 0\ \wedge\ \nabla^2 f(x^*)$ positiv definit $\Rightarrow\ x^*$ lokales Minimum

4.1 Abstiegsalgorithmen

1. $x_1 \in \mathbb{R}^n$ Startvektor, $k := 1$
2. $\nabla f(x_k)$ klein genug?, d.h. $||\nabla f(x_k)|| < \varepsilon$? Dann fertig
3. Wähle Abstiegsrichtung $d_k \in \mathbb{R}^n$ mit $d_k^T \nabla f(x_k) < 0$, wähle eine Schrittweite $\alpha_k > 0$ mit $f(x_k + \alpha_k d_k) < f(x_k),\ x_{k+1} := x_k + \alpha_k d_k,\ k \to k+1$, GOTO 2.

$d_k^T \nabla f(x_k) < 0$, da:

$$f(x_k + d_k) = f(x_k) + d_k^T \nabla f(x_k) + \text{ HOT}^1 \overset{!}{<} f(x_k)$$

Bestimmung der Abstiegsrichtung

- **Methode des steilsten Abstiegs**: $d_k := -\nabla f(x_k)$
- **Quadratische Modelle**: $f(x_k + d) \approx f(x_k) + d^T \nabla f(x_k) + \frac{1}{2} d^T B_k d$
  B_k etwa die Hesse-Matrix $\nabla^2 f(x_k)$ oder eine Näherung (B_k stets positiv definit!)
  Minimum des quadr. Modells kann explizit ausgerechnet werden:
  $0 = \nabla f(x_k) + B_k d \Rightarrow\ B_k d = -\nabla f(x_k)$

  $$d = -B_k^{-1} \nabla f(x_k)$$

  - $B_k = I$: Methode des steilsten Abstiegs
  - $B_k = \nabla^2 f(x_k)$: Newtonmethode $\quad x_{k+1} = x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$

Konvergenz

Seien $x_k \in \mathbb{R}^n,\ k \in \mathbb{N}$, konvergiere $x_k \xrightarrow{k\to\infty} x^*$. Die Konvergenz ist

Q-sublinear	falls $\lim_{k\to\infty} \frac{\|\|x_{k+1}-x^*\|\|}{\|\|x_k - x^*\|\|} = 0$
R-superlinear	falls $\lim_{k\to\infty} \sqrt[k]{\|\|x_k - x^*\|\|} = 0$
von Ordnung p	falls $\frac{\|\|x_{k+1}-x^*\|\|}{\|\|x_k - x^*\|\|^p} = o(1)$

[1]higher order terms

Lokale Konvergenz der Newtonmethode

Habe f ein lokales Minimum bei $x^* \in \mathbb{R}^n$, in einer Umgebung D existiere $\nabla^2 f$ mit Lipschitzbedingung $||\nabla^2 f(x) - \nabla^2 f(y)|| \leq L \cdot ||x - y|| \quad \forall x, y \in D, L > 0$ (Lipschitz-Konstante), $\nabla^2 f(x^*) > 0$. Dann gilt:

- $\exists \varepsilon > 0 :\ ||x_1 - x^*|| < \varepsilon \Rightarrow\ x_k \xrightarrow{k\to\infty} x^*$ quadratisch mit der Newton-Methode

  $$x_{k+1} = x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

- (Dennis & Moré) x_k konvergiert mit $x_{k+1} = x_k - B_k^{-1} \nabla f(x_k), \quad k \in \mathbb{N}$, gegen x^*, B_k positiv definit, $\lim_{k\to\infty} \frac{\left|\left|\left(B_k - \nabla^2 f(x_k)\right)(x_{k+1}-x_k)\right|\right|}{||x_{k+1}-x_k||} = 0 \ \Rightarrow\ \frac{||x_{k+1}-x^*||}{||x_k - x^*||} = o(1), \quad k \to \infty$ (superlineare Konvergenz)

Konvergenz der Methode des steilsten Abstiegs

Sei die Folge $x_k \in \mathbb{R}^n,\ k \in \mathbb{N}$ bestimmt durch die Methode des steilsten Abstiegs

$$x_{k+1} = x_k - \nabla f(x_k), \quad k \in \mathbb{N}$$

∇f stetig, $(x_k)_k$ beschränkt $\Rightarrow\ ||\nabla f(x_k)|| \xrightarrow{k\to\infty} 0$, d.h. Grenzwert ist statischer Punkt

4.2 Konjugierte Richtungen

$$\text{Zielfunktion: } F(x) = a + b^T x + \frac{1}{2} x^T G x,\ x, b \in \mathbb{R}^n, a \in \mathbb{R}, G \in \mathbb{R}^{n \times n}, G \text{ positiv definit}$$

$$\text{Konjugiertheit: } d_j \in \mathbb{R}^n :\ {d_j}^T G d_k = 0 \quad \forall j \neq k$$

$$\text{Abstiegsmethode: } x_{k+1} = x_k + \alpha_k d_k, \quad \alpha_k = \arg\min_{\alpha > 0} F(x_k + \alpha d_k)$$

- Konjugierte Richtungen sind linear unabhängig.
- ${d_i}^T \nabla F(x_{l+1}) = 0 \quad \forall i = 1, ..., l$
- $\exists k \leq n :\ \nabla F(x_{k+1}) = 0$

