文章编号:1672-058X(2013)07-0012-04

# 线性模型中的一种新的 k-d 类估计

汪 国 平
(重庆大学 数学与统计学院,重庆 401331)

**摘　要**:提出了线性模型中考虑样本信息中关于参数的先验信息的一种新的两参数估计,并给出了在均方误差矩阵准则下这种新的两参数估计优于 Liu 估计和修正 Liu 估计.

**关键词**:新的 k-d 类估计;修正 Liu 估计;修正岭估计;先验信息
**中图分类号**:O212.4　　**文献标志码**:A

## 1　引言与基础知识

考虑线性模型

$$Y = X\beta + \varepsilon \tag{1}$$

其中 $X_{n\times p}$ 是已知的列满秩设计矩阵,$Y_{n\times 1}$ 为观测向量,$\beta_{n\times 1}$ 为未知参数向量,$\varepsilon_{n\times 1}$ 为随机误差向量,并且 $E(\varepsilon)=0$, $\mathrm{Cov}(\varepsilon)=\sigma^2 I_n$, $I_n$ 是一个 $n\times n$ 阶的单位矩阵.根据 Gauss-Markov 定理,模型(1)的回归系数 $\beta$ 的最小二乘估计(LSE):

$$\hat{\beta}_{\mathrm{LSE}} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y \tag{2}$$

具有很多优良的性质,它是线性无偏估计中方差最小的.但当模型(1)存在复共线性时,即设计矩阵 $X$ 呈病态时,LSE 估计 $\hat{\beta}_{\mathrm{LSE}}$ 变得极不稳定,并且经常给出有误导性的信息.为了克服复共线性,一些学者提出了很多估计方法,其中很重要的一类估计就是有偏估计.

Hoerl 和 Kennard(1970)[2] 提出了岭估计(RE):

$$\hat{\beta}_{\mathrm{RE}}(k) = (X^{\mathrm{T}}X + kI)^{-1}X^{\mathrm{T}}Y \tag{3}$$

其中 $k\geqslant 0$ 是有偏参数.

Liu(1993) 结合 Stein [3] 估计和岭估计,提出 Liu 估计(LE):

$$\hat{\beta}_{\mathrm{LE}}(d) = (X^{\mathrm{T}}X + I)^{-1}(X^{\mathrm{T}}Y + dI)\hat{\beta}_{\mathrm{LSE}} \tag{4}$$

其中 $0<d<1$ 是有偏参数.

考虑先验信息,Swindel(1976)[4] 提出了一种修正的岭估计(MRE):

$$\hat{\beta}_{\mathrm{MRE}}(k,b_0) = (X^{\mathrm{T}}X + kI)^{-1}(X^{\mathrm{T}}Y + kb_0) \tag{5}$$

其中 $b_0$ 是给定的非随机的向量,被选择过来代表的就是关于 $\beta$ 的先验信息.很显然,当 $k$ 趋近于无穷大时,MRE 趋近于 $b_0$;特别是当 $k=0$ 时,MRE 就是 LSE.

同样考虑先验信息,Li 和 Yang(2012)[5] 提出了一种新的 Liu 类估计,即修正的 Liu 估计(MLE):

---

**收稿日期**:2013-03-06;**修回日期**:2013-04-07.
**作者简介**:汪国平(1986-),男,江西人,硕士研究生,从事线性模型研究.

$$\hat{\beta}_{\mathrm{MLE}}(d,b_0)=(X^{\mathrm{T}}X+I)^{-1}[(X^{\mathrm{T}}Y+dI)\hat{\beta}_{\mathrm{LSE}}+(1-d)b_0] \tag{6}$$

很容易看到,$\hat{\beta}_{\mathrm{MLE}}(1,b_0)=\hat{\beta}_{\mathrm{LSE}}$,$\hat{\beta}_{\mathrm{MLE}}(0,b_0)=(X^{\mathrm{T}}X+I)^{-1}(X^{\mathrm{T}}Y+b_0)$ 以及 $\hat{\beta}_{\mathrm{MLE}}(d,0)=\hat{\beta}_{\mathrm{LE}}(d)$. 正如 Swindel(1976)所指出的那样,在运用中考虑先验信息并且检验此时的回归系数的估计会更加的有用和合理.

