2007 年 4 月　　重庆师范大学学报(自然科学版)　　Apr. 2007
第 24 卷 第 2 期　　Journal of Chongqing Normal University(Natural Science Edition)　　Vol. 24 No. 2

非齐次等式约束线性回归模型参数的条件根方估计*

农秀丽[1,2]，刘万荣[1]
(1. 湖南师范大学 数学与计算机科学学院，长沙 410081；2. 南宁师范高等专科学校，广西 龙州 532400)

摘　要：对非齐次等式约束线性回归模型提出回归系数的条件根方估计，证明了在一定条件下，当回归自变量间存在复共线关系时，条件根方估计在均方误差意义下很好改进了约束最小二乘估计；讨论了条件根方估计的容许性；分别给出了在平均散布误差准则下条件根方估计优于约束最小二乘估计的充要条件与充分条件；对确定条件根方估计的参数值讨论了两种方法：根迹法、方差扩大因子法。

关键词：约束线性回归模型；条件根方估计；均方误差；容许性
中图分类号：O212　　**文献标识码**：A　　**文章编号**：1672-6693(2007)02-0024-05

The Conditional Root Square Estimation of Parameter of Restricted Linear Model

NONG Xiu-li[1,2], *LIU Wan-rong*[1]
(1. College of Mathematics and Computer Science, Hunan Normal University, Changsha 410081;
2. Nanning Teachers College, Longzhou Guangxi 532400, China)

Abstract: In this paper, the conditional root square estimation of parameter of restricted linear model is derived, when multicollinearity of explanatory variables exists. It is shown that it has smaller mean squares error than the RLSE, and the admissibility of the conditional root estimation is discussed. Under the MDE matrix comparisons criterion, the necessary and sufficient condition or sufficient condition, under which CRSE is superior to RLSE, is obtained. Two methods (Root Trace, Variance Inflation Factor) are used to evaluate the optimal value.

Key words: restricted linear model; the conditional root square; Mean squares error; admissibility

考虑带非齐次等式约束的线性回归模型

$$\begin{cases} Y = X\beta + e, E(e) = 0, \operatorname{cov}(e) = \sigma^2 I_n \\ R\beta = r \end{cases} \tag{1}$$

其中 Y 为 $n\times 1$ 的观察向量，X 为 $n\times p$ 的设计矩阵，e 为 $n\times 1$ 的随机误差向量，I_n 为 n 阶单位矩阵，$\sigma^2>0$ 为误差方差，R 为 $q\times p$ 的矩阵，r 为 $q\times 1$ 向量，$\beta\in B\triangleq\{\beta: R\beta=r\}$为待估参数向量。本文均假设下式成立

$$\operatorname{rank}(X) = p\text{(列满秩)}, \operatorname{rank}(R) = q\text{(行满秩)} \tag{2}$$

对模型(1)，可以求出使残差平方和 $S(\beta)$ 最小的 β^*，称为 β 的 RLSE(Restricted Least-Squares Estimator)。设 $S(\beta,\lambda) = (Y-X\beta)'(Y-X\beta) - 2\lambda'(R\beta - r)$ [1]，其中 λ 是 q 维向量。$S(\beta,\lambda)$ 分别关于 β,λ 求偏导有

$$\frac{\partial S(\beta,\lambda)}{\partial\beta} = -2X'Y + 2X'X\beta - 2R'\lambda = 0 \tag{3}$$

$$\frac{\partial S(\beta,\lambda)}{\partial\lambda} = -2(R\beta - r) = 0 \tag{4}$$

由(3)、(4)式可以求出 β 的 RLSE，即

$$\beta^* = \hat\beta + S^{-1}R'[RS^{-1}R']^{-1}(r - R\hat\beta) = [S^{-1} - S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}]X'Y + S^{-1}R'[RS^{-1}R']^{-1}r = WX'Y + V \tag{5}$$

---

* 收稿日期：2006-08-14　　修回日期：2006-12-21
资助项目：广西教育厅科研资助项目(No. 200607lx021)

