矩阵分析与应用

第五讲　　线性变换之三

信息工程学院
吕旋阳

2006-12-1

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

引入

有限维线性空间V中取定一组基后，V的任一线性变换都可以用矩阵来表示．为了研究线性变换性质，希望这个矩阵越简单越好，如对角矩阵．

从本节开始，我们主要讨论，如何选择一组适当的基，使V的某个线性变换在这组基下的矩阵就是一个对角矩阵?

约束极值与特征值

考虑一个几何约束条件的实对称二次型的极值问题

假定 $x \in R^n, x^T x = 1$ ，求 $x^T Ax$ 的极大值

其中 $A = A^T \in R^{n\times n}$ 。

Lagrauge函数 $L = x^T Ax - \lambda x^T x$

有极值的必要条件 $0 = \nabla L = 2(Ax - \lambda x) = 0$

也就是说，**A**的特征值、特征向量对是极值问题的解

一、特征值与特征向量

定义：设 T 是数域K上线性空间V的一个线性变换，

若对于K中的一个数 λ_0，存在一个V的非零向量 x，

使得

$$T(x)=\lambda_0 x \quad ,$$

则称 λ_0 为 T 的一个**特征值**，称 x 为 T 的属于特征值

λ_0 的**特征向量**.

注：① 几何意义：特征向量经线性变换后方向保持

相同 $(\lambda_0 > 0)$ 或相反 $(\lambda_0 < 0)$. $\lambda_0 = 0$ 时，$T(x) = 0$.

② 若 x 是 T 的属于特征值 λ_0 的特征向量，则

$kx \quad (k \in K, k \neq 0)$ 也是 T 的属于 λ_0 的特征向量.

$$\left(\because \quad T(kx) = kT(x) = k(\lambda_0 x) = \lambda_0 (kx) \right)$$

由此知，特征向量不是被特征值所唯一确定的，

但是特征值却是被特征向量所唯一确定的，即

若 $T(x) = \lambda x$ 且 $T(x) = \mu x$ ，则 $\lambda = \mu$.

二、特征值与特征向量的求法

分析：设 $\dim V = n$，$x_1, x_2, \cdots, x_n$ 是V的一组基，

线性变换 T 在这组基下的矩阵为A.

设 λ_0是 T 的特征值，它的一个特征向量 x在基

$x_1, x_2, \cdots, x_n$下的坐标记为 $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

则 $T(x)$在基 $x_1, x_2, \cdots, x_n$下的坐标为 $A\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

而$\lambda_0 x$的坐标是 $\lambda_0 \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$, 又 $T(x) = \lambda_0 x$

于是 $A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \lambda_0 \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$, 从而 $(\lambda_0 I - A) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = 0.$

即 $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ 是线性方程组 $(\lambda_0 I - A)X = 0$ 的解，

又 $\because x \neq 0$, $\therefore \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \neq 0$, $\therefore$ $(\lambda_0 I - A)X = 0$ 有非零解.

所以它的系数行列式 $|\lambda_0 I - A| = 0.$

以上分析说明：

若λ_0是T的特征值，则 $\left|\lambda_0 I - A\right| = 0.$

反之，若$\lambda_0 \in K$满足 $\left|\lambda_0 I - A\right| = 0,$

则齐次线性方程组 $(\lambda_0 I - A)X = 0$ 有非零解.

若 $(\xi_1, \xi_2, \cdots, \xi_n)^T$ 是 $(\lambda_0 I - A)X = 0$ 一个非零解，

则向量 $x = \xi_1 x_1 + \cdots + \xi_n x_n$ 就是T的属于λ_0的一个

特征向量.

设 $A \in K^{n\times n}$，λ是一个参数，矩阵 $\lambda I - A$ 称为

A的特征矩阵，它的行列式

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & & \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为A的特征多项式.

（ $|\lambda I - A|$ 是数域K上的一个n次多项式）

注：① 若矩阵A是线性变换 T 关于V的一组基的矩阵，而 λ_0 是 T 的一个特征值，则 λ_0 是特征多项式 $|\lambda I - A|$ 的根，即 $|\lambda_0 I - A| = 0.$

反之，若 λ_0 是A的特征多项式的根，则 λ_0 就是 T 的一个特征值.（所以，特征值也称特征根.）

② 矩阵A的特征多项式的根有时也称为A的特征值，而相应的线性方程组 $(\lambda I - A)X = 0$ 的非零解也就称为A的属于这个特征值的特征向量.

