应用概率统计 第二十四卷
第三期 2008年6月

Chinese Journal of Applied Probability
and Statistics Vol.24 No.3 Jun. 2008

线性混合模型参数的部分岭型谱分解估计

杨 虎 黎雅莲

(重庆大学数理学院统计与精算科学系, 重庆, 400030)

摘 要

对于随机效应部分为一般平衡多项分类的线性模型, 将王松桂等[1]提出的一种称之为谱分解估计(SDE)的参数估计新方法推广到具有病态设计阵的线性模型, 提出了部分岭型谱分解估计, 通过类似于主成分估计的降维模型变换, 可以很方便的研究它的抗干扰性和其它重要性质. 本文的结果可以很方便的应用于Panel模型.

关键词: 部分岭型谱分解估计, 线性混合模型, 可容许性.

学科分类号: O212.4.

§1. 引 言

线性混合效应模型由于在生物, 医学, 经济, 金融, 环境科学, 抽样调查及工程技术等领域的广泛应用, 近年来受到越来越广泛的关注[2, 3], 在许多方面获得了重要进展[4–7]. 对于随机效应部分为一般多项分类平衡数据模型的线性混合模型, 王松桂等[1]提出了一种新的估计方法, 称为谱分解估计方法, 此方法的突出特点是同时考虑方差分量和固定效应的估计且有显式解, 有利于进行假设检验、区间估计以及模型诊断等一系列更加深入的统计推断.

考虑一般的线性混合模型[1]

$$y = X\beta + U_1\xi_1 + \cdots + U_k\xi_k = X\beta + u. \tag{1.1}$$

其中y为$n\times 1$观测向量, X为$n\times p$设计矩阵, β为固定效应, U_i为$n\times q_i$的已知设计阵, $\xi_i\sim(0,\sigma_i^2 I)$, $i=1,2,\cdots,k$, $\mathsf{Cov}\,(\xi_i,\xi_j)=0$ $(i\neq j)$, $\theta=(\sigma_1^2,\cdots,\sigma_k^2)'$称为方差分量, 于是$V(\theta)=\mathsf{Cov}\,(y)=\mathsf{Cov}\,(u)=\sum\limits_{i=1}^{k}\sigma_i^2 U_i U_i'$.

若对$V(\theta)$进行谱分解: $V(\theta)=\sum\limits_{i=1}^{t}\lambda_i M_i$, 其中$\lambda_i$, M_i $(i=1,2,\cdots,t)$为$V(\theta)$的互不相等的非零特征根及其对应的特征向量矩阵, 且M_i为奇异的对称幂等阵, 当$i\neq j$, $M_iM_j=0$. 若用M_i左乘模型(1.1), 则得到

$$y^{(i)} = X_i\beta + \varepsilon_i, \qquad i=1,2,\cdots,t. \tag{1.2}$$

本文2005年12月14日收到, 2007年1月15日收到修改稿.

其中$y^{(i)}=M_iy$, $X_i=M_iX$, $\varepsilon_i=M_iu$, 易知$\varepsilon_i\sim(0,\lambda_iM_i)$.

文献[1]提出了可同时估计固定效应β和方差分量θ的谱分解估计. 其中固定效应β的谱分解估计有t个, 特别地, 针对每个特征值λ_i, 对应固定效应的谱分解估计表达式为

$$\widetilde{\beta}^i_{\mathrm{SDE}}=(X'M_iX)^-X'M_iy,\qquad i=1,2,\cdots,t,$$

且$c'\widetilde{\beta}^i_{\mathrm{SDE}}$是固定效应的最佳线性无偏估计(The best linear unbiased estimator简记为BLUE). 而方差分量θ的谱分解估计可以很容易的通过λ_i的估计

$$\widehat{\lambda}_i=(y_i-X\widetilde{\beta}^i_{\mathrm{SDE}})'M_i(y_i-X\widetilde{\beta}^i_{\mathrm{SDE}})/f,\qquad i=1,2,\cdots,t$$

