Uvod u numeričku matematiku

M. Klaričić Bakula

Ožujak, 2009.

1 Uvod

Jedan od osnovnih problema numeričke matematike je rješavanje linearnih sustava jednadžbi. U ovom poglavlju istraživat ćemo metode za rješavane kvadratnih sustava linearnih jednadžbi, tj. sustava s n jednadžbi i n nepoznanica,

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2j}x_j + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n &= b_i \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nj}x_j + \cdots + a_{nn}x_n &= b_n.
\end{aligned}$$

Matrica

$$A = [a_{ij}]_{i,j=1}^{n} \in \mathbb{R}^{n \times n}$$

je **matrica sustava**, a njeni elementi su **koeficijenti sustava**.

Vektor

$$\boldsymbol{b} = [b_i]_{i=1}^{n} \in \mathbb{R}^{n}$$

je vektor desne strane sustava ili **vektor slobodnih koeficijenata**.

Treba odrediti **vektor nepoznanica**

$$\boldsymbol{x} = [x]_{i=1}^{n} \in \mathbb{R}^{n}$$

tako da vrijedi

$$A\boldsymbol{x} = \boldsymbol{b}.$$

Naravno, što se teorije tiče, rješavanje ovakvog sustava je gotovo trivijalan problem, posebno ako je matrica sustava A regularna, u kojem slučaju je $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. Pri tom postoje eksplicitne formule za elemente matrice A^{-1}, kao i za samo rješenje $\boldsymbol{x}$. Osim toga, svima poznata Gaussova metoda eliminacije daje rješenje u $O\left(n^3\right)$ elementranih operacija, pa imamo i algoritam koji računa rješenje koristeći samo jednostavne aritmetičke operacije.

U primijenjenoj matematici, posebno u numeričkoj linearnoj algebri, situacija je puno složenija. U numeričkoj matematici riješiti neki problem danas znači *biti u stanju u konkretnoj situaciji s konkretnim podacima (koristeći računalo) brzo doći do dovoljno točne numeričke aproksimacije rješenja*. Npr., ako su matrica A i vektor $\boldsymbol{b}$ zapisani u nekim datotekama na disku, ili su dane neke procedure koje ih generiraju, onda je zadatak izračunati numeričke vrijednosti komponenata vektora $\boldsymbol{x}$. Današnja računala su vrlo brza, no osnovna značajka moderne numeričke matematike su problemi sve većih dimenzija.

Npr. ako je sustav veličine $n = 10^5$, onda Gaussova metoda traži broj operacija veličine 10^{15}, pa brzinom rada od oko 10^9 operacija u sekundi (što je standard za jednoprocesorsko računalo) dođemo do vremena izvršavanja od 10^6 sekunda, što je više od 10 dana.

Iz ovoga vidimo da problem koji je u matematičkoj praksi posve jednostavan može u stvarnoj praksi biti iznimno izazovan. *Posebno treba uzeti u obzir i razne greške koje se generiraju prikazom realnih brojeva i izvođenjem računskih operacija u računalu.*

2 Kako u praksi nastaje sustav linearnih jednadžbi

U praksi se rješavanje raznih problema svodi na rješavanje sustava linearnih jednadžbi. Pogledajmo ovaj primjer s kojim se često susrećemo.

PRIMJER. Zadani su parovi točaka $(x_i, y_i)\,, i = 0, 1, \ldots, n$, gdje su $y_i = f\,(x_i)$ izmjerene vrijednosti funkcije f koju želimo aproksimirati polinomom p stupnja n. Pretpostavimo da su svi dani čvorovi x_i različiti. Kriterij za odabir polinoma je da u čvorovima ima iste vrijednosti kao funkcija f (tj. da vrijedi $p\,(x_i) = f\,(x_i) = y_i$), a upravo zato i govorimo o **interpolacijskom polinomu**. Ako p prikažemo u **kanonskom obliku**

$$p\,(x) = a_0 + a_1 x + \cdots + a_n x^n = \sum_{j=0}^{n} a_j x^j,$$

onda za naći polimom p treba odrediti koeficijente $a_0, a_1, \ldots, a_n$.

Dakle, treba riješiti sustav

$$
\begin{aligned}
a_0 + a_1 x_0 + \cdots + a_j x_0^j + \cdots + a_n x_0^n &= y_0 \\
a_0 + a_1 x_1 + \cdots + a_j x_1^j + \cdots + a_n x_1 &= y_1 \\
&\vdots \\
a_0 + a_1 x_i + \cdots + a_j x_i^j + \cdots + a_n x_i^n &= y_i \\
&\vdots \\
a_0 + a_1 x_n + \cdots + a_j x_n^j + \cdots + a_n x_n^n &= y_n
\end{aligned}
$$

koji se matrično može zapisati kao

$$V\boldsymbol{a} = \boldsymbol{y},$$

pri čemu je

$$V = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_i & x_i^2 & \cdots & x_i^{n-1} & x_i^n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} & x_n^n \end{bmatrix}, \quad \boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}.$$

Matrica V zove se **Vandermondeova matrica**. Zahvaljujući njenom posebnom obliku moguće je odrediti $\boldsymbol{a} = V^{-1}\boldsymbol{y}$.

3 Gaussove eliminacije i trokutaste faktorizacije

Metoda Gaussovih eliminacija je svakako najstariji, najjednostavniji i najpoznatiji algoritam za rješavanje sustava linearnih jednadžbi. Ideja je jednostavna: da bismo riješili npr. sustav

$$\begin{aligned} 2x_1 - x_2 &= 1 \\ -x_1 + 2x_2 &= 1 \end{aligned}$$

dovoljno je primijetiti da zbog prve jednadžbe vrijedi $x_1 = (1 + x_2)/2$, pa je drugu jednadžbu moguće pisati kao

$$-\frac{1}{2}(1 + x_2) + 2x_2 = 1,$$

iz čega lako slijedi $x_2 = 1$ i $x_1 = 1$. Kažemo da smo nepoznanicu x_1 **eliminirali** iz druge jednadžbe. Ovu ideju lako možemo proširiti i na opće n-dimenzionalne sustave, $n > 1$, tako da sustavno eliminiramo neke nepoznanice iz nekih jednadžbi. Pokazuje se da takav algoritam ima zanimljivu strukturu i da ga se može lako zapisati u terminima matrica. Kvalitativno novi moment nastaje kada kada se sam proces eliminacija interpretira kao faktorizacija matrice sustava A na umnožak trokutastih matrica.

3.1 Matrični zapis metode eliminacija

Pogledajmo jedan primjer.
Pretpostavimo da želimo riješiti sustav jednadžbi

$$\begin{aligned} 5x_1 + x_2 + 4x_3 &= 19 \\ 10x_1 + 4x_2 + 7x_3 &= 39 \\ -15x_1 + 5x_2 - 9x_3 &= -32. \end{aligned}$$

Očito je da ako želimo eliminirati npr. x_1 iz druge jednadžbe, onda moramo prvu pomnožiti s -2 (uočimo da je $-2 = -10/5 = -a_{21}/a_{11}$) i pribrojiti drugoj. Ovo možemo matrično zapisati kao

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 1 & 4 \\ 10 & 4 & 7 \\ -15 & 5 & -9 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 4 \\ 0 & 2 & -1 \\ -15 & 5 & -9 \end{bmatrix},$$

odnosno

$$L^{(2,1)} A = A^{(1)}.$$

Analogno, nepoznanicu x_1 možemo eliminirati iz treće jednadžbe ako prvu pomnožimo s 3 (uočimo da je $3 = 15/5 = a_{21}/a_{11}$) i pribrojimo trećoj. U matričnom zapisu je to dano s

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 1 & 4 \\ 0 & 2 & -1 \\ -15 & 5 & -9 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 4 \\ 0 & 2 & -1 \\ 0 & 8 & 3 \end{bmatrix},$$

odnosno

$$L^{(3,1)} A^{(1)} = A^{(2)}.$$

Pri tom se mijenja i vektor s desne strane sustava

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 19 \\ 39 \\ -32 \end{bmatrix} = \begin{bmatrix} 19 \\ 1 \\ -32 \end{bmatrix} = \boldsymbol{b}^{(1)}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 19 \\ 1 \\ -32 \end{bmatrix} = \begin{bmatrix} 19 \\ 1 \\ 25 \end{bmatrix} = \boldsymbol{b}^{(2)}.$$

Uočimo da su matrice $L^{(2,1)}$ i $L^{(3,1)}$ donjetrokutaste. Nakon provedenih transformacija polazni sustav se svodi na ekvivalentni sustav (sustav s istim rješenjem)

$$\begin{aligned} 5x_1 + x_2 + 4x_3 &= 19 \\ 2x_2 - x_3 &= 1 \\ 8x_2 - 3x_3 &= 25. \end{aligned}$$

Analogno, želimo li eliminirati x_2 iz treće jednadžbe sustava, moramo drugu pomnožiti s -4 i pribrojiti trećoj. Matrično to zapisujemo kao

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 5 & 1 & 4 \\ 0 & 2 & -1 \\ 0 & 8 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 4 \\ 0 & 2 & -1 \\ 0 & 0 & 7 \end{bmatrix},$$

odnosno

$$L^{(3,2)} A^{(2)} = A^{(3)}.$$

Napravimo li i zadnji korak množenja vektora desne strane sustava dobijemo

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 19 \\ 1 \\ 25 \end{bmatrix} = \begin{bmatrix} 19 \\ 1 \\ 21 \end{bmatrix} = \boldsymbol{b}^{(3)},$$

pa nam ekvivalentni sustav u matričnom zapisu glasi

$$\begin{bmatrix} 5 & 1 & 4 \\ 0 & 2 & -1 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 19 \\ 1 \\ 21 \end{bmatrix}.$$

