NESTED ANALYSIS OF VARIANCE WITH AUTOCORRELATED ERRORS

Sastry G. Pantula and Kenneth H. Pollock

Department of Statistics, North Carolina State University
Box 8203, Raleigh, NC 27695-8203, U.S.A.

SUMMARY

In this paper we consider the problem where there is a randomized experimental design with several successive time measurements on each experimental unit. One approach to the analysis of such data is to treat time as the subplot treatment and use a split-plot analysis of variance. Alternatively, the problem may be considered in a more general multivariate framework. Here we recognize the time induced correlations and apply an autoregressive time series modelling approach. Estimation and testing are addressed. Two examples are presented to illustrate the practicality of our procedure. Some extensions are also considered briefly.

Key words: Time Series; Split-plot designs; Repeated measurement designs; Autoregressive processes; Multivariate analysis of variance.

1. Introduction

Frequently data arise from a randomized experimental design where several successive measurements over time are made on each experimental unit. There are many examples such as:

(i) In medical trials observations may be taken on each subject at regular time intervals,

(ii) In wildlife radio-telemetry study the distance moved by each animal may be tracked during successive periods of a day,

(iii) In plant (animal) experiments observations may be taken on the same plant (animal) as it grows and changes over time.

In all cases, there will be a correlation between observations taken on the same experimental unit at different time points. It is also likely that there is a decay in the correlation with increasing time lags. In other words, observations on the same experimental unit which are close together should be more highly correlated than those that are far apart in time. This is contrary to the assumption of constant correlation between all time points (subplots) within a unit (whole plot) made by the traditional split-plot analysis. An alternative approach is to consider a general multivariate analysis of variance (MANOVA) model, which assumes an arbitrary covariance matrix for the observations. (See Timm (1980)). However, we feel that the multivariate model is overparameterized and also it cannot easily incorporate the missing values. Therefore, we consider a time series modelling approach to the problem which serves as a compromise.

We now present the time series model that we consider for analyzing such data. We denote the variable under study by the letter y with two subscripts, where the first subscript refers to the individual experimental units and the second subscript refers to the successive measurements for the particular individual. We assume that n "individuals" are selected at random and that "measurements" are taken at t_i consecutive time points on the ith individual. We consider a linear model,

$$Y_{ij} = \sum_{k=1}^{r} G_{ijk}\theta_k + a_{ij}, \tag{1.1}$$

and

$$a_{ij} = v_i + \eta_{ij} \quad , \quad j = 1, 2, \ldots, t_i;\ i = 1, 2, \ldots, n, \tag{1.2}$$

where

Y_{ij} denotes the value of the jth measurement for the ith individual;

G_{ijk}, k = 1, 2, ..., r, denote the levels of the r control variables at which the observation Y_{ij} is obtained (the G_{ijk} are assumed to be fixed).

θ_k, k = 1, 2, ..., r, denote the unknown parameters to be estimated;

and

a_{ij}, the random error associated with Y_{ij}, is assumed to be the sum of the random effect associated with the ith sampled individual (v_i) and the random effect associated with the jth measurement for the ith individual in the sample (η_{ij}).

We assume that the random errors v_i are i.i.d. N $(0, \sigma_v^2)$ and that

$$\eta_{ij} = \alpha\eta_{i,j-1} + e_{ij}, \quad |\alpha|<1, \ j = 2, 3, \ldots, t_i \tag{1.3}$$

where $\{e_{ij}\}$ are i.i.d. N $(0, \sigma_e^2)$ and the initial value is sampled from N $(0, \sigma_e^2/(1 - \alpha^2))$. This is an autoregressive process of order 1 (AR(1)) (Fuller (1976) p. 36). We also assume that $\{\eta_{ij}, j = 1, 2, \ldots, t_i\}$ is independent of $\{\eta_{\ell j}, j = 1, 2, \ldots, t_i\}$ if $i \neq \ell$.

For the case $\alpha = 0$, which means that there is no autocorrelation, Fuller and Battese (1973) proposed a transformation to obtain an estimated generalized least squares (EGLS) estimator for $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_r)'$. Andersen et al. (1981) and Azzalini (1984) considered the case where α is not necessarily zero. Andersen et al. (1981) considered the effect of α in a two-way analysis of variance table constructed under the assumption that $\alpha = 0$. Azzalini (1984) assumed that $t_i = T$ for all i, and that $\underset{\sim}{G}_{ij} = (G_{ij1}, G_{ij2}, \ldots, G_{ijr}) = (\underset{\sim}{X}_i, \underset{\sim}{A}_j)$ where $\underset{\sim}{X}_i$ are some covariables observed on the ith individual (that do not depend on time j) and

$$A_{jk} = \begin{cases} 1 & \text{if } j = k < T \\ -(1-\alpha) & \text{if } j = T,\ k \neq 1 \\ -1 & \text{if } j = T,\ k = 1 \\ 0 & \text{otherwise} \end{cases} \quad , j = 1, 2, \ldots, T;\ k = 1, 2, \ldots, T-1, \qquad (1.4)$$

corresponds to the time dummy variable. (Note that $\underset{\sim}{G}_{ij}$ here depends on α.) Azzalini (1984) obtained algebraic expressions for the maximum likelihood (ML) estimators of $\underset{\sim}{\theta}$, σ_v^2 and σ_e^2 as functions of α and suggested that an estimator of α may be obtained by iteratively maximizing the likelihood.

In the present paper we follow the approach of Fuller and Battese (1973) to obtain EGLS estimates of $\underset{\sim}{\theta}$, and consistent estimators of σ_e^2, σ_v^2 and α. For the balanced case considered by Azzalini (1984), our estimators of $\underset{\sim}{\theta}$, σ_e^2 and σ_v^2 coincide with the "maximum likelihood estimators" (obtained as a function of α). We follow the approach of Andersen et al. (1981) to obtain a method of moments estimator of α.

