第 34卷 第 6期　福州大学学报(自然科学版)　Vol 34 No. 6
2006年 12月　Journal of Fuzhou University (Natural Science)　Dec. 2006

文章编号: 1000-2243(2006)06-0803-03

一般生长曲线模型中的简单投影预测

袁权龙

(贵州大学理学院, 贵州 贵阳　550025)

摘要: 考虑一般生长曲线模型 $Y = XBZ + \varepsilon$(其中, $E(\mathrm{Vec}(\varepsilon)) = 0$, $V(\mathrm{Vec}(\varepsilon)) = \sigma^2\Delta\otimes\Sigma$), 该模型的预测问题就是利用已观察值矩阵 Y 预测未观察值矩阵 $Y_0 = X_0BZ_0 + \varepsilon_0$. 作者研究了预测的最优性, 对任一线性可预测变量 $\theta = \mathrm{tr}(A'Y_0)$, 它的简单预测被定义为 $\hat{\theta}_{SPP} = \mathrm{Vec}'(A)(Z_0'\otimes X_0)[(Z\otimes X')T^-(Z'\otimes X)]^-(Z\otimes X')T^-\mathrm{Vec}(Y)$ (其中 $T = \Delta\otimes\Sigma + (Z'Z\otimes XX')$); 得到了 $\hat{\theta}_{SPP}$ 为 θ 的最优线性无偏预测的充要条件, 并研究了 $\hat{\theta}_{SPP}$ 关于协方差阵的稳健性, 推广了 Bolfarine H 等的有关结果.

关键词: 生长曲线模型; 简单投影预测; 稳健性; 最优预测

中图分类号: O212.4　　文献标识码: A

The simple projection predictor in the general growth curve model

YUAN Quan-long

(College of Science, Guizhou University, Guiyang Guizhou 550025, China)

Abstract: Considering the general growth curve model $Y = XBZ + \varepsilon$, where $E(\mathrm{Vec}(\varepsilon)) = 0$, $V(\mathrm{Vec}(\varepsilon)) = \sigma^2\Delta\otimes\Sigma$, the unknown observation matrix $Y_0 = X_0BZ_0 + \varepsilon_0$ is predicted using the known observation matrix Y. As to optimal predictor, on others' studies, for arbitrary linear predictable variable $\theta = \mathrm{tr}(A'Y_0)$, its SPP is then defined by $\hat{\theta}_{SPP} = \mathrm{Vec}'(A)(Z_0'\otimes X_0)[(Z\otimes X')T^-(Z'\otimes X)]^-(Z\otimes X')T^-\mathrm{Vec}(Y)$, where $T = \Delta\otimes\Sigma + (Z'Z\otimes XX')$. A number of necessary and sufficient conditions where $\hat{\theta}_{SPP}$ is also the best linear unbiased predictor are obtained, and the robustness of the $\hat{\theta}_{SPP}$ on the covariance matrix is investigated, and thus the relevant results drawn Bolfarine H *et al* are widely used.

Keywords: growth curve model; simple projection predictor; robustness; optimal predictor

考虑一般生长曲线模型[1-4]:

$$M_1: \begin{cases} Y = XBZ + \varepsilon \\ E(\mathrm{Vec}(\varepsilon)) = 0 \\ V(\mathrm{Vec}(\varepsilon)) = \sigma^2\Delta\otimes\Sigma \end{cases} \tag{1}$$

其中: Y 为 $n\times q$ 的随机观察阵, ε 为 $n\times q$ 的随机误差阵, X 和 Z 分别为 $n\times p$ 和 $k\times q$ 的已知设计阵, B 为 $p\times k$ 的未知参数阵, Δ 和 Σ 分别为已知的 q 阶和 n 阶对称非负定矩阵, σ 为未知参数且 $\sigma^2 > 0$, $\mathrm{Vec}(\varepsilon)$ 为把 ε 按列拉直得到的一列向量, $\Delta\otimes\Sigma$ 表示 Δ 与 Σ 的 Kronecker 乘积, $E(\cdot)$ 和 $V(\cdot)$ 分别表示随机向量的期望和协方差阵.

