Numerikus módszerek 2.

3. előadás: Hatványmódszer és inverz iteráció

Krebsz Anna

ELTE IK

Tartalomjegyzék

A Rayleigh-hányados tulajdonságai

Definíció: Rayleigh-hányados

Az $\frac{\langle Ax,x\rangle}{\langle x,x\rangle}$ hányadost $(x \neq 0)$ *Rayleigh-hányadosnak* nevezzük.

Tétel a Rayleigh-hányadosról

Ha $A = A^*$, akkor

1. $$\max_{x\neq 0} \frac{\langle Ax, x\rangle}{\langle x, x\rangle} = \lambda_{max} \quad \text{illetve} \quad \min_{x\neq 0} \frac{\langle Ax, x\rangle}{\langle x, x\rangle} = \lambda_{min}$$
2. Rögzített $x \neq 0$ vektor esetén
$$\min_{\lambda \in \mathbb{K}} \|Ax - \lambda x\|_2$$
megoldása $\lambda = \frac{\langle Ax,x\rangle}{\langle x,x\rangle}$.

Biz.: Mivel $A = A^*$, ezért $(x^*Ax)^* = x^*A^*x = x^*Ax$, tehát $\langle Ax, x\rangle = x^*Ax \in \mathbb{R}$.

1. Létezik U unitér mátrix, melyre $U^*AU = D$ diagonális mátrix $\Rightarrow$ $A = UDU^*$. $x \neq 0$ esetén

$$\frac{\langle Ax, x\rangle}{\langle x, x\rangle} = \frac{\langle UDU^*x, x\rangle}{\langle x, x\rangle} = \frac{\langle D\overbrace{U^*x}^{y}, \overbrace{U^*x}^{y}\rangle}{\langle x, x\rangle}.$$

Bevezetve az $y := U^*x$ jelölést az unitér traszformáció miatt $\langle y, y\rangle = \langle x, x\rangle$, így

$$\frac{\langle Dy, y\rangle}{\langle x, x\rangle} = \frac{\langle Dy, y\rangle}{\langle y, y\rangle} = \frac{\sum_{i=1}^n \lambda_i |y_i|^2}{\sum_{i=1}^n |y_i|^2}.$$

Becsüljük alulról és felülről

$$\lambda_{min} \leq \frac{\sum_{i=1}^n \lambda_i |y_i|^2}{\sum_{i=1}^n |y_i|^2} \leq \lambda_{max}.$$

Biz. folyt.: Belátjuk, hogy ezek a korlátok minimum illetve maximum értékek.

Legyen $x := v_{min} \neq 0$ a λ_{min}-hez tartozó sajátvektor, ekkor

$$\frac{\langle Av_{min}, v_{min}\rangle}{\langle v_{min}, v_{min}\rangle} = \frac{\langle \lambda_{min} v_{min}, v_{min}\rangle}{\langle v_{min}, v_{min}\rangle} = \lambda_{min}.$$

Hasonlóan $x := v_{max} \neq 0$ a λ_{max}-hoz tartozó sajátvektor, ekkor

$$\frac{\langle Av_{max}, v_{max}\rangle}{\langle v_{max}, v_{max}\rangle} = \frac{\langle \lambda_{max} v_{max}, v_{max}\rangle}{\langle v_{max}, v_{max}\rangle} = \lambda_{max}.$$

Biz. folyt.: 2. $Ax - \lambda x = r(x)$ az adott $x \neq 0$ vektorhoz és λ-hoz tartozó reziduális hiba.

$$\begin{aligned}\|Ax - \lambda x\|_2^2 &= \langle Ax - \lambda x, Ax - \lambda x\rangle = \langle Ax, Ax\rangle - 2\lambda\langle Ax, x\rangle + \lambda^2\langle x, x\rangle \\ &= \langle x, x\rangle \cdot \left[\left(\lambda - \frac{\langle Ax, x\rangle}{\langle x, x\rangle}\right)^2 - \left(\frac{\langle Ax, x\rangle}{\langle x, x\rangle}\right)^2\right] + \langle Ax, Ax\rangle\end{aligned}$$

