应用概率统计　第二十三卷
第一期　2007 年 2 月

Chinese Journal of Applied Probability
and Statistics Vol. 23 No. 1 Feb. 2007

错误先验假定下 Bayes 线性无偏估计的稳健性 *

张伟平　　韦来生
(中国科技大学统计与金融系，合肥，230026)

摘　要

本文基于错误的先验假定获得了一般线性模型下可估函数的 Bayes 线性无偏估计 (BLUE), 证明了在均方误差矩阵 (MSEM) 准则和后验 Pitman Closeness (PPC) 准则下 BLUE 相对于最小二乘估计 (LSE) 的优良性，并导出了它们的相对效率的界，从而获得 BLUE 的稳健性.

关 键 词: 错误先验假定，Bayes 线性无偏估计，最小二乘估计，相对效率，稳健性.
学 科 分 类 号: O212.1.

§1.　引　　言

考虑如下线性模型:

$$Y_{n\times 1} = X_{n\times p}\beta_{p\times 1} + e_{n\times 1}, \tag{1.1}$$

其中 Y 为观测向量，X 为非零的设计阵并且 $\mathrm{rank}(X)=r\le p$, β 为未知参数向量. e 为随机误差向量且假定 $\mathsf{E}e=0$, $\mathsf{Cov}\,(e)=\sigma^2 I_n$, σ^2 已知.

设满足下列条件

$$\mathsf{E}\beta=\mu, \qquad \mathsf{Cov}\,(\beta)=\tau^2 V \tag{1.2}$$

的先验分布 $\pi_0(\beta)$ 是错误指定的，正确的先验分布 $\pi(\beta)$ 满足下列条件

$$\mathsf{E}\beta=\theta, \qquad \mathsf{Cov}\,(\beta)=\tau^2 W, \tag{1.3}$$

其中 μ, τ^2, θ 已知，V, W 为已知的非负定矩阵，记为 $V\ge 0$, $W\ge 0$.

关于线性模型 (1.1) 的 Bayes 估计问题，韦来生 [1] 在假定设计阵列满秩和先验协方差矩阵正定的条件下考虑了 Bayes 线性无偏估计的小样本性质. 在实际应用中，设计阵 X 为病态及先验协方差矩阵为非负定的情形经常发生，这种情形给研究参数的 Bayes 估计方法及其性质带来一些困难. 本文就是针对这些情况讨论在设计阵 X 非列满秩且先验协方差矩阵 V, W 非负定的情形下求 Bayes 线性无偏估计的方法并探讨其关于错误先验假定的稳健性问题.

对模型 (1.1) 易知可估函数 $\eta = X\beta$ 的最小二乘估计 (LSE) 为

$$\widehat{\eta}_{LS} = X(X'X)^- X'Y, \tag{1.4}$$

其中 $(X'X)^-$ 表示矩阵 $X'X$ 的广义逆.

* 教育部博士点基金、中科院创新基金和中国科技大学青年基金资助项目.
本文 2004 年 8 月 20 日收到，2005 年 11 月 15 日收到修改稿.

设损失函数为二次损失，即

$$L(d,\theta)=(d-\theta)'D(d-\theta), \tag{1.5}$$

其中 θ 为参数， d 为行动， D 为正定方阵．相应的风险函数为： $R(d,\theta)=\mathsf{E}[L(d,\theta)]$, 此处及以下 E 表示关于 Y 和 β 的联合分布求均值．

考虑 η 的 Bayes 线性估计类：

$$\{\widehat{\eta}=AY+b:A \text{ 为 } n\times n \text{ 实矩阵}, b \text{ 为 } n\times 1 \text{ 维实向量}\}.$$

设 η 的 Bayes 线性无偏估计为 $\widehat{\eta}_B$, 即满足

$$R(\widehat{\eta}_B,X\beta)=\min_{A,b}R(\widehat{\eta},X\beta)=\min_{A,b}\mathsf{E}(\widehat{\eta}-X\beta)'D(\widehat{\eta}-X\beta) \tag{1.6}$$

和约束条件 $\mathsf{E}[\widehat{\eta}-X\beta]=0$. 由此解出 $b=(I_n-A)X\mu$, 为获得使 (1.6) 成立的矩阵 A 的最佳选择，计算 $\widehat{\eta}$ 的 Bayes 风险如下

$$\begin{aligned}R(\widehat{\eta},X\beta)&=\mathsf{E}(AY+b-X\beta)'D(AY+b-X\beta)\\&=\operatorname{tr}\{D\mathsf{E}[A(Y-X\beta)-(I_n-A)(X\beta-X\mu)][A(Y-X\beta)-(I_n-A)(X\beta-X\mu)]'\}\\&=\operatorname{tr}\{D[\sigma^2AA'+\tau^2(I_n-A)XVX'(I_n-A)']\}.\end{aligned}$$

令 $\partial R(\widehat{\eta},X\beta)/\partial A=0$, 利用矩阵微商法则得

$$A(\sigma^2I_n+\tau^2XVX')=\tau^2XVX'.$$

解此方程得到

$$A=XVX'(\rho I_n+XVX')^{-1}, \tag{1.7}$$

其中 $\rho=\sigma^2/\tau^2$. 从而在错误先验假定 (1.2) 之下 η 的 Bayes 线性无偏估计为

$$\widehat{\eta}_B=XVX'(\rho I_n+XVX')^{-1}Y+\rho(\rho I_n+XVX')^{-1}X\mu. \tag{1.8}$$

