Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Analysis of multivariate repeated measures data with a Kronecker product structured covariance matrix

Dayanand N. Naik [a] & Shantha S. Rao [b]
[a] Old Dominion University, USA
[b] PAREXEL International, USA

Available online: 02 Aug 2010

Journal of Applied Statistics, Vol. 28, No. 1, 2001, 91–105

Analysis of multivariate repeated measures data with a Kronecker product structured covariance matrix

DAYANAND N. NAIK[1] **& SHANTHA S. RAO**[2], [1]*Old Dominion University, USA and* [2]*PAREXEL International, USA*

ABSTRACT *In this article 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, assuming that the data on an individual has a covariance matrix which is a Kronecker product of two positive definite matrices, 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 multivariate normal variables. The multivariate tests using this approximation are developed for testing the usual hypotheses. Results are illustrated on a data set. A method for analysing unbalanced data is also discussed.*

1 Introduction

Measurements on a variable (or a characteristic) made at several occasions or under different treatment conditions on the same experimental unit, lead to repeated measures (or longitudinal) data. Repeated measures data routinely occur in many diverse fields, such as 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 is usually taken on n $(= n_1 + \ldots + n_g)$ individuals forming g groups over t occasions (time points). The problems of interest are to test for the (i) *time effect*, (ii) *group effect*, and (iii) the *effect of interaction between time and group*. Assuming that the vector of t measurements on each individual is a sample from a

Correspondence: Dayanand N. Naik, Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia 23529, USA.

ISSN 0266-4763 print; 1360-0532 online/01/010091-15
DOI: 10.1080/02664760120011626

t-variate normal distribution with a certain positive definite covariance matrix, say $\mathbf{V}$, these problems can be solved using *profile analysis*, a standard technique in multivariate statistical analysis, for example, see Morrison (1976).

In many practical problems, where the repeated measures occur, the covariance matrix $\mathbf{V}$ however, is found to have some structure. In that case, the use of univariate analysis of variance or regression models is recommended in place of multivariate tests. 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. If $\mathbf{V} = \sigma^2\mathbf{V}(\rho)$, where $\mathbf{V}(\rho) = (1 - \rho)\mathbf{I} + \rho\mathbf{1}\mathbf{1}'$, $\mathbf{I}$ is an identity matrix, $\mathbf{1}$ is a vector of ones and ρ is the correlation coefficient between any two repeated measures, then the analysis of repeated measures data is the essentially same as that of a typical split-plot design and the usual tests of hypotheses in split-plot design will address the problems (i)–(iii) (see Winer, 1971).

Many authors, for example, Baldessari (1965), Huynh & Feldt (1970), and Rouanet & Lepine (1970) have characterized the class of all covariance structures $\mathbf{V}$ for which the split-plot ANOVA remains invariant. This structure is termed as type H structure. A typical member of this class is $\mathbf{V} = \sigma^2(\mathbf{I} + \mathbf{a}\mathbf{1}' + \mathbf{1}\mathbf{a}') = \sigma^2 H(\mathbf{a})$, where $\mathbf{a}' = (a_1, \ldots, a_t)$ is an arbitrary vector.

Recently, Chaganty & Vaish (1997) noted that $\mathbf{V}$ need not be non-negative definite for some choices of $\mathbf{a}$ and provided a condition on the elements of $\mathbf{a}$ that would make the matrices in that class non-negative definite. Specifically, $H(\mathbf{a})$ is non-negative definite if it is of the form $H^\star(\mathbf{a}) = \mathbf{I} + (1/t)(\mathbf{a}\mathbf{1}' + \mathbf{1}\mathbf{a}') - (1/t)(1 + \bar{a})\mathbf{1}\mathbf{1}'$ with $(1/t)\sum_{i=1}^{t}(a_i - \bar{a})^2 \leqslant \bar{a}$, $\bar{a}$ being the average of the components of vector $\mathbf{a}$. The strict inequality corresponds to the positive definiteness of $H^\star(\mathbf{a})$.

We can summarize the strategy adopted for analysis of repeated measures data as follows. Test for type H covariance structure using the likelihood ratio test (see Huynh & Feldt, 1970); if accepted, adopt the univariate split-plot type of analysis for solving (i)–(iii), otherwise adopt the profile analysis. When the hypothesis of type H structure for $\mathbf{V}$ is rejected, one can still use the F statistics of the split-plot ANOVA. However, the distributions of these test statistics will no longer be exact. Box (1954) and Geisser & Greenhouse (1958) have developed usable approximations to the distributions of these statistics using a result of Satterthwaite (1941). Satterthwaite (1941) approximated (Satterthwaite approximation) the distribution of a quadratic form in normal variables, by a scale multiple of a χ^2 distribution such that the expected value and the variance of the quadratic form are equal to those of the approximating quantity.

Practical implementation of this method requires estimation of the degrees of freedoms of the approximate χ^2 or the F distributions. Various procedures to estimate the degrees of freedoms are available in the literature. For example, see Greenhouse & Geisser (1959) and Huynh & 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 studies. Some of these results can be found in the papers by Huynh & Feldt (1980), Jensen (1982), Keselman & Keselman (1990), and Quintana & Maxwell (1994). The analysis of repeated measures data, including the estimation of the degrees of freedom, has been successfully implemented in several statistical softwares. A detailed explanation of the analysis of repeated measures data using the SAS software can be found in Khattree & Naik (1999).

When the observations on n experimental units are made on a set of p variables

TABLE 1. MANOVA table for mixed effects model

Source	D.O.F.	SS&CP	Distribution under H_0
Between Groups			
Groups	$g-1$	$\mathbf{Q}_1$	$W_p(g-1, \Sigma)$
Individuals	$n-g$	$\mathbf{Q}_2$	$W_p(n-g, \Sigma)$
Within Groups			
Time	$t-1$	$\mathbf{Q}_3$	$W_p(t-1, \Sigma)$
Time*Group	$(g-1)(t-1)$	$\mathbf{Q}_4$	$W_p((t-1)(g-1), \Sigma)$
Error	$(t-1)(n-g)$	$\mathbf{Q}_5$	$W_p((t-1)(n-g), \Sigma)$
Total	$nt-1$	$\mathbf{Y}(\mathbf{I}_{nt}-\frac{1}{nt}\mathbf{J}_{nt})\mathbf{Y}'$	

(or characteristics) at t occasions, we have a set of *multivariate repeated measures* data. Analysis of such data is complicated by the existence of correlation among the measurements on different variables along with the correlation among measurements taken at different occasions. Several approaches to analyse these data exist in the literature. A brief review of these follows.