同时,研究者不仅要在 LSE 和有偏估计之间进行选择,还要在两个有偏估计之间进行选择.很显然,有偏估计之间的比较不管是理论上还是在实际应用中都是很有用的,一些学者已经比较了一些有偏估计. Sakallioglu(2001)比较了 RE,LE 和迭代估计(当出现复共线性时,对 LSE 采用迭代的方法构造出的估计);Akdeniz 和 Erol(2003)在均方误差矩阵准则下比较了几乎无偏广义岭估计和几乎无偏 Liu 估计;Sakallioglu 等(2008)获得了一种新的有偏估计,并且与 LSE,Liu 估计,两类 Liu 估计进行了对比;Li 和 Yang(2012)提出了一种修正的 Liu 估计(MLE),并且在均方误差矩阵准则下比较了 MLE,LSE,LE,RE,MRE.

此处考虑先验信息,借助于文献的方法,提出了一种新的两参数估计(MTPE),并且在均方误差矩阵准则下把其与 LE,RE,MRE 和 MLE 分别进行比较,讨论新估计的优良性.

## 2　基本引理

**引理 1**　假定矩阵 $M$ 正定,即 $M>0$. $\alpha$ 为任意给定向量,则

$$M-\alpha\alpha^{\mathrm{T}}>0 \Leftrightarrow \alpha^{\mathrm{T}}M^{-1}\alpha<1$$

**证明**　见文献[1].

**引理 2**　假定 $\hat{\beta}_1=A_1Y$,$\hat{\beta}_2=A_2Y$ 是参数 $\beta$ 的两个估计,设 $D=(A_1A_1^{\mathrm{T}}-A_2A_2^{\mathrm{T}})>0$. 则

$$\Delta(\hat{\beta}_1,\hat{\beta}_2)=\mathrm{MSEM}(\hat{\beta}_1)-\mathrm{MSEM}(\hat{\beta}_2)=\sigma^2D+b_1b_1^{\mathrm{T}}-b_2b_2^{\mathrm{T}}>0 \tag{7}$$

当且仅当

$$b_2^{\mathrm{T}}[\sigma^2D+b_1b_1^{\mathrm{T}}]b_2<1 \tag{8}$$

其中,$\mathrm{MSEM}(\hat{\beta}_i)=\mathrm{Cov}(\hat{\beta}_i)+b_ib_i^{\mathrm{T}}$,$b_i=\mathrm{Bias}(\hat{\beta}_i)=(A_iX-I)\beta$,$i=1,2$.

**证明**　见文献[6].

## 3　新估计的提出及其性质

假定 $T_k=(X^{\mathrm{T}}X+kI)^{-1}X^{\mathrm{T}}X=I-k(X^{\mathrm{T}}X+kI)^{-1}$,此时可以将 MRE(5)改写为

$$\begin{gathered}\hat{\beta}_{\mathrm{MRE}}(k,b_0)=(X^{\mathrm{T}}X+kI)^{-1}X^{\mathrm{T}}Y+k(X^{\mathrm{T}}X+kI)^{-1}b_0=\\ T_k\hat{\beta}_{\mathrm{LSE}}+(I-T_k)b_0\end{gathered} \tag{9}$$

同样假定 $F_d=(X^{\mathrm{T}}X+I)^{-1}(X^{\mathrm{T}}X+dI)=I-(1-d)(X^{\mathrm{T}}X+I)^{-1}$,此时可以将 MLE(6)改写为

$$\begin{gathered}\hat{\beta}_{\mathrm{MLE}}(d,b_0)=F_d\hat{\beta}_{\mathrm{LSE}}+(I-F_d)b_0=\\ (X^{\mathrm{T}}X+I)^{-1}(X^{\mathrm{T}}X+dI)\hat{\beta}_{\mathrm{LSE}}+(1-d)(X^{\mathrm{T}}X+I)^{-1}b_0=\\ (X^{\mathrm{T}}X+I)^{-1}[(X^{\mathrm{T}}X+dI)\hat{\beta}_{\mathrm{LSE}}+(1-d)b_0]\end{gathered} \tag{10}$$

