Algorithmische Mathematik

Kapitel 4

Nichtlineare Optimierung

Beispiel zur Motivation: Angebotsauswertung

Eine Fabrik benötigt eine Menge A eines bestimmten Rohstoffs, der von n verschieden Zulieferern angeboten wird.

M_i : maximale Liefermenge von Anbieter i
$f_i(x_i)$: Stückpreis bei Abnahme der Menge x_i

zugehörige Optimierungsaufgabe:

$$\begin{array}{rl} \text{min} & \displaystyle\sum_{i=1}^{n} x_i f_i(x_i) \\ \text{unter} & \displaystyle\sum_{i=1}^{n} x_i = A \\ & 0 \leq x \leq M \end{array}$$

Mengenrabatt $\Rightarrow f_i$ nicht konstant

$\Rightarrow$ ZF nicht linear !

z.B. $f_i(x_i) = p_i - a_i x_i$

$a_i = \frac{p_i - p_i^{\text{min}}}{M_i}$ Abschlagskonstante

ZF: $\sum_{i=1}^{n} p_i x_i - a_i x_i^2$

- Nichtlineare Optimierungsaufgabe:

$$\boxed{\min_{x \in S} f(x) \text{ mit } S \subset \mathbb{R}^n, f : \mathbb{R}^n \to \mathbb{R}.}$$

- $S = \mathbb{R}^n$:
  globales Minimum, "Kurvendiskussion"

- $S = \{x \in \mathbb{R}^n | h(x) = 0\}$ mit $h : \mathbb{R}^n \to \mathbb{R}^m$:
  Extrema unter Nebenbedingungen,
  "Lagrange-Multiplikatoren"

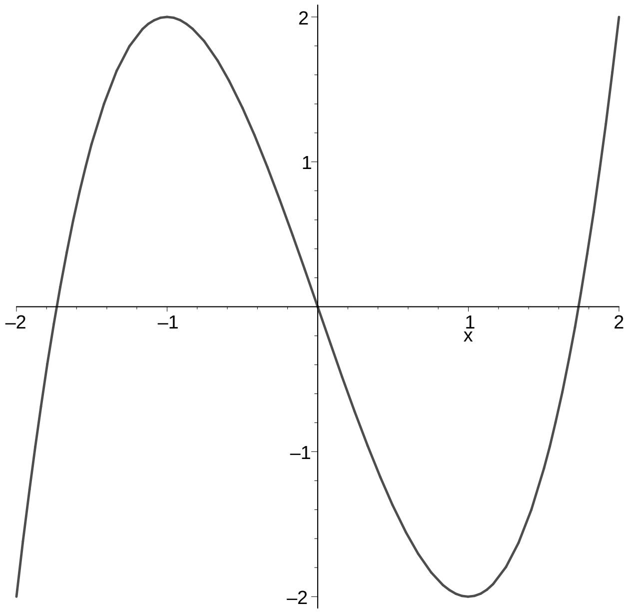
- $S = \{x \in \mathbb{R}^n | h(x) = 0, g(x) \geq 0\}$:
  "Kuhn-Tucker-Bedingungen"

- Lineares Programm als Spezialfall:

