

Numerikus módszerek 1 jegyzet

Toffalini Leonardo

2024. március 24.

ELTE

Tartalomjegyzék

1. Bevezetés	1
2. Numerikus modellezés	3
2.1. Numerikus modellezés lépései	3
2.2. Hibaforrások	4
2.3. Hibafogalmak	4
2.4. Az alpműveletek hibája	5
2.5. Korrekt kitűzésű feladatok	5
3. Normált terek	7
3.1. Normált tér	7
3.2. Fontos fogalmak normált terekben	8
3.3. Mátrixnormák	9
3.4. Kondíciós szám	10
4. Lineáris algebrai egyenletrendszerek megoldása	13
4.1. Gauss-elimináció	13
4.2. Főelem kiválasztás (pivoting)	16
4.3. Klasszikus iterációs módszerek	17
4.4. Richardson-iteráció	18
4.5. Jacobi-iteráció	18
4.6. Gauss-Seidel-iteráció	18
4.7. Stacionárius-iteráció	19
4.8. Stacionárius iteráció konvergenciája	20
4.9. SOR-módszer konvergenciája	21
5. Gradiens alapú módszerek	23
5.1. Gradiens módszer	25
5.2. Konjugált gradiens-módszer	25
6. Általános algebrai egyenletek megoldása	27
6.1. Gyökök stabilitása	27
6.2. Konvergencia sebesség	28
Irodalomjegyzék	31

1. fejezet

Bevezetés

Az alábbi egy jegyzet Havasi Ágnesnek a 2023/2024-es tavaszi félévében tartott Numerikus Módszerek 1 előadásáról. A jegyzet nem teljeskörű dokumentációja az előadáson elhangzottaknak és nem vállal felelősséget az esetleges hibákért.

2. fejezet

Numerikus modellezés

Ebben a fejezetben tárgyalni fogjuk az alapvető lépéseit és fogalmait a numerikus modellezésnek és a numerikus módszereknek.

2.1. Numerikus modellezés lépései

1. Valódi probléma

Halpopuláció időbeli fejlődése.

2. Tudományos modell

Vannak zsákmányhalak és ragadozó halak. A zsákmányhalak és a ragadozóhalak populációját befolyásolni, többek között:

- természetes szaporulat
- ragadozók esznek zsákmány halakat
- természetes pusztulás

3. Matematikai modell

- jelölje $x(t)$ a zsákmányhalak t időbeli össztömegét
- jelölje $y(t)$ a ragadozóhalak t időbeli össztömegét

Ezekkel a jelölésekkel felírhatjuk a változók közti összefüggést egy differenciálegyenlettel:

$$\begin{aligned}x' &= ax - bxy \\y' &= -cy + dxy \\x(0) &= x_0 \\y(0) &= y_0\end{aligned}$$

4. Numerikus modell

Közelítő módszert alkalmazunk az előző, úgy nevezett *Lotka Volterra* egyenletre.

5. Számítógépes modell

Lekódoljuk és futtatjuk a numerikus modellnek a programját.

2.2. Hibaforrások

1. Modellhiba

A tudományos és a matematikai modellben éltünk egyszerűsítésekkel, melyek nem pontosan ábrázolták a valóságot.

2. Képlethiba

A matematikai és a numerikus modellben egy egyszerűbb kifejezéssel helyettesítettünk egy bonyolultabb kifejezést. Tipikusan egy Taylor-sorral helyettesítettünk egy nehezen leírható függvényt.

Például:

$$\exp(2) = \sum_{k=0}^{\infty} \frac{2^k}{k!} \approx \sum_{k=0}^N \frac{2^k}{k!}$$

A képlethibának az egyik fajtája a *diszkretizációs hiba*, melynek tipikus esetei:

- folytonos függvényt helyettesítünk rácspont függvénnyel
- deriváltat helyettesítünk differenciáhányadossal
- integrált helyettesítünk egy véges összeggel
- végtelent helyettesítünk egy tetszőlegesen nagy termésszel számmal

3. A bemenő adatok hibája

Gyakran nem pontosan kapjuk meg az adatokat és így számolnunk kell ezzel a hibaforrással. Ez gyakran mérési hibából következik.

4. Számábrázolási hiba

A való életben nem szimbólikusan számolunk valós számokkal, hanem egy számítógépre hagyjuk a számításokat. A számítógépünk viszont csak egy véges részhalmozát képes ábrázolni a valósszámoknak, így ha egy valós számot adunk meg egy számítógépnek, akkor az a hozzá legközelebb álló ábrázolható számot fogja helyette használni.

2.3. Hibafogalmak

Szeretnénk számszerűen megfogalmazni, hogy mennyire pontosan számoltunk és, hogy mennyire tér el a számított érték a valódi értéktől. A továbbiakban jelölje $a \in \mathbb{R}$ a pontos értéket és $\tilde{a} \in \mathbb{R}$ a számított értéket.

Definíció 2.3.1 Az \tilde{a} abszolút hibájának a $\Delta a := a - \tilde{a}$ számot értjük.

Definíció 2.3.2 A $\Delta a \in \mathbb{R}_0^+$ számot az \tilde{a} egy abszolút hibakorlátjának nevezzük, ha $|\Delta a| \leq \Delta a$

Jelölésben $a = \tilde{a} \pm \Delta a$

Definíció 2.3.3 \tilde{a} relatív hibájának nevezzük a következőt: $\delta a = \frac{\Delta a}{|\tilde{a}|}$

Definíció 2.3.4 \tilde{a} relatív hibakorlátjának nevezzük a következőt: $\delta a \in \mathbb{R}_0^+$ szám melyre $|\delta a| \leq \delta a$

2.4. Az alpműveletek hibája

A következőkben keressük, hogy mennyire hibázunk, amikor számábrázolási hibából következően nem a pontos értékekkel végezzük el az alpműveleteket.

Tegyük fel, hogy $x, y \in \mathbb{R}$ helyett a hibás $\tilde{x}, \tilde{y} \in \mathbb{R}$ számokkal végezzük el az alpműveleteket.

1. Összeadás

$$\begin{aligned} |(x + y) - (\tilde{x} + \tilde{y})| &= |x - \tilde{x} + y - \tilde{y}| \\ &\leq |x - \tilde{x}| + |y - \tilde{y}| \\ &\leq \Delta_x + \Delta_y \end{aligned}$$

2. Kivonás

$$\begin{aligned} |(x - y) - (\tilde{x} - \tilde{y})| &= |x - \tilde{x} + \tilde{y} - y| \\ &\leq |x - \tilde{x}| + |\tilde{y} - y| = |x - \tilde{x}| + |y - \tilde{y}| \\ &\leq \Delta_x + \Delta_y \end{aligned}$$

3. Szorzás

$$\begin{aligned} |xy - \tilde{x}\tilde{y}| &= |xy + x\tilde{y} - x\tilde{y} - \tilde{x}\tilde{y}| \\ &= |x(y - \tilde{y}) + \tilde{y}(x - \tilde{x})| \\ &\approx |\tilde{x}(y - \tilde{y}) + \tilde{y}(x - \tilde{x})| \\ &\leq |\tilde{x}|\Delta_y + |\tilde{y}|\Delta_x := \Delta_{xy} \end{aligned}$$

4. Hányados

$$\left| \frac{x}{y} - \frac{\tilde{x}}{\tilde{y}} \right| \leq \frac{\Delta_{xy}}{\tilde{y}^2}$$

2.5. Korrekt kitűzésű feladatok

Mielőtt nekiállnánk egy feladatot megoldani érdemes elgondolkoznunk azon, hogy egyáltalán van-e értelme megoldani, vagy korrekten van-e kitűzve a feladat.