求特征值与特征向量的一般步骤

i) 在V中任取一组基 $x_1, x_2, \cdots, x_n$ 写出 T 在这组基下的矩阵A .

ii) 求A的特征多项式 $|\lambda I - A|$ 在K上的全部根它们就是 T 的全部特征值.

iii) 把所求得的特征值逐个代入方程组

$$(\lambda I - A)X = 0$$

并求出它的一组基础解系.(它们就是属于这个特征值的全部线性无关的特征向量在基 $x_1, x_2, \cdots, x_n$ 下的坐标.)

例:在线性空间V中，数乘变换K在任意一组基下的矩阵都是数量矩阵kI，它的特征多项式是

$$|\lambda I - kI| = (\lambda - k)^n.$$

故数乘法变换K的特征值只有数k，且

对 $\forall x \in V \quad (x \neq 0)$，皆有 $K(x) = kx.$

所以，V中任一非零向量皆为数乘变换K的特征向量.

例: 设线性变换 T 在基 x_1, x_2, x_3 下的矩阵是

$$A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix},$$

求 T 特征值与特征向量.

解：A的特征多项式

$$|\lambda I-A|=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&-2\\-2&-2&\lambda-1\end{vmatrix}=(\lambda+1)^2(\lambda-5)$$

故 T 的特征值为： $\lambda_1=-1$（二重），$\lambda_2=5$

把 $\lambda=-1$ 代入齐次方程组 $(\lambda I-A)X=0$，得

$$\begin{cases}-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\end{cases}$$

即 $x_1+x_2+x_3=0$

它的一个基础解系为：$(1,0,-1)$，$(0,1,-1)$

因此，属于-1的两个线性无关的特征向量为

$$y_1=x_1-x_3,\quad y_2=x_2-x_3$$

而属于-1的全部特征向量为

$$k_1y_1+k_2y_2,\qquad (k_1,k_2\in K\text{ 不全为零})$$

把 $\lambda=5$ 代入齐次方程组 $(\lambda I-A)X=0$，得

$$\begin{cases} 4x_1-2x_2-2x_3=0 \\ -2x_1+4x_2-2x_3=0 \\ -2x_1-2x_2+4x_3=0 \end{cases}$$

解得它的一个基础解系为：$(1,1,1)$

因此，属于5的一个线性无关的特征向量为

$$y_3=x_1+x_2+x_3$$

而属于5的全部特征向量为

$$k_3y_3,\qquad (k_3\in K,\ k_3\neq 0\ )$$

定义：设T为n维线性空间V的线性变换，λ_0为T的一个特征值，令V_{λ_0}为T的属于λ_0的全部特征向量再添上零向量所成的集合，即 $V_{\lambda_0}=\left\{x \mid Tx=\lambda_0 x\right\}$

则V_{λ_0}是V的一个子空间, 称之为T的一个特征子空间.

$$\because \quad T(x+y)=T(x)+T(y)=\lambda_0 x+\lambda_0 y=\lambda_0(x+y)$$

$$T(kx)=kT(x)=k(\lambda_0 x)=\lambda_0(kx)$$

$$\therefore \quad x+y\in V_{\lambda_0}, \qquad kx\in V_{\lambda_0}$$

若 T 在n维线性空间V的某组基下的矩阵为A，则

$$\dim V_{\lambda_0} = n - 秩(\lambda_0 I - A)$$

即特征子空间 V_{λ_0}的维数等于齐次线性方程组

$$(\lambda_0 I - A)X = 0 \qquad (*)$$

的解空间的维数，且由方程组(*)得到的属于λ_0的全部线性无关的特征向量就是 V_{λ_0}的一组基.

特征多项式的有关性质

1. 设 $A=\left(a_{ij}\right)\in K^{n\times n}$，则A的特征多项式

$$\left|\lambda I-A\right|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & & \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

$$=\lambda^{n}-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^{n}\left|A\right|$$

由多项式根与系数的关系还可得

① A的全体特征值的和$=\underline{a_{11}+a_{22}+\cdots+a_{nn}}$.

② A的全体特征值的积$=\left|A\right|$.

