中山大学学报（自然科学版）
第29卷 第4期 ACTA SCIENTIARUM NATURALIUM Vol.29 No.4
1990年 UNIVERSITATIS SUNYATSENI 1990

一类矩阵椭球等高分布在数据不完全的参数估计*

杨维权 黄亚川**

（数学系）

摘 要

对于一类矩阵椭球等高分布在数据不完全（缺失数据）情况下，本文导出了未知参数具有MLE条件的估计，并证明$\hat{\mu}_1$、$\hat{\mu}_2$分别为μ_1、μ_2的无偏估计、相合估计及较有效估计.

关键词 矩阵椭球等高分布族，不完全数据，参数估计

1 引 言

在运用回归分析、判别分析和因子分析等多元统计方法中，通常要假定n次独立观察数据矩阵$X=(x_{ti})_{n\times p}$为完全方阵。然而，由于种种原因，使得有些变量的观察数据缺失，X并非完全方阵。诚然，可以剔除那几次不完全观察的数据，剩下一个完全方阵$X=(x_{ti})n_{1\times p}$，$n_1<n$。但这样处理势必丢掉不少数据的信息。较为科学的方法是在保留到手的数据基础上，运用估计原理填补缺失数据，得到$n\times p$的完全方阵X。

文〔1〕、〔2〕对于指数分布族在数据不完全情形，导出了参数的MLE。文〔3〕对于矩阵椭球等高分布族在数据完全情形，给出了参数MLE的存在性证明。本文对于一类矩阵椭球等高分布族在数据不完全情形，导出未知参数具有MLE条件的估计，建立了3个定理。

2 引 理

设

$$X^{(i)}=\begin{pmatrix} X^{(i)}_{(1)} \\ \cdots\cdots \\ X^{(i)}_{(2)} \end{pmatrix}\begin{matrix} q \\ \\ p-q \end{matrix} \qquad \begin{matrix} 1\leqslant q<p,\ n>p \\ \\ i=1,2,\cdots,n \end{matrix}$$

$$\mathrm{X}=\begin{pmatrix} X^{(1)\prime} \\ \vdots \\ X^{(n)\prime} \end{pmatrix}=\begin{pmatrix} \mathrm{X}_{11} & \vdots & \mathrm{X}_{12} \\ \cdots & \vdots & \cdots \\ \mathrm{X}_{21} & \vdots & \mathrm{X}_{22} \end{pmatrix}\begin{matrix} n_1 \\ \\ n-n_1 \end{matrix} \tag{1}$$
$$\qquad\qquad q \qquad p-q$$

本文1989年3月10日收到

* 国家教委博士点基金资助项目. ** 1985级研究生

0114186

记$W=(X_{11}, X_{12}, X_{21})$，$Z=X_{22}$。假设$W$为已观察到的随机变量集合，$Z$为缺失随机变量集合，相应地有如下分块：

$$\mu=\begin{pmatrix}\mu_{(1)}\\ \cdots\cdots\\ \mu_{(2)}\end{pmatrix}, \qquad \Sigma=\begin{pmatrix}\Sigma_{11} & \vdots & \Sigma_{12}\\ \cdots & \vdots & \cdots\\ \Sigma_{21} & \vdots & \Sigma_{22}\end{pmatrix}$$

设 $X^{(1)},\cdots,X^{(n)}$为样本，若

$$V_{ec}(\mathbf{X})=\begin{pmatrix}X^{(1)}\\ \vdots\\ X^{(n)}\end{pmatrix}\sim EC_{np}(e_n\otimes\mu, I_n\otimes\Sigma,\varphi)$$

则记
$$\mathbf{X}=\begin{pmatrix}X^{(1)\prime}\\ \vdots\\ X^{(n)\prime}\end{pmatrix}_{n\times p}\sim LEC_{n\times p}(e_n\mu', I_n\otimes\Sigma,\varphi)$$

