# ANALIZĂ NUMERICĂ

Note de curs pentru anul III Informatică

de

lector dr. Daniel Stănică

(email: stanicad@fmi.unibuc.ro)

Analiza numerică (sau calculul numeric) studiază tehnici de aproximare pentru rezolvarea problemelor matematice, luând în considerare şi evaluarea posibilelor erori. Aceste tehnici de aproximare sunt transpuse în algoritmi implementaţi în computer prin intermediul unor diverse limbaje de programare (MATLAB, C++, JAVA, etc.), ţinându-se cont de puterea de calcul şi acuarteţea rezultatelor specifice calculatoarelor sau a limbajelor de programare.

Cursul va cuprinde următoarele **capitole**:

- Metode pentru aproximarea soluţiilor ecuaţiilor neliniare unidimensionale
- Metode pentru rezolvarea sau aproximarea soluţiilor sistemelor de ecuaţii liniare
- Metode pentru aproximarea valorilor proprii ale unei matrice
- Metode pentru aproximarea soluţiilor sistemelor de ecuaţii neliniare
- Aproximarea funcţiilor reale prin interpolare polinomială
- Construcţia curbelor de regresie
- Derivare numerică
- Integrare numerică: formule de cuadratură
- Metode pentru aproximarea soluţiilor ecuaţiilor diferenţiale

Bibliografie pentru lucrările de laborator:

Iuliana Paraschiv-Munteanu, Daniel Stănică - ANALIZĂ NUMERICĂ (Exerciţii şi teme de laborator), Editura Universităţii din Bucureşti, 2006.

Fenomenul de instabilitate manifestat în diverse procese matematice este important de studiat, dat fiind limitările tehnicii de calcul sau ale măsurătorilor de unde provin datele de calcul. Următorul exemplu scoate în evidenţă existenţa unui astfel de fenomen şi motivaţia studiului său.

Fie sistemul de ecuaţii $Ax = b$, unde

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \qquad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \qquad \text{cu soluţia} \qquad x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Considerăm sistemul perturbat $A(x + \delta x) = b + \delta b$, unde

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad b + \delta b = \begin{pmatrix} 32,1 \\ 22,9 \\ 33,1 \\ 30,9 \end{pmatrix}$$

Acesta are soluţia

$$x + \delta x = \begin{pmatrix} 9,2 \\ -12,6 \\ 4,5 \\ -1,1 \end{pmatrix} \tag{1}$$

Considerăm şi sistemul perturbat $(A + \Delta A)(x + \Delta x) = b$ unde

$$A + \Delta A = \begin{pmatrix} 10 & 7 & 8,1 & 7,2 \\ 7,08 & 5,04 & 6 & 5 \\ 8 & 5,98 & 9,89 & 9 \\ 6,99 & 4,99 & 9 & 9,98 \end{pmatrix} \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$$

Acesta are soluţia

$$x + \Delta x = \begin{pmatrix} -81 \\ 137 \\ -34 \\ 22 \end{pmatrix} \tag{2}$$

Sistemul (1) diferă de cel iniţial printr-o "mică" variaţie a coloanei termenilor liberi, iar sistemul (2) printr-o "mică" variaţie a elementelor matricei. După cum se observă, aceste "mici" variaţii antrenează după sine variaţii "mari" ale soluţiei iniţiale.

# Metode numerice pentru ecuaţii neliniare

În general, modelele matematice asociate fenomenelor fizice sau de altă natură conduc la ecuaţii de forma

$$f(x) = 0$$

Ne propunem să găsim aproximaţii ale soluţiei ecuaţiei anterioare, în cazul în care $f$ este o funcţie reală neliniară.

## Metoda bisecţiei

Această metodă constă în înjumătăţirea intervalului în care se află soluţia unei ecuaţii.

Fie $a < b \in R$ şi $f : [a, b] \to R$ o funcţie continuă cu $f(a) \cdot f(b) < 0$. Atunci există $z \in [a, b]$ astfel încât $f(z) = 0$.

Definim şirurile $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}, (c_n)_{n\geq 0}$ astfel:

- $a_0 := a$, $b_0 := b$, $c_0 := (a + b)/2$;
- Pentru $n \geq 1$
  - dacă $f(c_{n-1}) = 0$, atunci $\begin{cases} a_n := a_{n-1} \\ b_n := b_{n-1} \\ c_n := c_{n-1} \end{cases}$
  - dacă $f(a_{n-1}) \cdot f(c_{n-1}) < 0$, atunci $\begin{cases} a_n := a_{n-1} \\ b_n := c_{n-1} \\ c_n := (a_n + b_n)/2 \end{cases}$

$$- \text{ dacă } f(a_{n-1}) \cdot f(c_{n-1}) > 0, \text{ atunci } \begin{cases} a_n := c_{n-1} \\ b_n := b_{n-1} \\ c_n := (a_n + b_n)/2 \end{cases}$$

**Teorema -1.1** *Presupunem că funcţia $f$ are o singură rădăcină în $[a, b]$. Atunci şirul $(c_n)_{n \geq 0}$ construit mai sus converge la unica soluţie $z \in [a, b]$ a ecuaţiei $f(x) = 0$ şi*

$$|c_n - z| \leq \frac{b-a}{2^n}$$

Iteraţiile construite prin metoda bisecţiei se opresc la pasul $m$ pentru care

$$|f(c_m)| < \varepsilon$$

sau, pentru o mai mare acurateţe, la pasul $m$ pentru care

$$|f(c_m)| < \varepsilon \text{ şi } |c_m - c_{m-1}| < \varepsilon,$$

unde $\varepsilon$ este eroarea de aproximare.

**Exemplul -1.1** *Să se aproximeze, folosind metoda bisecţiei, soluţia ecuaţiei $x^3 + 4x^2 - 10 = 0$ (conţinută în intervalul $[-1, 2]$) cu eroarile $\varepsilon 1 = 0.01$, $\varepsilon 2 = 10^{-10}$.*

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon 1 = 0.01$ se obţine la pasul $m = 9$ şi are valoarea $c_{10} = 1.36132812500000$.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon = 10^{-10}$ se obţine la pasul $m = 35$ şi are valoarea $c_{37} = 1.36523001347086$.

## Regula falsi

"Regula falsi" este tot o metoda de "micsorare" a intervalului in care se afla solutia unei ecuatii.

Fie $a < b \in R$ şi $f : [a, b] \to R$ o funcţie continuă cu $f(a) \cdot f(b) < 0$. Atunci există $z \in [a, b]$ astfel încât $f(z) = 0$.

Definim şirurile $(a_n)_{n \geq 0}, (b_n)_{n \geq 0}, (c_n)_{n \geq 0}$ astfel:

- $a_0 := a$, $b_0 := b$, $c_0 := (a_0 f(b_0) - b_0 f(a_0))/(f(b_0) - f(a_0)$;
- Pentru $n \geq 1$
  - dacă $f(c_{n-1}) = 0$, atunci $\begin{cases} a_n := a_{n-1} \\ b_n := b_{n-1} \\ c_n := c_{n-1} \end{cases}$
  - dacă $f(a_{n-1}) \cdot f(c_{n-1}) < 0$, atunci $\begin{cases} a_n := a_{n-1} \\ b_n := c_{n-1} \\ c_n := (a_n f(b_n) - b_n f(a_n))/(f(b_n) - f(a_n) \end{cases}$
  - dacă $f(a_{n-1}) \cdot f(c_{n-1}) > 0$, atunci $\begin{cases} a_n := c_{n-1} \\ b_n := b_{n-1} \\ c_n := (a_n f(b_n) - b_n f(a_n))/(f(b_n) - f(a_n) \end{cases}$

Iteraţiile construite prin regula falsi se opresc la pasul $m$ pentru care

$$|f(c_m)| < \varepsilon$$

sau, pentru o mai mare acurateţe, la pasul $m$ pentru care

$$|f(c_m)| < \varepsilon \text{ şi } |c_m - c_{m-1}| < \varepsilon,$$

unde $\varepsilon$ este eroarea de aproximare.

**Teorema -1.2** *Presupunem că funcţia $f$ are o singură rădăcină în $[a, b]$. Atunci şirul $(c_n)_{n \geq 0}$ construit mai sus converge la unica soluţie $z \in [a, b]$ a ecuaţiei $f(x) = 0$.*

**Exemplul -1.2** *Să se aproximeze, folosind regula falsi, soluţia ecuaţiei $x^3 + 4x^2 - 10 = 0$ (conţinută în intervalul $[-1, 2]$) cu eroarile $\varepsilon 1 = 0.01$, $\varepsilon 2 = 10^{-10}$.*

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon 1 = 0.01$ se obţine la pasul $m = 8$ şi are valoarea $c_7 = 1.36506360624662$.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon = 10^{-10}$ se obţine la pasul $m = 21$ şi are valoarea $c_{21} = 1.36523001341145$.

## Metoda coardei

Fie $[a, b] \subset R$ şi $f : [a, b] \to R$ o funcţie. Definim recurent şirul $(x_n)_n$ astfel:

$$x_0 := a,\ \ x_1 := b \text{ şi}$$

$$x_{n+1} := \frac{x_0 f(x_n) - x_n f(x_0)}{f(x_n) - f(x_0)},\ \forall\ n \in \mathbf{N}$$

**Teorema -1.3** *Presupunem că $f$ este de două ori derivabilă pe $[a, b]$ şi*

$$a)\ f''(x) \neq 0,\ \forall\ x \in (a,b)$$

$$b)\ f(a) \cdot f''(a) > 0$$

$$c)\ f(a) \cdot f(b) < 0$$

*Atunci ecuaţia $f(x) = 0$, are o soluţie unică $z \in [a, b]$, iar şirul $(x_n)_n$ converge la $z$.*

Iteraţiile construite prin metoda coardei se opresc la pasul $m$ pentru care

$$|f(x_m)| < \varepsilon$$

sau, pentru o mai mare acurateţe, la pasul $m$ pentru care

$$|f(x_m)| < \varepsilon \text{ şi } |x_m - x_{m-1}| < \varepsilon,$$

unde $\varepsilon$ este eroarea de aproximare.

**Exemplul -1.3** *Să se aproximeze, folosind metoda coardei, soluţia ecuaţiei $x^3 - x + 3 = 0$ (conţinută în intervalul $[-3, 0]$) cu erorile $\varepsilon 1 = 0.01$, $\varepsilon 2 = 10^{-10}$.*

Se verifică ipotezele teoremei anterioare.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon 1 = 0.01$ se obţine la pasul $m = 10$ şi are valoarea $x_{15} = -1.66063905970749$.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon = 10^{-10}$ se obţine la pasul $m = 40$ şi are valoarea $x_{40} = -1.67169988158683$.

## Metoda secantei

Fie $[a, b] \subset R$ şi $f : [a, b] \to R$ o funcţie. Definim recurent şirul $(x_n)_n$ astfel:

$$x_0,\ x_1 \in [a, b] \text{ şi}$$

$$x_{n+1} := \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})},\ \forall\, n \in \mathbf{N}$$

**Teorema -1.4** *Presupunem că $f$ este derivabilă pe $[a, b]$ şi*

$$a)\ f'(x) \neq 0,\ \forall\, x \in [a,b]$$

*b) şirul $(x_n)_n$ are toate valorile în intervalul $[a, b]$*

$$c)\ f(a) \cdot f(b) < 0$$

*Atunci ecuaţia $f(x) = 0$, are o soluţie unică $z \in [a, b]$, iar şirul $(x_n)_n$ converge la $z$.*

Iteraţiile construite prin metoda secantei se opresc la pasul $m$ pentru care

$$|f(x_m)| < \varepsilon$$

sau, pentru o mai mare acurateţe, la pasul $m$ pentru care

$$|f(x_m)| < \varepsilon \text{ şi } |x_m - x_{m-1}| < \varepsilon,$$

unde $\varepsilon$ este eroarea de aproximare.

**Exemplul -1.4** *Să se aproximeze, folosind metoda secantei, soluţia ecuaţiei* $x^3 - x + 3 = 0$ *(conţinută în intervalul* $[-15, 15]$, *luând* $x_0 = 1$, $x_1 = 2$*) cu eroarile* $\varepsilon 1 = 0.01$, $\varepsilon 2 = 10^{-10}$.

Se verifică ipotezele teoremei anterioare. Punctul b) se verifică la fiecare pas.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon 1 = 0.01$ se obţine la pasul $m = 23$ şi are valoarea $x_{23} = -1.67109143768403$.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon = 10^{-10}$ (chiar $\varepsilon = 10^{-13}$) se obţine la pasul $m = 26$ şi are valoarea $x_{26} = -1.67169988165715$

## Metoda lui Newton

Fie $[a, b] \subset R$ şi $f : [a, b] \to R$ o funcţie.

**Teorema -1.5** *Presupunem că* $f$ *este de două ori derivabilă pe* $[a, b]$, *că* $f', f''$ *nu se anulează pe* $[a, b]$ *şi* $f(a) \cdot f(b) < 0$. *Fie* $x_0 \in [a, b]$ *astfel încât*

$$f(x_0) \cdot f''(x_0) > 0$$

*şi*

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \ \forall\, n \in \mathbf{N}$$

*Atunci ecuaţia* $f(x) = 0$, *are o soluţie unică* $z \in [a, b]$, *iar şirul* $(x_n)_n$ *converge la* $z$.

**Propoziţia -1.1** *Este adevărată următoarea formulă de evaluare a erorii:*

$$|x_n - z| \leq \frac{|f(x_n)|}{\min\limits_{x \in [a,b]} |f'(x)|}, \ \forall\, n \in \mathbf{N}$$

Iteraţiile construite prin metoda Newton se opresc la pasul $m$ pentru care

$$|f(x_m)| < \varepsilon$$

sau, pentru o mai mare acurateţe, la pasul $m$ pentru care

$$|f(x_m)| < \varepsilon \text{ şi } |x_m - x_{m-1}| < \varepsilon,$$

unde $\varepsilon$ este eroarea de aproximare (se poate folosi si evaluarea data de propozitia anterioara).

**Exemplul -1.5** *Să se aproximeze, folosind metoda secantei, soluţia ecuaţiei* $x^3 - x + 3 = 0$ *(conţinută în intervalul* $[-2, -1]$, *luând* $x_0 = -1$*) cu eroarile* $\varepsilon 1 = 0.01$, $\varepsilon 2 = 10^{-10}$.

Se verifică ipotezele teoremei anterioare.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon 1 = 0.01$ se obţine la pasul $m = 5$ şi are valoarea $x_5 = -1.67170038194364$.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon = 10^{-10}$ (chiar $\varepsilon = 10^{-11}$) se obţine la pasul $m = 7$ şi are valoarea $x_7 = -1.67169988165716$.

## Principiul contracţiei

Fie $\mathbf{I} \subset R$ un interval şi $f : \mathbf{I} \to R$ o funcţie. $f$ se numeşte contracţie dacă şi numai dacă:

$$a) \text{ există } q \in (0,1) \text{ astfel încât } |f(x) - f(y)| \le q\,|x - y|\,, \text{ pentru orice } x, y \in \mathbf{I}$$

$$b) \ f(\mathbf{I}) \subset \mathbf{I}$$

**Propoziţia -1.2** *Fie* $f : [a, b] \to R$ *o funcţie derivabilă.*

a) Dacă există $q \in (0, 1)$ astfel încât $|f\prime(x)| \le q$, pentru orice $x \in [a, b]$, atunci $|f(x) - f(y)| \le q\,|x - y|$, pentru orice $x, y \in [a, b]$.

b) Dacă $|f(x) - f(y)| \le q\,|x - y|$ şi $\left|f\left(\frac{a+b}{2}\right) - \frac{a+b}{2}\right| \le (1-q) \cdot \frac{b-a}{2}$, atunci $f([a, b]) \subset [a, b]$.

**Teorema -1.6** *Fie* $f : \mathbf{I} \to R$ *o contracţie şi* $x_0 \in \mathbf{I}$. *Definim şirul* $(x_n)_n$ *prin relaţia de recurenţă*

$$x_{n+1} := f(x_n), \ \forall \ n \in \mathbf{N}$$

*Atunci ecuaţia* $f(x) = x$ *are o soluţie unică* $z \in [a, b]$ , *iar şirul* $(x_n)_n$ *converge la* $z$, *cu următoarea formulă de evaluare a erorii:*

$$|x_n - z| \leq \frac{q}{1-q} |x_n - x_{n-1}| \leq \frac{q^n}{1-q} |x_1 - x_0| , \ \forall \ n \in \mathbf{N}.$$

Iteraţiile construite prin metoda coardei se opresc la pasul $m$ pentru care

$$\frac{q}{1-q} |x_m - x_{m-1}| < \varepsilon$$

sau la pasul $m$ pentru care

$$|f(x_m) - x_m| < \varepsilon \text{ şi } |x_m - x_{m-1}| < \varepsilon,$$

unde $\varepsilon$ este eroarea de aproximare.

**Exemplul -1.6** *Să se aproximeze, folosind principiul contracţiei, soluţia ecuaţiei*

$$x = \frac{8}{x+2}$$

*conţinută în intervalul* $[1, 4]$, *luând* $x_0 = 4$,*cu eroarile* $\varepsilon 1 = 0.01$, $\varepsilon 2 = 10^{-10}$.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon 1 = 0.01$ se obţine la pasul $m = 9$ şi are valoarea $x_9 = 1.99707174231332$.

Soluţia ecuaţiei, aproximată cu eroarea $\varepsilon = 10^{-10}$ se obţine la pasul $m = 36$ şi are valoarea $x_{36} = 2.00000000002183$.

# Metode pentru rezolvarea sau aproximarea soluţiilor sistemelor de ecuaţii liniare

## Norme de vectori şi norme de matrici

Fie $m \in N^*$. Pe spaţiul $R^m$ se consideră normele vectoriale uzuale $\|\cdot\|_1$, $\|\cdot\|_2$ şi $\|\cdot\|_\infty$ definite prin

$$\|x\|_1 = \sum_{i=1}^{m} |x_i| \,;$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^{m} |x_i|^2} \,;$$

$$\|x\|_\infty = \max_{1\le i\le m} |x_i| \,, \ \forall\, x = (x_i)_{1\le i\le m} \in \mathbf{R}^m.$$

Fie $M_m(R)$ spaţiul matricelor cu $m$ linii şi $m$ coloane cu elemente din $R$. Pe $M_m(R)$ se pot considera norme definite ca norme de operator liniar

$$\|A\|_{\alpha\beta} := \sup_{\|x\|_\alpha \le 1} \|Ax\|_\beta, \ \forall\, A \in \mathcal{M}_m(\mathbf{R})$$

unde $\|\cdot\|_\alpha$ şi $\|\cdot\|_\beta$ sunt norme pe $R^m$. În cazul în care $\alpha = \beta$ notăm $\|A\|_\alpha := \|A\|_{\alpha\alpha}$ şi o numim norma $\alpha$ a matricei $A$ (subordonată normei vectoriale $\|\cdot\|_\alpha$).

**Propoziţia -1.1** *Fie $A = (a_{ij})_{1\le i,j\le m} \in M_m(R)$. Atunci* $\|A\|_1 = \max\limits_{1\le j\le m} \sum\limits_{i=1}^{m} |a_{ij}|$.

**Propoziţia -1.2** *Fie $A = (a_{ij})_{1\le i,j\le m} \in M_m(R)$. Atunci* $\|A\|_\infty = \max\limits_{1\le i\le m} \sum\limits_{j=1}^{m} |a_{ij}|$.

**Propoziţia -1.3** *Fie $A \in M_m(R)$. Atunci $\|A\|_2 = \sqrt{\rho(A^tA)}$, unde $\rho(A^tA)$ reprezintă cea mai mare valoare proprie a matricei $A^tA$.*

**Exemplul -1.1** *Fie*

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & -5 & 6 \\ 7 & -8 & 9 \end{pmatrix}$$

*Atunci*

$$\|A\|_1 = \max_{1\le j\le 3} \sum_{i=1}^{3} |a_{ij}| =$$

$$= \max\left(|1| + |4| + |7|, |-2| + |-5| + |-8|, |3| + |6| + |9|\right) = \max(12, 15, 18) = 18$$

*şi*

$$\|A\|_\infty = \max_{1\le i\le 3} \sum_{j=1}^{3} |a_{ij}| =$$

$$= \max\left(|1| + |-2| + |3|, |4| + |-5| + |6|, |7| + |-8| + |9|\right) = \max(6, 15, 24) = 24.$$

## Condiţionarea unui sistem de ecuaţii liniare

Fenomenul de instabilitate manifestat în diverse procese matematice este important de studiat , dat fiind limitările tehnicii de calcul sau ale măsurătorilor de unde provin datele de calcul. Un exemplu cum este cel care urmează scoate în evidenţă existenţa unui astfel de fenomen şi motivaţia studiului său.

**Exemplul -1.2** *Fie sistemul de ecuaţii* $Ax = b$ *unde*

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \qquad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \qquad \textit{cu soluţia} \qquad x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

*Considerăm sistemul perturbat* $A(x + \delta x) = b + \delta b$ *unde*

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad b + \delta b = \begin{pmatrix} 32,1 \\ 22,9 \\ 33,1 \\ 30,9 \end{pmatrix}$$

$$\textit{cu soluţia } x+\delta x = \begin{pmatrix} 9,2 \\ -12,6 \\ 4,5 \\ -1,1 \end{pmatrix} \tag{1}$$

*Considerăm şi sistemul perturbat* $(A+\Delta A)(x+\Delta x)=b$ *unde*

$$A+\Delta A = \begin{pmatrix} 10 & 7 & 8,1 & 7,2 \\ 7,08 & 5,04 & 6 & 5 \\ 8 & 5,98 & 9,89 & 9 \\ 6,99 & 4,99 & 9 & 9,98 \end{pmatrix} \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 30 \end{pmatrix}$$

$$\textit{cu soluţia } x+\Delta x = \begin{pmatrix} -81 \\ 137 \\ -34 \\ 22 \end{pmatrix} \tag{2}$$

Sistemul (1) diferă de cel iniţial printr-o "mică" variaţie a coloanei termenilor liberi, iar sistemul (2) printr-o "mică" variaţie a elementelor matricei. După cum se observă aceste "mici" variaţii antrenează după sine variaţii "mari" ale soluţiei iniţiale.

**Exemplul -1.3** *Dacă considerăm sistemul perturbat* $(A+\Delta A)(x+\Delta x)=b+\Delta b$ *unde*

$$A+\Delta A = \begin{pmatrix} 10 & 7 & 8,1 & 7,2 \\ 7,08 & 5,04 & 6 & 5 \\ 8 & 5,98 & 9,89 & 9 \\ 6,99 & 4,99 & 9 & 9,98 \end{pmatrix} \quad b+\Delta b = \begin{pmatrix} 32,1 \\ 22,9 \\ 33,1 \\ 30,9 \end{pmatrix}$$

$$\textit{atunci acesta are soluţia } x+\Delta x = \begin{pmatrix} 332,90 \\ -550,49 \\ 143,88 \\ -84,58 \end{pmatrix} \tag{3}$$

În legătură cu această problemă de stabilitate a sistemelor de ecuaţii liniare sunt cunoscute următoarele rezultate:

**Definiţia -1.1** *Fie $A$ o matrice nesingulară şi $\| \ \|$ o normă a matricei $A$. Numărul $cond(A) := \|A\| \cdot \|A^{-1}\|$ se numeşte indice de condiţionare al matricei $A$ relativ la norma $\| \ \|$.*

**Propoziţia -1.4** *Fie sistemul de ecuaţii liniare $Ax = b$ şi sistemul perturbat $A(x + \delta x) = b + \delta b$. Presupunem $b \neq 0$. Atunci*

$$\frac{\|\delta x\|}{\|x\|} \leq cond(A) \cdot \frac{\|\delta b\|}{\|b\|} \,. \tag{4}$$

*şi*

$$\frac{\|\delta x\|}{\|x\|} \geq \frac{1}{cond(A)} \cdot \frac{\|\delta b\|}{\|b\|}$$

*În plus există $b \neq 0$ şi $\delta b \neq 0$ astfel încât inegalitatea (4) să devină egalitate.*

**Propoziţia -1.5** *Fie sistemul de ecuaţii liniare $Ax = b$ şi sistemul perturbat $(A + \Delta A)(x + \Delta x) = b$. Atunci*

$$\frac{\|\Delta x\|}{\|x + \Delta x\|} \leq cond(A) \cdot \frac{\|\Delta A\|}{\|A\|} \,. \tag{5}$$

*În plus există $b \neq 0$ şi $\Delta A \neq 0$ astfel încât inegalitatea (5) să devină egalitate.*

**Propoziţia -1.6** *Fie sistemul de ecuaţii liniare $Ax = b$ şi sistemul perturbat $(A + \Delta A)(x + \Delta x) = b + \Delta b$. Dacă $\|A^{-1}\Delta A\| < 1$, atunci*

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{cond(A)}{1 - \|A^{-1}\Delta A\|} \cdot \left( \frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right)$$

**Exemplul -1.4** *Vom considera cazul sistemului (1) şi vom lucra cu norma infinit.*

Avem:

$$\|\delta x\|_\infty = \|x + \delta x - x\|_\infty = 13,6$$

$$\|x\|_\infty = 1.$$

Atunci

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} = \frac{13,6}{1} = 13,6$$

$$A^{-1} = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}$$

$$\|A\|_\infty = 33, \ \left\|A^{-1}\right\|_\infty = 136.$$

Rezultă că

$$cond(A) := \|A\|_\infty \cdot \left\|A^{-1}\right\|_\infty = 33 \cdot 136 = 4488$$

Avem

$$\|\delta b\|_\infty = \|b + \delta b - b\|_\infty = 0,1.$$

$$\|b\|_\infty = 33..$$

Atunci

$$\frac{\|\delta b\|_\infty}{\|b\|_\infty} = \frac{0,1}{33} = \frac{1}{330}$$

şi

$$cond(A) \cdot \frac{\|\delta b\|_\infty}{\|b\|_\infty} = 4488 \cdot \frac{1}{330} = 13,6$$

Deci

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} = cond(A) \cdot \frac{\|\delta b\|_\infty}{\|b\|_\infty}$$

**Exemplul -1.5** *Fie sistemul de ecuaţii $Ax = b$ unde*

$$A = \begin{pmatrix} 10 & 1 & 4 & 0 \\ 1 & 10 & 5 & -1 \\ 4 & 5 & 10 & 7 \\ 0 & -1 & 7 & 9 \end{pmatrix} \qquad b = \begin{pmatrix} 15 \\ 15 \\ 26 \\ 15 \end{pmatrix} \qquad \text{cu soluţia} \quad x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

*Considerăm sistemul perturbat $A(x + \delta x) = b + \delta b$ unde*

$$A = \begin{pmatrix} 10 & 1 & 4 & 0 \\ 1 & 10 & 5 & -1 \\ 4 & 5 & 10 & 7 \\ 0 & -1 & 7 & 9 \end{pmatrix} \quad b + \delta b = \begin{pmatrix} 16 \\ 16 \\ 25 \\ 16 \end{pmatrix}$$

$$\text{cu soluţia } x + \delta x = \begin{pmatrix} 832 \\ 1324 \\ -2407 \\ 2021 \end{pmatrix}$$

**Exemplul -1.6** *Un alt exemplu de matrici cu un indice de condiţionare mare sunt matricele Hilbert $(H_n)$ de ordinul $n$, definite astfel:*

$$H_n = (h_{ij}^{(n)})_{i,j=\overline{1,n}} \text{ cu } h_{ij}^{(n)} = \frac{1}{i+j-1}, \ \forall\, i, j = \overline{1,n}$$

# Metode directe

## Metoda lui Gauss (cu pivotare parţială)

Se dau $m \in \mathbf{N}^*$ şi $A = (a_{ij}) \in M_{m,m+1}(\mathbf{R})$. Fie sistemul de ecuaţii liniare

$$\begin{cases} a_{11}x_1 + \cdots + a_{1m}x_m = a_{1,m+1} \\ \dots\dots\dots\dots\dots \\ a_{m1}x_1 + \cdots + a_{mm}x_m = a_{m,m+1} \end{cases}$$

cu necunoscutele $x_1, \dots, x_m \in \mathbf{R}$. Metoda lui Gauss este folosită atât pentru determinarea soluţiei sistemului, cât şi pentru calculul determinantului matricei sistemului..

Sistemul dat se transformă în $m-1$ etape astfel:

Iniţial se consideră determinantul matricei sistemului $\det = 1$

Pentru $n$ între 1 şi $m-1$ se efectuează următoarele:

- Se determină $max = |a_{sn}| = \max\limits_{n \le i \le m} |a_{in}|$ ($s \in \overline{n,m}$ reprezintă poziţia pe care s-a găsit maximul).
- Se ia $piv = a_{sn}$ (elementul $a_{sn}$ se numeşte pivot).
- Dacă $piv = 0$, atunci metoda nu se aplică.
- $det = det \cdot piv$.
- Dacă $s \neq n$, atunci $det = (-1) \cdot det$ şi se permută ecuaţia $s$ cu ecuaţia $n$.
- Coeficienţii ecuaţiei $n$ se împart la $piv$.
- Pentru $\forall i \in \overline{n+1,m}$ se elimină $x_n$ din ecuaţia $i$ astfel: $a_{ij} = a_{ij} - a_{in} \cdot a_{nj}$, $\forall\, j \in \overline{m,n}$.

Procedând astfel se obţine un sistem de forma:

$$\begin{cases} x_i + \sum\limits_{j=i+1}^{m} a_{ij} x_j = a_{i,m+1}, \quad 1 \le i \le m-1 \\ a_{mm} x_m = a_{m,m+1} \end{cases} \tag{6}$$

Determinantul matricei sistemului se calculează astfel: $det = det \cdot a_{mm}$.

Dacă $a_{mm} = 0$, atuci metoda nu se aplică.

Se rezolvă sistemul (6) astfel:

- $x_m = a_{m,m+1}/a_{m,m}$;

- $x_i = a_{i,m+1} - \sum\limits_{j=i+1}^{m} a_{ij}x_j,\ \forall\, i \in \overline{m-1, 1}.$

**Exemplul -1.7** *Fie sistemul de ecuaţii liniare*

$$\begin{cases} x + \frac{1}{2}y + \frac{1}{3}z &= 1 \\ \frac{1}{2}x + \frac{1}{3}y + \frac{1}{4}z &= 2 \\ \frac{1}{3}x + \frac{1}{4}y + \frac{1}{5}z &= 3 \end{cases}$$

*Să se calculeze determinantul matricei sistemului şi să se rezolve sistemul cu metoda lui Gauss.*

Folosind transformările date de metoda lui Gauss, obţinem următorul sistem de ecuaţii:

$$\begin{cases} x + \frac{1}{2}y + \frac{1}{3}z &= 1 \\ y + z &= 18 \\ \frac{1}{180}z &= \frac{7}{6} \end{cases}$$

Calculăm determinantul matricei sistemului (tot cu algoritmul dat de metoda lui Gauss) $\det = \frac{1}{2160}$ şi rezolvăm sistemul anterior. Obţinem $z = 210$, $y = -192$, $x = 27$.

**Observaţie:** Alegerea pivotului ca cel mai mare element de pe coloană se face pentru a minimiza erorile care apar dacă pivotul are valori mici. Dacă considerăm sistemul de ecuaţii liniare

$$\begin{cases} x + 592y = 437 \\ 592x + 4308y = 2251 \end{cases}$$

şi dacă considerăm că lucrăm numai cu 4 cifre exacte, atunci prin alegerea pivotului elementul 1,prin metoda lui Gauss se obţine soluţia

$$x = -1,6128,\ \ y = 0,7409$$

iar prin alegera pivotului elementul 592, prin metoda lui Gauss se obţine soluţia

$$x = -1,5891,\ \ y = 0,7409$$

pe când soluţia exactă a sistemului este

$$x = -1.58889055801431,\ y = 0.74085961242908.$$

## Metoda lui Gauss (cu pivotare totală)

Se dau $m \in \mathbf{N}^*$ şi $A = (a_{ij}) \in M_{m,m+1}(\mathbf{R})$. Fie sistemul de ecuaţii liniare

$$\begin{cases} a_{11}x_1 + \cdots + a_{1m}x_m = a_{1,m+1} \\ \dots\dots\dots\dots\dots \\ a_{m1}x_1 + \cdots + a_{mm}x_m = a_{m,m+1} \end{cases}$$

cu necunoscutele $x_1, \dots, x_m \in \mathbf{R}$. La metoda lui Gauss cu pivotare totală căutarea pivotului se face în toată matricea rămasă de transformat (reamintim că, la pivotarea parţială, pivotul se caută numai pe prima coloană a matricei rămase de transformat).

Sistemul dat se transformă în $m-1$ etape astfel:

Iniţial se consideră determinantul matricei sistemului $\det = 1$

Pentru $n$ între 1 şi $m-1$ se efectuează următoarele:

- Se determină $max = |a_{ps}| = \max\limits_{n \le i,j \le m} |a_{ij}|$ ($s, p \in \overline{n,m}$ reprezintă indicii de poziţie ai maximului determinat).
- Se ia $piv = a_{ps}$ (elementul $a_{ps}$ se numeşte pivot)
- Dacă $piv = 0$, atunci metoda nu se aplică.
- $det = det \cdot piv$.
- Dacă $p = n$ sau $s = n$ (dar nu $p = s = n$), atunci $det = (-1) \cdot det$ şi se permută coloana $s$ cu coloana $n$ (dacă $p = n$) sau linia $p$ cu linia $n$ (dacă $s = n$). Dacă $p \neq n$ şi $s \neq n$ se permută linia $p$ cu linia $n$ şi apoi coloana $s$ cu coloana $n$.
- Ecuaţia $n$ se împarte la $piv$.

- Pentru $\forall i \in \overline{n+1,m}$ se elimină $x_n$ din ecuaţia $i$ astfel: $a_{ij} = a_{ij} - a_{in} \cdot a_{nj}$, $\forall\, j \in \overline{m,n}$.

Procedând astfel se obţine un sistem de forma:

$$\begin{cases} x'_i + \sum\limits_{j=i+1}^{m} a_{ij}x'_j &= a_{i,m+1}, \quad 1 \le i \le m-1 \\ a_{mm}x'_m &= a_{m,m+1} \end{cases} \tag{7}$$

Determinantul matricei sistemului se calculează astfel: $det = det \cdot a_{mm}$.

Dacă $a_{mm} = 0$, atuci metoda nu se aplică.

Se rezolvă sistemul (7) astfel:

- $x'_m = a_{m,m+1}/a_{m,m}$;
- $x'_i = a_{i,m+1} - \sum\limits_{j=i+1}^{m} a_{ij}x'_j$, $\forall\, i \in \overline{m-1,1}$.

unde $(x'_1, \ldots, x'_m)$ este o permutare a soluţiei $(x_1, \ldots, x_m)$ sistemului iniţial (după cum s-au permutat coloanele matricei iniţiale în procesul de transformare a sistemului).

**Exemplul -1.8** *Folosind metoda lui Gauss cu pivotare totală, să se rezolve sistemul de ecuaţii liniare:*

$$\begin{cases} 5x + y + 2z = 29 \\ 3x - y + z = 10 \\ x + 2y + 4z = 31 \end{cases} \tag{8}$$

*Să se calculeze şi determinantul matricei sistemului.*

Folosind transformările date de metoda lui Gauss cu pivotare totală, obţinem următorul sistem de ecuaţii:

$$\begin{cases} x + \frac{2}{5}z + \frac{1}{5}y & = & \frac{29}{5} \\ \\ z + \frac{1}{2}y & = & 7 \\ \\ -\frac{3}{2}y & = & -6 \end{cases}$$

Calculăm determinantul matricei sistemului (tot cu algoritmul dat de metoda lui Gauss) det $= -27$ şi rezolvăm sistemul anterior. Obţinem $z = 5$, $y = 4$, $x = 3$.

## Metoda lui Gauss-Jordan de calcul a inversei unei matrice

Se dau $m \in N^*$ şi matricea

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mm} \end{pmatrix}$$

Metoda lui Gauss-Jordan este folosită pentru determinarea inversei matricei $A$ (dacă aceasta există). Se consideră ansamblul matriceal compus din matricea $A$ şi matricea unitate $I_m$

$$\left(\begin{array}{ccc|ccc} a_{11} & \dots & a_{1m} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mm} & 0 & \dots & 1 \end{array}\right)$$

Ansamblul anterior se transformă (analog metodei Gauss - cu pivotare parţială sau totală) astfel:

Pentru $n$ între 1 şi $m$ se efectuează următoarele:

- Se determină pivotul *piv* ca la metoda lui Gauss de rezolvare a unui sistem de ecuaţii liniare cu pivotare parţială (analog se poate descrie un algoritm folosind metoda lui Gauss cu pivotare totală).

- Dacă $piv = 0$, atunci metoda nu se aplică.
- Dacă $s \neq n$, atunci se permută ecuaţia $s$ cu ecuaţia $n$.
- Ecuaţia $n$ se împarte la $piv$.
- Pentru $\forall\ i \in \overline{n+1,m}$ se elimină $x_n$ din ecuaţia $i$ astfel: $a_{ij} = a_{ij} - a_{in} \cdot a_{nj}$, $\forall\ j \in \overline{2m,n}$.

Procedând astfel se obţine un ansamblu matricial de forma:

$$\left(\begin{array}{cccc|cccc} 1 & a_{12} & \dots & a_{1m} & a_{1,m+1} & a_{1,m+2} & \dots & a_{1,2m} \\ 0 & 1 & \dots & a_{2m} & a_{2,m+1} & a_{2,m+2} & \dots & a_{2,2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & a_{m,m+1} & a_{m,m+2} & \dots & a_{m,2m} \end{array}\right)$$

(matricea din partea dreaptă are cel puţin $n(n-1)/2$ zerouri).

Pentru $n$ între $m$ şi 2 se efectuează următoarele:

- Pentru $\forall\ i \in \overline{n-1,1}$ se recalculează elementele $a_{ij}$ astfel: $a_{ij} = a_{ij} - a_{in} \cdot a_{nj}$, $\forall\ j \in \overline{2m,n}$.

Precedând astfel se obţine un ansamblu matricial de forma:

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1m} \\ 0 & 1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & b_{m1} & b_{m2} & \dots & b_{mm} \end{array}\right).$$

Matricea

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{pmatrix}$$

are liniile matricei inverse $A^{-1}$ modulo permutarea liniilor descrisă în algoritm.

Rezolvarea unui sistem de ecuaţii liniare $Ax = b$ este echivalentă, după ce s-a calculat inversa matricei $A$, cu egalitatea $x = A^{-1}b$.

**Exemplul -1.9** *Folosind metoda Gauss-Jordan să se calculeze inversa matricei*

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$$

Formăm ansamblul matriceal

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & 0 & 0 & 1 \end{array}\right)$$

Cu transformările din algoritm obţinem:

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & 1 & 1 & -6 & 12 & 0 \\ 0 & 0 & 1 & 30 & -180 & 180 \end{array}\right)$$

iar, în final,

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 9 & -36 & 30 \\ 0 & 1 & 0 & -36 & 192 & -180 \\ 0 & 0 & 1 & 30 & -180 & 180 \end{array}\right)$$

Deci

$$A^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix}$$

## Factorizarea LU

Se numeşte factorizare $LU$ a unei matrice $A$ descompunerea matricei ca produs de două matrici, una inferior triunghiulară (notată $L$), alta superior triunghiulară (notată $U$), adică $A = LU$. Descompunerea, dacă este posibilă, nu este unică.

**Teorema -1.1** *Fie $A = (a_{ij})_{1\leq i,j\leq m}$ o matrice astfel încât determinanţii "de colţ" $\Delta_k := \det((a_{ij})_{1\leq i,j\leq k})$ să fie nenuli, pentru orice $k = \overline{1,m}$. Atunci $A$ se descompune unic sub forma $A = LU$ cu $L = (l_{ij})_{1\leq i,j\leq m}$ inferior triunghiulară şi $U = (u_{ij})_{1\leq i,j\leq m}$ superior triunghiulară cu elementele diagole egale cu* 1.

Calculul elementelor matricelor $L$ şi $U$ se face după formulele:

$$l_{i1} = a_{i1},\ i = \overline{1,m},\ u_{11} = 1,\ u_{1j} = \frac{a_{1j}}{l_{11}},\ j = \overline{2,m}. \tag{9}$$

iar, pentru $k = \overline{2,m}$,

$$l_{ik} = a_{ik} - \sum_{p=1}^{k-1} l_{ip}u_{pk},\ i = \overline{k,m},\ u_{kk} = 1,\quad u_{kj} = \frac{a_{kj} - \sum\limits_{p=1}^{k-1} l_{kp}u_{pj}}{l_{kk}},\ j = \overline{k+1,m}. \tag{10}$$

Fie un sistem de ecuaţii liniare $Ax = b$, pentru care $A$ admite factorizare $LU$. Atunci soluţia $y$ a sistemului $Ly = b$ se determină astfel:

$$y_i = \frac{b_i - \sum\limits_{k=1}^{i-1} l_{ik}y_k}{l_{ii}} \qquad \forall i = \overline{1,m}$$

Soluţia $x$ a sistemului iniţial se determină astfel:

$$x_i = y_i - \sum_{k=i+1}^{m} u_{ik}x_k \qquad \forall i = \overline{m,1}$$

**Exemplul -1.10** *Să se factorizeze sub forma $LU$ matricea*

$$A = \begin{pmatrix} 2 & 4 & -2 & 0 \\ -1 & -1 & 2 & 3 \\ 4 & 5 & -2 & -9 \\ 0 & 1 & 3 & 4 \end{pmatrix}$$

Se verifică mai întâi că determinanţii de colţ ai matricei $A$ sunt nenuli. Cu formulele (9) şi (10) se obţine:

$$l_{11} = 2,\ l_{21} = -1,\ l_{31} = 4,\ l_{41} = 0$$

$$u_{11} = 1,\ u_{12} = 2,\ u_{13} = -1,\ u_{14} = 0$$

$$l_{22} = 1,\ l_{32} = -3,\ l_{42} = 1$$

$$u_{22} = 1,\ u_{23} = 1,\ u_{24} = 3$$

$$l_{33} = 5,\ l_{43} = 2$$

$$u_{33} = 1,\ u_{34} = 0$$

$$l_{44} = 1$$

$$u_{44} = 1$$

Deci

$$L = \begin{pmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 4 & -3 & 5 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

## Metoda rădăcinii pătrate (Cholesky)

**Definiţia -1.2** *O matrice $A = (a_{ij})_{1 \le i,j \le m}$ se numeşte simetrică dacă $A = A^t$. Matricea $A$ este pozitiv definită dacă determinanţii "de colţ" $\Delta_k := \det((a_{ij})_{1 \le i,j \le k})$ sunt strict pozitivi pentru orice $k = \overline{1, m}$.*

**Teorema -1.2** *Fie $A$ simetrică şi pozitiv definită. Atunci $A$ se descompune unic sub forma $A = L \cdot L^t$ cu $L = (l_{ij})_{1 \le i,j \le m}$ inferior triunghiulară.*

Fie sistemul de ecuaţii liniare $Ax = b$, cu $A = (a_{ij}) \in M_m(\mathbb{R})$ simetrică şi pozitiv definită, iar $b = (b_i) \in \mathbb{R}^m$.

Descompunem $A = L \cdot L^t$ cu $L = (l_{ij}) \in M_n(\mathbb{R})$ matrice superior triunghiulară. Se determină mai întâi $y = (y_i) \in \mathbb{R}^n$ soluţia sistemului de ecuaţii $Ly = b$ şi apoi soluţia $x = (x_i) \in \mathbb{R}^n$ a sistemului iniţial se determină prin rezolvarea sistemului $L^t x = y$.

Se calculează elementele matricei $L$ astfel:

$$l_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2} \qquad \text{şi} \qquad l_{ij} = \frac{a_{ij} - \sum\limits_{k=1}^{j-1} l_{ik} l_{jk}}{l_{jj}}, \qquad \forall i = \overline{j+1, m}, \forall j = \overline{1, m} \tag{11}$$

Soluţia $y$ a sistemului $Ly = b$ se determină astfel:

$$y_i = \frac{b_i - \sum\limits_{k=1}^{i-1} l_{ik} y_k}{l_{ii}} \qquad \forall i = \overline{1, m} \tag{12}$$

Soluţia $x$ a sistemului iniţial se determină astfel:

$$x_i = \frac{y_i - \sum\limits_{k=i+1}^{m} l_{ki} x_k}{l_{ii}} \qquad \forall i = \overline{m, 1} \tag{13}$$

Dacă matricea $A$ nu este simetrică şi pozitiv definită, dar este inversabilă, atunci sistemul de ecuaţii liniare $Ax = b$ se transformă în sistemul echivalent $A^t A x = A^t b$, a cărui matrice este simetrică şi pozitiv definită.

**Exemplul -1.11** *Folosind metoda rădăcinii pătrate, să se rezolve sistemul de ecuaţii liniare $Ax = b$ cu*

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 10 & 4 \\ 2 & 4 & 6 \end{pmatrix}, b = \begin{pmatrix} 8 \\ 16 \\ 12 \end{pmatrix}$$

Se verifică imediat că matricea $A$ este simetrică şi pozitiv definită. Cu formulele (11) se obţine

$$l_{11} = 2, \; l_{12} = 1, \; l_{13} = 1$$

$$l_{22} = 3,\ l_{23} = 1$$

$$l_{33} = 2$$

Din (12) rezultă că

$$y_1 = 4,\ y_2 = 4,\ y_3 = 2$$

iar din (13) rezultă că

$$x_3 = 1,\ x_2 = 1,\ x_3 = 1.$$

## Descompunerea QR

Se consideră date $m \in \mathbf{N}^*$, $A = (a_{ij})_{i,j=\overline{1,m}} \in \mathcal{M}_m(\mathbf{R})$ şi $b = (b_i)_{i=\overline{1,m}} \in \mathbf{R}^m$. Vom calcula soluţia $x = (x_1, \ldots, x_m) \in \mathbf{R}^m$ a sistemului $Ax = b$ şi determinantul matricei sistemului folosind factorizarea $QR$.

Factorizarea $QR$ a matricei $A$ înseamnă descompunerea $A = QR$ cu $Q$ matrice ortogonală, adică $QQ^t = Q^tQ = \mathbf{I}_m$ şi $R$ matricea superior triunghiulară.