称之为A的迹,记作trA

定理：设 $A=(a_{ij})_{n\times n}, B=(b_{ij})_{n\times n}$ 则 $\mathrm{tr}(AB)=\mathrm{tr}(BA)$

证：令 $AB=(u_{ij}), BA=(v_{ij})$，于是有

$$u_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj},\quad v_{ij}=\sum_{l=1}^{n}b_{il}a_{lj}$$

$$\therefore\quad \mathrm{tr}(AB)=\sum_{i=1}^{n}u_{ii}=\sum_{i=1}^{n}\left[\sum_{k=1}^{n}a_{ik}b_{ki}\right]$$

$$=\sum_{k=1}^{n}\left[\sum_{i=1}^{n}b_{ki}a_{ik}\right]=\sum_{k=1}^{n}v_{kk}=\mathrm{tr}(BA)$$

定理：相似矩阵有相似的迹 $\quad \mathrm{tr}(AB)=\mathrm{tr}(BA)$

证：设 $A\sim B$，即 $\exists P\neq 0,\mathrm{st}\quad B=P^{-1}AP$

$$
\begin{aligned}
\mathrm{tr}(B)&=\mathrm{tr}(P^{-1}AP)\\
&=\mathrm{tr}(APP^{-1})\\
&=\mathrm{tr}(A)
\end{aligned}
$$

定理： 相似矩阵具有相同的特征多项式.

证:设 $A \sim B$, 则存在可逆矩阵P，使得

$$B = P^{-1}AP$$

于是， $|\lambda I - B| = |\lambda I - P^{-1}AP|$

$$= |\lambda P^{-1}IP - P^{-1}AP|$$

$$= |P^{-1}(\lambda I - A)P|$$

$$= |P^{-1}||\lambda I - A||P|$$

$$= |\lambda I - A|$$

注：① 由定理线性变换T的特征值与基的选择无关.

因此,矩阵A的特征多项式也说成是线性变换 T 的特征多项式；而线性变换 T 的特征值与特征向量有时也说成是矩阵A的特征值与特征向量.

② 有相同特征多项式的矩阵未必相似.

如 $A=\begin{pmatrix}1&0\\0&1\end{pmatrix}, B=\begin{pmatrix}1&1\\0&1\end{pmatrix}$

它们的特征多项式都是 $(\lambda-1)^2$; 但A、B不相似.

定理：任意n阶矩阵A与三角矩阵相似

证明:对阶数n利用数学归纳法证明. 当n=1时显然成立

设当阶数为n–1时定理成立。

设 $x_1, x_2, \cdots, x_n$ 是n个线性无关的列向量，

其中x_1 为A的特征值λ的特征向量, $Ax_1 = \lambda_1 x_1$

记 $P_1 = \left(x_1, x_2, \cdots, x_n\right)$

于是 $AP_1 = \left(Ax_1, Ax_2, \cdots, Ax_n\right) = \left(\lambda x_1, Ax_2, \cdots, Ax_n\right)$

由于$Ax_i \in C^n$,因此可以由 $x_1, x_2, \cdots, x_n$唯一地线性表示

即有 $Ax_i = b_{1i}x_1 + b_{2i}x_2 + \cdots + b_{ni}x_n$

于是 $AP_1 = (\lambda x_1, Ax_2, \cdots, Ax_n)$

$$= (x_1, x_2, \cdots, x_n)\begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$$

即

$$P_1^{-1}AP_1 = \begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$

由归纳假定，对于n–1阶矩阵A_1存在矩阵Q使得

$$Q^{-1}A_1Q = \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

记 $P_2=\begin{bmatrix}1 & 0^T\\ 0 & Q\end{bmatrix}$, $P=P_1P_2$

则有 $P^{-1}AP=(P_1P_2)^{-1}A(P_1P_2)=P_2^{-1}\left(P_1^{-1}AP_1\right)P_2$

$$=P_2^{-1}\begin{bmatrix}\lambda & b_{12} & \cdots & b_{1n}\\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}P_2$$

$$=\begin{bmatrix}\lambda_1 & * & \cdots & *\\ & \lambda_2 & \ddots & \vdots\\ & & \ddots & *\\ & & & \lambda_n\end{bmatrix}$$

#

Hamilton—Caylay定理 设 $A \in K^{n\times n}$ 其特征多项式

$$\varphi(\lambda)=\left|\lambda I - A\right| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

