Fall 1995

Analysis of Growth Curves Under Some Special Covariance Structures

Shobha Prabhala
Old Dominion University

ANALYSIS OF GROWTH CURVES UNDER SOME SPECIAL COVARIANCE STRUCTURES

by

Shobha Prabhala

M.Sc., May 1990, Indian Institute of Technology
Kanpur, India

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

DOCTOR OF PHILOSOPHY
Computational and Applied Mathematics

OLD DOMINION UNIVERSITY
September, 1995

Approved by:

Dr. D. N. Naik (Director)

ABSTRACT

ANALYSIS OF GROWTH CURVES UNDER SOME SPECIAL COVARIANCE STRUCTURES

Shobha Prabhala

Old Dominion University, 1995
Director: Dr. D. N. Naik

In this dissertation we consider the growth curve or generalized MANOVA model in its most general form given by,

$$Y_{ij_{n_i \times p_{ij}}} = A_{i_{n_i \times g}} \xi_{g \times k} B_{ij_{k \times p_{ij}}} + \epsilon_{ij_{n_i \times p_{ij}}}, \ i = 1, 2, \ldots, g, \ j = 1, 2, \ldots, n_i,$$

and develop statistical methodology for analyzing data using this model. Here g represents the number of groups, Y_{ij} is the observation matrix, ξ is a matrix of unknown parameters, A_i is a known matrix of rank g, and B_{ij} is a matrix of rank k. Further, the rows of the error matrix ϵ_{ij} are independent and each distributed as $N_{p_{ij}}(0, \Sigma_{ij})$. This model accommodates different kinds of unbalanced data, such as, monotone data, data missing from any occasion, and data observed at unequally spaced time points.

Our main results are: (1) derivation of the formulae for the maximum likelihood estimates (MLEs) of the parameters involved, (2) construction of the tests for testing general linear hypothesis of the form $H_o : E_{q \times g} \xi_{g \times k} F_{k \times v} = 0$, for known full rank matrices E and F, and (3) derivation of the formulae for prediction of (a) future observations corresponding to an individual, (b) the unobserved portion of

a partially observed data for a new individual, and (c) any missing value of an observation vector.

Deriving the maximum likelihood estimates and the prediction formulae for unbalanced data is a challenging problem. We have derived these results by taking two types of covariance structures for Σ_{ij}. These structures, namely equicorrelation structure and autoregressive structure, are most commonly used in the literature. For the autoregressive structure, the maximum likelihood estimator of the correlation parameter turns out to be a solution of a cubic equation. We prove that this cubic equation has a unique real root in $(-1, 1)$. This proves the uniqueness of the MLE. Further, we notice that the autoregressive structure leads to Markov structure when the data are observed at unequally spaced time intervals. For the model with Markov covariance structure, we derive a formula for estimating a missing value and show that the estimator based on this formula depends on only two neighboring data values. The results for equicorrelation structure are included in Chapter 2 and those for the autoregressive structure (Markov structure as well) are included in Chapter 3.

Finally, in the fourth chapter we point out some draw backs of fitting the linear growth curve models to biological data and suggest fitting nonlinear models to growth data. After reviewing the popular nonlinear models, we show the analysis of nonlinear models with different covariance structures using SAS software.

Dedicated

to

my parents

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor, Dr. D. N. Naik, for always finding time in his busy schedule to help me with my research. Without his guidance and support I could not have progressed this far. I would like to express my appreciation and gratitude to my dissertation committee, Dr. Ram Dahiya, Dr. Cynthia Jones and Dr. Larry Lee for their helpful suggestions and comments.

I acknowledge and am grateful to the strong friendship and support I received from Alka Dhakar, Surya Dhakar, Aditi, Aditya, Justine Shults and Cynthia Ghaem.

I owe my special thanks and gratitude to my dearest friend Sandeep Dhadda for his excellent advice and unfailing support during this difficult phase of my life.

Finally, my heartfelt thanks to my family for their love and support.

Contents

List of Tables

Chapter 1

Introduction

The analysis of growth data is important in many fields of study, like biology, medicine, agriculture, and education. Often in practice, data are of the form where several successive measurements over time are made on each subject (or experimental unit) and occur naturally. Especially in the field of medical and fisheries research, such data are common. Typically, data are obtained from studies that are designed to (a) describe the changes in an individual's response as time changes (b) compare mean responses (mean response curves) over time among several groups of individuals. These studies are called longitudinal studies, repeated measures data analysis, or analysis of growth curves.

Some examples where problems (a) and (b) are of interest are:

- Medical trials involving several groups of subjects, where measurements may be collected on each subject at regular time intervals.

• Growth (of animal or plant) experiments, where observations may be made on the same individual (animal or plant) as it grows and changes over time.

• In fisheries research, growth of certain species of fish may be observed over a period of time in the lakes of different geographical regions.

In many cases, especially if the number of measurements on each subject is small, these data can be analyzed by fitting polynomial growth curves to the repeated measures. This can be achieved using the generalized multivariate analysis of variance (generalized MANOVA) or the growth curve model introduced by Potthoff and Roy (1964). If the number of repeated measurements on each subject is the same, the growth curve model can be written as

$$Y_{n\times p} = A_{n\times g}\xi_{g\times k}B_{k\times p} + \in_{n\times p}, \tag{1.1}$$

where Y is an observation matrix, ξ is a matrix of unknown parameters, A is a known matrix of rank $g < n$, and 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 general, p represents the number of time points observed in each of the n cases, $(k-1)$ is the degree of the polynomial and g is the number of groups. Here B will be the matrix of polynomial terms.

The problems of interest relating to this model are :

• Estimation of ξ and Σ,

- Testing of general linear hypotheses of the form

$$H_o : D_{q\times g}\xi_{g\times k}F_{k\times v} = 0,$$

for known full rank matrices D and F, and

- Prediction of future observations.

Estimation and testing of hypothesis problems for this model have been discussed by many authors. Some of the important articles in this area are Potthoff and Roy (1964), Rao (1965, 1966, 1967), Khatri (1966), and Grizzle and Allen (1969). A survey of the analysis of this model is given by Timm (1980), and more recently, by von Rosen (1991). Applications of the growth curve models to different fields can be found in Nummi (1995). A recent book solely on growth curves, by Kshirsagar and Smith (1995), is an excellent collection of materials in this area with many real life applications. The prediction problem is discussed by Rao (1977). Recently Rao (1987) revived interest in this model by considering prediction problems using several methods.

1.1 Estimation and testing

Suppose our interest is to test $H_0 : D\xi F = 0$, where D is a $q \times g$ matrix of rank q and F is a $k \times v$ matrix of rank v. A likelihood ratio test for testing H_0, can be constructed, but the distribution of the test statistic is intractable. Hence, the following two methods have been suggested in the literature :

1. Potthoff & Roy's method.

2. Rao-Khatri analysis of covariance method.

We discuss only the Rao-Khatri method. In this method, a $p \times k$ matrix C_1 of rank k and a $p \times (p-k)$ matrix C_2 of rank $(p-k)$ are chosen such that $BC_1 = I_k$ and $BC_2 = 0$. Let $C = (C_1, C_2)$ be a $p \times p$ nonsingular matrix. Then make the transformation $Y_c = YC = (Y_1, Y_2)$ (say), where $Y_1 = YC_1$ and $Y_2 = YC_2$. Thus it is easy to see that $E(Y_1) = A\xi$ and $E(Y_2) = 0$, so the analysis of covariance model is

$$E[Y_1|Y_2] = A\xi + Y_2\Gamma = (A, Y_2)\,(\xi', \Gamma')'.$$

Since the rows of Y_1 are conditionally distributed as independent multivariate normals with a common covariance matrix, the theory of analysis of covariance can be applied. The estimates are:

$$\begin{aligned} \hat{\xi} &= (A'A)^{-1}A'YS^{-1}B'(BS^{-1}B')^{-1} \text{ and} \\ \hat{\Gamma} &= (C_2'SC_2)^{-1}C_2'SC_1, \end{aligned}$$

where

$$S = Y'[I - A(A'A)^{-1}A']Y/(n-p).$$

It can be shown that $\hat{\xi}$ is the maximum likelihood (ML) estimate of ξ w.r.t the conditional model as well as the original model. To test the hypothesis $H_0 : D\xi F = 0$, we define two matrices E and H as follows:

$$E = (n-g)F'(BS^{-1}B')^{-1}F \text{ and}$$

$$H = F'\hat{\xi}'D'[D(A'A)^{-1}D' + DLMLL'D']^{-1}D\hat{\xi}F,$$

where

$$LML' = \frac{1}{n-g}(A'A)^{-1}A'Y[S^{-1} - S^{-1}B'(BS^{-1}B')^{-1}BS^{-1}]Y'A(A'A)^{-1}.$$

When H_0 is true, E and H are independently distributed as v-dimensional Wishart with a common dispersion matrix and respective degrees of freedom $g+p-k$ and q . Since these Wishart distributions do not depend on the condition that Y_2 is fixed, these are also the unconditional distributions of these matrices under H_0. Using these matrices, the standard multivariate tests can be constructed. For example, Wilks' Λ for testing H_0 can be written using the E and H matrices as $\Lambda = |E+H|^{-1}|E|$. See Rao (1973) for a description of various multivariate tests.

1.2 Prediction

There are two types of prediction problems of interest that have been considered in the literature:

(1) Predict V_2 given V_1 and Y, where $V = (V_1, V_2)$ is a set of observations drawn from the growth curve model on a new individual (Lee and Geisser (1972, 1975), Fearn (1975), Rao (1975), and Reinsel (1984)). Here V has dimension $1 \times p$, V_1 has dimension $1 \times r$ and V_2 has dimension $1 \times (p-r)$. This problem is concerned with prediction of the unobserved portion of a partially observed vector corresponding to a new individual.

(2) The second problem is to predict y_f given Y, where y_f is a set of $N(\leq n)$ future q-dimensional observations whose current p-dimensional observations are a subset of Y (Rao (1977, 1984, 1987)). This problem is concerned with predicting future values on some or all of the already observed subjects.

Prediction of V_2 :

For simplicity assume that $r = p - 1$. Then V_2 is a scalar. Let $E(V) = E(V_1, V_2) = (E(V_1), E(V_2)) = A_f \xi B = (A_f \xi B_1, A_f \xi B_2)$, where A_f is an $1 \times m$ known vector, and $B = (B_1, B_2)$. Further, let

$$cov(V) = \Sigma = \begin{pmatrix} \Sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix},$$

where Σ is partitioned in accordance with the partition of V. Then the minimum predictive mean square error predictor for V_2 based on the conditional expectation of V_2 given Y and V_1 is given by

$$\hat{V}_2 = A_f \hat{\xi} B_2 + (V_1 - A_f \hat{\xi} B_1) \Sigma_{11}^{-1} \sigma_{12}. \tag{1.2}$$

Also

$$var(\hat{V}_2 - V_2) = \tau^2 = \sigma_{22} - \sigma_{21} \Sigma_{11}^{-1} \sigma_{12} + c_f d' D d,$$

where $c_f = A_f (A'A)^{-1} A_f', D = (B \Sigma^{-1} B')^{-1}$, and $d = B_2 - B_1 \Sigma_{11}^{-1} \sigma_{12}$. Since in practice, Σ in the prediction formula is unknown, the predictor is not computable.

However, if a consistent estimate of Σ is available, then the predictor can be evaluated as

$$\hat{V}_2 = A_f\hat{\xi}B_2 + (V_1 - A_f\hat{\xi}B_1)\hat{\Sigma}_{11}^{-1}\hat{\sigma}_{12}.$$

The estimated variance is

$$\hat{\tau}^2 = \hat{\sigma}_{22} - \hat{\sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\sigma}_{12} + c_f\hat{d}'\hat{D}\hat{d}.$$

Naik (1990) has constructed approximate prediction intervals for V_2. Since $(\hat{V}_2 - V_2)/\tau$ is approximately distributed as $N(0,1)$, a $100(1-\alpha)$ % prediction interval for V_2 can be constructed using the normal distribution. This can be summarized in the following theorem

THEOREM 1 : *If an estimate of* Σ,

$$\hat{\Sigma} = \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} \end{pmatrix},$$

is available, then an approximate $100(1-\alpha)$ *% prediction interval for* V_2 *is given by*

$$(\hat{V}_2 - z_{\alpha/2}\hat{\tau}, \hat{V}_2 + z_{\alpha/2}\hat{\tau}), \tag{1.3}$$

where

$$\begin{aligned} \hat{V}_2 &= A_f\hat{\xi}B_2 + (V_1 - A_f\hat{\xi}B_1)\hat{\Sigma}_{11}^{-1}\hat{\sigma}_{12}, \\ \hat{\tau}^2 &= \hat{\sigma}_{22} - \hat{\sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\sigma}_{12} + c_f\hat{d}'\hat{D}\hat{d}, \\ \hat{d} &= B_2 - B_1\hat{\Sigma}_{11}^{-1}\hat{\sigma}_{12}\ ,\ \hat{D} = (B\hat{\Sigma}^{-1}B')^{-1}, \end{aligned}$$

and $z_{\alpha/2}$ is the standard normal distribution cut off point.

Prediction of y_f:

Let us assume $N = n$ *and* $q = 1$ for simplicity. Then interest is to build prediction formula for y_{fi}, where y_{fi} is the future observation (at $(p+1)^{th}$ time point) of Y_i, the i^{th} row of Y. For the i^{th} individual we have

$$E\begin{pmatrix} Y_i \\ y_{fi} \end{pmatrix} = \begin{pmatrix} (A_i'\xi B)' \\ (A_i'\xi B_f)' \end{pmatrix}, cov\begin{pmatrix} Y_i \\ y_{fi} \end{pmatrix} = \Sigma_f = \begin{pmatrix} \Sigma & \sigma_f \\ \sigma_f' & \sigma_{2f} \end{pmatrix},$$

where A_i' is the i^{th} row of A, B_f is a $k \times 1$ known vector, and Σ_f is a $(p+1)\times(p+1)$ covariance matrix. Then the minimum predictive mean square error predictor based on the conditional expectation of y_{fi} given Y is

$$\hat{y}_{fi} = (A_i'\hat{\xi}B_f)' + \sigma_f'\Sigma^{-1}(Y_i - (A_i'\hat{\xi}B)') \tag{1.4}$$

and the variance of $\hat{y}_{fi}$ is

$$var(\hat{y}_{fi} - y_{fi}) = \tau_i^2 = \sigma_{2f} - \sigma_f\Sigma^{-1}\sigma_f + c_i g' D g, \tag{1.5}$$

where

$$c_i = A_i'(A'A)^{-1}A_i\,,\ g = B_f - B\Sigma^{-1}\sigma_f,\ and\, D = (B\Sigma^{-1}B')^{-1}.$$

If a certain consistent estimates of Σ_f and Σ are available then the predictors can be computed as

$$\hat{y}_{fi} \;=\; (A_i'\hat{\xi}B_f)' + \hat{\sigma}_f'\hat{\Sigma}^{-1}(Y_i - (A_i'\hat{\xi}B)'$$

As before, it is possible to construct an approximate prediction interval for y_{fi} (Naik (1990)). Since $(\hat{y}_{fi} - y_{fi})/\tau_i$ is distributed approximately as $N(0,1)$, we have the following theorem.

THEOREM **2** : *Approximate* $100(1-\alpha)\%$ *prediction interval for* y_{fi} *is given by*

$$(\hat{y}_{fi} - z_{\alpha/2}\hat{\tau}_i, \hat{y}_{fi} + z_{\alpha/2}\hat{\tau}_i),$$

where

$$\hat{y}_{fi} = (A_i'\hat{\xi}B_f)' + \hat{\sigma}_f'\hat{\Sigma}^{-1}(Y_i - (A_i'\hat{\xi}B)'),$$

$$\tau_i^2 = \hat{\sigma}_{2f} - \hat{\sigma}_f\hat{\Sigma}^{-1}\hat{\sigma}_f + c_i g' D g,$$

$$D = (B\hat{\Sigma}^{-1}B')^{-1},$$

assuming that a consistent estimate

$$\hat{\Sigma}_f = \begin{pmatrix} \hat{\Sigma} & \hat{\sigma}_f \\ \hat{\sigma}_f' & \hat{\sigma}_{2f} \end{pmatrix},$$

of Σ_f *is available.*

Note that an estimate of Σ in

$$\Sigma_f = \begin{pmatrix} \Sigma & \sigma_f \\ \sigma_f' & \sigma_{2f} \end{pmatrix},$$

can perhaps be obtained from the available data, but there is no data for estimating σ_f and σ_{2f}. This is so because these quantities refer to the covariances and variance of the future unobserved quantities. Hence, the formula for $\hat{y}_{fi}$ cannot be used in practice. To overcome this difficulty, certain structures on Σ (and hence on Σ_f) are assumed. We consider two special covariance structures, (1) equicorrelation structure in the second chapter, and (2) autoregressive structure in the third chapter.

1.3 Unbalanced data

So far, we have dealt with the situation where data are available for all the subjects or all occasions. This is called the balanced data. However, often in practice one does not have balanced data. In fact, unbalanced data are more common than balanced data. In general, there are two types of unbalanced data : (1) monotone or balanced incomplete data ; (2) unbalanced data. The first type of unbalancedness for n subjects at p occasions or time points has the following form:

Monotone or balanced incomplete

time

Subject	1	2	. . . .		p
1	y_{11}	y_{12}	. . .	y_{1p_1}	
2	y_{21}	y_{22}	. . . .		y_{2p_2}
.	.	.	. . . .		.
.	.	.	. . .		
n	y_{n1}	y_{n2}	. . . .		y_{np_n}

Note that for these type of data, observations are missing from the last few occasions for some of the subjects. Here p is the maximum of $p_i's$. These data occur naturally in practice. For example, some patients that are in a study at the beginning of an experiment might stop coming after some time for a variety of reasons.

We analyze these data by dividing the n subjects into g groups based on the number of measurements on each subject. Thus a model for balanced incomplete data or monotone data is as follows:

$$Y_{i_{n_i \times p_i}} = A_{i_{n_i \times g}} \xi_{g \times k} B_{i_{k \times p_i}} + \in_{i_{n_i \times p_i}}, i = 1, 2, ..., g. \quad (1.6)$$

Here g represents the number of groups in the model and each group has n_i units (not the same number in each group). In this model, rows of the error matrix $\in_i$ are independent each distributed as $N_{p_i}(0, \Sigma_i)$, where Σ_i is a $p_i \times p_i$ positive definite matrix. When $g = 1$ the model is similar to the growth curve model for the balanced data. Analysis of monotone data from a multivariate normal distribution has been discussed by Anderson (1957) and Bhargava (1975).

The second type of unbalancedness where data could be missing for any subject from any occasion (not just the last few occasions as in the monotone data case) can be represented such that the measurements are available at the checked ($\surd$) places.

Unbalanced

	time					
Subject	1	2	.	.	.	p
1	$\surd$	$\surd$	$-$	.	.	$\surd$
2	$-$	$\surd$	$-$	.	.	$\surd$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
n	$\surd$	$\surd$	$\surd$	.	.	$-$

A model for analyzing these data is

$$Y_{in_i \times p_i} = A_{in_i \times g} \xi_{g \times k} B_{k \times p} G_{i p \times p_i} + \in_{in_i \times p_i}, i = 1, 2, ..., g, \qquad (1.7)$$

where rows of $\in_i \sim N_{p_i}(0, G_i' \Sigma G_i)$, G_i is a matrix of 0's and 1's such that if the observations $i_1, ..., i_p$ are available then G_i has one in the $(k, i_k)^{th}$ position for $k = 1, 2, ..., p_i$ and zeros elsewhere. This model has an advantage over incomplete data model. In practice, one can have data missing from intermediate occasions as well. Thus this model is more general and can be used if data are missing from any occasion on any subject. Data with missing values are common in practice and a systematic approach to analyze such data is needed. There are some discussions about analyzing these types of data in the literature, for example, see Kleinbaum (1973) and Chi and Reinsel (1989).

There is a third type of unbalancedness that is possible. This is a more general case than the monotone or balanced incomplete data case. Here, for each group, data could be missing for some of the subjects from the last few occasions.

<u>Incomplete</u>

Group i

time

	1	2	. . .	.	p_i
1	y_{11}	y_{12}	. . .	$y_{1p_{i1}}$	
2	y_{21}	y_{22}	. . .	.	$y_{2p_{i2}}$
.	.	.	. . .	.	.
j	y_{j1}	y_{j2}	. . $y_{jp_{ij}}$		
.		.	. . .		
n_i	y_{n_i1}	y_{n_i2}	. . .	.	$y_{n_ip_{in_i}}$

A model for general incomplete data can be written as:

$$Y_{ij_{n_i \times p_{ij}}} = A_{i_{n_i \times g}} \xi_{g \times k} B_{ij_{k \times p_{ij}}} + \epsilon_{ij_{n_i \times p_{ij}}}, \; i = 1, 2, ..., g, j = 1, 2, ..., n_i. \qquad (1.8)$$

Here g represents the number of groups in the model and each group has n_i units (not the same number in each group). In this model, the rows of error matrix ϵ_{ij} are independent each distributed as $N_{p_{ij}}(0, \Sigma_{ij})$, where Σ_{ij} is a $p_{ij} \times p_{ij}$ positive definite matrix. When $g = 1$ the model reduces to the case where the number of measurements observed on each subject is different, that is, the first type of unbalancedness. The maximum likelihood analysis of the model is essentially the same as the much simpler incomplete (balanced) data model of equation (1.6).

Hence we will not pursue the analysis of the model (1.8) but restrict ourselves to (1.6).

1.3.1 Analysis of incomplete data under general covariance structure

Consider the model (1.6) , that is,

$$Y_{i_{n_i} \times p_i} = A_{i_{n_i} \times g} \xi_{g \times k} B_{i_{k \times p_i}} + \epsilon_{i_{n_i} \times p_i}, i = 1, 2, ..., g$$

with the rows of error matrix ϵ_i independent, each distributed as $N_{p_i}(0, \Sigma_i)$.

Case 1: When Σ_i is known for each i then the estimate of ξ can be obtained as

$$vec(\hat{\xi}) = [\sum_{i=1}^{g}(B_i \Sigma_i^{-1} B_i') \bigotimes (A_i' A_i)]^{-1} \sum_{i=1}^{g}(B_i \Sigma_i^{-1} \bigotimes A_i') vec(Y_i)$$

with the covariance matrix

$$Cov(vec(\hat{\xi})) = [\sum_{i=1}^{g}(B_i \Sigma_i^{-1} B_i') \bigotimes (A_i' A_i)]^{-1}$$

Case 2: When Σ_i are unknown then MLE of $vec(\hat{\xi})$ has the same form but the ML equations for Σ_i are intractable. Some suggestions about how to compute some alternate estimators of Σ_i are suggested by Crowder and Hand (1990).

1.4 Analysis of unbalanced data under general covariance structure

Consider the model (1.7). Kleinbaum (1973) called this model the Generalized

Growth Curve Multivariate (GGCM) Model. He has proposed best asymptotically normal (BAN) estimators for the parameters as follows

$$vec(\hat{\xi}) = [\sum_{i=1}^{g}(BG_i(G_i'\hat{\Sigma}G_i)^{-1}G_i'B')\bigotimes(A_i'A_i)]^{-1}\sum_{i=1}^{g}(BG_i(G_i'\hat{\Sigma}G_i)^{-1}\bigotimes A_i')vec(Y_i)$$

where $\hat{\Sigma}$ is a consistent estimator of Σ. He proposed an unbiased and consistent estimator of $\Sigma = ((\sigma_{rs}))$ by forming the usual pooled estimate of σ_{rs} as follows:

$$\hat{\sigma}_{rs} = \frac{1}{N_{rs} - R(D_{rs})}x_{rs}'[I - D_{rs}(D_{rs}'D_{rs})^{-1}D_{rs}']x_{rs} \; r, s = 1, 2, ..., p$$

where $N_{rs}(\geq 2)$ is the number of experimental units in which both response variates V_r and V_s are observed, $x_{rs}(N_{rs} \times 1)$ is the observation vector on V_r corresponding to those experimental units on which both response variates V_r and V_s are observed, $D_{rs}(N_{rs} \times g)$ is the design matrix, consisting of a row of A_i matrices, corresponding to x_{rs} and $V_1, V_2, ..., V_p$ are the p response variates corresponding to the p time points. It is important to note that this estimate of Σ is not necessarily positive definite for small samples.

In this thesis, we consider the two types of unbalanced data and using the ML method, under the above mentioned two types of structures for covariance matrix, show the analysis (estimation, prediction, and testing) of growth curves.

In the second and third chapters we consider the polynomial growth curve models under equicorrelation and autoregressive covariance structures respectively. In each chapter we consider the models for balanced data , for balanced incomplete data, and also for unbalanced data. For each covariance structure, the validity of assuming such a structure (goodness of fit of the model with the assumed structure) is discussed either by testing for the covariance structure or by using residual plots.

When the model (1.7) is used to analyze unbalanced data with an autoregressive covariance structure for Σ, the covariance matrix for the error, $G_i'\Sigma G_i$, becomes what is known as a Markov structure. Also, if balanced or balanced incomplete data are such that the repeated measurements are made at unequally spaced time points, then the natural covariance structure to use is Markov structure. Under this structure, the problem of predicting the unobserved portion of a partially observed vector V, when the unobserved portion is in the middle is also addressed in chapter 3. This result addresses some interesting questions relating to estimation of missing values in a time series context.

In each case computation of estimators have been illustrated either using the available software (for example SAS) or by writing FORTRAN programs.

The polynomial models can provide useful predictive information and may be the best approach if the growth information has been collected over a limited range of growth cycle. However, such models are often biologically unsatisfactory, as the parameters may not have a satisfactory biological interpretation. In that case we need to use nonlinear functions for fitting the growth processes. In the fourth chapter we consider the nonlinear growth curves and give computer programs using SAS software, to do the analysis of nonlinear models under variety of situations.

In summary, the new results developed in this dissertation can be found in chapters 2-4 and specifically in sections 2.3-2.5, 3.3-3.6, and 4.4-4.7.