从上面可以看出,MRE 实际上是先验信息和 LSE 的凸组合,而 MLE 是先验信息和 MLE 的凸组合.现在假定 $F_{k,d}=(X^{\mathrm{T}}X+I)^{-1}[X^{\mathrm{T}}X+(d+k)I](X^{\mathrm{T}}X+kI)^{-1}X^{\mathrm{T}}X$,于是类似地考虑得到了所要考虑的新估计,即修正的两参数估计(MTPE):

$$\hat{\beta}_{\mathrm{MTPE}}(k,d,b_0)=F_{k,d}\hat{\beta}_{\mathrm{LSE}}+(I-F_{k,d})b_0 \tag{11}$$

下面讨论新估计 $\hat{\beta}_{\mathrm{MTPE}}(k,d,b_0)$ 的几个性质:

**性质 1**　若 $d+k=1$,则 $\hat{\beta}_{\mathrm{MTPE}}(k,k-1,b_0)=\hat{\beta}_{\mathrm{MRE}}(k,b_0)$.

**性质2** 若 $k=0$,则 $\hat{\beta}_{\mathrm{MTPE}}(0,d,b_0)=\hat{\beta}_{\mathrm{MLE}}(d,b_0)$.

**性质3** 只要 $k\neq 0$ 且 $d+k\neq 1$,那么 $\hat{\beta}_{\mathrm{MTPE}}(k,d,b_0)$都是有偏估计.

**证明** 从新估计 $\hat{\beta}_{\mathrm{MTPE}}(k,d,b_0)$的定义中就可以很容易的得到性1、性质2、性质3.

## 4 均方误差矩阵准则下 MTPE 与 MRE 及 MLE 的性能比较

现在把模型(1)写成典型的形式

$$Y = Z\alpha + \varepsilon \tag{12}$$

其中,$Z=XQ$,$\alpha=Q^{\mathrm{T}}\beta$ 且 $Q$ 是正交矩阵,它的列向量构成了矩阵 $X^{\mathrm{T}}X$ 的特征向量.那么

$$Z^{\mathrm{T}}Z = Q^{\mathrm{T}}X^{\mathrm{T}}XQ = \Lambda = \mathrm{diag}\ (\lambda_1,\cdots,\lambda_p) \tag{13}$$

其中,$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p>0$ 是 $X^TX$ 的有序特征值.对于模型(12),可以得到下面的表达式

$$\hat{\alpha}_{\mathrm{LSE}} = \Lambda^{-1}Z^{\mathrm{T}}Y \tag{14}$$

$$\hat{\alpha}_{\mathrm{MRE}}(k,b) = (\Lambda + kI)^{-1}(Z^{\mathrm{T}}Y + kb) \tag{15}$$

$$\hat{\alpha}_{\mathrm{MLE}}(d,b) = (\Lambda + I)^{-1}[(\Lambda + dI)\Lambda^{-1}Z^{\mathrm{T}}Y + (1-d)b] \tag{16}$$

$$\begin{aligned}\hat{\alpha}_{\mathrm{MTPE}}(k,d,b_0) = {} & (\Lambda + I)^{-1}[\Lambda + (d+k)I](\Lambda + kI)^{-1}Z^{\mathrm{T}}Y + \\ & \{I - (\Lambda + I)^{-1}[\Lambda + (d+k)I](\Lambda + kI)^{-1}\Lambda\}b\end{aligned} \tag{17}$$

其中,$b=Q^{\mathrm{T}}b_0$.

由式(15)-(17),根据偏差向量、协方差矩阵和均方误差矩阵的定义,经简单计算可知,MTPE 的偏差向量和协方差矩阵为

$$B(\hat{\alpha}_{\mathrm{MTPE}}(k,d,b_0)) = A_{k,d}\alpha + (I - A_{k,d})b - \alpha = (A_{k,d} - I)(\alpha - b) = b_1 \tag{18}$$

$$\mathrm{Coc}(\hat{\alpha}_{\mathrm{MTPE}}(k,d,b_0)) = \sigma^2 A_{k,d}\Lambda^{-1}A_{k,d}^{\mathrm{T}} \tag{19}$$

其中,$A_{k,d}=(\Lambda+I)^{-1}[\Lambda+(d+k)I](\Lambda+kI)^{-1}\Lambda$.