求线性模型回归系数及其可估函数的 Bayes 估计有两种方法．一种是在正态线性模型下假定参数的先验分布为正态分布，易知其后验分布也为正态分布；故在二次损失下 Bayes 估计由后验均值给出 (见王松桂 [2] §6.4), 这是常用的求 Bayes 估计的方法．另一种方法是在 Gauss-Markov 模型下假定先验分布满足一定矩的条件，在二次损失下通过最小化 Bayes 风险获得 Bayes 线性无偏估计 (BLUE), 上面由 (1.8) 给出的估计就属于这一类．这一方法最早是由 Rao[3] 提出，Gruber[4] 作了具体的讨论．本文采用后一种方法获得有关参数的 BLUE, 并研究它的小样本性质．

在 Bayes 方法中，当先验信息积累不是足够多时先验分布被错误地指定是不可避免的．因此基于错误指定的先验假定而获得的 Bayes 估计的性质如何是人们所关心的．本文将讨论在均方误差矩阵 (MSEM) 准则和后验 Pitman Closeness (PPC) 准则下基于错误先验假定而获得的 BLUE 相对于 LSE 的优良性等小样本性质．

本文第二节将在 MSEM 准则下讨论错误先验假定下获得的 BLUE 相对于最小二乘估计 LSE 的优良性，第三节将考虑 BLUE 与 LSE 的相对效率，最后我们还将考虑 BLUE 相对于 LSE 的后验 PPC 特性．

§2. 均方误差矩阵准则下 Bayes 估计与 LS 估计的比较

本节我们在模型 (1.1) 和正确的先验假定 (1.3) 下计算 $\hat{\eta}_{LS}$ 和 $\hat{\eta}_B$ 的 MSEM, 然后讨论在 MSEM 准则下基于错误先验假定获得的 BLUE (1.8) 相对于 LSE (1.4) 的优良性.

定义 2.1 设参数向量 θ 的估计量为 $\hat{\theta}$, 则 $M(\hat{\theta})=\mathsf{E}(\hat{\theta}-\theta)(\hat{\theta}-\theta)'$ 称为 $\hat{\theta}$ 的均方误差矩阵 (MSEM); 而 $\mathrm{MSE}(\hat{\theta})=\mathsf{E}(\hat{\theta}-\theta)'(\hat{\theta}-\theta)$ 称为 $\hat{\theta}$ 的均方误差 (MSE). 设 $\hat{\theta}_1$, $\hat{\theta}_2$ 分别为参数向量 θ 的两个不同估计量， $\hat{\theta}_1$ 称为在 MSEM (或 MSE) 准则下优于 $\hat{\theta}_2$, 如果 $M(\hat{\theta}_2)-M(\hat{\theta}_1)\geq 0$ (或 $\mathrm{MSE}(\hat{\theta}_2)-\mathrm{MSE}(\hat{\theta}_1)\geq 0$).

记 $M_X=I-X(X'X)^-X'$, $A_V=XVX'(\rho I_n+XVX')^{-1}=I_n-\rho B_V^{-1}$, 其中 $B_V=\rho I_n+XVX'$. 在正确先验分布 (1.3) 下， η 的 BLUE (1.8) 的 MSEM 为

$$\begin{aligned}
M(\hat{\eta}_B) &= \mathsf{E}(\hat{\eta}_B-X\beta)(\hat{\eta}_B-X\beta)' \\
&= \mathsf{E}[A_V(Y-X\beta)-\rho B_V^{-1}(X\beta-X\mu)][A_V(Y-X\beta)-\rho B_V^{-1}(X\beta-X\mu)]' \\
&= \sigma^2(I_n-\rho B_V^{-1})(I_n-\rho B_V^{-1})'+\rho^2B_V^{-1}[\tau^2XWX'+X(\theta-\mu)(\theta-\mu)'X']B_V^{-1} \\
&= \sigma^2I_n-2\sigma^2\rho B_V^{-1}+\rho^2B_V^{-1}[\sigma^2I_n+\tau^2XWX'+X(\theta-\mu)(\theta-\mu)'X']B_V^{-1}. \qquad (2.1)
\end{aligned}$$

其中利用了 $I_n-\rho B_V^{-1}=XVX'B_V^{-1}$. 而 η 的 LSE (1.4) 的 MSEM 为

$$\begin{aligned}
M(\hat{\eta}_{LS}) &= \mathsf{E}(X(X'X)^-X'Y-X\beta)(X(X'X)^-X'Y-X\beta)' \\
&= X(X'X)^-X'\mathsf{E}(Y-X\beta)(Y-X\beta)'X(X'X)^-X' \\
&= \sigma^2X(X'X)^-X'. \qquad (2.2)
\end{aligned}$$

