

# Skript zur Vorlesung Direkte Suchverfahren

Florian Jarre  
Skript erstellt von Luca Püttmann

Sommersemester 2025

# Inhaltsverzeichnis

§2 Line-search-Verfahren	1
§3 Verfahren auf diskreten Gittern	8
§4 Testfunktionen	12
§5 Simplexverfahren (Nelder & Mead)	14
§6 Methoden der konjugierten Richtungen	26
§7 Lineare Approximationsmodelle	29

## §2 Line-search-Verfahren

Problem (1') minimiere  $f(x)$  mit  $l \leq x \leq u$ , wobei  $l_i \in \{-\infty\} \cup \mathbb{R}$ ,  $u_i \in \mathbb{R} \cup \{\infty\}$ ,  $l_i \leq u_i$  und die Ungleichung komponentenweise aufgefasst wird.

(**Bem.:**  $\min \leftrightarrow \max$ )

### Allgemeines line-search-Verfahren:

Startpunkt  $x^0 \in \mathbb{R}^n$

**Für**  $k = 0, 1, 2, \dots$

Bestimme eine Suchrichtung  $s^k \in \mathbb{R}^n$

Bestimme eine Schrittweite  $\alpha_k \in \mathbb{R}$ , s. d.  $F(x^k + \alpha_k s^k) < F(x^k)$

Setze  $x^{k+1} = x^k + \alpha_k s^k$

**End**

### Voraussetzung 1

Es gibt eine Zahl  $l \geq n$  so, dass stets gilt:

$l$  aufeinanderfolgende Suchrichtungen  $s^k$  spannen  $\mathbb{R}^n$  in folgendem strikten Sinn auf:

$$\exists c \in (0, 1) : \forall k \geq l \forall \nu \in \mathbb{R}^n : \max \left\{ \frac{|\nu^T s^j|}{\|s^j\|_2} \mid k - l + 1 \leq j \leq k \right\} \geq c \|\nu\|_2$$

**Bsp.:** Setze  $c := \frac{1}{\sqrt{n}}$  und  $s^1, s^2, \dots = e_1, e_2, \dots, e_n, e_1, e_2, \dots$ , d. h. die  $s^i$  durchlaufen zyklisch die kanonischen Einheitsvektoren. Dann ist Voraussetzung 1 mit  $l = n$  erfüllt:

$$\|s^j\|_2 = 1 \forall j \text{ und } \|\nu\|_2 \leq \sqrt{n} |\nu_{max}| = \sqrt{n} \max_j |\nu^T s^j|. \quad \#$$

**Bem.:** Falls Voraussetzung 1 verletzt ist, so sind die Iterierten  $x^k$  (zumindest für großes  $k$ ) "nahezu auf einem affinen Teilraum eingeschränkt". Somit kann die gesuchte Minimalstelle gar nicht oder nur sehr langsam approximiert werden.

Voraussetzung 1 ist aber auch bei "exakter line search" nicht hinreichend für

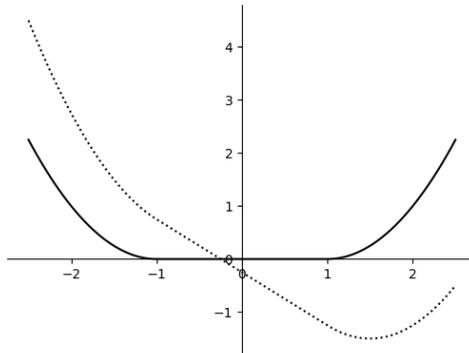
die Konvergenz gegen einen kritischen Punkt:

**Bsp.:** Powell (1973)

$$f(x, y, z) := -xy - xz - yz + (x - 1)_+^2 + (-x - 1)_+^2 + (y - 1)_+^2 + (-y - 1)_+^2 + (z - 1)_+^2 + (-z - 1)_+^2$$

wobei  $t_+ := \max\{t, 0\}$ .

Skizze für  $y = z = 0$ :  $y = z = 1/2$ :



Für gegebenes  $y, z$  wird  $f$  bezüglich  $x$  minimiert, falls

$$-x(y + z) + (x - 1)_+^2 + (-x - 1)_+^2 \rightarrow \min$$

Sei z. B.  $y + z > 0 \Rightarrow$  wähle  $x \geq 1$  (!)

$$\begin{aligned} \Rightarrow \text{minimiere } & -x(y + z) + (x - 1)^2 \\ \Leftrightarrow & -y - z + 2(x - 1) = 0 \\ \Leftrightarrow & x = \frac{y + z}{2} + 1 \end{aligned}$$

Analog, falls  $y + z < 0$ .

Allgemein gilt:  $x = \text{sign}(y + z)(1 + \frac{1}{2}|y + z|)$  (eindeutig, falls  $y + z \neq 0$ .)

Analog: Minimierung bez.  $y$  (oder  $z$ ) für festes  $x, z$  (oder festes  $x, y$ ).

Jetzt soll  $f$  alternierend entlang  $(\Delta x, 0, 0), (0, \Delta y, 0), (0, 0, \Delta z)$  minimiert werden "coordinate descent", Startpunkt sei  $(-1 - \varepsilon, 1 + \frac{\varepsilon}{2}, -1 - \frac{\varepsilon}{4}) =: \nu^{(0)}$ .

Die Minimierung bezüglich  $(\Delta x, 0, 0)$  liefert  $x^{\text{neu}} = 1 + \frac{1}{2} \cdot \frac{\varepsilon}{4}$ , also

$$\begin{aligned}\nu^{(1)} &= \left(1 + \frac{\varepsilon}{8}, 1 + \frac{\varepsilon}{2}, -1 - \frac{\varepsilon}{4}\right) \\ \nu^{(2)} &= \left(1 + \frac{\varepsilon}{8}, -1 - \frac{\varepsilon}{16}, -1 - \frac{\varepsilon}{4}\right) \\ \nu^{(3)} &= \left(1 + \frac{\varepsilon}{8}, -1 - \frac{\varepsilon}{16}, 1 + \frac{\varepsilon}{32}\right) \\ \nu^{(4)} &= \left(-1 - \frac{\varepsilon}{64}, -1 - \frac{\varepsilon}{16}, 1 + \frac{\varepsilon}{32}\right) \\ \nu^{(5)} &= \left(-1 - \frac{\varepsilon}{64}, 1 + \frac{\varepsilon}{128}, 1 + \frac{\varepsilon}{32}\right) \\ \nu^{(6)} &= \left(-1 - \frac{\varepsilon}{64}, 1 + \frac{\varepsilon}{128}, -1 - \frac{\varepsilon}{256}\right).\end{aligned}$$

Hier ist  $\nu^{(6)}$  genauso definiert wie  $\nu^{(0)}$  nur mit  $\frac{\varepsilon}{64}$  statt  $\varepsilon$ . Zyklen "entlang" von 6 Kanten des Würfels mit den Ecken  $(\pm 1, \pm 1, \pm 1)$ .

Dabei ist

$$\nabla f\left(\begin{matrix} x \\ y \\ z \end{matrix}\right) = \begin{pmatrix} -y - z + 2(x - 1)_+ - 2(-x - 1)_+ \\ -x - z + 2(y - 1)_+ - 2(-y - 1)_+ \\ -x - y + 2(z - 1)_+ - 2(-z - 1)_+ \end{pmatrix}$$

und

$$\nabla f\left(\begin{matrix} -1 \\ 1 \\ -1 \end{matrix}\right) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

dito  $\nabla f(\nu) \neq 0$  für die anderen Ecken.

Modifiziertes Verfahren (Lucidi, Sciandrone 1997):

Wähle  $\gamma > 0$ ,  $\delta \in (0, 1)$  und eine Folge  $(\beta_k)_{k \in \mathbb{N}}$  mit  $\beta_k > 0, \beta_k \searrow 0$ .

In jedem Schritt des allgemeinen line-search-Verfahrens werde  $\alpha = \alpha_k$  so gewählt, dass

$$(2) \begin{cases} f(x^k + \alpha s^k) \leq f(x^k) - \gamma \alpha^2 \|s^k\|_2^2 \\ \min\{f(x^k + \hat{\alpha} s^k), f(x^k - \hat{\alpha} s^k)\} \geq f(x^k) - \gamma \hat{\alpha}^2 \|s^k\|_2^2 \end{cases}$$

wobei  $\hat{\alpha} := \frac{\alpha}{\delta}$ .

Dabei wurde zuerst  $\alpha = \pm 1$  in die erste Ungleichung von (2) eingesetzt.

Falls diese erfüllt ist, aber die zweite nicht, so erfüllt auch  $\alpha := \frac{\alpha}{\delta}$  oder  $\alpha := -\frac{\alpha}{\delta}$  die erste Ungleichung.

Der Schritt  $\alpha \rightarrow \pm \frac{\alpha}{\delta}$  wird wiederholt. Entweder gilt dann  $|\hat{\alpha}| \rightarrow \infty$  und  $f(x^k + \hat{\alpha} s^k) \rightarrow -\infty$  (d. h. es gibt keine endliche globale Minimalstelle) oder es wird eine Schrittweite erzeugt, die beide Bed. aus (2) erfüllt.

Falls  $\alpha = \pm 1$  die erste Bedingung in (2) nicht erfüllt, so liefert  $\delta\alpha$  eine Schrittweite, die die zweite Bed. erfüllt. Wiederhole den Übergang  $\alpha \rightarrow \pm\delta\alpha$  so lange, bis entweder die erste Bed. erfüllt ist oder bis

$$\|\alpha s^k\| < \frac{1}{2}\beta_k \quad (3a)$$

gilt. ( $\beta_k$  ist hier fest und  $\alpha \searrow 0$ .)

