Linear Models for Multivariate Repeated Measures Data

Shantha S. Rao
Old Dominion University

LINEAR MODELS FOR MULTIVARIATE REPEATED MEASURES DATA

by

Shantha S. Rao

M.Sc. May 1988, Bangalore University
Bangalore, India

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirement for the Degree of

DOCTOR OF PHILOSOPHY

COMPUTATIONAL AND APPLIED MATHEMATICS

OLD DOMINION UNIVERSITY
May 1996

Approved by:

_____
Dayanand N. Naik (Director)

_____
Narasinga R. Chaganty (Member)

_____
John P. Morgan (Member)

_____
Ardythe L. Morrow (Member)

ABSTRACT

LINEAR MODELS FOR MULTIVARIATE REPEATED

MEASUREMENTS DATA

Shantha S. Rao

Old Dominion University, 1996

Director: Dr. Dayanand N. Naik

In this dissertation we focus mainly on the analysis of continuous multivariate repeated measurements data based on the assumption of multivariate normality. However certain aspects of the analysis of univariate repeated measures data are also considered. Typically, we have measurements on p variables (possibly correlated) in the form of $p \times 1$ vectors y_{ijk}, observed at $k = 1, 2, ..., t_{ij}$ occasions on $j = 1, 2, ..., n_i$ individuals from $i = 1, 2, ..., g$ groups. We assume a naturally occurring covariance structure $\mathbf{V}_{ij} \otimes \boldsymbol{\Sigma}$ among the p variables on the jth individual from i^{th} group made at t_{ij} occasions. Here $\mathbf{V}_{ij}$ and $\boldsymbol{\Sigma}$ are positive definite matrices of order $t_{ij} \times t_{ij}$ and $p \times p$ respectively. We develop a general linear model approach to accommodate both balanced and unbalanced repeated measures data.

Our main results are: (1) construction of Rao's score test for a simpler model with $p = 1$ (univariate case) and $\mathbf{V}_{ij}$ having a structure as in a mixed effects model, (2) comparison of all the methods for analyzing univariate repeated measures data

with time varying covariates, (3) derivation of the maximum likelihood estimates of the covariance matrices $\mathbf{V}$ and $\mathbf{\Sigma}$ in the balanced case, (4) derivation of Satterthwaite type approximation to the distribution of multivariate quadratic forms, (5) estimation of degrees of freedom for these approximations, and (6) derivation of the maximum likelihood estimates of the covariance parameters under certain specific covariance structures for unbalanced case.

Rao's score test is derived in Chapter 2. Analysis of repeated measures in the presence of time varying covariates is a useful but difficult problem. In Chapter 3, we review the existing methods for analyzing repeated mesaured data with time varying covariates and discuss their computational aspects using SAS software. We also point out that a linear model approach yields an unified tool to analyze these data. In Chapter 4, various results about balanced multivariate repeated measures models are derived. We present the entire scheme of analysis of balanced multivariate data including the computational details. Finally, the analysis of unbalanced multivariate repeated measures is discussed in Chapter 5. In this case we assume two commonly used covariance structures namely equicorrelation and autoregressive structures for $\mathbf{V}_{ij}$ and derive the maximum likelihood estimates of the unknown parameters.

Dedicated

to

My parents

iv

ACKNOWLEDGMENTS

It gives me great pleasure to acknowledge the guidance and support provided by my thesis adviser,Dr. Dayanand N. Naik throughout the course of this research. This dissertation would not have proceeded to completion without his constant encouragement and advice.

I am also indebted to Dr. N. R. Chaganty for his guidance and help during my stay at graduate school. I wish to thank him for agreeing to serve on my committee and offering valuable suggestions.

Dr. John P. Morgan and Dr. Ardythe L. Morrow are also thanked for taking time off thier busy schedules to serve on my dissertation committee.

Many thanks are due to all my friends who helped me get through difficult times. Above all I am proud to acknowledge the immense contribution made by members of my family. This dream would not have been realized without the constant support and encouragement received from them. The love and affection showered by each one of them went a long way in guiding me through times when I had almost given up hope. As a mark of affection I dedicate this work to my family.

v

Contents

PLEASE NOTE

Page(s) not included with original material
and unavailable from author or university.
Filmed as received.

UMI

List of Tables

Chapter 1

Introduction

Repeated measurements, as indicated by the name, are successive measurements made over time on each experimental unit. If measurements are made on n individuals over t time points on a single characteristic, we have a set of univariate repeated measurements data. On the other hand, if measurements are made on a set of t repeated measurements on p variables (or characteristics) on each of the n individuals, we have a set of multivariate repeated measurements data. Thus in this case the data on each individual is a $p \times t$ matrix.

The n individuals may be randomly divided into one of several groups to which the levels of an experimental factor are assigned. Repeated measurements occur very frequently in a variety of scientific fields where statistical models are used. The distinguishing characteristics of these measurements from those in the usual or more traditional regression setting is that one or more response variables are measured on

1

the same observational unit more than once, hence making the responses dependent.

The questions that one would generally like to ask and answer, in this context are

- Is there a group effect?

- Is there a time effect?

- Is there an interaction between time and group effects?

Let us first consider a set of univariate repeated measures. In an attempt to answer the above questions treat the t measurements $Y'_{ij} = (y_{ij1}, ..., y_{ijt})$, on the jth individual belonging to the ith group as a single multivariate observation Y_{ij} rather than as a set of separate univariate observations. Assuming a general covariance structure, we can use multivariate techniques to analyze these type of data. To be more specific, the above questions can be formulated as linear hypotheses of the form $H_0 : C\mu_{ij} = 0$, for appropriate choices of C, where μ_{ij} is the mean or the expected value of Y_{ij}. Then a *profile analysis* (MANOVA) can be used to test H_0.

The main reason that prompted us to use the multivariate approach is that we had no particular structure assumed for $cov(Y_{ij}) = V$. However in many practical situations, V is found to have some simpler structures. In that case, multivariate tests are less powerful. As an alternative to the multivariate approach, by considering each subject as a random block and time points as plots within blocks a split plot design model can be used to analyze repeated measures data.

2

The model in that case is

$$y_{ijk} = \mu + \alpha_i + \beta_k + (\alpha\beta)_{ik} + \delta_{j(k)} \quad + \quad \epsilon_{ijk}, i = 1, ..., g, \tag{1.1}$$

$$j = 1, ..., n_i, \quad k = 1, ..., t,$$

$$\delta_{j(k)} \sim N(0, \sigma_\delta^2),$$

$$\epsilon_{ijk} \sim N(0, \sigma^2),$$

where y_{ijk} is the kth observation on the jth individual from the ith group and μ, α_i, β_k, $(\alpha\beta)_{ik}$ and $\delta_{j(k)}$ have the usual meaning as in split plot design. The problems of interest are to test the following null hypothesis:

(i) $\alpha_1 = ... = \alpha_g$

(ii) $\beta_1 = ... = \beta_t$ and

(iii) $(\alpha\beta)'_{ik}s$ are all equal.

We know that the usual ANOVA technique can be applied here although the observations $y_{ij1}, ..., y_{ijt}$ on an individual are correlated. As before define, $Y'_{ij} = (y_{ij1}, ..., y_{ijt})$. Then the covariance matrix of $Y_{ij} = cov(Y_{ij}) = \mathbf{V} = \sigma^2 \mathbf{I} + \sigma_\delta^2 \mathbf{J}$, where $\mathbf{I}$ is an identity matrix of order t and $\mathbf{J}$ is a $t \times t$ matrix of ones. Further, if we define $Y' = (Y'_{11}, ..., Y'_{1n_1}, Y'_{21}, ..., Y'_{2n_2}, ..., Y'_{g1}, ..., Y'_{gn_g})$ then $cov(Y) = \Omega = diag(\mathbf{I}_{n_1} \otimes \mathbf{V}, ..., \mathbf{I}_{n_g} \otimes \mathbf{V})$.

The ANOVA table for the above model is

3

Source	D.O.F.	SS	F statistic
Between individuals			
Groups	$g-1$	$Q_1 = Y'AY$	$F_1 = (n-g)Q_1/(g-1)Q_2$
Individuals	$n-g$	$Q_2 = Y'BY$	
Within individuals			
Time	$t-1$	$Q_3 = Y'CY$	$F_2 = (n-g)Q_3/Q_5$
Time*Group	$(g-1)(t-1)$	$Q_4 = Y'DY$	$F_3 = (n-g)Q_4/(g-1)Q_5$
Error	$(t-1)(n-g)$	$Q_5 = Y'EY$	
Total	$n-1$	$Y'TY$	

Here $n = \sum_{i=1}^{g} n_i$. The quadratic forms Q_1 to Q_5 are

$$Q_1 = k\sum_{i=1}^{g} n_i(\bar{y}_{i..} - \bar{y}_{...})^2 = Y'AY$$

$$Q_2 = k\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\bar{y}_{ij.} - \bar{y}_{i..})^2 = Y'BY$$

$$Q_3 = n\sum_{k=1}^{t}(\bar{y}_{..k} - \bar{y}_{...})^2 = Y'CY$$

$$Q_4 = \sum_{i=1}^{g} n_i \sum_{k=1}^{t}(\bar{y}_{i.k} - \bar{y}_{i..} - \bar{y}_{..k} + \bar{y}_{...})^2 = Y'DY$$

$$Q_5 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\sum_{k=1}^{t}(y_{ijk} - \bar{y}_{ij.} - \bar{y}_{i.k} + \bar{y}_{i..})^2 = Y'EY$$

$$\text{and } Q_1 + ... + Q_5 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\sum_{k=1}^{t}(y_{ijk} - \bar{y}_{...})^2 = Y'TY$$

with appropriate choice of matrices A through E, and T. These $n \times n$ matrices are

all symmetric and can be easily derived.

4

The tests of hypotheses are carried out using the usual ANOVA F-tests. The split-plot design that we considered accommodates a simple correlation among repeated measures. However, there are certain other covariance structures which also admit the same F tests.

Huynh and Feldt (1970) and Rouanet and Le'pine (1970) have derived a set of necessary and sufficient conditions on the covariance structures such that the distributions of the F statistics remains invariant. Consequently, there are certain dependence among repeated measures which still admit the usual tests. The class of all covariance structures which admit the usual F tests is said to have type H structure. A typical member of this class is $V = \sigma^2(I + a1' + 1a')$, where $a' = (a_1, ..., a_t)$ is such that V is positive definite and 1 is a vector of all ones. Recently, Chaganty and Vaish (1995) studied this characterization more closely and produced the following form for $H(a)$:

$$H^*(a) = I + \frac{1}{t}(a1' + 1a') - \frac{1}{t}(1 + \bar{a})11',$$

where $a' = (a_1, ..., a_t)$ is such that

$$\frac{1}{t}\sum_{i=1}^{t}(a_i - \bar{a})^2 \leq \bar{a} \tag{1.2}$$

and $\bar{a}$ is the mean of the components of the vector a. This characterization by Chaganty and Vaish (1995) is most explicit in the sense that it gives easy condition (1.2) on the elements of a such that V is positive (semi) definite. Since the form of covariance structure determines the choice of the analysis, we have a need for

5

testing for the type H structure for $\mathbf{V}$. We have the following result due to Huynh and Feldt (1970) to test for type H structure.

Result1: Let $C_{t-1 \times t}$ be such that $CC' = \mathbf{I}_{t-1}$ and $C\mathbf{1} = 0$, and $U_{t-1 \times 1} = C_{t-1 \times t} Y_{t \times 1}$. Then $cov(Y) = \mathbf{V}$ is of type H structure iff $cov(U) = CVC' = \lambda \mathbf{I}_{t-1}$.

A likelihood ratio test for testing $cov(U) = \lambda \mathbf{I}_{t-1}$ can be constructed using a test of sphericity (Mauchly (1940)) on the transformed variables. The likelihood ratio is given by

$$\lambda = \frac{|CSC'|}{[1/(t-1)tr(CSC')]^{t-1}},$$

where S is the sample variance covariance matrix. Under the null hypothesis of sphericity, that is $cov(U) = \lambda \mathbf{I}_{t-1}$,

$$-\{(n-1) - (2(t-1)^2 + (t-1) + 2)/(6(t-1))\} ln \ \lambda$$

has an approximate χ^2 distribution with $\frac{1}{2}t(t-1) - 1$ degrees of freedom. Standard statistical software like SAS provides a test for sphericity. For details see Khattree and Naik (1995).

So far we have explored the possibility of using the usual F statistics under certain dependent covariance structures. That is, we noted that the ANOVA remains invariant as long as $\mathbf{V}$ is of type H structure. Now we need to address the situation where we reject the hypothesis of sphericity. Rejecting the hypothesis of sphericity implies the departure of $\mathbf{V}$ from type H structure. Under no structure

6

on $\mathbf{V}$, Geisser and Greenhouse (1958) have shown that $A\Omega B = A\Omega E = B\Omega E = C\Omega E = D\Omega E = 0$. Hence the quadratic forms Q_1 and Q_2, Q_1 and Q_5, Q_2 and Q_5, Q_3 and Q_5, and Q_4 and Q_5, are all pairwise independent. However, these quadratic forms no longer have exact χ^2 distributions. The results that follow enable us to find the distributions of these quadratic forms and their ratios.

Result 2: Consider $X_{t\times 1} \sim N_t(0, \mathbf{V})$. Let A be any symmetric matrix of order $t \times t$ with $rank(A) = r$. Then the distribution of any quadratic form $X'AX$ is the same as the distribution of certain linear combinations of independent χ^2 variables. That is, $X'AX \stackrel{d}{=} \sum_{j=1}^{r} \lambda_i \chi_j^2(1)$, where $\lambda_1, ..., \lambda_r$ are the eigenvalues of a certain matrix to be determined in the proof. Further $\chi_j^2(1)'s$ are all independent.

Proof:

Since $\mathbf{V}$ is a positive definite matrix we can express the quadratic form $X'AX$ as

$$
\begin{aligned}
X'AX &= X'\mathbf{V}^{-1/2}\mathbf{V}^{1/2}A\mathbf{V}^{1/2}\mathbf{V}^{-1/2}X \\
&= Z'BZ(\ Z = \mathbf{V}^{-1/2}X,\ Z \sim N(0,\mathbf{I})), \\
&= Z'\Gamma\Lambda\Gamma'Z \quad (B = \Gamma\Lambda\Gamma',\ \Gamma\Gamma' = \mathbf{I} \text{ and } \Lambda = diag(\lambda_1, ..., \lambda_r, 0, .., 0)) \\
&= U'\Lambda U,\ (U = \Gamma'Z,\ U \sim N(0,\mathbf{I})) \\
&= \sum_{j=1}^{r} \lambda_i u_i^2,\ u_j^2 \sim \chi_j^2(1),\ \text{and } \chi_1^2(1), ..., \chi_r^2(1) \text{ are independent.}
\end{aligned}
$$

We observe that $\lambda_1, ..., \lambda_r$ are the nonzero eigenvalues of B. (note that $rank\ (A) = rank\ (B) = r$). Hence the result.

We see that the quadratic form $X'AX$ is the weighted sum of independent χ^2 vari-

7

ates. Satterthwaite (1941) has developed the following approximate χ^2 distribution for $X'AX$.

Suppose the distribution of $X'AX$ is approximated by $a\chi_b^2$, where a and b are determined satisfying the following conditions:

$$(i) \quad E(X'AX) \quad = \quad E(a\chi_b^2), \tag{1.3}$$

$$(ii) \quad var(X'AX) \quad = \quad var(a\chi_b^2). \tag{1.4}$$

Using (1.3) and (1.4) we can show that $a = \dfrac{\sum\limits_{j=1}^{r} \lambda_j^2}{\sum\limits_{j=1}^{r} \lambda_j}$ and $b = \dfrac{(\sum\limits_{j=1}^{r} \lambda_j)^2}{\sum\limits_{j=1}^{r} \lambda_j^2}$.

From Result 2 we have the following:

$$Q_1 \quad \sim \quad \chi^2(g-1)$$

$$Q_2 \quad \sim \quad \chi^2(n-g)$$

$$Q_3 \quad \sim \quad \sum_{j=1}^{t-1} \lambda_j \chi_j^2(1)$$

$$Q_4 \quad \sim \quad \sum_{j=1}^{t-1} \lambda_j \chi_j^2(g-1)$$

$$Q_5 \quad \sim \quad \sum_{j=1}^{t-1} \lambda_j \chi_j^2(n-g),$$

where $\lambda_1, ..., \lambda_{t-1}$ are the eigenvalues of $(\mathbf{I} - 1/t\mathbf{J})\mathbf{V}$.

Box (1954) applied Satterthwaite approximation to obtain the distribution of the ratio of the quadratic forms. Writing $\epsilon = \dfrac{(\sum\limits_{j=1}^{r} \lambda_j)^2}{(t-1)\sum\limits_{j=1}^{r} \lambda_j^2}$, we can verify the

8

following

$$F_1 \sim F_{(g-1),(n-g)} \quad \text{(exact)} \tag{1.5}$$

$$F_2 \sim F_{\epsilon(t-1),\epsilon(t-1)(n-g)} \quad \text{(approx.)} \tag{1.6}$$

$$F_3 \sim F_{\epsilon(t-1)(g-1),\epsilon(t-1)(n-g)} \quad \text{(approx.)} \tag{1.7}$$

We use the above F statistics given by (1.5), (1.6) and (1.7) and test the hypotheses of interest. Notice that in applications we need to estimate ϵ. Greenhouse and Geisser (1959) suggested the estimate of ϵ given by

$$\hat{\epsilon}_{GG} = \frac{[tr(C'SC)]^2}{(t-1)tr(C'SC)^2},$$

where S is the sample variance covariance matrix. Huynh and Feldt (1976) have suggested an alternative estimate of ϵ given by

$$\hat{\epsilon}_{HF} = \frac{n(t-1)\hat{\epsilon}_{GG} - 2}{(t-1)[n-1-(t-1)]\hat{\epsilon}_{GG}}.$$

We summarize the scheme for analysis of univariate repeated measures using the ANOVA approach as follows:

- If ϵ is very small, which is an indication of lack of sphericity, use of the multivariate approach may be suggested.

- If ϵ is very close to 1, usual ANOVA for a split plot design may be carried out.

9

- Otherwise use of ANOVA with an approximate F (use of ϵ correction) may be appropriate.

Next, we give an outline of the subsequent chapters.

In Chapter 2 we study a generalization of the basic ANOVA model to analyze univariate repeated measures data. Laird and Ware (1982), utilizing the ideas introduced by Harville (1977), introduced mixed effects models to analyze these data. This modeling approach gives us a choice of various covariance structures for both individual random effects component and within individual errors. We consider a mixed effects model assuming within individual errors that follow an autoregressive (AR(1)) structure. We propose a score test for the parameter that accounts for the between individual random effects component. This test is very useful because it tests for the possibility of a simpler model.

In Chapter 3 we study methods of analyzing repeated measures data in the presence of covariates. We review methods of analyzing univariate repeated measures data in presence of two types of covariates: (a) covariates fixed over time and (b) time varying covariates. We provide computer programs, using SAS software to perform each of these analyses.

In Chapter 4 we consider analysis of balanced multivariate repeated measures data. Analysis of these data using a MANOVA model is considered. The well known Satterthwaite type approximation to the distribution of a quadratic form in normal variables is extended to the distribution of a multivariate quadratic form in

10

multivariate normal variates. The multivariate tests using this approximation are developed for testing the usual hypotheses.

In Chapter 5 we consider analysis of unbalanced multivariate repeated measures data. Use of multivariate linear models for analysis of these data is illustrated. Analysis of these models are discussed under some special covariance structures.

11

Chapter 2

Mixed Effects Models For

Univariate Repeated Measures

2.1 Introduction

In this chapter we study a generalization of the basic ANOVA model, to analyze repeated measures or longitudinal data. The defining characteristic of a repeated measures or longitudinal study is that the individuals are measured repeatedly over time. Longitudinal studies are designed to investigate changes in a characteristic which is measured over time. In medical experiments, the characteristic of interest might be blood pressure, cholesterol level, lung volume or serum glucose. Suppose we have t measurements taken repeatedly over time on n experimental units on a certain characteristic. Our main problem of interest is to make inference about the

12

differences between the levels of the characteristic at various time points. Stated in other words, our interest is to test a hypothesis about the time effect. To model longitudinal data collected over time for each member of a group of experimental units, one must recognize the correlation between serial observations on the same experimental unit.

In practice, longitudinal data are often unbalanced or incomplete, that is, all individuals are not observed the same number of times or with the same design matrix X. In these situations it is imperative to consider models that account for both unbalancedness and correlation among the observations on an individual.

Laird and Ware (1982), utilizing the ideas introduced by Harville (1977), introduced a generalization of the basic ANOVA model to analyze repeated measures data. Let β denote a $m \times 1$ vector of unknown population parameters and X_k be a known $t_k \times m$ design matrix. Let τ_k denote a $q \times 1$ vector of unknown individual random effects and C_k be a known $t_k \times q$ design matrix for individual effects. For each individual k, Laird and Ware (1982) considered the following model:

$$Y_k = X_k \beta + C_k \tau_k + \epsilon_k, \tag{2.1}$$

where ϵ_k is distributed as $N(0, \mathbf{V}_k)$. Here $\mathbf{V}_k$ is a $t_k \times t_k$ positive definite covariance matrix, depending on k through its dimension t_k. However, the set of unknown parameters in $\mathbf{V}_k$ do not depend on k. Next, the random vectors τ_k are distributed as $N(0, \boldsymbol{\Gamma})$, independently of each other and of ϵ_k. Here $\boldsymbol{\Gamma}$ is a $q \times q$ positive definite

13

matrix. To summarize we have model (2.1) with the following specifications.

$$E(\epsilon_k) = 0, \qquad V(\epsilon_k) = \mathbf{V}_k, \tag{2.2}$$

$$E(\tau_k) = 0, \qquad V(\tau_k) = \mathbf{\Gamma} \quad \text{and} \quad Cov(\tau_k, \epsilon_k) = 0. \tag{2.3}$$

Thus for the model (2.1) we have,

$$E(Y_k) = X_k\beta \quad \text{and}$$

$$V(Y_k) = C_k\mathbf{\Gamma}C_k' + \mathbf{V}_k.$$

For this model with $\mathbf{V}_k$ equal to an identity matrix, Laird and Ware (1982) utilized the EM algorithm to obtain the maximum likelihood (ML) and restricted maximum likelihood (REML) estimates of the variance components. However, the tests of hypothesis about the parameters of the covariance matrix were not considered by them.

Chi and Reinsel (1989) considered model (2.1) that contain both individual random effect component and within individual errors that follow an autoregressive (AR(1)) time series process. They developed a score test for the autocorrelation in the within individual errors for the 'conditional independence' random effects model. The AR(1) structure is particularly appealing because of the fact that it incorporates the natural decay in the correlation between observations taken on the same experimental unit with increasing time lags. In other words, observations on the same experimental unit which are closer together should be more highly correlated than those that are far apart in time.

14

Chi and Reinsel (1989) illustrated their methods on a data set with $C_k = \mathbf{1}_k$, where $\mathbf{1}_k$ is a $t_k \times 1$ vector of ones. This choice of C_k seems to be most useful in practical situations. With this choice of C_k, in the following section we derive a score test for the parameter that accounts for the between individual random effects component. Our model for individual k is

$$Y_k = X_k\beta + \mathbf{1}_k\tau_k + \epsilon_k, \qquad k = 1, ..., n, \qquad (2.4)$$

with

$$E(Y_k) = X_k\beta$$

$$\Sigma_k = cov(Y_k) = \sigma_\tau^2 \mathbf{1}\mathbf{1}' + \sigma^2 \mathbf{V}_k \qquad (2.5)$$

where the $(i,j)^{th}$ element of $\mathbf{V}_k$ is $\dfrac{1}{1-\rho^2}\rho^{|i-j|}$.

