

Eötvös Loránd Tudományegyetem

Természettudományi kar

Lineáris algebrai egyenletrendszerek iteratív megoldási módszerei

Szakdolgozat



Készítette:
Kis Ágnes

Matematika Bsc.
Matematikai elemző szakirány

Témavezető:
Dr. Faragó István
egyetemi tanár
Alkalmazott Analízis és
Számításmatematikai Tanszék

Budapest
2012

Tartalomjegyzék

Köszönetnyilvánítás	3
1. A lineáris egyenletrendszerek algebrai alapjai	4
1.1. Lineáris egyenletrendszerek megoldhatósága	4
1.2. Megoldási módszerek – Direkt módszerek	5
1.2.1. Gauss-elimináció	5
1.2.2. Cramer-szabály	8
1.3. A direkt módszerek műveletigénye	10
2. Az iterációs eljárások	12
2.1. Iterációs vagy direkt módszer?	12
2.2. Klasszikus iterációk	14
2.3. Az iterációs módszerek konvergenciája	17
3. Relaxációs módszerek	24
3.1. A JOR-módszer	24
3.2. A Gauss–Seidel-iteráció relaxált változatai	26
3.3. A módszerek összehasonlítása egy példán	28
Összefoglalás	33
Nyilatkozat	35

Köszönetnyilvánítás

Ezúton szeretném megköszönni témavezetőmnek, Faragó Istvánnak, hogy szakértelmével és hasznos tanácsaival hozzájárult dolgozatom elkészítéséhez, s lelkes hozzáállásával engem is ösztönzött arra, hogy a felmerülő problémákat továbbgondoljam, feldolgozzam.

Köszönettel tartozom továbbá családomnak és szeretteimnek, akik támogattak, bíztattak és nyugodt háttérrel biztosítottak számomra egyetemi éveim alatt.

1. fejezet

A lineáris egyenletrendszerek algebrai alapjai

1.1. Lineáris egyenletrendszerek megoldhatósága

A matematikában és az élet különböző területein is számos olyan problémával találkozhatunk, amelyek lineáris algebrai egyenletrendszerek megoldására vezetnek. Ilyen feladatok előfordulhatnak többek között a mechanikában, villamosságban, ökológiai, gazdasági és más vizsgálatok során. Tipikus példa az ipari termelés, mikor különböző termékeket szeretnénk előállítani különböző nyersanyagokból, vagy a különböző hálózatok (mint például villamos hálózatok, víz- vagy gázellátó csőrendszerek, vérkeringés) modellezése is példát ad arra, hogy milyen alkalmazásokból származnak a lineáris egyenletrendszerek.

A numerikus matematika több feladatát is ilyen lineáris rendszerek megoldására vezethetjük vissza. A nemlineáris rendszerek megoldásához lineáris egyenletrendszereket kell megoldanunk, a differenciál- és integrálegyenletek, valamint az interpolációs és optimalizációs feladatok közelítő megoldása is lineáris rendszerekkel kapcsolatos. A kétpontos peremérték feladatok és a parciális differenciálegyenletek közül az elliptikus egyenletek közelítésének módszere is visszavezethető egy lineáris algebrai egyenletrendszerre.

A lineáris algebrai egyenletrendszerek általános alakja a következő. Tegyük fel, hogy adottak a_{ij} és b_i számok, és keressük azon x_i számokat, melyekre teljesülnek a következő egyenletek, ha $i = 1, 2 \dots m$ és $j = 1, 2 \dots n$.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

A feladatot mátrixos alakban is megfogalmazhatjuk. Legyen \mathbf{b} m dimenziós vektor, és \mathbf{A} $m \times n$ -es mátrix. Keressük azt az \mathbf{x} n dimenziós vektort, mely kielégíti az

$\mathbf{Ax} = \mathbf{b}$ egyenletet. Az a_{ij} együtthatókból képzett \mathbf{A} mátrixot az egyenletrendszer *együtthatómátrixának* nevezzük. Ha ezt kibővítjük a \mathbf{b} vektor b_i komponenseiből képzett oszlopvektorral, akkor az egyenletrendszer $m \times (n + 1)$ -es *kibővített mátrixát* kapjuk, amit $\mathbf{A|\underline{b}}$ -vel jelölünk.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \mathbf{A|\underline{b}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{pmatrix}$$

Lineáris algebrából jól ismert a következő tétel:

1.1.1. Tétel. ([2] Freud, 2007) *Egy $\mathbf{Ax} = \mathbf{b}$ lineáris egyenletrendszer akkor és csak akkor megoldható, ha az \mathbf{A} együttható mátrix és az $\mathbf{A|\underline{b}}$ kibővített mátrix rangja megegyezik: $r(\mathbf{A}) = r(\mathbf{A|\underline{b}})$. Megoldhatóság esetén a megoldás akkor és csak akkor egyértelmű, ha a (közös) rang megegyezik az ismeretlenek számával: $r(\mathbf{A}) = r(\mathbf{A|\underline{b}}) = n$.*

A továbbiakban feltesszük, hogy az egyenletrendszer mátrixa négyzetes. Ekkor az, hogy az egyenletrendszernek pontosan egy megoldása van, azzal ekvivalens, hogy az együtthatómátrix determinánsa nem egyenlő nullával, illetve a mátrix teljes rangú. Azaz, ha $\mathbf{A} \in \mathbb{R}^{n \times n}$, akkor:

$$\exists! \text{ megoldás} \Leftrightarrow \det \mathbf{A} \neq 0 \Leftrightarrow r(\mathbf{A}) = n.$$

Ekkor az \mathbf{A} mátrixnak létezik inverze, és az egyenletrendszer megoldása $\underline{\mathbf{x}} = \mathbf{A}^{-1}\underline{\mathbf{b}}$.

1.2. Megoldási módszerek – Direkt módszerek

A lineáris egyenletrendszerek megoldási módszereit két nagy csoportba sorolhatjuk. Ezek az úgynevezett direkt módszerek, és az iterációs eljárások. A direkt módszerek esetén a megoldást véges sok aritmetikai művelet segítségével állítjuk elő. Ha számolásunk minden egyes lépésben pontos, akkor véges számú lépésben az egyenletrendszer pontos megoldásához jutunk. Ilyen direkt eljárások például a Cramer-szabály és a Gauss-elimináció.

1.2.1. Gauss-elimináció

A Gauss-elimináció vagy más néven Gauss-kiküszöbölés a leggyakrabban használt direkt megoldó módszer. A módszer két részből áll, egy úgynevezett eliminációs részből és a visszahelyettesítésből. Az első lépésben *elemi ekvivalens átalakításokkal* "kiküszöböljük", az egyenletrendszerből az ismeretleneket. Azaz az egyenletrendszert olyan alakra hozzuk, hogy az utolsó egyenletben csak az utolsó ismeretlen szerepeljen, az utolsó előttiben az utolsó kettő, és így tovább. Ekkor az együtthatómátrix egy felső háromszögmátrix lesz. Az *elemi ekvivalens átalakítások*, amiket az elimináció során használhatunk, a következők:

- Valamelyik egyenletet egy nullától különböző skalárral végigszorozzuk.
- Valamelyik egyenlethez hozzáadjuk egy másik skalárszorosát.
- Két egyenletet felcserélünk.
- Az olyan egyenleteket, ahol minden együttható és a jobboldali konstans is nulla, elhagyjuk.

Az eliminációs eljárás után a megoldást az utolsó egyenlettől kezdve, egyszerű visszahelyettesítésekkel kaphatjuk meg.

Nézzük az eljárást most részletesen! Tegyük fel hogy az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszert szeretnénk megoldani, ahol ismert a $\mathbf{b} \in \mathbb{R}^n$ vektor és az \mathbf{A} $n \times n$ -es mátrix, amire $\det \mathbf{A} \neq 0$, s az ismeretlen $\mathbf{x} \in \mathbb{R}^n$ vektort keressük. Ekkor az egyenletrendszerünk a következőképpen néz ki:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n. \end{aligned}$$

Először az első egyenlet segítségével küszöböljük ki a többi egyenletből x_1 -et a következőképpen: tegyük fel, hogy $a_{11} \neq 0$ és osszuk végig az első egyenletet a_{11} -el, majd minden további i -ik egyenletből ($i > 1$) vonjuk ki az így kapott első egyenlet a_{i1} -szeresét. Az átláthatóság kedvéért vezessünk be néhány új jelölést! Az első egyenlet együtthatói és jobboldala legyen

$$c_{1j} = \frac{a_{1j}}{a_{11}}, \quad \text{és} \quad y_1 = \frac{b_1}{a_{11}}.$$

A többi egyenletben az együtthatók és a jobboldalak legyenek

$$a_{ij}^{(1)} = a_{ij} - c_{1j}a_{i1}, \quad \text{és} \quad b_i^{(1)} = b_i - y_1a_{i1}.$$

A felső (1) indexek azt jelölik, hogy ezeket az értékeket az első eliminációs lépés elvégzése után kaptuk. Így az alábbi egyenletrendszerhez jutottunk:

$$\begin{aligned} x_1 + c_{12}x_2 + \cdots + c_{1n}x_n &= y_1 \\ 0x_1 + a_{22}^{(1)}x_2 + \cdots + a_{2n}^{(1)}x_n &= b_2^{(1)} \\ &\vdots \\ 0x_1 + a_{n2}^{(1)}x_2 + \cdots + a_{nn}^{(1)}x_n &= b_n^{(1)} \end{aligned}$$

Most a második egyenlet segítségével hasonlóan ejtjük ki x_2 -t minden $i > 2$ -re az i -ik egyenletekből, majd x_3 -at az utolsó $n - 3$ egyenletből és így tovább. Természetesen minden lépésnél feltesszük, hogy az aktuális $a_{kk}^{(k-1)}$ együttható, amivel osztunk nem egyenlő nullával.

Lássuk, hogyan alakul az együtthatómátrix az elimináció során.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \sim \begin{pmatrix} 1 & c_{12} & c_{13} & \dots & c_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & c_{12} & c_{13} & \dots & c_{1n} \\ 0 & 1 & c_{23} & \dots & c_{2n} \\ 0 & 0 & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix} \sim \begin{pmatrix} 1 & c_{12} & c_{13} & \dots & c_{1n} \\ 0 & 1 & c_{23} & \dots & c_{2n} \\ 0 & 0 & 1 & \dots & c_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} := \mathbf{C}$$

Így az eredeti $\mathbf{Ax} = \mathbf{b}$ egyenletrendszerből a Gauss-eliminációval az ekvivalens $\mathbf{Cx} = \mathbf{y}$ egyenletrendszert kaptuk. Mivel az utóbbi egyenletrendszer \mathbf{C} mátrixa speciális alakú (felső háromszögmátrix, aminek diagonálisában csupa egyesből áll), így az egyenletrendszerből az x_i értékek ($i = n, \dots, 1$) könnyen kifejezhetők.

$$\begin{aligned} x_n &= y_n \\ x_{n-1} &= y_{n-1} - c_{n-1,n}y_n \\ x_{n-2} &= y_{n-2} - c_{n-2,n}y_n - c_{n-2,n-1}y_{n-1} \\ &\vdots \\ x_k &= y_k - \sum_{l=k+1}^n c_{k,l}y_l \end{aligned}$$

Az eljárás során minden lépésnél kihangsúlyoztuk, hogy a *főelemek*, azaz azok az $a_{kk}^{(k-1)}$ együtthatók, melyek segítségével kinulláztuk az alattuk lévő, azonos oszlopban szereplő elemeket, nem lehetnek egyenlők nullával (hiszen ha valamelyik főelem nulla lenne, akkor az eljárás nullával osztás miatt nem tudna végigfutni). De nem csak az okozhat problémát, ha egy főelem nulla, hanem az is, ha túlságosan kis abszolútértékű. Az ilyen számmal való osztás során ugyanis olyan nagy lehet a kerekítési hiba, hogy az jelentősen torzíthatja a végeredményt. Ezeknek az elkerülésére szolgál az úgynevezett *főelemkiválasztás* módszere.

Tegyük fel, hogy a Gauss-elimináció során az $a_{kk}^{(k-1)}$ együttható nulla (vagy nagyon kis abszolútértékű). Ekkor alkalmazhatjuk a következő módszerek egyikét:

- **Részleges főelemkiválasztás.** Az $a_{kk}^{(k-1)}$ főelemmel azonos oszlopban lévő és nagyobb sorindexű elemek közül kiválasztjuk a legnagyobb abszolútértékűt, és ezt sorcserével a főátlóba hozzuk.

- **Teljes főelemkiválasztás.** Az $a_{ij}^{(k-1)}$ elemek közül (ahol $i, j \geq k$) kiválasztjuk a legnagyobb abszolútértékűt, majd sor- és oszlopcsérékkel a főátlóba transzformáljuk.

Tehát láthatjuk, hogy a főelemkiválasztás a mátrix sorainak és oszlopainak megfelelő átrendezését jelenti. Ezzel a módszerrel minden reguláris mátrix átrendezhető úgy, hogy a Gauss-elimináció végrehajtható legyen.

