应用概率统计　第二十八卷
第一期　2012年2月

Chinese Journal of Applied Probability
and Statistics Vol.28 No.1 Feb. 2012



# 多元超结构Berkson测量误差模型的分析 *

张赛茵　吴密霞　张忠占

(北京工业大学应用数理学院, 北京, 100124)

**摘　要**

Berkson测量误差模型在工业、农业、流行病学、经济学等领域有广泛的应用. 本文考虑了一类多元超结构Berkson测量误差模型, 给出了该模型中参数的相合估计, 推导了估计的渐近分布, 并把该方法应用到一元超结构Berkson测量误差模型中, 最后给出了模拟结果和实例分析来说明本文所提出的估计方法的表现.

**关键词**: 渐近性质, 相合估计, Berkson测量误差, 多元回归.

**学科分类号**: O212.4.

## §1. 引　　言

众所周知, 在EV(error-in-variable)模型中协变量的真实值与其测量误差独立, 那是因为假定误差发生在“试验之后”. 然而若假定误差发生在“试验之前”, 则协变量的观测值与测量误差独立, 即Berkson[1]提出的Berkson测量误差模型. 关于EV模型的研究, 目前已有很多突出的成果, 如崔恒建[2]讨论了线性EV模型的经验似然推断问题, 并通过Bartlett纠偏提高了置信域覆盖概率的精度. 王启华[3]研究了基于核实数据的部分线性EV模型, 构造了未知参数和未知函数的估计, 同时证明了估计的强相合性, 未知参数估计的渐近正态性, 并给出了估计的收敛速度. 薛留根[4]讨论了非线性EV模型中参数的估计问题, 给出了估计的渐近正态性和误差方差估计的收敛速度等等. 近年来, 许多学者将EV模型中的协变量独立同分布放宽, 研究协变量期望不相等的情形, 即超结构的EV模型, 如Patriota[5]考虑了带有方程误差的多元超结构EV模型, 给出了参数的相合估计及其渐近分布. Arellano-valle[6]给出了超结构椭球模型参数的相合估计及该估计的渐近性质. 关于Berkson测量误差的相关研究问题, 目前文献中也取得了很大的成果, 如Fuller[7]用最小二乘的方法讨论了一元线性Berkson测量误差模型的估计问题. Wang[8]考虑了协变量带有Berkson测量误差的非线性回归模型, 基于响应变量的前两阶条件矩给出了未知参数的最小距离估计, 且在一般正则条件下证明了该估计的相合性和渐近正态性. 刘强[9]考虑了协变量带有Berkson测量误差的非线性半参数模型, 采用核估计和最小距离方法给出了未知参数和未知函数的估计,

*北京市教委科技计划项目(JC006790201001)、北京市自然科学基金(1102010)、国家自然科学基金(10971007, 10801005)和北京工业大学第八届研究生科技基金(ykj-2010-3816)资助.

本文2011年3月24日收到, 2011年6月27日收到修改稿.

证明了未知参数估计的相合性和渐近正态性, 但是关于超结构的Berkson测量误差模型的研究目前还很少.

Berkson测量误差模型在医药、农业、经济、工业等方面有广泛的应用, 比如流行病学家研究一种肺病$Y$与某些空气污染物$X$的关系, $X$可以在城市中的某些污染物观测站测量, 但居民受到的实际污染量$x$是不可观测的; 再比如$X$是额定除草剂的数量, 但植物实际吸收量$x$是不可观测的. 上面的例子说明不可观测的$x$可以通过$X$加上一个与$X$独立的随机误差来代替. 如在肺病研究的例子中, 不同的观测站由于地理位置及周边环境不同, 其测量值$X$可以看成是来自不同期望分布的随机变量, 此时就是一个超结构Berkson测量误差模型.

本文将考虑一类多元超结构Berkson测量误差模型, 给出了该模型中参数的相合估计, 推导了估计的渐近分布, 并把该方法应用到一元超结构Berkson测量误差模型中. 其中第二部分是模型简介, 第三部分是本文的主要结果, 给出了参数的相合估计及参数估计的联合渐近分布. 第四部分把该方法应用到了一元超结构Berkson测量误差模型中, 并给出了一个实际例子. 第五部分是模拟研究, 模拟结果表明忽略模型的超结构特征会影响参数的区间估计.

## §2. 模型简介

本文考虑多元超结构Berkson测量误差模型:

$$\begin{cases} Y_i = \mathbf{a} + B\mathbf{x}_i + \mathbf{e}_i; \\ X_i = \mathbf{x}_i + \mathbf{u}_i, \end{cases} \tag{2.1}$$

$i = 1, \ldots, n$, 这里$\mathbf{x}_i$是$p_2 \times 1$的真实协变量向量, $\mathbf{a}$是$p_1 \times 1$的截距参数向量, $B$是$p_1 \times p_2$的斜率参数矩阵, $X_i$和$Y_i$是协变量和响应变量带有误差的观测值, 其测量误差分别为$\mathbf{u}_i$和$\mathbf{e}_i$, 假设向量$\mathbf{r}_i = (\mathbf{e}_i^T, \mathbf{u}_i^T, (X_i - \xi_i)^T)^T$, $i = 1, 2, \ldots, n$, 独立同对称分布(比如正态分布), 且$X_i, \mathbf{e}_i, \mathbf{u}_i$相互独立, 有

$$\begin{aligned} &\mathsf{E}(X_i) = \xi_i, &&\mathsf{E}(\mathbf{e}_i) = \mathbf{0}, &&\mathsf{E}(\mathbf{u}_i) = \mathbf{0}, \\ &\mathsf{Cov}\,(X_i - \xi_i) = \Sigma_X, &&\mathsf{Cov}\,(\mathbf{e}_i) = \Sigma_{\mathbf{e}}, &&\mathsf{Cov}\,(\mathbf{u}_i) = \Sigma_{\mathbf{u}}. \end{aligned}$$

记$Z_i = (Y_i^T, X_i^T)^T$, $i = 1, \ldots, n$, 则

$$\mathsf{E}(Z_i) = \mu_i = \begin{pmatrix} \mathbf{a} + B\xi_i \\ \xi_i \end{pmatrix}, \qquad \mathsf{Cov}\,(Z_i) = \Sigma = \begin{pmatrix} B\Gamma_{X+\mathbf{u}}B^T + \Sigma_{\mathbf{e}} & B\Sigma_X \\ \Sigma_X B^T & \Sigma_X \end{pmatrix}, \tag{2.2}$$

其中$\Gamma_{a+b} = \Sigma_a + \Sigma_b$.