Notice that (2.5) can be written as

$$\Sigma_k = cov(Y_k) = \sigma^2(\xi \mathbf{J}_k + \mathbf{V}_k), \qquad (2.6)$$

where $\xi = \sigma_\tau^2/\sigma^2$ and $\mathbf{J}_k$ is a $t_k \times t_k$ matrix of one's.

2.2 Score test for $H_0 : \xi = 0$.

We propose a score test for testing $H_0 : \sigma_\tau^2 = 0 \Leftrightarrow H_0 : \xi = 0$. The score test is based on the score criterion suggested by Rao (1948), and it has the advantage over other large sample tests, such as Wald's test and the likelihood ratio test, of requiring

15

estimation only under the null hypothesis. Computation of maximum likelihood estimates is quite tedious in most cases. However the score test enables considerable savings in terms of computation. It requires estimation only under the null model, and also yields a test asymptotically equivalent to the corresponding likelihood ratio test. Thus the null distribution of the score statistic is asymptotically χ^2 with degrees of freedom equal to the number of parametric restrictions imposed by the hypothesis.

As was mentioned before, Chi and Reinsel (1989) proposed a score test for testing the random effects model (2.1) with $\text{cov}(\epsilon_k) = \sigma^2 I$ against the same model with autocorrelated (AR(1)) errors for the $\epsilon'_k s$. Thus a score test for testing H_0 : $\rho = 0$ (the autoregressive parameter$=0$) provides a simple check for the presence of possible autocorrelation in the errors. The test for $H_0 : \xi = 0$, that we develop would provide a check for the possibility of a simpler model without the random effects .

Let $\alpha = (\beta', \sigma^2, \xi, \rho)'$ be the vector of parameters of the model (2.4) and let $\hat{\alpha}_0 = (\hat{\beta}'_0, \hat{\sigma}^2_0, \hat{\xi}_o, \hat{\rho}_0)'$ be the parameters of α replaced their corresponding ML estimates under the null model. The log-likelihood function of α, given $Y_1, Y_2, ..., Y_n$, for the model (2.4) can be written as

$$l(\beta, \sigma^2, \xi, \rho) = \text{constant} - \frac{1}{2} \sum_{k=1}^{n} \log | \Sigma_k | - \frac{1}{2} \sum_{k=1}^{n} (Y_k - X_k\beta)' \Sigma_k^{-1} (Y_k - X_k\beta),$$

where $\Sigma_k = cov(Y_k) = \sigma^2(\xi J_k + V_k)$. It is true that when evaluated at $\hat{\alpha}_0$, the score

16

vector $\frac{\partial \ell}{\partial \alpha}$ has all elements equal to zero except the derivative with respect to ξ de-

noted by $\frac{\partial \ell}{\partial \xi}_{|\hat{\alpha}_0}$. The $(i,j)^{th}$ element of information matrix is $I_{ij} = E\left(-\frac{\partial^2 \ell}{\partial \alpha_i \partial \alpha'_j}\right)$,

which has a form partitioned in accordance with $\alpha = (\beta, \sigma^2, \xi, \rho)$. Then the infor-

mation matrix denoted by $\mathbf{I}$ is given by

$$
\mathbf{I} = \begin{bmatrix} I_{\beta\beta'} & 0 & 0 & 0 \\[2mm] 0 & I_{\sigma^2\sigma^2} & I_{\sigma^2\xi} & I_{\sigma^2\rho} \\[2mm] 0 & I_{\xi\sigma^2} & I_{\xi\xi} & I_{\xi\rho} \\[2mm] 0 & I_{\rho\sigma^2} & I_{\rho\xi} & I_{\rho\rho} \end{bmatrix}
$$

The score statistic denoted by λ, then takes the form

$$
\lambda = \left(\frac{\partial \ell}{\partial \alpha}\right)'_{|\hat{\alpha}_0} (\mathbf{I})^{-1}_{|\hat{\alpha}_0} \left(\frac{\partial \ell}{\partial \alpha}\right)_{|\hat{\alpha}_0} \tag{2.7}
$$

which in the present context reduces to

$$
\lambda = \left(\frac{\partial \ell}{\partial \xi}\right)_{|\hat{\alpha}_0} \left(I^*_{\xi\xi}\right)_{|\hat{\alpha}_0} \left(\frac{\partial \ell}{\partial \xi}\right)_{|\hat{\alpha}_0} \tag{2.8}
$$

where $I^*_{\xi\xi}$ is the $(m+2, m+2)^{nd}$ element of the matrix $\mathbf{I}^{-1}$. The score test statistic

λ is asymptotically distributed as chi-square with 1 degree of freedom under the

null hypothesis. The second derivatives of $l(\theta^*)$ with respect to $\theta^* = (\sigma^2, \xi, \rho)$ have

the form

$$
\begin{aligned}
\frac{\partial^2 \ell}{\partial \theta^*_i \partial \theta^*_j} &= \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \theta^*_j} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \theta^*_i}) \\
&\quad - \sum_{k=1}^{n} (Y_k - X_k\beta)' \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \theta^*_j} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \theta^*_i} (Y_k - X_k\beta). \tag{2.9}
\end{aligned}
$$

17

Therefore, it is easy to see that

$$E(-\frac{\partial^2 \ell}{\partial \theta_i^* \partial \theta_j^*}) = \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \theta_i^*} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \theta_j^*}) \qquad (2.10)$$

The elements of the information matrix can be calculated using (2.10).

Result 1: *With the usual notations, the elements of the information matrix are given by*

$$(a) \quad E(-\frac{\partial^2 \ell}{\partial \sigma^4}) \quad = \quad \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \sigma^2} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \sigma^2})$$

$$(b) \quad E(-\frac{\partial^2 \ell}{\partial \sigma^2 \partial \xi}) \quad = \quad \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \sigma^2} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \xi})$$

$$(c) \quad E(-\frac{\partial^2 \ell}{\partial \sigma^2 \partial \rho}) \quad = \quad \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \sigma^2} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \rho})$$

$$(d) \quad E(-\frac{\partial^2 \ell}{\partial \xi^2}) \quad = \quad \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \xi} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \xi})$$

$$(e) \quad E(-\frac{\partial^2 \ell}{\partial \xi \partial \rho}) \quad = \quad \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \xi} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \rho})$$

$$(f) \quad E(-\frac{\partial^2 \ell}{\partial \rho^2}) \quad = \quad \frac{1}{2} \sum_{k=1}^{n} tr(\Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \rho} \Sigma_k^{-1} \frac{\partial \Sigma_k}{\partial \rho})$$

Proof: *Easy to see from (2.10)*

Result 2: *Under the null hypothesis, the expressions (a)-(f) reduce to the fol-*

18

lowing: Define $\mathbf{V}_k^* = \dfrac{\partial}{\partial \rho}(\mathbf{V}_k)$

$(a)\quad E(-\dfrac{\partial^2 \ell}{\partial \sigma^4}) \quad = \quad \dfrac{1}{2\sigma^4}\sum_{k=1}^{n} tr(\mathbf{I}_{t_k}) \quad\quad = \quad \sum_{k=1}^{n}\dfrac{1}{2\sigma^4}t_k$

$(b)\quad E(-\dfrac{\partial^2 \ell}{\partial \sigma^2 \partial \xi}) \quad = \quad \dfrac{1}{2\sigma^2}\sum_{k=1}^{n} tr(\mathbf{V}_k^{-1}\mathbf{J}_k) \quad = \quad \sum_{k=1}^{n}\dfrac{\rho^2(t_k - 2) - 2\rho(t_k - 1) + t_k}{2\sigma^2}$

$(c)\quad E(-\dfrac{\partial^2 \ell}{\partial \sigma^2 \partial \rho}) \quad = \quad \dfrac{1}{2\sigma^2}\sum_{k=1}^{n} tr(\mathbf{V}_k^{-1}\mathbf{V}_k^*) \quad = \quad \dfrac{n\rho}{\sigma^2(1 - \rho^2)}$

$(d)\quad E(-\dfrac{\partial^2 \ell}{\partial \xi^2}) \quad = \quad \dfrac{1}{2}\sum_{k=1}^{n} tr(\mathbf{V}_k^{-1}\mathbf{J}_k\mathbf{V}_k^{-1}\mathbf{J}_k) \quad = \quad \sum_{k=1}^{n}\dfrac{[\rho^2(t_k - 2) - 2\rho(t_k - 1) + t_k]^2}{2}$

$(e)\quad E(-\dfrac{\partial^2 \ell}{\partial \xi \partial \rho}) \quad = \quad \dfrac{1}{2}\sum_{k=1}^{n} tr(\mathbf{V}_k^{-1}\mathbf{J}_k\mathbf{V}_k^{-1}\mathbf{V}_k^*) \quad = \quad \sum_{k=1}^{n}(t_k - 1) - \rho(t_k - 2)$

$(f)\quad E(-\dfrac{\partial^2 \ell}{\partial \rho^2}) \quad = \quad \dfrac{1}{2}\sum_{k=1}^{n} tr(\mathbf{V}_k^{-1}\mathbf{V}_k^*\mathbf{V}_k^{-1}\mathbf{V}_k^*) \quad = \quad \sum_{k=1}^{n}\dfrac{(t_k - 1) - \rho^2(t_k - 3)}{(1 - \rho^2)^2}$

Proof: We now give proof of only (e), as the rest similarly follow. Under the null hypothesis

$$E(-\dfrac{\partial^2 \ell}{\partial \xi \partial \rho}) \quad = \quad \dfrac{1}{2}\sum_{k=1}^{n} tr(\mathbf{V}_k^{-1}\mathbf{J}_k\mathbf{V}_k^{-1}\mathbf{V}_k^*)$$

Consider $\quad tr(\mathbf{V}_k^{-1}\mathbf{J}_k\mathbf{V}_k^{-1}\mathbf{V}_k^*) \quad = \quad tr(\mathbf{V}_k^{-1}\mathbf{1}_k\mathbf{1}_k'\mathbf{V}_k^{-1}\mathbf{V}_k^*)$

$$= \quad (\mathbf{1}_k'\mathbf{V}_k^{-1}\mathbf{V}_k^*\mathbf{V}_k^{-1}\mathbf{1}_k)$$

$$\mathbf{V}_k^{-1} \quad = \quad \mathbf{I} + \rho^2 C_1 - \rho C_2$$

$$\mathbf{1}_k'\mathbf{V}_k^{-1} \quad = \quad \mathbf{1}_k' + \rho^2 \mathbf{1}_k' C_1 - \rho \mathbf{1}_k' C_2$$

$$= \quad \mathbf{1}_k' + \rho^2(0,1,1,...,1,1,0) - \rho(1,2,2,...,2,2,1)$$

$$= \quad (1 - \rho, 1 - 2\rho + \rho^2, ..., 1 - 2\rho + \rho^2, 1 - \rho)$$

$$= \quad (1 - \rho)(1, 1 - \rho, ..., 1 - \rho, 1) \quad\quad\quad (2.11)$$

19

Recall that $\mathbf{V}_k^* = \dfrac{\partial}{\partial \rho}(\mathbf{V}_k)$ and $\mathbf{V}_k$ is of order $t_k \times t_k$. For notational convenience, we carry out the further calculations by dropping the subscript k of t_k. Let us now examine the elements of the matrix $\mathbf{V}^* = (v_{ij}^*)$, for $i, j = 1, ..., t$

. Case (i)

$$\text{For } |\, i - j \,| \;=\; 0$$

$$v_{ij}^* \;=\; \frac{2\rho}{(1 - \rho^2)^2}$$

Case (ii)

$$\text{For } |\, i - j \,| \;>\; 0$$

$$v_{ij}^* \;=\; \frac{(1 - \rho^2)\,|\, i - j \,|\, \rho^{|i-j|-1} + 2\rho^{|i-j|+1}}{(1 - \rho^2)^2}$$

Note that $\mathbf{V}^*$ is of the form

$$\mathbf{V}^* = \begin{bmatrix} v_1 & v_2 & v_3 & \cdot & \cdot & \cdot & v_{t-1} & v_t \\ v_2 & v_1 & v_2 & \cdot & \cdot & \cdot & v_{t-2} & v_{t-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ v_{t-1} & v_{t-2} & v_{t-3} & \cdot & \cdot & \cdot & v_1 & v_2 \\ v_t & v_{t-1} & v_{t-2} & \cdot & \cdot & \cdot & v_2 & v_1 \end{bmatrix} \qquad (2.12)$$

20

Using (2.11) and (2.12) we can see that

$$\mathbf{V^*V^{-1}1} = \begin{bmatrix} v_1 + & (1-\rho) & \sum\limits_{i=2}^{t-1} v_i + v_t \\[2ex] v_2 + & (1-\rho) & \sum\limits_{i=1}^{t-2} v_i + v_{t-1} \\[2ex] & & \cdot \\ & & \cdot \\ & & \cdot \\[2ex] v_{t-1} + & (1-\rho) & \sum\limits_{i=1}^{t-2} v_i + v_2 \\[2ex] v_t + & (1-\rho) & \sum\limits_{i=2}^{t-1} v_i + v_1 \end{bmatrix}.$$

Then

$$
\begin{aligned}
\mathbf{1'V^{-1}V^*V^{-1}1} &= 2v_1 + 2v_t + 4(1-\rho)\sum_{i=2}^{t-1} v_i + (1-\rho)^2(t-2)v_1 \\
&+ (1-\rho)^2 \sum_{j=2}^{t-1}\sum_{i=2}^{j-1} v_i + +(1-\rho)^2 \sum_{j=2}^{t-1}\sum_{i=2}^{t-j} v_i
\end{aligned}
\tag{2.13}
$$

It can be verified that

$$\sum_{j=3}^{t-1}\sum_{i=2}^{j-1} v_i = \sum_{j=2}^{t-2}\sum_{i=2}^{t-j} v_i.$$

Therefore, the above equation (2.13) becomes

$$
\begin{aligned}
\mathbf{1'V^{-1}V^*V^{-1}1} &= 2v_1 + 2v_t + 4(1-\rho)\sum_{i=2}^{t-1} v_i \\
&+ (1-\rho)^2(t-2)v_1 + 2\sum_{j=2}^{t-1}\sum_{i=2}^{j-1} v_i.
\end{aligned}
\tag{2.14}
$$

21

We now evaluate

$$\sum_{i=2}^{t-1} v_i \quad \text{and} \quad \sum_{j=2}^{t-1}\sum_{i=2}^{j-1} v_i$$

in terms of the elments of $\mathbf{V}^*$.

. The elements of $\mathbf{V}^*$ for $i = 1, ..., t$ can be written in the form

$$v_i = \rho^{i-2}[(i-1) - (i-3)\rho^2]$$

Now,

$$
\begin{aligned}
\sum_{i=2}^{t-1} v_i &= \sum_{i=2}^{t-1}[\rho^{i-2}(i-1) - (i-3)\rho^i] \\
&= \sum_{i=2}^{t-1} i\rho^{i-2} - \sum_{i=2}^{t-1} \rho^{i-2} - \sum_{i=2}^{t-1} i\rho^i + 3\sum_{i=2}^{t-1} \rho^i \\
&= \frac{2\rho^2 + \rho + 1}{1-\rho} - \frac{\rho^{t-2}}{1-\rho}(4\rho^2 + \rho - 1) - t(1+\rho)\rho^{t-2} \quad (2.15)
\end{aligned}
$$

$$\sum_{j=2}^{t-1}\sum_{i=2}^{j-1} v_i = \frac{(t-2)(2\rho^2 + \rho + 1)}{(1-\rho)} - \frac{(4\rho^2 + \rho - 1)(1 - \rho^{t-2})}{(1-\rho^2)} \quad (2.16)$$

$$v_1 = 2\rho \quad (2.17)$$

$$v_t = \rho^{t-2}[(t-1) - (t-3)\rho^2] \quad (2.18)$$

From equations (2.15), (2.16), (2.17) and (2.18) we have

$$\mathbf{1}'\mathbf{V}^{-1}\mathbf{V}^*\mathbf{V}^{-1}\mathbf{1} = 2[(t-1) - \rho(t-2)]. \quad (2.19)$$

Now substituting $t = t_k$ back in (2.19), we have

$$E(-\frac{\partial^2 \ell}{\partial\xi\partial\rho}) = \frac{1}{2}\sum_{k=1}^{n} tr(\mathbf{V}_k^{-1}\mathbf{J}_k\mathbf{V}_k^{-1}\mathbf{V}_k^*) = \sum_{k=1}^{n}(t_k - 1) - \rho(t_k - 2).$$

22

Similarly, we calculate the other elements of the information matrix.

Partition the information matrix

$$\mathbf{I} = \begin{bmatrix} D & E \\ E' & J \end{bmatrix}$$

where

$$D = \frac{\partial^2 \ell}{\partial \beta \partial \beta'}$$

$$E = 0_{m \times 3}$$

$$J = \begin{bmatrix} \sum\limits_{k=1}^{n} \frac{1}{2\sigma^4} t_k & \sum\limits_{k=1}^{n} \frac{\rho^2(t_k-2)-2\rho(t_k-1)+t_k}{2\sigma^2} & \frac{n\rho}{\sigma^2(1-\rho^2)} \\ \sum\limits_{k=1}^{n} \frac{\rho^2(t_k-2)-2\rho(t_k-1)+t_k}{2\sigma^2} & \sum\limits_{k=1}^{n} \frac{[\rho^2(t_k-2)-2\rho(t_k-1)+t_k]^2}{2} & \sum\limits_{k=1}^{n}(t_k-1)-\rho(t_k-2) \\ \frac{n\rho}{\sigma^2(1-\rho^2)} & \sum\limits_{k=1}^{n}(t_k-1)-\rho(t_k-2) & \sum\limits_{k=1}^{n} \frac{(t_k-1)-\rho^2(t_k-3)}{(1-\rho^2)^2} \end{bmatrix}$$

Writing $\sum\limits_{k=1}^{n} t_i = A$ and $\sum\limits_{k=1}^{n} t_i^2 = B$, we have

$$J = \begin{bmatrix} \frac{1}{2\sigma^4}A & \frac{A(1-\rho)^2+2\rho n(1-\rho)}{2\sigma^2} & \frac{n\rho}{\sigma^2(1-\rho^2)} \\ \frac{A(1-\rho)^2+2\rho n(1-\rho)}{2\sigma^2} & \begin{array}{c} \frac{[(\rho^4-4\rho^3+5\rho^2-2\rho+1)B}{2\sigma^2} \\ +(-4\rho^4+12\rho^3-10\rho^2+2\rho)A \\ +(4\rho^4-4\rho^3+4\rho^2)] \end{array} & (1-\rho)A+n(2\rho-1) \\ \frac{n\rho}{\sigma^2(1-\rho^2)} & (1-\rho)A+n(2\rho-1) & \frac{[(1-\rho^2)A-n(1-3\rho^2)]}{(1-\rho^2)^2} \end{bmatrix}$$

Recall that the score statistic for testing $H_0 : \xi = 0$ is given by

$$\lambda = \left(\frac{\partial \ell}{\partial \xi}\right)_{|\hat{\alpha}_0} \left(I_{\xi\xi}^*\right)_{|\hat{\alpha}_0} \left(\frac{\partial \ell}{\partial \xi}\right)_{|\hat{\alpha}_0}, \qquad (2.20)$$

23

where

$$\left(\frac{\partial \ell}{\partial \xi}\right)_{|\hat{\alpha}_0} = -\frac{1}{2}\sum_{k=1}^{n} \mathbf{1}_k' \mathbf{V}_k^{-1} \mathbf{1}_k + \frac{1}{2\sigma^2}\sum_{k=1}^{n}(Y_k - X_k\beta)' \mathbf{V}^{-1} \mathbf{J}_k \mathbf{V}^{-1}(Y_k - X_k\beta),$$

$$\left(I^*_{\xi\xi}\right)_{|\hat{\alpha}_0} = J_{22} - (J_{21}\ J_{23})\begin{pmatrix} J_{11} & J_{13} \\ J_{31} & J_{33} \end{pmatrix}^{-1}\begin{pmatrix} J_{12} \\ J_{32} \end{pmatrix}$$

2.2.1 A special case

We now consider a special case of the model (2.4) with $t_k = t$ for all k and give the

score test for testing $H_0 : \xi = 0$. In this case it can be shown that

$$J = \begin{bmatrix} \dfrac{nt}{\sigma^2} & \dfrac{n[\rho^2(t-2)-2\rho(t-1)+t]}{2\sigma^2} & \dfrac{n\rho}{\sigma^2(1-\rho^2)} \\[2ex] \dfrac{n[\rho^2(t-2)-2\rho(t-1)+t]}{2\sigma^2} & \dfrac{n[\rho^2(t-2)-2\rho(t-1)+t]^2}{2} & n[(t-1)-\rho(t-2)] \\[2ex] \dfrac{n\rho}{\sigma^2(1-\rho^2)} & n[(t-1)-\rho(t-2)] & \dfrac{(t-1)-\rho^2(t-3)}{(1-\rho^2)^2} \end{bmatrix}$$

Now the components of the score test statistic are given by

$$\left(\frac{\partial \ell}{\partial \xi}\right)_{|\hat{\alpha}_0} = -\frac{n}{2}\mathbf{1}'\mathbf{V}^{-1}\mathbf{1} + \frac{1}{2\sigma^2}\sum_{k=1}^{n} U_k' \mathbf{V}^{-1} \mathbf{J} \mathbf{V}^{-1} U_k,$$

$$U_k = (Y_k - X_k\beta)$$

$$\text{and }\left(I^*_{\xi\xi}\right)_{|\hat{\alpha}_0} = \frac{2}{n(1-\rho)^4(t-2)(t-1)}.$$

Therefore the test statistic λ is given by

$$\lambda = \frac{n^2\left[(1-\hat{\rho}_0)(t(1-\hat{\rho}_0)+2\hat{\rho}_0)) + \frac{1}{2\hat{\sigma}_0^2}n(1-\hat{\rho}_0)^2\sum\limits_{k=1}^{n}(\sum\limits_{i=1}^{t} u_{ik} - \hat{\rho}_0\sum\limits_{i=2}^{t} u_{ik})^2\right]^2}{4(1-\hat{\rho}_0)^4(t-2)(t-1)} \quad (2.21)$$

Here u_{ik} is the ith component of the vector $\mathbf{u}_k$, and

24

$$
\mathbf{u}_k = \begin{pmatrix} y_{1k} - X_k \hat{\beta}_0 \\ y_{2k} - X_k \hat{\beta}_0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ y_{tk} - X_k \hat{\beta}_0 \end{pmatrix},
$$

vector of measurements on individual k. In order to compute the score test statistic

given by (2.21), we need $\hat{\beta}_0, \hat{\sigma}_0^2$ and $\hat{\rho}_0$, estimates of β, σ^2 and ρ respectively under

the null hypothesis $H_0 : \xi = 0$. We use the estimates obtained in the Appendix to

evaluate the test statistic. Once the statistic is evaluated it can be used to test H_0

by comparing its value with the χ^2 cutoff point.