First we present the estimation procedure and the asymptotic properties of the estimators in Section 2. This is followed in Section 3 by two examples to illustrate the practicality of the time series analysis and compare it to the split-plot analysis and the multivariate analysis. We conclude the paper with a discussion on mixed models and some extensions to two-fold nested error models and split block models.

2. Estimation

2.1 Background

We write the linear model (1.1) as

$$\underset{\sim}{Y} = \underset{\sim}{G}\,\underset{\sim}{\theta} + \underset{\sim}{a} \quad , \tag{2.1}$$

where $\underset{\sim}{Y} = (\underset{\sim}{Y}_1', \underset{\sim}{Y}_2', \ldots, \underset{\sim}{Y}_n')'$, $\underset{\sim}{Y}_i' = (Y_{i1}, Y_{i2}, \ldots, Y_{it_i})'$ and $\underset{\sim}{G}$ and $\underset{\sim}{a}$ are constructed similarly. Let $\underset{\sim}{V}_a$ denote the variance covariance matrix of $\underset{\sim}{a}$. Then,

$$\underset{\sim}{V}_a = \text{Block diagonal } \{\sigma_e^2 \underset{\sim}{\Gamma}_1 + \sigma_v^2 \underset{\sim}{J}_{t_1}, \ldots, \sigma_e^2 \underset{\sim}{\Gamma}_n + \sigma_v^2 \underset{\sim}{J}_{t_n}\} \tag{2.2}$$

where

$$\underset{\sim}{\Gamma}_i = (1 - \alpha^2)^{-1} \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{t_i-1} \\ \alpha & 1 & \alpha & \cdots & \alpha^{t_i-2} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \alpha^{t_i-1} & \alpha^{t_i-2} & \alpha^{t_i-3} & \cdots & 1 \end{bmatrix} : t_i \times t_i,$$

$$\underset{\sim}{J}_{t_i} = \underset{\sim}{1}_{t_i} \underset{\sim}{1}_{t_i}' \quad : \quad t_i \times t_i \; ,$$

and

$$\underset{\sim}{1}_{t_i} = (1, 1, \ldots, 1)' \; : \; t_i \times 1.$$

Note that $\underset{\sim}{V}_a$ depends on $\underset{\sim}{\delta} = (\sigma_e^2, \sigma_v^2, \alpha)'$ and is a square matrix of order $m = \sum_{i=1}^{n} t_i$. It is well known that the generalized least squares (GLS) estimator of $\underset{\sim}{\theta}$ is given by

$$\hat{\underset{\sim}{\theta}}_G = \hat{\underset{\sim}{\theta}}_G(\underset{\sim}{\delta}) = (\underset{\sim}{G}'\underset{\sim}{V}_a^{-1}\underset{\sim}{G})^{-1} \underset{\sim}{G}'\underset{\sim}{V}_a^{-1}\underset{\sim}{Y} . \tag{2.3}$$

First we present a simple way of computing the estimates $(\hat{\underset{\sim}{\theta}}_G)$ given in equation (2.3) assuming that the autoregressive parameter (α) and the variance components (σ_e^2 and σ_v^2) are known. Next, for known α we obtain estimators for the variance components that are unbiased and consistent. We then obtain a consistent estimator $\hat{\alpha}$ of α for the situation where the number of time points is bounded (i.e., $t_i \le T < \infty$ for some T), but the number of subjects (n) is large. Finally, under the same framework, we show that $\hat{\sigma}_e^2(\alpha)$, $\hat{\sigma}_v^2(\alpha)$ and $\hat{\underset{\sim}{\theta}}_G(\underset{\sim}{\delta})$ are asymptotically equivalent to $\hat{\sigma}_e^2(\hat{\alpha})$, $\hat{\sigma}_v^2(\hat{\alpha})$ and $\hat{\underset{\sim}{\theta}}_G(\hat{\underset{\sim}{\delta}})$. We also give a summary of the computational procedure which turns out to only involve simple transformations and regressions.

2.2 Computation of the GLS estimator of $\underset{\sim}{\theta}$

Note that the computation of the GLS estimator, using equation (2.3), involves inverting the matrix $\underset{\sim}{V}_a$ which typically is of large dimension. In this section we show that the generalized least squares problem can be transformed into an ordinary least squares problem which can then be easily computed using standard packages (e.g., SAS (1982) GLM procedure). Here we apply the transformation $\sigma_e \underset{\sim}{V}_a^{-1/2}$ to both $\underset{\sim}{Y}$ and $\underset{\sim}{G}$, following an approach similar to that of Fuller and Battese (1973). In fact, if $\alpha = 0$ then our transformation reduces to their result. In the following theorem we obtain a simple expression for $\underset{\sim}{V}_a^{-1/2}$.

<u>Theorem 1</u>: Consider the linear model (1.1) - (1.2). The GLS estimator $\hat{\underset{\sim}{\theta}}_G$ of $\underset{\sim}{\theta}$ may be obtained by regressing $Z_{ij}^{(1)} + \lambda_i Z_{ij}^{(2)}$ on $\underset{\sim}{H}_{ij}^{(1)} + \lambda_i \underset{\sim}{H}_{ij}^{(2)}$ using ordinary least squares, where

$$Z_{ij}^{(1)} = Z_{ij} - Z_{ij}^{(2)} ,$$

$$Z_{ij}^{(2)} = c_i^{-1} f_j d_i ,$$

$$d_i = \sum_{j=1}^{t_i} f_j Z_{ij},$$

$$c_i = \sum_{j=1}^{t_i} f_j^2 = (1-\alpha) [t_i - (t_i-2)\alpha] ,$$

$$\lambda_i = [(\sigma_e^2 + c_i \sigma_v^2)^{-1} \sigma_e^2]^{1/2} ,$$

$$f_j = \begin{cases} (1-\alpha^2)^{1/2} & , j = 1 \\ (1-\alpha) & , j \geq 2 \end{cases} ,$$

$$Z_{ij} = \begin{cases} f_1 Y_{i1} & , j = 1 \\ Y_{ij} - \alpha Y_{i,j-1} & , j \geq 2 \end{cases} , \qquad (2.4)$$

and $\underset{\sim}{H}_{ij}$, $\underset{\sim}{H}_{ij}^{(1)}$, $\underset{\sim}{H}_{ij}^{(2)}$ are defined similarly in terms of $\underset{\sim}{G}_{ij}$.

