CURS #13

CONȚINUTUL CURSULUI #13:

Fie $f : [a, b] \longrightarrow \mathbb{R}$ o funcție integrabilă și fie

$$I(f) := \int_a^b f(x)\,\mathrm{d}x \tag{1}$$

Definiția (VII.1.)

Se numește formulă de cuadratură a lui f o formulă de aproximare a integralei (1) de forma

$$\boxed{I_n(f) := \sum_{k=1}^{n+1} w_k\, f(x_k)} \tag{2}$$

unde x_k, $k = \overline{1, n+1}$ sunt astfel încât $a \le x_1 < x_2 < \ldots < x_{n+1} \le b$. $w_k \in \mathbb{R}$, $k = \overline{1, n+1}$, se numesc coeficienții/ponderile cuadraturii (2), iar x_k, $k = \overline{1, n+1}$ se numesc nodurile cuadraturii (2).

Definiția (VII.2.)

Mărimea $e_t(f)$ definită conform formulei

$$e_t(f) := I(f) - I_n(f) \tag{3}$$

se numește eroarea cuadraturii (2) a lui f.

Considerăm funcția $f : [a, b] \longrightarrow \mathbb{R}$ astfel încât $f \in C^{n+1}[a, b]$. Fie $P_n : [a, b] \longrightarrow \mathbb{R}$ polinomul de interpolare Lagrange:

$$P_n(x) = \sum_{k=1}^{n+1} L_{n,k}(x)\, f(x_k)\,, \quad x \in [a, b]$$

cu $L_{n,k}$ funcțiile de bază:

$$L_{n,k}(x) = \prod_{\substack{i=1\\ i\neq k}}^{n+1} \frac{x - x_i}{x_k - x_i}\,, \quad x \in [a, b]\,, \quad k = \overline{1, n+1}$$

Formula de cuadratură a lui f devine în acest caz:

$$\begin{aligned} I_n(f) &= \int_a^b P_n(x)\,\mathrm{d}x = \int_a^b \sum_{k=1}^{n+1} L_{n,k}(x)\, f(x_k)\,\mathrm{d}x \\ &= \sum_{k=1}^{n+1} \underbrace{\left(\int_a^b L_{n,k}(x)\,\mathrm{d}x\right)}_{=:w_k} f(x_k) \end{aligned}$$

sau

$$I_n(f) = \sum_{k=1}^{n+1} w_k f(x_k) \tag{4}$$

unde ponderile cuadraturii (4) sunt date de:

$$w_k = \int_a^b L_{n,k}(x)\,\mathrm{d}x\,, \quad k = \overline{1, n+1} \tag{5}$$

VII.2. Formulele de cuadratură Newton-Cotes

Dacă nodurile cuadraturii (2) sunt echidistante și $x_1 = a, x_{n+1} = b$ atunci formula (2) se numește formula de cuadratură Newton - Cotes închisă cu $(n+1)$ noduri/puncte. În acest caz avem:

$$\begin{cases} x_1 = a \\ h = \dfrac{b-a}{n} \\ x_i = a + (i-1)h, i = \overline{1, n+1} \end{cases} \tag{6}$$

Dacă nodurile cuadraturii (2) sunt echidistante și $x_1 > a, x_{n+1} < b$ atunci formula (2) se numește formula de cuadratură Newton - Cotes deschisă cu $(n+1)$ noduri/puncte. În acest caz vom considera discretizarea de forma:

$$\begin{cases} x_0 = a, x_{n+2} = b \\ h = \dfrac{b-a}{n+2} \\ x_i = a + ih, i = \overline{0, n+2} \end{cases} \tag{7}$$

Menționăm că pentru ambele metode formula de cuadratură rămâne de forma (4) cu ponderile date de (5).
Fie următoarele schimbări de variabile corespunzătoare celor două formule (închisă și deschisă):

(a) S.V. pentru formula de cuadratură Newton - Cotes închisă:

$$x = a + h(t-1), \quad t \in [1, n+1]; \quad \mathrm{d}x = h\,\mathrm{d}t \tag{8}$$

(b) S.V. pentru formula de cuadratură Newton - Cotes deschisă:

$$x = a + ht, \quad t \in [0, n+2]; \quad \mathrm{d}x = h\,\mathrm{d}t \tag{9}$$

În cazul schimbării de variabilă pentru formula de cuadratura Newton - Cotes închisă avem:

$$\begin{aligned} L_{n,k}(x) &= \prod_{\substack{i=1 \\ i \neq k}}^{n+1} \frac{x - x_i}{x_k - x_i} = \prod_{\substack{i=1 \\ i \neq k}}^{n+1} \frac{(a + h(t-1)) - (a + h(i-1))}{(a + h(k-1)) - (a + h(i-1))} \\ &= \prod_{\substack{i=1 \\ i \neq k}}^{n+1} \frac{t-i}{k-i}, \quad x \in [a, b], \quad k = \overline{1, n+1} \end{aligned}$$

Coeficienții/ponderile $w_k, k = \overline{1, n+1}$ cuadraturii Newton-Cotes se calculează după cum urmează:

(a) Pentru formula de cuadratură Newton - Cotes închisă avem

$$w_k = \int_a^b L_{n,k}(x)\,\mathrm{d}x = h \int_1^{n+1} \prod_{\substack{i=1 \\ i \neq k}}^{n+1} \frac{t-i}{k-i}\,\mathrm{d}t,$$

(b) Pentru formula de cuadratură Newton - Cotes deschisă avem

$$w_k = \int_a^b L_{n,k}(x)\,\mathrm{d}x = h \int_0^{n+2} \prod_{\substack{i=1 \\ i \neq k}}^{n+1} \frac{t-i}{k-i}\,\mathrm{d}t$$

VII.2.1. Formula de cuadratură a trapezului.

Considerăm cazul cuadraturii Newton-Cotes închisă (n=1).
Nodurile cuadraturii sunt:

$$x_1 = a, \quad x_2 = b, \quad h = b - a$$

Formula de cuadratură este:

$$I_1(f) = w_1 f(x_1) + w_2 f(x_2) = w_1 f(a) + w_2 f(b)$$

Ponderile cuadraturii sunt:

$$\begin{aligned} w_1 &= h \int_1^2 \frac{t-2}{-1}\,\mathrm{d}t = \frac{h}{2} \\ w_2 &= h \int_1^2 (t-1)\,\mathrm{d}t = \frac{h}{2} \end{aligned}$$

Astfel, obținem formula de cuadratură a trapezului:

$$I_1(f) = h\left[\frac{1}{2} f(a) + \frac{1}{2} f(b)\right] = \frac{b-a}{2}\left[f(a) + f(b)\right] \tag{10}$$

Estimarea erorii de cuadratură a trapezului:
Dacă $f \in C^2[a,b]$, se poate demonstra că

$$e_t(f) = I(f) - I_1(f) = -\frac{f''(\xi)}{12} h^3 = O(h^3), \quad \text{cu} \quad \xi \in (a,b)$$

VII.2.2 Formula de cuadratură Simpson.

Considerăm cazul cuadraturii Newton-Cotes închisă ($n = 2$). Nodurile cuadraturii sunt:

$$x_1 = a, \quad x_2 = \frac{a+b}{2} = a + h, \quad x_3 = b = a + 2h, \quad h = \frac{b-a}{2}$$

Formula de cuadratură este:

$$I_2(f) = w_1 f(x_1) + w_2 f(x_2) + w_3 f(x_3)$$

Ponderile cuadraturii sunt:

$$w_1 = h \int_1^3 \frac{1}{2}(t-2)(t-3)\,\mathrm{d}t = \frac{h}{3}$$
$$w_2 = h \int_1^3 -(t-1)(t-3)\,\mathrm{d}t = \frac{4h}{3}$$
$$w_3 = h \int_1^3 \frac{1}{2}(t-1)(t-2)\,\mathrm{d}t = \frac{h}{3}$$

Astfel, obținem formula de cuadratură Simpson:

$$\begin{aligned} I_2(f) &= h\left[\frac{1}{3} f(a) + \frac{4}{3} f\left(\frac{a+b}{2}\right) + \frac{1}{3} f(b)\right] \\ &= \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] \end{aligned} \tag{11}$$

Estimarea erorii de cuadratură Simpson: Se poate arăta că dacă $f \in C^4[a,b]$, atunci $\exists\, \xi \in (a,b)$ a.i.

$$e_t(f) = I(f) - I_2(f) = -\frac{f^{(4)}(\xi)}{90} h^5 = O(h^5) \tag{12}$$

VII.2.3. Formula de cuadratură a dreptunghiului

Considerăm cazul cuadraturii Newton-Cotes deschisă ($n = 0$). Nodurile cuadraturii sunt:

$$x_0 := a, \quad x_1 = \frac{a+b}{2} = a + h, \quad x_2 := b = a + 2h, \quad h = \frac{b-a}{2}$$

Formula de cuadratură este:

$$I_0(f) = w_1 f(x_1) = w_1 f\left(\frac{a+b}{2}\right) \tag{13}$$

Ponderea de cuadratură w_1 este:

$$w_1 = \int_a^b L_{0,1}(x)dx = b - a = 2h$$

Obs.: Convenție: Funcția de bază pentru $n = 0$ o vom considera $L_{0,1}(x) = 1$.
Obținem astfel formula de cuadratură a dreptunghiului

$$I_0(f) = 2\,h\,f\left(\frac{a+b}{2}\right) = (b-a)\,f\left(\frac{a+b}{2}\right) \tag{14}$$

Dacă $f \in C^2[a,b]$, atunci $\exists\, \xi \in (a,b)$ a.i. $e_t(f) = \dfrac{f''(\xi)}{3} h^3$

VII.3. Formule de cuadratură sumate

VII.3.1. Formula de cuadratură sumată a dreptunghiului.
Fie partiție/diviziune echidistantă $a = x_1 < x_2 < \ldots < x_{2m} < x_{2m+1} = b$, $m \geq 1$, a intervalului $[a, b]$:

$$[a,b] = \bigcup_{k=1}^{m} \left[x_{2k-1}, x_{2k+1}\right]; \quad x_k = a+(k-1)h, \quad k = \overline{1, 2m+1}; \quad h = \frac{b-a}{2m}$$

Are loc identitatea:

$$I(f) = \int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{m} \int_{x_{2k-1}}^{x_{2k+1}} f(x)\,\mathrm{d}x \tag{15}$$

Aplicăm formula de cuadratură a dreptunghiului pe fiecare subinterval $\left[x_{2k-1}, x_{2k+1}\right] \subset [a, b]$, $k = \overline{1, m}$

$$I_{0,m} = \sum_{k=1}^{m} I_0^k(f) = \sum_{k=1}^{m} f(x_{2k})(x_{2k+1} - x_{2k-2}) = 2h \sum_{k=1}^{m} f(x_{2k}) \tag{16}$$

unde $I_0^k(f)$ reprezintă formula de cuadratură a dreptunghiului scrisă pe intervalul $[x_{2k-1}, x_{2k+1}]$.
Obs.: Adunând erorile formulei de cuadratură de la fiecare subinterval , eroarea formulei de cuadratură sumată își micșorează ordinul cu o unitate. Fie $e_k = O(h^3)$, eroarea formulei de cuadratură a dreptunghiului pe subintervalul $[x_{2k-1}, x_{2k+1}]$, atunci eroarea formulei de cuadatură sumată a a dreptunghiului este:

$$e = \sum_{k=1}^{m} e_k = O(h^3) \sum_{k=1}^{m} 1 = m \cdot O(h^3) = \frac{b-a}{2h} \cdot O(h^3) = O(h^2) \tag{17}$$

VII.3.2. Formula de cuadratură sumată a trapezului.
Fie diviziunea echidistantă $a = x_1 < x_2 < \ldots < x_m < x_{m+1} = b$, $m \geq 1$, a intervalului $[a, b]$:

$$[a,b] = \bigcup_{k=1}^{m} \left[x_k, x_{k+1}\right]; \quad x_k = a + (k-1)h, \quad k = \overline{1, m+1}; \quad h = \frac{b-a}{m}$$

Are loc identitatea:

$$I(f) = \int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{m} \int_{x_k}^{x_{k+1}} f(x)\,\mathrm{d}x \tag{18}$$

În fiecare subinterval $\left[x_k, x_{k+1}\right] \subset [a, b]$, $k = \overline{1, m+1}$, considerăm nodurile de interpolare x_k și x_{k+1}.
Aplicăm formula de cuadratură a trapezului pe fiecare subinterval $\left[x_k, x_{k+1}\right] \subset [a, b]$, $k = \overline{1, m+1}$

$$\begin{aligned} I_{1,m}(f) &= \sum_{k=1}^{m} I_1^k(f) = \sum_{k=1}^{m} \frac{f(x_k) + f(x_{k+1})}{2} \cdot (x_{k+1} - x_k) \\ &= \frac{h}{2} \sum_{k=1}^{m} (f(x_k) + f(x_{k+1})) = \frac{h}{2}(f(x_1) + f(x_2) \\ &+ f(x_2) + f(x_3) + \ldots + f(x_m) + f(x_{m+1})) \\ &= \frac{h}{2}(f(x_1) + 2\sum_{k=2}^{m} f(x_k) + f(x_{m+1})) \end{aligned} \tag{19}$$

unde $I_1^k(f)$ reprezintă formula de cuadratură a trapezului pe intervalul $[x_k, x_{k+1}]$.
Eroarea formulei de cuadratură sumată a trapezului este de ordinul $O(h^2)$.

VII.3.3. Formula de cuadratură sumată Simpson.

Fie diviziune echidistantă $a = x_1 < x_2 < \ldots < x_{2m} < x_{2m+1} = b$, $m \geq 1$, a intervalului $[a, b]$:

$$[a,b] = \bigcup_{k=1}^{m} \left[x_{2k-1}, x_{2k+1}\right]; \quad x_k = a + (k-1)h, \quad k = \overline{1, 2m+1}$$
$$h := \frac{b-a}{2m}$$

Are loc identitatea:

$$I(f) = \int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{m} \int_{x_{2k-1}}^{x_{2k+1}} f(x)\,\mathrm{d}x \tag{20}$$

În fiecare subinterval $\left[x_{2k-1}, x_{2k+1}\right] \subset [a, b]$, $k = \overline{1, m}$, considerăm nodurile de interpolare x_{2k-1}, x_{2k} și x_{2k+1}

Aplicăm formula de cuadratură Simpson pe fiecare subinterval $\left[x_{2k-1}, x_{2k+1}\right] \subset [a, b]$, $k = \overline{1, m}$:

$$\begin{aligned} I_{2,m}(f) &= \sum_{k=1}^{m} I_2^k(f) = \sum_{k=1}^{m} \left(\frac{1}{3} f(x_{2k-1}) + \frac{4}{3} f(x_{2k}) + \frac{1}{3} f(x_{2k+1})\right) \times \\ &\quad \times \frac{x_{2k+1} - x_{2k-1}}{2} \\ &= \frac{h}{3}\left(f(x_1) + 4\sum_{k=1}^{m} f(x_{2k}) + 2\sum_{k=1}^{m-1} f(x_{2k+1}) + f(x_{2m+1})\right) \end{aligned} \tag{21}$$

unde $I_2^k(f)$ reprezintă formula de cuadratură Simpson pe intervalul $[x_{2k-1}, x_{2k+1}]$.

Eroarea formulei de cuadratură sumată Simpson este de ordinul $O(h^4)$.

CURS #12

CONȚINUTUL CURSULUI #12:

VI. Derivarea numerică.

VI.1. Diferențe finite progresive, regresive și centrale pentru $f'(x)$.

Fie $f \in C^2([a,b])$. Din Teorema lui Taylor rezultă, pentru $h > 0$:

$$f(x+h) = f(x) + f'(x)h + f''(\xi)\frac{h^2}{2}, \ \xi \in (x, x+h) \Rightarrow$$

$$f'(x) = \frac{f(x+h) - f(x)}{h} - f''(\xi)\frac{h}{2} = \frac{f(x+h) - f(x)}{h} + O(h) \Rightarrow$$

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} \tag{1}$$

Relația (1) se numește **formula de apoximare prin diferențe finite progresive** pentru $f'(x)$.
Are loc estimarea erorii de trunchiere:

$$|e_t| = \left| f'(x) - \frac{f(x+h) - f(x)}{h} \right| = \frac{h}{2}|f''(\xi)| \leq \frac{h}{2}M = O(h)$$

unde $M = \max_{t \in [x, x+h]} |f''(t)|$. Precizăm că orice funție continuă pe un interval închis este și mărginită.

Din Teorema lui Taylor rezultă, pentru $h > 0$:

$$f(x-h) = f(x) - f'(x)h + f''(\xi)\frac{h^2}{2}, \xi \in (x-h, x) \Rightarrow$$

$$f'(x) = \frac{f(x) - f(x-h)}{h} + f''(\xi)\frac{h}{2} = \frac{f(x) - f(x-h)}{h} + O(h)$$

Obținem astfel **formula de aproximare prin diferențe finite regresive** pentru $f'(x)$:

$$f'(x) \approx \frac{f(x) - f(x-h)}{h} \tag{2}$$

cu eroarea de trunchiere, e_t:

$$|e_t| = \left| f'(x) - \frac{f(x) - f(x-h)}{h} \right| = \frac{h}{2}|f''(\xi)| \leq \frac{h}{2}M = O(h)$$

unde $M = \max_{t \in [x-h, x]} |f''(t)|$.

Fie $f \in C^3[a,b]$. Din Teorema lui Taylor rezultă, pentru $h > 0$:

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2} + f^{(3)}(\xi_1)\frac{h^3}{6},$$
$$\xi_1 \in (x, x+h)$$
$$f(x-h) = f(x) - f'(x)h + f''(x)\frac{h^2}{2} - f^{(3)}(\xi_2)\frac{h^3}{6},$$
$$\xi_2 \in (x-h, x)$$

Scăzând a doua relație din prima și rearanjând termenii, obținem:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \left[f^{(3)}(\xi_1) + f^{(3)}(\xi_2)\right]\frac{h^2}{12},$$
$$\xi_1 \in (x, x+h), \quad \xi_2 \in (x-h, x)$$

Obținem astfel **formula de aproximare prin diferențe finite centrale** pentru $f'(x)$:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

cu erorea de trunchiere, e_t:

$$\begin{aligned} |e_t| &= \left| f'(x) - \frac{f(x+h) - f(x-h)}{2h} \right| = \frac{h^2}{12} |f^{(3)}(\xi_1) + f^{(3)}(\xi_2)| \\ &\leq \frac{h^2}{12} \left(|f^{(3)}(\xi_1)| + |f^{(3)}(\xi_2)| \right) \leq \frac{h^2}{12} M = O(h^2) \end{aligned}$$

unde $M = \max_{t \in [x, x+h]} |f'''(t)| + \max_{t \in [x-h, x]} |f'''(t)|$.

VI.2. Diferențe finite centrale pentru $f''(x)$.

Fie $f \in C^4[a, b]$. Din Teorema lui Taylor rezultă, pentru $h > 0$:

$$\begin{aligned} f(x+h) &= f(x) + f'(x)h + f''(x)\frac{h^2}{2} + f^{(3)}(x)\frac{h^3}{6} + f^{(4)}(\xi_1)\frac{h^4}{24}, \\ &\quad \xi_1 \in (x, x+h) \\ f(x-h) &= f(x) - f'(x)h + f''(x)\frac{h^2}{2} - f^{(3)}(x)\frac{h^3}{6} + f^{(4)}(\xi_2)\frac{h^4}{24}, \\ &\quad \xi_2 \in (x-h, x) \end{aligned}$$

Adunând relațiile de mai sus și rearanjând termenii, obținem:

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} - \left[f^{(4)}(\xi_1) + f^{(4)}(\xi_2) \right] \frac{h^2}{24},$$
$$\xi_1 \in (x, x+h), \quad \xi_2 \in (x-h, x)$$

Formula de aproximare prin diferențe finite centrale pentru $f''(x)$ este:

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

cu **eroarea de trunchiere**, e_t:

$$|e_t| = \left| f''(x) - \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \right| \leq \frac{h^2}{24} M = O(h^2)$$

unde $M = \max_{t \in [x, x+h]} |f^4(t)| + \max_{t \in [x-h, x]} |f^4(t)|$.

VI.3. Metoda de extrapolare Richardson.

Dacă avem dată o formulă de aproximare a derivatei $f'(x)$ de forma $f'(x) = \phi_1(x, h) + O(h)$, atunci în baza funcției ϕ_1 se poate construi recurent un șir de funcții $(\phi_n)_{n>1}$ care aproximează derivata $f'(x)$ cu ordinul de aproximare $O(h^n)$.