Before proceeding with the analysis with the other components of the model,

we provide a brief discussion of the Rao's score test with the other two aymptotic

tests, the likelihood ratio test and the Wald's test.

A discussion of the three large sample tests

We give a brief discussion of different tools for hypothesis testing in large samples.

Testing of hypotheses in large samples has a long history. Wilks (1938) introduced

the likelihood ratio test, Wald (1943) introduced what is now called Wald's test,

and Rao (1948) proposed the score test (now known as Rao's score test) for testing

in large samples.

<div align="center">25</div>

These three tests are known to have a common aymptotic distribution, which is chi-square with degrees of freedom equal to the dimension of the vector of parameters under the null hypothesis. In a text book on econometrics Amemiya (1985) has discussed these three tests. Hall and Mathiason (1990) showed that these tests are not just asymptotically equivalent in law (i.e., have the same asymptotic distribution) but are asymptotically pointwise equivalent. That is, the test statistics are indeed approximately equal. This equivalence is proved by these authors considering Pitman-type local alternatives.

By simulation and theoretical investigations, it has been shown by several authors that there are no strict dominance relationships among the Wald's, likelihood ratio and score test statistics. For example, see Amemiya (1985). However, Chandra and Mukherjee (1991) observed that Wald's and Rao's tests are not necessarily unbiased while the likelihood ratio test is. In this sense the three tests are not therefore strictly comparable. They suggested a Bartlett type modification for Wald's and score tests to correct for bias. They found that the modified score test has better local properties than the likelihood ratio or modified Wald test.

The conclusions from the above discussion and the fact that the score test is easier to compute than either the likelihood ratio or the Wald's test, prompted us to restrict out attention to the score test for testing $H_0 : \xi = 0$.

26

2.3 Analysis under the reduced model

Suppose we accept the null hypothesis $H_0 : \xi = 0$. Then the model (2.4) is reduced to

$$Y_k = X_k\beta + \epsilon_k, \qquad k = 1, ..., n \qquad (2.22)$$

That is, we now have a linear model with fixed effects and correlated errors. In the following sections, we discuss various methods of testing hypothesis of the form $H_0 : K\beta = 0$. In order to test the above linear hypothesis about the mean parameter vector β we can use one of the following alternatives.

(1) Generalized least squares method

(2) Likelihood ratio method

(3) Score test

(4) Approximate F test

While the methods (1) and (2) are commonly used, methods (3) and (4) are developed in this thesis.

2.3.1 Generalized least squares method

The generalized least squares estimator $\hat{\beta}$ is found by minimizing the quadratic form $\sum_{k=1}^{n} Q_k(\beta, \mathbf{V})$, where $Q_k(\beta, \mathbf{V}) = (Y_k - X_k\beta)'\mathbf{V}_k^{-1}(Y_k - X_k\beta)$. In the hypothetical case of known $\mathbf{V}_k$, the solution of β is given by

$\hat{\beta} = (\sum_{k=1}^{n} X_k'\mathbf{V}_k^{-1}X_k)^{-1}(\sum_{k=1}^{n} X_k'\mathbf{V}_k^{-1}Y_k)$. When $\mathbf{V}_k$ is unknown, the estimate of

27

β has the same form as above with $\mathbf{V}_k$ replaced by $\hat{\mathbf{V}}_k$, where $\hat{\mathbf{V}}_k$ is any consistent estimator of $\mathbf{V}_k$. The estimated covariance matrix of $\hat{\beta}$ is $var(\hat{\beta}) = \hat{\mathbf{V}}_\beta = (\sum_{k=1}^n X_k' \hat{\mathbf{V}}_k^{-1} X_k)^{-1}$. This is used for testing the hypothesis of the form $K\beta = 0$. The statistic $T_c = (K\hat{\beta})'(K\hat{\mathbf{V}}_\beta K')^{-1}(K\hat{\beta})$ approximately has a chi-square distribution with r degrees of freedom. This because as $n \to \infty$, $\hat{\mathbf{V}} \to \mathbf{V}$ in probability. Thus under the null hypothesis, $K\hat{\beta} \to N_r(0, K\mathbf{V}_\beta K')$ and so $T_c \to \chi_r^2$ in distribution.

2.3.2   Likelihood ratio method

Let $\Theta = (\beta, \sigma^2, \rho)$ be the vector of unknown parameters of the model. Let $L(\beta, \sigma^2, \rho)$ be the likelihood function obtained by evaluating the joint density of $Y_1, Y_2, ..., Y_n$ at their specified values $y_1, y_2, ..., y_n$. The likelihood ratio test for testing the linear hypothesis $H_0 : K\beta = 0$, where K is a $r \times m$ matrix of full rank $r \leq m$, rejects H_0 in favor of H_a if

$$\Lambda = \frac{\max_{\theta \in \Theta_0} L(\beta, \sigma^2, \rho)}{\max_{\theta \in \Theta} L(\beta, \sigma^2, \rho)} < C$$

with $\Theta = (\beta, \sigma^2, \rho)$ and $\Theta_0 = \{(\beta, \sigma^2, \rho) : K\beta = 0\}$.

Let $\hat{\theta} = (\hat{\beta}, \hat{\sigma}^2, \hat{\rho})$ be the MLE of $\theta \in \Theta$ and $\hat{\theta}_0 = (\hat{\beta}_0, \hat{\sigma}_0^2, \hat{\rho}_0)$ be the MLE of θ under H_0, that is when $K\beta = 0$. The method of estimation of θ under the entire parameter space Θ and under the null hypothesis H_0 is discussed in the Appendix to this chapter.

28

Using these estimates, the likelihood ratio statistic reduces to

$$\Lambda = \left(\frac{\mid \hat{\Sigma} \mid_\Theta}{\mid \hat{\Sigma} \mid_{\Theta_0}}\right)^{\frac{n}{2}}$$

$$= \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{\frac{nt}{2}} \left(\frac{1 - \hat{\rho}_0^2}{1 - \hat{\rho}^2}\right)^{\frac{n}{2}}$$

In applications of the likelihood ratio method, we must obtain the sampling distribution of the likelihood ratio statistic Λ. However, it is well known that when the sample size is large, the sampling distribution of $-2ln\ \Lambda$ is well approximated by a χ^2 distribution with degrees of freedom $(\nu - \nu_0)$. Here the degrees of freedom $(\nu - \nu_0)$=(dimension of Θ)-(dimension of Θ_0). Thus in the present situation $-2ln\ \Lambda \sim \chi^2$ with r degrees of freedom approximately.

2.3.3 Score test

Here we develop a score test for testing $H_0 : K\beta = 0$ vs. $H_a : K\beta \neq 0$. Let $\alpha = (\beta', \sigma^2, \rho)$ and let $\hat{\alpha}_0 = (\hat{\beta}', \hat{\sigma}^2, \hat{\rho})$ be the MLE's of α under the restriction of the null hypothesis. Let l denote the loglikelihood function of $Y_1, Y_2, ..., Y_n$, which can be written as

$$l(\beta, \sigma^2, \rho) = \text{constant} - \frac{1}{2}\sum_{k=1}^{n} log \mid \sigma^2 \mathbf{V}_k \mid$$

$$- \frac{1}{2\sigma^2}\sum_{k=1}^{n}(Y_k - X_k\beta)'\mathbf{V}_k^{-1}(Y_k - X_k\beta) \tag{2.23}$$

29

The components of the score vector are

$$S = \begin{pmatrix} \dfrac{\partial \ell}{\partial \beta} \\[2mm] \dfrac{\partial \ell}{\partial \sigma^2} \\[2mm] \dfrac{\partial \ell}{\partial \rho} \end{pmatrix}$$

$$= \begin{pmatrix} \dfrac{1}{\sigma^2} \sum_{k=1}^{n} (X_k' \mathbf{V}_k^{-1} Y_k - X_k' \mathbf{V}_k^{-1} X_k \beta) \\[3mm] -\dfrac{nt}{2\sigma^2} + \dfrac{1}{2\sigma^4} \sum_{k=1}^{n} (Y_k - X_k\beta)' \mathbf{V}_k^{-1} (Y_k - X_k\beta) \\[3mm] -\dfrac{n\rho}{1-\rho^2} - \dfrac{1}{2\sigma^2} \sum_{k=1}^{n} (Y_k - X_k\beta)'(2\rho C_1 - C_2)(Y_k - X_k\beta) \end{pmatrix}$$

We need to evaluate the score vector S at $\hat{\alpha}_0$, where $\hat{\alpha}_0 = (\hat{\beta}_0, \hat{\sigma}_0^2, \hat{\rho}_0)$, $\hat{\beta}_0, \hat{\sigma}_0^2, \hat{\rho}_0$

being the estimates of β, σ^2, and ρ under the null model.

It is true that $S_{\hat{\alpha}_0} = \begin{pmatrix} \dfrac{\partial \ell}{\partial \beta}\Big|_{\hat{\alpha}_0} \\[2mm] 0 \\[2mm] 0 \end{pmatrix}$

Let the information matrix of $\alpha = (\beta, \sigma^2, \rho)$ be partitioned as

$$\mathbf{I} = \begin{pmatrix} I_{\beta\beta'} & 0 & 0 \\[2mm] 0 & I_{\sigma^2\sigma^2} & I_{\sigma^2\rho} \\[2mm] 0 & I_{\rho\sigma^2} & I_{\rho\rho} \end{pmatrix}$$

where $I_{\alpha_i\alpha_j} = E\left(-\dfrac{\partial^2 \ell}{\partial \alpha_i \partial \alpha_j'}\right)$ for $\alpha_i, \alpha_j = \beta, \sigma^2$ and ρ. The score statistic for H_0 :

$K\beta = 0$ can be expressed in terms of the quadratic form as

$$\Lambda = \left(\dfrac{\partial \ell}{\partial \beta}\right)' (I_{\beta\beta})^{-1}_{|\hat{\alpha}_0} \left(\dfrac{\partial \ell}{\partial \beta}\right)$$

30

where $I_{\beta\beta} = -\frac{1}{\hat{\sigma}^2}\sum_{k=1}^n X_t'\ V_k^{-1}X_t\ |_{\hat{\alpha}_0}$. Estimation of α under $H_0 : K\beta = 0$ is discussed in the Appendix. For large n, the distribution of Λ is approximated by a χ^2-distribution with r degrees of freedom.

2.3.4 Approximate F test

We now discuss a new approach to testing hypotheses of the form $H_0 : K\beta = 0$ under the reduced model. Consider the model

$$Y_k = X_k\beta + \epsilon_k \quad k = 1,...,n$$

We now require $t_k = t$ for all k to develop this method. If we assume that $cov(\epsilon_k) = \sigma^2\mathbf{I}$, to test the above hypothesis we have the usual F statistic given by

$$F = \frac{(K\hat{\beta})'(K(X'X)^{-1}K')^{-1}(K\hat{\beta})}{r\hat{\sigma}^2}$$

with r and $n - r$ degrees of freedom, r being the rank of K. However we have dependent errors with $cov(\epsilon_k) = \mathbf{V} = \dfrac{\sigma^2}{1 - \rho^2}\rho^{|i-j|}$ instead of $cov(\epsilon_k) = \sigma^2\mathbf{I}$. In this situation the natural question that arises is about the distribution of the F statistic. Several authors including Box (1954) and Geisser and Greenhouse (1958) have tried to provide an answer to this question. We know that the sum of squares due to the hypothesis can be expressed as a quadratic form as $Y'HY$. Let the sum of squares due to error be expressed as $Y'EY$. In general to use the ratio $\dfrac{Y'HY}{Y'EY}$ as a test statistic, we require the two quadratic forms to be independent.

31

Stated in other words, we require $HVE = 0$. This condition may not hold true in general. However in some special situations, like the one-way ANOVA model, split plot designs, the above condition holds true for all the standard hypothesis. The distributions of the F statitics are considered in detail for the split plot design in Chapter 1. We now derive the expressions for the correction factor ϵ for the degrees of freedom under this autoregresive structure for $\mathbf{V}$.

We may recall from Chapter 1 that

$$\epsilon = (t-1)^{-1} \frac{[tr(\mathbf{V} - \frac{1}{t}J\mathbf{V})]^2}{tr(\mathbf{V} - \frac{1}{t}J\mathbf{V})^2}.$$

For our choice of $\mathbf{V}$

$$tr(\mathbf{V} - \frac{1}{t}J\mathbf{V}) = tr(\mathbf{V}) - \frac{1}{t}tr(J\mathbf{V})$$

$$tr(\mathbf{V} - \frac{1}{t}J\mathbf{V})^2 = tr(\mathbf{V}) - \frac{2}{t}tr(\mathbf{V}J\mathbf{V}) + \frac{1}{t^2}tr(J\mathbf{V}J\mathbf{V})$$

where

$$tr(\mathbf{V}) = \frac{\sigma^2 t}{1 - \rho^2}$$

$$tr(J\mathbf{V}) = \frac{\sigma^2[t(1-\rho^2) - 2\rho(1-\rho^t)]}{(1-\rho^2)(1-\rho)^2}$$

$$tr(\mathbf{V}\mathbf{V}) = \frac{\sigma^4}{(1-\rho^2)^4}[t(1-\rho^4) - 2\rho^2(1-\rho^{2t})]$$

$$tr(\mathbf{V}J\mathbf{V}) = \frac{\sigma^4\left[t(1+\rho)^2 + 2t\rho^{t+1} + 2\rho^2\frac{1-\rho^{2t}}{1-\rho^2} - 4\rho(1+\rho)\frac{1-\rho^t}{1-\rho}\right]}{(1-\rho^2)^2(1-\rho)^2}$$

$$tr(J\mathbf{V}J\mathbf{V}) = \left[\frac{\sigma^2[t(1-\rho^2) - 2\rho(1-\rho^t)]}{(1-\rho^2)(1-\rho)^2}\right]^2$$

In order to evaluate ϵ we need the estimates of σ^2 and ρ. The estimation of these is described in the Appendix.

32

2.4 Appendix

2.4.1 Model and estimators

Consider the model

$$Y_k = X_k\beta + \epsilon_k \qquad k = 1, ..., n,$$

where ϵ_k is a $t \times 1$ vector, X_k is the $t \times m$ design matrix for the mean vector and Y_k is a $t \times 1$ vector of dependent variables on individual k. It is assumed that $E(\epsilon_k) = 0$ and $cov(\epsilon_k) = \sigma^2 \mathbf{V}$, where

$$\mathbf{V} = \frac{1}{1-\rho^2}
\begin{bmatrix}
1 & \rho & \rho^2 & \cdots & \rho^{t-1} \\
\rho & 1 & \rho & \cdots & \rho^{t-2} \\
\cdot & \cdot & \cdot & \cdots & \\
\cdot & \cdot & \cdot & \cdots & \\
\rho^{t-1} & \rho^{t-2} & \rho^{t-3} & \cdots & 1
\end{bmatrix}$$

It can be verified that $\mid \mathbf{V} \mid = \dfrac{1}{1-\rho^2}$ and $\mathbf{V}^{-1} = I + \rho^2 C_1 - \rho^2 C_2$ where

$$C_1 =
\begin{bmatrix}
0 & 0 & \cdots & .0 & 0 \\
0 & 1 & \cdots & .0 & 0 \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
0 & 0 & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & 0
\end{bmatrix}$$

33

$$C_2 = \begin{bmatrix} 0 & 1 & . & . & .0 & 0 \\ 1 & 0 & . & . & .0 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & 0 & 1 \\ 0 & 0 & . & . & 1 & 0 \end{bmatrix}$$

The likelihood function for the model is given by

$$l = (\frac{1}{2\pi})^{\frac{n}{2}} \prod_{k=1}^{n} \mid \sigma^2 \mathbf{V} \mid^{-\frac{1}{2}} exp\{-\frac{1}{2\sigma^2} \sum_{k=1}^{n} (Y_k - X_k\beta)'\mathbf{V}^{-1}(Y_k - X_k\beta)\} \qquad (2.24)$$

and the loglikelihood function is given by

$$l\,(\beta, \sigma^2, \rho) = constant - \frac{nt}{2}ln\sigma^2 + \frac{n}{2}ln(1 - \rho^2)$$

$$- \frac{1}{2\sigma^2} \sum_{k=1}^{n} (Y_k - X_k\beta)'(I + \rho^2 C_1 - \rho C_2)(Y_k - X_k\beta). \qquad (2.25)$$

The equation (2.25) can be written as

$$l\,(\beta, \sigma^2, \rho) = constant - \frac{nt}{2}ln\sigma^2 + \frac{n}{2}ln(1 - \rho^2) - \frac{A_1}{2\sigma^2} - \frac{\rho^2 A_2}{2\sigma^2} + \frac{\rho A_3}{2\sigma^2} \qquad (2.26)$$

where

$$A_1 = \sum_{k=1}^{n}(Y_k - X_k\beta)'(Y_k - X_k\beta),$$

$$A_2 = \sum_{k=1}^{n}(Y_k - X_k\beta)'C_1(Y_k - X_k\beta),$$

$$A_3 = \sum_{k=1}^{n}(Y_k - X_k\beta)'C_2(Y_k - X_k\beta).$$

34

The maximum likelihood method yields the following non-linear maximum likelihood (ML) equations which can be solved iteratively for the unknown parameters β, σ^2 and ρ respectively.

$$\hat{\beta} = [\sum_{k=1}^{n} X_k' X_k + \hat{\rho}^2 \sum_{k=1}^{n} X_k' C_1 X_k - \hat{\rho} \sum_{k=1}^{n} X_k' C_2 X_k]^{-1} \tag{2.27}$$

$$[\sum_{k=1}^{n} X_k' Y_k + \hat{\rho}^2 \sum_{k=1}^{n} X_k' C_1 Y_k + \hat{\rho} \sum_{k=1}^{n} X_k' C_2 Y_k]$$

$$\hat{\sigma}^2 = \frac{1}{nt}[A_1 + \hat{\rho}^2 A_2 - \hat{\rho} A_3] \tag{2.28}$$

$$2(t-1)A_2\rho^3 - (t-2)A_3\rho^2 - 2(A_1 + tA_2)\rho + tA_3 = 0. \tag{2.29}$$

We note that only the equations (2.27) and (2.29) have to be solved iteratively.

2.4.2 Estimation under the null hypothesis $K\beta = 0$

Consider the model

$$Y_k = X_k \beta + \epsilon_k \qquad k = 1, ..., n \tag{2.30}$$

We need to estimate the parameters of the model under $H_0 : K\beta = 0$, where K is a matrix of order $r \times m$ with $rank\,(K) = r$. To do that choose a matrix L of order $m - r \times m$ such that $M = \begin{pmatrix} K \\ L \end{pmatrix}$ is a nonsingular matrix , and $MM^{-1} = I$. The model (2.30) can be written as

$$Y_k = X_k M^{-1} M\beta + \epsilon$$

35

$$= X_k M^{-1} \begin{pmatrix} K \\ \\ L \end{pmatrix} \beta + \epsilon$$

$$= (X_k^{(1)} \mid X_k^{(2)}) \begin{pmatrix} K\beta \\ \\ L\beta \end{pmatrix} + \epsilon \quad \text{(writing} \quad X_k M^{-1} = (X_k^{(1)} \mid X_k^{(2)}))$$

$$= X_k^{(1)} K\beta + X_k^{(2)} L\beta + \epsilon$$

Under H_0 the model reduces to

$$Y_k = Z_k \beta^* + \epsilon,$$

where $Z_k = X_k^{(2)}$ and $\beta^* = L\beta$. Now we need the estimates of β^*, σ^2 and ρ. The likelihood function under the null hypothesis yields the following non-linear maximum likelihood (ML) equations which can be solved for the unknown parameters β^*, ρ and σ^2. The equations are:

$$\hat{\beta}^* = [\sum_{k=1}^{n} Z_k' Z_k + \hat{\rho}^2 \sum_{k=1}^{n} Z_k' C_1 Z_k - \hat{\rho} \sum_{k=1}^{n} Z_k' C_2 Z_k]^{-1}$$
$$[\sum_{k=1}^{n} Z_k' Y_k + \hat{\rho}^2 \sum_{k=1}^{n} Z_k' C_1 Y_k + \hat{\rho} \sum_{k=1}^{n} Z_k' C_2 Y_k]$$
$$\hat{\sigma}^2 = \frac{1}{nt}[B_1 + \hat{\rho}^2 B_2 - \hat{\rho} B_3]$$

$$2(t-1)B_2 \rho^3 - (t-2)B_3 \rho^2 - 2(B_1 + tB_2)\rho + tB_3 = 0$$

where

$$B_1 = \sum_{k=1}^{n} (Y_k - Z_k \beta)'(Y_k - Z_k \beta),$$

36

$$B_2 = \sum_{k=1}^{n}(Y_k - Z_k\beta)'C_1(Y_k - Z_k\beta),$$

$$B_3 = \sum_{k=1}^{n}(Y_k - Z_k\beta)'C_2(Y_k - Z_k\beta).$$

37

Chapter 3

Analysis of Repeated Measures

With Covariates

In this chapter we study methods of analyzing repeated measures data in the presence of covariates. In the medical field we often come across repeated measures data along with covariates that influence the response variable. It is important to conduct the analysis of repeated measures data to account for covariate effects, when data contain covariates that might affect the analysis. In clinical trials, baseline measurement can be thought of as a useful covariate for analyzing response patterns at successive visits. In the following we describe two situations, cited by Patel (1986), where one or more covariates occur naturally. In therapies for the treatment of chronic stable angina, treadmill walking time (a covariate) is recorded just before the administration of a dose, and then at some post-dose time. The

38

effectiveness of the visit doses is evaluated relative to the corresponding baseline walking times. In another situation the effectiveness of a drug in the treatment of atherosclerosis is probably influenced by factors such as diet, exercise and smoking (covariates). If the influencing factors are measured we need to construct a model which makes use of this information.

The situations described above differ in the nature of covariate information available. In the first example, the value of the covariate is specific to the subject and does not vary with time. However, in the second situation the value of each covariate is being measured at each time point along with the response variable. Thus depending on the type of covariate information available, we have

- Covariates fixed over time

- Covariates changing over time, or time varying covariates

In a repeated measures data analysis the main problems of interest are

(i) To test for the interaction between time and the group effects (*parallel profiles*)

(ii) To compare the profiles of the means (over time) of different groups (*coincidental profiles*)

(iii) To test the difference between the means of the response variables at different time points (*horizontal profiles*)

In this chapter, we review the methods of analyzing repeated measures data in the presence of covariates. In the section 3.1 we discuss the methods of solving these

39

problems when the covariates are fixed over time. In section 3.2 we review various methods of analyzing the repeated measures data with time varying covariates. These methods are illustrated using numerical examples and the SAS software. The SAS programs are also provided for convenience.

3.1 Covariates fixed over time

Depending on the nature of the data and other considerations about the covariance structures of the repeated measures, one of the following two approaches are usually adopted in practice.

3.1.1 The multivariate approach

Let $Y_{ij}' = (y_{ij1}, ..., y_{ijt})$ be a vector of t measurements on the jth individual from the ith group. We also have a set of m covariates, $x_{1ij}, ..., x_{mij}$, on each individual. Let us assume $m = 1$ for simplicity of presentation and denote the value of the covariate by x_{ij}. We have the following model for vector Y_{ij},

$$Y_{ij} = \mu + \alpha_i + \lambda_i x_{ij} + \epsilon_{ij}, \quad j = 1, ..., n_i, \quad i = 1, ..., g, \tag{3.1}$$

where μ, α_i, and λ_i are all $t \times 1$ vectors, and $\epsilon_{ij} \sim N_t(0, V)$, V a positive definite matrix.

