1. Sisteme de ecuații liniare

Reamintim că un sistem de n ecuații algebrice liniare cu n necunoscute este de forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{1}$$

Dacă notăm cu A matricea coeficienților, cu x vectorul coloană format cu necunoscutele sistemului şi cu b coloana termenilor liberi, sistemul (1) se scrie sub formă matriceală :

$$Ax=b, \tag{2}$$

unde:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Metodele numerice de rezolvare a sistemelor algebrice de ecuații liniare sunt de două tipuri: *metode directe* şi *metode indirecte* (sau *iterative*).

Metodele directe constau în transformarea sistemului (1) într–un sistem triunghiular echivalent, care se rezolvă uşor. Cele mai cunoscute metode directe sunt: *metoda Gauss*, *metoda Cholesky* (utilizată pentru sistemele în care matricea *A* este simetrică şi pozitiv definită) şi *metoda Householder*.

Metodele directe permit determinarea soluției exacte a sistemului în cazul ideal, când nu avem erori de rotunjire. Numărul operațiilor aritmetice efectuate este de ordinul n^3. Pentru sisteme cu un număr de ecuații mai mare de 100, metodele directe devin inutilizabile datorită acumulării erorilor de rotunjire care alterează soluția.

Metodele indirecte (sau *iterative*) constau în construcția unui şir $\{x^{(k)}\}$ de vectori *n*–dimensionali, care converge la soluția exactă a sistemului. Se alege ca soluție aproximativă a sistemului un termen $x^{(s)}$ al şirului, al cărui ordin depinde de precizia impusă.

O iterație presupune efectuarea unui număr de operații aritmetice de ordinul n^2. Metodele iterative sunt utilizate la rezolvarea sistemelor mari de ecuații. Cele mai cunoscute metode iterative sunt: *Jacobi, Gauss–Seidel, metodele de relaxare*.

§1.1. Metoda Gauss. Factorizarea LU

Fie

$$m_r = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ m_{r+1,r} \\ \vdots \\ m_{n,r} \end{pmatrix} \quad \text{şi} \quad e_r = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(elementul 1 din e_r se află pe linia r).
O matrice de forma $M_r = I_n - m_r \cdot e_r^{\,T}$, unde $e_r^T = (0, \dots, 1, \dots, 0)$, se numeşte *matrice Frobenius*. O astfel de matrice are următoarea structură:

$$M_r = \begin{pmatrix} 1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & & & \\ 0 & & 1 & 0 & \cdots & \\ 0 & \cdots & -m_{r+1,r} & 1 & \cdots & 0 \\ & & \vdots & & & \\ 0 & \cdots & -m_{nr} & 0 & \cdots & 1 \end{pmatrix}$$

De exemplu, dacă $n = 4$ şi $r = 2$, avem:

$$M_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ m_{32} \\ m_{42} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & m_{32} & 0 & 0 \\ 0 & m_{42} & 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -m_{32} & 1 & 0 \\ 0 & -m_{42} & 0 & 1 \end{pmatrix}$$

Propoziția 1. *Orice matrice Frobenius* M_r *este inversabilă şi inversa sa este:*

$$M_r^{-1}=I_n+m_r\cdot e_r^T.$$

Demonstrație.

$$(I_n-m_r\cdot e_r^T)(I_n+m_r\cdot e_r^T)=I_n-m_r\cdot e_r^T+m_r\cdot e_r^T-m_r(e_r^T m_r)\,e_r^T.$$

Deoarece $e_r^T\cdot m_r=0$, rezultă:

$$M_r(I_n+m_r\cdot e_r^T)=I_n,\quad \text{şi deci}\quad M_r^{-1}=I_n+m_r\cdot e_r^T. \qquad \square$$

Teorema 1. *Fie A o matrice pătrată de ordinul n care satisface condiția:*

$$(*)\qquad \det\begin{pmatrix} a_{11} & \dots & a_{1r}\\ \vdots & & \vdots\\ a_{r1} & \dots & a_{rr}\end{pmatrix}\neq 0 \quad \textit{pentru orice}\quad r=\overline{1,n-1}.$$

Atunci există o matrice inferior triunghiulară $M\in \mathrm{M}_n(\mathrm{R})$ *astfel încât matricea* $U=MA$ *este superior triunghiulară.*

Demonstrație. Deoarece $a_{11}\neq 0$, putem considera matricea Frobenius

$$M_1=\begin{pmatrix} 1 & 0 & \dots & & 0\\ -\frac{a_{21}}{a_{11}} & 1 & 0 & \dots & 0\\ \vdots & & & & \vdots\\ -\frac{a_{n1}}{a_{11}} & 0 & \dots & & 1\end{pmatrix}.$$

Dacă notăm $A_1=A$ şi $A_2=M_1A_1$, atunci avem

$$A_2=\begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \dots & a_{1n}^{(2)}\\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)}\\ \vdots & \vdots & & \vdots\\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)}\end{pmatrix},$$

unde, notând cu $a_{ij}^{(1)}=a_{ij}$, pentru $i,j=\overline{1,n}$, avem: $a_{1j}^{(2)}=a_{1j}^{(1)}$ pentru $j=\overline{1,n}$; $a_{ij}^{(2)}=a_{ij}^{(1)}-\dfrac{a_{i1}^{(1)}a_{1j}^{(1)}}{a_{11}^{(1)}}$, pentru orice $i,j=\overline{2,n}$.

Observăm că

$$a_{22}^{(2)}=a_{22}-\frac{a_{21}a_{12}}{a_{11}}=\frac{1}{a_{11}}\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}\neq 0.$$

Dacă notăm

$$M_2=\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & -\dfrac{a_{32}^{(2)}}{a_{22}^{(2)}} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & -\dfrac{a_{n2}^{(2)}}{a_{22}^{(2)}} & 0 & \dots & 1 \end{pmatrix},$$

atunci

$$A_3=M_2A_2=\begin{pmatrix} a_{11}^{(3)} & a_{12}^{(3)} & a_{13}^{(3)} & \dots & a_{1n}^{(3)} \\ 0 & a_{22}^{(3)} & a_{23}^{(3)} & \dots & a_{2n}^{(3)} \\ 0 & 0 & a_{33}^{(3)} & \dots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_{nn}^{(3)} \end{pmatrix},$$

unde $a_{ij}^{(3)}=a_{ij}^{(2)}$ pentru i=1, 2, $j=\overline{1,n}$ şi $a_{ij}^{(3)}=a_{ij}^{(2)}-\dfrac{a_{i2}^{(2)}a_{2j}^{(2)}}{a_{22}^{(2)}}$, $i,j=\overline{3,n}$.

Un calcul simplu ne arată că

$$a_{33}^{(3)}=\frac{1}{a_{11}^{(1)}a_{22}^{(2)}}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}\neq 0.$$

În general, $a_{rr}^{(r)}\neq 0$ şi se poate considera matricea Frobenius:

$$M_r=\begin{pmatrix} 1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & & & \\ 0 & & 1 & 0 & \cdots & \\ 0 & \cdots & -\dfrac{a_{r+1,r}^{(r)}}{a_{rr}^{(r)}} & 1 & \cdots & 0 \\ & & \vdots & & & \\ 0 & \cdots & -\dfrac{a_{nr}^{(r)}}{a_{rr}^{(r)}} & 0 & \cdots & 1 \end{pmatrix}.$$

Dacă notăm cu $A_{r+1}=M_rA_r$, atunci

$$A_{r+1}=\begin{pmatrix} a_{11}^{(r+1)} & a_{12}^{(r+1)} & \dots & a_{1r}^{(r+1)} & \dots & a_{1n}^{(r+1)} \\ 0 & a_{22}^{(r+1)} & \dots & a_{2r}^{(r+1)} & \dots & a_{2n}^{(r+1)} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & a_{rr}^{(r+1)} & \dots & a_{rn}^{(r+1)} \\ 0 & 0 & \dots & 0 & a_{r+1,r+1}^{(r+1)} \;\dots & a_{r+1,n}^{(r+1)} \\ \vdots & \vdots & & \vdots & \vdots \quad . & \vdots \\ 0 & 0 & \dots & 0 & a_{n,r+1}^{(r+1)} \;\dots & a_{nn}^{(r+1)} \end{pmatrix},$$

unde $a_{ij}^{(r+1)}=a_{ij}^{(r)}$, pentru $i=\overline{1,r}$, $j=\overline{1,n}$, $a_{ij}^{(r+1)}=a_{ij}^{(r)}-\dfrac{a_{ir}^{(r)}a_{rj}^{(r)}}{a_{rr}^{(r)}}$,

$i,j=\overline{r+1,n}$.

În final se obține matricea superior triunghiulară

$$U=A_n=M_{n-1}...M_2M_1A=\begin{pmatrix} a_{11}^{(n)} & a_{12}^{(n)} & \dots & a_{1n}^{(n)} \\ 0 & a_{22}^{(n)} & \dots & a_{2n}^{(n)} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn}^{(n)} \end{pmatrix}.$$

Notăm cu $M=M_{n-1}M_{n-2}...M_2M_1$ şi demonstrația teoremei este completă. ❑

Exemplu.

$$A_1=A=\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix},\qquad M_1=\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{pmatrix},\qquad A_2=\begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & \frac{18}{5} & \frac{9}{5} \end{pmatrix},$$

$$M_2=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{9}{20} & 1 \end{pmatrix},\ U=A_3=\begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix},\ M=M_2M_1=\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{7}{20} & \frac{9}{20} & 1 \end{pmatrix}.$$

Considerăm sistemul

$$\begin{cases} 5x_1+2x_2+x_3=12 \\ 5x_1-6x_2+2x_3=-1 \\ -4x_1+2x_2+x_3=3 \end{cases},$$

a cărui soluție este $x_1=1$, $x_2=2$, $x_3=3$. Sub formă matriceală sistemul se scrie: $Ax=b$, unde $b=\begin{pmatrix}12\\-1\\3\end{pmatrix}$. Acest sistem este echivalent cu următorul sistem: $(M_2M_1A)x=(M_2M_1)b$. Efectuând calculele obținem

$$\begin{cases} 5x_1+2x_2+x_3=12 \\ \quad -8x_2+x_3=-13 \\ \qquad \dfrac{9}{4}x_3=\dfrac{27}{4} \end{cases}.$$

Numărul operațiilor pentru determinarea matricei U și a vectorului Mb

Pentru o linie fixată i se calculează $-\dfrac{a_{ir}^{(r)}}{a_{rr}^{(r)}}$, apoi se fac înmulțirile cu $a_{rj}^{(r)}$, $r+1\le j\le n$, și se adună $a_{ij}^{(r)}, r+1\le j\le n$. La fel și cu $b_i^{(r+1)}$. Sunt $2(n-r)+3$ operații elementare pentru fiecare linie i, $r+1\le i\le n$, și pentru fiecare etapă r vor fi $(n-r)[2(n-r)+3]$ operații. În total vor fi $\sum_{r=1}^{n}\left[2(n-r)^2+3(n-r)\right]=\frac{2}{3}n^3+\frac{1}{2}n^2-\frac{7}{6}n$ operații elementare. Dacă adăugăm și cele n^2 operații pentru rezolvarea sistemului triunghiular, rezultă că numărul de operații pentru rezolvarea sistemului $Ax=b$ este $\frac{2}{3}n^3+\frac{3}{2}n^2-\frac{7}{6}n$.

În continuare notăm cu $L_r=M_r^{-1}$. Din Propoziția 1 rezultă că L_r este de forma:

$$L_r=\begin{pmatrix} 1 & \dots & 0 & \dots & & 0 \\ \vdots & \ddots & \vdots & & & \vdots \\ 0 & & 1 & \dots & & 0 \\ 0 & & \dfrac{a_{r+1,r}^{(r)}}{a_{rr}^{(r)}} & \dots & & 0 \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & \dfrac{a_{nr}^{(r)}}{a_{rr}^{(r)}} & 0 & \dots & 1 \end{pmatrix}.$$

Dacă notăm cu $L=L_1L_2...L_{n-1}$, atunci L este o matrice inferior triunghiulară de tipul următor

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \ell_{21} & 1 & .0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \dots & 1 \end{pmatrix}.$$

Deoarece $A = M_1^{-1} M_2^{-1} ... M_{n-1}^{-1} U$, rezultă că:

$$A=LU \qquad (3)$$

Aşadar, orice matrice pătratică ce îndeplineşte condiția (*) din Teorema 1 admite o descompunere unică de forma (3), unde L este inferior triunghiulară având elementele de pe diagonala principală egale cu 1 şi U este superior triunghiulară. Descompunerea (3) este cunoscută sub numele de *factorizarea LU*.

Algoritmul pentru factorizarea LU

{ Determinarea matricelor U şi L cu păstrarea matricei A }

Pentru i:=1,n execută

  Pentru j:=1,n execută

    $u_{ij}:=a_{ij}$;

    dacă $i=j$ atunci $1_{ii}:=1$ altfel $1_{ij}:=0$;

  sfârşit pentru j ;

sfârşit pentru i ;

Pentru r:=1,n–1 execută

  Pentru i:=r+1,n execută

    Pentru j:=r+1,n execută

$$u_{ij} := u_{ij} - \frac{u_{ir}u_{rj}}{u_{rr}} \quad ;$$

    sfârşit pentru j ;

$$1_{ir} := \frac{u_{ir}}{u_{rr}} \quad ;$$

  sfârşit pentru i ;

sfârşit pentru r ;

Pentru i:=2,n execută

  Pentru j:=1,i–1 execută

    u_{ij}:=0 ;

  sfârşit pentru j ;

sfârşit pentru i .

Algoritmul se află programat în MATLAB şi poate fi apelat cu secvența:

$[L,U]=lu(A)$ { se afişează cele două matrice }

În exemplul precedent avem:

$$L_1=\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ -\frac{4}{5} & 0 & 1\end{pmatrix},\quad L_2=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -\frac{9}{20} & 1\end{pmatrix},$$

$$A=LU=\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ -\frac{4}{5} & -\frac{9}{20} & 1\end{pmatrix}\cdot\begin{pmatrix}5 & 2 & 1\\ 0 & -8 & 1\\ 0 & 0 & \frac{9}{4}\end{pmatrix}=\begin{pmatrix}5 & 2 & 1\\ 5 & -6 & 2\\ -4 & 2 & 1\end{pmatrix}$$

Observația 1. *Dacă pivotul este "foarte mic" , adică* $\left|a_{rr}^{(r)}\right| << 1$, *atunci împărțirile la acest pivot produc erori de rotunjire foarte mari, care alterează soluția. În acest caz se recomandă schimbarea pivotului. Se poate alege un nou pivot*

$$\pi_r = \left|a_{ir}^{(r)}\right| = \max\left\{ \left|a_{ij}^{(r)}\right| \;;\; r \le j \le n \right\}$$

sau $$\pi_r = \left|a_{jr}^{(r)}\right| = \max\left\{ \left|a_{k\ell}^{(r)}\right| \;;\; r \le k,\ \ell \le n \right\}$$

Aceasta presupune schimbarea între ele a două linii şi eventual şi a două coloane.

Algoritmul Gauss pentru rezolvarea sistemelor de ecuații liniare

Pentru r:=1,n–1 execută
- Pentru i:=r+1,n execută
  - Pentru j:=r+1,n execută
    - găseşte pivotul conform cu (β) ;
    - schimbă linia i cu linia pivotului şi coloana j cu coloana pivotului dacă este cazul ;
  - sfârşit pentru j
- sfârşit pentru i
- Pentru i:=r+1,n execută
  - $$b_i = b_i - \frac{a_{ir}b_i}{a_{rr}}$$
  - Pentru j:=r+1,n execută
    - $$a_{ij} = a_{ij} - \frac{a_{ir}a_{rj}}{a_{rr}};\ ;$$
  - sfârşit pentru j ;
- sfârşit pentru i ;

sfârşit pentru r ;

$$x_n := \frac{b_n}{a_{nn}}\ ;$$

Pentru i:=n–1,1,–1 execută
  s:=0 ;
  Pentru j:=i+1,n execută
    s:=s+$a_{ij} \cdot x_j$;
  sfârşit pentru j ;
  $x_i := \dfrac{(b_i - s)}{a_{ii}}$;
sfârşit pentru i .

§1.2. Matrice simetrice pozitiv definite

Reamintim că o matrice simetrică se numeşte pozitiv definită, dacă forma pătratică asociată este pozitiv definită. Mai precis, dacă A este o matrice simetrică, atunci A se numeşte *pozitiv definită* dacă

$$\varphi(x)=x^T Ax > 0 ,$$

pentru orice $x \neq 0$, unde $x = (x_1, x_2, ..., x_n)^T$.
Din Algebra Liniară, se ştie că o matrice simetrică A, este pozitiv definită dacă şi numai dacă Δ_r>0 pentru orice $r = \overline{1,n}$, unde

$$\Delta_r = \det \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{pmatrix} .$$

În practică aceste condiții sunt greu de verificat pentru matrice de dimensiuni mari. De aceea, în continuare vom prezenta unele condiții necesare, respectiv şi suficiente, pentru ca o matrice simetrică să fie pozitiv definită.

Propoziția 1. *Dacă A este o matrice simetrică pozitiv definită, atunci:*
  (a) $a_{ii} > 0$ *pentru orice* $i = \overline{1,n}$,
  (b) $a_{ii}a_{jj} > a_{ij}^2$ *pentru orice* $i, j = \overline{1,n}$.

Demonstrație.

$$\varphi(x) = x^T Ax = (x_1, ..., x_n) \cdot \begin{pmatrix} a_{11}x_1 + ... + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + ... + a_{nn}x_n \end{pmatrix} = (a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n)x_1 +$$

$$+ (a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n)x_2 + ... + (a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n)x_n$$

Ținând seama că $a_{ij} = a_{ji}$, în continuare avem

$$\varphi(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j = \begin{array}{r} a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + \\ + a_{22}x_2^2 + \ldots + 2a_{2n}x_2x_n + \\ \vdots \\ + a_{nn}x_n^2 \end{array}$$

În particular, pentru $x = e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ avem $\varphi(e_i) = a_{ii}$. Cum φ este pozitiv definită şi $e_i \neq 0$, rezultă că $a_{ii} = \varphi(e_i) > 0$, adică (a).

Pentru un număr real oarecare λ avem

$$\varphi(\lambda e_i + e_j) = a_{ii}\lambda^2 + 2a_{ij}\lambda + a_{jj} > 0 . \tag{1}$$

Pentru ca inegalitatea (1) să fie adevărată pentru orice $\lambda \in \mathrm{R}$, trebuie ca

$$\Delta = 4\left(a_{ij}^2 - a_{ii}a_{jj}\right) < 0 .$$

Aşadar am demonstrat că $a_{ij}^2 < a_{ii}a_{jj}$ pentru orice $i, j = \overline{1,n}$, adică (b). ❑

Observaţia 2. *Condiţiile care apar în Propoziţia* 1 *sunt doar necesare nu şi suficiente.*

Exemplu.

Matricea $A = \begin{pmatrix} 3 & 2 & -2 \\ 2 & 3 & 2 \\ -2 & 2 & 3 \end{pmatrix}$ satisface condiţiile din Propoziţia 1, dar nu este pozitiv definită.

Într–adevăr,

$$\varphi(x) = 3\left(x_1^2 + x_2^2 + x_3^2\right) + 4\left(x_1x_2 - x_1x_3 + x_2x_3\right).$$

Dacă $x = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, atunci $\varphi(x) = 9 - 12 = -3 < 0$, deci φ nu este pozitiv definită.

Definiţia 1. *Spunem că matricea* A *este tare diagonal dominantă dacă elementele sale satisfac inegalităţile:*

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| , \quad i = \overline{1,n} . \tag{d}$$

Dacă inegalitățile (d) *devin egalități pentru anumiți indici, dar nu pentru toți, matricea se numeşte slab diagonal dominantă.*

Teorema 1. *Fie A o matrice simetrică cu următoarele proprietăți:*

(i) *A este tare diagonal dominantă,*

(ii) *$a_{ii} > 0$ pentru $i=\overline{1,n}$.*

Atunci A este pozitiv definită.

Demonstrație.

Din condiția (i) rezultă că dacă $x \neq 0$, atunci:

$$\varphi(x)=\sum_{i=1}^{n} a_{ii}x_i^2+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_i x_j > \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right| x_i^2-\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|\cdot\left|x_i\right|\cdot\left|x_j\right|=$$

$$=\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|\cdot\left|x_i\right|\cdot\left(\left|x_i\right|-\left|x_j\right|\right)$$

Deoarece $a_{ij}=a_{ji}$ avem şi inegalitatea:

$$\varphi(x)>\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|\cdot\left|x_j\right|\left(\left|x_j\right|-\left|x_i\right|\right).$$

Adunând cele două inegalități rezultă

$$2\varphi(x)>\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|\cdot\left(\left|x_i\right|-\left|x_j\right|\right)^2 \geq 0.$$

Aşadar, $\varphi(x)>0$ pentru orice $x\neq 0$, deci φ este pozitiv definită. ❑

Definiția 2. *Fie $M=\{1,2,...,n\}$. O matrice A se numeşte reductibilă dacă există două submulțimi $S, T \subset M$ cu proprietățile:*

(i) *$S\neq\Phi$, $T\neq\Phi$*

(ii) *$S\cap T=\Phi$*

(iii) *$S\cup T=M$*

(iv) *$a_{ij}=0$ pentru orice $i\in S$ şi $j\in T$.*

Matricea A se numeşte ireductibilă dacă oricare ar fi submulțimile S şi T ale lui M cu proprietățile (i)–(iii)*, există $i_0\in S$ şi $j_0\in T$ astfel încât $a_{i_0 j_0}\neq 0$.*

Cel mai simplu exemplu de matrice reductibilă este matricea diagonală.

Teorema 2. *Fie A o matrice simetrică având următoarele proprietăți:*

(i) *A este slab diagonal dominantă,*

(ii) *A este ireductibilă,*

(iii) *$a_{ii}>0$ pentru orice $i=\overline{1,n}$.*

Atunci A este pozitiv definită.
Demonstrație. Procedând ca în demonstrația Teoremei 1, rezultă:

$$\varphi(x) \geq \frac{1}{2}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|\cdot\left(\left|x_i\right|-\left|x_j\right|\right)^2 \geq 0 .$$

Vom arăta că situația $\varphi(x)=0$ pentru $x\neq 0$ nu poate avea loc. Într–adevăr, $\varphi(x)$ se anulează în următoarele cazuri:

1) $a_{ij}=0$ pentru orice $i\neq j$. Atunci matricea A are forma diagonală şi este reductibilă.

2) $\left|x_i\right|=\left|x_j\right|=\alpha\neq 0$ pentru orice i şi j.

$$\varphi(x)=\sum_{i=1}^{n}a_{ii}\alpha^2+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}a_{ij}x_i x_j \geq \sum_{i=1}^{n}(a_{ii}-\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|)\cdot\alpha^2 \geq 0 .$$

Cum există cel puțin un indice i_0 astfel încât $\sum_{\substack{j=1\\ j\neq i_0}}^{n}\left|a_{i_0 j}\right|<a_{i_0 i_0}$, rezultă

$\varphi(x)>0$ pentru $x\neq 0$.

