PREDICTION IN GROWTH CURVE MODELS UNDER MARKOV COVARIANCE STRUCTURE

Dayanand N. Naik

Department of Mathematics and Statistics
Old Dominion University
Norfolk, Virginia 23529, U.S.A.
and

Shobha Prabhala

Searle, Inc.
Research and Development
4901 Searle Parkway, Skokie, Illinois 60077, U.S.A.

SUMMARY

The growth curve model $\mathbf{Y}_{i_{n_i \times p_i}} = \mathbf{A}_{i_{n_i \times q}} \xi_{q \times k} \mathbf{B}_{i_{k \times p_i}} + \in_{i_{n_i \times p_i}}, i = 1, ..., g$, where ξ is a matrix of unknown parameters and for the ith group, $\mathbf{Y}_i$ is the observation matrix, $\mathbf{A}_i$ is a known matrix of rank q, $\mathbf{B}_i$ is a known matrix of rank k, and the rows of $\in_i$ are independent each distributed as $N_{p_i}(0, \Sigma_i)$, Σ_i having a Markov structure is considered. The Markov structure on Σ, enables us to model the data, where the observations are made at unequally spaced time points or have missing values at any occasion. The maximum likelihood estimates of the parameters of the model are derived. Prediction of (a) the future observations for an individual, (b) the unobserved portion of a partially observed vector of measurements for an individual, and (c) any missing value for an individual, is discussed. Minimum mean square predictors as well as the predictors based on the maximum likelihood method are derived and their approximate variance in each case are given. Some of the results are illustrated using a data set.

Key words and phrases: Autoregressive structure, Markov covariance structure, maximum likelihood estimates, missing data, prediction, unbalanced data.

1. INTRODUCTION

The longitudinal data, repeated measures data, or growth curves data where several successive measurements over time are made on each subject (or experimental unit) occur naturally in the field of medical and other biological research. Typically, these data are obtained from the studies that are designed to (a) describe the changes in an individual's responses as time changes, and (b) compare mean responses over time among several groups of individuals.

If the number of measurements on each subject is small, the data can be analyzed by fitting polynomial growth curves to these repeated measurements. This can be achieved using the generalized multivariate analysis of variance (generalized MANOVA) or growth curve model introduced by Potthoff and Roy (1964) and further developed by Rao (1965, 1966, 1967), Khatri (1966), Grizzle and Allen (1969). An excellent review of these and other related models and their applications to various areas, can be found in Timm (1980), von Rosen (1991), Nummi (1995), and Kshirsagar and Smith (1995).

If the number of repeated measurements on each subject is same (balanced data; $p_i = p$), then the growth curve model is written as

$$\mathbf{Y}_{n\times p} = \mathbf{A}_{n\times q}\xi_{q\times k}\mathbf{B}_{k\times p} + \in_{n\times p}, \tag{1}$$

where $\mathbf{Y}$ is the observation matrix, ξ is a matrix of unknown parameters, $\mathbf{A}$ is a known matrix of rank $q < n$, and $\mathbf{B}$ is a known matrix of rank $k < p$. Further, the rows of the error matrix $\in$ are independent each distributed as $N_p(0, \Sigma)$, where Σ is a $p \times p$ positive definite matrix. In many applications, p represents the number of time points observed on each of the n cases, $(k-1)$ is the degree of the polynomial, and q the number of groups. The problems of interest are estimation of ξ and Σ, testing of general linear hypothesis of the form $H_0 : \mathbf{E}\xi\mathbf{F} = \mathbf{0}$, where $\mathbf{E}$ and $\mathbf{F}$ are known full rank matrices, and prediction of unobserved observations.

The two types of prediction problems that have been considered in the literature are: (i) Predict y_f, future observation for an individual, given $\mathbf{Y}$ and (ii) predict v_2, unobserved portion of a partially observed vector $\mathbf{v}'$ $(= (\mathbf{v}_1', v_2'))$ for an individual, given $\mathbf{v}_1$ and $\mathbf{Y}$. Another prediction problem that may be of interest to practitioners is (iii) prediction of unobserved portion v_m of a vector, given $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{Y}$, where $\mathbf{v}' = (\mathbf{v}_1', v_m', \mathbf{v}_2')$. This prediction also helps us to estimate the missing values in the growth curve model context. The prediction problem (i) has been considered by Rao (1977, 1984, 1987) and (ii) by Lee and Geisser (1972, 1975), Fearn (1975), Rao (1975), and Reinsel (1982). In the time series context, problem (iii) has been considered by Brubacher and Tunnicliffe-Wilson (1976) and Pena and Tiao (1991).