We can rewrite (3.1) as

$$Y_{ij}' = \mu' + \alpha_i' + x_{ij}\lambda_i' + \epsilon_{ij}'$$

40

$$= X'_{ij}B + \epsilon'_{ij},$$

where X'_{ij} is the $1 \times (2g + 1)$ vector $(1, \ 0, ...0, \ 1, \ 0, ..., 0, \ ..., x_{ij}, ..., 0)$ and B is the $(2g + 1) \times t$ matrix such that $B' = (\mu, \ \alpha_1, ..., \alpha_g, \ \lambda_1, ..., \lambda_g)$.

Next define $Y_{n \times t} = (Y_{11}, ..., Y_{1n_1}, Y_{21}, ..., Y_{2n_2}, ..., Y_{gn_g})'$, $n = \sum_{i=1}^{g} n_i$ and $X_{n \times (2g+1)} = (X_{11}, ..., X_{1n_1}, X_{21}, ..., X_{2n_2}, ..., X_{gn_g})'$. Then

$$Y = XB + E, \tag{3.2}$$

where E is a $n \times t$ matrix of errors. The rows of E are assumed to be independent and each distributed as $N_t(0, \mathbf{V})$. Under this model, our testing problems (i)-(iii) can be tested using the general linear hypothesis $H_0 : LBM = 0$. Any standard statistical package like SAS, can be utilized to do the above MANOVA.

3.1.2 The univariate approach

An alternative approach to the method described above is to adopt the analysis of a split plot design, but with covariates. Let y_{ijk} be an observed value of the response variable on the jth subject from the ith group at the kth occasion. Then we have the model

$$y_{ijk} = \mu + \alpha_i + \beta_k + (\alpha\beta)_{ik} + \lambda_k x_{ij} + \gamma_{j(i)} + \epsilon_{ijk}, \tag{3.3}$$

$$k = 1, ..., t, \ j = 1, ..., n_i, \ i = 1, ..., g$$

with $\epsilon_{ijk} \sim N(0, \sigma^2)$, $\gamma_{j(i)} \sim N(0, \sigma_\gamma^2)$. These assumptions on the random effects amounts to having an equicorrelation structure for $cov(Y_{ij}) = \mathbf{V} = \sigma^2 \mathbf{V}(\rho)$, that is,

41

$\mathbf{V}(\rho) = (1-\rho)\mathbf{I} + \rho\mathbf{11}'$, for the correlation matrix of an individual. Hence we can apply the usual linear model theory. For that, write $\mathbf{y}' = (Y_{11}', ..., Y_{1n_1}', ..., Y_{g1}', ..., Y_{gn_g}')$, with $\beta' = (\mu, \alpha_1, ..., \alpha_g, \beta_1, ..., \beta_t, (\alpha\beta)_{11}, ..., (\alpha\beta)_{gt}, \lambda_1, ..., \lambda_g)$. Then we have

$$\mathbf{y} = \mathbf{X}\beta + \epsilon, \tag{3.4}$$

where $\mathbf{X}$ is the appropriately chosen design matrix and $\epsilon \sim N_n(0, \Omega)$, $\Omega = \mathbf{I}_n \otimes \mathbf{V}_{t \times t} = \mathbf{I}_n \otimes \sigma^2 \mathbf{V}(\rho)$, $n = \sum_{i=1}^{g} n_i$. Under this model, testing of various hypotheses can be easily carried out by testing $H_0 : L\beta = 0$ for specific choices of L. Again softwares like SAS, can be utilized to do this testing.

The two approaches (the multivariate and the univariate) are different on the following two accounts:

(a) On one hand we assume a completely arbitrary correlation structure for the multivariate approach, and on the other hand we use a very simple equicorrelation structure for the univariate approach.

(b) We use the theory of multivariate statistical analysis for the first approach, and the ANOVA F-tests for the second.

Both of these approaches fail when the correlation structure is not of one of the two extremes considered above. To overcome the shortcomings of these methods the following alternative linear model approach can be adopted.

42

Suppose the linear model

$$y = \mathbf{X}\beta + \epsilon, \tag{3.5}$$

is used to model the repeated measures data with Y as a $nt \times 1$ vector of observations observed at t occasions on $n = \sum_{i=1}^{g} n_i$ patients, X as a $nt \times m$ design matrix and β as a $m \times 1$ vector of parameters. Suppose $cov(\epsilon) = \Omega = \mathbf{I}_n \otimes \mathbf{V}_{t \times t}$, where $\mathbf{V} = \mathbf{V}(\theta_1, ..., \theta_k)$, for some unknown parameters $\theta_1, ..., \theta_k$.

Assuming multivariate normality for ϵ, the maximum likelihood estimates of β, $\theta_1, ..., \theta_k$ can be derived. Then the likelihood ratio test, for testing any linear hypothesis of interest $H_0 : L\beta = 0$, can be utilized appealing to the asymptotic distribution of the test.

Another route taken is through generalized least squares estimates (GLE). For a known $\mathbf{V}$, since $\mathbf{V}$ is positive definite, GLE of β, $\hat{\beta} = (X'\mathbf{V}^{-1}X)^{-1}X'\mathbf{V}^{-1}Y$. Using this $\hat{\beta}$ the ANOVA test for testing $H_0 : L\beta = 0$ can be derived. But $\mathbf{V}$ is usually unknown. Using some consistent estimators of $\theta_1, ..., \theta_k$, $\mathbf{V}$ is estimated and the estimate $\hat{\mathbf{V}} = \mathbf{V}(\theta_1, ..., \theta_k)$ is used in place of $\mathbf{V}$ in the formula for $\hat{\beta}$ and the same ANOVA tests can still be used. But now the distribution of the test statistic will have to be approximated.

The maximum likelihood approach, when $\mathbf{V}$ is known, yields the same estimate of β as the GLE. Further, the likelihood ratio test for testing $L\beta = 0$ will still hold through if a $\sqrt{n}$-consistent estimate of $\mathbf{V}$, instead of the maximum likelihood

43

estimate, is utilized. This fact is useful since in many problems with structured covariances, finding the maximum likelihood estimates of $\theta_1, ..., \theta_k$ is generally hard, if not impossible. However, consistent estimates of $\theta_1, ..., \theta_k$ and hence of $\mathbf{V}$ can be obtained easily.

Another serious drawback of the multivariate and ANOVA methods is that they cannot be applied when data are unbalanced, that is, when different number of observations are available on different individuals. Further, the multivariate tests have been known to be less powerful, when the variance covariance matrix of the observation vector has a certain structure. Thus, either if data are unbalanced or the covariance matrix has a structure other than the equicorrelation structure, the maximum likelihood approach or the two stage GLE approach seems to be the right approach. Further, this general approach can be utilized when the covariates are time varying. Modification of this approach to accommodate unbalanced data will be discussed in section 3.2.4.

3.1.3   Example 1

In this section, we illustrate methods described in section 3.1.1 and 3.1.2 to analyze repeated measures data with fixed covariates. The data set in Table 1 is courtesy of Center for Pediatric Research, Eastern Virginia Medical School, Norfolk, Virginia and it has been discussed and analyzed by Pickering, Marrow, Herrara etal.,(1995). The purpose of the study was to see the effect of maternal immunization on the

44

serum anti-rotavirus antibody concentration. Thirty two women were randomized into three groups and received a single oral dose of either rhesus rotavirus monovalent reassortant vaccine (10^4 pfu), rhesus rotavirus tetravalent vaccine (10^4 pfu), or placebo. Measurements on the antibody concentration were taken. One measurement was taken before the vaccination and six other measurements were collected after a week, one month, two months, three months, four months and five months after the vaccination. Measurement taken before the vaccination is considered a baseline measurement or a covariate. Since there are many missing values for months four and five we have omitted them from our analysis.

Prior to carrying out statistical analysis it is important to verify if the data satisfy the assumption of multivariate normality. This was done using tests based on Mardia's *multivariate skewness* and *kurtosis* measures. See Khattree and Naik (1995) for SAS programs for computing these tests. For this data set with its small p-values, (p-$value(skewness) = 0$ and p-$value(kurtosis) = 0$), there is an indication that the assumption of multivariate normality is violated. We make a log-transformation on these data, and then again test for multivariate normality. We found that p-values for the tests based on skewness and kurtosis respectively are 0.1190 and 0.4180. This indicates that the log transformed data can be assumed to have come from a multivariate normal distribution. Consequently, we apply our methods to the log-transformed data.

Analysis using multivariate technique. Here we present our analysis using

45

Table 3.1: Serum Anti-Rotavirus Antibody Concentration

id	vac	pre	wk1	m1	m2	m3	m4	m5
5001	P	80	20	80	40	80	.	.
5007	P	320	40	40	40	40	40	40
5010	P	160	80	320	.	.	.	
5013	P	80	40	40	40	40	40	.
5014	P	640	80	80	80	160	1280	.
5017	P	1280	320	80	40	40	40	40
5022	P	80	40	20	40	20	.	.
5025	P	160	40	40	40	.	.	
5027	P	320	40	80	40	40	.	.
5031	P	80	40	40	20	40	.	.
5037	P	160	40	40	40	40	40	.
5005	R	160	80	80	80	80	160	160
5006	R	160	80	40	40	160	160	320
5008	R	160	80	80	80	1280	.	.
5012	R	320	320	80	160	.	.	
5016	R	1280	320	1280	320	1280	1280	.
5024	R	320	320	640	640	.	.	.
5026	R	640	320	80	160	80	80	.
5029	R	640	160	320	160	160	160	160
5032	R	1280	160	160	160	160	160	.
5033	R	40	40	40	20	40	.	.
5036	R	160	80	160	320	160	320	.
5040	R	80	320	320	320	320	.	.
5002	S	320	320	160	80	160	160	.
5004	S	160	160	80	80	80	320	160
5009	S	640	20	40	.	.	.	
5019	S	320	320	320	160	160	160	160
5020	S	80	40	80	80	640	.	.
5023	S	80	160	1280	.	.	.	
5034	S	1280	160	160	640	.	.	
5035	S	320	160	320	160	320	.	.
5041	S	160	40	40	40	40	.	.

46

multivariate technique. The following considerations are in order.

1. The effect of covariates may be different in various treatment groups. This can be tested by testing for statistical significance of the interactions of the treatment with covariates.

2. Covariates may influence the measurements taken at different time points very differently. This can be tested by examining the significance of the interaction between covariates and the time factor.

We summarize the results of the SAS output in the following table:

Hypothesis	Wilks' λ	Num DOF	Den DOF	p-value
Covariate*Group	0.60942	8	32	0.3741
Covariate	0.40258	4	16	0.0040
Group	0.60931	8	32	0.3738

It can be observed that the multivariate tests show that the effect of covariate over different treatment groups do not differ significantly (p-value=0.3741). However, the effect of the covariate itself is significant (p-value=0.0040). Therefore, it is important to include the covariate in the model. We also notice that there is no overall group effect.

Next we would like to examine whether the covariate influences the measurements taken at different time points differently. This can be achieved using RE-PEATED statement of SAS. The results are summarized in the following table:

47

Hypothesis	Wilks' λ	Num DOF	Den DOF	p-value
Time*Covariate	0.62301	3	17	0.0408

We see that the covariate does affect the measurements taken at different time points differently if 0.05 level of significance is used. However, if the level of significance is assumed to be 0.01 then the conclusion would be that the covariate does not affect differently.

Analysis using univariate approach. We present our results of the analysis using univariate model in the following table:

Hypothesis	Num DOF	Den DOF	F value	p-value
Covariate	1	19	7.74	0.0119
Group	2	19	0.21	0.8099
Group*Covariate	2	19	0.31	0.7381
Time	3	57	3.89	0.0135
Time*Covariate	3	57	3.71	0.0166
Time*Group	6	57	1.35	0.2489
Time*Covariate*Group	6	57	1.43	0.2191

We can see that the univariate tests indicate that the groups do not differ significantly (p-value=0.8099). The effect of the time factor is significant at 5% level of significance (p-value=0.0135). The interaction between time and group is

48

not significant (p-value=0.2489). We also see that the covariate does not affect different vaccination groups differently (p-value=0.7381).

3.2   Time varying covariates

As we mentioned in the introduction, in many repeated measures study, the covariates vary over time. Again assuming, for simplicity, that there is only one covariate, the data in this case can be represented by y_{ijk}, x_{ijk}, $k = 1, ..., t$, $j = 1, ..., n_i$, $i = 1, ..., g$. Analysis of these data seems to be a challanging problem. Only recently attempts have been made, in the literature, to analyze these data. We review these methods below.

3.2.1   The multivariate approach

Patel (1986) proposed the following multivariate model for analyzing repeated measures designs with time varying covariates. Let Y'_{ij} be $1 \times t$ vector of response variables and X'_{ij} be $1 \times t$ vector of covariates, both taken over t occasions on an individual. Next, define $Y_{n \times t} = (Y_{11}, ..., Y_{1n_1}, Y_{21}, ..., Y_{2n_2}, ..., Y_{gn_g})'$, $n = \sum_{i=1}^{g} n_i$ and $X_{n \times t} = (X_{11}, ..., X_{1n_1}, X_{21}, ..., X_{2n_2}, ..., X_{gn_g})'$. Then the following model

$$Y_{n \times t} = A_{n \times m} \xi_{m \times t} + X_{n \times t} \Gamma_{t \times t} + E_{n \times t}, \tag{3.6}$$

where A is a design matrix, ξ is a matrix of unknown parameters, Γ is a diagonal matrix with the unknown diagonal elements, $\gamma_1, ..., \gamma_t$, and E is a $n \times t$ error matrix,

49

can be utilized to represent the time varying covariates. Note that the model (3.6) is different from the usual multivariate analysis of covariance model in the sense that its parameter matrix Γ is known (to be zero) except the diagonal entries. This makes it hard to handle the analysis of this model in a routine way.

As usual rows of E are assumed to be independently distributed with a common multivariate normal distribution with a zero mean vector and a $t \times t$ covariance matrix $\mathbf{V}$. Patel (1986) has provided the maximum likelihood estimators of the parameters. He has also provided an iterative algorithm describing the computation of the unknown parameters and the likelihood ratio test for any general linear hypothesis of the form $L_{b \times m} \xi_{m \times t} M_{t \times c} = 0$.

The likelihood function under model (3.6) is

$$L = (2\pi)^{-\frac{1}{2}nt} |\mathbf{V}|^{-\frac{1}{2}n} exp - \frac{1}{2} tr(\mathbf{V}^{-1}R), \qquad (3.7)$$

where $R = (Y - A\xi - X\Gamma)'(Y - A\xi - X\Gamma)$. The maximum likelihood estimators of ξ and Γ are obtained by minimizing $\phi = ln|R|$. The idea is that, if ξ and Γ are known then the MLE of $\mathbf{V}$ would be $\frac{R}{n}$. In that case the log-likelihood function reduces to $ln|R|$ besides a constant. Hence the procedure iterates between $\hat{\mathbf{V}} = \frac{\hat{R}}{n}$ and minimizing $ln|R|$ with respect to ξ and Γ, $\hat{R}$ being R when ξ and Γ are replaced by their estimates, say $\hat{\xi}$ and $\hat{\Gamma}$.

Note that

$$\frac{\partial \phi}{\partial \xi} = 0 \quad \Rightarrow \quad \hat{\xi} = (A'A)^{-1}A'(Y - X\hat{\Gamma}) \qquad (3.8)$$

50

and $\quad \dfrac{\partial \phi}{\partial \Gamma} \quad = \quad -2 diag\{R^{-1}X'(Y - A\hat{\xi} - X\Gamma)\} = 0$

implies

$$diag\{\hat{R}^{-1}X'(I - A(A'A)^{-1}A')Y\} = diag\{\hat{R}^{-1}X'X\hat{\Gamma}\}. \tag{3.9}$$

Patel (1986) has suggested the following iterative steps to solve the ML equations:

Step 1. Taking $\hat{R} = I$, compute $\hat{\Gamma}$ from (3.9). This estimate of Γ is same as that obtained by minimizing $tr(R)$ w.r.t. Γ .

Step 2. Compute $\hat{\xi}$ from (3.9)

Step 3. Compute $\hat{R}$ using $\hat{\xi}$ and $\hat{\Gamma}$. Then compute $\hat{\phi} = log \ |\hat{R}|$.

Step 4. Compute a revised estimate of Γ from (3.9) using $\hat{R}$ obtained in step 3.

Step 5. Repeat steps 2, 3 and 4 until the absolute difference between two successive values of $\hat{\phi}$ is less than a pre-determined number.

Patel (1986) has also given similar algorithm for computing the ML estimates under the null hypothesis $H_0 : L\xi M = 0$. Using these MLE's the likelihood ratio test for testing H_0 can be obtained. Recall that the likelihood function under model (3.6) is given by

$$L = (2\pi)^{-\frac{1}{2}nt}|\mathbf{V}|^{-\frac{1}{2}n}exp-\frac{1}{2}tr(\mathbf{V}^{-1}R).$$

51

We compute the maximum of L as

$$max\,L = (2\pi)^{-\frac{1}{2}nt}|\hat{\mathbf{V}}|^{-\frac{1}{2}n}exp(-\frac{1}{2}nt), \qquad (3.10)$$

where $\hat{\mathbf{V}} = \hat{R}/n$. Similarly, writing the maximum likelihood estimator of $\mathbf{V}$ under H_0 as $\tilde{\mathbf{V}} = \tilde{R}/n$, we get the maximum of L under $H_0 : L\xi M = 0$ as

$$max\,L_{H_0} = (2\pi)^{-\frac{1}{2}nt}|\tilde{\mathbf{V}}|^{-\frac{1}{2}n}exp(-\frac{1}{2}nt). \qquad (3.11)$$

From (3.10) and (3.11) the likelihood ratio test for testing H_0 is to reject the null hypothesis if

$$\lambda = |\hat{\mathbf{V}}\tilde{\mathbf{V}}^{-1}|^{\frac{1}{2}n} \leq C_\alpha, \qquad (3.12)$$

where C_α is a constant to be determined satisfying $pr(\lambda \leq C_\alpha|H_0) = \alpha$. We use the standard result, that is, $-2\,log\lambda$ has asymptotically a chi-squared distribution with $mt - bc$ degrees of freedom, to test the hypothesis.

3.2.2 Example 2

In this section we illustrate the use of Patel's algorithm to analyze repeated measures data with time varying covariates. We use data given in Table 2, courtesy of Dr.Barbara Hargrave of Biological Sciences department, Old Dominion University. The effects of phenylephrine induced increase in arterial pressure on the secretion of atrial natriuretic peptide (ANP) in the ovine fetus are studied by Hargrave and Castle (1995). In this study 16 chronically cannulated fetal sheeps were divided into

52

two groups. Arterial pressure was increased by infusing phenylephrine to the fetus from each of the two groups. Systematic mean arterial pressure (MAP), plasma ANP concentrations and plasma renin activity (PRA) were measured at three time points (5 min, 15 min, and 30 min) after infusion. We take PRA as the response variable and MAP as time varying covariate. Here we have $g=2$, $t=3$, $n_1 =6$ and $n_2 =10$ and one covariate. We illustrate the computation of λ given by (3.12) for testing various hypotheses by choosing different L and M.

Let $L = I_{2\times2}$ and

$$M = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Then the hypothesis of possible differences between different time points can be tested using the hypothesis $L\xi M = 0$. The ML estimates of $n\Sigma$, under the nonnull and null hypotheses respectively are

$$\hat{R} = \begin{bmatrix} 19807.802 & -4186.271 & 4663.295 \\ -4186.271 & 48285.522 & 30775.957 \\ 4663.295 & 30775.957 & 24974.933 \end{bmatrix}$$

and

$$\tilde{R} = \begin{bmatrix} 20212.801 & -4704.252 & 5694.256 \\ -4704.252 & 48330.394 & 30810.573 \\ 5694.256 & 30810.573 & 28866.521 \end{bmatrix}.$$

53

Table 3.2: Plasma renin activity data

group	subject	time	PRA	MAP	ANP
young	528	5	66	135	-20
young	528	15	-12	545	-28
young	528	30	60	840	-24
young	502	5	59	21	9.8
young	502	15	59	97	20
young	502	30	46	224	45
young	505	5	47	55	20
young	505	15	6.7	78	41
young	505	30	6.7	210	39
young	24	5	-16	22	13
young	24	15	2.6	-28	36
young	24	30	-18.4	-33	49
young	599	5	-29	-22	19
young	599	15	-53	155	45
young	599	30	-61	629	50
young	49	5	6.9	7.6	16
young	49	15	-29	169	44
young	49	30	-35	754	60
old	717	5	-16	55	18
old	717	15	-40	158	18
old	717	30	-38	133	24
old	722	5	29	875	35
old	722	15	-40	158	18
old	722	30	-38	133	24
old	617	5	5.8	23	4.3
old	617	15	-29	56	-2.1
old	617	30	5.8	177	-2.1
old	618	5	-40	86	7.4
old	618	15	-60	121	22
old	618	30	-58	82	28
old	29	5	-80	-9.6	15
old	29	15	170	2.4	19
old	29	30	80	76	15
old	79	5	13	-16	-8.4
old	79	15	-45	0.2	-12
old	79	30	-25	-16	-6.7
old	610	5	-1.6	46	0.2
old	610	15	4.1	104	1.6
old	610	30	-7.1	25	3.2
old	623	5	60	31	5
old	623	15	6.6	81	0.2
old	623	30	0.2	442	0.2
old	332	5	16	15	4.9
old	332	15	-73	85	34
old	332	30	-41	830	36
old	20	5	-17	13	-17
old	20	15	-38	54	7.7
old	20	30	-50	205	0.2

54

The value of $\lambda=0.0025$ and $-2\,ln\lambda=11.9920$. Similarly we can test other hypotheses of interest with the appropriate choices of L and M. We summarize our results below:

Hypothesis	λ	D.O.F	Approximate χ^2 value	p-value
Group	0.1042	2	4.5225	0.1042
Time	0.0025	2	11.9920	0.0025
Time*Group	0.5993	2	1.0241	0.5993

From the above table we see that only the *time effect* is significant.

3.2.3 SUR model approach

Verbyla (1988) pointed out that model (3.6) can be expressed as Zellner's seemingly unrelated regression (SUR) model and a two stage estimation procedure can be utilized.

To illustrate, let $\xi = [\xi_1 : ... : \xi_t]$, $X = [X_1 : ... : X_t]$ and the elements of Γ be denoted by $\gamma_1, ..., \gamma_t$. Then model (3.6) can be written as

$$
\begin{aligned}
E(Y) &= A[\xi_1 : ... : \xi_t] + [X_1\gamma_1 : ... : X_t\gamma_t] \\
&= [A\xi_1 + X_1\gamma_1 : ... : A\xi_t + X_t\gamma_t] \\
&= [A_1\theta_1 : ... : A_t\theta_t],
\end{aligned}
$$

where $A_k = [A : X_k]$ and $\theta_k = \begin{pmatrix} \xi_k \\ \gamma_k \end{pmatrix}$.

