1. Tétel: Lebegőpontos számok és tulajdonságaik

Legyen $m = \sum_{i=1}^{t} m_i \cdot 2^{-i}$, ahol $t \in \mathbb{N}$, $m_1 = 1$, $m_i \in \{0,1\}$

Ekkor az $a = \pm m \cdot 2^k$ $(k \in \mathbb{Z})$ alakú számot normalizált lebegőpontos számnak nevezzük

m: mantissza

k: karakterisztika $k^- \leq k \leq k^+$

Jelölés: $a \pm [m_1, \dots m_t | k] = \pm 0, m_1 \dots m_t \cdot 2^k$

$$M(t, k^-, k^+) = \left\{ a = \pm 2^k \cdot \sum_{i=1}^{t} m_i \cdot 2^{-i} : \begin{array}{l} k^- \leq k \leq k^+ \\ m_i \in \{0,1\}, m_1 = 1 \end{array} \right\}$$

1. $\frac{1}{2} \leq m < 1$
2. M szimmetrikus 0-ra.
3. $\varepsilon_0 = [100\dots0|k^-] = \frac{1}{2} \cdot 2^{k^-} = 2^{k^- - 1}$
4. $\varepsilon_1 = [100\dots1|1] - [100\dots0|1] = 2^{-t} \cdot 2^1 = 2^{1-t}$
5. $M_\infty = [111\dots1|k^+] = 1.00\dots0 \cdot 2^{k^+} - 0.00\dots01 \cdot 2^{k^+} = (1 - 2^{-t}) \cdot 2^{k^+}$
6. $|M| = 2 \cdot 2^{t-1} \cdot (k^+ - k^- + 1)$

számegyenes?

"Tarosságok"

1. $fl(a+b) = fl(a)$, ahol $b \neq 0$

pl: $a = [1100|1]$, $b = [0001|-4]$ $\Rightarrow$

$$\begin{array}{r} [1100|1] \\ +\,[0001|1] \\ \hline [1100|1] \end{array}$$

2. Asszociativitás nem teljesül: $(a+b)+b \neq a+(b+b)$

pl: $a = [10011|4]$, $b = [10000|-1]$ $\rightarrow$ $[10011|4]$, $[00000|4]$

bal:

$(a+b) = [10011|4]$

$+b = [10011|4]$

jobb: $(b+b)$:

2. Tétel: A hibaszámítás alapjai

abszolút hiba: $|\Delta a| := |A - a|$

relatív hiba: $\delta a := \frac{\Delta a}{A} \approx \frac{\Delta a}{a}$

abszolút hibakorlát: $\Delta_a \geq |\Delta a|$

relatív hibakorlát: $\delta_a \geq |\delta a|$

műveletek:

$$\Delta_{a+b} = \Delta_a + \Delta_b$$

$$\Delta_{a \cdot b} = |b| \cdot \Delta_a + |a| \cdot \Delta_b$$

$$\Delta_{\frac{a}{b}} = \frac{|b| \cdot \Delta_a + |a| \cdot \Delta_b}{b^2}$$

$$\delta_{a \pm b} = \frac{|a| \cdot \delta_a + |b| \cdot \delta_b}{|a \pm b|}$$

$$\delta_{a \cdot b} = \delta_a + \delta_b$$

$$\delta_{\frac{a}{b}} = \delta_a + \delta_b$$

veszélyes: - $\pm$ (abszolút) ha két szám közel azonos nagyságú

- osztás (abszolút) ha a nevező közel nulla

Bizonyítás: 1. $a+b$

$$\Delta(a+b) = (A+B) - (a+b) = (A-a) \pm (B-b) = \Delta a + \Delta b$$

$$|\Delta(a+b)| = |\Delta a + \Delta b| \leq |\Delta a| + |\Delta b| \leq \Delta_a + \Delta_b = \Delta_{a+b}$$

relatív hiba:

$$\frac{\Delta(a+b)}{a+b} = \frac{\Delta a + \Delta b}{a+b} = \frac{a \cdot \delta a + b \cdot \delta b}{a+b}$$

$$\frac{|\Delta(a+b)|}{|a+b|} = \frac{|a \cdot \delta a + b \cdot \delta b|}{|a+b|} \leq \frac{|a| \cdot |\delta a| + |b| \cdot |\delta b|}{|a+b|} \leq \frac{|a| \cdot \delta_a + |b| \cdot \delta_b}{|a+b|} = \delta_{a+b}$$

2. $a \cdot b$

$$\Delta(ab) = AB - ab = AB - Ab + Ab - ab = A(B-b) + b(A-a) = A \cdot \Delta b + b \cdot \Delta a = \cancel{(a+\Delta a) \cdot \Delta b + b)}$$

$$= (a + \Delta a) \cdot \Delta b + b \cdot \Delta a \approx a \cdot \Delta b + b \cdot \Delta a$$ (Δa és Δb elhanyagolható)

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

A relatív hiba:

$$\delta(a+b) = \frac{\Delta(a+b)}{ab} \approx \frac{a \cdot \Delta b + b \cdot \Delta a}{ab} = \frac{\Delta b}{b} + \frac{\Delta a}{a} = \delta b + \delta a$$

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

3. Tétel: A Gauss-elimináció

A cél egy LER-t egyszerűbb alakra hozni.

1. lépés: balról jobbra a főátló alatti elemeket kinullázzuk.

2. lépés: jobbról balra a főátló felett nullázunk, amíg el nem érjük a $[I|x]$-t

Képlet: rekurzív módon

1. 1. egyenlet változatlan. Az i-edik egyenletből $(i=2,3,\dots,n)$ kivonjuk az 1. egyenletet $\left(\frac{a_{i1}^{(0)}}{a_{11}^{(0)}}\right)$-szeresét. $\Rightarrow a_{i1}^{(0)}=0$

$$a_{ij}^{(1)} = a_{ij}^{(0)} - \frac{a_{i1}^{(0)}}{a_{11}^{(0)}} \cdot a_{1j}^{(0)} \qquad \text{ahol } i=2,\dots,n \quad j=2,\dots,n,n+1$$

2.

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} \cdot a_{2j}^{(1)} \qquad \text{ahol } i=3,\dots,n \quad j=3,\dots,n,n+1$$

Általános:

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \cdot a_{kj}^{(k-1)}$$

Alkalmazások:

1. Azonos mátrixú LER $[A|b_1|b_2|b_3] \to$ GE $\to$ visszahely $\to [I|x_1|x_2|x_3]$

2. Determináns kiszámítása Mivel GE det. tartó (nincs sorcsere) $\Rightarrow$ det = főátlóbeli elemek szorzata.

$$\det(A) = \prod_{k=1}^{n} a_{kk}^{(k-1)}$$

3. Inverzmátrix $A \cdot X = I$ egyenlet. $[A|I] \to$ GE $\to$ visszahely $\to [I|A^{-1}]$

b)

4. Tétel: A GE és LU kapcsolata I.

GE: - előkészület: $[A|b]$ mátrix, ahol $a_{in+1} := b$

- k lépés ($k = 1, \dots, n-1$): $a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \cdot a_{kj}^{(k-1)}$ ahol $i = k+1, \dots, n$, $j = k+1, \dots, n, n+1$

- visszahelyettesítés: $x_n = \frac{a_{nn+1}^{(n-1)}}{a_{nn}^{(n-1)}}$ ahol $i = n-1, \dots, 1$

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

<u>főelemkiválasztás</u>: Előfordulhat hogy a LER megoldható, de GE nem működik (főátlóban 0) => sorcserével lehet megoldható

részleges

pontosan: a k-adik lépésben válasszunk egy olyan m indexet, melyre $|a_{mk}^{(k-1)}|$ maximális ($m \in \{k, k+1, \dots, n\}$), majd cseréljük ki a k és m-edik sort.