其中利用了 $X(X'X)^-X'X=X$. 注意到 $B_VM_X=M_XB_V=\rho M_X$, 所以

$$\begin{aligned}
& M(\hat{\eta}_{LS})-M(\hat{\eta}_B) \\
=\ & 2\sigma^2\rho B_V^{-1}-\rho^2B_V^{-1}[\sigma^2I_n+\tau^2XWX'+X(\theta-\mu)(\theta-\mu)'X']B_V^{-1}-\sigma^2M_X \\
=\ & B_V^{-1}\{2\sigma^2\rho B_V-\rho^2[\sigma^2I_n+\tau^2XWX'+X(\theta-\mu)(\theta-\mu)'X'+\sigma^2M_X]\}B_V^{-1} \\
=\ & \rho^2B_V^{-1}X[\sigma^2(X'X)^-+\tau^2(2V-W)-(\theta-\mu)(\theta-\mu)']X'B_V^{-1}. \qquad (2.3)
\end{aligned}$$

根据上式我们可以得到 $M(\hat{\eta}_{LS})-M(\hat{\eta}_B)\geq 0$ 的一个充分条件为

$$2V-W\geq 0,\qquad \tau^2(2V-W)-(\theta-\mu)(\theta-\mu)'\geq 0. \qquad (2.4)$$

当 $2V-W>0$ 时，利用代数事实：若 $A>0$, $d>0$ 为实数， x 为实的向量，则 $dA-xx'\geq 0$ $\Longleftrightarrow x'A^{-1}x\leq d$. 因此条件 (2.4) 等价于

$$2V-W>0,\qquad (\theta-\mu)'(2V-W)^{-1}(\theta-\mu)\leq\tau^2. \qquad (2.5)$$

综上，我们有如下结论

定理 2.1 设模型 (1.1) 和错误先验假定 (1.2) 下 η 的 LSE 和 BLUE 分别由 (1.4) 和 (1.8) 给出，则在 β 的正确先验分布 (1.3) 和条件 (2.4) 下，有

$$M(\hat{\eta}_{LS})-M(\hat{\eta}_B)\geq 0.$$

特别若条件 (2.5) 成立，则有 $M(\widehat{\eta}_{LS})-M(\widehat{\eta}_B)>0$.

利用 $\mathrm{MSE}(\widehat{\theta})=\mathrm{tr}\,M(\widehat{\theta})$, 立得下列推论

推论 2.1 设定理 2.1 的条件成立，则有

$$\mathrm{MSE}(\widehat{\eta}_{LS})-\mathrm{MSE}(\widehat{\eta}_B)>0.$$

注 2.1 若错误先验假定 (1.2) 是正确的，即 $\mu=\theta$, $V=W$, 从而 $\widehat{\eta}_B$ 是基于正确的先验假定而得到的 Bayes 线性无偏估计．此时条件 (2.4) 自然成立．因此 $M(\widehat{\eta}_{LS})-M(\widehat{\eta}_B)\geq 0$ 恒成立，若进一步还有 $V>0$, $\mathrm{rank}(X)=p$, 则 $M(\widehat{\eta}_{LS})-M(\widehat{\eta}_B)>0$. 这表明在 MSEM 准则下，由正确先验假定而得到的 BLUE 是必然优于 LSE 的．

定理 2.1 的结果可以推广到一般可估函数的情形．设 $G\beta$ (其中 G 为已知的 $m\times p$ 矩阵) 为可估函数，则存在矩阵 $m\times n$ 维矩阵 L, 使得 $G=LX$. 从而有

定理 2.2 设模型 (1.1) 和错误先验假定 (1.2) 下 η 的 LSE 和 BLUE 分别由 (1.4) 和 (1.8) 给出，此时 $G\beta$ 的 LSE 和 BLUE 分别为 $L\widehat{\eta}_{LS}$ 和 $L\widehat{\eta}_B$. 则在 β 的正确先验分布 (1.3) 和条件 (2.4) 下，有

$$M(L\widehat{\eta}_{LS})-M(L\widehat{\eta}_B)\geq 0.$$

证明: 由 $M(L\widehat{\eta}_{LS})-M(L\widehat{\eta}_B)=L[M(\widehat{\eta}_{LS})-M(\widehat{\eta}_B)]L'$ 和定理 2.1, 可知定理 2.2 成立． #

§3. LSE 与 BLUE 的相对效率

为获得 BLUE 相对于 LSE 在 MSEM 准则下优良性的定量指标，需讨论估计量的相对效率问题．参数的两个无偏估计量的相对效率有多种不同的定义方式．刘爱义、王松桂 [5] 是通过两个无偏估计量的协方差阵的 trace (迹) 之比来定义的．我们将对此定义作一定的修改，用估计量的均方误差矩阵代替协方差矩阵．其定义如下:

定义 3.1 设 $\widehat{\eta}_{LS}$ 和 $\widehat{\eta}_B$ 的 MSEM 分别由 (2.2) 和 (2.1) 给出．定义 $\widehat{\eta}_{LS}$ 和 $\widehat{\eta}_B$ 的相对效率如下:

$$e(\widehat{\eta}_{LS},\widehat{\eta}_B)=\frac{\mathrm{tr}\{M(\widehat{\eta}_B)\}}{\mathrm{tr}\{M(\widehat{\eta}_{LS})\}}.$$

