Calcul Numeric

Cursul 11

2020

Anca Ignat

Rezolvarea ecuaţiilor neliniare

Fie $f:[a,b]\to\mathbb{R}$ o **funcţie continuă** pe intervalul **[*a*,*b*]** astfel ca ***f(a)f(b)*** < ***0***. În aceste condiţii, există $\boldsymbol{x^*\in(a,b)}$ astfel ca ***f(x*)=0***. În cele ce urmează ne propunem să aproximăm soluţia $\boldsymbol{x^*}$ a ecuaţiei neliniare ***f(x)=0***.

Metoda bisecţiei (a înjumătăţirii intervalului)

Pp. ***f(a)f(b)*** < ***0!!!*** Pentru a aproxima soluţia $\boldsymbol{x^*}$ căutată, vom construi un şir de intervale $\boldsymbol{\{[a_k,b_k];k\geq 0\}}$ ce satsifac:

$$x^* \in [a_k, b_k]$$

$$[a_{k+1}, b_{k+1}] \subset [a_k, b_k]$$

$$b_{k+1} - a_{k+1} = \frac{b_k - a_k}{2}$$

Pentru primul interval vom considera:

$$a_0 = a \ , \ b_0 = b \ , \ k = 0.$$

Considerăm punctul ***c*** de mijloc al intervalului $[a_k, b_k]$:

$$c = \frac{a_k + b_k}{2}$$

Avem următoarele 3 variante:

1. $f(c)=0$ - soluţia căutată este $x^*=c$, algoritmul se opreşte;

2. $f(a_k)f(c) < 0 \rightarrow$ soluţia se găseşte în intervalul (a_k,c), continuăm procedeul cu intervalul $[a_{k+1} = a_k \ , \ b_{k+1} = c]$;

3. $f(b_k)f(c) < 0 \rightarrow$ soluţia se găseşte în intervalul $x^* \in (c,b_k)$ procedeul continuă cu intervalul $[a_{k+1} = c \ , \ b_{k+1} = b_k]$.

Dat $\varepsilon > 0$ există un interval $\left[a_{\bar{k}}, b_{\bar{k}}\right]$ astfel ca $x^* \in \left(a_{\bar{k}}, b_{\bar{k}}\right)$ şi $b_{\bar{k}} - a_{\bar{k}} < \varepsilon$ ($\bar{k} > \log_2 \dfrac{b-a}{\varepsilon}$). Pentru ε suficient de mic atât $a_{\bar{k}}$ cât şi $b_{\bar{k}}$ pot fi considerate aproximări ale soluţiei x^* ($a_{\bar{k}} \approx x^*$ prin lipsă iar $b_{\bar{k}} \approx x^*$ prin adaos).

Metoda tangentei (Newton-Raphson)

Vom presupune că funcţia $f \in C^1[a,b]$ este derivabilă pe **[*a,b*]** cu derivata continuă în acest interval şi satisface relaţia ***f(a)f(b)*** < ***0***. Pentru a aproxima soluţia x^* a ecuaţiei ***f(x)=0*** vom construi un şir $\{x_k\} \subseteq \mathbb{R}$ care să conveargă la x^*, $x_k \to x^*$, pentru $k \to \infty$. Primul element din şir, x_0, considerăm că este dat. Următorul element din şir se construieşte ca fiind punctul de intersecţie al tangentei la graficul funcţiei ***f*** în punctul $(x_0, f(x_0))$ cu axa absciselor. Procedeul se repetă cu x_1 pentru a-l obţine pe x_2, ş.a.m.d.

$$x_1 = \mathrm{O}x \cap \textbf{tangenta la graficul funcţiei } f \textbf{ în punctul} \left(x_0, f(x_0)\right)$$

$$x_2 = \mathrm{O}x \cap \textbf{tangenta la graficul funcţiei } f \textbf{ în punctul} \left(x_1, f(x_1)\right)$$

$$\vdots$$

$$x_{k+1} = \mathrm{O}x \cap \textbf{tangenta la graficul fcţ. } f \textbf{ în pt.} \left(x_k, f(x_k)\right), k = 0,1,2,$$

