Numerikus Analízis I.

Sövegjártó András

Jegyzet másodéves programozó és

programtervező matematikus szakos hallgatóknak

2003.

„A sikerhez és tudáshoz vezető út senki előtt sincs zárva,
akiben van bátorság és elszántság, hogy változzék,
nyitottság, hogy mások tapasztalataiból tanuljon és állhatatosság,
hogy gyakorlással elsajátítsa a sikeres cselekvés technikáját."

ELŐSZÓ

A numerikus analízis

1. az úgynevezett konstruktív matematika nagy fejezete,
2. a matematikai modellezés folyamatának fontos része.

1.

Az alkalmazott matematikában leggyakrabban nem elégszünk meg azzal, hogy tudjuk (belátjuk), a vizsgált problémának létezik megoldása, hanem meg is akarjuk annak számszerű értékét határozni, ez pedig kétféle módon lehetséges: vagy véges sok lépésben elvileg a pontos megoldást (számítási, kerekítési hibáktól eltekintve), vagy egy általunk előre megadott pontossággal, úgynevezett közelítő megoldások (elvileg végtelen) sorozatával.

2.

A matematikai modellezés egy több lépésből álló (kör)folyamat. A valóság egy részét vizsgálva, igyekezvén a jelenséget matematikailag leírni , elkészítjük annak egy lehetséges matematikai modelljét. Ez egyenletek és úgynevezett bemenő adatok (paraméterek) rendszere. (A numerikus analízisben mi már kész, adott matematikai modelleket vizsgálunk.) Sajnos, leggyakrabban a modellnek a pontos megoldása nem állítható elő úgynevezett analitikai eszközökkel véges sok lépésben. Ilyenkor numerikus módszerekre van szükség. (Egyszerű példa a $\sqrt{2}$ kiszámítása.) A numerikus módszer kidolgozása után annak programozása és a programkód tesztelése következik, majd a konkrét számítások elvégzése. A két utolsó lépése a matematikai modellezésnek a számítások, eredmények analízise, és az eredményeknek másokkal könnyen közölhető formában való megjelenítése (esetleg vizualizálása rajzok, képek formájában). Tengernyi számadatból a megoldást értelmezni kell tudni. Gyakran, az eredmények analízise során kiderül, hogy a kapott eredmény nem felel meg a valóságnak (pl. a kísérleti eredményeknek, vagy annak, ami várható lett volna). Ilyenkor a modellezést újra elölről kell kezdeni. A modell finomításával, a jelenség pontosabb leírásával a lépéseket újra meg kell ismételni. Ebben a (kör)folyamatban a numerikus analízis feladata numerikus módszerek kidolgozása és tanulmányozása.

A számítógépek megjelenésével olyan feladatok megoldása is lehetővé vált, amelyekre korábban gondolni sem lehetett. Ilyenek például az időjárás előrejelzése, űrkutatási és komplex műszaki feladatok megoldása stb.

Ha nyomon követjük a matematika fejlődésének történetét, azt láthatjuk, hogy a nagy matematikusok, természettudósok szinte mindegyike foglalkozott numerikus módszerek kidolgozásával is, hogy konkrét számítási eredményeket is kapjanak. Nevezetes példa erre Newton, akinek a fizikában, matematikában és numerikus módszerek területén elért eredményei is csodálatra méltóak! De megemlíthetnénk hasonló okból Gausst is.

2020-ra a számítógépek műveleti sebessége elérheti az emberi agy műveleti sebességét, azaz húsz millió milliárd ops-ot ($2 * 10^{16}$ művelet per sec-ot). Ez a tény a numerikus módszerek kidolgozásának és alkalmazásainak kérdését is új megvilágításba helyezi és újabb problémákat is vet fel a kutatásban a hatékony algoritmusokkal kapcsoaltban. De ez már a jövő és a fiatal nemzedék feladata lesz. A numerikus módszerekkel való ismerkedés első lépéseként ajánljuk a jegyzet anyagát a másodéves programozó és programtervező matematikus hallgatóknak, és nem utolsósorban mindazoknak, akiket a tárgy érdekel. Kívánjuk, hogy a matematikával való foglalatosság nagy örömet nyújtson (és az sem baj, ha hasznot is)!

Ezúton szeretnék köszönetet mondani Sipos Boglárkának, aki a jegyzet megírásának fáradtságos munkáját vállalta. Nélküle, lelkes hozzáállása és munkája nélkül az anyag nem készült volna el.

Budapest, 2003. szeptember 1.

Sövegjártó András

E kötet a programtervező matematikus szak nappali tagozatán a 2002/2003-as tanévben elhangzott I. félévi, Dr. Sövegjártó András egyetemi docens által tartott numerikus analízis előadássorozat alapján készült, kiegészítésekkel, Donald E. Knuth TEX szövegszedő rendszerével.

Eötvös Loránd Tudományegyetem Informatikai Kar, 2003.

Tartalomjegyzék

1. HIBÁK A MATEMATIKAI MODELLEZÉS FOLYAMATÁBAN

I. Öröklött hibák

1. A matematikai modell hibája
2. A bemenő adatok hibája

} Ezek a további számításokból nem kiküszöbölhetők, öröklődnek.

II. Numerikus módszer hibája

1. Diszkretizációs hiba
2. Approximációs hiba
3. Képlet hiba

III. Számítási hiba

1. Kerekítési hiba
2. Emberi, gépi tévedés

Természetes követelmény, hogy az előző hibák összemérhetők legyenek az utolsóval (kerekítési hibával).

Definíció :

Numerikus algoritmus: Aritmetikai és logikai műveletek sorozata.

Definíció :

Numerikus matematika: Numerikus módszerek és algoritmusok elmélete.

A numerikus módszer, algoritmus közelítő megoldást ad!
Természetes követelmény, hogy az algoritmus adott ε pontosságot véges sok lépésben rövid idő alatt érjen el. Stabil az algoritmus, ha a bemenő adatok kis változtatása a megoldásban is csak kis változtatást okoz, azaz, ha a számolás során a számítási hibák nem nőnek jelentősen.

Példák instabil algoritmusokra:

(1)

$$b_0 = 1\,, \quad b_1 = \frac{\sqrt{5}-1}{2}\,, \quad b_{n+1} = b_{n-1} - b_n$$

(2)

$$I_0 = \ln\frac{6}{5}\,, \quad I_n = -5\,I_{n-1} + \frac{1}{n}$$

1.1. SZÁMÁBRÁZOLÁS TETSZŐLEGES B BÁZISBAN

Tétel

Legyen $2 \leq B \in \mathbb{Z}\,, \quad 0 \neq x \in \mathbb{R}$.

Ekkor x mindig előállítható a következő alakban:

(⋆) $$\boxed{x = \sigma B^k \sum_{n=1}^{\infty} x_n B^{-n}}$$

ahol $\sigma \in \{\,-1\,,\,1\,\}\,, \quad k \in \mathbb{N}$ és $x_n \in \{\,0, 1, \ldots, B-1\,\}$.

Ez az előállítás egyértelmű is, ha $x_1 \neq 0$ (úgy is mondjuk, hogy az x_1 számjegy normált) és ha $\forall \quad N \in \mathbb{N}$ - re $\quad \exists \quad n \geq N$ úgy, hogy

(⋆⋆) $\quad x_n \neq B - 1.$

Megjegyzés :

1.
Tehát a B bázisban az x szám alakja:

$$\boxed{x = \sigma B^k \cdot 0.x_1 x_2 x_3 \ldots}$$

2.
Általában: $\quad B = 2, 8, 10, 16.$

3.
Mivel a számítógépen ábrázolható számok halmaza véges, így alakjuk *(⋆)* helyett a következő lesz:

(1) $$\boxed{x = \sigma B^k \sum_{n=1}^{t} x_n B^{-n}}$$

ahol $t \in \mathbb{N}$ fix

$$m := \sum_{n=1}^{t} x_n B^{-n} \qquad \text{neve: mantissza,}$$

t : a mantissza hossza, $\quad \sigma$: az x szám előjele és

k : az x szám kitevője (exponense), karakterisztikája.

1.1.1. Fixpontos ábrázolás :

($\star$) - ban k rögzített és $x_1 = 0$ is megengedett.

Példa :

$k = 0$ esetén *($\star$)* a $0 \leq |x| \leq 1$ számokat adja.

$k = t$ esetén azon x egészeket kapjuk, amelyekre

$$|x| \leq B^t - 1.$$

Ez tudományos számításokra nem alkalmas a fizikai állandók nagyságrendjének nagy különbözősége miatt. (Pl.: $m_{\text{el}} = 9.11 \cdot 10^{-28}$ g, $c = 2.998 \cdot 10^{10}$ cm/sec.)

1.1.2. Lebegőpontos ábrázolás :

($\star$) - ban a $t > 0$ mantissza hossza fix, és az exponensre alsó és felső korlát adott:

$$k_- \leq k \leq k_+ \qquad (k_-, k_+ \in \mathbb{Z}, \quad k_- < 0, \quad k_+ > 0)$$

Az így ábrázolható számok halmaza:

$$B^{k_- - 1} \leq |x| < B^{k_+}.$$

Ha $|x| < B^{k_- - 1}$, akkor nullával veszi a számítógép.

A B^{k_+} - nál nagyobb számok nem kezelhetők: exponenciális túlcsordulás.

$\varepsilon_0 := B^{k_- - 1}$ a nullához legközelebbi szám. (Ugyanis ***(1)*** - ben $x_1 = 1$.)

Megjegyzés :

Legtöbb számítógépen a B bázis 2 hatvány.

$B = 2$ bináris rendszer. Hátránya, hogy nagy k_-, k_+ számokat kell választani kellően nagy számhalmaz ábrázolásához.

$B = 8, 16$ oktális, hexadecimális rendszerek. Ezeknél minden x_n jegy bináris ábrázolására 3 illetve 4 bit szükséges.

Decimális számrendszer esetén :

$t = 8$ egyszeres pontosság,

$t = 16$ dupla pontosság mellett.

Az *(1)* számok tárolása számítógépen:

$[\pm, k, x_1, \ldots, x_t]$

A géptől és a pontosságtól függően az $m := (x_1, \ldots, x_t)$ mantissza tárolására 4, 8, 16 byte adott. Ezzel párhuzamosan nő k értékkészlete.

Mivel $k_- \leq k \leq k_+$,

a legnagyobb ábrázolható szám :

$$M_\infty := B^{k_+} \cdot \sum_{n=1}^{t} \frac{B-1}{B^n} = B^{k_+} \left(1 - B^{-t}\right) ,$$

a legkisebb :

$$-M_\infty .$$

TEHÁT :

A lebegőpontos számok a $[-M_\infty , M_\infty]$ intervallumbeli racionális számok diszkrét halmaza, ami a nullára szimmetrikus. Így a nullán kívül $(-\varepsilon_0 , \varepsilon_0)$ intervallumban nincs más lebegőpontos szám. A 0 ábrázolása: $[+, 0, 0, \ldots, 0]$.

Az ε_0 - hoz legközelebbi pozitív lebegőpontos szám:

$$B^k - \left(\frac{1}{B} + \frac{1}{B^t}\right) = \varepsilon_0 + B^{k_- - t} = \varepsilon_0 \left(1 + B^{1-t}\right) \qquad \text{és} \qquad \varepsilon_0 \cdot B^{1-t} \ll \varepsilon_0$$

Az 1 mindig lebegőpontos szám, ugyanis $1 = [+, 1, 1, 0, \ldots, 0]$.

Az 1 utáni lebegőpontos szám: $[+, 1, 1, 0, \ldots, 1] = 1 + \varepsilon_1$, ahol $\varepsilon_1 := B^{1-t}$.

ε_1 a gép relatív pontossága; gépi epszilon.

Példa :

Legyen $\quad B = 16 = 2^4.$

Egyszeres pontosságú lebegőpontos számok ábrázolására 32 bit = 4 byte hozzáférhető. 1 bit az úgynevezett előjelbit, 7 bit van az exponens számára. E 7 biten tárolódik a k exponens, amire

$$k_- = -64, \qquad k_+ = 63.$$

Így $\;0 \le k + 64 \le 127 = 2^7 - 1\;$ számok tároltak. A fennmaradó 3 byte-on $t = 6$ db

hexadecimális jegy tárolható.

Legyen $\;x = 123.75 = 7 \cdot 16^1 + 11 \cdot 16^0 + 12 \cdot 16^{-1} = 16^2 \left(7 \cdot 16^{-1} + 11 \cdot 16^{-2} + 12 \cdot 16^{-3}\right)$

ami a következő módon tárolódik:

0	1 0 0 0 0 1 0	0 1 1 1	1 0 1 1	1 1 0 0	0 0
+	66	7	11	12	

A dupla pontosságú lebegőpontos számábrázolásra 8 byte áll rendelkezésre. 7 byte a mantissza $t = 14$.

1.2. LEBEGŐPONTOS ARITMETIKA

Minden x valós számot gépen $\widetilde{x}$ véges előállításával reprezentálunk. Ezt a folyamatot kerekítésnek nevezzük, ami hibát okoz.

Definíció :

Legyen $\quad x, \widetilde{x} \in \mathbb{R}, \quad$ ahol $\widetilde{x}$ az x approximáltja.

Az $|x - \widetilde{x}|$ különbséget abszolút hibának nevezzük, jele: Δx.

Ha $x \neq 0$, az $\dfrac{x - \widetilde{x}}{x}$ hányadost relatív hibának nevezzük, jele: δx.

Tegyük fel, hogy minden számra és a számításokban teljesül a $k_- \le k \le k_+$ feltétel, azaz nincs túlcsordulás.

Megjegyzés :

Tudományos számításoknál a megoldás nagyságrendje nagy mértékben változhat, a relatív hiba érdekes, mivel skálafüggetlen, azaz $x \to \alpha x$, $\widetilde{x} \to \alpha\widetilde{x}$,,skálázás" a relatív hibát nem változtatja meg.

KEREKÍTÉSI SZABÁLY :

Legyen $B \geq 2$ páros egész, $t \in \mathbb{Z}^+$, $x \in \mathbb{R} \setminus \{0\}$ és x *(⋆)* előállítású. Ekkor az

$$Rd_t\left(x\right) := \begin{cases} \sigma B^k \cdot \sum\limits_{n=1}^{t} x_n B^{-n} & \text{ha } x_{t+1} < \dfrac{B}{2} \\ \sigma B^k \left(\sum\limits_{n=1}^{t} x_n B^{-n} + B^{-t} \right) & \text{ha } x_{t+1} \geq \dfrac{B}{2} \end{cases}$$

számot az x t - jegyre kerekített értékének nevezzük.

Megjegyzés :

$B = 10$ bázisban ez a szokásos kerekítési szabály.

Tétel

1. Érvényes a következő előállítás: $Rd_t\left(x\right) = \sigma B^{k'} \cdot \sum\limits_{n=1}^{t} x_n' B^{-n}$.

2. Az abszolút hibára teljesül, hogy: $\left|Rd_t\left(x\right) - x\right| \leq \dfrac{B^{k-t}}{2}$.

3. A relatív hibára érvényes, hogy: $\left|\dfrac{Rd_t\left(x\right) - x}{x}\right| \leq \dfrac{B^{-t+1}}{2}$.

4. Az is igaz, hogy: $\left|\dfrac{Rd_t\left(x\right) - x}{Rd_t\left(x\right)}\right| \leq \dfrac{B^{-t+1}}{2}$

(relatív hiba a kerekített értékre vonatkozóan).

Definíció :

A $\tau := \dfrac{B^{-t+1}}{2}$ számot a t - jegyű lebegőpontos aritmetika relatív pontosságának nevezzük.

Példa :

$\tau = 0.5 \cdot 16^{-5} < 0.5 \cdot 10^{-6}$.

Így nincs sok értelme a bemenő és kijövő adatokat több, mint 7 jeggyel a mantisszában megadni.

Duplapontos aritmetikában: $\tau = 0.5 \cdot 16^{-13} < 0.5 \cdot 10^{-15}$.

KIEGYSZERŰSÖDÉS :

Két közel egyenlő szám különbségének képzésekor lép fel a lebegőpontos aritmetikában. Először ugyanis az exponens azonos, de a mantissza kezdőjegyei nullává válnak. Az ezt követő normalizálásnál fontos jegyek elvesznek. (A mantissza jegyei balra tolódnak, az exponens csökken.)

Példa :

$a = \sqrt{9876} - \sqrt{9875} = 0.000503142 \cdot 10^1$, ugyanis:

$\sqrt{9876} = 9.937806599 \cdot 10^1$

$\sqrt{9875} = 9.937303457 \cdot 10^1$

Normalizálva :

$a = 5.031420000 \cdot 10^{-3}$

és itt a mantissza végén álló négy nulla értelmetlen.

Példa :

$$ax^2 + bx + c = 0 \qquad (a \neq 0)\,, \qquad (\star) \qquad x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Ha $b^2 \gg |4ac| \implies \sqrt{b^2 - 4ac} \approx |b|$ és így $(\star)$ képlet egyik előjelre súlyos kiegyszerűsödésre vezet. Ezért először a nagyobb abszolútértékű gyököt számoljuk a következő módon:

$$x_1 = \frac{-\left[b + \mathbf{sign}\,(b)\,\sqrt{b^2 - 4ac}\right]}{2a}\,, \quad \text{majd} \quad \implies \quad x_1 x_2 = \frac{c}{a} \quad \implies \quad x_2 = \frac{c}{a x_1}.$$

Ha $b^2 \approx 4ac$, akkor átrendezéssel sem kerülhető el a kiegyszerűsödés!

1.3. MŰVELETEK HIBÁI

Legyen x, y pontos érték, a, b az x, y közelítő értékei.

Definíció :

Az a közelítő érték hibája: $\Delta a := x - a$
A b közelítő érték hibája: $\Delta b := y - b$

Definíció :

Az a közelítő érték abszolút hibája: $|\,\Delta a\,| = |\,x - a\,|$
A b közelítő érték abszolút hibája: $|\,\Delta b\,| = |\,y - b\,|$

Definíció :

Az a közelítő érték abszolút hibakorlátja minden olyan Δ_a szám, amelyre teljesül, hogy $|\,\Delta a\,| \le \Delta_a$.

Hasonlóan: $|\,\Delta b\,| \le \Delta_b$

Definíció :

Az a közelítő érték relatív hibája: $\delta a := \dfrac{\Delta a}{x} \approx \dfrac{\Delta a}{a}$

A b közelítő érték relatív hibája: $\delta b := \dfrac{\Delta b}{y} \approx \dfrac{\Delta b}{b}$

Definíció :

Az a közelítő érték relatív hibakorlátja minden olyan δ_a szám, amelyre teljesül, hogy $\left|\dfrac{\Delta a}{a}\right| \le \delta_a$

Hasonlóan: $\left|\dfrac{\Delta b}{b}\right| \le \delta_b$

1.3.1. Összeg abszolút és relatív hibája, hibakorlátja

(a + b) hibája:

$$\Delta(a+b) = (x+y)-(a+b) = (x-a)+(y-b) = \Delta a+\Delta b \implies |\,\Delta(a+b)\,| \le |\,\Delta a\,|+|\,\Delta b\,| \le \Delta_a+\Delta_b$$
$$(\, y - b = \Delta b \implies y = b + \Delta b \,)$$

Összeg abszolút hibakorlátja: $\Delta_{(a+b)} = \Delta_a + \Delta_b$

Összeg relatív hibakorlátja: $\left|\dfrac{\Delta(a+b)}{a+b}\right| \le \dfrac{\Delta_a+\Delta_b}{|\,a+b\,|} = \dfrac{|\,a\,|\,\delta_a + |\,b\,|\,\delta_b}{|\,a+b\,|} \qquad \left(\delta_a = \dfrac{\Delta_a}{|\,a\,|}\right)$

1.3.2. Különbség abszolút és relatív hibája, hibakorlátja

(a − b) hibája:

$$\Delta(a-b)=(x-y)-(a-b)=(x-a)-(y-b)=\Delta a-\Delta b \implies |\,\Delta(a-b)\,|\le|\,\Delta a\,|+|\,\Delta b\,|\le\Delta_a+\Delta_b$$

Különbség abszolút hibakorlátja: $\Delta_{(a-b)}=\Delta_a+\Delta_b$

Különbség relatív hibakorlátja: $\left|\dfrac{\Delta(a-b)}{a-b}\right|\le\dfrac{\Delta_a+\Delta_b}{|\,a-b\,|}=\dfrac{|\,a\,|\,\delta_a+|\,b\,|\,\delta_b}{|\,a-b\,|}$

Közeli számok kivonását kerülni kell!

1.3.3. Szorzat abszolút és relatív hibája, hibakorlátja

(a · b) hibája:

$$\Delta(a\cdot b)=xy-ab=xy-ay+ay-ab=y(x-a)+a(y-b)=y\Delta a+a\Delta b=(b+\Delta b)\Delta a+a\Delta b=$$

$$=b\Delta a+a\Delta b+\underbrace{\Delta a\Delta b}_{\text{elhanyagolható}}\implies|\,\Delta(a\cdot b)\,|\le|\,b\Delta a\,|+|\,a\Delta b\,|\le|\,b\,|\,\Delta_a+|\,a\,|\,\Delta_b$$

Szorzat abszolút hibakorlátja: $\Delta_{ab}=|\,b\,|\,\Delta_a+|\,a\,|\,\Delta_b$

Szorzat relatív hibakorlátja: $\left|\dfrac{\Delta(a\cdot b)}{a\cdot b}\right|\le\dfrac{|\,b\,|\,\Delta_a+|\,a\,|\,\Delta_b}{|\,a\cdot b\,|}=\dfrac{|\,a\,|\,|\,b\,|\,\delta_a+|\,a\,|\,|\,b\,|\,\delta_b}{|\,a\cdot b\,|}=\delta_a+\delta_b$

1.3.4. Hányados abszolút és relatív hibája, hibakorlátja

$\left(\dfrac{\mathbf{a}}{\mathbf{b}}\right)$ **hibája:**

$$\Delta\left(\frac{a}{b}\right)=\frac{x}{y}-\frac{a}{b}=\frac{a}{b}\left(\frac{xb}{ya}-1\right)=\frac{a}{b}\left(\frac{xb-ya}{ya}\right)=\frac{a}{b}\left(\frac{xb-ab+ab-ya}{ya}\right)=\frac{a}{b}\left(\frac{b(x-a)-a(y-b)}{ya}\right)=$$

$$=\frac{a}{b}\left(\frac{b\Delta a-a\Delta b}{a(b+\Delta b)}\cdot\frac{b-\Delta b}{b-\Delta b}\right)=\frac{a}{b}\left(\frac{b^2\Delta a-b\Delta a\Delta b}{a(b^2-\Delta^2 b)}-\frac{ab\Delta b+a\Delta^2 b}{a(b^2-\Delta^2 b)}\right)\approx\frac{a}{b}\left(\frac{b^2\Delta a}{ab^2}-\frac{ab\Delta b}{ab^2}\right)=$$

$$=\frac{a}{b}\left(\frac{\Delta a}{a}-\frac{\Delta b}{b}\right)\implies\left|\Delta\left(\frac{a}{b}\right)\right|\le\left|\frac{a}{b}\right|\left(\left|\frac{\Delta a}{a}\right|+\left|\frac{\Delta b}{b}\right|\right)\le\left|\frac{a}{b}\right|\left(\frac{\Delta a}{|\,a\,|}+\frac{\Delta b}{|\,b\,|}\right)$$

Hányados abszolút hibakorlátja: $\Delta_{\frac{a}{b}}=\left|\dfrac{a}{b}\right|\left(\dfrac{\Delta a}{|\,a\,|}+\dfrac{\Delta b}{|\,b\,|}\right)$

Hányados relatív hibakorlátja: $\left|\dfrac{\Delta\left(\frac{a}{b}\right)}{\frac{a}{b}}\right|\le\dfrac{\left|\frac{a}{b}\right|\left(\frac{\Delta_a}{|\,a\,|}+\frac{\Delta_b}{|\,b\,|}\right)}{\left|\frac{a}{b}\right|}=\dfrac{\Delta_a}{|\,a\,|}+\dfrac{\Delta_b}{|\,b\,|}=\delta_a+\delta_b$

2. LINEÁRIS EGYENLETRENDSZEREK DIREKT MEGOLDÁSI MÓDSZEREI

Legyen

(1) $\quad A\underline{\mathbf{x}} = \underline{\mathbf{b}}$

Ahol A nem túl nagy rendű úgynevezett telemátrix (együttható mátrix), $A \in \mathbb{R}^{N\times N}$, $\underline{\mathbf{x}} \in \mathbb{R}^N$ ismeretlen vektor, $\underline{\mathbf{b}} \in \mathbb{R}^N$ ismert vektor.

Tétel **Cramer - szabály**

Ha $\det A \neq 0$, akkor ***(1)*** -nek $\exists!$ megoldása, amit a következő módon kapunk meg:

$$x_k = \frac{\triangle_k}{\triangle} \qquad k = 1, 2, \ldots, N.$$

Ahol $\det A = \triangle$.
Csak kis rendű mátrixok esetén célszerű alkalmazni. (N+1) mátrixot kell kiszámítani (N! rendű műveletet jelent.) Nagy környezethibára vezethet.

Tétel

Az $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$ egyenletnek $\exists$ $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ megoldása $\iff$ $\det A = 0$.

Következmény

$\det A \neq 0$ esetén csak triviális megodás létezik. [Az ***(1)*** megoldása ekkor $\underline{\mathbf{x}} = A^{-1}\,\underline{\mathbf{b}}$]

Tétel

Az ***(1)*** egyenletrendszernek $\exists!$ megoldása $\iff$ $\det A \neq 0$.

Alapkövetelmény a megoldási módszerrel szemben:
Minél kevesebb művelettel legyen meghatározható adott $\varepsilon > 0$ pontosságú közelítő megoldás. (Numerikus módszer hatékonysága.)

Definíció :

Direkt módszer: Pontos adatokat feltételezve, a számításokat pontosan végezve, véges sok művelettel a pontos eredmény meghatározható (A, $\underline{\mathbf{b}}$ - pontos bemenő adatok).
Természetesen a megoldás itt is pontatlan, függ a környezethibától, azaz a géptől és a megoldási módszer numerikus stabilitásától.

2.1. GAUSS - ELIMINÁCIÓ

Legyen

(1) $\quad A\underline{\mathbf{x}} = \underline{\mathbf{b}}$

$\det A \neq 0\,, \quad \underline{\mathbf{b}} \neq \underline{\mathbf{0}}$ $\quad$ ($\exists!$ megoldás), ahol A - kitöltött $N \times N$ - es mátrix.

Legyen $b_i := a_{i,\,N+1}$.

(1)

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1N}x_N &= a_{1,\,N+1}\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2N}x_N &= a_{2,\,N+1}\\
&\vdots\\
a_{N1}x_1 + a_{N2}x_2 + a_{N3}x_3 + \cdots + a_{NN}x_N &= a_{N,\,N+1}
\end{aligned}
$$

<u>Az algoritmus:</u>

1.

Tegyük fel, hogy $a_{11} \neq 0$. Adjuk ***(1)*** i - edik egyenletéhez az első egyenlet $\left(-\dfrac{a_{i1}}{a_{11}}\right)$ - szeresét $\Longrightarrow$

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1N}x_N &= a_{1,\,N+1}\\
\overset{(1)}{a_{22}}x_2 + \overset{(1)}{a_{23}}x_3 + \cdots + \overset{(1)}{a_{2N}}x_N &= \overset{(1)}{a_{2,\,N+1}}\\
&\vdots\\
\overset{(1)}{a_{N2}}x_2 + \overset{(1)}{a_{N3}}x_3 + \cdots + \overset{(1)}{a_{NN}}x_N &= \overset{(1)}{a_{N,\,N+1}}
\end{aligned}
$$

ahol: $\quad \boxed{\overset{(1)}{a_{ij}} = a_{ij} - \dfrac{a_{i1}}{a_{11}}a_{1j}} \qquad \begin{pmatrix} i = 2, 3, \ldots, N \\ j = 2, 3, \ldots, N+1 \end{pmatrix}$

Ha $a_{11} = 0$, akkor sorcserét alkalmazunk!

2.

Tegyük fel, hogy $\overset{(1)}{a}_{22} \neq 0$. Adjuk a második egyenlet $\left(-\dfrac{\overset{(1)}{a}_{i2}}{\overset{(1)}{a}_{22}}\right)$ - szeresét az i - edik egyenletéhez $(i = 3, 4, \ldots, N) \implies$

$$\begin{array}{rcl}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1N}x_N & = & a_{1,\,N+1} \\
\overset{(1)}{a}_{22}x_2 + \overset{(1)}{a}_{23}x_3 + \cdots + \overset{(1)}{a}_{2N}x_N & = & \overset{(1)}{a}_{2,\,N+1} \\
\overset{(2)}{a}_{33}x_3 + \cdots + \overset{(2)}{a}_{3N}x_N & = & \overset{(2)}{a}_{3,\,N+1} \\
 & \vdots & \\
\overset{(2)}{a}_{N3}x_3 + \cdots + \overset{(2)}{a}_{NN}x_N & = & \overset{(2)}{a}_{N,\,N+1}
\end{array}$$

ahol: $\boxed{\overset{(2)}{a}_{ij} = \overset{(1)}{a}_{ij} - \dfrac{\overset{(1)}{a}_{i2}}{\overset{(1)}{a}_{22}} \overset{(1)}{a}_{2j}}$ $\quad \begin{pmatrix} i = 3, 4, \ldots, N \\ j = 3, 4, \ldots, N+1 \end{pmatrix}$

3.

Végül az $(N+1)$ - edik lépés után: $\overset{(N-1)}{a}_{NN}x_N = \overset{N-1)}{a}_{N,N+1}$.

Legyen: $a_{ij} = \overset{(0)}{a}_{ij}$ $\quad \begin{pmatrix} i = 1, 2, \ldots, N \\ j = 1, 2, \ldots, N+1 \end{pmatrix}$.

Ekkor a k - adik lépésben a megváltozott együtthatók kiszámítása a következő módon történik:

(2) $\boxed{\overset{(k)}{a}_{ij} = \overset{(k-1)}{a}_{ij} - \dfrac{\overset{(k-1)}{a}_{ik}}{\overset{(k-1)}{a}_{kk}} \overset{(k-1)}{a}_{kj}}$ $\quad \begin{pmatrix} k = 1, \ldots, N-1 \\ i = k+1, \ldots, N \\ j = k+1, \ldots, N+1 \end{pmatrix}$

Így felső trianguláris mátrixot kapunk, ahol visszahelyettesítéssel az ismeretlenek kiszámolhatók:

(3) $\boxed{x_i = \dfrac{1}{\overset{(i-1)}{a}_{ii}} \left[\overset{(i-1)}{a}_{i,N+1} - \sum_{j=i+1}^{N} \overset{(i-1)}{a}_{ij}\; x_j \right]}$ $\quad (\; i = N, N-1, \ldots, 1 \;)$

2.1.1. A Gauss - elimináció műveletigénye

A ***(2)*** - ből következik, hogy $\forall\ k$ - ra $(N-k)$ osztást és $(N-k)(N-k+1)$ szorzást kell végeznünk, míg a visszahelyettesítésnél $\forall\ i$ - re ***(3)*** - ból $\Longrightarrow$ 1 osztás és $(N-i)$ szorzást kell végezni.

Így a műveletigény:

$$M_G = \sum_{k=1}^{N-1} \left[\ (N-k) + (N-k)\,(N-k+1)\ \right] + \sum_{i=1}^{N} \left[\,1 + (N-i)\ \right]$$

$$\boxed{M_G = \frac{1}{3}N^3 + O(N^2)}$$

A visszahelyettesítés műveletigénye elhanyagolható az eliminációnál!

$$M_G = \sum_{k=1}^{N-1} (N-k)^2 + 2\sum_{k=1}^{N-1} (N-k) + \sum_{i=1}^{N} [\,1 + (N-i)\,] = \frac{(N-1)\,N\,(2N-1)}{6} + N\,(N-1) +$$

$$+ N + \frac{N\,(N-1)}{2} = \frac{2N^3 + 3N^2 - 5N}{6} + \frac{N^2+N}{2} = \frac{N^3+3N^2-N}{3} = \frac{1}{3}N^3 + O(N^2)$$

Tétel

A Gauss - elimináció elvégezhető, azaz $\overset{(k-1)}{a_{kk}} \neq 0, \quad k = 1, \dots N-1 \iff$ az egyenlet mátrixának főminorai determinánsa nem nulla (bal felső sarokaldeterminánsok $\neq 0$).

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix} \neq 0 \qquad\qquad k = 1, \dots, N$$

Bizonyítás

Teljes indukcióval.
Hogy a folyamat ne akadjon el, kell, hogy $\overset{(k-1)}{a_{kk}} \neq 0$.

Az első lépésben $(k = 1)$:

$$\overset{(1)}{a_{22}} = a_{22} - \frac{a_{21}}{a_{11}}\,a_{12} = \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}} = \frac{\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}{\det\,[a_{11}]}, \qquad \text{ha } \det\,[a_{11}] \neq 0$$

Sőt, hogy az eljárás folytatható legyen, kell, hogy $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \neq 0$.

Mivel sorok lineáris kombinációja a determináns értékét nem változtatja meg, azért:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \det \begin{bmatrix} a_{11} & a_{12} \\ 0 & \overset{(1)}{a_{22}} \end{bmatrix} \qquad (\neq 0)$$

Tegyük fel, hogy a módszer a $(k-2)$ - edik lépésig elvégezhető, azaz $(k-1)$ - edrendig a főminorok nem nullák ($\det A_{k-1} \neq 0$).
Belátjuk, hogy a $(k-1)$ - edik lépésben $\overset{(k-1)}{a_{kk}} \neq 0 \iff \det A_k \neq 0$.

Tehát:

$$\det A_k = \det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ a_{21} & \cdots & a_{2k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix} = \det \begin{bmatrix} a_{11} & & \cdots & a_{1k} \\ 0 & \overset{(1)}{a_{22}} & \cdots & \overset{(1)}{a_{2k}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \overset{(k-1)}{a_{kk}} \end{bmatrix} = \prod_{i=1}^{k} \overset{(i-1)}{a_{ii}} = \overset{(k-1)}{a_{kk}} \cdot \prod_{i=1}^{k-1} \overset{(i-1)}{a_{ii}} = \overset{(k-1)}{a_{kk}} \cdot \det A_{k-1}.$$

Az indukciós feltétel miatt $\det A_{k-1} \neq 0$, ezért $\overset{(k-1)}{a_{kk}} \neq 0 \iff \det A_k \neq 0$.

■

2.2. GAUSS - JORDAN ELIMINÁCIÓ

Két fő dologban különbözik a Gauss - eliminációtól:

1. Minden lépésben először leosztjuk az egyenletet a megfelelő $\overset{(k-1)}{a_{kk}}$ együtthatóval.
2. Nem csak a főátló alatt, hanem felette is eliminálunk.

N lépés után:

$$\begin{aligned} x_1 &= \overset{(N)}{a_{1,N+1}} \\ x_2 &= \overset{(N)}{a_{2,N+1}} \\ &\vdots \\ x_n &= \overset{(N)}{a_{N,N+1}} \end{aligned}$$

A k - adik lépésben kapott együtthatók:

$$\textit{(1)} \qquad \boxed{\overset{(k)}{a_{ij}} = \overset{(k-1)}{a_{ij}} - \frac{\overset{(k-1)}{a_{ik}}}{\overset{(k-1)}{a_{kk}}} \overset{(k-1)}{a_{kj}}} \qquad \begin{pmatrix} k = 1, \ldots, N \\ i = 1, \ldots, N, \ i \neq k \\ j = 1, \ldots, N+1 \end{pmatrix}$$

Visszahelyettesíteni nem kell !

2.2.1. A Gauss - Jordan elimináció műveletigénye

(1) $\Longrightarrow$ $\forall\, k$ - ra $(N-k+1)$ osztás és $(N-k+1)(N-1)$ szorzás kell.

Így :

$$M=\sum_{k=1}^{N}\left[\ (N-k+1)+(N-k+1)\,(N-1)\ \right]=N\sum_{k=1}^{N}(N-k+1)=N\frac{(N+1)\,N}{2}=\frac{1}{2}N^3+O(N^2),$$

ami mintegy 50%- kal több a Gauss - eliminációénál.
Ezért lineáris egyenletrendszerek megoldására ezt nem célszerű használni.
DE: Programozási szempontból mátrix invertálásra előnyös.

Alkalmazások

1. Gauss - elimináció: egyenletrendszer megoldására telemátrix és kis rendszer (max 20×20) esetén.

2. Gauss - elimináció: determináns értékének kiszámítására,

$$\det A=(-1)^l\, \overset{(0)}{a_{11}}\ \overset{(1)}{a_{22}}\ \dots\ \overset{(N-1)}{a_{NN}}$$

ahol l - a sor- és oszlopcserék száma.

3. Gauss - Jordan elimináció: mátrix inverzének kiszámítására,

$$\left.\begin{array}{c} A\underline{\mathbf{x}}=\underline{\mathbf{e}}_1 \\ \vdots \\ A\underline{\mathbf{x}}=\underline{\mathbf{e}}_N \end{array}\right\}\quad \text{ahol } \underline{\mathbf{e}}_i \text{ az } i \text{ - edik egységvektor.}$$

$$A\underline{\mathbf{x}}=I\ \Longrightarrow\ \underline{\mathbf{x}}=A^{-1}$$

Belátható, hogy a Gauss - elimináció alkalmazásával kapott felső $\triangle$ mátrix alkalmas mátrix-szorzásokkal is megkapható.