Konstruktion von konjugierten Richtungen

$$d_1 = -\nabla F(x_1) = -g_1,\ d_k = -g_k + \sum_{j=1}^{k-1} \beta_{kj} d_j \quad \forall k > 1, \quad \beta_{kj} = \frac{{d_j}^T G g_k}{{d_j}^T G d_j}$$

Es gilt: ${g_i}^T g_j = 0 \quad \forall i \neq j$
$\Rightarrow\ \beta_{kj} = 0 \quad \forall j < k-1 \quad \Rightarrow\ \beta_k := \beta_{k,k-1} \quad \Rightarrow\ d_k = -g_k + \beta_k d_{k-1}$

$$\Rightarrow\ \beta_k = \frac{(g_k - g_{k-1})^T g_k}{(g_k - g_{k-1})^T d_{k-1}}$$

Varianten:

- Fletcher-Rieves: $\beta_k = \frac{||g_k||^2}{||g_{k-1}||^2}$
- Polak-Ribière: $\beta_k = \frac{(g_k - g_{k-1})^T g_k}{||g_{k-1}||^2}$

4.3 Konjugierte Richtungen aus quadratischen Modellen

$$d_k = \arg\min_{d \in \mathbb{R}^n} \{F(x_k) + d^T g_k + \frac{1}{2} d^T B_k d\}$$

$$B_k \text{ Näherung an die 2. Ableitung } \nabla^2 F(x_k), \quad H_k := B_k^{-1}$$

$$\text{Konjugiertheit: } (g_{i+1} - g_i)^T d_j = 0 \quad \forall i \neq j$$

Quasi-Newton-Bedingungen

- $H_{j+1}\gamma_j = \delta_j, \quad \gamma_j = g_{j+1} - g_j,\ \delta_j = x_{j+1} - x_j = \alpha_j d_j$
- $H_{j+1}\gamma_i = \delta_i \quad \forall i \leq j$

$$\text{Abstiegsrichnungen: } d_k = -H_k g_k, \quad g_k^T d_k = -g_k{}^T H_k g_k < 0$$

$$x_{k+1} = x_k + \alpha_k d_k, \quad \alpha_k = \arg\min_{\alpha > 0} F(x_k + \alpha d_k)$$

4.3.1 Broyden Linear Family (BLF)

$$H_{k+1} = H_k - \frac{H_k \gamma_k \gamma_k{}^T H_k}{\gamma_k{}^T H_k \gamma_k} + \frac{\delta_k \delta_k{}^T}{\delta_k{}^T \gamma_k} + \lambda_k \Big(\frac{\delta_k}{\delta_k{}^T \gamma_k} - \frac{H_k \gamma_k}{\gamma_k{}^T H_k \gamma_k} \Big) \Big(\frac{\delta_k}{\delta_k{}^T \gamma_k} - \frac{H_k \gamma_k}{\gamma_k{}^T H_k \gamma_k} \Big)^T$$

1. H_k symmetrisch $\Rightarrow$ H_{k+1} symmetrisch
2. $H_{k+1}\gamma_k = \delta_k$
3. $H_{k+1}\gamma_i = \delta_i \quad \forall i \leq k$

- DFP: $\lambda_k \equiv 0$
  $B_{k+1} = B_k - \frac{B_k \delta_k \delta_k^T}{\delta_k{}^T B_k \delta_k} + \frac{\gamma_k \gamma_k{}^T}{\delta_k{}^T \gamma_k} + \mu_k \Big(\frac{\gamma_k}{\delta_k{}^T \gamma_k} - \frac{B_k \delta_k}{\delta_k{}^T B_k \delta_k} \Big) \Big(\frac{\gamma_k}{\delta_k{}^T \gamma_k} - \frac{B_k \delta_k}{\delta_k{}^T B_k \delta_k} \Big)^T$
  - H_k positiv definit
  - F quadratische Funktion $\Rightarrow$ $||B_{k+1} - G||_{G^{-1}}^2 \leq ||B_k - G||_{G^{-1}}^2 - \frac{||G^{-\frac{1}{2}}(B_k \delta_k - \gamma_k)||_2^2}{||G^{\frac{1}{2}} \delta_k||_2^2}$
  - $(d_k = -B_k^{-1}, \alpha_k = 1 \Rightarrow)$ g_k konvergieren superlinear gegen 0
- BPGS $\lambda_k = \gamma_k{}^T H_k \gamma_k, \quad B_{k+1} = B_k - \frac{B_k \delta_k \delta_k{}^T B_k}{\delta_k{}^T B_k \delta_k} + \frac{\gamma_k \gamma_k{}^T}{\delta_k{}^T \gamma_k},$
  $G = G^{\frac{1}{2}} G^{\frac{1}{2}T},\ G^{-\frac{1}{2}} = (G^{\frac{1}{2}})^T, \quad ||A||_{G^{-1}} = ||G^{-\frac{1}{2}} A G^{-\frac{1}{2}}||_{\text{Frob.}}$
  $(||A||_{\text{Frob.}} = \sqrt{\sum_i \sum_j |a_{ij}|^2})$

4.3.2 Rang-1-Aufdatierung

$$B_{k+1} = B_k + \frac{(\gamma_k - B_k \delta_k)(\gamma_k - B_k \delta_k)^T}{(\gamma_k - B_k \delta_k)^T \delta_k}$$

