

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Lineáris algebrai egyenletrendszerek iterációs módszereinek konvergenciája

BSc Szakdolgozat

Készítette: Tujner Zoltán
Matematika BSc, Matematikai elemző szakirány

Témavezető: Dr. Havasi Ágnes
Alkalmazott Analízis és Számításmatematikai Tanszék



Budapest, 2015

Köszönetnyilvánítás

Ezúton szeretném megköszönni témavezetőmnek, Havasi Ágnesnek, hogy szakértelmével és hasznos tanácsaival hozzájárult dolgozatom elkészítéséhez, s lelkes hozzáállásával engem is ösztönzött arra, hogy a felmerülő problémákat továbbgondoljam, feldolgozzam. Emellett szeretném megköszönni családomnak és barátaimnak a sok biztatást és támogatást az egyetemi éveim alatt.

Tartalomjegyzék

1. Lineáris algebrai egyenletrendszerek	3
2. Iterációs módszerek	5
2.1. Az iterációs módszerek általános elve	5
2.2. Jacobi-iteráció	7
2.3. A Jacobi-iteráció konvergenciája	8
2.4. Gauss–Seidel-iteráció	9
2.5. A Gauss–Seidel-iteráció konvergenciája	10
2.6. A módszerek összehasonlítása egy példán	11
2.7. A relaxált Jacobi-iteráció	15
2.8. A relaxált Gauss–Seidel-iteráció	16
3. Richardson-iteráció	17
3.1. A Richardson-iteráció konvergenciája	17
3.2. Richardson-iteráció Csebisev-féle paramétermegválasztással	19
3.3. A Csebisev paramétermegválasztással kapott séma numerikus stabilitása	22
3.4. Számítási feladat	23
4. Összefoglalás	28
5. Függelék	29

Bevezetés

Szakedolgozatom a lineáris algebrai egyenletrendszerek iterációs megoldási módszereivel foglalkozik. Ezek egy vektorsorozatot generálnak, amely megfelelő feltételek teljesülése esetén tart a lineáris egyenletrendszer pontos megoldásához.

Az 1. fejezetben áttekintjük a lineáris algebrai egyenletrendszerek alapjait. Ezután felépítjük az iterációs módszerek általános alakját, majd bemutatjuk két speciális módszert, a Jacobi- illetve a Gauss–Seidel-iteráció különböző alakjait, és feltételeket adunk a konvergenciájukra. Miután egy feladaton keresztül összehasonlítjuk a két módszert, egy relaxációs paraméter bevezetésével finomítjuk ezeket.

A 3. fejezetben rátérünk a Richardson-iterációra, melyben először egy, az orosz irodalomban fellelhető [4], azonban másutt kevésbé ismert, úgynevezett Csebisev-féle paraméterválasztással optimalizáljuk a módszert, melyet később szintén egy orosz irodalomból [3] ismert eljárás segítségével numerikusan stabilizálunk. Végül a dolgozatban vizsgált iterációs módszereket egy egyszerű másodrendű közönséges differenciálegyenlethez tartozó peremértékfeladat numerikus megoldására alkalmazzuk, és összehasonlítjuk a pontosságukat.

1. fejezet

Lineáris algebrai egyenletrendszerek

A lineáris algebrai egyenletrendszerek általános alakját a következőképpen írhatjuk fel: Adottak az $a_{ij}, f_j \in \mathbb{R}$ ($i = 1, 2, \dots, m, j = 1, 2, \dots, n$) számok, és keressük azon x_j ($j = 1, 2, \dots, n$) számokat, amelyekre teljesülnek a következő egyenletek:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= f_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= f_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= f_m \end{aligned}$$

Az egyenletrendszer felírható mátrixosan, $Ax = f$ alakban is, ahol A az a_{ij} együtthatókból képzett $m \times n$ -es, úgynevezett együtthatómátrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

f az f_i szabad tagokból képzett oszlopvektor :

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$$

x pedig az ismeretlenek n -elemű oszlopvektora.

A lineáris algebrából ismeretes az ilyen alakú egyenletrendszerek megoldhatóságának szükséges és elégséges feltétele. Ehhez először definiáljuk az egyenletrendszer kibővített mátrixát.

1.0.1. Definíció. Az $Ax = f$ egyenletrendszer kibővített mátrixának azt a mátrixot nevezzük, amelyet A -ból úgy kapunk, hogy jobbról kiegészítjük az f oszlopvektorral. Jelölése: $A|f$

$$A|f = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & f_1 \\ a_{21} & a_{22} & \dots & a_{2n} & f_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & f_m \end{pmatrix}$$

1.0.2. Tétel (Kronecker–Capelli-tétel). Egy $Ax = f$ lineáris egyenletrendszer akkor és csak akkor oldható meg, ha az A együtthatómátrix és a f vektorral kibővített együtthatómátrix rangja megegyezik: $r(A) = r(A|f)$.

Ha a megoldhatóság ezen feltétele teljesül, akkor a megoldások számát a következő tétel segítségével határozhatjuk meg.

1.0.3. Tétel. Ha az egyenletrendszer megoldható és $r(A) < n$, akkor végtelen sok megoldás van, ha $r(A) = n$, akkor egyértelmű a megoldás.

2. fejezet

Iterációs módszerek

2.1. Az iterációs módszerek általános elve

A lineáris algebrai egyenletrendszerek megoldási módszereit két nagy csoportra bontjuk. Ezek a direkt illetve az iterációs módszerek.

Direkt módszereknek azokat a módszereket nevezzük, amelyekkel pontosan számolva az egyenletrendszer pontos megoldását kapnánk. Ezek közé tartoznak az eliminációs módszerek, a felbontási módszerek és a Cramer-szabály. Hátrányuk, hogy csak kisebb méretű egyenletrendszerek oldhatók meg velük reális időn belül. Így a gyakorlati feladatok esetén (amelyekben gyakran nagyméretű az együtthatómátrix) az iterációs módszereket használjuk, melyek nagy előnye a direkt módszerekkel szemben a kisebb tárigeny, illetve hogy kihasználhatók a gyakran meglévő hozzávetőleges információk is a megoldás várható értékeiről.

Tegyük fel, hogy az $Ax = f$ egyenletrendszer együtthatómátrixa négyzetes ($A \in \mathbb{R}^{m \times m}$) és determinánsa nullától különbözik ($\det A \neq 0$). Ekkor az (1.0.3) tételből következően az egyenletrendszernek létezik egyetlen megoldása. Az iterációs módszerek ezen egyértelmű megoldás közelítő meghatározására adnak lehetőséget.

Az iterációs módszerek általában olyan konvergens sorozatot konstruálnak, melyek határértéke az egyenletrendszer megoldása. Egylépésesnek nevezzük az iterációt, ha a következő közelítő megoldást mindig az utoljára kiszámolt eredmény felhasználásával számítjuk ki. Az iterációs módszerek általános alakjának a felírásához bontsuk fel az A mátrixot a következőképpen:

$$A = P - Q,$$

ahol Q reguláris. Ekkor

$$Ax = Px - Qx = f,$$

tehát

$$Px = f + Qx,$$

amiből az iterációs eljárás:

$$\begin{aligned} Px^{(k+1)} &= Qx^{(k)} + f \\ x^{(k+1)} &= P^{-1}Qx^{(k)} + P^{-1}f \end{aligned}$$

Ebből az egyenletből $B := P^{-1}Q$ és $v := P^{-1}f$ jelöléssel az egylépéses iterációs módszerek általános alakját kapjuk:

$$x^{(k+1)} = Bx^{(k)} + v, \quad k = 0, 1, \dots,$$

ahol $x^{(0)}$ adott, $x^{(k)}$ az iteráció k . lépésében kapott közelítés, és a B , úgynevezett iterációs mátrix függhet k -től, azaz minden egyes iterációs lépésben változhat.

Az egylépéses iterációs módszerek kanonikus alakja pedig a következőképpen írható fel:

$$B_k \frac{x^{(k+1)} - x^{(k)}}{\tau_k} + Ax^{(k)} = f,$$

ahol $B_k \in \mathbb{R}^{m \times m}$ adott reguláris mátrixok, τ_k pedig adott \mathbb{R} -beli paraméterek.

2.1.1. Definíció. Azt mondjuk, hogy az $x^{(k+1)} = Bx^{(k)} + v, k = 0, 1, \dots$ iteráció adott $x^{(0)}$ mellett konvergál, ha az $x^{(k)}$ sorozat konvergens valamely $\|\cdot\|$ vektor-normában. Ha bármely $x^{(0)}$ -ra konvergál, akkor az iterációs módszert konvergensnek nevezzük.

