

Tartalomjegyzék

1. Bevezetés	2
2. Broyden életrajza	2
3. Mátrixok alapvető jellemzése	3
3.1. Vektorok és mátrixok	3
3.2. Lineáris függetlenség	5
3.3. Nemszinguláris mátrixok	6
3.4. Sajátértékek és sajátvektorok	7
4. Lineáris egyenletrendszerek megoldása, osztályozása	7
5. Felső relaxációs eljárás (SOR)	7
6. Blokk konjugált gradiens módszer (BICG)	7
7. A konjugált gradiens módszerek új rendszertana	7
8. MATLAB tesztfeladatokon való összehasonlítás	7
9. Konklúzió	7
10. Appendix	7

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 Fletcher 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 módszerekben dolgozott. Később a numerikus lineáris algebrára fókuszált, ezen belül is a konjugált gradiens módszerekre és ezek rendszer-

tanára. 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 segítségért.

3. Mátrixok alapvető jellemzése

Ebben a fejezetben a további fejezetekhez szükséges fogalmakat vezetjük be és tisztázzuk a jelölésüket.

3.1. Vektorok és mátrixok

Definíció. A valós vektor a valós számok egy rendezett halmaza. A halmaz elemeinek a száma a vektor rendje, vagy más szóval a vektor dimenziója.

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

Definíció. Legyen $\mathbf{x} = [x_i]$ és $\mathbf{y} = [y_i]$ n -ed rendű vektor. A belső szorzata a sorvektor \mathbf{x}^T -nek és az oszlopvektor \mathbf{y} -nak

$$\mathbf{x}^T \mathbf{y} = \sum_{i=1}^n x_i y_i.$$

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

Definíció. A valós mátrix a valós vektorok egy rendezett halmaza.

A mátrixokat a szakdolgozatban vastag nagybetűvel jelöljük. Az $\mathbf{A} = [a_{ij}]$ jelölésben a_{ij} az \mathbf{A} mátrix i -edik sorának j -edik eleme. Ha a mátrix $m \times n$ -es, $i = 1, 2, \dots, m$ és $j = 1, 2, \dots, n$. Az $\mathbf{A} = [a_{ij}]$ mátrix transzponáltját \mathbf{A}^T -vel jelöljük, jelentése $\mathbf{A}^T = [a_{ji}]$, 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 csupa nullából álló mátrixot, vektort $\mathbf{0}$ -val jelöljük.

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

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

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

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

Definíció. Legyenek \mathbf{p}_i vektorok, y_i skalárok, $i = 1, 2, \dots, n$. A $\sum_{i=1}^n \mathbf{p}_i y_i$ vektor a \mathbf{p}_i vektorok lineáris kombinációja.

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

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

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

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

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 definiált.

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

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

A szorzás szerint \mathbf{E}_1 és \mathbf{F}_1 nem jól particionáltak, mert

$$\mathbf{E}_1 \mathbf{F}_1 \neq \begin{bmatrix} \mathbf{AJ} + \mathbf{CK} & \mathbf{BJ} + \mathbf{DK} \\ \mathbf{AL} + \mathbf{CM} & \mathbf{BL} + \mathbf{DM} \end{bmatrix}.$$

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

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

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

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

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 \mathbf{w} vektor konjugáltját $\overline{\mathbf{w}}$ -vel, az \mathbf{A} mátrix konjugáltját $\overline{\mathbf{A}}$ -val jelöljük. Defináljuk a belső szorzatot komplex vektorokra.

Definíció. A $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ komplex vektor Hermite-féle transzponáltja $\mathbf{z}^H = \overline{\mathbf{z}}^T = \mathbf{x}^T - i\mathbf{y}^T$.

Definíció. Az $\mathbf{A} = \mathbf{B} + i\mathbf{C}$ komplex mátrix Hermite-féle transzponáltja $\mathbf{A}^H = \overline{\mathbf{A}}^T = \mathbf{B}^T - i\mathbf{C}^T$.

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

Definiáljuk a szimmetrikus mátrix komplex megfelelőjét.

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

3.2. Lineáris függetlenség

Definíció. Az \mathbf{a}_i ($i = 1, 2, \dots, n$) vektorok lineárisan függetlenek, ha

$$\sum_{i=1}^n \mathbf{a}_i x_i = \mathbf{0}$$

csak $x_i = 0$ ($i = 1, 2, \dots, n$) skalárokkal teljesül. Ellenkező esetben az \mathbf{a}_i vektorok lineárisan összefüggők.

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

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

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

-ből 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 \mathbf{A} $m \times n$ -es mátrix oszlopai lineárisan összefüggők, ha létezik olyan n -ed rendű $\mathbf{x} \neq \mathbf{0}$ vektor, hogy $\mathbf{Ax} = \mathbf{0}$.

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

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

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.

3.3. Nemszinguláris mátrixok

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

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

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

Definíció. Legyen \mathbf{A} mátrix. Ha létezik olyan \mathbf{X} mátrix, hogy $\mathbf{XA} = \mathbf{I}$, akkor \mathbf{X} az \mathbf{A} mátrix baloldali inverze.

Definíció. Legyen \mathbf{A} mátrix. Ha létezik olyan \mathbf{A}^{-1} mátrix, hogy $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, akkor \mathbf{A}^{-1} az \mathbf{A} mátrix inverze.

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

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

Definíció. Legyen \mathbf{A} négyzetes mátrix. A következő állítások ekvivalensek.

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

Ha a fenti ekvivalens állítások teljesülnek, az \mathbf{A} mátrix nemszinguláris, egyébként szinguláris.

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

- 3.4. Sajátértékek és sajátvektorok
- 4. Lineáris egyenletrendszerek megoldása, osztályozása
- 5. Felső relaxációs eljárás (SOR)
- 6. Blokk konjugált gradiens módszer (BICG)
- 7. A konjugált gradiens módszerek új rendszertana
- 8. MATLAB tesztfeladatokon való összehasonlítás
- 9. Konklúzió
- 10. Appendix