SOME ASPECTS OF ESTIMATION AND HYPOTHESIS TESTING FOR GENERALIZED MULTIVARIATE LINEAR MODELS

by

Yang Chyuan Yuan

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1476

December 1984
SOME ASPECTS OF ESTIMATION AND HYPOTHESIS TESTING
FOR GENERALIZED MULTIVARIATE LINEAR MODELS

by

Yang Chyuan Yuan

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics.

Chapel Hill
1984

Approved by:

Advisor

Reader

Reader
Covariate adjustment procedures have been used to analyze the standard growth curve model. The procedure requires the selection of a non-singular matrix $H$ satisfying a set of conditions to reduce the model into a standard MANOCOVA. We show the resulting confidence procedure is invariant under different selections of the matrix $H$ satisfying the same set of conditions. We then extend the covariate adjustment procedure to the generalized growth curve model where in each disjoint set a subset of the total responses were measured. Again, the confidence procedure is invariant under different selections of the matrix $H$ when the set of conditions are satisfied in each disjoint set.

In a hierarchical multiresponse model, the $p$ responses are ordered. The model can be analyzed through a step-down procedure. With good estimates of the variances of the $p$ responses, the resulting confidence set can be constructed from $p$ quadratic forms of chi-square statistics. We optimize the confidence procedure by minimizing suitable norms of the confidence set.

The incomplete multiresponse (IM) model defined by Srivastava is based on a less-than-full-rank model in each of the disjoint datasets. For a design satisfying a set of specified conditions, this model can be analyzed by transforming the data into the framework of a standard multivariate general linear model. We propose a corresponding full rank IM model, show some simplification of the procedure, and compare these two models. We then study the relative performance of the proposed IM model and the corresponding complete multiresponse model.
Finally, we consider the intraclass correlation (IC) model. With the same p responses measured at each of t distinct time points, a natural extension of the IC model is the block version of the IC model. We examine the hypotheses of the block IC model as well as some special cases through likelihood ratio and Wald statistics. We then derive the null and non-null distributions for the corresponding test statistics.
ACKNOWLEDGEMENTS

I wish to express my sincere thanks to my advisor, Professor P. K. Sen, for suggesting this research topic. His constant guidance and support throughout the duration of this study are greatly appreciated. I also wish to thank Drs. R. W. Helms, J. L. Lamb, K. L. Lee, and M. J. Symons for serving on my doctoral committee. Special thanks to Dr. Symons with whom I have had the pleasure to serve as a research assistant during my graduate studies in the Department of Biostatistics.

The financial assistance from the Department of Biostatistics, Lipid Research Clinics Program, Cancer Research Center, Northrop Services, Inc., and Occupational Health Studies Group during my stay in Chapel Hill are gratefully acknowledged.

I would like to thank Dr. K. Boyle and Ms. V. Kasica for their friendship and many helpful comments for this study. Finally, for their excellent typing of the manuscript, Ms. R. Bahr and Ms. J. Capparella are warmly thanked.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I LITERATURE REVIEW AND OUTLINE OF THE RESEARCH</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Invariant Confidence Procedure</td>
<td>2</td>
</tr>
<tr>
<td>1.3 The Standard MGLM Model</td>
<td>4</td>
</tr>
<tr>
<td>1.3.1 Confidence Procedures in the Standard MGLM Model</td>
<td>7</td>
</tr>
<tr>
<td>1.4 The Growth Curve Models</td>
<td>9</td>
</tr>
<tr>
<td>1.4.1 The SGCM Model</td>
<td>9</td>
</tr>
<tr>
<td>1.4.2 The GGCM Model</td>
<td>12</td>
</tr>
<tr>
<td>1.5 The HM Model</td>
<td>13</td>
</tr>
<tr>
<td>1.6 The IM Model</td>
<td>16</td>
</tr>
<tr>
<td>1.7 Outline of the Research</td>
<td>19</td>
</tr>
<tr>
<td>II ON CONSTRUCTING THE CONFIDENCE SETS IN THE GENERALIZED MULTIVAIRATE LINEAR MODELS</td>
<td>22</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>22</td>
</tr>
<tr>
<td>2.2 The Standard Growth Curve Model</td>
<td>23</td>
</tr>
<tr>
<td>2.2.1 Covariate Adjustment in the SGCM Model</td>
<td>23</td>
</tr>
<tr>
<td>2.2.2 Invariant Confidence Procedure</td>
<td>25</td>
</tr>
<tr>
<td>2.3 The Generalized Growth Curve Model</td>
<td>28</td>
</tr>
<tr>
<td>2.3.1 Covariate Adjustment in the GGCM Model</td>
<td>28</td>
</tr>
<tr>
<td>2.3.2 Invariant Confidence Procedure</td>
<td>35</td>
</tr>
<tr>
<td>2.4 The Hierarchical Multiresponse Model</td>
<td>37</td>
</tr>
<tr>
<td>2.4.1 The HM Model with p=2</td>
<td>37</td>
</tr>
<tr>
<td>2.4.2 Optimum Confidence Procedure</td>
<td>45</td>
</tr>
<tr>
<td>2.5 Examples</td>
<td>54</td>
</tr>
<tr>
<td>2.5.1 An SGCM Model</td>
<td>56</td>
</tr>
<tr>
<td>2.5.2 A GGCM Model</td>
<td>58</td>
</tr>
<tr>
<td>2.5.3 An HM Model</td>
<td>62</td>
</tr>
</tbody>
</table>
CHAPTER III INCOMPLETE MULTIRESPONSE MODEL .... 70

3.1 Introduction ... 70

3.2 A full Rank Incomplete Multiresponse Model ... 71

3.2.1 Removing the Singularity of $LL'$ ... 75

3.3 Optimum Confidence Procedure ... 77

3.3.1 Invariance in the Optimum Confidence Procedure ... 82

3.4 Some Comparisons between Incomplete Multiresponse Models and Complete Multiresponse Models ... 84

3.4.1 Generalized Variance ... 85

3.4.2 Asymptotic Relative Efficiency ... 87

3.4.3 Cost ... 94

3.4.4 Minimum Risk ... 97

3.5 Numerical Examples ... 103

CHAPTER IV ON COVARIANCE STRUCTURES ... 109

4.1 Introduction ... 109

4.1.1 Asymptotic Null Distributions of Certain Likelihood Ratio Statistics ... 112

4.1.2 Asymptotic Distributions of the Wald Statistics ... 113

4.2 Testing Block Intraclass Covariance Structures for a Covariance Matrix: $H_1: \Sigma_0 = I \otimes \Sigma_1 + (J-I) \otimes \Sigma_2$ ... 115

4.2.1 The Asymptotic Null Distribution of the Likelihood Ratio Statistic ... 118

4.2.2 Asymptotic Non-null Distributions of the Likelihood Ratio Statistic when $t=2$ ... 119

4.3 Testing Proportionality for the Two Matrices, $\Sigma_1$ and $\Sigma_2$, in $H_1: \Sigma_1 = \omega \Sigma_2$ ... 120

4.3.1 The Wald Statistic When $p=2$ ... 124

4.3.2 Asymptotic Distributions of the Wald Statistic when $p=2$ ... 128
IV (continued)

4.4 Testing Intraclass Covariance Structures for the Two Matrices, $\Sigma_1$ and $\Sigma_2$, in $H_1$:

$H_3: \Sigma_1 = \sigma_1[I + \rho_1(J-I)]$, $\Sigma_2 = \sigma_2[I + \rho_2(J-I)]$ .... 129

4.4.1 The Asymptotic Null Distribution of the Likelihood Ratio Statistic .... 132

4.4.2 Asymptotic Non-null Distributions of the Likelihood Ratio Statistic when $p=2$ .... 133

4.5 Testing Equality of Elements of $\Sigma_2$ in $H_3$:

$H_4: \Sigma_1 = \sigma_1[I + \rho_1(J-I)]$, $\Sigma_2 = \sigma_1\omega J$ .... 138

4.5.1 The Asymptotic Null Distribution of the Modified Likelihood Ratio Statistic .... 140

4.5.2 Asymptotic Non-null Distributions of the Modified Likelihood Ratio Statistic .... 141

4.6 Numerical Examples .... 143

V SUGGESTIONS FOR FUTURE RESEARCH .... 149

BIBLIOGRAPHY .... 152
CHAPTER I

LITERATURE REVIEW AND OUTLINE OF THE RESEARCH

1.1 Introduction

In multireponse experiments, the responses are commonly assumed to have a multivariate normal distribution. With the same responses being measured on each experimental unit, the most commonly used model is the standard multivariate general linear model (MGLM).

Once the data of a repeated-measurements experiment have been collected at known times, a natural sequel to hypothesis testing is the fitting of a single polynomial function to the sample means by some form of least squares. In this way, the model for the responses might be tested more concisely by fewer coefficients as compared to the entire vector. A standard growth curve model (SGCM) can then be used to analyze this type of data.

Often the data in the SGCM model are missing either by chance or by design. In such a case, the data can be analyzed through a generalized growth curve model (GGCM).

Many situations occur where considerations based on prior information about the nature of responses and their relative importance in the investigation at hand enable the experimenter to decide which one, among a given pair of responses, needs to be measured on a larger number of experimental units. In such a case, he could arrange the responses in a descending order of importance into a hierarchical multiresponse (HM) model.
In some classes of multiresponse design, not all the responses are measured on each experimental unit. Such situations may arise when it is either physically impossible or uneconomical to study all responses on each unit. In this case, we turn to the incomplete multiresponse (IM) model.

The covariance matrix is of interest mostly as part of the model requirements that it be a known positive definite matrix or that it have a particular pattern with unknown elements. In some instances, it is natural to assume that the covariance matrix of the responses possesses certain symmetry properties that are suggested by the sampling situation.

The main purposes of this study are

1) To investigate the confidence procedures used in the generalized multivariate linear models. These models include SGCM, GGCM, HM, and IM models.

2) To test and estimate various covariance structures in blocks.

In this chapter, Sections 1.2 to 1.6 review the related literature and Section 1.7 outlines the research.

1.2 Invariant confidence procedure

Many statistical problems exhibit symmetries, which provide natural restrictions on the statistical procedures that are to be used. The invariance principle basically states that if two problems have identical formal structures (i.e. have the same sample space, parameter space, densities), then the same decision rule should be used in each problem. For a given problem, this principle is employed by considering transformations of the problem which result in transformed problems of identical structure.

The restriction that the decision rules in the original and trans-
formed problems be the same leads to the invariant decision rules. An optimal invariant decision rule may be approached from a viewpoint with proper specifications.

Let $X$ be a random variable, taking values in the sample space $X$, and having distribution $P_\theta$, $\theta \in \Theta$. Also let $g$ be a 1-1 transformation of the sample space $X$ onto $X$. Denote by $gX$ the random variable that takes on the value $gx$ when $X = x$, and the distribution of $gX$ is $P_{\theta'}$, with $\theta' = g\theta$, where $g$ is the corresponding induced transformation of $g$ on $\Theta$.

Suppose one desires inference about a function $\gamma$ of $\theta$, $\gamma(\theta)$, taking values in a space $\Gamma$. Then, with a constant probability $1-\alpha$ of covering the true value of $\gamma(\theta)$, a confidence procedure for $\gamma$, denoted $S(\cdot)$, is a function from $X$ into the collection of subsets of $\Gamma$. When $x$ is observed, $S(x)$ will be the suggested confidence set for $\gamma$ such that $P_\theta(\gamma \in S(x)) \geq 1-\alpha$ for all $\theta \in \Theta$.

A confidence procedure given by a class of confidence set $S(x)$ is invariant under the group of transformations $G$ if $gS(x) = S(gx)$ for all $x \in X$, $g \in G$, where $gS(x) = \{gy: y \in S\}$.

With a natural measure of the size of $S(x)$, an optimal invariant confidence procedure can be approached by choosing the invariant confidence procedure with the smallest measure of $S(x)$.

Let $\{\psi_i: i \in I\}$ be a family of parametric functions $\psi_i: \Gamma \rightarrow \Psi$, where $I$ is an index set and $\Psi$ is an arbitrary space. Then $\{\psi_i: i \in I\}$ can also be regarded as a single function $\psi: \Gamma \times I \rightarrow \Psi$. Similarly, a simultaneous confidence set (SCS) for $\{\psi_i\}$ can be regarded as a function $T$ on $X^*I$ whose values are subsets of $\Psi$.

Consider a confidence set $S(x)$ of $\gamma(\theta)$, when the two events $\{\gamma(\theta) \in S(x)\}$ and $\{\psi_i(\gamma(\theta)) \in T(x,i): i \in I\}$ are identical; then a family of exact SCS with
respect to $S$ is \{T(x,i); i \in I\}. If there exists a family of exact SCS with respect to $S_1$ \{T^*(x,i); i \in I\}, such that

$$T^*(x,i) \subset T(x,i) \text{ for all } i \in I;$$

then \{T^*(x,i); i \in I\} is a family of smallest exact SCS with respect to $S$.

Assume that an action of the induced group $\overline{G}$ on $\Gamma$ can be defined such that $\gamma(\overline{g} \theta) = \overline{g} \gamma(\theta)$ for all $\overline{g} \in \overline{G}$, $\theta \in \Theta$. Also assume that an action of $\overline{G}$ on $I$, and for each $i \in I$, an action of $\overline{G}$ on $\Psi$ (denoted $\psi \overline{g}_i \psi$) can be defined such that

\begin{equation}
\psi(\overline{g} \gamma, \overline{g}_i) = \overline{g}_i \psi(\gamma, i) \text{ for all } \gamma \in \Gamma, i \in I, \overline{g} \in \overline{G}.
\end{equation}

Then a simultaneous set estimator $T$ for $\psi$ is equivalent if

\begin{equation}
T(\overline{g} x, \overline{g}_i) = \overline{g}_i T(x, i) \text{ for all } x \in X, i \in I, \overline{g} \in \overline{G}.
\end{equation}

1.3 The standard MGLM model

The standard MGLM model is given by

\begin{equation}
E(Y) = X \beta,
\end{equation}

and

\begin{equation}
V(Y) = I_n \otimes \Sigma,
\end{equation}

where $Y_{n \times p}$ is a data matrix,

$X_{n \times r}$ is a design matrix,

$\beta_{r \times p}$ is a matrix of unknown parameters, and

the rows of $Y$ are independently normally distributed with a positive definite covariance matrix $\Sigma_{p \times p}$.

The likelihood function of the above model is
The maximum likelihood estimators (MLE's) of $\beta$ and $\Sigma$ are

\[ \hat{\beta} = (X'X)^{-1} X'Y \quad \text{and} \quad \hat{\Sigma} = \frac{1}{n} Y'[I - X(X'X)^{-1}X']Y. \]

It is of interest to estimate a linear set of $\beta$,

\[ \theta_{(g \times v)} = CBU, \quad \text{where} \quad C_{(g \times r)} \quad \text{is of rank} \quad g \quad \text{and} \quad U_{(p \times v)} \quad \text{is of rank} \quad v. \]

The MLE of $\theta$ is

\[ \hat{\theta} = CBU = C(X'X)^{-1} X'YU. \]

Three testing procedures are frequently used to test the null hypothesis $H_0: \theta = 0$. These procedures include Hotelling's generalized trace, Roy's largest root, and Wilks' likelihood ratio.

Letting

\[ S_H = \hat{\theta}'[C(X'X)^{-1}C']^{-1}\hat{\theta} \quad \text{and} \quad S_E = U'Y[I - X(X'X)^{-1}X']YU = n U'\hat{\Sigma}U, \]

then under $H_0$, $S_H$ has a Wishart distribution $W_v(q, U'\Sigma U)$, $S_E$ has a Wishart distribution $W_v(n-r, U'\Sigma U)$, and $S_H$ and $S_E$ are independent.

Together with $\lambda_j = \text{ch}_j(S_H S_E^{-1})$, the $j^{th}$ largest eigenvalue of $S_H S_E^{-1}$, $j=1, \ldots, v$, these three test procedures can be described as follows:

1. Hotelling's generalized trace statistics

\[ T_0^2 = \text{trace}(S_H S_E^{-1}) = \sum_{j=1}^{v} \lambda_j. \]
The percentage points of the distribution of $T_0^2$ has been tabulated by Davis (1970). Small sample $F$ approximations to $T_0^2$ and tables for their use have been found by Hughes and Saw (1972). For large value of $n$, $n T_0^2$ is asymptotically distributed as a $\chi^2$ variate with $gv$ d.f. We reject $H_0$ when $T_0^2$ is too large.

(2) Roy's largest root statistic

\begin{equation}
\lambda_1 = \text{ch}_1(S_H^{-1} S_E^{-1})
\end{equation}

The percentage points of the distribution of $\lambda_1$ can be found either from tables by Pillai and Bantegui (1959) or charts by Heck (1960) through the transformation

$$\theta_s = \frac{\lambda_1}{1+\lambda_1} \quad \text{and parameters} \quad s_0 = \min(g,v),$$

$$m_0 = (|g-v|-1)/2,$$

$$n_0 = (n-r-v-1)/2.$$  

We reject $H_0$ when $\lambda_1$ is too large.

(3) Wilks likelihood ratio statistic

\begin{equation}
\Lambda = \frac{|S_E|}{|S_H^{-1} S_E^{-1}|} = \prod_{j=1}^{v} (1+\lambda_j)^{-1}.
\end{equation}

For a wide range of values of $v$ and $s_0 = \min(g,v)$, transformations to exact chi-squared distributions and tables of the necessary multiplying factors have been given by Schatzoff (1966), Pillai and Gupta (1969), and Lee (1972). For large value of $n$, $-[n-r-\frac{1}{2}(v-g+1)] \ln\Lambda$ is distributed as a chi-squared variate with $gv$ degrees of freedom.

We reject $H_0$ when $\Lambda$ is too small.

When $s_0 = \min(g,v) = 1$, all three statistics are equivalent.
1.3.1 Confidence procedures in the standard MGLM model

In the standard MGLM model, a \((1-\alpha)\) confidence set for \(\theta\) can be constructed by replacing \(S_H\) with

\[(1.3.10) \quad S_H^* = (\hat{\theta} - \theta)' \left[ C(X'X)^{-1} C' \right]^{-1} (\hat{\theta} - \theta) \]

in the corresponding test statistic.

Of all three confidence procedures considered, only Roy's largest root gives the exact SCS for \(\theta\).

Letting \(\lambda_{1,1-\alpha}\) be the \((1-\alpha)\) point of the largest root distribution under \(H_0\), then a \((1-\alpha)\) confidence set for \(\theta\) is given by

\[(1.3.11) \quad \text{Ch}_1(S_H^*E) \leq \lambda_{1,1-\alpha}.\]

With

\[\text{Ch}_1(S_H^*E) = \max_a \frac{a' S_H^* a}{a' S_E a} = \max a \frac{a' (\hat{\theta} - \theta)' \left[ C(X'X)^{-1} C' \right]^{-1} (\hat{\theta} - \theta)a}{a' S_E a},\]

an equivalent confidence set is

\[(1.3.12) \quad [(\hat{\theta} - \theta)a]' \left[ C(X'X)^{-1} C' \right]^{-1} [(\hat{\theta} - \theta)a] \leq \lambda_{1,1-\alpha} a' S_E a \]

for all \((v \times 1)\) vector \(a\).

That is, \( |b' (\hat{\theta} - \theta)a| \leq \lambda_{1,1-\alpha} (a' S_E a) (b' \left[ C(X'X)^{-1} C' \right] b)^{1/2} \)

for all vectors \(a(v \times 1)\) and \(b(g \times 1)\).

Thus an exact \((1-\alpha)\) confidence set for \(\theta\) is

\[(1.3.13) \quad b' \hat{\theta} a - d \leq b' \theta a \leq b' \hat{\theta} a + d \]

for all vectors \(a(v \times 1)\) and \(b(g \times 1)\),

where \(d = \lambda_{1,1-\alpha} (a' S_E a) (b' \left[ C(X'X)^{-1} C' \right] b)^{1/2} \).

Schatzoff (1966) used simulation techniques to compare power among some tests of the general linear hypothesis. The results indicate that
Wilks' λ and Hotelling's $T^2_0$ provide about the same levels of protection over a wide spectrum of alternatives, whereas Roy's $\lambda_1$ is shown to be rather insensitive to alternatives involving more than one non-zero root. But for certain restricted alternatives (e.g. with single non-zero root), Roy's $\lambda_1$ is somewhat better. 

That is, the relative powers of the tests varied according to the type of alternative (the nature of the population). No single test procedure is best against all types of alternatives (uniformly most powerful).

From the null hypothesis $H_0: \Theta = 0$, we can construct an equivalent family of hypotheses $\Omega = \{H_{0i}, i \in I\}$, where $I$ is an index set and $H_{0i}: C_{i} \Theta U_{i} = 0$, $C_{i}(g_{i}xg)$ is of rank $g_{i}(sg)$, and $U_{i}(v\times v_{i})$ is of rank $v_{i}(sv)$.

Based on increasing root functions of $S_{K}^{-1}$, which include Hotelling's $T^2_0$, Roy's $\lambda_1$, and Wilks' λ. Gabriel (1968) showed some useful properties of these simultaneous test procedures (STP). These properties include

1. Coherence: If $H_{0i}, H_{0j} \in \Omega$, and $H_{0i}$ implies $H_{0j}$, then the test procedure rejects $H_{0i}$ whenever it rejects $H_{0j}$. Conversely, if $H_{0i}$ is accepted, so is $H_{0j}$.

2. The probability of making one or more type I errors in all the tests of an $\alpha$-level STP is exactly $\alpha$ if $H_0$ holds. Otherwise it is no greater than $\alpha$.

3. Among all STP given identical decisions on all one-root hypotheses (i.e. having equal critical values since on such hypotheses all the statistics are equal), Roy's $\lambda_1$ procedure is the most parsimonious in that it has the lowest level. In other words, it provides the same single root tests at the smallest probability of any type I error.

4. Among all STP of level $\alpha$, Roy's $\lambda_1$ procedure is the most resolvent
in the sense that in testing every one-root hypothesis it will reject every such hypothesis rejected by any other STP, and, possibly, some more.

(5) Among all simultaneous confidence bounds of joint confidence 1-\(\alpha\), Roy's \(\lambda_1\) procedure gives the narrowest confidence intervals for each parametric function. Roy's intervals are contained in the corresponding intervals of any other simultaneous confidence bounds.

For a family \(\{\psi_i(y)\}\) of parametric functions of the parameter of interest \(y = y(\theta)\), Wijsman (1979) presented a method for the construction of all families of the smallest SCS in a given class. The method is applied to the MANOVA problem (in its canonical form) of inference about \(M = E(Y_1)\), where \(Y_1(g \times p)\) has rows that are independent multivariate normal with common covariance matrix \(\Sigma\). Let \(S\) be the usual estimate of \(\Sigma\). Wijsman then showed that the smallest equivalent SCS for all \(a'M\), where \(a\) is a \((g \times 1)\) vector, are necessarily those that are exact with respect to the confidence set for \(M\) determined by \(Ch_1[M-Y_1]S^{-1}(M-Y_1)'\) ≤ constant, i.e. derived from the acceptance region of Roy's largest root statistic.

1.4 The growth curve models

1.4.1 The SGCM model

The SGCM model is given by

\[
(1.4.1) \quad E(Y_0) = XB_0,
\]

and

\[
V(Y_0) = I_n \otimes \Sigma_0,
\]

where

- \(Y_0(n \times g)\) is a data matrix,
- \(X(n \times r)\) is a design matrix,
- \(B(r \times p)\) is a matrix of unknown parameters,
\( p(\times q) \) is a known matrix of rank \( p \), and

the rows of \( Y_0 \) are independently normally distributed with

common variance matrix \( \Sigma_0(q\times q) \).

The hypothesis of interest is \( H_0: \theta = 0 \), where \( \theta(g\times v) = CBU \), \( C(g\times r) \) is of

rank \( g \), and \( U(p\times v) \) is of rank \( v \).

To analyze the above model, Potthoff and Roy (1964) suggested the transformation

\[
Y(n\times p) = Y_0 G^{-1}P'(PG^{-1}P')^{-1},
\]

where \( G(q\times q) \) is any positive definite symmetric matrix either nonstochastic

or independent of \( Y_0 \) such that \( PG^{-1}P' \) is nonsingular.

Using the above transformation, we can reduce the SGCM model to a standard MGLM model on the data matrix \( Y \).

\[
E(Y) = X\beta,
\]

and

\[
V(Y) = I_n \otimes \Sigma_G,
\]

where

\[
\Sigma_G(p\times p) = (PG^{-1}P')^{-1}PG^{-1} \Sigma_0(G^{-1}P'(PG^{-1}P')^{-1}.
\]

To avoid the arbitrary choice of the matrix \( G \) in the above procedure, Rao (1966, 1967) and Khatri (1966) used covariate adjustment to reduce the

SGCM model to a multivariate analysis of covariance (MANOCOVA) model. With

the construction of the nonsingular matrix \( H(q\times q) = (H_1(q\times p), H_2(q\times [q-p])) \)

such that

\[
PH_1 = I \quad \text{and} \quad PH_2 = 0,
\]

we have the following MANOCOVA model:

\[
E(Y|Z) = X\beta + Z\Gamma,
\]
and

\[ V(Y|Z) = I_n \otimes \left[ \Sigma_0^{-1} \right]^{-1}, \]

where

\[ Y_{(n \times p)} = Y_0 \cdot 1 \]

is a data matrix,

\[ Z_{(n \times [q-p])} = Y_0 \cdot H_2 \]

is a matrix of covariates,

\[ X_{(n \times r)} \]

is a design matrix,

\[ \beta_{(r \times p)} \]

is a matrix of unknown parameters, and

\[ \Gamma_{([q-p] \times p)} \]

is a matrix of unknown regression coefficients.

The procedure used in analyzing the SGCM model depends on

1) the method used to transform the data into the framework of a standard MGLM model, and

2) within each method, the statistic used to analyze the reduced standard MGLM model.

As described in Section 1.3.1, for each method used, no test statistic is uniformly most powerful.

Estimation of \( \beta \) in Rao-Khatri's method is more efficient than Potthoff and Roy's approach only if \( Y_0 \cdot H_1 \) and \( Y_0 \cdot H_2 \) are correlated. Selected subsets of the covariates may, in some cases, yield a better estimate for \( \beta \) than the complete set, as indicated by Grizzle and Allen (1969).

Gleser and Olkin (1970) introduced a canonical form for the SGCM model from which the maximum likelihood estimator \( \hat{\beta} \) can easily be obtained. In the process, they also considered transformations which leave the testing procedure invariant. Kariya (1978) then applied further invariance reduction to the canonical form for the SGCM model and derived a unique locally best invariant test.

Under certain conditions, Hooper (1982) derived a method to construct a confidence set having smallest expected measure within the class of invar-
iant level (1−α) confidence sets. He then applied the method to the SGCM model in the form of a MANOCOVA model.

1.4.2 The GGCM model

To incorporate possible missing responses in the response curve design, Kleinbaum (1973) defined the following GGCM model:

\[(1.4.6) \quad E(Y_j) = X_j\beta B_j, \]

and

\[V(Y_j) = I_n_j \odot \Sigma_j, \quad j=1, \ldots, u, \]

where

\[Y_j(n_j \times q_j) \]

is a data matrix,

\[X_j(n_j \times r) \]

is a design matrix,

\[\beta(r \times p) \]

is a matrix of unknown parameters,

\[P(p \times q) \]

is a known matrix of rank p,

\[B_j(q \times q_j) \]

is an incidence matrix of 0's and 1's, and

\[\Sigma_j(q_j \times q_j) = B_j^\prime \Sigma B_j \]

is a positive definite matrix.

The hypothesis of interest \(H_0: \theta = 0\), where \(\theta(g \times v) = CBu, C(g \times r)\) is of rank g, and \(U(p \times v)\) is of rank v.

Let \(\beta^*(r_p \times 1)\) be the vector constructed from \(\beta\) by putting the columns of \(\beta\) underneath each other. Then by also rolling out the columns of the observed data, the GGCM model can be analyzed through an equivalent univariate linear model.

Letting \(\theta^*(g_v \times 1) = HB^*\), where \(H(g_v \times rp) = U' \odot C, \theta^*\) and \(\theta\) have the same \(g_v\) elements. A best asymptotically normal estimate of \(\theta^*\) can be constructed from \(\hat{\theta}^* = HB^*\), where \(B^*\) is the weighted LSE of \(\beta^*\) in the univariate model.
Then the null hypothesis $H_0$ can be tested through a Wald statistic involving $\hat{\theta}^*$ and its estimated variance.

1.5 The HPM model

In an HPM model, the responses $V_1, \ldots, V_p$ are arranged in a descending order of importance. If $V_r$ is more important than $V_j$, then $V_r$ should also be observed on each experimental unit on which $V_j$ is observed.

Let $n_j$ be the number of experimental units on which the response $V_j$ is observed. Then the response data has the structure

$$ Y = \begin{pmatrix}
  Y_1^1 \\
  Y_2^1 \\
  \vdots \\
  Y_p^1 \\
  Y_1^2 \\
  Y_2^2 \\
  \vdots \\
  Y_p^2 \\
  \vdots \\
  Y_1^p \\
  Y_2^p \\
  \vdots \\
  Y_p^p
\end{pmatrix}, $$

where $Y_r^{(n_j' \times 1)}$ is a data matrix, $j = 1, \ldots, p$, $r = 1, \ldots, j$,

$$ n_j' = n_j - n_{j+1}, \quad j = 1, \ldots, p-1, $$

and $n_p' = n_p$.

Let $Y_r^{(n_j' \times j)} = (y_1^r, \ldots, y_j^r)$, $j = 1, \ldots, p$, and

$$ Y_r^{(n_r \times 1)} = \begin{pmatrix}
  y_1^r \\
  \vdots \\
  y_p^r
\end{pmatrix}, \quad r = 1, \ldots, p. $$

Then the HPM model can be described as follows:

(1.5.1) $E(Y_r) = X_r \beta_r$, \quad $r = 1, \ldots, p$,

and $V(Y_j) = n_j' \otimes \Sigma_j$, \quad $j = 1, \ldots, p$,

where $X_r^{(n_r \times m_r)}$ is a design matrix,
\( \beta_{r(m \times 1)} \) is a matrix of unknown parameters,
\( \Sigma_{j(j \times j)} \) is the \((j \times j)\) submatrix in the top left corner of \( \Sigma \),
\( \Sigma_{(p \times p)} = (\sigma_{r,s}) \) is the covariance matrix of the \( p \) responses,
and the rows of \( Y^j \) are independently normally distributed with common covariance matrix \( \Sigma_j \).

The hypothesis of interest \( H_0 \) is

\[
(1.5.2) \quad H_0: \bigcap_{j=1}^{p} H_{0j}, \text{ where } H_{0j}, \theta_j = 0, \text{ with } \theta_j(g_j \times 1) = C_j \beta_j;
\]

and
\( C_j(g_j \times m_j) \) is of rank \( g_j \).

Roy, Gnanadesikan and Srivastava (1971) described the following procedure in analyzing the above HM model:

For \( j = 1 \), we have a univariate linear model

\[
(1.5.3) \quad E(Y_1) = X_1 \beta_1,
\]
\( V(Y_1) = I_{n1} \otimes \sigma^2_{(1)}, \sigma^2_{(1)} = \sigma_{1,1} \).

For \( j < p \), \( Y_{j+1} \) can be analyzed through the conditional model

\[
(1.5.4) \quad E(Y_{j+1} | Y_j^*) = (X_{j+1}^* Y_j^*) \begin{bmatrix} Y_{j+1} \\ \xi_j \end{bmatrix},
\]
and
\( V(Y_{j+1} | Y_j^*) = I_{nj} \otimes \sigma^2_{(j+1)}, j = 1, \ldots, p-1, \)

where \( Y_{j+1(n_{j+1} \times 1)} \) is a data matrix,
\( Y^*_j(j(n_{j+1} \times j)) = \begin{bmatrix} Y_{j+1} & \cdots & Y_{j+1}^p \\ Y_1 & \cdots & Y_p \end{bmatrix} \) is a matrix of covariates,
and
\[ X^*_{j+1}(n_j \times p_{j+1} \sum_{k=1}^{m_k}) = \begin{bmatrix} x_{j+1}^{j+1} \\ x_{j+1}^{j+1} \\ \vdots \\ x_{j+1}^{j+1} \end{bmatrix} \]
is a design matrix with
\[ E(Y^*_r) = X_r^* \beta_r, \]
\[ \gamma_j + 1(\sum_{k=1}^{m_k} \times 1) = \begin{bmatrix} \beta_j^1 \\ \vdots \\ \beta_j^j \end{bmatrix} \]
is a matrix of unknown parameters,

\[ \beta_r^{j}(m_r \times 1) = -\beta_r \xi_{jr}, \quad r=1, \ldots, j, \]
\[ \xi_{j}(j \times 1) = \Sigma^{-1}_j \begin{bmatrix} \sigma_{1,j+1} \\ \vdots \\ \sigma_{j,j+1} \end{bmatrix} = \begin{bmatrix} \xi_{j1} \\ \vdots \\ \xi_{jj} \end{bmatrix} \]
is a matrix of unknown regression coefficients, and

\[ \sigma^2_{(j+1)} = \frac{\Sigma_{j+1}}{\Sigma_j}. \]

With models (1.5.3) and (1.5.4), the null hypothesis \( H_0 \) can be transformed into an equivalent hypothesis \( H_0^* \).

\[ (1.5.5) \quad H_0^*: \bigcap_{j=1}^{p} H_0^*, \text{ where } H_0^*: C_j \gamma_j = 0, \quad C_j(g_j \times \sum_{k=1}^{m_k}) \text{ is of rank } g_j. \]

Let \( \gamma_j^\hat{} \) be the least squares estimate (LSE) of \( \gamma_j \),
\[ \sigma^2_{(j)} \quad W_j \text{ be the covariance matrix of } C_j \gamma_j^\hat{}, \text{ and} \]
\[ s^2_{(j)} \quad \text{be the resulting sum of squares of residuals in the least squares estimation, } j=1, \ldots, p. \]

Then under \( H_0^* \), the statistic

\[ (1.5.6) \quad F_j = \frac{(C_j \gamma_j^\hat{})^T W_j^{-1} (C_j \gamma_j^\hat{}) / g_j}{s^2_{(j)} / (n_j - [(j-1) + \sum_{k=1}^{m_k}] )} \]
has an $F$ distribution with $g_j$ and $n_j - [(j-1) + \sum_{k=1}^{j} m_k]$ d.f.

Under $H^*_0: \bigcap_{j=1}^{p} H^*_0$, $F_j$'s are independently distributed. Thus, $H^*_0$ can be tested by the step-down procedure:

\[(1.5.7) \quad \text{Accept } H^*_0 \text{ if and only if } F_j \leq f_j, \quad j=1, \ldots, p, \]

where

\[\prod_{j=1}^{p} P(F_j \leq f_j | H^*_0) = 1 - \alpha.\]

1.6 The IM model

Assume that the experimental units can be divided into $u$ disjoint sets such that the same set of responses are measured on each unit in the same set. Srivastava (1968) defined the following IM model

\[(1.6.1) \quad E(Y_{j}) = X_j( \Gamma j), \beta^* B_j, \]

and

\[V(Y_{j}) = \sum_{j=1}^{u} \Sigma_j, \quad j=1, \ldots, u,\]

where

- $Y_j(n_j \times p_j)$ is a data matrix,
- $X_j(n_j \times [r_j + t])$ is a design matrix of rank $r_j + t - 1$,
- $\Gamma_j(r_j \times p_j)$ is a matrix of unknown parameters,
- $\beta^*(t \times p)$ is a matrix of common unknown parameters such that $J_{(1 \times t)} \beta^* = 0$, where $J_{(1 \times t)}$ is a (1 \times t) vector of 1's.
- $B_j(p \times p_j)$ is an incidence matrix of 0's and 1's,
- $\Sigma_j(p_j \times p_j) = B_j' \Sigma B_j$, and
- $\Sigma(p \times p)$ is a positive definite covariance matrix.

The hypothesis of interest is $H_0: C^* \beta^* U = 0$, where
Prior to analysis it is necessary for the data to be transformed into the framework of a standard MGLM which meets the specified conditions. Srivastava suggested the following procedure:

(i) First construct a linear set of $Y_j$, $Q_j(t \times p_j)$, such that

\begin{equation}
E(Q_j) = A_j \beta^* B_j,
\end{equation}

\begin{equation}
V(Q_j) = C_j \otimes \Sigma_j, \quad j=1,...,u.
\end{equation}

(ii) The design is homogeneous, i.e., there exists a set a known $(t \times t)$ matrices, \{${\mathbf{F}}_*^1, ..., {\mathbf{F}}_*^m$\}, such that

\begin{equation}
\sum_{k=1}^{m} {\mathbf{F}}_*^k = J_{(t \times t)} \quad \text{and}
\end{equation}

\begin{equation}
a_j = \sum_{k=1}^{m} a_{jk} {\mathbf{F}}_*^k \quad \text{for some known scalars } a_{j1}, ..., a_{jm}, \quad j=1,...,u.
\end{equation}

We then have

\begin{equation}
E(Q_j) = \sum_{k=1}^{m} \alpha_{jk} {\mathbf{F}}_*^k B_j = \sum_{k=1}^{m} (F_*^k \beta^*)(\alpha_{jk} B_j),
\end{equation}

where \( \alpha_{jk} = a_{jk} \theta_j \) for a fixed \( \theta_j \).

(iii) Combine $Q_1, ..., Q_u$.

\begin{equation}
Q_{(t \times u \times p_j)}(j=1,...,u) = (Q_1, ..., Q_u), \quad \text{we have}
\end{equation}

\begin{equation}
E(Q) = (E(Q_1), ..., E(Q_u))
\end{equation}

\begin{equation}
= \left( \sum_{k=1}^{m} (F_*^k \beta^*)(\alpha_{1k}^* B_1), ..., \sum_{k=1}^{m} (F_*^k \beta^*)(\alpha_{uk}^* B_u) \right)
\end{equation}

\begin{equation}
=(F_*^1 \beta^*, ..., F_*^m \beta^*) \mathbf{L}.
\end{equation}
where 
\[ L_{(mp \times \sum_{j=1}^{r} p_j)} = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix}, \]

and 
\[ L_k(p \times \sum_{j=1}^{u} p_j) = (\alpha_{1k} B_1, \ldots, \alpha_{uk} B_u). \]

(iv) Choose \( \{\theta_1, \ldots, \theta_u\} \), such that \( LL' \) is nonsingular.

