An extended growth curve model is considered which, among other things, is useful when linear restrictions exist on the mean in the ordinary growth curve model. The maximum likelihood estimators consist of complicated stochastic expressions. It is shown how, by the aid of fairly elementary calculations, the dispersion matrix for the estimator of the mean and the expectation of the estimated dispersion matrix are obtained. Results for Wishart, inverted Wishart, and inverse beta variables are utilized. Additionally, some asymptotic results are presented.

1. INTRODUCTION

In this paper the object is to discuss some properties of the maximum likelihood estimators in a multivariate linear normal model introduced by von Rosen [4]. The model is an extension of the ordinary growth curve model [2] and is, for example, useful when linear restrictions on the parameters describing the mean in the ordinary growth curve model exist. In the next we recall the model and restate two results presented by von Rosen [4].

DEFINITION. Multivariate linear normal model with mean \( \sum_{i=1}^{m} A_i B_i C_i \), referred to as MLNM(\( \sum_{i=1}^{m} A_i B_i C_i \)). Let \( X: p \times n, A_i: p \times m_i, m_i \leq p, \)
\( B_i: m_i \times k_i, C_i: k_i \times n, \Sigma: p \times p \) p.d. where \( \rho(C_i) + p \leq n \) and \( R(C_{m}) \subseteq \cdots \subseteq R(C_1) \). The columns of \( X \) are independently \( p \)-variate normally distributed with an unknown dispersion matrix \( \Sigma \) and \( E(X) = \sum_{i=1}^{m} A_i B_i C_i \), where the \( A \)'s and \( C \)'s are known and the \( B \)'s unknown.

Received July 10, 1989; revised March 17, 1990.


Key words and phrases: extended growth curve model, moments of estimators, inverted Wishart.
In the definition as well as in the sequel \( p(\cdot) \) and \( \mathbf{R}(\cdot) \) denote the rank and range space, respectively, and p.d. stands for positive definite. It is worth observing that instead of \( \mathbf{R}(C'_m) \subseteq \mathbf{R}(C'_{m-1}) \subseteq \cdots \subseteq \mathbf{R}(C'_1) \) we may assume \( \mathbf{R}(A_m) \subseteq \mathbf{R}(A_{m-1}) \subseteq \cdots \subseteq \mathbf{R}(A_1) \), since a reparametrization and a linear transformation in the latter case leads us back to a model of the type given by the definition.

Moreover, for arbitrary matrices \( G \) and \( F \), \( G^- \) denotes an arbitrary \( g \)-inverse in the sense of \( GG^- = G \) and \( F^9 \) any matrix of full rank spanning \( \mathbf{R}(F)^\perp \), the orthogonal complement to \( \mathbf{R}(F) \). Expectation of a matrix is defined elementwise and the covariance matrix between matrices with the help of a vectorized form, i.e., \( E[X] = (E[x]) \) and \( C[X, Y] = E[\text{vec}(X) \text{vec}(Y)^\prime] - E[\text{vec}(X)] E[\text{vec}(Y)^\prime] \). The dispersion matrix is given by \( D[X] = C[X, X] \).

**THEOREM 1.1** (von Rosen [3, Theorem 2.1]). Let

\[
P_r = T_{r-1} T_{r-2} T_{r-3} \times \cdots \times T_0, \quad T_0 = I, \ r = 1, 2, \ldots, m + 1
\]

\[
T_i = I - P_i A_i (A_i P_i S_i^{-1} P_i A_i)^{-1} A_i P_i S_i^{-1}, \quad i = 1, 2, \ldots, m
\]

\[
F_j = C_{j-1}(C_{j-1} C_j)^{-1} C_{j-1}(I - C_j(C_j C_j)^{-1} C_j), \quad C_0 = I, \ j = 1, 2, \ldots, m
\]

\[
S_i = \sum_{j=1}^i P_j X F_j F_j^\prime X^\prime P_j^\prime, \quad i = 1, 2, \ldots, m.
\]

Assuming \( S_1 \) to be p.d.,

\[
n\widehat{\Sigma} = \left( X - \sum_{i=1}^m A_i \hat{B}_i C_i \right) \left( X - \sum_{i=1}^m A_i \hat{B}_i C_i \right)^\prime = S_m + P_{m+1} X C_m^\prime (C_m C_m^\prime)^{-1} C_m X^\prime P_{m+1}
\]

and

\[
\widehat{E[X]} = \sum_{i=1}^m (I - T_i) X C_i (C_i C_i)^{-1} C_i = \sum_{i=1}^m (P_i - P_{i+1}) X C_i (C_i C_i)^{-1} C_i.
\]

The expressions for \( n\widehat{\Sigma} \) and \( \widehat{E[X]} \) will be discussed in this paper. They seem to be very natural when maximizing the likelihood, but since \( S_i \) is a stochastic matrix, the above given expressions are rather complicated. Note that \( \widehat{E[X]} \) is an unbiased estimator of \( E[X] \) which follows since \( XC_i \) is independent of \( P_i \) and \( P_{i+1} \), and \( \mathbf{R}(P_i') = \mathbf{R}(A_1: A_2: \cdots: A_{i-1})^\perp \). \( P_i \) is a projector where an inner product has been estimated by aid of \( S_i \). Other unbiased estimators for \( E[X] \) are, of course, available. For example, we
always obtain unbiased estimators as long as we estimate the inner product
with a function which is independent of $X'C_i$. To choose between different
estimators is similar to the problem of choosing between different unbiased
estimators for $\Sigma$. For the latter problem comments are given in Section 5.

The starting point of this paper is an observation which illuminates the
underlying stochastics (Lemma 2.1) and then, by aid of a suitable chosen
system of notations, we derive in Section 3 $D[\hat{E}[X]]$, in Section 4, $E[n\hat{\Sigma}]$,
and in Section 5, an unbiased estimator for $\Sigma$ which is a function of the
maximum likelihood estimator. Some asymptotic properties are discussed
in Section 7.

We believe that the MLNM($\sum_{i=1}^m A_i B_i C_i$) has many practical applica-
tions. Verbyla and Venables [5] have considered a model with mean
structure $\sum_{i=1}^m A_i B_i C_i$ but without the nested condition $R(C_m) \subseteq
R(C_{m-1}) \subseteq \cdots \subseteq R(C_1)$. As noted by the authors, the model belongs to a
class of models called seemingly unrelated regression models. Moreover,
Verbyla and Venables [5] gave some comments on the model discussed in
this paper and, among others, indicated how to estimate the parameters of
the mean. They also remarked on the importance of the model and gave
some examples. Other references for the MLNM($\sum_{i=1}^m A_i B_i C_i$) can be
found in von Rosen [4]. In von Rosen [4], vaguely speaking, the nested
condition was used for two reasons; first, it seems to be a necessary and
sufficient condition for obtaining explicit solutions to the likelihood equa-
tions in multivariate linear normal models and, second, as mentioned
above, the model may be a base for discussing interesting applications.