若$\mathbf{X}$有概率密度函数

$$|\Sigma|^{-n/2}g\{\operatorname{tr}[(\mathbf{X}-e_n\mu')\Sigma^{-1}(\mathbf{X}-e_n\mu')]\} \tag{2}$$

其中$g(\cdot)$为连续函数，此时记

$$\mathbf{X}\sim LEC_{n\times p}(e_n\mu', I_n\otimes\Sigma, g)$$

引理1 若$\mathbf{X}=(X^{(1)},\cdots,X^{(n)})'\sim LEC_{n\times p}(e_n\mu', I_n\otimes\Sigma,\varphi)$，则$X^{(i)}\sim EC_p(\mu, \Sigma, \varphi)$，$i=1,\cdots,n$。

引理2 若$\mathbf{X}=(X^{(1)},\cdots,X^{(n)})'\sim LEC_{n\times p}(e_n\mu', I_n\otimes\Sigma, g)$，$g(\cdot)$为单调下降的连续函数，则$(\mu,\Sigma)$的MLE为

$$\hat{\mu}=\frac{1}{n}\mathbf{X}'e_n \tag{3}$$

$$\hat{\Sigma}=\lambda_n\mathbf{X}'(I_n-\frac{1}{n}e_ne_n')\mathbf{X} \tag{4}$$

其中λ_n为使函数$\lambda^{-np/2}g(\frac{p}{\lambda})$达到极大值的$\lambda$值。

引理3 在引理2的假设下，若$E(\mathbf{X})$存在，则有

$$E\left(X^{(i)}_{(2)}|W,\mu,\Sigma\right)=\mu_{(2)}+\Sigma_{21}\Sigma^{-1}_{11}\left(X^{(i)}_{(1)}-\mu_{(1)}\right),(i=n_1+1,\cdots,n) \tag{5}$$

引理的证明参见文[3,4]。

3 定 理

定理1 设有引理2，记W为观察到的部分，Z为缺失部分。设$n_1>q$。则有

(i) $P\{S>\square\}=1$。其中

$$S=\mathbf{X}'_{11}(I_{n1}-\frac{1}{n_1}e_{n_1}e'_{n_1})\mathbf{X}_{11} \tag{6}$$

(ii) (μ,Σ)具有MLE条件的估计为

$$\hat{\mu}_{(1)}=\frac{1}{n}\mathbf{X}'_{(1)}e_n \tag{7}$$

$$\hat{\Sigma}_{11}=\lambda_n\mathbf{X}'_{(1)}\left(I_n-\frac{1}{n}e_ne'_n\right)\mathbf{X}_{(1)} \tag{8}$$

$$\hat{\Sigma}_{12}=\hat{\Sigma}_{11}S^{-1}\mathbf{X}'_{11}\left(I_{n_1}-\frac{1}{n_1}e_{n_1}c'_{n_1}\right)\mathbf{X}_{12} \tag{9}$$

$$\hat{\mu}_{(2)}=\frac{1}{n_1}\Big[X'_{12}e_{n_1}+\sum_{t=n_1+1}^{n}\hat{\Sigma}'_{12}\hat{\Sigma}^{-1}_{11}\Big(X^{(t)}_{(1)}-\hat{\mu}_{(1)}\Big)\Big] \quad (10)$$

$$\hat{\Sigma}_{22}=\lambda_n\Big[\sum_{t=1}^{n_1}\Big(X^{(t)}_{(2)}-\hat{\mu}_{(2)}\Big)\Big(X^{(t)}_{(2)}-\hat{\mu}_{(2)}\Big)'$$

$$+\sum_{i=n_1+1}^{n}\hat{\Sigma}'_{12}\hat{\Sigma}^{-1}_{11}\Big(X^{(t)}_{(1)}-\hat{\mu}_{(1)}\Big)\Big(X^{(t)}_{(1)}-\hat{\mu}_{(1)}\Big)'\hat{\Sigma}^{-1}_{11}\hat{\Sigma}_{12}\Big] \quad (11)$$

其中 $\mathbf{X}'_{(1)}=\Big(\mathbf{X}'_{11}\,\vdots\,\mathbf{X}'_{21}\Big)_{n\times q}$，$\lambda_n$如引理 2 中所定义。

