

This research was supported by the National Institutes of Health,  
Institute of General Medical Sciences, Grant No. GM-12868-04.

ANALYSIS OF GROWTH AND DOSE RESPONSE CURVES

James E. Grizzle  
Department of Biostatistics  
University of North Carolina

David M. Allen  
Department of Statistics  
University of Kentucky

Institute of Statistics Mimeo Series No. 576

March 1968

## ANALYSIS OF GROWTH AND DOSE RESPONSE CURVES

James E. Grizzle  
Department of Biostatistics  
University of North Carolina

David M. Allen  
Department of Statistics  
University of Kentucky

### Introduction

The experimental material and techniques that must be used in some fields make it advantageous to follow individuals for a period of time, perhaps over a sequence of doses, thus negating the assumption of independence among points of the response curve. Models appropriate for the analysis of such data are often discussed under the title of growth curve experiments, although the models also have application to other studies.

The objectives of this paper are to study the growth curve model from the point of view of the analysis of covariance begun by Rao [9], [10], [11]. This leads to a relatively simple derivation of the results produced by Khatri [6] and to easily implemented methods of estimating parameters and performing tests.

### The Model

Suppose there are  $N$  individuals who have been randomly assigned to one of the  $r$  cells of some design, and that the same characteristic has been measured at  $p$  times or under  $p$  different conditions for each individual. Further suppose that each individual's responses to treatment or at each time can be described by a linear model of the form

$$E(\underline{X}_{-ij}) = B \underline{\xi}_j \quad (1)$$

$p \times 1 \qquad p \times q \quad q \times 1$

where  $X_{ij}$  is the vector of observations made on the  $i$ th individual in the  $j$ th group and  $B$  is a known matrix of rank  $q$ . We shall call (1) the within individual model. The expected values are allowed to vary from group to group, but they must belong to the same family. That is, if the expected value is a  $q$ th degree polynomial in one group, it must be a  $q$ th degree polynomial in other groups, but the actual values of the parameters are allowed to be different.

The complete model expressing both the within individual and across individual designs can be expressed in matrix notation as

$$X \quad = \quad B \quad \xi \quad A \quad + \quad E \quad (2)$$

$(p \times N) \quad (p \times q) \quad (q \times r) \quad (r \times N) \quad (p \times N)$

where the columns of  $E$  are independently distributed as a  $p$ -variate normal with common covariance matrix  $\Sigma$  and mean vector  $\underline{0}$ ;  $B$  is the design matrix within individuals;  $A$  is the design matrix across individuals; and  $\xi$  is a matrix of unknown parameters. The matrices  $A$  and  $B$  can be assumed to be of full rank without loss of generality. This model can be extended in an obvious way to the case of  $p$  measurements on each of several characteristics, and Allen [1] has extended it to the case in which the response of each individual is described by a function which is nonlinear in the parameters. We shall not be concerned with these extensions.

This model has many similarities to the mixed model in which the individuals are considered to be a random sample from some population. The mixed model approach, discussed in [4], has limitations due to the assumption of uniform correlation among elements of the response vector. In the development to follow we require no special assumptions about the

pattern of covariance in  $\Sigma$ , but the development permits exploitation of the pattern when one exists. The data in Table 1 are presented as an example of the type of analysis problem we wish to discuss.

The data can be related to the model as follows: the elements of the vector  $X_{1j}$  are the responses over time of the  $i$ th dog,  $i = 1, \dots, N_j$ , to the  $j$ th treatment,  $j = 1, 2, 3, 4$ . Assume that a third degree polynomial is an adequate description of response of the  $i$ th dog over the period of observation. Then the orthogonal coefficients make up the matrix  $B$ , i.e.,

$$B = \begin{matrix} & b_0 & b_1 & b_2 & b_3 \\ \begin{matrix} 7 \times 4 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{bmatrix} 1 & -3 & 5 & -1 \\ 1 & -2 & 0 & 1 \\ 1 & -1 & -3 & 1 \\ 1 & 0 & -4 & 0 \\ 1 & 1 & -3 & -1 \\ 1 & 2 & 0 & -1 \\ 1 & 3 & 5 & 1 \end{bmatrix} \end{matrix}$$

The powers  $t$ ,  $t^2$ ,  $t^3$  ( $t$  = time after occlusion) could have been used.

