Wang and Zhang Journal of Inequalities and Applications 2013, 2013:502
http://www.journalofinequalitiesandapplications.com/content/2013/1/502

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

# The maximum likelihood estimations for a type of left ellipsoidal distribution

Yanling Wang[1*] and Yingshan Zhang[2]

*Correspondence: bigduckwyl@163.com
[1]College of Mathematics and Information Science, Henan Normal University, Xinxiang, 453007, China
Full list of author information is available at the end of the article

**Abstract**

In this paper, we prove that if the distribution of an $n \times mp$ random matrix $Y$ is a left ellipsoidal distribution with parameter $\mu^* = \mu(\mathrm{I}\mathrm{I}'_m \otimes \mathrm{I}_p)$, $\Sigma^* = m\Sigma$, and $Y_1, Y_2, \ldots, Y_m$ are independent and identical distributions, the maximum likelihood estimations of $\mu$, $\Sigma$ are $\overline{Y}$, $S^2$ if and only if $Y \sim N_{n\times mp}(\mu^*, \Sigma^*)$. If $Y_1, Y_2, \ldots, Y_m$ are not independent and identical distributions, then $Y$ may be not a normal distribution.

**Keywords:** left ellipsoidal distribution; normal distribution; maximum likelihood estimation

## 1 Introduction

Let $X = (x_1, x_2, \ldots, x_n)$, $x_1, x_2, \ldots, x_n$ be independent standard normal distributions. A random matrix $Y_{n\times p}$ is said to be a left stochastic ellipsoid matrix if it satisfies (see [1])

$$Y_{n\times p} = A_{n\times n}X_{n\times p} + \mu_{n\times p}, \qquad \Gamma_{n\times n}X_{n\times p} \stackrel{d}{=} X_{n\times p}, \qquad EXX' = I_n,$$

and its distribution, which is denoted by $E_{n\times p}(\mu, \Sigma)$, is called a left ellipsoidal distribution, where $\Sigma = AA'$ is reversible, and $E$ is ellipsoid. Besides, $d$ denotes identical distribution, $S$ denotes $XX'$, and all $\Gamma_{n\times n} \in o(n)$ are orthogonal matrices with order $n$. Then we have conclusions:

$$ES^kX = 0, \qquad ES^k = \alpha^k I_n,$$

where $k$ is an arbitrary integer and $k$ is a constant (see [2]).

If $X$ has the distribution density, the distribution density has the form $f(X'X)$, then the distribution density is said to be spherical distribution. If the distribution density of $Y$ is

$$|\Sigma|^{-\frac{p}{2}} f\bigl((Y-\mu)'\Sigma^{-1}(Y-\mu)\bigr),$$

then the distribution density is called left ellipsoidal distribution contour, and the contour surface is

$$(Y-\mu)'\Sigma^{-1}(Y-\mu) = D.$$

When $p = 1$, it is an ellipsoidal surface. The spectral decomposition of $\Sigma$ is $\Gamma\Lambda\Gamma'$, where $\Gamma$ is an orthogonal matrix, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\lambda_1 > \cdots > \lambda_n > 0$.

If the distribution density of $Y$ is

$$|\Sigma|^{-\frac{p}{2}} f\bigl[(Y-\mu)'\Sigma^{-1}(Y-\mu)\bigr] = \left(\sqrt{\frac{p}{2\pi}}\right)^{np} |\Sigma|^{-\frac{p}{2}} \exp\left\{-\frac{p}{2}\operatorname{tr}(Y-\mu)'\Sigma^{-1}(Y-\mu)\right\},$$

the elongated vector of $Y$, which is denoted by $\vec{Y}$, is normal $N_{np}(\vec{\mu}, \frac{1}{p}\mathrm{I}_p \otimes \Sigma)$ distribution. For the sake of consistency of symbols, we denote the normal distribution by $N_{n\times p}(\mu, \Sigma)$.