3) $a_{ij}=0$ pentru orice pereche de indici (i,j) pentru care $\left|x_i\right|\neq\left|x_j\right|$ şi $a_{ij}\neq 0$ dacă $\left|x_i\right|=\left|x_j\right|\neq 0$.

Fie $M=\{1,2, \dots , n\}$ şi $S=\left\{\, i,j\in M\,;\left|x_i\right|=\left|x_j\right|\neq 0 \,\right\}$.
Dacă $S=M$, atunci suntem în cazul 2).
Dacă $S=\Phi$, atunci $\left|x_i\right|\neq\left|x_j\right|$ pentru orice i şi j şi evident $\varphi(x)>0$ pentru $x\neq 0$.
Aşadar, putem presupune că $\Phi\neq S\subset M$ (incluziune strictă). Dacă notăm cu $T=M\setminus S$ atunci S şi T satisfac condițiile (i)–(iv) din Definiția 2, deci A este reductibilă.
❑

Exemplu.
Fie

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

Matricea A este simetrică, slab diagonal dominantă, ireductibilă şi are elementele de pe diagonala principală strict pozitive. Din Teorema 2 rezultă că A este pozitiv definită.

Oservația 2. *Teorema 2 este utilă la stabilirea faptului că anumite matrice care apar în rezolvarea numerică a ecuațiilor cu derivate parțiale de tip eliptic sunt pozitiv definite.*

§1.3. Metoda Cholesky

Fie A o matrice simetrică, pozitiv definită şi

$$\varphi(x) = x^T A x = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j$$

forma pătratică asociată. Deoarece $a_{11} > 0$ avem:

$$\varphi(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j = \left(\sqrt{a_{11}}x_1 + \sum_{j=2}^{n}\frac{a_{1j}}{\sqrt{a_{11}}}x_j\right)^2 +$$

$$+ \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}^{(1)}x_ix_j, \text{ unde } a_{ij}^{(1)} = a_{ij} - \frac{a_{1i}a_{1j}}{a_{11}}, \quad i,j = \overline{2,n}$$

Dacă notăm cu

$$\varphi_1(x) = \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}^{(1)} x_i x_j \ ,$$

atunci φ_1 este la rândul său o formă pătratică pozitiv definită.

Într–adevăr, să presupunem prin absurd că există $z = \begin{pmatrix} z_2 \\ \vdots \\ z_n \end{pmatrix} \neq 0$ astfel încât

$\varphi_1(z) \leq 0$.

Fie $z_1 = -\sum\limits_{j=2}^{n}\frac{a_{1j}}{a_{11}}z_j$ şi $\overline{z} = \begin{pmatrix} z_1 \\ z_2 \\ . \\ z_n \end{pmatrix}$.

În continuare avem

$$\varphi(\overline{z}) = \left(\sqrt{a_{11}}z_1 + \sum_{j=2}^{n}\frac{a_{1j}}{\sqrt{a_{11}}}z_j\right)^2 + \varphi_1(z) = 0 + \varphi_1(z) \leq 0\ ,$$

ceea ce contrazice faptul că φ este pozitiv definită.

Aşadar, am demonstrat că φ_1 este pozitiv definită. În particular, rezultă că $a_{22}^{(1)} > 0$. Mai departe procedăm cu φ_1 aşa cum am procedat cu φ şi obținem

$$\varphi_1(x)=\left(\sqrt{a_{22}^{(1)}}\,x_2+\sum_{j=3}^{n}\frac{a_{2j}^{(1)}}{\sqrt{a_{22}^{(1)}}}x_j\right)^2+\varphi_2(x)\quad ,$$

unde

$$\varphi_2(x)=\sum_{i=3}^{n}\sum_{j=3}^{n}a_{ij}^{(2)}x_ix_j$$

este pozitiv definită. În final $\varphi(x)$ se reprezintă ca o sumă de pătrate. Mai precis $\varphi(x)$ admite următoarea scriere:

$$\varphi(x)=\sum_{i=1}^{n}\left(\sqrt{a_{ii}^{(i-1)}}\,x_i+\sum_{j=i+1}^{n}\frac{a_{ij}^{(i-1)}}{\sqrt{a_{ii}^{(i-1)}}}x_j\right)^2,$$

unde

$$a_{ij}^{(0)}=a_{ij}\quad \text{şi}\quad a_{ij}^{(p)}=a_{ij}^{(p-1)}-\frac{a_{pi}^{(p-1)}a_{pj}^{(p-1)}}{a_{pp}^{(p-1)}},\quad p=\overline{1,n-1}\,.$$

Introducem notațiile:

$$\begin{aligned}
&r_{ii}=\sqrt{a_{ii}^{(i-1)}},\quad i=\overline{1,n}\\
&r_{ij}=\frac{a_{ij}^{(i-1)}}{r_{ii}},\quad i<j\\
&r_{ij}=0,\qquad j<i\\
&a_{ij}^{(p)}=a_{ij}^{(p-1)}-r_{pi}r_{pj},\ \ p=\overline{1,n-1},\ \ i,j=\overline{p+1,n}.
\end{aligned}\tag{1}$$

Cu aceste notații avem

$$\varphi(x)=\sum_{i=1}^{n}\left(\sum_{j=i}^{n}r_{ij}x_j\right)^2=(r_{11}x_1+r_{12}x_2+\dots+r_{1n}x_n)^2+(r_{22}x_2+\dots+r_{2n}x_n)^2+$$

$$+\dots+(r_{nn}x_n)^2$$

Dacă notăm cu R următoarea matrice superior triunghiulară

$$R=\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n}\\ 0 & r_{22} & \cdots & r_{2n}\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \dots & r_{nn}\end{pmatrix},$$

atunci $\varphi(x)=(x^TR^T)(Rx)=x^T(R^TR)x$. Pe de altă parte, $\varphi(x)=x^TAx$. Se obține astfel următoarea descompunere a matricei A

$$A=R^TR \tag{2}$$

unde R este o matrice superior triunghiulară.
Descompunerea (2) poartă numele de *factorizarea Cholesky* a matricei A şi are loc pentru matrice simetrice pozitiv definite.

Numărul de operații pentru determinarea matricei R
Pentru a calcula elementele liniei a i–a a matricei R sunt necesare $(n-i)(2i-1)+2i-2$ operații elementare şi o extragere de rădăcină pătrată. Pentru toate liniile sunt necesare

$$\sum_{i=1}^{n}\left[\ (n-i)(2i-1)+2i-2\ \right]=\frac{n^3}{3}+\frac{n^2}{2}-\frac{5n}{6}$$

operații elementare plus n extrageri de rădăcină pătrată.

Exemplu. Să se determine descompunerea Cholesky a matricei

$$A=\begin{pmatrix}3 & 2 & 2\\ 2 & 3 & 2\\ 2 & 2 & 3\end{pmatrix}$$

$$r_{11}=\sqrt{3}\,,\ \ r_{12}=\frac{2}{\sqrt{3}}\,,\ \ r_{13}=\frac{2}{\sqrt{3}}\,,\ \ a_{22}^{(1)}=a_{22}-r_{12}^2=\frac{5}{3}\,,\ r_{22}=\sqrt{\frac{5}{3}}\,,$$

$$a_{23}^{(1)}=a_{23}-r_{12}r_{13}=\frac{2}{3}\,,\ r_{23}=\frac{a_{23}^{(1)}}{r_{22}}=\frac{2}{\sqrt{15}}\,,\ \ a_{33}^{(1)}=a_{33}-r_{13}^2=\frac{5}{3}\,,$$

$$a_{33}^{(2)}=a_{33}^{(1)}-r_{23}^2=\frac{7}{5}\,,\ \ r_{33}=\sqrt{\frac{7}{5}}$$

$$R=\begin{pmatrix}\sqrt{3} & \dfrac{2}{\sqrt{3}} & \dfrac{2}{\sqrt{3}}\\ 0 & \sqrt{\dfrac{5}{3}} & \dfrac{2}{\sqrt{15}}\\ 0 & 0 & \sqrt{\dfrac{7}{5}}\end{pmatrix}.$$

Se verifică imediat că $A=R^TR$.

Rezolvarea sistemului $Ax=b$ cu metoda Cholesky, în cazul când matricea A este simetrică şi pozitiv definită, revine la rezolvarea a două sisteme triunghiulare şi anume

$$\begin{cases}R^T y=b\\ Rx=y\end{cases}$$

Algoritmul Cholesky pentru rezolvarea sistemelor de ecuații liniare
Pentru $p:=1,n-1$ execută

$r_{pp} := \sqrt{a_{pp}}$;
Pentru k:=p+1,n–1 execută
$r_{pk} := \dfrac{a_{pk}}{r_{pp}}$;
sfârşit pentru k ;
Pentru i:=p+1,n execută
Pentru j:=i,n execută
a_{ij}:=a_{ij}–$r_{pi}r_{pj}$;
sfârşit pentru j ;
sfârşit pentru i ;
sfârşit pentru p ;
{ Rezolvarea sistemului $R^T y=b$ }
$y_1 = \dfrac{b_1}{r_{11}}$;
Pentru i:=2,n execută
s:=0 ;
Pentru j:=1,i execută
s:=s+$r_{ij} \cdot y_j$;
sfârşit pentru j ;
$y_i := \dfrac{b_i - s}{r_{ii}}$;
sfârşit pentru i ;
{ Rezolvarea sistemului $Rx=y$ }
$x_n = \dfrac{y_n}{r_{nn}}$;
Pentru i:=n–1,1 execută
s:=0 ;
Pentru j:=i+1,n execută
s:=s+$r_{ij} \cdot x_j$;
sfârşit pentru j ;
$x_i := \dfrac{y_i - s}{r_{ii}}$;
sfârşit pentru i .

Algoritmul se află programat şi în MATLAB şi se apelează cu secvenţa:
R=chol(A);
x=R\R'\b { pentru afişarea soluţiei }

§1.4. Metoda Householder. Factorizarea QR

O *matrice Householder* este o matrice de forma $H = I_n - 2hh^T$, unde $h^T=(0,..., 0, h_i, ..., h_n)$ şi $\|h\|_2 = \sqrt{h_i^2 + ... + h_n^2} = 1$. Se observă imediat că o matrice Householder este simetrică şi are următoarea structură:

$$H = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ 0 & & & 1-2h_i^2 & -2h_ih_{i+1} & \dots & -2h_ih_n \\ & & & \vdots & \vdots & & \vdots \\ & & & -2h_nh_i & -2h_nh_{i+1} & \dots & 1-2h_n^2 \end{pmatrix}$$

Mai mult, constatăm că H este ortogonală. Într–adevăr,

$$H^2 = (I_n - 2hh^T)(I_n - 2hh^T) = I_n - 2hh^T - 2hh^T + 4h(h^Th)h^T.$$

Cum $h^Th = 1$, rezultă $H^2 = I_n$. Aşadar, avem $H^{-1} = H = H^T$.

Un calcul simplu ne arată că $(hh^T)x = (h^Tx)h$, pentru orice $x = (x_1, x_2, ..., x_n)^T$.

În continuare ne punem următoarea problemă: *dat fiind un vector coloană $x \neq 0$, se poate determina o matrice Householder H, astfel încât Hx să fie colinear cu e_1 ? (unde $e_1^T = (1,0,...,0)$).*

Cu alte cuvinte, se poate determina un vector coloană h, cu $\|h\|_2 = 1$ şi un număr real σ astfel încât $Hx = x - 2hh^Tx = \sigma \cdot e_1$?

Ținând seama de observația de mai sus, aceasta revine la $x - 2(h^Tx)h = \sigma \cdot e_1$, de unde rezultă $x - \sigma \cdot e_1 = 2(h^Tx)h$. Aşadar, h trebuie să fie colinear cu $x - \sigma \cdot e_1$. Cum $\|h\|_2 = 1$ rezultă

$$h = \frac{x - \sigma \cdot e_1}{\|x - \sigma \cdot e_1\|_2} \quad . \tag{1}$$

Pe de altă parte, H fiind ortogonală avem

$$\|x\|_2 = \|Hx\|_2 = |\sigma| \cdot \|e_1\| = |\sigma| .$$

Alegem $\sigma = -\mathrm{sgn}(x_1)\|x\|_2$ şi facem convenția $\mathrm{sgn}(x_1) = 1$ dacă $x_1 = 0$.

În continuare avem

$$x-\sigma\cdot e_1=\begin{pmatrix} x_1+\operatorname{sgn}(x_1)\cdot\|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} \left(|x_1|+\|x\|_2\right)\cdot\operatorname{sgn}(x_1) \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{şi}$$

$$\|x-\sigma e_1\|_2^2=2\|x\|_2^2+2|x_1|\cdot\|x\|_2=2\|x\|_2\left(\|x\|_2+|x_1|\right) .$$

Înlocuind în (1) obținem

$$h=\frac{1}{\sqrt{2\|x\|_2\left(\|x\|_2+|x_1|\right)}}\begin{pmatrix} \left(|x_1|+\|x\|_2\right)\operatorname{sgn}x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} . \tag{2}$$

Se obține astfel următorul algoritm pentru determinarea lui h şi deci a matricei H:

$$\begin{aligned} &H=I_n-\beta\, u\, u^T \\ &\beta=\left(\|x\|_2\left(\|x\|_2+|x_1|\right)\right)^{-1} \\ &u=\left(\left(|x_1|+\|x\|_2\right)\operatorname{sgn}(x_1), x_2, ..., x_n\right)^T \\ &\operatorname{sgn}(x_1)=1 \;\text{ dacă }\; x_1=0. \end{aligned} \tag{3}$$

Teorema 1. *Pentru orice matrice $A\in M_n(R)$ nesingulară există o matrice ortogonală H astfel încât matricea $R=HA$ este superior triunghiulară.*

Demonstrație.

Fie $a_1=\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$, prima coloană a matricei A. Din cele arătate mai înainte rezultă că există o matrice Householder H_1 astfel încât $H_1a_1=\sigma\cdot e_1$. Matricea H_1 se determină astfel:

$$s=\left(\sum_{j=1}^{n}a_{j1}^2\right)^{1/2} , \quad \beta=\left(s\left(s+|a_{11}|\right)\right)^{-1} , \quad u=\left(\left(|a_{11}|+s\right)\operatorname{sgn}(a_{11}), a_{21}, ..., a_{n1}\right)^T ,$$

$$\operatorname{sgn}(a_{11})=1 \;\text{ daca }\; a_{11}=0 , \quad H_1=I_n-\beta uu^T . \tag{4}$$

Dacă notăm cu $A_1=H_1A$, atunci A_1 are următoarea formă:

$$A_1=\begin{pmatrix} -\operatorname{sgn}(a_{11})s & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

În continuare considerăm vectorul $a_2^{(1)} = \begin{pmatrix} a_{22}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{pmatrix}$ și determinăm o matrice ortogonală $\widetilde{H}_2 \in M_{n-1}(\mathrm{R})$ astfel încât

$$\widetilde{H}_2 a_2^{(1)} = \sigma \cdot \widetilde{e}_1,$$

unde $\widetilde{e}_1^{\,T} = (1, 0, ..., 0) \in \mathrm{R}^{n-1}$.

Notăm cu $H_2 = \begin{pmatrix} 1 & 0 \\ 0 & \widetilde{H}_2 \end{pmatrix} \in M_n(\mathrm{R})$ și cu $A_2 = H_2 A_1$. Matricea A_2 va arăta astfel

$$A_2 = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \dots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix}, \text{ unde } a_{1j}^{(2)} = a_{1j}^{(1)},\ j = \overline{1,n}.$$

În continuare se determină o matrice Householder $\widetilde{H}_3 \in M_{n-2}(\mathrm{R})$ cu proprietatea că $\widetilde{H}_2 a_3^{(2)} = \sigma \widetilde{\widetilde{e}}_1$, unde $\widetilde{\widetilde{e}}_1^{\,T} = (1, 0, ..., 0) \in M_{n-2}(\mathrm{R})$. Vom nota cu $H_3 = \begin{pmatrix} I_2 & 0 \\ 0 & \widetilde{H}_3 \end{pmatrix} \in M_n(\mathrm{R})$ și cu $A_3 = H_3 A_2$. Matricea A_3 va avea toate elementele de sub diagonala principală, din primele trei coloane, zero. Procedeul continuă într-un mod evident. În final, obținem o matrice superior triunghiulară $A_{n-1} = H_{n-1} A_{n-2} = H_{n-1} \cdot ... \cdot H_2 H_1 A$. Dacă notăm $H = H_{n-1} \cdot ... \cdot H_2 H_1$ și cu $R = HA$, atunci H este ortogonală și R superior triunghiulară. ❑

Corolar. *Pentru orice matrice nesingulară $A \in M_n(\mathrm{R})$ există o matrice ortogonală Q și o matrice superior triunghiulară R astfel încât $A = QR$.*

Algoritmul Householder pentru rezolvarea sistemelor de ecuații liniare

Fie sistemul $Ax = b$ cu $A \in M_n(\mathrm{R})$. Notăm cu $C = (A|b) = (c_{ij}) \in M_{n,n+1}(\mathrm{R})$ matricea extinsă.

Pentru $i := 1, n-1$ execută

$$s := \sqrt{\sum_{j=i}^{n} c_{ji}^2}\ ;$$

dacă $s = 0$ atunci A este singulară . Stop!

altfel $\beta := (s(|c_{ii}| + s))^{-1}$; dacă $c_{ii} = 0$ atunci $\mathrm{sgn}(c_{ii}) := 1$;

$u := (0, ..., 0, (c_{ii} + s) \cdot \mathrm{sgn}(c_{ii}), c_{i+1,i}, ..., c_{ni})^T$;

$$H_i := I_n - \beta uu^T;\ \ C := H_i C;$$

sfârşit pentru i ;

Exemplu. Fie sistemul

$$\begin{cases} 5x_1 + 2x_2 + x_3 = 12 \\ 5x_1 - 6x_2 + 2x_3 = -1 \\ -4x_1 + 2x_2 + x_3 = 3 \end{cases}$$

Soluţia exactă este $x_1 = 1,\ x_2 = 2,\ x_1 = 3.$

Aplicăm metoda Householder.

$$A = \begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} \ ; \ b = \begin{pmatrix} 12 \\ -1 \\ 3 \end{pmatrix} \ , \ C = \begin{pmatrix} 5 & 2 & 1 & 12 \\ 5 & -6 & 2 & -1 \\ -4 & 2 & 1 & 3 \end{pmatrix} \ ;$$

Iteraţia I

$$a_1 = \begin{pmatrix} 5 \\ 5 \\ -4 \end{pmatrix} ;\ c_{11} = 5 \ ;\ s = \sqrt{66} = 8.124038405;$$

$$\beta = \frac{1}{8.124038405 \cdot 13.124038405} = 9.379086466 \cdot 10^{-3} \quad ; u = \begin{pmatrix} 13.124038405 \\ 5 \\ -4 \end{pmatrix} ;$$

$$H_1 = \begin{pmatrix} -0.615457455 & -0.615457455 & 0.492365964 \\ -0.615457455 & 0.765522838 & 0.187581729 \\ 0.492365964 & 0.187581729 & 0.849934617 \end{pmatrix}$$

$$A_1 = H_1 A = \begin{pmatrix} -8.12403840 & 3.44656174 & -1.35400640 \\ 0 & -5.44888848 & 1.10316995 \\ 0 & 1.55911078 & 1.71746403 \end{pmatrix} ;$$

$$b_1 = H_1 b = \begin{pmatrix} -5.292934112 \\ -7.588267109 \\ 8.270613687 \end{pmatrix} ;\ C = H_1 \cdot C = \left[A_1 | b_1\right]$$

Iteraţia a II–a

$$a_2^{(1)} = \begin{pmatrix} -5.448888481 \\ 1.559110785 \end{pmatrix} ;\ s = \sqrt{c_{22}^2 + c_{32}^2} = 5.667557862;$$

$$c_{22} = -5.448888481 \ ; \ ; \ \beta = 0.015872234 \ ; \ u = \begin{pmatrix} 0 \\ -11.116446343 \\ 1.559110785 \end{pmatrix} ;$$

$$H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.961417354 & 0.275093933 \\ 0 & 0.275093933 & 0.961417354 \end{pmatrix};$$

$$A_2 = H_2 A_1 = \begin{pmatrix} -8.124038405 & 3.446561747 & -1.354006401 \\ 0 & 5.667557862 & -0.588142797 \\ 0 & 0 & 1.954675092 \end{pmatrix};$$

$$b_2 = H_2 b_1 = \begin{pmatrix} -5.292934112 \\ 9.570687333 \\ 5.864025277 \end{pmatrix}; \quad C = H_2 \cdot C = \left[A_2 | b_2\right];$$

$$H = H_2 \cdot H_1 = \begin{pmatrix} -0.615457455 & -0.615457455 & 0.492365964 \\ 0.727158367 & -0.684384346 & 0.053467527 \\ 0.30406057 & 0.390935018 & 0.868744486 \end{pmatrix};$$

$$R = H \cdot A = \begin{pmatrix} -8.124038405 & 3.446561747 & -1.354006401 \\ 0 & 5.667557862 & -0.588142797 \\ 0 & 0 & 1.954675092 \end{pmatrix};$$

Soluția sistemului inițial este $x = R^{-1} H b$, unde:

$$R^{-1} = \begin{pmatrix} -0.123091491 & 0.074854538 & -0.062742657 \\ 0 & 0.176442839 & 0.053089941 \\ 0 & 0 & 0.511593975 \end{pmatrix}.$$

Se obține soluția $x_1 = 1$; $x_2 = 2$; $x_3 = 3.000000001$.

§1.5. Norme de matrice

Cele mai utilizate norme vectoriale pe R^n sunt:

1) $\|x\|_\infty = \max\{ |x_1|, |x_2|, \ldots, |x_n| \}$

2) $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \quad 1 \le p < \infty$

unde $x^T = (x_1, x_2, \ldots, x_n) \in \mathrm{R}^n$.

Definiția 1. *Se numeşte normă de matrice orice aplicație*

$$A \to \|A\|_M : \mathrm{M}_n(\mathrm{R}) \to \mathrm{R}_+$$

cu proprietățile:

(i) $\|A\|_M = 0$ *dacă şi numai dacă* $A = 0$,

(ii) $\|\lambda A\|_M = |\lambda| \|A\|_M$, ; $\lambda \in$R, $A \in$M$_n$(R) ,

(iii) $\|A + B\|_M \le \|A\|_M + \|B\|_M$,

(iv) $\|AB\|_M \le \|A\|_M \cdot \|B\|_M$, $A, B \in$M$_n$(R).

