Chapter 8

特征值和特征向量

在上一章中, 我们介绍了多项式对应的伴阵, 那么, 如何由伴阵恢复多项式呢?

假设多项式 $p(x) = x^2 + ax + b$, 则其伴阵为 $C[p(x)] = \begin{pmatrix} 0 & -b \\ 1 & -a \end{pmatrix}$,

$\det(xI - C[p(x)]) = \begin{vmatrix} x & b \\ -1 & x+a \end{vmatrix} = x^2 + ax + b$ 即恢复多项式.

同理, 对循环不变子空间 $\langle\langle v_{ij} \rangle\rangle$, 设其极小多项式为 $p_i^{e_{ij}}(x)$, 伴阵为 $C[p_i^{e_{ij}}(x)]$, 有 $\det(xI - C[p_i^{e_{ij}}]) = p_i^{e_{ij}}(x)$,
由定理 **??**, 线性算子 τ 的表示即为这些伴阵的直和, 故有

定义 8.1 <u>特征多项式</u>: $\det(xI - [\tau]_{\mathcal{B}}) = \prod_{ij} p_i^{e_{ij}}(x) \equiv C_\tau(x)$, 称为 τ 的**特征多项式**.

定理 8.1 <u>(课本第3 版引理7.17, 定理7.18)</u>: 特征多项式在相似操作下不变, 或线性算子的特征多项式唯一, 与线性算子的表示无关.

证: 设线性算子 τ 在定序基 $\mathcal{B}$ 和 $\mathcal{B}'$ 下的表示分别为 $[\tau]_{\mathcal{B}}$ 和 $[\tau]_{\mathcal{B}'}$,
则 $C_{[\tau]_{\mathcal{B}'}}(x) = \det(xI - [\tau]_{\mathcal{B}'}) = \det(xI - M_{\mathcal{B}\mathcal{B}'}[\tau]_{\mathcal{B}} M_{\mathcal{B}\mathcal{B}'}^{-1}) = \det(M_{\mathcal{B}\mathcal{B}'}(xI - [\tau]_{\mathcal{B}}) M_{\mathcal{B}\mathcal{B}'}^{-1}) = \det(M_{\mathcal{B}\mathcal{B}'}^{-1} M_{\mathcal{B}\mathcal{B}'}(xI - [\tau]_{\mathcal{B}})) = \det(xI - [\tau]_{\mathcal{B}}) = C_{[\tau]_{\mathcal{B}}}(x)$. □

$\because C_\tau(x) = \prod_{ij} p_i^{e_{ij}} = \prod_i p_i^{\sum_{j=1}^{t_i} e_{ij}}$, $m_\tau(x) = u p_1^{e_1} \cdots p_m^{e_m}$, 且 $e_i = e_{i1} \geq \cdots \geq e_{it_i}$ (不可约因子相同, 但幂次不等), $\therefore \deg C_\tau(x) = \sum_{ij} \deg p_i^{e_{ij}}(x) = \sum_{ij} d_{ij} = n \geq \deg m_\tau(x) \Longrightarrow m_\tau(x) \mid C_\tau(x)$.

在 $F[x]$ 中, 任一一次多项式 $x - r$ 都是不可约的.
在 $\mathbb{R}[x]$ 中, 有实根 $\Longleftrightarrow$ 可约.
$\because$ 实系数多项式的复根的共轭亦为该多项式的根, $\therefore$ 实系数多项式的复根总是成对出现[1], 故 $\mathbb{R}[x]$ 中的不可约多项式阶数 ≤ 2.

证: 当多项式阶数 > 2, 若阶数为奇数, 则多项式的根两两配对后必留下一个根, 该根必不为复数而为实数 $\Longrightarrow$ 可约; 若阶数为偶数, 以 4 阶为例, 假设多项式不可约, 设复根分别为 $z_1, \bar{z}_1, z_2, \bar{z}_2$, 则多项式可写为 $(x - z_1)(x - \bar{z}_1)(x - z_2)(x - \bar{z}_2) = (x - (z_1 + \bar{z}_1)x + z_1\bar{z}_1)(x - (z_2 + \bar{z}_2)x + z_2\bar{z}_2)$,

[1]

证: $0 \neq f(x) \in \mathbb{R}[x]$, 若 $z \in \mathbb{C}$, s.t. $f(z) = 0$, 即 $z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0$
$\Longrightarrow \bar{f(z)} = \bar{z}^n + a_{n-1}\bar{z}^{n-1} + \cdots + a_1\bar{z} + a_0 = f(\bar{z}) = 0$, 故得证. □

