文章编号：0465-7942(2009)03-0086-05

具有边界条件线性假设的检验问题

孙法省， 刘民千
（南开大学 数学科学学院，天津 300071）

摘要：对于正态线性模型的一般线性假设，讨论了实际中经常遇到的具有边界条件的线性假设的检验问题，利用几何投影工具得到了检验统计量. 该方法比一般的代数方法要简单.

关键词：约束最小二乘解；边界条件；似然比；正交投影；矩阵广义逆
中图分类号：O212 **文献标识码**：A

0 引 言

线性模型是现代统计学中应用最为广泛的模型之一，也是其它统计模型研究及应用的基础，其理论发展已比较成熟，在实际中也得到了较广泛的应用. 通常我们考虑的都是正态线性模型

$$y = X\beta + e,\ e \sim N_n(0,\sigma^2 I), \tag{1}$$

其中 y 为 $n\times 1$ 的观测向量，X 为 $n\times p$ 的设计阵，$\mathrm{rk}(X) = r \leqslant p$，$\beta$ 为 $p\times 1$ 的未知参数向量，e 为 $n\times 1$ 的随机误差向量，$\sigma^2 > 0$，这里 $\mathrm{rk}(X)$ 表示矩阵 X 的秩. 对于这类模型的一般线性假设 $H\beta = 0$，其中 $\mathrm{rk}(H_{m\times p}) = m \leqslant r$，$\mathscr{M}(H') \subset \mathscr{M}(X')$，即 $H\beta$ 为 m 个线性无关的可估函数，这里 $\mathscr{M}(A)$ 表示由矩阵 A 的列向量张成的子空间，我们有检验统计量

$$F = \frac{(SS_{He} - SS_e)/m}{SS_e/(n-r)},$$

其中

$$SS_e = \| y - X\hat\beta \|^2,\ \hat\beta = (X'X)^- X'y,$$
$$SS_{He} = \| y - X\hat\beta_H \|^2,\ \hat\beta_H = \hat\beta - (X'X)^- H'(H(X'X)^- H')^{-1}H\hat\beta.$$

如果线性假设 $H\beta = 0$ 成立，则 $F \sim F_{m,n-r}$，这里 $F_{m,n-r}$ 表示自由度为 $m, n-r$ 的 F 分布. 这个结论在许多文献中都可以找到，如文献[1]. 类似的问题也有不少人讨论过，见文献[2 − 3].

但是，我们经常会遇到这样的问题，即求模型

$$\begin{cases} y = X\beta + e,\ e \sim N_n(0,\sigma^2 I), \\ L\beta = 0, \end{cases} \tag{2}$$

其中 $\mathrm{rk}\begin{pmatrix} X \\ L \end{pmatrix} = \mathrm{rk}(X) + \mathrm{rk}(L)$，在线性假设 $H\beta = 0$ 下的检验统计量，即在约束条件 $L\beta = 0$ 下，再加约束条件 $H\beta = 0$ 的检验统计量，这里仍要求 $\mathscr{M}(H') \subset \mathscr{M}(X')$. 我们把约束条件

$$\mathrm{rk}\begin{pmatrix} X \\ L \end{pmatrix} = \mathrm{rk}(X) + \mathrm{rk}(L) = n, \quad L\beta = 0 \tag{3}$$

称为边界条件. 但是，在下面的证明中我们只要求

$$\mathrm{rk}\begin{pmatrix} X \\ L \end{pmatrix} = \mathrm{rk}(X) + \mathrm{rk}(L), \quad L\beta = 0.$$

很明显，模型(1) 是模型(2) 在条件 $L = 0$ 下的特殊情况. 对于该问题，可以将其转化为无约束的线性模型

收稿日期：2007-01-10
基金项目：国家自然科学基金(10671099)；高等学校博士学科点专项科研基金(20050055038)
作者简介：孙法省(1981−)，男，山东枣庄人，博士研究生.

(见文献[4]),也可以利用矩阵的广义逆求出约束最小二乘解,然后构造检验统计量(见文献[1]),还可以利用简约模型求检验统计量(见文献[5]). 由于约束最小二乘解的求解比较麻烦,所以检验统计量的推导比较复杂. 我们尝试利用几何投影的方法来推导它的检验统计量,相对于上述代数方法要简单的多.