Dacă notăm $Q = (q_{ij})$ cu $\sum_{k=1}^{m} q_{ik}q_{jk} = \delta_{ij}$ şi $\sum_{k=1}^{m} q_{ki}q_{kj} = \delta_{ij}$ pentru $i, j = \overline{1,m}$ şi $R = (r_{ij})$ cu $r_{ij} = 0$ pentru $1 \leq j < i \leq m$, atunci din identificarea $A = QR$ se obţin relaţiile următoare:

$$\begin{cases} r_{11} = \sqrt{a_{11}^2 + \cdots + a_{m1}^2}\,, \\[2ex] q_{i1} = \dfrac{a_{i1}}{r_{11}}\,, \quad \text{pentru } i = \overline{1,m} \end{cases}$$

şi pentru $k = \overline{2, m}$ avem:

$$\begin{cases} r_{jk} = \sum_{i=1}^{m} a_{ik} q_{ij}\,, \quad \text{pentru } j = \overline{1, k-1} \\ \\ r_{kk} = \sqrt{\sum_{i=1}^{m} a_{ik}^2 - \sum_{i=1}^{k-1} r_{ik}^2}\,, \\ \\ q_{ik} = \dfrac{1}{r_{kk}} \left( a_{ik} - \sum_{j=1}^{k-1} r_{jk} q_{ij} \right)\,, \quad \text{pentru } i = \overline{1, m}\,. \end{cases} \tag{14}$$

Rezolvarea sistemului după descompunerea matricei $A$ în produs $QR$ se face astfel:

- rezolvăm întâi sistemul $Qy = b$ a cărui soluţie este

$$y = Q^t b,$$

sau pe componente:

$$y_i = \sum_{j=1}^{m} q_{ji} b_j\,, \quad i = \overline{1, m}.$$

- apoi, rezolvăm sistemul triunghiular $Rx = y$, şi obţinem:

$$\begin{aligned} x_m &= \frac{y_m}{r_{mm}}\,, \\ x_i &= \frac{1}{r_{ii}} \left( y_i - \sum_{j=i+1}^{m} r_{ij} x_j \right) \quad , i = \overline{m-1, 1}\,. \end{aligned} \tag{15}$$

**Observaţie**. Descompunerea $QR$ nu există dacă există un $k \in \overline{1, m}$ astfel încât $r_{kk} = 0$ în (**??**) sau (**??**).

**Exemplul -1.12** *Pentru* $m = 3$ *, matricea* $A = \begin{pmatrix} 0 & 4 & 5 \\ -1 & -2 & -3 \\ 0 & 0 & 1 \end{pmatrix}$ *şi vectorul* $b = \begin{pmatrix} 23 \\ -14 \\ 3 \end{pmatrix}$*, rezolvaţi sistemul* $Ax = b$ *folosind o descompunere* $QR$*.*

Folosind relaţiile de mai sus obţinem:

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{şi} \quad R = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 1 \end{pmatrix},$$

iar pentru soluţie

$$y = Q^t b = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 23 \\ -14 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 23 \\ 3 \end{pmatrix},$$

iar din $Rx = y$ rezultă:

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

# Metode iterative de rezolvare a sistemelor de ecuaţii liniare

## Metoda lui Jacobi

Fie $m \in N^*,\ A \in M_m(R),\ b \in R^m$.

Considerăm sistemul de ecuaţii liniare $Ax = b$. Notăm cu $I$ matricea unitate din $M_m(R)$ şi cu $B := I - A$. Sistemul de ecuaţii $Ax = b$ se transformă echivalent astfel:

$$Ax = b \Longleftrightarrow (I - B)x = b \Longleftrightarrow x = Bx + b.$$

Pentru orice $x^{(0)} \in R^m$ definim şirul (numit şir Jacobi) $(x^{(n)})_{n \in \mathbf{N}}$ prin relaţia de recurenţă

$$x^{(n+1)} = Bx^{(n)} + b,\ \forall\, n \in \mathbf{N}$$

Fie $\|\cdot\|$ o normă pe $M_m(R)$. Dacă $\|B\| = q < 1$ atunci avem evaluările:

$$\|x^{(n)} - x\| \leq \frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\| \leq \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|,\ \forall\, n \in \mathbf{N}^*$$

(în particular $\lim\limits_{n\to\infty} x^{(n)} = x$).

Algoritm: Dacă $A = (a_{ij})_{1\le i,j\le m}$, $b = (b_i)_{1\le i\le m}$, $B = (b_{ij})_{1\le i,j\le m}$, $x^{(n)} = (x_i^{(n)})_{1\le i\le m}$, $\forall\, n \in N$, atunci $b_{ij} = -a_{ij}$ dacă $i \neq j$, $b_{ii} = 1 - a_{ii}$, $\forall\, 1 \le i, j \le m$. Dacă pe $R^m$ considerăm norma $\|\cdot\|_1$, atunci $q := \|B\|_1 = \max\limits_{1\le j\le m} \sum\limits_{i=1}^{m} |b_{ij}|$, iar dacă considerăm norma $\|\cdot\|_\infty$ atunci $q := \|B\|_\infty = \max\limits_{1\le i\le m} \sum\limits_{j=1}^{m} |b_{ij}|$. Se testează condiţia de aplicabilitate a metodei $q < 1$. În caz de aplicabilitate se atribuie iteratei iniţiale $x^{(0)}$ o valoare oarecare din $R^m$, iar calculul celorlalte iterate se face după formula:

$$x_i^{(n+1)} = \sum_{j=1}^{m} b_{ij} x_j^{(n)} + b_i,\ \forall\, 1 \le i \le m,\ \forall\, n \in \mathbf{N}.$$

Se opreşte calculul recursiv la iterata $x^{(n)}$ pentru care

$$\frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\|_p \le \boldsymbol{\varepsilon},$$

unde $p = 1$ sau $p = \infty$, iar $\varepsilon$ este eroarea de aproximaţie dorită.

**Exemplul -1.13** *Să se arate că se poate aplica metoda lui Jacobi (relativă la normele 1 şi $\infty$) pentru sistemul de ecuaţii liniare $Ax = b$ cu*

$$A = \begin{pmatrix} 0,9 & 0,1 & 0,2 & 0,3 \\ -0,2 & 0,8 & 0,1 & 0,4 \\ 0,1 & -0,3 & 0,7 & -0,1 \\ -0,3 & -0,2 & -0,1 & 0,9 \end{pmatrix};\ b = \begin{pmatrix} 3,6 \\ 2 \\ 2,4 \\ 1,5 \end{pmatrix}.$$

*Luând* $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ *să se determine numărul de iteraţii necesar pentru a aproxima soluţia sistemului cu o eroare mai mică de* $10^{-10}$.

Avem

$$B = I - A = \begin{pmatrix} 0,1 & -0,1 & -0,2 & -0,3 \\ 0,2 & 0,2 & -0,1 & -0,4 \\ -0,1 & 0,3 & 0,3 & 0,1 \\ 0,3 & 0,2 & 0,1 & 0,1 \end{pmatrix}$$

Atunci

$$\|B\|_{\mathbf{1}} = \max_{1 \leq j \leq 4} \sum_{i=1}^{4} |b_{ij}| = 0,9 < 1.,$$

şi

$$\|B\|_{\infty} = \max_{1 \leq i \leq 4} \sum_{j=1}^{4} |b_{ij}| = 0,9 < 1$$

Deci metoda lui Jacobi se aplică. Din formula de evaluare a erorii avem:

$$\|x^{(n)} - x\|_p \leq \frac{q^n}{1-q} \|x^{(1)} - x^{(0)}\|_p, \ \forall n \in \mathbf{N}^*$$

unde $p = 1$ sau $p = \infty$, $q = 0,9$, $x^{(1)} = Bx^{(0)} + b = b$. Deci, pentru a aproxima $x$ cu $x^{(n)}$ cu eroarea $\varepsilon = 10^{-10}$ este suficient ca $\frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_p < \varepsilon$. Avem

$$\|x^{(1)} - x^{(0)}\|_1 = \|b\|_1 = 9,5$$

şi

$$\|x^{(1)} - x^{(0)}\|_\infty = \|b\|_\infty = 3,6$$

Atunci

$$\frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_1 < \varepsilon \Longleftrightarrow n = \left[\log_{0,9} \frac{\varepsilon}{95}\right] + 1$$

şi

$$\frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_\infty < \varepsilon \Longleftrightarrow n = \left[\log_{0,9} \frac{\varepsilon}{36}\right] + 1$$

## Metoda lui Jacobi pentru matrice diagonal dominante pe linii

Fie $m \in N^*, \ A \in M_m(R), \ a \in R^m$.

Considerăm sistemul de ecuaţii liniare $Ax = a$. Notăm cu $I$ matricea unitate din $M_m(R)$ şi cu

$$D = \mathrm{diag}(A) = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{mm} \end{pmatrix}.$$

Spunem că $A$ este diagonal dominantă pe linii dacă

$$|a_{ii}| > \sum_{j=\overline{1,m}\ j\neq i} |a_{ij}|\,,\ \forall i \in \overline{1,m}$$

Atunci $a_{ii} \neq 0$, $\forall i \in \overline{1,m}$, deci $D$ este inversabilă. Notăm cu $B = I - D^{-1}A$ şi cu $b = D^{-1}a$. Sistemul de ecuaţii $Ax = a$ se transformă echivalent astfel:

$$Ax = a \Longleftrightarrow D^{-1}Ax = D^{-1}a \Longleftrightarrow (I - B)x = b \Longleftrightarrow x = Bx + b.$$

Pentru orice $x^{(0)} \in R^m$ definimm şirul $(x^{(n)})_{n\in\mathbf{N}}$ prin relaţia de recurenţă

$$x^{(n+1)} = Bx^{(n)} + b,\ \forall n \in \mathbf{N}.$$

**Teorema -1.3** *Fie $A$ o matrice diagonal dominantă pe linii, $B = I - D^{-1}A$, $q := \|B\|_\infty$ şi $(x^{(n)})_{n\in\mathbf{N}}$ definit ca mai sus. Atunci $q < 1$ şi avem evaluările:*

$$\|x^{(n)} - x\|_\infty \leq \frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\|_\infty \leq \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_\infty,\ \forall n \in \mathbf{N}^*$$

*(în particular $\lim\limits_{n\to\infty} x^{(n)} = x$).*

## Metoda lui Jacobi pentru matrice diagonal dominante pe coloane

Spunem că matricea $A$ este diagonal dominantă pe coloane dacă

$$|a_{jj}| > \sum_{i=\overline{1,m}\ i\neq j} |a_{ij}|\,,\ \forall j \in \overline{1,m}$$

Atunci $a_{ii} \neq 0$, $\forall i \in \overline{1,m}$, deci $D$ este inversabilă. Notăm cu $y = Dx$ (deci $x = D^{-1}y$), şi cu $B = I - AD^{-1}$. Sistemul de ecuaţii $Ax = a$ se transformă echivalent astfel:

$$Ax = a \Longleftrightarrow AD^{-1}y = a \Longleftrightarrow (I - B)y = a \Longleftrightarrow y = By + a.$$

Pentru orice $y^{(0)} \in R^m$ definimm şirul $(y^{(n)})_{n\in\mathbf{N}}$ prin relaţia de recurenţă

$$y^{(n+1)} = By^{(n)} + a, \ \forall\, n \in \mathbf{N}.$$

**Teorema -1.4** *Fie $A$ o matrice diagonal dominantă pe coloane, $B = I - AD^{-1}$, $q := \|B\|_{\mathbf{1}}$ şi $(y^{(n)})_{n\in\mathbf{N}}$ definit ca mai sus. Atunci $q < 1$ şi avem evaluările:*

$$\|y^{(n)} - y\|_{\mathbf{1}} \leq \frac{q}{1-q}\|y^{(n)} - y^{(n-1)}\|_1 \leq \frac{q^n}{1-q}\|y^{(1)} - y^{(0)}\|_{\mathbf{1}}, \ \forall\, n \in \mathbf{N}^*$$

*(în particular $\lim\limits_{n\to\infty} y^{(n)} = y$). Deci soluţia $x$ a sistemului $Ax = a$ este aproximată de $D^{-1}y^{(n)}$ $(\lim\limits_{n\to\infty} D^{-1}y^{(n)} = x)$.*

## Metoda Gauss-Seidel

Fie $m \in N^*,\ A \in M_m(R),\ b \in R^m$.

Considerăm sistemul de ecuaţii liniare $Ax = b$. Notăm cu $I$ matricea unitate din $M_m(R)$ şi cu $B := I - A$. Dacă elementele matricei $B$ sunt $(b_{ij})_{1\leq i,j\leq m}$, descompunem $B$ sub forma $B = L + R$, cu $L = (l_{ij})_{1\leq i,j\leq m}$, unde $l_{ij} := b_{ij}$, $\forall\, 1 \leq j < i \leq m$ şi $l_{ij} := 0$ altfel. Atunci $\det(I - L) = 1$ şi deci există $(I - L)^{-1}$.

Sistemul de ecuaţii $Ax = b$ se transformă echivalent astfel:

$$Ax = b \Longleftrightarrow (I - B)x = b \Longleftrightarrow (I - L - R)x = b \Longleftrightarrow$$

$$\Longleftrightarrow (I - L)^{-1}(I - L - R)x = (I - L)^{-1}b \Longleftrightarrow (I - (I - L)^{-1}R)x = (I - L)^{-1}b \Longleftrightarrow$$

$$\Longleftrightarrow x - (I - L)^{-1}Rx = (I - L)^{-1}b.$$

Dacă notăm $C := (I - L)^{-1}R$ şi $c := (I - L)^{-1}b$ atunci

$$Ax = b \Longleftrightarrow x = Cx + c.$$

Pentru orice $x^{(0)} \in C^m$ definim şirul Jacobi $(x^{(n)})_{n\in\mathbf{N}}$ prin relaţia de recurenţă

$$x^{(n+1)} = Cx^{(n)} + c \Longleftrightarrow x^{(n+1)} = Lx^{(n+1)} + Rx^{(n)} + b,\ \forall\, n \in \mathbf{N}$$

Deci

$$x_1^{(n+1)} = \sum_{j=1}^{m} b_{1j} x_j^{(n)} + b_1,\ \forall\, n \in \mathbf{N} \text{ şi}$$

$$x_i^{(n+1)} = \sum_{j=1}^{i-1} b_{ij} x_j^{(n+1)} + \sum_{j=i}^{m} b_{ij} x_j^{(n)} + b_i,\ \forall\, 2 \le i \le m,\ \forall\, n \in \mathbf{N}.$$

**Teorema -1.5** *Dacă* $q := \max\limits_{1\le i\le m} q_i$ *cu* $q_1 = \sum\limits_{j=1}^{m} |b_{1j}|$, $q_i = \sum\limits_{j=1}^{i-1} |b_{ij}| q_j + \sum\limits_{j=i}^{m} |b_{ij}|$, $2 \le i \le m$ *şi* $q < 1$ *atunci avem evaluările:*

$$\|x^{(n)} - x\|_\infty \le \frac{q}{1-q} \|x^{(n)} - x^{(n-1)}\|_\infty \le \frac{q^n}{1-q} \|x^{(1)} - x^{(0)}\|_\infty,\ \forall\, n \in \mathbf{N}^*$$

*(în particular* $\lim\limits_{n\to\infty} x^{(n)} = x$*).*

**Exemplul -1.14** *Fie sistemul de ecuaţii liniare*

$$\begin{cases} x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 = 1 \\ \frac{1}{5}x_1 + x_2 + \frac{1}{6}x_3 = 2 \\ \frac{1}{10}x_1 + \frac{1}{20}x_2 + x_3 = 3 \end{cases}$$

*Să se arate că se aplică metoda Gauss-Seidel.*

Soluţie Notăm cu $A$ matricea ceoficienţilor sistemului. Avem

$$B = I - A = \begin{pmatrix} 0 & -\frac{1}{2} & -\frac{1}{3} \\ -\frac{1}{5} & 0 & -\frac{1}{6} \\ -\frac{1}{10} & -\frac{1}{20} & 0 \end{pmatrix}$$

şi

$$q_1 = \frac{5}{6},\ q_2 = \frac{1}{3},\ q_3 = \frac{1}{10}.$$

Deci

$$q = \max(q_1, q_2, q_3) = \frac{5}{6} < 1$$

Rezultă că metoda Gauss-Seidel se aplică.

## Metoda relaxării simultane

Fie $m \in N^*$, $A \in M_m(R)$ simetrică şi pozitiv definită, $a \in R^m$.

Considerăm sistemul de ecuaţii liniare $Ax = a$. Notăm cu $I$ matricea unitate din $M_m(R)$ şi cu $D = diag(A)$. Atunci $D$ inversabilă. Fie $\sigma > 0$ un număr real. Notăm cu $C_\sigma := I - \sigma D^{-1}A$ şi cu $c_\sigma := \sigma D^{-1}a$. Sistemul de ecuaţii $Ax = a$ se transformă echivalent astfel:

$$Ax = a \Longleftrightarrow \sigma D^{-1}Ax = \sigma D^{-1}a \Longleftrightarrow (I - C_\sigma)x = c_\sigma \Longleftrightarrow x = C_\sigma x + c_\sigma.$$

Pentru orice $x^{(0)} \in R^m$ definim şirul Jacobi $(x^{(n)})_{n\in\mathbf{N}}$ prin relaţia de recurenţă

$$x^{(n+1)} = C_\sigma x^{(n)} + c_\sigma,\ \forall\, n \in \mathbf{N}$$

Fie $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$ valorile proprii ale matricei $D^{-1}A$. Dacă alegem $\sigma$ astfel încât $0 < \sigma < 2/\lambda_1$, şi notăm $q := \max\limits_{1\leq i\leq m} |1 - \sigma\lambda_i|$, atunci $q < 1$ şi avem evaluările:

$$\|x^{(n)} - x\|_D \leq \frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\|_D \leq \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_D,\ \forall\, n \in \mathbf{N}^*$$

(în particular $\lim\limits_{n\to\infty} x^{(n)} = x$), unde $\|x\|_D := \sqrt{\langle Dx, x\rangle}, \forall\, x \in R^m$, cu $\langle\cdot,\cdot\rangle$ produsul scalar uzual pe $R^m$.

Parametrul optim de relaxare $\sigma$ (pentru care $q$ corespunzător este minim) este

$$\sigma = \frac{2}{\lambda_1 + \lambda_m},$$

iar valoarea lui $q$ în acest caz este

$$q = \frac{\lambda_1 - \lambda_m}{\lambda_1 + \lambda_m}$$

În practică, de obicei, nu se calculează valoarea lui $\lambda_1$, ci se alege $\sigma \in (0, 2/\|D^{-1}A\|_\infty)$ (deoarece $\lambda_1 \leq \|D^{-1}A\|_\infty$), parcurgând intervalul cu un pas echidistant. Valoarea optimă a lui $\sigma$ se stabileşte atunci când numărul de iteraţii $n$, necesar pentru a aproxima soluţia sistemului de ecuaţii cu o eroare dată este minim.

Dacă matricea $A$ a sistemului de ecuaţii liniare $Ax = a$ nu este simetriică şi pozitiv definită, sistemul se transformă echivalent

$$Ax = a \Longleftrightarrow A^t Ax = A^t a$$

(unde $A^t = (a_{ji})_{1 \leq i,j \leq m}$ este transpusa matricei $A$), pentru care matricea $A^t A$ este simetrică şi pozitiv definită.

**Exemplul -1.15** *Fie sistemul de ecuaţii liniare:*

$$\begin{cases} 5x_1 + 3x_2 + 2x_3 = 1 \\ 3x_1 + 6x_2 + 3x_3 = 2 \\ 2x_1 + 3x_2 + 5x_3 = 3 \end{cases}$$

*Să se arate că se poate aplica metoda relaxării simultane. Să se determine parametrul optim de relaxare. Luând $x^{(0)} = (0, 0, 0)$ să se evalueze eroarea $x - x^{(n)}$.*

Soluţie: Notăm cu $A$ matricea sistemului. Se verifică că $A$ este simetrică şi pozitiv definită.

Fie $D = diag(A)$. Atunci

$$D^{-1}A = \begin{pmatrix} 1 & 3/5 & 2/5 \\ 1/2 & 1 & 1/2 \\ 2/5 & 3/5 & 1 \end{pmatrix}$$

şi $\det(D^{-1}A - \lambda I) = 0 \Longleftrightarrow (1 - \lambda)^3 - \frac{19}{25}(1 - \lambda) + \frac{6}{25} = 0$, de unde rezultă că valorile proprii ale matricei $D^{-1}A$ sunt $\lambda_1 = 2 > \lambda_2 = 3/5 > \lambda_3 = 2/5$. Parametrul optim de relaxare este

$$\sigma = \frac{2}{\lambda_1 + \lambda_3} = \frac{5}{6}$$

căruia îi corespunde

$$q = \frac{\lambda_1 - \lambda_3}{\lambda_1 + \lambda_3} = \frac{2}{3}.$$

Avem $x^{(1)} = C_\sigma x^{(0)} + c_\sigma = c_\sigma = \sigma D^{-1} b = 5/6 \cdot (1/5; 2/6; 3/5)^t$ unde $b = (1; 2; 3)^t$ este vectorul termenilor liberi din sistem. Atunci

$$\|x^{(1)}\|_D^2 = \left\langle Dx^{(1)}, x^{(1)} \right\rangle = \frac{2}{3} \cdot \frac{25}{9}$$

şi

$$\|x^{(n)} - x\|_D \le \frac{q^n}{1-q} \|x^{(1)} - x^{(0)}\|_D = 5 \cdot \left(\frac{2}{3}\right)^{n+\frac{1}{2}}.$$

## Metode relaxării succesive

Fie $m \in N^*$, $A \in M_m(R)$ simetrică şi pozitiv definită şi $b \in R^m$.

Considerăm sistemul de ecuaţii liniare $Ax = b$ şi $\sigma > 0$ un număr real numit parametru de relaxare. Notăm cu $I$ matricea unitate din $M_m(R)$ şi cu $D = diag(A)$. Atunci $D$ inversabilă. Dacă elementele matricei $A$ sunt $(a_{ij})_{1 \le i,j \le m}$, descompunem $A$ sub forma $B = L + D + R$, cu $L = (l_{ij})_{1 \le i,j \le m}$, unde $l_{ij} := a_{ij}$, $\forall\, 1 \le j < i \le m$ şi $l_{ij} := 0$ altfel. Atunci matricea $\sigma^{-1} D + L$ este inversabilă,iar sistemul de ecuaţii $Ax = b$ se transformă echivalent astfel:

$$Ax = b \Leftrightarrow (\sigma^{-1} D + L)^{-1}(L + D + R)x = (\sigma^{-1} D + L)^{-1} b \Leftrightarrow$$

$$(I - (\sigma^{-1} D + L)^{-1}((\sigma^{-1} - 1)D - R))x = (\sigma^{-1} D + L)^{-1} b$$

Notând cu $C_\sigma := (\sigma^{-1} D + L)^{-1}((\sigma^{-1} - 1)D - R)$ şi cu $c_\sigma := (\sigma^{-1} D + L)^{-1} b$ rezultă că

$$Ax = b \Leftrightarrow x = C_\sigma x + c_\sigma.$$

Pentru orice $x^{(0)} \in R^m$ definim şirul Jacobi $(x^{(n)})_{n \in \mathbf{N}}$ prin relaţia de recurenţă

$$x^{(n+1)} = C_\sigma x^{(n)} + c_\sigma,\ \forall\, n \in \mathbf{N}$$

Atunci

$$x^{(n+1)} = C_\sigma x^{(n)} + c_\sigma \Leftrightarrow$$

$$\Leftrightarrow x^{(n+1)} = (\sigma^{-1}D + L)^{-1}((\sigma^{-1} - 1)D - R)x^{(n)} + b) \Leftrightarrow$$

$$\Leftrightarrow \sigma^{-1}Dx^{(n+1)} = -Lx^{(n+1)} + ((\sigma^{-1} - 1) - R)x^{(n)} + b \Leftrightarrow$$

$$\Leftrightarrow x^{(n+1)} = (1-\sigma)x^{(n)} - \sigma D^{-1}(Lx^{(n+1)} + Rx^{(n)} - b) \Leftrightarrow$$

$$x_i^{(n+1)} = (1-\sigma)x_i^{(n)} - \frac{\sigma}{a_{ii}}(-b_i + \sum_{j=1}^{i-1} a_{ij}x_j^{(n+1)} + \sum_{j=i+1}^{m} a_{ij}x_j^{(n)}),\ \forall 1 \le i \le m$$

Dacă alegem $\sigma$ astfel încât $0 < \sigma < 2$, şi notăm $q := \|C_\sigma\|_A$ atunci $q < 1$ şi avem evaluările:

$$\|x^{(n)} - x\|_A \le \frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\|_A \le \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_A,\ \forall\, n \in \mathbf{N}^*$$

(în particular $\lim\limits_{n\to\infty} x^{(n)} = x$), unde $\|x\|_A := \langle Ax, x\rangle, \forall\, x \in R^m$, cu $\langle\cdot,\cdot\rangle$ produsul scalar uzual pe $\mathbf{R}^m$.

În practică, se alege $\sigma \in (0,2)$ parcurgând intervalul cu un pas echidistant. Valoarea optimă a lui $\sigma$ se stabileşte atunci când numărul de iteraţii $n$, necesar pentru a aproxima soluţia sistemului de ecuaţii cu o eroare dată, este minim.

Dacă matricea $A$ a sistemului de ecuaţii liniare $Ax = b$ nu este simetrică şi pozitiv definită, sistemul se transformă echivalent

$$Ax = b \Leftrightarrow A^tAx = A^tb$$

(unde $A^t$ este transpusa matricei $A$) pentru care matricea $A^tA$ este simetrică şi pozitiv definită.

## Metoda lui Ritz

Este o metoda de calcul a inversei unei matrice simetrice şi poyitiv definite. Fie $A \in M_m(\mathbf{R})$ o matrice simetrică şi pozitiv definită. Notăm cu $^*$ operaţia de transpunere (a unui vector sau a unei matrice). Definim succesiv vectorii $(w_k)_{k=0}^m \subset R^m$, $(v_k)_{k=1}^m \subset R^m$ şi matricele $(C_k)_{k=1}^m, (D_k)_{k=1}^m \subset M_m(\mathbf{R})$ astfel:

$$w_0 \neq 0,\ C_1 = \frac{w_0 w_0^*}{w_0^* A w_0},\ D_1 = I_m - C_1 A$$

iar pentru $k = \overline{1, m-1}$:

$$\text{Se alege } v_k \text{ astfel ca } D_k v_k \neq 0,$$

$$w_k = D_k v_k.$$

$$C_{k+1} := C_k + \frac{w_k w_k^*}{w_k^* A w_k},$$

$$D_{k+1} := I_m - C_{k+1} A,$$

Atunci $C_m = A^{-1}$ (inversa matricei $A$) şi $D_m = 0_m$.

**Exemplul -1.16** *Fie* $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 1 & 10 \\ 3 & 1 & 35 & 5 \\ 4 & 10 & 5 & 30 \end{pmatrix}$. *Folosind metoda lui Ritz să se calculeze* $A^{-1}$.

Se verifică mai întâi că matricea $A$ este simetrică şi pozitiv definita. Luăm $w_0 = (1,0,0,0)^*$. Atunci $w_0^* A w_0 = 1$ şi

$$C_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ D_1 = \begin{pmatrix} 0 & -2 & -3 & -4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Luăm $v_1 = (0,1,0,0)^*$. Rezultă că $w_1 = (-2,1,0,0)^*$ şi $w_1^* A w_1 = 1$, de unde

$$C_2 = \begin{pmatrix} 5 & -2 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ D_2 = \begin{pmatrix} 0 & 0 & -13 & 0 \\ 0 & 0 & 5 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Luăm $v_2 = (0,0,1,0)^*$. Rezultă că $w_2 = (-13,5,1,0)^*$ şi $w_2^* A w_2 = 1$, de unde

$$C_3 = \begin{pmatrix} 174 & -67 & -13 & 0 \\ -67 & 26 & 5 & 0 \\ -13 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ D_3 = \begin{pmatrix} 0 & 0 & 0 & 39 \\ 0 & 0 & 0 & -17 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Luăm $v_3 = (0,0,0,1)^*$. Rezultă că $w_3 = (39,-17,-3,1)^*$ şi $w_3^* A w_3 = 1$, de unde

$$C_4 = \begin{pmatrix} 1695 & -730 & -130 & 39 \\ -730 & 315 & 56 & -17 \\ -130 & 56 & 10 & -3 \\ 39 & -17 & -3 & 1 \end{pmatrix}, \ D_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

## Pseudoinversa unei matrice

Pseudoinversa unei matrice este o extensie naturală a noţiunii de matrice inversabilă. Se poate defini pseudoinversa atât pentru matrici pătratice (singulare sau nesingulare - caz în care pseudoinversa coincide cu inversa) cât şi pentru matrici dreptunghiulre (în care numărul de linii diferă de numărul de coloane). Pentru o matrice $T \in M_{m,n}(R)$,notăm cu $T^* \in M_{n,m}(R)$ transpusa matricei $T$ şi cu $T^+ \in M_{n,m}(R)$ pseudoinversa matricei matricei $T$.

Prin definiţie $T^+$ este singura soluţie $U \in M_{n,m}(R)$ a sistemului de ecuaţii matriceale:

$$TUT = T, \quad UTU = U, \qquad (TU)^* = TU, \quad (UT)^* = UT. \tag{16}$$

**Teorema -1.6** *Fie $T \in M_{m,n}(R)$ şi $b \in R^m$.*

*a) $x_0$ din $R^n$ verifică egalitatea*

$$\|Tx_0 - b\| = \inf_{x \in \mathbf{R}^n} \|Tx - b\| \tag{17}$$

*dacă şi numai dacă $T^*Tx_0 = T^*b$.*

*b) Elementul $\widetilde{x_0} := T^+b$ verifică egalitatea (17) şi, pentru orice alt element $x_0$ ce verifică aceeaşi egalitate, avem $\|\widetilde{x_0}\| \leq \|x_0\|$.*

*Reciproc, $T^+$ este unica soluţie $U \in M_{n,m}(R)$a problemei: $Ub$ satisface egalitatea (17) pentru orice $b \in R^m$, iar pentru orice alt element $x_0$, care verifică aceeaşi egalitate, avem $\|Ub\| \leq \|x_0\|$.*

Prin această teoremă, în cazul unui sistem de ecuaţii liniare $Tx = b$ compatibil nedeterminat, rezultă că $T^+b$ este o soluţie a acestui sistem (anume cea mai mică soluţie în normă), iar, în cazul unui sistem incompatibil, $T^+b$ este acel element $x$, minim în normă, pentru care $Tx - b$ este cel mai aproape de 0.

În continuare presupunem că rang $T = l \leq \min(m, n)$. Cu $R(A)$ am notăm imaginea matricei $A \in M_{m,n}(R)$, adică mulţimea $\{Ax \mid x \in R^n\}$.

Determinarea pseudoinversei unei matrice se poate face cu algoritmi cu un număr de paşi mai mic sau egal cu $l$. În această categorie intră următorii algoritmi

- (Grigore) Definim succesiv vectorii $(y_k)_{k=1}^l \subset X$, $(w_k)_{k=1}^l \subset R^n$ şi matricele $(H_k)_{k=0}^l \subset M_{n,m}(R)$ astfel:

$$H_0 = \mathbf{0}.$$

$$y_k \in R(T^*) \text{ (adică în imaginea transpusei), astfel ca } y_k - H_{k-1}Ty_k \neq 0,$$

$$w_k := y_k - H_{k-1}Ty_k \text{ şi}$$
$$H_k := H_{k-1} + \frac{w_k(Tw_k)^*}{\|Tw_k\|^2}, \ \forall k = \overline{1, l}.$$

Atunci $H_l = T^+$ (pseudoinversa matricei $T$).

**Exemplul -1.17** *Fie*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \quad \text{cu } rang(T) = 2 \text{ şi } H_0 = \mathbf{0}.$$

Luăm $y_1 = T^* * \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ (avem $y_1 - H_0 T y_1 = y_1 \neq 0$. Atunci $w_1 = y_1 - H_0 T y_1 = y_1$ şi $T w_1 = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 2 \end{pmatrix}$. Rezultă că $\|Tw_1\|^2 = 12$, iar $w_1 (Tw_1)^* = \begin{pmatrix} 2 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ -2 & 0 & -2 & -2 \end{pmatrix}$. Deci

$$H_1 = H_0 + \frac{w_1 (Tw_1)^*}{\|Tw_1\|^2} = \begin{pmatrix} 1/6 & 0 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 \\ -1/6 & 0 & -1/6 & -1/6 \end{pmatrix}.$$

Luăm $y_2 = T^* * \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ (avem $y_2 - H_1 T y_2 = \begin{pmatrix} -1/2 \\ 1 \\ 1/2 \end{pmatrix} \neq 0$. Atunci $w_2 = y_2 - H_1 T y_2$ şi $T w_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$. Rezultă că $\|Tw_2\|^2 = 3$, iar $w_2 (Tw_2)^* =$

$$\begin{pmatrix} 1/2 & -1/2 & 0 & -1/2 \\ -1 & 1 & 0 & 1 \\ -1/2 & 1/2 & 0 & 1/2 \end{pmatrix}. \text{ Deci}$$

$$H_2 = H_1 + \frac{w_2(Tw_2)^*}{\|Tw_2\|^2} = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix}$$

În concluzie

$$T^+ = H_2 = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix}$$

- (Gramm-Schmidt) Fie o bază $e_1, e_2, \ldots, e_l$ în $R(T^*)$ şi fie $f_i = Te_i$, $\forall\ i = \overline{1,l}$. Definim, succesiv, elementele $(e_i')_{i=1}^l, (x_i)_{i=1}^l \subset R^n$, $(f_i')_{i=1}^l (y_i)_{i=1}^l \subset R^m$, astfel:

$$x_1 = \frac{1}{\|f_1\|} e_1 \text{ şi } y_1 = Tx_1,$$

  iar pentru $i = \overline{2,l}$

$$e_i' = e_i - \sum_{j=1}^{i-1} \langle f_i, y_j \rangle x_j,\ f_i' = Te_i' \text{ şi } x_i = \frac{1}{\|f_i'\|} e_i',\ y_i = Tx_i$$

Definim matricea $U \in \mathcal{M}_{n,m}(R)$ prin

$$U := \sum_{i=1}^{l} x_i y_i^*$$

Atunci $U = T^+$.

**Exemplul -1.18** *Fie*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \quad \textit{cu} \ \dim R(T^*) = 2$$

*Luând* $e_1 = (1;0;-1)^*, e_2 = (0;1;0)^*$ *(bază în* $R(T^*)$*) obţinem* $f_1 = (2;0;2;2)^*, f_2 = (0;1;1;2)^*$. *Atunci*

$$x_1 = \frac{1}{2\sqrt{3}}(1;0;-1)^*,\ y_1 = \frac{1}{\sqrt{3}}(1;0;1;1)^*$$

*În continuare*

$$e_2' = e_2 - \langle f_2, y_1\rangle\, x_1 = \frac{1}{2}(-1;2;1)^* \text{ şi } f_2' = Te_2' = (-1;1;0;1)^*$$

*Atunci*

$$x_2 = \frac{1}{2\sqrt{3}}(-1;2;1)^*,\ y_2 = \frac{1}{\sqrt{3}}(-1;1;0;1)^*$$

*Rezultă că*

$$T^+ = x_1y_1^* + x_2y_2^* = \frac{1}{6}\begin{pmatrix} 2 & -1 & 1 & 0 \\ -2 & 2 & 0 & 2 \\ -2 & 1 & -1 & 0 \end{pmatrix}$$

- (Stanica) Definim, succesiv, vectorii $(u_k)_{k=0}^{l-1} \subset R^m$, $(v_k)_{k=0}^{l-1} \subset R^n$, $(w_k)_{k=0}^{l-1} \subset R^m$ şi matricele $(H_k)_{k=1}^{l} \subset M_{n,m}(R)$ astfel:

$$H_0 = T^*.$$

$$v_k \in R((I - TH_k)^*T),\ w_k \in R(H_k) \text{ astfel încît } \langle v_k, Tw_k\rangle \neq 0,$$
$$u_k := \frac{1}{\langle v_k, Tw_k\rangle}(I - H_kT)w_k \text{ şi}$$
$$H_{k+1} := H_k + u_k \cdot v_k^*,\ \forall k = \overline{0, l-1}.$$

Atunci $H_l = T^+$.

**Exemplul -1.19** *Fie*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \quad \textit{cu rang}(T) = 2 \textit{ şi } H_0 = T^*$$

Luând $a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ obţinem $v_0 := (I - TH_0)^* T a_1 = \begin{pmatrix} -5 \\ -3 \\ -8 \\ -11 \end{pmatrix}$,

$w_0 = H_0 b_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $u_0 := \dfrac{1}{\langle v_0, Tw_0 \rangle}(I - H_0 T)w_0 = \begin{pmatrix} 5/48 \\ 1/8 \\ -5/48 \end{pmatrix}$,

$$H_1 := H_0 + u_0 \cdot v_0^* = \begin{pmatrix} 23/48 & -5/16 & 1/6 & -7/48 \\ -5/8 & 5/8 & 0 & 5/8 \\ -23/48 & 5/16 & 1/6 & 7/48 \end{pmatrix}$$

Luând $a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ obţinem $v_1 := (I - TH_1)^* T a_2 = \begin{pmatrix} 7/8 \\ -7/8 \\ 0 \\ -7/8 \end{pmatrix}$, $w_1 =$

$H_1 b_2 = \begin{pmatrix} -5/16 \\ 5/8 \\ 5/16 \end{pmatrix}$, $u_1 := \dfrac{1}{\langle v_1, Tw_1 \rangle}(I - H_1 T)w_1 = \begin{pmatrix} -1/6 \\ 1/3 \\ 1/6 \end{pmatrix}$, deci

$$T^+ = H_2 := H_1 + u_1 \cdot v_1^* = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix}.$$

**Exemplul -1.20** *Fie sistemul de ecuaţii liniare $Tx = b$, unde*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \quad \textit{cu rang}(T) = 2 \textit{ şi } b = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}$$

*(acest sistem este compatibil nedetreminat). Atunci*

$$z = T^{+}b = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

*este o soluţie a sistemului (compatibil nedeterminat) de ecuaţii liniare $Tx = b$.*

**Exemplul -1.21** *Fie sistemul de ecuaţii liniare $Tx = b$, unde*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \text{ cu } rang(T) = 2 \text{ şi } b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

*(acest sistem este incompatibil). Atunci*

$$z = T^{+}b = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -1/3 \\ -1/3 \end{pmatrix}$$

*are proprietatea că*

$$\|Tz - b\| = \inf_{x \in \mathbf{R}^n} \|Tx - b\|$$

.

# Metode pentru aproximarea valorilor proprii ale unei matrice

În acest capitol prezentăm algoritmi pentru determinarea valorilor şi vectorilor proprii ai unei matrice $A \in \mathcal{M}_m(R)$, adică a scalarilor $\lambda \in C$ şi a vectorilor $x \in C^m \setminus \{0\}$ asfel ca:

$$Ax = \lambda x \,.$$

Valorile proprii $\lambda$ sunt rădăcinile polinomului caracteristic:

$$p_A(\lambda) = \det(\lambda I_m - A) \,.$$

Pentru o valoare proprie $\lambda$ dată, vectorul propriu se obţine ca o soluţie nenulă a sistemului liniar omogen:

$$(A - \lambda I_m)x = 0 \,.$$

Metodele numerice pentru determinarea valorilor proprii se împart în doua mari categorii:

- metode care utilizează polinomul caracteristic;
- metode care nu folosesc polinomul caracteristic.

Prima categorie de metode este mai rar folosită pentru că trebuie să calculăm rădăcinile unui polinom de grad $m$ arbitrar. Metodele care nu folosesc polinomul caracteristic aproximează, în general, prin iterare, valorile proprii ale unei matrice.

## Metode folosind polinomul caracteristic

Metodele numerice pe care le prezentăm în acest paragraf determină, de fapt, coeficienţii polinomului caracteristic. Dacă $A \in \mathcal{M}_m(R)$, atunci polinomul caracteristic va fi:

$$p_A(\lambda) = \lambda^m - c_1\lambda^{m-1} + c_2\lambda^{m-2} + \cdots + (-1)^m c_m \,,$$

unde $c_k$ este suma minorilor de ordin $k$ ai matricei $A$ pentru $k=\overline{1,m}$. Numărul minorilor de ordin $k$ ai matricei $A$ este $C_m^k = \dfrac{m!}{k!(m-k)!}$. Deci, calculând direct coeficienţii $c_k$, $k=\overline{1,m}$, ar trebui calculaţi $C_m^1 + C_m^2 + \cdots + C_m^m = 2^m - 1$ determinanţi. Acest lucru este dificil şi implică multe calcule pentru valori mari ale lui $m$. De aceea, s-au căutat metode care să determine mai uşor coeficienţii polinomului caracteristic. În subparagrafele următoare prezentăm câteva astfel de metode împreună cu algoritmii corespunzători şi exemple numerice.

## Metoda Krylov

Metoda Krylov se bazează pe teorema Cayley-Hamilton:

Dacă $A \in \mathcal{M}_m(R)$ şi $p_A(\lambda) = \lambda^m - p_1\lambda^{m-1} - p_2\lambda^{m-2} - \cdots - p_m$ , atunci $p_A(A) = 0_m$ , adică:

$$A^m - p_1 A^{m-1} - p_2 A^{m-2} - \cdots - p_{m-1}A - p_m I_m = 0_m \,.$$

Considerăm $y^{(0)} = \left(y_1^{(0)}, \ldots, y_m^{(0)}\right)^t \in R^m$ un vector arbitrar. Înmulţind relaţia (**??**) cu $y^{(0)}$, se obţine:

$$A^m y^{(0)} - p_1 A^{m-1} y^{(0)} - p_2 A^{m-2} y^{(0)} - \cdots - p_{m-1} A y^{(0)} - p_m y^{(0)} = 0 \,,$$

de unde, notând:

$$y^{(k)} = A^k y^{(0)} = A y^{(k-1)}, \quad \text{pentru} \quad k = \overline{1,m},$$

ne rezultă că $p_1, \ldots, p_m$ sunt soluţiile sistemului:

$$\begin{pmatrix} y^{(m-1)} & y^{(m-2)} & \ldots & y^{(1)} & y^{(0)} \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{m-1} \\ p_m \end{pmatrix} = y^{(m)} \,.$$

Deci determinarea soluţiilor polinomului caracteristic prin metoda Krylov se reduce la rezolvarea sistemului (**??**), în cazul în care are soluţie unică. Dacă sistemul nu are soluţie unică, atunci se poate încerca schimbarea vectorului iniţial $y^{(0)}$.

**Exemplul -1.1** *Determinaţi polinomul caracteristic pentru matricea*

$$A = \begin{pmatrix} 5 & -4 & -3 & -1 \\ 1 & 4 & 2 & -9 \\ 2 & 5 & 1 & -5 \\ 7 & 1 & 2 & -1 \end{pmatrix},$$

*folosind metoda Krylov.*

Alegem pentru aplicarea metodei Krylov vectorul $y^{(0)} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ şi calculăm

$$y^{(1)} = Ay^{(0)} = \begin{pmatrix} -16 \\ -21 \\ -5 \\ 11 \end{pmatrix}, \quad y^{(2)} = Ay^{(1)} = \begin{pmatrix} 8 \\ -209 \\ -197 \\ -154 \end{pmatrix},$$

$$y^{(3)} = Ay^{(2)} - \begin{pmatrix} 1621 \\ 164 \\ -456 \\ -393 \end{pmatrix}, \quad y^{(4)} = Ay^{(3)} = \begin{pmatrix} 9210 \\ 4902 \\ 5571 \\ 10992 \end{pmatrix}.$$

Coefienţii polinomului caracteristic se obţin rezolvâd sistemul:

$$\begin{pmatrix} 1621 & 8 & -16 & 1 \\ 164 & -209 & -21 & 2 \\ -456 & -197 & -5 & 3 \\ -393 & -154 & 11 & 4 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \begin{pmatrix} 9210 \\ 4902 \\ 5571 \\ 10992 \end{pmatrix}.$$

Folosind o metodă de rezolvare a sistemelor liniare , obţinem $p = \begin{pmatrix} 9 \\ -45 \\ 369 \\ 885 \end{pmatrix}$ şi polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 + 9\lambda^3 - 45\lambda^2 + 369\lambda + 885\,. \quad \square$$

## Metoda Leverrier

Determinarea coeficienţilor polinomului caracteristic prin metoda Leverrier se bazează pe următorul rezultat:

Dacă $P(x) = x^m + p_1x^{m-1} + p_2x^{m-2} + \cdots + p_m$ este un polinom de gradul $m$ cu rădacinile $\lambda_1, \ldots, \lambda_m$ şi $s_k = \lambda_1^k + \cdots + \lambda_m^k$ pentru $k \in N$ , atunci avem

$$-kp_k = \sum_{j=0}^{k-1} p_j s_{k-j} \quad \text{pentru} \quad k \in N^*\,,$$

cu $p_0 = 1$ .

Se ştie că, dacă $A \in \mathcal{M}_m(R)$ şi $\lambda_1, \ldots, \lambda_m$ sunt valori proprii distincte ale lui $A$, atunci $s_k = tr(A^k)$ pentru $k \in N^*$ .

Determinarea coeficienţilor polinomului carateristic cu metoda Leverrier constă în:

P1. se calculează $s_1 = tr(A)$ şi $p_1 = -s_1$ ;

P2. pentru $k = \overline{2, m}$ calculăm:

$$A^k\,, \quad s_k = tr(A^k)\,, \quad p_k = -\frac{1}{k}\,(s_k + p_1 s_{k-1} + p_2 s_{k-2} + \cdots p_{k-1} s_1).$$

P3. Polinomul caracteristic al matricei $A$ va fi

$$p_A(\lambda) = \lambda^m + p_1\lambda^{m-1} + p_2\lambda^{m-2} + \cdots + p_m\,.$$

**Observaţie**. Dacă $p_m = 0$, atunci $\lambda = 0$ este valoare proprie pentru $A$ şi nu există $A^{-1}$, iar dacă $p_m \neq 0$ atunci, conform teoremei Cayley-Hamilton, avem:

$$A^m + p_1A^{m-1} + p_2A^{m-2} - \cdots + p_{m-1}A + p_mI_m = 0_m\,,$$

de unde rezultă

$$A^{-1} = -\frac{1}{p_m}\left(A^{m-1} + p_1A^{m-2} + p_2A^{m-3} + \cdots + p_{m-1}I_m\right)\,.$$

Există un algoritm propus de Fadeev prin care se calculează simultan coeficienţii polinomului caracteristic şi $A^{-1}$ (dacă există). Astfel determinarea coeficienţilor polinomului carateristic cu metoda Leverrier constă în:

P1. se calculează $p_1 = -tr(A)$ şi $B_1 = A_1 + p_1I_m$, unde $A_1 = A$;

P2. pentru $k = \overline{2,m}$ calculăm:

$$A_k = AB_k\,,\quad p_k = -\frac{1}{k}tr(A_k)\,,\quad B_k = A_k + p_kI_m.$$

P3. Dacă $p_m \neq 0$ atunci avem $A^{-1} = -\dfrac{1}{p_m}B_{m-1}$.