De milyen esetekben nincs szükség főelemkiválasztásra? Azaz milyen mátrixok esetén teljesül automatikusan, hogy a Gauss-elimináció során egyik főelem sem lesz nulla? Erre a kérdésre ad választ a következő tétel.

1.2.1. Tétel. *A Gauss-módszer akkor és csak akkor hajtható végre főelemkiválasztás nélkül, ha az \mathbf{A} mátrix egyik főminorja sem nulla, azaz*

$$\det(\mathbf{A}(1:k, 1:k)) \neq 0 \quad \forall k = (1, \dots, n).$$

Bizonyítás. Mivel a Gauss-elimináció során az egyes sorokból más sorok számszorosait vonjuk ki, a determináns nem változik. Az egyes főminorok:

$$\begin{aligned} \det(a_{11}) &= a_{11} \neq 0, \\ \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \det \begin{pmatrix} 1 & c_{12} \\ 0 & a_{22}^{(1)} \end{pmatrix} = a_{22}^{(1)} \neq 0, \\ &\vdots \\ \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1k} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \dots & a_{kk} \end{pmatrix} &= \det \begin{pmatrix} 1 & c_{12} & c_{13} & \dots & c_{1k} \\ 0 & 1 & c_{23} & \dots & c_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{kk}^{(k-1)} \end{pmatrix} = a_{kk}^{(k-1)} \neq 0. \end{aligned}$$

Tehát az, hogy az egyes főminorok nem egyenlők nullával, éppen azt jelenti, hogy a főelemek nem nullák. És éppen ez kellett ahhoz, hogy a Gauss-elimináció végrehajtható legyen. ■

1.2.2. Megjegyzés. A fenti (1.2.1) tétel teljesül, azaz a Gauss-elimináció során nincs szükség főelemkiválasztásra, hogyha az \mathbf{A} mátrix:

- szigorúan diagonálisan domináns, vagy
- szimmetrikus, pozitív definit, vagy
- M-mátrix.

1.2.2. Cramer-szabály

Egy másik ismert direkt megoldási módszer a Cramer-szabály. Ez a szabály megmutatja, hogy egy olyan négyzetes mátrixú egyenletrendszernek, melynek pontosan egy megoldása van, azaz az együtthatómátrix determinánsa nem nulla, hogyan

kaphatjuk meg determinánsok segítségével a megoldását. Látni fogjuk, hogy ez a fajta megoldási módszer nagyon sok számolást igényel, ráadásul csak speciális esetekben alkalmazható, ezért a Cramer-szabálynak elsősorban elméleti jelentősége van, a gyakorlatban ritkán használható.

1.2.3. Tétel. (Cramer-szabály) *Ha $\mathbf{A} \in \mathbb{R}^{n \times n}$ és $\det \mathbf{A} \neq 0$, akkor az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszernek pontosan egy megoldása létezik. A megoldásban $x_j = \frac{\det \mathbf{A}_j}{\det \mathbf{A}}$, ahol az \mathbf{A}_j mátrixot úgy kapjuk, hogy az \mathbf{A} mátrixban a j -edik oszlop helyére a \mathbf{b} vektor komponenseit írjuk.*

Például:

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & \dots & a_{1n} \\ a_{21} & b_2 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \dots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}}$$

Bizonyítás. Az egyenletrendszernek valóban csak egy megoldása létezik, hiszen a feltétel szerint \mathbf{A} determinánsa nem nulla, amiből következik, hogy létezik \mathbf{A}^{-1} , tehát az egyenletrendszert balról \mathbf{A}^{-1} -zel szorozva ekvivalens egyenletrendszert nyerünk. Ez az $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ egyenletrendszer, ami tulajdonképpen már „meg is van oldva”. Ezzel beláttuk, hogy az eredeti egyenletrendszernek pontosan (ez az) egy megoldása van.

Tehát a tétel igazolásához azt kell megmutatni, hogy $\frac{\det \mathbf{A}_j}{\det \mathbf{A}}$ kifejezés valóban x_j -t adja. Legyen most $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$, ahol $\mathbf{a}_i = [a_{1i}, a_{2i}, \dots, a_{ni}]^T$ oszlopvektort jelöli. Ekkor

$$\mathbf{A}_j = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n] \quad \forall j \in \{1, \dots, n\}.$$

Az $\mathbf{Ax} = \mathbf{b}$ azt jelenti, hogy $\mathbf{a}_1 x_1 + \dots + \mathbf{a}_n x_n = \mathbf{b}$. Így (például) \mathbf{A}_2 determinánsa a következőképpen írható fel:

$$\begin{aligned} \det \mathbf{A}_2 &= \det[\mathbf{a}_1, \mathbf{b}, \dots, \mathbf{a}_n] = \det[\mathbf{a}_1, (\mathbf{a}_1 x_1 + \dots + \mathbf{a}_n x_n), \dots, \mathbf{a}_n] = \\ &= \det[\mathbf{a}_1, (\mathbf{a}_1 x_1), \dots, \mathbf{a}_n] + \det[\mathbf{a}_1, (\mathbf{a}_2 x_2), \dots, \mathbf{a}_n] + \dots + \det[\mathbf{a}_1, (\mathbf{a}_n x_n), \dots, \mathbf{a}_n] = \\ &= x_1 \det[\mathbf{a}_1, \mathbf{a}_1, \dots, \mathbf{a}_n] + x_2 \det[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n] + \dots + x_n \det[\mathbf{a}_1, \mathbf{a}_n, \dots, \mathbf{a}_n] = \\ &= x_2 \det[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n] = \underline{x_2 \det \mathbf{A}} \end{aligned}$$

(Az utolsó sort azért kaptuk, mert az előtte lévő összegben (a második kivételével) a determinánsok értéke nulla lett, mert a mátrixoknak volt egy-egy összefüggő oszlopa.)

A fenti egyenletből tehát megkaptuk, hogy $x_2 = \frac{\det \mathbf{A}_2}{\det \mathbf{A}}$, ami a többi indexre is hasonlóan látható. ■

A Cramer-szabálynál nagyon fontos feltétel, hogy az \mathbf{A} determinánsa ne legyen nulla. Eleve értelmetlen az x_j képletében $\det \mathbf{A} = 0$ -val osztani, de ennél több is igaz: ha $\det \mathbf{A} = 0$, akkor az egyenletrendszernek semmiképpen sem létezik egyértelmű megoldása.

1.2.4. Tétel. *Ha $\mathbf{A} \in \mathbb{R}^{n \times n}$ és $\det \mathbf{A} = 0$, akkor az $\mathbf{A}\mathbf{x} = \mathbf{b}$ egyenletrendszer vagy nem oldható meg, vagy pedig egynél több megoldása van.*

Most, hogy már tudjuk, hogyan lehet a Cramer-szabály segítségével egyenletrendszereket megoldani, láthatjuk, hogy ez a fajta megoldási módszer nagy mátrixok esetén rengeteg számolást igényel. Ám előfordulhat olyan szélsőséges példa, amikor mégis érdemesebb ezt a módszert választani, mint a Gauss-eliminációt; például, ha nem vagyunk kíváncsiak az egész \mathbf{x} vektorra, hanem csak néhány, esetleg egyetlen komponensét szeretnénk meghatározni.

1.3. A direkt módszerek műveletigénye

Fontos megvizsgálunk, hogy egy-egy eljárásnak mekkora a műveletigénye, hiszen a lépésszám ismerete segíthet abban, hogy egy konkrét feladatnál eldönthessük, hogy melyik megoldási módszert érdemes alkalmazni.

Vizsgáljuk meg, hogy mennyi a Gauss-módszer egyes lépéseinek műveletigénye! Vegyük először az eliminációs lépést. Tegyük fel, hogy a k -edik főelemmel nullázunk. Ekkor a k -edik sorban a j -edik tagokat ($j = k + 1, \dots, n$) el kell osztani $a_{kk}^{(k-1)}$ -el, ez $(n - k)$ lépés. Ezután ha az i -edik sorban akarunk nullázni, akkor először a k -edik sor nemnulla elemeit kell végigszorozni $\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$ -val, ez $(n - k + 1)$ szorzás, majd az i -edik sor nemnulla elemeiből kell kivonni a k -edik sor így kapott megfelelő elemeit, ez szintén $(n - k + 1)$ művelet. Tehát ez egy-egy soron $2(n - k + 1)$ műveletet jelent, amit összesen $(n - k)$ soron kell majd végrehajtani, mert $i = k + 1, \dots, n$.

Tehát amikor a k -edik főelemmel nullázunk, akkor összesen $(n - k) + 2(n - k + 1)(n - k)$ lépést hajtunk végre. Ezt kell szummázni az összes főelemre. Így az elimináció lépésszáma:

$$\begin{aligned} \sum_{k=1}^n (2(n - k + 1)(n - k) + n - k) &= \sum_{k=1}^n (2(n - k)^2 + 3(n - k)) = \\ &= 2 \frac{(n - 1)n(2n - 1)}{6} + 3 \frac{(n - 1)n}{2} = \frac{4n^3 + 3n^2 - 7n}{6} = \frac{2}{3}n^3 + O(n^2). \end{aligned}$$

Itt felhasználtuk az első n természetes szám összegére és négyzetösszegére vonatkozó képleteket. A visszahelyettesítés műveletigényéről pedig könnyen látható, hogy éppen $1 + 3 + \dots + 2n - 1 = n^2$. Tehát a Gauss-elimináció teljes műveletigénye:

$$\frac{2}{3}n^3 + O(n^2).$$

1.3.1. Megjegyzés. Nagy méretű egyenletrendszerek esetén a visszahelyettesítés lépésszáma elhanyagolható az elimináció műveletigényéhez képest.

A Cramer-szabály műveletigénye ennél jóval nagyobb, hiszen $(n + 1)$ különböző determinánst kell kiszámítanunk. Mivel egy-egy determináns kiszámításához $n!$ számú művelet elvégzése szükséges, így a Cramer-szabály teljes műveletigénye:

$$O(n \cdot n!).$$

2. fejezet

Az iterációs eljárások

2.1. Iterációs vagy direkt módszer?

A gyakorlati feladatoknál a direkt módszerek fő hibája, hogy megoldásuk nagyon sok számolással jár. Ezenkívül felmerül a kérdés, hogy érdemes-e a pontos megoldást kiszámolnunk, amikor általában a mátrix is és a jobb oldal is hibásan van megadva (mérési hibák, szándékosan rosszul közölt adatok).

A klasszikus iterációs módszerek lényege, hogy egy olyan konvergens sorozatot konstruálnak, melynek a határértéke az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer egyértelmű megoldása. Ezt \mathbf{x}^* -gal fogjuk jelölni. Lineáris egyenletrendszerek esetén lineáris iterációkkal foglalkozunk, melyek általában felírhatók

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}, \quad k = 0, 1, \dots \quad (2.1)$$

alakban, ahol $\mathbf{x}^{(0)}$ adott és a \mathbf{B} mátrix függhet k -től. Azt várjuk, hogy az $(\mathbf{x}^{(k)})$ sorozat tartson az \mathbf{x}^* megoldáshoz. A felírásból jól látszik, hogy egy iterációs lépés egy mátrix-vektor szorzásból és egy vektorösszeadásból áll, aminek a műveletigénye $2n^2$. Ez azt jelenti, hogy egy egyenletrendszer megoldása során maximum $\frac{n}{3}$ iterációs lépést hajthatunk végre, ha nem akarjuk meghaladni a Gauss-elimináció $\frac{2}{3}n^3 + O(n^2)$ -es lépésszámát. Ez például egy 100×100 -as mátrix esetében mindössze 33 lépést jelent. Ám annak az esélye, hogy tetszőleges kezdővektorral ennyi lépés alatt elég közel kerüljünk a megoldáshoz, igen kicsi. Emiatt az iterációs módszereket csak olyan egyenletrendszerek megoldására használjuk, melyek együtthatómátrixa ún. *ritka mátrix*, azaz olyan mátrix, melyben a nemnulla elemek száma $O(n)$ nagyságrendű. A Gauss-elimináció nem tudja kihasználni a ritka mátrixoknak ezt a jó tulajdonságát, mert az eljárás során a mátrix feltöltődik. Az iteratív módszereknél viszont a mátrix-vektor szorzást nagyban megkönnyíti a sok nullelem. A differenciálegyenletek megoldása során gyakran kapunk ritka mátrixokat, így ezekben az esetekben jól alkalmazhatók az iterációs módszerek.

Az iterációs eljárásokkal kapcsolatban felmerülhetnek a következő kérdések:

- Hogyan válasszuk meg a \mathbf{B} iterációs mátrixot?
- Hogyan válasszuk az \mathbf{f} vektort és az $\mathbf{x}^{(0)}$ kezdeti vektort?

- Mikor fog konvergálni a sorozat a megoldáshoz?
- Mekkora lesz a konvergencia sebessége?
- Mikor kell leállni az iterációval?

