PROFILE ANALYSIS FOR A GROWTH CURVE MODEL

Martin Ohlson* and Muni S. Srivastava**

In this paper, we consider profile analysis of several groups where the groups have partly equal means. This leads to a profile analysis for a growth curve model. The likelihood ratio statistics are given for the three hypotheses known in literature as parallelism, level hypothesis and flatness. Furthermore, exact and asymptotic distributions are given in the relevant cases.

Key words and phrases: Growth curve model, profile analysis.

1. Introduction

The profile analysis can be described for the two sample case as follows. Suppose that we wish to compare the performance of students from two different schools. Assume that we have $n_1$ students from school A and $n_2$ students from school B. We wish to compare the performance of the students in four different subjects such as English ($S_1$), History ($S_2$), Mathematics ($S_3$) and Science ($S_4$). Let $\mu_1 = (\mu_{11}, \mu_{12}, \mu_{13}, \mu_{14})'$ and $\mu_2 = (\mu_{21}, \mu_{22}, \mu_{23}, \mu_{24})'$ be the population means of the two groups for some tests in the four different subjects. The profiles for the two schools can be described as the lines connecting the points $(1, \mu_{i1}), \ldots, (4, \mu_{i4})$, for $i = 1, 2$.

Srivastava (1987) considered profile analysis for this type of problem but for several groups and derived the likelihood ratio statistics for the three hypotheses known in the literature as parallelism, level hypothesis and no condition variation.

The parallelism hypothesis, is the hypothesis that the two groups have similar profiles or that there is no interaction between the schools and the subjects, i.e.,

$$H_1 : \mu_1 - \mu_2 = \gamma 1 \quad \text{vs.} \quad A_1 : \mu_1 - \mu_2 \neq \gamma 1,$$

where $1$ is a vector of ones.

The level hypothesis is given by

$$H_2 \mid H_1 : \gamma = 0 \quad \text{vs.} \quad A_2 : \gamma \neq 0,$$

where $H_2 \mid H_1$ denotes the conditional hypothesis given that the hypothesis $H_1$ holds. Hence, the second hypothesis is, given $H_1$ is there any difference in mean between the two groups or are the level difference zero.

Received February 7, 2009. Revised November 18, 2009. Accepted April 9, 2010.

*The Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden. Email: martin.ohlson@mai.liu.se

**Department of Statistics, University of Toronto, 100 St. George Street Toronto, Ontario, M5S 3G3, Canada. Email: srivastava@utstat.toronto.edu
Finally the third hypothesis, the flatness (or no condition variation) hypothesis is given that the groups have similar profiles are the group means parallel to the $x$-axis, i.e., are they flat. The flatness hypothesis is equivalent to testing

$$H_3 | H_1 : \mu_2 = \delta 1 \quad \text{vs.} \quad \mu_2 \neq \delta 1.$$ 

More recently, profile analysis has been discussed under more general models. Okamoto et al. (2006) studied the asymptotic expansions of the distributions of some test statistics considering elliptical distributions. Others have extended this model and discussed the asymptotic expansions for the null distribution of some test statistics for profile analysis under non-normality, e.g., see Maruyama (2007) and Harrar and Xu (2008).

For a parallel profile model, i.e., assuming parallel profiles, it has also been proposed to consider different covariance structures. In Yokoyama and Fujikoshi (1993) and Yokoyama (1995) the random-effect covariance structure is considered and some tests for the random-effect and flatness are derived.

The above example however, assumes that all of the four subjects of school A are parallel to school B. But often all the schools pay great attention to the two main subjects such as English and History, and there may not be any difference in these subjects between the two schools. The difference may be between the other two subjects, Mathematics and Science. It would be reasonable to assume that

$$\mu_i = \begin{pmatrix} \theta_i \\ \nu \end{pmatrix} \quad i = 1, 2,$$

where $\theta_i$ and $\nu$ are unknown vectors given by $\theta_i = (\mu_{i1}, \mu_{i2})'$ and $\nu = (\mu_3, \mu_4)'$.

Thus, we need to consider profile analysis for the mean vectors $\theta_i$.

The organization of this paper is as follows. In Section 2, the three hypotheses are formulated and in Section 3, the general maximum likelihood estimators (MLEs) under the null hypothesis are presented. Furthermore, in Section 4, the likelihood ratio criteria are derived for the three hypotheses and their distributions and asymptotic distributions are obtained in the relevant cases. In Section 5, a numerical example is given and in Section 6, we give a conclusion of the paper. In Last, some results are given in an Appendix. For some examples of the ordinary profile analysis, i.e., without the restriction on partly equal means, see Srivastava and Carter (1983), Srivastava (1987, 2002).

## 2. The three hypotheses

Let $x_{ij}$ be $p$-dimensional random vectors independently distributed as $N_p(\mu_i, \Sigma)$, where $\Sigma > 0$, i.e. positive definite, $j = 1, \ldots, n_i$, $i = 1, \ldots, k$ and $N = n_1 + \cdots + n_k$. Furthermore, let the means be partly equal which is given by

$$\mu_i = \begin{pmatrix} \theta_i \\ \nu \end{pmatrix},$$

where $\theta_i$ and $\nu$ are unknown vectors given by $\theta_i = (\mu_{i1}, \mu_{i2})'$ and $\nu = (\mu_3, \mu_4)'$. Thus, we need to consider profile analysis for the mean vectors $\theta_i$.
where \( \theta_i \) are unknown \( r \)-vectors and \( \nu \) is a unknown \((p-r)\)-vector.

The first hypothesis, the parallelism hypothesis, can be formulated as

\[
H_1: \delta_i = \gamma_1, \quad i = 1, \ldots, k - 1 \quad \text{vs.} \quad A_1 \neq H_1
\]

where

\[
\delta_i = \theta_i - \theta_k,
\]

\[
1_r = (1, \ldots, 1)' : r \times 1
\]

and the scalars \( \gamma_i \) are unknown. Hence, we test if the first \( r \) part in the profiles are parallel while we know that the second \( p-r \) part are identical.

Furthermore, let \( \gamma = (\gamma_1, \ldots, \gamma_{k-1})' \) and

\[
\Delta = (\delta_1, \ldots, \delta_{k-1}) : r \times (k-1).
\]

The first hypothesis can also be written in a matrix form as

\[
H_1: \Delta = 1_r \gamma' \quad \text{vs.} \quad A_1 : \Delta \neq 1_r \gamma'.
\]

The level hypothesis is given by

\[
H_2 | H_1 : \gamma = 0 \quad \text{vs.} \quad A_2 | H_1 : \gamma \neq 0.
\]

The level hypothesis is testing the case when given parallel profiles, if the profiles are at the same level, i.e., the hypothesis is testing no column effect.