**Observaţie**. În algoritmul de mai sus, din modul de calculare a şirului $B_1, \ldots, B_{m-1}, B_m$ şi din teorema Cayley-Hamilton, se obţine că

$$B_m = A_m + p_mI_m = A^m + p_1A^{m-1} + \cdots + p_{m-1}A + p_mI_m = 0_m$$

**Exemplul -1.2** *Determinaţi, folosind metoda Leverrier, polinomul caracteristic şi matricea* $A^{-1}$, *dacă există, pentru matricea*

$$A = \begin{pmatrix} 11 & 0 & -3 & 12 \\ -9 & -5 & 2 & 5 \\ 2 & -4 & 15 & -11 \\ 2 & 6 & -7 & 10 \end{pmatrix},$$

Aplicând metoda Leverrier pentru matricea$A$ şi $m = 4$, obţinem:

$$s_1 = 31\,, \quad p_1 = -31\,,$$

- pentru $k = 2$:

$$A_2 = A^2 = \begin{pmatrix} 139 & 84 & -162 & 285 \\ -40 & 47 & 12 & -105 \\ 66 & -106 & 288 & -271 \\ -26 & 58 & -169 & 231 \end{pmatrix}, \quad s_2 = 705\,, \quad p_2 = 128\,,$$

- pentru $k = 3$:

$$A_3 = A^3 = \begin{pmatrix} 1019 & 1938 & -4674 & 6720 \\ -1049 & -913 & 1129 & -1427 \\ 7714 & -2248 & 5807 & -5616 \\ -684 & 1772 & -3958 & 4147 \end{pmatrix},$$

$$s_3 = 10060\,, \quad p_3 = 2609\,,$$

- pentru $k = 4$:

$$A_4 = A^4 = \begin{pmatrix} -2141 & 49326 & -116331 & 140532 \\ -3918 & -8513 & 28245 & -43842 \\ 39468 & -45684 & 116779 & -110709 \\ -23094 & 31854 & -82803 & 85660 \end{pmatrix},$$

$$s_4 = 191785\,, \quad p_4 = -12761\,,$$

şi polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 - 31\lambda^3 + 128\lambda^2 + 2609\lambda - 12761\,.$$

Cum $p_4 \neq 0$, rezultă că:

$$\begin{aligned} A^{-1} &= -\tfrac{1}{p_4}\left(A^3 + p_1 A^2 + p_2 A + p_3 I_4\right) \\ &= \begin{pmatrix} 0.0570 & -0.0522 & -0.0028 & -0.0454 \\ -0.0753 & -0.0314 & 0.0794 & 0.1934 \\ -0.0060 & 0.0412 & 0.1103 & 0.1079 \\ 0.0296 & 0.0581 & 0.0302 & 0.0686 \end{pmatrix}. \square \end{aligned}$$

**Exemplul -1.3** *Calculaţi coeficienţii polinomului caracteristic şi matricea* $A^{-1}$, *dacă există, pentru matricea*

$$A = \begin{pmatrix} 11 & 1 & -3 & 2 \\ -9 & -5 & 12 & 5 \\ 2 & -4 & 5 & -11 \\ 2 & 16 & -7 & 10 \end{pmatrix},$$

*folosind algoritmul Fadeev.*

Aplicând algoritmul Fadeev obţinem:

$$A_1 = A\,,$$

$$p_1 = -tr(A_1) = -21\,, \quad B_1 = A_1 + p_1 I_4 = \begin{pmatrix} -10 & 1 & -3 & 2 \\ -9 & -26 & 12 & 5 \\ 2 & -4 & -16 & -11 \\ 2 & 16 & -7 & -11 \end{pmatrix},$$

- pentru $k = 2$

$$A_2 = AB_1 = \begin{pmatrix} -121 & 29 & 13 & 38 \\ 169 & 153 & -260 & -230 \\ 4 & -90 & -57 & 50 \\ -158 & -226 & 228 & 51 \end{pmatrix},$$

$$p_2 = -\frac{1}{2}tr(A_2) = -13\,, \quad B_2 = A_2 + p_2 I_4 = \begin{pmatrix} -134 & 29 & 13 & 38 \\ 169 & 140 & -260 & -230 \\ 4 & -90 & -70 & 50 \\ -158 & -226 & 228 & 38 \end{pmatrix},$$

- pentru $k = 3$

$$A_3 = AB_2 = \begin{pmatrix} -1633 & 277 & 549 & 114 \\ -381 & -3171 & 1483 & 1598 \\ 814 & 1534 & -1792 & 828 \\ 828 & 668 & -1364 & -3574 \end{pmatrix},$$

$$p_3 = -\frac{1}{3}tr(A_3) = 3390\,, \quad B_3 = A_3 + p_3 I_4 = \begin{pmatrix} 1757 & 277 & 549 & 114 \\ -381 & 219 & 1483 & 1598 \\ 814 & 1534 & 1598 & 828 \\ 828 & 668 & -1364 & -184 \end{pmatrix},$$

- pentru $k=4$

$$A_4 = AB_3 = \begin{pmatrix} 18160 & 0 & 0 & 0 \\ 0 & 18160 & 0 & 0 \\ 0 & 0 & 18160 & 0 \\ 0 & 0 & 0 & 18160 \end{pmatrix},$$

$$p_4 = -\tfrac{1}{4}tr(A_4) = -18160\,, \quad B_4 = A_4 + p_4 I_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 - 21\lambda^3 - 13\lambda^2 + 3390\lambda - 18160\,.$$

Cum $p_4 \neq 0$, rezultă că:

$$A^{-1} = -\frac{1}{p_4}B_3 = \begin{pmatrix} 0.0968 & 0.0153 & 0.0302 & 0.0063 \\ -0.0210 & 0.0121 & 0.0817 & 0.0880 \\ 0.0448 & 0.0845 & 0.0880 & 0.0456 \\ 0.0456 & 0.0368 & -0.0751 & -0.0101 \end{pmatrix}. \quad \square$$

## Metoda coeficienţilor nedeterminaţi

Metoda coeficienţilor nedeterminati constă în determinarea coeficienţilor $p_1, \ldots, p_m$ ai polinomului caracteristic $p_A(\lambda) = \lambda^m + p_1\lambda^{m-1} + \cdots + p_{m-1}\lambda + p_m$ , asociat matriei

$A \in \mathcal{M}_m(R)$ , dintr-un sistem liniar format din $m$ ecuaţii liniare:

$$\begin{array}{rcl} p_m & = & D(0) \\ 1^m + p_1 1^{m-1} + \cdots + p_{m-1} 1 + p_m & = & D(1) \\ 2^m + p_1 2^{m-1} + \cdots + p_{m-1} 2 + p_m & = & D(2)) \\ & \vdots & \\ (m-1)^m + p_1 (m-1)^{m-1} + \cdots + p_{m-1}(m-1) + p_m & = & D(m-1) \end{array} \quad ,$$

unde $D(k) = \det(kI_m - A)$ pentru $k = \overline{0, m}$ .

Dacă notăm $p = (p_1, p_2, \ldots, p_{m-1})^t$ , sistemul (**??**) se reduce la $m-1$ ecuaţii liniare şi se poate scrie în forma:

$$Mp = b,$$

unde

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2^{m-1} & 2^{m-2} & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ (m-1)^{m-1} & (m-1)^{m-2} & \cdots & (m-1) \end{pmatrix}$$

şi

$$\begin{pmatrix} D(1) - D(0) - 1^m \\ D(2) - D(0) - 2^m \\ \vdots \\ D(m-1) - D(0) - (m-1)^m \end{pmatrix} .$$

Această metodă are dezavatajul că se calculează $m$ determinanţi de ordin $m$.

**Exemplul -1.4** *Determinaţi coeficienţii polinomului caracteristic pentru matricea*

$$A = \begin{pmatrix} 9 & -4 & 10 & 13 \\ 11 & -1 & -15 & 8 \\ 1 & -2 & -3 & 1 \\ 2 & 4 & 5 & 7 \end{pmatrix},$$

*folosind metoda coeficienţilor nedeterminaţi.*

Avem $p_4 = -2703$. Se obţine sistemul

$$\begin{pmatrix} 1 & 1 & 1 \\ 8 & 4 & 2 \\ 27 & 9 & 3 \end{pmatrix} p = \begin{pmatrix} 334 \\ 482 \\ 372 \end{pmatrix}$$

a cărui soluţie este:

$$p = \begin{pmatrix} -12 \\ -57 \\ 403 \end{pmatrix}.$$

Deci polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 - 12\lambda^3 - 57\lambda^2 + 403\lambda - 2703\,. \quad \square$$

## Metoda rotaţiilor

Se dau $m \in N,\ m \geq 2,\ A = (a_{ij})_{1\leq i,j\leq m} \in M_m(R)$ o matrice simetrică şi $\varepsilon > 0$ un număr real. Metoda rotaţiilor aproximează valorile proprii ale unei matrice simetrice prin construirea unui şir de matrice (obţinut cu ajutorul unor matrice de rotaţie) ale căror valori de pe diagonală converg către valorile proprii ale matricei $A$.

Dacă la pasul $n$ este cunoscută matricea $A_n$, matricea $A_{n+1}$ se determină cu formula

$$A_{n+1} = T_n^t A_n T_n,$$

unde $T_n$ este o matrice de rotaţie.

Calcul succesiv al matricelor $A_n$ $(n \geq 0)$, obţinute din matricea $A$ prin rotaţii plane se realizează după algoritmul:

$A_0 := A$

- Repetă (se calculează $Y = A_{n+1}$ în funcţie de $X = A_n$):

– se aleg $p < q \in \{1, \ldots, m\}$ astfel încât

$$|x_{pq}| = \max\{|x_{ij}|; i < j \in \{1, \ldots, m\}\};$$

– dacă $x_{pp} = x_{qq}$ atunci $\theta = \frac{\pi}{4}$, altfel $\theta = \frac{1}{2} \cdot \arctan \frac{2 \cdot x_{pq}}{x_{pp} - x_{qq}}$

– se iau $c = \cos(\theta)$ şi $s = \sin(\theta)$;

– Se calculează matricea $Y = T^t A T$ cu

$$T = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cos\theta & \cdots & -\sin\theta & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \sin\theta & \cdots & \cos\theta & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix},$$

unde elementele $cos\theta$, $-\sin\theta$, $\sin\theta$, $cos\theta$ se află la intersecţia liniilor $p$ şi $q$ cu coloanele $p$ şi $q$. Atunci

∗ $\forall\, i, j \notin \{p, q\},\ y_{ij} = x_{ij}$;

∗ $\forall\, j \notin \{p, q\},\ y_{pj} = y_{jp} = c \cdot x_{pj} + s \cdot x_{qj},\ y_{qj} = y_{jq} = -s \cdot x_{pj} + c \cdot x_{qj}$

∗ $y_{pq} = 0$; $y_{qp} = 0$; $y_{pp} = c^2 \cdot x_{pp} + 2 \cdot c \cdot s \cdot x_{pq} + s^2 \cdot x_{qq}$; $y_{qq} = s^2 \cdot x_{pp} - 2 \cdot c \cdot s \cdot x_{pq} + c^2 \cdot x_{qq}$;

– se face atribuirea $X = Y$;

– se calculează $modul = |X| := \sqrt{\sum_{i \neq j}^{m} x_{ij}^2}$ unde $i, j \in \overline{1, m}$.

pâna când $modul < \varepsilon$ .

- Se colectează în vectorul $z$ elementele diagonalei lui $X$, care reprezintă aproximări ale valorilor proprii ale matricei $A$ cu erori în modul mai mici decât $\varepsilon$ : $\forall\, i \in \overline{1, m}$ se ia $z_i = x_{ii}$.

- Se tipăreşte $z$.

**Exemplul -1.5** *Fie matricea*

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

*Folosind metoda rotaţiilor, să se determine valorile proprii ale matricei $A$.*

Se verifică imediat că matricea $A$ este simetrică. Se aleg $p < q$ astfel încât

$$|a_{pq}| = \max\{a_{ij}|; i < j \in \{1,2,3\}.$$

Atunci $p = 1$, $q = 3$. Se calculează

$$\theta = \frac{1}{2} \cdot \arctan \frac{2 \cdot a_{pq}}{a_{pp} - a_{qq}} = \frac{\pi}{6}.$$

Se ia

$$T = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$

şi se calculează matricea

$$B = T^t A T = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 8 & 4 \\ 0 & 4 & 8 \end{pmatrix}$$

Se aleg $p < q$ astfel încât

$$|b_{pq}| = \max\{b_{ij}|; i < j \in \{1,2,3\}.$$

Atunci $p = 2$, $q = 3$. Se calculează

$$\theta = \frac{\pi}{4} \text{ deoarece } b_{pp} = b_{qq}.$$

Se ia

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

şi se calculează matricea

$$B = T^t A T == \begin{pmatrix} 20 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Rezultă că valorile proprii ale matricei $A$ sunt $\lambda_1 = 20$, $\lambda_2 = 12$, $\lambda_3 = 4$. $\square$

## Metoda puterii

Este o metodă numerică pentru aproximarea celei mai mari valori proprii în modul şi a unui vector propriu corespunzător.

Vom considera cazul în care matricea $A \in \mathcal{M}_m(\mathbf{R})$ are toate valorile proprii pozitive. Fie $x^{(0)} \in \mathbf{R}^m$ astfel încât $Ax^{(0)} \neq 0$. Definim recurent şirul $(x^{(n)})_{n\in\mathbf{N}}$ prin relaţia $x^{(n+1)} := Ax^{(n)}$, $\forall n \in \mathbf{N}$. Notăm cu $(z^{(n)})_{n\in\mathbf{N}}$ şirul definit prin $z^{(n)} := \dfrac{1}{\|x^{(n)}\|} \cdot x^{(n)}$, $\forall n \in \mathbf{N}$. Atunci

$$\lim_{n\to\infty} \langle Az^{(n)}, z^{(n)} \rangle = \lambda,$$

unde $\lambda$ este cea mai mare valoare proprie a matricei $A$ şi

$$\lim_{n\to\infty} z^{(n)} = z,$$

unde $z$ este un vector propriu ataşat valorii proprii $\lambda$.

Metoda permite şi determinarea celorlalte valori proprii astfel:

- Se consideră matricea

$$T = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ x_2 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ x_m & 0 & \ldots & 1 \end{pmatrix},$$

unde $(1, x_2, \ldots, x_m)$ este un vector propriu corespunzător valorii proprii $\lambda$. Acest vector se obţine astfel:

$$x_i = \frac{z_i}{z_1},\ \forall i = \overline{2, m},$$

- Se calculează matricea $B = T^{-1}AT$.
- Se scrie matricea $B$ sub forma

$$B = \begin{pmatrix} b_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix},$$

cu $B_{22}$ matrice de ordinul $m-1$.

- Cu algoritmul descris mai sus se determină cea mai mare valoare proprie a matricei $C = B_{22}$. Aceasta este a doua valoare proprie (în ordine descrescătoare) a matricei $A$.
- Acest algoritm se aplică în continuare pentru matricea $C$.

### Algoritmul LU

Fie $A$ o matrice pătratică de ordin $m$. Descompunem $A = L_1U_1$ cu $L_1, U_1$ matrice triunghiulare (vezi factorizarea $LU$ din paragraful 1.2.4, capitolul 1). Construim şirul de matrice $(B_n)_n$ astfel:

$$B_0 = A = L_1U_1,\ B_1 = U_1L_1.$$

Descompunem $B_1 = L_2U_2$, luăm $B_2 = U_2L_2, \ldots$

Descompunem $B_n = L_{n+1}U_{n+1}$, luăm $B_{n+1} = U_{n+1}L_{n+1}, \ldots$

Când $n \to \infty$, elementele de pe diagonala lui $B_n$ converg la valorile proprii ale matricei $A$ (în practică, calculul iterativ se opreşte atunci când, pentru orice $i = \overline{1, m}$, avem $\left|b_{ii}^{(n)} - b_{ii}^{(n-1)}\right| < \varepsilon$, o eroare dată, unde cu $(b_{ij}^{(n)})_{i,j}$ am notat elementele matricei $B_n$).

## Algoritmul QR

Metoda QR este o metodă des folosită pentru calculul valorilor proprii pentru o matrice oarecare, în general, nesimetrică. Fie $A \in \mathcal{M}_m(\mathbf{R})$. Metoda QR constă în construirea unui şir de matrice $(A_k)_{k\geq 0}$ folosind factorizarea QR (definită la ssisteme liniare) astfel:

$$A_0 = A = Q_0 R_0\,, \quad A_1 = R_0 Q_0\,,$$
$$A_k = Q_k R_k\,, \quad A_{k+1} = R_k Q_k\,,\ k \geq 0\,.$$

Când $k \to \infty$, elementele de pe diagonala lui $A_k$ converg la valorile proprii ale matricei $A$. Calculul iterativ se opreşte atunci când, pentru orice $i = \overline{1,m}$, avem $\left|a_{ii}^{(k+1)} - a_{ii}^{(k)}\right| < \varepsilon$, pentru $\varepsilon > 0$ o eroare dată.

# Ecuaţii neliniare în $R^m$

## Principiul contracţiei

Fie $m \in \mathbf{N}^*$, $M \subset \mathbf{R}^m$ o mulţime şi $\| \cdot \|$ o normă pe $\mathbf{R}^m$. O funcţie $f : M \to \mathbf{R}$ $^m$ se numeşte $q$-contracţie a mulţimii $M$ relativă la norma $\| \cdot \|$ dacă are următoarele proprietăţi:

1) $\exists\, q \in (0, 1)$ astfel încât $\|f(x) - f(y)\| \leq q\|x - y\|$, $\forall\ x, y \in M$ şi

2) $f(M) \subset M$.

Principiul contracţiei se enunţă astfel:

Fie $\| \cdot \|$ o normă pe $\mathbf{R}^m$, $M$ o mulţime închisă din $\mathbf{R}^m$ şi $f : M \to \mathbf{R}$ $^m$ o $q$-contracţie a mulţimii $M$ relativă la norma $\| \cdot \|$. Atunci ecuaţia $f(x) = x$ are o soluţie unică $z \in M$, şi, pentru orice $x^{(0)} \in M$, şirul $(x^{(n)})_{n\in\mathbf{N}}$ definit prin relaţia de recurenţă $x^{(n+1)} = f(x^{(n)})$ converge la $z$. Avem evaluarea

$$\|x^{(n)} - z\| \leq \frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\| \leq \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|.$$

Fie $a \in \mathbf{R}^m$ şi $r > 0$. Dacă $M := \{x \in \mathbf{R}^m \mid \|x - a\| \leq r\}$, atunci:

1. Dacă $f$ diferenţiabilă pe $M$ şi există $q \in (0, 1)$ astfel încât $\|f'(x)\| \leq q$, $\forall\ x \in M$ rezultă că $\|f(x) - f(y)\| \leq q\|x - y\|$, $\forall\ x, y \in M$.

2. Dacă $\|f(x) - f(y)\| \leq q\|x - y\|$, $\forall\ x, y \in M$ şi $\|f(a) - a\| \leq (1 - q)r$ rezultă că $f(M) \subset M$.

Deci, în aceste condiţii, $f$ este $q$-contracţie a mulţimii $M$.

**Exemplul -1.1** *Fie sistemul de ecuaţii*

$$\begin{cases} x = (80 + x^3 - x^2 - 3yz)/100 \\ y = (60 - y^3 - z^2 + 4xz)/86 \\ z = (40 + z^3 + y^2 - 5xy)/72 \end{cases} \tag{1}$$

*Folosind principiul contracţiei în raport cu norma $\| \cdot \|_\infty$, să se arate că sistemul (1) are soluţie unică $v$ în $V := [0, 2]^3$. Pentru o eroare $\varepsilon$ dată şi pentru iterata iniţială*

$x^{(0)} := (1,1,1)$ *să se calculeze iterata* $x^{(n)}$ *(din principiul contracţiei) pentru care* $\|x^{(n)} - v\|_\infty \le \varepsilon$.

Fie $r = 1$ şi $a = (1,1,1)$. Atunci mulţimea $V = \{x \in \mathbf{R}^m \,|\, \|x - a\|_\infty \le r\}$. Fie funcţia $F = (f, g, h) : V \to \mathbf{R}^{\,3}$ cu $f(x,y,z) := (80 + x^3 - x^2 - 3yz)/100$, $g(x,y,z) := (60 - y^3 - z^2 + 4xz)/86$ şi $h(x,y,z) := (40 + z^3 + y^2 - 5xy)/72$, $\forall\ (x,y,z) \in V$. Sistemul de ecuaţii (1) este echivalent cu $F(x,y,z) = (x,y,z)$.

Avem

$$F'(x,y,z) = \left( \begin{array}{ccc} (3x^2 - 2x)/100 & -3z/100 & -3y/100 \\ 4z/86 & -3y^2/86 & (-2z + 4x)/86 \\ -5y/72 & (2y - 5x)/72 & 3z^2/72 \end{array} \right),$$

deci $\|F'(x,y,z)\|_\infty = \max((|3x^2 - 2x| + |-3z| + |-3y|)/100,\ (|4z| + |-3y^2| + |-2z + 4x|)/86,\ (|-5y| + |2y - 5x| + |3z^2|)/72) \le \max(28/100, 32/86, 36/72) = = 1/2$ (am folosit $0 \le x, y, z \le 2$). Putem lua $q = 1/2$. Pe de altă parte $\|F(a) - a\|_\infty = \max(28/100, 24/86, 35/72) = 35/72 < (1-q)r = 1/2$. Deci $F$ este $1/2$-contracţie a mulţimii $V$ în raport cu norma $\|\cdot\|_\infty$. Din principiul contracţiei rezultă că ecuaţia $F(x,y,z) = (x,y,z)$ are soluţie unică în $V$.

Pentru $x^{(0)} = (1,1,1)$ se calculează $x^{(n)}$ cu formula de recurenţă $x^{(n)} = F(x^{(n-1)})$, $\forall\ n \in N$. Se opreşte procesul iterativ atunci când $q/(1-q)\|x^{(n)} - x^{(n-1)}\|_\infty < \varepsilon$.

Pentru $\varepsilon = 10^{-10}$ se obţin următoarele rezultate:

| $n$ | $x$ | $y$ | $z$ |
|---|---|---|---|
| 1 | 0.7700000000 | 0.7209302326 | 0.5138888889 |
| 2 | 0.7875219889 | 0.7086511560 | 0.5261092736 |
| 3 | 0.7874973924 | 0.7095886570 | 0.5257974572 |
| 4 | 0.7874891632 | 0.7095640029 | 0.5257622709 |
| 5 | 0.7874902776 | 0.7095633762 | 0.5257631334 |
| 6 | 0.7874902723 | 0.7095634355 | 0.5257631103 |
| 7 | 0.7874902718 | 0.7095634337 | 0.5257631082 |
| 8 | 0.7874902719 | 0.7095634337 | 0.5257631083. □ |

## Metoda Gauss-Seidel neliniară

Fie $m \in N^*$, $(a_1, a_2, \ldots, a_m) \in \mathbf{R}^m$, $r_1, r_2, \ldots, r_m$ numere reale strict pozitive, $M := \{x = (x_1, x_2, \ldots, x_m) \in \mathbf{R}^m \mid |x_i - a_i| \leq r_i, \forall\ 1 \leq i \leq m\}$ şi $f = (f_1, f_2, \ldots, f_m) : M \to \mathbf{R}^{\ m}$ o $q$-contracţie a mulţimii $M$ relativă la norma $\|\cdot\|_\infty$.

Pentru orice $x^{(0)} \in M$ definim şirul $(x^{(n)})_{n\in\mathbf{N}}$ prin relaţia de recurenţă

$$x_1^{(n+1)} = f_1(x_1^{(n)}, x_2^{(n)}, \ldots, x_m^{(n)}) \text{ şi}$$

$$x_i^{(n+1)} = f_i(x_1^{(n+1)}, \ldots, x_{i-1}^{(n+1)}, x_i^{(n)}, \ldots, x_m^{(n)})\ \forall\, 2 \leq i \leq m.$$

Atunci şirul $(x^{(n)})_{n\in\mathbf{N}}$ converge către unica soluţie $z \in M$ a ecuaţiei $f(x) = x$. Avem evaluarea:

$$\|x^{(n)} - z\|_\infty \leq \frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\|_\infty \leq \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_\infty.$$

Dacă $f$ diferenţiabilă pe $M$ şi există $q \in (0,1)$ astfel încât $\|f'(x)\|_\infty \leq q$, $\forall\ x \in M$ rezultă că $\|f(x) - f(y)\|_\infty \leq q\|x - y\|_\infty$, $\forall\ x, y \in M$.

**Exemplul -1.2** *Fie sistemul de ecuaţii*

$$\begin{cases} x = (432 + x^4 + y^4 - 2z^4 + y^2z^2)/234 \\ y = (324 - x^4 + 2y^4 - 3z^4 + 3x^2z^2)/432 \\ z = (234 + 2x^4 - 3y^4 + 4z^4 - x^2y^2 + 3x + y)/324 \end{cases} \tag{2}$$

*Folosind metoda Gauss-Seidel neliniară, să se arate că sistemul (2) are soluţie unică $v$ în $V := [-2, 2] \times [-1, 1] \times [0, 1]$. Pentru o eroare $\varepsilon$ dată şi pentru iterata iniţială $x^{(0)} := (0, 0, 1/2)$ să se calculeze iterata $x^{(n)}$ pentru care $\|x^{(n)} - v\|_\infty \le \varepsilon$.*

Fie $r_1 = 2$, $r_2 = 1$ şi $r_3 = 1/2$, $a_1 = a_2 = 0$ şi $a_3 = 1/2$. Atunci mulţimea $V = \{(x, y, z) \in \mathbf{R}^3 \,|\, |x - a_1| \le r_1,\ |y - a_2| \le r_2,\ |z - a_3| \le r_3\}$. Fie funcţia $F = (f, g, h) : V \to \mathbf{R}^3$ cu $f(x, y, z) := (432 + x^4 + y^4 - 2z^4 + y^2z^2)/234$, $g(x, y, z) := (324 - x^4 + 2y^4 - 3z^4 + 3x^2z^2)/432$ şi $h(x, y, z) := (234 + 2x^4 - 3y^4 + 4z^4 - x^2y^2 + 3x + y)/324$, $\forall\ (x, y, z) \in V$. Sistemul de ecuaţii (2) este echivalent cu $F(x, y, z) = (x, y, z)$.

Avem $F'(x, y, z) =$

$$\begin{pmatrix} 4x^3/234 & (4y^3 + 2yz^2)/234 & (-8z^3 + 2y^2z)/234 \\ (-4x^3 + 6xz^2)/432 & 8y^3/432 & (-12z^3 + 6x^2z)/432 \\ (8x^3 - 2xy^2 + 3)/324 & (-12y^3 - 2x^2y + 1)/324 & 16z^3/324 \end{pmatrix},$$

deci $\|F'(x, y, z)\|_\infty = \max((|4x^3| + |4y^3 + 2yz^2| + |-8z^3 + 2y^2z|)/234, (|-4x^3 + 6xz^2| + |8y^3| + |-12z^3 + 6x^2z|)/432, (|8x^3 - 2xy^2 + 3| + |-12y^3 - 2x^2y + 1| + |16z^3|)/324) \le \max(48/234, 88/432, 108/234) = 1/3$ (am folosit $-2 \le x \le 2,\ -1 \le y \le 1,\ 0 \le z \le 1$). Putem lua $q = 1/3$. Pe de altă parte $|f(x, y, z) - a_1| \le (|432| + |x^4| + |y^4| + |2z^4| + |y^2z^2|)/234 \le 452/234 < 2$, $|g(x, y, z) - a_2| \le (|324| + |x^4| + |2y^4| + |3z^4| + |3x^2z^2|)/432 \le 357/432 < 1$ şi $|h(x, y, z) - a_3| \le (|234| + |2x^4| + |3y^4| + |4z^4| + |x^2y^2| + |3x| + |y|)/324 \le 122/324 < 1/2, \forall\ (x, y, z) \in V$. Deci $F(V) \subset V$ şi $F$ este 1/3-contracţie a mulţimii $V$ în raport cu norma $\|\cdot\|_\infty$. Din principiul contracţiei rezultă că ecuaţia $F(x, y, z) = (x, y, z)$ ar soluţie unică în $V$.

Pentru $x^{(0)} = (0, 0, 1/2)$ se calculează $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)})$ cu formula de recurenţă $x_1^{(n)} = f(x^{(n-1)})$, $x_2^{(n)} = g(x_1^{(n)}, x_2^{(n-1)}, x_3^{(n-1)})$, $x_3^{(n)} = h(x_1^{(n)}, x_2^{(n)}, x_3^{(n-1)})$ $\forall\ n \in N$. Se opreşte procesul iterativ atunci când $q/(1-q)\|x^{(n)} - x^{(n-1)}\|_\infty < \varepsilon$.

Pentru $\varepsilon = 10^{-10}$ se obţin următoarele rezultate:

| $n$ | $x$ | $y$ | $z$ |
|---|---|---|---|
| 1 | 1.8498931624 | 0.7283987914 | 0.8064491799 |
| 2 | 1.8952624827 | 0.7347217752 | 0.8182233410 |
| 3 | 1.9002520392 | 0.7348417419 | 0.8193883638 |
| 4 | 1.9008188414 | 0.7348467169 | 0.8195175386 |
| 5 | 1.9008834671 | 0.7348471269 | 0.8195321857 |
| 6 | 1.9008908388 | 0.7348471703 | 0.8195338544 |
| 7 | 1.9008916797 | 0.7348471752 | 0.8195340447 |
| 8 | 1.9008917756 | 0.7348471757 | 0.8195340664 |
| 9 | 1.9008917866 | 0.7348471758 | 0.8195340689 |
| 10 | 1.9008917878 | 0.7348471758 | 0.8195340692 |
| 11 | 1.9008917880 | 0.7348471758 | 0.8195340692 . □ |

## Metoda lui Newton

Fie $m \in \mathbf{N}^*$, $D$ o mulţime deschisă din $\mathbf{R}^m$ şi $f : D \to \mathbf{R}^m$ o funcţie de clasă $\mathcal{C}^2(D)$. Dacă presupunem că $f'(x)$ este inversabilă pentru orice $x \in D$, atunci ecuaţia $f(x) = 0$ se transformă echivalent astfel:

$$f(x) = 0 \Leftrightarrow f(x) + f'(x)(x - x) = 0 \Leftrightarrow x = x - (f'(x))^{-1} f(x).$$

În ipotezele de mai sus (fără a presupune că $f'(x)$ este inversabilă $\forall\, x \in D$), pentru orice $x^{(0)} \in D$ definim, dacă este posibil, şirul $(x^{(n)})_{n \in \mathbf{N}}$ prin relaţia de recurenţă

$$x^{(n+1)} = x^{(n)} - (f\prime(x^{(n)}))^{-1} f(x^{(n)}),\ \ \forall\, n \in \mathbf{N}. \tag{3}$$

Fie $\|\cdot\|$ o normă pe $\mathbf{R}^m$. Notăm tot cu $\|\cdot\|$ norma matricială asociată. Presupunem că există $a \in D$, $r > 0$, $M > 0$, $\alpha > 0$, $\beta > 0$, $\mu > 0$, $q \in (0,1)$ astfel încât:

1) $V := \overline{B_r(a)} := \{x \in \mathbf{R}^m \mid \|x-a\| \leq r\} \subset D$;

2) $\|f''(x)\| \leq M$, $\forall\, x \in V$;

3) Există $(f'(a))^{-1}$ şi $\|(f'(a))^{-1}\| \leq \alpha$;

4) $Mr\alpha < 1$ şi $\dfrac{\alpha}{1-Mr\boldsymbol{\alpha}} \leq \dfrac{1}{\mu}$;

5) $\|(f'(a))^{-1}f(a)\| \leq \beta \leq (1-q)r$;

6) $\dfrac{M}{\mu^2}\|f(x)\| \leq q$, $\forall\, x \in V$;

7) $\dfrac{Mr}{\mu} \leq q$.

Atunci ecuaţia $f(x) = 0$ are o soluţie unică $z \in V$ şi pentru orice $x^{(0)} \in V$ se poate construi şirul $(x^{(n)})_{n\in\mathbf{N}}$ (prin relaţia de recurenţă (3)) cu proprietăţile $x^{(n)} \in V$ şi

$$\|x^{(n)} - z\| \leq \frac{2\mu}{M} q^{2^n}, \ \forall\, n \in \mathbf{N}.$$

**Exemplul -1.3** *Fie sistemul de ecuaţii*

$$\begin{cases} x^3 + 3xy^2 + 3xz^2 - 183x + 2 = 0 \\ y^3 + +3x^2y + 3yz^2 - 193y + 1 = 0 \\ z^3 + 3x^2z + 3y^2z - 188(z-1) - 1 = 0 \end{cases} \tag{4}$$

*Folosind metoda Newton relativă la norma $\|\cdot\|_\infty$, să se arate că sistemul (4) are soluţie unică $v$ în $V := [-1,1]^2 \times [0,2]$. Pentru o eroare $\boldsymbol{\varepsilon}$ dată şi pentru iterata iniţială $x^{(0)} := (0,0,1)$ să se calculeze iterata $x^{(n)}$, pentru care $\|x^{(n)} - v\|_\infty \leq \boldsymbol{\varepsilon}$.*

Fie $r = 1$ şi $a = (0,0,1)$. Atunci mulţimea $V$ este $\overline{B_r(a)}$ în raport cu norma $\|\cdot\|_\infty$. Fie funcţia $F = (f,g,h) : V \to \mathbf{R}^3$ cu $f(x,y,z) := x^3 + 3xy^2 + 3xz^2 - 183x + 2$, $g(x,y,z) := y^3 + +3x^2y + 3yz^2 - 193y + 1$ şi $h(x,y,z) := z^3 + 3x^2z + 3y^2z - 188(z-1) - 1 = 0$, $\forall\,(x,y,z) \in V$. Sistemul de ecuaţii (4) este echivalent cu $F(x,y,z) = 0$.

Avem $F'(x,y,z) =$

$$\begin{pmatrix} 3x^2+3y^2+3z^2-183 & 6xy & 6xz \\ 6xy & 3x^2+3y^2+3z^2-193 & 6yz \\ 6xz & 6yz & 3x^2+3y^2+3z^2-188 \end{pmatrix},$$

$$F''(x,y,z) = \begin{pmatrix} 6x & 6y & 6z & 6y & 6x & 0 & 6z & 0 & 6x \\ 6y & 6x & 0 & 6x & 6y & 6z & 0 & 6z & 6x \\ 6z & 0 & 6x & 0 & 6z & 6y & 6x & 6y & 6z \end{pmatrix},$$

deci $\|F''(x,y,z)\|_\infty \le \max(|18x|+|12y|+|12z|, |12x|+|18y|+|12z|, |12x|+|12y|+|18z|) \le \max(54,54,60) = 60$ (am folosit $-1 \le x, y \le 1,\ 0 \le z \le 2$). Putem lua $M = 60$. Avem $F'(a) = \begin{pmatrix} -180 & 0 & 0 \\ 0 & -190 & 0 \\ 0 & 0 & -185 \end{pmatrix}$, deci $F'(a)$ inversabilă, $(F'(a))^{-1} =$ $\begin{pmatrix} -1/180 & 0 & 0 \\ 0 & -1/190 & 0 \\ 0 & 0 & -1/185 \end{pmatrix}$ şi $\|(F'(a))^{-1}\|_\infty = \max(1/180, 1/190, 1/185) = 1/180$.
Luăm $\alpha = 1/180$. Cum $Mr\alpha = 60/180 < 1$ şi $\alpha/(1 - MR\alpha) = 1/120$, putem lua $\mu = 120$. Avem $F(a) = (2,1,0)^t$ şi $(F'(a))^{-1}F(a) = (-1/90, -1/190, 0)^t$, deci $\|(F'(a))^{-1}F(a)\|_\infty = 1/90$. Luă m $\beta = 1/90$. Condiţia $\beta \le (1-q)r$ implică $q \le 89/90$. Din $\|F(x,y,z)\|_\infty \le \max(|x^3| + |3xy^2| + |3xz^2| + |183x| + 2, |y^3| + |3x^2y| + |3yz^2| + |193y| + 1, |z^3| + |3x^2z| + |3y^2z| + |188(z-1)| + 1) \le \max(201, 210, 209) = 210$ rezultă că $(M/\mu^2)\|F(x,y,z)\|_\infty \le q \Leftrightarrow q \ge 210/240 = 7/8$. Din $Mr/\mu \le q$ rezultă $q \ge 1/2$. Deci putem lua $q = 7/8$ şi atunci ecuaţia $F(x,y,z) = 0$ are o soluţie unică în $V$.

Pentru $x^{(0)} = (0,0,1)$ se calculează $x^{(n)}$ cu formula de recurenţă (4) (folosind la fiecare pas o procedură de inversare a matricei $(F'(x^{(n)}))^{-1}$). Se opreşte calculul iterativ atunci când $(2\mu)/M \cdot q^{2^n} \le \boldsymbol{\varepsilon} \Leftrightarrow 4 \cdot (7/8)^{2^n} \le \boldsymbol{\varepsilon}$.

Pentru $\varepsilon = 10^{-10}$ se obţin următoarele rezultate:

| $n$ | $x$ | $y$ | $z$ |
|---|---|---|---|
| 1 | $-0.0387039145$ | $-0.2345464953$ | 0.8497295394 |
| 2 | 0.0109754027 | 0.0044322958 | 1.0031237498 |
| 3 | 0.0111111051 | 0.0052632384 | 0.9999155652 |
| 4 | 0.0111111257 | 0.0052631697 | 1.0000048646 |
| 5 | 0.0111111247 | 0.0052631693 | 1.0000023842 |
| 6 | 0.0111111248 | 0.0052631693 | 1.0000024531 |
| 7 | 0.0111111248 | 0.0052631693 | 1.0000024512 . $\square$ |

## Metoda lui Newton simplificată

Fie $m \in \mathbf{N}^*$, $D$ o mulţime deschisă din $\mathbf{R}^m$ şi $f : D \to \mathbf{R}^m$ o funcţie de clasă $\mathcal{C}^2(D)$. Dacă presupunem că există $a \in D$ astfel încât $f'(a)$ este inversabilă, atunci ecuaţia $f(x) = 0$ se transformă echivalent astfel:

$$f(x) = 0 \Leftrightarrow f(x) + f'(a)(x - x) = 0 \Leftrightarrow x = x - (f'(a))^{-1} f(x).$$

Pentru orice $x^{(0)} \in D$ definim, dacă este posibil, şirul $(x^{(n)})_{n\in\mathbf{N}}$ prin relaţia de recurenţă

$$x^{(n+1)} = x^{(n)} - (f'(a))^{-1} f(x^{(n)}),\ \forall\, n \in \mathbf{N}. \tag{5}$$

Fie $\|\cdot\|$ o normă pe $\mathbf{R}^m$. Notăm tot cu $\|\cdot\|$ norma matricială asociată. Presupunem că există $r > 0$, $M > 0$, $\alpha > 0$, $\beta > 0$, $q \in (0, 1)$ astfel încât:

1) $V := \overline{B_r(a)} := \{x \in \mathbf{R}^m \mid \|x - a\| \leq r\} \subset D$;

2) $\|f''(x)\| \leq M$, $\forall\, x \in V$;

3) $\|(f'(a))^{-1}\| \leq \alpha$;

4) $\|(f'(a))^{-1} f(a)\| \leq \beta \leq (1 - q)r$;

5) $Mr\alpha \leq q$.

Atunci ecuaţia $f(x) = 0$ are o soluţie unică $z \in V$ şi pentru orice $x^{(0)} \in V$ se poate construi şirul $(x^{(n)})_{n\in\mathbf{N}}$ (prin relaţia de recurenţă (5)) cu proprietă ţile $x^{(n)} \in V$ şi

$$\|x^{(n)} - z\| \le \frac{q}{1-q}\|x^{(n)} - x^{(n-1)}\| \le \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|,\ \forall\, n \in \mathbf{N}.$$

**Exemplul -1.4** *Fie sistemul de ecuaţii*

$$\begin{cases} x^2 + y^2 + z^2 + x + y + z - 853/144 = 0 \\ 3x^2 - 4y^2 + z^2 + x - y - 3z + 1309/432 = 0 \\ 5x^2 + 2y^2 - 3z^2 + 2x - 4y - 6z + 37/9 = 0 \end{cases} \tag{6}$$

*Folosind metoda Newton simplificată relativă la norma $\|\cdot\|_\infty$, să se arate că sistemul (6) are soluţie unică $v$ în $V := [-0,8; 1,2]^3$. Pentru o eroare $\boldsymbol{\varepsilon}$ dată şi pentru iterata iniţială $x^{(0)} := (1,1,1)$, să se calculeze iterata $x^{(n)}$ pentru care $\|x^{(n)} - v\|_\infty \le \boldsymbol{\varepsilon}$.*

Fie $r := 0,2 = 1/5$ şi $a := (1,1,1)$. Atunci mulţimea $V$ este $\overline{B_r(a)}$ în raport cu norma $\|\cdot\|_\infty$. Fie funcţia $F = (f,g,h) : V \to \mathbf{R}^3$ cu $f(x,y,z) := x^2 + y^2 + z^2 + x + y + z - 853/144$, $g(x,y,z) := 3x^2 - 4y^2 + z^2 + x - y - 3z + 1309/432$ şi $h(x,y,z) := 5x^2 + 2y^2 - 3z^2 + 2x - 4y - 6z + 37/9$, $\forall\,(x,y,z) \in V$. Sistemul de ecuaţii (6) este echivalent cu $F(x,y,z) = 0$.

Avem

$$F'(x,y,z) = \begin{pmatrix} 2x+1 & 2y+1 & 2z+1 \\ 6x+1 & -8y-1 & 2z-3 \\ 10x+2 & 4y-4 & -6z-6 \end{pmatrix},$$

$$F''(x,y,z) = \begin{pmatrix} 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 6 & 0 & 0 & 0 & -8 & 0 & 0 & 0 & 2 \\ 10 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & -6 \end{pmatrix},$$

deci $\|F''(x,y,z)\|_\infty \le \max(|2|+|2|+|2|, |6|+|-8|+|2|, |10|+|4|+|-6|) = 20, \forall\,(x,y,z) \in V$. Putem lua $M = 20$. Avem $F'(a) = \begin{pmatrix} 3 & 9 & 3 \\ 7 & -9 & -1 \\ 12 & 0 & -12 \end{pmatrix}$, cu $\det(F'(a)) = 72$, deci

$F'(a)$ inversabilă, $(F'(a))^{-1} = \frac{1}{72}\begin{pmatrix} 9 & 3 & 2 \\ 6 & -6 & 2 \\ 9 & 3 & -4 \end{pmatrix}$ şi $\|(F'(a))^{-1}\|_\infty = \max(14/72, 14/72, 16/72) = 2/9$. Luă m $\alpha = 2/9$. Avem $F(a) = (11/144, 13/432, 1/9)^t$ şi $(F'(a))^{-1}F(a) = (1/72, 1/144, 1/216)^t$, deci $\|(F'(a))^{-1}F(a)\|_\infty = 1/72$. Luăm $\beta = 1/72$. Condiţia $\beta \leq (1-q)r$ implică $q \leq 67/72$. Din $\alpha M r \leq q$ rezultă $q \geq 8/9$. Deci putem lua $q = 8/9$ şi atunci ecuaţia $F(x, y, z) = 0$ are o soluţie unică în $V$.

Pentru $x^{(0)} = (1, 1, 1)$ se calculează $x^{(n)}$ cu formula de recurenţă (6). Se opreşte calculul iterativ atunci când $q/(1-q) \cdot \|x^{(n)} - x^{(n-1)}\|_\infty \leq \varepsilon$.

Pentru $\varepsilon = 10^{-10}$ se obţin următoarele rezultate:

| $n$ | $x$ | $y$ | $z$ |
|---|---|---|---|
| 1 | 0.9861111111 | 0.9930555556 | 0.9953703704 |
| 2 | 0.9860336380 | 0.9930399270 | 0.9953759520 |
| 3 | 0.9860328001 | 0.9930398800 | 0.9953760624 |
| 4 | 0.9860327911 | 0.9930398803 | 0.9953760635 |
| 5 | 0.9860327910 | 0.9930398803 | 0.9953760635 . □ |

# INTERPOLARE POLINOMIALĂ

## Formula lui Lagrange

Se consideră:

- $n \in \mathbf{N}^*$;
- $x_1, x_2, \ldots, x_n \in \mathbf{R}$ cu $x_i \neq x_j$ pentru orice $1 \leq i \neq j \leq n$ puncte (noduri) distincte două câte două;
- $y_1, y_2, \ldots, y_n \in \mathbf{R}$ numere reale.

Spunem că aceste elemente definesc un sistem de date notat cu $(SL)$.

Se numeşte polinom de interpolare Lagrange ataşat sistemului de date $(SL)$ un polinom $P \in \mathbf{R}[X]$ cu următoarele proprietăţi:

1) $\operatorname{grad} P \leq n-1$;

2) $P(x_i) = y_i$ pentru orice $1 \leq i \leq n$.

Polinomul de interpolare Lagrange ataşat sistemului de date $(SL)$ există, este unic şi se reprezintă sub forma:

$$P(x) = \sum_{i=1}^{n} y_i \prod_{j=1,\ j\neq i}^{n} \frac{x - x_j}{x_i - x_j}. \tag{1}$$

Formula (1) se numeşte formula lui Lagrange de reprezentare a polinomului de interpolare.

