5. fejezet

Lineáris egyenletrendszerek numerikus megoldása

Gauss-elimináció, szorzatfelbontásos módszerek, iterációs módszerek, legkisebb négyzetes módszerek.

5.1. A lineáris egyenletrendszerek alapproblémája

$A \in \mathbb{R}^{n\times m}, b \in \mathbb{R}^n$. Keressük $x \in \mathbb{R}^m$-et, hogy $Ax = b$. A lineáris algebrából tudjuk, hogy $\exists! A^{-1} \Leftrightarrow |A| \neq 0$. Ekkor a cél $x = A^{-1}b$ meghatározása.

5.1.1. Lemma. *Alsóháromszög-mátrixok (felsőháromszög-mátrixok) szorzata, illetve inverze alsóháromszög-mátrix (felsőháromszög-mátrix).*

5.2. Lineáris algebrai alapfogalmak

5.2.1. Definíció (Pozitív definitás). *Legyen A (szimmetrikus) mátrix, x vektor a megfelelő méretekkel. Ekkor a következők ekvivalensek:*

A pozitív definit	*A pozitív szemidefinit*
$\forall x \neq 0 : x^T A x > 0$	$\forall x : x^T A x \geq 0$
$i = 1, \dots, n : \lvert A_i \rvert > 0$	$\forall \{i_1, \dots, i_k\} \subset \{1, \dots, n\} :$ $\lvert A_{i_1,\dots,i_k} \rvert \geq 0$
$\forall \lambda_i(A) > 0$	$\forall \lambda_i(A) \geq 0$

A negatív definitás analóg módon definiálható.

5.2.2. Definíció (Ortogonális mátrix). *Ortogonális mátrixnak nevezünk egy Q mátrixot, ha oszlopvektorai páronként merőlegesek (skalárszorzatuk 0).*

5.2.3. Definíció (Sajátérték). *Adott A mátrixnak $u \neq 0$ sajátvektora, ha $\exists \lambda \in \mathbb{R}$, hogy $Au = \lambda u$. Ekkor λ-t a mátrix sajátértékének hívjuk.*

5.2.4. Definíció (Spektrálsugár). *Az A mátrix spektrálsugara a sajátértékei minimuma. Jele: $r(A) = \min_{i=1}^{n}(\lambda_i(A))$.*

5.3. Gauss-elimináció, szorzatfelbontásos módszerek

5.3.1. Gauss-elimináció és LU-felbontás

A Gauss-elimináció segítségével a fenti alakra hozott lineáris egyenletrendszerek oldhatók meg. A Gauss-elimináció célja precízen megfogalmazva egy LU szorzatfelbontás képzése, ha ez lehetséges.

A Gauss-elimináció tárgyalásában A alatt mindig $\mathbb{R}^{n\times n}$-beli mátrixot értünk.

5.3.1. Definíció (LU-felbontás). *Egy A mátrix $A = LU$ felbontását LU-felbontásnak nevezzük, ha:*

- *L alsóháromszög-mátrix és főátlójában csupa 1 található,*
- *U felsőháromszög-mátrix.*

5.3.2. Tétel (Az LU-felbontás unicitása és egzisztenciája).

- *Ha egy mátrixnak létezik inverze ($|A| \neq 0$), akkor legfeljebb egy LU-felbontása adható.*
- *Ha*
$$k = 1, \ldots, n-1 : |A_k| \neq 0,$$
akkor A-nak létezik LU-felbontása.

A módszernek háromféle megközelítését tárgyaljuk:

Elemi módszer: az elemi módszerben sorcserékkel és sorok skalárszorosának másik sorhoz való hozzáadásával (egyenlő együtthatók módszere) alakítjuk ki a felsőháromszög-alakot. Ebből L már egyszerűen meghatározható.

Oszloponkénti eliminálás: alkalmas alsóháromszög-mátrixokkal A-t balról szorozva egy-egy oszlop eliminálható.

Elemenkénti eliminálás: alkalmas mátrixokkal balról szorozva elemenként is haladhatunk az eliminálással.