Let $\mathbf{y}_{ijk}$, $k=1,\ldots,t$; $j=1,\ldots,n_i$; $i=1,\ldots,g$, be a $p\times 1$ vector of measurements on the jth individual in the ith group at the kth occasion and $\mathbf{y}_{ij}=(\mathbf{y}'_{ij1},\ldots,\mathbf{y}'_{ijt})'$. Then $\mathbf{y}_{ij}$ is $pt\times 1$ random observational vector corresponding to the jth individual in the ith group. Let $\text{cov}(\mathbf{y}_{ij})=\boldsymbol{\Omega}$, for $j=1,\ldots,n_i$; $i=1,\ldots,g$, where $\boldsymbol{\Omega}$ is a positive definite matrix. Using a multivariate linear model of the form $Y=XB+E$, where $n\times pt$, $n=\sum_{i=1}^{g} n_i$ matrix Y is the observation matrix obtained by taking each $\mathbf{y}'_{ij}$ in a row, X is a known design matrix of order $n\times k$, B is $k\times pt$ matrix of unknown parameters and, assuming rows of E independently follow multivariate normal distribution with zero mean vector and covariance matrix $\boldsymbol{\Omega}$, any linear hypothesis about the effect of the time factor, the group factor, or the interaction between them or any linear hypothesis about the elements of the rows of B can be formulated in the form of the general linear hypothesis $H_0:\mathbf{L}B\mathbf{M}=\mathbf{0}$, for known and full rank matrices $\mathbf{L}$ and $\mathbf{M}$. Using Wilks' Λ or any other standard multivariate tests, the null hypothesis H_0 can be tested. This approach for analysing the multivariate repeated measures data, known as doubly multivariate model (DMM) analysis, is commonly adopted in practice (see Timm, 1980).

Another approach taken to analyse multivariate repeated measures data is multivariate mixed model (MMM) analysis. Consider a mixed effects MANOVA model (similar to the split-plot design model of the univariate analysis of the usual repeated measures data) with the effects of the subjects within a group being random. Then the MANOVA table (that is similar to the ANOVA table for the split-plot design model) can be given as in Table 1.

Here, $\mathbf{Y}$, the $nt\times p$ matrix, is defined as $\mathbf{Y}=(\mathbf{y}_{111},\ldots,\mathbf{y}_{11t},\ldots,\mathbf{y}_{1n_11},\ldots,\mathbf{y}_{1n_1t},\ldots,\mathbf{y}_{gn_g1},\ldots,\mathbf{y}_{gn_gt})'$ and Σ, the variance covariance matrix of any row of $\mathbf{Y}$. The matrix quadratic forms $\mathbf{Q}_1$–$\mathbf{Q}_5$ are

$$\mathbf{Q}_1=t\sum_{i=1}^{g} n_i(\bar{\mathbf{y}}_{i..}-\bar{\mathbf{y}}_{...})(\bar{\mathbf{y}}_{i..}-\bar{\mathbf{y}}_{...})'=\mathbf{Y}'\mathbf{A}_1\mathbf{Y},$$

$$\mathbf{Q}_2=t\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\bar{\mathbf{y}}_{ij.}-\bar{\mathbf{y}}_{i..})(\bar{\mathbf{y}}_{ij.}-\bar{\mathbf{y}}_{i..})'=\mathbf{Y}'\mathbf{A}_2\mathbf{Y},$$

$$\mathbf{Q}_3 = n \sum_{k=1}^{t} (\bar{\mathbf{y}}_{..k} - \bar{\mathbf{y}}_{...})(\bar{\mathbf{y}}_{..k} - \bar{\mathbf{y}}_{...})' = \mathbf{Y}'\mathbf{A}_3\mathbf{Y}, \text{ and}$$

$$\mathbf{Q}_4 = \sum_{i=1}^{g} n_i \sum_{k=1}^{t} (\bar{\mathbf{y}}_{i.k} - \bar{\mathbf{y}}_{i..} - \bar{\mathbf{y}}_{..k} + \bar{\mathbf{y}}_{...})(\bar{\mathbf{y}}_{i.k} - \bar{\mathbf{y}}_{i..} - \bar{\mathbf{y}}_{..k} + \bar{\mathbf{y}}_{...})' = \mathbf{Y}'\mathbf{A}_4\mathbf{Y}$$

$$\mathbf{Q}_5 = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t} (\mathbf{y}_{ijk} - \bar{\mathbf{y}}_{ij.} - \bar{\mathbf{y}}_{i.k} + \bar{\mathbf{y}}_{i..})(\mathbf{y}_{ijk} - \bar{\mathbf{y}}_{ij.} - \bar{\mathbf{y}}_{i.k} + \bar{\mathbf{y}}_{i..})' = \mathbf{Y}'\mathbf{A}_5\mathbf{Y}$$

with the appropriate choice of symmetric matrices $\mathbf{A}_1$–$\mathbf{A}_5$ of order $nt \times nt$ and with the usual notations for the sample averages. The matrices $\mathbf{A}_1$–$\mathbf{A}_5$ can be easily derived (for example, see Geisser & Greenhouse, 1958). The matrix quadratic forms $\mathbf{Q}_1$–$\mathbf{Q}_5$ are independent of each other and under the appropriate null hypothesis each has a scale multiple of a Wishart distribution with appropriate degrees of freedom. See the works of Khatri (1962), Arnold (1979), Reinsel (1982), and Mathew (1989), in this regard.

Thomas (1983) considered a class of structures for $\mathbf{\Omega}$, members of which are sufficient to preserve the multivariate analysis of variance (MANOVA) obtained under the mixed effects model invariant, as well as the independence and distributions of the SS&CP matrices of Table 1. Pavur (1987) has characterized the class of all covariance structures for $\mathbf{\Omega}$ under which the MANOVA remains invariant. Vaish (1994), in his PhD thesis, pointed out, using counter examples, that Pavur's (1987) characterization may contain some matrices that are not non-negative definite. Further, the author has provided an elegant characterization of the class of all members of $\mathbf{\Omega}$, that are non-negative definite. Boik (1988) showed independently of Pavur (1987) that the structure considered by Thomas (1983) is necessary and sufficient for MANOVA to remain invariant. He has also developed the likelihood ratio test for testing this structure. Further, Boik (1988) considered the Satterthwaite type approximations (*multivariate Satterthwaite approximation*) to the distributions of various sums of squares and cross-products (SS&CP) matrices of the MANOVA table, when the covariance matrix $\mathbf{\Omega}$ is any general positive definite matrix. While approximating the distribution of a SS&CP matrix by a scale multiple of Wishart distribution, Boik (1988) assumed that the expected value and the trace of the covariance matrix of the SS&CP matrix are equal to those of the approximating matrix. Tan & Gupta (1983) and recently Khuri *et al.* (1994) also have considered the problems relating to multivariate Satterthwaite approximation to a SS&CP matrix. Their approximation is different from Boik's (1988) in the sense that they use a generalized variance (determinant of the covariance matrix) instead of a trace, for their approximations.

