ESTIMATION IN MULTIVARIATE ELLIPTICALLY CONTOURED LINEAR MODELS

Tonghui Wang
New Mexico State University
Las Cruces, New Mexico, 88003

ABSTRACT. For a multivariate elliptically contoured $n \times p$ random matrix $Y \sim \text{MEC}_{n \times p} (\mu, \sum_{j=1}^{k} V_j \otimes \Sigma_j, \phi)$, explicit formulae for the best quadratic unbiased estimators of the $\Sigma_j$'s and the MINQUE of $\Sigma \equiv \sum_{j=1}^{k} c_j \Sigma_j$, as well as the empirical estimator and maximum likelihood estimator of $\mu$ in $S_1 S_2'$ are obtained; here $\otimes$ is the Kronecker product, $S_1$ and $S_2$ are known linear subspaces of $\mathbb{R}^n$ and $\mathbb{R}^p$ respectively, and $S_1 S_2'$ is the linear space spanned by $xy'$ with $x \in S_1$ and $y \in S_2$. The distributions of the estimators are also studied by using a version of Cochran's Theorem. The multivariate components of variance models and certain growth curve models are special cases of this linear model. For illustration, our results are applied to the balanced multivariate mixed models.

1. Introduction.

Estimation of covariance matrices in multivariate components of variance models, under normal assumptions, has been discussed by many authors in recent years, see, e.g., Mathew, Niyogi and Sinha (1994), Kelly and Mathew (1993, 1994), and Calvin and Dykstra (1991), etc. In this paper, we are going to discuss some estimation problems on both mean and covariance matrices in multivariate components of variance models under the assumption of elliptically contoured settings, which include normal and certain non-normal distributions.

Let $M_{n \times p}$ be the set of all $n \times p$ matrices over the real field $\mathbb{R}$ and $\mathbb{R}^n$ be $M_{n \times 1}$. Let $\mu \in M_{n \times p}$ and $\phi$ be a function of $[0, \infty)$ into the complex field. Then a random matrix $Y$ in $M_{n \times p}$ is said to have a multivariate elliptically contoured distribution with parameters $\mu$, $\Sigma_Y$, and $\phi$ (written as $Y \sim \text{MEC}_{n \times p} (\mu, \Sigma_Y, \phi)$) if the characteristic function, $\hat{Y}$, of $Y$ is given by

$$\hat{Y}(T) = e^{i(T, \mu) \phi(u)}, \quad u \equiv (T, \Sigma_Y(T)), \quad T \in M_{n \times p},$$

1991 Mathematics Subject Classification. Primary 62J05; Secondary 62H12, 62H15.

Key words and phrases. Best quadratic unbiased estimator, Cochran theorem, empirical estimator, generalized Wishart distribution, growth curve model, MINQUE, multivariate components of variance model, multivariate elliptically contoured distribution.

- date d'arrivée : le 21.06.95
- version définitive : le 12.04.96

Typeset by AMS-TEX
where $\mu$ is the mean of $Y$, $\Sigma_Y = \sum_{j=1}^{k} V_j \otimes \Sigma_j$ is proportional to the covariance of $Y$, $V_j \in \mathbb{M}_{n \times n}$ and $\Sigma_j \in \mathbb{M}_{p \times p}$ are nonnegative definite matrices, and $(\cdot, \cdot)$ is the trace inner product: $(B, C) = \text{tr}(B'C) = \text{tr}(C'B)$. Here $\Sigma_Y$ in (1.1) is regarded as a linear transformation of $\mathbb{M}_{n \times p}$ into $\mathbb{M}_{n \times p}$. If $\Sigma_Y$ is viewed as a matrix in $\mathbb{M}_{np \times np}$, then $\Sigma_Y(T)$ can be written as $\Sigma_Y \text{vec} T$. (vec transforms the matrix $T$ with rows $t_1', t_2', \ldots, t_n'$ into the column vector $(t_1', t_2', \ldots, t_n')'$.)

In the special case $\phi(u) = e^{-u^2/2}$, $u \geq 0$, $Y \sim N_{n \times p}(\mu, \Sigma_Y)$, the multivariate normal distribution with mean $\mu$ and covariance $\Sigma_Y$. When $p = 1$ and $\Sigma_Y = \sigma^2 A$, $MEC_{n \times p}(\mu, \sigma^2 A, \phi)$ is referred to as the univariate elliptically contoured distribution and is written as $EC_n(\mu, \sigma^2 A, \phi)$. In general, it is not necessary that $Y$ have a density; if it does, then the density has the form

$$h(Y) = |\Sigma_Y|^{-\frac{1}{2}} f((Y - \mu, \Sigma_Y^{-1}(Y - \mu)))$$ (1.2)

where $f(.)$ is a non-increasing and continuous function. A special case of above definition of MEC distributions is given in Anderson and Fang (1982) where $\Sigma_Y$ is of the form of $A \otimes \Sigma$ with positive definite $A$ and $\Sigma$.

In this paper, we assume that

$$\mu \in S_1 S_2'$$ (1.3)

and

$$\Sigma_1, \Sigma_2, \ldots, \Sigma_k$$ are linearly independent, (1.4)

where $S_1$ and $S_2(\neq \{0\})$ are known linear subspaces of $\mathbb{R}^n$ and $\mathbb{R}^p$ respectively, and $S_1 S_2'$ is the linear span of $xy'$ with $x \in S_1$ and $y \in S_2$.

Even if $S_2 = \mathbb{R}^p$ and $\mu = ZB$ for some known $Z \in \mathbb{M}_{n \times q}$ and unknown $B \in \mathbb{M}_{q \times p}$, the above $Y$ still extends the multivariate components of variance models from the normal setting to the elliptically contoured setting; (see e.g., Anderson et al. (1986), and Rao and Kleffe (1988).) Also if $k = 1$ and $S_1$, $S_2$ are the image spaces for some known $Z_1 \in \mathbb{M}_{n \times q}$, $Z_2 \in \mathbb{M}_{p \times r}$, i.e. $\mu = Z_1 \Theta Z_2'$ for some unknown $\Theta \in \mathbb{M}_{q \times r}$, then $Y \sim MEC_{n \times p}(Z_1 \Theta Z_2', V \otimes \Sigma, \phi)$ becomes a growth curve model; (see e.g., Khatri (1988) and von Rosen (1991).) The model given in Khatri and Shah (1981) is the special case where $Y$ is normally distributed. For general multivariate elliptically contoured distributions, see Fang and Anderson (1990) or Fang and Zhang (1990).
In Section 3, an explicit formula for the best quadratic unbiased estimators, \( \hat{\Sigma}_i(Y) \)'s, of \( \Sigma_i \)'s are obtained under the assumptions (1.3) and (1.4). The MINQUE, minimum norm quadratic unbiased estimators, of \( \Sigma = \sum_{j=1}^{k} c_j \Sigma_j \), the linear combination of \( \Sigma_j \)'s is obtained. The empirical unbiased estimator and maximum likelihood estimator of \( \mu \) are also investigated.

In section 4, we investigate the distributions of \( \hat{\Sigma}_i(Y) \)'s by using a version of Cochran’s theorem.