λ-ra nézve egy másodfokú polinomot kaptunk, a teljes négyzetté alakításból látszik, hogy a kifejezés értéke a

$$\lambda = \frac{\langle Ax, x\rangle}{\langle x, x\rangle}$$

helyen lesz minimális. Tehát önadjungált mátrix esetén rögzített $x \neq 0$ vektor esetén a Rayleigh-hányados választásával lesz a reziduális hiba a legkisebb. $\square$

Az abszolút értékben legnagyobb (domináns) sajátértéket és sajátvektorát közelíti a következő iteráció:

$$x^{(0)} \neq 0 \ \ (\in \mathbb{R}^n) \quad \text{kezdővektor}$$
$$x^{(k+1)} := Ax^{(k)} \quad (k = 0, 1, \ldots).$$

Célszerű néhány lépés után vagy lépésenként normálni a vektorokat, hogy az alul- illetve túlcsordulást elkerüljük.

Tétel a hatványmódszerről

1. Legyen A normális, vagyis létezzen sajátvektorokból álló ortonormált bázisa: $(v_1, \ldots, v_n)$.
2. A sajátértékeire $|\lambda_n| > |\lambda_{n-1}| \geq \ldots \geq |\lambda_1|$.
3. Az $x^{(0)} = \sum_{i=1}^n c_i v_i$ kezdővektorra $c_n = \langle x^{(0)}, v_n \rangle \neq 0$. ($c_n$ az $x^{(0)}$-nak a v_n irányú komponense.)

Ekkor

1. $\lim_{k\to\infty} \frac{1}{\lambda_n^k} x^{(k)} = c_n v_n$.
2. $\lim_{k\to\infty} \frac{x_i^{(k+1)}}{x_i^{(k)}} = \lambda_n,$ ahol $|x_i^{(k)}| = \|x^{(k)}\|_\infty$.
3. Ha $A = A^*$, akkor $\lim_{k\to\infty} \frac{\left\langle x^{(k+1)}, x^{(k)} \right\rangle}{\left\langle x^{(k)}, x^{(k)} \right\rangle} = \lambda_n$.

Biz.:
A normális mátrix $\exists$ Q ortogonális mátrix, melyre
$Q^T AQ = D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.. Q oszlopai a sajátvektorok
$(Q = (v_1, \dots, v_n))$, továbbá

$$\langle x^{(0)}, v_n \rangle = \left\langle \sum_{j=1}^{n} c_j v_j, v_n \right\rangle = c_n \neq 0.$$

1. Fejtsük végig a rekurziót

$$x^{(k)} = Ax^{(k-1)} = \dots = A^k x^{(0)} = A^k \sum_{j=1}^{n} c_j v_j = \sum_{j=1}^{n} c_j A^k v_j =$$

$$= \sum_{j=1}^{n} c_j \lambda_j^k v_j = \lambda_n^k c_n v_n + \sum_{j=1}^{n-1} c_j \lambda_j^k v_j =$$

$$= \lambda_n^k \cdot \left(c_n v_n + \sum_{j=1}^{n-1} c_j \left(\frac{\lambda_j}{\lambda_n} \right)^k v_j \right)$$

Mivel $\left|\frac{\lambda_j}{\lambda_n}\right| < 1 \quad (j = 1, \ldots, n-1)$, ezért

$$\lim_{k\to\infty} \frac{1}{\lambda_n^k} x^{(k)} = \lim_{k\to\infty} \left(c_n v_n + \sum_{j=1}^{n-1} c_j \left(\frac{\lambda_j}{\lambda_n}\right)^k v_j \right) = c_n v_n.$$

A kapott $c_n v_n$ vektor λ_n-hez tartozó sajátvektor.

Ha $c_n = 0$ lenne, akkor nullvektort kapnánk, ami nem sajátvektor.