$(z_1+\bar{z}_1), z_1\bar{z}_1, (z_2+\bar{z}_2), z_2\bar{z}_2 \in \mathbb{R}$, $\therefore (x-(z_1+\bar{z}_1)x+z_1\bar{z}_1), (x-(z_2+\bar{z}_2)x+z_2\bar{z}_2) \in \mathbb{R}[x] \Longrightarrow$ 多项式可约, 故假设错误, 阶数 >2 的偶数阶实系数多项式必可约. □

例 8.1: $x^2+1=(x+i)(x-i)$ 无实根不可约. □

$\mathbb{Q}[x]$ 中的不可约多项式可以是任意次.
可用 **Eisenstein 方法** 判别 $\mathbb{Q}[x]$ 中的多项式的可约性: $\forall f(x)=a_0+a_1x+\cdots+a_nx^n$, 其中 $a_i\in\mathbb{Z}$, 若 $\exists$ 素数 p, s.t. $p\nmid a_n$, $p\mid a_i|_{i=0,\cdots,i=n-1}$, $p^2\nmid a_0$, 则 $f(x)$ 在 $\mathbb{Q}[x]$ 中不可约.

定义 8.2 特征值和特征向量: 线性算子 $\tau\in\mathcal{L}(V)$, $\lambda\in F$, 若 $\exists 0\neq v\in V$, s.t. $\tau(v)=\lambda v$, 则称 λ 为 τ 的**特征值**, v 为 τ 的**特征向量**.

定义 8.3: $\mathcal{E}_\lambda \equiv \{$特征值 λ 对应的特征向量$\} = \{v\in V \mid \tau(v)=\lambda v\}$.

$\mathcal{E}_\lambda$ 是 V 的子空间.

证: $\forall u,v\in\mathcal{E}_\lambda$, $\forall r,t\in F$, $\tau(ru+tv)=r\tau(u)+t\tau(v)=r(\lambda v)+t(\lambda v)=\lambda(ru)+\lambda(tv)=\lambda(ru+tv)\Longrightarrow ra+tv\in\mathcal{E}_\lambda$, 故得证. □

定义 8.4 特征谱: $\mathrm{spec}(\tau)\equiv\{\tau$ 的特征值$\}$.

(1) $\tau(v)=\lambda v \Longleftrightarrow \tau(v)-\lambda v=0$
$\Longleftrightarrow (\tau-\lambda)(v)=0$
$\Longleftrightarrow v\in\ker(\tau-\lambda)$
$(\because v\neq 0)$
$\Longleftrightarrow \ker(\tau-\lambda)\neq 0$
$\Longleftrightarrow (\tau-\lambda)$ 非单
$\Longleftrightarrow (\tau-\lambda)$ 是退化的.
(也有教材将 $\tau-\lambda$ 退化作为特征值的定义.)

(2) $v\neq 0$ 是 λ 对应的特征向量, $\tau(v)=\lambda \Longleftrightarrow (\tau-\lambda)(v)=0$,
若 $f(x)=x-\lambda$, 则 $f(x)$ 零化 v, 即 $f(\tau)(v)=0\Longrightarrow f(x)\in\mathrm{ann}(v)$,
又 $\because x-\lambda$ 不可约, $\therefore \mathrm{ann}(v)=\langle x-\lambda\rangle$,
$\because \mathrm{ann}(V)=\langle m_\tau(x)\rangle$, $\therefore \forall u\in V$, $m_\tau(x)(u)=0\Longrightarrow m_\tau(v)=0\Longrightarrow m_\tau(x)\in\mathrm{ann}(v)=\langle x-\lambda\rangle\Longrightarrow (x-\lambda)\mid m_\tau(x)$,
设 $m_\tau(x)=(x-\lambda)q(x)$, 其中 $q(x)\in F[x]$, 故特征值为极小多项式的根.

定理 8.2 (课本定理8.3): (1) $\mathrm{spec}(\tau)$ 是 $m_\tau(x)$ 或 $C_\tau(x)$ 的根的集合.

(2) 矩阵的特征值在相似操作下不变, 即相似矩阵具有相同的特征值.

(3) $\mathcal{E}_\lambda$ 是 $(\lambda I-A)x=0$ 的解空间.

定理 8.3 (课本定理8.4): $\tau \in \mathcal{L}(V)$, $\lambda_1, \cdots, \lambda_k$ 是 τ 的特征值且互不相等, 则

(1) 取 $0 \neq v_i \in \mathcal{E}_{\lambda_i}$, 则 $\{\lambda_1, \cdots, \lambda_k\}$ 线性无关;

(2) 若 $i \neq j$, $\mathcal{E}_i \neq \mathcal{E}_j = \{0\}$.