1 主要结果

记参数 $\theta = (\beta', \sigma^2)'$,则 θ 的似然函数为

$$L(\theta; y) = L(\beta', \sigma^2; y) = (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp\{-\frac{1}{2\sigma^2}\|y - X\beta\|^2\}.$$

设在模型(2) 下参数 θ 的约束极大似然估计为 $\hat{\theta}_L = (\hat{\beta}_L, \hat{\sigma}_L^2)'$,其中 $\hat{\sigma}_L^2 = \|y - X\hat{\beta}_L\|^2/n$,在模型

$$\begin{cases} y = X\beta + e, \quad e \sim N_n(0, \sigma^2 I), \\ L\beta = 0, \\ H\beta = 0, \end{cases} \tag{4}$$

下参数 θ 的约束极大似然估计 $\hat{\theta}_M = (\hat{\beta}_M, \hat{\sigma}_M^2)'$,其中 $M = \begin{pmatrix} L \\ H \end{pmatrix}$. 下面来导出似然比

$$\lambda(y) = \frac{L(\hat{\beta}_M, \hat{\sigma}_M^2; y)}{L(\hat{\beta}_L, \hat{\sigma}_L^2; y)} = \left(\frac{\|y - X\hat{\beta}_M\|^2}{\|y - X\hat{\beta}_L\|^2}\right)^{\frac{n}{2}} \triangleq \left(\frac{SS_M}{SS_L}\right)^{\frac{n}{2}} \tag{5}$$

若按照一般的代数方法推导模型(2) 在线性假设 $H\beta = 0$ 下的检验统计量将很麻烦,因为约束最小二乘解的求取及其表达式都很复杂,当然是可以求出来的,为了对比,在引理 1 中对它们的求解作了部分说明. 为此我们转换到从几何投影角度来研究该问题. 下面先看一下引理.

引理 1 $\begin{cases} \min \|y - X\beta\|^2 \\ L\beta = 0 \end{cases}$ 的解为 $\hat{\beta} = (T^- - T^- L'Q^- LT^-)X'y, T = X'X + L'L, Q = LT^- L'$.

证明 利用 Lagrange 乘子法,构造辅助函数

$$Q(\beta, \lambda) = \|y - X\beta\|^2 + 2\lambda' L\beta,$$

其中 λ 为 Lagrange 乘子. 将该函数关于 β 和 λ 求微商,并令所得微商等于零,得到

$$\begin{cases} X'X\beta + L'\lambda = X'y, \\ L\beta = 0, \end{cases}$$

即

$$\begin{pmatrix} X'X & L' \\ L & 0 \end{pmatrix}\begin{pmatrix} \beta \\ \lambda \end{pmatrix} = \begin{pmatrix} X'y \\ 0 \end{pmatrix}.$$

根据镶边矩阵广义逆的结果(见文献[1] 中定理 2.6),我们有

$$\begin{pmatrix} X'X & L' \\ L & 0 \end{pmatrix}^- \triangleq \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} = \begin{pmatrix} T^- - T^- L'Q^- LT^- & T^- L'Q^- \\ Q^- LT^- & Q^- Q - Q^- \end{pmatrix},$$

这里 $T = X'X + L'L$, $Q = LT^- L'$. 于是我们得到 β 的解为

$$\hat{\beta} = (T^- - T^- L'Q^- LT^-)X'y.$$

注 1) 此处 L 是任意的,不需满足 $\mathrm{rk}\begin{pmatrix} X \\ L \end{pmatrix} = \mathrm{rk}(X) + \mathrm{rk}(L)$;

2) 这里还需证明方程组有解及验证 $\hat{\beta}$ 在条件 $L\beta = 0$ 下的最小性,这些可以参阅文献[1];

3) 由上面可以看出 $\hat{\beta}$ 的表达式很复杂,如果再加上 $H\beta = 0$ 这一线性假设,则要更复杂的多. 因此先求约束最小二乘解,再来推导检验统计量的方法是非常麻烦的.

引理 2 $\{X\beta : L\beta = 0\} = \mathscr{M}(XL'^{\perp})$.