若$\Sigma_X = \mathbf{0}$ ($\mathbf{0}$是相应维数的零矩阵), 则模型(2.1)退化为函数型Berkson测量误差模型; 若$\xi_1 = \cdots = \xi_n = \xi$, 则模型变为典型的多元结构型Berkson测量误差模型; 若假设$\mathbf{x}_i$与$\mathbf{u}_i$相互独立, 而不是$X_i$与$\mathbf{u}_i$相互独立, 则模型(2.1)变为超结构EV模型.

本文讨论超结构Berkson测量误差模型(2.1). 为了模型(2.1)的可识别性, 假定$\Sigma_{\mathbf{u}}$已知. 在导出本文的主要结果之前, 我们做如下假定:

(A1) 存在一个$p_2 \times 1$的向量$\xi$和一个满足$\Sigma_X + \Sigma_\xi$为正定的$p_2 \times p_2$的矩阵$\Sigma_\xi$, 使得当$n$趋于无穷时, 有

$$\bar{\xi} \to \xi, \qquad S_\xi \to \Sigma_\xi,$$

其中

$$\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n}\xi_i \qquad \text{和} \qquad S_\xi = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - \bar{\xi})(\xi_i - \bar{\xi})^T;$$

(A2) $\mathsf{Cov}\,[\mathrm{Vech}(\mathbf{r}_1\mathbf{r}_1^T)] < \infty$.

这里, $\mathrm{Vech}(C)$为对称矩阵$C_{p\times p}$的上三角元素按列拉直所得的$[p(p+1)/2]\times 1$的向量. Magnus[10]证明了总存在唯一的列满秩矩阵$D$, 使得$\mathrm{Vec}(C) = D\cdot\mathrm{Vech}(C)$或者$\mathrm{Vech}(C) = D^+\cdot\mathrm{Vec}(C)$, 其中$D^+ = (D^TD)^{-1}D^T$是$D$的Moore-Penrose广义逆, 并且$D$与$C$的维数有关而与其元素无关.

## §3.　主要结果

记观测向量$Z_1, \ldots, Z_n$的样本均值和样本协方差矩阵分别为

$$\overline{Z} = \begin{pmatrix} \overline{Y} \\ \overline{X} \end{pmatrix}, \qquad S_Z = \begin{pmatrix} S_Y & S_{YX} \\ S_{XY} & S_X \end{pmatrix},$$

其中,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n}X_i, \qquad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n}Y_i,$$
$$S_X = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})X_i^T, \qquad S_Y = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})Y_i^T,$$
$$S_{XY} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})Y_i^T, \qquad S_{YX} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})X_i^T.$$

**引理 3.1**　记$Z_i = (Y_i^T, X_i^T)^T$, $i = 1, \ldots, n$是模型(2.1)的$n$个观测向量, 则有

(i) 在(A1)的假定下, 当$n \to \infty$时, 有

$$\overline{Z} = \begin{pmatrix} \overline{Y} \\ \overline{X} \end{pmatrix} \xrightarrow{\text{a.s.}} \mu = \begin{pmatrix} \mathbf{a} + B\xi \\ \xi \end{pmatrix},$$

$$S_Z = \begin{pmatrix} S_Y & S_{YX} \\ S_{XY} & S_X \end{pmatrix} \xrightarrow{\text{a.s.}} \Sigma + \Sigma_\mu = \begin{pmatrix} B\Gamma_{X+\mathbf{u}+\xi}B^T + \Sigma_{\mathbf{e}} & B\Gamma_{X+\xi} \\ \Gamma_{X+\xi}B^T & \Gamma_{X+\xi} \end{pmatrix},$$

其中,

$$\Sigma_\mu = \begin{pmatrix} B\Sigma_\xi B^T & B\Sigma_\xi \\ \Sigma_\xi B^T & \Sigma_\xi \end{pmatrix};$$

(ii) 在(A1)–(A2)的假定下, 当$n\to\infty$时有

$$\sqrt{n}(\overline{Z}-\mu) \xrightarrow{d} \mathcal{N}_p(\mathbf{0},\Sigma),$$
$$\sqrt{n}\cdot\mathrm{Vech}(S_Z-\Sigma-\Sigma_\mu) \xrightarrow{d} \mathcal{N}_{p(p+1)/2}(\mathbf{0},\Lambda+\Lambda_\mu),$$

且$\overline{Z}$与$S_Z$是渐近独立的, 这里, $p=p_1+p_2$, "$\xrightarrow{d}$"表示依分布收敛,

$$\Lambda_{\mathbf{r}} = \mathsf{Cov}\,[\mathrm{Vec}(\mathbf{r}_1\mathbf{r}_1^T)], \qquad \Lambda = D^+(A\otimes A)\Lambda_{\mathbf{r}}(A\otimes A)^T D^{+T},$$
$$\Lambda_\mu = D^+[(\Sigma\otimes\Sigma_\mu)+(\Sigma_\mu\otimes\Sigma)+(\Sigma\star\Sigma_\mu)+(\Sigma_\mu\star\Sigma)]D^{+T},$$
$$A = \begin{pmatrix} I_{p_1} & -B & B \\ 0 & 0 & I_{p_2} \end{pmatrix},$$

$\otimes$为Kronecher乘积, 定义"$\star$"为$\Sigma_\mu\star\Sigma=[\Sigma_\mu\otimes\Sigma_{\cdot 1},\ldots,\Sigma_\mu\otimes\Sigma_{\cdot p}]$, $\Sigma_{\cdot i}$表示$\Sigma$的第$i$列元素构成的向量, $D$的定义请参见(A2)后的说明.

