第十五讲 投影矩阵与 Moore-Penrose 逆

一、投影算子与投影矩阵

设 L，M 为$\mathbf{C}^n$的子空间并构成直和$\mathbf{L}+\mathbf{M}=\mathbf{L}\oplus\mathbf{M}=\mathbf{C}^n$.即

$\forall \mathbf{x}\in\mathbf{C}^n$，$\exists$唯一的$\mathbf{y}\in\mathbf{L},\mathbf{z}\in\mathbf{M}$使 x=y+z

称 y 为 x 沿着 M 到 L 的投影。

1. 定义：将任意$\mathbf{x}\in\mathbf{C}^n$变为其沿着 M 到 L 的投影的变换称为沿着 M 到 L 的投影算子，记为$\mathbf{P}_{\mathbf{L},\mathbf{M}}$即$\mathbf{P}_{\mathbf{L},\mathbf{M}}\ \mathbf{x}=\mathbf{y}\in\mathbf{L}$,投影算子是线性变换，其矩阵称为投影矩阵,仍记为$\mathbf{P}_{\mathbf{L},\mathbf{M}}$。

2. 充要条件

引理：设 n 阶方阵 E 为幂等矩阵，则$\mathbf{N(E)}=\mathbf{R(I-E)}$

证明：

$$\because E^2=E\rightarrow E(I-E)=O\rightarrow\forall x\in C^n,E[(I-E)x]=0$$

$$\rightarrow E[R(I-E)]=0\Rightarrow R(I-E)\subseteq N(E)$$

另一方面$\forall x\in N(E)$,即$Ex=0$，则

$$x=Ix-O=Ix-Ex=(I-E)x\in R（I-E）$$

$$\Rightarrow N（E）\subseteq R（I-E）$$

$$\therefore N（E）=R（I-E）$$

定理：n 阶方阵 P 成为投影矩阵的充要条件是 P 为幂等矩阵。

证明：充分性

$P^2=P,\forall x\in C^n$,令$y=Px\in R(P),z=(I-P)x\in R（I-P）=N（P）$。

若 $R（P）\bigcap N（P）=\{0\}$,则$P=P_{R（P）,N（P）}$确为投影矩阵，下面证之

$$\forall x\in R(P)\bigcap N(P),$$

一方面，因$\mathbf{x}\in\mathbf{R(P)}$,存在$\mathbf{u}\in\mathbf{C}^{\mathbf{n}}$使$\mathbf{x}=\mathbf{Pu}$

另一方面$\mathbf{x}\in\mathbf{N(P)}$,即$\mathbf{Px}=\mathbf{0}$.但$\mathbf{Px}=\mathbf{P}^2\mathbf{u}=\mathbf{Pu}=\mathbf{x}\to\mathbf{x}=\mathbf{0}$

$$\Rightarrow \mathbf{R(P)}\cap\mathbf{N(P)}=\{\mathbf{0}\}。$$ （P=I和P不等于I）

必要性 $\mathbf{P}=\mathbf{P}_{\mathbf{L},\mathbf{M}}$ 故 $\forall\mathbf{x}\in\mathbf{C}^{\mathbf{n}}$，$\exists$唯一分解 $\mathbf{y}\in\mathbf{L},\mathbf{z}\in\mathbf{M}$ 使

$$\mathbf{x}=\mathbf{y}+\mathbf{z}\text{且}\ \mathbf{Px}=\mathbf{y}$$

$$\to\mathbf{P}^2\mathbf{x}=\mathbf{P}\underset{\substack{\uparrow\\ \mathbf{y}=\mathbf{y}+\mathbf{0}}}{\mathbf{y}}=\mathbf{y}=\mathbf{Px}\xrightarrow{\mathbf{x}\text{任意}}\mathbf{P}^2=\mathbf{P}$$

