第5章　假设检验及其他

吴密霞

北京工业大学统计与数据科学系

E-mail: *wumixia@bjut.edu.cn*

教材

- 吴密霞, 王松桂. 2024. 线性模型引论（第2版）, 科学出版社.

线性假设的检验

先回顾以下似然比检验(likelihood ratio test).

设随机向量$\boldsymbol{y}$服从参数为$\boldsymbol{\theta} \in \Theta$ 的概率分布族, 考虑参数检验：

$$H_0\text{：}\ \boldsymbol{\theta} \in \Theta_0;\quad H_1\text{：}\ \boldsymbol{\theta} \bar{\in} \Theta_0,$$

这里Θ_0为Θ的一个子集.

记$L(\boldsymbol{\theta}; \boldsymbol{y})$ 为似然函数, $\widehat{\boldsymbol{\theta}}$为$\boldsymbol{\theta}$ 的ML估计. $\widehat{\boldsymbol{\theta}}_H$是原假设$H_0$：$\boldsymbol{\theta} \in \Theta_0$成立时$\boldsymbol{\theta}$的约束ML估计. 于是

$$\sup_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta};\ \boldsymbol{y}) = L(\widehat{\boldsymbol{\theta}};\ \boldsymbol{y}),$$

$$\sup_{\boldsymbol{\theta} \notin \Theta_0} L(\boldsymbol{\theta};\ \boldsymbol{y}) = L(\widehat{\boldsymbol{\theta}}_H;\ \boldsymbol{y}),$$

似然比定义为

$$\lambda(\boldsymbol{y}) = \frac{\sup\limits_{\boldsymbol{\theta}\in\Theta} L(\theta;\ \boldsymbol{y})}{\sup\limits_{\boldsymbol{\theta}\in\Theta_0} L(\boldsymbol{\theta};\ \boldsymbol{y})} = \frac{L(\widehat{\boldsymbol{\theta}};\ \boldsymbol{y})}{L(\widehat{\boldsymbol{\theta}}_H;\ \boldsymbol{y})}.$$

显然, $\lambda(\boldsymbol{y}) \geq 1$, 因为$L(\widehat{\boldsymbol{\theta}}_H; \boldsymbol{y})$ 是原假设成立时, 观察到样本$\boldsymbol{y}$ 的可能性的一个度量, 当在$\lambda(\boldsymbol{y})$ 比较大时, 则$L(\widehat{\boldsymbol{\theta}}_H;\ \boldsymbol{y})$ 相对较小, 即原假设成立观察到样本点y的可能性较小, 自然地, 在$\lambda(\boldsymbol{y})$ 较大时拒绝原假设, 于是取检验的拒绝域形为

$$\{\boldsymbol{y}:\ \ \lambda(\boldsymbol{y}) \geq c\},$$

这里c是一个待定常数. 称以上检验为似然比检验.

线性假设的检验

正态线性模型

$$y = X\beta + e, \qquad e \sim N_n(0,\ \sigma^2 I_n), \tag{5.1}$$

考虑线性假设

$$H_0 : H\beta = d \leftrightarrow H_1 : H\beta \neq d \tag{5.2}$$

的检验问题, 这里$\mathrm{rk}(\mathbf{X}) = r$, $\mathbf{H}$ 为$k \times p$ 的矩阵, 满足条件:

$$\mathrm{rk}(\mathbf{H}) = k,\ \ \mathcal{M}(\mathbf{H}') \subset \mathcal{M}(\mathbf{X}').$$

该检验问题可看作是对模型参数线性约束条件的$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{d}$ 检验, 也可被应用于回归方程显著性、回归系数显著性等问题的研究.

记$\boldsymbol{\mu} = \mathbf{X}\beta$. 似然函数$(\boldsymbol{\mu},\ \sigma^2)$为

$$L(\boldsymbol{\mu},\ \sigma^2) = (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{\mu}\|^2}{2\sigma^2}\right).$$

由定理4.1.5 和定理4.3.3 得$(\boldsymbol{\mu}, \sigma^2)$的ML估计和约束ML估计:

$$\widehat{\boldsymbol{\mu}} = \mathbf{X}\widehat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\boldsymbol{y}, \qquad \tilde{\sigma}^2 = \frac{\|\boldsymbol{y}-\mathbf{X}\widehat{\beta}\|^2}{n},$$

$$\widehat{\boldsymbol{\mu}}_{\mathbf{H}} = \mathbf{X}\widehat{\beta}_{\mathbf{H}} = \mathbf{X}\widehat{\beta} - \mathbf{A}(\mathbf{H}\widehat{\beta} - \boldsymbol{d}), \qquad \tilde{\sigma}^2_{\mathbf{H}} = \frac{\|\boldsymbol{y}-\mathbf{X}\widehat{\beta}_{\mathbf{H}}\|^2}{n},$$

这里, $\mathbf{A} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}$. 于是

$$\sup_{H_0} L(\boldsymbol{\mu}, \sigma^2) = L(\mathbf{X}\widehat{\beta}_{\mathbf{H}}, \tilde{\sigma}^2_{\mathbf{H}}) = \left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}} \|\boldsymbol{y}-\mathbf{X}\widehat{\beta}_{\mathbf{H}}\|^{-n},$$

$$\sup_{H_1} L(\boldsymbol{\mu}, \sigma^2) = L(\mathbf{X}\widehat{\boldsymbol{\beta}}, \tilde{\sigma}^2) = \left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}} \|\boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\|^{-n}.$$

似然比为

$$\lambda = \frac{\sup_{H_0} L(\boldsymbol{\mu}, \sigma^2)}{\sup_{H_1} L(\boldsymbol{\mu}, \sigma^2)} = \left(\frac{\|\boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_{\mathbf{H}}\|^2}{\|\boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\|^2}\right)^{n/2}.$$

记$\mathrm{SS}_e = \|\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta}\|^2$，$\mathrm{SS}_{He} = \|\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta}_{\mathbf{H}}\|^2$ 分别表示模型残差平方和与在约束$\mathbf{H}\boldsymbol{\beta} = \mathbf{0}$下的残差平方和. 记

$$F = \frac{n-r}{k}\left(\lambda^{2/n} - 1\right) = \frac{(\mathrm{SS}_{He} - \mathrm{SS}_e)/k}{\mathrm{SS}_e/(n-r)}.$$

统计量F 仅依赖于λ 且为λ的严增函数.

定理5.1.1

设$\mathbf{H}$为$k \times p$的矩阵, 满足条件: $\mathrm{rk}(\mathbf{H}) = k$, $\mathcal{M}(\mathbf{H}') \subset \mathcal{M}(\mathbf{X}')$. 则

(1) $\mathrm{SS}_e \sim \sigma^2 \chi^2_{n-r}$;

(2) $\mathrm{SS}_{\mathbf{H}e} - \mathrm{SS}_e = (\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d})'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}(\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d}) \sim \sigma^2 \chi^2_{k,\delta}$;

(3) $\mathrm{SS}_{\mathbf{H}e} - \mathrm{SS}_e$与$\mathrm{SS}_e$相互独立;

(4) 当线性假设$H_0 : \mathbf{H}\boldsymbol{\beta} = \boldsymbol{d}$为真时, $F \sim F_{k,\ n-r}$,

这里, $r = \mathrm{rk}(\mathbf{X})$, $\delta = (\mathbf{H}\boldsymbol{\beta} - \boldsymbol{d})'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}(\mathbf{H}\boldsymbol{\beta} - \boldsymbol{d})/\sigma^2$.