**定理 3.1** 在模型(2.1)下, 若条件(A1)成立, 则如下估计分别为$\mathbf{a},B,\xi$的相合估计:

$$\widehat{\mathbf{a}} = \overline{Y}-\widehat{B}\overline{X}, \qquad \widehat{B}=S_{YX}S_X^{-1}, \qquad \widehat{\xi}=\overline{X}.$$

由于上述参数估计是样本均值$\overline{Z}$和协方差矩阵$S_Z$的连续函数, 则由引理3.1 (i)可知, 它们都是相合估计.

**注记 1** $\widehat{\mathbf{a}},\widehat{B}$和$\widehat{\xi}$也是在典型的多元结构型Berkson测量误差模型下参数$\mathbf{a},B,\xi$的一个相合估计, 如Fuller (1987). 但它们的渐近方差在这两个不同的模型假设下是不等的.

**定理 3.2** 在模型(2.1)下, 如果假定(A1), (A2)都成立, 则

$$\begin{pmatrix} \sqrt{n}(\widehat{\mathbf{a}}-\mathbf{a}) \\ \sqrt{n}\cdot\mathrm{Vec}(\widehat{B}-B) \end{pmatrix} \xrightarrow{d} \mathcal{N}_{p_1+p_1p_2}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\Phi_{\mathbf{a}} & \Phi_{\mathbf{a}B} \\ \Phi_{B\mathbf{a}} & \Phi_B\end{pmatrix}\right), \tag{3.1}$$

这里,

$$\Phi_{\mathbf{a}} = P\Sigma P^T + (\xi^T\otimes I_{p_1})\Phi_B(\xi\otimes I_{p_1}),$$
$$\Phi_{\mathbf{a}B} = \Phi_{B\mathbf{a}}^T = -(\xi^T\otimes I_{p_1})\Phi_B, \qquad \Phi_B = QD(\Lambda+\Lambda_\mu)D^TQ^T,$$

其中,

$$Q = (\Gamma_{X+\xi}^{-1}\otimes I_{p_1})(H_2\otimes H_1) - (\Gamma_{X+\xi}^{-1}\otimes B)(H_2\otimes H_2),$$
$$H_1 = (I_{p_1},\mathbf{0}), \qquad H_2=(\mathbf{0},I_{p_2}), \qquad P=(I_{p_1},-B).$$

令定理3.2中的$\xi_1=\xi_2=\cdots=\xi_n=\xi$, 即$\Sigma_\xi=0$且$\Gamma_{X+\xi}=\Sigma_X$, $\Lambda_\mu=0$, 便得到典型的多元结构型Berkson测量误差模型下该估计的渐近分布. 下面我们在一元情形下给出超结构对Berkson模型带来的影响.

## §4.　模型简单应用

考虑模型(2.1)中$p_1=p_2=1$的情形, 即一元超结构Berkson测量误差模型

$$\begin{cases} Y_i=a+bx_i+e_i; \\ X_i=x_i+u_i, \end{cases} \tag{4.1}$$

$i=1,\ldots,n$. 记$\mathsf{Var}\,(X_i-\xi_i)=\sigma_X^2$, $\mathsf{Var}\,(e_i)=\sigma_e^2$, $\mathsf{Var}\,(u_i)=\sigma_u^2$. 这时(A1)假设中的$\Sigma_\xi$也是一个未知常数, 记作$\sigma_\xi^2$.

由定理3.1和定理3.2得到模型参数$a$和$b$的如下估计及其渐近分布:

$$\widehat{a}=\overline{Y}-\widehat{b}\overline{X},\qquad \widehat{b}=\frac{S_{YX}}{S_X},\qquad \widehat{\xi}=\overline{X}.$$

**推论 4.1**　在简化的(A1), (A2)条件下, 有

$$\begin{pmatrix}\sqrt{n}(\widehat{a}-a)\\ \sqrt{n}(\widehat{b}-b)\end{pmatrix}\xrightarrow{d}\mathcal{N}_2\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\Phi_a & \Phi_{ab}\\ \Phi_{ba} & \Phi_b\end{pmatrix}\right).$$

其中,

$$\Phi_b=\frac{b^2\sigma_u^2+\sigma_e^2}{\sigma_X^2+\sigma_\xi^2}, \tag{4.2}$$

$$\Phi_{ab}=\Phi_{ba}^T=-\frac{\xi(b^2\sigma_u^2+\sigma_e^2)}{\sigma_X^2+\sigma_\xi^2}, \tag{4.3}$$

$$\Phi_a=\frac{(\sigma_X^2+\sigma_\xi^2+\xi^2)(b^2\sigma_u^2+\sigma_e^2)}{\sigma_X^2+\sigma_\xi^2}. \tag{4.4}$$

若忽略一元超结构Berkson测量误差模型中的超结构特征, 即认为$\sigma_\xi^2=0$, 直接套用经典Berkson测量误差模型的渐近结果, 计算的$\widehat{b}$渐近方差$\widetilde{\Phi}_b$偏大, 这是因为经典Berkson测量误差模型下,

$$\widetilde{\Phi}_b=\frac{b^2\sigma_u^2+\sigma_e^2}{\sigma_X^2}>\frac{b^2\sigma_u^2+\sigma_e^2}{\sigma_X^2+\sigma_\xi^2}=\Phi_b.$$

因此在分析此类数据时, 要先考查一下数据的来源, 即自变量是否是独立同分布的, 从而决定是否采用超结构Berkson测量误差模型, 若是超结构模型其超结构的特性不能忽略.

