Tartalomjegyzék

1. Bevezetés **2**

2. Broyden életrajza **2**

3. Mátrixok alapvető jellemzése **3**

3.1. Vektorok és mátrixok 3

3.2. Lineáris függetlenség 5

4. Lineáris egyenletrendszerek megoldása, osztályozása **5**

5. Felső relaxációs eljárás (SOR) **5**

6. Blokk konjugált gradiens módszer (BlCG) **5**

7. A konjugált gradiens módszerek új rendszertana **5**

8. MATLAB tesztfeladatokon való összehasonlítás **5**

9. Konklúzió **5**

10. Appendix **5**

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

3.1. Vektorok és mátrixok

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

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.: $\boldsymbol{x} = [x_i]$, 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 vektor. A belső szorzata a sorvektor $\mathbf{x}^T$-nek és az oszlopvektor $\boldsymbol{y}$-nak

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

Definíció. Az $\boldsymbol{x}$ és $\boldsymbol{y}$ 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 $\boldsymbol{A} = [a_{ij}]$ 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 $j = 1, 2, ..., n$. 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.

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 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, ahol $y_i = \boldsymbol{a}_i^T\boldsymbol{x}$ az $\boldsymbol{A}$ mátrix és $\boldsymbol{x}$ vektor mátrix-vektor szorzata. 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 $c_{ij} = \boldsymbol{a}_i^T\boldsymbol{b}_j$.

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 $\sum_{i=1}^n \boldsymbol{p}_i y_i$ 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 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 particionáltak jól, mert

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

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 particionáltak jól, 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 $\boldsymbol{w}$ vektor konjugáltját $\overline{\boldsymbol{w}}$-vel, az $\boldsymbol{A}$ mátrix konjugáltját $\overline{\boldsymbol{A}}$-val jelöljük. Definiáljuk a belső szorzatot komplex vektorokra.

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

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

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

Definiáljuk a szimmetrikus mátrix komplex megfelelőjét.

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

3.2. Lineáris függetlenség

4. Lineáris egyenletrendszerek megoldása, osztályozása

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. Appendix