则 $\varphi(A) = A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI = 0.$

证明:A的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$ ，即

$$\varphi(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

必然存在可逆矩阵 $P_{n\times n}$ ，使得

$$P^{-1}AP=\begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix}$$

$$\varphi(P^{-1}AP) = (P^{-1}AP - \lambda_1 I)(P^{-1}AP - \lambda_2 I)\cdots(P^{-1}AP - \lambda_n I)$$

$$= \begin{bmatrix} 0 & * & \cdots & * \\ & \lambda_2-\lambda_1 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_1 \end{bmatrix} \begin{bmatrix} \lambda_1-\lambda_2 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_2 \end{bmatrix} \cdots (P^{-1}AP-\lambda_n I)$$

$$= \begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \ddots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & * \end{bmatrix} \begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ & \lambda_2-\lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_2-\lambda_3 \end{bmatrix} \cdots (P^{-1}AP-\lambda_n I)$$

$$= 0$$

即 $P^{-1}\varphi(A)P = 0 \quad \Rightarrow \varphi(A) = 0$

例: 设 $A=\begin{pmatrix}1&0&2\\0&-1&1\\0&1&0\end{pmatrix}$,求 $2A^8-3A^5+A^4+A^2-4I$

解：A的特征多项式 $\varphi(\lambda)=|\lambda I-A|=\lambda^3-2\lambda+1$

用$\varphi(\lambda)$去除 $2\lambda^8-3\lambda^5+\lambda^4+\lambda^2-4\lambda=g(\lambda)$,得

$$g(\lambda)=\varphi(\lambda)(2\lambda^5+4\lambda^3-5\lambda^2+9\lambda-14)+(24\lambda^2-37\lambda+10)$$

$\because\quad \varphi(A)=0,$

$\therefore\quad 2A^8-3A^5+A^4+A^2-4I=24A^2-37A+10I$

$$=\begin{pmatrix}-3&48&-26\\0&95&-61\\0&-61&34\end{pmatrix}$$

练习1：已知 $A\in P^{n\times n}$，λ 为A的一个特征值，则

（1） $kA \ (k\in P)$ 必有一个特征值为 $\underline{\quad k\lambda \quad}$；

（2） $A^m \ (m\in Z^+)$ 必有一个特征值为 $\underline{\quad \lambda^m \quad}$；

（3） A可逆时，A^{-1}必有一个特征值为 $\underline{\quad \lambda^{-1} \quad}$；

（4） A可逆时，A^*必有一个特征值为 $\underline{\quad \dfrac{|A|}{\lambda} \quad}$.

（5） $f(x)\in P[x]$，则 $f(A)$必有一个特征值为 $\underline{\quad f(\lambda) \quad}$.

练习2：已知3阶方阵A的特征值为：1、－1、2，

则矩阵 $\boldsymbol{B}=\boldsymbol{A}^3-2\boldsymbol{A}^2$ 的特征值为：$\underline{\quad -1,-3,0 \quad}$，

行列式 $|\boldsymbol{B}|=\underline{\quad 0 \quad}$.

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

由哈密尔顿—凯莱定理，$\forall A \in K^{n\times n}, \varphi(\lambda) = |\lambda I - A|$

是A的特征多项式，则 $\varphi(A) = 0.$

因此，对任定一个矩阵 $A \in K^{n\times n}$，总可以找到一个

多项式 $\varphi(x) \in P[x]$， 使 $\varphi(A) = 0.$ 此时，也称

多项式$\varphi(x)$以A为根.

本节讨论，以矩阵A为根的多项式的中次数最低的

那个与A的对角化之间的关系.

定义：设 $A\in K^{n\times n}$, 在数域K上的以A为根的多项式中，次数最低的首项系数为1的那个多项式，称为**A的最小多项式**.常记做 $m(\lambda)$

显然 $m(\lambda)$ 的次数不大于特征多项式 $\varphi(\lambda)$ 的次数

最小多项式的基本性质

定理：矩阵A的最小多项式$m(\lambda)$可以整除以A为根的任意首1多项式 $f(\lambda)$，且 $m(\lambda)$是唯一的.