All work on multivariate repeated measures done thus far in the literature has the basic assumption that $\text{cov}(\mathbf{y}_{ij}) = \mathbf{\Omega}$, where $\mathbf{\Omega}$ is a positive definite matrix. The structures similar to type H structure on $\mathbf{\Omega}$ were found, under which the MANOVA remains invariant. As we have noted in the previous section, the multivariate Satterthwaite approximations for SS&CP matrices of the MANOVA table (Table 1) were derived under the basic assumption that $\text{cov}(\mathbf{y}_{ij}) = \mathbf{\Omega}$.

However, Boik (1991) while attempting to study the efficiency of MMM analysis noted (a) that even a small departure from multivariate spericity inflates the size of MMM-based tests and (b) that ε-corrected MMM tests adequately control the test sizes unless the covariance matrix departs substantially from a multiplicative-Kronecker structure. Thus, Boik's analysis emphasizes the importance of using the Kronecker product structure for $\mathbf{\Omega}$ as a starting point for analysing multivariate repeated measures data. Also, Galecki (1994) has used the Kronecker product structure for modelling covariance structures for repeated measures specified by more than one repeated factor.

In the next section, we consider this Kronecker product structure as a starting point and derive various results. In Section 3, as an illustration, we apply these results to an example. Next, in Section 4, we discuss some methods to handle unbalanced multivariate repeated measures data. Some concluding remarks are made in the final section.

2 Covariance structure $\mathbf{V} \otimes \Sigma$

In this section, we begin with the assumption that

$$\operatorname{cov}(\mathbf{y}_{ij}) = \boldsymbol{\Omega} = \mathbf{V} \otimes \Sigma$$

where $\mathbf{V}$ and Σ, respectively, are $t \times t$ and $p \times p$ positive definite matrices. This structure has several advantages over the general covariance structure. First, it is well known (see Crowder & Hand, 1993; Jones, 1993) that the correlation matrix of the repeated measures usually has a simpler structure as opposed to a general structure. In our formulation, it is easier to accommodate different structures for the covariance matrix of repeated measures (via $\mathbf{V}$). Secondly, the number of unknown parameters of the 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 the general covariance structure. Next, as will be seen later, the multivariate Satterthwaite approximations to various SS&CP matrices is much simpler in the present formulation. Furthermore, this structure enables us to handle unbalanced multivariate repeated measures data more easily.

When $\operatorname{cov}(\mathbf{y}_{ij}) = \mathbf{V} \otimes \Sigma$ is considered, note that any column, say the lth column, of the $nt \times p$ matrix $\mathbf{Y}$, has the covariance matrix proportional to $\Delta = \mathbf{I}_n \otimes \mathbf{V}$. In fact, it is $\sigma_{ii}\Delta$, where σ_{ii} is the ith diagonal element of Σ. Suppose $\mathbf{V} = \mathbf{V}(\rho) = (1-\rho)\mathbf{I}_t + \rho\mathbf{J}_t$. Using the multivariate analogue of Cochran's theorem for quadratic forms (Siotani *et al.*, 1985) it can be shown that the MANOVA in Table 1 remains invariant for this case (see Theorem 4.6.3 of Vaish, 1994). Further, a characterization of the class of structures for $\mathbf{V}$ such that the MANOVA remains invariant yields type H structure for $\mathbf{V}$. Many interesting results about the structure $\mathbf{V} \otimes \Sigma$ are summarized in Vaish (1994).

2.1 Test for Type H structure

A likelihood ratio test for testing

$$\mathrm{H}_0 : \operatorname{cov}(\mathbf{y}_{ij}) = H(\mathbf{a}) \otimes \Sigma \text{ vs. } \mathrm{H}_a : \operatorname{cov}(\mathbf{y}_{ij}) = \mathbf{V} \otimes \Sigma$$

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

$$\mathrm{H}_0 : \operatorname{cov}(\mathbf{u}_{ij}) = \mathbf{I} \otimes \Sigma \text{ vs. } \mathrm{H}_a : \operatorname{cov}(\mathbf{u}_{ij}) = \mathbf{W} \otimes \Sigma \tag{1}$$

A likelihood ratio test for testing (1) is given by

$$\lambda = \frac{|\hat{\mathbf{W}}|^{np/2} |\hat{\Sigma}|^{n(t-1)/2}}{|\hat{\Sigma}_0|^{n(t-1)/2}} \tag{2}$$

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 the standard asymptotic results $-2\ln\lambda$ has a χ^2 distribution with $(t+1)(t-2)/2$ degrees of freedom. This is because, without loss of generality $\mathbf{W}\otimes\Sigma$ can be restricted so that $w_{11}=1$.

Derivation of the MLEs. We first derive the maximum likelihood estimators of $\mathbf{W}$ and Σ under (H_a). We have a random sample $\mathbf{u}_{ij}$, $j=1,\ldots,n$; $i=1,\ldots,g$ from a $p(t-1)$-variate normal distribution with a mean $\boldsymbol{\mu}_i'$ and covariance matrix $\mathbf{W}\otimes\Sigma$. Here, $\boldsymbol{\mu}_i$ is a $p(t-1)\times 1$ vector given by $\boldsymbol{\mu}_i'=(\mu_{i1}',\ldots,\mu_{i(t-1)}')$, 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. Then, the log-likelihood function of the parameters is given by

$$\ln l = -\frac{p(t-1)n}{2}\ln(2\pi) - \frac{n}{2}\ln|\mathbf{W}\otimes\Sigma| - \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\mathbf{u}_{ij}-\boldsymbol{\mu}_i)'\mathbf{W}^{-1}\otimes\Sigma^{-1}(\mathbf{u}_{ij}-\boldsymbol{\mu}_i)$$