**Exemplul -1.1** *Utilizând formula lui Lagrange de reprezentare să se determine polinomul de interpolare de grad 2 ce verifică condiţiile:*

$$P(0) = 2,\ P(1) = 1,\ P(4) = 4.$$

Din (1) rezultă că:

$$P(x) = P(0)L_1(x) + P(1)L_2(x) + P(4)L_3(x).$$

Avem:

$$L_1(x) = \frac{(x-1)(x-4)}{(0-1)(0-4)} = \frac{x^2-5x+4}{4},$$
$$L_2(x) = \frac{x(x-4)}{(1-0)(1-4)} = \frac{x^2-4x}{-3},$$
$$L_3(x) = \frac{x(x-1)}{(4-0)(4-1)} = \frac{x^2-x}{12}.$$

Atunci:

$$P(x) = 2 \cdot \frac{x^2-5x+4}{4} + 1 \cdot \frac{x^2-4x}{-3} + 4 \cdot \frac{x^2-x}{12} = \frac{1}{2}x^2 - \frac{3}{2}x + 2.\square$$

## Algoritmul lui Aitken

Fie $n \in \mathbf{N}^*$ şi $x_1, x_2, \ldots, x_n$ puncte distincte două câte două, conţinute într-un interval real $I$. Considerăm o funcţie continuă $f : I \to \mathbf{R}$. Se numeşte polinom de interpolare Lagrange asociat funcţiei $f$ şi nodurilor $x_1, x_2, \ldots, x_n$ polinomul $P$ de grad mai mic sau egal cu $n-1$ care verifică condiţiile $P(x_i) = f(x_i)$. Notăm valoarea acestui polinom într-un punct $x$ din $I$ cu $P(f; x_1, x_2, \ldots, x_n; x)$. Polinomul de interpolare Lagrange asociat funcţiei $f$ şi nodurilor $x_1, x_2, \ldots, x_n$ există şi este unic.

Dacă $n = 1$, atunci

$$P(f; x_1; x) = f(x_1), \tag{2}$$

iar dacă $n \geq 2$, atunci $P(f; x_1, x_2 \ldots, x_n; x) =$

$$= \frac{P(f; x_2, x_3 \ldots, x_n; x)(x - x_1) - P(f; x_1, x_2 \ldots, x_{n-1}; x)(x - x_n)}{x_n - x_1}. \tag{3}$$

Pentru orice $1 \leq i \leq n$ şi $i \leq j \leq n$ definim recurent polinoamele $P_{ij}$ astfel:

$$P_{1j}(x) := P(f; x_j; x) = f(x_j) \quad \forall 1 \leq j \leq n \text{ şi}$$

$$P_{ij}(x) := P(f; x_1, x_2, \ldots, x_{i-1}, x_j; x) \quad \forall\, 2 \le i \le n, \;\; \forall\, i \le j \le n. \tag{4}$$

Cu formula (3) rezultă că

$$P_{ij}(x) = \frac{(x - x_{i-1})P_{i-1j} - (x - x_j)P_{i-1i-1}}{x_j - x_{i-1}} \quad \forall\, 2 \le i \le n, \;\; \forall\, i \le j \le n. \tag{5}$$

Observăm că $P(f; x_1, x_2, \ldots, x_n; x) = P_{nn}(x)$.

Prezentăm în continuare un algoritm pentru calculul coeficienţilor polinomului de interpolare Lagrange $P(f; x_1, x_2, \ldots, x_n; x)$ (bazat pe algoritmul Aitken):

Cum grad $P_{ij} \le i - 1 \;\forall\, 1 \le i \le n$, $\forall\, i \le j \le n$, putem scrie

$$P_{ij}(x) = a_1^{ij} + a_2^{ij} x + \cdots + a_i^{ij} x^{i-1}.$$

Egalând coeficienţii lui $x^k$ cu $0 \le k \le i - 1$ din relaţile (4) şi (5), obţinem:

$$a_1^{1j} = f(x_j) \quad \forall\, 1 \le j \le n,$$

$$\begin{cases} a_1^{ij} = \dfrac{-a_1^{i-1\,j} x_{i-1} + a_1^{i-1\,i-1} x_j}{x_j - x_{i-1}} \\ a_k^{ij} = \dfrac{a_{k-1}^{i-1\,j} - a_k^{i-1\,j} x_{i-1} - a_{k-1}^{i-1\,i-1} + a_k^{i-1\,i-1} x_j}{x_j - x_{i-1}} \quad \forall\, 2 \le k \le i - 1 \\ a_i^{ij} = \dfrac{a_{i-1}^{i-1\,j} - a_{i-1}^{i-1\,i-1}}{x_j - x_{i-1}} \end{cases},$$

$\forall\, 2 \le i \le n$, $\forall\, i \le j \le n$. Deci se observă că determinarea coeficienţilor $a_k^{ij}$ este recursivă. În concluzie, avem:

$$P(f; x_1, x_2, \ldots, x_n; x) = a_1^{nn} + a_2^{nn} x + \cdots + a_n^{nn} x^{n-1}.$$

**Exemplul -1.2** *Fie funcţia* $f : [0, 12] \to R$ *cu* $f(x) = 6\sqrt{2x + 1} - 2x$, $\forall\, x \in [0, 12]$. *Să se determine polinomul de interpolare Lagrange* $P(f; 0, 4, 12; x)$ *utilizând algoritmul Aitken.*

Nodurile sunt $x_1 = 0, x_2 = 4, x_3 = 12$. Atunci $P_{11}(x) = f(0) = 6$, $P_{12}(x) = f(4) = 10$, $P_{13}(x) = f(12) = 6$. Rezultă că

$$P_{22}(x) = \frac{10x - 6(x - 4)}{4} = x + 6,$$

$$P_{23}(x) = \frac{6x - 6(x-12)}{12} = 6$$

şi astfel

$$P(f; 0, 4, 12; x) = P_{33}(x) = -\frac{x^2}{8} + \frac{3x}{2} + 6\,.$$

## Algoritmul lui Neville

Fie $n \in \mathbf{N}^*$ şi $x_1, x_2, \ldots, x_n$ puncte distincte două câte două conţinute într-un interval real $I$. Considerăm o funcţie continuă $f : I \to \mathbf{R}$.

Pentru orice $1 \leq i \leq n$ şi $i \leq j \leq n$ definim recurent polinoamele $P_{ij}$ astfel:

$$P_{1j}(x) := P(f; x_j; x) = f(x_j) \quad \forall\, 1 \leq j \leq n \text{ şi}$$

$$P_{ij}(x) = P(f; x_{j-i+1}, \ldots, x_{j-1}, x_j; x) \quad \forall\, 2 \leq i \leq n, \ \ \forall\, i \leq j \leq n. \tag{6}$$

Cu formula (3) rezultă că

$$P_{ij}(x) = \frac{(x - x_{j-i+1})P_{i-1\,j} - (x - x_j)P_{i-1\,j-1}}{x_j - x_{j-i+1}} \quad \forall\, 2 \leq i \leq n, \ \ \forall\, i \leq j \leq n. \tag{7}$$

Observăm că $P(f; x_1, x_2, \ldots, x_n; x) = P_{nn}(x)$.

Prezentăm în continuare un algoritm pentru calculul coeficienţilor polinomului de interpolare Lagrange $P(f; x_1, x_2, \ldots, x_n; x)$ (bazat pe algoritmul Neville):

Cum $\operatorname{grad} P_{ij} \leq i - 1 \ \forall\, 1 \leq i \leq n, \forall\, i \leq j \leq n$, putem scrie

$$P_{ij}(x) = a_1^{ij} + a_2^{ij} x + \cdots + a_i^{ij} x^{i-1}$$

Egalând coeficienţii lui $x^k$ cu $0 \leq k \leq i - 1$, din relaţile (6) şi (7) obţinem:

$$a_1^{1j} = f(x_j) \quad \forall\, 1 \leq j \leq n,$$

$$\begin{cases} a_1^{ij} = \dfrac{-a_1^{i-1\,j} x_{j-i+1} + a_1^{i-1\,j-1} x_j}{x_j - x_{j-i+1}} \\ a_k^{ij} = \dfrac{a_{k-1}^{i-1\,j} - a_k^{i-1\,j} x_{j-i+1} - a_{k-1}^{i-1\,j-1} + a_k^{i-1\,j-1} x_j}{x_j - x_{j-i+1}} & \forall\, 2 \leq k \leq i-1, \\ a_i^{ij} = \dfrac{a_{i-1}^{i-1\,j} - a_{i-1}^{i-1\,j-1}}{x_j - x_{j-i+1}} \end{cases}$$

$\forall\, 2 \leq i \leq n,\ \forall\, i \leq j \leq n$. Deci se observă că determinarea coeficienţilor $a_k^{ij}$ este recursivă. Astfel, avem:

$$P(f; x_1, x_2, \ldots, x_n; x) = a_1^{nn} + a_2^{nn} x + \cdots + a_n^{nn} x^{n-1}\,.$$

**Exemplul -1.3** *Fie funcţia* $f : [0, 12] \to R$ *cu* $f(x) = 6\sqrt{2x+1} - 2x$, $\forall\, x \in [0, 12]$. *Să se determine polinomul de interpolare Lagrange* $P(f; 0, 4, 12; x)$ *utilizând algoritmul Neville.*

Nodurile sunt $x_1 = 0, x_2 = 4, x_3 = 12$. Atunci $P_{11}(x) = f(0) = 6$, $P_{12}(x) = f(4) = 10$, $P_{13}(x) = f(12) = 6$. Rezultă că

$$P_{22}(x) = \frac{10x - 6(x-4)}{4} = x + 6,$$

$$P_{23}(x) = \frac{6(x-4) - 10(x-12)}{8} = -\frac{x}{2} + 12$$

şi astfel

$$P(f; 0, 4, 12; x) = P_{33}(x) = -\frac{x^2}{8} + \frac{3x}{2} + 6$$

## Formula lui Newton de reprezentare a polinomului de interpolare

Se consideră:

- $n \in \mathbf{N}^*$;
- $x_1, x_2, \ldots, x_n \in \mathbf{R}$ un sistem de puncte (noduri);
- $m \in \mathbf{N}^*$ numărul de puncte distincte;
- $y_1, y_2, \ldots, y_m \in \mathbf{R}$ punctele distincte ale sistemului;
- $n_1, n_2, \ldots, n_m \in \mathbf{N}^*$ ordinele de multiplicitate ale punctelor distincte în sistem;
- $z_{ij} \in \mathbf{R}$ cu $1 \leq i \leq m,\ 0 \leq j \leq n_i - 1$ familie de elemente reale.

Spunem că aceste elemente definesc un sistem de date notat cu $(SH)$.

Se numeşte polinom de interpolare Hermite ataşat sistemului de date $(SH)$ un polinom $P \in \mathbf{R}[x]$ cu următoarele proprietăţi:

1) $\operatorname{grad} P \leq n-1$;

2) $P^{(j)}(y_i) = z_{ij}$ pentru orice $1 \leq i \leq m,\ 0 \leq j \leq n_i - 1$.

Polinomul de interpolare Hermite ataşat sistemului dedate $(SH)$ există şi este unic.

Fie $I$ un interval real ce conţine punctele $x_1, \ldots, x_n$ şi $f : I \to \mathbf{R}$ o funcţie de $n_i - 1$ ori derivabilă în $y_i$ pentru orice $1 \leq i \leq m$. Luăm $z_{ij} := f^{(j)}(y_i),\ \forall\, 1 \leq i \leq m,\ 0 \leq j \leq n_i - 1$. Polinomul de interpolare Hermite ataşat acestui sistem de date se numeşte polinom de interpolare asociat funcţiei $f$ şi sistemului de date $(SH)$. Notăm valoarea acestui polinom într-un punct $x$ din $I$ cu $P(f; x_1, x_2, \ldots, x_n; x)$.

Se numeşte diferenţă divizată a funcţiei $f$ relativă la sistemul de noduri $x_1, x_2, \ldots, x_n$ coeficientul lui $x^{n-1}$ din reprezentarea canonică a polinomului de interpolare $P(f; x_1, x_2, \ldots, x_n; x)$. Notăm diferenţa divizată cu $f(x_1, x_2, \ldots, x_n)$. Diferenţele divizate se calculează cu ajutorul următoarelor formule:

Pentru $n \geq 2$ şi $x_1 \neq x_n$

$$f(x_1, x_2, \ldots, x_n) = \frac{f(x_2, \ldots, x_n) - f(x_1, \ldots, x_{n-1})}{x_n - x_1},$$

iar

$$f(\underbrace{a, a, \ldots, a}_{n \text{ ori}}) = \frac{f^{(n-1)}(a)}{(n-1)!}. \tag{8}$$

Următoarea egalitate este formula lui Newton de reprezentare a polinomului de interpolare:

$$P(f; x_1, x_2, \ldots, x_n; x) = f(x_1) + \sum_{i=2}^{n} f(x_1, x_2, \ldots, x_i) \prod_{j=1}^{i-1} (x - x_j).$$

Fie $f : I \to R$ o funcţie de $n$ ori derivabilă pe $I$. Atunci

$$|f(x) - P(f; x_1, x_2, \ldots, x_n; x)| \leq \frac{1}{n!} \cdot \max_{x \in I} \left| f^{(n)}(x) \right| \cdot \left| \prod_{i=1}^{n} (x - x_i) \right|, \forall x \in I.$$

**Algoritm:**

Notăm cu $a_{ij}$ diferenţele divizate $f(x_{j-i+1},\ldots,x_{j-1,}x_j)$, $\forall\ i \in \{2,\ldots,n\}, \forall\ j \in \{i,\ldots,n\}$ si cu $c_i = a_{ii}$. Atunci $P(x) := P(f;x_1,x_2,\ldots,x_n;x) = c_1 + \sum\limits_{i=2}^{n} c_i \prod\limits_{j=1}^{i-1}(x-x_j)$

Pentru $i \in \overline{1,n}$ luăm $a_{1i} = f(x_i)$ si $c_1 = a_1$.

Apoi pentru $i = \overline{2,n}$, calculăm diferenţele divizate $a_{ij} = \dfrac{a_{i-1:j} - a_{i-1:j-1}}{x_j - x_{j-i+1}}$ $\forall\ j = \overline{i,n}$ si luam $c_i = a_{ii}$.

Polinomul $P$ în punctul $x$ îl putem calcula cu o schemă de tip Horner astfel:

Are loc descompunerea:

$$P = [\ldots[[[c_n](x-x_{n-1}) - c_{n-1}](x-x_{n-2}) + c_{n-2}](x-x_{n-3}) + \cdots$$

$$\cdots + c_2](x-x_1) + c_1$$

Rezultă schema de calcul:

$P = c_n;\ P = P\cdot(x-x_{n-1})+c_{n-1};\ P = P\cdot(x-x_{n-2})+c_{n-2};\ldots;\ P = P\cdot(x-x_1)+c_1$

deci

- $P := c_n$;
- pentru $i = n-1$ până la 1 luăm $P = P\cdot(u-x_i)+c_i$;

**Exemplul -1.4** *Fie funcţia* $f : [0,4] \to R$ *cu* $f(x) = 6\sqrt{2x+1} - 2x$, $\forall\, x \in [0,4]$. *Utilizând formula de reprezentare Newton să se determine polinomul de interpolare Hermite* $P(x) = P(f;0,0,0,4,4;x)$. *Se se evalueze eroarea* $|f(x)-P(x)|$.

Avem $x_1 = x_2 = x_3 = 0$, $x_4 = x_5 = 4$. Din formula de reprezentare Newton avem:

$$P(f;0,0,0,4,4;x) = f(0) + f(0,0)x + f(0,0,0)x^2 +$$

$$+f(0,0,0,4)x^3 + f(0,0,0,4,4)x^3(x-4)$$

Vom calcula diferenţele divizate cu formulele de recurenţă date de (8). Calculele le vom organiza într-un tabel de forma:

| *nod* | *d.d.* 1$n$ | *d.d.* 2$n$ | *d.d.* 3$n$ | *d.d.* 4$n$ | *d.d.* 5$n$ |
|---|---|---|---|---|---|
| $x_1$ | $f(x_1)$ | $f(x_1,x_2)$ | $f(x_1,x_2,x_3)$ | $f(x_1,x_2,x_3,x_4)$ | $f(x_1,\dots,x_5)$ |
| $x_2$ | $f(x_2)$ | $f(x_2,x_3)$ | $f(x_2,x_3,x_4)$ | $f(x_2,x_3,x_4,x_5)$ | |
| $x_3$ | $f(x_3)$ | $f(x_3,x_4)$ | $f(x_3,x_4,x_5)$ | | |
| $x_4$ | $f(x_4)$ | $f(x_4,x_5)$ | | | |
| $x_5$ | $f(x_5)$ | | | | |

(am notat cu "d.d.$k$n " diferenţe divizate asociate la $k$ noduri).

Cum $f(x) = 6\sqrt{2x+1} - 2x = 6(2x+1)^{1/2} - 2x$, rezultă că $f'(x) = 6(2x+1)^{-1/2} - 2$ şi $f''(x) = -6(2x+1)^{-3/2}$. Atunci $f(0) = 6$, $f(4) = 10$, $f(0,0) = f'(0) = 4$, $f(0,4) = (f(4) - f(0))/(4-0) = 1$, $f(4,4) = f'(4) = 0$, $f(0,0,0) = \frac{1}{2}f''(0) = -3$, $f(0,0,4) = (f(0,4) - f(0,0))/(4-0) = -3/4$ etc.. Rezultă următorul tabel în cazul datelor problemei:

| *nod* | 1$n$ | 2$n$ | 3$n$ | 4$n$ | 5$n$ |
|---|---|---|---|---|---|
| 0 | 6 | 4 | $-3$ | $9/16$ | $-7/64$ |
| 0 | 6 | 4 | $-3/4$ | $2/16$ | |
| 0 | 6 | 1 | $-1/4$ | | |
| 4 | 10 | 0 | | | |
| 4 | 10 | | | | |

Atunci (cu formula de reprezentare Newton) avem

$$P(f;0,0,0,4,4;x) = 6 + 4x - 3x^2 + \frac{9}{16}x^3 - \frac{7}{64}x^3(x-4).$$

Din formula de evaluare a erorii rezultă că

$$|f(x) - P(f;0,0,0,4,4;x)| \le \frac{M_5}{5!}|x^3(x-4)^2| \quad \forall\, x \in [0,4],$$

unde $M_5 = \sup_{x\in[0,4]} |f^{(5)}(x)|$. Avem $f^{(3)}(x) = 18(2x+1)^{-5/2}$, $f^{(4)}(x) = -90(2x+1)^{-7/2}$, $f^{(5)}(x) = 630(2x+1)^{-9/2}$. Cum $f^{(6)}(x) = -630\cdot 9(2x+1)^{-11/2} < 0$, rezultă că $f^{(5)}$este descrescătoare şi cum $f^{(5)}(x) > 0\ \forall\, x \in [0,4]$ rezultă că $M_5 = f^{(5)}(0) = 630$. Atunci

$$|f(x) - P(x)| \leq \frac{630}{120}|x^3(x-4)^2| = \frac{21}{4}x^3(x-4)^2, \quad \forall\, x \in [0,4]. \ \square$$

## Interpolare Cebâşev

În cele ce urmează ne propunem să *diminuăm eroarea la interpolarea polinomială Lagrange* pe un interval dat $[a,b]$ în cazul unui număr dat $n \in N^*$ de noduri prin alegerea unor "noduri optimale". Definim "nodurile optimale" $y_1, y_2, \ldots, y_n$ prin relaţia:

$$\sup_{x\in[a,b]} \prod_{i=1}^{n} |x - y_i| \leq \sup_{x\in[a,b]} \prod_{i=1}^{n} |x - x_i|, \tag{9}$$

pentru orice puncte $x_1, x_2, \ldots, x_n$ din $[a,b]$. În continuare vVom construi aceste "noduri optimale".

Funcţia $T_n : R \to R$ dată de $T_n(x) = \cos(n \arccos x)$ este un polinom de gradul $n$ şi are $n$ rădăcini distincte situate în intervalul $[-1,1]$.

Fie $n \in N^*$ şi $a < b \in R$. Notăm

$$P_n(x) = 2\left(\frac{b-a}{4}\right)^n T_n\left(\frac{2x-a-b}{b-a}\right) \quad \forall\, x \in [a,b] \tag{10}$$

Atunci

$$\sup_{x\in[a,b]} |P_n(x)| \leq \sup_{x\in[a,b]} |Q(x)|,$$

pentru orice polinom $Q$ monic de grad $n$.

"Nodurile optimale", pe un interval dat $[a,b]$, la interpolarea polinomială Lagrange cu $n$ puncte, sunt:

$$y_i = \frac{b-a}{2}\cos\left(\frac{(2i-1)\pi}{2n}\right) + \frac{b+a}{2} \quad \forall\, i \in \overline{1,n}.$$

## Formula lui Newton de interpolare ascendentă

Fie $I$ un interval real, familia $(x_j)_{j\in\mathbf{Z}}$ o reţea de noduri (reale) echidistante cu pasul $h$ definită prin $x_0 \in I$ şi $x_j = x_0 + jh$, $\forall\, j \in \mathbf{N}$. Notăm $q := \dfrac{x - x_0}{h}$ unde $x \in I$.

Pentru $x \in I$ şi $n \in \mathbf{N}^*$ notăm $A_x^n = \prod\limits_{j=0}^{n-1}(x-j)$ numite aranjamente generalizate şi $C_x^n = \dfrac{A_x^n}{n!}$ numite combinări generalizate. Prin convenţie luăm $A_x^0 = 1$ şi $C_x^0 = 1$.

Fie $f : I \to \mathbf{R}$ o funcţie, $x \in I$ şi $h > 0$ un număr real. Numim **diferenţă nedivizată ascendentă a funcţiei $f$ în punctul $x$, corespunzătoare pasului $h$** (şi o notăm cu $\Delta_h f(x)$), expresia:

$$\Delta_h f(x) := f(x+h) - f(x).$$

Fie $n \in \mathbf{N}^*$. Numim **diferenţă nedivizată ascendentă de ordin $n$ a funcţiei $f$ în punctul $x$, corespunzătoare pasului $h$** (şi o notăm cu $\Delta_h^n f(x)$) compunerea:

$$\Delta_h^n f(x) := \Delta_h^{n-1}(\Delta_h f(x)),$$

unde, prin convenţie, $\Delta_h^0 f(x) = f(x)$, iar $\Delta_h^1 = \Delta_h$.

Cu datele şi notaţiile anterioare este adevărată egalitatea

$$P(f; x_0, x_1, \ldots, x_n; x) = \sum_{i=0}^{n} C_q^i \Delta_h^i f(x_0),$$

numită formula lui Newton de interpolare ascendentă.

**Exemplul -1.5** *Să se determine, folosind formula de interpolare Newton ascendentă, polinomul Lagrange ataşat datelor: $n = 3$, $x_0 = 4$, $h = 2$ (de unde $x_1 = 6$, $x_2 = 8$, $x_3 = 10$) şi $f : R \to R$ o funcţie continuă cu $f(x_0) = 1$, $f(x_1) = 3$, $f(x_2) = 8$, $f(x_3) = 20$ şi să se estimeze valoarea funcţiei $f$ în punctul $x = 7$.*

Organizăm calculul diferenţelor nedivizate ascendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\Delta_h^1 f(x_k)$ | $\Delta_h^2 f(x_k)$ | $\Delta_h^3 f(x_k)$ |
|---|---|---|---|---|---|
| 0 | 4 | 1 | 2 | 3 | 4 |
| 1 | 6 | 3 | 5 | 7 | |
| 2 | 8 | 8 | 12 | | |
| 3 | 10 | 20 | | | |

(De exemplu $\Delta_h^1 f(x_0) = \Delta_h f(x_0) = f(x_1) - f(x_0) = 3 - 1 = 2$, $\Delta_h^2 f(x_0) = \Delta_h^1 \Delta_h f(x_0) = \Delta_h f(x_1) - \Delta_h f(x_0) = \Delta_h^1 f(x_1) - \Delta_h^1 f(x_0) = 5 - 2 = 3$, etc.) Atunci $\Delta_h^0 f(x_0) = 1$, $\Delta_h^1 f(x_0) = 2$, $\Delta_h^2 f(x_0) = 3$, $\Delta_h^3 f(x_0) = 4$.

Fie $x \in \mathbf{R}$ Avem $q = \dfrac{x - x_0}{h} = \dfrac{x-4}{2}$şi atunci $C_q^0 = 1$, $C_q^1 = q = \dfrac{x-4}{2}$, $C_q^2 = \dfrac{q(q-1)}{2} = \dfrac{(x-4)(x-6)}{8}$, $C_q^3 = \dfrac{q(q-1)(q-2)}{6} = \dfrac{(x-4)(x-6)(x-8)}{48}$. Rezultă că: $P(f; x_0, x_1, x_2, x_3; x) =$

$$\sum_{i=0}^{3} C_q^i \Delta_h^i f(x_0) = 1 + (x-4) + \frac{3}{8}(x-4)(x-6) + \frac{4}{48}(x-4)(x-6)(x-8)$$

şi deci

$$P(f; x_0, x_1, x_2, x_3; 7) = 1 + 3 + \frac{9}{8} - \frac{1}{4} = \frac{39}{8} = 4.875\ .\square$$

## Formula lui Newton de interpolare descendentă

Cu notaţiile de la secţiunea anterioară, fie $f : I \to \mathbf{R}$ o funcţie, $x \in I$ şi $h > 0$ un număr real. Numim **diferenţă nedivizată descendentă a funcţiei $f$ în punctul $x$, corespunzătoare pasului** $h$ (şi o notăm cu $\nabla_h f(x)$), expresia:

$$\nabla_h f(x) := f(x) - f(x - h).$$

Fie $n \in \mathbf{N}^*$. Numim **diferenţă nedivizată descendentă de ordinul $n$ a funcţiei $f$ în punctul $x$, corespunzătoare pasului** $h$ (şi o notăm cu $\nabla_h^n f(x)$) compunerea:

$$\nabla_h^n f(x) := \nabla_h^{n-1}(\nabla_h f(x)),$$

unde, prin convenţie, $\nabla_h^0 f(x) = f(x)$, iar $\nabla_h^1 = \nabla_h$.. Atunci este adevărată egalitatea

$$P(f; x_0, x_{-1}, \ldots, x_{-n}; x) = \sum_{i=0}^{n} (-1)^i C_{-q}^i \nabla_h^i f(x_0),$$

numită formula lui Newton de interpolare descendentă.

**Exemplul -1.6** *Să se determine, folosind formula de interpolare Newton descendentă, polinomul Lagrange ataşat datelor:* $n = 3$, $x_0 = 10$, $h = 2$ *(de unde* $x_{-1} = 8$, $x_{-2} = 6$, $x_{-3} = 4$*) şi* $f : \mathbf{R} \to \mathbf{R}$ *o funcţie continuă cu* $f(x_0) = 20$, $f(x_{-1}) = 8$, $f(x_{-2}) = 3$, $f(x_{-3}) = 1$ *şi să se estimeze valoarea funcţiei* $f$ *în punctul* $x = 7$.

Organizăm calculul diferenţelor nedivizate descendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\nabla_h^1 f(x_k)$ | $\nabla_h^2 f(x_k)$ | $\nabla_h^3 f(x_k)$ |
|---|---|---|---|---|---|
| 0 | 10 | 20 | 12 | 7 | 4 |
| $-1$ | 8 | 8 | 5 | 3 | |
| $-2$ | 6 | 3 | 2 | | |
| $-3$ | 4 | 1 | | | |

(De exemplu $\nabla_h^1 f(x_0) = \nabla_h f(x_0) = f(x_0) - f(x_{-1}) = 20 - 8 = 12$, $\nabla_h^2 f(x_0) = \nabla_h^1 \nabla_h f(x_0) = \nabla_h f(x_0) - \nabla_h f(x_{-1}) = \nabla_h^1 f(x_0) - \nabla_h^1 f(x_{-1}) = 12 - 5 = 7$, etc.) Atunci $\nabla_h^0 f(x_0) = 20$, $\nabla_h^1 f(x_0) = 12$, $\nabla_h^2 f(x_0) = 7$, $\nabla_h^3 f(x_0) = 4$.

Fie $x \in \mathbf{R}$. Avem $q = \dfrac{x - x_0}{h} = \dfrac{x - 10}{2}$ şi atunci $C_{-q}^0 = 1$, $C_{-q}^1 = -q = -\dfrac{x-10}{2}$, $C_{-q}^2 = \dfrac{-q(-q-1)}{2} = \dfrac{(x-10)(x-8)}{8}$, $C_{-q}^3 = \dfrac{-q(-q-1)(-q-2)}{6} = \dfrac{(x-10)(x-8)(x-6)}{48}$. Rezultă că:

$$P(f; x_0, x_{-1}, x_{-2}, x_{-3}; x) = \sum_{i=0}^{3} (-1)^i C_{-q}^i \nabla_h^i f(x_0) =$$

$$= 20 + \frac{12}{2}(x - 10) + \frac{7}{8}(x - 10)(x - 8) + \frac{4}{48}(x - 10)(x - 8)(x - 6)$$

şi deci

$$P(f; x_0, x_{-1}, x_{-2}, x_{-3}; 7) = 20 - 18 + \frac{21}{8} + \frac{1}{4} = \frac{39}{8} = 4.875\ .\square$$

## Formula de interpolare Gauss înainte

Fie $f : I \to R$ o funcţie, $x \in I$ şi $h > 0$ un număr real. Numim **diferenţă nedivizată centrală a funcţiei** $f$ **în punctul** $x$ **corespunzătoare pasului** $h$ (şi o notăm cu $\delta_h f(x)$), expresia:

$$\delta_h f(x) := f\left(x + \frac{h}{2}\right) - f\left(x - \frac{h}{2}\right).$$

Fie $n \in N^*$. Numim **diferenţă nedivizată centrală de ordinul** $n$ **a funcţiei** $f$ **în punctul** $x$**, corespunzătoare pasului** $h$ (şi o notăm cu $\delta_h^n f(x)$) compunerea:

$$\delta_h^n f(x) := \delta_h^{n-1}(\delta_h f(x)),$$

unde, prin convenţie, $\delta_h^0 f(x) = f(x)$, iar $\delta_h^1 f(x) = \delta_h f(x)$.

Atunci este adevărată egalitatea

$$P(f; x_{-n} \dots, x_{-1}, x_0, x_1 \dots, x_n; x) =$$

$$= f(x_0) + \sum_{i=1}^{n} (C_{q+i-1}^{2i-1} \delta_h^{2i-1} f(x_{1/2}) + C_{q+i-1}^{2i} \delta_h^{2i} f(x_0))$$

(unde $x_{1/2} = x + \dfrac{1}{2}h$), numită formula de interpolare Gauss înainte.

**Exemplul -1.7** *Să se determine, folosind formula de interpolare Gauss înainte, polinomul Lagrange ataşat datelor:* $n = 2$, $x_0 = 6$, $h = 2$ *(de unde* $x_{-2} = 2$, $x_{-1} = 4$, $x_1 = 8$, $x_2 = 10$*) şi* $f : R \to R$ *o funcţie continuă cu* $f(x_{-2}) = -2$, $f(x_{-1}) = 1$, $f(x_0) = 3$, $f(x_1) = 8$, $f(x_2) = 20$ *şi să se estimeze valoarea funcţiei* $f$ *în punctul* $x = 7$.

Printr-un calcul elementar, se arată că, pentru orice $i = \overline{1, 2}$:

$$\delta_h^{2i-1} f(x_{1/2}) = \Delta_h^{2i-1} f(x_{-i+1}) \text{ şi } \delta_h^{2i} f(x_0)) = \Delta_h^{2i} f(x_{-i})).$$

Atunci: $P(f;x_{-2},x_{-1},x_0,x_1,x_2;x)=$

$$f(x_0)+\sum_{i=1}^{2}(C_{q+i-1}^{2i-1}\Delta_h^{2i-1}f(x_{-i+1})+C_{q+i-1}^{2i}\Delta_h^{2i}f(x_{-i}))=$$

$$=f(x_0)+C_q^1\Delta_h^1 f(x_0)+C_q^2\Delta_h^2 f(x_{-1})+C_{q+1}^3\Delta_h^3 f(x_{-1})+C_{q+1}^4\Delta_h^4 f(x_{-2}).$$

Organizăm calculul diferenţelor finite ascendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\Delta_h^1 f(x_k)$ | $\Delta_h^2 f(x_k)$ | $\Delta_h^3 f(x_k)$ | $\Delta_h^4 f(x_k)$ |
|---|---|---|---|---|---|---|
| $-2$ | 2 | $-2$ | 3 | $-1$ | 4 | 0 |
| $-1$ | 4 | 1 | 2 | 3 | 4 | |
| 0 | 6 | 3 | 5 | 7 | | |
| 1 | 8 | 8 | 12 | | | |
| 2 | 10 | 20 | | | | |

Avem: $f(x_0)=3, \Delta_h^1 f(x_0)=5, \Delta_h^2 f(x_{-1})=3, \Delta_h^3 f(x_{-1})=4, \Delta_h^4 f(x_{-2})=0$. Cum $q=\dfrac{x-6}{2}$ rezultă că $C_q^1=q=\dfrac{x-6}{2}$, $C_q^2=\dfrac{q(q-1)}{2}=\dfrac{(x-6)(x-8)}{8}$, $C_{q+1}^3=\dfrac{(q+1)q(q-1)}{6}=\dfrac{(x-4)(x-6)(x-8)}{48}$. Atunci:

$$P(f;x_{-2},x_{-1},x_0,x_1,x_2;x)=$$

$$=3+5\frac{x-6}{2}+3\frac{(x-6)(x-8)}{8}+4\frac{(x-4)(x-6)(x-8)}{48}$$

şi

$$f(7)\simeq P(f;x_{-2},x_{-1},x_0,x_1,x_2;7)=3+\frac{5}{2}-\frac{3}{8}-\frac{1}{4}=\frac{39}{8}=4.875\ .\square$$

## Formula de interpolare Gauss înapoi

Cu notaţiile din secţiunea anterioară este adevărată egalitatea:

$$P(f;x_{-n}\ldots,x_{-1},x_0,x_1\ldots,x_n;x)=$$

$$=f(x_0)+\sum_{i=1}^{n}(C_{q+i-1}^{2i-1}\delta_h^{2i-1}f(x_{-1/2})+C_{q+i}^{2i}\delta_h^{2i}f(x_0))$$

(unde $x_{-1/2}=x_0-\dfrac{1}{2}h$), numită formula de interpolare Gauss înapoi.

**Exemplul -1.8** *Să se determine, folosind formula de interpolare Gauss înapoi, polinomul Lagrange ataşat datelor:* $n = 2$, $x_0 = 6$, $h = 2$ *(de unde* $x_{-2} = 2$, $x_{-1} = 4$, $x_1 = 8$, $x_2 = 10$*) şi* $f : R \to R$ *o funcţie continuă cu* $f(x_{-2}) = -2$, $f(x_{-1}) = 1$, $f(x_0) = 3$, $f(x_1) = 8$, $f(x_2) = 20$ *şi să se estimeze valoarea funcţiei* $f$ *în punctul* $x = 7$.

Printr-un calcul elementar, se arată că, pentru orice $i = \overline{1,2}$:

$$\delta_h^{2i-1} f(x_{1/2}) = \Delta_h^{2i-1} f(x_{-i+1}) \text{ şi } \delta_h^{2i} f(x_0)) = \Delta_h^{2i} f(x_{-i})).$$

Atunci: $P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; x) =$

$$f(x_0) + \sum_{i=1}^{2} (C_{q+i-1}^{2i-1} \Delta_h^{2i-1} f(x_{-i}) + C_{q+i}^{2i} \Delta_h^{2i} f(x_{-i})) =$$

$$= f(x_0) + C_q^1 \Delta_h^1 f(x_{-1}) + C_{q+1}^2 \Delta_h^2 f(x_{-1}) + C_{q+1}^3 \Delta_h^3 f(x_{-2}) + C_{q+2}^4 \Delta_h^4 f(x_{-2}).$$

Organizăm calculul diferenţelor finite ascendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\Delta_h^1 f(x_k)$ | $\Delta_h^2 f(x_k)$ | $\Delta_h^3 f(x_k)$ | $\Delta_h^4 f(x_k)$ |
|---|---|---|---|---|---|---|
| $-2$ | 2 | $-2$ | 3 | $-1$ | 4 | 0 |
| $-1$ | 4 | 1 | 2 | 3 | 4 | |
| 0 | 6 | 3 | 5 | 7 | | |
| 1 | 8 | 8 | 12 | | | |
| 2 | 10 | 20 | | | | |

Avem: $f(x_0) = 3, \Delta_h^1 f(x_{-1}) = 2, \Delta_h^2 f(x_{-1}) = 3, \Delta_h^3 f(x_{-2}) = 4, \Delta_h^4 f(x_{-2}) = 0$. Cum $q = \dfrac{x-6}{2}$, rezultă că $C_q^1 = q = \dfrac{x-6}{2}$, $C_{q+1}^2 = \dfrac{(q+1)q}{2} = \dfrac{(x-4)(x-6)}{8}$, $C_{q+1}^3 = \dfrac{(q+1)q(q-1)}{6} = \dfrac{(x-4)(x-6)(x-8)}{48}$. Atunci:

$$P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; x) =$$

$$= 3 + 2\frac{x-6}{2} + 3\frac{(x-4)(x-6)}{8} + 4\frac{(x-4)(x-6)(x-8)}{48}$$

şi

$$f(7) \simeq P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; 7) = 3 + 1 + \frac{9}{8} - \frac{1}{4} = \frac{39}{8} = 4.875 \;. \square$$

## Interpolare cu funcţii spline cubice

Fie $I = [a,b] \subset \mathbf{R}$, $\Delta_n = (a = x_1 < x_2 < \ldots < x_n = b)$ o diviziune a intervalului $[a,b]$ şi $f_i \in R$, $i = \overline{1,n}$ pentru $n \in N^\star$.

Se numeşte **funcţie spline cubică, corespunzătoare diviziunii** $\Delta_n$ o funcţie $s : [a,b] \longrightarrow \mathbf{R}$ care satisface condiţiile:

$$\begin{cases} i)\ s|_{[x_i, x_{i+1}]} \text{ este polinom de gradul trei pentru } i = \overline{1, n-1} \\ \\ ii)\ s \in \mathcal{C}^2([a,b])\,. \end{cases}$$

Se notează cu $\mathcal{S}_3(\Delta_n)$ mulţimea funcţiilor spline cubice, corespunzătoare diviziunii $\Delta_n$.

**Funcţia spline cubică asociată sistemului de puncte** $(x_i, f_i)_{i=\overline{1,n}}$ este o funcţie cubică corespunzătoare diviziunii $\Delta_n$, $s : [a,b] \to R$, cu proprietatea

$$s\,(x_i) = f_i\,,\ i = \overline{1,n}\,.$$

În condiţile formulate mai sus, $\forall\, i = \overline{1, n-1}$, funcţia $s$ se reprezintă în mod unic sub forma:

$$s(x) = c_{1i} + c_{2i}\,(x - x_i) + \frac{c_{3i}}{2}\,(x - x_i)^2 + \frac{c_{4i}}{6}\,(x - x_i)^3\,,\ \forall x \in [x_i, x_{i+1}], \tag{11}$$

unde

$$\begin{cases} c_{1i} &=& f_i\,, \\ c_{2i} &=& s_i\,, \\ c_{3i} &=& \dfrac{1}{x_{i+1} - x_i}\left[6\dfrac{f_{i+1} - f_i}{x_{i+1} - x_i} - (4s_i + 2s_{i+1})\right], \\ \\ c_{4i} &=& \dfrac{6}{(x_{i+1} - x_i)^2}\left[(s_i + s_{i+1}) - 2\dfrac{f_{i+1} - f_i}{x_{i+1} - x_i} - \right], \end{cases} \tag{12}$$

iar $s_1, \ldots, s_n$ sunt soluţiile sistemului:

$$
\begin{gathered}
(x_{i+1} - x_i)\, s_{i-1} + 2\,(x_{i+1} - x_{i-1})\, s_i + (x_i - x_{i-1})\, s_{i+1} = \\
3\left[\frac{f_i - f_{i-1}}{x_i - x_{i-1}}\,(x_{i+1} - x_i) + \frac{f_{i+1} - f_i}{x_{i+1} - x_i}\,(x_i - x_{i-1})\right]\ ,\ i = \overline{2, n-1}\,.
\end{gathered} \tag{13}
$$

Pentru determinarea unică a spline-ului cubic, sistemul (12) trebuie completat cu condiţii la capete. De exemplu, putem presupune că sunt cunoscute valorile pentru $s\,''(x_1)$ şi $s\,''(x_n)$:

$$
s\,''(x_1) = \alpha \quad , \quad s\,''(x_n) = \beta\,, \tag{14}
$$

sau că sunt cunoscute valorile pentru $s\,'(x_1)$ şi $s\,'(x_n)$:

$$
s_1 = s\,'(x_1) = \alpha \quad , \quad s_n = s\,'(x_n) = \beta\,. \tag{15}
$$

Relaţiile (14) sunt echivalente cu:

$$
\begin{cases}
2s_1 + s_2 & = \quad 3\dfrac{f_2 - f_1}{x_2 - x_1} - \dfrac{\alpha}{2}\,(x_2 - x_1) \\[2em]
s_{n-1} + 2s_n & = \quad 3\dfrac{f_n - f_{n-1}}{x_n - x_{n-1}} + \dfrac{\beta}{2}\,(x_n - x_{n-1})\,.
\end{cases}
$$

**Remarcă**. Din (13) şi (**??**) sau din (13) şi (15) din se pot determina $s_1, \ldots, s_n$, apoi din (12) se determină $c_{1i}$, $c_{2i}$, $c_{3i}$, $c_{4i}$ pentru $i = \overline{1, n-1}$ şi deci, conform cu (11), funcţia $s$ este cunoscută pe $[a, b]$. Pentru $f : [a, b] \to \mathbf{R}$ de două ori derivabilă determinăm funcţia spline cubică de interpolare asociată diviziunii $\Delta_n$ şi functiei $f$, adică, o functie spline cubica $s \in \mathcal{S}_3(\Delta_n)$ cu proprietăţile $s(x_i) = f(x_i)$, $\forall\, i \in \overline{1, n}$ şi

$$
s\,''(a) = f\,''(a) \text{ şi } s\,''(b) = f\,''(b)\,.
$$

$$
\text{sau } \ s\,'(a) = f\,'(a) \text{ şi } s\,'(b) = f\,'(b)\,.
$$

# Derivare numerică

Derivarea numerică constă în aproximarea derivatelor de un anumit ordin ale unei funcţii, printr-o combinaţie liniară de valori ale funcţiei în anumite puncte.

**Formule pentru derivata de ordinul I**

Fie $f : \mathbf{R} \to \mathbf{R}$ o funcţie derivabilă, $x_0 \in \mathbf{R}$ şi $h > 0$ un număr real numit pas. Atunci

$$f'(x_0) = \frac{f(x_0+h) - f(x_0)}{h} + \mathcal{R}(f), \tag{1}$$

unde, dacă $f \in \mathcal{C}^2([x_0, x_0+h])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq \frac{h}{2} \cdot \max_{x \in [x_0, x_0+h]} |f''(x)|.$$

$$f'(x_0) = \frac{f(x_0) - f(x_0-h)}{h} + \mathcal{R}(f), \tag{2}$$

unde, dacă $f \in \mathcal{C}^2([x_0-h, x_0])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq \frac{h}{2} \cdot \max_{x \in [x_0-h, x_0]} |f''(x)|.$$

Formulele de derivare (1) @si (2) sunt exacte pe mulţimea polinoamelor de gradul 1.

$$f'(x_0) = \frac{f(x_0+h) - f(x_0-h)}{2h} + \mathcal{R}(f), \tag{3}$$

unde, dacă $f \in \mathcal{C}^3([x_0-h, x_0+h])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq \frac{h^2}{6} \cdot \max_{x \in [x_0-h, x_0+h]} |f'''(x)|.$$

Formula de derivare numerică anterioară se poate generaliza astfel:

$$f'(x_0) \simeq \frac{h_1^2 f(x_0+h_2) + (h_2^2 - h_1^2) f(x_0) - h_2^2 f(x_0-h_1)}{h_1 h_2 (h_1+h_2)},$$

unde $h_1, h_2$ sunt numere reale pozitive.

$$f'(x_0) = \frac{-3f(x_0) + 4f(x_0+h) - f(x_0+2h)}{2h} + \mathcal{R}(f), \tag{4}$$

unde, dacă $f \in \mathcal{C}^3([x_0, x_0+2h])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq \frac{h^2}{3} \cdot \max_{x \in [x_0, x_0+2h]} |f'''(x)| .$$

Formulele de derivare (3) @si (4) sunt exacte pe mulţimea polinoamelor de gradul 2.

$$f'(x_0) = \frac{-11f(x_0) + 18f(x_0+h) - 9f(x_0+2h) + 2f(x_0+3h)}{6h} + \mathcal{R}(f), \qquad (5)$$

unde, dacă $f \in \mathcal{C}^4([x_0, x_0+3h])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq \frac{h^3}{18} \cdot \max_{x \in [x_0, x_0+3h]} \left| f^{(4)}(x) \right|$$

Formula de derivare (5) este exactă pe mulţimea polinoamelor de gradul 3.

$$f'(x_0) = \frac{f(x_0-2h) - 8f(x_0-h) + 8f(x_0+h) - f(x_0+2h)}{12h} + \mathcal{R}(f), \qquad (6)$$

unde, dacă $f \in \mathcal{C}^5([x_0-2h, x_0+2h])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq \frac{2h^4}{5} \cdot \max_{x \in [x_0-2h, x_0+2h]} \left| f^{(5)}(x) \right| .$$

Formula de derivare (6) este exactă pe mulţimea polinoamelor de gradul 4.

bf Formule pentru derivata de ordinul II

Fie $f : \mathbf{R} \to \mathbf{R}$ o funcţie de două ori derivabilă, $x_0 \in \mathbf{R}$ şi $h > 0$ un număr real numit pas. Atunci

$$f''(x_0) = \frac{f(x_0) - 2f(x_0+h) + f(x_0+2h)}{h^2} + \mathcal{R}(f), \qquad (7)$$

unde, dacă $f \in \mathcal{C}^3([x_0, x_0+2h])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq h \cdot \max_{x \in [x_0, x_0+2h]} |f'''(x)| .$$

Formula de derivare (6) este exactă pe mulţimea polinoamelor de gradul 2.

$$f''(x_0) = \mathcal{R}(f) +$$

$$\frac{-f(x_0-2h)+16f(x_0-h)-30f(x_0)+16f(x_0+h)-f(x_0+2h)}{12h^2}, \qquad (8)$$

unde, dacă $f \in \mathcal{C}^5([x_0-2h, x_0+2h])$, restul $\mathcal{R}(f)$ poate fi evaluat prin

$$|\mathcal{R}(f)| \leq \frac{h^3}{5} \cdot \max_{x\in[x_0-2h,x_0+2h]} \left|f^{(5)}x)\right|.$$

Formula de derivare (6) este exactă pe mulţimea polinoamelor de gradul 4.