Ecuaţia tangentei la graficul funcţiei ***f*** într-un punct ***(a, f(a))*** este următoarea (pentru o funcţie derivabilă):

$$y = f(a) + f'(a)(x - a)$$

Pentru a calcula x_{k+1} din x_k vom considera ecuaţia tangentei:

$$y = f(x_k) + f'(x_k)(x - x_k)$$

unde luăm $y = 0$. Avem:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}\ ,\ k = 0,1,2,\ldots\ \ ,\ x_0 \text{ dat}$$

Formula de mai sus poate fi folosită doar dacă la fiecare pas $f'(x_k) \neq 0$. Dacă la un pas avem $f'(x_k) = 0$ putem calcula câteva iteraţii x_k $(k \geq \overline{k})$ folosind $f'(x_{\overline{k}-1})$.

Teoremă de convergenţă

Fie $f \in C^2[a,b]$, cu $f(a)f(b)<0$, $f'(x)\neq 0$ şi $f''(x)\neq 0$ $\forall x\in[a,b]$. Dacă alegem

$$x_0=\begin{cases} a & \textbf{pentru } f(a)f''(a)>0 \\ b & \textbf{pentru } f(b)f''(b)>0 \end{cases}$$

atunci şirul $\{x_k;k\geq 0\}$ construit cu metoda tangentei este monoton, mărginit şi convergent la unica soluţie x^* a ecuaţiei $f(x)=0$. Ordinul de convergenţă este mai mare decât ***2***.

Metoda falsei poziţii (a coardei)

Presupunem că funcţia f este continuă, $f \in C[a,b]$ şi satisface relaţia $f(a)f(b) < 0$. Vom construi un şir $\{x_k\} \subseteq \mathbb{R}$ care să convează la soluţia căutată x^*, $x_k \to x^*$, pentru $k \to \infty$. Considerăm date primul element din şir, x_0 şi un alt punct $\tilde{x}$. Procedeul de construire a şirului este următorul:

$$x_1 = \mathrm{O}x \cap \textbf{dreapta ce unește punctele } \left(\tilde{x}, f(\tilde{x})\right), \left(x_0, f(x_0)\right)$$

$$x_2 = \mathrm{O}x \cap \textbf{dreapta ce unește punctele } \left(\tilde{x}, f(\tilde{x})\right), \left(x_1, f(x_1)\right)$$

$$\vdots$$

$$x_{k+1} = \mathrm{O}x \cap \textbf{dreapta ce unește pt. } \left(\tilde{x}, f(\tilde{x})\right), \left(x_k, f(x_k)\right), k = 0,1,2,\ldots$$

Ecuaţia dreptei ce trece prin punctele ***(a, f(a))*** cu ***(b, f(b))*** este:

$$\frac{y - f(a)}{f(a) - f(b)} = \frac{x - a}{a - b}.$$

Pentru a-l obţine pe x_{k+1} din x_k avem:

$$\frac{y - f(x_k)}{f(x_k) - f(\tilde{x})} = \frac{x - x_k}{x_k - \tilde{x}} \quad \text{cu} \quad y = 0$$

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - \tilde{x})}{f(x_k) - f(\tilde{x})} = \frac{\tilde{x} f(x_k) - x_k f(\tilde{x})}{f(x_k) - f(\tilde{x})} \ ,$$

$$k = 0, 1, 2, \ldots, \quad x_0, \tilde{x} - \text{date}$$

Teoremă de convergenţă

Fie $f \in C^2[a,b]$, cu $f(a)f(b)<0$, $f'(x)\neq 0$ şi $f''(x)\neq 0$ $\forall x \in [a,b]$. Dacă alegem:

$$\begin{cases} \tilde{x}=a \text{ şi } x_0=b & \text{pentru } f(a)f''(a)>0 \\ \tilde{x}=b \text{ şi } x_0=a & \text{pentru } f(b)f''(b)>0 \end{cases}$$

atunci şirul $\{x_k; k\geq 0\}$ construit cu metoda falsei poziţii este monoton, mărginit deci convergent la unica soluţie x^* a ecuaţiei $f(x)=0$.

Metoda secantei

Presupunem că funcţia f este continuă, $f \in C[a,b]$ şi satisface relaţia $f(a)f(b) < 0$. Vom construi un şir $\{x_k\} \subseteq \mathbb{R}$ care să conveargă la soluţia căutată x^*, $x_k \to x^*$, pentru $k \to \infty$. Considerăm date primele două elemente din şir, x_0 şi x_1.