Using

$$B=\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\mathbf{u}_{ij}-\bar{\mathbf{u}}_{i.})(\mathbf{u}_{ij}-\bar{\mathbf{u}}_{i.})', \text{ with } \bar{\mathbf{u}}_{i.}=\sum_{j=1}^{n_i}\mathbf{u}_{ij}/n_i$$

the log-likelihood function can be rewritten as

$$\ln l = -\frac{np(t-1)}{2}\ln(2\pi) - \frac{n}{2}\ln|\mathbf{W}\otimes\Sigma| - \frac{1}{2}\mathrm{tr}(\mathbf{W}^{-1}\otimes\Sigma^{-1})B - \frac{1}{2}\sum_{i=1}^{g}n_i(\bar{\mathbf{u}}_{i.}-\boldsymbol{\mu}_i)'\mathbf{W}^{-1}\otimes\Sigma^{-1}(\bar{\mathbf{u}}_{i.}-\boldsymbol{\mu}_i) \tag{3}$$

Next, we partition the $p(t-1)\times 1$ vector $(\mathbf{u}_{ij}-\bar{\mathbf{u}}_{i.})$ into $(t-1)$ blocks of $p\times 1$ vectors such that $(\mathbf{u}_{ij}-\bar{\mathbf{u}}_{i.})=(\mathbf{u}_{ij1}-\bar{\mathbf{u}}_{i.1},\ldots,\mathbf{u}_{ij(t-1)}-\bar{\mathbf{u}}_{i.(t-1)})'$. Using this partition of $(\mathbf{u}_{ij}-\bar{\mathbf{u}}_{i.})$, we rewrite (3) and maximize it simultaneously with respect to Σ and $\mathbf{W}$. We get the following likelihood equations for estimating Σ and $\mathbf{W}$.

$$\hat{\Sigma}=\frac{1}{n(t-1)}\sum_{k=1}^{t-1}\sum_{l=1}^{t-1}\hat{w}_{kl}^{\star}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\mathbf{u}_{ijk}-\bar{\mathbf{u}}_{i.k})(\mathbf{u}_{ijl}-\bar{\mathbf{u}}_{i.l}) \tag{4}$$

$$\hat{\mathbf{W}}=\frac{1}{np}A \tag{5}$$

where $\hat{w}_{kl}^{\star}$ is the (kl)th element of $\hat{\mathbf{W}}^{-1}$, and the (kl)th element of matrix A is given by $A_{kl}=\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\mathbf{u}_{ijk}-\bar{\mathbf{u}}_{i.k})'\hat{\Sigma}^{-1}(\mathbf{u}_{ijl}-\bar{\mathbf{u}}_{i.l})$.

The above equations are to be solved iteratively to get the estimates of Σ 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) a matrix of difference between two successive solutions of (4), and (b) a matrix of difference between two successive solutions of (5). Continue the iterations until the maxima 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 PROC IML of SAS software is available from the authors to compute these estimates.

Mardia & Goodall (1993) considered the problem of modelling spatial-temporal multivariate environmental monitoring data. After removing the effect of time, their spatial multivariate data become the multivariate repeated measures data. For these data, Mardia & Goodall (1993) have indeed considered a Kronecker product covariance structure for modelling the covariance of the measurements at each monitoring site. They have further provided an algorithm for computing the maximum likelihood estimates of the parameters. For this alternative algorithm and other details, see Mardia & Goodall (1993).

Now the maximum likelihood estimate of Σ under H_0 is given by

$$\hat{\Sigma}_0 = \frac{1}{n(t-1)} \sum_{k=1}^{t-1} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (\mathbf{u}_{ijk} - \bar{\mathbf{u}}_{i.k})\ (\mathbf{u}_{ijk} - \bar{\mathbf{u}}_{i.k})' \tag{6}$$

If H_0 is accepted, one can use the MANOVA of Table 1 to analyse the data. Otherwise, multivariate Satterthwaite type approximation can be applied to find the approximation to the distributions of the SS&CP matrices. This approximation will be described next.

2.2 *Exact null distribution of* $\mathbf{Q}_i$

In the following, under no structures on $\mathbf{V}$, we derive the distribution of $\mathbf{Q}_i = \mathbf{Y}'\mathbf{A}_i\mathbf{Y}$, where $\mathbf{A}_i$ is an appropriately defined symmetric matrix of order $nt \times nt$ as defined in Section 1. First of all, it is easy to show that $\mathbf{A}_1\Delta\mathbf{A}_2 = \mathbf{A}_1\Delta\mathbf{A}_5 = \mathbf{A}_2\Delta\mathbf{A}_5 = \mathbf{A}_3\Delta\mathbf{A}_5 = \mathbf{A}_4\Delta\mathbf{A}_5 = \mathbf{0}$ (see Geisser & Greenhouse, 1958), where $\Delta = \mathbf{I}_n \otimes \mathbf{V}$. Hence, the matrix quadratic forms $\mathbf{Q}_1$ and $\mathbf{Q}_2$, $\mathbf{Q}_1$ and $\mathbf{Q}_5$, $\mathbf{Q}_2$ and $\mathbf{Q}_5$, $\mathbf{Q}_3$ and $\mathbf{Q}_5$, and $\mathbf{Q}_4$ and $\mathbf{Q}_5$, are all pairwise independent. Let the rank of $\mathbf{A}_i = v_i$ (for example, $v_1 = g - 1$ and so on). We observe that each column of $\mathbf{Y}$ is a multivariate normal vector of order $nt \times 1$ and has a covariate matrix proportional to $\Delta = \mathbf{I}_n \otimes \mathbf{V}$. Since, by assumption, Δ is a positive definite matrix, there exist $\Delta^{1/2}$ and $\Delta^{-1/2}$ such that $\Delta = \Delta^{1/2}\Delta^{1/2}$ and $\Delta^{-1} = \Delta^{-1/2}\Delta^{-1/2}$ and that $\Delta^{-1/2}\Delta^{1/2} = \mathbf{I}_{nt}$.

