Tartalomjegyzék

1. Bevezetés

2. Broyden életrajza

Charles George Broyden (1933. február 3. - 2011. május 20.) szerény családi háttérrel, Angliában született. Édesapja gyári munkásként, édesanyja háztartásbeliként dolgozott. Szülei ennek ellenére a kezdetektől tanulásra biztatták. Charles már gyerekként rengeteget olvasott, az iskolában jól teljesített. Édesapja sajnálatos módon meghalt tuberkulózisban, mikor Charles még csak 11 éves volt. Ez még jobban megnehezítette családi helyzetüket, de édesanyja így is arra biztatta, hogy egyetemre menjen. A King's College London egyetemen szerzett fizikus diplomát 1955-ben. A következő 10 évet az iparban töltötte. Ezután 1965-1967 között az Aberystwyth Egyetemen tanított, majd a University of Essex egyetemen 1967-ben professzor, később a matematika intézet dékánja lett. 1986-ban innen visszavonult, 1990-ben a Bolognai Egyetemen fogadott el professzori kinevezést. Jelentős szerepe volt a kvázi-Newton módszerek kifejlesztésében. A kvázi-Newton módszerek előtt a nemlineáris optimalizálási problémákat gradiens alapú módszerekkel oldották meg. Nemlineáris esetben az ehhez szükséges Hessian mátrix kiszámítása legtöbbször nem praktikus. A kvázi-Newton módszerek kifejlesztésére irányuló munka az 1960-as és 1970-es években zajlott, a nemlineáris optimalizálás egy izgalmas időszakában. A kutatásban részt vett még például Bill Davidon, Roger Flecher és Mike Powell. A kutatások eredménye valódi ipari alkalmazások problémáinak megoldására adott eszközöket.

Az iparban töltött évei alatt Broyden a Davidon-Fletcher-Powell (DFP) módszert adaptálta nemlineáris problémákra. Ez vezetett az 1965-ös klasszikus "A class of methods for solving nonlinear simultaneous equations" cikkéhez a Mathematics of Computation folyóiratban. Ez a munka széleskörűen elismert, mint a 20. század egyik legnagyobb numerikus analízis eredménye. A University of Essex egyetemen a DFP módszer optimalizálásra fókuszált. Feltűnt neki, hogy habár a módszer jól működik, néha furcsa eredményeket produkál. Kerekítési hibákra gyanakodott. A kutatása 1970-ben egy új, továbbfejlesztett módszerhez vezetett. Tőle függetlenül, nagyjából egy időben, Fletcher, Donald Goldfarb, és David Shanno is ugyanerre az eredményre jutott. Ezért az új módszert a neveik kezdőbetűiből BFGS módszernek nevezték el. Más kutatások folytatták a kvázi-Newton módszerek optimalizálását, de a BFGS módszer még ma is a leginkább választott, ha a Hessian mátrix kiszámítása túl költséges. 1981-től Abaffy Józseffel és Emilio Spedicato-val az ABS (Abaffy-Broyden-Spedicato) módszereken dolgozott.

Később a numerikus lineáris algebrára fókuszált, ezen belül is a konjugált gradiens módszerekre és ezek rendszertanára. A szakdolgozat kutatásainak a konjugált gradiens módszerekkel kapcsolatos részét öleli fel. 2011-ben 78 évesen, egy szívroham komplikációiba belehalt.

Feleségével, Joan-nal, 1959-ben házasodtak össze. Négy gyerekük született, a legidősebb, Robbie, sajnos 4 éves korában meghalt. Broyden nagy örömét lelte családjában, gyermekeiben, Christopher-ben, Jane-ben és Nicholas-ban. Szeretett madárlesre járni, zenével foglalkozni, kórusban énekelni, vitorlázni. A helyi közösség és az egyházi közösség aktív tagja volt. Hét unokája született. A legidősebb unokája, Tom, az Oxford egyetemen tanult matematikát, a második legidősebb unokája, Matt, a Warwick Egyetemen tanul matematikát, Ben pedig mérnöknek tanul a Swansea Egyetemen. Így nagyapjuk nyomában járnak.

Köszönet Joan Broyden-nek a életrajzban nyújtott segítségéért.

3. Mátrixok alapvető jellemzése

Ebben a fejezetben a további fejezetekhez szükséges fogalmakat vezetünk be. Definíciókat adunk meg és tételeket mondunk ki. Tisztázzuk a jelöléseket.

3.1. Vektorok és mátrixok

Definíció. A valós vektor a valós számok egy rendezett halmaza.

Definíció. Egy vektor elemeinek a száma a vektor rendje, vagy más szóval a vektor dimenziója.

A vektorokat a szakdolgozatban vastag kisbetűvel jelöljük. Például $\boldsymbol{x} = [x_i]$, ahol x_i a vektor i-edik elemét jelöli. Oszlopvektoron vektort, sorvektoron vektor transzponáltat értünk. Az $\boldsymbol{x}$ vektor transzponáltját $\boldsymbol{x}^T$-vel jelöljük.

Példa. Ha $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, akkor $\boldsymbol{x}^T = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$.

Definíció. Legyenek $\boldsymbol{x} = [x_i]$, $\boldsymbol{y} = [y_i]$ és $\boldsymbol{z} = [z_i]$ n-ed rendű vektorok. Legyen $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$. Ekkor $z_i = x_i + y_i$.

Definíció. Legyen $\boldsymbol{x} = [x_i]$ és $\boldsymbol{y} = [y_i]$ n-ed rendű valós vektor. A belső szorzata, vagy más néven skaláris szorzata a sorvektor $\boldsymbol{x}^T$-nek és az oszlopvektor $\boldsymbol{y}$-nak

$$\boldsymbol{x}^T\boldsymbol{y} = \sum_{i=1}^{n} x_i y_i \,. \tag{3.1}$$

Definíció. Az $\boldsymbol{x}$ és $\boldsymbol{y}$ vektorok egymásra ortogonálisak, ha a belső szorzatuk 0.

Definíció. Legyenek $\boldsymbol{p}_i$ vektorok és y_i skalárok, $i = 1,2,...,n$. A

$$\sum_{i=1}^{n} \boldsymbol{p}_i y_i \tag{3.2}$$

vektor a $\boldsymbol{p}_i$ vektorok lineáris kombinációja.

Definíció. A valós mátrix azonos rendű valós vektorok egy rendezett halmaza.

A mátrixokat a szakdolgozatban vastag nagybetűvel jelöljük. Egy $m{\times}n$-es mátrixot értelmezhetünk úgy, mint m darab sorvektor, vagy mint n darab oszlopvektor. Röviden azt mondjuk, hogy a mátrix sorai és oszlopai. Az $\boldsymbol{A} = [a_{ij}]$ jelölésben a_{ij} az $\boldsymbol{A}$ mátrix i-edik sorának j-edik eleme. A csupa nullából álló mátrixot vagy vektort $\boldsymbol{0}$-val jelöljük. Egy mátrixra akkor mondjuk, hogy ritka, ha viszonylag kevés nullától különböző eleme van. Egy mátrixra akkor mondjuk, hogy kitöltött, ha viszonylag kevés nulla eleme van.

Példa. Legyen $\boldsymbol{A} = [a_{ij}]$ $m{\times}n$-es mátrix. Ekkor $i = 1,2,...,m$ és $j = 1,2,...,n$.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Definíció. Ha egy mátrixnak ugyanannyi sora és oszlopa van, négyzetes mátrixnak nevezzük.

Definíció. A négyzetes mátrix rendje az oszlopainak száma.

Definíció. Az $\boldsymbol{A} = [a_{ij}]$ mátrix transzponáltját $\boldsymbol{A}^T$-vel jelöljük, jelentése $\boldsymbol{A}^T = [a_{ji}]$, azaz a mátrix sorainak és oszlopainak felcserélésével kapott mátrix.

Definíció. Legyen $\boldsymbol{A} = [a_{ij}]$ mátrix. A mátrix diagonális elemei azok az elemek, ahol $i = j$.

Definíció. Az $n{\times}n$-es $\boldsymbol{D} = [d_{ij}]$ mátrix diagonális mátrix, ha $d_{ij} = 0$ minden $i \neq j$ indexre. A diagonális mátrixot $diag(d_i)$-vel jelöljük, ahol d_i az i-edik diagonális elem.

Definíció. A $diag(d_i)$ mátrixot, ahol $d_i = 1$ minden i-re, egységmátrixnak nevezzük, és $\boldsymbol{I}$-vel jelöljük.

Definíció. Legyenek $\boldsymbol{A} = [a_{ij}]$, $\boldsymbol{B} = [b_{ij}]$ és $\boldsymbol{C} = [c_{ij}]$ $m{\times}n$-es mátrixok. Legyen $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$. Ekkor $c_{ij} = a_{ij} + b_{ij}$.

Definíció. Legyen $\boldsymbol{A}$ $m{\times}n$-es mátrix. Legyen $\boldsymbol{x} = [x_i]$ n-ed rendű vektor, $\boldsymbol{y} = [y_i]$ m-ed rendű vektor. Jelölje $\boldsymbol{a}_i^T$ az $\boldsymbol{A}$ mátrix i-edik sorát. Az $\boldsymbol{A}$ mátrix és $\boldsymbol{x}$ vektor mátrix-vektor szorzata $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ és

$$y_i = \boldsymbol{a}_i^T \boldsymbol{x} \,. \tag{3.3}$$

Definíció. Legyen $\boldsymbol{A}$ $m{\times}n$-es mátrix és jelöle $\boldsymbol{a}_i^T$ az $\boldsymbol{A}$ mátrix i-edik sorát. Legyen $\boldsymbol{B}$ $n{\times}p$-s mátrix és jelöle $\boldsymbol{b}_j$ a $\boldsymbol{B}$ mátrix j-edik oszlopát. Az $\boldsymbol{AB}$ mátrixszorzat eredménye a $\boldsymbol{C} = [c_{ij}]$ $m{\times}p$-s mátrix, ahol

$$c_{ij} = \boldsymbol{a}_i^T \boldsymbol{b}_j \,. \tag{3.4}$$

Az $\boldsymbol{AB}$ mátrixszorzat esetén az $\boldsymbol{A}$ mátrix oszlopainak száma meg kell egyezzen a $\boldsymbol{B}$ mátrix sorainak számával, hogy a megfelelő sorvektorok és oszlopvektorok belső szorzatai jól definiáltak legyenek. A mátrixszorzás általában nem kommutatív, $\boldsymbol{AB} \neq \boldsymbol{BA}$. Egy mátrixot az egységmátrixszal szorozva önmagát kapjuk, $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$.

Definíció. Az $\boldsymbol{A}$ mátrix szimmetrikus, ha $\boldsymbol{A} = \boldsymbol{A}^T$. Az $\boldsymbol{A}$ mátrix antiszimmetrikus, ha $\boldsymbol{A} = -\boldsymbol{A}^T$.

Mivel $(\boldsymbol{A}^T)^T = \boldsymbol{A}$ és $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$, ezért $(\boldsymbol{A}^T\boldsymbol{A})^T = \boldsymbol{A}^T(\boldsymbol{A}^T)^T = \boldsymbol{A}^T\boldsymbol{A}$. Azaz az $\boldsymbol{A}^T\boldsymbol{A}$ valós mátrix mindig szimmetrikus.

Definíció. Legyen $\boldsymbol{x}^T = [\boldsymbol{x}_1^T \ \boldsymbol{x}_2^T]$ n-ed rendű vektor, ahol $\boldsymbol{x}_1^T = [x_1, x_2, ..., x_r]$ és $\boldsymbol{x}_2^T = [x_{r+1}, x_{r+2}, ..., x_n]$ és $1 \leq r < n$. Az $\boldsymbol{x}_1$ és $\boldsymbol{x}_2$ vektorok az $\boldsymbol{x}$ vektor részvektorai, vagy partíciói.

Példa. Legyenek $\boldsymbol{x} = [\boldsymbol{x}_1 \ \boldsymbol{x}_2]$, $\boldsymbol{y} = [\boldsymbol{y}_1 \ \boldsymbol{y}_2]$, $\boldsymbol{u}$ n-ed rendű vektorok. Legyen $\boldsymbol{u} = \boldsymbol{x} + \boldsymbol{y}$. Ha $\boldsymbol{x}_1 = [x_1, x_2, ..., x_r]$, $\boldsymbol{x}_2 = [x_{r+1}, x_{r+2}, ..., x_n]$, $\boldsymbol{y}_1 = [y_1, y_2, ..., y_r]$, $\boldsymbol{y}_2 = [y_{r+1}, y_{r+2}, ..., y_n]$, $1 \leq r < n$, akkor $\boldsymbol{u} = [\ \boldsymbol{x}_1 + \boldsymbol{x}_2 \quad \boldsymbol{y}_1 + \boldsymbol{y}_2\ ]$. Igaz az is, hogy $\boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{x}_1^T\boldsymbol{y}_1 + \boldsymbol{x}_2^T\boldsymbol{y}_2$.