2.1.2. Tétel. Ha $\|B\|$ valamelyik indukált mátrixnormában < 1 , akkor konvergens az iteráció.

2.1.3. Tétel (A konvergencia szükséges és elégséges feltétele). Az

$$x^{(k+1)} = Bx^{(k)} + v \quad k = 0, 1, \dots$$

iteráció tetszőleges $x^{(0)} \in \mathbb{R}^m$ kezdeti érték mellett akkor és csak akkor konvergens, ha $\varrho(B) < 1$, ahol $\varrho(B)$ a B együtthatómátrix spektrálsugara, azaz a legnagyobb abszolút értékű sajátértéke.

2.1.4. Definíció. Azokat az iterációs eljárásokat, melyekben minden iterációs lépésnél ugyanazzal a B mátrixszal és τ paraméterrel számolunk, stacionárius iterációknak nevezzük:

$$B \frac{x^{(k+1)} - x^{(k)}}{\tau} + Ax^{(k)} = f \tag{2.1}$$

A következőkben bemutatjuk a legelterjedtebb iterációs módszereket, és megvizsgáljuk azok konvergenciáját.

2.2. Jacobi-iteráció

Legyen

$$Ax = f, \quad \text{ahol } A \in \mathbb{R}^{m \times m}, f \in \mathbb{R}^m, \det A \neq 0.$$

Az A mátrixot bontsuk fel a következő módon:

$$A = A_1 + D + A_2,$$

ahol A_1 az A mátrix szigorúan alsó háromszögű része, azaz a diagonális alatti rész, A_2 a mátrix szigorúan felső háromszögű része, D pedig olyan mátrix, melynek főátlójában az A mátrix diagonálisában lévő elemek szerepelnek (feltesszük, hogy ezek egyike sem 0). Vagyis az egyenletrendszer felírható az alábbi alakban:

$$\begin{aligned} Ax = f &\Leftrightarrow (A_1 + D + A_2)x = f \\ Dx &= -(A_1 + A_2)x + f \\ Dx^{(k+1)} &= -(A_1 + A_2)x^{(k)} + f \end{aligned} \quad (2.2)$$

Mivel D invertálható, kifejezhető $x^{(k+1)}$, és így a Jacobi-iteráció mátrixos alakja:

$$x^{(k+1)} = \underbrace{-D^{-1}(A_1 + A_2)}_{:=B_J} x^{(k)} + \underbrace{D^{-1}f}_v,$$

ahol B_J az iterációs mátrixot jelöli.

A (2.2)-ből némi átalakítással megkaphatjuk a módszer kanonikus alakját:

$$\begin{aligned} Dx^{(k+1)} &= -(A_1 + A_2)x^{(k)} + f \\ Dx^{(k+1)} - Dx^{(k)} + Dx^{(k)} + (A_1 + A_2)x^{(k)} &= f \\ D(x^{(k+1)} - x^{(k)}) + \underbrace{(D + A_1 + A_2)}_A x^{(k)} &= f \\ D(x^{(k+1)} - x^{(k)}) + Ax^{(k)} &= f \end{aligned}$$

Írjuk fel komponensenként az egyenletrendszert.

$$Ax = f \Leftrightarrow a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ii}x_i + \cdots + a_{im}x_m = f_i, \quad i = 1, 2, \dots, m \quad (2.3)$$

Tegyük fel, hogy $a_{ii} \neq 0$, ekkor a (2.3) átírható a_{ii} -vel leosztva és átrendezve:

$$x_i = - \left[\frac{a_{i1}}{a_{ii}}x_1 + \frac{a_{i2}}{a_{ii}}x_2 + \cdots + \frac{a_{ii-1}}{a_{ii}}x_{i-1} + \frac{a_{ii+1}}{a_{ii}}x_{i+1} + \cdots + \frac{a_{im}}{a_{ii}}x_m \right] + \frac{f_i}{a_{ii}} \quad (2.4)$$

$\forall i = 1, 2, \dots, m$

Ebből felépíthető egy iteráció, ahol $x_i^{(k)}$ az i -edik ismeretlen k -edik közelítése:

$$x_i^{(k+1)} = - \left[\frac{a_{i1}}{a_{ii}} x_1^{(k)} + \dots + \frac{a_{im}}{a_{ii}} x_m^{(k)} \right] + \frac{f_i}{a_{ii}}, \quad (2.5)$$

ahol $x_i^{(0)}$ adott tetszőleges érték és $k = 0, 1 \dots$

Összevont formában:

$$x_i^{(k+1)} = - \sum_{\substack{j=1 \\ j \neq i}}^m \frac{a_{ij}}{a_{ii}} x_j^{(k)} + \frac{f_i}{a_{ii}} \quad (2.6)$$

Ezt nevezzük a Jacobi-iteráció koordinátás alakjának.

2.3. A Jacobi-iteráció konvergenciája

Ebben a szakaszban elégséges feltételt adunk a Jacobi-iteráció konvergenciájára.

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

$$a_{ii} > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad i = 1, 2 \dots n$$

2.3.2. Tétel. Legyen $A \in \mathbb{R}^{n \times n}$ egy szimmetrikus és szigorúan diagonálisan domináns mátrix. Tegyük fel, hogy B invertálható. Ha τ olyan paraméter, amelyre $B - 0,5\tau A > 0$ akkor a (2.1) iteráció konvergens, és tart az egyenletrendszer megoldásához.

Ezt a tételt a Jacobi-iterációra alkalmazva az alábbi állításhoz jutunk.

2.3.3. Állítás. Ha $A \in \mathbb{R}^{m \times m}$ mátrix szimmetrikus és szigorúan diagonálisan domináns, akkor a Jacobi-módszer konvergens.

Bizonyítás: A Jacobi-iteráció esetén $D \frac{x^{(k+1)} - x^{(k)}}{1} + Ax^{(k)} = f$, azaz elegendő belátni a $D - 0,5A > 0$ egyenlőtlenséget a konvergencia teljesüléséhez.

$$D - 0,5A > 0 \Leftrightarrow 2D - A > 0$$

Szükség van még az alábbi becslésre:

A számtani és mértani közepek közötti összefüggésből tudjuk, hogy

$$\begin{aligned} \sqrt{x_i^2 x_j^2} &\leq \frac{x_i^2 + x_j^2}{2} \\ |a_{ij}| \sqrt{x_i^2 x_j^2} &\leq |a_{ij}| \frac{x_i^2 + x_j^2}{2} \\ 2a_{ij}x_i x_j &\leq 2|a_{ij}||x_i||x_j| \leq |a_{ij}|(x_i^2 + x_j^2) = |a_{ij}|x_i^2 + |a_{ij}|x_j^2 \\ a_{ij}x_i x_j &\leq \frac{1}{2}|a_{ij}|x_i^2 + \frac{1}{2}|a_{ij}|x_j^2 \\ \sum_{i,j=1}^m a_{ij}x_i x_j &\leq \frac{1}{2} \sum_{i,j=1}^m |a_{ij}|x_i^2 + \frac{1}{2} \sum_{i,j=1}^m |a_{ij}|x_j^2 \stackrel{\text{szimmetria}}{=} \frac{1}{2} \sum_{i,j=1}^m |a_{ij}|x_i^2 + \\ &+ \frac{1}{2} \sum_{i,j=1}^m |a_{ji}|x_i^2 = \frac{1}{2} \sum_{i,j=1}^m |a_{ij}|x_i^2 + \frac{1}{2} \sum_{i,j=1}^m |a_{ij}|x_i^2 = \sum_{i,j=1}^m |a_{ij}|x_i^2 \end{aligned}$$

$$\begin{aligned} \langle Ax, x \rangle &= \sum_{i,j=1}^m a_{ij}x_i x_j \leq \sum_{i=1}^m \sum_{j=1}^m a_{ij}x_i x_j = \\ &= \sum_{i=1}^m x_i^2 \sum_{j=1}^m |a_{ij}| = \sum_{i=1}^m x_i^2 (a_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^m |a_{ij}|) \end{aligned}$$

Mivel A -ról feltettük, hogy szigorúan diagonálisan domináns, ezért $a_{ii} > \sum_{\substack{j=1 \\ j \neq i}}^m |a_{ij}|$,

így:

$$\begin{aligned} \sum_{i=1}^m x_i^2 (a_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^m |a_{ij}|) &< \sum_{i=1}^m x_i^2 (a_{ii} + a_{ii}) = 2 \sum_{i=1}^m a_{ii} x_i^2 = 2 \langle Dx, x \rangle \\ \Rightarrow \langle Ax, x \rangle &< 2 \langle Dx, x \rangle = \langle 2Dx, x \rangle \end{aligned}$$