5.3.3. Definíció (Eliminációs mátrixok). *Legyen*

$$\mathbb{R}^{n\times n} \ni L_k(v) = I + v \cdot e_k^T,$$

illetve

$$\mathbb{R}^{n\times n} \ni L_{ik}(\alpha) = I + \alpha \cdot e_i \cdot e_k^T.$$

A k-adik lépésben válasszuk a mátrix aktuális állapotában a k-adik oszlop mínusz egyszeresét a felső k elem „kinullázásával” (jelölje ezt $v_k^{(k)}$) az eliminációs mátrix generáló vektorának. Ekkor igaz a következő állítás:

5.3.4. Tétel (Oszloponkénti elimináció).

$$A = L_1(v_1^{(1)})^{-1}, \dots, L_{n-1}(v_{n-1}^{(n-1)})^{-1} \cdot U,$$

továbbá

$$A = L_1(-v_1^{(1)}), \dots, L_{n-1}(-v_{n-1}^{(n-1)}) \cdot U.$$

Mivel minden L_i alsóháromszög-mátrix, ezért szorzatuk, illetve inverzeik szorzata is az.

5.3.5. Megjegyzés (Elemenkénti elimináció). *Az elemenkénti elimináció a következő formában formalizálható:*

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

$$\alpha ij = -\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}.$$

Gauss-Jordan módszer

A Gauss-elimináció egy variánsa a Gauss-Jordan módszer. A Gauss-elimináció során kialakuló U mátrixból (például felfelé haladó kiküszöböléssel) még elő kell állítani az egyenletrendeszer megoldását adó diagonális mátrixot.

A Gauss-Jordan módszer lényege, hogy egy lépésben nem csak a főátló alatti, hanem az a fölötti elemeket is eliminálja, így $n-1$ lépésben kialakul az egyenletrendszer megoldása. Ehhez csak kicsit kell módosítani az eliminációs mátrixokat.

Hatékonyság. Bár első pillantásra úgy tűnhet, a megoldás előállítására a Gauss-Jordan módszer műveletigényét tekintve fáradtságosabb a Gauss-eliminációnál.

A Gauss-elimináció invariáns tulajdonságai

A Gauss-elimináció során a mátrix egyes tulajdonságai megmaradnak a „jobb alsó" munkamátrix-részeken. Ezek:

- szimmetria,
- pozitív definitás,
- diagonális dominancia,
- profil (sáv).

5.3.2. LDU-felbontás

Az LU-felbontás egy variánsa az LDU-felbontás. Célunk, hogy U-ban is egyesek szerepeljenek a főátlóban, ezeket tehát „kiemeljük".

5.3.6. Definíció (LDU-felbontás). *Legyen $D \in \mathbb{R}^{n\times n}$ diagonális mátrix. Ekkor az LU-felbontás alapján $A = LD(D^{-1}U)$ képezhető úgy, hogy $\widetilde{U} = D^{-1}U$ diagonális elemei 1-esek legyenek.*

5.3.3. LDL-felbontás

5.3.7. Tétel (LDL-felbontás). *Ha A szimmetrikus mátrix, akkor LDU-felbontásában $U = L^T$, azaz a felbontás felírható $A = LDL^T$ alakban.*

5.3.8. Tétel (Cholesky-felbontás). *Ha A még pozitív definit is, akkor LDL-felbontásában D főátlóbeli elemei mind pozitívak. Ekkor képezhető $\sqrt{D}$, hogy $(\sqrt{D})^2 = D$, így $A = (L\sqrt{D})(L\sqrt{D})^T = \widetilde{L}\widetilde{L}^T$. Ezt* Cholesky-felbontás*nak nevezzük.*

5.3.9. Megjegyzés. *A Cholesky-felbontásban L főátlójában már nem biztos, hogy csak* 1*-esek szerepelnek.*

5.3.4. QR-felbontás

Ehhez a szorzatfelbántáshoz keresünk Q ortogonális és R felsőháromszögmátrixot, melyekre $A = QR$.