However, since the observations were made at equally spaced intervals we shall make use of the orthogonal coefficients. According to the model every dog in the  $j$ th group has the same expected response, i.e.,

$$E(X_{1j}) = \xi_{0j} b_0 + \xi_{1j} b_1 + \xi_{2j} b_2 + \xi_{3j} b_3$$

where  $b_k$ ,  $k = 0, 1, 2, 3$ , are the elements of the  $k$ th columns of  $B$  as shown above. Different groups of dogs may have different expected values but their response function is assumed to be a member of the same family.

Table 1

Group 1. Coronary sinus potassium mEq./l. in control dogs during the initial 13 minutes after coronary occlusion

Dog	Time - minutes after occlusion												
	1	3	5	7	9	11	13						
1	4.0	4.0	4.1	3.6	3.6	3.8	3.1						
2	4.2	4.3	4.7	4.7	4.8	5.0	5.2*						
3	4.3	4.2	4.3	4.3	4.5	5.8	5.4*						
4	4.2	4.4	4.6	4.9	5.3	5.6	4.9						
5	4.6	4.4	5.3	5.6	5.9	5.9	5.3						
6	3.1	3.6	4.9	5.2	5.3	4.2	4.1						
7	3.7	3.9	3.9	4.8	5.2	5.4	4.2						
8	4.3	4.2	4.4	5.2	5.6	5.4	4.7						
9	4.6	4.6	4.4	4.6	5.4	5.9	5.6						
Average	4.11	4.18	4.51	4.77	5.07	5.22	4.72						

\* Animals developing ventricular fibrillation

Table 1--continued

Group 2. Coronary sinus potassium mEq./l. in ten dogs having undergone extrinsic cardiac denervation three weeks prior to coronary occlusion

Dog	Time - minutes after occlusion									
	1	3	5	7	9	11	13			
10	3.4	3.4	3.5	3.1	3.1	3.7	3.3			
11	3.0	3.2	3.0	3.0	3.1	3.2	3.1			
12	3.0	3.1	3.2	3.0	3.3	3.0	3.0			
13	3.1	3.2	3.2	3.2	3.3	3.1	3.1			
14	3.8	3.9	4.0	3.9	3.5	3.5	3.4			
15	3.0	3.6	3.2	3.1	3.0	3.0	3.0			
16	3.3	3.3*	3.3	3.4	3.6	3.1	3.1			
17	4.2	4.0	4.2	4.1	4.2	4.0	4.0			
18	4.1	4.2	4.3	4.3	4.2	4.0	4.2			
19	4.5	4.4	4.3	4.5	4.3	4.4	4.4			
Average	3.54	3.63	3.62	3.56	3.56	3.50	3.46			

\* Denotes brief period of ectopic ventricular activity

Table 1--continued

Group 3. Coronary sinus potassium mEq./l. in ten dogs subjected to extrinsic cardiac denervation immediately prior to coronary occlusion

Dog	Time - <u>minutes</u> after occlusion									
	1	3	5	7	9	11	13			
20	3.2	3.3	3.8	3.8	4.4	4.2	4.0			
21	3.3	3.4	3.4	3.7	3.7	3.6	3.7			
22	3.1	3.3	3.2	3.1	3.2	3.1	3.1			
23	3.6	3.4	3.5	4.6	4.9	5.2	4.4			
24	4.5	4.5	5.4	5.7	4.9	4.0	4.0			
25	3.7	4.0	4.4	4.2	4.6	4.8	5.4*			
26	3.5	3.9	5.8	5.4	4.9	5.3	5.6			
27	3.9	4.0	4.1	5.0	5.4	4.4	3.9			
Average	3.60	3.73	4.20	4.44	4.50	4.53	4.26			

