# NUMERIKUS MÓDSZEREK I.
# TÉTELEK

Szerkesztette: *Balogh Tamás*

2014. január 19.

Ha hibát találsz, kérlek jelezd a *info@baloghtamas.hu* e-mail címen!

**1. A lebegőpontos számábrázolás egy modellje. A normalizált lebegőpontos szám fogalma, a legnagyobb, legkisebb pozitív szám, a relatív pontosság az $M(t, k^-, k^+)$ gépi számhalmazban. Az input függvény (fl) fogalma, tétel az ábrázolt szám hibájáról. Példák a véges számábrázolás miatt előforduló furcsaságokra.**

Az $a = \pm m 2^k$, ($m = \sum_{i=1}^{t} m_i 2^{-i}$, $m_i \in \{0,1\}, m_1 = 1, t \in \mathbb{N}, k \in \mathbb{Z}$) számot normalizált lebegőpontos számnak nevezzük, ahol $m_i$ a mantissza, $t$ a mantissza hossza, $k$ karakterisztika. Jelölése: $a = \pm[m_1 \dots m_t | k]$.

**Gépi számok halmaza:**

$$M = M(t, k^-, k^+) := \left\{ a = \pm m 2^k \mid m = \sum_{i=1}^{t} m_i 2^{-i}, m_i \in \{0,1\}, m_1 = 1,\, t \in \mathbb{N},\, k^- \le k \le k^+ \right\} \cup \{0\},$$

ahol $\frac{1}{2} \le m \le 1$, és $M$ a 0-ra szimmetrikus.

**Nevezetes gépi számok:**
A legnagyobb ábrázolható pozitív szám:

$$M_\infty = +[11\dots1|k^+] = \left(1 - \frac{1}{2^t}\right) \cdot 2^{k^+}$$

A legkisebb ábrázolható pozitív szám:

$$\varepsilon_0 = +[10\dots0|k^-] = \frac{1}{2} \cdot 2^{k^-}$$

A relatív korlát/pontosság:

$$\varepsilon_1 = \underbrace{[10\dots01 \mid 1]}_{\text{1 rákövetkezője}} - \underbrace{[10\dots00 \mid 1]}_{1} = \frac{1}{2^t} \cdot 2^1 = 2^{1-t}$$

**Input függvény:** A valós számok gépi számokkal való megfeleltetése. $\mathrm{fl}\colon \mathbb{R}^x \to M$, ahol

$$\mathrm{fl}(x) = \begin{cases} M_\infty, & (x > M_\infty) \\ -M_\infty, & (x < -M_\infty) \\ 0, & (0 \le |x| < \varepsilon_0) \\ \text{az } x\text{-hez legközelebbi gépi szám}, & (\varepsilon \le |x| \le M_\infty) \end{cases}$$

Az ábrázolt szám **abszolút hibakorlátja:**
Tétel:

$$|x - \mathrm{fl}(x)| \le \begin{cases} \varepsilon_0, & (0 \le |x| < \varepsilon_0) \\ \frac{1}{2}|x| \cdot \varepsilon_1, & (\varepsilon_0 \le |x| \le M_\infty) \end{cases}$$

Bizonyítás:

(1) $0 \le |x| < \varepsilon_0$ triviális

(2) $\varepsilon_0 \le |x| \le M_\infty$:

a, $x \in M$ : triviális: $|x - \mathrm{fl}(x)| = 0$

b, $x \notin M$ : Ekkor tegyük fel, hogy $x' \le x \le x''$ $\quad (x', x'' \in M$ szomszédjai)
$x' = [m_1 \dots m_t|k]$, az itervallum hossza $\frac{1}{2^t} \cdot 2^k$ (ezzel bevezethető a hiba. Így

$$|x - \mathrm{fl}(x)| \le \frac{1}{2} \cdot \frac{1}{2^t} \cdot 2^k = |x| \cdot \frac{1}{2^t} - \frac{1}{2}|x| \cdot \varepsilon_1. \quad \square$$