Un exemplu de normă de matrice este *norma euclidiană* de matrice, care se defineşte astfel

$$\|A\|_E = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2} . \quad (1)$$

Proprietățile (i) şi (ii) sunt evidente. Pentru a demonstra proprietățile (iii) şi (iv) se foloseşte inegalitatea Cauchy–Buniakovski–Schwarz pe R^n. Pentru exemplificare demonstrăm (iv). Fie $C = AB$. Atunci

$$c_{ij}^2 = \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)^2 \le \left(\sum_{k=1}^{n} a_{ik}^2 \right) \left(\sum_{k=1}^{n} b_{kj}^2 \right)$$

În continuare avem

$$\|AB\|_E^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^2 \le \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}^2 \right) \cdot \sum_{j=1}^{n} \left(\sum_{k=1}^{n} b_{kj}^2 \right) = \|A\|_E^2 \cdot \|B\|_E^2 ,$$

de unde rezultă $\|AB\|_E \le \|A\|_E \cdot \|B\|_E$.

Definiția 2. *O normă de matrice* $\|\cdot\|_M$ *se numeşte compatibilă cu norma vectorială* $\|\cdot\|_p$ *dacă* $\|Ax\|_p \le \|A\|_M \|x\|_p$ *pentru orice* x.

Observația 1. $\|Ax\|_2 \le \|A\|_E \|x\|_2$, $(\forall)\, x$.

Într–adevăr, $\|Ax\|_2^2 = \sum_{i=1}^{n} (a_{i1}x_1 + \ldots + a_{in}x_n)^2 \le \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}^2 \right) \left(\sum_{j=1}^{n} x_j^2 \right) = \|A\|_E^2 \cdot \|x\|_2^2$.

Observația 2. *Dacă* λ *este o valoare proprie a matricei A, atunci* $|\lambda| \le \|A\|_M$ *pentru orice normă de matrice compatibilă cu o normă vectorială.*

Într–adevăr, fie v un vector propriu al matricei A care corespunde valorii proprii λ . Atunci avem

$$|\lambda| \cdot \|v\| = \|\lambda v\| = \|Av\| \le \|A\|_M \|v\| ,$$

deci $|\lambda| \le \|A\|_M$.

După cum se ştie, între mulțimea $M_n(\mathrm{R})$ a matricelor pătratice cu elemente din R şi mulțimea $L(\mathrm{R}^n)$ a aplicațiilor liniare şi continue, $U : \mathrm{R}^n \to \mathrm{R}^n$, există o corespondență bijectivă. Mai precis, dacă A este matricea asociată transformării liniare U, atunci $U(e_i^T) = (a_{1i}, a_{2i}, ..., a_{ni})$,
unde $e_i^T = (0, ..., 1, ..., 0) \in \mathrm{R}^n$ şi $U(x^T) = (Ax)^T$. Pe de altă parte, spațiul $L(\mathrm{R}^n)$ este un spațiu normat în raport cu norma operatorială:

$$\|U\|_{\mathrm{O}} = \sup\left\{ \left\| U\left(x^T\right)\right\| ;\ \left\|x^T\right\| = 1 \right\} \tag{2}$$

unde cu $\|\cdot\|$ am notat o normă oarecare pe R^n.

Se ştie de asemenea că:

$$\|U\|_{\mathrm{O}} = \inf\left\{ c > 0 ;\ \left\|U\left(x^T\right)\right\| \le \mathrm{c}\left\|x^T\right\|,\ (\forall) x^T \in \mathrm{R}^n \right\} \tag{3}$$

Definiția 3. *Se numeşte norma matricei A subordonată normei vectoriale $\|\cdot\|$ următorul număr:*

$$\|A\| = \sup\left\{ \|Ax\| ;\ \|x\| = 1 \right\} = \sup_{x \ne 0} \frac{\|Ax\|}{\|x\|} \tag{4}$$

Ca şi în cazul normei operatoriale, avem

$$\|A\| = \inf\left\{ c > 0 ;\ \|Ax\| \le \mathrm{c}\|x\|,\ (\forall)\, x \right\} . \tag{5}$$

Din relația (5) rezultă în particular că $\|Ax\| \le \|A\| \cdot \|x\|$, $\forall x \in \mathrm{R}^n$, deci norma matriceală definită de (4) este compatibilă cu norma vectorială căreia îi este subordonată.

Este evident că aplicația $A \to \|A\|$ definită de (4) satisface proprietățile (i)–(iii) din definiția 1. De asemenea avem

$$\|ABx\| \le \|A\| \cdot \|Bx\| \le \|A\| \cdot \|B\| \cdot \|x\| ,$$

de unde rezultă $\|A \cdot B\| \le \|A\| \cdot \|B\|$.

Aşadar, formula (4) defineşte într–adevăr o normă de matrice.

Definiția 4. *Dacă $\lambda_1, ..., \lambda_n$ sunt valorile proprii ale matricei A , atunci se notează cu $\rho(A) = \max_{1 \le i \le n} |\lambda_i|$ şi $\rho(A)$ se numeşte raza spectrală a matricei A (în această definiție λ_i pot fi reale sau complexe)*

Teorema 1. *Pentru $A \in M_n(\mathrm{R})$ avem:*

(1) $$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}| ,$$

(2) $$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| \quad ,$$

(3) $$\|A\|_2 = \left(\rho\left(A^T \cdot A\right)\right)^{\frac{1}{2}} \quad ,$$

unde cu $\|A\|_p$ *am notat norma matricei* A *subordonată normei vectoriale* $\|x\|_p$.

Demonstrație.

$$\|Ax\|_\infty = \max_{1\le i\le n}\left|\sum_{j=1}^{n} a_{ij}x_j\right| \le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|\cdot|x_j| \le \|x\|_\infty \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$$

Rezultă $\|A\|_\infty \le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$. Rămâne să arătăm că există $\tilde{x}$ cu $\|\tilde{x}\|_\infty = 1$ astfel încât $\|A\tilde{x}\|_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$. Pentru aceasta, fie k astfel încât să avem

$$\sum_{j=1}^{n}|a_{kj}| = \max_{1\le i\le}\sum_{j=1}^{n}|a_{ij}| \tag{6}$$

şi fie $\tilde{x}_j = \begin{cases} 0 & \text{dacã} \quad a_{kj} = 0 \\ \dfrac{a_{kj}}{|a_{kj}|} & \text{dacã} \quad a_{kj} \ne 0 \end{cases}$.

Evident că $\|\tilde{x}\|_\infty = 1$ şi $\|A\tilde{x}\|_\infty = \sum_{j=1}^{n}|a_{kj}| = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$. Aşadar, am demonstrat (1). În continuare avem

$$\|Ax\|_1 = \sum_{i=1}^{n}|a_{i1}x_1 + \ldots + a_{in}x_n| \le \sum_{i=1}^{n}\left(|a_{i1}||x_1| + \ldots + |a_{in}||x_n|\right) \le$$

$$\le \left(\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|\right)\cdot\left(|x_1| + \ldots + |x_n|\right) = \max_{1\le j\le n}\left(\sum_{i=1}^{n}|a_{ij}|\right)\cdot\|x\|_1$$

de unde rezultă $\|A\|_1 \le \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$.

Pe de altă parte dacă $e^T_j=(0,...,1,...,0)$, atunci $\left\|e_j\right\|_1=1$ şi $\left\|Ae_j\right\|_1=\sum_{i=1}^{n}\left|a_{ij}\right|$, de unde rezultă $\left\|A\right\|_1\geq\sum_{i=1}^{n}\left|a_{ij}\right|$, pentru orice $j=\overline{1,n}$. Aşadar, $\left\|A\right\|_1\geq\max_{1\leq j\leq n}\sum_{i=1}^{n}\left|a_{ij}\right|$ şi cu aceasta afirmația (2) este dovedită.

Fie $B=A^TA$ şi fie

$$\mu_1=\sup\left\{\ x^TBx\,;\left\|x\right\|_2=1\ \right\} \tag{7}$$

Evident $\mu_1=\sup\left\{(Ax)^TAx\ ;\ \left\|x\right\|_2=1\right\}=\left\|A\right\|_2^2$, deci $\left\|A\right\|_2=\sqrt{\mu_1}$.

Deoarece mulțimea $S=\left\{x\ ;\ \left\|x\right\|_2=1\right\}$ este compactă, rezultă că există v cu proprietățile: $\mu_1=v^TBv$ şi $\left\|v\right\|_2=1$.

Vom arăta în continuare că $Bv=\mu_1v$, deci că v este un vector propriu pentru B şi corespunde valorii proprii μ_1.

Într-adevăr, pentru orice $z\neq0$ avem: $\left(\frac{z}{\left\|z\right\|_2}\right)^TB\left(\frac{z}{\left\|z\right\|_2}\right)\leq\mu_1$ şi deci

$$z^TBz\leq\mu_1\left\|z\right\|_2^2=\mu_1z^Tz \tag{8}$$

Pe de altă parte este evident că relația (8) este verificată şi pentru $z=0$. Deci relația (8) are loc pentru orice z. De asemenea avem:

$$v^TBv=\mu_1v^Tv \tag{9}$$

Dacă notăm cu $C=B-\mu_1I_n$, atunci avem:

$$z^TCz\leq0\ ,\quad (\forall)\ z\ \text{şi} \tag{8'}$$

$$v^TCv=0 \tag{9'}$$

Fie $z=v+ty$, unde $t\in\mathrm{R}$ este oarecare şi y este un vector oarecare. Din (8') şi din faptul că C este simetrică rezultă

$$v^TCv+2ty^T(Cv)+t^2y^TCy\leq0\ .$$

Ținând seama de (9') avem

$$t^2y^TCy+2ty^T(Cv)\leq0\quad . \tag{10}$$

Pentru ca (10) să fie adevărată pentru orice $t\in\mathrm{R}$ trebuie ca $y^TCv=0$. Cum y a fost arbitrar rezultă $0=Cv=(B-\mu_1I_n)v=Bv-\mu_1v$.

Aşadar, avem $Bv=\mu_1v$, deci μ_1 este valoare proprie pentru B şi în plus $\mu_1=\left\|A\right\|_2^2$.

Pe de altă parte, fie μ o altă valoare proprie a matricei B şi fie $u\neq0$, $\left\|u\right\|_2=1$, astfel încât $Bu=\mu u$. În continuare avem

$$\mu_1=\left\|A\right\|_2^2\geq\left\|Au\right\|_2^2=u^TBu=u^T\mu u=\mu\ .$$

Aşadar, μ_1 este cea mai mare valoare proprie a matricei B, deci am demonstrat şi afirmația (3). ❑

În particular dacă presupunem că matricea A este simetrică, rezultă că $B=A^2$. Fie $\lambda_1, \lambda_2, ..., \lambda_n$ valorile proprii ale matricei A, care în acest caz sunt reale. Se ştie că $\lambda_1^2, \lambda_2^2, ..., \lambda_n^2$ sunt valorile proprii ale matricei A^2. Să presupunem că $\lambda_1^2 = \max\limits_{1\le j\le n} \lambda_j^2$.

Din Teorema 1 rezultă că $\|A\|_2 = |\lambda_1|$. Dacă, în plus, A este pozitiv definită, atunci $\lambda_i>0$ pentru orice i. Să presupunem că: $\lambda_1 \ge \lambda_2 \ge ... \ge \lambda_n$. Din cele de mai sus rezultă $\|A\|_2 = \lambda_1$, unde λ_1 este cea mai mare valoare proprie a matricei simetrice şi pozitiv definite A.

§1.6. Perturbarea sistemelor liniare. Numărul de condiționare al unei matrice

Considerăm următorul sistem de ecuații liniare

$$\begin{cases} 10x_1 + 7x_2 + 8x_3 + 7x_4 = 32 \\ 7x_1 + 5x_2 + 6x_3 + 5x_4 = 23 \\ 8x_1 + 6x_2 + 10x_3 + 9x_4 = 33 \\ 7x_1 + 5x_2 + 9x_3 + 10x_4 = 31 \end{cases} \qquad (1)$$

a cărui soluție exactă este $x_1=x_2=x_3=x_4=1$.

Să considerăm acum sistemul (1') în care am modificat "puțin" termenii liberi

$$\begin{cases} 10x_1 + 7x_2 + 8x_3 + 7x_4 = 32.1 \\ 7x_1 + 5x_2 + 6x_3 + 5x_4 = 22.9 \\ 8x_1 + 6x_2 + 10x_3 + 9x_4 = 33.1 \\ 7x_1 + 5x_2 + 9x_3 + 10x_4 = 30.9 \end{cases} . \qquad (1')$$

Soluția sistemului (1') este $x_1= 9.2$; $x_2= -12.6$; $x_3= 4.5$; $x_4= -1.1$. Aşadar, o eroare mică, de ordinul 0.1, a termenilor liberi, produce o eroare mare, de ordinul 10, a soluției sistemului.

Fie acum sistemul (1") în care modificăm *puțin* coeficienții sistemului

$$\begin{cases} 10x_1 + 7x_2 + 8.1x_3 + 7.2x_4 = 32 \\ 7.08x_1 + 5.04x_2 + 6x_3 + 5x_4 = 23 \\ 8x_1 + 5.98x_2 + 9.89x_3 + 9x_4 = 33 \\ 6.99x_1 + 4.99x_2 + 9x_3 + 9.98x_4 = 31 \end{cases} . \qquad (1'')$$

Soluția sistemului (1") este: $x_1= -81$; $x_2= 137$; $x_3= -34$; $x_4= 22$.

Să analizăm acum efectul perturbării membrului drept asupra soluției unui sistem liniar $Ax=b$, în care matricea A este nesingulară.

Notăm cu δb perturbarea membrului drept şi cu δx perturbarea care rezultă pentru soluție. Avem: $A(x+\delta x)=b+\delta b$, de unde rezultă $A\delta x=\delta b$ şi deci $\delta x=A^{-1}\delta b$. Pentru orice normă de matrice compatibilă avem:

$$\|\delta x\| \le \left\|A^{-1}\right\| \cdot \|\delta b\| \tag{2}$$

Pe de altă parte $\|b\| = \|Ax\| \le \|A\| \cdot \|x\|$, de unde rezultă:

$$\frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|} \tag{3}$$

Din relațiile (2) şi (3) obținem $\dfrac{\|\delta x\|}{\|x\|} \le \|A\| \cdot \left\|A^{-1}\right\| \cdot \dfrac{\|\delta b\|}{\|b\|}$.

Numărul de condiționare al unei matrice se defineşte astfel

$$cond(A) = \|A\| \cdot \left\|A^{-1}\right\| \quad . \tag{4}$$

Aşadar, între eroarea relativă a membrului drept şi eroarea relativă a soluției sistemului avem următoarea inegalitate

$$\frac{\|\delta x\|}{\|x\|} \le cond(A) \cdot \frac{\|\delta b\|}{\|b\|} \quad . \tag{5}$$

Observăm că dacă numărul de condiționare al matricei coeficienților sistemului este mare, atunci la erori relativ mici ale termenilor liberi, pot apare erori relativ mari pentru soluția sistemului. În cazul exemplului (1) avem

$$A = \begin{pmatrix} 10 & 7 & 8 & 8 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}$$

şi $\mathrm{cond}_2(A) \cong 2984$. (S–a folosit norma de matrice $\| \ \|_2$). După cum se vede, numărul de condiționare este destul de mare, ceea ce explică instabilitatea soluției sistemului.

Numărul de condiționare are următoarele proprietăți:

(i) $\mathrm{cond}(I_n) \ge 1$

(ii) $\mathrm{cond}(A) = \mathrm{cond}(A^{-1})$

(iii) $\mathrm{cond}(\alpha A) = \mathrm{cond}(A)$ pentru orice $\alpha \ne 0$

(iv) $\mathrm{cond}_2(A) = \dfrac{\sqrt{\mu_1}}{\sqrt{\mu_n}}$, unde $\mu_1 \ge \mu_2 \ge \ldots \ge \mu_n > 0$ sunt valorile proprii ale matricei $B = A^T A$

(v) Dacă A este simetrică, atunci $\mathrm{cond}_2(A)=\dfrac{\max|\lambda_i|}{\min|\lambda_i|}$, unde $\lambda_1, ..., \lambda_n$ sunt valorile proprii ale matricei A

(vi) Dacă A este ortogonală, atunci $\mathrm{cond}(A)=1$.

Pentru a evalua eroarea soluției sistemului la o perturbare a coeficienților sistemului, avem nevoie de următoarele două leme.

Lema 1. *Dacă* $A\in M_n(R)$ *şi* $\|A\|<1$, *atunci:*

(*i*) $A+I_n$ *şi* $A-I_n$ *sunt nesingulare, şi*

(*ii*) $\dfrac{1}{\|A\|+1}\le\left\|(A\pm I_n)^{-1}\right\|\le\dfrac{1}{1-\|A\|}$.

Demonstrație.

Prezentăm demonstrația pentru $A+I_n$.

Presupunem prin absurd că $A+I_n$ este singulară. Atunci există $x\neq 0$, $\|x\|=1$ astfel încât $(A+I_n)\cdot x = 0$. În continuare avem $x=-Ax$, deci $\|x\|\le\|A\|\cdot\|x\|$.

Rezultă $\|A\|\ge 1$ ceea ce contrazice ipoteza $\|A\|<1$.

Pentru a demonstra (*ii*) observăm că

$$1=\|I_n\|=\left\|(I_n+A)(I_n+A)^{-1}\right\|\le\|I_n+A\|\cdot\left\|(I_n+A)^{-1}\right\|\le(1+\|A\|)\cdot\left\|(I_n+A)^{-1}\right\|$$

de unde rezultă

$$\frac{1}{\|A\|+1}\le\left\|(A+I_n)^{-1}\right\|.$$

Pe de altă parte avem

$$I_n=(I_n+A)^{-1}+A(I_n+A)^{-1} ,$$

de unde rezultă

$$\left\|(I_n+A)^{-1}\right\|=\left\|I_n-A(I_n+A)^{-1}\right\|\le 1+\|A\|\cdot\left\|(I_n+A)^{-1}\right\| ,$$

şi mai departe

$$\left\|(I_n+A)^{-1}\right\|\le\frac{1}{1-\|A\|}. \qquad \square$$

Lema 2 (a perturbării). *Fie* $A, B\in M_n(R)$ *cu proprietățile:*

(*i*) $\left\|A^{-1}\right\|\le\alpha$

(*ii*) $\left\|A^{-1}(B-A)\right\|\le k<1$.

Atunci B este nesingulară şi $\left\|B^{-1}\right\|\le\dfrac{\alpha}{1-k}$.

Demonstrație.

Din Lema 1 rezultă că $I_n+A^{-1}(B-A)=A^{-1}B$ este nesingulară.

Cum $\det(A^{-1}B) = \det(A^{-1})\cdot\det B$, va rezulta $\det B \neq 0$, deci B este nesingulară.
Tot din Lema 1 rezultă

$$\left\|B^{-1}A\right\| = \left\|\left[I_n + A^{-1}(B-A)\right]^{-1}\right\| \leq \frac{1}{1-\left\|A^{-1}(B-A)\right\|} \leq \frac{1}{1-k}.$$

Mai departe avem

$$\left\|B^{-1}\right\| = \left\|\left(B^{-1}A\right)A^{-1}\right\| \leq \left\|B^{-1}A\right\|\cdot\left\|A^{-1}\right\| \leq \frac{\alpha}{1-k}. \quad \square$$

Teorema 1. *Dacă perturbăm matricea coeficienților sistemului* $Ax = b$ *cu* δA *şi* $\left\|A^{-1}\delta A\right\| < 1$, *atunci între eroarea relativă a soluției şi eroarea relativă a matricei coeficienților are loc inegalitatea*

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\operatorname{cond}(A)}{1-\operatorname{cond}(A)\frac{\|\delta A\|}{\|A\|}}\frac{\|\delta A\|}{\|A\|}.$$

Demonstrație.
Din egalitățile $Ax=b$ şi $(A+\delta A)(x+\delta x)=b$ rezultă $A\delta x+\delta A x+\delta A\delta x=0$. În continuare avem $\delta x = -(A+\delta A)^{-1}\delta A\, x$ şi mai departe

$$\frac{\|\delta x\|}{\|x\|} \leq \left\|(A+\delta A)^{-1}\right\|\cdot\|\delta A\| \tag{6}$$

Dacă alegem în Lema 2 $\alpha = \left\|A^{-1}\right\|$ şi $B=A+\delta A$, atunci

$$\left\|A^{-1}(B-A)\right\| = \left\|A^{-1}\delta A\right\| < 1$$

şi va rezulta

$$\left\|B^{-1}\right\| = \left\|(A+\delta A)^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|}{1-\left\|A^{-1}\delta A\right\|} \leq \frac{\left\|A^{-1}\right\|}{1-\left\|A^{-1}\right\|\cdot\|\delta A\|}. \tag{7}$$

Din (6) şi (7) obținem

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\left\|A^{-1}\right\|\cdot\|\delta A\|}{1-\left\|A^{-1}\right\|\cdot\|\delta A\|} = \frac{\left\|A^{-1}\right\|\cdot\|A\|}{1-\left\|A^{-1}\right\|\cdot\|A\|\cdot\frac{\|\delta A\|}{\|A\|}}\cdot\frac{\|\delta A\|}{\|A\|}.$$

Ținând seama de definiția numărului de condiționare, ultima inegalitate devine

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\operatorname{cond}(A)}{1-\operatorname{cond}(A)\frac{\|\delta A\|}{\|A\|}}\frac{\|\delta A\|}{\|A\|}. \quad \square$$

Observația 1. *Dacă presupunem în plus că perturbăm şi membrul drept al sistemului cu* δb *atunci rezultă*

$$\frac{\|\delta x\|}{\|x\|} \le \frac{cond(A)}{1 - cond(A)\frac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right).$$

Observația 2. *Rezolvarea sistemului* $Ax=b$, *cu metoda Gauss revine la rezolvarea a două sisteme triunghiulare* $Uy=b$ *şi* $Lx=y$. *Rezolvarea fiecărui sistem necesită* n^2 *operații. Dacă unul din aceste sisteme este rău condiționat (ceea ce se poate întâmpla chiar dacă sistemul inițial* $Ax=b$ *este bine condiționat) metoda Gauss conduce la erori mari.*

Cu totul altfel stau lucrurile în cazul *metodei Householder*. Deoarece cond(Q) = 1 şi cond(QR) = cond(R), rezultă că sistemul $Qy = b$ este bine condiționat şi deci că sistemul $Ax = b$ are aceeaşi condiționare ca sistemul $Rx = y$. Aşadar, algoritmul Householder are proprietăți de stabilitate mai bune decât algoritmul Gauss.

Observația 3. *Pentru evaluarea numărului de condiționare* cond(A)*, este suficient să cunoaştem un majorant pentru* $\left\|A^{-1}\right\|$.

Calculul lui $\left\|(LU)^{-1}\right\|$ este mai uşor decât calculul lui $\left\|A^{-1}\right\|$, deoarece inversarea matricelor triunghiulare este uşoară. Să presupunem că

$$\left\|(LU)^{-1}\right\| = \alpha \quad \text{şi} \quad \|A - LU\| < \frac{k}{\alpha}$$

unde $0<k<1$. Atunci din Lema 2 rezultă $\left\|A^{-1}\right\| \le \dfrac{\alpha}{1-k}$.