\* Denotes animals developing ventricular fibrillation

Table 1--continued

Group 4. Coronary sinus potassium mEq./l. in ten dogs having undergone bilateral thoracic sympathectomy and stellatectomy three weeks prior to coronary occlusion

Dog	Time - <u>minutes</u> after occlusion									
	1	3	5	7	9	11	13			
28	3.1	3.5	3.5	3.2	3.0	3.0	3.2			
29	3.3	3.2	3.6	3.7	3.7	4.2	4.4*			
30	3.5	3.9	4.7	4.3	3.9	3.4	3.5			
31	3.4	3.4	3.5	3.3	3.4	3.2	3.4			
32	3.7	3.8	4.2	4.3	3.6	3.8	3.7			
33	4.0	4.6	4.8	4.9	5.4	5.6	4.8*			
34	4.2	3.9	4.5	4.7	3.9	3.8	3.7			
35	4.1	4.1	3.7	4.0	4.1	4.6	4.7			
36	3.5	3.6	3.6	4.2	4.8	4.9	5.0			
Average	3.64	3.78	4.01	4.07	3.98	4.06	4.04			

\* Denotes animals developing ventricular fibrillation



This does not preclude some of the  $\xi$ 's being zero. Now we can define the matrix

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \end{bmatrix}$$

to show how the expected value changes from group to group. Expansion of  $E(X) = B\xi A$  shows that this model gives the predicted value of the response of each dog at time  $t$  after coronary occlusion under the assumption that the expected functional form of the response in each treatment group is the same.

In this example  $A$  is the group indicator matrix of a simple design. However, this is not required.  $A$  could be a matrix of continuous variables as in multiple regression, or it could be made up of design variables and covariables. Thus  $A$  is allowed to be quite general. In contrast  $B$  is somewhat restricted. The present state of development of the theory requires that  $B$  be the same for each individual, i.e., the pattern of observations within individual be the same.

#### Review of the Literature

In the growth curve model as first formulated by Potthoff and Roy [8], an arbitrary matrix of weights  $G^{-1}$  was introduced into the analysis. They noted that their estimate of  $\xi$  was always unbiased, but that the variance of  $\hat{\xi}$  increased as  $G^{-1}$  departed from  $\Sigma^{-1}$ . They did not, however, develop the requisite theory for allowing  $G^{-1} = S^{-1}$ , where  $S$  is proportional to the estimate of  $\Sigma$  calculated from the data being used to estimate  $\xi$ . Khatri [6], using the method of maximum likelihood, obtained the estimate

of  $\xi$  in which  $S^{-1}$  is used as the weights, and he showed that the likelihood ratio, trace, or largest root test criterion could be used to test the hypothesis  $C \xi V = 0$ , where  $C$  is of rank  $c \leq q$  and  $V$  is of rank  $v \leq r$ . However, he did not give the details about any of the test criteria except the largest root.

Rao [9] (appendices 1 and 2) commented on Potthoff and Roy's approach. He pointed out that their method did not use all the information available in the sample for estimating  $\xi$  unless  $G^{-1}$  happened to equal  $\Sigma^{-1}$ . He showed further that the additional information could be utilized by incorporating into the model  $(p-q)$  covariables which are a basis of the set of linear functions of the columns of  $X$  given by  $[I - B(B'B)^{-1}B']X$ , i.e., the vectors which span the within individual error space. Rao continued the discussion in [9] in which he pointed out in an example that weighting did not necessarily produce shorter confidence intervals. Rao's approach [9], [10], [11], turns out to give results algebraically identical to those of Khatri.

Gleser and Olkin [5] studied the properties of a  $k$ -sample regression model with covariates. This model can be expressed as follows: if the  $i$ th vector observation is from the  $j$ th sample, the  $(j,i)$  element of  $A$  is one, and all other elements of the  $i$ th column of  $A$  are zero.

They obtained a form for the unconditional probability density function of the estimate of  $\xi$ . However, this density function is expressed as an integral, a representation which restricts its usefulness. They also derived tests of hypotheses and showed that their statistics were functions of the maximal invariant statistics. Their tests are identical to those we present when the models coincide.