<u>teljes</u>: A k-adik lépésben válasszunk egy olyan (m_1, m_2) indexpárt, melyre $|a_{m_1 m_2}^{(k-1)}|$ maximális ($m_1, m_2 \in \{k, k+1, \dots, n\}$), majd cseréljük ki a k és m_1-edik sort, valamint k és m_2-edik oszlopot.

Numerikus szempontból jobb a teljes, mert a GE hányadosai pontosabbak lesznek.

5. Tétel GE és LU kapcsolata II.

GE – előkészület: $[A|b]$ mátrix, ahol $a_{in+1} := b$

– k. lépés ($k = 1, \dots, n-1$): $a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \cdot a_{kj}^{(k-1)}$ ahol $i = k+1, \dots, n$, $j = k+1, \dots, n, n+1$

– visszahelyettesítés: $x_n = \frac{a_{nn+1}^{(n-1)}}{a_{nn}^{(n-1)}}$, $x_i = \frac{1}{a_{ii}^{(i-1)}} \left(a_{in+1}^{(i-1)} - \sum_{j=i+1}^{n} a_{ij}^{(i-1)} \cdot x_j \right)$

Feltételek: – Ha főátlóbeli elemek nem nullák $\Rightarrow$ GE működik sor-, oszlopcsere nélkül $\Rightarrow \exists LU$

– Ha $\det(A) \neq 0$ akkor $! \exists LU$

Ha $D_k = 0$ valamely k-ra ($a_{kk}^{(k-1)} = 0$) a GE elakad.

LER LU-val: 1. lépés: megoldjuk $Ly = b$ alsóháromszög egyenletrendszert y-ra

2. lépés: megoldjuk $Ux = y$ felsőháromszög egyenletrendszert x-re.

Így az eredeti $Ax = LUx = Ly = b$ teljesül.

Miért előnyös? Ha sokszor ugyanazt az A mátrixot használjuk különböző b vektorokra, akkor az LU-t csak egyszer kell elvégezni ($\approx \frac{2}{3} n^3$), és utána minden b-re csak két háromszögmátrixú rendszert kell megoldani ($\approx 2n^2$)

6. Tétel LU felbontás direkt módon

LU: A mátrix LU felbontása az az $L \cdot U$ szorzat, amire $A = LU$, hogy $L \in \mathcal{L}_1$ és $U \in \mathcal{U}$

elemenkénti meghatározás: Mivel $A = LU$ ezért $a_{ij} = \sum_{k=1}^{n} l_{ik} \cdot u_{kj}$

képletek: 1. Ha $i \leq j$ (főátló feletti vagy főátló)

Mivel $k > i$ esetén $l_{ik} = 0$ és $l_{ii} = 0$:

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

u_{ij}-t kifejezve:

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

2. Ha $i > j$ (főátló alatti)

Mivel $k > j$ esetén $u_{kj} = 0$:

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

Ha $u_{jj} \neq 0$, akkor:

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

somendek:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 3 & 3 \\ 4 & 4 & 5 & 5 \\ 6 & 6 & 6 & 7 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 3 & 5 & 7 \\ 2 & 4 & 5 & 7 \\ 2 & 4 & 6 & 7 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 3 & 3 \\ 2 & 4 & 5 & 5 \\ 2 & 4 & 6 & 7 \end{bmatrix}$$

sorfolytonos aszlopfolytonos parketta

műveletigény: (egyező a GE-vel): $\frac{2}{3} n^3 + O(n^2)$

részletezve: $\sum_{j=1}^{n} j^2 - \sum_{j=1}^{n} j + \sum_{i=1}^{n-1} s^2 \approx \frac{2}{3} n^3$

7. Tétel A Schur-komplementer

Definíció TFH $A_{11} \in \mathbb{R}^{k \times k}$ invertálható mátrix. Az A mátrix A_{11}-re vonatkozó Schur-komplementere $[A|A_{11}] := A_{22} - A_{21} A_{11}^{-1} A_{12}$ $(n-k)(n-k)$-s mátrix

Azt mutatja meg hogy A_{11}-gyel végzett GE után melyik mátrixon kell folytatni az eliminációt.

Megmaradási tételek

1. $\det(A) \neq 0 \Rightarrow \det([A|A_{11}]) \neq 0$
2. A szimmetrikus $\Rightarrow [A|A_{11}]$ szimmetrikus
3. A pozitív definit $\Rightarrow [A|A_{11}]$ pozitív definit
4. A szig. diag. dom. $\Rightarrow [A|A_{11}]$ szig. diag. dom.
5. $[A|A_{11}]$ fél sávszélessége $\leq$ A fél sávszélessége
6. GE során a profilnál a soronkénti és oszlopkénti nullák az első nem nulla elemig megmaradnak.

Bizonyítás: GE determináns tartó $\Rightarrow \det(A) = \det(A^{(1)}) \neq 0$

$$A^{(1)} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & [A|A_{11}] \end{array}\right]$$

Mivel bal alsó blokk 0, $\det(A^{(1)}) = \det(A_{11}) \cdot \det([A|A_{11}])$

Mivel A invertálható ($\det(A) \neq 0$) és A_{11} is, a szorzat csak akkor nem nulla ha $\det([A|A_{11}]) \neq 0$

$$0 \neq \det(A^{(1)}) = \underbrace{\det(A_{11})}_{\neq 0} \cdot \det([A|A_{11}]) \iff \det([A|A_{11}]) \neq 0$$

$\square$

8. Tétel LDU-felbontás

Fogalma: Az $A \in \mathbb{R}^{n\times n}$ mátrix LDU felbontásának nevezzük az $A = L\cdot D\cdot U$ szorzatot, ha $L \in \mathcal{L}_1$ alsó háromszög-, D diagonális-, és $U \in \mathcal{U}_1$ felső háromszögmátrix

Előállítás: a "hagyományos" $A = \tilde{L}\tilde{U}$-ból kiindulva $\tilde{L}$ egység alsóháromszög mátrix, $\tilde{U}$ pedig általános felsőháromszög. $\tilde{U}$ főátlóeleli elemeiből képezzük D-t:

$$D = \operatorname{diag}(\tilde{u}_{11}, \dots \tilde{u}_{nn})$$

Ekkor $\tilde{U}$ felírható $D\cdot U$ alakban $(U = D^{-1}\tilde{U})$

Az új U főátlójában 1-esek állnak (mivel a sorokat leosztottuk a diagonális elemmel)

Így

$$A = L\tilde{U} = LD\cdot\underbrace{(D^{-1}\tilde{U})}_{U} = LDU$$

"Közvetlen" meghatározás

Ha $i<j$ (felső)
$$u_{ij} = \frac{1}{d_{ii}}\left(a_{ij} - \sum_{k=1}^{i-1} l_{ik}\, d_{kk}\, u_{kj}\right)$$

