

Calcul Numeric

Cursul 10

2022

Anca Ignat

Rezolvarea ecuațiilor neliniare

Fie $f : [a, b] \rightarrow \mathbb{R}$ o funcție continuă pe intervalul $[a, b]$ astfel ca $f(a)f(b) < 0$. În aceste condiții, există $x^* \in (a, b)$ astfel ca $f(x^*) = 0$. În cele ce urmează ne propunem să aproximăm soluția x^* a ecuației neliniare $f(x) = 0$.

Metoda biseției (a înjumătățirii intervalului)

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

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

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

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

Avem următoarele 3 variante:

1. $f(c)=0$ - soluția căutată este $x^*=c$, algoritmul se oprește;
2. $f(a_k)f(c) < 0 \rightarrow$ soluția se găsește în intervalul (a_k, c) ,
continuăm procedeul cu intervalul $[a_{k+1} = a_k, b_{k+1} = c]$;
3. $f(b_k)f(c) < 0 \rightarrow$ soluția se găsește 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 $[a_{\bar{k}}, b_{\bar{k}}]$ astfel ca $x^* \in (a_{\bar{k}}, b_{\bar{k}})$ și

$b_{\bar{k}} - a_{\bar{k}} < \varepsilon$ ($\bar{k} > \log_2 \frac{b-a}{\varepsilon}$). Pentru ε suficient de mic atât

$a_{\bar{k}}$ cât și $b_{\bar{k}}$ pot fi considerate aproximări ale soluției x^*

($a_{\bar{k}} \approx x^*$ prin lipsă iar $b_{\bar{k}} \approx x^*$ prin adaos).

Metoda tangentei (Newton-Raphson)

Vom presupune că funcția $f \in C^1[a,b]$ este derivabilă pe $[a,b]$ cu derivata continuă în acest interval și satisface relația $f(a)f(b) < 0$. Pentru a aproxima soluția x^* a ecuației $f(x)=0$ vom construi un șir $\{x_k\} \subseteq \mathbb{R}$ care să convergă la x^* , $x_k \rightarrow x^*$, pentru $k \rightarrow \infty$. Primul element din șir, x_0 , considerăm că este dat. Următorul element din șir se construiește ca fiind punctul de intersecție al tangentei la graficul funcției f în punctul $(x_0, f(x_0))$ cu axa absciselor. Procedul se repetă cu x_1 pentru a-l obține pe x_2 , ș.a.m.d.

$x_1 = \mathbf{Ox} \cap \text{tangentă la graficul funcției } f \text{ în punctul } (x_0, f(x_0))$

$x_2 = \mathbf{Ox} \cap \text{tangentă la graficul funcției } f \text{ în punctul } (x_1, f(x_1))$

\vdots

$x_{k+1} = \mathbf{Ox} \cap \text{tangentă la graficul fct. } f \text{ în pt. } (x_k, f(x_k)), k = 0, 1, 2,$

Ecuția tangentei la graficul funcției f într-un punct $(a, f(a))$ este următoarea (pentru o funcție derivabilă):

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

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

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

unde luăm $y = 0$. Avem:

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

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 iterații x_k ($k \geq \bar{k}$) folosind $f'(x_{\bar{k}-1})$.

Teoremă de convergență

Fie $f \in C^2[a, b]$, cu $f(a)f(b) < 0$, $f'(x) \neq 0$ și $f''(x) \neq 0$

$\forall x \in [a, b]$. Dacă alegem

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

atunci șirul $\{x_k; k \geq 0\}$ construit cu metoda tangentei este monoton, mărginit și convergent la unica soluție x^* a ecuației $f(x)=0$. Ordinul de convergență este mai mare decât 2.

Metoda falsei poziții (a coardei)

Presupunem că funcția f este continuă, $f \in C[a, b]$ și satisface relația $f(a)f(b) < 0$. Vom construi un șir $\{x_k\} \subseteq \mathbb{R}$ care să convergă la soluția căutată x^* , $x_k \rightarrow x^*$, pentru $k \rightarrow \infty$. Considerăm date primul element din șir, x_0 și un alt punct \tilde{x} . Procedeu de construire a șirului este următorul:

$x_1 = \mathbf{Ox} \cap \text{dreapta ce unește punctele } (\tilde{x}, f(\tilde{x})), (x_0, f(x_0))$

$x_2 = \mathbf{Ox} \cap \text{dreapta ce unește punctele } (\tilde{x}, f(\tilde{x})), (x_1, f(x_1))$

\vdots

$x_{k+1} = \mathbf{Ox} \cap \text{dreapta ce unește pt. } (\tilde{x}, f(\tilde{x})), (x_k, f(x_k)), k = 0, 1, 2, \dots$

Ecuția 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 obține pe \mathbf{x}_{k+1} din \mathbf{x}_k avem:

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

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{f(\mathbf{x}_k)(\mathbf{x}_k - \tilde{\mathbf{x}})}{f(\mathbf{x}_k) - f(\tilde{\mathbf{x}})} = \frac{\tilde{\mathbf{x}}f(\mathbf{x}_k) - \mathbf{x}_kf(\tilde{\mathbf{x}})}{f(\mathbf{x}_k) - f(\tilde{\mathbf{x}})},$$
$$k = 0, 1, 2, \dots, \quad \mathbf{x}_0, \tilde{\mathbf{x}} - \text{date}$$

Teoremă de convergență

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

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

atunci șirul $\{x_k; k \geq 0\}$ construit cu metoda falsei poziții este monoton, mărginit deci convergent la unica soluție x^* a ecuației $f(x)=0$.