Într–adevăr, să alegem în Lema 2 matricea LU în loc de A şi matricea A în loc de B. Avem

$$\left\|A^{-1}(A - LU)\right\| \le \left\|A^{-1}\right\| \cdot \|A - LU\| \le \alpha \cdot \frac{k}{\alpha} = k < 1.$$

Atunci rezultă $\left\|A^{-1}\right\| \le \dfrac{\alpha}{1-k}$ şi deci $\mathrm{cond}(A) \le \dfrac{\alpha\|A\|}{1-k}$. ❑

§1.7. Metode iterative de rezolvare a sistemelor de ecuaţii liniare

Metodele directe de rezolvare numerică a sistemelor de ecuaţii liniare se utilizează pentru sisteme care au matricea coeficienţilor densă (aproape toţi coeficienţii sunt nenuli) şi cu un număr de ecuaţii moderat (până la 100 de ecuaţii). Pentru sisteme mari de ecuaţii de ordinul $10^3 \rightarrow 10^5$ şi care au matricea coeficienţilor rară (cu multe elemente nule), se utilizează metode iterative de rezolvare numerică.

Să presupunem că sistemul

$$Ax = b \tag{1}$$

se poate pune sub forma echivalentă

$$x = Bx + c \tag{2}$$

Forma echivalentă (2) ne sugerează următorul proces iterativ:

$$x^{(m+1)} = Bx^{(m)} + c, \ \ m \geq 0\,, \tag{3}$$

unde $x^{(0)}$ este un vector arbitrar.

Dacă notăm cu x^* soluţia exactă a sistemului, atunci avem

$$x^* = Bx^* + c \tag{4}$$

Fie $e^{(m)} = x^* - x^{(m)}$ vectorul eroare.

Din (3) şi (4) rezultă $e^{(m+1)} = Be^{(m)}$, $m \in \mathrm{N}^*$ şi mai departe

$$e^{(m)} = B^m e^{(0)} \tag{5}$$

Teorema 1. *Dacă* $\|B\| < 1$, *atunci şirul* $(x^{(m)})$ *este convergent şi* $\lim\limits_{m\to\infty} x^{(m)} = x^*$.

Demonstraţie.

Este suficient să arătăm că $\lim\limits_{m\to\infty} \mathrm{e}^{(\mathrm{m})} = 0$.

Din (5) avem

$$\left\|e^{(m)}\right\| \leq \left\|B^m\right\| \left\|e^{(0)}\right\| \leq \|B\|^m \cdot \left\|e^{(0)}\right\| \ .$$

Deoarece $\lim\limits_{m\to\infty} \|B\|^m = 0$, rezultă $\lim\limits_{m\to\infty} e^{(m)} = 0$. ❑

Teorema 2. *Condiţia necesară şi suficientă ca şirul* $(x^{(m)})$ *definit de* (3) *să fie convergent este ca* $\rho(B) < 1$ *, unde cu* $\rho(B)$ *s–a notat raza spectrală a matricei B.*

Demonstrație.

Este suficient să arătăm că $\lim\limits_{m\to\infty} B^m = 0$ dacă şi numai dacă $\rho(B) < 1$. Din Algebra liniară se ştie că matricea B se poate aduce la *forma canonică Jordan*, deci că există o matrice nesingulară C astfel încât

$$C^{-1}\cdot B\cdot C = J = \begin{pmatrix} J_{p_1}(\lambda_1) & \cdots & 0 \\ \vdots & J_{p_2}(\lambda_2) & \vdots \\ 0 & & J_{p_r}(\lambda_r) \end{pmatrix},$$

unde

$$J_p(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ & \lambda & 1 & & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

este o *celulă Jordan*, $\lambda_1, \dots, \lambda_r$ sunt valorile proprii ale matricei B şi $p_1, \dots, p_r$ sunt ordinele de multiplicitate ale acestor valori proprii. Deoarece $C^{-1}B^m C = J^m$, rezultă că $\lim\limits_{m\to\infty} B^m = 0$ dacă şi numai dacă $\lim\limits_{m\to\infty} J^m = 0$. Pe de altă parte, $J = D + N$, unde

$$D = \begin{pmatrix} \lambda_1 & & 0 & & & & 0 \\ & \ddots & & & & & \\ 0 & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_r & & 0 \\ & & & & & \ddots & \\ 0 & & & & 0 & & \lambda_r \end{pmatrix}$$

este o matrice diagonală de ordinul n, iar N este o *matrice nilpotentă de ordinul* n, adică $N^n = 0$.

În continuare avem $J^m = \sum\limits_{k=0}^{m} C_m^k D^{m-k} N^k$. Deoarece $N^k=0$ pentru $k \geq n$, vom avea

$$J^m = \sum_{k=0}^{n} C_m^k D^{m-k} N^k \quad . \tag{6}$$

Observăm că $\|D\|_\infty = \max\limits_{1\leq i\leq r} |\lambda_1| = \rho(B) < 1$. Din (6) rezultă:

$$\left\|J^m\right\|_\infty \leq \sum_{k=0}^{n} \frac{m(m-1)\dots(m-k+1)}{k!} \|D\|_\infty^{m-k} \|N\|_\infty^k < \sum_{k=0}^{n} \frac{m^k}{k!} \|N\|_\infty^k \cdot (\rho(B))^{m-k}$$

Cum $\lim\limits_{m\to\infty} m^k (\rho(B))^{m-k} = 0$, rezultă că $\lim\limits_{m\to\infty} \left\| J^m \right\|_\infty = 0$, deci că $\lim\limits_{m\to\infty} J^m = 0$.

Reciproc, să presupunem că $\lim\limits_{m\to\infty} B^m = 0$ şi că $\rho(B) \geq 1$. Atunci există un vector propriu $x \neq 0$ şi o valoare proprie λ, cu $|\lambda| \geq 1$, astfel încât $Bx=\lambda x$ şi deci $B^m x = \lambda^m x$.

Cum $(\lambda^m x)$ nu converge la 0, rezultă că B^m nu converge la 0, ceea ce contrazice ipoteza făcută. ❑

Una din cele mai cunoscute metode iterative este *metoda Jacobi*.
Să presupunem că matricea sistemului $Ax=b$ are proprietatea $a_{ii} \neq 0, i = \overline{1,n}$. Dacă notăm cu $D = \operatorname{diag}(a_{11}, ..., a_{nn})$ şi cu $E=D-A$, atunci obținem sistemul echivalent $(D-E)\cdot x=b$ şi mai departe

$$x=D^{-1}Ex + D^{-1}b \qquad (7)$$

Cum $D^{-1} = \operatorname{diag}\left(\frac{1}{a_{11}}, ..., \frac{1}{a_{nn}}\right)$, rezultă că $\left\| D^{-1}E \right\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{\substack{j=1 \\ j\neq i}}^{n} \left| \frac{a_{ij}}{a_{ii}} \right|$.

Observăm că dacă matricea A este tare diagonal dominantă, atunci $\left\| D^{-1}E \right\|_\infty < 1$ şi în virtutea Teoremei 1, şirul

$$x^{(m+1)} = \left(D^{-1}E\right)x^{(m)} + D^{-1}b, \quad m \geq 0 \qquad (8)$$

este convergent pentru orice aproximație inițială $x^{(0)}$. Aşadar, metoda Jacobi constă în următoarele:
Sistemul $Ax = b$ se pune sub forma echivalentă (7).
Scris pe componente, sistemul (7) arată astfel

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} x_j \right), \quad i = \overline{1,n} \; . \qquad (7')$$

Se obține şirul recurent $\left\{ x^{(m)} \right\}$ unde

$$x_i^{(m+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} x_j^{(m)} \right), \quad i = \overline{1,n} \; . \qquad (8')$$

Dacă matricea A este tare diagonal dominantă, şirul $(x^{(m)})$ converge la soluţia exactă a sistemului.

Exemplu. Fie sistemul

$$\begin{pmatrix} 5 & -1 & -1 & -1 \\ -1 & 10 & -1 & -1 \\ -1 & -1 & 5 & -1 \\ -1 & -1 & -1 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -4 \\ 12 \\ 8 \\ 34 \end{pmatrix}$$

Soluţia exactă este $x_1 = 1,\ x_2 = 2,\ x_3 = 3,\ x_4 = 4.$

$$D = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix},\ \mathrm{E} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix},\ \mathrm{D}^{-1}E = \begin{pmatrix} 0 & 0{,}2 & 0{,}2 & 0{,}2 \\ 0{,}1 & 0 & 0{,}1 & 0{,}1 \\ 0{,}2 & 0{,}2 & 0 & 0{,}2 \\ 0{,}1 & 0{,}1 & 0{,}1 & 0 \end{pmatrix},$$

$$D^{-1}b = \begin{pmatrix} -0{,}8 \\ 1{,}2 \\ 1{,}6 \\ 3{,}4 \end{pmatrix},\ \left\| D^{-1}E \right\|_\infty = 0.6 < 1\ .$$

Obţinem următorul proces iterativ:

$$\begin{cases} x_1^{(m+1)} = \phantom{0.1x_1^{(m)}+}\ 0.2x_2^{(m)} + 0.2x_3^{(m)} + 0.2x_4^{(m)} - 0.8 \\ x_2^{(m+1)} = 0.1x_1^{(m)} \phantom{+0.2x_2^{(m)}} + 0.1x_3^{(m)} + 0.1x_4^{(m)} - 1.2 \\ x_3^{(m+1)} = 0.2x_1^{(m)} + 0.2x_2^{(m)} \phantom{+0.2x_3^{(m)}} + 0.2x_4^{(m)} + 1.6 \\ x_4^{(m+1)} = 0.1x_1^{(m)} + 0.1x_2^{(m)} + 0.1x_3^{(m)} \phantom{+0.1x_4^{(m)}} + 3.4 \end{cases}$$

Dacă alegem aproximaţia iniţială $x_1^0 = x_2^0 = x_3^0 = x_4^0 = 0$ atunci după 5 iteraţii obţinem

$$x_1^{(5)} = 0{,}948;\quad x_2^{(5)} = 1{,}969;\quad x_3^{(5)} = 2{,}948;\quad x_4^{(5)} = 3{,}969$$

O altă metodă iterativă cunoscută este *metoda Gauss–Seidel* şi care corespunde următoarei *spargeri* a matricei coeficienţilor:

$$A = (D + L) + U \text{ unde } D = \mathrm{diag}(a_{11}, \ldots, a_{nn}),$$

$$L = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ a_{21} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & 0 \end{pmatrix},\ \text{iar } U = \begin{pmatrix} 0 & a_{12} & \ldots & a_{1n} \\ 0 & 0 & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Sistemul (1) devine $(D+L)\,x = -Ux + b$ şi mai departe obţinem următorul proces iterativ:

$$(D+L)\,x^{(m+1)} = -Ux^{(m)} + b \ . \qquad (9)$$

Pe componente obţinem

$$x_i^{(m+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(m)}\right), \quad i=\overline{1,n} \ . \qquad (10)$$

Din algoritmul (10) se observă că fiecare nouă componentă, $x_j^{(m+1)}$, este imediat utilizată la calculul următoarei componente.

Se poate arăta că procesul iterativ Gauss–Seidel este convergent dacă matricea A este tare diagonal dominantă.

În cazul exemplului precedent obţinem

$$\begin{pmatrix} 5 & 0 & 0 & 0 \\ -1 & 10 & 0 & 0 \\ -1 & -1 & 5 & 0 \\ -1 & -1 & -1 & 10 \end{pmatrix}\begin{pmatrix} x_1^{(m+1)} \\ x_2^{(m+1)} \\ x_3^{(m+1)} \\ x_4^{(m+1)} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1^{(m)} \\ x_2^{(m)} \\ x_3^{(m)} \\ x_4^{(m)} \end{pmatrix} + \begin{pmatrix} -4 \\ 12 \\ 8 \\ 34 \end{pmatrix}$$

sau

$$\begin{cases} x_1^{(m+1)} = \frac{1}{5}\left(x_2^{(m)} + x_3^{(m)} + x_4^{(m)} - 4\right) \\ x_2^{(m+1)} = \frac{1}{10}\left(x_1^{(m+1)} + x_3^{(m)} + x_4^{(m)} + 12\right) \\ x_3^{(m+1)} = \frac{1}{5}\left(x_1^{(m+1)} + x_2^{(m+1)} + x_4^{(m)} + 8\right) \\ x_4^{(m+1)} = \frac{1}{10}\left(x_1^{(m+1)} + x_2^{(m+1)} + x_3^{(m+1)} + 34\right) \end{cases} .$$

Pentru $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = x_4^{(0)} = 0,$ după 5 iteraţii obţinem

$$x_1^{(5)} = 0.995 \ ; \quad x_2^{(5)} = 1.998 \quad ; \quad x_3^{(5)} = 2.998 \quad ; \quad x_4^{(5)} = 3.999 \ .$$

§1.8. Metode de relaxare. Principiile de bază

Metodele de relaxare sunt metode iterative şi sunt utilizate pentru rezolvarea numerică a sistemelor liniare care au matricea coeficienţilor simetrică şi pozitiv definită.

Fie sistemul liniar

$$Ax - b = 0 \qquad (1)$$

unde matricea A este simetrică şi pozitiv definită. Dacă $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ este un vector de probă oarecare, atunci notăm cu

$$r = Av - b \ . \tag{2}$$

Vectorul r se numeşte *vectorul rezidual*.

Scopul oricărei metode de relaxare este ca prin schimbarea sistematică a vectorului v, vectorul rezidual corespunzător r să se micşoreze, eventual să se anuleze.

În cele ce urmează, pentru orice doi vectori

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \text{ şi } v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

vom nota produsul lor scalar cu

$$\langle u,v \rangle = v^T u = u_1 v_1 + u_2 v_2 + ... + u_n v_n \tag{3}$$

Asociem sistemului (1) funcţia pătratică

$$F(v) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_i v_j - \sum_{i=1}^{n} b_i v_i = \frac{1}{2} \langle Av,v \rangle - \langle b,v \rangle \tag{4}$$

Deoarece A este pozitiv definită, rezultă $Q(v)>0$ pentru orice $v\neq 0$, unde $Q(v) = \langle Av,v \rangle$. Observăm de asemenea că pentru orice $i = \overline{1,n}$ avem

$$\frac{\partial F}{\partial v_i} = \sum_{j=1}^{n} a_{ij} v_j - b_i \ ,$$

deci vectorul rezidual

$$r = \text{grad} F \ . \tag{5}$$

Teorema 1. *Problema determinării soluţiei sistemului* (1) *este echivalentă cu problema determinării punctului de minim al funcţiei pătratice* (4).

Demonstraţie.

Fie v_0 soluţia sistemului (1). Atunci $r_0=Av_0-b=0$. Cum $r_0=\text{grad}F(v_0)$, rezultă $\frac{\partial F}{\partial v_i}(v_0) = 0$. Aşadar, $v=v_0$ este punct critic pentru F. Pe de altă parte,

$$d^2 F(v_0) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} dv_i dv_j > 0 \ .$$

Rezultă că $v=v_0$ este un punct de minim global pentru F.

Reciproc, dacă $v=v_0$ este punct de minim pentru F atunci

$$\frac{\partial F}{\partial v_i}(v_0)=0\,,\quad i=\overline{1,n}\,.$$

Rezultă $\sum\limits_{j=1}^{n} a_{ij} v_j^0 - b_i = 0,\ i=\overline{1,n}$, deci $v=v_0$ este soluție pentru (1). ❑

În continuare prezentăm principiul de bază al metodei relaxării. Fie v un vector de probă oarecare, p o direcție dată şi $D=\{\, v'=v+tp;\ t\in \mathbb{R}\,\}$, dreapta care trece prin v şi este paralelă cu p. Ne propunem să determinăm $v_0'\in D$ astfel încât $F(v_0')=\min\{F(v');\ v'\in D\}$. Ținând seama de (4), rezultă

$$F(v')=\frac{1}{2}\langle A(v+tp),v+tp\rangle-\langle b,v+tp\rangle=$$

$$=\frac{1}{2}\langle Av,v\rangle-\langle b,v\rangle+\frac{t}{2}\langle Av,p\rangle+\frac{t^2}{2}\langle Ap,p\rangle+\frac{t}{2}\langle Ap,v\rangle-t\langle b,p\rangle=$$

$$=F(v)+\frac{t^2}{2}\langle Ap,p\rangle+t\langle Av,p\rangle-t\langle b,p\rangle=F(v)+\frac{t^2}{2}\langle Ap,p\rangle+t\langle Av-b,p\rangle=$$

$$=F(v)+\frac{t^2}{2}\langle Ap,p\rangle+t\langle r,p\rangle\ .$$

Folosim notația

$$f(t)=F(v')=F(v+tp)=F(v)+\frac{t^2}{2}\langle Ap,p\rangle+t\langle r,p\rangle \qquad (6)$$

Determinăm pe t astfel încât $f(t)=F(v')$ să fie minimă. Pentru aceasta trebuie să avem $f'(t)=0$, de unde rezultă $t\langle Ap,p\rangle+\langle r,p\rangle=0$. Aşadar, obținem:

$$t_{\min}=-\frac{\langle r,p\rangle}{\langle Ap,p\rangle} \qquad (7)$$

Cum $f''(t)=\langle Ap,p\rangle>0$, rezultă că vectorul $v_0'=v+t_{\min}p$ este un punct de minim pentru $\mathrm{F(v')}$.

În continuare avem

$$f(t_{\min})=F(v_0')=F(v)-\frac{1}{2}\frac{\langle r,p\rangle^2}{\langle Ap,p\rangle}$$

de unde rezultă

$$\Delta F=F(v_0')-F(v)=-\frac{1}{2}\frac{\langle r,p\rangle^2}{\langle Ap,p\rangle}\le 0\ .$$

Pentru ca $\Delta F<0$, trebuie ca $\langle r,p\rangle\neq 0$. Rezultă că direcția p se alege astfel încât p să nu fie perpendiculară pe r.

Observaţia 1. *Dacă* $r_0' = Av_0' - b$ *este vectorul rezidual corespunzător vectorului* $v_0' = v + t_{\min} p$, *atunci* $\langle r_0', p \rangle = 0$.

Într–adevăr,

$$\langle r_0', p \rangle = \langle Av\text{-}b, p \rangle + t_{\min} \langle Ap, p \rangle = \langle r, p \rangle - \frac{\langle r, p \rangle}{\langle Ap, p \rangle} \langle Ap, p \rangle = 0 \ .$$

Pentru interpretarea geometrică a principiului relaxării să considerăm cazul particular $n = 2$.

Ecuaţiile $F(v)$ = constant, reprezintă ecuaţiile unor elipse concentrice, al căror centru comun, coincide cu punctul de minim al funcţiei F. Într–adevăr, ecuaţia $F(v) = c$ revine la

$$a_{11}v_1^2 + 2a_{12}v_1v_2 + a_{22}v_2^2 - 2b_1v_1 - 2b_2v_2 = 2c \ . \tag{8}$$

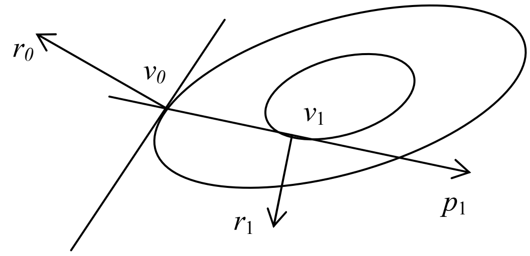
Deoarece A este pozitiv definită, rezultă că

$$\delta = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} > 0 \, ,$$

deci (8) reprezintă o elipsă.

Fie v_0 un vector de probă oarecare şi $c_0=F(v_0)$. Ecuaţia $F(v)=c_0$ reprezintă o elipsă şi $v=v_0$ aparţine acestei elipse. Deoarece $r_0=\mathrm{grad}F(v_0)$, rezultă că r_0 este perpendicular pe tangenta în $v=v_0$ la elipsă. Direcţia p_1 o alegem astfel încât să nu fie perpendiculară pe r_0. Fie $v_1=v_0+t_{\min}p_1$ şi fie $c_1=F(v_1)$.

Punctul $v=v_1$ aparţine elipsei $F(v)=c_1$ şi de asemenea aparţine dreptei ce trece prin v_0 şi are direcţia p_1. Fie $r_1=Av_1–b$. Din Observaţia 1, rezultă că r_1 este perpendicular pe direcţia p_1. Pe de altă parte $r_1=\mathrm{grad}F(v_1)$ este perpendicular pe tangenta în $v=v_1$ la elipsa $F(v)=c_1$. Rezultă că $v=v_1$, este punctul de tangenţă la elipsa $F(v)=c_1$ al dreptei care trece prin v_0 şi are direcţia p_1.



§1.9. Metoda relaxării simple

Este o metodă specifică calculelor de mână, având mai ales o semnificaţie istorică.

Fie v un vector de probă oarecare şi fie $r=Av - b$ vectorul rezidual corespunzător.

Dacă $\max\limits_{1\le i\le n}|r_i| = |r_j|$, atunci alegem $p=e_j$ unde $e_j^T = (0,...,1,...,0)$. Rezultă

$$t_{\min} = -\frac{\langle r, p\rangle}{\langle Ap, p\rangle} = -\frac{r_j}{a_{jj}} \quad \text{şi}$$

$$v' = v + t_{\min}\, p = v - \frac{r_j}{a_{jj}} e_j \tag{1}$$

Pe componente avem:

$$v_i' = \begin{cases} v_i & \text{daca} \quad i \ne j \\ v_j - \dfrac{r_j}{a_{jj}} & \text{daca} \quad i = j \end{cases} . \tag{2}$$

De asemenea vom avea $r' = Av' - b = r - \dfrac{r_j}{a_{jj}} Ae_j$ şi mai departe

$$\begin{cases} r_1' = r_1 - \dfrac{r_j}{a_{jj}} a_{1j} \\ \dots\dots\dots\dots\dots\dots \\ r_j' = 0 \\ \dots\dots\dots\dots\dots\dots \\ r_n' = r_n - \dfrac{r_j}{a_{jj}} a_{nj} \end{cases} \tag{3}$$

$$\Delta F = F(v) - F(v') = -\frac{1}{2}\frac{r_j^2}{a_{jj}} < 0\,,$$

ceea ce asigură convergența metodei. Deşi convergența este asigurată, experiențele numerice arată că aceasta este foarte lentă. Convergența este îmbunătățită dacă matricea A este tare diagonal dominantă.

