

Calcul Numeric

Cursul 12

2020

Anca Ignat

Metoda pantei maxime

$$\min\{ f(x) ; x \in \mathbb{R}^n \} , \quad f : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$x_0 \text{ dat} , \quad x_{k+1} = x_k + \alpha_k d_k , \quad k = 0, 1, \dots$$

d_k direcție de descreștere a lui f în x_k

$$\alpha_k > 0 \quad f(x_k + \alpha_k d_k) = \min\{ f(x_k + \alpha d_k) ; \alpha \in [0, \bar{\alpha}] \}$$

d_k direcție de descreștere dacă :

$$\nabla f(x_k)^T d_k < 0.$$

Metoda pantei maxime:

$$d_k = -\nabla f(x_k)$$

Cazul pătratic:

$$f(x) = \frac{1}{2}x^T A x - x^T b = \frac{1}{2}(Ax, x)_{\mathbb{R}^n} - (b, x)_{\mathbb{R}^n}$$

$b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$ simetrică și pozitiv definită

A pozitiv definită $\rightarrow \det A \neq 0$, f este strict convexă

$$\nabla f(x) = Ax - b \quad , \quad \nabla^2 f(x) = A$$

$$f(x^*) = \min\{f(x); x \in \mathbb{R}^n\} \quad x^* \text{ punct unic de minim} \Leftrightarrow$$

$$x^* \text{ soluția sistemului liniar } Ax = b, \quad x^* = A^{-1}b$$

$$g(x) = Ax - b$$

$$x_{k+1} = x_k - \alpha_k g_k, \quad g_k = Ax_k - b$$

$$\alpha_k = \operatorname{argmin}\{f(x_k - \alpha g_k); \alpha \in [0, \bar{\alpha}]\}$$

$$\begin{aligned}
f(x_k - \alpha g_k) &= \frac{1}{2}(x_k - \alpha g_k)^T A(x_k - \alpha g_k) - (x_k - \alpha g_k)^T b = \\
&= \frac{1}{2}(g_k^T A g_k) \alpha^2 - (g_k^T A x_k - g_k^T b) \alpha + f(x_k) = \\
&= \frac{1}{2}(g_k^T A g_k) \alpha^2 - (g_k^T g_k) \alpha + f(x_k)
\end{aligned}$$

$f(x_k - \alpha g_k)$ ecuație de gr. 2 în α , coef. lui α^2 , $g_k^T A g_k > 0$

$$\alpha_{\min} = \alpha_k = \frac{g_k^T g_k}{g_k^T A g_k}$$

Metoda pantei maxime pentru funcționale pătratice:

$$\mathbf{x}_0 - \text{dat}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left(\frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_k^T \mathbf{A} \mathbf{g}_k} \right) \mathbf{g}_k \quad , \quad k = 0, 1, \dots$$

$$\mathbf{g}_k = \mathbf{A} \mathbf{x}_k - \mathbf{b}$$

$$E(x) = \frac{1}{2}(x - x^*)^T A(x - x^*) = f(x) + \frac{1}{2}x^{*T} Ax^*$$

$$\nabla E(x) = \nabla f(x) = g(x)$$

Șirul construit cu metoda pantei maxime satisface:

$$E(x_{k+1}) = \left[1 - \frac{\left(g_k^T g_k \right)^2}{\left(g_k^T A g_k \right) \left(g_k^T A^{-1} g_k \right)} \right] E(x_k)$$

Inegalitatea lui Kantorovich

$A \in \mathbb{R}^{n \times n}$, $A = A^T$, $A > \mathbf{0}$ pozitiv definită

$$\frac{(x^T x)^2}{(x^T A x)(x^T A^{-1} x)} \geq \frac{4cC}{(c + C)^2}$$

c, C - cea mai mica și cea mai mare valoare proprie a lui A .

Teoremă

Cazul pătratic: Pentru orice iterație inițială \mathbf{x}_0 , șirul construit cu metoda pantei maxime converge la \mathbf{x}^* unicul punct de minim al funcției f . Avem:

$$E(\mathbf{x}_{k+1}) \leq \left(\frac{C - c}{C + c} \right)^2 E(\mathbf{x}_k)$$

Cazul general: $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2(\mathbb{R}^n)$, f are un punct de minim local \mathbf{x}^* . Presupunem că $F(\mathbf{x}^*) = \nabla^2 f(\mathbf{x}^*)$ are $c > 0$ cea mai mică valoare proprie și $C > 0$ cea mai mare valoare proprie. Dacă șirul $\{\mathbf{x}_k\}$ construit cu metoda pantei maxime converge la \mathbf{x}^* (la fiecare pas $A = \nabla^2 f(\mathbf{x}_k)$, $\mathbf{b} = \nabla f(\mathbf{x}_k)$), $\mathbf{x}_k \rightarrow \mathbf{x}^*$

atunci $f(\mathbf{x}_k) \rightarrow f(\mathbf{x}^*)$ converge liniar cu o rată de convergență $\leq \left(\frac{C-c}{C+c}\right)^2$.

