

## 7 Nichtlineare Optimierung

### Bisherige Optimierungsmodelle

$$\begin{array}{l} \min Z(\mathbf{x}) \\ \text{udN } \mathbf{x} \quad M \end{array}$$

charakterisiert durch

- lineare Zielfunktion
- Beschränkung zulässigen Bereichs
  - \* per Wertebereich Entscheidungsvariablen (zB  $\mathbf{x} \in \mathbb{R}_+^n$ )
  - \* per Nebenbedingungen, lineare Funktionen in  $\mathbf{x}$

Nichtlineare Optimierungsmodelle liegen vor, wenn

- Zielfunktion nichtlinear
- und / oder Nebenbedingungen nichtlinear

Entscheidungsvariablen durchgängig reellwertig

- $\mathbf{x} \in \mathbb{R}^n$  **unrestringiertes** Problem
- $\mathbf{x} \in M \subset \mathbb{R}^n$  **restringiertes** Problem, Restriktionen als Nebenbedingungen-Ungleichungssystem

Generell

- nichtlineare Nebenbedingungen  
"schwieriger" als nichtlineare Zielfunktion
- restringierte Probleme  
"schwieriger" als unrestringierte Probleme
- zwischen lokalen / relativen und globalen / absoluten Minima (Maxima) zu unterscheiden
- Hoffnung auf "generelle Methode"  
(angesichts der Vielfalt nichtlinearer Funktionen)  
offensichtlich müßig

## 7.1 Grundbegriffe und Optimalitätsbedingungen

### Optimierungsmodell

$$\begin{array}{ll} \min Z(\mathbf{x}) & \mathbf{x} \text{ n-dimensional,} \\ & Z \text{ auf } \mathbf{R}^n \text{ definiert} \\ \text{u.d.N } \mathbf{x} \in M & M \subseteq \mathbf{R}^n \end{array}$$

- $\mathbf{x} \in M$   **$\mathbf{x}$  ist zulässige Lösung**
- $Z(\mathbf{x}^*) \leq Z(\mathbf{x}) \quad \mathbf{x} \in M$   **$\mathbf{x}^*$  ist globaler Minimalpunkt**  
 $Z^* := \min_{\mathbf{x} \in M} Z(\mathbf{x})$   **$Z^*$  ist globales Minimum**
- Satz von **Weierstraß**:  
 $M$  kompakt,  $Z$  stetig auf  $M$   
 globaler Minimalpunkt von  $Z$  auf  $M$  existiert
- $\varepsilon > 0$   
 $U(\mathbf{x}) := \{\mathbf{x}' \in \mathbf{R}^n; |\mathbf{x} - \mathbf{x}'| < \varepsilon\}$   
 $U(\mathbf{x})$  ist (offene)  $\varepsilon$ -**Umgebung** von  $\mathbf{x}$
- $\varepsilon > 0: Z(\mathbf{x}) \leq Z(\mathbf{x}') \quad \mathbf{x}' \in M \cap U(\mathbf{x})$   
 $\mathbf{x}$  ist **lokaler Minimalpunkt**  
 $Z(\mathbf{x})$  ist **lokales Minimum**

Für lineare Optimierungsmodelle war

- lokaler Minimalpunkt auch globaler Minimalpunkt
- $M$  konvexes Polyeder (mit endlich vielen Ecken),  
 +: Existenz optimaler Lösung Lösung in Eckpunkt

Für nichtlineare Optimierungsmodelle gilt keine dieser Eigenschaften im allgemeinen

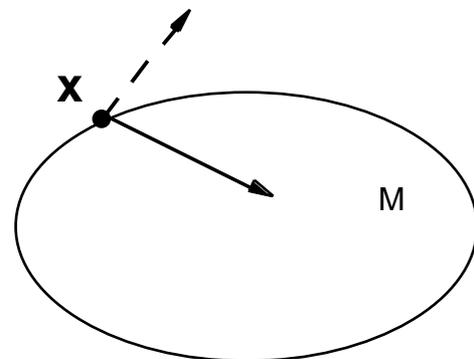
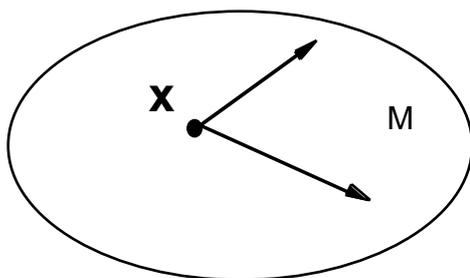
# LOKALES VORANSCHREITEN

Simplex bewegte sich  
 von zulässigem Punkt (Eck)  
 über zulässige Punkte (entlang Kante)  
 in Z-verbessernder Richtung (Z monoton fallend)  
 zu zulässigem Punkt (Eck)

Versuch der Übertragung der Idee benötigt Fassungen für

- "über zulässige Punkte"
- "in Z-verbessernder Richtung"
- uam

- $\mathbf{x} \in M, \mathbf{s} \in \mathbb{R}^n, \theta > 0$ :  $(\mathbf{x} + \theta \mathbf{s}) \in M$   $\iff \mathbf{s}$  ist **zulässige Richtung**
- $\mathbf{x}$  im Innern von  $M$ : alle  $\mathbf{s} \in \mathbb{R}^n$  zulässig
- $\mathbf{x}$  auf Rand von  $M$ : nicht alle  $\mathbf{s} \in \mathbb{R}^n$  zulässig



- $Z$  partiell differenzierbar nach  $\mathbf{x}, \mathbf{x} \in \mathbb{R}^n$   
 $\mathbf{g}(\mathbf{x}) := \text{grad } Z(\mathbf{x}) := ( \partial Z(\mathbf{x}) / \partial x_1, \dots, \partial Z(\mathbf{x}) / \partial x_n )^T$   
 $\mathbf{g}$  ist **Gradient** von  $Z$  in  $\mathbf{x}$   
 weist in Richtung stärksten Anstiegs  $Z$ ,  
 $\mathbf{g}(\mathbf{x})$  orthogonal zu Hyperfläche  $Z(\mathbf{x}) = z$   
 -  $\mathbf{g}$  ist Richtung größter Verbesserung  
 (Verkleinerung) von  $Z$ ,  $-\mathbf{g}$  orthogonal zu  $Z(\mathbf{x}) = z$

Bezeichne

$S(\mathbf{x}) := \{ \mathbf{s} ; \mathbf{s} \text{ zulässige Richtung in } \mathbf{x} \}$   
 die Menge aller in  $\mathbf{x}$  zulässigen Richtungen

