应用概率统计 第二十五卷
第一期 2009年2月

Chinese Journal of Applied Probability
and Statistics Vol.25 No.1 Feb. 2009

随机变量二次型的协方差在混合效应模型中的应用 *

马铁丰
(西南财经大学统计学院, 成都, 610074)

王松桂 尹素菊
(北京工业大学应用数理学院, 北京, 100124)

摘 要

本文提出方差分量ANOVA估计的一种改进方法, 证明了对于一般的方差分量模型, 只要方差分量的ANOVA估计存在就可以通过此方法给出其改进形式, 并且在均方误差意义下优于ANOVA估计. 特别地, 对于单向分类随机效应模型, Kelly和Mathew[1]对ANOVA估计的改进就是我们提出的改进方法的特殊形式, 这也给出了此类改进估计在均方误差意义下优于ANOVA估计的另一种合理的解释. 同时, 本文又将此思想应用到对谱分解估计的改进上. 本文应用协方差的简单性质证明了对带有一个随机效应的方差分量模型, 当随机效应的协方差阵只有一个非零特征值时, 随机效应方差分量谱分解估计在均方误差意义下总是优于ANOVA估计. 本文最后将第三节的结论推广到广义谱分解估计下, 同时给出广义谱分解估计待定系数的一个合理的取值.

关 键 词: 方差分量模型, 谱分解估计, ANOVA估计, 均方误差.

学 科 分 类 号: O212.1.

§1. 引 言

方差分量模型的一般形式为:

$$y = X\beta + U_1\xi_1 + U_2\xi_2 + \cdots + U_k\xi_k + \varepsilon. \tag{1.1}$$

这里y为$N\times 1$观测向量, X为已知$N\times t$设计阵, U_i是已知的设计矩阵, 随机效应ξ_i, $i=1,\cdots,k$互不相关, 且与ε也不相关, 为简单记$U_{k+1}=I$, $\xi_{k+1}=\varepsilon$. 一般假设$\xi_i\sim N(0,\sigma_i^2 I)$, 这时

$$\mathsf{Cov}\,(y) = \sum_{i=1}^{k+1} U_iU_i'\sigma_i^2 = V(\sigma^2). \tag{1.2}$$

其中$\sigma^2=(\sigma_1^2,\cdots,\sigma_{k+1}^2)'$, σ_i^2是未知参数, 称为方差分量. 关于方差分量的估计主要有以下五种常用的方法: 方差分析方法(ANOVA), 极大似然(ML)估计[2], 限制极大似然(RML)估计[3], 最小范数二次无偏(MINQU)估计[4–6]以及由王松桂和尹素菊[7]新近提出的谱分解估计, 其中方差分析方法拥有最长发展历史, 最早可追溯到1861年, Airy的工作[8]. ANOVA估计之所以被广泛应用, 一个重要原因就是在一定条件下ANOVA估计是最小方差无偏估计(Minimum Variance Unbiased Estimate, MVUE), 参见文献[9]. ANOVA估计的基本思

*北京市属市管高等学校人才强教计划(05006011200702)和国家自然科学基金数学天元青年基金(10726045)、国家自然科学基金青年基金(10801005)资助.

本文2006年3月6日收到, 2007年10月12日收到修改稿.

《应用概率统计》版权所有

想是将模型中的随机效应暂时视为固定效应, 即常数, 求得各效应的回归平方和, 令各回归平方和等于它们的期望, 所得方程组的解便是方差分量的ANOVA估计. 记

$$s_i = y'(P_{(X:U_1:\cdots:U_i)} - P_{(X:U_1:\cdots:U_{i-1})})y, \qquad i = 1, \cdots, k+1, \tag{1.3}$$

这里P_A为向$M(A)$的正交投影阵, $M(A)$为由矩阵A的列向量张成的子空间, U_0为零矩阵.

令

$$\mathsf{E}(s_i) = \sum_{j=1}^{k+1} c_{ij}\sigma_j^2. \tag{1.4}$$

记$C = (c_{ij})$, $s = (s_1, \cdots, s_{k+1})'$, 若$C$可逆, 则$\sigma^2$的ANOVA估计为

$$\widehat{\sigma}^2 = C^{-1}s. \tag{1.5}$$

本文利用随机效应方差分量ANOVA估计与随机误差方差分量ANOVA估计的相关性, 给出随机效应方差分量ANOVA估计的改进估计, 我们证明了对于一般的方差分量模型只要方差分量的ANOVA估计存在, 则我们给出的改进估计都会在均方误差意义下优于原估计; 在本文的第三部分, 我们首先将第二节的改进思想应用到对谱分解估计的改进上, 得到了类似的改进估计和优良性结论, 然后又证明了对于带有一个随机效应的方差分量模型, 当随机效应的协方差阵只有一个非零特征值时, 随机效应方差分量谱分解估计在均方误差意义下总是优于ANOVA估计. 在本文的第四部分, 我们将文中第三部分的结论推广到广义谱分解估计下, 找到了在广义谱分解估计类中在均方误差意义下一致优于ANOVA估计的广义谱分解估计.