证明 (1)的证明见定理4.1.5.

(2) 记$\mathbf{P}_{\mathbf{X}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$, $\mathbf{A} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}$.

注意到$(\mathbf{I}_n - \mathsf{P}_{\mathbf{X}})\mathbf{A} = \mathbf{0}$, $\boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}} = (\mathbf{I}_n - \mathsf{P}_{\mathbf{X}})\boldsymbol{y}$, 于是

$$\begin{aligned} \mathrm{SS}_{\mathbf{H}e} &= \|\boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_{\mathbf{H}}\|^2 \\ &= \|\boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{A}(\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d})\|^2 \\ &= \mathrm{SS}_e + (\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d})'\mathbf{A}'\mathbf{A}(\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d}). \end{aligned}$$

因为$\mathcal{M}(\mathbf{H}') \subset \mathcal{M}(\mathbf{X}')$, 所以$\mathbf{A}'\mathbf{A} = (\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}$. 故证得

$$\mathrm{SS}_{\mathbf{H}e} - \mathrm{SS}_e = (\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d})'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}(\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d}).$$

再利用正态随机向量的线性变换性质, 有

$$\mathbf{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{d} \sim N_k\left(\mathbf{H}\boldsymbol{\beta} - \boldsymbol{d}, \sigma^2\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}'\right), \tag{5.3}$$

结合$\mathrm{rk}(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}') = k$ 和推论3.4.3推出(2).

(3) 注意到$\mathrm{SS}_{\mathbf{H}e}-\mathrm{SS}_e$为$\mathbf{H}\widehat{\beta}$的函数,

$$\mathrm{SS}_e=\|\boldsymbol{y}-\mathbf{X}\widehat{\beta}\|^2=\boldsymbol{y}'(\mathbf{I}_n-\mathsf{P}_{\mathbf{X}})\boldsymbol{y},\quad \mathbf{H}\widehat{\beta}=\mathbf{H}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\boldsymbol{y},$$

$$\mathbf{H}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'(\mathbf{I}_n-\mathsf{P}_{\mathbf{X}})=\mathbf{0},$$

利用推论3.5.1 立得SS_e和$\mathbf{H}\widehat{\beta}$ 独立, 进而SS_e和$\mathrm{SS}_{He}-\mathrm{SS}_e$独立.

(4) 是(1)–(3)及F分布定义的直接推论. 定理证毕.

线性假设$\mathbf{H}\beta=\boldsymbol{d}$的似然比检验统计量的另一个表达式为

$$F=\frac{(\mathrm{SS}_{\mathbf{H}e}-\mathrm{SS}_e)/k}{\mathrm{SS}_e/(n-r)}=\frac{(\mathbf{H}\widehat{\beta}-\boldsymbol{d})'(\mathbf{H}(\mathbf{X}'\mathbf{X})^-\mathbf{H}')^{-1}(\mathbf{H}\widehat{\beta}-\boldsymbol{d})/k}{\mathrm{SS}_e/(n-r)}.$$

- 依似然比检验方法, 对于给定的显著性水平α, 若

$$F > F_{k,\ n-r}(\alpha),\ \text{或}\ \ P(F_{k,\ n-r} > F) < \alpha,$$

则拒绝原假设$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{d}$; 否则接受原假设$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{d}$, 这里$F_{k,\ n}(\alpha)$表示自由度为$k,\ n$ 的F分布的上侧α 分点.

- F检验的功效函数(power function)为

$$\psi = P\left(F > F_{k,\ n-r}(\alpha)\right) = 1 - \int_{-\infty}^{F_{k,\ n-r}(\alpha)} f_{k,\ n-r,\ \delta}(t)dt,$$

这里$f_{k,n-r,\delta}(x)$表示自由度为$k,\ n-r$、非中心参数为δ的F分布的概率密度函数, $\delta = (\mathbf{H}\boldsymbol{\beta} - \boldsymbol{d})'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}(\mathbf{H}\boldsymbol{\beta} - \boldsymbol{d})/\sigma^2$.

检验统计量F也可以由Wald方法直接构造.

- 当原假设$H_0 : \mathbf{H}\beta = \boldsymbol{d}$成立时, 则(5.3)成立, 故

$$\frac{(\mathbf{H}\widehat{\beta} - \boldsymbol{d})(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}(\mathbf{H}\widehat{\beta} - \boldsymbol{d})}{\sigma^2} \sim \chi_k^2,$$

- 用$\widehat{\sigma}^2 = \mathrm{SS}_e/(n-r)$替代$\sigma^2$, 再除以分子的自由度$k$, 便得统计量$F$.

令$\boldsymbol{d} = \mathbf{0}$, 由定理5.1.1 立得推论.

推论5.1.1

对于相容齐次线性假设$\mathbf{H}\beta = \mathbf{0}$, $\mathrm{rk}(\mathbf{H}) = k$, $\mathcal{M}(\mathbf{H}') \subset \mathcal{M}(\mathbf{X}')$, 当$\mathbf{H}\beta = \mathbf{0}$为真时, 有

$$F = \frac{\widehat{\beta}'\mathbf{H}'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}')^{-1}\mathbf{H}\widehat{\beta}/k}{\mathrm{SS}_e/(n-r)} \sim F_{k,\, n-r}.$$

F检验统计量的另一种形式

当$\mathbf{d}=\mathbf{0}$时，模型的残差平方和与约束条件$\mathbf{H}\beta=\mathbf{d}$下的残差平方和可表达为

$$\mathrm{SS}_e = \|\mathbf{y}-\mathbf{X}\beta\|^2 = \mathbf{y}'\mathbf{y}-\widehat{\beta}'\mathbf{X}'\mathbf{y},$$

$$\mathrm{SS}_{\mathbf{H}e} = \|\mathbf{y}-\mathbf{X}\beta_H\|^2 = \mathbf{y}'\mathbf{y}-\widehat{\beta}'_{\mathbf{H}}\mathbf{X}'\mathbf{y}.$$

F统计量具有形式

$$F = \frac{(\mathrm{RSS}(\beta)-\mathrm{RSS}_{\mathbf{H}}(\beta))/k}{\mathrm{SS}_e/(n-r)}. \tag{5.4}$$

于是F统计量的分子为增加了约束条件$\mathbf{H}\beta=\mathbf{0}$之后，回归平方和所减少的量除以$k$，而$k$作为分子的自由度，等于线性假设$\mathbf{H}\beta=\mathbf{0}$所含的独立方程的个数.