így $A < 2D$, azaz $2D - A > 0$. ■

2.4. Gauss–Seidel-iteráció

A Gauss–Seidel-iteráció abban különbözik a Jacobi-iterációtól, hogy a $(k+1)$ -edik közelítés i -edik komponensének kiszámításához felhasználjuk a $(k+1)$ -edik közelítés már addig kiszámolt komponenseit, vagyis az $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ értékeket:

$$x_i^{(k+1)} = - \left[\frac{a_{i1}}{a_{ii}} x_1^{(k+1)} + \dots + \frac{a_{i,i-1}}{a_{ii}} x_{i-1}^{(k+1)} \right] - \left[\frac{a_{i,i+1}}{a_{ii}} x_{i+1}^{(k)} + \dots + \frac{a_{im}}{a_{ii}} x_m^{(k)} \right] + \frac{f_i}{a_{ii}},$$

azaz

$$x_i^{(k+1)} = - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(k+1)} - \sum_{j=i+1}^m \frac{a_{ij}}{a_{ii}} x_j^{(k)} + \frac{f_i}{a_{ii}} \quad \forall i = 1, 2, \dots, m$$

A módszer a Jacobi-iterációhoz hasonló elven felírható mátrixos alakban:

$$(A_1 + D + A_2)x = f$$

$$(A_1 + D)x = -A_2x + f$$

$$(A_1 + D)x^{(k+1)} = -A_2x^{(k)} + f \quad (2.7)$$

$$x^{(k+1)} = \underbrace{-(A_1 + D)^{-1}A_2}_{:=B_{G-S}} x^{(k)} + \underbrace{(A_1 + D)^{-1}f}_v$$

Ezzel megkaptuk a Gauss–Seidel-iteráció mátrixos alakját, ahol B_{G-S} jelöli az iterációs mátrixot.

A (2.7)-ből kifejezhető a kanonikus alak:

$$(A_1 + D)x^{(k+1)} = -A_2x^{(k)} + f$$

$$(A_1 + D)x^{(k+1)} - (A_1 + D)x^{(k)} + (A_1 + D)x^{(k)} + A_2x^{(k)} = f$$

$$(A_1 + D)(x^{(k+1)} - x^{(k)}) + (A_1 + D + A_2)x^{(k)} = f$$

$$(A_1 + D)(x^{(k+1)} - x^{(k)}) + Ax^{(k)} = f$$

2.5. A Gauss–Seidel-iteráció konvergenciája

A Gauss–Seidel-iteráció konvergenciájára egy általánosabb tételből vezetünk le elégséges feltételt.

2.5.1. Tétel (Kahan-tétel). *Legyen A szimmetrikus, pozitív definit mátrix. Ekkor $\omega \in (0, 2)$ esetén a*

$$(D + \omega A_1) \frac{x^{(k+1)} - x^{(k)}}{\omega} + Ax^{(k)} = f \quad (2.8)$$

módszer konvergens.

Bizonyítás:

$$\begin{aligned} \langle Ax, x \rangle &= \langle (A_1 + D + A_2)x, x \rangle = \langle A_1x, x \rangle + \langle Dx, x \rangle + \langle A_2x, x \rangle = \langle A_1x, x \rangle + \\ &\langle Dx, x \rangle + \langle x, A_2^T x \rangle = 2\langle A_1x, x \rangle + \langle Dx, x \rangle \end{aligned}$$

$B = D + A_1\omega$ és $\tau = \omega$ megválasztása esetén a (2.1) alakú stacionárius iterációt kapjuk, melyről tudjuk, hogy konvergens olyan τ megválasztása esetén, melyre teljesül a $B - 0,5\tau A > 0$ egyenlőtlenség:

$$\underbrace{(D + A_1\omega)}_B - 0,5 \underbrace{\omega}_\tau A > 0 \Rightarrow B - 0,5\tau A > 0$$

Ebből:

$$\begin{aligned} \langle (B - 0,5\tau A)x, x \rangle &= \langle Bx, x \rangle - 0,5\tau \langle Ax, x \rangle = \\ &= \langle (D + A_1\omega)x, x \rangle - 0,5\omega \langle Ax, x \rangle = \\ &= \langle Dx, x \rangle + \omega \langle A_1x, x \rangle - 0,5\omega [2\langle A_1x, x \rangle + \langle Dx, x \rangle] = \\ &= \langle Dx, x \rangle + \omega \langle A_1x, x \rangle - \omega \langle A_1x, x \rangle - 0,5\omega \langle Dx, x \rangle = \\ &= (1 - 0,5\omega) \underbrace{\langle Dx, x \rangle}_{>0} \geq 0 \\ &\Rightarrow 1 - 0,5\omega > 0 \Rightarrow \omega < 2 \end{aligned}$$

■

2.5.2. Következmény. *Mivel a Gauss–Seidel-iteráció nem más, mint a (2.8) módszer $\omega = 1$ választása esetén, így szimmetrikus, pozitív definit mátrixokra konvergens.*

2.6. A módszerek összehasonlítása egy példán

Adott a következő lineáris egyenletrendszer:

$$\begin{aligned} x_1 - 2x_2 + 2x_3 &= 1 \\ -x_1 + x_2 - x_3 &= 1 \\ -2x_1 - 2x_2 + x_3 &= 1 \end{aligned}$$

amelyet $Ax = f$ alakban felírva:

$$A = \begin{pmatrix} 1 & -2 & 2 \\ -1 & 1 & -1 \\ -2 & -2 & 1 \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad f = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Mivel kisméretű az egyenletrendszer, ezért ki tudjuk számolni a pontos megoldását iterációs módszerek nélkül is. Használjuk a Gauss–Jordan-eliminációt:

$$\begin{array}{l}
\text{I} \\
\text{II} \\
\text{III} \\
\text{I}^* \\
\text{II}^* \\
\text{III}^* \\
\text{I}^{**} \\
\text{II}^{**} \\
\text{III}^{**} \\
\text{I}^{***} \\
\text{II}^{***} \\
\text{III}^{***}
\end{array}
\left(\begin{array}{ccc|c}
1 & -2 & 2 & 1 \\
-1 & 1 & -1 & 1 \\
-2 & -2 & 1 & 1 \\
1 & -2 & 2 & 1 \\
0 & -1 & 1 & 2 \\
0 & -6 & 5 & 3 \\
1 & 0 & 0 & -3 \\
0 & 1 & -1 & -2 \\
0 & 0 & -1 & -9 \\
1 & 0 & 0 & -3 \\
0 & 1 & 0 & 7 \\
0 & 0 & 1 & 9
\end{array} \right)
\begin{array}{l}
\\
\text{II} + \text{I} \\
\text{III} + 2 \cdot \text{I} \\
\text{I} + 2 \cdot \text{II} \\
-\text{II} \\
\text{III} + 6 \cdot \text{I} \\
\\
\text{II} + \text{III} \\
-\text{III} \\
\\
\\
\\
\end{array}$$

Ezzel megkaptuk a pontos megoldást, amely $x_1 = -3$, $x_2 = 7$ és $x_3 = 9$. Nézzük meg, hogyan oldható meg a Jacobi-iteráció segítségével. Első lépésben meg kell vizsgálnunk, hogy a feladatra konvergens-e a módszer. Ehhez először fel kell bontanunk az A együtthatómátrixot $A = A_1 + D + A_2$ alakban:

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -2 & -2 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Mint tudjuk, a Jacobi-iteráció mátrixos alakja:

$$x^{(k+1)} = \underbrace{-D^{-1}(A_1 + A_2)}_{:=B_J} x^{(k)} + \underbrace{D^{-1}f}_v$$

ahol B_J iterációs mátrixot kell vizsgálni a konvergencia teljesüléséhez.

$$B_J = -D^{-1}(A_1 + A_2) = - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{pmatrix}$$

Először a (2.1.2) tétel alapján vizsgáljuk meg a konvergenciát, azaz a kapott iterációs mátrix valamely indukált mátrixnormában kisebb-e, mint 1.

$$\| B_J \|_1 = \max_{j=1}^n \sum_{i=1}^n |b_{ij}| = \max\{3, 4, 3\} = 4 > 1$$

$$\|B_J\|_\infty = \max_{i=1}^n \sum_{j=1}^n |b_{ij}| = \max\{4, 2, 4\} = 4 > 1$$

$$\|B_J\|_2 = \sqrt{\lambda_{\max}(B_J^T B)} = \sqrt{12} > 1$$

Mint látható, ez a feltétel nem alkalmazható ebben az esetben.

Megjegyzés: Ha a mátrixban találhatóak abszolút értékben 1-nél nagyobb elemek, akkor a (2.1.1) tétel feltétele nem használható.