得到, 其中$f=\mathrm{rank}(M_i)-\mathrm{rank}(M_iX)$. 对于固定效应, 显然$\mathsf{E}(c'\widetilde{\beta}^i_{\mathrm{SDE}})=c'\beta$, 而且因为$X_i$的列向量空间包含在$M_i$的列向量空间, 即$\mathfrak{M}(X_i)\subset\mathfrak{M}(M_i)$, 根据最小二乘统一理论[8], 对任一可估函数$c'\beta$, 它的BLUE $c'\beta^*$与广义逆的选择无关, 因此可选Moore-Penrose广义逆, 仅仅考虑

$$\widetilde{\beta}^i_{\mathrm{SDE}}=(X'M_iX)^+X'M_iy,\qquad i=1,2,\cdots,t \tag{1.3}$$

即可, 易见

$$\begin{aligned}\mathsf{Cov}\,(\widetilde{\beta}^i_{\mathrm{SDE}})&=\mathsf{Cov}\,((X'M_iX)^+X'M_iy)=(X'M_iX)^+X'M_i\mathsf{Cov}\,(y^{(i)})M_iX(X'M_iX)^+\\&=(X'M_iX)^+X'M_i\cdot\lambda_iM_i\cdot M_iX(X'M_iX)^+=\lambda_i(X'M_iX)^+.\end{aligned}$$

若令$\delta_{i1}\geqslant\delta_{i2}\geqslant\cdots\geqslant\delta_{ir_i}>0$为$X'M_iX$的非零特征值, $r_i<p$, $\phi_{i1},\phi_{i2},\cdots,\phi_{ir_i}$为$(X'M_iX)$其对应的标准化特征向量, 并令$Q_i=(\phi_{i1},\phi_{i2},\cdots,\phi_{ir_i})$, 则$\delta_{i1}^{-1},\delta_{i2}^{-1},\cdots,\delta_{ir_i}^{-1}$为$(X'M_iX)^+$的非零特征值, 且有

$$X'M_iX=Q_i\mathrm{diag}(\delta_{i1},\delta_{i1},\cdots,\delta_{ir_i})Q_i'=Q_i\Lambda_iQ_i',$$

所以$(X'M_iX)^+=Q_i\mathrm{diag}(\delta_{i1}^{-1},\delta_{i2}^{-1},\cdots,\delta_{ir_i}^{-1})Q_i'=Q_i\Lambda_i^{-1}Q_i'$, 从而

$$\mathrm{tr}(X'M_iX)^+=\mathrm{tr}(Q_i\mathrm{diag}(\delta_{i1}^{-1},\delta_{i2}^{-1},\cdots,\delta_{ir_i}^{-1})Q_i')=\sum_{j=1}^{r_i}\delta_{ij}^{-1}.$$

于是得到$\widetilde{\beta}^i_{\mathrm{SDE}}$的均方误差

$$\mathrm{MSE}(\widetilde{\beta}^i_{\mathrm{SDE}})=\lambda_i\mathrm{tr}((X'M_iX)^+)=\lambda_i\sum_{j=1}^{r_i}\delta_{ij}^{-1},\qquad i=1,2,\cdots,t. \tag{1.4}$$

显然, 只要当$X'M_iX$存在某个接近于0的特征值δ_{ij}时, 谱分解估计$\widetilde{\beta}^i_{\mathrm{SDE}}$的均方误差就会变得异常大, 此时谱分解估计不再是一个优良估计. 为此, 借助于岭估计[7]的思想, 本文提出了部分岭型谱分解估计$\widetilde{\beta}^i_{\mathrm{SDE}}(k_i)$, 通过类似于主成分的降维模型变换, 可以很方便的研究了它的抗干扰性和其它重要性质.