关于 $e(\widehat{\eta}_{LS},\widehat{\eta}_B)$ 的界，我们有如下结论

定理 3.1 设 η 的 LSE 和 BLUE 分别由 (1.4) 和 (1.8) 给出，则在正确先验假定 (1.3) 条件 (2.4) 下，有

$$\frac{1}{r}\sum_{i=1}^{t}\frac{\lambda_i^2}{(\rho+\lambda_i)^2}\leq e(\widehat{\eta}_{LS},\widehat{\eta}_B)\leq\frac{1}{r}\sum_{i=1}^{t}\frac{\lambda_i^2+2\rho\lambda_i}{(\rho+\lambda_i)^2}.$$

其中 $r=\mathrm{rank}(X)$, $t=\mathrm{rank}(XVX')$, $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_t>0$ 为 XVX' 的所有非零特征根．

证明: 由于 $XVX'\geq 0$, 故存在正交矩阵 P_1 使得

$$XVX'=P_1'\Lambda P_1=\Lambda=\mathrm{diag}\{\lambda_1,\cdots,\lambda_t,\lambda_{t+1},\cdots,\lambda_n\},\tag{3.1}$$

此处 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_t>0$ 为 XVX' 的所有非零特征根， $\lambda_i=0$, $i=t+1,\cdots,n$. 进而

$$(\rho I_n+XVX')^{-1}=B_V^{-1}=P_1'\Delta P_1,\qquad 其中\ \ \Delta=(\rho I_n+\Lambda)^{-1}.\tag{3.2}$$

由 $\operatorname{rank}(X'X)=r$ 易知

$$\operatorname{tr}\{X(X'X)^{-}X'\}=\operatorname{tr}\{X'X(X'X)^{+}\}=r. \tag{3.3}$$

由 (2.1) 可知，有 $M(\widehat{\eta}_B)=\sigma^2XVX'B_V^{-2}XVX'+\rho^2B_V^{-1}[\tau^2XWX'+X(\theta-\mu)(\theta-\mu)'X']B_V^{-1}$. 故在条件 (2.4) 之下有

$$\operatorname{tr}\{\sigma^2XVX'B_V^{-2}XVX'\}\le\operatorname{tr}\{M(\widehat{\eta}_B)\}\le\operatorname{tr}\{\tau^2B_V^{-1}[\rho(XVX')^2+2\rho^2XVX']B_V^{-1}\}. \tag{3.4}$$

因此

$$\frac{1}{r}\operatorname{tr}\{XVX'B_V^{-2}XVX')\le e(\widehat{\eta}_{LS},\widehat{\eta}_B\}\le\frac{1}{r}\operatorname{tr}\{B_V^{-1}[(XVX')^2+2\rho XVX']B_V^{-1}\}. \tag{3.5}$$

由 (3.1) 和 (3.2) 可知

$$\operatorname{tr}\{XVX'B_V^{-2}XVX'\}=\operatorname{tr}\{B_V^{-1}(XVX')^2B_V^{-1}\}=\operatorname{tr}\{\Delta^2\Lambda^2\}=\sum_{i=1}^{t}\frac{\lambda_i^2}{(\rho+\lambda_i)^2}, \tag{3.6}$$

$$\operatorname{tr}\{B_V^{-1}[(XVX')^2+2\rho XVX']B_V^{-1}\}=\operatorname{tr}\{B_V^{-1}[B_V^2-\rho^2I_n]B_V^{-1}\}=\sum_{i=1}^{t}\frac{\lambda_i^2+2\rho\lambda_i}{(\rho+\lambda_i)^2}. \tag{3.7}$$

把 (3.6) 和 (3.7) 代入 (3.5), 从而定理得证.　　#

为考虑一般的可估函数 $G\beta=LX\beta$ 的 BLUE 和 LSE 的相对效率，我们需要如下引理

引理 3.1　设 A 与 B 为两个 n 阶 Hermite 阵，它们的特征值分别为 $\lambda_1\ge\cdots\ge\lambda_n$ 和 $\mu_1\ge\cdots\ge\mu_n$. 则

$$\sum_{i=1}^{n}\lambda_i\mu_{n-i+1}\le\operatorname{tr}\{AB\}\le\sum_{i=1}^{n}\lambda_i\mu_i,$$

并且左边不等式等号成立 $\Longleftrightarrow B=\sum_{i=1}^{n}\mu_{n-i+1}\varphi_i\varphi_i'$, 而右边不等式等号成立 $\Longleftrightarrow B=\sum_{i=1}^{n}\mu_i\varphi_i\varphi_i'$. 这里 $\varphi_1,\cdots,\varphi_n$ 为 A 的对应于 $\lambda_1,\cdots,\lambda_n$ 的标准正交特征向量.