证: (1) 若 $\sum_{i=1}^k r_i v_i = 0$, 则 $0 = \tau(0) = \tau\left(\sum_{i=1}^k r_i v_i\right) = \sum_{i=1}^k r_i \tau(v_i) = \sum_{i=1}^k r_i \lambda_i v_i$, 即 $r_1\lambda_1 v_1 + \cdots + r_k \lambda_k v_k = 0$,
又有 $\lambda_k \left(\sum_{i=1}^k r_i v_i\right) = 0$, 即 $r_1\lambda_k v_1 + r_2 \lambda_k v_2 + \cdots + r_k \lambda_k v_k = 0$,
以上两式相减消去 v_k 项得, $r(\lambda_1 - \lambda_k)v_1 + r_2(\lambda_2 - \lambda_k)v_2 + \cdots + r_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0$,
将 τ 作用于上式得, $\tau\left[r(\lambda_1 - \lambda_k)v_1 + r_2(\lambda_2 - \lambda_k)v_2 + \cdots + r_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1}\right] = r_1(\lambda_1 - \lambda_k)\lambda_1 v_1 + r_2(\lambda_2 - \lambda_k)\lambda_2 v_2 + \cdots + r_{k-1}(\lambda_{k-1} - \lambda_k)\lambda_{k-1} v_{k-1} = \tau(0) = 0$,
消去 v_{k-1} 项得, $\lambda_{k-1}\left[r(\lambda_1 - \lambda_k)v_1 + r_2(\lambda_2 - \lambda_k)v_2 + \cdots + r_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1}\right] - \tau\left[r(\lambda_1 - \lambda_k)v_1 + r_2(\lambda_2 - \lambda_k)v_2 + \cdots\right.$
$r_1(\lambda_k - \lambda_1)(\lambda_{k-1} - \lambda_1)v_1 + r_2(\lambda_k - \lambda_2)(\lambda_{k-1} - \lambda_2)v_2 + \cdots + r_{k-1}(\lambda_k - \lambda_{k-2})(\lambda_{k-1} - \lambda_{k-2})v_{k-2} = 0$,
重复以上操作 $k-1$ 次得, $r_1(\lambda_2 - \lambda_1)\cdots(\lambda_k - \lambda_1)v_1 = 0$,
$\because \lambda_1, \cdots, \lambda_k$ 互不相等且 $v_1 \neq 0$, $\therefore r_1 = 0$,
同理可得 $r_1 = \cdots = r_k = 0$, 故 $\{v_1, \cdots, v_k\}$ 线性无关.

(2) 取 $u \in \mathcal{E}_{\lambda_i} \cap \mathcal{E}_{\lambda_j} \Longrightarrow u \in \mathcal{E}_{\lambda_i} \Longrightarrow \tau(u) = \lambda_i u$,
且 $u \in \mathcal{E}_{\lambda_j} \Longrightarrow \tau(u) = \lambda_j u$
以上两式相减得, $0 = \tau(u) - \tau(u) = \lambda_i u - \lambda_j u = (\lambda_i - \lambda_j)u$,
又 $\because \lambda_i \neq \lambda_j$, $\therefore u = 0 \Longrightarrow \mathcal{E}_{\lambda_i} \cap \mathcal{E}_{\lambda_j} = \{0\}$.

□

若 $\tau \in \mathcal{L}(V)$, $p(x) \in F[x]$, 则 $p(\tau) \in \mathcal{L}(V)$, 那么 $p(\tau)$ 与 τ 的特征值有何关系?

定理 8.4 特征谱映射定理(课本第3 版定理8.3): F 为代数闭域[a], $p(x) \in F[x]$, 则 $\operatorname{spec}(p(\tau)) = p(\operatorname{spec}(\tau)) \equiv \{p(\lambda) \mid \lambda \in \operatorname{spec}(\tau)\}$.

[a]即 $F[x]$ 中不可约多项式阶数 $= 1$.

证: 首先证 $p(\operatorname{spec}(\tau)) \subseteq \operatorname{spec}(p(\tau))$: $\forall v \in \operatorname{spec}(\tau)$, $\tau(v) = \lambda v \Longrightarrow \tau^k(v) = \lambda^k v$,
设 $p(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$, 则 $p(\tau) = \tau^d + a_{d-1}\tau^{d-1} + \cdots + a_1\tau + a_0$,
$p(\tau)(v) = \tau^d(v) + a_{d-1}\tau^{d-1}(v) + \cdots + a_1\tau(v) + a_0 = \lambda^d v + a_{d-1}\lambda^{d-1} v + \cdots + a_1 \lambda v + a_0 = (\lambda^d + a_{d-1}\lambda_{d-1} + \cdots + a_1\lambda + a_0)v = p(\lambda)v$,
$\Longrightarrow p(\lambda)$ 是 $p(\tau)$ 的特征值, 即 $p(\lambda) \in \operatorname{spec}(p(\tau))$, 故 $p(\operatorname{spec}(\tau)) \subseteq \operatorname{spec}(p(\tau))$.