Odavde lako slijedi rješenje sustava

$$\begin{aligned} x_3 &= 21/7 = 3 \\ x_2 &= (1 + x_3)/2 = 2 \\ x_1 &= (19 - x_2 - 4x_3)/5 = 1. \end{aligned}$$

No lako se provjerava da vrijedi

$$\left(L^{(2,1)}\right)^{-1}\left(L^{(3,1)}\right)^{-1}\left(L^{(3,2)}\right)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 4 & 1 \end{bmatrix} = L,$$

pa je stoga zbog

$$L^{(3,2)}L^{(3,1)}L^{(2,1)}A = A^{(3)}$$

ispunjeno

$$A = \left(L^{(2,1)}\right)^{-1}\left(L^{(3,1)}\right)^{-1}\left(L^{(3,2)}\right)^{-1} A^{(3)} = LA^{(3)}.$$

Dakle, matricu A smo prikazali kao umnožak jedne donjetrokutaste (L) i jedne gornjetrokutaste ($A^{(3)}$) matrice. U ovakvom kontekstu gornjetrokutastu matricu $A^{(3)}$ označavamo s U, pa pišemo

$$A = LU.$$

Upravo stoga govorimo o LU faktorizaciji. Uočimo da je računaje inverza matrica $L^{(i,j)}$ jednostavno (samo promjenimo predznak netrivijalnim elementima u donjem trokutu), a cijeli umnožak se dobije stavljanjem tih elemenata na odgovarajuće pozicije u donjem trokutu matrice L. Vratimo li se na početni sustav imamo

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} = (LU)^{-1}\,\boldsymbol{b} = U^{-1}L^{-1}\boldsymbol{b}.$$

U terminima matrica A i $\boldsymbol{b}$ sustav iz ovog primjera riješen je metodom koja se sastoji od tri glavna koraka:

1. Matricu A faktoriziramo u oblik $A = LU$.
2. Rješavanjem donjetrokutastog sustava $L\boldsymbol{y} = \boldsymbol{b}$ treba odrediti vektor $\boldsymbol{y} = L^{-1}\boldsymbol{b}$.
3. Rješavanjem gornjetrokutastog sustava $U\boldsymbol{x} = \boldsymbol{y}$ treba odrediti vektor $\boldsymbol{x} = U^{-1}\boldsymbol{y} = U^{-1}\left(L^{-1}\boldsymbol{b}\right)$.

3.2 Rješavanje trokutastih sustava eliminacijama unaprijed i unazad

Trokutasti sustavi jednadžbi su najlakši za rješavanje. Pogledajmo jedan primjer dimenzije $n = 4$. Neka je

$$L\boldsymbol{x} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \boldsymbol{b},$$

te pretpostavimo da je matrica L regularna. Lako se provjeri da to znači $l_{ii} \neq 0$, $i = 1, 2, 3, 4$. U tom slučaju je

$$\begin{aligned} x_1 &= b_1/l_{11} \\ x_2 &= (b_2 - l_{21}x_1)\,/l_{22} \\ x_3 &= (b_3 - l_{31}x_1 - l_{32}x_2)\,/l_{33} \\ x_4 &= (b_4 - l_{41}x_1 - l_{42}x_2 - l_{43}x_3)\,/l_{44}. \end{aligned}$$

ALGORITAM za rješavanje linearnog sustava jednadžbi $L\boldsymbol{x} = \boldsymbol{b}$ s regularnom donjetrokutastom matricom $L \in \mathbb{R}^{n\times n}$.

$$x_1 = \frac{b_1}{l_{11}};$$
$$\text{for } i = 2,\ldots,n \ \left\{ x_i = \frac{1}{l_{ii}} \left(b_i - \sum_{j=1}^{i-1} l_{ij} x_j \right) \right\}.$$

Prebrojimo li operacije u gornjem algoritmu vidimo da imamo n dijeljenja, $1+2+\cdots+n = n\left(n+1\right)/2$ množenja i isto toliko zbrajanja (i oduzimanja). Dakle, ukupna složenost je $O\left(n^2\right)$, što je bolje od predviđene za n-dimenzionalni sustav linearnih jednadžbi.

ALGORITAM za rješavanje linearnog sustava jednadžbi $U\boldsymbol{x} = \boldsymbol{b}$ s regularnom gornjetrokutastom matricom $U \in \mathbb{R}^{n\times n}$.

$$\begin{aligned} x_n &= \frac{b_n}{u_{nn}}; \\ \text{for } i &= n-1,\ldots,1 \ \left\{ x_i = \frac{1}{u_{ii}} \left(b_i - \sum_{j=i+1}^{n} u_{ij} x_j \right) \right\}. \end{aligned}$$

I u ovom slučaju je ukupna složenost $O\left(n^2\right)$, što je bolje od predviđene za n-dimenzionalni sustav linearnih jednadžbi.

3.3 LU (LR) faktorizacija

Sada nam ostaje proučiti kako izvesti na najbolji mogući način faktorizaciju matrice $A \in \mathbb{R}^{n \times n}$ na umnožak donjetrokutaste i gornjetrokutaste matrice. Problem želimo riješiti za proizvoljan n, no radi jednostavnosti ćemo razmotriti slučaj $n = 5$. Neka je

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix}.$$

Sjetimo se da se eliminacija prve nepoznanice iz svih jednadžbi osim prve manifestira poništavanjem svih elemenata matrice A u prvom stupcu osim onog na dijagonali. To možemo napraviti u jednom potezu definiranjem matrice $L^{(1)}$ na način kao u prethodnom primjeru, dakle

$$L^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -a_{21}/a_{11} & 1 & 0 & 0 & 0 \\ -a_{31}/a_{11} & 0 & 1 & 0 & 0 \\ -a_{41}/a_{11} & 0 & 0 & 1 & 0 \\ -a_{51}/a_{11} & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Nakon množenja s A dobijemo

$$A^{(1)} = L^{(1)}A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} & a_{25}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} & a_{35}^{(1)} \\ 0 & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} & a_{45}^{(1)} \\ 0 & a_{52}^{(1)} & a_{53}^{(1)} & a_{54}^{(1)} & a_{55}^{(1)} \end{bmatrix}.$$

Primijetimo da je transformaciju $A \longmapsto A^{(1)}$ moguće izvesti samo ako je $a_{11} \neq 0$.

Lako se provjeri da je

$$\left(L^{(1)}\right)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ a_{21}/a_{11} & 1 & 0 & 0 & 0 \\ a_{31}/a_{11} & 0 & 1 & 0 & 0 \\ a_{41}/a_{11} & 0 & 0 & 1 & 0 \\ a_{51}/a_{11} & 0 & 0 & 0 & 1 \end{bmatrix},$$

te da iz $A = \left(L^{(1)}\right)^{-1} A^{(1)}$ slijedi

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_{21}/a_{11} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22}^{(1)} \end{bmatrix}.$$

Drugim riječima, u prvom koraku dobili smo faktorizaciju vodeće 2×2 podmatrice od A. Uvjet za izvod ove faktorizacije bio je $a_{11} \neq 0$.

Uočimo i da je

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \det \begin{bmatrix} 1 & 0 \\ a_{21}/a_{11} & 1 \end{bmatrix} \det \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22}^{(1)} \end{bmatrix} = 1 \cdot a_{11} a_{22}^{(1)}.$$

Označimo sada

$$\alpha_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} a_{22}^{(1)}.$$

Pretpostavimo li da je $\alpha_2 \neq 0$, onda je i $a_{22}^{(1)} \neq 0$, pa su dobro definirane matrice

$$L^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -a_{32}^{(1)}/a_{22}^{(1)} & 1 & 0 & 0 \\ 0 & -a_{42}^{(1)}/a_{22}^{(1)} & 0 & 1 & 0 \\ 0 & -a_{52}^{(1)}/a_{22}^{(1)} & 0 & 0 & 1 \end{bmatrix}, \quad \left(L^{(2)}\right)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & a_{32}^{(1)}/a_{22}^{(1)} & 1 & 0 & 0 \\ 0 & a_{42}^{(1)}/a_{22}^{(1)} & 0 & 1 & 0 \\ 0 & a_{52}^{(1)}/a_{22}^{(1)} & 0 & 0 & 1 \end{bmatrix}.$$

Vrijedi

$$A^{(2)} = L^{(2)}A^{(1)} = L^{(2)}L^{(1)}A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} & a_{25}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & a_{34}^{(2)} & a_{35}^{(2)} \\ 0 & 0 & a_{43}^{(2)} & a_{44}^{(2)} & a_{45}^{(2)} \\ 0 & 0 & a_{53}^{(2)} & a_{54}^{(2)} & a_{55}^{(2)} \end{bmatrix},$$

i također

$$\begin{aligned} A &= \left(L^{(1)}\right)^{-1}\left(L^{(2)}\right)^{-1}A^{(2)} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ a_{21}/a_{11} & 1 & 0 & 0 & 0 \\ a_{31}/a_{11} & a_{32}^{(1)}/a_{22}^{(1)} & 1 & 0 & 0 \\ a_{41}/a_{11} & a_{42}^{(1)}/a_{22}^{(1)} & 0 & 1 & 0 \\ a_{51}/a_{11} & a_{52}^{(1)}/a_{22}^{(1)} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} & a_{25}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & a_{34}^{(2)} & a_{35}^{(2)} \\ 0 & 0 & a_{43}^{(2)} & a_{44}^{(2)} & a_{45}^{(2)} \\ 0 & 0 & a_{53}^{(2)} & a_{54}^{(2)} & a_{55}^{(2)} \end{bmatrix}. \end{aligned}$$

Iz ovoga se vidi da vrijedi

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ a_{21}/a_{11} & 1 & 0 \\ a_{31}/a_{11} & a_{32}^{(1)}/a_{22}^{(1)} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix},$$

pa ako vrijedi $a_{11} \neq 0$ i $\alpha_2 \neq 0$ dobijamo trokutastu faktorizaciju vodeće 3×3 podmatrice od A. Sada stavimo

$$\alpha_3 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = 1 \cdot \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix} = a_{11} a_{22}^{(1)} a_{33}^{(2)}.$$

Očito postupak sada možemo nastaviti ako je ispunjen uvjet $\alpha_3 \neq 0$ (odnosno $a_{33}^{(2)} \neq 0$).

