第二章 基础知识补充

2.1 投影矩阵(projection matrix)

定义：矩阵 P称为投影矩阵，若：

1. P是对称的，即$P' = P$；
2. P是幂等的，即$P^2 = P$。

幂等矩阵的性质：

1). 特征值非 0 即 1；

2). P幂等则$tr(P) = rank(P)$；

3). P幂等$\Leftrightarrow rank(P) + rank(I_n - P) = n$。

投影矩阵的性质：

1. 投影矩阵是非负定的；
2. 若P_1, P_2是投影矩阵且$P_1 - P_2$非负定，则$P_1P_2 = P_2P_1 = P_2$，且$P_1 - P_2$也是投影矩阵。

2.2 广义逆

对相容性线性方程

$$A_{m\times n}x = b$$

若 $rank(A) = m = n$ 则方程有唯一解 $x = A^{-1}b$。若A不是方阵或者是奇异方阵，Penrose 指出方程的解可以先求解矩阵方程

$$A_{m\times n}B_{n\times m}A_{m\times n} = A_{m\times n}$$

矩阵B称为A的广义逆，记为A^-。

定义 1：对矩阵$A_{m\times n}$，一切满足方程$AXA = A$的矩阵X，称为矩阵A的广义逆，记为A^-，即$AA^-A = A$。

定理 1：设$rank(A_{m\times n}) = r$，若A表为

$$A = P_{m\times m}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q_{n\times n},$$

其中P, Q可逆，则

$$A^- = Q^{-1}\begin{pmatrix} I_r & B \\ C & D \end{pmatrix}P^{-1}$$

这里B, C, D为适当阶数的任意矩阵。

推论 1：

1. 任何矩阵A的广义逆A^-总存在；
2. A^-唯一$\Leftrightarrow A$可逆；
3. $rank(A^-) \ge rank(A) = rank(A^-A) = rank(AA^-)$

设矩阵$A_{m\times n}$，A的列向量所张成的空间记为$\mu(A)$，即$\mu(A) = \{Ax \mid x \in R^n\}$。

基本性质：

1.$\mu(A) \subset \mu(B) \Leftrightarrow \exists C, A = BC$；

2.$\dim \mu(A) = rank(A)$。

定理 2：$\mu(A') = \mu(A'A)$。

定理 3：对任何矩阵A

1). $A(A'A)^-A'$ 与 $(A'A)^-$ 的取值无关且 $rank\left(A(A'A)^-A'\right) = rank(A)$；

2). $A(A'A)^-A'A = A, \quad A'A(A'A)^-A' = A'$。

定理 4：设线性方程$A_{m\times n}x=b$是相容的，A^{-}为A给定的一个广义逆，则

1. $x=A^{-}b$即为方程的一个解；
2. 齐次方程 $Ax=0$ 的所有解为 $x=(I_n-A^{-}A)z$，其中z为任意$n\times 1$向量；
3. 方程 $Ax=b$ 的所有解为 $x=A^{-}b+(I_n-A^{-}A)z$。

推论 4：相容性方程$Ax=b(b\neq 0)$的所有解为$\{x\,|\,x=A^{-}b, A^{-}为A任一广义逆\}$。

2.3 正交投影

R^n 中两个向量 x,y 的内积定义为 $(x,y)=x'y$，若 $x'y=0$，则称 $x\perp y$。若 $S\subset R^n$为线性子空间，$\forall y\in S, x\perp y$，则称$x\perp S$。令$S^{\perp}=\{x\,|\,x\perp S\}$，则$S^{\perp}$也是线性子空间，称为$S$的正交补，易见 $S\cap S^{\perp}=\{0\}, S\oplus S^{\perp}=R^n, (S^{\perp})^{\perp}=S$。

例 1：令$S=\mu(A_{n\times m})\subset R^n, rank(A)=m$，则$S^{\perp}=\mu(B)$，这里$B=I_n-A(A'A)^{-1}A'$。

设$rank(A_{n\times m})=r$，若$n\times(n-r)$矩阵B满足1.$A'B=0$；2.$rank(B)=n-r$，则称矩阵B为A的正交补，记$B=A^{\perp}$。从定义易见$A^{\perp}$是所有使得$A'B=0$的B秩最大的矩阵。

例 2：设$n\times m$矩阵A，则

$$\mu(A^{\perp})=\mu(A)^{\perp}, \mu(A^{\perp})\oplus\mu(A)=R^n。$$

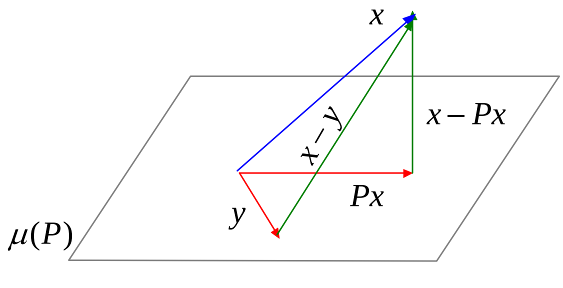
设$x\in R^n$，$S\subset R^n$为线性子空间，x有唯一分解$x=y+z$，$y\in S, z\in S^{\perp}$，称 y为 x在子空间 S 上的正交投影 (orthogonal projection)。若矩阵$P_{n\times n}$满足对$\forall x\in R^n$，其在子空间S上的正交投影$y=Px$，则称 P为S上的正交投影矩阵(简称投影矩阵)。