In letzterem Fall wähle  $\alpha = 0$ , d.h. die aktuelle Suchrichtung wird verworfen und das Verfahren fährt mit der nächsten Suchrichtung fort.

(Update  $x^{k+1} = x^k + \alpha s^k$  wie auf S. (1) und Wechsel  $k \mapsto k + 1$ .)

**Beh.:** Sei Voraussetzung 1 erfüllt und  $f$  stetig differenzierbar und  $x^*$  ein Häufungspunkt der Folge  $\{x^k\}$ . Dann gilt  $\nabla f(x^*) = 0$ .

**Bew.:** (schwächere Voraussetzung als bei den Abstiegsverfahren in Einf. Opti.)

$f(x^k)$  ist monoton fallend, nach unten beschränkt durch  $f(x^*)$ . Wegen (2) Teil 1 gilt

$$f(x^k) - f(x^{k+1}) \geq \gamma \|x^k - x^{k+1}\|_2^2. \quad (3)$$

(Trivial, falls  $\alpha = 0$  gewählt wurde)

Aufsummieren liefert, dass  $\|x^k - x^{k+1}\|$  eine Nullfolge ist (sonst ergäbe  $f(x^k) \rightarrow -\infty \leq f(x^*)$  einen Widerspruch).

Widerspruchsannahme:  $\|\nabla f(x^*)\|_2 =: \eta > 0$ .

Sei  $\varepsilon > 0$  so, dass

$$\|Df(x) - Df(x^*)\| < \frac{c\eta}{2} \quad \forall x \text{ mit } \|x - x^*\| \leq \varepsilon, \quad (\otimes)$$

wobei  $c$  aus Voraussetzung 1 sei. Sei nun  $\gamma$  aus (2).

Da  $\beta_k \rightarrow 0 \quad \exists \tilde{k}_0 : \beta_k < \min\{\frac{\varepsilon}{2\ell}, \frac{c\eta}{2\gamma}\}$  für  $k \geq \tilde{k}_0$  und  $\ell$  auch aus Vorauss. 1.

Da  $\|x^k - x^{k+1}\|$  eine Nullfolge ist, gibt es  $k_0 \geq \tilde{k}_0 : \|x^k - x^*\|_2 \leq \frac{\varepsilon}{2\ell} \quad \forall k \geq k_0$ .  
Ferner gilt  $\infty$  oft:  $\|x^k - x^*\| \leq \frac{\varepsilon}{2}$  (Häufungspunkt  $x^*$ ).

Für jedes solche  $k \geq k_0$  ist  $\|x^{k+j} - x^*\| \leq \varepsilon \quad (0 \leq j \leq \ell)$ .

Zu  $v = \nabla f(x^*)$  erfüllt mindestens eine der Suchrichtungen  $s^{k+j} \quad (0 \leq j \leq \ell - 1)$  nach Voraussetzung 1:

$$\frac{|v^T s^{k+j}|}{\|s^{k+j}\|_2} > c\|v\|_2 = c\eta. \quad (\otimes\otimes)$$

Setze  $\phi(\alpha) := f(x^{k+j} + \alpha s^{k+j})$  mit obigem  $j$  und  $|\alpha| \|s^{k+j}\| \leq \frac{\varepsilon}{2\ell}$ .

$$\begin{aligned}
\Rightarrow |\phi'(\alpha)| &= |Df(x^{k+j} + \alpha s^{k+j})s^{k+j}| \\
&\stackrel{\Delta\text{-Ungl.}}{\geq} |Df(x^*)s^{k+j}| - \underbrace{\|Df(x^{k+j} + \alpha s^{k+j}) - Df(x^*)\|_2}_{\leq \frac{c\eta}{2} \text{ (wegen } \circledast)} \|s^{k+j}\|_2 \\
&\stackrel{(\circledast\circledast)}{\geq} \frac{1}{2}c\eta \|s^{k+j}\|_2.
\end{aligned}$$

Insbesondere hat  $\phi'$  keine Nullstelle und das Vorzeichen ist für  $|\alpha| \leq \frac{\varepsilon}{2\ell\|s^{k+j}\|}$  konstant.

Mit  $\phi(\alpha) = \phi(0) + \int_0^\alpha \phi'(t) dt$  folgt

$$f(x^{k+j} + \alpha s^{k+j}) \leq f(x^{k+j}) - |\alpha| \frac{c\eta}{2} \|s^{k+j}\|_2 \text{ falls } \|\alpha s^{k+j}\|_2 \leq \frac{\varepsilon}{2\ell} \quad (5)$$

$\uparrow$   
 passendes Vorz. wählen  $\alpha\phi'(0) < 0$

Dabei beschreibt (5) einen gewissen Mindestanstieg.

**Beh.:** Für diese  $j$  wird  $|\alpha|$  nicht auf weniger als  $\frac{\beta_{k+j}}{\|s^{k+j}\|_2}$  reduziert.

**Bew.:** Falls nun  $\|\alpha s^{k+j}\| < \beta_{k+j}$  wäre, so wäre mit

$$\begin{aligned}
f(x^{k+j} + \alpha s^{k+j}) &\stackrel{(5)}{\leq} f(x^{k+j}) - \frac{c\eta}{2} \|\alpha s^{k+j}\|_2 \\
&\leq f(x^{k+j}) - \frac{c\eta}{2} \frac{\|\alpha s^{k+j}\|_2^2}{\beta_{k+j}} \\
&\leq \underset{\uparrow}{f(x^{k+j})} - \gamma \|\alpha s^{k+j}\|_2^2 \\
&\text{Def. } \tilde{k}_0 \Rightarrow \beta_{k+j} \leq \frac{c\eta}{2\gamma}
\end{aligned}$$

die 1. Bed. von (2) erfüllt, und  $|\alpha|$  würde nicht weiter reduziert. Die Schranke

$$\|\alpha s^{k+j}\| \stackrel{(3a)}{\leq} \frac{1}{2}\beta_{k+j} \text{ würde nicht erreicht.}$$

$\Rightarrow$  Der Fall  $\alpha = 0$  kann in maximal  $\ell - 1$  aufeinanderfolgenden Iterationen auftreten.

Für  $\alpha \neq 0$  impliziert die 2. Bed. aus (2) (mit  $k$  statt  $k + j$ )

$$-\gamma \|\hat{\alpha} s^k\|_2^2 \leq f(x^k + \hat{\alpha} s^k) - f(x^k) \stackrel{(-)}{\leq} \underset{\uparrow}{-\frac{1}{2}c\eta} \|\hat{\alpha} s^k\|_2$$

(5) gilt auch für  $\hat{\alpha}$ ,  $|\hat{\alpha}| \leq \frac{\varepsilon}{2\ell\|s^k\|}$

d.h.  $\gamma \|\hat{\alpha} s^k\| \geq \frac{1}{2}c\eta$  bzw.  $\|x^{k+1} - x^k\| \stackrel{\alpha = \delta \hat{\alpha}}{=} \|\delta \hat{\alpha} s^k\| \geq \underset{\uparrow}{\frac{\delta c\eta}{2\gamma}} > 0$  im Widerspruch  
wegen  $\gamma \|\hat{\alpha} s^k\| \geq \frac{1}{2}c\eta$

dazu, dass  $\|x^{k+1} - x^k\|$  eine Nullfolge ist.  $\#$

Anmerkungen zur line search:

- Gegeben  $x^k$ , setze  $\phi(t) := f(x^k + t\Delta x^k)$
- Gesucht ist eine lokale Minimalstelle  $\bar{t}$  von  $\phi$  mit  $\phi(\bar{t}) \leq \phi(0)$ .  
(Manche schlecht implementierten line-search-Algorithmen, wie z. B. `fminbnd` in Matlab, liefern auch Werte  $\bar{t}$  mit  $\phi(\bar{t}) > \phi(0)$ .)
- Das Verfahren des goldenen Schnitts (Einf. Opt.) konvergiert mit einer Rate von  $\frac{\sqrt{5}-1}{2} \approx 0.61803\dots$
- Zur lokalen Konvergenzbeschleunigung bietet sich quadratische Interpolation durch die letzten 3 Iterierte an, die Auswertung von deren Minimalstelle und verschiedene "safe guards".  
Recht technisch, kaum Veröffentlichungen, da zu trivial.
- Nutzung von kubischen Splines anstelle der quadrat. Interpolierenden (5 Stützpunkte)

1)) Zu  $t_0, t_1, \dots, t_n$  und  $\phi_0, \phi_1, \dots, \phi_n$  berechne die kubische Spline-Fkt.  $s$  mit

$$\begin{aligned} s(t_j) &= \phi_j & (0 \leq j \leq n) & \quad \text{und} \\ s'(t_0) &= 0 \\ s''(t_0) &= 0 \end{aligned}$$

$s$  kann ganz einfach berechnet werden:

Auf  $[t_0, t_1]$  gilt

$$s(t) = \phi_0 + s'(t_0)(t - t_0) + s''(t_0)\frac{(t - t_0)^2}{2} + \alpha(t - t_0)^3,$$

wobei  $s(t_0), s'(t_0), s''(t_0)$  gegeben sind und  $\alpha$  so zu bestimmen ist, dass  $s(t_1) = \phi_1$  erfüllt ist, also

$$\alpha = \frac{\phi(t_1) - s(t_0) - s'(t_0)(t_1 - t_0) - s''(t_0)\frac{(t_1 - t_0)^2}{2}}{(t_1 - t_0)^3}.$$

Damit ist  $s$  auf  $[t_0, t_1]$  vollständig bestimmt.