证明 要证两集合相等只需证两集合相互包含. 任取 $X\beta \in \{X\beta : L\beta = 0\}$,由于 $L\beta = 0$,所以 $\beta \in \mathscr{M}(L'^{\perp})$,因而存在 β_1 满足 $\beta = L'^{\perp}\beta_1$,即 $X\beta = XL'^{\perp}\beta_1 \in \mathscr{M}(XL'^{\perp})$,由 $X\beta$ 的任意性知 $\{X\beta : L\beta = 0\} \subseteq$

$\mathscr{M}(XL'^{\perp})$.

另一方面,任取 $\alpha \in \mathscr{M}(XL'^{\perp})$,则存在 α_1 使 $\alpha = XL'^{\perp}\alpha_1$. 令 $L'^{\perp}\alpha_1 = \beta_1$,则 $\alpha = X\beta_1$,且 $L\beta_1 = LL'^{\perp}\alpha_1 = 0$,所以 $\alpha \in \{X\beta: L\beta = 0\}$. 由 α 的任意性知 $\{X\beta: L\beta = 0\} \supseteq \mathscr{M}(XL'^{\perp})$.

综上可知 $\{X\beta: L\beta = 0\} = \mathscr{M}(XL'^{\perp})$. 证毕.

从几何意义上看 $X\hat{\beta}_L$ 为 y 向子空间 $\{X\beta: L\beta = 0\}$ 上的正交投影,由引理知,即是在 $\mathscr{M}(XL'^{\perp})$ 上的正交投影. 从而我们可以把 $X\hat{\beta}_L$ 表示成 $X\hat{\beta}_L = P_{XL'^{\perp}}y$,这里

$$P_A = A(A'A)^- A' \tag{6}$$

是向 $\mathscr{M}(A)$ 上的正交投影阵. 于是

$$SS_L = \| y - X\hat{\beta}_L \|^2 = y'(I - P_{XL'^{\perp}})y. \tag{7}$$

同理

$$X\hat{\beta}_M = P_{XM'^{\perp}}y, \tag{8}$$

$$SS_M = \| y - X\hat{\beta}_M \|^2 = y'(I - P_{XM'^{\perp}})y, \tag{9}$$

其中 $M = \begin{pmatrix} L \\ H \end{pmatrix}$. 由式(5), $\lambda(y) = \left(\dfrac{SS_M}{SS_L}\right)^{\frac{n}{2}}$,令 $F = \dfrac{n-r}{m}(\lambda(y)^{\frac{2}{n}} - 1)$,则

$$F = \frac{(SS_M - SS_L)/m}{SS_L/(n-r)}. \tag{10}$$

显然,F 仅依赖于 $\lambda(y)$,且为 $\lambda(y)$ 的严格增函数.

定理1 对模型(2),设 $H\beta$ 为 m 个线性无关的可估函数,即 $\mathscr{M}(H') \subset \mathscr{M}(X')$, $\mathrm{rk}(H) = m$. 对于由上式(7),(9)和(10)分别给出的 SS_L, SS_M 和 F,我们有

1) $SS_L \sim \sigma^2\chi^2_{n-r}$,其中 χ^2_{n-r} 表示自由度为 $n-r$ 的中心 χ^2 分布;

且在线性假设 $H\beta = 0$ 下,下面的结论成立

2) $SS_M \sim \sigma^2\chi^2_{n-(r-m)}$;

3) $SS_M - SS_L \sim \sigma^2\chi^2_m$ 且与 SS_L 相互独立;

4) $F \sim F_{m,n-r}$,其中 $F_{m,n-r}$ 表示自由度为 $m, n-r$ 的 F 分布.

证明 1) 由正交投影阵的定义(6),可知 $I - P_{XL'^{\perp}}$ 是对称幂等阵,从而

$$\mathrm{rk}(I - P_{XL'^{\perp}}) = n - \mathrm{rk}(P_{XL'^{\perp}}),$$

$$\begin{aligned}\mathrm{rk}(P_{XL'^{\perp}}) &= \mathrm{rk}\begin{pmatrix} X \\ L \end{pmatrix} - \mathrm{rk}(L) \\ &= \mathrm{rk}(X) + \mathrm{rk}(L) - \mathrm{rk}(L) = \mathrm{rk}(X) = r.\end{aligned}$$

由引理1知 $(I - P_{XL'^{\perp}}) \cdot X\beta = 0$,于是由文献[5]中关于 χ^2 分布的结论知 $SS_L \sim \sigma^2\chi^2_{n-r}$.