下面用Fuller (1987)中的一个实际例子来解释本文的结论.

**例 1** 假定农作物叶子中的磷含量$Y$与土壤中的磷元素$x$满足模型(4.1), 其中$(X_i, e_i, u_i)^T \sim \mathcal{N}_3\big((\xi_i, 0, 0)^T, \mathrm{diag}(\sigma_X^2, \sigma_e^2, 0.25)\big)$. $X$是土壤中磷元素的观测值. 18组观测值如表1所示.

表1 磷元素的18组观测值

| | | | | | | |
|---|---|---|---|---|---|---|
| $X$ | 1.18 | 1.18 | 2.02 | 1.26 | 2.39 | 1.64 |
| $Y$ | 64 | 60 | 71 | 61 | 54 | 77 |
| $X$ | 3.22 | 3.33 | 3.55 | 3.69 | 3.45 | 4.91 |
| $Y$ | 81 | 93 | 93 | 51 | 76 | 96 |
| $X$ | 4.91 | 4.75 | 4.91 | 1.70 | 5.27 | 5.56 |
| $Y$ | 77 | 93 | 95 | 54 | 68 | 99 |

这是一个Berkson测量误差模型. 由于农作物种类不同, 故其所需土壤中含磷量的平均水平不一定相同, 或由于环境因素导致土壤中含磷量平均水平不同, 即每组数据中$X$的期望不一定相同. 假定$\xi_i$已知(根据历史数据), 分别为

$$\begin{array}{llllll}
\xi_1 = 1.2, & \xi_2 = 1.2, & \xi_3 = 2.0, & \xi_4 = 1.2, & \xi_5 = 2.4, & \xi_6 = 1.6, \\
\xi_7 = 3.2, & \xi_8 = 3.2, & \xi_9 = 3.6, & \xi_{10} = 3.6, & \xi_{11} = 3.6, & \xi_{12} = 4.8, \\
\xi_{13} = 4.8, & \xi_{14} = 4.8, & \xi_{15} = 4.8, & \xi_{16} = 1.6, & \xi_{17} = 5.2, & \xi_{18} = 5.6.
\end{array}$$

由定理3.1得参数的估计为

$$\widehat{a} = \overline{Y} - \widehat{b}\overline{X} = 53.67, \qquad \widehat{b} = \frac{S_{YX}}{S_X} = 6.74.$$

由于推论4.1 $\Phi_a$、$\Phi_b$中$\sigma_\xi, \sigma_X$和$\sigma_e$未知, 我们用它们的经验估计来代替, 计算得

$$\widehat{\sigma}_\xi = \sqrt{\frac{1}{n-1}\sum_{i=1}^n (\xi_i - \overline{\xi})^2} = 1.53,$$

$$\widehat{\sigma}_X = \sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i - \xi_i)^2} = 0.08,$$

$$\widehat{\sigma}_e = \sqrt{\frac{1}{n-2}\sum_{i=1}^n (Y_i - \widehat{a} - \widehat{b}X_i)^2 - \widehat{b}^2\sigma_u^2} = 12.47.$$

故参数估计$\widehat{a}$和$\widehat{b}$渐近标准差的估计分别为

$$\sqrt{\widehat{\Phi_a}} = 7.15, \qquad \sqrt{\widehat{\Phi_b}} = 1.99.$$

从而得出参数$a$的95%渐近置信区间为$(41.94, 65.40)$, $b$的95%渐近置信区间为$(3.48, 10.00)$.

若忽视超结构特征, 认为$\sigma_\xi^2=0$, 直接套用经典Berkson测量误差模型的渐近结果, 得$\widehat{a}$和$\widehat{b}$渐近标准差的估计为

$$\widehat{\sqrt{\widetilde{\Phi}_a}}=121.71, \qquad \widehat{\sqrt{\widetilde{\Phi}_b}}=37.50.$$

此时$b$的95%渐近置信区间为$(-54.76,68.24)$. 由于0包含在该置信区间内, 表明$x$对$Y$的影响不显著, 这与前面的结论显然不同.

由此可以看出在做统计推断时, 忽略其超结构的特征直接套用经典Berkson测量误差模型的渐近结果, 往往会降低统计推断的精度.

## §5. 模拟研究

为了说明模型(4.1)的结论, 我们给出一个模拟的结果. 假定模型为

$$\begin{cases} Y_i=50+7x_i+e_i; \\ X_i=x_i+u_i, \end{cases} \tag{5.1}$$

$i=1,\ldots,100$. 这里,

$$X_i-\xi_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,0.0025), \qquad u_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,0.25), \qquad e_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1),$$

其中,

$$\begin{aligned} & (\xi_{1+k\times10},\xi_{2+k\times10},\xi_{3+k\times10},\xi_{4+k\times10},\xi_{5+k\times10},\xi_{6+k\times10},\xi_{7+k\times10}, \\ & \xi_{8+k\times10},\xi_{9+k\times10},\xi_{10+k\times10}) \\ = & (1.2,1.6,2.0,2.4,2.8,3.2,3.6,4.0,4.4,4.8), \qquad k=0,1,\ldots,9. \end{aligned}$$

重复5000次模拟, 得结果如表2所示, 其中平均值是5000次估计值的平均, Var是由参数的估计值计算的样本方差, $\Phi$是指用真实值计算(4.2)式和(4.4)式得到的值, $\widetilde{\Phi}$指忽略超结构令$\sigma_\xi^2=0$时计算(4.2)式和(4.4)式得到的值. 模拟得到的各参数方差与本文在超结构下推断的相应参数估计的渐近方差很接近, 但与忽略超结构直接应用Berkson测量误差模型下推断的渐近方差相差很大. 因此, 如果超结构存在, 我们处理数据时不能轻易忽视.