$\Rightarrow s(t_1) = \phi_1$  und  $s'(t_1)$  sowie  $s''(t_1)$  sind bekannt,

$$\begin{aligned} s'(t_1) &= s'(t_0) + s''(t_0)(t_1 - t_0) + 3\alpha(t_1 - t_0)^2, \\ s''(t_1) &= s''(t_0) + 6\alpha(t_1 - t_0) \end{aligned}$$

Analog kann jetzt  $s$  auf  $[t_1, t_2], [t_2, t_3], \dots$  berechnet werden.  
Die Ableitungen  $s'(t_0) = 0 = s''(t_0)$  waren dabei willkürlich gewählt.

Bestimme auf dieselbe Weise zwei Spline-Funktionen  $\hat{s}$  und  $\bar{s}$   
mit  $\hat{s}(t_j) = \bar{s}(t_j) = 0$  für  $0 \leq j \leq n$

$$\text{und } \hat{s}'(0) = 1 \quad \text{sowie} \quad \bar{s}'(0) = 0$$
$$\hat{s}''(0) = 0 \quad \bar{s}''(0) = 1.$$

Jede interpolierende Spline-Fkt. hat die Form  $t \mapsto s(t) + \alpha\hat{s}(t) + \beta\bar{s}(t)$   
mit festem  $(\alpha, \beta) \in \mathbb{R}^2$ . Möchte man nun z. B.  $\alpha, \beta$  so wählen, dass  
 $s'''(t_0)$  und  $s'''(t_{n-1})$  existieren (“*not-a-knot-spline*”) so ergibt sich mit  
der Notation  $s'''_{-}(t_j) := \lim_{t \nearrow t_j} s'''(t)$  und  $s'''_{+}(t_j) := \lim_{t \searrow t_j} s'''(t)$  die Forderung:

$$\underbrace{s'''_{-}(t_1)}_s + \alpha\hat{s}'''_{-}(t_1) + \beta\bar{s}'''_{-}(t_1) = s'''_{+}(t_1) + \alpha\hat{s}'''_{+}(t_1) + \beta\bar{s}'''_{+}(t_1)$$

s: kub. Polynom auf  $[t_0, t_1]$

und analog bei  $t = t_{n-1}$ .

Die Koeffizienten sind explizit gegeben, es ergibt sich ein lineares  
Gleichungssystem mit 2 Variablen & 2 Unbekannten.



Das System kann schlecht konditioniert sein, ggf. muss  $s$  mit  
den gegebenen Anfangswerten  $\alpha, \beta$  neu berechnet werden.

- Analog kann man auch den natürlichen Spline berechnen
- Analog kann man auch einen “*least-squares-spline*” berechnen, der  $\alpha, \beta$   
so bestimmt, dass eine gewichtete Quadratsumme der Sprünge der  
3. Ableitung an den Stellen  $t_1, t_2, \dots, t_{n-1}$  minimiert wird.  
(z. B. Gewichte  $w_i = \frac{1}{t_{i+1}-t_{i-1}}$ ). (Warum?)

### §3 Verfahren auf diskreten Gittern

Problem (1')  $\min f(x) \mid \ell \leq x \leq u$ . Diesmal  $\ell \leq u$ ,  $\ell, u \in \mathbb{R}^n$ . Wähle  $\nu_i \in \mathbb{N} \setminus \{0\}$  für  $1 \leq i \leq n$  und setze  $h := \frac{u_i - \ell_i}{\nu_i}$

$$\begin{aligned} \mathcal{G} &:= \{x \mid \frac{x_i - \ell_i}{h_i} \in \{0, 1, \dots, \nu_i\} \text{ für } 1 \leq i \leq n\} \\ &= \{\ell + \sum_{i=1}^n j_i h_i e_i \mid j_i \in \{0, 1, \dots, \nu_i\} \text{ für } 1 \leq i \leq n\} \end{aligned}$$

wobei  $e_i$  der  $i$ -te kanonische Einheitsvektor sei. "Gitter  $\mathcal{G}$ ". Zu  $x \in \mathcal{G}$  sei

$$\mathcal{N}(x) := \{x\} \cup \{x \pm h_i e_i\}_{1 \leq i \leq n} \cap [\ell, u]$$

die Nachbarschaft von  $x$  in  $\mathcal{G}$  bei der je maximal eine Komponente von  $x$  geändert wird.

Ansatz:

$x^1 \in \mathcal{G}$  gegeben.

**Für**  $k = 1, 2, \dots$

$$\bar{x} := x^k$$

**Für**  $i = 1, 2, \dots, n$

Ersetze  $\bar{x}$  durch  $\bar{x} + z h_i e_i \in \mathcal{G}$  (mit  $z \in \mathbb{Z}$ ).

**Falls**  $z \neq 0$

$$\text{so gelte dabei } f(\bar{x} + z h_i e_i) < f(\bar{x}) \quad \textcircled{*}$$

$$\text{und in jedem Fall gilt } f(\bar{x} + z h_i e_i) \leq f(\bar{x}^\pm)$$

$$\text{für } \bar{x}^\pm \in \{\bar{x} + h_i e_i, \bar{x} - h_i e_i\} \cap \mathcal{G}.$$

**End**

**Falls**  $\bar{x} = x^k$

stop

**sonst**

$$x^{k+1} := \bar{x}.$$

**End.**

In der Bed.  $\circledast$  ist eine “*line search*” entlang einer Koordinatenachse erlaubt, im Gegensatz zum obigen Bsp. von Powell (1973), ist dieser Ansatz aber in jedem Fall endlich.

Bei Abbruch gilt  $f(x^*) \leq f(x) \quad \forall x \in \mathcal{N}(x^k)$ .

Verfahren von Hooke & Jeeves (1961)

$x^0 := x^1 \in \mathcal{G}$  gegeben,  $k := 1$

**Falls**  $y^k := 2x^k - x^{k-1} \in [\ell, u]$  (dann ist  $y^k \in \mathcal{G}$ )!

und **falls**  $\min_{z \in \mathcal{N}(y^k)} f(z) < f(x^k)$

so setze  $x^{k+1} := \arg \min_{z \in \mathcal{N}(y^k)} f(z)$

**andernfalls**

definiere  $x^{k+1}$  wie in  $\circledast$  oben (S.  $\circledast$ ).

Konvergenzaussage:

Sei  $f: \mathbb{R}^k \rightarrow \mathbb{R}$  diff’bar und  $\nabla f$   $L$ -Lipschitz-stetig,

$$\|\nabla f(x + \Delta x) - \nabla f(x)\|_\infty \leq L \|\Delta x\|_\infty \quad \forall x, \Delta x : x, x + \Delta x \in [\ell, u].$$

Dann besitzt  $f$  auf  $[\ell, u]$  mind. eine lok. Minimalstelle  $\bar{x}$  und in jeden lok. Minimalstelle  $\bar{x}$  gilt:

$$\begin{cases} (\nabla f(\bar{x}))_i = 0 & \text{falls } \bar{x}_i \in (\ell, u) \\ (\nabla f(\bar{x}))_i \geq 0 & \text{falls } \bar{x}_i = \ell_i \\ (\nabla f(\bar{x}))_i \leq 0 & \text{falls } \bar{x}_i = u_i. \end{cases}$$

**Bew.:**  $\nabla \#$

Sei  $\bar{h} := \max_{1 \leq i \leq n} h_i$

Dann liefert das Verfahren von Hook & Jeeves bei Abbruch einen Punkt  $x^k$  mit

$$\begin{cases} |\nabla f(x^k)_i| \leq L\bar{h} & \text{falls } x_i^k \in (\ell_i, u_i) \\ \nabla f(x^k)_i \geq -L\bar{h} & \text{falls } x_i^k = \ell_i \\ \nabla f(x^k)_i \leq L\bar{h} & \text{falls } x_i^k = u_i. \end{cases}$$

**Bew.:**

Dass das Verfahren abbricht folgt, da es nur endlich viele Punkte in  $\mathcal{G}$  gibt und  $f$  bei jedem Wechsel von  $x^k$  zu  $x^{k+1} \neq x^k$  (strikt) reduziert wird. Gilt  $n$ -mal in Folge  $x^{k+1} = x^k$ , so bricht das Verfahren mit einem Punkt  $\bar{x} := x^k$  ab, für den gilt:  $f(\bar{x} \pm h_i e_i) \geq f(\bar{x})$  falls  $\ell_i < \bar{x}_i < u_i$ . Betrachte  $\ell(t) := f(\bar{x} + t e_i)$ . Nach dem Mittelwertsatz ist

$$\begin{aligned} f(\bar{x}) &= \ell(0) \\ &\leq \ell(h_i) \\ &= \ell(0) + h_i \ell'(t_+), \quad t_+ \in [0, h_i] \\ &= f(\bar{x}) + \frac{\partial}{\partial x_i} f(\underbrace{\bar{x} + t_+ e_i}_{=: \xi_{i_+} \in [\bar{x}, \bar{x} + h_i e_i]}) h_i. \end{aligned}$$

Analog:  $f(\bar{x}) \leq f(\bar{x}) - \frac{\partial}{\partial x_i} f(\xi_{i_-}) h_i$  mit  $\xi_{i_-} \in [\bar{x} - h_i e_i, \bar{x}]$

$$\begin{aligned} \Rightarrow \frac{\delta}{\delta x_i} f(\xi_{i_+}) &\geq 0 \geq \frac{\delta}{\delta x_i} f(\xi_{i_-}) \\ \Rightarrow \frac{\partial}{\partial x_i} f(\bar{x}) &\geq \frac{\partial}{\partial x_i} f(\xi_{i_+}) - \underbrace{L \|\bar{x} - \xi_{i_+}\|_\infty}_{\leq h_i \leq \bar{h}} \geq 0 - L\bar{h} \\ \frac{\partial}{\partial x_i} f(\bar{x}) &\leq \frac{\partial}{\partial x_i} f(\xi_{i_-}) + L \|\bar{x} - \xi_{i_-}\|_\infty \leq L\bar{h}. \end{aligned}$$