Classical statistical analysis is built on the basis of normal distribution. However, it is still an important problem whether these graceful properties can also be satisfied without the condition that it is a normal distribution.

Since the left ellipsoidal distribution family has much more members than the normal distribution family (in fact, it almost includes all common distributions), a large amount of scholars' fruitful research shows: on the one hand, the left ellipsoidal distribution as a multivariate normal distribution promotion is ideal; on the other hand, based on the research of the left ellipsoidal distribution, we can get many statistics used as solid properties of hypothesis test (see [3]). It is a trivial idea to extend the properties of the normal distribution family to the left ellipsoidal distribution family. Nevertheless, that is not always true. In this article, we prove that the maximum likelihood property of normal distribution cannot be extended to the left ellipsoidal distribution.

Let $Y_1, Y_2, \ldots, Y_m$ be the samples of independent identical distributions, and in the condition of normal distribution,

$$\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i, \qquad S^2 = \frac{1}{m}\sum_{i=1}^{m}(Y_i - \overline{Y})(Y_i - \overline{Y})'$$

are the maximum likelihood estimations of $\mu$, $\Sigma$, respectively, if $m$ is big enough and $S^2$ is positive definite. When $p = 1$, $S^2$ is positive definite with probability 1 if $m > n$ (see [4]). Additionally, the rank of $S^2$ does not decrease with the increasing of the columns in $Y$ so that $S^2$ is positive definite when $m > n$. Now we discuss them in the case of $m > \max(n, 2)$. When one focuses on the maximum likelihood estimation, the likelihood equations need to be deduced by the matrix differential method. Therefore, we add a trivial condition that the distribution density of $E_{n\times p}(\mu, \Sigma)$ is differentiable (see [5–8]).

We draw the following conclusions.

**Theorem 1.1** *Let $Y_1, Y_2, \ldots, Y_m$ be independent identically distributed, and $Y_i \sim E_{n\times p}(\mu, \Sigma)$. The maximum likelihood estimation of $\mu$, $\Sigma$ is $\overline{Y}$, $S^2$ if and only if $Y_i \sim N_{n\times p}(\mu, \Sigma)$.*

Note that for the proof of Theorem 1.1 one can refer to [9].

When the variables of $\mu^*$ are not independent, we have the solution as follows.

**Theorem 1.2** *Let $Y \sim E_{n\times mp}(\mu^*, \Sigma^*)$, $\mu^* = \mu(\mathrm{II}_m' \otimes \mathrm{I}_p)$, and $\Sigma^* = m\Sigma$. Besides, $Y = (Y_1, Y_2, \ldots, Y_m)$, where $Y_i$ is an $n \times p$ matrix. Then if $Y_1, Y_2, \ldots, Y_m$ are mutually independent identically distributed, the maximum likelihood estimations of $\mu$, $\Sigma$ are*

$$\overline{Y} = \frac{1}{m} Y\bigl(\mathrm{II}_m' \otimes \mathrm{I}_p\bigr), \qquad S^2 = \frac{1}{m} Y(I - P_{\mathrm{II}_m} \otimes \mathrm{I}_p)Y'$$

*if and only if $Y \sim N_{n\times mp}(\mu^*, \Sigma^*)$, where $P_{\mathrm{II}_m} = \frac{1}{m}\mathrm{II}_m\mathrm{II}_m'$, we note it $P = P_{\mathrm{II}_m} \otimes \mathrm{I}_p$ in the proof below.*

**Theorem 1.3** *Under the conditions of Theorem* 1.2, *if* $Y_1, Y_2, \ldots, Y_m$ *are not independent identically distributed, and the maximum likelihood estimations of* $\mu$, $\Sigma$ *are*

$$\overline{Y} = \frac{1}{m} Y\big(\mathrm{II}_m' \otimes \mathrm{I}_p\big), \qquad S^2 = \frac{1}{m} Y(I - P_{\mathrm{II}_m} \otimes \mathrm{I}_p)Y',$$