Exemplu.

$$\begin{cases} -x_1 & +0.2x_2 & +0.2x_3 & +0.6 & =0 \\ 0.2x_1 & -x_2 & +0.2x_3 & +0.5 & =0 \\ 0.2x_1 & +0.2x_2 & -x_3 & +0,4 & =0 \end{cases}$$

Dacă alegem $v^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, atunci

$$r^{(1)} = \begin{pmatrix} 0.6 \\ 0.5 \\ 0.4 \end{pmatrix} \text{ şi } \max(r_1^{(1)},\ r_2^{(1)},\ r_3^{(1)}) = r_1^{(1)} = 0.6 \ .$$

Aşadar

$$p_1 = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v^{(2)} = v^{(1)} - \frac{r_1^{(1)}}{a_{11}} e_1 \quad \text{şi} \quad r^{(2)} = r^{(1)} - \frac{r_1^{(1)}}{a_{11}} Ae_1 \, .$$

Pe componente avem

$$\begin{cases} v_1^{(2)} & = 0.6 \\ v_2^{(2)} & = 0 \\ v_3^{(2)} & = 0 \end{cases}, \quad \begin{cases} r_1^{(2)} & = 0 \\ r_2^{(2)} & = 0.62 \\ r_3^{(2)} & = 0.52 \end{cases},$$

$$r_2^{(2)} = \max\left(r_1^{(2)},\ r_2^{(2)}\ ,\ r_3^{(2)}\right) = 0.62 \, .$$

Rezultă

$$p_2 = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v^{(3)} = v^{(2)} - \frac{r_2^{(2)}}{a_{22}} e_2 \quad \text{şi} \quad r^{(3)} = r^{(2)} - \frac{r_2^{(2)}}{a_{22}} Ae_2 \ ;$$

$$\begin{cases} v_1^{(3)} & = 0.6 \\ v_2^{(3)} & = 0.62 \\ v_3^{(3)} & = 0 \end{cases}, \quad \begin{cases} r_1^{(3)} & = 0.124 \\ r_2^{(3)} & = 0 \\ r_3^{(3)} & = 0.644 \end{cases}$$

$$r_3^{(3)} = \max\left(r_1^{(3)},\ r_2^{(3)}\ ,\ r_3^{(3)}\right) = 0.644 \, .$$

În continuare

$$p_3 = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad v^{(4)} = v^{(3)} - \frac{r_3^{(3)}}{a_{33}} e_3 \quad \text{şi} \quad r^{(4)} = r^{(3)} - \frac{r_3^{(3)}}{a_{33}} Ae_3 \ .$$

$$\begin{cases} v_1^{(4)} & = 0.6 \\ v_2^{(4)} & = 0.62 \\ v_3^{(4)} & = 0.644 \end{cases}, \quad \begin{cases} r_1^{(4)} & = 0.2528 \\ r_2^{(4)} & = 0.1288 \\ r_3^{(4)} & = 0 \end{cases}, \text{ etc.}$$

§1.10. Metoda deplasărilor succesive (Gauss - Seidel)

În metoda deplasărilor succesive, direcţia de relaxare urmează ciclic direcţiile $e_1, e_2, \dots, e_n$, indiferent de reziduurile respective, după care ciclul se reia. Pentru simplificare să presupunem că avem sistemul

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 - b_1 &= 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 - b_2 &= 0 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 - b_3 &= 0 \end{cases}$$

Fie $v^{(0)}$ vectorul de probă iniţial şi fie $p' = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Conform formulei (1) din §9 rezultă $v' = v^{(0)} - \dfrac{r_1^{(0)}}{a_{11}} e_1$, iar pe componente

$$\begin{cases} v_1' = v_1^{(0)} - \dfrac{1}{a_{11}}\left(a_{11}v_1^{(0)} + a_{12}v_2^{(0)} + a_{13}v_3^{(0)} - b_1\right) \\ v_2' = v_2^{(0)} \\ v_3' = v_3^{(0)} \end{cases}.$$

În continuare alegem direcţia de relaxare

$$p'' = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

şi obţinem vectorul v'' de componente

$$\begin{cases} v_1'' = v_1' \\ v_2'' = v_2^{(0)} - \dfrac{1}{a_{22}}\left(a_{21}v_1' + a_{22}v_2' + a_{23}v_3' - b_2\right) \\ v_3'' = v_3^{(0)} \end{cases}.$$

În sfârşit, pentru direcţia de relaxare

$$p''' = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

obținem vectorul v''' de componente

$$\begin{cases} v_1''' = v_1'' \\ v_2''' = v_2'' \\ v_3''' = v_3^{(0)} - \dfrac{1}{a_{33}}\left(a_{31}v_1'' + a_{32}v_2'' + a_{33}v_3'' - b_3\right) \end{cases}.$$

După încheierea acestui ciclu, vectorul găsit va fi notat cu $v^{(1)}$ și va avea componentele:

$$\begin{cases} v_1^{(1)} = v_1''' = -\dfrac{a_{12}}{a_{11}}v_2^{(0)} - \dfrac{a_{13}}{a_{11}}v_3^{(0)} + \dfrac{b_1}{a_{11}} \\ v_2^{(1)} = v_2''' = -\dfrac{a_{21}}{a_{22}}v_1^{(1)} - \dfrac{a_{23}}{a_{22}}v_3^{(0)} + \dfrac{b_2}{a_{22}} \\ v_3^{(1)} = v_3''' = -\dfrac{a_{31}}{a_{33}}v_1^{(1)} - \dfrac{a_{32}}{a_{33}}v_2^{(1)} + \dfrac{b_3}{a_{33}} \end{cases} \tag{1}$$

Efectuând calculele obținem:

$$\begin{cases} a_{11}v_1^{(1)} + a_{12}v_2^{(0)} + a_{13}v_3^{(0)} = b_1 \\ a_{21}v_1^{(1)} + a_{22}v_2^{(1)} + a_{23}v_3^{(0)} = b_2 \\ a_{31}v_1^{(1)} + a_{32}v_2^{(1)} + a_{33}v_3^{(1)} = b_3 \end{cases}.$$

În general, pentru un sistem de n ecuații, după $(m+1)$ cicluri se obține vectorul $v^{(m+1)}$ care verifică ecuațiile:

$$\begin{cases} a_{11}v_1^{(m+1)} + a_{12}v_2^{(m)} + a_{13}v_3^{(m)} + \ldots + a_{1n}v_n^{(m)} = b_1 \\ a_{21}v_1^{(m+1)} + a_{22}v_2^{(m+1)} + a_{23}v_3^{(m)} + \ldots + a_{2n}v_n^{(m)} = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}v_1^{(m+1)} + a_{n2}v_2^{(m+1)} + a_{n3}v_3^{(m+1)} + \ldots + a_{nn}v_n^{(m+1)} = b_n \end{cases} \tag{2}$$

Observația 1. *Formulele* (1) *coincid cu formulele* (10) *din* §7.
Dacă notăm cu

$$E = \begin{pmatrix} 0 & 0 & 0\ldots & 0 \\ a_{21} & 0 & 0\ldots & 0 \\ a_{31} & a_{32} & 0\ldots & 0 \\ \ldots & \ldots & & \\ a_{n1} & a_{n2} & \ldots & 0 \end{pmatrix},\quad F = E^T \quad \text{și} \quad D = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix},$$

atunci matricea A admite descompunerea $A=E+D+F$ şi şirul de vectori $v^{(m)}$ verifică relația matricială

$$(D+E)v^{(m+1)} + Fv^{(m)} = b\ , \tag{3}$$

de unde rezultă

$$v^{(m+1)}= -(D+E)^{-1}Fv^{(m)}+(D+E)^{-1}b\ . \tag{4}$$

În sfârşit, notând

$$M= -(D+E)^{-1}F \quad \text{şi} \quad C=(D+E)^{-1}b \tag{5}$$

obținem procesul iterativ

$$v^{(m+1)}=Mv^{(m)}+C\ . \tag{6}$$

Exemplu. Să se găsească soluția aproximativă obținută după 5 iterații cu metoda deplasărilor succesive, luând vectorul inițial (0, 0, 0), pentru sistemul $Ax = b$, unde:

$$A=\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}, \qquad b=\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Rezolvare

$$(D+E)^{-1}=\frac{1}{36}\begin{pmatrix} 18 & 0 & 0 & 0 \\ 6 & 12 & 0 & 0 \\ 2 & 4 & 12 & 0 \\ 1 & 2 & 6 & 18 \end{pmatrix} \qquad F=\begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$M=-(D+E)^{-1}F=\frac{1}{36}\begin{pmatrix} 0 & 18 & 0 & 0 \\ 0 & 6 & 12 & 0 \\ 0 & 2 & 4 & 12 \\ 0 & 1 & 2 & 6 \end{pmatrix}$$

$$\det(M-\lambda I)=\lambda^2\left(\lambda^2-\frac{4}{9}\lambda+\frac{1}{36}\right)\ .$$

Valorile proprii ale matricei M sunt $\lambda_1 = 0.369$; $\lambda_2 = 0.077$; $\lambda_3 = 0$; $\lambda_4 = 0$.
$\rho(M) = 0.369 < 1$, deci procesul este convergent.

Pe componente algoritmul conduce la:

Iterația I

$$\begin{cases} x_1^{(1)} = \dfrac{1}{a_{11}}(b - a_{12}x_2^{(0)} - a_{13}x_3^{(0)} - a_{14}x_4^{(0)}) \\ x_2^{(1)} = \dfrac{1}{a_{22}}(b - a_{21}x_1^{(1)} - a_{23}x_3^{(0)} - a_{24}x_4^{(0)}) \\ x_3^{(1)} = \dfrac{1}{a_{33}}(b - a_{31}x_1^{(1)} - a_{32}x_2^{(1)} - a_{34}x_4^{(0)}) \\ x_4^{(1)} = \dfrac{1}{a_{44}}(b - a_{41}x_1^{(1)} - a_{42}x_2^{(1)} - a_{43}x_3^{(1)}) \end{cases} \quad \text{rezultă} \quad x^{(1)} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.75 \end{pmatrix}$$

Iterația a II –a

$$\begin{cases} x_1^{(2)} = \dfrac{1}{a_{11}}(b - a_{12}x_2^{(1)} - a_{13}x_3^{(1)} - a_{14}x_4^{(1)}) \\ x_2^{(2)} = \dfrac{1}{a_{22}}(b - a_{21}x_1^{(2)} - a_{23}x_3^{(1)} - a_{24}x_4^{(1)}) \\ x_3^{(2)} = \dfrac{1}{a_{33}}(b - a_{31}x_1^{(2)} - a_{32}x_2^{(2)} - a_{34}x_4^{(1)}) \\ x_4^{(2)} = \dfrac{1}{a_{44}}(b - a_{41}x_1^{(2)} - a_{42}x_2^{(2)} - a_{43}x_3^{(2)}) \end{cases} \quad \text{rezultă} \quad x^{(2)} = \begin{pmatrix} 0.75 \\ 0.75 \\ 0.83333 \\ 0.91667 \end{pmatrix}$$

Iterația a III –a

$$\begin{cases} x_1^{(3)} = \dfrac{1}{a_{11}}(b - a_{12}x_2^{(2)} - a_{13}x_3^{(2)} - a_{14}x_4^{(2)}) \\ x_2^{(3)} = \dfrac{1}{a_{22}}(b - a_{21}x_1^{(3)} - a_{23}x_3^{(2)} - a_{24}x_4^{(2)}) \\ x_3^{(3)} = \dfrac{1}{a_{33}}(b - a_{31}x_1^{(3)} - a_{32}x_2^{(3)} - a_{34}x_4^{(2)}) \\ x_4^{(3)} = \dfrac{1}{a_{44}}(b - a_{41}x_1^{(3)} - a_{42}x_2^{(3)} - a_{43}x_3^{(3)}) \end{cases} \quad \text{rezultă} \quad x^{(3)} = \begin{pmatrix} 0.875 \\ 0.90278 \\ 0.93981 \\ 0.96991 \end{pmatrix}$$

Iterația a IV –a

$$\begin{cases} x_1^{(4)} = \dfrac{1}{a_{11}}(b - a_{12}x_2^{(3)} - a_{13}x_3^{(3)} - a_{14}x_4^{(3)}) \\ x_2^{(4)} = \dfrac{1}{a_{22}}(b - a_{21}x_1^{(4)} - a_{23}x_3^{(3)} - a_{24}x_4^{(3)}) \\ x_3^{(4)} = \dfrac{1}{a_{33}}(b - a_{31}x_1^{(4)} - a_{32}x_2^{(4)} - a_{34}x_4^{(3)}) \\ x_4^{(4)} = \dfrac{1}{a_{44}}(b - a_{41}x_1^{(4)} - a_{42}x_2^{(4)} - a_{43}x_3^{(4)}) \end{cases} \quad \text{rezultă} \quad x^{(4)} = \begin{pmatrix} 0.95139 \\ 0.96373 \\ 0.97788 \\ 0.98894 \end{pmatrix}$$

Iterația a V –a

$$\begin{cases} x_1^{(5)} = \dfrac{1}{a_{11}}(b - a_{12}x_2^{(4)} - a_{13}x_3^{(4)} - a_{14}x_4^{(4)}) \\ x_2^{(5)} = \dfrac{1}{a_{22}}(b - a_{21}x_1^{(5)} - a_{23}x_3^{(4)} - a_{24}x_4^{(4)}) \\ x_3^{(5)} = \dfrac{1}{a_{33}}(b - a_{31}x_1^{(5)} - a_{32}x_2^{(5)} - a_{34}x_4^{(4)}) \\ x_4^{(5)} = \dfrac{1}{a_{44}}(b - a_{41}x_1^{(5)} - a_{42}x_2^{(5)} - a_{43}x_3^{(5)}) \end{cases} \quad \text{rezultă} \quad x^{(5)} = \begin{pmatrix} 0.98187 \\ 0.98658 \\ 0.99184 \\ 0.99592 \end{pmatrix}.$$

Teorema 1. *Dacă matricea A este simetrică şi pozitiv definită, procesul iterativ Gauss –Seidel este convergent.*
Demonstraţia rezultă din analiza descreşterii funcţiei pătratice F prin trecerea de la o iteraţie la alta. O altă demonstraţie se bazează pe faptul că se poate arăta că dacă A este simetrică şi pozitiv definită, atunci $\rho(M) < 1$ şi conform Teoremei 2 din §7, procesul iterativ este convergent.

§1.11. Metoda suprarelaxării

Pentru sisteme mari de ecuaţii, procesul iterativ Gauss –Seidel converge lent, deoarece raza spectrală $\rho(M)$ este în vecinătatea lui 1.

Metoda suprarelaxării este o generalizare a metodei Gauss –Seidel, care constă în introducerea unui parametru ω în vederea accelerării convergenţei. Ca şi în metoda Gauss –Seidel, direcţia de relaxare urmează ciclic direcţiile $e_1, e_2, ..., e_n$ dar se înlocuieşte $t_{\min}$ cu $t=\omega \cdot t_{\min}$. Exemplificăm metoda pe cazul particular al unui sistem de trei ecuaţii. Fie $v^{(0)}$ vectorul de probă iniţial. După un ciclu în care direcţia de relaxare urmează succesiv direcţiile e_1, e_2, e_3 obţinem vectorul $v^{(1)}$ de componente:

$$\begin{cases} v_1^{(1)} = v_1^{(0)} - \dfrac{\omega}{a_{11}}\left(a_{11}v_1^{(0)} + a_{12}v_2^{(0)} + a_{13}v_3^{(0)} - b_1\right) \\ v_2^{(1)} = v_2^{(0)} - \dfrac{\omega}{a_{22}}\left(a_{21}v_1^{(1)} + a_{22}v_2^{(0)} + a_{23}v_3^{(0)} - b_2\right) \\ v_3^{(1)} = v_3^{(0)} - \dfrac{\omega}{a_{33}}\left(a_{31}v_1^{(1)} + a_{32}v_2^{(1)} + a_{33}v_3^{(0)} - b_3\right) \end{cases} \qquad (1)$$

Dacă $\omega = 1$, obţinem din nou formulele (1) din §10. După efectuarea calculelor rezultă:

$$\begin{cases} \omega^{-1}a_{11}v_1^{(1)} + (1-\omega^{-1})a_{11}v_1^{(0)} + a_{12}v_2^{(0)} + a_{13}v_3^{(0)} = b_1 \\ a_{21}v_1^{(1)} + a_{22}\omega^{-1}v_2^{(1)} + (1-\omega^{-1})a_{22}v_2^{(0)} + a_{23}v_3^{(0)} = b_2 \\ a_{31}v_1^{(1)} + a_{32}v_2^{(1)} + a_{33}\omega^{-1}v_3^{(1)} + (1-\omega^{-1})a_{33}v_3^{(0)} = b_3 \end{cases} \quad (2)$$

Dacă introducem notațiile

$$E = \begin{pmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix}, \quad D = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \quad \text{şi} \quad F = E^T,$$

relațiile (2) capătă forma matricială

$$(E+\omega^{-1}D)v^{(1)} + [F+(1-\omega^{-1})D]v^{(0)} = b \ . \quad (3)$$

În general, pentru un sistem de n ecuații, şirul de vectori $v^{(m)}$ satisface relația:

$$(E+\omega^{-1}D)\, v^{(m+1)} + [F+(1-\omega^{-1})D]\, v^{(m)} = b \ . \quad (4)$$

În continuare avem

$$v^{(m+1)} = -(E+\omega^{-1}D)^{-1}\, [F+(1-\omega^{-1})D]\, v^{(m)} + (E+\omega^{-1}D)^{-1} b \ .$$

Notăm cu

$$\begin{cases} M(\omega) = -(E+\omega^{-1}D)^{-1}\left[F + (1-\omega^{-1})D\right] \\ C(\omega) = \left[E + \omega^{-1}D\right]^{-1} b \end{cases} . \quad (5)$$

Obținem astfel procesul iterativ

$$v^{(m+1)} = M(\omega)v^{(m)} + C(\omega) \ . \quad (6)$$

Pentru $\omega = 1$ obținem algoritmul Gauss –Seidel (vezi (4), (5), (6) din §10).

Parametrul optim, ω_{opt}, va fi acela pentru care raza spectrală a matricei $M(\omega)$ va fi minimă. Evident, pentru acest parametru se obține cea mai rapidă convergență.

Se poate demonstra următoarea teoremă.

Teorema 1. *Dacă matricea A este simetrică şi pozitiv definită, metoda suprarelaxării este convergentă pentru orice $0 < \omega < 2$.*

În particular, rezultă că metoda Gauss –Seidel este convergentă dacă A este simetrică şi pozitiv definită, deoarece corespunde cazului particular $\omega = 1$.

Determinarea parametrului optim, ω_{opt}, este posibilă în cazul *matricelor bloc tridiagonale.*

Definiția 1. *O matrice A se numeşte bloc tridiagonală, dacă are următoarea structură:*

$$A = \begin{pmatrix} D_1 & F_1 & 0 & 0 & \cdots & & & 0 \\ E_1 & D_2 & F_2 & 0 & \cdots & & & 0 \\ 0 & E_2 & D_3 & F_3 & \cdots & & & 0 \\ \vdots & & & & \ddots & & & \vdots \\ 0 & & & & \cdots & E_{m-2} & D_{m-1} & F_{m-1} \\ 0 & & & & \cdots & 0 & E_{m-1} & D_m \end{pmatrix},$$

unde D_i *sunt matrice pătratice de diferite ordine,* E_k *şi* F_k *sunt, în general, matrice dreptunghiulare.* F_k *are acelaşi număr de linii ca matricea* D_k *şi acelaşi număr de coloane ca matricea* D_{k+1}. E_k *are acelaşi număr de linii cu* D_{k+1} *şi acelaşi număr de coloane cu* D_k. *În afara matricelor care intră în bandă, toate elementele sunt nule. Dacă, în plus, matricele* D_i *sunt diagonale,* A *se numeşte diagonal bloc tridiagonală.*

Prezentăm de asemenea fără demonstrație următoarea teoremă.

Teorema 2. *Fie* A *o matrice simetrică, pozitiv definită şi diagonal tridiagonală. Atunci parametrul optim de relaxare este dat de relația* $\omega_{\text{opt}} = \dfrac{2}{1+\sqrt{1-\lambda_1^2}}$, *unde* λ_1 *este cea mai mare valoare proprie a matricei* $-D^{-1}(E+F)$.

Exemplu. Fie sistemul $Ax = b$ din exemplul din paragraful precedent, care are matricea diagonal bloc tridiagonală

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}, \quad E+F = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$

$$\mathrm{D} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad -D^{-1}(E+F) = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Ecuația caracteristică este

$$\begin{vmatrix} -\lambda & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & -\lambda & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & -\lambda & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & -\lambda \end{vmatrix} = 0,$$

care este echivalentă cu

$$\begin{vmatrix} 2\lambda & -1 & 0 & 0 \\ -1 & 3\lambda & -1 & 0 \\ 0 & -1 & 3\lambda & -1 \\ 0 & 0 & -1 & 2\lambda \end{vmatrix} = 36\lambda^4 + 16\lambda^2 + 1 = 0.$$

Rezultă $\lambda_i = \pm\sqrt{\dfrac{4 \pm \sqrt{7}}{18}}$ şi $\lambda_1 = 0.60763$.

Parametrul optim de relaxare $\omega_{opt} = \dfrac{2}{1 + \sqrt{1 - \lambda_1^2}} = 1.11469$.

Procedeul iterativ $x^{(m+1)} = M(\omega) x^{(m)} + C(\omega)$, unde

$$M(\omega) = \begin{pmatrix} -0.11469 & 0.55734 & 0 & 0 \\ -0.04261 & 0.0924 & 0.37156 & 0 \\ -0.01583 & 0.03433 & 0.02337 & 0.37156 \\ -0.00882482 & 0.01913 & 0.01303 & 0.0924 \end{pmatrix}, \quad C(\omega) = \begin{pmatrix} 0.55734 \\ 0.57865 \\ 0.58657 \\ 0.88426 \end{pmatrix},$$

conduce la următoarele valori ale vectorului soluțiilor pentru primele cinci iterații:

$$x^{(1)} = \begin{pmatrix} 0.55734 \\ 0.57865 \\ 0.58657 \\ 0.88426 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 0.81593 \\ 0.82631 \\ 0.93988 \\ 0.97976 \end{pmatrix}, \quad x^{(3)} = \begin{pmatrix} 0.92431 \\ 0.96946 \\ 0.98803 \\ 0.99565 \end{pmatrix}, \quad x^{(4)} = \begin{pmatrix} 0.99166 \\ 0.99595 \\ 0.99825 \\ 0.99953 \end{pmatrix},$$

$$x^{(5)} = \begin{pmatrix} 0.9987 \\ 0.99933 \\ 0.99978 \\ 0.99993 \end{pmatrix}.$$

§1.12. Metoda gradienților conjugați

Fie sistemul $Ax=b$, unde A este simetrică şi pozitiv definită.

Definiția 1. *Spunem că direcțiile p şi q sunt direcții conjugate în raport cu matricea A dacă* $\langle Ap,q\rangle = \langle p,Aq\rangle = 0$.