Legyen

$$(A\,|\,\underline{\mathbf{b}}) = \begin{bmatrix} a_{11} & \cdots & a_{1N} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{N1} & \cdots & a_{NN} & b_N \end{bmatrix}$$

Ekkor a Gauss - elimináció lépései:

1.

Azon r_1 index megkeresése, amelyre $a_{r_1 1} \neq 0\,, \qquad r_1 \in \{1, \ldots, N\}$.

Az első sor cseréje az r_1 - edik sorral az $(A\,|\,\underline{\mathbf{b}})$ mátrixban. Az eredménymátrix legyen $(\widehat{A}\,|\,\widehat{\underline{\mathbf{b}}})$.

2.

$\forall \quad i = 2, \ldots, N$ - re vonjuk ki az első sor $\quad l_{i1} = \dfrac{\widehat{a}_{i1}}{\widehat{a}_{11}}$ - szeresét az i - edik sorból $(\widehat{A}\,|\,\widehat{\underline{\mathbf{b}}})$ - ban.

Az eredménymátrix legyen $\qquad (A^{'}\,|\,\underline{\mathbf{b}}^{'})$, $\quad$ azaz

$$(A^{'}\,|\,\underline{\mathbf{b}}^{'}) = \begin{bmatrix} a^{'}_{11} & a^{'}_{12} & \cdots & a^{'}_{1N} & b^{'}_1 \\ 0 & a^{'}_{22} & \cdots & a^{'}_{2N} & b^{'}_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a^{'}_{N2} & \cdots & a^{'}_{NN} & b^{'}_N \end{bmatrix}$$

Az $\quad (A\,|\,\underline{\mathbf{b}}) \quad \Longrightarrow \quad (\widehat{A}\,|\,\widehat{\underline{\mathbf{b}}}) \quad \Longrightarrow \quad (A^{'}\,|\,\underline{\mathbf{b}}^{'})$ transzformációk leírhatók mátrixok szorzásával is a következő módon:

Az **1.** - beli sorcsere az $(A\,|\,\underline{\mathbf{b}})$ mátrixnak egy permutáció mátrixszal való szorzásnak felel meg.

Definíció :

Permutáció mátrix ($N \times N$ - es) :

$$
P_{ij} := \left[\begin{matrix}
1 & 0 & & & & \cdots & & & & & 0 \\
0 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & 0 & & \\
 & & & 0 & & \cdots & & 1 & & & \\
 & & & & 1 & & & & & & \\
\vdots & & & \vdots & & \ddots & & \vdots & & & \vdots \\
 & & & & & & 1 & & & & \\
 & & & 1 & & \cdots & & 0 & & & \\
 & & 0 & & & & & & 1 & & \\
 & & & & & & & & & \ddots & 0 \\
0 & & & & & \cdots & & & & 0 & 1
\end{matrix}\right]
\begin{matrix}
\\ \\ \\ \longleftarrow \; i \text{ - edik sor} \\ \\ \\ \\ \longleftarrow \; j \text{ - edik sor} \\ \\ \\ \\
\end{matrix}
$$

A permutáció mátrix felcseréli az i - edik és j - edik sort.
P_{ij} nem szinguláris, és belátható, hogy $P_{ij}^{-1} = P_{ij}$.

Definíció :

Frobenius mátrix ($N \times N$ - es) :

$$
\begin{array}{c}
\begin{matrix} & & & i \text{ - edik oszlop} & & & \\ & & & \downarrow & & & \end{matrix} \\
L_i := \begin{bmatrix}
1 & 0 & & \cdots & & & 0 \\
0 & & & & & & \\
 & \ddots & \ddots & \ddots & & & \\
\vdots & & 0 & 1 & 0 & & \vdots \\
 & & & -l_{i+1,\,i} & 1 & \ddots & \\
 & & \vdots & \vdots & 0 & \ddots & 0 \\
0 & \cdots & 0 & -l_{N,\,i} & 0 & 0 & 1
\end{bmatrix}
\end{array}
$$

A Frobenius mátrix az I - től csak egy oszlopban különbözik. L_i nem szinguláris.

Így a Gauss - elimináció **1.** lépése a következő módon írható:

$$(A' \,|\, \underline{\mathbf{b}}') = L_1(\widehat{A} \,|\, \widehat{\underline{\mathbf{b}}}) = L_1 P_{r_1 1}(A \,|\, \underline{\mathbf{b}})$$

Állítás

$$L_i^{-1} = \begin{bmatrix} 1 & 0 & & \cdots & & & 0 \\ 0 & & & & & & \\ & \ddots & \ddots & \ddots & & & \\ \vdots & & 0 & 1 & 0 & & \vdots \\ & & & l_{i+1,\,i} & 1 & \ddots & \\ & & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & l_{N,\,i} & 0 & 0 & 1 \end{bmatrix}$$

Bizonyítás

L_i - t írjuk a következő alakban: $L_i = I - m^{(i)} \left(e^{(i)}\right)^T$,

ahol $m^{(i)} := \left[0, \ldots, 0, l_{i+1\,i}, \ldots, l_{Ni}\right]^T$ ($m^{(i)}$ oszlopvektor, $e^{(i)}$ az i - edik egységvektor. $\Longrightarrow$

$$L_i \left(I + m^{(i)} e^{(i)\,T}\right) = \left(I - m^{(i)} e^{(i)\,T}\right) \cdot \left(I + m^{(i)} e^{(i)\,T}\right) = I + m^{(i)} e^{(i)\,T} - m^{(i)} e^{(i)\,T} - m^{(i)} e^{(i)\,T} m^{(i)} e^{(i)\,T} = I$$

∎

A stabilitás miatt célszerű r_1 - et úgy választani, hogy:

$$|a_{r_1 1}| := \max_{1 \le i \le N} |a_{i1}|$$

A k - adik lépés után a következő mátrixot kapjuk:

$$(A^{(k)} \,|\, \underline{\mathbf{b}}^{(k)}) = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} & b_1^{(k)} \\ & & \\ 0 & A_{22}^{(k)} & b_2^{(k)} \end{bmatrix}$$

ahol $A_{11}^{(k)} \in \mathbb{R}^{k \times k}$ egy felső $\triangle$ mátrix.

A $(k+1)$ - edik lépésben:

$\left(A^{(k+1)} \,|\, \underline{\mathbf{b}}^{(k+1)}\right) = L_{k+1} P_{r_k k+1} \left(A^{(k)} \,|\, b^{(k)}\right)$ és most csak az $\left(A_{22}^{(k)} \,|\, b_2^{(k)}\right)$ mátrix fog változni.

$(N-1)$ lépés után egy felső $\triangle$ mátrixot $R \in \mathbb{R}^{N \times N}$ és egy $\underline{\mathbf{c}} \in \mathbb{R}^N$ vektort kapunk, ahol

$$(R \,|\, \underline{\mathbf{c}}) = L_{N-1} P_{N-1} L_{N-2} P_{N-1} \ldots L_1 P_1 (A \,|\, \underline{\mathbf{b}})$$

2.3. *LU* - FELBONTÁS (TRIANGULÁRIS FELBONTÁS)

Legyen

(1) $\quad A\underline{\mathbf{x}} = \underline{\mathbf{b}}\,, \quad A \in \mathbb{R}^{N\times N}\,, \quad \det A \neq 0$

Bontsuk fel az A mátrixot két háromszögmátrix szorzatára!

$A = L \cdot U\,,$ ahol

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{2\,1} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ l_{N\,1} & \cdots & l_{N\,N-1} & 1 \end{bmatrix} \qquad U = \begin{bmatrix} u_{1\,1} & u_{1\,2} & \cdots & u_{1\,N} \\ 0 & u_{2\,2} & & \vdots \\ \vdots & & \ddots & u_{N-1\;N} \\ 0 & \cdots & 0 & u_{N\,N} \end{bmatrix}$$

Ekkor *(1)* megoldása: $LU\underline{\mathbf{x}} = \underline{\mathbf{b}}$

Ezt két lépésben végezhetjük:

$$L \cdot \underbrace{U\underline{\mathbf{x}}}_{\underline{\mathbf{y}}} = \underline{\mathbf{b}}$$

1. $\mathrm{L}\underline{\mathbf{y}} = \underline{\mathbf{b}}$

2. $\mathrm{U}\underline{\mathbf{x}} = \underline{\mathbf{y}}$

Tehát: $A = L \cdot U$ legyen!

Számítsuk ki L és U elemeit!

$$\begin{array}{rc} & \begin{array}{c} \hphantom{xxxxxxxxx} j \text{ - edik oszlop} \\ \hphantom{xxxxxxxxx}\downarrow \end{array} \\ i \text{ - edik sor} \;\longrightarrow & \underbrace{\begin{bmatrix} a_{1\,1} & & \cdots & & a_{1\,N} \\ & \ddots & & & \\ \vdots & & a_{i\,i} \;\cdots & a_{ij} & \vdots \\ & & & \ddots & \\ a_{N\,1} & & \cdots & & a_{N\,N} \end{bmatrix}}_{A} \end{array} =$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & & \cdots & & 0 \\ l_{2\,1} & 1 & & & & \\ \vdots & & \ddots & & & \vdots \\ l_{i\,1} & \cdots & l_{i\,i-1} & 1 & & \\ \vdots & & & & \ddots & \\ l_{N-1\,1} & & & & 1 & 0 \\ l_{N\,1} & & \cdots & & l_{N\,N-1} & 1 \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} u_{1\,1} & u_{1\,2} & \cdots & u_{1j} & \cdots & u_{1\,N} \\ 0 & u_{2\,2} & & & & \\ & & \ddots & \vdots & & \vdots \\ \vdots & & u_{i\,i} & u_{ij} & \cdots & u_{iN\,N} \\ & & & \ddots & & \vdots \\ & & & & u_{N-1\;N-1} & u_{N-1\;N} \\ 0 & & \cdots & & 0 & u_{N\,N} \end{bmatrix}}_{U}$$

(2) $a_{ij} = l_{i1}\, u_{1j} + l_{i2}\, u_{2j} + \cdots + l_{i\,i-1}\, u_{i-1\,j} + u_{ij}\ ,$ ha $i \le j$

(3) $a_{ij} = l_{i1}\, u_{1j} + l_{i2}\, u_{2j} + \cdots + l_{i\,i-1}\, u_{j-1\,j} + l_{ij}\, u_{jj}\ ,$ ha $i > j$

Ahonnan:

$i = 1$ - re: $a_{1j} = u_{1j}$ $j = 1, \ldots, N,$
$j = 1$ - re: $a_{i1} = l_{i1}\, u_{11}$ $i = 2, \ldots, N,$

és amikből u_{1j} , l_{i1} elemek számolhatók, ha $u_{11} \neq 0$.

$i = 2$ - re: $a_{2j} = l_{21}\, u_{1j} + u_{2j}$ $j = 2, \ldots, N,$
$j = 2$ - re: $a_{i2} = l_{i1}\, u_{12} + l_{i2}\, u_{22}$ $i = 3, \ldots, N,$

amikből u_{2j} és l_{i2} elemek számolhatók, ha $u_{22} \neq 0$.

Általánosan :

(2) $\Longrightarrow$ $$\boxed{u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik}\, u_{kj}} \qquad \boxed{i \le j} \quad (j = i, \ldots, N)$$

(3) $\Longrightarrow$ $$\boxed{l_{ij} = \frac{1}{u_{jj}} \left[a_{ij} - \sum_{k=1}^{j-1} l_{ik}\, u_{kj} \right]} \qquad \boxed{i > j} \quad (i = j+1, \ldots, N)$$

Most L és U ismeretében **1.** és **2.** egyenletrendszer megoldása következik, az alábbi formulák alapján:

$L\underline{\mathbf{y}} = \underline{\mathbf{b}}$ megoldása:

$$\boxed{y_i = b_i - \sum_{k=1}^{i-1} l_{ik}\, y_k} \qquad (i = 1. \ldots, N)$$

majd $U\underline{\mathbf{x}} = \underline{\mathbf{y}}$ megoldása:

$$\boxed{x_i = \frac{1}{u_{ii}} \left[y_i - \sum_{k=i+1}^{N} u_{ik}\, x_k \right]} \qquad (i = N, \ldots, 1)$$

Tétel

$$\left.\begin{array}{l}\det A \neq 0 \\ \exists\ LU = A \quad \text{felbontás}\end{array}\right\} \quad \Longrightarrow \quad \exists!\ \text{az } LU \text{ felbontás}$$

Bizonyítás

Indirekt.

$$\begin{array}{rcll} A = L_1\,U_1 & = & L_2\,U_2 & \quad / \cdot L_1^{-1} \\ U_1 & = & L_1^{-1}\,L_2\,U_2 & \quad / \cdot U_2^{-1} \\ U_1\,U_2^{-1} & = & L_1^{-1}\,L_2 & \end{array}$$

bal oldalon felső △ mátrix jobb oldalon alsó △ mátrix

$$\Downarrow$$

Ez csak az egységmátrix lehet,
azaz $L_1 = L_2$ és $U_1 = U_2$.

■

Tétel

Jelölje A_k $(k = 1, \ldots, N)$ az A mátrix k - adrendű főminorait.
Ha $\det A_k \neq 0$, $k = 1, \ldots, N$, akkor $\exists!$ $A = L \cdot U$ felbontás, ahol:

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{2\,1} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ l_{N\,1} & \cdots & l_{N\,N-1} & 1 \end{bmatrix} \quad \text{alsó △ mátrix, és} \qquad U = \begin{bmatrix} u_{1\,1} & u_{1\,2} & \cdots & u_{1\,N} \\ 0 & u_{2\,2} & & \vdots \\ \vdots & & \ddots & u_{N-1\,N} \\ 0 & \cdots & 0 & u_{N\,N} \end{bmatrix} \quad \text{felső △ mátrix.}$$

Bizonyítás

Teljes indukcióval k - ra.

$k = 1$ - re $\quad A = [a_{11}] \quad \Longrightarrow \quad L = [\,1\,]\,,\ U = [u_{11}]\,,\ u_{11} = a_{11}$

Tegyük fel, hogy $(k-1)$ - re a tétel érvényes $\Longrightarrow$ $\exists!$ felbontás, azaz $A_{k-1} = L_{k-1} \cdot U_{k-1}$

Bizonyítjuk k - ra, hogy $\exists!$ $A_k = L_k \cdot U_k$ felbontás.
Ehhez tekintsük az A_k - t !

$$(\star) \qquad A_k = \begin{bmatrix} A_{k-1} & \underline{\mathbf{b}} \\ & \\ \underline{\mathbf{c}}^T & a_{kk} \end{bmatrix}, \quad L_k = \begin{bmatrix} L_{k-1} & 0 \\ & \\ \underline{\mathbf{m}}^T & 1 \end{bmatrix}, \quad U_k = \begin{bmatrix} U_{k-1} & \underline{\mathbf{u}} \\ & \\ 0 & u_{kk} \end{bmatrix},$$

ahol $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$, $\underline{\mathbf{u}}$, $\underline{\mathbf{m}}$ vektorok $(k-1)$ komponensűek. $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ ismert.

A $(\star)$ felírást figyelembe véve a következőknek kell teljesülniük:

$\alpha.$ $\quad L_{k-1}\, U_{k-1} = A_{k-1}$ $\qquad$ $\gamma.$ $\quad \underline{\mathbf{m}}^T\, U_{k-1} = \underline{\mathbf{c}}^T$

$\beta.$ $\quad L_{k-1}\, \underline{\mathbf{u}} = \underline{\mathbf{b}}$ $\qquad$ $\delta.$ $\quad \underline{\mathbf{m}}^T\, \underline{\mathbf{u}} + u_{kk} = a_{kk}$

Az indukciós feltevés miatt L_{k-1} , U_{k-1} egyértelműen meghatározottak (α) .

$\det L_{k-1} \det U_{k-1} = \det A_{k-1} \neq 0$ a feltevés miatt $\Longrightarrow$ $\det L_{k-1} \neq 0$, $\det U_{k-1} \neq 0$,
tehát β, γ - ból $\underline{\mathbf{u}}$ és $\underline{\mathbf{m}}$ ugyancsak egyértelműen meghatározhatók.
Végül δ - ból u_{kk} egyértelműen meghatározható.

$\blacksquare$

2.3.1. Az LU - felbontás műveletigénye

U elemeinek kiszámítása

$$\boxed{u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik}\, u_{kj}} \qquad \boxed{i \leq j} \quad (j = i, \ldots, N)$$

alapján

$$\sum_{j=1}^{N} \left(\sum_{i=1}^{j} (i-1) \right) = \sum_{j=1}^{N} \frac{j\,(j-1)}{2} = \frac{1}{6} N^3 + O\left(N^2\right)$$

L elemeinek kiszámítása

$$\boxed{l_{ij} = \frac{1}{u_{jj}} \left[a_{ij} - \sum_{k=1}^{j-1} l_{ik}\, u_{kj} \right]} \qquad \boxed{i > j} \quad (i = j+1, \ldots, N)$$

alapján

$$\sum_{j=1}^{N} \left(\sum_{i=j+1}^{N} j \right) = \sum_{j=1}^{N} j\,(n-j) = \frac{1}{6} N^3 + O\left(N^2\right)$$

Így a műveletigény :

$$\boxed{M_{LU} = \frac{1}{3}N^3 + O\left(N^2\right)}$$

2.4. SÁVMÁTRIXOK, TRIDIAGONÁLIS MÁTRIXOK

Közönséges és parciális differenciálegyenletek diszkretizációval történő numerikus megoldása gyakran olyan lineáris egyenletrendszerre vezet, melynek mátrixa a fődiagonális egy sávján kívül csupa 0 - kat tartalmaz.

Definíció :

Az $A \in \mathbb{R}^{n\times n}$ mátrixot (m,k) sávmátrixnak vagy szalagmátrixnak nevezzük, ha $a_{ij} = 0$, $j-i > k$ és $i-j > m$ esetén.
Az $(1,1)$ sávmátrixot tridiagonális mátrixnak nevezzük.

Ha egy (m,k) szalagmátrixnak $\exists$ trianguláris felbontása, $A = LU$, akkor L $(m,0)$ sávú és U $(0,k)$ sávú mátrix lesz, ha sorcserére nincs szükség.
Tegyük fel, hogy meg akarjuk oldani az $A\underline{\mathbf{x}} = \underline{\mathbf{d}}$, $A \in \mathbb{C}^{n\times n}$, $\underline{\mathbf{d}} \in \mathbb{C}^n$ egyenletrendszert, ahol A tridiagonális és

$$\textit{(1)} \qquad A = \begin{bmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & b_{n-1} & a_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & b_n & a_n \end{bmatrix} = \text{tridiag}\,(b_i,\, a_i,\, c_i)$$

Tétel **Felbontási tétel tridiagonális mátrixokra**

Tegyük fel, hogy az *(1)* mátrix elemeire teljesülnek a következő egyenlőtlenségek:

$$(\star) \qquad \begin{cases} |a_1| > |c_1| > 0, & \\ |a_i| \geq |b_i| + |c_i|\,, & b_i \neq 0, \quad c_i \neq 0, \quad 2 \leq i \leq n-1 \\ |a_n| \geq |b_n| > 0. & \end{cases}$$

Ekkor :

i.)

$$\alpha_1 := a_1, \qquad \gamma_1 := c_1\, \alpha_1^{-1} \ ,$$

$$\alpha_i := \alpha_i - b_i\, \gamma_{i-1} \qquad (2 \le i \le n) \ ,$$

$$\gamma_i := c_i\, \alpha_i^{-1} \qquad (2 \le i \le n-1) \qquad \text{és teljesül, hogy:}$$

$$|\gamma_i| < 1, \quad (1 \le i \le n-1), \qquad 0 < |a_i| - |b_i| < |\alpha_i| < |a_i| + |b_i| \, , \quad (2 \le i \le n)$$

ii.)

A - nak $\exists$ $A = LU$ felbontása, ahol:

$L :=$ tridiag $(b_i,\, \alpha_i,\, 0)$, $\qquad R :=$ tridiag $(0,\, 1,\, \gamma_i)$

iii.)

A nem szinguláris.

Bizonyítás

i.)

A *(⋆)* - ból azonnal következik, hogy:

$$|\gamma_q| = |c_1| \cdot |a_1|^{-1} \le 1$$

Most legyen:

$$|\gamma_j| < 1, \qquad j = 1, \ldots, i-1$$

Ekkor:

$$|\gamma_i| = \left| \frac{c_i}{a_i - b_i\, \gamma_{i-1}} \right| \le \frac{|c_i|}{\big|\, |a_i| - |b_i|\, |\gamma_{i-1}| \,\big|} < \frac{|c_i|}{|a_i| - |b_i|} \le 1$$

Továbbá:

$$|a_i| + |b_i| > |a_i| + |b_i|\, |\gamma_{i-1}| \ge |\alpha_i| \ge |a_i| - |b_i|\, |\gamma_{i-1}| > |a_i| - |b_i| \ge |c_i| > 0$$

ii.)

Közvetlenül ellenőrizhető, hogy:

$A = \text{tridiag}\ (b_i,\ \alpha_i,\ 0) \cdot \text{tridiag}\ (0,\ 1,\ \gamma_i)$ érvényes, sor - oszlop szorzással:

$$\begin{array}{ll} a_{i\,i+1} = \alpha_i\, \gamma_i = \alpha_i \left(c_i\, \alpha_i^{-1}\right) = c_i, & 1 < i < n-1\,, \\ a_{i\,i} = b_i\, \gamma_{i-1} + \alpha_i = b_i\, \gamma_{i-1} + (\alpha_i - b_i\, \gamma_{i-1}) = a_i, & 2 \le i \le n\,, \\ a_{i+1\,i} = b_{i+1}, & 1 \le i \le n-1\,, \\ a_{11} = \alpha_1 = a_1. & \end{array}$$

iii.)

Az A mátrix reguláris, mivel:

$$\det A = \det L \cdot \det U = \prod_{i=1}^{n} \alpha_i \neq 0$$

■

Definíció :

A *(⋆)* tulajdonságú tridiagonális mátrixokat irreducibilisen diagonálisan domináns mátrixoknak nevezzük.

A fenti tétel tehát úgy is megfogalmazható, hogy egy irreducibilisen diagonálisan domináns A mátrixnak $\exists\ A = LU$ felbontása, ahol L $(1,0)$ - sávú, U $(0,1)$ - sávú (szalag)mátrix és ahol U fődiagonálisában 1 - esek állnak.

2.5. SCHUR - KOMPLEMENTER

A Gauss - algoritmusnál az 1. oszlop nullázása még a következő módon is leírható:

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$
$$L_1\, A\underline{\mathbf{x}} = L_1\underline{\mathbf{b}}\ , \quad \text{ahol:}$$

(1) $$L_1 := \begin{bmatrix} 1 & 0 & & \cdots & 0 \\ -l_{21} & 1 & 0 & & \vdots \\ \vdots & 0 & 1 & \ddots & \\ & \vdots & \ddots & \ddots & 0 \\ -l_{N1} & 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ & \\ -\underline{l}_1 & I_{N-1} \end{bmatrix}, \qquad A = \begin{bmatrix} a_{11} & \underline{\mathbf{a}}_1 \\ & \\ \underline{\mathbf{a}}_1 & A_{N-1\,N-1} \end{bmatrix}, \qquad \text{ahol:}$$

$$\underline{\mathbf{l}}_1 = \frac{1}{a_{11}}\,\overline{\mathbf{a}}_1\,, \qquad \overline{\mathbf{a}}_1 = [a_{21}, \ldots, a_{N1}]^T\,, \qquad \underline{\mathbf{a}}_1 = [a_{12}, \ldots, a_{1N}]$$

Így :

$$\textit{(2)} \qquad L_1 A = \begin{bmatrix} a_{11} & \underline{\mathbf{a}}_1 \\ & \\ 0 & A^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ & \\ 0 & A^{(1)} \end{bmatrix} \cdot \begin{bmatrix} 1 & \underline{\mathbf{u}}_1 \\ & \\ 0 & I_{N-1} \end{bmatrix}$$

Ilyen szorzatalakra bontható, ahol $\underline{\mathbf{u}}_1 := \dfrac{1}{a_{11}} \cdot \underline{\mathbf{a}}_1$, és $A^{(1)}$ a jobb alsó $(N-1) \times (N-1)$ - es mátrix az 1. oszlop nullázása után.

Ha $A_{N-1,N-1}$ az eredeti A mátrix jobb alsó sarka, akkor:

(1) $\Longrightarrow$ $\quad A^{(1)} = A_{N-1,N-1} - \underline{\mathbf{l}}_1 \underline{\mathbf{a}}_1 = A_{N-1,N-1} - \dfrac{1}{a_{11}} \overline{\mathbf{a}}_1 \underline{\mathbf{a}}_1$

(2) $\Longrightarrow$ tudva, hogy L_1 reguláris mátrix és inverze:
$$L_1^{-1} = \begin{bmatrix} 1 & 0 & & \cdots & 0 \\ l_{21} & 1 & 0 & & \vdots \\ \vdots & 0 & 1 & \ddots & \\ & \vdots & \ddots & \ddots & 0 \\ l_{N1} & 0 & \cdots & 0 & 1 \end{bmatrix}$$

$$\textit{(3)} \quad \Longrightarrow \quad A = \begin{bmatrix} 1 & 0 \\ & \\ \underline{\mathbf{l}}_1 & I_{N-1} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & 0 \\ & \\ 0 & A^{(1)} \end{bmatrix} \cdot \begin{bmatrix} 1 & \underline{\mathbf{u}}_1 \\ & \\ 0 & I_{N-1} \end{bmatrix}$$

Az A mátrix e felbontása három mátrix szorzatára tükrözi a Gauss - elimináció első lépését.
A k - adik lépésben $(N-k) \times (N-k)$ méretű $A^{(k)}$ mátrixot, és $\underline{\mathbf{u}}_k$, $\underline{\mathbf{l}}_k$ vektorokat állítunk elő.

Állítás

Az L_i Frobenius mátrix inverze:
$$L_i^{-1} = \begin{bmatrix} 1 & 0 & & \cdots & & & 0 \\ 0 & & & & & & \\ & \ddots & \ddots & \ddots & & & \\ \vdots & & 0 & 1 & 0 & & \vdots \\ & & & l_{i+1,\, i} & 1 & \ddots & \\ & & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & l_{N,\, i} & 0 & 0 & 1 \end{bmatrix}$$

Ugyanis ha $L_i = I - m^{(i)}\, e^{(i)\,T}$ $\quad \Longrightarrow$

$\Longrightarrow \qquad m := (0, \ldots, l_{i+1\,i}, \ldots, l_{N\,i})^T \qquad (I - m^{(i)}\, e^{(i)\,T})(I + m^{(i)}\, e^{(i)\,T}) = I$

Most (elimináljuk) nullázzuk ki az első két oszlopot az első lépésben.

Legyen

$$A = \begin{bmatrix} A_{11} & A_{12} \\ \\ A_{21} & A_{22} \end{bmatrix}, \qquad \underline{\mathbf{b}} = \begin{bmatrix} b_1 \\ \\ b_2 \end{bmatrix}, \qquad \text{és} \qquad A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ \\ a_{21} & a_{22} \end{bmatrix},$$

és ahol A_{22} $(N-2) \times (N-2)$ - es mátrix.

Általánosítva a fenti eljárást:

$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ - t szorozzuk balról az $L_2 = \begin{bmatrix} I_2 & 0 \\ \\ -l_2 & I_{N-2} \end{bmatrix}$ alsó $\triangle$ blokkmátrixszal.

Tegyük fel, hogy A_{11} reguláris mátrix. Ekkor:

$$(\star) \qquad L_2 A\underline{\mathbf{x}} = \begin{bmatrix} I_2 & 0 \\ \\ -l_2 & I_{N-2} \end{bmatrix} \underline{\mathbf{x}} = \begin{bmatrix} A_{11} & A_{12} \\ \\ \underline{\mathbf{0}} & A^{(2)} \end{bmatrix} \underline{\mathbf{x}} = \begin{bmatrix} b_1 \\ \\ \overline{b}_2 \end{bmatrix} \qquad (= L_2\underline{\mathbf{b}})$$

ahonnan $l_2 = A_{21} A_{11}^{-1}$ $(N-2) \times 2$ - es mátrix, továbbá:

$$(4) \qquad \boxed{A^{(2)} = A_{22} \cdot -A_{21} \cdot A_{11}^{-1} \cdot A_{12}} \qquad \text{ugyanis:} \qquad -l_2 A_{12} + I_{N-2} A_{22} = A^{(2)}$$

$$\overline{b}_2 = -A_{21} A_{11}^{-1} b_1 + b_2 \qquad \text{ugyanis:} \qquad L_2 \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ \overline{b}_2 \end{bmatrix}$$

Megjegyzés :

A_{11}^{-1} létezhet akkor is, ha $a_{11} = 0$, tehát ha a korábbi elimináció nem működik (csak sorcserével).

Így a *($\star$)* - beli $A^{(2)}$ - vel kapcsolatos egyenletrendszerrel kell tovább foglalkozni, ugyanis ha ennek megoldása:

$y := [x_3, \ldots, x_N]^T$, akkor:

$$(5) \qquad A_{11} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \overset{(\star)}{=} b_1 - A_{12}\, y \quad \text{egyenlet megoldása adja a hiányzó } x_1,\, x_2 \text{ - t.}$$

A *($\star$)* - beli $L_2 A$ felső $\triangle$ blokkmátrix, mint *(2)* - nél, szorzattá bontható.

$$(6) \qquad L_2 A = \begin{bmatrix} A_{11} & A_{12} \\ \\ 0 & A^{(2)} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ \\ 0 & A^{(2)} \end{bmatrix} \cdot \begin{bmatrix} I_2 & U_2 \\ \\ 0 & I_{N-2} \end{bmatrix},$$

ahol $U_2 = A_{11}^{-1} A_{12}$, ha A_{11} reguláris.

Belátható közvetlenül, hogy $L_2^{-1} = \begin{bmatrix} I_2 & 0 \\ l_2 & I_{N-2} \end{bmatrix}$, ezért ***(6)*** $\Longrightarrow$

(7) $$A = \begin{bmatrix} I_2 & 0 \\ l_2 & I_{N-2} \end{bmatrix} \cdot \begin{bmatrix} A_{11} & 0 \\ 0 & A^{(2)} \end{bmatrix} \cdot \begin{bmatrix} I_2 & U_2 \\ 0 & I_{N-2} \end{bmatrix}$$

Ennek és a ***(3)*** felbontásnak az általánosítása a következő:

(8) $$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A = \begin{bmatrix} I_k & 0 \\ L_k & I_{N-k} \end{bmatrix} \cdot \begin{bmatrix} D_k & 0 \\ 0 & D_{N-k} \end{bmatrix} \cdot \begin{bmatrix} I_k & U_k \\ 0 & I_{N-k} \end{bmatrix},$$

ahol:

$$\left.\begin{array}{ll} A_{11},\, D_k \in \mathbb{R}^{k\times k}, & A_{12},\, U_k \in \mathbb{R}^{k\times(N-k)} \\ A_{21},\, L_k \in \mathbb{R}^{(N-k)\times k}, & A_{22},\, D_{N-k} \in \mathbb{R}^{(N-k)\times(N-k)} \end{array}\right\} \quad 1 \le k \le N-1 \,.$$

Állítás

Ha $\det A_{11} \neq 0$, akkor a ***(8)*** blokkfelbontás létezik.

Bizonyítás

A ***(8)*** jobb oldalát kiszámítva $\Longrightarrow$ $A = \begin{bmatrix} D_k & D_k U_k \\ L_k D_k & D_{N-k} + L_k D_k U_k \end{bmatrix}$

Innen $\Longrightarrow$ $A_{11} = D_k, \quad A_{12} = D_k U_k, \quad A_{21} = L_k D_k, \quad A_{22} = D_{N-k} + L_k D_k U_k$

Mivel $\det A \neq 0$, azért $D_k,\, U_k,\, L_k$ és D_{N-k} egyértelműen meghatározottak, és

$$D_k = A_{11}, \quad U_k = A_{11}^{-1} A_{12}, \quad L_k = A_{21} A_{11}^{-1}\,,$$

$$\boxed{D_{N-k} = A_{22} \cdot - A_{21} \cdot A_{11}^{-1} \cdot A_{12}}$$

■

D_{N-k} az az almátrix, amely esetleg még további feldolgozásra vár.
Így a Gauss - elimináció 1. lépése után: $D_{N-k} = A^{(1)}$, illetve a blokklépés (x_1, x_2 egyidejű eliminációja) után: $D_{N-k} = A^{(2)}$.

Definíció :

A D_{N-k} mátrixot a ***(8)*** alakú blokkmátrix A_{11} blokkja Schur - komplementerének nevezzük.

Jele: $D_{N-k} = A_{22} \cdot - A_{21} \cdot A_{11}^{-1} \cdot A_{12} =: [A/A_{11}]$ (A létezését beláttuk, ha $\det A_{11} \neq 0$.)

2.6. PASSAGE (PROGONKA) MÓDSZER VAGY RÖVIDÍTETT GAUSS - ALGORITMUS TRIDIAGONÁLIS MÁTRIXÚ EGYENLETRENDSZERRE

Tekintsük a következő egyenleterndszert:

(1) $A\underline{\mathbf{x}} = \underline{\mathbf{f}}$, $A \in \mathbb{R}^{n\times n}$, $\underline{\mathbf{x}}\, , \underline{\mathbf{f}} \in \mathbb{R}^n$, ahol

$$A = \begin{bmatrix} a_1 & -c_1 & 0 & \cdots & 0 \\ -b_2 & a_2 & -c_2 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & -b_{n-1} & a_{n-1} & -c_{n-1} \\ 0 & \cdots & 0 & -b_n & a_n \end{bmatrix} = \text{tridiag}\,(-b_i,\, a_i,\, -c_i)$$

(2) $$\begin{cases} b_i,\, a_i,\, c_i > 0, \quad \forall i \text{ - re} \\ a_i \geq b_i + c_i, \quad \forall i \text{ - re} \\ \text{és } \exists\, j \text{ index, amelyre } a_j > b_j + c_j \end{cases}$$

Továbbá $b_1 := 0$, $c_n := 0$.

Az LU felbontásban a sávszélesség megmarad, így mindkét mátrixnak csak két (nem zérus) átlója van.

Így az ***(1)*** megoldását a következő alakban kereshetjük, ami a visszahelyettesítésnek felel meg:

(3) $x_{i-1} = \alpha_i x_i + \beta_i, \quad i = n, \ldots, 2$,

ahol az α_i, β_i együtthatókat az ***(1)*** i - edik sorából határozhatjuk meg:

$$-b_i x_{i-1} + a_i x_i - c_i x_{i+1} = f_i$$
$$-b_i\,(\alpha_i x_i + \beta_i) + a_i x_i - c_i x_{i+1} = f_i$$

azaz:

$$x_i\,(a_i - b_i\alpha_i) = c_i x_{i+1} + f_i + b_i\beta_i$$

Tegyük fel, hogy $(a_i - b_i\alpha_i) \neq 0$.

Ekkor:

(⋆) $$x_i = \frac{c_i}{a_i - b_i\alpha_i}\, x_{i+1} + \frac{f_i + b_i\beta_i}{a_i - b_i\alpha_i}, \text{ ahol}$$

(4) $$\alpha_{i+1} := \frac{c_i}{a_i - b_i\alpha_i} \qquad \text{és} \qquad \beta_{i+1} := \frac{f_i + b_i\beta_i}{a_i - b_i\alpha_i} \qquad i = 1, \ldots, n-1$$

α_1, β_1 kezdőértékek meghatározásához írjuk fel ***(1)*** első egyenletét:

$$a_i x_i - c_i x_2 = f_1 \qquad \Longrightarrow \qquad a_1 x_1 = c_1 x_2 + f_1$$

(⋆) alakra rendezve

$$x_1 = \frac{c_1}{a_1}\, x_2 + \frac{f_1}{a_1}$$

azt mutatja, hogy:

(5) $\alpha_1 = 0$, $\beta_1 = 0$ választható.

A ***(4)*** és ***(5)*** - ből minden α_i és β_i megkapható.

A ***(3)*** kiszámításához szükség van az x_n ,,kezdőértékre”. Ehhez tekintsük az ***(1)*** utolsó egyenletét:

$$-b_n x_{n-1} + a_n x_n = f_n$$

(3) $$\Longrightarrow \qquad -b_n\left(\alpha_n x_n + \beta_n\right) = f_n + b_n\beta_n$$

(6) $$\Longrightarrow \qquad x_n = \frac{f_n + b_n\beta_n}{a_n - b_n\alpha_n} =: \beta_{n+1}$$

Ebből indítva a visszahelyettesítési menetet, ***(3)*** alapján $x_{n-1}, \ldots, x_1$ megkapható.
A ***(3)*** - ***(6)*** algoritmust rövidített Gauss - eliminációnak vagy tridiagonális algoritmusnak is nevezik. (Progonka vagy passage módszer.)

2.6.1. Az algoritmus stabilitása

Tétel

A passage algoritmus ***(3)*** - ***(6)*** formulái jól definiáltak a ***(2)*** feltételek esetén, azaz $a_i - b_i\alpha_i \neq 0 \quad i = 1, \ldots, n-1$ és stabilak abban az értelemben, hogy $0 \leq \alpha_i \leq 1, \quad i = 1, \ldots, n$.