Most térjünk vissza a (2.1) iterációs eljáráshoz és vizsgáljuk meg, hogyan juthatunk $\mathbf{Ax} = \mathbf{b}$ egyenletrendszerből a \mathbf{B} iterációs mátrixhoz és \mathbf{f} vektorhoz. Legyen

$$\mathbf{A} = \mathbf{P} - \mathbf{Q},$$

ahol \mathbf{P} reguláris mátrix. Ekkor

$$\begin{aligned}\mathbf{b} &= \mathbf{Ax} = \mathbf{Px} - \mathbf{Qx} \\ \mathbf{Px} &= \mathbf{Qx} + \mathbf{b},\end{aligned}$$

amiből az iterációs eljárás

$$\begin{aligned}\mathbf{Px}^{(k+1)} &= \mathbf{Qx}^{(k)} + \mathbf{b} \\ \mathbf{x}^{(k+1)} &= \underbrace{\mathbf{P}^{-1}\mathbf{Q}}_{=\mathbf{B}}\mathbf{x}^{(k)} + \underbrace{\mathbf{P}^{-1}\mathbf{b}}_{=\mathbf{f}}.\end{aligned}\tag{2.2}$$

A (2.2) kifejezésből még

$$\mathbf{P}(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) + \mathbf{Ax}^{(k)} = \mathbf{b}$$

alakban is felírhatjuk az iterációt. A \mathbf{P} mátrixot *prekondicionálási mátrixnak* nevezük.

Tekintsük az alábbi differenciálegyenlet rendszert:

$$\mathbf{P}\frac{d\mathbf{x}}{dt} + \mathbf{Ax} = \mathbf{b} \quad (\mathbf{x} : \mathbb{R} \rightarrow \mathbb{R}^n)$$

Ha ennek létezik határértéke, azaz olyan $\mathbf{x}^* \in \mathbb{R}^n$, melyre $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{x}^*$, akkor ez az \mathbf{x}^* megoldása az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszernek. A fenti differenciálegyenlet rendszer diszkrétizált alakja:

$$\mathbf{P}\frac{\mathbf{x}(t + \tau) - \mathbf{x}(t)}{\tau} + \mathbf{Ax}(t) = \mathbf{b},\tag{2.3}$$

ahol τ az időegység.

Ebből a diszkrétizált alakból iterációs eljárást képezhetünk, ha a k iterációs szám-lálót úgy értelmezzük, mint a diszkrét t időt. Ekkor

$$\mathbf{x}(t) = \mathbf{x}^{(k)} \quad \text{és} \quad \mathbf{x}(t + \tau) = \mathbf{x}^{(k+1)}.$$

Így a (2.3) egyenletből a következő iterációt kapjuk:

$$\mathbf{P}\frac{\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}}{\tau} + \mathbf{Ax}^{(k)} = \mathbf{b}.$$

Az ilyen felírást az iterációs módszer *kanonikus alakjának* nevezik.

2.2. Klasszikus iterációk

Jacobi- és Gauss–Seidel-módszer

Ahhoz tehát, hogy iterációs módszerekkel dolgozni tudjunk, először keresni kell egy megfelelő konvergens sorozatot. De hogyan találhatunk ilyeneket?

Legyen most is $\mathbf{Ax} = \mathbf{b}$ az egyenletrendszer, aminek a megoldását közelíteni szeretnénk. Továbbra is feltesszük, hogy $\mathbf{A} \in \mathbb{R}^{n \times n}$ és $\det \mathbf{A} \neq 0$. Tudjuk, hogy ekkor az egyenletrendszernek létezik egyértelmű megoldása.

Írjuk fel most az egyenletrendszert komponensenként és rendezzük át az alábbi módon:

$$\begin{aligned} \mathbf{Ax} = \mathbf{b} &\Leftrightarrow a_{i1}x_1 + \dots + a_{in}x_n = b_i, & i = 1, \dots, n \\ &a_{ii} \neq 0 \quad \forall i = 1, \dots, n \\ x_i &= -\frac{a_{i1}}{a_{ii}}x_1 - \dots - \frac{a_{in}}{a_{ii}}x_n + \frac{b_i}{a_{ii}} \\ x_i &= -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}}x_j - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}}x_j + \frac{b_i}{a_{ii}} \end{aligned}$$

A fenti képletből adódik a következő iterációs eljárás: rögzítsük le az $\mathbf{x}^{(0)} = [x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}]$ kezdeti vektort, majd számoljuk ki $x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}$ értékeket a

$$x_i^{(k+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}}x_j^{(k)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}}x_j^{(k)} + \frac{b_i}{a_{ii}}, \quad i = 1, \dots, n$$

képlet segítségével (ahol most $k = 0$). Így megkapjuk $x^{(1)} = [x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}]$ vektort, azaz az első közelítést. Ennek segítségével $k = 1$ -et behelyettesítve kiszámoljuk a második közelítést, majd $k = 2$ -re a harmadik közelítést és így tovább.

Ezt az eljárást **Jacobi-iterációnak** nevezzük.

Egy másik klasszikus iterációs módszer a **Gauss–Seidel-iteráció**. Ez mindössze annyiban különbözik a Jacobi-iterációtól, hogy a $(k+1)$ -edik közelítés i -edik komponensének kiszámolásához felhasználjuk a $(k+1)$ -edik közelítés már addig kiszámolt komponenseit, azaz $j = 1, 2, \dots, (i-1)$ -et. Tehát a Gauss–Seidel-iteráció vektorkomponensenként kiírva:

$$x_i^{(k+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}}x_j^{(k+1)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}}x_j^{(k)} + \frac{b_i}{a_{ii}}, \quad i = 1, \dots, n$$

Az iterációs módszerek felírhatók mátrixos formában is. Írjuk fel az \mathbf{A} mátrixot $\mathbf{A} = \mathbf{A}_1 + \mathbf{D} + \mathbf{A}_2$ alakban, ahol \mathbf{D} egy olyan diagonális mátrix, melynek főátlójában az \mathbf{A} diagonálisában lévő elemek vannak (feltehetjük, hogy ezek közül egyik sem 0), \mathbf{A}_1 az \mathbf{A} mátrix diagonális alatti, míg \mathbf{A}_2 az \mathbf{A} mátrix diagonális feletti elemeinek mátrixa. Ekkor az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer felírható az alábbi alakban:

$$(\mathbf{A}_1 + \mathbf{D} + \mathbf{A}_2)\mathbf{x} = \mathbf{b}.$$

Ekkor azonos átalakításokkal a következő relációk érvényesek:

$$\begin{aligned} \mathbf{D}\underline{\mathbf{x}} &= -(\mathbf{A}_1 + \mathbf{A}_2)\underline{\mathbf{x}} + \underline{\mathbf{b}} \\ \underline{\mathbf{x}} &= -\mathbf{D}^{-1}(\mathbf{A}_1 + \mathbf{A}_2)\underline{\mathbf{x}} + \mathbf{D}^{-1}\underline{\mathbf{b}} \end{aligned} \quad (2.4)$$

A (2.4) kifejezésből származtatható a Jacobi-iteráció mátrixos alakja:

$$\underline{\mathbf{x}}^{(k+1)} = \underbrace{-\mathbf{D}^{-1}(\mathbf{A}_1 + \mathbf{A}_2)}_{:=\mathbf{B}_J} \underline{\mathbf{x}}^{(k)} + \underbrace{\mathbf{D}^{-1}\underline{\mathbf{b}}}_{\underline{\mathbf{f}}}, \quad (2.5)$$

ahol a Jacobi-iteráció iterációs mátrixát \mathbf{B}_J jelöli.

A koordinátás felírásnál láthattuk, hogy a Gauss–Seidel-módszer csak annyiban különbözik a Jacobi-iterációtól, hogy a $(k+1)$ -ik közelítésben felhasználjuk a már kiszámolt komponenseket. Ezt a különbséget a mátrixos felírás segítségével is kifejezhetjük:

$$\underline{\mathbf{x}}^{(k+1)} = -\mathbf{D}^{-1}(\mathbf{A}_1\underline{\mathbf{x}}^{(k+1)} + \mathbf{A}_2\underline{\mathbf{x}}^{(k)}) + \mathbf{D}^{-1}\underline{\mathbf{b}}. \quad (2.6)$$

Hogy megkapjuk a Gauss–Seidel-iteráció mátrixos alakját, rendezzük át ezt az implicit alakot explicitté!

$$\begin{aligned} \underline{\mathbf{x}}^{(k+1)} + \mathbf{D}^{-1}\mathbf{A}_1\underline{\mathbf{x}}^{(k+1)} &= -\mathbf{D}^{-1}\mathbf{A}_2\underline{\mathbf{x}}^{(k)} + \mathbf{D}^{-1}\underline{\mathbf{b}} \\ (\mathbf{D} + \mathbf{A}_1)\underline{\mathbf{x}}^{(k+1)} &= -\mathbf{A}_2\underline{\mathbf{x}}^{(k)} + \underline{\mathbf{b}} \\ \underline{\mathbf{x}}^{(k+1)} &= \underbrace{-(\mathbf{D} + \mathbf{A}_1)^{-1}\mathbf{A}_2}_{:=\mathbf{B}_{GS}} \underline{\mathbf{x}}^{(k)} + \underbrace{(\mathbf{D} + \mathbf{A}_1)^{-1}\underline{\mathbf{b}}}_{\underline{\mathbf{f}}}, \end{aligned} \quad (2.7)$$

ahol \mathbf{B}_{GS} a Gauss–Seidel-iteráció iterációs mátrixa. A fenti képletek segítségével az iterációk kanonikus alakja is könnyen felírható.

A Jacobi-iteráció kanonikus alakját az alábbi átalakításokkal kaphatjuk meg:

$$\begin{aligned} \underline{\mathbf{x}}^{(k+1)} &= -\mathbf{D}^{-1}(\mathbf{A}_1 + \mathbf{A}_2)\underline{\mathbf{x}}^{(k)} + \mathbf{D}^{-1}\underline{\mathbf{b}} \\ \mathbf{D}\underline{\mathbf{x}}^{(k+1)} + (\mathbf{A}_1 + \mathbf{A}_2)\underline{\mathbf{x}}^{(k)} &= \underline{\mathbf{b}} \\ \mathbf{D}\underline{\mathbf{x}}^{(k+1)} - \mathbf{D}\underline{\mathbf{x}}^{(k)} + (\mathbf{A}_1 + \mathbf{A}_2)\underline{\mathbf{x}}^{(k)} + \mathbf{D}\underline{\mathbf{x}}^{(k)} &= \underline{\mathbf{b}} \\ \mathbf{D}(\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}) + \underbrace{(\mathbf{A}_1 + \mathbf{A}_2 + \mathbf{D})}_{=\mathbf{A}}\underline{\mathbf{x}}^{(k)} &= \underline{\mathbf{b}} \\ \boxed{\mathbf{D}(\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}) + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}} \end{aligned}$$

A Gauss–Seidel-iteráció kanonikus alakját pedig a következő módon ál-

líthatjuk elő:

$$\begin{aligned}
 \underline{\mathbf{x}}^{(k+1)} &= -(\mathbf{D} + \mathbf{A}_1)^{-1} \mathbf{A}_2 \underline{\mathbf{x}}^{(k)} + (\mathbf{D} + \mathbf{A}_1)^{-1} \underline{\mathbf{b}} \\
 (\mathbf{D} + \mathbf{A}_1) \underline{\mathbf{x}}^{(k+1)} + \mathbf{A}_2 \underline{\mathbf{x}}^{(k)} &= \underline{\mathbf{b}} \\
 (\mathbf{D} + \mathbf{A}_1) \underline{\mathbf{x}}^{(k+1)} - (\mathbf{D} + \mathbf{A}_1) \underline{\mathbf{x}}^{(k)} + \mathbf{A}_2 \underline{\mathbf{x}}^{(k)} + (\mathbf{D} + \mathbf{A}_1) \underline{\mathbf{x}}^{(k)} &= \underline{\mathbf{b}} \\
 (\mathbf{D} + \mathbf{A}_1)(\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}) + \underbrace{(\mathbf{A}_2 + \mathbf{D} + \mathbf{A}_1)}_{=\mathbf{A}} \underline{\mathbf{x}}^{(k)} &= \underline{\mathbf{b}} \\
 \boxed{(\mathbf{D} + \mathbf{A}_1)(\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}) + \mathbf{A} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}}
 \end{aligned}$$

Vizsgáljuk meg, hogy a fenti módszerek által előállított vektorsorozatok valóban az $\mathbf{A}\underline{\mathbf{x}} = \underline{\mathbf{b}}$ egyenletrendszer megoldásához konvergálnak-e!

2.2.1. Állítás. Ha a Jacobi-iteráció által előállított $(\underline{\mathbf{x}}^{(k)})$ vektorsorozat konvergens, azaz létezik $\underline{\mathbf{x}}^*$, amire $\lim_{k \rightarrow \infty} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{x}}^*$, akkor $\underline{\mathbf{x}}^*$ megoldása az $\mathbf{A}\underline{\mathbf{x}} = \underline{\mathbf{b}}$ egyenletrendszernek.