§2. 方差分量ANOVA估计的改进

首先给出证明所需的引理.

引理 2.1 (二次投影定理) 设A, B是具有相同行数的矩阵, 记$C = (A:B)$, 则

$$P_C = P_A + Q_AB(B'Q_AB)^-B'Q_A, \tag{2.1}$$

这里$P_A = A(A'A)^-A'$是子空间上$M(A)$的正交投影阵, $Q_A = I - P_A$.

证明参见[11].

引理 2.2 设$Y \sim N_p(\mu, V)$, 则

$$\mathsf{Var}\,(Y'BY) = 2\mathrm{tr}[(BV)^2] + 4\mu'BVB\mu, \tag{2.2}$$

$$\mathsf{Cov}\,(Y'AY, Y'BY) = \mathrm{tr}[AVBV] + \mathrm{tr}[AVB'V] + \mu'(A+A')V(B+B')\mu. \tag{2.3}$$

证明参见[12].

记$D = (d_{ij}) = C^{-1}$, 则由(1.5)式显然有

$$\widehat{\sigma}_i^2 = \sum_{j=1}^{k+1} d_{ij}s_j = y'Ay + d_{i,k+1}y'Qy, \qquad i = 1, \cdots, k. \tag{2.4}$$

《应用概率统计》版权所有

其中

$$A = \sum_{j=1}^{k} d_{ij}(P_{(X:U_1:\cdots:U_j)} - P_{(X:U_1:\cdots:U_{j-1})}),$$
$$Q = P_{(X:U_1:\cdots:U_{k+1})} - P_{(X:U_1:\cdots:U_k)} = I - P_{(X:U_1:\cdots:U_k)}.$$

另外, 结合引理2.2有随机误差项的ANOVA估计为

$$\widehat{\sigma}^2_{k+1} = \frac{s_{k+1}}{\mathrm{tr}(Q)}. \tag{2.5}$$

其方差为$\mathsf{Var}\,(\widehat{\sigma}^2_{k+1}) = 2\sigma^4_{k+1}[\mathrm{tr}(Q^2)/\mathrm{tr}^2(Q)] = 2\sigma^4_{k+1}/\mathrm{tr}(Q)$.

$\widehat{\sigma}^2_{k+1}$与$\widehat{\sigma}^2_i$的协方差为

$$\mathsf{Cov}\,(\widehat{\sigma}^2_{k+1}, \widehat{\sigma}^2_i) = 2d_{i,k+1}\frac{\mathrm{tr}(Q^2)}{\mathrm{tr}(Q)}\sigma^4_{k+1} = 2d_{i,k+1}\sigma^4_{k+1}. \tag{2.6}$$

这是因为$QA = 0$. 据此, 我们在$\widehat{\sigma}^2_i$的基础上给出一种改进估计:

$$\widetilde{\sigma}^2_i = \widehat{\sigma}^2_i - m\widehat{\sigma}^2_{k+1}. \tag{2.7}$$

这里m为常数. 在下面讨论均方误差意义下$\widetilde{\sigma}^2_i$与$\widehat{\sigma}^2_i$的比较中将给出m的具体取值范围.

定理 2.1 若$d_{i,k+1} > 0$, 则有当$0 < m < 4\mathrm{tr}(Q)d_{i,k+1}/[\mathrm{tr}(Q)+2]$时, 在均方误差意义下$\widetilde{\sigma}^2_i$优于$\widehat{\sigma}^2_i$; 若$d_{i,k+1} < 0$, 则有当$4\mathrm{tr}(Q)d_{i,k+1}/[\mathrm{tr}(Q)+2] < m < 0$时, 在均方误差意义下$\widetilde{\sigma}^2_i$优于$\widehat{\sigma}^2_i$.

证明: $\widetilde{\sigma}^2_i$的均方误差(MSE)为:

$$\begin{aligned}
\mathrm{MSE}(\widetilde{\sigma}^2_i) &= \mathsf{E}(\widehat{\sigma}^2_i - m\widehat{\sigma}^2_{k+1} - \sigma^2_i)^2 \\
&= \mathsf{E}(\widehat{\sigma}^2_i - \sigma^2_i - m(\widehat{\sigma}^2_{k+1} - \sigma^2_{k+1}) - m\sigma^2_{k+1})^2 \\
&= \mathsf{E}(\widehat{\sigma}^2_i - \sigma^2_i)^2 - 2m\mathsf{E}((\widehat{\sigma}^2_i - \sigma^2_i)(\widehat{\sigma}^2_{k+1} - \sigma^2_{k+1})) \\
&\quad + m^2\mathsf{E}(\widehat{\sigma}^2_{k+1} - \sigma^2_{k+1})^2 + m^2\sigma^4_{k+1} \\
&= \mathrm{MSE}(\widehat{\sigma}^2_i) - 2m\mathsf{Cov}\,(\widehat{\sigma}^2_i, \widehat{\sigma}^2_{k+1}) + m^2\mathsf{Var}\,(\widehat{\sigma}^2_{k+1}) + m^2\sigma^4_{k+1} \\
&= \mathrm{MSE}(\widehat{\sigma}^2_i) - 4md_{i,k+1}\sigma^4_{k+1} + m^2\Big(1 + \frac{2}{\mathrm{tr}(Q)}\Big)\sigma^4_{k+1}.
\end{aligned}$$