Ha kapunk egy feladatot, akkor a következők korrekt elvárások:

- Létezzen megoldás (*egzisztencia*)
- Csak egy megoldás létezzen (*unicitás*)
- A feladat pontos megoldása folytonosan függjön a bemenő adatoktól.

Például az $Ax = b$ nem ilyen, mert ha egy kicsit megváltoztatjuk az A együttható mátrix elemét, akkor a megoldás nagy mértékben változhat.

3. fejezet

Normált terek

Eddig csak valós számokra alkalmaztuk az abszolútérték függvényt, amikor hibafogalmakról beszéltünk. Megeshet, hogy a keresett érték nem egy valós szám, hanem például egy mátrix vagy egy függvény vagy egy tetszőleges operátor. Ilyenkor nem tudjuk alkalmazni a szokásos abszolút érték függvényt, mert nem tudjuk, hogy mit jelent egy mátrix abszolútértéke.

Ennek érdekében bevezetünk egy olyan teret, melynek elemeire lehet a kiterjesztett abszolútérték függvényt használni.

3.1. Normált tér

Ahhoz, hogy kiterjesszük az abszolútérték függvényt tekintsük a tulajdonságait, hogy mit kéne örökölnie egy tágabb hossz fogalomnak:

1. $|x| \geq 0 \quad \forall x \in \mathbb{R}$ és $|x| = 0 \iff x = 0$
2. $|\lambda x| = |\lambda| \cdot |x|$ (abszolút homogenitás)
3. $|x + y| \leq |x| + |y| \quad \forall x, y \in \mathbb{R}$ (háromszög egyenlőtlenség)

Definíció 3.1.1 Legyen X tetszőleges vektortér, és $\|\cdot\| : X \rightarrow \mathbb{R}$ egy függvény a következő tulajdonságokkal:

1. $\|x\| \geq 0 \quad \forall x \in X$ és $\|x\| = 0 \iff x = 0_X$ (X nullvektora)
2. $\|\lambda x\| = |\lambda| \cdot \|x\| \quad \forall x \in X, \forall \lambda \in \mathbb{R}$
3. $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in X$

Ekkor ezen $\|\cdot\|$ függvényt normának nevezzük és a normált tér (N.T.) a következő rendezett pár: $(X, \|\cdot\|)$.

Definíció 3.1.2 Ha $(X, \|\cdot\|)$ Normált tér, akkor $x, y \in X$ elemek távolságán az $\|x - y\|$ számot értjük.

Megjegyzés 1 Ezt a $\|x - y\|$ távolságot szokás a norma által indukált metrikának nevezni.

Példa 1 Példák normákra és normált terekre:

1. $X = \mathbb{R}$ és $\|\cdot\| = |\cdot|$

2. $X = \mathbb{R}^n$ a következő normákkal:

(i) $\|x\|_1 := \sum |x_j|$

(ii) $\|x\|_2 := \sqrt{\sum |x_j|^2}$

(iii) $\|x\|_\infty := \max\{|x_j|\}$

(iv) $\|x\|_p := (\sum |x_j|^p)^{1/p}$

Ha $p \rightarrow \infty$ akkor $\|x\|_p \rightarrow \|x\|_\infty \quad \forall x \in X$

3. $X = C[a, b]$, azaz az $[a, b]$ intervallumon értelmezett folytonos függvények, a következő normákkal:

(i) $\|f\|_\infty := \max_{x \in [a, b]} |f(x)|$

(ii) $\|f\|_f := \int_a^b |f(x)| dx$

3.2. Fontos fogalmak normált terekben

Most hogy már kiterjesztettük a hossz fogalmát normált terekre, így képesek vagyunk az előző fejezetekben bevezetett fogalmakat analóg módon megfogalmazni a tér normájával.

1. Hibafogalmak

Legyen $(X, \|\cdot\|)$ egy tetszőleges Normált tér és $a, \tilde{a} \in X$. Ekkor

- \tilde{a} abszolút hibája: $a - \tilde{a} \in X$
- \tilde{a} abszolút hibakorlátja: $\Delta_a \in \mathbb{R}$ szám, melyre $\|a - \tilde{a}\| \leq \Delta_a$
- \tilde{a} relatív hibája: $\frac{a - \tilde{a}}{\|\tilde{a}\|} \in X$
- \tilde{a} relatív hibakorlátja: $\frac{\|a - \tilde{a}\|}{\|\tilde{a}\|} \leq \delta_a \in \mathbb{R}$

2. Konvergencia

Definíció 3.2.1 Azt mondjuk, hogy az $(x_n) \subset X$ sorozat konvergens, ha $\exists x \in X$, melyre $\|x_n - x\| \rightarrow 0$ ha $n \rightarrow \infty$.

3.3. Mátrixnormák

Tudjuk, hogy az $\mathbb{R}^{n \times n}$ -beli mátrixok a rajta értelmezett $+$ (összeadás) és λ -val való szorzás műveletekkel vektorteret alkotnak.

Kérdés 1 *Hogyan definiálható ezen a vektortéren norma?*

Definíció 3.3.1 *Legyen $\|\cdot\|_{\mathbb{R}^n}$ egy \mathbb{R}^n -beli vektornorma. Ekkor az $A \in \mathbb{R}^{n \times n}$ mátrix ezen vektornorma által indukált mátrixnormáján a következő számot értjük:*

$$\|A\| := \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}}$$

Magyarázó jelentések a definícióhoz:

- $\|Ax\|_{\mathbb{R}^n}$ - az Ax vektor "hossza"
- $\frac{\|Ax\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}}$ - hányszorosára nyújtotta az A mátrix az x vektort
- $\sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}}$ - lehetséges legnagyobb megnyújtásnak az értéke

Példa 2 *Tekintsük pár mátrixnak pár mátrixnormáját.*

1.

$$\|I\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ix\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}} = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|x\|}{\|x\|} = \sup 1 = 1$$

Tehát bármelyik \mathbb{R}^n -beli norma által indukált mátrixnormában az identitás mátrix normája 1, azaz $\|I\| = 1$.

2. A sup-norma kiszámítása a tanult vektornormák esetén: Ha $\|\cdot\|_{\mathbb{R}^n} = \|\cdot\|_1$, akkor:

$$\|A\| = \|A\|_1 = \max_{j \in \{1, \dots, n\}} \sum_{i=1}^n |a_{ij}|$$

max oszlopösszeg!