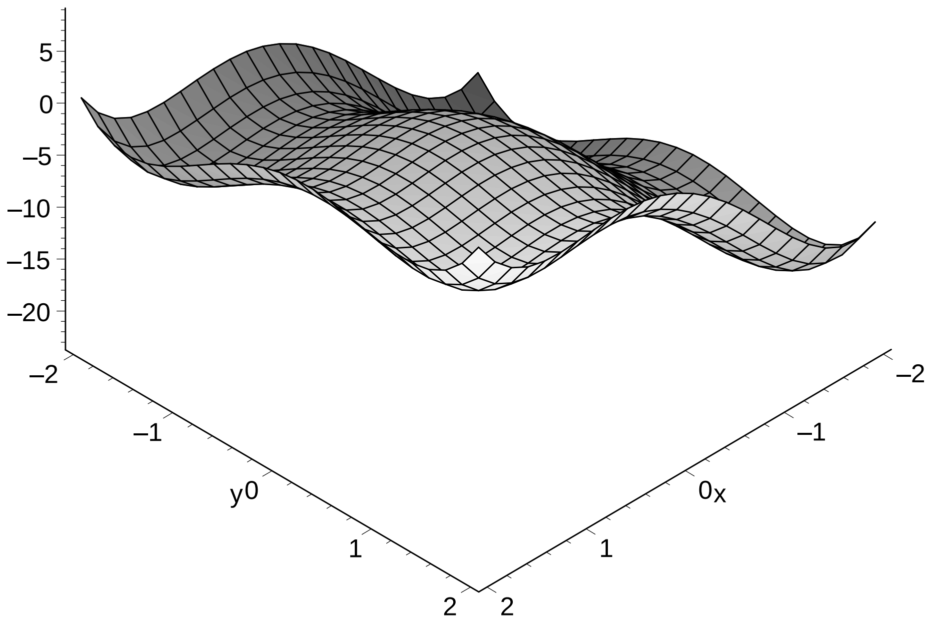
$$\begin{array}{rl} \text{max} & c^T x \\ \text{unter} & Ax = b \\ & x \geq 0 \end{array}$$

  Hier: $S = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ Polyeder

- $\min_{x\in[-2,2]} x^3 - 3x$



- $\min_{x,y\in\mathbf{R}} x^4 + y^4 - 5x^2 - 4y^2 + 5x + 2y - 1.5$



- **Definition :**
  Seien $S \subset \mathbb{R}^n, f : S \to \mathbb{R}$ und $x^* \in S$.

  x^* ist <u>lokales Minimum</u> von f über $S :\Longleftrightarrow$

  $\exists\ \epsilon > 0\ \forall x \in S \cap U_\epsilon(x^*) : f(x) \geq f(x^*)$

  x^* ist <u>striktes lokales Min.</u> von f über S:$\Longleftrightarrow$

  $\exists\ \epsilon > 0\ \forall x \in S \cap U_\epsilon(x^*), x \neq x^* : f(x) > f(x^*)$

  x^* ist <u>globales Minimum</u> von f über S:$\Longleftrightarrow$

  $\forall x \in S : f(x) \geq f(x^*)$

  x^* ist <u>striktes globales Min.</u> von f über S:$\Longleftrightarrow$

  $\forall x \in S x \neq x^* : f(x) > f(x^*)$

- **Satz** $(n = 1)$**:**
  Seien $x^* \in I = (a, b)$ und $f \in C^2(I)$.

  $f'(x^*) = 0, f''(x^*) > 0 \Longleftrightarrow$
  x^* ist striktes lokales Minimum von f.

- Verallgemeinerung auf $\mathbb{R}^n$?

- **Kurven :**

  Sei $I \subset \mathbb{R}$ ein Intervall

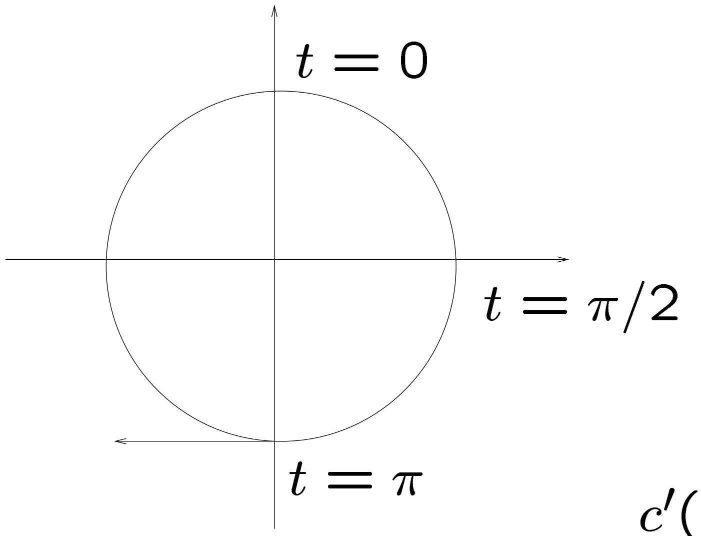
  $c : I \to \mathbb{R}^n$ heißt Kurve:$\Longleftrightarrow c_i$ stetig