表2　重复5000次的模拟结果

| | 真值 | 平均值 | Var | $\Phi$ | $\widetilde{\Phi}$ |
|---|---|---|---|---|---|
| $a$ | 50 | 50.01 | 1.02 | 1.03 | 477.13 |
| $b$ | 7 | 7.00 | 0.10 | 0.10 | 53.00 |
| $\sigma^2$ | 1 | 1.01 | | | |

## §6. 定理的证明

**引理3.1的证明:**

(i) 由模型(2.1)的假定可知

$$\mathsf{E}(\mathbf{r}_i)=0,\qquad \mathsf{Cov}\,(\mathbf{r}_i)=\Sigma_{\mathbf{r}}=\mathrm{diag}(\Sigma_{\mathbf{e}},\Sigma_{\mathbf{u}},\Sigma_X),$$

且$\epsilon_i=Z_i-\mu_i$, $i=1,\dots,n$是独立同分布的, 其均值是$\mathbf{0}$, 协方差矩阵$\Sigma$, 其中$\mu_i$和$\Sigma$已在(2.2)式中给出. 另外由(A1)有当$n\to\infty$时, $\overline{\xi}\to\xi$, $S_\xi\to\Sigma_\xi$. 因此, 对于$\overline{\mu}=(1/n)\sum\limits_{i=1}^{n}\mu_i$和$S_\mu=(1/n)\sum\limits_{i=1}^{n}(\mu_i-\overline{\mu})(\mu_i-\overline{\mu})^T$有

$$\overline{\mu}\xrightarrow{\text{a.s.}}\mu=\begin{pmatrix}\mathbf{a}+B\xi\\ \xi\end{pmatrix},\qquad S_\mu\xrightarrow{\text{a.s.}}\Sigma_\mu=\begin{pmatrix}B\Sigma_\xi B^T & B\Sigma_\xi\\ \Sigma_\xi B^T & \Sigma_\xi\end{pmatrix}.$$

由于$\epsilon_i=Z_i-\mu_i=A\mathbf{r}_i$, 其中$A=\begin{pmatrix}I_{p_1} & -B & B\\ 0 & 0 & I_{p_2}\end{pmatrix}$, $I_p$是$p\times p$单位矩阵, 故

$$\overline{Z}=\overline{\epsilon}+\overline{\mu}=\frac{1}{n}\sum_{i=1}^{n}A\mathbf{r}_i+\overline{\mu}\xrightarrow{\text{a.s.}}\mu,$$

$$\begin{aligned}S_Z&=\frac{1}{n}\sum_{i=1}^{n}(Z_i-\overline{Z})(Z_i-\overline{Z})^T\\&=\frac{1}{n}\sum_{i=1}^{n}(\epsilon_i+\mu_i-\overline{\epsilon}-\overline{\mu})(\epsilon_i+\mu_i-\overline{\epsilon}-\overline{\mu})^T\\&=\frac{1}{n}\sum_{i=1}^{n}(\epsilon_i-\overline{\epsilon})(\epsilon_i-\overline{\epsilon})^T+\frac{1}{n}\sum_{i=1}^{n}(\mu_i-\overline{\mu})(\mu_i-\overline{\mu})^T\\&\quad+\frac{1}{n}\sum_{i=1}^{n}(\epsilon_i-\overline{\epsilon})(\mu_i-\overline{\mu})^T+\frac{1}{n}\sum_{i=1}^{n}(\mu_i-\overline{\mu})(\epsilon_i-\overline{\epsilon})^T.\end{aligned}$$

易知

$$\frac{1}{n}\sum_{i=1}^{n}(\mu_i-\overline{\mu})(\epsilon_i-\overline{\epsilon})^T=\frac{1}{n}\sum_{i=1}^{n}(\mu_i-\overline{\mu})\epsilon_i^T\xrightarrow{\text{a.s.}}\mathbf{0}.$$

综上, 得

$$S_Z\xrightarrow{\text{a.s.}}\Sigma+\Sigma_\mu.$$

(ii) 由中心极限定理, 有

$$\sqrt{n}\overline{\epsilon}\xrightarrow{d}\mathcal{N}_p(\mathbf{0},\Sigma).$$

记$\omega_i=\epsilon_i\epsilon_i^T+\epsilon_i(\mu_i-\mu)^T+(\mu_i-\mu)\epsilon_i^T$, 则

$$\begin{aligned}\overline{\omega}-\Sigma&=\frac{1}{n}\sum_{i=1}^{n}\epsilon_i\epsilon_i^T+\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(\mu_i-\mu)^T+\frac{1}{n}\sum_{i=1}^{n}(\mu_i-\mu)\epsilon_i^T-\Sigma\\&=\Big[\frac{1}{n}\sum_{i=1}^{n}\epsilon_i\epsilon_i^T-\Sigma\Big]+\Big[\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(\mu_i-\mu)^T+\frac{1}{n}\sum_{i=1}^{n}(\mu_i-\mu)\epsilon_i^T\Big].\end{aligned}$$