例 分块模型

$$\boldsymbol{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{e}, \quad \boldsymbol{e} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I}_n),$$

其中$(\mathbf{X}_1, \mathbf{X}_2)$ 是列满秩的矩阵检验假设$H_0: \boldsymbol{\beta}_2 = \mathbf{0}$. 若$\boldsymbol{\beta}_2 = \mathbf{0}$ 成立, 原模型就可降维为$\boldsymbol{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{e}$, $\boldsymbol{e} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$. 此时$\boldsymbol{\beta}_1$的LS估计为$\widetilde{\boldsymbol{\beta}}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\boldsymbol{y}$.

检验统计量(5.4)为

$$F = \frac{\widehat{\boldsymbol{\beta}}'\mathbf{X}'\boldsymbol{y} - \widetilde{\boldsymbol{\beta}}_1'\mathbf{X}_1'\boldsymbol{y}}{\mathrm{SS}_e/(n-r)} = \frac{\widehat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{N}_{\mathbf{X}_1}\boldsymbol{y}/p_2}{\mathrm{SS}_e/(n-r)},$$

这里, $\mathbf{N}_{\mathbf{X}_1} = \mathbf{I}_n - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-}\mathbf{X}_1'$, $\widehat{\boldsymbol{\beta}}_2 = (\mathbf{X}_2'\mathbf{N}_{\mathbf{X}_1}\mathbf{X}_2')^{-1}\mathbf{X}_2\mathbf{N}_{\mathbf{X}_1}\boldsymbol{y}$.

例 模型同质性检验(homogeneity test).

检验两个模型

$$\boldsymbol{y}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{e}_1, \qquad \boldsymbol{e}_1 \sim N_{n_1}(\mathbf{0},\ \sigma^2\mathbf{I}_{n_1}),$$

$$\boldsymbol{y}_2 = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{e}_2, \qquad \boldsymbol{e}_2 \sim N_{n_2}(\mathbf{0},\ \sigma^2\mathbf{I}_{n_2})$$

是否为同模型, 即检验$H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, 问题归结为

$$\mathbf{H}\begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} = (\mathbf{I}_p, -\mathbf{I}_p)\begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} = \mathbf{0},$$

其中$\mathbf{H} = (\mathbf{I}_p, -\mathbf{I}_p)$.

各自模型下的LS估计, 即

$$\widehat{\beta}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\boldsymbol{y}_1, \quad \widehat{\beta}_2 = (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\boldsymbol{y}_2.$$

异质模型的残差平方和为

$$\mathrm{SS}_e = \boldsymbol{y}_1'\boldsymbol{y}_1 + \boldsymbol{y}_2'\boldsymbol{y}_2 - \widehat{\beta}_1'\mathbf{X}_1'\boldsymbol{y}_1 - \widehat{\beta}_2'\mathbf{X}_2'\boldsymbol{y}_2.$$

于是在H_0下同质模型的LS估计

$$\widehat{\beta}_{\mathbf{H}} = (\mathbf{X}_1'\mathbf{X}_1 + \mathbf{X}_2'\mathbf{X}_2)^{-1}(\mathbf{X}_1'\boldsymbol{y}_1 + \mathbf{X}_2'\boldsymbol{y}_2).$$

原模型的约束残差平方和为

$$\mathrm{SS}_{\mathbf{H}e} = \boldsymbol{y}_1'\boldsymbol{y}_1 + \boldsymbol{y}_2'\boldsymbol{y}_2 - \widehat{\beta}_{\mathbf{H}}'(\mathbf{X}_1'\boldsymbol{y}_1 + \mathbf{X}_2'\boldsymbol{y}_2).$$

于是

$$\begin{aligned}\mathrm{SS}_{\mathbf{H}e}-\mathrm{SS}_e &= \widehat{\boldsymbol{\beta}}_1'\mathbf{X}_1'\boldsymbol{y}_1+\widehat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\boldsymbol{y}_2-\widehat{\boldsymbol{\beta}}_{\mathbf{H}}'(\mathbf{X}_1'\boldsymbol{y}_1+\mathbf{X}_2'\boldsymbol{y}_2)\\ &=(\widehat{\boldsymbol{\beta}}_1-\widehat{\boldsymbol{\beta}}_{\mathbf{H}})'\mathbf{X}_1'\boldsymbol{y}_1+(\widehat{\boldsymbol{\beta}}_2-\widehat{\boldsymbol{\beta}}_{\mathbf{H}})'\mathbf{X}_2'\boldsymbol{y}_2,\end{aligned}$$

得到检验统计量

$$F=\frac{(\mathrm{SS}_{\mathbf{H}e}-\mathrm{SS}_e)/p}{\mathrm{SS}_e/(n-2p)}=\frac{((\widehat{\boldsymbol{\beta}}_1-\widehat{\boldsymbol{\beta}}_{\mathbf{H}})'\mathbf{X}_1'\boldsymbol{y}_1+(\widehat{\boldsymbol{\beta}}_2-\widehat{\boldsymbol{\beta}}_{\mathbf{H}})'\mathbf{X}_2'\boldsymbol{y}_2)/p}{\mathrm{SS}_e/(n-2p)}.$$

对给定的显著性水平α, 若

$$F>F_{p,\ n-2p}(\alpha),$$

则拒绝原假设$\boldsymbol{\beta}_1=\boldsymbol{\beta}_2$, 即认为两模型不是同一个线性回归模型.

置信域和同时置信区间

正态线性模型

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{e}, \quad \boldsymbol{e} \sim N_n(\mathbf{0},\ \sigma^2\mathbf{I}_n), \tag{5.5}$$

这里$\mathrm{rk}(\mathbf{X}) = r$. 讨论

(a) 单个可估函数$\boldsymbol{h}'\boldsymbol{\beta}$的置信区间;

(b) 可估函数向量$\mathbf{H}\boldsymbol{\beta} = (\boldsymbol{h}_1'\boldsymbol{\beta}, \cdots, \boldsymbol{h}_k'\boldsymbol{\beta})$的置信域以及多个可估函数的同时置信区间.

- 不失一般性, 假设$\mathbf{H} = (\boldsymbol{h}_1, \cdots, \boldsymbol{h}_k)'$ 为$k \times p$的行满秩矩阵,

$$\mathrm{rk}(\mathbf{H}) = k, \quad \mathcal{M}(\mathbf{H}') \subset \mathcal{M}(\mathbf{X}').$$

- 可估函数$\boldsymbol{h}'\boldsymbol{\beta}$ 置信区间
  - 由定理4.1.4知：$\boldsymbol{h}'\widehat{\boldsymbol{\beta}} \sim N(\boldsymbol{h}'\boldsymbol{\beta}, \sigma^2\boldsymbol{h}'(\mathbf{X}'\mathbf{X})^-\boldsymbol{h})$; $(n-r)\widehat{\sigma}^2/\sigma^2 \sim \chi^2_{n-r}$, 且两者独立. 于是

$$\frac{\boldsymbol{h}'\widehat{\boldsymbol{\beta}} - \boldsymbol{h}'\boldsymbol{\beta}}{\widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}}} \sim t_{n-r},$$

