CURSUL 1
METODE NUMERICE PENTRU ECUATII OPERATORIALE
(titular de curs: prof. univ. dr. Bica Alexandru Mihai)

ECUATII OPERATORIALE. REZOLVAREA NUMERICA A ECUATIILOR SCALARE PRIN METODA APROXIMATIILOR SUCCESIVE

Spatii metrice

Definitia 1: Fie $X \neq \emptyset$. Perechea (X,d) se numeşte spaţiu metric dacă funcţia $d : X \times X \to \mathbb{R}_+$ are proprietăţile :
(d_1) $d(x,y) = 0 \Longleftrightarrow x = y$
(d_2) $d(y,x) = d(x,y), \quad \forall x, y \in X$
(d_3) $d(x,z) \leq d(x,y) + d(y,z), \quad \forall x, y, z \in X$.
O asemenea funcţie se numeşte metrică pe X.

Definitia 2: Fie (X,d) spaţiu metric. Un şir $(x_n)_{n\in\mathbb{N}} \subset X$ se numeşte :
(i) convergent către $x \in X$ dacă şi numai dacă $\forall \varepsilon > 0,\ \exists n(\varepsilon) \in \mathbb{N}$, astfel încât $d(x_n, x) < \varepsilon,\ \forall n \in \mathbb{N},\ n \geq n(\varepsilon)$.
(ii) fundamental, dacă şi numai dacă $\forall \varepsilon > 0,\ \exists n(\varepsilon) \in \mathbb{N}$, astfel încât $d(x_n, x_m) < \varepsilon,\ \forall n, m \in \mathbb{N},\ n, m \geq n(\varepsilon)$.

Observaţie : Într-un spaţiu metric orice şir convergent este fundamental, dar reciproca nu este adevărată în general.

Definitia 3: Un spaţiu metric se numeşte complet dacă orice şir fundamental al său este convergent.

Definitia 4: Fie $(V, \mathbb{R}; +, \cdot)$ un spaţiu vectorial real. O funcţie $\|\cdot\| : V \to \mathbb{R}_+$ se numeşte normă dacă are proprietăţile :
(n_1) $\|x\| = 0 \Longleftrightarrow x = 0_V$
(n_2) $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|, \quad \forall x \in X, \quad \forall \alpha \in \mathbb{R}$;
(n_3) $\|x + y\| \leq \|x\| + \|y\|, \quad \forall x, y \in X$.
Perechea $(V, \|\cdot\|)$ se numeşte spaţiu normat.

Observaţie : Fie $(V, \|\cdot\|)$ un spaţiu normat. Se arată că funcţia $d : V \times V \to \mathbb{R}_+$ definită prin $d(x,y) = \|x - y\|$ este o metrică pe X, iar spaţiul (X,d) se numeşte spaţiul metric generat de norma $\|\cdot\|$.

Definitia 5: Un spaţiu normat pentru care spaţiul metric generat este complet se numeşte spaţiu Banach.

Exemple : 1) Pentru $X = \mathbb{R}$, funcţia modul este o normă care genereaza metrica naturală pe $\mathbb{R}$. Cu această metrica $\mathbb{R}$ este spaţiu metric complet (aşadar este şi spaţiu Banach). În schimb, multimea numerelor raţionale dotată cu aceeastă metrică este spaţiu metric incomplet. Astfel, spre exemplu, şirul cu termenul general $x_n = \left(1 + \frac{1}{n}\right)^n$, $n \in \mathbb{N}^*$, este convergent în $\mathbb{R}$, deci este fundamental, dar nu este convergent în mulţimea numerelor raţionale deoarece limita sa este un număr iraţional şi transcendent.

2) Pe $X = \mathbb{R}^n$, sunt definite normele :

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$$

$$\|x\|_1 = |x_1| + |x_2| + ... + |x_n|$$
$$\|x\|_\infty = \max(|x_i| : i = \overline{1,n})$$

pentru $x = (x_1, x_2, ..., x_n)$, numite respectiv, norma euclidiană, norma Minkowski şi norma Cebîşev. Acestea vor genera pe $\mathbb{R}^n$ metricile euclidiană ρ, Minkowski δ şi Cebîşev, d_∞. În raport cu oricare dintre aceste norme $\mathbb{R}^n$ este spaţiu Banach.