<u>Proof</u>: See Appendix.

2.3 Estimation of variance components

Here we present unbiased estimators of the variance components σ_e^2 and σ_v^2 assuming that α is known. Note that

$$\underset{\sim}{Z}_i^{(1)} \sim \text{NID } (\underset{\sim}{H}_i^{(1)} \underset{\sim}{\theta}, \ \sigma_e^2 \ (\underset{\sim}{I}_{t_i} - \underset{\sim}{D}_i))$$

and

$$c_i^{-1/2} \underset{\sim}{f}'_{(i)} \underset{\sim}{Z}_i \sim \text{NID} \, (c_i^{-1/2} \underset{\sim}{f}'_{(i)} \underset{\sim}{H}_i \, , \, \sigma_e^2 + c_i \sigma_v^2)$$

and that $\underset{\sim}{Z}_i^{(1)}$ is independent of $c_i^{-1/2} \underset{\sim}{f}'_{(i)} \underset{\sim}{Z}_i$. Therefore,

$$\hat{\sigma}_e^2 = \hat{\sigma}_e^2 \, (\alpha) = \hat{\underset{\sim}{e}}' \hat{\underset{\sim}{e}} / [\sum_{i=1}^{n} t_i - n - \text{rank} \, (\underset{\sim}{H}^{(1)})] \quad , \tag{2.5}$$

and

$$\hat{\sigma}_v^2 = \hat{\sigma}_v^2 \, (\alpha) = \frac{\hat{\underset{\sim}{u}}' \hat{\underset{\sim}{u}} - \hat{\sigma}_e^2 \, [n - \text{rank} \, (\underset{\sim}{F})]}{[\sum_{i=1}^{n} c_i - \text{tr} \, \{\underset{\sim}{F}' \underset{\sim}{C}^{-1} \underset{\sim}{F})^+ \, \underset{\sim}{F}' \underset{\sim}{F}\}]} \quad , \tag{2.6}$$

where $\hat{\underset{\sim}{e}}' \hat{\underset{\sim}{e}}$ denotes the residual sums of squares obtained by regressing $\underset{\sim}{Z}_i^{(1)}$ on $\underset{\sim}{H}_i^{(1)}$; $\hat{\underset{\sim}{u}}' \hat{\underset{\sim}{u}}$ denotes residual sums of squares obtained by regressing $c_i^{-1/2} \underset{\sim}{f}'_{(i)} \underset{\sim}{Z}_i$ on $c_i^{-1/2} \underset{\sim}{f}'_{(i)} \underset{\sim}{H}_i$;

$$\underset{\sim}{F} = \begin{bmatrix} \underset{\sim}{f}'_{(1)} \, \underset{\sim}{H}_1 \\ \underset{\sim}{f}'_{(2)} \, \underset{\sim}{H}_2 \\ \vdots \\ f'_{(n)} \, \underset{\sim}{H}_n \end{bmatrix} \quad ; \quad \underset{\sim}{C} = \text{diagonal} \, (c_1, c_2, \ldots, c_n);$$

$\underset{\sim}{A}^+$ denotes the Moore Penrose inverse of the matrix $\underset{\sim}{A}$ and

tr $(\underset{\sim}{A})$ denotes the trace of the matrix $\underset{\sim}{A}$.

The estimators (2.5) and (2.6) have variances given by,

$$\text{Var} \, (\hat{\sigma}_e^2) = 2\sigma_e^4 / [\sum_{i=1}^{n} t_i - n - \text{rank} \, (\underset{\sim}{H}^{(1)})] \tag{2.7}$$

and

$$\text{Var}\ (\hat{\sigma}_v^2) = \frac{\text{Var}\ (\hat{\underset{\sim}{u}}'\hat{\underset{\sim}{u}}) + [\text{n-rank}\ (\underset{\sim}{F})]^2\ \text{Var}\ (\hat{\sigma}_e^2)}{[\sum_{i=1}^{n} c_i - \text{tr}\ \{(\underset{\sim}{F}'\underset{\sim}{C}^{-1}\underset{\sim}{F})^+\ \underset{\sim}{F}'\underset{\sim}{F}\}]^2} \quad , \qquad (2.8)$$

where

$$\text{Var}\ (\hat{\underset{\sim}{u}}'\hat{\underset{\sim}{u}}) = 2\sigma_e^4\ [\text{n-rank}\ (\underset{\sim}{F})] + 4\sigma_e^2\sigma_v^2\ [\sum_{i=1}^{n} c_i - \text{tr}\ \{(\underset{\sim}{F}'\underset{\sim}{C}^{-1}\underset{\sim}{F})^+\ \underset{\sim}{F}'\underset{\sim}{F}\}]$$

$$+ 2\sigma_v^4\ [\sum_{i=1}^{n} c_i^2 - 2\ \text{tr}\{(\underset{\sim}{F}'\underset{\sim}{C}^{-1}\underset{\sim}{F})^+\ \underset{\sim}{F}'\underset{\sim}{C}\underset{\sim}{F}\} + \text{tr}\ (\{(\underset{\sim}{F}'\underset{\sim}{C}^{-1}\underset{\sim}{F})^+\ \underset{\sim}{F}'\underset{\sim}{F}\}^2)]\ .$$