Például:

$$\begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} \Rightarrow \|A\|_1 = \max\{|-2| + |0|, |1| + |3|\} = 3$$

3. Ha $\|\cdot\| = \|\cdot\|_2$, akkor:

$$\|A\| = \|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

ahol λ_{\max} a legnagyobb sajátértéket jelöli. Ezt a normát szokás *spektrálnormának* nevezni, mert a sajátértékek halmazát *spektrál*-nak nevezik.

4. Ha $\|\cdot\| = \|\cdot\|_\infty$, akkor:

$$\|A\| = \|A\|_\infty = \max_{i \in \{1, \dots, n\}} \sum_{j=1}^n |a_{ij}|$$

max sorösszeg!

Például:

$$\begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} \Rightarrow \|A\|_\infty = \max\{|-2| + |1|, |0| + |3|\} = 3$$

Állítás 3.3.1 Az indukált mátrix normákra igazak a következők:

1. $\|Ax\| \leq \|A\| \cdot \|x\| \quad \forall A \in \mathbb{R}^{n \times n}, \forall x \in \mathbb{R}^n.$
2. $\|I\| = 1$ (láttuk).
3. $\|A \cdot B\| \leq \|A\| \cdot \|B\| \quad \forall A, B \in \mathbb{R}^{n \times n}$ (szub multiplikatívitas).

Megjegyzés 2 Vannak egyéb, nem indukált, mátrix normák. például:

1. $\|A\|' = \max_{i,j} |a_{ij}|$ (maximális elem)
2. $\|A\|'' = \sum_{i,j=1}^n |a_{ij}|$ (elemösszeg)
3. $\|A\|_F = \sqrt{\sum_{i,j=1}^n a_{ij}^2}$ (Frobenius norma)

Ezekre a nem indukált mátrix normákra nem feltétlenül teljesülnek a 3.3.1-beli tulajdonságok.

3.4. Kondíciósám

Az előbb meggondoltuk, hogy egy lineáris egyenletrendszernek, $Ax = b$ -nek, az A együtthatómátrixának egy elemét kicsit perturbálva a megoldás drasztikusan változhat. Célunk, hogy megfogalmazzuk, hogy mennyire változhat a megoldás kis perturbációra.

A továbbiakban a következő egyenletrendszerrel fogunk foglalkozni.

$$Ax = b \tag{3.1}$$

Ahol $A \in \mathbb{R}^{n \times n}$, $\det A \neq 0$, $b \in \mathbb{R}^n$

Tegyük fel, hogy b helyett a perturbált \tilde{b} van a jobb oldalon:

$$A\tilde{x} = \tilde{b}$$

Jelölje:

$$\begin{aligned}\Delta x &= x - \tilde{x} \implies \tilde{x} = x - \Delta x \\ \Delta b &= b - \tilde{b} \implies \tilde{b} = b - \Delta b\end{aligned}$$

Ekkor:

$$\begin{aligned}A\tilde{x} &= \tilde{b} \\ A(x - \Delta x) &= b - \Delta b \\ Ax - A\Delta x &= b - \Delta b \\ A\Delta x &= \Delta b \\ \Delta x &= A^{-1}\Delta b\end{aligned}$$

Nézzük $\|\Delta x\|$ -át valamelyik \mathbb{R}^n -beli normában:

$$\|\Delta x\| = \|A^{-1}\Delta b\| \leq \|A^{-1}\| \cdot \|\Delta b\|$$

Most alkalmazzuk a 3.1-es egyenletrendszerre a normát.

$$b = Ax$$

$$\begin{aligned}\|b\| &= \|Ax\| \leq \|A\| \cdot \|x\| \\ \frac{1}{\|x\|} &\leq \|A\| \cdot \frac{1}{\|b\|} \\ \implies \frac{\|\Delta x\|}{\|x\|} &\leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\Delta b\|}{\|b\|}\end{aligned}$$

Tehát azt kaptuk, hogy minél nagyobb $\|A^{-1}\| \cdot \|A\|$ annál pontatlanabb a becslés.

Definíció 3.4.1 Az $\|A^{-1}\| \cdot \|A\|$ számot az A mátrix kondíció számának nevezzük és $\text{cond}(A)$ -val jelöljük.

Definíció 3.4.2 Azt mondjuk, hogy a 3.1-es egyenletrendszer rosszul kondicionált, ha $\text{cond}(A) \gg 1$.

Példa 3 Nezzük meg a már említett példának a kondíció számát.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1.01 \end{bmatrix}$$

Alkalmazzuk az $\|\cdot\|_1$ által indukált mátrix normát.

$$\|A\|_1 = \max\{1 + 1, 1 + 1.01\} = 2.01$$

$$A^{-1} = \begin{bmatrix} 101 & -100 \\ -100 & 100 \end{bmatrix} \implies \|A^{-1}\|_1 = \max\{101 + 100, 100 + 100\} = 201$$

$$\text{cond}(A) = 201 \cdot 2.01 = 404.01 \gg 1$$

Tehát valóban rosszul kondicionált volt az egyenlet rendszer.

4. fejezet

Lineáris algebrai egyenletrendszerek megoldása

Lineáris algebrai egyenletrendszerek megoldásaira két féle megoldási módszert fogunk tanulni. Direkt megoldókat és iterációs módszereket. Az előzőhöz tartozik például a Cramer-szabály vagy a Gauss-elimináció. Az iterációs módszereknek viszont a lényege az, hogy egy vektorsorozatot generálnak, melyek tartanak a pontos megoldáshoz.

4.1. Gauss-elimináció

Megoldandó egyenletrendszer: $Ax = b$, $A \in \mathbb{R}^{m \times m}$, $\det A \neq 0$, $b \in \mathbb{R}^m$

Lineáris algebrából tudjuk, hogy ezek a feltételek mellett egyértelműen létezik megoldás, tehát korrekt kitűzésű a feladat és van értelme nekiállni megoldani.

A lineáris egyenletrendszer teljes anyakönyvezet nevén a következő:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 & (1) \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mm}x_m &= b_m & (m) \end{aligned}$$

I. alakban: Átalakítjuk az egyenletrendszert normált felső háromszög mátrixúvá. Tehát a főátlóban legyenek egyesek és a főátló alatt csupa nulla.

1. lépés: Tegyük fel, hogy $a_{11} \neq 0$ ekkor

$$x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 + \frac{a_{1m}}{a_{11}}x_m = \frac{b_1}{a_{11}} = y_1 \quad (1) \quad (4.1)$$

2. lépés: 4.1 segítségével a másodiktól az m -edik egyenletekből elimináljuk x_1 -et, kivonva belőlük a 4.1-nek az a_{i1} -szerezését.

$$\begin{aligned} x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 + \frac{a_{1m}}{a_{11}}x_m &= \frac{b_1}{a_{11}} = y_1 \\ a_{22}^{(1)}x_2 + \dots &= y_2 \\ &\vdots \\ a_{m2}^{(1)}x_2 + \dots + a_{mm}^{(1)}x_m &= b_m \end{aligned}$$

3. lépés: Nem írom tovább mert mindenki tud Gauss-eliminálni...

Kérdés 2 Mikor hajtható végre a Gauss-elimináció?