Analog für  $\bar{x}_i = \ell_i$  oder  $\bar{x}_i = u_i$ .     #

Praktische Umsetzung:

- ggf. Skalierung auf  $-1 \leq x_i \leq 1$ \*
- Anwendung des Verfahrens auf einem groben Gitter  $\mathcal{G}$ , bei Abbruch Verfeinerung von  $\mathcal{G}$  und erneute Anwendung des Verfahrens.
- Modifizierung falls  $\ell_i = -\infty$  oder  $u_i = \infty$ :  
(unbeschränktes Gitter, Abbruch nach “maxit” Schritten)

---

\*Dabei ändert sich die Lipschitz-Konstante  $L$  bzw. die Konstanten “ $L_i \leq L$ ” zu  $\|\nabla f(x) - \nabla f(x + t e_i)\|_\infty \leq L_i |t|$ . ( $\forall x \in [\ell, u], \forall t : x + t e_i \in [\ell, u]$ )

- Falls  $\ell_i = -\infty, u_i = \infty \quad \forall i$  (Torczon 1997)
- Voraussetzung:  $S_1 := \{x \mid f(x) \leq f(x^1)\}$  sei beschränkt.
- wähle  $h_i \equiv h$  für  $1 \leq i \leq n$  (ggf. Minimierung von  $f(Dx)$  mit  $D = (\setminus)$  betrachten) und  $\underline{h} := \underline{1}$  zum Start.
- wähle  $\tau \in (0, 1) \cap \mathbb{Q}, \tau = \frac{\alpha}{\beta}, \alpha, \beta \in \mathbb{N}, (\alpha < \beta)$ .
- wenn bei  $\otimes$  ( $S. \textcircled{8}$ ) gilt  $f(x^{k+1}) = f(x^k)$ , so setze  $h := \tau h$ , andernfalls  $h := \frac{h}{\tau^{j_k}}$  mit  $j_k \in \{0, 1, \dots, \ell\}$  für ein festes  $\ell \in \mathbb{N}$ .

**Beh.:** Falls  $\nabla f$  auf  $S_1$  Lipschitz-stetig ist, so gilt  $\liminf_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$ .

**Bew.:**

Es gilt immer  $h = h^{(k)} \in \{\tau^j \mid j \in \mathbb{Z}\}$  da  $h$  stets mit Potenzen von  $\tau$  multipliziert wird. Da  $S_1$  beschränkt ist, kann  $h^{(k)}$  nicht vergrößert werden, wenn  $S_1$  nur einen Gitterpunkt enthält. ( $f(x^k) \leq f(x^1)$  gilt immer und somit  $f(x^k) \in S_1$ .)

$\Rightarrow \exists \underline{m} \in \mathbb{Z} : h^{(k)} \in \{\tau^j \mid j \geq \underline{m}, j \in \mathbb{Z}\} \quad (h^{(k)} \leq \tau^{\underline{m}})$

**Annahme:**  $\exists \overline{m} \in \mathbb{Z} : h^{(k)} \in \{\tau^j \mid j \leq \overline{m}, j \in \mathbb{Z}\}$

$\Rightarrow x^k$  hat stets die Form, dass die Komponenten gegeben sind durch

$$x_i^k = x_i^1 + \sum_{\underline{m} \leq j \leq \overline{m}} z_{ij} \tau^j,$$

wobei die  $z_{ij} \in \mathbb{Z}$  von  $k$  abhängen. Die Potenzen  $\tau^j$  sind stets Vielfache von  $\alpha^{\underline{m}} \beta^{-\overline{m}}$ . Seien  $\ell, u$  gegeben mit  $S_1 \subset [\ell, u]$ . Da es in  $[\ell_i, u_i]$  nur endlich viele Punkte aus  $x_i^1 + \lambda \alpha^{\underline{m}} \beta^{-\overline{m}}$  mit  $\lambda \in \mathbb{Z}$  gibt, bricht das Verfahren nach endlich vielen Schritten ab<sup>\*\*</sup>.  $\downarrow$

Also ist  $\liminf_{k \rightarrow \infty} h^{(k)} = 0$ . Jedes Mal, wenn  $h^{(k)}$  reduziert wird, gilt

$\|\nabla f(x^k)\|_\infty \leq L h^{(k)}$ .  $\Rightarrow$  Beh.  $\#$

---

<sup>\*\*</sup>Wenn z. B.  $\tau = \frac{1}{\sqrt{2}}$  ist, so legen unendlich viele Zahlen in der Menge  $\{k\tau + \ell \mid k, \ell \in \mathbb{Z}\} \cap [0, 1]$  und obiger Beweis ist nicht anwendbar.

## §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

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

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

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

$A \in \mathbb{R}^{m \times m}$  zufällig (z. B. iid.  $N(0, 1)$ )

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

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  $x$ ".

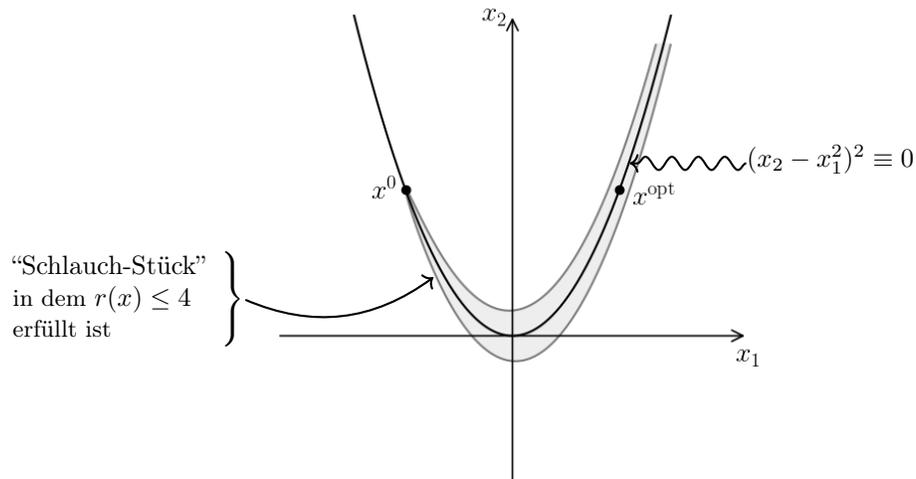
(Zusammenhang mit linearen Programmen – "so löst man die linearen Programme in der Praxis nicht".)

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

$$f(x) := \frac{1}{4}(x_1 - 1)^2 + \beta \underbrace{\sum_{i=1}^{n-1} (x_{i+1} - 2x_i^2 + 1)^2}_{\substack{\text{Nullstellenmenge, } x_n \text{ oszilliert} \\ \text{in } [-1, 1]^n, \text{ mit } 2^{n-1} - 1 \\ \text{Wechseln zwischen } -1 \text{ und } 1 \\ \text{(Tschebyscheff-Polynome)}}} \quad \beta \geq 1 \text{ fest}$$

“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) ¶

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

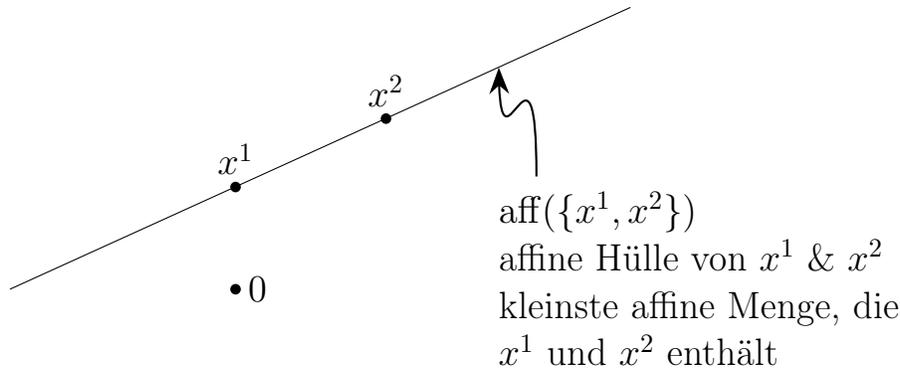
$$(1') \quad \min f(x) \mid x \in \mathbb{R}^n.$$

**Def.:** Die Punkte  $x^1, \dots, x^p \in \mathbb{R}^n$  sind affin unabhängig

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

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

**Bsp.:**



$$\begin{aligned} \text{aff}(\{x^1, x^2\}) &= \{x^1 + t(x^2 - x^1) \mid t \in \mathbb{R}\} \\ &= \left\{ x^1 + \sum_{i=1}^2 \lambda_i x^i \mid \sum_{i=1}^2 \lambda_i = 0 \right\} \end{aligned}$$

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

---

¶ “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  $\alpha > 0$  und bilde die Ecken  $x^1, \dots, x^{n+1}$  um  $x^0$  herum als gleichmäßiger Simplex mit Kantenlänge  $d$ . (Übung)
- Berechne  $f(x^1), \dots, f(x^{n+1})$  und sortiere die  $x^i$  so dass  $f(x^1) \leq f(x^2) \leq \dots \leq f(x^{n+1})$
- 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  $d \leq \varepsilon$ , 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  $d \rightarrow \frac{d}{2}$ .)