Biz. folyt.:

2. Az $x^{(k)}$ vektor i. komponense

$$x_i^{(k)} = \lambda_n^k \cdot \left(c_n v_n^{(i)} + \sum_{j=1}^{n-1} c_j \left(\frac{\lambda_j}{\lambda_n} \right)^k v_j^{(i)} \right).$$

Tegyük fel, hogy $|x_i^{(k)}| = \|x^{(k)}\|_\infty$, ezzel a megvalósítás során az osztás stabil lesz.

$$\frac{x_i^{(k+1)}}{x_i^{(k)}} = \lambda_n \cdot \frac{c_n v_n^{(i)} + \sum_{j=1}^{n-1} c_j \left(\frac{\lambda_j}{\lambda_n} \right)^{k+1} v_j^{(i)}}{c_n v_n^{(i)} + \sum_{j=1}^{n-1} c_j \left(\frac{\lambda_j}{\lambda_n} \right)^{k} v_j^{(i)}} =$$

$$= \lambda_n \cdot \frac{1 + \sum_{j=1}^{n-1} \frac{c_j}{c_n} \cdot \left(\frac{\lambda_j}{\lambda_n} \right)^{k+1} \frac{v_j^{(i)}}{v_n^{(i)}}}{1 + \sum_{j=1}^{n-1} \frac{c_j}{c_n} \cdot \left(\frac{\lambda_j}{\lambda_n} \right)^{k} \frac{v_j^{(i)}}{v_n^{(i)}}} \to \lambda_n \ \ (k \to \infty),$$

mivel $\left| \frac{\lambda_j}{\lambda_n} \right| < 1$.

Biz. folyt.: Belátható (lásd a nagy bizonyítások fájlja), hogy

$$\left|\frac{x_i^{(k+1)}}{x_i^{(k)}} - \lambda_n\right| = O\left(\left|\frac{\lambda_{n-1}}{\lambda_n}\right|^k\right).$$

Tetszőleges $y \in \mathbb{R}^n$ vektorra, melyre $\langle y, x^{(k)}\rangle \neq 0$ hasonlóan bizonyítható, hogy

$$\lim_{k\to\infty} \frac{\langle x^{(k+1)}, y\rangle}{\langle x^{(k)}, y\rangle} = \lambda_n.$$

Ebből az $y := e_i$ vektorral megkapjuk a tétel állítását. A tételbeli i index megválasztása garantálja a megvalósítás során az osztások stabilitását.

Biz. folyt.:
3. Felhasználjuk, hogy $A = A^T$ esetben a sajátvektorok ortonormált rendszert alkotnak, vagyis $\langle v_j, v_l \rangle = \delta_{jl}$.

$$\langle x^{(k)}, x^{(k)} \rangle = \left\langle \sum_{j=1}^{n} c_j \lambda_j^k v_j, \sum_{l=1}^{n} c_l \lambda_l^k v_l \right\rangle = \sum_{j=1}^{n} \sum_{l=1}^{n} c_j c_l \lambda_j^k \lambda_l^k \underbrace{\langle v_j, v_l \rangle}_{=\delta_{jl}} =$$

$$= \sum_{j=1}^{n} c_j^2 \lambda_j^{2k} = c_n^2 \lambda_n^{2k} + \sum_{j=1}^{n-1} c_j^2 \lambda_j^{2k} =$$

$$= c_n^2 \lambda_n^{2k} \cdot \left(1 + \sum_{j=1}^{n-1} \left(\frac{c_j}{c_n} \right)^2 \left(\frac{\lambda_j}{\lambda_n} \right)^{2k} \right)$$