由此显然, 当m介于0与$4\mathrm{tr}(Q)d_{i,k+1}/[\mathrm{tr}(Q)+2]$之间时, $\mathrm{MSE}(\widetilde{\sigma}^2_i) < \mathrm{MSE}(\widehat{\sigma}^2_i)$, 即在均方误差意义下$\widetilde{\sigma}^2_i$优于$\widehat{\sigma}^2_i$. 证毕. □

特别地, 取$k = 1$, 则模型(1.1)为两方差分量的线性混合模型:

$$y = X\beta + U_1\xi_1 + \varepsilon. \tag{2.8}$$

记$V = U_1U_1'$, Z为$N\times q$矩阵且满足$Z'X = 0$和$Z'Z = I_q$, 再记$V_1 = Z'VZ$, $q = N - \mathrm{rank}(X)$, $u = Z'y$. 令$s = \mathrm{rank}(V_1)$, 同时考虑V_1的谱分解

$$V_1 = \sum_{i=1}^{g} \lambda_i E_i. \tag{2.9}$$

《应用概率统计》版权所有

这里$\lambda_1 > \lambda_2 > \cdots > \lambda_g$为所有非零特征值, 它们的重数分别为$s_1, s_2, \cdots, s_g$, 对于任意$1 \le i \le g$, E_i为向特征值λ_i对应的特征子空间投影的正交投影阵. 显然$\mathrm{rk}(E_i) = s_i$, $s = \sum\limits_{i=1}^{g} s_i$, 记$E_{g+1} = I - \sum\limits_{i=1}^{g} E_i$, 则$\mathrm{rk}(E_{g+1}) = q - s$.

方差分量σ_1^2的ANOVA估计为

$$\widehat{\sigma}_1^2 = \frac{1}{\mathrm{tr}(V_1)}\Big(\sum_{i=1}^{g} u'E_i u - \frac{s}{q-s}u'E_{g+1}u\Big). \tag{2.10}$$

方差分量σ_2^2的ANOVA估计为

$$\widehat{\sigma}_2^2 = \frac{1}{q-s}u'E_{g+1}u. \tag{2.11}$$

特别地, 当$g = 1$时, 方差分量σ_1^2的ANOVA估计为

$$\widehat{\sigma}_1^2 = \frac{1}{\mathrm{tr}(V_1)}\Big(u'E_1 u - \frac{s}{q-s}u'E_2 u\Big). \tag{2.12}$$

方差分量σ_2^2的ANOVA估计为

$$\widehat{\sigma}_2^2 = \frac{1}{q-s}u'E_2 u. \tag{2.13}$$

Kelly和Mathew[1]给出ANOVA估计的一种改进形式:

$$\begin{aligned}
\widehat{\sigma}_{1(c,r)}^2 &= \frac{c}{\lambda_1}\Big(\frac{1}{s}u'E_1 u - \frac{r}{q-s}u'E_2 u\Big)\\
&= c\widehat{\sigma}_1^2 + \frac{c(1-r)}{\lambda_1(q-s)}u'E_2 u\\
&= c\widehat{\sigma}_1^2 + \frac{c(1-r)}{\lambda_1}\widehat{\sigma}_2^2.
\end{aligned} \tag{2.14}$$

这里c为区间$[0,1]$上的实数.

取$c = 1$, 比较(2.14)与(2.7)两式, 我们显然看出Kelly和Mathew给出的ANOVA估计的改进形式本质上就是我们给出的改进估计在特定模型下的形式, 其之所以能够在均方误差意义下改进ANOVA估计也就是源于两方差分量ANOVA估计的相关性.

例 1 单向分类模型

考虑下面的单向分类模型

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \qquad i = 1, 2, \cdots, n, \ j = 1, 2, \cdots, m. \tag{2.15}$$

将此模型表示成矩阵形式为

$$y = \mu\mathbf{1}_{nm} + (I_n \otimes \mathbf{1}_m)\alpha + \varepsilon. \tag{2.16}$$

显然此模型是模型(2.8)的特殊情况, 因此对方差分量ANOVA估计作相应的改进可以在均方误差意义下得到更优的估计.