If $\alpha = 0$ then our estimator of σ_e^2 coincides with that of Fuller and Battese (1973). However, our estimator of σ_v^2 is different from theirs. Note also that, for a fixed α, our model is equivalent to

$$\underset{\sim}{Z}_i(\alpha) = \underset{\sim}{H}_i(\alpha)\ \underset{\sim}{\theta} + \underset{\sim}{f}_{(i)}(\alpha)\ v_i + \underset{\sim}{e}_i \quad ,$$

which is a special case of a mixed model. For mixed models there are a variety of procedures available (see Giesbrecht (1983)) such as MINQUE and MIVQUE. It is beyond the scope of this article to do a detailed comparison of different estimators for σ_e^2 and σ_v^2. We chose the estimators (2.5) and (2.6) since they are simple to compute. Also, a simple test of $\sigma_v^2 = 0$ is possible because $\hat{\underset{\sim}{u}}'\hat{\underset{\sim}{u}}$ and $\hat{\sigma}_e^2$ are independent. We therefore can show that

$$\frac{\hat{\underset{\sim}{u}}'\hat{\underset{\sim}{u}}/\nu_1}{\hat{\sigma}_e^2} \sim F_{\nu_1,\ \nu_2} \qquad (2.9)$$

where $\nu_1 = [\text{n-rank}\ (\underset{\sim}{F})]$ and $\nu_2 = [\sum_{i=1}^{n} t_i - \text{n-rank}\ (\underset{\sim}{H}^{(1)})]$

with F denoting Fisher's F-distribution.

As we remarked in section 1, the model considered by Azzalini (1984) is a special case of our model (2.1) with $\underset{\sim}{G}_{ij} = (\underset{\sim}{X}_i, \underset{\sim}{A}_j)$, $\underset{\sim}{\theta} = (\underset{\sim}{\beta}', \underset{\sim}{\xi}')'$ and $t_i = T$ where $\underset{\sim}{A}_j$ is given in (1.4). For a fixed α, the likelihood estimators of $\underset{\sim}{\theta}$, σ_e^2 and σ_v^2 are obtained by maximizing,

$$\begin{aligned} 2\log(\text{likelihood}) = &- n(T-1)\log\sigma_e^2 - n\log(\sigma_e^2 + c_1\sigma_v^2) \\ &- \sigma_e^{-2}[\underset{\sim}{Z}^{(1)} - (\underset{\sim}{1} \otimes \underset{\sim}{B})\underset{\sim}{\xi}]'[\underset{\sim}{Z}^{(1)} - (\underset{\sim}{1} \otimes \underset{\sim}{B})\underset{\sim}{\xi}] \\ &- (\sigma_e^2 + c_1\sigma_v^2)^{-1}[\underset{\sim}{Z}^{(2)} - (\underset{\sim}{X}\otimes\underset{\sim}{f}_{(1)})\underset{\sim}{\beta}]'[\underset{\sim}{Z}^{(2)} - (\underset{\sim}{X}\otimes\underset{\sim}{f}_{(1)})\underset{\sim}{\beta}] \end{aligned}$$

where $\underset{\sim}{A} = (\underset{\sim}{A}_1', \underset{\sim}{A}_2', \ldots, \underset{\sim}{A}_T')'$, $\underset{\sim}{B} = \underset{\sim}{P}_1\underset{\sim}{A}$, $\underset{\sim}{P}_1$ is defined in (A.1) and $\underset{\sim}{C}_1 \otimes \underset{\sim}{C}_2$ denotes the Kronecker product of the matrices $\underset{\sim}{C}_1$ and $\underset{\sim}{C}_2$. Therefore, the ML estimator of $\underset{\sim}{\xi}$ may be obtained by regressing $\underset{\sim}{Z}^{(1)}$ on $(\underset{\sim}{1}\otimes\underset{\sim}{B})$, and the residual mean square error of this regression is the ML estimator of σ_e^2. Similarly, the ML estimator of $\underset{\sim}{\beta}$ may be obtained by regressing $\underset{\sim}{Z}^{(2)}$ on $(\underset{\sim}{X}\otimes\underset{\sim}{f}_{(1)})$, and the residual mean square error of this regression is the ML estimator of $\sigma_e^2 + c_1\sigma_v^2$. Also, the ANOVA table presented by Azzalini (1984) may be obtained by the above two regressions.

2.4 Estimation of α.

Using the properties of an autoregressive process of order 1 (Fuller (1976)), it is easy to show that for integers i, j and s,

$$E[a_{ij}a_{i,j+s}] = \sigma_v^2 + \sigma_\eta^2\alpha^s \tag{2.10}$$

where $\sigma_\eta^2 = (1-\alpha^2)^{-1}\sigma_e^2$. Assume $t_i > 2$ and let

$$N_1 = (m-2n)^{-1} \sum_{i=1}^{n}\sum_{j=1}^{t_i-2} a_{ij}(a_{ij} - a_{i,j+1}),$$

and

$$N_2 = (m-2n)^{-1} \sum_{i=1}^{n}\sum_{j=1}^{t_i-2} a_{ij}(a_{i,j+1} - a_{i,j+2})$$

where $m = \sum_{i=1}^{n} t_i$. Then,

$$E[N_1] = \sigma_\eta^2 (1 - \alpha)$$

and

$$E[N_2] = \alpha \sigma_\eta^2 (1 - \alpha).$$

Therefore, a ratio estimator for α is given by

$$\tilde{\alpha} = N_2/N_1 . \tag{2.11}$$

Under some regularity conditions, it can be shown that the estimator $\tilde{\alpha}$ is consistent and asymptotically normal for bounded t_i and n tending to infinity.