4.4 Trust Region Methoden

4.4.1 Algorithmus

Bei der Näherung durch ein quadratisches Modell, werden Regionen Δ_k eingeführt, in denen man der Näherung "traut".

$$Q_k(d) = F(x_k) + d^T g_k + \frac{1}{2} d^T B_k d \approx F(x_k + d) \quad \text{Quadratisches Modell}$$

1. Suchrichtung $||d_k||_2 \leq \Delta_k :\ Q_k(0) - Q_k(d_k) \geq c_0 \cdot \big(Q_k(0) - \min_{||d||_2 \leq \Delta_k} Q_k(d)\big)$
   (typisch: $c_0 = \frac{1}{2}$)
2. $\alpha_k = 1,\ x_{k+1} = x_k + d_k$
3. $F(x_k + d_k) \geq F(x_k) \Rightarrow\ x_{k+1} := x_k$
4. B_k mit üblichen Formeln (z.B. DFP) bilden
5. Wahl von $\Delta_k :\quad \varrho_k = \frac{F(x_k) - F(x_k + d_k)}{Q_k(0) - Q_k(d_k)} \quad$ (ideal: $\varrho_k = 1$)
   $$\Rightarrow\ \Delta_{k+1} = \begin{cases} \frac{1}{2}\Delta_k & \varrho_k \leq \frac{1}{10} \\ \Delta_k & \frac{1}{10} < \varrho_k \leq \frac{7}{10} \\ 2\Delta_k & \frac{7}{10} < \varrho_k \end{cases}$$

Die B_k müssen bei diesen Methoden nicht positiv definit sein: $(B_k + \lambda I) d_k = -g_k$

4.4.2 Eigenschaften

- $\exists c_1 \in \mathbb{R} :\ \Delta_k \geq c_1 \cdot ||g_k|| \quad \forall k \in \mathbb{N}$
- $\exists c_2 \in \mathbb{R} :\ Q_k(0) - Q_k(d_k) \geq c_2 ||g_k||^2$
- $||g_k|| \neq 0 \forall k \in \mathbb{N} \Rightarrow\ \limsup_{k \to \infty} ||g_k|| = 0$

5 Lineare Nebenbedingungen

$$F \in \mathcal{C}^1(\mathbb{R}^n), \quad F(x) \to \min!, \quad x \in \mathbb{R}^n$$

$$\text{Nebenbedingungen: } \begin{cases} a_i^T x - b_i = 0 & 1 \leq i \leq m' \\ a_i^T x - b_i \geq 0 & m' < i \leq m \end{cases} \quad a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$$

- **Kuhn-Tucker-Bedingungen** (KT, KKT): $F \in \mathcal{C}^1(\mathbb{R}^n), x^* \in \mathbb{R}^n$ zulässig (erfüllt NB): x^* lokales Minimum $\Rightarrow$ $\exists \lambda_i$ Lagrange-Multiplikatoren: $\nabla F(x^*) = \sum_{i=1}^m \lambda_i a_i, \lambda_i \geq 0\ \forall i > m', \lambda_i = 0\ \forall i : a_i^T x^* - b_i \neq 0$
- $F \in \mathcal{C}^2,\ x^*$ KT-Punkt: X^* lokales Minimum $\Rightarrow$ $d^T \nabla^2 F(x^*) d \geq 0\ \forall d \in \mathbb{R}^n :\ d^T a_i = 0\ \forall i \in I^* := \{j :\ a_j^T x^* - b_j = 0\}$
- $F \in \mathcal{C}^2, x^*$ KT-Punkt, $\hat{I} := \{i \in I^* :\ \lambda_i \neq 0 \vee i \leq m'\}$:
  $d^T \nabla^2 F(x^*) d > 0\ \forall d \neq 0 :\ d^T a_i = 0\ \forall i \in \hat{I} \Rightarrow\ x^*$ lokales Minimum

5.1 Methode der aktiven Mengen

Es wird die Dimension des Raums, in dem nach x^* gesucht wird, reduziert, indem man besonders die Nebenbedingung(en) betrachtet, die aktiv sind, d.h. die als Gleichung erfüllt sind.

$x_k \in \mathbb{R}^n$ aktueller, zulässiger Vektor, I_k aktuelle Schätzung der Menge der Indizes der aktiven Nebenbedingung, insb. $\{1,..,m'\} \subset I_k \; \forall k$
$g_k = \nabla F(x_k) \notin \text{span}\{a_i\}_{i\in I_k} \Rightarrow$ wähle Suchrichtung $d_k: \; {d_k}^T g_k < 0, {d_k}^T a_i = 0 \; \forall i \in I_k$
$\Rightarrow \; x_{k+1} = x_k + \alpha_k d_k$
α_k durch Nebenbedingung $c_p(x_{k+1}) = 0$ eingeschränkt $\Rightarrow \; I_{k+1} = I_k \cup \{p\}$
Wiederholung, bis $g_k \in \text{span}\{a_i\}_{i\in I_k}$, dann Berechnung der Lagrange-Multiplikatoren λ_i : $g_k = \sum_{i\in I_k} \lambda_i a_i$