Ha $i=j$ (diag)
$$d_{ii} = a_{ii} - \sum_{k=1}^{i-1} l_{ik}\, d_{kk}\, u_{ki}$$

Ha $i>j$ (alsó)
$$l_{ij} = \frac{1}{d_{jj}}\left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}\, d_{kk}\, u_{kj}\right)$$

A képleteket az $A = L\tilde{U}$ felbontás „közvetlen" képleteiből kapjuk:

$$\tilde{u}_{ii} \mapsto d_{ii}, \qquad \tilde{u}_{kj} \mapsto d_{kk}\, u_{kj}$$

Műveletigény

$$\frac{n^3}{3} + O(n^2) \qquad \text{(GE fele)}$$

9. Tétel Cholesky-féle LL^T felbontás

Definíció $A \in \mathbb{R}^{n \times n}$ szimmetrikus mátrix Cholesky-felbontásának nevezzük azt a $L \cdot L^T$ szorzatot, ami $A = LL^T$, hogy $L \in \mathbb{R}^{n \times n}$ alsó háromszögmátrix és $l_{ii} > 0$ $(i = 1, \dots, n)$

"Közvetlen" meghatározás:

$A = LL^T$-ből kiindulva A mátrix egy eleme

$$a_{ij} = \sum_{k=1}^{n} l_{ik} (L^T)_{kj} = \sum_{k=1}^{\min(i,j)} l_{ik} l_{jk}$$

Ha $j = i$:

$$a_{ii} = \sum_{k=1}^{i-1} l_{ik}^2 l_{ii}^2 \implies l_{ii} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2}$$

Ha $i > j$:

$$a_{ij} = \sum_{k=1}^{j-1} l_{ik} l_{jk} + l_{ij} l_{jj} \implies l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right)$$

Műveletigény $\frac{1}{3} n^3 + O(n^2)$ $\left(n \text{ gyökvonás}, \frac{n^3}{6} \text{ szorzás}, \frac{n^3}{6} \text{ összeadás} \right)$

Tárigény szimmetrikus $\implies$ elegendő az alsó háromszöget tárolni.

$$\frac{n(n+1)}{2}$$

10. Tétel QR-felbontás Gram-Schmidt ortogonalizációval

Definíció $A \in \mathbb{R}^{n\times n}$ QR felbontása az a $A = QR$ szorzat ahol $Q \in \mathbb{R}^{n\times n}$ ortogonális mátrix, $R \in \mathcal{U}$ pedig felsőháromszögmátrix

ortogonalizációs eljárás

Jelöljük $a_1, \dots, a_n$ az A mátrix oszlopait, $q_1, \dots, q_n$ a Q oszlopait.

Mivel R felső háromszög, $A = QR$ k-adik oszlopa:

$$a_k = \sum_{j=1}^{k} r_{jk} q_j = r_{1k} q_1 + r_{2k} q_2 + \dots + r_{kk} q_k$$

a cél r_{jk} és q_k meghatározása rekurzívan. $(k = 1, \dots, n)$

$\boxed{1.}$ r_{ik} $(i < k)$: Szorozzuk be az egyenletet balról q_i^T-vel $(i < k)$. Mivel Q oszlopai ortonormáltak $(q_i^T q_j = \delta_{ij})$:

$$q_i^T a_k = \sum_{j=1}^{k} r_{jk} (q_i^T q_j) = r_{ik} \cdot 1$$

Tehát $r_{ik} = q_i^T a_k$ (ez a skaláris szorzat: $\langle a_k, q_i \rangle$)

$\boxed{2.}$ q_k vektor és r_{kk}: Rendezzük át az eredeti egyenletet, levonva az ismert tagokat:

$$r_{kk} q_k = a_k - \sum_{j=1}^{k-1} r_{jk} q_j =: v_k$$

a v_k vektor a q_k irányába mutat, hossza pedig r_{kk}. Mivel q_k egységhosszú:

$$r_{kk} = \|v_k\|_2, \quad q_k = \frac{v_k}{r_{kk}}$$

Mi garantálja hogy nem akad el?

akkor nem akad el, ha minden lépésben $r_{kk} \neq 0$, azaz v_k vektor nem nulla.

Feltétel: A oszlopvektorai legyenek lineárisan függetlenek

Ellen az esetben a_k nem állítható elő a korábbi oszlopok lineáris kombinációjaként, így $v_k \neq 0$.

9. Tétel Cholesky-féle LL^T felbontás

Tétel: Ha Q_1 és Q_2 ortogonális mátrixok, akkor Q_1Q_2 ortogonális.

Bizonyítás: Vizsgáljuk az inverzét:

$$(Q_1Q_2)^T(Q_1Q_2) = (Q_2^T Q_1^T)(Q_1Q_2)$$

mátrixszorzás asszociativitását és Q_1 ortogonalitását ($Q_1^T Q_1 = 1$)-t kihasználva:

$$= Q_2^T(Q_1^T Q_1)Q_2 = Q_2^T \cdot 1 \cdot Q_2 = Q_2^T Q_2$$

Mivel Q_2 ortogonális:

$$= 1$$

Tehát a szorzat transzponáltja egyben az inverze is, így Q_1Q_2 ortogonális. □

11. Tétel A Householder transzformáció I.

Definíció: A Householder-mátrix a $H(v) \in \mathbb{R}^{n \times n}$ mátrix, melynek alakja:

$$H(v) = I - 2vv^T$$

ahol $v \in \mathbb{R}^n$ és $\|v\|_2 = 1$ (normált vektor). Az ehhez tartozó leképezés a Householder-transzformáció.

Geometriai tartalom: A $H(v)$ egy tükrözést valósít meg a v vektorra merőleges (azaz v normálvektorú), az origón átmenő $n-1$ dimenziós altérre.

Elemi tulajdonságok:

1. Szimmetrikus ($H^T = H$)

$$(I - 2vv^T)^T = I^T - 2(v^T)^T v^T = I - 2vv^T = H$$

2. Ortogonális és Involutorikus ($H^T = H^{-1} = H$, azaz $H^2 = I$)

$$H \cdot H = (I - 2vv^T)(I - 2vv^T) = I - 2vv^T - 2vv^T + 4v(v^Tv)v^T$$

Mivel $v^Tv = \|v\|^2 = 1$:

$$= I - 4vv^T + 4v(1)v^T = I$$

3. A normálvektort ellentettjébe viszi ($Hv = -v$)

$$Hv = (I - 2vv^T)v = v - 2v(v^Tv) = v - 2v(1) = -v$$

4. Tükrözési sík vektorait helyben hagyja ($\forall y \perp v: Hy = y$)

$$Hy = (I - 2vv^T)y = y - 2v(v^Ty)$$

Mivel $y \perp v$, ezért $v^Ty = 0$, így $Hy = y$

Zéró,

Alkalmazások:

- vektorra (x): $H(v)x = x - 2v(v^T x)$

Először kiszámoljuk a $v^T x$ skaláris szorzatot, majd ezzel és 2-vel megszorozzuk v-t, és kivonjuk x-ből.