证明 (i) $\mathbf{X}_{11}=\Big(X^{(1)}_{(1)},\cdots,X^{(n_1)}_{(1)}\Big)'$的特征函数为

$$E\Big\{\exp\Big[i\mathrm{tr}\Big(T'_{11}\mathbf{X}_{11}\Big)\Big]\Big\}=E\Big\{\exp\Big[i\mathrm{tr}\Big(\begin{matrix}T_{11}\square\\ \square\square\end{matrix}\Big)'\mathbf{X}\Big]\Big\}$$

$$=\exp\Big\{i\mathrm{tr}\Big[T'_{11}e_{n_1}\mu'_{(1)}\Big]\Big\}\varphi\Big\{\mathrm{tr}\Big[T_{11}\Sigma_{11}T'_{11}\Big]\Big\}$$

则 $\mathbf{X}_{11}\sim LEC_{n_1\times q}(e_{n_1}\mu_{(1)},\ I_{n_1}\otimes\Sigma_{11},\ \varphi)$。由文〔5〕中引理 7 及 $n_1>q$ 知，有 $P\{S>\square\}=1$。

(ii) 由引理 3 知，有

$$\hat{X}^{(t)}_{(2)}\hat{=}E\Big(X^{(t)}_{(2)}\Big|W,\mu,\Sigma,\Big)$$

$$=\mu_{(2)}+\Sigma_{21}\Sigma_{11}^{-1}\Big(X^{(t)}_{(1)}-\mu_{(1)}\Big),\quad (t=n_1+1,\cdots,n)$$

由引理 2 知，使似然函数达到极大值的参数应满足下述等式

$$\mu=\begin{pmatrix}\mu_{(1)}\\ \\ \mu_{(2)}\end{pmatrix}=\begin{pmatrix}\frac{1}{n}\sum\limits_{t=1}^{n}X^{(t)}_{(1)}\\ \\ \frac{1}{n}\sum\limits_{t=1}^{n_1}X^{(t)}_{(2)}+\frac{1}{n}\sum\limits_{t=n_1+1}^{n}\Big[\mu_{(2)}+\Sigma_{21}\Sigma^{-1}_{11}\Big(X^{(t)}_{(1)}-\mu_{(1)}\Big)\Big]\end{pmatrix} \quad (12)$$

当$t=n_1+1,\cdots,n$时，$X^{(t)}_{(2)}$用$\hat{X}^{(t)}_{(2)}$代替。

$$\mathbf{\Sigma}=\begin{pmatrix}\Sigma_{11}\ \Sigma_{12}\\ \\ \Sigma_{11}\ \Sigma_{22}\end{pmatrix}=\lambda_n\sum_{t=1}^{n}\begin{pmatrix}\Big(X^{(t)}_{(1)}-\mu_{(1)}\Big)\Big(X^{(t)}_{(1)}-\mu_{(1)}\Big)' & \vdots & \Big(X^{(t)}_{(1)}-\mu_{(1)}\Big)\Big(X^{(t)}_{(2)}-\mu_{(2)}\Big)'\\ \cdots & \vdots & \cdots\\ \Big(X^{(t)}_{(2)}-\mu_{(2)}\Big)\Big(X^{(t)}_{(1)}-\mu_{(1)}\Big)' & \vdots & \Big(X^{(t)}_{(2)}-\mu_{(2)}\Big)\Big(X^{(t)}_{(2)}-\mu_{(2)}\Big)'\end{pmatrix} \quad (13)$$

由(12)、(13)式得(7)、(8)两式。由(12)的第 2 项得

$$\mu_{(2)}=\frac{1}{n_1}\Big[\sum_{t=1}^{n_1}X^{(t)}_{(2)}+\sum_{t=n_1+1}^{n}\Sigma_{21}\Sigma^{-1}_{11}\Big(X^{(t)}_{(1)}-\mu_{(1)}\Big)\Big] \quad (14)$$

