METODE DE OPTIMIZARE UNI-DIMENSIONALĂ

În accepțiunea actuală, un algoritm de optimizare de tipul (1.4.1) este compunerea a două aplicații: prima determină direcția de deplasare către minim, în care se presupune că direcția este descendentă. A doua determină lungimea pasului de deplasare de-a lungul acestei direcții, care în esența ei este o metodă de minimizare pentru funcții care depind de o singură variabilă:

$$min\ \varphi(\alpha),\ \ \alpha \in R.$$

De aceea orice metodă de optimizare a funcțiilor de o singură variabilă se poate utiliza în acest scop. Totuși, vom vedea că în cadrul algoritmilor de optimizare se instituie o serie de proceduri care țin seama de valorile funcției de minimizat şi de gradientul acesteia, conducând astfel la algoritmi eficienți de optimizare. Metodele de determinare a lungimii pasului se grupează în două clase: metode de alegere optimă şi metode de alegere economică [Dennis şi Schnabel, 1983].

Metodele de alegere optimă, cunoscute încă ca şi metode de căutare exactă, urmăresc determinarea lui α^* din condiția de realizare a minimului funcției $f(x)$ în direcția d_k, adică $\varphi(\alpha) = f(x_k + \alpha d_k)$. Cele mai eficiente metode din această clasă se bazează fie pe explorare (căutare) directă, fie pe interpolare cu sau fără utilizarea derivatelor funcției de minimizat.

Metodele de explorare directă procedează la identificarea unui interval $[a,b] \subset R$ care conține punctul α^*, urmată imediat de o reducere iterativă a lungimii acestuia până la o valoare mai mică decât o toleranță impusă de localizare a lui α^*. Eficiența acestui mod de abordare depinde de strategia de construcție a lanțului de intervale $[a_k, b_k],\ \ k = 1,2,\cdots$ care conține α^*. Cele mai cunoscute metode de explorare directă sunt: metoda secțiunii de aur şi metoda Fibonacci, pe care le vom prezenta în acest capitol. Tot în această clasă de metode de explorare directă se consideră a fi şi metodele clasice Newton-Raphson şi a secantei pe ce care le vom discuta în capitolul 6 într-un context mai general.

Metodele de interpolare determină o aproximație a lui α^* utilizând fie valorile funcției de minimizat $f(x_k + \alpha d_k)$ în puncte din intervalul inițial de căutare $[0, L]$, fie valorile funcției şi ale derivatei acestei $\nabla f(x_k + \alpha d_k)^T d_k$ în două

puncte din $[0, L]$, unde L este o aproximație inițială a lui α^*, de exemplu $L = 1$. Cu aceste valori se construieşte polinomul de interpolare de gradul doi sau trei $q(\alpha)$ al funcției $f(x_k + \alpha d_k)$, pentru $\alpha \geq 0$, şi se determină punctul de minim $\hat{\alpha}$ al lui $q(\alpha)$. Cele mai cunoscute sunt metoda de interpolare pătratică (în două sau trei puncte) şi metoda de interpolare cubică.

Scopul acestui capitol este de a prezenta metode de determinare a lungimii pasului, cunoscute încă ca şi metode de căutare liniară, utilizând tehnici de alegere optimă a acestei lungimi.

Metodele de alegere economică a lungimii pasului, pe de altă parte, urmăresc determinarea lui α^* din condiția realizării unei reduceri semnificative a valorii funcției f în direcția d_k. De regulă acestea aleg pasul α pe baza unor majorări convenabile a diferenței $f(x_k + \alpha d_k) - f(x_k) = \varphi(\alpha) - \varphi(0)$. Aceste metode vor fi tratate în detaliu în capitolul 4.

3.1. METODA ÎNAINTE-ÎNAPOI. FUNCȚII UNIMODALE

O metodă simplă de determinare a unui interval inițial de căutare este dată de metoda înainte-înapoi. Ideea acestei metode este următoarea. Se consideră un punct inițial şi o lungime inițială a pasului. Deplasându-ne înainte se determină trei puncte din intervalul de căutare în care geometria funcției φ arată valorile „mare-mic-mare”. Dacă aceasta nu este posibil, atunci ne deplasăm înapoi. Altfel spus, dat punctul inițial α_0 şi lungimea inițială $h_0 > 0$, atunci dacă $\varphi(\alpha_0 + h_0) < \varphi(\alpha_0)$, atunci plecăm din punctul $\alpha_0 + h_0$ şi continuăm deplasarea înainte cu o valoare a pasului mai mare atât timp cât valorile funcției cresc. Dacă, $\varphi(\alpha_0 + h_0) > \varphi(\alpha_0)$, atunci din α_0 ne deplasăm înapoi atât timp cât valorile funcției cresc. În acest mod vom obține un interval inițial de căutare care conține minimul α^* [Bazaraa, Sheraly şi Shetty, 1993].

Algoritmul 3.1.1 (Algoritmul înainte-înapoi)

Pasul 1.	*Inițializare.* Se consideră: $\alpha_0 \in [0, +\infty)$, $h_0 > 0$, şi factorul $t > 1$ (de obicei $t = 2$). Se evaluează $\varphi(\alpha_0)$ şi se pune $k = 0$.
Pasul 2.	*Compararea valorilor funcției.* Se pune $\alpha_{k+1} = \alpha_k + h_k$ şi se calculează $\varphi_{k+1} = \varphi(\alpha_{k+1})$. Dacă $\varphi_{k+1} < \varphi_k$, atunci se continuă cu pasul 3; altfel se execută pasul 4.
Pasul 3.	*Pasul înainte.* Se pune $h_{k+1} = th_k$, $\alpha = \alpha_k$, $\alpha_k = \alpha_{k+1}$, $\varphi_k = \varphi_{k+1}$, $k = k + 1$ şi se continuă cu pasul 2.
Pasul 4.	*Pasul înapoi.* Dacă $k = 0$, atunci se inversează direcția de căutare. Se

	pune: $h_k = -h_k$, $\alpha_k = \alpha_{k+1}$ şi se continuă cu pasul 2, altfel, se pune: $a = min\{\alpha, \alpha_{k+1}\}$, $b = max\{\alpha, \alpha_{k+1}\}$ şi stop. ♦

Metodele de căutare liniară se bazează pe proprietatea de unimodalitate a funcției în direcția de căutare. Această proprietate este foarte naturală pentru funcțiile continuu diferențiabile, mai ales datorită faptului că intervalul de incertitudine este acceptabil de mic.