Biz. folyt.:

$$\begin{aligned}
\langle x^{(k+1)}, x^{(k)} \rangle &= \left\langle \sum_{j=1}^{n} c_j \lambda_j^{k+1} v_j, \sum_{l=1}^{n} c_l \lambda_l^{k} v_l \right\rangle = \\
&= \sum_{j=1}^{n} \sum_{l=1}^{n} c_j c_l \lambda_j^{k+1} \lambda_l^{k} \underbrace{\langle v_j, v_l \rangle}_{=\delta_{jl}} = \\
&= \sum_{j=1}^{n} c_j^2 \lambda_j^{2k+1} = c_n^2 \lambda_n^{2k+1} + \sum_{j=1}^{n-1} c_j^2 \lambda_j^{2k+1} = \\
&= c_n^2 \lambda_n^{2k+1} \cdot \left(1 + \sum_{j=1}^{n-1} \left(\frac{c_j}{c_n} \right)^2 \left(\frac{\lambda_j}{\lambda_n} \right)^{2k+1} \right)
\end{aligned}$$

Biz. folyt.:
A Rayleigh-hányadost felírva

$$\frac{\langle x^{(k+1)}, x^{(k)}\rangle}{\langle x^{(k)}, x^{(k)}\rangle} = \lambda_n \cdot \frac{1 + \sum_{j=1}^{n-1} \left(\frac{c_j}{c_n}\right)^2 \left(\frac{\lambda_j}{\lambda_n}\right)^{2k+1}}{1 + \sum_{j=1}^{n-1} \left(\frac{c_j}{c_n}\right)^2 \left(\frac{\lambda_j}{\lambda_n}\right)^{2k}} \ \to \lambda_n \ (k \to \infty),$$

mivel $\left|\frac{\lambda_j}{\lambda_n}\right| < 1$.

Belátható (lásd a nagy bizonyítások fájlja), hogy

$$\left|\frac{\langle x^{(k+1)}, x^{(k)}\rangle}{\langle x^{(k)}, x^{(k)}\rangle} - \lambda_n\right| = O\left(\left|\frac{\lambda_{n-1}}{\lambda_n}\right|^{2k}\right).$$

$\square$

Megjegyzések:

- A tétel nem csak normális mátrixra igaz, de így egyszerűbb a bizonyítás. Lásd Stoyan Gisbert: Numerikus módszerek 1.
- A $c_n \neq 0$ feltétel gépi számolások során nem fog öröklődni, ezért lassan, de konvergálni fog a módszer.
- A bizonyításból látszik, hogy a módszer konvergenciája a $\frac{|\lambda_{n-1}|}{|\lambda_n|}$ hányadostól függ. Szimmetrikus esetben a $\left(\frac{|\lambda_{n-1}|}{|\lambda_n|}\right)^2$ hányadostól függ.
- Az $\frac{x_i^{(k+1)}}{x_i^{(k)}}$ sajátérték közelítésben i értéke véges sok lépésben beáll egy állandó értékre.

Megjegyzések

- Az LER-k iterációs módszereinél az $x^{(k+1)} := Bx^{(k)} + c$ iterációk során használtuk az eltérés vektort (a továbblépés irányát):

$$s^{(k)} := x^{(k+1)} - x^{(k)} = (Bx^{(k)} + c) - (Bx^{(k-1)} + c) = \\ = B(x^{(k)} - x^{(k-1)}) = Bs^{(k-1)}.$$

  Látjuk, hogy az iteráció során az eltérés vektorra a B mátrix-szal hatványmódszert végzünk normálás nélkül.
- A nem invertálható $\Leftrightarrow$ 0 sajátérték.
- A^{-1} sajátértékei: $\lambda_i(A^{-1}) = \frac{1}{\lambda_i(A)}$ $(i = 1, \ldots, n)$.