Műveletigény: $4n$ művelet (szemben a mátrixszorzás $2n^2$-ével)

- Mátrixra (A):
  - Balról (HA): A transzformációt A minden oszlopvektorára alkalmazzuk (főátló kinullázódik, R előáll)
  - Jobbról (AH): – " " – sorvektorára alkalmazzuk (Q előáll: $Q_k = Q_{k-1} H_k$)

Összesített műveletigények:

- LER megoldása Householderrel: $\frac{4}{3}n^3 + O(n^2)$
- Teljes QR-felbontás: $\frac{8}{3}n^3 + O(n^2)$

12. Tétel: Householder-transzformáció II.

Definíció: Lásd 11 Tétel

Geometriai tartalom: Lásd 11. Tétel (bizonyítás nélkül)

Alkalmazások: Lásd 11. Tétel

LER megoldásához: A cél a $Ax=b$ megoldása

1. Felső háromszög alakra hozás: Householder-transzformációk sorozatával $(H_1, \dots, H_{n-1})$ az A mátrixot felső háromszögmátrixszá alakítjuk (R)

$$H_{n-1} \cdot \dots \cdot H_1 A = R$$

Ugyanezek a transzformációk a b vektorra:

$$Q^T b = H_{n-1} \cdots H_1 b$$

(ahol $Q^T = H_{n-1} \cdots H_1$)

2. visszahelyettesítés Megoldjuk az $Rx = Q^T b$ egyenletrendszert

Műveletigény: $\frac{4}{3} n^3 + O(n^2)$

13 Tétel: Mátrixnormák I.

Vektornorma: Legyen $n \in \mathbb{N}$ rögzített. Az $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ leképezést vektornormának nevezzük, ha teljesíti a következő axiómákat:

1. $\|x\| \geq 0 \quad (\forall x \in \mathbb{R}^n)$
2. $\|x\| = 0 \iff x = 0$
3. $\|\lambda x\| = |\lambda| \cdot \|x\| \quad (\forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}^n)$ – pozitív homogenitás
4. $\|x+y\| \leq \|x\| + \|y\| \quad (\forall x, y \in \mathbb{R}^n)$ – háromszög-egyenlőtlenség

Mátrixnorma: Legyen $n \in \mathbb{N}$ rögzített. Az $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ leképezést mátrixnormának nevezzük ha teljesíti a vektornorma axiómáit és szubmultiplikatív

1. $\|A\| \geq 0 \quad (\forall A \in \mathbb{R}^{n \times n})$
2. $\|A\| = 0 \iff A = 0$
3. $\|\lambda A\| = |\lambda| \cdot \|A\| \quad (\forall \lambda \in \mathbb{R}, \forall A \in \mathbb{R}^{n \times n})$
4. $\|A+B\| \leq \|A\| + \|B\| \quad (\forall A, B \in \mathbb{R}^{n \times n})$
5. $\|AB\| \leq \|A\| \cdot \|B\| \quad (\forall A, B \in \mathbb{R}^{n \times n})$ – szubmultiplikativitás

Indukált mátrixnorma: Legyen $\|\cdot\|_v : \mathbb{R}^n \to \mathbb{R}$ tetszőleges vektornorma. Ekkor a

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v} = \sup_{\|x\|_v = 1} \|Ax\|_v$$

függvényt a $\|\cdot\|_v$ vektornorma által indukált mátrixnormának nevezzük.

Képletek:

1-es norma (oszlopnorma): $\|A\|_1 = \max_{j=1}^{n} \sum_{i=1}^{n} |a_{ij}|$ (max absz. ért. oszlopösszeg)

∞-es norma (sornorma): $\|A\|_\infty = \max_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|$ (max absz. ért. sorösszeg)

2-es norma (spektrálnorma): $\|A\|_2 = \sqrt{\max_i \lambda_i (A^T A)}$ (az $A^T A$ max. sajátértékének gyöke)

Bizt,

Bizonyítás:

1. Nemnegativitás: $\|A\| \geq 0$, mert nemnegatív számok (normák) hányadosainak szuprémuma.

2. Nulla tulajdonság: Ha $A=0$, akkor $\forall x: Ax=0 \Rightarrow$ szuprémum $=0$.

Fordítva, ha szuprémum $=0$, akkor $\forall x: \|Ax\|_v = 0 \Rightarrow Ax=0$

ami csak $A=0$ esetén lehetséges

3. ~~Homogenitás~~ Háromszög egyenlőtlenség $\|A+B\|$ $= \sup \frac{\|(A+B)x\|_v}{\|x\|_v} \leq \sup \frac{\|Ax\|_v + \|Bx\|_v}{\|x\|_v} \leq \sup \frac{\|Ax\|_v}{\|x\|_v} + \sup \frac{\|Bx\|}{\|x\|}$

4. Homogenitás: $\|\lambda A\| = \sup_{x \neq 0} \frac{\|\lambda A x\|_v}{\|x\|_v} = |\lambda| \sup_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v} = |\lambda| \cdot \|A\|$

5. Szubmultiplikativitás: Ha $B \neq 0$, vezessük le a $y = Bx$ jelölést.

$$\|AB\| = \sup_{x \neq 0} \frac{\|A(Bx)\|_v}{\|x\|_v} = \sup_{x \neq 0, Bx \neq 0} \frac{\|A(Bx)\|_v}{\|Bx\|_v} \cdot \frac{\|Bx\|_v}{\|x\|_v}$$

Mivel a szorzat szuprémuma $\leq$ szuprémumok szorzatánál:

$$\leq \left(\sup_{y \neq 0} \frac{\|Ay\|_v}{\|y\|_v} \right) \left(\sup_{x \neq 0} \frac{\|Bx\|_v}{\|x\|_v} \right) = \|A\| \cdot \|B\|$$

14. Tétel Mátrixnormák II.

Mátrixnorma, indukált mátrixnorma: Lásd 13. tétel.

Illeszkedés fogalma: Egy mátrixnormát ($\|\cdot\|$) és egy vektornormát ($\|\cdot\|_v$) illeszkedőnek nevezünk, ha minden x vektorra és A mátrixra teljesül az

$$\|Ax\|_v \leq \|A\| \cdot \|x\|_v \qquad \text{egyenlőtlenség}$$

Bizonyítás: definíciókból következik. (~~hogy~~ Ha) $x \neq 0$:

$$\|A\| = \sup_{y \neq 0} \frac{\|Ay\|_v}{\|y\|_v} \geq \frac{\|Ax\|_v}{\|x\|_v}$$

átrendezve: $\|Ax\|_v \leq \|A\| \cdot \|x\|_v$. Ha $x = 0$ akkor $0 \leq 0$ triviálisan teljesül. $\square$

∞ vektornorma Állítás: $\|A\|_\infty = \max_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|$ (sornorma)

Bizonyítás: 1. Felső becslés: Ax vektor ∞ normájából:

$$\|Ax\|_\infty = \max_i |(Ax)_i| = \max_i \left| \sum_{j=1}^{n} a_{ij} x_j \right|$$

a háromszög egyenlőtlenség miatt:

$$\leq \max_i \sum_{j=1}^{n} |a_{ij}| |x_j|$$

Mivel $|x_j| \leq \|x\|_\infty$ minden j-re:

$$\leq \max_i \left(\sum_{j=1}^{n} |a_{ij}| \right) \cdot \|x\|_\infty = \left(\max_i \sum_{j=1}^{n} |a_{ij}| \right) \cdot \|x\|_\infty$$