*then* $Y$ *may be not a normal distribution.*

## 2 The proof of Theorem 1.2

Let $Y_1, Y_2, \ldots, Y_m$ be mutually identically independent distributions, and $Y_i \sim E_{n\times p}(\mu, \Sigma)$. Besides, we assume that $Y = Y_1, Y_2, \ldots, Y_m$, and $Y_i = AX_i + \mu$, then

$$\begin{aligned} &Y = mA\left(\frac{1}{m}X\right) + \mu\big(\mathrm{II}_m' \otimes \mathrm{I}_p\big), \\ &\Gamma_{n\times n}\left(\frac{1}{m}X\right) \stackrel{d}{=} \frac{1}{m}X, \\ &E\left(\frac{1}{m}X\right)\left(\frac{1}{m}X\right)' = I, \end{aligned}$$

where $\Gamma$ is a random orthogonal matrix with order $n$. As a result, $Y \sim E_{n\times mp}(\mu(\mathrm{II}_m' \otimes \mathrm{I}_p), m\Sigma)$.

Now the above proposition can be transformed to a research of the form of the left stochastic ellipsoid distribution with a degraded mean.

Notice that

$$\begin{aligned} Y &= \frac{1}{m}\sum_{i=1}^{m} Y_i = \frac{1}{m} Y(\mathrm{II}_m \otimes \mathrm{I}_p), \\ S^2 &= \frac{1}{m}\sum_{i=1}^{m}(Y_i - \mu)(Y_i - \mu)' \\ &= \frac{1}{m}\sum_{i=1}^{m}(Y_i - Y)(Y_i - Y)' = \frac{1}{m}(YP - Y)(YP - Y)' = \frac{1}{m}Y(I - P)Y', \end{aligned}$$

where $\mu = Y$, $P = P_{\mathrm{II}_m} \otimes \mathrm{I}_p$. Therefore, utilizing the single mean degraded left stochastic ellipsoid to express Theorem 1.1, we can obtain Theorem 1.2.

## 3 The proof of Theorem 1.3

Based on the above discussion, we get the question: If the condition of mutually identically independent distribution is discarded, can one obtain the same solution with Theorem 1.1 from $Y \sim E_{n\times mp}(\mu \cdot (\mathrm{II}_m' \otimes \mathrm{I}_p), m\Sigma)$? In other words, can the proposition be proved or not?

Let $Y \sim E_{n\times mp}(\mu \cdot (\mathrm{II}_m' \otimes \mathrm{I}_p), m\Sigma)$, and then the maximum likelihood estimations of $\mu$, $\Sigma$ are

$$\overline{Y} = \frac{1}{m} Y\big(\mathrm{II}_m' \otimes \mathrm{I}_p\big), \qquad S^2 = \frac{1}{m} Y(I - P)Y'.$$

Here, $P = P_{\mathrm{II}_m} \otimes \mathrm{I}_p$ if and only if $Y \sim N_{n\times mp}(\mu(\mathrm{II}_m' \otimes \mathrm{I}_p), m\Sigma)$.

The proposition is proved not to be true through the above discussion, and we can give the paradoxical instance. Now, we firstly deduce the necessary and sufficient conditions that the maximum likelihood estimation of the parameters $\mu$, $\Sigma$ in the left stochastic ellipsoid $E_{n\times mp}(\mu(\mathrm{II}'_m \otimes \mathrm{I}_p), m\Sigma)$ is $\overline{Y}$, $S^2$.