Definiția 3.1.1. *Fie funcția* $\varphi : R \to R$ *şi intervalul* $[a,b] \subset R$. *Dacă există* $\alpha^* \in [a,b]$ *astfel încât* $\varphi(\alpha)$ *este strict descrescătoare pe* $[a,\alpha^*]$ *şi strict crescătoare pe* $[\alpha^*,b]$, *atunci* $\varphi(\alpha)$ *se numeşte funcție unimodală pe* $[a,b]$. *Intervalul* $[a,b]$ *se numeşte interval de unimodalitate pentru funcția* $\varphi(\alpha)$.

Funcțiile unimodale se pot defini încă sub forma:

Definiția 3.1.2. *Dacă există un unic punct* $\alpha^* \in [a,b]$ *astfel încât pentru orice* $\alpha_1, \alpha_2 \in [a,b]$, $\alpha_1 < \alpha_2$, *următoarele condiții au loc:*
- dacă $\alpha_2 < \alpha^*$, *atunci* $\varphi(\alpha_1) > \varphi(\alpha_2)$;
- dacă $\alpha_1 > \alpha^*$, *atunci* $\varphi(\alpha_1) < \varphi(\alpha_2)$;
atunci funcția $\varphi(\alpha)$ *este unimodală pe intervalul* $[a,b]$.

Observăm imediat că funcțiile unimodale nu implică continuitatea şi diferențiabilitatea. Proprietatea importantă a funcțiilor unimodale constă în faptul că în procesul de căutare putem exclude subintervale din intervalul de incertitudine în care suntem siguri că minimul funcției nu există. Următoarea teoremă arată că dacă funcția φ este unimodală, atunci intervalul de incertitudine se poate reduce prin compararea valorilor funcției φ calculate numai în două puncte din interval.

Teorema 3.1.1. *Fie* $\varphi : R \to R$ *unimodală pe* $[a,b]$ *şi fie* $\alpha_1, \alpha_2 \in [a,b]$ *cu* $\alpha_1 < \alpha_2$. *Atunci*

(i) *Dacă* $\varphi(\alpha_1) \le \varphi(\alpha_2)$, *atunci* $[a,\alpha_2]$ *este un interval de unimodalitate pentru* φ;

(ii) *Dacă* $\varphi(\alpha_1) \ge \varphi(\alpha_2)$, *atunci* $[\alpha_1,b]$ *este un interval de unimodalitate pentru* φ.

Demonstrație*. (i)* Din definiția 3.1.2 observăm că există $\alpha^* \in [a,b]$ astfel încât $\varphi(\alpha)$ este strict descrescătoare pe $[a,\alpha^*]$ şi strict crescătoare pe $[\alpha^*,b]$. Deoarece $\varphi(\alpha_1) \le \varphi(\alpha_2)$, atunci $\alpha^* \in [a,\alpha_2]$. Deoarece $\varphi(\alpha)$ este unimodală pe

$[a,b]$, rezultă că este de asemenea unimodală pe $[a,\alpha_2]$. Deci $[a,\alpha_2]$ este un interval de unimodalitate pentru $\varphi(\alpha)$, ceea ce demonstrează prima parte a teoremei. Partea *(ii)* se demonstrează similar. ■

3.2. METODA SECȚIUNII DE AUR ŞI A LUI FIBONACCI

Metoda secțiunii de aur şi a lui Fibonacci sunt metode de secționare. Ideea acestor metode de minimizare a funcțiilor unimodale pe intervalul $[a,b]\subset R$ este de a reduce iterativ intervalul de incertitudine numai prin compararea valorilor funcției de minimizat. De îndată ce lungimea intervalului de incertitudine este mai mică decât un prag prestabilit, atunci punctele din acest interval se pot considera aproximații ale minimului funcției în direcția dată. Această clasă de metode utilizează numai valorile funcției şi sunt utilizate în cadrul algoritmilor de optimizare, în special pentru probleme nenetede sau a căror derivate a funcției de minimizat au expresii complicate [Gill, Murray şi Wright, 1981].

3.2.1. Metoda secțiunii de aur[1]

Să considerăm deci funcția $\varphi(\alpha)=f(x_k+\alpha d_k)$ unimodală pe intervalul $[a,b]$. La iterația k metoda secțiunii de aur a determinat intervalul $[a_k,b_k]$. În acest moment să considerăm două puncte $\lambda_k,\mu_k\in[a_k,b_k]$ cu $\lambda_k<\mu_k$ şi calculăm $\varphi(\lambda_k)$ şi $\varphi(\mu_k)$. Conform teoremei 3.1.1 avem;

(i) Dacă $\varphi(\lambda_k)\le\varphi(\mu_k)$, atunci punem $a_{k+1}=a_k$ şi $b_{k+1}=\mu_k$.

(ii) Dacă $\varphi(\lambda_k)>\varphi(\mu_k)$, atunci punem $a_{k+1}=\lambda_k$ şi $b_{k+1}=b_k$.

Rămâne să discutăm problema cum alegem punctele λ_k şi μ_k. Pentru aceasta impunem următoarele trei condiții:

1. Distanțele de la λ_k şi μ_k la capetele intervalului $[a_k,b_k]$ sunt echivalente în sensul că:

$$b_k-\lambda_k=\mu_k-a_k. \tag{3.2.1}$$

2. Rata de reducere a lungimii intervalelor de incertitudine la fiecare iterație este aceeaşi, adică

---

[1] Secțiunea de aur este numărul irațional dat de $(1+\sqrt{5})/2\cong 1.618033989$. Aceasta se mai numeşte încă *proporția de aur, proporția divină* sau *numărul de aur*. Structurile geometrice care respectă secțiunea de aur au o anumită frumusețe estetică, recunoscută încă din timpurile antice, şi care sugerează un anumit echilibru între simetrie şi asimetrie. Două cantități a şi b cu $a>b$ sunt într-o proporție de aur dacă raportul dintre suma lor la cantitatea mai mare este egal cu raportul dintre cantitatea mai mare la cea mică, adică $(a+b)/a=a/b$.

$$b_{k+1} - a_{k+1} = \tau(b_k - a_k), \text{ unde } \tau \in (0,1). \quad (3.2.2)$$

3. La fiecare iterație se execută numai o singură evaluare a funcției de minimizat.

Să considerăm acum cazul *(i)*. Introducând valorile cazului *(i)* în (3.2.2) obținem $\mu_k - a_k = \tau(b_k - a_k)$ şi apoi combinând aceasta cu (3.2.1) rezultă $b_k - \lambda_k = \mu_k - a_k$. Din aceste relații găsim:

$$\lambda_k = a_k + (1-\tau)(b_k - a_k), \quad (3.2.3)$$

$$\mu_k = a_k + \tau(b_k - a_k). \quad (3.2.4)$$

În acest caz noul interval este $[a_{k+1}, b_{k+1}] = [a_k, \mu_k]$. Pentru reducerea intervalului de incertitudine este necesar să alegem punctele λ_{k+1} şi μ_{k+1}. Din (3.2.4) găsim

$$\begin{aligned}\mu_{k+1} &= a_{k+1} + \tau(b_{k+1} - a_{k+1}) = a_k + \tau(\mu_k - a_k) \\ &= a_k + \tau(a_k + \tau(b_k - a_k) - a_k) = a_k + \tau^2(b_k - a_k). \quad (3.2.5)\end{aligned}$$

Dacă punem

$$\tau^2 = 1 - \tau, \quad (3.2.6)$$

atunci

$$\mu_{k+1} = a_k + (1-\tau)(b_k - a_k) = \lambda_k. \quad (3.2.7)$$

Rezultă deci că în noul punct μ_{k+1} nu este necesar să evaluăm funcția, deoarece acesta coincide cu λ_k în care deja cunoaştem valoarea funcției.