Tehát $\|A\|_\infty \leq \max_i \sum_j |a_{ij}|$

2. Válasszuk azt a sort, ahol a sorösszeg maximális (k). Legyen x vektor úgy hogy $x_j = \operatorname{sgn}(a_{kj})$

Ekkor $\|x\|_\infty = 1$. Az Ax vektor k-adik eleme:

$$(Ax)_k = \sum_{j=1}^{n} a_{kj} x_j = \sum_{j=1}^{n} a_{kj} \cdot \operatorname{sgn}(a_{kj}) = \sum_{j=1}^{n} |a_{kj}|$$

Mivel $\|Ax\|_\infty \geq |(Ax)_k|$, ezért:

$$\|Ax\|_\infty \geq \sum_{j=1}^{n} |a_{kj}| = \max_i \sum_{j=1}^{n} |a_{ij}| \qquad \square$$

15. Tétel Frobenius norma

Képlet: $$\|A\|_F := \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

Igazolás: Tudjuk hogy egységmátrix normája mindig 1:

$$\|I\| = \sup_{x\neq 0}\frac{\|Ix\|_v}{\|x\|_v} = \sup_{x\neq 0}\frac{\|x\|_v}{\|x\|_v} = 1$$

Ezzel szemben:

$$\|I\|_F = \sqrt{\sum_{i=1}^{n} 1^2} = \sqrt{n}$$

Ha $n>1$ akkor $\|I\|_F = \sqrt{n} \neq 1 \Rightarrow$ nem indukált norma.

Illeszkedés: Lásd 14. Tétel.

Indukált norma mindig illeszkedik: 14. Tétel

16. LER érzékenysége I.

Perturbáció: Tekintsük az $Ax=b$ LER-t ($A \in \mathbb{R}^{n\times n}$ invertálható).

TFH b helyett $b+\delta b$ perturbált vektorral számolunk. Ekkor a megoldás x helyett $x+\delta x$ lesz.

Perturbált egyenlet: $A(x+\delta x)=b+\delta b$

δb a jobb oldal hibavektora, δx pedig a megoldás hibavektora

Változás mértéke: 1. abszolút hiba: $\|\delta x\| \leq \|A^{-1}\| \cdot \|\delta b\|$

2. relatív hiba: Ha $b \neq 0$: $\dfrac{\|\delta x\|}{\|x\|} \leq \operatorname{cond}(A) \cdot \dfrac{\|\delta b\|}{\|b\|}$

Kondíciószám: Adott $A \in \mathbb{R}^{n\times n}$ invertálható mátrix és $\|\cdot\|$ mátrixnorma esetén

$$\operatorname{cond}(A) := \|A\| \cdot \|A^{-1}\|$$

Tulajdonságok: 1. $\operatorname{cond}(A) \geq 1$: Indukált norma esetén $\|I\|=1$

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \cdot \|A^{-1}\| = \operatorname{cond}(A)$$

2. Bármely $c \neq 0$ skalárra $\operatorname{cond}(cA) = \operatorname{cond}(A)$

$$\operatorname{cond}(cA) = \|cA\| \cdot \|(cA)^{-1}\| = |c| \cdot \|A\| \cdot \left|\frac{1}{c}\right| \cdot \|A^{-1}\| = \|A\| \cdot \|A^{-1}\|$$

3. Ha Q ortogonális, akkor $\operatorname{cond}_2(Q) = \cancel{2}\,1$

Mivel Q ortogonális, $\|Q\|_2 = 1$, és $Q^{-1} = Q^T$ is ortogonális, így $\|Q^{-1}\|_2 = 1$

$$\operatorname{cond}_2(Q) = 1 \cdot 1 = 1$$

17. Tétel LER érzékenysége II.

Perturbáció: Tekintsük $Ax=b$ LER-t. TFH bal oldali A helyett $A+\Delta A$ mátrix van

Ekkor a megoldás x helyett $x+\Delta x$ lesz.

$$(A+\Delta A)(x+\Delta x)=b$$

Változás mértéke: Ha A invertálható, $b\neq 0$, és $\|\Delta A\|\cdot\|A^{-1}\|<1$ akkor

$$\frac{\|\Delta x\|}{\|x\|}\leq\frac{\|A\|\cdot\|A^{-1}\|}{1-\|\Delta A\|\cdot\|A^{-1}\|}\cdot\frac{\|\Delta A\|}{\|A\|}$$

Kondíciószámok: Lásd 16. tétel.

15. 16. | LER megoldhatósága |
17. szóbeli | LER megoldhatósága ||

18. Tétel Iterációs módszerek konvergenciája

Ötlet: Az $Ax=b$ LER-t átalakítjuk a vele ekvivalens $x=Bx+c$ (fixpont) alakra, ahol $B\in\mathbb{R}^{n\times n}$ az átmenetmátrix. Ezután generálunk egy vektorsorozatot az alábbi rekurzióval:

$$x^{(0)}\in\mathbb{R}^n \text{ (tetszőleges kezdővektor}$$

$$x^{(k+1)}=Bx^{(k)}+c, \quad (k=0,1,2,\dots)$$

Ha a sorozat konvergens ($\lim_{k\to\infty} x^{(k)}=x^*$) és a leképezés folytonos, akkor a x^* határérték megoldása a LER-nek.

Összevetés direkt módszerekkel

Direkt módszer	Iterációs módszer
- véges lépésben ad megoldást	- Végtelen sorozat határértéke a megoldás
	- Konvergencia nem garantált

Kontrakció A $\phi:\mathbb{R}^n\to\mathbb{R}^n$ leképezés kontrakció, ha $\exists q\in[0,1)$, hogy $\forall x,y\in\mathbb{R}^n$:

$$\|\phi(x)-\phi(y)\|\le q\cdot\|x-y\|$$

Elégséges feltétel konvergenciára

Tétel: Az $x^{(k+1)}=Bx^{(k)}+c$ konvergens $\iff \rho(B)<1$ ($\rho(B)$ spektrálsugár)

Bizonyítás: $\Leftarrow$: minden mátrixhoz és $\varepsilon>0$-hoz található olyan indukált norma, hogy $\|B\|<\rho(B)+\varepsilon$. Ha $\rho(B)<1$ választhatunk olyan normát amiben $\|B\|<1 \Rightarrow$ az iteráció konvergens.

$\Rightarrow$: TFH $\rho(B)\geq 1$. Ekkor $\exists\lambda$, hogy $|\lambda|\geq 1$. Válasszunk olyan $x^{(0)}$ kezdővektort, hogy a kezdeti hiba $(x^{(0)}-x^*)$ legyen a λ-hoz tartozó sajátvektor. Ekkor a hiba a k. lépésben:

$$x^{(k)}-x^*=B(x^{(k-1)}-x^*)=\dots=B^k(x^{(0)}-x^*)=\lambda^k(x^{(0)}-x^*)$$

Vegyük mindkét oldal normáját:

$$\|x^{(k)}-x^*\|=|\lambda|^k\cdot\|x^{(0)}-x^*\|$$

Mivel $|\lambda|\geq 1$, a $|\lambda|^k$ sorozat nem tart nullához, így a hiba sem csökken nullára.

