第二章 矩阵的特征值与Jordan标准形

问: 给定的线性变换, 如何寻找一组基, 使的此线性变换在其下的矩阵具有最简单的形式?

本章目录

2.1 特征值与特征向量

2.1.1 线性变换的特征值和特征向量

2.1.2 特征子空间

本节总假设V是数域P上的n维线性空间, T是V上的线性变换.

2.1.1 线性变换的特征值与特征向量

设T在V的基$\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$下的矩阵为

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \ \lambda_i \in P,$$

则有

$$T\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i, \quad i = 1,2,\cdots,n.$$

定义1 若有P中的数λ_0和V中的**非零**向量$\boldsymbol{x}$满足

$$T\boldsymbol{x} = \lambda_0 \boldsymbol{x}$$

就称

①λ_0为T的一个**特征值**,

② $\boldsymbol{x}$为T的对应于特征值λ_0的**特征向量**.

注 若T在基$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$下的矩阵为对角阵$\boldsymbol{\Lambda}$, 则

① $\boldsymbol{\Lambda}$的每个对角元λ_i都是T的特征值,

② $\boldsymbol{x}_i$是T的对应于λ_i的特征向量.

定理1 设T 在基$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$下的矩阵为$\boldsymbol{A}$, 则

① $\boldsymbol{A}$特征值λ_0就是线性变换T的特征值,

② 如果$\boldsymbol{\xi}$是$\boldsymbol{A}$对应于λ的特征向量, 那么

$$\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{\xi}$$

是T 对应于λ_0的特征向量.

注 ① **线性变换的特征值就是矩阵的特征值**;

② **线性变换的特征向量的坐标是矩阵的特征向量**.

① n阶矩阵有n个特征值(重根按重数计算),因此n维空间上的线性变换有n个特征值.

② T 的特征值由T本身决定, 与基的选取无关.

(同一个线性变换在不同的基下的矩阵相似,而相似矩阵的特征值相同.)

③ 线性变换的特征值和特征向量的其它性质也类似于矩阵, 例如

1) 一个特征向量只能对应一个特征值,

2) 不同特征值的特征向量线性无关.

例1 设线性变换T在基$\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3$下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$$

求T的特征值和特征向量.

解 ① 特征值2,1,1.

② 对应2的特征向量为$c\boldsymbol{x}_3$, 其中$c\neq 0$.

③ 对应1的特征向量为$c(\boldsymbol{x}_1+2\boldsymbol{x}_2-\boldsymbol{x}_3)$, 其中$c\neq 0$.

2.1.2 特征子空间

定义2 设λ_0是T的特征值，易见

$$V_{\lambda_0}=\{\boldsymbol{x}\in V|T\boldsymbol{x}=\lambda_0\boldsymbol{x}\}$$

是V的子空间, 称为T属于λ_0的**特征子空间**.

① 称$\dim V_{\lambda_0}$为特征值λ_0的**几何重数**;

② 特征值λ_0作为特征多项式的根, 其重数称为λ_0的**代数重数**.

③ 矩阵的特征子空间: $(\lambda_0\boldsymbol{E}-\boldsymbol{A})\boldsymbol{\xi}=0$的解空间, **几何重数** $=n-R(\lambda_0\boldsymbol{E}-\boldsymbol{A})$.

定理2 特征值的几何重数不超过代数重数.

证明 ① 设$\dim V_{\lambda_0}=r$, 取V_{λ_0}的基$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_r$,
扩充为V的基$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_r,\boldsymbol{x}_{r+1},\cdots,\boldsymbol{x}_n$.

② T在$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_r,\boldsymbol{x}_{r+1},\cdots,\boldsymbol{x}_n$下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}\lambda_0\boldsymbol{E}_r & \boldsymbol{A}_{12}\\ O & \boldsymbol{A}_{22}\end{pmatrix}$$

③ 计算$|\lambda\boldsymbol{E}-\boldsymbol{A}|$得

$$(\lambda-\lambda_0)^r|\lambda\boldsymbol{E}-\boldsymbol{A}_{22}|,$$

因此λ_0的代数重数$\geq r$.

2.2 矩阵相似于对角阵的条件

2.2.1 线性变换的对角矩阵表示

2.2.2 Schur定理

2.2.3 正规矩阵

2.2.1 线性变换的对角矩阵表示

本小节总设T是n维线性空间V上的线性变换.

问 是否有V的基使T在其下的矩阵为**对角阵**?

设T在V的基$\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$下的矩阵为

$$\Lambda=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

则有$T\boldsymbol{y}_i=\lambda_i\boldsymbol{y}_i,\ i=1,2,\cdots,n.$

定理1 T 在某组基下的矩阵为对角阵的充要条件是T 有**n个线性无关**的**特征向量**.

推论 若T 有n个**互不相同**的特征值, 则T 在某组基下的矩阵为对角阵.

注 以上结论和矩阵的**相似对角化**类似.

设

$$T(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{A},$$

$$T(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n)=(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n)\boldsymbol{\Lambda},$$

$$(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{Q},$$

则$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda}$.

因此,

T在某组基下的矩阵为对角阵

$\Leftrightarrow$ $\boldsymbol{A}$相似于对角阵, 即$\boldsymbol{A}$可相似对角化.

设$\lambda_1,\cdots,\lambda_h$是$\boldsymbol{A}$所有互不相同的特征值，它们的几何重数为$k_i$，则$\boldsymbol{A}$有

$$k_1+\cdots+k_h$$

个线性无关的特征向量.