55

Thus writing $E(Y) = Y = [Y_1 : \ldots : Y_t]$ we have

$$E[Y_1 : \ldots : Y_t] = [A_1\theta_1 : \ldots : A_t\theta_t], \tag{3.13}$$

which is a SUR model with the covariance matrix of $vec(Y') = I \otimes V$. Now consider the maximum likelihood estimation based on (3.13) and the covariance $I \otimes V$ under the assumption of normality. Then the log of the likelihood function can be written as

$$log\ l = -\frac{n}{2}ln|V| - \frac{1}{2}\sum_{i=1}^{t}\sum_{j=1}^{t}tr\{v^{ij}(Y_i - A_i\theta_i)'(Y_j - A_j\theta_j)\}, \tag{3.14}$$

where v^{ij} is the $(i,j)^{th}$ element of V^{-1}. Differentiating $log\ l$ with respect to θ_k and equating the derivative to zero, we get

$$\hat{\theta}_k = (A_k'A_k)^{-1}A_k'Y_k + (A_k'A_k)^{-1}A_k'\sum_{i\neq k}\{(Y_i - A_i\theta_i)v^{ik}(v^{kk})^{-1}\}. \tag{3.15}$$

The solutions $\hat{\theta}_k, k = 1, \ldots, t$ can be obtained using the algorithm provided by Verbyla and Venables (1988). Equation (3.15) can be regarded as coming from the conditional model with the conditional expectation and variance respectively as

$$E[Y_k/Y_i, i \neq k] = A_k\theta_k + \sum_{i\neq k}(Y_i - A_i\theta_i)B_{ik} \tag{3.16}$$

$$var(Y_k/Y_i, i \neq k) = I_n \otimes (v^{kk})^{-1}, \tag{3.17}$$

where $B_{ik} = -v^{ik}(v^{kk})^{-1}$.

An algorithm for estimation is as follows:

Step 1. Estimate the initial θ_k marginally by $\theta_k^0 = (A_k'A_k)^{-1}A_k'Y_k$

56

Step 2. To find the jth iterate , begin with θ_t, and regress Y_t on $A_t, R_1^{j-1}, ..., R_{t-1}^{j-1}$, where θ_k^{j-1} is the $(j-1)^{st}$ step estimate of θ_k and $R_k^{j-1} = Y_k - A_k \theta_k^{j-1}$.

Step 3. For the jth iterate of θ_k, regress Y_k on $A_k, R_1^{j-1}, ..., R_{k-1}^{j-1}, R_{k+1}^{j-1}, ..., R_t^j$

Step 4. Continue until convergence.

Once the estimates of $\theta_1, ..., \theta_k$ are obtained using the algorithm, the maximum likelihood estimator of $\mathbf{V}$ can be derived using the residuals as $\hat{\mathbf{V}} = (Y - [A_1 \hat{\theta}_1 : ... : A_t \hat{\theta}_t])'(Y - [A_1 \hat{\theta}_1 : ... : A_t \hat{\theta}_t])/n$. Verbyla (1988) has also provided an algorithm to estimate the parameters under the null hypothesis $H_0 : L\xi M = 0$. Let the ML estimator of $\mathbf{V}$ under H_0 be $\tilde{\mathbf{V}}$. Then the likelihood ratio test can be constructed in the usual manner to test any linear hypothesis. The advantage of SUR model is in its simplicity of estimation. Only regression calculations are required making its implementation quite simple.

In the following, we use the statistic given in (3.12) for testing hypothesis of the form $L\xi M = 0$. The estimators of the covariance matrix under the null and the non-null hypotheses are obtained under SUR model set up. We apply this approach on the data set given in Example 2. The results are presented in the following table.

57

Hypothesis	λ	D.O.F	Approximate χ^2 value	p-value
Group	0.0596	2	5.6402	0.0596
Time	0.0255	2	7.3360	0.0255
Time*Group	0.2406	2	2.8488	0.2406

Notice from the above table, as before, that only the *time effect* is significant.

3.2.4 An alternative linear model approach

Although SUR model approach seems like a good alternative to the multivariate approach, both of these approaches have serious shortcomings. None of these approaches will handle unbalanced data. Suppose the ith individual has t_i, $i = 1, ..., n$ repeated measurements. Then the analysis of data is more complex and both of the approaches fail. However, an alternative modeling which can handle the unbalanced data will be described below. This alternative approach is also noted by Verbyla and Cullis (1990).

Consider model (3.6)

$$Y_{n \times t} = A_{n \times m} \xi_{m \times t} + X_{n \times t} \Gamma_{t \times t} + E_{n \times t},$$

where $\Gamma = diag\ (\gamma_1, ..., \gamma_t)$. Then the ith row of Y can be modeled as

$$y'_{i_1 \times t} = a'_i[\xi_1 : ... : \xi_t] + [x_{i1}, ..., x_{it}]diag(\gamma_1, ..., \gamma_t) + \epsilon'_i \tag{3.18}$$

$$= [a'_i \xi_1 : ... : a'_i \xi_t] + [\gamma_1, ..., \gamma_t]diag(x_{i1}, ..., x_{it}) + \epsilon'_i. \tag{3.19}$$

58

Taking transpose of both sides of (3.19) we can write

$$
y_i = \begin{bmatrix} a_i' & 0 & . & . & . & 0 \\ 0 & a_i' & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & a_i' \end{bmatrix} \begin{bmatrix} \xi_1 \\ . \\ . \\ . \\ \xi_t \end{bmatrix} + \begin{bmatrix} x_{i1} & 0 & . & . & . & 0 \\ 0 & x_{i2} & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & x_{it} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ . \\ . \\ \gamma_t \end{bmatrix} + \epsilon_i
$$

$$
= A_{i_t \times mt}\eta_{mt \times 1} + D_{i_t \times t}\gamma_{t \times 1} + \epsilon_i.
$$

where $A_i = diag(a_i', ..., a_i')$, $D_i = diag(x_{i1}, ..., x_{it})$, $\eta' = [\xi_1', .., \xi_t']$ and $\gamma' = (\gamma_1, ..., \gamma_t)$,

Further, the above equation can be rewritten in the form

$$
y_i = [A_i : D_i]\begin{bmatrix} \eta \\ \gamma \end{bmatrix} + \epsilon_i
$$

$$
y_{i_t \times 1} = B_{i_t \times (m+1)t}\beta_{(m+1)t \times 1} + \epsilon_i,
$$

where $B_i = [A_i : D_i]$, and $\beta = \begin{bmatrix} \eta \\ \gamma \end{bmatrix}$.

Now let $y' = (y_1', ..., y_n')$ and $\epsilon' = (\epsilon_1', ..., \epsilon_n')$. Then

$$
y = \begin{bmatrix} B_1 \\ . \\ . \\ . \\ B_t \end{bmatrix}\beta + \epsilon \tag{3.20}
$$

or

$$
y = B\beta + \epsilon, \tag{3.21}
$$

59

which is an usual linear model with $cov(\epsilon) = \Omega = I_n \otimes \mathbf{V}$, where $\mathbf{V}$ is a $t \times t$ covariance matrix of the repeated measures, and $B = (B'_1, ..., B'_t)'$.

Then the general least squares estimator of β is

$$\hat{\beta} = (B'\Omega^{-1}B)^{-1}B'\Omega^{-1}\mathbf{y}. \tag{3.22}$$

Writing $B' = (B'_1 : ... : B'_n)$ and $\Omega^{-1} = I \otimes \mathbf{V}^{-1}$, we have

$$\hat{\beta} = \{\sum_{i=1}^{n} B_i \mathbf{V}^{-1} B_i\}^{-1}\{\sum_{i=1}^{n} B_i \mathbf{V}^{-1} y_i\}. \tag{3.23}$$

We may use any two stage procedure to estimate β and $\mathbf{V}$. Further, any hypothesis of interest can be tested using the general linear hypothesis $H_0 : L\beta = 0$.

For any specific structured covariance matrix $\mathbf{V}$, this modeling approach can be taken. The SAS statistical procedure PROC MIXED can be adopted to find the estimates and to test any hypothesis of interest. See Khattree and Naik (1995) for some data analysis using PROC MIXED.

As it was pointed out earlier, this modeling technique enables us to deal with unbalanced data as well. Suppose y_i is vector of order $t_i \times 1$ containing the repeated measurements on the ith individual. Let $\mathbf{y}'$ and ϵ' be as defined before. The order of $\mathbf{y}$ and ϵ is now $\sum_{i=1}^{n} t_i \times 1$. Then the model is

$$\mathbf{y} = B\beta + \epsilon,$$

where $cov(\epsilon) = diag(\mathbf{V}_1, ..., \mathbf{V}_n)$, and $\mathbf{V}_i$ is of order $t_i \times t_i$.

Assuming a specific structure for $\mathbf{V}_i$ becomes almost inevitable in this case. It is common to assume that $\mathbf{V}_i$ depends only on a few parameters, say $\theta_1, ..., \theta_k$. See

60

Jennrich and Schluchter (1986) for a list of possible structures that can be used. Common structures for $\mathbf{V}_i$ are equicorrelation ($\theta_1 = \sigma^2, \theta_2 = \rho$) and autoregressive ($\theta_1 = \sigma^2, \theta_2 = \rho$) structures. The maximum likelihood estimation method is adopted to estimate the unknown parameters $\theta_1, ..., \theta_k$ and β. The form of the MLE of β remains as in (3.23) except that the matrix $\mathbf{V}$ has to be replaced by the matrix $\mathbf{V}_i$ that is of order $t_i \times t_i$ and B_i is of order $t_i \times (m+1)t_i$, for $i = 1, ..., n$. These (unbalanced) data can also be analyzed by PROC MIXED procedure of SAS software.

61

Chapter 4

Analysis Of Multivariate

Repeated Measurements

In this chapter we consider a set of t repeated measurements on p variables (or

characteristics) on each of the n individuals. Thus data on each individual is a $p \times t$

matrix. The n individuals themselves may be divided and randomly assigned to g

groups. Analysis of these data using a MANOVA model is considered. The well

known Satterthwaite type approximation to the distribution of a quadratic form in

normal variable is extended to the distribution of a multivariate quadratic form in

multivariate normal variates. The multivariate tests using this approximation are

developed for testing the usual hypotheses.

When measurements on a variable (or a characteristic) are made at several oc-

casions or under different treatment conditions on the same experimental unit, we

have repeated measures (or longitudinal) data. Repeated measures data routinely occur in many diverse fields like Medicine, Psychology, and Education. Analysis of these data needs special care since the measurements made at different occasions on the same individual may quite likely be correlated. A typical set of repeated measures data taken on $n(= n_1 ... + n_g)$ individuals forming g groups over t occasions (time points) are shown in Figure 4.1. The measurement taken on the jth individual belonging to the ith group at the tth occasion is denoted by y_{ijk}, where $k = 1, ..., t$, $j = 1, ..., n_i$, $i = 1, ..., g$.

The problems of interest are to test for (i) the *time effect*, (ii) the *group effect*, and (iii) the *time*group interaction*. Let us assume that the vector of t measurements on each individual is a sample from a $t-$variate normal distribution with a certain positive definite covariance matrix, say $\mathbf{V}$. The above three problems can be solved using *profile analysis*, a standard technique in the multivariate statistical analysis. For example, see Morrison (1976). In many practical problems where the repeated measures occur, the covariance matrix $\mathbf{V}$ is found to have some structure. In that case, the multivariate tests are found to be less powerful and use of univariate analysis of variance or regression models is recommended.

If $\mathbf{V}$ has the simple structure $\sigma^2 \mathbf{I}$, it is clear that the above three problems, (i)-(iii), can be written in the form of testing of hypothesis problems in a two-way analysis of variance model. Suppose $\mathbf{V} = \sigma^2 \mathbf{V}(\rho)$, where $\mathbf{V}(\rho) = (1 - \rho)\mathbf{I} + \rho \mathbf{1}\mathbf{1}'$, $\mathbf{I}$ is an identity matrix and $\mathbf{1}$ is a vector of ones. Then the analysis of repeated

63

measures data is essentially same as that of a typical split-plot design. The usual tests of hypotheses in the split-plot design will address the problems (i)-(iii) (see Winer (1971)).

Many authors, for example, Baldessari (1965), Huynh and Feldt (1970), and Rouanet and Lepine (1970) have characterized the class of all covariance structures for $\mathbf{V}$, such that the split-plot type of analysis will remain invariant. A typical member of this class, for a fixed vector $\mathbf{a}' = (a_1, ..., a_t)$, is

$$\mathbf{V} = \sigma^2(\mathbf{I} + \mathbf{a1}' + \mathbf{1a}') = \sigma^2 H(\mathbf{a}). \tag{4.1}$$

This structure is termed as type H structure. Recently, Chaganty and Vaish (1995) studied this characterization more closely and produced the following form for $H(\mathbf{a})$:

$$H^*(\mathbf{a}) = \mathbf{I} + \frac{1}{t}(\mathbf{a1}' + \mathbf{1a}') - \frac{1}{t}(1 + \bar{a})\mathbf{11}',$$

where $\mathbf{a}' = (a_1, ..., a_t)$ is such that

$$\frac{1}{t}\sum_{i=1}^{t}(a_i - \bar{a})^2 \le \bar{a} \tag{4.2}$$

and $\bar{a}$ is the mean of the components of the vector $\mathbf{a}$. This characterization by Chaganty and Vaish (1995) is most explicit in the sense that it gives easy condition (4.2) on the elements of $\mathbf{a}$ such that $\mathbf{V}$ is positive (semi) definite. The likelihood ratio test for testing for type H structure, using the results of Mauchly (1940), is given by Huynh and Feldt (1970). For certain efficiency studies involving these type of structures see Jensen (1982).

64

Group	Subject	1	2	. . .	t
	1	y_{111}	y_{112}	$\cdot\ \cdot\ \cdot$	y_{11t}
1	2	y_{121}	y_{122}	$\cdot\ \cdot\ \cdot$	y_{12t}
	$\cdot$	$\cdot$	$\cdot$	$\cdot\ \cdot\ \cdot\ \cdot$	
	n_1	y_{1n_11}	y_{1n_12}	$\cdot\ \cdot\ \cdot$	y_{1n_1t}
	1	y_{211}	y_{212}	$\cdot\ \cdot\ \cdot$	y_{21t}
2	2	y_{221}	y_{222}	$\cdot\ \cdot\ \cdot$	y_{22t}
	$\cdot$	$\cdot$	$\cdot$	$\cdot\ \cdot\ \cdot\ \cdot$	
	n_2	y_{2n_21}	y_{2n_22}	$\cdot\ \cdot\ \cdot$	y_{2n_2t}
	$\cdot$	$\cdot$	$\cdot$	$\cdot\ \cdot\ \cdot\ \cdot$	
	$\cdot$	$\cdot$	$\cdot$	$\cdot\ \cdot\ \cdot\ \cdot$	
	1	y_{g11}	y_{g12}	$\cdot\ \cdot\ \cdot$	y_{g1t}
g	2	y_{g21}	y_{g22}	$\cdot\ \cdot\ \cdot$	y_{g2t}
	$\cdot$	$\cdot$	$\cdot$	$\cdot\ \cdot\ \cdot\ \cdot$	
	n_g	y_{gn_g1}	y_{gn_g2}	$\cdot\ \cdot\ \cdot$	y_{gn_gt}

Table 4.1: Repeated measures data

A strategy for analysis of repeated measures data may be as follows: Test for the type H covariance structure; if accepted adopt the univariate split-plot type of analysis for solving (i)-(iii), else adopt the profile analysis. When the hypothesis of type H structure for V is rejected, one can still use the F-statistics of the split-plot analysis. However, the distributions of these test statistics will no longer be exact. Box (1954) and Geisser and Greenhouse (1958) have developed usable approximations to the distributions of these F-statistics using the result of Satterthwaite (1941). Satterthwaite (1941) approximated the distribution of a quadratic form in normal variables, by a scale multiple of a chi square distribution such that the expected value and the variance of the quadratic form are equal to those of the approximating quantity. This approximation is called Satterthwaite approximation in the literature. Practical implementation of this method requires estimation of the degrees of freedoms of the approximate chi square or the F distributions. Various procedures to estimate the degrees of freedoms are available in the literature. For example, see Greenhouse and Geisser (1959) and Huynh and Feldt (1976). Several authors have studied the effect of (a) heterogeneity of covariance, (b) estimating the degrees of freedom by different methods, and (c) sample size considerations, on the distributions of the F-statistics involved, using simulation experiments. For example, see Huynh and Feldt (1980), Keselman and Keselman (1990), and Quintana and Maxwell (1994). The complete analysis of repeated measures data, including the estimation of the degrees of freedom, has been successfully implemented

66

in several statistical softwares. A detailed explanation of the analysis of repeated measures data using the SAS software can be found in Khattree and Naik (1995).

When the observations on n experimental units are made on a set of p variables (or characteristics) at t occasions, we have what can be termed as a set of *multivariate repeated measures* data. Analysis of these data is further complicated by the existence of correlation among the measurements on different variables along with the correlation among measurements taken at different occasions. A typical set of multivariate repeated measures data would be in the same form as Figure 4.1, except that, y_{ijk} now is a $p \times 1$ vector of measurements on p characteristics. For clarity we denote this vector by Y_{ijk}. Analysis of these type of data is considered in this article. Considering a mixed effects MANOVA (Multivariate Analysis of Variance) model we formulate the usual testing of hypothesis problems. The approximate distributions of various SSCP (sums of squares and crossproducts) matrices are derived and used to test the hypotheses. The derivation of the approximate distribution of a SSCP matrix (a multivariate quadratic form) is in the spirit of Satterthwaite (1941), Box (1954) and Geisser and Greenhouse (1958).

4.1 General covariance structure

Suppose, as before, a set of repeated measurements at t occasions are taken on p variables on each of $n(= n_1 + ... + n_g)$ individuals belonging to g groups. The

67

problems of interest are (i) test for the difference between groups, (ii) test for the difference between the occasions and (iii) test for group and occasion interaction. It may also be of interest to test certain hypothesis about the mean vector of p variables. We adopt the multivariate analysis of variance (MANOVA) technique to analyze the data so that our questions of interest can be answered.

Let Y_{ijk}, $k = 1, ..., t$, $j = 1, ..., n_i$, $i = 1, ..., g$, be a $p \times 1$ vector of measurements on the jth individual in the ith group at the kth occasion. Let $Y'_{ij} = (Y'_{ij1}, ..., Y'_{ijt})$. Then Y_{ij} is $pt \times 1$ random observational vector corresponding to the jth individual in the ith group. Let $cov(Y_{ij}) = \Omega$, for $j = 1, ..., n_i, i = 1, ..., g$, where Ω is a positive definite matrix. Define the $n \times pt$ matrix $\mathbf{Y}$ as $\mathbf{Y} = (Y_{11}, ..., Y_{gn_g})'$. Then the multivariate repeated measures data can be modeled as

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E},$$

where $\mathbf{X}$ is a known design matrix, $\mathbf{B}$ is the matrix of unknown parameters, and $\mathbf{E}$ is the matrix of errors such that the rows of $\mathbf{E}$ are independently distributed as multivariate normal with zero mean vector and variance covariance matrix Ω. Any hypothesis about the effect of the time factor, the group factor, or the interaction between them or any linear hypothesis about the expected values of Y_{ijk} can be formulated in the form of the general linear hypothesis

$$\mathbf{LBM} = \mathbf{0} \qquad (4.3)$$

for known and full rank matrices $\mathbf{L}$ and $\mathbf{M}$. Using Wilks' Λ or any other standard

68

multivariate tests the null hypothesis (4.3) can be tested. This approach for analyzing the multivariate repeated measures data is commonly adopted in practice (see Timm, 1980). Suppose $\Omega = I_t \otimes \Sigma$, for a $p \times p$ positive definite matrix Σ, to indicate that the measurements taken over time are uncorrelated. Then the multivariate repeated measures data can be analyzed using two-way MANOVA model

$$Y_{ijk} = \mu + \alpha_i + \beta_k + (\alpha\beta)_{ik} + \epsilon_{ijk}, \qquad (4.4)$$

where μ, α_i, β_k, $(\alpha\beta)_{ik}$ are all $p \times 1$ vectors with the usual meaning as in the two-way MANOVA model. The hypotheses of interest are

(i)$\alpha_1 = ... = \alpha_g$ (no group effect), (ii)$\beta_1 = ... = \beta_t$ (no time or occasion effect) and

(iii)$(\alpha\beta)'_{ik}s$ are all equal (no group and time interaction). Under the assumption $\Omega = I_t \otimes \Sigma$, the errors ϵ_{ijk} are independently distributed with zero mean vector and variance covariance matrix Σ. Testing of hypothesis problems (i)-(iii), can be tackled using the usual MANOVA technique.

Next, to accommodate simple correlation among repeated measures, let us consider mixed effects MANOVA model (similar to the split-plot design model of the univariate analysis of the usual repeated measures data). The MANOVA table similar to the ANOVA table for split-plot design model, is shown in Table (4.2). Here the $p \times nt$ matrix Y is defined as

$Y = (Y_{111}, ,..., Y_{11t},..., Y_{1n_11},..., Y_{1n_1t},..., Y_{gn_g1},..., Y_{gn_gt})$ and $N = nt$. The ma-

69

Source	D.O.F.	SSCP Matrix	Dist. Under H_0
Between Groups			
Groups	$g - 1$	$Q_1 = Y A_1 Y'$	$W_p(g - 1, \Sigma)$
Individuals	$n - g$	$Q_2 = Y A_2 Y'$	$W_p(n - g, \Sigma)$
Within Groups			
Time	$t - 1$	$Q_3 = Y A_3 Y'$	$W_p(t - 1, \Sigma)$
Time*Group	$(g - 1)(t - 1)$	$Q_4 = Y A_4 Y'$	$W_p((t - 1)(g - 1), \Sigma)$
Error	$(t - 1)(n - g)$	$Q_5 = Y A_5 Y'$	$W_p((t - 1)(n - g), \Sigma)$
Total	$N - 1$	$Y(I_N - \frac{1}{N} J_N)Y'$	

Table 4.2: Manova table for mixed effects model

70

trix quadratic forms $Q_1 - Q_5$ are

$$Q_1 = t \sum_{i=1}^{g} n_i (\bar{Y}_{i..} - \bar{Y}_{...})(\bar{Y}_{i..} - \bar{Y}_{...})' = Y A_1 Y'$$

$$Q_2 = t \sum_{i=1}^{g} \sum_{j=1}^{n_i} (\bar{Y}_{ij.} - \bar{Y}_{i..})(\bar{Y}_{ij.} - \bar{Y}_{i..})' = Y A_2 Y'$$

$$Q_3 = n \sum_{k=1}^{t} (\bar{Y}_{..k} - \bar{Y}_{...})(\bar{Y}_{..k} - \bar{Y}_{...})' = Y A_3 Y'$$

$$Q_4 = \sum_{i=1}^{g} n_i \sum_{k=1}^{t} (\bar{Y}_{i.k} - \bar{Y}_{i..} - \bar{Y}_{..k} + \bar{Y}_{...})(\bar{Y}_{i.k} - \bar{Y}_{i..} - \bar{Y}_{..k} + \bar{Y}_{...})' = Y A_4 Y'$$

$$Q_5 = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t} (Y_{ijk} - \bar{Y}_{ij.} - \bar{Y}_{i.k} + \bar{Y}_{i..})(Y_{ijk} - \bar{Y}_{ij.} - \bar{Y}_{i.k} + \bar{Y}_{i..})' = Y A_5 Y'$$

with appropriate choice of matrices $A_1 - A_5$. These matrices are symmetric of order

$N \times N$ and can be easily derived. For example, see Geisser and Greenhouse (1958).

It is clear from the works of Khatri (1962), Arnold (1979), Reinsel (1982), and

Mathew (1989) that the above matrix quadratic forms are independent of each

other and under the appropriate null hypothesis each has a scale multiple of a

Wishart distribution with a certain degrees of freedom.