I. szakaszban $Ax = b \implies Ux = y$

Kérdés 3 Mi a kapcsolat y és b között?

$$\begin{aligned} b_1 &= a_{11}y_1 \\ b_2 &= a_{21}y_1 + a_{22}^{(1)}y_2 \\ &\vdots \\ b_m &= l_{j1}y_1 + l_{j2}y_2 + \dots + l_{mm}y_m \end{aligned}$$

Ahol $l_{jj} = a_{jj}^{(j-1)}$

Kompaktabb mátrix formába átírva:

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22}^{(1)} & \dots & 0 \\ & & \dots & a_{mm}^{(m-1)} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Ha a Gauss-elimináció elvégezhető akkor a fenti mátrix invertálható, azaz a főátlóban nincs 0, tehát $\exists L^{-1}$, ahol L a fenti alsó háromszög mátrix.

$$\text{Tehát } Ly = b \implies y = L^{-1}b \implies Ux = L^{-1}b \implies LUx = b$$

Ebből adódik egy új módszer (LU felbontás):

1. Felírjuk az A -t $A = LU$ alakban, ahol L invertálható alsó háromszög mátrix és U olyan felső háromszög mátrix melynek a főátlójában csak egyesek vannak.
2. Megoldjuk az $Ly = b$ egyenletrendszert, ebből kapunk egy értéket y -ra.
3. Megoldjuk az $Ux = y$ egyenletrendszert, amiből megkapjuk x -et.

Belátható, hogy az LU felbontás első és második lépse ekvivalens a Gauss-elimináció első szakaszával és harmadik lépés ekvivalens a Gauss-elimináció második szakaszával. Tehát ez a módszer a Gauss-elimináció módosított algoritmusa.

Ahhoz, hogy megválaszoljuk, hogy mikor végezhető el a Gauss-elimináció elég megválaszolnunk, hogy mikor létezik LU felbontás.

A következőképpen jelöljük a balfelső sarokdeterminánsokat (*főminorokat*):

$$\Delta_1 := a_{11}, \quad \Delta_2 := \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \dots, \Delta_m := \det A$$

Állítás 4.1.1 *Ha $\Delta_j \neq 0, \forall j \in \{1, \dots, m\}$, akkor létezik LU felbontása A -nak, és az egyértelmű.*

Bizonyítás: Csak az $A \in \mathbb{R}^{2 \times 2}$ esetre mutatjuk meg, magasabb dimenzióra teljes indukcióval lehet belátni az állítást.

Először bizonyítsuk a létezést.

$$A = L \cdot U = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \cdot \begin{bmatrix} 1 & u_{12} \\ 0 & 1 \end{bmatrix}$$

felbontás létezik \iff létezik $l_{11}, l_{21}, l_{22}, u_{12}$ ismeretlenekre nézve megoldása a következő egyenlet rendszernek.

$$\begin{aligned} l_{11} &= a_{11} \\ l_{11}u_{12} &= a_{12} \\ l_{21} &= a_{21} \\ l_{21}u_{12} + l_{22} &= a_{22} \end{aligned}$$

és a következő L mátrixnak létezen inverze

$$L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix}$$

azaz $l_{11} \neq 0, l_{22} \neq 0$.

Ha $a_{11} \neq 0$, akkor látható, hogy ennek az egyenletrendszernek egyértelműen létezik megoldása és az a következő:

$$l_{11} = a_{11}, \quad u_{12} = \frac{a_{12}}{a_{11}}, \quad l_{21} = a_{21}, \quad l_{22} = a_{22} - a_{21} \frac{a_{12}}{a_{11}}$$

Továbbá, $l_{11} \neq 0$, mert $a_{11} \neq 0$ és $l_{22} \neq 0$, mert $l_{22} = \frac{\det A}{a_{11}} \implies \exists L^{-1}$

Most lássuk be, hogy egyértelműen létezik.

Tegyük fel, hogy $A = L_1 U_1 = L_2 U_2$

$$\begin{aligned} L_2^{-1} L_1 U_1 &= U_2 \\ L_2^{-1} L_1 &= U_2 U_1^{-1} \end{aligned}$$

Mivel az alsóháromszög mátrixok és a felső háromszög mátrixok is egy-egy csoportot alkotnak, ezért a fenti csak akkor igaz, ha $L_2^{-1}L_1$ és $U_2U_1^{-1}$ is diagonális. Továbbá $U_{1,2}$ -nek a főátlójában egyesek vannak, tehát $U_2U_1^{-1}$ -nek is a főátlójában egyesek vannak. Tehát mindkét oldalon az egység mátrix van.

$$\implies L_2^{-1}L_1 = I = U_2U_1^{-1} \implies L_1 = L_2, \quad U_1 = U_2$$

Megmutatható, hogy ha $\Delta_j \neq 0$ valamely j -re, akkor \exists LU-felbontása A -nak. 2×2 esetben jól látszik:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\begin{bmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \cdot \begin{bmatrix} 1 & u_{12} \\ 0 & 1 \end{bmatrix} \implies l_{11} = 0 \implies \text{ekkor } L \text{ nem invertálható}$$

Következmény 1 A Gauss-elimináció pontosan akkor hajtható végre, ha A összes bal felső sarokdeterminánsa nem 0.

Megjegyzés 3 Pár észrevétel a Gauss-elimináció és az LU felbontással kapcsolatban:

1. $A \Delta_j \neq 0, \quad \forall j = 1, \dots, m$ teljesül, ha A szimmetrikus pozitív definit mátrix (szpd).
2. $A \Delta_j \neq 0, \quad \forall j = 1, \dots, m$ teljesül, ha A szigorúan domináns főátlójú, tehát $\forall i = 1, \dots, n$ -re $2|a_{ii}| > \sum_{j=1}^m |a_{ij}|$.
3. Ha $\det A \neq 0$, akkor mindig $\exists P \in \mathbb{R}^{n \times m}$ permutáló mátrix, hogy PA -nak \exists LU felbontása.
4. Ha A szimmetrikus pozitív definit mátrix, akkor létezik egy másik felbontása is: $A = G \cdot G^T$, ahol G alsó háromszög mátrix, pozitív főátlóval. (Cholesky-felbontás)

4.2. Főelem kiválasztás (pivoting)

A Gauss-elimináció során a j -edik lépésben a j -edik sort elosztjuk a_{jj} -vel. Tehát minél kisebb a_{jj} , annál pontatlanabb az osztás. Ennek orvosolására valahogyan meg kéne oldanunk, hogy egy nagyobb elemmel osszunk, de a Gauss-elimináció lényegét tartjuk meg.

Részleges főelem kiválasztás: Sorcserével a főátlóba hozzuk az a_{jj} alatti legnagyobb abszolútértékű elemet.

Teljes főelem kiválasztás: Sorcserével és oszlopcserével az $A[j : n, j : n]$ jobb alsó részmátrix legnagyobb abszolútértékű elemet visszük a főátlóba. Itt figyelni kell arra, hogy oszlop cserénél az x elemeket is cseréljük. Tehát ha egy P mátrixszal permutáljuk az oszlopait A -nak, akkor mikor visszaolvassuk x megoldást, akkor P^{-1} -el meg kell szorozni előtte.