Nakon računanja $L^{(3)}$ i $L^{(4)}$ dobijemo

$$A^{(4)} = L^{(4)} A^{(3)} = L^{(4)} L^{(3)} L^{(2)} L^{(1)} A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} & a_{25}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & a_{34}^{(2)} & a_{35}^{(2)} \\ 0 & 0 & 0 & a_{44}^{(3)} & a_{45}^{(3)} \\ 0 & 0 & 0 & 0 & a_{55}^{(4)} \end{bmatrix},$$

odnosno

$$\begin{aligned} A &= LU \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ a_{21}/a_{11} & 1 & 0 & 0 & 0 \\ a_{31}/a_{11} & a_{32}^{(1)}/a_{22}^{(1)} & 1 & 0 & 0 \\ a_{41}/a_{11} & a_{42}^{(1)}/a_{22}^{(1)} & a_{43}^{(2)}/a_{33}^{(2)} & 1 & 0 \\ a_{51}/a_{11} & a_{52}^{(1)}/a_{22}^{(1)} & a_{53}^{(2)}/a_{33}^{(2)} & a_{54}^{(3)}/a_{44}^{(3)} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} & a_{25}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & a_{34}^{(2)} & a_{35}^{(2)} \\ 0 & 0 & 0 & a_{44}^{(3)} & a_{45}^{(3)} \\ 0 & 0 & 0 & 0 & a_{55}^{(4)} \end{bmatrix}. \end{aligned}$$

Pri tom je

$$L = \left(L^{(1)}\right)^{-1} \left(L^{(2)}\right)^{-1} \left(L^{(3)}\right)^{-1} \left(L^{(4)}\right)^{-1},$$

a izvedivost postupka koji je doveo do faktorizacije $A = LU$ ovisila je o uvjetima

$$a_{11} \neq 0,\ a_{22}^{(1)} \neq 0,\ a_{33}^{(2)} \neq 0,\ a_{44}^{(3)} \neq 0.$$

Također, uočili smo da su ti uvjeti osigurani ako su determinante glavnih podmatrica matrice A različite od nule. Kod nas je to značilo uvjete

$$\alpha_1 = a_{11} \neq 0,\ \alpha_2 \neq 0,\ \alpha_3 \neq 0,\ \alpha_4 \neq 0.$$

Brojeve $a_{11},\ a_{22}^{(1)},\ a_{33}^{(2)},\ a_{44}^{(3)}$ nazivamo **pivotnim elementima** ili **pivotima**, a brojeve α_1, $\alpha_2,\ \alpha_3,\ \alpha_4$ nazivamo **glavnim minorama** matrice A.

3.4 Zaključak i algoritam

Ako je prvih $n-1$ minora matrice A različito od nule, onda su i svi pivotni elementi različiti od nule i Gaussove eliminacije daju LU faktorizaciju matrice A.

ALGORITAM $\left(U = A^{(n-1)} = \left[a_{ij}^{(n-1)}\right].\right)$

$L = I;$

for $k := 1$ **to** $n-1$

  begin

  for $j := k+1$ **to** n

    begin

    $l_{jk} = a_{jk}^{(k-1)} / a_{kk}^{(k-1)};$

    $a_{jk}^{(k)} = 0;$

    end

```
for j := k+1 to n
    begin
    for i := k+1 to n
        begin
        a_ij^(k) = a_ij^(k-1) - l_ik a_kj^(k-1);
        end
    end
end
```

Na osnovu prethodnog lako se vidi da vrijedi sljedeći teorem.

TEOREM. Neka je $A \in \mathbb{R}^{n\times n}$ i neka su determinante glavnih podmatrica $A\,(1:k, 1:k)$ različite od nule za $k = 1, 2, \ldots, n-1$. Tada postoji donjetrokutasta matrica L s jedinicama na dijagonali i gornjetrokutasta matrica U tako da vrijedi $A = LU$. Ako faktorizacija $A = LU$ postoji i ako je matrica A regularna, onda je ova faktorizacija jedinstvena. Tada je i

$$\det A = \prod_{i=1}^{n} u_{ii}.$$

DOKAZ. Dokaz egzistencije preskačemo jer je iz primjera jasno da se induktivno može dokazati za svaki $n \geq 2$. Dokažimo jedinstvenost pod navedenim uvjetima.

Pretpostavimo da postoje dvije takve faktorizacije

$$A = LU = L'U'.$$

Ako je A regularna, onda su regularne i matrice L, U, L', U', pa vrijedi

$$L^{-1}L' = U\left(U'\right)^{-1}.$$

U gornjoj jednakosti imamo jednakost donjetrokutaste i gornjetrokutaste matrice što znači da su obje dijagonalne. Kako i L i L' na dijagonali imaju jedinice, to isto ima i $L^{-1}L'$, pa je očigledno

$$L^{-1}L' = I,$$

to jest $L = L'$. Tada je i $U = U'$ ■

NAPOMENA. Primijetimo sljedeće: ako je A regularna i ako ima LU faktorizaciju, onda su nužno regularne i sve glavne podmatrice $A(1:k,1:k)$ za $k = 1, 2, \ldots, n-1$. To, naime, slijedi iz činjenice

$$\det A(1:k,1:k) = \prod_{i=1}^{k} u_{ii}, \quad k = 1, 2, \ldots, n.$$

3.5 LU faktorizacija s pivotiranjem

U prethodnom smo vidjeli da mogućnost provođenja LU faktorizacije direktno ovisi o tome da li su determinante glavnih podmatrica različite od nule. Problem koji se inače javlja ilustriran je u narednom primjeru.

PRIMJER. Neka je matrica A sustava $A\boldsymbol{x} = \boldsymbol{b}$ dana s

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Ova matrica je regularna i $\det A = -1$, pa dani sustav ima rješenje, no ne i LU faktorizaciju jer pretpostavka

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$

povlaći

$$\begin{aligned} 1 \cdot u_{11} &= 0 \\ 1 \cdot u_{12} &= 1 \\ l_{21} \cdot u_{11} &= 1 \\ l_{21} \cdot u_{12} + u_{22} &= 1. \end{aligned}$$

Očito iz ovoga slijedi $l_{21} \cdot 0 = 1$, što je nemoguće.

Lako se vidi da bi ovaj problem zamjenom prve i druge jednadžbe sustava bio uklonjen, jer matrica

$$A' = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

ima jednostavnu LU faktorizaciju $L = I,\ U = A'$. Veza između A i A' dana je matrično s

$$A' = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = PA.$$

Matricu P nazivamo **matricom permutacije** ili jednostavno **permutacijom**. Njeno djelovanje na matricu A je permutiranje redaka.

PRIMJER. Neka je matrica A sustava $A\boldsymbol{x} = \boldsymbol{b}$ dana s

$$A = \begin{bmatrix} 1 & 1 & 4 & 1 \\ 2 & 1 & 1 & 6 \\ 5 & 1 & 1 & 0 \\ 1 & 4 & 1 & 3 \end{bmatrix}.$$

Najveći element u prvom stupcu je na mjestu $(3,1)$, što znači da prvi pivot maksimiziramo ako zamijenimo prvi i treći redak matrice A. Tu zamjenu realizira permutacija $P^{(1)}$, gdje je

$$P^{(1)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad P^{(1)}A = \begin{bmatrix} 5 & 1 & 1 & 0 \\ 2 & 1 & 1 & 6 \\ 1 & 1 & 4 & 1 \\ 1 & 4 & 1 & 3 \end{bmatrix}.$$

Sada definiramo

$$L^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2/5 & 1 & 0 & 0 \\ -1/5 & 0 & 1 & 0 \\ -1/5 & 0 & 0 & 1 \end{bmatrix},$$

pa je

$$A^{(1)} = L^{(1)} P^{(1)} A = \begin{bmatrix} 5 & 1 & 1 & 0 \\ 0 & 3/5 & 3/5 & 6 \\ 0 & 4/5 & 19/5 & 1 \\ 0 & 19/5 & 4/5 & 3 \end{bmatrix}.$$

Naredni pivot je maksimiziran permutacijom $P^{(2)}$ gdje je

$$P^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad P^{(2)}A^{(1)} = \begin{bmatrix} 5 & 1 & 1 & 0 \\ 0 & 19/5 & 4/5 & 3 \\ 0 & 4/5 & 19/5 & 1 \\ 0 & 3/5 & 3/5 & 6 \end{bmatrix}.$$

Sljedeći korak eliminacije glasi

$$L^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -4/19 & 1 & 0 \\ 0 & -3/19 & 0 & 1 \end{bmatrix}, \; A^{(2)} = L^{(2)}P^{(2)}A^{(1)} = \begin{bmatrix} 5 & 1 & 1 & 0 \\ 0 & 19/5 & 4/5 & 3 \\ 0 & 0 & 69/19 & 7/19 \\ 0 & 0 & 9/19 & 105/19 \end{bmatrix}$$