2. Preliminaries.

For any \( T \in M_{n \times p} \), we let \( T', \text{Im} T, r(T), \) and \( T^+ \) denote, respectively, the transpose, the image set (column space), the rank, and the Moore-Penrose inverse of \( T \). The set of all \( n \times n \) nonnegative definite (n.n.d.) matrices is denoted by \( \mathcal{N}_n \) and the set of all \( n \times n \) symmetric matrices is denoted by \( \mathcal{S}_n \). For any \( T \in \mathcal{N}_n \) and \( \alpha > 0 \), \( T^\alpha \) denotes the \( \alpha \)th n.n.d. root of \( T \) and \( T^{-\alpha} \) denotes the \( \alpha \)th n.n.d. root of \( T^+ \). Then \( T^0 \) will be \( T^\alpha T^{-\alpha} = T^{-\alpha} T^\alpha \); in particular, \( T^0 = T^+ T = TT^+ \).

Also for \( A \in M_{n \times m} \) and \( B \in M_{p \times q} \), the Kronecker product, \( A \otimes B \), is taken as the linear transformation of \( M_{m \times q} \) into \( M_{n \times p} \) such that for all \( T \) in \( M_{m \times q} \),

\[
(A \otimes B)(T) = ATB'.
\]  

(2.1)

In vector form, \( (A \otimes B)\text{vec}T = \text{vec} (ATB') \), where \( \text{vec}T \) is the columnized vector of \( T \).

Let \( S \) be a linear subspace of \( \mathbb{R}^n \); let \( P_S \) be the orthogonal projection of \( \mathbb{R}^n \) onto \( S \) and \( P_{S^\perp} = I_n - P_S \). Then we have the following result.

Lemma 2.1. Let \( V \in \mathcal{N}_n \), \( X \in M_{n \times p} \), \( S = \text{Im}X \), \( T = V + XUX' \) for any \( U \in \mathcal{N}_p \) such that \( S \subseteq \text{Im} T \), and \( T^- \) be any symmetric generalized inverse of \( T \). Then

\[
T = X(X'T^-X)^{-X'} + V(P_{S^\perp}VP_{S^\perp})^+V.
\]  

(2.2)

Lemma 2.2. Let \( Y \sim MEC_{n \times p}(\mu, \Sigma_Y, \phi) \). Suppose that the first four order moments of \( Y \) exist. Then

(a) \( E(Y) = \mu \),

(b) \( \text{Cov}(Y) = -2\phi'(0) \sum_{j=1}^{k} V_j \otimes \Sigma_j \), and

(c) \( E(\otimes^4 Y) = \otimes^4 \mu - 2\phi'(0) \Delta_1 + 4\phi''(0) \Delta_2 \),
where
\[
\Delta_1 = V \otimes \mu \otimes \mu + \mu \otimes \mu \otimes V + (K_{n,n^2} \otimes I_n)(\mu \otimes \mu \otimes V)(K_{p^2,p} \otimes I_p)
\]
\[+ \mu \otimes V \otimes \mu + (I_n \otimes K_{nn} \otimes I_n)(V \otimes \mu \otimes \mu + \mu \otimes \mu \otimes V)(I_p \otimes K_{pp} \otimes I_p),
\]
\[
\Delta_2 = V \otimes V + (I_n \otimes K_{nn} \otimes I_n)(V \otimes V)(I_p \otimes K_{pp} \otimes I_p)
\]
\[+ (K_{n,n^2} \otimes I_n)(V \otimes V)(K_{p^2,p} \otimes I_p),
\]
\[V = \sum_{j=1}^k \text{vec} V_j(\text{vec} \Sigma_j)', \otimes^m Y \text{ denotes } Y \otimes Y \cdots \otimes Y \text{ (m times), and } K_{pn} \text{ in } M_{np \times np} \text{ is called commutation matrix defined by}
\]
\[K_{pn} \text{vec } T = \text{vec}(T'), \quad T \in M_{n \times p}.
\]
The proof of Lemma 2.2 is similar to the proof of Theorem 2.3 of Wong and Wang (1992), taking \( V = \sum_{j=1}^k \text{vec} V_j(\text{vec} \Sigma_j)' \) instead of \( V = \text{vec} A(\text{vec} \Sigma)' \).

**Lemma 2.3.** Suppose that \( Y \sim MEC_{n \times p}(\mu, \sum_{j=1}^k V_j \otimes \Sigma_j, \phi) \). Let \( W_1 \) and \( W_2 \) be symmetric matrices in \( M_{n \times n} \). Then
\[
(i) \quad E(Y'W_iY) = \mu'W_i\mu - 2\phi'(0) \sum_{j=1}^k \text{tr}(W_iV_j) \Sigma_j, \quad i = 1, 2, \quad \text{and}
\]
\[
(ii) \quad \text{Cov}(Y'W_1Y, Y'W_2Y) = 4\phi''(0)\nabla_1 + 4[\phi''(0) - (\phi'(0))^2]\nabla_2 - 2\phi'(0)\nabla_3,
\]

where
\[
\nabla_1 = \sum_{j,t=1}^k \{ \text{tr}(V_jW_1V_tW_2)(K_{pp} + I_p^2)(\Sigma_j \otimes \Sigma_t) \},
\]
\[
\nabla_2 = \sum_{j,t=1}^k \text{tr}(W_1V_j)\text{tr}(W_2V_t)\text{vec} \Sigma_j(\text{vec} \Sigma_t)',
\]

and
\[
\nabla_3 = \sum_{j=1}^k \{ \mu'W_1V_jW_2\mu \otimes \Sigma_j + \Sigma_j \otimes \mu'W_1V_jW_2\mu
\]
\[+ K_{pp}(\mu'W_2V_jW_1\mu \otimes \Sigma_j + \Sigma_j \otimes \mu'W_2V_jW_1\mu) \}.
\]

Lemma 2.3 can be proved using Lemma 2.2. The details of proofs of Lemma 2.2 and Lemma 2.3 are given in Wang (1995).

3. Estimation of parameters.

For the multivariate elliptically contoured linear model defined in (1.1) with restrictions (1.3) and (1.4), let
\[
f_i(W) = E(\text{tr}(Y'WY - \Sigma_i))^2, \quad W \in \mathcal{N}_n.
\]

Then \( Y'W_iY \) is called the best quadratic unbiased estimator of \( \Sigma_i \) if \( W_i \) minimizes \( f_i(W) \) subject to \( E(Y'WY) = \Sigma_i \).
Lemma 3.1. Consider the model $Y$ given in (1.1) with restriction (1.4). For each $i = 1, 2, \ldots, k$, let $W_i$ be a symmetric matrix in $M_{n \times n}$. Then $Y' W_i Y$ is unbiased of $\Sigma_i$ for all possible values of $\mu \in S_1 S_2$ if and only if

(a) $\mu' W_i \mu = 0$ and

(b) $\text{tr}(W_i V_j) = -\delta_{ij} (2\phi'(0))^{-1}$, $j = 1, 2, \ldots, k$, where $\delta_{ij}$ is the Kronecker symbol.