A felbontásnak háromféle módszere terjedt el:

1. Gram-Schmidt módszer,
2. Householder-módszer,
3. Jacobi-forgatás.

Gram-Schmidt módszer

A módszer először a $Q = AR^{-1}$ kiszámítását tűzi ki célul a Gram-Schmidt ortogonalizációs eljárás segítségével. Lépései

1. tegyük fel, hogy $(x_i)_1^n$ vektorrendszer lineárisan független,
2. Gram-Schmidt eljárással alakítsuk ki a megfelelő ortogonális rendszert: $(y_i)_1^n$, legyen az ebből képzett mátrix Q_0,
3. R_0 főátlójában legyenek egyesek, többi elemét pedig adják az ortogonalizációs eljárás együtthatói, ekkor $A = Q_0R_0$,
4. normáljuk Q_0-t egy D diagonális mátrix segítségével (D elemei Q_0 megfelelő oszlopainak normái), ekkor $A = (Q_0D)(D^{-1}R_0)$, ezzel megkaptuk a felbontást.
5.

5.3.10. Tétel (Gram-Schmidt ortogonalizációs eljárás). *Az eljárás célja egy lineárisan független vektorrendszerhez azonos alteret kifeszítő, ortogonális vektorrendszer konstruálása. Ehhez az*

$$y_i := x_i - \sum_{j=1}^{i-1} c_{ij} y_j$$

képletet használhatjuk, ahol az ortogonális tulajdonságot biztosító együtthatók:

$$c_{ki} := \frac{< y_k, x_i >}{< y_k, x_k >}$$

Householder-módszer

Ez és a következő (forgatásos) módszer a „modernebb" (azaz numerikusan hatékonyabb), $Q^T A = R$ megközelítésből vizsgálja a feladatot.

5.3.11. Definíció (Householder-mátrix). *Legyen* $||v|| = 1$*, ekkor a* v *által generált Householder-mátrix* $H(v) = I - 2vv^T$*.*

5.3.12. Megjegyzés. *A Householder-mátrix szimmetrikus és ortogonális.*

Az LU-felbontáshoz hasonlóan most olyan Householder-mátrixokat keresünk, melyekkel A-t balról szorozva végül felsőháromszög-alakot kapunk (ez lesz R) – az eljárás ebben az eliminációhoz hasonlít.

5.3.13. Tétel (A Householder-módszer generáló vektorai).
$v = \frac{a-te_1}{||a-te_1||}$*, ahol* $t = \pm||a||$ *megfelelő lesz a következő megfeleltetéssel:* a_1 *legyen a mátrix első oszlopa, minden további* a_i *pedig a mátrix aktuális állapotának* i*. oszlopa, a felső* $i-1$ *elemet* 0*-val helyettesítve.*

A végeredményben $Q^T = Q^{-1}$ a Householder-mátrixok szorzatából adódik, és $R = Q^T A$.

Jacobi-forgatás

Legyen $\varphi \in \mathbb{R}$, jelölje $c = cos(\varphi)$-t, $s = sin(\varphi)$-t. Ekkor az eljárás a következő lemmán alapul:

5.3.14. Lemma.

$$U = \begin{pmatrix} c & s \\ s & -c \end{pmatrix}$$

ortogonális, sőt, tetszőleges $n \times n$*-es egységmátrixot csak 4 helyen módosítva* $(U_{jj} = c, U_{ij} = s, U_{ji} = s, U_{ii} = -c)$ *is ortogonális mátrixot kapunk.*

5.3.15. Tétel. *Legyen* φ *olyan, hogy* $\mathrm{tg}(\varphi) = \frac{a_{ij}}{a_{jj}}$*. Ekkor* $U_{ij}(\varphi) \cdot A$ *az* A*-nak épp „kinullázza" az* A_{ij} *elemet.*

5.3.16. Megjegyzés. *A nullázás javasolt sorrendje:*
$(2,1) \to (3,1) \to \ldots \to (n,1)$
$(3,2) \to (4,2) \to \ldots \to (n,2)$
$\ldots \to (n, n-1)$

5.4. Iterációs módszerek

Az iterációs módszerek alapötlete a következő:

$$Ax = b \Leftrightarrow x = Mx + v,$$

, melyből iterációs eljárást kapunk a következő módon:

$$x_{n+1} = Mx_n + v$$

Keressünk azokat az M és v értékeket, melyekkel az iteráció folyamatosan közelíti a keresett x vektort.