Let 
\[ Z^*(t \times mp) = QL(LL')^{-1} = (Z_1^*, \ldots, Z_m^*) \]

and define
\[ Z(m[t-1] \times p) = \begin{pmatrix} Z_1 \\ \vdots \\ Z_m \end{pmatrix}, \]

where 
\[ Z_k([t-1] \times p) = (0([t-1] \times 1) I(t-1)) Z_k^*, \quad k = 1, \ldots, m. \]

Then we have
\[ E(Z) = \begin{pmatrix} F_1 \\ \vdots \\ F_m \end{pmatrix} \beta, \]

where
\[ F_k([t-1] \times [t-1]) = (0([t-1] \times 1) I(t-1)) F_k^* \begin{pmatrix} -I(t-1) \\ I(t-1) \end{pmatrix} \]
\[ \beta([t-1] \times p) = (0([t-1] \times 1) I(t-1) \beta^*. \]

(v) The factorization of \( V(Z) \) is possible, i.e. there exists a known positive definite matrix \( W(m[t-1] \times m(t-1)) \), such that
\[ V(Z) = W \otimes \Sigma^*, \]

where \( \Sigma^* \) is a positive definite matrix.

(vi) If the above conditions are met, then there exists a nonsingular matrix \( R(m[t-1] \times m(t-1)) \), such that \( RW R' = I_{m(t-1)}. \)

Thus we have a standard MGLM model
\[ E(RZ) = \begin{pmatrix} F_1 \\ \vdots \\ F_m \end{pmatrix} \beta, \]
\[ V(RZ) = I_{m(t-1)} \otimes \Sigma^*. \]
With \( C\*B^* = CB \), where \( C_{g\times[t-1]} = C \begin{bmatrix} 0(1\times[t-1]) \\ I(t-1) \end{bmatrix} \)

\[- [C \begin{bmatrix} J(1\times1) \\ 0([t-1]\times1) \end{bmatrix}] \otimes J(1\times[t-1]), \]

the null hypothesis \( H_0 \) is equivalent to \( H_0 : C\*U = 0. \)

Then the usual analysis procedures used in the standard MGLM model become available.

### 1.7 Outline of the research

Chapter II deals with the construction of confidence procedures for the generalized multivariate linear models. These models include SGCM, GGCM, and HM models.

In the SGCM model covariate adjustment procedures have been used to reduce the model into a standard MANOCOVA model. The procedure requires the selection of a nonsingular matrix \( H \) satisfying a set of conditions (1.4.4). Although the selection of the matrix \( H \) is not unique, the resulting confidence procedure is shown to be invariant under different selections of \( H \) satisfying the same set of conditions.

In the GGCM model, with enough observations and responses being measured in each disjoint set, we extend the above covariate adjustment procedure to each disjoint set. A univariate weighted least squares model is then constructed to simultaneously utilize the information from each disjoint set. Again, the resulting confidence procedure is shown to be invariant under different selections of the matrix \( H \) when the set of conditions is satisfied in each disjoint set.

In an HM model, the \( p \) responses are ordered. Using the correlation in-
formation among the p responses, the jth response can be analyzed through a covariate model with the covariates being the (j-1) previous responses, j=2,...,p. With good estimates of the variances of the p responses, the resulting confidence set can be constructed from p quadratic forms of χ² statistics. We demonstrate the confidence procedure when p=2. To optimize the confidence procedure, we propose two criteria involving suitable norms of the confidence set. These criteria are to minimize either the weighted sum of the p quadratic form lengths or the maximum of the p weighted quadratic form lengths. The weight is the inverse of the corresponding degrees of freedom. For the maximum weighted length criterion, the resulting optimal confidence procedure gives a larger confidence coefficient to the variate with a larger d.f. The result is also true for the weighted sum criterion when the degrees of freedom are less than five.

Chapter III considers the IM model. The model defined by Srivastava is based on a less-than-full-rank linear model in each of the disjoint datasets. For a design satisfying a set of specified conditions, this model can be analyzed by transforming the data into the framework of a standard MGCM model. We propose a corresponding full rank IM model and show some simplification of the procedure.

Consider a special IM design upon which the full rank model procedure can be performed. We construct an optimum confidence procedure to minimize the generalized variance of the parameter estimate \( \hat{\beta} \) and show some invariance property of the procedure. We then study the relative performance of the proposed special IM model and the corresponding complete multiresponse (CM) model. These comparisons include generalized variance, asymptotic relative efficiency, cost, and minimum risk. In each comparison, we assume the covariance matrix Σ has either an intraclass covariance structure or an autocorrelation with p=3.
Chapter IV examines various covariance structures in blocks. When observations in a design arise in a structured form, the resulting covariance matrix may have a special pattern. A common example of patterned covariance structure is the intraclass correlation (IC) model. With the same \( p \) responses measured at each of \( t \) distinct time points, a natural extension of the IC model is the block version of the IC model. We examine the hypotheses of the block IC model as well as some special cases in the standard MGLM model. These hypotheses are

\[
H_1: \Sigma_0 = I \otimes \Sigma_1 + (J-I) \otimes \Sigma_2,
\]

\[
H_2: \Sigma_2 = \omega \Sigma_1 \text{ in } H_1,
\]

\[
H_3: \Sigma_1 = \sigma_1 [I+\rho_1 (J-I)], \Sigma_2 = \sigma_2 [I+\rho_2 (J-I)] \text{ in } H_1, \text{ and}
\]

\[
H_4: \rho_2 = 1 \text{ in } H_3.
\]

In each hypothesis, the LR test is derived. For the hypotheses \( H_1, H_3, \) and \( H_4 \), we have a closed form expression for each test statistic. Moments of the test statistics are obtained and used to find asymptotic null and non-null distributions. For the hypothesis \( H_2 \), there is no closed form expression for the LR statistic. We construct the Wald statistic to test the hypothesis and then derive the asymptotic null and non-null distributions.

We outline proposals for further research in the concluding chapter of the dissertation.
CHAPTER II

ON CONSTRUCTING THE CONFIDENCE SETS IN THE
GENERALIZED MULTIVARIATE LINEAR MODELS

2.1 Introduction

In this chapter we construct confidence procedures for the generalized multivariate linear models. These models include SGCM, GGCM, and HM models. We then derive some optimum properties related to these constructed confidence procedures.

Section 2.2 examines the SGCM model. Using the covariate adjustment procedure, it is necessary to construct a nonsingular matrix $H_{(q \times q)} = (H_1(q \times p), H_2(q \times [q-p]))$, such that

$$PH_1 \text{ is nonsingular,}$$
$$PH_2 = 0.$$  

Section 2.2.2 shows that the covariate adjustment procedure is invariant under different selections of the nonsingular matrix $H$ satisfying the conditions (2.1.1).

Section 2.3 examines the GGCM model. Assume that we have enough observations in each disjoint set ($n_j \geq r+q_j$, $j=1,\ldots,u$) and enough responses on each observation ($q_j \geq p$, $j=1,\ldots,u$).

Section 2.3.1 extends the covariate adjustment procedure in the SGCM model to the GGCM model by using the covariate adjustment in each of the $u$ disjoint sets. A univariate weighted least squares model is constructed to simultaneously utilize the information from each of the $u$ disjoint sets.
The confidence set in the GGCM model can then be constructed from the Wald type statistic in the univariate model.

The covariate adjustment procedure in the GGCM model requires the construction of nonsingular matrices $H_j(q_j xq_j) = (H_{j1}(q_j x) , H_{j2}(q_j x[q_j-p]))$,  

(2.1.2)  

$\sum_j H_{j1} = I$,  

$\sum_j H_{j2} = 0, \quad j=1, \ldots, u$.  

Section 2.3.2 shows that the covariate adjustment procedure in the GGCM model is invariant under different selections of these nonsingular matrices $H_j, j=1, \ldots, u$, satisfying the conditions (2.1.2).

Section 2.4 examines the HM model. Using the correlation information among the p responses, the $j^{th}$ response can be analyzed through a covariates model with the covariates being the $(j-1)$ previous responses, $j=2, \ldots, p$. The resulting confidence set can then be constructed from the p independent confidence procedures involving the F statistics. Assuming we have a good estimate for $\Sigma$, the resulting confidence set can be constructed from the p independent confidence procedures involving the $\chi^2$ statistics.

Section 2.4.1 constructs confidence procedures in the HM model when $p=2$. Section 2.4.2 proposes two criteria by which the resulting p confidence sets can be optimized. We then derive the optimum solutions for these two criteria.

Section 2.5 gives numerical examples to illustrate the procedures derived in this chapter.

2.2 The standard growth curve model

2.2.1 Covariate adjustment in the SGCM model

With the construction of the nonsingular matrix $H(qxq) = (H_1(qxp))$,  

...
such that \( \Phi \) is nonsingular and \( \Phi^2 = 0 \). We can analyze the SGCM model through the following conditional model:

\[
(2.2.1) \quad E(Y|Z) = X\eta + Z\Gamma, \\
V(Y|Z) = I \otimes \Sigma,
\]

where \( Y_{(n \times p)} = Y_0H_1 \) is the new data matrix,

\( Z_{(n \times (q-p))} = Y_0H_2 \) is the matrix of covariates,

\( X_{(n \times r)} \) is the design matrix,

\( \eta_{(r \times p)} = \beta \Phi_1 \) is the matrix of unknown parameters,

\( \Gamma_{((q-p) \times p)} \) is the matrix of unknown regression coefficients, and

\( \Sigma_{(p \times p)} = H_1'P'(P_0^{-1}P')^{-1} \) \( \Phi_1 \) is the unknown conditional covariance.

The LSE's of \( \eta, \Gamma \) are given by

\[
(2.2.2) \quad \begin{pmatrix}
\hat{\eta} \\
\hat{\Gamma}
\end{pmatrix} = \begin{pmatrix}
X'X & X'Z \\
Z'X & Z'Z
\end{pmatrix}^{-1} \begin{pmatrix}
X'Y \\
Z'Y
\end{pmatrix}
\]

\[
= \begin{pmatrix}
(X'L_1X)^{-1} & -(X'L_1X)^{-1}X'Z(Z'Z)^{-1} \\
-(Z'Z)^{-1}Z'X(X'L_1X)^{-1} & (Z'Z)^{-1}Z'X(X'L_1X)^{-1}X'Z(Z'Z)^{-1}
\end{pmatrix} \begin{pmatrix}
X'Y \\
Z'Y
\end{pmatrix}
\]

\[
= \begin{pmatrix}
(X'L_1X)^{-1}X'L_1Y \\
(Z'Z)^{-1}Z'[I-X(X'L_1X)^{-1}X'L_1]Y
\end{pmatrix},
\]

where \( L_1 = I - Z(Z'Z)^{-1}Z' \).

For the standard MGLM model, three frequently used confidence procedures are Hotelling's trace, Roy's largest root, and Wilk's likelihood ratio. The resulting confidence sets for \( \theta = \Theta \cup = Cn(\Phi_1)^{-1} \) are functions of \( S_H^{-1} S_E^{-1} \), where

\[
(2.2.3) \quad S_H^* = (\hat{\theta} - \theta)' (CRC')^{-1} (\hat{\theta} - \theta),
\]
\[ S_E = [(PH_1)^{-1}U]' S[(PH_1)^{-1}U], \]
\[ \hat{\theta} = C \hat{\eta} (PH_1)^{-1}U, \]
\[ R = (X'X_1X)^{-1}, \]
\[ S = (Y - \hat{\eta} - Z\hat{\eta})' (Y - \hat{\eta} - Z\hat{\eta}) \]
\[ = Y'[I - (XZ) \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix}^{-1} \begin{bmatrix} X' \\ Z' \end{bmatrix}] Y, \]
\[ \hat{\eta} \text{ is given by (2.2.2).} \]

2.2.2 Invariant confidence procedure

The selection of the matrix \( H \) in the above procedure is not unique. The nonsingularity property of \( PH_1 \) includes the following special cases:

1. The columns of \( H_1 \) form a basis of the vector space generated by the rows of \( P \), which was used by Rao,

2. \( PH_1 = I \), by Khatvi.

Assume both matrices \( H = (H_1, H_2) \) and \( H^* = (H_1^*, H_2^*) \) satisfy the specified conditions, namely \( PH_1 \) and \( PH_1^* \) are nonsingular, and \( PH_2 = PH_2^* = 0 \). Then there exist matrices \( G_1(pxp), G_2((q-p)xp), \) and \( G_3((q-p)x(q-p)) \), such that \( G_1 \) and \( G_3 \) are nonsingular, and

\[ H_1^* = H_1 G_1 + H_2 G_2, \]
\[ H_2^* = H_2 G_3. \]

Thus we can examine the effect of different selections of the matrix \( H \) on the resulting confidence sets for \( \theta \) through the transformation

\[ (Y, Z) \rightarrow (YG_1 + ZG_2, ZG_3). \]
Letting

(2.2.6) \[ Y^* = YG_1 + ZG_2, \]

(2.2.7) \[ Z^* = ZG_3, \]

we then have

(2.2.8) \[ E(Z^*) = E(Z)G_3 = 0, \]

(2.2.9) \[ E(Y^*|Z^*) = E(YG_1 + ZG_3 | ZG_3) = E(YG_1 + ZG_2 | Z) \]

\[ = E(Y|Z)G_1 + ZG_2 = (Xn+Zp)G_1 + ZG_2 \]

\[ = XnG_1 + Z(\Gamma G_1 + G_2) = XnG_1 + ZG_3^{-1}(\Gamma G_1 + G_2), \]

(2.2.10) \[ V(Y^*|Z^*) = V(YG_1 + ZG_2 | Z) = V(YG_1 | Z) = I \otimes [G_1\Sigma G_1], \]

(2.2.11) \[ (nG_1)(PH_1)^{-1} = nG_1(PH_1G_1)^{-1} = n(PH_1). \]

**Lemma 2.2.1** With the transformation given by (2.2.5), in the parameter space, we have the following induced transformation:

(2.2.12) \[ (B,n,\Gamma,\Sigma) \rightarrow (B,nG_1,G_3^{-1}[\Gamma G_1 + G_2], G_1G_1). \]

From the transformation (2.2.5), we have

(2.2.13) \[ L_1^* = I - Z^*(Z^*Z^*)^{-1}Z^* = I - ZG_3[(ZG_3)'(ZG_3)]^{-1}(ZG_3)' \]

\[ = I - Z(Z'Z)^{-1}Z', \]

\[ = L_1, \]

\[ L_1^*Y^* = L_1(YG_1 + ZG_2) = L_1YG_1 + L_1ZG_2 \]

\[ = L_1YG_1 + [I - Z(Z'Z)^{-1}Z'] ZG_2 = L_1YG_1, \]

\[ \hat{n}^* = (X'L_1^*X)^{-1} X'L_1^*Y^* = (X'L_1X)^{-1} X'L_1YG_1 = \hat{n} G_1, \]
\[ P_{H_1}^* = P(H_1 G_1 + H_2 G_2) = P H_1 G_1 \]
\[ \hat{\theta}^* = C \hat{\beta}^* = C \eta^*(P H_1^*)^{-1} U = C \eta G_1 (P H_1 G_1)^{-1} U = \hat{\theta} \]

Thus \( S_H^* = (\hat{\theta} - \theta)'(C \Sigma C')^{-1}(\hat{\theta} - \theta) \) is invariant under the transformation.

Furthermore, after some simplification, we have

\[ (2.2.13) \quad I - (XZ) \left( \begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z \end{pmatrix}^{-1} \right) = L_1 - L_1 X (X' L_1 X)^{-1} X' L_1. \]

Thus, after the transformation,

\[ (2.2.14) \quad S^* = Y^* \left[ L_1 - L_1 X (X' L_1 X)^{-1} X' L_1 \right] Y^* \]
\[ = Y^* L_1 Y^* - Y^* L_1 X (X' L_1 X)^{-1} X' L_1 Y^* \]
\[ = Y^* L_1 Y G_1 - Y^* L_1 X (X' L_1 X)^{-1} X' L_1 Y G_1 \]
\[ = G_1 Y' L_1 Y G_1 - G_1 Y' L_1 X (X' L_1 X)^{-1} X' L_1 Y G_1 \]
\[ = G_1 S G_1, \]

\[ S_E^* = [(P H_1^*)^{-1} U]' S^* [(P H_1^*)^{-1} U] \]
\[ = [(P H_1 G_1)^{-1} U]' (G_1 S G_1) [(P H_1 G_1)^{-1} U] \]
\[ = S_E. \]

Since both matrices \( S_H^* \) and \( S_E \) are invariant under the transformations (2.2.5), (2.2.11), the confidence procedures based on the information from \( S_H S_E^{-1} \) are invariant under these transformations.

**Theorem 2.2.2.** The confidence procedures based on the information from \( S_H^* S_E^{-1} \), where \( S_H^* \) and \( S_E \) are given by (2.2.3), are invariant under different selections of the matrix \( H = (H_1, H_2) \) such that \( P H_1 \) is nonsingular and \( P H_2 = 0 \).
2.3 The generalized growth curve model

2.3.1 Covariate adjustment in the GGCM model

Recalling Section 1.4.2, the GGCM model is given by

\[(2.3.1) \quad E(Y_{0j}) = X_j \beta P B_j, \]
\[V(Y_{0j}) = I_{n_j} \otimes B_j^T \Sigma_0 B_j, \quad j=1, \ldots, u, \]

where

- \( Y_{0j} \) is \((n_j \times q_j)\),
- \( X_j \) is \((n_j \times r)\),
- \( \beta \) is \((r \times p)\),
- \( P \) is \((p \times q)\),
- \( B_j \) is \((q \times q_j)\),
- \( \Sigma_0 \) is \((q \times q)\).

The null hypothesis of interest is

\[(2.3.2) \quad H_0: \theta = 0, \quad \theta = C \theta U \quad \text{where} \]
\[C \in \mathbb{R}^{c \times r} \text{ is of rank } c, \quad U \in \mathbb{R}^{p \times v} \text{ is of rank } v. \]

Assuming \( q_j \geq p, \ n_j \geq r + q_j, \ j=1, \ldots, u \), then the covariate adjustment procedure used in the SGCM model can be extended to the GGCM model. For each \( j, \ j=1, \ldots, u \), we can construct a nonsingular matrix

\[(2.3.3) \quad H_j(q_j \times q_j) = (H_{j1}(q_j \times p), H_{j2}(q_j \times [q_j - p])), \quad \text{such that} \]
\[(2.3.4) \quad P B_j H_{j1} = I, \quad P B_j H_{j2} = 0. \]

Lemma 2.3.1. The GGCM model can be analyzed through the following conditional models:
(2.3.5) \[ E(Y_j | Z_j) = X_j \beta + Z_j \Gamma_j, \]
\[ V(Y_j | Z_j) = I_n_j \otimes \Sigma_j, \quad j=1, \ldots, u. \]

where

(2.3.6) \[ Y_j(n_j \times p) = Y_0 \mathbf{H}_j, \]
\[ Z_j(n_j \times [q_j - p]) = Y_0 \mathbf{H}_j, \]
\[ \Sigma_j = [PB_j (B_j \Sigma_j B_j)^{-1}(PB_j)^{'}]^{-1}, \]
\[ \mathbf{H}_j, \mathbf{H}_j^2 \text{are given by} (2.3.3). \]

Lemma 2.3.2. The conditional models given by (2.3.5) are equivalent to the following univariate models:

(2.3.7) \[ E(Y_j^v) = A_j \begin{pmatrix} \beta_j^v \\ \Gamma_j^v \end{pmatrix}, \]
\[ V(Y_j^v) = I_n_j \otimes \Sigma_j, \quad j=1, \ldots, u, \text{where} \]

(2.3.8) \[ A_j(n_j \times [r+q_j - p]) = (X_j \otimes I_p, Z_j \otimes I_p), \]

\( Y_j^v, \beta^v \) and \( \Gamma^v_j \) are vectors constructed from \( Y_j, \beta \) and \( \Gamma_j \) by putting the columns of the matrix underneath each other.

Since \( \Sigma_j \)'s are different from each other, the models given by (2.3.7) can be analyzed through a weighted least squares procedure.

Theorem 2.3.3. The GGCM model given by (2.3.1) can be analyzed by the following weighted least squares model:

(2.3.9) \[ E(Y^v) = A \xi, \]
\[ V(Y^v) = \Omega, \]

where
From the GGLM model (2.3.1), an unbiased and consistent estimator of 
\[ \Sigma_0 = (\sigma_{k,l}) \], \( \hat{\Sigma}_0 = (\hat{\sigma}_{k,l}) \), can be obtained by the pooled estimate of \( \sigma_{k,l} \) from those experimental units on which both responses \( V_k \) and \( V_\ell \) are observed. That is,

\[
(2.3.11) \quad \hat{\sigma}_{k,l} = \frac{1}{N_{k\ell} - r} \, \text{tr} \left[ I - X_{k\ell} (X_{k\ell}')^{-1} X_{k\ell} \right] \, Y_{k(l)}, \quad k, \ell = 1, \ldots, q,
\]

where \( Y_{k(l)} (N_{k\ell} \times 1) \) is the observation vector on \( V_k \) corresponding to those experimental units on which both \( V_k \) and \( V_\ell \) are observed, \( X_{k\ell} (N_{k\ell} \times r) \) is the design matrix corresponding to \( Y_{k(l)} \).
Denoting
\[ \hat{\Sigma}_j = \Sigma_j(\hat{\Sigma}_0) = [PB_j(B_j'\Sigma_0 B_j)^{-1}B_j']^{-1}, \]

\[ \hat{\Omega} = \Omega(\hat{\Sigma}_j = \hat{\Sigma}_j) = \begin{pmatrix} I_{n_1} \otimes \hat{\Sigma}_1 \\ \vdots \\ I_{n_u} \otimes \hat{\Sigma}_u \end{pmatrix}, \]

then a weighted least squares estimate of \( \xi \) is given by
\[ \hat{\xi} = (A'\hat{\Omega}^{-1}A)^{-1}A'\hat{\Omega}^{-1}Y. \]

Using the matrix formulae involving Kronecker products, we have
\[ A'\hat{\Omega}^{-1} = \begin{pmatrix} X_1' \otimes I_p & X_2' \otimes I_p & \cdots & X_u' \otimes I_p \\ Z_1' \otimes I_p & Z_2' \otimes I_p & \cdots & Z_u' \otimes I_p \end{pmatrix} \begin{pmatrix} I_{n_1} \\ \vdots \\ I_{n_u} \end{pmatrix}, \]

\[ = \begin{pmatrix} X_1' \otimes \hat{\Sigma}_1 & X_2' \otimes \hat{\Sigma}_2 & \cdots & X_u' \otimes \hat{\Sigma}_u \\ Z_1' \otimes \hat{\Sigma}_1 & Z_2' \otimes \hat{\Sigma}_2 & \cdots & Z_u' \otimes \hat{\Sigma}_u \end{pmatrix}, \]

\[ A'\hat{\Omega}^{-1}A = \begin{pmatrix} \sum_{j=1}^{u} (X_jX_j) \otimes \hat{\Sigma}_j^{-1} & X_1'Z_1 \otimes \hat{\Sigma}_1^{-1} & \cdots & X_u'Z_u \otimes \hat{\Sigma}_u^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_1'Z_1 \otimes \hat{\Sigma}_1^{-1} & Z_2'Z_2 \otimes \hat{\Sigma}_2^{-1} & \cdots & Z_u'Z_u \otimes \hat{\Sigma}_u^{-1} \end{pmatrix}, \]

(sym.)
\[ A' \hat{\Omega}^{-1} Y^v = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j ight) \]

\[ = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j \right) \]

\[ = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j \right) \]

\[ = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j \right) \]

\[ = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j \right) \]

\[ = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j \right) \]

\[ = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j \right) \]

\[ = \left( \sum_{j=1}^{u} (X_j' \otimes \hat{\Sigma}_j) Y^v_j \right) \]

Denoting

(2.3.15)

\[ R_{00} = \sum_{j=1}^{u} (X_j' X_j) \otimes \hat{\Sigma}_j, \]

\[ R_{0j} = (X_j' Z_j) \otimes \hat{\Sigma}_j, \]

\[ R_{jj} = Z_j' Z_j \otimes \hat{\Sigma}_j, \]

\[ D_{00} = \left( R_{00} - \sum_{j=1}^{u} R_{0j} R_{jj} R_{0j} \right)^{-1}, \]

\[ D_{0j} = -D_{00} R_{0j} R_{jj}, \]

then with appropriate \( D_{jj} \)'s we have

(2.3.16)

\[ A' \hat{\Omega}^{-1} A = \begin{pmatrix} R_{00} & R_{01} & R_{02} & \cdots & R_{0u} \\ & R_{11} & & & \\ & & R_{12} & \cdots & \\ & & & \text{(sym.)} & R_{uu} \end{pmatrix}, \]

\[ \{A' \hat{\Omega}^{-1} A\}^{-1} = \begin{pmatrix} D_{00} & D_{01} & D_{02} & \cdots & D_{0u} \\ & D_{11} & & & \\ & & D_{22} & \cdots & \\ & & & \text{(sym.)} & D_{uu} \end{pmatrix}. \]

With \( \hat{\xi} = \{A' \hat{\Omega}^{-1} A\}^{-1} A' \hat{\Omega}^{-1} Y^v \), we have

(2.3.17)

\[ \hat{B}^v = (D_{00} D_{01} D_{02} \cdots D_{0u}) A' \hat{\Omega}^{-1} Y^v. \]
Theorem 2.3.4  For the model given by (2.3.9), a weighted least squares estimate of $\beta^V$ is given by (2.3.17) and an estimate asymptotic variance for the estimator $\hat{\beta}^V$ is

\[
V(\hat{\beta}^V) = D_{00} = \sum_{j=1}^{u} (X_j' L_j X_j) \hat{\Sigma}_j^{-1},
\]

where $D_{00}$ is given by (2.3.15),

\[
L_j = I - Z_j (Z_j' Z_j)^{-1} Z_j',
\]

$\hat{\Sigma}_j$ is given by (2.3.12).

Letting $H = C \otimes U'$, the null hypothesis $H_0: C \delta U = 0$ is then equivalent to the hypothesis $H^*: H\beta^V = 0$. With large values of $n_1, \ldots, n_u$, the sample sizes in the $u$ disjoint sets, the resulting weighted least squares estimate $\hat{H}\beta^V$ is asymptotically distributed as a normal variate with mean $H\beta^V$ and variance $HV(\beta^V)H'$.

Corollary 2.3.5.  A $(1-\alpha)$ confidence set for $H\beta^V$ is given by

\[
(\hat{H}\beta^V - H\beta^V)' [H(V(\beta^V)) H']^{-1} (\hat{H}\beta^V - H\beta^V) \leq \chi^2_{cv, 1-\alpha}.
\]

With the estimate variance for $H\beta^V$, $V(H(\beta^V))$, an approximate $(1-\alpha)$ confidence set for $H\beta^V$ is given by

\[
W_n^* \leq \chi^2_{cv, 1-\alpha}.
\]
where
\[ W^*_n = (H\hat{\beta}^V - H\beta^V)' [HV(\hat{\beta}^V)H]'^{-1} (H\hat{\beta}^V - H\beta^V). \]

Putting $H\beta^V = 0$ in $W^*_n$, we have the following corollary:

**Corollary 2.3.6.** A test statistic for the null hypothesis $H_0$ given by (2.3.2) is

\[ (2.3.21) \quad W_n = (H\hat{\beta}^V)' [HV(\hat{\beta}^V)H]'^{-1} (H\hat{\beta}^V). \]

Under $H_0$, $W_n$ has an approximate $\chi^2$ distribution with cv d.f. We reject $H_0$ for large values of $W_n$.

When $u = 1$, the GGCM model reduces to the SGCM model. The above procedure then gives

\[ (2.3.22) \quad \hat{\Sigma}_0 = \frac{1}{n-r} S, \]
\[ \hat{\Sigma} = (P\hat{\Sigma}^{-1}P')^{-1}, \]
\[ D_{00} = \begin{pmatrix} R_{00} & R_{01} \\ \bar{R}_{10} & \bar{R}_{11} \end{pmatrix}^{-1} \]
\[ = \begin{pmatrix} (X'X) \otimes \hat{\Sigma}^{-1} - (X'Z) \otimes \hat{\Sigma}^{-1} \\ (Z'Z) \otimes \hat{\Sigma}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} (X'X) \otimes \hat{\Sigma}^{-1} \\ (Z'Z) \otimes \hat{\Sigma}^{-1} \end{pmatrix}^{-1} \]
\[ = (X'LY)^{-1} \otimes \hat{\Sigma}, \]

where
\[ S = Y_0'[I - X(X'X)^{-1}X']Y_0, \]
\[ L = I - Z(Z'Z)^{-1}Z'. \]

Thus we have

\[ (2.3.23) \quad \hat{\beta}^V = D_{00} [X'LY \otimes \hat{\Sigma}^{-1}]Y^V \]
\[ = [(X'LY)^{-1}X'LY \otimes I_p]Y^V, \]

which results in the same estimate for $\beta$ obtained by covariate adjustment procedure in the SGCM model.
With
\[
(2.3.24) \quad H^\ast (\hat{\beta}^\ast) H^\ast = [C \otimes U'][(X'LX)^{-1} \otimes \hat{\Sigma}][C' \otimes U]
\]
\[
= [C(X'LX)^{-1}C'] \otimes [U'U],
\]
we have
\[
W_n^\ast = (HB^\ast - HB^\ast)'[H(\hat{\beta}^\ast)H^\ast]^{-1}(HB^\ast - HB^\ast)
\]
\[
= ([C \otimes U'](\hat{\beta}^\ast - \beta^\ast))' \{[(C(X'LX)^{-1}C')^{-1} \otimes [U'U]^{-1}]

\cdot ([C \otimes U'](\hat{\beta}^\ast - \beta^\ast))
\]
\[
= (n-r) \text{trace } S_H^* S_E^{-1},
\]
where
\[
S_H^* = [C\hat{B}U - CBU]' [C(X'LX)^{-1}C']^{-1} [CBU - CBU]
\]
\[
S_E = (n-r) U'U = U'[\hat{P}^{-1}P']^{-1}U.
\]

Corollary 2.3.7. When \( u = 1 \), the covariate adjustment procedure in the GGCM model reduces to the trace criterion used in the covariate adjustment procedure in the SGCM model.

### 2.3.2 Invariant confidence procedure

In the above covariate adjustment procedure for the GGCM model, it is necessary to construct a set of nonsingular matrices, \( \{H_1,...,H_u\} \), such that

\[
(2.3.25) \quad H_j = (H_{j1},H_{j2}),
\]
\[
P_{B_jH_{j1}} = I,
\]
\[
P_{B_jH_{j2}} = 0, \quad j=1,...,u.
\]

The selection of the matrices \( \{H_1,...,H_u\} \) is not unique. Assuming both sets of matrices \( \{H_1,...,H_u\} \) and \( \{H_1^*,...,H_u^*\} \) satisfy the specified conditions (2.3.25), then there exist matrices \( G_{j2}, G_{j3} \), where \( G_{j3} \) is nonsingular,
j=1,...,u, such that

\[(2.3.26)\quad H_{j1}^* = H_{j1} + H_{j2} G_{j2},\]
\[(2.3.26)\quad H_{j2}^* = H_{j2} G_{j3}, \quad j=1,...,u.\]

Thus we can examine the effect of different selections of the set of matrices \(\{H_1,...,H_u\} \) on the resulting confidence sets for \(\theta\) through the transformations

\[(2.3.27)\quad (Y_j, Z_j) \rightarrow (Y_j + Z_j G_{j2}, Z_j G_{j3}), \quad j=1,...,u.\]

Letting

\[(2.3.28)\quad Y_j^* = Y_j + Z_j G_{j2},\]
\[(2.3.28)\quad Z_j^* = Z_j G_{j3}, \quad j=1,...,u,\]

then

\[(2.3.29)\quad E(Y_j^* | Z_j^*) = E(Y_j + Z_j G_{j2} | Z_j^*) = X_j \beta + Z_j G_{j2},\]
\[(2.3.29)\quad V(Y_j^* | Z_j^*) = V(Y_j + Z_j G_{j2} | Z_j^*) = V(Y_j | Z_j^*),\]
\[(2.3.29)\quad = I_n \otimes \Sigma_j, \quad j=1,...,u.\]

**Lemma 2.3.8.** With the transformations given by (2.3.27), in the parameter space, we have the following induced transformations

\[(2.3.30)\quad (\beta, \Gamma_j, \Sigma_0) \rightarrow (\beta, G_{j3}^{-1}(\Gamma_j + G_{j2}), \Sigma_0), \quad j=1,...,u.\]

From the transformations (2.3.27),

\[(2.3.31)\quad L_j^* = I - Z_j^* (Z_j^* Z_j^*)^{-1} Z_j^*,\]
\[ (Z_j G_j)^{-1}(Z_j G_j)^{\prime} = I - Z_j (Z_j^\prime Z_j)^{-1} Z_j \]

Thus the estimators

\( \hat{\Sigma}_j = [P B_j (B_j^\prime \hat{\Sigma}_0 B_j)^{-1} (P B_j)^{\prime}]^{-1}, \quad j = 1, \ldots, u, \)

are invariant under the transformations.

With \( L_j Z_j = 0, j = 1, \ldots, u, \)

\begin{align*}
[(X_j^\prime L_j^\prime) \otimes \hat{\Sigma}_j] Y_j^v &= [(X_j^\prime L_j^\prime) \otimes \hat{\Sigma}_j^{-1}] [Y_j^v + (Z_j G_j)^{\prime}] \\
&= [(X_j^\prime L_j^\prime) \otimes \hat{\Sigma}_j^{-1}] Y_j^v.
\end{align*}

Both the estimator \( \hat{\beta}^v = \left\{ \sum_{j=1}^{u} (X_j^\prime L_j X_j) \otimes \hat{\Sigma}_j^{-1} \right\}^{-1} \left\{ \sum_{j=1}^{u} [(X_j^\prime L_j) \otimes \hat{\Sigma}_j^{-1}] Y_j^v \right\} \) and its estimate variance \( V(\hat{\beta}^v) = \left\{ \sum_{j=1}^{u} (X_j^\prime L_j X_j) \otimes \hat{\Sigma}_j^{-1} \right\}^{-1} \) are invariant under the transformations (2.3.27).

**Theorem 2.3.9.** The covariate adjustment procedure for the GGCM model in Section 2.3.1 is invariant under different selections of the set of matrices \( \{H_1, \ldots, H_u\} \) satisfying the conditions (2.3.25).

### 2.4 The hierarchical multiresponse model

#### 2.4.1 The HM model with \( p = 2 \)

Section 1.5 describes the HM model. When \( p = 2 \) we have the response data

\begin{equation}
Y = \begin{pmatrix}
Y_{11} & Y_{12} \\
Y_{12} & Y_{22}
\end{pmatrix}, \text{ with } Y_{11}(n_1' \times 1), \ Y_{12}(n_2' \times 1), \ Y_{22}(n_2' \times 1), \ n_1' = n_1 - n_2.
\end{equation}
The resulting model is then given by

\[ E(Y) = \begin{bmatrix} E(Y_{11}) \\ E(Y_{12}) \\ E(Y_{22}) \end{bmatrix} = \begin{bmatrix} X_{11} \beta_1 \\ X_{12} \beta_1 \\ X_{22} \beta_2 \end{bmatrix}, \]

\[ V(Y_{11}) = I_{n_1} \otimes \sigma_{11}, \quad V(Y_{12}, Y_{22}) = I_{n_2} \otimes \Sigma_2, \]

where \( X_{11}(n_1 \times m_1), X_{12}(n_2 \times m_1), X_{22}(n_2 \times m_2) \) are design matrices,

\[ \beta_1(m_1 \times 1), \beta_2(m_2 \times 1) \] are matrices of unknown parameters, and

\[ \Sigma_2 = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \] is the covariance matrix.

The hypothesis of interest \( H_0 \) is

\[ (2.4.3) \quad H_0: \bigcap_{j=1}^{2} H_{0j}, \text{ where } H_{0j}: \theta_j = 0, \text{ with } \theta_j = C_j \beta_j, C_j(g_j \times m_j) \text{ is of rank } g_j. \]

Assuming the design matrices for \( Y_{12} \) and \( Y_{22} \) are the same, \( X_{12} = X_{22} = X_2 \), and utilizing the correlation information between \( Y_{12} \) and \( Y_{22} \), then the data \( Y_{22} \) can be analyzed by the conditional model

\[ (2.4.4) \quad E(Y_{22} | Y_{12}) = E(Y_{22}) + [Y_{12} - E(Y_{12})] \sigma_{11}^{-1} \sigma_{12} \]

\[ = X_2 \beta_2 + Y_{12} \sigma_{11}^{-1} \sigma_{12} - X_2 \beta_1 \sigma_{11}^{-1} \sigma_{12} \]

\[ = X_2 [\beta_2 - \beta_1 \sigma_{11}^{-1} \sigma_{12}] + Y_{12} \sigma_{11}^{-1} \sigma_{12} \]

\[ = (X_2 Y_{12}) \begin{bmatrix} \beta_2^* \\ Y_{2} \end{bmatrix}, \]

\[ V(Y_{22} | Y_{12}) = I_{n_2} \otimes \sigma_{(2)}^2, \]

where

\[ \beta_2^* = \beta_2 - \beta_1 \sigma_{11}^{-1} \sigma_{12} = \beta_2 - \beta_1 Y_2, \]

\[ Y_2 = \sigma_{11}^{-1} \sigma_{12}, \]

\[ \sigma_{(2)}^2 = \sigma_{22} - \sigma_{11}^{-1} \sigma_{12}^2. \]
Together with the unconditional model for the first variable

\[ (2.4.5) \]

\[
\begin{bmatrix}
Y_{11} \\
Y_{12}
\end{bmatrix} = \begin{bmatrix} X_{11} \\
X_2
\end{bmatrix} \beta_1,
\]

\[
V\left(\begin{bmatrix}
Y_{11} \\
Y_{12}
\end{bmatrix}\right) = I_n \otimes \sigma^2_1, \quad \sigma^2_1 = \sigma_{11},
\]

the unknown parameters \( \beta_1, \beta_2, \Sigma_2 \) can then be estimated.