证明：　见王松桂等 [6] 定理 6.4.1.　　#

定理 3.2　设模型 (1.1) 和错误先验假定 (1.2) 下 $G\beta$ 的 LSE 和 BLUE 分别为 $L\widehat{\eta}_{LS}$ 和 $L\widehat{\eta}_B$, 其中 $\widehat{\eta}_{LS}$ 和 $\widehat{\eta}_B$ 分别由 (1.4) 和 (1.8) 给出，则在正确先验 (1.3) 之下，当条件 (2.4) 成立时，有

$$\sum_{i=1}^{t}\frac{\lambda_i^2\mu_{n-i+1}}{(\rho+\lambda_i)^2}\Big/\sum_{i=1}^{r}\mu_i\le e(L\widehat{\eta}_{LS},L\widehat{\eta}_B)\le\sum_{i=1}^{t}\frac{(\lambda_i^2+2\rho\lambda_i)\mu_i}{(\rho+\lambda_i)^2}\Big/\sum_{i=1}^{r}\mu_{n-i+1}.$$

其中 $\mu_1\ge\cdots\ge\mu_n\ge0$ 为 LL' 的特征根，　$\lambda_i\ (i=1,\cdots,t)$ 如定理 3.1 所述.

证明：　由于 $G\beta$ 的 LSE $L\widehat{\eta}_{LS}$ 和 BLUE $L\widehat{\eta}_B$ 的均方误差矩阵的迹为

$$\operatorname{tr}\{M(L\widehat{\eta}_{LS})\}=\operatorname{tr}\{M(\widehat{\eta}_{LS})L'L\},\qquad\operatorname{tr}\{M(L\widehat{\eta}_B)\}=\operatorname{tr}\{M(\widehat{\eta}_B)L'L\}. \tag{3.8}$$

由于 $X(X'X)^{-}X'$ 的全部非零特征根为 r 个 1, 从而由 (2.2) 和引理 3.1 有

$$\sigma^2\sum_{i=1}^{r}\mu_{n-i+1}\le\operatorname{tr}\{M(\widehat{\eta}_{LS})L'L\}\le\sigma^2\sum_{i=1}^{r}\mu_i. \tag{3.9}$$

注意到事实：若 $B\ge0$, $C\le A\le D$, 则有 $\operatorname{tr}\{CB\}\le\operatorname{tr}\{AB\}\le\operatorname{tr}\{DB\}$, 再由 (3.1), (3.2) 和 (3.4) 及引理 3.1 有

$$\sigma^2\sum_{i=1}^{t}\frac{\lambda_i^2\mu_{n-i+1}}{(\rho+\lambda_i)^2}\le\operatorname{tr}\{M(\widehat{\eta}_B)L'L\}\le\sigma^2\sum_{i=1}^{t}\frac{(\lambda_i^2+2\rho\lambda_i)\mu_i}{(\rho+\lambda_i)^2}. \tag{3.10}$$

由 (3.9) 和 (3.10) 及定义 3.1 定理得证.　　#

§4. Bayes 估计相对于 LS 估计的 PPC 性质

Pitman Closeness (PC) 准则是 Pitman[7] 提出的，这一方法被忽略了相当长时间，直到 Rao[8,9] 有关 PC 准则的论文发表后， PC 准则作为比较不同估计量优劣的又一准则，受到了统计学者的广泛关注．Pitman Closeness 准则的定义如下：

定义 4.1 设有分布族 $\{\mathsf{P}_\theta(x), \theta\in\Theta\}$, Θ 为参数空间．$L(\widehat{\theta},\theta)$ 为损失函数．$\widehat{\theta}_1$, $\widehat{\theta}_2$ 为 θ 的两个不同估计量，称 $\widehat{\theta}_1$ 在 PC 准则下优于 $\widehat{\theta}_2$, 如果

$$\mathsf{P}(L(\widehat{\theta}_1,\theta)\le L(\widehat{\theta}_2,\theta))\ge 0.5,\qquad \text{对一切}\ \ \theta\in\Theta,$$

且严格不等号 $>$ 至少对某个 $\theta\in\Theta$ 成立．

Ghosh & Sen[10] 从 Bayes 角度引入 PC 准则的两个替代概念, 其中一个称为 Posterior Pitman Closeness (PPC) 准则．PPC 准则定义如下：

定义 4.2 设 π 为 θ 的一个先验分布， $\delta_1=\delta_1(x)$, $\delta_2=\delta_2(x)$ 为 θ 的两个不同估计量，称估计量 δ_1 在 PPC 准则下优于估计量 δ_2, 如果

$$\mathsf{P}_\pi(L(\delta_1,\theta)\le L(\delta_2,\theta)|x)\ge 0.5$$

对一切 $x\in\mathcal{X}$ 成立，右边不等号 “$>$” 至少对某个 $x\in\mathcal{X}$ 成立．此处 $\mathcal{X}$ 表示样本空间．

为讨论回归系数 β 的 Bayes 估计相对于 LS 估计的 PPC 特性, 需要假定误差 $e\sim N(0,\sigma^2 I_n)$, 还需对先验分布作进一步的假定．设先验分布 (1.2) 和 (1.3) 为正态，即

$$\pi_0(\beta):\ \ \beta\sim N(\mu,\tau^2 V) \tag{4.1}$$

为错误指定的先验分布．正确的先验分布为

$$\pi(\beta):\ \ \beta\sim N(\theta,\tau^2 W). \tag{4.2}$$

引理 4.1 设模型 (1.1) 随机误差 $e\sim N(0,\sigma^2 I_n)$ 且假定参数 β 的正确先验假定为 (4.2), 则 $\eta=X\beta$ 的 BLUE 为

$$\widehat{\eta}_r = XWX'(\rho I_n+XWX')^{-1}Y+\rho(\rho I_n+XWX')^{-1}X\theta.$$