Bizonyítás

Teljes indukcióval i - re.

Tudjuk, hogy $i = 1$ - re $\alpha_i = 0 \quad \Longrightarrow \quad a_1 - b_1\alpha_1 > 0$

Továbbá:

$$0 \leq \alpha_2 = \frac{c_1}{a_1 - b_1\alpha_1} \leq 1$$

Most legyen $0 \leq \alpha_i \leq 1$ rögzített i - re. Ekkor:

(2) $\quad \Longrightarrow \quad a_i - b_i\alpha_i \geq (1 - \alpha_i)\, b_i + c_i > 0,$

és így $\quad 0 < \alpha_{i+1} = \dfrac{c_i}{a_i - b_i\alpha_i} \leq 1$

$$a_i - b_i \geq c_i, \qquad \frac{1}{c_i} \geq \frac{1}{a_i - b_i} \qquad \Longrightarrow \qquad 1 \geq \frac{c_i}{a_i - b_i} \geq \frac{c_i}{a_i - b_i\alpha_i}$$

Tehát $\quad 0 \leq \alpha_i \leq 1 \quad \forall\, i$ - re.

Most még esetleg β_{n+1} - et nem tudnánk kiszámolni, ha $\alpha_n = 1 \quad a_n = b_n$ lenne.

De *(2)* alapján létezik j index, amelyre $a_j > b_j + c_j$, ami miatt $i = j+1, \ldots, n$ - re teljesül,

hogy $\quad 0 < \alpha_i < 1$.

Tehát $\alpha_n = 1$ nem állhat fenn.

■

2.7. SZIMMETRIKUS POZITÍV DEFINIT MÁTRIXOK CHOLESKY FELBONTÁSA

Legyen $\mathbb{R}^n$ euklideszi térben a skaláris szorzat:

$$(\underline{\mathbf{x}}, \underline{\mathbf{y}}) := \sum_{i=1}^{n} x_i\, y_i, \qquad \underline{\mathbf{x}},\ \underline{\mathbf{y}} \in \mathbb{R}^n \qquad (\underline{\mathbf{x}}, \underline{\mathbf{y}}) = \underline{\mathbf{x}}^T \underline{\mathbf{y}}$$

Állítás

$$A \in \mathbb{R}^{n\times n} \qquad \Longrightarrow \qquad (A\underline{\mathbf{x}}, \underline{\mathbf{y}}) = (\underline{\mathbf{x}}, A^T\underline{\mathbf{y}})$$

Definíció :

Az A mátrix szimmetrikus, ha $A = A^T$, azaz, ha $a_{ij} = a_{j\,i} \quad \forall\ i,\ j$ - re.

Megjegyzés :

A szimmetrikus $\iff$ $(A\underline{\mathbf{x}},\ \underline{\mathbf{y}}) = (\underline{\mathbf{x}},\ A\underline{\mathbf{y}})$

Ha A szimmetrikus $\implies$ A sajátértékei valós számok.

Definíció :

Az A szimmetrikus mátrix pozitív definit, ha $(A\underline{\mathbf{x}},\ \underline{\mathbf{x}}) > 0, \quad \forall\ \underline{\mathbf{x}} \neq \underline{\mathbf{0}}$.

Az A szimmetrikus mátrix pozitív szemidefinit, ha $(A\underline{\mathbf{x}},\ \underline{\mathbf{x}}) \geq 0, \quad \forall\ \underline{\mathbf{x}} \neq \underline{\mathbf{0}}$.

Tétel

Ha A pozitív definit $\implies$ $\det A \neq 0$, sőt: $\det A > 0$.

Bizonyítás

Indirekt.

Tegyük fel, hogy $\det A = 0 \implies \exists\ \underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, amelyre $A\underline{\mathbf{x}} = \underline{\mathbf{0}} \implies$
$(A\underline{\mathbf{x}},\ \underline{\mathbf{x}}) = (\underline{\mathbf{0}},\ \underline{\mathbf{x}}) = \underline{\mathbf{0}}$ lenne $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ vektorra, ami ellentmond annak, hogy A pozitív definit.

■

Állítás

Ha A pozitív definit $\implies$ $\exists\ A^{-1}$, ez is pozitív definit, sőt A $\forall$ főminora pozitív definit, és $a_{ii} > 0$.

Tétel **Kritériumok A pozitív definitségére**

Az A mátrix pozitív definit $\iff$ ha valamelyik az alábbiak közül teljesül:

1. $(A\underline{\mathbf{x}},\ \underline{\mathbf{x}}) \geq 0\,, \quad \forall \quad \underline{\mathbf{x}} \neq \underline{\mathbf{0}}$,

2. Az A mátrix sajátértékei pozitívak,

3. $\det A_k > 0$, ahol A_k főminora A - nak, $\quad k = 1, \ldots, n$.

Tétel

Ha $A = W \cdot W^T$ alakban írható $\Longrightarrow$ A szimmetrikus mátrix és pozitív definit.

Bizonyítás

Szimmetria :

$$A^T = \left(W \cdot W^T\right)^T = \left(W^T\right)^T \cdot W^T = W \cdot W^T = A$$

Pozitív definit :

$$(A\underline{\mathbf{x}}, \underline{\mathbf{x}}) = \left(W \cdot W^T\underline{\mathbf{x}}, \underline{\mathbf{x}}\right) = \left(W^T\underline{\mathbf{x}}, W^T\underline{\mathbf{x}}\right) = \|W^T\underline{\mathbf{x}}\|_2 > 0, \quad \text{ha} \quad \underline{\mathbf{x}} \neq \underline{\mathbf{0}}$$

■

Fontos kérdés : Milyen feltételek mellett létezik egy A mátrixnak $A = W \cdot W^T$ felbontása?

Definíció :

Az $A = W \cdot W^T$ felbontást, ahol W nem szinguláris mátrix, Cholesky felbontásnak nevezzük.

Belátjuk, hogy ha A szimmetrikus és pozitív definit mátrix, akkor a Cholesky felbontása létezik és a W trianguláris mátrix lehet.

Tétel

Legyen $A \in \mathbb{R}^{n \times n}$ szimmetrikus pozitív definit mátrix. Ekkor $\exists!$ $A = L \cdot L^T$ felbontás, ahol L alsó $\triangle$ mátrix, nem szinguláris és $l_{ii} > 0$ $(i = 1, \dots, n)$.

Bizonyítás

Teljes indukcióval n - re.

$n = 1$ - re $A = (a_{11})$, $a_{11} > 0$, így $L = L^T = \sqrt{a_{11}}$

Legyen $A \in \mathbb{R}^{n \times n}$ szimmetrikus pozitív definit és tegyük fel, hogy $(n-1)$ - re igaz az állítás,

tehát $A_{n-1} \in \mathbb{R}^{(n-1)\times(n-1)}$ - re $\exists!$ $A_{n-1} = L_{n-1} \cdot L_{n-1}^T$.

Tekintsük A_n - t:

$$A_n = \begin{bmatrix} A_{n-1} & \underline{\mathbf{b}} \\ \\ \underline{\mathbf{b}}^T & a_{nn} \end{bmatrix} \qquad \text{és keressük } L_n \text{ - t a következő alakban:} \qquad L_n = \begin{bmatrix} L_{n-1} & \underline{\mathbf{0}} \\ \\ \underline{\mathbf{c}}^T & \alpha \end{bmatrix}, \quad \text{ahol}$$

A_{n-1} főminor pozitív definit, $a_{nn} > 0$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}} \in \mathbb{R}^{n-1}$, $\alpha > 0$, és $l_{ii} > 0$ $i = 1, \dots, n-1$ (ezek L_{n-1} főátlójának elemei).

Ezt figyelembe véve:

$$\begin{bmatrix} A_{n-1} & \underline{\mathbf{b}} \\ \\ \underline{\mathbf{b}}^T & a_{nn} \end{bmatrix} = A_n = \begin{bmatrix} L_{n-1} & \underline{\mathbf{0}} \\ \\ \underline{\mathbf{c}}^T & \alpha \end{bmatrix} \cdot \begin{bmatrix} L_{n-1}^T & \underline{\mathbf{c}} \\ \\ 0 & \alpha \end{bmatrix} \qquad \Longrightarrow$$

$$A_{n-1} = L_{n-1} \cdot L_{n-1}^T \qquad \text{Az indukciós feltétel miatt } \exists\, L_{n-1} \text{ alsó } \triangle \text{ mátrix.}$$

$$\Longrightarrow \quad \begin{array}{l} L_{n-1} \cdot \underline{\mathbf{c}} = \underline{\mathbf{b}} \\ \\ \underline{\mathbf{b}}^T = \underline{\mathbf{c}}^T \cdot L_{n-1}^T \end{array} \qquad \Longrightarrow \qquad \det L_{n-1} \neq 0 \qquad \Longrightarrow \qquad \exists!\, \underline{\mathbf{c}} \quad \text{megoldás.}$$

$$a_{nn} = \underline{\mathbf{c}}^T\, \underline{\mathbf{c}} + \alpha^2$$

Mivel $\quad 0 < \det A_n = \det L_n \cdot \det L_n^T = \alpha \left(\det L_{n-1}\right) \cdot \alpha \left(\det L_{n-1}^T\right) = \alpha^2 \left(\det L_{n-1}\right)^2 \qquad \Longrightarrow$

$$\Longrightarrow \qquad \alpha^2 > 0 \quad \text{és} \quad \alpha^2 \in \mathbb{R}$$

Így $\quad \alpha^2 = a_{nn} - \underline{\mathbf{c}}^T \underline{\mathbf{c}} \qquad \Longrightarrow \qquad \boxed{\alpha = \sqrt{a_{nn} - \underline{\mathbf{c}}^T \underline{\mathbf{c}}}} \qquad \Longrightarrow \qquad \alpha > 0 \quad \text{és} \quad \alpha \in \mathbb{R}$

■

2.7.1. Az L_n mátrix l_{ij} elemeinek meghatározása

$A = L \cdot L^T$ - ból oszloponként haladva elég $i \geq j$ - re az L - beli elemeket meghatározni.

$$i \geq j \qquad a_{ij} = \sum_{k=1}^{j} l_{ik}\, l_{jk} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad \boxed{i = j} \qquad a_{jj} = \sum_{k=1}^{j} l_{jk}^2 \qquad \Longrightarrow \qquad (\star) \qquad \boxed{l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2\right)^{\frac{1}{2}}} \qquad > 0$$

$$(\star\star) \qquad \boxed{i > j} \qquad \boxed{l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}\, l_{jk}\right)}$$

A műveletigény : $\frac{1}{6}n^3 + O(n^2)$.

Állítás

Ha A szimmetrikus de nem pozitív definit, a Cholesky felbontás nem mindig lehetséges. Viszont, ha létezik a Cholesky felbontás, akkor az L mátrix nem lehet valós, ugyanis valós, nem szinguláris L esetén, ha $\exists\ A = L \cdot L^T \quad \Longrightarrow \quad A$ pozitív definit.

2.7.2. Egyenletrendszer megoldása Cholesky felbontással szimmetrikus mátrixra

(1) $\quad A\underline{\mathbf{x}} = \underline{\mathbf{b}}\,, \qquad A$ szimmetrikus

$\quad LL^T\underline{\mathbf{x}} = \underline{\mathbf{b}}\,, \qquad$ azaz

(2) $\quad L\underline{\mathbf{y}} = \underline{\mathbf{b}}\,, \qquad$ majd

(3) $\quad L^T\underline{\mathbf{x}} = \underline{\mathbf{y}}$.

A *(2)* j - edik egyenlete a visszahelyettesítés után:

(4) $$\boxed{y_j = \frac{1}{l_{jj}}\left(a_{j\,n+1} - \sum_{k=1}^{j-1} l_{jk}\, y_k\right)} \qquad \bigl(j = 1, \ldots, N\bigr)\ ,$$

ahol $\quad a_{j\,n+1} := b_j \qquad (j = 1, \ldots, n)$.

Majd bevezetve az $\quad y_j := l_{j\,n+1}$, *(4)* a következő alakú lesz:

$$L_{j\,n+1} = \frac{1}{l_{jj}}\left(a_{j\,n+1} - \sum_{k=1}^{j-1} l_{kj}\, l_{k\,n+1}\right) \qquad j = 1, \ldots n\ ,$$

ami éppen a *(⋆)* képletnek az $\boxed{i = n+1}$ - nek megfelelő speciális esete.

A *(3)* - at most már a szokásos visszahelyettesítéssel adjuk meg:

$$\boxed{x_i = \frac{1}{l_{ii}}\left(y_i - \sum_{k=i+1}^{N} l_{ki}\, x_k\right)} \qquad \bigl(i = N, \ldots, 1\bigr)\ .$$

A műveletigény :

$$\frac{1}{6}N^3 + O(N^2)\ .$$

Fele, mint az LU felbontásnál, de N darab gyökvonás is kell.

Megjegyzés :

A *(⋆⋆)* formulából $\Longrightarrow$

$$\sum_{k=1}^{j} l_{jk}^2 \le \max_{1\le j\le N} |a_{jj}|\ , \qquad \forall \quad 1 \le j \le N \text{ - re.}$$

Így az L mátrix minden eleme korlátos, a $\max\limits_{1\le j\le N} \sqrt{|a_{jj}|}$ korláttal, tehát nem nőhetnek túlságosan gyorsan. Ez a tulajdonság biztosítja a módszer (algoritmus) stabilitását.

Az l_{jk} számok az alatt a korlát alatt maradnak, amit a kiindulási adatok (az A mátrix elemei) meghatároznak!

2.8. HOUSEHOLDER (QR) FELBONTÁS

(1) $\quad A\underline{\mathbf{x}} = \underline{\mathbf{b}} \qquad A \in \mathbb{R}^{n\times n}$

Az A mátrixot felbontjuk $A = QR$ alakra, ahol Q ortogonális mátrix, R felső $\triangle$ mátrix.

Definíció :

Q mátrtix ortogonális, ha $\quad QQ^T = I$

Q tulajdonságai: $\qquad Q^T = Q^{-1} \qquad \det Q = \pm 1$

Ha $Q_1\ ,\ Q_2$ ortogonális $\qquad \Longrightarrow \qquad Q_1 \cdot Q_2$ is ortogonális

Láttuk már, hogy például a Gauss - eliminációnál a mátrixot felső $\triangle$ mátrixra lehet transzformálni ,,elemi” mátrixok (Frobenius, permutáció) szorzásával. Most is a Q ortogonális mátrixot úgynevezett Householder - mátrixok szorzásával állítjuk elő.

Ha $\quad A = QR \quad \overset{(1)}{\Longrightarrow} \quad QR\underline{\mathbf{x}} = \underline{\mathbf{b}}$ $\quad$ ezt balról Q^{-1} - gyel szorozva $\quad \Longrightarrow$

$$\Longrightarrow \qquad R\underline{\mathbf{x}} = Q^T\underline{\mathbf{b}} \qquad \text{felső } \triangle \text{ mátrixú rendszer.}$$

Definíció :

Householder - mátrix : $H \in \mathbb{R}^{k\times k}$, $\quad H = I - 2\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T$,

ahol $\underline{\mathbf{h}} \in \mathbb{R}^k$, $\quad \underline{\mathbf{h}} := [0,\dots,0,h_i,\dots,h_k]^T$, $\quad \|\,\underline{\mathbf{h}}\,\|_2 = 1$, $\quad \big(\underline{\mathbf{h}}^T, \underline{\mathbf{h}}\big) = 1$

H **alakja :**

$$H = \begin{bmatrix} I & 0 \\ 0 & I - 2\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ & \ddots & 1 & 0 & \dots & 0 \\ & & 0 & 1-2h_i^2 & -2h_ih_{i+1} & \cdots & -2h_ih_k \\ \vdots & & \vdots & -2h_ih_{i+1} & 1-2h_{i+1}^2 & \cdots & -2h_{i+1}h_k \\ & & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -2h_ih_k & -2h_{i+1}h_k & \cdots & 1-2h_k^2 \end{bmatrix}$$

H **tulajdonságai :**

1. H szimmetrikus: $H = H^T$,

2. H idempotens: $H^2 = I$, ugyanis: $H^2 = \big(I - 2\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T\big)^2 = I - 4\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T + 4\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T = I$,

3. H ortogonális: $I = H\cdot H = H\cdot H^{-1} \quad \Longrightarrow \quad H = H^{-1} = H^T$

H **geometriai jelentése :**

A H mátrtixszal való szorzás az $\mathbb{R}^k$ térnek a tükrözését jelenti a $H_{h,0} := \{\underline{\mathbf{z}} \in \mathbb{R}^k \ : \ \underline{\mathbf{h}}^T\underline{\mathbf{z}} = \underline{\mathbf{0}}\}$ hipersíkra (a $\underline{\mathbf{h}}$ vektorra merőleges hipersík).

$H\underline{\mathbf{z}} = \big(I - 2\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T\big)\,\underline{\mathbf{z}} = \underline{\mathbf{z}} - 2\,\underline{\mathbf{h}}\,\big(\underline{\mathbf{h}}^T\underline{\mathbf{z}}\big)$

Alapfeladat :

Adott $\underline{\mathbf{x}} \in \mathbb{R}^k$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ vektorhoz adjuk meg azt a $H \in \mathbb{R}^{k\times k}$ Householder mátrixot (azaz a $\underline{\mathbf{h}} \in \mathbb{R}^k$, $\|\,\underline{\mathbf{h}}\,\|_2 = 1$ vektort), amely az $\underline{\mathbf{x}}$ vektort az $\underline{\mathbf{e}}_1$ egységvektor számszorosába transzformálja.

Vagyis adott $\underline{\mathbf{x}} \in \mathbb{R}^k$ vektorhoz meg kell adni a H mátrixot meghatározó $\underline{\mathbf{h}} \in \mathbb{R}^k$ vektort, amelyre teljesül, hogy:

$H\underline{\mathbf{x}} = \big(I - 2\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T\big)\,\underline{\mathbf{x}} = \sigma\underline{\mathbf{e}}_1$

Mivel H ortogonális $\quad\Longrightarrow\quad \|\,\underline{\mathbf{x}}\,\|_2 = \|\,H\underline{\mathbf{x}}\,\|_2 = |\sigma| \quad\Longrightarrow\quad |\sigma| = \|\,\underline{\mathbf{x}}\,\|_2$,

Továbbá $H\underline{\mathbf{x}} = \sigma\underline{\mathbf{e}}_1$, $\quad H\underline{\mathbf{x}}7\,\big(I - 2\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T\big)\,\underline{\mathbf{x}} = \underline{\mathbf{x}} - 2\,\underline{\mathbf{h}}\,\underline{\mathbf{h}}^T\underline{\mathbf{x}} = \underline{\mathbf{x}} - 2\,\big(\underline{\mathbf{h}}^T\underline{\mathbf{x}}\big)\,\underline{\mathbf{h}} = \sigma\underline{\mathbf{e}}_1 \quad\Longrightarrow$

$\Longrightarrow \quad 2\,\big(\underline{\mathbf{h}}^T\underline{\mathbf{x}}\big)\,\underline{\mathbf{h}} = \underline{\mathbf{x}} - \sigma\underline{\mathbf{e}}_1$

Innen az következik, hogy a $\underline{\mathbf{h}}$ vektor az $(\underline{\mathbf{x}} - \sigma\underline{\mathbf{e}}_1)$ vektor számszorosa kell legyen. De mivel a $\underline{\mathbf{h}}$ egységvektor ($\|\,\underline{\mathbf{h}}\,\|_2 = 1$), azért $\underline{\mathbf{x}} - \sigma\underline{\mathbf{e}}_1$ - et osztva a normájával, megkapjuk a keresett $\underline{\mathbf{h}}$ egységvektort.

$$\underline{\mathbf{h}} = \frac{\underline{\mathbf{x}} - \sigma\underline{\mathbf{e}}_1}{\|\,\underline{\mathbf{x}} - \sigma\underline{\mathbf{e}}_1\,\|_2} \qquad \text{ahol} \qquad \sigma \in \mathbb{R} \quad \text{és} \quad |\sigma| = \|\,\underline{\mathbf{x}}\,\|_2$$

Minden feltételt kihasználtunk, de σ előjele határozatlan maradt.
Az algoritmus stabilitására és a hibafelhalmozódás elkerülésére válasszuk σ előjelét a következő módon:

$$\boxed{\sigma := -\mathsf{sign}(x_1)\,\|\,\underline{\mathbf{x}}\,\|_2} \qquad \text{és legyen} \qquad \mathsf{sign}\,(x_1) = 1, \quad \text{ha } x_1 = 0$$

$\underline{\mathbf{h}}$ kiszámításához kell:

$$\|\,\underline{\mathbf{x}} - \sigma\underline{\mathbf{e}}_1\,\|_2^2 = \|\,\underline{\mathbf{x}} + \mathsf{sign}(x_1)\,\|\,\underline{\mathbf{x}}\,\|_2\,\underline{\mathbf{e}}_1\,\|_2^2 = \Big|\,|x_1| + \|\,\underline{\mathbf{x}}\,\|_2\,\Big|^2 + \sum_{i=1}^{k} |x_i|^2 =$$

$$= \|\,\underline{\mathbf{x}}\,\|_2^2 + 2\,|x_1|\,\|\,\underline{\mathbf{x}}\,\|_2 + \|\,\underline{\mathbf{x}}\,\|_2^2 = 2\,\|\,\underline{\mathbf{x}}\,\|_2\,(\,\|\,\underline{\mathbf{x}}\,\|_2 + |x_1|)$$

$$\Longrightarrow \qquad \underline{\mathbf{h}} = \frac{\Big[\mathsf{sign}\,(x_1)\,\big(\,|x_1| + \|\,\underline{\mathbf{x}}\,\|_2\,\big), x_2, \ldots, x_k\Big]^T}{\Big[\,2\,\|\,\underline{\mathbf{x}}\,\|_2\,\big(\,|x_1| + \|\,\underline{\mathbf{x}}\,\|_2\,\big)\Big]^{\frac{1}{2}}}$$

Így a következő H mátrix megoldja az alapfeladatot:

i.) $\qquad H = I - \beta\,\underline{\mathbf{u}}\,\underline{\mathbf{u}}^T,$

ii.) $\qquad \beta := \Big[\,\|\,\underline{\mathbf{x}}\,\|_2\,\big(\,|x_1| + \|\,\underline{\mathbf{x}}\,\|_2\,\big)\Big]^{-1},$

iii.) $\qquad \underline{\mathbf{u}} := \Big[\mathsf{sign}\,(x_1)\,\big(\,|x_1| + \|\,\underline{\mathbf{x}}\,\|_2\,\big), x_2, \ldots, x_k\Big]^T.$

2.8.1. A Householder - algoritmus

Legyen $A \quad (n \times n)$ - es tetszőleges mátrix, továbbá $A^{(0)} := A$.

Az A mátrixot minden lépésben szorozzuk $H^{(i)} \in \mathbb{R}^{n\times n}$ Householder - mátrixszal.

$$H^{n-1} \cdot \ldots \cdot H^2 \cdot H^1 \cdot A = R$$

H^i - k ortogonálisak $\qquad \Longrightarrow \qquad A = \big(H^{n-1} \cdot \ldots \cdot H^2 \cdot H^1\big)^{-1} \cdot R = \big(H^1\big)^{-1} \cdot \ldots \cdot \big(H^{n-1}\big)^{-1} \cdot R$

$A = H^{(1)} \cdot H^{(2)} \cdot \ldots \cdot H^{(n-1)} \cdot R \qquad \Longrightarrow \qquad Q := H^{(1)} \cdot \ldots \cdot H^{(n-1)}$.

Legyen $A^{(0)} := A$.

i.)

Az első lépésben $A^{(0)}$ - t szorozzuk olyan $H^{(1)}$ Householder - mátrixszal, amelyre teljesül, hogy $H^{(1)}(\underline{\mathbf{a}}) = \sigma \underline{\mathbf{e}}_1$, ahol $\underline{\mathbf{a}}$ az A mátrix első oszlopvektora.

ii.)

$(i-1)$ - edik lépésben $H^{(i-1)} \cdot A^{(i-2)} = A^{(i-1)}$, ahol

$$A^{(i-1)} = \begin{bmatrix} B^{(i-1)} & C^{(i-1)} \\ & \\ 0 & \widetilde{A}^{(i-1)} \end{bmatrix} \text{ felső } \triangle \text{ mátrix, } \qquad \text{ahol}$$

$$B^{(i-1)} \in \mathbb{R}^{(i-1)\times(i-1)}, \qquad \widetilde{A}^{(i-1)} \in \mathbb{R}^{(n-i+1)\times(n-i+1)}, \qquad C^{(i-1)} \in \mathbb{R}^{(i-1)\times(n-i+1)}$$

iii.)

Az i - edik lépésben megkeressük azt a $\widetilde{H}^{(i)} \in \mathbb{R}^{(n-i+1)\times(n-i+1)}$ - es Householder - mátrixot, amelyre teljesül, hogy

$$\widetilde{H}^{(i)} \cdot a^{(i-1)} = \sigma \underline{\mathbf{e}}_1 , \qquad \underline{\mathbf{e}}_1 \in \mathbb{R}^{(n-i+1)}$$

Így kiegészítve $\widetilde{H}^{(i)}$ - ot $(n \times n)$ - esre a következő módon

$$H^{(i)} = \begin{bmatrix} I_{i-1} & 0 \\ & \\ 0 & \widetilde{H}^{(i)} \end{bmatrix}, \qquad H^{(i)} \in \mathbb{R}^{n\times n}$$

Ez is szimmetrikus és ortogonális, és ezzel szorozzuk $A^{(i-1)}$ - t $\Longrightarrow$ $H^{(i)} \cdot A^{(i-1)} = A^{(i)}$

iv.)

$(n-1)$ - edik lépésben:

$$H^{(n-1)} \cdot \ldots \cdot H^{(2)} \cdot H^{(1)} \cdots A = A^{(n-1)} = R ,$$

$$A = \left(H^{n-1} \cdot \ldots \cdot H^2 \cdot H^1\right)^{-1} \cdot R = \left(H^1\right)^{-1} \cdot \ldots \cdot \left(H^{n-1}\right)^{-1} \cdot R = H^{(1)} \cdot H^{(2)} \cdot \ldots \cdot H^{(n-1)} \cdot R$$

Így megkapjuk az $A^{(n-1)} = R$ felső $\triangle$ mátrixot és a Q ortogonális, szimmetrikus mátrixot,

amelyekkel $\boxed{A = Q \cdot R}$

Tétel

Tetszőleges $A \in \mathbb{R}^{n \times n}$ - es mátrix felbontható $A = Q \cdot R$ alakban, ahol Q ortogonális, R felső $\triangle$ mátrix.

Megjegyzés :

Ez, az úgynevezett QR felbontási tétel könnyen kiterjeszthető mind komplex -, illetve téglalap - mátrixokra is.

2.8.2. A Householder - algoritmus alkalmazása

Alkalmazás : $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ egyenletrendszer megoldására.

Input: $n \in \mathbb{N}$ $\quad C = \left[\, A|\underline{\mathbf{b}}\, \right] = [c_{ij}] \in \mathbb{R}^{n\times(n-i+1)}$

1.

Értékadás: $i := 1$

2.

Eliminációs lépés: $s := \left(\sum_{k=i}^{n} c_{ki}^2 \right)^{\frac{1}{2}}$

a.)

Ha $s = 0$, stop , egyébként:

$\beta := [s\,(|c_{ii}| + s)]^{-1}$,

$\underline{\mathbf{u}} := \left[0, \ldots, 0, c_{ii} + \mathsf{sign}\,(c_{ii})\, s, c_{i+1\,i}, \ldots, c_{nn}\right]^T$,

$\mathsf{sign}(c_{ii}) = 1$, ha $c_{ii} = 0$,

$H^{(i)} := I - \beta\, \underline{\mathbf{u}}\, \underline{\mathbf{u}}^T$

b.)

$C := H^{(i)} \cdot C =: (c_{ij})$

3.

Ellenőrzés: ha $i + 1 \leq n - 1$, akkor $i := i + 1$ és go to stop **2.** , egyébként stop. Majd jön a visszahelyettesítés.

2.8.3. A Householder - algoritmus műveletigénye

Az algoritmus műveletigénye : $\frac{4}{3}n^3 + O\left(n^2\right)$ összesen.

R előállítása: $\frac{2}{3}n^3 + O\left(n^2\right)$,

Q előállítása: $\frac{2}{3}n^3 + O\left(n^2\right)$,

de $Q^T\underline{\mathbf{b}}$ szorzat kiszámításához elég $O\left(n^2\right)$ művelet plusz $(n-1)$ darab gyökvonás.

Tehát $A\underline{\mathbf{x}} = \underline{\mathbf{b}} \quad \Longrightarrow \quad R\underline{\mathbf{x}} = Q^T\underline{\mathbf{b}}$ a műveletigény: $\frac{2}{3}n^3 + O\left(n^2\right)$,

kétszer annyi, mint a Gauss - eliminációnál.

Megjegyzés :

A Householder - transzformáció alkalmazható mátrixok sajátértékeinek kiszámítására is (és sajátvektorok meghatározására). Numerikusan rendkívül stabil!

2.8.4. $A = A_0$ szimmetrikus mátrix tridiagonális alakra való transzformációja Householder módszerével

Egyszerre egy egész sort és oszlopot teszünk zérussá, kivéve a tridiagonális elemeket.
Az eljárás $(n-2)$ lépésből áll. r - edikben kapjuk a 0 - kat az r - edik sorban és oszlopban. A korábbi 0 - k nem változnak.

Legyen az r - edik lépés előtt a mátrix:

$$A_{r-1} = \begin{bmatrix} x & x & & & & \\ x & x & x & & & \\ & x & x & x & x & x \\ & & x & x & x & x \\ & & x & x & x & x \\ & & x & x & x & x \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} x \\ x \\ x \end{matrix}}\right\} r \\ \left.\vphantom{\begin{matrix} x \\ x \\ x \end{matrix}}\right\} n-r \end{matrix} = \begin{bmatrix} & C_{r-1} & & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ & \underline{\mathbf{b}}_{r-1}^T & \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & & \underline{\mathbf{b}}_{r-1} & B_{r-1} \end{bmatrix}$$

ahol $\underline{\mathbf{b}}_{r-1}$ vektornak $(n-r)$ komponense van, C_{r-1} r - edrendű tridiagonális mátrix,

B_{r-1} $(n-r)$ - edrendű mátrix.

A transzformáció P_r $(n \times n)$ - es mátrixát következő módon adjuk meg:

$$P_r = \begin{bmatrix} I & 0 \\ & \\ 0 & Q_r \end{bmatrix} = \begin{bmatrix} I & 0 \\ & \\ 0 & I - 2\,\underline{\mathbf{v}}_r\,\underline{\mathbf{v}}_r^T \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} r \\ \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} n-r \end{array}$$

ahol $\underline{\mathbf{v}}_r$ $(n-r)$ komponensű és $\|\underline{\mathbf{v}}\| = 1$.

Egyszerű számolással $\Longrightarrow$

$$\textit{(1)} \qquad A_r = P_r A_{r-1} P_r = \begin{bmatrix} & & & 0 & 0 & 0 \\ & C_{r-1} & & 0 & 0 & 0 \\ & & & & \underline{\mathbf{c}}_r^T & \\ 0 & 0 & & & & \\ 0 & 0 & \underline{\mathbf{c}}_r & & Q_r B_{r-1} Q_r & \\ 0 & 0 & & & & \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} r \\ \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} n-r \end{array}$$

ahol $\underline{\mathbf{c}}_r = Q_r \underline{\mathbf{b}}_{r-1}$.

Ha $\underline{\mathbf{v}}_r$ - t úgy választjuk, hogy $\underline{\mathbf{c}}_r$ komponensei az első kivételével mind eltűnjenek, akkor A_r első $(r+1)$ sorában és oszlopában tridiagonális lesz.

(1) $\Longrightarrow$ triviális, hogy hasonlósági transzformáció, és $A_0, A_1, \ldots, A_{n-2}$ sajátértékei megegyeznek.

Kényelmesebb, ha a P_r mátrixot a következő módon választjuk meg:

$$P_r = I - 2\,\underline{\mathbf{w}}_r\,\underline{\mathbf{w}}_r^T \ ,$$

ahol $\underline{\mathbf{w}}_r$ n komponensű, első r ebből 0, $\|\,\underline{\mathbf{w}}\,\|_2 = 1$.

Ekkor:

$$\textit{(2)} \qquad P_r = I - 2\,\underline{\mathbf{w}}_r\,\underline{\mathbf{w}}_r^T = I - \frac{U_r\,U_r^T}{2K_r^2} \ , \qquad \text{ahol}$$

$$(\star)\quad \begin{cases} u_{ir} = 0\,, \qquad i = 1, \ldots, r\,, \\[2ex] u_{r+1\,r} = a_{r+1\,r} \mp S_r\,, \\[2ex] u_{ir} = a_{ir}\,, \qquad i = r+2, \ldots, n\,, \\[2ex] S_r = \left(\sum\limits_{i=r+1}^{n} a_{ir}^2 \right)^{\frac{1}{2}}\,, \\[2ex] 2K_r^2 = S_r^2 \mp a_{r+1\,r} S_r = S_r u_{r+1\,r} \end{cases}$$

Itt a_{ik} az A_{r-1} mátrix $(i,\,k)$ elemét jelöli (az r indexet elhagytuk).

(⋆) - ban S_r előjelét abból a feltételből határozzuk meg, hogy

$$|u_{r+1\,r}| = |a_{r+1\,r}| + S_r\ ,$$

mivel akkor kapjuk a legnagyobb pontosságot, ha *(2)* - beli nevező, $2K_r^2 = S_r u_{r+1\,r}$ értéke a lehető legnagyobb.

E triangularizációs eljárás műveletigénye: $\dfrac{2}{3}n^3 + O\left(n^2\right)$,

és emellett $(n-2)$ darab négyzetgyökvonás is kell!

Tétel

Legyen $A\quad (n \times n)$ - es nem szinguláris mátrix. Ekkor az A mátrix QR - felbontása egyértelmű az R mátrix diagonális elemeinek az előjelétől eltekintve.

Bizonyítás

Indirekt.

Tegyük fel, hogy:

(⋆) $\qquad A = Q_1 R_1 = Q_2 R_2$, $\qquad$ ahol Q_1, Q_2 ortogonális, $\quad R_1$, R_2 felső $\triangle$ mátrixok.

Mivel $\quad \det A = \det Q_i \cdot \det R_i = \det R_i\,, \qquad i = 1, 2$

azért $\qquad \Longrightarrow \qquad R_i \qquad (i = 1, 2) \qquad$ nem szinguláris mátrixok.

(⋆) $\qquad \Longrightarrow \qquad Q_2^T\, Q_1 = R_2\, R_1^{-1} = V$,

ahol $\quad V \quad$ szintén felső $\triangle$ mátrix.

Itt $Q_2^T Q_1$ ortogonális mátrixok, ezért

$$V^T V = Q_1^T Q_2 Q_2^T Q_1 = I$$

Tehát $V = D = \operatorname{diag}(d_{ii})$, ahol $d_{ii} = \pm 1$.

Mivel $V = R_2 R_1^{-1}$ is teljesül, az állítás bizonyított.

■

3. VEKTORNORMÁK, MÁTRIXNORMÁK

3.1. VEKTORNORMÁK

Definíció :

Vektornorma :

Legyen X n - dimenziós vektortér $\mathbb{K} := \mathbb{R}$ vagy $\mathbb{C}$ test felett.
A norma egy $\|\,.\,\| : X \to \mathbb{R}$, $\underline{\mathbf{x}} \to \|\underline{\mathbf{x}}\|$ leképezés a következő tulajdonságokkal:

$\forall \ \underline{\mathbf{x}}, \underline{\mathbf{y}} \in X$, $\forall \ \lambda \in \mathbb{C}$:

1. $\|\underline{\mathbf{x}}\| \geq 0$, $\|\underline{\mathbf{x}}\| = 0 \iff \underline{\mathbf{x}} = \underline{\mathbf{0}}$,

2. $\|\lambda \underline{\mathbf{x}}\| = |\lambda| \, \|\underline{\mathbf{x}}\|$,

3. $\|\underline{\mathbf{x}} + \underline{\mathbf{y}}\| \leq \|\underline{\mathbf{x}}\| + \|\underline{\mathbf{y}}\|$ $\triangle$ egyenlőtlenség.

Következmény

$(\star)$ $\big|\, \|\underline{\mathbf{x}}\| - \|\underline{\mathbf{y}}\| \,\big| \leq \|\underline{\mathbf{x}} - \underline{\mathbf{y}}\|$, $\forall \ \underline{\mathbf{x}}, \underline{\mathbf{y}} \in X$

Az $(X, \|\,.\,\|)$ párt normált vektortérnek nevezzük.

Leggyakoribb vektornormák :

Euklideszi norma: $\|\underline{\mathbf{x}}\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$, $\forall \ \underline{\mathbf{x}} \in X$, $1 \leq p < \infty$,

$\|\underline{\mathbf{x}}\|_\infty := \max_{i=1,\ldots,n} |x_i|$

Az $1, 2, \infty$ normákat használjuk leggyakrabban.