作者简介：农秀丽(1971-)，女，广西隆安人，硕士研究生，研究方向为数理统计。

其中 $\hat{\beta}=(X'X)^{-1}X'Y$ 为 β 的无约束 LSE,且 $\hat{\beta}$ 是 β 的 BLUE(Best Linear Unbiased Estimator),$Var(\hat{\beta})=\sigma^2S^{-1}$,$S=X'X$,$W=S^{-1}-S^{-1}R'[RS^{-1}R']RS^{-1}$,$V=S^{-1}R'[RS^{-1}R']r$,且 $E\beta^*=\beta+S^{-1}R'[RS^{-1}R']^{-1}(r-R\beta)=\beta$,$Var(\beta^*)=\sigma^2S^{-1}-\sigma^2S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}$,$Var(\hat{\beta})-Var(\beta^*)=\sigma^2S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}\geq 0$,这是因为 $\sigma^2>0$,$S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}\geq 0$。

综上可知,在估计类 $CY+Dr$ 中,β^* 是 β 的 BLUE。

在无约束线性模型下,当设计阵呈病态时,夏结来[2-3]提出了用根方估计来改善最小二乘估计(LSE)。根方估计和著名的岭估计虽然都是从同一角度提出的有偏估计,但和岭估计采取的途径不同,在一定条件下它也能很好地改进 LSE,根据 Monte Carlo 方法模拟结果,有时根方估计优于岭估计,有时岭估计优于根方估计。郑昌光[4]讨论在约束线性回归模型中,参数 β 的 RLSE β^* 的均方误差在一定条件下可以变得很大,因此效果也不理想,这就促使人们在 β 的有偏估计类中寻找一类合理的估计去改进 β^*,一种很自然的作法就是把无约束线性模型上的有偏估计推广到有约束的线性模型中,目前岭估计已有多种推广形式[5-6],史建红[7]将根方估计引入约束条件为 $R\beta=0$ 的线性模型。本文给出在约束条件 $R\beta=r$ 下线性模型回归系数 β 的根方估计,证明在一定条件下,当回归自变量间存在复共线关系时,条件根方估计在均方误差意义下很好改进参数的约束最小二乘估计(RLSE),并讨论条件根方估计的容许性。给出根方估计 k 值的选择采用两种方法 根迹法、方差扩大因子法。

1 引理、定义及其性质

引理1[7] 在模型(1)下,$WSW=W$,$WSW'=W$。

证明 $WSW=[S^{-1}-S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}]S[S^{-1}-S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}]=W$,又 $W'=W$,故 $WSW'=W$。 证毕

引理2 在模型(1)下,$W=S^{-1}-S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}$ 为半正定矩阵,且 $\mathrm{rank}(W)=p-q$。

证明 因 $WSW'=W$,又 S 为正定阵,故 W 为半正定阵,下证 W 的秩为 $p-q$。注意到 $W=S^{-\frac{1}{2}}[I_p-S^{-\frac{1}{2}}R'(RS^{-\frac{1}{2}}S^{-\frac{1}{2}}R')^{-1}RS^{-\frac{1}{2}}]S^{-\frac{1}{2}}=S^{-\frac{1}{2}}[I_p-P_{S^{-\frac{1}{2}}R'}]S^{-\frac{1}{2}}$,其中 P_A 表示向 $\mu(A)$ 的正交投影阵。所以,$\mathrm{rank}(W)=\mathrm{rank}[I_P-P_{S^{-\frac{1}{2}}R'}]=\mathrm{tr}(I_P-P_{S^{-\frac{1}{2}}R'})=p-\mathrm{rank}(P_{S^{-\frac{1}{2}}R'})=p-q$。 证毕

引理3 存在 p 阶正交阵 Q,使 $Q'WQ=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_{p-q},0,\cdots,0)\triangleq\Lambda$,$\lambda_i$ 为 W 的非零特征根,且 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_{p-q}>0$。

证明 由引理1结论显然。 证毕

引理4 在模型(1)下,β^* 的均方误差为 $MSE(\beta^*)=\sigma^2\sum_{i=1}^{p-q}\lambda_i$,$\lambda_i$ 为 W 的非零特征根,且 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_{p-q}>0$。

证明
$$MSE(\beta^*)=\mathrm{tr}(\mathrm{cov}(\beta^*))+\|E(\beta^*)-\beta\|^2=\mathrm{tr}(\mathrm{cov}(WX'Y+V))+\|\beta\text{-}\beta\|^2= \tag{6}$$
$$\sigma^2\mathrm{tr}(WSW)=\sigma^2\mathrm{tr}(W)=\sigma^2\sum_{i=1}^{p-q}\lambda_i$$
证毕