În mod similar se demonstrează că în cazul *(ii)* noul punct λ_{k+1} coincide cu μ_k. Deci la fiecare iterație este necesară numai o singură evaluare a funcției de minimizat, aşa după cum am precizat mai sus.

În continuare să vedem care este rata de reducere a lungimii intervalului de incertitudine la fiecare iterație. Rezolvând ecuația (3.2.6) găsim

$$\tau = \frac{-1 \pm \sqrt{5}}{2}.$$

Deoarece $\tau > 0$, considerăm

$$\tau = \frac{b_{k+1} - a_{k+1}}{b_k - a_k} = \frac{\sqrt{5}-1}{2} \cong 0.618. \quad (3.2.8)$$

Cu acestea, din (3.2.3) şi (3.2.4) obținem

$$\lambda_k = a_k + 0.382(b_k - a_k), \quad (3.2.9)$$

$$\mu_k = a_k + 0.618(b_k - a_k). \quad (3.2.10)$$

Deoarece la fiecare iterație rata de reducere a intervalului de incertitudine este $\tau = 0.618$, rezultă că dacă intervalul inițial este $[a_1, b_1]$, atunci lungimea intervalului după n iterații este $\tau^{n-1}(b_1 - a_1)$, ceea ce arată că *rata de convergență a metodei secțiunii de aur este liniară* [Avriel, 1976]

Cu acestea algoritmul secțiunii de aur este următorul:

Algoritmul 3.2.1 (Algoritmul secțiunii de aur)

Pasul 1.	*Inițializare.* Se determină intervalul inițial de incertitudine $[a_1, b_1]$ şi se consideră acuratețea $\delta > 0$. Se determină punctele inițiale λ_1 şi μ_1: $\lambda_1 = a_1 + 0.382(b_1 - a_1)$, $\mu_1 = a_1 + 0.618(b_1 - a_1)$. Se evaluează $\varphi(\lambda_1)$ şi $\varphi(\mu_1)$, şi se pune $k = 1$.
Pasul 2.	*Se compară valorile funcției de minimizat.* Dacă $\varphi(\lambda_k) > \varphi(\mu_k)$, atunci se continuă cu pasul 3; dacă $\varphi(\lambda_k) \le \varphi(\mu_k)$, atunci se trece la pasul 4.
Pasul 3.	*Cazul (ii).* Dacă $b_k - \lambda_k \le \delta$, stop cu ieşirea μ_k; altfel se calculează: $a_{k+1} = \lambda_k$, $b_{k+1} = b_k$, $\lambda_{k+1} = \mu_k$, $\varphi(\lambda_{k+1}) = \varphi(\mu_k)$, $\mu_{k+1} = a_{k+1} + 0.618(b_{k+1} - a_{k+1})$. Se evaluează $\varphi(\mu_{k+1})$ şi se continuă cu pasul 5.
Pasul 4.	*Cazul (i).* Dacă $\mu_k - a_k \le \delta$, stop cu ieşirea λ_k; altfel se calculează: $a_{k+1} = a_k$, $b_{k+1} = \mu_k$, $\mu_{k+1} = \lambda_k$, $\varphi(\mu_{k+1}) = \varphi(\lambda_k)$, $\lambda_{k+1} = a_{k+1} + 0.382(b_{k+1} - a_{k+1})$. Se evaluează $\varphi(\lambda_{k+1})$ şi se continuă cu pasul 5.
Pasul 5.	Se pune $k = k + 1$ şi se continuă cu pasul 2. ♦

3.2.2. Metoda lui Fibonacci

Fibonacci (1170-1250)

O metodă foarte asemănătoare cu metoda secțiunii de aur este metoda Fibonacci pe care acesta a prezentat-o în cartea sa *Liber Abaci* scrisă în 1202 unde a discutat şirul numerelor Fibonacci în legătură cu înmulțirea iepurilor. Principala diferență față de metoda secțiunii de aur constă în faptul că în metoda Fibonacci rata de reducere a lungimii intervalului de incertitudine nu se defineşte în funcție de secțiunea de aur dată de $\tau \cong 0.618$ ci în funcție de şirul numerelor Fibonacci. Cu alte cuvinte, reducerea lungimii intervalului de incertitudine variază de la o iterație la alta. Ambele aceste metode au aceeaşi rată de convergență, totuşi într-un anumit sens, metoda Fibonacci este optimă.

Şirul numerelor Fibonacci $\{F_k\}$ se defineşte sub forma:

$$F_0 = F_1 = 1, \tag{3.2.11}$$

$$F_{k+1} = F_k + F_{k-1}, \quad k = 1, 2, ... \tag{3.2.12}$$

Dacă în relaţiile (3.2.3)-(3.2.4) în locul lui τ se utilizează raportul F_{n-k} / F_{n-k+1}, atunci obţinem:

$$\lambda_k = a_k + \left(1 - \frac{F_{n-k}}{F_{n-k+1}}\right)(b_k - a_k) = a_k + \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k), k = 1, \ldots, n-1, \tag{3.2.13}$$

$$\mu_k = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k), \quad k = 1, \ldots, n-1, \tag{3.2.14}$$

care defineşte metoda Fibonacci.

Dacă $\varphi(\lambda_k) \le \varphi(\mu_k)$, atunci noul interval de incertitudine este $[a_{k+1}, b_{k+1}] = [a_k, \mu_k]$. Deci, utilizând (3.2.14) obţinem

$$b_{k+1} - a_{k+1} = \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k), \tag{3.2.15}$$

care furnizează reducerea la fiecare iteraţie. Se vede uşor că această formulă este valabilă şi în cazul în care $\varphi(\lambda_k) > \varphi(\mu_k)$.