《应用概率统计》版权所有

《应用概率统计》版权所有

§3. 谱分解估计与ANOVA估计的比较

对于模型(2.8), 假设V有谱分解:

$$V = \tau_1 M_1. \tag{3.1}$$

记$M = I - M_1$, 用M_1左乘模型(2.8), 得到

$$y^{(1)} = X_1\beta + \varepsilon_1, \qquad \varepsilon_1 \sim N(0, (\sigma^2 + \tau_1\sigma_1^2)M_1). \tag{3.2}$$

这里$y^{(1)} = M_1 y$, $X_1 = M_1 X$, $\varepsilon_1 = M_1 U_1 \xi_1 + M_1\varepsilon$.

文献[7]提出的方差分量σ_1^2和σ_2^2的谱分解估计为:

$$\widehat{\sigma}_{1S}^2 = \frac{1}{\tau_1}\Big[\frac{1}{r_1}y'(M_1 - M_1X(X'M_1X)^- X'M_1)y - \frac{1}{r}y'(M - MX(X'MX)^- X'M)y\Big], \tag{3.3}$$

$$\widehat{\sigma}_{2S}^2 = \frac{1}{r}y'[M - MX(X'MX)^- X'M]y. \tag{3.4}$$

这里$r_1 = \mathrm{rk}(M_1) - \mathrm{rk}(M_1X)$, $r = \mathrm{rk}(M) - \mathrm{rk}(MX)$.

上一节的改进思想也可以应用到谱分解估计上, 得到类似的改进结果.

类似于(2.7), 我们给出新估计:

$$\widetilde{\sigma}_{1S}^2 = \widehat{\sigma}_{1S}^2 - m\widehat{\sigma}_{2S}^2. \tag{3.5}$$

这里m为常数.

定理 3.1 当$-4/[\tau_1(r+2)] < m < 0$时, 在均方误差意义下$\widetilde{\sigma}_{1S}^2$优于$\widehat{\sigma}_{1S}^2$.

证明: 为表述简单, 我们记$B = M - MX(X'MX)^- X'M$.

应用引理2.2有

$$\mathsf{Cov}\,(\widehat{\sigma}_{1S}^2, \widehat{\sigma}_{2S}^2) = -\frac{2\mathrm{tr}B^2}{\tau_1 r^2}\sigma_2^4 = -\frac{2}{\tau_1 r}\sigma_2^4, \qquad \mathsf{Var}\,(\widehat{\sigma}_{2S}^2) = \frac{2\mathrm{tr}B^2}{r^2}\sigma_2^4 = \frac{2}{r}\sigma_2^4.$$

接下来计算$\widetilde{\sigma}_{1S}^2$的均方误差(MSE):

$$\begin{aligned}
\mathrm{MSE}(\widetilde{\sigma}_{1S}^2) &= \mathsf{E}(\widehat{\sigma}_{1S}^2 - m\widehat{\sigma}_{2S}^2 - \sigma_1^2)^2 \\
&= \mathsf{E}(\widehat{\sigma}_{1S}^2 - \sigma_1^2 - m(\widehat{\sigma}_{2S}^2 - \sigma_2^2) - m\sigma_2^2)^2 \\
&= \mathsf{E}(\widehat{\sigma}_{1S}^2 - \sigma_q^2)^2 - 2m\mathsf{E}((\widehat{\sigma}_{1S}^2 - \sigma_1^2)(\widehat{\sigma}_{2S}^2 - \sigma_2^2)) + m^2\mathsf{E}(\widehat{\sigma}_{2S}^2 - \sigma_2^2)^2 + m^2\sigma_2^4 \\
&= \mathrm{MSE}(\widehat{\sigma}_{1S}^2) - 2m\mathsf{Cov}\,(\widehat{\sigma}_{1S}^2, \widehat{\sigma}_{2S}^2) + m^2\mathsf{Var}\,(\widehat{\sigma}_{2S}^2) + m^2\sigma_2^4 \\
&= \mathrm{MSE}(\widehat{\sigma}_{1S}^2) + \frac{4m}{\tau_1 r}\sigma_2^4 + m^2\Big(1 + \frac{2}{r}\Big)\sigma_2^4.
\end{aligned}$$

由此显然, 当$-4/[\tau_1(r+2)] < m < 0$时, 在均方误差意义下$\widetilde{\sigma}_{1S}^2$优于$\widehat{\sigma}_{1S}^2$. 证毕. □

由于谱分解估计有着很好的统计性质, 因此受许多学者关注, 其中主要是将其与ANOVA估计进行比较. 目前的文献主要的比较工作都是集中在讨论谱分解估计与

ANOVA估计等价的问题, 如: 吴密霞[10]给出在带有一个随机效应的纵向平衡数据模型下, 随机效应方差分量谱分解估计与ANOVA估计相等的充要条件. 但是均没有讨论在何种情况下谱分解估计优于ANOVA估计. 本文证明了在更广范围的模型下谱分解估计优于ANOVA估计, 且不需任何条件.