Rao [11] also studied the properties of  $\hat{\xi}$  in a one-sample regression model with the specially chosen covariates. His discussion of the choice of covariates in the one-sample case is applicable in general.

In a recent paper Watson [13] discussed some time series models. This work turns out to have elements of similarity to the problem we are considering except our problem is somewhat simpler, since, under our model, we have enough data to obtain an independent estimate of  $\Sigma$ .

Williams [14] derived the variance of regression estimators weighted inversely as  $S^{-1}$  using a different technique from Rao [11], and investigated the effect of using the estimated weights on the variance of the estimator.

We note that the patterned dispersion matrices discussed by Rao [11] may also be appropriate for the Gleser and Olkin model. The effect of assuming one of these patterns is to reduce the model to another of the same type but with fewer covariates.

#### Derivation of Maximum Likelihood Estimate of $\xi$

Let  $B_1$  and  $B_2$  be matrices of full rank such that

$$\begin{matrix} (p \times q) & & p \times (p-q) \end{matrix}$$

$$B_1' B_1 = I \quad (q \times q)$$

and

$$B_2' B_2 = 0 \quad (p-q) \times q$$

It should be noted that possible choices for  $B_1$  and  $B_2$  are

$$B_1 = B(B'B)^{-1}$$

and  $B_2$  = a column basis of  $[I - B(B'B)^{-1}B']$ , or when the elements of  $B$  are the intercept and orthogonal polynomials up to degree  $q - 1 < p$ , then  $B_2$  can be chosen as the coefficients of order  $q, q + 1, \dots, p - 1$ .

Let

$$\begin{aligned} Y_1 &= B_1'X, \\ (q \times N) \end{aligned}$$

and

$$\begin{aligned} Y_2 &= B_2'X; \\ (p-q) \times N \end{aligned}$$

then

$$\begin{aligned} E(Y_1) &= B_1'B\xi A \\ &= \xi A. \end{aligned}$$

Hence,  $Y_1$  is composed of the vectors spanning the estimation space if each individual's expected response is given by equation (1).

The expected value of  $Y_2$  is

$$\begin{aligned} E(Y_2) &= B_2'B\xi A \\ &= 0, \end{aligned}$$

which shows that  $Y_2$  is composed of the vectors in the within individual error space.

The model (2) can be written in terms of  $Y_1$  and  $Y_2$  as

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \xi \\ \beta \end{pmatrix} A + E_1$$

where  $\beta = 0$  if our original model is adequate, and the columns of  $E_1$  are independently distributed as  $p$ -variate normal with covariance matrix

$$V = \begin{bmatrix} B_1' \Sigma B_1 & B_1' \Sigma B_2 \\ B_2' \Sigma B_1 & B_2' \Sigma B_2 \end{bmatrix},$$

and mean vector  $\underline{0}$ .

Assuming that our within individual model is correct ( $E(Y_2) = 0$ ), the joint density of  $Y_1$  and  $Y_2$  can be factored into the product of the marginal density of  $Y_2$ ,  $\psi_2(Y_2; V_2)$  and the conditional density of  $Y_1$  given  $Y_2$ ,  $\psi_1(Y_1 | Y_2; V_1, \eta)$ . Applying the technique of Anderson [2, p.28] we have

$$\psi_2(Y_2; V_2) \psi_1(Y_1 | Y_2; V_1, \eta) = \frac{|V_2^{-1}|^{\frac{N}{2}}}{(2\pi)^{\frac{N(p-q)}{2}}} \exp\left\{-\frac{1}{2} \text{tr}(V_2^{-1} Y_2 Y_2')\right\} \times$$

$$\frac{|V_1^{-1}|^{\frac{N}{2}}}{(2\pi)^{\frac{Nq}{2}}} \exp\left\{-\frac{1}{2} \text{tr}[V_1^{-1} (Y_1 - \xi A - \eta Y_2)(Y_1 - \xi A - \eta Y_2)']\right\} \quad (3)$$

where

$$V_2 \begin{matrix} (p-q) \times (p-q) \end{matrix} = B_2' \Sigma B_2,$$

$$V_1 \begin{matrix} (q \times q) \end{matrix} = B_1' \Sigma B_1 - B_1' \Sigma B_2 (B_2' \Sigma B_2)^{-1} B_2' \Sigma B_1$$

and

$$\eta \begin{matrix} q \times (p-q) \end{matrix} = B_1' \Sigma B_2 (B_2' \Sigma B_2)^{-1}.$$