2) 同理可证.

3) 由式(7)和式(9),得 $SS_M - SS_L = y'(P_{XL'^{\perp}} - P_{X(L'H')^{\perp}})y$. 显然 $\mathscr{M}(X(L'H')^{\perp}) \subset \mathscr{U}(XL'^{\perp})$,从而 $P_{XL'^{\perp}} \cdot P_{X(L'H')^{\perp}} = P_{X(L'H')^{\perp}} = P_{X(L'H')^{\perp}}P_{XL'^{\perp}}$,进而

$$\begin{aligned}(P_{XL'^{\perp}} - P_{X(L'H')^{\perp}})^2 &= P^2_{XL'^{\perp}} - P_{XL'^{\perp}}P_{X(L'H')^{\perp}} - P_{X(L'H')^{\perp}} \cdot P_{XL'^{\perp}} + P^2_{X(L'H')^{\perp}} \\ &= P_{XL'^{\perp}} - P_{X(L'H')^{\perp}} - P_{X(L'H')^{\perp}} + P_{X(L'H')^{\perp}} \\ &= P_{XL'^{\perp}} - P_{X(L'H')^{\perp}},\end{aligned}$$

即 $P_{XL'^{\perp}} - P_{X(L'H')^{\perp}}$ 是对称幂等阵. 又

$$\mathrm{rk}(P_{XL'^{\perp}} - P_{X(L'H')^{\perp}}) = \mathrm{rk}(P_{XL'^{\perp}}) - \mathrm{rk}(P_{X(L'H')^{\perp}}),$$

$$\mathrm{rk}(P_{XL'^{\perp}}) = \mathrm{rk}\begin{pmatrix} X \\ L \end{pmatrix} - \mathrm{rk}(L) = \mathrm{rk}(X) = r,$$

$$\begin{aligned}\mathrm{rk}(P_{X(L'H')^{\perp}}) &= \mathrm{rk}(X' \vdots L' \vdots H') - r\begin{pmatrix} L \\ H \end{pmatrix} \\ &= \mathrm{rk}(X) + \mathrm{rk}(L) - \mathrm{rk}(L) - \mathrm{rk}(H) = r - m,\end{aligned}$$

因而 $\mathrm{rk}(P_{XL'^\perp}-P_{X(L'H')^\perp})X\beta=0$，于是有 $SS_M-SS_L\sim\sigma^2\chi^2_m$. 又因为

$$(P_{XL'^\perp}-P_{X(L'H')^\perp})(I-P_{XL'^\perp})=0,$$

所以利用文献[5]中关于正态分布的二次型的定理3.5.2可得 SS_M-SS_L 与 SS_L 相互独立.

4) 由结论1)－3)及 F 分布的定义可直接得出. 定理证毕.

定理1给出了模型(2)下线性假设 $H\beta=0$ 的似然比检验统计量及其分布，即

$$F=\frac{(SS_M-SS_L)/m}{SS_L/(n-r)}\sim F_{m,n-r}.$$

依似然比检验原理知，对于给定的显著性水平 $\alpha(0<\alpha<1)$，若 $F>F_{m,n-r}(\alpha)$，则拒绝假设 $H\beta=0$；若 $F\leqslant F_{m,n-r}(\alpha)$，则接受假设 $H\beta=0$，这里 $F_{m,n-r}(\alpha)$ 表示自由度为 $m,n-r$ 的 F 分布的上侧 α 分位点.

注　定理1的证明中仅要求 $\mathrm{rk}\begin{pmatrix}X\\L\end{pmatrix}=\mathrm{rk}(X)+\mathrm{rk}(L)$. 在一般情况下，则常取 L 为边界条件(3).