- $\lambda_i \geq 0 \; \forall i > m' \Rightarrow$ KT-Punkt
- $\exists k > m': \; \lambda_k < 0 \Rightarrow \; -g_k = -\lambda_l a_l - \sum_{i\in I_k\setminus\{l\}} \lambda_i a_i, \; I_{k+1} = I_k\setminus\{l\}$

Entfernung/Hinzufügung eines Indizes in I_k nur bei $F(x_k) = \min\{F(x): \; c_i(x) = 0 \; \forall i \in I_k\}, \; F(x_{k+1}) < F(x_k) \; \forall k \Rightarrow \; \exists$ nur endliche Anzahl von Änderungen der Menge der aktiven Nebenbedingungen
Variante:
Nur in Kern$\{{A_k}^T\} = \{d \in \mathbb{R}^n: \; {A_k}^T d = 0\}$ rechnen:
Sei $Z_k \in \mathbb{R}^{n\times n-l}: \; {A_k}^T Z_k = 0$ mit linear unabhängigen Spalten. $\Rightarrow \; \nu_k = \arg\min\{F(x_k + Z_k\nu)\}, \; d_k = Z_k\nu_k = -Z_k({Z_k}^T G Z_k)^{-1}{Z_k}^T g_k$

6 Nichtlineare Nebenbedingungen

$$F(x) = \min_{x\in\mathbb{R}^n}!, \; F: \mathbb{R}^n \to \mathbb{R} \in \mathcal{C}^1(\mathbb{R}^n)$$

$$\text{Nebenbedingungen} \quad c_i(x) = 0 \; \forall 1 \leq i \leq m', \quad c_i(x) \geq 0 \; \forall m' < i \leq m$$

- x^* lokales Minimum $:\Leftrightarrow \; F(x) \geq F(x^*) \; \forall x \in U$ Umgebung von $x^*, \; x, x^*$ zulässig
- x^* KT-Punkt $:\Leftrightarrow \; x^*$ zulässig, $\exists \lambda_i: \; \nabla F(x^*) = \sum_{i=1}^m \lambda_i \nabla c_i(x), \; \lambda_i \geq 0 \; \forall i > m', \lambda_i = 0 \; \forall i: \; c_i(x^*) > 0$
- x^* zulässig, lokales Minimum, $\nabla c_i(x^*), \; i \in I^* = \{i: \; c_i(x^*) = 0\}$ lin. unabhängig $\Rightarrow \; x^*$ KT-Punkt
- Bedingungen 2. Ordnung: $L(x,\lambda) := F(x) - \sum_{i=1}^m \lambda_i c_i(x), \; \lambda = (\lambda_1,..,\lambda_m)^T \in \mathbb{R}^m$:
x^* KT-Punkt $\Rightarrow \; \nabla_x L(x,\lambda) = \nabla F(x) - \sum_{i=1}^m \lambda_i \nabla c_i(x), \; \nabla_x L(x^*,\lambda) = 0$
- x^* KT-Punkt, $\nabla c_i(x^*)$ linear unabh. $\forall i \in I^*, \; x^*$ lokales Minimum, λ^* Lagrange-Multiplikatoren $\Rightarrow \; d^T \nabla^2_{xx} L(x^*,\lambda^*) d \geq 0 \; \forall d: \; d^T \nabla c_i(x^*) = 0 \; \forall i \in I^*$
- x^* KT-Punkt, $\nabla c_i(x^*)$ linear unabh. $\forall i \in I^*, \; \lambda^*$ Lagrange-Multiplikatoren, $d^T \nabla^2_{xx} L(x^*,\lambda^*) d > 0 \; \forall d: \; d^T \nabla c_i(x^*) = 0 \; \forall i \in \hat{I}, d \neq 0 \Rightarrow \; x^*$ lokales Minimum

6.1 Penalty-Funktionen

L_1-Penalty-Funktion

$$\Phi_1(x) = F(x) + \sum_{i=1}^{m'} \mu_i |c_i(x)| + \sum_{i=m'+1}^{m} \mu_i \max\big(0, -c_i(x)\big)$$

- $\nabla c_i(x^*)$ lin. unabh. $\forall i \in I^*, \; x^*$ Lösung des Opt.problems, λ_i^* Lagrange-Multiplikatoren bei $x^*, \; \mu_i > |\lambda_i^*|, \; i \in I^* \Rightarrow \; x^*$ Minimum von Φ_1

Fletcher's Penalty-Funktion

Hier für $m = m', \nabla c_i(x)$ lin. uabhängig $\forall x \in \mathbb{R}^n$

$$\Phi_F(x) = F(x) - \sum_{i=1}^m \lambda_i(x) c_i(x) + \frac{r}{2} \cdot \sum_{i=1}^m c_i(x)^2, \quad \lambda_i(x) = \arg\min_{\lambda\in\mathbb{R}^m} ||\nabla F(x) - \sum_{i=1}^m \lambda_i \nabla c_i(x)||_2^2$$

- x^* KT-Punkt, $\nabla c_i(x^*)$ linear unabh. $\forall i \in I^*, \; \lambda^*$ Lagrange-Multiplikatoren, $d^T \nabla^2_{xx} L(x^*,\lambda^*) d > 0 \; \forall d: \; d^T \nabla c_i(x^*) = 0 \; \forall i \in \hat{I}, d \neq 0, \; r$ ausreichend groß $\Rightarrow \; x^*$ lokales Minimum von Φ_F