Procedeul de construire a şirului este următorul:

$$x_2 = \mathrm{O}x \cap \textbf{dreapta ce uneşte punctele } \left(x_0, f(x_0)\right), \left(x_1, f(x_1)\right),$$

$$x_3 = \mathrm{O}x \cap \textbf{dreapta ce uneşte punctele } \left(x_1, f(x_1)\right), \left(x_2, f(x_2)\right),$$

$$\vdots$$

$$x_{k+1} = \mathrm{O}x \cap \textbf{dreapta ce uneşte pct. } \left(x_{k-1}, f(x_{k-1})\right), \left(x_k, f(x_k)\right),$$

$$k = 1, 2, \ldots$$

Obţinem elementul x_{k+1} din x_k şi x_{k-1} astfel:

$$\frac{y - f(x_k)}{f(x_k) - f(x_{k-1})} = \frac{x - x_k}{x_k - x_{k-1}} \quad \textbf{cu} \quad y = 0$$

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} =$$

$$= \frac{x_{k-1} f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})} \quad ,$$

$$k = 1, 2, \ldots \ , x_0, x_1 \text{ daţi}$$

Teoremă de convergenţă

Fie x^* o soluţie a ecuţiei $f(x)=0$. Pp. că $f \in C^2\left[x^* - r, x^* + r\right]$, $f'(x) \neq 0$ şi $f''(x) \neq 0$ $\forall x \in \left[x^* - r, x^* + r\right]$. Atunci există $0 < r_0 \leq r$ pentru care, dacă $x_0, x_1 \in [x^* - r_0, x^* + r_0]$ atunci $x_k \in [x^* - r_0, x^* + r_0], \forall k \geq 2$ şi $x_k \to x^*$, pentru $k \to \infty$.

Ordinul de convergenţă este $q = \dfrac{1+\sqrt{5}}{2} \approx 1.61803$.

Metoda lui Laguerre

Fie polinomul:

$$p(x) = a_0(x - x_1)(x - x_2)\cdots(x - x_n)\ ,\ a_0 \neq 0$$

Metoda lui Laguerre propune construirea unui şir de numere care să conveargă la una din rădăcinile polinomului ***p***.
Considerăm derivata polinomului ***p***:

$$p'(x) = p(x)\left[\frac{1}{x-x_1} + \frac{1}{x-x_2} + \cdots + \frac{1}{x-x_n}\right]$$

Avem:

$$\ln|p(x)| = \ln|a_0| + \ln|x-x_1| + \ln|x-x_2| + \cdots + \ln|x-x_n|$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln|p(x)| = \frac{1}{x-x_1} + \frac{1}{x-x_2} + \cdots + \frac{1}{x-x_n} = \frac{p'(x)}{p(x)} = G(x)$$

$$-\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ln|p(x)| = \frac{1}{(x-x_1)^2} + \frac{1}{(x-x_2)^2} + \cdots + \frac{1}{(x-x_n)^2} =$$

$$= \frac{\left[p'(x)\right]^2 - p(x)p''(x)}{\left[p(x)\right]^2} = H(x)$$

Fie x_1 rădăcina pe care vrem s-o aproximăm şi y_k valoarea aproximativă curentă. Notăm cu $a = y_k - x_1$ şi facem presupunerea că y_k se află la acceeaşi distanţă de toate celelalte rădăcini, adică:

$$y_k - x_i = b \quad \forall i = 2,\ldots,n.$$

Prin urmare avem:

$$G(y_k)=\frac{p'(y_k)}{p(y_k)}=\frac{1}{a}+\frac{n-1}{b}$$

$$H(y_k)=\frac{\left[p'(y_k)\right]^2-p(y_k)p''(y_k)}{\left[p(y_k)\right]^2}=\frac{1}{a^2}+\frac{n-1}{b^2}$$

Rezolvăm acest sistem în raport cu ***a*** şi obţinem:

$$a=\frac{n}{\max\left[G(y_k)\pm\sqrt{(n-1)\left(nH(y_k)-G^2(y_k)\right)}\right]}$$