再设$|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_h)^{m_h}$，其中$m_i$是$\lambda_i$的代数重数，则$m_1+\cdots+m_h=n$.

从而$\boldsymbol{A}$可相似对角化$\Leftrightarrow$

$$\begin{aligned} n &= k_1+\cdots+k_h \\ &\le m_1+\cdots+m_h=n, \end{aligned}$$

即$k_i=m_i, i=1,\cdots,h$.

定理2 T 在某组基下的矩阵为对角阵 $\Leftrightarrow$ 其每个特征值的几何重数都等于代数重数.

注 仅需检验代数重数>1的特征值的几何重数.

例1 判断下列矩阵是否可相似对角化.

$$(1)\ A=\begin{pmatrix}1 & -1\\ 3 & -3\end{pmatrix},(2)\ A=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 2 & 0 & -1\end{pmatrix},$$

$$(3)\ A=\begin{pmatrix}2 & 0 & 0\\ 0 & 5 & -2\\ 0 & 6 & -2\end{pmatrix}.$$

例 2 在线性空间 $\mathbf{R}[x]_n$ 定义线性变换 *T*:

$$T(f(x)) = f'(x),\quad f(\mathrm{x}) \in \mathbf{R}[x]_n.$$

① 求线性变换 *T* 的特征值与特征向量;

② 当 *n*>1 时，是否存在 $\mathbf{R}[x]_n$ 的基，使得 *T* 在其下的矩阵为对角阵？

解 ① 取$\mathbf{R}[x]_n$中一组基$1, x, x^2, \cdots, x^{n-1}$, 则$T$在这组基下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & n-1 \\ & & & & 0 \end{pmatrix}.$$

计算知T的特征值$\lambda_1=\lambda_2=\cdots=\lambda_n=0$.

T 属于 0 的特征向量为$c, c\neq 0$.

② 0的几何重数为1，代数重数为n，当$n>1$时，不存在基使T 在其下的矩阵为对角阵.

例2 设$\boldsymbol{B}=\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\in P^{2\times2}$, 定义线性变换

$$T: P^{2\times2}\to P^{2\times2}, T(\boldsymbol{X})=\boldsymbol{BX}-\boldsymbol{XB},$$

并记子空间$W=\left\{\begin{pmatrix}a & b\\c & d\end{pmatrix}\in P^{2\times2}\middle| b+c=0\right\}$.

① 求W的一组基;

② 证明W是T的不变子空间;

③ 将T看作W上的线性变换, 求其特征值与特征向量;

④ 求W的基, 使T在其下的矩阵为对角阵.

2.2.2 Schur定理

定理3 (Schur) 设n阶矩阵$\boldsymbol{A}$的所有特征值为$\lambda_1, \lambda_2, \dots, \lambda_n$(重根按重数计算), 则存在**酉矩阵**$\boldsymbol{U}$, 使得

$$\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

注 上定理的结论是一种特殊的相似关系, 称为**酉相似**.

2.2.3 正规矩阵

定义1 设$\boldsymbol{A}\in\mathbf{C}^{n\times n}$, 若$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$, 则称$\boldsymbol{A}$为正规矩阵.

① 若$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$，则称$\boldsymbol{A}$为**Hermite矩阵**;

② 若$\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$, 则称$\boldsymbol{A}$为**反Hermite矩阵**.

③ Hermite矩阵, 反Hermite矩阵和酉矩阵都是正规矩阵, 但正规矩阵不限于此, 如

$$\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix},\quad \begin{pmatrix}2 & 0\\ 0 & i\end{pmatrix}.$$

④ 实(反)Hermite矩阵就是实(反)对称矩阵.

定理4 设$\boldsymbol{A}$是n阶正规矩阵, 则存在酉矩阵$\boldsymbol{U}$使得

$$\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

其中$\lambda_1, \lambda_2, \dots, \lambda_n$是$\boldsymbol{A}$的所有特征值(重根按重数计算).

注 正规矩阵酉相似于对角阵.

推论 ① Hermite矩阵的特征值均为实数.

② Hermite矩阵的不同特征值所对应的特征向量(在酉空间$\mathbf{C}^n$中)正交.

定义2 设$\boldsymbol{A}$为Hermite矩阵, $\boldsymbol{x} \in \mathbf{C}^n$, 称$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}$为Hermite二次型.

① Hermite二次型可通过酉变换化为标准形.

② 若对任意$\boldsymbol{x} \neq 0$, 都有$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x} > 0$, 则称$\boldsymbol{A}$正定, 记作$\boldsymbol{A} > 0$;

③ 若对任意$\boldsymbol{x}$, 都有$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x} \geq 0$, 则称$\boldsymbol{A}$半正定, 记作$\boldsymbol{A} \geq 0$.

定理 设$\boldsymbol{A}$是Hermite矩阵, 则以下等价:

① $\boldsymbol{A} > 0$;

② $\boldsymbol{A}$的特征值全都 > 0;

③ 存在可逆矩阵$\boldsymbol{B}$使得$\boldsymbol{A} = \boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}$.

定理 设$\boldsymbol{A}$是Hermite矩阵, 则以下等价:

① $\boldsymbol{A} \geq 0$;

② $\boldsymbol{A}$的特征值全都 ≥ 0;

③ 存在矩阵$\boldsymbol{B}$使得$\boldsymbol{A}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}$.

注 类似可定义负定和半负定, 且有类似结果.