Ellentmondás $\lightning$

$\square$

19. Tétel Jacobi-iteráció

Mátrixos alak: Legyen $A = L + D + U$

$Ax = b$ egyenlet ekkor $(L+D+U)x = b \Leftrightarrow Dx = -(L+U)x + b$ $\quad / \cdot D^{-1}$ (balról)

$$x = -D^{-1}(L+U)x + D^{-1}b$$

A fixpont alakból kapott iterációs formula

$$x^{(k+1)} = \underbrace{-D^{-1}(L+U)}_{B_J} x^{(k)} + \underbrace{D^{-1}b}_{c_J}$$

Koordinátás alak: $Ax=b$ i-edik sorából fejezzük ki x_i-t

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad \Rightarrow \quad a_{ii}x_i + \sum_{j\neq i} a_{ij}x_j = b_i$$

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

Ebből az iterációs képlet $(i=1,\dots,n)$

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

Reziduumvektoros alak: k-adik lépéshez tartozó maradékvektor: $r^{(k)} := b - Ax^{(k)}$

$$Dx^{(k+1)} = -(L+U)x^{(k)} + b \qquad / -Dx^{(k)}$$

$$Dx^{(k+1)} - Dx^{(k)} = b - (L+D+U)x^{(k)}$$

$$\Leftrightarrow$$

$$Dx^{(k+1)} - Dx^{(k)} = b - Ax^{(k)} = r^{(k)}$$

Új iterált: $\quad x^{(k+1)} = x^{(k)} + D^{-1}r^{(k)}$

Szerepe: Azt mutatja, hogy az új közelítést a régi korrekciával kapjuk. A korrekciót az $r^{(k)}$ D-vel való osztása adja. Numerikus szempontból előnyös lehet, ha a reziduumvektort számoljuk állási feltételnek.

20. Tétel: A csillapított Jacobi-iteráció

Mátrixos: $x^{(k+1)} := (1-\omega)x^{(k)} + \omega x_{jac}^{(k+1)} \Rightarrow x^{(k+1)} := (1-\omega)x^{(k)} + \omega(B_J x^{(k)} + c_J)$

rendezve $x^{(k)}$ szerint: $x^{(k+1)} = \underbrace{\left((1-\omega)I + \omega B_J\right)}_{B_{J(\omega)}} x^{(k)} + \underbrace{\omega c_J}_{c_{J(\omega)}} \cancel{\left(B_J x^{(k)} + c_J\right)}$

Koordinátás alak: súlyozott átlaggal

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

átrendezve:

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

Reziduumvektoros alak: $x^{(k+1)} - x^{(k)} = \omega\left(x_{jac}^{(k+1)} - x^{(k)}\right)$ $\quad \left(x_{jac}^{(k+1)} - x^{(k)} = D^{-1} r^{(k)}\right)$

$$\Downarrow$$

$$x^{(k+1)} = x^{(k)} + \omega D^{-1} r^{(k)}$$

Szerepe: A módszer a reziduumvektor irányába lép, de a lépés hosszát a ω paraméterrel szabályozzuk

- Ha $\omega = 1$: sima Jacobi iteráció
- Ha $\omega < 1$: a lépés kisebb. Segít olyan esetekben ahol az iteráció divergálna.
- Ha $\omega > 1$: a lépés nagyobb. Segíthet konvergenciában.

21. Tétel. Gauss-Seidel iteráció

Mátrixos: $(L+D+U)x=b \Rightarrow (L+D)x^{(k+1)} = -Ux^{(k)}+b \Rightarrow x^{(k+1)} = \underbrace{-(L+D)^{-1}x^{(k)}U}_{B_{GS}} + \underbrace{(L+D)^{-1}b}_{c_{GS}}$

Koordinátás alak:

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

Rezidumvektoros alak:

$$(L+D+U)x=b \Rightarrow (L+D)x^{(k+1)} = -Ux^{(k)}+b \Rightarrow (L+D)(x^{(k+1)}-x^{(k)}) = b - Ux^{(k)} - (L+D)x^{(k)} \Rightarrow$$

$$\Rightarrow (L+D)(x^{(k+1)}-x^{(k)}) = b - Ax^{(k)} = r^{(k)} \Rightarrow$$

$$\Rightarrow (L+D)(x^{(k+1)}-x^{(k)}) = r^{(k)} \Rightarrow \underline{\underline{x^{(k+1)} = x^{(k)} + (L+D)^{-1}r^{(k)}}}$$

Szerepe: Megmutatja, hogy az új iteráltat a régi korrekcióval kapjuk. A korrekciót egy alsóháromszög mátrixú LER megoldásával kapjuk. Ez a korrekció "jobb" mint a Jacobié, mert D^{-1} mellé L-t is használ (több információ).

22. Tétel Gauss-Seidel relaxációs

Mátrixos: $(L+D+U)x=b \Rightarrow (L+D)x=-Ux+b \qquad /\cdot\omega$

$$\omega(L+D)x=-\omega Ux+\omega b \qquad /+\left[(1-\omega)Dx=(1-\omega)Dx\right]$$

$$(D+\omega L)x=\left[(1-\omega)D-\omega U\right]x+\omega b \qquad /\cdot(D+\omega L)^{-1} \text{ balról}$$

$$x=(D+\omega L)^{-1}\left[(1-\omega)D-\omega U\right]x+\omega(D+\omega L)^{-1}b$$

$$x^{(k+1)}=(D+\omega L)^{-1}\left[(1-\omega)D-\omega U\right]x^{(k)}+\omega(D+\omega L)^{-1}b$$

Koordinátás: $x_i^{(k+1)}=(1-\omega)x_i^{(k)}+\omega x_{i,GS}^{(k+1)}$ ahol $x_{i,GS}^{(k+1)}$ hagyományos GS lépés eredménye

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

Reziduumvektoros alak: $(D+\omega L)x^{(k+1)}=\left((1-\omega)D-\omega U\right)x^{(k)}+\omega b$

$$Dx^{(k)}+\omega\underbrace{\left((-D-U)\right)}_{L-A}x^{(k)}+\omega b$$

$$Dx^{(k)}+\omega Lx^{(k)}+\omega\underbrace{\left(-Ax^{(k)}+b\right)}_{r^{(k)}}=(D+\omega L)x^{(k)}+\omega r^{(k)}$$

$$x^{(k+1)}=x^{(k)}+\omega(D+\omega L)^{-1}r^{(k)}$$

Szerepe: A módszer a reziduumvektor irányába lép, de a lépés hosszát és irányát a ω paraméter és a $(D+\omega L)^{-1}$ mátrix módosítja.

- $0<\omega<1$: alulrelaxálás (konvergencia javítása)
- $\omega=1$: Gauss-Seidel
- $1<\omega<2$: túlrelaxálás (gyorsítás)

23. Tétel Richardson-iteráció

Képlet:

$$Ax=b \qquad / \cdot p \quad (p \neq 0)$$

$$pAx = bp$$

$$0 = bp - pAx$$

$$x = x - pAx + pb = (I - pA)x + pb$$

$$x^{(k+1)} = \underbrace{(I - pA)}_{B_R} x^{(k)} + \underbrace{pb}_{c_R}$$

Reziduumvektoros alak:

$$x^{(k+1)} = x^{(k)} + p(\underbrace{b - Ax^{(k)}}_{r^{(k)}})$$

$$x^{(k+1)} = x^{(k)} + p r^{(k)}$$

Jelentősége: Az új közelítést ($x^{(k+1)}$) úgy kapjuk hogy a régiből ($x^{(k)}$) elindulunk a pillanatnyi "hiba" (reziduumvektor) irányába, a p paraméter által megadott lépéshosszal.