**Exemplul -1.1** *Să se aproximeze derivatele de ordinul I şi ordinul II ale funcţiei* $f(x) = e^{-x}sin(x)$ *în punctul* $x_0 = 1$, *cu metodele prezentate mai sus, luând pentru pasul* $h$ *valorile* $0.1; 0.01; 0.001; 0.0001$.

Cu formula de derivare (1) se obţin rezultatele: $h = 0.1 \Rightarrow f'(x_0) \simeq -0.1290271591; h = 0.01 \Rightarrow f'(x_0) \simeq -0.1127645338; h = 0.001 \Rightarrow f'(x_0) \simeq -0.1109923614; h = 0.0001 \Rightarrow f'(x_0) \simeq -0.1108136394$.

Cu formula de derivare (2) se obţin rezultatele: $h = 0.1 \Rightarrow f'(x_0) \simeq -0.0891707946, h = 0.01 \Rightarrow f'(x_0) \simeq -0.1087891083, h = 0.001 \Rightarrow f'(x_0) \simeq -0.1105948299, h = 0.0001 \Rightarrow f'(x_0) \simeq -0.1107738899$.

Cu formula de derivare (3) se obţin rezultatele: $h = 0.1 \Rightarrow f'(x_0) \simeq -0.1090989768, h = 0.01 \Rightarrow f'(x_0) \simeq -0.1107768211, h = 0.001 \Rightarrow f'(x_0) \simeq -0.1107935957, h = 0.0001 \Rightarrow f'(x_0) \simeq -0.1107937646$.

Cu formula de derivare (4) se obţin rezultatele: $h = 0.1 \Rightarrow f'(x_0) \simeq -0.1138788298, h = 0.01 \Rightarrow f'(x_0) \simeq -0.1108273447, h = 0.001 \Rightarrow f'(x_0) \simeq -0.1107941027, h = 0.0001 \Rightarrow f'(x_0) \simeq -0.1107937669$.

Cu formula de derivare (4) se obţin rezultatele: $h = 0.1 \Rightarrow f'(x_0) \simeq -0.1138788298, h = 0.01 \Rightarrow f'(x_0) \simeq -0.1108273447, h = 0.001 \Rightarrow f'(x_0) \simeq -0.1107941027, h = 0.0001 \Rightarrow f'(x_0) \simeq -0.1107937669$.

Cu formula de derivare (5) se obţin rezultatele: $h = 0.1 \Rightarrow f'(x_0) \simeq -0.1110871310, h = 0.01 \Rightarrow f'(x_0) \simeq -0.1107940736, h = 0.001 \Rightarrow f'(x_0) \simeq -0.1107937643, h = 0.0001 \Rightarrow f'(x_0) \simeq -0.1107937699$.

Cu formula de derivare (6) se obţin rezultatele: $h = 0.1 \Rightarrow f'(x_0) \simeq -0.1107952264, h = 0.01 \Rightarrow f'(x_0) \simeq -0.1107937655, h = 0.001 \Rightarrow f'(x_0) \simeq -0.1107937652, h = 0.0001 \Rightarrow f'(x_0) \simeq -0.1107937669$.

Valoarea exactă este $f'(x_0) = -0.1107937653$.

Cu formula de derivare (7) se obţin rezultatele: $h = 0.1 \Rightarrow f''(x_0) \simeq -0.3029665865, h = 0.01 \Rightarrow f''(x_0) \simeq -0.3874378126, h = 0.001 \Rightarrow f''(x_0) \simeq -0.3965174074, h = 0.0001 \Rightarrow f''(x_0) \simeq -0.3974491847$.

Cu formula de derivare (8) se obţin rezultatele: $h = 0.1 \Rightarrow f''(x_0) \simeq -0.3975339923, h = 0.01 \Rightarrow f''(x_0) \simeq -0.3975322253, h = 0.001 \Rightarrow f''(x_0) \simeq -0.3975308497, h = 0.0001 \Rightarrow f''(x_0) \simeq -0.3974643429$.

Valoarea exactă este $f''(x_0) = -0.3975322207$.

**Derivare numerică obţinută prin interpolare**

Fie $f : [a,b] \to \mathbf{R}$, $n \in \mathbf{N}^*, x_1, \ldots, x_n$ puncte distincte din $[a,b]$. Presupunem că $f$ este derivabilă. Se numeşte derivare numerică obţinută prin interpolare a lui $f$ aproximarea lui $f'(x)$ prin $P'(f; x_1, \ldots, x_n; x)$, $\forall x \in [a,b]$.

Din formula de reprezentare a lui Newton avem:

$$P(f; x_1, \ldots, x_n; x) = f(x_1) + \sum_{i=2}^{n} f(x_1, \ldots, x_i) \cdot \prod_{j=1}^{i-1} (x - x_j).$$

Rezultă că:

$$P'(f; x_1, \ldots, x_n; x) = f(x_1, x_2) + \sum_{i=3}^{n} f(x_1, \ldots, x_i) \cdot \left( \sum_{j=1}^{i-1} \prod_{k=1,\ k \neq j}^{i-1} (x - x_k) \right),$$

$\forall$ $(n \geq 3)$. Pentru calculul diferenţelor divizate se foloseşte algoritmul Neville: fie, prin notaţie, $F_{ij} = f[x_{j-i+1}, \ldots, x_{j-1,} x_j]$. Atunci $F_{ij}$ satisfac relaţiile de recurenţă: $F_{1i} = f(x_i), \forall\, i \in \overline{1,n}$ şi $F_{ij} = \dfrac{F_{i-1:j} - F_{i-1:j-1}}{x_j - x_{j-i+1}}, \qquad \forall\, i \in \{2, \ldots, n\}, \forall\, j \in \{i, \ldots, n\}$, deci $f(x_1, \ldots, x_i) = F_{ii}$, $\forall i \in \{2, \ldots, n\}$ .

# Cuadraturi numerice

Fie $[a,b] \subset \mathbf{R}$, $x_1, x_2, \ldots, x_n$ puncte distincte din $[a,b]$ şi $f : [a,b] \to \mathbf{R}$ o funcţie continuă. Se numeşte formulă de cuadratură egalitatea

$$\sigma(f) = \sum_{i=1}^{n} c_i f(x_i), \text{ unde } c_i \in \mathbf{R}, \forall\, i = \overline{1,n}.$$

Se numeşte integrare numerică aproximarea integralei $\int\limits_a^b f(x)dx$ cu $\sigma(f)$.

Fie $k \in N^*$, $h = (b-a)/k$, $x_i = a + ih$, $\forall\, i \in \overline{0,k}$. Dacă $\sigma_i(f)$ este o formulă de cuadratură a funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\, i \in \overline{0,k-1}$, atunci se defineşte formula de cuadratură sumată

$$\sigma^{(k)}(f) = \sum_{i=0}^{k-1} \sigma_i(f).$$

# Cuadraturi închise

## Formula de cuadratură a trapezului

Formula de cuadratură a trapezului este

$$\sigma(f) = \int\limits_a^b P\,(f; a, b; x)\, dx.$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\,(f; a, b; x) = f\,(a) + f(a,b)(x-a) = f(a) + \frac{f(b) - f(a)}{b-a}(x-a).$$

Atunci

$$\sigma(f) = \int\limits_a^b \left( f(a) + \frac{f(b)-f(a)}{b-a}(x-a) \right) dx =$$

$$= (b-a) f\,(a) + \frac{f(b)-f(a)}{b-a} \cdot \frac{(b-a)^2}{2} =$$

$$= \frac{b-a}{2} \cdot \left( f(a) + f(b) \right).$$

Dacă $f \in \mathcal{C}^2([a,b])$, din formula de evaluare a erorii la interpolare, avem

$$|f(x) - P(f; a, b; x)| \leq \frac{1}{2} |(x-a)(x-b)| \cdot \max_{x\in[a,b]} |f''(x)|, \ \forall\, x \in [a,b].$$

Atunci

$$\left| \int_a^b f(x)dx - \sigma(f) \right| = \left| \int_a^b (f(x) - P(f; a, b; x))\, dx \right| \leq$$

$$\leq \int_a^b |f(x) - P(f; a, b; x)|\, dx \leq \frac{1}{2} \max_{x\in[a,b]} |f''(x)| \cdot \int_a^b (x-a)(b-x)dx =$$

$$= \frac{1}{2} \max_{x\in[a,b]} |f'\prime(x)| \cdot \frac{(b-a)^3}{6} = \frac{(b-a)^3}{12} \cdot \max_{x\in[a,b]} |f''(x)|.$$

Formula de cuadratură a trapezului sumată este:

$$\sigma_T^{(k)}(f) = \frac{1}{2} \sum_{i=0}^{k-1} (x_{i+1} - x_i)(f(x_i) + f(x_{i+1})) = \frac{b-a}{2k} \sum_{i=0}^{k-1} (f(x_i) + f(x_{i+1})),$$

pentru care, dacă $f \in \mathcal{C}^2([a,b])$, avem evaluarea erorii:

$$\left| \int_a^b f(x)dx - \sigma_T^{(k)}(f) \right| \leq \sum_{i=0}^{k-1} \left| \int_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f) \right| \leq \frac{(b-a)^3}{12k^2} \cdot \max_{x\in[a,b]} |f''(x)|,$$

unde $\sigma_i(f)$ este formula de cuadratură a trapezului, aplicată funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\, i \in \overline{0, k-1}$

**Exemplul -1.1** *Folosind formula de cuadratură a trapezului sumată, să se aproximeze calculul integralei*

$$I = \int_1^2 \frac{1+\ln x}{x} dx$$

*pentru $k=5$, $k=10$, $k=50$ şi $k=100$, evaluând şi eroarea comisă.*

Se obţin următoarele rezultate:

| $k$ | $\sigma_D^{(k)}(f)$ | eroarea |
|---|---|---|
| 5 | 0.9328067513 | 0.0033333333 |
| 10 | 0.9332299640 | 0.0008333333 |
| 50 | 0.9333679124 | 0.0000333333 |
| 100 | 0.9333722435 | 0.0000083333 |

În stabilirea erorii am ţinut seama de $\max\limits_{x\in[a,b]} |f''(x)| = 1$. Facem observaţia că valoarea exactă a integralei este $I = 0.9333736875$ $\square$.

**Exemplul -1.2** *Să se determine numărul minim de subintervale $k$ în care trebuie împărţit intervalul $[0,1]$, pentru ca, folosind formula trapezului sumată, să putem calcula integrala*

$$I = \int_0^1 \frac{e^x}{x+1} dx$$

*cu o eroare mai mică de $10^{-5}$.*

Avem $\sup\limits_{x\in[0,1]} |f''(x)| = \frac{3e}{8}$. Din formula de evaluare a erorii rezultă că

$$k \geq \left[\sqrt{\frac{10^5}{12}}\right] + 1 = 92 \ \square$$

## Formula de cuadratură Simpson

Formula de cuadratură Simpson este

$$\sigma(f) = \int_a^b P\left(f; a, \frac{a+b}{2}, b; x\right) dx$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\left(f; a, \frac{a+b}{2} b; x\right) = f(a) + f\left(a, \frac{a+b}{2}\right)(x-a) +$$

$$+f\left(a,\frac{a+b}{2},b\right)(x-a)\left(x-\frac{a+b}{2}\right).$$

Atunci

$$\sigma(f)=\int_a^b\left(f(a)+f\left(a,\frac{a+b}{2}\right)(x-a)+\right.$$

$$\left.+f\left(a,\frac{a+b}{2},b\right)(x-a)\left(x-\frac{a+b}{2}\right)\right)dx=$$

$$=(b-a)f(a)+\frac{2(f((a+b)/2)-f(a))}{b-a}\cdot\frac{(b-a)^2}{2}+$$

$$+\frac{2(f(b)-2f((a+b)/2)+f(a))}{(b-a)^2}\cdot\frac{(b-a)^3}{12}=$$

$$=\frac{b-a}{6}\cdot\left(f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right).$$

Dacă $f\in\mathcal{C}^3([a,b])$, din formula de evaluare a erorii la interpolare avem

$$\left|f(x)-P\left(f;a,\frac{a+b}{2},b;x\right)\right|\leq$$

$$\leq\frac{1}{6}\left|(x-a)\left(x-\frac{a+b}{2}\right)(x-b)\right|\cdot\max_{x\in[a,b]}\left|f'''(x)\right|,\ \forall\ x\in[a,b].$$

Atunci

$$\left|\int_a^b f(x)dx-\sigma(f)\right|=\left|\int_a^b\left(f(x)-P\left(f;a,\frac{a+b}{2},b;x\right)\right)dx\right|\leq$$

$$\leq\int_a^b\left|f(x)-P\left(f;a,\frac{a+b}{2},b;x\right)\right|dx\leq$$

$$\leq\frac{1}{6}\max_{x\in[a,b]}|f''(x)|\cdot\int_a^b\left|(x-a)\left(x-\frac{a+b}{2}\right)(x-b)\right|dx=$$

$$=\frac{1}{6}\max_{x\in[a,b]}|f'''(x)|\cdot\frac{(b-a)^4}{32}=\frac{(b-a)^4}{192}\cdot\max_{x\in[a,b]}|f'''(x)|.$$

**Observaţie:** Dacă $f$ este derivabilă în punctul $(a+b)/2$, atunci formula de cuadratură Simpson se poate obţine şi cu formula

$$\sigma(f)=\int_a^b P\left(f;a,\frac{a+b}{2},\frac{a+b}{2},b;x\right)dx.$$

În acest caz, dacă $f \in \mathcal{C}^4([a,b])$, din formula de evaluare a erorii la interpolare, avem

$$\left| f(x) - P\left(f; a, \frac{a+b}{2}, \frac{a+b}{2}, b; x\right)\right| \leq$$

$$\leq \frac{1}{24}\left|(x-a)\left(x - \frac{a+b}{2}\right)^2 (x-b)\right| \cdot \max_{x\in[a,b]} \left|f^{(4)}(x)\right|, \ \forall\, x \in [a,b].$$

Atunci

$$\left|\int\limits_a^b f(x)dx - \sigma(f)\right| = \left|\int\limits_a^b \left(f(x) - P\left(f; a, \frac{a+b}{2}, \frac{a+b}{2}, b; x\right)\right) dx\right| \leq$$

$$\leq \int\limits_a^b \left|f(x) - P\left(f; a, \frac{a+b}{2}, \frac{a+b}{2}, b; x\right)\right| dx \leq$$

$$\leq \max_{x\in[a,b]} \left|f^{(4)}(x)\right| \cdot \int\limits_a^b \left|(x-a)\left(x - \frac{a+b}{2}\right)(x-b)\right|^2 dx =$$

$$= \frac{(b-a)^5}{2880} \cdot \max_{x\in[a,b]} \left|f^{(4)}(x)\right|.$$

Formula de cuadratură Simpson sumată este:

$$\sigma_S^{(k)}(f) = \frac{1}{6}\sum_{i=0}^{k-1}(x_{i+1} - x_i)\left(f\left(x_i\right) + 4f\left(\frac{x_i + x_{i+1}}{2}\right) + f(x_{i+1})\right) =$$

$$= \frac{b-a}{6k}\sum_{i=0}^{k-1}\left(f\left(x_i\right) + 4f\left(\frac{x_i + x_{i+1}}{2}\right) + f(x_{i+1})\right),$$

pentru care, dacă $f \in \mathcal{C}^3([a,b])$, avem evaluarea erorii:

$$\left|\int\limits_a^b f(x)dx - \sigma_S^{(k)}(f)\right| \leq \sum_{i=0}^{k-1}\left|\int\limits_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f)\right| \leq \frac{(b-a)^4}{192k^3} \cdot \max_{x\in[a,b]} \left|f\prime''(x)\right|,$$

unde $\sigma_i(f)$ este formula de cuadratură Simpson, aplicată funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\, i \in \overline{0, k-1}$, iar dacă $f \in \mathcal{C}^4([a,b])$, avem evaluara erorii:

$$\left|\int\limits_a^b f(x)dx - \sigma_S^{(k)}(f)\right| \leq \sum_{i=0}^{k-1}\left|\int\limits_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f)\right| \leq \frac{(b-a)^5}{2880k^4} \cdot \max_{x\in[a,b]} \left|f^{(4)}(x)\right|$$

unde $\sigma_i(f)$ este formula de cuadratură Simpson aplicată funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\, i \in \overline{0, k-1}$,

**Exemplul -1.3** *Folosind formula de cuadratură Simpson sumată să se aproximeze calculul integralei*

$$I = \int_1^2 \frac{1 + \ln x}{x} dx$$

*pentru* $k = 5$, $k = 10$ *şi* $k = 50$, *evaluând şi eroarea comisă.*

Se obţin următoarele rezultate:

| $k$ | $\sigma_D^{(k)}(f)$ | eroarea |
|---|---|---|
| 5 | 0.9333710349 | 0.0000005555 |
| 10 | 0.9333735173 | 0.0000000347 |
| 50 | 0.9333736875 | 0.0000000000 |

În stabilirea erorii am ţinut seama de $\max\limits_{x \in [a,b]} \left| f^{(4)}(x) \right| = 14$. Facem observaţia că valoarea exactă a integralei este $I = 0.9333736875$ $\square$.

**Exemplul -1.4** *Să se determine numărul minim de subintervale* $k$ *în care trebuie împărţit intervalul* $[0, 1]$, *pentru ca, folosind formula Simpson sumată, să putem calcula integrala*

$$I = \int_0^1 \frac{e^x}{x + 1} dx$$

*cu o eroare mai mică de* $10^{-5}$.

Avem $\sup\limits_{x \in [0,1]} |f^{(4)}(x)| = 9$. Din formula de evaluare a erorii rezultă că

$$k \geq \left[ \sqrt{\frac{10^5}{320}} \right] + 1 = 5. \square$$

## Formula de cuadratură Hermite

Presupunem că funcţia $f$ este derivabilă în punctele $a$ şi $b$. Formula de cuadratură Hermite este

$$\sigma(f) = \int_a^b P(f; a, a, b, b; x)\, dx.$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\left(f;a,a,b,b;x\right)=$$

$$=f\left(a\right)+f(a,a)(x-a)+f(a,a,b)(x-a)^2+f(a,a,b,b)(x-a)^2(x-b).$$

Atunci

$$\sigma(f)=\int_a^b(f\left(a\right)+f(a,a)(x-a)+$$

$$+f(a,a,b)(x-a)^2+f(a,a,b,b)(x-a)^2(x-b))dx=$$

$$=(b-a)f\left(a\right)+f'(a)\cdot\frac{(b-a)^2}{2}+\frac{\frac{f(b)-f(a)}{b-a}-f'(a)}{b-a}\cdot\frac{(b-a)^3}{12}+$$

$$-\frac{f'(b)-2\frac{f(b)-f(a)}{b-a}+f'(a)}{(b-a)^2}\cdot\frac{(b-a)^4}{12}=$$

$$=\frac{b-a}{2}(f(a)+f(b))+\frac{(b-a)^2}{12}(f'(a)-f'(b)).$$

Dacă $f\in\mathcal{C}^4([a,b])$, din formula de evaluare a erorii la interpolare avem

$$\left|f(x)-P\left(f;a,a,b,b;x\right)\right|\leq\frac{1}{24}\left|(x-a)^2(x-b)^2\right|\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|,$$

$\forall\ x\in[a,b]$. Atunci

$$\left|\int_a^b f(x)dx-\sigma(f)\right|=\left|\int_a^b\left(f(x)-P\left(f;a,a,b.b;x\right)\right)dx\right|\leq$$

$$\leq\int_a^b\left|f(x)-P\left(f;a,a,b,b;x\right)\right|dx\leq$$

$$\leq\frac{1}{24}\max_{x\in[a,b]}\left|f^{(4)}(x)\right|\cdot\int_a^b\left|(x-a)^2(x-b)^2\right|dx=$$

$$=\frac{1}{24}\max_{x\in[a,b]}\left|f^{(4)}(x)\right|\cdot\frac{(b-a)^5}{30}=\frac{(b-a)^5}{720}\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|.$$

Formula de cuadratură Hermite sumată este:

$$\sigma_H^{(k)}(f) = \sum_{i=0}^{k-1} \frac{x_{i+1} - x_1}{2}(f\,(x_{i+1}) +$$

$$+f(x_i)) + \frac{(x_{i+1} - x_i)^2}{12}(f\prime(x_{i+1}) - f'(x_i)) =$$

$$= \frac{b-a}{2k}\sum_{i=0}^{k-1}(f\,(x_{i+1}) + f(x_i)) + \frac{(b-a)^2}{12k^2}(f'(a) - f'(b)),$$

pentru care avem evaluarea erorii:

$$\left| \int_a^b f(x)dx - \sigma_H^{(k)}(f) \right| \leq \sum_{i=0}^{k-1} \left| \int_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f) \right| \leq \frac{(b-a)^5}{720k^4} \cdot \max_{x\in[a,b]} \left| f^{(4)}(x) \right|,$$

unde $\sigma_i(f)$ este formula de cuadratură Hermite, aplicată funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\ i \in \overline{0, k-1}$.

**Exemplul -1.5** *Folosind formula de cuadratură Hermite sumată, să se aproximeze calculul integralei*

$$I = \int_1^2 \frac{1 + \ln x}{x} dx$$

*pentru* $k = 5,\ k = 10,\ k = 50$ *şi* $k = 100,$ *evaluând şi eroarea comisă.*

Se obţin următoarele rezultate:

| $k$ | $\sigma_D^{(k)}(f)$ | eroarea |
|---|---|---|
| 5 | $0,9333843739$ | $0,000002222$ |
| 10 | $0,9333743697$ | $0,000000238$ |
| 50 | $0,9333736876$ | $0,0000000002$ |
| 100 | $0,9333736875$ | $0,0000000000$ |

În stabilirea erorii am ţinut seama de $\max_{x\in[a,b]} \left| f^{(4)}(x) \right| = 14$. Facem observaţia că valoarea exactă a integralei este $I = 0,9333736875$ $\square$.

**Exemplul -1.6** *Să se determine numărul minim de subintervale $k$ în care trebuie împărţit intervalul $[0,1]$, pentru ca, folosind formula Hermite sumată, să putem calcula integrala*

$$I=\int_0^1 \frac{e^x}{x+1}dx,$$

*cu o eroare mai mică de $10^{-5}$.*

Avem $\sup\limits_{x\in[0,1]} |f^{(4)}(x)| = 9$. Din formula de evaluare a erorii rezultă că

$$k \geq \left[\sqrt{\frac{10^5}{80}}\right]+1=6. \;\square$$

## Formula de cuadratură Newton

Formula de cuadratură Newton este

$$\sigma(f)=\int_a^b P\left(f;a,\frac{2a+b}{3},\frac{a+2b}{3},b;x\right)dx.$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\left(f;a,\frac{2a+b}{3},\frac{a+2b}{3},b;x\right)=f\left(a\right)+f\left(a,\frac{2a+b}{3}\right)(x-a)$$

$$+f\left(a,\frac{2a+b}{3},\frac{a+2b}{3}\right)(x-a)\left(x-\frac{2a+b}{3}\right)+$$

$$+f\left(a,\frac{2a+b}{3},\frac{a+2b}{3},b\right)(x-a)\left(x-\frac{2a+b}{3}\right)\left(x-\frac{a+2b}{3}\right)$$

Atunci

$$\sigma(f)=(b-a)f\left(a\right)+3\cdot\frac{f\left(\frac{2a+b}{3}\right)-f(a)}{b-a}\cdot\frac{(b-a)^2}{2}+$$

$$+\frac{9}{2}\cdot\frac{f\left(\frac{a+2b}{3}\right)-2f\left(\frac{2a+b}{3}\right)+f(a)}{(b-a)^2}\cdot\frac{(b-a)^3}{6}+$$

$$+\frac{9}{2}\cdot\frac{f(b)-3f\left(\frac{a+2b}{3}\right)+3f\left(\frac{2a+b}{3}\right)-f(a)}{(b-a)^3}\cdot\frac{(b-a)^4}{36}=$$

$$=\frac{b-a}{8}\left(f(a)+3f\left(\frac{a+2b}{3}\right)+3f\left(\frac{2a+b}{3}\right)+f(b)\right).$$

Dacă $f\in\mathcal{C}^4([a,b])$, din formula de evaluare a erorii la interpolare avem

$$\left|f(x)-P\left(f;a,\frac{2a+b}{3},\frac{a+2b}{3},b;x\right)\right|\le$$

$$\le\frac{1}{24}\left|(x-a)\left(x-\frac{2a+b}{3}\right)\left(x-\frac{a+2b}{3}\right)(x-b)\right|\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|,$$

$\forall\ x\in[a,b]$. Atunci

$$\left|\int_a^b f(x)dx-\sigma(f)\right|=\left|\int_a^b\left(f(x)-P\left(f;a,\frac{2a+b}{3},\frac{a+2b}{3},b;x\right)\right)dx\right|\le$$

$$\le\int_a^b\left|f(x)-P\left(f;a,\frac{2a+b}{3},\frac{a+2b}{3},b;x\right)\right|dx\le$$

$$\frac{1}{24}\max_{x\in[a,b]}\left|f^{(4)}(x)\right|\cdot\int_a^b\left|(x-a)\left(x-\frac{2a+b}{3}\right)\left(x-\frac{a+2b}{3}\right)(x-b)\right|dx=$$

$$=\frac{1}{24}\max_{x\in[a,b]}\left|f^{(4)}(x)\right|\cdot\frac{(b-a)^5}{270}=\frac{(b-a)^5}{6480}\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|.$$

Formula de cuadratură Newton sumată este: $\sigma_N^{(k)}(f)=$

$$\sum_{i=0}^{k-1}\frac{x_{i+1}-x_1}{8}\left(f\left(x_{i+1}\right)+3f\left(\frac{2x_{i+1}+x_i}{3}\right)+3f\left(\frac{x_{i+1}+2x_i}{3}\right)+f(x_i)\right)=$$

$$=\frac{b-a}{8k}\sum_{i=0}^{k-1}\left(f\left(x_{i+1}\right)+3f\left(\frac{2x_{i+1}+x_i}{3}\right)+3f\left(\frac{x_{i+1}+2x_i}{3}\right)+f(x_i)\right),$$

pentru care, dacă $f\in\mathcal{C}^4([a,b])$, avem evaluarea erorii:

$$\left|\int_a^b f(x)dx-\sigma_N^{(k)}(f)\right|\le\sum_{i=0}^{k-1}\left|\int_{x_i}^{x_{i+1}}f(x)dx-\sigma_i(f)\right|\le\frac{(b-a)^5}{6480k^4}\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|,$$

unde $\sigma_i(f)$ este formula de cuadratură Newton, aplicată funcţiei $f$ pe intervalul $[x_i,x_{i+1}]$, $\forall\ i\in\overline{0,k-1}$.

**Exemplul -1.7** *Folosind formula de cuadratură Newton sumată, să se aproximeze calculul integralei*

$$I = \int_1^2 \frac{1 + \ln x}{x} dx$$

*pentru $k = 5$, $k = 10$ şi $k = 50$, evaluând şi eroarea comisă.*

Se obţin următoarele rezultate:

| $k$ | $\sigma_D^{(k)}(f)$ | eroarea |
|---|---|---|
| 5 | $0,9333725034$ | $0,000000249$ |
| 10 | $0,9333736118$ | $0,000000015$ |
| 50 | $0,9333736875$ | $0,0000000000$ |

În stabilirea erorii am ţinut seama de $\max_{x \in [a,b]} \left| f^{(4)}(x) \right| = 1$. Facem observaţia că valoarea exactă a integralei este $I = 0,9333736875$ $\square$.

**Exemplul -1.8** *Să se determine numărul minim de subintervale $k$, în care trebuie împărţit intervalul $[0, 1]$, pentru ca, folosind formula Newton sumată, să putem calcula integrala*

$$I = \int_0^1 \frac{e^x}{x + 1} dx,$$

*cu o eroare mai mică de $10^{-5}$.*

Avem $\sup_{x \in [0,1]} |f^{(4)}(x)| = 9$. Din formula de evaluare a erorii rezultă că

$$k \geq \left[ \sqrt{\frac{10^5}{720}} \right] + 1 = 4. \ \square$$

# Cuadraturi deschise

## Formula de cuadratură a dreptunghiului

Formula de cuadratură a dreptunghiului este

$$\sigma(f) = \int_a^b P\left(f; \frac{a+b}{2}; x\right) dx \ .$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\left(f; \frac{a+b}{2}; x\right) = f\left(\frac{a+b}{2}\right)$$

Atunci

$$\sigma(f) = \int_a^b f\left(\frac{a+b}{2}\right) .dx = (b-a) f\left(\frac{a+b}{2}\right)$$

Dacă $f \in \mathcal{C}^1([a,b])$, din formula de evaluare a erorii la interpolare, avem:

$$\left| f(x) - P\left(f; \frac{a+b}{2}; x\right) \right| \leq \left| x - \frac{a+b}{2} \right| \cdot \max_{x\in[a,b]} |f'(x)| \, , \ \forall \ x \in [a,b].$$

Atunci

$$\left| \int_a^b f(x)dx - \sigma(f) \right| = \left| \int_a^b \left( f(x) - P\left(f; \frac{a+b}{2}; x\right) \right) dx \right| \leq$$

$$\leq \int_a^b \left| f(x) - P\left(f; \frac{a+b}{2}; x\right) \right| dx \leq \max_{x\in[a,b]} |f'(x)| \cdot \int_a^b \left| x - \frac{a+b}{2} \right| dx =$$

$$= \frac{(b-a)^2}{4} \cdot \max_{x\in[a,b]} |f'(x)| \, .$$

**Observaţie:** Dacă $f$ este derivabilă în punctul $(a+b)/2$, atunci formula de cuadratură a dreptunghiului se poate obţine şi cu formula

$$\sigma(f) = \int_a^b P\left(f; \frac{a+b}{2}, \frac{a+b}{2}; x\right) dx.$$

În acest caz, dacă $f \in \mathcal{C}^1([a,b])$, din formula de evaluare a erorii la interpolare, avem

$$\left| f(x) - P\left(f; \frac{a+b}{2}, \frac{a+b}{2}; x\right) \right| \leq \frac{1}{2} \left| x - \frac{a+b}{2} \right|^2 \cdot \max_{x\in[a,b]} |f''(x)| \, , \ \forall \ x \in [a,b].$$

Atunci

$$\left|\int_a^b f(x)dx - \sigma(f)\right| = \left|\int_a^b \left(f(x) - P\left(f; \frac{a+b}{2}, \frac{a+b}{2}; x\right)\right) dx\right| \leq$$

$$\leq \int_a^b \left|f(x) - P\left(f; \frac{a+b}{2}, \frac{a+b}{2}; x\right)\right| dx \leq \max_{x\in[a,b]} |f''(x)| \cdot \int_a^b \left|x - \frac{a+b}{2}\right|^2 dx =$$

$$= \frac{(b-a)^3}{24} \cdot \max_{x\in[a,b]} |f''(x)|.$$

Formula de cuadratură a dreptunghiului sumată este:

$$\sigma_D^{(k)}(f) = \sum_{i=0}^{k-1} (x_{i+1} - x_i) f\left(\frac{x_i + x_{i+1}}{2}\right) = \frac{b-a}{k} \sum_{i=0}^{k-1} f\left(\frac{x_i + x_{i+1}}{2}\right),$$

pentru care, dacă $f \in \mathcal{C}^1([a,b])$, avem evaluarea erorii:

$$\left|\int_a^b f(x)dx - \sigma_D^{(k)}(f)\right| \leq \sum_{i=0}^{k-1} \left|\int_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f)\right| \leq \frac{(b-a)^2}{4k} \cdot \max_{x\in[a,b]} |f'(x)|,$$

iar dacă $f \in \mathcal{C}^2([a,b])$, avem evaluarea erorii:

$$\left|\int_a^b f(x)dx - \sigma_D^{(k)}(f)\right| \leq \sum_{i=0}^{k-1} \left|\int_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f)\right| \leq \frac{(b-a)^3}{24k} \cdot \max_{x\in[a,b]} \left|f''(x)\right|,$$

unde $\sigma_i(f)$ este formula de cuadratură a dreptunghiului, aplicată funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\ i \in \overline{0, k-1}$.

**Exemplul -1.9** *Folosind formula de cuadratură a dreptunghiului sumată, să se aproximeze calculul integralei*

$$I = \int_1^2 \frac{1 + \ln x}{x} dx$$

*pentru $k = 5$, $k = 10$, $k = 50$ şi $k = 100$, evaluând şi eroarea comisă.*

Se obţin următoarele rezultate:

| $k$ | $\sigma_D^{(k)}(f)$ | eroarea |
|---|---|---|
| 5 | 0.9336531768 | 0.0016666666 |
| 10 | 0.9334452939 | 0.0004166666 |
| 50 | 0.9333765747 | 0.0000166666 |
| 100 | 0.9333744095 | 0.0000041666 |

În stabilirea erorii am ţinut seama de $\max\limits_{x\in[a,b]} |f''(x)| = 1$. Facem observaţia că valoarea exactă a integralei este $I = 0.9333736875$ $\square$.

**Exemplul -1.10** *Să se determine numărul minim de subintervale $k$ în care trebuie împărţit intervalul $[0,1]$, pentru ca, folosind formula dreptunghiului sumată, să putem calcula integrala*

$$I = \int_0^1 \frac{e^x}{x+1} dx$$

*cu o eroare mai mică de $10^{-5}$.*

Avem $\sup\limits_{x\in[0,1]} |f''(x)| = 1$. Din formula de evaluare a erorii rezultă că

$$k \geq \left[\sqrt{\frac{10^5}{24}}\right] + 1 = 65 \,\square.$$

n=2

Formula de cuadratură deschisă pentru $n = 2$ este

$$\sigma(f) = \int_a^b P\left(f; \frac{2a+b}{3}, \frac{a+2b}{3}; x\right) dx \,.$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\left(f; \frac{2a+b}{3}, \frac{a+2b}{3}; x\right) = f\left(\frac{2a+b}{3}\right) +$$

$$+\frac{3}{b-a}\left(f\left(\frac{a+2b}{3}\right)-f\left(\frac{2a+b}{3}\right)\right)\left(x-\frac{2a+b}{3}\right)$$

Atunci

$$\sigma(f)=\int_a^b f\left(\frac{2a+b}{2}\right)dx+\frac{3}{b-a}\left(f\left(\frac{a+2b}{3}\right)-f\left(\frac{2a+b}{3}\right)\right)\int_a^b\left(x-\frac{2a+b}{3}\right)dx=$$

$$=(b-a)f\left(\frac{2a+b}{3}\right)+\frac{b-a}{2}\left(f\left(\frac{a+2b}{3}\right)-f\left(\frac{2a+b}{3}\right)\right)=$$

$$=\frac{b-a}{2}\left(f\left(\frac{2a+b}{3}\right)+f\left(\frac{a+2b}{3}\right)\right)$$

Dacă $f\in\mathcal{C}^2([a,b])$, din formula de evaluare a erorii la interpolare, avem:

$$\left|f(x)-P\left(f;\frac{2a+b}{3},\frac{a+2b}{3};x\right)\right|\le$$

$$\le\frac{1}{2}\cdot\left|\left(x-\frac{2a+b}{3}\right)\left(x-\frac{a+2b}{3}\right)\right|\cdot\max_{x\in[a,b]}\left|f''(x)\right|,\forall x\in[a,b].$$

Atunci

$$\left|\int_a^b f(x)dx-\sigma(f)\right|=\left|\int_a^b\left(f(x)-P\left(f;\frac{2a+b}{3},\frac{a+2b}{3};x\right)\right)dx\right|\le$$

$$\le\int_a^b\left|f(x)-P\left(f;\frac{2a+b}{3},\frac{a+2b}{3};x\right)\right|dx\le$$

$$\le\frac{1}{2}\cdot\max_{x\in[a,b]}\left|f''(x)\right|\cdot\int_a^b\left|\left(x-\frac{2a+b}{3}\right)\left(x-\frac{a+2b}{3}\right)\right|dx=$$

$$=\frac{11(b-a)^3}{324}\cdot\max_{x\in[a,b]}\left|f''(x)\right|.$$

Formula de cuadratură deschisă sumată pentru $n=2$ este:

$$\sigma_D^{(k)}(f)=\sum_{i=0}^{k-1}\frac{x_{i+1}-x_i}{2}\left(f\left(\frac{2x_i+x_{i+1}}{3}\right)+f\left(\frac{x_i+2x_{i+1}}{3}\right)\right)=$$

$$= \frac{b-a}{2k}\sum_{i=0}^{k-1}\left(f\left(\frac{2x_i+x_{i+1}}{3}\right)+f\left(\frac{x_i+2x_{i+1}}{3}\right)\right),$$

pentru care, dacă $f \in \mathcal{C}^2([a,b])$, avem evaluarea erorii:

$$\left|\int_a^b f(x)dx - \sigma_D^{(k)}(f)\right| \leq \sum_{i=0}^{k-1}\left|\int_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f)\right| \leq \frac{11(b-a)^3}{324k^2}\cdot \max_{x\in[a,b]}\left|f''(x)\right|,$$

unde $\sigma_i(f)$ este formula de cuadratură deschisa pentru $n=2$, aplicată funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\ i \in \overline{0, k-1}$.

## n=3

Formula de cuadratură deschisa pentru $n=3$ este

$$\sigma(f) = \int_a^b P\left(f; \frac{3a+b}{4}, \frac{a+b}{2}, \frac{a+3b}{4}; x\right)dx.$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\left(f; \frac{3a+b}{4}, \frac{a+b}{2}, \frac{a+3b}{4}; x\right) = f\left(\frac{3a+b}{4}\right)+$$

$$+\frac{4}{b-a}\left(f\left(\frac{a+b}{2}\right)-f\left(\frac{3a+b}{4}\right)\right)\left(x-\frac{3a+b}{4}\right)+$$

$$+\frac{8}{(b-a)^2}\left(f\left(\frac{a+3b}{4}\right)-2f\left(\frac{a+b}{2}\right)+f\left(\frac{3a+b}{4}\right)\right)\left(x-\frac{3a+b}{4}\right)\left(x-\frac{a+b}{2}\right)$$

Atunci

$$\sigma(f) = \int_a^b f\left(\frac{3a+b}{4}\right)dx + \frac{4}{b-a}\left(f\left(\frac{a+b}{2}\right)-f\left(\frac{3a+b}{4}\right)\right)\int_a^b\left(x-\frac{3a+b}{4}\right)dx+$$

$$+\frac{8}{(b-a)^2}\left(f\left(\frac{a+3b}{4}\right)-2f\left(\frac{a+b}{2}\right)+f\left(\frac{3a+b}{4}\right)\right)\cdot$$

$$\cdot\int_a^b\left(x-\frac{3a+b}{4}\right)\left(x-\frac{a+b}{2}\right)dx =$$

$$= (b-a) f\left(\frac{3a+b}{4}\right) + (b-a)\left(f\left(\frac{a+b}{2}\right) - f\left(\frac{3a+b}{4}\right)\right) +$$

$$+\frac{2}{3}(b-a)\left(f\left(\frac{a+3b}{4}\right) - 2f\left(\frac{a+b}{2}\right) + f\left(\frac{3a+b}{4}\right)\right) =$$

$$= \frac{(b-a)}{3}\left(2f\left(\frac{3a+b}{4}\right) - f\left(\frac{a+b}{2}\right) + 2f\left(\frac{a+3b}{4}\right)\right)$$

Dacă $f \in \mathcal{C}^3([a,b])$, din formula de evaluare a erorii la interpolare, avem

$$\left| f(x) - P\left(f; \frac{3a+b}{4}, \frac{a+b}{2}, \frac{a+3b}{4}; x\right)\right| \leq$$

$$\leq \frac{1}{6}\left|\left(x - \frac{3a+b}{4}\right)\left(x - \frac{a+b}{2}\right)\left(x - \frac{a+3b}{4}\right)\right| \cdot \max_{x\in[a,b]}\left|f^{(3)}(x)\right|, \ \forall\, x \in [a,b].$$

Atunci

$$\left|\int_a^b f(x)dx - \sigma(f)\right| = \left|\int_a^b \left(f(x) - P\left(f; \frac{3a+b}{4}, \frac{a+b}{2}, \frac{a+3b}{4}; x\right)\right) dx\right| \leq$$

$$\leq \int_a^b \left| f(x) - P\left(f; \frac{3a+b}{4}, \frac{a+b}{2}, \frac{a+3b}{4}; x\right)\right| dx \leq$$

$$\leq \frac{1}{6} \cdot \max_{x\in[a,b]}\left|f^{(3)}(x)\right| \cdot \int_a^b \left|\left(x - \frac{3a+b}{4}\right)\left(x - \frac{a+b}{2}\right)\left(x - \frac{a+3b}{4}\right)\right| dx =$$

$$= \frac{1}{6} \cdot \max_{x\in[a,b]}\left|f^{(3)}(x)\right| \cdot \frac{5(b-a)^4}{256} = \frac{5(b-a)^4}{1536} \cdot \max_{x\in[a,b]}\left|f^{(3)}(x)\right|$$

Formula de cuadratură deschisa sumată pentru $n = 3$ este:

$$\sigma_T^{(k)}(f) = \frac{1}{3}\sum_{i=0}^{k-1}(x_{i+1} - x_i)\left(2f\left(\frac{3x_i + x_{i+1}}{4}\right) - f\left(\frac{x_i + x_{i+1}}{2}\right) + 2f\left(\frac{3x_i + x_{i+1}}{4}\right)\right) =$$

$$= \frac{b-a}{3k}\sum_{i=0}^{k-1}\left(2f\left(\frac{3x_i + x_{i+1}}{4}\right) - f\left(\frac{x_i + x_{i+1}}{2}\right) + 2f\left(\frac{3x_i + x_{i+1}}{4}\right)\right),$$

pentru care, dacă $f \in \mathcal{C}^3([a,b])$, avem evaluarea erorii:

$$\left|\int_a^b f(x)dx - \sigma_T^{(k)}(f)\right| \leq \sum_{i=0}^{k-1}\left|\int_{x_i}^{x_{i+1}} f(x)dx - \sigma_i(f)\right| \leq \frac{5(b-a)^4}{1536k^3} \cdot \max_{x\in[a,b]}\left|f^{(3)}(x)\right|,$$

unde $\sigma_i(f)$ este formula de cuadratură deschisa pentru $n = 3$, aplicată funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\ i \in \overline{0, k-1}$

## n=4

Formula de cuadratură deschisa pentru $n = 4$ este

$$\sigma(f) = \int_a^b P\left(f; \frac{4a+b}{5}, \frac{3a+2b}{5}, \frac{2a+3b}{5}, \frac{a+4b}{5}; x\right) dx.$$

Din formula lui Newton de reprezentare a polinomului de interpolare avem

$$P\left(f; \frac{4a+b}{5}, \frac{3a+2b}{5}, \frac{2a+3b}{5}, \frac{a+4b}{5}; x\right) = f\left(\frac{4a+b}{5}\right) +$$

$$+\frac{5}{b-a}\left(f\left(\frac{3a+2b}{5}\right) - f\left(\frac{4a+b}{5}\right)\right)\left(x - \frac{4a+b}{5}\right) +$$

$$+\frac{25}{2(b-a)^2}\left(f\left(\frac{2a+3b}{5}\right) - 2f\left(\frac{3a+2b}{5}\right) + f\left(\frac{4a+b}{5}\right)\right)\cdot$$

$$\cdot\left(x - \frac{4a+b}{5}\right)\left(x - \frac{3a+2b}{5}\right) +$$

$$+\frac{125}{6(b-a)^3}\left(f\left(\frac{a+4b}{5}\right) - 3f\left(\frac{2a+3b}{5}\right) + 3f\left(\frac{3a+2b}{5}\right) - f\left(\frac{4a+b}{5}\right)\right)\cdot$$

$$\cdot\left(x - \frac{4a+b}{5}\right)\left(x - \frac{3a+2b}{5}\right)\left(x - \frac{2a+3b}{5}\right)$$

Atunci

$$\sigma(f) = \int_a^b f\left(\frac{4a+b}{5}\right) dx + \frac{5}{b-a}\left(f\left(\frac{3a+2b}{5}\right) - f\left(\frac{4a+b}{5}\right)\right)\int_a^b \left(x - \frac{4a+b}{5}\right) dx +$$

$$+\frac{25}{2(b-a)^2}\left(f\left(\frac{2a+3b}{5}\right) - 2f\left(\frac{3a+2b}{5}\right) + f\left(\frac{4a+b}{5}\right)\right)\cdot$$

$$\cdot\int_a^b \left(x - \frac{4a+b}{5}\right)\left(x - \frac{3a+2b}{5}\right) dx +$$

$$+\frac{125}{6(b-a)^3}\left(f\left(\frac{a+4b}{5}\right)-3f\left(\frac{2a+3b}{5}\right)+3f\left(\frac{3a+2b}{5}\right)-f\left(\frac{4a+b}{5}\right)\right)\cdot$$

$$\cdot\int_a^b\left(x-\frac{4a+b}{5}\right)\left(x-\frac{3a+2b}{5}\right)\left(x-\frac{2a+3b}{5}\right)dx=$$

$$=(b-a)\int_a^b f\left(\frac{4a+b}{5}\right)+\frac{3}{2}(b-a)\left(f\left(\frac{3a+2b}{5}\right)-f\left(\frac{4a+b}{5}\right)\right)+$$

$$+\frac{17}{12}(b-a)\left(f\left(\frac{2a+3b}{5}\right)-2f\left(\frac{3a+2b}{5}\right)+f\left(\frac{4a+b}{5}\right)\right)+$$

$$+\frac{11}{24}(b-a)\left(f\left(\frac{a+4b}{5}\right)-3f\left(\frac{2a+3b}{5}\right)+3f\left(\frac{3a+2b}{5}\right)-f\left(\frac{4a+b}{5}\right)\right)=$$

$$=\frac{(b-a)}{24}\left(11f\left(\frac{4a+b}{5}\right)+f\left(\frac{3a+2b}{5}\right)+f\left(\frac{2a+3b}{5}\right)+11f\left(\frac{a+4b}{5}\right)\right)$$

Dacă $f\in\mathcal{C}^4([a,b])$, din formula de evaluare a erorii la interpolare, avem

$$\left|f(x)-P\left(f;\frac{4a+b}{5},\frac{3a+2b}{5},\frac{2a+3b}{5},\frac{a+4b}{5};x\right)\right|\leq$$

$$\leq\frac{1}{24}\left|\left(x-\frac{4a+b}{5}\right)\left(x-\frac{3a+2b}{5}\right)\left(x-\frac{2a+3b}{5}\right)\left(x-\frac{a+4b}{5}\right)\right|\cdot$$

$$\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|,\ \forall\, x\in[a,b].$$

Atunci

$$\left|\int_a^b f(x)dx-\sigma(f)\right|=\left|\int_a^b\left(f(x)-P\left(f;\frac{4a+b}{5},\frac{3a+2b}{5},\frac{2a+3b}{5},\frac{a+4b}{5};x\right)\right)dx\right|\leq$$

$$\leq\int_a^b\left|f(x)-P\left(f;\frac{4a+b}{5},\frac{3a+2b}{5},\frac{2a+3b}{5},\frac{a+4b}{5};x\right)\right|dx\leq\leq\frac{1}{24}\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|\cdot$$

$$\cdot\int_a^b\left|\left(x-\frac{4a+b}{5}\right)\left(x-\frac{3a+2b}{5}\right)\left(x-\frac{2a+3b}{5}\right)\left(x-\frac{a+4b}{5}\right)\right|dx=$$

$$=\frac{1}{24}\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|\cdot\frac{551(b-a)^4}{93750}=\frac{551(b-a)^4}{2250000}\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|$$

Formula de cuadratură deschisa sumată pentru $n=4$ este:

$$\sigma_T^{(k)}(f)=\frac{1}{24}\sum_{i=0}^{k-1}(x_{i+1}-x_i)\cdot$$

$$\cdot\left(11f\left(\frac{4x_i+x_{i+1}}{5}\right)+f\left(\frac{3x_i+2x_{i+1}}{5}\right)+f\left(\frac{2x_i+x_{i+1}}{5}\right)+11f\left(\frac{x_i+4x_{i+1}}{5}\right)\right)=$$

$$=\frac{b-a}{24k}\sum_{i=0}^{k-1}\left(11f\left(\frac{4x_i+x_{i+1}}{5}\right)+f\left(\frac{3x_i+2x_{i+1}}{5}\right)+f\left(\frac{2x_i+x_{i+1}}{5}\right)+11f\left(\frac{x_i+4x_{i+1}}{5}\right)\right),$$

pentru care, dacă $f\in\mathcal{C}^4([a,b])$, avem evaluarea erorii:

$$\left|\int_a^b f(x)dx-\sigma_T^{(k)}(f)\right|\le\sum_{i=0}^{k-1}\left|\int_{x_i}^{x_{i+1}} f(x)dx-\sigma_i(f)\right|\le\frac{551(b-a)^5}{2250000k^4}\cdot\max_{x\in[a,b]}\left|f^{(4)}(x)\right|,$$

unde $\sigma_i(f)$ este formula de cuadratură deschisa pentru $n=4$, aplicată funcţiei $f$ pe intervalul $[x_i,x_{i+1}]$, $\forall\ i\in\overline{0,k-1}$

## Formule de cuadratură Gauss

Fie $[a,b]\subset\mathbf{R}$ şi $n\in\mathbf{N}$. Notăm cu

$$L_n(x):=\frac{n!}{(2n)!}\cdot\left((x-a)^n\,(x-b)^n\right)^{(n)},\ \forall\ x\in[a,b]$$

($L_n$ se numeşte polinom Legendre de grad $n$).

a) $L_n$ este un polinim de grad $n$.

b) $L_n$ are $n$ rădăcini distincte $x_1,x_2,\ldots,x_n$ situate în intervalul $[a,b]$.

c) $\int_a^b L_n(x)C(x)dx=0$, pentru orice polinom $C$ de grad $\le n-1$.