In particular, if $W_i \in \mathcal{N}_n$ and $\mu = Z_1 \Theta Z_2'$ with $S_1 = \text{Im} Z_1$ and $S_2 = \text{Im} Z_2 \neq \{0\}$, then (a) and (b) are equivalent to

(a1) $W_i Z_1 = 0$,

(b1) $W_i V_j = 0$, $j \neq i$, $j = 1, 2, \ldots, k$, and

(c1) $\text{tr}(W_i V_i) = -1/(2\phi'(0))$.

Proof. The 'if' part is obvious. Suppose that $Y' W_i Y$ is an unbiased estimator of $\Sigma_i$ for all possible values of $\mu \in S_1 S_2$. Then by (i) of Lemma 2.3,

$$\mu' W_i \mu - 2\phi'(0) \sum_{j=1}^k \text{tr}(W_i V_j) \Sigma_j = \Sigma_i,$$

which implies (a) and

$$-2\phi'(0) \sum_{j \neq i}^k \text{tr}(W_i V_j) \Sigma_j - [2\phi'(0) \text{tr}(W_i V_i) + 1] \Sigma_i = 0.$$

So by (1.4), (b) follows.

Under the conditions on $W_i \in \mathcal{N}_n$, (a) is equivalent to

$$W_i \mu = 0, \quad \text{i.e.,} \quad WZ_1 \Theta Z_2' = 0.$$

Thus $0 = \text{vec}(WZ_1 \Theta Z_2') = (WZ_1 \otimes Z_2)\text{vec} \Theta$ for all possible $\Theta \in M_{q \times r}$, which implies $WZ_1 \otimes Z_2 = 0$. So (a1) follows as $Z_2 \neq 0$. It is easy to see that (b) is equivalent to (b1) and (c1) because $W_i$ is n.n.d. \(\square\)

For each $i \in \{1, 2, \ldots, k\}$, let $P_i$ be the orthogonal projection of $\mathbb{R}^n$ onto $(Z_1 : V_1 : \cdots : V_i-1 : V_{i+1} : \cdots : V_k)$. Assume that

$$\langle I_n - P_i \rangle V_i (I_n - P_i) \neq 0, \quad i = 1, 2, \ldots, k. \quad (3.2)$$

Remark 3.1. The assumption (3.2) was given by Rao and Kleffe (1988) for the univariate components of variance models and we apply it to multivariate components of variance models.
Theorem 3.2. Suppose that $Y \sim MEC_{n \times p}(\mu, \sum_{j=1}^{k} V_j \otimes \Sigma_j, \phi)$ with restrictions (1.4) and (3.2). For each $i \in \{1, \cdots, k\}$, let

$$\hat{\Sigma}_i(Y) = \alpha_i Y'W_i Y, \quad W_i = [(I_n - P_i) V_i (I_n - P_i)]^+, \quad (3.3)$$

where $\alpha_i = -1/(2m_i \phi'(0))$ and $m_i = r(W_i)$. Then $\hat{\Sigma}_i(Y)$ is a best quadratic unbiased estimator of $\Sigma_i$.

Proof. Let $Y'WY$ with $W \in \mathcal{N}_n$ be an unbiased estimator of $\Sigma_i$. Then by Lemma 3.1, (a1) - (c1) hold. By (a1) and (b1), (ii) of Lemma 2.3 is reduced to

$$\text{Cov}(Y'WY) = 4\phi''(0) \text{tr}(WV_iWV_i)(K_{pp} + I_p^2)(\Sigma_i \otimes \Sigma_i)$$

$$+ 4[\phi''(0) - (\phi'(0))^2](\text{tr}(WV_i))^2 \text{vec} \Sigma_i'(\text{vec} \Sigma_i)'.$$  

(3.4)

Thus by (c1) and (3.4), $f_i(W)$ in (3.1) becomes

$$f_i(W) = E[\text{tr}(Y'WY)^2] - 2E[\text{tr}(Y'WY \Sigma_i)] + \text{tr}(\Sigma_i^2)$$

$$= E[\text{tr}(Y'WY)^2] - \text{tr}(\Sigma_i^2) = \text{tr}[\text{Cov}(Y'WY)]$$

$$= 4\phi''(0) \text{tr}(WV_iWV_i) \left[\text{tr}(\Sigma_i^2) + (\text{tr} \Sigma_i)^2\right] + 4 \left[\frac{\phi''(0)}{(\phi'(0))^2} - 1\right] \text{tr}(\Sigma_i^2).$$

(3.5)

Note that $\phi''(0)$ in (3.4) is positive. Since $\text{tr}(\Sigma_i^2) + (\text{tr} \Sigma_i)^2 > 0$ and $\left[\frac{\phi''(0)}{(\phi'(0))^2} - 1\right] \text{tr}(\Sigma_i^2)$ do not depend on $W$, it suffices to minimize $\text{tr}(WV_iWV_i)$ subject to (a1) and (c1) of Lemma 3.1. Let $M_i = I_n - P_i$. Note (a1) and (b1) imply that $W = M_i CM_i$ for some $C \in \mathcal{N}_n$. Following Lagrange, let

$$g(C, \lambda) = \text{tr}(M_iCM_iV_iM_iCM_iV_i) - 2\lambda \left(\text{tr}(M_iCM_iV_i) + \frac{1}{2\phi'(0)} \right) \quad \text{for } C \in \mathcal{N}_n, \lambda \in \mathbb{R}.$$ 

Then we have

$$d g(C, \lambda)(dC, d\lambda) = 2\text{tr}(M_iV_iM_iCM_iV_iM_idC) - 2\lambda \text{tr}(M_iV_iM_idC)$$

$$- 2(\text{tr}(M_iV_iM_iC) + 1/(2\phi'(0))) d\lambda.$$

Setting the coefficient of $dC$ to zero yields

$$M_iV_iM_iCM_iV_iM_i = \lambda M_iV_iM_i. \quad (3.5)$$

From model (1.1), $M_iY \sim MEC_{n \times p}(0, M_iV_iM_i \otimes \Sigma_i, \phi)$. So by (b) of Lemma 2.2, $\text{Im}(M_iY) \subseteq \text{Im}(M_iV_i)$ with probability one. Hence by multiplying (3.5) on the
left by \( Y'(M_i V_i M_i)^+ \) and on the right by \( (M_i V_i M_i)^+ Y \), we get, with probability one,
\[
Y'M_i CM_i Y = \lambda Y'(M_i V_i M_i)^+ Y,
\]
i.e.,
\[
Y'WY = \lambda Y'W_i Y.
\]
(3.6)

Note that
\[
\text{tr} (W V_i) = \lambda \text{tr} (M_i (V_i M_i)^+ V_i) = \lambda \text{tr} (M_i (V_i M_i)^+ M_i V_i M_i) = \lambda m_i.
\]
(3.7)

By (c1) of Lemma 3.1 and (3.6), we have \( \lambda = -1/[2m_i \phi'(0)] = \alpha_i \). For \( dC \neq 0 \),
\[
\partial^2 g(C, \lambda)(dC, dC) = 2 \text{tr} (M_i dCM_i V_i M_i dCM_i V_i) = 2 \text{tr} \left( V_i^{1/2} M_i dCM_i V_i M_i dCM_i V_i^{1/2} \right) > 0,
\]
so that \( g(C, \lambda) \) is minimized at \( (\alpha_i W_i, \alpha_i) \) and hence \( \hat{\Sigma}_i(Y) \) is a best quadratic unbiased estimator of \( \Sigma_i \). □

Remark 3.2. For the univariate case where \( p = 1 \), Rao and Kleffe (1988) proved that each \( \hat{\sigma}_j(Y) \equiv \hat{\Sigma}_j(Y) \) is also a MINQUE of \( \theta_j \equiv \Sigma_j \) if it exists.