2　例　　子

下面把上一节得到的结论应用到单向分类模型

$$\begin{cases}y_{ij}=\alpha_i+e_{ij}, & e_{ij}\sim N(0,\sigma^2 I_{ab})\\ \sum\limits_{i=1}^{a}\alpha_i=0, & i=1,\cdots,a,j=1,\cdots,b,\end{cases}\tag{11}$$

的假设检验中，推导出在线性假设

$$H_0:\alpha_1=\alpha_2=\cdots=\alpha_a\tag{12}$$

下的检验统计量.

对该模型，把它与模型(2)相对应，我们有

$$y=(y_{11},\cdots,y_{1b},\cdots,y_{a1},\cdots,y_{ab})',\ e=(e_{11},\cdots,e_{1b},\cdots,e_{a1},\cdots,e_{ab})',$$

$$X=\begin{pmatrix}1_b & 1_b & & \\ 1_b & & 1_b & \\ \vdots & & & \ddots \\ 1_b & & & & 1_b\end{pmatrix}_{ab\times a},\ \beta=(\alpha_1,\alpha_2,\cdots,\alpha_a)',\ L=(1,1,\cdots,1)_{1\times a}.$$

而线性假设(12)对应有

$$H=\begin{pmatrix}1 & -1 & & \\ 1 & & -1 & \\ \vdots & & & \ddots \\ 1 & & & & -1\end{pmatrix}_{a\times 1}$$

这里矩阵 X 和 H 中空白处元素全为零，$\mathrm{rk}(X)=a$，$\mathrm{rk}(H)=a-1$. 利用前面的结论，即得检验统计量为

$$F=\frac{(SS_M-SS_L)/(a-1)}{SS_L/(ab-a)}\sim F_{a-1,ab-a},$$

其中

$$SS_M-SS_L=\sum_{i=1}^{a}b(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2,\ SS_L=\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{i\cdot})^2,$$

$$\bar{y}_{\cdot\cdot}=\frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}y_{ij},\bar{y}_{i\cdot}=\sum_{j=1}^{b}y_{ij},\ i=1,\cdots,a,$$

也即

$$F = \frac{\sum_{i=1}^{a} b(\bar{y}_{i.} - \bar{y}_{..})^2/(a-1)}{\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.})^2/(ab-a)}. \tag{13}$$

这样，对于给定的显著性水平 $\alpha(0<\alpha<1)$，若 $F > F_{a-1,ab-a}(\alpha)$，则拒绝线性假设(12)，即拒绝 $H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_a$，若 $F \leqslant F_{a-1,ab-a}(\alpha)$，则接受该线性假设，这里 $F_{a-1,ab-a}(\alpha)$ 表示自由度为 $a-1, ab-a$ 的 F 分布的上侧 α 分位点.

注意到本例中的约束条件 L 为边界条件，满足式(3). 由检验统计量(13) 的上述推导过程可以看到，对于具有边界约束条件的假设检验问题，本文所给出的几何投影的方法克服了传统的代数方法在辅导上较为繁琐的缺点，是比较容易实施的.

参 考 文 献

1 王松桂. 线性模型的理论及其应用[M]. 合肥：安徽教育出版社，1987.

2 Rao C R, Toutenburg H. Linear Models: Least Squares and Alternatives[M]. 2nd ed. New York: Springer-Verlag, 1999.

3 Sen A, Srivastava M. Regression Analysis: Theory, Methods and Applications[M]. New York: Springer-Verlag, 1997.

4 杨文礼. 线性模型引论[M]. 北京：北京师范大学出版社，1998.

5 王松桂，史建红，尹素菊，等. 线性模型引论[M]. 北京：科学出版社，2004.

The Problem of Testing Linear Hypothesis with Boundary Conditions

Sun Fasheng, Liu Minqian

(School of Mathematical Sciences, Nankai University, Tianjin 300071, China)

Abstract: Regarding the general linear hypothesis of normal linear model, the problem of testing linear hypothesis with boundary conditions is discussed, which is often encountered in practice. The tool of geometric projection is applied to obtain the testing statistic. This method is much simpler than the general algebra methods.

Key words: restricted least squares solution; boundary conditions; likelihood ratio; orthogonal projection; generalize inverse of matrices