定理 1：$\mu(A_{n\times m})$上的正交投影矩阵为$P_A=A(A'A)^{-}A'$。

定理 2：P为正交投影矩阵$\Leftrightarrow P$对称幂等。

定理 3：$P_{n\times n}$为投影矩阵$\Leftrightarrow\forall x\in R^n$，$\|x-Px\|=\inf\limits_{y\in\mu(P)}\|x-y\|$。

几何意义



2.4 多元正态分布

定义 1：随机向量$X=(x_1,\cdots,x_n)'$称为n维正态随机向量，如果其密度函数为

$$f(X)=\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)\right]$$

这里$\mu=(\mu_1,\cdots,\mu_2)'$，$\Sigma>0$。记$X\sim N_n(\mu,\Sigma)$。

容易计算若$X\sim N_n(\mu,\Sigma)$，则$EX=\mu, Var(X)=\Sigma$。

例 1：二维正态分布密度函数

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\cdot\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

相当于$\mu=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$，$\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$。

定义 2：设X为 n维随机向量，若存在非随机的$A_{n\times r}$，$rank(A)=r$以及$\mu_{n\times 1}$使得$X=AU+\mu$，其中$U\sim N_r(0,I_r)$，则称X服从均值为μ、协方差阵为$\Sigma=AA'$的多元正态分布，记为$X\sim N(\mu,\Sigma)$。

此定义把多元正态向量定义为若干个相互独立的一维标准正态分布的线性变换。协方差阵$\Sigma\geq 0$(不一定可逆)。

设$X\sim N(\mu,\Sigma)$，则X的特征函数为

$$\varphi(t)=Ee^{it'X}=\exp\left(it'\mu-\frac{1}{2}t'\Sigma t\right)$$

由于特征函数唯一决定分布函数，因此也可用特征函数来定义多元正态分布，避免$|\Sigma|=0$的情形。

性质：

1.多元正态分布的任意边际分布为相应的多元正态分布；

2.多元正态分布的任线性变换仍然是正态分布，即$X\sim N(\mu,\Sigma)$，则$Y=AX+b\sim N(A\mu+b,A\Sigma A')$；

3. 设$(X_1',X_2')'$为多元正态分布，$Cov(X_1,X_2)=0\Leftrightarrow X_1,X_2$独立；

4.$X\sim N(\mu,\Sigma)\Leftrightarrow\forall t, t'X\sim N(t'\mu,t'\Sigma t)$；

5.设$X\sim N(\mu,\Sigma)$，$\Sigma>0$作相应的分块

$$X=\begin{pmatrix}X_1\\ X_2\end{pmatrix},\mu=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$$

则给定X_2时X_1的条件分布仍是多元正态分布且条件期望

$$E\left(X_1\middle|X_2\right)=\mu_1+\Sigma_{12}\Sigma_{22}^{-1}(X_2-\mu_2),$$

条件方差

$$Var\left(X_1\middle|X_2\right)=\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}。$$

例 2：设$(x_1,x_2,x_3,x_4)'$联合分布为零均值的正态分布，则
$$Ex_1x_2x_3x_4 = Ex_1x_2Ex_3x_4 + Ex_1x_3Ex_2x_4 + Ex_1x_4Ex_2x_3$$

2.5 正态随机向量的二次型

设$X_{n\times1}$为随机向量，$A_{n\times n}$为对称矩阵，称$X'AX$为随机向量X的二次型。

定理 1：设随机向量X的均值为 μ，协方差阵为 Σ，则$E(X'AX) = tr(A\Sigma) + \mu'A\mu$。

定义 1：设$X \sim N_n(\mu, I_n)$，称随机变量$y = X'X$的分布是服从自由度为 n的非中心参数为$\lambda = \mu'\mu$的χ^2分布，记为$y \sim \chi^2_{n,\lambda}$；当$\lambda = 0$时称为中心$\chi^2$分布记为$y \sim \chi^2_n$。

$y \sim \chi^2_{n,\lambda}$的密度函数为
$$f_\lambda(y) = e^{-\frac{\lambda+y}{2}} \sum_{k=0}^{\infty} \frac{\lambda^k y^{\frac{n}{2}+k-1}}{k! 2^{\frac{n}{2}+2k} \Gamma\left(\frac{n}{2}+k\right)}, y > 0。$$

χ^2分布的基本性质：

1. 设$y_i \sim \chi^2_{n_i,\lambda_i}$，$i = 1,\cdots,k$且$y_i$相互独立，则$\sum_{i=1}^{k} y_i \sim \chi^2_{\sum_{i=1}^{k} n_i, \sum_{i=1}^{k}\lambda_i}$；
2. $E(\chi^2_{n,\lambda}) = n + \lambda$，$Var(\chi^2_{n,\lambda}) = 2n + 4\lambda$；
3. 设$y \sim \chi^2_{n,\lambda}$，则$y$的特征函数为
$$\varphi(t) = (1-2it)^{-\frac{n}{2}} e^{\frac{i\lambda t}{1-2it}}$$