§2.　部分岭型谱分解估计

考虑线性混合模型(1.1)中固定效应β的如下形式的有偏估计

$$\widetilde{\beta}^i_{\rm SDE}(k_i) = (X'M_iX + k_iQ_iQ_i')^+X'M_iy, \qquad i = 1, 2, \cdots, t. \tag{2.1}$$

我们称之为部分岭型谱分解估计. 因为

$$\|\widetilde{\beta}^i_{\rm SDE}(k_i)\|^2 = \|Q_i(\Lambda_i + k_iI)^{-1}\Lambda_iQ_i'\widetilde{\beta}^i_{\rm SDE}\|^2 < \|\widetilde{\beta}^i_{\rm SDE}\|^2,$$

因此部分岭型谱分解估计也是一种压缩型有偏估计. 由于$X'M_iX$是奇异的, 根据(1.4)式, 我们仅对非零特征值部分采用岭参数修正, 而不是简单地按岭估计的思路采用$X'M_iX + k_iI$对所有特征值都增加岭参数, 根据主成分估计的设计原理[7], 其合理性是显然的, 我们称之为部分岭参数方法.

和主成分估计面临的问题一样, 直接根据模型(1.2)讨论估计(2.1)无法回避较高的维数带来的奇异问题的均方误差计算, 为此我们作变换$\beta = Q_i\alpha$, $X_i = Z_iQ_i'$, 这实际上是一个类似于主成分的降维模型变换[8], 将参数空间的维数由p维降到了r_i维, 这个变换将模型(1.2)变为如下模型

$$y^{(i)} = Z_i\alpha + \varepsilon_i, \qquad i = 1, 2, \cdots, t, \quad \varepsilon_i \sim (0, \lambda_iM_i). \tag{2.2}$$

在此模型下, 根据最小二乘统一理论

$$\begin{aligned}\alpha^i_{\rm SDE} &= (Z_i'\lambda_i^{-1}M_i^-Z_i)^+Z_i'\lambda_i^{-1}M_i^-y^{(i)} = (Q_i'X_i'M_i^-X_iQ_i)^+Q_i'X_i'M_i^-y^{(i)}\\ &= Q_i'(X'M_iX)^+X'M_iy = Q_i'\widetilde{\beta}^i_{\rm SDE} = \Lambda_i^{-1}Z_i'y^{(i)}.\end{aligned}$$

相应的可以得到简化后的部分岭型谱分解估计

$$\begin{aligned}\widetilde{\alpha}^i_{\rm SDE}(k_i) &= Q_i'\widetilde{\beta}^i_{\rm SDE}(k_i) = Q_i'(X'M_iX + k_iQ_iQ_i')^+X'M_iy\\ &= Q_i'(X_i'X_i + k_iQ_iQ_i')^+X_i'y^{(i)} = (\Lambda_i + k_iI)^{-1}Z_i'y^{(i)}, \qquad i = 1, 2, \cdots, t.\end{aligned}$$

下面我们只需研究岭型谱分解估计$\widetilde{\alpha}^i_{\rm SDE}(k_i)$的性质.

定理 2.1　对线性模型(2.2), 当$0 < k_i < 2\lambda_i/(\alpha'\alpha)$时,

$$\text{MSEM}(\widetilde{\alpha}^i_{\rm SDE}(k_i)) < \text{MSEM}(\widetilde{\alpha}^i_{\rm SDE}), \qquad i = 1, 2, \cdots, t,$$

其中$\text{MSEM}(\widetilde{\alpha}^i_{\rm SDE}) = \mathsf{E}(\widetilde{\alpha}^i_{\rm SDE} - \alpha)(\widetilde{\alpha}^i_{\rm SDE} - \alpha)'$称为均方误差矩阵.

证明:　显然若记$A_i = (\Lambda_i + k_iI)^{-1}\Lambda_i$, 则$\widetilde{\alpha}^i_{\rm SDE}(k_i) = A_i\widetilde{\alpha}^i_{\rm SDE}$, 要使