Thomas (1983) considered a class of structures for Ω, members of which are

sufficient to keep the multivariate analysis of variance (MANOVA) obtained under

the mixed effects model invariant preserving the independence and distributions

of the SSCP matrices of Table 4.1. Pavur (1987) characterized the class of all

covariance structures for Ω under which the MANOVA remains invariant. Vaish

(1994) in a recent Ph.D. thesis pointed out using counter examples that Pavur

(1987)'s characterization may contain some matrices that are not non negative

71

definite. Further, he has provided a more elegant characterization of Ω, which is:

$$\Omega = (\mathbf{I}_t - \frac{1}{t}\mathbf{J}_t) \otimes \Sigma + \frac{1}{t}(\mathbf{1}_t \otimes \mathbf{I}_p)\mathbf{H}' + \frac{1}{t}\mathbf{H}(\mathbf{1}'_t \otimes \mathbf{I}_p) - \frac{1}{t}\mathbf{1}_t\mathbf{1}'_t \otimes \bar{\mathbf{H}}, \qquad (4.5)$$

where $\mathbf{H}$ is a $pt \times p$ arbitrary matrix of rank $(\mathbf{H}) \leq p$, and $\bar{\mathbf{H}} = \frac{1}{t}\sum_{i=1}^{t}\mathbf{H}_i$. Further, the $p \times p$ matrices $\mathbf{H}_i$'s are such that $\mathbf{H} = (\mathbf{H}'_1, ..., \mathbf{H}'_t)'$.

The structure (4.5) is slightly more general than the structure

$$\Omega = \begin{bmatrix} \Sigma_1 & \Sigma_2 & . & . & . & \Sigma_2 \\ \Sigma_2 & \Sigma_1 & . & . & . & \Sigma_2 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \Sigma_2 & \Sigma_2 & . & . & . & \Sigma_1 \end{bmatrix}$$

considered by Arnold (1979) and Mathew (1989). It may be noted that this structure is a multivariate analogue of the well known equicorrelation structure. Similarly, the covariance structure

$$\Omega = \mathbf{I}_t \otimes \Sigma_1 + \mathbf{1}_t\mathbf{1}'_t \otimes \Sigma_2$$

considered by Reinsel (1982) is also a particular case of (4.5).

Boik (1988) independently (of Pavur (1987)) showed that the structure considered by Thomas (1983) is necessary and sufficient for MANOVA to remain invariant. He also developed the likelihood ratio test for testing for this structure. Further, Boik (1988) considered the Satterthwaite type approximations (*multivariate Satterthwaite approximation*) to the distributions of various sums of squares and

72

crossproducts (SSCP) matrices of the MANOVA table, when the covariance matrix Ω is any general positive definite matrix. While approximating the distribution of a SSCP matrix by a scale multiple of Wishart distribution, Boik (1988) assumed that the expected value and the trace of the covariance matrix of the SSCP matrix are equal to those of approximating matrix.

Tan and Gupta (1983) and recently Khuri, Mathew, and Nel (1994) also have considered the problems relating to multivariate Satterthwaite approximation to a SSCP matrix. Their approximation is different from Boik (1988) in the sense that they use generalized variance (determinant of the covariance matrix) instead of trace, for their approximations.

4.2 Covariance structure $V \otimes \Sigma$

All the work done in the literature thus far, about multivariate repeated measures, has the basic assumption that $cov(Y_{ij}) = \Omega$, where Ω is a positive definite matrix. The structures similar to type H structure on Ω were found, under which the MANOVA remained invariant. As we have noted in the previous section, the multivariate Satterthwaite approximations for SSCP matrices of the MANOVA table (Table 4.1) were derived under the basic assumption that $cov(Y_{ij}) = \Omega$. In this section we start with the assumption that $cov(Y_{ij}) = V \otimes \Sigma$, where V is a $t \times t$ positive definite matrix. This structure has several advantages over the general covariance

structure. First, it is well known that (Crowder and Hand (1993), Jones (1993)) the correlation matrix of the repeated measures usually has a simple structure as opposed to a general structure. In our formulation it is easier to accommodate any structure for the correlation matrix of repeated measures (via V). Next, the number of unknown parameters of the variance covariance matrix, in our formulation, is much less, $(t(t+1)/2 + p(p+1)/2)$, as opposed to $(pt(pt+1)/2)$ in the case of general covariance structure. Furthermore, as we will see later, the multivariate Satterthwaite approximations to various SSCP matrices is much simpler in the present situation.

Note that the ith row of the $p \times nt$ matrix Y, where

$$Y = (Y_{111,}, ..., Y_{11t}, ..., Y_{1n_11}, ..., Y_{1n_1t}, ..., Y_{gn_g1}, ..., Y_{gn_gt}),$$ has the covariance matrix proportional to $\Delta = I_n \otimes V$. In fact it is $\sigma_{ii}\Delta$, where σ_{ii} is the ith diagonal element of Σ.

Suppose $V = I_t$, that is, the repeated measures are uncorrelated. Then using the two-way MANOVA model (4.4) analysis of the multivariate repeated measures data can be carried out as before. Now suppose $V = V(\rho) = (1-\rho)I_t + \rho J_t$. Using a result which is multivariate analogue of Cochran's theorem for quadratic forms it can be shown that the MANOVA in Table 1 will remain invariant for this case (see Theorem 4.6.3 of Vaish (1994)). Further, a characterization of the class of structures for V such that the MANOVA remains invariant yields type H structure (4.1) for V. Many interesting results about the structure $V \otimes \Sigma$ are summarized

74

in Vaish (1994).

4.2.1 Test for type H structure

In this section we construct the likelihood ratio test for testing

$$H_0 : cov(Y_{ij}) = H(\mathbf{a}) \otimes \Sigma \quad vs \quad H_a : cov(Y_{ij}) = \mathbf{V} \otimes \Sigma.$$

Make a transformation on the vector of measurements Y_{ij} to U_{ij}, such that $cov(U_{ij}) = \mathbf{W} \otimes \Sigma$, where $\mathbf{W} = \mathbf{CVC'}$. The choice of $\mathbf{C}$ is such that $\mathbf{C'1} = \mathbf{0}$ and $\mathbf{CC'} = \mathbf{I}_{t-1}$. One possible choice for $\mathbf{C}$ is the appropriately chosen submatrix of the Helmert's matrix. It is true that testing the above hypothesis is equivalent to testing

$$H_0 : cov(U_{ij}) = \mathbf{I} \otimes \Sigma \quad vs \quad H_a : cov(U_{ij}) = \mathbf{W} \otimes \Sigma. \qquad (4.6)$$

A likelihood ratio test for testing (4.6) is given by

$$\lambda = \frac{\mid \hat{\mathbf{W}} \mid^{np/2} \mid \hat{\Sigma} \mid^{n(t-1)/2}}{\mid \hat{\Sigma}_0 \mid^{n(t-1)/2}}, \qquad (4.7)$$

where $\hat{\mathbf{W}}$ and $\hat{\Sigma}$ are respectively the maximum likelihood estimates (MLEs) of $\mathbf{W}$ and Σ under the non-null case (H_a) and $\hat{\Sigma}_0$ is the MLE of Σ under the restriction of the null hypothesis (H_0). Then using standard results $-2ln\,\lambda$ has a χ^2 distribution with degrees of freedom equal to $t(t-1)/2$.

4.2.2 Derivation of the MLEs

We first derive the maximum likelihood estimators of $\mathbf{W}$ and Σ under (H_a). We have a random sample U_{ij}, $j = 1, ..., n_i$, $i = 1, ..., g$ from a $p(t\text{-}1)$-variate normal

75

distribution with a mean μ_i and variance covariance matrix $\mathbf{W} \otimes \mathbf{\Sigma}$. Here μ_i is a $p(t-1) \times 1$ vector given by $\mu_i' = (\mu_{i1}', ..., \mu_{i(t-1)}')$, where μ_{ik} is a $p \times 1$ vector representing the expected value of the transformed variable corresponding to the ith group and the jth time point.

The log-likelihood function is given by

$$ln\ l = -\frac{p(t-1)n}{2}ln(2\pi) - \frac{n}{2}ln \mid \mathbf{W} \otimes \mathbf{\Sigma} \mid$$
$$-\frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(U_{ij} - \mu_i)'\mathbf{W}^{-1} \otimes \mathbf{\Sigma}^{-1}(U_{ij} - \mu_i).$$

Let $\quad B = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(U_{ij} - \overline{U}_{i\cdot})(U_{ij} - \overline{U}_{i\cdot})', \text{with } \overline{U}_{i\cdot} = \sum_{j=1}^{n_i}U_{ij}/n_i.$

Then the likelihood function can be written as

$$ln\ l = -\frac{np(t-1)}{2}ln(2\pi) - \frac{n}{2}ln \mid \mathbf{W} \otimes \mathbf{\Sigma} \mid -\frac{1}{2}tr(\mathbf{W}^{-1} \otimes \mathbf{\Sigma}^{-1})B$$
$$-\frac{1}{2}\sum_{i=1}^{g}n_i(\overline{U}_{i\cdot} - \mu_i)'\mathbf{W}^{-1} \otimes \mathbf{\Sigma}^{-1}(\overline{U}_{i\cdot} - \mu_i). \quad\quad (4.8)$$

Next, we partition the $p(t-1) \times 1$ vector $(U_{ij} - \overline{U}_{i\cdot})$ into $(t-1)$ blocks of $p \times 1$ vectors such that $(U_{ij} - \overline{U}_{i\cdot}) = (U_{ij1} - \overline{U}_{i\cdot1}, ..., U_{ij(t-1)} - \overline{U}_{i\cdot(t-1)})'$. Using this partition of $(U_{ij} - \overline{U}_{i\cdot})$, we rewrite the likelihood (4.8) and maximize it simultaneously with respect to $\mathbf{\Sigma}$ and $\mathbf{W}$. We get the following likelihood equations for estimating $\mathbf{\Sigma}$ and $\mathbf{W}$.

$$\hat{\mathbf{\Sigma}} = \frac{1}{n(t-1)}\sum_{k=1}^{t-1}\sum_{l=1}^{t-1}\hat{w}_{kl}^{*}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(U_{ijk} - \overline{U}_{i\cdot k})(U_{ijl} - \overline{U}_{i\cdot l}) \quad\quad (4.9)$$

$$\hat{\mathbf{W}} = \frac{1}{np}A, \quad\quad (4.10)$$

76

where $\hat{w}_{kl}^*$ is the *(k l)-th* element of $\hat{\mathbf{W}}^{-1}$. The *(k l)-th* element of matrix A is given by $A_{kl} = \sum_{i=1}^{g} \sum_{j=1}^{n_i} (U_{ijk} - \overline{U}_{i.k})' \hat{\boldsymbol{\Sigma}}^{-1} (U_{ijl} - \overline{U}_{i.l})$.

The above equations are to be solved iteratively to get the estimates of $\boldsymbol{\Sigma}$ and $\mathbf{W}$. There is no general consensus as to when iterative methods should be stopped and the current values declared to be ML estimators. In our illustrative example, we have selected the following stopping rule. Compute two matrices: (a) matrix of difference between two successive solutions of (4.9), and (b) matrix of difference between two successive solutions of (4.10). Continue the iterations until the maximums of the absolute values of the elements of the matrices in (a) and (b) are smaller than the pre-specified quantities. A computer program using IML of SAS software is available with the author to compute these estimates.

In a manner analogous to what is described above, we can get the estimate of $\boldsymbol{\Sigma}$ under restriction of the null hypothesis (H_0). The Maximum likelihood estimate of $\boldsymbol{\Sigma}$ under the null hypothesis is given by

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{n(t-1)} \sum_{k=1}^{t-1} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (U_{ijk} - \overline{U}_{i.k})(U_{ijk} - \overline{U}_{i.k})'. \tag{4.11}$$

If H_0 is accepted, one can use the MANOVA of (4.2) to analyze the data. Otherwise multivariate Satterwaithe type approximation can be applied to find the approximation to the distributions of the SSCP matrices. This approximation will be described in the following sections.

77

4.2.3 A simulation study

In section 4.2.1 we derive a test for testing

$$H_0 : cov(Y_{ij}) = H(\mathbf{a}) \otimes \Sigma \quad vs \quad H_a : cov(Y_{ij}) = \mathbf{V} \otimes \Sigma,$$

using the likelihood ratio principle and derive the MLE's in section 4.2.2. Since the choice of the method of data analysis rests entirely on the outcome of this testing procedure, it is important that the test we are using is powerful enough. To study the power of our test given by (4.7) we conduct a small simulation study section. We know that the power function $\beta_\lambda(\theta)$ of the test statistic λ is defined as

$$\beta_\lambda(\theta) = P_\theta(Rejecting\ H_0) = P_\theta(\lambda > \lambda_c) \tag{4.12}$$

as a function of θ, where θ is the vector of all parameters of the covariance matrix $\mathbf{V} \otimes \Sigma$ and λ_c is such that under $H_0 : \theta = \theta_0$, $\beta_\lambda(\theta_0)$ is a specified value, say 0.05.

In order to calculate the power of the test, we generate random samples from multivariate normal distribution $N_{pt}(0, \mathbf{V} \otimes \Sigma)$, where $\mathbf{V}$ is the autoregressive covariance structure and Σ is a positive definite matrix. Note that autoregressive structure is not a member of class of all matrices of type $\mathbf{H}$. For $p=3$,

$$\Sigma = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 4 & 3 \\ 2 & 3 & 5 \end{bmatrix},$$

and for $t=3$ and 5 and for different choices of ρ, we generate multivariate normal samples. To represent both small samples and large samples we have selected

78

sample sizes of 12 and 30. We have selected $N=1000$ simulation runs of each case

and the power of the test was calculated as

$$P_\theta(Rejecting\ H_0) \doteq \frac{\#(-2\ ln\ \lambda > \chi_c^2)}{1000} \tag{4.13}$$

where χ_c^2 is the critical value under chi-square distribution with degrees of freedom

$t(t-1)/2$ and level of significance $\alpha=0.05$. We summarize the results of our study

for different choices of n, t and ρ, in the following table:

ρ	$n = 12$				$n = 30$			
	0.0	0.2	0.5	0.8	0.0	0.2	0.5	0.8
$t = 3$	0.04	0.075	0.248	0.691	0.03	0.107	0.620	0.990
$t = 5$	0.059	0.168	0.731	0.995	0.052	0.356	0.996	1.000

We make the following observations from the above table:

We see that power of the test is close to one as we depart from H_0. In other

words as ρ increases power of the test increases quite rapidly. Not surprisingly,

power is higher for larger values of n, ρ being the same.

When $\rho=0$, it amounts to sampling from the null distribution, and we expect

the power function to take values that meets the size condition. That is, the power

should be close to the level of significance, which is infact the case in the present

context.

79

4.2.4 Exact null distribution of Q_i

In the following, under no structures on $\mathbf{V}$, we derive the distribution of $Q_i = YA_iY'$, where A_i is an appropriately defined symmetric matrix of order $nt \times nt$. First of all, it is easy to show that $A_1 \Delta A_2 = A_1 \Delta A_5 = A_2 \Delta A_5 = A_3 \Delta A_5 = A_4 \Delta A_5 = 0$ (see Geisser and Greenhouse(1958)), where $\Delta = \mathbf{I}_n \otimes \mathbf{V}$. Hence the matrix quadratic forms Q_1 and Q_2, Q_1 and Q_5, Q_2 and Q_5, Q_3 and Q_5, and Q_4 and Q_5, are all pairwise independent. Let the rank of $(A_i) = \nu_i$ (for example, $\nu_1 = g - 1$ and so on). We observe that each row of Y is a multivariate normal vector of order $nt \times 1$ and has a covariance matrix proportional to $\Delta = \mathbf{I}_n \otimes \mathbf{V}$. Since by assumption Δ is a positive definite matrix, there exit $\Delta^{\frac{1}{2}}$ and $\Delta^{-\frac{1}{2}}$ such that $\Delta = \Delta^{\frac{1}{2}}\Delta^{\frac{1}{2}}$ and $\Delta^{-1} = \Delta^{-\frac{1}{2}}\Delta^{-\frac{1}{2}}$ and that $\Delta^{-\frac{1}{2}}\Delta^{\frac{1}{2}} = \mathbf{I}_{nt}$.

Consider $Q_i = YA_iY' = Y\Delta^{-\frac{1}{2}}\Delta^{\frac{1}{2}}A_i\Delta^{\frac{1}{2}}\Delta^{-\frac{1}{2}}Y' = ZB_iZ'$ with $Z = Y\Delta^{-\frac{1}{2}}$ and $B_i = \Delta^{\frac{1}{2}}A_i\Delta^{\frac{1}{2}}$. Now the rows of Z form a random sample of size nt from $N_p(0, \Sigma)$ distribution under certain null hypothesis. Since B_i is a symmetric matrix with rank ν_i, it can be written as $B_i = \Gamma\Lambda_i\Gamma'$, where $\Gamma\Gamma' = \Gamma'\Gamma = \mathbf{I}_{nt}$ and $\Lambda_i = Diag(\lambda_1, ..., \lambda_{\nu_i}, 0, .., 0)$, where $\lambda_1, ..., \lambda_{\nu_i}$ are the eigenvalues of B_i. Thus $Q_i = ZB_iZ' = Z\Gamma\Lambda_i\Gamma'Z' = U\Lambda_iU' = \sum_{j=1}^{\nu_i} \lambda_j U_j U_j'$, where $U_j \sim N_p(0, \Sigma)$ under H_0 and $U_1, ..., U_{\nu_i}$ are all independent.

It is well known that $U_j U_j' \sim W_p(1, \Sigma)$, which is a *pseudo-Wishart distribution*, since the degrees of freedom is less than the dimension, and has no *pdf*. Thus the distribution of Q_i is same as the distribution of linear combination of ν_i pseudo-

80

Wishart random matrices. In summary, we have the following:

$$Q_1 \sim W_p((g-1), \Sigma),$$

$$Q_2 \sim W_p((n-g), \Sigma),$$

$$Q_3 \sim \sum_{j=1}^{t-1} \lambda_j W_{pj}(1, \Sigma),$$

$$Q_4 \sim \sum_{j=1}^{t-1} \lambda_j W_{pj}(g-1, \Sigma),$$

$$Q_5 \sim \sum_{j=1}^{t-1} \lambda_j W_{pj}(n-g, \Sigma),$$

where $\lambda_1, ..., \lambda_{t-1}$ are the eigenvalues of $(\mathbf{I} - 1/t\mathbf{J})\mathbf{V}$ and $W_{pj}(\nu_j, \Sigma)$, for $j = 1, ..., t-1$ are Wishart random matrices with ν_j degrees of freedom, and mutually independent.

Suppose we want to test $H_{01} : \alpha_1 = ... = \alpha_g$. Then the Wilks' Λ for testing H_{01} is

$$\Lambda_1 = \frac{|Q_2|}{|Q_1 + Q_2|}.$$

The test is based on the usual asymptotic distribution (in some situations exact (Rao, 1973, p. 555)) of Λ_1. For example,

$$-[(n-g) - \frac{(p-g)}{2}] \, ln\, \Lambda_1 \sim \chi^2_{p(g-1)} \text{approximately}.$$

The approximate distributions of the test statistics for testing $H_{02} : \beta_1 = ... = \beta_t$ and $H_{03} : (\alpha\beta)'_{ik}s$ are all equal, will be derived in the next section.

81

4.2.5 Approximate null distributions of Q_3, Q_4, and Q_5

In this section, we approximate the distribution of each of SSCP matrices, Q_3, Q_4, and Q_5, to a scale multiple of a Wishart matrix, $gW(h, \Sigma)$, for some constants g and h. As in the univariate case the approximation is derived by equating the first two central moments. For that we first find the first two central moments of Q_i. It is well known (for example, see Eaton, 1983, p. 305) that for any $S \sim W_p(\nu, \Sigma)$, $E(S) = \nu\Sigma$ and $D(S) = 2\nu\Sigma \otimes \Sigma$. Here $D(S)$ denotes the variance covariance matrix of all the random quantities in S. Using these formulae we have the following:

$E(Q_3) = [tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})]\Sigma$ under $H_{02} : \beta_1 = \ldots = \beta_t$

$E(Q_4) = [(g-1)tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})]\Sigma$ under $H_{03} : (\alpha\beta)_{ik} = \gamma$ for all i, k.

$E(Q_5) = [(n-g)tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})]\Sigma$

also

$D(Q_3) = [2tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2]\Sigma \otimes \Sigma$ under H_{02}

$D(Q_4) = [2(g-1)tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2]\Sigma \otimes \Sigma$ under H_{03}

$D(Q_5) = [2(n-g)tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2]\Sigma \otimes \Sigma.$

Suppose we want to approximate the distribution of Q_3 by a random matrix having the distribution $g_1 W_p(h_1, \Sigma)$ so that the first two central moments of Q_3 and $g_1 W_p(h_1, \Sigma)$ are the same. Then,

$$tr[(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})]\Sigma = g_1 h_1 \Sigma \qquad (4.14)$$

82

$$2[tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2]\Sigma \otimes \Sigma \;=\; 2g_1^2 h_1 \Sigma \otimes \Sigma. \tag{4.15}$$

From (4.14) and (4.15)

$$
\begin{aligned}
g_1 &= \frac{tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2}{tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})} \\
h_1 &= \frac{[tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})]^2}{tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2}.
\end{aligned}
$$

Next, to approximate the distribution of Q_4 by a random matrix having the distribution $g_2 W_p(h_2, \Sigma)$ so that the first two central moments of Q_4 and $g_2 W_p(h_2, \Sigma)$ are the same we have,

$$(g-1)tr[(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})]\Sigma \;=\; g_2 h_2 \Sigma \tag{4.16}$$

$$2(g-1)[tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2]\Sigma \otimes \Sigma \;=\; 2g_2^2 h_2 \Sigma \otimes \Sigma. \tag{4.17}$$

From (4.16) and (4.17) we have,

$$
\begin{aligned}
g_2 &= \frac{tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2}{tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})} = g_1 \\
h_2 &= (g-1)\frac{[tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})]^2}{tr(\mathbf{V} - 1/t\mathbf{J}\mathbf{V})^2} = (g-1)h_1.
\end{aligned}
$$

Therefore, $Q_4 \sim g_1 W_p((g-1)h_1, \Sigma)$ approximately.

Similarly, it can shown that $Q_5 \sim g_1 W_p((n-g)h_1, \Sigma)$ approximately.