Now for model (1.1) with restriction (1.3) and (1.4), suppose that we want to estimate \( \Sigma = \sum_{j=1}^k c_j \Sigma_j \in \mathbb{N}_p \) with an unbiased quadratic estimator \( Y'WY \), using the MINQUE method introduced by Rao (1971).

**Theorem 3.3.** Consider the model (1.1) with restriction (1.3) and (1.4). Then a MINQUE of \( \Sigma \) is given by
\[
\hat{\Sigma}(Y) = \sum_{j=1}^k \lambda_j Y'(P_{S_j^+} V P_{S_j^+})^+ V_j (P_{S_j^+} V P_{S_j^+})^+ Y,
\]
(3.8)

where \( V = \sum_{j=1}^k V_j, \ P_{S_j^+} = I_n - Z_1(Z_1' Z_1)^+ Z_1' \), \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)' \) is any solution of \( H \lambda = -c/(2 \phi'(0)) \), \( c = (c_1, c_2, \ldots, c_k) \), and
\[
H = (h_{ij}) \in M_{k \times k}, \quad \text{with} \quad h_{ij} = \text{tr} [(P_{S_j^+} V P_{S_j^+})^+ V_i (P_{S_j^+} V P_{S_j^+})^+ V_j].
\]

The proof of Theorem 3.3 is similar to the proof of Theorem 5.2.1 of Rao and Kleffe (1988).
Corollary 3.4. In Theorem 3.3, let \( k = 1 \) and \( V = V_1 \). Then the MINQUE of \( \Sigma = \Sigma_1 \) given in (3.11) is reduced to
\[
\hat{\Sigma}(Y) = \lambda Y' (P_{s1}^TV P_{s1}^T)^+ Y,
\]
where \( \lambda = -(2m \phi'(0))^{-1} \) and \( m = r(P_{s1}^TV P_{s1}^T) \).

Remark 3.3. From Theorem 3.2 and Corollary 3.4, we know that the MINQUE in (3.9) and the best quadratic unbiased estimator in (3.3) of \( \Sigma \) are the same for the case where \( k = 1, V = V_1 \), and \( \Sigma = \Sigma_1 \). Also by Lemma 2.1, we obtain that with probability one, the \( \hat{\Sigma}(Y) \) in (3.9) can be rewritten as
\[
\lambda Y'[T^{-} - T^{-}Z_1(Z_1'T^{-}Z_1)^{-}Z_1'T^{-}]Y,
\]
where \( T = V + Z_1 U Z_1' \) for any \( U \in \mathcal{N}_n \) such that \( S_1 \subseteq \text{Im} T \).

Example 3.1. Consider the one-way balanced multivariate mixed model with \( m \) replications in each cell:
\[
Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad i = 1, \ldots, a, j = 1, \ldots, m.
\]
The unobservable \( p \)-dimensional random vectors \( \alpha_1, \ldots, \alpha_m, \epsilon_{11}, \epsilon_{12}, \ldots, \epsilon_{am} \) are uncorrelated with means 0 and covariances
\[
\text{Cov}(\alpha_i) = \Psi_1, \quad \text{Cov}(\epsilon_{ij}) = \Psi_2, \quad i = 1, \ldots, a, j = 1, \ldots, m.
\]
Let \( n = am, e_m = (1,1,\ldots,1)' \in \mathbb{R}^m, J_m = e_m e_m', Y_i = (Y_{i1}, Y_{i2}, \ldots, Y_{im})', \) and \( Y = (Y_1', Y_2', \ldots, Y_m'). \) If the multivariate elliptically contoured distribution is assumed, then the model (3.10) can be written as
\[
Y \sim \text{MEC}_{n \times p}(Z_1 \Theta, (I_a \otimes J_m) \otimes \Psi_1 + I_n \otimes \Psi_2, \phi).
\]
where \( Z_1 = e_n \) and \( \Theta = \mu' \). Let \( \Sigma_1 = m \Psi_1 + \Psi_2 \) and \( \Sigma_2 = \Psi_2 \). Then (3.11) can be rewritten as
\[
Y \sim \text{MEC}_{n \times p}(Z_1 \Theta, V_1 \otimes \Sigma_1 + V_2 \otimes \Sigma_2, \phi)
\]
with \( V_1 = I_a \otimes J_m/m \) and \( V_2 = I_n - I_a \otimes J_m/m \). Note here the assumption (3.2) is satisfied. In fact, \( P_1 = J_n/n + V_2, P_2 = V_1, \)
\[
(I_n - P_1)V_1(I_n - P_1) = V_1 - J_n/n, \quad (I_n - P_2)V_2(I_n - P_2) = V_2.
\]
since $V_1 + V_2 = I_n$, $V_1J_n = J_n$, $V_2J_n = 0$, and $J_n/n$, $V_1$, and $V_2$ are idempotent of rank 1, $a$, and $a(m - 1)$, respectively. Thus we have the following results:

(i) By Theorem 3.2, the best quadratic unbiased estimators of $\Sigma_1$ and $\Sigma_2$ are $\alpha_1 Y'(V_1 - J_n/n)Y$ and $\alpha_2 Y'V_2Y$ respectively, where $\alpha_1 = -1/[2(a - 1)\phi'(0)]$ and $\alpha_2 = -1/[2a(m - 1)\phi'(0)]$. Note that, for the usual normal MANOVA model, $\alpha_1 Y'(V_1 - J_n/n)Y$ is the between-group mean square matrix and $\alpha_2 Y'V_2Y$ is the within-group mean square matrix.

(ii) By Theorem 3.3, the MINQUE of $\Sigma = c_1\Sigma_1 + c_2\Sigma_2$ given in (3.8) is

$$
\hat{\Sigma}(Y) = \lambda_1 Y'P_{S_1}V_1 P_{S_1} Y + \lambda_2 Y'P_{S_2}V_2 P_{S_2} Y = c_1\alpha_1 Y'(V_1 - J_n/n)Y + c_2\alpha_2 Y'V_2Y,
$$

where $P_{S_1} = J_n/n$ and $V = I_n$. Thus the MINQUE of $\Sigma_i$'s are the same as their best quadratic unbiased estimators given in (i) above.

