§ 4 Testfunktionen

(für numerische Experimente, eindeutige Minimalstellen)

i) Generiere $\bar{x}_1, \dots, \bar{x}_m \in \mathbb{R}_+ = \{t \in \mathbb{R} \mid t \geq 0\}$ zufällig

$$\bar{s}_{m+1}, \dots, \bar{s}_n \in \mathbb{R}_+ \quad \text{dito.} \qquad (m < n,\ m, n \in \mathbb{N})$$

Setze $\bar{x}_{m+1} = \dots = \bar{x}_n = 0, \quad \bar{s}_1 = \dots = \bar{s}_m = 0.$

Generiere $\bar{y} \in \mathbb{R}^m$ zufällig.

$$A \in \mathbb{R}^{m \times n} \quad \text{"} \qquad (\text{z.B. iid. } N(0,1))$$

Setze $b := A \cdot \bar{x}, \quad c := A^T \bar{y} + \bar{s}.$

Für $z := (x^T, y^T, s^T)^T \in \mathbb{R}^{n+m+n}$ minimiere (ausgehend von $z^0 = 0$)

$$f(z) := \|Ax - b\|_2^2 + \|A^T y + s - c\|_2^2 + (c^T x - b^T y)^2 + \|x_-\|_2^2 + \|s_-\|_2^2$$

wobei $(x_-)_i := \min\{0, x_i\}$, d.h. x_- "enthält nur die negativen Komponenten von x!"

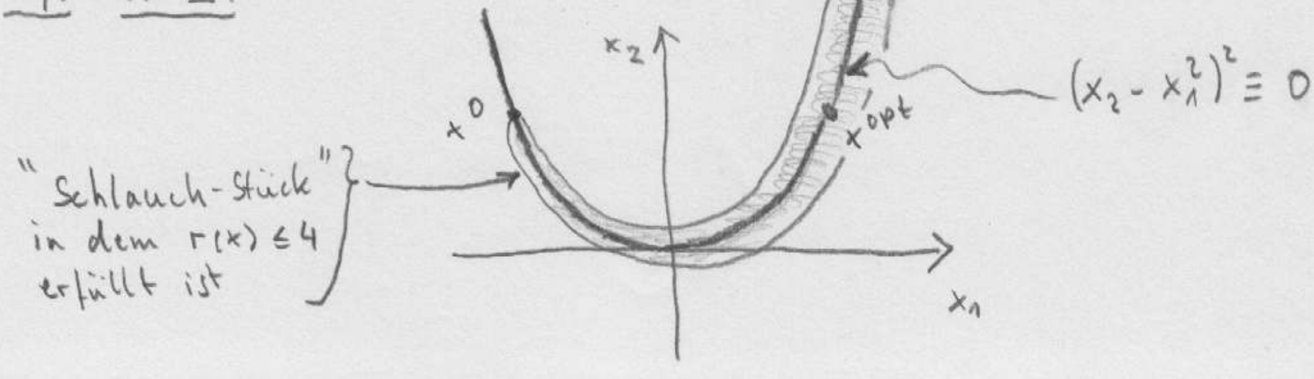
(Zusammenhang mit linearen Programmen – "so löst man die linearen Programme in der Praxis nicht".)

ii) "Rosenbrock-Funktion"

$$r(x) := (x_1 - 1)^2 + 100 \sum_{i=2}^{n} (x_i - x_{i-1}^2)^2$$

mit $x^0 := (-1, 1, \dots, 1)^T \in \mathbb{R}^n$.

Bsp. n=2:



iii) Nesterov's Rosenbrock-Funktion

$$f(x) := \frac{1}{4}(x_1-1)^2 + \beta \underbrace{\sum_{i=1}^{n-1} (x_{i+1} - 2x_i^2+1)^2}_{} \qquad \beta \geq 1 \text{ fest.}$$

Nullstellenmenge, x_n oszilliert
in $[-1,1]^n$ mit $2^{n-1}-1$ Wechseln
zwischen -1 und 1
(Tschebyscheff-Polynome)

"Jedes Abstiegsverfahren mit $x^{k+1} = x^k + \lambda_k \Delta x^k$ so dass f in $[0, \lambda_k]$ entlang $x^k + t\Delta x^k$ monoton fällt benötigt mind $1{,}618^{n-1}$ Schritte um f ausgehend von $x^0 = (-1, 1, \dots, 1)^T$ von $f(x^0) = 1$ auf $f(x^k) \leq \frac{1}{2}$ zu reduzieren, sofern $\beta \geq 400$ gewählt wird."

iv) viele weitere Beispiele ...

§5 Simplexverfahren (Nelder & Mead)[9]

(Nicht zu verwechseln mit dem Simplex-Verfahren zur Lösung linearer Programme).