Bizonyítás. A Jacobi-iteráció kanonikus képletében térjünk át limeszre!

$$\lim_{k \rightarrow \infty} [\mathbf{D}(\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}) + \mathbf{A}\underline{\mathbf{x}}^{(k)}] = \mathbf{D} \lim_{k \rightarrow \infty} \underbrace{(\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)})}_{\rightarrow \underline{\mathbf{x}}^* - \underline{\mathbf{x}}^* = 0} + \mathbf{A} \lim_{k \rightarrow \infty} \underbrace{\underline{\mathbf{x}}^{(k)}}_{\rightarrow \underline{\mathbf{x}}^*} = \mathbf{A}\underline{\mathbf{x}}^* = \underline{\mathbf{b}}$$

Tehát a vektorsorozat valóban az egyenletrendszer megoldásához tart. ■

2.2.2. Megjegyzés. Az állítás hasonlóan belátható a Gauss–Seidel-iterációra is:

$$\begin{aligned}
 &\lim_{k \rightarrow \infty} [(\mathbf{D} + \mathbf{A}_1)(\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}) + \mathbf{A}\underline{\mathbf{x}}^{(k)}] = \\
 &= (\mathbf{D} + \mathbf{A}_1) \lim_{k \rightarrow \infty} (\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}) + \mathbf{A} \lim_{k \rightarrow \infty} \underline{\mathbf{x}}^{(k)} = \mathbf{A}\underline{\mathbf{x}}^* = \underline{\mathbf{b}}.
 \end{aligned}$$

Egylépéses módszerek általános alakjai

Az eddig megismert két iterációs módszer az úgynevezett *egylépéses iterációk* közé tartozik. Az egylépéses iterációk általános kanonikus alakja:

$$\mathbf{P}_k \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\tau_k} + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}},$$

ahol $\mathbf{P}_k \in \mathbb{R}^{n \times n}$ adott reguláris mátrixok, τ_k pedig adott \mathbb{R} -beli paraméterek. A \mathbf{P}_k mátrixokat érdemes úgy megválasztani, hogy azok könnyen invertálhatóak legyenek, hiszen $\underline{\mathbf{x}}^{(k+1)}$ kiszámításához szükség van \mathbf{P}_k inverzének meghatározására. Ez jól látszik a fenti egyenlet következő átalakításából:

$$\begin{aligned}
 \mathbf{P}_k \underline{\mathbf{x}}^{(k+1)} &= \mathbf{P}_k \underline{\mathbf{x}}^{(k)} - \tau_k \mathbf{A}\underline{\mathbf{x}}^{(k)} + \tau_k \underline{\mathbf{b}} = (\mathbf{P}_k - \tau_k \mathbf{A}) \underline{\mathbf{x}}^{(k)} + \tau_k \underline{\mathbf{b}} \\
 \underline{\mathbf{x}}^{(k+1)} &= \mathbf{P}_k^{-1} (\mathbf{P}_k - \tau_k \mathbf{A}) \underline{\mathbf{x}}^{(k)} + \mathbf{P}_k^{-1} \tau_k \underline{\mathbf{b}}.
 \end{aligned}$$

Azokat az iterációs eljárásokat, melyekben minden iterációs lépésnél ugyanazzal a \mathbf{P} mátrixszal és τ paraméterrel számolunk, *stacionárius* iterációknak nevezzük. Képletük:

$$\mathbf{P} \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\tau} + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}.$$

Speciálisan a Jacobi-módszer $\mathbf{P} = \mathbf{D}$ és $\tau = 1$, a Gauss–Seidel-módszer pedig $\mathbf{P} = \mathbf{D} + \mathbf{A}_1$ és $\tau = 1$ választással a stacionárius iterációk közé tartoznak.

Azok az eljárások, melyeknél \mathbf{P} az \mathbf{I} egységmátrixszal egyenlő, az *explicit* iterációs eljárások. Ha egy explicit iteráció egyben stacionárius is, azaz

$$\frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\tau} + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}$$

képlettel írható fel, akkor „egyszerű iterációnak” nevezzük. Ha pedig az explicit iteráció nem stacionárius, azaz

$$\frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\tau_k} + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}$$

alakú, akkor pedig **Richardson-féle iterációról** beszélünk.

2.3. Az iterációs módszerek konvergenciája

A fejezet elején már említettük, hogy az iterációs módszerek legfontosabb tulajdonsága, hogy az általuk előállított vektorsorozat konvergens legyen és éppen az $\mathbf{A}\underline{\mathbf{x}} = \underline{\mathbf{b}}$ egyenletrendszer $\underline{\mathbf{x}}^*$ megoldásához tartson. De milyen feltételeknek kell teljesülnie ahhoz, hogy ez tetszőleges $\underline{\mathbf{x}}^{(0)}$ kezdővektor esetén megvalósuljon?

Először egy általános tételt fogunk kimondani a

$$\mathbf{P} \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\tau} + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}} \quad (2.8)$$

alakú stacionáris iterációkra. Ehhez vezessünk be néhány fogalmat!

Jelölje $\underline{\mathbf{e}}_k$ a k -adik közelítővektor és az $\underline{\mathbf{x}}^*$ megoldás különbségét. Ekkor az

$$\underline{\mathbf{e}}_k = \underline{\mathbf{x}}^{(k)} - \underline{\mathbf{x}}^*$$

vektort (a k -adik) *hibavektornak* nevezzük. Az, hogy az $(\underline{\mathbf{x}}^{(k)})$ vektorsorozat tart $\underline{\mathbf{x}}^*$ -hoz ekvivalens azzal, hogy a hibavektorok sorozata tart nullához:

$$\lim_{k \rightarrow \infty} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{x}}^* \quad \Leftrightarrow \quad \lim_{k \rightarrow \infty} (\underline{\mathbf{x}}^{(k)} - \underline{\mathbf{x}}^*) = 0 \quad \Leftrightarrow \quad \lim_{k \rightarrow \infty} \underline{\mathbf{e}}_k = 0.$$

Ha a (2.8) képletbe az $\underline{\mathbf{x}}$ vektorok limeszét helyettesítjük, majd kivonjuk az eredeti képletet, megkapjuk az ún. *hibaegyenletet*:

$$\begin{aligned}
 & \mathbf{P} \frac{\underline{\mathbf{x}}^* - \underline{\mathbf{x}}^*}{\tau} + \mathbf{A} \underline{\mathbf{x}}^* = \underline{\mathbf{b}} \\
 (-) \quad & \mathbf{P} \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\tau} + \mathbf{A} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}} \\
 \hline
 (=) \quad & \mathbf{P} \frac{\underline{\mathbf{e}}_{k+1} - \underline{\mathbf{e}}_k}{\tau} + \mathbf{A} \underline{\mathbf{e}}_k = 0, \tag{2.9}
 \end{aligned}$$

amit átrendezhetünk következő alakba:

$$\mathbf{P} \underline{\mathbf{e}}_{k+1} = \mathbf{P} \underline{\mathbf{e}}_k - \tau \mathbf{A} \underline{\mathbf{e}}_k = (\mathbf{P} - \tau \mathbf{A}) \underline{\mathbf{e}}_k. \tag{2.10}$$

2.3.1. Állítás. Tegyük fel, hogy \mathbf{A} szimmetrikus, szigorúan pozitív definit mátrix és \mathbf{P} invertálható, továbbá a $(\mathbf{P} - 0.5\tau\mathbf{A})$ mátrix szigorúan pozitív definit, azaz $\langle (\mathbf{P} - 0.5\tau\mathbf{A}) \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0 \quad \forall \underline{\mathbf{x}} \neq \underline{\mathbf{0}}$. Ekkor $\lim_{k \rightarrow \infty} \underline{\mathbf{e}}_k = \underline{\mathbf{0}}$.

Bizonyítás. Jelölje $J_k := \langle \mathbf{A} \underline{\mathbf{e}}_k, \underline{\mathbf{e}}_k \rangle$ skaláris szorzatot. Először megmutatjuk, hogy a (J_k) sorozat monoton csökkenő. A (2.10) egyenletet szorozzuk először balról \mathbf{P}^{-1} -zel, majd szintén balról \mathbf{A} -val!

$$\begin{aligned}
 \mathbf{P} \underline{\mathbf{e}}_{k+1} &= (\mathbf{P} - \tau \mathbf{A}) \underline{\mathbf{e}}_k \\
 \underline{\mathbf{e}}_{k+1} &= (\mathbf{I} - \tau \mathbf{P}^{-1} \mathbf{A}) \underline{\mathbf{e}}_k \tag{2.11}
 \end{aligned}$$

$$\mathbf{A} \underline{\mathbf{e}}_{k+1} = (\mathbf{A} - \tau \mathbf{A} \mathbf{P}^{-1} \mathbf{A}) \underline{\mathbf{e}}_k \tag{2.12}$$

Most a (2.12) egyenlet segítségével fejezzük ki J_{k+1} -et!

$$\begin{aligned}
 J_{k+1} &= \langle \mathbf{A} \underline{\mathbf{e}}_{k+1}, \underline{\mathbf{e}}_{k+1} \rangle \stackrel{(2.12)}{=} \langle (\mathbf{A} - \tau \mathbf{A} \mathbf{P}^{-1} \mathbf{A}) \underline{\mathbf{e}}_k, \underline{\mathbf{e}}_{k+1} \rangle \stackrel{(2.11)}{=} \\
 &= \langle (\mathbf{A} - \tau \mathbf{A} \mathbf{P}^{-1} \mathbf{A}) \underline{\mathbf{e}}_k, (\mathbf{I} - \tau \mathbf{P}^{-1} \mathbf{A}) \underline{\mathbf{e}}_k \rangle = \\
 &= \underbrace{\langle \mathbf{A} \underline{\mathbf{e}}_k, \underline{\mathbf{e}}_k \rangle}_{J_k} - \tau \underbrace{\langle \mathbf{A} \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k, \underline{\mathbf{e}}_k \rangle}_{\star} - \tau \underbrace{\langle \mathbf{A} \underline{\mathbf{e}}_k, \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k \rangle}_{\star} + \tau^2 \langle \mathbf{A} \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k, \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k \rangle
 \end{aligned}$$

A két \star -gal jelölt skaláris szorzat megegyezik, hiszen \mathbf{A} -ról feltettük, hogy szimmetrikus. Így a kapott egyenlet:

$$J_{k+1} = J_k - 2\tau \langle \mathbf{A} \underline{\mathbf{e}}_k, \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k \rangle + \tau^2 \langle \mathbf{A} \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k, \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k \rangle.$$

Vezessünk be egy új jelölést! Legyen

$$\underline{\mathbf{y}}_k := \mathbf{P}^{-1} \mathbf{A} \underline{\mathbf{e}}_k \quad \Leftrightarrow \quad \mathbf{P} \underline{\mathbf{y}}_k = \mathbf{A} \underline{\mathbf{e}}_k.$$

Helyettesítsük be $\underline{\mathbf{y}}_k$ -t és $\mathbf{P} \underline{\mathbf{y}}_k$ -t a megfelelő helyekre és alakítsuk tovább az egyenletet!

$$\begin{aligned}
 J_{k+1} &= J_k - 2\tau \langle \mathbf{P}\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle + \tau^2 \langle \mathbf{A}\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle = \\
 &= J_k - 2\tau (\langle \mathbf{P}\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle - 0.5\tau \langle \mathbf{A}\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle) = \\
 &= J_k - 2\tau \langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle
 \end{aligned} \tag{2.13}$$

Mivel az állítás feltételei alapján $(\mathbf{P} - 0.5\tau\mathbf{A})$ szigorúan pozitív definit mátrix, ezért $\langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle > 0$, amiből következik, hogy $J_{k+1} < J_k$. Továbbá minden J_k skaláris szorzat nagyobb nullánál, hiszen \mathbf{A} mátrix szigorúan pozitív definit. Tehát a (J_k) sorozat monoton csökkenő és alulról korlátos, ami azt jelenti, hogy létezik határértéke. Jelöljük ezt J^* -gal.

Térjünk most át a (2.13) képletben limeszre:

$$\begin{aligned}
 J^* &= \lim_{k \rightarrow \infty} J_{k+1} = \lim_{k \rightarrow \infty} [J_k - 2\tau \langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle] = \\
 &= \lim_{k \rightarrow \infty} J_k - 2\tau \lim_{k \rightarrow \infty} \langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle = \\
 &= J^* - 2\tau \lim_{k \rightarrow \infty} \langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle.
 \end{aligned}$$

A fenti átalakításból látható, hogy teljesülnie kell a következőnek:

$$\lim_{k \rightarrow \infty} \langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle = 0. \tag{2.14}$$

A következő lépésben szükségünk lesz az alábbi lemmára:

2.3.2. Lemma. Ha \mathbf{C} szigorúan pozitív definit mátrix, akkor létezik olyan $\delta > 0$ szám, melyre teljesül, hogy

$$\langle \mathbf{C}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle \geq \delta \|\underline{\mathbf{x}}\|^2$$

(Ilyen jó δ például a \mathbf{C} mátrix legkisebb sajátértéke is.)