这里, $\widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}} = \widehat{\sigma}\sqrt{\boldsymbol{h}'(\mathbf{X}'\mathbf{X})^-\boldsymbol{h}}$ 为$\boldsymbol{h}'\widehat{\boldsymbol{\beta}}$的标准差的估计. 由于

$$P\left(\frac{|\boldsymbol{h}'\widehat{\boldsymbol{\beta}} - \boldsymbol{h}'\boldsymbol{\beta}|}{\widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}}} \le t_{n-r}\left(\frac{\alpha}{2}\right)\right) = 1-\alpha,$$

可估函数$\boldsymbol{h}'\boldsymbol{\beta}$的置信系数为$1-\alpha$的置信区间

$$\left[\boldsymbol{h}'\widehat{\boldsymbol{\beta}} - t_{n-r}\left(\frac{\alpha}{2}\right)\widehat{\sigma}_{h'\widehat{\boldsymbol{\beta}}}\,,\quad \boldsymbol{h}'\widehat{\boldsymbol{\beta}} + t_{n-r}\left(\frac{\alpha}{2}\right)\widehat{\sigma}_{h'\widehat{\boldsymbol{\beta}}}\right]. \tag{5.6}$$

- 可估函数向量$\mathbf{H}\boldsymbol{\beta} = (\boldsymbol{h}_1'\boldsymbol{\beta}, \cdots, \boldsymbol{h}_k'\boldsymbol{\beta})'$的置信域
  - 记$\boldsymbol{\Phi} = \mathbf{H}\boldsymbol{\beta}$, 则$\widehat{\boldsymbol{\Phi}} = \mathbf{H}\widehat{\boldsymbol{\beta}}$ 为$\boldsymbol{\Phi}$的LS估计, 有
$$(\widehat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})'\mathbf{V}^{-1}(\widehat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}) \sim \sigma^2\chi^2_k.$$
  - 由于$\frac{(n-r)\widehat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-r}$, 且与$\widehat{\boldsymbol{\Phi}}$相互独立, 于是
$$\frac{(\widehat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})'\mathbf{V}^{-1}(\widehat{\boldsymbol{\Phi}} - \Phi)}{k\widehat{\sigma}^2} \sim F_{k,\ n-r}.$$

$\boldsymbol{\Phi}$的置信系数为$1-\alpha$的置信域

$$\mathcal{D} = \left\{\boldsymbol{\Phi}:\ (\widehat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})'\mathbf{V}^{-1}(\widehat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}) \leq k\widehat{\sigma}^2 F_{k,\ n-r}(\alpha)\right\} \tag{5.7}$$

多个可估函数的同时置信区间

考虑多个线性不相关的可估函数$\boldsymbol{h}_1'\boldsymbol{\beta}, \cdots, \boldsymbol{h}_k'\boldsymbol{\beta}$ 的同时置信区间

其中$\boldsymbol{h}_1, \cdots \boldsymbol{h}_k$线性不相关. 记$\mathbf{H} = (\boldsymbol{h}_1, \cdots, \boldsymbol{h}_k)'$, $\mathbf{V} = \mathbf{H}(\mathbf{X}'\mathbf{X})^{-}\mathbf{H}'$,

$\boldsymbol{\Phi} = (\boldsymbol{h}_1'\boldsymbol{\beta}, \cdots, \boldsymbol{h}_k'\boldsymbol{\beta})'$, $\boldsymbol{\Phi} = \mathbf{H}\boldsymbol{\beta}$, $\widehat{\boldsymbol{\Phi}} = \mathbf{H}\widehat{\boldsymbol{\beta}} = (\boldsymbol{h}_1'\widehat{\boldsymbol{\beta}}, \cdots, \boldsymbol{h}_k'\widehat{\boldsymbol{\beta}})$, 则

$$\mathrm{rk}(\mathbf{H}) = k, \quad \mathcal{M}(\mathbf{H}') \subset \mathcal{M}(\mathbf{X}'), \quad \mathbf{V} > \mathbf{0}.$$

引理5.2.1

设$\boldsymbol{a}$和$\boldsymbol{b}$均为$n \times 1$的向量, $\mathbf{A}$为$n \times n$正定方阵, 则

$$\sup_{\boldsymbol{b} \neq \mathbf{0}} \frac{(\boldsymbol{a}'\boldsymbol{b})^2}{\boldsymbol{b}'\mathbf{A}\boldsymbol{b}} = \boldsymbol{a}'\mathbf{A}^{-1}\boldsymbol{a}.$$

1. Scheffé 同时置信区间

所有可估函数$\boldsymbol{h}'\boldsymbol{\beta}$, $\boldsymbol{h} \in \mathcal{M}(\mathbf{H}')$的Scheffé置信区间$\mathcal{I}_{\boldsymbol{h}}$, 即其满足

$$P\left\{\boldsymbol{h}'\boldsymbol{\beta} \in \mathcal{I}_{\boldsymbol{h}},\ \text{对于一切}\boldsymbol{h} \in \mathcal{M}(\mathbf{H}')\right\} \geq 1-\alpha. \tag{5.8}$$

对一切可估函数$\boldsymbol{h}'\boldsymbol{\beta}$, 都有

$$T_{\boldsymbol{h}} = \frac{(\boldsymbol{h}'\widehat{\boldsymbol{\beta}} - \boldsymbol{h}'\boldsymbol{\beta})^2}{\widehat{\sigma}^2(\boldsymbol{h}'(\mathbf{X}'\mathbf{X})^-\boldsymbol{h})} \sim F_{1,n-r},$$

故可通过寻找$T_{\boldsymbol{h}}$的共同上界c 导出置信区间$\mathcal{I}_{\boldsymbol{h}} = \{\boldsymbol{h}'\boldsymbol{\beta};\ T_{\boldsymbol{h}} \leq c\}$, 其中$c$满足如下条件:

$$P\left\{T_{\boldsymbol{h}} \leq c,\quad \text{对于一切}\boldsymbol{h} \in \mathcal{M}(\mathbf{H}')\right\} = 1-\alpha, \tag{5.9}$$

等价于

$$P\left\{\sup_{\boldsymbol{h}\in\mathcal{M}(\mathbf{H}')} T_{\boldsymbol{h}} \le c\right\} = P\left\{\sup_{\boldsymbol{h}\in\mathcal{M}(\mathbf{H}')} \frac{(\boldsymbol{h}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}))^2}{\widehat{\sigma}^2(\boldsymbol{h}'(\mathbf{X}'\mathbf{X})^-\boldsymbol{h})} \le c\right\} = 1-\alpha.$$

因为$\boldsymbol{h} \in= \mathcal{M}(\mathbf{H}')$, 所以存在$k\times 1$的向量$\boldsymbol{b}$, 使得$\boldsymbol{h} = \mathbf{H}'\boldsymbol{b}$. 于是

$$\begin{aligned}\sup_{\boldsymbol{h}\in\mathcal{M}(\mathbf{H}')} \frac{(\boldsymbol{h}'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}))^2}{\widehat{\sigma}^2\boldsymbol{h}'(\mathbf{X}'\mathbf{X})^-\boldsymbol{h}} &= \sup_{\boldsymbol{b}\in\mathbb{R}^k} \frac{(\boldsymbol{b}'(\widehat{\boldsymbol{\Phi}}-\boldsymbol{\Phi}))^2}{\widehat{\sigma}^2\boldsymbol{b}'\mathbf{V}\boldsymbol{b}} \\ &= \frac{(\widehat{\boldsymbol{\Phi}}-\boldsymbol{\Phi})'\mathbf{V}^{-1}(\widehat{\boldsymbol{\Phi}}-\boldsymbol{\Phi})}{\widehat{\sigma}^2}.\end{aligned}$$

只要取$c = kF_{k,\, n-r}(\alpha)$, 就可是使得(5.9)成立.