Presupunem acum că dorim ca lungimea intervalului final de incertitudine să nu fie mai mare decât valoarea $\delta > 0$, adică $b_n - a_n \le \delta$, unde δ este o toleranţă dată, suficient de mică. Atunci, deoarece

$$b_n - a_n = \frac{F_1}{F_2}(b_{n-1} - a_{n-1}) = \frac{F_1}{F_2}\frac{F_2}{F_3} \cdots \frac{F_{n-1}}{F_n}(b_1 - a_1) = \frac{1}{F_n}(b_1 - a_1), \tag{3.2.16}$$

rezultă că

$$F_n \ge \frac{b_1 - a_1}{\delta}. \tag{3.2.17}$$

Deci, dat intervalul iniţial $[a_1, b_1]$, precum şi marginea superioară δ a lungimii intervalului final de incertitudine, rezultă că putem determina numărul Fibonacci F_n şi de aici din (3.2.17) numărul n al acestor numere din secvenţa Fibonacci. Algoritmul corespunzător metodei Fibonacci este următorul.

Algoritmul 2.2.2. (Algoritmul Fibonacci)

Pasul 1.	Se consideră F_n ca primul număr Fibonacci mai mare decât $(b_1 - a_1)/\delta$. Se iniţializează: $D_1 = 0$ şi $D_2 = F_n$.
Pasul 2.	Se consideră punctele: $\rho_1 = D_1 + F_{n-2}$, $\rho_2 = D_1 + F_{n-1}$, precum şi $\gamma_1 = \delta\rho_1$ şi $\gamma_2 = \delta\rho_2$. Se evaluează $\varphi_1 = \varphi(\gamma_1)$ şi $\varphi_2 = \varphi(\gamma_2)$.
Pasul 3.	Dacă $\varphi_2 > \varphi_1$, atunci se pune $D_2 = \rho_2$; altfel se pune $D_1 = \rho_1$.
Pasul 4.	Dacă $\rho_1 = \rho_2$, atunci stop; altfel se consideră $n = n - 1$ şi se continuă cu pasul 2. ♦

În continuare prezentăm o mică observație care precizează faptul că pentru $k \to \infty$ metoda Fibonacci şi cea a secțiunii de aur au aceeaşi rată de reducere a intervalului de incertitudine. Într-adevăr, dacă punem $F_k = r^k$, atunci din (3.2.11)-(3.2.12) obținem $r^2 - r - 1 = 0$, care rezolvată are soluțiile:

$$r_1 = \frac{1+\sqrt{5}}{2},\ r_2 = \frac{1-\sqrt{5}}{2}.$$

Soluția generală a ecuației cu diferențe finite $F_{k+1} = F_k + F_{k-1}$ este $F_k = Ar_1^k + Br_2^k$. Din condițiile inițiale $F_0 = F_1 = 1$, găsim $A = r_1/\sqrt{5}$ şi $B = -r_2/\sqrt{5}$. Deci

$$F_k = \frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1}\right\}.$$

Cu acestea

$$\lim_{k\to\infty} \frac{F_{k-1}}{F_k} = \frac{\sqrt{5}-1}{2} = \tau. \qquad (3.2.18)$$

Aceasta arată că ambele metode de secționare a intervalului de incertitudine, metoda secțiunii de aur şi metoda Fibonacci, au exact aceeaşi rată de reducere a lungimii intervalului de incertitudine când $k \to \infty$. Deci metoda Fibonacci converge cu rata de convergență τ. Din punctul de vedere al numărului de puncte în care funcția de minimizat trebuie evaluată, metoda Fibonacci este optimă, ea cerând cel mai mic număr de evaluări pentru a obține un interval final de lungime mai mică sau egală cu δ. În acest context şi metoda secțiunii de aur este aproximativ optimă, ea cerând un număr de evaluări ceva mai mare. Totuşi, fiind mai simplă, metoda secțiunii de aur este mai cunoscută şi mai utilizată în algoritmii de optimizare.

3.3. METODE DE INTERPOLARE

Metodele de interpolare pentru minimizarea mono-dimensională, pentru determinarea lungimii pasului, sunt o alternativă foarte serioasă la metodele secțiunii de aur şi a lui Fibonacci. Principiul acestei clase de metode constă în a aproxima funcția $\varphi(\alpha) = f(x_k + \alpha d_k)$ printr-un polinom de gradul doi sau trei, care împreună cu derivata acestuia ia aceleaşi valori în anumite puncte, şi de a considera acea valoare a lui α care minimizează acest polinom. În general când funcția are proprietăți bune de analiticitate, de exemplu este continuu diferențiabilă, atunci metodele de interpolare sunt superioare metodei secțiunii de aur şi a lui Fibonacci. De obicei, în algoritmii de optimizare, metodele de interpolare sunt mai des utilizate decât metodele secțiunii de aur sau Fibonacci.

3.3.1. Metode de interpolare pătratică

Vom prezenta mai multe variante de interpolare pătratică în două sau trei puncte, pentru care vom demonstra şi rezultate privind convergența acestora.

A) Metoda interpolării pătratice în două puncte (varianta 1)

Date două puncte α_1 şi α_2, precum şi valorile funcției $\varphi(\alpha_1)$, $\varphi(\alpha_2)$ şi a derivatei $\varphi'(\alpha_1)$ (sau $\varphi'(\alpha_2)$), se construieşte polinomul de interpolare

$$q(\alpha) = a\alpha^2 + b\alpha + c$$

care satisface condițiile:

$$\begin{aligned} q(\alpha_1) &= a\alpha_1^2 + b\alpha_1 + c = \varphi(\alpha_1), \\ q(\alpha_2) &= a\alpha_2^2 + b\alpha_2 + c = \varphi(\alpha_2), \\ q'(\alpha_1) &= 2a\alpha_1 + b = \varphi'(\alpha_1). \end{aligned} \tag{3.3.1}$$

Notăm: $\varphi_1 = \varphi(\alpha_1)$, $\varphi_2 = \varphi(\alpha_2)$, $\varphi_1' = \varphi'(\alpha_1)$ şi $\varphi_2' = \varphi'(\alpha_2)$. Atunci, din (3.3.1) obținem

$$a = \frac{\varphi_1 - \varphi_2 - \varphi_1'(\alpha_1 - \alpha_2)}{-(\alpha_1 - \alpha_2)^2},$$

$$b = \varphi_1' + 2\frac{\varphi_1 - \varphi_2 - \varphi_1'(\alpha_1 - \alpha_2)}{(\alpha_1 - \alpha_2)^2}\alpha_1.$$

Deci

$$\bar{\alpha} = -\frac{b}{2a} = \alpha_1 + \frac{1}{2}\frac{\varphi_1'(\alpha_1 - \alpha_2)^2}{\varphi_1 - \varphi_2 - \varphi_1'(\alpha_1 - \alpha_2)} \tag{3.3.2}$$

Cu acestea, obținem următoarea formula iterativă:

$$\alpha_{k+1} = \alpha_k + \frac{1}{2}\frac{\varphi_k'(\alpha_k - \alpha_{k-1})^2}{\varphi_k - \varphi_{k-1} - \varphi_k'(\alpha_k - \alpha_{k-1})}, \tag{3.3.3}$$

unde $\varphi_k = \varphi(\alpha_k)$, $\varphi_{k-1} = \varphi(\alpha_{k-1})$ şi $\varphi_k' = \varphi'(\alpha_k)$, cunoscută ca *formula interpolării pătratice*.

