线性代数 第21讲

11月22日

第五章第3讲　　相似

上一讲内容回顾

相似矩阵

相似变换，相似对角化

Hamilton-Cayley定理

矩阵的特征值，特征向量，特征多项式

$$Ax = \lambda x, \quad x \neq 0 \quad \Rightarrow \quad (\lambda I - A)x = 0$$

$$det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0$$

$$\begin{aligned} det(\lambda I - A) &= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\ &= \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n \\ &= \lambda^n - trace(A)\lambda^{n-1} + \cdots + (-1)^n det(A) \end{aligned}$$

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}, \quad \lambda_1 \lambda_2 \cdots \lambda_n = |A|$$

特征子空间，代数重数与几何重数

定理 5.2.4 设 n 阶方阵 A 的特征多项式为 $p_A(\lambda)$，那么

1. 数 λ_0 是 A 的特征值，当且仅当 $p_A(\lambda_0)=0$，即 λ_0 是特征多项式 $p_A(\lambda)$ 的根.
2. 向量 $\boldsymbol{x}_0\in\mathbb{C}^n$ 是 A 的属于 λ_0 的特征向量，当且仅当 $\boldsymbol{x}_0\in\mathcal{N}(\lambda_0 I_n-A)$ 且 $\boldsymbol{x}_0\neq\mathbf{0}$，即 $\boldsymbol{x}_0$ 是 $\lambda_0 I_n-A$ 的零空间中的非零向量.

解空间 $\mathcal{N}(\lambda_0 I_n-A)$ 称为 A 的属于 λ_0 的**特征子空间**. 注意到特征向量不为零，

$$\{A\text{ 的属于 }\lambda_0\text{ 的特征向量}\}=\mathcal{N}(\lambda_0 I_n-A)\backslash\{\mathbf{0}\}.$$

命题5.3.3　方阵的属于不同特征值的特征向量线性无关.

定义 5.2.10 (代数重数) 给定 n 阶方阵 A 及 A 的一个特征值 $\lambda_0\in\mathbb{C}$，如果 λ_0 是特征多项式 $p_A(\lambda)$ 的 n_0 重根，则称 n_0 为 λ_0 作为 A 的特征值的**代数重数**（简称**重数**），称 λ_0 是 A 的一个 n_0 重特征值.

一个 1 重特征值，又称为**单**特征值.

定义 5.3.6 (几何重数) 给定 n 阶方阵 A 及其特征值 λ_0，称特征子空间 $\mathcal{N}(\lambda_0 I_n-A)$ 的维数为 λ_0 作为 A 的特征值的**几何重数**.

任意特征值的几何重数都不小于 1，因此特征值至少对应一个特征向量.

可对角化与谱分解

定义5.3.1（谱分解） 对方阵 A，如果存在可逆矩阵 X 使得 $X^{-1}AX = \Lambda$ 是对角矩阵，则称 A 是（在 $\mathbb{C}$ 上）可对角化的，X 把 A 对角化，或 X 对角化 A.

如果方阵 A，X，Λ 都是实矩阵，则称A 在 $\mathbb{R}$ 上可对角化.

当 A 可对角化时，分解 $A = X\Lambda X^{-1}$ 称为 A 的谱分解。

之所以称为谱分解，是因为特征值也称为谱.

如果矩阵 A 是可对角化的，即 $A = X\Lambda X^{-1}$,则 $A^n = X\Lambda^n X^{-1}$

$A \in R^{n\times n}$, 假设其特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 满足 $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$,
且对应的特征向量为 $x_1, x_2, \cdots, x_n$, $Ax_k = \lambda x_k \,(k = 1, 2, \cdots, n)$

任取非零向量 $u \in R^n$, $u = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$

$$Au = A(c_1 x_1 + c_2 x_2 + \cdots + c_n x_n) = c_1\lambda_1 x_1 + c_2\lambda_2 x_2 + \cdots + c_n\lambda_n x_n$$

$$A^2 u = A \cdot Au = A(c_1\lambda_1 x_1 + c_2\lambda_2 x_2 + \cdots + c_n\lambda_n x_n) = c_1\lambda_1^2 x_1 + c_2\lambda_2^2 x_2 + \cdots + c_n\lambda_n^2 x_n$$