再证 $\operatorname{spec}(p(\tau)) \subseteq p(\operatorname{spec}(\tau))$: $\forall \lambda \in \operatorname{spec}(p(t))$, $0 \neq v \in \mathcal{E}_\lambda$, 即 $p(\tau)(v) = \lambda v \Longrightarrow (p(\tau) - \lambda)v = 0$,
令 $f(x) = p(x) - \lambda$, 则 $f(\tau)(v) = 0$,
$\because \lambda \in F$, $\therefore f(x) \in F[x]$,
$\because F[x]$ 是代数封闭的, $\therefore f(x)$ 可写为 $f(x) = (x - r_1)\cdots(x - r_m)$, 其中 $r_i \in F$, 或有重复
$\Longrightarrow (\tau - r_1)\cdots(\tau - r_m)(v) = 0$, 其中 $r_i \in \operatorname{spec}(\tau)$
其中必 $\exists r_i$, s.t. $(\tau - r_i)\cdots(t - r_m)(v) = 0 \Longrightarrow p(r_i) - \lambda = 0 \Longrightarrow p(r_i) = \lambda$,
$\Longrightarrow \lambda \in p(\operatorname{spec}(\tau))$, 故 $\operatorname{spec}(p(\tau)) \subseteq p(\operatorname{spec}(\tau))$.

综上, 得证. □

定义 8.5 **约当标准型:** F 为代数闭域, V 为 F 上的向量空间, $\dim V < \infty$, $\tau \in \mathcal{L}(V)$, 则 V 的最小多项式 $m_\tau(x) = (x-\lambda_1)^{e_1}\cdots(x-\lambda_m)^{e_m}$,
此时 $V = \bigoplus_{i=1}^m V_i$, 其中 $\operatorname{ann}(V_i) = \langle (x-\lambda_i)^{e_i}\rangle$,
$V_i = \bigoplus_{j=1}^{t_i}\langle\langle v_{ij}\rangle\rangle$, 其中 $\operatorname{ann}(v_{ij}) = \langle (x-\lambda_i)^{e_{ij}}\rangle$, $e_i = e_{i1} \geq \cdots \geq e_{it_i}$, $\dim\langle\langle v_{ij}\rangle\rangle = \deg(x-\lambda_i)^{e_{ij}} = e_{ij}$,
则定序基 $\mathcal{B}'_{ij} = \{v_{ij}, (\tau-\lambda_i)(v_{ij}), \cdots, (\tau-\lambda_i)^{e_{ij}-1}(v_{ij})\}$ 为 $\langle\langle v_{ij}\rangle\rangle$ 下, 线性算子 τ 在 $\langle\langle v_{ij}\rangle\rangle$ 中可表为 $[\tau]_{\mathcal{B}'_{ij}} = \begin{pmatrix}[\tau(b_1)]_{\mathcal{B}'_{ij}} & \cdots & [\tau(b_m)]_{\mathcal{B}'_{ij}}\end{pmatrix}$,

其中 $[\tau(b_1)]_{\mathcal{B}'_{ij}} = [\tau(v_{ij})]_{\mathcal{B}'_{ij}} = [(\tau-\lambda_i+\lambda_i)(v_{ij})]_{\mathcal{B}'_{ij}} = [(\tau-\lambda_i)(v_{ij})+\lambda_i(v_{ij})]_{\mathcal{B}'_{ij}} = \begin{pmatrix}\lambda_i\\1\\0\\\vdots\\0\end{pmatrix}$, $[\tau(b_2)]_{\mathcal{B}'_{ij}} = [\tau(\tau-\lambda_i)(v_{ij}))]_{\mathcal{B}'_{ij}} = [(\tau-\lambda_i+\lambda_i)((\tau-\lambda_i)(v_{ij}))]_{\mathcal{B}'_{ij}} = [(\tau-\lambda_i)^2(v_{ij})+\lambda_i\tau(v_{ij})]_{\mathcal{B}'_{ij}} = \begin{pmatrix}0\\\lambda_i\\1\\0\\\vdots\\0\end{pmatrix}$, $\cdots$, $[\tau((\tau-\lambda_i)^{e_{ij}-1}(v_{ij}))] = [(\tau-\lambda_i+\lambda_i)((\tau-\lambda_{ij})^{e_{ij}-1}(v_{ij}))]_{\mathcal{B}'_{ij}} = [(\tau-\lambda_i)^{e_{ij}}(v_{ij})+\lambda_i(\tau-\lambda_i)^{e_{ij}-1}(v_{ij})]_{\mathcal{B}'_{ij}} = [\lambda_i(\tau-\lambda_i)^{e_{ij}-1}(v_{ij})]_{\mathcal{B}'_{ij}} = \begin{pmatrix}0\\\vdots\\0\\\lambda_i\end{pmatrix}$,