Algoritmul este foarte simplu, odată determinat α_{k+1}, acesta este comparat cu α_k şi α_{k-1} astfel reducându-se intervalul de incertitudine. Această procedură continuă până când lungimea intervalului este mai mică decât o toleranță dată.

B) Metoda interpolării pătratice în două puncte (varianta 2)

Date două puncte α_1 şi α_2, precum şi valoarea funcție în unul dintre aceste puncte $\varphi(\alpha_1)$ (sau $\varphi(\alpha_2)$) şi valorile derivatelor $\varphi'(\alpha_1)$ şi $\varphi'(\alpha_2)$, se construieşte polinomul de interpolare $q(\alpha) = a\alpha^2 + b\alpha + c$ care satisface condițiile:

$$\begin{aligned} q(\alpha_1) &= a\alpha_1^2 + b\alpha_1 + c = \varphi(\alpha_1), \\ q'(\alpha_1) &= 2a\alpha_1 + b = \varphi'(\alpha_1), \end{aligned} \tag{3.3.4}$$

$$q'(\alpha_2) = 2a\alpha_2 + b = \varphi'(\alpha_2).$$

Ca mai sus obținem

$$\bar{\alpha} = -\frac{b}{2a} = \alpha_1 - \frac{\alpha_1 - \alpha_2}{\varphi_1' - \varphi_2'}\varphi_1'. \tag{3.3.5}$$

Deci, în acest caz obținem formula iterativă

$$\alpha_{k+1} = \alpha_k - \frac{\alpha_k - \alpha_{k-1}}{\varphi_k' - \varphi_{k-1}'}\varphi_k', \tag{3.3.6}$$

cunoscută ca *formula secantei*.

Observăm că dacă considerăm polinomul de interpolare Lagrange

$$L(\alpha) = \frac{(\alpha - \alpha_1)\varphi_2' - (\alpha - \alpha_2)\varphi_1'}{\alpha_2 - \alpha_1}$$

care interpolează $\varphi'(\alpha)$ în punctele α_1 şi α_2, atunci (3.3.5) se obține foarte simplu din $L(\alpha) = 0$.

Teorema 3.3.1. *Dacă* $\varphi : R \to R$ *este de trei ori continuă diferențiabilă şi există un* α^* *astfel încât* $\varphi'(\alpha^*) = 0$ *şi* $\varphi''(\alpha^*) \neq 0$*, atunci şirul* $\{\alpha_k\}$ *generat de (3.3.6) converge la* α^* *cu rata de convergență de ordin* $(1+\sqrt{5})/2 \cong 1.618$ *(secțiunea de aur).*

Demonstrație. Observăm că restul

$$R_2(\alpha) = \varphi'(\alpha) - L(\alpha) = \frac{1}{2}\varphi'''(\xi)(\alpha - \alpha_k)(\alpha - \alpha_{k-1}),\ \ \xi \in (\alpha_{k-1}, \alpha_k).$$

Pentru $\alpha = \alpha_{k+1}$, $L(\alpha_{k+1}) = 0$, deci

$$\varphi'(\alpha_{k+1}) = \frac{1}{2}\varphi'''(\xi)(\alpha_{k+1} - \alpha_k)(\alpha_{k+1} - \alpha_{k-1}),\ \xi \in (\alpha_{k-1}, \alpha_k).$$

Acum ținând seama de (3.3.6) obținem

$$\varphi'(\alpha_{k+1}) = \frac{1}{2}\varphi'''(\xi)\varphi_k'\varphi_{k-1}'\frac{(\alpha_k - \alpha_{k-1})^2}{(\varphi_k' - \varphi_{k-1}')^2},\ \xi \in (\alpha_{k-1}, \alpha_k).$$

Din teorema de medie rezultă că

$$\frac{\varphi_k' - \varphi_{k-1}'}{\alpha_k - \alpha_{k-1}} = \varphi''(\xi_0),\ \xi_0 \in (\alpha_{k-1}, \alpha_k).$$

În plus

$$\varphi_i' = \varphi_i' - \varphi'(\alpha^*) = (\alpha_i - \alpha^*)\varphi''(\xi_i),\ \xi_i \in (\alpha_i, \alpha^*),\ \ i = k-1, k, k+1.$$

Deci

$$\alpha_{k+1} - \alpha^* = \frac{1}{2}\frac{\varphi'''(\xi)\varphi''(\xi_k)\varphi''(\xi_{k-1})}{\varphi''(\xi_{k+1})[\varphi''(\xi_0)]^2}(\alpha_k - \alpha^*)(\alpha_{k-1} - \alpha^*).$$

Definim $e_i = \left|\alpha_i - \alpha^*\right|$, $i = k-1, k, k+1$, şi fie $0 < m_2 \le \left|\varphi'''(\alpha)\right| \le M_2$, $0 < m_1 \le \left|\varphi''(\alpha)\right| \le M_1$, $K_1 = m_2 m_1^2 / (2M_1^3)$ şi $K = M_2 M_1^2 / (2m_1^3)$. Atunci

$$K_1 \left|\alpha_k - \alpha^*\right| \left|\alpha_{k-1} - \alpha^*\right| \le \left|\alpha_{k+1} - \alpha^*\right| \le K \left|\alpha_k - \alpha^*\right| \left|\alpha_{k-1} - \alpha^*\right|.$$

Deoarece φ'' şi φ''' sunt continue în α^* rezultă că

$$\frac{\alpha_{k+1} - \alpha^*}{(\alpha_k - \alpha^*)(\alpha_{k-1} - \alpha^*)} \to \frac{1}{2} \frac{\varphi'''(\alpha^*)}{\varphi''(\alpha^*)}.$$

Deci

$$e_{k+1} = \left|\frac{\varphi'''(\eta_1)}{2\varphi''(\eta_2)}\right| e_k e_{k-1} \triangleq M e_k e_{k-1}, \quad \eta_1, \eta_2 \in (\alpha_{k-1}, \alpha_k), \quad M = \left|\varphi'''(\eta_1) / 2\varphi''(\eta_2)\right|.$$

Ca atare, există $\delta > 0$ astfel încât când punctele iniţiale $\alpha_0, \alpha_1 \in (\alpha^* - \delta, \alpha^* + \delta)$ şi $\alpha_0 \ne \alpha_1$, şirul $\{\alpha_k\}$ tinde la α^*.