Meg kell vizsgálni a szükséges és elégséges feltételt is, azaz a spektrálsugár értékét. A sajátértékek meghatározásához írjuk fel a karakterisztikus polinomot:

$$\det(B_J - \lambda I) = \begin{vmatrix} -\lambda & 2 & -2 \\ 1 & -\lambda & 1 \\ 2 & 2 & -\lambda \end{vmatrix} = -\lambda^3 + 4 - 4 - 4\lambda + 2\lambda + 2\lambda = -\lambda^3 = 0$$

azaz a sajátértékek $\lambda_{1,2,3} = 0$.

A mátrix spektrálsugara :

$$\rho(B_J) = \max |\lambda_i| = 0 < 1 \Rightarrow \text{Konvergencia a módszer}$$

Megjegyzés: Ha az iterációs mátrix spektrálsugara 0, akkor véges iterációra számíthatunk, legfeljebb n lépésben konvergál az iteráció, ahol n az ismeretlenek száma.

Most, hogy tudjuk, konvergencia a módszer, elkezdhetjük alkalmazni az iterációt.

Legyen az $x^{(0)}$ kiinduló vektor a $[0 \ 0 \ 0]^T$

$k = 0$

$$x^{(1)} = B_J x^{(0)} + v = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$k = 1$

$$x^{(2)} = B_J x^{(1)} + v = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$$

$k = 2$

$$x^{(3)} = B_J x^{(2)} + v = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 6 \\ 8 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 7 \\ 9 \end{pmatrix}$$

Ezzel meg is kaptuk a pontos megoldást, azonban nagyobb méretű mátrixok esetén, amikor nem vagyunk tisztában a megoldással, akkor a következő lépésben ugyanezt

az eredményt kell kapnunk, és megbizonyosodhatunk az eredményünk helyességéről.

$k = 3$

$$x^{(4)} = B_j x^{(3)} + v = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} -3 \\ 7 \\ 9 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 6 \\ 8 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 7 \\ 9 \end{pmatrix}$$

Oldjuk meg a feladatot a Gauss-Seidel-iterációval is.

$$x^{(k+1)} = \underbrace{-(A_1 + D)^{-1} A_2}_{:=B_{G-S}} x^{(k)} + \underbrace{(A_1 + D)^{-1} f}_v$$

Első lépésben meg kell határozunk a B_{G-S} iterációs mátrixot.

$$B_{G-S} = -(A_1 + D)^{-1} A_2 = - \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & -2 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & -8 \\ 0 & 2 & -1 \\ 0 & 8 & -6 \end{pmatrix}$$

Az (2.1.1) tétel feltétele biztosan nem teljesül, hiszen vannak 1-nél nagyobb elemek a mátrixban. Ezért a szükséges és elégséges feltétellel kell próbálkoznunk. A sajátértékek meghatározásához fel kell írunk a karakterisztikus polinomot:

$$\det(B_{G-S} - \lambda I) = \begin{vmatrix} -\lambda & 2 & -2 \\ 0 & 2 - \lambda & -1 \\ 0 & 8 & -6 - \lambda \end{vmatrix} = -\lambda(2 - \lambda)(-6 - \lambda) - 8\lambda = 0$$

$$-\lambda(2 - \lambda)(-6 - \lambda) - 8\lambda = -\lambda((2 - \lambda)(-6 - \lambda) - 8) = 0 \quad \Rightarrow \boxed{\lambda_1 = 0}$$

$$(2 - \lambda)(-6 - \lambda) - 8 = -12 - 2\lambda + 6\lambda + \lambda^2 + 8 = \lambda^2 + 4\lambda - 4 = 0 \Rightarrow \boxed{\lambda_{2,3} = -2 \pm 2\sqrt{2}}$$

A mátrix spektrálsugara:

$$\rho(B_{G-S}) = \max |\lambda_i| = 2 + 2\sqrt{2} > 1$$

Következésképpen, a módszer nem konvergens akármilyen kezdővektorra.

2.7. A relaxált Jacobi-iteráció

Vannak esetek, amikor az iterációs mátrix spektrálsugara nem kisebb 1-nél, és ilyenkor nem vagy csak nagyon lassan konvergál az iteráció. Ekkor azt tehetjük, hogy bevezetünk egy paramétert az iterációba, melyet úgy választunk meg, hogy az iteráció konvergens legyen.

Továbbra is a Jacobi-iterációval foglalkozunk, csak egy plusz ω , úgynevezett relaxációs paraméter bevezetésével próbáljuk finomítani a módszert.

Tekintsük a $Dx = -(A_1 + A_2)x + f$, illetve a $Dx = Dx$ egyenleteket. Ezeket rendre szorozzuk be ω , illetve $1 - \omega$ értékekkel:

$$\begin{aligned}\omega Dx &= -\omega(A_1 + A_2)x + \omega f \\ (1 - \omega)Dx &= (1 - \omega)Dx\end{aligned}$$

Ezután adjuk össze a két egyenletet:

$$Dx = (1 - \omega)Dx - \omega(A_1 + A_2)x + \omega f$$

Innen D^{-1} -gyel szorozva, majd felépítve az iterációt, megkapjuk a módszer mátrixos alakját:

$$\begin{aligned}x &= (1 - \omega)x - \omega D^{-1}(A_1 + A_2)x + \omega D^{-1}f \\ x^{(k+1)} &= \underbrace{\left((1 - \omega)I - \omega D^{-1}(A_1 + A_2) \right)}_{B_{J(\omega)}} x^{(k)} + \omega D^{-1}f\end{aligned}$$

A koordinátás alak a következőképpen írható fel:

$$\begin{aligned}x &= (1 - \omega)x - \omega (D^{-1}(A_1 + A_2)x + D^{-1}f) \\ x^{(k+1)} &= (1 - \omega)x^{(k)} - \omega \left(\sum_{\substack{j=1 \\ j \neq i}}^m \frac{a_{ij}}{a_{ii}} x_j^{(k)} + \frac{f_i}{a_{ii}} \right)\end{aligned}$$

Észrevehető, hogy $\omega = 1$ esetén éppen a Jacobi-iterációt kapjuk vissza.

Ha $\omega > 1$, akkor túlrelaxációról, ha $0 < \omega < 1$, akkor alulrelaxációról beszélünk.

2.7.1. Tétel. *Szimmetrikus, szigorúan diagonálisan domináns A mátrix esetén a relaxált Jacobi-módszer $\omega \in (0, 1]$ esetén konvergens.*

2.8. A relaxált Gauss–Seidel-iteráció

Induljunk el ugyanazon logika alapján, mint az előbb:
Tekintsük az $(A_1 + D)x = -A_2x + f$, illetve a $Dx = Dx$ egyenleteket. Ezeket rendre szorozzuk be ω , illetve $1 - \omega$ értékekkel:

$$\begin{aligned}\omega(A_1 + D)x &= -\omega A_2x + \omega f \\ (1 - \omega)Dx &= (1 - \omega)Dx\end{aligned}$$

Ezután adjuk össze a két egyenletet:

$$(\omega A_1 + D)x = (1 - \omega)Dx - \omega A_2x + \omega f$$

Innen $(\omega A_1 + D)^{-1}$ -gyel szorozva, majd felépítve az iterációt megkapjuk a módszer mátrixos alakját:

$$\begin{aligned}x &= (\omega A_1 + D)^{-1} [(1 - \omega)D - \omega A_2]x + \omega(\omega A_1 + D)^{-1}f \\ x^{(k+1)} &= \underbrace{(\omega A_1 + D)^{-1} [(1 - \omega)D - \omega A_2]}_{B_{G-S(\omega)}} x^{(k)} + \omega(\omega A_1 + D)^{-1}f\end{aligned}$$

A koordinátás alak felírásához átírjuk az iterációt:

$$\begin{aligned}(\omega A_1 + D)x^{(k+1)} &= (1 - \omega)Dx^{(k)} - \omega A_2x^{(k)} + \omega f \\ Dx^{(k+1)} &= -\omega A_1x^{(k+1)} - \omega A_2x^{(k)} + \omega f + (1 - \omega)Dx^{(k)} \\ x^{(k+1)} &= -\omega \left[\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(k+1)} + \sum_{j=i+1}^m \frac{a_{ij}}{a_{ii}} x_j^{(k)} + \frac{f_i}{a_{ii}} \right] + (1 - \omega)Dx^{(k)}\end{aligned}$$