模型(1)的预测问题就是利用已观察值矩阵 Y 预测未观察值矩阵 Y_0, 这里 Y_0 满足:

收稿日期: 2005-11-28
作者简介: 袁权龙(1980-), 男, 硕士, 讲师.
基金项目: 国家自然科学基金资助项目(10101006)

$$M_0:\begin{cases} Y_0 = X_0 B Z_0 + \varepsilon_0 \\ E(\mathrm{Vec}(\varepsilon_0)) = 0 \\ V(\mathrm{Vec}(\varepsilon_0)) = \sigma^2 \Delta \otimes \Sigma_0 \\ E(\mathrm{Vec}(\varepsilon)\mathrm{Vec}'(\varepsilon_0)) = \sigma^2 \Delta \otimes V \end{cases} \tag{2}$$

其中: X_0 和 Z_0 分别为 $m\times p$ 和 $k\times q$ 的已知矩阵, Σ_0 为已知的 m 阶对称非负定矩阵, ε_0 为 $m\times q$ 的随机误差阵, V 为 $n\times m$ 的已知矩阵.

文献[5]研究了具有任意秩的多元线性模型, 定义了简单投影预测 $\hat{\theta}_{SPP}$, 得到了 $\hat{\theta}_{SPP}$ 为线性可预测变量的最优线性无偏预测的充要条件, 并研究了 $\hat{\theta}_{SPP}$ 关于协方差阵的稳健性, 而本文的研究推广了文献[5]中的有关结果.

在应用上, 通常要预测 Y_0 的线性函数 $\theta(Y_0)$. 对于线性可预测变量 $\theta=\mathrm{tr}(A'Y_0)$ (其中 A 为 $m\times q$ 的已知矩阵, $\mathrm{tr}(\cdot)$ 表示矩阵的迹), 其最优线性无偏预测为:

$$\hat{\theta}_{BLUP} = \mathrm{tr}(F'Y)^{[4]}$$

其中, F 满足 $\mathrm{Vec}(F) = (T^- - T^-(Z'\otimes X)Q^-(Z\otimes X')T^-)(\Delta\otimes V)\mathrm{Vec}(A) + T^-(Z'\otimes X)Q^-(Z_0\otimes X'_0)\mathrm{Vec}(A)$, $T=\Delta\otimes\Sigma + (Z'Z\otimes XX')$, $Q=(Z\otimes X')T^-(Z'\otimes X)$, 这里 T^- 表示 T 的任一广义逆.

Bolfarine H 等强调预测必须具有简洁而直观的形式[1], 最简洁、直观的预测是简单投影预测(SPP), 线性可预测变量 $\theta=\mathrm{tr}(A'Y_0)$, 它的简单投影预测为:

$$\hat{\theta}_{SPP} = \mathrm{Vec}'(A)(Z'_0\otimes X_0)Q^-(Z\otimes X')T^-\mathrm{Vec}(Y) \tag{3}$$

1 $\hat{\theta}_{SPP}$ 的最优预测条件

在模型(1)下考虑 $\hat{\theta}_{SPP}$ 为最优线性无偏预测的条件.

引理 1[4]　对于模型(1), $\mathrm{tr}(A'Y_0)$ 为线性可预测变量, 当且仅当

$$(Z_0\otimes X'_0)\mathrm{Vec}(A)\in\mu(Z\otimes X')$$

其中, $\mu(A)$ 表示由矩阵 A 的列向量生成的线性子空间.

引理 2　$P((\mu(Y)\subset\mu(\Sigma\vdots X))=1$.