4.3. Klasszikus iterációs módszerek

Definíció 4.3.1 Azt mondjuk, hogy az $x^* \in \mathbb{R}^n$ az $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ függvény fixpontja, ha $f(x^*) = x^*$

Definíció 4.3.2 az $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ függvény kontrakció az $\|\cdot\|$ \mathbb{R}^n -beli normában, ha $\exists q \in [0, 1]$ melyre:

$$\|f(x) - f(y)\| \leq q \cdot \|x - y\| \quad \forall x, y \in D(f)$$

Tétel 4.3.1 (Banach fixpont tétel)

Ha $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ az egész \mathbb{R}^n -en értelmezett kontrakció (q -val), akkor:

1. f -nek egyértelműen létezik x^* fixpontja.
2. Tetszőleges $x^0 \in \mathbb{R}^n$ vektorból indítva $x^{n+1} = f(x^n)$ rekurzióval felépített $(x)_n$ sorozat konvergens, és $x_n \rightarrow x^*$.
3. $\|x^n - x^*\| \leq \frac{q^n}{1-q} \|x^1 - x^0\|$

Kérdés 4 Hogyan alkalmazhatjuk ezt a tételt lineáris algebrai egyenletrendszerek megoldására?

$$Ax = b, \quad A \in \mathbb{R}^{m \times m}, \quad \det A \neq 0, \quad b \in \mathbb{R}^m \quad (4.2)$$

Tegyük fel, hogy 4.2 átírható a következő alakra:

$$x = Qx + r, \quad Q \in \mathbb{R}^{m \times m}, \quad r \in \mathbb{R}^m \quad (4.3)$$

Ekkor az $f(x) := Qx + r$ ejöléssel a feladat megoldása az $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ függvény fixpontja. Ezt a fixpontot keressük iterációval.

Kérdés 5 Mikor lesz f kontrakció?

$$x, y \in \mathbb{R}^m, \quad f(x) - f(y) = Qx + r - Qy - r = Q(x - y)$$

$$\|f(x) - f(y)\|_{\mathbb{R}^m} = \|Q(x - y)\|_{\mathbb{R}^m} \leq \|Q\| \cdot \|x - y\|_{\mathbb{R}^m}$$

Tehát be kell látni, hogy $\|Q\| < 1$, akkor f kontrakció és $q = \|Q\|$.

Banach fixpont tételből következik, hogy a $x^{n+1} = Qx^n + r$ rekurzióval definiált vektor-sorozat konvergens (bármely \mathbb{R}^m -beli vektornormában), és $x_n \rightarrow x^*$, ahol x^* 4.2 megoldása.

Kérdés 6 Hogyan írhatjuk át 4.2-et olyan alakra amilyen 4.3?

Kérdés 7 Mikor fog teljesülni, hogy $\|Q\| < 1$ valamelyik indukált mátrixnorma szerint?

4.4. Richardson-iteráció

A *Richardson iteráció* vagy másnéven egyszerű iteráció, ahogyan a név is sugallja a legegyszerűbb módon alakítja át az $Ax = b$ egyenletet $f(x) = x$ alakúra. Pusztán annyi átalakítás történik, hogy nullára rendezzük az egyenletet és mindkét oldalhoz hozzáadunk x -et.

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

Tehát $f(x) = (I - A)x$ függvénynek fixpontjaként kapjuk az $Ax = b$ egyenlet megoldását a Banach-fixpont tétel alapján.

4.5. Jacobi-iteráció

A célunk még mindig, hogy egy függvénynek a fixpontjaként írjuk fel a lineáris egyenletrendszer megoldását. Ezt megtehetjük, ha a következőképpen felbontjuk az együttható mátrixot és egy kis algebrai manipulációt végzünk.

$$\begin{aligned} Ax &= b \\ A &= L + D + U \\ (L + D + U)x &= b \\ Dx &= -(L + U)x + b \\ x &= D^{-1}(b - (L + U)x) \\ &= -D^{-1}(L + U)x + D^{-1}b \\ Q_J &= -D^{-1}(L + U), \quad r_J = D^{-1}b \end{aligned}$$

Ekkor kapjuk, hogy a Jacobi fixpont iterációra rögzítsük $x^0 \in \mathbb{R}^m$ kezdőpontot és legyen az általános lépés:

$$x^{n+1} = -D^{-1}(L + U)x^n + D^{-1}b$$

Állítás 4.5.1

$$\|Q_J\|_\infty < 1 \iff A \text{ szigorúan domináns főátlójú}$$

Következmény 2 Ha A szigorúan domináns főátlójú, akkor a Jacobi-iteráció konvergens.

4.6. Gauss-Seidel-iteráció

A célunk még mindig, hogy egy függvénynek a fixpontjaként írjuk fel a lineáris egyenletrendszer megoldását. Ezt megtehetjük, ha a következőképpen felbontjuk az együttható mátrixot és egy kis algebrai manipulációt végzünk.

$$\begin{aligned}
Ax &= b \\
A &= L + D + U \\
(L + D + U)x &= b \\
(L + D)x &= -Ux + b \\
x &= -(L + D)^{-1}Ux + (L + D)^{-1}b \\
Q_{GS} &= -(L + D)^{-1}U, \quad r_{GS} = (L + D)^{-1}b
\end{aligned}$$

4.7. Stacionárius-iteráció

Észrevétel: A Jacobi-iteráció átírható a következő módon:

$$\begin{aligned}
x^{n+1} &= -D^{-1}(L + U)x^n + D^{-1}b \\
Dx^{n+1} &= -(L + U)x^n + b \\
Dx^{n+1} &= -(A - D)x^n + b \\
D(x^{n+1} - x^n) + Ax^n &= b
\end{aligned}$$

A fentit a Jacobi-iteráció kanonikus alakjának szokás nevezni.

Hasonló módon át tudjuk írni a Gauss-Seidel iterációt is:

$$(D + L)(x^{n+1} - x^n) + Ax^n = b \quad (\text{SI})$$

A fentit a Gauss-Seidel-iteráció kanonikus alakjának szokás nevezni.

Definíció 4.7.1 Legyen $B \in \mathbb{R}^{m \times m}$, és $\tau > 0$ szám. Ekkor a következő iterációt stacionárius-iterációnak nevezzük.

$$B \cdot \frac{x^{n+1} - x^n}{\tau} + Ax^n = b$$

Megjegyzés 4 Az előbb említett iterációs módszerek összegezve:

- *Jacobi:* $B = D, \quad \tau = 1$
- *Gauss-Seidel:* $B = D + L, \quad \tau = 1$
- *Még általánosabb:* $B \leftrightarrow B_n, \quad \tau \leftrightarrow \tau_n$

Említés szintjén még egy stacionárius iteráció a *Túlrelaxációs módszer* vagy angolul *Successive overrelaxation method (SOR)*:

$$B = D + \omega L, \quad \tau = \omega, \quad \text{ahol } \omega > 0 \text{ adott paraméter}$$

$$(D + \omega L) \cdot \frac{x^{n+1} - x^n}{\omega} + Ax^n = b$$

Megjegyzés 5 A SOR módszert $\omega = 1$ -el írva visszakapjuk a Gauss-Seidel-iterációt.