2.8.1. Tétel. *Ha A szimmetrikus, pozitív definit és $\omega \in (0, 2)$, akkor a relaxációs módszer konvergens.*

3. fejezet

Richardson-iteráció

Tekintsük az $Ax = f$ lineáris algebrai egyenletrendszert, ahol A szimmetrikus, pozitív definit mátrix (azaz minden sajátértéke valós, sőt pozitív). Szorozzuk meg az egyenletrendszer mindkét oldalát egy tetszőleges $\tau \neq 0$ számmal, majd rendezzük át:

$$\begin{aligned}Ax &= f \\ \tau Ax &= \tau f \\ 0 &= -\tau Ax + \tau f \\ x &= x - \tau Ax + \tau f = (I - \tau A)x + \tau f\end{aligned}$$

Ezek alapján definiálható a következő iteráció:

3.0.1. Definíció. *Richardson-iteráció τ paraméterrel - $R_{(\tau)}$*

$$x^{(k+1)} = \underbrace{(I - \tau A)}_{B_{R_{(\tau)}}} x^{(k)} + \underbrace{\tau f}_{C_{R_{(\tau)}}} = B_{R_{(\tau)}} x^{(k)} + C_{R_{(\tau)}}$$

A módszer kanonikus alakja:

$$\frac{x^{(k+1)} - x^{(k)}}{\tau_{k+1}} + Ax^{(k)} = f, \quad k = 0, 1, \dots$$

3.1. A Richardson-iteráció konvergenciája

3.1.1. Tétel. *Ha az $A \in \mathbb{R}^{n \times n}$ mátrix szimmetrikus, pozitív definit, és sajátértékeire $\lambda_{\min}(A) = \lambda_1 \leq \dots \leq \lambda_n = \lambda_{\max}(A)$ teljesül, akkor $R_{(\tau)}$ konvergens, ha $\tau \in \left(0, \frac{2}{\lambda_{\max}(A)}\right)$. Ekkor az optimális paraméter és a hozzá kapcsolódó optimális spektrálsugár pedig:*

$$\tau_{opt} = \frac{2}{\lambda_{\max}(A) + \lambda_{\min}(A)}, \quad \varrho_{opt} := \varrho(B_{R_{(\tau_{opt})}}) = \frac{\lambda_{\max}(A) - \lambda_{\min}(A)}{\lambda_{\max}(A) + \lambda_{\min}(A)}$$

Bizonyítás: $B_{R(\tau)} = I - \tau A$, aminek sajátértékei $\lambda_i(\tau) = 1 - \tau\lambda_i$, vagyis :

$$\begin{aligned}\lambda_1(\tau) &= 1 - \tau\lambda_1 = 1 - \tau\lambda_{\min}(A) \\ \lambda_2(\tau) &= 1 - \tau\lambda_2 \\ &\vdots \\ \lambda_n(\tau) &= 1 - \tau\lambda_n = 1 - \tau\lambda_{\max}(A)\end{aligned}$$

A $B_{R(\tau)}$ spektrálsugara így $\varrho(B_{R(\tau)}) = \max_{i=1}^n |1 - \tau\lambda_i|$.

Így $B_{R(\tau)}$ konvergens, ha $\varrho(B_{R(\tau)}) < 1$, azaz ha $\tau \in \left(0, \frac{2}{\lambda_{\max}(A)}\right)$.

Az optimális paramétert ott kapjuk meg, ahol a $\varrho(B_{R(\tau)})$ minimális, ezt a

$$\begin{aligned}|\lambda_n(\tau)| &= |\lambda_1(\tau)| \\ |1 - \tau\lambda_{\max}(A)| &= |1 - \tau\lambda_{\min}(A)|\end{aligned}$$

egyenlet megoldása adja:

$$\begin{aligned}-1 + \tau\lambda_{\max}(A) &= 1 - \tau\lambda_{\min}(A) \\ \tau\lambda_{\max}(A) + \tau\lambda_{\min}(A) &= 2 \\ \tau(\lambda_{\max}(A) + \lambda_{\min}(A)) &= 2 \\ \tau_{opt} &= \frac{2}{\lambda_{\max}(A) + \lambda_{\min}(A)}\end{aligned}$$

A τ_{opt} paraméterértékhez tartozó spektrálsugár:

$$\begin{aligned}\varrho(B_{R(\tau)}) &= |\lambda_1(\tau_{opt})| = |1 - \tau_{opt}\lambda_{\min}(A)| = \left|1 - \frac{2}{\lambda_{\max}(A) + \lambda_{\min}(A)}\lambda_{\min}(A)\right| = \\ &= \left|\frac{\lambda_{\max}(A) + \lambda_{\min}(A) - 2\lambda_{\min}(A)}{\lambda_{\max}(A) + \lambda_{\min}(A)}\right| = \frac{\lambda_{\max}(A) - \lambda_{\min}(A)}{\lambda_{\max}(A) + \lambda_{\min}(A)}\end{aligned}$$

■

3.2. Richardson-iteráció Csebisev-féle paramétermegválasztással

Tekintsük az

$$Ax = f$$

lineáris algebrai egyenletrendszert, ahol A szimmetrikus, pozitív definit mátrix.^[4] Ezt a rendszert a Richardson-iteráció alábbi, változó τ paraméterrel felírt változótáival szeretnénk megoldani:

$$\frac{x^{(k+1)} - x^{(k)}}{\tau_{k+1}} + Ax^{(k)} = f, \quad k = 0, 1, \dots \quad (3.1)$$

ahol x_0 adott.

Kérdés a paraméterek optimális megválasztása, vagyis azon $\tau_1, \tau_2, \dots, \tau_n$ értékek meghatározása, amelyek mellett az $x_n - x$ hiba normája minimális. (Az euklideszi normát fogjuk használni.)

A feladat pontos megfogalmazása és megoldása a következő tételben található.

3.2.1. Tétel. *Legyen A szimmetrikus, pozitív definit mátrix, $\lambda_{\min}(A) > 0$ és $\lambda_{\max}(A) > 0$ pedig a legkisebb és a legnagyobb sajátértéke. Legyen adva az n iterációs lépésszám. A (3.1) típusú módszerek közül a legkisebb $\|x^{(n)} - x\|$ hibával az a módszer rendelkezik, amelyre*

$$\tau_k = \frac{\tau_0}{1 + \rho_0 t_k}, \quad k = 1, 2, \dots, n, \quad (3.2)$$

ahol

$$\tau_0 = \frac{2}{\lambda_{\min}(A) + \lambda_{\max}(A)}, \quad \rho_0 = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)} \quad (3.3)$$

$$t_k = \cos \frac{(2k-1)\pi}{2n}, \quad k = 1, 2, \dots, n.$$

(Ezen t_k pontok a $[-1, 1]$ intervallumon értelmezett Csebisev-polinom zérushelyei.)

Ha a τ_k értékeket (3.2), (3.3) szerint választjuk, akkor a hibára

$$\|x^{(n)} - x\| \leq q_n \|x^{(0)} - x\| \quad (3.4)$$

adódik, ahol

$$q_n = \frac{2\rho_1^n}{1 + \rho_1^{2n}}, \quad \rho_1 = \frac{1 - \sqrt{\xi}}{1 + \sqrt{\xi}}, \quad \xi = \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)}. \quad (3.5)$$

Bizonyítás: A $z_k = x^{(k)} - x$ hiba kielégíti a következő egyenletet:

$$\frac{z^{(k+1)} - z^{(k)}}{\tau_{k+1}} + Az^{(k)} = 0, \quad k = 0, 1, \dots, n-1 \quad z^{(0)} = x^{(0)} - x. \quad (3.6)$$

(3.6)-ból kapjuk:

$$z^{(k)} = (I - \tau_k A)(I - \tau_{k-1} A) \dots (I - \tau_1 A)z^{(0)}, \quad k = 1, 2, \dots, n \quad z^{(0)} = x^{(0)} - x. \quad (3.7)$$

Speciálisan, $z^{(n)} = T_n z^{(0)}$, ahol

$$T_n = (I - \tau_n A)(I - \tau_{n-1} A) \dots (I - \tau_1 A). \quad k = 0, 1, \dots, n \quad z^{(0)} = x^{(0)} - x. \quad (3.8)$$

Mivel A szimmetrikus mátrix, így T_n is, 2-es normája a spektrálsugár, és teljesül a

$$\|z^{(n)}\| = \|T_n z^{(0)}\| \leq \|T_n\| \cdot \|z^{(0)}\| \Rightarrow \|z^{(n)}\| \leq |v| \cdot \|z^{(0)}\|, \quad (3.9)$$

ahol v a T_n legnagyobb abszolút értékű sajátértéke. A (3.9) becslés nem javítható, vagyis létezik olyan $z^{(0)}$ vektor, amelyre egyenlőség van. Már csak az a kérdés, hogy $\tau_1, \tau_2, \dots, \tau_n$ milyen megválasztása minimalizálja $|v|$ -t.