$\cdots$

$$A^m u = c_1\lambda_1{}^m x_1 + c_2\lambda_2{}^m x_2 + \cdots + c_n\lambda_n{}^m x_n$$

设 $c_1 \neq 0$, 则 $A^m u = \lambda_1{}^m \left(c_1 x_1 + c_2 \dfrac{\lambda_2{}^m}{\lambda_1{}^m} x_2 + \cdots + c_n \dfrac{\lambda_n{}^m}{\lambda_1{}^m} x_n \right) \to \lambda_1{}^m c_1 x_1$

任取非零向量 $u \in R^n$, 当 $m \to \infty$ 时, $A^m u \;\to\; \lambda_1{}^m c_1 x_1$

矩阵可对角化的条件

如果矩阵 A 是可对角化的，即 $A = X\Lambda X^{-1}$

$$A = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}^{-1}$$

$$A\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} AX_1 & AX_2 & \cdots & AX_n \end{bmatrix} = \begin{bmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \cdots & \lambda_n X_n \end{bmatrix}$$

命题5.3.2 对 n 阶方阵 A，A 可对角化，当且仅当 A 有 n 个线性无关的特征向量.

矩阵可对角化的条件

矩阵可对角化的一个充分条件

推论5.3.4 有 n 个不同特征值的 n 阶方阵，即特征值都是单特征值的方阵，可对角化.

命题5.3.7 方阵的特征值的几何重数不大于其代数重数.

定义5.3.8 几何重数和代数重数相等的特征值，称为半单特征值. 几何重数小于代数重数的特征值，称为亏损特征值.

如果一个特征值的代数重数是1，那么由几何重数不小于1。
可知，它是半单特征值，即单特征值是半单特征值.

定理5.3.9

1. n 阶方阵 A 可对角化，当且仅当其特征值都半单(几何重数和代数重数相等).
2. n 阶实方阵 A 在 $\mathbb{R}$ 上可对角化，当且仅当其特征多项式的根都是实根，

   且其特征值都半单.

设矩阵 $A=\begin{bmatrix}1&1&1\\a&b&c\\d&e&f\end{bmatrix}$，且 $\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}1\\-1\\0\end{bmatrix}$ 是 A 的特征向量，求 a,b,c,d,e,f.

解：因为三个特征向量 $\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}1\\-1\\0\end{bmatrix}$ 线性无关，所以矩阵 A 可对角化，

设特征向量 $\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}1\\-1\\0\end{bmatrix}$ 对应的特征值分别为 λ_1，λ_2，λ_3

则由 $\begin{bmatrix}1&1&1\\a&b&c\\d&e&f\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}=\lambda_1\begin{bmatrix}1\\1\\1\end{bmatrix}$，解得 $\lambda_1=3$，同理解得 $\lambda_2=0$，$\lambda_3=0$

所以 $\begin{bmatrix}1&1&1\\a&b&c\\d&e&f\end{bmatrix}=\begin{bmatrix}1&1&1\\1&0&-1\\1&-1&0\end{bmatrix}\begin{bmatrix}3&&\\&0&\\&&0\end{bmatrix}\begin{bmatrix}1&1&1\\1&0&-1\\1&-1&0\end{bmatrix}^{-1}=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$

不可对角化矩阵的例子

矩阵 $S_k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} (k \neq 0)$，只有一个特征值 1，其代数重数是 2，几何重数是 1，

该矩阵不可对角化.

矩阵 $J_n(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}_{n\times n}$ 只有一个特征值 λ，其代数重数是 n，几何重数是 1，当 $n > 1$ 时 λ 是亏损特征值. 形如 $J_n(\lambda)$ 的矩阵称为关于 λ 的 n 阶 **Jordan 块**. 当 $n > 1$ 时，它不可对角化. ☺

$(a_{22}-a_{11})\boldsymbol{e}_2+a_{12}\boldsymbol{e}_1$

勘误，教材197页

命题 5.2.5 上（下）三角矩阵的全部特征值就是其所有对角元素.