$$\begin{aligned}&\text{MSEM}(\widetilde{\alpha}^i_{\rm SDE}) - \text{MSEM}(\widetilde{\alpha}^i_{\rm SDE}(k_i))\\ =\ &\lambda_i\Lambda_i^{-1} - \lambda_iA\Lambda_i^{-1}A' - (A_i - I)\alpha\alpha'(A_i - I)'\\ =\ &\lambda_i\Lambda_i^{-1} - \lambda_i(\Lambda_i + k_iI)^{-1}\Lambda_i(\Lambda_i + k_iI)^{-1} - (-k_i(\Lambda_i + k_iI)^{-1})\alpha\alpha'(-k_i(\Lambda_i + k_iI)^{-1})'\\ =\ &(\Lambda_i + k_iI)^{-1}[\lambda_i(2k_iI + k_i^2\Lambda_i^{-1}) - k_i^2\alpha\alpha'](\Lambda + k_iI)^{-1} > 0,\end{aligned}$$

只需$\lambda_i(2k_iI+k_i^2\Lambda_i^{-1})-k_i^2\alpha\alpha'>0$. 上式成立的一个充分条件为$(2\lambda_i/k_i)I-\alpha\alpha'>0$. 利用代数事实: 设$A>0$, $d>0$, 则$dA-xx'>0\Leftrightarrow x'A^{-1}x<d$, 其中$A$为矩阵, d为实数, x为向量, 于是(2.1)等价于$0<k_i<2\lambda_i/(\alpha'\alpha)$, 由此定理获证. □

推论 2.1 当$0<k_i<2\lambda_i/(\alpha'\alpha)$时, 对一切$D\geq 0$和每一个$i$

$$\mathrm{GMSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))<\mathrm{GMSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i).$$

这里$\mathrm{GMSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i)=\mathsf{E}(\widetilde{\alpha}_{\mathrm{SDE}}^i-\alpha)'D(\widetilde{\alpha}_{\mathrm{SDE}}^i-\alpha)$称为广义均方误差. 特别取$D=I$, 则$\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))<\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i)$, $i=1,2,\cdots,t$.

定理 2.2 在对每一个i, $i=1,2,\cdots,t$, 在椭球$\alpha'C_i\alpha\leq\lambda_i$内,

$$\mathrm{GMSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))<\mathrm{GMSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i).$$

其中

$$C_i=\mathrm{diag}\Big(\frac{\delta_{i1}k_i^2}{2k_i\delta_{i1}+k_i^2},\frac{\delta_{i2}k_i^2}{2k_i\delta_{i2}+k_i^2},\cdots,\frac{\delta_{ir_i}k_i^2}{2k_i\delta_{ir_i}+k_i^2}\Big). \tag{2.3}$$

证明: 因为

$$\begin{aligned}&\mathrm{MSEM}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))=A_i\mathsf{Cov}\,(\widetilde{\alpha}_{\mathrm{SDE}}^i)A_i'+(A_i-I)aa'(A_i-I)',\\&\mathrm{MSEM}(\widetilde{\alpha}_{\mathrm{SDE}}^i)=\lambda_i\Lambda_i^{-1},\end{aligned}$$

故

$$\mathrm{MSEM}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))-\mathrm{MSEM}(\widetilde{\alpha}_{\mathrm{SDE}}^i)=\lambda_i\Delta_{i1}+\Delta_{i2}\alpha\alpha'\Delta_{i2}\leq 0. \tag{2.4}$$

其中,

$$\begin{aligned}&\Delta_{i1}=\mathrm{diag}\Big(\frac{\delta_{i1}}{(\delta_{i1}+k_i)^2}-\frac{1}{\delta_{i1}},\frac{\delta_{i2}}{(\delta_{i2}+k_i)^2}-\frac{1}{\delta_{i2}},\cdots,\frac{\delta_{ir_i}}{(\delta_{ir_i}+k_i)^2}-\frac{1}{\delta_{ir_i}}\Big),\\&\Delta_{i2}=\mathrm{diag}\Big(\frac{k_i}{\delta_{i1}+k_i},\frac{k_i}{\delta_{i2}+k_i},\cdots,\frac{k_i}{\delta_{ir_i}+k_i}\Big).\end{aligned}$$

