

§ 2 Line-search-Verfahren

①

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: $\min \Leftrightarrow \max$)

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

Allgemeines line-search-Verfahren:

Für $k = 0, 1, 2, \dots$

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

" " Schrittweite $\alpha_k \in \mathbb{R}$ so dass

$$f(x^k + \alpha_k s^k) < f(x^k)$$

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

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: $\exists c \in (0, 1)$:

$$\forall k \geq l \quad \forall v \in \mathbb{R}^n: \max \left\{ \frac{|v^T s^j|}{\|s^j\|_2} \mid k-l+1 \leq j \leq k \right\} \geq c \|v\|_2$$

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:

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

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)

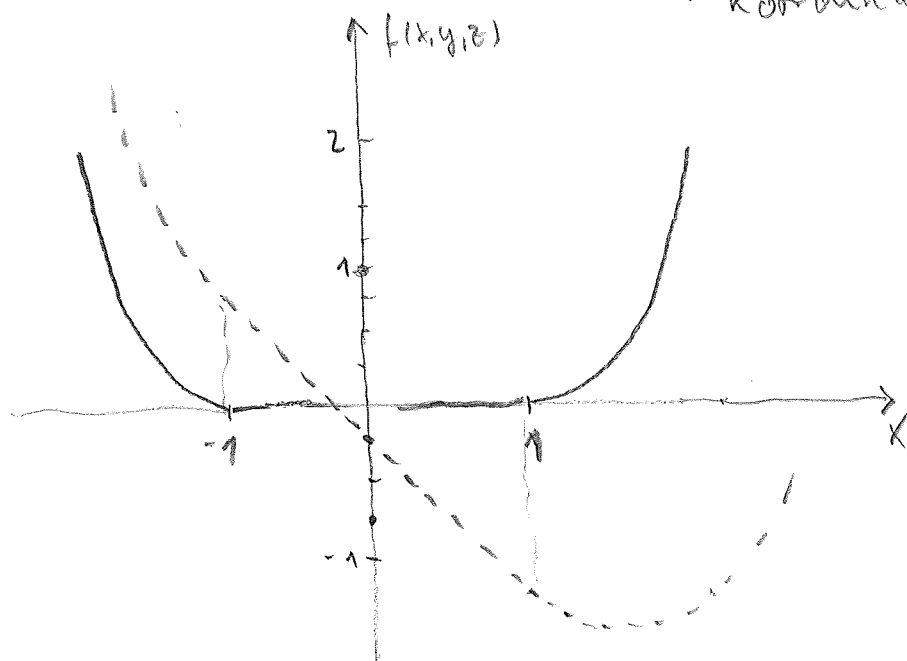
$$f(x, y, z) := -xy - xz - yz + (x-1)_+^2 + (-x-1)_+^2 \\ + (y-1)_+^2 + (-y-1)_+^2 + (z-1)_+^2 + (-z-1)_+^2$$

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

Skizze für $y = z = 0$:

$y = z = 1/2$:

|| konvex parallel zu den Koordinatenachsen.



③

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

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

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

$$\Rightarrow \text{minimiere } -x(y+z) + (x-1)^2$$

$$\Leftrightarrow -y+z + 2(x-1) = 0$$

$$\Leftrightarrow x = \frac{y+z}{2} + 1$$

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 $(-1-\varepsilon, 1+\frac{\varepsilon}{2}, -1-\frac{\varepsilon}{4}) =: v^{(0)}$

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

$$v^{(1)} = (1 + \frac{\varepsilon}{8}, 1 + \frac{\varepsilon}{2}, -1 - \frac{\varepsilon}{4})$$

$$v^{(2)} = (1 + \frac{\varepsilon}{8}, -1 - \frac{\varepsilon}{16}, -1 - \frac{\varepsilon}{4})$$

$$v^{(3)} = (1 + \frac{\varepsilon}{8}, -1 - \frac{\varepsilon}{16}, 1 + \frac{\varepsilon}{32})$$

$$v^{(4)} = (-1 - \frac{\varepsilon}{64}, -1 - \frac{\varepsilon}{16}, 1 + \frac{\varepsilon}{32})$$

$$v^{(5)} = (-1 - \frac{\varepsilon}{64}, 1 + \frac{\varepsilon}{128}, 1 + \frac{\varepsilon}{32})$$

$$v^{(6)} = (-1 - \frac{\varepsilon}{64}, 1 + \frac{\varepsilon}{128}, -1 - \frac{\varepsilon}{256})$$

$\equiv v^{(0)}$ mit $\frac{\varepsilon}{64}$ statt ε .

(4)

Zykeln "entlang" von 6 Kanten des Würfels mit den Ecken $(\pm 1, \pm 1, \pm 1)$.

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

und $\nabla f\left(\begin{smallmatrix} -1 \\ 1 \\ -1 \end{smallmatrix}\right) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, dito $\nabla f(v) \neq 0$ für die anderen Ecken.

Modifiziertes Verfahren (Lucidi Sciandrone 1997):

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

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 werde 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.