For the estimation of fixed effects $\mu = Z_1\Theta Z_2'$ in model (1.1) with restrictions of (1.3) and (1.4), let $S = S_1S_2'$ with $S_1 = \text{Im} Z_1$ and $S_2 = \text{Im} Z_2$ and $\Sigma_Y = \sum_{j=1}^{k} V_j \otimes \Sigma_j$. Let $K$ be a known matrix such that $K(\Theta)$ is estimable, i.e., there exists a known matrix $L$ such that $K(\Theta) = L(\mu)$. Then with respect to the known $\Sigma_Y$, by Theorem 2.1 of Wong (1993), $L\mu_*(Y)$ with

$$
\mu_*(Y) = [I_{np} - \Sigma_Y (P_{S_1} \Sigma_Y P_{S_1})^+] (Y)
$$

is a best linear unbiased estimator of $L(\mu)$. Since $\Sigma_i$'s are generally not known, one would replace them by their estimators, $\hat{\Sigma}_i(Y)$'s. Thus an empirical estimator of $L(\mu)$ is $L\hat{\mu}_*(Y)$ with

$$
\hat{\mu}_*(Y) = [I_{np} - \hat{\Sigma}_Y (P_{S_1} \hat{\Sigma}_Y P_{S_1})^+] (Y), \quad \hat{\Sigma}_Y = \sum_{j=1}^{k} V_j \otimes \hat{\Sigma}_j(Y).
$$

Note that $L\hat{\mu}_*(Y)$ is also an unbiased estimator of $L(\mu)$ (See, e.g., Khatri and Shah (1981)). So if the best quadratic unbiased estimators of $\Sigma_j$'s exist, one can simply use $\hat{\Sigma}_j(Y)$ given in (3.3) to obtain an unbiased empirical estimator of $L(\mu)$.

Remark 3.4. For the case where $k = 1$, $V = V_1$, and $\Sigma = \Sigma_1$, let $X = Z_1 \otimes Z_2$ and $P_S = X(X'X)^{-1}X'$. Suppose that $\Sigma_Y = V \otimes \Sigma$ is nonsingular. Then by Lemma 2.1,

$$
\mu_*(Y) \sim MEC_{n \times p}(\mu, \Sigma_Y - \Sigma_Y (P_{S_1} \Sigma_Y P_{S_1})^+ \Sigma_Y, \phi) = MEC_{n \times p}(\mu, X(X'\Sigma_Y^{-1}X)^{-1}X', \phi).
$$
Let \( \mu^*(Y) = X (X' \Sigma_Y^{-1} X)^{-1} X' \Sigma_Y^{-1} (Y) \). Then \( \mu^*(Y) \) and \( \mu_*(Y) \) have the same distribution. Note that

\[
\mu^*(Y) = Z_1 (Z'_1 V^{-1} Z_1) - Z'_1 V^{-1} Y \Sigma^{-1} Z_2 (Z'_2 \Sigma^{-1} Z_2) - Z'_2.
\]

With \( K = C \otimes D' \) and \( \hat{\Sigma}(Y) \) given in Remark 3.3, an empirical estimator of \( C \otimes D \) is given by

\[
C (Z'_1 V^{-1} Z_1) - Z'_1 V^{-1} Y \hat{\Sigma}^{-1} Z_2 (Z'_2 \hat{\Sigma}^{-1} Z_2) - D,
\]

which is the same as the one given in (5.6.5) of Srivastava and Khatri (1979).

If the best quadratic unbiased estimator of \( \Sigma_i \) does not exist, we might assume that \( V_1, V_2, \ldots, V_k \) are independent and commutative (See, e.g., Rogers and Young (1975)). Then there exists an orthogonal matrix \( \Gamma \in M_{n \times n} \) such that

\[
(\Gamma \otimes I_{p}) \Sigma Y (\Gamma' \otimes I_{p}) = diag \left( \sum_{j=1}^{k} d_{1j} \Sigma_j, \sum_{j=1}^{k} d_{2j} \Sigma_j, \ldots, \sum_{j=1}^{k} d_{nj} \Sigma_j \right),
\]

where \( \Gamma V \Gamma' \) is diagonal with elements \( \{d_{1j}, d_{2j}, \ldots, d_{nj}\}, j = 1, 2, \ldots, k \). If we let \( \Psi_i = \sum_{j=1}^{k} d_{ij} \Sigma_j \) for \( i = 1, 2, \ldots, n \), then

\[
\Gamma Y \sim MEC_{n \times p} (\Gamma \mu, diag \left( \Psi_1, \Psi_2, \ldots, \Psi_n \right), \phi).
\]

So for each \( i \in \{1, 2, \ldots, n\} \), we may use the MINQUE of \( \Psi_i = \sum_{j=1}^{k} d_{ij} \Sigma_j \) to construct \( \hat{\Sigma}_Y \) in (3.13).

Now for investigating the maximum likelihood estimator of \( \mu \), we assume that \( V_1, V_2, \ldots, V_k \) generate a commutative quadratic subspace, \( B \) of \( S_n \). Then by Lemma 6 of Seely (1971), there exists a basis \( R_1, R_2, \ldots, R_t \) for \( B \) such that \( R_i \) is idempotent and \( R_i R_j = 0 \) for \( i \neq j \). Thus

\[
\Sigma Y = \sum_{j=1}^{k} V_j \otimes \Sigma j = \sum_{j=1}^{k} \left( \sum_{i=1}^{t} s_{ij} R_i \right) \otimes \Sigma_j = \sum_{i=1}^{t} R_i \otimes \Lambda_i,
\]

where \( \Lambda_i = \sum_{j=1}^{k} s_{ij} \Sigma_j, i = 1, 2, \ldots, t \). Suppose that \( \Sigma Y \) and \( \Lambda_i 's \) are nonsingular. Then \( \sum_{i=1}^{t} R_i = I_n \) and

\[
\Sigma Y^{-1} = \sum_{i=1}^{t} R_i \otimes \Lambda_i^{-1}.
\]
Theorem 3.5. For model (1.1) with density function given in (1.2) and the restriction (1.3), suppose that (3.14) holds for nonsingular $\Lambda_i$'s. Then with $n > p$, the maximum likelihood estimator (MLE) of $\mu$ is given by

$$\hat{\mu}(Y) = Z_1(Z'_1Z_1)^{-1}Z'_1YZ_2(Z'_2Z_2)^{-1}Z'_2.$$ 

Proof. By (1.2), the likelihood function of $Y$ is given by

$$L(\mu, \Sigma_1, \ldots, \Sigma_k; Y) = \prod_{j=1}^{k} V_j \otimes \Sigma_j^{-\frac{1}{2}} f \left( Y - \mu, \left( \sum_{j=1}^{k} V_j \otimes \Sigma_j \right)^{-1} (Y - \mu) \right).$$

By (3.14),

$$L(\mu, \Sigma_1, \ldots, \Sigma_k; Y) = \prod_{i=1}^{t} R_i \otimes \Lambda_i^{-\frac{1}{2}} f \left( Y - \mu, \sum_{i=1}^{t} R_i (Y - \mu) \Lambda_i^{-1} \right).$$

Note that in (3.15), $\langle Y - \mu, R_i (Y - \mu) \Lambda_i^{-1} \rangle \geq 0$ for each $i = 1, 2, \ldots, t$. By the monotonicity of $f(\cdot)$, for any given positive definite $\Lambda_i$, $L(\mu, \Sigma_1, \ldots, \Sigma_k; Y)$, as a function of $\mu$, arrives at its maximum at

$$\langle Y - \mu, R_i (Y - \mu) \Lambda_i^{-1} \rangle = 0, \quad i = 1, 2, \ldots, t,$$

which is equivalent to

$$R_i (Y - \mu) = 0, \quad i = 1, 2, \ldots, t.$$  

Since $\mu = Z_1 \otimes Z'_2 \in S_1 S'_2$, we obtain from (3.16),

$$\sum_{i=1}^{t} P_{S_1} R_i Y P_{S_2} = \sum_{i=1}^{t} P_{S_1} R_i \mu P_{S_2},$$

which is equivalent to

$$P_{S_1} Y P_{S_2} = P_{S_1} \mu P_{S_2} = \mu,$$

as $\sum_{i=1}^{t} R_i = I_n$. Therefore the MLE of $\mu$, for nonsingular $\Lambda_i$, is given by

$$\hat{\mu}(Y) = P_{S_1} Y P_{S_2}$$

and the desired result follows. □
4. Distributions of Estimators.