Mátrixokat is részmátrixokra particionálhatunk. Ennek nagy jelentősége, hogy nagy mátrixokat egyszerűbben kezelhetünk.

Definíció. Legyen $\boldsymbol{A}$ $m\times n$-es mátrix. Az $\boldsymbol{A}$ mátrixból annak k számú sora ($1 \leq k \leq$ $\leq m-1$) és l számú oszlopa ($1 \leq l \leq n-1$) törlésével előállított $(m-k)\times(n-l)$-es részmátrixot az $\boldsymbol{A}$ mátrix minormátrixának nevezzük.

A műveletttartás a minormátrixokra két particionált mátrix között nem feltétlen jól definiált.

Definíció. Particionált mátrixok egy halmaza egy művelet szerint jól particionált, ha a mátrixok minormátrixai között a művelet jól definiált.

Példa. Az $\boldsymbol{E}_1$ és $\boldsymbol{F}_1$ mátrixok az összeadás szerint jól particionáltak.

$$\boldsymbol{E}_1 = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_{11} \\ a_{21} & a_{22} & a_{23} & b_{21} \\ \hline c_{11} & c_{12} & c_{13} & d_{11} \\ c_{21} & c_{22} & c_{23} & d_{21} \end{array}, \quad \boldsymbol{F}_1 = \begin{bmatrix} \boldsymbol{J} & \boldsymbol{K} \\ \boldsymbol{L} & \boldsymbol{M} \end{bmatrix} = \begin{array}{ccc|c} j_{11} & j_{12} & j_{13} & k_{11} \\ j_{21} & j_{22} & j_{23} & k_{21} \\ \hline l_{11} & l_{12} & l_{13} & m_{11} \\ l_{21} & l_{22} & l_{23} & m_{21} \end{array}$$

A szorzás szerint $\boldsymbol{E}_1$ és $\boldsymbol{F}_1$ nem jól particionáltak, mert a megfelelő minormátrixok szorzata (3.4) szerint nem képezhető.

Példa. Az $\boldsymbol{E}_2$ és $\boldsymbol{F}_2$ mátrixok a szorzás szerint jól particionáltak.

$$\boldsymbol{E}_2 = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_{11} \\ a_{21} & a_{22} & a_{23} & b_{21} \\ \hline c_{11} & c_{12} & c_{13} & d_{11} \\ c_{21} & c_{22} & c_{23} & d_{21} \end{array}, \quad \boldsymbol{F}_2 = \begin{bmatrix} \boldsymbol{J} & \boldsymbol{K} \\ \boldsymbol{L} & \boldsymbol{M} \end{bmatrix} = \begin{array}{c|ccc} j_{11} & k_{11} & k_{12} & k_{13} \\ j_{21} & k_{21} & k_{22} & k_{23} \\ j_{31} & k_{31} & k_{32} & k_{33} \\ \hline l_{11} & m_{11} & m_{12} & m_{13} \end{array}$$

Azonban az összeadás szerint $\boldsymbol{E}_2$ és $\boldsymbol{F}_2$ nem jól particionáltak, mert a megfelelő minormátrixok összege nem képezhető.

Definíció. Legyen $\boldsymbol{A}$ $n\times n$-es mátrix. Legyen $\boldsymbol{A}_{11}$ az $\boldsymbol{A}$ mátrix egy minormátrixa. Ha az $\boldsymbol{A}_{11}$ minormátrix elemei az $\boldsymbol{A}$ mátrix főátlójára szimmetrikusan helyezkednek el, az $\boldsymbol{A}_{11}$ minormátrixot főminormátrixnak nevezzük.

Definíció. Legyen $\boldsymbol{A}$ $n\times n$-es mátrix. Legyen $\boldsymbol{A}$ egy particionálása

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix},$$

ahol $\boldsymbol{A}_{11}$ $r\times r$-es ($r < n$) főminormátrix. Ekkor $\boldsymbol{A}_{11}$ az $\boldsymbol{A}$ mátrix bal felső r-ed rendű sarokminormátrixa.

Példa. Legyen

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Az $\boldsymbol{A}$ mátrix bal felső elsőrendű sarokminormátrixa $[a_{11}]$, a bal felső másodrendű sarokminormátrixa

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

és a bal felső harmadrendű sarokminormátrixa önmaga.

Az alkalmazásokban gyakran szükség van komplex vektorokra, mátrixokra. A belső szorzatot leszámítva a fenti definíciók érvényesek komplex vektorokra és mátrixokra is, azzal a különbséggel, hogy a komplex vektorok elemei komplex számok, a komplex mátrixok elemei komplex vektorok. A belső szorzat általánosabb definíciója megköveteli, hogy egy vektor saját magával vett belső szorzata valós, nem negatív, és csak akkor nulla, ha a vektor nulla. Definiáljuk a belső szorzatot komplex vektorokra is.

Definíció. A $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}$ komplex vektor konjugáltja $\overline{\boldsymbol{z}} = \boldsymbol{x} - i\boldsymbol{y}$.

Definíció. Az $\boldsymbol{A} = [a_{ij}]$ komplex mátrix konjugáltja $\overline{\boldsymbol{A}} = [\overline{a_{ij}}]$.

Definíció. A $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}$ komplex vektor Hermite-féle transzponáltja

$$\boldsymbol{z}^H = \overline{\boldsymbol{z}}^T = \boldsymbol{x}^T - i\boldsymbol{y}^T. \tag{3.5}$$

Definíció. Az $\boldsymbol{A} = \boldsymbol{B} + i\boldsymbol{C}$ komplex mátrix Hermite-féle transzponáltja

$$\boldsymbol{A}^H = \overline{\boldsymbol{A}}^T = \boldsymbol{B}^T - i\boldsymbol{C}^T. \tag{3.6}$$

Definíció. Az $\boldsymbol{A}$ komplex mátrix Hermite-mátrix, ha $\boldsymbol{A} = \boldsymbol{A}^H$.

Példa. Legyen $\boldsymbol{A} = \boldsymbol{B} + i\boldsymbol{C}$ és $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}$. Ekkor $\boldsymbol{A}^H\boldsymbol{z} = (\boldsymbol{B}^T - i\boldsymbol{C}^T)(\boldsymbol{x} + i\boldsymbol{y}) =$ $= \boldsymbol{B}^T\boldsymbol{x} + \boldsymbol{C}^T\boldsymbol{y} + i(\boldsymbol{B}^T\boldsymbol{y} - \boldsymbol{C}^T\boldsymbol{x})$.

A komplex Hermite-mátrixot így a valós szimmetrikus mátrix analógiájára definiáltuk. Hasonlóan a valós szimmetrikus esethez, a komplex esetre is igaz, hogy $(\boldsymbol{A}^H\boldsymbol{A})^H = \boldsymbol{A}^H\boldsymbol{A}$.

Definíció. A $\boldsymbol{w}$ komplex vektor és a $\boldsymbol{z}$ komplex vektor belső szorzata $\boldsymbol{z}^H\boldsymbol{w}$.

A $\boldsymbol{z}^H\boldsymbol{z} = (\boldsymbol{x}^T - i\boldsymbol{y}^T)(\boldsymbol{x} + i\boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{y}$ szorzat nem lehet se komplex, se negatív, és csak akkor nulla, ha $\boldsymbol{z}$ vektor nulla, így a definíció eleget tesz az általánosabb feltételeknek.

3.2. Lineáris függetlenség

A lineáris függetlenség alapvető fogalom. A következő fejezetekben nagyban fogunk annak a következményeire támaszkodni, hogy vektorok lineárisan függetlenek-e, vagy sem.

Definíció. Az $\boldsymbol{a}_i$, $i = 1{,}2, ..., n$ vektorok lineárisan összefüggők, ha léteznek olyan x_i skalárok, amelyek nem mindegyike zérus, és a velük képzett lineáris kombinációra fennáll, hogy

$$\sum_{i=1}^{n} \boldsymbol{a}_i x_i = \mathbf{0}. \tag{3.7}$$

Ellenkező esetben az $\boldsymbol{a}_i$ vektorok lineárisan függetlenek.

Példa. Az $\boldsymbol{a}_1^T = [\ 1 \quad 2 \quad -1\ ]$, $\boldsymbol{a}_2^T = [\ -2 \quad -1 \quad 1\ ]$, $\boldsymbol{a}_3^T = [\ -1 \quad 4 \quad -1\ ]$ vektorok lineárisan összefüggők, mert $3\boldsymbol{a}_1 + 2\boldsymbol{a}_2 - \boldsymbol{a}_3 = 0$.

Példa. Az $\boldsymbol{a}_1^T = [\ 1 \quad 0 \quad 0\ ]$, $\boldsymbol{a}_2^T = [\ 1 \quad 1 \quad 0\ ]$, $\boldsymbol{a}_3^T = [\ -1 \quad 1 \quad 1\ ]$ vektorok lineárisan függetlenek, mert

$$\sum_{i=1}^{n} \boldsymbol{a}_i x_i = \begin{bmatrix} x_1 + x_2 - x_3 \\ x_2 + x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Ebből következik, hogy $x_3 = 0$. Így $x_2 + x_3 = 0$ és $x_1 + x_2 + x_3 = 0$ akkor és csak akkor, ha x_1 és x_2 is nulla.

Definíció. Az $\boldsymbol{A}$ $m{\times}n$-es mátrix oszlopai lineárisan összefüggők, ha létezik olyan n-ed rendű $\boldsymbol{x} \neq \mathbf{0}$ vektor, hogy $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$. Ha nem létezik ilyen vektor, akkor a mátrix oszlopai lineárisan függetlenek.

Ha az $\boldsymbol{A}$ mátrix oszlopai lineárisan összefüggők, abból nem csak az következik, hogy az oszlopok megfelelő lineáris kombinációja nulla, hanem hogy létezik olyan $\boldsymbol{x} \neq \mathbf{0}$ vektor, amely az $\boldsymbol{A}$ mátrix minden sorára ortogonális. Az $\boldsymbol{x}$ vektor nem egyértelmű, mert a skalárral szorzása nem változtat az ortogonalitáson, így $\boldsymbol{x}$ tetszőlegesen méretezhető. Az $\boldsymbol{A}$ mátrix oszlopainak lineáris függetlenségét nem befolyásolja, ha az $\boldsymbol{A}$ mátrix sorait felcseréljük.

Definíció. A lineárisan független n-ed rendű vektorok halmazát, amelyből lineáris kombinációként bármely más n-ed rendű vektor kifejezhető, bázisnak nevezzük.

3.1. Tétel. *Az $n + 1$ darab n-ed rendű vektor lineárisan összefüggő. [1, 29. oldal, Theorem 2.2]*

Bizonyítás. Indukcióval bizonyítunk. Megmutatjuk, hogy ha bármely $(r-1)\times r$-es mátrix oszlopai lineárisan összefüggők, akkor bármely $r\times(r+1)$-es mátrix oszlopai is lineárisan összefüggők. Legyen $\boldsymbol{A}_1$ $r\times(r+1)$-es mátrix és

$$\boldsymbol{A}_1 = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{c}^T & \delta \end{bmatrix},$$

ahol $\boldsymbol{A}$ $(r-1)\times r$-es mátrix, $\boldsymbol{b}$ oszlopvektor, $\boldsymbol{c}^T$ sorvektor és δ skalár. Az $\boldsymbol{A}-\boldsymbol{b}\delta^{-1}\boldsymbol{c}^T$ egy $(r-1)\times r$-es mátrix, és az indukciós feltétel szerint bármely $(r-1)\times r$-es mátrix oszlopai lineárisan összefüggők. Ezért létezik egy olyan $\boldsymbol{x}\neq 0$ vektor, hogy

$$(\boldsymbol{A}-\boldsymbol{b}\delta^{-1}\boldsymbol{c}^T)\boldsymbol{x} = \boldsymbol{0}.$$

Ebből következik, hogy

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{c}^T & \delta \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \delta^{-1}\boldsymbol{c}^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}.$$

Mivel $\boldsymbol{x}\neq 0$, léteznie kell egy olyan $\boldsymbol{y}\neq 0$ vektornak, hogy $\boldsymbol{A}_1\boldsymbol{y}=\boldsymbol{0}$. Tehát $\boldsymbol{A}_1$ sorai is lineárisan összefüggők. □