It may be noted that the hypothesis \( H_1 \) is free of the parameter \( \nu \), the common \((p-r)\)-vector for all the \( k \) groups. Thus, the flatness hypothesis, will therefore be formulated as

\[
H_3 | H_1 : \theta_* = \gamma_k 1_r \quad \text{and} \quad \nu = 0 \quad \text{vs.} \quad A_3 | H_1 \neq H_3 | H_1,
\]

where

\[
(2.1)
\]

\[
\theta_* = \frac{1}{N} \sum_{i=1}^{k} n_i \theta_i.
\]

One can observe that under the flatness hypothesis there is no row effect.

3. The maximum likelihood estimators

Before we derive the MLEs will we give some notation and transformations.

Let \( X \) be the observation matrix of all the observations from the \( k \) groups

\[
X = (X_1, \ldots, X_k),
\]

where

\[
X_i = (x_{i1}, \ldots, x_{in_i}).
\]
i.e., $X_i$ is the $n_i$ observations from group $i$. The expectation of the observation matrix is given by

$$E(X) = \mu A,$$

where the matrix $\mu = (\mu_1, \ldots, \mu_k)$ is all the means of the $k$ different groups and $A$ is the between-individuals design matrix

$$A = \begin{pmatrix} 1_{n_1} & 0 \\ \vdots & \ddots \\ 0 & 1'_{n_k} \end{pmatrix}.$$ 

Furthermore, we will use the following standard notation

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij},$$

$$\bar{x} = \frac{1}{N} X A' (AA')^{-1} A 1_N = \frac{1}{N} \sum_{i=1}^{k} n_i \bar{x}_i,$$

$$\mu_* = \frac{1}{N} \mu A 1_N = \frac{1}{N} \sum_{i=1}^{k} n_i \mu_i,$$

$$\bar{X} = X A' (AA')^{-1} = (\bar{x}_1, \ldots, \bar{x}_k)$$

and the within sum of squares

$$V = X (I - A' (AA')^{-1} A) X' = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)' ,$$

which is a sufficient statistic for $\Sigma$. Define also $\eta$ to be

$$\eta = (\mu_1 - \mu_k, \ldots, \mu_{k-1} - \mu_k) = \begin{pmatrix} \delta_1 & \ldots & \delta_{k-1} \\ 0 & \ldots & 0 \end{pmatrix} = B \Delta$$

where

$$B = \begin{pmatrix} I_r \\ 0 \end{pmatrix} : p \times r.$$ 

We also know that a sufficient statistic for $\eta$ is

$$Y = (\bar{x}_1 - \bar{x}_k, \ldots, \bar{x}_{k-1} - \bar{x}_k).$$

Since $(\bar{x}_1, \ldots, \bar{x}_k, V)$ are sufficient statistics for $(\mu_1, \ldots, \mu_k, \Sigma)$ and equivalently $(Y, \bar{x}, V)$ are sufficient for $(\eta, \mu_*, \Sigma)$, the likelihood function, given in Srivastava (1987), is

$$c |\Sigma|^{-N/2} e^{tr}\left\{ -\frac{1}{2} \Sigma^{-1}[V + (Y - \eta) \Xi^{-1}(Y - \eta)' + N(\bar{x} - \mu_*)(\bar{x} - \mu_*)'] \right\}$$

$$= c |\Sigma|^{-N/2} e^{tr}\left\{ -\frac{1}{2} \Sigma^{-1}[V + (Z - B\Delta)(Z - B\Delta)'] + N(\bar{x} - \mu_*)(\bar{x} - \mu_*)'] \right\},$$

(3.2)
where \( c \) is a constant,
\[
Z = Y \Xi^{-1/2} : p \times (k - 1),
\]
\[
\tilde{\Delta} = \Delta \Xi^{-1/2} : r \times (k - 1)
\]
and
\[
\Xi = \begin{pmatrix}
  n_1^{-1} + n_k^{-1} & n_k^{-1} & \cdots & n_k^{-1} \\
  n_k^{-1} & n_2^{-1} + n_k^{-1} & \cdots & n_k^{-1} \\
  \vdots & \ddots & \ddots & \ddots \\
  n_k^{-1} & n_k^{-1} & \cdots & n_{k-1}^{-1} + n_k^{-1}
\end{pmatrix}.
\]
The mean structure in the likelihood function (3.2) is a growth curve model. We see directly that the MLE of \( \mu \) is \( \bar{x} \). From Corollary 1.10.3 of Srivastava and Khatri (1979) it follows that the MLE of \( \tilde{\Delta} \) is given by
\[
B \tilde{\Delta} = B(B'V^{-1}B)^{-1}B'V^{-1}Z
\]
and the MLE of the covariance matrix \( \Sigma \) is
\[
N \hat{\Sigma} = V + (I - B(B'V^{-1}B)^{-1}B'V^{-1})ZZ'(I - B(B'V^{-1}B)^{-1}B'V^{-1})'.
\]
4. The likelihood ratio statistics
In this section we derive the likelihood ratio statistics for the three hypotheses. We will start to present a very useful canonical reduction of the problem. This reduction will partition the mean into three parts and separate the likelihood ratio statistic into independent parts.

Let \( G : r \times r \) be an orthogonal matrix such that
\[
G = \left( \frac{1}{\sqrt{r}}, G_1 \right),
\]
i.e., we have \( GG' = G'G = I_r, G_1'1 = 0 \) and \( G_1'G_1 = I_{r-1} \). We can write the mean of \( Z \) as
\[
E(Z) = B\tilde{\Delta} = BGG'\tilde{\Delta} = B^*\Delta^*,
\]
where
\[
B^* = BG : p \times r
\]
and
\[
\Delta^* = G'\tilde{\Delta} = \frac{1}{r-1} \left( \frac{1}{\sqrt{r}}1'\tilde{\Delta} \right),
\]
\[
\tilde{\Delta} = (B'V^{-1}B)^{-1}B'V^{-1}Z
\]
i.e., we have separated the mean into two parts. One part is of dimension \(1 \times (k - 1)\) and the second part is of dimension \((r - 1) \times (k - 1)\). To partition the second part of the mean into two smaller parts let

\[
\Gamma = \begin{pmatrix}
B^{\ast'} \\
B_0^{\ast'}
\end{pmatrix} : p \times p,
\]

where \(B_0^{\ast'}\) are chosen such that \(\Gamma\) is orthogonal. Since \(\Gamma\) is known, consider the transformation

\[
Z^{\ast} = \Gamma Z = \begin{pmatrix} 1 \\ z_1^\ast' \\ Z_2^\ast \\ Z_3^\ast \end{pmatrix} = \begin{pmatrix} 1 \\ z_1^\ast' \\ Z_2^{(2)} \end{pmatrix},
\]

where the three partitions of \(Z^{\ast}\) are \(z_1^\ast' : 1 \times (k - 1)\), \(Z_2^\ast : (r - 1) \times (k - 1)\), \(Z_3^\ast : (p - r) \times (k - 1)\) and

\[
Z_2^{(2)} = \begin{pmatrix} Z_2^\ast \\ Z_3^\ast \end{pmatrix}.
\]

