北京航空航天大學
BEIHANG UNIVERSITY

矩阵理论

刘克新

自动化科学与电气工程学院

第二章 线性映射与矩阵

- 预备知识
- 线性映射
- 矩阵与同构基与坐标
- 特征值与特征向量
- 酉变换与酉矩阵
- 应用：图的矩阵表示

第二章 线性映射与矩阵

2.4 特征向量与特征值

定义2.4.1（**线性变换的特征值**和**特征向量**）设线性变换$T \in L(V)$，若存在$\lambda_0 \in F$及V的非零向量ξ使得$T\xi = \lambda_0 \xi$，则称λ_0是T的一个**特征值**，称ξ为T的属于特征值λ_0的一个**特征向量**.

例2.4.1 设V是数域F上的n维线性空间，定义恒等变换$T_1 \boldsymbol{x} = \boldsymbol{x}, \forall \boldsymbol{x} \in V$和零变换$T_2(\boldsymbol{x}) = \boldsymbol{\theta}$，$\forall \boldsymbol{x} \in V$，则$V$的任意非零向量$\xi$都是$T_1$的属于特征值$\lambda_0 = 1$的特征向量和$T_2$的属于特征值$\lambda_0 = 0$的特征向量.

注1： 从几何角度看，**特征向量在线性变换作用下保持共线**, 即在同一直线上（有可能反向）.

注2： 设ξ是T的属于特征值λ_0的一个特征向量，则$k\xi$也是T的属于特征值λ_0的特征向量, 其中$k \in F$且$k \neq 0$.

注3： 若$\xi \in N(T)$且$\boldsymbol{\xi \neq \theta}$, 则$\xi$是属于0的特征向量.

注4： 设$T \in L(V)$, $\xi_1, \cdots, \xi_n$是V的一组基, 且$T\xi_i = \lambda_i \xi_i$（$i = 1, \cdots, n$）, 则T在基$\xi_1, \cdots, \xi_n$下的矩阵为对角阵.

定义 2.4.2（**矩阵的特征值**和**特征向量**）设 $A \in F^{n\times n}$，λ为一文字，矩阵$\lambda I - A$称为A的**特征矩阵**，其行列式$|\lambda I - A|$称为A的**特征多项式**，方程$|\lambda I - A| = 0$的根称为A的**特征值**（或**特征根**）．方程$(\lambda I - A)\boldsymbol{\alpha} = 0$的非零解向量$\boldsymbol{\alpha}$称为属于特征值$\lambda$的**特征向量**.

注5：λ是线性变换T的特征值当且仅当λ是A的特征值；向量ξ是线性变换T的特征向量当且仅当$\boldsymbol{\alpha}$是A的特征向量，其中$\xi = [\xi_1, \cdots, \xi_n]\boldsymbol{\alpha}$，$\boldsymbol{\alpha}$是$\xi$在线性空间$V$的基$\xi_1, \cdots, \xi_n$下的坐标.

例2.4.2 设$T \in L(\mathbb{R}^3)$, 它在基ξ_1, ξ_2, ξ_3下的矩阵A为

$$\begin{bmatrix} 2 & -2 & 2 \\ -2 & -1 & 4 \\ 2 & 4 & -1 \end{bmatrix}$$

求T的全部特征值和特征向量.

思考：矩阵$A \in F^{n\times n}$有n个特征值吗？

例2.4.3 考查矩阵$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$，其特征多项式为$f_A(\lambda) = \begin{vmatrix} \lambda & 1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 + 1$. 若$F = \mathbb{C}$，则矩阵$A$有两个特征值，分别为±i；若$F = \mathbb{R}$，则方程$\lambda^2 + 1 = 0$无实根，即矩阵$A$在数域$\mathbb{R}$上无特征值. 因此, **矩阵的特征值依赖于$V$所在的数域$F$**.

注： 由一元n次多项式方程在复数域内有且仅有n个根知（代数基本定理），n阶矩阵A在复数域内必有n个特征值，记为$\lambda_1, \cdots, \lambda_n$，其中$\lambda_i$作为特征方程根的重数，称为$\lambda_i$的**代数重数**.