证明　由文献[3]中的引理 4.1, 有 $P(\mathrm{Vec}(Y)\in\mu(\Delta\otimes\Sigma\vdots Z'\otimes X))=1$. 因为 $\mu(\Delta\otimes\Sigma\vdots Z'\otimes X)\subset\mu(I\otimes\Sigma\vdots I\otimes X)$, 所以, $P(\mathrm{Vec}(Y)\in\mu(I\otimes\Sigma\vdots I\otimes X))=1$. 这等价于 $P((\mu(Y)\subset(\Sigma\vdots X))=1$.

引理 2 表明, 在几乎处处意义下, $\hat{\theta}_{SPP}$ 与 T^- 的选择无关.

定理 1　对于线性模型(1), 设 $\theta=\mathrm{tr}(A'Y_0)$ 为线性可预测变量, 则:

$$P(\hat{\theta}_{SPP}=\hat{\theta}_{BLUP})=1$$

当且仅当 $\mu((\Delta\otimes V)\mathrm{Vec}(A))\subset\mu(Z'\otimes X)$.

证明　$\hat{\theta}_{BLUP}=[\mathrm{Vec}'(A)(\Delta\otimes V')(T^- - T^-(Z'\otimes X)Q^-(Z\otimes X')T^-) + \mathrm{Vec}'(A)(Z'_0\otimes X_0)Q^-(Z\otimes X')T^-]\mathrm{Vec}(Y)$. 由式(3)可得, $P(\hat{\theta}_{SPP}=\hat{\theta}_{BLUP})=1$, 当且仅当

$$\mathrm{Vec}'(A)(\Delta\otimes V')(T^- - T^-(Z'\otimes X)^-Q^-(Z\otimes X')T^-)(\Delta\otimes\Sigma)=0$$

$\mathrm{Vec}'(A)(\Delta\otimes V')[I-T^-(Z'\otimes X)^-Q^-(Z\otimes X')]=0$, 这等价于:

$$(\Delta\otimes\bar{V})\mathrm{Vec}(A) = (Z'\otimes X)Q^-(Z\otimes X')T^-(\Delta\otimes V)\mathrm{Vec}(A)$$

即:

$$\mu((\Delta\otimes V)\mathrm{Vec}(A))\subset\mu(Z'\otimes X)$$

上述定理表明, $P(\hat{\theta}_{SPP}=\hat{\theta}_{BLUP})=1$ 的充要条件只与矩阵 Σ 有关, 而与 Σ_0 无关.

推论 1　在定理 1 的假设下, 若 ε 与 ε_0 不相关, 即 $V=0$, 则 $P(\hat{\theta}_{SPP}=\hat{\theta}_{BLUP})=1$.

推论 2　在定理 1 的假设下, $\Sigma>0$ 和 $\Delta>0$, 且 $\mu((\Delta\otimes V)\mathrm{Vec}(A))\subset\mu(Z'\otimes X)$, 则 θ 的唯一最优线性无偏预测为:

$$\hat{\theta}_{BLUP} = \mathrm{Vec}'(A)(Z'_0 \otimes X_0)[(Z\Delta^{-1}Z')^{-1}Z\Delta^{-1} \otimes (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}]\mathrm{Vec}(Y)$$

2 $\hat{\theta}_{SPP}$ 关于协方差矩阵的稳健性

考虑 2个多元线性模型：

$$d_i:\ Y = XBZ + \varepsilon_i \qquad (i = 1,\ 2)$$

其中：$E(\mathrm{Vec}(\varepsilon_i)) = 0$, $V(\mathrm{Vec}(\varepsilon_1)) = \sigma^2\Delta \otimes I$, $V(\mathrm{Vec}(\varepsilon_2)) = \sigma^2\Delta \otimes \Sigma$(这里 $\Delta > 0$).