By the use of Khatri's [6] lemma 1 we have

$$\begin{aligned} V_1 &= B_1' (\Sigma - \Sigma B_2 (B_2' \Sigma B_2)^{-1} B_2' \Sigma) B_1 \\ &= B_1' (B (B' \Sigma^{-1} B)^{-1} B') B_1 \\ &= (B' \Sigma^{-1} B)^{-1}. \end{aligned}$$

Thus  $V_1$  does not depend upon the particular choices of  $B_1$  and  $B_2$ . It is evident from (3) that the maximum likelihood estimator of  $\xi$  and the likelihood ratio statistics for testing hypotheses concerning  $\xi$  can be obtained using the conditional density of  $Y_1$ , given  $Y_2$ , since the marginal density of  $Y_2$  assumes the role of a constant. Notice also that the conditional distribution of  $Y_1$ , given  $Y_2$ , is a linear multivariate model:

$$Y_1 = \begin{bmatrix} \xi & \eta \end{bmatrix} \begin{bmatrix} A \\ Y_2 \end{bmatrix} + E_2 \quad (4a)$$

$(q \times N) \quad q \times (r+p-q) \quad (r+p-q) \times N \quad q \times N$

or for simplicity

$$Y_1 = \theta \quad D \quad + E_2, \quad (4b)$$

$q \times (r+p-q) \quad (r+p-q) \times N$

where the columns of  $E_2$  are independently distributed as  $q$ -variates normal with covariance matrix  $V_1$  and mean vector  $\underline{0}$ .

The matrix  $D = \begin{bmatrix} A \\ Y_2 \end{bmatrix}$  will have rank  $r+p-q$  with probability one.

Thus  $(DD')^{-1}$  exists with probability one, and the maximum likelihood estimator of  $\xi$  is given by the least squares estimate of  $\xi$  using model (4b), i.e.,

$$[\hat{\xi}, \hat{\eta}] = \hat{\theta} = Y_1 D' (DD')^{-1}, \quad (5)$$

and  $\hat{\xi}$  is the first  $r$  columns of  $\hat{\theta}$ .

The matrix  $DD'$  is

$$DD' = \begin{bmatrix} AA' & AX'B_2 \\ B_2'XA & B_2'XX'B_2 \end{bmatrix}.$$

If we invert  $DD'$  algebraically, write out the estimate of  $\xi$  and apply Khatri's [6] lemma 1 we find that

$$\hat{\xi} = (B'S^{-1}B)^{-1} B'S^{-1}X A'(AA')^{-1} \quad (6)$$

where

$$S = X[I - A'(AA')^{-1}A]X'. \quad (7)$$

Hence, the estimates obtained by the analysis of covariance, by maximum likelihood and by weighting inversely as the estimated variance are shown to be identical. A slight extension of the result given in equation 50 of Rao's paper [12] yields the unconditional covariance matrix the elements of  $\hat{\xi}$  as

$$\text{Var}(\hat{\xi}) = (AA')^{-1} \otimes \frac{N-r-1}{N-r-(p-q)-1} V_1, \quad (8)$$

where  $\otimes$  is the Kronecker product, and  $\text{Var}(\hat{\xi})$  denotes the covariance matrix of the elements of  $\hat{\xi}$  taken in a columnwise manner.