Az ábrázolt szám **relatív hibakorlátja:**

Következmény:

$$\frac{|x - \mathrm{fl}(x)|}{|x|} \le \frac{1}{2} \cdot \varepsilon_1 = 2^{-t} \quad (\varepsilon_0 \le |x| \le M_\infty).$$

A gépi ábrázolás miatt előforduló **furcsaságok**:

(1) $a \oplus b = a$, ahol $b \ne 0$. Például:

$$[1000 \mid -3] \quad \to \quad \begin{array}{r} [1100 \mid 1] \\ \oplus [0000 \mid 1] \\ \hline [1100 \mid 1] \end{array}$$

(2) $(a \oplus b) \oplus c \ne a \oplus (b \oplus c)$: asszociativitás nem mindig teljesül. Például:

$$\left.\begin{array}{l} a = [1100 \mid \ \ 1] \\ b = [1000 \mid -3] \end{array}\right\} \oplus : [1100 \mid 1] \oplus c = [1100 \mid 1]$$

$$\left.\begin{array}{l} b = [1000 \mid -3] \\ c = [1000 \mid -3] \end{array}\right\} \oplus : [1000 \mid -2] \oplus a = [1101 \mid 1]$$

(3) Kivonási jegyveszteség (relatív pontosság romlása). Például:

$$\begin{array}{r} [11011 \mid 2] \\ \ominus [11000 \mid 2] \\ \hline [00011 \mid 2] \end{array} = [11000 \mid -1]$$

(4) A részeredmény nem ábrázolható, de az eredmény igen. Például:
$\sqrt{a^2 + b^2}$, ahol $S = \max\{|a|, |b|\}$ nagy. Ekkor $S \cdot \sqrt{(\frac{a}{s})^2 + (\frac{b}{s})^2}$ alakban számolunk.

## 2. A hibaszámítás elemei. Az abszolút és relatív hiba, hibakorlát fogalma. Tétel az alapműveletek abszolút és relatív hibájáról. A függvényérték abszolút és relatív hibája. Függvény egy adott pontbeli kondíciószámának felírása.

Jelölés: $A$: a **pontos érték**, $a$: a **közelítő érték**. (Hibája csak a közelítő értéknek van.)

| | |
|---|---|
| A közelítő érték pontos hibája: | $\Delta a = A - a = a\delta a$ |
| A közelítő érték abszolút hibája: | $\|\Delta a\| = \|A - a\|$ |
| A közelítő érték egy abszolút hibakorlátja: | $\Delta_a \ge \|\Delta a\|$, $\Delta_a = \|a\|\delta_a$ |
| A közelítő érték relatív hibája: | $\delta a = \frac{\Delta a}{A} \approx \frac{\Delta a}{a}$ |
| A közelítő érték egy relatív hibakolátja: | $\delta_a \ge \|\delta a\|$ |

**Az alapműveletek abszolút és relatív hibakorlátjai**

TÉTEL:

$$\begin{array}{ll} \Delta_{a\pm b}=\Delta_a+\Delta_b & \delta_{a\pm b}=\frac{|a|}{|a\pm b|}\delta_a+\frac{|b|}{|a\pm b|}\delta_b \\ \Delta_{a\cdot b}=|b|\Delta_a+|a|\Delta_b & \delta_{a\cdot b}=\delta_a+\delta_b \\ \Delta_{\frac{a}{b}}=\frac{|b|\Delta_a+|a|\Delta_b}{b^2} & \delta_{\frac{a}{b}}=\delta_a+\delta_b \end{array}$$