The mean of \(Z^{\ast}\) is given by

\[
E(Z^{\ast}) = \Gamma B^{\ast} \Delta^{\ast} = \begin{pmatrix} B^{\ast'} \\ B_0^{\ast'} \end{pmatrix} B^{\ast} \Delta^{\ast} = \begin{pmatrix} \Delta^{\ast} \\ 0 \end{pmatrix} = B \Delta^{\ast} = \begin{pmatrix} m_1' \\ M_{(2)} \end{pmatrix},
\]

where

\[
m_1 = E(z_1^\ast')
\]

and

\[
M_{(2)} = E(Z_2^{(2)}) = \begin{pmatrix} r-1 \\ p-r \end{pmatrix} \begin{pmatrix} M_2 \\ 0 \end{pmatrix}.
\]

Furthermore, the transformation of the within sum of squares is

\[
V^{\ast} = \Gamma V \Gamma',
\]

and equivalently the transformation of the covariance matrix is

\[
\Sigma^{\ast} = \Gamma \Sigma \Gamma'.
\]
4.1. The first hypothesis $H_1$

The first hypothesis parallelism is given by

$$H_1: \Delta = 1, \gamma' \quad \text{vs.} \quad A_1 \neq H_1.$$ 

Under hypothesis $H_1$ the parameter matrix $\Delta^*$ can be written as

$$\Delta^* = \begin{pmatrix} \sqrt{\gamma' \Xi}^{-1/2} \\ 0 \end{pmatrix} = \begin{pmatrix} m_1' \\ 0 \end{pmatrix},$$

where

$$m_1 = \sqrt{\gamma' \Xi}^{-1/2} \gamma.$$ 

The expectation under $H_1$ is given by

$$E(Z^*) = 1_{p-1} \begin{pmatrix} m_1' \\ 0 \end{pmatrix} = bm_1',$n

where

$$b = \begin{pmatrix} 1 \\ 0_{p-1} \end{pmatrix} : p \times 1,$$

i.e., we have

$$M_2 = 0.$$ 

The maximum likelihood estimators for the transformed variables under the alternative hypothesis $A_1$ is given by

$$\hat{\Delta}^* = (B'V^{-1}B)^{-1}B'V^{-1}Z^*$$

and for the original variables

$$\hat{\Delta} = (B'V^{-1}B)^{-1}B'V^{-1}Z \Xi^{1/2}.$$ 

For the sake of completeness we will give the covariance matrices for the estimators. The covariance matrix for the estimator $\hat{\Delta}$ is given by

$$\text{Cov}(\hat{\Delta}) = E[(\Xi^{1/2} \otimes ((B'V^{-1}B)^{-1}B'V^{-1})) \text{Cov}(Z)]$$

$$= \Xi \otimes E[(B'V^{-1}B)^{-1}B'V^{-1} \Sigma V^{-1}B(B'V^{-1}B)^{-1}]$$

$$= \frac{N - k - 1}{N - k - p + r - 1} \Xi \otimes (B' \Sigma^{-1} B)^{-1},$$

where we have used $\text{Cov}(Z) = I \otimes \Sigma$ and Theorem 4.2.1 of Kollo and von Rosen (2005) with suitable modification for the last equality.
Under the null hypothesis $H_1$ we have for the transformed variables
\[
\hat{m}'_1 = (b'V'^{-1}b)^{-1}b'V'^{-1}Z^*
\]
and again for the original variables
\[
\hat{\gamma}' = (1' B' V^{-1} B 1)^{-1} 1' B' V^{-1} Z \Xi^{1/2}
\]
with the covariance matrix
\[
\text{Cov}(\hat{\gamma}) = \left(1 + \frac{p - 1}{N - k - p}\right) (1' B' \Sigma^{-1} B 1)^{-1} \Xi,
\]
where we have used Lemma 3.2 of Srivastava (1987). Furthermore, the likelihood ratio test is given by
\[
\lambda_{H_1} = \left| I + Z' (V^{-1} - V'^{-1}B (B' V'^{-1} B)^{-1} B' V'^{-1}) Z \right| = \left| I + Z' (V^{-1} - V'^{-1} B (V'^{-1} B)^{-1} B' V^{-1}) Z \right|
\]
in terms of the original variables, since
\[
\Gamma b = \frac{1}{\sqrt{p}} B 1.
\]
From the construction of the likelihood ratio statistic, reject the hypothesis $H_1$ for small values of $\lambda_{H_1}$.

4.1.1. The null distribution of $\lambda_{H_1}$

The way to find the null distribution of $\lambda_{H_1}$ is similar to the technique used in Chapter 6.4.4 of Srivastava and Khatri (1979). In this section the calculations are done in detail since it is an important step in the derivation of all the distributions in this paper.