例 5.2.6 观察二阶上三角矩阵 $A=\begin{bmatrix} a_{11} & a_{12} \\ & a_{22} \end{bmatrix}$. 显然, A 的特征多项式是 $(\lambda-a_{11})(\lambda-a_{22})$.

当 $a_{11}\neq a_{22}$ 时，A 的全部特征值是 a_{11},a_{22}，二者对应的特征子空间分别为

$$\mathcal{N}(a_{11}I_2-A)=\mathrm{span}(\boldsymbol{e}_1),\quad \mathcal{N}(a_{22}I_2-A)=\mathrm{span}((a_{11}-a_{22})\boldsymbol{e}_2+k\boldsymbol{e}_1).$$

特别地，A 有两个线性无关的特征向量 $\boldsymbol{e}_1,(a_{22}-a_{11})\boldsymbol{e}_2+a_{12}\boldsymbol{e}_1$.

当 $a_{11}=a_{22}$ 时，A 的全部特征值是 a_{11}. 如果 $a_{12}=0$，则其对应的特征子空间为 $\mathcal{N}(a_{11}I_2-A)=\mathbb{R}^2$. 特别地，$A$ 有两个线性无关的特征向量 $\boldsymbol{e}_1,\boldsymbol{e}_2$.

如果 $a_{12}\neq 0$，则其对应的特征子空间为 $\mathcal{N}(a_{11}I_2-A)=\mathrm{span}(\boldsymbol{e}_1)$. A 只有一个线性无关的特征向量 $\boldsymbol{e}_1$.

可以看到，对对角矩阵和上三角矩阵，尽管特征值都是对角元素，但对角元素是否相等对特征子空间的影响很大. 对角元素从不等到相等，对对角矩阵，特征子空间从两个一维子空间变成一个二维子空间；而对非对角的上三角矩阵，特征子空间从两个一维子空间变成一个一维子空间. ☺

命题 5.3.11 设分块对角矩阵 $A=\begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{bmatrix}$，其中 $A_i, i=1,\cdots,r$ 都是方阵，则 A 可对角化当且仅当所有 A_i 都可对角化.

证. 不妨设 $r=2$，否则反复应用 $r=2$ 的情形可得. $\boldsymbol{A}=\begin{bmatrix} \boldsymbol{A}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}_2 \end{bmatrix}$

充分性：如果 A_1, A_2 都可对角化，则存在可逆矩阵 X_1, X_2，

使得 $X_1^{-1}A_1X_1=D_1$，$X_2^{-1}A_2X_2=D_2$ 都是对角矩阵. 则

$$\begin{bmatrix} X_1^{-1} & \\ & X_2^{-1} \end{bmatrix}\begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}\begin{bmatrix} X_1 & \\ & X_2 \end{bmatrix}=\begin{bmatrix} D_1 & \\ & D_2 \end{bmatrix}$$ 也是对角矩阵，

而 $X=\begin{bmatrix} X_1 & \\ & X_2 \end{bmatrix}$ 可逆，且 $X^{-1}=\begin{bmatrix} X_1^{-1} & \\ & X_2^{-1} \end{bmatrix}$.

命题 5.3.11 设分块对角矩阵 $A=\begin{bmatrix}A_1 & & \\ & \ddots & \\ & & A_r\end{bmatrix}$，其中 $A_i, i=1,\cdots,r$ 都是方阵，则 A 可对角化当且仅当所有 A_i 都可对角化.

