第四讲 矩阵的对角化

	元素	坐标向量
加法	元素加法	坐标向量的加法
数乘	数与元素“乘”	数与坐标向量相乘
线性变换及其作用	对应关系	矩阵与坐标列向量的乘积

对任何线性空间，给定基后，我们对元素进行线性变换或线性运算时，只需用元素的坐标向量以及线性变换的矩阵即可，因此，在后面的内容中着重研究矩阵和向量。

对角矩阵的形式比较简单，处理起来较方便，比如求解矩阵方程$\boldsymbol{Ax=b}$时，将矩阵$\boldsymbol{A}$对角化后很容易得到方程的解。对角化的过程实际上是一个去耦的过程。以前我们学习过相似变化对角化。那么，一个方阵是否总可以通过相似变化将其对角化呢？或者对角化需要什么样的条件呢？如果不能对角化，我们还可以做哪些处理使问题变得简单呢？

一、特征值与特征向量

1. 定义1: 设$\boldsymbol{T}$是线性空间$\boldsymbol{V}$的线性变换，如果对于数域$\boldsymbol{K}$中的一个数λ，存在一非零向量$\boldsymbol{x}$，使得$\boldsymbol{T\,x} = \lambda\,\boldsymbol{x}$,则称$\lambda$为$\boldsymbol{T}$的一个特征值，$\boldsymbol{x}$为$\boldsymbol{T}$的属于特征值$\lambda$的特征向量。

- 特征向量不唯一

根据线性变换与矩阵的关系，设 $x_1, x_2 \cdots, x_n$ 是V^n的一组基，线性变换在这组基下的矩阵是A， λ 为T的特征值，它的一个特征向量在这组基下的坐标为 $(\xi_1,\xi_2\cdots,\xi_n)^T$ ，则Tx的坐标是：

$$A\begin{bmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{bmatrix}=Ax$$

λx的坐标是

$$\lambda\begin{bmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{bmatrix}=\lambda x$$

则由$Tx=\lambda x$有： $Ax=\lambda x$ $\quad\therefore (\lambda I-A)x=0$

$\because$ x非零 $\quad\therefore \det(\lambda I-A)=0$

定义2： 对阶方阵A，若存在数λ，及非零向量（列向量）x，使得$A x = \lambda x$,则称λ为A的特征值， x为A的属于特征值λ的特征向量。

- 特征向量不唯一
- 特征向量非零
- 称$(\lambda I - A)$ 为A的特征矩阵，称$\det(\lambda I - A)$ 为A的特征多项式

例 1. $A=\begin{bmatrix}1&2&2\\2&1&2\\2&2&1\end{bmatrix}$，求其特征值和特征向量。

解：$\det(\lambda I-A)=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&-2\\-2&-2&\lambda-1\end{vmatrix}=(\lambda+1)^2(\lambda-5)=0$ ∴ $\lambda_1=\lambda_2=-1$，$\lambda_3=5$

属于特征值$\lambda=-1$的特征向量满足$(-I-A)x=0$

$$\begin{bmatrix}2&2&2\\2&2&2\\2&2&2\end{bmatrix}\begin{bmatrix}\xi_1\\\xi_2\\\xi_3\end{bmatrix}=0\longrightarrow \xi_1+\xi_2+\xi_3=0\longrightarrow\begin{cases}\xi_1=\xi_1\\\xi_2=\xi_2\\\xi_3=-\xi_1-\xi_2\end{cases}$$

可取基础解系为 $x_1=(1\quad 0\quad -1)^T$， $x_2=(0\quad 1\quad -1)^T$，对应$\lambda=-1$的全部特征向量为：$k_1x_1+k_2x_2$

属于$\lambda=5$的特征向量满足$(5I-A)x=0$

$$\begin{bmatrix}4&-2&-2\\-2&4&-2\\-2&-2&4\end{bmatrix}\begin{bmatrix}\xi_1\\\xi_2\\\xi_3\end{bmatrix}=0\longrightarrow \xi_1=\xi_2=\xi_3$$

可取基础解系为$x_3=(1\quad 1\quad 1)^T$，对应$\lambda=5$的全部特征向量为：k_3x_3

例2. 平面上全部向量构成实数域上的一个二维线性空间。旋转变换在直角坐标系下的矩阵为

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

其特征多项式为：$\lambda^2 - 2\lambda\cos\theta + 1$

当 $\theta \neq k\pi$ 时，没有特征值。

从几何上看，这个结论是很明显的。

2. 矩阵的迹与行列式

矩阵的特征多项式det ($\lambda I - A$)的展开式中，有一项是主对角线上元素的乘积：

$$(\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

展开式中的其余各项，至多包含(n-2)个主对角线的元素，它们对λ的次数最多是(n-2)，因此，特征多项式中含λ的n次与n-1次的项只能在主对角线上元素的连乘中出现，它们是：

$$\lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1}$$

定义矩阵的迹：所有对角元素之和，即 $\mathrm{tr}A = \sum_{i=1}^{n} a_{ii}$

可以证明： $\det A = \prod_{i=1}^{n} \lambda_i$ $\quad$ $\mathrm{tr}A = \sum_{i=1}^{n} \lambda_i$