Konvergenciatétel:

Tétel: Legyen A szimmetrikus, pozitív definit, sajátértékei $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$.

$$\text{Richardson-iteráció konvergens} \iff 0 < p < \frac{2}{\lambda_n}$$

Optimális paraméter: $p_0 = \dfrac{2}{\lambda_1 + \lambda_n}$

24. Tétel Részleges LU és ILU

Részleges LU: Legyen $S \subseteq \{(i,j) : 1 \le i,j \le n\}$ pozícióhalmaz, amely tartalmazza a főátlót ($\{i,i\} \subseteq S$). Az $A \in \mathbb{R}^{n \times n}$ mátrix S pozícióhalmazú részleges LU-felbontásának nevezzük az $A = LU - R$ felbontást, ahol L alsó-, U felsőháromszögmátrix,

- $l_{ij} = u_{ij} = 0$, ha $(i,j) \notin S$
- $r_{ij} = 0$, ha $(i,j) \in S$

LU algoritmus: A GE módosítása. Csak akkor számítunk elemet, ha pozíciója $\in S$

- Inicializálás: $L = 0, U = 0, R = 0$
- Ciklus: $k = 1 \dots n$:
  - $l_{kk} = 1$
  - U kiszámítása: $j = k \dots n$
    - Ha $(k,j) \in S$: $u_{kj} = a_{kj} - \sum_{\mu=1}^{k-1} l_{k\mu} u_{\mu k}$
  - L kiszámítása: $i = k+1 \dots n$
    - Ha $(i,k) \in S$: $l_{ik} = \frac{1}{u_{kk}} \left(a_{ik} - \sum_{\mu=1}^{k-1} l_{i\mu} u_{\mu k} \right)$

Elégséges feltétel létezésre.

Tétel: Ha A M-mátrix, akkor bármely főátlót tartalmazó szimmetrikus S pozícióhalmaz esetén $\exists!$ az A mátrix S-hez tartozó részleges LU-felbontása.

Állítás: Tétel. Az előállított L, U-ra igaz, hogy $(LU)_{ij} = a_{ij}$ minden $(i,j) \in S$ esetén

Bizonyítás: Definíció szerint $(LU)_{ij} = \sum_{k=1}^{n} l_{ik} u_{kj}$. Mivel L, U háromszögek, összegzés $\min(i,j)$-ig megy

1. Ha $i \le j$: $(LU)_{ij} = \sum_{k=1}^{i} l_{ik} u_{kj} = \sum_{k=1}^{i-1} l_{ik} u_{kj} + l_{ii} u_{ij} \underset{(l_{ii}=1)}{=} \sum_{k=1}^{i-1} l_{ik} u_{kj} + u_{ij}$

2. Ha $i > j$: $(LU)_{ij} = \sum_{k=1}^{j} l_{ik} u_{kj} = \sum_{k=1}^{j-1} l_{ik} u_{kj} + l_{ij} u_{jj}$

Szintet olyan: $l_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right)$. Átrendezve: $l_{ij} u_{jj} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}$

Behelyettesítve:

$$= \sum_{k=1}^{j-1} l_{ik} u_{kj} + \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right) = a_{ij}$$

$\square$

25 Tétel: Nemlineáris egyenletek megoldásai.

Feladatok: egy $f: \mathbb{R} \to \mathbb{R}$ nemlineáris függvény gyökeinek meghatározása, azaz olyan x^* keresése, melyre $f(x^*)=0$. A feladat ekvivalens módon átfogalmazható fixponttproblémává: Keressünk meg egy $\phi: \mathbb{R} \to \mathbb{R}$ függvény fixpontját, azaz azt a x^* értéket, melyre $\phi(x^*)=x^*$.

Megoldás léte; Bolzano-tétel

Bolzano-tétel: Ha $f \in C[a,b]$ és $f(a) \cdot f(b) < 0$ (fv értékek különböző előjelűek) akkor $\exists$ legalább egy $x^* \in (a,b)$ gyök, melyre $f(x^*)=0$

Brouwer-féle fixponttétel:

Tétel: Ha $\phi: [a,b] \to [a,b]$ és $\phi \in C[a,b]$, akkor $\exists x^* \in [a,b]$ fixpont, melyre $\phi(x^*)=x^*$

Bizonyítás: 1. Legyen $g(x) = x - \phi(x)$

2. Mivel ϕ az $[a,b]$ intervallumot önmagába képezi, ezért $\phi(a) \in [a,b]$ és $\phi(b) \in [a,b]$.

Ebből következik:

$$g(a) = a - \phi(a) \leq 0$$
$$g(b) = b - \phi(b) \geq 0$$

így $g(a) \cdot g(b) \leq 0$

3. - Ha $g(a) \cdot g(b) = 0$, akkor $g(a)=0$ vagy $g(b)=0$, tehát a vagy b fixpont.
   - Ha $g(a) \cdot g(b) < 0$, akkor alkalmazzuk a Bolzano tételt g-re: $\exists x^* \in (a,b)$ melyre $g(x^*)=0$, ami ekvivalens azzal, hogy $x^* - \phi(x^*) = 0$, azaz $\phi(x^*) = x^*$ $\square$

Kontrakció fogalma $[a,b]$-n és Banach-féle fixponttétel

Kontrakció: A $\phi: [a,b] \to \mathbb{R}$ leképezés kontrakció az $[a,b]$-n, ha létezik $q \in [0,1)$ közös kontrakciós együttható, hogy minden $x, y \in [a,b]$ esetén:

$$|\phi(x) - \phi(y)| \leq q|x-y|$$

Banach-féle fixponttétel: Ha a $\phi: [a,b] \to [a,b]$ függvény kontrakció az $[a,b]$-n, q együtthatóval akkor:

1. $\exists! x^* \in [a,b]$ fixpont $(x^* = \phi(x^*))$
2. Bármely $x_0 \in [a,b]$ kezdőérték esetén az $x_{k+1} = \phi(x_k)$ sorozat konvergens és x^*-hoz tart.
3. Teljesülnek a hibabecslések:
   - $|x_k - x^*| \leq q^k |x_0 - x^*| \leq q^k (b-a)$
   - $|x_k - x^*| \leq \frac{q^k}{1-q} |x_1 - x_0|$

26. Tétel: Nemlineáris egyenletek megoldása II.

Ötlet: Mindkét esetben az, hogy a függvény görbéjét két ponton átmenő egyenessel közelítjük, és ennek az egyenesnek a zérushelyét választjuk a következő közelítő pontnak. A két ponton $(a, f(a)), (b, f(b))$ átmenő egyenes zérushelyének általános képlete:

$$y - f(a) = \frac{f(a) - f(b)}{a - b}(x - a)$$

zérushely esetén $y = 0$:

$$x = a - \frac{f(a)}{f(a) - f(b)}(a - b)$$

1. Húrmódszer: TFH $f \in C[a,b]$ és $f(a)f(b) < 0$. A sorozat $x_0 = a$, $x_1 = b$ pontokból indul. Minden lépésben az aktuális x_k-t egy korábbi x_s pontot kötünk össze, amelyre a fv-érték előjele ellentétes $f(x_k)$-val.

Képlet:

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

ahol s a legnagyobb olyan index, amelyre $f(x_k) \cdot f(x_s) < 0$

Konvergenciarend: A húrmódszer elsőrendben ($p = 1$) konvergál a gyökhöz.