BIZONYÍTÁS: Tegyük fel, hogy $A$ és $B$ azonos nagyságrendű. Ekkor

$$\Delta(a\cdot b)=AB-ab=AB-aB+aB-ab=B(A-a)+a(B-b)=B\Delta a+a\Delta b=$$

$$(b+\Delta b)\Delta a+\Delta b\cdot a\approx b\cdot\Delta a+a\Delta b,$$

mert $\Delta a\cdot\Delta b$ elhanyagolható. Ekkor

$$|\Delta(a\cdot b)|\le|b||\Delta a|+|a||\Delta b|\le|b|\Delta_a+|a|\Delta_b=\Delta_{a\cdot b},$$

$$\delta(a\cdot b)=\frac{\Delta(a\cdot b)}{ab}=\frac{b\Delta a+a\Delta b}{ab}=\frac{\Delta a}{a}+\frac{\Delta b}{b}=\delta_a+\delta_b,$$

$$|\delta(a\cdot b)|\le|\delta a|+|\delta b|\le\delta_a+\delta_b=\delta_{a\cdot b}.$$

Az összevonás, kivonás és osztás bizonyítása analóg módon. Kivonás és összeadás esetén feltesszük még, hogy $A$ és $B$ azonos előjelű. □

**A függvényérték abszolút hibakorlátja**

TÉTEL: Ha $f\in C^1(k(a))$ $\quad(k(a)=[a-\Delta_a,a+\Delta_a])$, akkor $\Delta_{f(a)}=M_1\cdot\Delta_a$, ahol $M_1=\max\{|f'(\xi)|:\xi\in k(a)\}$.

BIZONYÍTÁS: Lagrange-féle középérték-tételt alkalmazva:

$$\exists\,\xi\in k(a):\Delta f(a)=f(A)-f(a)=f'(\xi)\cdot(A-a)=f'(\xi)\cdot\Delta a,$$

így

$$|\Delta f(a)|=|f'(\xi)|\cdot|\Delta a|\le M_1\cdot\Delta_a=\Delta_{f(a)}.\quad\square$$

TÉTEL: Ha $f\in C^2(k(a))$ $\quad(k(a)=[a-\Delta_a,a+\Delta_a])$, akkor $\Delta_{f(a)}=|f'(a)|\cdot\Delta_a+\frac{M_2}{2}\cdot\Delta_a^2$, ahol $M_2=\max\{|f''(\xi)|:\xi\in k(a)\}$.

BIZONYÍTÁS: A Taylor-formula segítségével:

$$\exists\,\xi\in k(a):f(A)=f(a)+f'(a)(A-a)+\frac{f''(\xi)}{2}(A-a)^2\Rightarrow\Delta f(a)=f(A)-f(a)=$$

$$f'(a)\Delta a+\frac{f''(\xi)}{2}(A-a)^2\Rightarrow|\Delta f(a)|\le|f'(a)|\cdot|\Delta a|+\frac{M_2}{2}|\Delta a|^2\le|f'(a)|\Delta_a+\frac{M_2}{2}\Delta_a^2.\quad\square$$

KÖVETKEZMÉNY: A függvényérték relatív hibakorlátja:

$$\delta f(a)\approx\frac{f'(a)}{f(a)}\cdot\Delta_a,\quad|\delta f(a)|\le\frac{|a|\cdot|f'(a)|}{|f(a)|}\cdot\delta_a=\delta_{f(a)}$$

Az $f$ függvény $a$-beli kondíciószáma:

$$\mathrm{cond}(f,a)=\frac{|a|\cdot|f'(a)|}{|f(a)|},\text{ így }\delta_{f(a)}=\mathrm{cond}(f,a)\cdot\delta_a.$$

**3. Lineáris egyenletrendszerek (LER) megoldása Gauss-eliminációval. Az elimináció és a visszahelyettesítés műveletigénye. A sor-, illetve oszlopcsere szükségessége. A részleges és teljes főelemkiválasztás.**

*Feladat:* $\mathbf{Ax} = \mathbf{b} \to \mathbf{x} =?$, ahol $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$.
Az egyenletrendszer megoldható, ha

- $\mathbf{b}$ kifejezhető $\mathbf{A}$ oszlopvektorainak lineáris kombinációjaként,
- $\mathbf{A}$ oszlopvektorai lineárisan függetlenek.

*Cél:* Felsőháromszög alakra hozni az egyenletrendszert.