Jelölje $\lambda_k, k = 1, 2, \dots, m$ az A sajátértékeit. Az általánosság megszorítása nélkül feltehetjük, hogy

$$0 < \lambda_{\min}(A) = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_m = \lambda_{\max}(A). \quad (3.10)$$

(3.7) alapján

$$|v| = \max_{1 \leq k \leq m} |(1 - \tau_1 \lambda_k)(1 - \tau_2 \lambda_k) \dots (1 - \tau_n \lambda_k)|. \quad (3.11)$$

Látható, hogy

$$|v| \leq \max_{\lambda_{\min}(A) \leq \lambda \leq \lambda_{\max}(A)} |f_n(\lambda)|,$$

ahol $f_n(\lambda) = (1 - \tau_1 \lambda)(1 - \tau_2 \lambda) \dots (1 - \tau_n \lambda)$. Ilyen módon a következő minimax feladathoz jutunk:

$$\min_{\tau_1, \tau_2, \dots, \tau_n} \max_{\lambda_{\min}(A) \leq \lambda \leq \lambda_{\max}(A)} |f_n(\lambda)|. \quad (3.12)$$

A [4] könyv szerint éppen a (3.2),(3.3) szerinti τ_k értékek lesznek megfelelőek, és az eltérés nagysága az adott paraméterek esetén $|v| = q_n$, ahol q_n a (3.5) szerint számítható. Ezzel az állítást beláttuk. ■

A (3.2),(3.3) megválasztás melletti (3.1) iterációs módszer neve: *Richardson-iteráció Csebisev-féle paramétermegválasztással*.

Most határozzuk meg, hány iterációs lépés szükséges ahhoz, hogy ε pontossággal meg tudjuk oldani a feladatot a Csebisev-féle paramétermegválasztással. (3.4)-ből

$$\|x^{(n)} - x\| < \varepsilon \|x^{(0)} - x\|,$$

ha $q_n < \varepsilon$, ahol q_n a (3.5) szerinti. Így az

$$\frac{1 + \rho_1^{2n}}{\rho_1^n} > \frac{2}{\varepsilon}$$

egyenlőtlenséghez jutunk. Ezt megoldva $z = \rho_1^{-n} > 1$ -re:

$$\frac{1}{\rho_1^n} > \frac{1 + \sqrt{1 - \varepsilon^2}}{\varepsilon}$$

Az utóbbi egyenlőtlenség fennáll, ha $1/\rho_1^n \geq 2/\varepsilon$, azaz

$$n \geq n_0(\varepsilon) = \frac{\ln(2/\varepsilon)}{\ln(1/\rho_1)}. \quad (3.13)$$

A legkedvezőtlenebb esetben, amikor $\xi = \lambda_{\min}(A)/\lambda_{\max}(A)$ kicsi

$$\ln \frac{1}{\rho_1} = \ln \left(\frac{1 + \sqrt{\xi}}{1 - \sqrt{\xi}} \right) \approx 2\sqrt{\xi},$$

és ekkor az iterációs számra a következő közelítő összefüggést nyerjük:

$$n \geq n_0(\varepsilon) = \frac{\ln(2/\varepsilon)}{2\sqrt{\xi}}. \quad (3.14)$$

Kis ξ értéke esetén látszik ebből a Csebisev-féle paramétermegválasztással alkalmazott módszer előnye az egyszerűvel (sima Richardson) szemben. Ott $n_0(\xi) = \mathcal{O}(1/\xi)$, míg itt $n_0(\xi) = \mathcal{O}(1/\sqrt{\xi})$.

3.3. A Csebisev paramétermegválasztással kapott séma numerikus stabilitása

A Csebisev-féle paramétermegválasztás során a t_k alappontoktól függő τ_k paraméterek elvileg bármilyen sorrendben választhatók. Ha az iterációs lépések száma n , akkor a lehetséges sorrendek száma $n!$. Ha a lineáris egyenletrendszert számítógéppel oldjuk meg, óhatatlanul fellépnek kerekítési hibák. A kerekítési hiba hatása szempontjából már nem mindegy, hogy milyen sorrendben választjuk a paramétereket! Egy lehetőség a

$$t_k = \cos \frac{(2k-1)\pi}{2n}, \quad k = 1, 2, \dots, n.$$

csökkenő sorrend, de éppígy választhatnánk a

$$t_k = -\cos \frac{(2k-1)\pi}{2n}, \quad k = 1, 2, \dots, n.$$

növekvő sorrendet is, azonban ezen legegyszerűbb megválasztások numerikusan instabilak. Felmerül a kérdés: a paraméterek milyen sorrendje esetén lesz a legkisebb a kerekítési hibák hatása? Azaz: keressük az 1 és $2n-1$ közötti páratlan számok olyan $\{\theta_n(1), \theta_n(2), \dots, \theta_n(n)\}$ sorozatát, amely mellett a

$$\tau_k = \frac{\tau_0}{1 + \rho_0 \sigma_k}, \quad \sigma_k = -\cos \left[\frac{\pi}{2n} \theta_n(k) \right], \quad k = 1, 2, \dots, n$$

séma numerikusan stabil lesz. Ennek garantálásához alkalmas a következő algoritmus, amely a θ_m -ről a θ_{2m} -re és a θ_{2m} -ről a θ_{2m+1} -re való átlépés formuláját adja meg.^[3] A θ_m -ről a θ_{2m} -re léphetünk a következő két formula egyikével:

$$\theta_{2m}(2i-1) = \theta_m(i), \quad \theta_{2m}(2i) = 4m - \theta_{2m}(2i-1), \quad i = 1, 2, \dots, m, \quad (3.15)$$

$$\theta_{2m}(2i-1) = \theta_m(i), \quad \theta_{2m}(2i) = 4m + 2 - \theta_{2m}(2i-1), \quad i = 1, 2, \dots, m. \quad (3.16)$$

θ_{2m} -ről a θ_{2m+1} -re pedig a

$$\theta_{2m+1}(i) = \theta_{2m}(i), \quad i = 1, 2, \dots, 2m, \quad \theta_{2m+1}(2m+1) = 2m+1 \quad (3.17)$$

formula szerint léphetünk. Ha az n szám a 2 egész kitevős hatványa, azaz $n = 2^p$ ($p > 0$), akkor $\theta_1 = \{1\}$ -ből a fenti képletekkel meghatározzuk a $\theta_2, \theta_4, \theta_8$ és θ_{16} sorozatokat:

$$\theta_2 = \{1, 3\}, \quad \theta_4 = \{1, 7, 3, 5\}, \quad \theta_8 = \{1, 15, 7, 9, 3, 13, 5, 11\},$$

$$\theta_{16} = \{1, 31, 15, 17, 7, 25, 9, 23, 3, 29, 13, 19, 5, 27, 11, 21\}.$$

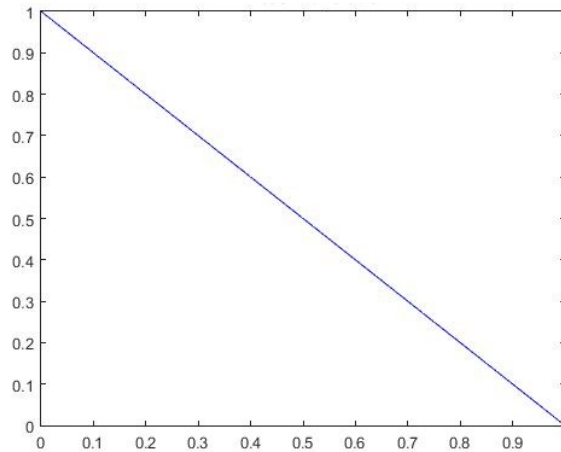
(Általánosítható tetszőleges $n > 0$ egész szám esetére.)

3.4. Számítási feladat

A következő megoldandó numerikus feladaton keresztül vizsgáljuk meg a Richardson-iteráció konvergenciáját:

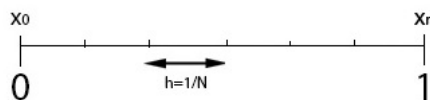
$$\begin{cases} u'' = 0 & 0 < x < 1 \\ u(0) = 1 \\ u(1) = 0 \end{cases}$$

melynek pontos megoldása $u(x) = 1 - x$.