证：（1）反证法。

假如 $m(\lambda)$ 不能整除 $f(\lambda)$，则有

$$f(\lambda)=m(\lambda)q(\lambda)+r(\lambda)$$

其中 $r(\lambda)$ 的次数小于 $m(\lambda)$ 的次数。于是

$$f(A)=m(A)q(A)+r(A)$$

由 $f(A)=0$ 和 $m(A)=0$ 可得 $r(A)=0$

这与 $m(\lambda)$是A的最小多项式相矛盾

设 $m_1(\lambda)$和$m_2(\lambda)$ 都是A的最小多项式，则

$$m_2(A)=0 \quad \Rightarrow m_1(\lambda)\,|\,m_2(\lambda)$$

$$m_1(A)=0 \quad \Rightarrow m_2(\lambda)\,|\,m_1(\lambda)$$

又 $m_1(x),m_2(x)$都是首1多项式,

故 $m_1(\lambda)=m_2(\lambda).$

定理：矩阵A的最小多项式$m(\lambda)$与其特征多项式 $\varphi(\lambda)$ 的零点相同（不计重数）。

证：显然 $\varphi(A)=0$，且 $m(\lambda)\,|\,\varphi(\lambda)$

所以 $m(\lambda)$ 的零点是 $\varphi(\lambda)$ 的零点

设 λ_0 是 $\varphi(\lambda)$ 的零点

$$Ax=\lambda_0 x(x\neq 0)\Rightarrow m(A)x=m(\lambda_0)x=0$$

所以 λ_0 也是 $m(\lambda)$ 的零点

推论：$m(\lambda)$ 一定含 $\varphi(\lambda)$ 的全部单因式

但：$m(\lambda)$不一定是$\varphi(\lambda)$的全部单因式的乘积

例：数量矩阵$k\mathbf{I}$的最小多项式是一次多项式 $x-k$;

特别地，单位矩阵的最小多项式是 $\lambda-1$ ；

零矩阵的最小多项式是 λ .

反之，若矩阵A的最小多项式是一次多项式，则

A一定是数量矩阵.

例2、求 $A=\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ 的最小多项式.

解：A的特征多项式为

$$\varphi(\lambda)=|\lambda I-A|=\begin{vmatrix}\lambda-1 & -1 & 0\\ 0 & \lambda-1 & 0\\ 0 & 0 & \lambda-1\end{vmatrix}=(\lambda-1)^3$$

又 $A-I\neq 0,$

$$(A-I)^2=A^2-2A+I$$

$$=\begin{pmatrix}1&2&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}2&2&0\\0&2&0\\0&0&2\end{pmatrix}+\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=0$$

$\therefore$ A的最小多项式为 $(\lambda-1)^2.$

定理：相似矩阵具有相同的最小多项式.

证：设矩阵A与B相似，$m_A(\lambda),m_B(\lambda)$分别为它们的最小多项式.

由A相似于B，存在可逆矩阵P， 使 $B=P^{-1}AP.$

从而 $m_A(B)=m_A(P^{-1}AP)=P^{-1}m_A(A)P=0$

$\therefore m_A(\lambda)$也以B为根， 从而 $m_B(\lambda)\mid m_A(\lambda).$

同理可得 $m_A(\lambda)\mid m_B(\lambda).$

又$m_A(\lambda),m_B(\lambda)$ 都是首1多项式，$\therefore m_A(\lambda)=m_B(\lambda).$

注：反之不然，即最小多项式相同的矩阵未必相似.

如：

$$A=\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&2\end{pmatrix},\quad B=\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&2&0\\0&0&0&2\end{pmatrix}$$

的最小多项式皆为$(x-1)^2(x-2)$，但A与B不相似.

$$\because\quad |\lambda I-A|=(\lambda-1)^3(\lambda-2),$$

$$|\lambda I-B|=(\lambda-1)^2(\lambda-2)^2$$

即 $|\lambda I-A|\neq|\lambda I-B|$.　　所以，A与B不相似.