Tétel

$\underline{\mathbf{x}} \in X\,, \qquad \dim X = n\ .$
Az $\|\,\underline{\mathbf{x}}\,\|$ vektornorma az $\underline{\mathbf{x}}$ vektor $x_1, \ldots, x_n$ koordinátáinak folytonos függvénye.

Bizonyítás

Ezt $\underline{\mathbf{x}} \in \mathbb{R}^n$ - re és $\|\,.\,\|_\infty$ normára bizonyítjuk.

$$\|\,\underline{\mathbf{x}}\,\|_\infty := \max_{1\le i\le n} |x_i|\ , \qquad \underline{\mathbf{x}} := (x_1, \ldots, x_n)$$

Legyen $\underline{\mathbf{z}} := (z_1, \ldots, z_n)$ és tekintsük $\underline{\mathbf{x}} + \underline{\mathbf{z}}\,,\ \ \underline{\mathbf{x}} \in X$ vektort.

$$(\star) \qquad \Longrightarrow \qquad \big|\,\|\,\underline{\mathbf{x}} + \underline{\mathbf{z}}\,\|_\infty - \|\,\underline{\mathbf{x}}\,\|_\infty\,\big| \le \big\|\,(\underline{\mathbf{x}} + \underline{\mathbf{z}}) - \underline{\mathbf{x}}\big\|_\infty = \|\,\underline{\mathbf{z}}\,\|_\infty = \max_{1\le i\le n} |z_i|$$

$$f\,(x_n) \to f\,(y)\ , \qquad \text{ha} \qquad x_n \to y$$

Tehát, ha $\underline{\mathbf{z}} \to \underline{\mathbf{0}}$ akkor $\max\limits_{1\le i\le n} |z_i| \to 0\ .$

Mivel végesdimenziós vektortérben bármely két vektornorma ekvivalens egymással (lásd később), ezért tetszőleges normára a folytonosság igaz.

■

Definíció :

Az X komplex vektorteret a $\mathbb{C}$ test felett komplex Euklideszi térnek nevezzük, ha létezik X -en $(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}) : X \to \mathbb{C}$ leképezés tetszőleges $\underline{\mathbf{x}}\,,\,\underline{\mathbf{y}} \in X$ - re, amit az $\underline{\mathbf{x}}\,,\underline{\mathbf{y}}$ elemek skaláris szorzatának nevezünk és amelyre teljesül, hogy $\forall\ \ \underline{\mathbf{x}}\,,\,\underline{\mathbf{y}} \in X\,, \quad \forall\ \ \lambda \in \mathbb{C}$:

1. $(\underline{\mathbf{x}}\,,\underline{\mathbf{x}}) \ge 0$ valós szám, $\quad (\underline{\mathbf{x}}\,,\underline{\mathbf{x}}) = 0 \quad \Longleftrightarrow \quad \underline{\mathbf{x}} = \underline{\mathbf{0}}$,

2. $(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}) = \overline{(\underline{\mathbf{y}}\,,\underline{\mathbf{x}})}$,

3. $(\lambda\underline{\mathbf{x}}\,,\underline{\mathbf{y}}) = \lambda\,(\underline{\mathbf{x}}\,,\underline{\mathbf{y}})$ első tényezőre nézve homogén,

4. $(\underline{\mathbf{x}}_1 + \underline{\mathbf{x}}_2\,,\underline{\mathbf{y}}) = (\underline{\mathbf{x}}_1\,,\underline{\mathbf{y}}) + (\underline{\mathbf{x}}_2\,,\underline{\mathbf{y}})$ első tényezőre nézve disztributív.

Következmény

a.) $(\underline{\mathbf{x}}\,,\lambda\underline{\mathbf{y}}) = \overline{\lambda}\,(\underline{\mathbf{x}}\,,\underline{\mathbf{y}})$,

b.) $\left(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}_1 + \underline{\mathbf{y}}_2\right) = \left(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}_1\right) + \left(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}_2\right)$ második tényezőre nézve disztributív.

Konkrét példa :

Legyen $\mathbb{C}^n$ komplex vektortér $\mathbb{C}$ felett, $\underline{\mathbf{x}} := (x_1, \ldots, x_n)$.

Skaláris szorzat:

$$\big(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}\big) := \overline{\underline{\mathbf{x}}}^T \underline{\mathbf{y}} = \sum_{i=1}^{n} \overline{x_i}\, y_i \qquad \text{és} \qquad (\underline{\mathbf{x}}\,,\underline{\mathbf{x}}) = \sum_{i=1}^{n} |x_i|^2 \geq 0$$

Erre a **3.** axióma helyett $\big(\underline{\mathbf{x}}\,,\lambda\underline{\mathbf{y}}\big) = \lambda\,\big(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}\big)$ érvényes.

További tulajdonságok :

Ha van skaláris szorzat, akkor X - en lehet normát definiálni:

$\|\,\underline{\mathbf{x}}\,\| := \sqrt{(\underline{\mathbf{x}}\,,\underline{\mathbf{x}})}$, erre teljesülnek a norma axiómái.

Cauchy - Bunyakovsky - Schwarz egyenlőtlenség :

$$\boxed{\;\big|\,\big(\underline{\mathbf{x}}\,,\underline{\mathbf{y}}\big)\,\big| \leq \|\,\underline{\mathbf{x}}\,\| \cdot \|\,\underline{\mathbf{y}}\,\|\;} \qquad \forall \quad \underline{\mathbf{x}}\,,\,\underline{\mathbf{y}} \in X$$

3.2. MÁTRIXNORMÁK

Definíció :

Legyen $\mathbb{K}^{n\times n}$ $(\mathbb{K} := \mathbb{R}$ vagy $\mathbb{C}$ az $(n \times n)$ - es mátrixok tere. $A \in \mathbb{K}^{n\times n}$.

Az $\|\,A\,\|\ : \mathbb{K}^{n\times n} \to \mathbb{R}$ leképezést mátrixnormának nevezzük, ha $\forall\ \ A\,,\,B \in \mathbb{K}^{n\times n}$, $\forall\ \ \lambda \in \mathbb{K}$ - ra teljesül, hogy:

1. $\|\,A\,\| \geq 0$, $\|\,A\,\| = 0 \iff A = 0$,

2. $\|\,\lambda A\,\| = |\lambda|\ \|\,A\,\|$ homogén,

3. $\|\,A + B\,\| \leq \|\,A\,\| + \|\,B\,\|$ $\triangle$ egyenlőtlenség,

4. $\|\,A \cdot B\,\| \leq \|\,A\,\| \cdot \|\,B\,\|$

A mátrixokat mint leképezéseket (operátorokat) tekintjük.

Legyen $\mathbb{K}^{m\times n}$ valós vagy komplex elemű $(m \times n)$ - es mátrixok $(m \cdot n)$ - dimenziós vektortere $\mathbb{R}$ vagy $\mathbb{C}$ felett, és $(X\,,\,\|\,.\,\|_x\,)$ n - dimenziós, $(Y\,,\,\|\,.\,\|_y\,)$ m - dimenziós vektortér.

Egy $A \in \mathbb{K}^{m\times n}$ - es mátrix egy lineáris leképezést definiál X - ből Y - ba.

Tudjuk, hogy az $X \to \| A\underline{\mathbf{x}} \|_y$ leképezés folytonos függvény az $\{\underline{\mathbf{x}} \in X : \| \underline{\mathbf{x}} \|_x = 1\}$ kompakt (korlátos és zárt) halmazon. A kompaktságot itt az $\| . \|_x$ norma szerint értjük.

Definíció :

Az $\| A \| := \sup\limits_{\underline{\mathbf{x}} \in X\setminus\{0\}} \dfrac{\| A\underline{\mathbf{x}} \|_y}{\| \underline{\mathbf{x}} \|_x} = \max\limits_{\| \underline{\mathbf{x}} \|_x = 1} \| A\underline{\mathbf{x}} \|_y$ véges számot

vektornorma által indukált mátrixnormának vagy természetes mátrixnormának is nevezik.

$$(\sup)\frac{\| A\underline{\mathbf{x}} \|_y}{\| \underline{\mathbf{x}} \|_x} = (\sup)\left\| A\left(\frac{\underline{\mathbf{x}}}{\| \underline{\mathbf{x}} \|_x}\right)\right\|_y = (\sup_{\| \underline{\mathbf{y}} \|_x = 1})\, \| A\underline{\mathbf{y}} \|_y = \max_{\| \underline{\mathbf{x}} \|_x = 1} \| A\underline{\mathbf{x}} \|_y$$

Következmény

A definícióból azonnal következik, hogy:

($\star$) $\quad \| A\underline{\mathbf{x}} \|_y \le \| A \| \cdot \| \underline{\mathbf{x}} \|_x$,

EZENTÚL: $\quad X = Y, \quad \| . \|_x = \| . \|_y =: \| . \|$

Így *($\star$)* $\Longrightarrow$ $\boxed{\| A\underline{\mathbf{x}} \| \le \| A \| \cdot \| \underline{\mathbf{x}} \|}$ *($\star\star$)*

Állítás

A fent definiált természetes (indukált) mátrixnorma ($A \to \| A \|$ leképezés) tényleg norma, azaz teljesülnek a mátrixnorma axiómái.

Bizonyítás

A **2.** homogenitás és a **3.** $\triangle$ egyenlőtlenség nyílvánvalóak, közvetlenül következnek a vektornorma hasonló tulajdonságaiból.

Az **1.** $\| A \| = 0 \iff A = 0$ is triviálisan teljesül, mivel:

$\boxed{\Rightarrow}$

($\star\star$) $\Longrightarrow$ $\| A\underline{\mathbf{x}} \| = 0 \quad \forall \ \underline{\mathbf{x}} \in X$ - re, $\quad A\underline{\mathbf{x}} = \underline{\mathbf{0}} \Longrightarrow A = 0$

$\boxed{\Leftarrow}$

A definíció alapján $\| A\underline{\mathbf{x}} \| = 0 \qquad \forall \quad \| \underline{\mathbf{x}} \| = 1$ - re $\qquad \Longrightarrow \qquad \| A \| = 0$.

A **4.** - hez vizsgáljuk:

$$\| AB\underline{\mathbf{x}} \| \overset{(\star\star)}{\leq} \| A \| \cdot \| B\underline{\mathbf{x}} \| \overset{(\star\star)}{\leq} \| A \| \cdot \| B \| \cdot \| \underline{\mathbf{x}} \| \qquad \overset{\text{def}}{\Longrightarrow} \qquad \| AB \|_{\leq} \| A \| \cdot \| B \|$$

■

Megjegyzés :

i.) Nyílván: $\| I \| = 1$

ii.) A $C := \| A \|$ az a legkisebb állandó szám, amelyre teljesül, hogy

$$\| A\underline{\mathbf{x}} \| \leq C \cdot \| \underline{\mathbf{x}} \| , \qquad \forall \quad \underline{\mathbf{x}} \in X$$

Az egyenlőség akkor teljesűl, ha $\underline{\mathbf{x}}$ az a vektor, amelyre $\| A\underline{\mathbf{x}} \|$ felveszi a maximális értékét.

iii.) $\mathbb{K}^{n \times n}$ - beli mátrixokra és a természetes normára érvényes a következő egyenlőtlenség:

$$\| A \cdot B \| \leq \| A \| \cdot \| B \|$$

Ugyanis: $\quad \| AB\underline{\mathbf{x}} \| \leq \| A \| \cdot \| B\underline{\mathbf{x}} \| \leq \| A \| \cdot \| B \| \cdot \| \underline{\mathbf{x}} \|$

Innen: $\quad \| AB \| \leq \| A \| \cdot \| B \|$ következik, mivel $\| AB \|$ a legkisebb olyan szám,

amelyre $\| AB\underline{\mathbf{x}} \| \leq C \cdot \| \underline{\mathbf{x}} \|$

Voltak :

$$\| \underline{\mathbf{x}} \|_1 := \sum_{i=1}^{n} |x_i| , \qquad \| \underline{\mathbf{x}} \|_2 := \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}} , \qquad \| \underline{\mathbf{x}} \|_\infty := \max_{1 \leq i \leq n} |x_i|$$

Kérdés :

A fenti vektornormák milyen mátrixnormákat definiálnak?

Definíció :

Legyen $A \in \mathbb{K}^{n\times n}$, $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ az A mátrix sajátértékei.

Ekkor a $\rho(A) := \max\limits_{1\le i\le n} |\lambda_i|$ számot az A mátrix spektrálsugarának nevezzük.

Jelölje $\|A\|_{1,2,\infty}$ az $\|\underline{\mathbf{x}}\|_{1,2,\infty}$ vektornormák által indukált mátrixnormákat.

A definíció alapján: $\|A\|_1 := \max\limits_{\|\underline{\mathbf{x}}\|=1} \|A\underline{\mathbf{x}}\|_1 = \sup\limits_{\underline{\mathbf{x}}\in X\setminus\{0\}} \dfrac{\|A\underline{\mathbf{x}}\|_1}{\|\underline{\mathbf{x}}\|_1}$

Tétel

Legyen $\|\,.\,\|_p$ az a mátrixnorma, amelyet a $\|\,.\,\|_p$ vektornorma indukál a $\mathbb{K}^{n\times n}$ mátrix-téren. Ekkor:

1. $\|A\|_1 := \max\limits_{1\le j\le n}\left(\sum\limits_{i=1}^{n}|a_{ij}|\right)$, Maximum oszlop-összeg norma

2. $\|A\|_\infty := \max\limits_{1\le i\le n}\left(\sum\limits_{j=1}^{n}|a_{ij}|\right)$, Maximum sor-összeg norma

3. $\|A\|_2 := \sqrt{\rho\left(\overline{A}^T A\right)}$, Spektrálnorma

Bizonyítás

1.

$$\|A\underline{\mathbf{x}}\|_1 = \max_{1\le j\le n}\left(\left|\sum_{i=1}^{n} a_{ij}\cdot x_i\right|\right) \le \max_{1\le j\le n}\left(\sum_{i=1}^{n}|a_{ij}|\cdot|x_i|\right) \le \max_{1\le j\le n}\left(\sum_{i=1}^{n}|a_{ij}|\right)\cdot\max_{1\le j\le n}|x_i|$$

$$\Longrightarrow \qquad \|A\|_1 \le \max_{1\le j\le n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$$

Tegyük fel, hogy a k - adik oszlop olyan, amelyre teljesül, hogy:

$$\|A\|_1 = \sum_{i=1}^{n}|a_{ik}| = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$$

Tekintsük az $\underline{\mathbf{e}}_k$ egységvektort. Ekkor:

$$\|A\underline{\mathbf{e}}_k\|_1 = \sum_{i=1}^{n}|a_{ik}| = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$$

Mivel $\|\underline{\mathbf{e}}_k\|_1 = 1$ $\Longrightarrow$ $\|A\|_1 \ge \max\limits_{1\le j\le n}\sum\limits_{i=1}^{n}|a_{ij}|$ $\Longrightarrow$ Tehát az egyenlőség fennáll!

2.

$$\| A\underline{\mathbf{x}} \|_\infty = \max_{1\le i\le n}\left(\left|\sum_{j=1}^{n} a_{ij}\cdot x_i\right|\right) \le \max_{1\le i\le n}\left(\sum_{j=1}^{n} |a_{ij}|\cdot |x_i|\right) \le \max_{1\le i\le n}\left(\sum_{j=1}^{n} |a_{ij}|\right)\cdot \max_{1\le i\le n} |x_i|$$

$$\Longrightarrow \qquad \| A \|_\infty \le \max_{1\le i\le n}\left(\sum_{j=1}^{n} |a_{ij}|\right)$$

Belátjuk, hogy itt egyenlőség is fennáll. Tegyük fel, hogy a k - adik sor olyan, amelyre teljesül, hogy:

$$\| A \|_\infty = \sum_{j=1}^{n} |a_{ik}| = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

Legyen $\widehat{\underline{\mathbf{x}}} \in \mathbb{K}^n$, amelyre $\widehat{x}_j := \begin{cases} 1 & \text{ha } a_{kj} = 0 \\ \dfrac{\overline{a}_{kj}}{|a_{kj}|} & \text{egyébként} \end{cases}$

Ekkor $\| \widehat{\underline{\mathbf{x}}} \|_\infty = 1$, $\quad a_{kj}\cdot \overline{a}_{kj} = |a_{kj}|^2 \quad \Longrightarrow \quad \| A\widehat{\underline{\mathbf{x}}} \|_\infty = \sum_{j=1}^{n} |a_{kj}| = \| A \|_\infty$

3.

$$\| A \|_2 = \sqrt{\rho\left(\overline{A}^T A\right)}$$

$\| A \|_2 = \max\limits_{\| \underline{\mathbf{x}} \| = 1} \| A \|_2 \quad \Longrightarrow \quad \exists \ \underline{\mathbf{y}} \in \mathbb{K}^n \quad \| \underline{\mathbf{y}} \|_2 = 1$ és $\| A\underline{\mathbf{y}} \|_2 = \| A \|_2$

úgy, hogy: $\quad \| A \|_2^2 = \overline{\underline{\mathbf{y}}}^T \overline{A}^T A\underline{\mathbf{y}}$

$\overline{A}^T A$ Hermitikus mátrix $\quad \Longrightarrow \quad \overline{\left(\overline{A}^T A\right)}^T = \overline{A}^T A$

Mivel $\overline{A}^T A$ Hermitikus $\quad \Longrightarrow \quad \exists$ az $\underline{\mathbf{x}}^1, \ldots, \underline{\mathbf{x}}^n$ sajátvektoroknak teljes ortonormált rendszere.

$\| \underline{\mathbf{x}}^i \|_2 = 1$, $\quad \left(\overline{\underline{\mathbf{x}}^i}\right)^T \cdot \underline{\mathbf{x}}^j = \delta_{ij}$, $\quad \lambda_1, \ldots, \lambda_n$ a nekik megfelelő sajátértékek.

Vizsgáljuk :

$$0 \le \| A\underline{\mathbf{x}}^i \|_2^2 = \overline{\underline{\mathbf{x}}^i}^T \overline{A}^T A\underline{\mathbf{x}}^i = \left(\overline{\underline{\mathbf{x}}^i}\right)^T \cdot \lambda_i \underline{\mathbf{x}}^i = \lambda_i \left(\overline{\underline{\mathbf{x}}^i}\right)^T \underline{\mathbf{x}}^i = \lambda_i$$

$\Longrightarrow \qquad \overline{A}^T A$ pozitív szemidefinit

Írjuk fel $\underline{\mathbf{y}}$ - t a sajátvektorok lineáris kombinációjaként:

$$\underline{\mathbf{y}} = \sum_{i=1}^{n} \alpha_i \underline{\mathbf{x}}^i$$

$$\Longrightarrow \qquad 1 = \| \underline{\mathbf{y}} \|_2^2 = \left(\sum_{i=1}^{n} \overline{\alpha}_i \left(\overline{\underline{\mathbf{x}}^i} \right)^T \right) \cdot \left(\sum_{j=1}^{n} \alpha_j \underline{\mathbf{x}}^j \right) = \sum_{i=1}^{n} |\alpha_i|^2$$

Vizsgáljuk : $\left(\overline{A}^T A \underline{\mathbf{x}}^i = \lambda_j \underline{\mathbf{x}}^j \right)$

$$\| A \|_2^2 = \| A \underline{\mathbf{y}} \|_2^2 = \overline{\underline{\mathbf{y}}}^T \overline{A}^T A \underline{\mathbf{y}} = \left(\sum_{i=1}^{n} \overline{\alpha}_i \left(\overline{\underline{\mathbf{x}}^i} \right)^T \right) \cdot \left(\overline{A}^T A \sum_{j=1}^{n} \alpha_j \underline{\mathbf{x}}^j \right) = \left(\sum_{i=1}^{n} \overline{\alpha}_i \left(\overline{\underline{\mathbf{x}}^i} \right)^T \right) \cdot \left(\sum_{j=1}^{n} \alpha_j \lambda_j \underline{\mathbf{x}}^j \right) =$$

$$= \sum_{i=1}^{n} |\alpha_i|^2 \cdot \lambda_i \le \max_{1 \le i \le n} \lambda_i \left(\sum_{i=1}^{n} |\alpha_i|^2 \right) \qquad \Longrightarrow \qquad \| A \|_2^2 \le \max_{1 \le i \le n} \lambda_i = \rho \left(\overline{A}^T A \right)$$

Itt egyenlőség is állhat, legyen ugyanis $\lambda_k = \rho \left(\overline{A}^T A \right)$, azaz a legnagyobb sajátértéke $\overline{A}^T A$ - nak.

A λ_k - hoz tartozó sajátvektor legyen $\underline{\mathbf{x}}^k$, $\| \underline{\mathbf{x}}^k \|_2 = 1$.

Vizsgáljuk :

$$\| A \|_2^2 \ge \| A \underline{\mathbf{x}}^k \|_2^2 = \left(\overline{\underline{\mathbf{x}}^k} \right)^T \overline{A}^T A \underline{\mathbf{x}}^k = \lambda_k \underline{\mathbf{x}}^k = \lambda_k \left(\overline{\underline{\mathbf{x}}^k} \right)^T \underline{\mathbf{x}}^k = \lambda_k = \rho \left(\overline{A}^T A \right) ,$$

$$\Longrightarrow \qquad \| A \|_2^2 \ge \rho \left(\overline{A}^T A \right)$$

Így $\Longrightarrow \qquad \| A \|_2 = \sqrt{\rho \left(\overline{A}^T A \right)}$

■

Definíció :

Az $\| \, . \, \|_\alpha$ és $\| \, . \, \|_\beta$ vektornormákat ekvivalensnek nevezzük, ha $\exists \ m , M \in \mathbb{R}^+$ állandók, hogy:

$$m \, \| \underline{\mathbf{x}} \|_\alpha \le \| \underline{\mathbf{x}} \|_\beta \le M \, \| \underline{\mathbf{x}} \|_\alpha \qquad \forall \ \ \underline{\mathbf{x}} \in X \text{ - re}$$

Nyílván $\Longrightarrow \quad \exists \ r , R \in \mathbb{R}^+$, hogy $r \, \| \underline{\mathbf{x}} \|_\beta \le \| \underline{\mathbf{x}} \|_\alpha \le R \, \| \underline{\mathbf{x}} \|_\beta \qquad r := \dfrac{1}{M} \qquad R := \dfrac{1}{m}$

Adott végesdimenziós vektortéren tetszőleges két norma ekvivalens egymással.
Belátva azt, hogy egy tetszőleges norma ekvivalens a $\|\,.\,\|_\infty$ normával:

$$\begin{array}{l} \|\,\underline{\mathbf{z}}\,\|_\alpha \leq M\,\|\,\underline{\mathbf{z}}\,\|_\infty \;\rightarrow\; 0 \\ \downarrow \\ 0 \end{array}$$

$\Longrightarrow$ A mátrixoknak egy adott vektorterén az összes természetes mátrixnorma ekvivalens egymással (ekvivalenciareláció).

3.2.1. Korlátok a normákra

Néha nehéz kiszámolni egy mátrix természetes normáját, így hasznos, ha a normára felső korlátot tudunk találni. Jó példa erre az $\|\,A\,\|_2$ spektrálnorma, mivel itt $\overline{A}^T A$ legnagyobb sajátértékét kellene ismernünk.

Definíció :

Az $\|\,A\,\|$ mátrixnorma illeszkedik az $\|\,\underline{\mathbf{x}}\,\|$ vektornormához, ha $\|\,A\underline{\mathbf{x}}\,\| \leq \|\,A\,\| \cdot \|\,\underline{\mathbf{x}}\,\| \;\; \forall \quad \underline{\mathbf{x}} \in \mathbb{K}^n$.

Természetes mátrixnormákra láttuk, hogy illeszkednek az őket indukáló vektornormákhoz:

$$\|\,A\underline{\mathbf{x}}\,\|_{1,\,2,\,\infty} \leq \|\,A\,\|_{1,\,2,\,\infty}\;\|\,\underline{\mathbf{x}}\,\|_{1,\,2,\,\infty}$$

Megjegyzés :

Az $\|\,A\,\|$ természetes mátrixnorma az egyenlőtlenségben a legkisebb azon C állandók közül, amelyekre teljesül, hogy $\|\,A\underline{\mathbf{x}}\,\| \leq C\,\|\,\underline{\mathbf{x}}\,\|$.

Definíció :

Frobenius - mátrixnorma :

$$\|\,A\,\|_F := \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}} = \sqrt{\operatorname{trace}\left(\overline{A}^T A\right)} \qquad A \in \mathbb{K}^{n\times n}$$

$\boxed{\textbf{Állítás}}$

$\|\,A\,\|_F$ mátrixnorma, és illeszkedik a $\|\,.\,\|_2$ vektornormához.

Bizonyítás

Valóban, ugyanis $\sqrt{\text{trace}\left(\overline{A}^T A\right)} = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}}$ úgy tekinthető, mint egy $\underline{\mathbf{y}} \in \mathbb{C}^{n \times n}$ vektor $\| \, . \, \|_2$ euklideszi normája, ahol $\underline{\mathbf{y}}$ komponensei a_{ij} . Így **1 - 3.** axiómák teljesülnek, továbbá:

$$\| AB \|_F^2 = \sum_{i,k=1}^{n} \left| \sum_{j=1}^{n} a_{ij} \, b_{jk} \right|^2 \le \sum_{i,k=1}^{n} \left[\left(\sum_{j=1}^{n} |a_{ij}|^2 \right) \cdot \left(\sum_{j=1}^{n} |b_{jk}|^2 \right) \right] =$$

$$= \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right) \cdot \left(\sum_{j,k=1}^{n} |b_{jk}|^2 \right) = \| A \|_F^2 \cdot \| B \|_F^2$$

Mivel:

$$\| A\underline{\mathbf{x}} \|_2^2 = \sum_{i=1}^{n} \left| \sum_{k=1}^{n} a_{ik} \, x_k \right|^2 \le \sum_{i=1}^{n} \left[\sum_{k=1}^{n} |a_{ik}|^2 \sum_{k=1}^{n} |x_k|^2 \right] = \| A \|_F^2 \cdot \| \underline{\mathbf{x}} \|_2^2 \quad \Longrightarrow$$

$\Longrightarrow$ $\| A \|_F$ illeszkedik az $\| \underline{\mathbf{x}} \|_2$ vektornormához. Ez egy hasznos felső korlátot ad a $\| A \|_2$ mátrixnormára.

■

Állítás

A Frobenius - mátrixnorma nem természetes (nem vektornorma által indukált) mátrixnorma.

Bizonyítás

Minden természetes mátrixnormára $\| I \| = 1$, ami nyilvánvaló, ha az

$$\| A \| := \max_{\underline{\mathbf{x}} \neq \underline{\mathbf{0}}} \frac{\| A\underline{\mathbf{x}} \|}{\| \underline{\mathbf{x}} \|} = \max_{\| \underline{\mathbf{x}} \| = 1} \| A\underline{\mathbf{x}} \|$$ definíciót az $A = I$ - re alkalmazzuk.

DE :

$\| I \|_F = \sqrt{n} \quad \Longrightarrow \quad n > 1$ esetén ellentmondás.

■

4. LINEÁRIS EGYENLETRENDSZEREK PERTURBÁCIÓJA

Az $A\underline{\mathbf{x}} = \underline{\mathbf{b}}\,,\quad A \in \mathbb{K}^{n\times n}\,,\quad \underline{\mathbf{b}} \in \mathbb{K}^n$ lineáris egyenletrendszer megoldásakor a gyakorlatban az $A\,,\,\underline{\mathbf{b}}$ elemeit csak közelítőleg ismerjük. Ilyenkor az eddigi módszerek is csak közelítő megoldást adnak. Ha a bemenő adatokat a számítógépben pontosan is tároljuk, a numerikus módszer mégsem a pontos megoldást fogja szolgáltatni, azaz a defektus (eltérés) $\neq 0$. Ezért szükség van a megoldásbeli hiba becslésére (korlátozására).

Feladat :

Legyen $A \in \mathbb{K}^{n\times n}$ nem szinguláris, $\underline{\mathbf{b}} \in \mathbb{K}^n$.

Vizsgáljuk meg az $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ lineáris egyenletrendszerbeli $A\,,\,\underline{\mathbf{b}}$ bemenő adatok hibájának hatását a megoldásra.

4.1. A JOBB OLDAL HIBÁJÁNAK (PERTURBÁCIÓJÁNAK) HATÁSA

$$A\,(\underline{\mathbf{x}} + \Delta\underline{\mathbf{x}}) = \underline{\mathbf{b}} + \Delta\underline{\mathbf{b}} \qquad \underline{\mathbf{x}}\ ,\ \Delta\underline{\mathbf{x}}\ ,\ \underline{\mathbf{b}}\ ,\ \Delta\underline{\mathbf{b}} \in \mathbb{K}^n\,,\quad A \in \mathbb{K}^{n\times n}\ ,$$

$$A\Delta\underline{\mathbf{x}} = \Delta\underline{\mathbf{b}} \quad\Longrightarrow\quad \Delta\underline{\mathbf{x}} = A^{-1}\Delta\underline{\mathbf{b}} \quad\Longrightarrow\quad \boxed{\|\,\Delta\underline{\mathbf{x}}\,\| \le \|\,A^{-1}\,\| \cdot \|\,\Delta\underline{\mathbf{b}}\,\|}$$

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}} \quad\Longrightarrow\quad 0 < \|\,\underline{\mathbf{b}}\,\| \le \|\,A\,\| \cdot \|\,\underline{\mathbf{x}}\,\| \quad\Longrightarrow\quad \boxed{\frac{1}{\|\,\underline{\mathbf{x}}\,\|} \le \frac{\|\,A\,\|}{\|\,\underline{\mathbf{b}}\,\|}}$$

Természetes mátrixnormát véve:

$$\boxed{\frac{\|\,\Delta\underline{\mathbf{x}}\,\|}{\|\,\underline{\mathbf{x}}\,\|} \le \|\,A^{-1}\,\| \cdot \|\,A\,\| \frac{\|\,\Delta\underline{\mathbf{b}}\,\|}{\|\,\underline{\mathbf{b}}\,\|}}$$

Így a megoldás relatív hibája a jobb oldal relatív hibájával becsülhető.

Definíció :

Legyen $A \in \mathbb{K}^{n\times n}$ nem szinguláris.

Az $\|\,A^{-1}\,\| \cdot \|\,A\,\|$ számot az A mátrix kondíciószámának nevezzük, és $\mathrm{cond}(A)$ - val jelöljük.

$$\mathrm{cond}(A) := \|\,A^{-1}\,\| \cdot \|\,A\,\|$$

Ha $\mathrm{cond}(A) \sim 1$ jól kondícionált,
ha $\mathrm{cond}(A) \gg 1$ rosszul kondícionált.

Természetes mátrixnormák esetén:

$$1 = \| I \| = \| A^{-1}A \| \leq \| A^{-1} \| \cdot \| A \| = \mathrm{cond}(A)$$

Tehát a kondíciószám ≥ 1. Persze értéke függ attól, milyen normát használunk.

Példa :

$$\left.\begin{aligned} 100x_1 + 500x_2 &= 1700 \\ 15x_1 + 75.01x_2 &= 255 \end{aligned}\right\} \qquad x_1 = 17\,, \qquad x_2 = 0\ ,$$

de ha:
$$\left.\begin{aligned} 100x_1 + 500x_2 &= 1700 \\ 15x_1 + 75.01x_2 &= 255.03 \end{aligned}\right\} \qquad x_1 = 2\,, \qquad x_2 = 3\ .$$

Lemma

Legyen $A \in \mathbb{K}^{n\times n}$, A nem szinguláris, $\| A \| < 1$ valamilyen természetes mátrixnormára.

Ekkor $(I + A)^{-1}$ mátrix létezik és érvényes a következő becslés:

$$\boxed{\frac{1}{1 + \| A \|} \leq \| (I + A)^{-1} \| \leq \frac{1}{1 - \| A \|}}$$

Bizonyítás

Az $(I + A)^{-1}$ létezik, ugyanis tekintsük:

$$\| (I + A)\,\underline{\mathbf{x}} \| = \| \underline{\mathbf{x}} + A\underline{\mathbf{x}} \| \geq |\, \| \underline{\mathbf{x}} \| - \| -A\underline{\mathbf{x}} \| \,| \geq \| \underline{\mathbf{x}} \| - \| A\underline{\mathbf{x}} \| \geq \| \underline{\mathbf{x}} \| - \| A \| \cdot \| \underline{\mathbf{x}} \| =$$

$$= \underbrace{(1 - \| A \|)}_{\text{pozitív, ha } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}} \| \underline{\mathbf{x}} \| > 0$$

Tehát $(I + A)\,\underline{\mathbf{x}} = \underline{\mathbf{0}} \iff \underline{\mathbf{x}} = \underline{\mathbf{0}}$, ami azt jelenti, hogy $(I + A)^{-1}$ létezik.

Legyen $C := (I + A)^{-1}$.

$$1 = \| (I + A)\,(I + A)^{-1} \| = \| (I + A)\,C \| = \| C + A\,C \| \geq \| C \| - \| A \| \cdot \| C \| = (1 - \| A \|)\, \| C \|$$

$$\Longrightarrow \qquad \frac{1}{1 - \| A \|} \geq \| C \| \qquad \text{a felső becslés kész.}$$

$$1 = \| \, (I+A)\,(I+A)^{-1} \, \| \; = \; \| \, (I+A)\,C \, \| \; = \; \| \, C + A\,C \, \| \; \leq \; \| \, C \, \| + \| \, A \, \| \cdot \| \, C \, \| \; = (1 + \| \, A \, \| \,) \, \| \, C \, \|$$

$$\Longrightarrow \qquad \frac{1}{1 + \| \, A \, \|} \leq \| \, C \, \| \; = \; \| \, (I+A)^{-1} \, \| \; , \quad \text{ami az alsó becslés.}$$

■

Következmény

Hasonló becslés érvényes az $(I - A)^{-1}$ mátrix esetén is, azaz:

$$\boxed{\frac{1}{1 + \| \, A \, \|} \leq \| \, (I - A)^{-1} \, \| \leq \frac{1}{1 - \| \, A \, \|}}$$

Perturbációs lemma

Legyen $A\,,\, B \in \mathbb{K}^{n \times n}$, A reguláris $\| \, A^{-1} \, \| \cdot \| \, B - A \, \| \leq \varkappa < 1$; természetes mátrixnormát veszünk.
Ekkor B^{-1} mátrix létezik és érvényes a következő becslés:

$$\boxed{\| \, B^{-1} \, \| \leq \frac{\| \, A^{-1} \, \|}{1 - \| \, A^{-1} \, \| \cdot \| \, B - A \, \|} \leq \frac{\alpha}{1 - \varkappa}}$$

Bizonyítás

Mivel $\| \, A^{-1} \, (B - A) \, \| \leq \| \, A^{-1} \, \| \cdot \| \, B - A \, \| < 1$

a lemmát alkalmazva $\Longrightarrow$ $\left[A^{-1} \, (B - A) \right]^{-1}$ létezik $\Longrightarrow$ $\left[A^{-1} B \right]^{-1}$ létezik

$\Longrightarrow$ B^{-1} is létezik.

$B^{-1} A = Y \,, \quad B^{-1} = Y A^{-1}$

$\| \, B^{-1} \, \| \; = \; \| \, B^{-1} A \, A^{-1} \, \| \; \leq \; \| \, B^{-1} A \, \| \cdot \| \, A^{-1} \, \|$

$$\| \, B^{-1} A \, \| \; = \; \| \, \left[A^{-1} B \right]^{-1} \, \| \; = \; \| \, \left[I + A^{-1} \, (B - A) \right]^{-1} \, \| \leq \frac{1}{1 - \| \, A^{-1} \, (B - A) \, \|} \leq \frac{1}{1 - \| \, A^{-1} \, \| \cdot \| \, B - A \, \|}$$

$$\Longrightarrow \qquad \boxed{\| \, B^{-1} \, \| \leq \frac{\| \, A^{-1} \, \|}{1 - \| \, A^{-1} \, \| \cdot \| \, B - A \, \|} \leq \frac{\alpha}{1 - \varkappa}}$$

■

4.2. AZ A MÁTRIX PERTURBÁCIÓJÁNAK HATÁSA

Tekintsük:

$(\star)$ $\quad (A + \Delta A)\,(\underline{\mathbf{x}} + \Delta\underline{\mathbf{x}}) = \underline{\mathbf{b}}\,, \quad A\,,\ \Delta A \in \mathbb{K}^{n\times n} \quad \underline{\mathbf{x}}\,,\ \Delta\underline{\mathbf{x}}\,,\ \underline{\mathbf{b}} \in \mathbb{K}^n.$

Tétel

Legyen $\|A^{-1}\| \cdot \|\Delta A\| < 1$.

Ekkor a $(\star)$ megoldásának hibája a következő módon becsülhető:

$$\boxed{\frac{\|\Delta\underline{\mathbf{x}}\|}{\|\underline{\mathbf{x}}\|} \le \frac{\|A^{-1}\| \cdot \|A\|}{1 - \|A^{-1}\| \cdot \|A\| \dfrac{\|\Delta A\|}{\|A\|}} \cdot \frac{\|\Delta A\|}{\|A\|}}$$

Bizonyítás

Legyen $B := A + \Delta A$.