Metoda Newton

\mathbf{x}_k - punctul curent de aproximare

$$f(\mathbf{x}) \simeq f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)^T (\mathbf{x} - \mathbf{x}_k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_k)^T \nabla^2 f(\mathbf{x}_k) (\mathbf{x} - \mathbf{x}_k) = g(\mathbf{y})$$

g – funcțională pătratică, $\mathbf{y} = \mathbf{x} - \mathbf{x}_k$,

$$g(y) = \frac{1}{2} y^T A y - y^T b + c$$

$$A = \nabla^2 f(x_k) \quad , \quad b = -\nabla f(x_k) \quad , \quad c = f(x_k)$$

$y^* = \arg \min \{g(y); y \in \mathbb{R}^n\}$ este unica soluție a sistemului liniar $Ay = b$, $y^* = A^{-1}b = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$

Metoda Newton

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) \quad , \quad k = 0, 1, \dots \quad , \quad x_0 \text{ - dat}$$

Teoremă

Fie $f \in C^3(\mathbb{R}^n)$ care are un punct de minim local \mathbf{x}^* astfel ca matricea $\nabla^2 f(\mathbf{x}^*) > \mathbf{0}$ este pozitiv definită. Dacă punctul de început \mathbf{x}_0 este suficient de aproape de \mathbf{x}^* , $\|\mathbf{x} - \mathbf{x}^*\| \leq r$, atunci șirul $\{\mathbf{x}_k\}$ generat cu metoda lui Newton converge la \mathbf{x}^* și ordinul de convergență este cel puțin 2.

Metoda gradientilor conjugați (a direcțiilor conjugate)

$$A \in \mathbb{R}^{n \times n}, \quad A = A^T, \quad A > 0 \text{ pozitiv definită}$$

$$\min \{ f(x) = \frac{1}{2} x^T A x - x^T b; x \in \mathbb{R}^n \}$$

Definiție

Pentru o matrice $A \in \mathbb{R}^{n \times n}$, $A = A^T$ simetrică, doi vectori $d_1, d_2 \in \mathbb{R}^n$ se numesc **A-ortogonali** sau **conjugați în raport cu A** dacă:

$$d_1^T A d_2 = (A d_2, d_1)_{\mathbb{R}^n} = (d_2, A d_1)_{\mathbb{R}^n} = (A d_1, d_2)_{\mathbb{R}^n} = 0$$

$A = \mathbf{0}_{n \times n} \Rightarrow \forall d_1, d_2$ sunt A – ortogonali

$A = I_n \Rightarrow$ ortogonalitate clasică $(d_1, d_2)_{\mathbb{R}^n} = \mathbf{0}$

Vectorii $\{d_0, d_1, \dots, d_k\}$ se numesc A -ortogonali sau A -conjugăți dacă:

$$d_i^T A d_j = (A d_j, d_i)_{\mathbb{R}^n} = \mathbf{0} , \quad \forall i \neq j , i, j = 0, 1, \dots, k$$

Propoziție

Fie $A \in \mathbb{R}^{n \times n}$, $A = A^T$, $A > 0$, și $\{d_0, d_1, \dots, d_k\}$ direcții A -conjugate, $d_i \neq 0$, $\forall i = 0, \dots, k$. Atunci vectorii $\{d_0, d_1, \dots, d_k\}$ sunt liniar independenți.