Let $X$ be a random matrix in $M_{n \times p}$ such that $X \sim MEC_{n \times p}(0, I_n \otimes \Sigma, \phi)$, where $\Sigma \in \mathcal{N}_p$. Partition $X$ into $\ell + 1$ parts:

$$X = (X'_1, X'_2, \ldots, X'_\ell, X'_{\ell+1})',$$

where each $X'_i \in M_{m_i \times p}$ and $\sum_{i=1}^{\ell} m_i = n - m_{\ell+1} \geq 0$. Then the distribution of $(X'_1X_1, X'_2X_2, \ldots, X'_\ell X_\ell)$, denoted by $GW_p(m_1, m_2, \ldots, m_\ell; m_{\ell+1}; \Sigma; \phi)$, is called the generalized Wishart distribution with parameters $m_1, m_2, \ldots, m_\ell, m_{\ell+1}, \Sigma,$ and $\phi$. In particular, if $\ell = 1$ and $m = m_1$, then

$$X'_1X_1 \sim GW_p(m; n - m; \Sigma; \phi). \quad (4.1)$$

Note that for the normal case where $X \sim N_{n \times p}(0, I_n \otimes \Sigma), GW_p(m_1, m_2, \ldots, m_\ell; m_{\ell+1}; \Sigma; \phi)$ is nothing but a joint distribution of independent Wishart random matrices $W_p(m_1, \Sigma), W_p(m_2, \Sigma), \ldots, W_p(m_\ell, \Sigma)$ and no longer depends on $m_{\ell+1}$. For the case where $\Sigma > 0$, the joint probability density function of $(X'_1X_1, X'_2X_2, \ldots, X'_\ell X_\ell)$ and its properties were obtained by Anderson and Fang (1982). We shall obtain a version of the Cochran theorem for our model (1.1) with restriction (1.4).

**Theorem 4.1.** Suppose that $Y \sim MEC_{n \times p}(\mu, \sum_{j=1}^{k} V_j \otimes \Sigma_j, \phi)$ with restriction (1.4) and $Pr(Y = \mu) < 1$, the probability of $Y = \mu$ is less than 1. Let $i = 1, 2, \ldots, \ell$, $W_i \in \mathcal{N}_n$, $\Sigma = \sum_{j=1}^{k} c_j \Sigma_j \in \mathcal{N}_p$, $m_i \in \{1, 2, \ldots\}$, and $Q_i(Y) = (Y - \mu)'W_i(Y - \mu)$. Then

$$(Q_1(Y), Q_2(Y), \ldots, Q_\ell(Y)) \sim GW_p(m_1, m_2, \ldots, m_\ell; m_{\ell+1}; \Sigma; \phi) \quad (4.2)$$

if and only if for any distinct $i, i' = 1, 2, \ldots, \ell$ and $j, j' = 1, 2, \ldots, k$,

(a) $c_j W_i V_j W_i = c_j W_i V_j W_i$,
(b) $W_i V_j W_i V_j = c_j W_i V_j$, \quad $r(W_i V_j) = m_i$, \quad and
(c) $W_i V_j W_i = 0$,

where (a) and (b) hold only for those $c_j, c_j' \neq 0$.

**Proof.** Suppose that (4.2) holds. Then by Theorem A of Wang and Wong (1995), there exists an $A \in \mathcal{N}_n$ such that for any distinct $i, i' = 1, 2, \ldots, \ell$,

(i) $(W_i \otimes I_p)(\Sigma_Y - A \otimes \Sigma)(W_i \otimes I_p) = 0$,
(ii) $AW_i AW_i = AW_i$, \quad $r(AW_i) = m_i$,
(iii) $W_i AW_i = 0$, \quad and
(iv) \((W_i \otimes I_p) \Sigma_Y (W_i' \otimes I_p) = 0\),

where \(Y \sim MEC_{n \times p}(\mu, \Sigma_Y, \phi)\). Thus, with \(\Sigma_Y = \sum_{j=1}^k V_j \otimes \Sigma_j\), it suffices to show that (i) - (iv) imply (a) - (c). By the assumption on \(\Sigma\), (i) and (iv) are reduced, respectively, to

\[
\sum_{j=1}^k W_i V_j W_i \otimes \Sigma_j - W_i AW_i \otimes \sum_{j=1}^k c_j \Sigma_j = 0
\]  

(4.3)

and

\[
\sum_{j=1}^k W_i V_j W_i' \otimes \Sigma_j = 0.
\]  

(4.4)

So by (1.4), (4.4) implies that \(W_i V_j W_i' = 0\), which is (c). Also by (1.4), (4.3) is reduced to

\[
W_i V_j W_i = c_j W_i AW_i.
\]  

(4.5)

Thus if \(c_j, c_j' \neq 0\), then we obtain from (4.5) that

\[
c_j W_i V_j W_i = c_j c_j' W_i AW_i = c_j W_i V_j W_i,
\]

which is (a). By (ii) and (4.5),

\[
W_i V_j W_i V_j W_i = c_j^2 W_i AW_i AW_i = c_j^2 W_i AW_i = c_j W_i V_j W_i,
\]

which implies the first condition of (b). Since

\[
m_i = r(W_i A) \geq r(W_i AW_i) = r(c_j^{-1} W_i V_j W_i) \geq r(c_j^{-1} W_i V_j W_i V_j) = r(W_i V_j)
\]

and

\[
m_i = r(W_i A) = r(W_i AW_i A) \leq r(W_i AW_i) = r(c_j^{-1} W_i V_j W_i) \leq r(W_i V_j),
\]

the second condition of (b) follows.