**Satz 7.1.01 : Verbessernde Richtungen**

Sei  $Z(\mathbf{x})$  stetig differenzierbar

$\mathbf{x}, \mathbf{s} \in \mathbb{R}^n$

$\mathbf{s}^T \mathbf{g}(\mathbf{x}) < 0 \quad ( \quad \mathbf{s}^T (-\mathbf{g}(\mathbf{x})) > 0 \quad )$

$\forall \epsilon > 0: \quad Z(\mathbf{x} + \epsilon \mathbf{s}) < Z(\mathbf{x}) \quad \forall \epsilon \in ]0, \epsilon_0[$

oB

Veranschaulichung:

wenn  $\mathbf{s}^T \mathbf{g}(\mathbf{x}) < 0$ , dann bilden  $\mathbf{s}$  und  $\mathbf{g}$  stumpfen Winkel  
 $\mathbf{s}$  und  $-\mathbf{g}$  spitzen Winkel

Fortschreiten im spitzen Winkel zu  $-\mathbf{g}$  verkleinert  $Z$

$\mathbf{x} \in M, \mathbf{s} \in S(\mathbf{x})$  zulässiges, verbesserndes  
 Voranschreiten

OPTIMALITÄTSBEDINGUNGEN

**Satz 7.1.02 : Notwendige Bedingungen**

Sei  $Z(\mathbf{x})$  stetig differenzierbar

$\mathbf{x}$  lokaler Minimalpunkt von  $Z$  auf  $M$

$\mathbf{s}^T \mathbf{g}(\mathbf{x}) \geq 0 \quad \forall \mathbf{s} \in S(\mathbf{x}) \quad (1)$

$\mathbf{x}$  innerer Punkt von  $M \quad \mathbf{g}(\mathbf{x}) = 0 \quad (2)$

entsprechend Def's lokaler Minimalpunkt  
 + zulässige Richtung:

In einer Umgebung von  $\mathbf{x}$  kann  $Z$  nicht verkleinert werden

$\mathbf{x}$  ist **stationärer Punkt**

$\mathbf{x}$  innerer Punkt von  $M \quad S(\mathbf{x}) = \mathbb{R}^n \quad (2)$

Für zweimal partiell differenzierbares  $Z(\mathbf{x})$   
 heißt (symmetrische) Matrix  
 der zweiten partiellen Ableitungen **Hesse-Matrix H**

$$\mathbf{H} := \begin{pmatrix} \frac{\partial^2 Z}{\partial x_1^2} & \cdots & \frac{\partial^2 Z}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 Z}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 Z}{\partial x_n^2} \end{pmatrix}$$

und  $\mathbf{H}(\mathbf{x})$  Hesse-Matrix am Punkt  $\mathbf{x}$

symmetrische  $n \times n$  Matrix  $\mathbf{A}$  heißt

**positiv semidefinit**      wenn  $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}: \mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$   
**positiv definit**        wenn  $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}: \mathbf{x}^T \mathbf{A} \mathbf{x} > 0$   
 ( $\mathbf{x}^T \mathbf{A} \mathbf{x}$  ist **quadratische Form**)

Beispiel:

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &= (x_1, x_2) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1 - x_2, -x_1 + x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1 - x_2)x_1 + (-x_1 + x_2)x_2 \\ &= (x_1 - x_2)^2 \quad \mathbf{x} \neq \mathbf{0}: > 0 \end{aligned}$$

Es wird im folgenden  
 auf positive Definitheit / Semidefinitheit von  $\mathbf{H}$  ankommen

Daher Hinweise auf Feststellung dieser Eigenschaft(en)

$n \times n$  Matrix  $\mathbf{A}$  lässt sich (zB mithilfe Gauß-Algorithmus)  
als Produkt  $\mathbf{A} = \mathbf{L} \mathbf{R}$  von Dreiecksmatrizen  $\mathbf{L}$ ,  $\mathbf{R}$  darstellen  
mit unterer / oberer Dreiecksform

$$\mathbf{L} = \begin{pmatrix} 1 & & & 0 \\ l_{21} & 1 & & \\ \vdots & & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \quad \mathbf{R} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{pmatrix}$$

Für symmetrisches  $\mathbf{A}$  ferner  $\mathbf{R}$  als Produkt  $\mathbf{R} = \mathbf{D} \mathbf{L}^T$   
mit Diagonalmatrix

$$\mathbf{D} = \text{diag}(d_{11}, \dots, d_{nn})$$

Aus  $\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{L}^T$   
folgt  $\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{L} \mathbf{D} \mathbf{L}^T \mathbf{x}$   
bzw mit  $\mathbf{y} := \mathbf{L}^T \mathbf{x}$   
auch  $\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{D} \mathbf{y} = \sum d_{ii} y_i^2$

$\mathbf{A}$  genau dann positiv definit, wenn alle  $d_{ii} > 0$   
positiv semidefinit, wenn alle  $d_{ii} \geq 0$

Zielführend folgender Satz

### Satz 7.1.03 : Hinreichende Bedingungen

Sei  $Z(\mathbf{x})$  zweimal stetig differenzierbar  
 $\mathbf{x}$  innerer Punkt von  $M$   
 $\mathbf{g}(\mathbf{x}) = \mathbf{0}$   
 $\mathbf{H}(\mathbf{x})$  positiv definit

$\mathbf{x}$  ist lokaler Minimalpunkt von  $Z$  auf  $M$

oB

Hinweis: positiv semidefinit ist nur notwendige Bedingung  
(Analogon im Eindimensionalen: "Wendepunkt")

Nach Satz 7.1.02 ist

$$\mathbf{s}^T \mathbf{g}(\mathbf{x}) = 0 \quad \mathbf{s} \in S(\mathbf{x})$$

notwendige Optimalitätsbedingung

- reduziert sich für unrestringierte Probleme auf  $\mathbf{g}(\mathbf{x}) = 0$
- ist für restringierte Probleme ( wo Optimalpunkte auf Rand von M liegen können / werden ) praktisch schlecht handhabbar ( wegen Schwierigkeit der expliziten Notierung von  $S(\mathbf{x})$  )

Methode der Lagrange'schen Multiplikatoren

- Integration Nebenbedingungen in Zielfunktion
- Optimierung unrestringierten Problems bietet Abhilfe,

ist

- von Nebenbedingungs-Gleichungen
- auf Nebenbedingungs-Ungleichungen zu übertragen

Für Optimierungsproblem

$$(O) \quad \min_{\mathbf{x} \in N} Z(\mathbf{x}) \quad z_i(\mathbf{x}) = 0 \quad i=1, \dots, m$$

werden **Lagrange-Multiplikatoren**  $u_i$  ( $i=1, \dots, m$ ) eingeführt und die **Lagrange-Funktion**