定理5.2.1

对于正态线性模型(5.5), 若$\mathrm{rk}(\mathbf{H}) = k$, $\mathcal{M}(\mathbf{H}') \subset \mathcal{M}(\mathbf{X}')$, 则对一切可估函数$\boldsymbol{h}'\beta$, $\boldsymbol{h} \in \mathcal{M}(\mathbf{H}')$, 其置信系数为$1-\alpha$的Scheffé 同时置信区间为

$$\left[\boldsymbol{h}'\widehat{\boldsymbol{\beta}} - \widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}}\sqrt{kF_{k,n-r}(\alpha)},\ \ \boldsymbol{h}'\widehat{\boldsymbol{\beta}} + \widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}}\sqrt{kF_{k,n-r}(\alpha)}\right]. \tag{5.10}$$

这里$\widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}} = \widehat{\sigma}\sqrt{\boldsymbol{h}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{h}}$.

注 若$k = r = \mathrm{rk}(\mathbf{X})$, 则我们得到所有可估函数$\boldsymbol{h}'\beta$的同时置信区间

$$\left[\boldsymbol{h}'\widehat{\boldsymbol{\beta}} - \widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}}\sqrt{rF_{r,n-r}(\alpha)},\ \ \boldsymbol{h}'\widehat{\boldsymbol{\beta}} + \widehat{\sigma}_{\boldsymbol{h}'\widehat{\boldsymbol{\beta}}}\sqrt{rF_{r,n-r}(\alpha)}\right].$$

对于$\mathbf{H} = (\boldsymbol{h}_1, \cdots, \boldsymbol{h}_k)'$非列满秩的情形.

推论5.2.1

对于正态线性模型(5.5), k个可估函数$\boldsymbol{h}_i'\beta$, $i = 1, \cdots, k$的Scheffé 同时置信区间为

$$\left[\boldsymbol{h}_i'\widehat{\boldsymbol{\beta}} - \widehat{\sigma}_{\boldsymbol{h}_i'\widehat{\boldsymbol{\beta}}}\sqrt{k_0 F_{k_0, n-r}(\alpha)},\ \boldsymbol{h}_i'\widehat{\boldsymbol{\beta}} + \widehat{\sigma}_{\boldsymbol{h}_i'\widehat{\boldsymbol{\beta}}}\sqrt{k_0 F_{k_0, n-r}(\alpha)}\right], \quad i = 1, \cdots, k,$$

这里k_0为$\boldsymbol{h}_1, \cdots, \boldsymbol{h}_k$的最大线性无关组的个数, 即$k_0 = \mathrm{rk}(\mathbf{H})$, 其中$\mathbf{H} = (\boldsymbol{h}_1, \cdots, \boldsymbol{h}_k)'$.

由于Scheffé同时置信区间是对所有可估函数$\boldsymbol{h}'\beta$, $\boldsymbol{h} \in \mathcal{M}(\mathbf{H}')$, 所以置信区间长度会偏大些.

2. Bonferroni同时置信区间

考虑k个可估函数$\boldsymbol{h}_i'\boldsymbol{\beta}$, $i=1,\cdots,k$的Bonferroni同时置信区间

$$\mathcal{I}_i(\alpha/k)=\left[\boldsymbol{h}_i'\widehat{\boldsymbol{\beta}}-t_{n-r}\left(\frac{\alpha}{2k}\right)\widehat{\sigma}_{\boldsymbol{h}_i'\widehat{\boldsymbol{\beta}}},\quad \boldsymbol{h}_i'\widehat{\boldsymbol{\beta}}+t_{n-r}\left(\frac{\alpha}{2k}\right)\widehat{\sigma}_{\boldsymbol{h}_i'\widehat{\boldsymbol{\beta}}}\right],\quad i=1,\cdots,k.$$

证明 应用Bonferroni不等式可证.

引理5.2.2 (Bonferroni不等式)

设$E_i,\cdots,E_k$为k个随机事件, $P(E_i)=1-\alpha_i$, $i\in\{1,\cdots,k\}$. 则

$$P\left(\bigcap_{i=1}^{k}E_i\right)=1-P\left(\bigcup_{i=1}^{k}\bar{E}_i\right)\geq 1-\sum_{i=1}^{k}P(\bar{E}_i)=1-\sum_{i=1}^{k}\alpha_i.$$

- 令$\alpha_i = \alpha/k$, $E_i = \{\boldsymbol{h}_i'\boldsymbol{\beta} \in \mathcal{I}_i(\alpha/k)\}$, 由Bonferroni不等式立得

$$P\left(\boldsymbol{h}_i'\boldsymbol{\beta} \in \mathcal{I}_i(\alpha),\ i=1,\cdots,k\right) = 1 - P\left(\bigcup_{i=1}^{k} \bar{E}_i\right) \geq 1 - kP\left(\bar{E}_1\right) = 1-\alpha.$$

注 对于固定的k个可估函数$\boldsymbol{h}_i'\boldsymbol{\beta}$, $i=1,\cdots,k$, 如果

$$t_{n-r}\left(\frac{\alpha}{2k}\right) < \sqrt{k_0 F_{k_0,n-r}(\alpha)}, \tag{5.11}$$

则采用Bonferroni同时置信区间，其区间长度更短, 精度更高, 反之, 采用Scheffé同时置信区间, 这里$k_0 = \mathrm{rk}(\boldsymbol{h}_1,\cdots,\boldsymbol{h}_k)$.

预测

- 所谓预测, 就是对指定的自变量的值, 预测对应的因变量所可能取的值.

  在线性模型中, 自变量往往代表一组试验条件或生产条件或社会经济条件, 由于试验或生产等方面的费用或试验周期长的原因, 在我们根据以往积累的数据获得经验模型后, 希望对一些感兴趣的试验、生产条件不真正去做试验, 而利用经验模型就对应的因变量的取值作出合理的估计和分析, 可见, 预测是普遍存在着的一个很有意义的实际问题.
- 和估计一样, 预测也有点预测和区间预测之分.

点预测

- 假设历史数据服从线性模型

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{e}, \quad E(\boldsymbol{e}) = \mathbf{0}, \quad \mathrm{Cov}(\boldsymbol{e}) = \sigma^2\boldsymbol{\Sigma} > \mathbf{0}, \tag{5.12}$$

这里$\boldsymbol{y}$为$n \times 1$观测向量, $\mathbf{X}$为$n \times p$的矩阵, $\mathrm{rk}(\mathbf{X}) = r$, $\boldsymbol{\Sigma}$已知.

