从约束型最小二乘估计到约束型岭估计的探索

刘冬喜

（娄底职业技术学院，湖南 娄底 417000）

摘　要：考虑非齐次约束线性回归模型回归系数的最小二乘估计及岭估计，当设计阵X为病态时，最小二乘估计不再是一个优良估计，提出了一种约束型岭估计，从约束条件入手，借助条件极值法研究$(Y-X\beta)'(Y-X\beta)$的最小值，使问题转化为在两个约束条件下的优化问题。

关键词：约束型最小二乘估计；约束型岭估计；条件极值；均方误差阵

中图分类号：O212　　　　**文献标识码**：A　　　　**文章编号**：1673-9833(2008)02-0040-03

Exploration on the Restricted Least Squares Estimation to the Restricted Ridge Estimation

Liu Dongxi

（Loudi Vocational and Technical College，Loudi Hunan 417000）

Abstract: We considered the least squares estimation and the ridge estimation of regression coefficient in the heterogeneous restricted linear regression model. When the design matrix X is all ill-conditioned matrix, the least squares estimation is no longer a good estimation, The paper proposes a new restricted ridge estimation, obtains from the restricted condition, with the aid of condition extreme value, we study the minimization of $(Y-X\beta)'(Y-X\beta)$,and cause the question to transform as optimized question under two restricted conditions.

Key words: restricted least squares estimation; restricted ridge estimation; condition extreme value; mean squares error matrix

考虑带非齐次线性等式约束的线性回归模型：

$$Y = X\beta+\varepsilon,\varepsilon \sim (0,\sigma^2 I_n), R\beta = r, \tag{1}$$

式中：Y为$n\times 1$向量；

X为$n\times p$设计矩阵；

R为$m\times p$非零已知矩阵；

ε为$n\times 1$误差向量；

I_n为n阶单位阵；

$\beta\in B\triangleq\{\beta : R\beta = r\}$，为$p\times 1$的未知回归系数向量；

$\sigma^2>0$为误差向量；

r为$m\times 1$向量。

$M(R')\subset M(X')$，于是$R\beta$是m个线性无关的可估函数。

假设秩$(X)=p$, 秩$(R)=m$, 简记$X'X\triangleq S$，$X'X+kI\triangleq S_k$，$(X'X)^{-1}X'Y\triangleq b$，约束型最小二乘估计及约束型岭估计值分别记为$b$（$R$）和$b$（$k$, R）。

1　约束型最小二乘估计

利用拉格朗日乘数法求模型（1）的最小二乘解，构造辅助函数：

$$L(\beta,\lambda)=(Y-X\beta)'(Y-X\beta)+2\lambda'(R\beta-r),$$

其中λ为$m\times 1$向量。

分别对β，λ求导，并令它们等于零向量，得到：

$$\frac{\partial L(\beta,\lambda)}{\partial\beta}=-2X'Y+2X'X\beta+2R'\lambda=0, \tag{2}$$

$$\frac{\partial L(\beta,\lambda)}{\partial\lambda}=2(R\beta-r)=0。\tag{3}$$

收稿日期：2007-10-08

作者简介：刘冬喜（1971-），女，湖南娄底人，娄底职业技术学院讲师，湖南师范大学硕士研究生，主要研究方向为概率论与数理统计.

式（3）显然与约束条件相符。根据秩（X)=p，可知$X'X$非奇异，故β的解为：

$$\beta=b+S^{-1}R'\left(RS^{-1}R'\right)^{-1}(r-Rb)@b(R)。\quad（4）$$

证明　b（R）确实是线性约束$R\beta=r$下β的最小二乘解，只需证明如下两点：

1）$Rb(R)=r$；

2）对一切满足$R\beta=r$的β，都有

$$(Y-X\beta)'(Y-X\beta)\geq\left(Y-Xb(R)\right)'\left(Y-Xb(R)\right)。$$

结论1）容易验证；为了证明2)，将平方和（$Y-X\beta$）'（$Y-X\beta$）作如下分解：

$$(Y-X\beta)'(Y-X\beta)=(Y-Xb)'(Y-Xb)+$$

$$(b-\beta)'X'X(b-\beta)=(Y-Xb)'(Y-Xb)+$$

$$\left(b-b(R)+b(R)-\beta\right)'S\left(b-b(R)+b(R)-\beta\right)=$$

$$(Y-Xb)'(Y-Xb)+\left(b-b(R)\right)'S\left(b-b(R)\right)+$$

$$\left(b(R)-\beta\right)'S\left(b(R)-\beta\right)$$

上面用到了式（4）及$M(R')\subset M(X')$导出的下述关系：

$$\left(b-b(R)\right)'S\left(b(R)-\beta\right)=$$

$$(Rb-r)'\left(RS^{-1}R'\right)^{-1}\left(Rb(R)-R\beta\right)=0。$$

这个等式对一切满足$R\beta=r$的β都成立。从而对一切满足$R\beta=r$的β，总有

$$(Y-X\beta)'(Y-X\beta)\geq$$

$$(Y-Xb)'(Y-Xb)+\left(b-b(R)\right)'S\left(b-b(R)\right)。\quad（5）$$

等号成立，当且仅当$\left(b(R)-\beta\right)'X'X\left(b(R)-\beta\right)=0$，即$X\beta=Xb$（$R$），在式（5）中用$Xb$（$R$）代替$X\beta$，等式成立，即