Pentru simplificare vom evita scrierea variabilei x ca argument al funcției ϕ_n. Avem astfel:

$$\begin{aligned} f'(x) &= \phi_1(h) + a_1 h + a_2 h^2 + a_3 h^3 + \ldots \\ &= \phi_1(h) + O(h) \end{aligned} \tag{3}$$

Cum (3) are loc pentru orice valoare $h > 0$, scriem formula de aproximare (3) pentru $h/2$:

$$f'(x) = \phi_1\left(\frac{h}{2}\right) + a_1\left(\frac{h}{2}\right) + a_2\left(\frac{h}{2}\right)^2 + a_3\left(\frac{h}{2}\right)^3 + \ldots \tag{4}$$

Efectuăm următoarea combinație: $2^1 \cdot$ (4) $-1 \cdot$ (3). Rezultă:

$$(2^1-1)f'(x) = \left[2^1\phi_1\left(\frac{h}{2}\right) - \phi_1(h)\right] + a_2\left(\frac{1}{2}-1\right)h^2 + a_3\left(\frac{1}{2^2}-1\right)h^3 + \dots$$

$$f'(x) = \phi_2(h) + b_2h^2 + b_3h^3 + b_4h^4 + \dots = \phi_2(h) + O(h^2) \tag{5}$$

unde

$$\begin{aligned}\phi_2(h) &:= \frac{1}{2^1-1}\left[2^1\phi_1\left(\frac{h}{2}\right) - \phi_1(h)\right] \\ &= \phi_1\left(\frac{h}{2}\right) + \frac{1}{2^1-1}\left[\phi_1\left(\frac{h}{2}\right) - \phi_1(h)\right]\end{aligned} \tag{6}$$

Cum relația (5) are loc pentru orice $h > 0$, scriem formula de aproximare (5) pentru $h/2$:

$$f'(x) = \phi_2\left(\frac{h}{2}\right) + b_2\left(\frac{h}{2}\right)^2 + b_3\left(\frac{h}{2}\right)^3 + b_4\left(\frac{h}{2}\right)^4 + \dots \tag{7}$$

Efectuăm următoarea combinație: $2^2\cdot$ (7) $-1\cdot$ (5). Rezultă:

$$f'(x) = \phi_3(h) + c_3h^3 + c_4h^4 + c_5h^5 + \dots = \phi_3(h) + O(h^3) \tag{8}$$

unde

$$\begin{aligned}\phi_3(h) &:= \frac{1}{2^2-1}\left[2^2\phi_2\left(\frac{h}{2}\right) - \phi_2(h)\right] \\ &= \phi_2\left(\frac{h}{2}\right) + \frac{1}{2^2-1}\left[\phi_2\left(\frac{h}{2}\right) - \phi_2(h)\right]\end{aligned} \tag{9}$$

Prin inducție după $n \geq 2$ se poate demonstra formula de aproximare pentru $f'(x)$:

$$f'(x) = \phi_n(h) + d_nh^n + d_{n+1}h^{n+1} + d_{n+2}h^{n+2} + \dots = \phi_n(h) + O(h^n) \tag{10}$$

unde

$$\begin{aligned}\phi_n(h) &:= \frac{1}{2^{n-1}-1}\left[2^{n-1}\phi_{n-1}\left(\frac{h}{2}\right) - \phi_{n-1}(h)\right] \\ &= \phi_{n-1}\left(\frac{h}{2}\right) + \frac{1}{2^{n-1}-1}\left[\phi_{n-1}\left(\frac{h}{2}\right) - \phi_{n-1}(h)\right]\end{aligned} \tag{11}$$

Vom adopta următoarea notație

$$Q_{ij} = \phi_j\left(\frac{h}{2^{i-j}}\right) \tag{12}$$

Cu această convenție, conform metodei inductive

$$\begin{aligned}Q_{ij} &= \phi_j\left(\frac{h}{2^{i-j}}\right) = \phi_{j-1}\left(\frac{h}{2^{i-j+1}}\right) \\ &+ \frac{1}{2^{j-1}-1}\left(\phi_{j-1}\left(\frac{h}{2^{i-j+1}}\right) - \phi_{j-1}\left(\frac{h}{2^{i-j}}\right)\right) \\ &= \phi_{j-1}\left(\frac{h}{2^{i-(j-1)}}\right) \\ &+ \frac{1}{2^{j-1}-1}\left(\phi_{j-1}\left(\frac{h}{2^{i-(j-1)}}\right) - \phi_{j-1}\left(\frac{h}{2^{i-1-(j-1)}}\right)\right)\end{aligned}$$

$$Q_{ij} = Q_{i,j-1} + \frac{1}{2^{j-1}-1}(Q_{i,j-1} - Q_{i-1,j-1}) \tag{13}$$

Vom da în continuare următorul tabel:

$h/2^n$	$O(h)$	$O(h^2)$	$O(h^3)$	$O(h^4)$	$O(h^5)$
h	$\boxed{\phi_1(h)}$ $\searrow$				
$h/2$	$\phi_1(h/2)$ $\longrightarrow$ $\searrow$	$\boxed{\phi_2(h)}$ $\searrow$			
$h/2^2$	$\phi_1(h/2^2)$ $\longrightarrow$ $\searrow$	$\phi_2(h/2)$ $\longrightarrow$ $\searrow$	$\boxed{\phi_3(h)}$ $\searrow$		
$h/2^3$	$\phi_1(h/2^3)$ $\longrightarrow$ $\searrow$	$\phi_2(h/2^2)$ $\longrightarrow$ $\searrow$	$\phi_3(h/2)$ $\longrightarrow$ $\searrow$	$\boxed{\phi_4(h)}$ $\searrow$	
$h/2^4$	$\phi_1(h/2^4)$ $\longrightarrow$ $\searrow$	$\phi_2(h/2^3)$ $\longrightarrow$ $\searrow$	$\phi_3(h/2^2)$ $\longrightarrow$ $\searrow$	$\phi_4(h)$ $\longrightarrow$ $\searrow$	$\boxed{\phi_5(h)}$ $\searrow$
...	...	...	...	...	...

Conform acestui tabel elementele de pe diagonala principală aproximează derivata $f'(x)$ cu ordinul de aproximare egal cu numărul coloanei.

$h/2^n$	$O(h)$	$O(h^2)$	$O(h^3)$	$O(h^4)$	$O(h^5)$
h	$\boxed{\phi_1(h) = Q_{11}}$				
$h/2$	$\phi_1(h/2) = Q_{21}$	$\boxed{\phi_2(h) = Q_{22}}$			
$h/2^2$	$\phi_1(h/2^2) = Q_{31}$	$\phi_2(h/2) = Q_{32}$	$\boxed{\phi_3(h) = Q_{33}}$		
$h/2^3$	$\phi_1(h/2^3) = Q_{41}$	$\phi_2(h/2^2) = Q_{42}$	$\phi_3(h/2) = Q_{43}$	$\boxed{\phi_4(h) = Q_{44}}$	
$h/2^4$	$\phi_1(h/2^4) = Q_{51}$	$\phi_2(h/2^3) = Q_{52}$	$\phi_3(h/2^2) = Q_{53}$	$\phi_4(h/2) = Q_{54}$	$\boxed{\phi_5(h) = Q_{55}}$
...	...	...	...	...	...

ALGORITM (Formula de extrapolare Richardson)

Date de intrare: f; x; h; n.

Date de ieșire: df.

STEP 1: Se defineşte funcţia $\phi = \phi(x, h)$;

for $i = 1 : n$ do

$Q_{i1} = \phi(x, h/2^{i-1})$;

endfor

STEP 2: for $i = 2 : n$ do

for $j = 2 : i$ do

Determină Q_{ij} conform (13);

endfor

endfor

STEP 3: $df = Q_{nn}$

Observaţie: Algoritmul Richardson poate fi aplicat şi pentru aproximarea derivatei de ordinul doi. Fiind dată o formulă de aproximare de ordinul doi pentru $f''(x)$ cu ordinul de aproximare $O(h^2)$, în calculul matricei Q se va suprima ultima coloană. Astfel, pentru evaluarea derivatei de ordinul 2 cu ordinul de aproximare $O(h^n)$ se va returna valoarea componentei $Q_{n-1,n-1}$.

CONȚINUTUL CURSULUI #11:

V. Interpolarea cu funcții spline.

V.3. Interpolare cu funcții spline cubice.

Definiția (V.3.)

Funcția $S : [a,b] \longrightarrow \mathbb{R}$ s.n. funcție spline cubică pentru funcția $f : [a,b] \longrightarrow \mathbb{R}$ dacă:

(a) *S este cubică pe porțiuni:*

$$S(x) = S_j(x), \quad \forall\, x \in I_j, \quad j = \overline{1,n} \tag{1}$$

unde

$$S_j : \bar{I}_j \longrightarrow \mathbb{R},$$

$$S_j(x) = a_j + b_j\,(x - x_j) + c_j\,(x - x_j)^2 + d_j\,(x - x_j)^3, j = \overline{1,n} \tag{2}$$

cu $a_j, b_j, c_j, d_j \in \mathbb{R}$, $j = \overline{1,n}$, ce trebuie determinate.

(b) *S interpolează f în x_j, $j = \overline{1,n+1}$:*

$$S(x_j) = f(x_j), \quad j = \overline{1,n+1} \tag{3}$$

Definiția (V.3. (continuare))

(c) *S este continuă în x_{j+1}, $j = \overline{1,n-1}$:*

$$S_j(x_{j+1}) = S_{j+1}(x_{j+1}), \quad j = \overline{1,n-1} \tag{4}$$

(d) *S' este continuă în x_{j+1}, $j = \overline{1,n-1}$:*

$$S_j'(x_{j+1}) = S_{j+1}'(x_{j+1}), \quad j = \overline{1,n-1} \tag{5}$$

(e) *S'' este continuă în x_{j+1}, $j = \overline{1,n-1}$:*

$$S_j''(x_{j+1}) = S_{j+1}''(x_{j+1}), \quad j = \overline{1,n-1} \tag{6}$$

(f) *Unul dintre următoarele seturi de condiții este îndeplinit*

$$(f)_1 : \quad S'(x_1) = f'(x_1), \quad S'(x_{n+1}) = f'(x_{n+1}) \tag{7}$$

$$(f)_2 : \quad S''(x_1) = f''(x_1), \quad S''(x_{n+1}) = f''(x_{n+1}) \tag{8}$$

Vom trata doar cazul $(f)_1$, cazul $(f)_2$ se abordează după același raționament. Conform condiției (b) rezultă

$$\begin{cases} a_j = f(x_j), \quad j = \overline{1,n} \\ a_n + b_n(x_{n+1} - x_n) + c_n(x_{n+1} - x_n)^2 + d_n(x_{n+1} - x_n)^3 = f(x_{n+1}) \end{cases} \tag{9}$$

Din (c) rezultă

$$a_j + b_j(x_{j+1} - x_j) + c_j(x_{j+1} - x_j)^2 + d_j(x_{j+1} - x_j)^3 = a_{j+1}, \quad j = \overline{1,n-1} \tag{10}$$

Relațiile (9) și (10) se rescriu

$$\begin{cases} a_j = f(x_j), \quad j = \overline{1,n} \\ b_j h_j + c_j h_j^2 + d_j h_j^3 = f(x_{j+1}) - f(x_j), \quad j = \overline{1,n} \end{cases} \tag{11}$$

Deoarece $S_j'(x) = b_j + 2c_j(x - x_j) + 3d_j(x - x_j)^2$ din (d) rezultă

$$b_j + 2c_j(x_{j+1} - x_j) + 3d_j(x_{j+1} - x_j)^2 = b_{j+1}, \quad j = \overline{1,n-1} \tag{12}$$

sau

$$b_j + 2c_j h_j + 3d_j h_j^2 = b_{j+1}, \quad j = \overline{1,n-1} \tag{13}$$

Deoarece $S_j''(x) = 2c_j + 6d_j(x - x_j)$, atunci conform (e) rezultă
$$c_j + 3d_jh_j = c_{j+1}, \quad j = \overline{1, n-1} \tag{14}$$
Din condițiile (f)$_1$ și ținând cont că $S'(x_1) = S_1'(x_1)$, $S'(x_{n+1}) = S_n(x_{n+1})$ rezultă
$$b_1 = f'(x_1) \tag{15}$$
$$b_n + 2c_n + 3d_nh_n^2 = f'(x_{n+1}) \tag{16}$$
Relația (16) poate fi înglobată în relațiile (13) dacă adoptăm notația $b_{n+1} = f'(x_{n+1})$. Se obține sistemul complet de determinare a coeficienților funcției spline cubice S
$$\begin{cases} a_j = f(x_j), \quad j = \overline{1, n} \\ b_jh_j + c_jh_j^2 + d_jh_j^3 = f(x_{j+1}) - f(x_j), \quad j = \overline{1, n} \\ b_j + 2c_jh_j + 3d_jh_j^2 = b_{j+1}, \quad j = \overline{1, n} \\ b_1 = f'(x_1), \quad b_{n+1} = f'(x_{n+1}) \\ c_{j-1} + 3d_{j-1}h_{j-1} = c_j, \quad j = \overline{2, n} \end{cases} \tag{17}$$
Obs.: Relațiile (17) formează un sistem de $4n+1$ ecuații și $4n+1$ necunoscute $a_j, c_j, d_j, j = \overline{1, n}$; $b_j, j = \overline{1, n+1}$.

Dacă cuplăm relațiile $(17)_2$ și $(17)_3$ obținem sistemul
$$\begin{cases} c_jh_j + d_jh_j^2 = \dfrac{f(x_{j+1}) - f(x_j)}{h_j} - b_j, \quad j = \overline{1, n} \\ 2c_jh_j + 3d_jh_j^2 = b_{j+1} - b_j, \quad j = \overline{1, n} \end{cases} \tag{18}$$
Combinațiile de forma: $(18)_1 \times 2 - (18)_2$ și $(18)_1 \times 3 - (18)_2$ furnizează expesii pentru coeficienții $d_j, c_j, j = \overline{1, n}$ exprimați în raport cu coeficienții $b_j, j = \overline{1, n+1}$. Astfel
$$\begin{cases} d_j = -\dfrac{2}{h_j^3}(f(x_{j+1}) - f(x_j)) + \dfrac{1}{h_j^2}(b_{j+1} + b_j), \quad j = \overline{1, n} \\ c_j = \dfrac{3}{h_j^2}(f(x_{j+1}) - f(x_j)) - \dfrac{b_{j+1} + 2b_j}{h_j}, j = \overline{1, n} \end{cases} \tag{19}$$
Introducând coeficienții c_j, d_j în relația $(17)_5$ se obține un sistem de $n+1$ ecuații, având drept necunoscute coeficienții, b_j, $j = \overline{1, n+1}$

$$\begin{cases} b_1 = f'(x_1) \\ \dfrac{1}{h_{j-1}}b_{j-1} + \left(\dfrac{2}{h_j} + \dfrac{2}{h_{j-1}}\right)b_j + \dfrac{1}{h_j}b_{j+1} \\ = -\dfrac{3}{h_{j-1}^2}f(x_{j-1}) + \left(\dfrac{3}{h_{j-1}^2} - \dfrac{3}{h_j^2}\right)f(x_j) + \dfrac{3}{h_j^2}f(x_{j+1}), \quad j = \overline{2, n} \\ b_{n+1} = f'(x_{n+1}) \end{cases} \tag{20}$$
În cazul unei diviziuni $(x_j)_{j=\overline{1,n+1}}$ echidistante cu pasul h sistemul de mai sus se rescrie sub forma
$$\begin{cases} b_1 = f'(x_1) \\ b_{j-1} + 4b_j + b_{j+1} = \dfrac{3}{h}(f(x_{j+1}) - f(x_{j-1})), \quad j = \overline{2, n} \\ b_{n+1} = f'(x_{n+1}) \end{cases} \tag{21}$$

cu matricea asociată, B, diagonal dominantă
$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 4 & 1 & 0 & \dots & 0 \\ 0 & 1 & 4 & 1 & \dots & 0 \\ 0 & 0 & 1 & 4 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \tag{22}$$

CURS #10

CONȚINUTUL CURSULUI #10:

V. Interpolarea cu funcții spline.
V.1. Interpolare cu funcții spline liniare.

Fie $f : [a, b] \to \mathbb{R}$ și $(x_i)_{i=\overline{1,n+1}}$ o diviziune a intervalului $[a, b]$, i.e. $a = x_1 < \ldots < x_{n+1} = b$. Fie $I_j = [x_j, x_{j+1})$ cu $\bar{I}_j = [x_j, x_{j+1}]$, $j = \overline{1, n-1}$, $I_n = \bar{I}_n = [x_n, x_{n+1}]$.

Definiția (V.1.)

Funcția $S : [a, b] \longrightarrow \mathbb{R}$ s.n. funcție spline liniară pentru funcția $f : [a, b] \longrightarrow \mathbb{R}$ dacă:

(a) *S este liniară pe porțiuni:*

$$S(x) = S_j(x), \quad \forall\, x \in I_j, \quad j = \overline{1, n} \tag{1}$$

unde

$$S_j : \bar{I}_j \longrightarrow \mathbb{R}, \quad S_j(x) = a_j + b_j\,(x - x_j), \quad j = \overline{1, n} \tag{2}$$

cu $a_j, b_j \in \mathbb{R}$, $j = \overline{1, n}$, ce trebuie determinate.

Definiția (V.1. (continuare))

(b) *S interpolează f în x_j, $j = \overline{1, n+1}$:*

$$S(x_j) = f(x_j), \quad j = \overline{1, n+1} \tag{3}$$

(c) *S este continuă în nodurile interioare, i.e. x_{j+1}, $j = \overline{1, n-1}$:*

$$S_j(x_{j+1}) = S_{j+1}(x_{j+1}), \quad j = \overline{1, n-1} \tag{4}$$

Relațiile (3)–(4) ne furnizează sistemul de ecuații liniare, i.e. $2n$ ecuații liniare pentru necunoscutele $a_j, b_j \in \mathbb{R}$, $j = \overline{1, n}$.
Conform condiției (b) și ținând cont de faptul că $x_j \in I_j, j = \overline{1, n}$ rezultă

$$S(x_j) = S_j(x_j) = f(x_j), \quad \text{deci} \quad a_j = f(x_j), \quad j = \overline{1, n}$$

Nodul $x_{n+1} \in I_n$, deci

$$S(x_{n+1}) = S_n(x_{n+1}) \Rightarrow a_n + b_n(x_{n+1} - x_n) = f(x_{n+1}) \Rightarrow$$

$$b_n = \frac{f(x_{n+1}) - f(x_n)}{x_{n+1} - x_n} \tag{5}$$

Conform condiției (c) se obțin succesiv următoarele relații:

$$(a_j + b_j(x - x_j))\,|_{x=x_{j+1}} = (a_{j+1} + b_{j+1}(x - x_{j+1}))\,|_{x=x_{j+1}}$$

$$a_j + b_j(x_{j+1} - x_j) = a_{j+1}, \quad j = \overline{1, n-1}$$

$$b_j = \frac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j}, \quad j = \overline{1, n-1} \tag{6}$$

Rezultă următoarea schemă numerică de determinare a coeficienților $a_j, b_j, j = \overline{1, n}$:

$$\begin{cases} a_j = f(x_j), & j = \overline{1, n} \\ b_j = \dfrac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j}, & j = \overline{1, n} \end{cases} \tag{7}$$

ALGORITM (Interpolarea spline liniară)

Date de intrare: $X; Y; x$;
Date de ieșire: y;

STEP 1: `Determină` n;

STEP 2: `for` $j = 1 : n$ `do`

$$a_j = Y_j; \quad b_j = \frac{Y_{j+1} - Y_j}{X_{j+1} - X_j};$$

`endfor`

STEP 3: `for` $j = 1 : n$ `do`

`if` $x \in [X_j, X_{j+1}]$ `do`

$S = a_j + b_j(x - X_j)$;

`STOP`

`endif`

`endfor`

$y = S$;

Obs.: Vectorul X conține nodurile de interpolare $x_1, \ldots, x_{n+1}$, iar vectorul Y conține valorile funcției în nodurile de interpolare, $f(x_1), \ldots, f(x_{n+1})$.
Exemplul # 1: Să se afle funcția spline liniară pentru funcția $f(x) = e^{2x}$ relativ la diviziunea $(x_1, \ x_2, \ x_3) = (-1, \ 0, \ 1)$.
Rezolvare:

$$S(x) = \begin{cases} S_1(x), & x \in [x_1, x_2) \\ S_2(x), & x \in [x_2, x_3] \end{cases}$$

unde $S_1(x) = a_1 + b_1(x - x_1)$ și $S_2(x) = a_2 + b_2(x - x_2)$. Se obține astfel

$$S(x) = \begin{cases} a_1 + b_1(x+1), & x \in [-1, 0) \\ a_2 + b_2 x, & x \in [0, 1] \end{cases}$$

Deoarece S interpolează f în cele trei noduri rezultă

$$S(x_1) = f(x_1), S(x_2) = f(x_2), S(x_3) = f(x_3)$$

echivalent:

$$S_1(-1) = e^{-2}, \ S_2(0) = 1, \ S_2(1) = e^2$$

de unde $a_1 = e^{-2}, \ a_2 = 1, \ a_2 + b_2 = e^2$, deci $b_2 = e^2 - 1$.
Pe de altă parte, S este continuă în nodul $x_2 \in (-1, 1)$, i.e.
$S_1(x_2) = S_2(x_2)$ sau $S_1(0) = S_2(0)$, deci $a_1 + b_1 = a_2$, de unde rezultă $b_1 = 1 - e^{-2}$. Obținem astfel, următoarea reprezentare:

$$S(x) = \begin{cases} e^{-2} + (1 - e^{-2})(x+1), & x \in [-1, 0) \\ 1 + (e^2 - 1)x, & x \in [0, 1] \end{cases}$$
$$= \begin{cases} 1 + (1 - e^{-2})x, & x \in [-1, 0) \\ 1 + (e^2 - 1)x, & x \in [0, 1] \end{cases}$$

V.2. Interpolare cu funcții spline pătratice.