3.1. ábra. Pontos megoldás

Diszkrétizáljuk a feladatot. Bontsuk fel a $[0, 1]$ intervallumot N részre, azonos h távolságokkal a következőképpen:



3.2. ábra. Intervallumfelosztás

Jelölje $u(x_i)$ közelítését v_i , és a második derivált közelítésére alkalmazzuk a szokásos sémát:

$$u''(x_i) \approx \frac{v_{i+1} - 2v_i - v_{i-1}}{h^2} = 0 \quad i = 1, \dots, N - 1$$

A peremfeltételekből tudjuk: $v_0 = 1, v_N = 0$, így az ismeretlen v_i elemekből álló v vektor legyen:

$$v := \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{N-1} \end{pmatrix}$$

Ennek elemeire felírható egy lineáris algebrai egyenletrendszer:

$$\begin{aligned} \frac{v_2 - 2v_1}{h^2} &= 0 \\ \frac{v_3 - 2v_2 - v_1}{h^2} &= 0 \\ &\vdots \\ \frac{0 - 2v_{N-1} - v_{N-2}}{h^2} &= 0 \end{aligned}$$

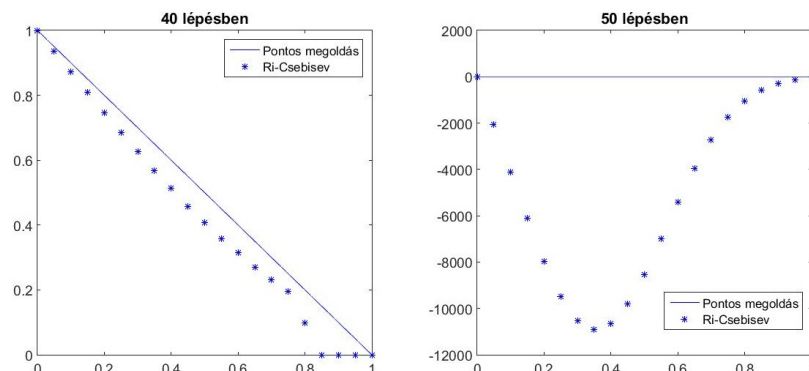
amely felírható $Av = f$ alakban a következő módon:

$$\underbrace{\frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \vdots & & & \ddots & & \\ 0 & \dots & \dots & & -1 & 2 \end{pmatrix}}_A \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{N-1} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{h^2} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_f$$

Vegyük észre, hogy A szimmetrikus, pozitív definit mátrix.

Oldjuk meg ezt a lineáris egyenletrendszert $N = 20$ -ra.

Először végezzünk 40 illetve 50 lépést Csebisev-féle paramétermegválasztással.

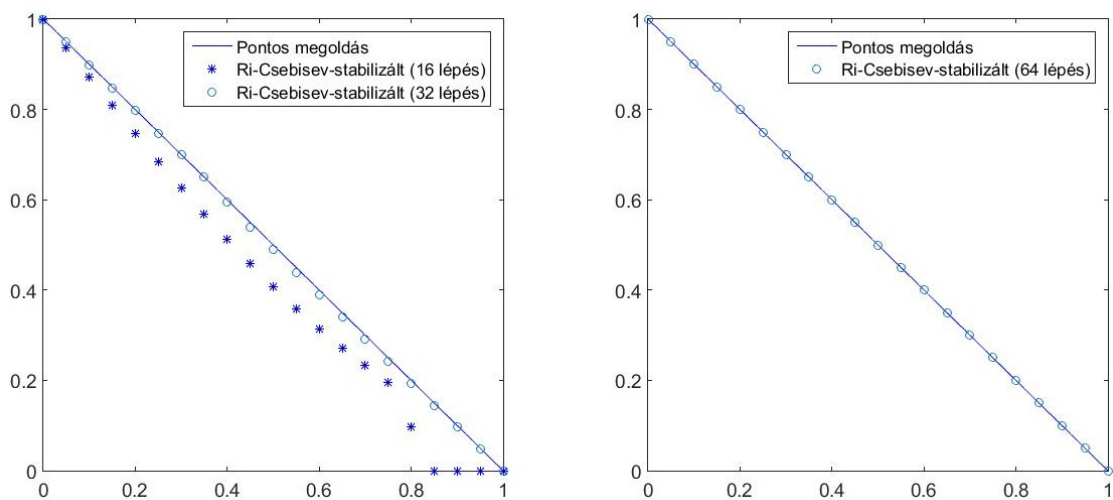


3.3. ábra. Richardson-iteráció 40 és 50 lépésben

Látható, hogy az iterációs lépések számát növelve, a kerekítési hibák hatására instabillá válik a módszer, és óriási eltérések alakulnak ki a pontos megoldástól.

Ennek kiküszöbölésére alkalmazzuk a stabilitáshoz szükséges paramétersorrendet.

Nézzük meg a θ_{16} , θ_{32} és θ_{64} eseteket:



3.4. ábra. Stabilizált Richardson-iteráció

Látható, hogy ez esetben az iterációs lépések számának növelésével a numerikus megoldás egyre közelebb kerül a pontoshoz.

Vizsgáljuk meg a hibákat először euklideszi normában ($\|x^{(n)} - x\|_2$), majd a numerikus megoldás számszerű értékeit a pontos megoldással összehasonlítva:

Richardson-Csebisev	Ri-Cseb-stabil	Ri-Cseb-stabil	Ri-Cseb-stabil
40	16	32	64
1.2984	0.3496	0.0289	1.7998e-04

3.1. táblázat. A pontos és az iterációs megoldás euklideszi távolsága

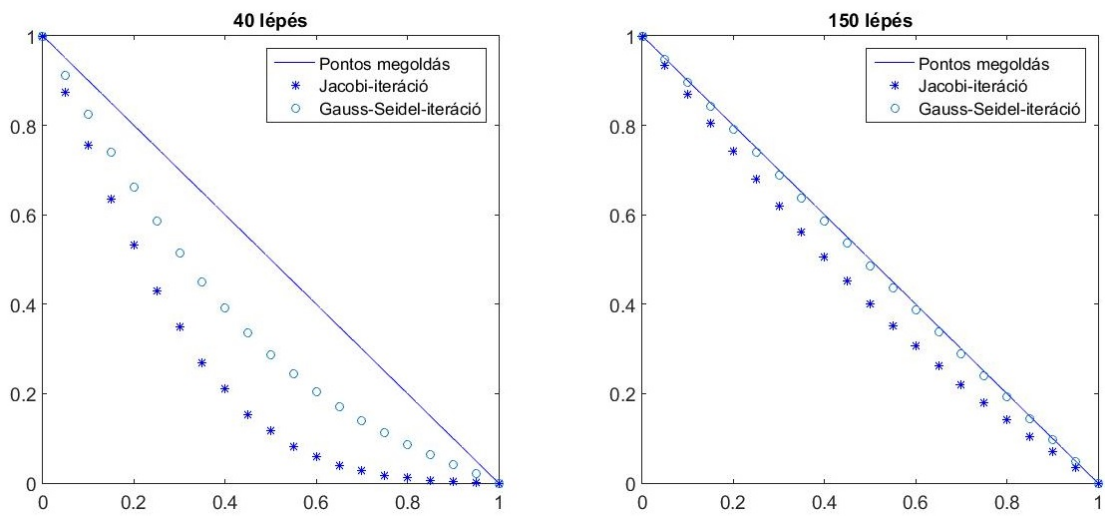
Pontos megoldás	Richardson-Csebisev 40	Ri-Cseb-stabil 16	Ri-Cseb-stabil 32	Ri-Cseb-stabil 64
1.0000	1.0000	1.0000	1.0000	1.0000
0.9500	1.1118	0.9357	0.9487	0.9500
0.9000	1.2112	0.8720	0.8977	0.9000
0.8500	1.2803	0.8082	0.8468	0.8500
0.8000	1.3080	0.7464	0.7969	0.8000
0.7500	1.2890	0.6846	0.7470	0.7500
0.7000	1.2243	0.6259	0.6989	0.7000
0.6500	1.1219	0.5671	0.6507	0.6500
0.6000	0.9944	0.5125	0.5954	0.5999
0.5500	0.8563	0.4578	0.5401	0.5499
0.5000	0.7209	0.4081	0.4897	0.4999
0.4500	0.5975	0.3584	0.4393	0.4499
0.4000	0.4908	0.3144	0.3896	0.3999
0.3500	0.4011	0.2703	0.3399	0.3499
0.3000	0.3260	0.2323	0.2908	0.3000
0.2500	0.2617	0.1942	0.2418	0.2500
0.2000	0.2044	0.0971	0.1932	0.2000
0.1500	0.1511	0	0.1446	0.1500
0.1000	0.1000	0	0.0964	0.1000
0.0500	0.0498	0	0.0481	0.0500
0	0	0	0	0

3.2. táblázat. Számszerű eltérések

Megoldottuk a feladatot Jacobi- és Gauss–Seidel-iterációval is:

Jacobi-iteráció 40	Jacobi-iteráció 150	Gauss–Seidel-iteráció 40	Gauss–Seidel-iteráció 150
1.2314	0.3133	0.6753	0.0441

3.3. táblázat. Euklideszi távolságok



3.5. ábra. Jacobi- és Gauss–Seidel-iteráció

Láthatjuk, hogy ezen iterációk esetén 150 lépésben sem érjük el azt a pontosságot, amelyet a Richardson-iterációnál 32 lépésben sikerült.