定理 3.2 对于模型(2.8), 若$V=U_1U_1'$只有一个非零特征值, 则随机效应的方差分量的谱分解估计均方误差不大于ANOVA估计的均方误差.

证明: 当$V=U_1U_1'$只有一个非零特征值时, 随机效应的方差分量的ANOVA估计仍用(2.10)表示. 虽然(2.10)式的得出没有V只有一个非零特征值这个条件, 但是当V只有一个非零特征值时, $V_1=Z'VZ$的非零特征值的个数可以不只有一个, 因此为行文简单我们仍记V_1的非零特征值的个数为g, g为正整数. 此时V_1的谱分解仍然用(2.9)表示, 随机效应的方差分量的ANOVA估计仍用(2.10)表示. 随机效应的方差分量的谱分解估计用由(3.3)给出.

根据引理2.2有:

$$\begin{aligned}
&\mathsf{Cov}\,(\widehat{\sigma}_{1S}^2,\widehat{\sigma}_1^2)\\
=\ &2\mathrm{tr}\Big[\frac{1}{\tau_1}\Big(\frac{1}{r_1}(M_1-P_{M_1X})-\frac{1}{r}(M-P_{MX})\Big)(\sigma_2^2I+\sigma_1^2V)\\
&\cdot\frac{1}{\mathrm{tr}(V_1)}\Big(\sum_{i=1}^{g}ZE_iZ'-\frac{s}{q-s}ZE_{g+1}Z'\Big)(\sigma_2^2I+\sigma_1^2V)\Big]\\
=\ &\frac{2}{\tau_1\mathrm{tr}(V_1)}\mathrm{tr}\Big[\Big(\sigma_2^2\Big(\frac{1}{r_1}(M_1-P_{M_1X})-\frac{1}{r}(M-P_{MX})\Big)+\sigma_1^2\frac{\tau_1}{r_1}(M_1-P_{M_1X})\Big)\\
&\cdot Z\Big(\sum_{i=1}^{g}E_i-\frac{s}{q-s}E_{g+1}\Big)Z'(\sigma_2^2I+\sigma_1^2V)\Big].
\end{aligned}$$

由引理2.1有$M_1-P_{M_1X}=M_1-(P_{(M:X)}-M)=I-P_{(M:X)}$, 进而有$(M_1-P_{M_1X})Z=(I-P_{(M:X)})Z=(I-P_M)Z=M_1Z$, 同理$(M-P_{MX})Z=MZ$. 记$G=\sum\limits_{i=1}^{g}E_i-[s/(q-s)]E_{g+1}$. 因此上式等价于

$$\begin{aligned}
=\ &\frac{2}{\tau_1\mathrm{tr}(V_1)}\mathrm{tr}\Big[\Big(\sigma_2^2\Big(\frac{1}{r_1}M_1ZGZ'-\frac{1}{r}MZGZ'\Big)+\sigma_1^2\frac{\tau_1}{r_1}M_1ZGZ'\Big)(\sigma_2^2I+\sigma_1^2V)\Big]\\
=\ &\frac{2}{\tau_1\mathrm{tr}(V_1)}\Big[\sigma_2^4\Big(\frac{1}{r_1}\mathrm{tr}(M_1ZGZ')-\frac{1}{r}\mathrm{tr}(MZGZ')\Big)+\sigma_1^4\frac{\tau_1^2}{r_1}\mathrm{tr}(M_1ZGZ')\\
&+2\sigma_1^2\sigma_2^2\frac{\tau_1}{r_1}\mathrm{tr}(M_1ZGZ')\Big].
\end{aligned}$$

又因为

$$\mathrm{tr}(M_1ZGZ')=\mathrm{tr}(Z'M_1ZG)=\frac{1}{\tau_1}\mathrm{tr}\Big(V_1\Big(\sum_{i=1}^{g}E_i-\frac{s}{q-s}E_{g+1}\Big)\Big)=\frac{1}{\tau_1}\mathrm{tr}(V_1),$$

$$\mathrm{tr}(MZGZ')=\mathrm{tr}(Z'(I-M_1)ZG)=\mathrm{tr}(G)-\frac{1}{\tau_1}\mathrm{tr}(V_1)=-\frac{1}{\tau_1}\mathrm{tr}(V_1).$$

《应用概率统计》版权所有

所以上式等价于

$$
\begin{aligned}
&= \frac{2}{\tau_1 \mathrm{tr}(V_1)}\Big[\sigma_2^4\Big(\frac{\mathrm{tr}(V_1)}{r_1\tau_1}+\frac{\mathrm{tr}(V_1)}{r\tau_1}\Big)+\sigma_1^4\frac{\tau_1\mathrm{tr}(V_1)}{r_1}+2\sigma_1^2\sigma_2^2\frac{\mathrm{tr}(V_1)}{r_1}\Big]\\
&= \frac{2}{\tau_1^2}\Big(\frac{1}{r_1}+\frac{1}{r}\Big)\sigma_2^4+\frac{2}{r_1}\sigma_1^4+\frac{4}{r_1\tau_1}\sigma_1^2\sigma_2^2\\
&= \mathsf{Var}\,(\widehat{\sigma}_{1S}^2).
\end{aligned}
$$