Now for testing $H_{02} : \beta_1 = \dots = \beta_t$ one can use the Wilks' Λ, which is, $\Lambda_2 = \dfrac{|Q_5|}{|Q_3 + Q_5|}$ and the fact that

$$- \left[(n - g)h_2 - \left(\frac{p - 1 - \frac{1}{2}}{2} \right) \right] \ln \Lambda_2 \sim \chi^2_{f_2}, \text{ approximately}. \qquad (4.18)$$

Similarly for testing H_{03} : $(\alpha_i)'_{ij}s$ are all equal, the Wilks' Λ is $\Lambda_3 = \frac{Q_3}{Q_4 + Q_3}$

and the distribution of the test statistic is

$$- \left[(n - g)h_1 - \left(\frac{p - 1 - (g - 1)h_1}{2} \right) \right] \ln \Lambda_3 \sim \chi^2_{p(g-1)h}, \text{approximately}. \qquad (4.19)$$

4.2.6 Estimation of the degrees of freedom

Since in practice V_{jk} is unknown the degrees of freedoms in the chi square approximations of (4.18) and (4.19) are unknown. One needs to estimate these so that the distributions in (4.18) and (4.19) can be utilized in applications. For estimating these degrees of freedoms, which are functions of $(V - 1/tJV)$, we simply need an estimate of V. One can use the maximum likelihood estimate of V that is obtained by simultaneously solving the following equations:

$$\hat{\Sigma} = \frac{1}{nt} \sum_{k=1}^{t} \sum_{l=1}^{t} \hat{v}_{kl}^{*} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (Y_{ijk} - \bar{Y}_{i.k})(Y_{ijl} - \bar{Y}_{i.l}) \qquad (4.20)$$

$$\hat{V} = \frac{1}{np} A, \qquad (4.21)$$

where $\hat{v}_{kl}^{*}$ is the (k l)-th element of $\hat{V}^{-1}$ and (k l)-th element of matrix A is given by $A_{kl} = \sum_{i=1}^{g} \sum_{j=1}^{n_i} (Y_{ijk} - \bar{Y}_{i.k})'\hat{\Sigma}^{-1}(Y_{ijl} - \bar{Y}_{i.l})$.

In the next section we consider an example and illustrate the estimation of the degrees of freedom using the maximum likelihood estimate of V.

84

4.3 Example

To illustrate our methods, we use data from Table 7.2 in Timm (1980), which was also analyzed by Thomas (1983) and Boik (1988) . These data were obtained by Timm from T. Zullo of the School of Dental Medicine at the University of Pittsburgh. The study concerns with the relative effectiveness of two orthopedic adjustments of the mandible. Nine subjects were assigned to each of two orthopedic treatments ($g = 2, n_1 = 9, n_2 = 9$), called activator treatments. The measurements were made on three characteristics ($p=3$) to assess the changes in the vertical position of the mandible at three time points ($t=3$) of activator treatment. Thus the data matrix $Y = (Y_{ij})$ is an 18×9 matrix. The data is presented in the following table:

The choice of method to analyze our data, rests on the test of hypothesis about the covariance structure described in section (3). Hence to test the null hypothesis (4.6), first we transform Y_{ij} to U_{ij} using the transformation $U_{ij} = (C \otimes I)Y_{ij}$, with

$$C = \begin{bmatrix} 0.7071 & -0.7071 & 0.0000 \\ 0.4082 & 0.4082 & -0.8165 \end{bmatrix}.$$

Then we have $cov(U_{ij}) = W \otimes \Sigma$.

The maximum likelihood estimates of Σ and W, simultaneously solving (4.9) and (4.10), are given by

85

Table 4.3: Dental data

	Subj	SOr-Me (mm)			ANS-ME (mm)			Pal-MP angle (degrees)		
	1	117.0	117.5	118.5	59.0	59.0	60.0	10.5	16.5	16.5
	2	109.0	110.5	111.0	60.0	61.5	61.5	30.5	30.5	30.5
	3	117.0	120.0	120.5	60.0	61.5	62.0	23.5	23.5	23.5
	4	122.0	126.0	127.0	67.5	70.5	71.5	33.0	32.0	32.5
T_1	5	116.0	118.5	119.5	61.5	62.5	63.5	24.5	24.5	24.5
	6	123.0	126.0	127.0	65.5	61.5	67.5	22.0	22.0	22.0
	7	130.5	132.0	134.5	68.5	69.5	71.0	33.0	32.5	32.0
	8	126.5	128.5	130.5	69.0	71.0	73.0	20.0	20.0	20.0
	9	113.0	116.5	118.0	58.0	59.0	60.5	25.0	25.0	24.5
	1	128.0	129.0	131.5	67.0	67.5	69.0	24.0	24.0	24.0
	2	116.5	120.0	121.5	63.5	65.0	66.0	28.5	29.5	29.5
	3	121.5	125.5	127.0	64.5	67.5	69.0	26.5	27.0	27.0
	4	109.5	112.0	114.0	54.0	55.5	57.0	18.0	18.5	19.0
	5	133.0	136.0	137.5	72.0	73.5	75.7	34.5	34.5	34.5
T_2	6	120.0	124.5	126.0	62.5	65.0	66.0	26.0	26.0	26.0
	7	129.5	133.5	134.5	65.0	68.0	69.0	18.5	18.5	18.5
	8	122.0	124.0	125.5	64.5	65.5	66.0	18.5	18.5	18.5
	9	125.0	127.0	128.0	65.5 86	66.5	67.0	21.5	21.5	21.6

$$\hat{\Sigma} = \begin{bmatrix} 0.3861535 & 0.2324518 & -0.16931 \\ 0.2324518 & 1.0859517 & -0.097388 \\ -0.16931 & -0.097388 & 0.4350231 \end{bmatrix}.$$

and

$$\hat{W} = \begin{bmatrix} 1.5931756 & 0.5033948 \\ 0.5033948 & 0.5951812 \end{bmatrix}.$$

Estimate of Σ under H_0 using (4.11) is

$$\hat{\Sigma}_0 = \begin{bmatrix} 0.4192387 & 0.2425412 & -0.240329 \\ 0.2425412 & 0.7572016 & -0.145422 \\ -0.240329 & -0.145422 & 0.6873045 \end{bmatrix}.$$

The value of the statistic given in (4.7) is λ=0.0000139. Then the test statistic value, $-2ln\,\lambda$, is equal to 22.37366. Comparing this with $\chi_3^2(0.05) = 7.815$, we clearly reject H_0 given in (4.6). Therefore, we use multivariate Satterwaithe type approximation described in sections (3.3-3.4), to find approximation to the distributions of the SSCP matrices. In order to use these methods we need to estimate g_1 and h_1 which in turn requires the estimate of V. We simultaneously solve (4.9)

87

Hypothesis	Wilks' λ	D.O.F	Approximate χ^2 value	p-value
Group	0.9083	3	1.4903	0.6845
Time	0.0619	3.7891	61.3910	0.0000
Time*Group	0.8145	3.7891	4.5279	0.3108

Table 4.4: Approximate MANOVA

and (4.10) to get the maximum likelihood estimate of V and it is found to be

$$\hat{V} = \begin{bmatrix} 2.4224556 & 2.2517394 & 2.2765762 \\ 2.2517394 & 2.2246306 & 2.2018579 \\ 2.2765762 & 2.2018579 & 2.2294622 \end{bmatrix}.$$

Using this we get $g_1 = 0.0636441$ and $h_1 = 1.5332644$.

We summarize our results in Table 4.3.

For computing the above p-values one can use any easily available softwares. For example, we have used PROBF function of SAS Software. Note from Table 4.3 that only the effect of time factor is significant.

88

Chapter 5

Analysis of Unbalanced

Multivariate Repeated

Measurements

In chapter 4 we have considered a set of balanced (same number of measurements on each individual) multivariate repeated measurements. These data can be represented by Y_{ijk}, $k = 1, ..., t$; $j = 1, ..., n_i$; $i = 1, ..., g$, which is a $p \times 1$ vector of measurements on the jth individual in the ith group at the kth occasion. Let $Y'_{ij} = (Y'_{ij1}, ..., Y'_{ijt})$. Then Y_{ij} is $pt \times 1$ random observational vector corresponding to the jth individual in the ith group. We have assumed that $cov(Y_{ij}) = V \otimes \Sigma$, where V is a $t \times t$ positive definite matrix. It was seen that if $V = V(\rho) = (1 - \rho)I_t + \rho J_t$, then the analysis of multivariate repeated measurements can be carried out using

89

the two-way MANOVA model. Further, it was noted that a characterization of the class of structures for V such that the MANOVA remains invariant yields type H structure. On the other hand, if V is completely arbitrary, we have suggested Satterthwaite type approximation to the distribution of multivariate quadratic forms of SS&CP matrices. Multivariate tests using this approximation were developed for testing the usual hypothesis.

5.1 Autoregressive structure for V

We now assume autoregressive structure for V, that is, $V = (\rho^{|i-j|})$. This structure is neither type H, nor completely arbitrary. With this structure on V, we derive the approximate distributions of SS&CP matrices. In chapter 4, assuming no structure on V we have shown the following:

$$Q_1 \sim W_p((g-1), \Sigma),$$

$$Q_2 \sim W_p((n-g), \Sigma),$$

$$Q_3 \sim g_1 W_p(h_1, \Sigma) \quad \text{approximately},$$

$$Q_4 \sim g_1 W_p((g-1)h_1, \Sigma) \quad \text{approximately},$$

$$Q_5 \sim g_1 W_p((n-g)h_1, \Sigma) \quad \text{approximately},$$

with

$$g_1 = \frac{tr(V - 1/tJV)^2}{tr(V - 1/tJV)}$$

90

$$h_1 = \frac{[tr(\mathbf{V} - 1/t\mathbf{JV})]^2}{tr(\mathbf{V} - 1/t\mathbf{JV})^2}.$$

Now, let $\mathbf{V} = (\rho^{|i-j|})$. Next we derive expressions for g_1 and h_1 in terms of ρ.

Notice that

$$tr(\mathbf{V} - 1/t\mathbf{JV}) = tr(\mathbf{V}) - 1/t\ tr(\mathbf{JV}), \tag{5.1}$$

$$tr(\mathbf{V} - 1/t\mathbf{JV})^2 = tr(\mathbf{VV}) - 2tr(\mathbf{VJV}) + 1/t^2\ tr(\mathbf{JVJV}). \tag{5.2}$$

Expressions (5.1) and (5.2) and in turn g_1 and h_1 can be evaluated using the following:

$$tr(\mathbf{V}) = t,$$

$$tr(\mathbf{JV}) = \frac{t(1-\rho^2) - 2\rho(1-\rho^t)}{(1-\rho)^2},$$

$$tr(\mathbf{VV}) = \frac{1}{(1-\rho^2)^2}\left[t(1-\rho^4) - 2\rho^2(1-\rho^{2t})\right],$$

$$tr(\mathbf{VJV}) = \frac{1}{(1-\rho)^2}\left[t(1+\rho)^2 + 2t\rho^{t+1} + 2\rho^2\frac{1-\rho^{2t}}{1-\rho^2} - 4\rho(1+\rho)\frac{1-\rho^t}{1-\rho}\right]$$

$$tr(\mathbf{JVJV}) = \left[\frac{t(1-\rho^2) - 2\rho(1-\rho^t)}{(1-\rho)^2}\right]^2.$$

5.1.1 Distributions of Q_3, Q_4 and Q_5

In order to find the distributions of the matrix quadratic forms Q_3, Q_4 and Q_5 we need to estimate g_1 and h_1. In the above discussion we have given expressions for

91

g_1 and h_1 as functions of ρ. However since ρ is unknown, we need to estimate the same. In the following we present maximum likelihood estimate of ρ.

Consider a random sample Y_{ij}, $j = 1, ..., n_i$; $i = 1, ..., g$ from a pt-variate normal distribution with mean μ_i (can be a linear function in a smaller dimension space) and variance covariance matrix $\mathbf{V} \otimes \mathbf{\Sigma}$. Here $\mathbf{V}$ is an autoregressive structure matrix and $\mathbf{\Sigma}$ is any positive definite matrix. Then the log-likelihood function can be written as,

$$
\begin{aligned}
ln\ l\ =\ & \text{constant} - \frac{n}{2}ln \mid \mathbf{V} \otimes \mathbf{\Sigma} \mid \\
& - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (Y_{ij} - \mu_i)' \mathbf{V}^{-1} \otimes \mathbf{\Sigma}^{-1} (Y_{ij} - \mu_i).
\end{aligned}
$$

$$
\text{Let}\quad B\ =\ \sum_{i=1}^{g} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_{i.})(Y_{ij} - \overline{Y}_{i.})', \text{with } \overline{Y}_{i.} = \sum_{j=1}^{n_i} Y_{ij}/n_i.
$$

Then

$$
\begin{aligned}
ln\ l\ =\ & \text{constant} - \frac{n}{2}ln \mid \mathbf{V} \otimes \mathbf{\Sigma} \mid - \frac{1}{2}tr(\mathbf{V}^{-1} \otimes \mathbf{\Sigma}^{-1})B \\
& - \frac{1}{2} \sum_{i=1}^{g} n_i (\overline{Y}_{i.} - \mu_i)' \mathbf{V}^{-1} \otimes \mathbf{\Sigma}^{-1} (\overline{Y}_{i.} - \mu_i).
\end{aligned}
\tag{5.3}
$$

Clearly the maximum likelihood estimate of $\mu_i = \overline{Y}_{i.}$. By plugging this value for μ_i in (5.3) we get

$$
ln\ l\ =\ -\frac{npt}{2}ln(2\pi) - \frac{n}{2}ln \mid \mathbf{V} \otimes \mathbf{\Sigma} \mid - \frac{1}{2}tr(\mathbf{V}^{-1} \otimes \mathbf{\Sigma}^{-1})\mathbf{B}.
\tag{5.4}
$$

Recall that $|\mathbf{V}| = (1 - \rho^2)^{t-1}$ and $\mathbf{V}^{-1} = \frac{1}{1 - \rho^2}[I + \rho^2 C_1 - \rho C_2]$. Differentiating (5.4) with respect to ρ and equating to zero we have,

$$
-\rho^3 2np(t - 1) + \rho^2 A_3 + \rho(2np(t - 1) - 2A_1 - 2A_2) + A_3 = 0,
\tag{5.5}
$$

92

where

$$A_1 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_{i.})'(\mathbf{I} \otimes \Sigma^{-1})(Y_{ij} - \overline{Y}_{i.}),$$

$$A_2 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_{i.})'(C_1 \otimes \Sigma^{-1})(Y_{ij} - \overline{Y}_{i.}),$$

$$A_3 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_{i.})'(C_2 \otimes \Sigma^{-1})(Y_{ij} - \overline{Y}_{i.}).$$

The cubic equation in ρ given by (5.5) can easily be shown to have a unique root in the interval (-1,1). Hence it is easy to compute the MLE $\hat{\rho}$, of ρ. Let $\hat{\mathbf{V}} = \mathbf{V}(\hat{\rho})$. Next, to find the maximum likelihood estimator of Σ, we need to maximize (5.4) with respect to Σ. In order to do that, partition the $pt \times 1$ vector $(Y_{ij} - \overline{Y}_{i.})$ into t blocks of $p \times 1$ vectors such that $(Y_{ij} - \overline{Y}_{i.}) = (Y_{ij1} - \overline{Y}_{i.1}, ..., Y_{ijt} - \overline{Y}_{i.t})'$. Using this partition of $(Y_{ij} - \overline{Y}_{i.})$, we rewrite the likelihood (5.4) and maximize with respect to Σ. This process yields,

$$\hat{\Sigma} = \frac{1}{nt}\sum_{k=1}^{t}\sum_{l=1}^{t}\hat{v}_{kl}^{*}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(Y_{ijk} - \overline{Y}_{i.k})(Y_{ijl} - \overline{Y}_{i.l}), \tag{5.6}$$

where $\hat{v}_{kl}^{*}$ is the $(kl)^{th}$ element of $\hat{\mathbf{V}}^{-1}$.

By solving equations (5.5) and (5.6) iteratively we find $\hat{\mathbf{V}}$ and $\hat{\Sigma}$, and then the estimates of g_1 and h_1. Thus we have approximate distributions of Q_3, Q_4 and Q_5.

5.2 Unbalanced multivariate repeated measures

So far, we have considered balanced case of multivariate repeated measures. That is, we have considered a set of t repeated measurements on p variables (or char-

93

acteristics) on each of the n individuals. Thus data on each individual is a $p \times t$ matrix. However, in practice we do come across data where we may not have measurements on the individuals at all the t time points. Stated in other words, we have t_{ij} repeated measurements on p variables on the jth individual from the ith group. Now data on each individual is a $p \times t_{ij}$ matrix. Let Y_{ijk} be a $p \times 1$ vector of observations on the jth individual from the ith group at the kth occasion. Let $Y_{ij}' = (Y_{ij1}', ..., Y_{ijt_{ij}}')$. Then Y_{ij} is a $pt_{ij} \times 1$ random observational vector corresponding to the jth individual from the ith group. Let us assume that $cov(Y_{ij}) = \mathbf{V}_{ij} \otimes \mathbf{\Sigma}$, where $\mathbf{V}_{ij}$ is a $t_{ij} \times t_{ij}$ positive definite matrix and $\mathbf{\Sigma}$ is a $p \times p$ positive definite matrix. It appears that the covariance matrix $\mathbf{V}_{ij}$ is different for every i, j. However this is not true. If $\mathbf{V}$ is a positive definite matrix of order $t \times t$, where t is the maximum of t_{ij}, then $\mathbf{V}_{ij}$ is the $t_{ij} \times t_{ij}$ submatrix of $\mathbf{V}$. In applications, we assume $\mathbf{V}_{ij}$ for every i and j, to have some structure depending on the same set of parameters, say $\theta_1, ..., \theta_k$. In other words, $\mathbf{V}_{ij}$ depends on i and j only through its dimension. Analysis of this type of data is not very easy. In the following section, we describe a general linear model approach to analyze unbalanced data, assuming certain structures for $\mathbf{V}_{ij}$.

94

5.3 A linear model approach

Consider the situation where we have a set of unbalanced multivariate repeated measurements. Let $\mathbf{Y}_{ij}$ be the $t_{ij} \times p$ matrix of observations on the jth individual from the ith group. Let us assume the following model for $\mathbf{Y}_{ij}$:

$$\mathbf{Y}_{ij} = X_{ij} B_i + E_{ij}, \quad i = 1, ..., g; \; j = 1, ..., n_i, \qquad (5.7)$$

where X_{ij} is $t_{ij} \times m$ design matrix and B_i is $m \times p$ matrix of unknown parameters, with $cov(vec\, E'_{ij}) = \mathbf{V}_{ij} \otimes \mathbf{\Sigma}$. Here $\mathbf{V}_{ij}$ is a positive definite matrix depending on only a few parameters, say $\theta_1, ..., \theta_k$. If $t_{ij} = t$ for all i and j (balanced data) then X_{ij} is a $t \times m$ matrix, say X_i. Further if X_i an identity matrix, then the model would contain a matrix of unstructured means, different for different groups.

In the unbalanced case, the interpretation of X_{ij} can be made as follows. Suppose G_{ij} is a $t_{ij} \times t$ matrix of 0's and 1's such that it has 1 at the $(k, l_k)^{th}$ position and 0 everywhere else, assuming that observations are available at the time points $l_1, ..., l_{t_{ij}}$. Then we take $X_{ij} = G_{ij} X_i$.

Let us stack all the $\sum_j^{n_i} t_{ij}$ observations from the ith group one below the other

95

and write

$$
Y_i = \begin{bmatrix} Y_{i1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ Y_{in_i} \end{bmatrix}, X_i = \begin{bmatrix} X_{i1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ X_{in_i} \end{bmatrix} \text{ and } E_i = \begin{bmatrix} E_{i1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ E_{in_i} \end{bmatrix}.
$$

Then we have the multivariate linear model

$$
\begin{bmatrix} Y_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ Y_g \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \ldots & 0 \\ 0 & X_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & X_g \end{bmatrix} \begin{bmatrix} B_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ B_g \end{bmatrix} + \begin{bmatrix} E_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ E_g. \end{bmatrix}
$$

By writing $n = \sum\limits_{i}^{g}\sum\limits_{j}^{n_i} t_{ij}$ and with other obvious notations we have the following

multivariate regression model

$$
\mathbf{Y}_{n\times p} = \mathbf{X}_{n\times mg}\mathbf{B}_{mg\times p} + \mathbf{E}_{n\times p}, \tag{5.8}
$$

with $cov(vec(\mathbf{Y}')) = \mathbf{W} = diag(\mathbf{V}_{11}, ..., \mathbf{V}_{1n_1}, ..., \mathbf{V}_{gn_g}) \otimes \mathbf{\Sigma}$.

From the usual multivariate regression theory it can be shown that

$$
\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{-1}\mathbf{Y}
$$

$$
= [\sum_{i=1}^{g}\sum_{j=1}^{n_i} X_{ij}'V_{ij}^{-1}X_{ij}]^{-1}[\sum_{i=1}^{g}\sum_{j=1}^{n_i} X_{ij}'V_{ij}^{-1}Y_{ij}]
$$

$$
\text{and} \quad \hat{\mathbf{\Sigma}} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'\mathbf{W}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})/(n - m).
$$

96

Using this $\hat{\mathbf{B}}$ (say $\theta_1, ..., \theta_k$ are known) any one of the MANOVA tests, for example Wilks' Λ for testing $H_0 : LBM = 0$ can be used. But $\mathbf{V}_{ij}$ and in turn $\mathbf{W}$ is usually unknown. Using some consistent estimators of $\theta_1, ..., \theta_k$, $\mathbf{V}_{ij}$ is estimated and the estimate $\hat{\mathbf{W}}$ is used in place of $\mathbf{W}$ in the formula for $\hat{\mathbf{B}}$. The same MANOVA tests can be used, even in this case for testing $H_0 : LBM = 0$. In the following sections we assume equicorrelation and autoregressive structures for $\mathbf{V}_{ij}$ and discuss estimation of parameters of variance covariance matrix. These two structures are commonly used in practice. Note that in both of these cases, $\mathbf{V}_{ij}$ is a function of only one unknown parameter ρ. We derive the maximun likelihood estimate of this unknown parameter and utilize it to perform the MANOVA.