Pentru a vedea rata de convergenţă fie $z_i = M e_i$, $y_i = \ln z_i$, pentru $i = k-1, k, k+1$. Atunci $z_{k+1} = z_k z_{k-1}$ şi $y_{k+1} = y_k + y_{k-1}$. Clar, şirul $\{y_k\}$ satisface condiţia Fibonacci care are ecuaţia caracteristică $t^2 - t - 1 = 0$ cu soluţiile $t_1 = (1+\sqrt{5})/2$ şi $t_2 = (1-\sqrt{5})/2$. Deci y_k se poate exprima sub forma $y_k = A t_1^k + B t_2^k$, $k = 1, 2, \ldots$, unde A şi B sunt coeficienţi determinaţi din condiţiile iniţiale. Când $k \to \infty$, $\ln z_k = y_k \cong A t_1^k$. Dar, deoarece

$$\frac{z_{k+1}}{z_k^{t_1}} \cong \frac{\exp(A t_1^{k+1})}{[\exp(A t_1^k)]^{t_1}} = 1,$$

rezultă că $e_{k+1} / e_k^{t_1} \cong M^{t_1 - 1}$ ceea ce arată că rata de convergenţă este chiar $t_1 = (1+\sqrt{5})/2 \cong 1.618$, adică secţiunea de aur. ■

Teorema arată că metoda bazată pe formula secantei este superliniar convergentă.

C) Metoda interpolării pătratice în trei puncte

Această metodă presupune că se cunosc trei puncte α_i, $i = 1, 2, 3$, precum şi valoarea funcţiei în acestea φ_i, $i = 1, 2, 3$. Condiţiile de interpolare sunt:

$$q(\alpha_i) = a\alpha_i^2 + b\alpha_i + c = \varphi_i, \quad i = 1, 2, 3 \tag{3.3.7}$$

Rezolvând acest sistem obţinem

$$a = -\frac{(\alpha_2 - \alpha_3)\varphi_1 + (\alpha_3 - \alpha_1)\varphi_2 + (\alpha_1 - \alpha_2)\varphi_3}{(\alpha_1 - \alpha_2)(\alpha_2 - \alpha_3)(\alpha_3 - \alpha_1)},$$

$$b = \frac{(\alpha_2^2 - \alpha_3^2)\varphi_1 + (\alpha_3^2 - \alpha_1^2)\varphi_2 + (\alpha_1^2 - \alpha_2^2)\varphi_3}{(\alpha_1 - \alpha_2)(\alpha_2 - \alpha_3)(\alpha_3 - \alpha_1)}.$$

Deci

$$\bar{\alpha} = -\frac{b}{2a} = \frac{1}{2}\frac{(\alpha_2^2 - \alpha_3^2)\varphi_1 + (\alpha_3^2 - \alpha_1^2)\varphi_2 + (\alpha_1^2 - \alpha_2^2)\varphi_3}{(\alpha_2 - \alpha_3)\varphi_1 + (\alpha_3 - \alpha_1)\varphi_2 + (\alpha_1 - \alpha_2)\varphi_3}, \qquad (3.3.8)$$

cunoscută ca *formula de interpolare pătratică în trei puncte*.
Analog ca mai sus (3.3.8) se poate obține din polinomul Lagrange de interpolare

$$L(\alpha) = \frac{(\alpha - \alpha_2)(\alpha - \alpha_3)}{(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)}\varphi_1 + \frac{(\alpha - \alpha_1)(\alpha - \alpha_3)}{(\alpha_2 - \alpha_1)(\alpha_2 - \alpha_3)}\varphi_2 + \frac{(\alpha - \alpha_1)(\alpha - \alpha_2)}{(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2)}\varphi_3$$

considerând $L'(\alpha) = 0$.

Algoritmul 3.3.1. (Căutare liniară prin interpolare pătratică în trei puncte)

Pasul 1.	Se consideră toleranța ε. Se determină trei puncte $\{\alpha_1, \alpha_2, \alpha_3\}$ între care se află α^*. Se calculează $\varphi_i = \varphi(\alpha_i)$, $i = 1, 2, 3$.
Pasul 2.	Utilizând (3.3.8) se calculează $\bar{\alpha}$.
Pasul 3.	Dacă $(\bar{\alpha} - \alpha_1)(\bar{\alpha} - \alpha_3) \geq 0$ se continuă cu pasul 4; altfel se execută pasul 5.
Pasul 4.	Din punctele $\alpha_1, \alpha_2, \alpha_3$ şi $\bar{\alpha}$ se determină un nou grup de trei puncte $\{\alpha_1, \alpha_2, \alpha_3\}$ cu care se execută pasul 2.
Pasul 5.	Dacă $\lvert\bar{\alpha} - \alpha_2\rvert < \varepsilon$, stop, altfel se continuă cu pasul 4. ♦

Următoarea teoremă arată că rata de convergența a algoritmului bazat pe formula de interpolare pătratică în trei puncte este numai 1.32, mai mică decât a celui bazat pe formula secantei [Sun şi Yuan, 2006].

Teorema 3.3.2. *Fie funcția $\varphi(\alpha)$ continuu diferențiabilă de patru ori cel puțin şi α^* astfel încât $\varphi'(\alpha^*) = 0$ şi $\varphi''(\alpha^*) \neq 0$. Atunci şirul $\{\alpha_k\}$ generat de formula (3.3.8) converge la α^* cu rata 1.32.*

Demonstrație. Utilizând polinomul de interpolare Lagrange putem scrie $\varphi(\alpha) = L(\alpha) + R_3(\alpha)$, unde

$$R_3(\alpha) = \frac{1}{6}\varphi'''(\xi)(\alpha - \alpha_1)(\alpha - \alpha_2)(\alpha - \alpha_3).$$

Deoarece $\varphi'(\alpha^*) = L'(\alpha^*) + R_3'(\alpha^*) = 0$, rezultă că

$$\varphi_1\frac{2\alpha^*-(\alpha_2+\alpha_3)}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)}+\varphi_2\frac{2\alpha^*-(\alpha_3+\alpha_1)}{(\alpha_2-\alpha_3)(\alpha_2-\alpha_1)}$$
$$+\varphi_3\frac{2\alpha^*-(\alpha_1+\alpha_2)}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)}+R_3'(\alpha^*)=0.$$