Definiția (V.2.)

Funcția $S : [a, b] \longrightarrow \mathbb{R}$ s.n. funcție spline pătratică pentru funcția $f : [a, b] \longrightarrow \mathbb{R}$ dacă:

(a) *S este pătratică pe porțiuni:*

$$S(x) = S_j(x), \quad \forall\, x \in I_j, \quad j = \overline{1, n} \tag{8}$$

unde

$$S_j : \bar{I}_j \longrightarrow \mathbb{R},$$
$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2, j = \overline{1, n} \tag{9}$$

cu $a_j, b_j, c_j \in \mathbb{R}$, $j = \overline{1, n}$, ce trebuie determinate.

(b) *S interpolează f în x_j, $j = \overline{1, n+1}$:*

$$S(x_j) = f(x_j), \quad j = \overline{1, n+1} \tag{10}$$

Definiția (V.2. (continuare))

(c) *S este continuă în nodurile interioare* x_{j+1}, $j = \overline{1, n-1}$:

$$S_j(x_{j+1}) = S_{j+1}(x_{j+1}), \quad j = \overline{1, n-1} \tag{11}$$

(d) *S' este continuă în nodurile interioare* x_{j+1}, $j = \overline{1, n-1}$:

$$S'_j(x_{j+1}) = S'_{j+1}(x_{j+1}), \quad j = \overline{1, n-1} \tag{12}$$

(e) *Una din următoarele condiții este satisfăcută*

$(e)_1$: $S'(x_1) = f'(x_1)$
$(e)_2$: $S'(x_{n+1}) = f'(x_{n+1})$

Conform condiției (b) rezultă

$$a_j = f(x_j), \quad j = \overline{1, n} \tag{13}$$

$$a_n + b_n(x_{n+1} - x_n) + c_n(x_{n+1} - x_n)^2 = f(x_{n+1}) \tag{14}$$

Conform condiției (c) rezultă

$$a_j + b_j(x_{j+1} - x_j) + c_j(x_{j+1} - x_j)^2 = a_{j+1}, \quad j = \overline{1, n-1} \tag{15}$$

sau

$$a_j + b_j(x_{j+1} - x_j) + c_j(x_{j+1} - x_j)^2 = f(x_{j+1}), \quad j = \overline{1, n-1} \tag{16}$$

Relațiile (14) și (16) pot fi cuplate și rescrise ca o singură relație pentru $j = \overline{1, n}$.
Cum $S'_j(x) = b_j + 2c_j(x - x_j)$, atunci conform condiției (d) rezultă

$$b_j + 2c_j(x_{j+1} - x_j) = b_{j+1}, \quad j = \overline{1, n-1} \tag{17}$$

Conform condiției (e) rezultă

$$S'_1(x_1) = f'(x_1) \Rightarrow b_1 = f'(x_1) \tag{18}$$

sau

$$S'_n(x_{n+1}) = f'(x_{n+1}) \Rightarrow b_n + 2c_n(x_{n+1} - x_n) = f'(x_{n+1}) \tag{19}$$

Dacă în (19) considerăm $b_{n+1} = f'(x_{n+1})$ atunci relațiile (19) și (17) pot fi cuplate și rescrise ca o singură relație pentru $j = \overline{1, n}$.

Fie $h_j = x_{j+1} - x_j, j = \overline{1, n}$ lungimea fiecărei subinterval $[x_j, x_{j+1}]$.
Obținem astfel, sistemele complete de ecuații necesare pentru determinarea coeficienților b_j, c_j:

$$\begin{cases} a_j + b_j h_j + c_j h_j^2 = f(x_{j+1}), & j = \overline{1, n} \\ b_1 = f'(x_1) \\ b_j + 2c_j h_j = b_{j+1}, & j = \overline{1, n-1} \end{cases} \tag{20}$$

sau

$$\begin{cases} a_j + b_j h_j + c_j h_j^2 = f(x_{j+1}), & j = \overline{1, n} \\ b_{n+1} = f'(x_{n+1}) \\ b_j + 2c_j h_j = b_{j+1}, & j = \overline{1, n} \end{cases} \tag{21}$$

Din $(20)_1$ rezultă

$$c_j = \frac{1}{h_j^2}\left(f(x_{j+1}) - f(x_j) - h_j b_j\right), \quad j = \overline{1, n} \tag{22}$$

Introducând (22) în $(20)_3$ obținem

$$b_{j+1} + b_j = \frac{2}{h_j}\left(f(x_{j+1}) - f(x_j)\right), \quad j = \overline{1, n-1} \tag{23}$$

Rezultă schemele numerice de calcul a coeficienților $b_j, c_j, j = \overline{1, n}$

$$\begin{cases} b_1 = f'(x_1) \\ b_{j+1} = \dfrac{2}{h_j}\left(f(x_{j+1}) - f(x_j)\right) - b_j, & j = \overline{1, n-1} \\ c_j = \dfrac{1}{h_j^2}\left(f(x_{j+1}) - f(x_j) - h_j b_j\right), & j = \overline{1, n} \end{cases} \tag{24}$$

sau

$$\begin{cases} b_{n+1} = f'(x_{n+1}) \\ b_j = \dfrac{2}{h_j}\left(f(x_{j+1}) - f(x_j)\right) - b_{j+1}, & j = \overline{n, 1} \\ c_j = \dfrac{1}{h_j^2}\left(f(x_{j+1}) - f(x_j) - h_j b_j\right), & j = \overline{1, n} \end{cases} \tag{25}$$

Exemplul #2: Să se afle funcția spline pătratică pentru funcția $f(x) = e^{2x}$ relativ la diviziunea $(x_1,\ x_2,\ x_3) = (-1,\ 0,\ 1)$.
Rezolvare:

$$S(x) = \begin{cases} S_1(x), & x \in [x_1, x_2) \\ S_2(x), & x \in [x_2, x_3] \end{cases}$$

unde

$$S_1(x) = a_1 + b_1(x - x_1) + c_1(x - x_1)^2$$

și

$$S_2(x) = a_2 + b_2(x - x_2) + c_2(x - x_2)^2$$

Se obține astfel

$$S(x) = \begin{cases} a_1 + b_1(x+1) + c_1(x+1)^2, & x \in [-1, 0) \\ a_2 + b_2 x + c_2 x^2, & x \in [0, 1] \end{cases}$$

Deoarece S interpolează f în cele trei noduri rezultă

$$S(x_1) = S_1(x_1) = f(x_1), S(x_2) = S_2(x_2) = f(x_2), S(x_2) = S_2(x_3) = f(x_3)$$

echivalent

$$S_1(-1) = e^{-2},\ S_2(0) = 1,\ S_2(1) = e^2$$

de unde $a_1 = e^{-2},\ a_2 = 1,\ a_2 + b_2 + c_2 = e^2$, deci

$$b_2 + c_2 = e^2 - 1. \tag{26}$$

Pe de altă parte, S este continuă în nodul $x_2 \in (-1, 1)$, i.e. $S_1(x_2) = S_2(x_2)$ sau $S_1(0) = S_2(0)$, deci $a_1 + b_1 + c_1 = a_2$, de unde rezultă

$$b_1 + c_1 = 1 - e^{-2}. \tag{27}$$

Derivatele funcțiilor S_1 și S_2 sunt:
$S_1'(x) = b_1 + 2c_1(x - x_1), S_2'(x) = b_2 + 2c_2(x - x_2)$. Funcția S' se exprimă prin formula

$$S'(x) = \begin{cases} b_1 + 2c_1(x+1), & x \in [-1, 0) \\ b_2 + 2c_2 x, & x \in [0, 1] \end{cases}$$

Derivata S' a funcției spline pătratice este continuă în nodul interior x_2, i.e. $S_1'(x_2) = S_2'(x_2)$ sau $S_1'(0) = S_2'(0)$ de unde rezultă

$$b_1 + 2c_1 = b_2 \tag{28}$$

Considerăm în plus satisfăcută condiția $S'(x_1) = f'(x_1)$ sau $S_1'(-1) = f'(-1)$, de unde $b_1 = 2e^{-2}$. Din relația (27) rezultă $c_1 = 1 - 3e^{-2}$, iar din (28) rezultă $b_2 = 2 - 4e^{-2}$. În final, din relația (26) rezultă $c_2 = e^2 + 4e^{-2} - 3$. Obținem astfel, următoarea reprezentare a funcției spline pătratice S :

$$S(x) = \begin{cases} e^{-2} + 2e^{-2}(x+1) + (1 - 3e^{-2})(x+1)^2, & x \in [-1, 0) \\ 1 + (2 - 4e^{-2})x + (e^2 + 4e^{-2} - 3)x^2, & x \in [0, 1] \end{cases}$$

CONȚINUTUL CURSULUI #8:

IV. Interpolarea Lagrange

Fie $f : [a, b] \to \mathbb{R}$ o funcție continuă și $(x_i)_{i=\overline{1,n+1}}$ o diviziune a intervalului $[a, b]$, i.e. $a = x_1 < x_2 < ... < x_{n+1} = b$.
Fie $\mathcal{P}_n$ mulțimea polinoamelor cel mult de grad $n \geq 0$:

$$\mathcal{P}_n = \left\{ P_n(x) = a_1 + a_2x + \ldots + a_nx^{n-1} + a_{n+1}x^n \;\middle|\; a_j \in \mathbb{R}, j = \overline{1, n+1} \right\}$$

Interpolarea Lagrange a funcției f relativ la diviziunea $(x_i)_{i=\overline{1,n+1}}$ constă în determinarea unui polinom $P_n \in \mathcal{P}_n$, numit polinom de interpolare Lagrange, care satisface relațiile:

$$P_n(x_i) = f(x_i), i = \overline{1, n+1} \tag{1}$$

Valorile $x_i, i = \overline{1, n+1}$ se numesc puncte sau noduri de interpolare.

IV.1. Metoda directă de determinare a polinomului Lagrange P_n.

Fie $P_n(x) = a_1 + a_2x + \ldots + a_nx^{n-1} + a_{n+1}x^n$ un polinom de interpolare al funcției f relativ la diviziunea $(x_i)_{i=\overline{1,N+1}}$. Din condițiile $P_n(x_i) = f(x_i)$, $y_i = f(x_i), i = \overline{1, n+1}$ rezultă următorul sistem de ecuații liniare

$$\begin{cases} a_1 + a_2x_1 + a_3x_1^2 + \ldots + a_{n+1}x_1^n = y_1 \\ a_1 + a_2x_2 + a_3x_2^2 + \ldots + a_{n+1}x_2^n = y_2 \\ a_1 + a_2x_3 + a_3x_3^2 + \ldots + a_{n+1}x_3^n = y_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_1 + a_2x_{n+1} + a_3x_{n+1}^2 + \ldots + a_{n+1}x_{n+1}^n = y_{n+1} \end{cases} \tag{2}$$

sau scris la formă matriceală

$$\underbrace{\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n+1} & x_{n+1}^2 & \ldots & x_{n+1}^n \end{pmatrix}}_{=:A} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n+1} \end{pmatrix} \tag{3}$$

Cum $x_i \neq x_j$, $1 \leq i < j \leq n+1$, rezultă

$$\det(A) = \prod_{1 \leq i < j \leq n+1} (x_i - x_j) \neq 0,$$

deci sistemul de ecuații liniare (3) este un sistem compatibil determinat cu soluția

$$\mathbf{a} = \left[a_1, a_2, \ldots, a_{n+1}\right]^T \in \mathbb{R}^{n+1}$$

Din unicitatea soluției rezultă că polinomul Lagrange se determină în mod unic.
Soluția sistemului de ecuații liniare (3) se poate obține, de exemplu, aplicând metoda Gauss cu pivotare totală.
Exemplu 1: Să se afle, prin metoda directă, polinomul de interpolare Lagrange $P_2(x)$ al funcției $f(x) = e^{2x}$ relativ la diviziunea $(-1; 0; 1)$.
Rezolvare: Fie $P_2(x) = a_1 + a_2x + a_3x^2$. Din condițiile $P_2(-1) = e^{-2}, P_2(0) = e^0, P_2(1) = e^2$ rezultă sistemul de ecuații liniare:

$$\begin{cases} a_1 + a_2 \cdot (-1) + a_3(-1)^2 = e^{-2} \\ a_1 + a_2 \cdot 0 + a_3 \cdot 0 = 1 \\ a_1 + a_2 \cdot 1 + a_3 \cdot 1 = e^2 \end{cases} \Rightarrow \begin{cases} a_1 = 1 \\ a_2 = \dfrac{e^2 - e^{-2}}{2} \\ a_3 = \dfrac{e^2 + e^{-2} - 2}{2} \end{cases} \quad (4)$$

Astfel, $P_2(x) = 1 + \dfrac{e^2 - e^{-2}}{2}x + \dfrac{e^2 + e^{-2} - 2}{2}x^2$.

IV.2. Metoda Lagrange de determinare a polinomului Lagrange P_n.

Se consideră următoarea reprezentare a polinomului Lagrange:

$$P_n(x) = \sum_{k=1}^{n+1} L_{n,k}(x)\, y_k, \quad x \in \mathbb{R} \quad (5)$$

unde $L_{n,k}$ sunt polinoame de gradul n ce urmează să fie determinate. Deoarece P_n interpolează funcția f în nodurile $\{x_i\}_{i=\overline{1,n+1}}$ atunci au loc relațiile, $P_n(x_i) = y_i$, de unde rezultă $L_{n,k}(x_i) = \delta_{ik}$. Deoarece $L_{n,k}$ sunt polinoame de gradul n și $L_{n,k}(x_i) = 0, i \neq k$ rezultă că $x_1, x_2, ..., x_{k-1}$, $x_{k+1}, ..., x_{n+1}$ sunt n rădăcini, deci $L_{n,k}$ se reprezintă:

$$L_{n,k}(x) = C_k(x - x_1)(x - x_2)\dots(x - x_{k-1})(x - x_{k+1})\dots(x - x_{n+1}), \quad (6)$$

iar din condiția $L_{n,k}(x_k) = 1$, rezultă relația pentru C_k :

$$C_k = \frac{1}{(x_k - x_1)(x_k - x_2)\dots(x_k - x_{k-1})(x_k - x_{k+1})\dots(x - x_{n+1})} \quad (7)$$

Se înlocuiesc C_k în (6) și se obțin expresiile

$$L_{n,k}(x) = \frac{(x - x_1)\dots(x - x_{k-1})(x - x_{k+1})\dots(x - x_{n+1})}{(x_k - x_1)\dots(x_k - x_{k-1})(x_k - x_{k+1})\dots(x_k - x_{n+1})}, \quad x \in \mathbb{R}, \quad k = \overline{1, n+1} \quad (8)$$

Funcțiile $L_{n,k}$ se numesc funcții de bază pentru interpolarea Lagrange și se vor rescrie sub o formă compactă

$$L_{n,k}(x) = \prod_{\substack{j=1 \\ j \neq k}}^{n+1} \frac{x - x_j}{x_k - x_j} \quad (9)$$

Vom nota în continuare $e_t(x) = f(x) - P_n(x)$ eroarea interpolării în fiecare punct.

Teorema (IV.1. Estimarea erorii de interpolare)

Fie $n \geq 1$, funcția $f : [a, b] \longrightarrow \mathbb{R}$ a.i. $f \in C^{n+1}[a, b]$ și diviziunea $(x_i)_{i=\overline{1,n+1}}$ a intervalului $[a, b]$. Atunci: $\forall\, x \in [a, b]$, $\exists\, \xi \in (a, b)$ astfel încât

$$e_t(x) := f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\pi_{n+1}(x)$$

unde

$$\pi_{n+1}(x) := (x - x_1)(x - x_2)\dots(x - x_{n+1}), \quad x \in [a, b]$$

Mai mult, are loc următoarea estimare a erorii de interpolare:

$$|f(x) - P_n(x)| \leq \frac{\max\limits_{\zeta \in [a,b]} |f^{(n+1)}(\zeta)}{(n+1)!} |\pi_{n+1}(x)|, \quad \forall\, x \in [a, b]$$

Exemplu 2: Să se afle, prin metoda Lagrange, polinomul de interpolare Lagrange $P_2(x)$ a funcţiei $f(x) = e^{2x}$ relativ la diviziunea $(-1; 0; 1)$.
Rezolvare: Polinomul $P_2(x)$ conform metodei Lagrange este $P_2(x) = L_{2,1}(x)y_1 + L_{2,2}(x)y_2 + L_{2,3}(x)y_3$, unde

$$L_{2,1}(x) = \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_2-x_3)} = \frac{x(x-1)}{2} = \frac{x^2-x}{2}$$
$$L_{2,2}(x) = \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} = \frac{(x+1)(x-1)}{-1} = 1-x^2$$
$$L_{2,3}(x) = \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)} = \frac{x(x+1)}{2} = \frac{x^2+x}{2}$$

Astfel,

$$\begin{aligned} P_2(x) &= \frac{x^2-x}{2}\cdot e^{-2} + (1-x^2) + \frac{x^2+x}{2}\cdot e^2 \\ &= 1 + \frac{e^2-e^{-2}}{2}x + \frac{e^2+e^{-2}-2}{2}x^2 \end{aligned}$$

IV.3. Metoda Newton de determinare a polinomului Lagrange P_n.

Se consideră următoarea reprezentare a polinomului Lagrange

$$P_n(x) = c_1 + c_2(x-x_1) + c_3(x-x_1)(x-x_2) + \ldots + c_{n+1}(x-x_1)(x-x_2)\ldots(x-x_n)$$

sau

$$P_n(x) = c_1 + \sum_{i=2}^{n+1} c_i \prod_{j=1}^{i-1}(x-x_j) \qquad (10)$$

Condiţiile $P_n(x_i) = y_i, i = \overline{1, n+1}$ ne furnizează sistemul de ecuaţii liniare necesar pentru determinarea coeficienţilor $c_i, i = \overline{1, n+1}$

$$\begin{cases} c_1 & = y_1 \\ c_1 + c_2(x_2-x_1) & = y_1 \\ c_1 + c_2(x_3-x_1) + c_3(x_3-x_1)(x_3-x_2) & = y_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ c_1 + c_2(x_{n+1}-x_1) + c_3(x_{n+1}-x_1)(x_{n+1}-x_2) + \ldots + & \\ + \ldots + c_{n+1}(x_{n+1}-x_1)\ldots(x_{n+1}-x_n) & = y_{n+1} \end{cases} \qquad (11)$$

Sistemul (11) este un sistem inferior triunghiular şi se rezolva conform metodei substituţii ascendente. Componentele matricei A asociată

sistemului sunt determinate conform relaţiilor:

$$\begin{aligned} a_{i1} &= 1, \quad i = \overline{1, n+1} \\ a_{ij} &= \prod_{k=1}^{j-1}(x_i - x_k), \quad i = \overline{2, n+1},\ j = \overline{2, i} \end{aligned} \qquad (12)$$

Exemplu 3: Să se afle, prin metoda Newton, polinomul de interpolare Newton $P_2(x)$ a funcţiei $f(x) = e^{2x}$ relativ la diviziunea $(-1; 0; 1)$.
Rezolvare: Polinomul $P_2(x)$ conform metodei Newton se reprezintă sub forma: $P_2(x) = c_1 + c_2(x-x_1) + c_3(x-x_1)(x-x_2)$. Din condiţiile $P_2(-1) = e^{-2}, P_2(0) = 1, P_2(1) = e^2$ rezultă sistemul

$$\begin{cases} c_1 & = e^{-2} \\ c_1 + c_2 & = 1 \\ c_1 + 2c_2 + 2c_3 & = e^2 \end{cases} \Leftrightarrow \begin{cases} c_1 & = e^{-2} \\ c_2 & = 1 - e^{-2} \\ c_3 & = \dfrac{e^{-2}+e^2-2}{2} \end{cases}$$

$$\begin{aligned} P_2(x) &= e^{-2} + (1-e^{-2})(x+1) + \frac{e^{-2}+e^2-2}{2}x(x+1) \\ &= 1 + \frac{e^2-e^{-2}}{2}x + \frac{e^2+e^{-2}-2}{2}x^2. \end{aligned}$$

IV.4. Metoda Newton cu diferenţe divizate de determinare a polinomului Lagrange P_n.