Fie

$$Q_i(x)=\prod_{j=1,\ j\ne i}^{n}\frac{x-x_j}{x_i-x_j}\text{ şi }\alpha_i=\int_a^b Q_i(x)dx,\ \forall\ i=\overline{1,n}.$$

Cu notaţiile anterioare, fie $f:[a,b]\to\mathbf{R}$ o funcţie continuă. Numim **formulă de cuadratură a lui Gauss de ordinul** $n$**, ataşată funcţiei** $f$, expresia

$$\sigma_G^n(f):=\sum_{i=1}^{n}\alpha_i f(x_i).$$

Pentru orice polinom $P$ de grad$\leq 2n-1$, avem

$$\int_a^b P(x)dx = \sum_{i=1}^n \alpha_i P(x_i),$$

(adică formula de cuadratură a lui Gauss este exactă de ordinul $2n-1$).

Dacă $f \in \mathcal{C}^{2n}([a,b])$ atunci este adevărată următoarea formulă de evaluare a erorii:

$$\left| \sigma_G^n(f) - \int_a^b f(x)dx \right| \leq \frac{(n!)^4}{((2n)!)^3} \cdot \frac{(b-a)^{2n+1}}{2n+1} \cdot \sup_{\xi\in[a,b]} \left| f^{(2n)}(\xi) \right|.$$

**Cazuri particulare**

**I.** $n=2$, $a=-1$, $b=1$

În acest caz,

$$L_2(x) = \frac{2!}{4!}\left(\left(x^2-1\right)^2\right)'' = \frac{1}{12}\left(12x^2-4\right) = x^2 - \frac{1}{3},$$

care are rădăcinile

$$x_1 = -\sqrt{\frac{1}{3}}, \mathrm{d}x_2 = \sqrt{\frac{1}{3}},$$

iar

$$\alpha_1 = \int_{-1}^{1} \frac{x-x_2}{x_1-x_2}dx = 1 \text{ şi } \alpha_2 = \int_a^b \frac{x-x_1}{x_2-x_1}dx = 1.$$

Deci, pentru $f$ o funcţie continuă pe $[-1,1]$, formula de cuadratură Gauss de ordinul 2 este

$$\sigma(f) = f\left(-\sqrt{\frac{1}{3}}\right) + f\left(\sqrt{\frac{1}{3}}\right).$$

Din formula de evaluare a erorii, în cazul în care $f$ este de 4 ori derivabilă, rezultă că

$$\left|\sigma(f) - \int_{-1}^{1} f(x)dx\right| \leq \frac{(2!)^4}{(4!)^3} \cdot \frac{2^5}{5} \cdot \sup_{\xi\in[a,b]} \left|f^{(4)}(\xi)\right| = \frac{1}{135} \cdot . \sup_{\xi\in[a,b]} \left|f^{(4)}(\xi)\right|.$$

**II.** $n=3$, $a=-1$, $b=1$

În acest caz,

$$L_3(x) = \frac{3!}{6!}\left(\left(x^2-1\right)^3\right)''' = \frac{1}{120}\left(120x^3 - 72x\right) = x^3 - \frac{3}{5}x,$$

care are rădăcinile

$$x_1 = -\sqrt{\frac{3}{5}},\ x_2 = 0,\ x_3 = \sqrt{\frac{3}{5}},$$

iar

$$\alpha_1 = \int_{-1}^{1} \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)}dx = \frac{5}{9},\ \alpha_2 = \int_a^b \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)}dx = \frac{8}{9},$$

$$\alpha_3 = \int_a^b \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_1)}dx = \frac{5}{9}.$$

Deci, pentru $f$ o funcţie continuă pe $[-1,1]$, formula de cuadratură Gauss de ordinul 3 este

$$\sigma(f) = \frac{1}{9}\left[5f\left(-\sqrt{\frac{3}{5}}\right) + 8f(0) + 5f\left(\sqrt{\frac{3}{5}}\right)\right].$$

Din formula de evaluare a erorii, în cazul în care $f$ este de 6 ori derivabilă, rezultă că

$$\left|\sigma(f) - \int_{-1}^{1} f(x)dx\right| \leq \frac{(3!)^4}{(6!)^3}\cdot\frac{2^7}{7}\cdot\sup_{\xi\in[a,b]}\left|f^{(6)}(\xi)\right| = \frac{1}{15750}\cdot.\sup_{\xi\in[a,b]}\left|f^{(6)}(\xi)\right|.$$

**Formule sumate:**

Fie $k \in N^*$, $h = (b-a)/k$, $y_i = a + ih$, $\forall\ i \in \overline{0,k}$. Dacă $\sigma_i(f)$ este o formulă de cuadratură Gauss a funcţiei $f$ pe intervalul $[x_i, x_{i+1}]$, $\forall\ i \in \overline{0,k-1}$, atunci se defineşte formula de cuadratură Gauss sumată

$$\sigma_G^{(k)}(f) = \sum_{i=0}^{k-1}\sigma_i(f).$$

- Formula lui Gauss sumată pentru două noduri:

  Pentru orice $j \in \{1,\ldots,k\}$ se consideră nodurile:

$$a_j = \frac{y_{j-1}+y_j}{2} - \frac{b-a}{2n\sqrt{3}},\quad b_j = \frac{y_{j-1}+y_j}{2} + \frac{b-a}{2n\sqrt{3}}.$$

Cuadratura este :

$$\sigma_G^{(k)} = \frac{h}{2} \cdot \sum_{j=1}^{n} (f(a_j) + f(b_j)),$$

pentru care (dacă $f \in \mathcal{C}^4([a,b])$) avem evaluarea erorii

$$\left| \sigma_G^{(k)} - \int_a^b f(x)dx \right| \leq \frac{7(b-a)^6}{2160k^5} \cdot \sup_{\xi \in [a,b]} \left| f^{(6)}(\xi) \right|.$$

- Formula lui Gauss sumată pentru trei noduri:

  Pentru orice $j \in \{1, \ldots, n\}$ se consideră nodurile:

$$a_j = \frac{y_{j-1} + y_j}{2} - \frac{(b-a)\sqrt{15}}{10n}, \quad b_j = \frac{y_{j-1} + y_j}{2},$$

$$c_j = \frac{y_{j-1} + y_j}{2} + \frac{(b-a)\sqrt{15}}{10n}.$$

  Cuadratura este:

$$\sigma_G^{(k)} = \frac{h}{18} \cdot \left[ \sum_{j=1}^{n} (5f(a_j) + 8f(b_j) + 5f(c_j)) \right],$$

  pentru care (dacă $f \in \mathcal{C}^6([a,b])$) avem evaluarea erorii

$$\left| \sigma_G^{(k)} - \int_a^b f(x)dx \right| \leq \frac{(b-a)^8}{2016000k^7} \cdot \sup_{\xi \in [a,b]} \left| f^{(6)}(\xi) \right|.$$

**Exemplul -1.11** *Folosind formulele de cuadratură ale lui Gauss sumate (cu două, respectiv cu trei noduri) să se aproximeze calculul integralei*

$$I = \int_1^2 \frac{1 + \ln x}{x} dx$$

*pentru $k = 5$, $k = 10$, $k = 50$ şi $k = 100$, evaluând şi eroarea comisă.*

Pentru formula de cuadratură Gauss sumată, relativă la două noduri, se obţin următoarele rezultate:

| $k$ | $\sigma_D^{(k)}(f)$ | eroarea |
|---|---|---|
| 5 | $0,9333754517$ | $0,000014518$ |
| 10 | $0,9333738009$ | $0,000000453$ |
| 50 | $0,9333736877$ | $0,0000000002$ |
| 100 | $0,9333736875$ | $0,0000000000$ |

În stabilirea erorii am ţinut seama de $\max\limits_{x\in[a,b]}\left|f^{(4)}(x)\right| = 14$.

Pentru formula de cuadratură Gauss sumată, relativă la trei noduri, se obţin următoarele rezultate:

| $k$ | $\sigma_D^{(k)}(f)$ | eroarea |
|---|---|---|
| 5 | $0,9333736921$ | $0,000000032$ |
| 10 | $0,9333736875$ | $0,000000000$ |

În stabilirea erorii am ţinut seama de $\max\limits_{x\in[a,b]}\left|f^{(6)}(x)\right| = 508$. Facem observaţia că valoarea exactă a integralei este $I = 0,9333736875$ $\square$.

# Metode numerice pentru ecuaţii diferenţiale

Studiul ecuaţiilor diferenţiale este o parte a matematicii, care face obiectul a numeroase cercetări şi care continuă să rămână în actualitate prin interesul particular pe care-l reprezintă pentru alte discipline, ca mecanica, astronomia, fizica si, mai recent, chimia, biologia, ştiinţe în care modelarea poate conduce la probleme matematice ce implică ecuaţii diferenţiale. De multe ori, chiar dacă rezultatele matematice ne permit să demostrăm existenţa şi unicitatea soluţiei problemei la care se ajunge, soluţia analitică nu poate fi obţinută, şi atunci se impune găsirea unei soluţii aproximative a problemei printr-o metodă de rezolvare numerică a problemei.

În acest capitol ne propunem să dăm o descriere a principalelor metode de rezolvare numerică a problemelor Cauchy (a problemelor cu condiţii iniţiale). Problema Cauchy se defineşte astfel: *dată fiind aplicaţia* $f : D \subset R \times R^m \longrightarrow R^m$, *care defineşte ecuaţia diferenţială:*

$$y'(t) = f(t, y(t)) \tag{1}$$

*şi punctul iniţial* $(t_0, y_0)$, *să se determine o soluţie* $y : I \subset R \longrightarrow R^m$ *a ecuaţiei diferenţiale (1) şi care verifică condiţia iniţială:*

$$y(t_0) = y_0. \tag{2}$$

Metodele pe care le vom prezenta în acest capitol se bazează pe aproximarea soluţiei problemei (1)-(2), definită pe un interval $[t_0, t_0 + T]$, printr-un şir finit de valori $(y_n)$ care aproximează soluţia $y(t)$ în nodurile $t_n = t_0 + nh$ , $n \in \{\overline{1, N}\}$ , adică $y_n \simeq y(t_n)$ . Se consideră că $N = \left[\frac{T}{h}\right]$ , unde $h$ este pasul, adică distanţa dintre două valori consecutive ale variabilei $t$, pentru care se face aproximarea.

Se cunosc două clase importante de metode numerice pentru rezolvarea problemei Cauchy:

1. **Metode într-un pas (metode directe)** sunt metodele în care $y_{n+1}$ se calculează printr-o relaţie de recurenţă în funcţie doar de $y_n$, calculat anterior. În această categorie intră metodele Euler, Euler îmbunătăţită, Euler modificată, Runge-Kutta.

2. **Metode în mai mulţi paşi (metode indirecte)** sunt metodele în care $y_{n+1}$ se calculează printr-o relaţie de recurenţă în funcţie de valorile precedente $y_n,\ y_{n-1}, \ldots,\ y_{n-r+1}$. În această categorie intră metodele Adams-Bashforth, Adams-Moulton, metoda predictor-corector, metoda diferenţialei retrograde.

# Metode într-un pas (directe)

Considerăm următoarea problemă Cauchy:

$$\begin{cases} y'(t) = f(t, y(t))\,, \quad \forall\, t \in I_0 = [t_0, t_0 + T] \\ y(t_0) = y_0\,. \end{cases} \tag{3}$$

O metoda numerică într-un pas, pentru aproximarea soluţiei acestei probleme, constă în determinarea aproximatiilor $(y_n)$ pentru $y(t_n)$ în $t_n = t_0 + nh$, cu $n = \overline{1, N}$, unde:

$$\begin{cases} y_0 = \alpha \in R^m\,, \\ y_{n+1} = y_n + h\Phi(t_n, y_n, h)\,. \end{cases} \tag{4}$$

Alegerea unei metode înseamnă să alegem funcţia $\Phi$.

**Observaţie**. Pasul $h$ se poate alege şi dependent de $n$ astfel încât $t_{n+1} = t_n + nh_n$ să nu depăşească $t_0 + T$. În acest caz, în definiţiile ce urmează se va considera $h = \max\limits_{n=\overline{0,N-1}} h_n$.

**Observaţie**. Noţiunea de stabilitate arată că o mică perturbare asupra datelor schemei numerice (4) produce o pertubare mică a aproximaţiei şi este independentă de $h$.

**Definiţia -1.1** *Spunem că metoda numerică (4) este de ordinul p, cu $p \in N^*$, dacă pentru orice y soluţie a problemei (3) avem:*

$$\max_{n=\overline{0,N-1}} \left\| \frac{1}{h} \left( y(t_{n+1}) - y(t_n) \right) - \Phi(t_n, y(t_n), h) \right\| = \mathcal{O}(h^p)\,.$$

## Metoda Euler explicită (progresivă)

Schema numerică pentru metoda Euler este următoarea:

$$\begin{cases} y_0 \in \mathbf{R}^m \,, \\ y_{n+1} = y_n + hf(t_n, y_n) \,. \end{cases}$$

**Interpretarea geometrică a metodei Euler**. Dacă s-a determinat $y_n$, pentru a determina $y_{n+1}$ se consideră soluţia ecuaţiei (1) care trece prin $(t_n, y_n)$, adică acea soluţie pentru care $y(t_n) = y_n$. Se duce tangenta la graficul acestei soluţii în $(t_n, y_n)$. Ecuaţia tangentei este:

$$y - y_n = f(t_n, y_n)(t - t_n) \,.$$

Pentru $t = t_{n+1}$ se oţine $y_{n+1} = y_n + (t_{n+1} - t_n)f(t_n, y_n)$, adică $y_{n+1}$ este intersecţia tangentei cu dreapta $t = t_{n+1}$. Din acest motiv, metoda Euler mai este numită şi metoda liniilor poligonale. Erorile se acumulează, ceea ce face ca, în această metodă, pentru pas $h$ nu prea mic, valorile numerice să fie mult diferite de valorile soluţiei exacte.

**Exemplul -1.1** *Folosind metoda Euler, determinaţi soluţia aproximativă a problemei Cauchy:*

$$\begin{cases} y' = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2} \,, \quad t \in [1, 2] \\ y(1) = 0.75 \end{cases}$$

*în $t = 2$, luând $h = 0.5$.*

Avem $f(t, y) = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}$ , $t_0 = 1$ şi $y_0 = 0.75$ . Se obţine $N = 2$ şi $y_{n+1} = y_n + 0.5f(t_n, y_n)$, pentru $n = 0, 1$ . Deci rezultă:

$$t_1 = 1.5 \,,\ y(t_1) \simeq y_1 = y_0 + 0,5f(t_0, y_0) = 0.75 + 0.5f(1, 0.75) =$$
$$0.75 + 0.5(-0.4375) = 0.5313 \,,$$
$$t_2 = 2 \,,\ y(t_2) \simeq y_2 = y_1 + 0,5f(t_1, y_1) = 0.5313 + 0.5f(1.5, 0.5313) =$$
$$0.5313 + 0.5(-0.1831) = 0.4397 \,.\square$$

Soluţia exactă a problemei din exerciţiul -1.1 este

$$y(t) = -\frac{t+5}{2t(t-5)}\,.$$

Valoarea exactă în $t = 2$ este $y(2) = 0.5833333333$, iar valoarea aproximativă, în funcţie de $h$, este dată în tabelul 1.

Table 1: *Date obtinute prin metoda Euler.*

| h | aproximarea pentru y(2) |
|---|---|
| 0.5 | 0.4397243924 |
| 0.2 | 0.5231702705 |
| 0.1 | 0.5526766962 |
| 0.05 | 0.5678404244 |

Se poate considera o metodă Euler implicită (retrogradă), în care schema numerică este:

$$\begin{cases} y_0 \in \mathbf{R}^m\,, \\ y_{n+1} = y_n + hf(t_{n+1}, y_{n+1})\,,\ n = \overline{0, N-1}\,. \end{cases}$$

Folosirea acestei metode este mai dificilă decât metoda Euler explicită, deoarece,la fiecare pas, calculul lui $y_{n+1}$ cu $y_n$ cunoscut, se face prin rezolvarea unei ecuaţii, în general, neliniare, şi care poate avea mai multe soluţii. Calculul soluţiei ecuaţiei din (**??**) în $y_{n+1}$ va necesita rezolvarea unui sistem liniar dacă $f$ este afină în $y$ sau, în cazul general, utilizarea unei metode pentru ecuaţii neliniare.

## Metode Runge-Kutta

Metodele Runge-Kutta se deosebesc de metoda Taylor prin faptul ca înlocuiesc calculul derivatelor lui $f$ prin evaluări pentru $f$ în puncte intermediare între $t_n$ şi $t_{n+1}$.

Metoda a fost introdusă de matematicianul german Carl David Runge în 1895 şi a fost dezvoltată de alt matematician german, Wilhelm Kutta, în 1901.

În definirea generală a metodelor Runge-Kutta, pentru $q \in \mathbf{N}^*$ se dau o listă de constante reale:

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ c_q & a_{q1} & \cdots & a_{qq} \\ \hline & b_1 & \cdots & b_q \end{array}$$

cu $c_1, \ldots, c_q$ nu neapărat distincte, şi formulele de cuadratură:

$$\int_0^{c_i} \psi(t)\,\mathrm{d}t \simeq \sum_{j=1}^{q} a_{ij}\,\psi(c_j)\,, \quad i = \overline{1,q}\,,$$
$$\int_0^{1} \psi(t)\,\mathrm{d}t \simeq \sum_{j=1}^{q} b_j\,\psi(c_j).$$

Vom considera cazul general, când punctele din $I_0$ nu sunt echidistante, adică $t_{n+1} - t_n = h_n$. Pentru fiecare $n = \overline{1,N}$, se consideră punctele intermediare

$$t_{n,i} = t_n + c_i h_n \text{ pentru } i = \overline{1,q}$$

şi determinăm $(y_{n,i})_{i=\overline{1,q}}$ prin rezolvarea sistemului:

$$y_{n,i} = y_n + h_n \sum_{j=1}^{q} a_{ij}\, f(t_{n,j}, y_{n,j})\,, \quad i = \overline{1,q}\,.$$

Aproximarea soluţiei în $t_{n+1}$ se determină din:

$$y_{n+1} = y_n + h_n \sum_{j=1}^{q} b_j\, f(t_{n,j}, y_{n,j})\,.$$

Relaţiile (**??**) şi (**??**) rezultă aplicând formulele de cuadratură (**??**). Fie $y$ o soluţie pentru (1). Folosind schimbarea de variabilă $u = \dfrac{t-t_n}{h_n}$, obţinem:

$$y(t_{n,i}) = y(t_n) + \int_{t_n}^{t_{n,i}} f(t, y(t))\mathrm{d}t =$$

$$y(t_n) + \int_0^{c_i} f(t_n + h_n u, y(t_n + h_n u))\mathrm{d}t\,, \quad i = \overline{1,q}$$

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t))\mathrm{d}t = y(t_n) + \int_0^1 f(t_n + h_n u, y(t_n + h_n u))\mathrm{d}t.$$

Aplicând formulele de cuadratură (**??**)se obţine:

$$y(t_{n,i}) \simeq y(t_n) + h_n \sum_{j=1}^{q} a_{ij}\ f(t_{n,j}, y(t_{n,j}))\,, \quad i = \overline{1,q}$$
$$y(t_{n+1}) \simeq y(t_n) + h_n \sum_{j=1}^{q} b_j\ f(t_{n,j}, y(t_{n,j}))\,,$$

de unde, pentru calculul valorilor aproximative avem (**??**) şi (**??**).

Schema numerică generală pentru metode Runge-Kutta se poate scrie sub forma:

$$y(t_{n+1}) = y(t_n) + h_n \Phi(t_n, y_n, h_n))\,, \quad n = \overline{0, N-1},$$

unde $\Phi(t,y,h) = \displaystyle\sum_{j=1}^{q} b_j\ f(t + c_j h, y_j)$ cu $y_1, \ldots, y_q$ rezultate prin rezolvarea sistemului:

$$y_i = y + h \sum_{j=1}^{q} a_{ij}\ f(t + c_j h, y_j)\ ,\ i = \overline{1,q}\,.$$

Notăm $A = (a_{ij})_{i,j=\overline{1,q}}$ , $C = \mathrm{diag}\,(c_1, c_2, \ldots, c_q)$, $b = (b_1, b_2, \ldots, b_q)^t$ şi $e = (1, 1, \ldots, 1)^t \in \mathbf{R}^q$.

În funcţie de coeficienţii $a_{ij}$ metodele Runge-Kutta pot fi clasificate astfel:

- dacă $a_{ij} = 0$, $\forall\ 1 \le i \le j \le q$, atunci metoda Runge-Kutta se numeşte *explicită.*

- dacă $a_{ij} = 0$, $\forall\ 1 \le i < j \le q$, atunci metoda Runge-Kutta se numeşte *semi-implicită.*

În cazul când nici una dintre condiţiile de mai sus nu este îndeplinită, metoda Runge-Kutta se numeşte *implicită*.

Dăm mai jos câteva rezultate importante în convergenţa metodelor Runge-Kutta.

**Teorema -1.1** *Dacă $f$ este lipschitziană în a doua varibilă, având constanta $L$, $\rho(|A|)$ este raza spectrală a lui $|A|$ şi $h_n$ este astfel încât $h_n L\rho(|A|) < 1$, atunci schema numerică Runge-Kutta (**??**) admite soluţie unică. De asemenea, pentru $h_n = h = const.$, astfel încât $0 < h \leq h^*$ cu $h^* L\rho(|A|) < 1$, metoda Runge-Kutta este stabilă.*

**Observaţie**. Pentru o metodă Runge-Kutta explicită avem $\rho(|A|) = 0$. Deci o metodă Runge-Kutta explicită întotdeauna este stabilă.

**Definiţia -1.2** *Vom spune că este îndeplinită condiţia $A(p)$ dacă sunt îndeplinite condiţiile $A'(l)$ pentru $l = \overline{1,p}$, unde $A'(l)$ sunt date în tabelul 2 pentru $l = \overline{1,5}$.*

**Teorema -1.2** *Condiţia necesară şi suficientă ca o metodă Runge-Kutta (pentru toate funcţiile suficient de regulate) să fie de ordin $p$ este să fie verificată condiţia $A(p)$.*

**Remarcă**. Condiţiile necesare pentru ca o metodă Runge-Kutta să fie de ordin $p$ sunt:

$$b^t e = 1 \;, \quad b^t C e = \frac{1}{2} \;, \ldots, b^t C^{p-1} e = \frac{1}{p} \;.$$

Condiţiile (**??**) sunt echivalente cu afirmaţia: formula de cuadratură $(\mathbf{??})_2$ este de ordin $p-1$, adică este exactă pentru polinoame până la gradul $p-1$. Se poate deduce că $p \leq 2q$.

Construcţia unei metode Runge-Kutta se face, în general, în modul următor:

- se determină relaţii pentru $b_1, \ldots, b_q$ şi $c_1, \ldots, c_q$ din condiţia ca formula de cuadratură $(\mathbf{??})_2$ să fie de ordin $p-1$;

- se obţin şi coeficienţii din matricea $A$, rezolvând restul de condiţii suficiente pentru metode Runge-Kutta, până la ordinul $p$, date în tabelul 3.

Pentru $p = 5$ se verifică şi condiţiile pentru $p = 3$ şi $p = 4$.

**Observaţie**. Pentru programarea metodelor Runge-Kutta este mai uşor să notăm:

$$k_{n,i} = f(t_{n,i}, y_{n,i}) \text{ pentru } i = \overline{1, q},$$

şi atunci avem:

$$k_{n,i} = f(t_{n,i}, y_n + h_n \sum_{j=1}^{q} a_{ij}\, k_{n,j}) \text{ pentru } i = \overline{1, q}$$

$$y_{n+1} = y_n + h_n \sum_{j=1}^{q} b_j\, k_{n,j}\,.$$

Rezolvând sistemul (**??**) în necunoscutele $(k_{n,i})_{i=\overline{1,q}}$ , calculăm mai uşor $y_{n+1}$ din (**??**).

De asemenea, obţinem:

$$y_{n,i} = y_n + h_n \sum_{j=1}^{q} a_{ij}\, k_{n,j}\,, \text{ pentru } i = \overline{1, q}\,.$$

**Definiţia -1.3** *O metodă* $RK_{pp'}$ *este o metodă Runge-Kutta de ordin* $p'$ *corespunzătoare tabloului de constante*

$$\begin{array}{c||ccc|c}
c_1 & a_{11} & \cdots & a_{1q} & 0 \\
\vdots & \vdots & \ddots & \vdots & \blacksquare \\
c_q & a_{q1} & \cdots & a_{qq} & 0 \\
\hline
1 & b_1 & \cdots & b_q & 0 \\
\hline\hline
 & b_1' & \cdots & b_q' & b_{q+1}' \\
 & & & &
\end{array}$$

*în care dacă excludem ultima linie şi ultima coloană, obţinem o metodă de ordin* $p$.

**Exemple particulare de metode Runge-Kutta.**

**Exemplul 1.** Pentru $q = 1$, metoda Runge-Kutta definită prin:

$$\begin{array}{c|c} \theta & \theta \\ \hline & 1 \end{array} \quad \text{cu } \theta \in \mathbf{R}\text{, se scrie:}$$

$$y_{n,1} = y_n + \theta f(t_n + \theta h_n, y_{n,1})$$
$$y_{n+1} = y_n + h_n f(t_n + \theta h_n, y_{n,1})\,.$$

Pentru $\theta = 0$ şi $\theta = 1$ se obţin metode de ordinul 1: **metoda Euler progresivă**, respectiv, **metoda Euler retrogradă**. Pentru $\theta = 1/2$ se obţine **metoda punctului de mijloc** şi este o metodă de ordinul 2.

**Exemplul 2.** Pentru $q = 2$, metoda Runge-Kutta definită prin:

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \alpha & \alpha & 0 \\ \hline & 1 - \dfrac{1}{2\alpha} & \dfrac{1}{2\alpha} \end{array}$$

cu $\alpha \in \mathbf{R}^*$, se scrie:

$$\begin{cases} y_{n,1} = y_n \\ \\ y_{n,2} = y_n + h_n\ \alpha\ f(t_n, y_{n,1}) \end{cases}$$

$$y_{n+1} = y_n + h_n \left( \left(1 - \frac{1}{2\alpha}\right) f(t_n, y_n) + \frac{1}{2\alpha} f\left(t_n + \alpha h_n, y_n + \alpha h_n f(t_n, y_n)\right)\right)$$

Dacă $\alpha = \dfrac{1}{2}$ , atunci se obţine **metoda Euler modificată (ameliorată)**:

$$y_{n+1} = y_n + h_n f\left(t_n + \frac{1}{2}\, h_n, y_n + \frac{1}{2}\, h_n f(t_n, y_n)\right).$$

Avem $A = \begin{pmatrix} 0 & 0 \\ 1/2 & 0 \end{pmatrix}$ , $C = \begin{pmatrix} 0 & 0 \\ 0 & 1/2 \end{pmatrix}$ , $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ şi este îndeplinită condiţia $A(2)$, deci metoda (**??**) este de ordin 2.

Dacă $\alpha = 1$, atunci se obţine **metoda Heun** sau **metoda Euler îmbunătăţită** sau **metoda** $RK_{12}$ :

$$y_{n+1} = y_n + \frac{1}{2}\, h_n\ \left(f(t_n, y_n) + f(t_n + h_n, y_n + h_n f(t_n, y_n))\right).$$

Avem $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ , $C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ , $b = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$ şi este îndeplinită condiţia $A(2)$, deci metoda (**??**) este de ordin 2.

**Exemplul 3.** Pentru $q = 4$ metoda Runge-Kutta definită prin:

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & 1/6 & 2/6 & 2/6 & 1/6 \\ & & & & \end{array}$$

este **metoda Runge-Kutta clasică** sau **metoda** $RK_{24}$ (Kutta - 1901). Sub această formă, metoda este cel mai des utilizată. Avem:

$$\begin{cases} y_{n,1} = y_n \\ y_{n,2} = y_n + \dfrac{h_n}{2} f(t_n, y_{n,1}) \\ y_{n,3} = y_n + \dfrac{h_n}{2} f(t_n + \dfrac{1}{2}\, h_n, y_{n,2}) \\ y_{n,4} = y_n + h_n f(t_n + \dfrac{1}{2}\, h_n, y_{n,3}) \end{cases}$$

si
$$\begin{aligned} y_{n+1} &= y_n + \frac{h_n}{6}\left( f(t_n, y_{n,1}) + 2f(t_n + \frac{1}{2}\, h_n, y_{n,2}) + \right. \\ &\left. 2f(t_n + \frac{1}{2}\, h_n, y_{n,3}) + f(t_{n+1}, y_{n,4})\right). \end{aligned} \tag{5}$$

În metoda $RK_{24}$ avem:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ o & 0 & 1 & 0 \end{pmatrix}, \ C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ b = \begin{pmatrix} 1/6 \\ 2/6 \\ 2/6 \\ 1/6 \end{pmatrix}$$ şi este îndeplinită condiţia $A(4)$, deci metoda (5) este de ordin 4.

**Exemplul 4.** Pentru $q = 5$, metoda Runge-Kutta definită prin:

$$\begin{array}{c|ccccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 2/7 & 2/7 & 0 & 0 & 0 & 0 \\ 4/7 & -8/35 & 4/5 & 0 & 0 & 0 \\ 6/7 & 29/42 & -2/3 & 5/6 & 0 & 0 \\ 1 & 1/6 & 1/6 & 5/12 & 1/4 & 0 \\ \hline & 11/96 & 28/96 & 35/96 & 14/96 & 8/96 \\ & & & & & \end{array}$$

este o metodă $RK_{34}$ .

**Exemplul 5.** Pentru $q = 6$, metoda Runge-Kutta definită prin:

$$\begin{array}{c|cccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 3/16 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 3/4 & 0 & -3/16 & 9/16 & 0 & 0 & 0 \\ 1 & 1/7 & 4/7 & 6/7 & -12/7 & 8/7 & 0 \\ \hline & 7/90 & 0 & 32/90 & 12/90 & 32/90 & 7/90 \\ & & & & & & \end{array}$$

este o metodă Runge-Kutta de ordin 6.

**Exemplul -1.2** *Se dă problema Cauchy:*

$$\begin{cases} y' = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}, & t \in [1, 2.9] \\ y(1) = 0.75 \end{cases}$$

*Să se determine valoarea aproximativă a soluţiei în $t = 2.9$, folosind metodele Euler progresivă, Euler modificată, $RK_{12}$ şi $RK_{24}$, pentru $N = 2$.*

Pentru $N = 2$ avem $h = 0.9500$. Luăm $y_0 = 0.75$, $t_0 = 1$. Pentru metoda $RK_{24}$, avem $f(t, y) = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}$. Se obţine:

- pentru $n = 1$:

$$y_{0,1} = y_0 = 1 \ , \quad t_{0,1} = t_0 = 1 \ ,$$

$$y_{0,2} = y_0 = 1 + 0.95 f(t_{0,1}, y_{0,1})/2 = 0.8812 \ , \quad t_{0,2} = t_0 + 0.95/2 = 1.4750 \ ,$$

$$y_{0,3} = y_0 = 1 + 0.95 f(t_{0,2}, y_{0,2})/2 = 1.0305 \ , \quad t_{0,3} = t_0 + 0.95/2 = 1.4750 \ ,$$

$$y_{0,4} = y_0 = 1 + 0.95 f(t_{0,3}, y_{0,3})/2 = 1.2360 \ , \quad t_{0,4} = t_0 + 0.95 = 1.9500 \ ,$$

$$y_1 = y_0 + 0.95(f(t_{0,1}, y_{0,1}) + 2f(t_{0,2}, y_{0,2}) + 2f(t_{0,3}, y_{0,3}) + f(t_{0,4}, y_{0,4}))/6 =$$

$$1 + 0.95(-0.2500 + 2 \cdot 0.0642 + 2 \cdot 0.2484 + 0.8281)/6 = 1.1905$$

$$t_1 = t_0 + h = 1.9500$$

- pentru $n = 2$:

$$y_{1,1} = y_1 = 1.1905 \ , \quad t_{1,1} = t_1 = 1.95 \ ,$$

$$y_{1,2} = y_1 = 1 + 0.95 f(t_{1,1}, y_{1,1})/2 = 1.5425 \ , \quad t_{1,2} = t_1 + 0.95/2 = 2.4250 \ ,$$

$$y_{1,3} = y_1 = 1 + 0.95 f(t_{1,2}, y_{1,2})/2 = 1.9984 \ , \quad t_{1,3} = t_1 + 0.95/2 = 2.4250 \ ,$$

$$y_{1,4} = y_1 = 1 + 0.95 f(t_{1,3}, y_{1,3})/2 = 4.1612 \ , \quad t_{1,4} = t_1 + 0.95 = 2.9 \ ,$$

$$y_2 = y_1 + 0.95(f(t_{1,1}, y_{1,1}) + 2f(t_{1,2}, y_{1,2}) + 2f(t_{1,3}, y_{1,3}) + f(t_{1,4}, y_{1,4}))/6 =$$

$$1.1905 + 0.95(0.7411 + 2 \cdot 1.7008 + 2 \cdot 3.1270 + 15.8509)/6 = 5.3464$$

$$t_2 = t_1 + h = 2.9 \ . \square$$

În tabelul 4 am dat valorile obţinute cu metodele cerute în enunţ pentru $N = 2$. Soluţia exactă este $y(t) = -\dfrac{t+3}{2t(t-3)}$ . Pentru aceleaşi cerinţe ca în exerciţiul -1.2, dar pentru $N = 7$ se obţin datele din tabelul 5. Pentru valori mai mari ale lui $N$ se obţin valori aproximative mult apropiate de soluţia exactă. De exemplu, pentru $N = 20$ valorile obţinute sunt în tabelul 6.

# Metode în mai mulţi paşi (indirecte)

Calculul valorii $y_{n+1}$ la $t_{n+1}$ se poate obţine folosind $r+1$ valori deja calculate $y_n, y_{n-1}, \ldots, y_{n-r+1}$, adică folosind $r+1$ paşi anteriori cu $r \in \mathbf{N}$. În general, se obţin metode cu caracteristici mai bune. Cel mai des întâlnite medote în $r+1$ paşi sunt metodele Adams. Se consideră problema (1)-(2), în care presupunem că $f$ este lipschitiziană în al doilea argument. Considerăm diviziunea:

$$t_0 < t_1, < \ldots, t_N = t_0 + T\,,$$

$$h_n = t_{n+1} - t_n\,.$$

Metodele Adams se bazează pe relaţia:

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t))\mathrm{d}t\,,$$

unde $y(\cdot)$ este soluţie a ecuaţiei (1).

Schema de aproximare într-o metodă Adams este:

$$y_{n+1} = y_n + \int_{t_n}^{t_{n+1}} P_{n,r}(t)\mathrm{d}t\,,$$

unde $P_{n,r}(\cdot)$ este o aproximare a funcţiei $f(\cdot,\cdot)$.

Se presupune că printr-o metodă directă (de exemplu, de tip Runge-Kutta) s-au determinat valorile $y_1, y_2, \ldots, y_r$ pornind de la o valoare $y_0$. Pentru $n \geq r$, $n \in \left\{\overline{1, N-1}\right\}$, presupunem cunoscute $y_n, y_{n-1}, \ldots, y_{n-r}$ şi notăm:

$$f_j = f(t_j, y_j)\,, \quad j = \overline{n-r, n}\,,$$

valorile aproximative pentru $f(t_j, y(t_j))$ , $j = \overline{n-r, n}$ , pentru $y(\cdot)$ soluţie a ecuaţiei (1).

## Metoda Adams-Bashforth

În metoda Adams-Bashforth considerăm, în (**??**), $P_{n,r}(t)$ polinomul de interpolare de grad $\leq r$, definit prin:

$$P_{n,r}(t_{n-i}) = f_{n-i}\,, \quad i = \overline{0,r}\,.$$

Dacă alegem pentru $P_{n,r}$ polinomul Lagrange, atunci se scrie sub forma:

$$P_{n,r}(t) = \sum_{i=0}^{r} f_{n-i} L_{n,i,r}(t)\,,$$

unde

$$L_{n,i,r}(t) = \prod_{j=0,j\neq i}^{r} \frac{t - t_{n-j}}{t_{n-i} - t_{n-j}}\,, \quad i = \overline{0,r}\,.$$

Avem

$$\int_{t_n}^{t_{n+1}} P_{n,r}(t)\mathrm{d}t = \sum_{i=0}^{r} f_{n-i} \int_{t_n}^{t_{n+1}} L_{n,i,r}(t)\mathrm{d}t\,,$$

unde notând:

$$b_{n,i,r} = \frac{1}{h_n} \int_{t_n}^{t_{n+1}} L_{n,i,r}(t)\mathrm{d}t\,, \quad i = \overline{0,r}\,,$$

schema Adams-Bashforth se poate scrie:

$$\begin{cases} y_{n+1} = y_n + \displaystyle\sum_{i=0}^{r} h_n b_{n,i,r} f_{n-i}\,, \\ f_{n+1} = f(t_{n+1}, y_{n+1})\,, \quad n = \overline{r, N-1}\,. \end{cases}$$

**Observaţie**. Pentru $r = 0$ se obţine metoda Euler progresivă, deoarece:

$$P_{n,0}(t) = f_n = f(t_n, y_n)\,, \quad L_{n,0,0}(t) = 1\,, \quad b_{n,0,0} = 1\,,$$

şi, deci

$$y_{n+1} = y_n + h_n f(t_n, y_n)\,.$$

Dacă considerăm $h_n = h_{n-1} = \ldots = h_{n-r} = h = \text{const.}$, atunci coeficienţii $b_{n,i,r}$ sunt independenţi de $n$. Din (**??**) prin schimbarea de variabilă $t = t_n + sh$ şi punând

$t_{n-j} = t_n - jh$ , $j = \overline{0, r}$ , se obţine:

$$b_{n,i,r} = b_{i,r} = \int_0^1 l_{i,r}(s)\mathrm{d}s\,, \qquad \text{unde } l_{i,r}(s) = \prod_{j=0, j\neq i}^{r} \frac{s+j}{j-i}\,,\ i = \overline{0, r}\,.$$

În tabelul 7 avem coeficienţii metodei Adams-Bashforth $(b_{i,r})_{i=\overline{0,r}}$ pentru $r = \overline{1, 5}$, astfel că:

- pentru $r = 1$ avem $y_{n+1} = y_n + \dfrac{h}{2}\,(3f_n - f_{n-1})$;
- pentru $r = 2$ avem $y_{n+1} = y_n + \dfrac{h}{12}\,(23f_n - 16f_{n-1} + 5f_{n-2})$;
- pentru $r = 3$ avem $y_{n+1} = y_n + \dfrac{h}{24}\,(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})$;
- pentru $r = 4$ avem

$$y_{n+1} = y_n + \frac{h}{720}\,(1901f_n - 2774f_{n-1} + 2616f_{n-2} - 1274f_{n-3} + 251f_{n-4})$$

l

**Exemplul -1.3** *Folosind metoda Adams-Bashforth pentru $r = 3$, determinaţi soluţia aproximativă a problemei Cauchy:*

$$\begin{cases} y' = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}\,, \quad t \in [1, 2.25] \\ y(1) = 0.75 \end{cases}$$

*în $t = 2.25$ pentru $N = 5$, determinând valorile de început $y_1, y_2, y_3$ cu metoda Euler progresivă.*

Avem $h = 0.25$ şi $f(t, y) = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}$. Aplicând metoda Euler progresivă se obţine:

$$\begin{aligned}
&t_0 = 1\,, && y_0 = 0.75\,, && f_0 = f(t_0, y_0) = -0.4375\,,\\
&t_1 = 1.25\,, && y_1 = 0.6406\,, && f_1 = f(t_1, y_1) = -0.2621\,,\\
&t_2 = 1.5\,, && y_2 = 0.5751\,, && f_2 = f(t_2, y_2) = -0.1638\,,\\
&t_3 = 1.75\,, && y_3 = 0.5342\,, && f_3 = f(t_3, y_3) = -0.1015\,,
\end{aligned}$$

şi, în continuare, cu metoda Adams-Bashforth pentru $r = 3$, rezultă:

$$t_4 = 2\,, \qquad y_4 = y_3 + \frac{h}{24}\left(55f_3 - 59f_2 + 37f_1 - 9f_0\right) = 0.5166\,,$$

$$f_4 = f(t_4, y_4) = -0.0539\,,$$

$$t_5 = 2.25\,, \qquad y_5 = y_4 + \frac{h}{24}\left(55f_4 - 59f_3 + 37f_2 - 9f_1\right) = 0.5096\,,$$

$$f_5 = f(t_5, y_5) = -0.0162\,.$$

Table 2: *Conditiile $A'(l)$ pentru $l = \overline{1,5}$.*

| $l$ | $A'(l)$ |
|---|---|
| 1 | $b^t e = 1$ |
| 2 | $b^t Ce = b^t Ae = \frac{1}{2}$ |
| 3 | $b^t C^2 e = b^t CAe = b^t (Ae)^{\cdot 2} = \frac{1}{3}$ , $\quad b^t ACe = b^t A^2 e = \frac{1}{6}$ |
| 4 | $b^t C^3 e = b^t C^2 Ae = b^t C (Ae)^{\cdot 2} = b^t (Ae)^{\cdot 3} = \frac{1}{4}$ , <br> $b^t CACe = b^t CA^2 e = b^t ((Ae).(ACe)) = b^t \big((Ae).(A^2 e)\big) = \frac{1}{8}$ <br> $b^t AC^2 e = b^t ACAe = b^t A(Ae)^{\cdot 2} = \frac{1}{12}$ , $\quad b^t A^2 Ce = b^t A^3 e = \frac{1}{24}$ |
| 5 | $b^t C^4 e = b^t C^3 Ae = b^t C^2 (Ae)^{\cdot 2} = b^t C (Ae)^{\cdot 3} = b^t (Ae)^{\cdot 4} = \frac{1}{5}$ , <br> $b^t C^2 ACe = b^t C^2 A^2 e = b^t C ((Ae).(ACe)) = b^t C \big((Ae).(A^2 e)\big) =$ <br> $b^t \big((Ae)^{\cdot 2}.(ACe)\big) = b^t C \big((Ae)^{\cdot 2}.(A^2 e)\big) = \frac{1}{10}$ , <br> $b^t CAC^2 e = b^t CACAe = b^t CA \big((Ae)^{\cdot 2}\big) = b^t \big((Ae).(AC^2 e)\big) =$ <br> $b^t ((Ae).(ACAe)) = b^t \big((Ae).A(Ae)^{\cdot 2}\big) = \frac{1}{15}$ , <br> $b^t AC^3 e = b^t AC^2 Ae = b^t AC(Ae)^{\cdot 2} = b^t A(Ae)^{\cdot 3} e = b^t (ACe)^{\cdot 2} e =$ <br> $b^t \big((ACe).(A^2 e)\big) = b^t \big(A^2 e)\big)^{\cdot 2} = \frac{1}{20}$ , <br> $b^t CA^2 Ce = b^t CA^3 e = b^t \big((Ae).(A^2 Ce)\big) = b^t \big((Ae).(A^3 e)\big) = \frac{1}{30}$ , <br> $b^t ACACe = b^t ACA^2 e = b^t A ((Ae).(ACe)) = b^t A \big((Ae).(A^2 e)\big) = \frac{1}{40}$ , <br> $b^t A^2 C^2 e = b^t A^2 CAe = b^t A^2 (Ae)^{\cdot 2} = \frac{1}{60}$ , $\quad b^t A^3 Ce = b^t A^4 e = \frac{1}{120}$ |