Observăm că (3.3.8) se poate scrie sub forma

$$\alpha_4=\frac{1}{2}\frac{\frac{\varphi_1(\alpha_2+\alpha_3)}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)}+\frac{\varphi_2(\alpha_3+\alpha_1)}{(\alpha_2-\alpha_3)(\alpha_2-\alpha_1)}+\frac{\varphi_3(\alpha_1+\alpha_2)}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)}}{\frac{\varphi_1}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)}+\frac{\varphi_2}{(\alpha_2-\alpha_3)(\alpha_2-\alpha_1)}+\frac{\varphi_3}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)}}.$$

Deci

$$\alpha^*-\alpha_4=\frac{1}{2}\frac{R_3'(\alpha^*)}{\frac{\varphi_1}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)}+\frac{\varphi_2}{(\alpha_2-\alpha_3)(\alpha_2-\alpha_1)}+\frac{\varphi_3}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)}}.$$

Definim $e_i=\alpha^*-\alpha_i$, $i=1,2,3,4$. Din relația de mai sus rezultă că

$$e_4[-\varphi_1(e_2-e_3)-\varphi_2(e_3-e_1)-\varphi_3(e_1-e_2)]$$
$$=-\frac{1}{2}R_3'(\alpha^*)(e_1-e_2)(e_2-e_3)(e_3-e_1).$$

Dar $\varphi'(\alpha^*)=0$, deci

$$\varphi_i=\varphi(\alpha^*)+\frac{1}{2}e_i^2\varphi''(\alpha^*)+O(e_i^3).$$

Neglijând termenii de ordin superior din relația de mai sus găsim

$$e_4=\frac{1}{\varphi''(\alpha^*)}R_3'(\alpha^*).$$

Din interpolarea Lagrange obținem

$$R_3'(\alpha)=\frac{1}{6}\varphi'''(\xi(\alpha))[(\alpha-\alpha_2)(\alpha-\alpha_3)+(\alpha-\alpha_1)(\alpha-\alpha_3)+(\alpha-\alpha_1)(\alpha-\alpha_2)]$$
$$+\frac{1}{24}\varphi^{(4)}(\eta)(\alpha-\alpha_1)(\alpha-\alpha_2)(\alpha-\alpha_3),$$

adică

$$R_3'(\alpha^*)=\frac{1}{6}\varphi'''(\xi(\alpha^*))(e_1e_2+e_2e_3+e_3e_1)+\frac{1}{24}\varphi^{(4)}(\eta)e_1e_2e_3.$$

Neglijând termenii de ordinul patru obținem

$$e_4=\frac{\varphi'''(\xi(\alpha))}{6\varphi''(\alpha^*)}(e_1e_2+e_2e_3+e_3e_1)=M(e_1e_2+e_2e_3+e_3e_1),$$

unde M este o constantă. În general

$$e_{k+2}=M(e_{k-1}e_k+e_ke_{k+1}+e_{k+1}e_{k-1}).$$

Deoarece când $e_k \to 0$, $e_{k+1} = O(e_k) = O(e_{k-1})$, rezultă că există $\bar{M} > 0$ astfel încât

$$|e_{k+2}| \le \bar{M}\,|e_k||e_{k-1}|,$$

adică

$$\bar{M}\,|e_{k+2}| \le \bar{M}\,|e_k|\,\bar{M}\,|e_{k-1}|.$$

Când $|e_i|$, $i = 1,2,3$, sunt suficient de mici astfel încât $\delta = max\{\bar{M}\,|e_1|, \bar{M}\,|e_2|, \bar{M}\,|e_3|\} < 1$, atunci $\bar{M}\,|e_4| \le \bar{M}\,|e_1|\,\bar{M}\,|e_2| \le \delta^2$. Punem $\bar{M}\,|e_k| \le \delta^{q_k}$, atunci

$$\bar{M}\,|e_{k+2}| \le \bar{M}\,|e_k|\,\bar{M}\,|e_{k-1}| \le \delta^{q_k}\delta^{q_{k-1}} \triangleq \delta^{q_{k+2}},$$

deci

$$q_{k+2} = q_k + q_{k-1}, \tag{3.3.9}$$

pentru $k \ge 2$, unde $q_1 = q_2 = q_3 = 1$. Ecuația caracteristică a lui (3.3.9) este $t^3 - t - 1 = 0$, care are o rădăcină reală $t_1 \cong 1.32$ şi două rădăcini complex conjugate $|t_2| = |t_3| < 1$. Soluția generală a lui (3.3.9) este de forma $q_k = At_1^k + Bt_2^k + Ct_3^k$, unde A, B, C sunt constante determinate din condițiile inițiale. Clar, când $k \to \infty$, atunci $q_{k+1} - t_1 q_k = Bt_2^k(t_2 - t_1) + Ct_3^k(t_3 - t_1) \to 0$. Deci, când k este suficient de mare, rezultă că $q_{k+1} - t_1 q_k \ge -0.1$. Dar $|e_k| \le (1/\bar{M})\delta^{q_k} \triangleq B_k$, pentru orice $k \ge 1$. Deci, când k este suficient de mare,

$$\frac{B_{k+1}}{B_k} = \frac{\delta^{q_{k+1}}/\bar{M}}{\delta^{t_1 q_k}/(\bar{M})^{t_1}} = \bar{M}^{1-t_1}\delta^{q_{k+1}-t_1 q_k} \le \delta^{-0.1}\bar{M}^{1-t_1},$$

care arată că rata de convergență este de ordinul $t_1 \cong 1.32$. ■

Observăm că metoda de căutare liniară prin interpolare pătratică în trei puncte are o rată de convergență mai mică decât a celei bazate pe interpolare pătratică cu formula secantei. Explicația constă în faptul că interpolarea pătratică în trei puncte nu utilizează informațiile date de derivata funcției φ în punctele din intervalul de căutare. Altfel spus, metoda nu ține seama de curbura funcției φ. În general în implementările avansate se utilizează căutarea liniară cu formula secantei.

3.3.2. Metode de interpolare cubică

Aceste metode aproximează funcția de minimizat $\varphi(\alpha)$ printr-un polinom cubic. Aceasta implică patru condiții de interpolare. Se cunosc mai multe variante. De exemplu, se pot utiliza valorile funcției în patru puncte, sau valorile funcției în trei puncte şi valoarea derivatei acesteia într-un punct, sau valoarea funcție şi a derivatei acesteia în două puncte. În general, interpolarea cubică are o rată de convergență mai bună decât interpolarea pătratică, dar aceasta implică calculul derivatelor şi deci computațional este mai costisitoare. În cele ce urmează

prezentăm interpolarea cubică care utilizează valorile funcţiei şi ale derivatei acesteia, fiecare în două puncte.