Already, for the MLNM($A_1 B_1 C_1$), the ordinary growth curve model,
the distributions are difficult to apply. For the general case we note
(follows from Lemma 2.1 in this paper) that the distribution for $E[X]$ con-
ists of a sum of $m$ non-independent stochastic variables where each
variable is of the same type as $E[X]$ in the MLNM($A_1 B_1 C_1$). Thus it
seems difficult to master the distributional problems and therefore, in this
paper, instead of discussing distributions, confidence regions, etc., we are
focusing on moments which will be of help when interpreting the maximum
likelihood estimators. Furthermore, in the more general model, discussed
by Verbyla and Venables [5], i.e., the unnested model, it was noted that
it appears difficult to find standard errors for the estimated parameters. In
this paper, it is shown that under the nested condition explicit moments
can be obtained. The moment relations may also be useful and illuminating
when considering alternatives to the maximum likelihood estimators. For
example, $c_0 S_1$ or if $m > 1$, $c_1 S_1 + c_2 S_2 + c_3 A_1 (A_1' \hat{\Sigma}^{-1} A_1)^{-1} A_1$, for an
appropriate choice of constants $c_0$, $c_1$, $c_2$, and $c_3$, are both unbiased
estimators of $\Sigma$. 
2. Preparation

In this section some important relations are established and notions are given for a fairly rudimentary treatment of obtaining $D[E[X]]$ and $E[n\hat{X}]$. All results will mainly rest on the next lemma which makes the stochastics in the maximum likelihood estimators interpretable.

**Lemma 2.1.** Let $G, w$, be defined by

\[ G_{r+1} = G_r(G_r^{-1} + W_r), \quad G_0 = I \]

\[ W_r = X(I - C_r(C_r)^{-1} C_r) X' \sim W_p(\Sigma, n - \rho(C_r)). \]

Then

\[ P_r^{-1} S_r^{-1} P_r = G_{r-1} (W_r G_{r-1})^{-1} G_r^{-1}. \]

**Proof.** A proof based on an induction argument will be given. $r - 1$ is trivial and for $r = 2$ we obtain

\[ P_2^{-1} S_2^{-1} P_2 = T_2^{-1} S_2^{-1} T_1 \]

\[ = T_1 S_1^{-1} T_1 - T_1 S_1^{-1} T_1 X F_2 \]

\[ \times (F_2' X' T_1 S_1^{-1} T_1 X F_2 + I)^{-1} F_2' X' T_1 S_1^{-1} T_1 \]

\[ = G_1((G_1 W_1 G_1)^{-1} \]

\[ -(G_1 W_1 G_1)^{-1} G_1' X F_2' X' G_1'(G_1 W_2 G_1)^{-1}) G_1' \]

\[ = G_1(G_1 W_2 G_1)^{-1} G_1', \]

since $X F_2' X' = W_2 - W_1$. Now suppose that the relation holds for $r - 1$. First note that

\[ P_{r-1}^{-1} S_{r-1}^{-1} P_r = P_{r-1}^{-1} (S_{r-1}^{-1} - S_{r-1}^{-1} P_{r-1} A_{r-1} ) \]

\[ \times (A_{r-1} P_{r-1}^{-1} S_{r-1}^{-1} P_{r-1} A_{r-1} )^{-1} A_{r-1} P_{r-1}^{-1} S_{r-1}^{-1} ) P_{r-1} \]

\[ = G_{r-1}((G_{r-1} W_{r-1} G_{r-1})^{-1} G_{r-1}' \]

Hence

\[ P_r^{-1} S_r^{-1} P_r = P_r((S_r^{-1} - S_r^{-1} P_r X F_r (F_r' X' P_r' S_r^{-1} P_r X F_r + I)^{-1} F_r' X' P_r' S_r^{-1} )) P_r \]

\[ = G_{r-1}((G_{r-1} W_{r-1} G_{r-1})^{-1} - (G_{r-1} W_{r-1} G_{r-1})^{-1} \]

\[ \times G_{r-1} X F_r F_r' X' G_{r-1}(G_{r-1} W_{r-1} G_{r-1})^{-1} G_{r-1} \]

\[ = G_{r-1}((G_{r-1} W_{r-1} G_{r-1})^{-1} G_{r-1}' \]

completing the proof. $\square$
Note that \( R(G_r) = R(A_1 : A_2 : \cdots : A_s)^{-1} \) and that \( P_r S_r^{-1} P_r \) in principle follows an inverted Wishart distribution. Moreover, \( G_r \) consists of \( p \) rows and \( q_r \) columns, where \( q_r = p - \rho(A_{r-1} : A_r) = p - \rho(A_1 : A_2 : \cdots : A_{r-1}) \), \( q_1 = p - \rho(A_1) \), and \( q_0 = p \). In Lemma 2.2 several relations and definitions have been collected which frequently will be used in the calculations of the next sections. The proofs are omitted since they include straightforward matrix manipulations. \( \Sigma^{1/2} \) stands for the symmetric square root of \( \Sigma \).

**Lemma 2.2.** Let the non-singular matrix \( H_r : q_r \times q_r \) and the orthogonal matrix \( (\Gamma') = ((\Gamma_1') : (\Gamma_2')')', \Gamma : q_{r-1} \times q_{r-1} \Gamma_1' : q_r \times q_{r-1} \), be defined by

\[
A_1^0 = H_1 (I_{p - \rho(A_1)} : 0) \Gamma_1^1 \Sigma^{-1/2} = H_1 \Gamma_1^1 \Sigma^{-1/2}
\]

\[
(G_{r-1} A_r)^0 H_{r-1} - H_r (I_{q_r} : 0) \Gamma_r' - H_r \Gamma_1', \quad \Gamma_0 = I, H_0 = I.
\]