Let $V \sim W_p(\Sigma, m)$, where $m = N - k$ and partition $V$ as
\[
V = \begin{pmatrix} v_{11} & v_{12}' & v_{13}' \\ v_{12} & V_{22} & V_{23} \\ v_{13} & V_{23}' & V_{33} \end{pmatrix} = \begin{pmatrix} v_{11} & v_{(12)}' \\ v_{(12)} & V_{(22)} \end{pmatrix},
\]
where $v_{11} : 1 \times 1$,
\[
v_{(12)} = (v_{12}', v_{13}') : 1 \times (p - 1),
\]
\[
V_{(22)} = \begin{pmatrix} V_{22} & V_{23} \\ V_{23}' & V_{33} \end{pmatrix} : (p - 1) \times (p - 1).
\]

Let also
\[
Z \sim N_{p,k-1}(M, \Sigma, I_{k-1}),
\]
independent of $V$ and partition $Z$ similar to $V$

\[
Z = \begin{pmatrix}
    z'_1 \\
    Z_2 \\
    Z_3
\end{pmatrix} = \begin{pmatrix}
    z'_1 \\
    \mathbf{Z}^{(2)}
\end{pmatrix} \quad \text{and} \quad
M = \begin{pmatrix}
    m' \\
    M_2 \\
    0
\end{pmatrix}.
\]

**Theorem 4.1.** The likelihood ratio statistic $\lambda_{H_1}$ given in equation (4.4), can be written as

\[
\lambda_{H_1} = \frac{|V_{2,3}|}{|V_{2,3} + W_1|}
\]

where $V_{2,3}$ is the Schur complement

\[
V_{2,3} = V_{22} - V_{23}V_{33}^{-1}V_{23}'
\]

and

\[
W_1 = (Z_2 - V_{23}V_{33}^{-1}Z_3)(I + Z_3'V_{33}^{-1}Z_3)^{-1}(Z_2 - V_{23}V_{33}^{-1}Z_3)'.
\]

**Proof.** Using Corollary 1.4.2(ii) of Srivastava and Khatri (1979) or Appendix A.2. in Siotani et al. (1985), we get

\[
V^{-1} = \begin{pmatrix}
    c_{22}^{-1} & -v_{11}'v_{12}'C_{11}^{-1} \\
    -c_{11}'v_{12}' & C_{11}^{-1}
\end{pmatrix},
\]

where

\[
c_{22} = v_{11}' - v_{12}'V_{22}^{-1}v_{12}
\]

and

\[
C_{11} = V_{22} - v_{12}'v_{11}^{-1}v_{12}'.
\]

Furthermore, using the definition of $b$ in equation (4.3) we have

\[
(b'V^{-1}b)^{-1} = c_{22},
\]

\[
b(b'V^{-1}b)^{-1}b'V^{-1} = \begin{pmatrix}
    1 - c_{22}v_{11}'v_{12}'C_{11}^{-1} \\
    0
\end{pmatrix}
\]

and

\[
V^{-1} - V^{-1}b(b'V^{-1}b)^{-1}b'V^{-1} = \begin{pmatrix}
    0 & 0' \\
    0 & V_{22}^{-1}
\end{pmatrix}.
\]
Then the second part of the denominator in (4.4) is given by

\[ Z'(V^{-1} - V^{-1}b'(bV^{-1}b)^{-1}b'V^{-1})Z = (z_1, Z'_{(2)}) \begin{pmatrix} 0 & 0' \\ 0 & V_{(22)}^{-1} \end{pmatrix} \begin{pmatrix} z_1' \\ Z'_{(2)} \end{pmatrix} = Z'_{(2)} V_{(22)}^{-1} Z_{(2)}. \]

Using similar calculations for the alternative hypothesis, the nominator in (4.4) is

\[ Z'(V^{-1} - V^{-1}B(BV^{-1}B)^{-1}B'V^{-1})Z = Z'_3 V^{-1}_{33} Z_3 \]

and the likelihood ratio can be written as

\[ \lambda_{H_1} = \frac{|I + Z'_3 V^{-1}_{33} Z_3|}{|I + Z'_{(2)} V_{(22)}^{-1} Z_{(2)}|} \]

Using Problem 1.3 (iii) of Srivastava and Khatri (1979) we have

\[ |V_{(22)}| = |V_{33}| |V_{22} - V_{23} V_{33}^{-1} V'_{23}|, \]
\[ |V_{(22)} + Z_{(2)} Z'_{(2)}| = |V_{33} + Z_3 Z'_3| \]
\[ |V_{22} + Z_2 Z'_2 - (V_{23} + Z_2 Z'_3)(V_{33} + Z_3 Z'_3)^{-1}(V_{23} + Z_2 Z'_3)'| \]

and this gives the likelihood ratio statistic as

\[ \lambda_{H_1} = \frac{|V_{22} - V_{23} V_{33}^{-1} V'_{23}|}{|V_{22} + Z_2 Z'_2 - (V_{23} + Z_2 Z'_3)(V_{33} + Z_3 Z'_3)^{-1}(V_{23} + Z_2 Z'_3)'|}. \]

Using Lemma A.1 in Appendix A, the denominator in the likelihood ratio (4.6) can be written as

\[ V_{22} + Z_2 Z'_2 - (V_{23} + Z_2 Z'_3)(V_{33} + Z_3 Z'_3)^{-1}(V_{23} + Z_2 Z'_3)' = V_{2,3} + (Z_2 - V_{23} V_{33}^{-1} Z_3)(I + Z'_3 V_{33}^{-1} Z_3)^{-1}(Z_2 - V_{23} V_{33}^{-1} Z_3)', \]

where

\[ V_{2,3} = V_{22} - V_{23} V_{33}^{-1} V'_{23}. \]

Hence, the likelihood ratio for the first hypothesis \( H_1 \) is given by

\[ \lambda_{H_1} = \frac{|V_{2,3}|}{|V_{2,3} + W_1|}, \]

where

\[ W_1 = (Z_2 - V_{23} V_{33}^{-1} Z_3)(I + Z'_3 V_{33}^{-1} Z_3)^{-1}(Z_2 - V_{23} V_{33}^{-1} Z_3)'. \]

and the proof is complete.
The likelihood ratio statistic is separated into independent parts. We will now derive the null distribution of \( \lambda_{H_1} \). Partition \( \Sigma \) similar to \( V \)

\[
\Sigma = \frac{1}{p-1} \begin{pmatrix}
\sigma_{11} & \sigma_{12}' \\
\sigma_{12} & \Sigma_{(22)} 
\end{pmatrix},
\]

where \( \sigma_{11} : 1 \times 1, \sigma_{12} : (p-1) \times 1 \) and \( \Sigma_{(22)} : (p-1) \times (p-1) \).