1. Példa: Hatványmódszer alkalmazása

Alkalmazzuk $x_0 = e_1$-ből indulva a hatványmódszer 3 lépését az alábbi mátrix abszolútértékben legnagyobb sajátértékének és hozzá tartozó sajátvektorának közelítésére.

$$A = \begin{bmatrix} 1 & 4 \\ -1 & 6 \end{bmatrix}$$

Készítsük el a sajátvektort közelítő x_k vektorokat. Keretezve jelöljük a vektor megfelelő elemét, melynek pozíciója szükséges a hányadosképzéshez, a sajátérték közelítéséhez.

$$x_1 = Ax_0 = \begin{bmatrix} 1 & 4 \\ -1 & 6 \end{bmatrix} \cdot \begin{bmatrix} \boxed{1} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \lambda_1^{(1)} = \frac{x_1^{(1)}}{x_0^{(1)}} = \frac{1}{1} = 1$$

$$x_2 = Ax_1 = \begin{bmatrix} 1 & 4 \\ -1 & 6 \end{bmatrix} \cdot \begin{bmatrix} \boxed{1} \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ -7 \end{bmatrix}, \quad \lambda_1^{(2)} = \frac{x_2^{(2)}}{x_1^{(2)}} = \frac{-3}{1} = -3$$

$$x_3 = Ax_2 = \begin{bmatrix} 1 & 4 \\ -1 & 6 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ \boxed{-7} \end{bmatrix} = \begin{bmatrix} -31 \\ -39 \end{bmatrix}, \quad \lambda_1^{(3)} = \frac{x_3^{(1)}}{x_2^{(1)}} = \frac{-39}{-7} = \frac{39}{7}$$

A 3. lépés után a legnagyobb abszolútértékű sajátérték közelítése

$$\lambda_1^{(3)} = \frac{39}{7} = 5\frac{4}{7}.$$

2. Példa: Hatványmódszer alkalmazása szimmetrikus mátrix esetén

Alkalmazzuk $x_0 = e_1$-ből indulva a hatványmódszer 3 lépését az alábbi mátrix abszolútértékben legnagyobb sajátértékének és hozzá tartozó sajátvektorának közelítésére.

$$A = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix}$$

Készítsük el a sajátvektort közelítő x_k vektorokat.

$$x_1 = Ax_0 = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} \boxed{1} \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$x_2 = Ax_1 = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} \boxed{4} \\ -1 \end{bmatrix} = \begin{bmatrix} 17 \\ -8 \end{bmatrix}$$

$$x_3 = Ax_2 = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} \boxed{17} \\ -8 \end{bmatrix} = \begin{bmatrix} 76 \\ -49 \end{bmatrix}$$

Példa hatványmódszerre

A maximális abszolútértékű sajátérték közelítésére a Rayleigh-hányadost számoljuk.

$$\lambda_1^{(1)} = \frac{\langle x_1, x_0 \rangle}{\langle x_0, x_0 \rangle} = \frac{4}{1} = 4$$

$$\lambda_1^{(2)} = \frac{\langle x_2, x_1 \rangle}{\langle x_1, x_1 \rangle} = \frac{76}{17} = 4,47$$

$$\lambda_1^{(3)} = \frac{\langle x_3, x_2 \rangle}{\langle x_2, x_2 \rangle} = \frac{1684}{353} = 4,77$$

A 3. lépés után a legnagyobb abszolútértékű sajátérték közelítése

$$\lambda_1^{(3)} = \frac{1684}{353} = 4,77.$$

Példa hatványmódszerre

Összehasonlításul, ha a nem szimmetrikus esetben alkalmazott hányadosokat használtuk volna (lásd korábbi bekeretezett elemek), akkor

$$\lambda_1^{(1)} = \frac{x_1^{(1)}}{x_0^{(1)}} = \frac{4}{1} = 4$$

$$\lambda_1^{(2)} = \frac{x_2^{(1)}}{x_1^{(1)}} = \frac{17}{4} = 4,25$$

$$\lambda_1^{(3)} = \frac{x_3^{(1)}}{x_2^{(1)}} = \frac{76}{17} = 4,47$$

Megjegyzés:

Javaslom mindkét $\lambda_1^{(3)}$ sajátérték és a hozzátartozó x_3 sajátvektor közelítésre kiszámolni a reziduum vektorból felírható becslést.