证明： 在先验假定 (4.2) 中，若 $\theta=0$, $W=I_p$, 则易知 β 的 Bayes 估计为

$$\widehat{\beta}=(X'X+\rho I_p)^{-1}X'Y. \tag{4.3}$$

对一般的 θ 和 $W\ge 0$, 由于 $\beta\sim N(\theta,\tau^2 W)$, $\mathrm{Rank}(W)=s$ 的充分必要条件为 $\beta=\theta+Q\alpha$, $\alpha\sim N_s(0,\tau^2 I_s)$, 其中 Q 为 $p\times s$ 满足 $QQ'=W$. 因此作变换 $\widetilde{Y}=Y-X\theta$, $\widetilde{X}=XQ$, 则得到

$$\begin{cases}\widetilde{Y}=\widetilde{X}\alpha+e, & e\sim N(0,\sigma^2 I_n),\\ \pi(\alpha):\ \ \alpha\sim N_s(0,\tau^2 I_s).\end{cases} \tag{4.4}$$

由 (4.3) 知 α 的 Bayes 估计为

$$\widehat{\alpha}=(\widetilde{X}'\widetilde{X}+\rho I)^{-1}\widetilde{X}'\widetilde{Y}. \tag{4.5}$$

因此 $\eta = X\beta$ 的 Bayes 估计为

$$\begin{aligned}
\widehat{\eta}_r &= X\theta + XQ\widehat{\alpha} \\
&= X\theta + \widetilde{X}(\widetilde{X}'\widetilde{X} + \rho I)^{-1}\widetilde{X}'\widetilde{Y} \\
&= X\theta + \widetilde{X}\widetilde{X}'(\widetilde{X}\widetilde{X}' + \rho I)^{-1}\widetilde{Y} \\
&= XWX'(\rho I_n + XWX')^{-1}Y + \rho(\rho I_n + XWX')^{-1}X\theta.
\end{aligned} \tag{4.6}$$

其中第三个等式利用了代数事实：$(P+BCB')^{-1} = P^{-1} - P^{-1}B(B'P^{-1}B + C^{-1})^{-1}B'P^{-1}$. 引理得证．　#

由此引理知在错误先验假定 (4.1) 下 η 的 BLUE $\widehat{\eta}_B$ 仍为 (1.8). 设 η 的 LSE $\widehat{\eta}_{LS}$ 仍由 (1.4) 表示，并记 $(\xi_1 - \xi_2)'D(\xi_1 - \xi_2) = \|\xi_1 - \xi_2\|_D^2$, 其中 ξ_1, ξ_2 均为向量，D 为非负定对称阵．当 D 为单位阵时，则用 $\|\xi_1 - \xi_2\|^2$ 表示．设损失函数为二次损失，即

$$L(\delta(x), \eta) = (\delta(x) - \eta)'(\delta(x) - \eta) = \|\delta(x) - \eta\|^2. \tag{4.7}$$

记

$$W(\widehat{\eta}_B, \widehat{\eta}_{LS}; \eta) = L(\widehat{\eta}_B, \eta) - L(\widehat{\eta}_{LS}, \eta).$$

我们有

定理 4.1　设模型 (1.1) 和错误先验假定 (1.2) 下 η 的 LSE 和 BLUE 分别由 (1.4) 和 (1.8) 给出，令 $A = B_W^{-1}P_X B_W^{-1}$, $C = B_W^{-1} - B_V^{-1}$, 其中 B_V 和 B_W 由第二节给出．如果

$$\|Y - X\theta\|_A^2 \geq (\|Y - X\theta\|_{C^2} + \|\mu - \theta\|_{X'B_V^{-2}X})^2, \tag{4.8}$$

则有

$$\mathsf{P}_\pi(W(\widehat{\eta}_B, \widehat{\eta}_{LS}; \eta) \leq 0 | Y) > 0.5.$$

此处 P_π 是关于正确先验分布 (4.2) 计算的．

注 4.1　条件 (4.8) 对随机误差 e 和先验误差之间做出了一定的限制，特别当 $V = W$ 时，此条件可以表示成 $\|P_X B_W^{-1} e\| \geq \|B_W^{-1}X(\mu - \theta)\|$, 说明先验误差不能太大而超过随机误差．