又因为$\mathsf{Cov}\,(\widehat{\sigma}_{1S}^2, \widehat{\sigma}_1^2) \le \sqrt{\mathsf{Var}\,(\widehat{\sigma}_{1S}^2)\mathsf{Var}\,(\widehat{\sigma}_1^2)}$, 所以有$\mathsf{Var}\,(\widehat{\sigma}_{1S}^2) \le \sqrt{\mathsf{Var}\,(\widehat{\sigma}_{1S}^2)\mathsf{Var}\,(\widehat{\sigma}_1^2)}$, 即$\mathsf{Var}\,(\widehat{\sigma}_{1S}^2) \le \mathsf{Var}\,(\widehat{\sigma}_1^2)$. 证毕. □

例 2　纵向数据模型

考虑带一个随机效应的纵向数据模型

$$y_{ij} = x'_{ij}\beta + u_i + \varepsilon_{ij}, \qquad i = 1, 2, \cdots, N,\ j = 1, 2, \cdots, T. \tag{3.6}$$

这里x_{ij}为$p \times 1$回归向量, β为$p \times 1$未知的固定效应, 假设随机效应u_i与误差ε_{ij}分别服从均值为0, 方差为σ_u^2和σ_ε^2的正态分布, 且彼此独立.

将此模型表示成矩阵形式为

$$y = X\beta + (I_N \otimes \mathbf{1}_T)u + \varepsilon. \tag{3.7}$$

显然$V = (I_N \otimes \mathbf{1}_T) \times (I_N \otimes \mathbf{1}_T)'$的非零特征值只有一个, 因此满足定理3.2的条件. 所以对于此类模型, 选择谱分解法估计方差分量比选择ANOVA法更好.

§4.　广义谱分解估计与ANOVA估计的比较

下面我们将第三节的结论推广到广义谱分解估计.

对于模型(2.8), 假设V有谱分解:

$$V = \sum_{i=1}^{k} \tau_i M_i. \tag{4.1}$$

记$M_0 = I - M_i$, 用M_i左乘模型(2.8), 得到$k+1$个新模型, 它们是

$$y^{(i)} = X_i\beta + \varepsilon_i, \qquad \varepsilon_i \sim N(0, (\sigma^2 + \tau_i\sigma_1^2)M_i). \tag{4.2}$$

这里$y^{(i)} = M_i y$, $X_i = M_i X$, $\varepsilon_i = M_i U\xi + M_i\varepsilon$, $\tau_0 = 0$.

记

$$\widehat{\sigma}_S^2 = \sum_{i=0}^{k} \frac{c_i}{r_i} y'(M_i - P_{M_iX})y. \tag{4.3}$$

文献[13]给出:

《应用概率统计》版权所有

(1) 若c_i $(i=0,1,\cdots,k)$满足条件

$$\sum_{i=0}^{k} c_i = 0, \qquad \sum_{i=0}^{k} \tau_i c_i = 1, \tag{4.4}$$

则$\widehat{\sigma}_S^2$为方差分量σ_1^2的广义谱分解估计.

(2) 若c_i $(i=0,1,\cdots,k)$满足条件

$$\sum_{i=0}^{k} c_i = 1, \qquad \sum_{i=0}^{k} \tau_i c_i = 0, \tag{4.5}$$

则$\widehat{\sigma}_S^2$为方差分量σ_2^2的广义谱分解估计.

这里$r_i=\mathrm{rk}(M_i)-\mathrm{rk}(M_iX)$.

此时显然方差分量的广义谱分解估计为一估计类. 然而文献[13]并未给出很好的确定c_i的方法, 只是通过两步的方法给出了非线性的估计, 小样本性质很难研究. 下面的定理将解决这个问题.

定理 4.1 对于模型(2.8), 在随机效应方差分量的广义谱分解估计类中存在一均方误差不大于ANOVA估计均方误差的广义谱分解估计.