Note, however, that $\tilde{\alpha}$ depends on the errors a_{ij} which are not observable. We denote $\hat{\alpha}$ as the estimator of α obtained by replacing $\{a_{ij}\}$ by its estimates $\{\tilde{a}_{ij}\}$ in (2.11) where $\{\tilde{a}_{ij}\}$ are the ordinary least squares residuals obtained by regressing $\underset{\sim}{Y}$ on $\underset{\sim}{G}$. Therefore, we have

$$\hat{\alpha} = \tilde{N}_2/\tilde{N}_1 \tag{2.12}$$

where $\tilde{N}_1$ and $\tilde{N}_2$ are estimators of N_1 and N_2 obtained using $\tilde{a}_{ij}$ in place of a_{ij}. The large sample variance of $\hat{\alpha}$ is given by

$$\mathrm{Var}(\hat{\alpha}) = \frac{2(1+\alpha)}{(m-2n)} + \frac{2n\sigma_v^2(1+\alpha)^2}{\sigma_e^2(m-2n)^2} + \frac{2n\alpha(1+\alpha)}{(1-\alpha)(m-2n)^2} + o\left(\frac{1}{n}\right) \tag{2.13}$$

and our estimator $\hat{\alpha}$ can be shown to be asymptotically equivalent to $\tilde{\alpha}$, under some regularity conditions.

Our estimator of α is similar to the estimator proposed by Andersen et al. (1983). For fixed n and t_i tending to infinity, Andersen et al. (1983) proposed a consistent estimator for α. Note that we have used the equation (2.10) only for s = 0, 1, 2. An alternative approach is to treat,

$$a_{ij}a_{i,j+s} = \sigma_v^2 + \sigma_e^2 (1-\alpha^2)^{-1} \alpha^s + \text{error},$$

as a nonlinear regression model, and obtain estimators for σ_v^2, σ_e^2 and α

simultaneously. This approach is very useful especially when some of the intermediate observations are missing, and will be considered elsewhere.

2.5 Summary of the computational procedure

The first step is to compute the ordinary least squares estimator of $\underset{\sim}{\theta}$ by regressing $\underset{\sim}{Y}$ on $\underset{\sim}{G}$ and use the least squares residuals to compute $\hat{\alpha}$ given in (2.12). Next we compute the estimates $\hat{\sigma}_e^2(\hat{\alpha})$ and $\hat{\sigma}_v^2(\hat{\alpha})$ from (2.5) and (2.6), which involve ordinary least squares regressions of $\underset{\sim}{Z}_i^{(1)}(\hat{\alpha})$ on $\underset{\sim}{H}_i^{(1)}(\hat{\alpha})$ and $c_i^{-1/2}\,\underset{\sim}{f}'_{(i)}\,Z_i(\hat{\alpha})$ on $c_i^{-1/2}\,\underset{\sim}{f}'_{(i)}\,\underset{\sim}{H}_i(\hat{\alpha})$. Finally, the EGLS estimate $\hat{\underset{\sim}{\theta}}_{EG}(\hat{\alpha})$ of $\underset{\sim}{\theta}$ is obtained by regressing $\underset{\sim}{Z}_i^{(1)}(\hat{\alpha}) + \lambda_i(\hat{\alpha})\,\underset{\sim}{Z}_i^{(2)}(\hat{\alpha})$ on $\underset{\sim}{H}_i^{(1)}(\hat{\alpha}) + \lambda_i(\hat{\alpha})\,\underset{\sim}{H}_i^{(2)}(\hat{\alpha})$. It is important to emphasize that the steps only involve simple regressions and simple transformations of the data. The transformations were chosen so that we do not have to invert $\underset{\sim}{V}_a$ which is an m x m matrix.

Using an approach similar to that of Fuller and Battese (1973), it can be shown that the estimators $\hat{\sigma}_e^2(\hat{\alpha})$, $\hat{\sigma}_v^2(\hat{\alpha})$ and $\hat{\underset{\sim}{\theta}}_G(\hat{\alpha})$ are asymptotically equivalent to $\hat{\sigma}_e^2(\alpha)$, $\hat{\sigma}_v^2(\alpha)$ and $\hat{\underset{\sim}{\theta}}_G(\alpha)$. Also, the transformations and the estimator considered in sections 2.1-2.4 can be extended easily to the case where some of the intermediate observations are missing on some of the individuals. The details, however, are not presented here. A program using the SAS (1982) MATRIX procedure that computes the estimates is available from the senior author.

3. Examples

3.1 Plant growth example

Here we consider data from a plant growth experiment conducted at North Carolina State University on cucumbers. The basic treatment design was a 2 x 3 factorial with the first factor (NUT) involving 2 levels of of Hoagland's solution and second factor (CON) involving 3 levels of Ferulic acid. Each treatment was applied to 9 seeds in a completely randomized design. After germination, the total leaf area was measured every two days for 14 days for a total of 7 measurements on each plant. The first measurements were discarded because they were very small, leaving a total of 6 measurements on each plant. Following the general approach of Box and Cox (1964), we transformed our data to the fourth root of (leaf area + 1) to achieve approximate normality and equal variances. Let

$Y_{\ell t}$ = (leaf area + 1)$^{1/4}$ of the ℓth plant on (2t+2) day

$\ell = 1, 2, \ldots, 54;$

$t = 1, 2, \ldots, 6;$

and

Y^*_{ijk} = (kth measurement of leaf area + 1)$^{1/4}$ of the jth plant treated with the ith treatment,

$i = 1, 2, \ldots, 6;\quad j = 1, 2, \ldots, 9;\quad k = 1, 2, \ldots, 6;$

A model for Y^*_{ijk} is

$$Y^*_{ijk} = G_1\mu + \sum_{s=1}^{5} G_{is}\beta_s + \sum_{t=1}^{5} G^*_{k,t}\xi_t + \sum_{s=1}^{5}\sum_{t=1}^{5} G_{is}G^*_{kt}\gamma_{st} + v_{ij} + \eta_{ijk}, \qquad (4.1)$$

where $G_1 \equiv 1$; G_{i1} = (NUT=2)-(NUT=1); G_{i2}= (CON=1)-(CON=3); G_{i3} = (CON=2)-(CON=3); $G_{i4} = G_{i1} \cdot G_{i2}$; $G_{i5} = G_{i1} \cdot G_{i3}$; G^*_{kt} = (k=t) - (k=6), t = 1, 2, ..., 5; and (a=b) is the Kronecker function which takes the value 1 if a=b and 0 otherwise. We assume that $v_{ij} \sim \text{NID}(0,\sigma_v^2)$ and that $\eta_{ijk} = \alpha\, \eta_{ij,\,k-1} + e_{ijk}$ where $|\alpha|<1$ and $e_{ijk} \sim \text{NID}(0,\sigma_e^2)$. Note that the β parameters correspond to the treatment contrasts, the ξ parameters correspond to the time contrasts and the γ parameters correspond to the treatment * time contrasts.