Chapter 2

Analysis of Growth Curves Under Equicorrelation Covariance Structure

2.1 Introduction

In this chapter, we consider the growth curve models under equicorrelation covariance structure. First, in section 2.1, we describe the structure itself by providing its determinant and the inverse. In section 2.2, we review the results for estimation and prediction problems for balanced data model. In section 2.3, we consider a model for monotone or balanced incomplete data. Further we point out that this model can handle any type of unbalanced data under equicorrelation structure. In section 2.4, we consider some goodness of fit tests. Finally in the last section, we give the computer program for estimation of parameters under unbalanced data model.

2.2 Equicorrelation structure

A step towards dependence from independent errors is equicorrelated errors. Under the equicorrelated errors model each pair of components of the error vector has the same correlation coefficient, say ρ. This structure is appropriate when the measurements are all made under similar conditions. For measurements of the same type made in the same way, it is usual to assume variance homogeneity. Thus Σ, the covariance matrix of the error vector, has the following structure:

$$\begin{aligned} \Sigma &= \sigma^2[(1-\rho)I_p + \rho J] \\ &= \sigma^2 V(\rho) \\ &= \sigma^2 \begin{pmatrix} 1 & \rho & \rho & . & . & \rho \\ \rho & 1 & \rho & . & . & \rho \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & \rho \\ \rho & . & . & . & . & 1 \end{pmatrix} \end{aligned}$$

The determinant and inverse respectively are

$$det(\Sigma) = |\Sigma| = \sigma^2(1 + (p-1)\rho)(\sigma^2(1-\rho))^{p-1}$$

and

$$\Sigma^{-1} = \frac{1}{\sigma^2(1-\rho)}[I - \frac{\rho}{(1-(p-1)\rho)}J].$$

Note that the inverse is also of the same form as Σ. We need the restriction, $-\frac{1}{p-1} < \rho < 1$, to maintain the positive definitness of Σ. The determinant and inverse can also be written in the following form for the reparameterization,

$$\tau_1 = \sigma^2(1+(p-1)\rho) \; and \; \tau_2 = \sigma^2(1-\rho) :$$

$$\begin{aligned}
|\Sigma| &= \sigma^2(1+(p-1)\rho)(\sigma^2(1-\rho))^{p-1} = \tau_1\tau_2^{p-1} \\
\Sigma^{-1} &= \frac{1}{\sigma^2(1-\rho)}[I - \frac{\rho}{(1-(p-1)\rho)}J] \\
&= \frac{1}{\tau_2}I - \frac{(\tau_1-\tau_2)}{p\tau_1\tau_2}J.
\end{aligned}$$

This reparametrization is useful to obtain the MLEs of σ^2 and ρ.

2.3 Balanced data model

For balanced data, the growth curve model (1.1) is

$$Y_{n\times p} = A_{n\times g}\xi_{g\times k}B_{k\times p} + \in_{n\times p},$$

where Y is an observation matrix, ξ is a matrix of unknown parameters, A is a known matrix of rank $g < n$, and B is a known matrix of rank $k < p$. Further, rows of error matrix $\in$ are independent each distributed as $N_p(0, \Sigma)$, where Σ is a $p \times p$ matrix. Here we are considering the equicorrelation structure for Σ. We discuss only estimation and prediction in this section since testing is done in exactly the same manner as discussed in the first chapter.

2.3.1 Estimation

Lee (1988) has discussed estimation of σ^2, ρ and ξ using the ML method. We first show estimation of the parameters ρ and σ^2. The likelihood function has the following form

$$L(\sigma^2,\rho) = (2\pi)^{-\frac{np}{2}}|\Sigma|^{-\frac{n}{2}}exp[-\frac{1}{2}tr\Sigma^{-1}(Y - A\xi B)'(Y - A\xi B)]$$

Rewriting the log of likelihood function in terms of τ_1 and τ_2 after substituting $\hat{\xi}$ for ξ, we get

$$\begin{aligned} lnL &= -\frac{np}{2}ln(2\pi) - \frac{n}{2}ln|\Sigma| - \frac{1}{2}tr[\Sigma^{-1}(Y - A\hat{\xi}B)'(Y - A\hat{\xi}B)] \\ &= -\frac{np}{2}ln(2\pi) - \frac{n}{2}ln(\tau_1\tau_2^{p-1}) - \frac{1}{2}tr[\frac{1}{\tau_2}I - \frac{1}{p\tau_2}J + \frac{1}{p\tau_1}J]\hat{E}'\hat{E} \\ &= -\frac{np}{2}ln(2\pi) - \frac{n}{2}ln(\tau_1) - \frac{n(p-1)}{2}ln(\tau_2) - \frac{1}{2\tau_2}tr\hat{E}'\hat{E} + \frac{1}{2p\tau_2}trJ\hat{E}'\hat{E} \\ &\quad -\frac{1}{2p\tau_1}J\hat{E}'\hat{E} \end{aligned}$$

where $\hat{E} = (Y - A\hat{\xi}B)$.

Taking the partial derivative of lnL with respect to τ_1 and setting it to zero we get

$$\hat{\tau_1} = \frac{1}{np}trJ\hat{E}'\hat{E} = \frac{1}{np}1'\hat{E}'\hat{E}1,$$

where 1 is a vector of ones. Also

$$\frac{\partial lnL}{\partial \tau_2} = 0$$

$$\Rightarrow \hat{\tau}_2 = \frac{1}{n(p-1)}[tr\hat{E}'\hat{E} - \frac{1'\hat{E}'\hat{E}1}{p}].$$

Since

$$\tau_1 = \sigma^2(1 + (p-1)\rho) \ \& \ \tau_2 = \sigma^2(1-\rho)$$

$$\sigma^2 = \tau_2 + \frac{\tau_1 - \tau_2}{p} = \frac{\overline{p-1}\tau_2 + \tau_1}{p}$$

$$\rho = \frac{\tau_1 - \tau_2}{p\sigma^2} = \frac{\tau_1 - \tau_2}{\overline{p-1}\tau_2 + \tau_1}.$$

This gives us the ML estimates of σ^2 and ρ as :

$$\hat{\sigma^2} = \frac{\frac{1}{n}tr\hat{E}'\hat{E} - \frac{1}{np}tr1'\hat{E}'\hat{E}1 + \frac{1}{np}tr1'\hat{E}'\hat{E}1}{p} = \frac{1}{np}tr\hat{E}'\hat{E}$$

and

$$\begin{aligned}\hat{\rho} &= \frac{\frac{1}{np}tr1'\hat{E}'\hat{E}1 - \frac{1}{n(p-1)}tr\hat{E}'\hat{E} + \frac{1}{n(p-1)p}tr1'\hat{E}'\hat{E}1}{\frac{1}{np}tr1'\hat{E}'\hat{E}1 + \frac{1}{n}tr\hat{E}'\hat{E} - \frac{1}{np}tr1'\hat{E}'\hat{E}1} \\ &= \frac{1'\hat{E}'\hat{E}1 - tr\hat{E}'\hat{E}}{(p-1)tr\hat{E}'\hat{E}}.\end{aligned}$$

The ML estimate of ξ as before is

$$\hat{\xi} = (A'A)^{-1}A'Y\Sigma^{-1}B'(B\Sigma^{-1}B')^{-1}.$$

But for the equicorrelation structure

$$Y\Sigma^{-1}B'(B\Sigma^{-1}B')^{-1} = YB'(BB')^{-1},$$

assuming that the matrix B has the vector of ones as one of the rows. Then the ML estimate of ξ can be written as

$$\hat{\xi} = (A'A)^{-1}A'YB'(BB')^{-1}. \tag{2.1}$$

2.3.2 Prediction

The predictions of V_2 and y_f have been considered by Lee (1988), under the equicorrelation structure.

Prediction of V_2 :

Using (1.2), the predictor of V_2 given Y and V_1 is

$$\begin{aligned}\hat{V}_2 &= A_f\hat{\xi}B_2 + (V_1 - A_f\hat{\xi}B_1)\Sigma_{11}^{-1}\sigma_{12} \\ &= A_f\hat{\xi}B_2 + \frac{\rho}{1+(p-2)\rho}(V_1 - A_f\hat{\xi}B_1)1_{p-1},\end{aligned}$$

since

$$\sigma_{12} = \sigma^2\rho 1_{p-1} \ , \ \sigma_{22} = 1$$

$$\Sigma_{11} = \sigma^2[(1-\rho)I_{p-1} + \rho J_{p-1}]$$

and

$$\Sigma_{11}^{-1}\sigma_{12} = \frac{\rho}{(1+(p-2)\rho)}1_{p-1}.$$

Also we have

$$\sigma_{21}\Sigma_{11}^{-1}\sigma_{12} = \frac{\rho^2(p-1)\sigma^2}{(1+(p-2)\rho)}.$$

From Theorem 1 of chapter 1, we get that the approximate $100(1-\alpha)$ % prediction interval for V_2 (Naik (1990)) as

$$(\hat{V}_2 - z_{\alpha/2}\hat{\tau}, \hat{V}_2 + z_{\alpha/2}\hat{\tau}) \tag{2.2}$$

where

$$\begin{aligned}
\hat{\tau}^2 &= \hat{\sigma}_{22} - \hat{\sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\sigma}_{12} + c_f\hat{d}'\hat{D}\hat{d} \\
&= \hat{\sigma}^2\frac{(1-\hat{\rho})(1+(p-1)\hat{\rho})}{1+(p-2)\hat{\rho}} + c_f\hat{d}'\hat{D}\hat{d} \\
\hat{d} &= B_2 - B_1\hat{\Sigma}_{11}^{-1}\hat{\sigma}_{12} \\
&= B_2 - B_1\frac{\hat{\rho}}{1+(p-2)\hat{\rho}}, \\
\hat{D} &= (B\hat{\Sigma}^{-1}B')^{-1}.
\end{aligned}$$

Prediction of y_f :

Let us assume $N = n$ and $q = 1$ here also. For the i^{th} individual we have

$$E\begin{pmatrix} Y_i \\ y_{fi} \end{pmatrix} = \begin{pmatrix} (A_i'\xi B)' \\ (A_i'\xi B_f)' \end{pmatrix}, cov\begin{pmatrix} Y_i \\ y_{fi} \end{pmatrix} = \Sigma_f = \begin{pmatrix} \Sigma & \sigma_f \\ \sigma_f' & \sigma_{2f} \end{pmatrix},$$

where A_i' is the i^{th} row of A, B_f is a $k \times 1$ known vector similar to the matrix B, and Σ_f is a $(p+1) \times (p+1)$ covariance matrix. For the equicorrelation structure we have

$$\Sigma_f = \begin{pmatrix} \Sigma & \sigma_f \\ \sigma_f & \sigma_{2f} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & . & . & \rho \\ \rho & 1 & \rho & . & . & \rho \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & \rho \\ \rho & . & . & . & . & 1 \end{pmatrix}$$

This gives

$$\begin{aligned} \hat{\sigma}_f'\Sigma^{-1} &= \sigma^2\rho 1_p' \frac{1}{\sigma^2(1-\rho)}[I_p - \frac{\rho}{(1-(p-1)\rho)}1_p 1_p'] \\ &= \frac{\rho}{(1+(p-1)\rho)}1_p' \\ \Rightarrow \sigma_f'\Sigma^{-1}\sigma_f &= \frac{\rho}{(1+(p-1)\rho)}1_p'\sigma^2\rho 1_p = \frac{\rho^2 p\sigma^2}{(1+(p-1)\rho)} \\ \Rightarrow \sigma^2 - \sigma_f'\Sigma^{-1}\sigma_f &= \sigma^2\left((1 - \frac{p\rho^2}{(1+(p-1)\rho)} \right) = \sigma^2\frac{(1-\rho)(1+p\rho)}{1+(p-1)\rho} \end{aligned}$$

From (1.4) we get the conditional predictor of y_{fi} given Y as

$$\begin{aligned} \hat{y}_{fi} &= (A_i'\xi B_f)' + \sigma_f'\Sigma^{-1}(Y_i - (A_i'\xi B_f)') \\ &= (A_i'\hat{\xi}B_f)' + \frac{\hat{\rho}}{1+(p-1)\hat{\rho}}1_p'(Y_i - (A_i'\hat{\xi}B_f)') \end{aligned}$$

and the variance of $(\hat{y}_{fi} - y_{fi})$ is given by

$$\begin{aligned} var(\hat{y}_{fi} - y_{fi}) = \hat{\tau}_i^2 &= \hat{\sigma}_{2f} - \hat{\sigma}_f\hat{\Sigma}^{-1}\hat{\sigma}_f + c_i\hat{g}'\hat{D}\hat{g} \\ &= \frac{\hat{\sigma}^2(1-\hat{\rho})(1+p\hat{\rho})}{1+(p-1)\hat{\rho}} + c_i\hat{g}'\hat{D}\hat{g}, \\ \text{with} \quad c_i &= A_i'(A'A)^{-1}A_i \end{aligned}$$

$$\hat{g} = B_f - \frac{\hat{\rho}}{(1+(p-1)\hat{\rho})} B\, 1_p$$
$$\hat{D} = (B\hat{\Sigma}^{-1}B')^{-1}$$

Using Theorem 2 of chapter 1, we have the approximate $100(1-\alpha)\%$ prediction interval for y_{fi} (Naik (1990)) as

$$(\hat{y}_{fi} - z_{\alpha/2}\hat{\tau}_i, \hat{y}_{fi} + z_{\alpha/2}\hat{\tau}_i)$$

where $\hat{y}_{fi}$ and τ_i^2 are as above.

2.4 Unbalanced data model

The model for unbalanced data considered in (1.7) is as follows:

$$Y_{in_i \times p_i} = A_{in_i \times g}\xi_{g\times k}B_{k\times p}G_{ip\times p_i} + \in_{in_i\times p_i}, i = 1, 2, ..., g, \tag{2.3}$$

where the rows of the error matrix $\in_i \sim N_p(0, G_i'\Sigma G_i)$, Σ having a equicorrelation structure. But $\Sigma_i = G_i'\Sigma G_i = \sigma^2 V_i(\rho)$, since $G_i'G_i = I_i$ and $G_i'JG_i = J_i$, where $J_i = 1_{p_i}1_{p_i}'$, $J = 1_p 1_p'$ and $V_i(\rho) = [(1-\rho)I_{p_i} + \rho J_{p_i}]$. Thus Σ_i also has equicorrelation structure. Therefore, for equicorrelation structure, the model for unbalanced data is the same as the model for the monotone or balanced incomplete data. Hence we consider estimation, testing, and prediction problems for the latter model only.

The model for balanced incomplete data is

$$Y_{in_i \times p_i} = A_{in_i \times m}\xi_{m\times k}B_{ik\times p_i} + \in_{in_i\times p_i}, i = 1, 2, ..., g. \tag{2.4}$$

The rows of error matrix $\in_i$ are independent each distributed as $N_{p_i}(0, \Sigma_i)$, where Σ_i is a $p_i \times p_i$ matrix with equicorrelation structure.

2.4.1 Estimation

For this model a solution to the maximum likelihood equation for ξ is not directly attainable. Hence we make the following transformation.

Let y_i be $vec(Y_i)$ and ϵ_i be $vec(\in_i)$. Let

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_g \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_g \end{pmatrix} \quad and$$

$$X = \begin{pmatrix} B_1' \otimes A_1 \\ B_2' \otimes A_2 \\ \cdot \\ \cdot \\ \cdot \\ B_g' \otimes A_g \end{pmatrix}$$

Then the model (2.4) can be written as

$$y = Xvec(\xi) + \epsilon. \tag{2.5}$$

Let $N_1 = \sum_{i=1}^{g} n_i p_i$. Then y, X and ϵ are matrices of dimensions $N_1 \times 1$, $N_1 \times gk$, and $N_1 \times 1$ respectively. Further $\epsilon \sim N(0, D)$, where

$$D = \begin{pmatrix} \Sigma_1 \otimes I_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 \otimes I_2 & & 0 \\ \cdot & & \cdot & 0 \\ \cdot & & \cdot & 0 \\ \cdot & & \cdot & 0 \\ 0 & 0 & \cdots & \Sigma_g \otimes I_g \end{pmatrix}$$

We know that if the model is $y = X\beta + \epsilon$ with $\epsilon \sim N(0, V)$, then MLE of β is $\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$. Thus for the model in (2.5), we have the ML estimate of $vec(\xi)$ as

$$vec(\hat{\xi}) = (X'D^{-1}X)^{-1}X'D^{-1}y. \tag{2.6}$$

To simplify (2.6), note that

$$D^{-1} = \begin{pmatrix} (\Sigma_1 \otimes I_1)^{-1} & 0 & \cdots & 0 \\ 0 & (\Sigma_2 \otimes I_2)^{-1} & & 0 \\ \cdot & & \cdot & 0 \\ \cdot & & \cdot & 0 \\ \cdot & & \cdot & 0 \\ 0 & 0 & \cdots & (\Sigma_g \otimes I_g)^{-1} \end{pmatrix}.$$

Next, using the properties that (a) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ (if inverses exist), (b) $(A \otimes B)(C \otimes D) = AC \otimes BD$, we can show that

$$X'D^{-1}X = \sum_{i=1}^{g}(B_i\Sigma_i^{-1}B_i') \bigotimes (A_i'A_i)$$

and

$$X'D^{-1}y = \sum_{i=1}^{g}(B_i\Sigma_i^{-1} \bigotimes A_i')vec(Y_i)$$

Thus for the model (2.5), we have the ML estimate of ξ as

$$\begin{aligned} vec(\hat{\xi}) &= (X'D^{-1}X)^{-1}X'D^{-1}y \\ &= [\sum_{i=1}^{g}(B_i\Sigma_i^{-1}B_i') \bigotimes (A_i'A_i)]^{-1} \sum_{i=1}^{g}(B_i\Sigma_i^{-1} \bigotimes A_i')vec(Y_i) \end{aligned} \quad (2.7)$$

It can be shown for $\Sigma_i = \sigma^2 V_i(\rho)$, the equicorrelation structure, that the expression for $\hat{\xi}$ does not depend on ρ and σ^2. However, if we consider the model (1.8) the ML estimate $\hat{\xi}$ of ξ will depend on ρ. However, we will not consider this model further since computationally it does not create any complexity.

Now, to find the ML estimates of ρ and σ^2 for the incomplete data model (2.4), with Σ_i having equicorrelation structure, the likelihood function is

$$\begin{aligned} L &= \prod_{i=1}^{g}[(2\pi)^{-\frac{n_i p_i}{2}}|\Sigma_i|^{-\frac{n_i}{2}} exp[-\frac{1}{2}tr\Sigma_i^{-1}(Y_i - A_i\hat{\xi}B_i)'(Y_i - A_i\hat{\xi}B_i)]] \\ &= (2\pi)^{-(\sum_{i=1}^{g}(\frac{n_i p_i}{2}))} \prod_{i=1}^{g}((|\Sigma_i|^{-\frac{n_i}{2}})exp[-\frac{1}{2}\sum_{i=1}^{g} tr\Sigma_i^{-1}\hat{E}_i'\hat{E}_i] \end{aligned} \quad (2.8)$$

where $\hat{E}_i = (Y_i - A_i\xi B_i)$. Recall that

$$\begin{aligned} |\Sigma_i| &= \sigma^2(1 + (p_i - 1)\rho)(\sigma^2(1 - \rho))^{p_i - 1} \\ \text{and } \Sigma_i^{-1} &= \frac{1}{\sigma^2(1-\rho)}[I_i - \frac{\rho}{(1 - (p_i - 1)\rho)}J_i]. \end{aligned}$$

Substituting these values for $|\Sigma_i|$ and Σ_i^{-1}, and taking the logarithm of likelihood function we get,

$$\begin{aligned} lnL &= -\frac{N_1}{2}ln(2\pi) - \frac{N_1}{2}ln\sigma^2 - \sum_{i=1}^{g}\frac{n_i(p_i-1)}{2}ln(1-\rho) - \sum_{i=1}^{g}\frac{n_i}{2}ln(1+(p_i-1)\rho) \\ &\quad -\frac{1}{2\sigma^2(1-\rho)}[\sum_{i=1}^{g} tr\hat{E}_i'\hat{E}_i - \sum_{i=1}^{g}\frac{\rho}{1+(p_i-1)\rho}trJ_i\hat{E}_i'\hat{E}_i], \qquad (2.9) \end{aligned}$$

where $N_1 = \sum_{i=1}^{g} n_i p_i$ and $\hat{E}_i = Y_i - A_i\hat{\xi}B_i$. Now we can get the ML estimate of σ^2 by taking the partial derivative of lnL w.r.t. σ^2 and setting it to zero, i.e:

$$\frac{\partial lnL}{\partial\sigma^2} = 0 \Rightarrow$$
$$-\frac{N_1}{2\sigma^2} + \frac{1}{2\sigma^4(1-\rho)}[\sum_{i=1}^{g} tr\hat{E}_i'\hat{E}_i - \sum_{i=1}^{g}\frac{\rho}{1+(p_i-1)\rho}trJ_i\hat{E}_i'\hat{E}_i] = 0.$$

This gives us the ML estimator of σ^2 as

$$\hat{\sigma^2} = \frac{1}{N_1(1-\rho)}[\sum_{i=1}^{g} tr\hat{E}_i'\hat{E}_i - \sum_{i=1}^{g}\frac{\rho}{1+(p_i-1)\rho}trJ_i\hat{E}_i'\hat{E}_i]. \qquad (2.10)$$

Note that $\hat{\sigma}^2$ is an explicit function of ρ (as it can be seen) since $\hat{E}_i$ is independent of ρ. Putting $\hat{\sigma}^2$ in the log of likelihood function (2.9), we can write lnL as a function of ρ:

$$\begin{aligned} lnL(\rho) &= -\frac{N_1}{2}ln(2\pi) - \frac{N_1}{2}ln\hat{\sigma}^2 - \sum_{i=1}^{g}\frac{n_i(p_i-1)}{2}ln(1-\rho) \\ &\quad -\sum_{i=1}^{g}\frac{n_i}{2}ln(1+(p_i-1)\rho) - \frac{N_1}{2} \end{aligned}$$

We have to minimize $-lnL(\rho)$ w.r.t. ρ to get the MLE of ρ. Taking partial derivative of $-lnL(\rho)$ w.r.t. ρ and setting it to zero gives us an equation which is difficult to solve. Instead we use numerical minimization methods to get the MLE. Consider the following function of ρ:

$$f(\rho) = -2lnL = N_1(1 + ln(2\pi) + ln\hat{\sigma}^2) + (N_1 - n)ln(1-\rho) + \sum_{i=1}^{g} n_i ln(1 + (p_i - 1)\rho). \tag{2.11}$$

For given values of N_1, $\hat{\sigma}^2$, n_1, n_2,...,n_g,g,p_1, p_2,...,p_g, we have to minimize $f(\rho)$ w.r.t. ρ and find a solution. Suppose this solution is $\hat{\rho}$. Use this $\hat{\rho}$ in $\hat{\sigma}^2$ of (2.10). Find the new $\hat{\sigma}^2$ and again find the minimum of $f(\rho)$. We continue these iterative steps until the solution ($\hat{\rho}$) is stabilized. These types of algorithms have been successfully implemented in popular software, like SAS. We illustrate the computation on a data set in section 2.5.

2.4.2 Testing of hypothesis

To test the hypothesis

$$H_O : E_{q\times g}\xi_{g\times k}F_{k\times v} = C$$

one can use the test statistic

$$W_n = (G\hat{\beta} - P)'[G(\sum_{i=1}^{g}(B_i\hat{\Sigma}_i^{-1}B_i') \otimes (A_i'A_i)^{-1})^{-1}G']^{-1}(G\hat{\beta} - P) \tag{2.12}$$

where $G = F' \otimes E$, $P = vec(C)$ and $\hat{\beta} = vec(\hat{\xi})$. This statistic W_n can be shown to be approximately distributed as χ^2 with qv degrees of freedom.

2.4.3 Prediction

Prediction of V_2

Let $E(V_{i_{1\times p_i}}) = E(V_{1i}, V_{2i}) = (A_{if}\xi B_{1i}, A_{if}\xi B_{2i})$ with $B_{i_{k\times p_i}} = (B_{1i}, B_{2i})$ and $Cov(V_i) = \Sigma_i$ having equicorrelation structure. The interest here is to predict V_{2i} given V_{1i} and Y_i, i stands for the i^{th} group. Using (1.2), a predictor of V_{2i} given V_{1i} and Y_i is

$$\begin{aligned}\hat{V}_{2i} &= A_{if}\hat{\xi}B_{2i} + (V_{1i} - A_{if}\hat{\xi}B_{1i})\Sigma_{11i}^{-1}\sigma_{12i} \\ &= A_{if}\hat{\xi}B_{2i} + \frac{\hat{\rho}}{1+(p_i-2)\hat{\rho}}(V_{1i} - A_{if}\hat{\xi}B_{1i})1_{p_i-1}.\end{aligned}$$

Also

$$\begin{aligned}\sigma_{21i}\Sigma_{11i}^{-1}\sigma_{12i} &= \sigma^2\rho 1'_{p_i-1}\frac{\rho}{(1+(p_i-2)\rho)}1_{p_i-1} \\ &= \frac{\rho^2(p_i-1)\sigma^2}{(1+(p_i-2)\rho)}.\end{aligned}$$

Since $(\hat{V}_{2i} - V_{2i})/\tau_i$ is approximately distributed as $N(0,1)$, using Theorem 1 of the first chapter, we can construct the approximate $100(1-\alpha)\%$ prediction interval for V_{2i} as

$$(\hat{V}_{2i} - z_{\alpha/2}\hat{\tau}_i, \hat{V}_{2i} + z_{\alpha/2}\hat{\tau}_i) \tag{2.13}$$

where

$$\begin{aligned}\hat{\tau}_i^2 &= \frac{\hat{\sigma}^2(1-\hat{\rho})(1+(p_i-1)\hat{\rho})}{1+(p_i-2)\hat{\rho}} + c_{fi}\hat{d}_i'\hat{D}_i\hat{d}_i, \\ \text{with } c_{fi} &= A_{if}(A_i'A_i)^{-1}A_{if}', \\ \text{and } \hat{\mathrm{d}}_{\mathrm{i}} &= B_{2i} - B_{1i}1_{p_i-1}\frac{\hat{\rho}}{1+(p_i-2)\hat{\rho}}.\end{aligned}$$

Prediction of y_f

For the j^{th} individual in the i^{th} group we have

$$E\begin{pmatrix} Y_{ij} \\ y_{if_j} \end{pmatrix} = \begin{pmatrix} (A'_{ij}\xi B_i)' \\ (A'_{ij}\xi B_{if})' \end{pmatrix}, cov\begin{pmatrix} Y_{ij} \\ y_{if_j} \end{pmatrix} = \Sigma_{fi} = \begin{pmatrix} \Sigma_i & \sigma_{fi} \\ \sigma'_{fi} & \sigma_{2fi} \end{pmatrix},$$

where A'_{ij} is the j^{th} row of A_i, B_{if} is a $k \times 1$ known vector defined similarly as B_i.

