In der Bed. ⊛ ist eine "line search" entlang einer Koordinatenachse erlaubt, im Gegensatz zum Bsp. von Powell (1973), S. 2 ist dieser Ansatz aber in jedem Fall endlich.

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

Verfahren von Hooke & Jeeves (1961)

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

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

und falls $\min\limits_{z \in \mathcal{N}(y^k)} f(z) < f(x^k)$ so setze

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

andernfalls definiere x^{k+1} wie in ⊛ oben (S. 11).

Konvergenzaussage:

Sei $f: \mathbb{R}^n \to \mathbb{R}$ diff'bar und ∇f L-Lipschitz-stetig,

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

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

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

Bew: ÜA #

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

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

$$\begin{cases} |\nabla f(x^k)_i| \le L\bar{h} & \text{falls } x_i^k \in (l_i, u_i) \\ \nabla f(x^k)_i \ge -L\bar{h} & \text{"} \quad x_i^k = l_i \\ \nabla f(x^k)_i \le L\bar{h} & \text{"} \quad x_i^k = u_i \end{cases}.$$

Bew.

Dass das Verfahren abbricht folgt, da es nur endlich viele Punkte in 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: $f(\bar{x} \pm h_i e_i) \ge f(\bar{x})$ falls $l_i < \bar{x}_i < u_i$

Betrachte $\ell(t) := f(\bar{x} + t e_i)$. Nach dem Mittelwertsatz ist

$$f(\bar{x}) = \ell(0) \le \ell(h_i) = \ell(0) + h_i \ell'(\underbrace{t_+}_{\in [0, h_i]}) = f(\bar{x}) + \frac{\partial}{\partial x_i} f(\underbrace{\bar{x} + t_+ e_i}_{=: \xi_{i_+} \in [\bar{x}, \bar{x} + h_i e_i]}) h_i .$$ Analog

$$f(\bar{x}) \le f(\bar{x}) - \frac{\partial}{\partial x_i} f(\xi_{i_-}) h_i \quad \text{mit } \xi_{i_-} \in [\bar{x} - h_i e_i, \bar{x}]$$

$$\Rightarrow \frac{\partial}{\partial x_i} f(\xi_{i_+}) \ge 0 \ge \frac{\partial}{\partial x_i} f(\xi_{i_-})$$

$$\Rightarrow \frac{\partial}{\partial x_i} f(\bar{x}) \ge \frac{\partial}{\partial x_i} f(\xi_{i_+}) - L \underbrace{\|\bar{x} - \xi_{i_+}\|_\infty}_{\le h_i \le \bar{h}} \ge 0 - L\bar{h}$$

$$\frac{\partial}{\partial x_i} f(\bar{x}) \le \frac{\partial}{\partial x_i} f(\xi_{i_-}) + L \|\bar{x} - \xi_{i_-}\|_\infty \le L\bar{h}.$$

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

#

Praktische Umsetzung:

- ggf. Skalierung auf $-1 \leq x_i \leq 1$ [a]
- Anwendung des Verfahrens auf einem <u>groben</u> <u>Gitter</u> G, bei Abbruch <u>Verfeinerung</u> von G und erneute Anwendung des Verfahrens.
- Modifizierung falls $l_i = -\infty$ oder $u_i = \infty$: (unbeschränktes Gitter, Abbruch nach "maxit" Schritten)

---

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

- Falls $l_i = -\infty$, $u_i = \infty$ $\forall i$ (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 $D = \left(\diagdown\right)$ betrachten) und <u>$h := 1$</u> zum Start.
  - wähle $\tau \in (0,1) \cap \mathbb{Q}$, $\tau = \frac{\alpha}{\beta}$, $\alpha, \beta \in \mathbb{N}$, $(\alpha < \beta)$.
  - wenn bei ⊛ (s. 11) gilt $f(x^{k+1}) = f(x^k)$ so setze $h := \tau h$, andernfalls $h := \frac{h}{\tau^{j_k}}$ mit $j_k \in \{0, 1, \dots, l\}$ für ein festes $l \in \mathbb{N}$.

<u>Beh.</u> Falls ∇f auf S_1 Lipschitz-stetig ist, so gilt

$$\liminf_{k \to \infty} \|\nabla f(x^k)\| = 0.$$

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

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

Annahme: $\exists\, \bar{m} \in \mathbb{Z} : h^{(k)} \in \{\tau^j \mid j \le \bar{m},\ j \in \mathbb{Z}\}$.

$\Rightarrow x^k$ hat stets die Form, dass die Komponenten gegeben sind durch $x_i^k = x_i^1 + \sum_{\underline{m} \le j \le \bar{m}} z_{ij} \tau^j$, wobei die $z_{ij} \in \mathbb{Z}$ von k abhängen. Die Potenzen τ^j sind stets Vielfache von $\alpha^{\underline{m}} \beta^{-\bar{m}}$. Seien l, u gegeben mit $S_1 \subset [l, u]$. Da es in $[l_i, u_i]$ nur endlich viele Punkte aus $x_i^1 + \lambda \alpha^{\underline{m}} \beta^{-\bar{m}}$ mit $\lambda \in \mathbb{Z}$ gibt bricht das Verfahren nach endlich vielen Schritten ab. ↯

Also ist $\liminf_{k \to \infty} h^{(k)} = 0$. Jedes Mal, wenn $h^{(k)}$ reduziert wird, gilt $\|\nabla f(x^k)\|_\infty \le L\, h^{(k)}$. $\Rightarrow$ Beh.

#

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