We assume that $Y$ has the distribution density, namely the likelihood function is

$$L(Y,\mu,\Sigma) = |\Sigma|^{-\frac{mp}{2}} f\bigl(\bigl(Y-\mu\bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr)\bigr)'(m\Sigma)^{-1}\bigl(Y-\mu\bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr)\bigr)\bigr).$$

Computing the logarithm difference of both sides, we can get

$$\begin{aligned} d\ln L(Y,\mu,\Sigma) &= -2\operatorname{tr} G\bigl(Y-\mu\bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr)\bigr)'\Sigma^{-1}\cdot d\mu\cdot\bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr) \\ &\quad - \operatorname{tr}\left(\Sigma^{-1}\bigl(Y-\mu\bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr)\bigr)'G\bigl(Y-\mu\bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr)\bigr)\Sigma^{-1} + \frac{mp}{2}\Sigma^{-1}\right)d\Sigma, \end{aligned}$$

where $A = (Y-\mu(\mathrm{II}'_m \otimes \mathrm{I}_p))'(m\Sigma)^{-1}(Y-\mu(\mathrm{II}'_m \otimes \mathrm{I}_p)) = (a_{ij})_{n\times n}$ and $A' = A$.

$$\begin{aligned} d\ln f(A) &= \sum_{i<j} \frac{\partial \ln f(A)}{\partial a_{ij}}\, da_{ij} \\ &= \sum_{i\cdot j} \ln f(A)_{ij}\, da_{ij} \\ &= \operatorname{tr} G(A)\, dA. \end{aligned}$$

We define

$$G(A)_{ii} = \frac{\partial \ln f(A)}{\partial a_{ij}}, \qquad G(A)_{ij} = \frac{1}{2}\frac{\partial \ln f(A)}{\partial a_{ij}}, \qquad G = G(A).$$

Let $P = P_{\mathrm{II}_m} \otimes \mathrm{I}_p$, and then the maximum likelihood estimation of $\mu$, $\Sigma$, which is denoted by $\hat{\mu}$, $\hat{\Sigma}$, satisfies

$$\begin{aligned} &\bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr)G\cdot P\cdot Y' = \bigl(\mathrm{II}'_m \otimes \mathrm{I}_p\bigr)G\cdot Y', \\ &Y\left((I-P)G(I-P) + \frac{P}{2}\mathrm{I}_p\right)Y' = 0. \end{aligned}$$

Namely,

$$\begin{aligned} &PGP = PG, \\ &(I-P)G(I-P) + \frac{P}{2}(I-P) = 0. \end{aligned}$$

Then we can obtain

$$\begin{aligned} &PGP = PG, \\ &\left(G + \frac{P}{2}\right)(I-P) = 0. \end{aligned}$$

As a result, we can conclude that

$$G = -\frac{P}{2}I + CP,$$

where $C$ is a random $mp \times mp$ matrix.

Since $G' = G$, $CP = PC'$. Besides, the 2-order differential of a likelihood function needs to satisfy: $\forall d\mu, d\Sigma, d^2 \ln L(Y,\mu,\Sigma) < 0$.

$$\because\ L = |\Sigma|^{-\frac{mp}{2}} f(A),$$
$$\therefore\ d\ln L = -\frac{mp}{2}\operatorname{tr}\Sigma^{-1}\,d\Sigma + d\ln f(A) = -\frac{mp}{2}\operatorname{tr}\Sigma^{-1}\,d\Sigma + \operatorname{tr} G\,d(A).$$