- 假设要预测m个点$\boldsymbol{x}_{0i} = (x_{0i1}, \cdots, x_{0ip})'$, 所对应的因变量$y_{0i}$, $i = 1, \cdots, k$的值, 且已知y_{0i}和历史数据服从同一个线性模型, 即

$$y_{0i} = \boldsymbol{x}_{0i}'\boldsymbol{\beta} + \varepsilon_{0i}, \qquad i = 1, \cdots, k.$$

采用矩阵形式, 则这个模型变为

$$\boldsymbol{y}_0 = \mathbf{X}_0\boldsymbol{\beta} + \boldsymbol{\varepsilon}_0, \quad E(\boldsymbol{\varepsilon}_0) = \mathbf{0}, \quad \mathrm{Cov}(\boldsymbol{\varepsilon}_0) = \sigma^2\boldsymbol{\Sigma}_0, \tag{5.13}$$

点预测

$$\boldsymbol{y}_0 = \begin{pmatrix} y_{01} \\ \vdots \\ y_{0k} \end{pmatrix}, \quad \mathbf{X}_0 = \begin{pmatrix} x_{011} & \cdots & x_{01p} \\ \vdots & & \vdots \\ x_{0k1} & \cdots & x_{0kp} \end{pmatrix}, \quad \boldsymbol{\varepsilon}_0 = \begin{pmatrix} \varepsilon_{01} \\ \vdots \\ \varepsilon_{0k} \end{pmatrix}.$$

记预测量与历史数据的协方差阵为Cov$(\boldsymbol{y},\boldsymbol{y}_0) = \sigma^2\mathbf{V}'$. 故

$$\text{Cov}\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{y}_0 \end{pmatrix} = \text{Cov}\begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{\varepsilon}_0 \end{pmatrix} = \sigma^2 \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{V}' \\ \mathbf{V} & \boldsymbol{\Sigma}_0 \end{pmatrix}. \quad (5.14)$$

- 假设$\mathcal{M}(\mathbf{X}_0') \subset \mathcal{M}(\mathbf{X}')$, 因为被预测量的均值$\mathbf{X}_0\beta$首先需可估.

点预测

- 用$E(\boldsymbol{y}_0) = \mathbf{X}_0\boldsymbol{\beta}$ 的估计来预测$\boldsymbol{y}_0$, 即

$$\boldsymbol{y}_0^* = \mathbf{X}_0\boldsymbol{\beta}^* = \mathbf{X}_0(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}.$$

可以证明: $\boldsymbol{y}_0^*$是$\boldsymbol{y}_0$的无偏预测, 即$E(\boldsymbol{y}_0^*) = E(\boldsymbol{y}_0)$.

注 点估计和点预测意义不同

- $\widehat{\boldsymbol{y}}_0$的预测误差: $\boldsymbol{z} = \boldsymbol{y}_0^* - \boldsymbol{y}_0$; $\mathbf{X}_0\boldsymbol{\beta}^*$的估计误差: $\boldsymbol{d} = \mathbf{X}_0\boldsymbol{\beta}^* - \mathbf{X}_0\boldsymbol{\beta}$. 假设被预测量与历史数据不相关, 即$\mathrm{Cov}(\boldsymbol{e}, \boldsymbol{\varepsilon}_0) = \mathbf{0}$, 此时

$$\mathrm{Cov}(\boldsymbol{y}_0^* - \boldsymbol{y}_0) = \mathrm{Cov}(\boldsymbol{y}_0^*) + \mathrm{Cov}(\boldsymbol{y}_0) = \sigma^2(\mathbf{X}_0(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}_0' + \boldsymbol{\Sigma}_0).$$

$$\mathrm{Cov}(\boldsymbol{d}) = \mathrm{Cov}(\mathbf{X}_0\boldsymbol{\beta}^*) = \sigma^2\mathbf{X}_0(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}_0 \leq \mathrm{Cov}(\boldsymbol{z}).$$

差别来源于被预测量$\boldsymbol{y}_0$为随机向量, 而被估计量$\boldsymbol{\mu}_0$为常数向量.

点预测

定义5.3.1

设$\tilde{y}_0$为y_0的一个预测, 则$\tilde{y}_0$的广义预测均方误差(PMSE)为

$$\mathrm{PMSE}(\tilde{y}_0) = E(\tilde{y}_0 - y_0)'\mathbf{A}(\tilde{y}_0 - y_0),$$

这里$\mathbf{A} > \mathbf{0}$. 特别地, 取$\mathbf{A} = \mathbf{I}_m$, 则

$$\mathrm{PMSE}(\tilde{y}_0) = E(\tilde{y}_0 - y_0)'(\tilde{y}_0 - y_0)$$

为$\tilde{y}_0$的预测均方误差.

- y_0的线性无偏预测类$\mathcal{D}_0$表达为如下形式:

$$\mathcal{D}_0 = \{\tilde{y}_0 = \mathbf{C}y;\ \text{其中}\mathbf{C}\text{满足条件}\mathbf{CX} = \mathbf{X}_0\}.$$

定理5.3.1

对于线性模型(5.12)和(5.13), 若$\mathbf{X}_0\beta$在模型(5.12)下可估, 则在广义预测均方误差意义下, y_0的最优线性无偏预测(BLUP)为

$$\widetilde{y}_0^* = \mathbf{C}y = \mathbf{X}_0\beta^* + \mathbf{V}\Sigma^{-1}(y - \mathbf{X}\beta^*). \qquad (5.15)$$

证明 对于任一 $\tilde{y}_0 \in \mathcal{D}_0$, 有$E(\tilde{y}_0 - y_0) = \mathbf{0}$ 且

$$\mathrm{Cov}(\tilde{y}_0 - y_0) = \sigma^2(\mathbf{C}\Sigma\mathbf{C}' + \Sigma_0 - \mathbf{C}\mathbf{V}' - \mathbf{V}\mathbf{C}').$$

结合定理3.2.1, 立得

$$\mathrm{PMSE}(\tilde{y}_0) = \mathrm{tr}\left(\mathbf{A}\mathrm{Cov}(\tilde{y}_0 - y_0)\right) = \sigma^2\mathrm{tr}\left(\mathbf{A}(\mathbf{C}\Sigma\mathbf{C}' + \Sigma_0 - 2\mathbf{C}\mathbf{V}')\right).$$

因此, 在广义预测均方误差意义下求y_0的BLUP等价于在条件$\mathbf{C}\mathbf{X} = \mathbf{X}_0$下求PMSE$(\tilde{y}_0)$的最小值.