It is easily shown that  $E(\hat{\xi}) = \xi$  and that the residual error sum of products matrix from fitting the model (4a)

is

$$\begin{aligned} S_{Y_1} &= B_1' S B_1 - B_1' S B_2 (B_2' S B_2)^{-1} B_2' S B_1^{-1} \\ &= (B' S^{-1} B)^{-1}. \end{aligned} \quad (9)$$

That

$$E(S_{Y_1}) = V_1 [N-r-(p-q)] \quad (10)$$

follows from the fact that  $S_{Y_1}$  is distributed as  $W_q [N-r-(p-q), V_1, 0]$

where  $W_{k_1} [k_2, \Sigma, \mu]$  denotes a  $k_1$  dimensional Wishart distribution with  $k_2$  degrees of freedom, covariance matrix  $\Sigma$ , and noncentrality matrix  $\mu$ .

From equation (8) and (10) it follows that an unbiased estimate of the variance of  $\hat{\xi}$  is

$$\hat{\text{Var}}(\hat{\xi}) = (AA')^{-1} \otimes \hat{V}_1 (N-r-1) / [N-r-(p-q)-1], \quad (11)$$

where  $\hat{V}_1 = \frac{1}{N-r-(p-q)} S_{Y_1}$ .

Even though the first two moments of  $\hat{\xi}$  are known, its complete distribution is not available in a simple form. The conditional distribution of  $\hat{\xi}$  given  $Y_2$  is, by linear multivariate theory, normally distributed with mean  $\xi$  and covariance matrix  $R_1 \otimes V_1$  where  $R_1$  is the leading  $r \times r$  minor of  $(DD')^{-1}$ .

#### Test of the Fit of the Model

The test of the fit of the model is, strictly speaking, a test of the fit of the within individual model. We have partitioned the observed vector into two parts:  $Y_1$  which spans the within individual estimation space, and  $Y_2$  which spans the within individual error space. Therefore if each individual's response over time or over doses is described



adequately by equation (1),  $E(Y_2) = 0$ , and a test of the fit of the model is given by testing

$$H_0: E(Y_2) = 0.$$

This test is produced easily from conventional multivariate analysis of variance in which

$$E(X) = \begin{matrix} \xi^* & A \\ p \times r & r \times n \end{matrix},$$

where  $A$  is the design matrix for individuals shown in equation (2), and  $\xi^*$  are the parameters associated with the analysis when the within individual aspect described by  $B$  is omitted. We assume this model and test

$$B_2' \xi^* I = 0.$$

The hypothesis matrix is

$$P_1 = \begin{matrix} = B_2' X A' (A A')^{-1} A X' B_2 \\ (p-q) \times (p-q) \end{matrix}$$

and the error matrix is

$$Q_1 = \begin{matrix} = B_2' X [I - A' (A A')^{-1} A] X' B_2 \\ (p-q) \times (p-q) \end{matrix}$$

These matrix forms will be produced automatically by any multivariate analysis of variance program which will test  $H_0: C \xi^* V = 0$  by setting  $C = B_2$  and  $V = I$ .

It is easy to show that  $P_1$  is distributed as  $W_{(p-q)} [r, V_2, B_2' \xi^* A A' \xi^* B_2]$  and  $Q_1$  is distributed as  $W_{p-q} [N-r, V_2, 0]$  independently of  $P_1$ . All of the

criteria in common use are functions of the characteristic roots  $\lambda_1$  of  $P_1 Q_1^{-1}$ .

The likelihood ratio test criteria is  $U = -m \log_e \prod_{i=1}^{p-q} \frac{1}{(1+\lambda_i)}$

where  $m = N-r - \frac{1}{2}(p-q-r+1)$ . This test statistic has asymptotically a  $\chi^2$ -distribution with  $r \times (p-q)$  degrees of freedom, or the exact percentage points can be found in [12].

The largest root test is given by

$$\theta = \frac{1}{1 + \lambda_{\max}}$$

and has degrees of freedom parameters

$$s = \min(r, p-q),$$

$$m = (|r-(p-q)| - 1)/2,$$

$$n = (N-r-(p-q)-1)/2,$$

using the notation given in Morrison [7].