引理5 在模型(1)下,令 $\alpha=Q'(\beta-V)=(\alpha_1,\alpha_2,\cdots,\alpha_p)'$,则 $\alpha_{p-q+1}=\alpha_{p-q+2}=\cdots=\alpha_p=0$,$Q$ 由引理4定义,$V=S^{-1}R'[RS^{-1}R']^{-1}r$。

证明 由引理3知,$Q'WQ=Q'[S^{-1}-S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}]Q=\Lambda$,故 $S^{-1}-S^{-1}R'[RS^{-1}R']^{-1}RS^{-1}=Q\Lambda Q'\Rightarrow I-S^{-1}R'[RS^{-1}R']^{-1}R=Q\Lambda Q'S$,从而 $\beta-S^{-1}R'[RS^{-1}R']^{-1}R\beta=Q\Lambda Q'S\beta$。因 $R\beta=r$,所以 $\beta-S^{-1}R'[RS^{-1}R']^{-1}r=Q\Lambda Q'S\beta$,则 $\beta-V=Q\Lambda Q'S\beta$,$Q'(\beta-V)=\Lambda Q'S\beta$。令 $\alpha=Q'(\beta-V)$,则 $\alpha=\Lambda Q'S\beta$,由 Λ 的定义知,$\alpha_{p-q+1}=\alpha_{p-q+2}=\cdots=\alpha_p=0$。 证毕

引理6 假设 $Q'V=(b_1\cdots b_{p-q}b_{p-q+1}\cdots b_p)'$,则有
$$Q'\beta=\alpha+Q'V=(\alpha_1+b_1\cdots\ \alpha_{p-q}+b_{p-q}\ \alpha_{p-q+1}+b_{p-q+1}\cdots\ \alpha_p+b_p)' \tag{7}$$
其中 $\alpha_{p-q+1}=\alpha_{p-q+2}=\cdots=\alpha_p=0$。

定义1 在模型(1)下,称由下式给定的 $\tilde{\beta}^{(k)}$ 为 β 的条件根方估计。
$$\tilde{\beta}^{(k)}=(W^+)^k\beta^*=(W^+)^k(WX'Y+V),0<k<1 \tag{8}$$

其中 $(W^{+})^{k}=Q\mathrm{diag}(\lambda_1^{-k},\cdots,\lambda_{p-q}^{-k},0,\cdots,0)Q'W$、$Q$、$V$ 的定义同上文。

定理 1　根方估计 $\tilde{\beta}^{(k)}$ 是 β 的有偏估计。

证明　因为 $E\tilde{\beta}^{(k)}=E(W^{+})^{k}\beta^{*}=(W^{+})^{k}\beta$,所以 $\tilde{\beta}^{(k)}$ 是 β 的有偏估计。　　证毕

定理 2　在模型(1)下,条件根方估计 $\tilde{\beta}^{(k)}$ 的均方误差为

$$MSE(\tilde{\beta}^{(k)})=\sigma^{2}\sum_{i=1}^{p-q}\lambda_i^{1-2k}+\sum_{i=1}^{p-q}(\alpha_i+b_i)^{2}(\lambda_i^{-k}-1)^{2}\triangleq H(k) \tag{9}$$

其中 α_i,b_i 的定义同上文。

证明　$MSE(\tilde{\beta}^{(k)})=\mathrm{tr}(\mathrm{cov}[(W^{+})^{k}\beta^{*}])+\|(W^{+})^{k}\beta-\beta\|^{2}=\mathrm{tr}((W^{+})^{k}\mathrm{cov}(\beta^{*})(W^{+})^{k})+$
$\beta'[(W^{+})^{k}-I][(W^{+})^{k}-I]\beta=\sigma^{2}\mathrm{tr}((\Lambda^{+})^{k}\Lambda(\Lambda^{+})k)+(Q'\beta)'[(\Lambda^{+})^{k}-I][(\Lambda^{+})^{k}-I]Q'\beta=$
$\sigma^{2}\mathrm{tr}((\Lambda^{+})^{k}\Lambda(\Lambda^{+})^{k})+(\alpha+Q'V)'[(\Lambda^{+})^{k}-I][(\Lambda^{+})^{k}-I](\alpha+Q'V)=$

$$\sigma^{2}\sum_{i=1}^{P-q}\lambda_i^{1-2k}+\sum_{i=1}^{p-q}(\alpha_i+b_i)^{2}(\lambda_i^{-k}-1)^{2}$$