证明: 根据引理2.2有:

$$\begin{aligned}
\mathsf{Cov}\,(\widehat{\sigma}_S^2,\widehat{\sigma}_1^2) &= 2\mathrm{tr}\Big[\Big(\sum_{i=0}^{k}\frac{c_i}{r_1}(M_i-P_{M_iX})\Big)(\sigma^2 I+\sigma_1^2 V)\\
&\quad \cdot\frac{1}{\mathrm{tr}(V_1)}\Big(\sum_{i=1}^{g} ZE_iZ'-\frac{s}{q-s}ZE_{g+1}Z'\Big)(\sigma^2 I+\sigma_1^2 V)\Big]\\
&= \frac{2}{\mathrm{tr}(V_1)}\mathrm{tr}\Big[\Big(\sigma^2\sum_{i=0}^{k}\frac{c_i}{r_1}(M_i-P_{M_iX})+\sigma_1^2\sum_{i=0}^{k}\frac{c_i\tau_i}{r_1}(M_i-P_{M_iX})\Big)\\
&\quad \cdot Z\Big(\sum_{i=1}^{g} E_i-\frac{s}{q-s}E_{g+1}\Big)Z'(\sigma^2 I+\sigma_1^2 V)\Big].
\end{aligned}$$

由引理2.1有$M_i-P_{M_iX}=M_i-(P_{(I-M_i:X)}-(I-M_i))=I-P_{(I-M_i:X)}$, 进而有$(M_i-P_{M_iX})Z=(I-P_{(I-M_i:X)})Z=(I-P_{I-M_i})Z=M_iZ$, 并且记$G=\sum\limits_{i=1}^{g}E_i-[s/(q-s)]E_{g+1}$. 因此上式等价于

$$\begin{aligned}
&= \frac{2}{\mathrm{tr}(V_1)}\mathrm{tr}\Big[\Big(\sigma^2\sum_{i=0}^{k}\frac{c_i}{r_i}M_iZGZ'+\sigma_1^2\sum_{i=0}^{k}\frac{c_i\tau_i}{r_i}M_iZGZ'\Big)(\sigma^2 I+\sigma_1^2 V)\Big]\\
&= \frac{2}{\mathrm{tr}(V_1)}\Big[\sigma_2^4\sum_{i=0}^{k}\frac{c_i}{r_i}\mathrm{tr}(M_iZGZ')+\sigma_1^4\sum_{i=0}^{k}\frac{c_i\tau_i^2}{r_i}\mathrm{tr}(M_iZGZ')\\
&\quad +2\sigma_1^2\sigma^2\sum_{i=0}^{k}\frac{c_i\tau_i}{r_i}\mathrm{tr}(M_iZGZ')\Big].
\end{aligned}$$

《应用概率统计》版权所有

同理可得

$$\begin{aligned}
\mathrm{Var}\,(\widehat{\sigma}_S^2) &= \mathrm{Cov}\,(\widehat{\sigma}_S^2,\widehat{\sigma}_S^2) = 2\mathrm{tr}\Big[\Big(\sum_{i=0}^{k}\frac{c_i}{r_1}(M_i-P_{M_iX})\Big)(\sigma^2 I+\sigma_1^2 V)\Big]^2 \\
&= 2\mathrm{tr}\Big[\Big(\sigma^2\sum_{i=0}^{k}\frac{c_i}{r_1}(M_i-P_{M_iX})+\sigma_1^2\sum_{i=0}^{k}\frac{c_i\tau_i}{r_1}(M_i-P_{M_iX})\Big)\Big]^2 \\
&= 2\Big[\sigma^4\sum_{i=0}^{k}\frac{c_i^2}{r_i^2}\mathrm{tr}(M_i-P_{M_iX})+\sigma_1^4\sum_{i=0}^{k}\frac{c_i^2\tau_i^2}{r_i^2}\mathrm{tr}(M_i-P_{M_iX}) \\
&\quad +2\sigma_1^2\sigma^2\sum_{i=0}^{k}\frac{c_i^2\tau_i}{r_i^2}\mathrm{tr}(M_i-P_{M_iX})\Big] \\
&= 2\Big[\sigma^4\sum_{i=0}^{k}\frac{c_i^2}{r_i}+\sigma_1^4\sum_{i=0}^{k}\frac{c_i^2\tau_i^2}{r_i}+2\sigma_1^2\sigma^2\sum_{i=0}^{k}\frac{c_i^2\tau_i}{r_i}\Big].
\end{aligned}$$

显然, 若取$c_i=\mathrm{tr}(M_iZGZ')/\mathrm{tr}(V_1)$, 则$\mathrm{Var}\,(\widehat{\sigma}_S^2)=\mathrm{Cov}\,(\widehat{\sigma}_S^2,\widehat{\sigma}_1^2)$.
又因为