Az A^{-1}-re alkalmazott hatványmódszer, csak invertálható mátrixra alkalmazható. Az abszolút értékben legkisebb sajátértéket és sajátvektorát közelíti a következő iteráció:

$$x^{(0)} \neq 0 \;\; (\in \mathbb{R}^n) \quad \text{kezdővektor}$$

$$x^{(k+1)} := A^{-1}x^{(k)} \quad \text{helyett az} \quad Ax^{(k+1)} = x^{(k)} \quad (k = 0, 1, \ldots).$$

LER-t oldjuk meg pl. az A mátrix LU-felbontásának felhasználásával. Ezzel a lépésenkénti műveletigény $\frac{2}{3}n^3 + \mathcal{O}(n^2)$-ről $n^2 + \mathcal{O}(n)$-re csökkenthető. Célszerű itt is néhány lépés után vagy lépésenként normálni a vektorokat, hogy az alul- illetve túlcsordulást elkerüljük.

Tétel az inverz iterációról

1. Legyen A normális, vagyis létezzen sajátvektorokból álló ortonormált bázisa: $(v_1, \ldots, v_n)$.
2. A sajátértékeire $|\lambda_1| < |\lambda_2| \leq \ldots \leq |\lambda_n|$.
3. Az $x^{(0)} = \sum_{i=1}^{n} c_i v_i$ kezdővektorra $c_1 = \langle x^{(0)}, v_1 \rangle \neq 0$. ($c_1$ az $x^{(0)}$-nak a v_1 irányú komponense.)

Ekkor

1. $\lim_{k\to\infty} \lambda_1^k x^{(k)} = c_1 v_1$.
2. $\lim_{k\to\infty} \frac{x_i^{(k+1)}}{x_i^{(k)}} = \frac{1}{\lambda_1}$, ahol $|x_i^{(k)}| = \|x^{(k)}\|_\infty$.
3. Ha $A = A^*$, akkor $\lim_{k\to\infty} \frac{\left\langle x^{(k+1)}, x^{(k)} \right\rangle}{\left\langle x^{(k)}, x^{(k)} \right\rangle} = \frac{1}{\lambda_1}$.

Biz.: A hatványmódszer tételéből következik.

Példa: Az inverz iteráció alkalmazása

Alkalmazzuk $x_0 = e_1$-ből indulva az inverz iteráció 3 lépését az alábbi mátrix legkisebb abszolútértékű sajátértékének és hozzá tartozó sajátvektorának közelítésére.

$$A = \begin{bmatrix} 1 & 4 \\ -1 & 6 \end{bmatrix}$$

Megoldás:
Az A^{-1} mátrixra kell alkalmaznunk a hatványmódszer 3 lépését.

A mátrix inverze:

$$A^{-1} = \frac{1}{10} \cdot \begin{bmatrix} 6 & -4 \\ 1 & 1 \end{bmatrix}.$$

Készítsük el a sajátvektort közelítő x_k vektorokat.

$$x_1 = Ax_0 = \frac{1}{10} \cdot \begin{bmatrix} 6 & -4 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{10} \cdot \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

$$x_2 = Ax_1 = \frac{1}{10} \cdot \begin{bmatrix} 6 & -4 \\ 1 & 1 \end{bmatrix} \cdot \frac{1}{10} \cdot \begin{bmatrix} 6 \\ 1 \end{bmatrix} = \frac{1}{100} \cdot \begin{bmatrix} 32 \\ 7 \end{bmatrix}$$

$$x_3 = Ax_2 = \frac{1}{10} \cdot \begin{bmatrix} 6 & -4 \\ 1 & 1 \end{bmatrix} \cdot \frac{1}{100} \cdot \begin{bmatrix} 32 \\ 7 \end{bmatrix} = \frac{1}{1000} \cdot \begin{bmatrix} 164 \\ 39 \end{bmatrix}$$

Keretezve jelöljük a vektor megfelelő elemét, melynek pozíciója szükséges a sajátérték közelítéséhez, a megfelelő hányadosképzéshez.