4.8. Stacionárius iteráció konvergenciája

Emlék:

$$B \frac{x^{n+1} - x^n}{\tau} + A \cdot x^n = b \quad (\text{SI})$$

$$Ax = b$$

Tegyük fel, hogy A szimmetrikus pozitív definit (szpd). Tehát $A = A^T$, $x^T Ax > 0$, ha $x \neq 0$. Másképpen, $\exists \delta > 0 : (Ax, x) \geq \delta \cdot \|x\|^2$.

Jelölje x^* a 3.1 egyenlet megoldását, azaz $Ax^* = b$ és $e_n := x^n - x^*$ (az n -edik iteráció hibáját).

Definíció 4.8.1 Azt mondjuk hogy a stacionárius iteráció (SI) konvergens, ha $\exists \lim x_n$ és $x_n \rightarrow x^*$, azaz $\lim_{n \rightarrow \infty} e_n = 0$.

Állítás 4.8.1 Tegyük föl, hogy A szpd. Ha $\exists B^{-1}$, és $\tau > 0$ paraméter olyan, hogy $B - 0.5\tau A$ szpd, akkor a stacionárius iteráció konvergens.

Bizonyítás:

$$x^n = e_n + x^*, \quad x^{n+1} = e_{n+1} + x^* \rightsquigarrow (\text{SI})\text{-be beírva}$$

$$B \frac{e_{n+1} + x^* - e_n - x^*}{\tau} + A e_n + A x^* = b$$

$$B \frac{e_{n+1} - e_n}{\tau} + A e_n = 0 \quad (3) \text{ hibaegyenlet}$$

Fejezzük ki e_{n+1} -el

$$B e_{n+1} = (B - \tau A) e_n$$

$$e_{n+1} = (I - \tau B^{-1} A) e_n$$

$$A e_{n+1} = (A - \tau A B^{-1} A) e_n$$

$$\implies (A e_{n+1}, e_{n+1}) = (A e_n - \tau A B^{-1} A e_n, e_n - \tau B^{-1} A e_n)$$

$$= (A e_n, e_n) - \tau (A B^{-1} A e_n, e_n) - \tau (A e_n, B^{-1} A e_n) + \tau^2 (A B^{-1} A e_n, B^{-1} A e_n)$$

Tudjuk, hogy

$$(A B^{-1} A e_n, e_n) = (B^{-1} A e_n, A^T e_n) = (B^{-1} A e_n, A e_n) = (A e_n, B^{-1} A e_n)$$

Tehát

$$(A e_{n+1}, e_{n+1}) = (A e_n, e_n) - 2\tau (A B^{-1} A e_n, e_n) + \tau^2 (A B^{-1} A e_n, B^{-1} A e_n)$$

Jelölje $J_n = (A e_n, e_n)$. Ezzel

$$J_{n+1} = J_n - 2\tau (A e_n, B^{-1} A e_n) + \tau^2 (A B^{-1} A e_n, \overbrace{B^{-1} A e_n}^{y_n})$$

Ezzel $B y_n = A e_n$

$$= J_n - 2\tau (B y_n, y_n) + \tau^2 (A y_n, y_n) = J_n - 2\tau \left((B y_n, y_n) - \frac{\tau}{2} (A y_n, y_n) \right)$$

$$\begin{aligned} \leadsto J_{n+1} &= J_n - 2\tau((B - 0.5\tau A)y_n, y_n) \\ \implies J_{n+1} &\leq J_n \end{aligned} \quad (4.4)$$

Mert feltétel szerint $\tau > 0$ és $(B - 0.5\tau A)$ szpd, tehát pozitív szor pozitív tagot vonunk ki, tehát egy pozitív számot vonunk ki. Ezért a jobb oldal kisebb mint J_n . Így a (J_n) sorozat monoton csökkenő, és $J_n \geq 0$, (mert $J_n = (Ae_n, e_n)$), tehát ez a sorozat alulról korlátos. Tehát (J_n) konvergens, jelölés $J^* := \lim_{n \rightarrow \infty} J_n$

4.4-ban vegyüünk limeszt \leadsto

$$\begin{aligned} J^* &= J^* - 2\tau \lim_{n \rightarrow \infty} ((B - 0.5\tau A)y_n, y_n) \\ \implies \lim_{n \rightarrow \infty} ((B - 0.5\tau A)y_n, y_n) &= 0 \end{aligned}$$

Mivel $B - 0.5\tau A$ szpd, ezért $\exists \delta > 0 : ((B - 0.5\tau A)y_n, y_n) \geq \delta \cdot \|y_n\|^2$. Rendőrelv:

$$\begin{aligned} ((B - 0.5\tau A)y_n, y_n) &\geq \delta \cdot \|y_n\|^2 \geq 0 \\ \lim_{n \rightarrow \infty} ((B - 0.5\tau A)y_n, y_n) &\geq \lim_{n \rightarrow \infty} \delta \cdot \|y_n\|^2 \geq 0 \\ \implies 0 &\geq \lim_{n \rightarrow \infty} \delta \cdot \|y_n\|^2 \geq 0 \implies \lim_{n \rightarrow \infty} \delta \cdot \|y_n\|^2 = 0 \end{aligned}$$

Mivel $y_n = B^{-1}Ae_n \leadsto e_n = A^{-1}By_n$, ezért

$$0 \leq \|e_n\| = \|A^{-1}By_n\| \leq \|A^{-1}B\| \cdot \|y_n\| \rightarrow 0 \implies \|e_n\| \rightarrow 0 \implies e_n \rightarrow 0$$

Ezzel beláttuk, hogy konvergens, mert a hiba 0-hoz tart.

4.9. SOR-módszer konvergenciája

Kérdés 8 *Hogyan válasszuk meg ω paramétert, hogy konvergáljon?*

Észrevétel: ω választása erősen függ A -tól.

Állítás 4.9.1 *Tetszőleges $A \in \mathbb{R}^{n \times m}$ esetén a SOR-módszer konvergenciájához szükséges, hogy $\omega \in (0, 2)$.*

Állítás 4.9.2 *Ha A szpd, akkor $\omega \in (0, 2)$ elégséges is a konvergenciához.*

Következmény 3 *Ha A szimmetrikus pozitív definit (szpd), akkor a Gauss-Seidel iteráció konvergens, mert a Gauss-Seidel iteráció pont a SOR-módszer $\omega = 1$ -el.*

5. fejezet

Gradiens alapú módszerek

Tekintsük megint a következő egyenletet.

$$Ax = b \quad (1)$$

Tegyük fel, hogy A szimmetrikus pozitív definit (szpd).

Definíció 5.0.1 *Definiáljuk a következő $\phi : \mathbb{R}^m \rightarrow \mathbb{R}$ függvényt:*

$$\phi(x) := \frac{1}{2}(x, Ax) - (x, b)$$

ez differenciálható \mathbb{R}^m -en.