$$\begin{aligned}
\sum_{i=1}^{k}\frac{\tau_i\mathrm{tr}(M_iZGZ')}{\mathrm{tr}(V_1)} &= \frac{\sum_{i=1}^{k}\tau_i\mathrm{tr}(M_iZGZ')}{\mathrm{tr}(V_1)} = \frac{\mathrm{tr}\Big(Z'\sum_{i=1}^{k}\tau_iM_iZG\Big)}{\mathrm{tr}(V_1)} \\
&= \frac{\mathrm{tr}V_1G}{\mathrm{tr}(V_1)} = \frac{\mathrm{tr}V_1}{\mathrm{tr}(V_1)} = 1, \\
\sum_{i=0}^{k}\frac{\mathrm{tr}(M_iZGZ')}{\mathrm{tr}(V_1)} &= \frac{\mathrm{tr}(M_0ZGZ')}{\mathrm{tr}(V_1)}+\sum_{i=1}^{k}\frac{\mathrm{tr}(M_iZGZ')}{\mathrm{tr}(V_1)} \\
&= \frac{\mathrm{tr}\Big(\Big(I-\sum_{i=1}^{k}M_i\Big)ZGZ'\Big)}{\mathrm{tr}(V_1)}+\sum_{i=1}^{k}\frac{\mathrm{tr}(M_iZGZ')}{\mathrm{tr}(V_1)} \\
&= \frac{\mathrm{tr}(ZGZ')}{\mathrm{tr}(V_1)} = \frac{\mathrm{tr}(G)}{\mathrm{tr}(V_1)} = 0.
\end{aligned}$$

所以$c_i=\mathrm{tr}(M_iZGZ')/\mathrm{tr}(V_1)$满足使得$\widehat{\sigma}_S^2$为广义谱分解估计的条件.

因为$\mathrm{Cov}\,(\widehat{\sigma}_S^2,\widehat{\sigma}_1^2)\leq\sqrt{\mathrm{Var}\,(\widehat{\sigma}_S^2)\mathrm{Var}\,(\widehat{\sigma}_1^2)}$, 所以有$\mathrm{Var}\,(\widehat{\sigma}_S^2)\leq\sqrt{\mathrm{Var}\,(\widehat{\sigma}_S^2)\mathrm{Var}\,(\widehat{\sigma}_1^2)}$, 即$\mathrm{Var}\,(\widehat{\sigma}_S^2)\leq\mathrm{Var}\,(\widehat{\sigma}_1^2)$. 又因为广义谱分解估计及ANOVA估计均为无偏估计, 因此它们的方差就是均方误差. 证毕. □

由此定理我们可以看出, 我们不但证明了广义谱分解估计可以在均方误差意义下优于ANOVA估计, 而且很好的解决了文献[13]给出的广义谱分解估计中系数c_i的选取问题.

参 考 文 献

[1] Kelly, R.J. and Mathew, T., Improved nonegative estimation of variance components in some mixed model with unbalanced data, *Journal of the Royal Statistical Society*, **55(B)**(1993), 897–911.

《应用概率统计》版权所有

[2] Hartley, H.O. and Rao, J.N.K., Maximum likelihood estimation for variance model, *Biometrika*, **54**(1967), 93–108.

[3] Patterson, H.D. and Thompson, R., Maximum likelihood estimation of components of variance, In *Proceeding of International Biometric Conference*, 1973, 197–207.

[4] Rao, C.R., Estimation of variance and covariance components-MINQUE theory, *Journal of Multivariate Analysis*, **1**(1971), 257–275.

[5] Rao, C.R., Minimum variance quadratic unbiased estimation of covariance components, *Journal of Multivariate Analysis*, **1**(1971), 445–457.

[6] Rao, C.R., Estimation of variance and covariance components in linear models, *Journal of American Statistical Association*, **67**(1972), 112–115.

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

[8] Airy, G.B., *On the Algebraical and Numerical Theory of Errors of Observations and the Combinations of Observation*, London: MacMillan, 1861.

[9] Graybill, F.A. and Hultquist, R.A., Theorems concerning Eisenhart's model Π, *Annals of Mathematical Statistics*, **32**(1961), 261–269.

[10] 吴密霞, 混合效应模型估计理论及方法, 北京工业大学博士论文, 2004.

[11] 张金槐, 线性混合模型参数估计及其改进, 武汉: 国防科技大学出版社, 1992.

[12] Searle, S.R., *Linear Models*, John Wiley, New York, 1971.

[13] 史建红, 王松桂, 方差分量的广义谱分解估计, 高校应用数学学报, **20(1)**(2005), 83–89.

The Covariances of Quadratic Forms of Random Variables and Their Applications in Linear Mixed Models

Ma Tiefeng

(*Statistics College, Southwestern University of Finance and Economics, Chengdu, 610074*)

Wang Songgui Yin Suju

(*College of Applied Sciences, Beijing University of Technology, Beijing, 100124*)

An improved ANOVA estimator is obtained in this paper. For the linear mixed model, the improved estimator can be obtained by this idea and dominates ANOVA estimator. For random effects model of one-way classification, the estimator considered by Kelly and Mathew[1] is a special case. This idea is also used to improve spectral decomposition estimator. This paper applys a simple property of variance-covariance to prove that the spectral decomposition estimator uniformly dominates ANOVA estimator for random effects model with one variance component when the covariance matrix of random effects only has one eigenvalue. Finally, this result is extended and we obtain a better feasible generalized spectral decomposition estimator.

Keywords: Variance component model, spectral decomposition estimator, ANOVA estimator, mean squared error.

AMS Subject Classification: 62F10.

《应用概率统计》版权所有