Nastavimo li dalje s permutiranjem redaka dobijemo da je $P^{(3)} = I$, pa je

$$L^{(3)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -9/69 & 1 \end{bmatrix}, \ A^{(3)} = L^{(3)} I A^{(2)} = \begin{bmatrix} 5 & 1 & 1 & 0 \\ 0 & 19/5 & 4/5 & 3 \\ 0 & 0 & 69/19 & 7/19 \\ 0 & 0 & 0 & 7182/1311 \end{bmatrix}.$$

Sada primijetimo da je

$$A^{(3)} = L^{(3)} I L^{(2)} P^{(2)} L^{(1)} P^{(1)} A,$$

gdje je

$$P^{(2)} L^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/5 & 0 & 0 & 1 \\ -1/5 & 0 & 1 & 0 \\ -2/5 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/5 & 1 & 0 & 0 \\ -1/5 & 0 & 1 & 0 \\ -2/5 & 0 & 0 & 1 \end{bmatrix} P^{(2)} = \widetilde{L}^{(1)} P^{(2)}.$$

Dakle,

$$U = A^{(3)} = L^{(3)} L^{(2)} \widetilde{L}^{(1)} P^{(2)} P^{(1)} A.$$

Ako stavimo $P = P^{(2)} P^{(1)}$, onda vrijedi

$$\begin{aligned} PA &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 4 & 1 \\ 2 & 1 & 1 & 6 \\ 5 & 1 & 1 & 0 \\ 1 & 4 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 1 & 0 \\ 1 & 4 & 1 & 3 \\ 1 & 1 & 4 & 1 \\ 2 & 1 & 1 & 6 \end{bmatrix} \\ &= \left(\widetilde{L}^{(1)}\right)^{-1} \left(L^{(2)}\right)^{-1} \left(L^{(3)}\right)^{-1} U \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/5 & 1 & 0 & 0 \\ 1/5 & 0 & 1 & 0 \\ 1/5 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4/19 & 1 & 0 \\ 0 & 3/19 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 9/69 & 1 \end{bmatrix} U. \end{aligned}$$

Dakle imamo,

$$PA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/5 & 1 & 0 & 0 \\ 1/5 & 4/19 & 1 & 0 \\ 1/5 & 3/19 & 9/69 & 1 \end{bmatrix} \begin{bmatrix} 5 & 1 & 1 & 0 \\ 0 & 19/5 & 4/5 & 3 \\ 0 & 0 & 69/19 & 7/19 \\ 0 & 0 & 0 & 7182/1311 \end{bmatrix}.$$

Možemo zaključiti sljedeće:

Za proizvoljnu $n \times n$ matricu A postoji permutacija P takva da Gaussove eliminacije daju LU faktorizaciju matrice PA. Pri tome je L donjetrokutasta matrica s jedinicama na dijagonali, a U je gornjetrokutasta matrica. Permutaciju P možemo odabrati tako da su svi elementi matrice L po apsolutnoj vrijednosti najviše jednaki jedinici.

TEOREM. Neka je $A \in \mathbb{R}^{n\times n}$ proizvoljna matrica. Tada postoj permutacija P takva da Gaussove eliminacije daju LU faktorizaciju $PA = LU$ matrice PA. Matrica $L = [l_{ij}]$ je donjetrokutasta s jedinicama na dijagonali, a $U = [u_{ij}]$ je gornjetrokutasta matrica. Pri tome, ako je P umnožak od p inverzija, vrijedi

$$\det A = (-1)^p \prod_{i=1}^{n} u_{ii}.$$

Ako su matrice $P^{(k)}$ odabrane tako da vrijedi

$$\left| \left(P^{(k)} A^{(k-1)} \right)_{kk} \right| = \max_{1\leq j\leq n} \left| \left(P^{(k)} A^{(k-1)} \right)_{jk} \right|$$

onda je

$$\max_{1\leq k\leq n} \max_{1\leq i,j\leq n} \left| \left(L^{(k)} \right)_{ij} \right| = \max_{1\leq i,j\leq n} |l_{ij}| = 1.$$

U tom slučaju faktorizaciju $PA = LU$ nazivamo LU **faktorizacijom s pivotiranjem redaka**.

4 Numerička svojstva Gaussovih eliminacija

Do sada nismo razmatrali praktične detalje realizacije izvedenih algoritama u računalu uz što su vezani mnogi problemi, kao npr.:

- računalo je ograničen stroj s konačnim memorijskim prostorom,
- ne raspolažemo s cijelim skupom $\mathbb{R}$ već s konačno mnogo brojeva,
- realne brojeve u stvarnosti aproksimiramo racionalnim brojevima.

U razmatranju događanja u računalu praktično je razdvojiti LU faktorizaciju od rješavanja trokutastog sustava, pa ćemo tako i napraviti.

4.1 Važnost pivotiranja

Prije nego što nastavimo naglasimo da je za spremanje elemenata matrice A u memoriji računala potrebno izdvojiti n^2 memorijskih lokacija. No zbog posebnog oblika matrica L i U za njih obje zajedno je također potrebno n^2 memorijskih lokacija, dakle onoliko koliko i za samu matricu A. Ako pažljivo promatramo proces računanja LU faktorizacije uočavamo da ga možemo izvesti tako da matrica U ostane zapisana u gornjem trokutu matrice A, a strogo donji trokut matrice L u strogo donjem trokutu matrice A (jedinice na dijagonali nije potrebno pamtiti!). Sve matrice $A^{(k)}$, $K = 1, \ldots, n-1$ pohranjujemo na istom $n \times n$ polju koje u početku sadrži matricu $A^{(0)} = A$. Zapis algoritma za faktorizaciju time postaje još jednostavniji.

ALGORITAM

```
for k = 1 to n − 1
    begin
    for j = k + 1 to n
        begin
        a_jk = a_jk / a_kk;
        end
    for j = k + 1 to n
        begin
        for i = k + 1 to n
            begin
            a_ij = a_ij − a_ik a_kj;
            end
        end
    end
```

PRIMJER. Neka je α mali parametar i neka je matrica A definirana s

$$A = \begin{bmatrix} \alpha & 1 \\ 1 & 1 \end{bmatrix}.$$

U egzaktnom računanju imamo

$$A = \begin{bmatrix} \alpha & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/\alpha & 1 \end{bmatrix} \begin{bmatrix} \alpha & 1 \\ 0 & 1 - 1/\alpha \end{bmatrix} = LU.$$

Pretpostavimo da ovaj račun provodimo na računalu u aritmetici s 8 decimalnih znamenki, tj. da je strojna preciznost $\varepsilon \approx 10^{-8}$. Neka je $|\alpha| < \varepsilon$, recimo $\alpha = 10^{-10}$. Vrijedi

$$\begin{aligned} \widetilde{l}_{21} &= = 1 \oslash \alpha = l_{21}\,(1 + \epsilon_1)\,, \quad |\epsilon_1| \leq \varepsilon \\ \widetilde{u}_{11} &= u_{11} \\ \widetilde{u}_{12} &= u_{12} \\ \widetilde{u}_{22} &= 1 \ominus 1 \oslash \alpha \stackrel{zaok.}{=} -1 \oslash \alpha. \end{aligned}$$

Čini se da su sve nastale greške male, no pokušamo li egzaktno riješiti problem vidjet ćemo da nije sve kako treba.

Imamo

$$\begin{aligned}\widetilde{L}\widetilde{U} &= \begin{bmatrix} 1 & 0 \\ 1 \oslash \alpha & 1 \end{bmatrix} \begin{bmatrix} \alpha & 1 \\ 0 & -1 \oslash \alpha \end{bmatrix} = \begin{bmatrix} \alpha & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \alpha & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} = A + \delta A.\end{aligned}$$

Primijetimo da δA ne možemo smatrati malom perturbacijom matrice A! Naime, jedan od najvećih elemenata matrice A, $a_{22} = 1$, je promijenjen u nulu. Ako bismo koristeći ovakvu faktorizaciju nastavili tražiti rješenje dobili bismo

$$\begin{bmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 - 2\alpha \end{bmatrix},$$

dok je egzaktno rješenje

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{-1}{\alpha-1} \\ \frac{2\alpha-1}{\alpha-1} \end{bmatrix} \approx \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Vidimo da do potpuno krivog rješenja nije došlo zbog akumuliranja velikog broja grešaka zaokruživanja. Problem je u samo jednoj aritmetičkoj operaciji (računaje $\widetilde{u}_{22}$) koja je u stvari izvršena jako točno i s vrlo malom greškom zaokruživanja. Odgovor leži u činjenici da je pivot α vrlo malen! Pogledajmo što bi se dogodilo da smo proveli pivotiranje.

$$A' = PA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \alpha & 1 \end{bmatrix}$$

$$A' = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 - \alpha \end{bmatrix} = LU$$

$$\widetilde{L} = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} = L, \ \ \widetilde{U} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \ \ |\alpha| > \varepsilon.$$

Dakle, imamo

$$\widetilde{L}\widetilde{U} = \begin{bmatrix} 1 & 1 \\ \alpha & 1+\alpha \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \alpha & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \alpha \end{bmatrix} = A' + \delta A', \quad |\delta A'| \leq \varepsilon \, |A'|$$

i umnožak

$$\left|\widetilde{L}\right| \left|\widetilde{U}\right| = \begin{bmatrix} 1 & 1 \\ |\alpha| & 1+\alpha \end{bmatrix}$$

je po elementima istog reda veličine kao i $|A'|$ ($|M|$ poznačava matricu kojoj su elementi aposlutne vrijednosti elemenata matrice M).
Dakle, pivotiranje redaka, koje osigurava da u matrici L svi elementi po apsolutnoj vrijednosti budu najviše jednaki jedinici, pridonosi numeričkoj stabilnosti. Naime, tada su i svi elementi matrice $\left|\widetilde{L}\right|$ manji ili jednaki jedan, pa veličina elemenata umnoška $\left|\widetilde{L}\right| \left|\widetilde{U}\right|$ bitno ovisi o elementima matrice $\left|\widetilde{U}\right|$, a ovi su, pak, dobiveni iz matrica $\widetilde{A}^{(k)}$.