In Table 1, we present the usual analysis of variance (α=0) table including the sums of squares for NUT and CON and their interaction with each other and with time. The estimate of α obtained using equation (2.12) is $\hat{\alpha} = 0.4067$ with an estimated standard error 0.2039. In Table 2 we present some selected contrast estimates with their standard errors computed under both assumptions $\alpha = 0$ and $\alpha = 0.4067$. The estimates are the same but the estimated standard errors are different. In Table 1, we also include the analysis of variance table suggested by Azzalini (1984) for $\alpha = 0.4067$. This corresponds to replacing $G^{*}_{k,t}$ in model (4.1) by $G^{**}_{k,t}$ where

$$G^{**}_{k,t} = \begin{cases} -(1-\alpha) & \text{if} \quad k = 6,\ t = 1 \\ -1 & \text{if} \quad k = 6,\ t = 1 \\ 1 & \text{if} \quad k = t \\ 0 & \text{otherwise} \quad , \quad t = 1, 2, \ldots, 5; \\ & \qquad\qquad\qquad k = 1, 2, \ldots, 6. \end{cases}$$

The above definition of $G^{**}_{k,t}$ is chosen so that the different partitioned sums of squares computed will be independent of each other if α is known. Notice however for our data the substantial conclusions remain the same. For example, NUT*CON interaction is not significant under either analysis. If $\alpha = 0$, then the autocorrelation between any two measurements on the same plant is constant and is given by $(\sigma_e^2 + \sigma_v^2)^{-1}\sigma_v^2 = 0.825$. For our model, the correlation between Y_{ijk} and Y_{ijs} is $(\sigma_\eta^2 + \sigma_v^2)^{-1}(\sigma_v^2 + \alpha^{|k-s|}\sigma_\eta^2)$ and for our example the estimated correlations of Y_{ij1} with $Y_{ij2}, \ldots, Y_{ij6}$ are 0.892, 0.849, 0.835, 0.826 and 0.821, respectively. The multivariate model assumes an arbitrary structure for the covariances of $(Y_{ij1}, Y_{ij2}, \ldots, Y_{ij6})$. For our example, the estimated correlations, using the multivariate model, of Y_{ij1} with $Y_{ij2}, \ldots, Y_{ij6}$ are 0.904, 0.813, 0.765, 0.712 and 0.664, respectively and that of Y_{ij2} with $Y_{ij3}, \ldots, Y_{ij6}$ are 0.914, 0.850, 0.848 and 0.807, respectively.

3.2 Bobcat Telemetry Example

A bobcat is caught, radiotagged and released back into the forest. The location of the bobcat is recorded every hour over 16 different days. Using the two consecutive locations of the bobcat, the distance and the rate at which the animal is moving are estimated. Missing values are generated at times because (i) the signal from the animal is weak, (ii) the animal could not be located or (iii) the investigator didn't take a reading.

A purpose of the study is to model the rate at which the animal travels during different time periods of a day. Let Y_{ij} denote the log of the rate at which the animal traveled during the jth hour of the ith day. The days are sufficiently apart so that we may assume that the observations taken on different days are independent. We consider the model (1.1)-(1.3) for Y_{ij} with $E[Y_{ij}] = \theta_1 + \theta_2 \text{ Cos } (\frac{\pi j}{12}) + \theta_3 \text{ Sin } (\frac{\pi j}{12}) + \theta_4 \text{ Cos } (\frac{\pi j}{6}) + \theta_5 \text{ Sin } (\frac{\pi j}{6})$, a trignometric function allowing up to two complete cycles in 24 hours. In Table 3 we present a comparison of parameter estimates and the standard error estimates for $\alpha = 0$ and estimated $\alpha = 0.48$ (S.E. = 0.24). Notice that there are substantial changes in some of the estimates (θ_5 and σ_v^2).

4. Discussion

For data that consist of measurements observed on each individual at several consecutive time points we considered a time series model given by (1.1)-(1.3). The main steps of our estimation procedure are

(i) obtain a consistent estimator $\hat{\alpha}$ of α ,

(ii) write the model (1.1) equivalently as

$$\underset{\sim}{Z}_i(\alpha) = \underset{\sim}{H}_i(\alpha)\ \underset{\sim}{\theta} + \underset{\sim}{f}_{(i)}(\alpha)\ v_i + e_i \qquad (4.1)$$

where $\underset{\sim}{Z}_i$, $\underset{\sim}{H}_i$ and $\underset{\sim}{f}_{(i)}$ are defined in section 2,

and

(iii) extend the results of Fuller and Battese (1973) to obtain consistent estimatorsof the parameters $\underset{\sim}{\theta}$ σ_e^2 and σ_v^2.

An alternative approach is to treat the equation (4.1) as a special case of general mixed-models and use any of the existing estimation procedures that are available for mixed-models. Giesbrecht (1983) provides an efficient program (SAS procedure MIXMOD) for computing the variance component estimates in a general mixed model. Also, since the procedure MIXMOD is written for general mixed models, it will be useful for computing the variance component estimates in nested-error models with autocovariance. Now we present an example where we have a two-fold nested-error model (split-split-plot).