**Theorem 4.2.** The distribution of \( \lambda_{H_1} \) is the same as the distribution of the product of \( r-1 \) independently distributed beta random variables with parameters

\[
\frac{1}{2}(N-k-p+r+1-i) \quad \text{and} \quad \frac{1}{2}(k-1), \quad \text{where} \quad i = 1, \ldots, r-1.
\]

**Proof.** Since \( V \sim W_p(\Sigma, m) \) then from Theorem 3.3.2 of Srivastava and Khatri (1979), \( V_{(22)} \sim W_{p-1}(\Sigma_{(22)}, m) \). Furthermore, from Theorem 3.5.5. of Srivastava and Khatri (1979), \( V_{2,3} \) and \( (V_{23}', V_{33}) \) are independently distributed as,

\[
V_{2,3} \sim W_{r-1}(\Sigma_{2,3}, m-p+r), \\
V_{23}'V_{33}^{-1} | V_{33} \sim N_{r-1,p-r}(\Sigma_{23}^{-1} \Sigma_{33}, V_{23}, V_{33}^{-1})
\]

and

\[
V_{33} \sim W_{p-r}(\Sigma_{33}, m),
\]

where

\[
\Sigma_{(22)} = \begin{pmatrix}
\Sigma_{22} & \Sigma_{23} \\
\Sigma_{23}' & \Sigma_{33}
\end{pmatrix}
\]

and

\[
\Sigma_{2,3} = \Sigma_{22} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{32}'.
\]

We also know that

\[
Z_{(2)} = \begin{pmatrix} Z_2 \\ Z_3 \end{pmatrix} \sim N_{p-1,k-1}(M_{(2)}, \Sigma, I_{k-1})
\]

and \( Z_{(2)} \) and \( V \) are independently distributed. Furthermore, given \( Z_3 \) the distribution for \( Z_2 \) is

\[
Z_2 | Z_3 \sim N_{r-1,k-1}(M_2 + \Sigma_{23} \Sigma_{33}^{-1} Z_3, \Sigma_{2,3}, I_{k-1}).
\]

From these equations

\[
Z_2 - V_{23} V_{33}^{-1} Z_3 | Z_3, V_{33} \sim N_{r-1,k-1}(M_2, \Sigma_{2,3}, I_{k-1} + Z_3' V_{33}^{-1} Z_3)
\]
and it is independent of \( \mathbf{V}_{2,3} \). Furthermore, under \( H_1 \), \( \mathbf{M}_2 = \mathbf{0} \),

\[
\mathbf{Y}_1 = (\mathbf{Z}_2 - \mathbf{V}_{23} \mathbf{V}_{33}^{-1} \mathbf{Z}_3)(\mathbf{I} + \mathbf{Z}_3' \mathbf{V}_{33}^{-1} \mathbf{Z}_3)^{-1/2}
\]

and \( \mathbf{V}_{2,3} \) are independently distributed as,

\[
Y_1 \sim \mathcal{N}_{r-1,k-1}(0, \Sigma_{2,3}, I_{k-1}),
\]

\[
V_{2,3} \sim \mathcal{W}_{r-1}(\Sigma_{2,3}, m - p + r)
\]

and the likelihood ratio is given by

\[
\lambda_{H_1} = \left| \frac{V_{2,3}}{V_{2,3} + Y_1'Y_1} \right|.
\]

Hence, using Theorem 6.3.2 of Srivastava and Khatri (1979) we see that the distribution of \( \lambda_{H_1} \) is the same as the distribution of the product of \( r - 1 \) independent beta random variables with parameters \( \frac{1}{2}(N - k - p + r + 1 - i) \) and \( \frac{1}{2}(k - 1) \), where \( i = 1, \ldots, r - 1 \). The proof is complete.

4.1.2. The asymptotic distribution of \( \lambda_{H_1} \)

The asymptotic approximation for a similar likelihood ratio statistic has been considered by several authors, see for example Rao (1948), Box (1949) or Srivastava and Khatri (1979, Chapter 6.3.7). Including the second and the third term, the expansion is given by

\[
P(-r_0 \ln \lambda_{H_1} \leq z) = P(\chi^2_{f} \leq z) + \frac{\alpha_2}{r_0^2}[P(\chi^2_{f+4} \leq z) - P(\chi^2_{f} \leq z)]
\]

\[
+ \frac{1}{r_0^4}(\alpha_4[P(\chi^2_{f+8} \leq z) - P(\chi^2_{f} \leq z)] - \alpha_2^2[P(\chi^2_{f+4} \leq z) - P(\chi^2_{f} \leq z)]) + \ldots
\]

where

\[
r_0 = N - p + \frac{1}{2}(k - r + 1),
\]

\[
f = (r - 1)(k - 1),
\]

\[
\alpha_2 = \frac{f}{48}((r - 1)^2 + (k - 1)^2 - 5),
\]

\[
\alpha_4 = \frac{1}{2}\alpha_2^2 + \frac{f}{1920}(3(r - 1)^4 + 3(k - 1)^4 + 10(r - 1)^2(k - 1)^2 - 50((r - 1)^2 + (k - 1)^2) + 159).
\]

For higher terms in the asymptotic expansion, see Rao (1948) and Box (1949). Hence, for large \( N \) the distribution can be given by

\[
-(N - p - \frac{1}{2}(k - r + 1)) \ln \lambda_{H_1} \sim \chi^2_{(r-1)(k-1)}.
\]
4.2. The second hypothesis

The second hypothesis is given by

\[ H_2 \mid H_1 : \gamma = 0 \quad \text{vs.} \quad A_2 \mid H_1 : \gamma \neq 0 \]

and we have the estimators

\[ N\hat{\Sigma}_{H_1} = V^* \]

\[ + \left( I - b(b'V^{*-1}b)^{-1}b'V^{*-1} \right) Z' Z'(I - b(b'V^{*-1}b)^{-1}b'V^{*-1})' \]

and

\[ N\hat{\Sigma}_{H_2 \mid H_1} = V^* + Z' Z'^* \]

Hence, the likelihood ratio statistic is

\[ \lambda_{H_2 \mid H_1} = \frac{|\hat{\Sigma}_{H_2}|}{|\hat{\Sigma}_{H_2 \mid H_1}|} \]

\[ = \frac{|V^*| \left| I + Z'(V^{*-1} - V^{*-1}b(b'V^{*-1}b)^{-1}b'V^{*-1})Z' \right|}{|V^* + Z' Z'^*|} \]

\[ = \frac{|I + Z'(V^{*-1} - V^{*-1}B1(1'B'V^{*-1}B1)^{-1}1'B'V^{*-1})Z|}{|I + Z' V^{*-1}Z|} \]

in terms of the original variables. Reject the hypothesis \( H_2 \mid H_1 \) for small values of \( \lambda_{H_2 \mid H_1} \). Furthermore, let \( V^* \) be partitioned as

\[ V^* = \begin{pmatrix} v_{11} & v'_{(12)} \\ v_{(12)} & V_{(22)} \end{pmatrix} \]