**Gauss elimináció** (az egyik háromszög alakra hozó módszer):

- 0. lépés: Legyen $a_{n+1} := b$ és $a_{ij}^{(0)} := a_{ij}$.
- 1. lépés: 1. egyenlet változatlan, következőekből elimináljuk $x_1$-t:
  új i. egyenlet := i. egyenlet $-\frac{a_{i1}}{a_{11}}\cdot$ 1. egyenlet, ahol $a_{11} \neq 0$, $i = 2, \ldots, n$. Így

$$a_{ij}^{(1)} := a_{ij}^{(0)} - \left(\frac{a_{i1}^{(0)}}{a_{11}^{(0)}}\right) \cdot a_{1j}^{(0)} \quad (i = 2, \ldots, n,\ j = 2, \ldots, n+1)$$

- k. lépés: k. egyenlet változatlan, következőekből elimináljuk $x_k$-t:
  új i. egyenlet := i. egyenlet $-\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}\cdot$k. egyenlet, azaz

$$a_{ij}^{(k)} := a_{ij}^{(k-1)} - \left(\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}\right) \cdot a_{kj}^{(k-1)} \quad (k = 1, \ldots, n-1,\ i = k+1, \ldots, n,\ j = k+1, \ldots, n+1).$$

- (n-1). lépés után felsőháromszög-mátrixot kapunk:

$$\begin{array}{rcccccl}
a_{1,1}^{(0)} \cdot x_1 + & \cdots & + \cdots + & a_{1,n}^{(0)} \cdot x_n & = & a_{1,n+1}^{(0)} \\
& \vdots & \vdots & & & \vdots \\
& a_{i,i}^{(i-1)} \cdot x_i + & \cdots + & a_{i,n}^{(i-1)} \cdot x_n & = & a_{i,n+1}^{(i-1)} \\
& & \vdots & & & \vdots \\
& & & a_{n,n}^{(n-1)} \cdot x_n & = & a_{n,n+1}^{(n-1)}
\end{array}$$

- Visszahelyettesítés a felsőháromszüg-mátrixú egyenletrendszerbe:

$$x_i = \underbrace{\frac{1}{a_{i,i}^{(i-1)}}}_{x_i\text{együtthatója}} \cdot \left( \underbrace{a_{i,n+1}^{(i-1)}}_{\text{ami } b_i \text{ helyén keletkezik}} - \underbrace{\sum_{j=i+1}^{n} (a_{i,j}^{(i-1)} \cdot x_j)}_{x_i \text{ utáni } x\text{-ek együtthatójukkal}} \right) \quad (i = n-1, \ldots, 1)$$

**Gauss-elimináció műveletigénye:**

- Eliminációs fázis (felsőháromszög-alak kialakítása):
  a k. lépésben:

$$\begin{array}{ll}
(n-k) & \text{db szorzás,}\\
(n-k)\cdot(n-k+1) & \text{db osztás,}\\
(n-k)\cdot(n-k+1) & \text{db összeadás.}\\
\hline
(n-k)\Big(2(n-k)+3\Big) & \text{db művelet}
\end{array}$$

$$\sum_{k=1}^{n-1}(n-k)\Big(n(n-k)+3\Big) = 2\cdot\sum_{k=1}^{n-1}(n-k)^2 + 3\cdot\sum_{k=1}^{n-1}(n-k) =$$

$$= 2\frac{(n-1)n(n+1)}{6} + 3\frac{(n-1)n}{2} = \boldsymbol{\frac{2}{3}n^3 + \mathcal{O}(n^2)}.$$