**Lemma 2.4.1.** The LSE's for the parameters in models (2.4.4) and (2.4.5) are

\[ (2.4.6) \]

\[
\hat{\beta}_1 = (X'_{11}X_{11} + X'_{12}X_{12})^{-1}(X'_{11}Y_{11} + X'_{12}Y_{12}),
\]

\[
\hat{\beta}_2 = (X'_{21}L_2X_{21})^{-1}X'_{21}L_2Y_{22},
\]

\[
\hat{\Sigma}_1 = \frac{1}{n_1 - m_1}[(Y_{11} - X_{11}\hat{\beta}_1)'(Y_{11} - X_{11}\hat{\beta}_1) + (Y_{12} - X_{12}\hat{\beta}_1)'(Y_{12} - X_{12}\hat{\beta}_1)],
\]

\[
\hat{\Sigma}_2 = \frac{1}{n_2 - m_2 - 1}(Y_{22} - X_{22}\hat{\beta}_2)'(Y_{22} - X_{22}\hat{\beta}_2),
\]

where

\[
L_1 = I - Y_{12}'(Y_{12}'Y_{12})^{-1}Y_{12},
\]

\[
L_2 = I - X_2'(X_2'X_2)^{-1}X_2.'
\]

If we further assume that there exists a matrix \( G_{(g_2 \times g_1)} \), such that \( G_2 = G_{C_1} \), then the testing of \( H_0 \) is equivalent to the testing of the hypothesis \( H_0^*, \)

\[ (2.4.7) \]

\[
H_0^* : \bigcap_{j=1}^{2} H_{0j}^*, \quad \text{where} \quad H_{0j}^* : \phi_j = 0,
\]

with \( \phi_1 = C_1\beta_1 \),

\[
\phi_2 = C_2\beta_2^*.
\]

**Theorem 2.4.2.** With a good estimate of \( \sigma^2_{(j)} \), \( W_j^* \) is approximately distributed
as a $\chi^2$ variate with $g_j$ d.f., $j=1,2$, where

\begin{equation}
(2.4.8) \quad W_j^* = (\hat{\phi}_j - \phi_j)' [\hat{\sigma}_j^2 (j)] V_j^{-1} (\hat{\phi}_j - \phi_j), \quad j=1,2,
\end{equation}

\begin{align*}
V_1 &= C_1 [X'_{11} X_{11} + X'_{22} X_{22}]^{-1} V_1, \\
V_2 &= C_2 [X'_{22} L_1 X_{22}]^{-1} V_2.
\end{align*}

Setting $\phi_j = 0$, $j=1,2$, in (2.4.8) we have the following corollary.

**Corollary 2.4.3.** Under $H^*_0$, $j=1,2$, the statistic

\begin{equation}
(2.4.9) \quad W_j = \hat{\phi}_j' [\hat{\sigma}_j^2 (j)] V_j^{-1} \hat{\phi}_j
\end{equation}

is approximately distributed as a $\chi^2$ variate with $g_j$ d.f. We reject $H^*_0$ at level $\alpha_j$ if $W_j > \chi^2_{g_j, 1-\alpha_j}$, the upper $\alpha_j$ point of the $\chi^2$ distribution with $g_j$ d.f.

Under $H^*_0$, $W_1^*$ and $W_2^*$ are independently distributed. With the constraint $1-\alpha = (1-\alpha_1)(1-\alpha_2)$, the $(1-\alpha)$ SCS for $\phi_1$ and $\phi_2$ can be constructed from

\begin{equation}
(2.4.10) \quad W_j^* \leq \chi^2_{g_j, 1-\alpha_j}, \quad j=1,2.
\end{equation}

That is, a $(1-\alpha_1)$ confidence set for $\phi_1$ and a $(1-\alpha_2)$ confidence set for $\phi_2$.

The hypothesis $H^*_0$ will then be rejected if we either reject $H^*_01$ at level $\alpha_1$ or reject $H^*_02$ at level $\alpha_2$.

The above confidence procedure for $\phi_1$ and $\phi_2$ does not directly give the SCS for $\theta_1$ and $\theta_2$. From (2.4.4) and $C_2 = GC_1$, we have

\begin{equation}
(2.4.11) \quad \theta_1 = \phi_1,
\end{equation}

\begin{equation}
\theta_2 = \phi_2 + G\theta_1 \gamma_2.
\end{equation}

Thus the confidence set for $\theta_2$ depends on the values of $\theta_1$, $\phi_2$ and $\gamma_2$. 
Two different procedures are used to construct the confidence set for $\theta_2$.

Procedure I:

Given the value of $\theta_1$, $\theta_1^0$, then from (2.4.4) and (2.4.11) we have

\begin{equation}
\theta_2 = \phi_2 + G\theta_1^0 y_2 = (C_2 \theta_1^0) \begin{pmatrix}
\beta_2^*
\gamma_2
\end{pmatrix},
\end{equation}

\begin{equation}
\hat{\theta}_2 = (C_2 \theta_1^0) \begin{pmatrix}
\hat{\beta}_2^*
\hat{\gamma}_2
\end{pmatrix},
\end{equation}

\begin{equation}
\hat{V}(\hat{\theta}_2) = (C_2 \theta_1^0) \begin{pmatrix}
\hat{\beta}_2^*
\hat{\gamma}_2
\end{pmatrix} (C_2 \theta_1^0)',
\end{equation}

\begin{align*}
\hat{V}(\hat{\theta}_2) &= \sigma_2^2 \begin{pmatrix}
(X_2^1 L_1 X_2) & -(X_2^1 L_1 X_2)^{-1} X_2^1 Y_12 (Y_12^* Y_12)\begin{pmatrix}
1
\end{pmatrix}

-(Y_12^* Y_12)^{-1} Y_12^* X_2 (X_2^1 L_1 X_2)\begin{pmatrix}
1
\end{pmatrix}

-(Y_2^* L_2 Y_2)^{-1} Y_2^* X_2 (X_2^1 L_1 X_2)\begin{pmatrix}
1
\end{pmatrix}

\end{pmatrix} (C_2 \theta_1^0)

\end{align*}

\begin{align*}
\hat{V}(\hat{\theta}_2) &= \sigma_2^2 \begin{pmatrix}
(C_2 (X_2^1 L_1 X_2)^{-1} C_2^* - G\theta_1^0 (Y_12^* Y_12)^{-1} Y_12^* X_2 (X_2^1 L_1 X_2)\begin{pmatrix}
1
\end{pmatrix}

-C_2 (X_2^1 L_1 X_2)^{-1} X_2^1 Y_12 (Y_12^* Y_12)^{-1}(G\theta_1^0)',

-G\theta_1^0 (Y_12^* L_2 Y_12)^{-1} (G\theta_1^0)',
\end{pmatrix}

\end{align*}

Theorem 2.4.4. Given the value of $\theta_1$, $\theta_1^0$, the quadratic form

\begin{equation}
W_3^* = (\hat{\theta}_2 - \theta_2)' \hat{V}(\hat{\theta}_2)^{-1} (\hat{\theta}_2 - \theta_2)
\end{equation}

is approximately distributed as a $\chi^2$ variate with $g_2$ d.f.

Corollary 2.4.5. Given the value of $\theta_1$, $\theta_1^0$, a $(1-\alpha_2)$ confidence set for $\theta_2$ is

\begin{equation}
W_3^* \leq \chi^2_{g_2,1-\alpha_2},
\end{equation}

where $W_3^*$ is given by (2.4.14).

Theorem 2.4.6. With $(1-\alpha) = (1-\alpha_1)(1-\alpha_2)$, a $(1-\alpha)$ SCS for $\theta_1$ and $\theta_2$ can be
constructed by the following steps:

1) A \((1-\alpha_1)\) confidence set for \(\theta_1\) given by (2.4.10),

2) For a given value of \(\theta_1\), a \((1-\alpha_2)\) confidence set for \(\theta_2\) given by (2.4.15).

When \(g_2 = 1\), with the value of \(\theta_1\), \(\theta_0\), a \((1-\alpha_2)\) confidence set for \(\theta_2\) is

\[
|\hat{\theta}_2 - \theta_2| \leq \left\{ \chi^2_{1,1-\alpha_2} V(\theta_2) \right\}^{1/2},
\]

where \(V(\theta_2)\) is given by (2.4.13).

Procedure II:

**Theorem 2.4.7.** With a good estimate of \(\sigma^2(2)\),

\[
W_3 = \left( \frac{Y_2 - \gamma_2}{Y_{12}L_2Y_{12}} \right) \left( \frac{Y_2 - \gamma_2}{\sigma^2(2)} \right)\]

is approximately distributed as a \(\chi^2\) variate with one d.f.

**Corollary 2.4.8.** A \((1-\alpha_3)\) confidence set for \(\gamma_2\) is

\[
|\hat{\gamma}_2 - \gamma_2|^2 \leq \left\{ \chi^2_{1,1-\alpha_3} \sigma^2(2) Y_{12}L_2Y_{12}^{-1} \right\}^{1/2}.
\]

**Theorem 2.4.9.** With \((1-\alpha) = (1-\alpha_1)(1-\alpha_2), \alpha_3 = \frac{\alpha_2}{2}\), a \((1-\alpha)\) conservative SCS for \(\theta_1\) and \(\theta_2\) can be constructed from the following confidence sets:

1) A \((1-\alpha_1)\) confidence set for \(\theta_1\);

2) A \((1-\alpha_3)\) confidence set for \(\phi_2\); and

3) A \((1-\alpha_3)\) confidence set for \(\gamma_2\).

When \(g_2 = 1\), a \((1-\alpha_3)\) confidence set for \(\phi_2\) is

\[
|\hat{\phi}_2 - \phi_2| \leq \left\{ \chi^2_{1,1-\alpha_3} \sigma^2(2) Y_2 \right\}^{1/2} = \left\{ \frac{\sigma^2(2) Y_{12}L_2Y_{12}^{-1}}{C_2} \right\}^{1/2}.
\]
Together with a \((1-\alpha_3)\) confidence set for \(Y_2\) given by (2.4.18) and a given value of \(\theta_1, \theta_0^0\), the resulting confidence set for \(\theta_2 = \phi_2 + \Theta_0^0 Y_2\) is

\[
(2.4.20) \quad |\hat{\phi}_2 + \Theta_0^0 \hat{Y}_2 - \theta_2| \leq |(\hat{\phi}_2 - \phi_2)| + |\Theta_0^0 (\hat{Y}_2 - Y_2)|
\]

\[
= \{\sigma^2 (2) \chi^2_{1, 1-\alpha_3} C_2 (X_2^L L X_2)_{-1} C_2\}^{1/2}
+ \{\sigma^2 (2) \chi^2_{1, 1-\alpha_3} \Theta_0^0 (Y_{12} L Y_{12})_{-1} (\Theta_0^0)_{1}^{1/2}
\]

\[
= \{\sigma^2 (2) \chi^2_{1, 1-\alpha_3}\}^{1/2} \{C_2 (X_2^L L X_2)_{-1} C_2\}^{1/2} + \{\Theta_0^0 (Y_{12} Y_{12} L Y_{12})_{-1} (\Theta_0^0)_{1}^{1/2}\}.
\]

**Corollary 2.4.10.** With \((1-\alpha) = (1-\alpha_1) (1-\alpha_2)\), a \((1-\alpha)\) conservative SCS for \(\theta_1\) and \(\theta_2\) can be constructed by the following steps:

1) A \((1-\alpha_1)\) confidence set for \(\theta_1\) given by (2.4.10);

2) Given the value of \(\theta_1, \theta_0^0\), a confidence set for \(\theta_2\) given by (2.4.20).

Both procedures (I) and (II) give a \((1-\alpha_1)\) confidence set for \(\theta_1\). However, for every given value of \(\theta_1, \theta_0^0\), the confidence procedure (II) ignores the correlation between \(\hat{\phi}_2\) and \(\hat{Y}_2\). This results in a larger confidence coefficient (\(\geq 1-\alpha_2\)) for the resulting confidence set of \(\theta_2\).

When \(g_2 = 1\), for every given value of \(\theta_1, \theta_0^0\), the length of the resulting confidence set for \(\theta_2\) from the confidence procedure (I) is

\[
(2.4.21) \quad L_1 = 2\{\chi^2_{1, 1-\alpha_2} \chi^2(\hat{\theta}_2)\}^{1/2}
\]

\[
= 2\{\sigma^2 (2) \chi^2_{1, 1-\alpha_2}\}^{1/2} \{C_2 (X_2^L L X_2)_{-1} C_2\}^{1/2} - 2\Theta_0^0 (Y_{12} Y_{12})_{-1} (\Theta_0^0)_{1}^{1/2} \{X_2^L L X_2\}_{-1} C_2
+ \Theta_0^0 (Y_{12} Y_{12} L Y_{12})_{-1} (\Theta_0^0)_{1}^{1/2}.
\]

Similarly, the length of the resulting confidence set for \(\theta_2\) from the confidence procedure (II) is
For a given value of $\theta_1^0$, $\theta_1^0$, the correlation between $\hat{\phi}_2$ and $G\theta_1^0$ is

\[
R = \frac{-G\theta_1^0(Y_1^{12}Y_1^{12})^{-1}Y_1^{12}X_2(X_2^{12}X_2^{12})^{-1}C_2^2}{\{[C_2(X_2^{12}L_2X_2^{12})^{-1}C_2^2][G\theta_1^0(Y_1^{12}L_2Y_1^{12})^{-1}G\theta_1^0]\}^{1/2}}.
\]

The comparison between $L_1$ and $L_2$ depends on the value of $R$.

1) For $R < 0$.

\[
L_1 = 2\sigma^2_{(2)}X_1^{2,1}\cdot\left\{\frac{X_2^{12}L_2X_2^{12}}{X_1^{2,1,1}}\right\}C_2^2 - 2G\theta_1^0(Y_1^{12}Y_1^{12})^{-1}Y_1^{12}X_2(X_2^{12}L_2X_2^{12})^{-1}C_2^2
\]

\[
+ G\theta_1^0(Y_1^{12}L_2Y_1^{12})^{-1}G\theta_1^0)^{1/2}
\]

\[
X_1^{2,1,1} \leq \frac{X_{1,1-2/2}}{X_1^{1,1-2/2}} \leq \frac{X_2^{12,1,1}}{X_2^{12,1-2/2}} L_2.
\]

2) For $R = 0$ ($\theta_1^0 = 0$).

\[
L_1 = 2\sigma^2_{(2)}X_1^{2,1}\cdot\left\{\frac{X_2^{12}L_2X_2^{12}}{X_1^{2,1,1}}\right\}C_2^2
\]

\[
= \left\{\frac{X_2^{12,1-2/2}}{X_1^{2,1,1}}\right\} L_2.
\]

3) For $R > 0$.

\[
L_1 = 2\sigma^2_{(2)}X_1^{2,1}\cdot\left\{\frac{X_2^{12}L_2X_2^{12}}{X_1^{2,1,1}}\right\}C_2^2 - 2G\theta_1^0(Y_1^{12}Y_1^{12})^{-1}Y_1^{12}X_2(X_2^{12}L_2X_2^{12})^{-1}C_2^2
\]

\[
+ G\theta_1^0(Y_1^{12}L_2Y_1^{12})^{-1}G\theta_1^0)^{1/2}
\]
Thus we have

\[ (2.4.24) \quad \frac{L_1}{L_2} \leq \left\{ \frac{X^2_{1,1-\alpha_2}}{X^2_{1,1-\alpha_2/2}} \right\}^{1/2} < 1. \]

That is, given the value for \( \theta_1 \), the confidence procedure (I) gives a smaller confidence set for \( \theta_2 \) than the confidence set from procedure (II).

2.4.2 Optimum confidence procedure

With a good estimate of \( \sigma^2_{(j)} \), \( j=1,\ldots,p \), a \( (1-\alpha_j) \) confidence set for \( \phi_j \) can be constructed from \( W^*_j \leq X^2_{g_j,1-\alpha_j} \), where \( W^*_j \) is given by (2.4.11).

Under \( H_0 \), \( W^*_j \)'s are independently distributed. A \( (1-\alpha) \) SCS for \( \phi_j \)'s can be constructed from

\[ (2.4.25) \quad W^*_j \leq X^2_{g_j,1-\alpha_j}, \quad j=1,2,\ldots,p, \]

with \( 1-\alpha = \Pi_{j=1}^p (1-\alpha_j) \).

The above confidence procedure is not invariant under different selections of \( \alpha_j \)'s satisfying the condition \( \Pi_{j=1}^p (1-\alpha_j) = 1-\alpha \). Letting \( y_j \) be the upper \( \alpha_j \) point of the \( \chi^2 \) distribution with \( g_j \) d.f., then the SCS for \( \phi_j \), \( j=1,\ldots,p \), can be optimized by either criterion:
1) Minimizing \( \max \frac{y_j}{g_j} \), the largest weighted length of the \( p \) quadratic forms, or

2) Minimizing \( \sum_{j=1}^{p} \frac{y_j}{g_j} \), the sum of weighted lengths of the \( p \) quadratic forms; with both being subject to the constraint \( \prod_{j=1}^{p} (1-\alpha_j) = 1-\alpha \),

with \( p[\chi^2_{g_j} \leq y_j] = 1-\alpha_j \).

\[ (2.4.26) \text{Criterion 1.} \quad \text{Minimizing } \max_{1 \leq j \leq p} \frac{y_j}{g_j} \text{, where } y_j = \chi^2_{g_j, 1-\alpha_j} \text{, subject to } \prod_{j=1}^{p} (1-\alpha_j) = 1-\alpha. \]

Let

\[ (2.4.27) \quad Z = \max_{1 \leq j \leq p} \frac{y_j}{g_j}, \]

where

\( y_j \) is a \( \chi^2 \) random variable with \( g_j \) d.f., \( j = 1, \ldots , p \),

\( y_j \)'s are independently distributed.

Then the distribution function of \( Z \) is

\[ (2.4.28) \quad F_Z(z) = p[Z \leq z] = p[\max_{1 \leq j \leq p} \frac{y_j}{g_j} \leq z] = p[y_j \leq zg_j, j = 1, \ldots , p] \]

\[ = \prod_{j=1}^{p} p[y_j \leq zg_j]. \]

Since \( Z \) has a continuous distribution, a unique value of \( Z, z_0 \), can be found such that \( F_Z(z_0) = 1-\alpha \).

Theorem 2.4.11. Under Criterion 1, the optimal procedure gives

\[ (2.4.29) \quad y_j = z_0 g_j, \quad \alpha_j = 1 - p[y_j \leq y_j] = 1 - p[y_j \leq z_0 g_j], \quad j = 1, \ldots , p, \]

where \( z_0 \) is the upper \( \alpha \) point of the random variable \( Z \) given by \( (2.4.27) \).
When \( g_i < g_j \), we have \( 1 - \alpha_i = F_{Y_i}(z_0 g_i) < F_{Y_j}(z_0 g_j) = 1 - \alpha_j \), the variate with a larger d.f. will have a larger confidence coefficient.

For the special case \( g_i = g_j \), \( F_{Y_i}(z_0 g) = F_{Y_j}(z_0 g) \), \( \alpha_i = \alpha_j \). That is, two variates with the same d.f. will have the same confidence coefficient.

Extending this special case to all \( p \) variates, we have

**Corollary 2.4.12.** When \( g_1 = g_2 = \ldots = g_p = g \), the optimum procedure gives

\[
\begin{align*}
\alpha_j &= \alpha^*, \\
y_j &= \chi^2_{g_j, 1-\alpha^*}, \quad j=1,\ldots,p.
\end{align*}
\]

where

\[
\alpha^* = 1 - (1-\alpha)^{1/p}.
\]

\[
\begin{align*}
0^* &= \chi^2_{g_j, 1-\alpha^*}, \\
\prod_{j=1}^{p} (1-\alpha_j) &= 1-\alpha.
\end{align*}
\]

**Criterion 2.** Minimizing \( \min \sum_{j=1}^{p} \frac{y_j}{g_j} \), where \( y_j = \chi^2_{g_j, 1-\alpha^*} \), subject to \( \prod_{j=1}^{p} (1-\alpha_j) = 1-\alpha \).

Letting \( Y_j \) be a \( \chi^2 \) random variable with \( g_j \) d.f., then the distribution function of \( Y_j \) is

\[
F_j(y) = 2 \Phi (\sqrt{y}) - 1 - \sum_{k=0}^{(g_j/2)-3/2} e^{-ky} \frac{(by)^k}{\Gamma(k+3/2)} \quad \text{for odd } g_j,
\]

\[
F_j(y) = 1 - \sum_{k=0}^{(g_j/2)-1} e^{-ky} \frac{(by)^k}{\Gamma(k+1)} \quad \text{for even } g_j,
\]

and the density function of \( Y_j \) is

\[
f_j(y) = \frac{1}{2 g_j} \frac{y^{g_j-2} e^{-y}}{\Gamma(g_j)}.
\]

By introducing a Lagrangian multiplier \( \lambda \) for the constraint \( \prod_{j=1}^{p} (1-\alpha_j) = 1-\alpha \),
criterion 2 is equivalent to minimizing

\[ D = D(\lambda, Y_1, \ldots, Y_p) = \frac{P}{J=1} \frac{Y_j}{g_j} - \lambda [\prod_{j=1}^{P} F_j(y_j) - (1-\alpha)]. \]

The optimum procedure can be obtained by solving the following system of first partial equations:

\[ \frac{\partial D}{\partial \lambda} : \prod_{j=1}^{P} F_j(y_j) - (1-\alpha) = 0, \quad \text{and} \]

\[ \frac{\partial D}{\partial y_k} : \lambda \frac{1}{g_k} - \lambda \left[ \frac{1-\alpha}{F_k(y_k)} \right] f_k(y_k) = 0, \quad k=1, \ldots, P. \]

Replacing \( \prod_{j=1}^{P} F_j(y_j) \) by \( (1-\alpha) \) in (2.4.37), we have the following \( P \) equations:

\[ \frac{1}{g_k} - \lambda \frac{1-\alpha}{F_k(y_k)} f_k(y_k) = 0, \quad k=1, \ldots, P. \]

From the derivative

\[ \frac{\partial F_k(y)}{\partial y} = \frac{1}{F_k(y)} \left[ f_k(y) \frac{\partial F_k(y)}{\partial y} - F_k(y) \frac{\partial f_k(y)}{\partial y} \right] 
\]

we know that \( F_k(y) \) is a monotonic increasing function of \( y \) under practical consideration (i.e. \( y > g_{k-2} \)).

From (2.4.5) we have

\[ \frac{1}{g_k} \frac{F_k(y_k)}{f_k(y_k)} = \lambda(1-\alpha), \quad \text{a constant for} \ k=1, \ldots, P. \]

Thus, for the special case \( g_i = g_j = g \), we have the same confidence coefficient.

\[ \frac{F_i(y_i)}{f_i(y_i)} = \frac{F_j(y_j)}{f_j(y_j)}, \]
\[ y_i = y_j, \]
\[ \alpha_i = \alpha_j. \]

Extending the above special case to all \( p \) variates, we have the following lemma:

**Lemma 2.4.13.** When \( g_1 = g_2 = \ldots = g_p = g \), the optimum procedure corresponding to criterion (2) gives

\[
\begin{align*}
\alpha_j &= \alpha^*, \\
y_i &= \chi_g^2,1-\alpha^*, \quad j=1,\ldots,p, \text{ where} \\
\alpha^* &= 1 - (1-\alpha)^{1/p}.
\end{align*}
\]

This is the same result as from optimum procedure (1).

Now we consider the case \( g_1 < g_j \). When \( p = 2 \), the system of equations (2.4.36), (2.4.37) reduces to

\[
\begin{align*}
\frac{F_1(y_1)}{g_1} \frac{F_2(y_2)}{g_2} &= 1-\alpha, \\
\frac{1}{g_1} \frac{F_1(y_1)}{f_1(y_1)} &= \lambda(1-\alpha), \text{ and} \\
\frac{1}{g_2} \frac{F_2(y_2)}{f_2(y_2)} &= \lambda(1-\alpha).
\end{align*}
\]

With the constraint \( F_1(y_1) F_2(y_2) = 1-\alpha \), \( y_2 \) is a decreasing function of \( y_1 \).

Defining the function

\[
(2.4.44) \quad h(y_1) = \frac{g_2}{g_1} \frac{F_1(y_1)}{f_1(y_1)} \frac{f_2(y_2)}{F_2(y_2)} - 1, \text{ where } F_1(y_1) F_2(y_2) = 1-\alpha,
\]

and applying (2.4.26) we have the following theorem:

**Theorem 2.4.14.** The function \( h(y_1) \) given by (2.4.44) is a monotonic increasing function of \( y_1 \). The optimal solution corresponding to criterion (2)
(2.4.32) can be obtained by solving the equation

\[ h(y_1) = 0. \]

If \( g_1 < g_2 \) and the function \( h(x^2) > 0 \), we have the optimal solution \( y_1^* < x^2 \).

That is, \( (1-\alpha)^{\frac{1}{2}} < (1-\alpha)^{\frac{3}{2}} \), the variate with a larger d.f. will have a larger confidence coefficient.

**Theorem 2.4.15.** When \( p=2, 1 \leq g_1 < g_2 \leq 4 \), the optimum procedure corresponding to criterion (2) results in a larger confidence coefficient for the variate with larger d.f.

**Proof:** Letting \( \alpha^0 = 1 - (1-\alpha)^{\frac{1}{2}} \), and

\[ y_j^0 = \frac{x^2}{g_j,1-\alpha^0}, \quad j=1,2, \quad \text{then} \]

\[ h(y_1^0) = \frac{g_2}{g_1} \frac{F_1(y_1^0)}{F_2(y_2^0)} \frac{f_2(y_2^0)}{f_1(y_1^0)} - 1 = \frac{g_2}{g_1} \frac{f_2(y_2^0)}{f_1(y_1^0)} - 1. \]

If we can show that \( h(y_1^0) > 0 \) for \( g_1 = 1,2,3, \ g_2 = g_1 + 1 \), then from

\[ \frac{2f_2(x_2^2,1-\alpha^0)}{f_1(x_2^2,1-\alpha^0)} > 1, \quad \frac{3f_2(x_3^2,1-\alpha^0)}{2f_1(x_2^2,1-\alpha^0)} > 1, \quad \frac{4f_2(x_4^2,1-\alpha^0)}{3f_1(x_3^2,1-\alpha^0)} > 1, \]

we have

\[ \frac{3f_2(x_3^2,1-\alpha^0)}{f_1(x_2^2,1-\alpha^0)} > 1, \quad \frac{4f_2(x_4^2,1-\alpha^0)}{2f_1(x_2^2,1-\alpha^0)} > 1, \quad \frac{4f_2(x_4^2,1-\alpha^0)}{f_1(x_2^2,1-\alpha^0)} > 1. \]

That is, \( \frac{g_2}{g_1} \frac{f_2(y_2^0)}{f_1(y_1^0)} - 1 > 0 \) for \( 1 \leq g_1 < g_2 \leq 4 \).

**Case 1)** \( g_1 = 1, \ g_2 = 2. \)

From \( F_1(y_1^0) = F_2(y_2^0) \) we have
\[ (2.4.48) \quad 2 \Phi ((y_1^0)^{1/2}) - 1 = 1 - e^{-y_2^0}, \]

\[ e^{y_2^0/2} = 2[1 - \Phi((y_1^0)^{1/2})] = 2 \int_{(y_1^0)^{1/2}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{y_1^0}^{\infty} y^{-1/2} e^{-ly} \, dy. \]

Then

\[ (2.4.49) \quad h(y_1^0) = \frac{2f_2(y_1^0)}{f_1(y_1^0)} - 1 = \frac{1}{2\Gamma(1)} e^{-y_2^0} = \frac{1}{2\Gamma(1)} e^{-y_1^0} - 1 \]

\[ = \frac{\sqrt{2\pi}}{\Gamma(1)} e^{-y_2^0} - 1 = \frac{\int_{y_1^0}^{\infty} y^{-1/2} e^{-ly} \, dy}{(y_1^0)^{-1/2} e^{-ly_1^0}} - 1. \]

Using integration by parts,

\[ (2.4.50) \quad \int_{y_1^0}^{\infty} y^{-1/2} e^{-ly} \, dy = 2(y_1^0)^{-1/2} e^{-ly_1^0} - \int_{y_1^0}^{\infty} y^{-3/2} e^{-ly} \, dy \]

\[ = 2(y_1^0)^{-1/2} e^{-ly_1^0} - 2(y_1^0)^{-1/2} e^{-ly_1^0} + 3 \int_{y_1}^{\infty} y^{-3/2} e^{-ly} \, dy \]

\[ > 2(y_1^0)^{-1/2} e^{-ly_1^0}[1 - (y_1^0)^{-1}]. \]

Thus

\[ (2.4.51) \quad h(y_1^0) = \frac{\int_{y_1^0}^{\infty} y^{-1/2} e^{-ly} \, dy}{(y_1^0)^{-1/2} e^{-ly_1^0}} - 1 \]

\[ > \frac{2(y_1^0)^{-1/2} e^{-ly_1^0}[1 - (y_1^0)^{-1}]}{(y_1^0)^{-1/2} e^{-ly_1^0}} - 1 \]

\[ = 2[1 - (y_1^0)^{-1}] - 1 \]

\[ = 2[1 - \Phi((y_1^0)^{1/2})] - 1 \]
Case 2) \( g_1 = 2, g_2 = 3 \).

From \( F_1(y_1^0) = F_2(y_2^0) \):

\[
1 - e^{-ky_1^0} = 2 \Phi((y_2^0)^{\frac{1}{2}}) - 1 - \frac{1}{\Gamma(3/2)} e^{-ky_2^0} (ky_2^0)^{\frac{1}{2}}
\]

\[
e^{-ky_1^0} = 2[1 - \Phi((y_2^0)^{\frac{1}{2}})] + \frac{2}{\sqrt{2\pi}} (y_2^0)^{\frac{1}{2}} e^{-ky_2^0}.
\]

Then

\[
(2.4.53) \quad h(y_1^0) = \frac{3f_2(y_2^0)}{2f_1(y_1^0)} = \frac{3}{2} \frac{1}{\Gamma(3/2)} \frac{(y_1^0)^{\frac{3}{2}} e^{-ky_2^0}}{1 + \Phi((y_2^0)^{\frac{1}{2}}) - 1 - \frac{1}{\Gamma(3/2)} e^{-ky_2^0} (ky_2^0)^{\frac{1}{2}}}
\]

\[
= \frac{3}{\sqrt{2\pi}} \frac{(y_2^0)^{\frac{1}{2}} e^{-ky_2^0}}{1 + \Phi((y_2^0)^{\frac{1}{2}}) - 1 - \frac{1}{\Gamma(3/2)} e^{-ky_2^0} (ky_2^0)^{\frac{1}{2}}}
\]

Using a change of variable and integration by parts, we have

\[
(2.4.54) \quad 1 - \Phi((y_2^0)^{\frac{1}{2}}) < \frac{(y_2^0)^{\frac{1}{2}} e^{-ky_2^0}}{\sqrt{2\pi}}.
\]

Thus,

\[
(2.4.55) \quad h(y_1^0) = \frac{3}{2} \frac{(y_2^0)^{\frac{1}{2}} e^{-ky_2^0}}{\sqrt{2\pi} \{1 - \Phi((y_2^0)^{\frac{1}{2}}) - (y_2^0)^{\frac{1}{2}} e^{-ky_2^0}\}} - 1
\]

\[
> \frac{3}{2} \frac{(y_2^0)^{\frac{1}{2}} e^{-ky_2^0}}{(y_2^0)^{\frac{1}{2}} e^{-ky_2^0} [1 + (y_2^0)^{-1}]} - 1
\]
\[
\frac{3}{2} \frac{1}{1 + (y_2^0)^{\frac{3}{2}}} - 1
\]

> 0 under practical consideration \((y_2^0 > 2)\).

Case 3) \(g_1 = 3, g_4 = 4\).

From \(F_1(y_1^0) = F_2(y_2^0)\), we have

\[
(2.4.56) \quad 2\Phi((y_1^0)^{\frac{1}{2}}) - 1 - \frac{1}{\Gamma(3/2)} e^{-by_1^0 - \frac{b}{2}y_1^0} = 1 - e^{-by_2^0 (1 + by_2^0)},
\]

Then

\[
(2.4.57) \quad h(y_1^0) = \frac{4 \frac{1}{2^2 \Gamma(2)} y_2^0 e^{-by_2^0}}{3 \frac{1}{2^3 \Gamma(3/2)} (y_1^0)^{\frac{1}{2}} e^{-by_1^0}} - 1 = \frac{\sqrt{2\pi} \ y_2^0 e^{-by_2^0}}{3 \ (y_1^0)^{\frac{1}{2}} e^{-by_1^0}} - 1
\]

\[
= \frac{1}{3} \frac{2}{(1/2 + 1/y_2^0)} \{\sqrt{2\pi} \ [1 - \Phi((y_1^0)^{\frac{1}{2}})] + (y_1^0)^{\frac{1}{2}} e^{-by_1^0}\} - 1
\]

\[
= \frac{4y_2^0}{3y_2^0 + 6} \{\sqrt{2\pi} \ 1 - \Phi((y_1^0)^{\frac{1}{2}})\} + 1 \} - 1
\]

> 0 under practical consideration \((y_2^0 > 6)\).

Thus we have
The optimal solution $Y^*_1 < \chi^2_{g_1,(1-\alpha)^2}$

That is, the variate with a larger d.f. will have a larger confidence coefficient.

Extending the above result to the general case $p > 2$, we have the following corollary:

**Corollary 2.4.16.** For $1 \leq g_j \leq 4$, $j=1, \ldots, p$. The optimum procedure corresponding to criterion (2) gives a larger confidence coefficient to the variate with a larger d.f.

The usual induction method is not applicable to larger values of $g_j$. The exact optimal values of $\{\alpha_1, \ldots, \alpha_p\}$ cannot be easily obtained from the above derivations. A numerical method of trial and error can be used to solve the system of equations (2.4.36) and (2.4.37).

### 2.5 Examples

In this section, three examples of SGCM, GGCM, and HM models are presented to illustrate the procedures derived in Sections 2.2 - 2.4.

The data set for the three models is the RATS data set collected by Gaynor and Grizzle (1980). The data was generated to assess the effect of neo-natal lead exposure in Sprague-Dawley rats on the growth of the myelin accumulation in the brain.

Four different treatments were used in the data:
1) Control  
2) Lead treated  
3) Starved control  
4) Super starved

The data set is comprised of 5 replications of the same experiment. For each replication 16 rat pups from 4 litters, all born on the same day, were randomized into 4 groups of 4 animals. Each group was then randomly allocated back to one of the four mothers for nursing. Each of the four treatments were then randomly assigned to one of the 4 maternal litters, so that the rat pups feeding from a particular mother would receive the same treatment. Four groups of one rat within a maternal litter were randomly chosen to be sacrificed at each of 30, 60, 90, and 120 days of age, respectively. Therefore, within a given replication of the experiment, one maternal litter received one of the four treatments, and then measurements were taken from one of the animals in that litter at each of 30, 60, 90, and 120 days of age. The measurement made on each animal was the natural logarithm of myelin content in the brain (mg/brain).

Assuming
1) the measurements from different litters are uncorrelated,  
2) the measurements from the same litter may be correlated, and  
3) there is no litter effect,
then we have the following standard MGLM model:

\[ E(Y_0) = X\beta, \]
\[ V(Y_0) = I_{20} \otimes \Sigma_0, \]

where \( Y_0 {20 \times 4} \), each row of \( Y_0 \) has 4 measurements from a litter,
The hypotheses of interest are the difference between the lead treated group and control group \((\beta_2 - \beta_1)\) and the difference between the lead treated group and starved control group \((\beta_2 - \beta_3)\).

2.5.1 An SGCM model

Using the RATS data, the SGCM model is

\[
(2.5.2) \quad E(Y_0) = X_0 \beta P, \\
V(Y_0) = I_{20} \otimes \Sigma_0,
\]

where  
\[
X = (I_4, I_4, I_4, I_4, I_4)', \\
P = \begin{bmatrix}
1 & 1 & 1 & 1 \\
-3 & -1 & 1 & 1 \\
1 & -1 & -1 & 1
\end{bmatrix},
\]

The hypothesis of interest is \(H_0: \theta = C \beta = 0\), where \(C = \begin{bmatrix}
-1 & 1 & 0 & 0 \\
0 & 1 & -1 & 0
\end{bmatrix}\).

Letting  
\[
(2.5.3) \quad H_1 = P' = \begin{bmatrix}
1 & -3 & 1 \\
1 & -1 & -1 \\
1 & 1 & -1 \\
1 & 3 & 1
\end{bmatrix}, \\
H_2 = \begin{bmatrix}
1 \\
-3 \\
3 \\
-1
\end{bmatrix},
\]

the resulting conditional model \((2.2.1)\) is

\[
(2.5.4) \quad E(Y|Z) = X_\gamma + Z_\Gamma, \\
V(Y|Z) = I \otimes \Sigma,
\]
where  
\[ \eta = \beta P \phi_1 = \beta \begin{pmatrix} 4 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \]

The LSE's of \( \eta \), \( \Gamma \) are

\[ \widehat{\eta} = \begin{pmatrix} 9.73541 & 3.29433 & -0.438172 \\ 8.21822 & 3.72421 & -0.326930 \\ 8.99111 & 3.56806 & -0.476797 \\ 7.66623 & 5.29237 & -0.860431 \end{pmatrix}, \]

\[ \widehat{\Gamma} = (0.085616 -0.0226278 0.0658045). \]

We then have

\[ \widehat{\theta} = \begin{pmatrix} -0.379298 & 0.0214941 & 0.0278105 \\ -0.193222 & 0.00780753 & 0.0374669 \end{pmatrix}, \]

\[ (CRC')^{-1} = \begin{pmatrix} 3.33328 & -1.66646 \\ -1.66646 & 3.33258 \end{pmatrix}, \]

\[ S_E = \begin{pmatrix} 0.0949042 & -0.0274339 & 0.0430596 \\ 0.0318341 & -0.0285718 \\ (sym.) & 0.0749477 \end{pmatrix}, \]

\[ S_H^{-1} = (\widehat{\theta} - \theta)'(CRC')^{-1}(\widehat{\theta} - \theta). \]

The confidence set can then be constructed from \( S_H^{-1} S_E \).