证毕

定理 3　在模型(1)下,当 $\prod_{i=1}^{p-q}\lambda_i>1$ 时,总存在 $k(0<k<1)$,使得

$$MSE(\tilde{\beta}^{(k)})\le MSE(\beta^{*}) \tag{10}$$

证明　记 $h_1(k)\triangleq\sigma^{2}\sum_{i=1}^{p-q}\lambda_i^{1-2k}$,$h_2(k)\triangleq\sum_{i=1}^{p-q}(\alpha_i+b_i)^{2}(\lambda_i^{-k}-1)^{2}$,则 $H(k)=h_1(k)+h_2(k)$。取 $k=0$ 时,有 $H(0)=\sigma^{2}\sum_{i=1}^{p-q}\lambda_i=MSE(\beta^{*})$,$H(k)$ 关于 k 求导得 $H'(k)=h'_1(k)+h'_2(k)$,$h'_1(k)=-2\sigma^{2}\sum_{i=1}^{p-q}\lambda_i^{1-2k}\ln\lambda_i$。当 $0\le k\le\frac{1}{2}$ 时,若 $\lambda_i\ge1$,则 $\ln\lambda_i\ge0$,$\lambda_i^{1-2k}\ge1$;若 $\lambda_i\le1$,则 $\ln\lambda_i\le0$,$\lambda_i^{1-2k}\le1$。故当 $\prod_{i=1}^{p-q}\lambda_i>1$ 时,有 $h'_1(k)=-2\sigma^{2}\sum_{i=1}^{p-q}\lambda_i^{1-2k}\ln\lambda_i\le-2\sigma^{2}\sum_{i=1}^{p-q}\ln\lambda_i=-2\sigma^{2}\ln\prod_{i=1}^{p-q}\lambda_i<0$,可知 $h_1(k)$ 是 $(0,\frac{1}{2})$ 上的单调减函数。

而 $h'_2(k)=-2\sum_{i=1}^{p-q}(\alpha_i+b_i)^{2}(\lambda_i^{-k}-1)\lambda_i^{-k}\ln\lambda_i\ge0$,且 $h'_2(0)=0$,即 $h'_2(k)\ge h'_2(0)$,因为 $h_2(k)$ 是 $(-\infty,\infty)$ 上的连续函数。因此,可知 $h_2(k)$ 是 $(0,\frac{1}{2})$ 上的单调增函数。所以总存在 $k_0(0<k_0\le\frac{1}{2})$,使得当 $0<k<k_0$ 时,$h'_2(k)<2\sigma^{2}\sum_{i=1}^{p-q}\ln\lambda_i$。

综上可知,当 $0<k<k_0$ 时,有 $H'(k)=h'_1(k)+h'_2(k)<0$,故 $H(k)$ 是 $(0,k_0)$ 上的单调减函数,因此有 $H(k)<H(0)$,$k\in(0,k_0)$,即 $MSE(\tilde{\beta}^{(k)})\le MSE(\beta^{*})$,用图形比较大小见图 1。

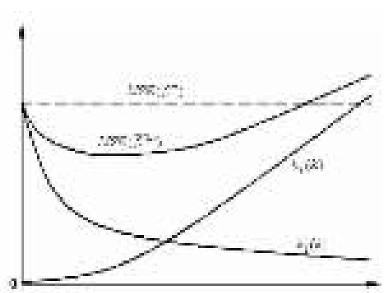

图 1　$MSE(\tilde{\beta}^{(k)})$ 与 $MSE(\beta^{*})$ 的大小比较

定理 3 说明当 W 的特征根很大时,从估计值的均方误差来看,总可以选择一个适当 k 的值使得条件根方估计优于 RLS 估计。

从定理 3 的证明过程知,$MSE(\tilde{\beta}^{(k)})=h_1(k)+h_2(k)$,$h_1(k)=E\|\tilde{\beta}^{(k)}-E\tilde{\beta}^{(k)}\|^{2}$,是 $\tilde{\beta}_i^{(k)}$ 的方差之和,且是 $(0,\frac{1}{2})$ 上的单调减函数;$h_2(k)=E\|E\tilde{\beta}^{(k)}-\beta\|^{2}$,是 $\tilde{\beta}_i^{(k)}$ 的偏差平方和,是 $(0,\frac{1}{2})$ 上的单调增函数。由此可知,根方估计就是用估计值的偏性来换取估计值方差的减小,从而使 $MSE(\tilde{\beta}^{(k)})\le MSE(\beta^{*})$。但是当 W 的非零特征根中有小于 1 的特征根时,条件根方估计将不是约束可容许估计。　　证毕