Since

$$\Sigma_{fi} = \begin{pmatrix} \Sigma_i & \sigma_{fi} \\ \sigma_{fi} & \sigma_{2fi} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & . & . & \rho \\ \rho & 1 & \rho & . & . & \rho \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & \rho \\ \rho & . & . & . & . & 1 \end{pmatrix}$$

and

$$\hat{\sigma}'_{fi}\Sigma_i^{-1} = \frac{\rho}{(1+(p_i-1)\rho)}1'_{p_i},$$

the predictor of y_{if_j} is given by

$$\begin{aligned} \hat{y}_{if_j} &= (A'_{ij}\hat{\xi}B_{if})' + \hat{\sigma}'_{fi}\Sigma_i^{-1}(Y_{ij} - (A'_{ij}\hat{\xi}B_i)') \\ &= (A'_{ij}\hat{\xi}B_{if})' + \frac{\hat{\rho}}{1+(p_i-1)\hat{\rho}}1'_{p_i}(Y_{ij} - (A'_{ij}\hat{\xi}B_i)') \end{aligned}$$

Now, using

$$\begin{aligned} \hat{\sigma}'_{fi}\hat{\Sigma}_i^{-1}\hat{\sigma}_{fi} &= \frac{\hat{\rho}}{(1+(p_i-1)\hat{\rho})}1'_{p_i}\sigma^2\hat{\rho}1_{p_i} = \frac{\hat{\rho}^2 p_i\sigma^2}{(1+(p_i-1)\hat{\rho})} \\ \hat{\sigma}^2 - \hat{\sigma}'_{fi}\hat{\Sigma}_i^{-1}\hat{\sigma}_{fi} &= \hat{\sigma}^2\left((1 - \frac{p_i\hat{\rho}^2}{(1+(p_i-1)\hat{\rho})} \right) = \hat{\sigma}^2\frac{(1-\hat{\rho})(1+p_i\hat{\rho})}{1+(p_i-1)\hat{\rho}} \end{aligned}$$

we have

$$
\begin{aligned}
var(\hat{y}_{if_j} - y_{if_j}) = \hat{\tau}_{ij}^2 &= \sigma_{2fi} - \hat{\sigma}'_{fi}\hat{\Sigma}_i^{-1}\hat{\sigma}_{fi} + c_{ij}\hat{g}'_i\hat{D}_i\hat{g}_i \\
&= \frac{\hat{\sigma}^2(1-\hat{\rho})(1+p_i\hat{\rho})}{1+(p_i-1)\hat{\rho}} + c_{ij}\hat{g}'_i\hat{D}_i\hat{g}_i,
\end{aligned}
$$

with

$$
\begin{aligned}
c_{ij} &= A'_{ij}(A'_iA_i)^{-1}A_{ij}, \\
\hat{g}_i &= B_{if} - B_i 1_{p_i}\frac{\hat{\rho}}{1+(p_i-1)\hat{\rho}}, \\
\hat{D}_i &= (B_i\hat{\Sigma}_i^{-1}B'_i)^{-1}.
\end{aligned}
$$

Since $(\hat{y}_{if_j} - y_{if_j})/\tau_{ij}$ is approximately distributed as $N(0,1)$ using Theorem 2 of chapter 1, we can construct the approximate $100(1-\alpha)$ % prediction interval for y_{if_j} as

$$(\hat{y}_{if_j} - z_{\alpha/2}\hat{\tau}_{ij}, \hat{y}_{if_j} + z_{\alpha/2}\hat{\tau}_{ij}).$$

Example 2.3.1

Just to illustrate that these methods can be implemented in practice, we provide an example with computer program. 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 shown in Table 2.1. Suppose for the treatment group we don't have data at time points 1.5 and 5. So

Table 2.1: Plasma Inorganic Phosphate data

	Hours after glucose challenge							
Patient	0	0.5	1	1.5	2	3	4	5
				Control				
1	4.3	3.3	3.0	2.6	2.2	2.5	3.4	4.4
2	3.7	2.6	2.6	1.9	2.9	3.2	3.1	3.9
3	4.0	4.1	3.1	2.3	2.9	3.1	3.9	4.0
4	3.6	3.0	2.2	2.8	2.9	3.9	3.8	4.0
5	4.1	3.8	2.1	3.0	3.6	3.4	3.6	3.7
6	3.8	2.2	2.0	2.6	3.8	3.6	3.0	3.5
7	3.8	3.0	2.4	2.5	3.1	3.4	3.5	3.7
8	4.4	3.9	2.8	2.1	3.6	3.8	4.0	3.9
9	5.0	4.0	3.4	3.4	3.3	3.6	4.0	4.3
10	3.7	3.1	2.9	2.2	1.5	2.3	2.7	2.8
11	3.7	2.6	2.6	2.3	2.9	2.2	3.1	3.9
12	4.4	3.7	3.1	3.2	3.7	4.3	3.9	4.8
13	4.7	3.1	3.2	3.3	3.2	4.2	3.7	4.3
				Obese				
1	4.3	3.3	3.0	2.6	2.2	2.5	2.4	3.4
2	5.0	4.9	4.1	3.7	3.7	4.1	4.7	4.9
3	4.6	4.4	3.9	3.9	3.7	4.2	4.8	5.0
4	4.3	3.9	3.1	3.1	3.1	3.1	3.6	4.0
5	3.1	3.1	3.0	2.6	2.6	1.9	2.3	2.7
6	4.8	5.0	2.9	2.8	2.2	3.1	3.5	3.6
7	3.7	3.1	3.3	2.8	2.9	3.6	4.3	4.4
8	5.4	4.7	3.9	4.1	2.8	3.7	3.5	3.7
9	3.0	2.5	2.3	2.2	2.1	2.6	3.2	3.5
10	4.9	5.0	4.1	3.7	3.7	4.1	4.7	4.9
11	4.8	4.3	4.7	4.6	4.7	3.7	3.6	3.9
12	4.4	4.2	4.2	3.4	3.5	3.4	3.9	4.0
13	4.9	4.3	4.0	4.0	3.3	4.1	4.2	4.3
14	5.1	4.1	4.6	4.1	3.4	4.2	4.4	4.9
15	4.8	4.6	4.6	4.4	4.1	4.0	3.8	3.8
16	4.2	3.5	3.8	3.6	3.3	3.1	3.5	3.9
17	6.6	6.1	5.2	4.1	4.3	3.8	4.2	4.8
18	3.6	3.4	3.1	2.8	2.1	2.4	2.5	3.5
19	4.5	4.0	3.7	3.3	2.4	2.3	3.1	3.3
20	4.6	4.4	3.8	3.8	3.8	3.6	3.8	3.8

$n_1 = 13, n_2 = 20$, $m = 2, k = 5, p = 8$, $p_1 = 8$ and $p_2 = 6$. A_1 and A_2 are of order 13×2 and 20×2 respectively. Using SAS, we fit the unbalanced model on this data set (PROGRAM 1). The results of this program are $\hat{\rho} = 0.33$, $\hat{\sigma}^2 = 0.61$ and

$$\hat{\xi} = \begin{pmatrix} 3.2 & -0.01 & 0.01 \\ 3.44 & 0.64 & -0.15 \end{pmatrix}$$

2.5 Goodness of fit tests

Consider the model (1.1), $Y = A\xi B + \in$, with the usual assumptions. One of the problems we face in practice is to determine an appropriate structure for the covariance matrix Σ. If there is only one group in the model, one can determine the appropriate structure for Σ by selecting several structures and testing hypotheses about these structures. For example, suppose we want to see whether equicorrelation structure fits well for the data. Using the data under the above model we test $H_0 : \Sigma = \sigma^2 V(\rho)$ (the equicorrelation structure) against $H_a : \Sigma$ is a positive definite matrix. This is a standard problem in multivariate analysis, perhaps under simpler multivariate linear model situation. See Anderson (1984) or Siotani, Hayakawa, and Fujikoshi (1985). Similarly, tests for testing other structures can be developed using the likelihood ratio test.

If there are several, say g, groups, then these type of testing of hypothesis problems become difficult. This is because, different groups may have different covariance matrices $\Sigma_i, i = 1, 2, ..., g$ and we may need to first test $H_0 : \Sigma_1 = \Sigma_2 = ... = \Sigma_g$ versus H_a : The covariance matrices Σ_i's are different. If H_0 is rejected and

2 or more structures are being considered for Σ_i, a variety of situations may occur in practice due to the fact that, only a certain parameters determining the structures may be different for different groups. The problem is further complicated if the data are unbalanced. For example, suppose we want to test $H_0 : \Sigma_{i_{p_i \times p_i}} = \sigma^2 V_i(\rho)$ against $H_1 : \Sigma_i$ is a positive definite submatrix of $\Sigma_{p \times p}$. In this case, finding the MLE of Σ_i under H_1 may be very difficult even if it exists.

In this section, we assume due to other practical considerations as well, that the structure for group i is $\Sigma_{i_{p_i \times p_i}} = \sigma_i^2 V(\rho_i)$, $i = 1, 2, ..., g$. The problem then is to test various hypotheses,

$H_{01} : \sigma_1^2 = \sigma_2^2 = ... = \sigma_g^2$, $\rho_1 = \rho_2 = ... = \rho_g$

$H_{02} : \sigma_1^2 = \sigma_2^2 = ... = \sigma_g^2 = \sigma^2$, ρ_i's are different

$H_{03} : \rho_1 = \rho_2 = ...\rho_g = \rho$, σ_i^2's are different

against an appropriate alternative hypothesis.

In a recent paper, Viana (1994) has derived the MLE's of the parameters under H_{02} and H_{03}. The MLE's of σ_i^2, ρ_i, $i = 1, 2, ..., g$ under the alternative $\Sigma_i = \sigma_i^2 V(\rho_i)$ are easily obtained using the MLE's for balanced data. The MLE's of σ^2 and ρ under H_{01} are derived in section 2.2.1. Using these MLE's various hypotheses can be tested using the likelihood ratio criterion.

If we consider an autoregressive structure instead of an equicorrelation structure, then tests for all of the above various hypotheses have been considered by Lee (1991).

Residual Plots

One can also suggest certain graphical procedures for visual and easy validation of the assumed structure. One graphical procedure we may suggest is as follows

Define $y = \text{vec}(Y)$, $\epsilon = \text{vec}(\in)$ and $\beta = \text{vec}(\xi)$. Then model (1.1) can be rewritten as

$$y = (B' \bigotimes A)\beta + \epsilon = X\beta + \epsilon \text{ (say)}$$

Further, under the assumption of rows of $\in$ being independently distributed as multivariate normal with mean 0 and covariance matrix Σ, we have ϵ to be distributed as np variate normal with mean 0 and covariance matrix $\Omega = \Sigma \otimes I$. Using standard generalized least squares theory we know that $\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$. The vectors of predicted values and the residuals respectively are given by $\hat{y} = X\hat{\beta}$ and $\hat{\epsilon} = y - \hat{y}$. It is standard practice in regression analysis to plot the respective elements of $\Omega^{-\frac{1}{2}}\hat{y}$ and $\Omega^{-\frac{1}{2}}\hat{\epsilon}$ as points in a plane. If the points follow a random pattern in the plane then the model is assumed to be valid. We adopt this plot for validating the assumed structure. Suppose Σ_i is the covariance matrix for the i^{th} group and it has an equicorrelation structure then the covariance of ϵ is of the form $\Omega = \text{Diag}(\Sigma_1, \Sigma_2, ..., \Sigma_g)$ instead of $I \otimes \Sigma$ and Ω is a function of the two parameters σ^2 and ρ. We estimate σ^2 and ρ using the ML method and estimate Ω. Suppose the estimator of Ω is $\hat{\Omega}$, then we compute $\hat{\Omega}^{-\frac{1}{2}}\hat{y}$ and $\hat{\Omega}^{-\frac{1}{2}}\hat{\epsilon}$ and plot the points in a plane as described. A program for plotting the residuals for a data set under the equicorrelation structure is provided in section 2.5.

2.6 Computer program

PROGRAM 1

```
options linesize=70;

Title 'Growth curve analysis of plasma data under equicorrelation structure';

data plasma;

infile 'plasma.data';

input subj1 y1 y2 y3 y4 y5 y6 y7 y8 subj2 y9 y10 y11 y12 y13 y14 y15 y16;

data new1(keep=subj1 y1-y8) new2(keep=subj2 y9-y16);

set plasma;

   /* Creating vec(Y) */

data new11;

set new1;

y=y1;time=1;output;

y=y2;time=2;output;

y=y3;time=3;output;

y=y4;time=4;output;

y=y5;time=5;output;

y=y6;time=6;output;

y=y7;time=7;output;

y=y8;time=8;output;

drop y1-y8;
```

```
data new22;

set new2;

y=y9;time=1;output;

y=y10;time=2;output;

y=y11;time=3;output;

y=y13;time=5;output;

y=y14;time=6;output;

y=y15;time=7;output;

drop y9-y16;

data b;

set new11(in=innew1 rename=(subj1=subj))

new22(rename=(subj2=subj));

if subj='.' then delete;

   /* We are creating the X matrix in this proc step. */

proc iml;

B = (1 1 1 1 1 1 1 1 ,

0 0.5 1 1.5 2 3 4 5 ,

0 0.25 1 2.25 4 9 16 25 );

G1 = (1 0 0 0 0 0 0 0 ,

0 1 0 0 0 0 0 0 ,

0 0 1 0 0 0 0 0 ,

0 0 0 1 0 0 0 0 ,
```

0 0 0 0 1 0 0 0 ,

0 0 0 0 0 1 0 0 ,

0 0 0 0 0 0 1 0 ,

0 0 0 0 0 0 0 1);

G2 = (1 0 0 0 0 0 ,

0 1 0 0 0 0 ,

0 0 1 0 0 0 ,

0 0 0 0 0 0 ,

0 0 0 1 0 0 ,

0 0 0 0 1 0 ,

0 0 0 0 0 1 ,

0 0 0 0 0 0);

A1 = (1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0 ,

1 0);

A2 = (0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

0 1 ,

```
0 1 ,
0 1 ,
0 1 );
Dz1 = (G1` * B`) @ A1;
Dz2 = (B2` * G`) @ A2;
D = Dz1//Dz2;
varnames=x1,x2,x3,x4,x5,x6;
create dd from D([colname = varnames]);
append from D;
close dd;
data last;
merge dd b;
```

/* We are fitting the model $y = Xvec(\xi) + \epsilon$ with equicorrelation covariance structure. */

```
proc mixed data=last method=ml;
class subj;
model y= x1 x2 x3 x4 x5 x6/s noint;
repeated intercept diag/subject=subj r;
make 'Predicted' out=pred;
make 'R' out=rmatrix;
run;
proc iml;
```

```
use rmatrix;
read all var{coll} into col;
sig1=col[1,1];
sig2=col[2,1];
rh=sig2/sig1;
rho=J(224,1,1);
rho=rh*rho;
varnames=rho;
create rhnew from rho([colname = varnames]);
append from rho;
close rhnew;
data final;
merge rhnew pred;
    /* In the following proc step we are transforming the predicted and residual
vectors.*/
proc iml;
use final;
read all var{rho,Pred,Resid} into aa;
rho1=aa[1,1];
g1=J(8,8,1.0);
g2=J(6,6,1.0);
c=1.0/sqrt(1.0-rho1);
```

```
bet1=(-rho1*8)/((1.0-rho1)*(1.0+7*rho1));

bet2=(-rho1*6)/((1.0-rho1)*(1.0+5*rho1));

d1=(1.0/8)*(-c+sqrt(c+bet1));

g1=c*I(8)+d1*J(8,8,1.0);

d2=(1.0/6)*(-c+sqrt(c+bet2));

g2=c*I(6)+d1*J(6,6,1.0);

pred1=(I(13) @ g1)*aa[1:104,2];

pred2=(I(20) @ g2)*aa[105:224,2];

tpred=pred1//pred2;

resd1=(I(13) @ g1)*aa[1:104,3];

resd2=(I(20) @ g2)*aa[105:224,3];

tresd=resd1//resd2;

zz=tpred || tresd;

varnames=tpred,tresd;

create last1 from zz(|colname = varnames|);

append from zz;

close last1;

    /*THE RESIDUAL PLOT*/

/* Plotting the transformed residuals on transformed predicted values.*/

proc plot data=last1;

plot tresd*tpred;
```

Chapter 3

Analysis of Growth Curves Under Autoregressive Covariance Structure

3.1 Introduction

In this chapter we consider the growth curve models for different situations with the autoregressive structure for the covariance matrix. In almost all cases, there will be correlation between the repeated measurements (measurements on the same subject or experimental unit) taken at different time points. It is also likely that there is a decay in the correlation with increasing time distance between the repeated measurements. In this case the natural structure for the correlation matrix is the autoregressive correlation structure. Moreover this structure enables us to model unbalanced (missing value) data. Hence in this chapter we restrict ourselves to the first order autoregressive correlation structure. While modeling the unbalanced data or the data observed at unequally spaced times, the first order autoregressive structure leads to the Markov structure as we will see later. In the next section, we present the autoregressive structure and its variance and the determinant. In section 3.3, we study the growth curve model for the balanced case and give the maximum likelihood estimators of the parameters involved (Lee (1988), Fujikoshi et al. (1990)) and give a test for testing H_0. We also consider the prediction of y_f given Y and prediction of V_2 given V_1 and Y.

Another prediction problem that may be of interest to scientists is the prediction of unobserved portion of a vector, when the unobserved portion is in the middle. For example, predict V_i, given V_1, V_2, and Y, where $V' = (V_1', V_i', V_2')$. We address this problem also in section 3.3.

In section 3.4, we consider the model to accommodate monotone or balanced incomplete data (that is, data that are missing only at the end) and derive the maximum likelihood estimators. Further, we consider the two prediction problems. The testing problem is similar to the one considered in the second chapter. Finally, in section 3.5, a model to analyze missing or unbalanced data (that is, data that are missing at any time of the observation) is considered.

3.2 Autoregressive structure

A widely used time series model is the autoregressive process and the structure of the correlation of a first order autoregressive process is called autoregressive correlation structure. In applications, we take a $p \times p$ autoregressive covariance matrix as

$$\begin{aligned}
\Sigma &= \sigma^2(\rho^{|i-j|}), i, j = 1, ..., p \qquad (3.1) \\
&= \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & . & . & \rho^{p-1} \\ \rho & 1 & \rho & \rho^2 & \rho^3 & . & \rho^{p-2} \\ & & . & & & & \\ & & & & . & & \\ \rho^{p-1} & \rho^{p-2} & & & & \rho & 1 \end{pmatrix} \\
&= \sigma^2 V(\rho).
\end{aligned}$$

The inverse of V can be written as

$$V^{-1} = (1 - \rho^2)^{-1}[\rho^2 C_1 - 2\rho C_2 + I_p],$$

where

$$C_1 = \begin{pmatrix} 0 & & & & & \\ & 1 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & 1 & \\ & & & & & & 0 \end{pmatrix}, C_2 = \frac{1}{2} \begin{pmatrix} 0 & 1 & & & & & \\ 1 & 0 & \cdot & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & 1 \\ & & & & & 1 & 0 \end{pmatrix} \qquad (3.2)$$

Also

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

3.3 Balanced data model

The growth curve model under the balanced case is

$$Y_{n \times p} = A_{n \times m}\xi_{m \times k}B_{k \times p} + \in_{n \times p},$$

where Y is an observation matrix, ξ is a matrix of unknown parameters, A is a known matrix of rank $m < n$, and B is a known matrix of rank $k < p$. Further, rows of error matrix $\in$ are independent each distributed as $N_p(0, \Sigma)$, where Σ is a $p \times p$ positive definite matrix.

The log of the likelihood function

$$L = (2\pi)^{-\frac{np}{2}}|\Sigma|^{-\frac{n}{2}}exp[-\frac{1}{2}tr\Sigma^{-1}(Y - A\xi B)'(Y - A\xi B)]$$

of ξ, ρ and σ^2 based on Y can be written as

$$lnL(\xi, \sigma^2, \rho) = -1/2[(np)ln\sigma^2 + n(p-1)ln(1-\rho^2) + (np)ln2\pi$$

$$+\sigma^2(1-\rho^2)^{-1}tr[(\rho^2 C_1 - 2\rho C_2 + I_p)(Y - A\xi B)'(Y - A\xi B)]$$

Taking the derivative of this equation with respect to σ^2 and setting it to zero we have

$$-\frac{np}{2\sigma^2} + \frac{1}{2\sigma^4(1-\rho^2)}tr[(\rho^2 C_1 - 2\rho C_2 + I_p)(Y - A\xi B)'(Y - A\xi B)] = 0$$

$$\Rightarrow \hat{\sigma^2} = \frac{1}{np(1-\rho^2)}tr[(\rho^2 C_1 - 2\rho C_2 + I_p)(Y - A\xi B)'(Y - A\xi B)]$$

Next, rewriting lnL as

$$lnL = -\frac{np}{2}ln2\pi - \frac{np}{2}ln\sigma^2 - \frac{n(p-1)}{2}ln(1-\rho^2) - \frac{\rho^2}{2\sigma^2(1-\rho^2)}trC_1(Y - A\xi B)'(Y - A\xi B)$$

$$+\frac{\rho}{\sigma^2(1-\rho^2)}trC_2(Y - A\xi B)'(Y - A\xi B) - \frac{1}{2\sigma^2(1-\rho^2)}tr(Y - A\xi B)'(Y - A\xi B)$$

and differentiating with respect to ρ we obtain

$$\frac{\partial lnL}{\partial \rho} = \frac{n(p-1)\rho}{(1-\rho^2)} - \frac{\rho}{\sigma^2(1-\rho^2)^2}trC_1(Y - A\xi B)'(Y - A\xi B) + \frac{1+\rho^2}{\sigma^2(1-\rho^2)^2}$$

$$trC_2(Y - A\xi B)'(Y - A\xi B) - \frac{\rho}{\sigma^2(1-\rho^2)^2}tr(Y - A\xi B)'(Y - A\xi B),$$

and

$$\frac{\partial lnL}{\partial \rho} = 0$$

gives

$$n(p-1)\rho(1-\rho^2)\sigma^2 - \rho tr[C_1(Y-A\xi B)'(Y-A\xi B)] + (1+\rho^2)$$

$$tr[C_2(Y-A\xi B)'(Y-A\xi B)] - \rho tr(Y-A\xi B)'(Y-A\xi B) = 0 \qquad (3.3)$$

Let $R = \frac{(Y-A\xi B)'(Y-A\xi B)}{n-m}$ *and* $a_i = tr(C_i R)$. Then we have the ML estimator of σ^2 as

$$\begin{aligned}\hat{\sigma^2} &= \frac{1}{np(1-\hat{\rho}^2)}[(n-m)\hat{\rho}^2 a_1 - 2(n-m)\hat{\rho} a_2 + (n-m)a_3] \\ &= \frac{n-m}{np(1-\hat{\rho}^2)}[a_1\hat{\rho}^2 - 2a_2\hat{\rho} + a_3]\end{aligned}$$

To get the ML estimator of ρ, let us substitute the value of $\hat{\sigma}^2$, R and a_i in (3.3). We get the following equation:

$$n(p-1)\rho(1-\rho^2)\frac{n-m}{np(1-\rho^2)}[a_1\rho^2 - 2a_2\rho + a_3] - \rho(n-m)a_1 +$$

$$(1+\rho^2)(n-m)a_2 - \rho(n-m)a_3 = 0$$

$$\Rightarrow \frac{p-1}{p}\rho[a_1\rho^2 - 2a_2\rho + a_3] - \rho a_1 + (1+\rho^2)a_2 - \rho a_3 = 0$$

$$\Rightarrow (p-1)a_1\rho^3 - 2(p-1)a_2\rho^2 + (p-1)a_3\rho - pa_1\rho$$

$$+pa_2 + pa_2\rho^2 - pa_3\rho = 0$$

Thus, ML estimate of ρ is the solution of the following cubic equation

$$(p-1)a_1\hat{\rho}^3 - (p-2)a_2\hat{\rho}^2 - (pa_1 + a_3)\hat{\rho} + pa_2 = 0.$$

We summarize the maximum likelihood (ML) estimates of ξ, σ^2, and ρ in the following theorem due to Fujikoshi et al. (1990). Also see Lee (1988).

THEOREM 3 *The ML estimates of ξ, σ^2 and ρ in the growth curve model with autoregressive covariance structure are the solutions of the following equations:*

$$(I)\ \hat{\xi} = \xi(\hat{\rho}) = (A'A)^{-1}A'Y\hat{V}^{1}B'(B\hat{V}^{1}B')^{-1}$$

$$(II)\ \hat{\sigma}^2 = \sigma^2(\hat{\xi},\hat{\rho}) = \frac{n-m}{n}(1-\hat{\rho}^2)^{-1}[a_1\hat{\rho}^2 - 2a_2\hat{\rho} + a_3]$$

$$(III)\ (p-1)a_1\hat{\rho}^3 - (p-2)a_2\hat{\rho}^2 - (pa_1 + a_3)\hat{\rho} + pa_2 = 0.$$

where

$$\hat{V} = V(\hat{\rho}), \hat{V}^{-1} = (1-\hat{\rho}^2)^{-1}[\hat{\rho}^2 C_1 - 2\hat{\rho}C_2 + C_3],$$

$$R = \frac{(Y - A\xi B)'(Y - A\xi B)}{n-m},\ a_i = tr(C_i R), i = 1,2,3,$$

with C_1 & C_2 as defined in equation 3.2 and $C_3 = I_p$.