3. 投影矩阵的构造

设已知$\mathbf{C}^{\mathbf{n}}$的子空间 L、M 构成直和$\mathbf{L}\oplus\mathbf{M}=\mathbf{C}^{\mathbf{n}}$,下面构造$\mathbf{P}_{\mathbf{L},\mathbf{M}}$。

取 L 的一个基$\left\{\mathbf{x_1,x_2\cdots x_r}\right\}$(设 L 为 r 维子空间)，M 的一个基

$\left\{\mathbf{y_1,y_2\cdots y_{n-r}}\right\}$（则 M 的维数为 n-r）。由直和关系知

$\left\{\mathbf{x_1,x_2\cdots x_r;\ y_1,y_2\cdots y_{n-r}}\right\}$即构成$\mathbf{C^n}$的一个基。故，如令

$$\mathbf{X}=\left[\mathbf{x_1,x_2\cdots x_r}\right],\quad \mathbf{Y}=\left[\mathbf{y_1,y_2\cdots y_{n-r}}\right]$$

则[X　Y]为可逆方阵。另一方面

$$\mathbf{x_i\in L\to P_{L,M}x_i=x_i;y_i\in M\to P_{L,M}y_i=0}$$

即　$\mathbf{P_{L,M}}\left[\mathbf{X\ Y}\right]=\left[\mathbf{X\ O}\right]\to\mathbf{P_{L,M}}=\left[\mathbf{X\ O}\right]\left[\mathbf{X\ Y}\right]^{-1}$

可见，$\mathbf{P_{L,M}}$的秩为 r（$\mathbf{rank}$（$\mathbf{P_{L,M}}$）$=\mathbf{dim\,R}$（$\mathbf{P_{L,M}}$）$=\mathbf{dim\,L}$）

二、正交投影算子与正交投影矩阵

L 为 $\mathbf{C}^n$ 的子空间，其正交补空间 $\mathbf{L}^{\perp}=\left\{\mathbf{x}\,\middle|\,(\mathbf{x},\ \mathbf{y})=\mathbf{0},\mathbf{x}\in\mathbf{C}^n,\mathbf{y}\in\mathbf{L}\right\}$（无特别声明取 $\mathbf{x}^H\mathbf{y}$）

1. 定义：设 L 是 $\mathbf{C}^n$ 的子空间，则称沿着 $\mathbf{L}^{\perp}$ 到 L 的投影算子 $\mathbf{P}_{\mathbf{L},\mathbf{L}^{\perp}}$ 为正交投影算子，简记为 $\mathbf{P}_{\mathbf{L}}$。正交投影算子的矩阵称为正交投影矩阵，仍记为 $\mathbf{P}_{\mathbf{L}}$。

2. 充要条件：n 阶方阵 P 为正交投影矩阵的充要条件是 P 为幂等的厄米矩阵。

证明：首先证明两个引理:

（1）对 n 阶方阵 A，$\forall\mathbf{x}\in\mathbf{C}^n$ 均有 $\mathbf{x}^H\mathbf{A}\mathbf{x}=\mathbf{0}$ 则 A=0， （2）N（$\mathbf{P}^H$）$=\mathbf{R}^{\perp}(\mathbf{P})$

(1)证明：设 A=$(a_{ij})_{n\times n}$取 $x=\left[0\cdots0\cdots1_{(\text{第}i\text{个})}\cdots0\right]^T$ 则 $x^HAx=a_{ii}=0$

再取 $x=\left[0\cdots0,\ \xi_i,0\cdots0,\ \xi_j,0\cdots0\right]^T$（$i\neq j$），注意到$a_{ii}=0,$

则$x^HAx=\overline{\xi_i}a_{ij}\xi_j+\overline{\xi_j}a_{ji}\xi_i$

$$\to\begin{cases}\xi_i=\xi_j=1,\text{则}a_{ij}+a_{ji}=0\\ \xi_i=1,\xi_j=\sqrt{-1},\text{则}a_{ij}-a_{ji}=0\end{cases}\to a_{ij}=a_{ji}=0\text{统一考虑即}A=0$$

