ЛЕКЦІЯ 10. НАБЛИЖЕНЕ ЗНАХОДЖЕННЯ ВЛАСНИХ ЧИСЕЛ І ВЕКТОРІВ МАТРИЦІ

10.1. Власні числа і власні вектори матриці

1. *Власним числом* матриці $A = (a_{ij})_{n\times n}$ називають число (дійсне або комплексне) λ, таке, що рівняння

$$A\vec{x} = \lambda\vec{x} \tag{10.1}$$

має нетривіальний розв'язок $\vec{x} \neq \vec{0}$, який називають *власним вектором* матриці A, який відповідає власному числу λ. Множину всіх власних чисел матриці A називають спектром матриці A. Рівняння (10.1) можна переписати як

$$(A - \lambda E_n)\vec{x} = \vec{0},$$

де E_n — одинична матриця n-го порядку. Ця однорідна система має нетривіальний розв'язок тоді й лише тоді, коли характеристичний визначник

$$\det(A - \lambda E_n) = 0.$$

Отже, правдива

Теорема 10.1. Власні числа матриці A є розв'язками λ характеристичного рівняння

$$\det(A - \lambda E_n) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

2. Обчислюючи характеристичний визначник, ми дістаємо характеристичний многочлен n-го степеня матриці A.

Для великих або дуже великих матриць може бути складно визначати власні числа, як взагалі, складно знаходити корені характеристичних многочленів високих степенів.

Сума власних чисел (деякі з них можуть бути рівними) дорівнює сумі діагональних елементів матриці A, яку називають *слідом* матриці A, тому

$$\operatorname{tr} A = \sum_{j=1}^{n} a_{jj} = \sum_{k=1}^{n} \lambda_k.$$

Добуток власних чисел дорівнює визначнику матриці A,

$$\det A = \lambda_1\lambda_2...\lambda_n.$$

3. Характеристичний многочлен матриці A, має вигляд

$$f(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n).$$

Зберімо однакові множники характеристичного многочлена і позначмо різні власні числа як $\lambda_1, \lambda_2, ..., \lambda_r$ $(r \le n)$:

$$f(\lambda) = (-1)^n(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_r)^{m_r}.$$

Показник m_j називають *алгебричною кратністю* власного числа λ_j. Найбільша кількість власних векторів, які відповідають власному числу λ_j називають *геометричною кратністю* власного числа λ_j.

4. Матрицю B розміром $n \times n$ називають *подібною* до матриці A, якщо існує невироджена матриця T така, що

$$B = T^{-1}AT.$$

Подібність є важливою для подальшого розгляду.

Теорема 10.2 (про подібні матриці). Подібні матриці мають ті самі власні числа. Якщо $\vec{x}$ є власним вектором матриці A, то $\vec{y} = T^{-1}\vec{x}$ є власним вектором матриці B, яке відповідає тому самому власному числу.

5.

Теорема 10.3 (спектральний зсув). Якщо матриця A має власні числа $\lambda_1, ..., \lambda_n$, то матриця $A - kE_n$ для будь-якого числа k має власні числа $\lambda_1 - k, \lambda_2 - k, ..., \lambda_n - k$.

Теорема 10.4 (про матричний многочлен). Якщо λ є власним числом матриці A, то

$q(\lambda) = \alpha_s\lambda^s + \alpha_{s-1}\lambda^{s-1} + ... + \alpha_1\lambda + \alpha_0$

є власним числом матричного многочлена

$q(\lambda) = \alpha_s A^s + \alpha_{s-1}A^{s-1} + ... + \alpha_1 A + \alpha_0 E_n$.

Теорема 10.5. Власні числа ермітових матриць, зокрема дійсних симетричних матриць $(A^{\mathrm{T}} = A)$, дійсні. Власні числа косоермітових матриць, зокрема кососиметричних матриць $(A^{\mathrm{T}} = -A)$, або суто уявні або 0. Власні числа ортогональних матриць $(A^{\mathrm{T}} = A^{-1})$ рівні за модулем 1.

6. Вибір чисельного методу для знаходження власних чисел залежить істотно від 2 обставин:

1) типу матриці (дійсна симетрична, дійсна загальна, комплексна, розріджена або ні);

2) типу інформації, який треба одержати, чи треба взнати всі власні числа чи тільки певні специфічні, зокрема, найбільше власне число, чи потрібно знаходити власні вектори тощо.

10.2. Включення власних чисел матриці

1. За винятком деяких простих випадків, неможливо визначити власні числа точно за скінченну кількість операцій, тому що вони є коренями многочленів високих степенів. Отже, переважно використовують ітераційний процес.