Furthermore, let

\[
V' = \Gamma_1^1 \Gamma_1'^{-1} \times \cdots \times \Gamma_1^1 \Sigma^{-1/2} W, \Sigma^{-1/2} (\Gamma_1')'
\]

\[
\times \cdots \times (\Gamma_1^1)^{-1}(\Gamma')' \sim W_{q_{r-1}}(I, n - \rho(C_r))
\]

\[
V' = \Gamma_1^1 \Gamma_1'^{-2} \times \cdots \times \Gamma_1^1 \Sigma^{-1/2} W, \Sigma^{-1/2} (\Gamma_1')'
\]

\[
\times \cdots \times (\Gamma_1^1)^{-2}(\Gamma'^{-1})', \quad V_1^1 - \Sigma^{-1/2} W_1 \Sigma^{-1/2}
\]

\[
Q_r = W_r G_r (G' W_r G_r)^{-1} G_r',
\]

\[
Z_{r,s} = \Gamma_1^1 \Gamma_1'^{-2} \times \cdots \times \Gamma_1^1 \Sigma^{-1/2} Q_r Q_{r+1} \times \cdots \times Q_s,
\]

\[
s \geq r \geq 2, \quad Z_{1,s} = Q_1 Q_2 \times \cdots \times Q_s
\]

\[
M_r = (I : (V_1')^{-1} V_{12}),
\]

where \( V_{11} \) and \( V_{12} \) refer to an ordinary partition of \( V' \). The following relations may be proved:

(i) \( G_1' = H_r \Gamma_r' \Gamma_1'^{-1} \times \cdots \times \Gamma_1^1 \Sigma^{-1/2} \)

(ii) \( P_r = Z_{1,r-1} (Z_{1,0} = I) \)

(iii) \( \Gamma_1^1 \Gamma_1'^{-2} \times \cdots \times \Gamma_1^1 \Sigma^{-1/2} Q_r = (\Gamma'^{-1})' M_r \Gamma_r' \Gamma_1'^{-1} \times \cdots \times \Gamma_1^1 \Sigma^{-1/2} \)

(iv) \( V_{11} = V_{11}' + \Gamma_1^1 \Gamma_1'^{-2} \times \cdots \times \Gamma_1^1 \Sigma^{-1/2} X F, F' X' \Sigma^{-1/2} (\Gamma_1')' \times \cdots \times (\Gamma_1^1)^{-1}(\Gamma_1')' \)

(v) \( \Sigma^{1/2}(\Gamma_1')' \times \cdots \times (\Gamma_1^1)^{-1}(\Gamma_1')' \Gamma_1^1 \Gamma_1'^{-1} \times \cdots \times \Gamma_1^1 \Sigma^{1/2} \)

\[= \Sigma G_r (G', \Sigma G_r)^{-1} G_1' \Sigma \]

(vi) \( \Sigma^{1/2}(\Gamma_1')' \times \cdots \times (\Gamma_1^1)^{-1}(\Gamma_1')' \Gamma_2^1 \Gamma_1'^{-1} \times \cdots \times \Gamma_1^1 \Sigma^{1/2} \)

\[= \Sigma G_{r-1} (G_r, \Sigma G_r)^{-1} G_{r-1} \Sigma - \Sigma G_r (G', \Sigma G_r)^{-1} G' \Sigma. \]
In the next section it turns our that some moment relations of quadratic forms in normally distributed variables are needed as well as the expectation for inverted Wishart variables and inverted beta variables. The multivariate beta distribution is defined in the following manner: let $A \sim W_p(\Sigma, n_1)$ be independent of $B \sim W_p(\Sigma, n_2)$. Put $A + B = T'T$, where $T$ is a unique upper triangular matrix with positive diagonal elements. Then there exists a symmetric non-singular matrix $U$ such that $A = T'UT$. $U$ is said to follow a multivariate beta distribution with parameters $\frac{1}{2}n_1$ and $\frac{1}{2}n_2$ (denoted $B_p(\frac{1}{2}n_1, \frac{1}{2}n_2)$). Observe that $U$ is independent of $T$ and if $\Gamma$ is an orthogonal matrix $\Gamma U \Gamma'$ and $U$ are identically distributed.

**Lemma 2.3.** Let $V \sim W_p(I, n)$ and let $Q$, of proper dimension, be independent of $V_{12}$. Furthermore, let $U \sim B_p(\frac{1}{2}n_1, \frac{1}{2}n_2)$, let $V', V^{-1}$, and $q_{\gamma-1}$ be the same as in Lemma 2.2 and let $M$ signify an arbitrary matrix of proper dimension.

(i) $E[V_{12}QV_{12}] = E[\text{tr}(QV_{11})]I$

(ii) $E[V^{-1}] = 1/(n - p - 1)I$ if $n - p - 1 > 0$

(iii) $E[M(M'V^{-1}M)M'V'] = E[M(M'VM)M']$

(iv) $E[M(M'V^{-1}M)M'V^{-1}V^{-1}M(M'V^{-1}M)M']$

(v) $E[V^{-1}M(M'V^{-1}M)M'V^{-1}]$

(vi) $E[U^{-1}] = (n_1 + n_2 - p - 1)/(n_1 - p - 1)I$ if $n - p - 1 > 0$

(vii) $E[\{(V'_{11})^{-1} h((V'_{11})^{-1})\}] = c_{\gamma-1} E[\{(V'_{11})^{-1} h((V'_{11})^{-1})\}]$ if $n - p - 1 > 0$, $\gamma-1 = (n - p(C_{\gamma-1}) - q_{\gamma-1} - 1)/(n - p(C_{\gamma-1}) - q_{\gamma-1} - 1)$ and $h(\cdot)$ is any measurable function of $(V'_{11})^{-1}$.

**Proof.** We just prove (vi) and (vii), since the other relations are either well known or follow more or less by straightforward manipulations. (vi) is based on the fact that $\Gamma U^{-1}\Gamma'$ and $U^{-1}$ follow the same distribution implying that $E[U^{-1}] = cI$ for some constant $c$ which then in turn is determined by the relation $(n_1 + n_2 - p - 1)/(n_1 - p - 1)I = E[(W')^{-1}]E[(W + W^2)^{-1}]^{-1}$, where $W' \sim W_p(\Sigma, n_1)$ and, independently, $W^2 \sim W_p(\Sigma, n_2)$. In (vii), remembering Lemma 2.2(iv), we apply that $V'_{11} = T'T$ and $V_{11}^{-1} = T'UT$, where $T$ is a unique upper triangular matrix with positive diagonal elements, independent of $U$ and $U \sim B_{q-1}(\frac{1}{2}(n - p(C_{\gamma-1})), \frac{1}{2}(\rho(C_{\gamma-1}) - \rho(C_{\gamma-1})))$.
3. The Dispersion of \( \hat{E}[X] \)

In this section we will show how to derive the dispersion matrix \( D[\hat{E}[X]] \), where \( \hat{E}[X] \) is given by Theorem 1.1. Before starting a rather technical presentation we observe that under suitable conditions for uniqueness we could also have discussed \( E[\hat{B}_i] \), \( D[\hat{B}_i] \), and \( D[\hat{B}_r] \), where an expression for \( \hat{B}_i \) is presented in von Rosen [4, Theorem 2.1].