Metoda secantei

Presupunem că funcția f este continuă, $f \in C[a, b]$ și satisface relația $f(a)f(b) < 0$. Vom construi un șir $\{x_k\} \subseteq \mathbb{R}$ care să convergă la soluția căutată x^* , $x_k \rightarrow x^*$, pentru $k \rightarrow \infty$. Considerăm date primele două elemente din șir, x_0 și x_1 .

Procedeul de construire a șirului este următorul:

$$\begin{aligned}
x_2 &= \mathbf{O}x \cap \text{dreapta ce unește punctele } (x_0, f(x_0)), (x_1, f(x_1)), \\
x_3 &= \mathbf{O}x \cap \text{dreapta ce unește punctele } (x_1, f(x_1)), (x_2, f(x_2)), \\
&\vdots \\
x_{k+1} &= \mathbf{O}x \cap \text{dreapta ce unește pct. } (x_{k-1}, f(x_{k-1})), (x_k, f(x_k)) , \\
&\qquad\qquad\qquad k = 1, 2, \dots
\end{aligned}$$

Obținem elementul x_{k+1} din x_k și x_{k-1} astfel:

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

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

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

$k = 1, 2, \dots$, x_0, x_1 dađı

Teoremă de convergență

Fie x^* o soluție a ecuației $f(x)=0$. Pp. că
 $f \in C^2[x^* - r, x^* + r]$, $f'(x) \neq 0$ și $f''(x) \neq 0$

$\forall x \in [x^* - r, x^* + r]$. Atunci există

$0 < r_0 \leq r$ pentru care, dacă $x_0, x_1 \in [x^* - r_0, x^* + r_0]$ atunci
 $x_k \in [x^* - r_0, x^* + r_0], \forall k \geq 2$ și $x_k \rightarrow x^*$, pentru $k \rightarrow \infty$.

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

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 șir de numere care să convergă la una din rădăcinile polinomului p .

Considerăm derivata polinomului p :

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

Avem:

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

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

$$\begin{aligned}
-\frac{d^2}{dx^2} \ln |p(x)| &= \frac{1}{(x-x_1)^2} + \frac{1}{(x-x_2)^2} + \dots + \frac{1}{(x-x_n)^2} = \\
&= \frac{[p'(x)]^2 - p(x)p''(x)}{[p(x)]^2} = H(x)
\end{aligned}$$

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

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

Prin urmare avem:

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

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

Rezolvăm acest sistem în raport cu a și obținem:

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

Semnul la numitor este ales astfel ca expresia să aibă magnitudine maximă. Dacă ținem cont de expresiile pentru \mathbf{G} și \mathbf{H} obținem pentru \mathbf{a} următoarea formulă:

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

Următorul element din șir va fi:

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

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

Procedeul se oprește când a devine suficient de mic. Metoda lui Laguerre se poate apl. și ptr. aprox. răd. complexe și de asemenea pentru polinoame cu coeficienți complecși. Pentru rădăcini simple metoda lui Laguerre are ordinul de convergență 3.

Sisteme de ecuații neliniare

Consideră sistemul neliniar:

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

Fie matricea jacobiană asociată funcției F (presupunem că funcțiile f_i sunt diferențiabile):

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

Pentru a găsi soluția X^* a sistemului de ecuații neliniare $F(X)=0$ se construiește un șir de vectori $X^{(k)} \in \mathbb{R}^n$ astfel:

$X^{(0)}$ – dat

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

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

Vectorul de corecție $\Delta^{(k)}$ poate fi calculat și ca soluție a sistemului liniar:

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

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

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

Optimizare numerică

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

Un punct $x^* \in \mathbb{R}^n$ se numește **punct de minim global** pentru funcția f dacă $f(x^*) \leq f(x) \quad \forall x \in \mathbb{R}^n$.