- 应用Lagrange乘子法求解这个极值问题. 构造辅助函数

$$F(\mathbf{C}, \boldsymbol{\Lambda}) = \sigma^2 \mathrm{tr}(\mathbf{A}\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}' - 2\mathbf{A}\mathbf{C}\mathbf{V}') - 2\mathrm{tr}((\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\Lambda}),$$

这里$\boldsymbol{\Lambda}$为$p \times m$的Lagrange乘子. 于是, 对$F(\mathbf{C}, \boldsymbol{\Lambda})$关于$\mathbf{C}, \boldsymbol{\Lambda}$求微商, 并令其为零, 得到

$$\boldsymbol{\Sigma}\mathbf{C}'\mathbf{A} = \mathbf{V}'\mathbf{A} + \mathbf{X}\boldsymbol{\Lambda}/\sigma^2, \tag{5.16}$$

$$\mathbf{C}\mathbf{X} = \mathbf{X}_0. \tag{5.17}$$

由(5.16)可得

$$\mathbf{C} = \mathbf{V}\mathbf{\Sigma}^{-1} + \mathbf{A}^{-1}\mathbf{\Lambda}'\mathbf{X}'\mathbf{\Sigma}^{-1}/\sigma^2. \tag{5.18}$$

将其代入(5.17)整理得 $\mathbf{\Lambda}'\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X} = \mathbf{A}(\mathbf{X}_0 - \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{X})\sigma^2$. 因为 $\mathcal{M}(\mathbf{X}_0') \subset \mathcal{M}(\mathbf{X}')$, 即$\mathbf{X}_0\beta$ 为可估函数, 此方程相容, 其解为

$$\mathbf{\Lambda}' = \sigma^2\mathbf{A}(\mathbf{X}_0 - \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{X})(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-}.$$

代入(5.18), 得到

$$\mathbf{C} = \mathbf{X}_0(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{\Sigma}^{-1} + \mathbf{V}\mathbf{\Sigma}^{-1}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{\Sigma}^{-1}).$$

故证得y_0 的BLUP为 $\tilde{y}_0^* = \mathbf{C}y = \mathbf{X}_0\beta^* + \mathbf{V}\mathbf{\Sigma}^{-1}(y - \mathbf{X}\beta^*)$.

推论5.3.1

对于线性模型(5.12)和(5.13)，若$\mathbf{X}_0\boldsymbol{\beta}$在模型(5.12)下可估，且$\mathrm{Cov}(\boldsymbol{e},\varepsilon_0)=\mathbf{0}$, 则在广义预测均方误差意义下, $\boldsymbol{y}_0$的BLUP为

$$\boldsymbol{y}_0^* = \mathbf{X}_0\boldsymbol{\beta}^*,$$

其中$\boldsymbol{\beta}^* = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^-\mathbf{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$. 进一步，假设$\mathrm{Cov}(\boldsymbol{e}) = \sigma^2\mathbf{I}_n$, 则$\boldsymbol{y}_0$的BLUP为

$$\widehat{\boldsymbol{y}}_0 = \mathbf{X}_0\widehat{\boldsymbol{\beta}},$$

其中$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^-\mathbf{X}'\boldsymbol{y}$.

$k=1$ 是一个重要的特殊情形. 欲预测$y_0=\boldsymbol{x}_0'\boldsymbol{\beta}+e$, 记

$$\text{Cov}\begin{pmatrix} \boldsymbol{y} \\ y_0 \end{pmatrix} = \sigma^2\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{\sigma}_{12} \\ \boldsymbol{\sigma}_{12}' & \sigma_{22} \end{pmatrix},$$

则y_0 的BLUP为 $\tilde{y}_0^*=\boldsymbol{x}_0'\boldsymbol{\beta}^*+\sigma_{12}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\mathbf{X}\boldsymbol{\beta}^*)$.

- 若y_0与$\boldsymbol{y}$不相关, 则$y_0^*=\boldsymbol{x}_0'\boldsymbol{\beta}^*$ 为y_0 的BLUP, 且

$$\text{Var}(y_0^*-y_0)=\sigma^2\left(\boldsymbol{x}_0'(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^-\boldsymbol{x}_0+\sigma_{22}\right),$$

- 进一步, 若Cov($\boldsymbol{y}$) $=\sigma^2\mathbf{I}_n$, Var$(y_0)=\sigma^2$, 则$\widehat{y}_0=\boldsymbol{x}_0'\widehat{\boldsymbol{\beta}}$ 为y_0 的BLUP,

$$\text{Var}(\widehat{y}_0-y_0)=\sigma^2(\boldsymbol{x}_0'(\mathbf{X}'\mathbf{X})^-\boldsymbol{x}_0+1).$$

点预测

在实际应用中, 往往$\mathbf{V}$ 和Σ 是未知的, 一种常用的作法是用它们的某种估计$\widehat{\mathbf{V}}$和$\widehat{\Sigma}$替代.

$$\widetilde{y}_0 = \mathbf{C}\boldsymbol{y} = \mathbf{X}_0\widetilde{\boldsymbol{\beta}} + \widehat{\mathbf{V}}\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y} - \mathbf{X}\widetilde{\boldsymbol{\beta}}),$$

这里

$$\widetilde{\boldsymbol{\beta}} = (\mathbf{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{X})^{-}\mathbf{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y}.$$

此时, $\widetilde{y}_0$ 尽管不再是BLUP, 甚至于它根本不是线性的, 但是为方便计, 人们把它称为经验BLUP.

区间预测

所谓区间预测, 就是找一个区间, 使得被预测量的可能取值落在这个区间内的概率达到预先给定的值.

区间预测具有重要的应用, 例如, 在工程技术中, 设计者想知道新产品的某项性能指标大概会落在多大的区间范围内等.

- 假设$(\boldsymbol{y}', \boldsymbol{y}_0')'$的联合分布为正态分布,

$$\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{y}_0 \end{pmatrix} \sim N_{n+k}\left(\begin{pmatrix} \mathbf{X}\beta \\ \mathbf{X}_0\beta \end{pmatrix}, \sigma^2 \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{V}' \\ \mathbf{V} & \boldsymbol{\Sigma}_0 \end{pmatrix}\right). \tag{5.19}$$

和前面一样, 假设$\mathcal{M}(\mathbf{X}_0') \subset \mathcal{M}(\mathbf{X}')$. 记

$$\beta^* = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}, \quad \sigma^{2*} = (\boldsymbol{y} - \mathbf{X}\beta^*)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \mathbf{X}\beta^*)/(n-r).$$

1. 被预测量$\boldsymbol{y}_0$与历史数据$\boldsymbol{y}$不相关，$\text{Cov}(\boldsymbol{y}_0, \boldsymbol{y}) = \mathbf{V} = \mathbf{0}$.