2　MDE 准则比较 CRSE 与 RLSE

为下文需要,先给出以下定义。

定义 2[8]　设 $\hat{\beta}$ 是参数 β 的估计,则称 $M(\hat{\beta},\beta)=E(\hat{\beta}-\beta)(\hat{\beta}-\beta)'$ 为估计 $\hat{\beta}$ 的平均散布误差。由定义 1 有

$$M(\hat{\beta},B)=E(\hat{\beta}-\beta)(\hat{\beta}-\beta)'=Var(\hat{\beta})+Bias(\hat{\beta},\beta)(Bias(\hat{\beta},\beta))'。 \tag{11}$$

定义 3[8]　设 $\hat{\beta}_1,\hat{\beta}_2$ 是参数 β 的两个估计,若它们的平均散布误差矩阵之差是非负定的,即 $\Delta(\hat{\beta}_1,\hat{\beta}_2)=$

$M(\hat{\beta}_1,\beta)-M(\hat{\beta}_2,\beta)\geq 0$,则称 $\hat{\beta}_2$ 是在 MDE-I 准则下优于 $\hat{\beta}_1$。

定义4[8] 设 $\hat{\beta}_1,\hat{\beta}_2$ 是参数 β 的两个估计,若

$$E(\hat{\beta}_1-\beta)'(\hat{\beta}_1-\beta)-E(\hat{\beta}_2-\beta)'(\hat{\beta}_2-\beta)=\mathrm{tr}\{\Delta(\hat{\beta}_1,\hat{\beta}_2)\}\geq 0 \tag{12}$$

则称 $\hat{\beta}_2$ 是在 MDE-Ⅱ准则下优于 $\hat{\beta}_1$。

定理4 条件根方估计 $\tilde{\beta}^{(k)}$ 在 MDE-I 准则下优于约束最小二乘估计 β^* 的充分必要条件是

$$\sigma^{-2}\beta'Q(\Lambda^k-I)(\Lambda^{2k+1}-\Lambda)^{-1}(\Lambda^k-I)Q'\beta\leq 1 \tag{13}$$

证明 $\beta^*=WX'Y+V,\tilde{\beta}^{(k)}=(W^+)^k\beta^*=(W^+)^k(WX'Y+V),0<k<1$,其中 $W=S^{-1}-S^{-1}R'[RS^{-1}R']^{-1}RS^{-1},V=S^{-1}R'[RS^{-1}R']^{-1}r$,因为 $E\beta^*=\beta,Var(\beta^*)=\sigma^2W,E\tilde{\beta}^{(k)}=(W^+)^k\beta,Var(\tilde{\beta}^{(k)})=\sigma^2(W^+)^kW(W^+)^k$,所以 $M(\beta^*,\beta)=Var(\beta^*)+Bias(\beta^*,\beta)(Bias(\beta^*,\beta))'=\sigma^2W$ (14)

$$\begin{aligned}M(\tilde{\beta}^{(k)},\beta)&=Var(\tilde{\beta}^{(k)})+Bias(\tilde{\beta}^{(k)},\beta)(Bias(\tilde{\beta}^{(k)},\beta))'=\sigma^2(W^+)^kW(W^+)^k+((W^+)^k\beta-\beta)((W^+)^k\beta-\beta)'=\\&\sigma^2(W^+)^kW(W^+)^k+((W^+)^k-I)\beta\beta'((W^+)^k-I)=\sigma^2(W^+)^kW(W^+)^k+(W^+)^k\beta\beta'(W^+)^k-\\&(W^+)^k\beta\beta'(W^+)^{-k}-(W^+)^k(W^+)^k(W^+)^{-k}\beta\beta'(W^+)^k+(W^+)^k(W^+)^{-k}\beta\beta'(W^+)^{-k}(W^+)^k\end{aligned} \tag{15}$$