Recently, focus has been given to solving the above testing and prediction problems by adopting the maximum likelihood method and assuming certain correlation structures on Σ. For example, see Azzalini (1987), Lee (1988), Fujikoshi, Kanda and Tanimura (1990), Keramidas and Lee (1990), and Naik (1990). The correlation that exists between the repeated measurements (measurements on the same subject or experimental unit) taken at different time points is likely to decay with increasing time distance between the repeated measurements, leading to the first order autoregressive correlation structure for Σ. That is,

$$\Sigma = \sigma^2(\rho^{|i-j|}),\ \ i,j = 1,...,p.$$

For model (1) with the first order autoregressive structure for Σ, Lee (1988), and Fujikoshi, Kanda and Tanimura (1990) have derived the maximum likelihood estimators of the parameters involved and have suggested test for testing H_0. Further the two types of prediction problems (i) and (ii) are considered by Lee (1988). Recently, Naik and Prabhala (1997) considered the problem of fitting growth curve model with autoregressive covariance structure for monotone unbalanced data, and obtained the maximum likelihood estimators of the parameters involved and also discussed the three types of prediction problems (i)-(iii).

If the first order autoregressive structure for Σ is used for modeling an unbalanced data (other than monotone) or data observed at unequally spaced time periods, it leads to the Markov covariance structure. This case will be considered in this article. The derivation of the maximum likelihood (ML) estimators of the parameters will be considered in Section 2. The three prediction problems mentioned above will be considered in Section 3. The minimum mean square predictors in each of the three cases, using the conditional expectation formulae, along with their respective estimated approximate variances will be provided. Finally, we will illustrate our methods on a data set by computing the ML estimators and the predictors. Chi and Reinsel (1989) have considered Markov structure for the error along with certain random effects term in the model. For their model, Chi and Reinsel (1989) have considered prediction problem (ii).

2. MAXIMUM LIKELIHOOD ESTIMATES

Balanced data observed at unequally spaced time points, $t_1,...,t_p$, balanced incomplete data (monotone data) observed at unequally spaced time points, $t_{i1},...,t_{ip_i},\ i = 1,...,g$, and unbalanced (or missing) data are most common in medical fields, where the designed studies are hard or too expensive to perform, except in the cases of few planned clinical trials.

To analyze these data we consider the following generalized growth curve model,

$$\mathbf{Y}_{i_{n_i} \times p_i} = \mathbf{A}_{i_{n_i} \times q} \xi_{q \times k} \mathbf{B}_{k \times p} \mathbf{G}_{i_{p \times p_i}} + \in_{i_{n_i} \times p_i}, i = 1,...,g, \tag{2}$$

where $\mathbf{G}_i'$ is a matrix of 0's and 1's such that if the observations at time points $\alpha_1,...,\alpha_{p_i}$ are available then $\mathbf{G}_i$ has one in the (l,α_l)th position for $l = 1,...,p_i$ and zeros elsewhere. This model is quite general. Note that if $n_i = 1$ for all i and $g = n$, it allows unequal number of measurements p_i on every individual. Kleinbaum (1973) considered ML estimation of parameters under this model with a general covariance structure Σ for the error. We consider an autoregressive structure for Σ and derive the ML estimators. Under AR(1) structure for Σ, the rows of error matrix $\in_i$ in (2) are independent and identically distributed as $N_{p_i}(0, \mathbf{G}_i'\Sigma\mathbf{G}_i)$ and it is easy to see that

$$\mathbf{G}_i'\Sigma\mathbf{G}_i = \Sigma_i = \sigma^2(\rho^{|t_{ij}-t_{ij'}|}) = \sigma^2\mathbf{V}_i(\rho)\ , j, j' = 1, ..., p_i, \tag{3}$$

where t_{ij}, $j = 1, ..., p_i$ are the consecutive times points where the observations on an individual of the ith group are made. The structure (3) is called a Markov structure.