Következmény. *Az n-ed rendű vektorok bázisa mindig n darab vektor.*

3.2. Tétel. *Legyen r darab lineárisan független vektorunk, amik egy új vektor hozzáadásával lineárisan összefüggővé válnak. Ekkor az új vektor kifejezhető az eredeti r vektorok egyértelmű lineáris kombinációjaként. [1, 30. oldal, Lemma 2.2]*

Bizonyítás. Legyen $\boldsymbol{A}$ $n\times r$-es mátrix és $\boldsymbol{b}$ n-ed rendű vektor. Legyenek az $\boldsymbol{A}$ mátrix oszlopai lineárisan függetlenek. Legyen a $\boldsymbol{b}$ vektor az $\boldsymbol{A}$ mátrix oszlopaival lineárisan összefüggő. Mivel ez összesen $r+1$ darab lineárisan összefüggő vektor, létezik olyan r-ed rendű $\boldsymbol{y}$ vektor és η skalár, hogy $\boldsymbol{A}\boldsymbol{y}+\boldsymbol{b}\eta=\boldsymbol{0}$. A skalár η nem lehet nulla, mert akkor $\boldsymbol{A}$ mátrix oszlopai lineárisan összefüggők lennének, ami ellentmond a hipotézisnek. Így az $\boldsymbol{y}\eta^{-1}$ vektor létezik, és ha $\boldsymbol{x}=\boldsymbol{y}\eta^{-1}$, akkor $\boldsymbol{b}=\boldsymbol{A}\boldsymbol{x}$. Azt kell belátni, hogy az $\boldsymbol{x}$ vektor egyértelmű. Tegyük fel, hogy létezik olyan $\boldsymbol{z}$ vektor, hogy $\boldsymbol{b}=\boldsymbol{A}\boldsymbol{z}$. Ha ezt kivonjuk az előző egyenlőségből, azt kapjuk, hogy $\boldsymbol{0}=\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{z})$. Mivel $\boldsymbol{A}$ mátrix oszlopai lineárisan függetlenek, ez csak akkor lehetséges, ha $\boldsymbol{x}=\boldsymbol{z}$. □

Definíció. Az $\boldsymbol{A}$ mátrix lineárisan független oszlopainak maximum darabszáma az mátrix rangja. Jelölése $r(\boldsymbol{A})$.

3.3. Nemszinguláris mátrixok

Definiáljuk a nemszinguláris mátrixot. Ehhez megvizsgáljuk a mátrix oszlopainak illetve sorainak lineáris függőségét, definiáljuk a mátrix inverzét, és megvizsgáljuk a lineáris függőség és a mátrix inverze közötti kapcsolatot.

3.3. Tétel. *Nem létezik olyan $\boldsymbol{X}$ mátrix, hogy $\boldsymbol{AX} = \boldsymbol{I}$, ha $\boldsymbol{A}$ mátrix sorai lineárisan összefüggők. [1, 30. oldal, Lemma 2.3]*

Bizonyítás. Tegyük fel az ellenkezőjét. Mivel $\boldsymbol{A}$ mátrix lineárisan összefüggő, létezik olyan $\boldsymbol{y} \neq \boldsymbol{0}$ vektor, hogy $\boldsymbol{y}^T\boldsymbol{A} = \boldsymbol{0}^T$. Így $\boldsymbol{y}^T\boldsymbol{AX} = \boldsymbol{0}^T$, de mivel $\boldsymbol{AX} = \boldsymbol{I}$, ez csak akkor teljesül, ha $\boldsymbol{y} = \boldsymbol{0}$. Így ellentmondáshoz jutottunk. $\square$

Definíció. Legyen $\boldsymbol{A}$ négyzetes mátrix. Ha létezik olyan $\boldsymbol{X}$ mátrix, hogy

$$\boldsymbol{AX} = \boldsymbol{I}, \tag{3.8}$$

akkor $\boldsymbol{X}$ az $\boldsymbol{A}$ mátrix jobboldali inverze. Ha létezik olyan $\boldsymbol{Y}$ mátrix, hogy

$$\boldsymbol{YA} = \boldsymbol{I}, \tag{3.9}$$

akkor $\boldsymbol{Y}$ az $\boldsymbol{A}$ mátrix baloldali inverze.

Definíció. Legyen $\boldsymbol{A}$ négyzetes mátrix. Ha létezik olyan $\boldsymbol{X}$ mátrix, hogy

$$\boldsymbol{AX} = \boldsymbol{XA} = \boldsymbol{I}, \tag{3.10}$$

akkor $\boldsymbol{X}$ az $\boldsymbol{A}$ mátrix inverze. Az $\boldsymbol{A}$ mátrix inverzének jelölése $\boldsymbol{A}^{-1}$. Egy mátrix invertálható, ha létezik inverze.

3.4. Tétel. *Legyen $\boldsymbol{A}$ n-ed rendű négyzetes mátrix. Ha az $\boldsymbol{A}$ mátrix oszlopai lineárisan függetlenek, létezik olyan $\boldsymbol{X}$ mátrix, hogy $\boldsymbol{AX} = \boldsymbol{I}$. [1, 30. oldal, Theorem 2.3]*

Bizonyítás. Legyen $\boldsymbol{A}$ $n \times n$-es négyzetes mátrix, és legyenek az oszlopai lineárisan függetlenek. Legyenek $\boldsymbol{A}_i$ $(i = 1,2, ..., n-1)$ $(n-1)\times n$-es mátrixok. Az $\boldsymbol{A}_i$ mátrix mindig legyen az $\boldsymbol{A}$ mátrix i-edik sorának elhagyásával kapott mátrix. Azaz $\boldsymbol{A}_i$ j-edik sora az $\boldsymbol{A}$ mátrix j-edik sora, ha $1 \leq j \leq i-1$, de a $(j+1)$-edik sora, ha $i \leq j \leq n-1$. A 3.1-es tételből következik, hogy mindig létezik olyan $\boldsymbol{x}_i \neq \boldsymbol{0}$, hogy $\boldsymbol{A}_i\boldsymbol{x}_i = \boldsymbol{0}$. Mivel létezik olyan $\boldsymbol{x}_i$, hogy $\boldsymbol{A}_i\boldsymbol{x}_i = \boldsymbol{0}$, de $\boldsymbol{Ax}_i \neq \boldsymbol{0}$, így az $\boldsymbol{Ax}_i$ vektornak egyedül az i-edik eleme nem nulla. Minden i-re választhatjuk $\boldsymbol{x}_i$-t úgy, hogy $\boldsymbol{Ax}_i = \boldsymbol{e}_i$ legyen, ahol $\boldsymbol{e}_i$ az n-ed rendű egységmátrix i-edik oszlopa. Legyen $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n]$, így $\boldsymbol{AX} = \boldsymbol{I}$. $\square$

3.5. Tétel. *Ha egy négyzetes mátrix oszlopai lineárisan függetlenek, akkor jobboldali inverze egyértelmű. [1, 31. oldal, Theorem 2.4]*

Bizonyítás. Legyen $\boldsymbol{A}$ négyzetes mátrix, és legyenek az oszlopai lineárisan függetlenek. Tegyük fel, hogy az $\boldsymbol{A}$ mátrix jobboldali inverze nem egyértelmű, azaz léteznek olyan $\boldsymbol{X} \neq \boldsymbol{Y}$ mátrixok, hogy $\boldsymbol{AX} = \boldsymbol{I}$ és $\boldsymbol{AY} = \boldsymbol{I}$. Vonjuk ki a két egyenlőséget egymásból, így kapjuk, hogy $\boldsymbol{A}(\boldsymbol{X} - \boldsymbol{Y}) = \boldsymbol{0}$. Ez azonban az $\boldsymbol{A}$ mátrix oszlopainak lineáris függetlensége miatt csak akkor lehetséges, ha $\boldsymbol{X} = \boldsymbol{Y}$. □

3.6. Tétel. *Ha egy négyzetes mátrixnak létezik egyértelmű jobboldali inverze, akkor az megegyezik a baloldali inverzével. [1, 31. oldal, Theorem 2.5]*

Bizonyítás. Legyen $\boldsymbol{A}$ négyzetes mátrix, és legyenek az oszlopai lineárisan függetlenek. Legyen $\boldsymbol{AX} = \boldsymbol{I}$. Az egyenlőséget jobbról $\boldsymbol{A}$ mátrixszal szorozva kapjuk, hogy $\boldsymbol{AXA} = \boldsymbol{A}$, átrendezve $\boldsymbol{A}(\boldsymbol{XA} - \boldsymbol{I}) = \boldsymbol{0}$. Ebből az $\boldsymbol{A}$ mátrix oszlopainak lineáris függetlensége miatt következik, hogy $\boldsymbol{XA} = \boldsymbol{I}$. □

Definíció. Legyen $\boldsymbol{A}$ négyzetes mátrix. Ha a következő ekvivalens állítások teljesülnek, akkor az $\boldsymbol{A}$ mátrix nemszinguláris, egyébként szinguláris.

1. $\boldsymbol{A}$-nak vannak lineárisan független oszlopai.
2. $\boldsymbol{A}$-nak vannak lineárisan független sorai.
3. $\boldsymbol{A}$ invertálható.

3.7. Tétel. *Legyen $\boldsymbol{A}$ négyzetes mátrix, és legyen*

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{0} & \boldsymbol{A}_{22} \end{bmatrix},$$

ahol $\boldsymbol{A}_{11}$ és $\boldsymbol{A}_{22}$ minormátrixok is négyzetes mátrixok. Ekkor az $\boldsymbol{A}$ mátrix akkor és csak akkor nemszinguláris, ha $\boldsymbol{A}_{11}$ és $\boldsymbol{A}_{22}$ minormátrixok nemszingulárisak. [1, 33. oldal, Lemma 2.4]

Bizonyítás. Legyenek $\boldsymbol{A}_{11}$ és $\boldsymbol{A}_{22}$ minormátrixok nemszingulárisak. Ekkor

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\ \boldsymbol{0} & \boldsymbol{A}_{22}^{-1} \end{bmatrix},$$

tehát az $\boldsymbol{A}$ mátrix nemszinguláris.

Legyen $\boldsymbol{A}_{\boldsymbol{11}}$ minormátrix szinguláris. Ekkor mindig létezik olyan $\boldsymbol{x} \neq \boldsymbol{0}$ vektor, hogy $\boldsymbol{A}_{11}\boldsymbol{x} = \boldsymbol{0}$. Ekkor az $\boldsymbol{A}$ mátrix szingularitása csak $\boldsymbol{A}_{11}$ minormátrixtól függ, mert

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{0} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}.$$

Ezért, ha $\boldsymbol{A_{11}}$ minormátrix szinguláris, akkor az $\boldsymbol{A}$ mátrix szinguláris. Hasonlóan, ha $\boldsymbol{A}_{22}$ minormátrix szinguláris, mindig létezik olyan $\boldsymbol{y} \neq \boldsymbol{0}$ vektor, hogy $\boldsymbol{y}^T\boldsymbol{A}_{22} = \boldsymbol{0}^T$. □

Most definiálunk és megvizsgálunk néhány elméleti vagy gyakorlati szempontból fontos nemszinguláris mátrixot.

Definíció. Az ortogonális mátrix olyan valós mátrix, melynek az inverze a transzponáltja.

Definíció. Az $\boldsymbol{U} = [u_{ij}]$ négyzetes mátrix felső háromszögmátrix, ha $u_{ij} = 0$ minden $i > j$-re.

Definíció. Az $\boldsymbol{L} = [l_{ij}]$ négyzetes mátrix alsó háromszögmátrix, ha $l_{ij} = 0$ minden $i < j$-re.

Definíció. Az alsó háromszögmátrix alsó egység háromszögmátrix, ha minden diagonális eleme egy.

Definíció. A felső háromszögmátrix felső egység háromszögmátrix, ha minden diagonális eleme egy.