$L: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$  definiert gemäß

$$L(\mathbf{x}, \mathbf{u}) := Z(\mathbf{x}) + \sum_{i=1}^m u_i z_i(\mathbf{x}) = Z(\mathbf{x}) + \mathbf{u}^T \mathbf{z}(\mathbf{x})$$

Punkt  $(\mathbf{x}^* \in \mathbb{R}^n, \mathbf{u}^* \in \mathbb{R}_+^m)$  heißt **Sattelpunkt** von  $L$ , wenn  
 (vgl. Spieltheorie)  

$$L(\mathbf{x}^*, \mathbf{u}) \leq L(\mathbf{x}^*, \mathbf{u}^*) \leq L(\mathbf{x}, \mathbf{u}^*) \quad \mathbf{x} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}_+^m$$

**Satz 7.1.04 : Sattelpunkt und Optimalität**

$(\mathbf{x}^* \in \mathbb{R}^n, \mathbf{u}^* \in \mathbb{R}_+^m)$  **Sattelpunkt** von  $L$   
 $\mathbf{x}^*$  ist optimale Lösung von (O)

- $Z(\mathbf{x}^*) + \mathbf{u}^T \mathbf{z}(\mathbf{x}^*) = L(\mathbf{x}^*, \mathbf{u})$   
 (a)  $\mathbf{u}^T \mathbf{z}(\mathbf{x}^*) \leq \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*) = L(\mathbf{x}^*, \mathbf{u}^*) = Z(\mathbf{x}^*) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*)$   
 $\mathbf{u} \in \mathbb{R}_+^m$
- (a) ist für  $\mathbf{z}(\mathbf{x}^*) > \mathbf{0}$ ,  $\mathbf{u} > \mathbf{u}^*$  nicht erfüllbar  
 $\mathbf{z}(\mathbf{x}^*) \geq \mathbf{0}$   $\mathbf{x}^*$  zulässig für (O)
- (a) liefert für  $\mathbf{u} = \mathbf{0}$   
 $\mathbf{0} \leq \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*)$   
 zusammen mit  
 $Z(\mathbf{x}^*) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*) = L(\mathbf{x}^*, \mathbf{u}^*)$   
 $Z(\mathbf{x}^*) \leq Z(\mathbf{x}) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}) = L(\mathbf{x}, \mathbf{u}^*) \leq L(\mathbf{x}^*, \mathbf{u}^*) = Z(\mathbf{x}^*) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*)$   
 $\mathbf{x} \in \mathbb{R}^n$
- $\mathbf{u} \geq \mathbf{0}$ ,  $\mathbf{z}(\mathbf{x}) \geq \mathbf{0}$  im zulässigen Bereich  
 $Z(\mathbf{x}^*) \leq Z(\mathbf{x})$  zulässigen  $\mathbf{x}$   
 $\mathbf{x}^*$  ist optimal

"Schön" wäre jetzt,  
 wenn auch optimaler Punkt Sattelpunkt gälte

Ist problem(-klassen)-abhängig zu untersuchen

## 7.2 Konvexe Optimierungsprobleme

Sätze 7.1.02/.03 formulieren Bedingungen für lokale Minimalpunkte fest

bei linearen Optimierungsproblemen waren

- lokale Minimalpunkte
- auch globale Minimalpunkte

Kann nicht allgemein gelten

Aber über lineare Zielfunktionen hinaus?

Erinnerung:

- **Menge**  $K \subset \mathbb{R}^n$  heißt **konvex** wenn

$$\mathbf{x}_1, \mathbf{x}_2 \in K; 0 \leq r \leq 1 \quad r \mathbf{x}_1 + (1-r)\mathbf{x}_2 \in K$$

Durchschnitt konvexer Mengen ist konvex

- Sei  $K \subset \mathbb{R}^n$  konvex. **Funktion**  $F: K \rightarrow \mathbb{R}$  heißt **konvex** wenn

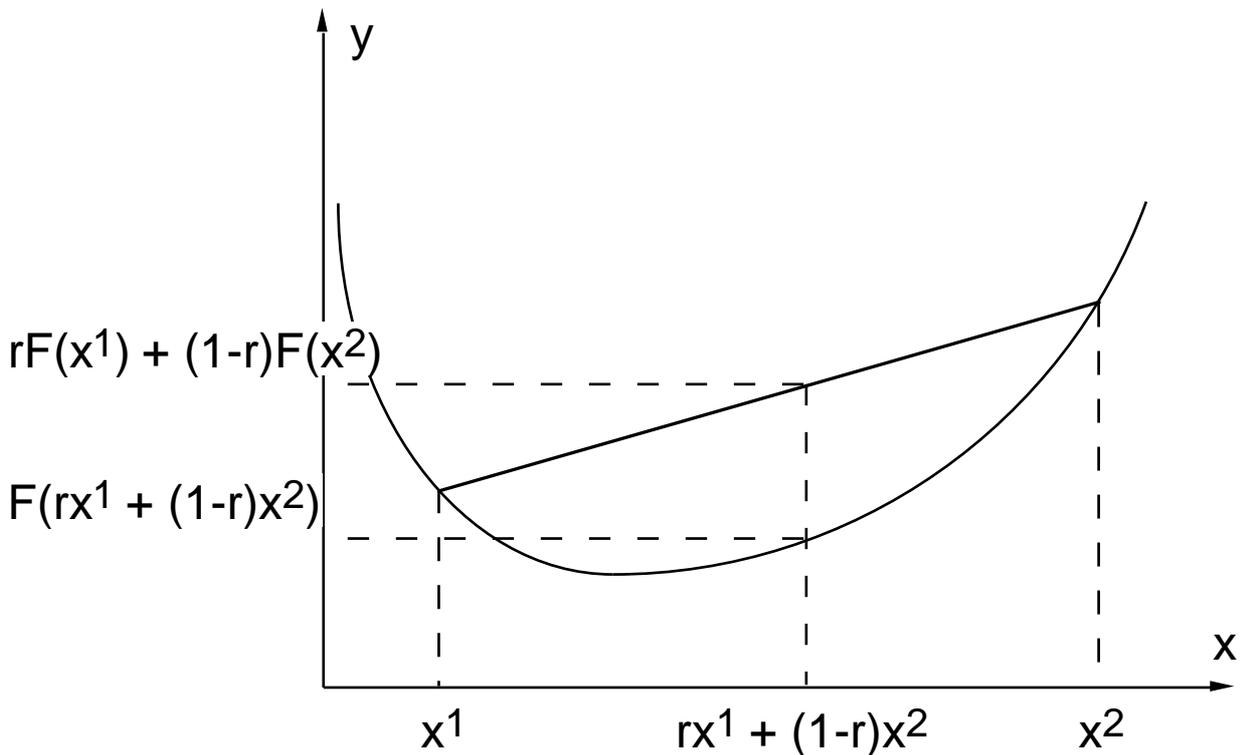
$$\mathbf{x}_1, \mathbf{x}_2 \in K; 0 < r < 1 \quad F(r \mathbf{x}_1 + (1-r)\mathbf{x}_2) \leq r F(\mathbf{x}_1) + (1-r) F(\mathbf{x}_2)$$