Since \( E[\hat{E}[X]] = E[X] \) and

\[
D[\hat{E}[X]] = D[\hat{E}[X] - E[X]] = D \left[ \sum_{i=1}^{m} (I - T_i)(X - E[X]) C_i'(C_i C_i')^{-1} C_i \right]
\]

we consider \( D[(I - T_i)(X - E[X]) C_i'(C_i C_i')^{-1} C_i] \) and \( C[(I - T_i)(X - E[X]) C_i'(C_i C_i')^{-1} C_i, (I - T_i)(X - E[X]) C_i'(C_i C_i')^{-1} C_i] \) for arbitrary \( r \) and \( s \).

Hence, \( D[\hat{E}[X]] \) is obtained by summing up to \( m \) which is left to the reader. Let \( \otimes \) stand for the right Kronecker product defined by \( A \otimes B = (a_{ij}B) \) for \( A = (a_{ij}) \).

**Theorem 3.1.** Let \( c_i \) be defined in Lemma 2.3(vii), \( q_r \) as in Lemma 2.2 and put \( d_{k, r} = c_i c_i \times \cdots \times c_{i, r-1} \) \( i = 1, 2, ..., r-1 \), where \( e = (p - q_r) (n - \rho(C_r) - 1)/((n - \rho(C_r) - q_r - 1)(n - \rho(C_r) - q_r - 1 + p - q_r - 1)) \), \( f_r = (n - \rho(C_r) - 1)/(n - \rho(C_r) - q_r - 1) \), and

\[
K_i = \Sigma G_{i-1} (G_i G_{i-1})^{-1} G_i^{-1} \Sigma - \Sigma G_i (G_i \Sigma G_i)^{-1} G_i^{-1} \Sigma,
\]

\[
L_i = \Sigma G_{i-1} (G_i G_{i-1})^{-1} G_i^{-1} A_i (A_i G_i (G_i \Sigma G_i)^{-1} G_i^{-1} A_i)^{-1} G_i^{-1} A_i
\]

Then

\[
D[(I - T_i)(X - E[X]) C_i'(C_i C_i')^{-1} C_i] = C_i'(C_i C_i')^{-1} C_i \otimes \left( \sum_{i=1}^{r-1} d_{k, r} K_i + f_r L_r \right).
\]

**Proof.** Let

\[
R_{i-1} = A_i (A_i G_{i-1} (G_{i-1} W_i G_{i-1})^{-1} G_{i-1} A_i)^{-1} A_i G_{i-1} (G_{i-1} W_i G_{i-1})^{-1} G_{i-1}
\]

and, since

\[
D[(I - T_i)(X - E[X]) C_i'(C_i C_i')^{-1} C_i] = C_i'(C_i C_i')^{-1} C_i \otimes E[(I - T_i) \Sigma (I - T_i)],
\]

(3.1)
we will discuss, with the help of the relations in Lemma 2.2 and Lemma 2.3,

\[
E[(I - T_r) \Sigma (I - T'_r)] \\
= E[Z_{1,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{1,r-1}] \\
= \Sigma^{1/2}(\Gamma_1) E[M_1 Z_{2,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{2,r-1} M'_1] \Gamma_1 \Sigma^{1/2} \\
= \Sigma^{1/2}(\Gamma_1) E[Z_{2,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{2,r-1}] \Gamma_1 \Sigma^{1/2} \\
+ E[\text{tr}((V'_{11})^{-1} Z_{2,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{2,r-1})] \Sigma^{1/2}(\Gamma_2) \Gamma_1 \Sigma^{1/2}. \quad (3.2)
\]

Note that (3.2) has the nice property of being a recurrence relation and we proceed by deriving the expectation of the trace function in the righthand side of (3.2). Utilizing Lemma 2.2, Lemma 2.3(vii), and the fact that \( V_{11}^2 \) is included in \( Z_{2,r-1} \) and \( R_{r-1} \) but that otherwise these expressions are functionally independent of \( V_{11}^1 \), we obtain

\[
E[\text{tr}((V'_{11})^{-1} Z_{2,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{2,r-1})] \\
= c_1 E[\text{tr}((V''_{11})^{-1} V''_{11} (\Gamma_2) (V''_{11})^{-1} \\
\times Z_{3,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{3,r-1} (V''_{11})^{-1} (\Gamma_2) (V''_{11})^{-1})] \\
= c_1 E[\text{tr}((V''_{11})^{-1} Z_{3,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{3,r-1})] \quad (3.3)
\]

which also is a recurrence relation. Thus, proceeding in the same manner,

\[
E[\text{tr}((V''_{11})^{-1} Z_{2,r-1} R_{r-1} \Sigma R'_{r-1} Z'_{2,r-1})] \\
= c_1 c_2 \times \cdots \times c_{r-1} E[\text{tr}((V''_{11})^{-1} B_r (B'_r (V''_{11})^{-1} B_r) - B'_r (V''_{11})^{-1})] = d_{r,r}, \quad (3.4)
\]

where

\[
B_r = \Gamma_1^{-1} \Gamma_1^{-2} \times \cdots \times \Gamma_1^{1} \Sigma^{-1/2} A_r
\]

and, for the last equality in (3.4), Lemma 2.3(v) has been applied.