Így $A^{-1}(\Delta A) = A^{-1}(B - A) \quad \Longrightarrow \quad \|A^{-1}\| \cdot \|B - A\| < 1$

A perturbációs lemmát alkalmazva $\quad \Longrightarrow \quad \exists \quad B^{-1} = (A + \Delta A)^{-1}$ és igaz, hogy:

$$\boxed{\|(A + \Delta A)^{-1}\| \le \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|}}$$

Most $(\star)$ - ot rendezve $\quad (A + \Delta A)\,\Delta\underline{\mathbf{x}} = \underline{\mathbf{b}} - (A + \Delta A)\,\underline{\mathbf{x}} = \Delta A\underline{\mathbf{x}}$

$$\Delta\underline{\mathbf{x}} = (A + \Delta A)^{-1}\,\Delta A\underline{\mathbf{x}}$$

$$\frac{\|\Delta\underline{\mathbf{x}}\|}{\|\underline{\mathbf{x}}\|} \le \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|}\,\|\Delta A\| = \frac{\|A^{-1}\| \cdot \|A\|}{1 - \|A^{-1}\| \cdot \|A\| \dfrac{\|\Delta A\|}{\|A\|}} \cdot \frac{\|\Delta A\|}{\|A\|} =$$

$$= \frac{\operatorname{cond}(A)}{1 - \operatorname{cond}(A)\dfrac{\|\Delta A\|}{\|A\|}} \cdot \frac{\|\Delta A\|}{\|A\|}$$

■

Megjegyzés :

Ehhez a becsléshez kell $\| A^{-1} \|$ - ra felső korlát, ami általában nem ismert vagy nehezen számolható. Ezért hasznos a következő

ÉSZREVÉTEL :

Ha a korábbiak alapján már van egy LU vagy QR felbontásunk A - ra, amelyekre a $D := LU - A$ vagy $D := QR - A$ általában kicsi: $\| D \| \ll 1$, már könnyen meghatározhatjuk az inverzeiket. De akkor a perturbációs lemma szerint

$$\| A^{-1} \| \leq \frac{\alpha}{1-\varkappa} ,$$

ahol α olyan, hogy $\| D \| \leq \dfrac{\varkappa}{\alpha}$ és még $\| (LU)^{-1} \| \leq \alpha$, vagy $\| (QR)^{-1} \| \leq \alpha$

De :

Az LU felbontásban előfordulhat, hogy mindkét mátrix rosszul kondícionált lesz, holott A jól kondícionált ($\mathrm{cond}\,(A) \sim 1$). Ez azt jelenti, hogy a Gauss - elimináció még jól kondícionált feladatra is adhat gyenge eredményt (nagy hibával).

A Householder - féle QR felbontásnál más a helyzet! Például a $\| . \|_2$ vektornormával és a neki megfelelő mátrixnormával dolgozva, a Q mátrix ortogonalitása miatt

$$\| Q\underline{\mathbf{x}} \|_2 = \| \underline{\mathbf{x}} \|_2 , \quad \forall \quad \underline{\mathbf{x}} \in \mathbb{C}^n , \quad \text{és így} \quad \| Q \|_2 = 1$$

Hasonlóan: $\| Q^T \|_2 = \| Q^{-1} \|_2 = 1,$ és ezért $\mathrm{cond}\,(Q) = 1$

Másrészt $\| A \|_2 = \| Q^T Q A \|_2 \leq \| QA \|_2 \leq \| A \|_2$, és ezért $\| A \|_2 = \| QA \|_2$

Hasonlóan: $\| A^{-1} Q^{-1} \|_2 = \| A^{-1} \|_2$, amivel a következőt láttuk be:

Állítás

Az A mátrix kondíciószáma változatlan marad, ha a QR algoritmust használjuk, azaz:

$$\mathrm{cond}_2\,(A) = \mathrm{cond}_2\,(QA)$$

Ez is mutatja a QR algoritmus stabilitását!

Megjegyzés :

A fenti tétel általánosítható.

Az $\quad A\underline{\mathbf{x}} = \underline{\mathbf{b}} \qquad (A + \Delta A)\,(\underline{\mathbf{x}} + \Delta \underline{\mathbf{x}}) = \underline{\mathbf{b}} + \Delta \underline{\mathbf{b}}$

egyenletek megoldásaira igazolható a következő becslés:

$$\boxed{\frac{\|\,\Delta \underline{\mathbf{x}}\,\|}{\|\,\underline{\mathbf{x}}\,\|} \leq \frac{\mathrm{cond}\,(A)}{1 - \mathrm{cond}\,(A)\,\dfrac{\|\,\Delta A\,\|}{\|\,A\,\|}} \cdot \left(\frac{\|\,\Delta A\,\|}{\|\,A\,\|} + \frac{\|\,\Delta \underline{\mathbf{b}}\,\|}{\|\,\underline{\mathbf{b}}\,\|} \right)}$$

Példa :

A Hilbert - mátrix rosszul kondícionált!

$$A = (a_{ij})\ , \quad a_{ij} := \frac{1}{i+j-1} \qquad (i, j = 1, \ldots, n)$$

Ha $\quad b_i = \displaystyle\sum_{k=1}^{n} \frac{1}{i+k-1}\ ,$ akkor $\ A\underline{\mathbf{x}} = \underline{\mathbf{b}}\ $ megoldása $\ \underline{\mathbf{x}} = [1, \ldots, 1]^T\ .$

$n = 10$ - re már minden megoldás rossz eredményt ad.

5. LINEÁRIS EGYENLETRENDSZEREK ITERÁCIÓS MEGOLDÁSI MÓDSZEREI

Főleg akkor hasznosak, ha

(1) $\qquad A\underline{\mathbf{x}} = \underline{\mathbf{b}}\,, \quad A \in \mathbb{K}^{n\times n}\,, \quad \underline{\mathbf{x}}\ ,\ \underline{\mathbf{b}} \in \mathbb{K}^n$

egyenletrendszerben A ritka és nagy méretű mátrix.

Az *(1)* - et először vele ekvivalens

(2) $\qquad \underline{\mathbf{x}} = C\underline{\mathbf{x}} + \underline{\mathbf{c}}$

alakra hozzuk, úgynevezett ,,fixpont egyenlet”. Ez

(3) $\qquad \underline{\mathbf{x}} = \Phi\,(\underline{\mathbf{x}})$ alakú, ahol $\ \Phi\,(\underline{\mathbf{x}}) = C\underline{\mathbf{x}} + \underline{\mathbf{c}}\ .$

AZ ITERÁCIÓ ELVE :

Kiindulva egy tetszőleges $\underline{\mathbf{x}}_0 \in \mathbb{K}^n$ kezdővektorból, képezzük az

$\underline{\mathbf{x}}_{k+1} = C\underline{\mathbf{x}}_k + \underline{\mathbf{c}}$ úgynevezett ,,közelítő megoldások" sorozatát.

Ha $\underline{\mathbf{x}}_k$ sorozat konvergens és tart $\underline{\mathbf{x}}^*$ - hoz, akkor a *(3)* - nak, *(2)* - nek és így *(1)* - nek is a megoldása lesz.

KÉRDÉSEK :

A.)
Milyen feltételek biztosítják, hogy tetszőleges $\underline{\mathbf{x}}_0$ - ból kiindulva az $\{\underline{\mathbf{x}}_k\}$ sorozat konvergens? (KONVERGENCIA KÉRDÉSE)

B.)
Hogyan becsülhető a közelítő $(\underline{\mathbf{x}}_k)$ megoldások hibája? (HIBABECSLÉS)

C.)
Milyen gyors a konvergencia? (KONVERGENCIASEBESSÉG)

5.1. KONVERGENCIA

Tekintsük az $\underline{\mathbf{x}}_{k+1} = C\underline{\mathbf{x}}_k + \underline{\mathbf{c}}$ sorozatot tetszőleges $\underline{\mathbf{x}}_0$ esetén.

Legyen $\underline{\mathbf{x}}^* = C\underline{\mathbf{x}}^* + \underline{\mathbf{c}}$ pontos megoldás.

Vizsgáljuk meg:

$$\underline{\mathbf{x}}_{k+1} - \underline{\mathbf{x}}^* = C\underline{\mathbf{x}}_k + \underline{\mathbf{c}} - C\underline{\mathbf{x}}^* - \underline{\mathbf{c}} = C\left(\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*\right) = C^2\left(\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}^*\right) = \ldots = C^{k+1}\left(\underline{\mathbf{x}}_0 - \underline{\mathbf{x}}^*\right) \overset{?}{\to} 0$$

$$\Longrightarrow \qquad \{\underline{\mathbf{x}}_k\} \text{ sorozat konvergens} \qquad \Longleftrightarrow \qquad \lim_{k\to\infty} C^k = 0$$

$\boxed{\textbf{Tétel}}$ **Konvergenciakritérium**

Legyen $C \in \mathbb{K}^{n\times n}$, $\lim\limits_{k\to\infty} C^k = 0 \iff \rho(C) < 1$, azaz a C mátrix spektrálsugara < 1.

$\boxed{\textbf{Bizonyítás}}$

$\boxed{\Rightarrow}$

Tegyük fel, hogy $\lim\limits_{k\to\infty} C^k = 0$ és bizonyítjuk, hogy $\rho(C) < 1$, azaz, hogy a C mátrix minden sajátértéke kisebb, mint 1 (abszolútértékben).

Indirekt.

Tegyük fel, hogy létezik olyan λ sajátértéke a C mátrixnak, amelyre $|\lambda| \geq 1$.

Mivel λ a C sajátértéke $\Longrightarrow$ $\exists$ $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, hogy:

$$\begin{aligned} C\underline{\mathbf{x}} &= \lambda\underline{\mathbf{x}} \\ C^2\underline{\mathbf{x}} &= \lambda C\underline{\mathbf{x}} = \lambda^2\underline{\mathbf{x}} \\ C^k\underline{\mathbf{x}} &= \lambda^k\underline{\mathbf{x}} \end{aligned}$$

Mivel $\lim \lambda^k \neq 0$ $\Longrightarrow$ $\left\{C^k\right\}$ nem lehet nullsorozat, tehát $\rho(C) < 1$ szükséges feltétel.

$\boxed{\Leftarrow}$

Tegyük fel, hogy $\rho(C) < 1$ és bizonyítjuk, hogy $\lim\limits_{k\to\infty} C^k = 0$.

Tudjuk, hogy létezik olyan T reguláris mátrix, hogy a C mátrix a T segítségével Jordan - alakra transzformálható:

$$J = TCT^{-1} = \begin{bmatrix} J_1 & 0 & \cdots & & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & J_i & & \\ & & & \ddots & 0 \\ 0 & \cdots & & 0 & J_m \end{bmatrix}, \quad \text{ahol}$$

J_i $(n_i \times n_i)$ - es mátrix, amelynek alakja:

$$J_i := \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & & \ddots & & \vdots \\ \vdots & & \ddots & & 0 \\ & & & & 1 \\ 0 & \cdots & & 0 & \lambda_i \end{bmatrix}$$

Mivel:

$$J^k = \left(TCT^{-1}\right)^k = TC^kT^{-1} \qquad \Longrightarrow \qquad J^2 = \left(TCT^{-1}\right)\left(TCT^{-1}\right) = TC^2T^{-1}$$

$$\Longrightarrow \qquad C^k = T^{-1}J^kT$$

Ezért $\lim\limits_{k\to\infty} C^k = 0$ $\Longleftrightarrow$ $\lim\limits_{k\to\infty} J^k = 0$

Viszont:

$$J^k = \begin{bmatrix} J_1^k & 0 & \cdots & & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & J_i^k & & \\ & & & \ddots & 0 \\ 0 & \cdots & & 0 & J_m^k \end{bmatrix},$$

ezért elég belátni, hogy $\lim\limits_{k\to\infty} J_i^k = 0$, azaz az egyes Jordan - blokkok 0 - hoz való tartását $\forall \quad i = 1, \ldots, m$ - re.

A J_i $(n_i \times n_i)$ - es mátrix a következő alakban írható:

$$J_i = \lambda_i I + S_i \;, \quad \text{ahol} \quad I\,,\, S_i \;\, (n_i \times n_i) \text{ - es mátrixok és} \qquad S_i := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & & \ddots & & \vdots \\ \vdots & & \ddots & & 0 \\ & & & & 1 \\ 0 & & \cdots & & 0 \end{bmatrix}$$

Tudjuk (Gyapjas jegyzet 125. oldal), hogy S_i mátrix n_i - edrendben nilpotens mátrix, azaz $S_i^{n_i} = 0$.

Így J_i^k alakja (Newton binomiális formulája alapján):

$$J_i^k = (\lambda_i I + S_i)^k = \sum_{j=0}^{k} \binom{k}{j} \lambda_i^{k-j} S_i^j = \sum_{j=0}^{n_i - 1} \binom{k}{j} \lambda_i^{k-j} S_i^j$$

Fix j - re a $\sum$ n_i darab tagból áll, és fix i - re az egyes tagokban S_i^j mátrix nem változik $k \to \infty$ esetén.

Kérdés : $\lim\limits_{k\to\infty} J^k \stackrel{?}{=} 0$

Ha belátjuk, hogy $\lim\limits_{k\to\infty} \binom{k}{j} \lambda_i^{k-j} = 0$, akkor a $\sum$ minden tagja tart 0 - hoz, és ekkor $\lim\limits_{k\to\infty} J^k = 0$ - t is beláttuk.

Továbbá $\left| \binom{k}{j} \lambda_i^{k-j} \right| \leq \left| k^j \right| \cdot \left| \lambda_i^{k-j} \right| \xrightarrow[k\to\infty]{} 0 \,, \qquad \lambda_i < 1$

$\boxed{\text{HF.}}$ $\lim\limits_{k\to\infty} \lambda^k k^j = 0 \,, \quad 0 < \lambda < 1 \;, \quad \text{fix} \quad \lambda \,,\, j$ - re.

■

Tétel **Konvergencia - tétel**

Legyen $A \in \mathbb{K}^{n\times n}$.

Ekkor az $\underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$ alakú iterációs sorozat tetszőleges $\underline{\mathbf{x}}_0$ kezdővektorból kiindulva konvergál

a ***(2)*** $(\underline{\mathbf{x}} = C\underline{\mathbf{x}} + \underline{\mathbf{c}})$ illetve ***(1)*** $(A\underline{\mathbf{x}} = \underline{\mathbf{b}})$ pontos $\underline{\mathbf{x}}^*$ megoldásához $\iff$ $\rho\,(C) < 1$

Lemma

Legyen $C \in \mathbb{K}^{n\times n}$.

Ekkor $\rho\,(C) \leq \|\,C\,\|$ valamilyen természetes mátrixnormára.

Bizonyítás

Legyen λ a C mátrix tetszőleges sajátértéke és a λ - hoz tartozó sajátvektor $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$.

Ekkor $C\underline{\mathbf{x}} = \lambda\underline{\mathbf{x}}$.

Természetes mátrixnormát véve:

$\|\,\lambda\underline{\mathbf{x}}\,\| = \|\,C\underline{\mathbf{x}}\,\|$

$|\lambda| \cdot \|\,\underline{\mathbf{x}}\,\| = \|\,C\underline{\mathbf{x}}\,\| \leq \|\,C\,\| \cdot \|\,\underline{\mathbf{x}}\,\|$

$\underline{\mathbf{x}} \neq \underline{\mathbf{0}} \quad\Longrightarrow\quad \|\,\underline{\mathbf{x}}\,\| \neq 0 \quad\Longrightarrow\quad |\lambda| \leq \|\,C\,\|$

Tetszőleges λ sajátértékre igaz, így $\Longrightarrow$ $\rho\,(C) \leq \|\,C\,\|$

■

Tétel **Elégséges feltétel**

Az $\underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$ iterációs sorozat konvergenciájának elégséges feltétele az, hogy valamilyen természetes mátrixnormára teljesüljön, hogy $\|\,C\,\| < 1$ tetszőleges $\underline{\mathbf{x}}_0$ - ból kiindulva.

5.2. HIBABECSLÉS

Legyen

$$\left.\begin{array}{ll} \textit{(1)} & A\underline{\mathbf{x}} = \underline{\mathbf{b}} \\[2ex] \textit{(2)} & \underline{\mathbf{x}} = C\underline{\mathbf{x}} + \underline{\mathbf{c}} \end{array}\right\} \quad \underline{\mathbf{x}}^* \quad \text{legyen a pontos megoldás}$$

Kicsit általánosabban: Legyen $C := C_1 + C_2$ és tekintsük a következő iterációs sorozatot:

(3) $\underline{\mathbf{x}}_k = -c_1\underline{\mathbf{x}}_k + C_2\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}} \qquad (\underline{\mathbf{x}}^* = C_1\underline{\mathbf{x}}^* + C_2\underline{\mathbf{x}}^* + \underline{\mathbf{c}})$

Tegyük fel, hogy tetszőleges $\underline{\mathbf{x}}_0$ - ból kiindulva a *(3)* konvergens sorozat.

Tétel **A - POSTERIORI HIBABECSLÉS**

Ha $\| C \| < 1$ valamilyen természetes mátrixnormára, akkor érvényes a következő hibabecslés:

(4) $$\boxed{\| \underline{\mathbf{x}}_k - \underline{\mathbf{x}}^* \| \leq \frac{\| C_2 \|}{1 - \| C \|} \| \underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1} \|}$$

Bizonyítás

Vizsgáljuk:

$$\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^* = C_1\underline{\mathbf{x}}_k + C_2\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}} - (c_1\underline{\mathbf{x}}^* + C_2\underline{\mathbf{x}}^* + \underline{\mathbf{c}}) = C_1 (\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*) + C_2 \left(\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}^*\right)$$

$$(I - C_1 - C_2) (\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*) = C_2 \left(\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}^*\right) - C_2 (\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*) = C_2 \left(\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}_k\right)$$

$$(I - C) (\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*) = C_2 \left(\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}_k\right)$$

Mivel $\| C \| < 1 \quad \Longrightarrow \quad \exists \quad (I - C)^{-1}$

$\Bigg($Lemma és következménye a 60. oldalon. $\quad \Longrightarrow \quad \| (I - C)^{-1} \| \leq \dfrac{1}{1 - \| C \|}\Bigg)$

$\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^* = (I - C)^{-1} C_2 \left(\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}_k\right)$ Normára térve:

$$\| \underline{\mathbf{x}}_k - \underline{\mathbf{x}}^* \| \leq \| (I - C)^{-1} \| \cdot \| C_2 \| \cdot \| \underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1} \| \leq \| \underline{\mathbf{x}}_k - \underline{\mathbf{x}}^* \| \leq \frac{\| C_2 \|}{1 - \| C \|} \| \underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1} \|$$

$\blacksquare$

Megjegyzés :

A *(4)* hibabecslés érvényes az $\underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$ alakú iterációs eljárásokra is, ugyanis ez következik *(3)* - ból $C_1 = 0\,,\quad C_2 = C$ választásával.

Ekkor *(4)* alakja:

(5)
$$\boxed{\|\,\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*\,\| \le \frac{\|\,C\,\|}{1-\|\,C\,\|}\,\|\,\underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1}\,\|}$$

A *(4)* és *(5)* jól alkalmazható, azonban gyakorlati esetben figyelembe kell vennünk a kerekítési hibákat is.

Következmény **A - PRIORI HIBABECSLÉS**

$\underline{\mathbf{x}} = C\underline{\mathbf{x}} + \underline{\mathbf{c}}$ alakú iterációra, $\|\,C\,\| < 1$.

$$\boxed{\|\,\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*\,\| \le \frac{\|\,C\,\|^{\,k}}{1-\|\,C\,\|}\,\|\,\underline{\mathbf{x}}_1 - \underline{\mathbf{x}}_0\,\|}$$

Bizonyítását lásd később, a nemlineáris egyenletek megoldásánál, a Banach - féle fixponttétel után.

5.3. JACOBI ITERÁCIÓ

Legyen

(1) $A\underline{\mathbf{x}} = \underline{\mathbf{b}}\,,\quad A \in \mathbb{R}^{n\times n}\,,\quad A$ nem szinguláris, $\underline{\mathbf{b}} \in \mathbb{R}^n$.

Bontsuk fel A - t a következő módon:

$A = -L + D - R\,,$ ahol

$$L := -\begin{bmatrix} 0 & \cdots & & 0 \\ a_{21} & \ddots & & \vdots \\ \vdots & & & \\ a_{n1} & \cdots & a_{n\,n-1} & 0 \end{bmatrix}, \quad D := \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}, \quad R := -\begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \ddots & & \vdots \\ & & & a_{n-1\,n} \\ 0 & \cdots & & 0 \end{bmatrix}$$

(1) $\Longrightarrow \quad (-L + D - R)\,\underline{\mathbf{x}} = \underline{\mathbf{b}}$

$D\underline{\mathbf{x}} = (L + R)\,\underline{\mathbf{x}} + \underline{\mathbf{b}}$

Ha A reguláris mátrix, akkor mindig elérhető, hogy D nem szinguláris (sorcserékkel).

(2) $\quad \underline{\mathbf{x}} = D^{-1}(L+R)\,\underline{\mathbf{x}} + D^{-1}\underline{\mathbf{b}} \qquad (\underline{\mathbf{x}} = \Phi(\underline{\mathbf{x}})$ alakú$)$

Így tetszőleges $\underline{\mathbf{x}}_0$ - ból kiindulva képezhetjük a közelítő megoldások sorozatát.

(3) $\quad \boxed{\underline{\mathbf{x}}_k = D^{-1}(L+R)\,\underline{\mathbf{x}}_{k-1} + D^{-1}\underline{\mathbf{b}}}$ **Jacobi iteráció**

Koordinátánként írva :

$$\boxed{x_i^{[k]} = \frac{1}{a_{ii}}\left[b_i - \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{[k-1]}\right]} \qquad (i = 1, \dots, n)$$

„SZIMULTÁN VÁLTOZTATÁSOK MÓDSZERE”

Most $\quad C := D^{-1}(L+R)\ , \quad \underline{\mathbf{c}} := D^{-1}\underline{\mathbf{b}}$

$\underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$ alakú.

5.3.1. Konvergencia kritérium

(3) Jacobi iteráció tetszőleges $\underline{\mathbf{x}}_0$ kezdőérték esetén konvergens $\iff \rho\left(D^{-1}(L+R)\right) < 1$

Elégséges feltétel a módszer konvergenciájára : $\| D^{-1}(L+R) \| < 1$ valamilyen természetes mátrixnormára

Definíció :

Az A mátrix

erős sor-összeg tulajdonságú, ha: $\displaystyle\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| < |a_{ii}| \quad \forall \quad i$ - re,

erős-oszlopösszeg tulajdonságú, ha: $\displaystyle\sum_{\substack{i=1\\ i\neq j}}^{n} |a_{ij}| < |a_{jj}| \quad \forall \quad j$ - re.

Tétel **Elégséges feltétel a Jacobi módszer konvergenciájára :**

Ha az A mátrixra az erős sor- , vagy oszlop-összeg tulajdonság teljesül, akkor *(3)* Jacobi iteráció konvergál tetszőleges $\underline{\mathbf{x}}_0$ esetén.

Ugyanis ekkor a $D^{-1}(L+R)$ mátrix sor- vagy oszlopnormája < 1 lesz!

5.4. GAUSS - SEIDEL ITERÁCIÓ

A *(3)* formulát vizsgálva látható, hogy a k - adik iterált vektor i - edik koordinátájának kiszámításához már felhasználhatjuk az előző $1, \ldots, i-1$ koordinátáit a k - adik iteráltnak. Így kapjuk a Gauss - Seidel iterációt.

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

$$(-L + D - R)\,\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

$$D\underline{\mathbf{x}} = (L + R)\,\underline{\mathbf{x}} + \underline{\mathbf{b}}$$

$$\underline{\mathbf{x}} = D^{-1}\,(L\underline{\mathbf{x}} + R\underline{\mathbf{x}}) + D^{-1}\underline{\mathbf{b}}$$

(4) $\boxed{\underline{\mathbf{x}}_k = D^{-1}\left(L\underline{\mathbf{x}}_k + R\underline{\mathbf{x}}_{k-1}\right) + D^{-1}\underline{\mathbf{b}}}$ **Gauss - Seidel iteráció**

Koordinátánként írva :

$$\boxed{x_i^{[k]} = \frac{1}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{[k]} - \sum_{j=i+1}^{n} a_{ij}x_j^{[k-1]}\right]}$$

„FOKOZATOS VÁLTOZTATÁSOK MÓDSZERE”

5.4.1. Konvergencia kritérium

A Gauss - Seidel iteráció vizsgálatához *(4)* - et átírjuk:

$$D\underline{\mathbf{x}}_k - L\underline{\mathbf{x}}_k = R\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{b}}$$

$$(D - L)\,\underline{\mathbf{x}}_k = R\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{b}}$$

$$\underline{\mathbf{x}}_k = (D - L)^{-1}\,\underline{\mathbf{x}}_{k-1} + (D - L)^{-1}\,\underline{\mathbf{b}}\ ,$$

ahol $\quad C := (D - L)^{-1}\,R\,, \qquad \underline{\mathbf{c}} := (D - L)^{-1}\,\underline{\mathbf{b}}$

A Gauss - Seidel iteráció tetszőleges $\underline{\mathbf{x}}_0$ kezdőérték esetén konvergál $\iff$ $\rho\Big((D - L)^{-1}\,R\Big) < 1$

Elégséges feltétel a Gauss - Seidel iteráció konvergenciájára :

$\|\,(D - L)^{-1}\,R\,\| < 1$ valamilyen természetes mátrixnormára.

Tétel

Ha az A mátrixra az erős sor- vagy oszlop-összeg tulajdonság teljesül, akkor a Gauss - Seidel iteráció is konvergál tetszőleges $\underline{\mathbf{x}}_0$ - ból kiindulva és a konvergencia legalább olyan gyors, mint a Jacobi iterációnál.

5.5. RELAXÁCIÓS MÓDSZEREK

5.5.1. Jacobi relaxáció

A Jacobi iteráció:

$$\underline{\mathbf{x}}_k = D^{-1}(L+R)\,\underline{\mathbf{x}}_{k-1} + D^{-1}\underline{\mathbf{b}} \qquad / \pm \left(\underline{\mathbf{x}}_{k-1}\right)$$

$$\underline{\mathbf{x}}_k = \underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}_{k-1} + D^{-1}(L+R)\,\underline{\mathbf{x}}_{k-1} + D^{-1}\underline{\mathbf{b}}$$

$$\underline{\mathbf{x}}_k = \underline{\mathbf{x}}_{k-1} + D^{-1}(L-D+R)\,\underline{\mathbf{x}}_{k-1} + D^{-1}\underline{\mathbf{b}}$$

$$\underline{\mathbf{x}}_k = \underline{\mathbf{x}}_{k-1} - D^{-1}\underbrace{\left[A\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{b}}\right]}_{\text{ez a } (k-1) \text{ - edik lépés defektus (eltérés) vektora}}$$

A Jacobi relaxáció interpretációja: a k - adik iteráltat a $(k-1)$ - edik iterált vektorból a defektus vektor segítségével kapjuk. Természetes dolog egy úgynevezett ω relaxációs paraméter bevezetése.

$$\Longrightarrow \qquad \underline{\mathbf{x}}_k = \underline{\mathbf{x}}_{k-1} - \omega D^{-1}\left[A\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{b}}\right]$$

$$\boxed{\underline{\mathbf{x}}_k = \omega D^{-1}[L+R]\,\underline{\mathbf{x}}_{k-1} + (1-\omega)\,\underline{\mathbf{x}}_{k-1} + \omega D^{-1}\underline{\mathbf{b}}} \qquad \textbf{Jacobi relaxáció}$$

Koordinátás alak :

$$\boxed{x_i^{[k]} = \frac{\omega}{a_{ii}}\left[b_i - \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{[k-1]}\right] + (1-\omega)\,x_j^{[k-1]}}$$

Tétel

Legyen a Jacobi iteráció konvergens.
Ekkor a Jacobi relaxáció is konvergens $\quad \forall \quad 0 < \omega \leq 1$ - re.

5.5.2. Gauss - Seidel relaxáció

Tekintsük a következő lineáris egyenletrendszert:

(1) $A\underline{\mathbf{x}} = \underline{\mathbf{b}} \quad / \cdot \omega \quad \omega$ paraméter $\quad (A = -L + D - R)$

$\omega A\underline{\mathbf{x}} = \omega\underline{\mathbf{b}} \quad / + D\underline{\mathbf{x}}$

$D\underline{\mathbf{x}} = D\underline{\mathbf{x}} - \omega A\underline{\mathbf{x}} + \omega\underline{\mathbf{b}} = D\underline{\mathbf{x}} + \omega\left(\underline{\mathbf{b}} - A\underline{\mathbf{x}}\right)$

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

$\underline{\mathbf{x}}_k = \underline{\mathbf{x}}_{k-1} + \omega D^{-1}\left(\underline{\mathbf{b}} - A\underline{\mathbf{x}}_k\right)$ Jacobi - relaxáció

$\underline{\mathbf{x}} = \underline{\mathbf{x}} + \omega D^{-1}\left(\underline{\mathbf{b}} - A\underline{\mathbf{x}}\right) = \underline{\mathbf{x}} + \omega D^{-1}\left(\underline{\mathbf{b}} + L\underline{\mathbf{x}} - D\underline{\mathbf{x}} + R\underline{\mathbf{x}}\right)$

$\underline{\mathbf{x}} = \omega D^{-1}\left(\underline{\mathbf{b}} + L\underline{\mathbf{x}} + R\underline{\mathbf{x}}\right) + (1-\omega)\,\underline{\mathbf{x}} \quad \Longrightarrow$ újra visszakapjuk a Jacobi relaxációt $\Longrightarrow$

$$\boxed{\underline{\mathbf{x}}_k = \omega D^{-1}\left(\underline{\mathbf{b}} + L\underline{\mathbf{x}}_{k-1} + R\underline{\mathbf{x}}_{k-1}\right) + (1-\omega)\,\underline{\mathbf{x}}_{k-1}} \qquad \textbf{Jacobi relaxáció}$$

Koordinátás alak :

$$\boxed{x_i^{[k]} = \frac{\omega}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{[k-1]} - \sum_{j=i+1}^{n} a_{ij}x_j^{[k-1]}\right] + (1-\omega)\,x_i^{[k-1]}} \qquad i = 1, \ldots, n$$

Ha az első $\sum$ - ba a k - adik iterált vektor már kiszámolt koordinátáit írjuk, megkapjuk az úgynevezett **Gauss - Seidel relaxációt**:

$$(2) \qquad \boxed{\underline{\mathbf{x}}_k = \omega D^{-1}\left[\underline{\mathbf{b}} + L\underline{\mathbf{x}}_k + R\underline{\mathbf{x}}_{k-1}\right] + (1-\omega)\,\underline{\mathbf{x}}_{k-1}}$$

Koordinátás alak :

$$\boxed{x_i^{[k]} = \frac{\omega}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{[k]} - \sum_{j=i+1}^{n} a_{ij}x_j^{[k-1]}\right] + (1-\omega)\,x_i^{[k-1]}} \qquad i = 1, \ldots, n$$

A Gauss - Seidel relaxáció konvergencia vizsgálatához *(2)* - t átalakítjuk $\underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$ formára. *(2)* - t rendezve:

$$\left(I - \omega D^{-1}L\right)\underline{\mathbf{x}}_k = \omega D^{-1}R\underline{\mathbf{x}}_{k-1} + (1-\omega)\,I\underline{\mathbf{x}}_{k-1} + \omega D^{-1}\underline{\mathbf{b}}$$

$$\left(I - \omega D^{-1}L\right)\underline{\mathbf{x}}_k = \left[(1-\omega)\,I + \omega D^{-1}R\right]\underline{\mathbf{x}}_{k-1} + \omega D^{-1}\underline{\mathbf{b}}$$

$$\Longrightarrow \qquad \underline{\mathbf{x}}_k = \left(I - \omega D^{-1}L\right)^{-1} \left[(1-\omega)\, I + \omega D^{-1}R\right] \underline{\mathbf{x}}_{k-1} + \left(I - \omega D^{-1}L\right)^{-1} \omega D^{-1}\underline{\mathbf{b}}$$

Ez most már $\underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$ alakú.

Konvergencia kritérium :

A Gauss - Seidel relaxáció konvergens tetszőleges $\underline{\mathbf{x}}_0$ - ból kiindulva $\Longleftrightarrow$

$$\Longleftrightarrow \qquad \rho\left(\left(I - \omega D^{-1}L\right)^{-1} \left[(1-\omega)\, I + \omega D^{-1}R\right]\right) < 1$$

Tétel

Legyen A szimmetrikus $(n \times n)$ - es mátrix.

$\forall \quad a_{ii} > 0\,, \quad i = 1, \ldots, n$

Ekkor a Gauss - Seidel relaxáció konvergens $\Longleftrightarrow$ A pozitív definit mátrix és $0 < \omega < 2$

Elégséges feltétel :

Ha az A szimmetrikus, pozitív definit mátrix és $0 < \omega < 2$ $\Longrightarrow$ a Gauss - Seidel relaxáció konvergens.

Ha $0 < \omega < 1$ „alulrelaxálás”,

ha $1 < \omega < 2$ „túlrelaxálás”.

Tétel

Legyen A szimmetrikus pozitív definit, tridiagonális mátrix. Ekkor a Gauss - Seidel relaxáció optimális ω paramétere a következő:

$$\boxed{\omega_{\text{opt}} = \frac{2}{1 + \sqrt{1 - \left[\rho\left(C_j\right)\right]^2}}}$$

ahol $\rho\left(C_j\right)$ az „eredeti” Jacobi iteráció C mátrixának spektrálsugara.

5.6. KÉTRÉTEGŰ ITERÁCIÓS ELJÁRÁSOK

Tekintsük az

(1) $\quad A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ lineáris egyenletrendszert.

Ebből az

(2) $\quad \underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$ iterációs eljáráshoz a következő módon is eljuthatunk:

Legyen $\quad A = P - Q$, ahol P reguláris mátrix.

(1) $\quad \Longrightarrow \quad P\underline{\mathbf{x}} = Q\underline{\mathbf{x}} + \underline{\mathbf{b}} \quad \Longrightarrow \quad \underline{\mathbf{x}} = P^{-1}Q\underline{\mathbf{x}} + P^{-1}\underline{\mathbf{b}}$, ami *(2)* alakú.

$$\underline{\mathbf{x}}_k = P^{-1}Q\underline{\mathbf{x}}_{k-1} + P^{-1}\underline{\mathbf{b}}, \quad C := P^{-1}Q, \quad \underline{\mathbf{c}} := P^{-1}\underline{\mathbf{b}}$$

$$\Longrightarrow \quad P\underline{\mathbf{x}}_k = Q\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{b}} \qquad / - P\underline{\mathbf{x}}_{k-1}$$

$$P\left(\underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1}\right) = -\left(P - Q\right)\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{b}}$$

(3) $$\boxed{P\left(\underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1}\right) = -A\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{b}}} \quad \overset{(2)}{\Longrightarrow} \quad \boxed{\underline{\mathbf{x}}_k = \left(I - P^{-1}A\right)\underline{\mathbf{x}}_{k-1} + P^{-1}\underline{\mathbf{b}}}$$

Az ilyen alakú iterációs módszereket kétrétegű iterációs módszereknek, P mátrixot prekondicionálási mátrixnak nevezzük $\quad \left(C := I - P^{-1}A\right)$.

A P mátrix megválasztásánál két, gyakran ellentétes szempontot kell figyelembe venni:

i.) A P mátrix legyen közel az A mátrixhoz, mert ekkor a $\quad C = I - P^{-1}A \quad$ normája kicsi, a konvergencia pedig gyors lesz.

ii.) A P mátrix legyen könnyen invertálható, mert különben minden iterációs lépés túl nagy műveletigénnyel járna.

Néhány példa P megválasztására :

1.

$P = D$, akkor visszakapjuk a Jacobi iterációt.

$$D\left(\underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1}\right) = -\left(-L + D - R\right)\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{b}}$$

$$\underline{\mathbf{x}}_k = D^{-1}\left(\underline{\mathbf{b}} + L\underline{\mathbf{x}}_{k-1} + R\underline{\mathbf{x}}_{k-1}\right)$$

2.

$P = (D - L)$, akkor a Gauss - Seidel iterációt kapjuk.

3.

$P = \dfrac{D}{\omega}$ vagy $P = \dfrac{D - L}{\omega}$, akkor pedig a Jacobi relaxációt, illetve a Gauss - Seidel relaxációt kapjuk vissza.

5.6.1. Tompított Jacobi módszer

A *(3)* iterációs módszer általánosításaként kapható, ugyanis a módszer kapcsolatba hozható a következő differenciálegyenlet-rendszerrel:

$$(4) \qquad P\frac{\mathrm{d}\underline{\mathbf{y}}}{\mathrm{d}t} = -A\underline{\mathbf{y}} + \underline{\mathbf{b}}, \quad A \in \mathbb{R}^{n\times n}, \quad \underline{\mathbf{y}}\ ,\ \underline{\mathbf{b}} \in \mathbb{R}^n, \quad \underline{\mathbf{y}}(0) = \underline{\mathbf{y}}_0 \in \mathbb{R}^n, \quad \underline{\mathbf{y}}(t) = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Belátható, hogy a *(4)* úgynevezett „stacionárius" megoldása (amikor $\underline{\mathbf{y}}$ már nem függ a t időtől, azaz amikor $\dfrac{\mathrm{d}\underline{\mathbf{y}}}{\mathrm{d}t} = 0 \quad (t \to \infty)$) az *(1)* lineáris egyenletrendszer megoldását adja.