streng konvex für  $<$

konkav für

streng konkav für  $>$

- lineare Funktionen sind sowohl konvex als auch konkav weder streng konvex noch streng konkav
- für  $K=\mathbb{R}$ : streng konvexe Funktionen "hängen nach unten durch"



Direkt aus Definitionen:

### Satz 7.2.01 : Eigenschaften konvexer Funktionen

Seien  $K \subseteq \mathbb{R}^n$  konvexe Menge

$F_1, \dots, F_m : K \rightarrow \mathbb{R}$  konvexe Funktionen

$a \in \mathbb{R}$

dann

(a) nichtnegative Linearkombination der  $F_i$

$$G := \sum_{i=1}^m r_i F_i \quad r_i \geq 0$$

ist konvexe Funktion auf  $K$

(b) Mengen

$\{\mathbf{x} \in K \mid F(\mathbf{x}) \leq a\}$  (abgeschlossen)

$\{\mathbf{x} \in K \mid F(\mathbf{x}) < a\}$  (offen)

sind konvexe Mengen

Es gibt diverse nützliche Kriterien zur Feststellung der Konvexität einer Funktion, so

**Satz 7.2.02 : Konvexitätskriterium I**

Sei  $K \subset \mathbb{R}^n$  konvexe Menge mit inneren Punkten

$F: K \rightarrow \mathbb{R}$  zweimal stetig differenzierbar

$F$  ist genau dann konvex (streng konvex) auf  $K$  wenn Hesse-Matrix  $\mathbf{H}(\mathbf{x})$  positiv semidefinit (positiv definit)  $\circ B$

**Satz 7.2.03 : Konvexitätskriterium II**

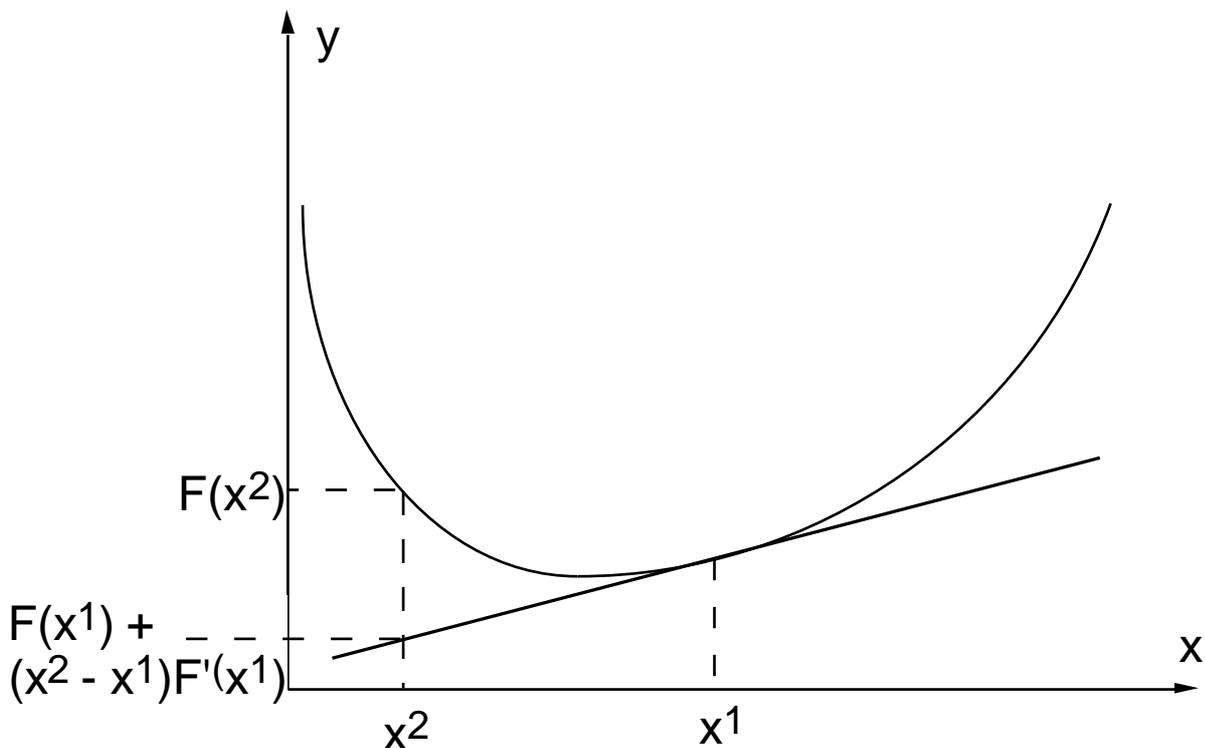
Sei  $K \subset \mathbb{R}^n$  konvexe Menge

$F: K \rightarrow \mathbb{R}$  stetig differenzierbar

$F$  ist genau dann konvex (streng konvex) auf  $K$  wenn

$$\mathbf{x}_1, \mathbf{x}_2 \in K: F(\mathbf{x}_2) \geq F(\mathbf{x}_1) + (\mathbf{x}_2 - \mathbf{x}_1)^T \text{grad } F(\mathbf{x}_1) \quad (>) \quad \circ B$$

Tangente an  $F(\mathbf{x})$  "niemals oberhalb Kurve":



Mit konvexer Zielfunktion  $Z(\mathbf{x})$ ,  
 konvexen Nebenbedingungen  $z_i(\mathbf{x}) \leq 0 \quad i=1, \dots, m$   
 und damit konvexer Lösungsmenge  
 (Durchschnitt von Mengen gemäß S.7.2.01b)  
 liegt ein sog. **konvexes Optimierungsproblem** (K) vor

$$\begin{aligned}
 \text{(K)} \quad & \min \quad Z(\mathbf{x}) \\
 & \text{udN} \quad z_i(\mathbf{x}) \leq 0 \quad i=1, \dots, m \\
 & \quad Z, z_i \ (i=1, \dots, m) \text{ konvex}
 \end{aligned}$$

für konvexe Optimierungsprobleme gilt (erwünschterweise:)