Alkalmazzuk a lemmát $\mathbf{C} = \mathbf{P} - 0.5\tau\mathbf{A}$ és $\underline{\mathbf{x}} = \underline{\mathbf{y}}_k$ választással.

$$\langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{y}}_k, \underline{\mathbf{y}}_k \rangle \geq \delta \|\underline{\mathbf{y}}_k\|^2 \geq 0$$

Az egyenlőtlenség bal oldala (2.14) szerint tart nullához (ha $k \rightarrow \infty$), a jobb oldalon pedig nulla áll, tehát a rendőr-elv miatt $\delta \|\underline{\mathbf{y}}_k\|^2$ határértéke is nulla, amiből rögtön következik, hogy $\lim_{k \rightarrow \infty} \|\underline{\mathbf{y}}_k\| = 0$ szintén.

Eredetileg $\underline{\mathbf{y}}_k$ -t úgy választottuk, hogy $\underline{\mathbf{y}}_k = \mathbf{P}^{-1}\mathbf{A}\underline{\mathbf{e}}_k$, amiből $\underline{\mathbf{e}}_k = \mathbf{A}^{-1}\mathbf{P}\underline{\mathbf{y}}_k$. Ez normában:

$$\begin{aligned}
 \|\underline{\mathbf{e}}_k\| &= \|\mathbf{A}^{-1}\mathbf{P}\underline{\mathbf{y}}_k\| \leq \|\mathbf{A}^{-1}\mathbf{P}\| \|\underline{\mathbf{y}}_k\| \\
 \lim_{k \rightarrow \infty} \|\underline{\mathbf{e}}_k\| &= \|\mathbf{A}^{-1}\mathbf{P}\| \underbrace{\lim_{k \rightarrow \infty} \|\underline{\mathbf{y}}_k\|}_{=0} = 0.
 \end{aligned}$$

Tehát ezzel beláttuk, hogy $\lim_{k \rightarrow \infty} \|\underline{\mathbf{e}}_k\| = 0$, azaz a stacionárius iteráció konvergens és az $\mathbf{A}\underline{\mathbf{x}} = \underline{\mathbf{b}}$ egyenletrendszer $\underline{\mathbf{x}}^*$ megoldásához tart, ha a (2.3.1) állítás feltételei teljesülnek. ■

A Jacobi-iteráció konvergenciája

A Jacobi-iteráció konvergenciájára vonatkozó tétel megfogalmazásához szükségünk lesz a szigorúan diagonálisan domináns mátrixok fogalmának definiálására.

2.3.3. Definíció. Egy $\mathbf{A} \in \mathbb{R}^{n \times n}$ mátrixot szigorúan diagonálisan dominánsnak nevezünk, ha minden sorban a főátlóban lévő elem nagyobb, mint az adott sorban lévő összes többi elem abszolútértékének összege, azaz

$$a_{ii} > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad \forall i = 1, \dots, n.$$

2.3.4. Tétel. Ha az \mathbf{A} mátrix szimmetrikus és szigorúan diagonálisan domináns, akkor a Jacobi-iteráció konvergens.

Bizonyítás. A tételt a (2.3.1) tétel segítségével fogjuk bebizonyítani. Tudjuk, hogy az általános stacionárius iterációból a Jacobi-iterációt a $\mathbf{P} = \mathbf{D}$ és $\tau = 1$ választással kapjuk. Ezt és a tétel feltételeit figyelembe véve láthatjuk, hogy teljesül a (2.3.1) tétel két feltétele: \mathbf{A} szimmetrikus és szigorúan pozitív definit, illetve \mathbf{P} invertálható. Tehát csak azt kell belátnunk, hogy $(\mathbf{P} - 0.5\tau\mathbf{A})$ egy szigorúan pozitív definit mátrix. Ez most azt jelenti, hogy

$$\langle (\mathbf{D} - 0.5\mathbf{A})\mathbf{x}, \mathbf{x} \rangle > 0 \quad \Leftrightarrow \quad \langle (2\mathbf{D} - \mathbf{A})\mathbf{x}, \mathbf{x} \rangle > 0 \quad \Leftrightarrow \quad \langle 2\mathbf{D}\mathbf{x}, \mathbf{x} \rangle > \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle.$$

A számtani és mértani közepek közötti összefüggésből tudjuk, hogy

$$\begin{aligned} \sqrt{x_i^2 x_j^2} &\leq \frac{x_i^2 + x_j^2}{2} \\ |a_{ij}| \sqrt{x_i^2 x_j^2} &\leq |a_{ij}| \frac{x_i^2 + x_j^2}{2} \\ 2a_{ij}x_i x_j &\leq 2|a_{ij}||x_i||x_j| \leq |a_{ij}|(x_i^2 + x_j^2) = |a_{ij}|x_i^2 + |a_{ij}|x_j^2 \\ a_{ij}x_i x_j &\leq 0.5|a_{ij}|x_i^2 + 0.5|a_{ij}|x_j^2. \end{aligned} \tag{2.15}$$

Most írjuk fel $\mathbf{A}\mathbf{x}$ és \mathbf{x} vektorok skaláris szorzatát!

$$\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \sum_{i,j=1}^n a_{ij}x_i x_j \stackrel{(2.15)}{\leq} 0.5 \sum_{i,j=1}^n |a_{ij}|x_i^2 + 0.5 \sum_{i,j=1}^n |a_{ij}|x_j^2$$

Az összeg második tagjában felcseréljük a szummákat:

$$\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \leq 0.5 \sum_{i,j=1}^n |a_{ij}|x_i^2 + 0.5 \sum_{j,i=1}^n |a_{ji}|x_i^2.$$

Mivel \mathbf{A} szimmetrikus, ezért $a_{ji} = a_{ij}$.

$$\begin{aligned} \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle &\leq 0.5 \sum_{i,j=1}^n |a_{ij}|x_i^2 + 0.5 \sum_{j,i=1}^n |a_{ij}|x_i^2 = \sum_{i,j=1}^n |a_{ij}|x_i^2 = \\ &= \sum_{i=1}^n x_i^2 \left(\sum_{j=1}^n |a_{ij}| \right) = \sum_{i=1}^n x_i^2 \left(|a_{ii}| + \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right) = \sum_{i=1}^n x_i^2 \left(a_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right). \end{aligned}$$

Mivel \mathbf{A} -ról feltettük, hogy szigorúan diagonálisan domináns, ezért $\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < a_{ii}$, így:

$$\langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle \leq \sum_{i=1}^n x_i^2 \left(a_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right) < \sum_{i=1}^n x_i^2 (2a_{ii}) = 2 \sum_{i=1}^n a_{ii} x_i^2 = 2 \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle.$$

Tehát $\langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle < \langle 2\mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle$ és ezzel a tételt beláttuk. ■

Az egyszerű iteráció konvergenciája

Most azt fogjuk megvizsgálni, hogyan kell megválasztani az egyszerű iterációban a τ paramétert úgy, hogy az iteráció konvergens legyen. Az egyszerű iteráció kanonikus alakját a következőképpen írtuk fel:

$$\frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\tau} + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}.$$

Továbbra is azt az esetet fogjuk vizsgálni, amikor \mathbf{A} szimmetrikus, szigorúan pozitív definit mátrix.

Ismét a (2.3.1) tételt fogjuk használni a konvergencia bizonyításához. Mivel az \mathbf{A} mátrixot szimmetrikus, szigorúan pozitív definitnek választjuk, és a \mathbf{P} mátrixnak most az \mathbf{I} egységmátrix felel meg, amiről tudjuk, hogy invertálható, így a (2.3.1) tétel két feltétele teljesül. Tehát már csak azt kell megvizsgálnunk, hogy τ milyen megválasztása esetén lesz az $(\mathbf{I} - 0.5\tau\mathbf{A})$ mátrix szigorúan pozitív definit.

2.3.5. Tétel. *Szimmetrikus, szigorúan pozitív definit \mathbf{A} mátrixokra az egyszerű iteráció $\tau < \frac{2}{\lambda_{max}}$ esetén konvergens, ahol λ_{max} az \mathbf{A} mátrix legnagyobb sajátértéke.*

Bizonyítás. Mivel az \mathbf{A} mátrix szigorúan pozitív definit és szimmetrikus, ezért az összes sajátértéke pozitív és valós. Legyenek $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ ($:= \lambda_{max}$) az \mathbf{A} mátrix sajátértékei. Az $(\mathbf{I} - 0.5\tau\mathbf{A})$ mátrix sajátértékeinek felírásához szükségünk lesz az alábbi lemmára:

2.3.6. Lemma. Legyenek \mathbf{A} és \mathbf{B} $n \times n$ -es mátrixok. \mathbf{A} sajátértékeit jelölje λ_i , \mathbf{B} sajátértékeit pedig μ_j . Ekkor a két mátrix lineáris kombinációjának sajátértékei megegyeznek a mátrixok sajátértékeinek lineáris kombinációjával, azaz

$$a\mathbf{A} + b\mathbf{B} \text{ sajátértékei: } a\lambda_i + b\mu_j \quad a, b \in \mathbb{R},$$

akkor és csak akkor, ha a két mátrix kommutál egymással, azaz $\mathbf{AB} = \mathbf{BA}$. Ez a feltétel ekvivalens azzal, hogy \mathbf{A} és \mathbf{B} mátrixok sajátvektorai megegyeznek.

Jelen esetben ez nyilvánvalóan teljesül \mathbf{I} és \mathbf{A} mátrixokra, hiszen

$$\mathbf{A} = \mathbf{IA} = \mathbf{AI} = \mathbf{A}.$$

Tehát az $(\mathbf{I} - 0.5\tau\mathbf{A})$ mátrix sajátértékei felírhatók a következő alakban:

$$1 - 0.5\tau\lambda_i, \quad i = 1, \dots, n.$$

Az iteráció akkor konvergens, ha az $(\mathbf{I} - 0.5\tau\mathbf{A})$ mátrix szigorúan pozitív definit, ami azt jelenti, hogy az összes sajátértéke pozitív, azaz

$$1 - 0.5\tau\lambda_i > 0, \quad i = 1, \dots, n.$$

Ez a feltétel ekvivalens az

$$1 - 0.5\tau\lambda_{max} > 0$$

feltétellel, azaz

$$\tau < \frac{2}{\lambda_{max}}.$$

Tehát valóban $\tau < \frac{2}{\lambda_{max}}$ a konvergencia elégséges feltétele. ■

2.3.7. Állítás. A fenti tétel feltétele nem csak elégséges, hanem szükséges feltétele is a konvergenciának.

Bizonyítás. Most azt kell belátnunk, hogy ha $\tau < \frac{2}{\lambda_{max}}$ nem teljesül, akkor az iteráció nem lehet konvergens. Mivel a konvergencia azt jelenti, hogy tetszőleges $\underline{\mathbf{x}}^{(0)}$ kezdeti vektor esetén tart az iteráció az $\underline{\mathbf{x}}^*$ megoldáshoz, ezért most elég megmutatnunk azt, hogy létezik olyan $\underline{\mathbf{x}}^{(0)}$ vektor, amire ez nem igaz.

Jelöljük $\underline{\mathbf{v}}_n$ -nel az \mathbf{A} mátrix legnagyobb, azaz a $\lambda_n (= \lambda_{max})$ sajátértékéhez tartozó sajátvektort. Ez azt jelenti, hogy $\mathbf{A}\underline{\mathbf{v}}_n = \lambda_n\underline{\mathbf{v}}_n$. Továbbá legyen a kezdeti vektor $\underline{\mathbf{x}}^{(0)} = \underline{\mathbf{x}}^* + \underline{\mathbf{v}}_n$. Ekkor a nulladik hibavektor, $\underline{\mathbf{e}}_0$ éppen $\underline{\mathbf{v}}_n$ -nel lesz egyenlő, hisz

$$\underline{\mathbf{e}}_0 = \underline{\mathbf{x}}^{(0)} - \underline{\mathbf{x}}^* = (\underline{\mathbf{x}}^* + \underline{\mathbf{v}}_n) - \underline{\mathbf{x}}^* = \underline{\mathbf{v}}_n.$$

Az egyszerű iteráció hibaegyenlete felírható a (2.9) képletből $\mathbf{P} = \mathbf{I}$ behelyettesítéssel:

$$\frac{\underline{\mathbf{e}}_{k+1} - \underline{\mathbf{e}}_k}{\tau} + \mathbf{A}\underline{\mathbf{e}}_k = 0.$$

Ebből átrendezéssel a $(k + 1)$ -ik hibavektorra kapjuk:

$$\underline{\mathbf{e}}_{k+1} = \underline{\mathbf{e}}_k - \tau\mathbf{A}\underline{\mathbf{e}}_k = (\mathbf{I} - \tau\mathbf{A})\underline{\mathbf{e}}_k.$$

A fenti képlet segítségével fejezzük ki $\underline{\mathbf{e}}_k$ -t!