由于$\mathsf{E}(\varpi-\Sigma)=\mathsf{E}[\epsilon_i\epsilon_i^T+\epsilon_i(\mu_i-\mu)^T+(\mu_i-\mu)\epsilon_i^T-\Sigma]=0$, 且

$$\begin{aligned}
&\mathsf{Cov}\,[\mathrm{Vech}(\omega_i-\Sigma)]\\
=\;&\mathsf{Cov}\,[\mathrm{Vech}(\epsilon_i\epsilon_i^T+\epsilon_i(\mu_i-\mu)^T+(\mu_i-\mu)\epsilon_i^T-\Sigma)]\\
=\;&\Lambda+D^+\mathsf{Cov}\,[\epsilon_i\otimes(\mu_i-\mu)]D^{+T}+D^+\mathsf{Cov}\,[(\mu_i-\mu)\otimes\epsilon_i]D^{+T}\\
&+D^+\mathsf{Cov}\,[\epsilon_i\otimes(\mu_i-\mu),(\mu_i-\mu)\otimes\epsilon_i]D^{+T}+D^+\mathsf{Cov}\,[(\mu_i-\mu)\otimes\epsilon_i,\epsilon_i\otimes(\mu_i-\mu)]D^{+T}\\
=\;&\Lambda+D^+(\Sigma\otimes((\mu_i-\mu)(\mu_i-\mu)^T))D^{+T}+D^+(((\mu_i-\mu)(\mu_i-\mu)^T)\otimes\Sigma)D^{+T}\\
&+D^+(\Sigma\star((\mu_i-\mu)(\mu_i-\mu)^T))D^{+T}+D^+(((\mu_i-\mu)(\mu_i-\mu)^T)\star\Sigma)D^{+T}.
\end{aligned}$$

根据假定条件中(A2) $\Lambda_{\mathbf{r}}=\mathsf{Cov}\,[\mathrm{Vec}(\mathbf{r}_1\mathbf{r}_1^T)]<\infty$, 有

$$\Lambda=\mathsf{Cov}\,[\mathrm{Vech}(\epsilon_1\epsilon_1^T)]=\mathsf{Cov}\,[\mathrm{Vech}(A\mathbf{r}_1\mathbf{r}_1^TA^T)]=D^+(A\otimes A)\Lambda_{\mathbf{r}}(A\otimes A)^TD^{+T}<\infty,$$

其中$D$是$p^2\times p(p+1)/2$维相应的变换矩阵, $D^+=(D^TD)^{-1}D^T$. 由中心极限定理, 有

$$\sqrt{n}\cdot\mathrm{Vech}(\varpi-\Sigma)\xrightarrow{d}\mathcal{N}_{p(p+1)/2}(\mathbf{0},\Lambda+\Lambda_\mu).$$

由于$\{\sqrt{n}(\overline{Z}-\mu),\sqrt{n}\cdot\mathrm{Vech}(S_Z-\Sigma-\Sigma_\mu)\}$与$\{\sqrt{n}\bar{\epsilon},\sqrt{n}\cdot\mathrm{Vech}(\varpi-\Sigma)\}$是渐近同分布的, 所以

$$\begin{aligned}
&\sqrt{n}(\overline{Z}-\mu)\xrightarrow{d}\mathcal{N}_p(\mathbf{0},\Sigma),\\
&\sqrt{n}\cdot\mathrm{Vech}(S_Z-\Sigma-\Sigma_\mu)\xrightarrow{d}\mathcal{N}_{p(p+1)/2}(\mathbf{0},\Lambda+\Lambda_\mu).
\end{aligned}$$

又由于

$$\begin{aligned}
&\mathsf{Cov}\,[\bar{\epsilon},\mathrm{Vech}(\varpi-\Sigma)]\\
=\;&\mathsf{Cov}\left[\frac{1}{n}\sum_{i=1}^n\epsilon_i,\mathrm{Vech}\Big(\frac{1}{n}\sum_{i=1}^n\epsilon_i\epsilon_i^T+\frac{1}{n}\sum_{i=1}^n\epsilon_i(\mu_i-\mu)^T+\frac{1}{n}\sum_{i=1}^n(\mu_i-\mu)\epsilon_i^T\Big)\right]\\
=\;&\frac{1}{n^2}\sum_{i=1}^n\mathsf{Cov}\,[\epsilon_i,\mathrm{Vech}(\epsilon_i\epsilon_i^T+\epsilon_i(\mu_i-\mu)^T+(\mu_i-\mu)\epsilon_i^T)]\\
=\;&\frac{1}{n^2}\sum_{i=1}^n\mathsf{Cov}\,[\epsilon_i,\mathrm{Vech}(\epsilon_i\epsilon_i^T)]+\frac{1}{n^2}\sum_{i=1}^n\mathsf{Cov}\,[\epsilon_i,\mathrm{Vech}(\epsilon_i(\mu_i-\mu)^T)]\\
&+\frac{1}{n^2}\sum_{i=1}^n\mathsf{Cov}\,[\epsilon_i,\mathrm{Vech}((\mu_i-\mu)\epsilon_i^T)],
\end{aligned}$$

且$\mathbf{r}_i$的分布是对称的, $\mathsf{E}[\epsilon_i\mathrm{Vech}(\epsilon_i\epsilon_i^T)^T]=\mathbf{0}$, 再结合假定(A1), (A2), 有

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^n\mathsf{Cov}\,[\epsilon_i,\mathrm{Vech}(\epsilon_i\epsilon_i^T)]\\
=\;&\frac{1}{n^2}\sum_{i=1}^n\mathsf{E}[\epsilon_i\mathrm{Vech}(\epsilon_i\epsilon_i^T)^T]-\frac{1}{n^2}\sum_{i=1}^n\mathsf{E}(\epsilon_i)\mathsf{E}[\mathrm{Vech}(\epsilon_i\epsilon_i^T)^T]\\
=\;&\mathbf{0},
\end{aligned}\tag{6.1}$$