5.4.1. Definíció (Additív felbontás). *Legyen adott egy S mátrix: ekkor A-nak az S-re vonatkozó additív felbontása az $A = S - T$ felbontás.*

Ekkor az additív felbontáson alapuló iterációs képletet konstruálhatunk:

$$(S - T)x = b$$

$$Sx = Tx + b$$

$$x = \underbrace{S^{-1}T}_{M}x + \underbrace{S^{-1}b}_{v}$$

5.4.2. Megjegyzés. *Minél kisebb $||M||$, az iteráció konvergenciája annál gyorsabb (feltéve, hogy konvergál).*

Most vezessük be az $A = L + D + U$ *additív* felbontást úgy, hogy L az A-nak az alsó-, U felsőháromszög-része, D pedig A diagonális mátrixa.

5.4.1. Jacobi-iteráció

5.4.3. Definíció (Jacobi-iteráció). *A fenti jelölésekkel az iterációt Jacobi-iterációnak nevezzük, ha $S = D$ és $T = -(L + U)$.*

5.4.4. Tétel. *A definícióból triviálisan következik, hogy*

$$M_J = -D^{-1}(L + U), \text{ s}$$

$$v_J = D^{-1}b.$$

Relaxált alak

Relaxáció alatt a módszer kis módosítását értjük új paraméterek bevezetésével, mellyel a hatékonyság növelhető.

5.4.5. Definíció (Jacobi-iteráció relaxált alakja). *Vezessük be az*

$$x_{k+1} = M(\omega)x_k + v(\omega)$$

iterációs formulát úgy, hogy

$$M_J(\omega) = (1-\omega)I - \omega D^{-1}(L+U) \text{ s}$$

$$v_J(\omega) = \omega D^{-1}b.$$

Ezt a Jacobi-iteráció relaxált alakjának nevezzük.

5.4.6. Tétel (Jacobi-iteráció konvergenciatétele).

- *Ha A diagonálisa nem nullmátrix és A soronként vagy oszloponként szigorúan diagonálisan domináns, akkor a Jacobi-iteráció konvergens.*
- *Ha az $\omega = 1$ eseten a relaxált Jacobi-iteráció konvergens, akkor minden $0 < \omega < 1$ esetben is konvergens.*

5.4.2. Gauss-Seidel iteráció

5.4.7. Definíció (Gauss-Seidel iteráció). *A fenti jelölésekkel az iterációt Gauss-Seidel iterációnak nevezzük, ha $S = D + L$ és $T = -U$).*

5.4.8. Tétel. *A definícióból triviálisan következik, hogy*

$$M_{GS} = -(D+L)^{-1}U, \text{ s}$$

$$v_{GS} = (D+L)^{-1}b.$$

5.4.9. Definíció (Gauss-Seidel iteráció relaxált alakja). *Vezessük be az*

$$x_{k+1} = M(\omega)x_k + v(\omega)$$

iterációs formulát úgy, hogy

$$M_{GS}(\omega) = (D+\omega L)^{-1}[(1-\omega D - \omega U] \text{ s}$$

$$v_{GS}(\omega) = (D+\omega L)^{-1}\omega b.$$

Ezt a Gauss-Seidel iteráció relaxált alakjának nevezzük.

5.4.10. Tétel (Gauss-Seidel iteráció konvergenciatétele).

- *Ha A diagonálisa nem nullmátrix és A soronként vagy oszloponként szigorúan diagonálisan domináns, akkor a Gauss-Seidel iteráció konvergens.*
- *Ha A szimmetrikus és pozitív definit, akkor a Gauss-Seidel iteráció konvergens minden* $0 < \omega < 2$ *paraméterre.*
- *Ha a Gauss-Seidel iteráció konvergens valamely* ω *paraméterre, akkor* $0 < \omega < 2$.

5.4.3. Richardson-iteráció

5.4.11. Definíció (Richardson-iteráció). *Ha A szimmetrikus és pozitív definit,* $0 \neq p \in \mathbb{R}$*, akkor a fenti iterációs képletben legyen*

$$M = M_p = I - pA,$$

$$v = v_p = vp.$$

Ezt az iterációt Richardson-iterációnak nevezzük.

5.4.12. Tétel (A Richardson-iteráció paraméterei). *Jelölje A sajátértékeit rendre* $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$*. A Richardson-iteráció pontosan akkor:*

- *konvergens, ha* $p \in \left(0, \frac{2}{r(A)}\right)$,
- *optimális, ha* $p = \frac{2}{\lambda_1 + \lambda_n}$.