3) Fie spaţiul de funcţii

$$X = \{f : [a,b] \to \mathbb{R},\ f \text{ continuă }\}.$$

Pe acest spaţiu se definesc metrica Cebîşev, $d_C : X \times X \to \mathbb{R}_+$

$$d_C(f,g) = \max\{|f(x) - g(x)| : x \in [a,b]\}$$

şi metrica Bielecki, $d_B : X \times X \to \mathbb{R}_+$

$$d_B(f,g) = \max\{|f(x) - g(x)| \cdot e^{-\tau|x-x_0|} : x \in [a,b]\}$$

pentru orice $f, g \in X$, unde $\tau > 0$ este constant, iar $x_0 \in [a,b]$ este fixat. Spaţiul metric (X, d_C) se noteazaa prin $C[a,b]$, iar $(X.d_B)$ prin $B[a,b]$. Ambele spaţii au metrica provenită din normă, induc convergenţa uniformă şi sunt complete. Metrica Bielecki a fost introdusă de A. Bielecki în 1956 pentru a servi la studiul ecuaţiilor integrale şi integro-diferenţiale Volterra prin tehnica punctului fix.

Definitia 6: Fie (X,d) spaţiu metric. Pentru $x \in X$ fixat şi $r > 0$, mulţimile

$$B(x;r) = \{y \in X : d(x,y) < r\}$$

şi

$$\overline{B}(x;r) = \{y \in X : d(x,y) \leq r\}$$

se numesc bila deschisă şi respectiv bila închisă cu centrul în x şi raza r. O mulţime $A \subset X$ se numeşte deschisă dacă $\forall x \in A$, $\exists r > 0$ astfel încât $B(x;r) \subset A$. Topologia astfel construită pe X se numeşte topologia indusă de metrica d.

Definitia 7: Fie (X,d) spaţiu metric. O aplicaţie $T : X \to X$ se numeşte :
(i) Lipschitziană (sau $L-$Lipschitz), dacă există $L > 0$ astfel încât

$$d(T(x), T(y)) \leq L \cdot d(x,y), \quad \forall x, y \in X$$

(ii) contracţie, dacă există $\alpha \in (0,1)$ astfel încât T este $\alpha-$Lipschitz
(iii) neexpansivă, dacă este $1-$Lipschitz
(iv) contractivă, dacă

$$d(T(x), T(y)) \leq d(x,y), \quad \forall x, y \in X, \quad x \neq y$$

(v) izometrie, dacă

$$d(T(x), T(y)) = d(x,y), \quad \forall x, y \in X.$$

Teorema de punct fix a lui Banach

Teorema (principiul de punct fix al lui Banach): Fie (X, d) un spaţiu metric complet şi $f : X \to X$ o aplicaţie pentru care există $\alpha \in (0, 1)$ astfel încât

$$d(f(x), f(y)) \leq \alpha d(x, y), \quad \forall x, y \in X. \tag{1}$$

Atunci :
(i) f are un unic punct fix x^* şi dacă x_0 este un element oarecare din X, considerând şirul

$$x_0, \quad x_1 = f(x_0), ... \quad x_n = f(x_{n-1}), ...$$

numit şirul aproximaţiilor succesive, acest şir converge către x^*;
(ii) are loc estimarea apriori,

$$d(x_n, x^*) \leq \frac{\alpha^n}{1-\alpha} d(x_0, x_1), \quad \forall n \in \mathbb{N}^*. \tag{2}$$

(iii) Dacă $g : X \to X$ este astfel încât $\exists \eta \in \mathbb{R}$, cu proprietatea $d(f(x), g(x)) \leq \eta$, $\forall x \in X$ şi considerăm şirul $y_n = g^n(x_0)$, atunci

$$d(y_n, x^*) \leq \frac{\eta}{1-\alpha} + \frac{\alpha^n}{1-\alpha} d(x_0, x_1).$$