- Visszahelyettesítési fázis:
  $x_i$ kifejezésénél:

$$\begin{array}{lll}
1 & \text{db osztás,} & \\
(n-i)\cdot(n-k+1) & \text{db szorzás,} & \\
(n-i)\cdot(n-k+1) & \text{db összeadás.} & +x_n \text{ esetén 1 db osztás}\\
\hline
2(n-i)+1 & \text{db művelet} &
\end{array}$$

$$\Big(\sum_{i=1}^{n-1} 1 + 2(\underbrace{n-i}_{s})\Big) + 1 = \sum_{s=1}^{n-1}(2s+1) + 1 = 2\sum_{s=1}^{n-1} s + (n-1) + 1 =$$

$$= 2\frac{(n-1)n}{2} + n = \boldsymbol{n^2 + \mathcal{O}(n)}.$$

A Gauss-elimináció **elvégezhető sor- és oszlopcsere nélkül** $\Leftrightarrow a_{kk}^{(k-1)} \neq 0 \quad (k = 1 \dots n-1)$.
Ha a k. lépésben mégis $a_{k,k}^{(k-1)} = 0$:

- lehetséges a sorcsere (egyenlet), a megoldás nem változik,
- oszlopcsere (a megoldásvektor komponensei cserélődnek).

Kézi számításnál csak akkor cserélünk, ha muszáj. **Gépi számolás esetén főelemkiválasztást** alkalmazunk.
**Részleges főelemkiválasztás:** A k. lépésnél: $\{a_{k,k}^{(k-1)}, \dots, a_{n,k}^{(k-1}\}$ közül a maximális abszolútértékű elem sorát cseréljük a k. sorral, a megoldás nem változik.

**Teljes főelemkiválasztás:** A k. lépésnél a [k.-n.] sorok és oszlopok által meghatározott részmátrixban keressük a legnagyobb abszolútértékű elemet, ennek a sorát a k. sorral, illetve oszlopát a k. oszloppal cseréljük. A megoldás változik az oszlopcsere miatt, ezt nyomon kell követni.

Mindkét főelemkiválasztós eljárásban a sor és oszlopcserét nem végezzük el, helyette az induláskor felvett sor és oszlopindexelő vektorban cserélünk.
Nem kell az elimináció előtt ellenőrizni, hogy megoldható-e az egyenletrendszer, mert az algoritmus közben eldől a megoldhatóság.
Ha nem tudjuk az $a_{k,k}^{(k-1)} = 0$ elemet cserélni, mert mindenhol $(k \to n-1)\,0$ maradt, és az utolsó oszlopban is csak 0 marad, akkor végtelen sok megoldás van, egyébként, ha az utolsó oszlopban maradnak nemnulla elemek, akkor nincs megoldás.

**4. A GE alkalmazásai: determináns számítása, azonos mátrixú lineáris egyenletrendszerek megoldása, mátrix inverz számítás. A GE felírása speciális mátrix szorzásokkal. Kapcsolata az LU felbontással.**

**Determináns számítása:**

A Gauss-elimináció után a determináns a főátlóbeli elemek szorzatával számítható, mivel a Gauss-elminiáció átalakításai determinánstartóak.
$\det(\mathbf{A}) = a_{11}^{(0)} \cdot a_{22}^{(1)} \dots a_{nn}^{(n-1)} \cdot (-1)^t$, ahol a $t$ a sor- és oszlopindexek összes inverziószáma.

**Azonos mátrixú lineáris egyenletrendszerek megoldása:**

$\mathbf{A}\mathbf{x}_1 = \mathbf{b}_1, \mathbf{A}\mathbf{x}_2 = \mathbf{b}_2, \mathbf{A}\mathbf{x}_3 = \mathbf{b}_3$
$\mathbf{A}|\mathbf{b}_1|\mathbf{b}_2|\mathbf{b}_3 \xrightarrow[\text{visszahelyettesítés}]{\text{Gauss-elimináció}} \mathbf{I}|\mathbf{x}_1|\mathbf{x}_2|\mathbf{x}_3.$

**Inverz számítása:**

$\mathbf{A}\mathbf{x} = \mathbf{I} \quad (\mathbf{x} = \mathbf{A}^{-1}$ egyenletből).
$\mathbf{A}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$, így $\mathbf{A}\mathbf{x}_i = \mathbf{e}_i$.
$\mathbf{A}|\underbrace{\mathbf{e}_1|\mathbf{e}_2|\dots|\mathbf{e}_n}_{\mathbf{I}} \xrightarrow[\text{visszahelyettesítés}]{\text{Gauss-elimináció}} \mathbf{I}|\underbrace{\mathbf{x}_1|\mathbf{x}_2|\dots|\mathbf{x}_n}_{\mathbf{x}=\mathbf{A}^{-1}}.$
(a bol oldai oszlopcsere az eredményben sorcserét jelent!)