A *(4)* - ből pedig *(3)* alakú iterációs módszert kaphatunk!

A *(4)* - ben a deriváltat különbségi hányadossal helyettesítve $\Longrightarrow$

$$\Longrightarrow \qquad P\frac{\underline{\mathbf{y}}(t_k) - \underline{\mathbf{y}}(t_{k-1})}{\Delta t} = -A\underline{\mathbf{y}}(t_{k-1}) + \underline{\mathbf{b}}$$

Legyen:

$$\left.\begin{aligned} &\underline{\mathbf{y}}(t_k) =: \underline{\mathbf{x}}_k \\ &\underline{\mathbf{y}}(t_{k-1}) =: \underline{\mathbf{x}}_{k-1} \\ &\Delta t := 1 \end{aligned}\right\} \quad \Longrightarrow \quad \boxed{P\left(\underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1}\right) = -A\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{b}}} \qquad (3)$$

Ha $\Delta t = \omega \quad \Longrightarrow \quad \boxed{P\left(\underline{\mathbf{x}}_k - \underline{\mathbf{x}}_{k-1}\right) = -\omega A\underline{\mathbf{x}}_{k-1} + \omega\underline{\mathbf{b}}}$

iterációs módszert kapjuk, amit tompított Jacobi módszernek nevezünk.

Ha $P = I$, akkor EXPLICIT TOMPÍTOTT JACOBI módszernek (RICHARDSON ITERÁCIÓ) nevezzük, egyébként IMPLICIT TOMPÍTOTT JACOBI módszernek.

Tétel

Legyen A szimmetrikus pozitív definit mátrix, $P = I$, $\underline{\mathbf{y}}(t)$ a *(4)* megoldása $\underline{\mathbf{y}}(0) = \underline{\mathbf{y}}_0$ kezdeti érték esetén.

Ekkor $\lim\limits_{t\to\infty} \| \underline{\mathbf{y}}(t) - \underline{\mathbf{x}} \| = 0$, ahol $\underline{\mathbf{x}}$ az *(1)* lineáris egyenletrendszer megoldása.

Bizonyítás

Legyen $\underline{\mathbf{z}}(t) := \underline{\mathbf{y}}(t) - \underline{\mathbf{x}}$ a hiba, ahonnan $\underline{\mathbf{y}}(t) = \underline{\mathbf{z}}(t) + \underline{\mathbf{x}}$

Ezt *(4)* - be írva $\Longrightarrow$ $\dfrac{\mathrm{d}\underline{\mathbf{z}}}{\mathrm{d}t} + A\underline{\mathbf{z}} = 0 \qquad t > 0, \quad \underline{\mathbf{z}}(0) = \underline{\mathbf{y}}(0) - \underline{\mathbf{x}}$

Majd $\underline{\mathbf{z}}$ - vel skalárisan szorozva: $\left(\dfrac{\mathrm{d}\underline{\mathbf{z}}}{\mathrm{d}t}, \underline{\mathbf{z}}\right) + (A\underline{\mathbf{z}}, \underline{\mathbf{z}}) = 0 \qquad (\star)$

Itt az első tag: $\left(\dfrac{\mathrm{d}\underline{\mathbf{z}}}{\mathrm{d}t}, \underline{\mathbf{z}}\right) = \dfrac{1}{2} \cdot \dfrac{\mathrm{d}}{\mathrm{d}t} (\underline{\mathbf{z}}, \underline{\mathbf{z}}) = \dfrac{1}{2} \cdot \dfrac{\mathrm{d}}{\mathrm{d}t} \| \underline{\mathbf{z}} \|^2$,

és mivel A pozitív definit $\Longrightarrow$ $(A\underline{\mathbf{z}}, \underline{\mathbf{z}}) \geq \delta \| \underline{\mathbf{z}} \|^2$ ahol $\delta > 0$,

(⋆) - ból $\Longrightarrow$ $\dfrac{\mathrm{d}}{\mathrm{d}t} \| \underline{\mathbf{z}}(t) \|^2 + 2\delta \| \underline{\mathbf{z}}(t) \|^2 \leq 0$

Ezt szorozva $e^{2\delta t} > 0$ - val $\Longrightarrow$ $\dfrac{\mathrm{d}}{\mathrm{d}t} e^{2\delta t} \| \underline{\mathbf{z}}(t) \|^2 \leq 0$,

ahonnan integrálva 0 - tól t - ig $\Longrightarrow$ $e^{2\delta t} \| \underline{\mathbf{z}}(t) \|^2 \leq \| \underline{\mathbf{z}}(0) \|^2$

Most gyököt vonva és $\underline{\mathbf{z}} = \underline{\mathbf{y}} - \underline{\mathbf{x}}$ - et visszaírva, $e^{\delta t}$ - vel osztva

$$\Longrightarrow \quad \| \underline{\mathbf{y}}(t) - \underline{\mathbf{x}} \| \leq e^{-\delta t} \| \underline{\mathbf{y}}(0) - \underline{\mathbf{x}} \| \xrightarrow[t\to\infty]{} 0$$

■

Tétel

Legyen A szimmetrikus pozitív definit, $P = I$.
Ekkor az egyszerű iteráció (Richardson iteráció) konvergens, ha

$\boxed{0 < \omega < \dfrac{2}{\rho(A)}}$ tetszőleges $\underline{\mathbf{x}}_0$ kezdővektor esetén.

Az optimális ω paraméter megválasztására igaz a következő tétel.

Tétel

Ha A szimmetrikus pozitív definit mátrix, akkor az egyszerű iterációra egyértelműen létezik ω_{opt} optimális paraméter, amelynek értéke:

$$\boxed{\omega_{\text{opt}} = \frac{2}{m+M}} \qquad \text{ahol} \qquad \begin{array}{l} m := \min \lambda(A) \\ M := \max \lambda(A) \end{array}$$

5.6.2. A kerekítési hibák hatása az iterációs eljárásokra

Vizsgáljuk meg, hogyan befolyásolják a kerekítési hibák az

(1) $\quad \underline{\mathbf{x}}_k = C\underline{\mathbf{x}}_{k-1} + \underline{\mathbf{c}}$

alakú iterációs eljárások közelítő megoldásainak sorozatát. Az iterációs módszerek gyakorlati alkalmazása során is kerekítjük a számított értékeket.

Tetszőleges $\underline{\mathbf{x}}_0$ - ból kiindulva az $\underline{\mathbf{x}}_1 = C\underline{\mathbf{x}}_0 + \underline{\mathbf{c}}$ pontos érték helyett a kerekítési hiba miatt $\underline{\mathbf{y}}_1 = \underline{\mathbf{x}}_1 + \underline{\varepsilon}_1$ értéket kapjuk, ahol $\underline{\varepsilon}_1$ az első lépésben elkövetett kerekítési hibavektor.

A második lépésben

$\underline{\mathbf{y}}_2 = C\underline{\mathbf{y}}_1 + \underline{\mathbf{c}} + \underline{\varepsilon}_2$, ahol $\underline{\varepsilon}_2$ a második lépésben elkövetett kerekítés hibája.

Így az *(1)* közelítő megoldássorozat helyett a következő sorozatot kapjuk:

(2) $\quad \underline{\mathbf{y}}_k = C\underline{\mathbf{y}}_{k-1} + \underline{\mathbf{c}} + \underline{\varepsilon}_k$, ahol $\underline{\varepsilon}_k$ a k - adik lépésbeli hibavektor, $\quad \underline{\mathbf{y}}_0 = \underline{\mathbf{x}}_0$, $\quad \underline{\varepsilon}_0 = \underline{\mathbf{0}}$

Legyen $\quad \|\,\underline{\varepsilon}_k\,\| < \varepsilon \quad$ és vezessük be a

$\underline{\mathbf{z}}_k := \underline{\mathbf{y}}_k - \underline{\mathbf{x}}_k \qquad k = 1, 2, \ldots \qquad$ jelölést. Itt $\underline{\mathbf{z}}_k$ a hibavektor.

Ekkor $\quad \underline{\mathbf{y}}_k = \underline{\mathbf{z}}_k + \underline{\mathbf{x}}_k$ - t *(2)* - be írva *(1)* $\quad \Longrightarrow$

$\Longrightarrow \qquad \underline{\mathbf{z}}_k = C\underline{\mathbf{z}}_{k-1} + \underline{\varepsilon}_k \qquad$ és innen $\qquad \Longrightarrow$

$$\|\,\underline{\mathbf{z}}_k\,\| \le \|\,C\,\| \cdot \|\,\underline{\mathbf{z}}_{k-1}\,\| + \|\,\underline{\varepsilon}_k\,\| \le \|\,C\,\| \left(\|\,C\,\| \cdot \|\,\underline{\mathbf{z}}_{k-2}\,\| + \|\,\underline{\varepsilon}_{k-1}\,\|\right) + \|\,\underline{\varepsilon}_k\,\| \le \|\,C\,\|^2 \cdot \|\,\underline{\mathbf{z}}_{k-2}\,\| +$$

$$+ \|\,C\,\| \cdot \varepsilon + \varepsilon \le \ldots \le \|\,C\,\|^k \cdot \|\,\underline{\mathbf{z}}_0\,\| + \varepsilon \left(\|\,C\,\|^{k-1} + \|\,C\,\|^{k-2} + \ldots + \|\,C\,\| + 1\right)$$

$$\Longrightarrow \qquad \|\,\underline{\mathbf{z}}_k\,\| \le \|\,C\,\|^k \cdot \|\,\underline{\mathbf{z}}_0\,\| + \varepsilon \frac{1}{1 - \|\,C\,\|}$$

$$\Longrightarrow \qquad \boxed{\|\,\underline{\mathbf{z}}_k\,\| \le \frac{1}{1 - \|\,C\,\|} \cdot \varepsilon} \, , \qquad \text{ha } \|\,C\,\| < 1$$

Az ε értéke n^2 szorzás és n^2 összeadás utáni kerekítési hibát jelenti. Így ennek értéke nagy is lehet. Ezért, ha még $\| C \|$ szám is 1 - hez közeli, a kerekítési hibák a megoldást nagymértékben torzíthatják. A gyakorlati alkalmazásoknál ezért arra kell törekedni, hogy az $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ lineáris egyenletrendszer $\underline{\mathbf{x}} = C\underline{\mathbf{x}} + \underline{\mathbf{c}}$ ekvivalens alakra való átalakítását úgy csináljuk, hogy $\| C \|$ minél kisebb legyen.

6. NEMLINEÁRIS EGYENLETEK ITERÁCIÓS MEGOLDÁSI MÓDSZEREI

Az általános iteráció elvét most alkalmazzuk nemlineáris egyenletek és egyenletrenszerek megoldására.

FELADAT :

Keressük az

(1) $\quad F(x) = 0 \quad$ egyenlet egy x megoldását (gyökét), ahol x^* az ***(1)*** megoldása, ha $F(x^*) = 0$.

Itt az F operátor egy $F : D \to X \quad D \subset X$ leképezés, $(X, \|.\|)$ Banach - tér (lineáris, normált, teljes).

Tehát ***(1)*** megoldását D - ben (X - ben) keressük. Véges sok lépésben meghatározni ***(1)*** gyökét azonban csak nagyon ritka esetben lehet, ezért van szükség iterációs módszerekre.

Ehhez írjuk át ***(1)*** - et vele ekvivalens alakra:

(2) $\quad x = \Phi(x) \quad$ „fixpontegyenlet”

Ha x^* a ***(2)*** megoldása, úgynevezett fixpontja, $x^* = \Phi(x^*)$, akkor x^* az ***(1)*** - nek is megoldása.

A ***(2)*** megoldásának meghatározására képezzük a következő iterációs sorozatot (közelítő megoldások sorozata):

(3) $\quad x_{k+1} = \Phi(x_k)$, ahol Φ „iterációs függvény”

KÉRDÉSEK :

1. Hogyan választható alkalmas Φ iterációs függvény?

2. Milyen feltételek esetén konvergál az $\{x_i\}$ közelítő megoldások sorozata?

3. Milyen gyors a konvergencia?

4. Az x_0 kezdőértéket hogyan kell megválasztani?

6.1. KONVERGENCIA

Definíció :

Azt mondjuk, hogy az $\{x_k\}$ iterációs sorozat konvergens, ha:

$$\lim_{k\to\infty} x_k = x^* , \qquad \text{vagy} \qquad \lim_{k\to\infty} \| x_k - x^* \| = 0 ,$$

ahol x^* a ***(2)*** megoldása, fixpontja, ami egyben megoldása ***(1)*** - nek is.

Tétel **Bolzano - tétel**

$$\text{Ha} \quad \left.\begin{array}{l} f \in C\,[a\,,\,b] \\ f\,(a) < 0 < f\,(b) \end{array}\right\} \quad \Longrightarrow \quad \exists \quad x^* \in (a\,,\,b)\ , \quad \text{hogy} \quad f\,(x^*) = 0$$

Bizonyítás

Ezt intervallumfelezési eljárással bizonyítjuk (Leindler - Schipp: Analízis I. jegyzet 151. oldal).

Legyen $x_0 = a\,,\ \ y_0 = b$.

Tegyük fel, hogy $x_k\,,\ \ y_k$ - t $k = 1, \ldots, n$ - re már meghatároztuk.

Ekkor legyen $z_n := \dfrac{y_n - x_n}{2}$ és

$$x_{n+1} = x_n\,, \quad y_{n+1} = z_n\,, \quad \text{ha} \quad f\,(z_n) \geq 0,$$
$$x_{n+1} = z_n\,, \quad y_{n+1} = y_n\,, \quad \text{ha} \quad f\,(z_n) < 0.$$

Így egymásba skatulyázott intervallumok sorozatát kapjuk, amelyre:

$$\lim x_n = \lim y_n = x^*\,, \qquad y_n - x_n = \frac{b-a}{2^n} \xrightarrow[n\to\infty]{} 0$$

■

Definíció :

Legyen $(X\,,\ \|\,.\,\|\,)$ Banach - tér, $\Phi\ :\ X \to X$ X - nek önmagába való leképezése.

A Φ leképezés kontrakció (összehúzódás), ha $\exists\ \ 0 < \alpha < 1$ szám, hogy $\forall\ \ x\,,\,y \in X$ - re

$\|\,\Phi\,(x) - \Phi\,(y)\,\| \leq \alpha\,\|\,x - y\,\|$.

Elégséges feltétel a *(3)* iteráció konvergenciájára :

$\boxed{\textbf{Tétel}}$ Banach - féle fixponttétel

Legyen $(X\,,\ \|\,.\,\|)$ Banach - tér, $\Phi\ :\ X \to X$ leképezés kontrakció X - en.

Ekkor az $x = \Phi\,(x)$ egyenletnek $\exists\,!$ $x^* \in X$ megoldása, ,,fixpontja" $(x^* = \Phi\,(x^*))$, továbbá az $x_{k+1} = \Phi\,(x_k)$ iterációs sorozat tetszőleges x_0 esetén konvergál x^* - hoz, azaz $\lim\limits_{k\to\infty} x_k = x^*$.

SŐT: A tétel bizonyításából a közelítő megoldásra hibabecslést is kapunk!

Bizonyítás

1.)

$\{x_k\}$ Cauchy - sorozat, ugyanis:

$$\|\,x_{k+1} - x_k\,\| \ = \ \|\,\Phi\,(x_k) - \Phi\,(x_{k-1})\,\| \ \le \alpha\,\|\,x_k - x_{k-1}\,\| \ \le \alpha^k\,\|\,x_1 - x_0\,\| \ ,$$

ahonnan $s > k$ esetén:

$$(\star)\qquad \begin{cases} \|\,x_s - x_k\,\| \ \le \ \|\,x_s - x_{s-1} + x_{s-1} - \ldots - x_{k+1} + x_{k+1} - x_k\,\| \ \le \\ \le \|\,x_s - x_{s-1}\,\| + \|\,x_{s-1} - x_{s-2}\,\| + \ldots + \|\,x_{k+1} - x_k\,\| \ \le \\ \le \alpha^{s-1}\,\|\,x_1 - x_0\,\| + \alpha^{s-2}\,\|\,x_1 - x_0\,\| + \ldots + \alpha^k\,\|\,x_1 - x_0\,\| \ = \\ = \left(\alpha^{s-1} + \alpha^{s-2} + \ldots + \alpha^k\right)\,\|\,x_1 - x_0\,\| = \alpha^k\left(\alpha^{s-k-1} + \ldots + \alpha + 1\right)\,\|\,x_1 - x_0\,\| \ \le \\ \le \alpha^k\,\dfrac{1}{1-\alpha}\,\|\,x_1 - x_0\,\| \end{cases}$$

$$(\star)\qquad \Longrightarrow \qquad \|\,x_s - x_k\,\| \le \frac{\alpha^k}{1-\alpha}\,\|\,x_1 - x_0\,\|$$

Tehát $\{x_k\}$ tényleg Cauchy - sorozat $\Longrightarrow$ $\exists$ x^* hogy $\lim\limits_{k\to\infty} x_k = x^*\,,$ $x^* \in X$

2.)

x^* a *(2)* megoldása (fixpontja), ugyanis:

$$\| x^* - \Phi(x^*) \| = \| x^* - x_k + x_k - \Phi(x^*) \| \leq \| x^* - x_k \| + \| \Phi(x_{k-1}) - \Phi(x^*) \| \leq$$

$$\leq \| x^* - x_k \| + \alpha \| x_{k-1} - x^* \| \xrightarrow[k \to \infty]{} 0$$

$$\Longrightarrow \quad \| x^* - \Phi(x^*) \| = 0 \quad \Longrightarrow \quad x^* - \Phi(x^*) = 0 \quad \Longrightarrow \quad x^* = \Phi(x^*)$$

3.)

x^* egyértelmű megoldás (fixpont), ugyanis:

ha $\eta = \Phi(\eta)$ is teljesül, akkor

$$\| \eta - x^* \| = \| \Phi(\eta) - \Phi(x^*) \| \leq \alpha \| \eta - x^* \|$$

$(1-\alpha) \| \eta - x^* \| \leq 0$ és mivel $\alpha < 1 \quad \Longrightarrow \quad x^* = \eta$

■

Megjegyzés :

Ha $x_{k+1} = x_k$, akkor megállítjuk az iterációt és $x_k = x_{k+1} = \Phi(x_k)$ a keresett x^* megoldás (fixpont).

Általánosítás :

Sok gyakorlati esetben a Φ leképezés a $D \subset X$ zárt halmazon definiált.

Ha $\Phi : D \to D$ és Φ kontrakció D - n, a tétel továbbra is érvényben marad.

Ha $x_0 \in D$ (így választjuk), akkor $x_1 = \Phi(x_0) \in D$, és általában $x_k \in D \quad k \geq 2$, tehát az iteráltak D - ben maradnak $\Longrightarrow \quad \exists! \; x^* \in D, \quad \lim x_k = x^*$.

6.2. HIBABECSLÉS

A - PRIORI HIBABECSLÉS

$$(\star) \qquad \| x_s - x_k \| \leq \frac{\alpha^k}{1-\alpha} \| x_1 - x_0 \|$$

Tudjuk, hogy $\lim\limits_{s \to \infty} x_s = x^*$.

(⋆) - ban limeszre térve $\Longrightarrow$ $\boxed{\| x^* - x_k \| \leq \frac{\alpha^k}{1-\alpha} \| x_1 - x_0 \|}$

Speciális esetben ez kiadja a lineáris egyenletrendszerekre vonatkozó a - priori hibabecslést:

$\underline{\mathbf{x}} = C\underline{\mathbf{x}} + \underline{\mathbf{c}}\,,\quad \Phi = C\underline{\mathbf{x}} + \underline{\mathbf{c}}\,,\quad \|\,C\,\| < 1 \quad \Longrightarrow \quad \Phi$ lineáris leképezés kontrakció, és

$\alpha = \|\,C\,\| < 1$

A - POSTERIORI HIBABECSLÉS

Vizsgáljuk:

$$\|\,x^* - x_k\,\| = \|\,x^* - x_{k+1} + x_{k+1} - x_k\,\| \leq \|\,x^* - x_{k+1}\,\| + \|\,x_{k+1} - x_k\,\| \leq$$

$$\leq \|\,\Phi(x^*) - \Phi(x_k)\,\| + \|\,\Phi(x_k) - \Phi(x_{k-1})\,\| \leq \alpha\,\|\,x^* - x_k\,\| + \alpha\,\|\,x_k - x_{k-1}\,\|$$

Rendezve $\Longrightarrow$ $(1-\alpha)\,\|\,x^* - x_k\,\| \leq \alpha\,\|\,x_k - x_{k-1}\,\|$,

ahol $\alpha < 1 \quad \Longrightarrow \quad \boxed{\|\,x^* - x_k\,\| \leq \dfrac{\alpha}{1-\alpha}\,\|\,x_k - x_{k-1}\,\|}$

6.2.1. Az egyszerű iteráció konvergenciasebessége

VOLT :

(1) $F(x) = 0$

(2) $x = \Phi(x)$

(3) $x_{k+1} = \Phi(x_k)$

$$\text{Legyen}\quad \left.\begin{array}{l} \Phi\ :\ [a\,,\,b] \to [a\,,\,b] \\[2ex] \Phi \in C^1\,[a\,,\,b] \\[2ex] k := \max\limits_{x\in[a\,,\,b]} \left|\Phi'(x)\right| < 1 \end{array}\right\} \quad \text{ekkor } \Phi \text{ kontrakció } [a\,,\,b] \text{ - n.}$$

Vizsgáljuk a hibát:

$$x_{k+1} - x^* = \Phi(x_k) - \Phi(x^*) \overset{\text{Lagrange - féle középértéktétel}}{=} \Phi'(\eta)\,(x_k - x^*)\quad ,$$

ahol $\eta = x_k + \Theta\,\varepsilon_k\,,\quad 0 < \Theta < 1\,,\quad \varepsilon_k := x_k - x^*$.

Bevezetve:

(4) $\varepsilon_k := x_k - x^* \qquad \Longrightarrow \qquad \varepsilon_{k+1} = \Phi'(\eta)\,\varepsilon_k$

Ha $\Phi'(x) < 0 \quad [a\,,\,b]$ - n, akkor az $\{x_k\}$ sorozat alternáló,

ha $\Phi'(x) > 0 \quad [a\,,\,b]$ - n, akkor az $\{x_k\}$ sorozat monoton.

(4) $\Longrightarrow \qquad \varepsilon_k \to 0 \qquad k \to \infty$ esetén és

$\eta \to x^* \quad (k \to \infty)$ így $\qquad \Longrightarrow \qquad \boxed{\lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \Phi'(x^*)}$

Definíció :

Ha $\Phi'(x^*) \neq 0$, akkor azt mondjuk, hogy az $\{x_k\}$ sorozat lineárisan konvergál az x^* - hoz,

$|\varepsilon_{k+1}| \leq K \cdot |\varepsilon_k| \qquad$ vagy $\qquad |\varepsilon_{k+1}| = O(|\varepsilon_k|)$.

Ha $\Phi'(x^*) = 0$, akkor azt mondjuk, hogy az $\{x_k\}$ sorozat szuperlineárisan konvergál az x^* - hoz.

Ekkor a Lagrange - féle középértéktétel nem mond semmit, de a Taylor - formula alapján:

$$x_{k+1} - x^* = \Phi(x_k) - \Phi(x^*)$$

$$\Phi(x_k) = \Phi(x^*) + \frac{\Phi'(x^*)}{1!}(x_k - x^*) + \frac{\Phi''(\eta)}{2!}(x_k - x^*)^2$$

$$\Longrightarrow \qquad x_{k+1} - x^* = \frac{\Phi''(\eta)}{2!}(x_k - x^*)^2$$

$$\Longrightarrow \qquad \varepsilon_{k+1} = \widetilde{K} \cdot \varepsilon_k^2 \qquad |\varepsilon_{k+1}| = O\left(|\varepsilon|^2\right) \qquad \text{vagy} \qquad \lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k^2} = \frac{1}{2}\Phi''(x^*)\ ,$$

ami azt jelenti, hogy a konvergencia legalább másodrendű, de ehhez kell, hogy $\Phi \in C^2[a\,,\,b]$.

Definíció :

Az $\{x_k\}$ sorozat konvergenciája p - edrendű, ha $\displaystyle\lim_{k\to\infty} \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} = c$ (állandó)

vagy $\qquad \|\,\varepsilon_{k+1}\,\| = O(\,\|\,\varepsilon_k\,\|^{\,p})$.

Kérdés : Alkalmas Φ függvény választásával a szuperlineáris konvergencia elérhető - e?

6.3. NEWTON - MÓDSZER (ITERÁCIÓ)

Legyen $f \in C^1[a\,,\,b]$ és keressük

(1) $f(x) = 0$ egyenlet megoldását $[a\,,\,b]$ - n.

Tegyük fel, hogy $\exists\ x^* \in [a\,,\,b]$, melyre $f(x^*) = 0$.

(1) - et átírjuk vele ekvivalens alakra, amelyre a $\Phi'(x^*) = 0$ lesz.

Legyen $g \in C^1[a\,,\,b]$, $g \neq 0$ $[a\,,\,b]$ - n.

$$f(x) = 0 \qquad / \cdot g(x)$$

(2) $g(x)\,f(x) = 0 \qquad / + x$

(3) $x = x + f(x)\,g(x) \qquad (\Phi = x + f \cdot g)$

Hogy szuperlineáris legyen a konvergencia, kell, hogy $\Phi'(x^*) = 0$

$$\Phi(x) = x + f(x)\,g(x)$$

$$\Longrightarrow \quad \Phi'(x) = 1 + f'(x)\,g(x) + f(x)\,g'(x)$$

$$\Longrightarrow \quad \Phi'(x^*) = 1 + f'(x^*)\,g(x^*) + \overbrace{f(x^*)\,g'(x^*)}^{0}$$

$$\Longrightarrow \quad 1 + f'(x^*)\,g(x^*) = 0$$

$$\Longrightarrow \quad g(x^*) = -\frac{1}{f'(x)}\ , \qquad f'(x^*) \neq 0\,, \quad x \in [a\,,\,b]$$

Ez elérhető akkor, ha a g függvényt az x^* hely környezetében

$g(x) = -\dfrac{1}{f'(x)}$ - nek választjuk.

Így a szuperlineáris konvergenciájú módszerünk a következő:

$$\boxed{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}} \qquad k = 0, 1, 2, \ldots$$

Tétel **Konvergencia - tétel**

Legyen $f \in C^2[a\,,\,b]$ $f(x^*) = 0\,,$ $x^* \in [a\,,\,b]\,,$ $f'(x^*) \neq 0$ (x^* egyszeres gyök)

$0 < m := \inf \left| f'(x^*) \right|$ és $M := \sup \left| f''(x) \right|$ az x^* valamely környezetében.

Ekkor $\exists$ $U(x^*)$ környezet, amelyre tetszőleges $x_0 \in U(x^*)$ kezdőérték esetén a Newton - módszer konvergens és ez a konvergencia (lokálisan) másodrendű.

Bizonyítás

Legyen $\Phi(x) := x - \dfrac{f(x)}{f'(x)}$.

a.)

Belátjuk, hogy Φ kontrakció egy alkalmas környezetében x^* - nak.

$|\Phi(x) - \Phi(y)| \leq \alpha\,|x - y|\,,$ $\alpha < 1$

$\Phi(x) - \Phi(y) = \Phi'(i)\,(x - y)$

Vizsgáljuk:

$$\Phi'(x) = 1 - \frac{-f(x)\,f''(x) + \left[f'(x)\right]^2}{\left[f'(x)\right]^2} = \frac{f(x)\,f''(x)}{\left[f'(x)\right]^2}$$

Mivel $f(x^*) = 0\,,$ f folytonos és $\exists$ $m\,,M$ állandók, azért $\exists$ $U(x^*) := [x^* - \delta\,,\,x^* + \delta]$

környezete x^* - nak, amelyre $\left|\Phi'\right| < 1$

$\Longrightarrow$ Φ az $U(x^*)$ környezetében kontrakció.

b.)

Belátjuk, hogy $\Phi\,:\,U(x^*) \to U(x^*)$.

Legyen $x \in U(x^*)\,,$ azaz $|x - x^*| < \delta$.

Tekintsük: $|\Phi(x) - x^*| = |\Phi(x) - \Phi(x^*)| \leq \alpha\,|x - x^*| < |x - x^*| < \delta\,,$ $\alpha < 1$

$\Longrightarrow$ $\forall$ $x \in U(x^*)$ - ra a $\Phi(x) \in U(x^*)$ - nak.

A Banach - féle fixponttétel miatt a Newton - módszer $U(x^*)$ - on konvergens tetszőleges $x_0 \in U(x^*)$ - ra.

c.)

Belátjuk, hogy a Newton - módszer konvergenciája $U(x^*)$ - on másodrendű $\left(|\varepsilon_{k+1}| = O\left(|\varepsilon_k|^2\right)\right)$.

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \qquad / \cdot (-1) \, , \qquad / + x^*$$

$$(\star) \qquad \underbrace{x^* - x_{k+1}}_{\varepsilon_{k+1}} = x^* - x_k + \frac{f(x_k)}{f'(x_k)}$$

Írjuk fel f - re a Taylor - formulát az x_k pontban.

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{f''(\xi)}{2!}(x - x_k)^2$$

$x = x^*$ - ra: $\underbrace{f(x^*)}_{0} = f(x_k) + f'(x_k)(x^* - x_k) + \frac{f''(\xi)}{2!}(x^* - x_k)^2$

$$0 = f(x_k) + f'(x_k)(x^* - x_k) + \frac{f''(\xi)}{2!}(x^* - x_k)^2 \qquad / : f'(x_k)$$

$$0 = \frac{f(x_k)}{f'(x_k)} + (x^* - x_k) + \frac{f''(\xi)}{2f'(x_k)}(x^* - x_k)^2$$

Ezt $(\star)$ - ba írva $\implies$ $x^* - x_{k+1} = -\frac{f''(\xi)}{2f'(x_k)} \underbrace{(x^* - x_k)^2}_{\varepsilon_k^2}$

$\implies$ $|\varepsilon_{k+1}| = O\left(|\varepsilon_k|^2\right)$, azaz a konvergencia valóban másodrendű $U(x^*)$ - on.

■

Megjegyzés :

A Newton - módszer másik neve érintőmódszer.

A többszörös gyökök esete

Sokszor nem ismerjük az x^* gyök multiplicitását. Ezekre hasznosak az olyan módszerek, amelyek nem függnek a gyök multiplicitásától.

Ötlet : ha $f(x) = 0$ - nak x^* valahányszoros gyöke, akkor

(1) $u(x) := \frac{f(x)}{f'(x)}$ - nek az x^* egyszeres gyöke lesz.

$$f(x) = (x - x^*)^r \cdot g(x)$$

$$f'(x) = r(x - x^*)^{r-1} \cdot g(x) + (x - x^*)^r \cdot g'(x)$$

$$u'(x) = \frac{\left[f'(x)\right]^2 - f(x) f''(x)}{[f'(x)]^2}$$

Most $u(x) = 0$ - ra alkalmazzuk a Newton - módszert:

$$x_{k+1} = x_k - \frac{u(x_k)}{u'(x_k)} = x_k - \frac{f(x_k)\left[f'(x_k)\right]^2}{f'(x_k)\left\{[f'(x_k)]^2 - f(x_k) f''(x_k)\right\}} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad \boxed{x_{k+1} = x_k - \frac{f(x_k) f'(x_k)}{[f'(x_k)]^2 - f(x_k) f''(x_k)}} \qquad k = 0, 1, 2, \ldots$$

Belátható, hogy ha a jobb oldal kellően sima, akkor e módszer konvergenciája másodrendű marad, de ez a módszer kisebb hatékonyságú általában, mint az „eredeti” , mivel a második deriváltat is számolni kell minden egyes lépésben.

Ha a gyök multiplicitása ismert, mondjuk r, akkor igaz a következő:

Tétel

A tett feltevések mellett az

(2) $$\boxed{x_{k+1} = x_k - r\frac{f(x_k)}{f'(x_k)}}$$

formula konvergenciája is másodrendű, ha a gyök multiplicitása r.

Bizonyítás

(2) $\Longrightarrow \quad x^* - x_{k+1} = x^* - x_k + r\dfrac{f(x_k)}{f'(x_k)} \qquad / \cdot f'(x_k)$

(3) $\quad (x^* - x_{k+1}) \cdot f'(x_k) = (x^* - x_k) \cdot f'(x_k) + r \cdot f(x_k)$

Legyen $\quad g(x) := (x^* - x) \cdot f'(x) + r \cdot f(x)$

Tudjuk, hogy x^* az f - nek r - szeres gyöke $\quad \Longrightarrow \quad g(x^*) = 0$

Vizsgáljuk:

$$g'(x) = (x^* - x) \cdot f''(x) - f'(x) + r \cdot f'(x)$$

$$g''(x) = (x^* - x) \cdot f'''(x) - 2 \cdot f''(x) + r \cdot f''(x)$$

$$\vdots$$

$$g^{(j)}(x) = (x^* - x) \cdot f^{(j+1)}(x) - j \cdot f^{(j)}(x) + r \cdot f^{(j)}(x)$$

$$\Longrightarrow \qquad g(x^*) = g'(x^*) = \ldots = g^{(r)} = 0$$

Írjuk fel g - re a Taylor - formulát x^* - ban.

$$g(x) = g(x^*) + g'(x^*)(x - x^*) + \frac{g''(x^*)}{2!}(x - x^*)^2 + \ldots + \frac{g^{(r)}(x^*)}{r!}(x - x^*)^r + \frac{g^{(r+1)}(\xi)}{(r+1)!}(x - x^*)^{r+1}$$

$$\Longrightarrow \qquad \boxed{g(x) = \frac{g^{(r+1)}(\xi)}{(r+1)!}(x - x^*)^{r+1}} \qquad \textit{(4)}$$

Továbbá f Taylor - sora:

$$f(x) = \frac{f^{(r)}(\eta)}{r!}(x - x^*)^r \qquad \Longrightarrow \qquad \boxed{f'(x) = \frac{f^{(r)}(\eta)}{(r-1)!}(x - x^*)^{r-1}} \qquad \textit{(5)}$$

$$\textit{(3)} \qquad \Longrightarrow \qquad (x^* - x_{k+1}) \cdot f'(x_k) = g(x_k) \qquad \Longrightarrow \qquad (x^* - x_{k+1}) = \frac{g(x_k)}{f'(x_k)}$$

Ide beírva *(4)* - et, *(5)* - öt:

$$(x^* - x_{k+1}) = \frac{\dfrac{g^{(r+1)}(\xi)}{(r+1)!}(x_k - x^*)^{r+1}}{\dfrac{f^{(r)}(\eta)}{(r-1)!}(x_k - x^*)^{r-1}}$$

$$\Longrightarrow \qquad x^* - x_{k+1} = \frac{g^{(r+1)}(\xi)}{r(r+1) \cdot f^{(r)}(\eta)}(x_k - x^*)^2$$

$x^* - x_{k+1} = C \cdot (x_k - x^*)^2 \qquad \Longrightarrow \qquad |\varepsilon_{k+1}| \leq C \cdot |\varepsilon_k|^2$ a deriváltak korlátosságát feltéve.

■

6.4. SZELŐ MÓDSZER

A Newton - módszer egyik hátránya lehet, hogy deriváltat kell számolni. Ezt küszöböli ki a szelő módszer. A deriváltat különbségi hányadossal helyettesítjük.

Legyen $f'(x_k) \approx \dfrac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$.

$$\Longrightarrow \qquad x_{k+1} = x_k - \frac{f(x_k)}{\dfrac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}} \qquad \Longrightarrow \qquad x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}$$

$$\Longrightarrow \qquad \boxed{x_{k+1} = \frac{x_{k-1} f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}} \qquad k = 1, 2, \ldots$$

Megjegyzés :

i.)

A szelő módszer nem írható $x_{k+1} = \Phi(x_k)$ alakban, mint a Newton - módszer, most ugyanis két kezdőértékre van szükség : x_0 , x_1

ii.)

A szelő módszer konvergenciasebessége kisebb, mint a Newton - módszeré.

A konvergenciasebesség: $\dfrac{1+\sqrt{5}}{2} \approx 1.618$

6.5. MÓDOSÍTOTT NEWTON - MÓDSZER

Most az érintő helyett mindig $f'(x_0)$ meredekségű egyenes gyökét határozzuk meg.

$$\Longrightarrow \qquad \boxed{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}}$$

6.6. KONVERGENCIAGYORSÍTÁS : AITKEN TRANSZFORMÁCIÓ (AITKEN Δ^2 MÓDSZERE)

Tekintsünk egy konvergens $\{x_k\}$ sorozatot $\lim\limits_{k\to\infty} x_k = x^*$, amelyik lineárisan konvergál x^* - hoz:

(1) $\quad x_{k+1} - x^* \approx \alpha\,(x_k - x^*)$, ahol $0 < \alpha < 1$

Az ***(1)*** segítségével α és x^* kifejezhető a következő módon:

(1) $\Longrightarrow \quad x_{k+2} - x^* = \alpha\,(x_{k+1} - x^*)$ ***(2)***

$$\Longrightarrow \quad \alpha = \frac{x_{k+2} - x_{k+1}}{x_{k+1} - x_k}$$

Ezt ***(1)*** - be írva $\Longrightarrow$

(3)
$$\boxed{x^* = \frac{x_k\,x_{k+2} - x_{k+1}^2}{x_{k+2} - 2x_{k+1} + x_k}}$$

Várható, hogy az x_k , x_{k+1} , x_{k+2} segítségével x^* jobb közelítése kapható meg, mint az eredeti $\{x_k\}$ sorozat.

$$x_k^{'} = \frac{x_k\,x_{k+2} - x_{k+1}^2}{x_{k+2} - 2x_{k+1} + x_k} \qquad \textit{(3)}$$

De ***(3)*** numerikusan instabil formula, ezért átírjuk:

(3) $\Longrightarrow \quad (x_{k+2} - 2x_{k+1} + x_k)\; x^* = x_k\,x_{k+2} - x_{k+1}^2 = (x_{k+2} - 2x_{k+1} + x_k)\; x_k - (x_{k+1} - x_k)^2$

$$\Longrightarrow \quad \boxed{x^* = x_k - \frac{(x_{k+1} - x_k)^2}{x_{k+2} - 2x_{k+1} + x_k}}$$

Ez már numerikusan stabil formula.