Această teoremă pune în evidenţă şirul aproximaţiilor succesive $(x_n)_{n \in \mathbb{N}}$ care converge către unicul punct fix al contracţiei f, definind o procedură iterativă, numită procedura Picard, iar rata de convergenţă a acestui şir este dată în inegalitatea

$$d(x_n, x^*) \leq a \cdot d(x_{n-1}, x^*) \leq a^n \cdot d(x_0, x^*), \quad \forall n \in \mathbb{N}^*.$$

Aceasta ne arată că rata de convergenţă a iteraţiilor Picard este liniară pentru orice contracţie. Deasemenea, procedura Picard este cel puţin atât de rapidă ca seria geometrică $\sum a^n$.

Ecuatii operatoriale

Definitie: Fie X, Y doua multimi nevide si consideram functia $f : X \to Y$. Pentru $y \in Y$ fixat, se numeste ecuatie operatoriala problema determinarii multimii

$$\{x \in X : y = f(x)\}.$$

Exemple de ecuatii operatoriale

1. Daca $X = Y = \mathbb{R}$ si $f : X \to Y$ este definita prin $f(x) = ax$, unde a este constanta, ecuatia operatoriala $y = ax$ se numeste ecuatie liniara.

2. Daca $X = Y = \mathbb{R}$ si $f : A \subset X \to Y$ este o functie continua atunci ecuatia operatoriala $y = f(x)$ se numeste ecuatie scalara neliniara. Aici se incadreaza si ecuatiile algebrice de grad $n \geq 2$, precum si ecuatiile exponentiale, logaritmice, irationale, trigonometrice.

3. Daca $X = Y = \mathbb{R}^n$ si fie $A \in M_n(\mathbb{R})$ o matrice, $b = (b_1, ..., b_n) \in \mathbb{R}^n$ un vector dat, iar functia $f : X \to Y$ este definita prin $f(x) = A \cdot x$, pentru $x = (x_1, ..., x_n) \in \mathbb{R}^n$. Ecuatia operatoriala

$$A \cdot x = b \Longleftrightarrow \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ ... \\ b_n \end{pmatrix}$$

se numeste sistem de ecuatii liniare.

4. Daca $X = Y = C[a, b]$, $f \in C[a, b]$ si $g : [a, b] \times [a, b] \times \mathbb{R} \to \mathbb{R}$ sunt functii date, iar operatorul $F : X \to Y$ este definit prin

$$F(x(t)) = f(t) + \int_a^b g(t, s, x(s))\, ds, \quad t \in [a, b]$$

pentru fiecare $x \in C[a, b]$, atunci ecuatia operatoriala $f = F(x)$ se numeste ecuatie integrala Fredholm de speta a doua. Cand operatorul este definit prin

$$F(x(t)) = f(t) + \int_a^t g(t, s, x(s))\, ds, \quad t \in [a, b]$$

ecuatia operatoriala se va numi ecuatie integrala Volterra de speta a doua.

5. Daca $X = C^n[a, b] = \{f \in C[a, b] : f$ este derivabila de n ori cu $f^{(n)}$ continua$\}$, $Y = C[a, b]$, iar $f : [a, b] \times \mathbb{R}^n \to \mathbb{R}$ este o functie continua, atunci pentru operatorul $F : X \to Y$ definit prin

$$F(x(t)) = x^{(n)}(t) - f\left(t, x(t), x'(t), ..., x^{(n-1)}(t)\right), \quad t \in [a, b]$$

ecuatia operatoriala $F(x) = 0$ se numeste ecuatie diferentiala de ordinul n, si poate fi scrisa sub forma

$$x^{(n)}(t) = f\left(t, x(t), x'(t), ..., x^{(n-1)}(t)\right), \quad t \in [a, b].$$

Rezolvarea numerica a ecuatiilor neliniare scalare

Consideram problema determinarii radacinilor reale ale ecuaţiei $f(x) = 0$, unde $f : I \to \mathbb{R}$, cu $I \subset \mathbb{R}$ un interval. Putine sunt situatiile in care radacinile reale ale acestei ecuatii se pot determina exact. După cum se stie, chiar in cazul particular al ecuatiilor algebrice cu coeficienti reali radacinile nu se pot determina exact pentru ecuatii de grad mai mare decat patru (in conformitate cu teorema lui Abel-Ruffini). Asadar, are sens si este pe deplin justificata problema aproximarii radacinilor reale ale ecuatiei $f(x) = 0$.