The determinant and the inverse of Σ_i are given below. Let $d_{ij} = t_{i(j+1)} - t_{ij}$, and $f_{ij} = \frac{1}{1-\rho^{2d_{ij}}}$, $j = 1, ..., p_i - 1$. Then

$$|\Sigma_i| = |\sigma^2\mathbf{V}_i(\rho)| = \sigma^{2p_i}(f_{i1}f_{i2}...f_{ip_i-1})^{-1}$$

and

$$\Sigma_i^{-1} = \sigma^{-2}\mathbf{V}_i(\rho)^{-1} = \sigma^{-2}\mathbf{V}_i^{-1} =$$
$$\sigma^{-2}\begin{pmatrix} f_{i1} & -f_{i1}\rho^{d_{i1}} & . & . & 0 \\ -f_{i1}\rho^{d_{i1}} & f_{i1}+f_{i2}-1 & . & . & 0 \\ 0 & -f_{i2}\rho^{d_{i2}} & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & f_{ip_i-2}+f_{ip_i-1}-1 & -f_{ip_i-1}\rho^{d_{ip_i-1}} \\ 0 & . & 0 & -f_{ip_i-1}\rho^{d_{ip_i-1}} & f_{ip_i-1} \end{pmatrix}$$

If $t_{i1}, t_{i2}, ..., t_{ip}$ are integers then $\mathbf{V}_i(\rho)$ is positive definite for $-1 < \rho < 1$, otherwise for $0 < \rho < 1$. Then we have the following theorem for ML estimators of ξ, σ^2 and ρ.

THEOREM **1** *The ML estimators of ξ, σ^2 and ρ in the unbalanced model with autoregressive covariance structure are the solutions of the following equations:*

$$(I)\ vec(\hat{\xi}) = [\sum_{i=1}^{g}(\mathbf{B}\mathbf{G}_i\mathbf{V}_i(\hat{\rho})^{-1}\mathbf{G}_i'\mathbf{B}') \otimes (\mathbf{A}_i'\mathbf{A}_i)]^{-1}\sum_{i=1}^{g}(\mathbf{B}\mathbf{G}_i\mathbf{V}_i(\hat{\rho})^{-1} \otimes \mathbf{A}_i')vec(\mathbf{Y}_i)$$

$$(II)\ \hat{\sigma}^2 = \frac{1}{\sum_{i=1}^{g} p_i n_i}\sum_{i=1}^{g} tr\hat{\mathbf{V}}_i(\hat{\rho})^{-1}\hat{\in}_i'\hat{\in}_i$$

$$(III)\ \hat{\sigma}^2\sum_{i=1}^{g} n_i \sum_{j=1}^{p_i-1} 2d_{ij}f_{ij}\hat{\rho}^{2d_{ij}-1} - \sum_{i=1}^{g}\sum_{j=1}^{p_i-1}[f_{ij}{}^2(2d_{ij})\hat{\rho}^{2d_{ij}-1}\sum_{k=1}^{n_i}(\hat{\epsilon}_{kj}^2 + \hat{\epsilon}_{kj+1}^2)$$
$$-2d_{ij}\hat{\rho}^{d_{ij}-1}f_{ij}{}^2(1+\hat{\rho}^{2d_{ij}})\sum_{k=1}^{n_i}(\hat{\epsilon}_{kj}\hat{\epsilon}_{kj+1})] = 0,$$

where $\hat{\in}_{i_{n_i\times p_i}} = (\hat{\epsilon}_{kj}),\ k = 1, ..., n_i;\ j = 1, ..., p_i,$ *is the matrix of residuals for the ith group.*

Note that the equation (III) for estimating ρ involves $\hat{\sigma}^2$. However, this does not create any problems for numerically evaluating the ML estimates using the three equations $(I)-(III)$. Solutions to these equations can be obtained using, for example, PROC MIXED procedure of SAS software. Using the estimates provided in Theorem 1 and the asymptotic distributions of the test statistics any linear hypothesis of the form $H_0 : \mathbf{E}\xi\mathbf{F} = \mathbf{0}$, for known $\mathbf{E}$ and $\mathbf{F}$ matrices, is usually tested.

3. PREDICTION

The minimum (predictive) mean square predictors for y_{if}, the future observation for the ith individual, v_2, the unobserved portion of the vector $\mathbf{v}'(= (\mathbf{v}_1', v_2'))$, and v_m, the missing portion of $\mathbf{v}'(= (\mathbf{v}_1', v_m', \mathbf{v}_2'))$ will be given next. For simplicity of presentation we assume that y_{if}, v_2, and v_m are all scalars. However, in general, these can be vectors and the formulae given below can easily be adopted to that case.