$$\underline{\mathbf{e}}_k = (\mathbf{I} - \tau\mathbf{A})\underline{\mathbf{e}}_{k-1} = (\mathbf{I} - \tau\mathbf{A})^2\underline{\mathbf{e}}_{k-2} = \dots = (\mathbf{I} - \tau\mathbf{A})^k\underline{\mathbf{e}}_0 = (\mathbf{I} - \tau\mathbf{A})^k\underline{\mathbf{v}}_n$$

Mivel

$$(\mathbf{I} - \tau\mathbf{A})\underline{\mathbf{v}}_n = \underline{\mathbf{v}}_n - \tau\mathbf{A}\underline{\mathbf{v}}_n = \underline{\mathbf{v}}_n - \tau\lambda_n\underline{\mathbf{v}}_n = (1 - \tau\lambda_n)\underline{\mathbf{v}}_n,$$

ezért

$$\underline{\mathbf{e}}_k = (\mathbf{I} - \tau\mathbf{A})^k\underline{\mathbf{v}}_n = (1 - \tau\lambda_n)^k\underline{\mathbf{v}}_n.$$

Az iteráció pontosan akkor konvergens, ha $\lim_{k \rightarrow \infty} \underline{\mathbf{e}}_k = 0$. Ez most akkor és csak akkor fog teljesülni, ha

$$|1 - \tau\lambda_n| < 1 \quad \Leftrightarrow \quad -1 < 1 - \tau\lambda_n < 1.$$

Az egyértelmű, hogy $1 - \tau\lambda_n < 1$ teljesül, hiszen τ egy pozitív paraméter és $\lambda_n > 0$, hiszen \mathbf{A} -ról feltettük, hogy szigorúan pozitív definit. Tehát ahhoz, hogy az iteráció konvergens legyen, teljesülnie kell a $(-1 < 1 - \tau\lambda_n)$ feltételnek is. Ez a

$$\tau < \frac{2}{\lambda_n} = \frac{2}{\lambda_{max}}$$

feltételt eredményezi. Tehát ha $\tau \geq \frac{2}{\lambda_{max}}$, akkor létezik olyan kezdeti vektor (az $\underline{\mathbf{x}}^{(0)} = \underline{\mathbf{x}}^* + \underline{\mathbf{v}}_n$), ami esetén az egyszerű iteráció nem konvergens. Ezzel beláttuk, hogy $\tau < \frac{2}{\lambda_{max}}$ valóban szükséges feltétele is a konvergenciának. ■

3. fejezet

Relaxációs módszerek

3.1. A JOR-módszer

Ebben a fejezetben olyan iterációs módszerekről lesz szó, melyeket a már eddig megismert klasszikus iterációkból kapunk valamilyen kis módosítással.

Az alábbi triviális azonosság tetszőleges iterációs módszerre felírható:

$$\underline{\mathbf{x}}^{(k+1)} = \underline{\mathbf{x}}^{(k)} + (\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}). \quad (3.1)$$

Képezzünk most egy új iterációs eljárást! Vegyük a fenti egyenlőséget és módosítsuk a következőképpen: a jobb oldali $\underline{\mathbf{x}}^{(k+1)}$ helyére írjunk egy olyan vektort, melyet a már ismert iterációs módszerek közül valamelyik állít elő a $(k+1)$ -ik lépésben, továbbá a jobb oldali zárójeles tagot szorozzuk meg egy tetszőleges ω pozitív paraméterrel.

$$\underline{\mathbf{x}}^{(k+1)} = \underline{\mathbf{x}}^{(k)} + \omega (\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)})$$

Tekintsük például a Jacobi-iterációt, és tegyük fel, hogy a k -adik lépést már végrehajtottuk, tehát ismerjük $\underline{\mathbf{x}}^{(k)}$ vektort. Azt a vektort, amit a Jacobi-iteráció következő lépésében állítanánk elő, jelöljük $\underline{\mathbf{x}}_J^{(k+1)}$ -el. Így definiálhatjuk a következő iterációt:

$$\underline{\mathbf{x}}^{(k+1)} = \underline{\mathbf{x}}^{(k)} + \omega (\underline{\mathbf{x}}_J^{(k+1)} - \underline{\mathbf{x}}^{(k)}). \quad (3.2)$$

Nyilván ha $\omega = 1$, akkor a (3.2) szerint $\underline{\mathbf{x}}^{(k+1)} = \underline{\mathbf{x}}_J^{(k+1)}$, azaz az új iteráció a Jacobi-iterációt eredményezi. Ha viszont $\omega \neq 1$, akkor a fenti módosítás azt jelenti, hogy az adott $\underline{\mathbf{x}}^{(k)}$ vektorból most is a Jacobi-iteráció által adott $\underline{\mathbf{x}}_J^{(k+1)}$ vektor irányába lépünk tovább, de az $\underline{\mathbf{x}}_J^{(k+1)}$ vektortól egy kicsit közelebbi vagy kicsit távolabbi vektorba fogunk jutni attól függően, hogy ω -t kisebbnek vagy nagyobbak választottuk-e 1-nél. Ha $0 < \omega < 1$, akkor *alulrelaxálásról*, ha pedig $1 < \omega$, akkor *túlrelaxálásról* beszélünk. A relaxálás szó itt most arra utal, hogy a Jacobi-iteráció képletét nem vesszük olyan szigorúan, kicsit módosítunk rajta. Ezt a módosított iterációt **JOR-módszernek** nevezzük, iterációs mátrixát $\mathbf{B}_{J(\omega)}$ jelöli.

A JOR-módszer mátrixos alakját tehát úgy kaphatjuk meg, hogy a fenti képletben $\underline{\mathbf{x}}_J^{(k+1)}$ helyére a Jacobi-iteráció által előállított $(k+1)$ -ik vektor képletét írjuk:

$$\underline{\mathbf{x}}^{(k+1)} = \underline{\mathbf{x}}^{(k)} + \omega \left(-\mathbf{D}^{-1}(\mathbf{A}_1 + \mathbf{A}_2)\underline{\mathbf{x}}^{(k)} + \mathbf{D}^{-1}\underline{\mathbf{b}} - \underline{\mathbf{x}}^{(k)} \right) \quad (3.3)$$

$$\underline{\mathbf{x}}^{(k+1)} = \underbrace{\left[(1-\omega)\mathbf{I} - \omega\mathbf{D}^{-1}(\mathbf{A}_1 + \mathbf{A}_2) \right]}_{\mathbf{B}_{J(\omega)}} \underline{\mathbf{x}}^{(k)} + \omega\mathbf{D}^{-1}\underline{\mathbf{b}}. \quad (3.4)$$

A (3.3) képletből egyszerű átrendezéssel pedig megkapjuk a JOR-módszer kanonikus alakját is:

$$\mathbf{D} \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\omega} = -(\mathbf{A}_1 + \mathbf{A}_2)\underline{\mathbf{x}}^{(k)} + \underline{\mathbf{b}} - \mathbf{D}\underline{\mathbf{x}}^{(k)}$$

$$\boxed{\mathbf{D} \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\omega} + \mathbf{A}\underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}.}$$

Az egyértelmű, hogy a módszer $\omega = 1$ esetén milyen feltételek mellett lesz konvergens, hiszen ekkor éppen a Jacobi-iterációt kapjuk vissza. Vizsgáljuk meg, hogy milyen más ω esetén lehet a JOR-módszer konvergens!

3.1.1. Tétel. *Szimmetrikus, szigorúan diagonálisan domináns \mathbf{A} mátrixokon a JOR-módszer $\omega \in (0, 1]$ esetén konvergens.*

Bizonyítás. A Jacobi-iteráció konvergenciájára vonatkozó (2.3.4) tétel bizonyítása során beláttuk, hogy

$$\langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle < 2 \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle. \quad (3.5)$$

A konvergencia bizonyításához most is a (2.3.1) tételt fogjuk használni, $\mathbf{P} = \mathbf{D}$ és $\tau = \omega$ szereposztásban. Az hogy az \mathbf{A} mátrix szimmetrikus, szigorúan pozitív definit, és \mathbf{D} invertálható, teljesül, hiszen \mathbf{A} -ról feltettük, hogy szimmetrikus, és szigorúan diagonálisan domináns. Tehát csak azt kell megvizsgálunk, hogy milyen ω esetén lesz $(\mathbf{D} - 0.5\omega\mathbf{A})$ mátrix szigorúan pozitív definit.

$$\begin{aligned} \langle (\mathbf{D} - 0.5\omega\mathbf{A})\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle &= \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - 0.5\omega \langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle \stackrel{(3.5)}{>} \\ &\stackrel{(3.5)}{>} 0.5 \langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - 0.5\omega \langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = (0.5 - 0.5\omega) \langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle \end{aligned}$$

Mivel \mathbf{A} szigorúan pozitív definit, ezért $\langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0$, tehát a $\langle (\mathbf{D} - 0.5\omega\mathbf{A})\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle$ skaláris szorzat akkor lesz pozitív, ha $(0.5 - 0.5\omega) \geq 0$. Ez nyilván az

$$\omega \leq 1$$

feltételt jelenti.

Mivel ω pozitív paraméter, így a tételt beláttuk, a JOR-módszer $0 < \omega \leq 1$ esetén konvergens. ■

3.2. A Gauss–Seidel-iteráció relaxált változatai

A Jacobi-módszerhez hasonlóan a Gauss–Seidel-iterációnak is létezik relaxált változata. Induljunk el ugyanazon logika alapján, mint az előbb, csak most $\underline{\mathbf{x}}_J^{(k+1)}$ helyett $\underline{\mathbf{x}}_{GS}^{(k+1)}$ vektort helyettesítsük be a (3.2) egyenletbe, ami a Gauss–Seidel-iteráció által $(k+1)$ -ik lépésben előállított vektort jelöli. Ekkor a következő iterációt kapjuk:

$$\underline{\mathbf{x}}^{(k+1)} = \underline{\mathbf{x}}^{(k)} + \omega \left(\underline{\mathbf{x}}_{GS}^{(k+1)} - \underline{\mathbf{x}}^{(k)} \right). \quad (3.6)$$

A fenti képletbe az $\underline{\mathbf{x}}_{GS}^{(k+1)}$ helyére először a (2.7) explicit egyenlettel felírt Gauss–Seidel-iteráció által előállított $(k+1)$ -ik vektort helyettesítettem be. Így az alábbi egyenletet kaptam:

$$\underline{\mathbf{x}}^{(k+1)} = \underline{\mathbf{x}}^{(k)} + \omega \left(-(\mathbf{D} + \mathbf{A}_1)^{-1} \mathbf{A}_2 \underline{\mathbf{x}}^{(k)} + (\mathbf{D} + \mathbf{A}_1)^{-1} \underline{\mathbf{b}} - \underline{\mathbf{x}}^{(k)} \right),$$

amiből az új iteráció mátrixos alakja:

$$\underline{\mathbf{x}}^{(k+1)} = \underbrace{[(1 - \omega)\mathbf{I} - \omega(\mathbf{D} + \mathbf{A}_1)^{-1} \mathbf{A}_2]}_{:=\mathbf{B}_{GSOR}} \underline{\mathbf{x}}^{(k)} + \omega(\mathbf{D} + \mathbf{A}_1)^{-1} \underline{\mathbf{b}}, \quad (3.7)$$

a kanonikus alakja pedig

$$(\mathbf{D} + \mathbf{A}_1) \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\omega} = -\mathbf{A}_2 \underline{\mathbf{x}}^{(k)} + \underline{\mathbf{b}} - (\mathbf{D} + \mathbf{A}_1) \underline{\mathbf{x}}^{(k)}$$

átrendezés következtében:

$$\boxed{(\mathbf{D} + \mathbf{A}_1) \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\omega} + \mathbf{A}_2 \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}}. \quad (3.8)$$

Ám a szakirodalomban mégsem a fenti módszert szokták a Gauss–Seidel-iteráció relaxált változataként emlegetni. A (3.8) kanonikus képlettel leírható módszert a továbbiakban **GSOR-módszernek** fogom nevezni.