从而 $[\tau]_{\mathcal{B}'_{ij}} = \begin{pmatrix}\lambda_i & 0 & \cdots & \cdots & 0 & 0\\ 1 & \lambda_i & \cdots & \cdots & 0 & 0\\ 0 & 1 & \cdots & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & \cdots & \lambda_i & 0\\ 0 & 0 & \cdots & \cdots & 1 & \lambda_i\end{pmatrix} \equiv g(\lambda_i, e_{ij})$, 称为**约当块**,
定序基 $\mathcal{B} = \cup_{ij}\mathcal{B}'_{ij}$ 下, 线性算子 τ 可表为

$$[\tau]_{\mathcal{B}} = \begin{pmatrix}g(\lambda_1,e_{11}) &&&&&&\\ & \ddots &&&&&\\ && g(\lambda_1,e_{1t_1}) &&&&\\ &&& \ddots &&&\\ &&&& g(\lambda_m,e_{m1}) &&\\ &&&&& \ddots &\\ &&&&&& g(\lambda_m,e_{mt_m})\end{pmatrix},$$

称为**约当标准型**.
[a]

[a] $\mathcal{B}_{ij} = \{v_{ij}, \tau(v_{ij}), \cdots, \tau^{e_{ij}-1}(v_{ij})\}$ 亦为 $\langle\langle v_{ij}\rangle\rangle$ 的一组基, $[\tau]_{\mathcal{B}=\cup_{ij}\mathcal{B}_{ij}}$ 为有理标准型.

$\mathcal{B}'_{ij} = \{v_{ij}, (\tau-\lambda_i)(v_{ij}), \cdots, (\tau-\lambda_i)^{e_{ij}-1}(v_{ij})\}$ 为 $\langle\langle v_{ij}\rangle\rangle$ 的一组基.

证: 先证 $\mathcal{B}'_{ij}$ 线性无关: 设 $l_0 v + l_1(\tau-\lambda_i)(v) + \cdots + l_{m-1}(\tau-\lambda_i)^{m-1}(v) = 0$,
令 $h(x) = l_0 + l_1 x + \cdots + l_{m-1}x^{m-1}$, 则 $h(x)v = 0 \Longrightarrow h(x) \in \operatorname{ann}(v_{ij}) = \langle (x-\lambda_i)^{e_{ij}}\rangle$,
然而 $\because \deg(x-\lambda_i)^{e_{ij}} = e_{ij} \geq \deg h(x) = e_{ij}-1$, $\therefore l_0 = l_1 = \cdots = k_{e_{ij}-1}$, 故 $\mathcal{B}'_{ij}$ 线性无关.

再证 $\mathcal{B}'_{ij}$ 生成 $\langle\langle v_{ij}\rangle\rangle$: $\langle\langle v_{ij}\rangle\rangle = \{h(x)v \mid h(x) \in F[x]\} = \{h(\tau)(v) \mid h(x) \in F[x]\}$,
$\forall h(\tau)v \in \langle\langle v_{ij}\rangle\rangle$, 若 $\deg h(x) \leq e_{ij}-1$, 则 $h(x)v =$, 即 $h(\tau)v = (l_0 + l_0\tau + \cdots + l_{e_{ij}-1}\tau^{e_{ij}-1})(v) = (l_0 + l_0((\tau-\lambda_i)+\lambda_i) + \cdots + l_{e_{ij}-1}((\tau-\lambda_i)+\lambda_i)^{e_{ij}-1})(v)$ 可由 $\mathcal{B}'_{ij}$ 表示,
若 $\deg h(x) > e_{ij}-1$, $h(x) = h((x-\lambda_i)+\lambda_i) = h'(x-\lambda_i) = q(x-\lambda_i)(x-\lambda_i)^{e_{ij}-1} + r(x-\lambda_i)$, 其中 $q(x-\lambda_i)$ 为商多项式, 余多项式 $r(x-\lambda_i) = 0$ 或 $\deg r(x) < \deg(x-\lambda_{ij})^{e_{ij}-1} = e_{ij}-1$,
$\Longrightarrow h(x)v_{ij} = h'(x-\lambda_i)v_{ij} = (q(\tau-\lambda_i)(\tau-\lambda_i)^{e_{ij}} + r(\tau-\lambda_i))(v_{ij}) = q(\tau-\lambda_i)p(\tau-\lambda_i)(v_{ij}) + r(\tau-\lambda_i)(v_{ij})$, 其中 $\because (x-\lambda_i)^{e_{ij}-1} \in \operatorname{ann}(v_{ij})$, $\therefore (\tau-\lambda_i)^{e_{ij}}(v_{ij}) = 0 \Longrightarrow h(x)v_{ij} = r(\tau)(v_{ij})$ 可由 $\mathcal{B}'_{ij}$ 表示.