3.1 Prediction of y_f

For the jth individual in the ith group suppose y_{if_j} is to be predicted. Extending model (2) we have have

$$E\begin{pmatrix} \mathbf{Y}_{ij} \\ y_{if_j} \end{pmatrix} = \begin{pmatrix} (\mathbf{A}_{ij}'\xi\mathbf{B}\mathbf{G}_i)' \\ (\mathbf{A}_{ij}'\xi\mathbf{B}\mathbf{G}_{if})' \end{pmatrix} \text{ with}$$
$$cov\begin{pmatrix} \mathbf{Y}_{ij} \\ y_{if_j} \end{pmatrix} = \Sigma_{fi} = \begin{pmatrix} \Sigma_i & \sigma_{fi} \\ \sigma_{fi}' & \sigma_{2fi} \end{pmatrix},$$

where Σ_{fi} is a $(p_i+1)\times(p_i+1)$ positive definite covariance matrix $\mathbf{A}_{ij}'$ is the jth row of $\mathbf{A}_i$ and $\mathbf{BG}_{if}$ is a $k\times 1$ known vector. For Markov covariance structure it is easy to show that

$$\sigma_{fi}'\Sigma_i^{-1} = (0_{p_i-1}', \frac{\rho^{d_{ip_i}}}{1-\rho^{2d_{ip_i-1}}}[1-\rho^{(d_{ip_i}+d_{ip_i-1})}]).$$

Then the minimum mean square predictor for y_{if_j}, after substituting the (ML) estimators of the parameters, is given by

$$\hat{y}_{if_j} = (\mathbf{A}_{ij}'\hat{\xi}\mathbf{B}\mathbf{G}_{if})' + (0_{p_i-1}', \frac{\hat{\rho}^{d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{(d_{ip_i}+d_{ip_i-1})}])(\mathbf{Y}_{ij} - (\mathbf{A}_{ij}'\hat{\xi}\mathbf{B}\mathbf{G}_i)')$$

with the estimated variance of the predictor,

$$\hat{\tau}_{ij}^2 = \hat{\sigma}^2(1 - \frac{\hat{\rho}^{2d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}]) + c_{ij}\hat{\mathbf{h}}_i'\hat{\mathbf{D}}_i\hat{\mathbf{h}}_i,$$

where $c_{ij} = \mathbf{A}_{ij}'(\mathbf{A}_i'\mathbf{A}_i)^{-1}\mathbf{A}_{ij}$, $\hat{\mathbf{D}}_i = (\mathbf{B}\mathbf{G}_i\hat{\Sigma}_i^{-1}\mathbf{G}_i'\mathbf{B}')^{-1}$ and $\hat{\mathbf{h}}_i = \mathbf{B}\mathbf{G}_{if} - \mathbf{B}\mathbf{G}_i(0_{p_i-1}', \frac{\hat{\rho}^{d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}])'$.

3.2 Prediction of v_2

Next, let $E(\mathbf{v}_{i_{1\times p_i}}) = E(\mathbf{v}_{1i}, v_{2i}) = (\mathbf{A}_{if}\xi\mathbf{B}\mathbf{G}_{1i}, \mathbf{A}_{if}\xi\mathbf{B}\mathbf{G}_{2i})$, where $\mathbf{G}_{i_{k\times p_i}} = (\mathbf{G}_{1i}, \mathbf{G}_{2i})$. Let $cov(\mathbf{v}_i) = \Sigma_i = \mathbf{G}_i'\Sigma\mathbf{G}_i$, where Σ has an autoregressive structure. As before, this implies that Σ_i has a Markov structure. Then

$$\hat{v}_{2i} = \mathbf{A}_{if}\hat{\xi}\mathbf{B}\mathbf{G}_{2i} + (\mathbf{v}_{1i} - \mathbf{A}_{if}\hat{\xi}\mathbf{B}\mathbf{G}_{1i})(0'_{p_i-2}, \frac{\hat{\rho}^{d_{p_i-1}}}{1-\hat{\rho}^{2d_{p_i-2}}}[1-\hat{\rho}^{d_{ip_i-1}+d_{ip_i-2}}])'$$

with its estimated variance

$$\hat{\tau}_i^2 = \hat{\sigma}^2(1 - \frac{\hat{\rho}^{2d_{ip_i-1}}}{1-\hat{\rho}^{2d_{ip_i-2}}}[1-\hat{\rho}^{d_{ip_i-1}+d_{ip_i-2}}]) + c_{fi}\hat{\mathbf{d}}_i'\hat{\mathbf{D}}_i\hat{\mathbf{d}}_i,$$

where $c_{fi} = \mathbf{A}_{if}(\mathbf{A}_i'\mathbf{A}_i)^{-1}\mathbf{A}_{if}'$, $\hat{\mathbf{D}}_i = (\mathbf{B}\mathbf{G}_i\hat{\Sigma}_i^{-1}\mathbf{G}_i'\mathbf{B}')^{-1}$, and $\hat{\mathbf{d}}_i = \mathbf{B}\mathbf{G}_{2i} - \mathbf{B}\mathbf{G}_{1i}(0'_{p_i-2}, \frac{\hat{\rho}^{d_{ip_i-1}}}{1-\hat{\rho}^{2d_{ip_i-2}}}[1-\hat{\rho}^{d_{ip_i-1}+d_{ip_i-2}}])$.