The likelihood ratio can then also be written as

\[ \lambda_{H_2 \mid H_1} \]

\[ = \frac{|V^*| \left| I + Z'(V^{*-1} - V^{*-1}b(b'V^{*-1}b)^{-1}b'V^{*-1})Z' \right|}{|V^* + Z' Z'^*|} \]

\[ = \frac{|v_{11} - z_1' z_1 - (v'_{(12)} - z_1' Z'_{(2)}) (V_{(22)} - Z'_{(2)} Z'^{-1} (v'_{(12)} - z_1' Z'_{(2)}))|}{v_{1,23} + w_2} \]

since

\[ |V^*| = v_{1,23} |V_{(22)}| \]

where we again have the Schur complement

\[ v_{1,23} = v_{11} - v'_{(12)} V_{(22)}^{-1} v_{(12)} \]

and used Lemma A.1 for the last equality, i.e.,

\[ w_2 = (z_1' - v'_{(12)} V_{(22)}^{-1} Z'_{(2)}) (I - Z'_{(2)} V_{(22)}^{-1} Z'_{(2)})^{-1} (z_1' - v'_{(12)} V_{(22)}^{-1} Z'_{(2)})' \]

The distribution of \( \lambda_{H_2 \mid H_1} \) is derived in a similar way as the distribution for \( \lambda_{H_1} \).
Theorem 4.3. Rejecting the hypothesis $H_2 | H_1$ for small values of $\lambda_{H_2|H_1}$ is equal to rejecting the hypothesis for large values of

$$F = \frac{y_2^\top y_2}{v_{1.23}}$$

where

$$y_2'' = (z_1'' - v_{(12)'(22)} V^{-1}_{(22)} Z_{(2)}') (I - Z_{(2)}' V^{-1}_{(22)} Z_{(2)})^{-1/2}.$$

The null distribution of $F$ is given by

$$\frac{N - k - p + 1}{k - 1} F \sim F_{k - 1, N - k - p + 1}.$$

Proof. Since $V^* \sim W_p(\Sigma^*, m)$, where $m = N - k$ then from Theorem 3.5.5. of Srivastava and Khatri (1979), $v_{1.23}$ and $(v_{(12)'(22)}, V_{(22)})$ are independently distributed as,

$$v_{1.23} \sim W_1(\sigma_{1.23}, m - p + 1),$$

$$v_{(12)'(22)} V^{-1}_{(22)} | V_{(22)} \sim N_{1, p - 1}(\sigma'_{(12)} \Sigma^{-1}_{(22)}, \sigma_{1.23}, V^{-1}_{(22)}).$$

and

$$V_{(22)} \sim W_{p - 1}(\Sigma_{(22)}, m),$$

where

$$\sigma_{1.23} = \sigma_{11} - \sigma'_{(12)} \Sigma^{-1}_{(22)} \sigma_{(12)}.$$

We also know that

$$Z^* = \begin{pmatrix} z_1'' \\ Z_{(2)}' \end{pmatrix} \sim N_{p, k - 1}(M, \Sigma^*, I_{k - 1})$$

and $Z^*$ and $V^*$ are independently distributed. Furthermore, given $Z_{(2)}'$, the distribution for $z_1''$ is

$$z_1'' | Z_{(2)}' \sim N_{1, k - 1}(m_1' + \sigma'_{(12)} \Sigma^{-1}_{(22)} (Z_{(2)}' - M_{(2)}), \sigma_{1.23}, I_{k - 1}).$$

From these equations

$$z_1'' - v_{(12)'(22)} V^{-1}_{(22)} Z_{(2)}' | Z_{(2)}', V_{(22)} \sim N_{1, k - 1}(m_1' - \sigma'_{(12)} \Sigma^{-1}_{(22)} M_{(2)}, \sigma_{1.23}, I_{k - 1} + Z_{(2)}' V^{-1}_{(22)} Z_{(2)})$$

and it is independent of $v_{1.23}$. Furthermore, under hypothesis $H_2 | H_1$, $m_1' = 0'$ and $M_{(2)} = 0$,

$$y_2'' = (z_1'' - v_{(12)'(22)} V^{-1}_{(22)} Z_{(2)}') (I - Z_{(2)}' V^{-1}_{(22)} Z_{(2)})^{-1/2}.$$
and $v_{1.23}$ are independently distributed as,

$$y_2^* \sim N_{k-1}(0, \sigma^{-1}_{1.23}I_{k-1}),$$

$$\sigma^{-1}_{1.23}v_{1.23} \sim \chi^2(N-k-p+1).$$

Finally, the likelihood ratio is given by

$$\lambda_{H_2|H_1} = \frac{v_{1.23}}{v_{1.23} + y_2^*y_2^*} = \frac{1}{1 + F},$$

where

$$F = \frac{y_2^*y_2^*}{v_{1.23}}.$$

Hence, rejecting the hypothesis $H_2 | H_1$ for small values of $\lambda_{H_2|H_1}$ is equal to rejecting for large values of $F$.

Under $H_2 | H_1$ the likelihood ratio statistic $\lambda_{H_2|H_1}$ is beta distributed with the two parameters $(\frac{1}{2}(N-k-p+1), \frac{1}{2}(k-1))$ since (4.7) and

$$\sigma^{-1/2}_{1.23}y_2^* \sim N_{k-1}(0, I_{k-1}).$$

More interesting is the distribution of $F$. From (4.7) and (4.8) the null distribution of $F$ is given by

$$\frac{N-k-p+1}{k-1}F \sim F_{k-1, N-k-p+1}$$

and the proof is complete.

4.3. The third hypothesis

Since the hypothesis $H_1$ is free of the parameter $\nu$, the common $(p-r)$-vector for all the $k$ groups, the flatness hypothesis is equivalent to testing

$$H_3 | H_1 : \theta = \gamma_kl_r \quad \text{and} \quad \nu = 0 \quad \text{vs.} \quad A_3 | H_1 \neq H_3 | H_1.$$ The hypothesis $H_3 | H_1$ can also be written as

$$\mu_\bullet = \gamma_kB1_r,$$

where $B$ is given in equation (3.1). If we use the same transformations as before, (4.1) and (4.2), we have

$$\mu_\bullet = \gamma_kB1_r = m_kb,$$

where $m_k = \sqrt{\gamma_k}$ and $b$ is given in (4.3). From the likelihood function (3.2) we see that under $H_3|H_1$ the estimator of $m_k$ is given by