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

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

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

Dacă funcția $f \in C^1(\mathbb{R}^n)$ se numește **gradient** al funcției f în punctul \mathbf{x} vectorul:

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

Dacă funcția $f \in C^2(\mathbb{R}^n)$ se numește **matrice hessiană** a funcției f în punctul \mathbf{x} matricea:

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

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

Teorema lui Taylor

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

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

Dacă $f \in C^2(\mathbb{R}^n)$ este de două ori continuu diferențiabilă 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 (condiții necesare de optim de ordinul întâi)

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

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

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

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

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

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

și este **pozitiv definită** dacă

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

Teorema 3 (condiții suficiente de ordinul al doilea)

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

O funcție f se numește **convexă** dacă:

$$f(ax_1 + (1-a)x_2) \leq af(x_1) + (1-a)f(x_2), \forall x_1, x_2 \in \mathbb{R}^n, \forall a \in [0,1].$$

Teorema

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

Metode de descreștere

Se numește **direcție de descreștere** a funcției f în punctul x un vector $d \in \mathbb{R}^n$ pentru care

$$f(x + \alpha d) \leq f(x) \quad , \quad \forall \alpha \in [0, \bar{\alpha}).$$

Teoremă

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

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

Algoritm de descreștere

1. alege $x \in \mathbb{R}^n$
 2. do
 - găsește o direcție de descreștere d a lui f în x
 - găsește $\tilde{\alpha} > 0$ astfel ca
$$f(x + \tilde{\alpha}d) = \min\{f(x + \alpha d); \alpha \in (0, \bar{\alpha})\}$$
(ajustarea pasului – line search)
 - $x = x + \tilde{\alpha}d$
- while (nu am găsit soluția)

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 soluției problemei de minimizare de mai sus. Vom folosi dezvoltarea în serie Taylor a funcției g

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

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

Pentru $a \approx a_k$ putem considera că funcția q aproximează funcția g , $q(a) \approx g(a)$ și

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

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

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

a_{k+1} este soluția ecuației

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

Metoda secantei

Dacă în relațiile de mai sus se face următoarea aproximare:

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

obținem următoarea metodă de aproximare, numită și metoda secantei:

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

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 funcția g cu un polinom de grad 3

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

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

$$\begin{aligned} q(a_{k-1}) &= g(a_{k-1}) & , & & q(a_k) &= g(a_k) \\ q'(a_{k-1}) &= g'(a_{k-1}) & , & & q'(a_k) &= g'(a_k) \end{aligned}$$

a_{k+1} este punctul de minim al funcției 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

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

- nu se obține $\tilde{\alpha}$ optimal
- pentru reducerea timpului de calcul optimizarea se oprește înainte de a ajunge la soluție, în funcție de anumite criterii/teste de oprire

Regula lui Armijo

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

\bar{a} este acceptabil după regula lui Armijo dacă

$$(1) \quad g(\bar{a}) \leq \bar{g}(\bar{a})$$

$$(2) \quad g(\sigma \bar{a}) \geq \bar{g}(\sigma \bar{a})$$

$k = 0;$

se alege a_0

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

$\cdot a_{k+1} = \frac{1}{\sigma} a_k$

$\cdot k = k + 1$

$\sigma = 2 \text{ sau } 10, \quad \varepsilon = 0.2$

Testul Goldstein

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

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

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

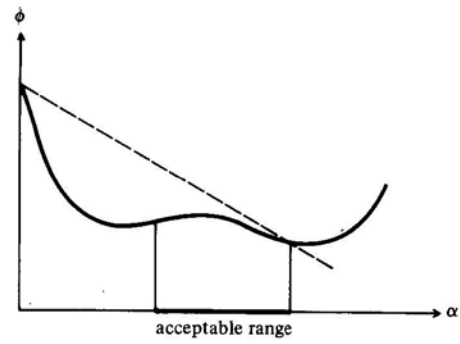
$\bar{\alpha}$ este acceptat dacă

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

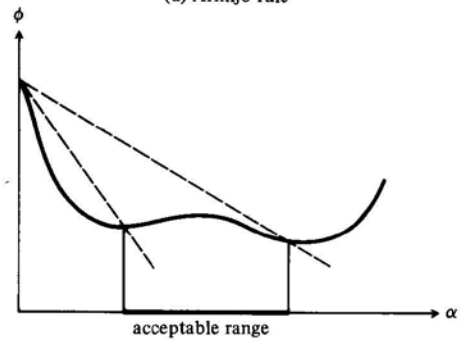
Testul Wolfe

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

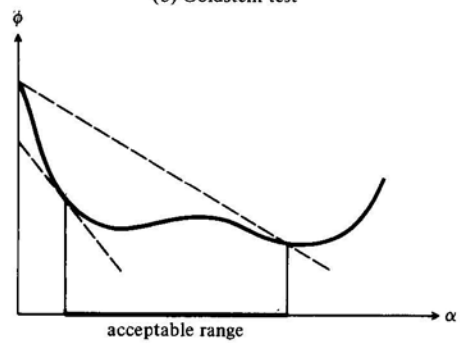
$$\begin{aligned}\bar{\alpha} : \quad g(\bar{\alpha}) &\leq \bar{g}(\bar{\alpha}) = g(0) + \varepsilon \bar{\alpha} g'(0) \\ g'(\bar{\alpha}) &\geq (1 - \varepsilon)g'(0)\end{aligned}$$



(a) Armijo rule



(b) Goldstein test



(c) Wolfe test