将(14)式代入(13)式得

$$\Sigma_{12}=\lambda_n\Big\{\sum_{t=1}^{n_1}\Big(X^{(t)}_{(1)}-\mu_{(1)}\Big)\Big(X^{(t)}_{(2)}-\frac{1}{n_1}\sum_{t=1}^{n}X^{(t)}_{(2)}\Big)'$$

$$-\frac{1}{n_1}\left[\sum_{t=1}^{n_1}\left(X_{(1)}^{(t)}-\mu_{(1)}\right)\right]\left[\sum_{t=n_1+1}^{n}\left(X_{(1)}^{(t)}-\mu_{(1)}\right)'\Sigma_{11}^{-1}\Sigma_{12}\right]$$

$$+\sum_{t=n_1+1}^{n}\left(X_{(1)}^{(t)}-\mu_{(1)}\right)\left(X_{(1)}^{(t)}-\mu_{(1)}\right)'\Sigma_{11}^{-1}\Sigma_{12}\Bigg\} \tag{15}$$

用$\hat{\mu}_{(1)}$、$\hat{\Sigma}_{11}$代替(15)式中的$\mu_{(1)}$、Σ_{11}，由于

$$\sum_{t=n_1+1}^{n}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)=-\sum_{t=1}^{n_1}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)$$

$$\Sigma_{12}=\hat{\Sigma}_{11}\hat{\Sigma}_{11}^{-1}\Sigma_{12}=\lambda_n\sum_{t=1}^{n}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)'\hat{\Sigma}_{11}^{-1}\Sigma_{12}$$

于是(15)式可写成

$$\left[\sum_{t=1}^{n_1}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)'-\frac{1}{n_1}\sum_{t=1}^{n_1}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)\right.$$

$$\left.\cdot\sum_{t=1}^{n_1}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)'\right]\hat{\Sigma}_{11}^{-1}\Sigma_{12}$$

$$=\sum_{t=1}^{n_1}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)\left(X_{(1)}^{(t)}-\frac{1}{n_1}\sum_{t=1}^{n_1}X_{(2)}^{(t)}\right)' \tag{16}$$

(16)式的左端可表示为

$$\left(\mathbf{X}_{11}-e_{n_1}\hat{\mu}_{(1)}'\right)'\left(\mathbf{X}_{11}-e_{n_1}\hat{\mu}_{(1)}'\right)-\frac{1}{n_1}\left(\mathbf{X}_{11}-e_{n_1}\hat{\mu}_{(1)}'\right)'e_{n_1}e_{n_1}'\left(\mathbf{X}_{11}-e_{n_1}\hat{\mu}_{(1)}'\right)$$

$$=\mathbf{X}_{11}'\left(I_{n_1}-\frac{1}{n_1}e_{n_1}e_{n_1}'\right)\mathbf{X}_{11}-\hat{\mu}_{(1)}e_{n_1}'\left(I_{n_1}-\frac{1}{n_1}e_{n_1}e_{n_1}'\right)\mathbf{X}_{11}$$

$$-\mathbf{X}_{11}'\left(I_{n_1}-\frac{1}{n_1}e_{n_1}e_{n_1}'\right)e_{n_1}\hat{\mu}_{(1)}'+\hat{\mu}_{(1)}'e_{n_1}\left(I_{n_2}-\frac{1}{n_1}e_{n_1}e_{n_1}'\right)e_{n_1}\hat{\mu}_{(1)}$$

$$=X_{11}'\left(I_{n_1}-\frac{1}{n_1}{}_{n_1}e_{n_1}'\right)\mathbf{X}_{11}=S$$

于是

$$\hat{\Sigma}_{12}=\hat{\Sigma}_{11}S^{-1}\left[\sum_{t=1}^{n_1}\left(X_{(1)}^{(t)}-\hat{\mu}_{(1)}\right)\left(X_{(2)}^{(t)}-\frac{1}{n_1}\sum_{t=1}^{n_1}X_{(2)}^{(t)}\right)'\right]$$

$$=\hat{\Sigma}_{11}S^{-1}\left[\left(\mathbf{X}_{11}-\frac{1}{n}e_{n_1}e_{n_1}'\mathbf{X}_{(1)}\right)'\left(I_{n_1}-\frac{1}{n_2}e_{n_1}e_{n_1}'\right)\mathbf{X}_{12}\right]$$

$$=\hat{\Sigma}_{11}S^{-1}\mathbf{X}_{11}'\left(I_{n_1}-\frac{1}{n_1}e_{n_1}e_{n_1}'\right)\mathbf{X}_{12}$$