Suppose n locations are selected at random from a collection of locations. From each of n locations, b plots each are selected at random. Different levels of fertilizer are applied to the plots, at the beginning of the experiment. Yield from each of the plots is observed on t successive years. Let Y_{ijk} denote the yield in the kth year from the jth plot of the ith location. The yield, a function of the composition of the soil in each plot, the rainfall at each location, dose of the fertilizer used, etc., may be modeled by

$$Y_{ijk} = \underset{\sim}{G}_{ijk} \underset{\sim}{\theta} + v_i + e_{ij} + \eta_{ijk} , \qquad (4.2)$$

$$i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, b;\ k = 1, 2, \ldots, t;$$

where we assume

$$\eta_{ijk} = \alpha\ \eta_{ij,k-1} + \varepsilon_{ijk} , \qquad |\alpha|<1$$

$$v_i \sim \text{NID}(0, \sigma_v^2) ,\ e_{ij} \sim \text{NID}(0, \sigma_e^2),\ \varepsilon_{ijk} \sim \text{NID}(0, \sigma_\varepsilon^2)$$

and v_i, e_{ij} and ε_{ijk} are independent.

The case $\alpha = 0$ is considered by Fuller and Battese (1973).

Note that (4.2) may also be written as,

$$\underset{\sim}{Z}_{ij}(\alpha) = \underset{\sim}{H}_{ij}(\alpha)\ \underset{\sim}{\theta} + \underset{\sim}{f}_{(i)}(\alpha)\ v_i + \underset{\sim}{f}_{(i)}(\alpha)\ e_{ij} + \varepsilon_{ijk} \qquad (4.3)$$

where $\underset{\sim}{Z}_{ij} = \underset{\sim}{P}_i \underset{\sim}{Y}_{ij}$, $\underset{\sim}{H}_{ij} = \underset{\sim}{P}_i \underset{\sim}{G}_{ij}$ and $\underset{\sim}{P}_i$ and $\underset{\sim}{f}_{(i)}$ are as defined in the Appendix.

First fix α. We can extend again the results of Fuller and Battese (1973) to obtain simple consistent estimators of $\underset{\sim}{\theta}$, σ_e^2, σ_v^2 and σ_ε^2. Also, as in section 2 we can find a consistent estimator $\hat{\alpha}$ of α and use $\hat{\alpha}$ in place of α in (4.3) to obtain the estimates of the parameters $\underset{\sim}{\theta}$, σ_e^2, σ_v^2 and σ_ε^2. Another approach is to consider $\underset{\sim}{Z}_{ij}(\hat{\alpha})$ in (4.3) as a special case of mixed models and use the MIXMOD procedure.

Finally we also mention that other extensions are possible. For example a traditional split-block analysis (Steel and Torrie 1980, p. 390) could be extended in a similar fashion to that developed in section 2.

ACKNOWLEDGEMENTS

We wish to thank Professor Richard Lancia for providing us with the bobcat radio-telemetry data and for some helpful observations. We also wish to thank Professor Udo Blum for allowing us to use his cucumber plants data. We thank Professor Wayne Fuller for his comments.

REFERENCES

Andersen, A. H., Jensen, E. B., and Geart, S. (1981). Two-way analysis of variance with correlated errors. International Statistical Review 49, 153-169.

Azzalini, A. (1984). Estimation and hypothesis testing for collections of autoregressive time series. Biometrika 71, 85-90.

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society Series B 26, 211-252.

Fuller, W. A. (1976). Introduction to Statistical Time Series. Wiley, New York.

Fuller, W. A. and Battese, G. E. (1973). Transformations for estimation of linear models with nested-error structure. Journal of the American Statistical Association 68, 626-632.

Giesbrecht, F. (1983). An efficient procedure for computing minque of variance components and generalized least squares estimates of fixed effects. Communications in Statistics-Theory and Methods 12(18):2169-2177.

SAS Institute, Inc. (1982). SAS User's Guide: Statistics, SAS Institute, Inc., Cary, North Carolina.

Steel, R. G. D. and Torrie, J. H. (1980). Principles and procedures of Statistics: A Biometrical Approach. McGraw Hill, New York.

Timm, M. H. (1980). Multivariate analysis of variance of repeated measures. In Handbook of Statistics, Vol 1, P. R. Krishnaiah (ed.), 41-87: North Holland, Amsterdam.

Table 1. Comparison of analysis of variance tables for $\alpha = 0$ and $\alpha = 0.4067$ (S.E. = 0.2039) for the transformed data.

Source	df	$\alpha = 0$		$\alpha = 0.4067$	
		MS	F	MS	F
Mean (μ)	1	1622		693	
Treatments (β)	5	7.80	72	3.41	230
NUT (β_1)	(1)	16.75	155	7.59	513
CON (β_2, β_3)	(2)	11.10	102	4.72	319
NUT*CON (β_4, β_5)	(2)	0.03	0.24	0.01	0.68
Error (a)	48	0.108	29	0.015	16
Time (ξ)	5	16.27	4412	10.60	3386
Time*Treatment (γ)	25	1.07	290	0.13	43
NUT*Time	(5)	0.82	222	0.52	165
CON*Time	(10)	0.10	28	0.07	21
NUT*CON*Time	(10)	0.02	5.69	0.01	4.3
Error (b)	240	0.0037		0.0031	
		$\hat{\sigma}_v^2 = 0.0174$ (S.E. = 0.0037)		$\sigma_v^2 = 0.0174$ (S.E. = 0.0038)	

Table 2. Comparison of selected parameter estimates and standard error estimates for $\alpha = 0$ and $\hat{\alpha} = 0.4067$ (S.E. = 0.2039) for the transformed data.