Semnul la numitor este ales astfel ca expresia să aibă magnitudine maximă. Dacă ţinem cont de expresiile pentru ***G*** şi ***H*** obţinem pentru ***a*** următoarea formulă:

$$a = \frac{n\, p(y_k)}{\max\left[p'(y_k) \pm \sqrt{(n-1)^2 \left[p'(y_k) \right]^2 - n(n-1)\, p(y_k)\, p''(y_k)} \right]}$$

Următorul element din şir va fi:

$$y_{k+1} = y_k - a = y_k - \frac{n}{\max\left[G(y_k) \pm \sqrt{(n-1)\left(nH(y_k) - G^2(y_k)\right)}\right]}$$

$$y_{k+1} = y_k - \frac{n\,p(y_k)}{\max\left[p'(y_k) \pm \sqrt{(n-1)^2\left[p'(y_k)\right]^2 - n(n-1)p(y_k)p''(y_k)}\right]}$$

Procedeul se opreşte când ***a*** devine suficient de mic. Metoda lui Laguerre se poate apl. şi ptr. aprox. răd. complexe şi de asemenea pentru polinoame cu coeficienţi complecşi. Pentru rădăcini simple metoda lui Laguerre are ordinul de convergenţă ***3***.

Sisteme de ecuaţii neliniare

Consideră sistemul neliniar:

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n) = 0 \\ f_2(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \ldots, x_n) = 0 \end{cases} \Leftrightarrow \quad F(X) = 0, \; F = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}, \; X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Fie matricea jacobiană asociată funcţiei $\boldsymbol{F}$ (presupunem că funcţiile $\boldsymbol{f_i}$ sunt diferenţiabile):

$$\nabla F(X)=\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}(x_1, x_2, \dots, x_n)$$

Pentru a găsi soluţia X^* a sistemului de ecuaţii neliniare $F(X) = 0$ se construieşte un şir de vectori $X^{(k)} \in \mathbb{R}^n$ astfel:

$$\boldsymbol{X}^{(0)} - \text{dat}$$

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{X}^{(k)} - \left[\nabla \boldsymbol{F}(\boldsymbol{X}^{(k)})\right]^{-1} \boldsymbol{F}(\boldsymbol{X}^{(k)}) = \boldsymbol{X}^{(k)} + \boldsymbol{\Delta}^{(k)},$$

$$\boldsymbol{\Delta}^{(k)} = -\left[\nabla \boldsymbol{F}(\boldsymbol{X}^{(k)})\right]^{-1} \boldsymbol{F}(\boldsymbol{X}^{(k)}) \in \mathbb{R}^{n} \quad , \quad \boldsymbol{k = 0,1,\ldots}$$

Vectorul de corecţie $\boldsymbol{\Delta}^{(k)}$ poate fi calculat şi ca soluţie a sistemului liniar:

$$\nabla \boldsymbol{F}(\boldsymbol{X}^{(k)})\boldsymbol{\Delta} = -\boldsymbol{F}(\boldsymbol{X}^{(k)})$$

unde matricea sistemului este matricea jacobiană calculată în punctul $\boldsymbol{X}^{(k)}$ iar vectorul termenilor liberi este $\left(-\boldsymbol{F}(\boldsymbol{X}^{(k)})\right)$.

Metoda descrisă mai sus poartă numele de metoda Newton. Pentru ***n=1*** metoda Newton este chiar metoda tangentei descrisă anterior.

Optimizare numerică

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

Un punct $x^* \in \mathbb{R}^n$ se numeşte **punct de minim global** pentru funcţia f dacă $f(x^*) \le f(x)\ \ \forall x \in \mathbb{R}^n$.

Un punct $x^* \in \mathbb{R}^n$ se numeşte **punct de minim local** pentru funcţia f dacă există o vecinătate V a punctului x^* pentru care $f(x^*) \le f(x)\ \ \forall x \in V$.