Now suppose that (a) - (c) hold. Let \(X \sim MEC_{n \times p}(0, I_n \otimes \Sigma, \phi)\). Since \(\Sigma \in \mathcal{N}_p\), without loss of generality, we may assume that \(c_1 > 0\). Thus by (i) of Lemma 2.3, \(\text{tr}(W_i V_i) > 0\) for all \(i = 1, 2, \ldots, \ell\). Now consider the random matrices \(U_i = W_i^{1/2} (Y - \mu)\) and \(U_{s_i} = c_i^{-1/2} W_i^{1/2} V_i^{1/2} X, i = 1, 2, \ldots, \ell\). Then by (a) and (c),

\[
(U_1', U_2', \ldots, U_\ell')' \sim MEC_{n \times p}(0, \Sigma_U, \phi).
\]
and

\[(U'_{s1}, U'_{s2}, \ldots, U'_{s\ell})' \sim MEC_{\nu \times p}(0, \Sigma_{U*}, \phi),\]

where

\[
\Sigma_U = \begin{pmatrix}
W_1^{\frac{1}{2}} \otimes I_p \\
W_2^{\frac{3}{2}} \otimes I_p \\
\vdots \\
W_{\ell}^{\frac{3}{2}} \otimes I_p
\end{pmatrix}
\left(\sum_{j=1}^{k} V_j \otimes \Sigma_j \right)
\begin{pmatrix}
W_1^{\frac{1}{2}} \otimes I_p \\
W_2^{\frac{3}{2}} \otimes I_p \\
\vdots \\
W_{\ell}^{\frac{3}{2}} \otimes I_p
\end{pmatrix}
\]

\[= \text{diag} \left( \sum_{j=1}^{k} W_1^{\frac{3}{2}} V_1, \sum_{j=1}^{k} W_2^{\frac{3}{2}} V_2, \ldots, \sum_{j=1}^{k} W_{\ell}^{\frac{3}{2}} V_{\ell} \right) \otimes \Sigma \]

and

\[
\Sigma_{U*} = c_1^{-1}
\begin{pmatrix}
W_1^{\frac{1}{2}} V_1^{\frac{1}{2}} \otimes I_p \\
W_2^{\frac{3}{2}} V_1^{\frac{1}{2}} \otimes I_p \\
\vdots \\
W_{\ell}^{\frac{3}{2}} V_1^{\frac{1}{2}} \otimes I_p
\end{pmatrix}
(1_n \otimes \Sigma)(V_1^{\frac{1}{2}} W_1^{\frac{1}{2}} \otimes I_p, V_1^{\frac{3}{2}} W_2^{\frac{1}{2}} \otimes I_p, \ldots, V_{\ell}^{\frac{3}{2}} W_{\ell}^{\frac{1}{2}} \otimes I_p)
\]

\[= c_1^{-1} \text{diag} \left( W_1^{\frac{3}{2}} V_1 W_1^{\frac{1}{2}} \otimes \Sigma, W_2^{\frac{3}{2}} V_2 W_2^{\frac{1}{2}} \otimes \Sigma, \ldots, W_{\ell}^{\frac{3}{2}} V_{\ell} W_{\ell}^{\frac{1}{2}} \otimes \Sigma \right).\]

Since for each i,

\[
\sum_{j=1}^{k} W_i V_j W_i \otimes \Sigma_j = c_1^{-1} W_i V_i W_i \otimes \sum_{j=1}^{k} c_j \Sigma_j = c_1^{-1} W_i V_i W_i \otimes \Sigma,
\]

\[
\Sigma_U = c_1^{-1} \text{diag} \left( W_1^{\frac{3}{2}} V_1 W_1^{\frac{1}{2}} \otimes \Sigma, W_2^{\frac{3}{2}} V_2 W_2^{\frac{1}{2}} \otimes \Sigma, \ldots, W_{\ell}^{\frac{3}{2}} V_{\ell} W_{\ell}^{\frac{1}{2}} \otimes \Sigma \right).
\]

So \( \Sigma_U = \Sigma_{U*} \) and therefore by definition (1.1),

\[(U'_1, U'_2, \ldots, U'_\ell)' \overset{d}{=} (U'_{s1}, U'_{s2}, \ldots, U'_{s\ell})', \quad (4.6)\]

where \( x \overset{d}{=} y \) means that \( x \) and \( y \) have the same distribution. Let

\[g(u_1, u_2, \ldots, u_\ell) = (u'_1 u_1, u'_2 u_2, \ldots, u'_\ell u_\ell), \quad u_i \in M_{n \times p}.\]

Since \( g \) is a Borel function, we obtain from (4.6),

\[(U'_1 U_1, U'_2 U_2, \ldots, U'_\ell U_\ell) \overset{d}{=} (U'_{s1} U_{s1}, U'_{s2} U_{s2}, \ldots, U'_{s\ell} U_{s\ell}),\]
i.e.,

\[(Q_1(Y), Q_2(Y), \ldots, Q_\ell(Y)) \overset{d}{=} (X'\tilde{W}_1X, X'\tilde{W}_2X, \ldots, X'\tilde{W}_\ell X),\]

(4.7)

where \(\tilde{W}_i = c_i^{-1}V_i^{1/2}W_iV_i^{1/2}\), \(i = 1, 2, \ldots, \ell\). By (b) and (c), \(\tilde{W}_i\) is idempotent of rank \(m_i\) and \(\tilde{W}_i\tilde{W}_i' = 0\) for all distinct \(i, i'\). So there exists an orthogonal matrix \(\Gamma \in M_{n \times n}\) such that

\[\Gamma'\tilde{W}_i\Gamma = \text{diag}(0, \ldots, 0, I_{m_i}, 0, \ldots, 0), \quad i = 1, 2, \ldots, \ell.\]

Let \(X_* = \Gamma X\). Then \(X_* \sim MEC_{n \times p}(0, I_n \otimes \Sigma, \phi)\). Thus with

\[X_* = (X'_1, X'_2, \ldots, X'_{t+1}), \quad X_i \in M_{m_i \times p},\]

\[(X'\tilde{W}_1X, X'\tilde{W}_2X, \ldots, X'\tilde{W}_\ell X) \overset{d}{=} (X'_1\Gamma'\tilde{W}_1\Gamma X_*, X'_2\Gamma'\tilde{W}_2\Gamma X_*, \ldots, X'_{t+1}\Gamma'\tilde{W}_\ell \Gamma X_*),\]

(4.8)

By (4.7) and (4.8),

\[(Q_1(Y), Q_2(Y), \ldots, Q_\ell(Y)) \overset{d}{=} (X'_1X_1, X'_2X_2, \ldots, X'_{t+1}X_\ell),\]

proving (4.2). \(\square\)

**Corollary 4.2.** In Theorem 4.1, let \(k = 1\), \(V = V_1\), and \(\Sigma_1 = \Sigma\), i.e., \(Y \sim MEC_{n \times p}(\mu, V \otimes \Sigma, \phi)\). Then

\[(Q_1(Y), Q_2(Y), \ldots, Q_\ell(Y)) \sim GW_p(m_1, m_2, \ldots, m_\ell; m_{\ell+1}; \Sigma; \phi)\]

if and only if for any distinct \(i, j = 1, 2, \ldots, \ell,\)

(a1) \(W_iVW_jV = W_iV, \quad r(W_iV) = m_i; \quad \text{and}\)

(b1) \(W_iVW_j = 0.\)

Anderson and Fang (1982) obtained Corollary 4.2 for the case where \(V = I_n\) and \(\Sigma\) is positive definite.