Parameter	Estimate	Standard errors	
		$\alpha = 0$	$\alpha = 0.4067$
Mean (μ)	2.2376	0.0183	0.0186
(NUT=2) vs (NUT=1): β_1	0.2274	0.0183	0.0186
(CON=1) vs (CON=3): β_2	-0.0407	0.0259	0.0263
(CON=2) vs (CON=3): β_3	-0.2983	0.0259	0.0263
(Time=1) vs (Time=6): ξ_1	-0.8094	0.0075	0.0073
(Time=2) vs (Time=6): ξ_2	-0.3928	0.0075	0.0067
(Time=3) vs (Time=6): ξ_3	-0.0958	0.0075	0.0064
(Time=4) vs (Time=6): ξ_4	0.1660	0.0075	0.0064
(Time=5) vs (Time=6): ξ_5	0.4675	0.0075	0.0067

Table 3. Comparison of parameter estimates and standard error estimates for $\alpha = 0$ and $\hat{\alpha} = 0.48$ (S.E. = 0.24) for the bobcat activity data.

Parameter	$\alpha = 0$		$\alpha = 0.48$	
	Estimate	Standard error	Estimate	Standard error
Mean (θ_1)	2.59	0.14	2.57	0.14
$\mathrm{Cos}(\frac{\pi j}{12})$: θ_2	0.04	0.15	0.07	0.19
$\mathrm{Sin}(\frac{\pi j}{12})$: θ_3	0.03	0.15	-0.10	0.20
$\mathrm{Cos}(\frac{\pi j}{6})$: θ_4	-0.07	0.14	-0.08	0.17
$\mathrm{Sin}(\frac{\pi j}{6})$: θ_5	-0.22	0.14	-0.14	0.18
σ_e^2	1.16	0.16	1.12	0.15
σ_v^2	0.14	0.13	0.00	0.13

APPENDIX

Here we present an outline of proof of Theorem 1.

Proof of Theorem 1: Note that $\underset{\sim}{\Gamma}_i = \underset{\sim}{P}_i^{-1} \underset{\sim}{P}_i^{-1'}$, where

$$\underset{\sim}{P}_i = \begin{bmatrix} (1-\alpha^2)^{1/2} & 0 & 0 & \cdots & 0 & 0 \\ -\alpha & 1 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & & & \\ 0 & 0 & 0 & \cdots & -\alpha & 1 \end{bmatrix} . \qquad (A.1)$$

Now,

$$\begin{aligned} \underset{\sim}{Z}_i &= \underset{\sim}{P}_i \underset{\sim}{Y}_i \\ &= \underset{\sim}{P}_i \underset{\sim}{G}_i \underset{\sim}{\theta} + \underset{\sim}{P}_i \underset{\sim}{a}_i \\ &= \underset{\sim}{H}_i \underset{\sim}{\theta} + \underset{\sim}{u}_i . \end{aligned}$$

Observe that

$$\begin{aligned} \text{Var}(\underset{\sim}{u}_i) &= \underset{\sim}{V}_i \\ &= \sigma_e^2 \underset{\sim}{I}_{t_i} + \sigma_v^2 \underset{\sim}{f}_{(i)} \underset{\sim}{f}'_{(i)} \\ &= \sigma_e^2 (\underset{\sim}{I}_{t_i} - \underset{\sim}{D}_i) + (\sigma_e^2 + c_i \sigma_v^2) \underset{\sim}{D}_i , \end{aligned}$$

where

$$\begin{aligned} \underset{\sim}{f}_{(i)} &= \underset{\sim}{P}_i \underset{\sim}{1}_{t_i} = (f_1, f_2, \ldots, f_{t_i})' , \\ \underset{\sim}{D}_i &= c_i^{-1} \underset{\sim}{f}_{(i)} \underset{\sim}{f}'_{(i)} \end{aligned}$$

and $\underset{\sim}{I}_{t_i}$ denotes the identity matrix of order t_i. Note that $(\underset{\sim}{I}_{t_i} - \underset{\sim}{D}_i)$ and $\underset{\sim}{D}_i$ are symmetric idempotent matrices that are mutually orthogonal. Therefore,

$$\sigma_e \underset{\sim}{V}_i^{-1/2} = (\underset{\sim}{I}_{t_i} - \underset{\sim}{D}_i) + \lambda_i \underset{\sim}{D}_i ,$$

where $\lambda_i = (\sigma_e^2 + c_i \sigma_v^2)$.

Now consider

$$\sigma_e \underset{\sim}{V}_i^{-1/2} \underset{\sim}{Z}_i = \underset{\sim}{Z}_i^{(1)} + \lambda_i \underset{\sim}{Z}_i^{(2)}$$

$$= [\underset{\sim}{H}_i^{(1)} + \lambda_i \underset{\sim}{H}_i^{(2)}] \underset{\sim}{\theta} + \sigma_e \underset{\sim}{V}_i^{-1/2} \underset{\sim}{u}_i ,$$

where

$$\underset{\sim}{Z}_i^{(1)} = (\underset{\sim}{I}_{t_i} - \underset{\sim}{D}_i) \underset{\sim}{Z}_i ,$$

$$\underset{\sim}{Z}_i^{(2)} = \underset{\sim}{D}_i \underset{\sim}{Z}_i ,$$

and $\underset{\sim}{H}_i^{(1)}$, $\underset{\sim}{H}_i^{(2)}$ are defined similarly. Note that

$$\text{Var} (\sigma_e \underset{\sim}{V}_i^{-1/2} \underset{\sim}{u}_i) = \sigma_e^2 \underset{\sim}{I}_{t_i} ,$$

and hence the GLS estimate of $\underset{\sim}{\theta}$ for model (1.1)-(1.3) may be obtained by regressing $(\underset{\sim}{Z}_i^{(1)} + \lambda_i \underset{\sim}{Z}_i^{(2)})$ on $(\underset{\sim}{H}_i^{(1)} + \lambda_i \underset{\sim}{H}_i^{(2)})$ using ordinary least squares.