### Satz 7.2.04 : Hauptsatz der konvexen Optimierung

Sei  $M \subseteq \mathbb{R}^n$  konvexe (Lösungs-)Menge,  $M$   
 $Z: M \rightarrow \mathbb{R}$  konvexe (Ziel-)Funktion

dann

- (a) Menge  $M^*$  aller globalen Minimalpunkte  
 von  $Z$  auf  $M$  ist konvex
- (b) jeder lokale Minimalpunkt von  $Z$  auf  $M$   
 ist auch globaler Minimalpunkt

zu (a):  $\mathbf{x}^*$  globaler Minimalpunkt mit Minimum  $Z^* := Z(\mathbf{x}^*)$   
 S.7.2.01b:  $\{\mathbf{x} \in M \mid Z(\mathbf{x}) = Z^*\}$  ist konvex

zu (b):  $\mathbf{x}'$  lokaler Minimalpunkt mit Minimum  $Z' := Z(\mathbf{x}')$   
 Annahme:  $\exists \mathbf{x} : Z(\mathbf{x}) < Z'$  ( $\mathbf{x}'$  nicht global. Min.Pkt)  
 $Z$  konvex: für  $0 < r < 1$  gilt  

$$\begin{aligned}
 Z(r\mathbf{x} + (1-r)\mathbf{x}') & \leq rZ(\mathbf{x}) + (1-r)Z(\mathbf{x}') \\
 Z(\mathbf{x}' + r(\mathbf{x} - \mathbf{x}')) & < Z'
 \end{aligned}$$
 in  $\epsilon$ -Umgebung von  $\mathbf{x}'$  finden sich  $\mathbf{x}'' : Z(\mathbf{x}'') < Z'$   
 Widerspruch

Insbesondere werden

- die notwendigen Bedingungen für lokalen Minimalpunkt  
**(B) Satz 7.1.02:**  $\mathbf{s}^T \mathbf{g}(\mathbf{x}) = 0$  für  $\mathbf{s} \in S(\mathbf{x})$   
 $\mathbf{g}(\mathbf{x}) = \mathbf{0}$  für innere Punkte v. M
- für konvexe Optimierungsprobleme zu hinreichenden Bedingungen für globalen Minimalpunkt

**Satz 7.2.05 : Hinreichende Optimalitätsbedingung**

Sei  $M \subset \mathbb{R}^n$  konvexe (Lösungs-)Menge  
 $Z: M \rightarrow \mathbb{R}$  konvexe, stetig differenzierb. (Ziel-)Funktion  
 $\mathbf{x}^* \in M$

(B) gilt für  $\mathbf{x}^*$   $\mathbf{x}^*$  ist globaler Minimalpunkt von Z auf M

$\mathbf{x} \in M, \mathbf{x} \in M$

M konvex: Verbindungsstrecke  $\mathbf{x}^* \rightarrow \mathbf{x}$  in M  
 $\mathbf{s} = \mathbf{x} - \mathbf{x}^* \in S(\mathbf{x}^*)$  (ist zulässige Richtung)

$\mathbf{s} \in S(\mathbf{x}^*)$ :  $\mathbf{s}^T \mathbf{g}(\mathbf{x}^*) > 0$ :  $\mathbf{s} = \mathbf{x} - \mathbf{x}^*$

(B) gilt:  $(\mathbf{x} - \mathbf{x}^*)^T \mathbf{g}(\mathbf{x}^*) \geq 0$  für  $\mathbf{x} \in M$

S. 7.2.03:  $F(\mathbf{x}) \geq F(\mathbf{x}^*) + (\mathbf{x} - \mathbf{x}^*)^T \mathbf{g}(\mathbf{x}^*)$  für  $\mathbf{x} \in M$   
 Behauptung

Auf Umkehrung von S. 7.1.04 (Sattelpunkt Minimalpunkt) für konvexe Probleme zielend, Formulierung der sog. **Slater-Bedingung:**

$\mathbf{x} \in \mathbb{R}^n$ : nichtlin. Nebenbedingungen:  $z_i(\mathbf{x}) < 0$

( bei ausschließlich nichtlin. Nebenbedingungen impliziert Slater-Bedingung, daß M innere Punkte enthält )

### Satz 7.2.06 : Satz von Kuhn / Tucker

Sei im Rahmen eines konvexen Optimierungsproblems gemäß (K) die Slater-Bedingung erfüllt.

Dann gilt:  $\mathbf{x}^*$  ist optimale Lösung von (K)

Lagrange-Funktion  $L(\mathbf{x}, \mathbf{u})$  besitzt

Sattelpunkt  $(\mathbf{x}^*, \mathbf{u}^*)$  mit  $\mathbf{u}^* \geq \mathbf{0}$

oB

### Globale Sattelpunkteigenschaft

ist schlecht explizit nachprüfbar

Für (wie hier gegeben)

differenzierbare, konvexe  $Z, z_i$

durch äquivalente **lokale** Bedingungen ersetzbar:

### Karush-Kuhn-Tucker-Bedingungen

für stetig differenzierbare, konvexe  $Z, z_i$  ( $i=1, \dots, m$ )

(aus:  $L/ \mathbf{x}=\mathbf{0}, L/ \mathbf{u}=\mathbf{0}$   
unter Nutzung  $\mathbf{z}(\mathbf{x}) \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0}$  )

$$(7.2.07) \quad \text{grad } Z(\mathbf{x}^*) + \sum_{i=1}^m u^*_i \text{ grad } z_i(\mathbf{x}^*) = \mathbf{0}$$

$$u^*_i z_i(\mathbf{x}^*) = 0 \quad i = 1, \dots, m$$

$$z_i(\mathbf{x}^*) \geq 0 \quad i = 1, \dots, m$$

$$\mathbf{u}^* \geq \mathbf{0}$$

### KKT-Bedingungen

- lassen sich für spezifische (Unter-)Klassen weiter konkretisieren
- so zB (breit untersucht) für **quadratische Opt.Probleme**  
"=" quadratische Zielfunktion  
+ lineare Nebenbedingungen (M ist konv. Polyeder)

## 7.3 Übersicht über Lösungsverfahren

Skizzen zu Fällen

- EINDIMENSIONALE OPTIMIERUNG
- UNRESTRINGIERTE OPTIMIERUNGSPROBLEME IM  $\mathbf{R}^n$
- RESTRINGIERTE OPTIMIERUNGSPROBLEME IM  $\mathbf{R}^n$

