

Bem 3

(31)

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

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

Mit Bem 2. folgt daraus und aus $d^j{}^T A d^{j+1} = 0$

auch $d^j{}^T \nabla f(x^{j+2}) = 0$ 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

$$d^i{}^T A d^j = (A^{\frac{1}{2}} d^i)^T (A^{\frac{1}{2}} d^j) \quad \#$$

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

$$d^T \nabla f(\bar{x}) = d^T \nabla f(\tilde{x}) \quad \text{erfüllen, so ist } \bar{d} := \bar{x} - \tilde{x} \quad \text{A-konjugiert zu } d:$$

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

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

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

Sei s^{l+1} lin. unabh. von $d^1, \dots, d^l$ (illustrativ z.B. $s^i = e_i$ die Einheitsvekt.)

Setze $s_{l+1} := \operatorname{argmin} \{ \varphi | f(x + \varphi s^{l+1}) \}$ und $\hat{x}^{l+1} := x + s_{l+1} s^{l+1}$.

Minimiere f ausgehend von $\hat{x}^{l+1}$ sukzessive entlang $d^1, \dots, d^l$ um $\tilde{x}^{l+1}$ zu erhalten.

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

Ebenso: $(d^j)^T \nabla f(x^{l+1}) = 0$ $j=1, \dots, l$ (s.v.)

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

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

Jede Line Search benötigt bei quadrat. Fkt 3 Funktionsauswertungen, sodass der Gesamtaufwand für n Schritte $\frac{3}{2}(n+3n-2)$ 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 $x \mapsto a^T x + \beta$
mit $a^T v^i + \beta = f(v^i) \quad (0 \leq i \leq n)$

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

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

Setze $\bar{v} := v^0 - \delta \frac{a}{\|a\|_2}$

ii) Wenn $f(\bar{v}) < f(v^0)$ nutze die Sherman-Morrison-Formel um ein v^i gegen $\bar{v}$ auszutauschen:

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

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 δ (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: $S := \min \left\{ \min_{1 \leq i \leq n} \|v^0 - v^i\|_2, \delta \right\}$
z.B. neuer regulärer Simplex mit Kantenlänge δ .
- Andernfalls reduziere δ .

Erweiterung auf Probleme mit Ungleichungsrestriktionen
Powell 1994.

Erweiterung auf quadratische interpolierende Funktionen
"Übung" Powell 2009. "Uobyqa"

(deutlich schneller als Nelder-Mead)