证明：　因为

$$\begin{aligned}
W(\widehat{\eta}_B, \widehat{\eta}_{LS}; \eta) &= \|(\widehat{\eta}_B - \widehat{\eta}_{LS}) + (\widehat{\eta}_{LS} - \eta)\|^2 - \|\widehat{\eta}_{LS} - \eta\|^2 \\
&= \|\widehat{\eta}_B - \widehat{\eta}_{LS}\|^2 + 2(\widehat{\eta}_B - \widehat{\eta}_{LS})'(\widehat{\eta}_{LS} - \eta) \\
&= \|(\widehat{\eta}_B - \widehat{\eta}_r) + (\widehat{\eta}_r - \widehat{\eta}_{LS})\|^2 + 2(\widehat{\eta}_B - \widehat{\eta}_r + \widehat{\eta}_r - \widehat{\eta}_{LS})'(\widehat{\eta}_{LS} - \widehat{\eta}_r + \widehat{\eta}_r - \eta) \\
&= \|\widehat{\eta}_B - \widehat{\eta}_r\|^2 - \|\widehat{\eta}_r - \widehat{\eta}_{LS}\|^2 + 2(\widehat{\eta}_B - \widehat{\eta}_{LS})'(\widehat{\eta}_r - \eta).
\end{aligned} \tag{4.9}$$

由 (4.9) 可知 $W(\widehat{\eta}_B, \widehat{\eta}_{LS}; \eta) \leq 0$ 等价于

$$2(\widehat{\eta}_B - \widehat{\eta}_{LS})'(\widehat{\eta}_r - \eta) \leq \|\widehat{\eta}_{LS} - \widehat{\eta}_r\|^2 - \|\widehat{\eta}_B - \widehat{\eta}_r\|^2. \tag{4.10}$$

因为 $B_V = \rho I_n + XVX'$, $P_X = X(X'X)^- X'$, $B_W = \rho I_n + XWX'$, 而且 $P_X B_W = \rho P_X + XWX'$ 及 $P_X B_W^{-1} = \rho^{-1}(P_X - I_n) + B_W^{-1}$, 从而 $\rho B_W^{-1}X = \rho P_X B_W^{-1}X$. 我们有

$$\begin{aligned}
\widehat{\eta}_{LS} - \widehat{\eta}_r &= X(X'X)^- X'Y - XWX'B_W^{-1}Y - \rho B_W^{-1}X\theta \\
&= (P_X - I_n + I_n)Y - (B_W - \rho I_n)B_W^{-1}Y - \rho B_W^{-1}X\theta \\
&= \rho P_X B_W^{-1}(Y - X\theta),
\end{aligned} \tag{4.11}$$

$$\begin{aligned}\widehat{\eta}_B-\widehat{\eta}_r &= XVX'B_V^{-1}Y+\rho B_V^{-1}X\mu-XWX'B_W^{-1}Y-\rho B_W^{-1}X\theta\\ &=(B_V-\rho I_n)B_V^{-1}Y+\rho B_V^{-1}X\mu-(B_W-\rho I_n)B_W^{-1}Y-\rho B_W^{-1}X\theta\\ &=\rho(B_W^{-1}-B_V^{-1})(Y-X\theta)+\rho B_V^{-1}X(\mu-\theta). \end{aligned}\tag{4.12}$$

将 (4.11) 和 (4.12) 代入 (4.10) 得

$$\begin{aligned}&2(\widehat{\eta}_B-\widehat{\eta}_{LS})'(\widehat{\eta}_r-\eta)\\ \le\ &\rho^2\{\|Y-X\theta\|_A^2-\|Y-X\theta\|_{C^2}^2-\|\mu-\theta\|^2_{X'B_V^{-2}X}-2(Y-X\theta)'CB_V^{-1}X(\mu-\theta)\}.\end{aligned}\tag{4.13}$$

利用 Cauchy-Schwarz 不等式: $x'DEy\le(x'DD'x)^{1/2}(y'E'Ey)^{1/2}$, 可知对上式右边最后一项有

$$2\rho^2(Y-X\theta)'CB_V^{-1}X(\mu-\theta)\le 2\rho^2\|Y-X\theta\|_{C^2}\cdot\|\mu-\theta\|_{X'B_V^{-2}X},$$

故知 (4.13) 式蕴含了下式

$$2(\widehat{\eta}_B-\widehat{\eta}_{LS})'(\widehat{\eta}_r-\eta)\le\rho^2\{\|Y-X\theta\|_A^2-(\|Y-X\theta\|_{C^2}+\|\mu-\theta\|_{X'B_V^{-2}X})^2\}.\tag{4.14}$$

因为 $W(\widehat{\eta}_B,\widehat{\eta}_{LS};\eta)\le 0\Longleftrightarrow$ (4.10) $\Longleftrightarrow$ (4.13), 而 (4.13) 蕴含了 (4.14), 因此

$$\begin{aligned}&\mathsf{P}_\pi(W(\widehat{\eta}_B,\widehat{\eta}_{LS};\eta)\le 0|Y)\\ \ge\ &\mathsf{P}_\pi(2(\widehat{\eta}_{LS}-\widehat{\eta}_B)'(\eta-\widehat{\eta}_r)\le\rho^2\{\|Y-X\theta\|_A^2-(\|Y-X\theta\|_{C^2}+\|\mu-\theta\|_{X'B_V^{-2}X})^2\}|Y).\end{aligned}$$

由于 β 的正确先验为正态分布，故给定 Y 时 β 的后验分布仍为正态分布．而 $\widehat{\eta}_r$ 就是 η 的后验均值，故 $(\eta-\widehat{\eta}_r)|Y$ 的分布为均值为 0 的正态分布．因此当条件 (4.8) 成立时，上式概率大于等于 1/2. 定理得证． #