Returning to (3.2) and utilizing the recurrence property together with Lemma 2.2(v) and (vi) and Lemma 2.3(iv), we obtain

\[
E[(I - T_r) \Sigma (I - T'_r)] \\
= \sum_{i=1}^{r-1} d_{i,r} K_i + \Sigma^{1/2}(\Gamma_1) \times \cdots \times (\Gamma_1^{-2}) \times (\Gamma_1^{-1}) E[B_r (B'_r (V''_{11})^{-1} B_r) - B'_r (V''_{11})^{-1}] \Gamma_1^{-1} \Gamma_1^{-2} \times \cdots \times \Gamma_1^{1} \Sigma^{1/2} \\
= \sum_{i=1}^{r-1} d_{i,r} K_i + f_r L_r. \quad (3.6)
\]
THEOREM 3.2. Let $r < s$ and let $K_i$ and $L_i$ be defined in Theorem 3.1;

\[
C[(I - T_r)(X - E[X]) C_r(C_r) - C_r, (I - T_s)(X - E[X]) C_s(C_s) - C_s] = \begin{cases} 
-C'_r(C_s C'_r) - C_s \otimes d_{r,s} K_r & \text{if } r > 1, \\
-C'_r(C_s C'_r) - C_s \otimes (d_{1,s} K_1 + L_1) & \text{if } r = 1.
\end{cases}
\]

\[= \begin{cases} 
C'_r(C_s C'_r) - C_s \otimes d_{r,s} K_r & \text{if } r > 1, \\
C'_r(C_s C'_r) - C_s \otimes (d_{1,s} K_1 + L_1) & \text{if } r = 1.
\end{cases}
\]

**Proof.** Put $P_{C_i} = C'_r(C_s C'_r) - C_f$. The first observation is that $(X - E[X]) P_{C_i}$ can be written $(X - E[X]) (P_{C_i} + P_{C_r} - P_{C_i})$ and that $(X - E[X]) P_{C_i}$ is independent of $I - T_r$, $I - T_s$, and $(X - E[X]) (P_{C_r} - P_{C_i})$. Hence,

\[
C[(I - T_r)(X - E[X]) P_{C_i}, (I - T_s)(X - E[X]) P_{C_i}] = C[(I - T_r)(X - E[X]) P_{C_i}, (I - T_s)(X - E[X]) P_{C_i}] = C'_r(C_s C'_r) - C_s \otimes E[(I - T_r) \Sigma (I - T_s)].
\]

From now on we will discuss $E[(I - T_r) \Sigma (I - T_s)]$ in a manner similar to the case when $r = s$ and it turns out that when $r < s$, $r > 1$, there exists a recurrence property which is somewhat easier to handle. It will be shown that

\[
E[(I - T_r) \Sigma (I - T_s)] = \Sigma^{1/2} (\Gamma^r_1)^Y E[Z_{1,r-1} R_{r-1} \Sigma R_{s-1} Z_{1,s-1} \Sigma] \Gamma^r_1 \Sigma^{1/2}
\]

\[
= \cdots = \Sigma^{1/2} (\Gamma^r_1)^Y \cdot \cdots \cdot (\Gamma^{r-2}_1)^Y (\Gamma^{r-1}_1)^Y E[\Gamma^{r-1}_1 \Gamma^{r-2}_1 \Sigma^{1/2} R_{r-1} \Sigma R_{s-1} Z_{r-1,s-1} \Sigma] \Gamma^{r-1}_1 \Gamma^{r-2}_1 \Sigma^{1/2} \cdot \cdots \cdot \Sigma^{1/2} \cdot 1/2
\]

(3.7)

and this is true if

\[
E[\text{tr}((V^r_{11})^{-1} Z_{k+1,r-1} R_{r-1} \Sigma R_{s-1} Z_{k+1,s-1})] = 0, \quad k + 1 < r < s. \quad (3.8)
\]

However, the expectation in (3.8) equals

\[
c_k c_{k+1} \times \cdots \times c_{r-2} E[\text{tr}((V^r_{11})^{-1} \Gamma^{r-1}_1 \Gamma^{r-2}_1 \\
\times \cdots \times \Gamma^{1}_1 \Sigma^{1/2} (I - Q_r) \Sigma R_{r-1} Z_{r-1,s-1})] = c_k c_{k+1} \times \cdots \times c_{r-1} E[\text{tr}((V^r_{11})^{-1} \Gamma^{r-1}_1 \Gamma^{r-1}_1 \\
\times \cdots \times \Gamma^{1}_1 \Sigma^{1/2} (I - I) \Sigma R_{r-1} Z_{r-1,s-1})] = 0
\]
and, hence, we just have to explore the right-hand side of (3.7) which equals
\[
\Sigma^{1/2}(\Gamma_1') \times \cdots \times (\Gamma_1'-2)'(\Gamma_1'-1)' E[\Gamma_1'^{-1} \Gamma_1'^{-2} \\
\times \cdots \times \Gamma_1' \Sigma^{-1/2}(I - Q_r) \Sigma R_{r-1} Z_{r,s-1} ] \Gamma_1'^{-1} \Gamma_1'^{-2} \times \cdots \times \Gamma_1' \Sigma^{1/2}. 
\]
Doing some calculations, this expression is seen to be identical to (compare with (3.4))
\[
- E[\text{tr}((V_{11}^s)^{-1} \Gamma_1'^{-1} \times \cdots \times \Gamma_1'^{-1} \Sigma^{-1/2} \Sigma R_{r-1} Z_{r+1,s-1})] \Sigma^{1/2}(\Gamma_1')' \\
\times \cdots \times (\Gamma_1'^{-1})'(\Gamma_2')' \Gamma_2'^{-1} \times \cdots \times \Gamma_1' \Sigma^{1/2} \\
= -c_r c_{r+1} \times \cdots \times c_{s-1} E[\text{tr}((V_{11}^s)^{-1} B_s (V_{11}^s)^{-1} B_s) \times \cdots \times (\Gamma_1'^{-1})'(\Gamma_2')' \Gamma_2'^{-1} \Gamma_1'^{-1} \\
\times \cdots \times \Gamma_1' \Sigma^{1/2} = -d_{r,s} K_r,
\]
where \( B_s \) is given by (3.5).

Finally, we look at the case \( r = 1 \),
\[
E(I - T_1) \Sigma (I - T_1^s) \\
= E[Q_1 \Sigma R_{r-1} Z_{1,s-1}] \\
= \Sigma^{1/2}(\Gamma_1')' E[\Gamma_1' \Sigma^{-1/2} \Sigma R_{r-1} Z_{2,s-1}] \Gamma_1' \Sigma^{1/2} \\
+ E[\text{tr}((V_{11}^s)^{-1} \Gamma_1'^{-1} \Sigma^{-1/2} \Sigma R_{r-1} Z_{2,s-1})] \Sigma^{1/2}(\Gamma_1')' \Gamma_1' \Sigma^{1/2} \\
= \Sigma^{1/2}(\Gamma_1')' \Gamma_1' \Sigma^{-1/2} \Sigma E[\Sigma^{-1/2}(\Gamma_1')' \\
\times \cdots \times (\Gamma_1'^{-1})'(\Gamma_1')'^{-1} B_s (B_s)^{-1} B_s - B_s] \Gamma_1'^{-1} \\
\times \cdots \times \Gamma_1' \Sigma^{1/2} + d_{1,s} K_1 = L_s + d_{1,s} K_1,
\]
where the last identity is obtainable from Lemma 2.3(iii).