Fie $v^{(0)}$ un vector de probă şi $r^{(0)}=Av^{(0)} - b$ vectorul rezidual corespunzător. În metoda gradienților conjugați pentru rezolvarea sistemului $Ax=b$, prima direcție de relaxare se alege $p^{(1)} = - r^{(0)}$. În continuare avem:

$$t_{\min} = -\frac{\langle r^{(0)},p^{(1)}\rangle}{\langle Ap^{(1)},p^{(1)}\rangle} = \frac{\langle r^{(0)},r^{(0)}\rangle}{\langle Ap^{(1)},p^{(1)}\rangle} \quad , \tag{1}$$

$$v^{(1)} = v^{(0)} + q_1 p^{(1)} \quad , \tag{2}$$

unde

$$q_1 = t_{\min} = -\frac{\langle r^{(0)},p^{(1)}\rangle}{\langle Ap^{(1)},p^{(1)}\rangle} \tag{3}$$

Valoarea minimă a funcției $F=F(v)$, când v parcurge dreapta ce trece prin $v^{(0)}$ şi are direcția $p^{(1)}= - r^{(0)}$ este $v^{(1)}$.

Fie

$$r^{(1)}=\mathrm{grad}F(v^{(1)})=Av^{(1)} - b.$$

Din Observația 1 din §8 rezultă că

$$\langle r^{(1)},p^{(1)}\rangle = -\langle r^{(1)},r^{(0)}\rangle = 0\ . \tag{4}$$

Următoarea direcție de relaxare $p^{(2)}$ se alege de forma $p^{(2)}= - r^{(1)}+c_1p^{(1)}$ şi în plus să fie conjugată direcției $p^{(1)}$ în raport cu A.

Aşadar, avem:

$$0 = \langle Ap^{(2)},p^{(1)}\rangle = \langle p^{(2)},Ap^{(1)}\rangle = -\langle r^{(1)},Ap^{(1)}\rangle + c_1\langle Ap^{(1)},p^{(1)}\rangle$$

de unde rezultă

$$c_1 = \frac{\langle r^{(1)},Ap^{(1)}\rangle}{\langle Ap^{(1)},p^{(1)}\rangle} \quad . \tag{5}$$

Avem de asemenea

$$v^{(2)} = v^{(1)} + t_{\min} p^{(2)} = v^{(1)} - \frac{\left\langle r^{(1)}, p^{(2)} \right\rangle}{\left\langle Ap^{(2)}, p^{(2)} \right\rangle} p^{(2)} \quad . \tag{6}$$

În general, pentru orice $k \geq 2$ obținem

$$c_{k\text{-}1} = \frac{\left\langle r^{(k-1)}, Ap^{(k-1)} \right\rangle}{\left\langle Ap^{(k-1)}, p^{(k-1)} \right\rangle} \quad , \tag{7}$$

$$p^{(k)} = -r^{(k-1)} + c_{k-1} p^{(k-1)} \quad , \tag{8}$$

$$q_k = -\frac{\left\langle r^{(k-1)}, p^{(k)} \right\rangle}{\left\langle Ap^{(k)}, p^{(k)} \right\rangle} , \tag{9}$$

$$v_k = v^{(k-1)} + q_k p^{(k)} . \tag{10}$$

Metoda gradienților conjugați este definită de formulele (7) – (10).

În continuare prezentăm unele simplificări şi proprietăți suplimentare. Deoarece $r^{(k-1)}$ este ortogonal pe direcția $p^{(k-1)}$ rezultă

$$\left\langle r^{(k-1)}, p^{(k)} \right\rangle = \left\langle r^{(k-1)}, -r^{(k-1)} + c_{k-1} p^{(k-1)} \right\rangle = -\left\langle r^{(k-1)}, r^{(k-1)} \right\rangle$$

şi deci

$$q_k = \frac{\left\langle r^{(k-1)}, r^{(k-1)} \right\rangle}{\left\langle Ap^{(k)}, p^{(k)} \right\rangle} > 0 . \tag{9'}$$

Pe de altă parte, din (10) avem

$$r^{(k)} = Av^{(k)} - b = Av^{(k-1)} + q_k Ap^{(k)} - b = r^{(k-1)} + q_k Ap^{(k)} \quad .$$

Obținem deci următoarea relația de recurență

$$r^{(k)} = r^{(k-1)} + q_k Ap^{(k)} \quad . \tag{11}$$

Observăm că

$$\left\langle r^{(k)}, r^{(k-1)} \right\rangle = 0 \quad . \tag{12}$$

Într –adevăr, din (11) rezultă

$$\left\langle r^{(k)}, r^{(k-1)} \right\rangle = \left\langle r^{(k-1)}, r^{(k-1)} \right\rangle + q_k \left\langle r^{(k-1)}, Ap^{(k)} \right\rangle \quad . \tag{13}$$

Pe de altă parte, ținând seama de (8), de (9') şi de faptul că $p^{(k)}$ şi $p^{(k-1)}$ sunt A – conjugate, rezultă

$$q_k \left\langle r^{(k-1)}, Ap^{(k)} \right\rangle = \left\langle r^{(k-1)}, r^{(k-1)} \right\rangle \frac{\left\langle r^{(k-1)}, Ap^{(k)} \right\rangle}{\left\langle Ap^{(k)}, -r^{(k-1)} + c_{k-1} p^{(k-1)} \right\rangle} = -\left\langle r^{(k-1)}, r^{(k-1)} \right\rangle$$

.

Din (13) şi (14) rezultă acum (12) .

Deoarece $Ap^{(k)}$ oricum trebuie calculat, rezultă că vectorul rezidual $r^{(k)}$ se va calcula din relația de recurență (11) şi nu prin înlocuirea directă a lui $v^{(k)}$ în expresia $Av=b$.

În continuare vom stabili o altă formulă pentru coeficientul c_{k-1}.
Din (11) şi (12) rezultă

$$\left\langle r^{(k-1)}, Ap^{(k-1)} \right\rangle = \left\langle r^{(k-1)}, \frac{1}{q_{k-1}}\left(r^{(k-1)} - r^{(k-2)}\right) \right\rangle = \frac{1}{q_{k-1}}\left\langle r^{(k-1)}, r^{(k-1)} \right\rangle .$$

Ținând seama de (7) şi de (9') obținem

$$c_{k-1} = \frac{\left\langle r^{(k-1)}, Ap^{(k-1)} \right\rangle}{\left\langle Ap^{(k-1)}, p^{(k-1)} \right\rangle} = \frac{\left\langle r^{(k-1)}, r^{(k-1)} \right\rangle}{\left\langle r^{(k-2)}, r^{(k-2)} \right\rangle} .$$

Algoritm pentru rezolvarea sistemelor de ecuații liniare cu metoda gradienților conjugați
Calculează

$$r^{(0)} = Av^{(0)} - b\ ;\ p^{(1)} = -r^{(0)};$$

$$q_1 = \frac{\left\langle r^{(0)}, r^{(0)} \right\rangle}{\left\langle Ap^{(1)}, p^{(1)} \right\rangle}\ ;\ v^{(1)} = v^{(0)} + q_1 p^{(1)}\ ;\ r^{(1)} = r^{(0)} + q_1 Ap^{(1)}\ ;$$

Pentru k:=2,n calculează

$$c_{k-1} = \frac{\left\langle r^{(k-1)}, r^{(k-1)} \right\rangle}{\left\langle r^{(k-2)}, r^{(k-2)} \right\rangle}\ ;\ p^{(k)} = -r^{(k-1)} + c_{k-1} p^{(k-1)}\ ;$$

$$q_k = \frac{\left\langle r^{(k-1)}, r^{(k-1)} \right\rangle}{\left\langle Ap^{(k)}, p^{(k)} \right\rangle}\ ;\ v^{(k)} = v^{(k-1)} + q_k p^{(k)}\ ;\ r^{(k)} = r^{(k-1)} + q_k Ap^{(k)}\ ;$$

sfârşit pentru k .
În Mathcad algoritmul de mai sus este aplicat unui exempu.

Metoda gradientilor conjugati

Folosind metoda gradientilor conjugati sa se gaseasca solutia sistemului de ecuatii liniare Ax=b

$$A := \begin{bmatrix} 5 & 1 & -1 & 1 \\ 1 & 6 & 2 & -2 \\ -1 & 2 & 7 & 1 \\ 1 & -2 & 1 & 8 \end{bmatrix} \qquad b := \begin{bmatrix} 6 \\ 7 \\ 9 \\ 8 \end{bmatrix}$$

Vectorul de proba $x := \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ $n := 4$

Algoritmul metodei gradientilor conjugati

$$\mathrm{GrCon}(A, b, x, n) := \left| \begin{array}{l} r^{<0>} \leftarrow A \cdot x - b \\ p \leftarrow -r^{<0>} \\ q \leftarrow \dfrac{r^{<0>} \cdot r^{<0>}}{A \cdot p \cdot p} \\ x \leftarrow x + q \cdot p \\ r^{<1>} \leftarrow r^{<0>} + q \cdot A \cdot p \\ \text{for } k \in 2..n \\ \quad \left| \begin{array}{l} c \leftarrow \dfrac{r^{<k-1>} \cdot r^{<k-1>}}{r^{<k-2>} \cdot r^{<k-2>}} \\ p \leftarrow -r^{<k-1>} + c \cdot p \\ q \leftarrow \dfrac{r^{<k-1>} \cdot r^{<k-1>}}{A \cdot p \cdot p} \\ x \leftarrow x + q \cdot p \\ r^{<k>} \leftarrow r^{<k-1>} + q \cdot A \cdot p \end{array} \right. \\ x \end{array} \right.$$

Apelarea programului si afisarea rezultatelor

$$\mathrm{GrCon}(A, b, x, n) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Teorema 1. *În metoda gradienților conjugați direcțiile de relaxare* $p^{(k)}$*, (*k*=1,2,...) sunt conjugate două câte două în raport cu matricea A, iar vectorii reziduali* $r^{(k)}$*, (*k*=0,1,...) sunt ortogonali doi câte doi.*

Demonstrație.

Demonstrația se face prin inducție relativ la k. Pentru k=1 avem $\left\langle r^{(1)}, r^{(0)} \right\rangle = 0$ din (4), iar pentru prima afirmație nu avem ce arăta. Ipoteza de inducție este:

$$\left\langle r^{(i)}, r^{(j)} \right\rangle = 0 \qquad \text{pentru} \quad i \neq j, \quad 0 \leq i, \quad j \leq k, \tag{15}$$

$$\left\langle p^{(i)}, Ap^{(j)} \right\rangle = 0 \qquad \text{pentru} \quad i \neq j, \quad 1 \leq i \leq j \leq k. \tag{16}$$

Va trebui să arătăm că

$$\left\langle r^{(k+1)}, r^{(j)} \right\rangle = 0, \qquad \text{pentru} \ j = \overline{0,k} \tag{17}$$

$$\left\langle p^{(k+1)}, Ap^{(j)} \right\rangle = 0, \quad \text{pentru} \quad j = \overline{1,k}. \tag{18}$$

Fie $j=k$, atunci

$$\left\langle p^{(k+1)}, Ap^{(k)} \right\rangle = \left\langle Ap^{(k+1)}, p^{(k)} \right\rangle = 0$$

deoarece p^k și p^{k+1} sunt A –conjugate.

Fie $1 \leq j < k$, atunci

$$\left\langle p^{(k+1)}, Ap^{(j)} \right\rangle = \left\langle -r^{(k)} + c_k p^{(k)}, Ap^{(j)} \right\rangle = -\left\langle r^{(k)}, Ap^{(j)} \right\rangle \quad ,$$

deoarece $\left\langle p^{(k)}, Ap^{(j)} \right\rangle = 0$ conform ipotezei (16).

Pe de altă parte,

$$Ap^{(j)} = \frac{1}{q_j}\left(r^{(j)} - r^{(j-1)}\right)$$

și din (15) rezultă $\left\langle r^{(k)}, Ap^{(j)} \right\rangle = 0$.

Așadar am demonstrat (18).

Pentru $j=k$, (17) este adevărată din (12).

Fie $0 \leq j < k$. Din (11) și (15) rezultă:

$$\left\langle r^{(k+1)}, r^{(j)} \right\rangle = \left\langle r^{(k)} + q_{k+1} Ap^{(k+1)}, r^{(j)} \right\rangle = q_{k+1}\left\langle Ap^{(k+1)}, r^{(j)} \right\rangle =$$

$$= q_{k+1}\left\langle Ap^{(k+1)}, -p^{(j)} + c_{j-1} p^{(j-1)} \right\rangle$$

Ținând seama de (18) și de faptul că A este simetrică, rezultă $\left\langle r^{(k+1)}, r^{(j)} \right\rangle = 0$ și cu aceasta teorema este demonstrată. ❑

Din Teorema 1 rezultă că vectorii reziduali $r^{(k)}$ sunt ortogonali doi câte doi și deci sunt liniar independenți (dacă sunt nenuli). Așadar, nu pot exista (n+1) vectori reziduali nenuli. Rezultă că în metoda gradienților conjugați soluția exactă se găseşte în cel mult n paşi. Teoretic ar trebui ca vectorul rezidual $r^{(n)}$ să fie zero şi deci $v=v^{(n)}$ să fie soluţia exactă a sistemului.

În practică acest lucru nu se întâmplă, deoarece în determinarea vectorilor $r^{(k)}$ intervin erori de calcul, care fac ca aceştia să nu formeze un sistem ortogonal. Deoarece în general $r^{(n)}\neq 0$, continuăm să calculăm $r^{(k)}$, $k>n$ până obţinem un vector rezidual nul sau foarte mic $\left(\left\|r^{(k)}\right\|<<1\right)$. Această atitudine se justifică prin faptul că metoda gradienţilor conjugaţi este o metodă de relaxare prin care valoarea funcţiei pătratice $F=F(v)$ se micşorează la fiecare pas.

Metoda gradienţilor conjugaţi se dovedeşte foarte utilă pentru sistemele în care matricea A are multe zerouri şi fiecare ecuaţie are o anumită regularitate internă.

Astfel de sisteme apar în procesul de discretizare a problemei la limită a ecuaţiilor cu derivate parţiale de tip eliptic.

§1.13. Metoda celor mai mici pătrate

În procesul de prelucrare şi ajustare a datelor, apar sisteme de ecuaţii liniare supradimensionate sau subdimensionate. Abordăm pentru început problema sistemelor supradimensionate. Fie sistemul

$$\begin{cases} a_{11}x_1+...+a_{1n}x_n=b_1 \\ \dots\dots\dots\dots \\ a_{m1}x_1+...+a_{mn}x_n=b_m \end{cases} \quad , \quad m>n \quad . \tag{1}$$

Evident, un asemenea sistem poate să nu aibă soluţie. Fie

$$r=Ax-b=(r_1,r_2,...,r_m)^T \tag{2}$$

vectorul rezidual , unde

$$r_i=a_{i1}x_1+...+a_{in}x_n-b_i \quad , \; i=\overline{1,m} \quad .$$

Considerăm funcţia pătratică

$$f(x)=f(x_1,...,x_n)=\langle r,r\rangle=r_1^2+r_2^2+...+r_m^2 \tag{3}$$

Definiţia 1. *Se numeşte soluţie în sensul celor mai mici pătrate a sistemului* (1)*, acel vector* x^**, pentru care funcţia* (2) *are valoarea minimă.*

Dacă: $\min\limits_{x\in R^n} f(x)=f(x^*)=0$, atunci $r_i(x^*)=0$, pentru orice $i=\overline{1,m}$. Rezultă că sistemul (1) este compatibil şi $x=x^*$ este soluţia exactă a sa.

În general, sistemul (1) nu este compatibil şi $\min\limits_{x\in \mathrm{R}^n} f(x)= f(x^*) > 0$, iar $x=x^*$ este un substitut pentru *soluția sistemului* şi anume soluția în sensul celor mai mici pătrate.

Funcția f se poate pune sub forma

$$f(x)=\langle r,r\rangle = \langle Ax\text{-}b,Ax\text{-}b\rangle = \langle Ax,Ax\rangle - 2\langle Ax,b\rangle + \langle b,b\rangle$$

şi mai departe

$$f(x)=\left\langle A^T Ax,x\right\rangle - 2\left\langle A^T b,x\right\rangle + \langle b,b\rangle \qquad (4)$$

Teorema 1. *Dacă* $\operatorname{rang}A = n$, *atunci sistemul (1) admite o singură soluție în sensul celor mai mici pătrate şi aceasta este soluția (unică) a sistemului.*

$$A^T Ax=A^T b \qquad (5)$$

(Sistemul (5) *se numeşte sistemul normal al lui Gauss).*

Demonstrație.

Punctele de extrem ale funcției pătratice f dată de (4), se caută printre punctele sale critice, iar acestea, se află rezolvând sistemul:

$$\operatorname{grad} f = 0$$

Cum $\operatorname{grad} f=A^TAx - A^Tb$, obținem sistemul $A^TAx=A^Tb$. Pe de altă parte se ştie că:

$$\operatorname{rang}A=\operatorname{rang}A^T=\operatorname{rang}\left(A^T A\right)=\operatorname{rang}\left(AA^T\right).$$

Matricea $B=A^TA$ este o matrice pătratică de ordinul n şi $\operatorname{rang}B=n$, conform celor de mai sus. Rezultă că sistemul (5) admite o soluție unică, $x=x^*$, care este punct critic pentru f.

Matricea B este evident simetrică şi semipozitiv definită. Mai mult, în ipoteza noastră, matricea B este pozitiv definită. Într –adevăr, dacă presupunem că $\langle Bx,x\rangle = 0$, atunci rezultă $\langle Ax,Ax\rangle = 0$ şi deci $Ax=0$. Cum $\operatorname{rang}A=n<m$ rezultă $x=0$.

Pe de altă parte avem

$$d^2 f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\,dx_i\,dx_j > 0,$$

de unde rezultă că $x=x^*$ este punct de minim pentru f şi cu aceasta teorema este demonstrată. ❑

Aşadar, în ipoteza $\operatorname{rang}A=n$, soluția sistemului (1), în sensul celor mai mici pătrate, este unică şi se află rezolvând sistemul (5). Acest sistem este simetric pozitiv definit. Rezolvarea sa se poate face prin metoda Cholesky sau una din metodele de relaxare.

Observația 1. *Teoretic, soluția sistemului* (5) *este* $x^*=(A^TA)^{-1}A^Tb$. *Matricea* $P=(A^TA)^{-1}A^T$ *se numeşte pseudoinversa matricei (dreptunghiu –lare)* A.

Se observă că dacă A este pătratică, atunci $P=A^{-1}(A^T)^{-1}A^T=A^{-1}$, deci noțiunea de matrice pseudoinversă generalizează noțiunea de matrice inversă (pentru matrice dreptunghiulare).
Rezolvarea practică a sistemului (5) ridică probleme din cauza faptului că numărul de condiționare al matricei $B=A^TA$ este mare. Fie $\lambda_1 \geq \lambda_2 \geq .. \geq \lambda_n > 0$ valorile proprii ale matricei B. Atunci:

$$\mathrm{cond}(B) = \frac{\lambda_1}{\lambda_n} \quad . \tag{6}$$

Cum

$$\lambda_1 = \sup_{x\neq 0}\frac{\langle Bx,x\rangle}{\langle x,x\rangle} \geq \max_i\langle Be_i,e_i\rangle = \max_i b_{ii} \quad \text{şi} \quad \lambda_n = \inf_{x\neq 0}\frac{\langle Bx,x\rangle}{\langle x,x\rangle} \leq \min_i\langle Be_i,e_i\rangle = \min_i b_{ii},$$

rezultă

$$\mathrm{cond}(B) \geq \frac{\max b_{ii}}{\min b_{ii}} \quad . \tag{7}$$

Exemplul 1. *(Dreapta de regresie)*
Să presupunem că vrem să găsim o dreaptă $y=mx+n$ care să treacă prin punctele:
$M_1(0,0)$; $M_2(2,1)$; $M_3(5,3)$; $M_4(8,5)$; $M_6(10,6)$
Se obține astfel sistemul

$$\begin{cases} 0\cdot m+n=0 \\ 2\cdot m+n=1 \\ 5\cdot m+n=3 \\ 8\cdot m+n=5 \\ 10\cdot m+n=6 \end{cases} \quad . \tag{8}$$

Evident, sistemul (8) este supradimensionat şi incompatibil. Avem:

$$A = \begin{pmatrix} 0 & 1 \\ 2 & 1 \\ 5 & 1 \\ 8 & 1 \\ 10 & 1 \end{pmatrix}; \ b = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 5 \\ 6 \end{pmatrix}; \ B = A^T A = \begin{pmatrix} 193 & 25 \\ 25 & 5 \end{pmatrix}; \ A^T b = \begin{pmatrix} 117 \\ 15 \end{pmatrix} \quad .$$

Ecuațiile normale ale lui Gauss sunt

$$\begin{cases} 193m + 25n = 117 \\ 25m + 5n = 15 \end{cases} \quad .$$

Soluția exactă este $m = \frac{21}{34}, n = -\frac{3}{34}$, iar valorile proprii sunt $\lambda_1 = 196.268$ şi $\lambda_2 = 1.732$. Rezultă $\mathrm{cond}(B) \cong 113$. Dacă folosim estimarea (7) obținem

$$\frac{\max b_{ii}}{\min b_{ii}} = \frac{193}{5} = 38{,}6 \,.$$

Dreapta de regresie în acest caz este: $y = \frac{21}{34}x - \frac{3}{34}$. Această dreaptă nu trece prin punctele M_i, dar este acea dreaptă din plan care trece *cel mai aproape* de aceste puncte.

Să presupunem că vrem să determinăm dreapta de regresie corespunzătoare punctelor $M_i(x_i, y_i)$, $i = \overline{1,n}$.

Matricele

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ n & 1 \end{pmatrix} \quad \text{şi} \quad B = A^T A = \begin{pmatrix} \frac{n(n+1)(2n+1)}{6} & \frac{n(n+1)}{2} \\ \frac{n(n+1)}{2} & n \end{pmatrix}$$

conduc la

$$\frac{\max b_{ii}}{\min b_{ii}} = \frac{(n+1)(2n+1)}{6} > \frac{n^2}{3} \,.$$

Pentru n=100, cond(B)>13333, deci sistemul normal al lui Gauss este prost condiționat.

Sistemele subdimensionate apar în probleme legate de ajustarea datelor. Să presupunem că măsurând n cantități $x_1, x_2, \ldots, x_n$, găsim valorile $l_1, l_2, \ldots, l_n$. Pe de altă parte, necunoscutele $x_1, x_2, \ldots, x_n$ satisfac anumite ecuații şi anume

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases} , \quad m < n \,. \qquad (9)$$

Datorită lipsei de acuratețe a măsurătorilor se pune condiția ca suma pătratelor corecțiilor $\sum_{i=1}^{n}(l_i - x_i)^2$ să fie minimă. Se obține astfel o problemă de extrem cu legături.

Exemplul 2.

Să presupunem că măsurând unghiurile x_1, x_2, x_3 ale unui triunghi găsim valorile l_1, l_2, l_3. Evident avem legătura

$$x_1 + x_2 + x_3 = 180° \,.$$

Impunem condiția ca abaterile datorate impreciziei măsurătorilor să fie cât mai mici, deci ca expresia

$$(x_1 - l_1)^2 + (x_2 - l_2)^2 + (x_3 - l_3)^2$$

să fie minimă.