Un punct $x^* \in \mathbb{R}^n$ se numeşte **punct de minim strict local** pentru funcţia f dacă există o vecinătate V a punctului x^* pentru care $f(x^*) < f(x)\ \ \forall x \in V,\ x \ne x^*$.

$$V = S(x^*, r) = \left\{ x \in \mathbb{R}^n ; \left\| x - x^* \right\| \le r \right\}.$$

Dacă funcţia $f \in C^1(\mathbb{R}^n)$ se numeşte **gradient** al funcţiei f în punctul x vectorul:

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right)^T .$$

Dacă funcţia $f \in C^2(\mathbb{R}^n)$ se numeşte **matrice hessiană** a funcţiei f în punctul x matricea:

$$\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i,j=1,\dots,n}$$

$$\nabla^2 f(x) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \dfrac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ & \vdots & & \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

Teorema lui Taylor

Dacă funcţia $f \in C^1(\mathbb{R}^n)$ este continuu diferenţiabilă atunci există $t \in (0,1)$ astfel ca:

$$f(x+p) = f(x) + \nabla f(x+tp)^T p. \qquad (1)$$

Dacă $f \in C^2(\mathbb{R}^n)$ este de două ori continuu diferenţiabilă atunci există $t \in (0,1)$ astfel ca:

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x+tp) p. \quad (2)$$

Teorema 1 (condiţii necesare de optim de ordinul întâi)

Fie $\boldsymbol{f} \in \boldsymbol{C}^{1}(\mathbb{R}^{\boldsymbol{n}})$ şi $\boldsymbol{x^*} \in \mathbb{R}^{\boldsymbol{n}}$ un punct de minim local pentru ***f***. Atunci $\nabla \boldsymbol{f}(\boldsymbol{x^*}) = \boldsymbol{0}$.

Un punct $\boldsymbol{x^*}$ pentru care $\nabla \boldsymbol{f}(\boldsymbol{x^*}) = \boldsymbol{0}$ se numeşte **punct staţionar**. Teorema de mai sus spune că orice punct de minim local trebuie să fie punct staţionar.

Teorema 2 (condiţii necesare de optim de ordinul al doilea)

Fie $\boldsymbol{f} \in \boldsymbol{C}^2(\mathbb{R}^n)$ şi $\boldsymbol{x^*} \in \mathbb{R}^n$ un punct de minim local pentru $\boldsymbol{f}$. Atunci $\nabla \boldsymbol{f}(\boldsymbol{x^*}) = \mathbf{0}$ şi matricea hessiană $\nabla^2 \boldsymbol{f}(\boldsymbol{x^*})$ este pozitiv semidefinită.

O matrice $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ este **pozitiv semidefinită** dacă

$$\left(\boldsymbol{Ax}, \boldsymbol{x}\right)_{\mathbb{R}^n} \geq \mathbf{0} \quad \forall \boldsymbol{x} \in \mathbb{R}^n$$

şi este **pozitiv definită** dacă

$$\left(\boldsymbol{Ax}, \boldsymbol{x}\right)_{\mathbb{R}^n} > \mathbf{0} \quad \forall \boldsymbol{x} \in \mathbb{R}^n,\ \boldsymbol{x} \neq \mathbf{0}.$$

Teorema 3 (condiţii suficiente de ordinul al doilea)

Fie $\boldsymbol{f} \in \boldsymbol{C}^2(\mathbb{R}^n)$ şi $\boldsymbol{x^*} \in \mathbb{R}^n$ un punct pentru care $\boldsymbol{\nabla f(x^*) = 0}$ şi matricea hessiană $\boldsymbol{\nabla^2 f(x^*)}$ este pozitiv definită. Atunci $\boldsymbol{x^*}$ este punct de minim strict local pentru $\boldsymbol{f}$.

O funcţie $\boldsymbol{f}$ se numeşte **convexă** dacă:

$$\boldsymbol{f\left(a\,x_1 + (1-a)x_2\right) \le af(x_1) + (1-a)f(x_2) \ , \forall x_1, x_2 \in \mathbb{R}^n, \forall a \in [0,1].}$$

Teorema

Dacă $\boldsymbol{f}$ este funcţie convexă orice punct de minim local este punct de minim global. Dacă în plus $\boldsymbol{f} \in \boldsymbol{C}^{\mathbf{1}}(\mathbb{R}^{\boldsymbol{n}})$ atunci orice punct staţionar este punct de minim global.