Separarea radacinilor prin metoda sirului lui Rolle

Primul pas in aproximarea radacinilor reale ale ecuatiei $f(x) = 0$ este izolarea fiecarei radacini reale intr-un interval al axei reale de lungime cat mai mica. Un procedeu de izolare a radacinilor este dat de metoda sirului lui Rolle, iar pentru micsorarea fiecarui interval astfel obtinut se continua cu aplicarea catorva pasi din algoritmul injumatatirii intervalului. Odata izolata radacina intr-un interval de lungime mica se aplica pentru aproximarea acesteia o metoda iterativa care furnizeaza un sir de aproximatii ale radacinii. Oprirea algoritmului se realizeaza la acel termen al sirului care aproximeaza radacina cu precizia dorita.

Teorema: Fie $I \subset \mathbb{R}$ un interval si $f : I \to \mathbb{R}$ o functie continuă. Atunci, $\forall a, b \in I$, $a < b$, cu $f(a) \cdot f(b) < 0$ exista $x_0 \in (a, b)$ astfel incat $f(x_0) = 0$. Daca, in plus, f este strict monotona pe (a, b) atunci ecuatia $f(x) = 0$ are in intervalul (a, b) o solutie unica.

Demonstratie: Deoarece functia f este continua, va avea proprietatea lui Darboux si intrucat $f(a) \cdot f(b) < 0$ deducem ca $f(a) < 0 < f(b)$ sau $f(b) < 0 < f(a)$. Atunci intervalul $[f(a), f(b)]$ sau $[f(b), f(a)]$ contine pe zero si din proprietatea lui Darboux rezulta ca exista $x_0 \in (a, b)$ astfel incat $f(x_0) = 0$. Daca functia f este si strict monotona atunci va fi injectiva si nu poate avea mai mult de o solutie in intervalul (a, b).

Definitie: Fie $I = (a, b)$ cu a, b finite sau infinite, $a < b$ si $f : I \to \mathbb{R}$ o functie derivabila astfel incat functia f' este continua si se anuleaza intr-un numar finit de puncte distincte $x_1 < x_2 < ... < x_{n-1}$. Atunci sirul

$$\lim_{x \to a, x > a} f(x), \quad f(x_1), ..., f(x_{n-1}), \quad \lim_{x \to b, x < b} f(x)$$

se numeste sirul lui Rolle asociat functiei f.

Teorema: Fie $I = (a, b)$ cu a, b finite sau infinite, $a < b$ si $f : I \to \mathbb{R}$ o functie derivabila astfel incat functia f' este continua si se anuleaza intr-un numar finit de puncte distincte $x_1 < x_2 < ... < x_{n-1}$. Daca notam $x_0 = a, x_n = b$ si daca exista $i \in \{1, ..., n\}$ astfel incat $f(x_{i-1}) \cdot f(x_i) < 0$ (respectiv $f(x_{i-1}) \cdot f(x_i) > 0$), atunci in intervalul (x_{i-1}, x_i) exista o singura radacina a ecuatiei $f(x) = 0$ (respectiv, in intervalul (x_{i-1}, x_i) nu exista nici o radacina a ecuatiei $f(x) = 0$).