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

**Bem. 1:** Für  $\varepsilon = 0, n = 2$  und beschränkter Menge  $\mathcal{N}^0 := \{x \mid f(x) \leq f(x^3)\}$  wird  $d$  unendlich oft reduziert.

**Bew.:**

Für feste Kantenlänge  $d$  bilden die möglichen Simplices ein unendliches Gitter über  $\mathcal{N}^0$ , (Bienenwaben in gleichseitige Dreiecke unterteilt.)

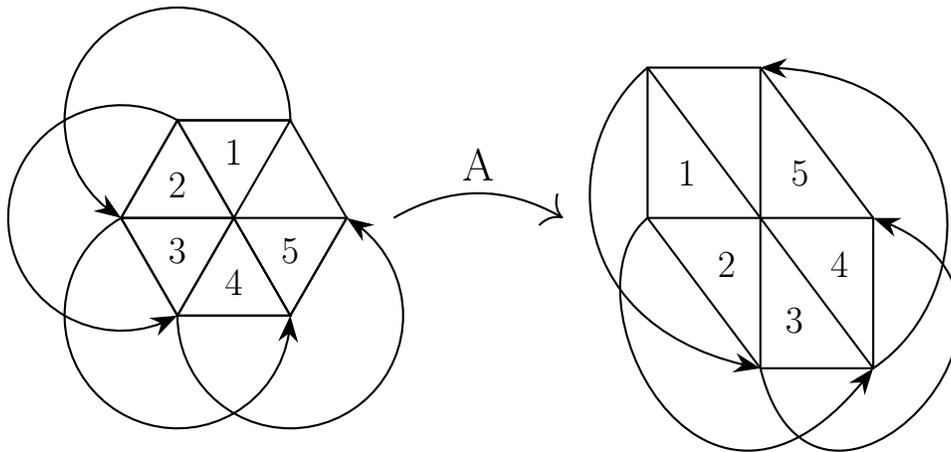
Da die Funktionswerte wegen  $f(\hat{x}) < f(x^{n+1})$  streng monoton fallen wird  $d$  spätestens dann reduziert, wenn  $f$  an allen Gitterpunkten in  $\mathcal{N}^0$  ausgewertet wurde. (Ähnlich, wenn zu Beginn z. B. der Simplex mit den Ecken  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  gewählt wurde. Die Ecken bleiben bei jeder Spiegelung ganzzahlig.)

**Bem. 2:**

i) Induktiv lässt sich zeigen, dass die Methode die Iterierten  $A^{-1}v^{(i)}$  erzeugt, wenn sie nicht auf  $f$  und  $v^{(0)}, \dots, v^{(n)}$  angewendet wird, sondern auf  $\tilde{f}$  und  $\tilde{v}^{(0)}, \dots, \tilde{v}^{(n)}$  mit  $\tilde{f}(y) \equiv f(Ay)$  und  $\tilde{v}^{(i)} := A^{-1}v^{(i)}$

( $0 \leq i \leq n$ ) mit einer invertierbaren Matrix  $A$ . (Die  $\tilde{v}^{(i)}$  bilden dann kein regelmäßiges Simplex mehr.)

**Bsp.:**



**Bem. 3:** Für  $\varepsilon = 0, n = 2$  und beschränkter Menge  $\mathcal{N}^0 := \{x \mid f(x) \leq f(x^3)\}$  kann es (ohne weitere Annahmen an  $f$ ) passieren, dass unendlich viele Simplices (Tetraeder) mit derselben Kantenlänge  $d$  generiert werden.



tung in Real- und Imaginärteil):

$$\begin{aligned}
 \cos(kx) &= \cos^k(x) - \binom{k}{2} \cos^{k-2}(x) \sin^2(x) + \binom{k}{4} \cos^{k-4} \sin^4(x) \\
 &\quad - \binom{k}{6} \cos^{k-6}(x) \sin^6(x) + \dots \\
 &= \sum_{j=1}^{\lfloor \frac{k}{2} \rfloor} (-1)^j \binom{k}{2j} \underbrace{\sin^{2j}(x)}_{(1-\cos^2(x))^j} \cos^{k-2j}(x) \\
 &= \cos^k(x) \underbrace{\left( 1 + \binom{k}{2} + \binom{k}{4} + \dots + \binom{k}{2\lfloor \frac{k}{2} \rfloor} \right)}_{\stackrel{(*)}{=} 2^{k-1}} + \cos^{k-2}(x) \cdot \underbrace{\hspace{2cm}}_{\in \mathbb{Z}} + \dots \\
 &= \left(\frac{1}{3}\right)^k \cdot \underbrace{2^{k-1}}_{\text{kein Vielfaches von 3}} + \left(\frac{1}{3}\right)^{k-2} \cdot \text{“ganzzahlig”} \notin \mathbb{Z}
 \end{aligned}$$

$$(*) : \left\{ \sum_{j=0}^n \binom{k}{j} = 2^k, \sum_{j=0}^k (-1)^j \binom{k}{j} = 0 \right\}$$

Also sind die Simplices, die durch wiederholte Drehungen um  $\alpha$  entstehen, alle verschieden. Das Argument der *“endlich vielen Gitterpunkte”* ist nicht anwendbar, auch wenn sich bei obigem Bsp. keine stetige Funktion findet, die eine solche wiederholte Rotation ad infinitum bedingen würde.

- Modifizierung von Dannis & Torczon (1997):

Spiegle alle Ecken an der “besten” Ecke:  $v^{(i)} \rightarrow \hat{v}^{(i)} = 2v^{(0)} - v^{(i)}, 1 \leq i \leq n$ .

Wenn  $\min_{i \geq 1} f(\hat{v}^i) > f(v^0)$ , schrumpfe den Simplex,  $v^i \rightarrow \frac{1}{2}(v^1 + v^i)$  für  $i \geq 1$ .

Wenn nicht geschrumpft wird, teste eine Verdopplung der Kantenlänge  $\hat{v}^i \rightarrow -v^0 + 2\hat{v}^i$  für  $i \geq 1$ .

**Konvergenz:** (Abbruch bei  $\|v^0 - v\| \leq \delta$ .)

O.B.d.A alle Ecken zu Beginn ganzzahlig.

Induktiv folgt, dass die Ecken ganzzahlig bleiben, solange nicht geschrumpft wird.  $\Rightarrow$  Wenn  $\{x \mid f(x) \leq f(v^0)\}$  beschränkt ist, muss der Simplex wiederholt geschrumpft werden.  $\#$

$$\begin{aligned}
 \text{Bei Abbruch gilt } f(v^0) &\leq f(v^i) & 1 \leq i \leq n \\
 f(v^0) &\leq f(\hat{v}^i) & 1 \leq i \leq n \\
 v^0 &= \frac{1}{2}(v^i + \hat{v}^i) & 1 \leq i \leq m
 \end{aligned}$$

Wie bei den Gitterverfahren folgt  $\|\nabla f(v^0)\| = \mathcal{O}(\delta)$ .

Konvergenzbetrachtung zum Verfahren von Spendley et al (S. (15)).

Bei Abbruch sei o.B.d.A  $f(\hat{v}) > f(v^n)$ , d. h.  $\hat{v}$  ist der größte der Funktionswerte der beiden letzten Simplices.

Setze  $v^* := \frac{1}{n} \sum_{i=0}^{n-1} v^i$  und  $g^* := \nabla f(v^*)$ .

**Annahme:**  $f$  sei stetig diff'bar.

Kantenlänge bei Abbruch des Verfahrens:  $\delta$

Es gilt:  $f(v^i) = f(v^*) + g^{*T}(v^i - v^*) + o(\delta)$

Aufsummiert und durch  $n$  dividiert:

$$\underbrace{\frac{1}{n} \sum_{i=0}^{n-1} f(v^i)}_{\leq f(v^n)} = f(v^*) + o(\delta) \rightsquigarrow$$

ferner:  $f(v^n) = f(v^*) + g^{*T}(v^n - v^*) + o(\delta)$

$$\begin{aligned}
 &\quad \quad \quad \downarrow \\
 &\quad \quad \quad f(v^*) + o(\delta)
 \end{aligned}$$

$\Rightarrow g^{*T}(v^n - v^*) \geq o(\delta)$

Da  $f(\hat{v}) \geq f(v^n)$  folgt analog  $g^{*T}(\hat{v} - v^*) \geq o(\delta)$

Aus  $\hat{v} - v^* = -(v^n - v^*)$  folgt  $g^{*T}(\hat{v} - v^*) = o(\delta)$ .

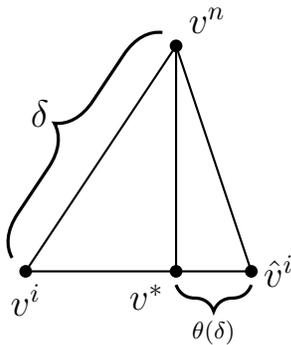
Für die übrigen Ecken folgt

$$\begin{aligned}
 g^{*T}(v^i - v^*) &= f(v^i) - f(v^*) + o(\delta) \\
 &\leq f(v^n) - f(v^*) + o(\delta) \\
 &= g^{*T}(v^n - v^*) + o(\delta) \\
 &= o(\delta). \quad \boxtimes
 \end{aligned}$$

Sei nun  $\hat{v}^i$  der Punkt, an dem die Halbgerade  $v^i + \lambda(v^* - v^i)$ ,  $\lambda \geq 0$ , den

Simplex verlässt,  $\hat{v}^i = \sum_{j=0}^{n-1} \vartheta_j v^j$  mit  $\vartheta_j \geq 0$ ,  $\sum_{j=0}^{n-1} \vartheta_j = 1$ .

$$\Rightarrow g^{*T}(\hat{v}^i - v^*) = \sum_{j=0}^{n-1} \vartheta_j g^{*T}(v^j - v^*) \leq \underbrace{o(\delta)}_{\text{s. o.}}$$



Da der Simplex regulär ist,  
ist  $\|\hat{v}^i - v^*\| = \theta(\delta)$   
(Übung  $\approx \frac{\delta}{n-1}$ )

Da  $\hat{v}^i - v^*$  entgegengesetzt zu  $v^i - v^*$  gerichtet ist folgt  $g^*(v^i - v^*) \geq o(\delta)$ .  
Insgesamt  $g^*(v^i - v^*) = o(\delta)$

Damit sind  $n$  linear unabhängige Richtungen  $r^i$  der Länge  $\theta(\delta)$  gegeben, für die  $g^{*T} r = o(\delta)$  gilt.

$$\Rightarrow \|g^*\| = o(1).$$

(Beim Verfahren von Torczon & Dennis vereinfacht sich obige Argumentation etwas.)