Allgemeine Penalty-Funktionen

$$\Phi_r(x) = F(x) + r \cdot P\big(c(r)\big)$$

$$P: \mathbb{R}^m \to \mathbb{R}_0^+, \; P(c) = 0 \; \forall c \in \mathbb{R}^m: \; c_1 = .. = c_{m'} = 0, \; c_{m'} = .. = c_m \geq 0; \text{ sonst: } P(c) > 0$$

- F stetig, Nebenbedingungen widerspruchsfrei, $(r_k)_k$ monoton wachsend, $r_k \stackrel{k\to\infty}{\longrightarrow} \infty, x(r_k)$ Minima von $\Phi_r(x) \Rightarrow$ alle Häufungspunkte von $\big(x(r_k)\big)_k$ sind Lösungen der Aufgabe
- P stetig: $\Phi(x) = F(x) + P\big(c(x)\big)$ besitzt Minimum in $\hat{x} \Rightarrow \; \hat{x}$ minimiert $F(x)$ mit den Nb. $c_i(x) = c_i(\hat{x})$

6.2 Augmentierte Lagrange-Methode

Wir wollen $\Phi(x,\lambda,r)$ minimieren:

$$m = m': \quad \Phi(x,\lambda,r) = F(x) - \sum_{i=1}^m \lambda_i c_i(x) + \frac{r}{2} \sum_{i=1}^m c_i(x)^2$$

Sei $x(\lambda, r) \in \mathbb{R}^n$ Minimum ($\lambda \in \mathbb{R}^m, r > 0$)

- $f: \mathbb{R}^{m_1} \to \mathbb{R}^n, g: \mathbb{R}^{m_1} \to \mathbb{R}^m: \; ||f(y)|| \sim ||g(y)|| :\Leftrightarrow \; \frac{||f(y)||}{||g(y)||} < \infty, \; \frac{||g(y)||}{||f(y)||} < \infty$ für einen gegebenen Bereich der y
- x^* KT-Punkt, $\nabla c_i(x^*)$ linear unabh. $\forall i \in I^*, \; \lambda^*$ Lagrange-Multiplikatoren, $d^T \nabla^2_{xx} L(x^*,\lambda^*) d > 0 \; \forall d: \; d^T \nabla c_i(x^*) = 0 \; \forall i \in \hat{I}, d \neq 0$
  - $\Rightarrow \; ||x(\lambda,r) - x^*|| \sim ||c\big(x(\lambda,r)\big)||$ in einer Umgebung von x^*, in der die Hesse-Matrix von $L(x) = F(x) - \sum_{i=1}^m \lambda_i^* c_i(x) + \frac{r}{2}\sum_{i=1}^m c_i(x)^2$ positiv definit ist
  - $||x(\lambda,r) - x^*|| = O(\frac{||\lambda-\lambda^*||}{r-r^*}) \; \forall r > r^*$

- r groß genug, damit $x(\lambda)$ in einer Umgebung von $x^* \Rightarrow \quad ||\lambda^+ - \lambda^*|| = O(\frac{||\lambda-\lambda^*||}{r-r^*})$, wobei $\lambda^+ = (\lambda - r) \cdot c\big(x(\lambda)\big)$
- r groß genug $\Rightarrow \quad \lambda \to \lambda^*$

- $\phi(\lambda) := \Phi\big(x(\lambda), \lambda, r\big): \quad \phi(\lambda) \leq \Phi(\lambda^*) \quad \forall \lambda \in \mathbb{R}^m$
- $\phi(\lambda)$ maximieren:
  - Methode des steilsten Anstiegs: $\lambda^+ = \lambda + r\nabla c$
  - $\nabla^2\phi = -A^T B^{-1} A > 0$, wobei $A = (\nabla c_1, .., \nabla c_m), B = \nabla^2\Phi$
  - Newton-Methode: $\lambda^+ = \lambda - (\nabla^2\phi)^{-1}$, für große $r: \ \lambda^+ \approx \lambda - r \cdot c$

6.3 Lineare Ungleichungs- und allgemeine Gleichungs-Nb.

$$c_i(x) = 0 \quad \forall 1 \leq i \leq m \quad \text{allgemeine Gleichheits-Nb.}$$

$$x \in S \subset \mathbb{R}^n,\ S \text{ Inneres+Rand eines Polyeders; Lineare Ungleichungs-Nb.}$$

$$\Rightarrow \ x(\lambda) = \arg\min_{x\in S} \Phi(x, \lambda, r) = \arg\min_{x \in S} F(x) - \sum_{i=1}^{m} \lambda_i c_i(x) + \frac{r}{2} \sum_{i=1}^{m} {c_i}^2(x)$$

$\lambda \in \mathbb{R}^m$ sind die Lagrange-Multiplikatoren der Gleichheitsnebenbedingugen und es wird bei $\Phi(\lambda) = \Phi\big(x(\lambda), \lambda, r\big)$ maximiert.