Припускатимемо, що матриці є дійсними, а їх власні числа (для несиметричних матриць) можуть бути комплексно-спряженими.

Важливість теореми Ґершґоріна полягає в тому, що вона дає область, яка складається із замкнених кругів на комплексній площині і включає всі власні числа заданої матриці. А саме, для кожного $j = \overline{1,n}$ нерівність у формулюванні теореми визначає замкнений круг комплексної λ-площини із центром у точці a_{jj} і певним радіусом.

Теорема 10. 6 (Ґершґоріна). Нехай λ — власне число будь-якої матриці $A = (a_{ij})_{n\times n}$. Для будь-якого індекса i $(i = \overline{1,n})$ правдива нерівність

$$\left|a_{ii} - \lambda\right| \le \sum_{\substack{k=1,\\ k\ne i}}^{n} \left|a_{ik}\right|.$$

Приміром, для симетричної матриці

$$A = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 5 & 1 \\ 1/2 & 1 & 1 \end{pmatrix}$$

маємо такі круги Ґершґоріна (рис. 10.1):

$$D_1 = \left\{ z \in \mathbb{C} \middle| |z| \le 1 \right\},$$
$$D_2 = \left\{ z \in \mathbb{C} \middle| |z - 5| \le \frac{3}{2} \right\},$$
$$D_3 = \left\{ z \in \mathbb{C} \middle| |z - 1| \le \frac{3}{2} \right\}.$$

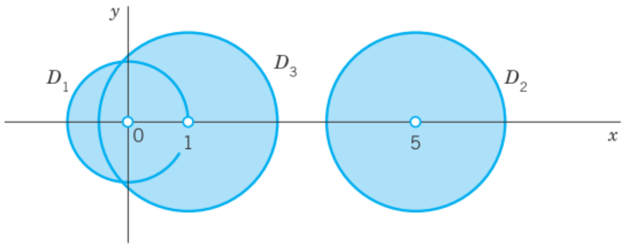


Рис. 10.1

2.

Теорема 10.7 (узагальнена Ґершґоріна). Якщо p кругів Ґершґоріна утворюють множину S таку, що відокремлена від $n - p$ інших кругів заданої матриці A, то множина S містить точно p власних чисел матриці A (кожен порахований відповідно до алгебричної кратності).

3. За означенням матриця $A = (a_{ij})_{n \times n}$ є *діагонально панівною*, якщо

$$|a_{ii}| \ge \sum_{k \ne i} |a_{ik}|.$$

Матрицю називають *строго діагонально панівною*, якщо ця нерівність строга.

Теорема 10.8. Строго діагонально панівна матриця невироджена.

Теорема 10.9 (Шура). Нехай $A = (a_{ij})_{n \times n}$. Для кожного власного числа $\lambda_1, \dots, \lambda_n$ виконано нерівність

$$|\lambda_k|^2 \le \sum_{i=1}^{n} |\lambda_i|^2 \le \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2.$$

Друга нерівність стає рівністю тоді й лише тоді, коли матриця A справджує умову (такі матриці називають нормальними)

$$A^{\mathrm{T}} A = A A^{\mathrm{T}},$$

що виконано для симетричних, кососиметричних і ортогональних матриць.

4.

Теорема 10.10 (Перрона). Нехай $A = (a_{ij})_{n\times n}$ дійсна матриця з додатними елементами. Нехай A має додатне власне значення $\lambda = \rho$ (спектральному радіусу) з алгебричною кратністю 1. Відповідний йому власний вектор можна вибрати з усіма додатними координатами.

10.3. Степеневий метод знаходження власних чисел

1. Нехай треба знайти наближені значення власних чисел матриці $A = (a_{ij})_{n\times n}$. У степеневому методі починають з вектора $\vec{x}^{(0)} \neq \vec{0}$ з n координатами і послідовно обчислюють

$$\vec{x}^{(1)} = A\vec{x}^{(0)}, \vec{x}^{(2)} = A\vec{x}^{(1)}, \ldots, \vec{x}^{(s)} = A\vec{x}^{(s-1)}.$$

Для спрощення позначмо $\vec{x}^{(s-1)} = \vec{x}, \vec{x}^{(s)} = \vec{y}$, так що

$$\vec{y} = A\vec{x}.$$

Цей метод застосовують до будь-якої квадратної матриці $A_{n\times n}$ з панівним власним числом λ, тобто існує таким власним числом, що $|\lambda|$ більше, ніж решта модулів власних чисел.