Fujikoshi et al. (1990) have studied the asymptotic properties of the ML estimators $\hat{\xi}$, $\hat{\sigma^2}$ and $\hat{\rho}$ and have established that, for fixed p and g, as $n \to \infty$,

$$vec(((A'A)^{\frac{1}{2}}(\hat{\xi} - \xi))') \sim N_{gk}(0, I \otimes (B\hat{\Sigma}^{-1}B')^{-1}), \tag{3.4}$$

where $\hat{\Sigma} = \hat{\sigma}^2 V(\hat{\rho})$.

3.3.1 Testing of hypothesis

Consider the problem of testing $H_0 : E\xi F = C$ versus $H_1 : E\xi F \neq C$, for known matrices E, F and C of order $q \times g$, $k \times \nu$ and $q \times \nu$ respectively, with E of rank q and F of rank ν.

Using the asymptotic distribution of $\hat{\xi}$ given in (3.4) a test criterion can be easily constructed. We note that the asymptotic variance of $E\hat{\xi}F$ using (3.4) is

$$var(vec(E\hat{\xi}F)') = (E(A'A)^{-1}E') \otimes F'(B\hat{\Sigma}^{-1}B')^{-1}F.$$

Hence a test criterion to test H_0 is to use the trace of

$$(E\hat{\xi}F - C)'(E(A'A)^{-1}E')^{-1}(E\hat{\xi}F - C)(F'(B\hat{\Sigma}^{-1}B')^{-1}F)^{-1} \qquad (3.5)$$

whose asymptotic distribution is χ^2 with $q\nu$ degrees of freedom. The reader is referred to Azzalini (1987) for a similar result. Other criteria based on the eigenvalues of the matrix (3.5) can also be used to test H_0.

3.3.2 Prediction

In the following we present minimum (predictive) mean square predictors for y_{if}, the future observation for the ith individual, and V_2, the unobserved portion of the vector $V'(= (V_1', V_2'))$. For simplicity of presentation we assume that both y_{if} and V_2 are scalars. However, in general, these can be vectors.

Prediction of V_2:

The minimum mean square predictor of V_2 given V_1 and Y is obtained as follows: Let $E(V) = E(V_1, V_2) = (A_f\xi B_1, A_f\xi B_2)$, where A_f is $1 \times g$ known vector, $B =$

(B_1, B_2), and $cov(V) = \Sigma$. Then $\hat{V}_2 = A_f\hat{\xi}B_2 + (V_1 - A_f\hat{\xi}B_1)\Sigma_{11}^{-1}\sigma_{12}$ with the variance, $\tau^2 = var(\hat{V}_2 - V_2) = \sigma_{22} - \sigma_{21}\Sigma_{11}^{-1}\sigma_{12} + c_f d'Dd$,

where

$$cov\,(V) = \Sigma = \begin{bmatrix} \Sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & . & . & \rho^{p-1} \\ \rho & 1 & \rho & \rho^2 & \rho^3 & . & \rho^{p-2} \\ & & . & & & & \\ & & & & . & & \\ \rho^{p-1} & \rho^{p-2} & & & & \rho & 1 \end{pmatrix},$$

$c_f = A_f(A'A)^{-1}A'_f$, and $d = B_2 - B_1\Sigma_{11}^{-1}\sigma_{12}$.

For the autoregressive structure,

$$\sigma_{12} = \sigma^2(\rho^{p-1}, \rho^{p-2}, ..., \rho)'\ ,\ \sigma_{22} = \sigma^2$$

$$\Sigma_{11} = \sigma^2 V(\rho)_{p-1 \times p-1}$$

$$\begin{aligned} \Sigma_{11}^{-1}\sigma_{12} &= (1-\rho^2)^{-1}[\rho^2 C_1 - 2\rho C_2 + I_{p-1}](\rho^{p-1}, \rho^{p-2}, ..., \rho, 1)' \\ &= (1-\rho^2)^{-1} \begin{bmatrix} 1 & -\rho & 0 & . & . & . & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & . & . & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & . & 0 \\ . & . & . & & & & \\ . & . & . & & & & \\ 0 & 0 & . & . & . & -\rho & 1 \end{bmatrix} \begin{pmatrix} \rho^{p-1} \\ \rho^{p-2} \\ \rho^{p-3} \\ . \\ . \\ \rho \end{pmatrix} \end{aligned}$$

$$= (1-\rho^2)^{-1} \begin{pmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ \rho - \rho^3 \end{pmatrix}$$

$$= (0'_{p-2}, \rho)'.$$

Also,

$$\begin{aligned} \sigma_{21}\Sigma_{11}^{-1}\sigma_{12} &= \sigma_{12}(1-\rho^2)^{-1}[\rho^2 C_1 - 2\rho C_2 + I_{p-1}](\rho^{p-1}, \rho^{p-2}, ..., \rho, 1)' \\ &= \sigma^2(\rho^{p-1}, \rho^{p-2}, ..., \rho)(0'_{p-2}, \rho)' \\ &= \sigma^2\rho^2 \end{aligned}$$

and

$$\sigma_{22} - \sigma_{12}\Sigma_{11}^{-1}\sigma_{12} = \sigma^2 - \sigma^2\rho^2 = \sigma^2(1-\rho^2)$$

Thus we have

$$\begin{aligned} \hat{V}_2 &= A_f\hat{\xi}B_2 + (V_1 - A_f\hat{\xi}B_1)(0'_{p-2}, \rho)' \\ \tau^2 &= var(\hat{V}_2 - V_2) = \sigma^2(1-\rho^2) + c_f d' D d, \\ d &= B_2 - B_1(0'_{p-2}, \rho). \end{aligned}$$

Using the maximum likelihood estimates of ρ and σ^2 that are obtained using Y, we can compute the predictor and its variance. Note that the observed portion of V

is used only for prediction purposes and not for estimation of the parameters. One can suggest an iterative procedure for prediction of V_2 utilizing V_1 for estimation as well as prediction. This procedure is discussed later in this section.

Prediction of y_f:

For the i^{th} individual we have

$$E\begin{pmatrix} Y_i \\ y_{if} \end{pmatrix} = \begin{bmatrix} (A_i'\xi B)' \\ (A_i'\xi B_f)' \end{bmatrix}, \ and\ cov \begin{pmatrix} Y_i \\ y_{if} \end{pmatrix} = \Sigma_f = \begin{bmatrix} \Sigma & \sigma_f \\ \sigma_f' & \sigma_{2f} \end{bmatrix},$$

where Σ_f is a $(p+1) \times (p+1)$ positive definite matrix, A_i' is the i^{th} row of A and B_f is a $k \times 1$ known vector. Using this and the formula of conditional expectation of y_{if} given Y_i , the predictor is given by

$$\hat{y}_{fi} = (A_i'\hat{\xi}B_f)' + \sigma_f'\Sigma^{-1}(Y_i - (A_i'\hat{\xi}B_f)'),$$

and the variance of the predictor is

$$\tau_i^2 = var(\hat{y}_{fi} - y_{fi}) = \sigma_{2f} - \sigma_f'\Sigma^{-1}\sigma_f + c_i g' D g,$$

where $c_i = A_i'(A'A)^{-1}A_i$, $g = B_f - B\Sigma^{-1}\sigma_f$, and $D = (B\Sigma^{-1}B')^{-1}$.

For the autoregressive structure

$$\Sigma_f = \begin{bmatrix} \Sigma & \sigma_f \\ \sigma_f & \sigma_{2f} \end{bmatrix} = \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & . & . & \rho^p \\ \rho & 1 & \rho & \rho^2 & \rho^3 & . & \rho^{p-1} \\ & & . & & & & \\ & & & & . & & \\ \rho^p & \rho^{p-2} & & & & \rho & 1 \end{pmatrix}.$$

Then doing calculations in a similar way (as for V_2), we get

$$\sigma_f' \Sigma^{-1} = (0_{p-1}', \rho)$$

$$\sigma_{2f} - \sigma_f' \Sigma^{-1} \sigma_f = \sigma^2(1 - \rho^2).$$

Using this we have

$$\hat{y}_{fi} = (A_i' \hat{\xi} B_f)' + (0_{p-1}', \rho)(Y_i - (A_i' \hat{\xi} B)')$$

$$\tau_i^2 = var(\hat{y}_{fi} - y_{fi}) = \sigma^2(1 - \rho^2) + c_i g' D g$$

$$g = B_f - B(0_{p-1}', \rho)'.$$

In applications, we replace ρ and σ^2 in the formulae by their maximum likelihood estimates obtained using the data matrix Y. It is interesting to note here that we have used the observation corresponding to the i^{th} individual for both estimation of the parameters and prediction of y_{fi}. As in the equicorrelation case one can construct approximate prediction intervals for y_{fi} and also for V_{2i}. See Naik (1990) for details.

3.3.3 Prediction of missing values

Prediction of V_i :

Here we consider the prediction of unobserved portion of a vector, when the unobserved portion is in the middle. That is, we predict V_i, given V_1, V_2, and Y, where $V' = (V_1', V_i', V_2')$. For convenience, suppose V_i is a scalar and it corresponds to the i^{th} position of the vector V. Rearrange the elements of V so that $V' =$

(V_i, V_1', V_2'). Let $E(V') = (E(V_i), E(V_1'), E(V_2')) = (A_f\xi B_i, A_f\xi B_1, A_f\xi B_2)$, where A_f etc. are appropriately defined vectors or matrices. Let

$$cov\begin{pmatrix} V_i \\ V_1 \\ V_2 \end{pmatrix} = \begin{bmatrix} \omega_{ii} & \omega_{i1} & \omega_{i2} \\ \omega_{1i} & \Omega_{11} & \Omega_{12} \\ \omega_{2i} & \Omega_{21} & \Omega_{22} \end{bmatrix} \equiv \begin{bmatrix} \omega_{ii} & \omega_i \\ \omega_i' & \Omega \end{bmatrix}$$

The minimum mean square predictor of V_i then is

$$E(V_i/V_1, V_2, Y) = A_f\xi B_i + \omega_i'\Omega^{-1}\left[\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} - \begin{pmatrix} A_f\xi B_1 \\ A_f\xi B_2 \end{pmatrix}\right]'$$

Computation of $\omega_i'\Omega^{-1}$. 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

$$\begin{aligned} \Omega^{-1} &= \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \\ &= \begin{bmatrix} \Omega_{11}^{-1} + FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{bmatrix}, \end{aligned}$$

where $F = \Omega_{11}^{-1}\Omega_{12}$ and $E = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}$.

Note that the $(i-1)\times(i-1)$ matrix Ω_{11}, the $(p-i)\times(p-i)$ matrix Ω_{22}, and the $(i-1)\times(p-i)$ matrix Ω_{12} for the structure in hand are

$$\Omega_{11} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & . & . & \rho^{i-2} \\ \rho & 1 & \rho & . & . & \rho^{i-3} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{i-2} & \rho^{i-3} & . & & . & 1 \end{pmatrix}$$

$$\Omega_{22} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & . & . & \rho^{p-i-1} \\ \rho & 1 & \rho & . & . & \rho^{p-i-2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{p-i-1} & . & . & . & . & 1 \end{pmatrix}$$

$$\Omega_{12} = \sigma^2 \begin{pmatrix} \rho^{i} & \rho^{i+1} & . & . & . & \rho^{p-1} \\ \rho^{i-1} & \rho^{i} & . & . & . & \rho^{p-2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{2} & \rho^{3} & . & . & . & \rho^{p-i+1} \end{pmatrix} = \Omega_{21}'.$$

Further, $\omega_i' = (\omega_{1i}, \omega_{2i})' = \sigma^2(\rho^{i-1}, \rho^{i-2}, ..., \rho^2, \rho, \rho, \rho^2, ..., \rho^{p-i})$ and $\omega_{ii} = \frac{\sigma^2}{1-\rho^2}$.

$$\Omega_{21} = \sigma^2 \begin{pmatrix} \rho^i & \rho^{i-1} & . & . & \rho^2 \\ \rho^{i+1} & \rho^i & . & . & \rho^3 \\ . & . & . & . & . \\ . & . & . & . & . \\ \rho^{p-1} & \rho^{p-2} & . & . & \rho^{p-i+1} \end{pmatrix}$$

$$= \begin{pmatrix} \rho a' \\ \rho^2 a' \\ . \\ . \\ \rho^{p-i} a' \end{pmatrix}$$

Here $a' = (\rho^{i-1}, \rho^{i-2}, ..., \rho)'$ and $\Omega_{11}^{-1} = \frac{1}{\sigma^2(1-\rho^2)}[\rho^2 C_1 - \rho C_2 + I_p]$.

$$\Omega_{21}\Omega_{11}^{-1} = \begin{pmatrix} \frac{\rho^2}{1-\rho^2}\rho a' C_1 - \frac{\rho}{1-\rho^2}\rho a' C_2 + \frac{1}{1-\rho^2}\rho a' \\ \frac{\rho^2}{1-\rho^2}\rho^2 a' C_1 - \frac{\rho}{1-\rho^2}\rho^2 a' C_2 + \frac{1}{1-\rho^2}\rho^2 a' \\ . \\ . \\ \frac{\rho^2}{1-\rho^2}\rho^{p-1} a' C_1 - \frac{\rho}{1-\rho^2}\rho^{p-1} a' C_2 + \frac{1}{1-\rho^2}\rho^{p-1} a' \end{pmatrix}$$

Here

$$
\begin{aligned}
a'C_1 &= (\rho^{i-1}, \rho^{i-2}, ..., \rho)' \begin{pmatrix} 0 & & & & & & \\ & 1 & & & & & \\ & & . & & & & \\ & & & . & & & \\ & & & & . & & \\ & & & & & 1 & \\ & & & & & & 0 \end{pmatrix} \\
&= (0, \rho^{i-2}, \rho^{i-3}, ..., \rho^2, 0)
\end{aligned}
$$

$$
\begin{aligned}
a'C_2 &= (\rho^{i-1}, \rho^{i-2}, ..., \rho)' \begin{pmatrix} 0 & 1 & & & & & \\ 1 & 0 & . & & & & \\ & . & . & . & & & \\ & & . & . & . & & \\ & & & . & . & . & \\ & & & & . & . & 1 \\ & & & & & 1 & 0 \end{pmatrix} \\
&= (\rho^{i-2}, \rho^{i-1} + \rho^{i-3}, \rho^{i-2} + \rho^{i-4}, ..., \rho^3 + \rho, \rho^2)
\end{aligned}
$$

$$
\begin{aligned}
& \frac{\rho}{1-\rho^2}[\rho^2 a'C_1 - \rho a'C_2 + a'] \\
= & \frac{\rho}{1-\rho^2}[(0, \rho^i, \rho^{i-1}, ..., \rho^4, 0) - (\rho^{i-1}, \rho^i + \rho^{i-2}, \rho^{i-1} + \rho^{i-3}, ..., \rho^4 + \rho^2, \rho^3) \\
& + (\rho^{i-1}, \rho^{i-2}, ..., \rho)'] \\
= & \frac{\rho}{1-\rho^2}[(0, 0, 0, ..., 0, \rho(1-\rho^2))]
\end{aligned}
$$

$$= (0,0,0,...,0,\rho^2)$$

Therefore,

$$\Omega_{21}\Omega_{11}^{-1} = \begin{pmatrix} 0 & 0 & . & . & . & 0 & \rho^2 \\ 0 & 0 & . & . & . & 0 & \rho^3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 & \rho^{p-i+1} \end{pmatrix} = F'$$

$$\Omega_{21}\Omega_{11}^{-1}\Omega_{12} = \begin{pmatrix} 0 & 0 & . & . & . & 0 & \rho^2 \\ 0 & 0 & . & . & . & 0 & \rho^3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 & \rho^{p-i+1} \end{pmatrix} \sigma^2 \begin{pmatrix} \rho^i & \rho^{i+1} & . & . & . & \rho^{p-1} \\ \rho^{i-1} & \rho^i & . & . & . & \rho^{p-2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^2 & \rho^3 & . & . & . & \rho^{p-i+1} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \rho^4 & \rho^5 & . & . & . & \rho^{p-i+3} \\ \rho^5 & \rho^6 & . & . & . & \rho^{p-i+4} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{p-i+3} & \rho^{p-i+4} & . & . & . & \rho^{2(p-i+1)} \end{pmatrix}$$

$$= db'$$

$$= \begin{pmatrix} \rho \\ \rho^2 \\ . \\ . \\ . \\ \rho^{p-i} \end{pmatrix} \sigma^2\rho^2(\rho, \rho^2, ..., \rho^{p-i})$$

Therefore, $E = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12} = \Omega_{22} - db'$ and $E^{-1} = \Omega_{22}^{-1} + \frac{\Omega_{22}^{-1}db'\Omega_{22}^{-1}}{1-b'\Omega_{22}^{-1}d}$. And $\Omega_{22}^{-1} = \frac{1}{\sigma^2(1-\rho^2)}[\rho^2 C_1 - \rho C_2 + I_{p-i}]$, so

$$\begin{aligned}
&(\rho, \rho^2, ..., \rho^{p-i})\Omega_{22}^{-1} \\
= \; & \frac{\rho^2}{\sigma^2(1-\rho^2)}(0, \rho^2, .., \rho^{p-i-1}, 0) - \frac{\rho}{\sigma^2(1-\rho^2)}(\rho^2, \rho + \rho^3, .., \rho^{p-i-2} + \rho^{p-i}, \rho^{p-i-1}) \\
& + \frac{1}{\sigma^2(1-\rho^2)}(\rho, \rho^2, ..., \rho^{p-i}) \\
= \; & \frac{1}{\sigma^2(1-\rho^2)}(\rho - \rho^3, 0, ..., 0) \\
= \; & \frac{1}{\sigma^2}(\rho, 0, ..., 0).
\end{aligned}$$

Thus,

$$b'\Omega_{22}^{-1}d = (\rho^3, 0, ..., 0) \begin{pmatrix} \rho \\ \rho^2 \\ . \\ . \\ \rho^{p-i} \end{pmatrix} = \rho^4$$

Hence we have

$$E^{-1} = \Omega_{22}^{-1} + \frac{1}{\sigma^2 1 - \rho^4} \begin{pmatrix} \rho \\ 0 \\ . \\ . \\ 0 \end{pmatrix} (\rho^3, 0, .., 0)$$

$$= \Omega_{22}^{-1} + \frac{1}{\sigma^2} \begin{pmatrix} \frac{\rho^4}{1-\rho^4} & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \end{pmatrix}$$

Next,

$$FE^{-1} = \begin{pmatrix} \underline{0}' \\ \underline{0}' \\ . \\ . \\ \underline{0}' \\ \rho d' \end{pmatrix} \Omega_{22}^{-1} + \frac{1}{\sigma^2} \begin{pmatrix} \underline{0}' \\ \underline{0}' \\ . \\ . \\ \underline{0}' \\ \rho \underline{d}' \end{pmatrix} \begin{pmatrix} \frac{\rho^4}{1-\rho^4} & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ \rho d' \Omega_{22}^{-1} & & & & & \end{pmatrix} + \frac{1}{\sigma^2} \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ \frac{\rho^6}{1-\rho^4} & 0 & . & . & . & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ \frac{\rho^2}{\sigma^2} & 0 & . & . & . & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ \frac{\rho^6}{\sigma^2 1-\rho^4} & 0 & . & . & . & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ \frac{\rho^2}{\sigma^2 1-\rho^4} & 0 & . & . & . & 0 \end{pmatrix}$$

So,

$$FE^{-1}F' = \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ \frac{\rho^2}{\sigma^2 1-\rho^4} & 0 & . & . & . & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & . & . & . & 0 & \rho^2 \\ 0 & 0 & . & . & . & 0 & \rho^3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 & \rho^{p-i+1} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & \frac{\rho^4}{\sigma^2 1-\rho^4} \end{pmatrix}$$

Thus,

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

$$= \begin{bmatrix} \Omega_{11}^{-1} + FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \Omega_{11}^{-1} + \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & \frac{\rho^4}{\sigma^2(1-\rho^4)} \end{pmatrix} & - \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & \frac{\rho^4}{\sigma^2(1-\rho^4)} \end{pmatrix} \\ - \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & \frac{\rho^4}{\sigma^2(1-\rho^4)} \end{pmatrix} & \Omega_{22}^{-1} + \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & \frac{\rho^4}{\sigma^2(1-\rho^4)} \end{pmatrix} \end{bmatrix}$$

Finally, we need

$$
\begin{aligned}
w_i'\Omega^{-1} &= \sigma^2(\rho^{i-1},\rho^{i-2},...,\rho^2,\rho:\rho,\rho^2,...,\rho^{p-i})\Omega^{-1} \\
&= (a':d')\begin{bmatrix} \Omega_{11}^{-1}+FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{bmatrix} \\
&= [a'\Omega_{11}^{-1}+(0,0,...,0,\frac{\rho^5}{1-\rho^4})-(0,0,...,0,\frac{\rho^3}{1-\rho^4}):-(\frac{\rho^3}{1-\rho^4},0,...,0) \\
&\quad +d'\Omega_{22}^{-1}+(\frac{\rho^5}{1-\rho^4},0,...,0)] \\
&= [(0,0,...,0,\rho)+(0,0,...,0,\frac{\rho^5}{1-\rho^4})-(0,0,...,0,\frac{\rho^3}{1-\rho^4}):-(\frac{\rho^3}{1-\rho^4},0,...,0) \\
&\quad +(\rho,0,...,0)+(\frac{\rho^5}{1-\rho^4},0,...,0)] \\
&= (0,...,0,\frac{\rho}{1+\rho^2}:\frac{\rho}{1+\rho^2},0,...,0).
\end{aligned}
$$

Thus a minimum mean square predictor of V_i given V_1, V_2 and Y is

$$
\hat{V}_i = A_f\hat{\xi}B_i+\frac{\rho}{1+\rho^2}(V_{i-1}-A_f\hat{\xi}B_{i-1})+\frac{\rho}{1+\rho^2}(V_{i+1}-A_f\hat{\xi}B_{i+1}), \tag{3.6}
$$

where the vectors B_{i-1} and B_{i+1} are all appropriate part of the matrix B and V_{i-1} and V_{i+1} are the $(i-1)^{th}$ and $(i+1)^{th}$ elements of V respectively. Also

$$
\begin{aligned}
w_i'\Omega^{-1}w_i &= (0,...,0,\frac{\rho}{1+\rho^2}:\frac{\rho}{1+\rho^2},0,...,0)\sigma^2(\rho^{i-1},\rho^{i-2},...,\rho^2,\rho:\rho,\rho^2,...,\rho^{p-i})' \\
&= \frac{\sigma^2}{1-\rho^2}\frac{2\rho}{1+\rho^2}.
\end{aligned}
$$

And

$$
w_{ii}-w_i'\Omega^{-1}w_i = \frac{\sigma^2}{1-\rho^2}-\frac{\sigma^2}{1-\rho^2}\frac{2\rho}{1+\rho^2} = \frac{\sigma^2}{1+\rho^2}
$$

So, we have the var$(\hat{V}_i - V_i)$ as

$$
\tau_i^2 = \frac{\sigma^2}{1+\rho^2}+c_fd_i'Dd_i,
$$

where $d_i = B_i - (\frac{\rho}{1+\rho^2}B_{i-1} + \frac{\rho}{1+\rho^2}B_{i+1})$.

An alternative predictor of V_i :

In finding $\hat{V}_i$ ($\hat{V}_2$ as well) above, we have used the observed part of the vector V for predicting V_i but ignored it for estimating various parameters involved. We suggest an iterative prediction method in which the observed portion of V will be used not only for prediction but also for estimation. The procedure essentially makes use of the formulae developed in (3.6) for $\hat{V}_i$ and the maximum likelihood estimates in Theorem 3. It can be described as follows:

1. Compute $\tilde{V}_i$ using the formula given in (3.6) for an initial value of ρ, say $\rho = 0.5$.
2. Using $\tilde{V}_i$, form the vector $\tilde{V}' = (\tilde{V}_i, V_1, V_2)$ and the new data matrix

$$Y_{new} = \begin{pmatrix} Y \\ \tilde{V}' \end{pmatrix}$$

3. Use Theorem 3 to evaluate ξ, ρ and σ^2 replacing Y by Y_{new} and A by

$$A_{new} = \begin{pmatrix} A \\ A_f \end{pmatrix}$$

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

To compare the predictor $\tilde{V}_i$ with $\hat{V}_i$ we conduct the following simulation experiment.

First, we generate $p = 12$, normal random variable z_j, $j = 1, ..., p$ with mean zero and variance 1. Next, for $\rho = 0.5$ we find a $p \times p$ matrix $\Gamma(\rho)$ such that

$V(\rho) = \Gamma(\rho)\Gamma'(\rho)$, where $V(\rho)$ is a $p \times p$ matrix of autoregressive structure. Then we find a $p \times 1$ vector $u = \Gamma z$, where $z' = (z_1, z_2, ..., z_p)$, and $u' = (u_1, u_2, ..., u_p)$. Let $\xi = (\xi_0, \xi_1, \xi_2)$ and $\mu_j = \xi_0 + \xi_1 j + \xi_2 j^2$, $j = 1, 2, ..., p$. We set $y_j = \mu_j + u_j$, $j = 1, 2, ..., p$ to obtain a vector $y' = (y_1, y_2, ..., y_p)$. For our simulation we take $\xi = (0, 0, 0)$.