6.4 Allgemeine Ungleichungs- und Gleichheits-Nb.

Es werden Schlupfvariablen eingeführt:

$$z_i \geq 0 \ \ \forall m' < i \leq m, \quad c_i(x) = 0 \ \ \forall 1 \leq i \leq m' \quad c_i(x) - z_i = 0 \ \ \forall m' < i \leq m$$

Wir minimieren

$$\Phi(x, \lambda, r) = F(x) - \sum_{i=1}^{m'} \lambda_i c_i(x) + \frac{r}{2} \sum_{i=1}^{m'} c_i(x)^2 - \sum_{i=m'+1}^{m} \lambda_i\big(c_i(x) - z_i\big) + \frac{r}{2} \sum_{i=m'+1}^{m} \big(c_i(x) - z_i\big)^2$$

unter der Voraussetzung $z_i \geq 0 \,\forall m' < i \leq m$:
Für $m' + 1 \leq i \leq m$ gilt:

$$\min_{z_i \geq 0} -\lambda_i\big(c_i(x) - z_i\big) + \frac{r}{2}\big(c_i(x) - z_i\big)^2 = -\frac{{\lambda_i}^2}{2r} \quad \text{falls } z_i = c_i(x) - \frac{\lambda_i}{r} \geq 0,$$

$$\min_{z_i \geq 0} -\lambda_i\big(c_i(x) - z_i\big) + \frac{r}{2}\big(c_i(x) - z_i\big)^2 = -\frac{{\lambda_i}^2}{2r} + \frac{r}{2}\Big(c_i(x) - \frac{\lambda_i}{r}\Big)^2 \quad \text{falls } c_i(x) - \frac{\lambda_i}{r} < 0,\ z_i = 0$$

Es folgt eine Aufdatierungsformel für die λ_i:

$$\lambda_i \to \lambda_i^* = \begin{cases} \lambda_i - rc_i(x) & 1 \leq i \leq m' \\ \max\{0, \lambda_i - rc_i(x)\} & m' + 1 \leq i \leq m \end{cases}$$

6.5 Lineare Näherung an nichtlineare Nebenbedingungen

$$x_{k+1} = x_k + \alpha_k \cdot d_k, \quad k = 0, 1, 2, ..$$

wobei x_0 Startvektor und d_k Suchrichtungen sind mit

$$c_i(x_k) + {d_k}^T \nabla c_i(x_k) = 0 \quad \forall 1 \leq i \leq m'; \quad c_i(x_k) + {d_k}^T \nabla c_i(x_k) \geq 0 \quad \forall m' < i \leq m$$

6.6 Newton-Methode zur Berechnung von KT-Punkten

Wende das Standard-Newton Verfahren an auf

$$\nabla F(x) = \sum_{i=1}^{m} \lambda_i \nabla c_i(x), \quad c_i(x) = 0 \ \forall i = 1, .., m$$

$$\Rightarrow \ \nabla F(x^k) + \nabla^2 L(x^k)\delta^k - \sum_{i=1}^{m} \lambda_i^k - \sum_{i=1}^{m} \nu_i^k \nabla c_i(x^k) = 0, \quad \nabla^2 L = \nabla^2 F - \sum_{i=1}^{m} \eta_i^k \nabla^2 c_i$$

Es wird die folgende Gleichung nach $(\delta^k, -\lambda^k - \eta^k) \in \mathbb{R}^{n+m}$ aufgelöst:

$$\begin{pmatrix} \nabla^2 L(x^k) & A^k \\ A^{kT} & 0 \end{pmatrix} \begin{pmatrix} \delta^k \\ -\lambda^k - \eta^k \end{pmatrix} = \begin{pmatrix} -\nabla F(x^k) \\ -c(x^k) \end{pmatrix}$$

wobei $A^k := \big(\nabla c_1(x^k) \quad \dots \quad \nabla c_m(x^k)\big)$, $c(x^k) := \big(c_i(x^k)\big)_{i=1}^m$

- $x^{k+1} = x^k + \delta^k, \quad \lambda^{k+1} = \lambda^k + \eta_k, \quad x^k, x^{k+1}, \delta^k \in \mathbb{R}^n,\ \lambda^k, \lambda^{k+1}, \eta^k \in \mathbb{R}^m$
- $\nabla^2 L(x^k)$ positiv definit $\Rightarrow \ \delta^k = \arg\min_{\delta \in \mathbb{R}^n} F(x^k) + \delta^T \nabla F(x^k) + \frac{1}{2}\delta^T \nabla^2 L(x^k)\delta$ mit Nebenbedingung $c_i(x^k) + \delta^T \nabla c_i(x^k) = 0 \ \forall i = 1, .., m$

6.7 Quasi-Newton-Methoden für Minimierung mit Nb.

1. $x_0 \in \mathbb{R}^n$ Startvektor, $k = 0$
2. $B_k \approx \nabla^2 L(x^k)$ mit Quasi-Newton-Verfahren berechnen
3. $\delta_k = \arg\min_{\delta \in \mathbb{R}^n} F(x_k) + \delta^T \nabla F(x_k) + \frac{1}{2}\delta^T B_k \delta$ mit Nb. $c_i(x_k) + \delta_k^T \nabla c_i(x_k) = 0 \ \forall i = 1, .., m'$, $c_i(x_k) + \delta_k^T \nabla c_i(x_k) \geq 0 \ \forall i = m' + 1, .., m$
   Kontrolle mit Gütefunktion, ob $\Phi(x_{k+1}) < \Phi(x_k)$
4. $x_{k+1} = x_k + \alpha_k \delta_k$, $\alpha_k > 0$
5. x_{k+1} nahe genug an der Lösung $\Rightarrow$ fertig
   sonst: $k \to k + 1$, GOTO 2.