3.8. Tétel. *Az $\boldsymbol{U} = [u_{ij}]$ felső háromszögmátrix akkor és csak akkor nemszinguláris, ha $u_{ii} \neq 0$ minden i-re. [1, 34. oldal, Theorem 2.6]*

Bizonyítás. Legyen $\boldsymbol{U}$ n-ed rendű felső háromszögmátrix. Legyen $\boldsymbol{U}_i$ a bal felső i-ed rendű sarokminormátrixa $\boldsymbol{U}$-nak., azaz $\boldsymbol{U} = \boldsymbol{U}_n$. Legyen

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_{i-1} & \boldsymbol{v}_i \\ \boldsymbol{0} & u_{ii} \end{bmatrix}, \quad 2 \leq i \leq n$$

és $\boldsymbol{v}_i^T = [u_{1i,}u_{2i}, ..., u_{(i-1)i}]$. A 3.7-es tételből következik, hogy $\boldsymbol{U}$ akkor és csak akkor nemszinguláris, ha minden $\boldsymbol{U}_i$ sarokminormátrix nemszinguláris. Ha $u_{ii} \neq 0$ bármely $1 \leq i \leq n$ esetén, akkor $\boldsymbol{U}_1, \boldsymbol{U}_2, ..., \boldsymbol{U}_n$ sarokminormátrixok nemszingulárisak, azaz $\boldsymbol{U}$ nemszinguláris. □

3.9. Tétel. *Négyzetes mátrixok szorzata akkor és csak akkor nemszinguláris, ha a szorzat minden tényezője nemszinguláris.*

Bizonyítás. Legyenek $\boldsymbol{A}$, $\boldsymbol{B}$ és $\boldsymbol{C}$ négyzetes mátrixok. Legyen $\boldsymbol{C} = \boldsymbol{AB}$. Tegyük fel, hogy $\boldsymbol{B}$ és $\boldsymbol{C}$ mátrixok nemszingulárisak, de $\boldsymbol{A}$ mátrix szinguláris. Mivel $\boldsymbol{C}$ nemszinguláris, létezik inverze, és $\boldsymbol{CC}^{-1} = \boldsymbol{I}$. Ebből következik, hogy $(\boldsymbol{AB})\boldsymbol{C}^{-1} = \boldsymbol{I}$. De ez csak akkor lehetséges, ha $\boldsymbol{C}^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, amivel ellentmondáshoz jutottunk. □

Definíció. Az $\boldsymbol{A}$ valós mátrix pozitív definit, ha $\boldsymbol{A}$ szimmetrikus, és $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$ minden $\boldsymbol{x} \neq \boldsymbol{0}$-ra.

3.10. Tétel. *Legyen $\boldsymbol{A}$ $m\times n$-es valós mátrix lineárisan független oszlopokkal. Ekkor $\boldsymbol{A}^T\boldsymbol{A}$ pozitív definit. [1, 34. oldal, Lemma 2.5]*

Bizonyítás. Legyen $\boldsymbol{A}$ $m\times n$-es valós mátrix lineárisan független oszlopokkal. Az $\boldsymbol{A}^T\boldsymbol{A}$ mátrix szimmetrikus. Legyen $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, így $\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^T\boldsymbol{y} > 0$, akkor és csak akkor, ha $\boldsymbol{y} \neq \boldsymbol{0}$. Mivel az $\boldsymbol{A}$ mátrix oszlopai lineárisan függetlenek, $\boldsymbol{y}$ akkor és csak akkor nulla, ha $\boldsymbol{x}$ is nulla. Az $\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}$ akkor és csak akkor nulla, ha $\boldsymbol{x} = \boldsymbol{0}$. □

3.4. Vektornormák és mátrixnormák

Gyakran akarunk vektorokat vagy mátrixokat a nagyságuk alapján összehasonlítani. Például ha egy iteratív eljárással közelítünk egy vektorhoz, a hibavektor (azaz a vektor és a közelítő vektor közötti különbség) nagysága jó ha gyorsan csökken. A norma a vektorokhoz és a mátrixokhoz egy skalárt rendel.

Definíció. Az $\|\boldsymbol{x}\|$ skalár az $\boldsymbol{x}$ vektor normája, ha kielégíti a következő három feltételt.

1. $\|\boldsymbol{x}\| = 0$, ha $\boldsymbol{x} = \boldsymbol{0}$, egyébként $\|\boldsymbol{x}\| > 0$.
2. $\|\boldsymbol{x}\theta\| = \|\boldsymbol{x}\|\,|\theta|$, ahol θ skalár.
3. $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

Definíció. Az $\boldsymbol{x}$ vektor l_1, l_2, l_∞ normáinak definíciója:

1. $\|\boldsymbol{x}\|_1 = \sum_i |x_i|$ az l_1 norma,
2. $\|\boldsymbol{x}\|_2 = (\sum_i |x_i^2|)^{\frac{1}{2}}$ az l_2 (euklideszi) norma,
3. $\|\boldsymbol{x}\|_\infty = \max_i |x_i|$ az l_∞ (maximum) norma.

3.11. Tétel. *(Cauchy-egyenlőtlenség) Legyen $\boldsymbol{x}$ és $\boldsymbol{y}$ n-ed rendű nem nulla vektor. Az*

$$|\boldsymbol{x}^T\boldsymbol{y}| \leq \|\boldsymbol{x}\|_2\,\|\boldsymbol{y}\|_2 \tag{3.11}$$

egyenlőtlenség akkor és csak akkor igaz, ha $\boldsymbol{y}$ vektor skalárszorosa $\boldsymbol{x}$ vektornak. [1, 42. oldal, Cauchy's Inequality]

A Cauchy-egyenlőtlenség komplex vektorokra is igaz, ha a transzponáltat Hermite-féle transzponáltra cseréjük. A tételt itt nem bizonyítjuk.

Definíció. Az $\|\boldsymbol{A}\|$ skalár az $\boldsymbol{A}$ mátrix normája, ha kielégíti a következő négy feltételt.

1. $\|\boldsymbol{A}\| = 0$, ha $\boldsymbol{A} = \boldsymbol{0}$, egyébként $\|\boldsymbol{A}\| > 0$.

2. $\|\boldsymbol{A}\theta\| = \|\boldsymbol{A}\| \, |\theta|$, ahol θ skalár.

3. $\|\boldsymbol{A}+\boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$.

4. $\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\| \, \|\boldsymbol{B}\|$.

Definíció. Az $\boldsymbol{A}$ mátrix $\|\boldsymbol{A}\|_p$ indukált normájának definíciója

$$\boldsymbol{A} = \max_{x\neq 0} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}, \tag{3.12}$$

ahol $p = 1{,}2$ vagy ∞.

Az indukált mátrixnormákra mindig igaz, hogy $\|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\| \, \|\boldsymbol{x}\|$. Ez a gyakorlatban sokszor hasznos. Az l_1 indukált mátrixnormát explicit az

$$\|\boldsymbol{A}\|_1 = \max_j \sum_i |a_{ij}| \tag{3.13}$$

képlettel számolhatjuk. Az l_∞ indukált mátrixnormát az

$$\|\boldsymbol{A}\|_\infty = \max_i \sum_j |a_{ij}| \tag{3.14}$$

képlettel számolhatjuk. Az l_2 által indukált mátrixnormát spektrális normának hívjuk. A spektrális norma elméleti fontossága mellett nagy gyakorlati hátránya, hogy nincs egyszerű explicit képlet a kiszámolására.

Definíció. Az $\boldsymbol{A}$ mátrix Frobenius normájának definíciója

$$\|\boldsymbol{A}\|_F = (\sum_{i,j} |a_{ij}|^2)^{\frac{1}{2}}. \tag{3.15}$$

A Frobenius norma az euklideszi vektornorma mátrix megfelelője, de nem az euklideszi vektornorma által indukált mátrixnorma. Nem vektornorma által indukált mátrixnorma, ezért a gyakorlatban sokszor szükségtelenül pontatlan eredményekhez vezet.

Definíció. A nemszinguláris $\boldsymbol{A}$ mátrix kondíciószámának jelölése $k(\boldsymbol{A})$, és definíciója

$$k(\boldsymbol{A}) = \|\boldsymbol{A}\| \left\|\boldsymbol{A}^{-1}\right\|, \tag{3.16}$$

ahol tetszőleges norma választható.

A kondíciószám mérete jellemzi a mátrix szingularitásának mértékét. Ha a kondíciószám kicsi, szokás a mátrixot jól kondicionáltnak hívni, ha a kondíciószám nagy, szokás a mátrixot rosszul kondicionáltnak hívni.

4. Lineáris egyenletrendszerek megoldása, osztályozása

A megoldása az

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{4.1}$$

alakú egyenletrendszernek, ahol $\boldsymbol{A}$ n-ed rendű nemszinguláris mátrix, $\boldsymbol{x}$ ismeretlen n-ed rendű vektor, $\boldsymbol{b}$ tetszőleges n-ed rendű vektor, az egyik leggyakoribb feladat. Az $\boldsymbol{A}$ mátrix az egyenletrendszer együtthatómátrixa. A (4.1) egyenletrendszernek akkor és csak akkor létezik megoldása, ha a $\boldsymbol{b}$ vektor lineárisan kifejezhető az $\boldsymbol{A}$ mátrix oszlopvektoraiból. Vagy másként fogalmazva, a $\boldsymbol{b}$ vektor és az $\boldsymbol{A}$ mátrix oszlopvektorai nem lineárisan függetlenek. Ekkor az egyenletrendszer megoldása

$$\boldsymbol{A}^{-1}\boldsymbol{b}\,. \tag{4.2}$$

Az $\boldsymbol{A}$ mátrix inverzének számolása a gyakorlatban nem célszerű, mert túl műveletigényes. A megoldási módszerek két fő osztálya az iteratív eljárások és a direkt eljárások. Egy megoldási módszer hatékonysága két fő szempont alapján bírálható el:

1. A megoldási módszer mennyire gyors, azaz mennyire műveletigényes?
2. Mennyire pontos a kiszámított megoldás? [2]

A gyakorlatban előforduló együtthatómátrixok általában vagy kitöltöttek és alacsony rendszámúak (pl. a rendszám kisebb mint 30 [2]), vagy ritkák és nagy rendszámúak. A direkt eljárások általában előnyösebbek a kis rendszámú kitöltött mátrixoknál, míg az iteratív eljárások általában előnyösebbek a nagy rendszámú ritka mátrixoknál. [2]

Az egyenletrendszer megoldásánál számolnunk kell különböző hibaforrásokkal. Hibaforrás lehet az együtthatómátrix vagy a jobboldali $\boldsymbol{b}$ vektor elemeinek hibája, a számítások közben keletkező kerekítési hibák, és a megoldási módszer képlethibája. [2]

4.1. Direkt eljárások

A direkt eljárások mentesek a képlethibától, és a megoldást véges sok lépésben állítják elő. Elméletben a megoldás pontos, a gyakorlatban azonban ez mégsem teljesül, a lépések közben keletkező kerekítési hibák miatt. A direkt eljárások alapja az (4.1) egyenletrendszer sorozatos transzformációja, amíg a megoldás könnyen számolhatóvá válik.

4.1.1. Az LU-felbontás

Definíció. Az $\boldsymbol{A}$ mátrix LU-felbontásán az

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} \tag{4.3}$$

felbontást értjük, ahol $\boldsymbol{L}$ alsó háromszögmátrix és $\boldsymbol{U}$ felső háromszögmátrix.

Az LU-felbontással az (4.1) egyenletrendszer megoldása az

$$\boldsymbol{L}\boldsymbol{y} = \boldsymbol{b} \tag{4.4}$$

és

$$\boldsymbol{U}\boldsymbol{x} = \boldsymbol{y} \tag{4.5}$$

háromszög alakú egyenletrendszerek megoldására egyszerűsíthető, ahol $\boldsymbol{y}$ vektor a felbontás egy köztes eredménye. Először a (4.5) egyenletrendszert oldjuk meg. Legyen $\boldsymbol{U} = [u_{ij}]$, $\boldsymbol{x} = [x_i]$ és $\boldsymbol{y} = [y_i]$ $(i, j = 1{,}2, ..., n)$ ekkor

$$x_i = (y_i - \sum_{j=i+1}^{n} u_{ij}x_j)/u_{ii}\,, \tag{4.6}$$

ahol $u_{ii} \neq 0$. Ahhoz, hogy a (4.5) egyenletrendszernek tetszőleges $\boldsymbol{y}$ vektorra legyen megoldása, feltételezzük, hogy $\boldsymbol{U}$ mátrix nemszinguláris. Így a 3.8-es tételből következik, hogy $u_{ii} \neq 0$. A (4.4) egyenletrendszer $\boldsymbol{y}$ vektor ismeretében hasonlóan oldható meg. [1]

4.1. Tétel. *Az $\boldsymbol{A}$ nemszinguláris mátrixnak akkor és csak akkor létezik LU-felbontása, ha az $\boldsymbol{A}$ mátrix minden bal felső sarokminormátrixa nemszinguláris. [1, 56. oldal, Theorem 4.1]*

Bizonyítás. Legyen $\boldsymbol{A}$ mátrix nemszinguláris. Megmutatjuk, hogy ha az $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ felbontás létezik, akkor az $\boldsymbol{A}$ mátrix bal felső sarokminormátrixai szükségképpen nemszingulárisak. Particionáljuk $\boldsymbol{A}$ mátrixot úgy, hogy

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix},$$

ahol $\boldsymbol{A}_{11}$ egy tetszőleges rendű bal felső sarokminormátrixa az $\boldsymbol{A}$ mátrixnak. Particionáljuk az $\boldsymbol{L}$ és $\boldsymbol{U}$ mátrixokat hasonlóan. A hipotézis szerint $\boldsymbol{A} = \boldsymbol{LU}$, azaz

$$\begin{bmatrix} \boldsymbol{L}_{11} & \boldsymbol{0} \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{11} & \boldsymbol{U}_{12} \\ \boldsymbol{0} & \boldsymbol{U}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}$$

és

$$\boldsymbol{L}_{11}\boldsymbol{U}_{11} = \boldsymbol{A}_{11} \,.$$

Az $\boldsymbol{A}$ mátrix nemszinguláris, így a 3.9-es tétel miatt $\boldsymbol{L}$ és $\boldsymbol{U}$ mátrixok is nemszingulárisak. A 3.7-es tétel miatt $\boldsymbol{L}_{11}$ és $\boldsymbol{U}_{11}$ szintén nemszingulárisak, így $\boldsymbol{A}_{11}$ nemszinguláris. Mivel $\boldsymbol{A}_{11}$ egy tetszőleges rendű bal felső sarokminormátrixa az $\boldsymbol{A}$ mátrixnak, így az $\boldsymbol{A}$ mátrix szükségképpen nemszinguláris.