Metode de descreştere

Se numeşte **direcţie de descreştere** a funcţiei $\boldsymbol{f}$ în punctul $\boldsymbol{x}$ un vector $\boldsymbol{d} \in \mathbb{R}^{\boldsymbol{n}}$ pentru care

$$\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{\alpha d}) \leq \boldsymbol{f}(\boldsymbol{x}) \ , \ \forall \boldsymbol{\alpha} \in [\mathbf{0}, \overline{\boldsymbol{\alpha}}).$$

Teoremă

Fie $f \in C^1(\mathbb{R}^n)$. Un vector d este direcţie de descreştere pentru f în x dacă şi numai dacă:

$$\nabla f(x)^T d = \big(d, \nabla f(x)\big)_{\mathbb{R}^n} < 0.$$

Algoritm de descreştere

1. alege $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}$
2. do
    - găseşte o direcţie de descreştere $\boldsymbol{d}$ a lui $\boldsymbol{f}$ în $\boldsymbol{x}$
    - găseşte $\boldsymbol{\tilde{\alpha} > 0}$ astfel ca

      $$\boldsymbol{f(x + \tilde{\alpha} d) = \min\{ f(x + \alpha d);\ \alpha \in (0, \overline{\alpha})\}}$$

      (ajustarea pasului – line search)
    - $\boldsymbol{x = x + \tilde{\alpha} d}$

   while (nu am găsit soluţia)

Ajustarea pasului (line search)

Trebuie rezolvată o problemă de minimizare unidimensională

$$\min\{g(a)\,;\, a \in [b,c]\}$$

Metoda Newton

Fie a_k aproximarea curentă a soluţiei problemei de minimizare de mai sus. Vom folosi dezvoltarea în serie Taylor a funcţiei g

$$g(a) = g(a_k) + g'(a_k)(a-a_k) + \frac{1}{2} g''(a_k)(a-a_k)^2 + \frac{1}{3!} g'''(b_k)(a-a_k)^3$$

$$q(a) = g(a_k) + g'(a_k)(a - a_k) + \frac{1}{2} g''(a_k)(a - a_k)^2$$

Pentru $a \approx a_k$ putem considera că funcţia q aproximează funcţia g , $q(a) \approx g(a)$ şi

$$\min\{g(a)\ ;\ a \in [c,b]\} \simeq \min\{q(a)\ ;\ a \in [c,b]\}.$$

Construim elementul a_{k+1} ca fiind soluţia problemei:

$$q(a_{k+1}) = \min\{q(a)\ ;\ a \in [c,b]\}$$

a_{k+1} este soluţia ecuaţiei

$$q'(a) = 0 \ \Leftrightarrow q'(a) = g'(a_k) + g''(a_k)(a - a_k) \Rightarrow \boxed{a_{k+1} = a_k - \frac{g'(a_k)}{g''(a_k)}}$$

Metoda secantei

Dacă în relaţiile de mai sus se face următoarea aproximare:

$$g''(a_k) \approx \frac{g'(a_k) - g'(a_{k-1})}{a_k - a_{k-1}}$$

obţinem următoarea metodă de aproximare, numită şi metoda secantei:

$$\boxed{a_{k+1} = a_k - g'(a_k)\frac{a_k - a_{k-1}}{g'(a_k) - g'(a_{k-1})}}$$

Aproximare spline cubică

Vom folosi pentru a construi a_{k+1} pe a_{k-1} , a_k , $g(a_{k-1})$, $g(a_k)$, $g'(a_{k-1})$, $g'(a_k)$. Aproximăm funcţia g cu un polinom de grad 3

$$g(a) \approx q(a) = c_0 a^3 + c_1 a^2 + c_2 a + c_3$$

Funcţia q (respectiv constantele c_i) se calculează a.î.

$$q(a_{k-1}) = g(a_{k-1}) \quad , \quad q(a_k) = g(a_k)$$
$$q'(a_{k-1}) = g'(a_{k-1}) \quad , \quad q'(a_k) = g'(a_k)$$