Mögliche Gütefunktion:

- $$\Phi(x) = F(x) + \sum_{i=1}^{m'} \mu_i |c_i(x)| + \sum_{i=m'+1}^{m} \mu_i \max\big(0, -c_i(x)\big)$$

- $$\Phi_1(x) = F(x) + \sum_{i=1}^{m} \mu_i |c_i(x)|$$

Sätze:

- $\lambda \in \mathbb{R}^m$ Lagrangemultiplikator von $F(x_k)+\delta^T \nabla F(x_k)+\frac{1}{2}\delta^T B_k \delta \to \min$ mit Nb. $c_i(x_k)+\delta_k^T \nabla c_i(x_k) = 0 \;\forall i = 1,..,m',\; c_i(x_k)+\delta_k^T \nabla c_i(x_k) \geq 0 \;\forall i = m'+1,..,m$ und $\mu_i \geq |\lambda_i| \;\forall i = 1,..,m \Rightarrow\; \delta_k = \arg\min_{\delta \in \mathbb{R}^n} F(x_k)+\delta^T \nabla F(x_k)+\frac{1}{2}\delta^T B_k \delta + \sum_{i=1}^{m'} \mu_i |c_i(x_k)+\delta^T \nabla c_i(x_k)| + \sum_{i=m'+1}^{m} \mu_i \max\big(0, -c_i(x_k) - \delta^T \nabla c_i(x_k)\big)$
- $\mu_i \geq |\lambda_i|$ wie eben, $\delta_k \neq 0 \Rightarrow\; \exists \alpha > 0 :\; \Phi(x_k + \alpha\delta_k) < \Phi(x_k)$
- $(x_k)_k, (\delta_k)_k$ beschränkt $\Rightarrow\; \exists \alpha_k :\; \Phi_i(x_k + \alpha_k \delta_k) \leq \Phi_i(x_k) + \frac{1}{10}\alpha_k \sigma_k,\; \sigma_k := \delta_k \nabla F(x_k) - \sum_{i=1}^{m} \mu_i |c_i(x)|$ und alle Häufungspunkte von $(x_k)_k$ sind KT-Punkte
- $B_k = \nabla^2 L(x^*) \Rightarrow$ quadratische Konvergenz : $||x_k + \delta_k - x^*|| = O\big(||x_k - x^*||^2\big)$
- In Praxis: etwa B_k durch BFGS, wobei γ_k durch $ga\hat{m}ma_k := \vartheta_k \gamma_k + (1-\vartheta_k) B_k \delta_k,\; 0 < \vartheta_k \leq 1$ ersetzt wird, etwa ϑ_k so, dass $\hat{\gamma_k}^T \delta_k \geq \frac{1}{10}\delta_k^T B_k \delta_k > 0$
- x^* zulässiges, lokales Minimum mit hinreichenden Bedingungen 2. Ordnung, Quasi-Newton-Methode mit $p^T B_k q = p^T \nabla_{xx}^2 L(x^*, \lambda^*) q \;\forall p, q :\; p^T c_i(x^*) = q^T \nabla c_i(x^*) = 0 \;\forall i \in \hat{I} \Rightarrow\; ||x_{k+2} - x^*|| = 0\big(||x_k - x^*||^2\big)$

6.8 Der Maratos-Effekt

Manchmal wird ein Schritt durchgeführt, bei dem die Zielfunktion F vergrößert wird. Dieser führt eventuell zu einer schnelleren Konvergenz. Verwirft man diesen Schritt (etwa durch Kontrolle mit einer Gütefunktion), so verringert man die Konvergenzgeschwindigkeit.

Abhilfen:

- Man erlaubt einige Vergrößerungen der Gütefunktion $\Phi_1(x)$.
- Man verwendet die Gütefunktion $\Phi_F(x)$, da dort dieser Effekt nicht auftritt.

7 Eindimensionale Minimierungsverfahren

Bei vielen Abstiegsmethoden muss α als eindimensionales Minimum gewählt werden:

$$\alpha_k = \arg\min_{\alpha > 0} F(x_k + \alpha d_k)$$

Im Folgenden suchen wir $x^* = \arg\min_{x \in [a,b]} f(x)$

7.1 Verfahren des goldenen Schnitts

1. $k = 0,\; a_0 := a,\; b_0 := b$
2. $v_k := a_k + \frac{3-\sqrt{5}}{2}(b_k - a_k),\; w_k := a_k + \frac{\sqrt{5}-1}{2}(b_k - a_k)$
3. $f(v_k) > f(w_k) \Rightarrow\; a_{k+1} := v_k,\; b_{k+1} := b_k$
   $f(v_k) \leq f(w_k) \Rightarrow\; a_{k+1} := a_k,\; b_{k+1} := w_k$
4. $|b_{k+1} - a_{k+1}| < \varepsilon \Rightarrow$ fertig
   $|b_{k+1} - a_{k+1}| \geq \varepsilon \Rightarrow\; k \to k+1$, GOTO 2.

7.2 DSCP (lokale quadratische Approximation)

Sei $f \in \mathcal{C}^1\big((a,b)\big)$ und strikt konvex.