因此

$$\begin{aligned}\Delta(\beta^*,\tilde{\beta}^{(k)})&=M(\beta^*,\beta)-M(\tilde{\beta}^{(k)},\beta)=\\&\sigma^2(W^+)^k(W^+)^{-k}W(W^+)^{-k}(W^+)^k-\sigma^2(W^+)^kW(W^+)^k-(W^+)^k\beta\beta'(W^+)^k+\\&(W^+)^k\beta\beta'(W^+)^{-k}(W^+)^k+(W^+)^k(W^+)^{-k}\beta\beta'(W^+)^k-(W^+)^k(W^+)^{-k}\beta\beta'(W^+)^{-k}(W^+)^k=\\&(W^+)^k(\sigma^2(W^+)^{-k}W(W^+)^{-k}-\sigma^2W-\beta\beta'+\beta\beta'(W^+)^{-k}+(W^+)^{-k}\beta\beta'-(W^+)^{-k}\beta\beta'(W^+)^{-k}(W^+)^k=\\&(W^+)^k(\sigma^2((W^+)^{-k}W(W^+)^{-k}-W)+\beta\beta'((W^+)^{-k}-I)+(W^+)^{-k}\beta\beta'(I-(W^+)^{-k}(W^+)^k=\\&(W^+)^k(\sigma^2((W^+)^{-k}W(W^+)^{-k}-W)+\beta\beta'((W^+)^{-k}-I)-(W^+)^{-k}\beta\beta'((W^+)^{-k}-I)(W^+)^k=\\&(W^+)^k\sigma^2((W^+)^{-k}W(W^+)^{-k}-W)-((W^+)^{-k}-I)\beta\beta'((W^+)^{-k}-I))(W^+)^k\end{aligned} \tag{16}$$

因为 $(W^+)^k\geq 0$,要使 $\Delta(\beta^*,\tilde{\beta}^{(k)})\geq 0$,只要使

$$\sigma^2((W^+)^{-k}W(W^+)^{-k}-W)-((W^+)^{-k}-I)\beta\beta'((W^+)^{-k}-I))\geq 0 \tag{17}$$

而

$$\begin{aligned}(W^+)^{-k}W(W^+)^{-k}-W&=\\&Q\mathrm{diag}(\lambda_1^k,\cdots,\lambda_{p-q}^k,0,\cdots,0)Q'Q\mathrm{diag}(\lambda_1,\cdots,\lambda_{p-q},0,\cdots,0)Q'Q\mathrm{diag}(\lambda_1^k,\cdots,\lambda_{p-q}^k,0,\cdots,0)Q'-\\&Q\mathrm{diag}(\lambda_1,\cdots,\lambda_{p-q},0,\cdots,0)Q'=Q\mathrm{diag}(\lambda_1^{2k+1},\cdots,\lambda_{p-q}^{2k+1},0,\cdots,0)Q'-Q\mathrm{diag}(\lambda_1,\cdots,\lambda_{p-q},0,\cdots,0)Q'=\\&Q\mathrm{diag}(\lambda_1^{2k+1}-\lambda_1,\cdots,\lambda_{p-q}^{2k+1}-\lambda_{p-q},0,\cdots,0)Q'=Q(\Lambda^{2k+1}-\Lambda)Q'\end{aligned} \tag{18}$$

$$((W^+)^{-k}-I)\beta\beta'((W^+)^{-k}-I))=Q(\Lambda^k-I)Q'\beta\beta'Q(\Lambda^k-I)Q'$$

因此要使

$$\begin{aligned}&(\sigma^2(W^+)^{-k}W(W^+)^{-k}-W)-((W^+)^{-k}-I)\beta\beta'((W^+)^{-k}-I))=\sigma^2Q(\Lambda^{2k+1}-\Lambda)Q'-Q(\Lambda^k-\\&I)Q'\beta\beta'Q(\Lambda^k-I)Q'=Q(\sigma^2(\Lambda^{2k+1}-\Lambda)-(\Lambda^k-I)Q'\beta\beta'Q(\Lambda^k-I)Q'\geq 0\end{aligned}$$

只要 $(\sigma^2(\Lambda^{2k+1}-\Lambda)-(\Lambda^k-1)Q'\beta\beta'Q(\Lambda^k-I))\geq 0$。从而有

$$\sigma^{-2}\beta'Q(\Lambda^k-I)(\Lambda^{2k+1}-\Lambda)^{-1}(\Lambda^k-I)Q'\beta\leq 1。$$

证毕

定理5 条件根方估计 $\tilde{\beta}^{(k)}$ 在 MDE-Ⅱ准则下优于约束最小二乘估计 β^* 的充分条件是

$$\sigma^{-2}\beta'Q(\Lambda^k-I)(\Lambda^+)^{2k}(\Lambda^k-I)Q'\beta\leq\mathrm{tr}(\Lambda-\Lambda^{1-2k}) \tag{19}$$