### Das Verfahren von Nelder & Mead (1965)

(Lange Zeit das weltweit populärste Verfahren)

Wie bisher,  $\hat{v} := \left( \frac{2}{n} \sum_{i=0}^{n-1} v^i \right) - v^n \quad \circledast$

1)) Falls  $f(v^0) \leq f(\hat{v}) < f(v^{n-1})$  so ersetze  $v^n$  durch  $\hat{v}$ .  
Sortiere neu und gehe zu  $\circledast$  "wie gehabt".

2)) Falls  $f(\hat{v}) < f(v^0)$  setze  $\hat{\hat{v}} := \hat{v} + \frac{1}{2}(\hat{v} - v^*)$   
falls  $f(\hat{\hat{v}}) < f(\hat{v})$  ersetze  $v^n$  durch  $\hat{\hat{v}}$ , sonst durch  $\hat{v}$ .  
Sortiere neu und gehe zu  $\circledast$  "Extrapolation".

3)) Falls  $f(\hat{v}) \geq f(v^{n-1})$   
(dann würde man im nächsten Schritt zurückspiegeln)

3a)) falls  $f(\hat{v}) < f(v^n)$  setze  $\hat{\hat{\hat{v}}} := \frac{3}{4}\hat{v} + \frac{1}{4}v^n$

- 3b)) falls  $f(\hat{v}) \geq f(v^n)$  setze  $\hat{v} := \frac{3}{4}v^n + \frac{1}{4}\hat{v}$  "kürzerer Schritt"
- falls nun  $f(\hat{v}) < f(v^n)$  ersetze  $v^n$  durch  $\hat{v}$ , sortiere neu und gehe zu  $\otimes$  (in direkt nachfolgenden Schritt wird sicher nicht identisch zurückgespiegelt, da die Schrittlänge verkürzt worden war.)
- sonst schrumpfe den Simplex zu  $v^0$  hin.

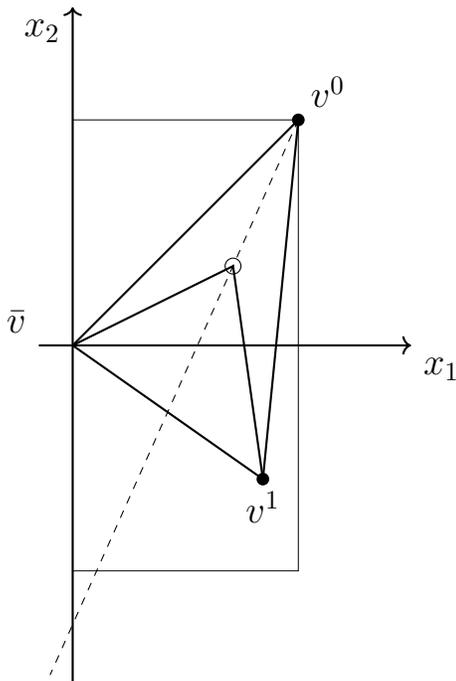
**Bsp.:** Mc Kinnon 1997 (SIOPT)

1)) Konstruktion einer Folge von Simplices, die gegen  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  konvergiert. Nachfolgend ist die Nummerierung dahingehend geändert, dass immer eine neue Ecke  $v^{k+1}$  dazu kommt und zum Start

$$\bar{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, v^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v^1 = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \text{ mit } \lambda_1 = \frac{1 + \sqrt{33}}{8} \approx 0.843$$

$$\lambda_2 = \frac{1 - \sqrt{33}}{8} \approx -0.593$$

gegeben sind und  $f(\bar{v}) < f(v^1) < f(v^0)$  gelte.  
 Es soll wiederholt Schritt 3b durchgeführt werden.  
 Die Dreiecke sollen dabei einer Rekursion folgen:



Erster Schritt:

$$\hat{v} = \frac{3}{2}(\bar{v} + v^1) - v^0 = v^1 - v^0$$

$$\hat{\hat{v}} = \frac{3}{4}v^0 + \frac{1}{4}\hat{v} = \frac{1}{2}v^0 + \frac{1}{4}v^1$$

**Beh.:**  $v^n = \begin{pmatrix} \lambda_1^n \\ \lambda_2^n \end{pmatrix}$  mit obigen  $\lambda_1, \lambda_2$      $e_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $e_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

**Bew.:** (durch Induktion)

IA:

$n = 0$  ✓

$n = 1$  ✓

IS:  $(n - 1, n) \rightarrow n + 1$  :

$$\begin{aligned} \hat{\hat{v}} &= \frac{1}{2}v^n + \frac{1}{4}v^{n-1} \\ &= \frac{1}{4}(\lambda_1^n e_1 + \lambda_2^n e_2) + \frac{1}{2}(\lambda_1^{n-1} e_1 + \lambda_2^{n-1} e_2) \\ &= \lambda_1^{n-1} \left( \frac{1}{4}\lambda_1 + \frac{1}{2} \right) e_1 + \lambda_2^{n-1} \left( \frac{1}{4}\lambda_2 + \frac{1}{2} \right) e_2 \\ &= \lambda_1^{n-1} \left( \frac{1}{32}(1 + \sqrt{33} + 16) \right) e_1 + \lambda_2^{n-1} \left( \frac{1}{32}(1 - \sqrt{33} + 16) \right) e_2 \\ &= \lambda_1^{n-1} \left( \frac{1}{64}(1 + \sqrt{33})^2 \right) e_1 + \lambda_2^{n-1} \left( \frac{1}{64} \underbrace{(1 - \sqrt{33})^2}_{\substack{(1 \pm \sqrt{33})^2 = 1 \pm 2\sqrt{33} + 33 \\ = 2(\pm\sqrt{33} + 17)}} \right) e_2 \\ &= \lambda_1^{n-1} \lambda_1^2 e_1 + \lambda_2^{n-1} \lambda_2^2 e_2 \end{aligned}$$

Also  $\hat{\hat{v}} = \lambda_1^{n+1} e_1 + \lambda_2^{n+1} e_2$ .    #

2)) Eine streng konvexe Funktion, die ein Verhalten wie oben erzeugt:

Wähle  $\theta > 0$ ,  $\phi > 0$ ,  $\tau > 1$  und

$$f(x_1, x_2) := \begin{cases} \theta x_1^\tau + x_2^2 + x_2 & \text{für } x_1 \geq 0 \\ \theta \phi |x_1|^\tau + x_2^2 + x_2 & \text{für } x_1 < 0 \end{cases}$$

$\Rightarrow f$  ist stetig diff'bar (für  $\tau > 2$  sogar 2-mal stetig diff'bar)

Auf den Halbräumen  $x_1 \geq 0$  bzw.  $x_1 \leq 0$  ist  $f$  konvex und aufgrund der Differenzierbarkeit ist  $f$  nach Übungsblatt 1 auf ganz  $\mathbb{R}^2$  konvex.

Sei  $1 < \tau < \hat{\tau} := \frac{\ln |\lambda_2|}{\ln(\lambda_1)} \approx 3,0605 \dots$

$$\begin{aligned} \text{Es gilt: } \lambda_1^{\hat{\tau}} &= \lambda_1^{\frac{\ln |\lambda_2|}{\ln(\lambda_1)}} \\ &= e^{\ln(\lambda_1) \frac{\ln |\lambda_2|}{\ln(\lambda_1)}} \\ &= e^{\ln(\lambda_1) \frac{\ln |\lambda_2|}{\ln(\lambda_1)}} \\ &= e^{\ln |\lambda_2|} \\ &= |\lambda_2| \end{aligned}$$

$$\Rightarrow \lambda_1^{\tau} > |\lambda_2| \quad (\text{da } \lambda_1 \in (0, 1)) \quad (\ast)$$

$$\Rightarrow f(v^n) = f(\underbrace{\lambda_1^n}_{>0}, \lambda_2^n) = \theta \lambda_1^{n\tau} + \lambda_2^{2n} + \lambda_2^n \quad \text{und}$$

$$f(v^{n+1}) = f(\lambda_1^{n+1}, \lambda_2^{n+1}) = \theta \lambda_1^{(n+1)\tau} + \lambda_1^{2n+2} + \lambda_2^{n+1}$$

Reflektierter Punkt:

$$\hat{v}^n = \frac{1}{2}v^{n+1} + (-v^n + \frac{1}{2}v^{n+1}) = v^{n+1} - v^n$$

Der ‘vorletzte’ Punkt wird reflektiert an  $\frac{1}{2}(\bar{v} + v^{n+1})$

Dabei ist

$$\hat{v}^n = v^{n+1} - v^n = \begin{pmatrix} \lambda_1^{n+1} \\ \lambda_2^{n+1} \end{pmatrix} - \begin{pmatrix} \lambda_1^n \\ \lambda_2^n \end{pmatrix} = \begin{pmatrix} -\lambda_1^n(1 - \lambda_1) \\ -\lambda_2^n(1 - \lambda_2) \end{pmatrix} \quad (\ast\ast)$$

$$\Rightarrow f(\hat{v}^n) = \theta \underbrace{\phi \lambda_1^{n\tau} (1 - \lambda_1)^\tau}_{|x_1|^\tau \text{ (Def. von } f)} + \lambda_2^{2n}(1 - \lambda_2)^2 + \lambda_2^n(1 - \lambda_2)$$

$$\alpha)) \quad f(v^{n+1}) - f(v^n) = \theta \lambda_1^{n\tau} \underbrace{(\lambda_1^\tau - 1)}_{<0} + \underbrace{\lambda_2^{2n}}_{>0} \underbrace{(\lambda_2^2 - 1)}_{<0} + \lambda_2^n \underbrace{(\lambda_2 - 1)}_{<0} \stackrel{!}{<} 0$$

↑  
wechselndes Vorzeichen

(Damit  $v^n$  ‘als nächstes’ gespiegelt wird.)