1. Wähle $x_0 > x_1 > x_2 :\; f(x_0) > f(x_1), f(x_1) < f(x_2)$
2. Interpoliere mit quadratischem Polynom an den Stellen $\big(x_0, f(x_0)\big),\; \big(x_1, f(x_1)\big),\; \big(x_2, f(x_2)\big)$ und berechne das Minimum $\overline{x}$
3. Ergebnis nicht gut genug $\Rightarrow\; x_1 := \overline{x}$, GOTO 1.

8 Globale Minimierungs- und Optimierungsverfahren

Bei den bisherigen Vefahren wurden lokale Extrema, stationäre Punkte und KT-Punkte mit Hilfe von Taylor-Entwicklungen und quadratischen Modellen berechnet.
Wir suchen nun globale Minima.

8.1 Globale Näherungsverfahren

- Polynome, Polynominterpolation und -approximation
  $+$: einfaches Verfahren
  $-$: kaum geeinget für globale Näherung und höhere Dimensionen
- Splines (stückweise Polynome), z.B. zur Interpolation
  $+$: geeignet für globale Näherung
  $-$: nicht geeinget für Dimensionen $n \gg 10$
- rationale Approximation
  $+$: geeignet für globale Näherung
  $-$: nicht geeinget für Dimensionen $n \gg 10$

8.2 Eine neue multivariable Näherungsmethode

Zielfunktion: $f : \Omega \Rightarrow \mathbb{R},\quad \Omega \subset \mathbb{R}^n,\; n > 1$ **Näherung:** $s : \Omega \Rightarrow \mathbb{R},\quad s \approx f,\; s$ “einfach”

s soll von $f(x_1), .., f(x_m),\; x_1, .., x_m \in \Omega$ pw. verschieden, abhängen, etwa $s(x_i) = f(x_i) \;\forall i = 1, .., m$ (Interpolation)

$$s(x) = \sum_{i=1}^{m} \lambda_i \varphi(||x - x_i||_2) \quad \text{mit } \lambda_i \in \mathbb{R},\; \varphi : \mathbb{R} \to \mathbb{R} \text{ stetig}$$

- $\varphi(r) := e^{-c^2r^2}, \quad c > 0$
- $\varphi(r) := \frac{1}{\sqrt{c^2+r^2}}, \quad c > 0$
- $\varphi(r) := r^{2k+1}, \quad k \in \mathbb{N}_0$
- $\varphi(r) := \sqrt{r^2+c^2}, \quad c > 0$
- $\varphi(r) := r^2 \log r$

Wir verlangen:

$$s(x_j) = \sum_{i=1}^{m} \lambda_i \varphi\big(||x_j - x_i||_2\big) \stackrel{!}{=} f(x_j) \quad \forall j = 1,..,m$$

Dazu wird das folgende lineare Gleichnugssystem gelöst:

$$A\lambda = b, \quad A = \big(\varphi(||x_j - x_i||_2)\big)_{j,i=1}^{m},\ b = \big(f(x_j)\big)_{j=1}^{m}$$

λ ist für reguläre A eindeutig bestimmt.

Seien $x_1,..,x_m$ pw. verschieden:

- $\varphi(r) := e^{-c^2r^2},\ c > 0,\ f \in \mathcal{C}(\mathbb{R}^n)$
  $\Rightarrow\ \exists!\ \lambda_1,..,\lambda_m:\ s(x_j) = \sum_{i=1}^{m} \lambda_i\varphi\big(||x_j - x_i||_2\big) = f(x_j)\ \forall j = 1,..,m$
- $\varphi(r) := \frac{1}{\sqrt{c^2+r^2}},\ c > 0,\ f \in \mathcal{C}(\mathbb{R}^n)$
  $\Rightarrow\ \exists!\ \lambda_1,..,\lambda_m:\ s(x_j) = \sum_{i=1}^{m} \lambda_i\varphi\big(||x_j - x_i||_2\big) = f(x_j)\ \forall j = 1,..,m$
- $g(y) := \big|\sum_{j=1}^{m} \lambda_j e^{ix_j\cdot y}\big|$
  - $\varphi(r) = r: \quad g(y) = O(||y||)$
    $\Rightarrow\ \exists!\ \lambda_1,..,\lambda_m:\ s(x_j) = \sum_{i=1}^{m} \lambda_i\varphi\big(||x_j - x_i||_2\big) = f(x_j)\ \forall j = 1,..,m$
  - $\varphi(r) = \sqrt{r^2+c^2},\ c > 0: \quad g(y) = O(||y||)$
    $\Rightarrow\ \exists!\ \lambda_1,..,\lambda_m:\ s(x_j) = \sum_{i=1}^{m} \lambda_i\varphi\big(||x_j - x_i||_2\big) = f(x_j)\ \forall j = 1,..,m$
  - $\varphi(r) = r^2 \log r: \quad g(y) = O(||y||^2)$
    $\Rightarrow\ \exists!\ \lambda_1,..,\lambda_m:\ s(x_j) = \sum_{i=1}^{m} \lambda_i\varphi\big(||x_j - x_i||_2\big) = f(x_j)\ \forall j = 1,..,m$
  - $\sum_{j=1}^{m} \lambda_i q(x_j) = 0\ \forall q \in \mathbb{P}_n^k \Rightarrow\ |g(y)| = O(||y||^{k+1}),\ ||y|| \leq 1$