### EINDIMENSIONALE OPTIMIERUNG

$Z: \mathbf{R} \rightarrow \mathbf{R}$  Zielfunktion

$\min_{x \in \mathbf{R}} Z(x)$       bzw  $x \in \mathbf{R}_+$       bzw  $x \in [a,b], b>a$   
 udN

- einfachste Optimierungsaufgabe mit eigenem Interesse
- Teilproblem in höherdimensionalen Optimierungsaufgaben

Problem besitzt genau eine optimale Lösung,  
wenn  $Z$  im zulässigen Bereich  $[a,b]$

**unimodal**  $x^* \in [a,b]$ :  $Z$  streng monoton fallend auf  $[a,x^*]$ ,  
 $Z$  streng monoton steigend auf  $[x^*,b]$

streng konvex      unimodal

Einfache Lösungsverfahren ("ableitungsfrei")

- nutzen nur Funktionswerte
- verkleinern "Suchintervall" schrittweise

Typisch: 2 "Stützstellen"  $x_1, x_2$  mit  $a < x_1 < x_2 < b$ ,

$Z(x_1) < Z(x_2)$ : weiter mit  $[a, x_2]$

$Z(x_1) > Z(x_2)$ : weiter mit  $[x_1, b]$

Teilung zB "goldener Schnitt"

Aufwendigere (idR bessere) Lösungsverfahren

- nutzen Funktionswerte + Werte 1. (+höherer) Ableitungen
- konstruieren Folge (hoffentlich:) konvergierender Punkte

Typisch: **Newton-Verfahren**

- Annahme:  $Z$  zweimal stetig differenzierbar

Minimalpunkt  $x^*$  (a,b)

Näherungspunkt  $x^i$  (a,b);  $i=0,1,\dots$

(Vorbereitung durch "einfaches" Verfahren,  
Grenzpunkte a,b immer separat zu prüfen)

- Ziel: Bestimmung  $x^*: Z'(x^*)=0$  Nullstelle  $Z'$   
( $Z$  konvex: notwendig und hinreichend  
für globale Minimalstelle)

- Vorgang: Berechnung  $Z'(x^i)$ ,  $Z''(x^i)$   
Linearisierung der Ableitungsfunktion  $Z'$  in  $x^i$   
Bestimmung der Tangentennullstelle  $=: x^{i+1}$

$$x^{i+1} := x^i - \frac{Z'(x^i)}{Z''(x^i)}$$

- Newton verwendet quadratische (parabolische)  
Approximation von  $Z$   
bei quadratischem  $Z$  Optimalstelle in 1 Schritt gefunden

Konvergenz bei beliebigem Startpunkt  $x^0$  nicht gesichert,  
analytische Bedingungen für sicheres  $x^0$  kompliziert

- sind  $Z'$ ,  $Z''$  nicht analytisch bekannt,  
kann mit numerischen Näherungen  
auf Basis der Funktionswerte "benachbarter" Punkte  
gearbeitet werden



## Schritte von Abstiegsverfahren

- Initialisierung  
wähle  $\mathbf{x}^0 \in \mathbf{R}^n$ ,  $i:=0$
- Iteration über
  - Abbruchprüfung  
berechne  $\mathbf{g}^i := \mathbf{g}(\mathbf{x}^i)$   
terminiere falls  $\mathbf{g}^i = \mathbf{0}$  (bzw  $|\mathbf{g}^i| < \epsilon$ )
  - Abstiegsrichtung, Suchrichtung  
bestimme  $\mathbf{s}^i \in \mathbf{R}^n : (\mathbf{s}^i)^T \mathbf{g}^i < 0$
  - Schrittweite  
berechne  $d^i > 0 : Z(\mathbf{x}^i + d^i \mathbf{s}^i) < Z(\mathbf{x}^i)$   
optimale Schrittweite ist eindimensionales Subproblem
  - Fortschritt  
 $\mathbf{x}^{i+1} := \mathbf{x}^i + d^i \mathbf{s}^i$ ;  $i := i+1$

Konkrete Verfahren unterscheiden sich durch Wahl von

- Abstiegsrichtung
- Schrittweite

- klassisches **Gradientenverfahren**,  
**Verfahren des steilsten Abstiegs**

$$\mathbf{s}^i := -\mathbf{g}^i$$

- idR deutlich schneller konvergierend ist mehrdimensionales **Newtonverfahren**
  - zusätzlich benötigt: zweite Ableitungen  
Forderung  $Z$  (mindest.) 2-mal stetig diff'bar  
Notation  $\mathbf{H}^i := \mathbf{H}(\mathbf{x}^i)$  Hesse-Matrix

$\mathbf{H}^i$  positiv definit, Satz 7.2.02

Z ist in Umgebung von  $\mathbf{x}^i$  streng konvex

- Standard-Newton setzt

$$\mathbf{s}^i := - (\mathbf{H}^i)^{-1} \mathbf{g}^i$$

$$d^i := 1 \quad (\text{ia nicht optimal, optimieren!})$$

so daß

$$\mathbf{x}^{i+1} := \mathbf{x}^i - (\mathbf{H}^i)^{-1} \mathbf{g}^i$$

- $\mathbf{H}^i$  muß nicht explizit invertiert werden, da wegen

$$\mathbf{H}^i \mathbf{x}^{i+1} := \mathbf{H}^i \mathbf{x}^i - \mathbf{g}^i$$

$\mathbf{x}^{i+1}$  als Lösung linearen Gl'systems ermittelbar

- Nachteil ist hoher Rechenaufwand:  
Hesse-Matrix + lin. Gl'system  
abgemildert in vereinfachtem Newton-Verfahren,  
welches  $\mathbf{H}^i$  für "einige" Schritte  $i, i+1, \dots$  verwendet

- "zwischen"

- Gradientenverfahren (schlechte Konvergenz)

- Newtonverfahren (hoher Aufwand)

sind **Verfahren der konjugierten Gradienten**

- mit besserer Konvergenz als Gradientenverfahren

- mit geringerem Aufwand als Newtonverfahren

...

- ...