**Theorem 4.3.** In Theorem 3.2, assume that for each \(i = 1, 2, \ldots, k\), the best quadratic unbiased estimator, \(\hat{\Sigma}_i(Y)\), of \(\Sigma_i\) exists. Then we have the following results:

(i) \(\alpha_i^{-1}\hat{\Sigma}_i(Y) = YW_iY \sim GW_p(m_i; n - m_i; \Sigma_i; \phi),\)

where \(W_i = [(I_n - P_i)V_i(I_n - P_i)]^+\) and \(P_i\) is given in (3.3).
(ii) For any distinct \( i, i' \in \{1, 2, \ldots, k\} \), \( \hat{\Sigma}_i(Y) \) and \( \hat{\Sigma}_{i'}(Y) \) are uncorrelated if 
\[
\phi''(0) - (\phi'(0))^2 = 0.
\]
They are independent random matrices if \( Y \sim N_{n \times p}(\mu, \Sigma_Y) \) with 
\[
\Sigma_Y = \sum_{j=1}^{k} V_j \otimes \Sigma_j.
\]
(iii) If each \( \Sigma_i \) is positive definite and \( m_i > p \), then 
\[
(Q_{*1}(Y), Q_{*2}(Y), \ldots, Q_{*k}(Y)) \sim GW_p(m_1, m_2, \ldots, m_k; n - \sum_{j=1}^{k} m_j; p, \phi),
\]
where 
\[
Q_{*i}(Y) = \alpha_{-1}^{-1} \Sigma_i^{-1} Y' W_i Y \Sigma_i^{-1/2} \quad \text{with} \quad W_i = [(I_n - P_i) V_i (I_n - P_i)]^+.
\]

Proof. (i) Let \( M_i = I_n - P_i \). Then \( M_i \mu = 0 \) and \( M_i V_j = 0 \) for all \( j \neq i \), \( j = 1, 2, \ldots, k \). Thus for any distinct \( i, j \),
\[
W_i \mu = 0 \quad \text{and} \quad W_i V_j = 0. \quad (4.9)
\]

By (3.4),
\[
Y' W_i Y = Y' (M_i V_i M_i)^+ Y = (Y - \mu)' (M_i V_i M_i)^+ (Y - \mu) = (Y - \mu)' W_i (Y - \mu).
\]

By Theorem 4.1, it suffices to show that, with \( \Sigma_i = \Sigma \), (a) and (b) of Theorem 4.1 hold. Note here \( c_i = 1 \) and \( c_j = 0 \), \( j \neq i \), \( j = 1, 2, \ldots, k \). Thus we need only show that
\[
W_i V_i W_i V_i = W_i V_i \quad \text{and} \quad r(W_i V_i) = m_i.
\]

In fact, we have
\[
W_i V_i W_i V_i = (M_i V_i M_i)^+ V_i (M_i V_i M_i)^+ V_i
\]
\[
= (M_i V_i M_i)^+ M_i V_i M_i (M_i V_i M_i)^+ V_i = W_i V_i
\]
and
\[
m_i = r(W_i) \geq r(W_i V_i) \geq r(W_i V_i W_i) = r(W_i) = m_i,
\]
the desired result follows.

(ii) By (4.9) and (ii) of Lemma 2.3, the covariance of \( \hat{\Sigma}_i(Y) \) and \( \hat{\Sigma}_{i'}(Y) \) is reduced to
\[
\text{Cov} \left( \hat{\Sigma}_i(Y), \hat{\Sigma}_{i'}(Y) \right) = 4 \alpha_i \alpha_{i'} [\phi''(0) - (\phi'(0))^2] \text{tr} (W_i V_i) \text{tr} (W_{i'} V_{i'}) \text{vec} \Sigma_i (\text{vec} \Sigma_{i'})'
\]
\[
= \frac{\phi''(0) - (\phi'(0))^2}{(\phi'(0))^2} \text{vec} \Sigma_i (\text{vec} \Sigma_{i'}).'
\]
So if $\phi''(0) - (\phi'(0))^2 = 0$, then $\text{Cov} \left( \hat{\Sigma}_i(Y), \hat{\Sigma}_j(Y) \right) = 0$, i.e., they are uncorrelated. Furthermore, if $Y \sim N_{n \times p}(\mu, \sum_{j=1}^{k} V_j \otimes \Sigma_j)$, then by Theorem 2.2 of Wong et al. (1991), $\hat{\Sigma}_i(Y)$ and $\hat{\Sigma}_j(Y)$ are independent.

(iii) Let $U_i = W_i^{-\frac{1}{2}} Y \Sigma_i^{-\frac{1}{2}}$. Then by (4.9),

$$U_i \sim MEC_{n \times p}(0, W_i^{-\frac{1}{2}} V_i W_i^{-\frac{1}{2}} \otimes I_p, \phi) = MEC_{n \times p}(0, W_i^0 \otimes I_p, \phi)$$

and

$$(U_1', U_2', \ldots, U_k') \sim MEC_{n_k \times p}(0, \text{diag}(W_1^0 \otimes I_p, W_2^0 \otimes I_p, \ldots, W_k^0 \otimes I_p), \phi)$$

where $W_i^0 = W_i W_i^+$. Let $X \sim MEC_{n \times p}(0, I_n \otimes I_p, \phi)$ and $U_{si} = W_i^0 X$. Then

$$(U_{s1}, U_{s2}, \ldots, U_{sk}) \overset{d}{=} (U_{s1}', U_{s2}', \ldots, U_{sk}')$$

and hence

$$(U_1' U_1, U_2' U_2, \ldots, U_k' U_k) \overset{d}{=} (U_{s1} U_{s1}, U_{s2} U_{s2}, \ldots, U_{sk} U_{sk})$$

$$\overset{d}{=} (X' W_1^0 X, X' W_2^0 X, \ldots, X' W_k^0 X).$$

Since for any distinct $i, j$, $W_i^0 W_j^0 = 0$, $r(W_i^0) = m_i$, and $W_i^0 W_j^0 = 0$, i.e, (a1) and (b1) of Corollary 4.2 hold, we obtain by Corollary 4.2,

$$(X' W_1^0 X, X' W_2^0 X, \ldots, X' W_k^0 X) \sim GW_p(m_1, m_2, \ldots, m_k; n - \sum_{j=1}^{k} m_j; I_p; \phi).$$

(4.11)

Note that

$$(Q_{s1}(Y), Q_{s2}(Y), \ldots, Q_{sk}(Y)) \overset{d}{=} (U_1' U_1, U_2' U_2, \ldots, U_k' U_k).$$

(4.12)

The desired result follows from (4.10) (4.11).

Remark 4.1. Note that from (ii) of Theorem 4.3, if $\phi''(0) - (\phi'(0))^2 \neq 0$, then $\hat{\Sigma}_i(Y)$ and $\hat{\Sigma}_j(Y)$ are correlated. For example, let $U$ be an $n \times p$ random matrix such that $\text{vec} U \overset{d}{=} v(np)$, the uniform random vector over the $np$-dimensional unit sphere. By a result of Wong and Wang (1992), $\phi'(0) = (np)^{-1}$ and $\phi''(0) = (np(np+2))^{-1}$. Thus $\text{Cov}(\hat{\Sigma}_i(Y), \hat{\Sigma}_j(Y)) = 2(np)^{-1} \text{vec} I_n(\text{vec} I_p)'$.

**Acknowledgements**

The author would like to thank the referee and Professor Nguyen, H. T. for their helpful comments which led to the improvement of this paper.
REFERENCES