TEOREM. Neka je LU faktorizacija $n \times n$ matrice A izračunata s pivotiranjem redaka u aritmetici s preciznošću ε i neka su matrice $\widetilde{L}$ i $\widetilde{U}$ dobivene aproksimacije matrica L i U. Ako je pri tome korištena permutacija P, onda je

$$\widetilde{L}\widetilde{U} = P\left(A + \delta A\right), \quad |\delta A| \leq \frac{2n\varepsilon}{1 - 2n\varepsilon} P^T \left|\widetilde{L}\right| \left|\widetilde{U}\right|.$$

PROPOZICIJA. Ako LU faktorizaciju računamo s pivotiranjem redaka u aritmetici s maksimalnom greškom zaokruživanja ε, onda je

$$\rho = \frac{\max_{i,j,k} \widetilde{a}_{ij}^{(k)}}{\max_{i,j,k} a_{ij}^{(k)}} \leq 2^{n-1} \left(1 + \varepsilon\right)^{2(n-1)},$$

gdje je $A^{(k)} = \left[a_{ij}^{(k)}\right]$, a $\widetilde{A}^{(k)} = \left[\widetilde{a}_{ij}^{(k)}\right]$ strojno izračunata $A^{(k)}$.

4.2 Analiza numeričkog rješenja trokutastog sustava

PROPOZICIJA. Neka je T donjetrokutasta (gornjetrokutasta) matrica reda n i neka je sustav $T\boldsymbol{v} = \boldsymbol{d}$ riješen supstitucijama unaprijed (unazad). Ako je $\widetilde{v}$ rješenje dobiveno primjenom aritmetike računala s preciznošću ε, onda postoji donjetrokutasta (gornjetrokutasta) matrica δT takva da vrijedi

$$(T + \delta T)\,\widetilde{\boldsymbol{v}} = \boldsymbol{d},$$

gdje je

$$|\delta T| \leq \eta\,|T|\,, \quad 0 \leq \eta \leq \frac{n\varepsilon}{1 - n\varepsilon}.$$

Zaključak ove propozicije je vrlo važan: izračunato rješenje zadovoljava trokutasti sustav s matricom koeficijenata koja se po elementima malo razlikuje od zadane. Npr. radimo li s preciznošću $\varepsilon = 10^{-8}$ i ako je $n = 1000$, onda izračunato rješenje $\widetilde{\boldsymbol{v}}$ zadovoljava jednadžbu $\widetilde{T}\widetilde{\boldsymbol{v}} = \boldsymbol{d}$, gdje se elementi matrice $\widetilde{T}$ i T poklapaju na barem 5 decimala (od osam mogućih).

4.3 Točnost izračunatog rješenja sustava

Sada ćemo ocijeniti koliko točno možemo na računalu riješiti linearni sustav $A\boldsymbol{x} = \boldsymbol{b}$ u kojem smo izračunali LU faktorizaciju $PA = LU$ i supstitucijama naprijed i nazad izračunali rješenje

$$\boldsymbol{x} = U^{-1}\left(L^{-1}(P\boldsymbol{b})\right).$$

Kao što smo vidjeli u prethodnom, numeričku analizu možemo provesti bez pivotiranja jer možemo pretpostaviti da smo permutaciju odmah primijenili na polazne podatke. Dakle, radi jednostavnosti formula pretpostavimo da su na matrice A i $\boldsymbol{b}$ već primijenjene zamjene redaka, tako da su formule jednostavno

$$A = LU$$

i

$$\boldsymbol{x} = U^{-1}\left(L^{-1}\boldsymbol{b}\right).$$

Neka su $\widetilde{L}$ i $\widetilde{U}$ strojno izračunate matrice, pri čemu je

$$\widetilde{L}\widetilde{U} = A + \delta A.$$

Prema prethodnoj propoziciji za izračunato rješenje $\widetilde{\boldsymbol{y}}$ sustava $\widetilde{L}\boldsymbol{y} = \boldsymbol{b}$ vrijedi

$$\left(\widetilde{L} + \delta\widetilde{L}\right)\widetilde{\boldsymbol{y}} = \boldsymbol{b}, \quad \left|\delta\widetilde{L}\right| \leq \frac{n\varepsilon}{1 - n\varepsilon}\left|\widetilde{L}\right|,$$

a na isti način za rješenje sustava $\widetilde{U}\boldsymbol{x} = \widetilde{\boldsymbol{y}}$ vrijedi

$$\left(\widetilde{U} + \delta\widetilde{U}\right)\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{y}}, \quad \left|\delta\widetilde{U}\right| \leq \frac{n\varepsilon}{1 - n\varepsilon}\left|\widetilde{U}\right|.$$

Dakle,

$$\left(\widetilde{L} + \delta\widetilde{L}\right)\left(\widetilde{U} + \delta\widetilde{U}\right)\widetilde{\boldsymbol{x}} = \boldsymbol{b},$$

to jest

$$(A + \delta A + E)\,\widetilde{\boldsymbol{x}} = \boldsymbol{b}, \quad E = \widetilde{L}\delta\widetilde{U} + \delta\widetilde{L}\widetilde{U} + \delta\widetilde{L}\delta\widetilde{U}.$$

Time smo dokazali sljedeći teorem.

TEOREM. Neka je $\widetilde{\boldsymbol{x}}$ rješenje regularnog $n \times n$ sustava $A\boldsymbol{x} = \boldsymbol{b}$ dobiveno Gaussovim eliminacijama s pivotiranjem redaka u aritmetici s preciznošću ε takvom da je $2n\varepsilon < 1$. Tada postoji perturbacija ΔA matrice A za koju vrijedi

$$(A + \Delta A)\,\widetilde{\boldsymbol{x}} = \boldsymbol{b},$$

pri čemu je

$$|\Delta A| \leq \frac{5n\varepsilon}{1 - 2n\varepsilon} P^T \left|\widetilde{L}\right| \left|\widetilde{U}\right|.$$

DOKAZ. Znamo da vrijedi

$$\Delta A = \delta A + P^T E = \delta A + P^T \widetilde{L} \delta \widetilde{U} + P^T \delta \widetilde{L} \widetilde{U} + P^T \delta \widetilde{L} \delta \widetilde{U},$$

pa je

$$|\Delta A| \leq |\delta A| + P^T \left|\widetilde{L}\right| \left|\delta \widetilde{U}\right| + P^T \left|\delta \widetilde{L}\right| \left|\widetilde{U}\right| + P^T \left|\delta \widetilde{L}\right| \left|\delta \widetilde{U}\right|,$$

gdje je

$$\begin{aligned} |\delta A| &\leq \frac{2n\varepsilon}{1 - 2n\varepsilon} P^T \left|\widetilde{L}\right| \left|\widetilde{U}\right|, \\ \left|\delta \widetilde{L}\right| &\leq \frac{n\varepsilon}{1 - n\varepsilon} \left|\widetilde{L}\right|, \\ \left|\delta \widetilde{U}\right| &\leq \frac{n\varepsilon}{1 - n\varepsilon} \left|\widetilde{U}\right|. \end{aligned}$$

Dakle, imamo

$$|\Delta A| \leq \left[\frac{2n\varepsilon}{1 - 2n\varepsilon} + \frac{2n\varepsilon}{1 - n\varepsilon} + \left(\frac{n\varepsilon}{1 - n\varepsilon}\right)^2\right] P^T \left|\widetilde{L}\right| \left|\widetilde{U}\right|.$$

Kako je

$$0 \leq \frac{2n\varepsilon}{1-n\varepsilon} \leq \frac{2n\varepsilon}{1-2n\varepsilon},$$

a zbog $0 < 2n\varepsilon < 1$ (iz čega je posebno $0 < n\varepsilon < 1$) je također

$$0 \leq \left(\frac{n\varepsilon}{1-n\varepsilon}\right)^2 = \frac{(n\varepsilon)^2}{1-2n\varepsilon+n^2\varepsilon^2} \leq \frac{(n\varepsilon)^2}{1-2n\varepsilon} \leq \frac{n\varepsilon}{1-2n\varepsilon}.$$

Odmah slijedi

$$\begin{aligned} |\Delta A| &\leq \left[\frac{2n\varepsilon}{1-2n\varepsilon} + \frac{2n\varepsilon}{1-2n\varepsilon} + \frac{n\varepsilon}{1-2n\varepsilon}\right] P^T \left|\widetilde{L}\right| \left|\widetilde{U}\right| \\ &= \frac{5n\varepsilon}{1-2n\varepsilon} P^T \left|\widetilde{L}\right| \left|\widetilde{U}\right| \qquad \blacksquare \end{aligned}$$

5 Faktorizacija Choleskog

Kažemo da je simetrična $n \times n$ matrica **pozitivno definitna** ako za sve $\boldsymbol{x} \in \mathbb{R}^n$ različite od $\mathbf{0}$ vrijedi

$$\boldsymbol{x}^T A \boldsymbol{x} > 0.$$

Uzmemo li npr. da je $\boldsymbol{x} = \boldsymbol{e}_i$ (i-ti stupac jedinične matrice reda n), onda je

$$\boldsymbol{e}_i^T A \boldsymbol{e}_i = a_{ii} > 0,$$

što znači da su dijagonalni elementi pozitivno definitne matrice nužno pozitivni. Nadalje, ako je S bilo koja regularna matrica i $\boldsymbol{x} \neq \mathbf{0}$, onda je i $\boldsymbol{y} = S\boldsymbol{x} \neq \mathbf{0}$ i vrijedi

$$\boldsymbol{x}^T \left(S^T A S\right) \boldsymbol{x} = (S\boldsymbol{x})^T A \, (S\boldsymbol{x}) = \boldsymbol{y}^T A \boldsymbol{y} > 0,$$

iz čega slijedi da je i $S^T A S$ pozitivno definitna matrica.