证. 不妨设 $r=2$，否则反复应用 $r=2$ 的情形可得. $A=\begin{bmatrix}A_1 & 0\\ 0 & A_2\end{bmatrix}$

必要性：首先 $p_A(\lambda)=p_{A_1}(\lambda)p_{A_2}(\lambda)$，因此 A_1，A_2的特征值必为 A的特征值 。

λ_0 是 A 的特征值，其代数重数为 n_0,因为 A可对角化，所以其几何重数 $m_0=n_0$,

λ_0 作为 A_1 的特征值，其代数重数与几何重数分别为 n_{01}和 m_{01},

λ_0 作为 A_2 的特征值，其代数重数与几何重数分别为 n_{02} 和 m_{02},

$n_0=n_{01}+n_{02}$, $\quad m_{01}\le n_{01}, m_{02}\le n_{02}$

$$m_0\overset{?}{=}m_{01}+m_{02}$$

$$m_0=\dim N\left(\begin{bmatrix}\lambda_0 I_{r_1}-A_1 & \\ & \lambda_0 I_{r_2}-A_2\end{bmatrix}\right)=r_1+r_2-rank\left(\lambda_0 I_{r_1}-A_1\right)-rank\left(\lambda_0 I_{r_2}-A_2\right)$$

$$=\dim N\left(\lambda_0 I_{r_1}-A_1\right)+\dim N\left(\lambda_0 I_{r_2}-A_2\right)=m_{01}+m_{02}$$

$$\left.\begin{array}{c} m_{01}\le n_{01},\quad m_{02}\le n_{02}\\ m_0=m_{01}+m_{02}\le n_{01}+n_{02}=n_0\\ m_0=n_0\end{array}\right\}\Rightarrow m_{01}=n_{01}, m_{02}=n_{02}$$

所以 A_1, A_2 的任意特征值，几何重数等于代数重数 (半单)，所以 A_1, A_2都是可对角化的 。

例5.3.12（利用矩阵的谱分解求 Fibonacci 数列的通项公式）

给定$F_0 = 0$，$F_1 = 1$ 和递推关系$F_{n+1} = F_n + F_{n-1}$. 试给出 F_n 的通项公式.

这是一个两步递推，因此难以直接得到通项. 然而，可以写出矩阵形式的一步递推：

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix}.$$

$\boldsymbol{u}_{n+1} = A\boldsymbol{u}_n$，那么 $\boldsymbol{u}_n = A^{n-1}\boldsymbol{u}_1$. 而 $\boldsymbol{u}_1 = \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

设 $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \boldsymbol{X\Lambda X}^{-1}$，则 $\boldsymbol{A}^{n-1} = \boldsymbol{X\Lambda}^{n-1}\boldsymbol{X}^{-1}$

$$\begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix}$$

$$A = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix},$$

其中 $\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}$ 是两个特征值.

$$F_n = \boldsymbol{e}_1^{\mathrm{T}} \boldsymbol{u}_n = \boldsymbol{e}_1^{\mathrm{T}} \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^{n-1} & \\ & \lambda_2^{n-1} \end{bmatrix} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2}.$$

相似

定义5.4.1（相似）对方阵 A，B，如果存在可逆矩阵 X 使得 $X^{-1}AX = B$，则称 A 和 B 相似，或 A 相似于 B，$A \backsim B$.

一个矩阵可对角化（$X^{-1}AX = \Lambda$）当且仅当它相似于对角矩阵.

例5.4.2 数量矩阵只与自己相似：$X^{-1}(kI_n)X = kI_n$.

对角矩阵 $\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}$ 与 $\begin{bmatrix} \lambda_2 & \\ & \lambda_1 \end{bmatrix}$ 相似：$\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$.

命题 5.4.4 方阵的相似关系有如下不变量：

1. 秩；
2. 特征多项式、特征值、特征值的代数重数、迹、行列式；
3. 特征值的几何重数.

命题5.4.4 方阵的相似关系有如下不变量：

1. 秩； 2. 特征多项式、特征值、特征值的代数重数、迹、行列式；

3. 特征值的几何重数.

证. 1. 显然 $\mathrm{rank}(X^{-1}AX) = \mathrm{rank}(A)$，秩是不变量.

2. $\det(\lambda I_n - X^{-1}AX) = \det(\lambda I_n - A)$，特征多项式是不变量.