Legyen

$$\boldsymbol{A}_r = \boldsymbol{L}_r\boldsymbol{U}_r \,, \tag{4.7}$$

ahol $\boldsymbol{A}_r, \boldsymbol{L}_r$ és $\boldsymbol{U}_r$ az $\boldsymbol{A}, \boldsymbol{L}$ és $\boldsymbol{U}$ mátrixok r-ed rendű bal felső sarokminormátrixai. Megmutatjuk, hogy ha $\boldsymbol{A}_r$ nemszinguláris, akkor $\boldsymbol{A}_{r+1}$ is szükségképpen nemszinguláris. Indukcióval bizonyítunk. Particionáljuk $\boldsymbol{A}_{r+1}$-t úgy, hogy

$$\boldsymbol{A}_{r+1} = \begin{bmatrix} \boldsymbol{A}_r & \boldsymbol{b}_r \\ \boldsymbol{c}_r^T & \delta_r \end{bmatrix}, \tag{4.8}$$

ahol $\boldsymbol{b}_r$ és $\boldsymbol{c}_r$ r-ed rendű vektorok és δ_r skalár. Particionáljuk $\boldsymbol{L}_{r+1}$-et és $\boldsymbol{U}_{r+1}$-et úgy, hogy

$$\boldsymbol{L}_{r+1} = \begin{bmatrix} \boldsymbol{L}_r & \boldsymbol{0} \\ \boldsymbol{p}_r & \lambda_r \end{bmatrix} \tag{4.9}$$

és

$$\boldsymbol{U}_{r+1} = \begin{bmatrix} \boldsymbol{U}_r & \boldsymbol{v}_r \\ \boldsymbol{0}^T & \mu_r \end{bmatrix}. \tag{4.10}$$

Ekkor található olyan $\boldsymbol{p}_r, \lambda_r, \boldsymbol{v}_r$ és μ_r, hogy

$$\boldsymbol{L}_{r+1}\boldsymbol{U}_{r+1} = \boldsymbol{A}_{r+1} \,. \tag{4.11}$$

A particionált (4.8), (4.9) és (4.10) mátrixokkal felírva a (4.11) egyenlőséget

$$\begin{bmatrix} \boldsymbol{L}_r & \boldsymbol{0} \\ \boldsymbol{p}_r & \lambda_r \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_r & \boldsymbol{v}_r \\ \boldsymbol{0}^T & \mu_r \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_r & \boldsymbol{b}_r \\ \boldsymbol{c}_r^T & \delta_r \end{bmatrix}.$$

A szorzás elvégzése után a (4.7) egyenlőséget és az

$$\boldsymbol{L}_r\boldsymbol{v}_r = \boldsymbol{b}_r \,, \tag{4.12}$$

$$\boldsymbol{p}_r^T \boldsymbol{U}_r = \boldsymbol{c}_r^T \tag{4.13}$$

és

$$\boldsymbol{p}_r^T \boldsymbol{v}_r + \lambda_r \mu_r = \delta_r \,, \tag{4.14}$$

egyenlőségeket kapjuk. Mivel a hipotézis szerint $\boldsymbol{A}_r$ nemszinguláris, a (4.7) egyenlőségből következik, hogy $\boldsymbol{L}_r$ és $\boldsymbol{U}_r$ szintén nemszingulárisak. Ezért léteznek olyan egyértelmű $\boldsymbol{p}_r$ és $\boldsymbol{v}_r$ vektorok, amelyek kielégítik a (4.12) és (4.13) egyenlőségeket. Mivel léteznek olyan λ_r és μ_r skalárok, amik kielégítik a (4.14) egyenlőséget, létezik olyan $\boldsymbol{L}_{r+1}$ és $\boldsymbol{U}_{r+1}$, amik kielégítik a (4.11) egyenlőséget. Tehát, ha az $\boldsymbol{A}$ mátrix bal felső r-ed rendű sarokminormátrixa nemszinguláris, és az $\boldsymbol{A}$ mátrixnak létezik LU-felbontása, akkor az $\boldsymbol{A}$ mátrix bal felső $r+1$-ed rendű sarokminormátrixa is nemszinguláris. □

Következmény. *Ha $\boldsymbol{L}$ alsó egység háromszögmátrix, az LU-felbontás egyértelmű. [1, 57. oldal, Corollary]*

Bizonyítás. Ha $\boldsymbol{L}$ alsó egység háromszögmátrix, akkor a (4.9) felbontásban $\lambda_r = 1$ minden r-re. Ebből következik, hogy (4.10) felbontásban μ_r egyértelmű minden r-re. Mivel $\boldsymbol{p}_r$ és $\boldsymbol{v}_r$ egyértelmű minden r-re, következik az eredmény. □

4.1.2. A Gauss-féle elimináció

Az egyik legrégebbi és legjobb módszer a (4.1) egyenletrendszer megoldására a Gauss-féle elimináció. Az első, második, stb. egyenlőség megfelelő skalárszorosát kivonjuk a többi egyenlőségből, úgy hogy ezzel az ismeretlen változókat elimináljuk. Ezt addig folytatjuk, amíg az utolsó egyenlőségben már csak egy ismeretlen változó marad. ekkor a kapott egyenlőségek felső háromszög alakúak. A módszert egy példán mutatjuk be.

Példa. Oldjuk meg a

$$\begin{array}{rrrrcr} 2x_1 & +x_2 & +3x_3 & -x_4 & = & 3 \\ -4x_1 & -3x_2 & -4x_3 & +5x_4 & = & 2 \\ 6x_1 & +4x_2 & +4x_3 & -5x_4 & = & -1 \\ -4x_1 & -3x_2 & +2x_3 & +4x_4 & = & 7 \end{array}$$

egyenletrendszert! Vonjuk ki az 1. egyenlőség $(-2)-$szeresét a második és negyedik egyenlőségből, és a 3-szorosát a harmadik egyenlőségből. Így ezekből az egyenlőségekből az x_1 ismeretlent elimináltuk, és a

$$\begin{array}{rrrrcr} 2x_1 & +x_2 & +3x_3 & -x_4 & = & 3 \\ & -x_2 & +2x_3 & +3x_4 & = & 8 \\ & x_2 & -5x_3 & -2x_4 & = & -10 \\ & -3x & +8x_3 & +2x_4 & = & 13 \end{array}$$

új egyenletrendszert kaptuk. Az új egyenletrendszerben az x_2 ismeretlen eliminálásához vonjuk ki a 2. egyenlőség (-1)-szeresét a harmadik egyenlőségből, és az 1-szeresét a negyedik egyenlőségből. Így ezekből az egyenlőségekből az x_2 ismeretlent elimináltuk, és a

$$\begin{array}{rrrrcr} 2x_1 & +x_2 & +3x_3 & -x_4 & = & 3 \\ & -x_2 & +2x_3 & +3x_4 & = & 8 \\ & & -3x_3 & x_4 & = & -2 \\ & & +6x_3 & x_4 & = & 5 \end{array}$$

új egyenletrendszert kaptuk. Végül vonjuk ki a 3. egyenlőség (-2)-szeresét a negyedik egyenlőségből. Így a negyedik egyenlőségből x_2-t elimináltuk, és a

$$\begin{array}{rrrrcr} 2x_1 & +x_2 & +3x_3 & -x_4 & = & 3 \\ & -x_2 & +2x_3 & +3x_4 & = & 8 \\ & & -3x_3 & x_4 & = & -2 \\ & & & x_4 & = & 1 \end{array}$$

új egyenletrendszert kaptuk. Visszahelyettesítve az $x_4 = 1$-et kapjuk, hogy $x_3 = 1$, $x_2 = -3$ és $x_1 = 2$. [1, 60. oldal, Example 4.3]

Legyen $\boldsymbol{A}^{(r-1)} = [a_{ij}^{(r-1)}]$, $1 \leq r \leq n$ a (4.1) egyenletrendszer Gauss-féle eliminációval való megoldásakor az $(r-1)$-edik eliminációs lépés után kapott együtthatómátrix. Legyen $\boldsymbol{M}_r = [m_{ij}]$ és

$$\boldsymbol{A}^{(r)} = \boldsymbol{M}_r \boldsymbol{A}^{(r-1)}. \tag{4.15}$$

A (4.15) egyenlőségben az $\boldsymbol{M}_r$ mátrixot mindig úgy választjuk, hogy a szorzás eredményeként az $\boldsymbol{A}^{(r)}$ együtthatómátrix r-edik oszlopában az r-edik sor alatt az együtthatókra nullákat kapjunk. Ehhez a (4.15) egyenlőségben

$$\boldsymbol{M}_r = \boldsymbol{I} - \boldsymbol{m}_r \boldsymbol{e}_r^T\,, \tag{4.16}$$

ahol

$$\boldsymbol{m}_r = [0,0,...,0, m_{r+1,r}, m_{r+2,r}, .., m_{nr}]^T$$

és

$$m_{ir} = \frac{a_{ir}^{(r-1)}}{a_{rr}^{(r-1)}}, \qquad a_{rr}^{(r-1)} \neq 0, \tag{4.17}$$

az $\boldsymbol{e}_r$ vektor pedig az r-edik oszlopa az $\boldsymbol{I}$ egységmátrixnak. Ekkor az $\boldsymbol{A}^{(r)}$ mátrixot úgy számolhatjuk, hogy kivonjuk az $\boldsymbol{A}^{(r-1)}$ mátrix r-edik sorának m_{ir}-szeresét az $\boldsymbol{A}^{(r-1)}$ mátrix i-edik sorából minden i-re, ahol $i = r+1, r+2, ..., n$. [1]

Definíció. A (4.15) egyenlőségben az m_{ir} elemeket multiplikátoroknak hívjuk, az $a_{rr}^{(r-1)}$ elemeket pedig r-edik pivot, vagy r-edik főelemnek hívjuk.

A módszer az alkalmazásakor az

$$\boldsymbol{A}^{(r)}\boldsymbol{x} = \boldsymbol{b}^{(r)} \tag{4.18}$$

egyenlőségrendszereket generálja, ahol $\boldsymbol{b}^{(r)} = \boldsymbol{M}_r\boldsymbol{b}^{(r-1)}$.

$$\boldsymbol{A}^{(n-1)} = \boldsymbol{M}_{n-1}\boldsymbol{M}_{n-2}...\boldsymbol{M}_1\boldsymbol{A}, \tag{4.19}$$

ahol $\boldsymbol{A}^{(n-1)}$ felső háromszögmátrix, és $\boldsymbol{M}_r$ a (4.16) egyenlőség miatt mindig alsó egység háromszögmátrix. Alsó egység háromszögmátrix inverze alsó egység háromszögmátrix, és alsó egység háromszögmátrixok szorzata is alsó egység háromszögmátrix, így ha

$$\boldsymbol{U} = \boldsymbol{A}^{(n-1)} \tag{4.20}$$

és

$$\boldsymbol{L} = (\boldsymbol{M}_{n-1}\boldsymbol{M}_{n-2}...\boldsymbol{M}_1)^{-1}, \tag{4.21}$$

a módszer eredménye az $\boldsymbol{A}$ együtthatómátrix LU-felbontását adja. Mivel $\boldsymbol{L}$ alsó egység háromszögmátrix, az LU-felbontás egyértelmű. Ha $\boldsymbol{x}$ kielégíti a (4.18) egyenlőséget, akkor kielégíti a (4.1) egyenlőséget is. [1]

A (4.17) egyenlőségből látszik, hogy a pivot elemek nem lehetnek nullák. Ha bármelyik pivot elem nulla, akkor az $\boldsymbol{A}$ mátrix bal felső sarokminormátrixa szinguláris. Ebből következik, hogy az $\boldsymbol{A}$ mátrixnak nem létezik LU-felbontása (4.1-es tétel). [1]

A nulla értékű r-edik pivot elem kiküszöbölésének egy módja, ha az $\boldsymbol{A}^{(r-1)}$ mátrix sorait felcseréljük, úgy hogy az r-edik pivot elem helyén ne nullát kapjunk. Ezt úgy érhetjük el, ha az $\boldsymbol{A}^{(r-1)}$ mátrix r-edik oszlopában az $a_{rr}^{(r-1)}$ elemtől lefelé megkeressük az első nem nulla elemet, legyen ez pl. $a_{ir}^{(r-1)}$, majd felcseréljük az $\boldsymbol{A}^{(r-1)}$ mátrix r-edik sorát az i-edik sorával. Ezt megtehetjük, mert $i > r$. [1]

4.2. Iteratív eljárások

Az iteratív megoldási módszereknél a (4.1) egyenletrendszer $\boldsymbol{A}$ mátrixa lényegében változatlan marad és közelítő $\boldsymbol{x}_i$ $(i = 1,2, ...)$ megoldásokat generálunk. Azt reméljük, hogy az $\{\boldsymbol{x}_i\}$ sorozat a megoldáshoz konvergál. A közelítő megoldások generálásának leállítását, azaz az iterációk leállítását, kilépési feltételhez, feltételekhez kötjük.

