Tartalomjegyzék

1. Bevezetés

2. Broyden életrajza

Charles George Broyden (1933. február 3. - 2011. május 20.) szerény családi háttérrel rendelkezett. Angliában született, édesapja gyári munkásként, édesanyja háztartásbeliként dolgozott. Szülei ennek ellenére a kezdetektől tanulásra biztatták. Charles már gyerekként rengeteget olvasott, jó tanuló volt. Édesapja sajnálatos módon meghalt tuberkulózisban, mikor ő még csak 11 éves volt. Ez még jobban megnehezítette családi helyzetüket, de édesanyja még 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 kutatás eredményeként valódi ipari alkalmazások nemlineáris optimalizálási problémáinak megoldására adtak 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 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 módszereken dolgozott. Később a numerikus lineáris algebrára fókuszált, ezen belül is a konjugált gradiens módszerekre és ezek rendszertanára. A szakdolgozat legfőképp kutatásának 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, 4 éves korában meghalt. Broyden nagy örömét lelte családjában, gyermekeiben, Christopher, Jane 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 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. A 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. Az $\boldsymbol{x}$ vektor transzponáltját $\boldsymbol{x}^T$-vel jelöljük. Oszlopvektoron vektort, sorvektoron vektor transzponáltat értünk.

Definíció. Legyen $\boldsymbol{x} = [x_i]$ és $\boldsymbol{y} = [y_i]$ n-ed rendű valós vektor. A belső szorzata vagy skalár 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. A valós mátrix azonos rendű valós vektorok egy rendezett halmaza.

A mátrixokat a szakdolgozatban vastag nagybetűvel jelöljük. Az $\boldsymbol{A} = [a_{ij}]$ jelölésben a_{ij} az $\boldsymbol{A}$ mátrix i-edik sorának j-edik eleme. Ha a mátrix $m \times n$-es, $i = 1{,}2, ..., m$ és $j = 1{,}2, ..., n$. Egy $m \times n$-es mátrixot értelmezhetünk, mint m darab sorvektor, vagy mint n darab oszlopvektor. Röviden azt mondjuk, hogy a mátrix sorai és oszlopai. 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. Az $i = j$ elemek a mátrix diagonális elemei. Ha a mátrixnak ugyanannyi sora és oszlopa van, négyzetes mátrixnak nevezzük. A négyzetes mátrix rendje az oszlopainak száma. A csupa nullából álló mátrixot, vektort $\mathbf{0}$-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.

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

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

$$y_i = \boldsymbol{a}_i^T \boldsymbol{x}. \tag{3.2}$$

A mátrix-vektor szorzás disztributív.

Definíció. Legyen $\boldsymbol{A}$ $m \times n$-es mátrix. Jelöle $\boldsymbol{a}_i^T$ az $\boldsymbol{A}$ mátrix i-edik sorát. Legyen $\boldsymbol{B}$ $n \times p$-s mátrix. 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

$$c_{ij} = \boldsymbol{a}_i^T \boldsymbol{b}_j. \tag{3.3}$$

Az $\boldsymbol{AB}$ esetén az $\boldsymbol{A}$ mátrix oszlopainak száma meg kell egyezzen a $\boldsymbol{B}$ mátrix sorainak számával, hogy a megfelelő sorvektorok és oszlopvektorok belső szorzatai definiálva legyenek. A mátrixszorzás nem kommutatív, $\boldsymbol{AB} \neq \boldsymbol{BA}$.

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

$$\sum_{i=1}^{n} \boldsymbol{p}_i y_i \tag{3.4}$$

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

Definíció. Az $n \times n$-es $\boldsymbol{D} = [d_{ij}]$ mátrix diagonális, 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. 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 és $\boldsymbol{u} = \boldsymbol{x} +$ $+ \boldsymbol{y}$. 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. A művelettartás a részmátrixokra két particionált mátrix között azonban nem feltétlenül definiált.