Additionally,

$$\begin{aligned} d^2\ln L &= \frac{mp}{2}\operatorname{tr}\Sigma^{-1}\,d\Sigma\,\Sigma^{-1}\,d\Sigma + \operatorname{tr} G\cdot d^2A + \operatorname{tr}(dG)\,dA \\ &= \operatorname{tr}\biggl(\frac{mp}{2} + \Sigma^{-1}\bigl(\bigl(Y - \mu\bigl(\mathrm{II}_m'\otimes \mathrm{I}_p\bigr)\bigr)\bigl(Y - \mu\bigl(\mathrm{II}_m'\otimes \mathrm{I}_p\bigr)\bigr)'\bigr)\biggr)\Sigma^{-1}\,d\Sigma\,\Sigma^{-1}\,d\Sigma \\ &\quad + \operatorname{tr} G(\mathrm{II}_m\otimes \mathrm{I}_p)\,d\mu'(m\Sigma)^{-1}\,d\mu\bigl(\mathrm{II}_m'\otimes \mathrm{I}_p\bigr) + \operatorname{tr}(dG)\,dA \\ &= \operatorname{tr} G(\mathrm{II}_m\otimes \mathrm{I}_p)\,d\mu'(m\Sigma)^{-1}\,d\mu\bigl(\mathrm{II}_m'\otimes \mathrm{I}_p\bigr) + \operatorname{tr}(dG)\,dA \\ &= -\frac{mp}{2}\operatorname{tr}\,d\mu'(m\Sigma)^{-1}\,d\mu + \operatorname{tr} C\cdot P(\mathrm{II}_m\otimes \mathrm{I}_p)\,d\mu'(m\Sigma)^{-1}\,d\mu\bigl(\mathrm{II}_m'\otimes \mathrm{I}_p\bigr) \\ &\quad + \operatorname{tr}(dC)\cdot P\,dA \\ &= -\frac{mp}{2}\operatorname{tr}\,d\mu'(m\Sigma)^{-1}\,d\mu + \operatorname{tr}\bigl(\mathrm{II}_m'\otimes \mathrm{I}_p\bigr)\cdot C\cdot(\mathrm{II}_m\otimes \mathrm{I}_p)\cdot d\mu'(m\Sigma)^{-1}\,d\mu \\ &\quad + \operatorname{tr}(dC)\cdot P\,dA. \end{aligned}$$

Consequently, we have the conclusion as follows.

Let $Y \sim E_{n\times mp}(\mu\cdot(\mathrm{II}_m'\otimes \mathrm{I}_p), m\Sigma)$, and then the maximum likelihood estimations of $\mu$, $\Sigma$ are $\overline{Y}$, $S^2$ if and only if

(1) $G = -\frac{p}{2}I + CP$, $CP = PC'$.
(2) $\operatorname{tr}(\mathrm{II}_m'\otimes \mathrm{I}_p)\cdot C\cdot(\mathrm{II}_m'\otimes \mathrm{I}_p)\cdot d\mu'(m\Sigma)^{-1}\,d\mu + \operatorname{tr}(dC)\cdot P\,dA < \frac{mp}{2}\operatorname{tr}\,d\mu'(m\Sigma)^{-1}\,d\mu$ $\forall d\mu, d\Sigma$.

Since $Y \sim E_{n\times mp}(\mu\cdot(\mathrm{II}_m'\otimes \mathrm{I}_p), m\Sigma)$ if and only if $C = 0$, $Y$ may not be a normal random matrix in general.

For instance, let $C = cI$, $c = -\frac{p}{2}\operatorname{tr} PA$, $P = P_{\mathrm{II}_m}\otimes \mathrm{I}_p$, and $A = (Y - \mu\cdot(\mathrm{II}_m'\otimes \mathrm{I}_p))'(m\Sigma)^{-1}(Y - \mu\cdot(\mathrm{II}_m'\otimes \mathrm{I}_p))$.

Then the distribution density of $Y$ is

$$L(Y,\mu,\Sigma) = C|\Sigma|^{-\frac{mp}{2}}\exp\biggl\{-\frac{p}{2}\bigl(\operatorname{tr} A + \operatorname{tr}^2 PA\bigr)\biggr\},$$

where $C = (\int |\Sigma|^{-\frac{mp}{2}}\exp\{-\frac{p}{2}(\operatorname{tr} A + \operatorname{tr}^2 PA)\}\,dY)^{-1}$.