Table 3: *Conditiile suficiente pentru metode de ordin $p \in \{3,4,5\}$.*

| $p$ | Conditii suficiente pentru metoda Runge-Kutta de ordin $p$ |
|---|---|
| 3 | $Ae = Ce$ , $b^t e = 1$ , $b^t Ce = \frac{1}{2}$ , $b^t C^2 e = \frac{1}{3}$ , $b^t ACe = \frac{1}{6}$ |
| 4 | $b^t C^3 e = \frac{1}{4}$ , $b^t CACe = \frac{1}{8}$ , $b^t AC^2 e = \frac{1}{12}$ , $b^t A^2 Ce = \frac{1}{24}$ |
| 5 | $b^t C^4 e = \frac{1}{5}$ ,<br>$b^t C^2 ACe = \frac{1}{10}$ , $b^t CAC^2 e = \frac{1}{15}$ , $b^t (AC)^{\cdot 2} e = \frac{1}{20}$ , $b^t CA^2 Ce = \frac{1}{30}$ |

Table 4: *Valori aproximative pentru $N = 2$.*

| n | $t_n$ | Solutia exacta | Metoda Euler | Metoda $RK_{12}$ | Metoda Euler mod. | Metoda $RK_{24}$ |
|---|---|---|---|---|---|---|
| 0 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1 | 1.9500 | 1.2088 | 0.7625 | 0.9405 | 1.0610 | 1.1905 |
| 2 | 2.9000 | 10.1724 | 0.8809 | 1.6339 | 2.1295 | 5.3464 |

Table 5: *Valori aproximative pentru $N = 7$.*

| n | $t_n$ | Solutia exacta | Metoda Euler | Metoda $RK_{12}$ | Metoda Euler mod. | Metoda $RK_{24}$ |
|---|---|---|---|---|---|---|
| 0 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1 | 1.2714 | 0.9718 | 0.9321 | 0.9635 | 0.9698 | 0.9717 |
| 2 | 1.5429 | 1.0103 | 0.9270 | 0.9934 | 1.0048 | 1.0103 |
| 3 | 1.8143 | 1.1190 | 0.9687 | 1.0881 | 1.1058 | 1.1188 |
| 4 | 2.0857 | 1.3335 | 1.0578 | 1.2736 | 1.3011 | 1.3332 |
| 5 | 2.3571 | 1.7677 | 1.2083 | 1.6274 | 1.6731 | 1.7667 |
| 6 | 2.6286 | 2.8825 | 1.4532 | 2.3968 | 2.4825 | 2.8745 |
| 7 | 2.9000 | 10.1724 | 1.8665 | 4.7272 | 4.9158 | 9.2669 |

Table 6: *Valori aproximative pentru $N = 20$.*

| n | $t_n$ | Solutia exacta | Metoda Euler | Metoda $RK_{12}$ | Metoda Euler mod. | Metoda $RK_{24}$ |
|---|---|---|---|---|---|---|
| 0 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 | 1.00000 |
| 1 | 1.09500 | 0.98155 | 0.97625 | 0.98114 | 0.98149 | 0.98155 |
| 2 | 1.19000 | 0.97265 | 0.96228 | 0.97187 | 0.97251 | 0.97265 |
| 3 | 1.28500 | 0.97219 | 0.95666 | 0.97105 | 0.97195 | 0.97219 |
| 4 | 1.38000 | 0.97960 | 0.95849 | 0.97808 | 0.97922 | 0.97960 |
| 5 | 1.47500 | 0.99472 | 0.96732 | 0.99278 | 0.99417 | 0.99472 |
| 6 | 1.57000 | 1.01777 | .982995 | 1.01535 | 1.01700 | 1.01777 |
| 7 | 1.66500 | 1.04936 | 1.00568 | 1.04636 | 1.04828 | 1.04936 |
| 8 | 1.76000 | 1.09054 | 1.03581 | 1.08682 | 1.08905 | 1.09054 |
| 9 | 1.85500 | 1.14290 | 1.07416 | 1.13825 | 1.14083 | 1.14290 |
| 10 | 1.95000 | 1.20879 | 1.12186 | 1.20289 | 1.20588 | 1.20879 |
| 11 | 2.04500 | 1.29162 | 1.18052 | 1.28401 | 1.28749 | 1.29162 |
| 12 | 2.14000 | 1.39644 | 1.25240 | 1.38638 | 1.39046 | 1.39643 |
| 13 | 2.23500 | 1.53090 | 1.34062 | 1.51721 | 1.52199 | 1.53089 |
| 14 | 2.33000 | 1.70713 | 1.44962 | 1.68770 | 1.69333 | 1.70712 |
| 15 | 2.42500 | 1.94532 | 1.58577 | 1.91626 | 1.92284 | 1.94529 |
| 16 | 2.52000 | 2.28175 | 1.75851 | 2.23503 | 2.24252 | 2.28169 |
| 17 | 2.61500 | 2.78861 | 1.98225 | 2.70542 | 2.71305 | 2.78847 |
| 18 | 2.71000 | 3.63278 | 2.28004 | 3.45972 | 3.46390 | 3.63226 |
| 19 | 2.80500 | 5.30646 | 2.69075 | 4.83886 | 4.82237 | 5.30322 |
| 20 | 2.90000 | 10.17241 | 3.28441 | 8.01230 | 7.86910 | 10.11166 |

Table 7: *Valorile coeficien@tilor* $b_{i,r}$ *pentru* $r = \overline{1,5}$.

| | 6 | | | | | |
|---|---|---|---|---|---|---|
| | $i = 0$ | $i = 1$ | $i = 2$ | $i = 3$ | $i = 4$ | $i = 5$ |
| 1 | 3/2 | -1/2 | - | - | - | - |
| 2 | 23/12 | -4/3 | 5/12 | - | - | - |
| 3 | 55/24 | -59/24 | 37/24 | -3/8 | - | - |
| 4 | 1901/720 | -1387/360 | 109/30 | -637/360 | 251/720 | - |
| 5 | 4277/1440 | -2641/480 | 4991/720 | -3649/720 | 959/480 | -95/288 |

# Metode numerice pentru ecuaţii neliniare

## Metoda bisecţiei

**Exemplul -1.1** *Să se aproximeze soluţia ecuaţiei* $x^3 + 4x^2 - 10 = 0$ *(conţinută în intervalul* $[-1, 2]$*).*

Soluţie Avem

$$a_0 = a = -1, b_0 = b = 2, f(a) = -7, f(b) = 14.$$

Atunci $c_0 = (a_0 + b_0)/2 = 1/2$. Cum $f(c_0) < 0$ rezultă că

$$a_1 = c_0 = 1/2,\ b_1 = b_0 = 2,\ c_1 = (a_1 + b_1)/2 = 5/4 = 1,25$$

Cum $f(c_1) < 0$ rezultă că

$$a_2 = c_1 = 5/4,\ b_2 = b_1 = 2,\ c_2 = (a_2 + b_2)/2 = 13/8 = 1,625$$

Cum $f(c_2) > 0$ rezultă că

$$a_3 = a_2 = 5/4,\ b_3 = c_2 = 13/8,\ c_3 = (a_3 + b_3)/2 = 23/16 = 1,4375$$

Cum $f(c_3) > 0$ rezultă că

$$a_4 = a_3 = 5/4,\ b_4 = c_3 = 23/16,\ c_4 = (a_4 + b_4)/2 = 43/32 = 1,34375$$

etc.

## Regula falsi

**Exemplul -1.2** *Să se aproximeze soluţia ecuaţiei* $x^3 + 4x^2 - 10 = 0$ *(conţinută în intervalul* $[-1, 2]$*).*

Soluţie Avem

$$a_0 = a = -1, b_0 = b = 2, f(a) = -7, f(b) = 14.$$

Atunci $c_0 = (a_0 f(b_0) - b_0 f(a_0))/(f(b_0) - f(a_0) = 0$. Cum $f(c_0) < 0$ rezultă că

$$a_1 = c_0 = 0,\ b_1 = b_0 = 2,\ c_1 = (a_1 f(b_1) - b_1 f(a_1))/(f(b_1) - f(a_1) = 5/6 = 0,83333$$

Cum $f(c_1) < 0$ rezultă că

$$a_2 = c_1 = 5/6,\ b_2 = b_1 = 2,\ c_2 = (a_2 f(b_2) - b_2 f(a_2))/(f(b_2) - f(a_2) = 110/91 = 1,20879$$

etc.

## Metoda coardei

**Exemplul -1.3** *Să se aproximeze soluţia ecuaţiei* $2x^3 - 4x + 1 = 0$ *(conţinută în intervalul* $[1/4, 1]$*), folosind metoda coardei.*

Solutie: Fie $f(x) = 2x^3 - 4x + 1$. Avem: $f'(x) = 6x^2 - 8x$ si $f''(x) = 12x$. Luam $a = 1/4$ si $b = 1$. Evident, $f''$ nu se anuleaza pe $[a, b]$, $f(a) \cdot f''(a) = 3/32 > 0$ si $f(a) \cdot f(b) = -1/32 < 0$. Deci conditiile de aplicabilitate ale metodei coardei sunt indeplinite. Luam $x_0 = 1/4$ si $x_1 = 1$. Atunci

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)} = \frac{3}{11},\ x_3 = \frac{x_0 f(x_3) - x_3 f(x_0)}{f(x_3) - f(x_0)} = \frac{899}{3475},\ \text{etc.}\square$$

## Metoda secantei

**Exemplul -1.4** *Să se aplice metoda secantei pe intervalul* $[0, 1]$ *pentru rezolvarea ecuaţiei* $3e^x - 4\cos x = 0$ *pentru* $x_0 = 1$ *şi* $x_1 = 0,9$.

Soluţie: Evident $f$ este de două ori derivabilă pe $[0,1]$ şi $f\prime(x) = 3e^x + 4\sin x,$, funcţie care evident nu se anulează pe $[0,1]$. Cum $f(0)\cdot f(1) = -3e + 4\cos 1 < 0$ rezultă că ecuaţia are o radăcină în intervalul $[0,1]$.

Avem

$$x_2 := \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)} = \frac{3e^{9/10} - 4\cos(9/10) - 9/10(3e - 4\cos 1)}{3e^{9/10} - 4\cos(9/10) - 3e + 4\cos 1} = 0,45575$$

## Metoda lui Newton

**Exemplul -1.5** *Să se arate că se aplică metoda lui Newton pe intervalul* $[0,1]$ *pentru rezolvarea ecuaţiei* $3e^x - 4\cos x = 0$. *Luând* $x_0 = 1$, *să se calculeze valoarea primei iteraţii.*

Soluţie: Evident $f$ este de două ori derivabilă pe $[0,1]$ şi $f\prime(x) = 3e^x + 4\sin x$, $f\prime\prime(x) = 3e^x + 4\cos x$, funcţii care evident nu se anulează pe $[0,1]$. Cum $f(0)\cdot f(1) = -3e + 4\cos 1 < 0$ şi $f(1)\cdot f\prime\prime(1) = 9e^2 - 16\cos 1 > 0$, rezultă că metoda lui Newton se aplică.

Avem

$$x_1 := x_0 - \frac{f(x_0)}{f\prime(x_0)} = 1 - \frac{3e - 4\cos 1}{3e + 4\sin 1} = \frac{4\sin 1 + 4\cos 1}{3e + 4\sin 1} = 0,47975$$

## Principiul contracţiei

**Exemplul -1.6** *Fie funcţia*

$$f : \left[-\frac{3}{4}, -\frac{1}{2}\right] \to \mathbf{R},\ f(x) = x^2 + x - \frac{7}{16}$$

*Să se arate că* $f$ *este o contracţie. Luând* $x_0 = -\frac{5}{8}$, *să se determine numărul de iteraţii necesari pentru a aproxima soluţia ecuaţiei* $f(x) = x$ *cu o eroare* $\varepsilon$.

Soluţie Funcţia $f$ este derivabilă şi

$$f'(x) = 2x + 1$$

Avem

$$\max_{x\in\left[-\frac{3}{4},-\frac{1}{2}\right]} |f'(x)| = \max_{x\in\left[-\frac{3}{4},-\frac{1}{2}\right]} |2x+1| = \max_{x\in\left[-\frac{3}{4},-\frac{1}{2}\right]} (-2x-1) = -2\cdot\left(-\frac{3}{4}\right) - 1 = \frac{1}{2}$$

Deci

$$|f'(x)| \leq q := \frac{1}{2}.$$

Notăm $a = -\frac{3}{4}, b = -\frac{1}{2}$. Atunci

$$\left|f\left(\frac{a+b}{2}\right) - \frac{a+b}{2}\right| \leq (1-q)\cdot\frac{b-a}{2} \Longleftrightarrow \left|f\left(-\frac{5}{8}\right)+\frac{5}{8}\right| \leq (1-\frac{1}{2})\cdot\frac{1}{8} \Longleftrightarrow$$

$$\Longleftrightarrow \left|\frac{25}{64} - \frac{5}{8} - \frac{7}{16} + \frac{5}{8}\right| \leq \frac{1}{16} \Longleftrightarrow \frac{3}{64} \leq \frac{4}{64},\ \text{ceea ce este adevărat.}$$

Deci $f$ este o contracţie. Fie $z$ soluţia ecuaţiei $f(x) = x$ pe intervalul $[a, b]$.

Din formula de evaluare a erorii avem

$$|x_n - z| \leq \frac{q^n}{1-q}\,|x_1 - x_0|$$

Dacă

$$\frac{q^n}{1-q}\,|x_1 - x_0| \leq \varepsilon,\ \text{atunci}\ |x_n - z| \leq \varepsilon,\ \text{deci}\ x_n\ \text{aproximează pe}\ z\ \text{cu eroarea}\ \varepsilon.$$

Avem

$$x_1 = f(x_0) = f\left(-\frac{5}{8}\right) = \frac{25}{64} - \frac{5}{8} - \frac{7}{16} = -\frac{43}{64}$$

şi

$$|x_1 - x_0| = \left|-\frac{43}{64} + \frac{5}{8}\right| = \frac{3}{64}$$

Atunci

$$\frac{q^n}{1-q}\,|x_1 - x_0| \leq \varepsilon \Longleftrightarrow \frac{\left(\frac{1}{2}\right)^n}{1-\frac{1}{2}}\cdot\frac{3}{64} \leq \varepsilon \Longleftrightarrow \frac{3}{2^{n+5}} \leq \varepsilon \Longleftrightarrow$$

$$\Longleftrightarrow 2^{n+5} \geq \frac{3}{\varepsilon} \Longleftrightarrow n \geq \log_2\left(\frac{3}{\varepsilon}\right) - 5$$

Deci putem lua

$$n = \left[\log_2\left(\frac{3}{\varepsilon}\right)\right] - 4, \text{ unde } [\cdot] \text{ reprezintă partea întreagă.}$$

# Metode pentru rezolvarea sau aproximarea soluţiilor sistemelor de ecuaţii liniare

## Norme de vectori şi norme de matrici

**Exemplul -1.7** *Fie*

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & -5 & 6 \\ 7 & -8 & 9 \end{pmatrix}$$

*Atunci*

$$\|A\|_1 = \max_{1\leq j\leq 3} \sum_{i=1}^{3} |a_{ij}| =$$

$$= \max(|1| + |4| + |7|, |-2| + |-5| + |-8|, |3| + |6| + |9|) = \max(12, 15, 18) = 18$$

*şi*

$$\|A\|_\infty = \max_{1\leq i\leq 3} \sum_{j=1}^{3} |a_{ij}| =$$

$$= \max(|1| + |-2| + |3|, |4| + |-5| + |6|, |7| + |-8| + |9|) = \max(6, 15, 24) = 24.$$

## Condiţionarea unui sistem de ecuaţii liniare

**Exemplul -1.8**

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \quad \textit{cu soluţia} \quad x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

*Considerăm sistemul perturbat* $A(x + \delta x) = b + \delta b$ *unde*

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad b + \delta b = \begin{pmatrix} 32,1 \\ 22,9 \\ 33,1 \\ 30,9 \end{pmatrix}$$

$$\textit{cu soluţia } x + \delta x = \begin{pmatrix} 9,2 \\ -12,6 \\ 4,5 \\ -1,1 \end{pmatrix} \tag{1}$$

*Considerăm şi sistemul perturbat* $(A + \Delta A)(x + \Delta x) = b$ *unde*

$$A = \begin{pmatrix} 10 & 7 & 8,1 & 7,2 \\ 7,08 & 5,04 & 6 & 5 \\ 8 & 5,98 & 9,89 & 9 \\ 6,99 & 4,99 & 9 & 9,98 \end{pmatrix} \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 30 \end{pmatrix}$$

$$\textit{cu soluţia } x + \Delta x = \begin{pmatrix} -81 \\ 137 \\ -34 \\ 22 \end{pmatrix} \tag{2}$$

**Exemplul -1.9** *Sa se arate ca* $\frac{\|\delta x\|_\infty}{\|x\|_\infty} = cond(A) \cdot \frac{\|\delta b\|_\infty}{\|b\|_\infty}$ *(Vom considera cazul sistemului (1) şi vom lucra cu norma infinit).*

Avem:

$$\|\delta x\|_\infty = \|x + \delta x - x\|_\infty = \|(9,2 - 1; -12,6 - 1; 4,5 - 1; -1,1 - 1)\|_\infty =$$

$$= \|(8,2; -13,6; 3,5; -2,1)\|_\infty = \max(|8,2|\,; |-13,6|\,; |3,5|\,; |-2,1|) = 13,6.$$

$$\|x\|_\infty = \|(1;1;1;1)\|_\infty = \max(|1|\,;|1|\,;|1|\,;|1|) = 1.$$

Atunci

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} = \frac{13,6}{1} = 13,6$$

$$A^{-1} = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}$$

$$\|A\|_\infty = \max(|10|+|7|+|8|+|7|\,,|7|+|5|+|6|+|5|\,,|8|+|6|+|10|+|9|\,,|7|+|5|+|9|+|10|) =$$

$$= \max(32, 23, 33, 31) = 33$$

$$\left\|A^{-1}\right\|_\infty = \max(|25| + |-41| + |10| + |-6|\,, |-41| + |68| + |-17| + |10|\,,$$

$$, |10| + |-17| + |5| + |-3|\,, |-6| + |10| + |-3| + |2|) =$$

$$= \max(82, 136, 35, 21) = 136.$$

Rezultă că

$$cond(A) := \|A\|_\infty \cdot \left\|A^{-1}\right\|_\infty = 33 \cdot 136 = 4488$$

Avem

$$\|\delta b\|_\infty = \|b + \delta b - b\|_\infty = \|(32,1 - 32; 22,9 - 23; 33,1 - 33; 30,9 - 31)\|_\infty =$$

$$= \|(0,1; -0,1; 0,1; -0,1)\|_\infty = \max(|0,1|\,; |-0,1|\,; |0,1|\,; |-0,1|) = 0,1.$$

$$\|b\|_\infty = \|(32;23;33;31)\|_\infty = \max(|32|\,;|23|\,;|33|\,;|31|) = 33..$$

Atunci

$$\frac{\|\delta b\|_\infty}{\|b\|_\infty} = \frac{0,1}{33} = \frac{1}{330}$$

şi

$$cond(A) \cdot \frac{\|\delta b\|_\infty}{\|b\|_\infty} = 4488 \cdot \frac{1}{330} = 13,6$$

Deci

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} = cond(A) \cdot \frac{\|\delta b\|_\infty}{\|b\|_\infty}$$

# Metode directe

# Metoda lui Gauss (cu pivotare parţială)

**Exemplul -1.10** *Fie sistemul de ecuaţii liniare*

$$\begin{cases} x + \frac{1}{2}y + \frac{1}{3}z &= 1 \\ \frac{1}{2}x + \frac{1}{3}y + \frac{1}{4}z &= 2 \\ \frac{1}{3}x + \frac{1}{4}y + \frac{1}{5}z &= 3 \end{cases}$$

*Să se calculeze determinantul matricei sistemului şi să se rezolve sistemul cu metoda lui Gauss.*

Soluţie Notăm cu $(a_{ij})_{1\le i,j\le 3}$ matricea sistemului. Iniţializăm det $= 1$. La etapa $n = 1$ căutăm cel mai mare element în modul de pe coloana $n$ a matricei sistemului. Acesta este max $= |a_{11}| = 1$ şi se află pa linia 1. Deci $s = 1$ şi $piv = a_{11} = 1$. Calculăm det $=$ det $*piv$, deci det $= 1$. Cum $s = n$ se trece la pasul împărţirii la pivot. Cum pivotul este egal cu 1, prin împărţire la pivot linia $n$ rămâne nemodificată. Eliminăm $x$ din liniile 2 şi 3 ca în algoritm. Obţinem următorul sistem de ecuaţii:

$$\begin{cases} x + \frac{1}{2}y + \frac{1}{3}z &= 1 \\ \frac{1}{12}y + \frac{1}{12}z &= \frac{3}{2} \\ \frac{1}{12}y + \frac{4}{45}z &= \frac{8}{3} \end{cases}$$

La etapa $n = 2$ căutăm cel mai mare element în modul de pe coloana $n$ a matricei sistemului (de la elementul $a_{22}$ în jos). Acesta este max $= |a_{22}| = 1/12$ şi se află pa linia 2. Deci $s = 2$ şi $piv = a_{22} = 1/12$. Calculăm det = det $*piv$, deci det $= 1/12$. Cum $s = n$ se trece la pasul împărţirii la pivot. Ecuaţia 2 devine $y + z \;=\; 18$ . Eliminăm $y$ din linia 3 ca în algoritm. Obţinem următorul sistem de ecuaţii:

$$\begin{cases} x + \frac{1}{2}y + \frac{1}{3}z &= 1 \\ y + z &= 18 \\ \frac{1}{180}z &= \frac{7}{6} \end{cases}$$

Calculăm determinantul matricei sistemului det = det $*a_{33} = \frac{1}{12} * \frac{1}{180} = \frac{1}{2160}$ şi rezolvăm sistemul anterior. Obţinem $z = 210$, $y = -192$, $x = 27$.

# Metoda lui Gauss-Jordan de calcul a inversei unei matrice

**Exemplul -1.11** *Folosind metoda Gauss-Jordan să se calculeze inversa matricei*

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$$

Soluţie Formăm ansamblul matriceal

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & 0 & 0 & 1 \end{array}\right)$$

Cu transformările din algoritm obţinem:

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & \frac{1}{12} & \frac{1}{12} & -\frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{12} & \frac{4}{45} & -\frac{1}{3} & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & 1 & 1 & -6 & 12 & 0 \\ 0 & \frac{1}{12} & \frac{4}{45} & -\frac{1}{3} & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & 1 & 1 & -6 & 12 & 0 \\ 0 & 0 & \frac{1}{180} & \frac{1}{6} & -1 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & 1 & 1 & -6 & 12 & 0 \\ 0 & 0 & 1 & 30 & -180 & 180 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 0 & -9 & 60 & -60 \\ 0 & 1 & 0 & -36 & 192 & -180 \\ 0 & 0 & 1 & 30 & -180 & 180 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 9 & -36 & 30 \\ 0 & 1 & 0 & -36 & 192 & -180 \\ 0 & 0 & 1 & 30 & -180 & 180 \end{array}\right)$$

Deci

$$A^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix}$$

Rezolvarea unui sistem de ecuaţii liniare $Ax = b$ este echivalentă, după ce s-a calculat inversa matricei $A$, cu egalitatea $x = A^{-1}b$.

## Factorizarea LU

**Exemplul -1.12** *Să se factorizeze sub forma LU matricea*

$$A = \begin{pmatrix} 2 & 4 & -2 & 0 \\ -1 & -1 & 2 & 3 \\ 4 & 5 & -2 & -9 \\ 0 & 1 & 3 & 4 \end{pmatrix}$$

Soluţie Se verifică mai întâi că determinanţii de colţ ai matricei $A$ sunt nenuli. Cu formulele (**??**) şi (**??**) se obţine:

$$l_{11} = 2,\ l_{21} = -1,\ l_{31} = 4,\ l_{41} = 0$$

$$u_{11} = 1,\ u_{12} = 2,\ u_{13} = -1,\ u_{14} = 0$$

$$l_{22} = 1,\ l_{32} = -3,\ l_{42} = 1$$

$$u_{22} = 1,\ u_{23} = 1,\ u_{24} = 3$$

$$l_{33} = 5,\ l_{43} = 2$$

$$u_{33} = 1,\ u_{34} = 0$$

$$l_{44} = 1$$

$$u_{44} = 1$$

Deci

$$L = \begin{pmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 4 & -3 & 5 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

# Metoda rădăcinii pătrate

**Exemplul -1.13** *Folosind metoda rădăcinii pătrate, să se rezolve sistemul de ecuaţii liniare $Ax = b$ cu*

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 10 & 4 \\ 2 & 4 & 6 \end{pmatrix}, b = \begin{pmatrix} 8 \\ 16 \\ 12 \end{pmatrix}$$

Soluţie Se verifică imediat că matricea $A$ este simetrică. Deoarece

$$\Delta_1 = |4| > 0,\ \Delta_2 = \begin{vmatrix} 4 & 2 \\ 2 & 10 \end{vmatrix} = 36 > 0,\ \Delta_3 = \begin{vmatrix} 4 & 2 & 2 \\ 2 & 10 & 4 \\ 2 & 4 & 6 \end{vmatrix} = 144 > 0$$

rezultă că matricea $A$ este pozitiv definită. **Cu formulele din curs** se obţine

$$l_{11} = 2,\ l_{21} = 1,\ l_{31} = 1$$

$$l_{22} = 3,\ l_{32} = 1$$

$$l_{33} = 2$$

Rezultă că

$$y_1 = 4,\ y_2 = 4,\ y_3 = 2$$

si

$$x_3 = 1,\ x_2 = 1,\ x_3 = 1.$$

# Metode iterative de rezolvare a sistemelor de ecuaţii liniare

## Metoda lui Jacobi

**Exemplul -1.14** *Să se arate că se poate aplica metoda lui Jacobi (relativă la normele 1 şi $\infty$) pentru sistemul de ecuaţii liniare $Ax = b$ cu*

$$A = \begin{pmatrix} 0,9 & 0,1 & 0,2 & 0,3 \\ -0,2 & 0,8 & 0,1 & 0,4 \\ 0,1 & -0,3 & 0,7 & -0,1 \\ -0,3 & -0,2 & -0,1 & 0,9 \end{pmatrix};\ b = \begin{pmatrix} 3,6 \\ 2 \\ 2,4 \\ 1,5 \end{pmatrix}.$$

*Luând* $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ *să se determine numărul de iteraţii necesar pentru a aproxima soluţia sistemului cu o eroare mai mică de* $10^{-10}$.

Soluţie: Avem

$$B = I - A = \begin{pmatrix} 0,1 & -0,1 & -0,2 & -0,3 \\ 0,2 & 0,2 & -0,1 & -0,4 \\ -0,1 & 0,3 & 0,3 & 0,1 \\ 0,3 & 0,2 & 0,1 & 0,1 \end{pmatrix}$$

Atunci

$$\|B\|_{\mathbf{1}} = \max_{1\le j\le 4} \sum_{i=1}^{4} |b_{ij}| = \max(|0,1|+|0,2|+|-0,1|+|0,3|\,, |-0,1|+|0,2|+|0,3|+|0,2|\,,$$

$$|-0,2|+|-0,1|+|0,3|+|0,1|\,, |-0,3|+|-0,4|+|0,1|+|0,1|) = \max(0,7;0,8;0,7;0,9) = 0,9 < 1.$$

şi

$$\|B\|_{\infty} = \max_{1\le i\le 4} \sum_{j=1}^{4} |b_{ij}| = \max(|0,1|+|-0,1|+|-0,2|+|-0,3|\,, |0,2|+|0,2|+|-0,1|+|-0,4|\,,$$

$$|-0,1|+|0,3|+|0,3|+|0,1|\,, |0,3|+|0,2|+|0,1|+|0,1|) = \max(0,7;0,9;0,8;0,7) = 0,9 < 1.$$

Deci metoda lui Jacobi se aplică. Din formula de evaluare a erorii avem:

$$\|x^{(n)} - x\|_p \le \frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_p, \ \forall\, n \in \mathbf{N}^*$$

unde $p = 1$ sau $p = \infty$, $q = 0,9$, $x^{(1)} = Bx^{(0)} + b = b$. Deci, pentru a aproxima $x$ cu $x^{(n)}$ cu eroarea $\varepsilon = 10^{-10}$ este suficient ca $\frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_p < \varepsilon$. Avem

$$\|x^{(1)} - x^{(0)}\|_1 = \|b\|_1 = |3,6| + |2| + |2,4| + |1,5| = 9,5$$

şi

$$\|x^{(1)} - x^{(0)}\|_\infty = \|b\|_\infty = \max(|3,6|\,,|2|\,,|2,4|\,,|1,5|) = 3,6$$

Atunci

$$\frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_1 < \varepsilon \Longleftrightarrow \frac{0,9^n}{1-0,9}9,5 < \varepsilon \Longleftrightarrow 0,9^n < \frac{\varepsilon}{95} \Longleftrightarrow n = \left[\log_{0,9}\frac{\varepsilon}{95}\right] + 1$$

şi

$$\frac{q^n}{1-q}\|x^{(1)} - x^{(0)}\|_\infty < \varepsilon \Longleftrightarrow \frac{0,9^n}{1-0,9}3,6 < \varepsilon \Longleftrightarrow 0,9^n < \frac{\varepsilon}{36} \Longleftrightarrow n = \left[\log_{0,9}\frac{\varepsilon}{36}\right] + 1$$

# Metoda Gauss-Seidel

**Exemplul -1.15** *Fie sistemul de ecuaţii liniare*

$$\begin{cases} x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 = 1 \\ \frac{1}{5}x_1 + x_2 + \frac{1}{6}x_3 = 2 \\ \frac{1}{10}x_1 + \frac{1}{20}x_2 + x_3 = 3 \end{cases}$$

*Să se arate că se aplică metoda Gauss-Seidel.*

Soluţie Notăm cu $A$ matricea ceoficienţilor sistemului. Avem

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{5} & 1 & \frac{1}{6} \\ \frac{1}{10} & \frac{1}{20} & 1 \end{pmatrix}$$

Atunci

$$B = I - A = \begin{pmatrix} 0 & -\frac{1}{2} & -\frac{1}{3} \\ -\frac{1}{5} & 0 & -\frac{1}{6} \\ -\frac{1}{10} & -\frac{1}{20} & 0 \end{pmatrix}$$

şi

$$q_1 = |b_{11}| + |b_{12}| + |b_{13}| = \left|-\frac{1}{2}\right| + \left|-\frac{1}{3}\right| = \frac{5}{6}$$

$$q_2 = |b_{21}|\, q_1 + |b_{22}| + |b_{23}| = \left|-\frac{1}{5}\right| \cdot \frac{5}{6} + \left|-\frac{1}{6}\right| = \frac{1}{3}$$

$$q_3 = |b_{31}|\, q_1 + |b_{32}|\, q_2 + |b_{33}| = \left|-\frac{1}{10}\right| \cdot \frac{5}{6} + \left|-\frac{1}{20}\right| \cdot \frac{1}{3} = \frac{1}{10}$$

Deci

$$q = \max(q_1, q_2, q_3) = \max\left(\frac{5}{6}, \frac{1}{3}, \frac{1}{10}\right) = \frac{5}{6} < 1$$

Rezultă că metoda Gauss-Seidel se aplică.

## Metoda relaxării simultane

**Exemplul -1.16** *Fie sistemul de ecuaţii liniare:*

$$\begin{cases} 5x_1 + 3x_2 + 2x_3 = 1 \\ 3x_1 + 6x_2 + 3x_3 = 2 \\ 2x_1 + 3x_2 + 5x_3 = 3 \end{cases}$$

*Să se arate că se poate aplica metoda relaxării simultane. Să se determine parametrul optim de relaxare. Luând* $x^{(0)} = (0, 0, 0)$ *să se evalueze eroarea* $x - x^{(n)}$.

Soluţie: Notăm cu $A$ matricea sistemului. Avem

$$A = \begin{pmatrix} 5 & 3 & 2 \\ 3 & 6 & 3 \\ 2 & 3 & 5 \end{pmatrix} = A^t \text{ transpusa matricei } A.$$

Deci $A$ simetrică. Deoarece

$$\Delta_1 := |5| > 0, \quad \Delta_2 := \begin{vmatrix} 5 & 3 \\ 3 & 6 \end{vmatrix} > 0, \quad \Delta_3 := \begin{vmatrix} 5 & 3 & 2 \\ 3 & 6 & 3 \\ 2 & 3 & 5 \end{vmatrix} > 0,$$

cu teorema lui Sylvester rezultă că matricea $A$ este şi pozitiv definită. în concluzie, se poate aplica metoda relaxării simultane.

Fie $D = diag(A)$. Atunci

$$D^{-1}A = \begin{pmatrix} 1 & 3/5 & 2/5 \\ 1/2 & 1 & 1/2 \\ 2/5 & 3/5 & 1 \end{pmatrix}$$

şi $\det(D^{-1}A - \lambda I) = 0 \iff (1-\lambda)^3 - \frac{19}{25}(1-\lambda) + \frac{6}{25} = 0$. Dacă notăm $1 - \lambda = t$, atunci ecuaţia $25t^3 - 19t + 6 = 0$ are rădăcinile $t_1 = -1$, $t_2 = 2/5$, $t_3 = 3/5$, de unde rezultă că valorile proprii ale matricei $D^{-1}A$ sunt $\lambda_1 = 2 > \lambda_2 = 3/5 > \lambda_3 = 2/5$. Parametrul optim de relaxare este

$$\sigma = \frac{2}{\lambda_1 + \lambda_3} = \frac{5}{6}$$

căruia îi corespunde

$$q = \frac{\lambda_1 - \lambda_3}{\lambda_1 + \lambda_3} = \frac{2}{3}.$$

Avem $x^{(1)} = C_\sigma x^{(0)} + c_\sigma = c_\sigma = \sigma D^{-1} b = 5/6 \cdot (1/5; 2/6; 3/5)^t$ unde $b = (1; 2; 3)^t$ este vectorul termenilor liberi din sistem. Atunci

$$\|x^{(1)}\|_D^2 = \left\langle Dx^{(1)}, x^{(1)} \right\rangle = \frac{25}{36}\left(\frac{1}{5} + \frac{2}{3} + \frac{9}{5}\right) = \frac{2}{3} \cdot \frac{25}{9}$$

şi

$$\|x^{(n)} - x\|_D \le \frac{q^n}{1-q} \|x^{(1)} - x^{(0)}\|_D = 5 \cdot \left(\frac{2}{3}\right)^{n+\frac{1}{2}}.$$

## Metoda lui Ritz

**Exemplul -1.17** *Fie* $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 1 & 10 \\ 3 & 1 & 35 & 5 \\ 4 & 10 & 5 & 30 \end{pmatrix}$. *Folosind metoda lui Ritz să se calculeze* $A^{-1}$.

Se verifică mai întâi că matricea $A$ este simetrică şi pozitiv definita. Luăm $w_0 = (1,0,0,0)^*$. Atunci $w_0^* A w_0 = 1$ şi

$$C_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ D_1 = \begin{pmatrix} 0 & -2 & -3 & -4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Luăm $v_1 = (0,1,0,0)^*$. Rezultă că $w_1 = (-2,1,0,0)^*$ şi $w_1^* A w_1 = 1$, de unde

$$C_2 = \begin{pmatrix} 5 & -2 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ D_2 = \begin{pmatrix} 0 & 0 & -13 & 0 \\ 0 & 0 & 5 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Luăm $v_2 = (0,0,1,0)^*$. Rezultă că $w_2 = (-13,5,1,0)^*$ şi $w_2^* A w_2 = 1$, de unde

$$C_3 = \begin{pmatrix} 174 & -67 & -13 & 0 \\ -67 & 26 & 5 & 0 \\ -13 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ D_3 = \begin{pmatrix} 0 & 0 & 0 & 39 \\ 0 & 0 & 0 & -17 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Luăm $v_3 = (0,0,0,1)^*$. Rezultă că $w_3 = (39,-17,-3,1)^*$ şi $w_3^* A w_3 = 1$, de unde

$$C_4 = \begin{pmatrix} 1695 & -730 & -130 & 39 \\ -730 & 315 & 56 & -17 \\ -130 & 56 & 10 & -3 \\ 39 & -17 & -3 & 1 \end{pmatrix}, \ D_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

## Pseudoinversa unei matrice

**Exemplul -1.18** *Fie*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \quad cu\ rang(T) = 2 \text{ şi } H_0 = \mathbf{0}.$$

Luăm $y_1 = T^* * \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ (avem $y_1 - H_0 T y_1 = y_1 \neq 0$. Atunci $w_1 = y_1 - H_0 T y_1 = y_1$ şi $T w_1 = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 2 \end{pmatrix}$. Rezultă că $\|T w_1\|^2 = 12$, iar $w_1 (T w_1)^* = \begin{pmatrix} 2 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ -2 & 0 & -2 & -2 \end{pmatrix}$. Deci

$$H_1 = H_0 + \frac{w_1 (T w_1)^*}{\|T w_1\|^2} = \begin{pmatrix} 1/6 & 0 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 \\ -1/6 & 0 & -1/6 & -1/6 \end{pmatrix}.$$

Luăm $y_2 = T^* * \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ (avem $y_2 - H_1 T y_2 = \begin{pmatrix} -1/2 \\ 1 \\ 1/2 \end{pmatrix} \neq 0$. Atunci

$w_2 = y_2 - H_1Ty_2$ şi $Tw_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$. Rezultă că $\|Tw_2\|^2 = 3$, iar $w_2(Tw_2)^* =$

$\begin{pmatrix} 1/2 & -1/2 & 0 & -1/2 \\ -1 & 1 & 0 & 1 \\ -1/2 & 1/2 & 0 & 1/2 \end{pmatrix}$. Deci

$$H_2 = H_1 + \frac{w_2(Tw_2)^*}{\|Tw_2\|^2} = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix}$$

În concluzie

$$T^+ = H_2 = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix}$$

.

**Exemplul -1.19** *Fie*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \quad cu \ \dim R(T^*) = 2$$

*Luând* $e_1 = (1;0;-1)^*, e_2 = (0;1;0)^*$ *(bază în* $R(T^*)$*) obţinem* $f_1 = (2;0;2;2)^*, f_2 = (0;1;1;2)^*$. *Atunci*

$$x_1 = \frac{1}{2\sqrt{3}}(1;0;-1)^*, \ y_1 = \frac{1}{\sqrt{3}}(1;0;1;1)^*$$

*În continuare*

$$e_2' = e_2 - \langle f_2, y_1 \rangle x_1 = \frac{1}{2}(-1;2;1)^* \ şi \ f_2' = Te_2' = (-1;1;0;1)^*$$

*Atunci*

$$x_2 = \frac{1}{2\sqrt{3}}(-1;2;1)^*,\ y_2 = \frac{1}{\sqrt{3}}(-1;1;0;1)^*$$

*Rezultă că*

$$T^+ = x_1y_1^* + x_2y_2^* = \frac{1}{6}\begin{pmatrix} 2 & -1 & 1 & 0 \\ -2 & 2 & 0 & 2 \\ -2 & 1 & -1 & 0 \end{pmatrix}$$

**Exemplul -1.20** *Fie*

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \quad cu\ rang(T) = 2\ şi\ H_0 = T^*$$

Luând $a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ obţinem $v_0 := (I - TH_0)^*Ta_1 = \begin{pmatrix} -5 \\ -3 \\ -8 \\ -11 \end{pmatrix}$,

$$w_0 = H_0b_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},\ u_0 := \frac{1}{\langle v_0, Tw_0\rangle}(I - H_0T)w_0 = \begin{pmatrix} 5/48 \\ 1/8 \\ -5/48 \end{pmatrix},$$

$$H_1 := H_0 + u_0 \cdot v_0^* = \begin{pmatrix} 23/48 & -5/16 & 1/6 & -7/48 \\ -5/8 & 5/8 & 0 & 5/8 \\ -23/48 & 5/16 & 1/6 & 7/48 \end{pmatrix}$$

Luând $a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ obţinem $v_1 := (I - TH_1)^*Ta_2 = \begin{pmatrix} 7/8 \\ -7/8 \\ 0 \\ -7/8 \end{pmatrix}$, $w_1 =$

$$H_1 b_2 = \begin{pmatrix} -5/16 \\ 5/8 \\ 5/16 \end{pmatrix},\ u_1 := \frac{1}{\langle v_1, Tw_1\rangle}(I - H_1 T)w_1 = \begin{pmatrix} -1/6 \\ 1/3 \\ 1/6 \end{pmatrix},\ \text{deci}$$

$$T^+ = H_2 := H_1 + u_1 \cdot v_1^* = \begin{pmatrix} 1/3 & -1/6 & 1/6 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \\ -1/3 & 1/6 & -1/6 & 0 \end{pmatrix}.$$

# Metode pentru aproximarea valorilor proprii ale unei matrice

## Metode folosind polinomul caracteristic

### Metoda Danilevsky

.