特征多项式决定了特征值、特征值的代数重数、迹、行列式，因此它们都是不变量.

3. 考虑几何重数. 如果 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_r$ 是 A 的属于 λ_0 的特征子空间的一组基，

则 $X^{-1}\boldsymbol{u}_1, \cdots, X^{-1}\boldsymbol{u}_r$ 就是 $X^{-1}AX$ 的属于 λ_0 的 r 个线性无关的特征向量.

事实上，容易验证 $X^{-1}AX(X^{-1}\boldsymbol{u}_i) = X^{-1}(\lambda_0\boldsymbol{u}_i) = \lambda_0(X^{-1}\boldsymbol{u}_i)$，$i = 1, \cdots, r$，

即它们是属于 λ_0 的特征向量.

另一方面，由于 $X^{-1}[\boldsymbol{u}_1 \cdots \boldsymbol{u}_r] = [X^{-1}\boldsymbol{u}_1 \cdots X^{-1}\boldsymbol{u}_r]$，

而 X^{-1} 可逆，因此 $[\boldsymbol{u}_1 \cdots \boldsymbol{u}_r]$ 和 $[X^{-1}\boldsymbol{u}_1 \cdots X^{-1}\boldsymbol{u}_r]$ 这两个矩阵的秩相等，

因此 $X^{-1}\boldsymbol{u}_1, \cdots, X^{-1}\boldsymbol{u}_r$ 线性无关.

于是 λ_0 作为 A 的特征值的几何重数不大于 λ_0 作为 $X^{-1}AX$ 的特征值的几何重数。

相似变换

命题5.4.3 方阵的相似关系是等价关系. 满足：反身性，对称性，传递性

等价关系：向量组的线性表示（生成子空间），矩阵相抵（秩）

矩阵相似（特征值, 几何重数, 即是否可对角化）

对应于相似关系的等价变换称为**相似变换**，即 $\boldsymbol{T}$: $A \mapsto X^{-1}AX$.

如果 A 可对角化，则称对角化得到的对角矩阵为 A 的相似标准形.

在 A 可对角化时，
它的相似标准形在对角元素相差一个排列次序的意义下唯一.

$$\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&0&&1&\\&&&\ddots&&\\&&1&&0&\\&&&&&\ddots\\&&&&&&1\end{bmatrix}\begin{bmatrix}\lambda_1&&&&&\\&\ddots&&&&\\&&\lambda_i&&&\\&&&\ddots&&\\&&&&\lambda_j&\\&&&&&\ddots\\&&&&&&\lambda_n\end{bmatrix}\begin{bmatrix}1&&&&&\\&\ddots&&&&\\&&0&&1&\\&&&\ddots&&\\&&1&&0&\\&&&&&\ddots\\&&&&&&1\end{bmatrix}^{-1}=\begin{bmatrix}\lambda_1&&&&&\\&\ddots&&&&\\&&\lambda_j&&&\\&&&\ddots&&\\&&&&\lambda_i&\\&&&&&\ddots\\&&&&&&\lambda_n\end{bmatrix}$$

任何方阵一定可以相似变换为上三角矩阵

命题5.4.6 对 n 阶方阵 A，存在可逆矩阵 X，使得 $X^{-1}AX = T$，
T 是上三角矩阵，且 T 的对角元素是 A 的 n 个特征值（计重数）.
进一步地，通过选择特定的 X，能够令 T 的对角元素是 A 的特征值的任意排列.

证. 采用数学归纳法. $n = 1$ 时，显然.

假设任意 $n - 1$ 阶方阵都有如上分解，观察 n 阶方阵.

任取 A 的一个特征对 $(\lambda_1, \boldsymbol{x}_1)$，把 $\boldsymbol{x}_1$ 扩充成 $\mathbb{C}^n$ 的一组基 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$.