最小多项式求法

定理：设n阶矩阵A特征多项式φ (λ)，特征矩阵的λI–A的全体n–1阶子式的最大公因式为d(λ)，则A最小多项式为

$$m(\lambda)=\frac{\det(\lambda I-A)}{d(\lambda)}$$

设 $A=\begin{pmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{pmatrix}$，求 $m(\lambda)$

解：$\lambda I-A=\begin{pmatrix} \lambda-3 & -3 & 2 \\ -1 & \lambda-5 & -2 \\ -1 & 3 & \lambda \end{pmatrix}$

$$\varphi(\lambda)=|\lambda I-A|=(\lambda-2)^2(\lambda-4)$$

$M_{11}=\lambda^2-5\lambda+6$ $\quad M_{21}=3\lambda-6$ $\quad M_{31}=2\lambda-4$

$M_{12}=\lambda-2$ $\quad M_{22}=\lambda^2-3\lambda+2$ $\quad M_{32}=2\lambda-4$

$M_{13}=-\lambda+2$ $\quad M_{23}=-3(\lambda-2)$ $\quad M_{33}=(\lambda-2)(\lambda-6)$

$d(\lambda)=\lambda-2$

$$\therefore m(\lambda)=\frac{\varphi(\lambda)}{d(\lambda)}=(\lambda-2)(\lambda-4)$$

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

一、可对角化的概念

定义1：设T是n维线性空间V的一个线性变换，如果存在V的一个基，使T在这组基下的矩阵为对角矩阵，则称线性变换T可对角化.

定义2：矩阵A是数域K上的一个n阶方阵. 如果存在一个K上的n阶可逆矩阵P，使$P^{-1}AP$为对角矩阵，则称矩阵A可对角化.

二、可对角化的条件

定理：设T为n维线性空间V的一个线性变换，

则T可对角化 $\Leftrightarrow$ T有n个线性无关的特征向量.

证：设T在基$x_1, x_2, \cdots x_n$下的矩阵为对角矩阵

$$diag\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)$$

则有$Tx_i = \lambda_i x_i, \ \ i = 1, 2, \cdots n.$

$\therefore x_1, x_2, \cdots x_n$就是$T$的$n$个线性无关的特征向量.

反之，若T有n个线性无关的特征向量$y_1, y_2, \cdots, y_n$，那么就取$y_1, y_2, \cdots, y_n$为基，则在这组基下T的矩阵是对角矩阵.

定理： 设 T 为 n 维线性空间V的一个线性变换，

如果 $x_1, x_2, \cdots x_k$ 分别是 T 的属于互不相同的特征值

$\lambda_1, \lambda_2, \cdots \lambda_k$ 的特征向量，则 $x_1, x_2, \cdots x_k$ 线性无关．

证：对 k 作数学归纳法．

当 $k=1$ 时，$\because x_1 \neq 0$，$\therefore x_1$ 线性无关．命题成立．

假设对于 $k-1$ 来说，结论成立．现设 $\lambda_1, \lambda_2, \cdots \lambda_k$ 为

T 的互不相同的特征值，x_i 是属于 λ_i 的特征向量，

即 $\quad Tx_i = \lambda_i x_i, \quad i = 1, 2, \cdots, n.$

设 $a_1 x_1 + a_2 x_2 + \cdots + a_k x_k = 0, \quad a_i \in K \qquad ①$

以 λ_k 乘①式的两端，得

$$a_1\lambda_k x_1 + a_2\lambda_k x_2 + \cdots a_k\lambda_k x_k = 0. \qquad ②$$

又对①式两端施行线性变换 T，得

$$a_1\lambda_1 x_1 + a_2\lambda_2 x_2 + \cdots a_k\lambda_k x_k = 0. \qquad ③$$

③式减②式得

$$a_1(\lambda_1 - \lambda_k)x_1 + a_2(\lambda_2 - \lambda_k)x_2 + \cdots a_{k-1}(\lambda_{k-1} - \lambda_k)x_{k-1} = 0$$

由归纳假设,$x_1, x_2, \cdots, x_{k-1}$线性无关，所以

$$a_i(\lambda_i - \lambda_k) = 0,\ \ i = 1,2,\cdots,k-1.$$

但 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 互不相同，所以 $a_1 = a_2 = \cdots = a_{k-1} = 0.$

将之代入①，得 $a_k x_k = 0.$

$\because \quad x_k \neq 0, \qquad \therefore \quad a_k = 0$

故 $x_1, x_2, \cdots, x_k$ 线性无关.

定理25： n阶矩阵与对角矩阵相似的充要条件：A有n个线性无关的特征向量，或A有完备的特征向量系

对角化的一般方法

设 T 为维线性空间V的一个线性变换，$x_1, x_2, \cdots, x_n$ 为V的一组基， T 在这组基下的矩阵为A.