在误差正态条件下，预测误差

$$\boldsymbol{z} = \boldsymbol{y}_0^* - \boldsymbol{y}_0 = \mathbf{X}_0\boldsymbol{\beta}^* - \boldsymbol{y}_0 \sim N_k(\mathbf{0},\ \sigma^2\boldsymbol{\Omega}), \tag{5.20}$$

其中$\boldsymbol{\Omega} = (\omega_{ij}) = \boldsymbol{\Sigma}_0 + \mathbf{X}_0(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-}\mathbf{X}_0'$. 记$\boldsymbol{\Sigma}_0 = (\sigma_{ij}^{(0)})$, $\boldsymbol{z} = (z_1, \cdots, z_k)'$, 其中，$z_i = y_{0i}^* - y_{0i}, \quad i = 1, \cdots, k.$

由定理4.3.3和$\boldsymbol{y}_0$与$\boldsymbol{y}$的独立性, 可推得

$$\frac{z_i}{\sigma^*\sqrt{\omega_{ii}}} \sim t_{n-r}, \qquad \frac{\boldsymbol{z}'\boldsymbol{\Omega}^{-1}\boldsymbol{z}}{k\sigma^{2*}} \sim F_{k,n-r}.$$

记$\mathbf{X}_0 = (\boldsymbol{x}_{01}, \cdots, \boldsymbol{x}_{0k})'$, $\omega_{ii} = \sigma_{ii}^{(0)} + \boldsymbol{x}_{0i}'(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-}\boldsymbol{x}_{0i}$.

定理5.3.2

在正态假设(5.19)下, 若$\mathbf{X}_0\beta$在模型(5.12)下可估, Cov$(\boldsymbol{y}_0, \boldsymbol{y}) = \mathbf{0}$, 则

(1) 单个因变量y_{0i}的置信系数为$1-\alpha$的预测区间为

$$\boldsymbol{x}'_{0i}\boldsymbol{\beta}^* \mp t_{n-r}\left(\frac{\alpha}{2}\right)\sqrt{\omega_{ii}}\sigma^*, \quad i \in \{1, \cdots, k\};$$

(2) $y_{01}, \cdots, y_{0k}$的置信系数为$1-\alpha$的Bonferroni同时预测区间为

$$\boldsymbol{x}'_{0i}\boldsymbol{\beta}^* \mp t_{n-r}\left(\frac{\alpha}{2k}\right)\sqrt{\omega_{ii}}\sigma^*, \quad i = 1, \cdots, k;$$

(3) $y_{0i}, \cdots, y_{0m}$置信系数为$1-\alpha$的Scheffé同时预测区间为

$$\boldsymbol{x}'_{0i}\boldsymbol{\beta}^* \mp \sqrt{kF_{k,\, n-r}(\alpha)\omega_{ii}}\sigma^*, \quad i = 1, \cdots, k.$$

注1 当Cov$(\boldsymbol{y}_0, \boldsymbol{y}) = \mathbf{0}$时, y_{0i}置信系数为$1-\alpha$的预测区间总是比其均值$\boldsymbol{x}_{0i}'\beta$的相应区间估计的长, 因为

$$\omega_{ii} = \sigma_{ii}^{(0)} + \mathrm{Var}(\boldsymbol{x}_i'\boldsymbol{\beta}^*) > \mathrm{Var}(\boldsymbol{x}_i'\boldsymbol{\beta}^*).$$

注2 不难发现k个可估函数$\boldsymbol{x}_{01}'\boldsymbol{\beta}, \cdots, \boldsymbol{x}_{0k}'\boldsymbol{\beta}$ 的Scheffé 同时置信区间中$k_0 = \mathrm{rk}(\mathbf{X}_0)$, 而$y_{0i}, \cdots, y_{0k}$的Scheffé同时预测区间中的$k$是被预测点的个数, $k \geq k_0$.

注3 当$\boldsymbol{\Sigma} = \mathbf{I}_n$, $\boldsymbol{\Sigma}_0 = \mathbf{I}_k$时, 用LS解$\widehat{\boldsymbol{\beta}}$ 和$\widehat{\sigma}$分别替换定理5.3.2 中的GLS解$\boldsymbol{\beta}^*$和σ^*, 即可得相应结果, 此时$\omega_{ii} = 1 + \boldsymbol{x}_{0i}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{x}_{0i}$.

2. 被预测量y_0与历史数据y相关($\mathrm{Cov}(y_0, y) = \mathbf{V} \neq \mathbf{0}$)

引理5.3.1

在正态假设(5.19)下，若$\mathbf{X}_0\beta$在模型(5.3.1)下可估，且$\mathcal{M}(\mathbf{V}') \subset \mathcal{M}(\mathbf{X}')$, 则

(1) $\widetilde{y}_0^* - y_0$ 和σ^{2^*} 独立;

(2) $\frac{\widetilde{y}_{0i}^* - y_{0i}}{\omega_{ii}\sigma^*} \sim t_{n-r}, \quad i = 1, \cdots, k,$

(3) $\frac{(\widetilde{y}_0^* - y_0)'\Omega^{-1}(\widetilde{y}_0^* - y_0)/k}{\sigma^{2*}} \sim F_{r_0, n-r},$

其中$r_0 = \mathrm{rk}(\mathbf{X}_0)$, $\omega_{ii} = \mathrm{Var}(\widetilde{y}_{0i}^* - y_{0i}) = \mathbf{\Omega}_{ii}$, 见(5.20).

区间预测

定理5.3.3

若引理5.3.1的条件成立, 则

(1) 单个因变量y_{0i}的置信系数为$1-\alpha$的预测区间为

$$\boldsymbol{x}_{0i}'\boldsymbol{\beta}^* \mp t_{n-r}\left(\frac{\alpha}{2}\right)\sqrt{\omega_{ii}}\sigma^*, \quad i \in \{1, \cdots, k\};$$

(2) $y_{0i}, \cdots, y_{0k}$的置信系数为$1-\alpha$的Bonferroni型同时预测区间为

$$\boldsymbol{x}_{0i}'\boldsymbol{\beta}^* \mp t_{n-r}\left(\frac{\alpha}{2m}\right)\sqrt{\omega_{ii}}\sigma^*, \quad i = 1, \cdots, k;$$

(3) $y_{0i}, \cdots, y_{0k}$置信系数为$1-\alpha$的Scheffé型同时预测区间为

$$\boldsymbol{x}_{0i}'\boldsymbol{\beta}^* \mp \sigma^*\sqrt{\omega_{ii} m F_{m,\, n-r}(\alpha)}, \quad i = 1, \cdots, k.$$

区间预测

注1 当$\mathbf{V}=\mathbf{0}$时, 一定有$\mathcal{M}(\mathbf{V}')\subset\mathcal{M}(\mathbf{X}')$. 另外易证若$\mathbf{V}\neq\mathbf{0}$, 则在定理**??**中的该条件等价于$\mathbf{\Sigma}_0=\mathbf{X}_0\mathbf{M}\mathbf{X}_0'$, 其中$\mathbf{M}\geq\mathbf{0}$.

注2 若条件$\mathcal{M}(\mathbf{V}')\subset\mathcal{M}(\mathbf{X}')$不成立, 则$\widetilde{\boldsymbol{y}}_0^*$与$\sigma^{2*}$不独立, 故定理5.3.3给出的预测区间不再适用. 此时需要基于渐近性质去构造.

小结

- 熟练掌握线性假设检验方法
- 熟练掌握置信域和同时置信区间
- 熟练掌握点预测和区间预测
  - 被预测量y_0与历史数据y不相关情形
  - 被预测量y_0与历史数据y相关情形
- 初步了解后续几节中最优设计、测量误差、逆回归、缺失数据分析的内容