记 $X_1 = [\boldsymbol{x}_1\ X_{12}] = [\boldsymbol{x}_1\ \boldsymbol{x}_2 \cdots \boldsymbol{x}_n]$. 则 AX_1 的列能被 X_1 的列线性表示，

$$AX_1 = \begin{bmatrix} A\boldsymbol{x}_1 & AX_{12} \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{x}_1 & AX_{12} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & X_{12} \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & A_1 \end{bmatrix} = X_1 M_1,$$

$$AX_1 = \begin{bmatrix} A\boldsymbol{x}_1 & AX_{12} \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{x}_1 & AX_{12} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & X_{12} \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & A_1 \end{bmatrix} = X_1M_1,$$

即 M_1 是分块上三角矩阵. 而 A_1 是 $n-1$ 阶方阵，根据归纳假设，存在可逆矩阵 X_2，使得 $A_1 = X_2T_2X_2^{-1}$，其中 T_2 上三角. 于是

$$A = X_1 \begin{bmatrix} \lambda_1 & * \\ 0 & X_2T_2X_2^{-1} \end{bmatrix} X_1^{-1} = X_1 \begin{bmatrix} 1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ 0 & T_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & X_2 \end{bmatrix}^{-1} X_1^{-1}.$$

记 $X = X_1 \begin{bmatrix} 1 & 0 \\ 0 & X_2 \end{bmatrix}, T = \begin{bmatrix} \lambda_1 & * \\ 0 & T_2 \end{bmatrix}$，则 X 可逆，T 上三角，且 $A = XTX^{-1}$.

综上，命题对任意 n 成立.

考察特征多项式，立得 T 的对角元素就是 A 的 n 个特征值.

通过上述证明能看到，按照特征值的排列选择特征对，

就能使 T 的对角元素满足条件.

计算矩阵特征值的QR算法

$$A = Q_1 R_1$$

$$A_1 = Q_1 R_1, \quad A_2 = R_1 Q_1$$

$$A_2 = Q_2 R_2, \quad A_3 = R_2 Q_2$$

……

$$A_k = Q_k R_k, \quad A_{k+1} = R_k Q_k, \quad \cdots, \quad A_m \to \text{上三角矩阵}$$

$$A_1 = Q_1 R_1, \quad A_2 = R_1 Q_1 \quad \Rightarrow \quad A_2 = Q_1^T A_1 Q_1 = Q_1^{-1} A_1 Q_1$$

$$A_2 = Q_2 R_2, \quad A_3 = R_2 Q_2 \quad \Rightarrow \quad A_3 = Q_2^T A_2 Q_2 = Q_2^T Q_1^T A_1 Q_1 Q_2 = \left(Q_1 Q_2\right)^T A_1 Q_1 Q_2$$

……

$$A_k = Q_k R_k, \quad A_{k+1} = R_k Q_k$$

$$A_{k+1} = \left(Q_1 \cdots Q_k\right)^T A \left(Q_1 \cdots Q_k\right) = \left(Q_1 \cdots Q_k\right)^{-1} A \left(Q_1 \cdots Q_k\right)$$

定理5.4.7 (Hamilton-Cayley 定理)

设方阵 A 的特征多项式为 $p_A(\lambda)$，则 $p_A(A) = O$.

证. 设 A 有分解 $A = XTX^{-1}$，其中 T 是上三角矩阵.

而 $p_A(A) = Xp_A(T)X^{-1}$，只需证明 $p_A(T) = O$.

若 $\lambda_1, \cdots, \lambda_n$ 是 A 的 n 个特征值，则 $p_A(\lambda) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$，

于是 $p_A(T) = (T - \lambda_1 I_n)\cdots(T - \lambda_n I_n)$.