Definiţia (IV.1.)

Fie funcţia $f : [a, b] \longrightarrow \mathbb{R}$ *şi o diviziune* $(x_i)_{i=\overline{1,n+1}}$.

(i) *S.n. diferenţa divizată (DD) de ordin* 0 *a lui* f *în raport cu nodul* x_1*:*

$$f[x_1] := f(x_1)$$

(ii) *S.n. DD de ordin* 1 *a lui* f *în raport cu nodurile* x_1, x_2*:*

$$f[x_1, x_2] := \frac{f[x_2]-f[x_1]}{x_2-x_1}$$

(iii) *S.n. DD de ordin* 2 *a lui* f *în raport cu nodurile* x_1, x_2, x_3*:*

$$f[x_1, x_2, x_3] := \frac{f[x_2, x_3]-f[x_1, x_2]}{x_3-x_1}$$

Definiția (IV.1. continuare)

(iv) *S.n. DD de ordin n a lui f în raport cu nodurile $x_1, x_2, \dots, x_{n+1}$:*

$$f[x_1, x_2, \dots, x_{n+1}] := \frac{f[x_2, x_3, \dots, x_{n+1}] - f[x_1, x_2, \dots, x_n]}{x_{n+1} - x_1}$$

Teorema (IV.2. formula de interpolare a lui Newton cu DD)

Polinomul de interpolare Lagrange de gradul n asociat funcției $f : [a, b] \longrightarrow \mathbb{R}$ și nodurilor de interpolare $\{x_1, x_2, \dots, x_{n+1}\}$ este dat de formula

$$P_n(x) = f[x_1] + f[x_1, x_2](x - x_1) + \dots \tag{13}$$

$$+f[x_1, x_2, \dots, x_{n+1}](x - x_1)(x - x_2)\dots(x - x_n) \tag{14}$$

$$= f[x_1] + \sum_{i=2}^{n+1} f[x_1, \dots, x_i] \prod_{j=1}^{i-1}(x - x_j), \qquad x \in [a, b] \tag{15}$$

Construim în continuare următorul tabel cu diferențe divizate:

x_i	DD ordin 0	DD ordin 1	DD ordin 2	DD ordin 3	...
x_1	$\boxed{f[x_1] = f(x_1)}$ $\searrow$				
x_2	$f[x_2] = f(x_2)$ $\longrightarrow$ $\searrow$	$\boxed{f[x_1, x_2]}$ $\searrow$			
x_3	$f[x_3] = f(x_3)$ $\longrightarrow$ $\searrow$	$f[x_2, x_3]$ $\longrightarrow$ $\searrow$	$\boxed{f[x_1, x_2, x_3]}$ $\searrow$		
x_4	$f[x_4] = f(x_3)$ $\longrightarrow$ $\searrow$	$f[x_3, x_4]$ $\longrightarrow$ $\searrow$	$f[x_2, x_3, x_4]$ $\longrightarrow$ $\searrow$	$\boxed{f[x_1, x_2, x_3, x_4]}$ $\searrow$	
...	...	...	...	...	...

Fie matricea Q, matricea inferior triunghiulară definită astfel:

$$Q_{ij} = f[x_{i-j+1}, ..., x_i] \tag{16}$$

Se observă că elementele matricei coincid cu diferențele divizate din tabel.

x_i	DD ordin 0	DD ordin 1	DD ordin 2	DD ordin 3	...
x_1	$\boxed{f[x_1] = Q_{11}}$				
x_2	$f[x_2] = Q_{21}$	$\boxed{f[x_1, x_2] = Q_{22}}$			
x_3	$f[x_3] = Q_{31}$	$f[x_2, x_3] = Q_{32}$	$\boxed{f[x_1, x_2, x_3] = Q_{33}}$		
x_4	$f[x_4] = Q_{41}$	$f[x_3, x_4] = Q_{42}$	$f[x_2, x_3, x_4] = Q_{43}$	$\boxed{f[x_1, x_2, x_3, x_4] = Q_{44}}$	
...	...	...	...	...	...

Au loc următoarele relații:

$$f[x_{i-j+1}, ..., x_i] = \frac{f[x_{i-j+2}, ..., x_i] - f[x_{i-j+1}, ..., x_{i-1}]}{x_i - x_{i-j+1}}$$

$$= \frac{f[x_{i-(j-1)+1}, ..., x_i] - f[x_{(i-1)-(j-1)+1}, ..., x_{i-1}]}{x_i - x_{i-j+1}}$$

Obținem astfel o relație de recurență pentru componentele matricei Q:

$$Q_{ij} = \frac{Q_{i,j-1} - Q_{i-1,j-1}}{x_i - x_{i-j+1}}, \quad j = \overline{2, n+1}, i = \overline{j, n+1} \tag{17}$$

Prima coloană a matricei Q se calculează conform formulei:

$$Q_{i1} = f(x_i), i = \overline{1, n+1}.$$

Exemplu 2: Să se afle, prin metoda Newton cu DD, polinomul de interpolare Lagrange $P_2(x)$ al funcției $f(x) = e^{2x}$ relativ la diviziunea $(-1; 0; 1)$.

Rezolvare: Construim tabelul diferențelor divizate:

x_i	DD ordin 0	DD ordin 1	DD ordin 2
-1	e^{-2}		
0	1	$1 - e^{-2}$	
1	e^2	$e^2 - 1$	$\frac{e^{-2} + e^2 - 2}{2}$

Pentru reprezentarea polinomului $P_2(x)$ păstrăm din tabel doar elementele de pe diagonala principală, i.e., $e^{-2}, 1 - e^{-2}$ și $\frac{e^{-2} + e^2 - 2}{2}$. Se obține

$$P_2(x) = e^{-2} + (1 - e^{-2})(x + 1) + \frac{e^{-2} + e^2 - 2}{2}(x + 1)x.$$

ALGORITM (Metoda Newton cu diferențe divizate)

Date de intrare: $(x_i)_{i=\overline{1,n+1}}; (y_i)_{\overline{1,n+1}}; x;$
Date de ieșire: y;

STEP 1: Se determină matricea Q

$$Q_{i1} = f(x_i), i = \overline{1, n+1}$$
$$Q_{ij} = \frac{Q_{i,j-1} - Q_{i-1,j-1}}{x_i - x_{i-j+1}}, i = \overline{2, n+1}, j = \overline{2, i};$$

STEP 2: Determină $P_n = Q_{11} + \sum_{k=2}^{n+1} Q_{kk}(x - x_1)...(x - x_{k-1})$

STEP 3: $y = P_n$.

CURS #7

CONȚINUTUL CURSULUI #7:

II.2. Valori proprii. Metoda Jacobi de aproximare a valorilor proprii unei matrice simetrice.

Definiția (II.7.)

Fie $A \in \mathcal{M}_n(\mathbb{R})$. Un număr complex λ se numește valoare proprie a matricei A dacă $\exists\ v \in C^n \setminus \{0\}$ astfel încât $Av = \lambda v$. Vectorul v se numește vector propriu asociat valorii λ.

O formă echivalentă a relației $Av = \lambda v$ este:

$$(A - \lambda I_n)v = 0 \tag{1}$$

Se obține astfel un sistem omogen care depinde de parametrul λ și are drept necunoscute componentele $v_1, v_2, ..., v_n$ ale vectorului v. Acest sistem admite soluție nenulă dacă și numai dacă

$$det(A - \lambda I_n) = 0 \tag{2}$$

Definiția (II.8.)

Fie $A \in \mathcal{M}_n(\mathbb{R})$. Polinomul de grad n, $P_n(\lambda) = det(A - \lambda I_n)$ se numește polinomul caracteristic al matricei A.

Rădăcinile polinomului $P_n(\lambda)$ sunt valorile proprii ale matricei A. Mulțimea valorilor proprii ale matricei A se numește spectrul matricei A și se notează cu $\sigma(A)$.

Propoziție (II.3.)

Fie $A \in \mathcal{M}_n(\mathbb{R})$ și $\lambda_i, i = \overline{1,n}$, valorile proprii asociate matricei A.

a) *Dacă A este simetrică, atunci $\lambda_i \in \mathbb{R}, i = \overline{1,n}$;*
b) *Dacă A este simetrică și semipozitiv definită, atunci $\lambda_i \in \mathbb{R}_{\geq 0}, i = \overline{1,n}$;*
c) *Dacă A este simetrică și pozitiv definită, atunci $\lambda_i \in \mathbb{R}_{>0}, i = \overline{1,n}$.*

Definiția (II.9.)

Fie $A \in \mathcal{M}_n(\mathbb{R})$. Se definește raza spectrală $\rho(A)$ a matricei A astfel:

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i| \tag{3}$$

unde $\lambda_i \in \sigma(A), i = \overline{1,n}$. Dacă $\lambda = a + bi \in C$ atunci $|\lambda| = \sqrt{a^2 + b^2}$.

Definiția (II.10.)

Matricea $R(\theta) = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix}$ cu $\theta \in \mathbb{R}$ se numește matrice de rotații în două dimensiuni.

Obs.: Matricea $R(\theta)$ rotește vectorii în planul xOy cu unghiul θ în sensul acelor de ceasornic.

Exemplu 2: Vectorul $e_1 = (1,0)^T$ este vectorul $e_2 = (0,1)^T$ rotit cu $\theta = \dfrac{\pi}{2}$ conform acelor de ceasornic. Într-adevăr,

$$R\left(\frac{\pi}{2}\right) e_2 = \begin{pmatrix} cos(\pi/2) & sin(\pi/2) \\ -sin(\pi/2) & cos(\pi/2) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4}$$

$$= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e_1 \tag{5}$$

Definiția (II.11.)

Fie $n \in \mathbb{N}, p < q = \overline{1,n}$ și un unghi $\theta \in \mathbb{R}$. Cu notația $c = cos\theta, s = sin\theta$, se definește matricea de rotație Givens, matricea ortogonală

$$R^{pq}(\theta) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & c & \cdots & s & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & -s & \cdots & c & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \tag{6}$$

unde elementele c, s se află la intersecția liniilor p și q cu coloanele p și q.

Dacă $r_{k\ell}, k, \ell = \overline{1,n}$, sunt componentele matricei $R^{(pq)}(\theta)$, atunci acestea se exprimă prin:

$$r_{pp} = c, \quad r_{pq} = s, \quad r_{qp} = -s, \quad r_{qq} = c \tag{7}$$

$$r_{kl} = \delta_{kl}, \quad k, l = \overline{1,n}, \quad k, l \neq p, q. \tag{8}$$

Obs.: Pentru $n = 3$ matricea de rotație Givens rotește vectorii $u \in \mathbb{R}^3$ în planul generat de vectorii $e_p, e_q, p < q$.

Exemplu 3: Să se afle vectorul rotit în planul Oe_1e_3 al vectorului $e_1 = (1,0,0)^T$ cu un unghi $\theta = \dfrac{\pi}{2}$.

Se observă că $c = cos\dfrac{\pi}{2} = 0, s = sin\dfrac{\pi}{2} = 1$, de unde rezultă matricea de rotație

$$R^{(13)}\left(\frac{\pi}{2}\right) = \begin{pmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

Astfel că

$$R^{(13)}\left(\frac{\pi}{2}\right) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = -e_3$$

Dacă aplicăm matricea $R^{(pq)}(\theta)$ unui vector $a \in \mathbb{R}^n$, acesta își va schimba doar elementele p și q.
Fie

$$b = R^{(pq)}(\theta)a$$

sau scris pe componente

$$b_k = \sum_{s=1}^{n} r_{ks}a_s, k = \overline{1,n}$$

de unde rezultă

$$\begin{cases} b_p = r_{pp}a_p + r_{pq}a_q = ca_p + sa_q \\ b_q = r_{qp}a_p + r_{qq}a_q = -sa_p + ca_q \\ b_k = a_k, k = \overline{1,n}, k \neq p, q \end{cases} \tag{9}$$

Dacă aplicăm matricea $R^{(pq)}(\theta)$ unei matrice $A \in \mathcal{M}(\mathbb{R})$, atunci matricea A își va schimba doar liniile p și q. Fie

$$B = R^{(pq)}(\theta)A$$

sau scris pe componente

$$b_{k\ell} = \sum_{s=1}^{n} r_{ks} a_{s\ell}, k, \ell = \overline{1, n}$$

de unde rezultă

$$\begin{cases} b_{p\ell} = \sum_{s=1}^{n} r_{ps} a_{s\ell} = r_{pp} a_{p\ell} + r_{pq} a_{q\ell} = c a_{p\ell} + s a_{q\ell},\ \ \ell = \overline{1, n} \\ b_{q\ell} = \sum_{s=1}^{n} r_{qs} a_{s\ell} = r_{qp} a_{p\ell} + r_{qq} a_{q\ell} = -s a_{p\ell} + c a_{q\ell},\ \ \ell = \overline{1, n} \\ b_{k\ell} = a_{k\ell},\ \ k, \ell = \overline{1, n},\ \ k, \ell \neq p, q \end{cases} \tag{10}$$

Fie în continuare matricea B de forma

$$B = (R^{(pq)}(\theta)^T) A (R^{(pq)}(\theta) \tag{11}$$

Această transformare afectează atât liniile cât și coloanele p, q ale matricei A. În urma unui calcul elementar rezultă componentele matricei B :

$$\begin{cases} b_{ij} = a_{ij}, i, j \neq p, q, \\ b_{pj} = b_{jp} = c a_{pj} - s a_{qj}, j \neq p, q \\ b_{qj} = b_{jq} = s a_{pj} + c a_{qj}, j \neq p, q \\ b_{pp} = c^2 a_{pp} - 2cs a_{pq} + s^2 a_{qq} \\ b_{qq} = s^2 a_{pp} + 2cs a_{pq} + c^2 a_{qq} \\ b_{pq} = b_{qp} = sc(a_{pp} - a_{qq}) + (c^2 - s^2) a_{pq} \end{cases} \tag{12}$$

Ideea metodei Jacobi este să se aplice matricei A rotații succesive de forma (11) până se obține o matrice diagonală.
La fiecare rotație de forma $B = (R^{(pq)}(\theta)^T) A (R^{(pq)}(\theta))$ vom impune condiția ca elementele nediagonale b_{pq}, b_{qp} să fie nule, astfel

$$b_{pq} = b_{qp} = sc(a_{pp} - a_{qq}) + (c^2 - s^2) a_{pq} = 0$$

sau echivalent

$$a_{pq} cos2\theta + \frac{1}{2}(a_{pp} - a_{qq}) sin2\theta = 0$$

Dacă $a_{pp} \neq a_{qq}$ obținem

$$tg2\theta = \frac{2a_{pq}}{a_{qq} - a_{pp}}$$

de unde rezultă

$$\theta = \frac{1}{2} arctg \left(\frac{2a_{pq}}{a_{qq} - a_{pp}} \right)$$

Se observă că $\theta \in (-\frac{\pi}{4}, \frac{\pi}{4})$. Dacă $a_{pp} = a_{qq}$ rezultă $cos2\theta = 0$, deci $\theta = \frac{\pi}{4}$.

Metoda Jacobi aproximează valorile proprii ale unei matrice simetrice prin construirea unui șir de matrice, $(A_m)_{m \geq 0}$, obținut cu ajutorul matricilor de rotație, ale căror valori de pe diagonală converg către valorile proprii ale matricei A. Șirul $(A_m)_{m \geq 0}$ este construit conform schemei numerice:

$$\begin{cases} A_0 = A \\ A_m = (R^{pq}(\theta))^T A_{m-1} R^{pq}(\theta),\ \ m \geq 1 \end{cases} \tag{13}$$

Metoda Jacobi clasică presupune alegerea perechii (p, q) cu proprietatea ca

$$|a_{pq}^{(m)}| = \max_{i<j} |a_{ij}^{(m)}| \tag{14}$$

unde $a_{ij}^{(m)}, i, j = \overline{1, n}$ sunt elementele matricei curente A_m. În această manieră se vor elimina două elemente care sunt cele mai mari în valoarea absolută. Trebuie să remarcăm că elementele care s-au anulat la o iterație dată sunt în general înlocuite cu elemente nenule în timpul rotațiilor succesive. Vom repeta procedeul până când toate elementele nediagonale sunt mai mici decât o valoare ε (numită toleranța) sub care un număr este considerat 0.

Definiția (II.12.)

Se numește modul al matricei A numărul

$$|A| = \sqrt{\sum_{i \neq j = 1}^{n} a_{ij}^2} \tag{15}$$

ALGORITM (Metoda Jacobi de aproximare a valorilor proprii)

Date de intrare: $A = (a_{ij})_{i,j=\overline{1,n}}-$ simetrică, ε;

Date de ieșire: $\lambda_i \in \mathbb{R}, i = \overline{1,n}$.

STEP 1: while $|A| \geq \varepsilon$ do

Determină p, q a.î. $|a_{pq}| = \max\limits_{1 \leq i < j \leq n} |a_{ij}|$;

if $a_{pp} = a_{qq}$ then

$\theta = \dfrac{\pi}{4}$

else

$\theta = \dfrac{1}{2} arctg \dfrac{2a_{pq}}{a_{qq} - a_{pp}}$;

endif

$c = cos\theta, s = sin\theta$;

for $j = 1 : n$ do

if $j \neq p, q$ then

$u = a_{pj}c - a_{qj}s$; $v = a_{pj}s + a_{qj}c$;

$a_{pj} = u$; $a_{qj} = v$; $a_{jp} = u$; $a_{jq} = v$;

endif

endfor

$u = c^2 a_{pp} - 2csa_{pq} + s^2 a_{qq}$;

$v = s^2 a_{pp} + 2csa_{pq} + c^2 a_{qq}$;

$a_{pp} = u$; $a_{qq} = v$;

$a_{pq} = 0$; $a_{qp} = 0$;

endwhile

STEP 2: $\lambda_i = a_{ii}, i = \overline{1,n}$.

Teorema (II.3.)

Fie $n \geq 3, \lambda_n \geq \lambda_{n-1} \geq ... \geq \lambda_1$ valorile proprii ale matricei simetrice A și $\alpha_n^{(m)} \geq \alpha_{n-1}^{(m)} \geq ... \geq \alpha_1^{(m)}$ elementele diagonale ale matricei A_m construită iterativ conform formulei (13) unde p, q, θ sunt calculați conform algoritmului (Metoda Jacobi). Atunci

$$|\lambda_i - \alpha_i^{(m)}| \leq |A_m| \leq q^m |A|, \forall i = \overline{1,n} \tag{16}$$

cu $q = \sqrt{1 - \dfrac{2}{n^2 - n}}$.

Se observă că $\lim\limits_{m \to \infty} \alpha_i^{(m)} = \lambda_i, i = \overline{1,n}$.

Exemplu 4: Fie matricea

$$A = \begin{pmatrix} 17 & -2 & 3\sqrt{3} \\ -2 & 8 & 2\sqrt{3} \\ 3\sqrt{3} & 2\sqrt{3} & 11 \end{pmatrix} \tag{17}$$

Folosind metoda Jacobi, să se determine valorile proprii ale matricei A. Evident că A este simetrică. Se determină $p < q$, astfel încât $|a_{pq}| = \max\limits_{1 \leq i < j \leq 3} |a_{ij}|$. Se observă că

$$|a_{pq}| = \max\{|a_{12}|, |a_{13}|, |a_{23}|\} = |a_{13}| = 3\sqrt{3} \Rightarrow p = 1, q = 3$$

Deoarece $a_{11} \neq a_{33}$ rezultă

$$\theta = \frac{1}{2} arctg \frac{2a_{13}}{a_{33} - a_{11}} = \frac{1}{2} arctg(-\sqrt{3}) = -\frac{\pi}{6}$$

Matricea de rotație este:

$$R^{(13)}\left(\frac{\pi}{6}\right) = \begin{pmatrix} cos\frac{\pi}{6} & 0 & -sin\frac{\pi}{6} \\ 0 & 1 & 0 \\ sin\frac{\pi}{6} & 0 & cos\frac{\pi}{6} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{pmatrix}$$

Se recalculează A :

$$A = \left(R^{(13)}\left(\frac{\pi}{6}\right)\right)^T AR^{(13)}\left(\frac{\pi}{6}\right) = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 8 & 4 \\ 0 & 4 & 8 \end{pmatrix}$$

Se reia algoritmul.

$$|a_{pq}| = \max\{|a_{12}|, |a_{13}|, |a_{23}|\} = |a_{23}| = 4 \Rightarrow p = 2, q = 3$$

Deoarece $a_{22} = a_{33}$ rezultă $\theta = \frac{\pi}{4}$. Atunci

$$R^{(23)}\left(\frac{\pi}{4}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\frac{\pi}{4} & sin\frac{\pi}{4} \\ 0 & -sin\frac{\pi}{4} & cos\frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Se recalculează matricea A :

$$A = \left(R^{(23)}\left(\frac{\pi}{4}\right)\right)^T AR^{(23)}\left(\frac{\pi}{4}\right) = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 12 \end{pmatrix}$$

Procesul iterativ se oprește datorită faptului că toate elementele nediagonale ale matricei A sunt nule.