要使$\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)$在均方误差矩阵的意义下优于$\widetilde{\alpha}_{\mathrm{SDE}}^i$, (2.4)必须为半负定方阵, 即

$$\lambda_i\Delta_{i1}+\Delta_{i2}\alpha\alpha'\Delta_{i2}\leq 0.$$

这只需要$\lambda_i\mu_i'\Delta_{i1}\mu_i\leq-\mu_i'\Delta_{i2}\alpha\alpha'\Delta_{i2}\mu_i=-(\mu_i'\Delta_{i2}\alpha)^2$对一切$\mu_i$成立, 由于$\Delta_{i1}<0$, 根据Cauchy-Schwarz不等式, 可知

$$-\alpha'\Delta_{i2}\Delta_{i1}^{-1}\Delta_{i2}\alpha\geq-\frac{(\mu_i'\Delta_{i2}\alpha)^2}{\mu_i'\Delta_{i1}\mu_i}.$$

从而得到$\lambda_i\geq-(\mu_i'\Delta_{i2}\alpha)^2/(\mu_i'\Delta_{i1}\mu_i)$的一个充分条件为$\lambda_i\geq\alpha'C_i\alpha$. 这里$C_i$由(2.3)所定义, 注意到

$$\mathrm{GMSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))<\mathrm{GMSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i)\Leftrightarrow\mathrm{MSEM}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))<\mathrm{MSEM}(\widetilde{\alpha}_{\mathrm{SDE}}^i),$$

故在椭球$\alpha' C_i \alpha \leq \lambda_i$内

$$\mathrm{GMSE}(\widetilde{\alpha}^i_{\mathrm{SDE}}(k_i)) < \mathrm{GMSE}(\widetilde{\alpha}^i_{\mathrm{SDE}}). \tag{2.5}$$

□

定理 2.3 在线性估计类中, $\widetilde{\alpha}^i_{\mathrm{SDE}}(k_i)$均为$\alpha$的可容许估计.

证明: 因为$\widetilde{\alpha}^i_{\mathrm{SDE}}(k_i) = A_i \widetilde{\alpha}^i_{\mathrm{SDE}}$, 这里

$$A_i = (\Lambda_i + k_i I)^{-1}\Lambda_i = \mathrm{diag}\Big(\frac{\delta_{i1}}{\delta_{i1}+k_i}, \frac{\delta_{i2}}{\delta_{i2}+k_i}, \cdots, \frac{\delta_{ir_i}}{\delta_{ir_i}+k_i}\Big). \tag{2.6}$$

而

$$A_i\Lambda_i^{-1}A_i = \mathrm{diag}\Big(\frac{\delta_{i1}}{(\delta_{i1}+k_i)^2}, \frac{\delta_{i2}}{(\delta_{i2}+k_i)^2}, \cdots, \frac{\delta_{ir_i}}{(\delta_{ir_i}+k_i)^2}\Big),$$
$$A_i\Lambda_i^{-1} = \mathrm{diag}\Big(\frac{1}{\delta_{i1}+k_i}, \frac{1}{\delta_{i2}+k_i}, \cdots, \frac{1}{\delta_{ir_i}+k_i}\Big).$$

由于$\delta_{i1}/(\delta_{i1}+k_i)^2 < 1/(\delta_{i1}+k_i)$, $i=1,2,\cdots,p$, 故$A_i\Lambda_i^{-1}A_i < A_i\Lambda_i^{-1}$, 根据文献[8]的推论4.7.3知$\widetilde{\alpha}^i_{\mathrm{SDE}}(k_i)$为$\alpha$的可容许估计. □

§3. 部分岭型谱分解估计抗干扰性的改进

引理 3.1 对线性方程组$BX=b$, 用$\tau(B)=\|B\|\|B^{-1}\|$表示矩阵B的条件数, 则当$1-\tau(B)\|\Delta B\|\|B\|^{-1} < 1$时,

$$\frac{\|\Delta X\|}{\|X\|} \leq \frac{\tau(B)}{1-\tau(B)\|\Delta B\|\|B\|^{-1}}\Big(\frac{\|\Delta B\|}{\|B\|} + \frac{\|\Delta b\|}{\|b\|}\Big). \tag{3.1}$$

$\|\Delta X\|/\|X\|$, $\|\Delta B\|/\|B\|$, $\|\Delta b\|/\|b\|$分别表示X, B, b的相对误差. 此引理的证明参见文献[9].