Definíció. Mátrixok egy halmaza egy művelet szerint jól particionált, ha a mátrixok részmátrixai 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

$$\boldsymbol{E}_1\boldsymbol{F}_1 \neq \begin{bmatrix} \boldsymbol{AJ} + \boldsymbol{BL} & \boldsymbol{AK} + \boldsymbol{BM} \\ \boldsymbol{CJ} + \boldsymbol{DL} & \boldsymbol{CK} + \boldsymbol{DM} \end{bmatrix}.$$

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

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

Az összeadás szerint $\boldsymbol{E}_2$ és $\boldsymbol{F}_2$ nem jól particionáltak, mert

$$\boldsymbol{E}_2 + \boldsymbol{F}_2 \neq \begin{bmatrix} \boldsymbol{A} + \boldsymbol{J} & \boldsymbol{K} + \boldsymbol{B} \\ \boldsymbol{C} + \boldsymbol{L} & \boldsymbol{D} + \boldsymbol{M} \end{bmatrix}.$$

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. A $\boldsymbol{w}$ komplex vektor konjugáltját $\overline{\boldsymbol{w}}$-vel, az $\boldsymbol{A}$ komplex mátrix konjugáltját $\overline{\boldsymbol{A}}$-val jelöljük.

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

Definíció. A $\boldsymbol{w}$ 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 nem csak a mátrix algebrában. 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

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

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

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

Példa. Az $\boldsymbol{a}_1^T = [\ 1\ \ 0\ \ 0\ ]$, $\boldsymbol{a}_2^T = [\ 1\ \ 1\ \ 0\ ]$, $\boldsymbol{a}_3^T = [\ -1\ \ 1\ \ 1\ ]$ 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 teljesül, 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 $\boldsymbol{x} \neq \boldsymbol{0}$ vektor, hogy $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$.

Ha az $\boldsymbol{A}$ mátrix oszlopai lineárisan összefüggők, abból nem csak az következik, hogy az oszlopok megfelelő lineáris kombinációja nulla, hanem hogy létezik olyan $\boldsymbol{x} \neq \boldsymbol{0}$ vektor, amely az $\boldsymbol{A}$ mátrix minden sorára ortogonális. Az $\boldsymbol{x}$ vektor nem egyértelmű, mert 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.

A következő tételeket csak kimondjuk és hivatkozunk a bizonyításukra.

3.1. Tétel. *Az $n+1$ darab n-ed rendű vektor lineárisan összefüggő.*

A bizonyítást lásd, C.G. Broyden - Basic Matrices, Theorem 2.2, 29. oldal.

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.*

A bizonyítást lásd, C.G. Broyden - Basic Matrices, Lemma 2.2, 30. oldal.

3.3. Tétel. *Legyen $\boldsymbol{A}$ tetszőleges mátrix. Mindig létezik olyan négyzetes $\boldsymbol{G}$ és $\boldsymbol{H}$ mátrix, hogy a $\boldsymbol{G}$ mátrix oszlopai és a $\boldsymbol{H}$ mátrix sorai lineárisan függetlenek, és $\boldsymbol{A} = \boldsymbol{G}\boldsymbol{H}$.*

A bizonyítást lásd C.G. Broyden - Basic Matrices, Theorem 7.1, 124. oldal.

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

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

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.

3.4. Tétel. *Nem létezik olyan $\boldsymbol{X}$ mátrix, hogy $\boldsymbol{AX} = \boldsymbol{I}$, ha $\boldsymbol{A}$ mátrix sorai lineárisan összefüggők.*

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. □

Definíció. Legyen $\boldsymbol{A}$ mátrix. Ha létezik olyan $\boldsymbol{X}$ mátrix, hogy $\boldsymbol{AX} = \boldsymbol{I}$, akkor $\boldsymbol{X}$ az $\boldsymbol{A}$ mátrix jobboldali inverze. Ha létezik olyan $\boldsymbol{Y}$ mátrix, hogy $\boldsymbol{YA} = \boldsymbol{I}$, akkor $\boldsymbol{Y}$ az $\boldsymbol{A}$ mátrix baloldali inverze.

Definíció. Legyen $\boldsymbol{A}$ mátrix. Ha létezik olyan $\boldsymbol{X}$ mátrix, hogy $\boldsymbol{AX} = \boldsymbol{XA} = \boldsymbol{I}$, akkor $\boldsymbol{X}$ az $\boldsymbol{A}$ mátrix inverze. Az $\boldsymbol{A}$ mátrix inverzének jelölése $\boldsymbol{A}^{-1}$.

Egy mátrix invertálható, ha létezik inverze.

3.5. Tétel. *Ha egy négyzetes mátrixnak vannak lineárisan független oszlopai, akkor jobboldali inverze egyértelmű.*

3.6. Tétel. *Ha egy négyzetes mátrixnak létezik egyértelmű jobboldali inverze, akkor az megegyezik a baloldali inverzével.*

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.

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.

