矩阵多项式

定义 设 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 t + a_0$ 是关于x在$P[x]$上的x的多项式. 若A是方阵, 则称

$$f(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I$$

为方阵A的多项式.

定义 设 A是方阵, 定义如下映射φ:

$$\varphi_A(f(x)) = f(A), \ \forall f(x) \in P[x], \ A \in P^{n\times n}.$$

则 φ_A 是由$P[x]$到方阵集$\{A \mid A \in R^{n\times n}\}$上的线性映射.

令 $\ker\varphi_A = \{f(x) \mid f(A) = 0\}$.

定义 (最低多项式) 若$\ker\varphi_A \neq 0$, 以$d_A(x)$表示$\ker\varphi_A$中次数最低的首一多项式, 称为A的最低多项式.

则我们有:

Claim: $\boxed{\begin{array}{l} f(x) \in \ker\varphi(A) \text{ 当且仅当 } d_A(x) \mid f(x). \\ \text{, 即 } \ker\varphi_A = \{d_A(x) g(x) \mid g(x) \in P[x]\}. \end{array}}$

证明: 以$q(x)$, $r(x)$表示$f(x)$除以$d_A(x)$的商式与余式,

由$f(A) = d_A(A) q(A) + r(A) = r(A)$.

所以$f(x) \in \ker\varphi_A$当且仅当$f(x) \in \ker\varphi_A$.

但 $d_A(x)$ 为$\ker\varphi_A$的次数最低者, 若$r(x) \neq 0$,

则由$\deg r(x) < \deg d_A(x)$. 这不可能, 故$r(x) = 0$.

定理 (Hamilton-Caylay) 设 $A\in P^{n\times n}$, $f(\lambda)=\det(\lambda I_n-A)$ $=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$, 则

$$f(A)=A^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0I_n=0$$

证明：略：

详见《高等代数与解析几何》下册

孟道骥 P33

推论：$d_A(\lambda)$ 是 A 的最低多项式，则我们有：$\deg d_A(\lambda)\le n$

下面开始我们开始探讨在线性代数中留下的问题：

定理：记 $S_i=Nul\{A-\lambda_i I\}$ 其中 λ_i 为 A 的特征值，$i=1,\cdots,k$.

$\lambda_1,\cdots,\lambda_k$ 互不相等，A 是数域 P 上 $n\times n$ 方阵

则若 V 是 P 上 n 维线性空间，则

$$V=\bigoplus_{i=1}^{k}S_i=S_1\oplus S_2\oplus\cdots\oplus S_k\ (\text{注}\ \oplus\ \text{这里指直和})$$

$\Leftarrow$ 当且仅当 $\Rightarrow$

A 的最低多项式 $d_A(\lambda)$ 为不同的一次因式的积，

即 $d_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$

证明：① 设 $V=\bigoplus_{i=1}^{k}S_i$

令 $d(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$.

Let $d_i(\lambda)=\prod_{j\ne i}(\lambda-\lambda_j)$，于是 $d(\lambda)=(\lambda-\lambda_i)d_i(\lambda)$

设 $\forall\vec{x}\in V$，于是 $\vec{x}=\vec{x}_1+\cdots+\vec{x}_k$，$\vec{x}_i\in S_i$

因而 Let $d(A)=(A-\lambda_iI)d_i(A)$，$d_i(A)=\prod_{j\ne i}(A-\lambda_jI)$

则 $d(A)\vec{x}=\sum_{i=1}^{k}d_i(A)(A-\lambda_iI)\vec{x}_i=0$

$\therefore d(\lambda)\in\operatorname{Ker}\varphi_A$. 故 $d_A(\lambda)\mid d(\lambda)$.

又因为 $\forall i,1\le i\le k$，取 $\vec{x}_i\in S_i$，$\vec{x}_i\ne 0$，则

$$d_i(A)\vec{x}_i=\prod_{j\ne i}(A-\lambda_jI)\vec{x}_i=\prod_{j\ne i}(\lambda_i-\lambda_j)\vec{x}_i\ne 0$$

$\therefore d_A(\lambda)=d(\lambda)=(\lambda-\lambda_1)\cdots(\lambda-\lambda_k)$.

② 若 $d_A(\lambda)=(\lambda-\lambda_1)\cdots(\lambda-\lambda_k)$.

令 $d_i(\lambda)=\prod_{j\neq i}(\lambda-\lambda_j)$，$1\le i\le k$，

于是 $d_1(\lambda),\cdots,d_k(\lambda)$ 互素. 故存 $u_i(\lambda)\in P[\lambda]$ 使得

$u_1(\lambda)d_1(\lambda)+\cdots+u_k(\lambda)d_k(\lambda)=1$，即 $d_1(\lambda)u_1(\lambda)+\cdots+d_k(\lambda)u_k(\lambda)=1$.

令 $V_i=\{\ d_i(A)u_i(A)\vec{\alpha}\ \mid \vec{\alpha}\in V\}$.