由(14)式得(10)式，并得(11)式。

定理2　在定理1的假设下，则有

(i) (7)式的$\hat{\mu}_{(1)}\sim EC_q\left(\mu_{(1)},\frac{1}{n}\Sigma_{11},\varphi\right)$

(ii) 当Σ为已知时，(14)式中$\mu_{(1)}$用$\hat{\mu}_{(1)}$代替，则

$$\hat{\mu}_{(2)} \sim EC_{p-q}\left(\mu_{(2)}, \frac{1}{n_1}\Sigma_{22} - \frac{n_2}{nn_1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}, \varphi\right)$$

证明 (i) $\hat{\mu}_{(1)} = \left(\frac{1}{n}I_q, \square, \cdots, \frac{1}{n}I_q, \square\right)\text{Vec}(\mathbf{X}) \hat{=} C_1\text{Vec}(\mathbf{X})$，于是

$C_1(e_n \otimes \mu) = \mu_{(1)}$，

$$C_1(I_n \otimes \Sigma)C_1' = \left(\frac{1}{n}\Sigma_{11}, \frac{1}{n}\Sigma_{12}, \cdots, \frac{1}{n}\Sigma_{11}, \frac{1}{n}\Sigma_{12}\right)C_1' = \frac{1}{n}\Sigma_{11}\text{，从而}$$

$\hat{\mu}_{(1)} = C_1\text{Vec}(\mathbf{X}) \sim EC_q\left(\mu_{(1)}, \frac{1}{n}\Sigma_{11}, \varphi\right)$

(ii) 当Σ为已知时，记$n_2 = n - n_1$，则有

$$\hat{\mu}_{(2)} = \frac{1}{n_1}\left[\mathbf{X}_{12}' e_{n_1} + \Sigma_{21}\Sigma_{11}^{-1}\left(\mathbf{X}_{21}e_{n_2} - \frac{n_2}{n}\mathbf{X}_{(1)}' e_n\right)\right]$$

$$= \left(\frac{-n_2}{n\,n_1}\Sigma_{21}\Sigma_{11}^{-1}, \frac{1}{n_1}I_{p-q}, \cdots, \frac{-n_2}{n\,n_1}\Sigma_{21}\Sigma_{11}^{-1}, \frac{1}{n_1}I_{p-q},\right.$$

$$\left.\frac{1}{n}\Sigma_{21}\Sigma_{11}^{-1}, \square, \cdots, \frac{1}{n}\Sigma_{21}\Sigma_{11}^{-1}, \square\right)\text{Vec}(\mathbf{X}) \hat{=} C_2\text{Vec}(\mathbf{X})$$

于是，$C_2(e_n \otimes \mu) = \mu_{(2)}$，

$$C_2(I_n \otimes \Sigma)C_2' = \frac{1}{n_i}\Sigma_{22} - \frac{n_2}{n\,n_1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

从而 $\hat{\mu}_{(2)} = C_2\text{Vec}(\mathbf{X}) \sim EC_{p-q}\left(\mu_{(12)}, \frac{1}{n}\Sigma_{22} - \frac{n_2}{n\,n_1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}, \varphi\right)$