Demonstratie: Deoarece pe intervalele (x_{i-1}, x_i), $i = \overline{1, n}$ functia f' nu se anuleaza, in virtutea continuitatii sale avem $f'(x) \neq 0$, $\forall x \in (x_{i-1}, x_i)$ si astfel f este strict monotona pe intervalele (x_{i-1}, x_i), $i = \overline{1, n}$. Daca $f(x_{i-1}) \cdot f(x_i) < 0$ atunci in virtutea teoremei precedente ecuatia $f(x) = 0$ are in intervalul (x_{i-1}, x_i) exact o solutie. Daca $f(x_{i-1}) \cdot f(x_i) > 0$ atunci $f(x_{i-1}), f(x_i) > 0$ sau $f(x_{i-1}), f(x_i) < 0$ si deoarece f este strict monotona pe intervalul (x_{i-1}, x_i) rezulta ca nu se poate anula in interiorul intervalului (x_{i-1}, x_i).

Observatia 1. Teorema precedenta se poate reformula astfel: Intre doua radacini consecutive ale derivatei exista cel mult o radacina a functiei. Acest enunt se mai numeste consecinta teoremei lui Rolle, iar aplicarea sa creaza metoda sirului lui Rolle de separare a radacinilor.

Consecinta: In conditiile teoremei precedente, numarul variatiilor de semn din sirul lui Rolle (din care am exclus termenii nuli) este egal cu numarul radacinilor reale ale ecuatiei $f(x) = 0$ situate in multimea $(a, b) \setminus \{x_1, ..., x_{n-1}\}$.

Observatia 2. Odata cu aplicarea teoremei precedente, fiecare radacina reala a ecuatiei $f(x) = 0$ este izolata intr-un interval de forma (x_{i-1}, x_i) pentru care avem inegalitatea $f(x_{i-1}) \cdot f(x_i) < 0$ pentru doi termeni consecutivi din sirul lui Rolle. Acest interval se poate micsora prin aplicarea in continuare a metodei injumatatirii intervalului. Scopul acestei micsorarii a intervalului este acela de a izola radacina intr-un interval pe care sa poata fi indeplinite conditiile de convergenta a metodei iterative care se va aplica in continuare pentru a aproxima cu acuratete cat mai mare radacina ecuatiei.

Metoda tangentei

Aceasta metoda iterativa de aproximare a radacinilor ecuatiei $f(x) = 0$ provine din aproximarea locala liniara a graficului unei functii continue. Astfel, daca dorim sa aproximam printr-un segment de dreapta graficul unei functii in vecinatatea unui punct avem doua alternative: tangenta la grafic in punctul respectiv si coarda la grafic cu unul din capete in acest punct. Luand prima varianta obtinem metoda tangentei datorata lui Newton si luand a doua varianta obtinem metoda coardei.

In cazul metodei tangentei se obtine un sir $(x_n)_{n\in\mathbb{N}}$ de aproximatii ale radacinii x^* cu $x_n \in (a, b)$, $\forall n \in \mathbb{N}^*$, dat in mod recurent prin

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, \quad n \in \mathbb{N}^*. \tag{3}$$

Pentru convergenta catre x^* a sirului dat prin recurenta in (3) are loc:

Teorema 1 (de existenta, unicitate si convergenta): Fie $f : [a, b] \to \mathbb{R}$ o functie de clasa $C^2[a, b]$ cu $f'(x) \neq 0, f''(x) \neq 0, \forall x \in [a, b]$ si $f(a) \cdot f(b) < 0$. Atunci ecuatia $f(x) = 0$ are o unica solutie in intervalul $[a, b]$, obtinuta ca limita a sirului $(x_n)_{n\in\mathbb{N}}$ dat in mod recurent in (3) unde $x_0 \in [a, b]$ este fixat arbitrar astfel incat $f(x_0) \cdot f''(x_0) > 0$.