步骤:

1° 求出矩阵A的全部特征值 $\lambda_1, \lambda_2, \cdots, \lambda_k$.

2° 对每一个特征值 λ_i ，求出齐次线性方程组

$$\left(\lambda_i I - A\right) X = 0, \qquad i = 1.2.\cdots k$$

的一个基础解系（此即 T 的属于 λ_i 的全部线性无关的特征向量在基 $x_1, x_2, \cdots, x_n$ 下的坐标）.

3° 若全部基础解系所含向量个数之和等于n，则

T 有n个线性无关的特征向量 $y_1, y_2, \cdots, y_n$，从而 T

（或矩阵A）可对角化．以这些解向量为列，作一个

n阶方阵C，则C可逆，$C^{-1}AC$ 是对角矩阵．而且

C就是基 $x_1, x_2, \cdots, x_n$到基 $y_1, y_2, \cdots, y_n$的过渡矩阵．

例：设复数域上线性空间V的线性变换 T在某组基

x_1, x_2, x_3 下的矩阵为

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

问T是否可对角化. 在可对角化的情况下，写出

基变换的过渡矩阵.

解：A的特征多项式为

$$\left|\lambda I - A\right| = \begin{vmatrix} \lambda & 0 & -1 \\ 0 & \lambda - 1 & 0 \\ -1 & 0 & \lambda \end{vmatrix} = \left(\lambda - 1\right)^2 \left(\lambda + 1\right)$$

得A的特征值是1、1、－1.

解齐次线性方程组 $\left(1 \cdot I - A\right) X = 0$，得 $x_1 = x_3$

故其基础解系为：(1,0,1),(0,1,0)

所以， $y_1 = x_1 + x_3$， $y_2 = x_2$

是T 的属于特征值1的两个线性无关的特征向量.

再解齐次线性方程组 $(-1\cdot I-A)X=0$，得 $\begin{cases} x_1=-x_3 \\ x_2=0 \end{cases}$

故其基础解系为：$(1,0,-1)$

所以， $y_3=x_1-x_3$

是T的属于特征值－1的线性无关的特征向量.

y_1,y_2,y_3 线性无关，故T可对角化，且

T 在基 y_1,y_2,y_3下的矩阵为对角矩阵

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix};$$

$$(y_1, y_2, y_3) = (x_1, x_2, x_3)\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

即基x_1, x_2, x_3到 y_1, y_2, y_3的过渡矩阵为

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix},$$

$$C^{-1}AC = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

例： 问A是否可对角化？若可，求可逆矩阵C，使

$C^{-1}AC$ 为以角矩阵. 这里 $A=\begin{pmatrix} 3 & 2 & -1 \\ -2 & -2 & 2 \\ 3 & 6 & -1 \end{pmatrix}$

解: A的特征多项式为

$$|\lambda I-A|=\begin{vmatrix} \lambda-3 & -2 & 1 \\ 2 & \lambda+2 & -2 \\ -3 & -6 & \lambda+1 \end{vmatrix}$$

$$=\lambda^3-12\lambda+16=(\lambda-2)^2(\lambda+4)$$

得A的特征值是2、2、-4 .

对于特征值2，求出齐次线性方程组

$$\begin{pmatrix} -1 & -2 & 1 \\ 2 & 4 & -2 \\ -3 & -6 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的一个基础解系：（－2、1、0），（1、0、1）

对于特征值－4，求出齐次方程组

$$\begin{pmatrix} -7 & -2 & 1 \\ 2 & -2 & -2 \\ -3 & -6 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的一个基础解系： $(\frac{1}{3}, -\frac{2}{3}, 1)$

所以A可对角化.

令 $$C=\begin{pmatrix} -2 & 1 & \frac{1}{3} \\ 1 & 0 & -\frac{2}{3} \\ 0 & 1 & 1 \end{pmatrix}$$

则 $$C^{-1}AC=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

例： 在 $P[x]_n(n>1)$ 中，求微分变换D的特征多项式.并证明：D在任何一组基下的矩阵都不可能是对角矩阵（即D不可对角化）.