条件数$\tau(B)$反映了$BX=b$的解X的相对误差对于B和b的相对误差的依赖程度, 当条件数$\tau(B)$很大时, 即使B和b有较小的扰动, 也会引起X的较大偏差, 要改善X受B和b扰动的影响程度, 降低条件数是比较直接的方法, 因为条件数较小, 相应的解X抗干扰的能力更强.

定义 3.1 降维后模型(2.2)的设计阵满秩, 且$\widetilde{\alpha}^i_{\mathrm{SDE}}(k_i) = A_i\widetilde{\alpha}^i_{\mathrm{SDE}}$, 我们称

$$\tau(\widetilde{\alpha}^i_{\mathrm{SDE}}(k_i)) = \|A_i^{-1}\Lambda_i\|\|A_i\Lambda_i^{-1}\|$$

为估计$\widetilde{\alpha}^i_{\mathrm{SDE}}(k_i)$的条件数.

对线性模型(2.2)的正则方程$Z_i'Z_i\beta = Z_i'y^{(i)}$, 同样有

$$\frac{\|\Delta\alpha\|}{\|\alpha\|} \leq \frac{\tau(Z_i'Z_i)}{1-\tau(Z_i'Z_i)\|\Delta(Z_i'Z_i)\|\|Z_i'Z_i\|^{-1}}\Big(\frac{\|\Delta(Z_i'Z_i)\|}{\|Z_i'Z_i\|} + \frac{\|\Delta(Z_i'y^{(i)})\|}{\|Z_i'y^{(i)}\|}\Big).$$

而$\tau(Z_i'Z_i)=\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i)$, 可见, $\widetilde{\alpha}_{\mathrm{SDE}}^i$的条件数反映了$\widetilde{\alpha}_{\mathrm{SDE}}^i$的抗干扰程度. 因此, 我们自然希望部分岭型谱分解估计既改进谱分解估计, 又不降低抗干扰能力, 就得要求

$$\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))\leq\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i).$$

我们称满足这个条件的估计$\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)$具有相对于$\widetilde{\alpha}_{\mathrm{SDE}}^i$更强的抗干扰性.

定理 3.1 对模型(2.2), 总有$\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))\leq\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i)$. 因而$\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)$改进了$\widetilde{\alpha}_{\mathrm{SDE}}^i$的抗干扰性.

证明: 因为

$$\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))=\|A_i^{-1}\Lambda_i\|\|A\Lambda_i^{-1}\|=\tau(\Lambda_i+k_iI)=\frac{\delta_{i1}+k_i}{\delta_{ir_i}+k_i},$$

同样

$$\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i)=\|\Lambda_i\|\|\Lambda_i^{-1}\|=\frac{\delta_{i1}}{\delta_{ir_i}},$$

所以

$$\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))=\frac{\delta_{i1}+k_i}{\delta_{ir_i}+k_i}\leq\frac{\delta_{i1}}{\delta_{ir_i}}=\tau(\widetilde{\alpha}_{\mathrm{SDE}}^i).$$

□

§4. 部分岭型谱分解估计的优效性

比较两个估计的另一个重要方法是相对效率, 文献[10]给出了线性模型参数估计的主要相对效率形式, 对于本文讨论的有偏估计, 我们尝试引入如下的相对效率定义:

定义 4.1 对模型(2.2), 设$\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)$为$\alpha$的部分岭型谱分解估计, 我们用

$$1-\frac{\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))}{\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i)}$$

来表示$\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)$相对于$\widetilde{\alpha}_{\mathrm{SDE}}^i$的效率, 记为$\rho(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))$.

如果$\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))\leq\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i)$, 则效率不会为负, 这在设计阵病态时是很容易满足的, 这时$0\leq\rho(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))\leq 1$. 效率越接近于1, $\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)$改进$\widetilde{\alpha}_{\mathrm{SDE}}^i$的程度也就越高.