证明 因为

$$\begin{aligned}&\mathrm{tr}\{\Delta(\beta^*,\tilde{\beta}^{(k)})\}=\mathrm{tr}\{(W^+)^k(\sigma^2((W^+)^{-k}W(W^+)^{-k}-W)-((W^+)^{-k}-\\&I)\beta\beta'((W^+)^{-k}-I))(W^+)^k\}=\sigma^2\mathrm{tr}\{(W^+)^kQ(\Lambda^{2k+1}-\Lambda)Q'(W^+)^k\}-\mathrm{tr}\{(W^+)^kQ(\Lambda^k-I)Q'\beta\beta'Q(\Lambda^k-I)Q'(W^+)^k\}=\\&\sigma^2\mathrm{tr}\{Q(\Lambda^+)^kQ'Q(\Lambda^{2k+1}-\Lambda)Q'Q(\Lambda^+)^kQ'\}-\mathrm{tr}\{Q(\Lambda^+)^kQ'Q(\Lambda^k-I)Q'\beta\beta'Q(\Lambda^k-I)Q'Q(\Lambda^+)^kQ'\}=\\&\sigma^2\mathrm{tr}(\Lambda-\Lambda^{1-2k})-\beta'Q(\Lambda^k-I)(\Lambda^+)^{2k}(\Lambda^k-I)Q'\beta=\\&\sigma^2\mathrm{tr}(\Lambda-\Lambda^{1-2k})-\beta'Q(\Lambda^k-I)(\Lambda^+)^{2k}(\Lambda^k-I)Q'\beta\end{aligned} \tag{20}$$

要使 $\sigma^2\mathrm{tr}(\Lambda-\Lambda^{1-2k})-\beta'Q(\Lambda^k-I)(\Lambda^+)^{2k}(\Lambda^k-I)Q'\beta\geq 0$,只要 $\sigma^{-2}\beta'Q(\Lambda^k-I)(\Lambda^+)^{2k}(\Lambda^k-I)Q'\beta\leq\mathrm{tr}(\Lambda-\Lambda^{1-2k})$

证毕

3 k 值的确定

定理3指出选择适当的 k 值能使根方估计 $\tilde{\beta}^{(k)}$ 的均方误差。小于 RLSE β^* 的均方误差。以下给出两种选择根方估计 $\tilde{\beta}^{(k)}$ 中 k 值的方法。

1）根迹法[9-10]（Root Trace）。$\tilde{\beta}^{(k)}$ 是 k 的函数，$\tilde{\beta}_i^{(k)}$ 随 k 变化的轨迹称之为根迹。由 $\tilde{\beta}^{(k)}=(W^+)^k\beta^*=(W^+)^k(WX'Y+V)$，$0<k<1$ 知，当 k 趋向于 0 时，$\tilde{\beta}^{(k)}\to\beta^*$；当 $k\to1$ 时，$\tilde{\beta}^{(k)}\to X'Y+W^+V$。当 k 值大于某个正常数时，$\tilde{\beta}_i^{(k)}$ 就逐渐稳下来，在坐标纸上绘出根迹图后，按下列原则来确定 k 值，选取使根迹达到稳定的最小 k 值，k 值一旦确定下来，就可按 $\tilde{\beta}^{(k)}=(W^+)^k(WX'Y+V)$ 求出回归系数的估计值。如此给出的回归系数的估计值与传统的残差分析法求出的回归系数相比较，其思想是截然不同的。由根迹法确定的 k 值，一般能克服 RLSE 的明显不足之处，如校正回归系数 RLSE 中的一些‘不正确的符号”。但根迹法比较直观，难以令人信服。

根迹图与岭迹图（Ridge Trace）相比有着明显的优点，根迹图在[0，1]上有两种极端的情况，因此只须在[0，1]上考察 $\tilde{\beta}_i^{(k)}$ 的变化趋势，而岭迹图最好是在[0，∞]上考察 $\tilde{\beta}_i^{(k)}$ 的变化趋势。

2）方差扩大因子法[9-10]。根据 Marguardt Snee 方差扩大因子（Variance Inflation Factor）的定义法，可以定义根方估计 $\tilde{\beta}^{(k)}$ 的方差扩大因子。