解：在 $P[x]_n$中取一组基：$1,\ x,\ \dfrac{x^2}{2!},\ \cdots,\ \dfrac{x^{n-1}}{(n-1)!}$

则D在这组基下的矩阵为

$$A=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

于是

$$|\lambda I-A|=\begin{pmatrix}\lambda & -1 & 0 & \cdots & 0\\ 0 & \lambda & -1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & -1\\ 0 & 0 & 0 & \cdots & \lambda\end{pmatrix}=\lambda^n$$

$\therefore$ **D**的特征值为0（n重）.

又由于对应特征值**0**的齐次线性方程组 $-AX=0$

的系数矩阵的秩为n－1，从而方程组的基础解系

只含有一个向量，它小于$P[x]_n$的维数n（>1）.

故**D**不可对角化 .

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

1、定义

设 T 是数域K上线性空间V的线性变换，V_1是V的子空间，若 $\forall x \in V_1$, 有 $T(x) \in V_1$

则称V_1是 T 的不变子空间

注：

V的平凡子空间（V及零子空间）对于V的任意一个变换 T 来说，都是不变子空间.

不变子空间的简单性质

1）两个不变子空间的交与和仍是不变子空间.

2）设 $V_1 = L(x_1, x_2, \cdots x_s)$， 则$V_1$是不变子空间

$\Leftrightarrow T(x_1), T(x_2), \cdots, T(x_s) \in V_1$.

证：" $\Rightarrow$ " 显然成立.

" $\Leftarrow$ " 任取 $x \in V_1$，设 $x = k_1 x_1 + k_2 x_2 + \cdots + k_s x_s$,

则 $T(x) = k_1 T(x_1) + k_2 T(x_2) + \cdots + k_s T(x_s)$.

由于 $T(x_1), T(x_2), \cdots, T(x_s) \in V_1$， $\therefore$ $T(x) \in V_1$.

故V_1为T 的不变子空间.

一些重要不变子空间

1) 线性变换T的值域 $R(T)$与核 $N(T)$都是T的不变子空间.

证：$\because$ $R(T)=\{T(x)|x\in V\}\subseteq V$,

$\therefore \forall x\in R(T)$, 有$T(x)\in R(T)$.

故 $R(T)$为T的不变子空间.

又任取 $x\in N(T)$, 有 $T(x)=0\in N(T)$.

$\therefore$ $N(T)$也为T的不变子空间.

2） 任何子空间都是数乘变换K的不变子空间.

$$(\because \forall x \in V_1,\ \mathrm{K}x = kx \in V_1)$$

3） 线性变换T的特征子空间 V_{λ_0} 是T的不变子空间.

$$(\because \forall x \in V_{\lambda o},\ \text{有} T(x) = \lambda_o x \in V_{\lambda o}.)$$

4） 由T的特征向量生成的子空间是T的不变子空间.

证：设 $x_1, x_2, \cdots, x_s$ 是T的分别属于特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 的特征向量．任取 $x \in L(x_1, x_2, \cdots, x_s)$,

设 $x = k_1 x_1 + k_2 x_2 + \cdots + k_s x_s$, 则

$$T(x) = k_1\lambda_1 x_1 + k_2\lambda_2 x_2 + \cdots + k_s\lambda_s x_s \in L(x_1, x_2, \cdots, x_s)$$

$\therefore$ $L(x_1, x_2, \cdots, x_s)$ 为T的不变子空间.

注：

特别地，由 T 的一个特征向量生成的子空间是一个一维不变子空间. 反过来，一个一维不变子空间必可看成是 T 的一个特征向量生成的子空间.

事实上，若 $V_1 = L(x) = \{kx \mid k \in K, x \neq 0\}$.

则 x 为 $L(x)$ 的一组基.　因为V_1为不变子空间，

$\therefore$ $T(x) \in V_1$，即必存在 $\lambda \in K$，使 $T(x) = \lambda x$.

$\therefore$ x 是T 的特征向量.

定理27：设T是线性空间 V^n 的线性变换，且V^n

可分解为s个T的不变子空间的直和

$$V^n = V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

在每个不变子空间 V_i 中取基

$$x_{i1}, x_{i2}, \cdots, x_{in_i} (i = 1, 2, \cdots, s) \qquad (*)$$

将其合并作为V^n 的基，则T在该基下的矩阵为

$$A = \mathrm{diag}(A_1, A_2, \cdots, A_s)$$

其中 A_i $(i = 1, 2, \cdots, s)$ 是T在V_i 的基(*)下的矩阵

作业

- **P79：16、17、18**