Célunk, hogy a $\phi(x)$ függvényt minimalizáljuk, tehát nézzük meg, hogy hol lesz 0 a gradiense.

$$\phi'(x) = \nabla \phi(x) = Ax - b \quad (\text{számolással ellenőrizhető})$$

Ekkor pont az $r := b - Ax$ maradékvektor -1 szeresét kapjuk.

Hol 0 a gradiens?

$$\phi'(x) = Ax - b = 0 \leadsto x = A^{-1}b$$

ez éppen a 3.1 megoldása, tehát a $\phi(x)$ függvényt minimalizálni ekvivalens azzal, hogy megoldjuk a 3.1 egyenletet.

$$\phi''(x) = A$$

Mivel feltettük, hogy A szpd, ezért $\phi''(x)$ pozitív definit, tehát ahol a gradiens nulla ott lokális minimum hely van.

$\implies x^*$ az egyetlen lokális minimum hely / globális minimum helye ϕ -nek.

Kérdés 9 *Hogy néz ki a ϕ függvény?*

Példa 4 *Tekintsünk egy két dimenziós példát, ahol már a következő egyenletrendszerrel tartunk:*

$$2x_1 = 4$$

$$8x_2 = 8$$

Megoldás:

Ránézésre látszik, hogy a megoldás $x_1^* = 2$, $x_2^* = 1$

Írjuk ki A és b teljes alakját.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

Helyettesítsük be A -t és b a $\phi(x)$ függvénybe.

$$\begin{aligned} \phi(x) &= \frac{1}{2}(x, Ax) - (x, b) \\ \phi(x) &= \frac{1}{2} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} 2x_1 \\ 8x_2 \end{bmatrix} \right) - \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \end{bmatrix} \right) \\ &= \frac{1}{2}x_1 2x_1 + \frac{1}{2}x_2 8x_2 - 4x_1 - 8x_2 = (x_1 - 2)^2 + 4(x_2 - 1)^2 - 8 \end{aligned}$$

Vizsgáljuk a szintvonalait ennek a függvénynek.

$Ac = 0$ -hoz tartozó szintvonal:

$$(x_1 - 2)^2 + 4(x_2 - 1)^2 - 8 = 0$$

$$\frac{(x_1 - 2)^2}{8} + \frac{(x_2 - 1)^2}{2} = 1$$

Tehát azt kaptuk, hogy ez egy $(2, 1)$ középpontú ellipszis $\sqrt{8}, \sqrt{2}$ hosszú főtengelyekkel. Azaz valóban $(2, 1)$ a megoldás.

Tehát a függvény szintvonalai koncentrikus hiperellipszoidok!

Először gondoljuk meg, hogy egy $x \in \mathbb{R}^m$ pontot és egy $p \neq 0$ vektort rögzítve p irány mentén hol veszi fel a ϕ a legkisebb értéket?

Jelölés: $g(\alpha) := \phi(x + \alpha p)$

Kérdés 10 Mely α -ra lesz $g(\alpha)$ függvény értéke minimális?

Állítás 5.0.1 A $g(\alpha) = \phi(x + \alpha p)$ függvény egyértelmű minimumát az

$$\alpha = \frac{(p, r)}{(p, Ap)}$$

megvalósítás esetén veszi föl!

Bizonyítás: Faragó I. Numerikus módszerek jegyzet 83. oldalán található.[1]

Kérdés 11 Hogyan válasszuk meg p_1, p_2, \dots keresési irányokat?

5.1. Gradiens módszer

Tudjuk: A $\nabla\phi$ -vel ellentétes irányban a legmeredekebb a lejtés.

x_i pontban p_{i+1} -el jelölve a keresési irányt:

$$\begin{aligned} p_{i+1} &:= -\nabla\phi(x_i) \\ \nabla\phi(x) &= Ax - b = -r \\ \implies p_{i+1} &:= -\nabla\phi(x_i) = b - Ax_i = r_i \end{aligned}$$

ami éppen az x_i pontbeli maradékvektor.

$$x_i \rightsquigarrow x_{i+1} = x_i + \alpha \cdot p_{i+1} = x_i + \frac{(p_{i+1}, r_i)}{(p_{i+1}, Ap_{i+1})} \cdot p_{i+1} = x_i + \frac{(r_i, r_i)}{(r_i, Ar_i)} \cdot r_i$$

Kérdés 12 *Mi lesz x_{i+1} -ben a maradékvektor?*

$$r_{i+1} = b - Ax_{i+1} = b - A \cdot \left(x_i + \frac{(r_i, r_i)}{(r_i, Ar_i)} \cdot r_i \right)$$

Vegyük észre: $r_i \perp r_{i+1}$, mert addig megyünk p_i irányban ameddig nem érintjük a következő szintvonalat, amire a következő gradiens merőleges.

Ez előző vizuálisan magyarázza az egymást követő irányok merőlegességét, de bizonyítsuk be formálisabban. Írjuk fel a skaláris szorzatát az egymást követő irányoknak!

$$\begin{aligned} (r_i, r_{i+1}) &\stackrel{?}{=} 0 \\ (r_i, r_{i+1}) &= \left(r_i, b - A \left(x_i + \frac{(r_i, r_i)}{(r_i, Ar_i)} \cdot r_i \right) \right) \\ &= (r_i, r_i) - (r_i, A \frac{(r_i, r_i)}{(r_i, Ar_i)} \cdot r_i) \\ &= (r_i, r_i) - \frac{(r_i, r_i)}{(r_i, Ar_i)} (r_i, Ar_i) \\ &= (r_i, r_i) - (r_i, r_i) = 0 \end{aligned}$$

Megjegyzés 6 *Ha $\text{cond}_2(A)$ nagy, akkor lassú a konvergencia.*

5.2. Konjugált gradiens-módszer

Az előbb láttuk be, hogy a gradiens módszernél a p_1 keresési irány $(r_0) \perp r_1$. Azaz:

$$0 = (p_1, r_1) = (p_1, b - Ax_1) = (p_1, Ax^* - Ax_1) = (p_1, A(x^* - x_1))$$

Definíció 5.2.1 Legyen $A \in \mathbb{R}^{m \times m}$ szimmetrikus pozitív definit (szpd). Azt mondjuk, hogy x és $y \in \mathbb{R}^m$ vektorok A -konjugáltak/ortogonálisak, ha $(x, Ay) = 0$.

Tehát olyan keresési irányt lenne érdemes választani, amely p_1 -re A -ortogonális!
Keressük p_2 -t a következő alakban:

$$\begin{aligned} p_2 &= r_1 - \beta_1 \cdot p_1 \\ (p_1, A(r_1 - \beta_1 \cdot p_1)) &= 0 \\ \beta_1 &= ? \end{aligned}$$

$$\begin{aligned} (p_1, Ar_1) - \beta_1(p_1, Ap_1) &= 0 \\ \implies \beta_1 &= \frac{(p_1, Ar_1)}{(p_1, Ap_1)} \end{aligned}$$

Ezen β_1 -et választva, a $p_2 = r_1 - \beta_1 \cdot p_1$ irányba lépve az x^* minimum helybe lépünk! Tehát $m = 2$ esetén 2 lépésben meg tudjuk határozni a lineáris egyenletrendszer megoldását.