Consider $\mathbf{Q}_i = \mathbf{Y}'\mathbf{A}_i\mathbf{Y} = \mathbf{Y}'\Delta^{-1/2}\Delta^{1/2}\mathbf{A}_i\Delta^{1/2}\Delta^{-1/2}\mathbf{Y} = \mathbf{Z}'\mathbf{B}_i\mathbf{Z}$ with $\mathbf{Z} = \Delta^{-1/2}\mathbf{Y}$ and $\mathbf{B}_i = \Delta^{1/2}\mathbf{A}_i\Delta^{1/2}$. Now the rows of $\mathbf{Z}$ form a random sample of size nt from $N_p(0, \Sigma)$ under certain null hypotheses. Since $\mathbf{B}_i$ is a symmetric matrix with rank v_i, it can be written as $\mathbf{B}_i = \Gamma\Lambda_i\Gamma'$, where $\Gamma\Gamma' = \Gamma'\Gamma = \mathbf{I}_{nt}$ and $\Lambda_i = \mathrm{Diag}(\lambda_1, \ldots, \lambda_{v_i}, 0, \ldots, 0)$, $\lambda_1 \geqslant \ldots \geqslant \lambda_{v_i}$ being the eigenvalues of $\mathbf{B}_i$. Thus, $\mathbf{Q}_i = \mathbf{Z}\mathbf{B}_i\mathbf{Z} = \mathbf{Z}'\Gamma\Lambda_i\Gamma'\mathbf{Z} = \mathbf{U}'\Lambda_i\mathbf{U} = \sum_{j=1}^{v_i} \lambda_j \mathbf{U}_j\mathbf{U}_j'$, $\mathbf{U}_j \sim N_p(0, \Sigma)$ under H_0 and $\mathbf{U}_1, \ldots, \mathbf{U}_{v_i}$ are all independent, where $\mathbf{U} = \Gamma'\mathbf{Z}$ and $\mathbf{U}_j$ is the jth row of $\mathbf{U}$.

It is well known that $\mathbf{U}_j\mathbf{U}_j' \sim W_p(1, \Sigma)$, which is a *pseudo-Wishart distribution* (since the number of degrees of freedom is less than or equal to the dimension) and has no *probability density function*. Thus, the distribution of $\mathbf{Q}_i$ is the same as the distribution of the linear combination of v_i pseudo-Wishart random matrices. In summary, we have the following:

$$\mathbf{Q}_1 \sim W_p((g-1), \Sigma), \mathbf{Q}_2 \sim W_p((n-g), \Sigma),$$

$$\mathbf{Q}_3 \sim \sum_{j=1}^{t-1} \lambda_j W_p^{(j)}(1, \Sigma), \mathbf{Q}_4 \sim \sum_{j=1}^{t-1} \lambda_j W_p^{(j)}(g-1, \Sigma),$$

$$\mathbf{Q}_5 \sim \sum_{j=1}^{t-1} \lambda_j W_p^{(j)}(n-g, \Sigma)$$

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

Suppose we want to test H_{01} of no group effect. Then the Wilks' Λ for testing H_{01} is

$$\Lambda_1 = \frac{|\mathbf{Q}_2|}{|\mathbf{Q}_1 + \mathbf{Q}_2|}$$

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

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

The approximate distributions of the test statistics for testing H_{02}, that there is no time effect, and H_{03}, that there is no time and group interaction, will be derived next.

2.3 Approximate null distributions of $\mathbf{Q}_3$, $\mathbf{Q}_4$ *and* $\mathbf{Q}_5$

In this section, we approximate the distribution of each of SS&CP matrices, $\mathbf{Q}_3$, $\mathbf{Q}_4$ and $\mathbf{Q}_5$, 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 $\mathbf{Q}_i$. It is well known (for example, see Eaton, 1983, p. 305) that for any $\mathbf{S} \sim W_p(v, \Sigma)$, $E(\mathbf{S}) = v\Sigma$ and $D(\mathbf{S}) = 2v\Sigma \otimes \Sigma$. Here, $D(\mathbf{S})$ denotes the variance covariance matrix of all the random quantities in $\mathbf{S}$. Using these formulae we have the following:

$$E(\mathbf{Q}_3) = \left[\mathrm{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{J}\mathbf{V}\right)\right]\Sigma \text{ under } \mathrm{H}_{02}$$

$$E(\mathbf{Q}_4) = \left[(g-1)\mathrm{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{J}\mathbf{V}\right)\right]\Sigma \text{ under } \mathrm{H}_{03}$$

$$\mathrm{E}(\mathbf{Q}_5) = \left[(n-g)\mathrm{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{J}\mathbf{V}\right)\right]\Sigma$$

also

$$D(\mathbf{Q}_3) = \left[2\mathrm{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{J}\mathbf{V}\right)^2\right]\Sigma \otimes \Sigma \text{ under } \mathrm{H}_{02}$$

$$D(\mathbf{Q}_4) = \left[2(g-1)\mathrm{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{J}\mathbf{V}\right)^2\right]\Sigma \otimes \Sigma \text{ under } \mathrm{H}_{03}$$

$$\mathrm{D}(\mathbf{Q}_5) = \left[2(n-g)\mathrm{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{J}\mathbf{V}\right)^2\right]\Sigma \otimes \Sigma$$

Suppose we want to approximate the distribution of $\mathbf{Q}_3$ by a random matrix having the distribution $g_1 W_p(h_1, \Sigma)$ so that the first two central moments of $\mathbf{Q}_3$ and $g_1 W_p(h_1, \Sigma)$ are the same. Then,

$$\operatorname{tr}\left[\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)\right]\Sigma = g_1 h_1 \Sigma \tag{7}$$

$$2\left[\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)^2\right]\Sigma \otimes \Sigma = 2g_1^2 h_1 \Sigma \otimes \Sigma \tag{8}$$

From (7) and (8)

$$g_1 = \frac{\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)^2}{\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)} \quad \text{and} \quad h_1 = \frac{\left[\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)\right]^2}{\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)^2}$$

Next, to approximate the distribution of $\mathbf{Q}_4$ by a random matrix having the distribution $g_2 W_p(h_2, \Sigma)$ so that the first two central moments of $\mathbf{Q}_4$ and $g_2 W_p(h_2, \Sigma)$ are the same we have,

$$(g-1)\operatorname{tr}\left[\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)\right]\Sigma = g_2 h_2 \Sigma \tag{9}$$

$$2(g-1)\left[\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)^2\right]\Sigma \otimes \Sigma = 2g_2^2 h_2 \Sigma \otimes \Sigma \tag{10}$$

From (9) and (10) we have,

$$g_2 = \frac{\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)^2}{\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)} = g_1 \quad \text{and} \quad h_2 = (g-1)\frac{\left[\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)\right]^2}{\operatorname{tr}\left(\mathbf{V} - \frac{1}{t}\mathbf{JV}\right)^2} = (g-1)h_1.$$

Therefore, $\mathbf{Q}_4 \sim g_1 W_p((g-1)h_1, \Sigma)$ approximately. Similarly, it can be shown that $\mathbf{Q}_5 \sim g_1 W_p((n-g)h_1, \Sigma)$ approximately.