CURS #5

CONȚINUTUL CURSULUI #5:

II.1.9. Metoda Cholesky.

Definiția (II.5.)

Fie $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$. Numim descompunerea Cholesky a matricei A, descompunerea de forma

$$A = LL^T \tag{1}$$

unde $L = (\ell_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$ este o matrice inferior triunghiulară.

Definiția (II.6.)

Fie $A \in \mathcal{M}_n(\mathbb{R})$.

a) *A se numește simetrică dacă și numai dacă $A^T = A$;*

b) *A se numește semipozitiv definită dacă și numai dacă $< Av, v >\geq 0, \forall v \in \mathbb{R}^n$;*

c) *A se numește pozitiv definită dacă și numai dacă $< Av, v >> 0, \forall v \in \mathbb{R}^n \setminus \{0\}$.*

$< \cdot, \cdot >$ reprezintă produsul scalar pe $\mathbb{R}^n$ definit astfel: $< u, v >= \sum_{i=1}^n u_i v_i, \forall u, v \in \mathbb{R}^n$.

Teorema (II.2.)

Dacă $A \in \mathcal{M}_n(\mathbb{R})$ este o matrice simetrică și pozitiv definită, atunci descompunerea Cholesky există.

Obs.:Pentru a arăta că A este pozitiv definită se va folosi criteriul lui Sylvester și anume: Matricea simetrică $A \in \mathcal{M}_n(\mathbb{R})$ este pozitiv definită dacă și numai dacă toți minorii principali, i.e. $detA_k > 0, A_k = (a_{ij})_{i,j=\overline{1,k}}$.

Propoziție (II.2.)

Fie $A \in \mathcal{M}_n(\mathbb{R})$ simetrică și pozitiv definită. Dacă componentele $\ell_{kk}, k = \overline{1,n}$ de pe diagonala principală a matricei L din descompunerea Cholesky sunt strict pozitive, atunci descompunerea este unică.

CALCULUL MATRICEI L : Relația $A = LL^T$ se va scrie astfel:

$$\begin{pmatrix} a_{kk} \cdots a_{ki} \\ \vdots \quad\quad \vdots \\ a_{ik} \cdots a_{ii} \end{pmatrix} = \begin{pmatrix} \ell_{kk} \quad\quad 0 \\ \vdots \ddots \\ \ell_{ik} \cdots \ell_{ii} \end{pmatrix} \begin{pmatrix} \ell_{kk} \cdots \ell_{ik} \\ \ddots \vdots \\ 0 \quad\quad \ell_{ii} \end{pmatrix} \tag{2}$$

Presupunem că printr-o anumită metodologie au fost calculate primele $k-1$ coloane din L, deci și primele linii $k-1$ din L^T.

ETAPA 1: Calculăm elementul ℓ_{kk} de pe diagonala principală, scriind expresia pentru a_{kk} :

$$a_{kk} = \sum_{s=1}^{n} \ell_{ks}\ell^T_{sk} = \sum_{s=1}^{k} \ell_{ks}\ell^T_{sk} = \sum_{s=1}^{k} \ell^2_{ks} = \ell^2_{kk} + \sum_{s=1}^{k-1} \ell^2_{ks}$$

Cum $\ell_{kk} > 0$ va rezulta

$$\ell_{kk} = \sqrt{a_{kk} - \sum_{s=1}^{k-1} \ell^2_{ks}} \tag{3}$$

ETAPA 2: Calculăm restul elementelor de pe coloana k, i.e. $\ell_{ik}, i > k$, scriind expresia pentru a_{ik} :

$$a_{ik} = \sum_{s=1}^{n} \ell_{is}\ell_{sk}^T = \sum_{s=1}^{k} \ell_{is}\ell_{sk}^T = \sum_{s=1}^{k} \ell_{is}\ell_{ks} = \ell_{ik}\ell_{kk} + \sum_{s=1}^{k-1} \ell_{is}\ell_{ks} \Rightarrow$$

$$\ell_{ik} = \frac{1}{\ell_{kk}}\left(a_{ik} - \sum_{s=1}^{k-1} \ell_{is}\ell_{ks}\right) \quad (4)$$

ALGORITM (Metoda Cholesky)

Date de intrare: $A = (a_{ij})_{i,j} = \overline{1,n}$; $b = (b_i)_{i=\overline{1,n}}$;
Date de ieșire: $L = (\ell_{ij})_{i,j=\overline{1,n}}$; $x = (x_i)_{i=\overline{1,n}}$

STEP 1: $\alpha = a_{11}$;
    if $\alpha \leq 0$ then
      OUTPUT('A nu este pozitiv definită');
      STOP.
    endif

$\ell_{11} = \sqrt{a_{11}}$;
for $i = 2 : n$ do
  $\ell_{i1} = \dfrac{a_{i1}}{\ell_{11}}$;
endfor

STEP 2: for $k = 2 : n$ do
  $\alpha = a_{kk} - \sum_{s=1}^{k-1} \ell_{ks}^2$;
  if $\alpha \leq 0$ then
    OUTPUT('A nu este pozitiv definită');
    STOP.
  endif
  $\ell_{kk} = \sqrt{\alpha}$;

    for $i = k + 1 : n$ do
      $\ell_{ik} = \dfrac{1}{\ell_{kk}}\left(a_{ik} - \sum_{s=1}^{k-1} \ell_{is}\ell_{ks}\right)$;
    endfor
  endfor

STEP 3: $y =$SubsAsc(L, b);
STEP 4: $x =$SubsDesc(L^T, y).

Exemplu 2: Fie $A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 10 & 4 \\ 2 & 4 & 6 \end{pmatrix}$

a) Să se verifice dacă A este simetrică și pozitiv definită;
b) În caz afirmativ, să se determine factorizarea Cholesky.
c) Să se rezolve sistemul $Ax = b$, $b = (12\ 30\ 10)^T$ prin metoda Cholesky.

a) $4 > 0, \begin{vmatrix} 4 & 2 \\ 2 & 10 \end{vmatrix} = 36 > 0, det(A) = 144 > 0 \Rightarrow$ conform criteriului Sylvester rezultă că matricea A este pozitiv definită. Conform Th. (II.2.) matricea A admite descompunere Cholesky. Astfel $\exists L$ inferior triunghiulară astfel încât $A = LL^T$.

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 10 & 4 \\ 2 & 4 & 6 \end{pmatrix} = \begin{pmatrix} \ell_{11} & 0 & 0 \\ \ell_{21} & \ell_{22} & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} \end{pmatrix} \begin{pmatrix} \ell_{11} & \ell_{21} & \ell_{31} \\ 0 & \ell_{22} & \ell_{32} \\ 0 & 0 & \ell_{33} \end{pmatrix}$$

$$= \begin{pmatrix} \ell_{11}^2 & \ell_{11}\ell_{21} & \ell_{11}\ell_{31} \\ \ell_{21}\ell_{11} & \ell_{21}^2 + \ell_{22}^2 & \ell_{21}\ell_{31} + \ell_{22}\ell_{32} \\ \ell_{31}\ell_{11} & \ell_{31}\ell_{21} + \ell_{32}\ell_{22} & \ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2 \end{pmatrix}$$

Aflăm mai întâi elementul ℓ_{11} : $\ell_{11}^2 = 4 \Rightarrow \ell_{11} = 2$ ($\ell_{11} > 0$). Calculăm în continuare elementele de pe prima coloană din $L, (\ell_{21}, \ell_{31})$: $\ell_{21}\ell_{11} = 2$ și $\ell_{31}\ell_{11} = 2$ de unde rezultă $\ell_{21} = 1$ și $\ell_{31} = 1$. Continuăm procesul calculând elementul ℓ_{22} : $\ell_{21}^2 + \ell_{22}^2 = 10 \Rightarrow \ell_{22} = 3$ ($\ell_{22} > 0$). Aflăm elementul rămas pe coloana a II-a, i.e., ℓ_{32} : $\ell_{31}\ell_{21} + \ell_{32}\ell_{22} = 4 \Rightarrow \ell_{32} = 1$. În final calculăm elemntul ℓ_{33} : $\ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2 = 6 \Rightarrow \ell_{33} = 2$ ($\ell_{33} > 0$).

Am obținut $L = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 2 \end{pmatrix}$.

Rezolvăm sistemul $Ly = b$:

$$\begin{cases} 2y_1 = 12 \\ y_1 + 3y_2 = 30 \\ y_1 + y_2 + 2y_3 = 10 \end{cases} \Rightarrow \begin{cases} y_1 = 6 \\ y_2 = 8 \\ y_3 = -2 \end{cases}$$

În final se rezolvă sistemul $L^T x = y$:

$$\begin{cases} 2x_1 + x_2 + x_3 = 6 \\ 3x_2 + x_3 = 8 \\ x_1 + x_2 + 2x_3 = -2 \end{cases} \Rightarrow \begin{cases} x_1 = 2 \\ x_2 = 3 \\ x_3 = -1 \end{cases}$$

CURS #4

CONȚINUTUL CURSULUI #4:

II. Metode numerice de rezolvare a sistemelor liniare.
  II.1. Metode directe de rezolvare a sistemelor de ecuații liniare.
  II.1.7. Calculul rangului unei matrice cu ajutorul metodei Gauss cu pivotare parțială.
  II.1.8. Decompunerea LU.

II.1.7. Calculul rangului unei matrice cu ajutorul metodei Gauss cu pivotare parțială.

Definiția (II.3.)

Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$ o matrice nenulă. Spunem că matricea A are rangul r și notăm $rangA = r$, dacă A are un minor nenul de ordin r, iar toți minorii lui A de ordin mai mare decât r sunt nuli.

Fiind dat sistemul

$$Ax = b,$$

cu $A \in \mathcal{M}_n(\mathbb{R})$ și $b, x \in \mathbb{R}^n$ se disting următoarele cazuri:

- Sisemul $Ax = b$ este compatibil determinat, i.e. admite o soluție unică dacă și numai dacă $rangA = rang\bar{A} = n$;
- Sistemul $Ax = b$ este compatibil nedeterminat, i.e. admite o infinitate de soluții dacă și numai dacă $rangA = rang\bar{A} < n$;
- Sistemul $Ax = b$ este incompatibil, i.e. nu admite soluții, dacă și numai dacă $rangA \neq rang\bar{A}$.

ALGORITM (Rangul unei matrice folosind metoda de eliminare Gauss cu pivotare parțială)

Date de intrare: $A \in \mathcal{M}_{m,n}(\mathbb{R})$, tol;
Date de ieșire: rang

STEP 1: Se iniţializează linia, coloana şi rangul:
  $h = 1, k = 1, rang = 0$;
STEP 2: while $h \leq m$ and $k \leq n$ do
  - Se caută pivotul a_{pk} :
    $|a_{pk}| = \max\limits_{j=\overline{h,m}} |a_{jk}|$;
  - if (maximul este mai mic egal decat tol) then
    Se trece la următoarea coloană: $k = k + 1$;
    Se trece la următorul pas al buclei while;
  endif

- if $p \neq h$ then
  $L_p \leftrightarrow L_h$ (Se intershimbă liniile);
endif
- Se elimină elementele sub pivot:
  for $l = h + 1 : m$
    $m_{lk} = \frac{a_{lk}}{a_{hk}}$;
    $L_l \leftarrow L_l - m_{lk} L_h$
  endfor
- Se avansează pe linie
  $h = h + 1$;
- Se avansează pe coloană
  $k = k + 1$;
- Se creşte rangul
  $rang = rang + 1$;

endwhile

Exemplul # 1

Să se afle rangul matricei A folosind metoda GPP

$$A = \begin{pmatrix} 4 & -1 & 5 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ 3 & -1 & 8 & 0 \end{pmatrix}$$

Răspuns: Indicație: Se transformă matricea A conform algoritmului de mai sus și se calculează rangul fie numărând liniile nenule, fie conform definiției rangului.

$$A \sim \begin{pmatrix} 4 & -1 & 5 & 0 \\ 0 & \frac{1}{4} & -\frac{17}{4} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

Am găsit un minor $\begin{vmatrix} 4 & -1 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{vmatrix}$, de ordinul 3 nenul, iar unicul minor de ordinul 4, $|A|$, este zero. Concludem că $rang = 3$.

II.1.8. Decompunerea LU.

Am văzut, în secțiunea precedentă, că mai multe sisteme cu aceeași matrice pot fi tratate simultan aplicând metode de tip Gauss. În multe situații nu toți termenii din membrul drept sunt disponibili de la început. Putem dori, de exemplu, să rezolvăm sistemele de ecuații

$$Ax^{(1)} = b^{(1)}, Ax^{(2)} = b^{(2)}, ..., Ax^{(n)} = b^{(n)}$$

unde $b^{(2)}$ este o funcție de $x^{(1)}$, $b^{(3)}$ este o funcție de $x^{(2)}$ ș.a.m.d. Prin urmare, rezolvarea lor simultană numai este posibilă, urmând ca algoritmii Gauss sa fie aplicați pentru fiecare sistem în parte, mărindu-se considerabil numărul de iterații. În asemenea situații ne vin în ajutor metode de factorizare.

Definiția (II.4.)

Se numește descompunere (sau factorizarea) LU a unei matrice $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$, scrierea matricei A ca produs de două matrice, una inferior triunghiulară, notată cu L și alta superior triunghiulară, notată cu U, i.e.

$$A = LU \tag{1}$$

Propoziție (II.1.)

Fie $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$ o matrice care admite descompunerea LU cu $L = (\ell_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$ inferior triunghiulară, $\ell_{kk} = 1$, $k = \overline{1,n}$ și $U = (u_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$ superior triunghiulară. Atunci descompunerea este unică.

Odată calculate matricele L, U, sistemul $Ax = b$ se rezolvă imediat și anume:

$$Ax = b \Leftrightarrow LUx = b \Leftrightarrow \begin{cases} Ux =: y \\ Ly = b \end{cases} \Leftrightarrow \begin{cases} x = SubsDesc(U, y) \\ y = SubsAsc(L, b) \end{cases} \tag{2}$$

Teorema (II.1.)

Matricile se obțin prin metodele de eliminare Gauss și anume:

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{n1} & m_{n2} & m_{n3} & \cdots & 1 \end{pmatrix}, U = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{pmatrix} \tag{3}$$

unde:

$$m_{\ell k} = \frac{a_{\ell k}^{(k)}}{a_{kk}^{(k)}}, k = \overline{1, n-1}, l = \overline{k+1, n}$$

$$u_{kj} = a_{kj}^{(k)}, k = \overline{1, n-1}, j = \overline{k, n}, \;\; u_{n,n} = a_{n,n}^{(n-1)}$$

Notăm că $a_{kj}^{(k)}$ reprezintă componenta cu indici kj a matricei A la etapa k, conform algoritmului de eliminare Gauss.

Aplicând metoda Gauss cu sau fără pivotare, liniile vor fi permutate. Prin acest proces de schimbare a liniilor se obțin L și U astfel încât $A' = LU$, unde A' va fi matricea A cu liniile permutate.

Fie vectorul w vectorul cu pozițiile inițiale ale liniilor matricei A, i.e. $w = (1, 2, ..., n)$. În procesul de schimbare a liniilor vom reține aceste interschimbări în vectorul w. Mai exact, la interschimbarea a două linii $L_l \leftrightarrow L_k$ vom interschimba si elementele $w_l \leftrightarrow w_k$.

Deasemenea, în urma interschimbării de linii va fi afectată și matricea L și anume, se vor interschimba subliniile situate sub diagonala principală.

Vectorul b se va modifica după cum urmează:

$$b'_k = b_{w_k}, k = \overline{1, n}$$

În final, rezolvăm două sisteme triunghiulare $Ly = b', Ux = y$.

ALGORITM (Factorizarea LU cu GPP)

Date de intrare: $A \in \mathcal{M}_n(\mathbb{R})$;
Date de ieșire: $L = (\ell_{ij})_{i,j=\overline{1,n}}, U, w$

STEP 1: Iniţializăm $L = I_n$; $w = 1 : n$;
STEP 2: for $k = 1 : n-1$ do
- Se calculează p astfel încât $|a_{pk}| = \max\limits_{j=\overline{k,n}} |a_{jk}|$;
- if $a_{pk} = 0$ then
  OUTPUT('A nu admite factorizarea LU')
  STOP
  endif
- if $p \neq k$ then
  $L_p \leftrightarrow L_k$ (schimbă linia p cu linia k în A);
  $w_p \leftrightarrow w_k$;

  if $k > 1$ then
  $\ell_{p,1:k-1} \leftrightarrow \ell_{k,1:k-1}$; (schimbă sublinii în L)
  endif
  endif
- for $\ell = k+1 : n$ do
  $\ell_{\ell k} = \dfrac{a_{\ell k}}{a_{kk}}$;
  $L_\ell \leftarrow L_\ell - \ell_{\ell k} L_k$;
  endfor

STEP 3: • if $a_{nn} = 0$ then
  OUTPUT('A nu admite factorizarea LU')
  STOP
  endif
STEP 4: $U = A$.

Exemplul # 2

Să se rezolve prin metoda LU cu GFP sistemul $Ax = b$, unde

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 3 & 2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ 4 \\ 10 \end{pmatrix}$$

Aplicăm metoda Gauss fără pivotare. Inițializăm vectorul $w = (1, 2, 3)$. $k = 1$: Se caută primul $p = \overline{1,3}$ cu $|a_{p1}| \neq 0 \Rightarrow p = 2$. Interschimbăm $L_2 \leftrightarrow L_1$, deasemenea efectuăm aceeași interschimbare și în vectorul w, obținându-se astfel $w = (2, 1, 3)$. Se obține o matrice echivalentă cu A

$$A \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$

Multiplicatorii $m_{21} = \dfrac{a_{21}^{(1)}}{a_{11}^{(1)}} = 0, m_{31} = \dfrac{a_{31}^{(1)}}{a_{11}^{(1)}} = 3$. În urma transformării elementare $L_3 \leftarrow L_3 - 3L_1$ se obține

$$A \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & -2 \end{pmatrix}$$

$k=2$: Se caută primul $p=\overline{2,3}$ cu $|a_{p2}|\neq 0 \Rightarrow p=2$. Multiplicatorii $m_{32}=\dfrac{a_{32}^{(2)}}{a_{22}^{(2)}}=\dfrac{2}{1}=2$.

Se aplică în final transformarea $L_3 \leftarrow L_3 - 2L_2$

$$A\sim\begin{pmatrix}1&0&1\\0&1&2\\0&0&-6\end{pmatrix}$$

Obținem astfel matricele

$$L=\begin{pmatrix}1&0&0\\0&1&0\\3&2&1\end{pmatrix},\quad U=\begin{pmatrix}1&0&1\\0&1&2\\0&0&-6\end{pmatrix}$$

Obs.: Produsul matricelor L,U este:

$$LU=\begin{pmatrix}1&0&0\\0&1&0\\3&2&1\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&2\\0&0&-6\end{pmatrix}=\begin{pmatrix}1&0&1\\0&1&2\\3&2&1\end{pmatrix}=:A'$$

Matricea A' este matricea A cu liniile permutate.

Pentru a afla soluția sistemului $Ax=b$ vom aplica schimbarea pozițiilor elementelor vectorului b după cum urmează:

$$b_1'=b_{w_1}, b_2'=b_{w_2}\Rightarrow b'=(4\ 8\ 10)^T$$

Sistemul $Ly=b'$ se rescrie

$$\begin{cases}y_1=4\\y_2=8\\3y_1+2y_2+y_3=10\end{cases}\Rightarrow\begin{cases}y_1=4\\y_2=8\\y_3=-18\end{cases}$$

Din relația $Ux=y$ sau echivalent cu sistemul

$$\begin{cases}x_1+x_3=4\\x_2+2x_3=8\\-6x_3=-18\end{cases}\Rightarrow\begin{cases}x_1=1\\x_2=2\\x_3=3\end{cases}$$

Exemplu #3 Să se rezolve prin metoda LU cu GPP sistemul de la Exemplul #2.

Răspuns: $L=\begin{pmatrix}1&0&0\\0&1&0\\\frac{1}{3}&-\frac{2}{3}&1\end{pmatrix},\ U=\begin{pmatrix}3&3&1\\0&1&2\\0&0&2\end{pmatrix}, w=\begin{pmatrix}3&1&2\end{pmatrix}$

Exemplul # 4 Fie sistemul

$$\begin{cases}x_2+x_3=3\\2x_1+x_2+5x_3=5\\4x_1+2x_2+x_3=1\end{cases}$$

Să se afle facorizarea LU a matricei asociate A asociate sistemului, utilizând metoda GPP. Să se afle soluția sistemului conform factorizării LU.