则显然 V_i 是 V 的子空间，且由 $\forall\vec{\alpha}\in V$

$$\vec{\alpha}=I\vec{\alpha}=\sum_{i=1}^{k} d_i(A)u_i(A)\vec{\alpha}$$

知 $\vec{\alpha}\in V_1+V_2+\cdots+V_k$

又因为

$$(A-\lambda_i I)(d_i(A)u_i(A))\vec{\alpha}=d_A(A)u_i(A)\vec{\alpha}=0$$

于是 $V_i\subseteq S_i$ ，$i=1,\cdots,k$

因此我们有

$$V=\sum_{i=1}^{k}V_i\subseteq S_1\oplus S_2\oplus\cdots\oplus S_k\subseteq V$$

故 $V=S_1\oplus S_2\oplus\cdots\oplus S_k$

推论：(1) $V=\bigoplus_{i=1}^{k}S_i \overset{\text{推}}{\Longleftrightarrow}$ A 是可以被对角化

(2) 若 λ_i 为 A 的特征值，则 λ_i 为 A 的特征多项式的 $\dim S_i$ 重根

(3) V 为 P 上的线性空间，A 是 P 数域上的方阵，若 A 的特征多项式在 P 中有 $\dim V$ 个不同的根，则 A 可以对角化.

上述定理依赖于最低多项式，即最低多项式是

$$d(\lambda) = (\lambda-\lambda_1)\cdots(\lambda-\lambda_s),$$

λ_s 为矩阵A所有不同值的特征值。

若最低多项式不能写成那样，又会怎样？

<u>定理</u>：设 V 是复数域 C 上 n 维线性空间，$f(\lambda)$ 为矩阵A的特征多项式，且有因式分解：

$$f(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$$

其中 $\lambda_i \in C$，$i\neq j$ 时，$\lambda_i \neq \lambda_j$，则 V 可分解为以下子空间 $R_{\lambda_i}(A)$ 的直和：

$$V = R_{\lambda_1}(A) \oplus R_{\lambda_2}(A) \oplus \cdots \oplus R_{\lambda_s}(A).$$

其中 $R_{\lambda_i}(A) = \cancel{\mathrm{Ker}}\ \mathrm{Nul}\,(A-\lambda_i I)^{n_i}$

$$= \{\vec{\alpha} \in V \mid (A-\lambda_i I)^{n_i}\vec{\alpha} = 0\}.$$

证明：令 $f_i(\lambda) = \prod_{j\neq i}(\lambda-\lambda_j)^{n_j}$，$i=1,\cdots,s$，

则 $f(\lambda) = (\lambda-\lambda_i)^{n_i} f_i(\lambda)$.

可证 $(f_1(\lambda), f_2(\lambda), \cdots, f_s(\lambda)) = 1$，因而 $\exists\, u_i(\lambda) \in C[\lambda]$，

s.t. $\sum_{i=1}^{s} u_i(\lambda) f_i(\lambda) = 1$，从而 $\forall \vec{\alpha} \in V$，有

$$\vec{\alpha} = I\vec{\alpha} = \sum_{i=1}^{s} u_i(A) f_i(A)\vec{\alpha}$$

而 $(A-\lambda_i I)^{n_i} u_i(A) f_i(A)\vec{\alpha} = u_i(A) f(A)\vec{\alpha} = 0$

即 $u_i(A) f_i(A)\vec{\alpha} \in R_{\lambda_i}(A)$，因而

$$V = R_{\lambda_1}(A) + R_{\lambda_2}(A) + \cdots + R_{\lambda_s}(A).$$

设 $\vec{\beta}_i \in R_{\lambda_i}(A)$，$1\le i\le s$，且

$$\vec{\beta}_1 + \vec{\beta}_2 + \cdots + \vec{\beta}_s = 0$$

由 $(\lambda-\lambda_j)^{n_j} \mid f_i(\lambda)$，$i\neq j$，故 $f_i(A)\vec{\beta}_j = 0$，故 $f_i(A)\vec{\beta}_i = 0$

又 $(f_i(\lambda), (\lambda-\lambda_i)^{n_i}) = 1$，故有 $u(\lambda), v(\lambda) \in C[\lambda]$, s.t.

$$u(\lambda) f_i(\lambda) + v(\lambda)(\lambda-\lambda_i)^{n_i} = 1$$

因而 $\vec{\beta}_i = I\vec{\beta}_i$

$$= u(A) f_i(A)\vec{\beta}_i + v(A)(A-\lambda_i I)^{n_i}\beta_i$$

$$= 0$$

于是 $V = R_{\lambda_1}(A) \oplus R_{\lambda_2}(A) \oplus \cdots \oplus R_{\lambda_s}(A)$

定理 2. $Nul(A-\lambda_i I) \subseteq Nul(A-\lambda_i I)^{n_i}$

其中 A为C上矩阵(方阵)，其特征多项式为

$$f(\lambda) = (\lambda-\lambda_1)^{n_1} \cdots (\lambda-\lambda_s)^{n_s}, \quad \lambda_i \in C$$

定理 $\dim R_{\lambda_i}(A) = n_i$