Bevezetve a következő differenciaoperátorokat:

$$\Delta x_k := x_{k+1} - x_k$$

$$\Delta^2 x_k := \Delta x_{k+1} - \Delta x_k = x_{k+2} - 2x_{k+1} + x_k$$

Így ***(3)*** alakja: $\quad x^* = x_k - \dfrac{(\Delta x_k)^2}{\Delta^2 x_k} \qquad$ ***(4)***

A jobb oldal számolásával $\left\{x_k^{'}\right\}$ transzformált sorozatot kapjuk, ami gyorsabban konvergál x^* - hoz, mint az eredeti sorozat.

(5) $$\boxed{x_k^{'} := x_k - \frac{(\Delta x_k)^2}{\Delta^2 x_k}} \qquad k = 1, 2, \ldots$$

Tétel

Legyen $0 < \alpha < 1$ olyan szám, amellyel az $\{x_k\}$ sorozat a következő módon írható:

(6) $\quad x_{k+1} = (\alpha + \delta_k)\, x_k$, ahol $\lim\limits_{k\to\infty} \delta_k = 0$.

Ekkor az ***(5)*** transzformált sorozat elég nagy k - ra jól definiált és $\left\{x_k^{'}\right\}$ gyorsabban konvergál x^* - hoz, mint az $\{x_k\}$ sorozat abban az értelemben, hogy:

$$\lim_{k\to\infty} \frac{x_k^{'} - x^*}{x_k - x^*} = 0$$

Bizonyítás

Legyen $\quad \varepsilon_k := x_k - x^*$.

Ekkor ***(6)*** $\quad \Longrightarrow \quad \varepsilon_{k+1} = (\alpha + \delta_k)\, \varepsilon_k \qquad$ ***(7)***

Vizsgáljuk ***(5)*** - öt:

$$x_{k+1} - x_k = (x_{k+1} - x^*) - (x_k - x^*) = \varepsilon_{k+1} - \varepsilon_k \overset{(7)}{=} \varepsilon_k\, [(\alpha - 1) + \delta_k]$$

Továbbá:

$$x_{k+2} - 2x_{k+1} + x_k = (x_{k+2} - x^*) - (2x_{k+1} - 2x^*) + (x_k - x^*) = \varepsilon_{k+2} - 2\varepsilon_{k+1} + \varepsilon_k \overset{(7)}{=}$$

$$= \varepsilon_{k+1} (\alpha + \delta_{k+1}) - 2 (\alpha + \delta_k)\, \varepsilon_k + \varepsilon_k = \varepsilon_k (\alpha + \delta_k)(\alpha + \delta_{k+1}) - 2\varepsilon_k (\alpha + \delta_k) + \varepsilon_k = \varepsilon_k \left[(\alpha - 1)^2 + \mu_k\right],$$

ahol μ_k a többi tagot tartalmazza, amelyekben δ_k , δ_{k+1} szerepel. $\quad \mu_k \to 0 \quad (k \to \infty)$

Ezeket figyelembe véve ***(5)*** $\quad \Longrightarrow \quad x_k^{'} - x^* = (x_k - x^*) - \dfrac{\varepsilon_k^2\, [(\alpha - 1) + \delta_k]^2}{\varepsilon_k \left[(\alpha - 1)^2 + \mu_k\right]}$

$$\Longrightarrow \quad x_k^{'} - x^* = \varepsilon_k \left[1 - \frac{[(\alpha - 1) + \delta_k]^2}{(\alpha - 1)^2 + \mu_k}\right]$$

$$\Longrightarrow \qquad \frac{x_k' - x^*}{x_k - x^*} = 1 - \frac{[(\alpha-1)+\delta_k]^2}{(\alpha-1)^2+\mu_k} \xrightarrow[k\to\infty]{} 0$$

■

6.7. AITKEN - STEFFENSEN TRANSZFORMÁCIÓ

Tekintsük az $x_{k+1} = \Phi(x_k)$ iterációs sorozatot, amely az $f(x) = 0$ egyenlet x^* megoldásához tart. Képezzük a következő transzformált sorozatot:

x_0	$x_1 = x_0 - \dfrac{(y_0-x_0)^2}{z_0-2y_0+x_0}$	$x_2 = x_1 - \dfrac{(y_1-x_1)^2}{z_1-2y_1+x_1}$	$\cdots$
$y_0 = \Phi(x_0)$	$y_1 = \Phi(x_1)$	$y_2 = \Phi(x_2)$	$\cdots$
$z_0 = \Phi(y_0)$	$z_1 = \Phi(y_1)$	$z_2 = \Phi(y_2)$	$\cdots$

(8)
$$\boxed{x_k = x_{k-1} - \frac{(y_{k-1}-x_{k-1})^2}{z_{k-1}-2y_{k-1}+x_{k-1}}}$$

Állítás

Ha az $x_{k+1} = \Phi(x_k)$ iterációs sorozat lineárisan konvergál x^* - hoz, akkor a *(8)* sorozat legalább másodrendben konvergál x^* - hoz $\left(\Phi'(x^*) \neq 0\right)$.

6.8. BECSLÉS POLINOM GYÖKEINEK ELHELYEZKEDÉSÉRE

Legyen

$$f(x) = 0 \quad \text{helyett most} \quad P_n(x) = 0 \qquad \Longrightarrow \qquad x_{n+1} = x_n - \frac{P_n(x_n)}{P_n'(x_n)}$$

a Newton - módszer polinomokra.

Tétel

Tekintsük a $P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ polinomot,

és legyen $a_n \neq 0\,,\quad a_0 \neq 0$ és

$A := \max\{|a_{n-1}|\,,\,|a_{n-2}|\,,\,\ldots\,,\,|a_0|\}\,,$

$B := \max\{|a_n|\,,\,|a_{n-1}|\,,\,\ldots\,,\,|a_1|\}.$

Ekkor a $P_n(x)$ polinom gyökeinek elhelyezkedésére igaz a következő becslés:

$$0 < \frac{1}{1 + \dfrac{B}{|a_0|}} < |x_k| < 1 + \frac{A}{|a_n|}$$

Bizonyítás

Vizsgáljuk $P_n(x)$ - et $|x| > 1$ esetén. $P_n(x)$ - et rendezve:

$P_n(x) - a_{n-1}x^{n-1} - \ldots - a_1 x - a_0 = a_n x^n$

$$\Longrightarrow \qquad |P_n(x)| + \left|a_{n-1}x^{n-1}\right| + \ldots + |a_1 x| + |a_0| \geq |a_n| \cdot |x|^n \geq |a_n|\,|x|^n - A\left(|x|^{n-1} + \ldots + |x| + 1\right) =$$

$$= |a_n| \cdot |x|^n - A\frac{|x|^{n-1} - 1}{|x| - 1} > |a_n| \cdot |x|^n - A\frac{|x|^n}{|x| - 1} = |x|^n\left(|a_n| - A\frac{1}{|x| - 1}\right)$$

Legyen x_k a polinom gyöke $\Longrightarrow$ $|P_n(x_k)| = 0$ $\Longrightarrow$ $|a_n| - \dfrac{A}{|x_k| - 1} < 0$ $\Longrightarrow$

$$\Longrightarrow \qquad |a_n| < \frac{A}{|x_k| - 1}\,, \qquad |a_n| \cdot |x_k| - |a_n| < A\,, \qquad |x_k| < 1 + \frac{A}{|a_n|}$$

Tehát a $P_n(x) = 0$ egyenlet gyökei a komplex számsík $R = 1 + \dfrac{A}{|a_n|}$ sugarú nyílt körlemezén helyezkednek el, speciálisan a valós gyökök a $(-R\,,\,R)$ intervallumban.

Alsó becslés :

Mivel $a_0 \neq 0$ $\Rightarrow$ a 0 nem gyöke a $P_n(x)$ polinomnak. Ezért, ha x_k a $P_n(x)$ polinom gyöke, akkor $\dfrac{1}{x_k}$ a $P_n(x)$ úgynevezett reciprok polinomjának a gyöke lesz, ahol a reciprok polinom:

$$\widetilde{P}_n(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

Tekintsük:

$$P_n(x_k) = a_n x_k^n + a_{n-1}x_k^{n-1} + \ldots + a_1 x_k + a_0 = x_k^n \left(a_n + a_{n-1}\frac{1}{x_k} + \ldots + a_1 \frac{1}{x_k^{n-1}} + a_0 \frac{1}{x_k^n} \right) = 0$$

$$\Longrightarrow \qquad \widetilde{P}_n\left(\frac{1}{x_k}\right) = 0 \qquad \text{tehát} \qquad \frac{1}{x_k} \ \text{ a } \ \widetilde{P}_n \ \text{ reciprok polinom gyöke.}$$

Most alkalmazzuk a felső becslést a reciprok polinom gyökeire.

$$\left|\frac{1}{x_k}\right| < 1 + \frac{B}{|a_0|} \qquad \Longrightarrow \qquad r := \frac{1}{1 + \dfrac{B}{|a_0|}} < |x_k|$$

A negatív valós gyökök a $[-R\,,\,-r]$, a pozitív valós gyökök az $[r\,,\,R]$ intervallumban helyezkednek el.

■

6.9. NEWTON - MÓDSZER POLINOMOKRA

Legyen $f(x) = 0 \qquad x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$

$$\boxed{P_n(x) = 0}$$

Ha polinomokra alkalmazzuk a Newton - módszert, akkor a feladat $P_n(x_0)$ és $P_n'(x_0)$ helyettesítési értékek meghatározása.

(⋆) $P_n(x) = (x - x_0)\,Q_{n-1}(x) + b_0$ alakban írható, ugyanis $P_n(x)$ - re a Taylor - formulát x_0 - ban felírva:

$$P_n(x) = P_n(x_0) + P_n'(x_0)(x - x_0) + \ldots + \frac{P^{(n)}(x_0)}{n!}(x - x_0)^n$$

$$P_n(x) = (x - x_0)\,Q_{n-1}(x) + P_n(x_0)\,, \qquad \text{ahol}$$

$$Q_{n-1}(x) = b_n x^{n-1} + b_{n-1}x^{n-2} + \ldots + b_2 x + b_1 \qquad \text{és} \qquad \boxed{P_n(x_0) = b_0}$$

Q_{n-1} - et *(⋆)* - ba írva $\quad\Longrightarrow$

$$\Longrightarrow \qquad P_n(x) = (x - x_0)\left(b_n x^{n-1} + b_{n-1}x^{n-2} + \ldots + b_1\right) + b_0 =$$

$$= b_n x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0 - b_n x_0 x^{n-1} - b_{n-1}x_0 x^{n-2} - \ldots - b_1 x_0 =$$

$$= b_n x^n + (b_{n-1} - b_n x_0)\,x^{n-1} + \ldots + (b_1 - b_2 x_0)\,x + (b_0 - b_1 x_0)$$

Ezt $P_n(x)$ eredeti alakjával összehasonlítva $\Longrightarrow$

$b_n = a_n$

$a_{n-1} = b_{n-1} - b_n x_0 \qquad \Longrightarrow \qquad b_{n-1} = a_{n-1} + b_n x_0$

$$\boxed{b_k = a_k + b_{k+1} x_0}$$

$b_0 = a_0 + b_1 x_0 = P_n(x_0)$

(Horner - algoritmus)

Derivált kiszámolása :

$$P_n'(x) = (x - x_0)\, Q_{n-1}'(x) + Q_{n-1}(x) \qquad \Longrightarrow \qquad P_n'(x_0) = Q_{n-1}(x_0)$$

Most a Horner - algoritmust a $Q_{n-1}(x_0)$ kiszámítására lehet alkalmazni:

$c_n = b_n$

$$\boxed{c_k = b_k + c_{k+1} x_0} \qquad k = n-1, \ldots, 1$$

$c_1 = Q_{n-1}(x_0) = P_n'(x_0)$

7. NEMLINEÁRIS EGYENLETRENDSZEREK ITERÁCIÓS MEGOLDÁSI MÓDSZEREI

Tekintsük a következő egyenletrendszert:

$$\textit{(1)} \qquad \left.\begin{aligned} f_1(x_1, \ldots, x_n) &= 0 \\ f_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \ldots, x_n) &= 0 \end{aligned}\right\} \qquad \Longrightarrow \qquad F(\underline{\mathbf{x}}) = \underline{\mathbf{0}}\,, \quad \underline{\mathbf{x}} \in \mathbb{R}^n\,, \quad F : D \to \mathbb{R}^n\,, \quad D \subset \mathbb{R}^n$$

Az *(1)* - et is megoldhatjuk iterációval, illetve a Newton - módszerrel.
Igaz a Banach - féle fixponttétel következő speciális esete:

Tétel

Írjuk *(1)* - et vele ekvivalens alakba:

(2) $\underline{\mathbf{x}} = \Phi(\underline{\mathbf{x}})$

Tegyük fel, hogy Φ függvény folytonos és a parciális deriváltjai is folytonosak.

$\Phi : D \to D$ alakú, ahol $D \subset \mathbb{R}^n$ zárt halmaz.

Ekkor a $\underline{\mathbf{x}} = \Phi(\underline{\mathbf{x}})$ egyenletnek létezik $\underline{\mathbf{x}}^*$ fixpontja D - ben, továbbá, ha

$$\left| \frac{\partial \Phi_i}{\partial x_j} \right| < \frac{\alpha}{n}, \quad \text{ahol} \quad 0 < \alpha < 1 \quad \text{állandó,}$$

akkor az $\underline{\mathbf{x}}^*$ megoldás egyértelmű és az $\underline{\mathbf{x}}_{k+1} = \Phi(\underline{\mathbf{x}}_k)$ iteráció tetszőleges $\underline{\mathbf{x}}_0 \in D$ esetén $\underline{\mathbf{x}}^*$ - hoz tart.

$$\| \underline{\mathbf{x}}_k - \underline{\mathbf{x}}^* \| \le \frac{\alpha^k}{1-\alpha} \| \underline{\mathbf{x}}_1 - \underline{\mathbf{x}}_0 \|$$

Megjegyzés :

A Φ kontrakciója elégséges, de nem szükséges feltétele a megoldás létezésének.

Tétel **Brouwer - féle fixponttétel**

Legyen $D \subset \mathbb{R}^n$, $\Phi : D \to D$ folytonos függvény.

$D = \{ \underline{\mathbf{x}} \in \mathbb{R}^n : \| \underline{\mathbf{x}} \| \le 1 \}$, D zárt egységgömb $\Longrightarrow$ $\exists$ $\underline{\mathbf{x}}^* \in D$, $\underline{\mathbf{x}}^* = \Phi(\underline{\mathbf{x}}^*)$

7.1. NEWTON - MÓDSZER

Legyen $F(\underline{\mathbf{x}}) = \underline{\mathbf{0}}$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Az $F(\underline{\mathbf{x}})$ függvényt linearizáljuk $\underline{\mathbf{x}}_0$ környezetében.

$F(\underline{\mathbf{x}}) \approx F(\underline{\mathbf{x}}_0) + F'(\underline{\mathbf{x}}_0)(\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) = \underline{\mathbf{0}}$, ahol $F'(\underline{\mathbf{x}}_0)$ az F függvény úgynevezett Jacobi - mátrixa (derivált mátrix).

$$F'(\underline{\mathbf{x}}) := \left(\frac{\partial f_i(\underline{\mathbf{x}})}{\partial x_j} \right), \qquad i, j = 1, \ldots, n$$

$$F'(\underline{\mathbf{x}}) = \begin{bmatrix} \dfrac{\partial f_1(\underline{\mathbf{x}})}{\partial x_1} & \cdots & \dfrac{\partial f_1(\underline{\mathbf{x}})}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_n(\underline{\mathbf{x}})}{\partial x_1} & \cdots & \dfrac{\partial f_n(\underline{\mathbf{x}})}{\partial x_n} \end{bmatrix}$$

(1) - re a Newton - módszer:

$$F(\underline{\mathbf{x}}_1) + F'(\underline{\mathbf{x}}_1)(\underline{\mathbf{x}} - \underline{\mathbf{x}}_1) = \underline{\mathbf{0}}$$

(3)
$$\boxed{F(\underline{\mathbf{x}}_k) + F'(\underline{\mathbf{x}}_k)(\underline{\mathbf{x}} - \underline{\mathbf{x}}_k) = \underline{\mathbf{0}}} \qquad k = 0, 1, 2, \ldots$$

A Newton - módszer realizálásához szükséges feltétel, hogy létezzen az $\left[F'(\underline{\mathbf{x}}_k)\right]^{-1}$ inverz mátrix, ugyanis ***(3)*** - ból $\underline{\mathbf{x}}_{k+1}$ - re rendezve $\Longrightarrow$

$$\boxed{\underline{\mathbf{x}}_{k+1} = \underline{\mathbf{x}}_k - \left[F'(\underline{\mathbf{x}}_k)\right]^{-1} \cdot F(\underline{\mathbf{x}}_k)} \qquad k = 0, 1, 2, \ldots$$

A gyakorlatban sokszor minden egyes lépésben a ***(3)*** lineáris egyenletrendszert oldjuk meg, majd az így kapott $\underline{\mathbf{x}}_{k+1} - \underline{\mathbf{x}}_k$ különbséget $\underline{\mathbf{x}}_k$ - hoz adva kapjuk a $(k+1)$ - edik közelítő megoldást.

Koordinátás alakja :

$$\boxed{f_i(\underline{\mathbf{x}}_k) + \sum_{j=1}^{n} \frac{\partial f_i(\underline{\mathbf{x}}_k)}{\partial x_j} \left(x_j^{k+1} - x_j^k \right) = 0} \qquad i = 1, \ldots, n$$

7.2. MÓDOSÍTOTT NEWTON - MÓDSZER

Alakja vagy

$$\boxed{F(\underline{\mathbf{x}}_k) + F'(\underline{\mathbf{x}}_0)\left(\underline{\mathbf{x}}_{k+1} - \underline{\mathbf{x}}_k\right) = \underline{\mathbf{0}}} \qquad k = 0, 1, 2, \ldots$$

vagy pedig

$$\boxed{\underline{\mathbf{x}}_{k+1} = \underline{\mathbf{x}}_k - \left[F'(\underline{\mathbf{x}}_0)\right]^{-1} F(\underline{\mathbf{x}}_k)} \qquad k = 0, 1, 2, \ldots$$

Így nem kell minden egyes lépésben a Jacobi - mátrixot, illetve inverzét számolni, hanem csak egyszer. Ezt általában akkor használjuk, ha a Jacobi - mátrix kiszámítása munkaigényes. Így azonban a konvergenciasebesség csökken, lineáris lesz csak.

Megjegyzés :

A Newton - módszer egyenletrendszerre levezethető abból az elvből is, mint egyváltozós esetben, hogy másodrendű konvergenciagyorsaságot nyerjünk. Most a $g(x)$ függvény helyett egy $A(\underline{\mathbf{x}})$ mátrixszal kell szorozni $F(\underline{\mathbf{x}}) = \underline{\mathbf{0}}$ - t, majd levezethető, hogy

$$A(\underline{\mathbf{x}}) = -\left[F'(\underline{\mathbf{x}})\right]^{-1},$$

amivel az $\underline{\mathbf{x}} = \underline{\mathbf{x}} + A(\underline{\mathbf{x}})\,F(\underline{\mathbf{x}})$ $\quad (\underline{\mathbf{x}} = G(\underline{\mathbf{x}}))$ iterációs módszer a másodrendű Newton - módszert adja.

$\boxed{\textbf{Tétel}}$

Tegyük fel, hogy $\underline{\mathbf{x}}^*$ az $\underline{\mathbf{x}} = G(\underline{\mathbf{x}})$ egy megoldása, ahol $G = (g_1, \dots, g_n) \in \mathbb{R}^n \to \mathbb{R}^n$.

Ha létezik $\delta > 0$ szám a következő tulajdonságokkal:

i.)

$$\frac{\partial g_i(\underline{\mathbf{x}})}{\partial x_j} \in C(N_\delta)\,, \quad \text{ahol} \quad N_\delta := \{\,\underline{\mathbf{x}} \;:\; \|\,\underline{\mathbf{x}} - \underline{\mathbf{x}}^*\,\| < \delta\}\,, \qquad i\,,\, j = 1, \dots, n$$

ii.)

$$\frac{\partial^2 g_i(\underline{\mathbf{x}})}{\partial x_j\,\partial x_k} \quad \text{folytonos és} \quad \left|\frac{\partial^2 g_i(\underline{\mathbf{x}})}{\partial x_j\,\partial x_k}\right| \le M\,,$$

valamilyen M állandóval, $\underline{\mathbf{x}} \in N_\delta$ esetén $(i\,,\, j\,,\, k = 1, \dots, n)$.

iii.)

$$\frac{\partial g_i(\underline{\mathbf{x}}^*)}{\partial x_j} = 0\,, \qquad i\,,\, j = 1, \dots, n$$

akkor az $\underline{\mathbf{x}}_k = G\left(\underline{\mathbf{x}}_{k-1}\right)$ sorozat négyzetesen konvergál $\underline{\mathbf{x}}^*$ - hoz tetszőleges $\underline{\mathbf{x}}_0 \in N_\delta$ esetén, és

$$\|\,\underline{\mathbf{x}}_k - \underline{\mathbf{x}}^*\,\|_\infty \le \frac{n^2 M}{2}\,\|\,\underline{\mathbf{x}}_{k-1} - \underline{\mathbf{x}}^*\,\|_\infty^2 \qquad \forall \quad k \ge 1 \text{ - re.}$$

8. SAJÁTÉRTÉKFELADATOK

Sajátértékek meghatározásának feladata gyakran merül fel a matematikában és alkalmazásaiban. Például fizikában rezgési problémák vizsgálatában, ahol a rendszer sajátfrekvenciái a matematikai modell (lineáris egyenletrendszer) mátrixa sajátértékeinek felelnek meg.

FELADAT :

Legyen $A \in \mathbb{K}^{n \times n}$ mátrix.

Határozzuk meg azt a $\lambda \in \mathbb{K}$ számot, és $\underline{\mathbf{x}} \in \mathbb{K}^n$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ vektort, amelyre teljesül, hogy

(1) $A\underline{\mathbf{x}} = \lambda \underline{\mathbf{x}}$

Ha ez teljesül, akkor λ az A mátrix sajátértéke és $\underline{\mathbf{x}}$ a λ - hoz tartozó sajátvektor.

Ha $\underline{\mathbf{x}}_1$, $\underline{\mathbf{x}}_2$ sajátvektora λ - nak $\Longrightarrow$

$$\Longrightarrow \quad A(\underline{\mathbf{x}}_1 + \underline{\mathbf{x}}_2) = \lambda(\underline{\mathbf{x}}_1 + \underline{\mathbf{x}}_2) \quad \text{és} \quad A(c \cdot \underline{\mathbf{x}}_1) = \lambda \cdot c \cdot \underline{\mathbf{x}}_1$$

Ha $\lambda \in \mathbb{K}$ az A mátrix sajátértéke, akkor az

$$E(\lambda) := \{\underline{\mathbf{x}} \in \mathbb{K}^n \ : \ A\underline{\mathbf{x}} = \lambda \underline{\mathbf{x}}\} \quad \text{halmaz}$$

lineáris altere $\mathbb{K}^n$ - nek, amit a λ sajátértékhez tartozó sajátaltérnek nevezünk.

Legyen $E(\lambda)$ dimenziója $d(\lambda)$.

$$d(\lambda) = n - \operatorname{rang}(A - \lambda I)$$

Ha λ az A sajátértéke, akkor $d(\lambda) > 0$.

$d(\lambda)$ - t a λ sajátérték geometriai multiplicitásának nevezzük.

Állítás

$$d(\lambda) > 0 \quad \Longleftrightarrow \quad (A - \lambda I) \text{ mátrix szinguláris.}$$

Ez azt jelenti, hogy a λ az A mátrix sajátértéke $\Longleftrightarrow$ ha λ a

(2) $P(\lambda) := \det(A - \lambda I)$

karakterisztikus polinom gyöke.

$$P(\lambda) = (\lambda - \lambda_1)^{n_1} \cdot \ldots \cdot (\lambda - \lambda_i)^{n_i} \cdot \ldots \cdot (\lambda - \lambda_k)^{n_k}$$

A λ_i sajátérték algebrai multiplicitása: $v(\lambda_i) = n_i$

Állítás

$0 \leq d(\lambda_i) \leq v(\lambda_i) \leq n$

Ha $d(\lambda_i) = v(\lambda_i)$ minden λ_i - re, akkor az A mátrixot egyszerű struktúrájúnak nevezzük és ekkor létezik $\mathbb{K}$ - ben sajátvektorokból álló bázis. Ekkor azt mondjuk, hogy A - nak létezik teljes sajátvektor - rendszere. Az ilyen A mátrixok diagonalizálhatók, azaz létezik olyan T mátrix, amelyre teljesül, hogy

$T^{-1}AT = \Lambda$, Λ diagonális mátrix

Λ elemei a főátlóban az A mátrix sajátértékei és a T mátrix oszlopvektorai az A sajátvektorai. Ez a numerikus módszerek szempontjából is fontos tulajdonság, tekintve, hogy tetszőleges $\underline{\mathbf{x}} \in \mathbb{K}^n$ vektor az A mátrix sajátvektorainak lineáris kombinációjaként kifejezhető.

A karakterisztikus polinom gyökeinek meghatározásával numerikus szempontból nem ajánlatos egy mátrix sajátértékeit számolni, ugyanis a karakterisztikus polinom együtthatói általában közelítőleg adottak, és $P(\lambda)$ gyökei, főleg ha többszörös gyökök, igen érzékenyek a polinom együtthatóinak változására, így pontatlan eredményt kaphatunk. Numerikus módszerre van szükség!

Példa :

Az A mátrix és az $A_T := T^{-1}AT$ transzformált mátrix sajátértékei megegyeznek, ugyanis:

$$\det\left(T^{-1}AT - \lambda I\right) = \det\left(T^{-1}(A - \lambda I)\,T\right) = \det(A - \lambda I)$$

De a λ - hoz tartozó sajátvektor megváltozik, ugyanis:

$$A\underline{\mathbf{x}} = \lambda\underline{\mathbf{x}}$$

$$T^{-1}ATT^{-1}\underline{\mathbf{x}} = T^{-1}\lambda\underline{\mathbf{x}}$$

$$\Longrightarrow \qquad T^{-1}AT\left(T^{-1}\underline{\mathbf{x}}\right) = \lambda\left(T^{-1}\underline{\mathbf{x}}\right) ,$$

tehát a $T^{-1}AT$ mátrix λ sajátértékéhez tartozó sajátvektor $T^{-1}\underline{\mathbf{x}}$, ahol $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ az A mátrix λ sajátértékéhez tartozó sajátvektor.

8.1. SAJÁTÉRTÉKEK LOKALIZÁCIÓJA

Tudjuk, hogy $A\underline{\mathbf{x}} = \lambda\underline{\mathbf{x}}$, $A \in \mathbb{K}^{n\times n}$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ esetén $|\lambda| \leq \|A\|$, azaz $\rho(A) \leq \|A\|$,

ahol $\|A\|$ tetszőleges vektornorma által indukált mátrixnorma.

Hermitikus (szimmetrikus) mátrix sajátértékei valósak és különböző λ_i , λ_j $(i \neq j)$ sajátértékekhez tartozó sajátvektorok merőlegesek egymásra $\left(\underline{\mathbf{x}}^i \perp \underline{\mathbf{x}}^j\right)$.
Ferdén hermitikus mátrix sajátértékei a képzetes tengelyen helyezkednek el: $\mathrm{Re}\,\lambda_i = 0 \quad \forall \quad i$ - re.

Tétel **Gerschgorin - tétel**

Legyen $A \in \mathbb{K}^{n\times n}$.

Ekkor az A mátrix sajátértékei a $|\lambda - a_{ii}| \leq R_i$ úgynevezett Gerschgorin körök egyesített halmazában helyezkednek el, ahol az a_{ii} középpontú körök sugara:

$$R_i := \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|$$

Bizonyítás

Legyen az A mátrix λ sajátértékéhez tartozó sajátvektor $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$.

Ekkor $A\underline{\mathbf{x}} = \lambda\underline{\mathbf{x}} \qquad (\underline{\mathbf{x}} \neq \underline{\mathbf{0}})$.

Koordinátánként írva: $\displaystyle\sum_{j=1}^{n} a_{ij}\, x_j = \lambda x_i\,, \qquad i = 1,\ldots,n$

Rendezve: $\displaystyle\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}\, x_j = (\lambda - a_{ii})\, x_i\,, \qquad i = 1,\ldots,n$

Legyen $|x_i| = \max\limits_{1\leq j\leq n} |x_j|$, azaz x_i az $\underline{\mathbf{x}}$ sajátvektor legnagyobb abszolútértékű koordinátája

$$\Longrightarrow \qquad \lambda - a_{ii} = \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}\frac{x_j}{x_i}$$

$$\Longrightarrow \qquad |\lambda - a_{ii}| = \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\,\frac{|x_j|}{|x_i|} \leq \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| = R_i$$

■

Megjegyzés :

Legyen A^T mátrixra $R_j' := \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$.

Ekkor az A mátrix sajátértékei az

$$\left(\bigcup_{i=1}^{n} K_{R_i}\right) \bigcap \left(\bigcup_{j=1}^{n} K_{R_j'}\right)$$

halmazban találhatók, ami általában szűkebb halmaz.

Lemma

Ha az A mátrix erős - sorösszeg tulajdonságú $\Longrightarrow$ A nem szinguláris.

Bizonyítás

Indirekt.

Tegyük fel, hogy A szinguláris mátrix, azaz $\exists$ $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, amelyre $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$

$\Longrightarrow$ $A\underline{\mathbf{x}} = 0 \cdot \underline{\mathbf{x}}$

Ez azt jelenti, hogy a 0 az A mátrix sajátértéke.

Ekkor a Gerschgorin - tétel megismétlésével azt kapjuk, hogy

$$|a_{ii}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

Ez pedig ellentmondás, ugyanis az A mátrix erős - sorösszeg tulajdonságú, azaz

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| < |a_{ii}|$$

$\blacksquare$

Tétel

Ha a Gerschgorin körök közül j darab $(1 \leq j < n)$ diszjunkt a többi körtől, akkor ezen j - kör uniójában pontosan j darab sajátérték van.

8.2. HATVÁNY - MÓDSZER LEGNAGYOBB ABSZOLÚTÉRTÉKŰ SAJÁTÉRTÉK MEGHATÁROZÁSÁRA

Sok gyakorlati esetben nincs szükségünk egy mátrix összes sajátértékére, hanem csak néhányra, például a legnagyobb és legkisebb abszolútértékűre. Erre alkalmas a hatvány - módszer.

Tegyük fel, hogy az $A \in \mathbb{K}^{n\times n}$ mátrix diagonalizálható, azaz A - nak létezik sajátvektorokból álló bázisa $\mathbb{K}^n$ - ben : $\underline{\mathbf{x}}_1, \ldots, \underline{\mathbf{x}}_n$

($\star$) $\quad A\underline{\mathbf{x}}_i = \lambda_i \underline{\mathbf{x}}_i$

Ekkor tetszőleges $\underline{\mathbf{x}}_0 \in \mathbb{K}^n$ vektor a következő alakban írható:

$$\underline{\mathbf{v}}_0 = \sum_{i=1}^{n} \alpha_i \underline{\mathbf{x}}_i$$

Tegyük fel, hogy az A mátrix sajátértékei csökkenő sorrendben vannak rendezve:

$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$, ahol λ_i az $\underline{\mathbf{x}}_i$ sajátvektorhoz tartozó sajátérték.

Képezzük $\underline{\mathbf{x}}_0$ - ból kiindulva a következő sorozatot:

(1) $$\underline{\mathbf{v}}_1 = A\underline{\mathbf{v}}_0 = A\left(\sum_{i=1}^{n} \alpha_i \underline{\mathbf{x}}_i\right) = \sum_{i=1}^{n} \alpha_i A\underline{\mathbf{x}}_i \overset{(\star)}{=} \sum_{i=1}^{n} \alpha_i \lambda_i \underline{\mathbf{x}}_i$$

(2) $$\underline{\mathbf{v}}_2 = A\underline{\mathbf{v}}_1 = A\left(\sum_{i=1}^{n} \alpha_i \lambda_i \underline{\mathbf{x}}_i\right) = \sum_{i=1}^{n} \alpha_i \lambda_i A\underline{\mathbf{x}}_i = \sum_{i=1}^{n} \alpha_i \lambda_i^2 \underline{\mathbf{x}}_i$$

(3) $$\underline{\mathbf{v}}_k = A\underline{\mathbf{v}}_{k-1} = \sum_{i=1}^{n} \alpha_i \lambda_i^k \underline{\mathbf{x}}_i$$

(4) $$\underline{\mathbf{v}}_{k+1} = A\underline{\mathbf{v}}_k = \sum_{i=1}^{n} \alpha_i \lambda_i^{k+1} \underline{\mathbf{x}}_i$$

I. ESET :

Legyen $\alpha_1 \neq 0$ és $|\lambda_1| > |\lambda_2|$

(3) $$\Longrightarrow \qquad \underline{\mathbf{v}}_k = \alpha_1 \lambda_1^k \underline{\mathbf{x}}_1 + \sum_{i=2}^{n} \alpha_i \lambda_i^k \underline{\mathbf{x}}_i$$

$$\underline{\mathbf{v}}_k = \lambda_1^k \left[\alpha_1 \underline{\mathbf{x}}_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \underline{\mathbf{x}}_i \right] \qquad \Longrightarrow$$

$$\Longrightarrow \qquad \boxed{\lim_{k\to\infty} \frac{1}{\lambda_1^k \underline{\mathbf{v}}_k} = \alpha_1 \underline{\mathbf{x}}_1} \qquad \textbf{\textit{(5)}}$$

(6) $$\lim_{k\to\infty} \frac{1}{\lambda_1^{k+1} \underline{\mathbf{v}}_{k+1}} = \alpha_1 \underline{\mathbf{x}}_1$$

(5), ***(6)*** - ot szorozzuk skalárisan $\underline{\mathbf{y}}$ tetszőleges vektorral, amely nem merőleges $\underline{\mathbf{x}}_1$ - re.

(7) $$\frac{1}{\lambda_1^k} \left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{v}}_k \right) = \alpha_1 \left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{x}}_1 \right)$$

(8) $$\frac{1}{\lambda_1^{k+1}} \left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{v}}_{k+1} \right) = \alpha_1 \left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{x}}_1 \right)$$

Majd ***(8)*** - at osztva ***(7)*** - tel $\Longrightarrow$ $\boxed{\lim_{k\to\infty} \frac{\left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{v}}_{k+1} \right)}{\left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{v}}_k \right)} = \lambda_1}$ ***(9)***

Megjegyzés :

1.

Ha $A \in \mathbb{R}^{n\times n}$, akkor $|\lambda_1| > |\lambda_2|$ $\Longrightarrow$ $\lambda_1 \in \mathbb{R}$, ugyanis ha $\lambda \in \mathbb{C}$ sajátérték $\Longrightarrow$ $\overline{\lambda}$ is az. Így, ha $\underline{\mathbf{v}}_1 \in \mathbb{R}^n$, akkor az iterációs folyamat is valós értéket ad.

2.

Célszerű $\underline{\mathbf{v}}_k$ - t minden iterációs lépésben normalizálni, így elkerülhetők az iteráltak méreteinek nagy változásai.

3.

Az $\underline{\mathbf{y}}$ vektort célszerű úgy választani, hogy $\underline{\mathbf{v}}_k$ legnagyobb abszolútértékű helyén 1 , másutt pedig 0 álljon. Ez csökkenti a szükséges számításokat és csökkenti az osztás közben fellépő kerekítési hibákat. A gyakorlati számításnál így egymás utáni $\underline{\mathbf{v}}_k$ vektorok maximum abszolútértékű komponenseinek hányadosával közelítjük λ_1 - et.

4.

Az $\alpha_1 \neq 0$ feltétel gyakorlatban felesleges, ugyanis ha $\alpha_1 = 0$ eleinte, a kerekítési hibák miatt később úgyis $\alpha_1 \neq 0$ lesz. Ez az iterációt végülis a λ_1 értékre állítja be.

A λ sajátértékhez tartozó sajátvektor is számolható.