通过式子(18)和(19),可以得到

$$\mathrm{MSEM}(\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)) = \sigma^2 A_{k,d}\Lambda^{-1}A_{k,d}^{\mathrm{T}} + b_1 b_1^{\mathrm{T}} \tag{20}$$

MLE 的偏差向量和协方差矩阵为

$$B(\hat{\alpha}_{\mathrm{MLE}}(d,b)) = A_d\alpha + (I - A_d)b - \alpha = (A_d - I)(\alpha - b) = b_2 \tag{21}$$

$$\mathrm{Cov}(\hat{\alpha}_{\mathrm{MLE}}(d,b)) = \sigma^2 A_d\Lambda^{-1}A_d^{\mathrm{T}} \tag{22}$$

其中,$A_d=(\Lambda+I)^{-1}(\Lambda+dI)$.

通过式子(21)和(22),可以得到

$$\mathrm{MSEM}(\hat{\alpha}_{\mathrm{MLE}}(d,b)) = \sigma^2 A_d\Lambda^{-1}A_d^{T} + b_2 b_2^{\mathrm{T}} \tag{23}$$

MRE 的的偏差向量和协方差矩阵为

$$B(\hat{\alpha}_{\mathrm{MRE}}(k,b)) = A_k\alpha + (I - A_k)b - \alpha = (A_k - I)(\alpha - b) = b_3 \tag{24}$$

$$\mathrm{Cov}(\hat{\alpha}_{\mathrm{MRE}}(d,b)) = \sigma^2 A_k\Lambda^{-1}A_k^{\mathrm{T}} \tag{25}$$

其中,$A_k=(\Lambda+kI)^{-1}\Lambda$.

通过式子(24)和(25),可以得到

$$\mathrm{MSEM}(\hat{\alpha}_{\mathrm{MRE}}(d,b)) = \sigma^2 A_k\Lambda^{-1}A_k^{\mathrm{T}} + b_3 b_3^{\mathrm{T}} \tag{26}$$

现在考虑如下差分:

$$\Delta_1 = \mathrm{MSEM}(\hat{\alpha}_{\mathrm{MLE}}(d,b)) - \mathrm{MSEM}(\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)) = \sigma^2 D_1 + b_2 b_2^{\mathrm{T}} - b_1 b_1^{\mathrm{T}} \tag{27}$$

$$\Delta_2 = \mathrm{MSEM}(\hat{\alpha}_{\mathrm{MRE}}(d,b)) - \mathrm{MSEM}(\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)) = \sigma^2 D_2 + b_3 b_3^{\mathrm{T}} - b_1 b_1^{\mathrm{T}} \tag{28}$$