定理 4.1 对每一个i, 当$\alpha'a\leq(\delta_{ir_i}^{-1}-1)\lambda_i$时,

$$\rho(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))\geq\frac{1-\delta_{ir_i}-\alpha'\alpha\lambda_i^{-1}\delta_{ir_i}}{1+(r_i-1)\delta_{ir_i}}.$$

证明:

$$\begin{aligned}\rho(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)) &= 1-\frac{\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))}{\mathrm{MSE}(\widetilde{\alpha}_{\mathrm{SDE}}^i)}\\ &= 1-\frac{\lambda_i\mathrm{tr}(A_i\Lambda_i^{-1}A_i')+\alpha'(A_i-I)^2\alpha}{\lambda_i\mathrm{tr}\Lambda_i^{-1}}\geq 1-\frac{\lambda_i\mathrm{tr}(A\Lambda_i^{-1})+\alpha'\alpha}{\lambda_i\mathrm{tr}\Lambda_i^{-1}}.\end{aligned}$$

因为$(x+a)/(x+b)$当$a<b$是x的增函数, 而由已知条件可得$(1+\alpha' a\lambda_i^{-1})\delta_{ir_i}\le 1$, 所以有

$$\rho(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))\ge 1-\frac{r_i+\alpha'\alpha\lambda_i^{-1}}{r_i-1+\delta_{ir_i}^{-1}}=\frac{1-\delta_{ir_i}-\alpha'\alpha\lambda_i^{-1}\delta_{ir_i}}{1+(r_i-1)\delta_{ir_i}}.$$

可见, 当设计阵病态, 即δ_{ir_i}很小且使$\alpha'\alpha\lambda_i^{-1}\delta_{ir_i}$相对较小时, $\rho(\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i))$可以很接近于1, 即是说$\widetilde{\alpha}_{\mathrm{SDE}}^i(k_i)$可以很好的改进$\widetilde{\alpha}_{\mathrm{SDE}}^i$. □

§5. 部分岭型谱分解估计在Panel模型中的应用

考虑含有一个随机效应的Panel模型

$$y_{it}=\beta_0+x_{it_1}\beta_1+\cdots+x_{it_p}\beta_p+\mu_i+\varepsilon_{it},\qquad i=1,2,\cdots,N;\ t=1,2,\cdots,T, \tag{5.1}$$

这里y_{it}表示第i个个体第t次观测值, μ_i是第i个个体的随机效应, μ_i和ε_{it}相互独立, $\mu_i\sim N(0,\sigma_\mu^2)$, $\varepsilon_{it}\sim N(0,\sigma_\varepsilon^2)$. 可以将模型(5.1)用矩阵表示为[1]

$$y=1_{NT}\beta_0+X\beta+(I_N\otimes 1_T)\mu+\varepsilon=1_{NT}\beta_0+X\beta+u. \tag{5.2}$$

模型(5.2)右边随机部分u的均值为0, 协方差阵为$V=\sigma_\mu^2(I_N\otimes J_T)+\sigma_\varepsilon^2 I_{NT}$. 文献[1]指出这个模型的固定效应$\beta$的两种重要估计Between估计和Within估计都是谱分解估计, 事实上对模型(5.2)的协方差进行谱分解: $\lambda_1=T\sigma_\mu^2+\sigma_\varepsilon^2$, $\lambda_2=\sigma_\varepsilon^2$, 相应得到标准特征向量所构成的矩阵$M_1=I_N\otimes\overline{J}_T$, $M_2=I_N\otimes(I_T-\overline{J}_T)$. 采用文献[1]类似的处理, 消去$\beta_0$, 将$M_1$修正为$P_Z=(I_N-\overline{J}_N)\otimes\overline{J}_T$, 它依旧具有$M_1$的所有性质, 在$\mathrm{rank}(X'P_ZX)=\mathrm{rank}(X'M_2X)=p$的条件下, 利用文献[1]同样的方法, 可以获得$\beta$的两个谱分解估计

$$\widetilde{\beta}_{\mathrm{SDE}}^B=(X'P_ZX)^{-1}X'P_Zy,\qquad \widetilde{\beta}_{\mathrm{SDE}}^W=(X'M_2X)^{-1}X'M_2y.$$