Now, for testing H_{02} one can use the Wilks' Λ, which is, $\Lambda_2 = (|\mathbf{Q}_5|/|\mathbf{Q}_3 + \mathbf{Q}_5|)$ and the fact that

$$-\left[(n-g)h_1 - \left(\frac{p+1-h_1}{2}\right)\right]\ln \Lambda_2 \sim \chi^2_{ph_1} \text{ approximately.} \tag{11}$$

Similarly, for testing H_{03} the Wilks' Λ is $\Lambda_3 = (|\mathbf{Q}_5|/|\mathbf{Q}_4 + \mathbf{Q}_5|)$ and the distribution of the test statistic is

$$-\left[(n-g)h_1 - \left(\frac{p+1-(g-1)h_1}{2}\right)\right]\ln \Lambda_3 \sim \chi^2_{p(g-1)h_1} \text{ approximately.} \tag{12}$$

2.4 Estimation of degrees of freedom

Since, in practice, $\mathbf{V}_{t\times t}$ is unknown, the degrees of freedom in the χ^2 approximations of (11) and (12) are unknown. One needs to estimate these so that the distributions in (11) and (12) can be utilized in applications. For estimating these degrees of freedoms, which are functions of $(\mathbf{V} - (1/t)\mathbf{J}\mathbf{V})$, we simply need an estimate of $\mathbf{V}$. One can use the maximum likelihood estimate of $\mathbf{V}$ that is obtained by simultaneously solving the following equations:

$$\hat{\Sigma} = \frac{1}{nt} \sum_{k=1}^{t} \sum_{l=1}^{t} \hat{v}_{kl}^{\star} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (\mathbf{y}_{ijk} - \bar{\mathbf{y}}_{i.k})\,(\mathbf{y}_{ijl} - \bar{\mathbf{y}}_{i.l}) \tag{13}$$

$$\hat{\mathbf{V}} = \frac{1}{np} A \tag{14}$$

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

The effects of estimated degrees of freedom and other aspects, such as heterogeneity of covariance and sample size considerations, on the approximated distributions have to be studied using simulation experiments. We defer these types of discussion to a later date. In the next section, we consider an example and illustrate the estimation of the degrees of freedom using the maximum likelihood estimate of $\mathbf{V}$.

3 An example

To illustrate our methods, we use data from Table 7.2 in Timm (1980), which was also analysed by Thomas (1983) and Boik (1988). These data were obtained by Neil Timm from T. Zullo of the School of Dental Medicine at the University of Pittsburgh. The study is concerned with the relative effectiveness of two orthopaedic adjustments of the mandible. Nine subjects were assigned to each of two orthopaedic 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 $\mathbf{Y} = (\mathbf{y}_{ij})$ is an 18×9 matrix.

The choice of method to analyse our data, rests on the test of hypothesis about the covariance structure described in the previous section. Hence, to test the null hypothesis (1), first we transform $\mathbf{y}_{ij}$ to $\mathbf{u}_{ij}$ using the transformation $\mathbf{u}_{ij} = (\mathbf{C} \otimes \mathbf{I})\mathbf{y}_{ij}$, with

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

Then we have $\mathrm{cov}(\mathbf{u}_{ij}) = \mathbf{W} \otimes \Sigma$.

The maximum likelihood estimates of Σ and $\mathbf{W}$, simultaneously solving (4) and (5), are given by

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

TABLE 2. Approximate MANOVA

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	0.0001
Time*Group	0.8145	3.7891	4.5279	0.3108

and

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

The estimate of Σ under H_0 using (6) 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 (2) is $\lambda = 0.0000139$. Then the test statistic value, $-2\ln\lambda$, is equal to 22.37366. Comparing this with $\chi^2_3(0.05) = 7.815$, we clearly reject H_0 given in (1). Therefore, we use a multivariate Satterthwaite type approximation described in Sections (2.3–2.4), to find an approximation to the distributions of the SS&CP matrices. In order to use these methods we need to estimate g_1 and h_1, which in turn requires the estimate of $\mathbf{V}$. We simultaneously solve (13) and (14) to get the maximum likelihood estimate of $\mathbf{V}$ and it is found to be

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

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

We summarize our results in Table 2.

For computing the above p-values one can use any easily available software. For example, we have used the PROBF function of SAS Software. Note from Table 2 that only the effect of the time factor is significant.

4 Unbalanced data

Thus far we have considered only a balanced case of multivariate repeated measures. However, in practice, measurements on the individuals at all the t time points may not be available. This leads to unbalanced data. Suppose we have t_{ij} repeated measurements on p variables on the jth individual ($j = 1, \ldots, n_i$) from the ith group ($i = 1, \ldots, g$). Let $\mathbf{y}'_{ij} = (\mathbf{y}'_{ij1}, \ldots, \mathbf{y}'_{ijt_{ij}})$ be the observational vector representing these measurements. Assume $\mathrm{cov}(\mathbf{y}_{ij}) = \mathbf{V}_{ij} \otimes \Sigma$, where $\mathbf{V}_{ij}$ is a $t_{ij} \times t_{ij}$ positive definite matrix and, as before, Σ is a $p \times p$ positive definite matrix. The covariance matrix $\mathbf{V}_{ij}$ is not completely different for every i, j, but it is the $t_{ij} \times t_{ij}$ submatrix of a $t \times t$ positive definite matrix $\mathbf{V}$, t being the maximum of t_{ij}. In applications, $\mathbf{V}_{ij}$ is assumed to have some structure depending on the same set of r parameters, say

$\theta_1, \ldots, \theta_r$. Exact small sample inference for this type of data is not very easy. In the following, we present a general multivariate linear model approach to analyse unbalanced data, assuming $\mathbf{V}_{ij}$ is a function of $\theta_1, \ldots, \theta_r$.