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{Cov}\,[\epsilon_i,\mathrm{Vech}(\epsilon_i(\mu_i-\mu)^T)]\\
=\;&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}[\epsilon_i(\mathrm{Vech}(\epsilon_i(\mu_1-\mu)^T))^T]-\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}(\epsilon_i)\mathsf{E}[(\mathrm{Vech}(\epsilon_i(\mu_i-\mu)^T))^T]\\
=\;&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}[\epsilon_i((\mu_i-\mu)^T\otimes\epsilon_i^T)D^{+T}]-\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}(\epsilon_i)\mathsf{E}[((\mu_i-\mu)^T\otimes\epsilon_i^T)D^{+T}]\\
=\;&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}[(\mu_i-\mu)^T\otimes(\epsilon_i\epsilon_i^T)D^{+T}]\\
=\;&\frac{1}{n^2}\sum_{i=1}^{n}(\mu_i-\mu)^T\otimes\mathsf{E}(\epsilon_1\epsilon_1^T)D^{+T}\\
\to\;&\mathbf{0},
\end{aligned}\tag{6.2}$$

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{Cov}\,[\epsilon_i,\mathrm{Vech}((\mu_i-\mu)\epsilon_i^T)]\\
=\;&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}[\epsilon_i(\mathrm{Vech}((\mu_i-\mu)\epsilon_i^T))^T]-\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}(\epsilon_i)\mathsf{E}[(\mathrm{Vech}((\mu_i-\mu)\epsilon_i^T))^T]\\
=\;&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}[\epsilon_i(\epsilon_i^T\otimes(\mu_i-\mu)^T)D^{+T}]-\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}(\epsilon_i)\mathsf{E}[(\epsilon_i^T\otimes(\mu_i-\mu)^T)D^{+T}]\\
=\;&\frac{1}{n^2}\sum_{i=1}^{n}\mathsf{E}[(\epsilon_i\epsilon_i^T)\otimes(\mu_i-\mu)^TD^{+T}]\\
=\;&\mathsf{E}(\epsilon_1\epsilon_1^T)\otimes\Big[\frac{1}{n^2}\sum_{i=1}^{n}(\mu_i-\mu)^TD^{+T}\Big]\\
\to\;&\mathbf{0}.
\end{aligned}\tag{6.3}$$

由(6.1)–(6.3)式, 有

$$\mathsf{Cov}\,(\overline{\epsilon},\mathrm{Vech}(\overline{\varpi}-\Sigma))\xrightarrow{\text{a.s.}}\mathbf{0},$$

故$\sqrt{n}\overline{\epsilon}$与$\sqrt{n}\cdot\mathrm{Vech}(\overline{\varpi}-\Sigma)$是渐近独立, 则有$\overline{Z}$与$S_Z$是渐近独立的. 证毕. □

**定理3.2的证明:**

首先, 记$S_{YX}=H_1S_ZH_2^T$, $S_X=H_2S_ZH_2^T$, 其中$H_1=(I_{p_1},0)$, $H_2=(0,I_{p_2})$, 则有$\widehat{B}=\widehat{B}(S_Z)=H_1S_ZH_2^T(H_2S_ZH_2^T)^{-1}$是样本协方差矩阵$S_Z$的函数. 定义$s_k$为$\mathrm{Vech}(S_Z)$的第$k$个元素. 通过矩阵微分的方法, 如Magnus (2007), 有

$$\frac{\partial\widehat{B}}{\partial s_k}=H_1\Big(\frac{\partial S_Z}{\partial s_k}\Big)H_2^T(S_X)^{-1}-S_{YX}(S_X)^{-1}H_2\Big(\frac{\partial S_Z}{\partial s_k}\Big)H_2^T(S_X)^{-1},$$

则

$$\frac{\partial\mathrm{Vec}(\widehat{B})}{\partial s_k}=[(S_X^{-1}H_2)\otimes H_1]D\frac{\partial\mathrm{Vech}(S_Z)}{\partial s_k}-[(S_X^{-1}H_2)\otimes(S_{YX}S_X^{-1}H_2)]D\frac{\partial\mathrm{Vech}(S_Z)}{\partial s_k},$$

在$\Sigma+\Sigma_\mu$处的雅可比矩阵为

$$\frac{\partial\{\mathrm{Vec}(\widehat{B})\}}{\partial\{\mathrm{Vech}(S_Z)^T\}}\Big|_{S_Z=\Sigma+\Sigma_\mu}=(\Gamma_{X+\xi}^{-1}\otimes I_{p_1})(H_2\otimes H_1)D-(\Gamma_{X+\xi}^{-1}\otimes B)(H_2\otimes H_2)D=QD,$$

其中,

$$Q=(\Gamma_{X+\xi}^{-1}\otimes I_{p_1})(H_2\otimes H_1)-(\Gamma_{X+\xi}^{-1}\otimes B)(H_2\otimes H_2).$$

因此, 由delta方法, 有

$$\sqrt{n}\cdot\mathrm{Vec}(\widehat{B}-B)\xrightarrow{d}\mathcal{N}_{p_1p_2}(\mathbf{0},\Phi_B),$$

其中$\Phi_B=QD(\Lambda+\Lambda_\mu)D^TQ^T$.

容易看出, 其中

$$\widehat{\mathbf{a}}-\mathbf{a}=\widehat{P}(\overline{Z}-\mu)-(\xi^T\otimes I_{p_1})\mathrm{Vec}(\widehat{B}-B),\qquad \widehat{P}=(I_{p_1},-\widehat{B}),$$

有

$$\begin{aligned}\begin{pmatrix}\sqrt{n}(\widehat{\mathbf{a}}-\mathbf{a})\\ \sqrt{n}\cdot\mathrm{Vec}(\widehat{B}-B)\end{pmatrix}&=\begin{pmatrix}\widehat{P}&-(\xi^T\otimes I_{p_1})\\ \mathbf{0}&I_{p_1p_2}\end{pmatrix}\begin{pmatrix}\sqrt{n}(\overline{Z}-\mu)\\ \sqrt{n}\cdot\mathrm{Vec}(\widehat{B}-B)\end{pmatrix}\\ &\xrightarrow{d}\mathcal{N}_{p_1+p_1p_2}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix},\begin{pmatrix}\Phi_{\mathbf{a}}&\Phi_{\mathbf{a}B}\\ \Phi_{B\mathbf{a}}&\Phi_B\end{pmatrix}\right).\end{aligned}$$