5.3.1 Equicorrelation structure for $\mathbf{V}_{ij}$

Suppose we have an initial estimate of $\mathbf{B}$. For example, an initial estimate of $\mathbf{B}$ may be obtained by performing regression under $\rho = 0$. Then let $\hat{E}_{ij} = \mathbf{Y}_{ij} - X_{ij}\hat{B}_i$. We can assume that $\hat{e}'_{ij} = vec(E'_{ij}),\quad i = 1, ..., g;\quad j = 1, ..., n_i$, form a random sample from pt_{ij}-variate normal distribution with zero mean vector and covariance matrix $\mathbf{V}_{ij} \otimes \boldsymbol{\Sigma}$. Then the log-likelihood function can be written as

$$ln\, l = \text{constant} - \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i} ln\mid \mathbf{V}_{ij} \otimes \boldsymbol{\Sigma} \mid$$
$$- \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i} \hat{e}'_{ij}\mathbf{V}_{ij}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\hat{e}_{ij}. \tag{5.9}$$

97

Next, partition the $pt_{ij} \times 1$ vector $\hat{\epsilon}_{ij}$ into t_{ij} blocks of $p \times 1$ vectors such that

$\hat{\epsilon}_{ij} = (\hat{\epsilon}_{ij1}, ..., \hat{\epsilon}_{ijt_{ij}})'$. Using this partition rewrite the likelihood (5.9) as

$$
\begin{aligned}
ln\ l\ =\ & \text{constant} - \frac{p}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} ln|\mathbf{V}_{ij}| - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} t_{ij} ln|\mathbf{\Sigma}| \\
& - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \sum_{l=1}^{t_{ij}} v_{kl}^* \hat{\epsilon}_{ijk}' \mathbf{\Sigma}^{-1} \hat{\epsilon}_{ijl}.
\end{aligned}
\tag{5.10}
$$

Differentiating (5.10) with respect to $\mathbf{\Sigma}$ and equating to zero we get

$$
\hat{\mathbf{\Sigma}} = \frac{\sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \sum_{l=1}^{t_{ij}} v_{kl}^* (\hat{\epsilon}_{ijk})(\hat{\epsilon}_{ijl})'}{\sum_{i=1}^{g} \sum_{j=1}^{n_i} t_{ij}}.
\tag{5.11}
$$

Recall that $|\mathbf{V}_{ij}| = (1-\rho)^{t_{ij}-1}(1 + (t_{ij}-1)\rho)$ and $v_{kk}^* = \dfrac{1 + (t_{ij}-2)\rho}{[1 + (t_{ij}-1)\rho](1-\rho)}$

and $v_{kl}^* = \dfrac{\rho}{[1 + (t_{ij}-1)\rho](1-\rho)}$. Now using this we can rewrite (5.10) as

$$
\begin{aligned}
ln\ l\ =\ & \text{constant} - \frac{p}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (t_{ij}-1) ln(1-\rho) - \frac{p}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} ln[1 + (t_{ij}-1)\rho] \\
& - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} t_{ij} ln|\mathbf{\Sigma}| - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \frac{1 + (t_{ij}-2)\rho}{[1 + (t_{ij}-1)\rho](1-\rho)} (\hat{\epsilon}_{ijk})' \mathbf{\Sigma}^{-1} (\hat{\epsilon}_{ijk}) \\
& - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \sum_{l=1}^{t_{ij}} - \frac{\rho}{[1 + (t_{ij}-1)\rho](1-\rho)} (\hat{\epsilon}_{ijk})' \mathbf{\Sigma}^{-1} (\hat{\epsilon}_{ijl})
\end{aligned}
\tag{5.12}
$$

Differentiating (5.12) with respect to ρ and equating to zero yields,

$$
\begin{aligned}
& \frac{p}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \frac{(t_{ij}-1)t_{ij}\rho}{[1 + (t_{ij}-1)\rho](1-\rho)} \\
& - \frac{\rho}{2(1-\rho)^2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \sum_{l=1}^{t_{ij}} \frac{(t_{ij}-1)[2 + (t_{ij}-2)\rho]}{[1 + (t_{ij}-1)\rho]} (\hat{\epsilon}_{ijk})' \hat{\mathbf{\Sigma}}^{-1} (\hat{\epsilon}_{ijl}) \\
& + \frac{1}{2(1-\rho)^2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \frac{1 + (t_{ij}-1)\rho^2}{[1 + (t_{ij}-1)\rho]^2} (\hat{\epsilon}_{ijk})' \hat{\mathbf{\Sigma}}^{-1} (\hat{\epsilon}_{ijk}) = 0
\end{aligned}
\tag{5.13}
$$

98

Solving (5.11) and (5.13) simultaneously, we get the estimates of Σ and ρ. Now, with this new estimate of ρ find a revised estimate of **B**. This two-stage estimation procedure of finding ρ and **B**, is continued until the estimates are stabilized.

5.3.2 Autoregressive structure for $\mathbf{V}_{ij}$

In this case, unlike in the equicorrelation case, the basic structure of $\mathbf{V}_{ij}$ is distorted in the presence of missing observations. We may now consider two types of missing data. If the missing observations are only at the end (monotone data) then the basic form of $\mathbf{V}_{ij}$ remains the same. That is $V_{ij} = (\rho^{|i-j|})$ where $i, j = 1, ..., t_{ij}$. However in the other forms of missing values, where the observations may be missing for any occassion the autoregressive structure leads to what is called a Markov structure. In the following we describe the method of estimating ρ in both situations.

Monotone data: As noted before, monotone data refers to the unbalanced case where the data are missing from the last few occassions. As before, let us suppose that an initial estimate of **B** is available. Then the form of the MLE of Σ is same as that in (5.11). But now, v_{kl}^* is the $(kl)^{th}$ element of $\mathbf{V}_{ij}^{-1}$ given by

$$\mathbf{V}_{ij}^{-1} = \frac{1}{(1-\rho^2)}[I_{t_{ij}} + \rho^2 C_{1ij} - \rho C_{2ij}]$$

where C_{1ij} and C_{2ij} are $t_{ij} \times t_{ij}$ matrices with similar structures as described in the Appendix of Chapter 2. Next, using, $|\mathbf{V}_{ij}| = (1 - \rho^2)^{t_{ij}-1}$ and $\mathbf{V}_{ij}^{-1}$ rewrite the

99

log-likelihood function given by (5.10) as

$$
= \text{constant} - \frac{p}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (t_{ij} - 1) ln(1 - \rho^2) - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} t_{ij} ln|\Sigma|
$$

$$
- \frac{1}{2(1 - \rho^2)} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \hat{\epsilon}'_{ij} [(\mathbf{I}_{t_{ij}} \otimes \Sigma^{-1} + \rho^2 C_{1ij} \otimes \Sigma^{-1} - \rho C_{2ij} \otimes \Sigma^{-1}] \hat{\epsilon}_{ij} \quad (5.14)
$$

Differentiating (5.14) with respect to ρ and equating to zero we get,

$$
-2\rho^3 pn + \rho^2 A_3 + 2\rho(pn - A_1 - A_2) + A_3 = 0, \quad (5.15)
$$

where

$$
n = \sum_{i=1}^{g} \sum_{j=1}^{n_i} (t_{ij} - 1)
$$

$$
A_1 = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \hat{\epsilon}'_{ij} (\mathbf{I}_{ij} \otimes \Sigma^{-1}) \hat{\epsilon}_{ij}
$$

$$
A_2 = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \hat{\epsilon}'_{ij} (C_{1ij} \otimes \Sigma^{-1}) \hat{\epsilon}_{ij}
$$

$$
A_3 = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \hat{\epsilon}'_{ij} (C_{2ij} \otimes \Sigma^{-1}) \hat{\epsilon}_{ij}.
$$

The cubic equation (5.15) is similar to the one studied by Prabhala (1995) in the context of growth curve models. It has been proved that the equation has a unique root in the interval (-1,1). Solving (5.15) and (5.11) with appropriate v^*_{kl} simlutaneously we get the estimates of ρ and Σ.

General unbalanced data: We now consider a model that can be used for analyzing data that are missing from any of the intermediate occassions as well. Let $\mathbf{Y}_{ij}$ be the $t_{ij} \times p$ matrix of observations on the jth individual from the ith group

100

such that $\mathbf{Y}_{ij} = G_{ij}\check{\mathbf{Y}}_{ij}$. Here $\check{\mathbf{Y}}_{ij}$ is a $t \times p$ matrix of measurements, $t = max(t_{ij})$ and G_{ij} is a matrix of order $t_{ij} \times t$ as defined in the beginning of section 5.3.

Let us assume the following model for $\mathbf{Y}_{ij}$

$$\mathbf{Y}_{ij} = X_{ij}B_i + E_{ij}, \quad i = 1,...,g; \ j = 1,...,n_i, \tag{5.16}$$

where X_{ij} is $t_{ij} \times m$ design matrix and B_i is $m \times p$ matrix of unknown parameters. Then $cov(vec(E'_{ij})) = G_{ij}\mathbf{V}G'_{ij} \otimes \mathbf{\Sigma}$, where $\mathbf{V}$ is a $t \times t$ matrix having an autoregressive structure. Then we can verify that $G_{ij}\mathbf{V}G'_{ij}$ has a Markov structure, that is $G_{ij}\mathbf{V}G'_{ij} = \mathbf{V}_{ij} = (\rho^{|t_{ijk}-t_{ijk'}|})$, $k, k' = 1,...,t_{ij}$, where t_{ijk}'s are the consecutive time points where the observations are made on the jth individual from the ith group. Then $\mathbf{V}_{ij}$ is of the form

$$V_i(\rho) = \begin{pmatrix} 1 & \rho^{t_{ij2}-t_{ij1}} & \rho^{t_{ij3}-t_{ij1}} & . & . & \rho^{t_{ijt_{ij}}-t_{ij1}} \\ \rho^{t_{ij2}-t_{ij1}} & 1 & \rho^{t_{ij3}-t_{ij2}} & . & . & \rho^{t_{ijt_{ij}}-t_{ij2}} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{t_{ijt_{ij}}-t_{ij1}} & \rho^{t_{ijt_{ij}}-t_{ij2}} & . & & . & . & 1 \end{pmatrix}$$

The determinant and the inverse of this matrix are given below. Let $d_{ijk} = t_{ij(k+1)} - t_{ijk}$, and $f_{ijk} = \frac{1}{1-\rho^{2d_{ijk}}}$, $k = 1,...,t_{ij}-1$. Then

$$|\mathbf{V}_{ij}| = (f_{ij1}f_{ij2}...f_{ijt_{ij}-1})^{-1}$$

and $\mathbf{V}_{ij}^{-1} =$

101

$$
\begin{pmatrix}
f_{ij1} & -f_{ij1}\rho^{d_{ij1}} & 0 & \cdot \; \cdot & 0 \\[1em]
-f_{ij1}\rho^{d_{ij1}} & f_{ij1}+f_{ij2}-1 & -f_{ij2}\rho^{d_{ij2}} & \cdot \; \cdot & 0 \\[1em]
0 & -f_{ij2}\rho^{d_{ij2}} & f_{ij2}+f_{ij3}-1 & \cdot \; \cdot & 0 \\[1em]
\cdot & \cdot & \cdot & \cdot \; \cdot & \cdot \\[1em]
\cdot & \cdot & \cdot & \cdot \; \cdot & \cdot \\[1em]
0 & \cdot & \cdot & \cdot \;\; f_{ijt_{ij}-2}+f_{ijt_{ij}-1}-1 & -f_{ijt_{ij}-1}\rho^{d_{ijt_{ij}-1}} \\[1em]
0 & \cdot & \cdot & 0 \;\; -f_{ijt_{ij}-1}\rho^{d_{ijt_{ij}-1}} & f_{ijt_{ij}-1}
\end{pmatrix}
$$

Let $\hat{E}_{ij} = \mathbf{Y}_{ij} - X_{ij}B_i$. The log-likelihood function of $\hat{\epsilon}_{ij} = vec((\hat{E}_{ij})$, $i = 1,...,g$, $j = 1,...,n_i$, is

$$
\begin{aligned}
ln\, l \;=\; & \text{constant} - \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i} ln\mid \mathbf{V}_{ij}\otimes \mathbf{\Sigma}\mid \\
& - \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\epsilon}'_{ij}\mathbf{V}_{ij}^{-1}\otimes \mathbf{\Sigma}^{-1}\hat{\epsilon}_{ij}.
\end{aligned} \tag{5.17}
$$

Following steps similar to what we did in the previous scetion, it is easy to see that

$$
\hat{\mathbf{\Sigma}} = \frac{\displaystyle\sum_{i=1}^{g}\sum_{j=1}^{n_i}\sum_{k=1}^{t_{ij}}\sum_{l=1}^{t_{ij}} v_{kl}^{*}(\hat{\epsilon}_{ijk})(\hat{\epsilon}_{ijl})'}{\displaystyle\sum_{i=1}^{g}\sum_{j=1}^{n_i} t_{ij}}. \tag{5.18}
$$

Next, differentiating $ln\, l$ with respect to ρ and equating to zero, we get

$$
-\frac{p}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}\sum_{k=1}^{t_{ij}-1}\frac{2d_{ijk}\rho^{2d_{ijk}-1}}{1-\rho^{2d_{ijk}}} - \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\epsilon}'_{ij}\frac{\partial}{\partial\rho}[\mathbf{V}_{ij}^{-1}\otimes \mathbf{\Sigma}^{-1}]\hat{\epsilon}'_{ij} = 0
$$

that is,

$$
-\frac{p}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}\sum_{k=1}^{t_{ij}-1}\frac{2d_{ijk}\rho^{2d_{ijk}-1}}{1-\rho^{2d_{ijk}}} - \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}tr(\frac{\partial}{\partial\rho}[\mathbf{V}_{ij}^{-1}\otimes \mathbf{\Sigma}^{-1}]\hat{\epsilon}_{ij}\hat{\epsilon}'_{ij}) = 0
$$

102

i.e.,

$$- \frac{p}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}-1} \frac{2d_{ijk}\rho^{2d_{ijk}-1}}{1-\rho^{2d_{ijk}}} - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} v_{kk} tr(\Sigma^{-1} \epsilon_{ijk} \epsilon_{ijk}')$$

$$+ \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}-1} v_{kk+1} tr(\Sigma^{-1} \epsilon_{ijk} \epsilon_{ijk+1}') \qquad (5.19)$$

where v_{kk} is the kth diagonal element of $\frac{\partial}{\partial \rho} V_{ij}^{-1}$ and v_{kk+1} is the $(k, k+1)^{st}$ off

diagonal element of $\frac{\partial}{\partial \rho} V_{ij}^{-1}$. These elements are calculated below.

To evaluate $\frac{\partial V_{ij}^{-1}}{\partial \rho}$, consider

$$\frac{\partial f_{ijk}}{\partial \rho} = \frac{\partial}{\partial \rho}\left(\frac{1}{1-\rho^{2d_{ijk}}}\right) = \frac{2d_{ijk}\rho^{2d_{ijk}-1}}{(1-\rho^{2d_{ijk}})^2} = f_{ijk}^2 2d_{ijk}\rho^{2d_{ijk}-1}.$$

Thus the diagonal elements of $\frac{\partial}{\partial \rho} V_{ij}^{-1}$ are:

$$(f_{ij1}^2 2d_{ij1}\rho^{2d_{ij1}-1}, f_{ij1}^2 2d_{ij1}\rho^{2d_{ij1}-1} + f_{ij2}^2 2d_{ij2}\rho^{2d_{ij2}-1}, ...,$$

$$f_{ijt_{ij}-2}^2 2d_{ijt_{ij}-2}\rho^{2d_{ijt_{ij}-2}-1} + f_{ijt_{ij}-1}^2 2d_{ijt_{ij}-1}\rho^{2d_{ijt_{ij}-1}-1}, f_{ijt_{ij}-1}^2 2d_{ijt_{ij}-1}\rho^{2d_{ijt_{ij}-1}-1}).$$

Also,

$$\frac{\partial(f_{ijk}\rho^{d_{ijk}})}{\partial \rho} = f_{ijk}d_{ijk}\rho^{d_{ijk}-1} + \rho^{d_{ijk}}\frac{\partial f_{ijk}}{\partial \rho}$$

$$= f_{ijk}d_{ijk}\rho^{d_{ijk}-1} + \rho^{d_{ijk}}f_{ijk}^2 2d_{ijk}\rho^{2d_{ijk}-1}$$

$$= d_{ijk}\rho^{d_{ijk}-1}f_{ijk}[1 + \frac{2\rho^{2d_{ijk}}}{1-\rho^{2d_{ijk}}}]$$

$$= d_{ijk}\rho^{d_{ijk}-1}f_{ijk}[\frac{1+\rho^{2d_{ijk}}}{1-\rho^{2d_{ijk}}}]$$

$$= d_{ijk}\rho^{d_{ijk}-1}f_{ijk}^2(1+\rho^{2d_{ijk}}).$$

Therefore, the off-diagonal elements of $\frac{\partial}{\partial \rho} V_{ij}^{-1}$ are:

103

$$(d_{ij1}\rho^{d_{ij1}-1}f_{ij1}^2(1+\rho^{2d_{ij1}}), d_{ij2}\rho^{d_{ij2}-1}f_{ij2}^2(1+\rho^{2d_{ij2}}), ..., d_{ijt_{ij}-1}\rho^{d_{ijt_{ij}-1}-1}f_{ijt_{ij}-1}^2(1+$$

$$\rho^{2d_{ijt_{ij}-1}})).$$

Plugging these values in (5.19) we can solve for ρ. As before an iterative scheme can be developed to find the ML estimates of B, ρ and Σ.

104

Bibliography

[1] Amemiya, T. (1985). *Advanced Econometrics.* Harvard Univ. Press, Cambridge, Massachusettes.

[2] Arnold, S.F. (1979). Linear models with exchangeably distributed errors. *Journal of American Statistical Association,* **74**, 194-199.

[3] Baldessari, B. (1965). Remarque sur le Rapport de Combinaisons Lineaires de χ^2. *Publications de l'Institut de Statistique de l'Université Paris,* **14**, 393-406.

[4] Boik, J.B. (1988). The mixed model for multivariate repeated measures: validity conditions and an approximate test. *Psychometrika,* **53**, 469-486.

[5] Box, G.E.P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems -I. *Annals of Mathematical Statistics,* **25**, 290-302.

[6] Chaganty, N.R. and Vaish, A.K. (1994). An invariance property of common statistical tests. *To Appear in Linear Algebra and Its Applications.*

[7] Chandra, T.K. and Mukherjee, R. (1991). Bartlett-type modification for Rao's efficient score statistic. *Journal of Multivariate Analysis*, **36**, 103-112.

[8] Chi, E.M. and Reinsel, C. (1989). Models for longitudinal data with random effects and AR(1) errors. *Journal of Americal Statistical Association*, **84**, 452-459.

[9] Crowder, M.J. and Hand, D.J. (1990). *Analysis of Repeated Measures*. Chapman and Hall, London.

[10] Eaton, M.L. (1983). *Multivariate Statistics - A Vector Space Approach*. John Wiley, New York.

[11] Geisser, S. and Greenhouse, S. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. *Annals of Mathematical Statistics*, **29**, 885-891.

[12] Greenhouse, S. and Geisser, S. (1959). On methods of profile data. *Psychometrika*, **24**, 95-112.

[13] Hall, W. J. and Mathaison, D.J. (1990). On large sample estimation and testing in parametric models. *International Statistical Review*, **58**, 1, 77-97.

[14] Hargrave, B.Y. and Castle, C.C. (1995). Effects of phenylephrine induced increase in arterial pressure and closure of the ductus arteriosus on the secretion

of atrial natriuretic peptide (ANP) and renin in the ovine fetus. *Life Sciences*, **57**, 1, 31-43.

[15] Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of Americal Statistical Association*, **72**, 320-340.

[16] Huynh, H. and Feldt, L.S. (1970). Conditions under which mean square ratios in repeated measures designs have exact F-distribution. *Journal of American Statistical Association*, **65**, 1582-1589.

[17] Huynh, H. and Feldt, L.S. (1976). Estimation of Box correction for degrees of freedom from sample data in randomised block and split plot designs. *Journal of Educational Statistics*, **1**, 69-82.

[18] Huynh, H. and Feldt, L.S. (1980). Performance of traditional F tests in repeated measures design under covariance heterogeneity. *Communications in Statistics- Theory and Methods*, **9**, 61-74.

[19] Jenrich, R.I. and Schluchter, M.D. (1986). Unbalanced repeated measures models with structured covariance matrices. *Biometrics*, **42**, 805-820.

[20] Jensen, D.R. (1982). Efficiency and robustness in the use of repeated measurements. *Biometrics*, **38**, 813-825.

[21] Jones, R.H. (1993). *Longitudinal Data with Serial Correlation: A State Space Approach*. Chapman and Hall, London.

[22] Keselman, J.C. and Keselman, H.J. (1990). Analysing unbalanced repeated measures designs. *British Journal of Mathematical and Statistical Psychology*, **43**, 265-282.

[23] Khatri, C.G. (1962). Conditions for Wishartness and independence of second degree polynomials in normal vector. *Annals of Mathematical Statistics*, **33**, 1002-1007.

[24] Khattree, R. and Naik, D.N. (1995). *Applied Multivariate Statistics with SAS Software*. SAS Institute, Cary, North Carolina.

[25] Khuri, I.K., Mathew, T. and Nel, D.G. (1994). A test to determine closeness of multivariate Satterthwaite's approximation. *Journal of Multivariate Analysis*, **50**.

[26] Laird, N.M. and Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**, 963-974.

[27] Mathew, T. (1989). MANOVA in the multivariate components of variance model. *Journal of Multivariate Analysis*, **29**, 30-38.

[28] Mauchly, J.W. (1940). Significance test for sphericity of a normal-n-variate distribution. *Annals of Mathematical Statistics*, **11**, 204-209.

[29] Morrison, D.F. (1976). *Multivariate Statistical Methods.* McGraw- Hill, New York.

[30] Patel, H.I. (1986). Analysis of repeated measures designs with changing covariates in clinical trials. *Biometrika,* **73**, 707-715.

[31] Pavur, R.J. (1987). Distribution of multivariate quadratic forms under certain covariance structures. *Canadian Journal of Statistics,* **15**, 169-176.

[32] Pickering, L.K., Morrow, A.L. and Herrera, I. etal., (1995). Effect of maternal rotavirus immunization on milk and serum antibody titers. *To Appear in Journal of Infectious Diseases.*

[33] Prabhala, S. (1995). *Growth curve models under special covariance structures* PhD Thesis, Old Dominion University.

[34] Quintana, S.M. and Maxwell, S.E. (1994). A Monte Carlo comparison of seven ϵ-adjustment procedures in repeated measures designs with small sample sizes. *Journal of Educational Statistics,* **19**, 57-71.

[35] Rao, C.R. (1948). Tests of significance in multivariate analysis. *Biometrika,* **35**, 58-79.

[36] Rao, C.R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley, New York.

109

[37] Reinsel, G. (1982). Multivariate repeated measurements or growth curve models with multivariate random-effects covariance structure. *Journal of American Statistical Association,* **77**, 190-195.

[38] Rouanet, H. and Lepine, D. (1970). Comparison between treatments in a repeated measures design: ANOVA and multivariate methods. *British Journal of Mathematical and Statistical Psychology,* **23**, 147-163.

[39] Satterthwaite, F.E. (1941). Synthesis of variance. *Psychometrika,* **6**, 309-316.

[40] Tan, W.Y. and Gupta, R.P. (1983). On approximating a linear combination of central wishart matrices with positive coefficients. *Communications in Statistics- Theory and Methods,* **12**, 2589-2600.

[41] Timm, N.H. (1980). Multivariate analysis of variance of repeated measurements. In P. R. Krishnaiah (Ed.), *Handbook of Statistics, Vol. 1 (pp. 41-87),* North-Holland, New York.

[42] Thomas, D.R. (1983). Univariate repeated measures techniques applied to multivariate data. *Psychometrika,* **48**, 451-464.

[43] Vaish, A.K. (1994). *Invariance properties of statistical tests for dependent observations.* PhD Thesis, Old Dominion University.

[44] Verbyla, A.P. (1988). Analysis of repeated measures designs with changing covariates. *Biometrika,* **75**, 172-174.

[45] Verbyla, A.P. and Cullis, B.R. (1990). The analysis of multistratum and spati ally correlated repeated measures data. *Biometrics,* **48**, 1015-1032.

[46] Verbyla, A.P. and Venables, W.N. (1988). An extension of growth curve model. *Biometrika,* **75**, 129-138.

[47] Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of American Mathematical Society,* **54**, 426-482.

[48] Wilks, S.S. (1938). The large sample distribution for the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics,* **9**, 60-62.

[49] Winer, B.J. (1971). *Statistical Principles in Experimental Design (2nd ed.).* McGraw-Hill, New York.

Biographical Statement

Place of Birth

    Mysore, India

Date of Birth

    June 07, 1965

Education

    B.Sc. in Statistics, 1986,
    Bangalore University, Bangalore, India

    M.Sc. in Statistics, 1988,
    Bangalore University, Bangalore, India

    M.S. in Applied and Computational Mathematics, 1994
    Old Dominion University, Norfolk, VA

Employment

    Department of Mathematics and Statistics, Old Dominion University,
    1990-1995, Teaching Assistant

    M.E.S. College of Arts, Science & Commerce, Bangalore, India,
    1988-1989, lecturer

    Department of Statistics, Bangalore University, India,
    1989-1990, Research Assistant