A convenient form for computing the non-zero roots of  $P_1 Q_1^{-1}$  which avoids specification of  $B_1$  and  $B_2$  is to use the identity

$$\text{ch. } P_1 Q_1^{-1} = \text{ch. } XA'(AA')^{-1} AX' S^{-1} [S - B(B'S^{-1}B)^{-1}B'] S^{-1},$$

where ch. denotes non-zero characteristic roots. A more suggestive form for ease of interpretation is

$$\text{ch. } P_1 Q_1^{-1} = \text{ch. } (\hat{X}_A S^{-1} \hat{X}'_A - \hat{X}_{A,B} S^{-1} \hat{X}'_{A,B})$$

where  $\hat{X}_A$  denotes the predicted value of  $X$  when using  $A$  only as the basis of prediction and  $\hat{X}_{A,B}$  denotes that both  $A$  and  $B$  are to be used in prediction.

Test of Linear Hypotheses About  $\xi$  and  $\eta$

Testing the hypotheses

$$H_0: \begin{matrix} C & \xi & V = 0, \\ c \times q & q \times r & r \times v \end{matrix}$$

where rank  $C = c \leq q$  and rank  $V = v \leq r$ , for model (2) is accomplished by testing the hypothesis

$$H_0: C \begin{bmatrix} \xi & \eta \\ \eta & \xi \end{bmatrix} \begin{bmatrix} V \\ 0 \end{bmatrix} = C \xi V = 0$$

for model (4a). The error matrix produced is the correct form, namely,

$$Q_2 = CY_1(I - D'(DD')^{-1}D)Y_1'C;$$

and its distribution given  $D$  is  $W_c[N-r-p+q, CV_1C', 0]$ .

Since the distribution  $Q_2$  depends only on the rank of  $D$ , the unconditional distribution of  $Q_2$  is the same as its conditional distribution. The hypothesis matrix is

$$P_2 = (C \hat{\xi} V) (V'R_1V)^{-1} (C \hat{\xi} V)',$$

where  $R_1$  is the leading  $r \times r$  minor of  $(DD')^{-1}$ . The conditional distribution of  $P_2$  is  $W_c[v, CV_1C', (C \hat{\xi} V) (V'R_1V)^{-1} (C \hat{\xi} V)']$ . If the null hypothesis is true, the distribution of  $P_2$  does not depend on  $D$ , and its conditional and unconditional distributions are the same. If the null hypothesis is false the noncentrality parameter depends upon  $Y_2$ .

There are at least two approaches to estimating  $\xi$  and to testing  $C \xi V = 0$ . If we choose to estimate  $\xi$  by equation (6) we can compute

$$P_2 = (C \hat{\xi} V) (V'R_1V)^{-1} (C \hat{\xi} V)',$$

where

$$R_1 = (AA')^{-1} + (AA')^{-1} AX' [S^{-1} - S^{-1} B(B'S^{-1}B)^{-1} B'S^{-1}] XA' (AA')^{-1},$$

and

$$Q_2 = C(B'S^{-1}B)^{-1} C'.$$

This method avoids finding  $B_1$  and  $B_2$ , but it has the disadvantage of requiring a special computer program. (We assume that the majority of analysis of this type will be performed with the assistance of an electronic computer.)

An alternative method which requires less special programming is to use model given by (4b) which leads to the estimate of  $\xi$  given in equation (5). Then any general multivariate linear hypothesis program can be used to analyze the data. When we want to test the hypothesis  $C \xi V = 0$  we must remember that  $\eta$  has been estimated also. The matrix of estimated parameters is  $(\hat{\xi} \hat{\eta})$ , and we must eliminate  $\hat{\eta}$  from the test. This is done easily by augmenting  $V$  by the null matrix  $0$  to insure that  $\hat{\eta}$  is eliminated. Thus to test  $C \xi V$  we actually test  $C(\xi \eta) \begin{pmatrix} V \\ 0 \end{pmatrix}$  in the model (4a). Then the correct  $P_2$  and  $Q_2$  are produced automatically. The only difficulty in this approach is finding  $B_2$ . When equation (1) can be expressed in terms of orthogonal polynomials,  $B_2$  is easily found as the coefficients of the polynomials of degree  $q, q+1, \dots, p-1$ , which avoids finding  $B_2 =$  a column basis of  $[I - B(B'B)^{-1}B']$ .