Demonstratie: Din continuitatea functiei f si din conditia $f(a) \cdot f(b) < 0$ deducem ca ecuatia $f(x) = 0$ are o solutie in intervalul $[a, b]$. Deoarece $f'(x) \neq 0, \forall x \in [a, b]$ rezulta ca functia f este strict monotona si atunci ecuatia $f(x) = 0$ are o unica solutie in intervalul $[a, b]$. Fie aceasta $x^* \in (a, b)$. Avand $f(a) \cdot f(b) < 0$ deducem ca ori $f(a) < 0, f(b) > 0$, ori $f(b) < 0, f(a) > 0$. Presupunem ca suntem in primul caz (in celalalt caz demonstratia se realizeaza analog). Atunci f este strict crescatoare si $f'(x) > 0, \forall x \in [a, b]$. Din conditia $f(x_0) \cdot f''(x_0) > 0$ deducem ca ori $f(x_0) > 0, f''(x_0) > 0$, ori $f(x_0) < 0, f''(x_0) < 0$. Presupunem ca suntem in prima situatie (in cealalta situatie demonstratia se realizeaza analog). Atunci $f''(x) > 0, \forall x \in [a, b]$ si $a < x^* < x_0 \leq b$. Vom arata prin inductie ca pentru termenii sirului dat in (3) avem $x_n > x^*$, $\forall n \in \mathbb{N}^*$.

Deoarece $x_0 > x^*$, vom presupune ca $x_n > x^*$ si vom arata ca $x_{n+1} > x^*$. Folosind formula lui Taylor pentru x_n si x^* avem,

$$0 = f(x^*) = f(x_n) + f'(x_n)(x^* - x_n) + \frac{(x^* - x_n)^2 \cdot f''(\xi)}{2} \tag{4}$$

si deoarece $f''(\xi) > 0$, deducem ca $f(x_n) + f'(x_n)(x^* - x_n) < 0$, adica,

$$x^* < x_n - \frac{f(x_n)}{f'(x_n)} = x_{n+1}.$$

Deci, $x_n > x^*$, $\forall n \in \mathbb{N}$. Din faptul ca functia f este strict crescatoare rezulta ca $f(x_n) > f(x^*) = 0$, $\forall n \in \mathbb{N}$ si atunci

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} < x_n, \forall n \in \mathbb{N} \tag{5}$$

adica sirul $(x_n)_{n\in\mathbb{N}}$ este strict descrescator si in plus, $x^* < x_n < x_0, \quad \forall n \in \mathbb{N}^*$, Prin urmare, acest sir este monoton si marginit, deci convergent. Fie z limita sa. Prin trecere la limita in egalitatea (3), in virtutea continuitatii funcţiilor f si f' se obtine,

$$z = z - \frac{f(z)}{f'(z)}$$

adica $f(z) = 0$. Datorita unicitatii solutiei ecuatiei $f(x) = 0$ deducem $z = x^*$. Deci, $\lim\limits_{n\to\infty} x_n = x^*$.

Datorita convergentei mentionate in teorema precedenta putem aproxima solutia x^* cu termenii sirului $(x_n)_{n\in\mathbb{N}}$ dat in (3). Asupra erorii acestei aproximari avem:

Consecinta 2 (estimarea erorii): In conditiile teoremei precedente au loc inegalitatile:

$$|x^* - x_n| \le \frac{|f(x_n)|}{m_1}, \quad \forall n \in \mathbb{N}^* \tag{6}$$

$$|x^* - x_n| \le \frac{M_2}{2m_1} \cdot (x_n - x_{n-1})^2, \quad \forall n \in \mathbb{N}^* \tag{7}$$

$$|x^* - x_n| \le \frac{M_2}{2m_1} \cdot (x^* - x_{n-1})^2, \quad \forall n \in \mathbb{N}^* \tag{8}$$

unde

$$m_1 = \min_{x\in[a,b]} |f'(x)|, \qquad M_2 = \max_{x\in[a,b]} |f''(x)|.$$

Demonstratie: Fie $x \in [a,b]$ fixat arbitrar. Din teorema cresterilor finite a lui Lagrange rezulta ca exista c cuprins intre x si x^* astfel incat