$\boldsymbol{a_{k+1}}$ este punctul de minim al funcţiei $\boldsymbol{q}$

$$a_{k+1} = a_k - (a_k - a_{k-1})\left[\frac{g'(a_k) + u_2 - u_1}{g'(a_k) - g'(a_{k-1}) + 2u_2}\right]$$

$$u_1 = g'(a_{k-1}) + g'(a_k) - 3\frac{g(a_k) - g(a_{k-1})}{a_k - a_{k-1}}$$

$$u_2 = \sqrt{u_1^2 - g'(a_{k-1})g'(a_k)}$$

Ajustarea inexactă a pasului

$$f(x+\tilde{\alpha}d) \cong \min\{f(x+\alpha d);\ \alpha \in (0,\overline{\alpha})\}$$

- nu se obţine $\tilde{\alpha}$ optimal
- pentru reducerea timpului de calcul optimizarea se opreşte înainte de a ajunge la soluţie, în funcţie de anumite criterii/teste de oprire

Regula lui Armijo

$$g(a) = f(x_k + a\,d_k) \quad , \quad \overline{g}(a) = g(0) + \varepsilon a\, g'(0) \; , \; \varepsilon \in (0,1)$$

$\overline{a}$ este acceptabil după regula lui Armijo dacă

$$\text{(1)} \qquad g(\overline{a}) \leq \overline{g}(\overline{a})$$

$$\text{(2)} \qquad g(\sigma\overline{a}) \geq \overline{g}(\sigma\overline{a})$$

$k = 0;$

se alege a_0

while $(g(a_k) > \bar{g}(a_k))$

- $a_{k+1} = \dfrac{1}{\sigma} a_k$
- $k = k + 1$

$\sigma = 2$ sau 10 , $\varepsilon = 0.2$

Testul Goldstein

Dat $\varepsilon \in (0,2), \bar{\alpha}$ este considerat acceptabil dacă:

$$g(\bar{\alpha}) \leq \bar{g}(\bar{\alpha}) \text{ si } g(\bar{\alpha}) > g(0) + (1-\varepsilon) g'(0)$$

$$x_{k+1} = x_k + \bar{\alpha} d_k$$

$\bar{\alpha}$ este acceptat dacă

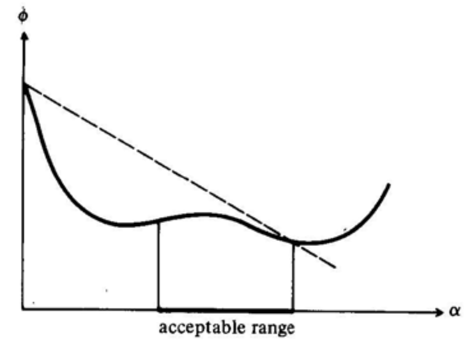
$$\varepsilon \leq \frac{f(x_{k+1}) - f(x_k)}{\alpha \nabla f(x_k) d_k} \leq 1 - \varepsilon$$

Testul Wolfe

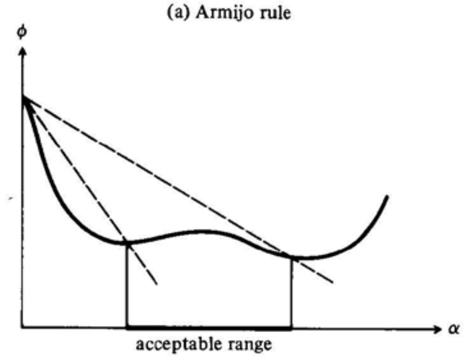
$$\varepsilon \in (0,2)$$

$$\bar{\alpha}: \quad g(\bar{\alpha}) \le \bar{g}(\bar{\alpha}) = g(0) + \varepsilon \bar{\alpha} g'(0)$$

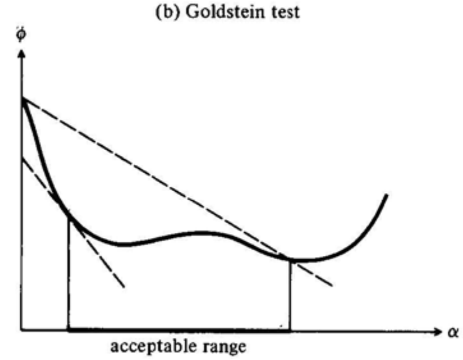
$$g'(\bar{\alpha}) \ge (1-\varepsilon) g'(0)$$



(a) Armijo rule



(b) Goldstein test



(c) Wolfe test