§ 5.4 特征值与特征向量.

本节主要研究线性变换在何种情况下最简单.

1. 特征值与特征向量.

Def 1、设 V 是 $\mathbb{P}$ 上 n 维线性空间, $\mathscr{A}\in \mathrm{End}\,V$, $\lambda_0\in\mathbb{P}$.

若 $\exists$ V 中 $\xi\neq 0$, s.t. $A\xi=\lambda_0\xi$

则称 λ_0 是 $\mathscr{A}$ 的特征值, ξ 为 $\mathscr{A}$ 的属于 λ_0 的特征向量.

• 几何上, 特征向量在变换前后是共线的.

Th1、设 V 是 $\mathbb{P}$ 上线性空间, $\mathscr{A}\in\mathrm{End}\,V$, $\lambda_0\in\mathbb{P}$, 令

$$E_{\lambda_0}(\mathscr{A}):=\{\xi\in V \mid A\xi=\lambda_0\xi\}.$$

则 $E_{\lambda_0}(\mathscr{A})$ 是 V 的子空间. 且 $\dim E_{\lambda_0}(\mathscr{A})=\dim V-\mathrm{rank}(\lambda_0\,\mathrm{id}-\mathscr{A})$.

又 $E_{\lambda_0}(\mathscr{A})$ 与下面三命题等价.

1) λ_0 是 $\mathscr{A}$ 特征值

2) $E_{\lambda_0}(\mathscr{A})\neq\{0\}$, 或 $\dim E_{\lambda_0}(\mathscr{A})>0$.

3) $\det(\lambda_0\,\mathrm{id}-\mathscr{A})=0$ …特征值

称 $E_{\lambda_0}(\mathscr{A})$ 为 $\mathscr{A}$ 属于 λ_0 的特征子空间.

Prf 1° $E_{\lambda_0}(\mathscr{A})$ 为子空间. 且 $0\in E_{\lambda_0}(\mathscr{A})$,

若 $\xi,\eta\in E_{\lambda_0}(\mathscr{A})$, $k,l\in\mathbb{P}$,

$$\mathscr{A}(k\xi+l\eta)=k\mathscr{A}\xi+l\mathscr{A}\eta=\lambda_0\underbrace{(k\xi+l\eta)}_{\in E_{\lambda_0}(\mathscr{A})}.$$

2° 对 $E_{\lambda_0}(\mathscr{A})=\{\xi\in V\mid \mathscr{A}\xi=\lambda_0\xi\}$

$$=\{\xi\in V\mid (\lambda_0\,\mathrm{id}-\mathscr{A})\xi=0\}$$

$$=\ker(\lambda_0\,\mathrm{id}-\mathscr{A}).$$

则 $\dim E_{\lambda_0}(\mathscr{A})=\dim\ker(\lambda_0\,\mathrm{id}-\mathscr{A})$

$=\dim V-\dim(\lambda_0\,\mathrm{id}-\mathscr{A})V$ ← 满

$=\dim V-\mathrm{rank}(\lambda_0\,\mathrm{id}-\mathscr{A})$

3° 条件等价性. 1⇒2 by def 2⇒3

2⇒3 由 $\dim E_{\lambda_0}(\mathcal{A})>0$, 存 $\xi\neq 0$, $\exists\, \xi\in E_{\lambda_0}(\mathcal{A})$ s.t.

$\mathcal{A}\xi=\lambda_0\xi$. 即 $(\lambda_0 \mathrm{id}-\mathcal{A})\xi=0$.

在V中取基 $\alpha_1\cdots\alpha_n$ $M(\lambda_0\mathrm{id}-\mathcal{A};\alpha_1\cdots\alpha_n)\underbrace{\mathrm{crd}(\xi;\alpha_1\cdots\alpha_n)}_{\neq 0}=0$

由

$\Rightarrow \det(\lambda_0\mathrm{id}-\mathcal{A})=\det M(\lambda_0\mathrm{id}-\mathcal{A};\alpha_1\cdots\alpha_n)=0$.

3⇒1 在V中取基 $\alpha_1\cdots\alpha_n$, 于是 $\det M(\lambda_0\mathrm{id}-\mathcal{A};\alpha_1\cdots\alpha_n)=\cancel{\det(\lambda_0\mathrm{id}}$

$=\det(\lambda_0\mathrm{id}-\mathcal{A})=0$.