2. Нехай матриця A — симетрична матриця розміром $n \times n$ і вектор $\vec{x} \neq \vec{0}$ — будь-який дійсний вектор з n координат. Позначмо

$$\vec{y} = A\vec{x}, m_0 = \overleftarrow{x} \cdot \vec{x}, m_1 = \overleftarrow{x} \cdot \vec{y}, m_2 = \overleftarrow{y} \cdot \vec{y}.$$

Відношення Релея

$$q = \frac{m_1}{m_0}$$

є наближенням для власного числа λ матриці A.

Якщо покласти $q = \lambda - \varepsilon$, то для похибки знаходження q маємо оцінку

$$|\varepsilon| \leq \delta = \sqrt{\frac{m_2}{m_0} - q^2}.$$

Головна перевага методу є його простота. Він може бути застосований для дуже великих розріджених матриць. Основний недолік — повільна збіжність. Якщо ми шукаємо збіжну послідовність власних векторів, кожен крок треба починати з масштабування вектора, ділячи всі координати вектора на найбільший за модулем елемент.

3. Знайдімо наближене значення власного числа і власного вектора матриці

$$A=\begin{pmatrix}0,49 & 0,02 & 0,22\\ 0,02 & 0,28 & 0,20\\ 0,22 & 0,20 & 0,40\end{pmatrix}.$$

❍Вибираємо $\overleftarrow{x}^0=\begin{pmatrix}1 & 1 & 1\end{pmatrix}$. Тоді

$$\vec{x}^{(0)}=\begin{pmatrix}1\\1\\1\end{pmatrix},\vec{x}^{(1)}=\begin{pmatrix}0,890244\\0,609756\\1\end{pmatrix},\vec{x}^{2}=\begin{pmatrix}0,931193\\0,541284\\1\end{pmatrix},$$
$$\vec{x}^{(5)}=\begin{pmatrix}0,990663\\0,504682\\1\end{pmatrix},\vec{x}^{(10)}=\begin{pmatrix}0,999707\\0,500146\\1\end{pmatrix},\vec{x}^{(15)}=\begin{pmatrix}0,999991\\0,500005\\1\end{pmatrix}.$$

Тут

$$A\vec{x}^{(0)}=\begin{pmatrix}0,73\\0,5\\0,82\end{pmatrix}.$$

Після масштабування

$$\vec{x}^{(1)}=\begin{pmatrix}0,73/0,82\\0,5/0,82\\1\end{pmatrix}$$

і так далі. Панівне власне значення дорівнює $0,72$. Власний вектор

$$\begin{pmatrix}1\\0,5\\1\end{pmatrix}.$$

Відповідні значення q та δ обчислюють кожного разу до масштабування. Тому на першому кроці:

$$q=\frac{m_1}{m_0}=\frac{\overleftarrow{x}^{(0)}A\vec{x}^{(0)}}{\overleftarrow{x}^{(0)}\cdot\vec{x}^{(0)}}=\frac{2,05}{3}=0,683333;$$
$$\delta=\sqrt{\frac{m_2}{m_0}-q^2}=\sqrt{\frac{(A\vec{x}^{(0)})^{\mathrm{T}}A\vec{x}^{(0)}}{\overleftarrow{x}^{(0)}\cdot\vec{x}^{(0)}}-q^2}=\sqrt{\frac{1,4553}{3}-q^2}=0,134743.$$

Маємо таку таблицю:

j	1	2	5	10
q	0.683333	0.716048	0.719944	0.720000
δ	0.134743	0.038887	0.004499	0.000141
ϵ	0.036667	0.003952	0.000056	$5 \cdot 10^{-8}$

10.4. Тридіагоналізація та QR-факторизація

Розгляньмо задачу знайти всі власні числа дійсної симетричної матриці $A = (a_{ij})_{n\times n}$. На першому етапі зведімо задану матрицю послідовно до трьохдіагональної матриці — матриці, у якої всі елементи нульові, за винятком головної діагоналі і діагоналей вище і нижче головної (метод Гаусголдера). Тридіагоналізація Гаусголдера спрощує матрицю не міняючи її власні значення.