$$\left(Y-Xb(R)\right)'\left(Y-Xb(R)\right)=(Y-Xb)'(Y-Xb)+$$

$$\left(b-b(R)\right)'S\left(b-b(R)\right)。$$

综上，证明了结论2）。

定理1[1]　对于线性模型（1），设R为$m\times p$矩阵，秩(R)=m，$M(R')\subset M(X')$且$R\beta=r$相容，若秩（X)=p，则b（R）=$b+S^{-1}R'$（$RS^{-1}R'$）$^{-1}$（$r-Rb$）为β的约束最小二乘估计。

一般地，对于线性模型（1），若秩（X）=$r<p$时，上述结论仍然成立，只需把逆改为广义逆即可，这时

$$b=(X'X)^{-1}X'Y,$$

$$b(R)=b+S^{-1}R'\left(RS^{-1}R'\right)^{-1}(r-Rb),$$

是β在约束条件$R\beta=r$下的约束LS解。

b（R）为一般的带约束的线性模型（1）的最小二乘估计，简称为ORLSE，也是最好线性无偏估计，但当设计阵X存在复共线性时，ORLSE往往不稳定或与实际结果大相径庭，主要原因在于$X'X$的病态性，使得（$X'X$）$^{-1}$的某些元素变大，最终估计出的参数β在P维空间的某些方向上严重偏离实际值。

为了克服这种病态现象，很多学者对此进行了探讨，提出了约束型岭估计。

2　约束型岭估计

在无线性约束$R\beta=r$的情况下，Hoerl和Kennard于1970年提出了应用广泛的岭回归方法[2]：

$$b(k)=\left(X'X+kI\right)^{-1}X'Y=S_K^{-1}X'Y，K\geq0。$$

当kI替换为对角阵K时，称为广义岭回归。通过取适当的k或K，可以用来减少病态的影响，这是一种有偏估计。

在处理约束条件的情况下，Sarkar于1992年引进了一种新的估计：

$$\beta_s(k)=T_k b(R),T(k)=\left(I_p+k(X'X)^{-1}\right)^{-1}，其中b（R）$$

为ORLSE。

通过引入T_k使得回归参数的方差降低，使估计值变得更加稳定，但用T_k进行修正不能彻底克服共线性带来的误差。最好的方法是从约束条件入手，找到一种类似于ORLSE参数估计方法，在过程中修正。

Swamy et al（1978）and Swamy and Mehta（1977）提到岭回归可以通过条件极值得到：

$$\min_\beta\left\{(Y-X\beta)'(Y-X\beta)\mid\beta'\beta\leq\phi^2\right\}^{[3]},$$

其中$\beta'\beta\leq\phi^2$是为了克服共线性而对参数估计范围的强制性约束。

同样，在约束条件$R\beta=r$下，很容易利用拉格朗日乘数法得到：

$$\min_\beta\left\{(Y-X\beta)'(Y-X\beta)\mid\beta'\beta\leq\phi^2\right\},$$

即问题转化为求解在两个约束条件下的最小二乘问题：

$$\min_\beta(Y-X\beta)'(Y-X\beta),$$

$$s.t\quad\beta\beta\leq\phi^2，R\beta=r。$$

对此最优问题,引入函数：

$$L(\beta,k,\lambda)=(Y-X\beta)'(Y-X\beta)+k\left(\beta'\beta-\phi^2\right)+2\lambda(R\beta-r),$$

其中$k\geq0$为常数，λ为$m\times1$向量。

分别对β,k,λ求导，并令它们等于零向量，得到：

$$\frac{\partial L(\beta,k,\lambda)}{\partial\beta}=-2X'Y+2X'X\beta+2k\beta+2R'\lambda=0，\quad（6）$$

$$\frac{\partial L(\beta,k,\lambda)}{\partial k}=\beta'\beta-\phi^2=0，\quad（7）$$

$$\frac{\partial L(\beta,k,\lambda)}{\partial\lambda}=2(R\beta-r)=0。\quad（8）$$

显然式（7）、（8）与条件相符。根据秩（$\boldsymbol{X}$）=p，$k\geq 0$，可知$\boldsymbol{X'X}+k\boldsymbol{I} @ \boldsymbol{S}_k$非奇异，故$\boldsymbol{\beta}$的解为

$$b(k,\boldsymbol{R})=\boldsymbol{S}_k^{-1}(\boldsymbol{X'Y}-\boldsymbol{R'}\boldsymbol{\lambda})。 \tag{9}$$