注意 $T - \lambda_i I_n$ 的第 i 个对角元素是零，逐步计算矩阵乘法就得到 $p_A(T) = O$.

$$\begin{bmatrix} O_r & a_{r\times 1} & b_{r\times(n-r-1)} \\ & u_{r+1} & c_{1\times(n-r-1)} \\ & & U_{n-r-1} \end{bmatrix}\begin{bmatrix} U_r & a'_{r\times 1} & b'_{r\times(n-r-1)} \\ & 0 & c'_{1\times(n-r-1)} \\ & & U'_{n-r-1} \end{bmatrix}$$

$$= \begin{bmatrix} O_r & O_{r\times 1} & a_{r\times 1}c'_{1\times(n-r-1)} + b_{r\times(n-r-1)}U'_{n-r-1} \\ & 0 & u_{r+1}c'_{1\times(n-r-1)} + c_{1\times(n-r-1)}U'_{n-r-1} \\ & & U_{n-r-1}U'_{n-r-1} \end{bmatrix}$$

Hamilton - Cayley定理证法2

设 $p_A(\lambda)=|\lambda I-A|=\lambda^n+c_{n-1}\lambda^{n-1}+\cdots+c_1\lambda+c_0,$

则 $p_A(A)=A^n+c_{n-1}A^{n-1}+\cdots+c_1A+c_0I=\mathbf{0}.$

证明 设 $B(\lambda)=(\lambda I-A)^*$，则有 $B(\lambda)(\lambda I-A)=|\lambda I-A|I=p_A(\lambda)I.$

记 $B(\lambda)=\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\cdots+B_0,$

因为 $p_A(\lambda)=\lambda^n+c_{n-1}\lambda^{n-1}+\cdots+c_1\lambda+c_0,$

所以 $\left(\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\cdots+B_0\right)(\lambda I-A)$

$$=\left(\lambda^n+c_{n-1}\lambda^{n-1}+\cdots+c_1\lambda+c_0\right)I$$

$$\left(\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\cdots+B_0\right)(\lambda I-A)=\left(\lambda^n+c_{n-1}\lambda^{n-1}+\cdots+c_1\lambda+c_0\right)I$$

展开比较两边同次项系数得

$$\begin{cases} B_{n-1}=I \\ B_{n-2}-B_{n-1}A=a_{n-1}I \\ \vdots \\ B_0-B_1A=a_1I \\ -B_0A=a_0I \end{cases} \qquad \text{所以}\begin{cases} B_{n-1}A^n=A^n \\ B_{n-2}A^{n-1}-B_{n-1}AA^{n-1}=c_{n-1}A^{n-1} \\ \vdots \\ B_0A-B_1AA=c_1A \\ -B_0A=c_0I \end{cases}$$

所以 $\mathbf{0}=A^n+c_{n-1}A^{n-1}+\cdots+c_1A+c_0=p_A(A).$

设 A 为 n 阶矩阵， $p(A)=c_mA^m+\cdots+c_1A+c_0$ 也为 n 阶矩阵，$p(A)=c_mA^m+\cdots+c_1A+c_0I=0$,则称 $p(x)$ 是 A 的化零多项式。

特征多项式是化零多项式.

Jordan标准型

定理 5.4.8 (Jordan 分解) 对 n 阶方阵 A，存在可逆矩阵 X，使得

$$X^{-1}AX = J = \begin{bmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_r}(\lambda_r) \end{bmatrix},$$

其中 $J_{n_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}_{n_i \times n_i}$ 是 n_i 阶 Jordan 块，而 $n_1 + \cdots + n_r = n$，且除了这些 Jordan 块的排列次序外，J 被 A 唯一确定.

同时对角化

定义5.4.10（同时对角化） 设 A，B 是 n 阶方阵，如果存在可逆矩阵 X 使得 $X^{-1}AX = \Lambda_1$ 和 $X^{-1}BX = \Lambda_2$ 都是对角矩阵，则称 A，B 可以同时对角化.

命题 5.4.11 对可对角化的 n 阶方阵 A, B，以下叙述等价：

1. A, B 可以同时对角化；
2. 存在 n 个线性无关的向量，同时是 A, B 的特征向量；
3. A, B 可交换，即 $AB = BA$.

证. 显然，1 和 2 等价.

"1 ⇒ 3"：设 $X^{-1}AX = \Lambda_1, X^{-1}BX = \Lambda_2$，因为对角矩阵可交换，于是 $AB = XX^{-1}AXX^{-1}BXX^{-1} = X\Lambda_1\Lambda_2X^{-1} = X\Lambda_2\Lambda_1X^{-1} = XX^{-1}BXX^{-1}AXX^{-1} = BA$.