1. Метод тридіагоналізації Гаусголдера. Нехай задано дійсно симетричну матрицю $A = (a_{ij})_{n\times n}$. Зведімо її $(n-2)$ перетвореннями подібності за допомогою матриць $P_1, P_2, \dots, P_{n-2}$ до трьохдіагональної форми. Ці матриці є ортогональними і симетричними. Тому

$$P_i^{-1} = P_i^{\mathrm{T}} = P_i, i = \overline{1, n-2}.$$

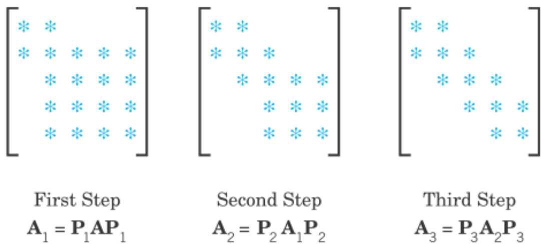
Ці перетворення виконують, починаючи із заданої матриці

$$A_0 = A = (a_{ij})_{n\times n},$$

а саме:

$$A_1 = P_1 A_0 P_1 = (a_{ij}^{(1)})_{n\times n};$$
$$A_2 = P_2 A_1 P_2 = (a_{ij}^{(2)})_{n\times n};$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$B = A_{n-2} = P_{n-2} A_{n-3} P_{n-2} = (a_{ij}^{(n-2)})_{n\times n}.$$

Приміром, для матриці 5-го порядку перетворення матимуть вигляд:



Усі P_k матриці визначають співвідношенням:

$$P_k = E_n - 2\vec{v}_k \cdot \overleftarrow{v}_k,$$

де E_n — одинична матриця, та $\vec{v}_k = (v_{ik})$ є одиничний вектор с першими k компонентами нулями, тому:

$$\vec{v}_1 = \begin{pmatrix} 0 \\ * \\ * \\ \dots \\ * \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 0 \\ 0 \\ * \\ \dots \\ * \end{pmatrix}, \dots, \vec{v}_{n-2} \begin{pmatrix} 0 \\ 0 \\ \dots \\ * \\ * \end{pmatrix},$$

де зірочками позначено інші координати (взагалі кажучи ненульові).

Крок 1. $\vec{v}_1$ має координати:

$$v_{11} = 0;$$

$$v_{21} = \sqrt{\frac{1}{2}\left(1 + \frac{|a_{21}|}{S_1}\right)},$$

$$v_{i1} = \frac{a_{i1} \operatorname{sgn} a_{21}}{2v_{21}S_1}, i = \overline{3, n},$$

де $S = \sqrt{a_{21}^2 + a_{31}^2 + \dots + a_{n1}^2} > 0.$

Отже, ми можемо визначити P_1 та A_1.

Крок 2. Ми обчислюємо $\vec{v}_2$ збільшуючи всі індекси на 1 і замінюючи a_{ij} на $a_{ij}^{(1)}$. Тому

$$v_{12} = v_{22} = 0;$$

$$v_{32} = \sqrt{\frac{1}{2}\left(1 + \frac{|a_{32}^{(1)}|}{S_2}\right)},$$

$$v_{i2} = \frac{a_{i2}^{(1)} \operatorname{sgn} a_{32}^{(1)}}{2v_{32}S_2}, i = \overline{4, n},$$

де $S_2 = \sqrt{a_{32}^{(1)2} + a_{42}^{(1)2} + \dots + a_{n2}^{(1)2}}.$

Тоді знаходять P_2 та A_2.

Крок 3...Знаходимо $\vec{v}_3$, збільшуючи всі індекси на 1 і заміняючи елементи $a_{ij}^{(1)}$ на елементи $a_{ij}^{(2)}$ матриці A і так далі.

10.5. QR-факторизація

1. У цьому методі спершу перетворюють задану дійсну симетричну матрицю $A = (a_{n\times n})$ у трьохдіагональну матрицю $B_0 = B$ методом Гаусголдера. Це утворює велику кількість нулів і зменшує кількість подальшої роботи. Потім обчислюють послідовно матриці $B_1, B_2, \ldots$.

Крок 1. Факторизують $B_0 = Q_0 R_0$ з ортогональною Q_0 і верхньою трикутною R_0. Обчислюють $B_1 = R_0 Q_0$.

Крок 2. Факторизують $B_1 = Q_1 R_2$. Обчислюють $B_2 = R_1 Q_1$.

Загальний крок $s+1$. Факторизують $B_s = Q_s R_s$. Обчислюють $B_{s+1} = R_s Q_s$.

Збіжність до діагональної матриці. Доведімо, що B_{s+1} подібно до B. Маємо $R_s = Q_s^{-1} B_s$. Тоді

$$B_{s+1} = R_s Q_s = Q_s^{-1} B_s Q_s.$$

Тому B_{s+1} подібна до B_s. Отже, B_{s+1} подібна до $B_0 = B$ для всіх s. А це означає, що матриця B_{s+1} має ті самі власні числа що й B.