$$f(x^*) - f(x) = (x^* - x) \cdot f'(c)$$

si atunci

$$|x^* - x| = \frac{|f(x)|}{|f'(c)|} \le \frac{|f(x)|}{m_1}.$$

De aici rezulta,

$$|x^* - x_n| \leq \frac{|f(x_n)|}{m_1}, \quad \forall n \in \mathbb{N}^*.$$

Aplicand formula lui Taylor pentru x_n si x_{n-1} rezulta

$$f(x_n) = f(x_{n-1}) + f'(x_{n-1})(x_n - x_{n-1}) + \frac{(x_n - x_{n-1})^2 \cdot f''(c)}{2}$$

cu c cuprins intre x_{n-1} si x_n. Din formula (3) deducem ca

$$f(x_{n-1}) + f'(x_{n-1})(x_n - x_{n-1}) = 0$$

si atunci

$$f(x_n) = \frac{(x_n - x_{n-1})^2 \cdot f''(c)}{2}$$

ceea ce conduce la,

$$|x^* - x_n| \leq \frac{|f(x_n)|}{m_1} = \frac{\left|(x_n - x_{n-1})^2 \cdot f''(c)\right|}{2m_1} \leq \frac{M_2}{2m_1} \cdot (x_n - x_{n-1})^2, \quad \forall n \in \mathbb{N}^*.$$

Atunci, din

$$0 = f(x^*) = f(x_n) + f'(x_n)(x^* - x_n) + \frac{(x^* - x_n)^2 \cdot f''(\xi)}{2}$$

deducem

$$x^* - x_n = -\frac{f(x_n)}{f'(x_n)} - \frac{(x^* - x_n)^2 \cdot f''(\xi)}{2f'(x_n)}$$

iar din aceasta obtinem

$$x^* - x_{n+1} = x^* - x_n + \frac{f(x_n)}{f'(x_n)} = -\frac{(x^* - x_n)^2 \cdot f''(\xi)}{2f'(x_n)}$$

si atunci,

$$|x^* - x_{n+1}| \leq \frac{|f''(\xi)|}{2|f'(x_n)|} \cdot (x^* - x_n)^2 \leq \frac{M_2}{2m_1} \cdot (x^* - x_n)^2, \quad \forall n \in \mathbb{N}.$$

Definitie: Sirul iterativ $(x_n)_{n \in \mathbb{N}^*}$ cu $\lim_{n \to \infty} x_n = x^*$ are ordinul de convergenta $p \in \mathbb{R}$, $p \geq 1$, daca exista $c > 0$ astfel incat

$$\lim_{n \to \infty} \frac{|x^* - x_{n+1}|}{|x^* - x_n|^p} = c.$$

Teorema 3 (ordinul de convergenta): Metoda iterativa a tangentei are ordinul de convergenta $p = 2$.

Demonstratie: Considerand un termen x_n al sirului recurent $(x_n)_{n\in\mathbb{N}}$ construit in baza formulei recurente (3), aplicam formula lui Taylor functiei $f \in C^2[a,b]$ in jurul punctului x_n si obtinem:

$$f(x^*) = f(x_n) + f'(x_n)\cdot(x^* - x_n) + \frac{f''(c_n)}{2}\cdot(x^* - x_n)^2$$

unde punctul c_n este situat in intervalul (a,b) intre x^* si x_n.

Tinand seama de faptul ca $f(x^*) = 0$ deducem,

$$f(x_n) + f'(x_n)\cdot(x^* - x_n) = -\frac{f''(c_n)}{2}\cdot(x^* - x_n)^2$$

si impartind cu $f'(x_n) \neq 0$ vom obtine

$$\frac{f(x_n)}{f'(x_n)} + x^* - x_n = -\frac{f''(c_n)}{2f'(x_n)}\cdot(x^* - x_n)^2.$$

Deoarece

$$x^* - x_{n+1} = x^* - x_n + \frac{f(x_n)}{f'(x_n)}$$

va rezulta

$$x^* - x_{n+1} = -\frac{f''(c_n)}{2f'(x_n)}\cdot(x^* - x_n)^2$$

si atunci

$$\frac{|x^* - x_{n+1}|}{|x^* - x_n|^2} = \frac{|f''(c_n)|}{2|f'(x_n)|}.$$

Intrucat $\lim\limits_{n\to\infty} x_n = x^*$, din faptul ca c_n este situat intre x^* si x_n , deducem ca $\lim\limits_{n\to\infty} c_n = x^*$. Folosind continuitatea functiilor f' si f'', trecand la limita vom obtine

$$\lim_{n\to\infty} \frac{|x^* - x_{n+1}|}{|x^* - x_n|^2} = \frac{|f''(x^*)|}{2|f'(x^*)|} > 0.$$