## RESTRINGIERTE OPTIMIERUNGSPROBLEME IM $\mathbb{R}^n$

$\min_{\mathbf{x} \in M \subset \mathbb{R}^n} Z(\mathbf{x})$       idR stetig differenzierbar  
 udN

$$LM := \{ \mathbf{x}^* \in M \mid \mathbf{s}^T \mathbf{g}(\mathbf{x}^*) = 0 \quad \mathbf{s} \in S(\mathbf{x}^*) \}$$

(Hoffnung auf) Näherung für stationäre Punkte  $\mathbf{x}^* \in LM$   
 $Z, M$  konvex (Satz 7.2.05)       $\mathbf{x}^*$  sind glob. Minimalpunkte

### Verfahrensklassen:

- Verfahren zulässiger Richtungen
- Verfahren der Straffunktionen (Barrierefunktionen)
- Schnittebenenverfahren
- Karush-Kuhn-Tucker-Verfahren
- ...

### • Verfahren zulässiger Richtungen

Anwendung iW für lineare Restriktionen

Vorgang in Abwandlung von nichtrestringiertem Vorgang

- Initialisierung: Wahl  $\mathbf{x}^0 \in M$ ,  $i:=0$
- Iteration über
  - \* Abstiegsrichtung: Bestimmung  $\mathbf{s}^i \in \mathbb{R}^n : (\mathbf{s}^i)^T \mathbf{g}^i < 0$   
(ohne Erfolg: Abbruch)
  - \* Schrittweite: Berechnung  $d^i > 0 : Z(\mathbf{x}^i + d^i \mathbf{s}^i) < Z(\mathbf{x}^i)$   
mit optimaler Schrittweite,  
aber **unter Beachtung Restriktionen**
  - \* Fortschritt:  $\mathbf{x}^{i+1} := \mathbf{x}^i + d^i \mathbf{s}^i$ ;  $i := i+1$

Bei linearen Restriktionen Nebenbeding'gen gewohnter Form:

$$\begin{aligned} \min \quad & Z(\mathbf{x}) \\ \text{udN} \quad & \mathbf{A} \mathbf{x} \leq \mathbf{b} \quad \mathbf{b} \in \mathbb{R}^m, \mathbf{A}: m \times n \end{aligned}$$

Iterationspunkt  $\mathbf{x}^i \in M$  (immer gewahrt) liegt

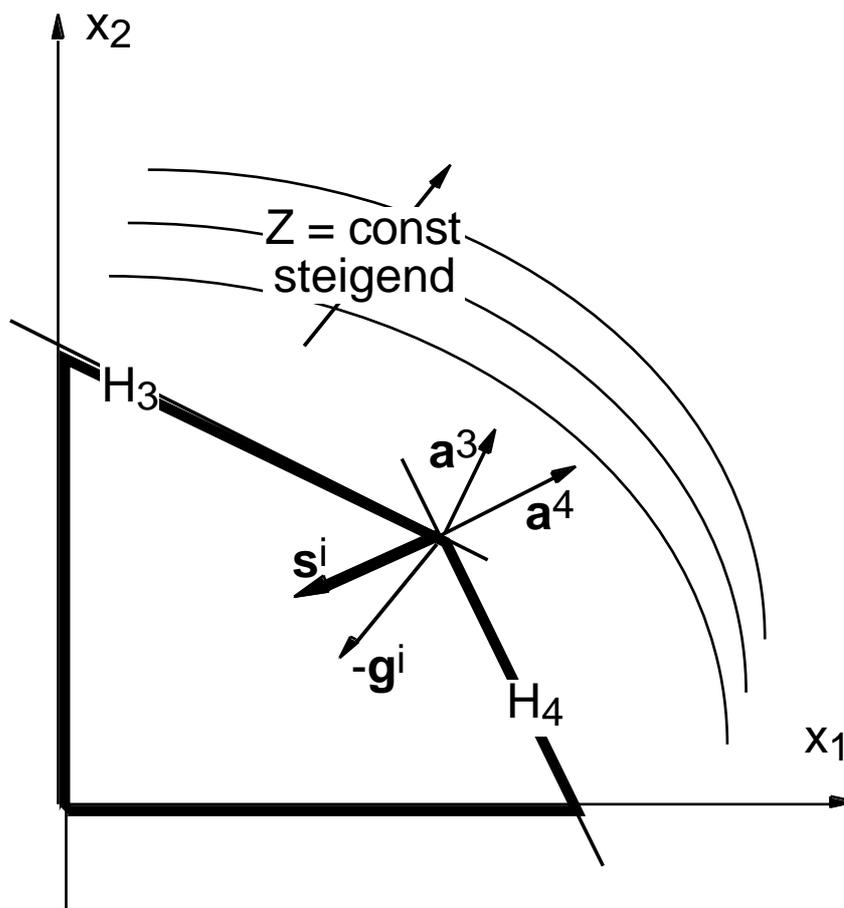
- auf  $(0,1,\dots,m)$  begrenzenden Hyperebenen

$$H_k: \quad (\mathbf{a}^k)^T \mathbf{x} = b^k \quad k \in I^i$$

wo  $\mathbf{a}^k$  k-ter Zeilenvektor  $\mathbf{A}$  (Vektor  $\mathbf{a}^k$  orth.  $H_k$ ,

dh "nach außen")

- auf keiner Grenzebene (im Inneren von  $M$ ) für  $I^i =$
- oder auf Menge **in  $\mathbf{x}^i$  aktiver Hyperebenen**



Bedingungen Fortschrittsrichtung  $\mathbf{s}^i$ :

- im stumpfen Winkel zu  $\mathbf{a}^k, k \in I^i$  "zulässig"
- im spitzen Winkel zu  $-\mathbf{g}^i$  "Z besser (kleiner)"

Bedingungen formal:

- stumpfe, äußerstenfalls rechte Winkel  
 $(\mathbf{a}^k)^T \mathbf{s}^i \geq 0 \quad k \in I^i$
- spitzer Winkel  $-\mathbf{g}^i$                       stumpfer Winkel  $\mathbf{g}^i$   
 $(-\mathbf{g}^i)^T \mathbf{s}^i > 0$                        $(\mathbf{g}^i)^T \mathbf{s}^i < 0$

Verschiedene Regeln zur Wahl von  $\mathbf{s}^i$

- Rosen-Verfahren
  - \* berechnet  $\mathbf{s}^i$  als Projektion  $\mathbf{g}^i$  auf  $(H_k; k \in I^i)$  für  $I^i$
  - nutzt  $-\mathbf{g}^i$  als  $\mathbf{s}^i$  für  $I^i =$
  - \* in beiden Fällen ist  $|\mathbf{s}^i| < \epsilon$  Abbruchbedingung  
 (stationärer Punkt erreicht)
- vgl auch Karmarkars Projektionsmethode?  
 blieb "im Inneren" !