Using Roy's largest root procedure, the \((1-\alpha)\) confidence set for \( \theta \) is given by

\[ \text{Ch}_1(S_H^{-1}S_E) \leq \lambda_{\alpha; s, m, n}, \]

where \( \lambda_{\alpha; s, m, n} \) is the upper \( \alpha \) point of the distribution of the largest root of \( S_H^{-1}S_E \) with the parameters...
(2.5.8) \[ s = 2, \]
\[ m = \frac{|2-3| - 1}{2} = 0, \]
\[ n = \frac{20-5-3-1}{2} = 5.5. \]

The exact SCS for \( \theta \) can then be constructed from

\[ \frac{a' S^* a}{a' S_E a} \leq \lambda_{\alpha; s, m, n} \quad \text{for all vectors} \ a, \]

or from

\[ |b'(\theta-\theta)a| \leq \{\lambda_{\alpha; s, m, n} [a' S_E a][b'(CRC')b]\}^{\frac{1}{2}} \]

for all vectors \( a, b \).

From Heck's chart, the upper 5% point of the distribution of the largest

root of \( S_H (S_H + S_E)^{-1} \) under \( H_0 \) is

\[ \lambda_{\alpha; s, m, n} = 0.537. \]

Thus \( \lambda_{\alpha; s, m, n} = \frac{\lambda_{\alpha; s, m, n}}{1-\lambda_{\alpha; s, m, n}} = 1.160. \)

When \( \theta = 0, \)

\[ S_H = \begin{bmatrix} 0.359703 & -0.0203465 & -0.02664980 \\ 0.00118379 & 0.00126350 \\ (sym.) & 0.00378339 \end{bmatrix}, \quad \text{and} \]

\[ \text{Ch}_1(S_H S_E^{-1}) = \text{Ch}_1(S_H S_E^{-1}) = 5.7242 > 1.160. \]

Thus we reject the hypothesis \( H_0 \) by Roy's largest root procedure.

2.5.2 A GGCM model

The data for this model were obtained from the RATS data with the
4th response for the first 12 observations and the 3rd response for the other 8 observations deleted. The resulting GGCM model is given by

\[(2.5.14)\]

\[
E(Y_{0j}) = X_j \beta P B_j,
\]

\[
V(Y_{0j}) = I_{n_j} \otimes B_j^T \Sigma_0 B_j, \quad j=1,2,
\]

where

\[n_1 = 12, \quad n_2 = 8,
\]

\[
X_1 = (I_4, I_4, I_4)^T,
\]

\[
X_2 = (I_4, I_4)^T,
\]

\[
P = \begin{bmatrix}
1 & 1 & 1 & 1 \\
-3 & -1 & 1 & 3 \\
1 & -1 & -1 & 1 
\end{bmatrix},
\]

\[
B_1 = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1 
\end{bmatrix},
\]

\[
B_2 = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1 
\end{bmatrix}.
\]

The hypothesis of interest is \(H_0: C_3 = 0, C = \begin{bmatrix}
-1 & 1 & 0 & 0 \\
0 & 1 & -1 & 1 
\end{bmatrix}.\)

Since both matrices \(PB_1\) and \(PB_2\) are square matrices, no covariates will be used in the model. With

\[(2.5.15)\]

\[
H_{11} = (PB_1)^{-1} = \frac{1}{2} \begin{bmatrix}
1 & 0 & 1 \\
-1 & -1 & -2 \\
2 & 1 & 1 
\end{bmatrix},
\]

\[
H_{21} = (PB_2)^{-1} = \frac{1}{6} \begin{bmatrix}
1 & -1 & 2 \\
3 & 0 & -3 \\
2 & 1 & 1 
\end{bmatrix}.
\]
the weighted least squares model in (2.3.9) gives

(2.5.16)

\[ E(Y^V) = \Lambda \beta^V \]

\[ V(Y^V) = \Omega, \]

where

\[ \Lambda = \begin{pmatrix} X_1 \otimes I_3 \\
X_2 \otimes I_3 \end{pmatrix}, \]

\[ \Omega = \begin{pmatrix} I_{12} \otimes \Sigma_1 \\
I_8 \otimes \Sigma_2 \end{pmatrix}. \]

An unbiased estimate of \( \Sigma_0 \) can be constructed from (2.3.11). This is

(2.5.17)

\[ \hat{\Sigma}_0 = 0.01 \begin{pmatrix} 6.077180 & 0.125344 & 1.27238 & -1.684370 \\
0.125344 & 2.166910 & -1.00445 & -0.460083 \\
& & 1.14083 & \\
& & & 0.267248 \end{pmatrix} \]

We then have

(2.5.18)

\[ \hat{\Sigma}_1 = 0.01 \begin{pmatrix} 5.41601 & 2.15224 & 5.28900 \\
& 1.32916 & 2.40876 \\
& & 5.48671 \end{pmatrix}, \]

\[ \hat{\Sigma}_2 = 0.01 \begin{pmatrix} 4.20609 & -0.155961 & -0.374414 \\
& 0.269810 & -0.328200 \\
& & 1.07214 \end{pmatrix}, \]

\[ \hat{\beta}^V = \left\{ \left\{ \sum_{j=1}^{2} (X_j^\prime X_j) \otimes \hat{\Sigma}_j^{-1} \right\}^{-1} \left\{ \sum_{j=1}^{2} (X_j^\prime \otimes \hat{\Sigma}_j^{-1}) Y^V_j \right\} \right\} \]
\[
\hat{H}^\beta = \begin{bmatrix}
-0.346787 \\
-0.205938 \\
0.0394105 \\
0.00398122 \\
0.0192067 \\
-0.000045205
\end{bmatrix}
\]

\[
H^V (\hat{\beta}^V) H^* = \left\{ \frac{2}{j=1} (X_j^T X_j) \otimes \hat{\Sigma}_j^{-1} \right\}^{-1}
\]

\[
\begin{pmatrix}
1.00849 & 0.504244 & -0.105946 & -0.52973 & -0.186797 & -0.093399 \\
1.00849 & -0.52973 & -1.05946 & -0.093399 & -0.186797 \\
& & 0.792057 & 0.396029 & -0.371661 & -0.18583 \\
& & & 0.792057 & -0.18583 & -0.371661 \\
& & & & 0.171074 & 0.0855371 \\
& & & & & 0.171074
\end{pmatrix}
\]

At \( \alpha = 0.05 \), a \((1-\alpha)\) confidence set can then be constructed from (2.3.19).

That is,

\[
W_n^* = (H^V \hat{\beta}^V - H^V \beta^V)' [H^V (\hat{\beta}^V) H^*]^{-1} (H^V \hat{\beta}^V - H^V \beta^V) \leq \chi^2_{6, .95} = 12.592.
\]

When \( H^V \beta^V = 0 \), we have
Thus we reject $H_0$ at level 0.05.

2.5.3 An HM model with $p=2$

The data for this model were obtained from the first two responses (30 days, 60 days) of the RATS data with the first 4 observations of the second response (60 days) deleted.

As per the notation used in Section 2.4.2, the model is given by

\[
E(Y_{11}(4 \times 1)) = X_{11} \beta_1, \\
E(Y_{12}(16 \times 1)) = X_{2} \beta_1, \\
E(Y_{22}(16 \times 1)) = X_{2} \beta_2.
\]

\[
V(Y_{11}) = I_4 \otimes \sigma_{11}, \\
V(Y_{12}, Y_{22}) = I_{16} \otimes \Sigma_2,
\]

where \( X_{11} = I_4, \ X_2 = (I_4, I_4, I_4, I_4)' \).

The hypothesis of interest is

\[
H_0: \sum_{j=1}^{2} H_{0j}, \text{ with } H_{0j}: \theta_j = C_j \beta_j = 0,
\]

where

\[
C_1 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}, \quad C_2 = \begin{pmatrix} -1 & 2 & -1 & 0 \end{pmatrix}.
\]

With the models (2.4.4) and (2.4.5), we have the following LSE's:
\begin{align*}
\hat{\beta}_1 &= \begin{pmatrix} 1.8084 \\ 1.39318 \\ 1.56946 \\ 0.825473 \end{pmatrix}, \\
\hat{\sigma}^2_{(1)} &= 0.0607718, \\
\hat{\beta}_2 &= \begin{pmatrix} 2.2028 \\ 1.02011 \\ 2.08189 \\ 1.89668 \end{pmatrix}, \\
\hat{y}_2 &= 0.0832716, \\
\hat{\sigma}^2_{(2)} &= 0.019233.
\end{align*}

At \( \alpha = 0.05 \), an optimal \((1-\alpha)\) confidence set for \( \phi_1 \) and \( \phi_2 \) corresponding to criterion (1) given by (2.4.26) is a \((1-\alpha_j)\) confidence set for \( \phi_j \), \( j=1,2 \), where

\begin{equation}
\alpha_1 = 0.0131213, \\
\alpha_2 = 0.0373690.
\end{equation}

Assuming \( \hat{\sigma}^2_{(j)} \) is a good estimator for \( \sigma^2_{(j)} \), \( j=1,2 \), then an optimal \((1-\alpha)\) confidence set for \((\phi_1, \phi_2)\) can be constructed from a \((1-\alpha_1)\) confidence set for \( \phi_1 \) and a \((1-\alpha_2)\) confidence set for \( \phi_2 \). That is, from (2.4.10), we have

\begin{equation}
W^* = \left[ \begin{pmatrix} -0.415218 \\ -0.176277 \end{pmatrix} - \phi_1 \right]' \left( 0.4 \hspace{1cm} 0.2 \right) \left( 0.4 \hspace{1cm} 0.2 \right)^{-1} \left[ \begin{pmatrix} -0.415218 \\ -0.176277 \end{pmatrix} - \phi_1 \right] \leq \chi^2_{2,1-\alpha_1},
\end{equation}

\begin{equation}
\begin{pmatrix} .415218 \\ .176277 \end{pmatrix} + \phi_1 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} .415218 \\ .176277 \end{pmatrix} + \phi_1 \leq .316027.
\end{equation}
(2.5.24) \[ W_2^* = [-.644472 - \phi_2]((.019233) (2.09948))^{-1} [-.644472 - \phi_2] \]
\[ \leq \chi_{1,1-\alpha_2}^2, \]
\[ |.644472 + \phi_2| \leq .418311, \]
\[ -1.06303 \leq \phi_2 \leq -.22641. \]

Setting \( \phi_1 = \phi_2 = 0 \), we have

(2.5.25) \[ W_1^* = \hat{\phi}_1 [\hat{\sigma}^2 (1) V_1]^{-1} \hat{\phi}_1 \]
\[ = \begin{pmatrix} -0.415218 \\ -0.176277 \end{pmatrix} \]
\[ = \begin{pmatrix} 0.0607718 \\ 0.2 \end{pmatrix} \]
\[ = 7.14621 \]
\[ < 8.66704 = \chi_{2,1-\alpha_1}^2 \], and

\[ W_2^* = \hat{\phi}_2 [\hat{\sigma}^2 (2) V_2]^{-1} \hat{\phi}_2 \]
\[ = (-0.644472) [(0.019233) (2.09948)]^{-1} (-0.644472) \]
\[ = 10.2861 \]
\[ > 4.33352 = \chi_{1,1-\alpha_2}^2. \]

Thus \( \phi_1 = \phi_2 = 0 \) is not in the above \((1-\alpha)\) SCS for \( \phi_1, \phi_2 \). At level \( \alpha \), we will reject the hypothesis \( H_0^* : \phi_1 = \phi_2 = 0 \).

A \((1-\alpha_1)\) confidence set for \( \theta_1 \) is given by (2.5.23). That is,

(2.5.26) \[ \begin{pmatrix} .415218 + \theta_1 \\ .176277 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} .415218 \\ .176277 \end{pmatrix} \leq .316027. \]

Both \( \theta_1 = 0 \) and \( \theta_1 = \hat{\theta}_1 \) are in the above confidence set for \( \theta_1 \). For a given value of \( \theta_1, \theta_0^1 \), we have
Two different procedures are used to construct the confidence set for $\theta_2$.

Procedure I:

A $(1-\alpha_2)$ confidence set for $\theta_2$ is

\[(2.5.28) \quad \left| \hat{\theta}_2 - \theta_2 \right| \leq \chi^2_{1,1-\alpha_2} \hat{V}(\hat{\theta}_2)^{1/2}. \]

At $\theta_0^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,

\[(2.5.29) \quad \hat{\theta}_2 = -0.644472, \quad \left| \hat{\theta}_2 - \theta_2 \right| \leq \left\{ \chi^2_{1,0.962631} \hat{\sigma}^2(\hat{\theta}_2) \right\}^{1/2} = \left\{ (4.33352)(0.019233)(2.09948) \right\}^{1/2} = 0.418311. \]

Thus $-1.06278 \leq \theta_2 \leq -0.226162$.

At $\theta^0_1 = \hat{\theta}_1 = \begin{pmatrix} -0.415218 \\ -0.176277 \end{pmatrix}$,

\[(2.5.30) \quad \hat{\theta}_2 = -0.693727, \quad \left| \hat{\theta}_2 - \theta_2 \right| \leq 0.354989. \]

Thus $-1.04872 \leq \theta_2 \leq 0.338738$.

Procedure II:

At $\alpha_2 = 0.037369$, a $(1-\alpha_2/2)$ confidence set for $\gamma_2$ is

\[(2.5.31) \quad \left| \hat{\gamma}_2 - \gamma_2 \right| \leq 0.366596, \quad -0.283324 \leq \gamma_2 \leq 0.449868. \]

Similarly, a $(1-\alpha_2/2)$ confidence set for $\phi_2$ is
At $\theta_1^0 = 0$, the confidence set for $\theta_2$ is the same as the confidence set for $\phi_2$. That is,

\begin{equation}
-1.11705 \leq \theta_2 \leq -0.171895.
\end{equation}

At $\theta_1^0 = \hat{\theta}_1$,

\begin{equation}
\hat{\theta}_2 = -0.693727,
\end{equation}

\begin{equation}
-1.38314 \leq \theta_2 \leq -0.0043099.
\end{equation}

Procedure (II) gives a longer confidence interval for $\theta_2$. Similarly, with criterion (2) given by (2.4.32),

\begin{equation}
\hat{\theta}_2 = .0187951,
\end{equation}

\begin{equation}
\alpha_1 = .0318027,
\end{equation}

\begin{equation}
\chi^2_{2,1-\alpha_1} = 7.94832,
\end{equation}

\begin{equation}
\chi^2_{1,1-\alpha_2} = 4.6091,
\end{equation}

\begin{equation}
\chi^2_{1,1-\alpha_2/2} = 5.81375.
\end{equation}

A $(1-\alpha_1)$ confidence set for $\phi_1$ is

\begin{equation}
[\begin{bmatrix} .415218 \\ .176277 \end{bmatrix} + \phi_1] [\begin{bmatrix} \frac{1}{2} & -1 \\ -1 & 2 \end{bmatrix} [\begin{bmatrix} .415218 \\ .176277 \end{bmatrix} + \phi_1] \leq .28982.
\end{equation}

A $(1-\alpha_2)$ confidence set for $\phi_2$ is
Thus criterion (2) gives a smaller confidence set for \( \phi_1 \) and a larger confidence set for \( \phi_2 \). Since \( \phi_1 = \phi_2 = 0 \) is not in the above \((1-\alpha)\) SCS for \( \phi_1 \) and \( \phi_2 \), we will reject the hypothesis \( H_0 \) at level \( \alpha \).

A \((1-\alpha_1)\) confidence set for \( \theta_1 \) is given by (2.6.36). That is,

\[
(2.5.38) \quad \left[ \begin{array}{c} .415218 \\ .176277 \end{array} \right] + \theta_1 \leq .28982.
\]

Procedure I:

At \( \theta_1^0 = (0) \),

\[
(2.5.39) \quad \hat{\theta}_2 = -.644472, \\
|\hat{\theta}_2 - \theta_2| \leq .431407, \\
-1.07588 \leq \theta_2 \leq -.213066.
\]

At \( \theta_1^0 = \hat{\theta}_1 = \begin{bmatrix} -.415218 \\ -.176277 \end{bmatrix} \),

\[
(2.5.40) \quad \hat{\theta}_2 = -.693727, \\
|\hat{\theta}_2 - \theta_2| \leq .366103, \\
-1.05983 \leq \theta_2 \leq -.327624.
\]

Procedure II:

At \( \alpha_2 = .0318027 \), a \((1-\alpha_2/2)\) confidence set for \( \gamma_2 \) is
\[(2.5.41) \quad \left| \gamma_2 - \gamma_2' \right| \leq .375856, \]
\[-.292585 \leq \gamma_2 \leq .459128. \]

A \((1-(\alpha_2/2))\) confidence set for \(\phi_2\) is given by

\[(2.5.42) \quad \left| \phi_2 - \phi_2' \right| \leq .484515, \]
\[-1.12899 \leq \phi_2 \leq -.159958. \]

At \(\theta_0^0 = (0,0)\), \(\phi_2 = \phi_2'\), we then have

\[(2.5.43) \quad \hat{\theta}_2 = -.644472, \]
\[\left| \hat{\theta}_2 - \theta_2 \right| \leq .484515, \]
\[-1.12899 \leq \phi_2 \leq -.159958. \]

At \(\theta_0^0 = \hat{\theta}_1\),

\[(2.5.44) \quad \hat{\theta}_2 = -.693727, \]
\[\left| \hat{\theta}_2 - \theta_2 \right| \leq .706832, \]
\[-1.40056 \leq \phi_2 \leq 0.0131048. \]

For both criteria, the exact confidence set for \(\theta_2\) depends on the value of \(\theta_1\). By both criteria, procedure (II) gives a larger confidence set for \(\theta_2\) than procedure (I).

Criterion (1) minimizes the maximum of the weighted lengths of the two quadratic forms. Criterion (2) minimizes the sum of the weighted lengths of the two quadratic forms, where the weight is the inverse of the corresponding d.f.

At \(\alpha = .05\), from criterion (1), we have
\[(2.5.45)\]
\[
\begin{align*}
\chi^2_{2,1-\alpha_1} &= 8.66704, \\
\chi^2_{1,1-\alpha_2} &= 4.33352,
\end{align*}
\]
\[
\max\{\frac{\chi^2_{2,1-\alpha_1}}{2}, \chi^2_{1,1-\alpha_2}\} = 4.33352,
\]
\[
\frac{\chi^2_{2,1-\alpha_1}}{2} + \chi^2_{1,1-\alpha_2} = 8.66704.
\]

From criterion (2), we have

\[(2.5.46)\]
\[
\begin{align*}
\chi^2_{2,1-\alpha_1} &= 7.74832, \\
\chi^2_{1,1-\alpha_2} &= 4.6091,
\end{align*}
\]
\[
\max\{\frac{\chi^2_{2,1-\alpha_1}}{2}, \chi^2_{1,1-\alpha_2}\} = 4.6091,
\]
\[
\frac{\chi^2_{2,1-\alpha_1}}{2} + \chi^2_{1,1-\alpha_2} = 8.58326.
\]

Thus for the set with larger d.f. (two d.f.) criterion (1) gives a larger confidence coefficient (.9868787) than criterion (2) (.9812049). And for the set with smaller d.f. (one d.f.), criterion (1) gives a smaller confidence coefficient (.9681973) than criterion (2) (.9681973).
CHAPTER III

INCOMPLETE MULTIRESPONSE MODEL

3.1 Introduction

In multiresponse design, if the same $p$ responses are measured on each unit of the experiment, we may have a standard MGLM. However, in some cases it is either physically impossible or uneconomical to examine all the responses on each experimental unit. In these cases, it is important to have a procedure that can be used to analyze the experiment where observations on some of the responses are missing not by accident, but by design.

One class of designs with missing responses is the IM model proposed by Srivastava (1968). The model definition and the general procedure to analyze the data are outlined in Section 1.6.

The IM model defined by Srivastava is based on a less-than-full-rank model in each of the disjoint data sets. In Section 3.2, we propose a corresponding full rank IM model. We then derive the procedure to analyze the data and show some simplification of the procedure.

In Section 3.3, we consider a special type of IM full rank model upon which the analyzing procedure can be simplified. We then construct an optimal procedure to minimize the generalized variance of the estimated parameters, and demonstrate the invariance property of the proposed optimal procedure.

To evaluate the IM design, Section 3.4 compares the IM model derived in Section 3.3 to the corresponding complete multiresponse (CM) model with respect to generalized variance, asymptotic relative efficiency, cost, and minimum risk.
Section 3.5 gives a numerical example to illustrate the IM full rank procedure derived in this chapter.

3.2 A full rank incomplete multiresponse model

Recalling Section 1.6, the given IM model has the restriction that the sum of all elements in each column of the unknown matrix of parameters $\beta^*$ is zero. To simplify the procedure in analyzing this type of data, we construct an IM model without the restriction on the unknown parameters. The resulting full rank model is given by

\[(3.2.1)\]

$$
\begin{align*}
E(Y_j) &= (X_{j1} X_{j2}) \begin{bmatrix} \Gamma_j \\ \beta B_j \end{bmatrix}, \\
V(Y_j) &= I_{n_j} \otimes \Sigma_j, \quad j=1, \ldots, u,
\end{align*}
$$

where

- $Y_j(n_j \times p_j)$ is a data matrix,
- $(X_{j1}(n_j \times r_j) X_{j2}(n_j \times r_j))$ is a design matrix of rank $(r_j \times r)$,
- $\Gamma_j(r_j \times p_j)$ is a matrix of unknown parameters,
- $\beta(r \times p)$ is a matrix of common unknown parameters,
- $B_j(p \times p_j)$ is an incidence matrix of 0's and 1's,
- $\Sigma_j(p_j \times p_j) = B_j' \Sigma B_j$, is the covariance matrix for the $p_j$ responses in $Y_j$.
- $\Sigma(p \times p)$ is a positive definite covariance matrix.

The hypothesis of interest is

$$H_0: \Gamma B U = 0,$$

where $C(g \times r)$ is of rank $g$, $U(p \times v)$ is of rank $v$.

Prior to analysis it is necessary for the data to be transformed into the
framework of a standard MGLM which meets the specified conditions. A procedure by which this may be performed is as follows:

(i) First construct a linear set of \( Y_j \).

\[
Q_j (r \times p_j) = M_j Y_j, \quad j = 1, \ldots, u,
\]

where \( M_j (n_j \times r) \) is a known matrix of orthogonal columns such that

\[
M_j^t M_j = I_r, \quad M_j^t X_j = 0.
\]

Then we have

\[
E(Q_j) = A_j B B_j, \quad V(Q_j) = I_r \otimes \Sigma_j,
\]

where \( A_j (r \times r) = M_j^t X_j \), a known matrix.

(ii) The design is homogeneous, i.e., there exists a set of known \((r \times r)\) matrices, \( \{F_1, \ldots, F_m\} \), such that \( A_j = \sum_{k=1}^m \alpha_{jk} F_k \) for some known scalars \( \alpha_{j1}, \ldots, \alpha_{jm}, \quad j = 1, \ldots, u \). We then have

\[
E(Q_j) = \left[ \sum_{k=1}^m \alpha_{jk} F_k \right] B B_j = \sum_{k=1}^m (F_k \beta)(\alpha_{jk} B_j).
\]

(iii) Combine \( Q_1, \ldots, Q_u \),

\[
Q (r \times u) = (Q_1, \ldots, Q_u), \quad \text{then we have}
\]

\[
E(Q) = (E(Q_1), \ldots, E(Q_u))
\]

\[
= (\sum_{k=1}^m (F_k \beta)(\alpha_{1k} B_1), \ldots, \sum_{k=1}^m (F_k \beta)(\alpha_{uk} B_u))
\]

\[
= (F_1 \beta, \ldots, F_m \beta)L,
\]

\[
V(Q) = I_r \otimes \Sigma(0),
\]
where \( L = \begin{pmatrix} l_1 \\ \vdots \\ l_m \end{pmatrix} \) 

\[
L_k(p \times \sum_{j=1}^{m} p_j) = (\alpha_{1k} B_1, \ldots, \alpha_{uk} B_u),
\]

\[
\Sigma_{(0)}(\sum_{j=1}^{u} p_j \times \sum_{j=1}^{m} p_j) = \begin{pmatrix} \Sigma_1 \\ \vdots \\ \Sigma_u \end{pmatrix}
\]

(iv) \( LL' \) is nonsingular.

Denote \( H_{jk} = (H_{jk}) = L'(LL')^{-1} \)

where \( H_{jk} \) is \((p_j \times p), \ j=1, \ldots, u, \ k=1, \ldots, m\).

(v) Let

\[
Z^*_{(r \times mp)} = QH = \left( \sum_{j=1}^{u} Q_j l_1, \ldots, \sum_{j=1}^{u} Q_j l_m \right), \text{ then}
\]

\[
E(Z^*) = (F_1 \beta, \ldots, F_m \beta),
\]

\[V(Z^*) = I_r \otimes [H' \Sigma_{(0)}] H \]

\[= I_r \otimes (\Sigma(k_1, k_2)),\]

where

\[
\Sigma(k_1, k_2) = \sum_{j=1}^{u} H'_{jk_1} \Sigma_{h} H_{jk_2}, \quad k_1, k_2 = 1, \ldots, m.
\]

(vi) Rearrange the elements in \( Z^* \).

\[
Z_{(mr \times p)} = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}
\]
where
\[ Z_k = \sum_{j=1}^{u} Q_j H_{jk}, \quad k=1, \ldots, m. \]

Then we obtain the expectation of \( Z \)
\[
E(Z) = \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix} \beta.
\] (3.2.9)

(vii) The factorization of \( V(Z) \) is possible, i.e., there exists a known positive definite matrix \( W_{(mr \times mr)} \), such that
\[
(3.2.10) \quad V(Z) = W \otimes \Sigma^*,
\]
where \( \Sigma^*_{(p \times p)} \) is a positive definite matrix.

(viii) If the above conditions are met, then there exists a nonsingular matrix \( R_{(mr \times mr)} \), such that \( RWR' = I \). Then we have
\[
E(RZ) = \begin{bmatrix} R \\ \vdots \\ R \end{bmatrix} \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix} \beta,
\]
\[
V(RZ) = I_{mr} \otimes \Sigma^*.
\] (3.2.11)

The rows of \( RZ \) are independently normally distributed with common covariance matrix \( \Sigma^* \). Thus we have a standard MGLM model, by which the usual analysis procedures are applicable.

A critical condition in the above procedure is the factorization of \( V(Z) \).

With
\[
Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_m \end{bmatrix}, \quad Z_k = \sum_{j=1}^{u} Q_j H_{jk}, \quad k=1, \ldots, m,
\]
the covariance between \( Z_{k_1} \) and \( Z_{k_2} \) is given by
\[
(3.2.12) \quad \text{Cov}(Z_{k_1}, Z_{k_2}) = I_r \otimes \Sigma(k_{1,k_2}), \quad k_1, k_2 = 1, \ldots, m.
\]
If all the $m^2$ covariance matrices $\Sigma(k_1,k_2)$, $k_1k_2 = 1, \ldots, m$, are proportional to a fixed matrix, $\Sigma^*$, then we can factorize $V(Z)$ into matrices with proper orders.

**Theorem 3.2.1.** If there exist a positive definite ($p \times p$) matrix $\Sigma^*$ and known scalars $\lambda_{k_1k_2}$, $k_1k_2 = 1, \ldots, m$, such that

\begin{equation}
\Sigma(k_1,k_2) = \lambda_{k_1,k_2} \Sigma^*,
\end{equation}

then we have a suitable factorization of $V(Z)$.

\begin{equation}
V(Z) = W \otimes \Sigma^*,
\end{equation}

where

\[ W = I_r \otimes \Lambda, \]

\[ \Lambda_{(m \times m)} = \begin{bmatrix}
\lambda_{11} & \cdots & \lambda_{1m} \\
\vdots & \ddots & \vdots \\
\lambda_{m1} & \cdots & \lambda_{mm}
\end{bmatrix} \]

### 3.2.1 Removing the singularity of $LL'$

For the IM model with the restriction on the matrix of unknown parameters $\beta^*$, application of the procedure transforming the data back into the framework of a standard MGLM model gives

\begin{equation}
E(Q^*_j([r+1] \times p_j)) = \sum_{k=1}^{m^*} F_k^* \beta^* (\alpha_{jk} + \theta_j) B^*_j, \quad j = 1, \ldots, u,
\end{equation}

where

\[ \theta_j \text{ is a fixed constant}, \]

\[ \sum_{k=1}^{m^*} F_k = J([r+1] \times [r+1]). \]

\begin{equation}
E\left(Q^* \left(\sum_{j=1}^{u} \left( [r+1] \times [r+1] \right) \right) \right) = (F^*_1 \beta^*_1, \ldots, F^*_m \beta^*_m) L^*,
\end{equation}

where
By using suitable choices of $\theta_j$'s, we aim to remove the singularity of $L^*L^{**}$. Once done, the resulting matrix $L^*$ is of full row rank $mp$.

A matrix of unrestricted unknown parameters $\beta$ can be constructed from the last $r$ rows of $\beta^*$. The resulting full rank model procedure then gives

\begin{align}
\beta_{(r \times p)} &= (O_{(r \times 1)}, I_{r})\beta^*, \\
Q_j_{(r \times r)} &= (O_{(r \times 1)}, I_{r})Q_j^*, \quad j=1, \ldots, u.
\end{align}

Thus

\begin{align}
E(Q_j) &= \sum_{k=1}^{m^*} F_k \beta (\alpha_{jk} + \theta_j) B_j, \quad j=1, \ldots, u, \\
E(Q_{(r \times \sum_{j=1}^{u} P_j)}) &= (F_1 \beta, \ldots, F_{m^*} \beta)L^*.
\end{align}

With $m = m^* - 1$, the restriction of $\sum_{k=1}^{m^*} F_k = 0$ can be removed by replacing $F_{m^*}$ with $-\sum_{k=1}^{m} F_k$. Then

\begin{align}
E(Q_j) &= \sum_{k=1}^{m} F_k \beta (\alpha_{jk} + \theta_j) B_j - \sum_{k=1}^{m} F_k \beta (\alpha_{jk} + \theta_j) B_j \\
&= \sum_{k=1}^{m} F_k \beta (\alpha_{jk} - \alpha_{jm^*}) B_j, \quad \text{and} \\
E(Q_{(r \times \sum_{j=1}^{u} P_j)}) &= (F_1 \beta, \ldots, F_{m^*} \beta)L^*.
\end{align}
where
\[
L = \begin{bmatrix}
L_1 \\ \\
\vdots \\ \\
L_m
\end{bmatrix},
\]
\[
L = \begin{bmatrix}
\alpha_{1k} - \alpha_{1m^*} \\
\vdots \\
\alpha_{uk} - \alpha_{um^*}
\end{bmatrix} B_1, \ldots, \begin{bmatrix}
\alpha_{1k} - \alpha_{1m^*} \\
\vdots \\
\alpha_{uk} - \alpha_{um^*}
\end{bmatrix} B_p
\]
\[
= L^+_{k*} - L^*_{m^*}, \quad k=1,\ldots,m.
\]

Thus \( LL' \) is nonsingular if \( L^*L'^* \) is nonsingular.

**Theorem 3.2.2.** Without the restriction on the set of matrices \( \{F_1, \ldots, F_m\} \) in the full rank model procedure, \( \sum_{k=1}^{m} F_k = 0 \), the nonsingularity of \( LL' \) cannot be removed by a suitable selection of \( \{\theta_1, \ldots, \theta_u\} \) in the corresponding restricted model procedure.

### 3.3 Optimum confidence procedure

For simplicity, we consider an IM design upon which the full rank model procedure can be performed. The IM design satisfies the following conditions:

\[(3.3.1)\]

1. \( u \) is even, even number of disjoint sets;
2. \( B_{j+u/2} = B_j, \quad j=1,\ldots,u/2, \)
3. \( A_j = F_1, \quad A_{j+u/2} = F_2, \quad j=1,\ldots,u/2, \)
4. All responses are symmetric w.r.t. the \( u \) disjoint sets involved. That is, the number of disjoint sets for which a specific set of exactly \( j \) responses are measured is \( \ell_j \), the same for all \( \binom{p}{j} \) different combinations of \( j \) responses, \( j=1,\ldots,p. \)

From the condition (4) we have

\[(3.3.2)\]

1. \( \sum_{j=1}^{p} \ell_j \binom{p}{j} = u, \)
Each response is measured on \( k_0 \) disjoint sets where
\[
k_0 = \frac{P}{2} \ell_j (p-1) , \quad \text{and}
\]

Each pair of responses is measured on \( k_1 \) disjoint sets
where
\[
k_1 = \sum_{j=2}^{\ell} \ell_j (p-2).
\]

The resulting full rank model procedure gives

\[
E(Q_j) = F_1 \beta B_j ,
\]

\[
E(Q_{j+u/2}) = F_2 \beta B_j , \quad j=1,...,u/2.
\]

We then have

\[
E(Q) = \left( \begin{array}{c}
F_1 \beta F_2 \beta \\
L \\
\end{array} \right) L ,
\]

\[
V(Q) = I_r \otimes I_2 \otimes \left( \begin{array}{c}
B_1 \Sigma B_1 \\
\vdots \\
B_{u/2} \Sigma B_{u/2}
\end{array} \right). 
\]

where \( L = I_2 \otimes (B_1,\ldots,B_{u/2}) \).

With
\[
LL' = [I_2 \otimes (B_1,\ldots,B_{u/2})][I_2 \otimes (B_1,\ldots,B_{u/2})]' 
\]

\[
= I_2 \otimes \sum_{j=1}^{u/2} B_j B_j' 
\]

\[
= I_2 \otimes \left[ \frac{k_0}{2} I_p \right] 
\]

\[
= \frac{k_0}{2} I_{2p} , \quad \text{and}
\]

\[
Z_{(r \times 2p)}^* = QL'(LL')^{-1}
\]

\[
= \left( \sum_{j=1}^{u/2} Q_j B_j', \sum_{j=u/2+1}^{u} Q_j B_j' \right) \left( \frac{k_0}{2} I_{2p} \right)^{-1}
\]
we have

(3.3.5) \[ E(Z^*) = (F_1 \beta, F_2 \beta), \]

\[ V(Z^*) = I_r \otimes \frac{4}{k_0^2} I_2 \otimes \frac{u}{2} \left( \sum_{j=1}^{u/2} (B_j B_j^T) \sum_{j=u/2+1}^{u} (B_j B_j^T) \right) \]

\[ = \frac{2}{k_0} I_r \otimes [I_2 \otimes \Sigma^*], \]

where

\[ \Sigma^* = \frac{2}{k_0} \sum_{j=1}^{u/2} (B_j B_j^T) \sum_{j=u/2+1}^{u} (B_j B_j^T) \]

\[ = \text{Diag}(\Sigma) + (k_1/k_0) \text{Off Diag}(\Sigma). \]

With

\[ Z = \frac{2}{k_0} \begin{bmatrix} \sum_{j=1}^{u/2} Q_j B_j^T \\ \sum_{j=(u/2)+1}^{u} Q_j B_j^T \end{bmatrix}, \]

we have a standard MGUM model.

(3.3.6) \[ E(Z) = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \beta, \]

\[ V(Z) = \frac{2}{k_0} I_{2r} \otimes \Sigma^*. \]

The confidence set for C8U can be constructed through the usual confidence procedures used in the standard MGUM model. A usual estimate of \( \beta \) is given by

(3.3.7) \[ \hat{\beta} = \left[ \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \right]^{-1} \left[ \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \right]' Z \]

\[ = \frac{2}{k_0} (F_1^T F_1 + F_2^T F_2)^{-1} \left[ \frac{u}{2} \sum_{j=1}^{u/2} Q_j B_j^T + F_2^T \sum_{j=u/2+1}^{u} Q_j B_j^T \right], \]

with

(3.3.8) \[ V(\hat{\beta}) = \left[ \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \right]' \left[ \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \right]^{-1} \otimes \left[ \frac{2}{k_0} \Sigma^* \right]. \]
The full rank IM model procedure begins with the selection of the set of matrices \(\{M_1, \ldots, M_u\}\) such that

\[
\begin{align*}
M_jM_j^T &= I_r, \\
M_jX_{jl} &= 0, \ j=1, \ldots, u.
\end{align*}
\]

The above selection of \(\{M_1, \ldots, M_u\}\) is not unique.

An optimal confidence procedure can be used to derive the set of matrices \(\{M_1, \ldots, M_u\}\) such that the generalized variance of \(\hat{\beta}\) is minimized.

With \(V(\hat{\beta}) = \frac{2}{k_0} (F_1'F_1 + F_2'F_2)^{-1} \otimes \Sigma^*\), the generalized variance is

\[
(3.3.10) \quad |V(\hat{\beta})| = \left(\frac{2}{k_0}\right)^{r_p} (|F_1'F_1 + F_2'F_2|)^{-r_p} |\Sigma^*|^r.
\]

Since \(\left(\frac{2}{k_0}\right)^{r_p}|\Sigma^*|^r\) is invariant under different selections of the set \(\{M_1, \ldots, M_u\}\) satisfying the conditions (3.3.9), the optimal criterion is equivalent to maximizing \(|F_1'F_1 + F_2'F_2|\).

Letting \(U = \{u_1, \ldots, u_{r_1}\}\) be an orthogonal basis of the vector space generated by the columns of \(X_{11}\); then an additional set of orthogonal vectors \(V = \{v_1, \ldots, v_r\}\) can be constructed such that \(\{U, V\}\) is an orthogonal basis of the vector space generated by the columns of \(X_1 = (X_{11}, X_{12})\).

Denoting the matrices \(U_{0(n_1 \times r_1)} = (u_1, \ldots, u_{r_1})\) and \(V_{0(n_1 \times r)} = (v_1, \ldots, v_r)\), we then have

\[
(3.3.11) \quad X_{12} = U_{0}T_u + V_{0}T_v \text{ for some known matrices } T_u(r_1 \times r), \ T_v(r \times r).
\]

Comparing the column space of \(M_1\) with that of \(V_0\), we have two different situations:
1) Type I: the column spaces of $M_1$ and $V_0$ are the same.

2) Type II: the column spaces of $M_1$ and $V_0$ are not the same.