A hibabecslés: $|x_k - x^*| \leq \frac{1}{M}(M|x_0 - x^*|)^k$

2. Szelőmódszer: A Newton-módszerből származtatható úgy, hogy a deriváltat a fv-görbe két legutóbbi pontján (x_k, x_{k-1}) átmenő szelő meredekségével közelítjük. Itt nem fekszik meg az előjelváltás megőrzése.

Képlet:

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

($k = 1, 2, \ldots$ és x_0, x_1 adott kezdőpontok.)

Konvergenciarend: a szelőmódszer $p = \frac{1 + \sqrt{5}}{2} \approx 1{,}618$ rendben konvergál.

Konvergenciarend fogalma: Módszerek kapcsán a következő hibabecslési alakokat használják a rend definiálására/szemléltetésére: Egy (x_k) sorozat konvergenciája p-rendű, ha nagy k-ra érvényes egy $|x_{k+1} - x^*| \leq C \cdot |x_k - x^*|^p$ alakú becslés (ahol C konstans).

- Elsőrendű: $p = 1$. A hiba minden lépésben konstans szorzóval csökken.
- Másodrendű: $p = 2$. A hiba négyzetes taggal arányosan csökken, ami nagyon gyors konvergenciát jelent.

27. Tétel Nemlineáris egyenletek megoldása III.

Intervallumfelezés: Feltétel: $f \in C[a,b]$, és $f(a)f(b) < 0$ (mint Bolzano-tételnél)

1. Legyen $c_1 = \frac{a+b}{2}$ az intervallum felezőpontja.
2. Vizsgáljuk $f(c_1)$-et:
   - Ha $f(c_1)=0$: megtaláltuk a gyököt. Vége.
   - Ha $f(a)f(c_1)<0$: a gyök a $[a, c_1]$ bal oldali intervallumban van. Új intervallum: $[a_1, b_1] := [a, c_1]$
   - Ha $f(c_1)f(b)<0$: a gyök a $[c_1, b]$ jobb oldali intervallumban van. Új intervallum: $[a_1, b_1] := [c_1, b]$
3. Ismételjük, amíg az intervallum hossza nem csökken kívánt pontosság alá.

Hibabecslés: Minden lépésben felezzük az intervallum hosszát. k-adik lépésben a hossz

$$\frac{b-a}{2^k}$$

Ha a közelítőértéknek (x_k) az intervallum felezőpontját választjuk, a hiba legfeljebb az intervallum fele:

$$|x_k - x^*| \leq \frac{b-a}{2^{k+1}}$$

Tehát a módszer lineárisan konvergens.

Newton módszer: ötlet (geometriai megközelítés): a f görbét egy adott x_k pontban az érintőegyenesével helyettesítjük. A következő közelítőpontnak (x_{k+1}) az érintőegyenes x-tengellyel való metszéspontját választjuk.

Levezetés:

1. f érintőjének egyenlete $(x_k, f(x_k))$ pontban:
$$y - f(x_k) = f'(x_k)\cdot(x - x_k)$$
2. Legyen $y=0$ és $x = x_{k+1}$:
$$0 - f(x_k) = f'(x_k)\cdot(x_{k+1} - x_k)$$
3. Ha $f'(x_k) \neq 0$:
$$-\frac{f(x_k)}{f'(x_k)} = x_{k+1} - x_k$$
4. átrendezve:
$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Többváltozós Newton módszer

Legyen $F: \mathbb{R}^n \to \mathbb{R}^n$ egy nemlineáris vektorfüggvény. Keressük az $x^* \in \mathbb{R}^n$ vektort, melyre $F(x^*)=0$

1. F elsőrendű Taylor-polinomja $x^{(k)}$ körül:

$$F(x) \approx F(x^{(k)}) + F'(x^{(k)}) \cdot (x - x^{(k)})$$

ahol $F'(x^{(k)}) \in \mathbb{R}^{n \times n}$ a Jacobi mátrix

2. Keressük azt a $x^{(k+1)}$ vektort, melyre a linearizált fv 0-t ad.

$$0 = F(x^{(k)}) + F'(x^{(k)}) \cdot (x^{(k+1)} - x^{(k)})$$

3. Rendezve

$$F'(x^{(k)}) \cdot (x^{(k+1)} - x^{(k)}) = -F(x^{(k)})$$

Legyen $s^{(k)} = x^{(k+1)} - x^{(k)}$ irányvektor. Ekkor minden lépésben egy $F'(x^{(k)}) \cdot s^{(k)} = -F(x^{(k)})$ LER-t kell megoldani.

4. Formálisan kifejezve (ha a Jacobi mátrix invertálható)

$$x^{(k+1)} = x^{(k)} - \left(F'(x^{(k)})\right)^{-1} F(x^{(k)})$$

28. Tétel. Polinomok becsléseinek gyöke

Tétel: Legyen $P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ $(a_0 \neq 0)$ valós (vagy komplex) együtthatós polinom.

Legyen M az együtthatók abszolút értékeinek maximuma a főegyütthatót kivételével

$$M = \max_{0 \leq i \leq n-1} |a_i|$$

Ekkor egy polinom bármely x^* gyökére teljesül: $|x^*| < 1 + \frac{M}{|a_n|}$

Tehát a gyökök egy origó középpontú $R = 1 + \frac{M}{|a_n|}$ sugarú körlemezen belül helyezkednek.

Bizonyítás: Indirekt mutassuk be, hogy ha $|x| \geq 1 + \frac{M}{|a_n|}$ akkor $|P(x)| > 0$, tehát x nem lehet gyök.

1. Háromszög egyenlőtlenség "fordított" irányban ($|A-B| \geq |A| - |B|$):

$$|P(x)| = \left| a_n x^n + \sum_{i=0}^{n-1} a_i x^i \right| \geq |a_n x^n| - \left| \sum_{i=0}^{n-1} a_i x^i \right|$$

2.

$$\geq |a_n||x|^n - \sum_{i=0}^{n-1} |a_i||x|^i$$

3. Mivel $M \geq |a_i|$:

$$\geq |a_n||x|^n - M \sum_{i=0}^{n-1} |x|^i$$

4. A szumma egy mértani sorozat összege (hányadosa $|x|$). Mivel feltettük, hogy $|x| \geq 1 - \frac{M}{|a_n|} > 1$:

$$\sum_{i=0}^{n-1} |x|^i = \frac{|x|^n - 1}{|x| - 1} < \frac{|x|^n}{|x| - 1}$$

5. Behelyettesítve:

$$|P(x)| > |a_n||x|^n - M \frac{|x|^n}{|x| - 1} = |x|^n \left(|a_n| - \frac{M}{|x| - 1} \right)$$

6. Ahhoz hogy $|P(x)| > 0$ legyen, a zárójeles tagnak pozitívnak kell lennie.

$$|a_n| - \frac{M}{|x| - 1} > 0 \iff |a_n| > \frac{M}{|x| - 1}$$

átrendezve:

$$|x| - 1 > \frac{M}{|a_n|} \implies |x| > 1 + \frac{M}{|a_n|}$$

Tehát ha $|x| \geq 1 + \frac{M}{|a_n|}$, akkor a polinom értéke biztos nem 0. Így minden gyökének ennél szigorúan kisebbnek kell lennie abszolút értékben.