$$\hat{m}_k = (b'V^{-1}b)^{-1}b'V^{-1}x^*.$$
where
\[ \bar{x}^* = \Gamma \bar{x} \sim N_p \left( m \bar{b}, \frac{1}{N} \Sigma^* \right) \]
and \( \Sigma^* = \Gamma \Sigma \Gamma' \). The estimator for the covariance matrix under \( H_3 \mid H_1 \) is given by
\[
N \hat{\Sigma}_{H_3 \mid H_1} = V^* + (I - b(b'V^{*-1}b)^{-1}b'V^{*-1})(Z^*Z^{*-1} + N\bar{x}^*\bar{x}^{**})
\]
\[ \times (I - V^{*-1}b(b'V^{*-1}b)^{-1}b') \]
\[ = V^* + V^*C(C'V^*C)^{-1}(C'Z^*Z^{*-1}C + NC'\bar{x}^*\bar{x}^{*}) \]
\[ \times (C'V^*C)^{-1}CV^* \]  
(4.9)
and the estimator under \( H_1 \) is given by
\[
N \hat{\Sigma}_{H_1} = V^* + (I - b(b'V^{*-1}b)^{-1}b'V^{*-1})
\]
\[ \times Z^*Z^{*-1}(I - V^{*-1}b(b'V^{*-1}b)^{-1}b') \]
\[ = V^* + V^*C(C'V^*C)^{-1}C'Z^*Z^{*-1}C(C'V^*C)^{-1}CV^* \]  
(4.10)
where we have used Lemma 1 of Khatri (1966), i.e.,
\[
V^{*-1} - V^{*-1}b(b'V^{*-1}b)^{-1}b'V^{*-1} = C'(C'V^*C)^{-1}C' \]
where \( C' : (p - 1) \times p \) is a matrix such that \( C'b = 0 \). We choose \( C' = (0, I_{p-1}) \).

**Theorem 4.4.** Rejecting \( H_3 \mid H_1 \) for small values of \( \lambda_{H_3 \mid H_1} \) is equivalent to rejecting the hypothesis for large values of
\[ T^2 = N\bar{x}^*C(C'V^*C + C'Z^*Z^{*-1}C)^{-1}C'\bar{x}^*. \]
Furthermore, the null distribution of \( T^2 \) is given by
\[ \frac{N - p + 1}{p - 1} T^2 \sim F_{p-1, N-p+1}. \]

**Proof.** From equation (4.9) and (4.10) the likelihood ratio is given by
\[
\lambda_{H_3 \mid H_1} = \frac{\left| \hat{\Sigma}_{H_1} \right|}{\left| \hat{\Sigma}_{H_3 \mid H_1} \right|} = \frac{|I + (C'V^*C)^{-1}C'Z^*Z^{*-1}C|}{|I + (C'V^*C)^{-1}(C'Z^*Z^{*-1}C + NC'\bar{x}^*\bar{x}^{*})|}
\]
\[ = \frac{|C'V^*C + C'Z^*Z^{*-1}C|}{|C'V^*C + C'Z^*Z^{*-1}C + NC'\bar{x}^*\bar{x}^{*}|} \]
\[ = \left( \frac{1}{1 + N\bar{x}^*C(C'V^*C + C'Z^*Z^{*-1}C)^{-1}C'\bar{x}^*} \right) = \frac{1}{1 + T^2}, \]
where
\[ T^2 = N\bar{x}^*C(C'V^*C + C'Z^*Z^{*-1}C)^{-1}C'\bar{x}^*. \]
We reject $H_3 | H_1$ for small values of $\lambda_{H_3 | H_1}$, i.e., for large values of $T^2$. But, since

$$C'x^* \sim N_{p-1} \left( 0, \frac{1}{N} \Sigma_{(22)} \right),$$

$$C'V^*C = V_{(22)} \sim W_{p-1}(\Sigma_{(22)}, N - k),$$

$$Z_{(2)}^2 \sim N_{p-1,k-1}(0, \Sigma_{(22)}, I_{k-1})$$

and $V_{(22)}$ and $Z^*_{(2)}$ are independent, we have

$$Z^*_{(2)}Z^*_{(2)} \sim W_{p-1}(\Sigma_{(22)}, k - 1),$$

$$C'V^*C + C'Z^*Z^*C \sim W_{p-1}(\Sigma_{(22)}, N - 1).$$

Finally the distribution of $T^2$ is given by

$$\frac{N - p + 1}{p - 1} T^2 \sim F_{p-1, N-p+1}$$

and the proof is complete. See Chapter 4.3.3 of Srivastava and Khatri (1979) for more details.

5. Example with fictitious data

We wish to compare the performance of students from four different schools. Assume that we have $n_i$ students from school $i = 1, 2, 3, 4$. We wish to compare the performance of the students in four different subjects such as Mathematics ($S_1$), Science ($S_2$), English ($S_3$) and History ($S_4$). We assume that all the schools pay great attention to the two main subjects such as English and History, and there may not be any difference in these subjects between the four schools. The difference may be between the other two subjects Mathematics and Science. Students were required to solve problems in each subject. All the problem were planned to be of the same difficulty and the time to solve each problem was recorded. From the fictitious data we obtain the following means and estimators

$$\bar{x}_1 = ( 40.8401 \ 50.0870 \ 66.0990 \ 54.4465 )', \quad n_1 = 10$$
$$\bar{x}_2 = ( 19.2984 \ 29.7353 \ 64.8225 \ 54.7835 )', \quad n_2 = 15$$
$$\bar{x}_3 = ( 29.5027 \ 39.2996 \ 64.6787 \ 54.4949 )', \quad n_3 = 20$$
$$\bar{x}_4 = ( 15.1663 \ 24.9324 \ 65.2049 \ 54.6189 )', \quad n_4 = 25$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{k} n_i \bar{x}_i = ( 23.8155 \ 33.6600 \ 65.1003 \ 54.5941 )'$$

and

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{k} \bar{x}'_i$$

and

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{k} n_i \bar{x}_i$$

and

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{k} n_i \bar{x}_i$$
\[ V = \begin{pmatrix} 383.6149 & 344.8670 & 257.2851 & 200.2245 \\ 344.8670 & 415.1372 & 280.9100 & 207.3783 \\ 257.2851 & 280.9100 & 398.2755 & 181.5067 \\ 200.2245 & 207.3783 & 181.5067 & 349.9283 \end{pmatrix}, \]

\[ Y = \begin{pmatrix} 25.6738 & 4.1322 & 14.3364 \\ 25.1547 & 4.8029 & 14.3673 \\ 0.8941 & -0.3824 & -0.5262 \\ -0.1724 & 0.1646 & -0.1240 \end{pmatrix}, \]

and the estimator of \( \gamma \) as

\[ \hat{\gamma} = (24.9911 \; 4.6172 \; 14.6716)' . \]

This gives

\[ \lambda_{H_1} = \frac{|I + Z'(V^{-1} - V^{-1}B(V^{-1}B1)'-1B'V^{-1})Z|}{I + Z'(V^{-1} - V^{-1}B1(V^{-1}B1)'-1B'V^{-1})Z|} = 0.9159 \]

and

\[ -(N - p - \frac{1}{2}(k - r + 1)) \ln \lambda_{H_1} = 5.6655 < \chi^2_{(r-1)(k-1), 0.05} = 7.8147. \]
We reject $H_1$ if $-(N - p - \frac{1}{2}(k - r + 1)) \ln \lambda_{H_1}$ is big, i.e., the hypothesis $H_1$ is not rejected, see Figure 1.