Доведімо, що матриця B_{s+1} симетрична за методом математичної індукції. А саме, $B_0 = B$ — симетрична. Припускаючи, що B_s — симетрична, маємо $B_s^{\mathrm{T}} = B_s$ і користуючись тим, що $Q_s^{-1} = Q_s^{\mathrm{T}}$, дістаємо

$$B_{s+1}^{\mathrm{T}} = (Q_s^{\mathrm{T}} B_s Q_s)^{\mathrm{T}} = Q_s^{\mathrm{T}} B_s^{\mathrm{T}} Q_s = Q_s^{\mathrm{T}} B_s Q_s = B_{s+1}.$$

Якщо всі власні числа матриці B різні за модулем, скажімо

$$|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|,$$

то

$$\lim_{s\to\infty} B_s = D,$$

де D — діагональна матриця, з елементами $\lambda_1, \lambda_2, \ldots, \lambda_n$ на головній діагоналі.

2. Як одержати QR-факторизацію, скажімо для

$$B = B_0 = (b_{ij}) = Q_0 R_0.$$

Трьохдіагональна матриця B має $n-1$ узагалі ненульових елементів нижче головної діагоналі. Це елементи $b_{21}, b_{32}, \ldots, b_{nn-1}$. Помножмо матрицю B на матрицю C_2 таку що

$$C_2 B = (b_{ij}^{(2)})$$

має $b_{21}^{(2)=0.}$ Множачи цю матрицю на матрицю C_3 таку, що

$$C_3 C_2 B = (b_{ij}^{(3)})$$

має $b_{32}^{(3)} = 0$. Після $n-1$ таких множень залишиться верхня трикутна матриця R_0, а саме

$$C_n C_{n-1} ... C_3 C_2 = R_0.$$

3. Матриці C_j мають просту структуру. C_j має 2×2 підматрицю

$$\begin{pmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{pmatrix}$$

у рядках $j-1$ та j і стовпцях $j-1$ та j. Інші елементи головної діагоналі матриці C_j рівні 1, а всі решта елементів нулі. Приміром, якщо $n = 4, c_j = \cos\theta_j, s_j = \sin\theta_j$, маємо

$$C_2 = \begin{pmatrix} c_2 & s_2 & 0 & 0 \\ -s_2 & c_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, C_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c_3 & s_3 & 0 \\ 0 & -s_3 & c_3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, C_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c_4 & s_4 \\ 0 & 0 & -s_4 & c_4 \end{pmatrix}.$$

Ці матриці C_j є ортогональними. Тому їх добуток та обернена матриця до добутку є також ортогональними. Позначмо цю обернену як Q_0. Тому

$$B_0 = Q_0 R_0,$$

де $C_j^{-1} = C_j^{\mathrm{T}}$,

$$Q_0 = (C_n C_{n-1} ... C_3 C_2)^{-1} = C_2^{\mathrm{T}} C_3^{\mathrm{T}} ... C_{n-1}^{\mathrm{T}} C_n^{\mathrm{T}}.$$

Це і є QR*-факторизація* матриці B_0. Далі маємо

$$B_1 = R_0 Q_0 = R_0 C_2^{\mathrm{T}} C_3^{\mathrm{T}} ... C_{n-1}^{\mathrm{T}} C_n^{\mathrm{T}}.$$

При чому не має потреби знаходити Q_0 у явному вигляді: сперши знаходимо $R_0 C_2^T$, потім $(R_0 C_2^{\mathrm{T}}) C_3^{\mathrm{T}}$ тощо. Так само на подальших кроках дістаємо матриці $B_2, B_3, ...$.

4. Визначення $\cos\theta_j$ **та** $\sin\theta_j$. Покажімо як знаходити кути повороту. Значення $\cos\theta_2$ та $\sin\theta_2$ у матриці C_2 повинні бути такими, що $b_{21}^{(2)} = 0$ у добутку

$$C_2B = \begin{pmatrix} c_2 & s_2 & . & . \\ -s_2 & c_2 & . & . \\ . & . & . & . \\ . & . & . & . \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & . & . \\ b_{21} & b_{22} & . & . \\ . & . & . & . \\ . & . & . & . \end{pmatrix}$$

Отже,

$$b_{21}^{(2)} = -s_2 b_{11} + c_2 b_{21} = -(\sin\theta_2) b_{11} + (\cos\theta_2) b_{21} = 0.$$

Звідки

$$\operatorname{tg}\theta_2 = \frac{s_2}{c_2} = \frac{b_{21}}{b_{11}},$$

і

$$\cos\theta_2 = \frac{1}{\sqrt{1 + \operatorname{tg}^2\theta_2}} = \frac{1}{\sqrt{1 + \left(\dfrac{b_{21}}{b_{11}}\right)^2}}.$$