$$x_0 = \begin{bmatrix} \boxed{1} \\ 0 \end{bmatrix}, x_1 = \frac{1}{10} \cdot \begin{bmatrix} 6 \\ 1 \end{bmatrix} \to \frac{1}{\lambda_1^{(1)}} = \frac{x_1^{(1)}}{x_0^{(1)}} = \frac{6}{10} = \frac{3}{5}$$
$$x_1 = \frac{1}{10} \begin{bmatrix} \boxed{6} \\ 1 \end{bmatrix}, x_2 = \frac{1}{100} \begin{bmatrix} 32 \\ 7 \end{bmatrix} \to \frac{1}{\lambda_1^{(2)}} = \frac{x_2^{(1)}}{x_1^{(1)}} = \frac{32}{60} = \frac{8}{15}$$
$$x_2 = \frac{1}{100} \begin{bmatrix} \boxed{32} \\ 7 \end{bmatrix}, x_3 = \frac{1}{1000} \begin{bmatrix} 164 \\ 39 \end{bmatrix} \to \frac{1}{\lambda_1^{(3)}} = \frac{x_3^{(1)}}{x_2^{(1)}} = \frac{164}{320} = \frac{41}{80}$$

A 3. lépésben a legkisebb abszolútértékű sajátérték közelítése

$$\lambda_1^{(3)} = \frac{80}{41}.$$

Tartalomjegyzék

Állítás

Az A és $A_p := A - p \cdot I$ ($p \in \mathbb{R}$) sajátvektorai azonosak a sajátértékekre:

$$\lambda_i(A_p) = \lambda_i - p.$$

Biz.: $A_p v_i = (A - p \cdot I)v_i = Av_i - pv_i = (\lambda_i - p)v_i$.

Megj.:

- Az eltolással a hatványmódszer szélső sajátértékek meghatározására alkalmas illetve a konvergencia gyorsítható.
- Ha A-ra alkalmazunk eltolást, majd inverz iterációt, akkor az $\frac{1}{\lambda_i - p}$ domináns sajátértékét közelítjük. Ha p-t úgy adjuk meg, hogy a λ_i (egyszeres) sajátérték legyen a legközelebb hozzá, akkor ehhez konvergál a módszer. Vagyis inverz iterációval közbülső sajátértékeket is közelíthetünk. Lásd gyakorlat.

Példa shiftelésre

Példa: Shiftelés alkalmazása

Tegyük fel, hogy az A mátrix sajátértékei a következők:

$$\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \lambda_4 = 4, \lambda_5 = 5.$$

Mennyi lesz a hatványmódszer gyorsaságát jellemző hányados? A-nak melyik sajártértékét tudjuk közelíteni az $A - pI$-re felírt hatványmódszerrel?

1. Ha nem alkalmazunk shiftet.
2. Ha $p = 2$ shiftet alkalmazunk.
3. Ha $p = 4$ shiftet alkalmazunk.
4. Ha $p = -95$ shiftet alkalmazunk.

Megoldás:

1. Ha nem alkalmazunk shiftet, akkor a hányados $\frac{4}{5} = 0.8$, az A mátrix 5 sé-t közelítjük vele.
2. $p = 2$ esetén az $A - 2 \cdot I$ mátrix sajátértékei: -1,0,1,2,3. Ekkor a hányados $\frac{2}{|3|} \approx 0.66.$, az A mátrix 5 sé-t közelítjük vele.
3. $p = 4$ esetén az $A - 4 \cdot I$ mátrix sajátértékei: -3,-2,-1,0,1. Ekkor a hányados $\frac{2}{|-3|} \approx 0.66$, az A mátrix 1 sé-t közelítjük vele.
4. $p = -95$ esetén az $A + 95 \cdot I$ mátrix sajátértékei: 96,97,98,99,100. Ekkor a hányados $\frac{99}{100} = 0.99$, az A mátrix 5 sé-t közelítjük vele.

Tartalomjegyzék

Kérdés: Egy sajátérték illetve sajátvektor ismeretében hogyan határozható meg a többi?