Definíció. Az $\boldsymbol{A}$ mátrixot $\boldsymbol{P} - \boldsymbol{Q}$ alakban kifejezve, ahol $\boldsymbol{P}$ nemszinguláris, az $\boldsymbol{A}$ mátrix felosztásának nevezzük.

Legyen az $\boldsymbol{A} = \boldsymbol{P} - \boldsymbol{Q}$ olyan felosztás, hogy a $\boldsymbol{P}\boldsymbol{y} = \boldsymbol{c}$ egyenletrendszer egyszerűen megoldható. Írjuk a (4.1) egyenletrendszert

$$\boldsymbol{P}\boldsymbol{x} = \boldsymbol{b} + \boldsymbol{Q}\boldsymbol{x} \tag{4.22}$$

alakban. Az $\boldsymbol{x}$ vektort úgy próbáljuk meghatározni, hogy egy tetszőleges $\boldsymbol{x}_0$ becslésből kiindulva egy $\{\boldsymbol{x}_i\}$ sorozatot generálunk, a

$$\boldsymbol{P}\boldsymbol{x}_{i+1} = \boldsymbol{b} + \boldsymbol{Q}\boldsymbol{x}_i \tag{4.23}$$

egyenletrendszer sorozatos megoldásával. [1]

4.2. Tétel. *A (4.23) egyenlőségből képzett $\{\boldsymbol{x}_i\}$ sorozat a (4.1) egyenletrendszer megoldásához konvergál, ha $\left\|\boldsymbol{P}^{-1}\boldsymbol{Q}\right\| < 1$, tetszőleges normával. [1, 68. oldal, Theorem 4.3]*

Bizonyítás. Legyen a hibavektor $\boldsymbol{e}_i = \boldsymbol{x}_i - \boldsymbol{x}$, ahol $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$, a (4.1) egyenletrendszer megoldása. A (4.22) összefüggést kivonva a (4.23) összefüggésből és balról $\boldsymbol{P}^{-1}$-el szorozva kapjuk, hogy $\boldsymbol{e}_{i+1} = \boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{e}_i$.

Legyen $\left\|\boldsymbol{P}^{-1}\boldsymbol{Q}\right\| = \alpha$, ekkor $\|\boldsymbol{e}_{i+1}\| \leq \alpha \|\boldsymbol{e}_i\|$, és ebből következik, hogy $\|\boldsymbol{e}_i\| \leq$ $\leq \alpha^i \|\boldsymbol{e}_0\|$. Tehát ha $\alpha < 1$, elég nagy i-re $\|\boldsymbol{e}_i\|$ tetszőlegesen kicsi. Azaz $\lim_{i\to\infty} \boldsymbol{x}_i =$ $= \boldsymbol{x}$. □

A fentiekben feltételeztük, hogy az $\boldsymbol{A}$ mátrix felosztása, a $\boldsymbol{P}$ és a $\boldsymbol{Q}$ mátrixok i-től függetlenek.

Definíció. Az iterációs eljárást stacionáriusnak nevezzük, ha az iterációs felosztás nem függ i-től.

Legyen $\boldsymbol{A}$ nemszinguláris mátrix és legyen $\boldsymbol{A} = \boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U}$, ahol $\boldsymbol{D}$ diagonális mátrix, $\boldsymbol{L}$ alsó háromszögmátrix és $\boldsymbol{U}$ felső háromszögmátrix. Néhány példa stacionárius iterációs eljárásokra:

- Jacobi-iteráció: $\boldsymbol{P} = \boldsymbol{D}, \quad \boldsymbol{Q} = \boldsymbol{L} + \boldsymbol{U}$,
- Gauss-Seidel eljárás: $\boldsymbol{P} = \boldsymbol{D} - \boldsymbol{L}, \quad \boldsymbol{Q} = \boldsymbol{U}$,
- felső relaxációs eljárás: $\boldsymbol{P} = \omega^{-1}\boldsymbol{D} - \boldsymbol{L}, \quad \boldsymbol{Q} = \boldsymbol{U} + (\omega^{-1} - 1)\boldsymbol{D}$. [1]

Az ω skalárral a konvergenciát gyorsítjuk. Ez általában azt jelenti, hogy $1 \leq \omega < 2$. Ha $\omega = 1$, a módszer pontosan a Gauss-Seidel eljárás.

A következő fejezetekben két iterációs eljárással fogunk részletesen foglalkozni, Broyden kutatásainak eredményeire támaszkodva. A felső relaxációs eljárással és a blokk konjugált gradiens módszerrel.

4.3. Sajátértékek és sajátvektorok

Ha az $\boldsymbol{A}$ mátrix négyzetes és nemszinguláris, akkor az $\boldsymbol{Ax} = \boldsymbol{b}$ egyenletrendszernek egyértelmű a megoldása. Ilyen egyenletrendszerek sokszor előfordulnak, amikor valamilyen rendszer statikus viselkedését vizsgáljuk, ahol a rendszer válasza a valamilyen stimulusra a $\boldsymbol{b}$ vektor. Azonban, ha az ilyen rendszerek dinamikus viselkedését akarjuk vizsgálni, akkor meg kell határoznunk azokat a λ skalárokat, amelyekre az $(\boldsymbol{A} - \boldsymbol{I}\lambda)$ mátrix szinguláris. Ezek a λ skalárok a vizsgált rendszer belső tulajdonságaival függenek össze. [1]

Definíció. Legyen $\boldsymbol{A}$ négyzetes mátrix. A λ skalár, ami mellett $\boldsymbol{A} - \lambda\boldsymbol{I}$ szinguláris, az $\boldsymbol{A}$ mátrix sajátértéke.

Definíció. Legyen $\boldsymbol{A}$ négyzetes mátrix. Az $\boldsymbol{x} \neq \boldsymbol{0}$ vektor, ami mellett

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}, \tag{4.24}$$

az $\boldsymbol{A}$ mátrix jobboldali sajátvektora. Legyen $\boldsymbol{B}$ négyzetes mátrix. Az $\boldsymbol{y} \neq \boldsymbol{0}$ vektor, ami mellett

$$\boldsymbol{y}^T(\boldsymbol{B} - \lambda\boldsymbol{I}) = \boldsymbol{0}^T, \tag{4.25}$$

a $\boldsymbol{B}$ mátrix baloldali sajátvektora.

A (4.24) egyenlőségből

$$\boldsymbol{Ax} = \lambda\boldsymbol{x}. \tag{4.26}$$

Definíció. Legyen $\boldsymbol{A}$ n-ed rendű négyzetes mátrix és $\boldsymbol{x} \neq \boldsymbol{0}$ tetszőleges n-ed rendű vektor. Ekkor az

$$\boldsymbol{x}, \boldsymbol{Ax}, \boldsymbol{A}^2\boldsymbol{x}, ...$$

sorozatot Krylov sorozatnak nevezzük.

4.3. Tétel. *Minden négyzetes mátrixnak van legalább egy sajátértéke. [1, 76. oldal, Theorem 5.1]*

Bizonyítás. Legyen $\boldsymbol{A}$ n-ed rendű négyzetes mátrix és $\boldsymbol{x} \neq \boldsymbol{0}$ tetszőleges n-ed rendű vektor. Az $\boldsymbol{x}, \boldsymbol{Ax}, \boldsymbol{A}^2\boldsymbol{x}, ..., \boldsymbol{A}^k\boldsymbol{x}$ Krylov sorozat vektorai $k > n$ esetén lineárisan összefüggők a 3.1-es tétel miatt, de nem zárhatjuk ki, hogy a sorozat ennél kevesebb vektora is lineárisan összefüggő. Tegyük fel, hogy a sorozat első $r \leq n$ vektora lineárisan független, de az első $r+1$ darab vektora lineárisan összefüggő. Ekkor léteznek olyan α_i, $i = 0,1, ..., r$, nem mind nulla, $r+1$ darab skalárok, hogy

$$\boldsymbol{x}\alpha_0 + \boldsymbol{A}\boldsymbol{x}\alpha_1 + ... + \boldsymbol{A}^r\boldsymbol{x}\alpha_r = \boldsymbol{0}. \tag{4.27}$$

Különösen $\alpha_r \neq 0$, ugyanis ebből a sorozat első $r+1$ darab vektorának lineáris függetlensége következne, ami ellentmond a hipotézisnek. Ezért a (4.27) egyenlőség átirtható a

$$\left(\frac{\alpha_0}{\alpha_r}\boldsymbol{I} + \frac{\alpha_1}{\alpha_r}\boldsymbol{A} + \frac{\alpha_2}{\alpha_r}\boldsymbol{A}^2 + ... + \boldsymbol{A}^r\right)\boldsymbol{x} = \boldsymbol{0} \tag{4.28}$$

alakra. Egyszerűbben

$$p(\boldsymbol{A})\boldsymbol{x} = \boldsymbol{0},$$

ahol

$$p(\boldsymbol{A}) = \frac{\alpha_0}{\alpha_r}\boldsymbol{I} + \frac{\alpha_1}{\alpha_r}\boldsymbol{A} + ... + \boldsymbol{A}^r\,. \tag{4.29}$$

Legyen

$$p(\xi) = \frac{\alpha_0}{\alpha_r} + \frac{\alpha_1}{\alpha_r}\xi + ... + \xi^r \tag{4.30}$$

polinom, ahol ξ skalár. Az algebra alaptételéből tudjuk, hogy

$$p(\xi) = (\xi - \lambda_1)(\xi - \lambda_2)...(\xi - \lambda_r)\,, \tag{4.31}$$

ahol $\lambda_1, \lambda_1, ..., \lambda_r$ a (nem feltétlen különböző) gyökei a $p(\xi) = 0$ egyenletnek. Láthatjuk, ha a (4.31) és (4.32) egyenlőségek jobboldalait kifejtjük, és összehasonlítjuk a ξ^i és $\boldsymbol{A}^i$ együtthatókat, ahol $0 \leq i \leq r$, hogy

$$p(\boldsymbol{A}) = (\boldsymbol{A} - \lambda_1\boldsymbol{I})(\boldsymbol{A} - \lambda_2\boldsymbol{I})...(\boldsymbol{A} - \lambda_r\boldsymbol{I})\,. \tag{4.32}$$

Most megmutatjuk, hogy $(\boldsymbol{A} - \lambda_1\boldsymbol{I})$ szinguláris. Tegyük fel az ellenkezőjét. Ekkor a (4.28) egyenlőséget balról $(\boldsymbol{A} - \lambda_1\boldsymbol{I})^{-1}$-el megszorozva kapjuk, a (4.29) és (4.32) egyenlőséget felhasználva, hogy

$$(\boldsymbol{A} - \lambda_2\boldsymbol{I})(\boldsymbol{A} - \lambda_3\boldsymbol{I})...(\boldsymbol{A} - \lambda_r\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}\,, \tag{4.33}$$

ahol most $\boldsymbol{x}$ $(r-1)$-ed rendű. A (4.33) egyenlőség úgy is írható, hogy

$$(\beta_0\boldsymbol{I} + \beta_1\boldsymbol{A} + ... + \boldsymbol{A}^{r-1})\boldsymbol{x} = \boldsymbol{0},$$

vagy

$$\boldsymbol{x}\beta_0 + \boldsymbol{A}\boldsymbol{x}\beta_1 + ... + \boldsymbol{A}^{r-1}\boldsymbol{x} = \boldsymbol{0},$$

ahol $0 \leq i \leq r-2$, és β_i-k a megfelelő együtthatók. Ebből az következik, hogy az $\boldsymbol{x}, \boldsymbol{Ax}, ...$ sorozat első r vektora lineárisan összefüggő, ami ellentmond a hipotézisnek. Ez az ellentmondás garantálja az $(\boldsymbol{A} - \lambda_1 \boldsymbol{I})$ mátrix szingularitását, és az $\boldsymbol{A}$ mátrix legalább egy sajátértékének létezését. $\square$