Repeating this procedure, $n = 20$ times, we obtain the $n \times p$ data matrix Y such that

$$Y = A\xi B + E, \tag{3.7}$$

where A is a $n \times 1$ vector of ones, B is a matrix of second degree polynomial terms and rows of E are independently distributed as p-variate normal with zero means and covariance matrix $\sigma^2 V(\rho)$, $V(\rho)$ having the autoregressive structure. Next, we generate one more $1 \times p$ vector V in the same way. We pretend that V_5, the 5^{th} element of V, is missing. For each simulation run (that is, for each generated set of data), we compute $\hat{V}_5$ and $\tilde{V}_5$ using (3.6) and the steps (1)-(4) of the algorithm respectively. During this process we also have $\hat{\rho}$ and $\tilde{\rho}$, the estimates of ρ using the respective procedures. We repeat the whole procedure $N = 2000$ times and obtain $\frac{1}{N}\sum(\hat{V}_5 - V_5)^2$, $\frac{1}{N}\sum(\tilde{V}_5 - V_5)^2$, $\frac{1}{N}\sum(\hat{\rho} - \rho)^2$, $\frac{1}{N}\sum(\tilde{\rho} - \rho)^2$, $\frac{1}{N}\sum \hat{\rho} - \rho$ and $\frac{1}{N}\sum \tilde{\rho} - \rho$. The above quantities respectively are the predictive mean squared errors for estimating ρ for the two procedures (using formula (3.6), and using the algorithm) , and biases for the two procedures. These quantities for the simulation are 0.6216, 0.6201, 0.0036, 0.0034, -0.0141, and -0.0123. We have run this experiment for different choices of n, p, and ρ. The results were not different from the pattern we see in this example. It is clear that there is a slight improvement if we use the algorithm, although the improvement is very minimal.

Since this algorithm is easy to adopt and in fact, leads to the maximum likelihood predictor of V_i based on the unconditional likelihood (note that $\hat{V}_i$ is the predictor based on conditional likelihood), we recommend that this algorithm be used in practice. Of course, for large sample ($n \rightarrow \infty$) the two predictors are essentially the same.

3.4 Model for monotone or balanced incomplete data

The model for incomplete data is as follows:

$$Y_{i_{n_i \times p_i}} = A_{i_{n_i \times m}} \xi_{m \times k} B_{i_{k \times p_i}} + \in_{i_{n_i \times p_i}}, i = 1, 2, ..., g, \quad (3.8)$$

where g represents the number of groups in the model and each group has n_i units (not the same number in each group). Here the rows of the error matrix $\in_i$ are independent each distributed as $N_p(0, \Sigma_i)$, where Σ_i is a $p_i \times p_i$ matrix with autoregressive covariance structure. We discuss estimation and prediction for this model. Testing of hypothesis $H_0 : E\xi F = C$ can be done the same way as explained in chapter 2.

3.4.1 Estimation

We have seen in chapter 2 that for this model the ML estimate of ξ is

$$vec(\hat{\xi}) = [\sum_{i=1}^{g}(B_i\Sigma_i^{-1}B_i')\bigotimes(A_i'A_i)]^{-1}\sum_{i=1}^{g}(B_i\Sigma_i^{-1}\bigotimes A_i')vec(Y_i). \tag{3.9}$$

Now we consider an autoregressive covariance structure for Σ_i as follows:

$$\Sigma_i = \sigma^2 V_i(\rho) = \sigma^2(\rho^{|j-k|}), j,k = 1,2,...,p_i, i = 1,....,g. \tag{3.10}$$

We have

$$\begin{aligned}
|\Sigma_i| &= (\sigma^2)^{p_i}(1-\rho^2)^{p_i-1} \\
\Sigma_i^{-1} &= (\sigma^2)^{-1}V_i^{-1} = (\sigma^2)^{-1}(1-\rho^2)^{-1}[\rho^2 C_{1i} - 2\rho C_{2i} + I_{p_i}]
\end{aligned}$$

Here C_{1i} and C_{2i} are $p_i \times p_i$ matrices with the same structure as C_1 and C_2 of section 3.2. Taking the log of likelihood function of (2.8) and substituting for $|\Sigma_i|$ and Σ_i^{-1} from the above equations, we get

$$\begin{aligned}
lnL &= -\frac{N_1}{2}ln(2\pi) - \sum_{i=1}^{g} ln(|\Sigma_i|^{\frac{n_i}{2}}) - \frac{1}{2}\sum_{i=1}^{g} tr\Sigma_i^{-1}\hat{E}_i'\hat{E}_i \\
&= -\frac{N_1}{2}ln(2\pi) - \frac{1}{2}N_2 ln(1-\rho^2) - \frac{1}{2}N_1 ln\sigma^2 + \frac{\sigma^2}{1-\rho^2} \\
&\quad \sum_{i=1}^{g} tr(\rho^2 C_{1i} - 2\rho C_{2i} + I_{p_i})\hat{E}_i'\hat{E}_i.
\end{aligned}$$

Hence

$$\frac{\partial lnL}{\partial \sigma^2} = -\frac{N_1}{2\sigma^2} + \frac{1}{2\sigma^4(1-\rho^2)}\sum_{i=1}^{g} tr(\rho^2 C_{1i} - 2\rho C_{2i} + I_{p_i})E_i'E_i$$
$$= 0 \Rightarrow$$
$$\hat{\sigma}^2 = \frac{1}{N_1(1-\rho^2)}\sum_{i=1}^{g} tr(\rho^2 C_{1i} - 2\rho C_{2i} + I_{p_i})E_i'E_i,$$

where $N_1 = \sum_{i=1}^{g} n_i p_i$, $N_2 = \sum_{i=1}^{g} n_i(p_i - 1)$. Let us write $a_1 = \sum_{i=1}^{g} a_{1i}$, $a_2 = \sum_{i=1}^{g} a_{2i}$, $a_3 = \sum_{i=1}^{g} a_{3i}$, $a_{ki} = tr(C_{ki}R_i)$, $R_i = \hat{E}_i'\hat{E}_i$. So we can write the MLE of σ^2 as

$$\hat{\sigma}^2 = \frac{1}{N_1(1-\hat{\rho}^2)}(\hat{\rho}^2 a_1 - 2\hat{\rho}a_2 + a_3).$$

Now to get the MLE of ρ, we take partial derivative of lnL w.r.t ρ.

$$\frac{\partial lnL}{\partial \rho} = \frac{\sum_{i=1}^{g} n_i(p_i-1)\rho}{(1-\rho^2)} - \frac{\rho}{\hat{\sigma}^2(1-\rho^2)^2}\sum_{i=1}^{g} trC_{1i}\hat{E}_i'\hat{E}_i + \frac{1+\rho^2}{\hat{\sigma}^2(1-\rho^2)^2}\sum_{i=1}^{g} trC_{2i}\hat{E}_i'\hat{E}_i - \frac{\rho}{\hat{\sigma}^2(1-\rho^2)^2}\sum_{i=1}^{g} tr\hat{E}_i'\hat{E}_i$$
$$= 0$$

$$\Rightarrow \frac{N_2\hat{\rho}}{(1-\hat{\rho}^2)} - \frac{\hat{\rho}}{\hat{\sigma}^2(1-\hat{\rho}^2)^2}a_1 + \frac{1+\hat{\rho}^2}{\hat{\sigma}^2(1-\hat{\rho}^2)^2}a_2 - \frac{\hat{\rho}}{\hat{\sigma}^2(1-\hat{\rho}^2)^2}a_3 = 0$$
$$\Rightarrow N_2\hat{\rho}(1-\hat{\rho}^2)\hat{\sigma}^2 - \hat{\rho}a_1 + (1+\hat{\rho}^2)a_2 - \hat{\rho}a_3 = 0$$
$$\Rightarrow \frac{N_2}{N_1}\hat{\rho}(\hat{\rho}^2 a_1 - 2\hat{\rho}a_2 + a_3) - \hat{\rho}a_1 + (1+\hat{\rho}^2)a_2 - \hat{\rho}a_3 = 0$$

$$\Rightarrow \hat{\rho}^3 N_2 a_1 - 2\hat{\rho}^2 N_2 a_2 + \hat{\rho} N_2 a_3 - \hat{\rho} N_1 a_1 + N_1 a_2 + \hat{\rho}^2 N_1 a_2 - \hat{\rho} N_1 a_3 = 0$$

$$\Rightarrow \hat{\rho}^3 N_2 a_1 - \hat{\rho}^2 a_2(2N_2 - N_1) + \hat{\rho}(N_2 a_3 - N_1 a_1 - N_1 a_3) + N_1 a_2 = 0$$

$$\Rightarrow \hat{\rho}^3 N_2 a_1 - \hat{\rho}^2 a_2(2N_2 - N_1) - \hat{\rho}(N_1 a_1 + (N_1 - N_2)a_3) + N_1 a_2 = 0$$

So we have the following theorem for the ML estimates under the incomplete (balanced) data model similar to the one given by Fujikoshi et al. (1990) for balanced data model:

THEOREM 4 *The MLE's of ξ, σ^2 and ρ in the model (3.8) with covariance structure (3.1) are the solutions of the following equations:*

$$\begin{aligned}
&(I) vec(\hat{\xi}) = [\sum_{i=1}^{g}(B_i \hat{V}_i^{-1} B_i') \bigotimes (A_i' A_i)]^{-1} \sum_{i=1}^{g}(B_i \hat{V}_i^{-1} \bigotimes A_i') vec(Y_i) \\
&(II) \hat{\sigma}^2 = \frac{1}{N_1(1-\hat{\rho}^2)}(\hat{\rho}^2 a_1 - 2\hat{\rho} a_2 + a_3) \\
&(III) \hat{\rho}^3 N_2 a_1 - \hat{\rho}^2 a_2(2N_2 - N_1) - \hat{\rho}(N_1 a_1 + (N_1 - N_2)a_3) + N_1 a_2 = 0,
\end{aligned}$$

where

$$\hat{V}_i = V_i(\hat{\rho}), N_1 = \sum_{i=1}^{g} n_i p_i, N_2 = \sum_{i=1}^{g} n_i(p_i - 1), a_1 = \sum_{i=1}^{g} a_{1i}, a_2 = \sum_{i=1}^{g} a_{2i},$$

$$a_3 = \sum_{i=1}^{g} a_{3i}, a_{ki} = tr(C_{ki} R_i), k = 1,2,3, C_{3i} = I_{p_i} \, and \, R_i = (Y_i - A_i \xi B_i)'(Y_i - A_i \xi B_i).$$

Note that the equation for estimating ρ is a cubic equation in ρ. Although this equation is different from the cubic equation in Dahiya and Korwar (1980), the ideas from there can be used to show, using Descartes' rule of signs, that the equation has unique root in the interval $(-1, 1)$. A proof follows.

Proof of uniqueness of $\hat{\rho}$:

The cubic equation is

$$\hat{\rho}^3 N_2 a_1 - \hat{\rho}^2 a_2(2N_2 - N_1) - \hat{\rho}(N_1 a_1 + (N_1 - N_2)a_3) + N_1 a_2 = 0.$$

Let us write this as f(ρ) = 0

We note that $N_1 > N_2$, $a_1 > 0$, and $a_3 > 0$. For a_2 we have the following cases:

Case 1 : $a_2 > 0$

By Descartes' Rule, the number of positive roots of this equation is at most 2 and the number of negative roots is at most 1, provided $p_i \geq 2$ with at least one p_i greater than 2.

Case 2: $a_2 < 0$

Again by Descartes' Rule, the number of negative roots is at most 2 and the number of positive roots is at most 1 , provided $p_i \geq 2$ with at least one p_i greater than 2.

Case 3: $a_2 = 0$

For this case, it can be easily shown that there is only one root in the interval (-1,1) and it is equal to zero.

We prove for the case when $a_2 > 0$, the cubic equation $f(\rho) = 0$ has a unique root between (-1,1). It can be similarly proved for the case when $a_2 < 0$.

Case when $a_2 > 0$:

Consider

$$f(-1) = -N_2 a_1 - a_2(2N_2 - N_1) + (N_1 a_1 + (N_1 - N_2)a_3) + N_1 a_2$$

$$= (N_1 - N_2)(a_1 + 2a_2 + a_3).$$

Since $a_2 > 0$ and $N_1 - N_2 > 0$, $f(-1) > 0$. Also $f(0) = N_1 a_2 > 0$. Now

$$\begin{aligned} f(1) &= N_2 a_1 - a_2(2N_2 - N_1) - (N_1 a_1 + (N_1 - N_2)a_3) + N_1 a_2 \\ &= (N_2 - N_1)(a_1 - 2a_2 + a_3) \end{aligned}$$

But $((a_1 - 2a_2 + a_3) > 0$, (Note that $\sum_{k=2}^{p_i} \sum_{j=1}^{p_i} (\epsilon_{jk} - \epsilon_{jk-1})^2 > 0$ gives $((a_1 - 2a_2 + a_3) > 0)$, hence $f(1) < 0$.

Since $f(-1) > 0$ and $f(0) > 0$, and by Descartes' Rule there can be only one negative root, there are no roots in the interval (-1,0). Thus if we prove that there is a unique root in the interval (0,1), we are done. Since $f(1) < 0$, there is at least one root in the interval (0,1). There are in total two positive roots and one negative root. So there can be only one root between (0,1) since $f(0) > 0$ and $f(1) < 0$. Thus the above cubic equation has a unique root in the interval (-1,1). This root can be easily evaluated using simple computer programs.

3.4.2 Prediction

Prediction of y_f

For the j^{th} individual in the i^{th} group we have

$$E\begin{pmatrix} Y_{ij} \\ y_{if_j} \end{pmatrix} = \begin{pmatrix} (A'_{ij}\xi B_i)' \\ (A'_{ij}\xi B_{if})' \end{pmatrix},$$

$$cov\begin{pmatrix} Y_{ij} \\ y_{if_j} \end{pmatrix} = \Sigma_{fi} = \begin{pmatrix} \Sigma_i & \sigma_{fi} \\ \sigma'_{fi} & \sigma_{2fi} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & . & . & \rho^{p_i} \\ \rho & 1 & \rho & \rho^2 & \rho^3 & . & \rho^{p_i-1} \\ & & . & & & & \\ & & & & . & & \\ \rho^{p_i} & \rho^{p_i-1} & & & & \rho & 1 \end{pmatrix}$$

A'_{ij} is the j^{th} row of A_i and B_{if} is a $k \times 1$ known vector. Using this and the conditional expectation of y_{if_j} given Y_{ij}, the predictor is given by

$$\hat{y}_{if_j} = (A'_{ij}\hat{\xi}B_{if})' + (0'_{p_i-1}, \hat{\rho})(Y_{ij} - (A'_{ij}\hat{\xi}B_i)')$$

Also,

$$var(\hat{y}_{if_j} - y_{if_j}) = \hat{\tau}^2_{ij} = \sigma_{2fi} - \sigma'_{fi}\Sigma_i^{-1}\sigma_{fi} + c_{ij}\hat{g}'_i\hat{D}_i\hat{g}_i.$$

But

$$\begin{aligned} \sigma'_{fi}\Sigma_i^{-1}\sigma_{fi} &= (0'_{p_i-1}, \rho)\sigma_{fi} \\ &= (0'_{p_i-1}, \rho)(\rho^{p_i}, \rho^{p_i-1}, ..., \rho)' \\ &= \sigma^2\rho^2 \end{aligned}$$

which gives

$$\sigma_{2fi} - \sigma'_{fi}\Sigma_i^{-1}\sigma_{fi} = \sigma^2 - \sigma^2\rho^2 = \sigma^2(1-\rho^2)$$

Thus we get the variance as

$$\hat{\tau}^2_{ij} = \hat{\sigma}^2(1-\hat{\rho}^2) + c_{ij}\hat{g}'_i\hat{D}_i\hat{g}_i,$$

where

$$\begin{aligned} c_{ij} &= A'_{ij}(A'_iA_i)^{-1}A_{ij}, \\ \hat{g}_i &= B_{if} - B_i(0'_{p_i-1}, \hat{\rho})', \\ \hat{D}_i &= (B_i\hat{\Sigma}_i^{-1}B'_i)^{-1}, \\ \hat{\Sigma}_i^{-1} &= \frac{1}{\hat{\sigma}^2(1-\hat{\rho}^2)}[\hat{\rho}^2C_{1i} - 2\hat{\rho}C_{2i} + I_{p_i}]. \end{aligned}$$

We observe that $(\hat{y}_{if_j} - y_{if_j})/\hat{\tau}_{ij}$ is approximately distributed as normal with mean zero and variance one. Hence a $100(1-\alpha)$ % prediction interval for y_{if_j} can be constructed using the normal distribution as:

$$(\hat{y}_{if_j} - z_{\alpha/2}\hat{\tau}_{ij}, \hat{y}_{if_j} + z_{\alpha/2}\hat{\tau}_{ij}) \tag{3.11}$$

where $\hat{y}_{if_j}$ and $\hat{\tau}_{ij}$ are as given above.

Prediction of V_2

Let $E(V_{i_{1\times p_i}}) = E(V_{1i}, V_{2i}) = (A_{if}\xi B_{1i}, A_{if}\xi B_{2i})$. Also $B_{i_{k\times p_i}} = (B_{1i}, B_{2i})$ and $Cov(V) = \Sigma_i$ having autoregressive structure. Σ_i is partitioned in accordance with the partition of V_i

$$\Sigma_i = \begin{pmatrix} \Sigma_{11i} & \sigma_{12i} \\ \sigma_{21i} & \sigma_{22i} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & . & . & \rho^{p_i-1} \\ \rho & 1 & \rho & . & . & \rho^{p_i-2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & \rho \\ \rho^{p_i-1} & . & . & . & \rho & 1 \end{pmatrix}$$

and

$$\sigma_{12i} = \sigma^2(\rho^{p_i-1}, \rho^{p_i-2}, ..., \rho)'$$

$$\sigma_{22i} = \sigma^2, \ \Sigma_{11i} = \sigma^2 V_i(\rho)_{p_i-1 \times p_i-1}$$

Following the same procedure as we did for y_{if_j}, we get

$$\begin{aligned} \Sigma_{11i}^{-1}\sigma_{12i} &= (0'_{p_i-2}, \rho)' \\ \sigma_{22i} - \sigma_{21i}\Sigma_{11i}^{-1}\sigma_{12i} &= \sigma^2(1-\rho^2). \end{aligned}$$

Thus a conditional predictor of V_{2i} given V_{1i} and Y is

$$\begin{aligned} \hat{V}_{2i} &= A_{if}\hat{\xi}B_{2i} + (V_{1i} - A_{if}\hat{\xi}B_{1i})\Sigma_{11i}^{-1}\sigma_{12i} \\ &= A_{if}\hat{\xi}B_{2i} + (V_{1i} - A_{if}\hat{\xi}B_{1i})(0'_{p_i-2}, \hat{\rho})'. \end{aligned}$$

And also the variance of $(\hat{V}_{2i} - V_{2i})$ is

$$\begin{aligned} \hat{\tau}_i^2 &= \sigma_{22i} - \sigma_{21i}\Sigma_{11i}^{-1}\sigma_{12i} + c_{fi}\hat{d}_i'\hat{D}_i\hat{d}_i \\ &= \hat{\sigma}^2(1-\hat{\rho}^2) + c_{fi}\hat{d}_i'\hat{D}_i\hat{d}_i, \end{aligned}$$

where

$$\begin{aligned} c_{fi} &= A_{if}(A_i'A_i)^{-1}A_{if}', \\ \hat{d}_i &= B_{2i} - B_{1i}(0'_{p_i-2}, \hat{\rho}), \\ \hat{D}_i &= (B_i\hat{\Sigma}_i^{-1}B_i')^{-1}. \end{aligned}$$

Since $(\hat{V}_{2i} - V_{2i})/\hat{\tau}_{ij}$ is approximately distributed as $N(0,1)$, a $100(1-\alpha)$ % prediction interval for V_{2i} can be constructed using the normal distribution as:

$$(\hat{V}_{2i} - z_{\alpha/2}\hat{\tau}_{ij}, \hat{V}_{2i} + z_{\alpha/2}\hat{\tau}_{ij}). \tag{3.12}$$

3.5 Unbalanced data model

The growth curve model for analyzing unbalanced data is :

$$Y_{in_i \times p_i} = A_{in_i \times m} \xi_{m \times k} B_{k \times p} G_{i p \times p_i} + \in_{in_i \times p_i}, i = 1, 2, ..., g, \tag{3.13}$$

where G_i is the matrix of 0's and 1's such that if the observations $i_1, ..., i_{p_i}$ are available, then G_i has 1 in the $(k, i_k)^{th}$ position for $k = 1, ..., p_i$ and zeros elsewhere. In this chapter, rows of error matrix $\in_i$ are independent, each distributed as $N_p(0, G_i'\Sigma G_i)$, where Σ is a $p \times p$ matrix with autoregressive structure, that is, $\Sigma = \sigma^2(\rho^{|i-j|})$. It is easy to see that $G_i'\Sigma G_i$ has a Markov structure, that is,

$$G_i'\Sigma G_i = \Sigma_i = \sigma^2(\rho^{|t_{ij} - t_{ij'}|}) = \sigma^2 V_i(\rho) \ , j, j' = 1, ..., p_i,$$

where t_{ij}, $j = 1, ..., p_i$ are the consecutive times points where the observations on an individual of the i^{th} group is made. Then $V_i(\rho)$ can be written as

$$V_i(\rho) = \begin{pmatrix} 1 & \rho^{t_{i2}-t_{i1}} & \rho^{t_{i3}-t_{i1}} & . & . & \rho^{t_{ip_i}-t_{i1}} \\ \rho^{t_{i2}-t_{i1}} & 1 & \rho^{t_{i3}-t_{i2}} & . & . & \rho^{t_{ip_i}-t_{i2}} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \rho^{t_{ip_i}-t_{i1}} & \rho^{t_{ip_i}-t_{i2}} & . & & . & 1 \end{pmatrix}$$

The determinant and the inverse of this matrix 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

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

and

$V_i(\rho)^{-1} = V_i^{-1} =$

$$\begin{pmatrix} f_{i1} & -f_{i1}\rho^{d_{i1}} & 0 & . & . & 0 \\ -f_{i1}\rho^{d_{i1}} & f_{i1}+f_{i2}-1 & -f_{i2}\rho^{d_{i2}} & . & . & 0 \\ 0 & -f_{i2}\rho^{d_{i2}} & f_{i2}+f_{i3}-1 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & . & . & . & f_{ip_i-2}+f_{ip_i-1}-1 & -f_{ip_i-1}\rho^{d_{ip_i-1}} \\ 0 & . & . & 0 & -f_{ip_i-1}\rho^{d_{ip_i-1}} & f_{ip_i-1} \end{pmatrix}$$

If $t_{i1}, t_{i2}, ..., t_{ip}$ are integers then $V_i(\rho)$ is p.d. for $-1 < \rho < 1$, otherwise for $0 < \rho < 1$.

3.5.1 Estimation

The log of the likelihood function is written as

$$\begin{aligned} lnL &= -\frac{N_1}{2}ln(2\pi) - \sum_{i=1}^{g}\frac{n_i}{2}ln(|\Sigma_i|) - \frac{1}{2\sigma^2}\sum_{i=1}^{g} trV_i^{-1}E_i'E_i \\ &= -\frac{N_1}{2}ln(2\pi) - \frac{N_1}{2}ln(\sigma^2) - \sum_{i=1}^{g}\frac{n_i}{2}\sum_{j=1}^{p_i-1} ln(1-\rho^{2d_{ij}}) - \frac{1}{2\sigma^2}\sum_{i=1}^{g} trV_i^{-1}E_i'E_i \end{aligned}$$

Here $N_1 = \sum_{i=1}^{g} n_i p_i$ and $E_i = Y_i - A_i\xi B_i$. Taking the partial derivative of lnL w.r.t. σ^2 and setting it to zero gives the ML estimate of σ^2 as

$$\hat{\sigma}^2 = \frac{1}{N_1}\sum_{i=1}^{g} tr\hat{V}_i^{-1}\hat{E}_i'\hat{E}_i,$$

where $\hat{E}_i = Y_i - A_i\hat{\xi}B_i$ and $\hat{V}_i = V_i(\hat{\rho})$, $\hat{\xi}$ and $\hat{\rho}$ being the MLEs of ξ and ρ respectively. The MLE $\hat{\xi}$ of ξ will be given in the theorem that follows and the MLE $\hat{\rho}$ of ρ is derived next.

Take the partial derivative of lnL with respect to ρ, and set it to zero, to get

$$-\sum_{i=1}^{g}\frac{n_i}{2}\sum_{j=1}^{p_i-1}\frac{2d_{ij}\rho^{2d_{ij}-1}}{1-\rho^{2d_{ij}}} - \frac{1}{2\sigma^2}\sum_{i=1}^{g} tr\frac{\partial}{\partial\rho}V_i^{-1}E_i'E_i = 0. \tag{3.14}$$

Now to evaluate $\frac{\partial V_i^{-1}}{\partial\rho}$, consider

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

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

$$(f_{i1}^2 2d_{i1}\rho^{2d_{i1}-1}, f_{i1}^2 2d_{i1}\rho^{2d_{i1}-1} + f_{i2}^2 2d_{i2}\rho^{2d_{i2}-1}, \ldots,$$
$$f_{ip_i-2}^2 2d_{ip_i-2}\rho^{2d_{ip_i-2}-1} + f_{ip_i-1}^2 2d_{ip_i-1}\rho^{2d_{ip_i-1}-1}, f_{ip_i-1}^2 2d_{ip_i-1}\rho^{2d_{ip_i-1}-1}).$$

Also,

$$\begin{aligned}
\frac{\partial(f_{ij}\rho^{d_{ij}})}{\partial\rho} &= f_{ij}d_{ij}\rho^{d_{ij}-1} + \rho^{d_{ij}}\frac{\partial f_{ij}}{\partial\rho} \\
&= f_{ij}d_{ij}\rho^{d_{ij}-1} + \rho^{d_{ij}}f_{ij}^2 2d_{ij}\rho^{2d_{ij}-1} \\
&= d_{ij}\rho^{d_{ij}-1}f_{ij}[1 + \frac{2\rho^{2d_{ij}}}{1-\rho^{2d_{ij}}}] \\
&= d_{ij}\rho^{d_{ij}-1}f_{ij}[\frac{1+\rho^{2d_{ij}}}{1-\rho^{2d_{ij}}}] \\
&= d_{ij}\rho^{d_{ij}-1}f_{ij}^2(1+\rho^{2d_{ij}}).
\end{aligned}$$

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

$$(d_{i1}\rho^{d_{i1}-1}f_{i1}^2(1+\rho^{2d_{i1}}), d_{i2}\rho^{d_{i2}-1}f_{i2}^2(1+\rho^{2d_{i2}}), \ldots, d_{ip_{i-1}}\rho^{d_{ip_{i-1}}-1}f_{ip_{i-1}}^2(1+\rho^{2d_{ip_{i-1}}})).$$

Thus we see that $\frac{\partial}{\partial\rho}V_i^{-1}$ is a tridiagonal matrix with diagonal and off-diagonal elements as given above. Substituting this in equation(3.14), we get that the ML estimate of ρ as a solution of the equation:

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

Then we have the following theorem for ML estimators of ξ, σ^2 and ρ.