Scriem matricea asociată sistemului și vectorul termenilor liberi:

$$A=\begin{pmatrix}0&1&1\\2&1&5\\4&2&1\end{pmatrix},\ b=\begin{pmatrix}3\\5\\1\end{pmatrix}$$

Inițializăm matricea L și vectorul w:

$$L=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\ w=\begin{pmatrix}1\\2\\3\end{pmatrix}$$

$k=1$: Se calculează $|a_{p1}|=\max\limits_{j=\overline{1,3}}|a_{jk}|=max\{|a_{11}|,|a_{21}|,|a_{31}|\}=|a_{31}|$, rezultă $p=3$. Se interschimbă atât liniile $L_1\leftrightarrow L_3$, cât și elementele $w_1\leftrightarrow w_3$. Se obține:

$$A\sim\begin{pmatrix}4&2&1\\2&1&5\\0&1&1\end{pmatrix},\ w=\begin{pmatrix}3\\2\\1\end{pmatrix}$$

Se efectuează transformarea elementară $L_2\leftarrow L_2-\dfrac{2}{4}L_1$, în timp ce multiplicatorii $\ell_{21}=\dfrac{1}{2}, \ell_{31}=0$. Se obțin matricile A și L actualizate:

$$A\sim\begin{pmatrix}4&2&1\\0&0&\frac{9}{2}\\0&1&1\end{pmatrix},\ L=\begin{pmatrix}1&0&0\\\frac{1}{2}&1&0\\0&0&1\end{pmatrix}$$

$k=2$: Se calculeză $|a_{p2}|=\max\limits_{j=\overline{2,3}}|a_{jk}|=max\{|a_{22}|,|a_{32}|\}=|a_{32}|$, rezultă $p=3$. Se interschimbă atât liniile $L_2\leftrightarrow L_3$, cât și elementele $w_2\leftrightarrow w_3$. La acest pas, deoarece $k>1$, se vor interschimba sublinii în matricea L și anume: $\ell_{p,1:k-1}\leftrightarrow \ell_{k,1:k-1}$ sau echivalent $\ell_{3,1}\leftrightarrow \ell_{2,1}$.

Se obține:

$$A \sim \begin{pmatrix} 4 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \frac{9}{2} \end{pmatrix}, L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}, \; w = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

Matricea $U = A = \begin{pmatrix} 4 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \frac{9}{2} \end{pmatrix}$

Cu ajutorul vectorului w se calculează elementele vectorului b' :

$$b' = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$$

Sistemul $Ly = b'$ se rescrie

$$\begin{cases} y_1 = 1 \\ y_2 = 3 \\ \frac{1}{2}y_1 + y_3 = 5 \end{cases} \Rightarrow \begin{cases} y_1 = 1 \\ y_2 = 3 \\ y_3 = \frac{9}{2} \end{cases}$$

Din relația $Ux = y$ sau echivalent cu sistemul

$$\begin{cases} 4x_1 + 2x_2 + x_3 = 1 \\ x_2 + x_3 = 3 \\ \frac{9}{2}x_3 = \frac{9}{2} \end{cases} \Rightarrow \begin{cases} x_1 = -1 \\ x_2 = 2 \\ x_3 = 1 \end{cases}$$

CONȚINUTUL CURSULUI #3:

II.1.1. Sisteme liniare superior triunghiulare

Definiția (II.1.)

a) *Matricea* $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$ *se numește matrice superior triunghiulară dacă și numai dacă elementele sub diagonala principală sunt nule, i.e.* $a_{ij} = 0, \forall i > j$;

b) *Un sistem liniar a cărui matrice asociată este superior triunghiulară se numește sistem superior triunghiular.*

Fie sistemul liniar $Ax = b, A \in \mathcal{M}_n(\mathbb{R})$ superior triunghiulară cu $a_{kk} \neq 0, k = \overline{1,n}$ și $b \in \mathbb{R}^n$. Sistemul superior triunghiular $Ax = b$ se scrie sub forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1k}x_k + \ldots + a_{1n}x_n = b_1 & (E_1) \\ \qquad a_{22}x_2 + \ldots + a_{2k}x_k + \ldots + a_{2n}x_n = b_2 & (E_2) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ \qquad\qquad a_{kk}x_k + \ldots + a_{kn}x_n = b_k & (E_k) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ \qquad\qquad\qquad a_{nn}x_n = b_n & (E_n) \end{cases} \tag{1}$$

Din (E_n) rezultă

$$x_n = \frac{b_n}{a_{nn}}. \tag{2}$$

Fie ecuația (E_k): $a_{kk}x_k + \sum_{j=k+1}^{n} a_{kj}x_j = b_k$. Dacă din ultimele $n-k$ ecuații sunt calculate componentele $x_j, j = \overline{k+1,n}$, atunci din (E_k) rezultă

$$x_k = \frac{1}{a_{kk}}\left(b_k - \sum_{j=k+1}^{n} a_{kj}x_j\right) \tag{3}$$

ALGORITM (Metoda substituției descendente)

Date de intrare: $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$; $b \in \mathbb{R}^n$;
Date de ieșire: $x \in \mathbb{R}^n$;

1. $x_n = \frac{1}{a_{nn}} b_n$; $k = n - 1$;
2. `while` $k > 0$ `do`

   $x_k = \frac{1}{a_{kk}}\left(b_k - \sum_{j=k+1}^{n} a_{kj}x_j\right)$;

   $k = k - 1$;

   `endwhile`

Definim în continuare conform Algoritmului (Metoda substituției descendente) procedura **SubsDesc** având sintaxa $x =$**SubsDesc**(A, b), unde x este soluția sistemului $Ax = b$.

II.1.2. Metoda Gauss fără pivotare.

Fie sistemul liniar $Ax = b, A \in \mathcal{M}_n(\mathbb{R}),\ b \in \mathbb{R}^n$. Sistemul $Ax = b$ se scrie sub forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1k}x_k + \ldots + a_{1n}x_n = b_1 & (\mathrm{E}_1) \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2k}x_k + \ldots + a_{2n}x_n = b_2 & (\mathrm{E}_2) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kk}x_k + \ldots + a_{kn}x_n = b_k & (\mathrm{E}_k) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nk}x_k + \ldots + a_{nn}x_n = b_n & (\mathrm{E}_n) \end{cases} \tag{4}$$

Se numește **pivot** al matricei $A = \left(a_{ij}\right)_{i,j=\overline{1,n}}$ orice element de pe diagonala principală a matricei A, i.e. $a_{kk},\ k \in \overline{1,n}$.
Metoda Gauss fără pivotare transformă matricea extinsă $\overline{A}$ folosind transformările elementare într-o matrice superior triunghiulară, obținându-se astfel un sistem compatibil cu sistemul inițial.
La fiecare pas $k = \overline{1, n-1}$ al metodei lui Gauss fără pivotare se alege drept pivot corespunzător coloanei k primul element nenul $a_{pk} \neq 0, p \geq k$ de pe coloana k a matricei transformate. Apoi se elimină toate elementele de pe coloana k situate sub pivot (folosind transformările elementare).

ALGORITM (Metoda Gauss fără pivotare)

Date: $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R}), \quad b \in \mathbb{R}^n$

1. $A = (A \mid b) = (a_{ij})_{i=\overline{1,n};j=\overline{1,n+1}}$ (matricea extinsă);
2. for $k = 1 : n-1$ do
   - Se caută primul p cu $k \leq p \leq n$ a.î. $a_{pk} \neq 0$;
   - if (nu a fost gasit p) then
     - OUTPUT('Sist. incomp. sau sist. comp. nedet.')
     - STOP.
   - endif
   - if $p \neq k$ then
     - $L_p \leftrightarrow L_k$ (schimbă linia p cu linia k)
   - endif
   - for $\ell = k+1 : n$ do

       - $m_{\ell k} = \dfrac{a_{\ell k}}{a_{kk}}$;
       - $L_\ell \leftarrow L_\ell - m_{\ell k} L_k$;
     - endfor
   - endfor
3. if $a_{nn} = 0$ then
   - OUTPUT('Sistem incomp. sau sist. comp. nedet.')
   - STOP.
   - endif
4. $x =$ **SubsDesc**$\left((a_{ij})_{i,j=\overline{1,n}}, (a_{i,n+1})_{i=\overline{1,n}}\right)$.

Exemplul #1 Să se rezolve, folosind metoda lui Gauss fără pivotare, sistemul liniar:

$$\begin{cases} x_2 + 2x_3 = 8 \\ x_1 + x_3 = 4 \\ 3x_1 + 2x_2 + x_3 = 10 \end{cases} \tag{5}$$

Matricea extinsă $\overline{A}$ asociată sistemului este:

$$\bar{A} = [A|b] = \left[\begin{array}{ccc|c} 0 & 1 & 2 & 8 \\ 1 & 0 & 1 & 4 \\ 3 & 2 & 1 & 10 \end{array}\right]$$

$k = 1 : a_{21} \neq 0 \Rightarrow p = 2$. Deoarece $p \neq k$ interschimbăm $L_p \leftrightarrow L_k$. Se obține o matrice echivalentă cu matricea $\overline{A}$.

$$\bar{A} \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 8 \\ 3 & 2 & 1 & 10 \end{array}\right]$$

Eliminăm toate elementele de pe prima coloană situate sub elementul $a_{11} = 1$ al matricei echivalente. Aplicăm următoarea transformare elementară $L_3 \leftarrow L_3 - \dfrac{3}{1}L_1$:

$$\bar{A} \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 8 \\ 0 & 2 & -2 & -2 \end{array}\right]$$

$k = 2 : a_{22} = 1 \neq 0$. Eliminăm elementul situat sub pivotul curent a_{22} aplicând transformarea $L_3 \leftarrow L_3 - 2L_2$

$$\bar{A} \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & -6 & -18 \end{array}\right]$$

Matricea finală este o matrice superior triunghiulară și reprezintă matricea asociată unui sistem compatibil cu sistemul inițial. Soluția sistemului este: $x_1 = 1, x_2 = 2, x_3 = 3$.

Exemplul # 2 Să se rezolve, folosind metoda lui Gauss fără pivotare, sistemul liniar:

$$\begin{cases} \varepsilon x_1 + x_2 = 1 & (\mathrm{E}_1) \\ x_1 + x_2 = 2 & (\mathrm{E}_2) \end{cases} \tag{6}$$

unde $\varepsilon = \mathrm{O}(10^{-20}) \ll 1$.
Transformăm matricea $A \in \mathcal{M}_2(\mathbb{R})$ într-o matrice superior triunghiulară

$$\bar{A} = \left[\begin{array}{cc|c} \varepsilon & 1 & 1 \\ 1 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} \varepsilon & 1 & 1 \\ 0 & 1 - \frac{1}{\varepsilon} & 2 - \frac{1}{\varepsilon} \end{array}\right]$$

Cum $a_{11} \neq 0$, s-a efectuat transformarea $L_2 \leftarrow L_2 - \dfrac{a_{21}}{a_{11}} L_1$

Obținem sistemul liniar superior triunghiular echivalent:

$$\begin{cases} \varepsilon x_1 + \qquad x_2 = 1 \\ \left(1 - \dfrac{1}{\varepsilon}\right) x_2 = 2 - \dfrac{1}{\varepsilon} \end{cases} \tag{7}$$

Sistemul liniar (7) se rezolvă prin metoda substituției descendente

$$\begin{cases} x_2 = \dfrac{2 - 1/\varepsilon}{1 - 1/\varepsilon} = \dfrac{2\varepsilon - 1}{\varepsilon - 1} \approx \dfrac{-1}{-1} = 1 \\ x_1 = \dfrac{1 - x_2}{\varepsilon} = \dfrac{1 - 1}{\varepsilon} = 0 \end{cases} \Longrightarrow \tag{8}$$

$$x_1 = 0 \quad \& \quad x_2 = 1$$

Verificare:

$$\begin{cases} \varepsilon x_1 + x_2 = 0 + 1 = 1 & (\mathrm{E}_1) \\ x_1 + x_2 = 0 + 1 = 1 & (\mathrm{E}_2) \end{cases} \tag{9}$$

Relațiile (9) implică faptul că soluția (8) a sistemului liniar (11), obținută prin metoda lui Gauss fără pivotare, conține o eroare foarte mare.

II.1.3. Metoda Gauss cu pivotare parțială.

La fiecare pas $k = \overline{1, n-1}$ al Algoritmului metodei Gauss fără pivotare se alege ca pivot corespunzător coloanei k elementul a_{pk} cu valoarea absolută cea mai mare de pe coloana k, aflat sub sau pe diagonala principală a matricei curente A, i.e.

$$|a_{pk}| = \max_{j=\overline{k,n}} |a_{jk}|, \quad p \in \overline{k, n} \tag{10}$$

ALGORITM (Metoda Gauss cu pivotare parțială)

Date: $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$, $b \in \mathbb{R}^n$

1. $A = (A \mid b) = (a_{i,j})_{i=\overline{1,n}, j=\overline{1,n+1}}$ `(matricea extinsă)`
2. `for` $k = 1 : n-1$ `do`

   `Determină primul indice` p, $(k \le p \le n)$

   `a.î.` $|a_{pk}| = \max_{j=\overline{k,n}} |a_{jk}|$

   `if` $a_{pk} = 0$ `then`

   `OUTPUT('Sist. incomp. sau comp. nedet.')`

   `STOP.`

   `endif`

```
      if p ≠ k then
          L_p ↔ L_k (schimbă linia p cu linia k)
      endif
        for ℓ = k+1 : n do
            m_ℓk = a_ℓk / a_kk;
            L_ℓ ← L_ℓ − m_ℓk L_k;
        endfor
    endfor
3. if a_nn = 0 then
      OUTPUT('Sist.  incomp.  sau comp.
    nedet.')
      STOP.
    endif
4. x = SubsDesc((a_ij)_{i,j=1,n}, (a_{i,n+1})_{i=1,n})
```


Exemplul #3 Să se rezolve, folosind metoda lui Gauss cu pivotare parțială, sistemul liniar:

$$\begin{cases} x_1 + x_3 = 4 \\ x_2 + 2x_3 = 8 \\ 3x_1 + 2x_2 + x_3 = 10 \end{cases} \tag{11}$$

Matricea extinsă $\overline{A}$ asociată sistemului este:

$$\bar{A} = [A|b] = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 8 \\ 3 & 2 & 1 & 10 \end{array}\right]$$

$k = 1 : |a_{p1}| = \max_{j=\overline{1,3}} |a_{j1}| = |a_{31}| \Rightarrow p = 3$. Interschimbăm $L_3 \leftrightarrow L_1$. Se obține matricea echivalentă cu $\overline{A}$

$$\bar{A} \sim \left[\begin{array}{ccc|c} 3 & 2 & 1 & 10 \\ 0 & 1 & 2 & 8 \\ 1 & 0 & 1 & 4 \end{array}\right]$$

În urma transformării elementare $L_3 \leftarrow L_3 - \dfrac{1}{3}L_1$ se obține:

$$\bar{A} \sim \left[\begin{array}{ccc|c} 3 & 2 & 1 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{array}\right]$$

$k = 2 : |a_{p2}| = \max_{j=\overline{2,3}} |a_{j2}| = |a_{22}| \Rightarrow p = 2$. Alpicăm transformarea elementară $L_3 \leftarrow L_3 - \dfrac{-2/3}{1} = L_3 + \dfrac{2}{3}L_2$

$$\bar{A} \sim \left[\begin{array}{ccc|c} 3 & 2 & 1 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 2 & 6 \end{array}\right]$$

Soluția sistemului este: $x_1 = 1, x_2 = 2, x_3 = 3$.

Exemplul #4 Să se rezolve, folosind metoda lui Gauss cu pivotare parțială, sistemul liniar:

$$\begin{cases} \varepsilon x_1 + x_2 = 1 & (\mathrm{E}_1) \\ x_1 + x_2 = 2 & (\mathrm{E}_2) \end{cases} \tag{12}$$

unde $\varepsilon = \mathrm{O}(10^{-20}) \ll 1$.
Scriem matricea extinsă asociată sistemului:

$$\bar{A} = \left[\begin{array}{cc|c} \varepsilon & 1 & 1 \\ 1 & 1 & 2 \end{array}\right] \tag{13}$$

Determinăm pivotul corespunzător coloanei $k = 1$ a matricei curente A:

$$|a_{p1}| = \max_{j=\overline{1,2}} |a_{j1}| = \max\left\{|\varepsilon|, 1\right\} = 1 \quad \Longrightarrow \quad p = 2 \tag{14}$$

Interschimbăm liniile $k = 1$ și $p = 2$, i.e. $\mathrm{L}_2 \longleftrightarrow \mathrm{L}_1$

Se obține matricea echivalenta

$$\bar{A} \sim \left[\begin{array}{cc|c} 1 & 1 & 2 \\ \varepsilon & 1 & 1 \end{array}\right]$$

În urma transformării $L_2 \leftarrow L_2 - \frac{a_{21}}{a_{11}} L_1$, i.e., $L_2 \leftarrow L_2 - \varepsilon L_1$ se obține:

$$\bar{A} \sim \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1-\varepsilon & 1-2\varepsilon \end{array}\right]$$

Obținem sistemul liniar superior triunghiular echivalent:

$$\begin{cases} x_1 + \quad x_2 = 2 \\ \quad (1-\varepsilon)\, x_2 = 1 - 2\varepsilon \end{cases} \tag{15}$$

Sistemul liniar (15) se rezolvă prin metoda substituției descendente

$$\begin{cases} x_2 = \dfrac{1-2\varepsilon}{1-\varepsilon} \approx \dfrac{1}{1} = 1 \\ x_1 = 2 - x_2 = 2 - 1 = 1 \end{cases} \implies \tag{16}$$

$$x_1 = x_2 = 1$$

Verificare:

$$\begin{cases} \varepsilon\, x_1 + x_2 = \varepsilon + 1 \approx 1 & (\mathrm{E}_1) \\ x_1 + x_2 = 1 + 1 = 2 & (\mathrm{E}_2) \end{cases} \tag{17}$$

Relațiile (17) implică faptul că soluția (16) a sistemului liniar (11), obținută prin metoda lui Gauss cu pivotare parțială, este acurată.

Exemplul $\#5$

Să se rezolve, folosind metoda lui Gauss cu pivotare parțială, sistemul liniar:

$$\begin{cases} x_1 + C\, x_2 = C & (\mathrm{E}_1) \\ x_1 + x_2 = 2 & (\mathrm{E}_2) \end{cases} \tag{18}$$

unde $C = \mathrm{O}(10^{20}) \gg 1$.

Transformăm matricea $A \in \mathcal{M}_2(\mathbb{R})$ într-o matrice superior triunghiulară, ținând seama de alegerea pivotului corespunzător coloanelor matricelor respective:

$$\bar{A} = \left[\begin{array}{cc|c} 1 & C & C \\ 1 & 1 & 2 \end{array}\right] \quad (\mathrm{L}_2) \longleftarrow \left(\mathrm{L}_2 - \frac{1}{2}\,\mathrm{L}_1\right) \quad \implies$$

$$\bar{A} \sim \left[\begin{array}{cc|c} 1 & C & C \\ 0 & 1-C & 2-C \end{array}\right]$$

Obținem sistemul liniar superior triunghiular echivalent:

$$\begin{cases} x_1 + C\, x_2 = C \\ \quad (1-C)\, x_2 = 2 - C \end{cases} \tag{19}$$

Sistemul liniar (19) se rezolvă prin metoda substituției descendente

$$\begin{cases} x_2 = \dfrac{2-C}{1-C} \approx \dfrac{-C}{-C} = 1 \\ x_1 = C - C\, x_2 = C - C = 0 \end{cases} \implies$$

$$x_1 = 0 \quad \& \quad x_2 = 1$$

Verificare:

$$\begin{cases} x_1 + C\, x_2 = 0 + C = C & (\mathrm{E}_1) \\ x_1 + x_2 = 0 + 1 = 1 & (\mathrm{E}_2) \end{cases} \tag{20}$$

Relațiile (20) implică faptul că soluția (20) a sistemului liniar (18), obținută prin metoda lui Gauss cu pivotare parțială, conține o eroare foarte mare.

Cauza erorii se datorează faptului că nu se ține seama de valoarea pivotului în raport cu valorile elementelor liniei sale. Se introduce astfel pivotărea totală.

II.1.4. Metoda Gauss cu pivotare totală.

La fiecare pas $k = \overline{1, n-1}$ al Algoritmului metodei Gauss fără pivotare alegem ca pivot elementul curent a_{pm} cu valoarea absolută cea mai mare din submatricea $\left(a_{ij}\right)_{i,j=\overline{k,n}}$, i.e.

$$\boxed{|a_{pm}| = \max_{i,j=\overline{k,n}} |a_{ij}|\,, \quad p, m \in \overline{k, n}} \tag{21}$$

Dacă $m \neq k$, atunci interschimbăm coloanele k și m. Dacă $p \neq k$, atunci interschimbăm liniile k și ℓ.
Obs.: La interschimbarea a două coloane se schimbă ordinea necunoscutelor în vectorul x.