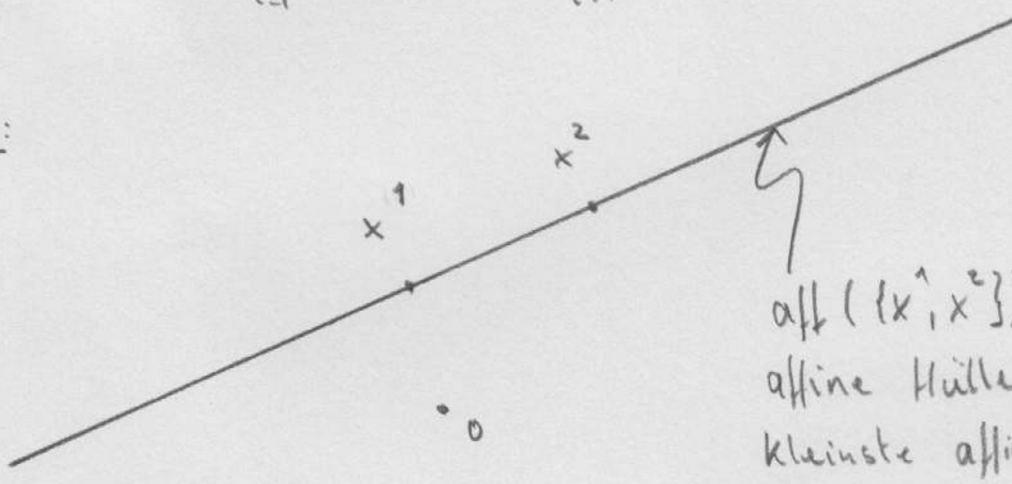
(1') $\min f(x) \mid x \in \mathbb{R}^n$.

<u>Def:</u> Die Punkte $x^1, \dots, x^p \in \mathbb{R}^n$ sind <u>affin unabhängig</u>

$$:\Leftrightarrow \binom{x^1}{1}, \dots, \binom{x^p}{1} \in \mathbb{R}^{n+1} \text{ sind linear unabhängig}$$

$$:\Leftrightarrow \left\{ \left\{ \sum_{i=1}^{p} \lambda_i x^i = 0 \ \& \ \sum_{i=1}^{p} \lambda_i = 0 \right\} \Rightarrow \lambda_i \equiv 0 \ (1 \le i \le p) \right\}.$$

<u>Bsp:</u>



$\mathrm{aff}(\{x^1, x^2\})$
affine Hülle von x^1 & x^2
kleinste affine Menge, die x^1 und x^2 enthält

$$\mathrm{aff}(\{x^1,x^2\}) = \{x^1 + t(x^2 - x^1) \mid t \in \mathbb{R}\}$$
$$= \left\{ x^1 + \sum_{i=1}^{2} \lambda_i x^i \ \Big| \ \sum_{i=1}^{2} \lambda_i = 0 \right\}.$$

<u>Simplex:</u> Zu affin unabhängigen Punkten $x^1, \dots, x^{n+1} \in \mathbb{R}^n$ ist $\mathrm{conv}(\{x^1, \dots, x^{n+1}\})$ die konvexe Hülle von $x^1, \dots, x^{n+1}$ (kleinste konvexe Menge, die $x^1, \dots, x^{n+1}$ enthält) ein Simplex.

[9] "Anfangszeit der Nutzung von Computer-Programmen", 1962 bzw 1965

Die Punkte $x^1, \dots, x^{n+1}$ sind die Ecken (engl. vertices) und der Schwerpunkt ist gegeben durch $\bar{x} = \frac{1}{n+1} \sum_{i=1}^{n+1} x^i$.

Simplex-Verfahren (nach Spendley, Hext, Himsworth (1962))

- Gegeben $x^0 \in \mathbb{R}^n$. Wähle $d > 0$ und bilde die Ecken $x^1, \dots, x^{n+1}$ um x^0 herum als gleichmäßigen Simplex mit Kantenlänge d. (Übung)
- Berechne $f(x^1), \dots, f(x^{n+1})$ und sortiere die x^i so dass
  $$f(x^1) \le f(x^2) \le \dots \le f(x^{n+1})$$
- Wiederhole:
  - Reflektiere x^{n+1} am Schwerpunkt von $x^1, \dots, x^n$:
    $$\hat{x} := 2\left(\frac{1}{n} \sum_{i=1}^{n} x^i\right) - x^{n+1}$$
  - Falls $f(\hat{x}) < f(x^{n+1})$ so ersetze x^{n+1} durch $\hat{x}$ und sortiere die x^i erneut (d.h. füge $\hat{x}$ entsprechend des Wertes $f(\hat{x})$ in $x^1, \dots, x^n$ ein).
  - Falls $f(\hat{x}) \geq f(x^{n+1})$ und $d \le \varepsilon$ stop
  - Falls $f(\hat{x}) \geq f(x^{n+1})$ und $d > \varepsilon$ schrumpfe den Simplex, $x^i := \frac{1}{2}(x^i + x^1)$ für $i = 2, \dots, n+1$ und sortiere die x^i erneut. (Damit auch $d \to \frac{d}{2}$.)

(Modifizierung: Falls $f(\hat{x}) \in \left[f(x^n), f(x^{n+1})\right)$ so reflektiere im nächsten Schritt x^n am Schwerpunkt der übrigen Punkte - und nicht das neue $x^{n+1} = \hat{x}$.)