3.3 Prediction of v_m

In the following, we derive a formula for predicting v_m, based on the conditional expectation. As in the previous cases, for convenience, suppose v_m is a scalar and it corresponds to the mth position (that is, the mth element) of the vector $\mathbf{v}$. Again for easy presentation we suppress the suffix 'i' for the group. Also let $\mathbf{B}\mathbf{G}_m = \mathbf{b}_m, \mathbf{B}\mathbf{G}_1 = \mathbf{B}_1$ and $\mathbf{B}\mathbf{G}_2 = \mathbf{B}_2$. Re-arrange the elements of $\mathbf{v}$ so that $\mathbf{v}' = (v_m, \mathbf{v}_1', \mathbf{v}_2')$. Let $E(\mathbf{v}') = (E(v_m), E(\mathbf{v}_1'), E(\mathbf{v}_2')) = (A_f\xi\mathbf{b}_m, A_f\xi\mathbf{B}_1, A_f\xi\mathbf{B}_2)$, where A_f etc. are appropriately defined vectors or matrices. We similarly partition the covariance matrix of $\mathbf{v}$ as

$$cov\begin{pmatrix} v_m \\ \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} = \begin{bmatrix} \omega_{mm} & \omega_{m1} & \omega_{m2} \\ \omega_{1m} & \Omega_{11} & \Omega_{12} \\ \omega_{2m} & \Omega_{21} & \Omega_{22} \end{bmatrix} \equiv \begin{bmatrix} \omega_{mm} & \omega_m \\ \omega_m' & \Omega \end{bmatrix}.$$

The minimum mean square predictor of v_m then is

$$E(v_m/\mathbf{v}_1, \mathbf{v}_2, Y) = A_f\xi\mathbf{b}_m + \omega_m'\Omega^{-1}\left[\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} - \begin{pmatrix} A_f\xi\mathbf{B}_1 \\ A_f\xi\mathbf{B}_2 \end{pmatrix}'\right]$$

Computation of $\omega_m'\Omega^{-1}$ and $\omega_m'\Omega^{-1}\omega_m$ for calculating the variance of the predictor are provided in an Appendix. The predictor of v_m given $\mathbf{v}_1$, $\mathbf{v}_2$, and Y is given by

$$\begin{aligned} \hat{v}_m = A_f\hat{\xi}\mathbf{b}_m &+ \frac{\hat{\rho}^{d_{m-1}}(1-\hat{\rho}^{2d_m})}{1-\hat{\rho}^{2(d_{m-1}+d_m)}}(v_{m-1} - A_f\hat{\xi}\mathbf{b}_{m-1}) \\ &+ \frac{\hat{\rho}^{d_m}(1-\hat{\rho}^{2d_{m-1}})}{1-\hat{\rho}^{2(d_{m-1}+d_m)}}(v_{m+1} - A_f\hat{\xi}\mathbf{b}_{m+1}) \end{aligned} \tag{4}$$

with the estimated variance

$$\hat{\tau_i}^2 = \frac{\hat{\sigma}^2}{1-\hat{\rho}^2}\frac{[1-\hat{\rho}^{2d_{m-1}} - \hat{\rho}^{2d_m} + \hat{\rho}^{2(d_{m-1}+d_m)}]}{1-\hat{\rho}^{2(d_{m-1}+d_m)}} + c_f\hat{\mathbf{d}}'\hat{\mathbf{D}}\hat{\mathbf{d}},$$

where

$$\hat{\mathbf{d}} = \mathbf{b}_m - [\frac{\hat{\rho}^{d_{m-1}}(1-\hat{\rho}^{2d_m})}{1-\hat{\rho}^{2(d_{m-1}+d_m)}}\mathbf{b}_{m-1} + \frac{\hat{\rho}^{d_m}(1-\hat{\rho}^{2d_{m-1}})}{1-\hat{\rho}^{2(d_{m-1}+d_m)}}\mathbf{b}_{m+1}]$$

4. MAXIMUM LIKELIHOOD PREDICTOR

Recently Bjornstad (1996), extended the likelihood principle to include the prediction problems. It is argued that the correct likelihood to consider, when the prediction is of interest, is the joint probability density function of all the random variables including the ones that are being predicted and not the likelihood function based on the conditional distribution of observed variables given the unobserved variables (being predicted). In this section we derive the predictors of missing observations by maximizing the joint probability density function of all the random quantities including that those are being predicted.