定理 2：设$X \sim N_n(\mu, I_n)$，$A_{n\times n}$对称，则$X'AX \sim \chi^2_{r,\mu'A\mu} \Leftrightarrow A$幂等且$rank(A) = r$。

推论 2：设$X \sim N_n(\mu, \Sigma)$，$\Sigma > 0$，$A_{n\times n}$对称，则$X'AX \sim \chi^2_{r,\mu'A\mu}$当且仅当下述之一成立

1. $A\Sigma$幂等且$rank(A) = r$；
2. ΣA幂等且$rank(A) = r$；
3. Σ为A的一个广义逆且$rank(A) = r$。

例 3：设$X \sim N_n(A\beta, \sigma^2 I_n)$，$rank(A) = r$，则$\frac{X'(I_n - A(A'A)^- A')X}{\sigma^2} \sim \chi^2_{n-r}$。

定理 3：设$X \sim N_n(\mu, I_n)$，A，A_1对称且$X'AX = X'A_1X + X'A_2X \sim \chi^2_{r,\mu'A\mu}$，$X'A_1X \sim \chi^2_{s,\mu'A_1\mu}$，$A_2 \ge 0$则：

1. $X'A_2X \sim \chi^2_{r-s,\mu'A_2\mu}$；
2. $X'A_1X$与$X'A_2X$独立；
3. $A_1A_2 = 0$。

推论 3：

1. 设 $X \sim N_n(\mu, I_n)$，A_1, A_2 对称且 $X'A_1X, X'A_2X$为χ^2分布，则$X'A_1X, X'A_2X$相互独立$\Leftrightarrow A_1A_2 = 0$；
2. 设 $X \sim N_n(\mu, \Sigma), \Sigma > 0$ A, A_1 对称且 $X'AX = X'A_1X + X'A_2X \sim \chi^2_{r,\lambda}$，$X'A_1X \sim \chi^2_{s,\lambda_1}$，$A_2 \ge 0$则$X'A_2X \sim \chi^2_{r-s,\lambda_2}$，$X'A_1X$与$X'A_2X$独立且$A_1\Sigma A_2 = 0$。

2.6 正态随机向量的线性形式与二次型之间的独立性

设$X \sim N_n(\mu,\Sigma)$，A,B为对称矩阵，C为$m\times n$矩阵，本节研究二次型$X'AX$，$X'BX$之间以及与线性形式CX之间独立性问题。

定理 1：设$X \sim N_n(\mu,I_n)$，A对称，若$CA=0$，则CX与$X'AX$独立。

推论 1：设$X \sim N_n(\mu,\Sigma),\Sigma>0$，$A$对称，若$C\Sigma A=0$，则$CX$与$X'AX$独立。

例 4：设$x_i \sim N(\mu_i,\sigma^2)$彼此相互独立，令$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$，则$\bar{x}$与$s^2$独立。

定理 2：设$X \sim N_n(\mu,I_n)$，A,B对称，若$AB=0$，则$X'AX$与$X'BX$独立。

推论 2：设$X \sim N_n(\mu,\Sigma),\Sigma>0$，$A,B$对称，若$A\Sigma B=0$，则$X'AX$与$X'BX$独立。

2.7 多元 t 分布(Multivariate t)

一维 t 分布：设x,y独立且$x\sim N(\delta,1)$，$y\sim\chi_n^2$，称$t=x/\sqrt{\frac{y}{n}}$的分布称为自由度为 n 非中心参数为δ的 t 分布，记为$t\sim t_{n,\delta}$；$\delta=0$称为中心的 t 分布；$t\sim t_{n,\delta}$的密度函数为：

$$f(t|n,\delta)=\frac{n^{\frac{n}{2}}e^{-\frac{\delta^2}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)(n+t^2)^{\frac{n+1}{2}}}\sum_{i=0}^{\infty}\Gamma\left(\frac{n+i+1}{2}\right)\frac{(\delta t)^i}{i!}\left(\frac{2}{n+t^2}\right)^{\frac{i}{2}}$$

当$\delta=0$时，t_n的密度为

$$f(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}。$$

定义 1：设随机向量$X_{p\times1}$有密度

$$f(X)=\frac{\Gamma\left(\frac{n+p}{2}\right)|B|^{\frac{1}{2}}}{(n\pi)^{\frac{p}{2}}\Gamma\left(\frac{n}{2}\right)}\left[1+\frac{(X-\mu)'B(X-\mu)}{n}\right]^{-\frac{n+p}{2}}$$

则称X的分布为p元 t 分布或称 t 向量，记为$X\sim t_p(\mu,B^{-1},n)$。

注：n为自由度，μ为位置参数，B相当于刻度参数。

例 5：设$X\sim N_p(0,\Sigma)$，$y\sim\chi_n^2$且X与Y独立，则$t=X/\sqrt{\frac{y}{n}}\sim t_p(0,\Sigma,n)$。