$\exists\, X_0\in P^{n\times 1}$, $X_0\neq 0$, s.t.

$$M(\lambda_0\mathrm{id}-\mathcal{A};\ldots)X_0=0.$$

因而 $\xi=(\alpha_1\cdots\alpha_n)X_0\neq 0$, 且 $\mathcal{A}\xi=\lambda_0\xi$. ∴ 特征子空间.

Def 2、设 $A\in P^{n\times n}$, λ是一个文字, 称 $\det(\lambda I_n-A)$ 为A的多项式

它的根称为A的特征值/特征根. 若λ_0为A的特征值, 称

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

的非零解为λ_0的特征向量.

由此给出特征值与特征多项式的求法.

1、在V中取一组基 $\alpha_1,\cdots,\alpha_n$, 求$\mathcal{A}$在此基下的矩阵 $A:=M(\mathcal{A};\alpha_1\cdots\alpha_n)$.

2、对特征多项式 $\det(\lambda I_n-A)=f(\lambda)$ 的根. 即$\mathcal{A}$的特征值

3、对每个特征值λ_0 解齐次方程组 $(\lambda_0 I_n-A)X=0$. 其基础解系是 $E_{\lambda_0}(\mathcal{A})$ 的基在 $\alpha_1,\cdots,\alpha_n$ 下的坐标.

例1. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 为V的基, $\mathcal{A}\in\mathrm{End}\,V$, 且 $M(\mathcal{A};\varepsilon_1,\varepsilon_2,\varepsilon_3)=A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$.

求$\mathcal{A}$的特征值与特征向量.

解、由计算机, 特征多项式 $|\lambda I_3-A|=(\lambda-5)(\lambda+1)^2$. 则根为 -1 (二重) 与 5.

特征值为 -1, 5.

当 $\lambda=-1$ 时，$\begin{pmatrix}-2&-2&-2\\-2&-2&-2\\-2&-2&-2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=0$，有 $\begin{pmatrix}1\\-1\\0\end{pmatrix}\begin{pmatrix}1\\0\\-1\end{pmatrix}$ 作为基础解系.

故 $E_{-1}(\mathscr{A})$ 有基 $\varepsilon_1-\varepsilon_2$，$\varepsilon_1-\varepsilon_3$.

当 $\lambda=5$ 时，$\begin{pmatrix}4&-2&-2\\-2&4&-2\\-2&-2&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=0$，有基础解系 $\begin{pmatrix}1\\1\\1\end{pmatrix}$.

故 $E_5(\mathscr{A})$ 有基 $\varepsilon_1+\varepsilon_2+\varepsilon_3$.

注意到 $\varepsilon_1-\varepsilon_2$，$\varepsilon_1-\varepsilon_3$，$\varepsilon_1+\varepsilon_2+\varepsilon_3$ 亦为 V 的基，$\mathscr{A}$ 在此基下矩阵为

$$\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}\begin{pmatrix}1&1&1\\-1&0&1\\0&-1&1\end{pmatrix}=\begin{pmatrix}-1&&\\&-1&\\&&5\end{pmatrix}$$，非常简单.

Eg 2. 求 $\mathbb{P}[x]_n$ 线性变换 $\mathscr{D}=\frac{d}{dx}$ 特征值与特征向量.

取 $\mathbb{P}[x]_n$ 中基 $1, x, x^2, \cdots, x^{n-1}$，$\mathscr{D}$ 在此基下矩阵为

$$D=\begin{pmatrix}0&1&0&\cdots&0\\&0&2&\cdots&0\\&&\ddots&\ddots&\vdots\\&&&0&n-1\\&&&&0\end{pmatrix},\ \det(\lambda I_n-D)=\lambda^n.$$

故 $-DX=0$. 基础解系为 $\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$，故 $\mathscr{D}=\frac{d}{dx}$ 特征向量为常数 $k\neq0$.

Eg 3. 求旋转 $I_\theta=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}$ 特征向量.

$\lambda^2-2\lambda\cos\theta+1$，当 $\cos\theta=\pm1$ 时，特征值为 1，-1. 平面上所有非0向量均为特征向量.

2. 矩阵多项式.

设 $\mathbb{P}$ 为数域，λ 为文字，$\mathbb{P}[\lambda]^{n\times n}$ 表示以 λ 为多项式为元素的 n 阶方阵集合，且 $\mathbb{P}^{n\times n}\subset\mathbb{P}[\lambda]^{n\times n}$. 在 $\mathbb{P}[\lambda]^{n\times n}$ 亦可定义加法、乘法且具有相同规律. $A(\lambda)\in\mathbb{P}[\lambda]^{n\times n}$，$A(\lambda)$ 可逆 $\Leftrightarrow$ $\det A(\lambda)$ 为非0常数.