No zašto je nama važan pojam pozitivne definitnosti matrice? Pokazuje se da pozitivna definitnost matrice sustava osigurava egzistenciju LU faktorizacije bez pivotiranja. Pogledajmo kako.

Ako pozitivno definitnu matricu A razdijelimo u blokmatrice tako da je

$$A = \begin{bmatrix} a_{11} & \boldsymbol{a}^T \\ \boldsymbol{a} & \hat{A} \end{bmatrix}, \quad \hat{A} \in \mathbb{R}^{(n-1)\times(n-1)}, \quad \boldsymbol{a} \in \mathbb{R}^{n-1},$$

onda je $a_{11} > 0$ i prvi korak eliminacija je

$$\begin{bmatrix} 1 & \mathbf{0}^T \\ -\frac{1}{a_{11}}\boldsymbol{a} & I_{n-1} \end{bmatrix} \begin{bmatrix} a_{11} & \boldsymbol{a}^T \\ \boldsymbol{a} & \hat{A} \end{bmatrix} = \begin{bmatrix} a_{11} & \boldsymbol{a}^T \\ \mathbf{0} & \hat{A} - \frac{1}{a_{11}}\boldsymbol{a}\boldsymbol{a}^T \end{bmatrix}.$$

Sada primijetimo da vrijedi i

$$\begin{aligned} &\begin{bmatrix} 1 & \mathbf{0}^T \\ -\frac{1}{a_{11}}\boldsymbol{a} & I_{n-1} \end{bmatrix} \begin{bmatrix} a_{11} & \boldsymbol{a}^T \\ \boldsymbol{a} & \hat{A} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{a_{11}}\boldsymbol{a}^T \\ \mathbf{0} & I_{n-1} \end{bmatrix} \\ &= \begin{bmatrix} 1 & \mathbf{0}^T \\ -\frac{1}{a_{11}}\boldsymbol{a} & I_{n-1} \end{bmatrix} \begin{bmatrix} a_{11} & \boldsymbol{a}^T \\ \boldsymbol{a} & \hat{A} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ -\frac{1}{a_{11}}\boldsymbol{a} & I_{n-1} \end{bmatrix}^T \\ &= \begin{bmatrix} a_{11} & \mathbf{0}^T \\ \mathbf{0} & \hat{A} - \frac{1}{a_{11}}\boldsymbol{a}\boldsymbol{a}^T \end{bmatrix}. \end{aligned}$$

Vidimo da je novodobivena matrica oblika S^TAS, pa je sigurno pozitivno definitna. Iz ovog se lako dobije i da je matrica $\hat{A} - \frac{1}{a_{11}}\boldsymbol{a}\boldsymbol{a}^T$ pozitivno definitna, dakle je i njen prvi element na glavnoj dijagonali pozitivan, što znači da se potupak eliminacija može nastaviti na isti način. Time je dokazana egzistencija faktorizacije matrice A u oblik R^TR, gdje je matrica R gornjetrokutasta. Ovakvu faktorizaciju $A = R^TR$ nazivamo faktorizacijom Choleskog ili trokutastom faktorizacijom pozitivno definitne matrice. Elemente matrice R možemo izračunati jednostavnim nizom formula. Raspisivanjem faktorizacije po komponentama

$$A = R^TR = \begin{bmatrix} r_{11} & 0 & \cdots & \cdots & 0 \\ r_{12} & r_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ r_{1n} & r_{2n} & \cdots & r_{n-1,n} & r_{nn} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & \cdots & r_{1n} \\ 0 & r_{22} & 0 & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & r_{n-1,n-1} & r_{n-1,n} \\ 0 & 0 & \cdots & 0 & r_{nn} \end{bmatrix}$$

lako dobijemo

$$a_{ij} = \sum_{k=1}^{i} r_{ki}r_{kj}, \qquad i \leq j.$$

Iz prethodnog direktno slijedi algoritam za računanje faktorizacije Choleskog pozitivno definitne matrice $A \in \mathbb{R}^{n\times n}$.

ALGORITAM

$r_{11} = \sqrt{a_{11}};$
$r_{21} = a_{12}/r_{11};$
for $i = 2$ **to** n
    begin
    $r_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2};$
    for $j = i+1$ **to** n
        begin
        $r_{ij} = \left(a_{ij} - \sum_{k=1}^{i-1} r_{ki}r_{kj}\right)/r_{ii}$
        end
    end

Promotrimo li kvadratni sustav linearnih jednadžbi u kojem je matrica sustava $A\boldsymbol{x} = \boldsymbol{b}$ pozitivno definitna i $A = R^T R$ njena trokutasta faktorizacija, onda rješenje sustava

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} = \left(R^T R\right)^{-1} \boldsymbol{b} = R^{-1} \left(R^T\right)^{-1} \boldsymbol{b}$$

možemo dobiti $\boldsymbol{y}$ tako da prvo nađemo rješenje $\boldsymbol{y}$ sustava $R^T \boldsymbol{y} = \boldsymbol{b}$, a zatim još riješimo sustav $R\boldsymbol{x} = \boldsymbol{y}$. Kako je R gornjetrokutasta matrica, cijeli postupak je vrlo jednostavan (pogledati u prethodnom kako se rješavaju trokutasti sustavi!).

Naš sljedeći cilj je ispitati numerička svojstva takvog postupka ako se njegove operacije izvode na računalu u aritmetici s preciznošću ε. Podsjetimo se da treba izvesti sljedeće postupke:

- trokutastu faktorizaciju $A = R^T R$;
- supstitucije unaprijed za rješavanje sustava $R^T \boldsymbol{y} = \boldsymbol{b}$;
- supstitucije unazad za rješavanje sustava $R\boldsymbol{x} = \boldsymbol{y}$.

Ako je $\widetilde{\boldsymbol{x}}$ izračunata aproksimacija točnog rješenja sustava $A\boldsymbol{x} = \boldsymbol{b}$, što možemo reći o $\widetilde{\boldsymbol{x}}$? Znamo da provodeći račun u računalu u aritmetici s preciznošću ε dobivamo približni rastav perturbirane matrice A, tj. znamo da vrijedi

$$\widetilde{R}^T \widetilde{R} = A + \delta A.$$

Dalje rješavamo dva trokutasta sustava $\widetilde{R}^T \boldsymbol{y} = \boldsymbol{b}$ i $\widetilde{R}\boldsymbol{x} = \boldsymbol{y}$ za koja ponovno dobivamo tek približna rješenja $\widetilde{\boldsymbol{x}}$ i $\widetilde{\boldsymbol{y}}$. Prema jednoj od prethodnih propozicija postoje gornjetrokutaste matrice $\delta_1 \widetilde{R}$ i $\delta_2 \widetilde{R}$ takve da vrijedi

$$\left(\widetilde{R} + \delta_1 \widetilde{R}\right)^T \widetilde{\boldsymbol{y}} = \boldsymbol{b}, \quad \left(\widetilde{R} + \delta_2 \widetilde{R}\right) \widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{y}},$$

pri čemu je

$$\left|\delta_1 \widetilde{R}\right| \leq \eta \left|\widetilde{R}\right|, \quad \left|\delta_2 \widetilde{R}\right| \leq \eta \left|\widetilde{R}\right|, \quad \text{gdje je } \eta \leq \frac{n\varepsilon}{1 - n\varepsilon}.$$

Iz ovoga dobijemo sljedeće

$$\left(\widetilde{R} + \delta_1 \widetilde{R}\right)^T \widetilde{\boldsymbol{y}} = \left(\widetilde{R} + \delta_1 \widetilde{R}\right)^T \left(\widetilde{R} + \delta_2 \widetilde{R}\right) \widetilde{\boldsymbol{x}} = \boldsymbol{b},$$

to jest

$$\left(\widetilde{R}^T \widetilde{R} + \widetilde{R}^T \delta_2 \widetilde{R} + \left(\delta_1 \widetilde{R}\right)^T \widetilde{R} + \left(\delta_1 \widetilde{R}\right)^T \delta_2 \widetilde{R}\right) \widetilde{\boldsymbol{x}} = \boldsymbol{b}.$$

Označimo li

$$E = \widetilde{R}^T \delta_2 \widetilde{R} + \left(\delta_1 \widetilde{R}\right)^T \widetilde{R} + \left(\delta_1 \widetilde{R}\right)^T \delta_2 \widetilde{R},$$

imamo

$$\left(\widetilde{R}^T \widetilde{R} + E\right) \widetilde{\boldsymbol{x}} = \boldsymbol{b},$$

pri čemu je

$$|E| \leq \left(2\eta + \eta^2\right) \left|\widetilde{R}\right|^T \left|\widetilde{R}\right|.$$

Dakle, možemo zaključiti da dobiveno rješenje $\widetilde{\boldsymbol{x}}$ zadovoljava sustav

$$\left(\widetilde{R}^T\widetilde{R}+E\right)\widetilde{\boldsymbol{x}}=\boldsymbol{b},$$

u kojem je E po elementima mala perturbacija približne matrice sustava $\widetilde{A}=\widetilde{R}^T\widetilde{R}$. Pri tom vrijedi sljedeća veza između izračunatog rješenja i polaznog sustava:

$$(A+F)\,\widetilde{\boldsymbol{x}}=\boldsymbol{b},\qquad F=\delta A+E.$$

Matrice pogreške δA i E ocijenjene su po elementima dovoljno malim ogradama, pa je dobivena matrica $A+F$ "blizu" matrice A (da smo npr. nakon faktorizacije sustave rješavali egzaktno, imali bismo $F=\delta A$). No ovi zaključci nisu zadovoljavajući! Naime, tijekom ovog postupka izgubili smo simetričnost matrice sustava A jer matrica E, a time ni matrica F, općenito ne mora biti simetrična. Simetrija matrice A je najčešće posljedica strukture problema kojeg opisujemo linearnim sustavom $R\boldsymbol{x}=\boldsymbol{b}$, pa nam je važno da je izračunato rješenje $\widetilde{\boldsymbol{x}}$ rješenje bliskog problema s istom strukturom.