证毕. □

**推论4.1的证明:**

若记$\widetilde{\Lambda}=(A\otimes A)\Lambda_{\mathbf{r}}(A\otimes A)^T$, 则$\Lambda=D^+\widetilde{\Lambda}D^{+T}$, 由定理3.2得

$$\begin{aligned}&\widetilde{\Lambda}_{1,1}=\sigma_e^2+4b^2\sigma_u^2\sigma_e^2+4b^2\sigma_X^2\sigma_e^2+b^4\sigma_u^2+4b^4\sigma_X^2\sigma_u^2+b^4\sigma_X^2,\\ &\widetilde{\Lambda}_{1,2}=\widetilde{\Lambda}_{1,3}=\widetilde{\Lambda}_{2,1}=\widetilde{\Lambda}_{3,1}=2b\sigma_X^2\sigma_e^2+2b^3\sigma_X^2\sigma_u^2+b^3\sigma_X^2,\\ &\widetilde{\Lambda}_{1,4}=\widetilde{\Lambda}_{4,1}=b^2\sigma_X^2,\qquad \widetilde{\Lambda}_{2,2}=\widetilde{\Lambda}_{2,3}=\widetilde{\Lambda}_{3,2}=\widetilde{\Lambda}_{3,3}=\sigma_X^2\sigma_e^2+b^2\sigma_X^2\sigma_u^2+b^2\sigma_X^2,\\ &\widetilde{\Lambda}_{2,4}=\widetilde{\Lambda}_{4,2}=\widetilde{\Lambda}_{3,4}=\widetilde{\Lambda}_{4,3}=b\sigma_X^2,\qquad \widetilde{\Lambda}_{4,4}=\sigma_X^2\end{aligned}$$

和

$$\begin{aligned}&\Sigma\otimes\Sigma_\mu\\ =&\begin{pmatrix}b^4(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+b^2\sigma_e^2\sigma_\xi^2&b^3(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+b\sigma_e^2\sigma_\xi^2&b^3\sigma_X^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2\\ b^3(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+b\sigma_e^2\sigma_\xi^2&b^2(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+\sigma_e^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2\\ b^3\sigma_X^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2\\ b^2\sigma_X^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2&\sigma_X^2\sigma_\xi^2\end{pmatrix},\\ &\Sigma\star\Sigma_\mu\\ =&\begin{pmatrix}b^4(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+b^2\sigma_e^2\sigma_\xi^2&b^3\sigma_X^2\sigma_\xi^2&b^3(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+b\sigma_e^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2\\ b^3(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+b\sigma_e^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2&b^2(\sigma_X^2+\sigma_u^2)\sigma_\xi^2+\sigma_e^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2\\ b^3\sigma_X^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2&b^2\sigma_X^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2\\ b^2\sigma_X^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2&b\sigma_X^2\sigma_\xi^2&\sigma_X^2\sigma_\xi^2\end{pmatrix}.\end{aligned}$$

$\Sigma_\mu \otimes \Sigma$和$\Sigma_\mu \star \Sigma$同理可计算, 把上述结果代入(3.1)式$\Phi_a$, $\Phi_{ab}$及$\Phi_b$的表达式中即得(4.2)–(4.4)式结果. □

**致谢** 非常感谢审稿者提出的宝贵意见.

## 参 考 文 献


[1] Berkson, J., Are there two regressions, *J. Am. Statist. Assoc.*, **45**(1950), 164–180.

[2] Cui, H.J. and Chen, S.X., Empirical likelihood confidence region for parameter in the errors-in-variables models, *J. Multi. Anal.*, **84**(2003), 101–115.

[3] Wang, Q.H., Estimation of partial linear error-in-variables models with validation data, *J. Multi. Anal.*, **69**(1999), 30–64.

[4] 薛留根, 非线性EV回归模型中参数估计的渐近性质, 数学年刊, **26A**(2005), 351–360.

[5] Patriota, A.G., Bolfarine, H. and Arellano-Valle, R.B., A multivariate ultrastructural errors-in-variables model with equation error, *J. Multi. Anal.*, **102**(2011), 386–392.

[6] Arellano-Valle, R.B., Bolfarine, H. and Vilca-Labra, R., Ultrastructural elliptical model, *Can. J. Stat.*, **24**(1996), 207–216.

[7] Fuller. W.A., *Measurement Error Models*, New York: Wiley, 1987.

[8] Wang, L., Estimation of nonlinear models with berkson measurement errors, *Ann. Statist.*, **32**(2004), 2559–2579.

[9] 刘强, 薛留根, 带有Berkson测量误差的非线性半参数模型的渐近性质, 北京工业大学学报, **35**(2009), 1567–1572.

[10] Magnus, J.R. and Neudecker, H., *Matrix Differential Calculus with Applications in Statistics and Econometrics*, New York: Wiley, 2007.


# The Analysis of Multivariate Ultrastructural Berkson Measurement Errors Model

ZHANG SAIYIN  WU MIXIA  ZHANG ZHONGZHAN

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

Berkson measurement error regression model has a wide range of applications in the areas of industry, agriculture, epidemiology, economics, and so on. This paper deals with a multivariate ultrastructural Berkson measurement error regression model, and presents consistent estimators for model parameters. The asymptotic distributions for the estimators are derived, and an application to a simple ultrastructural Berkson measurement error regression model is given. A real example and simulation results are provided for the illustration of the method proposed in this paper.

**Keywords:** Asymptotic property, consistent estimator, Berkson measurement error, multivariate regression.

**AMS Subject Classification:** 62F10, 62F12.