而应用本文的方法, 我们可以相应的定义如下两个有偏估计

$$\widetilde{\beta}_{\mathrm{SDE}}^B(k_i)=(X'P_ZX+k_iQ_iQ_i')^{-1}X'P_Zy,\qquad \widetilde{\beta}_{\mathrm{SDE}}^W(k_i)=(X'M_2X+k_iQ_iQ')^{-1}X'M_2y,$$

分别称为部分岭型Between估计和部分岭型Within估计, 而方差分量的谱分解估计可以很容易的通过λ_i的估计

$$\widehat{\lambda}_i^B=(y_i-X\widetilde{\beta}_{\mathrm{SDE}}^B(k_i))'M_i(y_i-X\widetilde{\beta}_{\mathrm{SDE}}^B(k_i))/f,\qquad i=1,2,\cdots,t$$

和

$$\widehat{\lambda}_i^W=(y_i-X\widetilde{\beta}_{\mathrm{SDE}}^W(k_i))'M_i(y_i-X\widetilde{\beta}_{\mathrm{SDE}}^W(k_i))/f,\qquad i=1,2,\cdots,t$$

得到, 其中$f=R(M_i)-R(M_iX)$.

类似的可以由上文的方法得到它们的相应优良性质, 尤其在设计阵病态时用于各种统计推断会取得明显的应用效果, 限于篇幅, 这里不再赘述.

参 考 文 献

[1] 王松桂, 尹素菊, 线性混合模型参数的一种新估计, 中国科学, **32(5)**(2002), 434–443.

[2] Verbeke, G., Molenberghs, G., *Linear Mixed Models in Pactice: A SAS-Orented Approach, Lecture Notes in Statistics*, Vol.126, New York: Springer-Verlag, 1997.

[3] Verbeke, G., Molenberghs, G., *Linear Mixed Models for Longitudinal Data*, New York: Springer-Verlag, 2000.

[4] Wang, S.G., Chow, S.C., *Advanced Linear Models*, New York: Marcel Dekker Inc, 1994.

[5] Khunri, A.I., Mathew, T., Sinha, B.K., *Statistica Tests for Mixed Linear models*, New York: Wiley, 1998.

[6] Searle, S.R., Casella, G., McCulloch, C.E., *Variance Components*, New York: Wiley, 1998.

[7] Horel, A.E., Kennard, R.W., Ridge regression: biased estimation for non-orthogonal problems, *Technometrics*, **12**(1970), 55–88.

[8] 王松桂, 线性模型的理论及应用, 安徽教育出版社, 1987.

[9] 孙继广, 矩阵扰动分析, 科学出版社, 1987.

[10] Yang, H., Two new classes of the generalized kantorovich inequalities and their applications, *Sci. Bulletin*, **36(22)**(1991), 1849–1851.

Partial Ridge-Type Spectral Decomposition Estimate in Linear Mixed Model

Yang Hu Li Yalian

(*Department of Statistics and Actuarial Science, College of Mathematics and Physics,*)

(*Chongqing University, Chongqing, 400030*)

As to the linear model with its random effect being common balance multiple classification, Wang and Yin[1] introduced a new parameter estimation method called the spectral decomposition estimate (SDE). In this paper, we generalize the SDE into the situation when the design matrix is ill-conditioned and propose the partial ridge-type spectral decomposition estimate. Through the model transformation of dimensionality reduction which is similar to the principal component estimate, we can easily get that the new estimate has stronger noise-rejection ability and some other important properties. Furthermore, the results we get in the paper can also be easily used in Panel model.

Keywords: Partial ridge-type spectral decomposition estimate, linear mixed model, admissibility.

AMS Subject Classification: 62J07.