Tekintsük:

$$\frac{A\underline{\mathbf{v}}_k}{\| \underline{\mathbf{v}}_k \|} = \frac{\underline{\mathbf{v}}_{k+1}}{\| \underline{\mathbf{v}}_k \|} \approx \frac{\alpha_1 \lambda_1^{k+1} \underline{\mathbf{x}}_1}{\| \underline{\mathbf{v}}_k \|} = \frac{\lambda_1 \left(\alpha_1 \lambda_1^k \underline{\mathbf{x}}_1 \right)}{\| \underline{\mathbf{v}}_k \|} \approx \lambda_1 \frac{\underline{\mathbf{v}}_k}{\| \underline{\mathbf{v}}_k \|}$$

$$A \left(\frac{\underline{\mathbf{v}}_k}{\| \underline{\mathbf{v}}_k \|} \right) \approx \lambda_1 \left(\underline{\mathbf{v}}_k \right) \| \underline{\mathbf{v}}_k \|$$

$$\boxed{\lim_{n \to \infty} \frac{\underline{\mathbf{v}}_k}{\| \underline{\mathbf{v}}_k \|} = \underline{\mathbf{x}}_1}$$

Megjegyzés :

A *(9)* konvergenciasebessége a $\left| \dfrac{\lambda_2}{\lambda_1} \right|$ hányados nagyságától függ.

<u>II. ESET :</u>

Legyen $|\lambda_1| = \ldots = |\lambda_r| > |\lambda_{r+1}| > \ldots$

Az előzőekhez hasonló gondolatmenettel $\Longrightarrow$ $\boxed{\lambda_1 = \lim_{k \to \infty} \frac{\left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{v}}_{k+1} \right)}{\left(\underline{\mathbf{y}} \, , \, \underline{\mathbf{v}}_k \right)}}$

A konvergencia sebessége a $\left| \dfrac{\lambda_{r+1}}{\lambda_1} \right|$ hányadostól függ.

Ekkor azonban a $\dfrac{\underline{\mathbf{v}}_k}{\| \underline{\mathbf{v}}_k \|}$ vektorsorozat a λ_1 - hez tartozó r darab lineárisan független sajátvektorok egyikéhez fog tartozni. Ha azonban rendre különböző $\underline{\mathbf{v}}_0$ vektorból indulunk ki, akkor megkaphatjuk a többi λ_1 - hez tartozó sajátvektort is.

III. ESET :

Legyen $\lambda_1 = \ldots = \lambda_r = -\lambda_{r+1} = \ldots = -\lambda_{r+s}$

Ekkor $$\underline{\mathbf{v}}_k = \left(\sum_{i=1}^{r} \alpha_i \underline{\mathbf{x}}_i + (-1)^k \sum_{i=r+1}^{r+s} \alpha_i \underline{\mathbf{x}}_i \right) \lambda_1^k + \sum_{i=r+s+1}^{n} \alpha_i \lambda_i^k \underline{\mathbf{x}}_i \,.$$

Legyen $k = 2m$.

$$\Longrightarrow \qquad \lambda_1^2 = \lim_{m\to\infty} \frac{\left(\underline{\mathbf{y}}\,,\, \underline{\mathbf{v}}_{2m+2}\right)}{\left(\underline{\mathbf{y}}\,,\, \underline{\mathbf{v}}_{2m}\right)} \,.$$

Belátható, hogy

$$\underline{\mathbf{v}}_{2m+1} + \lambda_1 \underline{\mathbf{v}}_{2m} \qquad \text{és} \qquad \underline{\mathbf{v}}_{2m+1} - \lambda_1 \underline{\mathbf{v}}_{2m}$$

vektorok normáltjai rendre a λ_1 illetve a $-\lambda_1$ sajátértékeknek megfelelő sajátvektorokhoz tartanak.

IV. ESET :

Legyen $\lambda_1 = r \cdot e^{i\varphi}$, $\quad \lambda_2 = r \cdot e^{-i\varphi} \qquad \Longrightarrow \qquad |\lambda_1| = |\lambda_2| > |\lambda_3| \ldots$

Ekkor fennállnak a következő közelítő egyenlőségek:

1. $\underline{\mathbf{v}}_{m+2} = \alpha_1 \lambda_1^{m+2} \underline{\mathbf{x}}_1 + \alpha_2 \lambda_2^{m+2} \underline{\mathbf{x}}_2$

2. $\underline{\mathbf{v}}_{m+1} = \alpha_1 \lambda_1^{m+1} \underline{\mathbf{x}}_1 + \alpha_2 \lambda_2^{m+1} \underline{\mathbf{x}}_2$

3. $\underline{\mathbf{v}}_{m} = \alpha_1 \lambda_1^{m} \underline{\mathbf{x}}_1 + \alpha_2 \lambda_2^{m} \underline{\mathbf{x}}_2$

A **2.** egyenletet $-(\lambda_1 + \lambda_2)$ - vel, a **3** - at $\lambda_1\lambda_2$ - vel szorozva:

$$\underline{\mathbf{v}}_{m+2} = \alpha_1 \lambda_1^{m+2} \underline{\mathbf{x}}_1 + \alpha_2 \lambda_2^{m+2} \underline{\mathbf{x}}_2$$

$$-(\lambda_1 + \lambda_2)\, \underline{\mathbf{v}}_{m+1} = -(\lambda_1 + \lambda_2)\, \alpha_1 \lambda_1^{m+1} \underline{\mathbf{x}}_1 - (\lambda_1 + \lambda_2)\, \alpha_2 \lambda_2^{m+1} \underline{\mathbf{x}}_2$$

$$\lambda_1\lambda_2 \underline{\mathbf{v}}_{m} = \lambda_1\lambda_2 \alpha_1 \lambda_1^{m} \underline{\mathbf{x}}_1 + \lambda_1\lambda_2 \alpha_2 \lambda_2^{m} \underline{\mathbf{x}}_2$$

$+$

$$\underline{\mathbf{v}}_{m+2} - (\lambda_1 + \lambda_2)\, \underline{\mathbf{v}}_{m+1} + \lambda_1\lambda_2 \underline{\mathbf{v}}_{m} = \alpha_1 \lambda_1^{m} \underline{\mathbf{x}}_1 \left[\lambda_1^2 - (\lambda_1 + \lambda_2)\, \lambda_1 + \lambda_1\lambda_2\right] + \alpha_2 \lambda_2^{m} \underline{\mathbf{x}}_2 \left[\lambda_2^2 - (\lambda_1 + \lambda_2)\, \lambda_2 + \lambda_1\lambda_2\right]$$

Bevezetve:

$$A := -\left(\lambda_1 + \lambda_2\right)\,, \qquad B := \lambda_1\lambda_2 \qquad \Longrightarrow$$

$$(\star) \qquad \underline{\mathbf{v}}_{m+2} + A\underline{\mathbf{v}}_{m+1} + B\underline{\mathbf{v}}_m = \alpha_1\lambda_1^m\underline{\mathbf{x}}_1\left[\lambda_1^2 + A\lambda_1 + B\right] + \alpha_2\lambda_{12}^m\underline{\mathbf{x}}_2\left[\lambda_2^2 + A\lambda_2 + B\right] = \underline{\mathbf{0}}$$

A $\lambda_1\,,\lambda_2$ a $z^2 + Az + B = 0$ alakú másodfokú polinom gyöke.

Ha a számítások során $\underline{\mathbf{v}}_{m+2} + A\underline{\mathbf{v}}_{m+1} + B\underline{\mathbf{v}}_m = \underline{\mathbf{0}}$ egyenlőséget tapasztalunk, akkor a $\lambda_1\,,\lambda_2$ sajátértékeket a $(\star)$ - ból kiszámolhatjuk.

Feladat : A és B együtthatók meghatározása.

A $\underline{\mathbf{v}}_{m+2} + A\underline{\mathbf{v}}_{m+1} + B\underline{\mathbf{v}}_m = \underline{\mathbf{0}}$ egyenletet felírva $r\,,s$ koordinátákon $\Longrightarrow$

$$\left.\begin{aligned} v_{r\,,m+2} + Av_{r\,,m+1} + Bv_{r\,,m} &= 0 \\ v_{s\,,m+2} + Av_{s\,,m+1} + Bv_{s\,,m} &= 0 \end{aligned}\right\} \quad \text{Innen } A \text{ és } B \text{ meghatározható.}$$

$$A = -\frac{v_{r\,,m}v_{s\,,m+2} - v_{s\,,m+1}v_{r\,,m+2}}{v_{r\,,m}v_{s\,,m+1} - v_{s\,,m}v_{r\,,m+1}}$$

$$B = \frac{v_{r\,,m+1}v_{s\,,m+2} - v_{s\,,m}v_{r\,,m+2}}{v_{r\,,m}v_{s\,,m+1} - v_{s\,,m}v_{r\,,m+1}}$$

A és B ismeretében $(\star)$ - ból $\lambda_1\,,\lambda_2$ számolható.

A sajátvektorok meghatározása :

$$\underline{\mathbf{v}}_{m+1} - \lambda_1\underline{\mathbf{v}}_m \approx \alpha_2\lambda_2^m\left(\lambda_2 - \lambda_1\right)\underline{\mathbf{x}}_2\,,$$

$$\underline{\mathbf{v}}_{m+1} - \lambda_2\underline{\mathbf{v}}_m \approx \alpha_1\lambda_1^m\left(\lambda_1 - \lambda_2\right)\underline{\mathbf{x}}_1$$

vektorok normáltjai rendre az $\underline{\mathbf{x}}_1$ és $\underline{\mathbf{x}}_2$ sajátvektorok közelítései lesznek, ugyanis a bal oldali vektorok az $\underline{\mathbf{x}}_1$ és $\underline{\mathbf{x}}_2$ sajátvektorokkal megegyező irányúak, így normálás után felhasználhatók azok közelítésére.

8.3. JACOBI MÓDSZERE SZIMMETRIKUS (VALÓS) MÁTRIX SAJÁTÉRTÉKEINEK, SAJÁTVEKTORAINAK MEGHATÁROZÁSÁRA

E módszer közvetlenül számolja egy A szimmetrikus mátrix összes sajátértékét (sajátvektorát) végtelen sok iterációs lépésben.

Legyen $A \in \mathbb{R}^{n \times n}$ - es szimmetrikus mátrix $\Longrightarrow$ A minden sajátértéke valós, továbbá létezik olyan Q ortogonális mátrix, amellyel A diagonális alakra transzformálható.

$Q^T \cdot A \cdot Q = D$ a D diagonális elemei az A sajátértékei, a Q pedig megadja a sajátvektorokat.

Alapgondolat :

Olyan $\{\, \Omega_k \,\}$ ortogonális mátrixokból álló sorozatot konstruálunk, amelyre teljesül, hogy:

$\lim\limits_{k \to \infty} \Omega_1 \cdot \ldots \cdot \Omega_k = Q$ Legyen:

$$A = A_0$$

$$A_1 = \Omega_1^T \cdot A_0 \cdot \Omega_1$$

$$\vdots$$

$$A_k = \Omega_k^T \cdot A_{k-1} \cdot \Omega_k = \Omega_k^T \cdot \ldots \cdot \Omega_1^T \cdot A_0 \underbrace{\Omega_1 \cdot \ldots \cdot \Omega_k}_{Q}$$

Itt Ω_k ortogonális mátrix. Belátható, hogy ha A szimmetrikus mátrix, akkor A_k is szimmetrikus mátrix lesz.

Definíció :

Jacobi - forgatásnak (rotációnak) nevezzük a következő $(n \times n)$ - es mátrixot:

$$\Omega_{ij}(\varphi) := \begin{array}{cl} \begin{bmatrix} 1 & 0 & & & & \cdots & & & & 0 \\ 0 & \ddots & & & & & & & & \\ & & 1 & & & & & & 0 & \\ & & & \cos\varphi & & \cdots & & -\sin\varphi & & \\ & & & & 1 & & & & & \\ \vdots & & & \vdots & & \ddots & & \vdots & & \vdots \\ & & & & & & 1 & & & \\ & & & \sin\varphi & & \cdots & & \cos\varphi & & \\ & & 0 & & & & & & 1 & \\ & & & & & & & & \ddots & 0 \\ 0 & & & & & \cdots & & & 0 & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \longleftarrow \ i \text{ - edik sor} \\ \\ \\ \\ \longleftarrow \ j \text{ - edik sor} \\ \\ \\ \\ \end{array} \end{array}$$

ahol $|\varphi| \leq \pi$.

Ez a mátrix ortogonális és az általa meghatározott lineáris transzformáció az i - edik és j - edik koordináta tengely φ szögű elforgatását jelenti. Hasonlósági transzformációk sorozatával „kinullázzuk" az A mátrix főátlón kívüli elemeit.

Feladat :

Határozzuk meg azt a φ szöget, amelyikkel az $\Omega_{ij}(\varphi)$ Jacobi - forgatást alkalmazva az A mátrix legnagyobb abszolútértékű főátlón kívüli eleme „kinullázódik".

Legyen $a_{i(0)j(0)}$ a maximális elem A - ban.

$\Omega^T_{i(0)j(0)} \cdot A \cdot \Omega_{i(0)j(0)} = A_1$ és az A_1 mátrix $a_{i(0)j(0)}$ eleme nulla legyen!

Mivel az $\Omega_{i(0)j(0)}$ mátrixszal szorozva az A mátrix i - edik és j - edik sora illetve oszlopa változik meg, azért elég a következőt vizsgálnunk:

$$\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} a_{i(0)i(0)} & a_{i(0)j(0)} \\ a_{i(0)j(0)} & a_{j(0)j(0)} \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} \overset{1}{a}_{i(0)i(0)} & \overset{1}{a}_{i(0)j(0)} \\ \overset{1}{a}_{i(0)j(0)} & \overset{1}{a}_{j(0)j(0)} \end{bmatrix}$$

Innen $\Longrightarrow$

$$\overset{1}{a}_{i(0)j(0)} = -a_{i(0)i(0)} \sin\varphi\cos\varphi + a_{i(0)j(0)}\cos^2\varphi - a_{i(0)j(0)}\sin^2\varphi + a_{j(0)j(0)}\sin\varphi\cos\varphi =$$

$$= \sin\varphi\cos\varphi\left[a_{j(0)j(0)} - a_{i(0)i(0)}\right] + a_{i(0)j(0)}\left[\cos^2\varphi - \sin^2\varphi\right] =$$

$$= \frac{1}{2}\left[a_{j(0)j(0)} - a_{i(0)i(0)}\right]\cdot\sin 2\varphi + a_{i(0)j(0)}\cdot\cos 2\varphi$$

Hogy $\overset{1}{a}_{i(0)j(0)} = 0$ legyen, kell, hogy:

$$(\star) \qquad \boxed{\mathbf{tg}\, 2\varphi = \frac{2a_{i(0)j(0)}}{a_{i(0)i(0)} - a_{j(0)j(0)}}} \qquad |\varphi| \le \frac{\pi}{4}$$

ha a nevező nem nulla.

Ha a nevező nulla, akkor $\cos 2\varphi = 0 \Longrightarrow \varphi = \pm\frac{\pi}{4}$ és az előjelet az $a_{i(0)j(0)}$ előjelének megfelelően választjuk.

Ez a lépés ismételhető, **k** - adik lépésben a φ szöget úgy választjuk, hogy az $\overset{k}{a}_{i(k-1)j(k-1)} = 0$ legyen.

$$(\star) \Longrightarrow \qquad \boxed{\mathbf{tg}\, 2\varphi = \frac{2a_{i(k-1)j(k-1)}}{a_{i(k-1)i(k-1)} - a_{j(k-1)j(k-1)}}} \qquad |\varphi| \le \frac{\pi}{4}$$

A forgatás végrehajtásához a φ szög kiszámítása helyett elég $\cos\varphi$, $\sin\varphi$ értékét ismernünk, ugyanis csak ezekre van szükség.

Legyen $c := \cos\varphi\,,\quad s := \sin\varphi \quad$ és $\quad r := \cos 2\varphi\ .$

Figyelembe véve, hogy

$$\cos\varphi = \frac{1}{\sqrt{1+\mathbf{tg}^2\,\varphi}} \qquad \Longrightarrow \qquad \cos\varphi = \frac{1}{\sqrt{1+\mathbf{tg}^2\,2\varphi}}$$

$$(\star) \qquad \Longrightarrow \qquad r = \cos 2\varphi = \frac{1}{\sqrt{1+\dfrac{4\left(\overset{k-1}{a}_{i(k-1)j(k-1)}\right)^2}{\left(\overset{k-1}{a}_{i(k-1)i(k-1)} - \overset{k-1}{a}_{j(k-1)j(k-1)}\right)^2}}} =$$

$$= \frac{\left|\overset{k-1}{a}_{i(k-1)i(k-1)} - \overset{k-1}{a}_{j(k-1)j(k-1)}\right|}{\sqrt{\left(\overset{k-1}{a}_{i(k-1)i(k-1)} - \overset{k-1}{a}_{j(k-1)j(k-1)}\right)^2 + 4\left(\overset{k-1}{a}_{i(k-1)j(k-1)}\right)^2}}$$

Ha ezt ismerjük, akkor
$$\begin{matrix} r = c^2 - s^2 \\ \\ r = 1 - 2s^2 \end{matrix} \qquad \Longrightarrow \qquad r = 2c^2 - 1$$

ahonnan $\quad \Longrightarrow \quad c = \sqrt{\dfrac{r+1}{2}} \quad$ és $\quad s = \sigma\sqrt{\dfrac{1-r}{2}}$

ahol $\quad \sigma = \operatorname{sign}\left[\left(\overset{k-1}{a}_{i(k-1)i(k-1)} - \overset{k-1}{a}_{j(k-1)j(k-1)}\right)\cdot \overset{k-1}{a}_{i(k-1)j(k-1)}\right]$

A $(\star)$ képlet levezetésében és c, s kiszámításában nem tettük fel, hogy $\overset{k-1}{a}_{i(k-1)j(k-1)}$ az A_{k-1} mátrix legnagyobb abszolútértékű főátlón kívüli eleme. Az eljárás működik, ha $\overset{k-1}{a}_{i(k-1)j(k-1)} \neq 0$.

Az iteráció előrehaladtával általában $\mathbf{tg}\,2\varphi \to 0\,,\quad r = \cos 2\varphi \to 1\,,\quad c = \cos\varphi \to 1\quad$ és $s = \sin\varphi \to 0\ .$

Amikor $\mathbf{tg}\,2\varphi \ll 1$ akkor $\cos 2\varphi \approx 1\,,\quad \cos\varphi \approx 1\quad$ és $\quad \sin\varphi \approx 0\ .$

Ekkor megáll az iteráció, hiszen ekkor $\overset{k}{a}_{ij} = \overset{k-1}{a}_{ij}$ és nem $\overset{k}{a}_{ij} = 0\ .$

Ezért indokoltabb más képletek használata:

Legyen $\Theta := \mathbf{ctg}\, 2\varphi$, amit *($\star$)* - ból számolhatunk:

$$\Theta = \frac{1}{\mathbf{tg}\, 2\varphi} = \frac{\overset{k-1}{a}_{i(k-1)i(k-1)} - \overset{k-1}{a}_{j(k-1)j(k-1)}}{2\ \overset{k-1}{a}_{i(k-1)j(k-1)}} ,$$

majd $\mathbf{tg}\, 2\varphi = \dfrac{2\mathbf{tg}\,\varphi}{1-\mathbf{tg}^2\,\varphi}$ alapján, ha

$$t := \mathbf{tg}\,\varphi\ , \text{ akkor } \quad \Longrightarrow \quad \frac{1}{\Theta} = \frac{2t}{1-t^2}$$

ahonnan t - re másodfokú egyenletet kapunk: $t^2 + 2\Theta t - 1 = 0$

Innen a kisebbik abszolútértékű gyököt meghatározva:

$$t = \frac{1}{\Theta + \operatorname{sign}\Theta\sqrt{1+\Theta^2}} .$$

Ezzel számolva nem akad el az iteráció.

Tétel

Legyen $A \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix.

Ekkor az

$$\Omega^T_{i(k-1)j(k-1)} \cdot A_{k-1} \cdot \Omega_{i(k-1)j(k-1)} \qquad k = 1, 2, \ldots$$

sorozat elemenként olyan diagonális mátrixhoz konvergál, amelynek főátlójában az A mátrix sajátértékei állnak.

Bizonyítás

Mivel A szimmetrikus, azért létezik olyan C ortogonális és D diagonális mátrix, hogy $A = C^T \cdot D \cdot C$ alakban írható, ahol D diagonális mátrix az A sajátértékeivel.

Definíció :

Az A mátrix nyoma: $\quad \operatorname{trace}(A) = \displaystyle\sum_{i=1}^{n} a_{ii}$

Állítás

$\text{trace } (\alpha \cdot A) = \alpha \cdot \text{trace } (A) \qquad (\alpha \in \mathbb{R})$

$\text{trace } (A \pm B) = \text{trace } (A) \pm \text{trace } (B)$

$\text{trace } (A \cdot B) = \text{trace } (B \cdot A)$

Most tekintsük az A mátrix elemeinek négyzetösszegét:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 = \text{trace } \left(A^T \cdot A\right) = \text{trace } \left(C^T \cdot D \cdot C \cdot C^T \cdot D \cdot C\right) =$$
$$= \text{trace } \left(C^T \cdot D^2 \cdot C\right) = \text{trace } \left(C \cdot C^T \cdot D^2\right) = \text{trace } \left(D^2\right) = \sum_{i=1}^{n} \lambda_i^2$$

Vezessük be: $\qquad N(A) := \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2$.

Most tekintsük az A_{k-1} - ről az A_k - ra való átmenetet.

$$N(A_k) = \sum_{i=1}^{n} \lambda_i^2 - \sum_{i=1}^{n} \left[\overset{k}{a}_{i(k-1)i(k-1)}\right]^2$$

$$N(A_{k-1}) = \sum_{i=1}^{n} \lambda_i^2 - \sum_{i=1}^{n} \left[\overset{k-1}{a}_{i(k-1)i(k-1)}\right]^2$$

Vizsgáljuk:

$$(\star) \qquad N(A_k) - N(A_{k-1}) = \sum_{i=1}^{n} \left[\overset{k-1}{a}_{i(k-1)i(k-1)}\right]^2 - \sum_{i=1}^{n} \left[\overset{k}{a}_{i(k-1)i(k-1)}\right]^2 =$$

$$= \left[\overset{k-1}{a}_{i(k-1)i(k-1)}\right]^2 + \left[\overset{k-1}{a}_{j(k-1)j(k-1)}\right]^2 - \left[\overset{k}{a}_{i(k-1)i(k-1)}\right]^2 - \left[\overset{k}{a}_{j(k-1)j(k-1)}\right]^2$$

Felírva az átmenetet:

$$\begin{bmatrix} \overset{k}{a}_{i(k-1)i(k-1)} & \overset{k}{a}_{i(k-1)j(k-1)} \\ \overset{k}{a}_{i(k-1)j(k-1)} & \overset{k}{a}_{j(k-1)j(k-1)} \end{bmatrix} = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} \overset{k-1}{a}_{i(k-1)i(k-1)} & \overset{k-1}{a}_{i(k-1)j(k-1)} \\ \overset{k-1}{a}_{i(k-1)j(k-1)} & \overset{k-1}{a}_{j(k-1)j(k-1)} \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

Tudjuk, hogy a hasonlósági transzformáció nem változtatja meg a mátrix nyomát és a determinánsának értékét, ezért tekintsük:

$$(\text{trace } (A_k))^2 - 2\det(A_k) = (\text{trace } (A_{k-1}))^2 - 2\det(A_{k-1})$$

$$\left(\overset{k}{a}_{i(k-1)i(k-1)} + \overset{k}{a}_{j(k-1)j(k-1)}\right)^2 - 2\left(\overset{k}{a}_{i(k-1)i(k-1)} \cdot \overset{k}{a}_{j(k-1)j(k-1)} - \left[\overset{k}{a}_{i(k-1)j(k-1)}\right]^2\right) =$$

$$= \left(\overset{k-1}{a}_{i(k-1)i(k-1)} + \overset{k-1}{a}_{j(k-1)j(k-1)}\right)^2 - 2\left(\overset{k-1}{a}_{i(k-1)i(k-1)} \cdot \overset{k-1}{a}_{j(k-1)j(k-1)} - \left[\overset{k-1}{a}_{i(k-1)j(k-1)}\right]^2\right)$$

$$\left[\overset{k}{a}_{i(k-1)i(k-1)}\right]^2 + \left[\overset{k}{a}_{j(k-1)j(k-1)}\right]^2 + 2\left[\overset{k}{a}_{i(k-1)j(k-1)}\right]^2 = \left[\overset{k-1}{a}_{i(k-1)i(k-1)}\right]^2 + \left[\overset{k-1}{a}_{j(k-1)j(k-1)}\right]^2 + 2\left[\overset{k-1}{a}_{i(k-1)j(k-1)}\right]^2$$

$$(\star) \quad \Longrightarrow \quad N(A_k) - N(A_{k-1}) = 2\underbrace{\left[\overset{k}{a}_{i(k-1)j(k-1)}\right]}_{0}^2 - 2\left[\overset{k-1}{a}_{i(k-1)j(k-1)}\right]^2$$

$$\Longrightarrow \quad N(A_k) = N(A_{k-1}) - 2\left[\overset{k-1}{a}_{i(k-1)j(k-1)}\right]^2$$

Igaz a következő becslés :

$$\left(\frac{n(n-1)}{2} \cdot 2\right)\left(\overset{k-1}{a}_{i(k-1)j(k-1)}\right)^2 \geq N(A_{k-1})$$

$$\Longrightarrow \quad \left(\overset{k-1}{a}_{i(k-1)j(k-1)}\right)^2 \geq \frac{N(A_{k-1})}{n(n-1)}$$

$$\Longrightarrow \quad N(A_k) \leq N(A_{k-1}) - 2 \cdot \frac{N(A_{k-1})}{n(n-1)} = N(A_{k-1})\left(1 - \frac{2}{n(n-1)}\right) =$$

$$= N(A_0) \cdot \left(1 - \frac{2}{n(n-1)}\right)^k \xrightarrow[k \to \infty]{} 0$$

■

8.4. SZIMMETRIKUS TRIDIAGONÁLIS MÁTRIX SAJÁTÉRTÉKÉNEK MEGHATÁROZÁSA

Tekintsük a következő $(n \times n)$ - es mátrixot:

$$T := \begin{bmatrix} \alpha_1 & \beta_2 & 0 & \cdots & 0 \\ \beta_2 & \alpha_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & & & \alpha_{n-1} & \beta_n \\ 0 & \cdots & 0 & \beta_n & \alpha_n \end{bmatrix}$$

T sajátértékei meghatározhatók a Newton - módszer segítségével. Tudjuk ugyanis, hogy a T sajátértékei a $P(\lambda) := \det(T - \lambda I)$ λ - ban n - edfokú polinom gyökei.

Mivel T szimmetrikus $\Longrightarrow$ T sajátértékei valósak. A $P(\lambda)$ polinom gyökeinek meghatározására pedig felhasználhatjuk a Newton - módszert anélkül, hogy a $P(\lambda)$ polinom együtthatóit kiszámolnánk.

Legyen:

$$T_i := \begin{bmatrix} \alpha_1 & \beta_2 & 0 & \cdots & 0 \\ \beta_2 & \alpha_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & & & \alpha_{i-1} & \beta_i \\ 0 & \cdots & 0 & \beta_i & \alpha_i \end{bmatrix} \quad \text{és} \quad P_i(\lambda) := \det(T_i - \lambda I) = \det \begin{bmatrix} \alpha_1 - \lambda & \beta_2 & 0 & \cdots & 0 \\ \beta_2 & \alpha_2 - \lambda & \ddots & & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & & & \alpha_{i-1} & \beta_i \\ 0 & \cdots & 0 & \beta_i & \alpha_i - \lambda \end{bmatrix}$$

Newton - módszerrel $P(\lambda)$ gyöke: $\qquad \lambda_{n+1} = \lambda_n - \dfrac{P(\lambda_n)}{P'(\lambda_n)}$.

Ez utóbbi determinánst az utolsó oszlopa szerint kifejtve $\Longrightarrow$

$$P_i(\lambda) = (\alpha_i - \lambda)\, P_{i-1}(\lambda) - \beta_i^2\, P_{i-2}(\lambda) \qquad i = 2, \ldots, n$$

$(\star)$ $\qquad P_n(\lambda) = P(\lambda) \qquad$ tetszőleges $\lambda \in \mathbb{R}$ - re

Legyen: $P_0(\lambda) = 1$, $\quad P_1(\lambda) := (\alpha_1 - \lambda)$, $\quad$ ugyanis $P_2(\lambda)$ - ra ekkor lesz jó a képlet.

$(\star)$ - ból $P(\lambda)$ értéke számolható tetszőleges $\lambda \in \mathbb{R}$ - re a $P(\lambda)$ polinom együtthatói nélkül.

A Newton - módszerhez még kell $P^{'}(\lambda)$ értéke is.

$(\star)$ - ot deriválva $\quad \Longrightarrow$

$$P_i'(\lambda) = -P_{i-1}(\lambda) + (\alpha_i - \lambda)\, P_{i-1}'(\lambda) - \beta_i^2\, P_{i-2}'(\lambda)$$

$$P_0'(\lambda) = 0\,, \quad P_1'(\lambda) = -1$$

A Newton - módszer alkalmazásához jó kezdőértékre van szükség. Ehhez általában választhatók a következő tridiagonális mátrix sajátértékei:

$$\overline{T} := \begin{bmatrix} \overline{\alpha} & \overline{\beta} & 0 & \cdots & 0 \\ \overline{\beta} & \overline{\alpha} & \ddots & & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & & & \overline{\alpha} & \overline{\beta} \\ 0 & \cdots & 0 & \overline{\beta} & \overline{\alpha} \end{bmatrix}$$

ahol $\quad \overline{\alpha} := \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \alpha_i \quad$ és $\quad \overline{\beta} := \dfrac{1}{n-1}\displaystyle\sum_{i=2}^{n} \beta_i\ .$

A $\overline{T}$ mátrix sajátértékeit pedig ismerjük, sőt a sajátvektorokat is, igaz ugyanis a következő tétel:

Tétel

Legyen D $(n \times n)$ - es tridiagonális mátrix a következő alakú:

$$D := \begin{bmatrix} a & c & 0 & \cdots & 0 \\ b & a & \ddots & & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & & & a & c \\ 0 & \cdots & 0 & b & a \end{bmatrix} \qquad \boxed{b \cdot c > 0}$$

Ekkor a D mátrix sajátértékei a következő módon számolhatók:

$$\lambda_i = a + \text{sign}\,(b) \cdot 2\sqrt{b \cdot c} \cdot \cos\left(\frac{i\pi}{n+1}\right)\,, \qquad i = 1, \ldots, n$$

és a megfelelő $\underline{\mathbf{x}}^i$ sajátvektor j - edik koordinátája:

$$x_j^i = \left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi j}{n+1}\right)\,, \qquad i\,,\, j = 1, \ldots, n$$

$\boxed{\textbf{Bizonyítás}}$

Belátjuk, hogy $D\underline{\mathbf{x}}^i = \lambda_i \underline{\mathbf{x}}^i$.

A j **- edik koordinátára felírva az egyenlőséget :**

$$Dx_j^i = b\left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi\,(j-1)}{n+1}\right) + a\left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi j}{n+1}\right) + c\left(\frac{b}{c}\right)^{\frac{j}{2}} \sin\left(\frac{i\pi\,(j+1)}{n+1}\right) =$$

$$= a\left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi j}{n+1}\right) + c\left(\frac{b}{c}\right)^{\frac{j}{2}} \cdot \left[\sin\left(\frac{i\pi\,(j-1)}{n+1}\right) + \sin\left(\frac{i\pi\,(j+1)}{n+1}\right)\right] =$$

$$= a\left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi j}{n+1}\right) + c\left(\frac{b}{c}\right)^{\frac{j}{2}} \cdot 2\sin\left(\frac{i\pi j}{n+1}\right)\cos\left(\frac{i\pi}{n+1}\right) =$$

$$= a\left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi j}{n+1}\right) + c\left(\frac{b}{c}\right)^{\frac{1}{2}} \left(\frac{b}{c}\right)^{\frac{j-1}{2}} \cdot 2\sin\left(\frac{i\pi j}{n+1}\right)\cos\left(\frac{i\pi}{n+1}\right) =$$

$$= a\left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi j}{n+1}\right) + \operatorname{sign}\,(b)\sqrt{b\cdot c} \cdot 2\sin\left(\frac{i\pi j}{n+1}\right)\cos\left(\frac{i\pi}{n+1}\right)\left(\frac{b}{c}\right)^{\frac{j-1}{2}} =$$

$$= \underbrace{\left[a + \operatorname{sign}\,(b)\cdot 2\cdot\sqrt{b\cdot c}\cdot\cos\left(\frac{i\pi}{n+1}\right)\right]}_{\lambda_i} \cdot \underbrace{\left(\frac{b}{c}\right)^{\frac{j-1}{2}} \sin\left(\frac{i\pi j}{n+1}\right)}_{x_j^i} = \lambda_i x_j^i$$

$\blacksquare$

Megjegyzés :

Gyakori eset differenciálegyenletek numerikus megoldásánál, hogy

$T = \text{tridiag}\,(-1\,,\,2\,,\,-1)$,   tehát $b = c = -1$ $\operatorname{sign}\,(b) = -1$

és ekkor a mátrix sajátértékei:

$$\lambda_i = 2 - 2\cos\left(\frac{i\pi}{n+1}\right) \qquad i = 1, \ldots, n$$

és a megfelelő $\underline{\mathbf{x}}^i$ sajátvektor j - edik koordinátája:

$$x_j^i = \sin\left(\frac{i\pi j}{n+1}\right)\,, \quad i\,,\, j = 1, \ldots, n$$

8.5. RANGSZÁMCSÖKKENTÉS

Egy sajátérték és sajátvektor ismeretében egy $(n \times n)$ - es mátrix többi sajátértékének és sajátvektorának meghatározása $(n-1)$ - edrendű (és így tovább) mátrix sajátértékeinek és sajátvektorainak a meghatározására vezethető vissza.

Legyen A $(n \times n)$ - es mátrix és $\lambda_1, \ldots, \lambda_n$ a sajátértékei és a nekik megfelelő sajátvektorok $\underline{\mathbf{x}}_1, \ldots, \underline{\mathbf{x}}_n$.

Tudjuk, hogy hasonlósági transzformáció nem változtatja meg a sajátértékeket.

Legyen:

(1) $\quad A\underline{\mathbf{x}}_1 = \lambda\underline{\mathbf{x}}_1 \quad \Longrightarrow \quad HAH^{-1}(H\underline{\mathbf{x}}_1) = \lambda_1 (H\underline{\mathbf{x}}_1)$

H reguláris $\quad \Longrightarrow \quad H\underline{\mathbf{x}}_1 \neq 0$

Ha a H transzformáció olyan, hogy:

$H\underline{\mathbf{x}}_1 = \sigma\underline{\mathbf{e}}_1$, (például H Householder - mátrix esetén)

akkor ezt *(1)* - be írva $\Longrightarrow$

$$HAH^{-1}(\sigma\underline{\mathbf{e}}_1) = \lambda_1(\sigma\underline{\mathbf{e}}_1)$$

$$HAH^{-1}\underline{\mathbf{e}}_1 = \lambda_1\underline{\mathbf{e}}_1$$

Így HAH^{-1} a következő alakban írható:

(2) $\quad \begin{bmatrix} \lambda_1 & \underline{\mathbf{b}}^T \\ 0 & B \end{bmatrix}$ ahol B $(n-1) \times (n-1)$ - es mátrix.

Most legyen $A\underline{\mathbf{x}}_2 = \lambda_2\underline{\mathbf{x}}_2$. Ekkor:

(3) $\quad HAH^{-1}(H\underline{\mathbf{x}}_2) = \lambda_2 (H\underline{\mathbf{x}}_2) \quad \Longrightarrow \quad \lambda_2$ a HAH^{-1} - nek is sajátértéke.

Írjuk a $H\underline{\mathbf{x}}_2$ - t a következő alakban:

$H\underline{\mathbf{x}}_2 := \begin{bmatrix} \alpha \\ \underline{\mathbf{x}}_2' \end{bmatrix} = \begin{bmatrix} \alpha & \underline{\mathbf{x}}_2' \end{bmatrix}^T \qquad \underline{\mathbf{x}}_2'$ $(n-1)$ dimenziós.

Így $\stackrel{(2)\ (3)}{\Longrightarrow}$ $\begin{bmatrix} \lambda_1 & \underline{\mathbf{b}}^T \\ & \\ 0 & B \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \\ \underline{\mathbf{x}}_2^{'} \end{bmatrix} = \lambda_2 \begin{bmatrix} \alpha \\ \\ \underline{\mathbf{x}}_2^{'} \end{bmatrix}$ $\Longrightarrow$

$$\alpha\lambda_1 + \underline{\mathbf{b}}^T \underline{\mathbf{x}}_2^{'} = \lambda_2\alpha$$

$$\alpha\left(\lambda_1 - \lambda_2\right) = -\underline{\mathbf{b}}^T \underline{\mathbf{x}}_2^{'}$$

$$B\underline{\mathbf{x}}_2^{'} = \lambda_2\underline{\mathbf{x}}_2^{'}$$

Innen $\Longrightarrow$ A B λ_2 sajátértéke egyúttal A - nak is sajátértéke. Tehát A második λ_2 sajátértékének meghatározásához elég B sajátértékét keresni.