Let the joint density function of Y, and $\mathbf{v}$ be $f(Y, \mathbf{v})$. Then

$$f(Y, \mathbf{v}) = f(v_m | Y, \mathbf{v}_1, \mathbf{v}_2)\, f(Y, \mathbf{v}_1, \mathbf{v}_2)$$

Under the normality assumption, the ML predictor of v_m is nothing but (4), except that the estimators of ξ, ρ, and σ^2 are to be obtained by simultaneously maximizing $f(Y, \mathbf{v})$. In the following we will provide an algorithm for computing the estimators of all the fixed parameters and the predictor.

The procedure essentially goes back and forth from the formulae developed in (4) for $\hat{v}_m$ and the maximum likelihood estimates in Theorem 1. It can be described in several steps as follows:

Step 1. Compute $\hat{v}_m$ using the formula given in (4) for an initial value of ρ, say $\rho = 0.5$. Denote this predictor by $\tilde{v}_m$.

Step 2. Using $\tilde{v}_m$, form the vector $\tilde{\mathbf{v}}' = (\tilde{v}_m, \mathbf{v}_1, \mathbf{v}_2)$ and the new data matrix

$$\mathbf{Y}_{new} = \begin{pmatrix} \mathbf{Y} \\ \tilde{\mathbf{v}}' \end{pmatrix}$$

Step 3. Use Theorem 1 to evaluate ξ, ρ and σ^2 replacing $\mathbf{Y}$ by $\mathbf{Y}_{new}$ and $\mathbf{A}$ by

$$\mathbf{A}_{new} = \begin{pmatrix} \mathbf{A} \\ \mathbf{A}_f \end{pmatrix}$$

Step 4. Repeat steps (1)-(3) until the estimators are stabilized.

It is an interesting fact to note that the algorithm given above is nothing but the EM algorithm (Little and Rubin (1987)) applied to the multivariate normal data in the growth curve setup.

The prediction formulae developed in this paper can be easily extended when the covariance matrix is a slight generalization of Markov covariance structure considered by Nu′ñez-Anto′n and Woodworth (1994).

5. AN EXAMPLE

In a study of the association of hyperglycemia and relative hyperinsulinemia, standard glucose tolerance tests were administered to 13 control and 20 obese patients on the Pediatric Clinical Research Ward of the University of Colorado Medical Center (Zerbe, 1979). Plasma inorganic phosphate measurements determined from blood samples withdrawn 0, 0.5, 1, 1.5, 2, 3, 4 and 5 hours after a standard dose oral glucose challenge are observed.

Suppose interest is to predict the value of plasma inorganic phosphate at an unobserved time period in the middle, say at time point 2.5, for an individual. Note that measurements have been made at the time periods 2 as well as 3, but not at 2.5. We can use formula (4) for this prediction. In the following we will illustrate the prediction for the first individual only. For this prediction, according to formula (4), data corresponding to the first individual is used only for prediction and not for estimation of ξ. Further, not all but only two neighboring measurements, 2.2 at the location 2 and 2.5 at the location 3 are used in the prediction formula. We first fit model (2) for these data with $n_1 = 12$ (leaving out the first individual in group 1), $n_2 = 20$, $g = 2$, $q = 2$, $p_1 = p_2 = 8$, and $k = 3$, referring to a second degree polynomial fit in eight time periods. Because the time periods at which the observations are made are unequally spaced, we use Markov covariance structure to fit the covariances between them. Also based on some initial study, we fit different intercept but common slope model for both the groups.

The maximum likelihood estimate of ξ is

$$\hat{\xi} = \begin{bmatrix} 4.1090 & -.8366 & .1556 \\ 4.4308 & -.8366 & .1556 \end{bmatrix},$$

and of ρ is $\hat{\rho} = .5689$. From the assumed model and the data the other quantities required to use formula (4) are, $A_f = (1, 0)$, $\mathbf{b}_m = (1, 2.5, 6.25)'$, $\mathbf{b}_{m-1} = (1, 2, 4)'$, $\mathbf{b}_{m+1} = (1, 3, 9)'$, $d_{m-1} = t_m - t_{m-1} = 2.5 - 2 = .5$, and $d_m = t_{m+1} - t_m = 3 - 2.5 = .5$. Also as noted above $v_{m-1} = 2.2$ and $v_m = 2.5$. Putting these quantities in the prediction formula and simplifying we get $\hat{v}_{2.5} = 2.3372$. Similarly, prediction at any time point or for any individual can be carried out.