（2）证明：$\forall x\in N(P^H)$,即$P^Hx=0\to x^HP=0\to$

$\forall y\in C^n$，$x^H(Py)=0$

由y的任意性，知$x\perp R$（P）$\Rightarrow N$（P^H）$\subseteq R^{\perp}(P)$

另一方面，设$x\in R^{\perp}(P)$即$\forall y\in C^n$,均有$x^H(Py)=0\to x^HP=0$

$$\to P^H x = 0 \to x \in N(P^H) \Rightarrow N(P^H) \supseteq R^{\perp}(P)$$

所以 $N(P^H) = R^{\perp}(P)$

现在证明该充要条件。

充分性：

$$P^2 = P,\quad P^H = P \to P = P_{R(p),\,N(p)} = P_{R(p),\,N(p^H)} = P_{R(p),\,R^{\perp}(p)} = P_{R(p)}$$

必要性：

$$P = P_L$$

$$\forall x \in C^n, \text{可唯一地分解成} y = Px \in L, z = (I - P)x \in L^{\perp} \text{使} x = y + z$$

又 $y \in L, z \in L^{\perp} \to y^H z = 0 \to x^H P^H (I - P)x = 0 \xrightarrow{x\text{任意}} P^H (I - P) = 0$

$P^H = P^H P = (P^H P)^H = (P^H)^H = P$， P 为厄米矩阵。

幂等已由上一定理得知。

3. 正交投影矩阵的构造

设 L 的一个基为 $\{\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_r\}$， $\mathbf{L}^{\perp}$ 的一个基为 $\{\mathbf{y}_1,\mathbf{y}_2\cdots\mathbf{y}_{n-r}\}$。则

$$\mathbf{x}_i^H\mathbf{y}_j=\mathbf{0}$$

令 $\mathbf{X}=\left[\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_r\right]$， $\mathbf{Y}=\left[\mathbf{y}_1,\mathbf{y}_2\cdots\mathbf{y}_{n-r}\right]$ 则 $\mathbf{X}^H\mathbf{Y}=\mathbf{0}$

$$(\mathbf{A}^{-1}=(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H)$$

$$\mathbf{P_L = [X\ O][X\ Y]^{-1} = [X\ O]\left\{[X\ Y]^H[X\ Y]\right\}^{-1}[X\ Y]^H}$$

$$\mathbf{= [X\ O]\begin{bmatrix} X^HX & X^HY \\ Y^HX & Y^HY \end{bmatrix}^{-1}[X\ Y]^H = [X\ O]\begin{bmatrix} X^HX & 0 \\ 0 & Y^HY \end{bmatrix}^{-1}\begin{bmatrix} X^H \\ Y^H \end{bmatrix}}$$

$$\mathbf{= \begin{bmatrix} X(X^HX)^{-1} & O \end{bmatrix}\begin{bmatrix} X^H \\ Y^H \end{bmatrix} = X(X^HX)^{-1}X^H}$$

三、投影矩阵与广义逆矩阵

(i) $\mathbf{AXA = A \rightarrow AX = P_{R(A),N(AX)}}$

（ii） $\mathbf{XAX = X \rightarrow XA = P_{R(X),N(XA)}}$

$$\mathbf{\begin{cases} AXA = A \\ (AX)^H = AX \end{cases} \rightarrow AX = P_{R(A)}}$$

$$\begin{cases} XAX = X \\ (XA)^H = XA \end{cases} \rightarrow XA = P_{R(X)}$$

Moore 定义：设 $A \in C^{m\times n}, X \in C^{n\times m}$ 且 $AX = P_{R(A)}$，$XA = P_{R(X)}$

则 X 为 A 的 Moore 广义逆矩阵。事实上，Moore 广义逆矩阵正是 $A^\dagger$

作业 P295 1、4