4. EXPECTATION OF \( \hat{\Sigma} \)

\( \hat{\Sigma} \) was defined in Theorem 1.1 and now \( E[\hat{\Sigma}] \) will be derived. The necessary calculations for doing this will be similar, but easier, to those presented in Section 3 because the matrix \( R_i \) is not included in the expressions below.

**Theorem 4.1.** Let \( K_i \) be defined in Theorem 3.1, \( c_i \) be defined in
Lemma 2.3(vii) and put \( g_{ij} = c_i c_{i+1} \times \cdots \times c_{j-1} q_j / (n - \rho(C_j) - q_j - 1) \) if \( i < j \),
\( g_{ji} = q_j / (n - \rho(C_j) - q_j - 1) \):

\[
E[n\hat{\Sigma}] = \sum_{j=1}^{m} \left( \rho(C_{j-1}) - \rho(C_j) \right)
\times \left( \sum_{i=1}^{j-1} g_{i,j-1} K_i + \Sigma G_{j-1} (G_{j-1}' \Sigma G_{j-1})^{-1} G_{j-1}' \Sigma \right)
\]
\[+ \rho(C_m) \left( \sum_{i=1}^{m} g_{i,m} K_i + \Sigma G_{m} (G_{m}' \Sigma G_{m})^{-1} G_{m}' \Sigma \right). \]

**Proof:** From Theorem 1.1, it follows that the moments for \( P_j X F_j F_j' X' P_j \), \( j = 1, 2, \ldots, m \), and \( P_{m+1} X P_{C_m} X' P_{m+1} \) are needed. Since \( X F_j F_j' X' \) and \( XP_{C_m} X' \) are independent of \( P_j \) and \( P_{m+1} \), respectively, and \( E[X F_j F_j' X'] = (\rho(C_{j-1}) - \rho(C_j)) \Sigma \), \( E[XP_{C_m} X'] = \rho(C_m) \Sigma \), the theorem is verified, since by copying the arguments in Section 3, the following lemma may be proved.

**Lemma 4.1.** \( E[P_j \Sigma P_j'] = \sum_{r=1}^{r-1} g_{i,r-1} K_i + \Sigma G_{r-1} (G_{r-1}' \Sigma G_{r-1})^{-1} G_{r-1}' \Sigma, \)
\( r = 1, 2, \ldots, m + 1. \)

5. **Unbiased Estimators for \( \Sigma \)**

In a standard MANOVA model \( S = 1/(n - \rho(C_1)) S_1 \) is an unbiased estimator for \( \Sigma \) and fairly natural, since \( S \) may be regarded as a restricted maximum likelihood estimator. Therefore, because the MANOVA model is a special case of the MLNM(\( \sum_{i=1}^{m} A_i B_i C_i \)), it may also be relevant to discuss unbiased estimators in the MLNM(\( \sum_{i=1}^{m} A_i B_i C_i \)).

In the previous section, \( \hat{\Sigma} \) was shown to be a biased estimator for \( \Sigma \) in the MLNM(\( \sum_{i=1}^{m} A_i B_i C_i \)) and the bias was a function of the design matrices \( A_1, A_2, \ldots, A_m \). In this section we are going to correct for the bias. Of course, \( S \) is an unbiased estimator for \( \Sigma \) but we will see how to obtain an unbiased estimator which solely is a function of the maximum likelihood estimator \( \hat{\Sigma} \).

**Lemma 5.1.** Let \( k_{ij} = n (g_{i,j-1} - g_{i,j-1}) / (n - \rho(C_i) - q_i) \), where \( g_{i,j-1} \), \( i = 1, 2, \ldots, j - 1 \), is defined in Theorem 4.1 and \( g_{j,j-1} = 1. \):

\[
E[P_j \Sigma P_j'] + E \left[ \sum_{i=1}^{j-1} k_{i,j} A_i (A_i' P_j' \hat{\Sigma}^{-1} P_j A_i)^{-1} A_i' \right] = \Sigma.
\]
Proof. From the proof of Theorem 2.1 in von Rosen [4] follows that 

\[ A'_i P'_i \Sigma^{-1} = n A'_i P'_i S_i^{-1} \]

which, by the aid of some calculations, implies that

\[
E[k_{i,j} A_i (A'_i P'_i \Sigma^{-1} P_i A_i)^{-1} A'_i] = (g_{i+1,j-1} - g_{i,j-1}) (\Sigma - \Sigma G_i (G_i \Sigma G_i)^{-1} G_i \Sigma)
\]

and summing over \( i \) confirms the lemma.

By combining Theorem 4.1, Lemma 4.1, and Lemma 5.1, the next theorem is established.

**Theorem 5.1.**

\[
\hat{\Sigma} + 1/n \sum_{j=1}^{m} (\rho(C_{j-1}) - \rho(C_j)) \sum_{i=1}^{j-1} k_{i,j} A_i (A'_i P'_i \hat{\Sigma}^{-1} P_i A_i)^{-1} A'_i
\]

\[
+ 1/n \rho(C_m) \sum_{i=1}^{m} k_{i,m+1} A_i (A'_i P'_i \hat{\Sigma}^{-1} P_i A_i)^{-1} A'_i
\]

is an unbiased estimator of \( \Sigma \).

Note that by utilizing the ideas in this section it is also fairly easy to write an unbiased estimator of \( D[\hat{E}[X]] \). It is difficult to argue for the superiority of unbiased estimators based on the maximum likelihood estimators of \( \Sigma \), although some reasons exist. To make things clearer, look first at the special case MLNM(\( A_1 B_1 C_1 \)). Hence,

\[
n \hat{\Sigma} - S_1 + V V',
\]

\[
V = X C'_1 (C'_1 C'_1)^{-1} C_1 - A_1 (A'_1 S_1^{-1} A_1) A'_1 S_1^{-1} X C'_1 (C'_1 C'_1)^{-1} C_1.
\]

An unbiased estimator according to Theorem 5.1 is

\[
\hat{\Sigma} + 1/n \rho(C) K_{1,2} A_1 (A'_1 \hat{\Sigma}^{-1} A_1) A'_1.
\]

However, a second unbiased estimator is given, as in the standard MANOVA model, by \( S = 1/(n - \rho(C_1)) S_1 \). Which one should be used? It is easily seen that everything depends on the choice of design matrices \( A_1 \) and \( C_1 \) and this has, to some extent, been confirmed by simulations (von Rosen [3]). It has also been shown that (5.1) may be better than \( S \), as well as that the opposite may hold. However, we believe that in many realistic situations (5.1) is the best choice. This is because \( \hat{\Sigma} \) is build up with the aid of two source of information, namely, the difference between the mean and the observations (\( S_1 \)) and the difference between the mean and the
estimated growth curve \((V'V')\), whereas \(S\) just uses the difference between the mean and the observations.