**A Gauss-elimináció felírása mátrix szorzásokkal:**

$$\mathbf{L_k} = \begin{bmatrix} 1 & & & & & \\ 0 & \ddots & & 0 & & \\ 0 & 0 & \ddots & & & \\ 0 & 0 & -l_{k+1,k} & \ddots & & \\ 0 & 0 & \vdots & 0 & \ddots & \\ 0 & 0 & -l_{n,k} & 0 & 0 & 1 \end{bmatrix} = \mathbf{I} - \mathbf{l}_k \mathbf{e}_k^\top, \text{ ahol } \mathbf{l}_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ l_{k+1,k} \\ \vdots \\ l_{n,k} \end{bmatrix}, \text{ és } l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}.$$

$\mathbf{L}_k \mathbf{A}^{(k-1)} = \mathbf{A}^{(k)}$ a Gauss-elimináció egy lépése, ha

- az első k db sor nem változik,
- i. sor = i. sor - $l_{ik}\cdot$ k. sor.

$\mathbf{L}_{n-1} \cdot \mathbf{L}_{n-2} \cdots \mathbf{L}_2 \cdot \mathbf{L}_1 \cdot \mathbf{A} =: \mathbf{U}$ felsőháromszög-mátrix.

Bizonyítás: $\mathbf{A}^{(k)}$ első $k$ darab sora nem változik ✓
$i > k$-ra: $\mathbf{A}^{(k-1)}$ $i$. sora $- \mathbf{l}_{ik} \cdot \mathbf{A}^{(k-1)}$ $k$. sora.

**Kapcsolata az LU-felbontással:**

Az **LU**-felbontás:

$$\mathbf{A} = \mathbf{L}_1^{-1}\mathbf{L}_2^{-1} \dots \mathbf{L}_{n-1}^{-1} \cdot \mathbf{U} = \mathbf{L}\mathbf{U}$$

Tétel:

$$\mathbf{L}_k^{-1} = \mathbf{I} + \mathbf{l}_k \mathbf{e}_k^\top.$$

Bizonyítás:
$$\mathbf{L}_k \cdot \mathbf{L}_k^{-1} = (\mathbf{I} - \mathbf{l}_k\mathbf{e}_k^\top)(\mathbf{I} + \mathbf{l}_k\mathbf{e}_k^\top) = \mathbf{I} - \mathbf{l}_k(\underbrace{\mathbf{e}_k^\top\mathbf{l}_k}_{0})\mathbf{e}_k^\top = \mathbf{I}$$

Tétel :
$$\mathbf{L}_1^{-1}\mathbf{L}_2^{-1}\dots\mathbf{L}_{n-1}^{-1} = \mathbf{I} + \mathbf{l}_1\mathbf{e}_1^\top + \mathbf{l}_2\mathbf{e}_2^\top + \dots + \mathbf{l}_{n-1}\mathbf{e}_{n-1}^\top.$$

Bizonyítás: Teljes indukcióval.

· $k = 1$ : ✓ (előző tétel)

· Tegyük fel, hogy $k < n-1$-re igaz.