综上, 得证. □

有理标准型的存在无需附加条件, 而约当标准型的存在是需要附加条件的: 仅当极小多项式能分解成一次多项式的乘积, 即 $m_\tau(x) = (x-\lambda_1)^{e_1}\cdots(x-\lambda_m)^{e_m}$ 时, 约当标准型才存在.

上面我们看到, $\tau((\tau-\lambda_i)^{e_{ij}-1}(v_{ij})) = \lambda(\tau-\lambda_i)^{e_{ij}}(v_{ij})$, 即 $(\tau-\lambda_i)^{e_{ij}-1}(v_{ij}) \in \mathcal{E}_{\lambda_i}$, 那么, 是否还有其他向量 $\in \mathcal{E}_{\lambda_i}$?

$\{(\tau-\lambda_i)^{e_{ij}-1}(v_{ij}) \mid j = 1, \cdots, t_i\} \subseteq \mathcal{E}_{\lambda_i}$, $\because V_{p_i} = \bigoplus_{j=1}^{t_i}\langle\langle v_{ij}\rangle\rangle$, $\therefore \{(\tau-\lambda_i)^{e_{ij}-1}(v_{ij}) \mid j\}$ 线性无关.

定义 8.6 代数重数: 在 $\mathcal{C}_\tau(x)$ 中 λ_i 作为根的重数, 即 $\dim V_{p_i} = \sum_{j=1}^{t_i} e_{ij}$.

定义 8.7 几何重数: $\dim \mathcal{E}_{\lambda_i} = t_i$.

定理 8.5 (课本定理8.5): 几何重数 $\dim V_{p_i} = \sum_{j=1}^{t_i} e_{ij} \geq$ 代数重数 $\dim \mathcal{E}_{\lambda_i} = t_i$

证: $e_i \geq e_{i1} \geq \cdots \geq e_{it_i} \geq 0$, 故得证. □

几何重数 = 代数重数的特殊情况下, $[\tau]$ 的约当标准型何如?

若几何重数 = 代数重数, 即 $\dim V_{p_i} = \dim \mathcal{E}_{\lambda_i}$, 则 $e_{ij} = 1 \forall j \Longrightarrow e_i = e_{i1} = 1$
$\Longrightarrow m_\tau(\lambda) = (x-\lambda_1)\cdots(x-\lambda_m)$,
此时 $[\tau]_{\mathcal{B}'_{ij}} = \begin{pmatrix}\lambda_i\end{pmatrix}$, $[\tau]_{\mathcal{B}'} = \operatorname{diag}(\lambda_1, \cdots, \lambda_m)$.

定理 8.6 (课本定理8.10, 8.11, 8.18): 下列叙述等价:

τ 可对角化, 即 $\exists$ 一组基 $\mathcal{B}$, s.t. $[\tau]_{\mathcal{B}}$ 为对角阵.

(2) 几何重数 = 代数重数.

(3) $m_\tau(x) = (x-\lambda_1)\cdots(x-\lambda_m)$, λ_i 互不相等.

(4) $V_{p_i} = \mathcal{E}_i$, $V = \mathcal{E}_{\lambda_1} \oplus \cdots \oplus \mathcal{E}_{\lambda_m}$.

(5) 特征向量构成 V 的基.

(6) $\tau = \lambda_1\rho_1 + \cdots + \lambda_k\rho_k$, 其中 $\rho_1 + \cdots + \rho_k = 1$ 为单位分解, $\operatorname{spec}(\tau) = \{\lambda_1, \cdots, \lambda_k\}$, $\operatorname{Im}\rho_i = \mathcal{E}_{\lambda_i}$, $\ker\rho_i = \bigoplus_{j\neq i}\mathcal{E}_{\lambda_j}$.

8.1 投影算子

定义 8.8 投影(算子): 向量空间 $V = S \oplus T$, 映射 $\rho_{ST}: V \to V$, $u_S + u_T \mapsto u_S$, 其中 $u_S \in S$, $u_T \in T$, 则 ρ_{ST} 称为在 S 上沿 T 的**投影(算子)**.