4.4. Tétel. *Legyen $\boldsymbol{A}$ komplex négyzetes mátrix. Ha λ sajátértéke és $\boldsymbol{x}$ sajátvektora az $\boldsymbol{A}$ mátrixnak, akkor $\overline{\lambda}$ is sajátértéke és $\overline{\boldsymbol{x}}$ is sajátvektora $\boldsymbol{A}$ mátrixnak. [1, 77. oldal, Theorem 5.2]*

Bizonyítás. Mivel $\boldsymbol{Az} = \boldsymbol{z}\lambda$, a komplex konjugáltakat véve $\overline{(\boldsymbol{Az})} = \overline{(\boldsymbol{z}\lambda)}$. $\overline{(\boldsymbol{Az})} =$ $= \overline{\boldsymbol{A}}\overline{z}$ és $\overline{(\boldsymbol{z}\lambda)} = \overline{\lambda}\overline{z}$. $\square$

Következmény. *Ha $\boldsymbol{A}$ valós mátrix, és ha van komplex sajátértéke és sajátvektora, ezek mindig komplex konjugált párokban jelentkeznek.*

Bizonyítás. Ha $\boldsymbol{A}$ valós mátrix, akkor $\boldsymbol{A} = \overline{\boldsymbol{A}}$. Az eredmény a tételből közvetlenül következik. $\square$

Bevezetjük a valós ortogonális mátrix komplex megfelelőjét, az unitér mátrixot.

Definíció. Az $\boldsymbol{A}$ komplex mátrix unitér mátrix, ha $\boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{I}$.

A Hermite-féle mátrixot ezért a valós szimmetrikus mátrix komplex megfelelőjének tekintjük.

Definíció. Legyen $\boldsymbol{A}$ és $\boldsymbol{P}$ n-ed rendű mátrix, és legyen $\boldsymbol{P}$ nemszinguláris. A $\boldsymbol{PAP}^{-1}$ transzformációt hasonlósági transzformációnak, az $\boldsymbol{A}$ és $\boldsymbol{PAP}^{-1}$ mátrixokat hasonlónak nevezzük. Ezenfelül, ha $\boldsymbol{P}$ mátrix ortogonális (unitér), a $\boldsymbol{PAP}^{-1}$ hasonlósági transzformációt ortogonális (unitér) transzformációnak nevezzük.

4.5. Tétel. *A hasonlósági transzformáció nem változtatja meg a mátrix sajátértékeit. [1, 79. oldal, Theorem 5.3]*

Bizonyítás. Legyen $\boldsymbol{A}$ és $\boldsymbol{P}$ n-ed rendű mátrix, és legyen $\boldsymbol{P}$ nemszinguláris. Legyen $\boldsymbol{Ax} = \boldsymbol{x}\lambda$. Ekkor $\boldsymbol{PAP}^{-1}\boldsymbol{Px} = \boldsymbol{Px}\lambda$. Ezért, ha λ az $\boldsymbol{A}$ mátrixhoz tartozó sajátérték, és $\boldsymbol{x}$ a hozzátartozó sajátvektor, akkor λ a $\boldsymbol{PAP}^{-1}$ mátrixnak is sajátértéke, és $\boldsymbol{Px}$ a hozzátartozó sajátvektora. $\square$

4.6. Tétel. *Minden k-ad rendű $\boldsymbol{A}$ mátrixhoz létezik olyan $\boldsymbol{X}$ unitér mátrix, hogy $\boldsymbol{X}^H\boldsymbol{AX}$ első oszlopa a k-ad rendű egységmátrix első oszlopának többszöröse. [1, 79. oldal, Lemma 5.2]*

Bizonyítás. Legyen $\boldsymbol{Ax} = \lambda\boldsymbol{x}$, ahol $\boldsymbol{x}$ normalizált, azaz $\boldsymbol{x}^H\boldsymbol{x} = 1$. Legyen $\boldsymbol{X}$ olyan unitér mátrix, aminek az első oszlopa $\boldsymbol{x}$. Ekkor $\boldsymbol{x} = \boldsymbol{Xe}_1$, így $\boldsymbol{AXe}_1 = \boldsymbol{Xe}_1\lambda$. Balról szorozva $\boldsymbol{X}^H$ mátrixszal kapjuk, hogy

$$\boldsymbol{X}^H\boldsymbol{AXe}_1 = \boldsymbol{e}_1\lambda\,. \tag{4.34}$$

$\square$

Megmutatjuk, hogy minden n-ed rendű négyzetes mátrixot hasonlósági transzformációval felső háromszögmátrixszá alakíthatunk. A felső háromszögmátrix sajátértékei pedig a főátló elemei.

4.7. Tétel. *Minden n-ed rendű $\boldsymbol{A}$ mátrixhoz létezik egy olyan $\boldsymbol{Q}$ unitér mátrix, hogy*

$$\boldsymbol{Q}^H \boldsymbol{A} \boldsymbol{Q} = \boldsymbol{U}, \tag{4.35}$$

ahol $\boldsymbol{U}$ felső háromszögmátrix. [1, 80. oldal, Theorem 5.4 (Schur's theorem)]

Bizonyítás. Indukcióval bizonyítunk. Tegyük fel, hogy létezik olyan $\boldsymbol{Q}_r$ unitér mátrix, hogy

$$\boldsymbol{Q}_r^H \boldsymbol{A} \boldsymbol{Q}_r = \begin{bmatrix} \boldsymbol{U}_r & \boldsymbol{B}_r \\ \boldsymbol{0} & \boldsymbol{C}_r \end{bmatrix}, \tag{4.36}$$

ahol $\boldsymbol{U}_r$ r-ed rendű felső háromszögmátrix. Megmutatjuk, hogy ekkor az egyenlőség $(r+1)$-re is igaz. Mivel $\boldsymbol{C}_r$ négyzetes, a 4.6-os tétel miatt létezik olyan $\boldsymbol{X}_r$ unitér mátrix, hogy

$$\boldsymbol{X}_r^H \boldsymbol{C}_r \boldsymbol{X}_r \boldsymbol{e}_1 = \boldsymbol{e}_1 \lambda \,, \tag{4.37}$$

ahol λ a $\boldsymbol{C}_r$ részmátrix sajátértéke. Legyen

$$\boldsymbol{P}_r = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}_r \end{bmatrix}. \tag{4.38}$$

Mivel $\boldsymbol{P}_r$ unitér mátrix, a (4.36) és (4.38) egyenlőségekből következik, hogy

$$\boldsymbol{P}_r \boldsymbol{Q}_r^H \boldsymbol{A} \boldsymbol{Q}_r \boldsymbol{P}_r = \begin{bmatrix} \boldsymbol{U}_r & \boldsymbol{B}_r \boldsymbol{X}_r \\ \boldsymbol{0} & \boldsymbol{X}_r \boldsymbol{C}_r \boldsymbol{X}_r \end{bmatrix}. \tag{4.39}$$

Így a (4.37) egyenlőség miatt a (4.39) egyenlőség jobb oldala írható úgy, hogy

$$\begin{bmatrix} \boldsymbol{U}_{r+1} & \boldsymbol{B}_{r+1} \\ \boldsymbol{0} & \boldsymbol{C}_{r+1} \end{bmatrix}.$$

Legyen $\boldsymbol{Q}_{r+1} = \boldsymbol{Q}_r \boldsymbol{P}_r$, ekkor $\boldsymbol{Q}_{r+1}$ unitér, mert két unitér mátrix szorzata unitér. Így a (4.36) egyenlőség $(r+1)$-re is igaz. Mivel a (4.36) egyenlőség $r = 1$ esetén igaz a 4.6-os tétel miatt, így minden r-re, $1 \leq r \leq n-1$, igaz. Ha $r = n-1$ a (4.36) egyenlőség jobb oldala felső háromszög mátrix, ezt $\boldsymbol{U}$-val jelölve és $\boldsymbol{Q}_{n-1}$-et $\boldsymbol{Q}$-val jelölve, kapjuk a tétel állítását. $\square$

A tétel következményeit bizonyítás nélkül közöljük, a bizonyítások megtalálhatók [1] 80-82. oldalán.

Következmény. *Az n-ed rendű $\boldsymbol{A}$ négyzetes mátrixnak n darab sajátértéke van.*

A tétel következő következménye előtt bevezetjük a normálmátrix fogalmát.

Definíció. Az $\boldsymbol{A}$ mátrix normálmátrix, ha $\boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^H$.

Következmény. *Az n-ed rendű $\boldsymbol{A}$ négyzetes mátrix diagonális mátrixszá transzformálható egy unitér transzformációval akkor és csak akkor, ha az $\boldsymbol{A}$ mátrix normálmátrix.*

Következmény. *Ha az $\boldsymbol{A}$ Hermite-féle mátrix, akkor az $\boldsymbol{U}$ mátrix a (4.35) egyenlőségben valós diagonális mátrix, és ezért az $\boldsymbol{A}$ mátrix sajátértékei valósak.*

Következmény. *Ha az $\boldsymbol{A}$ mátrix valós, és minden sajátértéke valós, akkor az $\boldsymbol{A}$ mátrix valós felső háromszögmátrixszá alakítható egy ortogonális transzformációval.*

Következmény. *Ha az $\boldsymbol{A}$ mátrix valós és szimmetrikus, akkor az $\boldsymbol{A}$ mátrix valós diagonális mátrixszá alakítható egy ortogonális transzformációval*

Megmutattuk, hogy az n-ed rendű $\boldsymbol{A}$ négyzetes mátrixnak n darab sajátértéke van, de ezek a sajátértékek nem feltétlen különbözőek. Most megvizsgáljuk hány darab különböző, azaz lineárisan független sajátvektora van egy mátrixnak. Kiemelt fontosságú, hogy egy n-ed rendű négyzetes mátrixnak van-e n darab lineárisan független sajátvektora. Ha igen, akkor ezek a lineárisan független sajátvektorok bázist alkotnak.

Definíció. Ha egy mátrix sajátvektorai nem alkotnak bázist, akkor a mátrixot defektív mátrixnak nevezzük.

Definíció. A mátrixot amelynek oszlopai egy nemdefektív mátrix sajátvektorai, modálmátrixnak nevezzük.

A modálmátrix nem egyértelmű. Egyrészt nincs meghatározva a mátrix oszlopainak sorrendje, másrészt a sajátvektorok tetszőlegesen méretezhetők.

4.8. Tétel. *Legyen $\boldsymbol{A}$ n-ed rendű mátrix és legyen m darab különböző sajátértéke, $m \leq n$. Az $\boldsymbol{A}$ mátrixnak legalább m darab lineárisan független sajátvektora van. [1, 86. oldal, Theorem 5.6]*

Bizonyítás. A tételt először az $\boldsymbol{U} = [u_{ij}]$ felső háromszögmátrixra bizonyítjuk. Jelölje λ_i az $\boldsymbol{U}$ mátrix sajátértékeit, λ_k pedig a többszörös sajátértékeit. Legyen

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_{11} & \boldsymbol{U}_{12} \\ \boldsymbol{0} & \boldsymbol{U}_{22} \end{bmatrix}, \tag{4.40}$$

ahol $\boldsymbol{U}_{11}$ és $\boldsymbol{U}_{22}$ négyzetes mátrixok. Particionáljuk úgy az $\boldsymbol{U}$ mátrixot, hogy $\boldsymbol{U}_{11}$ mátrixnak ne legyen diagonális eleme λ_k, az $\boldsymbol{U}_{22}$ mátrixnak pedig legyen az első diagonális eleme λ_k. Tekintsük az $(\boldsymbol{U} - \lambda_k \boldsymbol{I}) = \boldsymbol{0}$ egyenlőséget, amit írhatunk az

$$\begin{bmatrix} \boldsymbol{U}_{11} - \lambda_k \boldsymbol{I} & \boldsymbol{U}_{12} \\ \boldsymbol{0} & \boldsymbol{U}_{22} - \lambda_k \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \tag{4.41}$$

vagy a

$$(\boldsymbol{U}_{11} - \lambda_k \boldsymbol{I})\boldsymbol{x}_1 = -\boldsymbol{U}_{12}\boldsymbol{x}_2 \tag{4.42}$$

és

$$(\boldsymbol{U}_{22} - \lambda_k \boldsymbol{I})\boldsymbol{x}_2 = \boldsymbol{0} \tag{4.43}$$

alakban. A particionálás miatt az $(\boldsymbol{U}_{22} - \lambda_k \boldsymbol{I})$ mátrix első oszlopa nulla, ezért $\boldsymbol{x}_2 = = \boldsymbol{e}_1$, ahol $\boldsymbol{e}_1$ a megfelelő rendű egységmátrix első oszlopa. A particionálás miatt az $(\boldsymbol{U}_{11} - \lambda_k \boldsymbol{I})\boldsymbol{x}_1$ nemszinguláris. Ekkor az (4.42) egyenlőségből

$$\boldsymbol{x}_1 = -(\boldsymbol{U}_{11} - \lambda_k \boldsymbol{I})^{-1}\boldsymbol{U}_{12}\boldsymbol{e}_2\,. \tag{4.44}$$

A (4.41) egyenlőségből látszik, hogy ha $\boldsymbol{U}_{11}$ $(r-1)$-ed rendű, akkor az $\boldsymbol{x}$ sajátvektor az r-edik oszlopa valamilyen n-ed rendű felső egység háromszögmátrixnak. Így az $\boldsymbol{U}$ mátrix m darab különböző sajátértékéhez tartozik különböző sajátvektor, amelyek egy felső egység háromszögmátrix különböző oszlopai. Mivel a felső egység háromszögmátrix nemszinguláris, a sajátvektorok lineárisan függetlenek.