定理 3 在定理 1 的假设下，若$\mathbf{X}$的二阶矩存在，则有

(i) 定理 1 中的$\hat{\mu}_{(1)}$及$\hat{\mu}_{(2)}$分别为$\mu_{(1)}$及$\mu_{(2)}$的相合估计。

(ii) $\hat{\theta}^{(n)} = \left(\sqrt{n}\left(\hat{\mu}_{(1)} - \mu_{(1)}\right)', \sqrt{n_1}\left(\hat{\mu}_{(2)} - \mu_{(2)}\right)'\right)$的分布为

$EC_p(\square, \Sigma^{(n)}, \varphi)$，其中

$$\Sigma^{(n)} = \begin{pmatrix} \Sigma_{11} & \vdots & (n_1/n)^{1/2}\Sigma_{21} \\ (n_1/n)^{1/2}\Sigma_{21} & \vdots & \Sigma_{22} - (n_2/n)\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}$$

(iii) 当且仅当$\varphi(t) = e^{-bt/2}$，$\hat{\theta}^{(n)}$渐近于正态分布$N_p(\square, b\Sigma^{(n)})$。

证明从略。

上述定理的意义在于用$\mathbf{X}$数据矩阵建立参数的估计，优于用$\mathbf{X}_1$数据矩阵建立参数的估计，如μ的估计，前者用$\hat{\mu}$表示，后者用$\tilde{\mu}$表示，记作

$$\hat{\mu} = \left(\hat{\mu}_1, \cdots, \hat{\mu}_q, \hat{\mu}_{q+1}, \cdots, \hat{\mu}_p\right)'$$

$$\tilde{\mu} = \left(\tilde{\mu}_1, \cdots, \tilde{\mu}_q, \tilde{\mu}_{q+1}, \cdots, \tilde{\mu}_p\right)'$$

$\Sigma = (\sigma_{ij})_{p\times p}$，$\alpha = -2\varphi'(o)$，用$\rho_{q+i;1,\cdots,q}$表示分量$x_{q+j}$同分量$x_1, \cdots, x_q$的多重相关系

数，$j=1,\cdots,p-q$。由定理2知，则有

$$D(\hat{\mu}_i)=\frac{\alpha}{n}\sigma_{ii}<\frac{\alpha}{n_1}\sigma_{ii}=D(\tilde{\mu}_i),\quad i=1,\cdots,q.$$

$$D(\hat{\mu}_{q+j})=\alpha\left(\frac{1}{n_1}\sigma_{q+j,q+j}-\frac{n_2}{n\,n_1}\sigma_{q+j,q+j}\rho^2_{q+j;1,\cdots,q}\right)$$

$$\leqslant\frac{\alpha}{n_1}\sigma_{q+j,q+j}=D(\tilde{\mu}_{q+j}),\quad j=1,\cdots,p-q.$$

当n_2值较大或多重相关系数绝对值较大时，定理2建立的估计量更为有效。

参考文献

〔1〕 Dempster A P et al., *J.Roy. Statist. Soc.*, B.39(1977), 1～38

〔2〕 Orchard T et al., *Proc.6th Berkeley Symp.. Math. Statist*, prob.1,1972, 697～715

〔3〕 张尧庭等，数学物理学报，1985，341～353

〔4〕 杨维权等，多元统计分析，高等教育出版，1989

〔5〕 Mairhead R J, *Aspects of Multiveriate Statistical Theory*, Wiley,1982

The Estimation of Parameters for A Class of Matrix Elliptically Contoured Distributions in Incomplete Data

*Yang Weiquan** *Huang Yachuan*

Abstract

The estimation of parameters for a class of matrix elliptically contoured distributions (MECD) in incomplete data is discussed. In the case of i.i.d. exponential-family, EM-algorirhm is feasible and effective, but not so good in MECD.

Using the conditional expectation of missing values given all known data as the estimators of missing values, by the results of MECD,we give estimators in Theorem 1. Under certain general regularity conditions, we prove that $\hat{\mu}_1$、$\hat{\mu}_2$ are unbised and consistent estimators of μ_1、μ_2 respectively and that the potential gain in precision by using the incomplete data is attainable.

Keywords matrix elliptically Contoured distribution, incomplete data, estimation of parameters

* Department of Mathematics