A Gauss–Seidel-iteráció ismert relaxált változatát, vagyis az úgynevezett **SOR-módszert** akkor kapjuk meg, ha a (3.6) képletben $\underline{\mathbf{x}}_{GS}^{(k+1)}$ helyére a Gauss–Seidel-iteráció (2.6) implicit képletének jobb oldalát helyettesítjük.

$$\begin{aligned} \underline{\mathbf{x}}^{(k+1)} &= \underline{\mathbf{x}}^{(k)} + \omega \left(-\mathbf{D}^{-1} (\mathbf{A}_1 \underline{\mathbf{x}}^{(k+1)} + \mathbf{A}_2 \underline{\mathbf{x}}^{(k)}) + \mathbf{D}^{-1} \underline{\mathbf{b}} - \underline{\mathbf{x}}^{(k)} \right) \\ \underline{\mathbf{x}}^{(k+1)} + \omega \mathbf{D}^{-1} \mathbf{A}_1 \underline{\mathbf{x}}^{(k+1)} &= \underline{\mathbf{x}}^{(k)} + \omega \left(-\mathbf{D}^{-1} \mathbf{A}_2 \underline{\mathbf{x}}^{(k)} + \mathbf{D}^{-1} \underline{\mathbf{b}} - \underline{\mathbf{x}}^{(k)} \right) \\ \mathbf{D} \underline{\mathbf{x}}^{(k+1)} + \omega \mathbf{A}_1 \underline{\mathbf{x}}^{(k+1)} &= \mathbf{D} \underline{\mathbf{x}}^{(k)} + \omega \left(-\mathbf{A}_2 \underline{\mathbf{x}}^{(k)} + \underline{\mathbf{b}} - \mathbf{D} \underline{\mathbf{x}}^{(k)} \right) \end{aligned} \quad (3.9)$$

A (3.9) egyenletből a következő átrendezéssel kapjuk a SOR-módszer mátrixos alakját:

$$\begin{aligned} (\mathbf{D} + \omega \mathbf{A}_1) \underline{\mathbf{x}}^{(k+1)} &= [(1 - \omega) \mathbf{D} - \omega \mathbf{A}_2] \underline{\mathbf{x}}^{(k)} + \omega \underline{\mathbf{b}} \\ \underline{\mathbf{x}}^{(k+1)} &= \underbrace{-(\mathbf{D} + \omega \mathbf{A}_1)^{-1} [(\omega - 1) \mathbf{D} + \omega \mathbf{A}_2]}_{:=\mathbf{B}_{GS(\omega)}} \underline{\mathbf{x}}^{(k)} + \omega (\mathbf{D} + \omega \mathbf{A}_1)^{-1} \underline{\mathbf{b}}, \end{aligned} \quad (3.10)$$

ahol $\mathbf{B}_{GS(\omega)}$ a SOR-módszer iterációs mátrixa.

A (3.9) egyenletből a kanonikus alakot is kifejezhetjük:

$$\begin{aligned}
 (\mathbf{D} + \omega \mathbf{A}_1) \underline{\mathbf{x}}^{(k+1)} &= \underbrace{\mathbf{D} \underline{\mathbf{x}}^{(k)}}_{+\omega \mathbf{A}_1 \underline{\mathbf{x}}^{(k)}} + \underbrace{\omega(-\mathbf{A}_2 - \mathbf{D}) \underline{\mathbf{x}}^{(k)}}_{-\omega \mathbf{A}_1 \underline{\mathbf{x}}^{(k)}} + \omega \underline{\mathbf{b}} \\
 (\mathbf{D} + \omega \mathbf{A}_1) \underline{\mathbf{x}}^{(k+1)} &= (\mathbf{D} + \omega \mathbf{A}_1) \underline{\mathbf{x}}^{(k)} - \omega(\mathbf{A}_2 + \mathbf{D} + \mathbf{A}_1) \underline{\mathbf{x}}^{(k)} + \omega \underline{\mathbf{b}} \\
 \boxed{(\mathbf{D} + \omega \mathbf{A}_1) \frac{\underline{\mathbf{x}}^{(k+1)} - \underline{\mathbf{x}}^{(k)}}{\omega} + \mathbf{A} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{b}}}. & \quad (3.11)
 \end{aligned}$$

A Gauss–Seidel-módszer relaxált változatainál is fontos kérdés, hogy ω milyen értéke esetén konvergensek. Először lássuk a SOR-módszer konvergenciájára vonatkozó ismert tételt.

3.2.1. Tétel. (Kahan-tétel) *Tegyük fel, hogy \mathbf{A} szimmetrikus, szigorúan pozitív definit mátrix. Ekkor a SOR-módszer $\omega \in (0, 2)$ esetén konvergens.*

Bizonyítás. Tekintsük az $\mathbf{A} = \mathbf{A}_1 + \mathbf{D} + \mathbf{A}_2$ felbontást. Mivel \mathbf{A} szimmetrikus, ezért $\mathbf{A}_2^T = \mathbf{A}_1$ és mivel \mathbf{A} szigorúan pozitív definit, ezért a diagonálisában lévő elemek pozitívak, azaz $\mathbf{D} > 0$. Ezek segítségével fejezzük ki az $\langle \mathbf{A} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle$ skaláris szorzatot:

$$\begin{aligned}
 \langle \mathbf{A} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle &= \langle (\mathbf{A}_1 + \mathbf{D} + \mathbf{A}_2) \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = \langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{A}_2 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = \\
 &= \langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \underline{\mathbf{x}}, \mathbf{A}_2^T \underline{\mathbf{x}} \rangle = \langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle.
 \end{aligned}$$

Tehát

$$\langle \mathbf{A} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = 2\langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle. \quad (3.12)$$

A bizonyításhoz használjuk a (2.3.1) tétel konvergenciára vonatkozó elégséges feltételét, azaz

$$\mathbf{P} - 0.5\tau \mathbf{A} > 0 \quad \Leftrightarrow \quad \langle (\mathbf{P} - 0.5\tau \mathbf{A}) \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0,$$

ahol most $\mathbf{P} = (\mathbf{D} + \omega \mathbf{A}_1)$, $\tau = \omega$. Lássuk, mikor fog ez teljesülni!

$$\begin{aligned}
 \langle (\mathbf{P} - 0.5\tau \mathbf{A}) \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle &= \langle \mathbf{P} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - 0.5\tau \langle \mathbf{A} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = \\
 &= \langle (\mathbf{D} + \omega \mathbf{A}_1) \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - 0.5\omega [2\langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle] = \\
 &= \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \omega \langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - \omega \langle \mathbf{A}_1 \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - 0.5\omega \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = (1 - 0.5\omega) \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle
 \end{aligned}$$

Tehát a konvergenciához az kell, hogy

$$(1 - 0.5\omega) \langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0$$

teljesüljön. Mivel $\mathbf{D} > 0$, így $\langle \mathbf{D} \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0$. $(1 - 0.5\omega)$ pedig akkor lesz pozitív, ha $\omega < 2$. Tehát a SOR-módszer valóban akkor lesz konvergens, ha $\omega \in (0, 2)$. ■

3.2.2. Következmény. Szimmetrikus, szigorúan pozitív definit \mathbf{A} mátrixokra a Gauss–Seidel-módszer konvergens.

Most pedig vizsgáljuk meg, hogy vajon a GSOR-módszer konvergenciájának is ugyanez a feltétele, vagy ez az iteráció más ω értékek esetén konvergens.

Továbbra is tegyük fel, hogy \mathbf{A} szimmetrikus, szigorúan pozitív definit mátrix. Mint az előbb, most is azt kell vizsgálnunk, hogy

$$\langle (\mathbf{P} - 0.5\tau\mathbf{A})\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0$$

feltétel teljesül-e. Itt $\mathbf{P} = \mathbf{D} + \mathbf{A}_1$ és $\tau = \omega$.

$$\langle (\mathbf{D} + \mathbf{A}_1 - 0.5\omega\mathbf{A})\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \langle \mathbf{A}_1\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - 0.5\omega \langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle$$

Felhasználva a Kahan-tétel bizonyításában levezetett (3.12) egyenlőséget

$$\langle \mathbf{A}_1\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle = \frac{\langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle}{2}$$

teljesül. Így a fenti egyenlőség tovább írható a következőképpen:

$$\begin{aligned} \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \frac{\langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle - \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle}{2} - 0.5\omega \langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle &= \\ = 0.5 \langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + (0.5 - 0.5\omega) \langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle \end{aligned}$$

Mivel \mathbf{A} szigorúan pozitív definit, így $\langle \mathbf{A}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle$ és $\langle \mathbf{D}\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle$ skaláris szorzatok pozitívak. Tehát ahhoz, hogy $\langle (\mathbf{D} + \mathbf{A}_1 - 0.5\omega\mathbf{A})\underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0$ legyen, a

$$0.5 - 0.5\omega \geq 0$$

feltételnek kell teljesülnie. Ezt átrendezve az

$$1 \geq \omega$$

feltételt nyerjük. Tehát a GSOR-módszer $\omega \in (0, 1]$ esetén konvergens.

3.3. A módszerek összehasonlítása egy példán

Ebben a részben az eddig megismert három relaxációs módszert egy konkrét példán fogjuk összehasonlítani. A példát David M. Young ([4]) könyvének mintafeladata és a hozzá tartozó 3. számú gyakorlat alapján állítottam össze.

Jelölések: Legyen $\bar{\Omega} := [0, 1] \times [0, 1]$ tartomány, azaz az egységnyezet. Jelöljük $\bar{\Omega}$ belsejét Ω -val, határát pedig $\partial\Omega$ -val.

Feladat. Keresünk egy olyan $u(x, y)$ függvényt, amely folytonos az $\bar{\Omega}$ tartományon, kétszer folytonosan differenciálható Ω -n és kielégíti az alábbi Poisson-egyenletet:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = G(x, y) := x + y + 1 \quad (x, y) \in \Omega \quad (3.13)$$

$$u(x, y) = g(x, y) := 1 + x^2 \quad (x, y) \in \partial\Omega. \quad (3.14)$$

Határozzuk meg a numerikus megoldást $h = \frac{1}{3}$ rácsállandóval, ha a kezdeti közelítőértékek nullák!

Először nézzük meg, hogyan adódik lineáris egyenletrendszer a feladat numerikus megoldásából a véges differenciák módszerével. Helyezzünk négyzetrácsot a tartományra úgy, hogy az egyenesek párhuzamosak legyenek a tengelyekkel, a négyzetek oldala, azaz a rácsállandó pedig legyen h . Ekkor az (x, y) rácspontban a másodrendű parciális deriváltakat a következő differenciahányadosokkal közelítjük:

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u(x+h, y) + u(x-h, y) - 2u(x, y)}{h^2}$$

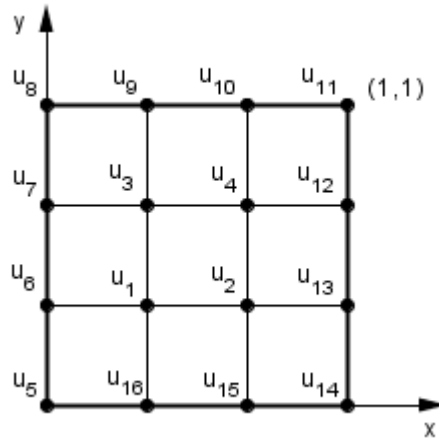
$$\frac{\partial^2 u}{\partial y^2} \approx \frac{u(x, y+h) + u(x, y-h) - 2u(x, y)}{h^2}.$$

Ha a (3.13) egyenletben a parciális deriváltak helyére a differenciahányadosokat írjuk, majd $(-h^2)$ -tel végigszorozunk, a következő differenciaegyenletet kapjuk:

$$4u(x, y) - u(x+h, y) - u(x-h, y) - u(x, y+h) - u(x, y-h) = -h^2 G(x, y). \quad (3.15)$$

Ez az egyenlet a (3.14) peremfeltétellel együtt a feladat diszkrét megfelelője.

Feladatunkban a rácsállandó $h = \frac{1}{3}$. Az ehhez tartozó rácshálót az alábbi ábra szemlélteti:



3.1. ábra.

A 3.1. ábra és a (3.15) egyenlet segítségével minden belső pontra fel tudunk írni egy lineáris egyenletet. A perempontok értékét a peremfeltétel szabja meg. Írjunk fel tehát a belső rácspontokra egy négy ismeretlenes lineáris egyenletrendszert!

$$4u_1 - u_2 - u_6 - u_3 - u_{16} = G(u_1)$$

$$4u_2 - u_{13} - u_1 - u_4 - u_{15} = G(u_2)$$

$$4u_3 - u_4 - u_7 - u_9 - u_1 = G(u_3)$$

$$4u_4 - u_{12} - u_3 - u_{10} - u_2 = G(u_4)$$

Helyettesítsünk u_i helyére $g(u_i)$ -t, ha u_i perempont (vagyis $i = \{5, 6, \dots, 16\}$), majd rendezzük bal oldalra az ismeretleneket, a jobb oldalakat pedig számítsuk is ki!

$$\begin{aligned} 4u_1 - u_2 - u_3 &= G(u_1) + g(u_6) + g(u_{16}) = G\left(\frac{1}{3}, \frac{1}{3}\right) + g\left(0, \frac{1}{3}\right) + g\left(\frac{1}{3}, 0\right) = \frac{29}{9} \\ 4u_2 - u_1 - u_4 &= G(u_2) + g(u_{13}) + g(u_{15}) = G\left(\frac{2}{3}, \frac{1}{3}\right) + g\left(1, \frac{1}{3}\right) + g\left(\frac{2}{3}, 0\right) = \frac{49}{9} \\ 4u_3 - u_4 - u_1 &= G(u_3) + g(u_7) + g(u_9) = G\left(\frac{1}{3}, \frac{2}{3}\right) + g\left(0, \frac{2}{3}\right) + g\left(\frac{1}{3}, 1\right) = \frac{37}{9} \\ 4u_4 - u_3 - u_2 &= G(u_4) + g(u_{12}) + g(u_{10}) = G\left(\frac{2}{3}, \frac{2}{3}\right) + g\left(1, \frac{2}{3}\right) + g\left(\frac{2}{3}, 1\right) = \frac{28}{9} \end{aligned}$$

Ezek alapján fel tudjuk írni, az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszert:

$$\underbrace{\begin{pmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{pmatrix}}_{=\mathbf{A}} \underbrace{\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}}_{=\mathbf{x}} = \frac{1}{9} \underbrace{\begin{pmatrix} 29 \\ 49 \\ 37 \\ 28 \end{pmatrix}}_{=\mathbf{b}}.$$

Ennek az egyenletrendszernek persze könnyen kiszámolható a pontos megoldása:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{216} \begin{pmatrix} 403 \\ 494 \\ 422 \\ 397 \end{pmatrix} \approx \begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}.$$

Most pedig vizsgáljuk meg, hogyan közelítik a relaxációs eljárások a pontos megoldást! A (3.4), (3.10) és (3.7) egyenleteket felhasználva mind a három módszerre írtam egy eljárást Matlabban. Ezeket az eljárásokat futtattam le különböző ω értékekre, majd a kapott eredményeket összefoglaló táblázatokba írtam.