Hogy a tételt az általános $\boldsymbol{A}$ mátrixra belássuk, legyen

$$\boldsymbol{Q}^H \boldsymbol{A}\boldsymbol{Q} = \boldsymbol{U}, \tag{4.45}$$

ahol $\boldsymbol{Q}$ unitér mátrix. Legyen $\boldsymbol{V}$ $n{\times}m$-es mátrix a $\lambda_1, \lambda_2, ..., \lambda_m$ különböző sajátértékekhez tartozó lineárisan független sajátvektorok mátrixa. Legyen $\boldsymbol{\Lambda}_1 = diag(\lambda_i)$ az m-ed rendű diagonális mátrix a $\lambda_1, \lambda_2, ..., \lambda_m$ különböző sajátértékkel. Ekkor

$$\boldsymbol{U}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda}_1\,. \tag{4.46}$$

A (4.45) és a (4.46) egyenlőségekből következik, hogy

$$\boldsymbol{A}\boldsymbol{Q}\boldsymbol{V} = \boldsymbol{Q}\boldsymbol{V}\boldsymbol{\Lambda}_1\,.$$

Azaz a $\boldsymbol{Q}\boldsymbol{V}$ mátrix oszlopai az $\boldsymbol{A}$ mátrixnak a $\lambda_1, \lambda_2, ..., \lambda_m$ különböző sajátértékeihez tartozó lineárisan független sajátvektorok mátrixa. Mivel $\boldsymbol{V}$ mátrix oszlopai lineárisan függetlenek és $\boldsymbol{Q}$ unitér mátrix, így a $\boldsymbol{Q}\boldsymbol{V}$ mátrix nemszinguláris, tehát a sajátvektorok lineárisan függetlenek. □

A tétel következményeit bizonyítás nélkül közöljük, a bizonyítások megtalálhatók [1] 88. oldalán.

Következmény. *Ha egy n-ed rendű mátrixnak n darab különböző sajátértéke van, akkor nem defektív mátrix.*

Következmény. *Ha egy mátrix sajátértéke egyszeres, a hozzátartozó sajátvektor egyértelmű, a méretezést leszámítva.*

Következmény. *Ha egy mátrix sajátértéke r-szeres, akkor legfeljebb r darab lineárisan független sajátvektor tartozik hozzá.*

A Iteratív eljárások fejezetben láttuk, hogy az iteratív eljárások konvergenciája összefügg a mátrixok normájával. Megvizsgálunk néhány, a sajátérték és a mátrix norma közötti összefüggést.

Definíció. Legyen $\boldsymbol{A}$ n-ed rendű mátrix, λ_i sajátvektorokkal, $1 \leq i \leq n$. Az $\boldsymbol{A}$ mátrix spektrális sugara $\rho(\boldsymbol{A}) = \max_i |\lambda_i|$.

4.9. Tétel. *Legyen $\boldsymbol{A}$ négyzetes mátrix. Ekkor $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|$ tetszőleges normával. [1, 89. oldal, Theorem 5.7]*

Bizonyítás. Legyen $\boldsymbol{x}$ az $\boldsymbol{A}$ mátrix egy sajátvektora és legyen λ a hozzátartozó sajátérték, azaz $\boldsymbol{Ax} = \lambda\boldsymbol{x}$. Így a normák tulajdonságai alapján

$$\|\lambda\boldsymbol{x}\| = \|\lambda\| \|\boldsymbol{x}\| = \|\boldsymbol{Ax}\| \leq \|\boldsymbol{A}\| \|\boldsymbol{x}\| .$$

Mivel $\boldsymbol{x} \neq \boldsymbol{0}$, így $|\lambda| \leq \|\boldsymbol{A}\|$. □

Definíció. Legyen $\boldsymbol{A} = [a_{ij}]$ n-ed rendű mátrix. Ekkor az $\boldsymbol{A}$ mátrix nyoma $tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$.

4.10. Tétel. *A mátrixhoz tartozó sajátértékek összege egyenlő a mátrix nyomával. [1, 91. oldal, Theorem 5.9]*

Bizonyítás. Bármilyen $\boldsymbol{B}$ és $\boldsymbol{C}$ mátrixokra, amelyeknek a $\boldsymbol{BC}$ és $\boldsymbol{CB}$ szorzatuk definiálva van, $tr(\boldsymbol{BC}) = tr(\boldsymbol{CB})$. A 4.7-es tételből tudjuk, hogy $\boldsymbol{Q}^H\boldsymbol{AQ} = \boldsymbol{U}$. Ezért

$$tr(\boldsymbol{U}) = tr((\boldsymbol{Q}^H\boldsymbol{A})\boldsymbol{Q}) = tr(\boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{A})) = tr(\boldsymbol{A}),$$

az $\boldsymbol{U}$ mátrix nyoma pedig megegyezik az $\boldsymbol{U}$ mátrix sajátvektorainak összegével. □

4.4. Kovergens mátrixok

A különböző iteratív eljárások konvergenciája gyakran egy $\boldsymbol{A}^r$ mátrix viselkedésétől függ, miközben r tart a végtelenhez. Bizonyos iteratív eljárások akkor és csak akkor konvergensek, ha $\boldsymbol{A}^r$ a nullmátrixhoz konvergál, ahogy növekszik r. [1]

Definíció. Az n-ed rendű $\boldsymbol{A}$ mátrix konvergens, ha $\lim_{r\to\infty} \|\boldsymbol{A}^r\| = 0$.

4.11. Tétel. *Legyen $\boldsymbol{U}$ felső háromszögmátrix és $\rho(\boldsymbol{U}) < 1$. Ekkor létezik $\boldsymbol{B}$ diagonális mátrix, amelyre $\left\|\boldsymbol{B}\boldsymbol{U}\boldsymbol{B}^{-1}\right\|_\infty < 1$ és $\left\|\boldsymbol{B}\boldsymbol{U}\boldsymbol{B}^{-1}\right\|_1 < 1$. [1, 92. oldal, Theorem 5.10]*

Bizonyítás. Legyen $\boldsymbol{U} = [u_{ij}]$ n-ed rendű felső háromszögmátrix. Legyen $\boldsymbol{B} = diag(1, \beta, \beta^2, ..., \beta^{n-1})$, ahol válasszunk $\beta > 1$ skalárt úgy, hogy

$$\beta^{j-i} > \frac{(n-1)\,|u_{ij}|}{1-\rho(\boldsymbol{U})}, \qquad j > i \tag{4.47}$$

teljesüljön. Legyen $\boldsymbol{V} = [v_{ij}] = \boldsymbol{B}\boldsymbol{U}\boldsymbol{B}^{-1}$. Ekkor $j > i$-re $v_{ij} = u_{ij}/\beta^{j.i}$ és a (4.47) egyenlőtlenség miatt

$$|v_{ij}| < \frac{1-\rho(\boldsymbol{U})}{n-1}, \qquad j > i. \tag{4.48}$$

De $|v_{ii}| = |u_{ii}| \leq \rho(\boldsymbol{U})$ és $v_{ij} = 0$ minden $i > j$-re. Ezért

$$\|\boldsymbol{V}\|_\infty \leq \rho(\boldsymbol{U}) + \max_i \sum_{j=i+1}^{n} |v_{ij}| \ ,$$

és így az (4.48) egyenlőtlenség miatt

$$\|\boldsymbol{V}\|_\infty \leq \rho(\boldsymbol{U}) + (n-1)\left[\frac{1-\rho(\boldsymbol{U})}{n-1}\right] < 1\,.$$

Hasonlóan, $\|\boldsymbol{V}\|_1 < 1$. □

4.12. Tétel. *Az n-ed rendű $\boldsymbol{A}$ mátrix konvergenciájának szükséges és elégséges feltétele, hogy $\rho(\boldsymbol{A}) < 1$. [1, 93. oldal, Theorem 5.11]*

Bizonyítás. A tételt először egy $\boldsymbol{U} = [u_{ij}]$ felső háromszögmátrixra bizonyítjuk, majd a 4.7-es tétellel kiterjesztjük a bizonyítást tetszőleges négyzetes mátrixra. Az i-edik diagonális eleme $\boldsymbol{U}^r$ mátrixnak u_{ii}^r, ezért $\boldsymbol{U}$ nem lehet konverges, ha $u_{ii} \geq 1$ bármilyen i-re. Ekkor ugyanis az $\boldsymbol{U}$ mátrix hatványozásakor az aktuális u_{ii}^r hatványok folyamatosan nőnek. Ha $u_{ii} \leq 1$, az $\boldsymbol{U}$ mátrix hatványozásakor az aktuális u_{ii}^r hatványok folyamatosan csökkennek. Mivel $\rho(\boldsymbol{U}) = \max_i |u_{ii}|$, ezért $\rho(\boldsymbol{U}) < 1$ szükséges feltétele a konvergenciának. Bizonyítjuk a feltétel elégségességet. A 4.11-es tétel következménye, hogy ha $\rho(\boldsymbol{U}) < 1$, akkor létezik olyan $\boldsymbol{B}$ diagonális mátrix, hogy $\boldsymbol{V} = \boldsymbol{B}\boldsymbol{U}\boldsymbol{B}^{-1}$ és $\|\boldsymbol{V}\|_2 < 1$. Ezért $\boldsymbol{U} = \boldsymbol{B}^{-1}\ \boldsymbol{V}\boldsymbol{B}$ és $\boldsymbol{U}^r = \boldsymbol{B}^{-1}\ \boldsymbol{V}^r\boldsymbol{B}$, mivel $\|\boldsymbol{B}\|_2 \left\|\boldsymbol{B}^{-1}\right\|_2 = \beta^{n-1}$. Ezekből következik, hogy $\|\boldsymbol{U}^r\|_2 \leq \beta^{n-1} \|\boldsymbol{V}\|_2^r$. Mivel $\|\boldsymbol{V}\|_2 < 1$ és $\lim_{r\to\infty} \|\boldsymbol{U}^r\| = 0$, az $\boldsymbol{U}$ mátrix konvergens.

Általános esetben, a 4.7-es tételből $\boldsymbol{Q}^H\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{U}$, ezért $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{U}\boldsymbol{Q}^H$ és $\boldsymbol{A}^r = \boldsymbol{Q}\boldsymbol{U}^r\boldsymbol{Q}^H$. Mivel $\|\boldsymbol{Q}\|_2 \left\|\boldsymbol{Q}^H\right\|_2 = 1$, $\|\boldsymbol{A}^r\|_2 = \|\boldsymbol{U}^r\|_2$. Ebből az egyenlőségből következik, hogy a tétel általános esetben is igaz. □

5. Felső relaxációs eljárás (SOR)

6. Blokk konjugált gradiens módszer (BlCG)

7. A konjugált gradiens módszerek új rendszertana

8. MATLAB tesztfeladatokon való összehasonlítás

9. Konklúzió

10. Függelék

Hivatkozások

[1] Charles George Broyden. *Basic Matrices.* The Macmillan Press Ltd, 1975.

[2] Anthony Ralston. *Bevezetés a numerikus analízisbe.* Műszaki Könyvkiadó, Budapest, 1968.