To nas vodi do sljedećeg problema: ako je $(A+F)\,\widetilde{\boldsymbol{x}} = \boldsymbol{b}$, postoji li simetrična perturbacija ΔA za koju postoje dovoljno dobre ocjene, a takva da je $(A+\Delta A)\,\widetilde{\boldsymbol{x}} = \boldsymbol{b}$? Naredni teorem daje potvrdan odgovor na to pitanje.

TEOREM. Neka je $(A+F)\,\widetilde{\boldsymbol{x}} = \boldsymbol{b}$, gdje je A pozitivno definitna matrica, te neka vrijedi

$$\max_{i,j} \frac{|F_{ij}|}{\sqrt{a_{ii}a_{jj}}} \leq \xi.$$

Tada postoji simetrična perturbacija ΔA takva da je $(A+\Delta A)\,\widetilde{\boldsymbol{x}} = \boldsymbol{b}$. Pri tome je

$$\max_{i\neq j} \frac{|\Delta a_{ij}|}{\sqrt{a_{ii}a_{jj}}} \leq \xi, \quad \max_{i} \frac{|\Delta a_{ij}|}{a_{ii}} \leq (2n-1)\,\xi.$$

ZAKLJUČAK. Pozitivno definitne sustave možemo na računalu riješiti s pogreškom koja je ekvivalentna malim promjenama koeficijenata u matrici sustava, a da se pri tom očuva struktura polaznog problema (simetrija).

DOKAZ. (Skica) Primijetimo da perturbacija ΔA mora zadovoljavati jednadžbu $\Delta A\widetilde{\boldsymbol{x}} = F\widetilde{\boldsymbol{x}}$ koja daje n uvjeta za $n(n+1)/2$ stupnjeva slobode u ΔA (zbog uvjeta simetričnosti). Stavimo

$$D = \text{diag}\left(\sqrt{a_{ii}}\right)_{i=1}^{n}$$

i promotrimo skalirani sustav

$$D^{-1}(A+F)\,D^{-1}D\widetilde{\boldsymbol{x}} = D^{-1}\boldsymbol{b}$$

koji ćemo zapisati u obliku

$$(A_s+F_s)\,\boldsymbol{z} = D^{-1}\boldsymbol{b}, \quad A_s = D^{-1}AD^{-1}, \quad F_s = D^{-1}FD^{-1}, \quad \boldsymbol{z} = D\widetilde{\boldsymbol{x}}.$$

Neka je permutacija P takva da vektor $\widetilde{\boldsymbol{z}} = P^T\boldsymbol{z}$ zadovoljava

$$|\widetilde{z}_1| \le |\widetilde{z}_2| \le \cdots \le |\widetilde{z}_n|$$

(tj. permutacija P je takva da komponente vektora posloži u rastućem poretku). Uočimo i da vrijedi $PP^T = I$. Gornji sustav zapišemo u ekvivalentnom obliku

$$P^T(A_s+F_s)\,PP^T\boldsymbol{z} = P^T(A_s+F_s)\,P\widetilde{\boldsymbol{z}} = P^TD^{-1}\boldsymbol{b}.$$

Ponovno uvedemo kraće oznake

$$A_{s,p} = P^T A_s P, \quad F_{s,p} = P^T F P,$$

pa sustav možemo zapisati u obliku

$$(A_{s,p} + F_{s,p})\, \widetilde{\boldsymbol{z}} = P^T D^{-1} \boldsymbol{b}.$$

Konstruirat ćemo simetričnu matricu M za koju vrijedi

$$M\widetilde{\boldsymbol{z}} = F_{s,p}\widetilde{\boldsymbol{z}}.$$

Definiramo

$$m_{ij} = \begin{cases} (F_{s,p})_{ij}\,, & i < j \\ (F_{s,p})_{ji}\,, & j < i \end{cases},$$

dok dijagonalne elemente $m_{ii},\ i = 1, \ldots, n$, odredimo tako da vrijedi

$$m_{ii}\widetilde{z}_i + \sum_{i \neq j} m_{ij}\widetilde{z}_j = (F_{s,p})_{ii}\, \widetilde{z}_i + \sum_{i \neq j} (F_{s,p})_{ij}\, \widetilde{z}_j.$$

Za ovako definiranu simetričnu matricu M pokazuje se da vrijedi

$$(A_{s,p} + M)\, \widetilde{\boldsymbol{z}} = P^T D^{-1} \boldsymbol{b},$$

ili ekvivalentno

$$\left(A_{s,p} + PMP^T\right) \boldsymbol{z} = D^{-1} \boldsymbol{b}.$$

Pri tome je

$$\max_{i \neq j} \left| \left(PMP^T\right)_{ij} \right| \leq \xi,$$

i

$$\max_{i} \left| \left(PMP^T\right)_{ii} \right| \leq (2n-1)\, \xi.$$

Skaliranjem sustava (unatrag) dobijemo da za simetričnu perturbaciju $\Delta A = D \left(PMP^T\right) D$ vrijedi

$$(A + \Delta A)\, \widetilde{\boldsymbol{x}} = \boldsymbol{b}$$

■

6 Iterativne metode

U prethodnom smo vidjeli da se rješenje linearnog sustava $A\boldsymbol{x} = \boldsymbol{b}$ općenito ne može izračunati potpuno točno: najčešće dobivamo rješenje $\widetilde{\boldsymbol{x}}$ koje zadovoljava sustav

$$(A + \delta A)\,\widetilde{\boldsymbol{x}} = \boldsymbol{b}$$

blizak polaznom sustavu. Dakle, Gaussove eliminacije ne garantiraju idealnu točnost.

Osim toga, u praksi moramo biti svjesni da je računalo ograničeno ne samo po pitanju numeričke točnosti, već i raspoloživim vremenom i memorijom. U primijenjenoj matematici najčešće se javljaju sustavi velikih dimenzija $\left(n > 10^5\right)$ kod kojih je proces Gaussovih eliminacija iz više razloga praktički neprovediv. Npr. spremanje matrice sustava s $n = 10^5$ nepoznanica zahtijeva 10^{10} memorijskih lokacija, pa već i to može predstavljati poteškoću. No važno je istaknuti da su matrice takvih sustava često **rijetko popunjene** (tj. velika većina elemenata im je jednaka nuli, a elementi koji nisu nula su obično pravilno raspoređeni).

Rijetko popunjena matrica po blokovima može izgledati npr. ovako (praznine su blokovi nula):

$$\begin{bmatrix} \blacksquare & \blacksquare & & & & \\ \blacksquare & \blacksquare & & & & \\ & & \blacksquare & & & \\ & & & \blacksquare & \blacksquare & \\ & & & \blacksquare & \blacksquare & \\ & & & & & \blacksquare \end{bmatrix}.$$

Kada je matrica velika i gusto popunjena ono što preostaje jest učitavanje djelova matrice iz vanjske u radnu memoriju.

No ponekad je poznat način na koji se generiraju elementi matrice A, pa nas zanima kako postupiti u takvom slučaju pod uvjetom da ne tražimo egzaktno rješenje $\boldsymbol{x}$, već dovoljno dobru aproksimaciju $\widetilde{\boldsymbol{x}}$. Stoga ima smisla konstruirati niz $\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(k)}, \ldots$ vektora iz $\mathbb{R}^n$ sa sljedećim svojstvima:

- za svaki $k \in \mathbb{N}_0$ formula za računanje $\boldsymbol{x}^{(k)}$ je jednostavna;
- $\boldsymbol{x}^{(k)}$ teži prema $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ za neki k (obično je takav $k << n$).

Kako točno konstruirati takav niz ovisit će o konkretnom problemu kojeg rješavamo.

6.1 Jacobijeva metoda

Jacobijeva metoda je jedna od najjednostavnijih klasičnih iterativnih metoda za rješavanje linearnih sustava. Ideju same metode ilustrirat ćemo na jednostavnom primjeru 2×2 sustava.