"3 ⇒ 1"：由于 A 可对角化，存在可逆矩阵 X_1，使得 $X_1^{-1}AX_1 = \widetilde{\Lambda}_1$ 是对角矩阵. 再利用置换矩阵 P，可将 $\widetilde{\Lambda}_1$ 变为如下形式：$P^{-1}\widetilde{\Lambda}_1P = \Lambda_1 = \begin{bmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_r I_{n_r} \end{bmatrix}$，其中 $\lambda_1, \cdots, \lambda_r$ 两两不同.

此时 $P^{-1}X_1^{-1}AX_1P=\Lambda_1$. 令 $\widetilde{B}=P^{-1}X_1^{-1}BX_1P$, 则由 $AB=BA$ 可知 $\Lambda_1\widetilde{B}=\widetilde{B}\Lambda_1$.

把 $\widetilde{B}$ 写成与 Λ_1 对应的分块矩阵 $\widetilde{B}=\begin{bmatrix}\widetilde{B}_{11} & \cdots & \widetilde{B}_{1r}\\ \vdots & & \vdots\\ \widetilde{B}_{r1} & \cdots & \widetilde{B}_{rr}\end{bmatrix}$, 由 $\Lambda_1\widetilde{B}=\widetilde{B}\Lambda_1$ 可知

$(\lambda_iI_{n_i})\widetilde{B}_{ij}=\widetilde{B}_{ij}(\lambda_jI_{n_j})$, $(\lambda_i-\lambda_j)\widetilde{B}_{ij}=0$. 对 $i\neq j$, 有 $\widetilde{B}_{ij}=0$, 即 $\widetilde{B}=\begin{bmatrix}\widetilde{B}_{11} & & \\ & \ddots & \\ & & \widetilde{B}_{rr}\end{bmatrix}$

也是分块对角矩阵.

由 B 可对角化可得 $\widetilde{B}$ 可对角化, 根据命题 5.3.11 , $\widetilde{B}_{ii}$ 都可对角化. 设 $X_{ii}^{-1}\widetilde{B}_{ii}X_{ii}=\Lambda_{ii}$, 其中 X_{ii} 可逆, Λ_{ii} 对角. 令 $X_2=\begin{bmatrix}X_{11} & & \\ & \ddots & \\ & & X_{rr}\end{bmatrix}$, 那么 X_2 可逆, 且 $X_2^{-1}\widetilde{B}X_2=\begin{bmatrix}\Lambda_{11} & & \\ & \ddots & \\ & & \Lambda_{rr}\end{bmatrix}=:\Lambda_2$. 又有 $X_2^{-1}\Lambda_1X_2=\begin{bmatrix}X_{11}^{-1}\lambda_1I_{n_1}X_{11} & & \\ & \ddots & \\ & & X_{rr}^{-1}\lambda_rI_{n_r}X_{rr}\end{bmatrix}=\Lambda_1$,

因此

$$X_2^{-1}P^{-1}X_1^{-1}AX_1PX_2=\Lambda_1,\quad X_2^{-1}P^{-1}X_1^{-1}BX_1PX_2=\Lambda_2,$$

即 X_1PX_2 将 A,B 同时对角化. $\square$

作 业（11月22日）

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

练习5.3. 6, 15, 20

练习5.4. 1, 2, 3, 5, 6

11月29日提交

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

11	22	23	24	25	26	27	28	相似	5.4	实对称矩阵	6.1
12	29	30	1	2	3	4	5	正定矩阵	6.2	奇异值分解	6.3
12	6	7	8	9	10	11	12	奇异值分解	6.3	线性空间，基和维数	7.1 7.2
12	13	14	15	16	17	18	19	基和维数，线性映射	7.2 7.3	向量的坐标	7.4
12	20	21	22	23	24	25	26	线性映射的矩阵	7.5	复习	