THEOREM 5 *The MLE's of ξ, σ^2 and ρ in the unbalanced model with autoregressive covariance structure are the solutions of the following equations:*

$$(I)\,vec(\hat{\xi})=[\sum_{i=1}^{g}(BG_i\Sigma_i^{-1}G_i'B')\bigotimes(A_i'A_i)]^{-1}\sum_{i=1}^{g}(BG_i\Sigma_i{}^{-1}\bigotimes A_i')vec(Y_i)$$
$$(II)\,\hat{\sigma}^2=\frac{1}{\sum_{i=1}^{g}n_ip_i}\sum_{i=1}^{g}tr\hat{V}_i^{-1}\hat{E}_i'\hat{E}_i$$
$$(III)\,\hat{\sigma}^2\sum_{i=1}^{g}n_i\sum_{j=1}^{p_i-1}2d_{ij}f_{ij}\hat{\rho}^{2d_{ij}-1}-\sum_{i=1}^{g}\sum_{j=1}^{p_i-1}[f_{ij}{}^2(2d_{ij})\hat{\rho}^{2d_{ij}-1}\sum_{k=1}^{n_i}(\hat{\epsilon}_{kj}^2+\hat{\epsilon}_{kj+1}^2)$$
$$-2d_{ij}\hat{\rho}^{d_{ij}-1}f_{ij}{}^2(1+\hat{\rho}^{2d_{ij}})\sum_{k=1}^{n_i}(\hat{\epsilon}_{kj}\hat{\epsilon}_{kj+1})]=0,$$

where $\hat{E}_{i_{n_i\times p_i}}=(\hat{\epsilon}_{kj})$ $k=1,2,...,n_i$ $j=1,2,...,p_i$, is the matrix of residuals for the i^{th} group.

Note, unlike in Theorems 3 and 4, that the third equation (III) for estimating ρ involves $\hat{\sigma}^2$. However, this does not create any problem for numerically evaluating the ML estimates using the three equations $(I)-(III)$. We have successfully implemented this algorithm using a FORTRAN program. This program is provided in section 3.6 (PROGRAM 1). We fitted the plasma data under Markov structure. The results are $\hat{\rho} = 0.71$, $\hat{\sigma}^2 = 0.47$ and

$$\hat{\xi} = \begin{pmatrix} 4.45 & -0.52 & 0.045 \\ 5.01 & -0.53 & 0.04 \end{pmatrix}$$

Using these ML estimates one can suggest a test statistic of the form (2.12) to test the hypothesis $H_0 : E\xi F = C$.

3.5.2 Prediction

Prediction of y_f

For the j^{th} individual in the i^{th} group we have

$$E\begin{pmatrix} Y_i \\ y_{if_j} \end{pmatrix} = \begin{pmatrix} (A'_{ij}\xi BG_i)' \\ (A'_{ij}\xi BG_{if})' \end{pmatrix} \quad cov\begin{pmatrix} Y_{ij} \\ y_{if_j} \end{pmatrix} = \Sigma_{fi} = \begin{pmatrix} \Sigma_i & \sigma_{fi} \\ \sigma'_{fi} & \sigma_{2fi} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} 1 & \rho^{t_{i2}-t_{i1}} & \rho^{t_{i3}-t_{i1}} & \cdots & & \rho^{t_{ip_i+1}-t_{i1}} \\ \rho^{t_{i2}-t_{i1}} & 1 & \rho^{t_{i3}-t_{i2}} & \cdots & & \rho^{t_{ip_i+1}-t_{i2}} \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ \rho^{t_{ip_i}-t_{i1}} & \rho^{t_{ip_i}-t_{i2}} & \cdot & \cdot\ \cdot & & \\ \rho^{t_{ip_i+1}-t_{i1}} & \rho^{t_{ip_i+1}-t_{i2}} & \cdot & \cdot\ \cdot & \rho^{t_{ip_i+1}-t_{ip_i}} & 1 \end{pmatrix}$$

Since

$$\sigma'_{fi} = \sigma^2(\rho^{t_{ip_i+1}-t_{i1}}, \rho^{t_{ip_i+1}-t_{i2}}, \ldots, \rho^{t_{ip_i+1}-t_{ip_i-1}}, \rho^{t_{ip_i+1}-t_{ip_i}}),$$

$$\sigma_{2fi} = \sigma^2, \text{ and}$$

$$\Sigma_i^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} f_{i1} & -f_{i1}\rho^{d_{i1}} & 0 & . \quad . & 0 \\ -f_{i1}\rho^{d_{i1}} & f_{i1}+f_{i2}-1 & -f_{i2}\rho^{d_{i2}} & 0 \quad . & 0 \\ 0 & -f_{i2}\rho^{d_{i2}} & f_{i2}+f_{i3}-1 & 0 \quad 0 & 0 \\ . & . & . & . \quad . & . \\ . & . & . & . \quad . & . \\ 0 & . & . & 0 \quad f_{p_i-2}+f_{p_i-1}-1 & -f_{ip_i-1}\rho^{d_{ip_i-1}} \\ 0 & . & . & 0 \quad -f_{ip_i-1}\rho^{d_{ip_i-1}} & f_{ip_i-1} \end{pmatrix},$$

we have

$$\sigma'_{fi}\Sigma_i^{-1} = (f_{i1}\rho^{t_{ip_i+1}-t_{i1}} - f_{i1}\rho^{d_{i1}}\rho^{t_{ip_i+1}-t_{i2}},$$

$$-f_{i1}\rho^{d_{i1}}\rho^{t_{ip_i+1}-t_{i1}} + f_{i1}\rho^{t_{ip_i+1}-t_{i2}} + f_{i2}\rho^{t_{ip_i+1}-t_{i2}} - \rho^{t_{ip_i+1}-t_{i2}} - f_{i2}\rho^{d_{i2}}\rho^{t_{ip_i+1}-t_{i3}},$$

$$-f_{i2}\rho^{d_{i2}}\rho^{t_{ip_i+1}-t_{i2}} + f_{i2}\rho^{t_{ip_i+1}-t_{i2}} + f_{i3}\rho^{t_{ip_i+1}-t_{i2}} - \rho^{t_{ip_i+1}-t_{i2}} - f_{i3}\rho^{t_{ip_i+1}-t_{i3}},$$

$$\ldots, -f_{ip_i-1}\rho^{d_{ip_i-1}}\rho^{t_{ip_i+1}-t_{ip_i-1}} + f_{ip_i-1}\rho^{t_{ip_i+1}-t_{ip_i}})$$

$$= (f_{i1}\rho^{t_{ip_i+1}-t_{i1}} - f_{i1}\rho^{t_{ip_i+1}-t_{i2}+t_{i2}-t_{i1}},$$

$$-f_{i1}\rho^{t_{ip_i+1}-t_{i1}+t_{i2}-t_{i1}} + f_{i1}\rho^{t_{ip_i+1}-t_{i2}} + f_{i2}\rho^{t_{ip_i+1}-t_{i2}} - \rho^{t_{ip_i+1}-t_{i2}} - f_{i2}\rho^{t_{i3}--t_{i2}+t_{ip_i+1}-t_{i3}},$$

$$\ldots, -f_{ip_i-1}\rho^{t_{ip_i-1}-t_{ip_i-1}+t_{ip_i+1}-t_{ip_i-1}} + f_{ip_i-1}\rho^{t_{ip_i+1}-t_{ip_i}})$$

$$= (0, 0, \ldots, f_{ip_i-1}[\rho^{t_{ip_i+1}-t_{ip_i}}[1 - \rho^{t_{ip_i+1}-t_{ip_i-1}}]])$$

$$= (0, 0, \ldots, f_{ip_i-1}[\rho^{d_{ip_i}}[1 - \rho^{d_{ip_i}+d_{ip_i-1}}]])$$

$$= (0, 0, \ldots, \frac{\rho^{d_{ip_i}}}{1-\rho^{2d_{ip_i-1}}}[1 - \rho^{d_{ip_i}+d_{ip_i-1}}])$$

$$= (0', \frac{\rho^{d_{ip_i}}}{1-\rho^{2d_{ip_i-1}}}[1 - \rho^{d_{ip_i}+d_{ip_i-1}}])$$

Thus a predictor of y_{if_j} based on conditional expectation of y_{if_j} given Y is

$$\begin{aligned}\hat{y}_{if_j} &= (A'_{ij}\hat{\xi}BG_{if})' + \sigma'_{fi}\Sigma_i^{-1}(Y_{ij} - (A'_{ij}\hat{\xi}BG_i)') \\ &= (A'_{ij}\hat{\xi}BG_{if})' + (0'_{p_i-1}, \frac{\hat{\rho}^{d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}])(Y_{ij} - (A'_{ij}\hat{\xi}BG_i)')\end{aligned}$$

The variance of $\hat{y}_{if_j} - y_{if_j}$ is

$$var(\hat{y}_{if_j} - y_{if_j}) = \tau_{ij}^2 = \sigma_{2fi} - \sigma'_{fi}\Sigma_i^{-1}\sigma_{fi} + c_{ij}\hat{g}'_i\hat{D}_i\hat{g}_i$$

But

$$\begin{aligned}\sigma'_{fi}\Sigma_i^{-1}\sigma_{fi} &= (0', \frac{\rho^{d_{ip_i}}}{1-\rho^{2d_{ip_i-1}}}[1-\rho^{d_{ip_i}+d_{ip_i-1}}])\sigma^2(\rho^{t_{ip_i+1}-t_{i1}}, \rho^{t_{ip_i+1}-t_{i2}}, \\ &\quad \ldots, \rho^{t_{ip_i+1}-t_{ip_i-1}}, \rho^{t_{ip_i+1}-t_{ip_i}})' \\ &= \sigma^2\frac{\rho^{d_{ip_i}}}{1-\rho^{2d_{ip_i-1}}}[1-\rho^{d_{ip_i}+d_{ip_i-1}}]\rho^{t_{ip_i+1}-t_{ip_i}} \\ &= \sigma^2\frac{\rho^{2d_{ip_i}}}{1-\rho^{2d_{ip_i-1}}}[1-\rho^{d_{ip_i}+d_{ip_i-1}}]\end{aligned}$$

Hence

$$\begin{aligned}\hat{\tau}_{ij}^2 &= \hat{\sigma}^2(1 - \frac{\hat{\rho}^{2d_{iip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}] + c_{ij}\hat{g}'_i\hat{D}_i\hat{g}_i, \\ c_{ij} &= A'_{ij}(A'_iA_i)^{-1}A_{ij} \\ \hat{g}_i &= BG_{if} - BG_i(0'_{p_i-1}, \frac{\hat{\rho}^{d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}])' \\ \hat{D}_i &= (BG_i\hat{\Sigma}_i^{-1}G'_iB')^{-1}.\end{aligned}$$

Prediction of V_2

Let $E(V_{i_1 \times p_i}) = E(V_{1i}, V_{2i}) = (A_{if}\xi BG_{1i}, A_{if}\xi BG_{2i})$. Also $G_{i_k \times p_i} = (G_{1i}, G_{2i})$ and $Cov(V) = G'_i\Sigma G_i$ where Σ has autoregressive structure. As before $G'_i\Sigma G_i$ has a Markov structure. The predictor of V_{2i} based on the conditional expectation of V_{2i} given V_{1i} and Y is given by

$$\hat{V}_{2i} = A_{if}\hat{\xi}BG_{2i} + (V_{1i} - A_{if}\hat{\xi}BG_{1i})\Sigma_{11i}^{-1}\sigma_{12i}.$$

Following the steps similar to that for prediction of y_{if_j}, we get a predictor of V_{2i} and variance of $(\hat{V}_{2i} - V_{2i})$ as follows:

$$\begin{aligned}
\hat{V}_{2i} &= A_{if}\hat{\xi}BG_{2i} + (V_{1i} - A_{if}\hat{\xi}BG_{1i})(0'_{p_i-2}, \frac{\hat{\rho}^{d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}])' \\
\hat{\tau}_i^2 &= \hat{\sigma}^2(1 - \frac{\hat{\rho}^{2d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}] + c_{fi}\hat{d}_i'\hat{D}_i\hat{d}_i \\
c_{fi} &= A_{if}(A_i'A_i)^{-1}A_{if}' \\
\hat{d}_i &= BG_{2i} - BG_{1i}(0'_{p_i-2}, \frac{\hat{\rho}^{d_{ip_i}}}{1-\hat{\rho}^{2d_{ip_i-1}}}[1-\hat{\rho}^{d_{ip_i}+d_{ip_i-1}}]) \\
\hat{D}_i &= (BG_i\hat{\Sigma}_i^{-1}G_i'B')^{-1}.
\end{aligned}$$

3.5.3 Prediction of missing values

As in the cases of balanced and monotone data, it is possible to develop a formula for predicting the missing value in the present case also. In the following, we derive a formula for predicting V_i, based on the conditional expectation.

Prediction of V_i. As in the previous cases, for convenience, suppose V_i is a scalar and it corresponds to the i^{th} position (that is, the i^{th} element) of the vector V. Again for easy presentation we suppress the suffix (i) for the group. Also let

$BG_1 = B_1$ and $BG_2 = B_2$. Re-arrange the elements of V so that $V' = (V_i, V_1', V_2')$. Let $E(V') = (E(V_i), E(V_1'), E(V_2')) = (A_f\xi B_i, A_f\xi B_1, A_f\xi B_2)$, where A_f etc. are appropriately defined vectors or matrices. We similarly partition the covariance matrix of V as

$$cov\begin{pmatrix} V_i \\ V_1 \\ V_2 \end{pmatrix} = \begin{bmatrix} \omega_{ii} & \omega_{i1} & \omega_{i2} \\ \omega_{1i} & \Omega_{11} & \Omega_{12} \\ \omega_{2i} & \Omega_{21} & \Omega_{22} \end{bmatrix} \equiv \begin{bmatrix} \omega_{ii} & \omega_i \\ \omega_i' & \Omega \end{bmatrix}.$$

The minimum mean square predictor of V_i then is

$$E(V_i/V_1, V_2, Y) = A_f\xi B_i + \omega_i'\Omega^{-1}\left[\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} - \begin{pmatrix} A_f\xi B_1 \\ A_f\xi B_2 \end{pmatrix}\right]'$$

Computation of $\omega_i'\Omega^{-1}$. 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 $i-1 \times i-1$ matrix Ω_{11}, the $p-i \times p-i$ matrix Ω_{22}, and the $i-1 \times p-i$ matrix Ω_{12} for Markov structure are

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

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

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

Further,

$$\omega_i' = (\omega_{i1}, \omega_{i2}) =$$

$$\sigma^2(\rho^{t_i-t_1}, \rho^{t_i-t_2}, \ldots, \rho^{t_i-t_{i-2}}, \rho^{t_i-t_{i-1}}, \rho^{t_{i+1}-t_i}, \rho^{t_{i+2}-t_i}, \ldots, \rho^{t_p-t_i}).$$

With some algebra it can be shown that

$$\omega_i'\Omega^{-1} = (0,0,\ldots,0,\frac{\hat{\rho}^{d_{i-1}}(1-\hat{\rho}^{2d_i})}{1-\hat{\rho}^{2(d_{i-1}+d_i)}},\frac{\hat{\rho}^{d_i}(1-\hat{\rho}^{2d_{i-1}})}{1-\hat{\rho}^{2(d_{i-1}+d_i)}},0,\ldots,0)$$

and

$$\omega_i'\Omega^{-1}\omega_i = \sigma^2\frac{\rho^{2d_{i-1}}(1-\rho^{2d_i})+\rho^{2d_i}(1-\rho^{2d_{i-1}})}{1-\rho^{2(d_{j-1}+d_j)}}.$$

Thus the predictor of V_i given V_1, V_2, and Y is given by

$$\begin{aligned}\hat{V}_i = A_f\hat{\xi}B_i &+ \frac{\hat{\rho}^{d_{i-1}}(1-\hat{\rho}^{2d_i})}{1-\hat{\rho}^{2(d_{i-1}+d_i)}}(V_{t_{i-1}} - A_f\hat{\xi}B_{i-1})\\ &+ \frac{\hat{\rho}^{d_i}(1-\hat{\rho}^{2d_{i-1}})}{1-\hat{\rho}^{2(d_{i-1}+d_i)}}(V_{t_{i+1}} - A_f\hat{\xi}B_{i+1})\end{aligned}$$

with the estimated variance

$$\hat{\tau}_i^2 = \frac{\hat{\sigma}^2}{1-\hat{\rho}^2}\frac{[1-\hat{\rho}^{2d_{i-1}}-\hat{\rho}^{2d_i}+\hat{\rho}^{2(d_{i-1}+d_i)}]}{1-\hat{\rho}^{2(d_{i-1}+d_i)}} + c_f\hat{d}'\hat{D}\hat{d}.$$

where

$$\hat{d} = B_i - [\frac{\hat{\rho}^{d_{i-1}}(1-\hat{\rho}^{2d_i})}{1-\hat{\rho}^{2(d_{i-1}+d_i)}}B_{i-1} + \frac{\hat{\rho}^{d_i}(1-\hat{\rho}^{2d_{i-1}})}{1-\hat{\rho}^{2(d_{i-1}+d_i)}}B_{i+1}]$$

3.6 Computer programs

<u>PROGRAM 1</u>

```
integer n,p,n1,n2,p1,p2,g,k,np1,np2,kg
parameter (p1=7,p2=6,n1=13,n2=20)
parameter (n=20,p=16,g=2,k=3,kg=6,
```

```
c npl=91,np2=120)
  real rho,rhat,sigma,rho1,tr1,tr2
  dimension u(n,p),y1(n1,p1),y2(n2,p2),a1(n1,g),a2(n2,g)
  dimension t1(p1,1),t2(p2,1),b1(k,p1),b2(k,p2),sum1(k * g,k * g)
  dimension sum2(k * g,1),a1t(g,n1),a2t(g,n2),b1t(p1,k)
  dimension b2t(p2,k),aa1(g,g),aa2(g,g),sum1i(k * g,k * g)
  dimension bvi1(k,p1),bvi2(k,p2),bvib1(k,k),bvib2(k,k)
  dimension sz11(k * g,k * g),sz12(k * g,k * g),sz21(k * g,n1 * p1)
  dimension sz22(k * g,n2 * p2),vecy1(n1 * p1,1),vecy2(n2 * p2,1)
  dimension sz31(k * g,1),sz32(k * g,1),veczhat(k * g,1)
  dimension zhat(g,k),res1(p1,p1),res2(p2,p2)
  dimension a1z(n1,k),azb1(n1,p1),e1(n1,p1),e1t(p1,n1)
  dimension a2z(n2,k),azb2(n2,p2),e2(n2,p2),e2t(p2,n2)
dimension ft11(p1-1,1),ft12(p2-1,1),vi1(p1,p1),vi2(p2,p2)
  common /first/ ev1(p1-1,1),ev2(p1-1,1),ev3(p2-1,1),ev4(p2-1,1)
  common /secnd/ id1(p1-1,1),id2(p2-1,1)
c     external f
  rho=0.5
  do i=1,n
  read(1, * )(u(i,j),j=1,p)
  end do
  do i=1,n1
```

```
      do j=1,6
      y1(i,j)=u(i,j)
      end do
      y1(i,7)=u(i,8)
      end do

      do i=1,n2
      y2(i,1)=u(i,9)
      y2(i,2)=u(i,10)
      y2(i,3)=u(i,11)
      y2(i,4)=u(i,13)
      y2(i,5)=u(i,14)
      y2(i,6)=u(i,16)
      end do

c     Creating ai matrices
      call zer(n1,g,a1)
      call zer(n2,g,a2)
      do i=1,n1
      a1(i,1)=1.0
      end do

      do i=1,n2
      a2(i,2)=1.0
```

```
      end do
      do i=1,5
      t1(i,1)=i
      end do
      t1(6,1)=7.0
      t1(7,1)=11.0
      do i=1,3
      t2(i,1)=i
      end do
      t2(4,1)=5.0
      t2(5,1)=7.0
      t2(6,1)=11.0
      do i=1,p1-1
      id1(i,1)=t1(i+1,1)-t1(i,1)
      end do
      do i=1,p2-1
      id2(i,1)=t2(i+1,1)-t2(i,1)
      end do
c     Creating bi matrices
      do j=1,p1
      b1(1,j)=1.0
```

```
        b1(2,j)=t1(j,1)
        b1(3,j)=t1(j,1) * * 2
        end do

        do j=1,p2
        b2(1,j)=1.0
        b2(2,j)=t2(j,1)
        b2(3,j)=t2(j,1) * * 2
        end do

        call xp(a1,n1,g,a1t)
        call xp(a2,n2,g,a2t)
        call xp(b1,k,p1,b1t)
        call xp(b2,k,p2,b2t)
        call xxp(a1,n1,g,aa1)
        call xxp(a2,n2,g,aa2)

print * ,'Beginning do loop'
1000      continue
        call vin(p1,rho,id1,ft11,vi1)
        call vin(p2,rho,id2,ft12,vi2)
        call mm(k,p1,b1,vi1,p1,bvi1)
        call mm(k,p1,bvi1,b1t,k,bvib1)
        call mm(k,p2,b2,vi2,p2,bvi2)
```

```
      call mm(k,p2,bvi2,b2t,k,bvib2)
      call kr(k,k,bvib1,g,g,aa1,sz11)
      call kr(k,k,bvib2,g,g,aa2,sz12)
      call kr(k,p1,bvi1,g,n1,a1t,sz21)
      call kr(k,p2,bvi2,g,n2,a2t,sz22)
      call vec(n1,p1,y1,vecy1)      call vec(n2,p2,y2,vecy2)
      call mm(kg,np1,sz21,vecy1,1,sz31)
      call mm(kg,np2,sz22,vecy2,1,sz32)
c     Creating zhat
      do i=1,kg
      do j=1,kg
      sum1(i,j)=sz11(i,j)+sz12(i,j)
      end do
      end do
      do i=1,kg
      sum2(i,1)=sz31(i,1)+sz32(i,1)
      end do
      call linrg(kg,sum1,kg,sum1i,kg)
      call mm(kg,kg,sum1i,sum2,1,veczhat)
      ki=1
      do i=1,k
```

```
do j=1,g
zhat(j,i)=veczhat(ki,1)
ki=ki+1
end do
end do
print * ,'The estimates: zhat'
do i=1,g
write (6, * )(zhat(i,j),j=1,k)
end do
call mm(n1,g,a1,zhat,k,a1z)
call mm(n1,k,a1z,b1,p1,azb1)
do i=1,n1
do j=1,p1
e1(i,j)=y1(i,j)-azb1(i,j)
end do
end do
call mm(n2,g,a2,zhat,k,a2z)
call mm(n2,k,a2z,b2,p2,azb2)
do i=1,n2
do j=1,p2
e2(i,j)=y2(i,j)-azb2(i,j)
```

```
end do
end do
do j=1,p1-1
summ2=0.0
summ3=0.0
do i=1,n1
summ2=summ2+e1(i,j) * * 2+e1(i,j+1) * * 2
summ3=summ3+e1(i,j) * e1(i,j+1)
end do
ev1(j,1)=summ2
ev2(j,1)=summ3
end do
do j=1,p2-1
summ4=0.0
summ5=0.0
do i=1,n2
summ4=summ4+e2(i,j) * * 2+e2(i,j+1) * * 2
summ5=summ5+e2(i,j) * e2(i,j+1)
end do
ev3(j,1)=summ4
ev4(j,1)=summ5
```

```
end do
call xp(e1,n1,p1,e1t)
call mm(p1,n1,e1t,e1,p1,res1)
call trab(p1,vi1,res1,tr1)
call xp(e2,n2,p2,e2t)
call mm(p2,n2,e2t,e2,p2,res2)
call trab(p2,vi2,res2,tr2)
sigma=(tr1+tr2)/(n1 * p1+n2 * p2)
print * ,'sigma=',sigma
rhat=rho
call rhohat(sigma,rho)
comp=abs(rhat-rho)
if (comp.ge.10e-5) go to 1000
print * ,'rhat=',rhat
stop
end
```

```
c     Subroutines start from here
c     Creating rohat
      subroutine rhohat(sig,rr)
      integer itmax,info,nrt,ier
      real f,eps,eta,r,rr,err,err2,rguess
      external f
      common sig1
      sig1=sig
      err=1.0e-5
      err2=1.0e-5
      eps=1.0e-5
      eta=1.0e-2
      itmax=100
      nrt=1
      rguess=rr
      call zreal(f,err,err2,eps,eta,nrt,itmax,rguess,r,info)
      rr=r
      print * ,'r=',r
      return
      end

      function f(r)
```

```
real r,sumn1,sig1,sumn2,sumn3,sumn4,sumn5,sumn6
common sig1
parameter (p1=7,p2=6,n1=13,n2=20)
common /first/ ev1(6,1),ev2(6,1),ev3(5,1),ev4(5,1)
common /secnd/ id1(6,1),id2(5,1)
sumn1=0.0
sumn2=0.0
sumn3=0.0
sumn4=0.0
sumn5=0.0
sumn6=0.0
do j=1,p1-1
sumn1=sumn1+2.0 * id1(j,1) * (1.0/(1.-r * * (2 * id1(j,1)))) *
c (r * * (2 * id1(j,1)-1))
end do
do j=1,p2-1
sumn2=sumn2+2.0 * id2(j,1) * (1.0/(1.-r * * (2 * id2(j,1)))) *
c (r * * (2 * id2(j,1)-1))
end do
do j=1,p1-1
sumn3=sumn3+(ev1(j,1) * ((1.0/(1.-r * * (2 * id1(j,1)))) * * 2) *
```

```
c (2.0 * id1(j,1)) * (r * * (2 * id1(j,1)-1)))
 end do
 do j=1,p1-1
 sumn5=sumn5+(ev2(j,1) * ((1.0/(1.-r * * (2 * id1(j,1)))) * * 2) *
c (2.0 * id1(j,1)) * (r * * (id1(j,1)-1)) * (1+r * * (2 * id1(j,1))))
 end do
 do j=1,p2-1
 sumn4=sumn4+(ev3(j,1) * ((1.0/(1.-r * * (2 * id2(j,1)))) * * 2) *
c (2.0 * id2(j,1)) * (r * * (2 * id2(j,1)-1)))
 end do
 do j=1,p2-1
 sumn6=sumn6+(ev4(j,1) * ((1.0/(1.-r * * (2 * id2(j,1)))) * * 2) *
c (2.0 * id2(j,1)) * (r * * (id2(j,1)-1)) * (1+r * * (2 * id2(j,1))))
 end do
 f=sig1 * (n1 * sumn1+n2 * sumn2)-(sumn3-sumn5+sumn4-sumn6)
 return
 end

c    Multiplication xs=x * s
 subroutine mm(n,ip,x,s,iq,xs)
 dimension x(n,ip),s(ip,iq),xs(n,iq)
 do i=1,n
```

```
      do k=1,iq
      sum=0.0
      do j=1,ip
      sum=sum+x(i,j) * s(j,k)
      end do
      xs(i,k)=sum
      end do
      end do
      return
      end

c     pp is transpose of x
      subroutine xp(x,n,ip,pp)
      dimension x(n,ip),pp(ip,n)
      do i=1,n
      do j=1,ip
      pp(j,i)=x(i,j)
      end do
      end do
      return
      end

c     s is transpose(x) * x
```

```
      subroutine xxp(x,n,ip,s)
      dimension x(n,ip),s(ip,ip)
      do i=1,ip
      do j=1,ip
      sum=0.0
      do k=1,n
      sum=sum+x(k,i) * x(k,j)
      end do
      s(i,j)=sum
      s(j,i)=sum
      end do
      end do
      return
      end

c     Kronecker product
      subroutine kr(m,l,a,p,q,b,c)
      integer m,l,p,q
      dimension a(m,l),b(p,q),c(m * p,l * q)
      do i=1,m
      do j=1,l
      do ik=1,p
```

```
      do jk=1,q
      c((i-1) * p+ik,(j-1) * q+jk)=a(i,j) * b(ik,jk)
      end do
      end do
      end do
      end do
      return
      end

c     Creating inverse matrix
      subroutine vin(m,x,id,f,vi)
      integer m
      real x
      dimension vi(m,m),id(m-1,1),f(m-1,1)
      external zer
      do i=1,m-1
      f(i,1)=1.0/(1.0-x * * (2 * id(i,1)))
      end do
      call zer(m,m,vi)
      do i=2,m-1
      vi(i,i)=f(i-1,1)+f(i,1)-1.0
      vi(i,i-1)=-1.0 * f(i-1,1) * x * * id(i-1,1)
```

```
      vi(i,i+1)=-1.0 * f(i,1) * x * * id(i,1)
      end do
      vi(1,1)=f(1,1)
      vi(m,m)=f(m-1,1)
      vi(1,2)=-1.0 * f(1,1) * x * * id(1,1)
      vi(m,m-1)=-1.0 * f(m-1,1) * x * * id(m-1,1)
      return
      end

c     Creating a m by m matrix of zeroes
      subroutine zer(m,l,z)
      integer m,l
      real z(m,l)
      do i=1,m
      do j=1,l
      z(i,j)=0.0
      end do
      end do
      return
      end
c     Creating subroutine for vec operation
```

```
      subroutine vec(m,l,a,b)
      integer m,l
      dimension a(m,l),b(m * l,1)
      do j=1,l
      do i=1.m
      b((j-1) * m+i,1)=a(i,j)
      end do
      end do
      return
      end
c     Creating trace(a * b)
      subroutine trab(m,a.b,tr)
      real tr
      integer m
      dimension a(m,m),b(m,m)
      sum=0.0
      do i=1,m
      do j=1,m
      sum=sum+a(i,j) * b(j,i)
      end do
      end do
```

```
tr=sum
return
end
```