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

3.8. Tétel. *Legyen $\boldsymbol{A}$ négyzetes mátrix, ahol*

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

és $\boldsymbol{A}_{11}$ és $\boldsymbol{A}_{22}$ is négyzetes mátrixok. Ekkor $\boldsymbol{A}$ akkor és csak akkor nemszinguláris, ha $\boldsymbol{A}_{11}$ és $\boldsymbol{A}_{22}$ nemszinguláris.

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

3.9. Tétel. *Ha az $\boldsymbol{A}$ mátrixnak vannak lineárisan független oszlopai, akkor $\boldsymbol{A}^T\boldsymbol{A}$ pozitív definit.*

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.

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

$$|\boldsymbol{x}^T\boldsymbol{y}| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2 \tag{3.8}$$

egyenlőtlenség akkor és csak akkor igaz, ha $\boldsymbol{y}$ vektor skalárszorosa $\boldsymbol{x}$ vektornak.

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

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

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

$$\boldsymbol{A} = \max_{x \neq 0} \frac{\|\boldsymbol{Ax}\|_p}{\|\boldsymbol{x}\|_p}, \tag{3.9}$$

ahol $p = 1{,}2$ vagy ∞.

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

$$\|\boldsymbol{A}\|_1 = \max_j \sum_i |a_{ij}| \tag{3.10}$$

képlettel számolhatjuk. Az l_∞ indukált mátrixnormát az

$$\|\boldsymbol{A}\|_\infty = \max_i \sum_j |a_{ij}| \tag{3.11}$$

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

$$\|\boldsymbol{A}\|_F = (\sum_{i,j} |a_{ij}|^2)^{\frac{1}{2}}. \tag{3.12}$$

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.

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 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?

A gyakorlatban előforduló együtthatómátrixok általában vagy kitöltöttek és alacsony rendszámúak (pl. a rendszám kisebb mint 30), 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.

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.

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.

A Gauss-féle elimináció

Az LU felbontás

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. Ezután 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 egyszerűsödik, 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 3.7-es tételből következik, hogy $u_{ii} \neq 0$. A (4.4) egyenletrendszer $\boldsymbol{y}$ vektor ismeretében hasonlóan oldható meg.

4.1. Tétel. *Az $\boldsymbol{A}$ nemszinguláris mátrix felbontható $\boldsymbol{L}\boldsymbol{U}$ mátrixok szorzatára, ahol $\boldsymbol{L}$ alsó háromszögmátrix, $\boldsymbol{U}$ felső háromszögmátrix, ha $\boldsymbol{A}$ minden főátlójában lévő részmátrixa nemszinguláris.*

Bizonyítás. Ide kell bizonyítás. □

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.

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

$$\boldsymbol{P}\boldsymbol{x} = \boldsymbol{b} + \boldsymbol{Q}\boldsymbol{x} \tag{4.7}$$

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

$$\boldsymbol{P}\boldsymbol{x}_{i+1} = \boldsymbol{b} + \boldsymbol{Q}\boldsymbol{x}_i \tag{4.8}$$

egyenletrendszer sorozatos megoldásával.

4.2. Tétel. *A (4.8) egyenletrendszerből képzett $\{\boldsymbol{x}_i\}$ sorozat a (4.1) egyenletrendszer megoldásához konvergál, ha $\left\|\boldsymbol{P}^{-1}\boldsymbol{Q}\right\| < 1$, tetszőleges normával.*

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 (4.7) összefüggést kivonva a (4.8) 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 $\|\boldsymbol{e}_i\| \leq \alpha^i \|\boldsymbol{e}_0\|$. 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 $\boldsymbol{Q}$ mátrix 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: $\boldsymbol{P} = \omega^{-1}\boldsymbol{D} - \boldsymbol{L}, \quad \boldsymbol{Q} = \boldsymbol{U} + (\omega^{-1} - 1)\boldsymbol{D}$.

A felső relaxációs eljárással külön fejezetben részletesen foglalkozunk. 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.

4.3. Sajátértékek és sajátvektorok

Fontos jellemzője még az egyenletrendszernek az $\boldsymbol{A}$ mátrix a sajátértékei és sajátvektorai

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

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}, \tag{4.9}$$

az $\boldsymbol{A}$ mátrix jobboldali sajátvektora.

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

$$\boldsymbol{y}^T(\boldsymbol{A} - \lambda \boldsymbol{I}) = \boldsymbol{0}, \tag{4.10}$$

az $\boldsymbol{A}$ mátrix baloldali sajátvektora.