Under Type I situation, there exists a known orthogonal matrix $T_1 (r \times r)$ such that $M_1 = V_0 T_1$. We then have

\begin{equation}
F_{1(I)} = Z_1^T X_{12} = T_1^T V_0 (U_0^T u + V_0^T v) = T_1^T v.
\end{equation}

\begin{equation}
F_{1(I)}^T F_{1(I)} = T_1^T T_1^T T_1^T V = T_1^T v.
\end{equation}

Letting $K = \{k_1, \ldots, k_{n_1-r_1-r}\}$ be an orthogonal basis of the vector space which is orthogonal to the column space of $X_1$, we then denote

\begin{equation}
K_0(n_1 \times [n_1-r_1-r]) = (k_1, \ldots, k_{n_1-r_1-r}).
\end{equation}

Under Type II situation, there exist known matrices $T_2 (r \times r)$ and $T_3 ([n_1-r_1-r] \times r)$ such that

\begin{equation}
M_1 = V_0 T_2 + K_0 T_3.
\end{equation}

We then have

\begin{equation}
F_{1(II)} = M_1^T X_{12} = (T_2^T V_0^T + T_3^T K_0^T) (U_0^T u + V_0^T v) = T_2^T v.
\end{equation}

\begin{equation}
F_{1(II)}^T F_{1(II)} = T_2^T T_2^T v.
\end{equation}

From $M_1^T M_1 = (V_0 T_2 + K_0 T_3)^T (V_0 T_2 + K_0 T_3)$

\begin{equation}
= T_2^T T_2 + T_3^T T_3
\end{equation}

and the condition $M_1^T M_1 = I$, we have

\begin{equation}
T_2^T T_2 + T_3^T T_3 = I.
\end{equation}

Together with (3.3.13), we have
Extending the above result to all \( u \) disjoint sets,
\[
|F'_{1(II)}F_{1(II)}| \leq |F'_{1(I)}F_{1(I)}|,
\]
\[
|F'_{2(II)}F_{2(II)}| \leq |F'_{2(I)}F_{2(I)}|,
\]
\[
|F'_{1(II)}F_{1(II)} + F'_{2(II)}F_{2(II)}| \leq |F'_{1(I)}F_{1(I)} + F'_{2(II)}F_{2(II)}| \leq |F'_{1(I)}F_{1(I)} + F'_{2(I)}F_{2(I)}|.
\]

**Theorem 3.3.1.** To meet the criterion of minimizing the generalized variance of the parameter estimator \( \hat{\beta} \) given by (3.3.7), an optimum confidence procedure for the IM model satisfying the conditions (3.3.1) can be achieved by using the Type I procedure to construct the set \( \{M_1, \ldots, M_u\} \) in the full rank model procedure.

**3.3.1 Invariance in the optimum confidence procedure**

In the above optimum confidence procedure, the selection of the set \( \{M_1, \ldots, M_u\} \) is not unique. The effect of different Type I selections of the set \( \{M_1, \ldots, M_u\} \) on the resulting confidence set can be examined through the transformations.
\[(3.3.17) \quad M_j \rightarrow M_j P_1 \]

\[ M_{j+u/2} \rightarrow M_{j+u/2} P_2, \quad j=1, \ldots, u/2, \]

where \( P_1 \) and \( P_2 \) are \((r \times r)\) orthogonal matrices.

Under these transformations,

\[(3.3.18) \quad Q_j \rightarrow P_1^T Q_j, \]

\[ Q_{j+u/2} \rightarrow P_2^T Q_{j+u/2}, \quad j=1, \ldots, u/2 \]

\[ Z_1 \rightarrow P_1^T Z_1, \]

\[ Z_2 \rightarrow P_2^T Z_2, \]

\[ F_1 \rightarrow P_1^T F_1, \]

\[ F_2 \rightarrow P_2^T F_2. \]

The induced transformations for \( \beta \) and \( \Sigma^* \) on the parameter space are identity transformations.

\[(3.3.19) \quad \beta \rightarrow \beta, \]

\[ \Sigma^* \rightarrow \Sigma^*. \]

From \((3.3.18)\)

\[ (P_1^T F_1)' (P_1^T F_1) = F_1^T F_1, \]

\[ (P_1^T F_2)' (P_1^T F_2) = F_1^T F_2, \]

\[ (P_1^T Z_1)' (P_1^T Z_1) = F_1^T Z_1, \]

\[ (P_2^T Z_2)' (P_2^T Z_2) = F_2^T Z_2. \]

Thus both \( \hat{\beta} \) and \( V(\hat{\beta}) \) are invariant under the transformations \((3.3.17)\), and
we have the following result.

**Theorem 3.3.2.** The optimum confidence procedure given by Theorem 3.3.1 is invariant under different selections of the set \( \{M_1, \ldots, M_u\} \) satisfying the Type I conditions.

### 3.4 Some comparisons between incomplete multiresponse model and complete multiresponse model

The full rank model satisfying the set of conditions (3.3.1) in Section 3.3. is

\[
\begin{align*}
E(Q_j) &= F_1 \beta B_j, \\
E(Q_{j+u/2}) &= F_2 \beta B_j,
\end{align*}
\]

and

\[
V(Q_{j+u/2}) = V(Q_j) = I_r \otimes \Sigma_j, \quad j=1, \ldots, u/2.
\]

The resulting estimate of \( \beta \) and its variance are

\[
\hat{\beta}_{IM} = \frac{2}{k_0} (F_1'F_1 + F_2'F_2)^{-1} \left[ \sum_{j=1}^{u/2} Q_j \beta_j' + \sum_{j=u/2+1}^{u} Q_j B_j' \right], \quad \text{and}
\]

\[
V(\hat{\beta}_{IM}) = \frac{2}{k_0} (F_1'F_1 + F_2'F_2)^{-1} \otimes \Sigma^*,
\]

where

\[
\Sigma^* = \text{Diag}(\Sigma) + \left(\frac{k_1}{k_0}\right) \text{Off Diag}(\Sigma).
\]

The corresponding CM model is given by

\[
\begin{align*}
E(Q_j) &= F_1 \beta, \\
E(Q_{j+u/2}) &= F_2 \beta, \\
V(Q_j) &= I_r \otimes \Sigma, \quad j=1, \ldots, u.
\end{align*}
\]

A resulting standard MGLM is
(3.4.4) \quad \mathbb{E}(Y) = X\beta,

and \quad V(Y) = I_{ur} \otimes \Sigma,

where \quad Y_{(ur \times p)} = \begin{pmatrix} Q_1 \\ \vdots \\ Q_u \end{pmatrix},

\quad X_{(ur \times r)} = \begin{pmatrix} F_1 \otimes J \left( \left( \frac{u}{2} \right) \times 1 \right) \\ F_2 \otimes J \left( \left( \frac{u}{2} \right) \times 1 \right) \end{pmatrix},

and \quad J \left( \left( \frac{u}{2} \right) \times 1 \right) \text{ is a } \left( \left( \frac{u}{2} \right) \times 1 \right) \text{ column of 1's.}

The estimate of \( \beta \) and its variance are

(3.4.5) \quad \hat{\beta}_{CM} = (X'X)^{-1}X'Y = \frac{2}{u} \left( F_1'F_1 + F_2'F_2 \right)^{-1} \left[ \sum_{j=1}^{u/2} Q_j + \sum_{j=u/2+1}^{u} Q_j \right],

and \quad V(\hat{\beta}_{CM}) = \frac{2}{u} \left( F_1'F_1 + F_2'F_2 \right)^{-1} \otimes \Sigma.

3.4.1 Generalized variance

The ratio of the generalized variance for \( \hat{\beta}_{IM} \) to that for \( \hat{\beta}_{CM} \) is

(3.4.6) \quad \frac{|V(\hat{\beta}_{IM})|}{|V(\hat{\beta}_{CM})|} = \frac{|\frac{2}{u} \left( F_1'F_1 + F_2'F_2 \right)^{-1} \otimes \Sigma|}{|\frac{2}{u} \left( F_1'F_1 + F_2'F_2 \right)^{-1} \otimes \Sigma|} = \left( \frac{u}{k_0} \right)^{p} \frac{|\Sigma^*|}{|\Sigma|} r.

Thus, we can compare these two models by

(3.4.7) \quad R = \left\{ \frac{|V(\hat{\beta}_{IM})|}{|V(\hat{\beta}_{CM})|} \right\}^{1/r} = \left( \frac{u}{k_0} \right)^{p} \frac{|\Sigma^*|}{|\Sigma|}.

Depending on the structure of \( \Sigma \), we consider the following two special cases:

1) Intraclass covariance structure
In this case,
\[ \Sigma = \sigma_0^2 [I + \rho (J - I)] , \]
and
\[ \Sigma^* = \sigma_0^2 [I + r_1 \rho (J - I)] , \]
where
\[ r_1 = \frac{k_1}{k_0} < 1, \quad -1/(p-1) < \rho < 1 . \]

Then we have
\[ (3.4.8) \quad R = R(\rho) = \frac{u_p}{k_0} \frac{(1-r_1 \rho)^{p-1} [1+(p-1)r_1 \rho]}{(1-\rho)^{p-1} [1+(p-1)\rho]} \]

After some simplification,
\[ \frac{3R(\rho)}{3\xi} = \frac{u_p}{k_0} \frac{p(p-1)(1-\rho)^{p-2}(1-r_1 \rho)^{p-2}}{[(1-\rho)^{p-1}[1+(p-1)\rho]]^2} (1-r_1) \{ 1 + r_1 \{ 1 + (p-2)\rho \} \}^2 \]
which has the same sign as \( \rho \).

As \( \rho \) increases from \(-1/(p-1)\) to 0, \( R(\rho) \) decreases from \( R(-1/(p-1)) \) to \( +\infty \) to \( R(0) = \frac{u_p}{k_0} > 1 \). As \( \rho \) increases from 0 to 1, \( R(\rho) \) increases from \( \frac{u_p}{k_0} \) to \( R(1) = +\infty \). That is, the generalized variance from the IM model is greater than the generalized variance from the CM model. As the absolute value of \( \rho \) increases, \( R(\rho) \) increases, making the IM model less desirable.

2) Autocorrelation when \( p=3 \)

In this case,
\[ \Sigma = \sigma_0^2 \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} \]
\[ \Sigma^* = \sigma_0^2 \begin{pmatrix} 1 & r_1 \rho & r_1 \rho^2 \\ r_1 \rho & 1 & r_1 \rho \\ r_1 \rho^2 & r_1 \rho & 1 \end{pmatrix} , \quad -1 < \rho < 1 . \]
Then

\[ R(\rho) = \left( \frac{u}{k_0} \right)^3 \frac{1 - 2r_1^2 \rho^2 - r_1^4}{1 - 2 \rho^2 + \rho^4}. \]

**Theorem 3.4.1.** A comparison of the generalized variance between the IM model and its corresponding CM model is given by \( R \) in (3.4.6). Considering the intraclass covariance structure and autocorrelation when \( p = 3 \), \( R \) attains its minimum when there is no correlation among all \( p \) responses. \( R(\rho) \) is an increasing function of \( |\rho| \). As \( |\rho| \) increases, the IM model is less desirable.

### 3.4.2 Asymptotic relative efficiency

A practical comparison of two different procedures is through a measure of their relative efficiency. When both procedures are at the same Type I error level \( (\alpha) \), the relative efficiency is the inverse ratio of the two sample sizes required to achieve the same power against a specified alternative.

In this section we compare the IM model with the corresponding CM model through a measure of asymptotic relative efficiency. For both models, the null hypothesis is \( H_0: C_{(g \times r)} \beta J_{(p \times 1)} = 0 \), where \( C \) is a matrix of rank \( g \). The Wilks likelihood ratio statistic will be used to calculate the power under the fixed alternative \( H_a: C \beta J - a \beta J_{(g \times 1)} \), where \( a \beta \) is a fixed nonzero constant.

To improve the precision of the IM design (3.3.6) we repeat the same design. By repeating the IM design \( n_1 \) times independently, we have the following standard MGLM.

\[ (3.4.10) \quad E(Z_0) = F_0 \beta, \]

\[ V(Z_0) = \frac{2}{k_0} \mathbf{I}_{2n_1} \otimes \Sigma^*, \]
where
\[ Z_0(2n_1 \times p) = \begin{pmatrix} Z(1) \\ \vdots \\ \vdots \\ Z(n_1) \end{pmatrix}, \]

and
\[ F_0(2n_1 \times r) = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \otimes J(n_1 \times 1). \]

The estimate of \( \beta \) and its corresponding variance are

\[
(3.4.11) \quad \hat{\beta}_{IM} = (F_0'F_0)^{-1}F_0'Z_0 = \frac{1}{n_1} (F_1'F_1 + F_2'F_2)^{-1} \begin{pmatrix} F_1' \\ F_2' \end{pmatrix} \sum_{j=1}^{n_1} Z(j),
\]

and
\[
V(\hat{\beta}_{IM}) = (F_0'F_0)^{-1} \otimes \frac{2}{k_0} \Sigma^* = \frac{2}{n_1k_0} (F_1'F_1 + F_2'F_2)^{-1} \otimes \Sigma^*.
\]

For the null hypothesis \( H_0 : C\beta U = 0 \), the likelihood ratio statistic is

\[
(3.4.12) \quad W_1 = \frac{|S_{E_1}|}{|S_{H_1} + S_{E_1}|},
\]

where
\[
S_{H_1} = (C\hat{\beta}_{IM})' [C(F_0'F_0)^{-1}C']^{-1} (C\hat{\beta}_{IM}),
\]
\[
S_{E_1} = U'Z_0[I - F_0(F_0'F_0)^{-1}F_0'] Z_0 U.
\]

Letting \( \tau_1 = 1 - \frac{1}{2n_1r} [r-r+\frac{1}{2}(v+r+1)] = 1 - \frac{v+r+1}{4n_1r} \),

then for large value of \( N_1 = (2n_1r)\tau_1 = 2n_1r - \frac{1}{2}(v+r+1) \), the asymptotic power function of the statistic \( W_1 \) at Type I level \( \alpha \) against the alternative \( H_a \) is
where $y_{f, 1-\alpha}$ is the $(1-\alpha)$ point of the $\chi^2$ distribution with $f$ d.f. $f = g\gamma$.

$\chi^2_f(\sigma_1)$ is the noncentral $\chi^2$ distribution function with noncentrality $\sigma_1$.

$\sigma_1 = \text{trace } \Omega_1,$

and

$$
\Omega_1 = [U^*(\frac{2}{n_1k_0} \Sigma^*)U]^{-1} (CBU)' [C(F_1^tF_1 + F_2^tF_2)^{-1}C']^{-1} (CBU)
$$

$$
= \frac{n_1k_0}{2} [U^*\Sigma^*U]^{-1} (CBU)' [C(F_1^tF_1 + F_2^tF_2)^{-1}C']^{-1} (CBU).
$$

A similar procedure can be performed for the corresponding CM design (3.4.4).

By repeating the design $n_2$ times, we have the following standard MGLM.

(3.4.14) $E(Y_0) = X_0\beta, \quad V(Y_0) = I_{n_2ur} \otimes \Sigma,$

where

$$
Y_0(\text{n}_2\text{urx}\text{p}) = \begin{bmatrix}
Y(1) \\
\vdots \\
Y(n_2)
\end{bmatrix},
$$

and

$$
X_0(\text{n}_2\text{urx}\text{r}) = \begin{bmatrix}
F_1 \otimes J((u/2)\times1) \\
F_2 \otimes J((u/2)\times1)
\end{bmatrix} \otimes J(n_2\times1).
$$

We then have

(3.4.15) $\hat{\beta}_{CM} = (X_0^tX_0)^{-1}X_0^tY_0,$

and

$$
V(\hat{\beta}_{CM}) = (X_0^tX_0)^{-1} \otimes \Sigma = \frac{2}{n_2u} (F_1^tF_1 + F_2^tF_2)^{-1} \otimes \Sigma.
$$
For the hypothesis $H_0 : \text{CBU} = 0$, the likelihood ratio statistic is

\[
W_2 = \frac{|S_{H_2}|}{|S_{H_2} + S_{E_2}|},
\]

where

\[
S_{H_2} = (\hat{\beta}_{CM})' [C(X'_0X_0)^{-1}C']^{-1}(\hat{\beta}_{CM})
\]

and

\[
S_{E_2} = U'Y_0[I - X'_0(X'_0X_0)^{-1}X_0]Y_0U.
\]

Letting \( \tau_2 = 1 - \frac{1}{n_2ur} [r-r_k(r+v+1)] = 1 - \frac{r+v+1}{2n_2ur} \),

then for large values of \( N_2 = (n_2ur) \tau_2 = n_2ur - \frac{(r+v+1)}{2} \), the asymptotic power function of the statistic \( W_2 \) at Type I level \( \alpha \) against the alternative \( H_a \) is

\[
1 - \beta_2 = p(-2\tau_2 \ln W_2 \geq \chi^2_{\alpha,1-\alpha}|H_a) \]

\[
= p(\chi^2_{\alpha,2} \geq y_{\alpha,1-\alpha}) + O(1/n_2),
\]

where

\[
\Omega_2 = \frac{n_2u}{2} (U'\Sigma U)^{-1} (\text{CBU})' [C(F'_1 F'_1 + F'_2 F'_2)^{-1}C']^{-1} (\text{CBU})
\]

For the two designs to have asymptotically the same power, it is necessary that \( \sigma_1 = \sigma_2 \). The asymptotic relative efficiency for the IM model against the corresponding complete multiresponse model is then

\[
\text{ARE} = \frac{n_2}{n_1} = k \frac{1}{u} \frac{\text{trace}([U'\Sigma U]^{-1} (\text{CBU})' [C(F'_1 F'_1 + F'_2 F'_2)^{-1}C']^{-1} (\text{CBU}))}{\text{trace}([U'\Sigma U]^{-1} (\text{CBU})' [C(F'_1 F'_1 + F'_2 F'_2)^{-1}C']^{-1} (\text{CBU}))}
\]

Under the fixed alternative \( H_a : \text{CBU} = a_0 J (g \times 1) \),

\[
(\text{CBU})' [C(F'_1 F'_1 + F'_2 F'_2)^{-1}C']^{-1} (\text{CBU}) = a_0^2 b_0 \text{ for a fixed positive constant } b_0.
\]
It follows that
\[
\text{ARE} = \frac{k_0}{u} \frac{\text{trace}\{[U'\Sigma U]^{-1} a_0^2 b_0^2\}}{\text{trace}\{[U'U]^{-1} a_0^2 b_0^2\}} = \frac{k_0}{u} \frac{\text{trace}\{[U'\Sigma U]^{-1}\}}{\text{trace}\{[U'U]^{-1}\}}.
\]

We consider the following two special cases of the structure of $\Sigma$.

(I) Intraclass covariance structure, $U = J_{(p \times 1)}$.

In this case we have
\[
\Sigma = \sigma_0^2 [1 + \rho (J-I)],
\]
\[
\Sigma^* = \sigma_0^2 [1 + r_1 \rho (J-I)], -\frac{1}{p-1} < \rho < 1, \ 0 < r_1 < 1,
\]
\[
\text{trace}\{[U'\Sigma U]^{-1}\} = \frac{1}{\rho \sigma_0^2} \frac{1}{[1 + (p-1)\rho]} ,
\]
and
\[
\text{trace}\{[U'\Sigma^* U]^{-1}\} = \frac{1}{\rho \sigma_0^2} \frac{1}{[1 + (p-1)r_1\rho]} .
\]

Thus the asymptotic relative efficiency is given by
\[
(3.4.19) \ \text{ARE}(\rho) = \frac{k_0}{u} \frac{\text{trace}\{[U'\Sigma U]^{-1}\}}{\text{trace}\{[U'U]^{-1}\}} = \frac{k_0}{u} \frac{[1 + (\rho-1)\rho]}{[1 + (\rho-1)r_1\rho]} ,
\]
which is a continuous monotonic increasing function of $\rho$, $-\frac{1}{p-1} \leq \rho \leq 1$.

At $\rho = -\frac{1}{p-1}$ we have the lower bound of the function $\text{ARE}(-\frac{1}{p-1}) = 0$.

That is, as $\rho$ approaches $-\frac{1}{p-1}, V(C\hat{\beta}^CM'_{(p \times 1)})$ approaches 0. This makes the IM model less desirable. When $\rho = 0$, there is no correlation among all $p$ responses and $\text{ARE}(0) = \frac{k_0}{u} < 1$. At $\rho = 1$, the upper bound of the function is $\text{ARE}(1) = \frac{k_0}{u} \frac{p}{1 + (p-1)r_1}$.

From (3.3.2) we have
\[
u = \sum_{j=1}^{p} \ell_j \binom{p}{j}, \ k_0 = \sum_{j=1}^{p} \ell_j \binom{p-1}{j-1}, \text{ and}
\]
\[ k_1 = \sum_{j=2}^{P} \ell_j (P-2). \]

Together with the identity \( \binom{P-1}{j-1} = \binom{P}{j} \), we then have

\[
\text{ARE}(1) = \frac{k_0}{u} \frac{P}{1+(p-1)r_{11}} \]

\[
= \frac{\sum_{j=1}^{P} \ell_j (P-1)}{\sum_{j=1}^{P} \ell_j (P)} \frac{P}{1+(p-1) \sum_{j=2}^{P} \ell_j (P-2)} \]

\[
= \frac{p \left[ \sum_{j=1}^{P} \ell_j (P) \right] \left[ \sum_{j=1}^{P} \ell_j (P) \right]}{\left[ \sum_{j=1}^{P} \ell_j (P) \right] \left[ \sum_{j=1}^{P} \ell_j (P) \right]^2} \]

\[
= \frac{p \left[ \sum_{j=1}^{P} \ell_j (P) \right]^2}{\left[ \sum_{j=1}^{P} \ell_j (P) \right]^2}.
\]

But

\[
\left[ \sum_{j=1}^{P} \ell_j (P) j \right]^2 = \sum_{j_1, j_2} \left[ \ell_{j_1} (P) \ell_{j_2} (P) j_1 j_2 \right]
\]

\[
= \sum_{j} \left[ \ell_{j} (P) j \right]^2 + \sum_{j_1 \neq j_2} \left[ \ell_{j_1} (P) \ell_{j_2} (P) j_1 j_2 \right]
\]

\[
\leq \sum_{j} \left[ \ell_{j} (P) j \right]^2 + \sum_{j_1 \neq j_2} \left[ \ell_{j_1} (P) \ell_{j_2} (P) \right] j_1^2 j_2^2
\]

\[
= \sum_{j} \left[ \ell_{j} (P) j \right]^2 + \sum_{j_1 \neq j_2} \left[ \ell_{j_1} (P) \ell_{j_2} (P) \right] j_1^2
\]

\[
= \sum_{j_1, j_2} \left[ \ell_{j_1} (P) \ell_{j_2} (P) \right] j_1^2
\]

\[
= \left[ \sum_{j} \ell_{j} (P) \right] \left[ \sum_{j} \ell_{j} (P) j^2 \right].
\]
so,
\[
\frac{\left( \sum j \ell_j(P_j) \right)^2}{\left( \sum j \ell_j(P_j) \right)^2 \sum j \ell_j(P_j)} \leq 1,
\]
and
\[
(3.4.20) \quad \text{ARE}(1) = \frac{k_0}{u} \frac{p}{1+(p-1)r_1} \leq 1.
\]

**Theorem 3.4.2.** Assume intraclass covariance structure. Consider the hypothesis \( H_0: \Sigma J(p \times 1) = 0 \) against the alternative \( H_a: \Sigma J = a_0 J(g \times 1) \), where \( a_0 \neq 0 \). Then the asymptotic relative efficiency between the IM model and its corresponding CM model is given by ARE in (3.4.19), where ARE is an increasing function of \( \rho \) and has value less than one for \(-1/(p-1) < \rho < 1\).

\[
\text{ARE}(-\frac{1}{p-1}) = 0, \quad \text{ARE}(0) = \frac{k_0}{u}, \quad \text{ARE}(1) = \frac{k_0}{u} \frac{p}{1+(p-1)r_1} \leq 1.
\]

(II) Autocorrelation when \( p=3, \ U = J(3 \times 1) \).

In this case, we have
\[
\Sigma = \sigma_0^2 \begin{pmatrix}
1 & \rho & \rho^2 \\
1 & \rho \\
1 & 1
\end{pmatrix},
\]
and \( \Sigma^* = \sigma_0^2 \begin{pmatrix}
1 & r_1 \rho & r_1 \rho^2 \\
1 & r_1 \rho \\
1 & 1
\end{pmatrix}, \ -1 < \rho < 1, \ 0 < r_1 < 1.
\]

Furthermore,
\[
\text{trace}([U^t \Sigma U]^{-1}) = [J(1 \times 3 \Sigma J(3 \times 1)]^{-1} = \frac{1}{\sigma_0^2} \frac{1}{3 + 4 \rho + 2 \rho^2},
\]
and
\[
\text{trace}([U^t \Sigma^* U]^{-1}) = [J(1 \times 3 \Sigma^* J(3 \times 1)]^{-1} = \frac{1}{\sigma_0^2} \frac{1}{3 + 4 r_1 \rho + 2 r_1 \rho^2}.
\]
Thus we have

$$\text{ARE}(\rho) = \frac{k_0}{u} \frac{\text{trace}([U'SU]^{-1})}{\text{trace}([U'SU]^{-1})} = \frac{k_0}{u} \frac{3+4\rho+2\rho^2}{3+4r_1^2+2r_1^2},$$

which is an increasing function of $\rho$, $-1 < \rho < 1$.

Theorem 3.4.3. Assume autocorrelation when $p=3$. Consider the hypothesis $H_0: CBJ(p \times 1) = 0$ against the alternative $H_a: CBJ = a_0 J(g \times 1)$, where $a_0 \neq 0$. Then the asymptotic relative efficiency between the IM model and its corresponding CM model is given by $\text{ARE}$ in (3.4.21). $\text{ARE}$ is an increasing function of $\rho$ and has value less than one for $-1 < \rho < 1$.

$$\text{ARE}(-1) = \frac{k_0}{u} \frac{1}{3-2r_1}, \quad \text{ARE}(0) = \frac{k_0}{u}, \quad \text{ARE}(1) = \frac{k_0}{u} \frac{3}{1+2r_1} \leq 1.$$

3.4.3 Cost

Following the comparison of the asymptotic relative efficiency between the IM model and its corresponding CM model, we now consider the cost factors involved in the two models.

Assuming that in the IM model, $q$ individual units were used on each of the $u$ disjoint sets; then we have a total of $uq$ experimental units. Letting $p$ be the number of complete responses, then the total number of response measurements in the experiment is

$$N_1 = q \sum_{j=1}^{p} \ell_j(p_j) = qp \sum_{j=1}^{p} \ell_j(p_j-1) = qpk_0.$$

For the corresponding CM model, we have the same total number of experimental units, $uq$. With $p$ responses for each unit, the total number of response measurements is

$$N_2 = uqp.$$
Let $c_0$ be the participant cost for each unit in the experiment and $c_j$ be the additional cost per measurement for each unit on which exactly $j$ responses were measured, $j=1,\ldots,p$. Then for $n_1$ repetitions of the IM design, the total cost is

$$T_1 = n_1\{uqc_0 + q \sum_{j=1}^{p} j \ell_j^{(p)}c_j\}.$$  

And for $n_2$ repetitions of the corresponding CM model, the total cost is

$$T_2 = n_2\{uqc_0 + uqpc_p\}.$$  

Assuming the cost per measurement increases as the number of responses measured on the same unit increases, then we have

$$0 < c_1 \leq c_2 \leq \ldots \leq c_p \quad \text{and} \quad c_1 \leq c_A \leq c_p,$$

where

$$c_A = \frac{\sum_{j=1}^{p} j \ell_j^{(p)}c_j}{\sum_{j=1}^{p} j \ell_j^{(p)}} = \frac{\sum_{j=1}^{p} j \ell_j^{(p)}c_j}{pk_0},$$

is the average cost per measurement for the IM model.

Assume also that in comparing with the cost per measurement the unit participant cost is negligible. Then, practically,

$$T_1 = n_1q \sum_{j=1}^{p} j \ell_j^{(p)}c_j = n_1q \ c_A \ pk_0,$$

and

$$T_2 = n_2q uqpc_p.$$  

We now examine the asymptotic relative cost (ARC) for the IM model and its corresponding CM model, where ARC is the relative cost of the two models when both have asymptotically the same power. The relative cost is
When both models have asymptotically the same power, \( \frac{n_2}{n_1} \) is the asymptotic relative efficiency. Thus

\[
(3.4.28) \quad \text{ARC} = \frac{T_1}{T_2} = \frac{1}{\text{ARE}} \frac{c_A k_0}{c_p u}.
\]

As in Section 3.4.2, we consider the null hypothesis \( H_0 : \beta J = 0 \) against the alternative \( H_a : \beta J = a_0 J \), where \( a_0 \) is a nonzero constant. Following Section 3.4.2, we have the following special cases:

(I) Intraclass covariance structure

In this case, \( \text{ARE}(\rho) = \frac{k_0}{u} \frac{1+(p-1)\rho}{[1+(p-1)r_1\rho]} \).

The asymptotic relative cost is then given by

\[
(3.4.29) \quad \text{ARC}(\rho) = \frac{1}{\text{ARE}(\rho)} \frac{c_A k_0}{c_p u}
= \frac{1+(p-1)r_1\rho}{1+(p-1)\rho} \frac{c_A}{c_p},
\]

a continuous decreasing function of \( \rho, -1/(p-1) < \rho < 1 \).

Thus, \( \text{ARC}(\frac{1}{p-1}) = +\infty \), \( \text{ARC}(0) = \frac{c_A}{c_p} \leq 1 \), and \( \text{ARC}(1) = \frac{1+(p-1)r_1\rho}{p} \frac{c_A}{c_p} \leq \frac{c_A}{c_p} \leq 1 \).

Thus when \( \rho > 0 \), the cost for the IM model is less than the cost for the corresponding CM model.

(II) Autocorrelation when \( p=3 \)

In this case, \( \text{ARE}(\rho) = \frac{k_0}{u} \frac{3+4\rho+2\rho^2}{3+4r_1\rho+2r_1\rho^2} \). The asymptotic relative cost is
a continuous decreasing function of \( \rho, -1 < \rho < 1 \). It follows that

\[
ARC(-1) = (3-2r_1) \frac{c_A}{c_p} > \frac{c_A}{c_p}, \quad ARC(0) = \frac{c_A}{c_p} \leq 1,
\]

and

\[
ARC(1) = \frac{1+2r_1}{3} \frac{c_A}{c_p} < \frac{c_A}{c_p} \leq 1.
\]

As in case (I), when \( \rho > 0 \), the cost for the IM model is less than the cost for the corresponding CM model.

3.4.4 Minimum risk

In parametric estimation, the loss function often can be used as a measure of the discrepancy between the parameter value and its estimate value. Consider a p-variate multinormal population with unknown mean \( \mu \) and covariance matrix \( \Sigma \). Having recorded a sample \( \{y_1, \ldots, y_n\} \), the loss function in estimating \( \mu \) by \( \overline{y} \) can be defined as

\[
L_n = (\overline{y} - \mu)' A(\overline{y} - \mu),
\]

where

\[
\overline{y} = \frac{1}{n} \sum_{j=1}^{n} y_j,
\]

and \( A \) is a known \( p \times p \) positive definite matrix.

If we also consider the cost factor involved, then the loss incurred in estimating \( \mu \) by \( \overline{y} \) is given by

\[
L_n = (\overline{y} - \mu)' A(\overline{y} - \mu) + c_n,
\]

where \( c(> 0) \) is the known cost per unit sample.

Thus the risk function is
(3.4.33) \[ R_n = E(L_n) = n^{-1} \text{trace}(AE) + cn. \]

For a given $\Sigma$, $R_n$ attains a minimum at $n^*$, where

(3.4.34) \[ n^* = \left\{ \frac{1}{c} \text{trace}(AE) \right\}^{\frac{1}{2}}, \]

and \[ R^*_n = 2cn^*. \]

Extending the above derivation to the general case of estimating a matrix of unknown parameters, $\beta_{(r \times p)}$, the loss function in estimating $\beta$ by $\hat{\beta}$ is

(3.4.35) \[ L_n = \text{trace}\{(\hat{\beta} - \beta)'A(\hat{\beta} - \beta)\} + c^*(n), \]

where $c^*(n)$ is the corresponding total cost in the experiment.

In IM design, the estimate $\hat{\beta}_{IM}$ given by (3.4.11) is an unbiased estimate of $\beta$, and its variance is

\[ V(\hat{\beta}_{IM}) = \frac{2}{n_1k_0} (F'F_1 + F_2'F_2)^{-1} \otimes \Sigma^* \]

Consider the total cost $T_1$ given by (3.4.28), $T_1 = n_1q \, c_A \, pk_0$. Then we have the risk

(3.4.36) \[ R_{IM}(n_1) = E\{\text{trace}[(\hat{\beta}_{IM} - \beta)'A(\hat{\beta}_{IM} - \beta)]\} + n_1q \, c_A \, pk_0 \]

\[ = E\{\text{trace}[A(\hat{\beta}_{IM} - \beta)'(\hat{\beta}_{IM} - \beta)]\} + n_1q \, c_A \, pk_0 \]

\[ = \text{trace}(A \frac{2}{n_1k_0} \text{trace}[(F_1'F_1 + F_2'F_2)^{-1}]\Sigma^*) + n_1q \, c_A \, pk_0 \]

\[ = \frac{h_0}{n_1k_0} \text{trace}(AE^*) + n_1q \, c_A \, pk_0, \]

where $h_0 = 2 \, \text{trace} [(F_1'F_1 + F_2'F_2)^{-1}]$.

The above risk $R_{IM}(n_1)$ attains a minimum at
Note that $R_{IM}(\hat{n}_1) = 2q c_A p k_0 \hat{n}_1 = 2\left(\frac{h_0}{q c_A p} \text{trace}(A\Sigma^*)\right)^{\frac{1}{2}}$.

Similarly, for the corresponding CM design, the estimate $\hat{\beta}_{CM}$ given by (3.4.15) is also an unbiased estimate of $\beta$, and its variance is

$$V(\hat{\beta}_{CM}) = \frac{2}{n_2 u} (F_1'F_1 + F_2'F_2)^{-1} \otimes \Sigma.$$ 

The total cost $T_2$ is given by (3.4.26), $T_2 = n_2 q u p c_p$. Thus, the risk is

$$R_{CM}(n_2) = E\{\text{trace}[(\hat{\beta}_{CM} - \beta)' A (\hat{\beta}_{CM} - \beta)]\} + n_2 q u p c_p$$

$$= \text{trace}\{A\Sigma[(\hat{\beta}_{IM} - \beta)'(\hat{\beta}_{CM} - \beta)]\} + n_2 q u p c_p$$

$$= \frac{h_0}{n_2 u} \text{trace}(A\Sigma) + n_2 q u p c_p.$$ 

The risk $R_{CM}(n_2)$ attains a minimum at

$$(3.4.39) \quad \hat{n}_2 = \left\{ \frac{h_0}{u^2 q p c_p} \text{trace}(A\Sigma) \right\}^{\frac{1}{2}} = \frac{1}{4} \left\{ \frac{h_0}{q p c_p} \text{trace}(A\Sigma) \right\}^{\frac{1}{2}},$$

$$R_{CM}(\hat{n}_2) = 2q u p c_p \hat{n}_2 = 2\left(\frac{h_0}{q p c_p} \text{trace}(A\Sigma)\right)^{\frac{1}{2}}.$$ 

Since both designs have $q$ experimental units in each replication, the numbers of repetitions, $n_1$ and $n_2$, can then be used as a comparison of the relative sample size.

The relative risk can be evaluated by comparing the optimal sample size together with the corresponding relative minimum risk. Assuming $c_1 = c_2 = \ldots = c_p = c_A$, the relative optimal sample size is

$$R_{CM}(\hat{n}_1, \hat{n}_2) = \frac{\hat{n}_1}{\hat{n}_2} = \frac{u}{k_0} \left\{ \frac{\text{trace}(A\Sigma^*)}{\text{trace}(A\Sigma)} \right\}^{\frac{1}{2}},$$

with the relative minimum risk
Now we assume that the matrix $A$ has an intraclass covariance structure,

$$A = [I + s(J-I)], \quad \frac{1}{p-1} < s < 1,$$

and we consider the following special cases based on the structure of $\Sigma$.

(I) Intraclass covariance structure

In this case

$$\Sigma = \sigma_0^2 [1 + \rho(J-I)],$$

$$\Sigma^* = \sigma_0^2 [1 + r_1 \rho(J-I)], \quad \frac{1}{p-1} < \rho < 1,$$

$$\text{trace}(A\Sigma) = \sigma_0^2 [1 + (p-1) s \rho],$$

and

$$\text{trace}(A\Sigma^*) = \sigma_0^2 [1 + (p-1) r_1 s \rho].$$

We then have

\[3.4.42\]

$$\begin{aligned}
\hat{R}_N(n_1, n_2) &= \frac{\frac{1+(p-1)r_1 s \rho}{1+(p-1)s \rho}}{k_0}, \\
\hat{R}_R(n_1, n_2) &= \frac{1+(p-1)r_1 s \rho}{1+(p-1)s \rho}. \\
\end{aligned}$$

Both $R_N$ and $R_R$ are decreasing functions of $s \rho$, $\frac{1}{p-1} < s \rho < 1$, where

$$R_R(s \rho = -\frac{1}{p-1}) = +\infty, \quad R_N(s \rho = -\frac{1}{p-1}) = +\infty,$$

$$R_R(s \rho = 0) = 1, \quad R_N(s \rho = 0) = \frac{u}{k_0} > 1,$$

$$R_R(s \rho = 1) = \left\{ \frac{1+(p-1)r_1 s \rho}{p} \right\} < 1,$$

and

$$R_N(s \rho = 1) = \frac{u}{k_0} \left\{ \frac{1+(p-1)r_1 s \rho}{p} \right\}.$$

From (3.4.20),

$$\frac{k_0}{u} \frac{p}{1+(p-1)r_1} \leq 1.$$
Thus
\[ RN(s\rho = 1) = \frac{1+(p-1)r_{1}}{k_0} \left( \frac{1}{p} \right) \geq \frac{1}{k_0} \frac{1+(p-1)r_{1}}{p} \geq 1. \]

That is, when \( s\rho < 0 \), we have \( RR > 1 \), \( RN > 1 \). The IM model has a large minimum risk with a larger sample size.

When \( s\rho = 0 \), which includes the case \( \rho = 0 \), i.e. no correlation among all \( p \) responses, we have \( RR = 1 \), \( RN > 1 \). Both models have the same minimum risk, but a larger sample size is required for the IM model.