$\ker\rho_{ST} = T$, $\operatorname{Im}\rho_{ST} = S$, $V = \ker\rho_{ST} + \operatorname{Im}\rho_{ST}$.

定理 8.7 (课本第3 版定理2.21): (1) $V = S \oplus T$, 则 $\rho_{ST} + \rho_{TS} = 1_V$ (V 上的恒等变换) 且 $\rho_{ST} \circ \rho_{TS} = \rho_{TS} \circ \rho_{ST} = 0_V$ (V 上的恒等变换).

(2) $\sigma \in \mathcal{L}(V)$, 若 $V = \ker\sigma \oplus \operatorname{Im}\sigma$ 且 $\sigma|_{\operatorname{Im}\sigma} = 1|_{\operatorname{Im}\sigma}$, 其中 $|_{\operatorname{Im}\sigma}$ 代表算子定义域为 $\operatorname{Im}\sigma$, 则 σ 是在 $\operatorname{Im}\sigma$ 上沿 $\ker\sigma$ 的投影.

证: (1) $\forall v \in V$, $\because V = S \oplus T$, $\therefore v = v_S + v_T$, 其中 $v_S \in S$, $v_T \in T$
$\Longrightarrow (\rho_{ST} + \rho_{TS})(v) = (\rho_{ST} + \rho_{TS})(v_S + v_T) = \rho_{ST}(v_S) + \rho_{ST}(v_T) + \rho_{TS}(v_S) + \rho_{TS}(v_T) = v_S + 0 + 0 + v_T = v \Longrightarrow \rho_{ST} + \rho_{TS} = 1$;
$\Longrightarrow \rho_{ST} \circ \rho_{TS}(v) = \rho_{ST}(\rho_{TS}(v_S + v_T)) = \rho_{ST}(\rho_{TS}(v_S) + \rho_{TS}(v_T)) = \rho_{ST}(0 + v_T) = 0 \Longrightarrow \rho_{ST} \circ \rho_{TS} = 0$,
同理, $\rho_{TS} \circ \rho_{ST} = 0$.

(2) $\forall v \in V$, $\because V = \ker\sigma \oplus \operatorname{Im}\sigma$, $\therefore v = v_{\ker} + v_{\operatorname{Im}}$, 其中 $v_{\ker} \in \ker\sigma$, $v_{\operatorname{Im}} \in \operatorname{Im}\sigma$
$\Longrightarrow \sigma(v) = \sigma(v_{\ker} + v_{\operatorname{Im}}) = \sigma(v_{\ker}) + \sigma(v_{\operatorname{Im}}) = \sigma(v_{\ker}) + 1(v_{\operatorname{Im}}) = 0 + v_{\operatorname{Im}}$, 故 σ 为 $\operatorname{Im}\sigma$ 上沿 $\ker\sigma$ 的投影.

□

定理 8.8 (课本定理第3 版2.22): $\rho \in \mathcal{L}(V)$ 为投影 $\Longleftrightarrow \rho^2 = \rho$.

证: "$\Longrightarrow$": $\forall v \in V$, $v = u_S + u_T$, 其中 $u_S \in S$, $u_T \in T$,
$\rho(v) = u_S$, $\rho(\rho(v)) = \rho(u_S) = u_S \Longrightarrow \rho^2 = \rho$.

"$\Longrightarrow$": 首先将 V 分解成 $V = \ker\rho \oplus \operatorname{Im}\rho$:
一方面, $\forall v \in \ker\rho \cap \operatorname{Im}\rho \Longrightarrow v \in \ker\rho \Longleftrightarrow \rho(v) = 0$
且 $v \in \operatorname{Im}\rho \Longleftrightarrow \exists u \in V$, s.t. $v = \rho(u)$
$\Longleftrightarrow 0 = \rho(v) = \rho(\rho(u)) = \rho^2(u) = \rho(u) = v \Longrightarrow \ker\rho \cap \operatorname{Im}\rho = \{0\}$.
另一方面, $\forall v \in V$, $\rho(v) \in \operatorname{Im}\rho$,
$\because \rho(v - \rho(v)) = \rho(v) - \rho(\rho(v)) = \rho(v) - \rho^2(v) = 0$, $\therefore v = (v - \rho(v)) + \rho(v)$, 其中 $v - \rho(v) \in \ker\rho$, $\rho(v) \in \operatorname{Im}\rho \Longrightarrow V = \ker\rho \oplus \operatorname{Im}\rho$.
故 $V = \ker\rho + \operatorname{Im}\rho$.