对一般的可估函数 $G\beta=L\eta$, 记 $M=L'L$, $A_M=B_W^{-1}P_XMP_XB_W^{-1}$, 我们有

定理 4.2 设模型 (1.1) 和错误先验假定 (1.2) 下 η 的 LSE 和 BLUE 由 (1.4) 和 (1.8) 给出，若

$$\|Y-X\theta\|_{A_M}^2\ge(\|Y-X\theta\|_{CMC}+\|\mu-\theta\|_{X'B_V^{-1}MB_V^{-1}X})^2,\tag{4.15}$$

则有

$$\mathsf{P}_\pi(W(L\widehat{\eta}_B,L\widehat{\eta}_{LS};L\eta)\le 0|Y)\ge 0.5.$$

此处 P_π 是关于正确先验分布 (4.2) 计算的．

证明: 因为

$$W(L\widehat{\eta}_B,L\widehat{\eta}_{LS};L\eta)=\|\widehat{\eta}_B-\widehat{\eta}_r\|_M^2-\|\widehat{\eta}_r-\widehat{\eta}_{LS}\|_M^2+2(\widehat{\eta}_B-\widehat{\eta}_{LS})'M(\widehat{\eta}_r-\eta).\tag{4.16}$$

由 (4.11) 和 (4.12) 易知

$$\begin{aligned}&\|\widehat{\eta}_r-\widehat{\eta}_{LS}\|_M^2-\|\widehat{\eta}_B-\widehat{\eta}_r\|_M^2\\ =\ &\rho^2\{\|Y-X\theta\|^2_{A_M-CMC}-\|\mu-\theta\|^2_{X'B_V^{-1}MB_V^{-1}X}-2(Y-X\theta)'CMB_V^{-1}X(\mu-\theta)\}.\end{aligned}$$

由 Cauchy-Schwarz 不等式知

$$(Y-X\theta)'CMB_V^{-1}X(\mu-\theta)\le\|Y-X\theta\|_{CMC}\cdot\|\mu-\theta\|_{X'B_V^{-1}MB_V^{-1}X}.\tag{4.17}$$

因此当条件 (4.15) 满足时

$$
\begin{aligned}
& \mathsf{P}_\pi(W(L\widehat{\eta}_B, L\widehat{\eta}_{LS}; L\eta) \le 0) \\
\ge\ & \mathsf{P}_\pi(2(\widehat{\eta}_{LS} - \widehat{\eta}_B)' M(\eta - \widehat{\eta}_r) \\
& \le \rho^2[\|Y - X\theta\|^2_{A_M} - (\|Y - X\theta\|_{CMC} + \|\mu - \theta\|_{X'B_V^{-1}MB_V^{-1}X})^2] | Y) \\
\ge\ & 0.5.
\end{aligned}
$$

定理得证．　　#

若先验分布 (4.1) 是正确的，定理 4.1 的条件必然成立，因此有如下结论：

推论 4.1　若先验分布 (4.1) 是正确的，即为正态分布 $N(\theta, \tau^2 W)$，则必有

$$\mathsf{P}_\pi(W(L\widehat{\eta}_B, L\widehat{\eta}_{LS}; L\eta) \le 0|Y) > 0.5.$$

参 考 文 献

[1] 韦来生，错误先验假定下回归系数 Bayes 估计的小样本性质，应用概率统计，**16(1)**(2000), 71–80.

[2] 王松桂，线性模型的理论及其应用，安徽教育出版社，合肥，1987.

[3] Rao, C.R., *Linear Statistical Inference and Its Applications*, Second Edition, New York: Wiley, 1973.

[4] Gruber, M.H.J., *Regression Estimators, A Comparative Study*, Boston: Academic Press, 1990.

[5] 刘爱义，王松桂，线性模型中最小二乘估计的一种新的相对效率，应用概率统计，**5(1)**(1989), 97–104.

[6] 王松桂，贾忠贞，矩阵论中不等式，安徽教育出版社，合肥，1994.

[7] Pitman, E.J.G., The closest estimates of statistical parameters, *Proc. Camb. Phil. Soc.*, **33**(1937), 212–222.

[8] Rao, C.R., Some comments on the minimum mean square error as criterion of estimation, In *Statistics and Related Topics*, North Holland, Ameserdam, 1981,

[9] Rao, C. R., Keating, J.P. and Mason, R.L., The Pitman nearness criterion and its determination, *Commun. Statist. Theory and Methods*, **15**(1986), 3173–3191.

[10] Ghosh, M. and Sen, P.K., Bayesian Pitman closeness, *Comm. Statist. Theory and Methods*, **20**(1991), 3659–3678.

The Robustness of Bayes Linear Unbiased Estimations under Misspecified Prior Assumption

ZHANG WEIPING　　　WEI LAISHENG

(*Department of Statistics and Finance, University of Science and Technology of China, Hefei, 230026*)

The Bayes linear unbiased estimator (BLUE) of parameters is derived for general linear model under misspecified prior assumption. The robustness of BLUE over ordinary least square estimator (LSE) are shown in terms of mean square error matrix criterion and Pitman closeness criterion, the bounds of the relative efficiency of estimators are obtained.