其中,

$$
\begin{aligned}
D_1 = {} & \sigma^2 A_d \Lambda^{-1} A_d^{\mathrm{T}} - \sigma^2 A_{k,d} \Lambda^{-1} A_{k,d}^{\mathrm{T}} = \\
& \sigma^2 (\Lambda + I)^{-1} (\Lambda + dI) \Lambda^{-1} (\Lambda + dI)(\Lambda + I)^{-1} - \sigma^2 (\Lambda + I)^{-1} \\
& [\Lambda + (d+k)I](\Lambda + kI)^{-1} \Lambda \Lambda^{-1} \Lambda (\Lambda + kI)^{-1} [\Lambda + (d+k)I](\Lambda + I)^{-1} = \\
& \sigma^2 (\Lambda + I)^{-1} (\Lambda + dI) \Lambda^{-1} (\Lambda + dI)(\Lambda + I)^{-1} - \sigma^2 (\Lambda + I)^{-1} (\Lambda + kI)^{-1} \\
& [\Lambda + (d+k)I] \Lambda \Lambda^{-1} \Lambda [\Lambda + (d+k)I](\Lambda + kI)^{-1} (\Lambda + I)^{-1} = \\
& \sigma^2 (\Lambda + I)^{-1} (\Lambda + kI)^{-1} [2kd\Lambda^2 + 2kd(k+d)\Lambda + k^2 d^2 I] \Lambda^{-1} (\Lambda + kI)^{-1} (\Lambda + I)^{-1}
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
D_2 = {} & \sigma^2 A_k \Lambda^{-1} A_k^{\mathrm{T}} - \sigma^2 A_{k,d} \Lambda^{-1} A_{k,d}^{\mathrm{T}} = \\
& \sigma^2 (\Lambda + kI)^{-1} \Lambda \Lambda^{-1} \Lambda (\Lambda + kI)^{-1} - \sigma^2 (\Lambda + I)^{-1} \\
& [\Lambda + (d+k)I](\Lambda + kI)^{-1} \Lambda \Lambda^{-1} \Lambda (\Lambda + kI)^{-1} [\Lambda + (d+k)I](\Lambda + I)^{-1} = \\
& \sigma^2 (\Lambda + kI)^{-1} \Lambda \Lambda^{-1} \Lambda (\Lambda + kI)^{-1} - \sigma^2 (\Lambda + kI)^{-1} (\Lambda + I)^{-1} \\
& [\Lambda + (d+k)I] \Lambda \Lambda^{-1} \Lambda [\Lambda + (d+k)I](\Lambda + I)^{-1} (\Lambda + kI)^{-1} = \\
& \sigma^2 (\Lambda + kI)^{-1} (\Lambda + I)^{-1} \{2(1-d-k)\Lambda^2 + [1-(d+k)^2]\Lambda\} (\Lambda + I)^{-1} (\Lambda + kI)^{-1}
\end{aligned} \tag{30}
$$

下面给出在均方误差矩阵准则下,$\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)$优于$\hat{\alpha}_{\mathrm{MLE}}(d,b)$与$\hat{\alpha}_{\mathrm{MRE}}(d,b)$的充要条件.

**定理 1**　在均方误差矩阵准则下,$\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)$优于$\hat{\alpha}_{\mathrm{MLE}}(d,b)$的充要条件为

$$\mathrm{MSEM}(\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)) \leqslant \mathrm{MSEM}(\hat{\alpha}_{\mathrm{MLE}}(d,b)) \Leftrightarrow b_1^{\mathrm{T}} D_1^{-1} b_1 \leqslant 1$$

**证明**　因为 $k>0$ 和 $d>0$,并且 $\Lambda$ 是正交矩阵,所以从式(29)可知,$D_1>0$.由引理 2,从式(27)可知,

$$\Delta_1 \geqslant 0 \Leftrightarrow b_1^{\mathrm{T}} D_1^{-1} b_1 \leqslant 1$$

故定理 1 得证.

**定理 2**　若 $k>0$,$d>0$ 且 $k+d\leqslant 1$ 时,则在均方误差矩阵准则下,$\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)$优于$\hat{\alpha}_{\mathrm{MRE}}(d,b)$的充要条件为

$$\mathrm{MSEM}(\hat{\alpha}_{\mathrm{MTPE}}(k,d,b)) \leqslant \mathrm{MSEM}(\hat{\alpha}_{\mathrm{MRE}}(d,b)) \Leftrightarrow b_1^{\mathrm{T}} D_2^{-1} b_1 \leqslant 1$$

**证明**　因为 $k>0$,$d>0$ 且 $k+d\leqslant 1$,并且 $\Lambda$ 是正交矩阵,所以从式(30)可知,$D_2>0$.由引理 2,从式(28)可知,

$$\Delta_2 \geqslant 0 \Leftrightarrow b_1^{\mathrm{T}} D_2^{-1} b_1 \leqslant 1$$

故定理 2 得证.

**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] HOERL A E, Kennard R W.Ridge regression: biased estimation for non-orthogonal problems[J]. Technometrics,1970(12):55-68

[3] LIU K . Anewclass of biased estimate in linear regression[J].CommunStat Theory Methods,1993(22):393-402

[4] SWINDEL B F. Good ridge estimators based on prior information[J]. Commun Stat Theory Methods 1976,A5(11):1065-1075

[5] LI Y L,YANG H. A new Liu-type estimator in linear regression model[J].Stat Papers,2012(53):427-437

[6] 王松桂,吴密霞,贾忠贞. 矩阵不等式[M]. 北京: 科学出版社, 2006


(下转第 23 页)