Tétel

Tegyük fel, hogy λ_n, v_n sajátpár ismert, $\|v_n\|_2 = 1$. Vegyük azt a $H := H(v)$ Householder-transzformációt, melyre $v := \frac{v_n - e_1}{\|v_n - e_1\|_2}$. Ekkor

$$HAH = \lambda_n e_1, \quad \text{azaz} \quad HAH = \begin{bmatrix} \lambda_n & c^T \\ 0 & B \end{bmatrix}$$

és B sajátértékei: $\lambda_1, \ldots, \lambda_{n-1}$.

Biz.: $He_1 = v_n$ és $Hv_n = e_1$, innen

$$HAHe_1 = HAv_n = \lambda_n Hv_n = \lambda_n e_1.$$

$$\det(A - \lambda I) = (\lambda_n - \lambda) \cdot \det(B - \lambda I)$$

Vagyis B sajátértékei: $\lambda_1, \ldots, \lambda_{n-1}$.

Kérdés:
Ha meghatároztuk B sajátértékeit és sajátvektorait, akkor hogyan kapjuk meg A sajátvektorait? Ezzel a technikával csak néhány sajátértéket érdemes meghatározni.

Tétel

Ha a B sajátvektorai $u_1, \dots, u_{n-1}$, akkor az A sajátvektorai $\lambda_i \neq \lambda_n$ esetén a következők

$$v_i = H \cdot \begin{bmatrix} b_i \\ u_i \end{bmatrix}, \quad \text{ahol} \quad b_i = \frac{c^T u_i}{\lambda_i - \lambda_n} \;\; (i = 1, \dots, n-1).$$

Biz.: Minden vektort kiterjesztünk egy $b_i \in \mathbb{R}$ értékkel, ezeket kell megadnunk.

$$HAH \cdot \begin{bmatrix} b_i \\ u_i \end{bmatrix} = \begin{bmatrix} \lambda_n & c^T \\ 0 & B \end{bmatrix} \cdot \begin{bmatrix} b_i \\ u_i \end{bmatrix} = \lambda_i \cdot \begin{bmatrix} b_i \\ u_i \end{bmatrix}$$

$$\begin{aligned} \lambda_n b_i + c^T u_i &= \lambda_i b_i \\ c^T u_i &= (\lambda_i - \lambda_n) b_i \\ b_i &= \frac{c^T u_i}{\lambda_i - \lambda_n} \;\; (i = 1, \ldots, n-1). \end{aligned}$$

Ezután a hasonlósági transzformációból $v_i = H \cdot \begin{bmatrix} b_i \\ u_i \end{bmatrix}$ az A sajátvektora $(i = 1, \ldots, n-1)$-re.

- Ha $\exists i: \; \lambda_i = \lambda_n$ és $c^T u_i = 0$, akkor $b_i \in \mathbb{R}$ tetszőleges lehet.
- Ha $\exists i: \; \lambda_i = \lambda_n$ és $c^T u_i \neq 0$, akkor nem állítható elő új lineárisan független sajátvektor.

Tétel Hotelling-féle rangszámcsökkentés

Legyen A szimmetrikus és tegyük fel, hogy λ_n, v_n sajátpár ismert, $\|v_n\|_2 = 1$. Ekkor a

$$B := A - \lambda_n \cdot v_n v_n^T$$

transzformációt elvégezve a sajátvektorok nem változnak.
B sajátértékei: $\lambda_1, \ldots, \lambda_{n-1}, 0$.

Biz.:

- $Bv_n = Av_n - \lambda_n v_n \underbrace{(v_n^T v_n)}_{=1} = \lambda_n v_n - \lambda_n v_n = 0 = 0 \cdot v_n$
- A szimmetrikus, ezért létezik SONB: $(v_1, \ldots, v_n)$.
  $Bv_i = Av_i - \lambda_n v_n \underbrace{(v_n^T v_i)}_{=0} = Av_i = \lambda_i v_i$, $(i = 1, \ldots, n-1)$

Köszönöm a figyelmet!