When \( s\rho > 0 \), we have \( RR < 1 \), \( RN > 1 \). The IM model has a smaller minimum risk, but needs a larger sample size.

(II) Autocorrelation when \( p=3 \)

In this case
\[ \Sigma = \begin{pmatrix} 1 & \rho & \rho^2 \\ 1 & \rho \\ \rho^2 & 1 \end{pmatrix}, \]

\[ \Sigma^* = \begin{pmatrix} 1 & r_1\rho & r_1\rho^2 \\ 1 & r_1\rho \\ r_1\rho & 1 \end{pmatrix}, \quad -1 < \rho < 1, \]

\[ \text{trace}(\Sigma) = \rho \sigma_0^2 \left[ 1 + \sigma_0^2 \left( r_1 \rho + r_1 \rho^2 \right) \right], \]

and
\[ \text{trace}(\Sigma^*) = \rho \sigma_0^2 \left[ 1 + r_1 \sigma_0 + r_1 \rho^2 \right]. \]

We then have
\[ (3.4.43) \quad RN(\hat{n}_1, \hat{n}_2) = \frac{1}{k_0} \left( \frac{1+r_1 s\rho + r_1 s\rho^2}{1+s\rho+s\rho^2} \right), \quad \text{and} \]

\[ RR(\hat{n}_1, \hat{n}_2) = \left( \frac{1+r_1 s\rho + r_1 s\rho^2}{1+s\rho+s\rho^2} \right). \]

From
\[ \frac{\partial RR}{\partial \rho} = -\frac{1}{2} \left( \frac{1+r_1 s\rho + r_1 s\rho^2}{1+s\rho+s\rho^2} \right)^{-\frac{3}{2}} \left( 1+s\rho+s\rho^2 \right)^{-2} \left( 1-r_1 \right) s \left( 1+2\rho \right), \]

we know that \( RR \) is an increasing function of \( \rho \) if \( s(1+2\rho) < 0 \) and is a
decreasing function of \( p \) if \( s(1+2p) > 0 \).

\[
RR(p = -1) = 1, \quad RN(p = -1) = \frac{u}{k_0},
\]

\[
RR(p = - \frac{1}{2}) = \left( \frac{4-r_1s}{4-s} \right)^{\frac{1}{2}}, \quad RN(p = - \frac{1}{2}) = \frac{u}{k_0} \left( \frac{4-r_1s}{4-s} \right)^{\frac{1}{2}},
\]

\[
RR(p = 0) = 1, \quad RN(p = 0) = \frac{u}{k_0},
\]

\[
RR(p = 1) = \left( \frac{1+2r_1s}{1+2s} \right)^{\frac{1}{2}}, \quad \text{and} \quad RN(p = 1) = \frac{u}{k_0} \left( \frac{1+2r_1s}{1+2s} \right)^{\frac{1}{2}}.
\]

When \( s > 0 \), \( RR \) is an increasing function of \( p \) for \( p < - \frac{1}{2} \), and is a decreasing function of \( p \) for \( p > - \frac{1}{2} \). Thus \( RR \) attains its minimum at \( p = 1 \):

\[
RR(p = 1) = \left( \frac{1+2r_1s}{1+2s} \right)^{\frac{1}{2}} < 1; \quad \text{and}
\]

\[
RN(p = 1) = \frac{u}{k_0} \left( \frac{1+2r_1s}{1+2s} \right)^{\frac{1}{2}} > \frac{u}{k_0} \left( \frac{1+2r_1s}{1+2s} \right)^{\frac{1}{2}}.
\]

We see that when \( s > 0 \), the IM model has a smaller minimum risk only for large values of \( p \) (near one), but a large sample size is needed.

When \(- \frac{1}{2} < s < 0\), \( RR \) attains its minimum at \( p = - \frac{1}{2} \):

\[
RR(p = - \frac{1}{2}) = \left( \frac{4-r_1s}{4-s} \right)^{\frac{1}{2}} < 1; \quad \text{and}
\]

\[
RN(p = - \frac{1}{2}) = \frac{u}{k_0} \left( \frac{4-r_1s}{4-s} \right)^{\frac{1}{2}} > \frac{u}{k_0} \left( \frac{4-r_1s}{4-s} \right)^{\frac{1}{2}}.
\]

\[
= \left( \frac{u}{k_0} \frac{1+5r_1s}{3} \right)^{\frac{1}{2}} > \left( \frac{u}{k_0} \frac{1+2r_1s}{3} \right)^{\frac{1}{2}} > 1.
\]

We see that when \( s < 0 \), the IM model has a smaller minimum risk only for \( p \)
near \( \frac{1}{2} \), but a larger sample size is needed.

When \( p = 0 \), both models have the same minimum risk, but a larger sample size is required for the IM model.

3.5 Numerical example

Fisher's IRIS dataset consists of random samples of 50 flowers from the iris species: virginica, versicolor, and setosa. For numerical illustration of the IM procedure, we use the responses of three measurements: petal length (V1), sepal length (V2), and sepal width (V3). Assume the responses for each flower are multivariate normally distributed with a common covariance matrix. We will test the hypothesis of a common mean vector for the three responses in the three species populations.

With suitable arrangement of the data, we divide the data into \( u = 10 \) disjoint sets where each set \( S_j \) consists of samples of 5 flowers from each species. We then assume the following response-wise design:

<table>
<thead>
<tr>
<th>Set:</th>
<th>( S_1 ), ( S_6 )</th>
<th>( S_2 ), ( S_7 )</th>
<th>( S_3 ), ( S_8 )</th>
<th>( S_4 ), ( S_9 )</th>
<th>( S_5 ), ( S_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses:</td>
<td>2,3</td>
<td>1,3</td>
<td>1,2</td>
<td>1,2,3</td>
<td>1,2,3</td>
</tr>
</tbody>
</table>

That is, each response is measured on \( k_0 = 8 \) disjoint sets, each pair of responses is measured on \( k_1 = 6 \) disjoint sets, and \( r_1 = \frac{k_1}{k_0} = 0.75 \).

In each disjoint set, we have a standard MGLM.

\[
(3.5.1) \quad E(Y_{ij}) = X_{ij}\beta_j, \quad \text{and} \\
V(Y_{ij}) = I_{15} \otimes \Sigma_j, \quad j=1,\ldots,10,
\]

where

\[
X_{ij}(15 \times 3) = I_3 \otimes J_{(5 \times 1)}', \\
\Sigma_j(p_j \times p_j) = B_j'\Sigma B_j',
\]
Following the procedure derived in Section 3.2, we first construct

\[ Q_j(3 \times p_j) = M_j^T Y_j, \quad j = 1, \ldots, 10, \]

where

\[ M_j(15 \times 3) = \frac{1}{\sqrt{5}} X_j = \frac{1}{\sqrt{5}} I_3 \otimes J(5 \times 1). \]

Then we have

\[ (3.5.2) \quad E(Q_j) = \sqrt{5} \beta B_j, \quad \text{and} \]
\[ V(Q_j) = I_3 \otimes \Sigma_j, \quad j = 1, \ldots, 10. \]

Thus, the above IM design satisfies the set of conditions (3.3.1).

Letting \( Q = (Q_1, \ldots, Q_{10}) \), we have

\[ E(Q) = \sqrt{5} (\beta, \beta)L, \quad \text{where} \quad L = I_2 \otimes (B_1, \ldots, B_5). \]

Now

\[ LL' = \frac{k_0}{2} I_6 = 4 I_6, \quad \text{and} \]
\[ Z^{*}(3 \times 6) = (Z_1, Z_2) = QL'(LL')^{-1} \]
\[ = \begin{bmatrix}
\end{bmatrix} \]

Letting \( Z(6 \times 3) = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \), we have a standard MGLM with

\[ (3.5.3) \quad E(Z) = \sqrt{5} \begin{bmatrix} I_3 \\ I_3 \end{bmatrix} \beta, \]
and \[ V(Z) = \frac{2}{k_0} I_6 \otimes \Sigma^* = \frac{1}{4} I_6 \otimes \Sigma^*, \]

where \[ \Sigma^* = \text{Diag}(\Sigma) + (0.75) \text{Off diag}(\Sigma). \]

Thus it follows that

\[
\beta_{IM} = \begin{pmatrix}
1.4600 & 5.0050 & 3.4400 \\
4.2025 & 5.9975 & 2.7800 \\
5.5525 & 6.5400 & 2.9725
\end{pmatrix}, \text{ and}
\]

\[
V(\beta_{IM}) = \frac{2}{k_0} (21_3)^{-1} \otimes \Sigma^* = \frac{1}{40} I_3 \otimes \Sigma^*. 
\]

In testing the hypothesis of species effect,

\[ H_0: \theta = \mathbf{0}, \text{ where } C = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}, \text{ we have}
\]

\[
\hat{\theta} = \mathbf{C} \beta_{IM} = \begin{pmatrix} -2.7425 & -0.9925 & 0.6600 \\ -4.0925 & -1.5350 & 0.4675 \end{pmatrix}.
\]

The matrix of sums of squares and products among species is

\[
S_{II} = \begin{pmatrix}
86.9745 & 32.4543 & -11.5447 \\
12.1186 & -4.2274 & -2.3040 \\
2.3040 & -4.2274 & -6.3040
\end{pmatrix}.
\]

The error matrix is

\[
S_E = \begin{pmatrix}
0.089875 & 0.0491875 & 0.0118125 \\
0.1105630 & 0.0698750 & 0.0513125
\end{pmatrix}.
\]

The Wilks LR statistic is

\[
\Lambda_1 = \frac{|S_E|}{|S_{II} + S_E|} = 0.00002579.
\]

To evaluate the significance of the statistic \( \Lambda_1 \), we use the equivalent F statistic
\[ F_1 = \frac{1 - \Lambda_1^{1/2}}{3\Lambda_1^{1/2}}. \]

Under \( H_0 \), \( F_1 \) has an \( F \) distribution with 6 and 3 d.f. When
\[
\Lambda_1 = 0.000002579, \quad F_1 = 207.231 \quad \text{and the corresponding p-value = .0048.}
\]
Thus at a level .05, we reject the hypothesis of a common mean vector for the three species of iris.

With all 3 responses measured on each flower, we have a corresponding CM model:

\[ (3.5.7) \quad E(Y) = \mathbf{x}\beta, \quad \text{and} \]
\[ V(Y) = I_{30} \otimes \Sigma, \]
where
\[
Y_{(30 \times 3)} = \begin{bmatrix} Q_1 \\ \vdots \\ Q_{10} \end{bmatrix}, \quad X_{(30 \times 3)} = \sqrt{5} [I_3 \otimes J_{(10 \times 1)}].
\]

The estimates of interest are:

\[ (3.5.8) \quad \hat{\beta}_{CM} = \begin{pmatrix} 1.462 & 5.006 & 3.428 \\ 4.260 & 5.936 & 2.770 \\ 5.552 & 6.588 & 2.974 \end{pmatrix}, \quad \text{and} \]
\[ V(\hat{\beta}_{CM}) = \frac{2}{u} (2I_3)^{-1} \otimes \Sigma = \frac{1}{50} I_3 \otimes \Sigma. \]

In testing the same null hypothesis of species effect, we have

\[ (3.5.9) \quad \hat{\theta} = \begin{pmatrix} -2.798 & -0.930 & 0.658 \\ -4.090 & -1.582 & 0.454 \end{pmatrix}, \]
\[ \quad S_H = \begin{pmatrix} 437.103 & 165.2480 & -57.2396 \\ 63.2121 & -19.9527 & 11.3449 \end{pmatrix}, \]
\[ \quad S_E = \begin{pmatrix} 4.9386 & 4.4986 & 0.8028 \\ 7.6002 & 2.6960 & 2.4300 \end{pmatrix}. \]
The Wilks LR statistic is

\[(3.5.10) \quad \Lambda_2 = \frac{|S_E|}{|S_H + S_E|} = 0.002681.\]

The corresponding F statistic is

\[F_2 = \frac{25}{3} \frac{1 - \Lambda_2^{1/2}}{\Lambda_2^{1/2}} = 152.623.\]

Under H_0, the statistic F_2 has an F distribution with 6 and 50 d.f. The corresponding p-value is less than .0001, and we reject the hypothesis H_0 at a level .05.

Assuming intraclass covariance structure for Σ and ρ = 0.6, we have the following comparisons:

1. Generalized variance

The comparison is given by R in (3.4.8). Here

\[R = \left(\frac{4}{k_0}\right)^p \frac{(1-r_1^2)\rho^{p-1}}{(1-\rho)^{p-1}} \frac{[1+(p-1)r_1^2\rho]}{[1+(p-1)\rho]}\]

\[= \left(\frac{10}{8}\right)^3 \frac{[1-(0.75)(0.6)]^2}{[1-0.6]^2} \frac{1+2(0.95)(0.6)}{1+2(0.6)}\]

\[= 3.8704 > 1.\]

The generalized variance for \(\hat{\beta}_{IM}\) is larger than the generalized variance for \(\hat{\beta}_{CM}\).

2. Asymptotic relative efficiency for H_0: CβJ = 0 against

H_a: CβJ = a_0J, \quad a_0 \neq 0.

The comparison is given by ARE in (3.4.19) where
\[ ARE(0.6) = \frac{k_0}{u} \frac{[1+(p-1)\rho]}{[1+(p-1)r_1\rho]} = \frac{8}{10} \frac{[1+2(0.6)]}{[1+2(0.95)(0.6)]]} = 0.9263 < 1. \]

(3) Cost with \( C_1 = C_2 = \ldots = C_p \).

We estimate the cost comparison given by ARC in (3.4.29) to be

\[ ARC(0.6) = \left( \frac{1+(p-1)r_1\rho}{1+(p-1)\rho} \right) \frac{C_A}{C_p} = \frac{1+2(0.95)(0.6)}{1+2(0.6)} = 0.8636 < 1. \]

The cost for the IM model is less than the cost for the corresponding CM model.

(4) Minimum risk with \( s = 0.6 \).

From (3.4.42) we have

\[ RR = \left\{ \frac{1+(p-1)\rho}{1+(p-1)\rho} \right\}^{1/2} = \left\{ \frac{1+2(0.95)(0.6)}{1+2(0.6)} \right\}^{1/2} = 0.9462 < 1, \]

\[ RN = \frac{u}{k_0} \left\{ \frac{1+(p-1)\rho}{1+(p-1)\rho} \right\}^{1/2} = 1.1828 > 1. \]

The IM model has a smaller minimum risk (\( RR < 1 \)), but a larger sample size is needed (\( RN > 1 \)).
4.1 Introduction

In multiresponse designs, if the observations arise in a structured form, such as with the same p responses at each of t different time points, then the resulting covariance matrix may contain special patterns. In this chapter, we study hypothesis testing of some particular patterns of covariance structures in the model.

A standard MGLM model with tp responses is given by

\[(4.1.1)\quad H_0: E(Y_0) = X\beta \text{ and } V(Y_0) = I_N \otimes \Sigma_0,\]

where \(Y_0(N \times tp)\) is an observed data matrix, 
\(X(N \times q)\) is a design matrix of rank q, 
\(\beta(q \times tp)\) is a matrix of unknown parameters, and 
\(\Sigma_0(tp \times tp) = (\Sigma_{i,j}),\) is an unknown positive definite covariance matrix with \(\Sigma_{i,j}(p \times p), i,j = 1,2, \ldots, t.\)

The null hypotheses considered in this chapter are

\(H_1: \Sigma_0 = I \otimes \Sigma_1 + (J-I) \otimes \Sigma_2,\)
\(H_2: \omega \Sigma_1 \text{ in } H_1,\)
\(H_3: \Sigma_1 = \sigma_1[I + \rho_1(J-I)], \Sigma_2 = \sigma_2[I + \rho_2(J-I)] \text{ in } H_1 \text{, and}\)
\(H_4: \rho_2 = 1 \text{ in } H_3.\)

In each case, the LR test statistic is derived. For the test statistic expressed in closed-form, its moments are used to find the
asymptotic distributions. For the test statistic without a closed-form expression, the Wald test statistic is used.

The asymptotic non-null distribution of a test statistic depends in general upon the type of alternative being considered. For the statistic with closed-form expression, we study asymptotic non-null distribution of the statistic under a fixed alternative. As the number of observations approaches infinity, the power of the test will approach one against any fixed alternative. Thus, for each test statistic, we also consider the asymptotic non-null distributions under a sequence of local alternatives.

The log likelihood function under $H_0$ is

$$L(\beta, \Sigma_0) = \frac{1}{2} N p \ln(2\pi) - \frac{1}{2} N \ln(|\Sigma_0|) - \frac{1}{2} \text{tr}(\Sigma_0^{-1}(Y-X\beta)'(Y-X\beta)).$$

(4.1.2)

The MLE of $\beta$ does not depend on the pattern structure of $\Sigma_0$ and is

$$\hat{\beta} = (X'X)^{-1}X'Y.$$  

(4.1.3)

With the MLE of $\Sigma_0$,

$$\hat{\Sigma}_0 = \frac{1}{N} S,$$  

(4.1.4)

where

$$S = Y'[I-X(X'X)^{-1}X']Y,$$

the resulting maximum log-likelihood under $H_0$ is then

$$L_0 = L(\hat{\beta}, \hat{\Sigma}_0) = \frac{1}{2} N p \ln(2\pi) + \frac{1}{2} N p \ln(N) - \frac{1}{2} N \ln(|S|) - \frac{1}{2} N p.$$  

(4.1.5)

Asymptotic distributions of certain LR statistics satisfying a set of specific conditions are contained in Section 4.1.1. Asymptotic distributions of the Wald statistic are contained in Section 4.1.2.

The hypothesis $H_1$ of block intraclass covariance structure in a given standard MGLM model is examined in Section 4.2. With suitable transformations on the parameter space and the observed data, the MLE's
of $\Sigma_1$ and $\Sigma_2$ can be solved directly from the system of partial equations. The LR statistic for testing $H_1$ can then be derived. An asymptotic expansion for the distribution of the LR statistic under $H_1$ is contained in Section 4.2.1 and asymptotic non-null distribution of the LR statistic when $t = 2$ under both the fixed alternative and a sequence of local alternatives are contained in Section 4.2.2.

The hypothesis $H_2: \Sigma_2 = \omega \Sigma_1$ in $H_1$ is examined in Section 4.3. Under $H_2$, the MLE of $\Sigma_1$ can be derived as an explicit function of $\hat{\omega}$, the MLE of $\omega$. But a closed form solution for $\hat{\omega}$ cannot be solved from the system of partial equations. However, by using numerical methods, the unique solution for $\hat{\omega}$ can be estimated. Thus, the LR statistic for testing $H_2$ vs. $H_1$ can then be evaluated.

The above LR statistic does not have a closed-form expression. Its value depends on the numerical estimate of the parameter $\omega$. We can test the hypothesis $H_2$ vs. $H_1$ without using the MLE of the parameters under $H_2$. A test which is asymptotically equivalent to the LR test is the Wald type test. The Wald statistic for testing $H_2$ vs. $H_1$ when $p = 2$ is contained in Section 4.3.1. Asymptotic null distribution and asymptotic non-null distributions under a sequence of local alternatives for the test statistic are contained in Section 4.3.2.

The hypothesis $H_3$ of intraclass covariance structures for the two matrices $\Sigma_1$ and $\Sigma_2$ in $H_1$ is examined in Section 4.4. The MLE's of $\Sigma_1$ and $\Sigma_2$ under $H_3$ can be solved directly from the system of partial equations. The LR statistic for testing $H_3$ vs. $H_1$ can then be derived. An asymptotic expansion for the distribution of the LR statistic under the null hypothesis $H_3$ is contained in Section 4.4.1. Asymptotic non-null distributions of the LR statistic when $p = 2$ under both the fixed
alternative and a sequence of local alternatives are contained in Section 4.4.2.

The hypothesis $H_4$ of equality of elements in the matrix $\Sigma_2$ in $H_3$ is examined in Section 4.5. As in Section 4.4, the MLE's of $\Sigma_1$ and $\Sigma_2$ under $H_4$ can be solved directly from the system of partial equations. The modified LR statistic for testing $H_4$ vs. $H_3$ can then be derived. Asymptotic null and non-null distributions of the LR statistic are contained in Section 4.5.1 and 4.5.2, respectively.

Numerical examples based on a data set from the Lipid Research Clinics Program are given in Section 4.6.

4.1.1 Asymptotic Null Distributions of Certain Likelihood Ratio Statistics

The distribution of a LR statistic can be obtained from its moments. Let $W$ ($0 \leq W \leq 1$) be a random variable whose moments are

\[
E(W^h) = K_0 \left( \frac{b}{n} \sum_{k=1}^{b} y_k^{h} \left( \frac{a}{n} \sum_{j=1}^{a} x_j \right) \right) \frac{a}{n} \sum_{j=1}^{a} \Gamma[x_j^{(1+h)} + e_j],
\]

where

\[
K_0 = \frac{b}{n} \sum_{k=1}^{b} \Gamma[y_k x_k] = \frac{a}{n} \sum_{j=1}^{a} \Gamma[x_j e_j],
\]

\[
\sum_{j=1}^{a} x_j = \sum_{k=1}^{b} y_k.
\]

Define $Z = -2\rho \ln W$, where $0 \leq \rho \leq 1$. Then the log characteristic function of $Z$ is

\[
\ln \phi_2(s) = \sum_{j=1}^{a} \left\{ -\ln \Gamma[\rho x_j g_j e_j] + 2is\rho x_j \ln x_j + \ln \Gamma[\rho x_j (1-2is) + g_j + e_j] \right\}
\]
Assume that \(x_j \) and \(y_k \) are of the same order, i.e., \(O(x_j) = O(y_k) = O(n)\). Also assume that \(\rho \) depends on \(n\) so that \(1-\rho = O(n^{-1})\). Then by using an asymptotic expansion of the log gamma function according to Barnes [see Muirhead (1982), page 305], we have the following result.

**Theorem 4.1.1.** Let \(W\) be a random variable satisfying the conditions specified in (4.1.6). Then the distribution of \(-2\rho \ln W\) can be expanded for large \(n\) as

\[
P(-2\rho \ln W \leq z) = P(\chi^2_F \leq 0(n^{-2})),
\]

where

\[
f = -2[\sum_{j=1}^{a} e_j - \sum_{k=1}^{b} f_k - (a-b)]
\]

\[
\rho = 1 - \frac{1}{f}[\sum_{j=1}^{a} \frac{B_2(e_j)}{x_j} - \sum_{k=1}^{b} \frac{B_2(f_k)}{y_k}],
\]

\(B_2(d) = d^2 - d + \frac{1}{6}\) is the Bernoulli polynomial of degree 2.

4.1.2. Asymptotic Distributions of the Wald Statistic

Let \(y_1, y_2, \ldots, y_N\) be a random sample from a population with density function \(f(y; \theta)\), where \(\theta' = (\theta_1, \ldots, \theta_q)\) is a \(q\)-dimensional parameter. Then the log-likelihood of the \(N\) observations is given by

\[
L(\theta) = \sum_{j=1}^{N} L_j(\theta), \text{ where}
\]

\[
L_j(\theta) = \ln f(y_j, \theta) \text{ is the log-likelihood from the } j^{th} \text{ observation.}
\]

Define

\[
I_0(\theta) = -E[\frac{\partial^2 L_j(\theta)}{\partial \theta \partial \theta'}], \text{ the information matrix in } y_j.
\]
Then under certain regularity conditions, the MLE of $\theta$, $\hat{\theta}_N$, is asymptotically distributed as a Normal variate with mean $\theta$ and variance $\frac{1}{N} I_0^{-1}(\theta)$.

In testing the hypothesis $H: C\theta = \gamma_0$, where $C(r \times q)$ is of rank $r$. The Wald statistic is given by

$$W_N = N(C\hat{\theta}_N - \gamma_0)' [CI_0^{-1}(\theta)C']^{-1}(C\hat{\theta}_N - \gamma_0).$$

Under the null hypothesis $H$, $C\hat{\theta}_N$ is asymptotically distributed as a Normal variate with mean $\gamma_0$ and variance $\frac{1}{N}[CI_0^{-1}(\theta)C']$. Thus, the Wald statistic $W_N$ is asymptotically distributed $\chi^2$ with $r$ d.f. We reject $H$ for large values of $W_N$.

When the hypothesis $H$ is not true, we consider a sequence of local alternatives $H_N: C\theta - \gamma_0 = \frac{1}{\sqrt{N}} \Delta$, where $\Delta$ is a fixed constant not involving $N$. Under the alternative $H_N$, the statistic $W_N$ is asymptotically distributed as a noncentral $\chi^2$ with $r$ d.f. with noncentrality parameter

$$\lambda^2 = \Delta' [CI_0^{-1}(\theta)C']^{-1} \Delta.$$

The above information matrix $I_0(\theta)$ depends on the unknown parameter $\theta$. A consistent estimator for $I_0(\theta)$ can be derived by replacing $\theta$ with its MLE $\hat{\theta}_N$. That is

$$I_0(\hat{\theta}_N) = -\left\{E\left[\frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'}\right]\right\} \theta = \hat{\theta}_N$$

$$= -\frac{1}{N}\left\{E\left[\frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'}\right]\right\} \theta = \hat{\theta}_N$$

The estimate for $I_0^{-1}(\theta)$, $I_0^{-1}(\hat{\theta}_N)$, may not be unbiased. To reduce the possible bias in the estimator, we apply the jackknife method.

For a random sample of size $N$, there are $N$ different subsamples for size $N-1$. The $j$th subsample can be constructed by deleting the $j$th
observation from the original sample. Let \( \hat{\theta}_{N(j)} \) be the consistent estimator for \( \hat{\theta}_N(\theta) \) from the \( j \)th subsample. That is,

\[
(4.1.16) \quad I_0(\hat{\theta}_{N(j)}) = \left\{ \frac{\partial^2 L_j(\theta)}{\partial \theta \partial \theta} \right\} \theta = \hat{\theta}_{N(j)}
\]

\[
= - \frac{1}{N-1} \left\{ \frac{\partial^2 [L(\theta) - L_j(\theta)]}{\partial \theta \partial \theta} \right\} \theta = \hat{\theta}_{N(j)}
\]

where \( \hat{\theta}_{N(j)} \) is the MLE of \( \theta \) from the \( j \)th subsample.

Then the jackknife estimator of \( I^{-1}_0(\theta) \) is given by

\[
(4.1.17) \quad I^{-1}_0(\hat{\theta}) = N I^{-1}_0(\hat{\theta}_N) - (N-1) \frac{1}{N} \sum_{j=1}^{N} I^{-1}_0(\hat{\theta}_{N(j)}).
\]

We replace \( I^{-1}_0(\theta) \) with its jackknife estimator \( I^{-1}_0J(\hat{\theta}) \) in the Wald test procedure above.

4.2. Testing Block Intraclass Covariance Structure for a Covariance Matrix:

H1: \( \Sigma = I \otimes \Sigma_1 + (J-I) \otimes \Sigma_2 \).

Under the hypothesis \( H_1 \), the covariance matrix can be written as

\[
(4.2.1) \quad \Sigma = I \otimes \Sigma_1 + (J-I) \otimes \Sigma_2,
\]

where

- \( I(t\times t) \) is an identity matrix,
- \( J(t\times t) \) is a matrix of 1's,
- \( \Sigma_1(p\times p) \) and \( \Sigma_2(p\times p) \) are symmetric.

Since the matrix \( J-I \) is symmetric, there exists an orthogonal matrix \( \Gamma(t\times t) \), such that \( \Gamma'(J-I)p = D \), where \( D(t\times t) = \text{Diag}(t-1,-1,-1,\ldots,-1) \), with elements being the eigenvalues of the matrix \( J-I \). Then we have

\[
(4.2.2) \quad \Sigma = (\Gamma \Gamma') \otimes \Sigma_1 + (\Gamma D\Gamma') \otimes \Sigma_2
\]

\[
= (\Gamma \otimes I_p)(I_t \otimes \Sigma_1)(\Gamma' \otimes I_p) + (\Gamma \otimes I_p)(D \otimes \Sigma_2)(\Gamma' \otimes I_p)
\]

\[
= (\Gamma \otimes I_p)(I_t \otimes \Sigma_1 + D \otimes \Sigma_2)(\Gamma' \otimes I_p)
\]
where Block(A_1, A_2, ..., A_t) is a block diagonal matrix with the matrices A_1, A_2, ..., A_t on the main diagonal, \( \Omega_1 = \Sigma_1 + (t-1)\Sigma_2 \), and \( \Omega_2 = \Sigma_1 - \Sigma_2 \).

From Section 4.1, we know that \( \hat{\beta} = (X'X)^{-1}X'Y \) is the MLE of \( \beta \) when \( \Sigma_0 \) is positive definite. Thus, under \( H_1 \), the log-likelihood function, after being maximized with respect to \( \beta \), is

\[
L(\beta, \Omega_1, \Omega_2) = -\frac{1}{2} N \text{tr} \ln(2\pi) - \frac{1}{2} N \ln \left| \left( \Gamma \otimes I_p \right) \text{Block}(\Omega_1, \Omega_2, \ldots, \Omega_2) \left( \Gamma' \otimes I_p \right) \right| \\
- \frac{1}{2} \text{trace} \left[ \left( \Gamma \otimes I_p \right) \text{Block}(\Omega_1^{-1}, \Omega_2^{-1}, \ldots, \Omega_2^{-1}) \left( \Gamma' \otimes I_p \right) S \right] \\
= -\frac{1}{2} N \text{tr} \ln(2\pi) - \frac{1}{2} N \ln \left| \Omega_1 \right| - \frac{1}{2} N (t-1) \ln \left| \Omega_2 \right| \\
- \frac{1}{2} \text{trace} \left[ \text{Block}(\Omega_1^{-1}, \Omega_2^{-1}, \ldots, \Omega_2^{-1}) \right] \\
= -\frac{1}{2} N \text{tr} \ln(2\pi) - \frac{1}{2} N \ln \left| \Omega_1 \right| - \frac{1}{2} N (t-1) \ln \left| \Omega_2 \right| \\
- \frac{1}{2} \text{trace} \left( \Omega_1^{-1} T_1 \right) - \frac{1}{2} \text{trace} \left( \Omega_2^{-1} T_2 \right),
\]

where \( S = (Y - X\hat{\beta})' (Y - X\hat{\beta}) = Y' [1 - \frac{X(X'X)^{-1}X'}{1}] Y \),

\[
T = \left( \Gamma' \otimes I_p \right) S \left( \Gamma \otimes I_p \right) = (T_{ij}), \quad i, j = 1, 2, \ldots, t,
\]

\[
T_1 = T_{11}, \quad T_2 = \sum_{j=2}^{t} T_{jj}.
\]

In the above log-likelihood, the two matrices \( \Omega_1 \) and \( \Omega_2 \) do not appear in the same term. Thus, the MLE's of \( \Omega_1 \) and \( \Omega_2 \) can be derived separately as

\[
\hat{\Omega}_1 = \frac{1}{N} T_1 \quad \text{and} \quad \hat{\Omega}_2 = \frac{1}{N(t-1)} T_2.
\]

With \( \Omega_1 = \Sigma_1 + (t-1)\Sigma_2 \), \( \Omega_2 = \Sigma_1 - \Sigma_2 \), we then have the following results.
Theorem 4.2.1. Under $H_1$, the MLE's of $\Sigma_1$ and $\Sigma_2$ are given by

\begin{align}
(4.2.6) \quad \hat{\Sigma}_1 &= \frac{1}{t} \left[ \hat{\Omega}_1 + (t-1)\hat{\Omega}_2 \right] = \frac{1}{Nt}(T_1 + T_2) \quad \text{and} \\
(4.2.7) \quad \hat{\Sigma}_2 &= \frac{1}{t} \left[ \hat{\Omega}_1 - \hat{\Omega}_2 \right] = \frac{1}{Nt}(T_1 - \frac{1}{t-1}T_2).
\end{align}

The resulting maximum log-likelihood function is

\begin{align}
(4.2.8) \quad L_1 &= L(\hat{\beta}, \hat{\Omega}_1, \hat{\Omega}_2) = -\frac{1}{2}Nt \ln(2t) + \frac{1}{2}Nt \ln|N| + \frac{1}{2}N(t-1)p \ln|t-1| \\
&\quad \quad -\frac{1}{2}Nt \ln(|T_1|) - \frac{1}{2}N(t-1) \ln(|T_2|) - \frac{1}{2}Nt p.
\end{align}

Corollary 4.2.2. The LR statistic used in testing $H_1$ vs. $H_0$ is

\begin{align}
(4.2.9) \quad \Lambda_1 &= \left\{ \frac{(t-1)(t-1)p}{|T_1|} \right\}^{\frac{1}{2}N}.
\end{align}

When $t = 2$, the above LR test statistic reduces to

\begin{align}
(4.2.10) \quad \Lambda_1 &= \left\{ \frac{|T|}{|T_1| |T_2|} \right\}^{\frac{1}{2}N},
\end{align}

which is equivalent to a LR statistic for testing the independence of 2 sets of variables.

The maximum log-likelihood under $H_1$ can also be derived without the use of above orthogonal transformations because the inverse of an intraclass correlation matrix also has an intraclass correlation structure. In this case the resulting LR statistic is

\begin{align}
(4.2.11) \quad \Lambda_1^* &= \left\{ \frac{tp(t-1)(t-1)p}{|S|} \right\}^{\frac{1}{2}N}.
\end{align}

Equating $\Lambda_1$ and $\Lambda_1^*$, we have

\begin{align}
(4.2.12) \quad \left[ \frac{1}{N}(s_1+s_2) \right] | \left[ \frac{1}{N}(t-1)s_1-s_2 \right] |^{(t-1)} = |T_1| |T_2|^{(t-1)}.
\end{align}

We then have the following useful result.
Theorem 4.2.3. Let $S_{(tp \times tp)} = (S_{ij}(p \times p))$ be a positive definite symmetric matrix and $T_{(tp \times tp)} = (T_{ij}(p \times p)) = (\Gamma' \otimes I_p) S (\Gamma \otimes I_p)$, where $\Gamma_{(t \times t)}$ is an orthogonal matrix such that 

\[ \Gamma' [J_{(t \times t)} - I_t] \Gamma = \text{Diag}(t-1, -1, \ldots, -1). \]

Then we have 

\[ |T_1| = \frac{1}{t} (S_1 + S_2), \quad \text{and} \]

\[ |T_2| = \frac{1}{t} [(t-1) S_1 - S_2], \]

where $T_1 = T_{11}$ and $T_2 = \sum_{j=2}^t T_{jj}$. 

Proof:

Denoting $\Gamma_{(t \times t)} = (\Gamma_1(t \times 1) \Gamma_2(t \times [t-1]))$, then we have 

\[ \Gamma_1 = \frac{1}{\sqrt{t}} J_{(t \times 1)} \quad \text{and} \]

\[ T = \left[ \begin{array}{c} \Gamma_1' \otimes I_p \\ \Gamma_2' \otimes I_p \end{array} \right] S \left[ \begin{array}{c} \Gamma_1 \otimes I_p \\ \Gamma_2 \otimes I_p \end{array} \right]. \]

Thus, 

\[ T_1 = (\Gamma_1' \otimes I_p) S (\Gamma_1 \otimes I_p) = \frac{1}{t} [(J_{(1 \times p)} \otimes I_p) S (J_{(p \times 1)} \otimes I_p)] \]

\[ = \frac{1}{t} \sum_{i,j} S_{ij} = \frac{1}{t} (S_1 + S_2), \]

\[ |T_1| = \frac{1}{t} (S_1 + S_2). \]

Together with (4.2.12), we have 

\[ |T_2| = \frac{1}{t} [(t-1) S_1 - S_2]. \]

4.2.1 The Asymptotic Null Distribution of the LR Statistic

In this section we derive an asymptotic expansion for the LR statistic $A_1$ as sample size $N$ increases, beginning with a lemma regarding the moments of the LR statistic when the hypothesis $H_1$ is true.
Lemma 4.2.4. When the hypothesis $H_1$ is true, the $h^{th}$ moment of $A_1$ is given by

$$E(A_1^h) = k_0 \left( t-1 \right)^{\frac{hN(t-1)p}{2}} \prod_{j=1}^{tp} \Gamma \left[ \frac{hN(1+h)}{2} + \frac{1}{2}(1-q-j) \right] \prod_{k=1}^{p} \Gamma \left[ \frac{hN(1+h)}{2} + \frac{1}{2}(1-q-k) \right] \prod_{k=1}^{p} \Gamma \left[ \frac{hN(t-1)}{2} + \frac{1}{2}(1-(t-1)q-k) \right]$$

where $k_0$ is a constant not involving $h$.

The $h^{th}$ moment of $A_1$ has the same form as (4.1.6) with

$$a = tp, x_j = \frac{N}{2}, e_j = \frac{1}{2}(1-q-j), j = 1,2,\ldots,a,$$

$$b = 2p, y_k = \frac{N}{2}, t_k = \frac{1}{2}(1-q-k), k = 1,2,\ldots,p,$$

$$y_k = \frac{N}{2}(t-1), t_k = \frac{1}{2}(1-(t-1)q+p-k), k = p+1,\ldots,2p.$$

Thus, analogous to Theorem 4.1.1, we have

Theorem 4.2.5. When the null hypothesis $H_1$ is true, the distribution function of $-2p \ln A_1$, where $p$ is given by (4.1.9), can be expanded for large $M = pN$ as

$$p(-2p \ln A_1 \leq z) = p(\chi^2_f \leq z) + O(M^{-2}),$$

where

$$f = -2 \left[ \sum_{j=1}^{a} e_j - \sum_{k=1}^{b} f_k - \frac{1}{2}(a-b) \right] = \frac{p}{2} \left[ (t^2-2)p + (t-2) \right].$$

4.2.2. Asymptotic Non-null Distributions of the LR Statistic When $t=2$

When $t=2$, the LR statistic, $A_1$, is equal to

$$\sum_{i=1}^{N} \left| T_{ii} \right|^2$$

The random matrix, $T$, is distributed as a Wishart distribution $W_{2p}(n,\Omega)$, where $n = N-q$ and $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12} & \Omega_{22} \end{bmatrix}$. When the hypothesis $H_1$ is true, $\Omega_{12} = 0$. 
The power function of the LR test of level $\alpha$ is $P(-2p \ln \Lambda_1 > k_\alpha)$, where $k_\alpha$ is the upper $100\alpha \%$ point of the distribution of $-2p \ln \Lambda$, when the hypothesis $H_1 (\Omega_{12} = 0)$ is true. Let $p^2 = \text{Diag}(\omega_1, \ldots, \omega_p)$, where $\omega_1, \ldots, \omega_p$ are the eigenvalues of $\Omega_{11}^{-1} \Omega_{12} \Omega_{22}^{-1} \Omega_{21}$. Then, under $H_1 p^2 = 0$.