4. fejezet

Összefoglalás

Szakedolgozatomban megvizsgáltam néhány alapvető iterációs módszert lineáris algebrai egyenletrendszerek megoldására. Egy iterációs módszerrel szemben a leg-alapvetőbb elvárás az, hogy a kapott vektorsorozat tetszőleges kezdeti vektor esetén konvergáljon a lineáris egyenletrendszer megoldásához. A dolgozatban bemutattam a két legelterjedtebb iterációs módszert, a Jacobi- és a Gauss–Seidel-iterációt, és ismertettem néhány, konvergenciára vonatkozó tételt és azok bizonyítását. Láttuk, hogy a Jacobi-iteráció szimmetrikus és szigorúan diagonálisan domináns mátrixú egyenletrendszerek esetén konvergens, illetve hogy a Gauss–Seidel-iteráció konvergenciájához elégséges, ha az együtthatómátrix szimmetrikus és pozitív definit.

Nagyobb hangsúlyt kapott a Richardson-iteráció, amelyet egy kevésbé ismert eljárással, a Csebisev-féle paraméterválasztással optimalizáltunk. A kapott módszer a paraméterek csökkenő vagy növekvő megválasztása esetén numerikusan instabil, ami azt jelenti, hogy a kerekítési hibák hatása felhalmozódik, és elrontja a megoldást. Az orosz irodalomban fellelhető a probléma megoldása, amelynek lényege, hogy az egymást követő iterációs lépésekben egy speciális algoritmus szerint kell megválasztani a paramétereket. E módszer hatékonyságát megfigyelhetjük a dolgozat végén található számolási feladaton keresztül is, ahol látható, hogy ötödannyi iterációs lépés alatt elérjük azt a pontosságot, mint amelyet a Jacobi- és a Gauss–Seidel-iterációval sikerült.

Függelék

A felhasznált programok Matlab kódja

Jacobi- és Gauss–Seidel-iteráció

```
1 clear all
2 close all
3 N=20;
4 h = 1/N;
5 A = -1/h^2*(-2*eye(N-1) + diag(ones(1,N-2),1) +
6   diag(ones(1,N-2),-1));
7 f = zeros(N-1,1);
8 f(1) = 1/h^2;
9 x = A\f;
10 x = [1; x; 0];
11 %Jacobi 40
12 D=diag(diag(A));
13 T=D-A;
14 b=D*f;
15 y(:,1) = zeros(N-1,1);
16 for k=1:40
17     y(:,k+1) = D\(T*y(:,k)+b);
18 end
19 y(:,end);
20 yteljes = [1; y(:,end); 0];
21 %Jacobi 150
22 w(:,1) = zeros(N-1,1);
23 for k=1:150
24     w(:,k+1) = D\(T*w(:,k)+b);
```

```

25 end
26 w(:,end);
27 wteljes = [1; w(:,end); 0];
28 %gauss 40
29 S=tril(A);
30 U=S-A;
31 z(:,1) = zeros(N-1,1);
32 for k=1:40
33     z(:,k+1) = S\ (U*z(:,k)+f);
34 end
35 z(:,end);
36 zteljes = [1; z(:,end); 0];
37 %Gauss 150
38 p(:,1) = zeros(N-1,1);
39 for k=1:150
40     p(:,k+1) = S\ (U*p(:,k)+f);
41 end
42 p(:,end);
43 pteljes = [1; p(:,end); 0];
44 t_abr = 0:0.01:1;
45 u = 1-t_abr;
46 x_abr = 0:h:1;
47 subplot(1,2,1)
48 plot(t_abr,u,'b',x_abr,yteljes,'b*',x_abr,zteljes,'o' )
49 legend('Pontos megoldas','Jacobi-iteracio',
50 'Gauss-Seidel-iteracio')
51 title('\fontsize{14}40 lepes')
52 subplot(1,2,2)
53 plot(t_abr,u,'b',x_abr,wteljes,'b*',x_abr,pteljes,'o' )
54 legend('Pontos megoldas','Jacobi-iteracio',
55 'Gauss-Seidel-iteracio')
56 title('\fontsize{14}150 lepes')

```


Richardson-iteráció

```
1 clear all
2 close all
3 N = 20;
4 h = 1/N;
5 A = -1/h^2* (-2*eye(N-1)+diag(ones(1,N-2),1)+
6 diag(ones(1,N-2),-1));
7 b = zeros(N-1,1);
8 b(1) = 1/h^2;
9 x = A\b;
10 x = [1; x; 0];
11 % Az egyenletrendszer megoldása Ri-iterációval
12 gamma1 = 4/h^2*(sin(pi*h/2))^2;
13 gamma2 = 4/h^2*(cos(pi*h/2))^2;
14 ksi = (tan(pi*h/2))^2;
15 ro_0 = (1-ksi)/(1+ksi);
16 tau_0 = 2/(gamma1 + gamma2);
17 %16 lépés
18 y(:,1) = zeros(N-1,1);
19 n2 = 16;
20 theta1 = [1 31 15 17 7 25 9 23 3 29 13 19 5 27 11 21];
21 for k = 1:16
22     sigma(k)=-cos(theta1(k)/2/n2*pi);
23     tau(k) = tau_0/(1+ro_0*sigma(k));
24     y(:,k+1) = y(:,k)+ tau(k)*(b-A*y(:,k));
25 end
26 y(:,end);
27 yteljes = [1; y(:,end); 0];
28 euklidesz1=norm(abs(yteljes-x),2)
29 %32 lépés
30 z(:,1) = zeros(N-1,1);
31 n3 = 32;
32 theta2 = [1 63 31 33 15 49 17 47 7 57 25 39 9 55
33 23 41 3 61 29 35 13 51 19 45 5 59 27 37 11 53 21 43];
34 for k = 1:32
```

```

35     sigma(k)=-cos(theta2(k)/2/n3*pi);
36     tau(k) = tau_0/(1+ro_0*sigma(k));
37     z(:,k+1) = z(:,k)+ tau(k)*(b-A*z(:,k));
38 end
39 z(:,end);
40 zteljes = [1; z(:,end); 0];
41 euklidesz2=norm(abs(zteljes-x),2)
42 %64 lepes
43 w(:,1) = zeros(N-1,1);
44 n4 = 64;
45 theta3 = [1 127 63 65 31 97 33 95 15 113 49 79 17 111
46 47 81 7 121 57 71 25 103 39 89 9 119 55 73 23 105 41
47 87 3 125 61 67 29 99 35 93 13 115 51 77 19 109 45 83
48 5 123 59 69 27 101 37 91 11 117 53 75 21 107 43 85];
49 for k = 1:64
50     sigma(k)=-cos(theta3(k)/2/n4*pi);
51     tau(k) = tau_0/(1+ro_0*sigma(k));
52     w(:,k+1) = w(:,k)+ tau(k)*(b-A*w(:,k));
53 end
54 w(:,end);
55 wteljes = [1; w(:,end); 0];
56 euklidesz3=norm(abs(wteljes-x),2)
57 t_abr = 0:0.01:1;
58 u = 1-t_abr;
59 x_abr = 0:h:1;
60 subplot(1,2,1)
61 plot(t_abr,u,'b',x_abr,yteljes,'b*',x_abr,zteljes,'o')
62 legend('Pontos megoldas','Ri-Csebisev-stabilizalt
63 (16 lepes)','Ri-Csebisev-stabilizalt (32 lepes)')
64 subplot(1,2,2)
65 plot(t_abr,u,'b',x_abr,wteljes,'o')
66 legend('Pontos megoldas',
67 'Ri-Csebisev-stabilizalt (64 lepes)')

```

Irodalomjegyzék

- [1] Faragó István, Horváth Róbert: *Numerikus módszerek*, Typotex kiadó (2013)
- [2] Stoyan Gispert, Takó Galina: *Numerikus módszerek I.*, Typotex (1993)
- [3] A.A Szamarszkij: *Véges különbséges sémák elmélete (oroszul)*, Nauka (1977)
- [4] A.A Szamarszkij, A.V. Gulin: *Numerikus módszerek (oroszul)*, Nauka (1989)
- [5] Freud Róbert: *Lineáris algebra*, ELTE Eötvös Kiadó (2007)