Consider the multivariate linear model $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$, where $\mathbf{Y}$ is the $n \times p$ observation matrix, obtained by stacking all $n = \sum_{i=1}^{g} \sum_{j=1}^{n_i} t_{ij}$, $p \times 1$ vectors, $\mathbf{y}_{ijk}$, as rows and $\mathbf{B}$ is the $m \times p$ matrix of unknown parameters. Next, the design matrix $\mathbf{X}$ is obtained by stacking all the $t_{ij} \times m$ matrices $\mathbf{X}_{ij}$ one below the other, where $\mathbf{X}_{ij}$ is the design matrix corresponding to the jth individual in the ith group and is obtained as follows. Suppose $\mathbf{X}_i$ is the $t \times m$ design matrix for the jth individual (same for all j) in group i, when data are balanced, that is, when $t_{ij} = t$. Let $\mathbf{G}_{ij}$ be the $t_{ij} \times m$ matrix of 0s and 1s 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 $l_1, \ldots, l_{t_{ij}}$. Then $\mathbf{X}_{ij} = \mathbf{G}_{ij}\mathbf{X}_i$. Further, $\mathrm{cov}(\mathrm{vec}(\mathbf{E}')) = \mathbf{F} = \mathrm{diag}(\mathbf{V}_{11}, \ldots, \mathbf{V}_{1n_1}, \ldots, \mathbf{V}_{gn_g}) \otimes \Sigma$.

Then, from the usual multivariate regression theory

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

$$= \left[\sum_{i=1}^{g} \sum_{j=1}^{n_1} \mathbf{X}_{ij}'\mathbf{V}_{ij}^{-1}\mathbf{X}_{ij} \right]^{-1} \left[\sum_{i=1}^{g} \sum_{j=1}^{n_i} \mathbf{X}_{ij}'\mathbf{V}_{ij}^{-1}\mathbf{Y}_{ij} \right]$$

and

$$\hat{\Sigma} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'\mathbf{F}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})/(n - m)$$

Using this $\hat{\mathbf{B}}$ (assuming $\theta_1, \ldots, \theta_r$ are known), any one of the multivariate tests, for example Wilks' Λ, for testing $\mathrm{H}_0 : \mathbf{LBM} = \mathbf{0}$ can be used to test the usual hypothesis. But $\mathbf{V}_{ij}$ and, in turn, $\mathbf{F}$ is usually unknown. Using some consistent estimators of $\theta_1, \ldots, \theta_r$ one can proceed to test these hypotheses.

ML Estimation of $\mathbf{V}_{ij}$ *and* Σ. Let $\boldsymbol{\varepsilon}_{ij} = \mathbf{y}_{ij} - E(\mathbf{y}_{ij})$. Then the log-likelihood function, assuming that $\boldsymbol{\varepsilon}_{ij}$, $i = 1, \ldots, g; j = 1, \ldots, n_i$, forms a random sample from pt_{ij}-variate normal distribution with zero mean vector and covariance matrix $\mathbf{V}_{ij} \otimes \Sigma$ is proportional to

$$-\frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \ln |\mathbf{V}_{ij} \otimes \Sigma| - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \boldsymbol{\varepsilon}_{ij}'\mathbf{V}_{ij}^{-1} \otimes \Sigma^{-1}\boldsymbol{\varepsilon}_{ij} \tag{15}$$

Next, partition the $pt_{ij} \times 1$ vector $\boldsymbol{\varepsilon}_{ij}$ into t_{ij} blocks of $p \times 1$ vectors such that $\boldsymbol{\varepsilon}_{ij} = (\varepsilon_{ij1,\ldots,}\varepsilon_{ijt_{ij}})'$. Using this partition, rewrite (15) as

$$-\frac{p}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \ln |\mathbf{V}_{ij}| - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} t_{ij} \ln |\Sigma| - \frac{1}{2} \sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \sum_{l=1}^{t_{ij}} v_{kl}^{\star} \varepsilon_{ijk}' \Sigma^{-1} \varepsilon_{ijl} \tag{16}$$

where $v_{kl}^{\star}$ is the (k,l)th element of $\mathbf{V}_{ij}^{-1}$. Differentiating (16) with respect to Σ and equating to zero we get

$$\hat{\Sigma} = \frac{\sum_{i=1}^{g} \sum_{j=1}^{n_i} \sum_{k=1}^{t_{ij}} \sum_{l=1}^{t_{ij}} v_{kl}^{\star} (\varepsilon_{ijk}) \; (\varepsilon_{ijl})'}{\sum_{i=1}^{g} \sum_{j=1}^{n_i} t_{ij}} \tag{17}$$

Suppose a first-order autoregressive (AR(1)) structure $(\rho^{|k-l|})$, ρ being the correlation coefficient between any two successive measurements, is assumed for $\mathbf{V}$ and data are missing only at the last few occasions (monotone data). Then the basic structure of the covariance matrix for the jth individual in the ith group remains the same, that is, $\mathbf{V}_{ij} = (\rho^{|k-l|})$, $k, l = 1, \ldots, t_{ij}$. Now $v^{\star}_{kl}$ in (17) is the (kl)th element of $\mathbf{V}_{ij}^{-1}$, where

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

$\mathbf{C}_{1ij}$ and $\mathbf{C}_{2ij}$ are $t_{ij} \times t_{ij}$ matrices such that $\mathbf{C}_1 = \mathrm{diag}(0, 1, \ldots, 1, 0)$ and $\mathbf{C}_2$ is a tridiagonal matrix with 0 on the diagonal and 1 on the upper and lower diagonals. Next, using $|\mathbf{V}_{ij}| = (1-\rho^2)^{t_{ij}-1}$ and $\mathbf{V}_{ij}^{-1}$, rewrite (15) as

$$\begin{aligned} &-\frac{p}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(t_{ij}-1)\ln(1-\rho^2) - \frac{1}{2}\sum_{i=1}^{g}\sum_{j=1}^{n_i} t_{ij}\ln|\Sigma| \\ &-\frac{1}{2(1-\rho^2)}\sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\varepsilon}'_{ij}[(\mathbf{I}_{t_{ij}}\otimes\Sigma^{-1} + \rho^2\mathbf{C}_{1ij}\otimes\Sigma^{-1} - \rho\mathbf{C}_{2ij}\otimes\Sigma^{-1}]\hat{\varepsilon}_{ij} \end{aligned} \tag{18}$$

Differentiating (18) with respect to ρ and equating to zero yields,

$$-2\rho^3 p n_d + \rho^2 A_3 + 2\rho(p n_d - A_1 - A_2) + A_3 = 0 \tag{19}$$

where

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

$$A_2 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\varepsilon}'_{ij}(\mathbf{C}_{1ij}\otimes\Sigma^{-1})\hat{\varepsilon}_{ij}, \text{ and } A_3 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\varepsilon}'_{ij}(\mathbf{C}_{2ij}\otimes\Sigma^{-1})\hat{\varepsilon}_{ij}.$$

The cubic equation (19) can be studied further to show that there is a unique root in the interval $(-1, 1)$. Solving (19) and (17) with appropriate $v^{\star}_{kl}$ simultaneously we get the estimates of ρ and Σ. Note that the form of $\mathbf{V}_{ij}$ will be distorted if data were missing from any occasions (not necessarily at the end) and will lead to what is called a Markov structure. The estimation of ρ in this case can be performed, although things are notationally more complicated.