· ekkor $(k+1)$-re:

$$(\mathbf{L}_1^{-1}\dots\mathbf{L}_k^{-1}) \cdot \mathbf{L}_{k+1}^{-1} = (\mathbf{I} + \mathbf{l}_1\mathbf{e}_1^\top + \dots + \mathbf{l}_k\mathbf{e}_k^\top) \cdot (\mathbf{I} + \mathbf{l}_{k+1}\mathbf{e}_{k+1}^\top) =$$
$$= \mathbf{I} + \mathbf{l}_1\mathbf{e}_1^\top + \dots + \mathbf{l}_k\mathbf{e}_k^\top + \mathbf{l}_{k+1}\mathbf{e}_{k+1}^\top + \mathbf{l}_1\underbrace{\mathbf{e}_1^\top\mathbf{l}_{k+1}}_{0}\mathbf{e}_{k+1}^\top + \dots + \mathbf{l}_k\underbrace{\mathbf{e}_k^\top\mathbf{l}_{k+1}}_{0}\mathbf{e}_{k+1}^\top =$$
$$= \mathbf{I} + \mathbf{l}_1\mathbf{e}_1^\top + \dots + \mathbf{l}_{k+1}\mathbf{e}_{k+1}^\top. \quad \square$$

Következmény: Ha a Gauss-elimináció elvégezhető sor és oszlopcsere nélkül, akkor $\exists \mathbf{A} = \mathbf{LU}$ alakú felbontás, ahol $\mathbf{L} \in \mathcal{L}^{(1)}, \mathbf{U} \in \mathcal{U}$. Igaz a megfordítás is.

## 5. Az LU felbontás, tétel a ∃!-ről. A főminorok és az LU felbontás kapcsolata. L és U elemeinek meghatározásának menete, sorrendek az elemek kifejezésére. Műveletigénye.

$\mathbf{A} = \mathbf{LU}$, ahol
$$\mathbf{L} = \begin{bmatrix} 1 & & 0 \\ \cdot & 1 & \\ \cdot & \cdot & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \cdot & \cdot & \cdot \\ & \cdot & \cdot \\ 0 & & \cdot \end{bmatrix}$$

Vagyis $\mathbf{L}$ alsóháromszög mátrix, diagonálisában 1-esek, $\mathbf{U}$ pedig felsőháromszög mátrix.
$\mathbf{Ax} = \mathbf{b} \Leftrightarrow \mathbf{L}(\underbrace{\mathbf{Ux}}_{\mathbf{y}}) = \mathbf{b} \Leftrightarrow$

(1) $\mathbf{Ly} = \mathbf{b} \to \mathbf{y}$,

(2) $\mathbf{Ux} = \mathbf{y} \to \mathbf{x}$.

Tétel: Ha a Gauss-elimináció elvégezhető sor és oszlopcsere nélkül, akkor $\exists\, \mathbf{LU} = \mathbf{A}$.

Tétel: Ha $\det(\mathbf{A}) \neq 0 \Rightarrow \exists!\, \mathbf{LU} = \mathbf{A}$.

Bizonyítás: Indirekt tegyük fel hogy ∃ olyan különböző

· $\mathbf{L}_1\mathbf{U}_1$ és

· $\mathbf{L}_2\mathbf{U}_2$,

hogy
$\mathbf{A} = \mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$. Mind a 4 mátrix intvertálható, hiszen

$$\det(\mathbf{L}_1) = \det(\mathbf{L}_2) = 1$$

és a det szorzástétel miatt

$$\det(\mathbf{U}_1) \neq 0, \quad \det(\mathbf{U}_2) \neq 0,$$

így

$$\mathbf{L}_2^{-1}\mathbf{L}_1 = \mathbf{U}_2\mathbf{U}_1^{-1}.$$

Tudjuk, hogy két alsó háromszögmátrixot megszorozva alsó háromszögmátrixot és két felső háromszögmátrixot megszorozva felső háromszögmátrixot kapunk, és invertálható háromszög mátrix szorzata is ugyan olyan háromszög lesz. Tehát $\mathbf{L}_2^{-1}\mathbf{L}_1$ alsó háromszögmátrixot ad, és $\mathbf{U}_2\mathbf{U}_1^{-1}$ felső háromszögmátrixot ad. Így egyenlőség csak akkor áll fent, ha mindkét oldalon diagonális mátrix áll, de akkor a bal mátrix $\mathbf{I}$, így akkor a jobb is $\mathbf{I}$. Tehát

$$\mathbf{L}_2^{-1}\mathbf{L}_1 = \mathbf{I} \Rightarrow \mathbf{L}_2 = \mathbf{L}_1,$$

$$\mathbf{U}_2\mathbf{U}_1^{-1} = \mathbf{I} \Rightarrow \mathbf{U}_1 = \mathbf{U}_2. \quad \square$$

Tétel: Ha az $\mathbf{A}$ mátrix főminoraira: $\mathbf{D}_k \neq 0 \quad (k = 1, \ldots, n-1)$, akkor az $\mathbf{LU} = \mathbf{A}$ felbontás előállítható.