Chapter 4

Non-Linear Growth Curves

4.1 Introduction

We have considered the problems of fitting the polynomial curves to the growth data in the previous two chapters. The polynomial model can provide useful predictive information and may be the best approach if the growth information has been collected over a limited range of growth cycle. However, in general, the parameters of such models may have unsatisfactory biological interpretations. Hence, in practical problems, nonlinear functions are used to fit the growth data.

The problem of fitting nonlinear functions to growth data has been considered by many in the literature. One of the first models is by Gompertz (1825), followed by Verhulst (1845), von Bertalanffy (1957), and Richards (1959). In fact, the functions considered by these authors are recognized by their names, except the logistic function introduced by Verhulst.

A nonlinear growth curve model can be written in the form

$$\begin{aligned} y &= f(\underline{x}, \underline{\theta}) + \epsilon \\ &= f(x_1, x_2, ..., x_k, \theta_1, \theta_2, ..., \theta_p) + \epsilon, \end{aligned}$$

where y is the observed response variable, f is a specified nonlinear function, $x_1, x_2, ..., x_k$ are predictor variables and $\theta_1, \theta_2, ..., \theta_p$ are the unknown parameters.

We give the forms of some of the popular nonlinear functions in section 4.3. For more details and other interpretations, see Ratkowsky (1983) and Seber and Wild (1989). Among these curves, the von Bertalanffy growth curve has been widely used in fisheries science. Kimura (1980) considered the statistical analysis of this curve using the likelihood method and illustrated the analysis on a fisheries growth data. In the next section, we present this important data set in Table 4.1 and utilize this in the later sections to illustrate our methods. The analysis done by Kimura (1980) has been adopted using simple SAS codes by Lakkis and Jones (1992). These codes can be easily modified to fit any of the nonlinear models. For example, we provide these codes for Richards curve at the end of this chapter.

We note that the sample sizes for different age groups, in the data set of Table 4.1 are different. In section 4.4, we present the analysis (using SAS codes) by taking the sample sizes for different ages being different into consideration. In section 4.5, we consider a more general model, where the fish are assumed to have been randomly selected from a population. This amounts to using a random effects term in the model.

Finally, in section 4.6, we consider a nonlinear growth curve in a multivariate setup. Assuming certain popular covariance structures for the covariance matrix of the error vector, we show how the ML estimates of the parameters of the covariance matrix can be computed using SAS programs.

4.2 Fish data set

The Pacific hake, Merluccius productus, is a common gadid fish that ranges from the gulf of California to the gulf of Alaska. Dark (1975) collected data on Pacific hake, off California, Oregon, and Washington during 1965-69, and studied the age and growth of these fish. Biological data were collected from both commercial fishery and aboard research vessels.

The specimens were first dissected to determine the sex of a fish. Then they were measured from snout to the fork of the tail to determine the length. An Otolith was removed for age determination. The specimens were also weighed to determine the weight. The number of fish caught for study each year were different. Further, the number of fish in each age group were very different. The average body lengths at various ages (with the corresponding sample sizes) for male and female hakes taken off California, Oregon, and Washington coast are given in Table 4.1. For more details about these data and for other considerations see Dark (1975). For a statistical analysis of these data see Kimura (1980) and Lakkis and Jones (1992).

Table 4.1: Average length at various ages for male and female Pacific hake taken off California, Oregon, and Washington during 1965-69 (adopted from Dark 1975).

Age (years)	Female		Male	
	Sample size	Mean length (cm)	Sample size	Mean length (cm)
1.0	385	15.40	385	15.40
2.0	36	28.03	28	26.93
3.3	17	41.18	13	42.23
4.3	135	46.20	83	44.59
5.3	750	48.23	628	47.93
6.3	1073	50.26	1134	49.67
7.3	1459	51.82	1761	50.87
8.3	626	54.27	432	52.30
9.3	199	56.98	93	54.77
10.3	97	58.93	21	56.43
11.3	44	59.00	8	55.88
12.3	11	60.91	-	-
13.3	6	61.83	-	-

4.3 Different nonlinear curves

For many types of growth data, the growth curve is an S-shaped or is of what is called Sigmoidal pattern. This is because, the current growth rate is proportional to the current size and the remaining growth. The four curves which we are going to discuss are most popular in the literature. These are, Logistic curve, von Bertalanffy curve, Gompertz curve, and the Richards curve.

Logistic Curve: With x denoting time and f denoting the size,

$$f(x) = \frac{\alpha}{1 + \beta \exp(-Kx)}. \tag{4.1}$$

Here α is the limiting size and K is the constant of proportionality. This curve can be used to describe growth in the size of a population or an organ or the biochemical nature of a certain growth processes.

von Bertalanffy curve: The function is of the form

$$f(x) = l_\infty(1 - exp(-K(x - \gamma))), \tag{4.2}$$

where l_∞ is the asymptotic size, K acts as a scale parameter on x thus governing the rate of growth, and γ is the point of inflection. This curve is used extensively in fisheries research.

Gompertz curve: The function is of the form

$$f(x) = \alpha\, exp(-exp(-K(x - \gamma))), \tag{4.3}$$

where α, K, and γ have the same meaning as above. The Gompertz curve has the property that any power of f again is a Gompertz curve. The Gompertz curve is used for population studies and in the growth model for the heart of a chicken.

Richards curve: Richards curve is the most general of Sigmoidal types of curves. It has the form

$$f(x) = \alpha[1 - (\delta - 1)exp(-K(x - \gamma))]^{\frac{1}{1-\delta}}, \delta \neq 1, \tag{4.4}$$

where α is the final size, the point of inflection is at the time point $x = \gamma$, K is a scale parameter on x, and δ is a measure for degree of asymmetry to the growth curve. If $\delta < 1$, then one has to put the additional restriction, $(1 - \delta)exp(K\gamma) \leq 1$, to ensure that $0 \leq f \leq \alpha$. All the previous three curves are special cases of the Richards curve. For example, $\delta = 2$ gives the logistic curve, $\delta = 0$ gives the von Bertalanffy curve, and the curve obtained by taking the limit as $\delta \to 1$, is the Gompertz curve.

The data we come across in this area generally are average growth of a particular characteristic like, the length measured at different ages (the time factor) for male and female fishes (the groups). A model for fitting these data can be written as

$$y_i = \mu(\alpha_i, K_i, \gamma_i, \delta_i, x) + \epsilon_i, \; i = 1, 2, ..., g, \tag{4.5}$$

where y_i is a $m_i \times 1$ vector of average lengths for the i^{th} group, x is a $m_i \times 1$ vector of ages, $\mu(\alpha_i, K_i, \gamma_i, \delta_i, x)$ is a nonlinear function and the error vector ϵ_i is distributed as $N_{p_i}(0,\sigma^2 I)$. If $\mu(\alpha_i, K_i, \gamma_i, \delta_i, x) = \alpha_i[1-(\delta_i-1)exp(-K_i(x-\gamma_i))]^{\frac{1}{1-\delta_i}}$, $\delta_i \neq 1$ (Richards curve), then letting

$$\begin{aligned} S(\alpha_1,..,\alpha_g,K_1,..,K_g,\gamma_1,..,\gamma_g,\delta_1,..,\delta_g) &= \sum_{i=1}^{g}\sum_{k=1}^{p_i}(y_{ik}-\alpha_i[1-(\delta_i-1) \\ &\quad exp(-K_i(x_k-\gamma_i))]^{\frac{1}{1-\delta_i}})^2, \end{aligned}$$

we can write the likelihood function as :

$$L = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} exp(-S(\alpha_1,..,\alpha_g,K_1,..,K_g,\gamma_1,..,\gamma_g,\delta_1,..,\delta_g)/2\sigma^2),$$

where $N = \sum_{i=1}^{g} m_i$. For a fixed σ^2, maximizing L with respect to the parameters $\alpha_i, K_i, \gamma_i, \delta_i$, $i = 1, 2, ..., g$ is same as minimizing S with respect to the corresponding parameters. Hence, it follows that the maximum likelihood estimators of $\alpha_i, K_i, \gamma_i, \delta_i, i = 1, 2, ..., g$ are the least squares estimators. Next, taking the first derivative of the log likelihood function with respect to σ^2, setting it to zero and solving the equation for σ^2, we get the MLE of σ^2. That is,

$$lnL = -\frac{N}{2}ln(2\pi\sigma^2) - \frac{S(\alpha_1,..,\alpha_g,K_1,..,K_g,\gamma_1,..,\gamma_g,\delta_1,..,\delta_g)}{2\sigma^2},$$

$$\frac{\partial lnL}{\partial\sigma^2} = 0 \Rightarrow$$

the maximum likelihood of σ^2 is given by

$$\hat{\sigma^2} = \frac{1}{N}S(\hat{\alpha}_1,..,\hat{\alpha}_g,\hat{K}_1,..,\hat{K}_g,\hat{\gamma}_1,..,\hat{\gamma}_g,\hat{\delta}_1,..,\hat{\delta}_g), \tag{4.6}$$

where $\hat{\theta}_i$ is the maximum likelihood (equivalently least equares) estimator of the parameter θ_i.

In section 4.7, we give a SAS program (PROGRAM 1) for calculating the MLE (4.6) for the fish data (Table 4.1). The results of this program are: $\hat{\alpha}_1 = 56.5$, $\hat{\alpha}_2 = 66.2$, $K_1 = 0.38$, $K_2 = 0.14$, $\delta_1 = -0.3$, $\delta_2 = -1.28$, $\gamma_1 = -0.3659$ and $\gamma_2 = -5.0$.

4.4 von Bertalanffy model with unequal sample sizes under incomplete data

Consider the von Bertalanffy model in the following form:

$$y_{ik} = \mu(l_{\infty i}, K_i, t_{0i}, x_k) + \epsilon_{ik}, k = 1, 2, ..., m_i, \ i = 1, 2, ..., g, \tag{4.7}$$

with $\mu(l_{\infty i}, K_i, t_{0i}, x_k) = l_{\infty i}(1 - exp(-K_i(x_k - t_{0i}))$ and the errors ϵ_{ik} having $N(0, \sigma^2)$ distribution. Here g represents the number of groups.

In fisheries, the data generally are such that, for every age group, the sample size is different. Thus it is essential that the sample sizes be part of the model and its analysis. To accommodate the sample size we suggest the following model:

Let $y_i = (y_{i1}, y_{i2}, ..., y_{im_i})'$, and the model be written as

$$y_i = \mu(l_{\infty i}, K_i, t_{0i}, x) + \epsilon_i, \tag{4.8}$$

where y_i is a $m_i \times 1$ data vector, ϵ_i is a $m_i \times 1$ vector of errors, and the $m_i \times 1$ vector

$$\mu(x) = \mu(l_{\infty i}, K_i, t_{0i}, x) = \begin{pmatrix} \mu(l_{\infty i}, K_i, t_{0i}, x_1) \\ \mu(l_{\infty i}, K_i, t_{0i}, x_2) \\ . \\ . \\ . \\ \mu(l_{\infty i}, K_i, t_{0i}, x_{m_i}) \end{pmatrix}$$

such that $\mu(l_{\infty i}, K_i, t_{0i}, x_k) = l_{\infty i}(1 - exp(-K_i(x_k - t_{0i})))$. We assume that ϵ_i is distributed as $N_{m_i}(0, \sigma^2 V_i)$, where

$$V_i = \begin{pmatrix} \frac{1}{n_{i1}} & 0 & 0 & . & . & . & 0 \\ 0 & \frac{1}{n_{i2}} & 0 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 & \frac{1}{n_{im_i}} \end{pmatrix},$$

n_{ik} being the sample size in the i^{th} group at age x_k. Let $z_i = V_i^{-\frac{1}{2}} y_i$. Then

$$cov(z_i) = V_i^{-\frac{1}{2}}(\sigma^2 V_i) V_i^{-\frac{1}{2}} = \sigma^2 I$$

and

$$E(z_i) = V_i^{-\frac{1}{2}} \mu(x)$$

$$= \begin{pmatrix} \sqrt{(n_{i1})} & 0 & 0 & . & . & 0 \\ 0 & \sqrt{(n_{i2})} & 0 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & \sqrt{(n_{im_i})} \end{pmatrix} \begin{pmatrix} \mu(l_{\infty i}, K_i, t_{0i}, x_1) \\ \mu(l_{\infty i}, K_i, t_{0i}, x_2) \\ . \\ . \\ . \\ \mu(l_{\infty i}, K_i, t_{0i}, x_m) \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{(n_{i1})}\mu(l_{\infty i}, K_i, t_{0i}, x_1) \\ \sqrt{(n_{i2})}\mu(l_{\infty i}, K_i, t_{0i}, x_2) \\ . \\ . \\ . \\ \sqrt{(n_{ip_i})}\mu(l_{\infty i}, K_i, t_{0i}, x_{m_i}) \end{pmatrix} = \nu_i \ (say)$$

Thus the transformed model can be written as

$$z_{ik} = \sqrt{n_{ik}} l_{\infty i}(1 - exp(-K_i(x_k - t_{0i}))) + \eta_{ik}\,,\ k = 1, 2, ..., m_i,\ i = 1, 2, ..., g \quad (4.9)$$

where z_{ik} is transformed data and η_{ik} are distributed independently as $N(0, \sigma^2)$. This model now is similar to (4.7). Letting

$$\begin{aligned} S &= S(l_{\infty 1}, .., l_{\infty g}, K_1, .., K_g, t_{01}, .., t_{0g}) \\ &= \sum_{i=1}^{g} \sum_{k=1}^{p_i} (z_i - \sqrt{n_{ik}} l_{\infty i}.(1 - exp(-K_i(x_k - t_{0i}))))^2 \end{aligned}$$

The likelihood function can be written as:

$$L = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} exp(-S(l_{\infty 1},..,l_{\infty g},K_1,..,K_g,t_{01},..,t_{0g})/2\sigma^2)$$

where $N = \sum_{i=1}^{g} m_i$. The ML estimators of $l_{\infty i}, K_i, t_{0i}$, for $i = 1,2,...,g$ are the least squares estimators minimizing (4.10). And the maximum likelihood estimator of σ^2 is given by

$$\hat{\sigma^2} = \frac{1}{N} S(\hat{l}_{\infty 1},..,\hat{l}_{\infty g},\hat{K}_1,..,\hat{K}_g,\hat{t}_{01},..,\hat{t}_{0g}), \tag{4.10}$$

where $\hat{l}_{\infty 1},..,\hat{l}_{\infty g},\hat{K}_1,..,\hat{K}_g,\hat{t}_{01},..,\hat{t}_{0g}$ are the estimates obtained by nonlinear least squares minimization of (4.10).

Once we have the ML estimators for all the parameters we can perform the likelihood ratio tests as given in Kimura (1980).

<u>Example 4.4.1:</u>

For the fish data in Table 4.1, we use PROC NLIN (NLIN procedure of SAS software) to fit a von Bertalanffy model with different sample sizes. From Table 4.1 we see that the data on the average lengths at different ages for male and female fishes are given. The sample sizes are also provided at different ages, for both, the male and the female fishes. The procedure PROC NLIN requires that the initial estimates of the parameters $l_{\infty 1}, l_{\infty 2}, K_1, K_2, t_{01}, t_{02}$ be provided. For that, we first plot, for each group (male and female), the length as a function of age, and get the initial estimates for $l_{\infty 1}, l_{\infty 2}, t_{01}, t_{02}$ by visually examining the plot (see Lakkis and Jones (1992)). The intial estimates for K_1 and K_2 are obtained by substituting the corresponding initial estimates for $l_{\infty i}, t_i$, the average $\bar{y}_i$ for y_{ik} and the average x for x_k in (4.7). The SAS program is provided in PROGRAM 2 of section 4.7. The results of this program are $\hat{l}_{\infty 1} = 54.17$, $\hat{l}_{\infty 2} = 58.23$, $\hat{K}_1 = 0.4$, $\hat{K}_2 = 0.32$, $\hat{t}_{01} = 0.17$ and $\hat{t}_{02} = 0.03$.

4.5 von Bertalanffy model under equicorrelation structure

The motivation for adopting the model (4.11) shown below, comes from the special way in which the fish data presented in Table 4.1 are collected. It is clear that we don't have data for the same fish at different occasions. Rather, at every occasion, certain number of fish are collected from a pool and they are classified as belonging to a certain age group. Further, the male and female fishes are separated to form two groups. The number of fish caught at different occasions is different and this leads to unequal sample sizes. Although the fish on which the data are collected at different occasions are different, they share a common habitat. This leads us to believe that the following model would be more appropriate to analyze the fish data of Table 4.1. Consider the model,

$$y_i = \mu(l_{\infty i}, K_i, t_{0i}, x) + \epsilon_i \tag{4.11}$$

such that y_i is a $m_i \times 1$ data vector (a vector of lengths), the k^{th} element of the mean vector $\mu(l_{\infty i}, K_i, t_{0i}, x_k)$ is $l_{\infty i}(1 - exp(-K_i(x_k - t_{0i})))$. The error vector ϵ_i is assumed to be distributed as $N_{m_i}(0, \sigma^2 V_i)$, where

$$V_i = \begin{pmatrix} \frac{1}{n_{i1}} & \frac{\rho}{\sqrt{n_{i1}n_{i2}}} & \frac{\rho}{\sqrt{n_{i1}n_{i3}}} & \cdot & \cdot & \cdot & \frac{\rho}{\sqrt{n_{i1}n_{im_i}}} \\ \frac{\rho}{\sqrt{n_{i2}n_{i1}}} & \frac{1}{n_{i2}} & \frac{\rho}{\sqrt{n_{i2}n_{i3}}} & \cdot & \cdot & \cdot & \frac{\rho}{\sqrt{n_{i2}n_{im_i}}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{\rho}{\sqrt{n_{im_i}n_{i1}}} & \frac{\rho}{\sqrt{n_{im_i}n_{i2}}} & \cdot & \cdot & \cdot & \cdot & \frac{1}{n_{im_i}} \end{pmatrix}$$

The matrix V_i has equicorrelation structure except for the different sample sizes. We rewrite the model (4.11) by transforming the data so that the covariance matrix of the error vector is $\sigma^2[(1-\rho)I_{m_i} + \rho J_{m_i}]$. For that let

$$D_i = \begin{pmatrix} \frac{1}{n_{i1}} & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \frac{1}{n_{i2}} & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & \frac{1}{n_{im_i}} \end{pmatrix} \text{ and } W_i = \begin{pmatrix} 1 & \rho & \rho & \cdot & \cdot & \cdot & \rho \\ \rho & 1 & \rho & \cdot & \cdot & \cdot & \rho \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho & \rho & \cdot & \cdot & \cdot & \rho & 1 \end{pmatrix}$$

Now we make the transformation $z_i = C_i^{-\frac{1}{2}} y_i$. Then

$$E(z_i) = C_i^{-\frac{1}{2}} \mu_i$$

$$var(z_i) = C_i^{-\frac{1}{2}}(\sigma^2 V_i)C_i^{-\frac{1}{2}} = C_i^{-\frac{1}{2}}(\sigma^2 C_i^{\frac{1}{2}} W_i C_i^{\frac{1}{2}})C_i^{-\frac{1}{2}} = \sigma^2 W_i$$

The model for transformed data is

$$z_{im_i \times 1} = \nu_i + \eta_{im_i \times 1}, i = 1, 2, ..., g, \tag{4.12}$$

where $\nu_i = C_i^{-\frac{1}{2}} \mu_i$ and the error vector η_i is distributed as $N_{m_i}(0, \sigma^2 W_i)$.

Glasbey (1979) considered the following model for a single group,

$$y_t = f_t(\theta) + \epsilon_t, \ t = 1, 2, ..., m, \tag{4.13}$$

where $f_t(\theta)$ is a nonlinear growth function depending on certain unknown parameters θ. Assuming the first order autoregressive process for the errors, he has provided an algorithm for computing the maximum likelihood estimates of the parameters involved. Our problem is similar, but with the equicorrelated errors.

Suppose we consider the model (4.13) with the observations y_t made at time points $t = t_1, t_2, ..., t_m$ (unequally spaced time points). Then the errors $(\epsilon_{t_1}, \epsilon_{t_2}, ..., \epsilon_{t_m})$ will have a Markov covariance structure. Estimating the parameters under this structure has been discussed in chapter 3, when the polynomial models were fit for growths. For the model (4.13) with Markov structure, we provide a SAS macro program for fitting a nonlinear growth curve in PROGRAM 3 (on fish data). This program includes Glasbey (1979)'s model as a particular case. The results of this program are $\hat{\rho} = -0.045$, $\hat{l}_{\infty 1} = 52$, $\hat{l}_{\infty 2} = 53$, $\hat{K}_1 = 0.73$, $\hat{K}_2 = 0.65$, $\hat{t}_{01} = 0.52$ and $\hat{t}_{02} = 0.49$.

4.6 Analysis of nonlinear models for multivariate data

So far in this chapter, we have considered the model of the form

$$y_{ij} = \mu(\theta_i, x_{ij}) + \epsilon_{ij},\ j = 1, 2, ..., m_i,\ i = 1, 2, ..., g \tag{4.14}$$

where g is the number of groups, m_i is the number of subjects in the i^{th} group, μ is a certain nonlinear function, and the vector $(\epsilon_{i1}, ..., \epsilon_{im_i})'$ possibly has a covariance structure, like equicorrelation, autoregressive, or a Markov structure. In the following, we consider a model to accommodate multivariate observations on each subject or individual.