$\omega=0.5$	JOR-módszer		SOR-módszer		GSOR-módszer	
	$\mathbf{x}^{(k)}$	\mathbf{e}_k	$\mathbf{x}^{(k)}$	\mathbf{e}_k	$\mathbf{x}^{(k)}$	\mathbf{e}_k
$k=5$	$\begin{pmatrix} 1.3941 \\ 1.8104 \\ 1.4875 \\ 1.3672 \end{pmatrix}$	$\begin{pmatrix} 0.4716 \\ 0.4767 \\ 0.4662 \\ 0.4707 \end{pmatrix}$	$\begin{pmatrix} 1.4426 \\ 1.9140 \\ 1.5911 \\ 1.5227 \end{pmatrix}$	$\begin{pmatrix} 0.4231 \\ 0.3730 \\ 0.3626 \\ 0.3152 \end{pmatrix}$	$\begin{pmatrix} 1.4966 \\ 2.0297 \\ 1.7068 \\ 1.6876 \end{pmatrix}$	$\begin{pmatrix} 0.3692 \\ 0.2573 \\ 0.2469 \\ 0.1503 \end{pmatrix}$
$k=10$	$\begin{pmatrix} 1.7539 \\ 2.1750 \\ 1.8420 \\ 1.7261 \end{pmatrix}$	$\begin{pmatrix} 0.1119 \\ 0.1120 \\ 0.1117 \\ 0.1118 \end{pmatrix}$	$\begin{pmatrix} 1.7871 \\ 2.2202 \\ 1.8872 \\ 1.7816 \end{pmatrix}$	$\begin{pmatrix} 0.0787 \\ 0.0668 \\ 0.0665 \\ 0.0564 \end{pmatrix}$	$\begin{pmatrix} 1.8207 \\ 2.2613 \\ 1.9283 \\ 1.8244 \end{pmatrix}$	$\begin{pmatrix} 0.0451 \\ 0.0258 \\ 0.0254 \\ 0.0136 \end{pmatrix}$
$k=50$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-5} \begin{pmatrix} 0.1125 \\ 0.1125 \\ 0.1125 \\ 0.1125 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-7} \begin{pmatrix} 0.9266 \\ 0.7812 \\ 0.7812 \\ 0.6586 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-9} \begin{pmatrix} 0.3607 \\ 0.1804 \\ 0.1804 \\ 0.0902 \end{pmatrix}$
$k=100$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-12} \begin{pmatrix} 0.6364 \\ 0.6364 \\ 0.6368 \\ 0.6368 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-14} \begin{pmatrix} 0.3553 \\ 0.2665 \\ 0.2665 \\ 0.2442 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Először $\omega = 0.5$ érték mellett teszteltem. A táblázatból jól látszik, hogy az összes módszer viszonylag gyorsan konvergál a pontos megoldáshoz. Ez nem is meglepő, hiszen beláttuk, hogy a JOR- és a GSOR-módszer $\omega \in (0, 1]$, a SOR-módszer pedig $\omega \in (0, 2)$ esetén konvergens. A táblázatból az is leolvasható, hogy ezen a konkrét feladaton a GSOR-módszer konvergál a leggyorsabban a megoldáshoz.

$\omega=1$	JOR-módszer		SOR-módszer		GSOR-módszer	
	$\underline{\mathbf{x}}^{(k)}$	$\underline{\mathbf{e}}_k$	$\underline{\mathbf{x}}^{(k)}$	$\underline{\mathbf{e}}_k$	$\underline{\mathbf{x}}^{(k)}$	$\underline{\mathbf{e}}_k$
$k=5$	$\begin{pmatrix} 1.7995 \\ 2.2292 \\ 1.8958 \\ 1.7717 \end{pmatrix}$	$\begin{pmatrix} 0.0663 \\ 0.0579 \\ 0.0579 \\ 0.0663 \end{pmatrix}$	$\begin{pmatrix} 1.8601 \\ 2.2842 \\ 1.9509 \\ 1.8365 \end{pmatrix}$	$\begin{pmatrix} 0.0057 \\ 0.0028 \\ 0.0028 \\ 0.0014 \end{pmatrix}$	$\begin{pmatrix} 1.8601 \\ 2.2842 \\ 1.9509 \\ 1.8365 \end{pmatrix}$	$\begin{pmatrix} 0.0057 \\ 0.0028 \\ 0.0028 \\ 0.0014 \end{pmatrix}$
$k=10$	$\begin{pmatrix} 1.8639 \\ 2.2850 \\ 1.9516 \\ 1.8362 \end{pmatrix}$	$\begin{pmatrix} 0.0018 \\ 0.0021 \\ 0.0021 \\ 0.0018 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-5} \begin{pmatrix} 0.5528 \\ 0.2764 \\ 0.2764 \\ 0.1382 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-5} \begin{pmatrix} 0.5528 \\ 0.2764 \\ 0.2764 \\ 0.1382 \end{pmatrix}$
$k=50$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-14} \begin{pmatrix} 0.1554 \\ 0.1332 \\ 0.1554 \\ 0.1776 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$
$k=100$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Másodszor az $\omega = 1$ esetet teszteltem, tehát a JOR-módszer itt valójában a Jacobi-iterációt, a SOR- és a GSOR-módszerek pedig a Gauss–Seidel-iterációt adták. A táblázatban is látszik, hogy utóbbi kettő lefuttatása után valóban ugyanazokat a közelítő vektorokat kaptuk.

$\omega=1.5$	JOR-módszer		SOR-módszer		GSOR-módszer	
	$\underline{\mathbf{x}}^{(k)}$	$\underline{\mathbf{e}}_k$	$\underline{\mathbf{x}}^{(k)}$	$\underline{\mathbf{e}}_k$	$\underline{\mathbf{x}}^{(k)}$	$\underline{\mathbf{e}}_k$
$k=5$	$\begin{pmatrix} 1.4545 \\ 2.7000 \\ 2.3563 \\ 1.4259 \end{pmatrix}$	$\begin{pmatrix} 0.4112 \\ 0.4130 \\ 0.4026 \\ 0.4121 \end{pmatrix}$	$\begin{pmatrix} 1.9812 \\ 2.3583 \\ 2.0145 \\ 1.8667 \end{pmatrix}$	$\begin{pmatrix} 0.1154 \\ 0.0712 \\ 0.0608 \\ 0.0287 \end{pmatrix}$	$\begin{pmatrix} 1.9254 \\ 2.2680 \\ 1.9243 \\ 1.8502 \end{pmatrix}$	$\begin{pmatrix} 0.0596 \\ 0.0190 \\ 0.0294 \\ 0.0122 \end{pmatrix}$
$k=10$	$\begin{pmatrix} 3.1161 \\ 1.0365 \\ 0.7035 \\ 3.0884 \end{pmatrix}$	$\begin{pmatrix} 1.2504 \\ 1.2506 \\ 1.2502 \\ 1.2504 \end{pmatrix}$	$\begin{pmatrix} 1.8615 \\ 2.2858 \\ 1.9528 \\ 1.8381 \end{pmatrix}$	$\begin{pmatrix} 0.0043 \\ 0.0013 \\ 0.0009 \\ 0.0001 \end{pmatrix}$	$\begin{pmatrix} 1.8582 \\ 2.2876 \\ 1.9546 \\ 1.8376 \end{pmatrix}$	$\begin{pmatrix} 0.0076 \\ 0.0006 \\ 0.0009 \\ 0.0004 \end{pmatrix}$
$k=50$	$\begin{pmatrix} 9408.7 \\ -9404.6 \\ -9404.9 \\ 9408.7 \end{pmatrix}$	$\begin{pmatrix} 9406.9 \\ 9406.9 \\ 9406.9 \\ 9406.9 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-14} \begin{pmatrix} 0.2665 \\ 0.1332 \\ 0.1554 \\ 0.0444 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$10^{-13} \begin{pmatrix} 0.4796 \\ 0.0133 \\ 0.0133 \\ 0 \end{pmatrix}$
$k=100$	$10^8 \begin{pmatrix} 6.5909 \\ -6.5909 \\ -6.5909 \\ 6.5909 \end{pmatrix}$	$10^8 \begin{pmatrix} 6.5909 \\ 6.5909 \\ 6.5909 \\ 6.5909 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1.8657 \\ 2.2870 \\ 1.9537 \\ 1.8380 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Végül $\omega = 1.5$ esetén futtattam le az eljárásokat. A korábban leírt konvergenciára vonatkozó bizonyítások alapján csak a SOR-módszerről mondható el biztosan, hogy emellett az ω érték mellett konvergens, amit a kísérlet ki is mutatott. A JOR-módszer láthatóan használhatatlan ilyen ω -val, már a századik iterációs lépésben 10^8 nagyságrendű a megoldástól való eltérés. Ellenben a GSOR-módszer ezen a feladaton, az adott kezdeti vektorral nagyon jól közelíti a megoldást.

A fenti eredmények alapján úgy gondolom, érdemes lehet a továbbiakban is foglalkozni a GSOR-módszerrel; megvizsgálni, hogyan működik a módszer nagyobb méretű mátrixokon, illetve hogy lehet-e olyan konvergencia-tételt megfogalmazni és bizonyítani, ami szerint nem csak az $\omega \in (0, 1]$, hanem ennél tágabb intervallumon is konvergens az iteráció.

Összefoglalás

Az élet különböző területein számos olyan probléma merül fel, mely lineáris algebrai egyenletrendszerek megoldására vezet vissza. Általában igen nagy méretű egyenletrendszereket kell ezek során megoldanunk, ezért nagyon fontos az, hogy ismerjük azokat a numerikus eljárásokat, melyek gyorsan és kis hibával megtalálják ezeknek az egyenletrendszereknek a megoldását.

Szakdolgozatomban bemutattam néhány olyan módszert is, amikkel direkt módon lehet megkeresni a pontos megoldást, és azt is megmutattam, hogy ezek műveletigénye igen nagy, s emiatt van szükség a numerikus módszerekre. Ezután bemutattam a klasszikus iteratív eljárásokat, amikkel kapcsolatban néhány konvergenciára vonatkozó tételt és azok bizonyítását is ismertettem.

Végül a Jacobi-iteráció és a Gauss–Seidel-iteráció relaxálásával foglalkoztam. Megmutattam, hogyan állnak elő az ismert relaxációs módszerek, a JOR- és a SOR-módszer. Szintén leírtam ezeknek az eljárásoknak a konvergenciájáról szóló tételeket és azok bizonyítását. A munka közben témavezetőmmel felismertük a Gauss–Seidel-iteráció egy másik fajta relaxálását, amit GSOR-módszernek neveztünk el. Leírtam, hogy ez a módszer hogyan áll elő a Gauss–Seidel-iterációból, s megfogalmaztam egy, a módszerre vonatkozó konvergencia-kritériumot. Egy konkrét példán összehasonlítottam a három relaxációs módszer működését. Ez a kísérlet arra ösztönöz, hogy az új, GSOR-módszerrel a továbbiakban is érdemes lesz foglalkozni, s megvizsgálni, hogy más példákon, nagyobb egyenletrendszereken is olyan jól működik-e, mint az általam vizsgált feladat megoldása során.

Irodalomjegyzék

- [1] Faragó István, Horváth Róbert: *Numerikus módszerek*, Typotex (2011)
- [2] Freud Róbert: *Lineáris algebra*, ELTE Eötvös Kiadó (2007)
- [3] Stoyan Gisbert, Takó Galina: *Numerikus módszerek I.*, Typotex (2005)
- [4] David M. Young: *Nagy lineáris rendszerek iterációs megoldása*, Műszaki Könyvkiadó (1979)
- [5] Faragó István: *Alkalmazott analízis 1-2*, előadás jegyzet

NYILATKOZAT

Név: Kis Ágnes

ELTE Természettudományi Kar, szak: Matematika BSc

ETR azonosító: KIARABT.ELTE

Szakedolgozat címe:

Lineáris algebrai egyenletrendszerek iteratív megoldási módszerei

A **szakedolgozat** szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló munkám eredménye, saját szellemi termékem, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2012. május 31.

a hallgató aláírása