  $t_0 \in I, c'(t_0) := \Big(c_1'(t_0), \ldots, c_n'(t_0)\Big)$
  $c'(t_0)$ heißt Tangentialvektor an c in t_0

  Beispiel :

  $c :\ [0, 2\pi] \to \mathbb{R}^2$

  $\quad t \longmapsto (\sin(t), \cos(t))$

  $c'(\pi) = (-1, 0)$

  Proposition :

  Sei $x^* \in S$ ein lokales Min. und $c : [0, a] \to S$ ein Weg mit $c(0) = x^*$ und $c'(0) \neq 0$, dann

  $\exists\ \delta \leq a\ \forall \alpha \in (0, \delta] : f(c(\alpha)) \geq f(x^*)$

- **Partielle Ableitungen :**

  Sei $f : S \to \mathbb{R}$, sowie $c : I \to \mathbb{R}^n$ stetig differenzierbare Kurve mit $c(t_0) = p \in S$ und $c'(t_0) = d$.

  - $\frac{\partial f}{\partial d}(p) := (f \circ c)'(t_0)$ heißt

    Richtungsableitung von f in Richtung d

  - $\frac{\partial f}{\partial x_i}(p) := \frac{\partial f}{\partial e_i}(p)$ heißt

    i–te partielle Ableitung von f

  - $\nabla f(p) = \left(\frac{\partial f}{\partial x_1}(p), \ldots, \frac{\partial f}{\partial x_n}(p)\right)^T =: Df(p)^T$

    heißt Gradient von f

  - $D^2 f(p) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(p) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(p) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(p) \end{pmatrix}$

    heißt Hesse-Matrix von f

  - Für vektorwertiges $h : U \subset \mathbb{R}^n \to \mathbb{R}^k$ heißt

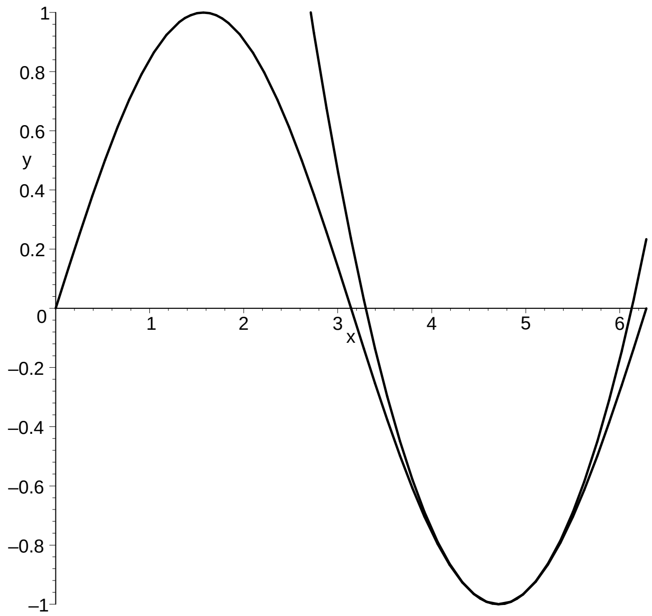
    $Dh(x) := \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(x) & \cdots & \frac{\partial h_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_k}{\partial x_1}(x) & \cdots & \frac{\partial h_k}{\partial x_n}(x) \end{pmatrix}$ die

    Jacobi-Matrix von h.