根据约束条件$\boldsymbol{R\beta}=\boldsymbol{r}$，秩（$\boldsymbol{R}$）=$m$，$\boldsymbol{RS}_K^{-1}\boldsymbol{R'}$非奇异，因而，

$$\boldsymbol{\lambda}=(\boldsymbol{RS}_K^{-1}\boldsymbol{R'})^{-1}(\boldsymbol{RS}_K^{-1}\boldsymbol{X'Y}-\boldsymbol{r}),$$

$$b(k,\boldsymbol{R})=\boldsymbol{S}_K^{-1}\left\{\boldsymbol{X'Y}-\boldsymbol{R'}\left(\boldsymbol{RS}_K^{-1}\boldsymbol{R'}\right)^{-1}\left(R\boldsymbol{S}_K^{-1}\boldsymbol{X'Y}-\boldsymbol{r}\right)\right\}=$$

$$b(k)+\boldsymbol{S}_k^{-1}\boldsymbol{R'}\left(\boldsymbol{RS}_K^{-1}\boldsymbol{R'}\right)^{-1}\left(\boldsymbol{r}-\boldsymbol{R}b(k)\right), \tag{10}$$

其中$b(k)=\boldsymbol{S}_K^{-1}\boldsymbol{X'Y}$为无条件约束的岭估计表达式。

对式（10）与一般带约束的最小二乘估计方法式（4）比较易发现，它们的表达方式完全一致，只是将一般最小二乘估计b换为岭估计$b(k)$，将$\boldsymbol{S}=\boldsymbol{X'X}$换为$\boldsymbol{S}_k=\boldsymbol{X'X}+k\boldsymbol{I}$，从而整个表达式消除了（$\boldsymbol{X'X}$）$^{-1}$，克服了$\boldsymbol{X'X}$的病态带来估计参数的不稳定现象。

定理 2 对于线性模型（1），设$\boldsymbol{R}$为$m\times p$矩阵，秩（$\boldsymbol{R}$）=m，$\boldsymbol{M}(\boldsymbol{R'})\subset \boldsymbol{M}(\boldsymbol{X'})$且$\boldsymbol{R\beta}=\boldsymbol{r}$相容，则$\boldsymbol{\beta}$的约束型岭估计为：

$$b(k,\boldsymbol{R})=b(k)+\boldsymbol{S}_k^{-1}\boldsymbol{R'}\left(\boldsymbol{RS}_K^{-1}\boldsymbol{R'}\right)^{-1}\left(\boldsymbol{r}-\boldsymbol{R}b(k)\right),$$

其中，$b(k)=\boldsymbol{S}_K^{-1}$,

$\boldsymbol{S}_k=\boldsymbol{X'X}+k\boldsymbol{I}$。

定理 3[4] 在偏序意义下，约束型岭估计b（k，$\boldsymbol{R}$）的协方差一致不大于约束型最小二乘估计b（$\boldsymbol{R}$）的协方差，即对任意$k>0$，式子

$\boldsymbol{D}=cov\left(b(\boldsymbol{R})\right)-cov\left(b(k,\boldsymbol{R})\right)\geq 0$都成立，并且$\boldsymbol{D}$随$k$值在Löwner偏序意义下单调递增。

定理 4[5] 设

$$\Delta=MSEM\left(\boldsymbol{\beta},b(\boldsymbol{R})\right)-MSEM\left(\boldsymbol{\beta},b(k,\boldsymbol{R})\right),$$

则满足$\Delta\geq 0$的充要条件是关于k的式子

$k\left(\boldsymbol{\sigma}^{-2}\boldsymbol{\beta\beta'}-\boldsymbol{UXX'U}\right)\leq 2\boldsymbol{U}$成立，一个充分条件是$k$在下列范围内取值：

$$0<k\leq 2\boldsymbol{\sigma}^2 \Big/ \boldsymbol{\beta'U^+\beta},$$

其中$\boldsymbol{U}=\boldsymbol{I}_p-\boldsymbol{R'}\left(\boldsymbol{RR'}\right)^{-1}\boldsymbol{R}$，$\boldsymbol{U}^+$为$\boldsymbol{U}$的加号逆。

3 结论

由上可知，当$0<k\leq 2\boldsymbol{\sigma}^2 \big/ \boldsymbol{\beta'U^+\beta}$时，在均方误差意义下，约束型岭估计$b$（$k$，$\boldsymbol{R}$）优于约束型最小二乘估计$b$（$\boldsymbol{R}$），特别地，当回归模型不带约束$\boldsymbol{R\beta}=r$时，$\boldsymbol{R}=0$，$\boldsymbol{U}^+=\boldsymbol{U}=\boldsymbol{I}_p$，充分条件变为$0<k\leq 2\boldsymbol{\sigma}^2 \big/ \boldsymbol{\beta'U^+\beta}$，与C. R. Rao的结果一致。

参考文献:

[1] 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.

[2] 张金槐. 线性模型参数估计及其改进[M]. 长沙：国防科技大学出版社，1999.

[3] Rao C R， Toutenburg H. Linear Models[M]. 北京：世界图书出版公司北京公司，1998.

[4] 钟 震. 线性模型中的约束型有偏估计的研究[D]. 重庆：重庆大学数理学院（应用数学），2005.

[5] 史建红. 约束线性回归模型回归系数的条件岭型估计[J]. 山西师范大学学报，2001，15(4)：10－16.

（责任编辑：罗立宇）