要预测 $Y_0 = X_0BZ_0 + \varepsilon_0$, $E(\mathrm{Vec}(\varepsilon_0)) = 0$, $V(\mathrm{Vec}(\varepsilon_0)) = \sigma^2\Delta \otimes \Sigma_0$, $E(\mathrm{Vec}(\varepsilon_i)\mathrm{Vec}'(\varepsilon_0)) = \sigma^2\Delta \otimes V_i$(其中 $i = 1,\ 2$)，设 $\theta = \mathrm{tr}(A'Y_0)$为线性可预测变量，$\hat{\theta}_{SPP_i}$为 θ在模型 d_i中的简单投影预测，即：

$$\hat{\theta}_{SPP_1} = \mathrm{Vec}'(A)(Z'_0 \otimes X_0)[(Z \otimes X')(\Delta \otimes I)^{-1}(Z' \otimes X)]^{-1}(Z \otimes X')(\Delta \otimes I)^{-1}\mathrm{Vec}(Y)$$

$$\hat{\theta}_{SPP_2} = \mathrm{Vec}'(A)(Z'_0 \otimes X_0)Q^{-}(Z \otimes X')T^{-}\mathrm{Vec}(Y)$$

假设模型 d_1是所选用的模型，而真正的模型为 d_2，这里研究 $\hat{\theta}_{SPP_1}$也是 θ在模型 d_2下的简单投影预测的条件.

定理 2　设 $\theta = \mathrm{tr}(A'Y_0)$为线性可预测变量，则 $P(\hat{\theta}_{SPP_1} = \hat{\theta}_{SPP_2}) = 1$，当且仅当

$$\mathrm{Vec}'(A)(Z'_0 \otimes X_0)\{[(Z \otimes X')(\Delta \otimes I)^{-1}(Z' \otimes X)]^{-1}(Z \otimes X')(\Delta \otimes I)^{-1}T - Q^{-}(Z \otimes X')\} = 0$$

证明　由 $\hat{\theta}_{SPP_i}$的表达式$(i = 1,\ 2)$可知，$P(\hat{\theta}_{SPP_1} = \hat{\theta}_{SPP_2}) = 1$，当且仅当

$$\mathrm{Vec}'(A)(Z'_0 \otimes X_0)\{[(Z \otimes X')(\Delta \otimes I)^{-1}(Z' \otimes X)]^{-1}(Z \otimes X')(\Delta \otimes I)^{-1} - Q^{-}(Z \otimes X')T^{-}\}(\Delta \otimes \Sigma) = 0 \qquad (4)$$

由于 $\theta = \mathrm{tr}(A'Y_0)$为线性可预测变量，故$(Z_0 \otimes X'_0)\mathrm{Vec}(A) \in \mu(Z \otimes X')$. 所以可化为：

$$\mathrm{Vec}'(A)(Z'_0 \otimes X_0)\{[(Z \otimes X')(\Delta \otimes I)^{-1}(Z' \otimes X)]^{-1}(Z \otimes X')(\Delta \otimes I)^{-1}T - Q^{-}(Z \otimes X')\} = 0$$

可见，该定理得证.

推论 3　若$(Z'(Z\Delta^{-1}Z')^{-1}Z\Delta^{-1} \otimes P_X)T = T(\Delta^{-1}Z'(Z\Delta^{-1}Z')^{-1}Z \otimes P_X)$，则对一切线性可预测变量 $\theta = \mathrm{tr}(A'Y_0)$，都有：

$$P(\hat{\theta}_{SPP_1} = \hat{\theta}_{SPP_2}) = 1$$

其中：$P_X = X(X'X)^{-}X'$.

参考文献：

[1] Bolfarine H, Rodrigues J. On the simple projection predictor in finite populations[J]. Aust Jour Statist, 1988, 30: 338－341.

[2] Bolfarine H, Zacks S, Elian S N, et al. Optimal prediction of the finite populations regression coefficient[J]. Sankhya: B, 1994, 56: 1－10.

[3] 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.

[4] 喻胜华. 一般生长曲线模型中的最优预测[J]. 中南大学学报：自然科学版，2004(10)：280－284.

[5] 喻胜华，何灿芝. 任意秩多元线性模型中的简单投影预测[J]. 中南工业大学学报：自然科学版，2002, 33(4)：438－440