- **Satz von Taylor** (n=1):

  Sei $x_0 \in I \subset \mathbb{R}$ und $f \in C^2(I, \mathbb{R})$. Dann:

$$\begin{aligned} f(x) = & f(x_0) + f'(x_0)\ (x - x_0) \\ & + \frac{1}{2} f''(x_0)\ (x - x_0)^2 + o((x - x_0)^2) \end{aligned}$$



- **Satz :**

  Sei $p \in S \subset \mathbb{R}^n$, $f \in C^2(S, \mathbb{R})$, $c \in C^2(I, \mathbb{R}^n)$ mit $c(t_0) = p$ und $c'(t_0) = d$. Dann:

$$\begin{aligned} (f \circ c)(t) = & f(p) + Df(p)d(t - t_0) \\ & + \frac{1}{2} Df(p)c''(t_0)\ (t - t_0)^2 \\ & + \frac{1}{2} d^T D^2 f(p) d(t - t_0)^2 + o((t - t_0)^2) \end{aligned}$$

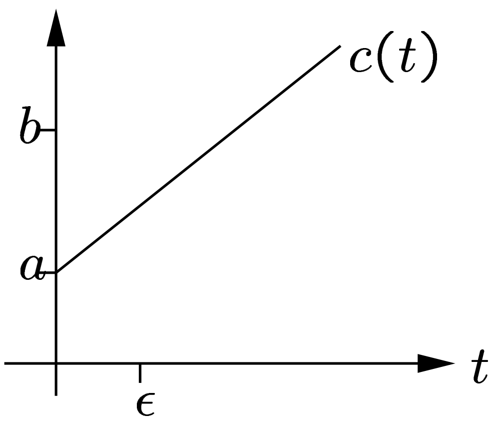
- **Notwendige und hinreichende Bedingungen für Extremwerte**

  Definition :

  Sei $S \subset \mathbb{R}^n, x \in S$ und $d \in \mathbb{R}^n$.
  d heißt zulässige Richtung für x bzgl. $S :\Longleftrightarrow$
  $\exists\ \epsilon > 0, c \in C^1([0,\epsilon], S) : c(0) = x, c'(0) = d$

  Beispiel :

  Sei $S = [a,b], x = a$ und $c(t) := a + dt$, dann ist $c'(0) = d$ und es gilt:
  d zulässig $\Longleftrightarrow d \geq 0$.



  Proposition :

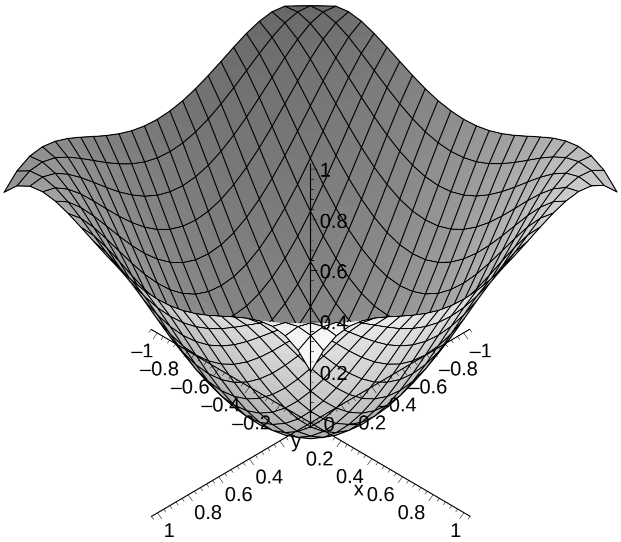
  Sei $S \subset U \subset \mathbb{R}^n, U$ offen und $f \in C^1(U)$. Dann:
  x^* relatives Minimum von f in $S \Longrightarrow$
  $Df(x^*)d \geq 0$ für alle zulässigen Richtungen d.

  Korollar :

  x^* rel. Min. im Innern von $S \Rightarrow Df(x^*) = 0$.