3.定理：

定理1：相似矩阵有相同的特征多项式，因此，有相同的特征值、迹和行列式。

证明：设A~B，即存在可逆矩阵P，使得：

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

于是： $|\lambda I - B| = |\lambda I - P^{-1}AP| = |P^{-1}(\lambda I - A)P|$

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

定理2：设 A、B 分别为 $m \times n$ 和 $n \times m$ 阶矩阵，则

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

sylvster 定理：设 A、B 分别为 $m \times n$ 和 $n \times m$ 阶矩阵，则

$$\det(\lambda I_m - AB) = \lambda^{m-n} \det(\lambda I_n - BA)$$

即：AB 与 BA 的非零特征值相同。

二、矩阵对角化的充要条件

定理1：n阶方阵A可通过相似变换对角化的充要条件是它具有n个线性无关的特征向量。

[证明]充分性：已知A具有n个线性无关的特征向量$x_1, x_2, \cdots, x_n$，则

$$Ax_i = \lambda_i x_i \qquad i = 1, 2, \cdots, n$$

$$A[x_1 \ \ x_2 \ \ \cdots \ \ x_n] = [\lambda_1 x_1 \ \ \lambda_2 x_2 \ \ \cdots \ \ \lambda_n x_n]$$

$$= [x_1 \ \ x_2 \ \ \cdots \ \ x_n]\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

$x_1, x_2, \cdots, x_n$线性无关，故$P = [x_1 \ \ x_2 \ \ \cdots \ \ x_n]$为满秩矩阵，

令$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$，则有 $AP = P\Lambda$， 即$P^{-1}AP = \Lambda$

必要性：已知存在可逆方阵P，使$P^{-1}AP=\Lambda=\begin{bmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{bmatrix}$

将P写成列向量$P=[P_1\ \ P_2\ \ \cdots\ \ P_n]$，$P_n$为$n$维列向量。

$$[AP_1\ \ AP_2\ \ \cdots\ \ AP_n]=[\lambda_1P_1\ \ \lambda_2P_2\ \ \cdots\ \ \lambda_nP_n]$$

可见，λ_i为A的特征值，P_i为A的特征向量。所以，A具有n个线性无关的特征向量。

定理2：属于不同特征值的特征向量是线性无关的。

推论：n阶方阵有n个互异的特征值，则必可对角化。

三、内积空间

1. Euclid空间： 设V是实线性空间（$k\in R$），对于V中任意两个元素x，y均按某一规则定义一个实数，记为(x, y),若它满足：

(1)交换律：$(x, y)= (y, x)$

(2)分配律：$(x, y+z)= (x, y)+ (x, z)$

(3)齐次性：$(kx, y)= k(x, y)$

(4)非负性：$(x, x) \geq 0$，当且仅当$x=0$时，$(x, x) =0$

则称(x, y)为x与y的内积。定义了内积的实线性空间称为Euclid空间。

对于一个给定的线性空间，可以定义多种内积，较典型的，如三维向量空间的数量积就满足以上四条性质，构成内积

例 2. 在线性空间$\mathbf{R}^n$中，对于向量 $x=(\xi_1 \quad \xi_2 \quad \cdots \quad \xi_n)^T$ $y=(\eta_1 \quad \eta_2 \quad \cdots \quad \eta_n)^T$ 可定义