定义 5　回归系数 β 的估计值 $\tilde{\beta}^{(k)}$ 的方差扩大因子为矩阵 $\frac{\operatorname{cov}(\tilde{\beta}^{(k)})}{\sigma^2}$ 对角线元素中的最大者，记为 $VIF(\tilde{\beta}^{(k)})$，即 $VIF(\tilde{\beta}^{(k)})=\max\left(\frac{Var(\tilde{\beta}_i^{(k)})}{\sigma^2}\right)$。

证 明　因 $\operatorname{cov}(\tilde{\beta}^{(k)})=\operatorname{cov}((W^+)^k\beta^*)=\sigma^2(W^+)^kW(W^+)^k$，故 $=\frac{\operatorname{cov}(\tilde{\beta}^{(k)})}{\sigma^2}=(W^+)^kW(W^+)^k$。$VIF(\tilde{\beta}^{(k)})$ 就是矩阵 $(W^+)^kW(W^+)^k$ 对角线中的最大者。当 $k=0$ 时，$VIF(\tilde{\beta}^{(k)})=VIF(\tilde{\beta}^*)$，随着 k 值的增大，有 $VIF(\tilde{\beta}^{(k)})\le VIF(\tilde{\beta}^*)$。这是因为 $\operatorname{tr}\left(\frac{\operatorname{cov}(\tilde{\beta}^{(k)})}{\sigma^2}\right)=\operatorname{tr}\left(\frac{(W^+)^kW(W^+)k}{\sigma^2}\right)=\sum_{i=1}^{p-q}\lambda_i^{1-2k}$ 是 $(0,\frac{1}{2})$ 的单调减函数。因此，当 $k=0$ 时，$\operatorname{tr}\left(\frac{\operatorname{cov}(\tilde{\beta}^{(k)})}{\sigma^2}\right)=\sum_{i=1}^{p-q}\lambda_i=\operatorname{tr}(W)=\operatorname{tr}\left(\frac{\operatorname{cov}(\beta^*)}{\sigma^2}\right)$；当 $k=\frac{1}{2}$ 时，$\operatorname{tr}\left(\frac{\operatorname{cov}(\tilde{\beta}^{(k)})}{\sigma^2}\right)=p-q$；当 $0\le k\le\frac{1}{2}$ 时，$p-q\le\operatorname{tr}\left(\frac{\operatorname{cov}(\tilde{\beta}^{(k)})}{\sigma^2}\right)\le\operatorname{tr}\left(\frac{\operatorname{cov}(\beta^*)}{\sigma^2}\right)=\operatorname{tr}(W)$，从上式可以看出，随着 $k\in(0,\frac{1}{2})$ 增大，有 $VIF(\tilde{\beta}^{(k)})\le VIF(\tilde{\beta}^*)$。　**证毕**

夏结来[2]针对无约束的根方估计提出了一个选择 k 值的原则：选择最小 k 的值使根方估计的方差扩大因子不大于 10。在等式约束下，k 值的选取也采用类似的原则，具体选取 k 值的原则是：取使 $(W^+)^kW(W^+)^k$ 的对角线元素最大者不大于 10 的最小的 k 值。

由于选择 k 值的标准不同，用以上两种方法选取的 k 值很可能是不等的。

参考文献：

[1] 王松桂，陈敏，陈丽萍. 线性统计模型[M]. 北京：高等教育出版社，1999.

[2] 夏结来，郭祖超，胡琳. 回归系数根方有偏估计及其应用[J]. 数理统计与应用概率，1988，3（1）：21-30.

[3] 夏结来，郭祖超，胡琳. 回归系数的广义根方估计及其模拟[J]. 应用数学，1994，7（2）：187-192.

[4] 郑昌光. 约束条件下的线性估计[J]. 应用概率统计，1986（1）：5-12.

[5] 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.

[6] 田保光，王建新，常桂娟. 广义岭型主成分估计在降维估计类中的方差最优性质[J]. 四川师范大学学报（自然科学版），2005，28（4）：426-428.

[7] 史建红. 约束线性模型下回归系数的条件根方估计[J]. 沈阳师范学院学报，1999，15（4）：13-19.

[8] RAO C R. Linear Models[M]. New York：Springer-Verlag，1998.

[9] 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.

[10] 杨文礼. 线性模型引论[M]. 北京：北京师范大学出版社，1998.

（责任编辑　游中胜）