Zusätzlich zu beachten:

- Schrittweite " $d^i$ " so daß M nicht verlassen  
 (wegen linearer NB leicht formalisierbar)  
 $u^i$  zu optimieren  
 (eindimensionale Optimierung)
- Begrenzung Schrittweite führt implizit  
 auch zur Entdeckung "unbegrenzten Problems"

- **Verfahren der Straffunktionen (Barrierefunktionen)**

Restringiertes Optimierungsproblem

- durch unrestringiertes Optimierungsprobleme approximiert
- unter Veränderung der Zielfunktion derart daß
  - \* Verlassen des zulässigen Bereichs mit "Strafkosten" belastet (Veränderung  $Z$  außerhalb  $M$ )
  - \* Rand des zulässigen Bereichs in "Barriere" verwandelt (Veränderung  $Z$  bei Annäherung an Rand  $M$ )

Skizze **Straffunktionen**:

$$\begin{aligned} \min Z(\mathbf{x}) \quad \text{udN } \mathbf{z}(\mathbf{x}) &= \mathbf{0} \\ \min Y(\mathbf{x}) \quad \text{udN } \mathbf{x} &\in \mathbf{R}^n \end{aligned}$$

$$\text{wo } Y(\mathbf{x}) := Z(\mathbf{x}) + p(\mathbf{x}) \quad \text{mit } \begin{aligned} p(\mathbf{x}) &= 0 & \mathbf{x} &\in M \\ &> 0 & \mathbf{x} &\notin M \end{aligned}$$

Minimalpunkt  $\mathbf{x} \in M$  für  $Y$  ist Minimalpunkt für  $Z$

verschiedene Straffunktionen, oft verwendet:

$$\begin{aligned} p(\mathbf{x}) &:= \mathbf{z}^+(\mathbf{x})^T \mathbf{z}^+(\mathbf{x}) \\ \text{wo } \mathbf{z}^+(\mathbf{x}) &:= \max(\mathbf{0}, \mathbf{z}(\mathbf{x})) \quad (\text{komponentenweise}) \end{aligned}$$

Folgt Behandlung à la unrestringiertes Problem, Idee idR verfeinert in Form

$$Y(\mathbf{x}) := Z(\mathbf{x}) + r^i p(\mathbf{x})$$

- dh mit
- gewichteter Strafe
  - + im Iterationsverlauf angepaßter Gewichtung (idR  $r^i$  wachsend mit Iterationsindex)

## Skizze **Barrierefunktionen**:

Voraussetzung:  $M$  besitzt innere Punkte  $M^\circ \subset M$

Barrierefunktionen  $b: M^\circ \rightarrow \mathbf{R}$  besitzen die Eigenschaften

- $b$  stetig auf  $M^\circ$
- $b(\mathbf{x}) \rightarrow \infty$  bei Annäherung von  $\mathbf{x}$  an den Rand von  $M$

verschiedene Barrierefunktionen,

im Falle von Nebenbedingungen

$\mathbf{z}(\mathbf{x}) < \mathbf{0}$  (Rand **nicht** erlaubt)  
oft verwendet sind

$$b_1(\mathbf{x}) := -\ln(-z_i(\mathbf{x})) \qquad b_2(\mathbf{x}) := -\frac{1}{z_i(\mathbf{x})}$$

Folgt Behandlung à la unrestringiertes Problem,  
unter Minimierung der Ersatzfunktion

$$Y(\mathbf{x}) := Z(\mathbf{x}) + 1/r^i b(\mathbf{x})$$

mit  $r^i > 0$ , Vergrößerung der  $r^i$  im Iterationsverlauf

### • **Schnittebenenverfahren**

Vorrangig bei nichtlinearen Restriktionen eingesetzt

Zunächst lineare Zielfunktion:

$$\begin{array}{ll} \min & \mathbf{c}^T \mathbf{x} \\ \text{u.d.N} & \mathbf{x} \in M \subset \mathbf{R}^n \end{array} \quad \begin{array}{l} M \text{ nichtleer, kompakt, konvex,} \\ \text{kein Polytop (sonst: bekannt)} \end{array}$$

"Um M" werde konvexes Poytop  $P_1$  M gelegt

Optimierung auf  $P_1$  (zB mit Simplex)  
Lösung  $x^1$

- $x^1 \in M$   $x^1$  ist optimale Lösung
- $x^1 \in M$  Bestimmung Hyperebene, welche  $P_1$  beschneidet  
konvexes Poytop  $P_2$   
wo  $P_2 \subset M, x^1 \in P_2$

usf

Verschiedene Verfahren zur Konstruktion der Schnittebenen im Gebrauch

vgl auch Gomory-Verfahren

(bis hier angenommene)

Linearität der Zielfunktion ist nicht wesentlich:

$$\begin{array}{ll} \min & Z(\mathbf{x}) \\ \text{udN} & \mathbf{x} \in M \subset \mathbf{R}^n \end{array} \quad \begin{array}{l} Z \text{ nichtlinear, konvex} \\ M \text{ nichtleer, kompakt, konvex} \end{array}$$

$$\begin{array}{ll} \text{Einführung} & \text{zusätzlicher Variabler} \\ & + \text{zusätzlicher Nebenbedingung} \end{array} \quad \begin{array}{l} x_{n+1} \in \mathbf{R} \\ Z(\mathbf{x}) - x_{n+1} = 0 \end{array}$$

Betrachtung Optimierungsproblem

$$\begin{array}{ll} \min & x_{n+1} \\ \text{udN} & \mathbf{x} \in M \subset \mathbf{R}^n \end{array} \quad \begin{array}{l} \text{lineare Zielfunktion} \\ Z(\mathbf{x}) - x_{n+1} = 0 \end{array}$$

$x^*$  genau dann optimale Lösung Ursprungsproblem,  
wenn  $(x^*, x_{n+1}^*)$  Lösung Ersatzproblem, wo  $x_{n+1}^* = Z(x^*)$

- **Karush-Kuhn-Tucker-Verfahren**

Analytische Verfolgung der KKT-Bedingungen  
(hinreichend für Optimalität)

für konkrete Problemklassen

(zB quadratische Zielfunktion,

in Abschn.7.2 erwähnt, nicht explizit ausgeführt)

führt ggf zu einfacher behandelbaren Ersatzproblemen

(zB mit Methoden der linearen Optimierung)

hier "Abbruch",

viel weiteres Wissen vorhanden!

- ( Lagrange-Dualität,  
Separable Optimierung,  
Quotientenoptimierung,  
konkave Optimierung,  
B&B-Verfahren,  
Deflations- und Tunnel-Techniken,  
nichtdeterministische Verfahren,  
evolutionäre Algorithmen  
... )

LEER

LEER