One of the interest is to empirically assess these predicted values. For that let us presume that the values at the time period 2 are not available. Using the calculations similar to the above, predict the values at time period 2 for all the individuals in the data set. Then the average squared difference between the predicted and observed values can be taken as a measure of the performance of the predictor. For the data set in hand,

$$\frac{1}{33}\sum_{1}^{33}(\hat{v}_m - v_m)^2 = 0.3706.$$

Although we cannot make judgement based on absolute quantity like the one derived above, the number 0.3706 seems to be quite small indicating that the prediction is good.

Now for a comparison of the above predictor with the maximum likelihood predictor computed using the algorithm we have calculated the measure $\frac{1}{33}\sum_{1}^{33}(\hat{v}_m(ML) - v_m)^2$ and this value turns out to be 0.234. This indicates that the ML predictor may be slightly better, evaluated using this empirical measure, in the present context.

REFERENCES

Azzalini, A. (1987). Growth curve analysis for patterned covariance matrices. *New Perspectives in Theoretical and Applied Statistics*, Ed. M.L. Puri, J.P. Vilaplana, and W. Wertz, pp. 63-73, Wiley, New York.

Bjornstad, J.F. (1996). On the generalization of the likelihood function and the likelihood principle. *Journal of American Statistical Association*, 91, 791-806.

Brubacher, S.R. and Tunnicliffe-Wilson, G. (1976). Interpolating time series with application to the estimation of holiday effects on the electricity demand. *Applied Statistics*, 25, 107-116.

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

Fearn, T. (1975). A bayesian approach to growth curves. *Biometrika*, 62, 89-100.

Fujikoshi, Y., Kanda, T. and Tanimura, N. (1990). The growth curve model with an autoregressive covariance structure. *Annals of Institute of Statistical Mathematics* , 42, 533-542.

Grizzle, J.E. and Allen, D.M. (1969). Analysis of growth and dose response curves. *Biometrics*, 25, 413-424.

Keramidas, E.M. and Lee, J.C. (1990). Forecasting technological substitutions with short time series. *Journal of American Statistical Association*, 85, 625-632.

Khatri, C.G. (1966). A note on a MANOVA model applied to problems in growth curve. *Annals of Institute of Statistical Mathematics* , 18, 75-86.

Kleinbaum, D.G. (1973). A generalization of the growth curve model which allows missing data. *Journal of Multivariate Analysis* , 3, 117-124.

Kshirsagar, A.M. and Smith, W.B. (1995). *Growth Curves* . Marcel Dekker, New York.

Lee, J.C. (1988). Prediction and estimation of growth curves with special covariance structures. *Journal of American Statistical Association*, 83, 432-440.

Lee, J.C. and Geisser, S. (1972). Growth curve prediction. *Sankhya*, ser. A, 34, 393-412.

Lee, J.C. and Geisser, S. (1975). Applications of growth curve prediction. *Sankhya*, ser. A, 37, 239-256.

Little, R.J.A. and Rubin, D.B. (1987). *Statistical Analysis With Missing Data* . Wiley, New York.

Naik, D.N. (1990). Prediction intervals for growth curves. *Journal of Applied Statistics*, 17, 245-254.

Naik, D.N. and Prabhala, S. (1997). Analysis of growth curves under autoregressive covariance structure. *Journal of Applied Statistical Sciences* , 5, 291-301.

Nummi, T. (1995). *Estimation and Prediction in Growth Curve Models with Applications.* University of Tampere, Tampere.

Nu'ñez-Anto'n, V. and Woodworth, G.G. (1994). Analysis of longitudinal data with unequally spaced observations and time-dependent correlated errors. *Biometrics*, 50, 445-456.

Pena, D. and Tiao, G.C. (1991). A note on likelihood estimation of missing values in time-series. *The American Statistician*, 45, 212-213.

Potthoff, R.F. and Roy, S.N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, 51, 313-326.

Rao, C.R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, 52, 447-458.

Rao, C.R. (1966). Covariance adjustment and related problems in multivariate analysis. *Multivariate Analysis*, Vol. 1, ed. P. R. Krishnaiah, pp. 87-103, Academic Press, New York.

Rao, C.R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. *Proc.* 5^{th} *Berkeley Symp. on Math. Statist. Prob.*, Vol. 1, pp. 355-372, Univ. of California Press, Berkeley.

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

Rao, C.R. (1975). Simultaneous estimation of parameters in different linear models and applications to biometric problems. *Biometrics*, 31, 545-554.