Az optimális paraméterrel az iterációs mátrix spektrálsugara $r_opt = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$.

5.5. Legkisebb négyzetes módszerek

5.5.1. Az általánosított inverz

5.5.1. Definíció (Szinguláris felbontás). *Legyen* $A, D \in \mathbb{C}^{n\times m}$, $U \in \mathbb{C}^{n\times n}$, $V \in \mathbb{C}^{m\times m}$*. Ekkor az* $A = UDV^*$ *felbontást az A szinguláris felbontásának hívjuk, ha U és V unitérek (valós esetben ortogonálisak), valamint D diagonális és nincs negatív eleme.*

5.5.2. Megjegyzés. *Egy mátrixnak általában több szinguláris felbontása létezik.*

5.5.3. Definíció (Általánosított inverz).

- *A D diagonális mátrix D^+ általánosított inverzének mérete legyen azonos D^T méretével, és minden d elemére*

$$d^+ = \begin{cases} \frac{1}{d} & , \text{ ha } d \neq 0 \\ 0 & , \text{ ha } d = 0 \end{cases} .$$

- *Az A mátrix általánosított inverzének nevezzük az $A^+ = VD^+U^*$ mátrixot, ha UDV^* az A egy szinguláris felbontása.*

5.5.4. Definíció (Általánosított megoldás). *Az $Ax = b$ lineáris egyenletrendszer általánosított megoldása $x^+ = A^+b$.*

5.5.5. Tétel (Az általánosított inverz algebrai tulajdonságai).

1. $AA^+ = (AA^+)^T$ *és* $A^+A = (A^+A)^T$ *(szimmetrikus/önadjungált),*
2. $AA^+A = A$,
3. $A^+AA^+ = A^+$.

5.5.6. Tétel (Az általánosított inverz geometriai tulajdonságai).

1. $||Ax^+ - b|| = \min \left\{ ||Ax - b|| \middle| x \in \mathbb{R}^m \right\}$,
2. $x = (x - x^+) + x^+$.

5.5.7. Tétel (Az általánosított inverz teljes rangú mátrixokra). *Legyen $A \in \mathbb{R}^{n \times m}$ teljes rangú mátrix. Ekkor ha:*

- $n \geq m$ *(a mátrix túlhatározott), akkor* $A^+ = (A^TA)^{-1}A^T$,
- $n \leq m$ *(a mátrix alulhatározott), akkor* $A^+ = A^T(AA^T)^{-1}$.

5.5.8. Tétel (Az inverz és az általánosított inverz kapcsolata). *Ha A-nak van inverze, akkor általánosított inverze megeyezik az inverzével.*

5.5.2. Diszkrét legkisebb négyzetek módszere

A legkisebb négyzetek módszerével általunk választott alakú polinomfüggvényt illeszthetünk diszkrét pontokra a síkon úgy, hogy a függvénygörbe és az alappontok közötti távolságok négyzetösszege minimális legyen. A módszer túlhatározott egyenletrendszerek általánosított megoldását használja fel – jó tulajdonsága az általánosított inverz geometriai tulajdonságából adódik.

5.5.9. Definíció (Normálegyenlet). *Legyen $Xc = y$ túlhatározott lineáris egyenletrendszer. Ekkor az $X^T Xc = X^T b$ egyenletrendszert az eredeti mátrixegyenlet normálegyenletének nevezzük.*

5.5.10. Megjegyzés. *Egy túlhatározott egyenletrendszer normálegyenlete mindig megoldható, és eredménye az eredeti egyenlet általánosított megoldása.*

A módszer. Legyenek adottak az $(x_i, y_i)_{i=1}^n$ alappontok és a $(\varphi_i)_{i=1}^m$ alapfüggvények. Ekkor keressük a

$$\varphi(x) = \sum_{i=1}^{m} c_i \varphi_i(x)$$

polinom együtthatóit úgy, hogy az általa leírt görbe négyzetesen legjobban közelítse az alappontokat.

Ehhez képezzük a következő struktúrákat:

$$M = \begin{pmatrix} \varphi_1(x_1) & \dots & \varphi_m(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(x_n) & \dots & \varphi_m(x_n) \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}.$$

A szükséges együtthatókat az $M^T Mc = M^T y$ normálegyenlet megoldásából kapjuk.