Neka je dan sustav

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

pri čemu je $a_{11} \neq 0$ i $a_{22} \neq 0$. Uočimo da rješenje $\boldsymbol{x}$ zadovoljava uvjete

$$\begin{aligned} x_1 &= \frac{1}{a_{11}} (b_1 - a_{12}x_2) \\ x_2 &= \frac{1}{a_{22}} (b_2 - a_{21}x_1) \, . \end{aligned}$$

Te nas relacije motiviraju da neku približnu vrijednost rješenja $\boldsymbol{x}^{(0)} = \begin{bmatrix} x_1^{(0)} & x_2^{(0)} \end{bmatrix}^T$ korigiramo pomoću formula

$$
\begin{aligned}
x_1^{(1)} &= \frac{1}{a_{11}} \left(b_1 - a_{12} x_2^{(0)} \right) \\
x_2^{(1)} &= \frac{1}{a_{22}} \left(b_2 - a_{21} x_1^{(0)} \right).
\end{aligned}
$$

Naravno, nadamo se da je $\boldsymbol{x}^{(1)}$ bolja aproksimacija egzaktnog rješenja $\boldsymbol{x}$ nego $\boldsymbol{x}^{(0)}$. Postupak možemo nastaviti tako da pomoću $\boldsymbol{x}^{(1)}$ izračunamo na isti način $\boldsymbol{x}^{(2)}$ itd. Pitanje je pod kojim uvjetima tako dobivene iteracije teže prema rješenju $\boldsymbol{x}$?

Uočimo da vrijedi

$$
\begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} 1/a_{11} & 0 \\ 0 & 1/a_{22} \end{bmatrix} \left(\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} 0 & -a_{12} \\ -a_{21} & 0 \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} \right).
$$

Dakle, ako stavimo

$$A = D - N, \quad D = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}, \quad N = \begin{bmatrix} 0 & -a_{12} \\ -a_{21} & 0 \end{bmatrix},$$

možemo jednostavno pisati

$$\boldsymbol{x}^{(k+1)} = D^{-1} \left(\boldsymbol{b} + N \boldsymbol{x}^{(k)} \right) = D^{-1} N \boldsymbol{x}^{(k)} + D^{-1} \boldsymbol{b}.$$

Upravo ovom relacijom definirana je **Jacobijeva iterativna metoda**.

PRIMJER. Neka je

$$A = \begin{bmatrix} 2 & 0.1 \\ -0.1 & 2 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 19.9 \\ -3 \end{bmatrix}.$$

Lako se provjeri da je

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} = \begin{bmatrix} 10 \\ -1 \end{bmatrix}.$$

Za početnu iteraciju uzmimo vektor

$$\boldsymbol{x}^{(0)} = D^{-1}\boldsymbol{b} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 19.9 \\ -3 \end{bmatrix} = \begin{bmatrix} 9.49999999999999 \\ -1.5 \end{bmatrix}.$$

Naš izbor je rezultat jednostavne ideje da matricu A aproksimiramo matricom D jer su joj elementi na dijagonali veći od izvandijagonalnih. Ovo je gruba aproksimacija, no ima smisla.

Iteriranjem dobijemo

$$\boldsymbol{x}^{(5)} = \begin{bmatrix} 1.000000015625000e+001 \\ -1.000000015625000e+000 \end{bmatrix},$$

dok je relativna greška

$$e_k = \left\| \boldsymbol{x} - \boldsymbol{x}^{(k)} \right\|_\infty / \|\boldsymbol{x}\|_\infty$$

jednaka

$$e_5 = 1.562499996055067e-007.$$

Lako se vidi da je

$$\boldsymbol{e}^{(k+1)} = \boldsymbol{x}^{(k+1)} - \boldsymbol{x} = D^{-1}N\left(\boldsymbol{x}^{(k)} - \boldsymbol{x}\right) = D^{-1}N\boldsymbol{e}^{(k)}, \quad k \in \mathbb{N}_0.$$

Naravno, isto se može napraviti i za veće sustave pod analognim uvjetima na koeficijente matrice A $\left(\prod_{i=1}^{n} a_{ii} \neq 0\right)$.

Iz relacije

$$\boldsymbol{e}^{(k+1)} = D^{-1}N\boldsymbol{e}^{(k)}, \quad k \in \mathbb{N}_0,$$

lako se dobije

$$\boldsymbol{e}^{(k)} = \left(D^{-1}N\right)^k \boldsymbol{e}^{(0)}, \quad k \in \mathbb{N},$$

gdje je $\boldsymbol{e}^{(0)} = \boldsymbol{x}^{(0)} - \boldsymbol{x}$. Uzimanjem proizvoljne vektorske i odgovarajuće matrične norme dobivamo

$$\left\|\boldsymbol{e}^{(k)}\right\| \leq \left\|\left(D^{-1}N\right)^k\right\| \left\|\boldsymbol{e}^{(0)}\right\| \leq \|D^{-1}N\|^k \left\|\boldsymbol{e}^{(0)}\right\|.$$

Iz ove relacije zaključujemo da će $\left\|\boldsymbol{e}^{(k)}\right\|$ težiti k nuli kada $k \to \infty$ za svaki početni $\boldsymbol{x}^{(0)}$ ako $\left\|D^{-1}N\right\|^k$ teži k nuli kada $k \to \infty$. Npr., ako je $\left\|D^{-1}N\right\| < 1$, onda $\left\|D^{-1}N\right\|^k$ teži k nuli kada $k \to \infty$. No ako $\left\|D^{-1}N\right\|^k$ ne teži k nuli kada $k \to \infty$ ne možemo izvesti zaključak o konvergenciji.
Iz ovih argumenata lako slijede dvije naredne propozicije.

PROPOZICIJA. Ako je u rastavu $A = D - N$ u nekoj matričnoj normi ispunjeno

$$\left\| D^{-1} N \right\| < 1,$$

onda za svaku početnu iteraciju $\boldsymbol{x}^{(0)}$ niz

$$\boldsymbol{x}^{(k+1)} = D^{-1} \left(\boldsymbol{b} + N \boldsymbol{x}^{(k)} \right), \quad k \in \mathbb{N}_0,$$

konvergira rješenju $\boldsymbol{x}$ sustava $A\boldsymbol{x} = \boldsymbol{b}$.

PROPOZICIJA. Ako je matrica A dijagonalno dominantna u smislu da je

$$|a_{ii}| > \sum_{j=1,\ j \neq i}^{n} |a_{ij}| \ , \quad i = 1, \ldots, n,$$

onda za svaku početnu iteraciju $\boldsymbol{x}^{(0)}$ niz

$$\boldsymbol{x}^{(k+1)} = D^{-1} \left(\boldsymbol{b} + N \boldsymbol{x}^{(k)} \right), \quad k \in \mathbb{N}_0,$$

konvergira rješenju $\boldsymbol{x}$ sustava $A\boldsymbol{x} = \boldsymbol{b}$.

6.2 Gauss-Seidelova metoda

Vidjeli smo da se u primjeru danom za Jacobijevu metodu $x_1^{(1)}$ i $x_2^{(1)}$ računaju neovisno pomoću $x_1^{(0)}$ i $x_2^{(0)}$. No imalo bi smisla u formuli za $x_2^{(1)}$ koristiti upravo izračunatu vrijednost $x_1^{(1)}$ jer je ona vjerojatno bolja od $x_1^{(0)}$. Općenito, Jacobijevu formulu za iteraciju modificiramo tako da prilikom računanja svake komponente vektora $\boldsymbol{x}^{(k+1)}$ koristimo najsvježije izračunate vrijednosti. Npr. u slučaju $n = 4$ imali bismo

$$\begin{aligned}
x_1^{(k+1)} &= \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - a_{14}x_4^{(k)}\right)\\
x_2^{(k+1)} &= \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - a_{24}x_4^{(k)}\right)\\
x_3^{(k+1)} &= \frac{1}{a_{33}}\left(b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} - a_{34}x_4^{(k)}\right)\\
x_4^{(k+1)} &= \frac{1}{44}\left(b_4 - a_{41}x_1^{(k+1)} - a_{42}x_2^{(k+1)} - a_{43}x_3^{(k+1)}\right).
\end{aligned}$$

U općenitom slučaju imali bismo

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right).$$

No vratimo se primjeru $n = 4$. Stavimo li

$$L = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}, \qquad U = - \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

vrijedi $A = L - U$, pa uz uvjet regularnosti matrice L Gauss-Seidelovu metodu možemo zapisati kao

$$\boldsymbol{x}^{(k+1)} = L^{-1} \left(\boldsymbol{b} + U \boldsymbol{x}^{(k)} \right), \quad k \in \mathbb{N}_0,$$

a kao i u analizi Jacobijeve metode imamo

$$\boldsymbol{e}^{(k)} = \left(L^{-1} U \right)^k \boldsymbol{e}^{(0)}, \ \ k \in \mathbb{N}.$$

PROPOZICIJA. Ako je matrica A pozitivno definitna, onda za svaku početnu iteraciju $\boldsymbol{x}^{(0)}$ niz

$$\boldsymbol{x}^{(k+1)} = L^{-1}\left(\boldsymbol{b} + U\boldsymbol{x}^{(k)}\right), \quad k \in \mathbb{N}_0,$$

konvergira rješenju $\boldsymbol{x}$ sustava $A\boldsymbol{x} = \boldsymbol{b}$.

Uočimo da su Jacobijeva i Gauss-Seidelova metoda vrlo slične: matrica sustava A se zapiše u obliku $A = M - S$, gdje je M regularna matrica, a iteracije su dane formulom

$$\boldsymbol{x}^{(k+1)} = M^{-1}\left(\boldsymbol{b} + S\boldsymbol{x}^{(k)}\right), \quad k \in \mathbb{N}_0.$$

Pri tom je matrica M odabrana tako da ju je lako invertirati (u slučaju $M = D$ je dijagonalna, a u slučaju $M = L$ je donjetrokutasta). Konvergencija prema rješenju je za proizvoljan odabir početne iteracije $\boldsymbol{x}^{(0)}$ osigurana ako je $\left\|M^{-1}S\right\| < 1$ za neku matričnu normu $\|\cdot\|$. Vidjeli smo da će u slučaju Jacobijeve metode to sigurno biti ispunjeno ako je A dijagonalno dominantna, a u slučaju Gauss-Seidelove metode ako je A pozitivno definitna.