ALGORITM (Metoda Gauss cu pivotare totală)

Date de intrare: $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$, $b \in \mathbb{R}^n$;
Date de ieșire: $x \in \mathbb{R}^n$;

1. $A = (A \mid b) = (a_{i,j})_{i=\overline{1,n}, j=\overline{1,n+1}}$ `(matricea extinsă);`
   $index_i = i, i = \overline{1, n}$;
2. `for` $k = 1 : n-1$ `do`
   `Determină primii indici` p, m $(k \leq p, m \leq n)$
   `a.î.` $|a_{pm}| = \max_{i,j=\overline{k,n}} |a_{ij}|$;
   `if` $a_{pm} = 0$ `then`
   `OUTPUT('Sist. incomp. sau comp. nedet.')`
   `STOP`
   `endif`
   `if` $p \neq k$ `then`
   $L_p \leftrightarrow L_k$ `(schimbă linia` p `cu linia` k`);`
   `endif`

   `if` $m \neq k$ `then`
   $C_m \leftrightarrow C_k$ `(schimbă coloana` m `cu coloana` k`);`
   $index_m \leftrightarrow index_k$ `(schimbă indicii nec.);`
   `endif`
   `for` $\ell = k+1 : n$ `do`
   $m_{\ell k} = \dfrac{a_{\ell k}}{a_{kk}}$;
   $L_\ell \leftarrow L_\ell - m_{\ell k} L_k$;
   `endfor`
   `endfor`
3. `if` $a_{nn} = 0$ `then`
   `OUTPUT('Sist. incomp. sau comp. nedet.')`
   `STOP.`
   `endif`
4. $y =$ **SubsDesc**$\left((a_{ij})_{i,j=\overline{1,n}}, (a_{i,n+1})_{i=\overline{1,n}}\right)$;
   $x_{index_i} = y_i, i = \overline{1, n}$ `(renumerotare nec.).`

Exemplul #6 Să se rezolve, folosind metoda lui Gauss cu pivotare totală, sistemul liniar:

$$\begin{cases} x_1 + C\,x_2 = C & (\mathrm{E}_1) \\ x_1 + x_2 = 2 & (\mathrm{E}_2) \end{cases} \tag{22}$$

unde $C = \mathrm{O}(10^{20}) \gg 1$.

$$\widetilde{A} = \left[\begin{array}{cc|c} 1 & C & C \\ 1 & 1 & 2 \end{array}\right] \tag{23}$$

Determinăm pivotul corespunzător coloanei $k = 1$ a matricei curente A căutând maximul elementelor matricei A:

$$|a_{pm}| = \max_{i,j=\overline{1,2}} |a_{ij}| = |C| = |a_{12}| \quad \Longrightarrow \quad p = 1\,,\ m = 2 \tag{24}$$

Cum $p = k$ și $m \neq k$, iterschimbăm coloanele $k = 1$ și $m = 2$.

$$\overline{A} \sim \left[\begin{array}{cc|c} C & 1 & C \\ 1 & 1 & 2 \end{array}\right] \quad \left(\mathrm{E}_2 - \frac{1}{C}\mathrm{E}_1\right) \longrightarrow (\mathrm{E}_2) \quad \Longrightarrow \tag{25}$$

$$\overline{A} \sim \left[\begin{array}{cc|c} C & 1 & C \\ 0 & 1 - \dfrac{1}{C} & 1 \end{array}\right] \tag{26}$$

Obținem sistemul liniar superior triunghiular echivalent:

$$\begin{cases} C\,x_2 + \quad x_1 = 2\,C \\ \left(1 - \dfrac{1}{C}\right) x_1 = 1 \end{cases} \Longleftrightarrow \begin{cases} x_2 + \quad x_1 = C \\ \dfrac{C-1}{C} x_1 = 1 \end{cases} \tag{27}$$

Sistemul liniar (27) se rezolvă prin metoda substituției descendente

$$\begin{cases} x_1 = \dfrac{C}{C-1} \approx \dfrac{C}{C} = 1 \\ x_2 = \dfrac{C - x_1}{C} = \dfrac{C-1}{C} \approx \dfrac{C}{C} = 1 \end{cases} \Longrightarrow \tag{28}$$

$$x_1 = x_2 = 1$$

Verificare:

$$\begin{cases} x_1 + Cx_2 = 1 + C \approx C & (\mathrm{E}_1) \\ x_1 + x_2 = 1 + 1 = 2 & (\mathrm{E}_2) \end{cases} \tag{29}$$

Relația (29) implică faptul că soluția (28) a sistemului liniar (22), obținută prin metoda lui Gauss cu pivotare totală, este acurată.

II.1.5. Sisteme liniare inferior triunghiulare

Definiția (II.2.)

a) *Matricea $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$ se numește inferior triunghiulară dacă și numai dacă elementele sub diagonala principală sunt nule, i.e. $a_{ij} = 0, \forall i > j$;*

b) *Un sistem liniar a cărui matrice asociată este inferior triunghiulară se numește sistem inferior triunghiular.*

Fie sistemul liniar $Ax = b$, unde $A \in \mathcal{M}_n(\mathbb{R})$ este inferior triunghiulară cu $a_{kk} \neq 0, k = \overline{1,n}$ și $b \in \mathbb{R}^n$. Sistemul inferior triunghiular $Ax = b$ se scrie sub forma

$$\begin{cases} a_{11}\,x_1 & = b_1 & (\mathrm{E}_1) \\ a_{21}\,x_1 + a_{22}\,x_2 & = b_2 & (\mathrm{E}_2) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\ a_{k1}\,x_1 + a_{k2}\,x_2 + \ldots + a_{kk}\,x_k & = b_k & (\mathrm{E}_k) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\ a_{n1}\,x_1 + a_{n2}\,x_2 + \ldots + a_{nk}\,x_k + \ldots + a_{nn}\,x_n & = b_n & (\mathrm{E}_n) \end{cases} \tag{30}$$

Din (E_1) rezultă

$$x_1 = \frac{b_1}{a_{11}}. \tag{31}$$

Fie ecuația (E_k): $a_{kk}x_k + \sum\limits_{j=1}^{k-1} a_{kj}x_j = b_k$. Dacă din primele $k-1$ ecuații sunt calculate componentele $x_j, j = \overline{1, k-1}$, atunci din (E_k) rezultă

$$x_k = \frac{1}{a_{kk}} \left(b_k - \sum_{j=1}^{k-1} a_{kj}x_j\right) \tag{32}$$

ALGORITM (Metoda substituției ascendente)

Date de intrare: $A = (a_{ij})_{i,j=\overline{1,n}}$; $b = (b_i)_{i=\overline{1,n}}$;
Date de ieșire: $x = (x_i)_{i=\overline{1,n}}$

STEP 1: $x_1 = \dfrac{1}{a_{11}} b_1$;

STEP 2: for $2:n$ do

$$x_k = \frac{1}{a_{kk}}\left(b_k - \sum_{j=1}^{k-1} a_{kj}\, x_j\right);$$

endfor

Definim în continuare conform Algoritmului (Metoda substituției ascendente) procedura **SubsAsc** având sintaxa $x =$**SubsAsc**(A, b), procedură care returnează soluția x a sistemului $Ax = b$.

II.1.6. Inversarea unei matrice aplicând metodele Gauss cu pivotare. Determinantul unei matrice.

Fie $A = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{R})$ inversabilă și A^{-1} invesa matricei. Inversa A^{-1} verifică relația

$$AA^{-1} = A^{-1}A = I_n.$$

Fie $x^{(k)} \in \mathbb{R}^n$, $k = \overline{1,n}$ coloana k a matricei A^{-1}, i.e.,

$$A^{-1} = cols(x^{(1)}, ..., x^{(k)}, ..., x^{(n)}).$$

Deasemenea, fie $e^{(k)} = (0, ..., 1, ..., 0)^T$, cu 1 pe poziția k, coloana k din matricea I_n. Atunci

$$AA^{-1} = I_n \Longleftrightarrow Ax^{(k)} = e^{(k)}, k = \overline{1,n} \tag{33}$$

Am obținut n sisteme liniare în care vectorii necunoscutelor sunt pe rând coloanele inversei și vor fi calculați conform unei metode de pivotare, fie de exemplu, metoda Gauss cu pivotare totală.

Sistemele (33) se pot rezolva și simultan dacă se consideră drept matrice extinsă, matricea formată din matricea A la care se adaugă cele n coloane ale matricei I_n.

Fie $a_{11}^{(1)}, a_{22}^{(2)}, ..., a_{n-1,n-1}^{(n-1)}, a_{nn}^{(n-1)}$ pivoții la fiecare etapă din algoritmii Gauss, atunci

$$|A| = (-1)^s a_{11}^{(1)} a_{11}^{(2)} a_{n-1,n-1}^{(n-1)} a_{nn}^{(n-1)} \tag{34}$$

unde s este numărul de interschimbări de linii și coloane, în funcție de metodă. Matricea A se modifică pe parcursul iterațiilor, din acest motiv s-a folosit notația cu indici sus pentru a face distincție între elementele matricei A la fiecare pas.

CURS #2

CONȚINUTUL CURSULUI #2:

I. Metode de aproximare a soluțiilor ecuațiilor neliniare.
   - I.3. Metoda secantei.
   - I.4. Metoda poziției false.

Notă: Textele scrise cu roșu reprezintă material suplimentar.

La pasul k, aproximarea x_k a soluției exacte x^* a ecuației $f(x) = 0, x \in [a, b]$ se obține prin intersecția cu axa Ox a secantei AB la graficul lui f, prin punctele $A\big(x_{k-1}, f(x_{k-1})\big)$ și $B\big(x_{k-2}, f(x_{k-2})\big)$. Prin urmare, nu se mai folosește tangenta la graficul lui f, deci nu mai este necesar caculul derivatei lui f.

$$AB : \frac{x - x_{k-1}}{x_{k-2} - x_{k-1}} = \frac{y - f(x_{k-1})}{f(x_{k-2}) - f(x_{k-1})} \tag{1}$$

$$\{x_k\} = AB \cap Ox \Rightarrow \frac{x_k - x_{k-1}}{x_{k-2} - x_{k-1}} = -\frac{f(x_{k-1})}{f(x_{k-2}) - f(x_{k-1})} \Rightarrow$$

$$x_k = x_{k-1} - f(x_{k-1})\frac{x_{k-1} - x_{k-2}}{f(x_{k-1}) - f(x_{k-2})}$$

sau

$$x_k = \frac{x_{k-2}f(x_{k-1}) - x_{k-1}f(x_{k-2})}{f(x_{k-1}) - f(x_{k-2})}, k \geq 2 \tag{2}$$

unde $x_0, x_1 \in [a, b]$

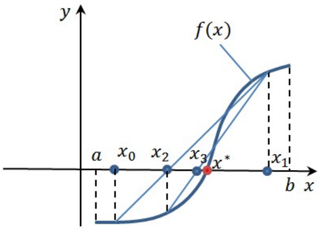

Figure: Metoda secantei

Teorema ((I.3.) Convergența metodei secantei)

Presupunem că $f \in C^1([a, b]), f(a)f(b) < 0, f'(x) \neq 0, \forall x \in [a, b]$. Atunci $\exists! x^*$ astfel încât $f(x^*) = 0$. Mai mult, $\exists \delta > 0$, astfel încât, dacă $x_0, x_1 \in [x^* - \delta, x^* + \delta] \subseteq [a, b]$, atunci șirul $(x_k)_{k \geq 0}$ construit prin metoda secantei rămâne în intervalul $[x^* - \delta, x^* + \delta]$ și converge către x^*.

Demonstrație: Existența și unicitatea este asigurată de faptul că $f(a)f(b) < 0$ și $f'(x) \neq 0, \forall x \in [a, b]$.
Deoarece $f'(x^*) \neq 0$, putem considera $f'(x^*) = \mu > 0$.
Din continuitatea derivatei f' rezultă că, pentru $\forall \varepsilon > 0, \exists \delta > 0$ astfel încât

$$|f'(x) - f'(x^*)| < \varepsilon, \quad \forall x \in [x^* - \delta, x^* + \delta] \subseteq [a, b] \tag{3}$$

sau

$$-\varepsilon + \mu < f'(x) < \varepsilon + \mu$$

Fie $\varepsilon = \frac{\mu}{4}$, atunci

$$\frac{3}{4}\mu < f'(x) < \frac{5}{4}\mu, \quad \forall x \in [x^* - \delta, x^* + \delta] \tag{4}$$

Conform metodei secantei

$$x_{k+1} = x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \tag{5}$$

Din dezvoltarea în serie Taylor a funcției f în vecinătatea punctului x_k și evaluată în x^* rezultă:

$$f(x^*) = f(x_k) + (x^* - x_k)f'(\xi_k), \quad \xi_k \in [x^*, x_k]$$

sau

$$f(x_k) = -(x^* - x_k)f'(\xi_k) \tag{6}$$

Mai mult, aplicând teorema Lagrange pe intervalul $[x_{k-1}, x_k]$ rezultă că $\exists \eta_k \in (x_{k-1}, x_k)$ astfel încât:

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} = f'(\eta_k) \tag{7}$$

Din (7) în (5) rezultă:

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

sau

$$x^* - x_{k+1} = x^* - x_k + \frac{f(x_k)}{f'(\eta_k)}, \tag{8}$$

iar conform cu (6) avem:

$$x^* - x_{k+1} = x^* - x_k - \frac{(x^* - x_k)f'(\xi_k)}{f'(\eta_k)}$$

sau

$$x^* - x_{k+1} = (x^* - x_k)\left(1 - \frac{f'(\xi_k)}{f'(\eta_k)}\right) \tag{9}$$

Din (4) rezultă următoarea estimare:

$$-\frac{2}{3} < 1 - \frac{f'(x)}{f'(y)} < \frac{2}{3} \tag{10}$$

Fie $x_0, x_1 \in [x^* - \delta, x^* + \delta]$. Presupunem că $x_k \in [x^* - \delta, x^* + \delta]$ și vom demonstra că și $x_{k+1} \in [x^* - \delta, x^* + \delta]$. Se observă că $\eta_k, \xi_k \in [x^* - \delta, x^* + \delta]$, iar conform relației (10), din (9) rezultă

$$|x^* - x_{k+1}| \leq \frac{2}{3}|x^* - x_k| \leq \frac{2}{3}\delta \tag{11}$$

Astfel că, $x_{k+1} \in [x^* - \delta, x^* + \delta]$, deci șirul $(x_k)_{k\geq 0}$ rămâne în intervalul $[x^* - \delta, x^* + \delta]$. Mai mult,

$$|x^* - x_{k+1}| \leq \frac{2}{3}|x^* - x_k| \leq ... \leq \left(\frac{2}{3}\right)^{k+1}|x^* - x_0| \tag{12}$$

rezultă că șirul $(x_k)_{k\geq 0}$ este convergent la x^*. □

Obs.: Se poate arăta că

$$lim_{k\to\infty}\frac{|x^* - x_{k+1}|}{|x^* - x_k|^r} = \alpha, \alpha > 0 \tag{13}$$

unde $r = \frac{1}{2}(1 + \sqrt{5}) \approx 1,62$, astfel că metoda secantei este mai rapidă decât metoda liniară dar mai lentă decât cea pătratică.

Din punct de vedere computațional valorile inițiale x_0, x_1 se aleg din vecinătatea soluției x^*, astfel încât la fiecare iterație se testează ca termenul x_k să rămână în intervalul $[a, b]$. Pentru optimizarea metodei se va alege intervalul maxim $[a, b]$ pe care funcția f este definită, nu-și schimbă monotonia (i.e. $f'(x) \neq 0, \forall x \in [a, b]$) și $f(a)f(b) < 0$.

ALGORITM (Metoda secantei)

Date de intrare: $f, a, b, x_0, x_1, \varepsilon$;

Date de ieșire: x_{aprox};

STEP1 : Se aleg $x_0, x_1 \in [a, b]$; $k = 1$;

STEP2 : while $\frac{|x_k - x_{k-1}|}{|x_{k-1}|} \geq \varepsilon$ do

$k = k + 1$;

$x_k = \frac{x_{k-2}f(x_{k-1}) - x_{k-1}f(x_{k-2})}{f(x_{k-1}) - f(x_{k-2})}$;

if $x_k < a$ or $x_k > b$ then

OUTPUT('Introduceţi alte valori pentru x_0, x_1');

STOP.

endif

endwhile;

STEP3 : $x_{aprox} = x_k$.

I.4. Metoda poziției false

Metoda poziției false construiește șirurile $(a_k)_{k\geq 0}, (b_k)_{k\geq 0}, (x_k)_{k\geq 0}$ conform următoarei scheme grafice: la pasul k, aproximarea x_k a soluției exacte x^* a ecuației $f(x) = 0$ se obține prin intersecția dreptei AB cu axa Ox, unde $A(a_k, f(a_k)), B(b_k, f(b_k))$. Intervalul $[a_k, b_k]$ se construiește conform metodei bisecției.

$$AB : \frac{x - a_k}{b_k - a_k} = \frac{y - f(a_k)}{f(b_k) - f(a_k)} \tag{14}$$

$$\{x_k\} = AB \cap Ox \Rightarrow \frac{x_k - a_k}{b_k - a_k} = \frac{-f(a_k)}{f(b_k) - f(a_k)} \Rightarrow \tag{15}$$

$$x_k = a_k - f(a_k)\frac{b_k - a_k}{f(b_k) - f(a_k)} \tag{16}$$

sau

$$x_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)} \tag{17}$$

Avem astfel următoarea schemă generală:

$$(a_k, b_k, x_k) \tag{18}$$

$$= \begin{cases} a_k = a_{k-1}, b_k = b_{k-1}, x_k = x_{k-1}, \text{ dacă } f(x_{k-1}) = 0 \\ a_k = a_{k-1}, b_k = x_{k-1}, x_k = \dfrac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}, \text{ dacă } f(a_{k-1})f(x_{k-1}) < 0 \\ a_k = x_{k-1}, b_k = b_{k-1}, x_k = \dfrac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}, \text{ dacă } f(a_{k-1})f(x_{k-1}) > 0, \end{cases} \tag{19}$$

unde $a_0 = a, b_0 = b, x_0 = \dfrac{a_0 f(b_0) - b_0 f(a_0)}{f(b_0) - f(a_0)}$.

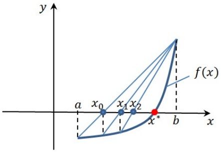

Figure: Metoda poziției false

Teorema (I.4. Teorema de convergență a metodei poziției false)

Presupunem că $f \in C^2([a, b]), f(a)f(b) < 0$ și f', f'' nu se anulează pe $[a, b]$. Atunci ecuația $f(x) = 0$ are o soluție unică $x^* \in (a, b)$, iar șirul $(x_k)_{k\geq 0}$ construit prin metoda poziției false converge la x^*.

ALGORITM (Metoda poziției false)

Date de intrare: f, a, b, ε; **Date de ieșire:** x_{aprox};

STEP1 : $k = 0; a_0 = a; b_0 = b; x_0 = \dfrac{a_0 f(b_0) - b_0 f(a_0)}{f(b_0) - f(a_0)}$;

STEP2 : do

$k = k + 1$;

if $f(x_{k-1}) = 0$ then

$x_k = x_{k-1}$;

STOP.

elseif $f(a_{k-1})f(x_{k-1}) < 0$ then

$a_k = a_{k-1}$; $b_k = x_{k-1}$; $x_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}$;

elseif $f(a_{k-1})f(x_{k-1}) > 0$ then

$a_k = x_{k-1}$; $b_k = b_{k-1}$; $x_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}$;

endif

while $\dfrac{|x_k - x_{k-1}|}{|x_{k-1}|} \geq \varepsilon$;

STEP3 : $x_{aprox} = x_k$.

CURS #1

CONȚINUTUL CURSULUI #1:

0.Erori

0.1. Erori de trunchiere

Fie x un parametru reprezentat prin valoarea sa exactă, F o funcție în baza căreia se evaluează exact o formulă matematică și F_t o funcție obținută în urma operației de trunchiere a formulei exacte.

Definiția (0.1.)

Definim $e_t(x) = F(x) - F_t(x)$ și numim eroare de trunchiere.

Teorema (0.1. Dezvoltarea în serie Taylor)

Fie $f : I \to \mathbb{R}$ o funcție de infinit ori derivabilă pe intervalul $I, x_0 \in I$ fixat. Atunci,

$$f(x) = \quad f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \ldots$$

Teorema (0.1. (reprezentarea restului sub forma Lagrage))

Mai mult, dacă f este derivabilă de $n+1$ ori, pentru orice $x \in I$, $\exists\, \xi$ cuprins între x_0 și x, astfel încât:

$$f(x) = \quad f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

Exemplul #1

Fie $x = e$, numărul Euler și formula matematică exactă de calcul a valorii acestui număr, obținută în baza dezvoltării în serie Taylor a funcției e^x în jurul punctului $x_0 = 0$ și evaluată în $x = 1$.