下面考虑 $a_{ij}(\lambda)=\sum_{k=0}^{m}a_{ijk}\lambda^k$，$1\leq i,j\leq n$ 构成的 $A(\lambda)=(a_{ij}(\lambda))\in\mathbb{P}[\lambda]^{n\times n}$，按多项式之系数排，$A_k=(a_{ijk})$. 于是 $A(\lambda)=\sum_{k=0}^{m}A_k\lambda^k$，因此 $A_0\cdots A_m$ 由 $A(\lambda)$ 唯一决定.

下面定义它的若干运算 设 $A(\lambda)=\sum_{k=0}^{m} A_k\lambda^k$， $B(\lambda)=\sum_{k=0}^{m} B_k\lambda^k$.

$$A(\lambda)+B(\lambda)=\sum_{k=0}^{m}(A_k+B_k)\lambda^k$$

$$A(\lambda)B(\lambda)=\underbrace{\sum_{t=0}^{2m}}_{\text{枚举次数}}\sum_{k+l=t}A_kB_l\lambda^t.$$

Eg 4. $\begin{pmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & -2\\ -2 & -2 & \lambda-1\end{pmatrix}=\lambda\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$

Th2.（Hamilton-Caylay）设 $A\in\mathbb{P}^{n\times n}$，$f(\lambda)=\det(\lambda I_n-A)$
$=\lambda^n+a_1\lambda^{n-1}+\cdots+a_n$

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

Prf. 设 $B(\lambda)=(\lambda I_n-A)^*$ ↖伴随. 于是 $\mathrm{ent}_{ij}B(\lambda)\in\mathbb{P}[\lambda]$ 是 $\mathrm{ent}_{ij}(\lambda I_n-A)$ 的代数余子式. 因而 $\deg(\mathrm{ent}_{ij}B(\lambda))\le n-1$.

于是
$$B(\lambda)=\lambda^{n-1}B_0+\lambda^{n-2}B_1+\cdots+B_{n-1},\ B_i\in\mathbb{P}^{n\times n}$$

又由于 $B(\lambda)(\lambda I_n-A)=f(\lambda)I_n$

代入 $B(\lambda)$，对比系数： $(\lambda^{n-1}B_0+\lambda^{n-2}B_1+\cdots+B_{n-1})(\lambda I_n-A)$
$=\lambda^nI_n+\lambda^{n-1}a_1I_n+\cdots+a_nI_n$

得
$$\begin{cases}B_0=I_n & \xrightarrow{\times A^n}\ \cancel{A^n}B_0A^n=A^n\\ B_1-B_0A=a_1I_n & \xrightarrow{\times A^{n-1}}\ A^{n-1}B_0A^n=a_1I_n\\ \cdots\cdots & \quad\vdots\\ B_{n-1}-B_{n-2}A=a_{n-1}I_n & \\ -B_{n-1}A=a_nI_n & \xrightarrow{\times I_n}\ -B_{n-1}A=a_nI_n\end{cases}$$

↓ 相加 ↓

$0=f(A)$.

推论1. 设 $\mathscr{A}$ 是 n 维空间的线性变换, $f(\lambda)$ 是 $\mathscr{A}$ 的特征多项式, 则 $f(\mathscr{A})=0$.

推论2. 设 $\mathscr{A}$ 是 n 维线性空间的线性变换, $d_{\mathscr{A}}(\lambda)$ 是 $\mathscr{A}$ 的最小多项式, 则 $\deg d_{\mathscr{A}}(\lambda) \leq n$.

◁ $f(\lambda)$ 为 $\mathscr{A}$ 的特征多项式, 而 $d_{\mathscr{A}}(\lambda) \mid f(\lambda)$, 故 $\deg d_{\mathscr{A}}(\lambda) \leq n$.

(或可看作矩阵).

附. 关于 Cayley-Hamilton 直观理解.

将其因式分解后每一个因子当成线性变换.

$$(\lambda_1 \mathrm{id} - \mathscr{A})(\lambda_2 \mathrm{id} - \mathscr{A})\cdots$$

↓

把每一个特征向量拉到 0 的位置.