5 Concluding remarks

In this paper, we have illustrated the use of the Kronecker product covariance matrix to analyse multivariate repeated measures data. The advantages of using this covariance include, flexibility in using the structured covariance matrices for the repeated measures, savings in degrees of freedom for estimation and testing of hypothesis, and analysis of unbalanced repeated measures data. An algorithm for calculation of the maximum likelihood estimates is also given. Modelling the covariance by the Kronecker product structure with more than two matrices also has applications. For example, for modelling the covariance of multivariate environmental monitoring data obtained repeatedly over time and space, Mardia & Goodall (1993) have used a Kronecker product of three matrices.

Acknowledgements

We wish to thank Professor N. Rao Chaganty for his many helpful discussions and support. We also want to thank an anonymous referee who drew our attention to the paper by Mardia & Goodall (1993). This research is partially supported by SDIO/IST and managed by the US Army research office, grant number DAAL03-91-G-0179.

REFERENCES

ARNOLD, S. F. (1979) Linear models with exchangeably distributed errors, *Journal of American Statistical Association*, 74, pp. 194–199.

BALDESSARI, B. (1965) Remarque sur le rapport de combinaisons lineaires de χ^2, *Publications de l'Institut de Statistique de l'Université Paris*, 14, pp. 393–406.

BOIK, J. B. (1988) The mixed model for multivariate repeated measures: validity conditions and an approximate test, *Psychometrika*, 53, pp. 469–486.

BOIK, J. B. (1991) Scheffé's mixed model for multivariate repeated measures: A relative efficiency evaluation, *Communications in Statistics, Theory and Methods*, 20, pp. 1233–1255.

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, pp. 290–302.

CHAGANTY, N. R. & VAISH, A. K. (1997) An invariance property of common statistical tests, *Linear Algebra and its Applications*, 264, pp. 421–437.

CROWDER, M. J. & HAND, D. J. (1993) *Analysis of Repeated Measures* (London, Chapman & Hall).

EATON, M. L. (1983) *Multivariate Statistics—a Vector Space Approach* (New York, Wiley).

GALECKI, A. T. (1994) General class of covariance structures for two or more repeated factors in longitudinal data analysis, *Communications in Statistics, Theory and Methods*, 23, pp. 3105–3120.

GEISSER, S. & GREENHOUSE, S. (1958) An extension of Box's results on the use of the *F* distribution in multivariate analysis, *Annals of Mathematical Statistics*, 29, pp. 885–891.

GREENHOUSE, S. & GEISSER, S. (1959) On methods of profile data, *Psychometrika*, 24, pp. 95–112.

HUYNH, H. & FELDT, L. S. (1970) Conditions under which mean square ratios in repeated measures designs have exact *F*-distribution, *Journal of American Statistical Association*, 65, pp. 1582–1589.

HUYNH, H. & 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, pp. 69–82.

HUYNH, H. & FELDT, L. S. (1980) Performance of traditional *F* tests in repeated measures design under covariance heterogeneity, *Communications in Statistics, Theory and Methods*, 9, pp. 61–74.

JENSEN, D. R. (1982) Efficiency and robustness in the use of repeated measurements, *Biometrics*, 38, pp. 813–825.

JONES, R. H. (1993) *Longitudinal Data with Serial Correlation: A State-Space Approach* (London, Chapman & Hall).

KESELMAN, J. C. & KESELMAN, H. J. (1990) Analysing unbalanced repeated measures designs, *British Journal of Mathematical and Statistical Psychology*, 43, pp. 265–282.

KHATRI, C. G. (1962) Conditions for Wishartness and independence of second degree polynomials in normal vector, *Annals of Mathematical Statistics*, 33, pp. 1002–1007.

KHATTREE, R. & NAIK, D. N. (1999) *Applied Multivariate Statistics with SAS Software*, 2nd edn (Cary, North Carolina, SAS Institute and New York, Wiley).

KHURI, I. K., MATHEW, T. & NEL, D. G. (1994) A test to determine closeness of multivariate Satterthwaite's approximation, *Journal of Multivariate Analysis*, 51, pp. 201–209.

MARDIA, K. V. & GOODALL, C. (1993) Spatial-temporal analysis of multivariate environmental monitoring data. *In:* G. P. PATIL & C. R. RAO (eds), *Multivariate Environmental Statistics*, vol. 6, pp. 347–385.

MATHEW, T. (1989) MANOVA in the multivariate components of variance model, *Journal of Multivariate Analysis*, 29, pp. 30–38.

MORRISON, D. F. (1976) *Multivariate Statistical Methods* (New York, McGraw Hill).

PAVUR, R. J. (1987) Distribution of multivariate quadratic forms under certain covariance structures, *Canadian Journal of Statistics*, 15, pp. 169–176.

QUINTANA, S. M. & 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, pp. 57–71.

RAO, C. R. (1973) *Linear Statistical Inference and Its Applications*, 2nd edn (New York, Wiley).

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

ROUANET, H. & LEPINE, D. (1970) Comparison between treatments in a repeated measures design: ANOVA and multivariate methods, *British Journal of Mathematical and Statistical Psychology*, 23, pp. 147–163.

SATTERTHWAITE, F. E. (1941) Synthesis of variance, *Psychometrika*, 6, pp. 309–316.

SIOTANI, M., HAYAKAWA, T. & FUJIKOSHI, Y. (1985) *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook* (Columbus, Ohio, American Science Press).

TAN, W. Y. & GUPTA, R. P. (1983) On approximating a linear combination of central Wishart matrices with positive coefficients, *Communications in Statistics, Theory and Methods*, 12, pp. 2589–2600.

TIMM, N. H. (1980) Multivariate analysis of variance of repeated measurements. *In:* P. R. KRISHNAIAH (ed), *Handbook of Statistics*, vol. 1 (pp. 41–87) (New York, North-Holland).

THOMAS, D. R. (1983) Univariate repeated measures techniques applied to multivariate data, *Psychometrika*, 48, pp. 451–464.

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

WINER, B. J. (1971) *Statistical Principles in Experimental Design* (2nd edn) (New York, McGraw-Hill).