Rao, C.R. (1977). Prediction of future observations with special reference to linear models. *Multivariate Analysis*, Vol. 4, ed. P.R. Krishnaiah, pp. 193-208, Academic Press, New York.

Rao, C.R. (1984). Prediction of future observations in polynomial growth curve models. *Proceedings of the Indian Statistical Institute Golden Jubilee International Conference on Statistics: Applications and New Directions,* pp. 512-520, Indian Statistical Institute, Calcutta.

Rao, C.R. (1987). Prediction of future observations in growth curve models. *Statistical Science*, 2, 434-471.

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

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

von Rosen D. (1991). The growth curve model: A review. *Communications in Statistics-Theory Method*, 20, 2791-2822.

Zerbe, G.O. (1979). Randomization analysis of the completely randomized design extended to growth and response curves. *Journal of American Statistical Association*, 74, 215-221.

APPENDIX

Computation of $\omega'_m\Omega^{-1}$. First we compute the inverse of Ω. For that we use the formula

$$\begin{bmatrix} A & B \\ B' & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1}+A^{-1}BE^{-1}B'A^{-1} & -A^{-1}BE^{-1} \\ -E^{-1}B'A^{-1} & E^{-1} \end{bmatrix}$$

where $E = D - B'A^{-1}B$ (see Rao, 1973, p. 33).

We use this formula to compute

$$\Omega^{-1} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1}.$$

Note that the $m-1\times m-1$ matrix Ω_{11}, the $p-m\times p-m$ matrix Ω_{22}, and the $m-1\times p-m$ matrix Ω_{12} for Markov structure are

$$\Omega_{11} = \sigma^2 \begin{pmatrix} 1 & \rho^{t_2-t_1} & \rho^{t_3-t_1} & . & . & \rho^{t_{m-1}-t_1} \\ \rho^{t_2-t_1} & 1 & \rho^{t_3-t_2} & . & . & \rho^{t_{m-1}-t_2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{t_{m-1}-t_1} & \rho^{t_{m-1}-t_2} & \rho^{t_{m-1}-t_3} & . & . & 1 \end{pmatrix}$$

$$\Omega_{22} = \sigma^2 \begin{pmatrix} 1 & \rho^{t_{m+2}-t_{m+1}} & \rho^{t_{m+3}-t_{m+1}} & . & . & \rho^{t_p-t_{m+1}} \\ \rho^{t_{m+2}-t_{m+1}} & 1 & \rho^{t_{m+3}-t_{m+2}} & . & . & \rho^{t_p-t_{m+2}} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{t_p-t_{m+1}} & \rho^{t_p-t_{m+2}} & \rho^{t_p-t_{m+3}} & . & . & 1 \end{pmatrix},$$

$$\Omega_{12} = \sigma^2 \begin{pmatrix} \rho^{t_{m+1}-t_1} & \rho^{t_{m+2}-t_1} & . & . & . & \rho^{t_p-t_1} \\ \rho^{t_{m+1}-t_2} & \rho^{t_{m+2}-t_2} & . & . & . & \rho^{t_p-t_2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{t_{m+1}-t_{m-1}} & \rho^{t_{m+2}-t_{m-1}} & . & . & . & \rho^{t_p-t_{m-1}} \end{pmatrix} = \Omega'_{21}.$$

Further,

$$\omega'_m = (\omega_{1m}, \omega_{2m}) = \\ \sigma^2(\rho^{t_m-t_1}, \rho^{t_m-t_2}, ..., \rho^{t_m-t_{m-2}}, \rho^{t_m-t_{m-1}}, \rho^{t_{m+1}-t_m}, \rho^{t_{m+2}-t_m}, ..., \rho^{t_p-t_m}).$$

With some algebra it can be shown that

$$\omega'_m\Omega^{-1} = (0, ..., 0, \frac{\hat{\rho}^{d_{m-1}}(1-\hat{\rho}^{2d_m})}{1-\hat{\rho}^{2(d_{m-1}+d_m)}}, \frac{\hat{\rho}^{d_m}(1-\hat{\rho}^{2d_{m-1}})}{1-\hat{\rho}^{2(d_{m-1}+d_m)}}, 0, ..., 0)$$

and

$$\omega'_m\Omega^{-1}\omega_m = \sigma^2 \frac{\rho^{2d_{m-1}}(1-\rho^{2d_m}) + \rho^{2d_m}(1-\rho^{2d_{m-1}})}{1-\rho^{2(d_{j-1}+d_j)}}.$$