Observatia 1. Estimarea aposteriori (7) ne arata ca ordinul erorii acestei metode este 2, adica, $|x^* - x_n| \leq K\cdot(x_n - x_{n-1})^2$, $\forall n \in \mathbb{N}^*$, unde K este o constanta ce depinde doar de functia f si de intervalul $[a,b]$ (nedepinzand de termenii sirului), si ne permite sa stabilim un criteriu de oprire a algoritmului, util implementarii sale.

Aproximarea radicalilor

Observatia 2. Metoda tangentei a lui Newton permite si construirea unui algoritm de aproximare a radicalilor aritmetici. Astfel, folosind functia $f : (0,b) \to \mathbb{R}$, $f(x) = x^k - c$, cu $c > 0, k \in \mathbb{N}^*, k \geq 2$, unde $b = \max\{1, c\}$ si luand

$x_0 = b$, avem $f'(x_0) > 0, f''(x_0) > 0$ si $f(x_0) > 0$. Se obtine faptul ca sirul dat recurent prin

$$x_n = x_{n-1} - \frac{x_{n-1}^k - c}{k x_{n-1}^{k-1}} = \frac{1}{k}\left[(k-1)\, x_{n-1} + \frac{c}{x_{n-1}^{k-1}}\right], \quad n \in \mathbb{N}^*$$

este strict descrescator, marginit si convergent, $\lim\limits_{n\to\infty} x_n = x^*$. Trecand la limita in relatia de recurenta obtinem

$$x^* = x^* - \frac{(x^*)^k - c}{k\,(x^*)^{k-1}} \Longrightarrow (x^*)^k - c = 0 \Longrightarrow x^* = \sqrt[k]{c}.$$

Deci, $\lim\limits_{n\to\infty} x_n = \sqrt[k]{c}$. Pentru $k = 2$ obtinem sirul

$$x_n = x_{n-1} - \frac{x_{n-1}^2 - c}{2x_{n-1}} = \frac{1}{2}\left(x_{n-1} + \frac{c}{x_{n-1}}\right), \quad n \in \mathbb{N}^*$$

adica cunoscuta formula a lui Heron de aproximare a radacinii patrate, iar pentru $n = 3$ se obtine,

$$x_n = x_{n-1} - \frac{x_{n-1}^3 - c}{3x_{n-1}^2} = \frac{2}{3} x_{n-1} + \frac{c}{3x_{n-1}^2}, \quad n \in \mathbb{N}^*.$$

Observatia 3. Daca $f' > 0, f'' > 0$, sau $f' < 0, f'' < 0$ atunci sirul $(x_n)_{n\in\mathbb{N}}$ dat in (3) este strict descrescator convergent catre solutia ecuatiei, iar daca $f' < 0, f'' > 0$ sau $f' > 0, f'' < 0$ atunci sirul $(x_n)_{n\in\mathbb{N}}$ dat in (3) este strict crescator convergent catre solutia ecuatiei.
Algoritmul a fost dat pentru cazul în care soluţia ecuaţiei a fost izolată în intervalul $[a, b]$, iar funcţiile f' şi f'' au semn constant pe acest interval. Alegerea lui x_0 la pasul 1 este evidentă,

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

deoarece acoperă toate situaţiile posibile. Această alegere va conduce la convergenţa metodei.

Observatia 4. Estimarea aposteriori (7) genereaza urmatorul criteriu practic de oprire a algoritmului: pentru $\varepsilon > 0$ dat, sa se determine primul numar natural n pentru care

$$|x_n - x_{n-1}| < \varepsilon,$$

algoritmul oprindu-se la această iteratie.