Wegen  $\lambda_1^\tau > |\lambda_2|$  ist die Ungleichung erfüllt, falls

$$\theta(1 - \lambda_1^\tau) > |\lambda_2 - 1| \quad (\ast\ast\ast)$$

$$\beta)) \quad f(v^{n+1}) - f(\bar{v}) > \theta \lambda_1^{(n+1)\tau} + \lambda_2^{n+1} = \theta (\lambda_1^\tau)^{n+1} + \lambda_2^{n+1} \stackrel{\ast}{>} 0 \quad \text{für } \theta \geq 1.$$

$\gamma))$

$$\begin{aligned}
 & f(\hat{v}^n) \stackrel{!}{>} f(v^n) \\
 \Leftrightarrow^{**} & \theta \underbrace{\phi(\lambda_1^n(1-\lambda_1))^\tau}_{|x_1|} + \underbrace{\lambda_2^{2n}(1-\lambda_2)^2}_{x_2^2} - \underbrace{\lambda_2^n(1-\lambda_2)}_{x_2} > \theta \underbrace{\lambda_2^{n\tau}}_{x_1^\tau} + \underbrace{\lambda_2^{2n}}_{x_2^2} + \underbrace{\lambda_2^n}_{x_2} \\
 \Leftrightarrow & \theta \underbrace{\lambda_1^{n\tau}}_{\stackrel{*}{>}|\lambda_2|^n} (\underbrace{\phi(1-\lambda_1)^\tau}_{>0} - 1) > \underbrace{\lambda_2^{2n}}_{>0} (\underbrace{1 - (1-\lambda_2)^2}_{<0}) + \lambda_2^n(1 + (1-\lambda_2))
 \end{aligned}$$

Hinreichend:

$$\theta(\phi(1-\lambda_1)^\tau - 1) > 2 - \lambda_2 \quad (***)$$

Werte, die  $***$  und  $****$  erfüllen:

- $\tau = 2, \theta = 6, \phi = 60$     "gleichmäßig streng konvex"
- $\tau = 3, \theta = 6, \phi = 400$     "2-mal stetig diff'bar & konvex".

Aber  $\nabla f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ .    #

**Einschub:**

Simplex mit Ecken  $v^i = \frac{e_i}{\sqrt{2}}$  ( $1 \leq i \leq n$ ) und  $v^{n+1} = -\beta e$ ,  
wobei  $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$  und  $e = (1, \dots, 1)^T$ .

$$\Rightarrow \|v^i - v^j\|_2^2 = \frac{1}{2} + \frac{1}{2} = 1 \quad \text{für } 1 \leq i \leq j \leq n$$

$$\|v^i - v^{n+1}\|_2^2 = \underbrace{\left(\frac{1}{\sqrt{2}} + \beta\right)^2}_{\text{Komponente } i} + (n-1)\beta^2 \stackrel{!}{=} 1 \quad \text{für } i \leq n$$

$$\Rightarrow \frac{1}{2} + \sqrt{2}\beta + n\beta^2 = 1 \quad \Rightarrow \beta^2 + \frac{\sqrt{2}}{n}\beta - \frac{1}{2n} = 0$$

$$\Rightarrow \beta = -\frac{1}{\sqrt{2n}} + \sqrt{\frac{1}{2n^2} + \frac{1}{2n}} = \frac{1}{\sqrt{2n}}(-1 + \sqrt{1+n}) \quad (\text{ganz grob } \approx \frac{1}{\sqrt{2n}})$$

wähle positive Wurzel

$$\text{Probe: } \|v^i - v^{n+1}\|_2^2 = \left(\frac{1}{\sqrt{2}}\left(1 + \frac{1}{n}(-1 + \sqrt{1+n})\right)\right)^2 + (n-1)\frac{1}{2n^2}\underbrace{(-1 + \sqrt{1+n})^2}_{=:\rho}$$

$$\rho^2 = 1 - 2\sqrt{1+n} + 1 + n = 2 + n - 2\sqrt{n+1} = \underbrace{n - 2\rho}$$

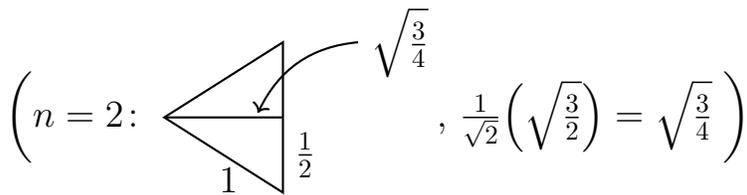
$$\begin{aligned} \|v^i - v^{n+1}\|_2^2 &= \frac{1}{2}\left(1 + \frac{\rho}{n}\right)^2 + \frac{n-1}{2n^2}\rho^2 \\ &= \frac{1}{2} + \frac{\rho}{n} + \frac{\rho^2}{2n^2} + \frac{n-1}{2n^2}\rho^2 \\ &= \frac{1}{2} + \frac{\rho}{n} + \frac{n}{2n^2}\rho^2 \\ &= \frac{1}{2} + \frac{\rho}{n} + \frac{1}{2n}\underbrace{(n - 2\rho)} \\ &= \frac{1}{2} + \frac{\rho}{n} + \frac{1}{2} - \frac{\rho}{n} \\ &= 1 \quad \# \end{aligned}$$

$$\text{Dieser Simplex erfüllt: } \left(\sum_{i=1}^n v^i\right) + v^{n+1} = \frac{1}{\sqrt{2}}e - \beta e = \underbrace{\frac{1}{\sqrt{2}}\left(1 - \frac{\sqrt{n+1}-1}{n}\right)}_{=:\mu} e$$

Bilde  $v^i \rightarrow v^i - \frac{\mu e}{n+1}$  um einen Simplex mit Schwerpunkt 0 zu erhalten.  $\#$

Höhe des Simplex (mit Kantenlänge 1)

$$\begin{aligned}
 \underbrace{\left\| \frac{1}{n} \sum_{i=0}^{n-1} v^i - \overbrace{v^n}^{\text{gegenüberliegende Ecke}} \right\|_2}_{\text{Schwerpunkt } v^0, \dots, v^{n-1}} &= \left\| \frac{1}{\sqrt{2n}} e + \beta e \right\|_2 \\
 &= \|e\| \left( \frac{1}{\sqrt{2n}} + \frac{1}{\sqrt{2n}} (-1 + \sqrt{n+1}) \right) \\
 &= \sqrt{n} \left( \frac{\sqrt{n+1}}{\sqrt{2n}} \right) \\
 &= \frac{1}{\sqrt{2}} \left( \sqrt{\frac{n+1}{n}} \right)
 \end{aligned}$$

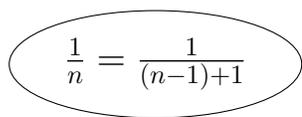


Der Schnitt der Geraden durch ein  $v^i$  und den Schwerpunkt mit dem Simplex hat die Länge  $\sqrt{\frac{n+1}{2n}}$ .

Der Schwerpunkt teilt diese Strecke im Verhältnis  $1 : n$  auf.

Auf S. (20) ¶ geschieht dies in Dimension  $n - 1$  sodass die Größe  $\theta(\delta)$  dort den

exakten Wert  $\delta \cdot \sqrt{\frac{n}{2(n-1)}} \cdot \frac{1}{n} = \frac{\delta}{\sqrt{2n(n-1)}}$  hat.

$\left( \approx \frac{\delta}{\sqrt{2n}} \right)$   

 $\frac{1}{n} = \frac{1}{(n-1)+1}$

---

¶ Konvergenz des Simplexverfahren von Spendley et al.

## §6 Methoden der konjugierten Richtungen

Sei  $A = A^T \in \mathbb{R}^{n \times n}$  pos. definit,  $A \succ 0$  ( $\succ$  bedeutet positiv definit; Löwner Halbordnung).

Eigenwertzerlegung:  $A = UDU^T$  mit  $D = (\diagdown)$  und  $U^T U = U U^T = I$ .

Sei  $A^{\frac{1}{2}} := U D^{\frac{1}{2}} U^T \Rightarrow A^{\frac{1}{2}} \succ 0$  und  $(A^{\frac{1}{2}})^2 = A$ .

( $A^{\frac{1}{2}}$  hängt nicht von der speziellen Wahl von  $U$  ab.)