$$\left. \begin{array}{l} A \in \mathbb{R}^{n \times n}, \quad A = A^T, \quad A > 0 \\ \{d_0, d_1, \dots, d_{n-1}\} \text{ - direcții } A \text{ - conjugate} \end{array} \right\} \Rightarrow \begin{array}{l} \{d_0, d_1, \dots, d_{n-1}\} \\ \text{bază în } \mathbb{R}^n \end{array}$$

$$x^* = \operatorname{argmin}\{f(x); x \in \mathbb{R}^n\} \Leftrightarrow x^* \text{ soluția sist. } Ax = b$$

$$x^* = \alpha_0 d_0 + \alpha_1 d_1 + \cdots + \alpha_{n-1} d_{n-1}$$

$$\left(\underbrace{Ax^*}_{=b}, d_i \right)_{\mathbb{R}^n} = d_i^T Ax^* = \alpha_i d_i^T Ad_i$$

$$\alpha_i = \frac{d_i^T b}{d_i^T Ad_i} = \frac{(b, d_i)_{\mathbb{R}^n}}{(Ad_i, d_i)_{\mathbb{R}^n}}$$

$$x^* = \sum_{i=0}^{n-1} \frac{(b, d_i)_{\mathbb{R}^n}}{(Ad_i, d_i)_{\mathbb{R}^n}} d_i$$

Considerăm procesul iterativ

$$\mathbf{x}_0 \in \mathbb{R}^n - \text{dat}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad k = 0, 1, 2, \dots, n-1$$

$$\alpha_k = -\frac{\mathbf{g}_k^T \mathbf{d}_k}{\mathbf{d}_k^T \mathbf{A} \mathbf{d}_k}, \quad \mathbf{g}_k = \mathbf{A} \mathbf{x}_k - \mathbf{b}$$

are proprietatea că $\mathbf{x}_n = \mathbf{x}^*$.

$$\mathbf{x}_k = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \dots + \alpha_{k-1} \mathbf{d}_{k-1}$$

$$\mathbf{x}^* = \mathbf{x}_0 + \beta_0 \mathbf{d}_0 + \beta_1 \mathbf{d}_1 + \dots + \beta_{n-1} \mathbf{d}_{n-1}$$

$$\alpha_i = \frac{d_i^T A(x_k - x_0)}{d_i^T A d_i} \quad , \quad \beta_i = \frac{d_i^T A(x^* - x_0)}{d_i^T A d_i}$$

$$d_k^T A(x_k - x_0) = 0$$

$$x^* - x_k = (\beta_0 - \alpha_0)d_0 + \cdots + (\beta_{k-1} - \alpha_{k-1})d_{k-1} + \beta_k d_k + \cdots + \beta_{n-1}d_{n-1}$$

Corolar

$$g_k^T d_i = 0 \quad \forall i < k$$

Algoritmul gradientilor conjugați

$$\mathbf{x}_0 \in \mathbb{R}^n, \mathbf{g}_0 = A\mathbf{x}_0 - \mathbf{b}$$

$$\mathbf{d}_0 = -\mathbf{g}_0 = \mathbf{b} - A\mathbf{x}_0$$

$$\alpha_k = -\frac{\mathbf{g}_k^T \mathbf{d}_k}{\mathbf{d}_k^T A \mathbf{d}_k}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$$

$$\mathbf{g}_{k+1} = A\mathbf{x}_{k+1} - \mathbf{b} \text{ sau } \mathbf{g}_{k+1} = \mathbf{g}_k + \alpha_k A \mathbf{d}_k$$

$$\beta_k = \frac{\mathbf{g}_{k+1}^T A \mathbf{d}_k}{\mathbf{d}_k^T A \mathbf{d}_k}$$

$$\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_k \mathbf{d}_k$$

$$y_1, y_2, \dots, y_p \in \mathbb{R}^n$$

$$\begin{aligned} \text{span}\{y_1, y_2, \dots, y_p\} &= \{y = a_1 y_1 + \dots + a_p y_p \in \mathbb{R}^n; a_i \in \mathbb{R}, i = 1, \dots, p\} \\ &= \text{subspațiul generat de vectorii } y_1, y_2, \dots, y_p \end{aligned}$$

Teoremă

Presupunem că $x_k \neq x^*$. Avem următoarele relații:

$$(1) \quad g_k^T g_i = 0 \quad \forall i = 0, 1, \dots, k-1$$

$$(2) \quad \text{span}\{g_0, g_1, \dots, g_k\} = \text{span}\{g_0, Ag_0, \dots, A^k g_0\}$$

$$(3) \quad \text{span}\{d_0, d_1, \dots, d_k\} = \text{span}\{g_0, Ag_0, \dots, A^k g_0\}$$

$$(4) \quad d_k^T A d_i = 0 \quad \forall i = 0, 1, \dots, k-1$$

Șirul $x_k \rightarrow x^*$ în cel mult n pași.

$\mathcal{K}(g_0, k) = \text{span}\{g_0, Ag_0, \dots, A^k g_0\}$ - se numește **subspațiu Krylov de grad k** pentru g_0 .