General multivariate nonlinear models under the mixed effects model setup have been considered in the literature and algorithms to compute the MLEs have been provided by many authors. For example, see the works of Lindstrom and Bates (1990), Palmer, Phillips, and Smith (1991), Hirst, Boyle, Zerbe and Wilkening (1991), Frey and Muller (1992), and Vonesh and Carter (1992). Our aim in this section however is to provide easy computational tools using easily available software, rather than giving another general model. For our discussion we consider the model of the form

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk},$$
$$i = 1, 2, ..., g; j = 1, 2, ..., n_i; k = 1, 2, ..., p_{ij}.$$

Defining $y_{ij} = (y_{ij1}, ..., y_{ijp_{ij}})'$ and $\epsilon_{ij} = (\epsilon_{ij1}, ..., \epsilon_{ijp_{ij}})'$, we can rewrite the model as

$$y_{ij_{p_{ij}\times 1}} = \mu_i + \epsilon_{ij_{p_{ij}\times 1}}, \tag{4.15}$$

with ϵ_{ij} distributed as $N(0, \sigma^2 V_{ij}(\rho))$ and μ_i is the appropriate vector of means. We provide SAS macro programs to compute the ML estimates of the parameters under the model $\mu_{ij} = l_{\infty i}(1 - exp(-K_i(x_{jk} - t_{0i})))$ in PROGRAM 4 for plasma data of Table 2.1. The results of this program are $\hat{\rho} = 0.81$, $\hat{\sigma}^2 = 2.29$, $\hat{l}_{\infty 1} = 66.95$, $\hat{l}_{\infty 2} = 300.25$, $\hat{K}_1 = 0.003$, $\hat{K}_2 = 017.79$, $\hat{t}_{01} = -9.1$ and $\hat{t}_{02} = -11$.

4.7 Computer programs

PROGRAM 1

/ * The following code fits Richards model * /

data fish;

input length d1 d2 age;

cards;

proc nlin data=fish maxiter=100;

parms α_1=55 K_1=0.276 γ_1=0 δ_1=0 α_2=60 K_2=0.23 γ_2=0 δ_2=0;

model $length = \alpha_1 d1[(1 + (\delta_1 - 1)exp(-K_1(age - \gamma_1)))]^{(\frac{1}{(\delta_1-1)})}+$

$\alpha_2 d2[(1 + (\delta_2 - 1)exp(-K_2(age - \gamma_2)))]^{(\frac{1}{(\delta_2-1)})}$;

PROGRAM 2

```
options ls=78 ps=45 nodate nonumber;

title 'Nonlinear Growth Curve Analysis under variance sigsq/n';

data fish;

input length size d1 d2 age@@;

sqsize=sqrt(size);tlength=sqsize*length;

lines;

15.40 385 1 0 1.0 26.93 28 1 0 2.0 42.23 13 1 0 3.3 44.59 83 1 0 4.3 47.63 628 1 0 5.3

49.67 1134 1 0 6.3 50.87 1761 1 0 7.3 52.30 432 1 0 8.3 54.77 93 1 0 9.3 56.43 21 1 0

10.3 55.88 8 1 0 11.3

15.40 385 0 1 1.0 28.03 36 0 1 2.0 41.18 17 0 1 3.3 46.20 135 0 1 4.3 48.23 750 0 1

5.3

50.26 1073 0 1 6.3 51.82 1459 0 i 7.3 54.27 626 0 1 8.3 56.98 199 0 1 9.3 58.93 97 0

1

10.3 59.00 44 0 1 11.3 60.91 11 0 1 12.3 61.83 6 0 1 13.3;



/* The following code fits von Bertalanffy model under variance sigma-square/n*/

proc nlin data=fish maxiter=400;

parms l1=55 k1=.276 t1=0 l2=60 k2=.23 t2=0;

model tlength=l1*d1*sqsize*(1-exp(-k1*(age-t1)))+l2*d2*sqsize*(1-exp(-k2*(age-t2)));
```

```
output out=new p=plength r=rlength;

run;



proc iml;

use new;

read all varrlength into res;

sum=0.0;

do i=1 to 24;

sum=sum+res[i]**2;

end;

print sum;

sigsq=sum/24;

print sigsq;
```

PROGRAM 3

In this program, we are fitting von Bertalanffy curve to fish data with Markov covariance structure.

```
/* The following macro estimates the value of rho iteratively till it's value stabi-
lizes.*/

%macro rhohat1;

proc iml;
use &ldata;
/* Giving the initial value of rho as 0.5*/
rho=J(24,1,0.5);
varnames={rho };
create nn from rho(|colname = varnames|);
append from rho;
close nn;

%ttt: data nn1;
merge nn &ldata;
```

/* In the this proc step, we are transforming the data set Y to z so that z has covariance structure $\sigma^2 I$. And also we are creating rhat so that at every iteration, we can use it to compare it with the current value of rho.*/

```
    proc iml;
use nn1;
read all var {tlength,rho } into aa;
rho1=aa[1,2];
print rho1;
/*p1=11 and p2=13*/
dist1=J(10,1,1.0);
dist1[2,1]=1.3;
g1=J(11,11,0.0);
g1[1,1]=1.0;
%do i=2 %to 11;
g1[&i,&i-1]=-(rho1**(dist1[&i-1,1]))/sqrt(1-rho1**(2*dist1[&i-1,1]));
g1[&i,&i]=1.0/sqrt(1-rho1**(2*dist1[&i-1,1]));
%end;
dist2=J(12,1,1.0);
dist2[2,1]=1.3;
g2=J(13,13,0.0);
g2[1,1]=1.0;
g2[&i,&i-1]=-(rho1**(dist2[&i-1,1]))/sqrt(1-rho1**(2*dist2[&i-1,1]));
g2[&i,&i]=1.0/sqrt(1-rho1**(2*dist2[&i-1,1]));
```

```
z1=g1*aa[1:11,1];

z2=g2*aa[12:24,1];

z=z1 || z2;

rhat=J(24,1,1);

rhat = rho1*rhat;

za=z || rhat;

varnames={ z,rhat };

create dd from za(|colname = varnames|);

append from za;

close dd;


data last;

merge dd &ldata nn;


/*We are using proc nlin to fit the von Bertalanffy curve to Z*/

   proc nlin data=last maxiter=400;

parms l1=55 k1=.276 t1=0 l2=60 k2=.23 t2=0;

model z=l1*d1*sqsize*(age1*(1-exp(-k1*(age-t1)))

+age2*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))

+age2*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))

+age3*(-rho**1.3/sqrt(1-rho**(2*1.3)))*(1-exp(-k1*((age-1.3)-t1)))

+age3*(1.0/sqrt(1-rho**(2*1.3)))*(1-exp(-k1*(age-t1)))
```

```
+age4*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))
+age4*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))
+age5*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))
+age5*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))
+age6*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))
+age6*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))
+age7*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))
+age7*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))
+age8*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))
+age8*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))
+age9*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))
+age9*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))
+age10*(-rho/sqrt(1-rho**2))*(1-exp(-k1*((age-1)-t1)))
+age10*(1.0/sqrt(1-rho**2))*(1-exp(-k1*(age-t1)))
+age11*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age11*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2))))
+l2*d2*sqsize*(age1*(1-exp(-k2*(age-t2)))
+age2*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age2*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age3*(-rho**1.3/sqrt(1-rho**(2*1.3)))*(1-exp(-k2*((age-1.3)-t2)))
+age3*(1.0/sqrt(1-rho**(2*1.3)))*(1-exp(-k2*(age-t2)))
```

```
+age4*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age4*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age5*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age5*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age6*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age6*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age7*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age7*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age8*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age8*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age9*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age9*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age10*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age10*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age11*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age11*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age12*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age12*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2)))
+age13*(-rho/sqrt(1-rho**2))*(1-exp(-k2*((age-1)-t2)))
+age13*(1.0/sqrt(1-rho**2))*(1-exp(-k2*(age-t2))));
output out=newdata p=pz r=rz;
```

```
run;

data newdd;

merge newdata nn;

run;

/* The data set newdd has the residuals after fitting the von Bertalanffy curve on
Z. In the next proc iml, we are transforming the residuals so that now they have
the Markov structure*/

    proc iml;

use newdd;

read all var {rho,rz } into aa;

rho1=aa[1,1];

run;

dist1=J(10,1,1.0);

dist1[2,1]=1.3;

g1=J(11,11,0.0);

g1[1,1]=1.0;

%do i=2 %to 11;

g1[&i,&i-1]=-(rho1**(dist1[&i-1,1]))/sqrt(1-rho1**(2*dist1[&i-1,1]));

g1[&i,&i]=1.0/sqrt(1-rho1**(2*dist1[&i-1,1]));

%end;
```

```
dist2=J(12,1,1.0);

dist2[2,1]=1.3;

g2=J(13,13,0.0);

g2[1,1]=1.0;

%do i=2 %to 13;

g2[&i,&i-1]=-(rho1**(dist2[&i-1,1]))/sqrt(1-rho1**(2*dist2[&i-1,1]));

g2[&i,&i]=1.0/sqrt(1-rho1**(2*dist2[&i-1,1]));

%end;

ry1=g1*aa[1:11,2];

ry2=g2*aa[12:24,2];

ry=ry1//ry2;

varnames={ res };

create ndat from ry(|colname = varnames|);

append from ry;

close ndat;

data ndata;

merge ndat &ldata;

/* We are using proc mixed to find the ML estimate of rho.*/

    proc mixed data=ndata method=ml;

class group;

model res= /s;
```

```
repeated/type=sp(pow)(age) subject=group r;
MAKE 'CovParms' OUT=final;
run;

/* Data set final has the new value of rho*/

proc iml;
use final;
read all var {est } into est;
rh=est[1,1];
rho=J(24,1,1);
rho=rh*rho;
varnames={rho };
create nn from rho([colname = varnames]);
append from rho;
close nn;

data the1(keep=rhat);
set last;
```

/* Here we are comparing the new value of ρ which is in rho to the previous value of *rho* which is in rhat. So if the difference between the two is stored in compos.*/

```
data tlast;
merge the1 nn;
```

```
comp=abs(rhat-rho);

proc iml;
use tlast;
read all var{comp } into comp;
compos=comp[1,1];
print compos;
%if compos ≥ 0.01 %then %goto ttt;

%mend rhohat1;

/* Main program starts here*/
options ls=78 ps=45 nodate nonumber mprint mlogic symbolgen;
title 'Nonlinear Growth Curve Analysis under Markov covariance structure';
data fisht;
input length size d1 d2 age@@;
sqsize=sqrt(size);tlength=sqsize*length;
lines;
15.40 385 1 0 1.0 26.93 28 1 0 2.0 42.23 13 1 0 3.3 44.59 83 1 0 4.3 47.63 628 1 0 5.3
49.67 1134 1 0 6.3 50.87 1761 1 0 7.3 52.30 432 1 0 8.3 54.77 93 1 0 9.3 56.43 21 1 0
10.3 55.88 8 1 0 11.3
15.40 385 0 1 1.0 28.03 36 0 1 2.0 41.18 17 0 1 3.3 46.20 135 0 1 4.3 48.23 750 0 1
5.3
```

```
50.26 1073 0 1 6.3 51.82 1459 0 1 7.3 54.27 626 0 1 8.3 56.98 199 0 1 9.3 58.93 97 0
1
10.3 59.00 44 0 1 11.3 60.91 11 0 1 12.3 61.83 6 0 1 13.3;

data fish;
set fisht;
if age=1.0 then age1=1;
else age1=0;
if age=2.0 then age2=1;
else age2=0;
if age=3.3 then age3=1;
else age3=0;
if age=4.3 then age4=1;
else age4=0;
if age=5.3 then age5=1;
else age5=0;
if age=6.3 then age6=1;
else age6=0;
if age=7.3 then age7=1;
else age7=0;
if age=8.3 then age8=1;
else age8=0;
```

```
if age=9.3 then age9=1;
else age9=0;
if age=10.3 then age10=1;
else age10=0;
if age=11.3 then age11=1;
else age11=0;
if age=12.3 then age12=1;
else age12=0;
if age=13.3 then age13=1;
else age13=0;
if d1=1 then group=1;
else group=2;

%let ldata=fish; %rhohat1;
```

PROGRAM 4

This program is for a data set with two groups, each group having n_i observations at p_i occasions, $i = 1, 2, ..., g$. Here we are fitting a von Bertalanffy curve with autoregressive covariance structure. We have used the plasma data.

```
/* Given a initial value of rho(ρ), we estimate the mean. Using this estimate of mean, we find the ML estimate of rho. With this new value of rho, we again find th ML estimate of mean. We do this iteratively till the value of rho stabilizes. All of this is done inside this macro.*/

%macro rhohat;
proc iml;
use &ldata;
/* Giving the initial value of rho as -0.5*/
rho=J(198,1,-0.5);
varnames={rho};
create nn from rho(|colname = varnames|);
append from rho;
close nn;

%ttt: data nn1;
data nn1;
merge nn &ldata;
```

/* In the this proc step, we are transforming the data set Y to z so that z has covariance structure $\sigma^2 I$. And also we are creating f11 and f22 to be used to transform the the mean of Y. We are creating rhat so that at every iteration, we can use it to compare it with the current value of rho.*/

```
proc iml;
use nn1;
read all var{y,rho} into aa;
rho1=aa[1,2];
print rho1;
g=J(6,6,0.0);
g[1,1]=1.0;
%do i=2 %to 6;
g[&i,&i]=1.0/(sqrt(1.0-rho1**2));
g[&i,&i-1]= - rho1/(sqrt(1.0-rho1**2));
%end;
f1=1.0/(sqrt(1.0-rho1**2));
f2=(- rho1)/(sqrt(1.0-rho1**2));
f11=J(198,1,1);
f11=(f1*f11);
f22=J(198,1,1);
f22=(f2*f22);
```

```
rhat=J(198,1,1);

rhat = rho1*rhat;

tm1=I(33);

tm2=tm1 @ g;

z= tm2 * aa[,1];

za=z || f11 || f22 || rhat;

varnames={ z,f11,f22,rhat };

create dd from za(|colname = varnames| );

append from za;

close dd;


data last;

merge dd &ldata;


/*We are using proc nlin to fit the von Bertalanffy curve to Z*/

proc nlin data=last maxiter=400;

parms l1=5.0 k1=0 tl1=0 l2=6.7 k2=0 tl2=0;

model z =g1*(t1*l1*(1-exp(-k1*(time-tl1)))+f11*t2*l1*(1-exp(-k1*(time-tl1)))+

f22*t2*l1*(1-exp(-k1*((time-1)-tl1))) + f11*t3*l1*(1-exp(-k1*(time-tl1)))+

f22*t3*l1*(1-exp(-k1*((time-1)-tl1))) + f11*t4*l1*(1-exp(-k1*(time-tl1)))+

f22*t4*l1*(1-exp(-k1*((time-1)-tl1))) + f11*t5*l1*(1-exp(-k1*(time-tl1)))+

f22*t5*l1*(1-exp(-k1*((time-1)-tl1))) + f11*t6*l1*(1-exp(-k1*(time-tl1)))+
```

```
f22*t6*l1*(1-exp(-k1*((time-1)-tl1)))) +
g2*(t1*l2*(1-exp(-k2*(time-tl2)))+f11*t2*l2*(1-exp(-k2*(time-tl2)))+
f22*t2*l2*(1-exp(-k2*((time-1)-tl2))) + f11*t3*l2*(1-exp(-k2*(time-tl2)))+
f22*t3*l2*(1-exp(-k2*((time-1)-tl2))) + f11*t4*l2*(1-exp(-k2*(time-tl2)))+
f22*t4*l2*(1-exp(-k2*((time-1)-tl2))) + f11*t5*l2*(1-exp(-k2*(time-tl2)))+
f22*t5*l2*(1-exp(-k2*((time-1)-tl2))) + f11*t6*l2*(1-exp(-k2*(time-tl2)))+
f22*t6*l2*(1-exp(-k2*((time-1)-tl2))));
output out=newdata p=pz r=rz;
run;

data newdd;
merge newdata nn;
run;

/* The data set newdd has the residuals after fitting the von Bertalanffy curve on
Z. In the next proc iml, we are transforming the residuals so that now they have
the autoregressive covariance structure.*/

proc iml;
use newdd;
read all var{rho,rz} into aa;
rho1=aa[1,1];
g=J(6,6,0.0);
```

```
g[1,1]=1.0;
%do i=2 %to 6;
g[&i,&i]=1.0/(sqrt(1.0-rho1**2));
g[&i,&i-1]= - rho1/(sqrt(1.0-rho1**2));
%end;
tm1=I(33);
tm2=tm1 @ g;
ry=inv(tm2) * aa[,2];
varnames={ res };
create ndat from ry(|colname = varnames|);
append from ry;
close ndat;
data ndata;
merge ndat &ldata;

/* We are using proc mixed to find the ML estimate of rho.*/
proc mixed data=ndata method=ml;
class group subj;
model res= group/s;
repeated/type=AR(1) subject=subj r;
MAKE 'CovParms' OUT=final;
run;
```

```
/* Data set final has the new value of rho*/

proc iml;

use final;

read all var{est} into est;

rh=est[1,1];

rho=J(198,1,1);

rho=rh*rho;

varnames={rho};

create nn from rho(|colname = varnames|);

append from rho;

close nn;

data thel(keep=rhat);

set last;

/* Here we are comparing the new value of ρ which is in rho to the previous value

of rho which is in rhat. So if the difference between the two is stored in compos.*/

data tlast;

merge thel nn;

comp=abs(rhat-rho);

proc iml;

use tlast;

read all var{comp } into comp;
```

```
compos=comp[1,1];

print compos;


%if compos ≥ 0.01 %then %goto ttt;

%mend rhohat;

   /* Main program starts here.*/

options linesize=70 mprint mlogic symbolgen;

title 'Fitting plasma data using proc mixed with von Bertalanffy curve and
autoregressive covariance structure';

data plasma;

infile 'plasma.data';

input subj1 y1 y2 y3 y4 y5 y6 y7 y8 subj2 y9 y10 y11 y12 y13 y14 y15 y16;

data new1(keep=subj1 y1-y8) new2(keep=subj2 y9-y16);

set plasma;

data new11;

set new1;

y=y1;time=1;t1=1;t2=0;t3=0;t4=0;t5=0;t6=0;output;

y=y2;time=2;t1=0;t2=1;t3=0;t4=0;t5=0;t6=0;output;

y=y3;time=3;t1=0;t2=0;t3=1;t4=0;t5=0;t6=0;output;

y=y4;time=4;t1=0;t2=0;t3=0;t4=1;t5=0;t6=0;output;

y=y5;time=5;t1=0;t2=0;t3=0;t4=0;t5=1;t6=0;output;
```

```
y=y6;time=6;t1=0;t2=0;t3=0;t4=0;t5=0;t6=1;output;
y=y7;time=7;t1=0;t2=0;t3=0;t4=0;t5=0;t6=0;output;
y=y8;time=8;t1=0;t2=0;t3=0;t4=0;t5=0;t6=0;output;
drop y1-y8;
data new22;
set new2;
y=y9;time=1;t1=1;t2=0;t3=0;t4=0;t5=0;t6=0;output;
y=y10;time=2;t1=0;t2=1;t3=0;t4=0;t5=0;t6=0;output;
y=y11;time=3;t1=0;t2=0;t3=1;t4=0;t5=0;t6=0;output;
y=y12;time=4;t1=0;t2=0;t3=0;t4=1;t5=0;t6=0;output;
y=y13;time=5;t1=0;t2=0;t3=0;t4=0;t5=1;t6=0;output;
y=y14;time=6;t1=0;t2=0;t3=0;t4=0;t5=0;t6=1;output;
y=y15;time=7;t1=0;t2=0;t3=0;t4=0;t5=0;t6=0;output;
y=y16;time=8;t1=0;t2=0;t3=0;t4=0;t5=0;t6=0;output;
drop y9-y16;
data b11;
set new11(in=innew1 rename=(subj1=subj))
new22(rename=(subj2=subj));
if innew1 then group=1;
else group=2;
if subj='.' then delete;
```

```
if time='7' then delete;
if time='8' then delete;


data b;
set b11;
if group=1 then g1=1;
else g1=0;
if group=2 then g2=1;
else g2=0;

%let ldata=b;
%rhohat;
```

Bibliography

[1] Anderson, T. W. (1957). Maximum likelihood estimators for a multivariate normal distribution when some observations are missing, *J. Amer. Statist. Assoc.*, 52, 200-203.

[2] Anderson, T. W. (1984). *Introduction to Multivariate Statistical Analysis.* John Wiley, New york.

[3] Anderson, A. H., Jensen, E. B. and Schou, G. (1981). Two way analysis of variance with correlated errors. *Inter. Stat. Review*, 49, 153-167.

[4] Azzalini, A. (1987). Growth curve analysis for patterned covariance matrices, *New Perspectives in Applied Statistics* , (eds. M. L. Puri et al.), 63-73.

[5] Bhargava, R. P. (1975). Some one-sample hypothesis testing problems when there is a monotone sample from a multivariate normal population, *Ann. Inst. Statist. Math.*, 27, 327-339.

[6] 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.

[7] Chi, E. M. and Reinsel, G. C. (1989). Models for longitudinal data with random effects and AR(1) errors, *J. Amer. Statist. Assoc.*, 84, 452-459.

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

[9] Dahiya, R. C. and Korwar, R. M. (1980). Maximum likelihood estimators for a bivariate normal distribution with missing data, *Ann. Statist.*, 8, 687-692.

[10] Dark, T. A. (1975). Age and growth of Pacific Hake, Merluccius Productus, *Fishery Bulletin*, 73, 336-355.

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

[12] Frey, C. M. and Muller, K. E. (1992). Analysis methods for nonlinear models with compound symmetric covariance, *Commun. Statist. -Theory Meth.*, 21, 1163-1182.

[13] Fujikoshi, Y., Kanda, T. and Tanimura, N. (1990). The growth curve model with an autoregressive covariance structure, *Ann. Inst. Statist. Math.*, 42, 533-542.

[14] Glasbey, C. A. (1979). Correlated residuals in non-linear regression applied to growth data, *Appl. Statist.*, 28, 251-259.

[15] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115, 513-585.

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

[17] Hirst, K., Boyle, D. W., Zerbe, G. O. and Wilkening, R. B. (1991). On nonlinear random effects models for repeated measurements, *Commun. Statist. - Simula.*, 20, 463-478.

[18] Khatri, C. G. (1966). A note on a MANOVA model applied to problems in growth curve, *Ann. Inst. Statist. Math.*, 18, 75-86.

[19] Kimura, D. K. (1980). Likelihood methods for the von Bertalanffy growth curves, *U. S. Fishery Bulletin*, 77, 765-776.

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

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

[22] Lakkis, H. D. and Jones, C. M. (1992). Comparing von Bertalanffy growth curves with SAS using the likelihood methods developed by Kimura, *Preprint*.

[23] Lee, J. C. (1988). Prediction and estimation of growth curves with special covariance structures, *J. Amer. Statist. Assoc.*, 83, 432-440.

[24] Lee, J. C. (1991). Tests and model selection for the general growth curve model, *Biometrics*, 47, 147-159.

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

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

[27] Lindstrom, M. J. and Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data, *Biometrics*, 46, 673-687.

[28] Naik, D. N. (1990). Prediction intervals for growth curves, *J. App. Stat.*, 17, 245-254.

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

[30] Pantula, S. G. and Pollock, K. H. (1985). Nested analysis of variance with autocorrelated errors, *Biometrics*, 41, 909-920.

[31] Palmer, M. J., Phillips, B. F. and Smith, G. T. (1991). Application of nonlinear models with random coefficients to growth data, *Biometrics*, 47, 623-635.

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

[33] 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.

[34] 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.

[35] Rao, C. R. (1966). Covariance adjustment and related problems in multivariate analysis, *Multivariate Analysis*, 1, 87-103.

[36] Rao, C. R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals, *Proc. Fifth Berkeley Symp. on Math. Statist. Prob.*, Vol. 1, 355-372, Univ. of California Press, Berkeley.

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

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

[39] Rao, C. R. (1977). Prediction of future observations with special reference to linear models, *Multivariate Analysis*, 4, 193-208.

[40] 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.

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

[42] Ratkowsky, D. A. (1983). *Nonlinear Regression Modeling.* Marcel Dekker, New York.

[43] Reinsel, G. (1982). Multivariate repeated measurement or growth curve models with multivariate random effects covariance structure. *J. Amer. Statist. Assoc.*, 77, 190-195.

[44] Reinsel, G. (1984). Estimation and prediction in a multivariate random effects generalized linear model, *J. Amer. Statist. Assoc.*, 79, 406-414.

[45] Richards, F. J. (1959). A flexible growth function for emperical use, *J. Exp. Botany*, 10, 290-300.

[46] Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression.* John Wiley, New York.

[47] Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook.* American Sciences Press, Columbus, Ohio.

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

[49] Verhulst, P. F. (1845). Recherches mathematiques sur la loi d'accroissement de la population, *Nouvelles Memoires Academic Royale, Science et Lettres, Bruxelles, Series 2*, 18, 1-38.

[50] Viana, M. A. G. (1994). Combined maximum likelihood estimates for the equicorrelation coefficient, *Biometrics*, 50, 813-820.

[51] von Bertalanffy, L. (1957). Quantitative laws in metabolism and growth, *Quart. Rev. Biol.*, 32, 217-231.

[52] von Rosen D. (1991). The growth curve model: A review, *comm. statist. theory. meth.*, 20, 2791-2822.

[53] Vonesh, E. F. and Carter, R. L. (1992). Mixed-Effects Nonlinear Regression for Unbalanced Repeated Measures, *Biometrics*, 48, 1-17.

[54] Zerbe, G. O. (1979). Randomization analysis of the completely randomized design extended to growth and response curves, *J. Amer. Statist. Assoc.*, 74, 215-221.

Biographical Statement

Place of Birth

Vijayawada, India

Date of Birth

January 04, 1967

Education

B.Sc. in Statistics, 1988,
Banaras Hindu University, Varanasi, India

M.Sc. in Statistics, 1990,
Indian Institute of Technology, Kanpur, India

Employment

Department of Mathematics and Statistics, Old Dominion University,
1991-1995, Teaching Assistant.