We consider two different alternatives

(A) A fixed alternative $H_k: \Omega_{12} \neq 0$.

(B) A sequence of local alternatives $H_M: p^2 = \frac{1}{M^2} \psi$,
where $\psi$ is a fixed diagonal matrix.

Define the random variable $Z$ by

$$Z = \frac{-2p \ln \Lambda_1}{M^2} + M^2 \ln(|I-p^2|).$$

Then from Muirhead (1982), we have the following results.

**Theorem 4.2.6.** Under the fixed alternative $H_k: \Omega_{12} \neq 0$, the distribution function of the random variable $Z$ given by (4.2.17), can be expanded as

$$P\left(\frac{Z}{\tau} \leq z\right) = \Phi(z) - M^{-k} \left[\frac{p^2}{\tau} \phi(z) + \frac{4}{\tau^2} (\sigma_1 - q_2) \phi(2)(z)\right] + O(M^{-1}),$$

where $\tau = \text{trace } p^2$, $\tau^2 = 4\sigma_1$, and $\Phi$ and $\phi$ denote the standard normal distribution and density functions, respectively.

**Theorem 4.2.7.** Under the sequence of local alternatives $H_M: p^2 = \frac{1}{M^2} \psi$,

The distribution function of $-2p \ln \Lambda_1$ can be expanded as

$$P(-2p \ln \Lambda_1 \leq z) = P(X_f^2 \leq z) + \frac{\sigma_1}{M} \left[P(X_{f+2}^2 \leq z) - p(X_f^2 \leq z)\right] + O(M^{-1}),$$

where $f = p^2$, $\rho = 1 - \frac{1}{N}(p^2 + q)$, and $\sigma_1 = \text{trace}(\psi)$.

**4.3. Testing Proportionality for the Two Matrices, $E_1$ and $E_2$, in $H_1$**

$H_2: E_2 = \omega E_1$

When $H_1$ is true, the log-likelihood function, maximized with respect to $\beta$, is given by (4.2.3). That is...
(4.3.1) \[ L(\hat{\beta}, \Omega_1, \Omega_2) = -\frac{1}{2} N tp \ln(2\pi) - \frac{1}{2} N \ln(|\Omega_1|) - \frac{1}{2} N(t-1) \ln(|\Omega_2|) \\
- \frac{1}{2} \text{trace}(\Omega_1^{-1}T_1) - \frac{1}{2} \text{trace}(\Omega_2^{-1}T_2), \]

where \[ \Omega_1 = \Sigma_1 + (t-1)\Sigma_2, \quad \Sigma_2 = \Omega_1 - \Omega_2. \]

When \( H_2 \) is true, \( \Sigma_2 = \omega E_1, \quad \Omega_1 = [1 + (t-1)\omega] \Sigma_1 \) and \( \Sigma_2 = [1-\omega] \Sigma_1 \).

Thus, under \( H_2 \), the log-likelihood function, maximized with respect to \( \beta \), is

(4.3.2) \[ L(\hat{\beta}, \omega, \Sigma_1) = -\frac{1}{2} N tp \ln(2\pi) - \frac{1}{2} Np \ln[1 + (t-1)\omega] - \frac{1}{2} N(t-1)p \ln[1-\omega] \\
- \frac{1}{2} Nt \ln(|\Sigma_1|) - \frac{1}{2} \text{trace} \left\{ \Sigma_1^{-1} \left[ \frac{1}{1+(t-1)\omega} T_1 + \frac{1}{1-\omega} T_2 \right] \right\}. \]

The matrix \( \frac{1}{1+(t-1)\omega} T_1 + \frac{1}{1-\omega} T_2 \) is symmetric. Thus, the function \( L(\hat{\beta}, \omega, \Sigma_1) \) can be maximized with respect to \( \Sigma_1 \) at

(4.3.3) \[ \hat{\Sigma}_1 = \frac{1}{Nt} \left[ \frac{1}{1+(t-1)\omega} T_1 + \frac{1}{1-\omega} T_2 \right]. \]

The resulting log-likelihood function is

(4.3.4) \[ L(\hat{\beta}, \omega, \hat{\Sigma}_1) = -\frac{1}{2} N tp \{ \ln(2\pi) - \ln(Nt) + 1 \} + \frac{1}{2} N(t-1)p \ln[1+(t-1)\omega] \\
+ \frac{1}{2} Np \ln[1-\omega] \\
- \frac{1}{2} Nt \ln\{ |T_1 + T_2| - \omega_1 |T_1 - (t-1)T_2| \}. \]

To simplify the above expression, we use the following result from Rao (1973).

Lemma 4.3.1. Let \( A \) and \( B \) be \( p \times p \) symmetric matrices of which \( B \) is positive definite. Then there exists a matrix \( R \) such that \( A = R'AR \) and \( B = R'R \), where \( A = \text{Diag}(\lambda_1, \ldots, \lambda_p) \) with \( \lambda_k \) being the \( k^{\text{th}} \) eigenvalue of \( AB^{-1} \).

Replacing \( A \) by \( T_1 - (t-1)T_2 \), \( B \) by \( T_1 + T_2 \) in the above lemma, we have
(4.3.5) \[ |[T_1+T_2] - \omega_1[T_1-(t-1)T_2]| = |R'R - \omega R'\Lambda R| \]
\[ = |R'| |I-\omega \Lambda| |R| = |R'R| \prod_{k=1}^{p} (1-\omega \lambda_k) \]
\[ = |T_1+T_2| \prod_{k=1}^{p} (1-\omega \lambda_k), \]

where \( \lambda_k \) is the \( k^{th} \) eigenvalue of \([T_1-(t-1)T_2][T_1+T_2]^{-1} \).

Then, the log-likelihood is

(4.3.6) \[ L(\hat{\theta}, \omega, \hat{\Sigma}_1) = -\frac{1}{2} N \ln(2\pi) - \ln(N_t) + 1} + \frac{1}{2} N \ln(t-1) \ln(1+(t-1)\omega] \]
\[ + \frac{1}{2} N \ln|1-\omega| - \frac{1}{2} N \ln(|T_1+T_2|) - \frac{1}{2} N \sum_{k=1}^{p} \ln(1-\omega \lambda_k) \],

and the MLE of \( \omega, \hat{\omega} \), can be solved from the partial equation

\[ \frac{\partial L(\hat{\theta}, \omega, \hat{\Sigma}_1)}{\partial \omega} = 0. \]
That is

(4.3.7) \[ \frac{\partial L(\hat{\theta}, \omega, \hat{\Sigma}_1)}{\partial \omega} = -\frac{1}{2} N \omega \left( \frac{(t-1)^2}{1+(t-1)\omega} + \frac{1}{2} N \frac{p(-(t-1))}{1-\omega} \right) - \frac{1}{2} N \sum_{k=1}^{p} \frac{\lambda_k}{1-\omega \lambda_k} \left( 1+(t-1)\omega \right) \left[ 1-\omega \right] \]
\[ = \frac{N t}{2\left[ 1+(t-1)\omega \right]} \left\{ \frac{p\left( (t-2) - \omega(t-1) \right)}{1-\omega \lambda_k} + \sum_{k=1}^{p} \frac{\lambda_k}{1-\omega \lambda_k} \left[ 1+(t-1)\omega \right] \left[ 1-\omega \right] \right\} \]
\[ = \frac{N t}{2\left[ 1+(t-1)\omega \right]} \left\{ \sum_{k=1}^{p} \frac{(t-2) - \omega(t-1) + \lambda_k}{1-\omega \lambda_k} \right\}. \]

Since \( \frac{1}{t-1} < \omega < 1 \), \( [1-\omega] [1+(t-1)\omega] \neq 0 \), \( \hat{\omega} \) can be solved from \( g(\omega) = 0 \), where

(4.3.8) \[ g(\omega) = \sum_{k=1}^{p} \frac{(t-2) - (t-1)\omega + \lambda_k}{1-\omega \lambda_k}. \]

The function \( g(\omega) \) is a continuous function of \( \omega \) with

\[ g(-\frac{1}{t-1}) = p(t-1) > 0 \]
and \[ g(1) = -p < 0. \]
There exists at least one solution of $\omega$ between $-\frac{1}{t-1}$ and 1. To show the uniqueness of the solution, we evaluate

$$(4.3.9) \quad \frac{\partial g(\omega)}{\partial \omega} = \sum_{k=1}^{p} \frac{(-1)}{[1-\lambda_k \omega]^2} (1-\lambda_k) \left[ (t-1) + \lambda_k \right],$$

and use the following lemma.

**Lemma 4.3.2.** If matrices $T_1 (p \times p)$ and $T_2 (p \times p)$ are positive definite, then the eigenvalues of $[T_1 - (t-1)T_2][T_1 + T_2]^{-1}$ are between $-(t-1)$ and 1.

**Proof:**

Since $[T_1 + T_2] - [T_1 - (t-1)T_2] = tT_2$ and

$$(t-1)[T_1 + T_2] + [T_1 - (t-1)T_2] = tT_1$$

are positive definite.

We know that both $I - [T_1 - (t-1)T_2][T_1 + T_2]^{-1}$ and

$$(t-1)I + [T_1 - (t-1)T_2][T_1 + T_2]^{-1}$$

are positive definite. Then the eigenvalues of $[T_1 - (t-1)T_2][T_1 + T_2]^{-1}$ are less than 1 and greater than $-(t-1)$.

Thus,

$$(4.3.10) \quad \frac{\partial g(\omega)}{\partial \omega} = \sum_{k=1}^{p} \frac{(-1)}{[1-\lambda_k \omega]^2} (1-\lambda_k)[(t-1) + \lambda_k] < 0.$$

Since $g(\omega)$ is a strictly decreasing function of $\omega$, there exists a unique solution for $g(\omega) = 0$ between $-\frac{1}{(t-1)}$ and 1. The value of $\hat{\omega}$ can be evaluated through numerical methods, e.g. the halving-interval method.

Summarizing these results, we have

**Theorem 4.3.3.** Under $H_2$, the MLE of $\hat{\omega}_1$ is
where \( \hat{\omega} \) is the unique solution of \( g(\hat{\omega}) = 0 \), and \( g(\hat{\omega}) \) is given by (4.3.8).

The resulting maximum log-likelihood function is

\[
L_2 = L(\hat{\beta}, \hat{\omega}, \hat{\Sigma}_1)
= -\frac{1}{2}Nt \ln \{ \ln(2\pi) - \ln(\text{det}(\Sigma_1)) + 1 \} + \frac{1}{2}N(t-1)p \ln[1+(t-1)\hat{\omega}] + \frac{1}{2}Np \ln[1-\hat{\omega}] - \frac{1}{2}Nt \ln(\text{det}(\Sigma_2)) - \frac{1}{2}N(t-1)\hat{\omega}.
\]

\[\text{Corollary 4.3.4.} \text{ The LR statistic used in testing } H_0 \text{ vs. } H_1 \text{ is given by}
\]

\[
\Lambda_2 = \exp(L_2 - L_1)
= \left\{ \frac{\ln(2\pi)}{p(t-1)} - \frac{1}{2}Nt \ln(\text{det}(\Sigma_1)) + \frac{1}{2}N(t-1)\hat{\omega} \right\}.
\]

Under \( H_0 \), the statistic \(-2 \ln \Lambda_2\) is asymptotically distributed as a \( \chi^2 \) variable with \( f \) degrees of freedom, where \( f = 2 \left[ \frac{p(p+1)}{2} - \left( \frac{p(p+1)}{2} + 1 \right) \right] = \frac{1}{2}(p-1)(p+2) \). We reject \( H_0 \) for large values of \(-2 \ln \Lambda_2\).

\[\text{4.3.1 The Wald statistic when } P = 2\]

Under the hypothesis \( H_1 \), the log-likelihood maximized with respect to \( \beta \), is given by (4.3.1). Since \( \Sigma_1 \) is a positive definite symmetric matrix, there exists a symmetric matrix \( \Sigma_1^{-\frac{1}{2}} \) such that \( \Sigma_1^{-\frac{1}{2}} \Sigma_1^{\frac{1}{2}} = \Sigma_1^{\frac{1}{2}} \Sigma_1^{\frac{1}{2}} = I \). Letting \( R = \Sigma_1^{-\frac{1}{2}} \Sigma_2 \Sigma_1^{-\frac{1}{2}} \), where \( \Sigma_1^{-\frac{1}{2}} = (\Sigma_1^{\frac{1}{2}})^{-1} \), then we have \( \Sigma_2 = \Sigma_1^{\frac{1}{2}} R \Sigma_1^{\frac{1}{2}} \).

Under \( H_1 \), the MLE's of \( \Sigma_1 \) and \( \Sigma_2 \) are

\[
\hat{\Sigma}_1 = \frac{1}{Nt} (T_1 + T_2) \quad \text{and} \quad \hat{\Sigma}_2 = \frac{1}{Nt} (T_1 - \frac{1}{t-1} T_2).
\]
With a one-to-one transformation between \( \{\Sigma_1, \Sigma_2\} \) and \( \{\Sigma_1, R\} \), the MLE of \( R \) is

\[
\hat{R} = \hat{\Sigma}_1^{-\frac{1}{2}} \hat{\Sigma}_2 \hat{\Sigma}_1^{-\frac{1}{2}}.
\]

When \( p = 2 \), we can express \( R = \begin{bmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_3 \end{bmatrix}, \Sigma_1^{-\frac{1}{2}} = \begin{bmatrix} d_1 & d_2 \\ d_2 & d_3 \end{bmatrix}, T_1 = \begin{bmatrix} u_{11} & u_{12} \\ u_{12} & u_{22} \end{bmatrix}, \) and

\[
T_2 = \begin{bmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{bmatrix}.
\]

Furthermore, by denoting \( \theta' = (\theta_1, \theta_2, \theta_3) \) and \( d' = (d_1, d_2, d_3) \), the log-likelihood can be written as a function of \( \theta \) and \( d \).

\[
L(\theta, d) = -\frac{1}{2} \ln(2\pi) + N_t \ln(D(d)) - \frac{1}{2} N_t \ln(A_1(\theta)) - \frac{1}{2} N(t-1) \ln(A_2(\theta))
\]

where

\[
A_1(\theta) = |I + (t-1)R| = |1 + (t-1)\theta_1| |1 + (t-1)\theta_3| - (t-1)^2 \theta_2^2,
\]

\[
A_2(\theta) = |I - R| = (1-\theta_1)(1-\theta_3) - \theta_2^2,
\]

\[
B_1(d, \theta) = [1 + (t-1)\theta_3][d_1^2 u_{11} + 2d_1 d_2 u_{12} + d_2^2 u_{22}]
- 2(t-1)\theta_2[d_1^2 u_{11} + d_2^2 u_{12} + d_1 d_3 u_{12} + d_2 d_3 u_{22}]
+ [1 + (t-1)\theta_1][d_2^2 u_{11} + 2d_2 d_3 u_{12} + d_3^2 u_{22}],
\]

and

\[
B_2(d, \theta) = (1-\theta_3)[d_1^2 v_{11} + 2d_1 d_2 v_{12} + d_2^2 v_{22}] + 2\theta_2[d_1^2 v_{11} + d_2^2 v_{12} + d_1 d_3 v_{12} + d_2 d_3 v_{22}]
+ d_2 d_3 v_{22} + (1-\theta_1)[d_2^2 v_{11} + 2d_2 d_3 v_{12} + d_3^2 v_{22}].
\]

Under \( H_2: \Sigma_2 = \omega \Sigma_1 \), we have \( R = \omega I \).
Thus, the hypothesis \( H_2 \) is equivalent to the hypothesis that:

\[
\begin{bmatrix}
\hat{\theta}_N \\
\hat{d}_N
\end{bmatrix}
= 0, \hspace{1em} \text{where } \begin{bmatrix}
C_{(2 \times 6)}
\end{bmatrix} = \begin{bmatrix}
1 & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0
\end{bmatrix}.
\]

As in Section 4.1.2, a Wald statistic can be defined

\[
W_N = N \left[ C \left( \begin{bmatrix}
\hat{\theta}_N \\
\hat{d}_N
\end{bmatrix} \right) \right]' \left[ C I_{0J}^{-1} (\hat{\theta}, \hat{d}) C' \right]^{-1} \left[ C \left( \begin{bmatrix}
\hat{\theta}_N \\
\hat{d}_N
\end{bmatrix} \right) \right],
\]

where \( I_{0J}^{-1}(\hat{\theta}, \hat{d}) \) is a jackknife estimator of \( I_0^{-1}(\theta, d) \) given by (4.1.17).

Omitting the lengthy derivation, we have the following matrix expectations.

\[
\begin{aligned}
\frac{\partial^2 L(\theta, d)}{\partial \theta \partial \theta'} &= -\frac{1}{2}(N-2n) \left\{ \frac{[\Delta_\theta \theta, A_1]}{[A_1(\theta)]} - \frac{[\Delta_\theta A_1][\Delta_\theta A_1]'}{[A_1(\theta)]^2} \right. \\
&\quad \left. + (t-1) \frac{[\Delta_\theta \theta, A_2]}{[A_2(\theta)]} - (t-1) \frac{[\Delta_\theta A_2][\Delta_\theta A_2]'}{[A_2(\theta)]^2} \right\} \\
\frac{\partial^2 L(\theta, d)}{\partial d \partial \theta'} &= N_t \frac{[\Delta_\theta d, D]}{[D(d)]} - N_t \frac{[\Delta_d D][\Delta_d D]'}{[D(d)]^2} - \frac{E[\Delta_\theta d, B_1]}{[A_1(\theta)]}, \quad \text{and} \\
\frac{\partial^2 L(\theta, d)}{\partial d \partial d'} &= -\frac{1}{2} \frac{E[\Delta_\theta d, B_1]}{[A_1(\theta)]} + \frac{1}{2} \frac{E[\Delta_d B_1][\Delta_\theta A_1]'}{[A_1(\theta)]^2} - \frac{1}{2} \frac{E[\Delta_d B_2]}{[A_2(\theta)]} \\
&\quad + \frac{1}{2} \frac{E[\Delta_\theta B_2][\Delta_\theta A_2]'}{[A_2(\theta)]^2},
\end{aligned}
\]

where

\[
\begin{aligned}
\Delta_\theta A_1 &= (t-1) \begin{bmatrix} 1+(t-1) \theta_3, & -2(t-1) \theta_2, & 1+(t-1) \theta_1 \end{bmatrix}', \\
\Delta_d D &= \begin{bmatrix} d_3, & -2d_2, & d_1 \end{bmatrix}',
\end{aligned}
\]
\[
\Delta_{d B_1} = 2 \begin{bmatrix}
[1+(t-1)\theta_3 \delta z] & \left[ d_1 u_{11} + d_2 u_{12} \right] - (t-1)\theta_2 \left[ d_2 u_{11} + d_3 u_{12} \right] \\
[1+(t-1)\theta_3 \delta z] & \left[ \left[ d_1 u_{11} + d_2 u_{12} \right] - (t-1)\theta_2 \left[ d_1 u_{11} + 2d_2 u_{12} + d_3 u_{22} \right] + [1+(t-1)\theta_1] \right] \\
\left[ d_1 u_{11} + d_2 u_{12} \right] & \left[ \left[ d_1 u_{11} + d_2 u_{12} \right] + [1+(t-1)\theta_1] \right] \left[ d_2 u_{11} + d_3 u_{22} \right] \\
- (t-1)\theta_2 \left[ d_1 u_{12} + d_2 u_{22} \right] + [1+(t-1)\theta_1] \left[ d_2 u_{11} + d_3 u_{22} \right]
\end{bmatrix},
\]

\[
\Delta_{e A_1} = (t-1)^2 \begin{bmatrix}
0 & 0 & 1 \\
0 & -2 & 0 \\
1 & 0 & 0
\end{bmatrix}, \quad \Delta_{d d^'} D = \begin{bmatrix}
0 & 0 & 1 \\
0 & -2 & 0 \\
1 & 0 & 0
\end{bmatrix},
\]

\[
\Delta_{d d^', B_1} =
\begin{bmatrix}
\left[ 1+(t-1)\theta_3 \right] u_{11} & \left[ 1+(t-1)\theta_3 \right] u_{12} - (t-1)\theta_2 u_{11} & -(t-1)\theta_2 u_{12} \\
\left[ 1+(t-1)\theta_3 \right] u_{22} - 2(t-1)\theta_2 u_{12} + (1+(t-1)\theta_1) u_{11} & -(t-1)\theta_2 u_{22} & + [1+(t-1)\theta_1] u_{12} \\
\left[ 1+(t-1)\theta_1 \right] u_{22} & \left[ 1+(t-1)\theta_1 \right] u_{12}
\end{bmatrix},
\]

\[
\Delta_{d o, B_1} = 2(t-1) \begin{bmatrix}
0 & (d_2 u_{11} + d_3 u_{12}) & (d_1 u_{11} + d_2 u_{12}) \\
(d_1 u_{11} + d_3 u_{22} + 2d_2 u_{12}) & (d_1 u_{12} + d_2 u_{22}) & 0 \\
\end{bmatrix},
\]

The matrices $\Delta o A_2$, $\Delta d B_2$, $\Delta o o A_2$, $\Delta dd^, B_2$, and $\Delta d o, B_2$ can be derived from $\Delta o A_1$, $\Delta d B_1$, $\Delta o o A_1$, $\Delta dd^, B_1$, and $\Delta d o, B_1$ be replacing $(t-1)$ with $-1$ and $u_{ij}$ with $v_{ij}$.

The statistic $T_1$ has a Wishart distribution, $W_2(n, \Sigma_1 \left[ 1+(t-1)R \right] \Sigma_1^t)$ with expectation:

\[
E(T_1) = n \Sigma_1^t \left[ 1+(t-1)R \right] \Sigma_1^t. \quad \text{That is,}
\]
(4.3.20) \[ E(u_{11}) = \frac{n}{[D(d)]^2} \left\{ d_3^2 \left[ 1+(t-1)\theta_1 \right] - 2(t-1)d_2d_3\theta_2 + d_2^2 \left[ 1+(t-1)\theta_3 \right] \right\}, \]

\[ E(u_{12}) = \frac{-n}{[D(d)]^2} \left\{ d_2d_3 \left[ 1+(t-1)\theta_1 \right] - (t-1)(d_2^2 + d_1d_3)\theta_2 + d_1d_2 \left[ 1+(t-1)\theta_3 \right] \right\}, \]

\[ E(u_{22}) = \frac{n}{[D(d)]^2} \left\{ d_2^2 \left[ 1+(t-1)\theta_1 \right] - 2(t-1)d_1d_2\theta_2 + d_1^2 \left[ 1+(t-1)\theta_3 \right] \right\}. \]

Similar expectations for \( v_{11}, v_{12}, \) and \( v_{22} \) can also be obtained from \( E(u_{11}), E(u_{12}), \) and \( E(u_{22}) \) by replacing \( n \) with \( n(t-1) \) and \( (t-1) \) with -1.

The expected values of \( \Delta dB_1', \Delta dB_2', \Delta dd'B_1', \Delta dd'B_2', \Delta d\theta'B_1', \) and \( \Delta d\theta'B_2' \) are obtained by replacing \( \{u_{11}, u_{12}, u_{22}, v_{11}, v_{12}, v_{22}\} \) with corresponding expectations in the expression.

With the above expected matrices, the matrix \( I_0(\hat{\theta}_N, \hat{d}_N) \) can then be evaluated. The matrices \( I_0(\hat{\theta}_N(j), \hat{d}_N(j)) \), \( j = 1, 2, \ldots, N \), can also be evaluated in a similar manner. Thus, we have a jackknife estimator of the inverse of the information matrix, \( I_0^{-1}(\hat{\theta}, \hat{d}) \), and the Wald statistic.

4.3.2. Asymptotic Distributions of the Wald Statistic when \( p=2 \)

When \( p=2 \), the Wald statistic \( W_N \) can be obtained from (4.3.18). Under the hypothesis \( H_2: E_2 = \omega E_1 \), the statistic \( W_N \) is asymptotically distributed as a \( \chi^2 \) variate with 2 degrees of freedom. We reject \( H_2 \) for large values of \( W_N \).

The power function of the Wald test of level \( \alpha \) is \( p(W_N > k_\alpha) \), where \( k_\alpha \) is the upper 100 \( \alpha \)% point of the distribution of \( W_N \) when \( H_2 \) is true.

Letting \( \Delta'_\theta = (\theta_1-\theta_3, \theta_2) \), under the hypothesis \( H_2 \), we have \( \Delta_\theta = 0 \). Now consider a sequence of local alternatives \( H_N: \Delta_\theta = \frac{1}{\sqrt{N}} \Delta_0 \), where \( \Delta_0 \) is a vector of fixed constants. When \( H_N \) is true, the Wald statistic \( W_N \) is
asymptotically distributed as a non-central $\chi^2$ variate with 2 degrees of freedom and non-centrality parameter

\[
(4.3.21) \quad \lambda^2 = \Delta_0 \left[ C T_0^{-1} (\hat{\theta}, \hat{\alpha}) C' \right]^{-1} \Delta_0.
\]

Since the non-central $\chi^2$ distribution is a mixture of $\chi^2$ distributions with Poisson weights, we have the power function

\[
(4.3.22) \quad p(W_n > k_\alpha | W_n \sim \chi^2_2(\lambda^2)) = \sum_{k=0}^{\infty} P_k(\lambda) \cdot P(\chi^2_{2+2k} > k_\alpha),
\]

where

\[
P_k(\lambda) = \left( \frac{\lambda^2}{2} \right)^k e^{-k\lambda^2} / k!.
\]

Many approximations to the non-central $\chi^2$ distribution, $\chi^2_\mu(\lambda^2)$, have been suggested. Among them, Pearson (1959) used a linear function of a central $\chi^2$ distribution, $c \chi^2 + b$, so that the first three moments of $\chi^2_\mu(\lambda^2)$ and $c \chi^2 + b$ agree. The appropriate values of $b$, $c$ and $f$ are

\[
b = - \frac{\lambda^4}{\mu + 3\lambda^2}, \quad \quad c = \frac{\mu + 3\lambda^2}{\mu + 2\lambda^2}, \quad \quad \text{and} \quad \quad f = \mu + \frac{\lambda^4(3\mu + 8\lambda^2)}{(\mu + 3\lambda^2)^2}.
\]

4.4. Testing Intraclass Covariance Structures for the Two Matrices, $\Sigma_1$ and $\Sigma_2$, in $H_1$:

$H_3$: $\Sigma_1 = \sigma_1 I + \sigma_1 \rho_1 (J-I)$, $\Sigma_2 = \sigma_2 I + \sigma_2 \rho_2 (J-I)$.

Under the hypothesis $H_1$ of block intraclass covariance structure, the log-likelihood function, maximized with respect to $B$, is given by (4.2.3). That is
(4.4.1) \begin{align*}
L(\beta, \Omega_1, \Omega_2) &= -\frac{1}{2} N \ln(2\pi) - \frac{1}{2} N \ln(|\Omega_1|) - \frac{1}{2} N(t-1) \ln(|\Omega_2|) \\
&\quad - \frac{1}{2} \text{trace}(\Omega_1^{-1} T_1) - \frac{1}{2} \text{trace}(\Omega_2^{-1} T_2),
\end{align*}

where \( \Omega_1 = \Sigma_1 + (t-1) \Sigma_2 \) and \( \Omega_2 = \Sigma_1 - \Sigma_2 \).

Under \( H_2 \), we have

(4.4.2) \begin{align*}
\Omega_1 &= [\sigma_1 + (t-1)\sigma_2] I + [\sigma_1 \rho_1 + (t-1)\sigma_2 \rho_2] (J-I) \\
\text{and} \quad \Omega_2 &= [\sigma_1 - \sigma_2] I + [\sigma_1 \rho_1 - \sigma_2 \rho_2] (J-I).
\end{align*}

-Both \( \Omega_1 \) and \( \Omega_2 \) have intraclass covariance structures. Without loss of generality, \( \Omega_1^{-1} \) and \( \Omega_2^{-1} \) can be written as

(4.4.3) \begin{align*}
\Omega_1^{-1} &= d_1 I + d_1 \gamma_1 (J-I) \\
\text{and} \quad \Omega_2^{-1} &= d_2 I + d_2 \gamma_2 (J-I).
\end{align*}

The resulting log-likelihood is

(4.4.4) \begin{align*}
L(\hat{\beta}, d_1, d_2, \gamma_1, \gamma_2) &= -\frac{1}{2} N \ln(2\pi) + \frac{1}{2} N \ln(d_1) + \frac{1}{2} N(p-1) \ln(1-\gamma_1) + \frac{1}{2} N \ln[(p-1)\gamma_1] \\
&\quad + \frac{1}{2} N(t-1)p \ln(d_2) + \frac{1}{2} N(t-1)(p-1) \ln(1-\gamma_2) + \frac{1}{2} N(t-1) \ln[(p-1)\gamma_2] \\
&\quad - \frac{1}{2} d_1 u_1 u_2 - \frac{1}{2} d_1 \gamma_1 u_2 - \frac{1}{2} d_2 v_1 v_2 - \frac{1}{2} d_2 \gamma_2 v_2,
\end{align*}

where \( u_1 = \sum_{j=1}^{p} T_1(j,j) \), \( u_2 = \sum_{i\neq j}^{p} T_1(i,j) \)

(4.4.4) \begin{align*}
&\quad - \frac{1}{2} d_1 u_1 u_2 - \frac{1}{2} d_1 \gamma_1 u_2 - \frac{1}{2} d_2 v_1 v_2 - \frac{1}{2} d_2 \gamma_2 v_2,
\end{align*}

where \( u_1 = \sum_{j=1}^{p} T_1(j,j) \), \( u_2 = \sum_{i\neq j}^{p} T_1(i,j) \)

(4.4.4) \begin{align*}
&\quad - \frac{1}{2} d_1 u_1 u_2 - \frac{1}{2} d_1 \gamma_1 u_2 - \frac{1}{2} d_2 v_1 v_2 - \frac{1}{2} d_2 \gamma_2 v_2,
\end{align*}

where \( v_1 = \sum_{j=1}^{p} T_2(j,j) \), \( v_2 = \sum_{i\neq j}^{p} T_2(i,j) \).
Solving the partial equations of the above log-likelihood yields the estimates

\[
\hat{d}_1 = \frac{N\{p(1)u_1 + (p-2)u_2\}}{\{(p-1)u_1 - u_2\}[u_1 + u_2]}, \quad \hat{d}_2 = \frac{N(t-1)p\{(p-1)v_1 + (p-2)v_2\}}{\{(p-1)v_1 - v_2\}[v_1 + v_2]},
\]

\[
\hat{Y}_1 = -\frac{u_2}{(p-1)u_1 + (p-2)u_2}, \quad \hat{Y}_2 = -\frac{v_2}{(p-1)v_1 + (p-2)v_2}.
\]

We then have the following results.

**Theorem 4.4.1.** Under the hypothesis \(H_3\), the MLE's of \(\Sigma_1\) and \(\Sigma_2\) are

\[
\hat{\Sigma}_1 = \frac{1}{t}[\hat{\Omega}_1 + (t-1)\hat{\Omega}_2] \text{ and } \hat{\Sigma}_2 = \frac{1}{t}[\hat{\Omega}_1 - \hat{\Omega}_2],
\]

where

\[
\hat{\Omega}_1^{-1} = \hat{d}_1 I + \hat{d}_1 \hat{Y}_1 (J-I) \text{ and } \hat{\Omega}_2^{-1} = \hat{d}_2 I + \hat{d}_2 \hat{Y}_2 (J-I).
\]

The resulting maximum log-likelihood is

\[
L_3 = -\frac{1}{2}N\ln(2\pi) + \frac{1}{2}N\ln(Np) + \frac{1}{2}N(t-1)\ln(p(t-1)) + \frac{1}{2}N(t-1)\ln\ln(t-1)
\]

\[
-\frac{1}{2}N\ln(p-1)\ln\{p(1)u_1 - u_2\} - \frac{1}{2}N\ln[u_1 + u_2]
\]

\[
-\frac{1}{2}N(t-1)(p-1)\ln\{p(1)v_1 - v_2\} - \frac{1}{2}N(t-1)\ln[v_1 + v_2].
\]

**Corollary 4.4.2.** The LR test statistic used in testing \(H_3\) vs. \(H_1\) is

\[
\Lambda_3 = \exp \left( L_3 - L_1 \right) = \Lambda_{31} \Lambda_{32}, \text{ where }
\]

\[
\Lambda_{31} = \left\{ \left( \frac{(p-1)}{(p-1)u_1 - u_2} \right) \left[ \frac{|T_1|}{\{u_1 + u_2\}} \right] \right\} \quad \Lambda_{32} = \left\{ \left( \frac{(p-1)}{(p-1)v_1 - v_2} \right) \left[ \frac{|T_2|}{\{v_1 + v_2\}} \right] \right\}
\]

\[
H_3 \text{ is rejected for small values of } \Lambda_3.
\]
4.4.1. The Asymptotic Null Distribution of the LR Statistics

The statistics, $\Lambda_{31}$ and $\Lambda_{32}$, given by (4.4.10) have the same form as the LR statistic $\Lambda_1$ in Section 4.2. Applying Lemma 4.2.4 to $\Lambda_{31}$ and $\Lambda_{32}$, we have

Lemma 4.4.3. Under $H_3$, the $h^{th}$ moment of $\Lambda_3$ is

\[
E(\Lambda_3^h) = k_0 (p-1)^{(p-1)h} \prod_{j=1}^{p} \frac{\Gamma[\frac{k_0}{2}(1+h)+\frac{1}{2}(1-q-j)]}{\Gamma[\frac{k_0}{2}(1+h)-\frac{1}{2}(1-q-j)]} \prod_{j=1}^{p} \frac{\Gamma[\frac{k_0}{2}N(t-1)(1+h)+\frac{1}{2}(1-q-j)]}{\Gamma[\frac{k_0}{2}N(t-1)(1+h)-\frac{1}{2}(1-q-j)]},
\]

where $k_0$ is a constant not involving $h$.

The above moment has the same form as (4.1.6) with

\[
\begin{align*}
a &= 2p, \quad x_j = \frac{k_0}{2}N, \quad e_j = \frac{k_0}{2}(1-q-j), \quad j = 1,2,\ldots,p, \\
x_j &= \frac{k_0}{2}N(t-1), \quad e_j = \frac{k_0}{2}(1-q-p-j), \quad j = p+1,\ldots,2p, \\
b &= 4, \quad y_1 = \frac{k_0}{2}N, \quad y_2 = \frac{k_0}{2}N(p-1), \quad y_3 = \frac{k_0}{2}N(t-1), \quad y_4 = \frac{k_0}{2}N(t-1)(p-1), \\
f_1 &= f_3 = -\frac{1}{2}q, \quad f_2 = f_4 = -\frac{1}{2}(p-1)q.
\end{align*}
\]

The asymptotic distribution of $\Lambda_3$ can then be derived by using Theorem 4.4.1.

Theorem 4.4.4. When the null hypothesis $H_3$ is true, the distribution of $-2p \ln \Lambda_3$ can be expanded for large value of $M = \rho N t$ as

\[
p(-2p \ln \Lambda_3 \leq z) = p(\chi^2_f \leq z) + o(M^{-2}),
\]

where

\[
f = p^2+p-4,
\]

\[
\rho = 1 - \frac{1}{f} \left[ \sum_{j=1}^{2p} x_j^{-1} (e_j^2 - e_j + \frac{1}{6}) \right] - \sum_{k=1}^{4} y_k^{-1} (f_k^2 - f_k + \frac{1}{6}) \].
4.4.2. Asymptotic Non-null Distributions of the LR Statistic when p=2

When p=2,

\( \Lambda_{31} = \left\{ \frac{|T_1|}{\sqrt{2(v_1-u_2)(v_1+u_2)}} \right\}^{\frac{N}{2}} \), 
\( \Lambda_{32} = \left\{ \frac{|T_2|}{\sqrt{2(v_1-v_2)(v_1+v_2)}} \right\}^{\frac{N}{2}} \),

where \( f = 2 \), and \( \rho = 1 - \frac{t}{2N(t-1)} (q + \frac{3}{2}) \).

Applying Theorem 4.2.3 to \( T_1 \) and \( T_2 \),

\( \Lambda_{31} = \left\{ \frac{|V_1|}{\sqrt{V_1(1,1)V_1(2,2)}} \right\}^{\frac{N}{2}} \) and \( \Lambda_{32} = \left\{ \frac{|V_2|}{\sqrt{V_2(1,1)V_2(2,2)}} \right\}^{\frac{N}{2}} \),

where \( V_1 = \Gamma_2 \text{'} T_1 \Gamma_2 \) and \( V_2 = \Gamma_2 \text{'} T_2 \Gamma_2 \) with \( \Gamma_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \).

If we denote \( \Psi_1 = \Gamma_2 \text{'} \Omega_1 \Gamma_2 \) and \( \Psi_2 = \Gamma_2 \text{'} \Omega_2 \Gamma_2 \), then \( \Psi_1 \) and \( \Psi_2 \) are diagonal matrices under \( H_3 \). We also express \( P^2 \) as \( \text{Diag}(\Psi_1, \Psi_2) \), where

\( \Psi_1 = \frac{[\Psi_1(1,2)]^2}{\Psi_1(1,1)\Psi_1(2,2)} \) and \( \Psi_2 = \frac{[\Psi_2(1,2)]^2}{\Psi_2(1,1)\Psi_2(2,2)} \).

Two types of alternatives are considered here.

(A) A sequence of local alternatives \( H_M^\rho \): \( P^2 = \frac{1}{M^2} \text{Diag}(\theta_1, \theta_2) \), where \( \theta_1 \) and \( \theta_2 \) are fixed constants.

(B) A fixed alternative \( H_\kappa^\rho \): \( P^2 \neq 0 \).