Ha az $\boldsymbol{A}$ mátrixszal szorozzuk a sajátvektorát, a sajátvektor mérete változhat, iránya azonban nem változik, legfeljebb ellenkezőjére fordul. Azaz

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \tag{4.11}$$

a (4.9) jelölésekkel. A sajátértékek és sajátvektorok a mátrixok fontos jellemzői.

4.3. Tétel. *Minden négyzetes mátrixnak van legalább egy sajátértéke.*

Valós mátrixokhoz gyakran tartozik komplex sajátérték és sajátvektor.

4.4. Tétel. *Legyen $\boldsymbol{A}$ komplex négyzetes mátrix. Ha λ sajátértéke, $\boldsymbol{x}$ sajátvektora $\boldsymbol{A}$-nak, akkor $\overline{\lambda}$ sajátértéke, $\overline{\boldsymbol{x}}$ sajátvektora $\boldsymbol{A}$-nak.*

4.5. Tétel. *A valós szimmetrikus mátrixnak minden sajátértéke valós.*

4.6. Tétel. *A Hermite-féle mátrixnak minden sajátértéke valós.*

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

Definíció. Legyen $\boldsymbol{A}$ és $\boldsymbol{P}$ n-ed rendű mátrix. Legyen $\boldsymbol{P}$ nemszinguláris. A $\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}$ transzformáció hasonlósági transzformáció, az $\boldsymbol{A}$ és $\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}$ mátrixok hasonlóak.

4.7. Tétel. *A hasonlósági transzformáció nem változtatja meg a mátrix sajátértékeit.*

Definíció. Legyen $\boldsymbol{A}$ és $\boldsymbol{P}$ n-ed rendű mátrix. Legyen $\boldsymbol{P}$ nemszinguláris. A $\boldsymbol{PAP}^{-1}$ hasonlósági transzformáció ortogonális transzformáció, ha $\boldsymbol{P}$ mátrix ortogonális.

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

Definíció. Legyen $\boldsymbol{A}$ és $\boldsymbol{P}$ n-ed rendű mátrix. Legyen $\boldsymbol{P}$ nemszinguláris. A $\boldsymbol{PAP}^{-1}$ hasonlósági transzformáció unitér transzformáció, ha $\boldsymbol{P}$ mátrix unitér..

4.8. Tétel. *Minden n-ed rendű $\boldsymbol{A}$ mátrixhoz létezik egy olyan $\boldsymbol{Q}$ unitér mátrix, hogy $\boldsymbol{Q}^H\boldsymbol{AQ} = \boldsymbol{U}$, ahol $\boldsymbol{U}$ felső háromszögmátrix.*

Következmény. *Minden n-ed rendű mátrixhoz n darab sajátérték tartozik.*

Az n darab sajátérték lehet nem különböző. Minden n-ed rendű négyzetes mátrixot hasonlósági transzformációval felső háromszögmátrixszá alakíthatunk, melynek sajátértékei a főátló elemei. A felső háromszögmátrix sajátértékei a főátló elemei.

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

4.9. Tétel. *Legyen $\boldsymbol{A}$ n-ed rendű mátrix és tartozzon hozzá m darab különböző sajátérték, $m \leq n$. Az $\boldsymbol{A}$ mátrixhoz legalább m lineárisan független sajátvektor tartozik.*

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

4.10. Tétel. *Legyen $\boldsymbol{A}$ négyzetes mátrix. Ekkor $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|$ tetszőleges normával.*

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

4.11. Tétel. *A mátrixhoz tartozó sajátértékek összege egyenlő a mátrix nyomával.*

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.

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

4.12. Tétel. *Legyen $\boldsymbol{U}$ felső háromszögmátrix és $\rho(\boldsymbol{U}) < 1$. Ekkor létezik $\boldsymbol{B}$ diagonális mátrix, amelyre $\left\|\boldsymbol{BUB}^{-1}\right\|_\infty < 1$ és $\left\|\boldsymbol{BUB}^{-1}\right\|_1 < 1$.*

4.13. Tétel. *Az $\boldsymbol{A}$ mátrix konvergenciájának szükséges és elégséges feltétele, hogy $\rho(\boldsymbol{A}) < 1$.*

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

11. Felhasznált irodalom

[1] C.G. Broyden, *Basic Matrices*, The Macmillan Press Ltd, 1975.