Întrucat, derivata funcției e^x rămâne la fel, dezvoltarea în serie Taylor a funcției $f(x) = e^x$ în jurul punctului x_0 este:

$$e^x \quad = e^{x_0} + e^{x_0}(x - x_0) + \frac{e^{x_0}}{2!}(x - x_0)^2 + \ldots + \frac{e^{x_0}}{n!}(x - x_0)^n + \ldots = \sum_{k=0}^{\infty} \frac{e^{x_0}}{k!}(x - x_0)^k$$

Notăm cu $F(x)$ formula exactă prin intermediul căreia se evaluează valoarea e^x pentru un x dat, astfel $e = F(1)$ sau

$$e = \sum_{k=0}^{\infty} \frac{1}{k!}$$

Fie $F_t(x)$ formula trunchiată de forma:

$$F_t(x) = \sum_{k=0}^{n} \frac{e^{x_0}}{k!}(x - x_0)^k \tag{1}$$

deci, numărul e poate fi aproximat cu

$$e \approx F_t(1) = \sum_{k=0}^{n} \frac{1}{k!}$$

Eroarea de trunchiere a acestei aproximări se determină prin formula:

$$e_t(1) = F(1) - F_t(1) = \sum_{k=n+1}^{\infty} \frac{1}{k!}$$

sau, folosind formula de reprezentare Lagrange a restului din dezvoltarea Taylor, obținem: $e_t(1) = \dfrac{e^{\xi}}{(n+1)!}$ cu ξ între 0 și 1.

Exemplul #2 Fie $g \in C^2([a,b])$, $x \in (a,b)$ și $h > 0$ astfel încât $x + h$ rămâne în intervalul (a,b). Conform formulei de dezvoltare în serie Taylor, cu reprezentarea restului sub forma Lagrange, putem scrie:

$$g(x+h) = g(x) + g'(x)h + g''(\xi)\frac{h^2}{2},\ \xi \in (x, x+h)$$

Din expresia de mai sus, putem deduce o formulă prin intermediul căreia se poate calcula prima derivată

$$g'(x) = \frac{g(x+h) - g(x)}{h} - g''(\xi)\frac{h^2}{2}$$

Întrucât, nu întotdeauna putem calcula analitic derivata funcției, formula de mai sus poate fi trunchiată la:

$$g'(x) \approx \frac{g(x+h) - g(x)}{h}$$

cu eroarea de trunchiere $e_t(x) = -g''(\xi)\frac{h^2}{2}$

0.2. Erori de rotunjire

Un *număr mașină* poate fi reprezentat în baza 10 cu virgulă mobilă normalizată sub forma:

$$x^* = \pm 0.d_1 d_2 ... d_k \times 10^n \tag{2}$$

$$0 \leq d_1, d_2, ..., d_k \leq 9,\ \ d_1 \neq 0 \tag{3}$$

cifrele $d_1, d_2, ..., d_k$ se numesc cifre semnificative.
Observație: Reprezentarea se numește normalizată deoarece cifra care precede virgula este zero.
Un număr real x, din punct de vedere matematic, se reprezintă cu o infinitate de cifre semnificative

$$x = \pm 0.d_1 d_2 ... d_k ... \times 10^n \tag{4}$$

Acest număr poate fi asimilat cu numărul mașină cu k cifre semnificative după următoarea regulă:

$$x^* = \begin{cases} \pm 0.d_1 d_2 ... d_k \times 10^n, \text{ dacă } d_{k+1} < 5 \\ \pm 0.d_1 d_2 ... (d_k + 1) \times 10^n, \text{ dacă } d_{k+1} \geq 5 \end{cases} \tag{5}$$

Exemplul #3 Să se afle numărul mașină cu 5 cifre semnificative al numărului $\pi = 3,14159265...$
Numărul π scris în reprezentarea normalizată are forma

$$\pi = +0.31415965... \times 10^1$$

iar numărul mașină cu 5 cifre semnificative asociat este

$$+0.31416 \times 10^1$$

Definiția (0.2.)

Presupunem că x^* reprezintă aproximarea numărului x. Notăm cu

$$e_a := |x - x^*|$$

și numim eroarea absolută a aproximării. Notăm cu

$$e_r := \frac{|x - x^*|}{|x|}$$

și numim eroarea relativă.

Exemplul #4 Să se determine erorile absolută și relativă, dacă x este aproximarea lui x^*, unde $x = 0,3000 \times 10^4$, $x^* = 0,3100 \times 10^4$.
Eroarea absolută $e_a = 0,01 \times 10^4 = 100$, iar eroarea relativă $e_r = \dfrac{100}{3 \times 10^3} = 0,1$.
Eroarea absolută nu este un indice suficient pentru stabilirea exactității calcului, pentru că acesta ne dă doar abaterea fără a ține cont de ordinul mărimilor implicate în calcul. Pentru a determina exactitatea calculului, eroarea absolută se raportează la mărimea valorii exacte, obținându-se eroarea relativă, sau dacă se dorește să se exprime în procente, aceasta se înmulțește cu 100%.

Definiția (0.3.)

Spunem că numărul x^* aproximează numărul x cu k cifre semnificative, dacă k este cel mai mare număr cu proprietatea

$$e_r = \frac{|x - x^*|}{|x|} \leq 5 \times 10^{-k} \tag{6}$$

Observație: Dacă x^* este numărul mașină asociat numărului x după regula (5), atunci are loc estimarea (6). Într-adevăr, fie $x = \pm 0.d_1 d_2 ... d_k ... \times 10^n$ și x^* definit prin:

$$x^* = \begin{cases} \pm 0.d_1 d_2 ... d_k \times 10^n, \text{ dacă } d_{k+1} < 5 \\ \pm 0.d_1 d_2 ... (d_k + 1) \times 10^n, \text{ dacă } d_{k+1} \geq 5 \end{cases}$$

Vom considera cele două cazuri separat.

Cazul 1. $d_{k+1} < 5$:

$$\frac{|x - x^*|}{|x|} = \frac{|0.0 \cdots 0 d_{k+1} \cdots| \times 10^n}{|0.d_1 d_2 \cdots d_k \cdots| \times 10^n} \leq \frac{d_{k+1} \times 10^{-(k+1)}}{0.1} = d_{k+1} \times 10^{-k} \leq 5 \times 10^{-k} \tag{7}$$

Cazul 2. $d_{k+1} \geq 5$: (sau, echivalent $d'_{k+1} := 10 - d_{k+1} \leq 5$)

$$\frac{|x - x^*|}{|x|} = \frac{|0.0 \cdots 0(10 - d_{k+1}) \cdots| \times 10^n}{|0.d_1 d_2 \cdots d_k \cdots| \times 10^n} \leq \frac{d'_{k+1} \times 10^{-(k+1)}}{0.1} = d'_{k+1} \times 10^{-k} \leq 5 \times 10^{-k} \tag{8}$$

I. Metode de aproximare a soluțiilor ecuațiilor neliniare

I.1. Metoda bisecției

Fie $f : [a, b] \to \mathbb{R}$ o funcție continuă, astfel încât $f(a)f(b) < 0$. Atunci $\exists x^* \in (a, b)$, astfel încât $f(x^*) = 0$.

Metoda bisecției generează un șir de aproximări $(x_k)_{k \geq 0}$ convergent către soluția exactă x^* a ecuației $f(x) = 0$ (i.e. $lim_{k\to\infty} x_k = x^*$, unde x^* verifică ecuația $f(x) = 0$).

Metoda bisecției constă în înjumătățirea la fiecare pas k a intervalului $[a, b]$ și selectarea acelui interval notat prin $[a_k, b_k]$ în care se află x^*.

Șirurile $(a_k)_{k\geq 0}$, $(b_k)_{k\geq 0}$ și $(x_k)_{k\geq 0}$ se construiesc conform schemei:

$$(a_k, b_k, x_k) = \begin{cases} a_k = a_{k-1}, b_k = b_{k-1}, x_k = x_{k-1}, & \text{dacă} \quad f(x_{k-1}) = 0 \\ a_k = a_{k-1}, b_k = x_{k-1}, x_k = \dfrac{a_k + b_k}{2}, & \text{dacă} \quad f(a_{k-1})f(x_{k-1}) < 0 \\ a_k = x_{k-1}, b_k = b_{k-1}, x_k = \dfrac{a_k + b_k}{2}, & \text{dacă} \quad f(a_{k-1})f(x_{k-1}) > 0, \end{cases} \tag{9}$$

unde $a_0 = a, b_0 = b, x_0 = \dfrac{a_0 + b_0}{2}$.

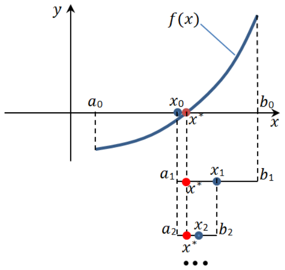

Figure: Metoda bisecției

Teorema (I.1.)

Fie $f : [a, b] \to \mathbb{R}$ continuă, $f(a)f(b) < 0$. Dacă f admite soluție unică $x^* \in (a, b)$ atunci șirul $(x_k)_{k\geq 0}$ este convergent la x^* și

$$|x^* - x_k| \leq \frac{b - a}{2^{k+1}}, \forall k \geq 0 \tag{10}$$

Demonstrație:

$$|x^* - x_k| \leq \frac{1}{2}|a_k - b_k| = \begin{cases} \frac{1}{2}|a_{k-1} - x_{k-1}|, f(a_{k-1})f(x_{k-1}) < 0 \\ \frac{1}{2}|x_{k-1} - b_{k-1}|, f(a_{k-1})f(x_{k-1}) > 0 \end{cases} \tag{11}$$

Constatăm că

$$\frac{1}{2}|a_{k-1} - x_{k-1}| = \frac{1}{2}|a_{k-1} - \frac{a_{k-1} + b_{k-1}}{2}| = \frac{1}{4}|a_{k-1} - b_{k-1}| \tag{12}$$

Analog

$$\frac{1}{2}|x_{k-1} - b_{k-1}| = \frac{1}{4}|a_{k-1} - b_{k-1}| \tag{13}$$

Astfel că, din (11) rezultă

$$0 \leq |x^* - x_k| \leq \frac{1}{4}|a_{k-1} - b_{k-1}| = \frac{1}{8}|a_{k-2} - b_{k-2}| = ... = \frac{1}{2^{k+1}}|a_0 - b_0| \tag{14}$$

sau $|x^* - x_k| \leq \frac{1}{2^{k+1}}|a-b|$ de unde rezultă $lim_{k\to\infty} x_k = x^*$. $\square$

Criteriul de oprire: Fiind dat $\varepsilon > 0$, se caută $N \in \mathbb{N}$ astfel încât $\frac{b-a}{2^{N+1}} < \varepsilon \Leftrightarrow N > log_2(\frac{b-a}{\varepsilon}) - 1 \Leftrightarrow N = \left[log_2(\frac{b-a}{\varepsilon})\right]$;

Definiția (I.1.)

Fie șirul $(x_k)_{k\geq 0}$ convergent la x^*. Spunem că șirul $(x_k)_{k\geq 0}$ converge ***cel puțin liniar*** *la x^*, dacă există șirul de numere reale pozitive $(\varepsilon_k)_{k\geq 0}$ convergent la zero și $\alpha \in (0,1)$ astfel încât*

$$|x_k - x^*| \leq \varepsilon_k, \quad k \geq 0 \quad \text{și} \quad lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \alpha \tag{15}$$

- Dacă relația (15) are loc pentru $\alpha = 0$, atunci spunem că șirul $(x_k)_{k\geq 0}$ converge **superliniar**;
- Dacă relația (15) are loc pentru $\alpha \in (0,1)$ și $\varepsilon_k = |x_k - x^*|$, $k \geq 0$, atunci spunem că șirul $(x_k)_{k\geq 0}$ converge **liniar**;
- Dacă (15) are loc pentru $\alpha = 1$ și $\varepsilon_k = |x_k - x^*|$, atunci viteza de convergență este mai lentă decât cea liniară și spunem că șirul $(x_k)_{k\geq 0}$ converge **subliniar**.

Definiția (I.2.)

Fie șirul $(x_k)_{k\geq 0}$ convergent la x^*. Spunem că șirul $(x_k)_{k\geq 0}$ converge la x^* cu ordinul de convergență cel puțin egal cu $r > 1$, dacă există un șir $(\varepsilon_k)_{k\geq}$ de numere reale pozitive convergent la 0 și $\alpha > 0$ astfel încât

$$|x_k - x^*| \leq \varepsilon_k, \quad k \geq 0 \quad \text{și} \quad lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k^r} = \alpha \tag{16}$$

Dacă (16) are loc pentru $\varepsilon_k = |x_k - x^*|$, $k \geq 0$, atunci spunem că șirul $(x_k)_{k\geq 0}$ converge la x^* cu **ordinul r de convergență**. În particular, dacă $r = 2$ atunci spunem că $(x_k)_{k\geq 0}$ converge **pătratic.**

Obs.: Datorită faptului că în cazul metodei bisecției avem estimarea $|x^* - x_k| \leq \frac{1}{2^{k+1}}(b-a)$ putem considera $\varepsilon_k = \frac{1}{2^{k+1}}(b-a)$. Atunci

$$lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \frac{1}{2} \in (0,1), \tag{17}$$

deci convergența este **cel puțin liniară**.

ALGORITM (Metoda bisecției)

Date de intrare: f, a, b, ε;

Date de ieșire: x_{aprox};

STEP 1: $a_0 = a$; $b_0 = b$; $x_0 = \frac{a_0+b_0}{2}$;

$N = \left[log_2(\frac{b-a}{\varepsilon}) - 1\right] + 1$;

STEP 2: for $k = 1 : N$ do

if $f(x_{k-1}) = 0$ then

$x_k = x_{k-1}$;

break

elseif $f(a_{k-1})f(x_{k-1}) < 0$ then

$a_k = a_{k-1}$; $b_k = x_{k-1}$; $x_k = \frac{a_k+b_k}{2}$;

elseif $f(a_{k-1})f(x_{k-1}) > 0$ then

$a_k = x_{k-1}$; $b_k = b_{k-1}$; $x_k = \frac{a_k+b_k}{2}$;

endif

endfor

$x_{aprox} = x_k$.

I.2.Metoda Newton-Raphson

Fie $f : [a,b] \to \mathbb{R}$ o funcție derivabilă astfel încât $f(a)f(b) < 0$. Metoda N-R presupune construcția șirului $(x_k)_{k\geq 0}$ conform următoarei scheme grafice: la pasul k, aproximarea x_k a soluției exacte x^* a ecuației $f(x) = 0$ se obține prin intersecția cu axa Ox a tangentei T la graficul funcției f în punctul $(x_{k-1}, f(x_{k-1}))$.

$$T : y = f'(x_{k-1})(x - x_{k-1}) + f(x_{k-1}) \tag{18}$$

$$\{x_k\} = T \cap Ox \Rightarrow f'(x_{k-1})(x_k - x_{k-1}) + f(x_{k-1}) = 0 \Rightarrow$$

$$x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})} \tag{19}$$

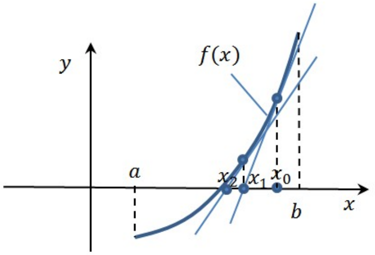

Figure: Metoda Newton

Teorema (I.2)

Presupunem că $f \in C^2([a,b])$, f', f'' nu se anulează pe $[a,b]$ și $f(a)f(b) < 0$. Fie $x_0 \in [a,b]$ astfel încât să aibă loc condiția

$$f(x_0)f''(x_0) > 0 \tag{20}$$

Atunci ecuația $f(x) = 0$ are o soluție unică $x^* \in (a,b)$, iar șirul $(x_k)_{k\geq 0}$ construit prin metoda Newton-Raphson, rămâne în $[a,b]$ și converge pătratic la x^*.

Demonstrație:
EXISTENȚA: Existența soluției ecuației $f(x) = 0$ este asigurată de condiția $f(a)f(b) < 0$.
UNICITATEA: Presupunem că $\exists y^* \in (a,b)$ cu $x^* \neq y^*$ și $f(y^*) = 0$. Cum $f(x^*) = f(y^*) = 0$, atunci conform Teoremei lui Rolle rezultă că $\exists c \in (x^*, y^*)$ astfel încât $f'(c) = 0$, contradicție, deoarece am presupus cu a f' este nenulă pe intervalul $[a,b]$.

CONVERGENȚA: Fără a restrânge generalitatea vom considera f', f'' strict pozitive, i.e. $f'(x) > 0, f''(x) > 0, \forall x \in [a,b]$. Celelalte cazuri se tratează în mod analog.
Fie $x_0 \in [a,b]$ cu proprietatea (20), atunci $f(x_0) > 0 = f(x^*)$. Deoarece $f'(x) > 0, \forall x \in [a,b]$ rezultă că f este strict crescătoare, astfel că $x^* < x_0 \leq b$ sau $x_0 \in (x^*, b]$.
Presupunem în continuare $x_k \in (x^*, b]$, i.e. $x^* < x_k \leq b$. Dezvoltăm în serie Taylor funcția f în jurul punctului x_k și evaluăm funcția în punctul x^* :

$$f(x^*) = f(x_k) + (x^* - x_k)f'(x_k) + \frac{1}{2}(x^* - x_k)^2 f''(\xi_k), \quad \xi_k \in [x^*, x_k] \tag{21}$$

Împărțim această relație la $f'(x_k)$, ținem cont că $f(x^*) = 0$ și $x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$. Obținem:

$$x_{k+1} - x^* = \frac{1}{2}(x^* - x_k)^2 \frac{f''(\xi_k)}{f'(x_k)}, \quad \xi_k \in [x^*, x_k] \tag{22}$$

Din monotonia funcției f rezultă $f(x_k) > 0 = f(x^*)$. Din (19) rezultă $x_{k+1} < x_k$, iar conform formulei (22) rezultă $x_{k+1} > x*$,

deci $x^* < x_{k+1} < x_k \leq b$. Am obținut că șirul $(x_k)_{k\geq 0}$ este descrescător și mărginit, deci convergent.
Fie $y^* = lim_{k\to\infty} x_k$, atunci trecând la limită în formula (19) rezultă:

$$y^* = y^* - \frac{f(y^*)}{f'(y^*)} \Rightarrow f(y^*) = 0, \tag{23}$$

deci y^* este soluție a ecuației $f(x) = 0$, iar din unicitatea soluției avem $x^* = y^*$.
Din relația (22) rezultă

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|^2} = \frac{1}{2}\frac{f''(\xi_k)}{f'(x_k)} \tag{24}$$

Dacă $\varepsilon_k = |x_k - x^*|$ atunci

$$lim_{k\to\infty}\frac{\varepsilon_{k+1}}{\varepsilon_k^2} = lim_{k\to\infty}\frac{|x_{k+1} - x^*|}{|x_k - x^*|^2} = lim_{k\to\infty}\frac{1}{2}\frac{|f''(\xi_k)|}{|f'(x_k)|} \tag{25}$$

$$= \frac{1}{2}\frac{|f''(x^*)|}{|f'(x^*)|} \in (0, \infty) \tag{26}$$

Rezultă că $(x_k)_{k\geq 0}$ converge **pătratic** la x^*. □

Deoarece f', f'' nu se anulează pe intervalul $[a, b]$, atunci funcția trebuie să fie monotonă (crescătoare sau descrescătoare) și să nu-și schimbe concavitatea pe intervalul dat.

Strategie de lucru: Din punct de vedere computațional se alege conform graficului funcției un interval în care funcția să fie monotonă și să nu-și schimbe concavitatea. Valoarea x_0 se alege în modul următor:

1. Dacă f este convexă ($f''(x_0) > 0$), atunci $f(x_0) > 0$;
2. Dacă f este concavă ($f''(x_0) < 0$), atunci $f(x_0) < 0$.

Pentru metoda N-R ca și criteriu de oprire vom alege una din următoarele condiții:

- $|f(x_k)| < \varepsilon$;
- $\dfrac{|x_k - x_{k-1}|}{|x_{k-1}|} < \varepsilon$.

ALGORITM (Metoda Newton-Raphson)

Date de intrare: f, f', x_0, ε;
Date de ieșire: x_{aprox};

STEP 1: $k = 0$;

STEP 2: do

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

while $\dfrac{|x_k - x_{k-1}|}{|x_{k-1}|} \geq \varepsilon$;

$x_{aprox} = x_k$.