定理2.4.1 设$\lambda_1,\cdots,\lambda_n$是$A=(a_{ij})\in\mathbb{C}^{n\times n}$的特征值，则

$$\prod_{i=1}^{n}\lambda_i=|A|$$

$$\sum_{i=1}^{n}\lambda_i=\sum_{i=1}^{n}a_{ii}=\mathrm{tr}(A)$$

式中，$\mathrm{tr}(A)$称为矩阵A的**迹**.

例2.4.4 设λ是可逆复方阵A的特征值, 试证明

（1）λ^{-1}是A^{-1}的特征值;

（2）$\lambda^{-1}|A|$是A^*的特征值.

$$AA^* = A^*A = |A|E$$

例2.4.4 设λ是可逆复方阵A的特征值, 试证明

（1）λ^{-1}是A^{-1}的特征值;

（2）$\lambda^{-1}|A|$是A^{*}的特征值.

$$AA^{*}=A^{*}A=|A|E$$

$$r(A^{*})=\begin{cases} n, r(A)=n \\ 1, r(A)=n-1 \\ 0, r(A)\le n-2 \end{cases}$$

定义2.4.3（**特征子空间**）设λ是矩阵$A \in \mathbb{C}^{n\times n}$的一个特征值, 定义集合

$$E(\lambda_i) = \left\{\boldsymbol{x} \in \mathbb{C}^n \middle| A\boldsymbol{x} = \lambda_i \boldsymbol{x}\right\}$$

则$E(\lambda_i)$是$\mathbb{C}^n$的线性子空间, 称为属于特征值λ_i的**特征子空间**, $\dim\big(E(\lambda_i)\big)$为特征值$\lambda_i$的**几何重数**.

$$|\lambda I - A| = (\lambda - \lambda_1)^{\mu_1} \cdots (\lambda - \lambda_s)^{\mu_s}$$

μ_i称为特征值λ_i的**代数重数**

注6： 特征子空间$E(\lambda)$也是齐次线性方程组$(\lambda I - A)\boldsymbol{x} = 0$的解空间, 还是矩阵$(\lambda I - A)$的零空间.

A的属于特征值λ的所有特征向量构成线性空间吗？

$$E(\lambda) = \{\boldsymbol{x} \in \mathbb{C}^n | A\boldsymbol{x} = \lambda\boldsymbol{x}\}$$

特征值λ的**特征子空间，维数至少是1维，最高是多少维？**

例2.4.5 求如下矩阵特征值的代数重数和几何重数

$$J_1=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},J_2=\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix},J_3=\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}$$

代数重数都是3

几何重数分别是3、2、1

定理2.4.2 复方阵的任一特征值的几何重数不超过它的代数重数.

定理2.4.2 复方阵的任一特征值的几何重数不超过它的代数重数.

若复方阵的任一特征值的几何重数等于它的代数重数，则矩阵可相似对角化，反之亦成立

命题2.4.1 若n阶方阵A与B相似, 则

（1）A与B有相同的特征多项式与特征值;

（2）A与B有相同的秩与行列式;

（3）A与B有相同的迹.

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

注8： 线性变换的矩阵的特征多项式与基的选取无关, 它直接由线性变换决定, 故可称之为**线性变换的特征多项式**.

定理2.4.4 矩阵A的属于不同特征值的特征向量线性无关.

注9： 形如D的矩阵称为Vardermonde矩阵，它的行列式称为Vardermonde行列式.

例2.4.6 设$A=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$，$\boldsymbol{x}^{(0)}=\begin{bmatrix}0\\ 1\end{bmatrix}$，试利用幂法求$A$的绝对值最大的特征值及其特征向量的近似值（$k$取到5）.

思考：若初试向量$\boldsymbol{x}^{(0)}$与$\boldsymbol{x}_1$正交对结果会有影响吗？

思考：在知道特征值的一个较好的估计值后，如何进一步提高精度？

思考:阅读论文"Eigenvectors from eigenvalues"[25],试比较文献中所提及的计算矩阵特征向量的方法与定义法、幂法的优劣.