In the general case we have to go back to von Rosen [4] and look at the scheme for finding the maximum likelihood estimators. The scheme consists of \(m\) steps and, in each step, a projection on a tensor product is made, as well as a projection on its orthogonal complement. However, what is essential in that in the \(m\) steps the information from the design matrices \(A_1, A_2, \ldots, A_m\), by turns, is used and in each step the likelihood increases. The maximum likelihood estimators are of course based on the complete set of design matrices but, in order to construct unbiased estimators for \(\Sigma\), one may consider sunsets. Hence, a whole class of unbiased estimators is obtained. For example, from Lemma 5.1 it follows that

\[
1/(\rho(c_{j-1}) - \rho(c_j)) P_j X F_j F_j' X' P_j' + \sum_{j=1}^{j-1} k_{i,j} A_i (A_i' P_j' \Sigma^{-1} P_i A_i)^{-1} A_i'
\]

generates a class of unbiased estimators which relies on the design matrices up to \(A_{j-1}\) (\(j = 1\) corresponds to the special case \(S\) where no design matrices are included). It is a delicate problem to choose between the estimators and we just note here that the problem is closely connected with the choice of models, i.e., the choice of constant \(m\) in \(E[X] = \sum_{i=1}^{m} A_i B_i C_i\). However, since the unbiased estimator in Theorem 5.1 takes care of information from all \(A_i\)'s we believe that it is advantageous to apply this estimator in most realistic situations. By some cumbersome algebra it seems possible to obtain dispersion matrices for the various estimators but once again all results depend on the structures in the \(A_i\)'s and \(C_i\)'s. In von Rosen [3] such calculations have been performed for the MLNM\((A_1 B_1 C_1)\).

6. Application

The MLNM\((ABC)\) with restrictions \(DBE = 0\) will be discussed. This set up is equivalent (see von Rosen [4]) to the MLNM\((A_1 B_1 C_1 + A_2 B_2 C_2)\), where \(A_1 = A(D')^0\), \(A_2 = AD'\), \(C_1 = C\), and \(C_2 = E^0 C\). The likelihood ratio test was firstly derived by Khatri [1] and is in common use and available on standard statistical packages. The results presented below give additional information about the estimators and how to do the inference in the growth curve model under the restrictions \(DBE = 0\). From Theorem 3.1 and Theorem 3.2 it follows that

\[
D[A \hat{B} C] = D_1 + D_2 + 2C_{1,2},
\]
where

\[ D_1 = D[(I - T_1)(X - E[X]) C_1'(C_1 C_1')^{-1} C_1] \]
\[ = C_1'(C_1 C_1')^{-1} C_1 \otimes f_1 L_1 \]
\[ D_2 = D[(I - T_2)(X - E[X]) C_2'(C_2 C_2')^{-1} C_2] \]
\[ = C_2'(C_2 C_2')^{-1} C_2 \otimes (d_{1,2} K_1 + f_2 L_2) \]
\[ C_{1,2} = C[(I - T_1)(X - E[X]) C_1'(C_1 C_1')^{-1} C_1, \]
\[ (I - T_2)(X - E[X]) C_2'(C_2 C_2')^{-1} C_2] \]
\[ = C_2'(C_2 C_2')^{-1} C_2 \otimes (d_{1,2} K_1 + L_2). \]

Moreover, applying the results in Sections 4 and 5, we obtain
\[
E[n \hat{\Sigma}] = (n - \rho(C_1)) \Sigma + (\rho(C_1) - \rho(C_2))(g_{1,1} K_1 + \Sigma G_1 (G_1' \Sigma G_1)^{-1} G_1' \Sigma) + \rho(C_2)(g_{1,2} K_1 + g_{2,2} K_2 + \Sigma G_2 (G_2' \Sigma G_2)^{-1} G_2' \Sigma)
\]
and
\[
\hat{\Sigma} + 1/n (\rho(C_1) - \rho(C_2))K_{1,2} A_1 (A_1' \hat{\Sigma}^{-1} A_1)^{-1} A_1'
\]
\[ + 1/n \rho(C_2)(K_{1,3} A_1 (A_1' \hat{\Sigma}^{-1} A_1)^{-1} A_1' + K_{2,3} A_2 (A_2' \hat{\Sigma}^{-1} P_2 A_2)^{-1} A_2') \]
is an unbiased estimator for \( \Sigma \).

### 7. Asymptotic Considerations

We have seen in the previous sections that the distributions for the maximum likelihood estimators are complicated. Therefore, in order to supplement the use of the moment relations, we will derive asymptotic equivalent expressions for \( \hat{\Sigma} \). Furthermore, consistency for \( \hat{\Sigma} \) will be verified. The asymptotic situation we are thinking of is the one with increasing number of columns in \( X: p \times n \). This will be denoted \( n \to \infty \). Moreover, it will be assumed that for all \( n \) the parameter matrices as well as the \( A \) matrices will be the same. This means, among other things, that the number of rows in the \( C \) matrices will be constant. Before giving a fairly easy treatment we need some more definitions. Convergence in probability is denoted \( \to^P \). "\( X \to^P Y \), as \( n \to \infty \), for any sequence of random matrices "\( X \) and some matrix \( Y \), all with the same dimensions, signifies that "\( x_{ij} \to^P y_{ij} \), as \( n \to \infty \), for all \( i \) and \( j \). Equivalently, by using a distance, \( \| \cdot \| \), between two matrices defined by
\[
\| Q - P \| = \text{vec}(Q - P)' \text{vec}(Q - P) = \text{tr}((Q - P)(Q - P)').
\]
for any matrices \( Q \) and \( P \) of the same dimension, "\( X \rightarrow^n Y \) means \( \|X - Y\| \rightarrow^p 0 \). It is worth observing that \( \|A(Q - P)B\| > \|C(Q - P)D\| \) if \((BB' \otimes AA') - (DD' \otimes CC')\) is p.d. Further, for any two sequences of matrices "\( X \) and "\( Y \) the concept of asymptotic equivalence \( \rightarrow^a \) is defined by "\( X \rightarrow^n Y \) \( \rightarrow^p 0 \) or equivalently by \( \|X - Y\| \rightarrow^p 0 \). Note that the superscript \( n \) will usually be omitted in the rest of this section. Moreover, \( \rightarrow^a \) is only fruitful if "\( X \rightarrow^n Y \) \( \rightarrow^p 0 \) is not trivially true, i.e., "\( X \rightarrow^n Y \) \( \rightarrow^p 0 \) if and only if "\( X \rightarrow^n Y \) \( \rightarrow^p 0 \).