Bizonyítás: Ha $\forall\, \mathbf{D}_k : \det(\mathbf{D}_k) \neq 0$ (ugyanis $\mathbf{D}_k = a_{11}^{(0)} \ldots a_{kk}^{(k-1)}$), úgy a GE elvégezhető, tehát $\exists\, \mathbf{LU}$ felbontás. $\square$

Tétel: Ha az $\mathbf{A}$ mátrix főminoraira $\mathbf{D}_k \neq 0 \quad (k = 1, \ldots, n)$, akkor az $\mathbf{LU}$ felbontás egyértelmű.

Bizonyítás: Ha $\underset{\text{előző tétel}}{\det(\mathbf{A})} \neq 0 \Rightarrow \exists!\, \mathbf{LU} = \mathbf{A}$. $\square$

**L és U elemeinek meghatározása mátrixszorzással:**

Általános képlet:

$$a_{ij} = \sum_{k=1}^{\min(i,j)} l_{ik} \cdot u_{kj}.$$

Az $a_{ij}$ elem $(i, j)$ poziciója határozza meg, hogy $l_{ij}$ vagy $u_{ij}$ elemet számol.

- Felsőháromszög mátrix elemei: $i \leq j$

$$a_{ij} = l_{ii} \cdot u_{ij} + \sum_{k=1}^{i-1} l_{ik} \cdot u_{kj} \Rightarrow$$

$$\Rightarrow u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} \cdot u_{kj} \quad (l_{ii} = 1 \text{ miatt}).$$

- Alsó háromszög mátrix elemei: $i > j$

$$a_{ij} = l_{ij} \cdot u_{jj} + \sum_{k=1}^{j-1} l_{ik} \cdot u_{kj} \Rightarrow$$

$$\Rightarrow l_{ij} = \frac{1}{u_{jj}} \cdot (a_{ij} - \sum_{k=1}^{j-1} l_{ik} \cdot u_{kj}).$$

**Sorrendek az elemek kifejezésére:**

- sorfolytonosan $(i, j)$ szerint,
- oszlopfolytonosan $(i, j)$ szerint,
- parkettás módszer - sor/oszlop felváltva.

$\mathbf{U}$ első sora = $\mathbf{A}$ első sora,
$\mathbf{L}$ első oszlopa = $\mathbf{A}$ első oszlopa$/a_{11}$.

**Az LU-felbontás műveletigénye:**

$u_{ij}$: rögzített $i$-re $(i-1)$ szorzás, $(i-1)$ összeadás $= 2(i-1)$,
$l_{ij}$: rögzített $j$-re $(j-1)$ szorzás, $(j-1)$ öszeadás, 1 osztás $= 2j-1$,

$$\sum_{i=2}^{n}\sum_{j=1}^{n} 2(i-1) + \sum_{j=1}^{n-1}\sum_{i=j+1}^{n-1} (2j+1) = \boldsymbol{\frac{2}{3}n^3 + \mathcal{O}(n^2)}.$$

## 6. Fogalmak: A szimmetrikus, pozitív definit, szigoróan diagonálisan domináns a sorokra illetve oszlopokra, fél sávszélesség, profil, Schur-komplementer. A GE (LU felbontás) megmaradási tételei.

**A szimmetrikus**, ha $\mathbf{A}^\top = \mathbf{A}$.
**A pozitív definit**, ha

$$\forall\, \lambda_i > 0 \quad (i = 1 \ldots n) \Leftrightarrow$$