**Exemplul -1.21** *Determinaţi polinomul caracteristic pentru matricea*

$$A = \begin{pmatrix} -12 & -4 & 3 & 4 \\ -23 & -8 & 6 & 8 \\ -60 & -39 & 21 & 28 \\ -3 & -2 & 1 & 2 \end{pmatrix},$$

*folosind metoda Danilevski.*

Aplicând paşii algoritmului se obţine:

$$A_0 = \begin{pmatrix} -12 & -4 & 3 & 4 \\ -23 & -8 & 6 & 8 \\ -60 & -39 & 21 & 28 \\ -3 & -2 & 1 & 2 \end{pmatrix},$$

$$S_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 2 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow A_1 = S_3^{-1} A_0 S_3 = \begin{pmatrix} -3 & 2 & 3 & -2 \\ -5 & 4 & 6 & -4 \\ 22 & -11 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$S_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -\frac{1}{11} & \frac{2}{11} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow A_2 =$$

$$= S_2^{-1} A_1 S_2 = \begin{pmatrix} 1 & -\frac{2}{11} & \frac{37}{11} & -2 \\ -11 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 170 & \end{pmatrix},$$

$$S_1 = \begin{pmatrix} -\frac{1}{11} & \frac{2}{11} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow A_3 = S_1^{-1} A_2 S_1 \begin{pmatrix} 3 & 0 & -37 & 22 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

Deci polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 - 3\lambda^3 + 37\lambda - 22\,. \quad \square$$

## Metoda Krylov

**Exemplul -1.22** *Determinaţi polinomul caracteristic pentru matricea*

$$A = \begin{pmatrix} 5 & -4 & -3 & -1 \\ 1 & 4 & 2 & -9 \\ 2 & 5 & 1 & -5 \\ 7 & 1 & 2 & -1 \end{pmatrix},$$

*folosind metoda Krylov.*

Alegem pentru aplicarea metodei Krylov vectorul $y^{(0)} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ şi calculăm

$$y^{(1)} = Ay^{(0)} = \begin{pmatrix} -16 \\ -21 \\ -5 \\ 11 \end{pmatrix}, \quad y^{(2)} = Ay^{(1)} = \begin{pmatrix} 8 \\ -209 \\ -197 \\ -154 \end{pmatrix},$$

$$y^{(3)} = Ay^{(2)} - \begin{pmatrix} 1621 \\ 164 \\ -456 \\ -393 \end{pmatrix}, \quad y^{(4)} = Ay^{(3)} = \begin{pmatrix} 9210 \\ 4902 \\ 5571 \\ 10992 \end{pmatrix}.$$

Coefienţii polinomului caracteristic se obţin rezolvâd sistemul:

$$\begin{pmatrix} 1621 & 8 & -16 & 1 \\ 164 & -209 & -21 & 2 \\ -456 & -197 & -5 & 3 \\ -393 & -154 & 11 & 4 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \begin{pmatrix} 9210 \\ 4902 \\ 5571 \\ 10992 \end{pmatrix}.$$

Folosind o metodă de rezolvare a sistemelor liniare , obţinem $p = \begin{pmatrix} 9 \\ -45 \\ 369 \\ 885 \end{pmatrix}$ şi polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 + 9\lambda^3 - 45\lambda^2 + 369\lambda + 885\,. \quad \Box$$

## Metoda Leverrier

**Exemplul -1.23** *Determinaţi, folosind metoda Leverrier, polinomul caracteristic şi matricea* $A^{-1}$ *, dacă există, pentru matricea*

$$A = \begin{pmatrix} 11 & 0 & -3 & 12 \\ -9 & -5 & 2 & 5 \\ 2 & -4 & 15 & -11 \\ 2 & 6 & -7 & 10 \end{pmatrix},$$

Aplicând metoda Leverrier pentru matricea$A$ şi $m = 4$, obţinem:

$$s_1 = 31\,, \quad p_1 = -31\,,$$

- pentru $k = 2$:

$$A_2 = A^2 = \begin{pmatrix} 139 & 84 & -162 & 285 \\ -40 & 47 & 12 & -105 \\ 66 & -106 & 288 & -271 \\ -26 & 58 & -169 & 231 \end{pmatrix}, \quad s_2 = 705\,, \quad p_2 = 128\,,$$

- pentru $k = 3$:

$$A_3 = A^3 = \begin{pmatrix} 1019 & 1938 & -4674 & 6720 \\ -1049 & -913 & 1129 & -1427 \\ 7714 & -2248 & 5807 & -5616 \\ -684 & 1772 & -3958 & 4147 \end{pmatrix},$$

$s_3 = 10060\,, \quad p_3 = 2609\,,$

- pentru $k = 4$:

$$A_4 = A^4 = \begin{pmatrix} -2141 & 49326 & -116331 & 140532 \\ -3918 & -8513 & 28245 & -43842 \\ 39468 & -45684 & 116779 & -110709 \\ -23094 & 31854 & -82803 & 85660 \end{pmatrix},$$

$s_4 = 191785\,, \quad p_4 = -12761\,,$

şi polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 - 31\lambda^3 + 128\lambda^2 + 2609\lambda - 12761\,.$$

Cum $p_4 \neq 0$, rezultă că:

$$\begin{aligned} A^{-1} &= -\tfrac{1}{p_4}\left(A^3 + p_1A^2 + p_2A + p_3I_4\right) \\ &= \begin{pmatrix} 0.0570 & -0.0522 & -0.0028 & -0.0454 \\ -0.0753 & -0.0314 & 0.0794 & 0.1934 \\ -0.0060 & 0.0412 & 0.1103 & 0.1079 \\ 0.0296 & 0.0581 & 0.0302 & 0.0686 \end{pmatrix} . \square \end{aligned}$$

**Exemplul -1.24** *Calculaţi coeficienţii polinomului caracteristic şi matricea $A^{-1}$, dacă există, pentru matricea*

$$A = \begin{pmatrix} 11 & 1 & -3 & 2 \\ -9 & -5 & 12 & 5 \\ 2 & -4 & 5 & -11 \\ 2 & 16 & -7 & 10 \end{pmatrix},$$

*folosind algoritmul Fadeev.*

Aplicând algoritmul Fadeev obţinem:

$$A_1 = A\,,$$

$$p_1 = -tr(A_1) = -21\,, \quad B_1 = A_1 + p_1I_4 = \begin{pmatrix} -10 & 1 & -3 & 2 \\ -9 & -26 & 12 & 5 \\ 2 & -4 & -16 & -11 \\ 2 & 16 & -7 & -11 \end{pmatrix},$$

- pentru $k = 2$

$$A_2 = AB_1 = \begin{pmatrix} -121 & 29 & 13 & 38 \\ 169 & 153 & -260 & -230 \\ 4 & -90 & -57 & 50 \\ -158 & -226 & 228 & 51 \end{pmatrix},$$

$$p_2 = -\frac{1}{tr(A_2)} = -13\,, \quad B_2 = A_2 + p_2 I_4 = \begin{pmatrix} -134 & 29 & 13 & 38 \\ 169 & 140 & -260 & -230 \\ 4 & -90 & -70 & 50 \\ -158 & -226 & 228 & 38 \end{pmatrix},$$

- pentru $k = 3$

$$A_3 = AB_2 = \begin{pmatrix} -1633 & 277 & 549 & 114 \\ -381 & -3171 & 1483 & 1598 \\ 814 & 1534 & -1792 & 828 \\ 828 & 668 & -1364 & -3574 \end{pmatrix},$$

$$p_3 = -\frac{1}{tr(A_3)} = 3390\,, \quad B_3 = A_3 + p_3 I_4 = \begin{pmatrix} 1757 & 277 & 549 & 114 \\ -381 & 219 & 1483 & 1598 \\ 814 & 1534 & 1598 & 828 \\ 828 & 668 & -1364 & -184 \end{pmatrix},$$

- pentru $k = 4$

$$A_4 = AB_3 = \begin{pmatrix} 18160 & 0 & 0 & 0 \\ 0 & 18160 & 0 & 0 \\ 0 & 0 & 18160 & 0 \\ 0 & 0 & 0 & 18160 \end{pmatrix},$$

$$p_4 = -\tfrac{1}{tr(A_4)} = -18160\,, \quad B_4 = A_4 + p_4 I_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 - 21\lambda^3 - 13\lambda^2 + 3390\lambda - 18160\,.$$

Cum $p_4 \neq 0$, rezultă că:

$$A^{-1} = -\frac{1}{p_4} B_3 = \begin{pmatrix} 0.0968 & 0.0153 & 0.0302 & 0.0063 \\ -0.0210 & 0.0121 & 0.0817 & 0.0880 \\ 0.0448 & 0.0845 & 0.0880 & 0.0456 \\ 0.0456 & 0.0368 & -0.0751 & -0.0101 \end{pmatrix}. \quad \square$$

## Metoda coeficienţilor nedeterminaţi

**Exemplul -1.25** *Determinaţi coeficienţii polinomului caracteristic pentru matricea*

$$A = \begin{pmatrix} 9 & -4 & 10 & 13 \\ 11 & -1 & -15 & 8 \\ 1 & -2 & -3 & 1 \\ 2 & 4 & 5 & 7 \end{pmatrix},$$

*folosind metoda coeficienţilor nedeterminaţi.*

Avem $p_4 = -2703$. Se obţine sistemul

$$\begin{pmatrix} 1 & 1 & 1 \\ 8 & 4 & 2 \\ 27 & 9 & 3 \end{pmatrix} p = \begin{pmatrix} 334 \\ 482 \\ 372 \end{pmatrix}$$

a cărui soluţie este:

$$p = \begin{pmatrix} -12 \\ -57 \\ 403 \end{pmatrix}.$$

Deci polinomul caracteristic este:

$$p_A(\lambda) = \lambda^4 - 12\lambda^3 - 57\lambda^2 + 403\lambda - 2703\,. \quad \square$$

## Metoda rotaţiilor

**Exemplul -1.26** *Fie matricea*

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

*Folosind metoda rotaţiilor, să se determine valorile proprii ale matricei $A$.*

Se verifică imediat că matricea $A$ este simetrică. Se aleg $p < q$ astfel încât

$$|a_{pq}| = \max\{a_{ij}|; i < j \in \{1, 2, 3\}.$$

Atunci $p = 1$, $q = 3$. Se calculează

$$\theta = \frac{1}{2} \cdot \arctan \frac{2 \cdot a_{pq}}{a_{pp} - a_{qq}} = \frac{\pi}{6}.$$

Se ia

$$T = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$

şi se calculează matricea

$$B = T^t A T = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 8 & 4 \\ 0 & 4 & 8 \end{pmatrix}$$

Se aleg $p < q$ astfel încât

$$|b_{pq}| = \max\{b_{ij}|; i < j \in \{1,2,3\}.$$

Atunci $p = 2$, $q = 3$. Se calculează

$$\theta = \frac{\pi}{4} \text{ deoarece } b_{pp} = b_{qq}.$$

Se ia

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

şi se calculează matricea

$$B = T^t A T == \begin{pmatrix} 20 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Rezultă că valorile proprii ale matricei $A$ sunt $\lambda_1 = 20$, $\lambda_2 = 12$, $\lambda_3 = 4$. □

# Ecuaţii neliniare în $R^m$

## Principiul contracţiei

**Exemplul -1.27** *Fie sistemul de ecuaţii*

$$\begin{cases} x = (80 + x^3 - x^2 - 3yz)/100 \\ y = (60 - y^3 - z^2 + 4xz)/86 \\ z = (40 + z^3 + y^2 - 5xy)/72 \end{cases} \tag{3}$$

*Folosind principiul contracţiei în raport cu norma $\|\cdot\|_\infty$, să se arate că sistemul (3) are soluţie unică $v$ în $V := [0,2]^3$. Pentru o eroare $\varepsilon$ dată şi pentru iterata iniţială $x^{(0)} := (1,1,1)$ să se calculeze iterata $x^{(n)}$ (din principiul contracţiei) pentru care $\|x^{(n)} - v\|_\infty \le \varepsilon$.*

Fie $r = 1$ şi $a = (1,1,1)$. Atunci mulţimea $V = \{x \in \mathbf{R}^m \mid \|x-a\|_\infty \le r\}$. Fie funcţia $F = (f,g,h) : V \to \mathbf{R}^3$ cu $f(x,y,z) := (80 + x^3 - x^2 - 3yz)/100$, $g(x,y,z) := (60 - y^3 - z^2 + 4xz)/86$ şi $h(x,y,z) := (40 + z^3 + y^2 - 5xy)/72$, $\forall\ (x,y,z) \in V$. Sistemul de ecuaţii (3) este echivalent cu $F(x,y,z) = (x,y,z)$.

Avem

$$F'(x,y,z) = \begin{pmatrix} (3x^2-2x)/100 & -3z/100 & -3y/100 \\ 4z/86 & -3y^2/86 & (-2z+4x)/86 \\ -5y/72 & (2y-5x)/72 & 3z^2/72 \end{pmatrix},$$

deci $\|F'(x,y,z)\|_\infty = \max((|3x^2-2x| + |-3z| + |-3y|)/100,\ (|4z| + |-3y^2| + |-2z+4x|)/86,\ (|-5y| + |2y-5x| + |3z^2|)/72) \le \max(28/100, 32/86, 36/72) = = 1/2$ (am folosit $0 \le x,y,z \le 2$). Putem lua $q = 1/2$. Pe de altă parte $\|F(a)-a\|_\infty = \max(28/100, 24/86, 35/72) = 35/72 < (1-q)r = 1/2$. Deci $F$ este $1/2$-contracţie a mulţimii $V$ în raport cu norma $\|\cdot\|_\infty$. Din principiul contracţiei rezultă că ecuaţia $F(x,y,z) = (x,y,z)$ are soluţie unică în $V$.

Pentru $x^{(0)} = (1,1,1)$ se calculează $x^{(n)}$ cu formula de recurenţă $x^{(n)} = F(x^{(n-1)})$, $\forall\ n \in N$. Se opreşte procesul iterativ atunci când $q/(1-q)\|x^{(n)} - x^{(n-1)}\|_\infty < \varepsilon$.

## Metoda Gauss-Seidel neliniară

**Exemplul -1.28** *Fie sistemul de ecuaţii*

$$\begin{cases} x = (432 + x^4 + y^4 - 2z^4 + y^2z^2)/234 \\ y = (324 - x^4 + 2y^4 - 3z^4 + 3x^2z^2)/432 \\ z = (234 + 2x^4 - 3y^4 + 4z^4 - x^2y^2 + 3x + y)/324 \end{cases} \tag{4}$$

*Folosind metoda Gauss-Seidel neliniară, să se arate că sistemul (4) are soluţie unică $v$ în $V := [-2,2] \times [-1,1] \times [0,1]$. Pentru o eroare $\varepsilon$ dată şi pentru iterata iniţială $x^{(0)} := (0,0,1/2)$ să se calculeze iterata $x^{(n)}$ pentru care $\|x^{(n)} - v\|_\infty \leq \varepsilon$.*

Fie $r_1 = 2$, $r_2 = 1$ şi $r_3 = 1/2$, $a_1 = a_2 = 0$ şi $a_3 = 1/2$. Atunci mulţimea $V = \{(x,y,z) \in \mathbf{R}^3 \,|\, |x - a_1| \leq r_1,\ |y - a_2| \leq r_2,\ |z - a_3| \leq r_3\}$. Fie funcţia $F = (f,g,h) : V \to \mathbf{R}^{\,3}$ cu $f(x,y,z) := (432 + x^4 + y^4 - 2z^4 + y^2z^2)/234$, $g(x,y,z) := (324 - x^4 + 2y^4 - 3z^4 + 3x^2z^2)/432$ şi $h(x,y,z) := (234 + 2x^4 - 3y^4 + 4z^4 - x^2y^2 + 3x + y)/324$, $\forall\ (x,y,z) \in V$. Sistemul de ecuaţii (4) este echivalent cu $F(x,y,z) = (x,y,z)$.

Avem $F'(x,y,z) =$

$$\begin{pmatrix} 4x^3/234 & (4y^3 + 2yz^2)/234 & (-8z^3 + 2y^2z)/234 \\ (-4x^3 + 6xz^2)/432 & 8y^3/432 & (-12z^3 + 6x^2z)/432 \\ (8x^3 - 2xy^2 + 3)/324 & (-12y^3 - 2x^2y + 1)/324 & 16z^3/324 \end{pmatrix},$$

deci $\|F'(x,y,z)\|_\infty = \max((|4x^3| + |4y^3 + 2yz^2| + |-8z^3 + 2y^2z|)/234,\ (|-4x^3 + 6xz^2| + |8y^3| + |-12z^3 + 6x^2z|)/432,\ (|8x^3 - 2xy^2 + 3| + |-12y^3 - 2x^2y + 1| + |16z^3|)/324) \leq \max(48/234, 88/432, 108/234) = 1/3$ (am folosit $-2 \leq x \leq 2,\ -1 \leq y \leq 1,\ 0 \leq z \leq 1$). Putem lua $q = 1/3$. Pe de altă parte $|f(x,y,z) - a_1| \leq (|432| + |x^4| + |y^4| + |2z^4| + |y^2z^2|)/234 \leq 452/234 < 2$, $|g(x,y,z) - a_2| \leq (|324| + |x^4| + |2y^4| + |3z^4| + |3x^2z^2|)/432 \leq 357/432 < 1$ şi $|h(x,y,z) - a_3| \leq (|234| + |2x^4| + |3y^4| + |4z^4| + |x^2y^2| + |3x| + |y|)/324 \leq 122/324 < 1/2$, $\forall\ (x,y,z) \in V$. Deci $F(V) \subset V$ şi $F$ este 1/3-contracţie a mulţimii $V$ în raport cu norma $\|\cdot\|_\infty$. Din principiul contracţiei rezultă că ecuaţia $F(x,y,z) = (x,y,z)$ ar soluţie unică în $V$.

Pentru $x^{(0)} = (0,0,1/2)$ se calculează $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)})$ cu formula de recurenţă $x_1^{(n)} = f(x^{(n-1)})$, $x_2^{(n)} = g(x_1^{(n)}, x_2^{(n-1)}, x_3^{(n-1)})$, $x_3^{(n)} = h(x_1^{(n)}, x_2^{(n)}, x_3^{(n-1)})$ $\forall\ n \in N$. Se opreşte procesul iterativ atunci când $q/(1-q)\|x^{(n)} - x^{(n-1)}\|_\infty < \varepsilon$.

## Metoda lui Newton

**Exemplul -1.29** *Fie sistemul de ecuaţii*

$$\begin{cases} x^3 + 3xy^2 + 3xz^2 - 183x + 2 = 0 \\ y^3 + +3x^2y + 3yz^2 - 193y + 1 = 0 \\ z^3 + 3x^2z + 3y^2z - 188(z-1) - 1 = 0 \end{cases} \tag{5}$$

*Folosind metoda Newton relativă la norma $\|\cdot\|_\infty$, să se arate că sistemul (5) are soluţie unică $v$ în $V := [-1,1]^2 \times [0,2]$. Pentru o eroare $\boldsymbol{\varepsilon}$ dată şi pentru iterata iniţială $x^{(0)} := (0,0,1)$ să se calculeze iterata $x^{(n)}$, pentru care $\|x^{(n)} - v\|_\infty \leq \boldsymbol{\varepsilon}$.*

Fie $r = 1$ şi $a = (0,0,1)$. Atunci mulţimea $V$ este $\overline{B_r(a)}$ în raport cu norma $\|\cdot\|_\infty$. Fie funcţia $F = (f,g,h) : V \to \mathbf{R}^3$ cu $f(x,y,z) := x^3 + 3xy^2 + 3xz^2 - 183x + 2$, $g(x,y,z) := y^3 + +3x^2y + 3yz^2 - 193y + 1$ şi $h(x,y,z) := z^3 + 3x^2z + 3y^2z - 188(z-1) - 1 = 0$, $\forall\,(x,y,z) \in V$. Sistemul de ecuaţii (5) este echivalent cu $F(x,y,z) = 0$.

Avem $F'(x,y,z) =$

$$\begin{pmatrix} 3x^2+3y^2+3z^2-183 & 6xy & 6xz \\ 6xy & 3x^2+3y^2+3z^2-193 & 6yz \\ 6xz & 6yz & 3x^2+3y^2+3z^2-188 \end{pmatrix},$$

$$F''(x,y,z) = \begin{pmatrix} 6x & 6y & 6z & 6y & 6x & 0 & 6z & 0 & 6x \\ 6y & 6x & 0 & 6x & 6y & 6z & 0 & 6z & 6x \\ 6z & 0 & 6x & 0 & 6z & 6y & 6x & 6y & 6z \end{pmatrix},$$

deci $\|F''(x,y,z)\|_\infty \leq \max(|18x| + |12y| + |12z|, |12x| + |18y| + |12z|, |12x| + |12y| + |18z|) \leq \max(54, 54, 60) = 60$ (am folosit $-1 \leq x, y \leq 1,\ 0 \leq z \leq 2$). Putem lua $M = 60$. Avem $F'(a) = \begin{pmatrix} -180 & 0 & 0 \\ 0 & -190 & 0 \\ 0 & 0 & -185 \end{pmatrix}$, deci $F'(a)$ inversabilă, $(F'(a))^{-1} =$ $\begin{pmatrix} -1/180 & 0 & 0 \\ 0 & -1/190 & 0 \\ 0 & 0 & -1/185 \end{pmatrix}$ şi $\|(F'(a))^{-1}\|_\infty = \max(1/180, 1/190, 1/185) = 1/180.$

Luăm $\alpha = 1/180$. Cum $Mr\alpha = 60/180 < 1$ şi $\alpha/(1 - MR\alpha) = 1/120$, putem lua $\mu = 120$. Avem $F(a) = (2, 1, 0)^t$ şi $(F'(a))^{-1}F(a) = (-1/90, -1/190, 0)^t$, deci $\|(F'(a))^{-1}F(a)\|_\infty = 1/90$. Luă m $\beta = 1/90$. Condiţia $\beta \leq (1-q)r$ implică $q \leq 89/90$. Din $\|F(x,y,z)\|_\infty \leq \max(|x^3| + |3xy^2| + |3xz^2| + |183x| + 2, |y^3| + |3x^2y| + |3yz^2| + |193y| + 1, |z^3| + |3x^2z| + |3y^2z| + |188(z-1)| + 1) \leq \max(201, 210, 209) = 210$ rezultă că $(M/\mu^2)\|F(x,y,z)\|_\infty \leq q \Leftrightarrow q \geq 210/240 = 7/8$. Din $Mr/\mu \leq q$ rezultă $q \geq 1/2$. Deci putem lua $q = 7/8$ şi atunci ecuaţia $F(x,y,z) = 0$ are o soluţie unică în $V$.

Pentru $x^{(0)} = (0,0,1)$ se calculează $x^{(n)}$ cu formula de recurenţă (5) (folosind la fiecare pas o procedură de inversare a matricei $(F'(x^{(n)}))^{-1}$). Se opreşte calculul iterativ atunci când $(2\mu)/M \cdot q^{2^n} \leq \boldsymbol{\varepsilon} \Leftrightarrow 4 \cdot (7/8)^{2^n} \leq \boldsymbol{\varepsilon}$.

## Metoda lui Newton simplificata

**Exemplul -1.30** *Fie sistemul de ecuaţii*

$$\begin{cases} x^2 + y^2 + z^2 + x + y + z - 853/144 = 0 \\ 3x^2 - 4y^2 + z^2 + x - y - 3z + 1309/432 = 0 \\ 5x^2 + 2y^2 - 3z^2 + 2x - 4y - 6z + 37/9 = 0 \end{cases} \tag{6}$$

*Folosind metoda Newton simplificată relativă la norma* $\|\cdot\|_\infty$*, să se arate că sistemul (6) are soluţie unică* $v$ *în* $V := [-0,8; 1,2]^3$*. Pentru o eroare* $\boldsymbol{\varepsilon}$ *dată şi pentru iterata iniţială* $x^{(0)} := (1,1,1)$*, să se calculeze iterata* $x^{(n)}$ *pentru care* $\|x^{(n)} - v\|_\infty \leq \boldsymbol{\varepsilon}$*.*

Fie $r := 0,2 = 1/5$ şi $a := (1,1,1)$. Atunci mulţimea $V$ este $\overline{B_r(a)}$ în raport cu norma $\|\cdot\|_\infty$. Fie funcţia $F = (f,g,h) : V \to \mathbf{R}^3$ cu $f(x,y,z) := x^2 + y^2 + z^2 + x + y + z - 853/144$, $g(x,y,z) := 3x^2 - 4y^2 + z^2 + x - y - 3z + 1309/432$ şi $h(x,y,z) := 5x^2 + 2y^2 - 3z^2 + 2x - 4y - 6z + 37/9$, $\forall\,(x,y,z) \in V$. Sistemul de ecuaţii (6) este echivalent cu $F(x,y,z) = 0$.

Avem

$$F'(x,y,z)=\begin{pmatrix} 2x+1 & 2y+1 & 2z+1 \\ 6x+1 & -8y-1 & 2z-3 \\ 10x+2 & 4y-4 & -6z-6 \end{pmatrix},$$

$$F''(x,y,z)=\begin{pmatrix} 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 6 & 0 & 0 & 0 & -8 & 0 & 0 & 0 & 2 \\ 10 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & -6 \end{pmatrix},$$

deci $\|F''(x,y,z)\|_\infty \le \max(|2|+|2|+|2|, |6|+|-8|+|2|, |10|+|4|+|-6|) = 20, \forall\, (x,y,z) \in V$. Putem lua $M=20$. Avem $F'(a)=\begin{pmatrix} 3 & 9 & 3 \\ 7 & -9 & -1 \\ 12 & 0 & -12 \end{pmatrix}$, cu $\det(F'(a))=72$, deci $F'(a)$ inversabilă, $(F'(a))^{-1}=\dfrac{1}{72}\begin{pmatrix} 9 & 3 & 2 \\ 6 & -6 & 2 \\ 9 & 3 & -4 \end{pmatrix}$ şi $\|(F'(a))^{-1}\|_\infty = \max(14/72, 14/72, 16/72) = 2/9$. Luă m $\alpha = 2/9$. Avem $F(a)=(11/144, 13/432, 1/9)^t$ şi $(F'(a))^{-1}F(a) = (1/72, 1/144, 1/216)^t$, deci $\|(F'(a))^{-1}F(a)\|_\infty = 1/72$. Luăm $\beta = 1/72$. Condiţia $\beta \le (1-q)r$ implică $q \le 67/72$. Din $\alpha M r \le q$ rezultă $q \ge 8/9$. Deci putem lua $q = 8/9$ şi atunci ecuaţia $F(x,y,z)=0$ are o soluţie unică în $V$.

Pentru $x^{(0)}=(1,1,1)$ se calculează $x^{(n)}$ cu formula de recurenţă (6). Se opreşte calculul iterativ atunci când $q/(1-q)\cdot\|x^{(n)}-x^{(n-1)}\|_\infty \le \boldsymbol{\varepsilon}$.

# Interpolare polinomială

## Formula lui Lagrange

**Exemplul -1.31** *Utilizând formula lui Lagrange de reprezentare să se determine polinomul de interpolare de grad 2 ce verifică condiţiile:*

$$P(0)=2,\ P(1)=1,\ P(4)=4.$$

Din curs rezultă că:

$$P(x) = P(0)L_1(x) + P(1)L_2(x) + P(4)L_3(x).$$

Avem:

$$L_1(x) = \frac{(x-1)(x-4)}{(0-1)(0-4)} = \frac{x^2-5x+4}{4},$$
$$L_2(x) = \frac{x(x-4)}{(1-0)(1-4)} = \frac{x^2-4x}{-3},$$
$$L_3(x) = \frac{x(x-1)}{(4-0)(4-1)} = \frac{x^2-x}{12}.$$

Atunci:

$$P(x) = 2\cdot\frac{x^2-5x+4}{4} + 1\cdot\frac{x^2-4x}{-3} + 4\cdot\frac{x^2-x}{12} = \frac{1}{2}x^2 - \frac{3}{2}x + 2.\Box$$

## Algoritmul lui Aitken

**Exemplul -1.32** *Fie funcţia* $f : [0, 12] \to R$ *cu* $f(x) = 6\sqrt{2x+1} - 2x$, $\forall\, x \in [0, 12]$. *Să se determine polinomul de interpolare Lagrange* $P(f; 0, 4, 12; x)$ *utilizând algoritmul Aitken.*

Nodurile sunt $x_1 = 0, x_2 = 4, x_3 = 12$. Atunci $P_{11}(x) = f(0) = 6$, $P_{12}(x) = f(4) = 10$, $P_{13}(x) = f(12) = 6$. Rezultă că

$$P_{22}(x) = \frac{10x - 6(x-4)}{4} = x + 6,$$

$$P_{23}(x) = \frac{6x - 6(x-12)}{12} = 6$$

şi astfel

$$P(f; 0, 4, 12; x) = P_{33}(x) = -\frac{x^2}{8} + \frac{3x}{2} + 6\,.$$

## Algoritmul lui Neville

**Exemplul -1.33** *Fie funcţia* $f : [0, 12] \to R$ *cu* $f(x) = 6\sqrt{2x+1} - 2x$, $\forall\, x \in [0, 12]$. *Să se determine polinomul de interpolare Lagrange* $P(f; 0, 4, 12; x)$ *utilizând algoritmul Neville.*

Nodurile sunt $x_1 = 0, x_2 = 4, x_3 = 12$. Atunci $P_{11}(x) = f(0) = 6$, $P_{12}(x) = f(4) = 10$, $P_{13}(x) = f(12) = 6$. Rezultă că

$$P_{22}(x) = \frac{10x - 6(x-4)}{4} = x + 6,$$

$$P_{23}(x) = \frac{6(x-4) - 10(x-12)}{8} = -\frac{x}{2} + 12$$

şi astfel

$$P(f; 0, 4, 12; x) = P_{33}(x) = -\frac{x^2}{8} + \frac{3x}{2} + 6$$

## Formula lui Newton de reprezentare a polinomului de interpolare

**Exemplul -1.34** *Fie funcţia* $f : [0, 4] \to R$ *cu* $f(x) = 6\sqrt{2x+1} - 2x$, $\forall\, x \in [0, 4]$. *Utilizând formula de reprezentare Newton să se determine polinomul de interpolare Hermite* $P(x) = P(f; 0, 0, 0, 4, 4; x)$. *Se se evalueze eroarea* $|f(x) - P(x)|$.

Avem $x_1 = x_2 = x_3 = 0$, $x_4 = x_5 = 4$. Din formula de reprezentare Newton avem:

$$P(f; 0, 0, 0, 4, 4; x) = f(0) + f(0,0)x + f(0,0,0)x^2 +$$

$$+ f(0,0,0,4)x^3 + f(0,0,0,4,4)x^3(x-4)$$

Vom calcula diferenţele divizate cu formulele de recurenţă date de (**??**). Calculele le vom organiza într-un tabel de forma:

| *nod* | *d.d.* 1$n$ | *d.d.* 2$n$ | *d.d.* 3$n$ | *d.d.* 4$n$ | *d.d.* 5$n$ |
|---|---|---|---|---|---|
| $x_1$ | $f(x_1)$ | $f(x_1,x_2)$ | $f(x_1,x_2,x_3)$ | $f(x_1,x_2,x_3,x_4)$ | $f(x_1,\ldots,x_5)$ |
| $x_2$ | $f(x_2)$ | $f(x_2,x_3)$ | $f(x_2,x_3,x_4)$ | $f(x_2,x_3,x_4,x_5)$ | |
| $x_3$ | $f(x_3)$ | $f(x_3,x_4)$ | $f(x_3,x_4,x_5)$ | | |
| $x_4$ | $f(x_4)$ | $f(x_4,x_5)$ | | | |
| $x_5$ | $f(x_5)$ | | | | |

(am notat cu "d.d.$k$n " diferenţe divizate asociate la $k$ noduri).

Cum $f(x) = 6\sqrt{2x+1} - 2x = 6(2x+1)^{1/2} - 2x$, rezultă că $f'(x) = 6(2x+1)^{-1/2} - 2$ şi $f''(x) = -6(2x+1)^{-3/2}$. Atunci $f(0) = 6$, $f(4) = 10$, $f(0,0) = f'(0) = 4$, $f(0,4) = (f(4) - f(0))/(4-0) = 1$, $f(4,4) = f'(4) = 0$, $f(0,0,0) = \frac{1}{2}f''(0) = -3$, $f(0,0,4) = (f(0,4) - f(0,0))/(4-0) = -3/4$ etc.. Rezultă următorul tabel în cazul datelor problemei:

| *nod* | 1$n$ | 2$n$ | 3$n$ | 4$n$ | 5$n$ |
|---|---|---|---|---|---|
| 0 | 6 | 4 | $-3$ | 9/16 | $-7/64$ |
| 0 | 6 | 4 | $-3/4$ | 2/16 | |
| 0 | 6 | 1 | $-1/4$ | | |
| 4 | 10 | 0 | | | |
| 4 | 10 | | | | |

Atunci (cu formula de reprezentare Newton) avem

$$P(f;0,0,0,4,4;x) = 6 + 4x - 3x^2 + \frac{9}{16}x^3 - \frac{7}{64}x^3(x-4).$$

Din formula de evaluare a erorii rezultă că

$$|f(x) - P(f;0,0,0,4,4;x)| \le \frac{M_5}{5!}|x^3(x-4)^2| \quad \forall\, x \in [0,4],$$

unde $M_5 = \sup\limits_{x\in[0,4]} |f^{(5)}(x)|$. Avem $f^{(3)}(x) = 18(2x+1)^{-5/2}$, $f^{(4)}(x) = -90(2x+1)^{-7/2}$, $f^{(5)}(x) = 630(2x+1)^{-9/2}$. Cum $f^{(6)}(x) = -630\cdot 9(2x+1)^{-11/2} < 0$, rezultă că $f^{(5)}$este descrescătoare şi cum $f^{(5)}(x) > 0\ \forall\, x \in [0,4]$ rezultă că $M_5 = f^{(5)}(0) = 630$. Atunci

$$|f(x) - P(x)| \leq \frac{630}{120}|x^3(x-4)^2| = \frac{21}{4}x^3(x-4)^2, \quad \forall\, x \in [0,4]. \square$$

## Formula lui Newton de interpolare ascendentă

**Exemplul -1.35** *Să se determine, folosind formula de interpolare Newton ascendentă, polinomul Lagrange ataşat datelor:* $n = 3$, $x_0 = 4$, $h = 2$ *(de unde* $x_1 = 6$, $x_2 = 8$, $x_3 = 10$*) şi* $f : R \to R$ *o funcţie continuă cu* $f(x_0) = 1$, $f(x_1) = 3$, $f(x_2) = 8$, $f(x_3) = 20$ *şi să se estimeze valoarea funcţiei* $f$ *în punctul* $x = 7$.

Organizăm calculul diferenţelor nedivizate ascendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\Delta_h^1 f(x_k)$ | $\Delta_h^2 f(x_k)$ | $\Delta_h^3 f(x_k)$ |
|---|---|---|---|---|---|
| 0 | 4 | 1 | 2 | 3 | 4 |
| 1 | 6 | 3 | 5 | 7 | |
| 2 | 8 | 8 | 12 | | |
| 3 | 10 | 20 | | | |

(De exemplu $\Delta_h^1 f(x_0) = \Delta_h f(x_0) = f(x_1) - f(x_0) = 3-1 = 2$, $\Delta_h^2 f(x_0) = \Delta_h^1 \Delta_h f(x_0) = \Delta_h f(x_1) - \Delta_h f(x_0) = \Delta_h^1 f(x_1) - \Delta_h^1 f(x_0) = 5-2 = 3$, etc.) Atunci $\Delta_h^0 f(x_0) = 1$, $\Delta_h^1 f(x_0) = 2$, $\Delta_h^2 f(x_0) = 3$, $\Delta_h^3 f(x_0) = 4$.

Fie $x \in \mathbf{R}$ Avem $q = \dfrac{x-x_0}{h} = \dfrac{x-4}{2}$şi atunci $C_q^0 = 1$, $C_q^1 = q = \dfrac{x-4}{2}$, $C_q^2 = \dfrac{q(q-1)}{2} = \dfrac{(x-4)(x-6)}{8}$, $C_q^3 = \dfrac{q(q-1)(q-2)}{6} = \dfrac{(x-4)(x-6)(x-8)}{48}$. Rezultă că: $P(f; x_0, x_1, x_2, x_3; x) =$

$$\sum_{i=0}^{3} C_q^i \Delta_h^i f(x_0) = 1 + (x-4) + \frac{3}{8}(x-4)(x-6) + \frac{4}{48}(x-4)(x-6)(x-8)$$

şi deci

$$P(f; x_0, x_1, x_2, x_3; 7) = 1 + 3 + \frac{9}{8} - \frac{1}{4} = \frac{39}{8} = 4.875\ .\square$$

## Formula lui Newton de interpolare descendentă

**Exemplul -1.36** *Să se determine, folosind formula de interpolare Newton descendentă, polinomul Lagrange ataşat datelor:* $n = 3$, $x_0 = 10$, $h = 2$ *(de unde* $x_{-1} = 8$, $x_{-2} = 6$, $x_{-3} = 4$*) şi* $f : \mathbf{R} \to \mathbf{R}$ *o funcţie continuă cu* $f(x_0) = 20$, $f(x_{-1}) = 8$, $f(x_{-2}) = 3$, $f(x_{-3}) = 1$ *şi să se estimeze valoarea funcţiei* $f$ *în punctul* $x = 7$.

Organizăm calculul diferenţelor nedivizate descendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\nabla_h^1 f(x_k)$ | $\nabla_h^2 f(x_k)$ | $\nabla_h^3 f(x_k)$ |
|---|---|---|---|---|---|
| 0 | 10 | 20 | 12 | 7 | 4 |
| $-1$ | 8 | 8 | 5 | 3 | |
| $-2$ | 6 | 3 | 2 | | |
| $-3$ | 4 | 1 | | | |

(De exemplu $\nabla_h^1 f(x_0) = \nabla_h f(x_0) = f(x_0) - f(x_{-1}) = 20 - 8 = 12$, $\nabla_h^2 f(x_0) = \nabla_h^1 \nabla_h f(x_0) = \nabla_h f(x_0) - \nabla_h f(x_{-1}) = \nabla_h^1 f(x_0) - \nabla_h^1 f(x_{-1}) = 12 - 5 = 7$, etc.) Atunci $\nabla_h^0 f(x_0) = 20$, $\nabla_h^1 f(x_0) = 12$, $\nabla_h^2 f(x_0) = 7$, $\nabla_h^3 f(x_0) = 4$.

Fie $x \in \mathbf{R}$. Avem $q = \dfrac{x - x_0}{h} = \dfrac{x - 10}{2}$ şi atunci $C_{-q}^0 = 1$, $C_{-q}^1 = -q = -\dfrac{x-10}{2}$, $C_{-q}^2 = \dfrac{-q(-q-1)}{2} = \dfrac{(x-10)(x-8)}{8}$, $C_{-q}^3 = \dfrac{-q(-q-1)(-q-2)}{6} = \dfrac{(x-10)(x-8)(x-6)}{48}$. Rezultă că:

$$P(f; x_0, x_{-1}, x_{-2}, x_{-3}; x) = \sum_{i=0}^{3} (-1)^i C_{-q}^i \nabla_h^i f(x_0) =$$

$$= 20 + \frac{12}{2}(x - 10) + \frac{7}{8}(x-10)(x-8) + \frac{4}{48}(x-10)(x-8)(x-6)$$

şi deci

$$P(f; x_0, x_{-1}, x_{-2}, x_{-3}; 7) = 20 - 18 + \frac{21}{8} + \frac{1}{4} = \frac{39}{8} = 4.875\ .\square$$

## Formula de interpolare Gauss înainte

**Exemplul -1.37** *Să se determine, folosind formula de interpolare Gauss înainte, polinomul Lagrange ataşat datelor:* $n = 2$, $x_0 = 6$, $h = 2$ *(de unde* $x_{-2} = 2$, $x_{-1} = 4$, $x_1 = 8$, $x_2 = 10$*) şi* $f : R \to R$ *o funcţie continuă cu* $f(x_{-2}) = -2$, $f(x_{-1}) = 1$, $f(x_0) = 3$, $f(x_1) = 8$, $f(x_2) = 20$ *şi să se estimeze valoarea funcţiei* $f$ *în punctul* $x = 7$.

Printr-un calcul elementar, se arată că, pentru orice $i = \overline{1, 2}$:

$$\delta_h^{2i-1} f(x_{1/2}) = \Delta_h^{2i-1} f(x_{-i+1}) \text{ şi } \delta_h^{2i} f(x_0)) = \Delta_h^{2i} f(x_{-i})).$$

Atunci: $P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; x) =$

$$f(x_0) + \sum_{i=1}^{2} (C_{q+i-1}^{2i-1} \Delta_h^{2i-1} f(x_{-i+1}) + C_{q+i-1}^{2i} \Delta_h^{2i} f(x_{-i})) =$$

$$= f(x_0) + C_q^1 \Delta_h^1 f(x_0) + C_q^2 \Delta_h^2 f(x_{-1}) + C_{q+1}^3 \Delta_h^3 f(x_{-1}) + C_{q+1}^4 \Delta_h^4 f(x_{-2}).$$

Organizăm calculul diferenţelor finite ascendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\Delta_h^1 f(x_k)$ | $\Delta_h^2 f(x_k)$ | $\Delta_h^3 f(x_k)$ | $\Delta_h^4 f(x_k)$ |
|---|---|---|---|---|---|---|
| $-2$ | 2 | $-2$ | 3 | $-1$ | 4 | 0 |
| $-1$ | 4 | 1 | 2 | 3 | 4 | |
| 0 | 6 | 3 | 5 | 7 | | |
| 1 | 8 | 8 | 12 | | | |
| 2 | 10 | 20 | | | | |

Avem: $f(x_0) = 3, \Delta_h^1 f(x_0) = 5, \Delta_h^2 f(x_{-1}) = 3, \Delta_h^3 f(x_{-1}) = 4, \Delta_h^4 f(x_{-2}) = 0$. Cum $q = \dfrac{x-6}{2}$ rezultă că $C_q^1 = q = \dfrac{x-6}{2}$, $C_q^2 = \dfrac{q(q-1)}{2} = \dfrac{(x-6)(x-8)}{8}$, $C_{q+1}^3 = \dfrac{(q+1)q(q-1)}{6} = \dfrac{(x-4)(x-6)(x-8)}{48}$. Atunci:

$$P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; x) =$$

$$= 3 + 5\frac{x-6}{2} + 3\frac{(x-6)(x-8)}{8} + 4\frac{(x-4)(x-6)(x-8)}{48}$$

şi

$$f(7) \simeq P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; 7) = 3 + \frac{5}{2} - \frac{3}{8} - \frac{1}{4} = \frac{39}{8} = 4.875 \ .\square$$

## Formula de interpolare Gauss înapoi

**Exemplul -1.38** *Să se determine, folosind formula de interpolare Gauss înapoi, polinomul Lagrange ataşat datelor:* $n = 2$, $x_0 = 6$, $h = 2$ *(de unde* $x_{-2} = 2$, $x_{-1} = 4$, $x_1 = 8$, $x_2 = 10$*) şi* $f : R \to R$ *o funcţie continuă cu* $f(x_{-2}) = -2$, $f(x_{-1}) = 1$, $f(x_0) = 3$, $f(x_1) = 8$, $f(x_2) = 20$ *şi să se estimeze valoarea funcţiei* $f$ *în punctul* $x = 7$.

Printr-un calcul elementar, se arată că, pentru orice $i = \overline{1,2}$:

$$\delta_h^{2i-1} f(x_{1/2}) = \Delta_h^{2i-1} f(x_{-i+1}) \text{ şi } \delta_h^{2i} f(x_0)) = \Delta_h^{2i} f(x_{-i})).$$

Atunci: $P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; x) =$

$$f(x_0) + \sum_{i=1}^{2} (C_{q+i-1}^{2i-1} \Delta_h^{2i-1} f(x_{-i}) + C_{q+i}^{2i} \Delta_h^{2i} f(x_{-i})) =$$

$$= f(x_0) + C_q^1 \Delta_h^1 f(x_{-1}) + C_{q+1}^2 \Delta_h^2 f(x_{-1}) + C_{q+1}^3 \Delta_h^3 f(x_{-2}) + C_{q+2}^4 \Delta_h^4 f(x_{-2}).$$

Organizăm calculul diferenţelor finite ascendente în tabelul următor:

| $k$ | $x_k$ | $f(x_k)$ | $\Delta_h^1 f(x_k)$ | $\Delta_h^2 f(x_k)$ | $\Delta_h^3 f(x_k)$ | $\Delta_h^4 f(x_k)$ |
|---|---|---|---|---|---|---|
| $-2$ | 2 | $-2$ | 3 | $-1$ | 4 | 0 |
| $-1$ | 4 | 1 | 2 | 3 | 4 | |
| 0 | 6 | 3 | 5 | 7 | | |
| 1 | 8 | 8 | 12 | | | |
| 2 | 10 | 20 | | | | |

Avem: $f(x_0) = 3, \Delta_h^1 f(x_{-1}) = 2, \Delta_h^2 f(x_{-1}) = 3, \Delta_h^3 f(x_{-2}) = 4, \Delta_h^4 f(x_{-2}) = 0$. Cum $q = \dfrac{x-6}{2}$, rezultă că $C_q^1 = q = \dfrac{x-6}{2}$, $C_{q+1}^2 = \dfrac{(q+1)q}{2} = \dfrac{(x-4)(x-6)}{8}$, $C_{q+1}^3 = \dfrac{(q+1)q(q-1)}{6} = \dfrac{(x-4)(x-6)(x-8)}{48}$. Atunci:

$$P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; x) =$$

$$= 3 + 2\frac{x-6}{2} + 3\frac{(x-4)(x-6)}{8} + 4\frac{(x-4)(x-6)(x-8)}{48}$$

şi

$$f(7) \simeq P(f; x_{-2}, x_{-1}, x_0, x_1, x_2; 7) = 3 + 1 + \frac{9}{8} - \frac{1}{4} = \frac{39}{8} = 4.875\ . \square$$

# Metode numerice pentru ecuaţii diferenţiale

# Metode într-un pas (directe)

## Metoda Euler explicită (progresivă)

**Exemplul -1.39** *Folosind metoda Euler, determinaţi soluţia aproximativă a problemei Cauchy:*

$$\begin{cases} y' = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}\,, & t \in [1,2] \\ y(1) = 0.75 \end{cases}$$

*în $t = 2$, luând $h = 0.5$.*

Avem $f(t,y) = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}$ , $t_0 = 1$ şi $y_0 = 0.75$ . Se obţine $N = 2$ şi $y_{n+1} = y_n + 0.5 f(t_n, y_n)$, pentru $n = 0, 1$ . Deci rezultă:

$$t_1 = 1.5\ ,\ y(t_1) \simeq y_1 = y_0 + 0,5 f(t_0, y_0) = 0.75 + 0.5 f(1, 0.75) =$$
$$0.75 + 0.5(-0.4375) = 0.5313\,,$$

$$t_2 = 2\ ,\ y(t_2) \simeq y_2 = y_1 + 0,5 f(t_1, y_1) = 0.5313 + 0.5 f(1.5, 0.5313) =$$
$$0.5313 + 0.5(-0.1831) = 0.4397\ .\square$$

## Metode Runge-Kutta

**Exemplul -1.40** *Se dă problema Cauchy:*

$$\begin{cases} y' = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}, & t \in [1, 2.9] \\ y(1) = 0.75 \end{cases}$$

*Să se determine valoarea aproximativă a soluţiei în $t = 2.9$, folosind metoda şi $RK_{24}$, pentru $N = 2$.*

Pentru $N = 2$ avem $h = 0.9500$. Luăm $y_0 = 0.75$, $t_0 = 1$. Pentru metoda $RK_{24}$, avem $f(t, y) = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}$. Se obţine:

- pentru $n = 1$:

$$y_{0,1} = y_0 = 1\ , \quad t_{0,1} = t_0 = 1\ ,$$

$$y_{0,2} = y_0 = 1 + 0.95 f(t_{0,1}, y_{0,1})/2 = 0.8812\ , \quad t_{0,2} = t_0 + 0.95/2 = 1.4750\ ,$$

$$y_{0,3} = y_0 = 1 + 0.95 f(t_{0,2}, y_{0,2})/2 = 1.0305\ , \quad t_{0,3} = t_0 + 0.95/2 = 1.4750\ ,$$

$$y_{0,4} = y_0 = 1 + 0.95 f(t_{0,3}, y_{0,3}) = 1.2360\ , \quad t_{0,4} = t_0 + 0.95 = 1.9500\ ,$$

$$y_1 = y_0 + 0.95(f(t_{0,1}, y_{0,1}) + 2f(t_{0,2}, y_{0,2}) + 2f(t_{0,3}, y_{0,3}) + f(t_{0,4}, y_{0,4}))/6 =$$

$$1 + 0.95(-0.2500 + 2 \cdot 0.0642 + 2 \cdot 0.2484 + 0.8281)/6 = 1.1905$$

$$t_1 = t_0 + h = 1.9500$$

- pentru $n = 2$:

$$y_{1,1} = y_1 = 1.1905\ , \quad t_{1,1} = t_1 = 1.95\ ,$$

$$y_{1,2} = y_1 = 1 + 0.95 f(t_{1,1}, y_{1,1})/2 = 1.5425\ , \quad t_{1,2} = t_1 + 0.95/2 = 2.4250\ ,$$

$$y_{1,3} = y_1 = 1 + 0.95 f(t_{1,2}, y_{1,2})/2 = 1.9984\ , \quad t_{1,3} = t_1 + 0.95/2 = 2.4250\ ,$$

$$y_{1,4} = y_1 = 1 + 0.95 f(t_{1,3}, y_{1,3}) = 4.1612\ , \quad t_{1,4} = t_1 + 0.95 = 2.9\ ,$$

$$y_2 = y_1 + 0.95(f(t_{1,1}, y_{1,1}) + 2f(t_{1,2}, y_{1,2}) + 2f(t_{1,3}, y_{1,3}) + f(t_{1,4}, y_{1,4}))/6 =$$

$$1.1905 + 0.95(0.7411 + 2 \cdot 1.7008 + 2 \cdot 3.1270 + 15.8509)/6 = 5.3464$$

$$t_2 = t_1 + h = 2.9\ .\ \square$$

# Metode în mai mulţi paşi (indirecte)

## Metoda Adams-Bashforth

1

**Exemplul -1.41** *Folosind metoda Adams-Bashforth pentru $r = 3$, determinaţi soluţia aproximativă a problemei Cauchy:*

$$\begin{cases} y' = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}, & t \in [1, 2.25] \\ y(1) = 0.75 \end{cases}$$

*în $t = 2.25$ pentru $N = 5$, determinând valorile de început $y_1, y_2, y_3$ cu metoda Euler progresivă.*

Avem $h = 0.25$ şi $f(t, y) = y^2 - \dfrac{y}{t} - \dfrac{1}{4t^2}$. Aplicând metoda Euler progresivă se obţine:

$$\begin{aligned}
&t_0 = 1, && y_0 = 0.75, && f_0 = f(t_0, y_0) = -0.4375, \\
&t_1 = 1.25, && y_1 = 0.6406, && f_1 = f(t_1, y_1) = -0.2621, \\
&t_2 = 1.5, && y_2 = 0.5751, && f_2 = f(t_2, y_2) = -0.1638, \\
&t_3 = 1.75, && y_3 = 0.5342, && f_3 = f(t_3, y_3) = -0.1015,
\end{aligned}$$

şi, în continuare, cu metoda Adams-Bashforth pentru $r = 3$, rezultă:

$$t_4 = 2, \qquad y_4 = y_3 + \frac{h}{24}(55f_3 - 59f_2 + 37f_1 - 9f_0) = 0.5166,$$

$$f_4 = f(t_4, y_4) = -0.0539,$$

$$t_5 = 2.25, \qquad y_5 = y_4 + \frac{h}{24}(55f_4 - 59f_3 + 37f_2 - 9f_1) = 0.5096,$$

$$f_5 = f(t_5, y_5) = -0.0162.$$