Furthermore, for the second hypothesis we have

$$\lambda_{H_2 | H_1} = \frac{|I + Z'(V^{-1} - V^{-1}B_1(1'B'V^{-1}B_1)^{-1}1'B'V^{-1})Z|}{|I + Z'V^{-1}Z|} = 0.0302$$

and

$$\frac{N - k - p + 1}{k - 1} (1 - \lambda_{H_2 | H_1}) \lambda_{H_2 | H_1}^{-1} = 673.4902 > F_{k-1,N-k-p+1,0.05} = 2.7505.$$ 

Hence the second hypothesis $H_2$ given $H_1$ is rejected at level 0.05.

For testing the third hypothesis $H_3 | H_1$ we calculate

$$T^2 = 924.7916 > \frac{p - 1}{N - p + 1} F_{p-1,N-p+1,0.05} = 0.1228,$$

i.e., the hypothesis $H_3 | H_1$ is also rejected at the level 0.05.

6. Conclusion

In this paper, a generalization of the profile analysis for several groups considered by Srivastava (1987) is presented. We assume that the different groups have partly equal means, i.e., a part of the profiles are identical. This lead to a profile analysis for a growth curve model. Three hypotheses for profile analysis known as parallelism, level hypothesis and flatness are considered, and the likelihood ratio statistic in each case are calculated and the null distributions are discussed. It is shown that these distributions are similar to the ones given in Srivastava (1987) modified for the identical means.

Acknowledgements

The work of Martin Ohlson was supported by The Royal Swedish Academy of Science, project number FOA08V-060.

Appendix A

In this appendix we give and prove a Lemma with some cumbersome calculations.

**Lemma A.1.** Let the following matrices be $A_{11} : p \times p$, $A_{12} : p \times q$, $A_{22} : q \times q$ and $X_1 : p \times n$, $X_2 : q \times n$. Then

$$A_{11} + X_1X_1' - (A_{12} + X_1X_1')(A_{22} + X_2X_2')^{-1}(A_{12} + X_1X_1')'$$

(A.1)$$= A_{11} - (A_{12}A_{22}^{-1}X_2)(I + X_2A_{22}^{-1}X_2)^{-1}(X_1 - A_{12}A_{22}^{-1}X_2)' ,$$

where

$$A_{12} = A_{11} - A_{12}A_{22}^{-1}A_{12}' .$$
Proof. Using Theorem 1.4.1(iv) of Srivastava and Khattri (1979) we have

\[
(A_{12} + X_1X'_1)(A_{22} + X_2X'_2)^{-1}(A_{12} + X_1X'_2)' = (A_{12} + X_1X'_2)(A_{22}^{-1} - A_{22}^{-1}X_2(I + X_2^{'2}A_{22}^{-1}X_2)^{-1}X_2^{'2}A_{22}^{-1})
\]

\[
\times (A_{12} + X_1X'_2)'
\]

and the left hand side in equation (A.1) can be written as

\[
\begin{align*}
A_{11} + X_1X'_1 - (A_{12} + X_1X'_2)(A_{22} + X_2X'_2)^{-1}(A_{12} + X_1X'_2)' &= A_{11} + X_1X'_1 - (A_{12} + X_1X'_2)A_{22}^{-1}(A_{12} + X_1X'_2)' \\
&+ (A_{12} + X_1X'_2)A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}(A_{12} + X_1X'_2)'
\end{align*}
\]

\[
(A.2) = A_{12} + F_1 = F_2 - F_3 + A_{12} A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}A_{12}^{-1},
\]

where

\[
F_1 = X_1X'_1 - X_1X'_2 A_{22}^{-1}X_2 X'_1
\]

\[
+ X_1X'_2 A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}X_2 X'_1
\]

\[
= X_1(I - X_2^{'}A_{22}^{-1}X_2 + X_2^{'}A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}X_2)X'_1
\]

\[
= X_1(I - X_2^{'}A_{22}^{-1/2}(I - A_{22}^{-1/2}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1/2}))
\]

\[
\times A_{22}^{-1/2}X_2 X'_1
\]

\[
(A.3) = X_1(I + X_1^{'}A_{22}^{-1}X_2)^{-1}X'_1,
\]

\[
F_2 = A_{12}A_{22}^{-1}X_2 X'_1 - A_{12} A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}X_2 X'_1
\]

\[
= A_{12} A_{22}^{-1}X_2(I - (I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}X_2)X'_1
\]

\[
= A_{12} A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}(I + X_2^{'}A_{22}^{-1}X_2 - X_2^{'}A_{22}^{-1}X_2)X'_1
\]

\[
(A.4) = A_{12} A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X'_1
\]

and similarly for \( F_3 \).

\[
F_3 = X_1X'_2 A_{22}^{-1}A_{12}^{-1} - X_1X'_2 A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}A_{12}^{-1}
\]

\[
(A.5) = X_1(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}A_{12}^{-1}
\]

Insert equation (A.3)-(A.5) in (A.2) and we have

\[
A_{11} + X_1X'_1 - (A_{12} + X_1X'_2)(A_{22} + X_2X'_2)^{-1}(A_{12} + X_1X'_2)'
\]

\[
= A_{12} + X_1(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X'_1 - A_{12} A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X'_1
\]

\[
- X_1(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}A_{12}^{-1}
\]

\[
+ A_{12} A_{22}^{-1}X_2(I + X_2^{'}A_{22}^{-1}X_2)^{-1}X_2^{'}A_{22}^{-1}A_{12}^{-1}
\]

\[
= A_{12} + (X_1 - A_{12} A_{22}^{-1}X_2)(I + X_2^{'}A_{22}^{-1}X_2)^{-1}(X_1 - A_{12} A_{22}^{-1}X_2)',
\]

which concludes the proof.
REFERENCES