Revenind la cazul general, se pune problema să determinăm n necunoscute x_1, x_2, ... , x_n care satisfac legăturile (9) şi care minimizează funcția:

$$f(x_1,...,x_n) = (x_1 - l_1)^2 + ... + (x_n - l_n)^2 \ .$$

Considerăm funcția auxiliară

$$\Phi(x_1,...,x_n,\lambda_1,...,\lambda_m) = (x_1 - l_1)^2 + ... + (x_n - l_n)^2 -$$
$$-2\lambda_1(a_{11}x_1 + ... + a_{1n}x_n - b_1)... - 2\lambda_m(a_{m1}x_1 + ... + a_{mn}x_n - b_m)$$

Punctele critice ale funcției ϕ sunt date de sistemul

$$\begin{cases} x_1 - l_1 - (a_{11}\lambda_1 + ... + a_{m1}\lambda_m) = 0 \\ .... \\ x_n - l_n - (a_{1n}\lambda_1 + ... + a_{mn}\lambda_m) = 0 \\ a_{11}x_1 + ... + a_{1n}x_n - b_1 = 0 \\ .... \\ a_{m1}x_1 + ... + a_{mn}x_n - b_m = 0 \end{cases} . \qquad (10)$$

Dacă notăm cu $l = (l_1,...,l_n)^T$ şi cu $\lambda = (\lambda_1,...,\lambda_m)^T$, atunci sistemul (10) se scrie sub forma matricială

$$x = A^T\lambda + l \qquad (10')$$
$$Ax = b \qquad (10'')$$

Înlocuind (10') în (10'') obținem sistemul

$$AA^T\lambda = b - Al \ . \qquad (11)$$

Dacă $\operatorname{rang} A = m$, atunci $\operatorname{rang}(AA^T) = m$ şi sistemul (11) are soluție unică. Mai mult, matricea AA^T este pătratică, simetrică şi pozitiv definită, deci rezolvarea sistemului (11) se poate face cu metoda Cholesky sau una din metodele de relaxare. Rezolvând sistemul (11) găsim multiplicatorii lui Lagrange $\lambda_1,...,\lambda_m$, iar din relația $x = A^T\lambda + l$ determinăm punctul de minim condiționat al funcției *f.*
(Cum $d^2\phi = 2(dx_1^2 + ... + dx_n^2) > 0$ rezultă că avem într –adevăr un punct de minim).

Exerciții

Folosind metoda Gauss să se rezolve următoarele sisteme de ecuații liniare:

1. $$\begin{cases} 6x_1 + x_2 - x_3 + x_4 = 2 \\ x_1 + 7x_2 + 2x_3 - x_4 = -1 \\ -x_1 + 2x_2 + 8x_3 + x_4 = 12 \\ x_1 - x_2 + x_3 + 9x_4 = -5 \end{cases}$$

R. Matricea triunghiulară $\widetilde{A} = \begin{pmatrix} 1.0 & 0.1667 & -0.1667 & 0.1667 \\ 0 & 1.0 & 0.3171 & -0.1707 \\ 0 & 0 & 1.0 & 0.2150 \\ 0 & 0 & 0 & 1.0 \end{pmatrix}$,

vectorul termenilor liberi transformat $\widetilde{b} = \begin{pmatrix} 0.333 \\ -0.1951 \\ 1.7850 \\ -1.0 \end{pmatrix}$.

Soluția sistemului $x = \begin{pmatrix} 1 \\ -1 \\ 2 \\ -1 \end{pmatrix}$.

2. $$\begin{cases} 4x_1 + x_2 - x_3 + x_4 = 1 \\ x_1 + 8x_2 + x_3 + x_4 = 7 \\ x_1 + x_2 + 6x_3 + x_4 = -10 \\ x_1 + x_2 + x_3 + 7x_4 = 12 \end{cases}$$

R. Matricea triunghiulară $\widetilde{A} = \begin{pmatrix} 1.0 & 0.2500 & -0.2500 & 0.2500 \\ 0 & 1.0 & 0.1613 & 0.0968 \\ 0 & 0 & 1.0 & 0.1105 \\ 0 & 0 & 0 & 1.0 \end{pmatrix}$,

vectorul termenilor liberi transformat $\widetilde{b} = \begin{pmatrix} 0.2500 \\ 0.8710 \\ -1.7789 \\ 2.0 \end{pmatrix}$.

Soluția sistemului $x = (-1\ 1\ -2\ 2)'$.

3. $$\begin{cases} 5x_1 + x_2 + x_3 + x_4 & = 14 \\ x_1 + 8x_2 + 2x_3 + x_4 & = 127 \\ -x_1 + x_2 + 6x_3 - x_4 & = 15 \\ x_1 + x_2 + x_3 + x_4 & = 10 \end{cases}$$

R. Matricea triunghiulară $\widetilde{A} = \begin{pmatrix} 1.0 & 0.2000 & -0.2000 & 0.2000 \\ 0 & 1.0 & 0.2308 & 0.1026 \\ 0 & 0 & 1.0 & -0.1558 \\ 0 & 0 & 0 & 1.0 \end{pmatrix}$,

vectorul termenilor liberi transformat $\widetilde{b} = \begin{pmatrix} 2.8000 \\ 3.1026 \\ 2.3766 \\ 4.0 \end{pmatrix}$.

Soluția sistemului $x = (1\ 2\ 3\ 4)'$.

Să se rezolve următoarele sisteme de ecuații liniare folosind metoda Cholesky:

4. $$\begin{cases} 10x_1 + x_2 - x_3 - x_4 & = 9 \\ x_1 + 9x_2 + x_3 - x_4 & = -6 \\ -x_1 + x_2 + 11x_3 + x_4 & = 8 \\ -x_1 - x_2 + x_3 + 8x_4 & = -7 \end{cases}$$

R. $R^T = \begin{pmatrix} 3.1623 & 0 & 0 & 0 \\ 0.3162 & 2.9833 & 0 & 0 \\ -0.3162 & 0.3687 & 0.32809 & 0 \\ -0.3162 & -0.3017 & 0.3082 & 2.7774 \end{pmatrix}$. Soluția sistemului $R^Ty=b$ este $y = \begin{pmatrix} 2.846 \\ -2.3129 \\ 2.9726 \\ -2.7774 \end{pmatrix}$, iar cea a sistemului $Rx=y$, şi deci soluția sistemului inițial este $x = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$.

5. $$\begin{cases} 8x_1 - x_2 - x_3 + x_4 & = 7 \\ -x_1 + 6x_2 + x_3 - 2x_4 & = 6 \\ -x_1 + x_2 + 7x_3 + x_4 & = -8 \\ x_1 - 2x_2 + x_3 + 9x_4 & = -11 \end{cases}$$

R. $$R^T = \begin{pmatrix} 2.82843 & 0 & 0 & 0 \\ -0.35355 & 2.42384 & 0 & 0 \\ -0.35355 & 0.361 & 2.59705 & 0 \\ 0.35355 & -0.77357 & 0.54071 & 2.82564 \end{pmatrix}.$$

Soluția sistemului $R^Ty=b$ este $y = \begin{pmatrix} 2.47487 \\ 2.83641 \\ -3.13776 \\ -2.82564 \end{pmatrix}$ iar cea a sistemului $Rx=y$, şi deci soluția sistemului inițial este $x=\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$.

6. $$\begin{cases} 6x_1 + x_2 - 2x_3 - x_4 & = 19 \\ x_1 + 5x_2 + x_3 - x_4 & = 15 \\ -2x_1 + x_2 + 11x_3 - x_4 & = 8 \\ -x_1 - x_2 - x_3 + 4x_4 & = -10 \end{cases}$$

R. $$R^T = \begin{pmatrix} 2.44949 & 0 & 0 & 0 \\ 0.40825 & 2.19848 & 0 & 0 \\ -0.8165 & 0.60648 & 3.15682 & 0 \\ -0.40825 & -0.37905 & -0.34954 & 1.88878 \end{pmatrix}.$$

Soluția sistemului $R^Ty=b$ este $y = \begin{pmatrix} 7.75672 \\ 5.3825 \\ 3.50636 \\ -1.88878 \end{pmatrix}$, iar cea a sistemului $Rx=y$, şi deci soluția sistemului inițial este $x=(3\ 2\ 1\ -1)'$.

Folosind metoda Householder, să se rezolve sistemele:

7. $$\begin{cases} 4x_1+3x_2-2x_3+x_4 &= 2 \\ 5x_1+6x_2+7x_3-8x_4 &= -14 \\ 9x_1-8x_2+7x_3-6x_4 &= -30 \\ x_1+x_2-x_3-x_4 &= 0 \end{cases}$$

R. det(A)= –1112; descompunerea $A=QR$ este dată de

$$Q=\begin{pmatrix} 0.3607 & -0.3882 & 0.7877 & -0.3143 \\ 0.4508 & -0.7068 & -0.5445 & 0.0257 \\ 0.8115 & 0.5787 & -0.0784 & 0.02 \\ 0.0902 & -0.1217 & 0.2774 & 0.9487 \end{pmatrix},$$

$$R=\begin{pmatrix} 11.0905 & -2.6148 & 8.0249 & -8.0249 \\ 0 & -10.1569 & 0.0016 & 1.9155 \\ 0 & 0 & -6.2130 & 5.83368 \\ 0 & 0 & 0 & -1.58889 \end{pmatrix}.$$

Soluția sistemului $Qy=b$ este $y=\begin{pmatrix} 29.9354 \\ -8.2430 \\ 11.5498 \\ -1.5888 \end{pmatrix}$,

iar cea sistemului $Rx=y$ este $x=(-1\ 1\ -1\ 1)'$.

8. $$\begin{cases} x_1+2x_2-3x_3-4x_4 &= 16 \\ 5x_1-6x_2+7x_3-8x_4 &= 2 \\ 9x_1+8x_2+7x_3+6x_4 &= 6 \\ 4x_1-3x_2+2x_3-x_4 &= -2 \end{cases}.$$

R. det(A)= –2808; descompunerea $A=QR$ este dată

$$Q=\begin{pmatrix} 0.09017 & 0.17005 & 0.78843 & -0.58424 \\ 0.45083 & -0.71359 & -0.2494 & -0.47469 \\ 0.8115 & 0.55307 & -0.18504 & 0.03651 \\ 0.36067 & -0.39494 & 0.53098 & 0.65727 \end{pmatrix},$$

$$R=\begin{pmatrix} 11.09054 & 2.88534 & 9.2872 & 0.541 \\ 0 & 10.23107 & -2.42367 & 8.7419 \\ 0 & 0 & -4.3444 & -2.79972 \\ 0 & 0 & 0 & 5.69631 \end{pmatrix}.$$

Soluția sistemului $Qy=b$ este $y=\begin{pmatrix} 6.49202 \\ 5.40201 \\ 9.94384 \\ -11.39263 \end{pmatrix}$,

iar cea sistemului $Rx=y$ este $x=\begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}$.

9. $$\begin{cases} 4x_1+3x_2+2x_3+x_4 &= 2 \\ x_1+2x_2+3x_3+4x_4 &= -2 \\ 9x_1-8x_2+6x_3-7x_4 &= 30 \\ x_1-x_2+x_3-x_4 &= 4 \end{cases}.$$

R. det(A)= –80 ; descompunerea $A=QR$ este dată de

$$Q=\begin{pmatrix} 0.40202 & -0.82257 & 0.3961 & 0.0697 \\ 0.1005 & -0.39663 & -0.90448 & -0.12039 \\ 0.90453 & 0.4028 & -0.05925 & -0.12673 \\ 0.1005 & 0.06173 & -0.14663 & 0.98213 \end{pmatrix},$$

$$R=\begin{pmatrix} 9.94987 & -5.92972 & 6.63325 & -5.62821 \\ 0 & -6.5451 & -0.3565 & -5.29041 \\ 0 & 0 & -2.42341 & -2.66043 \\ 0 & 0 & 0 & -0.50691 \end{pmatrix}.$$

Soluția sistemului $Qy=b$ este $y=\begin{pmatrix} 28.14106 \\ 11.47901 \\ 0.23702 \\ 0.50691 \end{pmatrix}$,

iar cea sistemului $Rx=y$ este $x=\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$.

10. Pentru matricea $A=\begin{pmatrix}1 & 2 & 1 & -1\\ 2 & 1 & -1 & 3\\ 2 & 1 & -1 & 1\\ 1 & -1 & -1 & 1\end{pmatrix}$, să se calculeze:

a) $\det(A)$, A^{-1}, $\det(A^{-1})$;

b) $\|A\|_1$, $\|A\|_2$, $\|A\|_\infty$; $\|A^{-1}\|_1$, $\|A^{-1}\|_2$, $\|A^{-1}\|_\infty$;

c) $\mathrm{cond}_1(A)$, $\mathrm{cond}_2(A)$, $\mathrm{cond}_\infty(A)$.

R. a) $\det(A)=6$, $A^{-1}=\begin{pmatrix}0.6667 & 0 & -0.3333 & 1\\ -0.3333 & 0 & 0.6667 & -1\\ 1 & 0.5 & -1.5 & 1\\ 0 & 0.5 & -0.5 & 0\end{pmatrix}$,

$\det(A^{-1})=-0.1667$;

b) $\begin{cases}\|A\|_1=6,\ \|A\|_2=5.7446,\ \|A\|_\infty=7;\\ \|A^{-1}\|_1=3,\ \|A^{-1}\|_2=2.848,\ \|A^{-1}\|_\infty=4.\end{cases}$

c) $\mathrm{cond}_1(A)=18$, $\mathrm{cond}_2(A)=12.7511$, $\mathrm{cond}_\infty(A)=28$.

11. Să se calculeze:

a) $\det(A)$, A^{-1}, $\det(A^{-1})$;

b) $\|A\|_1$, $\|A\|_2$, $\|A\|_\infty$; $\|A^{-1}\|_1$, $\|A^{-1}\|_2$, $\|A^{-1}\|_\infty$;

c) $\mathrm{cond}_1(A)$, $\mathrm{cond}_2(A)$, $\mathrm{cond}_\infty(A)$

pentru matricea $A=\begin{pmatrix}4 & -1 & 5\\ 3 & 2 & -2\\ 6 & 1 & 3\end{pmatrix}$.

R. a) $\det(A)=8$, $A^{-1}=\begin{pmatrix}1 & 1 & -1\\ -2.625 & -2.25 & 2.875\\ -1.125 & -1.25 & 1.375\end{pmatrix}$, $\det(A^{-1})=0.125$;

b) $\begin{cases}\|A\|_1=6,\ \|A\|_2=5.7446,\ \|A\|_\infty=7;\\ \|A^{-1}\|_1=3,\ \|A^{-1}\|_2=2.848,\ \|A^{-1}\|_\infty=4.\end{cases}$

c) $\mathrm{cond}_1(A)=68.25$, $\mathrm{cond}_2(A)=48.28525$, $\mathrm{cond}_\infty(A)=77.5$.

Folosind metoda Iacobi să se găsească soluția aproximativă pentru următoarele sisteme de ecuații liniare :

12. $\begin{cases} 5x_1 - x_2 - x_3 & = 5 \\ x_1 + 6x_2 + x_3 & = -4 \\ x_1 - x_2 + 7x_3 & = 9 \end{cases}$

R. Sistemul se poate pune sub forma $x = Bx+c$, unde $B = D^{-1} \cdot (D–A)$, iar $D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{pmatrix}$, $E = D–A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$,

$B = \begin{pmatrix} 0 & 0.2 & 0.2 \\ -0.16667 & 0 & -0.16667 \\ -0.14286 & 0.14286 & 0 \end{pmatrix}$, $c = D^{-1} \cdot b = \begin{pmatrix} 1 \\ -0.66667 \\ 1.28571 \end{pmatrix}$, $\|B\|_2$=0.41997,

$\rho(B)$= 0.29277. Soluția la fiecare iterație, pornind cu $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, este:

$$x^{(1)} = \begin{pmatrix} 1 \\ -0.66667 \\ 1.28571 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 1.12381 \\ -1.04762 \\ 1.04762 \end{pmatrix},$$

$$x^{(3)} = \begin{pmatrix} 1 \\ -1.02857 \\ 0.97551 \end{pmatrix}, \quad x^{(4)} = \begin{pmatrix} 0.98939 \\ -0.99592 \\ 0199592 \end{pmatrix}.$$

Soluția exactă fiind $x^* = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, eroarea care se face dacă se reține ca soluție aproximativă $x^{(4)}$, este $\left\| x^* - x^{(4)} \right\| = 0.01208$.

13. $\begin{cases} 6x_1 + x_2 - x_3 + x_4 & = 2 \\ x_1 + 7x_2 + x_3 - x_4 & = -4 \\ x_1 + 2x_2 + 8x_3 + x_4 & = 6 \\ x_1 + x_2 + x_3 + 9x_4 & -8 \end{cases}$ (să se scrie soluția obținută după patru iterații).

R. Se poate pune sistemul sub forma $x = Bx+c$, unde $B = D^{-1} \cdot E$, iar

$$D = \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}, \quad E = D–A = \begin{pmatrix} 0 & -1 & 1 & 2 \\ -1 & 0 & -1 & 1 \\ -1 & -2 & 0 & -1 \\ 1 & 1 & 1 & 0 \end{pmatrix},$$

$$B=\begin{pmatrix} 0 & -0.16667 & 0.16667 & -0.33333\\ -0.14286 & 0 & -0.14286 & 0.14286\\ -0.125 & -0.25 & 0 & -0.125\\ -0.11111 & -0.11111 & -0.11111 & 0 \end{pmatrix},$$

$c=D^{-1}\cdot b=\begin{pmatrix} 0.33333\\ -0.57143\\ 0.75\\ -0.88889 \end{pmatrix}$, $\|B\|_2$=0.5989, $\rho(B)$= 0.27926. Soluția pentru primele 4 iterații, pornind cu $x^{(0)}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$, este: $x^{(1)}=\begin{pmatrix} 0.33333\\ -0.57143\\ 0.75\\ -0.88889 \end{pmatrix}$,

$$x^{(2)}=\begin{pmatrix} 0.84987\\ -0.85317\\ 0.9623\\ -0.94577 \end{pmatrix},\ x^{(3)}=\begin{pmatrix} 0.95117\\ -0.96542\\ 0.97528\\ -0.99544 \end{pmatrix},\ x^{(4)}=\begin{pmatrix} 0.9886\\ -0.98884\\ 0.99689\\ -0.99567 \end{pmatrix}.$$

Soluția exactă fiind $x^*=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}$, eroarea care se face dacă se reține ca soluție aproximativă $x^{(4)}$, este $\left\|x^*-x^{(4)}\right\|=0.01682$.

14. $$\begin{cases} 4x_1-x_2-x_3+x_4 &=3\\ x_1+7x_2+x_3-x_4 &=14\\ x_1-x_2+6x_3+x_4 &=21\\ x_1-x_2-x_3+9x_4 &=38 \end{cases}$$

R. Se poate pune sistemul sub forma $x=Bx+c$, unde $B=D^{-1}\cdot E$, iar

$$D=\begin{pmatrix} 4&0&0&0\\ 0&7&0&0\\ 0&0&6&0\\ 0&0&0&9 \end{pmatrix},\quad E=D-A=\begin{pmatrix} 0&1&1&-1\\ -1&0&-1&1\\ -1&1&0&-1\\ -1&1&-1&0 \end{pmatrix},$$

$$B=\begin{pmatrix} 0 & 0.25 & 0.25 & -0.25 \\ -0.14286 & 0 & -0.14286 & 0.14286 \\ -0.16667 & 0.16667 & 0 & -0.16667 \\ -0.11111 & 0.11111 & -0.11111 & 0 \end{pmatrix},$$

$c=D^{-1}\cdot b=\begin{pmatrix} 0.75 \\ 2 \\ 3.5 \\ 4.22222 \end{pmatrix}$, $\|B\|_2$=0.60753, $\rho(B)$= 0.22758. Pentru primele 4 iterații,

pornind cu $x^{(0)}=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, soluția este: $x^{(1)}=\begin{pmatrix} 0.75 \\ 2 \\ 3.5 \\ 4.22222 \end{pmatrix}$,

$$x^{(2)}=\begin{pmatrix} 1.06944 \\ 1.99603 \\ 3.00463 \\ 3.97222 \end{pmatrix},\ x^{(3)}=\begin{pmatrix} 1.00711 \\ 1.98545 \\ 2.99239 \\ 3.99133 \end{pmatrix},\ x^{(4)}=\begin{pmatrix} 0.99663 \\ 1.99883 \\ 2.99784 \\ 3.99844 \end{pmatrix}.$$

Soluția exactă fiind $x^*=\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, eroarea care se face dacă se reține ca soluție aproximativă $x^{(4)}$, este $\left\|x^*-x^{(4)}\right\|=0.00417$.

Să se rezolve cu metoda Gauss–Seidel următoarele sisteme:

15. $\begin{cases} 5x_1-2x_2-2x_3 & =-3 \\ -x_1+6x_2-x_3 & =2 \\ -x_1-3x_2+4x_3 & =8 \end{cases}$

R. Se observă că matricea coeficienților sistemului $A=\begin{pmatrix} 5 & -2 & -2 \\ -1 & 6 & -1 \\ -1 & -3 & 4 \end{pmatrix}$ este tare diagonal dominantă şi atunci algoritmul Gauss–Seidel este convergent.