Since \( \hat{E}[X] \) is an estimator for the mean of all \( n \) columns in \( X \), it is meaningless to discuss asymptotic properties unless one's intention is to discuss a fixed subset of columns. Therefore we are going to show an asymptotic result for the first \( v \) columns in \( X \), denoted \( X_v \). Similarly, let \( C_{iv} \) denote the first \( v \) columns in \( C_i \). It may be worth observing that since we are working with finite dimensional matrices there exists a fixed number of columns \( \delta \), such that the information added by new columns is of the same type as the information from previous columns, i.e. \( R(C_{1\delta}) = R(C_{1n}) \) for \( n \geq \delta \). Let

\[
P_{r\Sigma} = T_{r-1\Sigma} T_{r-2\Sigma} T_{r-3\Sigma} \times \cdots \times T_{0\Sigma}, \quad T_{0\Sigma} = I, \quad r = 1, 2, \ldots, m + 1
\]

\[
T_{i\Sigma} = I - P_{i\Sigma} A_i (A'_i P'_{i\Sigma} \Sigma^{-1} P_{i\Sigma} A_i)^{-1} A'_i P'_{i\Sigma} \Sigma^{-1}, \quad i = 1, 2, \ldots, m.
\]

Our first observation is that, for Theorem 1.1, \( S_i/n \rightarrow^p \Sigma \) and \( X F_i F_i' X'/n \rightarrow^p 0 \), \( i = 2, 3, \ldots, m \). Hence, we obtain the chain \( S_1/n \rightarrow^p \Sigma \Rightarrow T_1 \rightarrow^p T_{1\Sigma} \Rightarrow P_2 \rightarrow^p P_{2\Sigma} \Rightarrow S_2/n \rightarrow^p \Sigma \Rightarrow T_2 \rightarrow^p T_{2\Sigma} \Rightarrow P_3 \rightarrow^p P_{3\Sigma} \Rightarrow \cdots \Rightarrow S_m/n \rightarrow^p \Sigma \Rightarrow T_m \rightarrow^p T_{m\Sigma} \Rightarrow P_{m+1} \rightarrow^p P_{m+1\Sigma} \). Let

\[
\hat{E}_{\Sigma} [X_v] = \sum_{i=1}^{m} (I - T_{i\Sigma}) X C_i (C_i C'_i)^{-1} C_{iv}, \quad v \leq \delta
\]

\[
S_{i\Sigma} = \sum_{j=1}^{i} P_{j\Sigma} X F_j F_j' X' F_j, \quad i = 1, 2, \ldots, m,
\]

and the following theorem will be proved.

**Theorem 7.1.**

(i) \( \hat{E} [X_v] \rightarrow^a F_{\Sigma} [X_v] \)

(ii) \( n \hat{S} \rightarrow^a S_{\Sigma} + P_{m+1\Sigma} X C'_m (C_m C'_m)^{-1} C_m X' P_{m+1\Sigma} \)

(iii) \( \hat{\Sigma} \rightarrow^p \Sigma \).

**Proof.** (ii) and (iii) follow immediately from the above implication chain. For (i) it is enough to consider \( (I - T_i) X C'_i (C_i C'_i)^{-1} C_i \). Let \( \epsilon \) be
any small quantity, \( M \) be any arbitrary non-stochastic matrix, and set \( Y = X - \sum_{j=1}^{r-1} A_j B_j C_j \). Then

\[
P(\|T_{12} - T_i\| X C_i' (C_i C_i')^{-1} C_i  > \varepsilon)
= P(\|T_{12} - T_i\| Y C_i' (C_i C_i')^{-1} C_i  > \varepsilon, \quad MM' - YC_i'(C_i C_i')^{-1} C_i Y' \text{ is p.d.})
+ P(\|T_{12} - T_i\| Y C_i'(C_i C_i')^{-1} C_i  > \varepsilon, \quad MM' - YC_i'(C_i C_i')^{-1} C_i Y' \text{ is not p.d.})
< P(\|T_{12} - T_i\| M  > \varepsilon)
+ P(MM' - YC_i'(C_i C_i')^{-1} C_i Y' \text{ is not p.d.}). \tag{7.1}
\]

Since \( T_i \to^p T_{12} \) implies that \( P(\|T_{12} - T_i\| M  > \varepsilon) \) converges to zero we will discuss the second expression. This has to be somewhat carefully done, since the number of columns in \( C_i \) is increasing as \( n \to \infty \). Note that for some vector \( \alpha: p \times 1 \) we obtain

\[
P(MM' - YC_i'(C_i C_i')^{-1} C_i Y' \text{ is not p.d.})
= P(\alpha'C_i'(C_i C_i')^{-1} C_i Y' \geq \alpha'M M' \alpha)
\leq \frac{\text{tr}(C_i'(C_i C_i')^{-1} C_i Y')}{\alpha'M M' \alpha},
\]

where the Markov inequality has been applied. Now, since \( R(C_{m') \subseteq R(C_{m-1}) \subseteq \cdots \subseteq R(C_i) \) implies that \( C_i = Q_i C_i', i < j \), where \( Q_i \) is independent of \( n \), \( C_j C_i'(C_i C_i')^{-1} C_i = Q_j C_{m'} \) which obviously also is independent of \( n \) and \( \text{tr}(C_{m'}'(C_i C_i')^{-1} C_i) \leq \text{tr}(C_{m'}'(C_i C_i')^{-1} C_i) = \rho(C_{m'}) \) and since \( M \) is arbitrary, \( M \) is chosen so that the probability in (7.1) is smaller than any prerequested quantity. \( \Box \)

Remarks. \( E_{\Sigma}[X_v] \) is normally distributed and then by aid of standard results it is possible to construct asymptotic confidence regions for \( E[X_v] \). Furthermore, in (ii) we have a sum of independent Wishart distributed terms.

**REFERENCES**