**Def.:** Richtungen  $d^1, \dots, d^k$  heißen  $A$ -konjugiert, falls  $d^i \neq 0$  für  $1 \leq i \leq k$  und  $(d^i)^T A d^j = 0$  für  $i \neq j$ .

**Bem. 1:**  $A$ -konjugierte Richtungen sind linear unabhängig.

**Bew.:**

$$\begin{aligned} \text{Sei } z &:= \sum_{i=1}^k \mu_i d^i \\ \Rightarrow z^T A z &= \left( \sum_{i=1}^k \mu_i d^i \right)^T A \left( \sum_{j=1}^k \mu_j d^j \right) \\ &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j (d^i)^T A d^j \\ &= \sum_{i=1}^k \mu_i^2 \underbrace{(d^i)^T A d^i}_{>0} \\ &> 0 \end{aligned}$$

falls  $\mu \neq 0$ .  $\#$

Sei  $c \in \mathbb{R}^n$  und  $A \succ 0$  und  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $f(x) \equiv \frac{1}{2} x^T A x + c^T x$ .

Gegeben  $x, d \in \mathbb{R}^n$  finde  $\alpha^* := \arg \min f(x + \alpha d)$ :

$\ell(\alpha) := f(x + \alpha d)$  eine konvexe quadratische Funktion

$$\begin{aligned} \ell'(\alpha) &= \nabla f(x + \alpha d)^T d \\ &= (A(x + \alpha d) + c)^T d \\ &\stackrel{!}{=} 0 \end{aligned}$$

d. h.  $\alpha^* = \frac{-x^T A d - c^T d}{d^T A d}$ .

Übung: Man gebe ein einfaches Verfahren an, um  $\alpha^*$  anhand von  $\ell(\alpha_1), \ell(\alpha_2)$  und  $\ell(\alpha_3)$  mit  $\alpha_1 < \alpha_2 < \alpha_3$  zu bestimmen.

**Bem. 2:** Falls  $\alpha$  und  $\bar{d}$   $A$ -konjugiert sind, so ist

$$\underbrace{d^T \nabla f(x + \alpha \bar{d})}_{Ax+c+\alpha A\bar{d}} \equiv \text{const.} \quad \forall \alpha \in \mathbb{R} \quad \#$$

**Bem. 3:** Falls  $d^1, \dots, d^k$   $A$ -konjugiert sind und ausgehend von  $x^0$  die Punkte  $x^{j+1} = x^j + \alpha_j d^j$  mittels exakter Line-Search bestimmt werden, so gilt

$$(d^j)^T \nabla f(x^{k+1}) = 0 \quad \text{für } j = 1, \dots, k$$

**Bew.:**

$k = j$ :  $d^{jT} \nabla f(x^{j+1}) = d^{jT} \nabla f(x^j + \alpha_j d^j) = 0$  (exakte line Search)

Mit Bem. 2 folgt daraus und aus  $d^{jT} A d^{j+1} = 0$  auch  $d^{jT} \nabla f(x^{j+2}) = 0$  und dito für  $h \geq j + 1$ .  $\#$

**Bem. 4:** Die Richtungen  $d^i$  sind  $A$ -konjugiert  $\Leftrightarrow A^{\frac{1}{2}} d^i$  sind orthogonal.  
 $d^{iT} A d^j = (A^{\frac{1}{2}} d^i)^T (A^{\frac{1}{2}} d^j)$   $\#$

Wenn also  $\nabla f(x^{k+1}) \neq 0$ , so können  $d^1, \dots, d^k$  nicht den ganzen Raum aufspannen (Bem. 3) und dann ex. (Bem. 4) ein  $d^{k+1} \neq 0$  das zu  $d^1, \dots, d^k$   $A$ -konjugiert ist.

**Bem. 5:** Wenn  $\bar{x}$  und  $\tilde{x}$  ( $\bar{x} \neq \tilde{x}$ ) und  $d \neq 0$  die Beziehung  $d^T \nabla f(\bar{x}) = d^T \nabla f(\tilde{x})$  erfüllen, so ist  $\bar{d} := \bar{x} - \tilde{x}$   $A$ -konjugiert zu  $d$ :

$$0 = d^T \nabla f(\bar{x}) - d^T \nabla f(\tilde{x}) = d^T (A\bar{x} + c) - d^T (A\tilde{x} + c) = d^T A(\bar{x} - \tilde{x}). \quad \#$$

Seien nun  $d^1, \dots, d^\ell$   $A$ -konjugiert und  $x^{\ell+1}$  durch sukzessive Minimierung bez.  $d^1, \dots, d^\ell$  gegeben.

Nach Bem. 3 folgt  $d^{jT} \nabla f(x^{\ell+1}) = 0$  für  $1 \leq j \leq \ell$ .

Sei  $s^{\ell+1}$  lin. unabh. von  $d^1, \dots, d^\ell$ . (iterativ z.B.  $s^i = e_i$  die Einheitsvektoren.) Setze  $\rho_{\ell+1} := \arg \min\{\rho \mid f(x^{\ell+1} + \rho s^{\ell+1})\}$  und  $\hat{x}^{\ell+1} := x^{\ell+1} + \rho_{\ell+1} s^{\ell+1}$ .

Minimiere  $f$  ausgehend von  $\hat{x}^{\ell+1}$  sukzessiv entlang  $d^1, \dots, d^\ell$ . Um  $\tilde{x}^{\ell+1}$  zu erhalten.

Nach Bem. 3 gilt dann  $(d^j)^T \nabla f(\tilde{x}^{\ell+1}) = 0. \quad j = 1, \dots, \ell$

Ebenso  $(d^j)^T \nabla f(x^{\ell+1}) = 0 \quad j = 1, \dots, \ell$  (s.o.)

Nach Bem. 5 ist  $d^{\ell+1} := \tilde{x}^{\ell+1} - x^{\ell+1}$  auch  $A$ -konjugiert zu  $d^1, \dots, d^\ell$ .

Löse  $f(\underbrace{\tilde{x}^{\ell+1} + \tilde{\rho}d^{\ell+1}}_{=:x^{\ell+2}}) \longrightarrow \min$

(Dies ist äquivalent zu  $f(\underbrace{x^{\ell+1} + \tilde{\rho}d^{\ell+1}}_{=:x^{\ell+2}}) \longrightarrow \min.$ )  $\#$

Jede Line Search benötigt bei quadrat. Fkt. 3 Funktionsauswertungen, sodass der Gesamtaufwand für  $n$  Schritte  $\frac{3}{2}(n^2 + 5n - 4)$  Funktionsauswertungen beträgt.

- Anwendung auch auf nicht-quadratische Funktionen.
- In der Regel ist die Berechnung eines Gradienten genauso "teuer" wie die Funktionsauswertung, sodass dieser Ansatz bei quadratischen Funktionen nicht effizient ist.

## §7 Lineare Approximationsmodelle

$v^0, \dots, v^n$  gegeben mit  $f(v^0) \leq \dots \leq f(v^n)$ . Finde eine affin-lineare Funktion

$$s \mapsto a^T x + \beta \quad \text{mit} \quad a^T v^i + \beta = f(v^i) \quad (0 \leq i \leq n)$$

$$\Leftrightarrow \underbrace{\begin{pmatrix} (v^0)^T & 1 \\ \vdots & \vdots \\ (v^n)^T & 1 \end{pmatrix}}_{=: A} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} f(v^0) \\ \vdots \\ f(v^n) \end{pmatrix} \quad \text{d. h.} \quad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = A^{-1} \begin{pmatrix} f(v^0) \\ \vdots \\ f(v^n) \end{pmatrix}$$

Wähle  $\rho > 0$  "Trust-region-radius"

Setze  $\bar{v} := v^0 - \rho \frac{a}{\|a\|_2}$ . (Die Näherung  $f(v) \approx a^T v + \beta$  motiviert den Ansatz  $\nabla f(v) \approx a$ , sodass  $\bar{v}$  grob genähert ein Schritt des steilsten Abstiegs ist.)

i) Wenn  $f(\bar{v}) < f(v^0)$ , nutze die *Sharma-Morrisson-Formel* um ein  $v^i$  gegen  $\bar{v}$  auszutauschen:

$$(A + pq^T)^{-1} = A^{-1} - \frac{A^{-1}pq^T A^{-1}}{1 + q^T A^{-1}p} \quad \text{mit} \quad p = e_i, \quad q = \begin{pmatrix} \bar{v} - v^i \\ 0 \end{pmatrix}$$

Wähle  $i$  "groß" d.h.  $f(v^i)$  "groß" aber so, dass

$$|1 + q^T A^{-1}p| = |1 + ((\bar{v} - v^i)^T, 0) A^{-1} e_i|$$

ebenfalls nicht "zu klein" ist. Wenn

$$|f(\bar{v}) - (a^T \bar{v} + \beta)| \leq \frac{1}{2} a^T (v^0 - \bar{v}),$$

so vergrößere  $\rho$  (zusätzlich).

ii) Wenn  $f(\bar{v}) \geq f(v^0)$

- Wenn  $\|A\|_F \|A^{-1}\|_F > \kappa n$  für ein vorgegebenes  $\kappa > 2$ , so regularisiere den Simplex:  $\delta := \min\{\min_{1 \leq i \leq n} \|v^0 - v^i\|_2, \rho\}$  z. B. neuer regelmäßiger Simplex mit Kantenlänge  $\delta$ .
- Andernfalls reduziere  $\rho$ .

Erweiterung auf Probleme mit Ungleichungsrestriktionen Powell 1994.

Erweiterung auf quadratische interpolierende Funktionen

"Übung" Powell 2009. "Uobyqa" (deutliche schneller als Nelder Mead)