Forma practică a metodei gradientilor conjugați

$$x_0 \in \mathbb{R}^n - \text{dat}, \quad k = 0$$

$$g_0 = Ax_0 - b, \quad d_0 = -g_0$$

$$\text{while } (g_k \neq 0)$$

$$\left\{ \begin{array}{l} \alpha_k = \frac{g_k^T g_k}{d_k^T A d_k} \\ x_{k+1} = x_k + \alpha_k d_k \\ g_{k+1} = g_k + \alpha_k A d_k \\ \beta_k = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \\ d_{k+1} = -g_{k+1} + \beta_k d_k \\ k = k + 1; \end{array} \right.$$

Teoremă

Dacă matricea A are doar r valori proprii distincte, algoritmul gradientilor conjugați calculează soluția \mathbf{x}^* în cel mult r iterații.

Teoremă

Dacă $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ sunt valorile proprii ale matricei A atunci:

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_A^2 \leq \left(\frac{\lambda_{n-k} - \lambda_1}{\lambda_{n-k} + \lambda_1} \right)^2 \|\mathbf{x}_0 - \mathbf{x}^*\|_A^2$$

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_A^2 = 2E(\mathbf{x}_{k+1})$$

$$E(\mathbf{x}_{k+1}) \leq \left(\frac{\lambda_{n-k} - \lambda_1}{\lambda_{n-k} + \lambda_1} \right)^2 E(\mathbf{x}_0).$$

Metodele gradientilor conjugați neliniare

$$f(x) \simeq f(x_k) + \nabla f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k)$$

$$g_k \leftrightarrow \nabla f(x_k)$$

$$A \leftrightarrow \nabla^2 f(x_k)$$

Varianta pentru funcții oarecare a metodei gradientilor conjugați:

$$\mathbf{x}_0 \in \mathbb{R}^n - \text{dat}, \quad k = 0$$

$$\mathbf{g}_0 = \nabla f(\mathbf{x}_0), \quad \mathbf{d}_0 = -\mathbf{g}_0$$

$$\text{while}(\mathbf{g}_k \neq \mathbf{0})$$

$$\left\{ \begin{array}{l} \mathbf{A} = [\nabla^2 f(\mathbf{x}_k)] \\ \alpha_k = -\frac{\mathbf{g}_k^T \mathbf{d}_k}{\mathbf{d}_k^T \mathbf{A} \mathbf{d}_k} \\ \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \\ \mathbf{g}_{k+1} = \nabla f(\mathbf{x}_{k+1}) \\ \beta_k = \frac{\mathbf{g}_{k+1}^T \mathbf{A} \mathbf{d}_k}{\mathbf{d}_k^T \mathbf{A} \mathbf{d}_k} \\ \mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_k \mathbf{d}_k \\ k = k + 1; \end{array} \right.$$

Metoda Fletcher-Reeves

α_k se calculează folosind metoda ajustării pasului

$$\beta_k = \frac{\mathbf{g}_{k+1}^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{g}_k}$$

$\mathbf{x}_0 \in \mathbb{R}^n$ – dat,

$\mathbf{g}_0 = \nabla f(\mathbf{x}_0)$, $\mathbf{d}_0 = -\mathbf{g}_0$, $k = 0$

***while* ($\mathbf{g}_k \neq \mathbf{0}$)**

$$\left\{ \begin{array}{l} \alpha_k = \min\{f(\mathbf{x}_k + \alpha \mathbf{d}_k); \alpha \in [0, \bar{\alpha})\} \\ \text{(exact sau inexact cu testul lui Wolfe)} \\ \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \\ \mathbf{g}_{k+1} = \nabla f(\mathbf{x}_{k+1}) \\ \beta_k^{FR} = \frac{\mathbf{g}_{k+1}^T \mathbf{g}_k}{\mathbf{g}_k^T \mathbf{g}_k} \\ \mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_k^{FR} \mathbf{d}_k \\ k = k + 1; \end{array} \right.$$

Se pune problema dacă d_k sunt direcții de descreștere?

$$d_k = -g_k + \beta_{k-1} d_{k-1}$$

$$g_k^T d_k = -g_k^T g_k + \beta_{k-1} g_k^T d_{k-1}$$

Dacă se folosește ajustarea pasului exactă:

α_{k-1} este punct de minim local pentru f pe direcția d_{k-1} prin urmare $g_k^T d_{k-1} = 0$ ($g_k = \nabla f(x_k)$).