Sistemul se poate pune sub forma $x^{(m+1)}=(D+L)^{-1}(-Ux^{(m)}+b)$, $m\geq 0$, unde:

$$L=\begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & -3 & 0 \end{pmatrix},\ U=\begin{pmatrix} 0 & -2 & -2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix},\ D=\begin{pmatrix} 5 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Atunci

$$(D+L)^{-1} = \begin{pmatrix} 0.2 & 0 & 0 \\ 0.03333 & 0.16667 & 0 \\ 0.075 & 0.125 & 0.25 \end{pmatrix} ,$$

$$-(D+L)^{-1}U = \begin{pmatrix} 0 & 0.4 & 0.4 \\ 0 & 0.06667 & 0.23333 \\ 0 & 0.15 & 0.275 \end{pmatrix} ,$$

$(D+L)^{-1}b = \begin{pmatrix} -0.6 \\ 0.23333 \\ 2.025 \end{pmatrix}$. Se obțin pornind cu $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, vectorii:

$$x^{(1)} = \begin{pmatrix} -0.6 \\ 0.23333 \\ 2.025 \end{pmatrix} , \; x^{(2)} = \begin{pmatrix} 0.30333 \\ 0.72139 \\ 2.61687 \end{pmatrix} , \; x^{(3)} = \begin{pmatrix} 0.73531 \\ 0.89203 \\ 2.85285 \end{pmatrix} ,$$

$$x^{(4)} = \begin{pmatrix} 0.89795 \\ 0.95847 \\ 2.94334 \end{pmatrix} .$$

16. $\begin{cases} 6x_1 - x_2 - 2x_3 + x_4 = 2 \\ 2x_1 + 6x_2 - x_3 - x_4 = 8 \\ -x_1 - 3x_2 + 8x_3 + 2x_4 = 2 \\ x_1 + x_2 + x_3 + 4x_4 = -1 \end{cases}$

R. Se observă că matricea coeficienților sistemului

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

este tare diagonal dominantă şi algoritmul Gauss–Seidel este convergent.
Sistemul se poate pune sub forma $x^{(m+1)} = (D+L)^{-1}(-Ux^{(m)}+b)$, $m \geq 0$, unde:

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

$$D=\begin{pmatrix}6&0&0&0\\0&6&0&0\\0&0&8&0\\0&0&0&4\end{pmatrix}.$$

Atunci $(D+L)^{-1}=\begin{pmatrix}0.16667&0&0&0\\-0.05556&0.16667&0&0\\0&0.0625&0.125&0\\-0.02778&-0.05729&-0.03125&0.25\end{pmatrix}$,

$$-(D+L)^{-1}U=\begin{pmatrix}0&0.16667&0.33333&-0.16667\\0&-0.05556&0.05556&0.22222\\0&0&0.0625&-0.1875\\0&-0.02778&-0.11285&0.03299\end{pmatrix},$$

$$(D+L)^{-1}b=\begin{pmatrix}0.33333\\1.22222\\0.75\\-0.82639\end{pmatrix}.$$

Alegând $x^{(0)}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$, obținem: $x^{(1)}=\begin{pmatrix}0.5\\1\\0.9375\\-0.85938\end{pmatrix}$, $x^{(2)}=\begin{pmatrix}0.95573\\1.02778\\0.96973\\-0.98831\end{pmatrix}$,

$$x^{(3)}=\begin{pmatrix}0.99259\\0.99937\\0.99592\\-0.99697\end{pmatrix},\ x^{(4)}=\begin{pmatrix}0.99803\\1.00048\\0.99918\\-0.99942\end{pmatrix},\ x^{(5)}=\begin{pmatrix}0.99971\\1.00006\\0.99984\\-0.9999\end{pmatrix}.$$

17. $\begin{cases}10x_1-x_2-x_3+x_4=7\\x_1+11x_2-x_3-x_4=-10\\-x_1-x_2+12x_3+x_4=22\\x_1+x_2+x_3+13x_4=-24\end{cases}$

R. Se observă că matricea coeficienților sistemului

$$A=\begin{pmatrix}10&-1&-1&1\\1&11&-1&-1\\-1&-1&12&1\\1&1&1&13\end{pmatrix}$$

este tare diagonal dominantă şi algoritmul Gauss–Seidel este convergent.

Sistemul se poate pune sub forma $x^{(m+1)}=(D+L)^{-1}(-Ux^{(m)}+b)$, $m\geq 0$, unde:

$$L=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \ , \ U=\begin{pmatrix} 0 & -1 & -1 & 1 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \ ,$$

$$D=\begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 11 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$

Atunci

$$(D+L)^{-1}=\begin{pmatrix} 0.1 & 0 & 0 & 0 \\ -0.00909 & 0.09091 & 0 & 0 \\ 0.00758 & 0.00758 & 0.08333 & 0 \\ -0.00758 & -0.00758 & -0.00641 & 0.07692 \end{pmatrix} \ ,$$

$$-(D+L)^{-1}U=\begin{pmatrix} 0 & 0.1 & 0.1 & -0.1 \\ 0 & -0.00909 & 0.08182 & 0.1 \\ 0 & 0.00758 & 0.01515 & -0.08333 \\ 0 & -0.00758 & -0.01515 & 0.00641 \end{pmatrix} \ ,$$

$$(D+L)^{-1}b=\begin{pmatrix} 0.7 \\ -0.97273 \\ 1.81061 \\ -1.96445 \end{pmatrix} \ .$$

Alegând $x^{(0)}=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, obținem: $x^{(1)}=\begin{pmatrix} 0.7 \\ -0.97273 \\ 1.81061 \\ -1.96445 \end{pmatrix}$,

$$x^{(2)}=\begin{pmatrix} 0.98023 \\ -1.01219 \\ 1.99437 \\ -1.99711 \end{pmatrix} \ , \ x^{(3)}=\begin{pmatrix} 0.99793 \\ -1.00001 \\ 1.99958 \\ -1.99999 \end{pmatrix} \ , \ x^{(4)}=\begin{pmatrix} 0.99993 \\ -1.00001 \\ 1.99998 \\ -1.99999 \end{pmatrix} \ ,$$

$$x^{(5)}=\begin{pmatrix} 1 \\ -1 \\ 2 \\ -2 \end{pmatrix} \ .$$

Folosind metoda relaxării simple să se scrie soluția aproximativă pentru următoarele sisteme:

18. $\begin{cases} 8x_1 - x_2 + x_3 & = 3 \\ -x_1 + 6x_2 - x_3 & = 14 \\ x_1 - x_2 + 9x_3 & = -28 \end{cases}$

R. Matricea sistemului este $A = \begin{pmatrix} 8 & -1 & 1 \\ -1 & 6 & -1 \\ 1 & -1 & 9 \end{pmatrix}$.

Luând vectorul de probă $x^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ se obține $r^{(1)} = \begin{pmatrix} -3 \\ -14 \\ 28 \end{pmatrix}$, deci prima direcție de relaxare este $p^{(1)} = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Rezultă $x^{(2)} = \begin{pmatrix} 0 \\ 0 \\ -3.11111 \end{pmatrix}$.

Analog :

$$r^{(2)} = \begin{pmatrix} -6.11111 \\ -10.88889 \\ 0 \end{pmatrix}, p^{(2)} = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, x^{(3)} = \begin{pmatrix} 0 \\ 1.81481 \\ -3.11111 \end{pmatrix};$$

$$r^{(3)} = \begin{pmatrix} -7.92593 \\ 0 \\ -1.81481 \end{pmatrix}, p^{(3)} = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, x^{(3)} = \begin{pmatrix} 0.99074 \\ 1.81481 \\ -3.11111 \end{pmatrix};$$

$$r^{(4)} = \begin{pmatrix} 0 \\ -0.99074 \\ -0.824070 \end{pmatrix}, p^{(4)} = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, x^{(4)} = \begin{pmatrix} 0.99074 \\ 1.91994 \\ -3.11111 \end{pmatrix}.$$

ş. a. m. d.

19. $\begin{cases} 10x_1 + x_2 - x_3 - x_4 & = 9 \\ x_1 + 9x_2 - x_3 + x_4 & = -24 \\ -x_1 - x_2 + 11x_3 + x_4 & = 30 \\ -x_1 + x_2 + x_3 + 8x_4 & = -32 \end{cases}$

R. Matricea sistemului este $A = \begin{pmatrix} 10 & 1 & -1 & -1 \\ 1 & 9 & -1 & 1 \\ -1 & -1 & 11 & 1 \\ -1 & 1 & 1 & 8 \end{pmatrix}$.

Alegând $x^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ se obține $r^{(1)} = \begin{pmatrix} -9 \\ 24 \\ -30 \\ 32 \end{pmatrix}$.

Deci $p^{(1)} = e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $x^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -4 \end{pmatrix}$ $r^{(2)} = \begin{pmatrix} -5 \\ 20 \\ -34 \\ 0 \end{pmatrix}$.

Mai departe obținem :

$$p^{(2)} = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \; x^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 3.090901 \\ -4 \end{pmatrix} ; \; r^{(3)} = \begin{pmatrix} -8.09091 \\ 16.90909 \\ 0 \\ 3.09091 \end{pmatrix} , \; p^{(3)} = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} ,$$

$$x^{(4)} = \begin{pmatrix} 0 \\ -1.87879 \\ 3.09091 \\ -4 \end{pmatrix} ; \; p^{(4)} = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \; x^{(5)} = \begin{pmatrix} 0.99697 \\ -1.87879 \\ 3.09091 \\ -4 \end{pmatrix} ; \; r^{(5)} = \begin{pmatrix} 0 \\ 0.99697 \\ 0.88182 \\ 0.21515 \end{pmatrix}$$

$$, \; p^{(5)} = e_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad x^{(6)} = \begin{pmatrix} 0.99697 \\ -1.98956 \\ 3.09091 \\ -4 \end{pmatrix} .$$

20. $$\begin{cases} 4x_1 - 2x_2 + x_3 & = 9 \\ -2x_1 + 6x_2 - x_3 & = 5 \\ x_1 - x_2 + 5x_3 & = 6 \end{cases}$$

R. Matricea $A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 6 & -1 \\ 1 & -1 & 5 \end{pmatrix}$ este simetrică, tare diagonal dominantă: $|4| > |-2| + |1|$, $|6| > |-2| + |-1|$, $|5| > |1| + |-1|$, Δ_1=4 , Δ_2=20 , Δ_3=94 , deci pozitv definită şi se poate aplica metoda. Luând vectorul de probă

$$x^{(1)}=\begin{pmatrix}0\\0\\0\end{pmatrix}\quad \text{se obțin: } r^{(1)}=\begin{pmatrix}-9\\-5\\-6\end{pmatrix},\ p^{(1)}=e_1=\begin{pmatrix}1\\0\\0\end{pmatrix},\ x^{(2)}=\begin{pmatrix}2.25\\0\\0\end{pmatrix};$$

$$r^{(2)}=\begin{pmatrix}0\\-9.5\\-3.75\end{pmatrix},\ \text{este } p^{(2)}=e_2=\begin{pmatrix}0\\1\\0\end{pmatrix},\ x^{(3)}=\begin{pmatrix}2.25\\1.58333\\0\end{pmatrix};\ r^{(3)}=\begin{pmatrix}-3.16667\\0\\-5.33333\end{pmatrix}$$

$$,\ p^{(3)}=e_3=\begin{pmatrix}0\\0\\1\end{pmatrix},\ x^{(3)}=\begin{pmatrix}2.25\\1.5833\\1.06667\end{pmatrix};\ r^{(4)}=\begin{pmatrix}-2.1\\-1.06667\\0\end{pmatrix},\ p^{(4)}=e_1=\begin{pmatrix}1\\0\\0\end{pmatrix},$$

$$x^{(4)}=\begin{pmatrix}2.775\\1.58333\\1.06667\end{pmatrix}.$$

Să se scrie soluția aproximativă pentru următoarele sisteme de ecuații liniare obținută cu metoda suprarelaxării:

21. $$\begin{cases}2x_1-x_2 & =3\\ -x_1+3x_2-x_3 & =-5\\ \quad -x_2+4x_3 & =5\end{cases}$$

R. Matricea coeficienților sistemului este bloc tridiagonală, simetrică şi pozitiv definită, deci metoda suprarelaxării este convergentă. Sistemul se poate pune sub forma:

$$x^{(m+1)}=Mx^{(m)}+C,\quad m\geq 0,\ \text{unde: } M=-\left(E+\frac{1}{\omega}D\right)^{-1}\cdot\left[F+\left(1-\frac{1}{\omega}\right)D\right],$$

$$C=\left(E+\frac{1}{\omega}\right)^{-1}\cdot b,\ E=\begin{pmatrix}0&0&0\\-1&0&0\\0&-1&0\end{pmatrix},\ D=\begin{pmatrix}2&0&0\\0&3&0\\0&0&4\end{pmatrix},\ F=E^T,$$

$\omega=\dfrac{2}{1+\sqrt{1-\lambda_1^2}}$, λ_1 fiind cea mai mare valoare proprie a matricei $-D^{-1}(E+F)$, $\lambda_1=0.5$, $\omega=1.0718$ şi considerând vectorul inițial $x^{(0)}=b$, rezultă

$$x^{(1)}=\begin{pmatrix}-1.28719\\-0.10088\\0.95373\end{pmatrix},\quad x^{(2)}=\begin{pmatrix}1.64605\\-0.85027\\1.04344\end{pmatrix},\quad x^{(3)}=\begin{pmatrix}1.03385\\-0.98314\\1.0014\end{pmatrix},$$

$$x^{(4)}=\begin{pmatrix}1.00661\\-0.99835\\1.00034\end{pmatrix}.$$

22. $\begin{cases}14x_1 + x_2 = 13\\ x_1 + 5x_2 - x_3 = -3\\ -x_2 + 14x_3 - 2x_4 = -15\\ -2x_3 + 5x_4 = 7\end{cases}$

R. Matricea coeficienților sistemului este bloc tridiagonală, simetrică şi pozitiv definită, deci metoda suprarelaxării este convergentă. Sistemul se poate pune sub forma

$x^{(m+1)}=Mx^{(m)}+C,\quad m\geq 0,$ unde $M=-\left(E+\frac{1}{\omega}D\right)^{-1}\cdot\left[F+\left(1-\frac{1}{\omega}\right)D\right]$,

$$C=\left(E+\frac{1}{\omega}\right)^{-1}\cdot b,\quad E=\begin{pmatrix}0&0&0&0\\1&0&0&0\\0&-1&0&0\\0&0&-2&0\end{pmatrix},\quad D=\begin{pmatrix}14&0&0&0\\0&5&0&0\\0&0&14&0\\0&0&0&5\end{pmatrix},\quad F=E^T,$$

$\omega=\dfrac{2}{1+\sqrt{1-\lambda_1^2}}$, λ_1 fiind cea mai mare valoare proprie a matricei $-D^{-1}(E+F)$,

$\lambda_1=0.2735$, $\omega=1.01943$ şi considerând vectorul inițial
$x^{(0)}=b$, rezultă:

$$x^{(1)}=\begin{pmatrix}0.91242\\-3.79769\\-0.05784\\1.26758\end{pmatrix},\quad x^{(2)}=\begin{pmatrix}1.20542\\-0.79542\\-0.96444\\1.0093\end{pmatrix},\quad x^{(3)}=\begin{pmatrix}0.98111\\-0.99288\\-0.99882\\1.0003\end{pmatrix},$$

$$x^{(4)}=\begin{pmatrix}0.99985\\-0.99987\\-0.99997\\1.00001\end{pmatrix},\quad x^{(5)}=\begin{pmatrix}0.99999\\-0.99999\\-1\\1\end{pmatrix}.$$

Să se găsească soluția aproximativă obținută cu metoda Gauss–Seidel pentru sistemele de ecuații liniare următoare:

23. $$\begin{cases} 14x_1 + x_2 = 13 \\ x_1 + 5x_2 - x_3 = -3 \\ -x_2 + 14x_3 - 2x_4 = -15 \\ -2x_3 + 5x_4 = 7 \end{cases}$$

Să se precizeze numărul de iterații necesare pentru ca eroarea să se micşoreze de 10 ori.

R. Matricea A fiind simetrică şi pozitiv definită procedeul iterativ Gauss–Seidel este convergent. Sistemul se poate pune sub forma

$$x^{(m+1)}=Mx^{(m)}+C,\ m\geq 0,$$

unde

$$M = -(D+E)^{-1}F,\ C = (D+E)^{-1}b,\ E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \end{pmatrix},\ F = E^T,$$

$$D = \begin{pmatrix} 14 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Deci $M = \begin{pmatrix} 0 & -0.07143 & 0 & 0 \\ 0 & 0.01429 & 0.2 & 0 \\ 0 & 0.00102 & 0.01429 & 0.14286 \\ 0 & 0.0004 & 0.000571 & 0.05714 \end{pmatrix}$, $C = \begin{pmatrix} 0.92857 \\ -0.78571 \\ -1.12755 \\ 0.94898 \end{pmatrix}$.

Dacă se porneşte cu $v^{(0)}= b$ obțin vectorii:

$$v^{(1)} = \begin{pmatrix} 1.14286 \\ -3.82857 \\ -0.3449 \\ 1.26204 \end{pmatrix},\ v^{(2)} = \begin{pmatrix} 1.20204 \\ -0.90939 \\ -0.95609 \\ 1.01756 \end{pmatrix},\ v^{(3)} = \begin{pmatrix} 0.99353 \\ -0.98992 \\ -0.99677 \\ 1.00129 \end{pmatrix},$$

$$v^{(4)} = \begin{pmatrix} 0.99928 \\ -0.99921 \\ -0.99976 \\ 1.0001 \end{pmatrix},\ v^{(5)} = \begin{pmatrix} 0.99994 \\ -0.99994 \\ -0.99998 \\ 1.00001 \end{pmatrix}.$$

Raza spectrală este $\rho(M)=0.0748$, rata de convergență este $R(M)= -\lg(\rho(M))=1.12609$ şi numărul de iterații după care eroarea scade de 10 ori este $K\approx \dfrac{1}{R(M)}=0.88803$, adică la fiecare iterație eroarea scade de 10 ori.

24. $$\begin{cases} 5x_1 + x_2 - x_3 & = 5 \\ x_1 + 8x_2 + x_3 & = 10 \\ -x_1 + x_2 + 15x_3 & = 15 \end{cases}$$

Să se precizeze numărul de iterații necesare pentru ca eroarea să se micşoreze de 10 ori.

R. Matricea A fiind simetrică şi pozitiv definită procedeul iterativ Gauss–Seidel este convergent. Sistemul se poate pune sub forma

$$x^{(m+1)}=Mx^{(m)}+C\,,\ \ m\geq 0,$$

unde

$$M = -(D+E)^{-1}F\,,\ \ C = (D+E)^{-1}b\,,\ \ E = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix}\,,\ F = E^T,\ \ D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 15 \end{pmatrix}.$$

Deci $$M = \begin{pmatrix} 0 & -0.2 & 0.2 \\ 0 & 0.02 & -0.12 \\ 0 & -0.01467 & 0.02133 \end{pmatrix}\,,\ C = \begin{pmatrix} 1 \\ 0.9 \\ 1.0067 \end{pmatrix}.$$

Dacă se porneşte cu $v^{(0)} = b$ obțin vectorii:

$$v^{(1)} = \begin{pmatrix} 2 \\ -0.7 \\ 1.18 \end{pmatrix}\,,\ v^{(2)} = \begin{pmatrix} 1.376 \\ 0.7444 \\ 1.04211 \end{pmatrix}\,,\ v^{(3)} = \begin{pmatrix} 1.05954 \\ 0.78984 \\ 1.01798 \end{pmatrix}\,,\ v^{(4)} = \begin{pmatrix} 1.04563 \\ 0.79364 \\ 1.0168 \end{pmatrix}\,,$$

$$v^{(5)} = \begin{pmatrix} 1.04463 \\ 0.79386 \\ 1.01672 \end{pmatrix}.$$

Raza spectrală este $\rho(M)$=0.06262, rata de convergență este $R(M)= -\lg(\rho(M))$=1.20326 şi numărul de iterații după care eroarea scade de 10 ori este $K \approx \dfrac{1}{R(M)}$=0.83108, adică la fiecare iterație eroarea scade de 10 ori.

Să se scrie soluția aproximativă obținută cu metoda gradienților conjugați aplicată sistemelor de ecuații liniare:

25. $$\begin{cases} 6x_1 & + x_2 & - x_3 & = 6 \\ x_1 & + 8x_2 & + x_3 & = -8 \\ -x_1 & + x_2 & + 10x_3 & = -12 \end{cases}.$$

R. Matricea coeficienților sistemului este simetrică şi pozitiv definită şi deci se poate aplica metoda. Luând $x^{(0)}=\begin{pmatrix}0\\0\\0\end{pmatrix}$ ca vector de probă, se obțin rezultatele.

Iterația I.

$r^{(0)}=Ax_0-b=\begin{pmatrix}-6\\8\\12\end{pmatrix}$, prima direcție de relaxare, $p^{(1)}=-r^{(0)}$,

$$q_1=-\frac{\langle r^{(0)},p^{(1)}\rangle}{\langle Ap^{(1)},p^{(1)}\rangle}=0.10133,\quad x^{(1)}=x^{(0)}+q_1p^{(1)}=\begin{pmatrix}0.60797\\-0.81063\\-1.21595\end{pmatrix}.$$

Iterația a II-a .

$$r^{(1)}=r^{(0)}+q_1Ap^{(1)}=\begin{pmatrix}-1.94684\\0.90698\\-1.57807\end{pmatrix},\quad c_1=\frac{\langle r^{(1)},r^{(1)}\rangle}{\langle r^{(0)},r^{(0)}\rangle}=0.02911$$

$$p^{(2)}=-r^{(1)}+c_1p^{(1)}=\begin{pmatrix}2.12151\\-1.13987\\1.22874\end{pmatrix},\quad q_2=-\frac{\langle r^{(1)},p^{(2)}\rangle}{\langle Ap^{(2)},p^{(2)}\rangle}=0.17916,$$

$$x^{(2)}=x^{(1)}+q_2p^{(2)}=\begin{pmatrix}0.98807\\-1.01485\\-0.9958\end{pmatrix}.$$

Iterația a III-a.

$$r^{(2)}=r^{(1)}+q_2Ap^{(2)}=\begin{pmatrix}-0.09062\\-0.12656\\0.03906\end{pmatrix},\quad c_2=\frac{\langle r^{(2)},r^{(2)}\rangle}{\langle r^{(1)},r^{(1)}\rangle}=3.62608\cdot 10^{-3}$$

$$p^{(3)}=-r^{(2)}+c_2p^{(2)}=\begin{pmatrix}0.09832\\0.12243\\-0.03461\end{pmatrix},\quad q_3=-\frac{\langle r^{(2)},p^{(3)}\rangle}{\langle Ap^{(3)},p^{(3)}\rangle}=0.12133,$$

$$x^{(3)}=x^{(2)}+q_3p^{(3)}=\begin{pmatrix}1\\-1\\-1\end{pmatrix}.$$

26. Să se determine traseul optim pentru o conductă de gaze naturale care să treacă prin "apropierea" localităților L_i , $i=1,...,10$, care raportate la un sistem cartezian de referință au coordonatele următoare:

$L_1(1,2)$, $L_2(2,2)$, $L_3(5,3)$, $L_4(7,4)$, $L_5(10,2)$, $L_6(11,3)$, $L_7(15,4)$, $L_8(16,5)$, $L_9(18,1)$, $L_{10}(20,4)$.

R. Luând traseul după o dreaptă, se obține sistemul:

$$\begin{cases} a+b &= 2 \\ 2a+b &= 2 \\ 5a+b &= 3 \\ 7a+b &= 4 \\ 10a+b &= 2 \\ 11a+b &= 3 \\ 15a+b &= 4 \\ 16a+b &= 5 \\ 18a+b &= 1 \\ 20a+b &= 4 \end{cases}$$

care este supradimensionat. Se formează sistemul normal al lui Gauss $Bu=c$, unde $B=A^TA$, $c=A^Tb$, adică: $B=\begin{pmatrix} 1505 & 105 \\ 105 & 10 \end{pmatrix}$, $c=\begin{pmatrix} 340 \\ 30 \end{pmatrix}$, a cărui soluție este $u_1=0.06211$, $u_2=2.34783$. Raportat la acel sistem de coordonate traseul conductei trebuie să urmeze dreapta $y = 0.06211x+2.34783$.