$$(x,y)=\sum_{i=1}^{n}\xi_i\eta_i$$

则$\mathbf{R}^n$就成为一个Euclid空间。

例 3. 在闭区间[a,b]上的所有连续函数所构成的实线性空间C(*a*,*b*)中，对其中任意两个函数$f(x)$, $g(x)$定义：

$$(f(x),g(x))=\int_a^b f(x)g(x)dx$$

可以证明它满足内积的四个条件，所以C(*a*,*b*)是Euclid空间。

2.酉空间： 设V是复线性空间（$k\in C$），对于V中任意两个元素x，y均按某一规则定义一个复数，记为(x, y),若它满足：

（1）交换律 $(x,y)=\overline{(y,x)}$

（2）分配律 $(x,y+z)=(x,y)+(x,z)$

（3）齐次性 $(kx,y)=k(x,y)$ or $(x,ky)=\bar{k}(x,y)$

（4）非负性 $(x,x)\geq 0$，当且仅当$x=0$时，$(x,x)=0$

则称(x, y)为x与y的内积。定义了内积的复线性空间称为酉空间。

例 4. 在线性空间C^n中，对于任意两个向量 $x=(\xi_1 \quad \xi_2 \quad \cdots \quad \xi_n)^T$

$y=(\eta_1 \quad \eta_2 \quad \cdots \quad \eta_n)^T$ 可定义

$$(x,y)=\sum_{i=1}^{n}\xi_i\overline{\eta_i}$$

则C^n就成为一个酉空间。

3. 度量矩阵

设$\{x_1, x_2, \cdots, x_n\}$是欧氏空间$V^n$的一个基，$\forall \boldsymbol{x}, \boldsymbol{y} \in V^n$，

$\boldsymbol{x} = \xi_1 \boldsymbol{x_1} + \xi_2 \boldsymbol{x_2} + \cdots + \xi_n \boldsymbol{x_n}$，$\boldsymbol{y} = \eta_1 \boldsymbol{x_1} + \eta_2 \boldsymbol{x_2} + \cdots + \eta_n \boldsymbol{x_n}$，则

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i,j=1}^{n} \xi_i \eta_j (\boldsymbol{x}_i, \boldsymbol{x}_j) = (\xi_1, \xi_2, \cdots, \xi_n) \boldsymbol{A} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}$$

其中，
$$\boldsymbol{A} = \begin{bmatrix} (\boldsymbol{x}_1, \boldsymbol{x}_1) & \cdots & (\boldsymbol{x}_1, \boldsymbol{x}_n) \\ (\boldsymbol{x}_2, \boldsymbol{x}_1) & \cdots & (\boldsymbol{x}_2, \boldsymbol{x}_n) \\ \vdots & & \vdots \\ (\boldsymbol{x}_n, \boldsymbol{x}_1) & \cdots & (\boldsymbol{x}_n, \boldsymbol{x}_n) \end{bmatrix}$$

称 $\boldsymbol{A}$ 为V^n对于基$\{x_1, x_2, \cdots, x_n\}$的度量矩阵（或 **Gram** 矩阵）。

$\boldsymbol{A}$ 为对称正定矩阵。

4. 正交性： 若$(x,y)=0$，则称x与y正交。称$|x|=\sqrt{(x,x)}$为向量x的模（或长度）。

两个非零向量x与y之间的夹角定义为：$\alpha=\arccos\left(\dfrac{(x,y)}{|x|\cdot|y|}\right)$

5. Gram-Schmidt正交化手续

设$x_1,x_2,\cdots,x_n$为一组线性无关的元素或向量，可以进行如下正交归一化操作（正交规范化或正交单位化）：

$1°\quad y_1=\dfrac{x_1}{|x_1|}$

$2°\quad x_2'=x_2+k_{21}y_1$ 选择合适的k_{21}使x_2'与y_1正交：

$$(x_2',y_1)=(x_2,y_1)+k_{21}(y_1,y_1)=0$$

$$k_{21}=-(x_2,y_1)$$

$$y_2=\frac{x_2'}{|x_2'|}$$

$3°$ $x_3' = x_3 + k_{31}y_1 + k_{32}y_2$ 选择k_{31}、k_{32}使x_3'与y_1和y_2均正交

$$(x_3', y_1) = (x_3', y_2) = 0$$

$$(x_3', y_1) = (x_3, y_1) + k_{31} = 0 \quad \to k_{31} = -\left(x_3, y_1\right)$$

$$(x_3', y_2) = (x_3, y_2) + k_{32} = 0 \quad \to k_{32} = -\left(x_3, y_2\right)$$

$$y_3 = \frac{x_3'}{|x_3'|}$$

一般地， $x_i' = x_i + \sum_{j=1}^{i-1} k_{ij} y_j \quad i = 1,2,\cdots,n$

$$k_{ij} = -\left(x_i, y_j\right)$$

$$y_i = \frac{x_i'}{|x_i'|}$$

$y_1, y_2, \cdots, y_n$成为一组正交归一化向量：$(y_i, y_j) = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

若$x_1, x_2, \cdots, x_n$为一组基元素，则$y_1, y_2, \cdots, y_n$成为标准正交基。

在标准正交基下，向量 x 的坐标可用内积表示出来：

$$x=(x_1,x)x_1+(x_2,x)x_2+\cdots+(x_n,x)x_n$$

事实上，设 $x=\xi_1x_1+\xi_2x_2+\cdots+\xi_nx_n$，以 $x_i \quad (i=1,2,\cdots,n)$
与上式两端作内积，便得：

$$\xi_i=(x_i,x) \quad (i=1,2,\cdots,n)$$

作业：P106-107 1(1)(2),2,4,5,10,11