Under \( H_3 \), the characteristic function of \(-2\rho \ln \Lambda_3\) is

\( \phi(M,s,\Psi_1,\Psi_2) = \text{E} \{ \exp[is(-2\rho \ln \Lambda_3)] \} \)

\( = \text{E} \{ \Lambda_3^{-2is\rho} \} \)

\( = \text{E} \{ \Lambda_{31}^{-2is\rho} \} \text{E} \{ \Lambda_{32}^{-2is\rho} \} \).
where $\Lambda_{31}$ and $\Lambda_{32}$ are statistics for testing independence between two variables. The following result from Sugiura and Fujikoshi (1969) was used to derive the above characteristic function.

**Lemma 4.4.5.** Given $A(2\times2)$, a random matrix with a Wishart distribution, $W_2(n,\Sigma)$, where $A = (a_{ij})$ and $\Sigma(2\times2) = (\sigma_{ij})$, and denoting

$$A = \frac{|A|}{a_{11} a_{22}},$$

then the $h$th moment of $A$ is

$$(4.4.19) \quad E(A^n) = \frac{\Gamma(\frac{1}{2}m) \Gamma(\frac{1}{2}(n-h)+\frac{1}{2})}{\Gamma(\frac{1}{2}k+n-h+\frac{1}{2}) \Gamma(\frac{1}{2}(n-k-1)+\frac{1}{2})} \left(1-\gamma\right)^{\frac{n}{2}} \frac{n}{2} F_1 \left(\frac{1}{2}m, \frac{1}{2}k; \frac{1}{2}h+\frac{1}{2}; \gamma\right).$$

In (4.4.19) $\gamma = \frac{\sigma_{12}^2}{\sigma_{11} \sigma_{22}}$, and $2 F_1$ is a generalized hypergeometric function defined by

$$(4.4.20) \quad 2 F_1 (a_1, a_2; b_1; \gamma) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k}{(b_1)_k} \frac{\gamma^k}{k!},$$

with $(a)_k = \frac{(a+k-1)!}{(a-1)!}.$

Applying Lemma 4.4.5 to the statistics $\Lambda_{31}$ and $\Lambda_{32}$,

$$(4.4.21) \quad \phi(M, s, \psi_1, \psi_2) = E\{A^{-2is} \} \cdot E\{A^{-2is} \}$$

$$= \phi(M, s, 0, 0) \ G_1(M, s, \psi_1) \ G_2(M, s, \psi_2),$$

where

$$(4.4.22) \quad \phi(M, s, 0, 0)$$

is the characteristic function of $-2s \ln \Lambda_3$.

Under $H_3$: $\psi_1 = \psi_2 = 0$,

$$G_1(M, s, \psi_1) = (1-\psi_1)^{\frac{1}{2}M+\frac{1}{2}} \frac{1}{2} F_1 \left(\frac{1}{2}t M+\sigma; \frac{1}{2}t M+\frac{\sigma}{2}; \frac{1}{2}t M+\frac{\sigma}{2}; \frac{1}{2}t M+\frac{\sigma}{2} \right) \text{ is } \frac{1}{2} M; \psi_1),$$

and
\[
G_2(M, s, \psi_2) = (1 - \psi_2)^{\frac{(t-1)M + (t-1)\sigma}{2t}}
\]

\[
2F_1\left(\frac{(t-1)M + (t-1)\sigma}{2t}, \frac{(t-1)M + (t-1)\sigma}{2t}, \frac{(t-1)M + (t-1)\sigma}{2t}, \frac{is(t-1)M}{t}; \psi_2\right).
\]

Under \( H_3 \), \(-2p \ln \Lambda_3\) has an asymptotic \( \chi^2 \) distribution with 2 d.f.

Thus,
\[
(4.4.23) \quad \phi(M, s, 0, 0) = \left(1 - \frac{2i\sigma}{\sigma}\right)^{\frac{1}{2}} + 0(M^{-2}) = (1 - 2i\sigma)^{-1} + 0(M^{-2}).
\]

Using the asymptotic expansion of the generalized hypergeometric function \( 2F_1 \) given by Muirhead (1982, page 724), we have the following expansions.

\[
(4.4.24) \quad G_1(M, s, \psi_1) = \exp\left\{ \frac{1}{2t} \left(1 - \frac{1}{2i\sigma}\right) \left(\theta_1\right)\right\} [1 + 0(M^{-2})]
\]

\[
= 1 - \frac{\theta_1}{2t} (1 - \frac{1}{1 - 2i\sigma}) M^{-1} + 0(M^{-2}),
\]

and
\[
G_2(M, s, \psi_2) = 1 - \frac{(t-1)\theta_2}{2t} (1 - \frac{1}{1 - 2i\sigma}) M^{-1} + 0(M^{-2}).
\]

Combining the above results,
\[
(4.4.25) \quad \phi(M, s, \psi_1, \psi_2) = \left(1 - 2i\sigma\right)^{\frac{1}{2}} \left(-\frac{\theta_1 + (t-1)\theta_2}{2tM}\right)
\]

\[
\cdot \left[ (1 - 2i\sigma)^{-\frac{f}{2}} - (1 - 2i\sigma)^{-\frac{f+2}{2}} \right] + 0(M^{-2}).
\]

Inverting this expansion, we have the following result:

**Theorem 4.4.6.** Under the sequence of local alternatives

\( H_M: p^2 = \frac{1}{M^2} \text{Diag}(\theta_1, \theta_2) \), the distribution function of \(-2p \ln \Lambda_3\) can be expanded as

\[
(4.4.26) \quad p(-2p \ln \Lambda_3 \leq z) = p(X_f^2 \leq z) - \frac{\theta_1 + (t-1)\theta_2}{2tM} [p(X_{f+2}^2 \leq z) - p(X_f^2 \leq z)]
\]

\[+ 0(M^{-2}).\]
Let $M_1 = N_0 = \frac{1}{t} M$, then from Muirhead (1982), the statistic

$$-2\rho M^{-1} \ln \Lambda_1$$

has the asymptotic mean $-\ln(1-\psi_1)$, and the asymptotic variance of order $O(M^{-1}) = 0(M^{-1})$. Similarly, for $M_2 = N(t-1)\rho = \frac{t-1}{t} M$, the statistic $-2\rho M_2^{-1} \ln \Lambda_2$ has the asymptotic mean $-\ln(1-\psi_2)$, and the asymptotic variance of order $O(M_2^{-1}) = 0(M^{-1})$. Thus, the statistic

$$(-4.4.27) -2\rho M^{-1} \ln \Lambda_3 = M^{-1}(M_1[-2\rho M_{1}^{-1} \ln \Lambda_{31}] + M_2[-2\rho M_{2}^{-1} \ln \Lambda_{32}])$$

has the asymptotic mean

$$M^{-1}(M_1[-\ln(1-\psi_1)] + M_2[-\ln(1-\psi_2)]) = -\frac{1}{t}(-\ln(1-\psi_1) + (t-1)(1-\psi_2)),$$

and the asymptotic variance of order $O(M^{-1})$. The variable, $Z$, defined by

$$(-4.4.28) Z = -2\rho M^{-1} \ln \Lambda_3 + \frac{1}{t} \{\ln(1-\psi_1) + (t-1) \ln(1-\psi_2)]},$$

has the asymptotic mean 0, and the asymptotic variance of order $O(1)$. The characteristic function of $Z$ is

$$(-4.4.29) \phi_Z(s) = E\{\exp(isZ)\}$$

$$= E\{\exp[-2is\rho M^{-1} \ln \Lambda_3]\} \cdot \exp\{isM^{-1}\frac{1}{t} \{\ln(1-\psi_1) + (t-1) \ln(1-\psi_2)]\}$$

$$= \phi(M, sM^{-1}, \psi_1, \psi_2) \cdot [(1-\psi_1)(1-\psi_2)\frac{1}{t}] is\frac{1}{t} M^{\frac{1}{2}}$$

$$= \phi(M, sM^{-1}, 0, 0) \cdot (T_1(M, sM^{-1}, \psi_1) G_2(M, sM^{-1}, \psi_2)$$

$$\cdot (1-\psi_1) is\frac{1}{t} M^{\frac{1}{2}} \cdot (1-\psi_2) is\frac{1}{t} \frac{(t-1)}{M^{\frac{1}{2}}},$$

where the functions $\phi$, $G_1$, and $G_2$ are given by (4.4.21) - (4.4.22).

To simplify (4.4.29), we use the following two lemmas.

Lemma 4.4.7. The hypergeometric function, \( {}_2F_1(a_1, a_2; b_1; \gamma) \) is equal to

$$(-4.4.30) {}_2F_1(b_1-a_1, b_1-a_2; b_1; \gamma).$$
Lemma 4.4.8. [From Sugiura and Fujikoshi (1969)]

The hypergeometric function \( _2F_1(-isM^\frac{1}{2}, -isM^\frac{1}{2}; \frac{1}{2}(M+\sigma) - isM^\frac{1}{2}, \gamma) \) has the asymptotic expansion

\[
(4.4.31) \quad \exp(-2s^2\gamma) \{1+4M^{-\frac{1}{2}}(is)^3 \gamma(1-\gamma) + O(M^{-1})\}
\]

The expansion of the characteristic function of \( Z \) is therefore:

\[
(4.4.32) \quad \phi_Z(s) = \exp\{-2s^2\frac{1}{\tau} [\psi_1+(t-1)\psi_2]\} \\
\cdot [1+M^{-\frac{1}{2}}(2is + \frac{4}{\tau} [\psi_1(1-\psi_1) + (t-1)\psi_2(1-\psi_2)](is)^3)] + O(M^{-1}).
\]

Letting

\[
(4.4.33) \quad \tau^2 = 4 \frac{[\psi_1+(t-1)\psi_2]}{\frac{s^2}{\tau}},
\]

we have

\[
(4.4.34) \quad \phi_Z(s) = \phi_Z\left(\frac{s}{\tau}\right) \\
= \exp\{-\frac{1}{2}s^2\{1+M^{-\frac{1}{2}}[a_1 \frac{1}{\tau} (is) + a_2 \frac{1}{\tau^3} (is)^3]\} + O(M^{-1}),
\]

where

\[
a_1 = 2, \quad a_2 = 4 \frac{1}{\tau} [\psi_1(1-\psi_1) + (t-1)\psi_2(1-\psi_2)].
\]

Inverting (4.4.34), and using the fact that

\[
\int e^{isx} \phi^{(k)}(x) dx = -(is)^k e^{-\frac{s^2}{2}} \quad \text{for} \quad k = 1,3,
\]

where \( \phi^{(k)}(x) \) is the \( k \)th derivative of the standard normal density, we have

Theorem 4.4.9. Under the fixed alternative \( H_k: \tau^2 \neq 0 \), the distribution function of the random variable \( Z \) given by (4.4.28) can be expanded as

\[
(4.4.35) \quad P\left(\frac{Z}{\tau} \leq z\right) = \Phi(z) - M^{-\frac{1}{2}}[a_1 \frac{1}{\tau} \phi(z) + a_2 \frac{1}{\tau^3} \phi^{(2)}(z)] + O(M^{-1}),
\]

where \( \Phi \) and \( \phi \) denote the standard normal distribution and density functions, respectively. \( \tau^2 \) is given by (4.4.33), \( a_1 \) and \( a_2 \) are given by (4.4.34).
4.5. Testing Equality of Elements of $\Sigma_2$ in $H_3$:

$H_4$: $\Sigma_1 = \sigma_1[I + \rho_1(J-I)], \Sigma_2 = \sigma_1\omega J$.

Under the hypothesis $H_4$,

\[(4.5.1)\quad \Omega_1 = \Sigma_1 + (t-1)\Sigma_2 = \sigma_1\{(1+(t-1)\omega)I + [\rho_1+(t-1)\omega](J-I)\}\]

and

\[(4.5.2)\quad \Omega_2 = \Sigma_1 - \Sigma_2 = \sigma_1\{(1-\omega)I + (\rho_1-\omega)(J-I)\}\]

Denoting

\[(4.5.3)\quad \Omega_1^{-1} = d((1-\gamma_1)I + \gamma_1(J-I))\]

and

\[(4.5.4)\quad \Omega_2^{-1} = d((1-\gamma_2)I - \gamma_2(J-I))\]

The log-likelihood, maximized with respect to $\beta$, is therefore

\[(4.5.5)\quad L(\hat{\beta}, d, \gamma_1, \gamma_2) = \frac{t}{2} \ln (2\pi) + \frac{1}{2} \ln (d) + \frac{1}{2} \ln (1-\rho_1) + \frac{1}{2} \ln (1-\rho_2)\]

where

\[u_1 = \frac{1}{p} \sum_{j=1}^{p} T_1(j,j), u_2 = \frac{1}{p} \sum_{i \neq j} T_1(i,j), v_1 = \frac{1}{p} \sum_{j=1}^{p} T_2(j,j), v_2 = \frac{1}{p} \sum_{i \neq j} T_2(i,j)\]

Solving the partial equations of the above log-likelihood yields

\[(4.5.6)\quad d = \frac{N\rho(p-1)}{[(p-1)u_1-u_2] + [(p-1)v_1-v_2]}\]

\[\gamma_1 = \frac{1}{p} - \frac{N}{u_1+u_2} \frac{1}{d} = \frac{1}{p} \left\{1 - \frac{[(p-1)u_1-u_2] + [(p-1)v_1-v_2]}{t(p-1)(u_1+u_2)}\right\}\]
\[ \hat{\gamma}_2 = \frac{1}{p} \cdot \frac{N(t-1)}{v_1 + v_2} \quad \hat{d} = \frac{1}{p} \left\{ 1 - \frac{(t-1)}{t(p-1)} \cdot \frac{[(p-1)u_1 - u_2] + [(p-1)v_1 - v_2]}{v_1 + v_2} \right\} \]

We then have the following results.

**Theorem 4.5.1.** Under the hypothesis \( H_4 \), the MLE's of \( \Sigma_1 \) and \( \Sigma_2 \) are

\[ \hat{\Sigma}_1 = \frac{1}{t} \left[ \hat{\Omega}_1 + (t-1)\hat{\Omega}_2 \right] \quad \text{and} \quad \hat{\Sigma}_2 = \frac{1}{t} \left[ \hat{\Omega}_1 + \hat{\Omega}_2 \right], \quad \text{where} \]

\[ \hat{\Omega}_1^{-1} = \hat{d} \left[ (1-\hat{\gamma}_1)I - \hat{\gamma}_1(J-I) \right] \quad \text{and} \quad \hat{\Omega}_2^{-1} = \hat{d} \left[ (1-\hat{\gamma}_2)I - \hat{\gamma}_2(J-I) \right]. \]

The resulting maximum log-likelihood is

\[ L_4 = \frac{-\frac{1}{2}N\text{tr} \ln(2\pi) + \frac{1}{2}N\text{tr} \ln[np] + \frac{1}{2}N\text{tr}[(p-1)\ln[t(p-1)] + \frac{1}{2}N\text{tr}[(t-1)]}{\frac{1}{2}N\text{tr}[(p-1)\ln[(p-1)u_1 - u_2] + [(p-1)v_1 - v_2]]} \]

\[ -\frac{1}{2}N\text{tr}[u_1 + u_2] + \frac{1}{2}N\text{tr}[v_1 + v_2]. \]

**Corollary 4.5.2.** The LR test statistic for testing \( H_4 \) vs. \( H_3 \) is given by

\[ \Lambda_4 = \exp(L_4 - L_3) = \frac{\frac{N}{2}t(p-1)}{\frac{N}{2}(t-1)(p-1)} \cdot \frac{W_1}{W_2}, \]

where \( W_1 = \frac{1}{p} \left[ [(p-1)u_1 - u_2] \right], \quad W_2 = \frac{1}{p} \left[ [(p-1)v_1 - v_2] \right]. \)

\( H_3 \) is rejected for small values of \( \Lambda_4 \).

Under the hypothesis \( H_3 \), both \( \Omega_1 \) and \( \Omega_2 \) have intraclass covariance structures. Thus, there exists an orthogonal transformation \( \Gamma_p(p \times p) \), such that

\[ \Psi_1(p \times p) = \Gamma_p' \Omega_1 \Gamma_p = \text{Diag}(\psi_{11}, \psi_{12}, \ldots, \psi_{12}) \quad \text{and} \]

\[ \Psi_2(p \times p) = \Gamma_p' \Omega_2 \Gamma_p = \text{Diag}(\psi_{21}, \psi_{22}, \ldots, \psi_{22}). \]
Under \( H_3 \), \( T_1 \) has a Wishart \( \mathcal{W}_p(n, \Omega_1) \) distribution.

Denoting \( Q_1(p \times p) = \frac{\Gamma'(p-1)}{\Gamma_p} \), and applying Theorem 4.2.3,

\[
(4.5.11) \quad W_1 = \frac{1}{p} \left[ (p-1)u_1 - u_2 \right] = \sum_{j=2}^{p} Q_1(j,j)
\]

has a Wishart \( \mathcal{W}_1(n(p-1), \psi_{12}) \) distribution.

Similarly,

\[
(4.5.12) \quad W_2 = \frac{1}{p} \left[ (p-1)v_1 - v_2 \right] \text{ has a Wishart } \mathcal{W}_1(n(t-1)(p-1), \psi_{22}) \text{ distribution.}
\]

Further applying Theorem 4.2.3 to \( \Omega_j, \psi_j, j = 1,2 \), yields

\[
(4.5.13) \quad \psi_{12} = \psi_{22} = \sigma_1(1-\rho_1) \text{ under } H_4.
\]

Since \( W_1 \) and \( W_2 \) have different degrees of freedom, the LR test statistic \( \Lambda_4 \) of Lemma 4.5.2 is biased. That is, the probability of rejecting \( H_4 \) when \( H_4 \) is false can be smaller than the probability of rejecting \( H_4 \) when \( H_4 \) is true.

We therefore use the following modified LR statistic:

\[
(4.5.14) \quad \Lambda_4^* = \Lambda_4 = \frac{n \frac{n}{2} (p-1)}{\frac{n}{2}(t-1)(p-1)} \frac{W_1^{n/2(p-1)-1} W_2^{n/2(t-1)(p-1)-1}}{(W_1 + W_2)^{nt/2-1}}.
\]

4.5.1. The Asymptotic Null Distribution of the Modified LR Statistic

When the hypothesis \( H_4 \) is true, the \( h \text{th} \) moment of \( \Lambda_4^* \) is

\[
(4.5.15) \quad E[(\Lambda_4^*)^h] = \frac{\frac{n}{2}(p-1)^h}{(t-1)^h \frac{n}{2}(t-1)(p-1)^h} E \left\{ \frac{W_1^{n/2(p-1)-1} W_2^{n/2(t-1)(p-1)-1}}{(W_1 + W_2)^{nt/2-1}} \right\}.
\]
\[
\frac{\text{nt}(p-1)h}{(t-1)\text{nt}(t-1)(p-1)h} = \frac{(2\psi)^{\text{nt}(p-1)h}}{\Gamma[\text{nt}(p-1)(1+h)] \Gamma[\text{nt}(t-1)(p-1)(1+h)]/\Gamma[\text{nt}(p-1)] \Gamma[\text{nt}(t-1)(p-1)]}
\]
\[
= k_0 \frac{\text{nt}(p-1)h}{(t-1)\text{nt}(t-1)(p-1)h} \frac{\Gamma[\text{nt}(p-1)(1+h)] \Gamma[\text{nt}(t-1)(p-1)(1+h)]}{\Gamma[\text{nt}(p-1)(1+h)]},
\]
where \(k_0\) is a constant not involving \(h\).

This \(h\)th moment has the same form as (4.1.6) with

\[(4.5.16)\quad a = 2, \quad x_1 = \text{nt}(p-1), \quad x_2 = \text{nt}(t-1)(p-1), \quad e_1 = e_2 = 0, \quad b = 1, \quad y_1 = \text{nt}(p-1), \quad f_1 = 0.\]

Applying Theorem 4.4.1, we have

**Theorem 4.5.3.** When the hypothesis \(H_4\) is true, the distribution function

\[
p(-2\rho \ln \Lambda_4^*)
\]

can be expanded for large \(M = \rho nt p\) as

\[(4.5.17)\quad p(-2\rho \ln \Lambda_4^* \leq z) = p(X_1^2 \leq z) + O(M^{-2}),\]

where

\[
\rho = 1 - \frac{1}{3n(p-1)} \frac{(t^2-t-1)}{t(t-1)}.
\]

4.5.2. Asymptotic Non-null Distributions of the Modified LR Statistic

If we denote

\[(4.5.18)\quad \Delta = \frac{\psi_12}{\psi_22},\]

where \(\psi_12\) and \(\psi_22\) are given by (4.5.10), then \(\Lambda = 1\) under the hypothesis \(H_4\). In this section, the following alternatives are considered:

(A) A general fixed alternative \(H_k: \Delta \neq 1\).

(B) A sequence of local alternatives \(H_M: \Delta = 1 + \frac{1}{M} \theta\), where \(\theta\) is a fixed constant and \(M = \rho nt p\).
Define the random variable $Z$ by

$$Z = -2\rho \ln A_4^* M^{-1/2} + \ln \left[ \frac{\Delta}{k_1 k_2} \right],$$

where

$$k_2 = \frac{n(p-1)}{n(p-1)+n(t-1)(p-1)} = \frac{1}{t},$$

$$k_2 = 1 - k_1 = \frac{t-1}{t},$$

and $\rho$ is given by (4.5.17).

Then, from Muirhead (1982), we have the following theorems.

**Theorem 4.5.4.** Under the fixed alternative, $H_k: \Delta \neq 1$, the distribution function of the random variable $Z$ can be expanded asymptotically as

$$p\left( \frac{Z}{\rho} \leq z \right) = \Phi(z) + \frac{1}{\mu^2} \left[ a_1 \phi(z) - a_2 \phi''(z) \right] + O(M^{-1}),$$

where $\Phi$ and $\phi$ denote the standard normal distribution and density functions respectively, and

$$\tau^2 = k_2 k_1 k_2 \sigma^2,$$

$$a_1 = \frac{1}{\tau} \left( k_1 k_2 (\sigma^2 \phi^2(z)) - z \right) = \frac{1}{\tau} \left[ k_1 k_2 - 1 \right],$$

$$a_2 = \frac{2k_1 k_2}{\tau^2} \left[ \frac{1}{3} \sigma^2 + \frac{2}{3} (k_1 - k_2) \sigma^3 - k_1 k_2 \sigma^4 \right],$$

$$\sigma = \frac{1-\Delta}{k_1 k_2}.$$

**Theorem 4.5.5.** Under the sequence of local alternatives $H_M: \Delta = 1 + \frac{\theta}{M}$, the distribution function of $-2\rho \ln A_4^*$ can be expanded as

$$p(-2\rho \ln A_4^* \leq z) = p(x_1^2 \leq z) + \frac{(t-1)\theta}{4M^2} \left[ p(x_3^2 \leq x) - p(x_1^2 \leq z) \right] + O(M^{-1}).$$
4.6. Numerical Examples

The data set used in this section is from the Prevalence Study of the Lipid Research Clinics Program. This data set consists of 57 white participants (36 males, 21 females) between ages 50 and 59 from one of the clinic centers. For each participant, the variables used are sex, plasma total cholesterol levels (mg/dl) at Visit 1 (CHOL1) and Visit 2 (CHOL2), and plasma triglyceride levels (mg/dl) at Visit 1 (TRIG1) and Visit 2 (TRIG2).

Assuming standard MGLM for the data, let

\[ Y_0 (57 \times 4) \] is the observed data on CHOL1, TRIG1, CHOL2, TRIG2,

\[ X(57 \times 4) = \begin{bmatrix} 1 & 36 & 0 \\ 0 & 1 & 21 \end{bmatrix} \] is the design matrix,

\[ \beta(2 \times 4) \] is the matrix of unknown parameters, and

\[ \Sigma_0 (4 \times 4) \] is the unknown positive definite covariance matrix.

Under \( H_0 \), then log-likelihood is

\[
L(\beta, \Sigma_0) = -\frac{1}{2} N tp \ln(2\pi) - \frac{1}{2} N \ln(|\Sigma_0|) - \frac{1}{2} \text{trace}(\Sigma_0^{-1}(Y-X\beta)'(Y-X\beta)),
\]

where \( N = 57, t = p = 2. \)

The MLE's of \( \beta \) and \( \Sigma_0 \) are

\[
\hat{\beta} = \begin{bmatrix} 228.667 \\ 235.048 \end{bmatrix}, \quad \hat{\Sigma} = \begin{bmatrix} 1153.56 & -797.472 & 1024.38 & -768.868 \\ -797.472 & 7503.29 & -1272.68 & 6068.37 \\ 1024.38 & -1272.68 & 1357.54 & -917.09 \\ -768.868 & 6068.37 & -917.09 & 7506.74 \end{bmatrix}.
\]

The resulting maximum log-likelihood is \( L_0 = -1167.72. \)
4.6.1. Testing Block Intraclass Covariance Structure for $\Sigma_0$

Under the hypothesis $H_1: \Sigma = \begin{pmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_2 & \Sigma_1 \end{pmatrix}$, where $\Sigma_1$ and $\Sigma_2$ are $2 \times 2$ matrices, the MLE's of $\Sigma_1$ and $\Sigma_2$ are

$$\hat{\Sigma}_1 = \begin{pmatrix} 1255.55 & -857.281 \\ -857.281 & 7405.01 \end{pmatrix} \quad \text{and} \quad \hat{\Sigma}_2 = \begin{pmatrix} 1024.38 & -1020.77 \\ 1020.77 & 6068.37 \end{pmatrix}.$$  

The resulting maximum log-likelihood is $L_1 = -1169.08$, and the LR statistic for testing $H_1$ vs. $H_0$ is

$$\Lambda_1 = \exp(L_1 - L_0) = 0.2555565.$$  

From Section 4.2.1. with $\rho = .921053$, $M = n\rho = 52.5$, the distribution function of $-2\rho \ln \Lambda$ can be expanded for large $M$ as

$$p(-2\rho \ln \Lambda \leq z) = p(x^2_f \leq z) + O(M^{-2}), \text{ where } f = p^2 = 4.$$  

For $\Lambda_1 = 0.2555565$, $-2\rho \ln \Lambda_1 = 2.51315$, and the asymptotic p-value of the test is .642283. Thus, we accept the hypothesis $H_1$ of block intraclass covariance structure for $\Sigma_0$.

When the hypothesis $H_1$ is not true, we first consider the fixed alternative $H_k: \Sigma = \begin{pmatrix} 12 & -8 & 10 & -8 \\ 72 & -12 & 60 \\ (\text{Sym.}) & 14 & -9 \\ \end{pmatrix}$

According to Section 4.4.2.,

$$p^2 = \begin{pmatrix} 0.0241015 & 0 \\ 0 & 0.0038207 \end{pmatrix}, \quad t = .334199, \text{ and}$$

the statistic $Z = \frac{2\rho \ln \Lambda_1}{M^2} + M^2 \ln(1 - p^2) = 0.14234$. 

Thus, at the observed \( z = \frac{\bar{Z}}{t} = 0.425914 \), the asymptotic power is

\[
(4.4.6) \quad P\left( \frac{\bar{Z}}{t} > z \right) = 1 - \Phi(z) + \frac{1}{M^2} \phi(z) + \frac{4}{\tau^3} (\sigma_1 - \sigma_2) \phi(2)(z) = 0.716401.
\]

Now consider the local alternative \( H_M : p^2 = \frac{1}{M^2} I_2 \). Under \( H_M \), the asymptotic power at the observed statistic \( z = -2p \ln \Lambda_1 = 2.51315 \) is

\[
(4.4.7) \quad P(-2p \ln \Lambda_1 > z) = 1 - \left\{ p\left( \chi_f^2 \leq z \right) + \frac{\sigma_1}{M} \left[ p\left( \chi_f^2 < z \right) - p\left( \chi_f^2 \leq z \right) \right] \right\} = 0.650844,
\]

which is slightly larger than the p-value of the statistic \(-2p \ln \Lambda_1 \) under the null hypothesis \( H_1 \).

4.6.2. Testing Proportionality for \( \Sigma_1 \) and \( \Sigma_2 \) in \( H_1 \)

Under the hypothesis \( H_2 : \Sigma_2 = \omega \Sigma_1 \), the MLE's of \( \omega \) and \( \Sigma_1 \) are estimated numerically as

\[
(4.6.8) \quad \hat{\omega} = 0.815071 \quad \text{and} \quad \hat{\Sigma}_1 = \begin{bmatrix} 1253.08 & -75.3128 \\ -75.3128 & 7325.45 \end{bmatrix}.
\]

The resulting maximum log-likelihood is \( L_2 = -1175.11 \).

The MLE's of \( \theta_N \) and \( d_N \) are

\[
(4.6.9) \quad \hat{\theta}_N = \begin{bmatrix} 0.770236 \\ -0.112842 \\ 0.801107 \end{bmatrix},
\]

\[
\hat{d}_N = \begin{bmatrix} 0.0293072 \\ 0.00243191 \\ 0.0118626 \end{bmatrix}.
\]

For \( C = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \),

\[
(4.6.10) \quad C \begin{bmatrix} \hat{\theta}_N \\ \hat{d}_N \end{bmatrix} = \begin{bmatrix} -0.308711 \\ -0.112842 \end{bmatrix}.
\]
The jackknife estimate of the variance of \( C \begin{bmatrix} \hat{\theta}_N \\ \hat{d}_N \end{bmatrix} \) is

\[
\frac{1}{N} C I^{-1}_{0J}(\hat{\theta}, \hat{d}) C' = .0001 \begin{bmatrix} 16.3036 & 6.72915 \\ 6.72915 & 7.53537 \end{bmatrix}.
\]

The Wald Statistic has the value

\[
W_N = N \left[ C \begin{bmatrix} \hat{\theta}_N \\ \hat{d}_N \end{bmatrix} \right]' \left[ C L_{0J}^{-1}(\hat{\theta}, \hat{d}) C' \right]^{-1} \left[ C \begin{bmatrix} \hat{\theta}_N \\ \hat{d}_N \end{bmatrix} \right] = 21.6441.
\]

Under \( H_2 \), \( W_N \) is asymptotically distributed as a \( \chi^2 \) variate with 2 d.f. The asymptotic p-value of \( W_N \) is

\[
P(X_2^2 > 21.6441) = 0.00001995,
\]

the hypothesis \( H_2 \) is rejected at level \( \alpha = 0.05 \).

When \( H_2 \) is not true, we consider the local alternative

\[
H_N: \Delta_0 = \begin{bmatrix} \theta_1 - \theta_3 \\ \theta_2 \end{bmatrix} = \frac{1}{\sqrt{N}} \begin{bmatrix} -0.05 \\ -0.1 \end{bmatrix}.
\]

Under \( H_N \), the statistic \( W_N \) is asymptotically distributed as a non-central \( \chi^2 \) variate with 2 d.f. and non-centrality parameter

\[
\lambda^2 = \begin{bmatrix} -0.05 \\ -0.1 \end{bmatrix}' \left[ C L_{0J}^{-1}(\hat{\theta}, \hat{d}) C' \right]^{-1} \begin{bmatrix} -0.05 \\ -0.1 \end{bmatrix} = 0.980587.
\]

Thus, the asymptotic power at \( W_N = 21.6441 \) is

\[
P(X_2^2(\lambda^2) > 21.6441) = \sum_{k=0}^{\infty} \left( \frac{\lambda^2}{2} \right)^k \frac{e^{-\lambda^2/2}}{k!} P(X_2^2 > 21.6441).
\]

We can estimate the above power by summing terms with Poisson weights

\[
\left( \frac{\lambda^2}{2} \right)^k \frac{e^{-\lambda^2/2}}{k!} > .000001. \text{ In this case, we have}
\]

\[
P(X_2^2(\lambda^2) > 21.6441) = .000286.
\]
With Pearson's approximation, we use \( b = -\frac{\lambda^4}{2+3\lambda^2} = -0.194577 \), 
\( c = \frac{2+3\lambda^2}{2+2\lambda^2} = 1.24755 \), and \( f = 2 + \frac{\lambda^4(6+8\lambda^2)}{(2+3\lambda^2)^2} = 2.54512 \).

Also, 
\[
(4.6.17) \quad P(X_2^2(\lambda^2) > 21.6441) = P(cX_2^2 + b > 21.6441) = 0.000326.
\]

### 4.6.3. Testing Intraclass Covariance Structures for \( \Sigma_1 \) and \( \Sigma_2 \) in \( H_1 \)

Under the hypothesis \( H_3 \), both \( \Sigma_1 \) and \( \Sigma_2 \) have intraclass covariance structures. The MLE's of \( \Sigma_1 \) and \( \Sigma_2 \) are

\[
(4.6.18) \quad \hat{\Sigma}_1 = \begin{bmatrix} 4330.28 & -857.281 \\ -857.281 & 4330.28 \end{bmatrix} \quad \text{and} \quad \hat{\Sigma}_2 = \begin{bmatrix} 3546.36 & -1020.77 \\ -1020.77 & 3546.38 \end{bmatrix}.
\]

The resulting maximum log-likelihood is \( L_3 = -1211.83 \), and the LR statistic for testing \( H_3 \) vs. \( H_1 \) is

\[
(4.6.19) \quad \Lambda_3 = \exp(L_3 - L_1) = 2.7242 \times 10^{-19}.
\]

According to Section 4.4.1, the distribution function of \(-2p \ln \Lambda_3\) with \( p = 0.938596 \), \( M = Ntp = 107 \), can be expanded for large \( M \) as

\[
(4.6.20) \quad P(-2p \ln \Lambda_3 \leq z) = p(X_f^2 \leq z) + O(M^{-2}), \text{ where } f = p^2 + p - 4 = 2.
\]

For \( \Lambda_3 = 2.7242 \times 10^{-19} \), \(-2p \ln \Lambda_3 = 80.2445 \), the asymptotic p-value of the LR test is 0. Thus, the hypothesis \( H_3 \) of intraclass covariance structures for \( \Sigma_1 \) and \( \Sigma_2 \), is rejected.

When the hypothesis \( H_3 \) is not true, we first consider the fixed alternative \( H_k : \Sigma_1 = 100 \begin{bmatrix} 12 & -8 \\ -8 & 72 \end{bmatrix}, \Sigma_2 = 100 \begin{bmatrix} 10 & -10 \\ -10 & 60 \end{bmatrix} \).

According to Section 4.4.2., we have
\[ p^2 = \begin{pmatrix} 0.539697 & 0 \\ 0 & 0.555556 \end{pmatrix}, \tau = 1.48004, a_1 = 2, a_2 = 0.973558, \]

and the statistic \( Z = -2pM^{-1/2} \ln \Lambda_3 + M^{1/2} \frac{1}{\tau} \ln(1-p^2) = -0.449484. \)

For \( z = \frac{Z}{\tau} = -0.303698, \) the asymptotic power is therefore

\[ (4.6.21) \quad P(\frac{Z}{\tau} > z) = 1 - \Phi(z) + M^{-1/2} \left[ a_1 \frac{1}{\tau} \phi(z) + a_2 \frac{1}{\tau^3} \phi(2)(z) \right] = 0.659049. \]

Next, we consider the local alternative \( H_M: \) \( p^2 = \frac{1}{M^2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \) \( \theta_1 = 0, \theta_2 = 1. \)

Under \( H_M, \) the asymptotic power at \( z = -2p \ln \Lambda_2 = 80.2443 \) is

\[ (4.6.22) \quad P(-2p \ln \Lambda_2 > z) = 1 - p(X_2^2 \leq z) + \frac{\theta_1 + (t-1)\theta_2}{2tM} [p(X_t^2 \leq z) - p(X_{t+2}^2 \leq z)] \]

\[ = 1 - p(X_2^2 \leq 80.2443) + \frac{1}{2M} [p(X_2^2 \leq 80.2443) - p(X_4^2 \leq 80.2443)] \]

\[ = 0. \]
CHAPTER V
SUGGESTIONS FOR FUTURE RESEARCH

In this dissertation, we derive optimal confidence procedures for generalized multivariate linear models. In an HM model with p responses, we assume that the covariance matrix of the responses is known or that a good estimate is available from the experiment with a large sample size. Considering the correlation information among the responses, a step-down procedure is used to construct the confidence set for the unknown parameters. The confidence procedure involves p quadratic forms of chi-square statistics. We optimize the confidence procedure by minimizing suitable norms of the confidence set. These norms include the weighted sum of the P quadratic form lengths and the maximum weighted quadratic form length. In each case, the weight is the inverse of the degrees of freedom for the corresponding chi-square statistic. For the weighted sum criterion, when all the degrees of freedom are less than 5, the optimal confidence procedure gives a larger confidence coefficient to the quadratic form with larger degrees of freedom. The usual induction method is not applicable to extend this result to a large degrees of freedom. The development of some alternative procedures to accommodate all possible values of the degrees of freedom is another area of future research.

When the sample size is small, the individual F statistics should be used in place of the chi-square statistics to construct the confidence set. In this case, the confidence procedure involves complicated F distributions.
It would be helpful to develop similar optimal criteria and, hence, the corresponding optimal procedures related to these F statistics.

In Chapter III, we examine the relative performance of the proposed special IM model and its corresponding complete multiresponse model. These comparisons include generalized variance, asymptotic relative efficiency, cost, and minimum risk. The result of each comparison depends on the covariance structure of the p responses. We assume a covariance matrix with an intraclass correlation structure or an autocorrelation structure with p=3. Extension of these comparisons to autocorrelation-structured covariance matrix with a general value of p as well as other structured patterns of the covariance matrix is desirable. The power comparison depends on the alternative hypothesis. We consider linear hypotheses of the form CBU = 0 versus the alternative CBU = θ₀, where U is a column vector of 1's and θ₀ is a column vector with equal elements. Future developments in the power comparison include an extension to a more general setting of linear hypotheses with different values of the matrices C, U, and θ₀.

Finally, we consider the situation when the same p responses are measured at each of t distinct time points. We examine the hypotheses of the block intraclass correlation model as well as some special cases in the standard MGLM model. In each case, the LR test is derived. When the LR statistic has a closed form expression, its moments are used to find asymptotic null and non-null distributions. When a closed form expression for the LR statistic does not exist, we construct the Wald statistic to test the hypothesis and then derive the asymptotic null and non-null distributions. After testing of block covariance structures, the next step in this research effort is the testing of the linear
hypothesis $CBU = 0$ when the covariance matrix has one of the specified block intraclass correlation structures. Suppose the LR procedure were used to test the hypothesis. Then, the LR statistic would be the ratio of the maximum likelihood under the specified block covariance structure and the linear hypothesis $CBU = 0$ to the maximum likelihood under the specified block covariance structure.


