第四讲 矩阵的对角化

基

	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 $Ax=b$ 时，将矩阵 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：对 m 阶方阵 A，若存在数 λ，及非零向量（列向量）x，使得 $Ax=\lambda x$,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
- $(\lambda I-A)x=0$ 有非零解，则 $\det(\lambda I-A)=0$，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] $$\det(\lambda I-A)=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&-2\\-2&-2&\lambda-1\end{vmatrix}=0$$

$$(\lambda+1)^2(\lambda-5)=0$$

$$\lambda_1=\lambda_2=-1 \qquad \lambda_3=5$$

属于特征值$\lambda=-1$的特征向量 $\qquad (-I-A)x=0$

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

可取基础解系为 $x_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ $x_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$

属于$\lambda = 5$的特征向量 $(5I - A)x = 0$

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

可取基础解系为 $x_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

2. 矩阵的迹与行列式

$$\mathrm{tr}A = \sum_{i=1}^{n} a_{ii}$$ 所有对角元素之和

$$\det A = \prod_{i=1}^{n} \lambda_i \qquad \mathrm{tr}A = \sum_{i=1}^{n} \lambda_i$$

3. 两个定理

（1）设 A、B 分别为 $m \times n$ 和 $n \times m$ 阶矩阵，则

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

（2）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 的特征值只差零特征值的个数，非零特征值相同。

二、 矩阵对角化的充要条件

定理：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

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

$$A\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \end{bmatrix}$$

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

$x_1, x_2, \cdots, x_n$线性无关，故$P = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$为满秩矩阵，

令$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$，则有

$$AP = P\Lambda$$

$$P^{-1}AP=\Lambda$$

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$P=\begin{bmatrix}P_1 & P_2 & \cdots & P_n\end{bmatrix}$，$P_n$为$n$维列向量

$$\begin{bmatrix}AP_1 & AP_2 & \cdots & AP_n\end{bmatrix}=\begin{bmatrix}\lambda_1P_1 & \lambda_2P_2 & \cdots & \lambda_nP_n\end{bmatrix}$$

可见，$\boldsymbol{\lambda}_i$为A的特征值，P_i为A的特征向量，

$\therefore$ A具有n个线性无关的特征向量。

推论：n阶方阵有n个互异的特征值，则必可对角化。（充分条件）

三、 内积空间

1. Euclid 空间

设V是实线性空间（$k\in\mathbf{R}$），对于V中任何两个元素x、y均按某一规则

存在一个实数与之对应，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)齐次律 $(kx,y)=k(x,y)$

（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。以n维向量空间为例：

$$x=\begin{bmatrix}\xi_1 & \xi_2 & \cdots & \xi_n\end{bmatrix}^T,\quad y=\begin{bmatrix}\eta_1 & \eta_2 & \cdots & \eta_n\end{bmatrix}^T$$

可定义内积$\left(x,y\right)=\sum_{i=1}^{n} w_i\xi_i\eta_i$，它满足内积的四条性质：

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

（2）$\left(x,y+z\right)=\sum_{i=1}^{n} w_i\xi_i(\eta_i+\varsigma_i)=\sum_{i=1}^{n} w_i\xi_i\eta_i+\sum_{i=1}^{n} w_i\xi_i\varsigma_i=\left(x,y\right)+\left(x,z\right)$

（3）$\left(kx,y\right)=\sum_{i=1}^{n}w_i(k\xi_i)\eta_i=k\sum_{i=1}^{n}w_i\xi_i\eta_i=k\left(x,y\right)$

（4）$\left(x,x\right)=\sum_{i=1}^{n}w_i\xi_i^2\geq 0$　当且仅当 $x_i=0$ 时，$\left(x,x\right)=0$

该内积可写为：$\left(x,y\right)=x^{\mathrm{T}}Wy$，其中 $W=\begin{bmatrix} w_1 & & & 0 \\ & w_2 & & \\ & & \ddots & \\ 0 & & & w_n \end{bmatrix}$

更一般的，对实对称正定矩阵 A，$\left(x,y\right)=x^{\mathrm{T}}Ay$ 也满足内积的定义。

（正定：（1）特征值全为正（2）各阶顺序主子式大于 0）

2. 酉空间：

设V是复线性空间（$k \in \mathbf{C}$），对于V中任何两个元素x、y均按某一规则

存在一个复数与之对应，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）齐次律 $(kx, y) = k(x, y)$ 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。

以n维向量空间为例，A为厄米（$A^{\mathrm{H}} = A$）正定（$x^{\mathrm{H}} A x > 0$）矩阵，

$$(x,y)=x^{\mathrm{T}}A\overline{y}=\sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i a_{ij}\overline{\eta}_j$$

较常见的比如$A=\operatorname{diag}[w_1 \quad w_2 \quad \cdots \quad w_n]$，$w_i>0$

最简单：实　$(x,y)=x^{\mathrm{T}}y$

复　$(x,y)=x^{\mathrm{T}}\overline{y}$

3. 正交性：若$(x,y)=0$，则称x与y正交。

x与y的夹角：$\cos\alpha=\dfrac{(x,y)}{|x||y|}$，$\alpha$称为$x$与$y$的夹角。

4. 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）：

$1^\circ \quad y_1 = \dfrac{x_1}{|x_1|}$

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

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

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

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

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

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

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

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

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

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

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

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

$y_1, y_2, \cdots, y_n$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$x = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n$，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