Since $G = -\frac{p}{2}I + CP$ and $CP = PC'$, and $(m\operatorname{tr}\,d\mu'(m\Sigma)^{-1}\,d\mu)(-\frac{p}{2}\operatorname{tr} PA) + (\operatorname{tr} P\,dA)\times(-\frac{p}{2}\operatorname{tr} P\,dA) < 0 < \frac{mp}{2}\operatorname{tr}\,d\mu'(m\Sigma)^{-1}\,d\mu$ so that the maximum likelihood estimations of $\mu$, $\Sigma$ are $\overline{Y}$, $S^2$, then the distribution of $Y$ is not a normal distribution.

## 4 Conclusion

In this paper, we proved that if the distribution of an $n\times mp$ random matrix $Y$ is a left ellipsoidal distribution with parameter $\mu^* = \mu(\mathrm{II}_m'\otimes \mathrm{I}_p)$, $\Sigma^* = m\Sigma$, and $Y_1, Y_2, \ldots, Y_m$ are independent and identical distributions, and $Y \sim N_{n\times mp}(\mu^*, \Sigma^*)$, then the maximum likelihood estimations of $\mu$, $\Sigma$ are $\overline{Y}$, $S^2$. If $Y_1, Y_2, \ldots, Y_m$ are not independent and identical distributions, then the maximum likelihood estimations of $\mu$, $\Sigma$ are $\overline{Y}$, $S^2$. We used the matrix differential method to deduce that only and only if $Y \sim E_{n\times mp}(\mu \cdot (\mathrm{II}'_m \otimes \mathrm{I}_p), m\Sigma)$, which needs to satisfy two conditions.

**Competing interests**
The authors declare that they have no competing interests.

**Authors' contributions**
All authors contributed equally to the manuscript, read and approved the final manuscript.

**Author details**
[1]College of Mathematics and Information Science, Henan Normal University, Xinxiang, 453007, China. [2]School of Finance and Statistics, East China Normal University, Shanghai, 200062, China.

**Acknowledgements**
This work is supported by the Foundation and Frontier Technology Research Project of Henan Province of China (102300410264) and the Education Department Fundamental Research Project of Henan Province of China (2010A110010). The authors would like to thank the anonymous referees for several useful interesting comments and suggestions about the paper.

Received: 22 March 2013 Accepted: 15 October 2013 Published: 08 Nov 2013

**References**

1. Zhang, Y: Some Distribution Results for Spherical Distribution and Their Applications, 405. Econometrics and Statistics Colloquium, Rosenwald (1980)
2. Wang, L: On the moments of a kind of elliptical matrix distributions. Chinese J. Appl. Probab. Statist. **36**(3), 231-242 (1987)
3. Zhang, Y, Fang, K, Chen, H: The matrix distribution family regression. J. Math. Phys. **5**(3), 341-353 (1985)
4. Ahsanullan, M: Some characterizations of the bivariate normal distribution. Metrika **32**, 215-218 (1985)
5. Nguyen, TT: A note on matrix variate normal distribution. Technical report, Dept. of Math. and Statist., Bowling Green State University, Bowling Green, Ohio (1993)
6. Zhang, Y, Fang, K: In Multivariate Statistical Analysis, 81. Science Press (1982)
7. Zhang, Y: Matrix differential - the means to deducing exact distribution. J. Math. Res. Expo. **3**(1), 179-184 (1983)
8. Anderson, TW, Fang, KT: Maximum likelihood estimators and likelihood ratio criteria for multivariate elliptically contoured distributions. Technical report No. 1, ARO Contract DAAG 29-82-K-0156, Department of Statistics, Stanford University (1982)
9. Wang, Y, Zhang, Y: The maximum likelihood estimations for left ellipsoidal distribution. J. Henan Norm. Univ. Nat. Sci. **36**(4), 17-19 (2008)

10.1186/1029-242X-2013-502
**Cite this article as:** Wang and Zhang: The maximum likelihood estimations for a type of left ellipsoidal distribution. *Journal of Inequalities and Applications* 2013, 2013:502