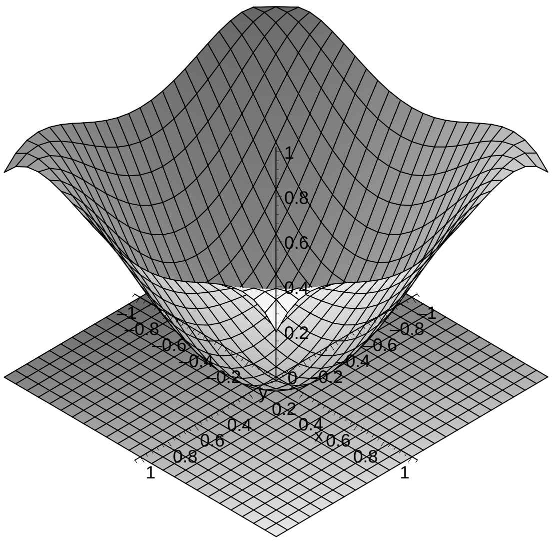
Beispiel :

$$\begin{array}{ll} \text{min} & f(x_1,x_2) = \sin(x_1^2 + x_2^2) \\ \text{unter} & -1 \le x_1 \le 1, -1 \le x_2 \le 1 \end{array}$$



Tangentialebene in $(0,0)$:

$$u(x_1,x_2) = f(0,0) + Df(0,0)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$



Beispiel :

min $\quad f(x_1, x_2) = x_1^2 - x_1 + x_2 + x_1 x_2$
unter $x_1 \geq 0, x_2 \geq 0$

Es ist $Df(x_1, x_2) = (2x_1 - 1 + x_2, x_1 + 1)$.

- $x_1 \geq 0 \Rightarrow Df(x_1, x_2) \neq 0 \Rightarrow f$ hat kein lokales Minimum im Innern von S

- Minimum am Rand $\Rightarrow x_1 = 0 \vee x_2 = 0$

$x = \begin{pmatrix} 0 \\ x_2 \end{pmatrix}, d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ zulässige Richtung
$\Rightarrow d_1 \geq 0$

$d = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \Rightarrow Df(x)d = (x_2 - 1, 1) \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -1$
$\Rightarrow d$ zulässige Abstiegsrichtung
$\Rightarrow x$ kein relatives Minimum

$x = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}, d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ zulässige Richtung
$\Rightarrow d_2 \geq 0$

$Df(x)d = (1 - 2x_1)d_1 + (x_1 + 1)d_2$
$Df(x)d \geq 0 \Longleftrightarrow x_1 = 1/2$

$\Rightarrow x^* = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$ relatives Minimum !

Proposition :

Sei $S \subset U \subset \mathbb{R}^n, U$ offen und $f \in C^2(U)$. Dann gilt für alle zulässigen Richtungen d:

x^* relatives Minimum von f in $S \Longrightarrow$

- $Df(x^*)d \geq 0$

- $Df(x^*)d = 0 \Rightarrow d^T D^2 f(x^*)d \geq 0$.

Korollar :

Sei x^* relatives Min. im Innern von S. Dann gilt für alle $d \in \mathbb{R}^n$:

$Df(x^*) = 0$ und $d^T D^2 f(x^*)d \geq 0$

Proposition :

Sei $U \subset \mathbb{R}^n$ offen, $f \in C^2(U)$ und $x^* \in U$. Dann gilt:

$Df(x^*) = 0$ und $D^2 f(x^*)$ positiv definit $\Rightarrow$

x^* striktes lokales Minimum von f.

Beispiel :

min $f(x,y) = x^2 + \frac{2}{9}y^3 - 2xy$ über $\mathbb{R}^2$

$$Df(x,y) = (2x - 2y, \frac{2}{3}y^2 - 2x) = 0$$

$$\iff \quad x = y, \frac{2}{3}y^2 = 2x$$

$$\iff \quad (x,y) = (0,0) \vee (x,y) = (3,3)$$

Hesse Matrix:

$$D^2 f(x,y) = \begin{pmatrix} 2 & -2 \\ -2 & \frac{4}{3}y \end{pmatrix}$$

Für $x = y = 3$ sind die Eigenwerte $\lambda_{1,2} = 3 \pm \sqrt{5}$, die Hesse-Matrix ist positiv definit.

$\Rightarrow$ $(3,3)$ striktes Minimum.

Für $x = y = 0$ sind die Eigenwerte $\lambda_{1,2} = 1 \pm \sqrt{5}$, die Hesse-Matrix ist indefinit.

$\Rightarrow$ $(0,0)$ kein Minimum, sondern Sattelpunkt