Megjegyzés 7 $A \in \mathbb{R}^{m \times m}$ esetén is általánosítható az eljárás. Ekkor legfeljebb m lépésben megkapjuk a megoldást.

6. fejezet

Általános algebrai egyenletek megoldása

Ebben a fejezetben egyismeretlenes valós egyenletekkel foglalkozunk. Egy ilyen egyenlet mindig felírható a következő alakban:

$$f(x) = 0 \tag{6.1}$$

ahol $f : \mathbb{R} \rightarrow \mathbb{R}$ függvény.

Ezzel 6.1-nek a megoldása ugyanaz mint f zérushelye. Ezt keressük a továbbiakban!

6.1. Gyökök stabilitása

Kérdés 13 *Mennyire érzékeny a megoldás f kis megváltoztatására?*

Tegyük fel, hogy 6.1 helyett az

$$\tilde{f}(x) = 0 \tag{6.2}$$

Egyenletet oldjuk meg, és tegyük fel, hogy 6.1-nek és 6.2-nek is $\exists!$ megoldása, melyek x^* illetve \tilde{x}^* rendre.

A következő legyen a mérőszámunk az eltérésre:

$$|x^* - \tilde{x}^*| \leq ?$$

Ha f és \tilde{f} csak *kicsit* tér el egymástól, akkor legfeljebb mennyire tér el x^* és \tilde{x}^* ? Mérje $\max_{[a,b]} |f - \tilde{f}|$ az f és \tilde{f} eltérését.

Tegyük fel, hogy $f \in C[a, b] \cap D(a, b)$

Ismétlés: (Lagrange-középérték tétel) Tegyük fel, hogy $f \in C[a, b] \cap D(a, b)$. Ekkor $\exists c \in (a, b)$ úgy, hogy

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Továbbá tegyük fel, hogy x^* és $\tilde{x}^* \in [a, b]$, és $\max_{[a, b]} |f - \tilde{f}| < \varepsilon$. Alkalmazzuk a Lagrange-középérték tételt az $[x^*, \tilde{x}^*]$ intervallumon (feltéve, hogy $x^* < \tilde{x}^*$):

$$\exists c \in (x^*, \tilde{x}^*) : f(\tilde{x}^*) - f(x^*) = f'(c)(\tilde{x}^* - x^*)$$

Tegyük fel, hogy $f'(x) \neq 0 \quad \forall x \in (x^*, \tilde{x}^*)$.

$$\iff |\tilde{x}^* - x^*| = \left| \frac{f(\tilde{x}^*)}{f'(c)} \right| = \frac{|f(\tilde{x}^*) - \tilde{f}(\tilde{x}^*)|}{|f'(c)|} < \frac{\varepsilon}{\min_{[a, b]} |f'|}$$

Definíció 6.1.1 Az $M := \frac{1}{\min_{[a, b]} |f'|}$ számot a 6.1 egyenlet kondicionáltsági számának nevezzük.

Tehát ha $\max_{[a, b]} |f - \tilde{f}| < \varepsilon$, akkor $|\tilde{x}^* - x^*| < M \cdot \varepsilon$.

6.2. Konvergencia sebesség

Tegyük fel, hogy $\lim_{k \rightarrow \infty} x_k = x^*$, és legyen $e_k := x_k - x^*$. ($\lim_{k \rightarrow \infty} e_k = 0$ vagy $\lim_{k \rightarrow \infty} |e_k| = 0$)

Definíció 6.2.1 Azt mondjuk, hogy az (x_k) sorozat konvergencia rendje $p \geq 1$, ha

$$\lim_{k \rightarrow \infty} \frac{\log |e_k|}{\log |e_{k-1}|} = p$$

- Ha $p = 1$, akkor lineáris vagy elsőrendű konvergenciáról beszélünk.
- Ha $p = 2$, akkor másodrendű vagy kvadratis konvergenciáról beszélünk.

Példa 5 Elsőrendű és másodrendű konvergens sorozatok hibatagjainak lecsengésére példák.

Elsőrendű:

	$ e_k $	$\frac{\log e_k }{\log e_{k-1} }$
$k = 1$	10^{-3}	N/A
$k = 2$	10^{-4}	1.33
$k = 3$	10^{-5}	1.25

Másodrendű:

	$ e_k $	$\frac{\log e_k }{\log e_{k-1} }$
$k = 1$	10^{-3}	N/A
$k = 2$	10^{-6}	2
$k = 3$	10^{-12}	2

Állítás 6.2.1 Tegyük fel, hogy $|e_k| = c_k \cdot |e_{k-1}|$, $k = 1, 2, \dots$ ahol $0 < \underline{c} \leq c_k \leq \bar{c} < 1$. Valamilyen \underline{c} és \bar{c} konstansokra.

Ekkor $x_k \rightarrow x^*$ monoton módon és elsőrendben.

Bizonyítás: Monotonan, mivel $0 < c_k < 1 \implies |e_k| < |e_{k-1}| \quad \forall k = 1, 2, \dots$
 $\implies (|e_k|)$ sorozat monoton csökkenő.

Konvergál, mivel $|e_k| = c_k \cdot |e_{k-1}| \leq \bar{c} \cdot |e_{k-1}| \leq \bar{c} \cdot \bar{c} \cdot |e_{k-2}| \leq \dots \leq \bar{c}^k \cdot |e_0|$. Mivel $\bar{c} < 1$ ezért tényleg $\lim_{k \rightarrow \infty} |e_k| = 0$.

A feltételben lévő egyenletnek mindkét oldalán logaritmust véve:

$$\begin{aligned} \log|e_k| &= \log c_k + \log|e_{k-1}| \\ \implies \frac{\log|e_k|}{\log|e_{k-1}|} &= \frac{\log c_k}{\log|e_{k-1}|} + 1 \end{aligned}$$

Látszik, hogy $\log|e_{k-1}| \rightarrow -\infty$. Mostmár elegendő lenne belátni, hogy $\log c_k$ korlátos.

$$\log \underline{c} < \log c_k \leq 0$$

Tehát $\frac{\log c_k}{\log|e_{k-1}|} \rightarrow 0 \implies$ a jobb oldal $\rightarrow 1 \implies p = 1$ a konvergenca rendje, azaz elsőrendű a konvergenca.

Állítás 6.2.2 Tegyük fel, hogy $|e_k| = c_k \cdot |e_{k-1}|^p$ $k = 1, 2, \dots$ ahol $p > 1$ és $0 < \underline{c} \leq c_k \leq \bar{c} < +\infty$. Valamilyen \underline{c} és \bar{c} konstansokra. Továbbá $\bar{c}^{1/p-1} \cdot |e_0| < 1$. Ekkor (x_k) konvergens és a konvergenca rendje p .

Megjegyzés 8 Az utóbbi feltétel azt jelelnti, hogy a konvergenca csak akkor következik, ha x_0 elég közel van x^* -hoz. Ugyanakkor $\bar{c} < +\infty$, és nem kell teljesülnie, hogy $\bar{c} < 1$.

Irodalomjegyzék

- [1] Faragó István, H.R.: Numerikus módszerek. Typotex (2016)