$\Rightarrow g_k^T d_k = -g_k^T g_k = -\|g_k\|_2^2 < 0 \Rightarrow d_k$ direcție de descreștere

Dacă se folosește ajustarea pasului inexactă am putea avea $\mathbf{g}_k^T \mathbf{d}_k > \mathbf{0}$ (\mathbf{d}_k direcție de creștere!!) dar folosind testul lui Wolfe deducem:

$$\begin{aligned} f(\mathbf{x}_k + \alpha_k \mathbf{d}_k) &\leq f(\mathbf{x}_k) + \varepsilon \alpha_k \mathbf{g}_k^T \mathbf{d}_k \\ |\nabla f(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k| &\leq (1 - \varepsilon) |\mathbf{g}_k^T \mathbf{d}_k| \\ \varepsilon \in \left(0, \frac{1}{2}\right) &\Rightarrow \mathbf{g}_k^T \mathbf{d}_k < \mathbf{0} \end{aligned}$$

Metoda Polak-Ribière

- variantă a metodei Fletcher-Reeves

$$\beta_k^{PR} = \frac{\mathbf{g}_{k+1}^T (\mathbf{g}_{k+1} - \mathbf{g}_k)}{\mathbf{g}_k^T \mathbf{g}_k}$$

Dacă se face ajustarea pasului inexactă cu testul lui Wolfe nu putem deduce că \mathbf{d}_k sunt direcții de descreștere.

Se folosește $\beta_k^+ = \max\{\beta_k^{PR}, 0\}$ și un test Wolfe adaptat pentru a obține \mathbf{d}_k direcții de descreștere.

Varianta Hestenes-Stiefel

$$\beta_k^{HS} = \frac{\mathbf{g}_{k+1}^T (\mathbf{g}_{k+1} - \mathbf{g}_k)}{(\mathbf{g}_{k+1} - \mathbf{g}_k)^T \mathbf{d}_k}$$

Precondiționare

Se consideră norma:

$$\|x\|_A = \sqrt{(Ax, x)_{\mathbb{R}^n}}$$

Evaluarea erorii în metoda pantei maxime:

$$\|x^{(k)} - x^*\|_A \leq \left(\frac{k(A) - 1}{k(A) + 1} \right)^k \|x^{(0)} - x^*\|_A$$

$$k(A) = \frac{\lambda_n}{\lambda_1} - \text{numărul de condiționare spectral}$$

$$0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \text{ valorile proprii ale matricii } A$$

Avem convergență rapidă dacă numărul de condiționare spectrală al matricei A este apropiat de 1 ($k(A) \geq 1$ întotdeauna).

Ideea preconditionării este de a transforma sistemul $Ax=b$ astfel încât să îmbunătățim proprietățile spectrale.

$$Ax = b \quad \Leftrightarrow \quad \tilde{A}x = \tilde{b} \quad , \quad \text{cu } k(\tilde{A}) \ll k(A)$$

Precondiționare

$$Ax = b \rightarrow M^{-1}Ax = M^{-1}b \text{ (la stânga)}$$

$$\rightarrow AM^{-1}y = b \text{ , } x = M^{-1}y \text{ (la dreapta)}$$

$$\rightarrow M_1^{-1}AM_2^{-1}y = M_1^{-1}b \text{ , } x = M_2^{-1}y \text{ (split) , } M = M_1M_2$$

cu M matrice nesingulară , M “ \approx ” A . Matricea M sau M^{-1} poartă numele de *matrice de precondiționare*.

Cum trebuie să alegem matricea M ?

- sistemul preconditionat ($\tilde{A}x = \tilde{b}$) să fie ușor de rezolvat (convergență rapidă)
- matricea de preconditionare să fie economic de construit și aplicat – ietrațiile să nu fie costisitor de construit

Matricea de preconditionare Jacobi

$$M = \text{diag}[a_{11}, a_{22}, \dots, a_{nn}]$$

Matrice de preconditionare SSOR

$$M = (D + L) D^{-1} (D + L)^T \quad (A = L + D + L^T)$$

$$M(\omega) = \frac{1}{2-\omega} \left(\frac{1}{\omega} D + L \right) \left(\frac{1}{\omega} D \right)^{-1} \left(\frac{1}{\omega} D + L \right)^T, \quad \omega \in (0, 2)$$

Pentru ω - optimal, în anumite cazuri:

$$k(M(\omega_{opt})^{-1} A) = O(\sqrt{k(A)})$$

(ω_{opt} - foarte costisitor de calculat)