Fie deci două puncte a şi b în care cunoaştem valorile $\varphi(a)$ şi $\varphi(b)$, precum şi valorile derivatei $\varphi'(a)$ şi $\varphi'(b)$ cu care construim polinomul cubic de interpolare de forma

$$p(\alpha) = c_1(\alpha - a)^3 + c_2(\alpha - a)^2 + c_3(\alpha - a) + c_4, \tag{3.3.10}$$

unde c_i, $i = 1,2,3,4$ sunt coeficienţi care se determină din condiţiile:

$$p(a) = c_4 = \varphi(a),$$
$$p(b) = c_1(b-a)^3 + c_2(b-a)^2 + c_3(b-a) + c_4 = \varphi(b),$$
$$p'(a) = c_3 = \varphi'(a),$$
$$p'(b) = 3c_1(b-a)^2 + 2c_2(b-a) + c_3 = \varphi'(b).$$

După cum ştim, condiţiile suficiente de minim sunt

$$p'(\alpha) = 3c_1(\alpha - a)^2 + 2c_2(\alpha - a) + c_3 = 0 \tag{3.3.11}$$

şi

$$p''(\alpha) = 6c_1(\alpha - a) + 2c_2 > 0. \tag{3.3.12}$$

Rezolvând (3.3.11) obţinem

$$\alpha = a + \frac{-c_2 \pm \sqrt{c_2^2 - 3c_1c_3}}{3c_1}, \text{ dacă } c_1 \neq 0, \tag{3.3.13}$$

$$\alpha = a - \frac{c_3}{2c_2}, \text{ dacă } c_1 = 0. \tag{3.3.14}$$

Pentru a garanta condiţia suficientă (3.3.12), vom considera rădăcina cu semnul + din (3.3.13) pe care combinând-o cu (3.3.14) obţinem

$$\alpha - a = \frac{-c_2 + \sqrt{c_2^2 - 3c_1c_3}}{3c_1} = \frac{-c_3}{c_2 + \sqrt{c_2^2 - 3c_1c_3}}. \tag{3.3.15}$$

Când $c_1 = 0$, (3.3.15) este chiar (3.3.14). Atunci, minimul lui $p(\alpha)$ este

$$\bar{\alpha} = a - \frac{c_3}{c_2 + \sqrt{c_2^2 - 3c_1c_3}}, \tag{3.3.16}$$

care după cum se vede este exprimat în funcţie de c_1, c_2 şi c_3. Problema este că acesta trebuie exprimat în funcţie de $\varphi(a)$, $\varphi(b)$, $\varphi'(a)$ şi $\varphi'(b)$. Pentru aceasta fie:

$$s = 3\frac{\varphi(b) - \varphi(a)}{b - a}, \quad z = s - \varphi'(a) - \varphi'(b), \quad w^2 = z^2 - \varphi'(a)\varphi'(b).$$

Din condiţiile de interpolare găsim

$$s = 3\frac{\varphi(b) - \varphi(a)}{b - a} = 3[c_1(b-a)^2 + c_2(b-a) + c_3],$$

$$z = s - \varphi'(a) - \varphi'(b) = c_2(b-a) + c_3,$$
$$w^2 = z^2 - \varphi'(a)\varphi'(b) = (b-a)^2(c_2^2 - 3c_1c_3).$$

Atunci

$$(b-a)c_2 = z - c_3, \quad \sqrt{c_2^2 - 3c_1c_3} = \frac{w}{b-a},$$

şi deci

$$c_2 + \sqrt{c_2^2 - 3c_1c_3} = \frac{z + w - c_3}{b-a}. \tag{3.3.17}$$

Dar $c_3 = \varphi'(a)$ şi substituind (3.3.17) în (3.3.16) obținem

$$\bar{\alpha} - a = \frac{-(b-a)\varphi'(a)}{z + w - \varphi'(a)}$$

care se mai poate scrie ca:

$$\bar{\alpha} - a = \frac{-(b-a)\varphi'(a)\varphi'(b)}{(z+w-\varphi'(a))\varphi'(b)} = \frac{-(b-a)(z^2 - w^2)}{\varphi'(b)(z+w) - (z^2 - w^2)} = \frac{(b-a)(w-z)}{\varphi'(b) - z + w}.$$

Totuşi, relația de mai sus nu este potrivită implementării deoarece numitorul ei poate fi zero sau foarte aproape de zero. De aceea vom căuta o expresie echivalentă sub forma următoare.

$$\bar{\alpha} - a = \frac{-(b-a)\varphi'(a)}{z+w-\varphi'(a)} = \frac{(b-a)(w-z)}{\varphi'(b) - z + w} = \frac{(b-a)(-\varphi'(a) + w - z)}{\varphi'(b) - \varphi'(a) + 2w}$$
$$= (b-a)\left(1 - \frac{\varphi'(b) + z + w}{\varphi'(b) - \varphi'(a) + 2w}\right).$$

Deci

$$\bar{\alpha} = a + (b-a)\frac{w - \varphi'(a) - z}{\varphi'(b) - \varphi'(a) + 2w}. \tag{3.3.18}$$

În (3.3.18) numitorul $\varphi'(b) - \varphi'(a) + 2w \neq 0$. Mai mult, deoarece $\varphi'(a) < 0$ şi $\varphi'(b) > 0$, rezultă că $w^2 = z^2 - \varphi'(a)\varphi'(b) > 0$. Considerând $w > 0$, rezultă că numitorul $\varphi'(b) - \varphi'(a) + 2w > 0$. Utilizând aceleaşi argumente ca în cazul interpolării pătratice putem stabili rata de convergență a metodei de interpolare cubică. La fel ca în cazul teoremei 3.3.1, obținem

$$e_{k+1} = M(e_k e_{k-1}^2 + e_k^2 e_{k-1}),$$

unde M este o constantă. În acest caz se poate arăta că ecuația caracteristică este $t^2 - t - 2 = 0$, care are soluția pozitivă $t = 2$. Deci, metoda de interpolare cubică în două puncte este convergentă cu rata de convergență de ordinul 2.

Metodele prezentate în acest capitol, deşi au importanța lor, nu sunt utilizate în pachetele de optimizare fără restricții. De aceea nici nu am insistat în privința lor. *Metodele de căutare liniară utilizate în algoritmii de optimizare fără restricții sunt cele economice pe care le discutăm în capitolul 4.*

Martie 21, 2007