又有 $\forall \rho(u) \in \operatorname{Im}\rho$, $\rho(\rho(u)) = \rho^2(u) = \rho(u) \Longrightarrow \rho|_{\operatorname{Im}\rho} = 1|_{\operatorname{Im}\rho}$,
由定理 **??** 得, ρ 为在 $\operatorname{Im}\rho$ 上沿 $\ker\rho$ 的投影.

综上, 得证. □

定义 8.9 正交: $\rho, \sigma \in \mathcal{L}(V)$ 为投影, 若 $\rho\sigma = \sigma\rho = 0$, 则称 ρ 与 σ **正交**, 记作 $\rho \perp \sigma$.

$\rho \perp \sigma \Longleftrightarrow \forall v \in V$, $\rho\sigma(v) = \rho(\sigma(v)) = 0$ 且 $\sigma\rho(v) = \sigma(\rho(v)) = 0 \Longleftrightarrow \operatorname{Im}\sigma \subseteq \ker\rho$ 且 $\operatorname{Im}\rho \subseteq \ker\sigma$.

定义 8.10 **单位分解:** $\rho_1+\cdots+\rho_k=1$, 其中 ρ_i 为投影且互相 $\perp$, 则称该式为**单位分解**.

定理 8.9 (课本第3 版定理2.25): (1) $\rho_1+\cdots+\rho_k=1$ 为单位分解, 则 $\operatorname{Im}\rho_1\oplus\cdots\oplus\operatorname{Im}\rho_k=V$.

(2) 若 $V=S_1\oplus\cdots\oplus S_k$, 令 ρ_i 是在 S_i 上沿 $\sum_{j\neq i}S_j$ 的投影, 则 $\rho_1+\cdots+\rho_k=1$ 为单位分解.

证: (1) 先证 V 由 $\{\operatorname{Im}\rho_1,\cdots,\operatorname{Im}\rho_k\}$ 生成: $\forall v\in V$, $v=1(v)=(\rho_1+\cdots+\rho_k)(v)=\rho_1(v)+\cdots+\rho_k(v)$, 其中 $\rho_i(v)\in\operatorname{Im}\rho_i$.

再证 $\operatorname{Im}\rho_i\cap(\cup_{j\neq i}\operatorname{Im}\rho_j)=\{0\}$: $\forall v\in\operatorname{Im}\rho_i\cap(\cup_{j\neq i}\operatorname{Im}\rho_j)\iff v\in\operatorname{Im}\rho_i\iff\rho_i(v)=v$

且 $\exists j\neq i$, s.t. $v\in\operatorname{Im}\rho_j\iff\exists u\in V$, s.t. $\rho_j(u)=v$

$\Longrightarrow v=\rho_i(v)=\rho_i(\rho_j(u))=\rho_i\rho_j(u)=0\Longrightarrow\operatorname{Im}\rho_i\cap\left(\sum_{j\neq i}\operatorname{Im}\rho_j\right)=\{0\}$.

综上, 得证.

(2) $\forall v\in V$, $\because V=S_1\oplus\cdots\oplus S_k$, $\therefore v=v_1+\cdots+v_k$, 其中 $v_i=\rho_i(v)\in S_i$

$\Longrightarrow v=\rho_1(v)+\cdots+\rho_k(v)=(\rho_1+\cdots+\rho_k)(v)\Longrightarrow\rho_1+\cdots+\rho_k=1$.

$\forall v\in V$, $\rho_j(v)\in\operatorname{Im}\rho_j=S_j\subseteq\sum_{j\neq i}S_i\Longrightarrow\rho_j\rho_i(v)=\rho_j(\rho_i(v))=0$, 其中 $i\neq j\Longrightarrow\rho_j\rho_i=0$,

同理, $\rho_j\rho_i=0$, 即 ρ_i 与 ρ_j 正交, 故 $\rho_1+\cdots\rho_k=1$ 为单位分解.

□

若 $V=\mathcal{E}_{\lambda_1}\oplus\cdots\oplus\mathcal{E}_{\lambda_k}$, 令 ρ_i 为在 $\mathcal{E}_{\lambda_i}$ 上沿 $\sum_{j\neq i}\mathcal{E}_{\lambda_i}$ 的投影, 则 $\rho_1+\cdots+\rho_k=1$ 是单位分解, $\rho_i|_{\mathcal{E}_{\lambda_i}}=1|_{\mathcal{E}_{\lambda_i}}\Longrightarrow\forall v\in\mathcal{E}_{\lambda_i}$, $\tau(v)=\lambda_i v=\lambda\rho_i(v)$, 即在 $\mathcal{E}_{\lambda_i}$ 上, $\tau|_{\mathcal{E}_{\lambda_i}}=\lambda\rho_i$, 从而引出定理 **??** (6).