In either case once  $B_2$  has been determined the analysis could proceed as follows.  $B_1' = (B'B)^{-1}B'$  is formed, and as the data are read into the computer the linear transformations

$$Y_1 = B_1' X \text{ and } Y_2 = B_2' X$$

are made, and we proceed in a straightforward way using model (4b) for the remainder of the analysis.

Another advantage is that in case we do not want to use all vectors in the error space as covariables, which is equivalent to changing the weights used to something other than  $S^{-1}$ , those selected can be easily included in this method of calculation. In contrast deletion of some of the variables could not be accomplished so easily by the first method described.

Both the likelihood ratio and largest root test criteria are functions of the characteristic roots of  $P_2 Q_2^{-1}$ . The likelihood ratio test criterion is  $U = -m \log_e \prod_{i=1}^c \left( \frac{1}{1+\lambda_i} \right)$ , where  $m = N-r-(p-q) - \frac{1}{2}(c-v+1)$ , and has a central  $\chi^2$ -distribution with  $cv$  degrees of freedom if the null hypothesis is true.

The largest root test criterion is  $\theta = \frac{1}{1+\lambda_{\max}}$  has degrees of freedom parameters

$$s = \min (v, c)$$

$$m = (|v-c| - 1)/2$$

$$n = (N-r-(p-q)-c-1)/2$$

If all the vectors in the error space are not used as covariables the  $(p-q)$  shown in parentheses should be reduced to the number actually used. The circumstances in which it is desirable to use a smaller number are discussed in the next section.

We may also test the contributions made by the inclusion of any set of linear combinations of the columns of  $\eta$  in the model. Testing the hypothesis

$$H_0: \eta L = 0$$

where  $L$  is a known matrix of order  $(p-q) \times \ell$  is done by testing the hypothesis

$$H_0: [\xi \ \eta] \begin{pmatrix} 0 \\ L \end{pmatrix} = \eta L = 0$$

for model (4a). The error matrix for this test is

$$Q_2 = Y_1 (I - D'(DD')^{-1}D) Y_1'$$

and its conditional and unconditional distribution is  $W_q [N-r-p+q, V_1, 0]$ .

The hypothesis matrix is

$$P_3 = (\hat{\eta}L) (L'R_2L)^{-1} (\hat{\eta}L)'$$

where  $R_2$  is the lower right  $(p-q) \times (p-q)$  minor of  $(DD')^{-1}$ . The conditional distribution of  $P_3$  given  $D$  is  $W_q [\ell, V_1, (\eta L)'(L'R_2L)^{-1}(\eta L)]$ .

If the null hypothesis is true, the distribution of  $P_3$  does not depend on  $D$ , and its conditional and unconditional distributions are the same.

It can be shown that the noncentrality parameter depends upon  $Y_2$  and the particular choice of  $B_2$  unless  $L$  is square. If  $L$  is square the noncentrality parameter depends upon  $Y_2$  but not on the particular choice of  $B_2$ . When  $L$  is square the hypothesis being tested is equivalent to

$$H_0: \eta = 0,$$

or all of the covariances between the elements of  $Y_1$  and  $Y_2$  are zero.

The likelihood ratio test has the usual form in which

$m = N - r - (p - q) - \frac{1}{2}(q - \ell + 1)$  and  $U$  has  $q\ell$  degrees of freedom. The degrees of freedom for the largest root test criterion are

$$s = \min(\ell, q)$$

$$m = (|\ell - q| - 1)/2$$

$$n = (N - r - (p - q) - q - 1)/2.$$

It should be noted that when  $\eta = 0$ , the implication is that covariates are not needed, that is we might as well have worked with the simpler model. Rao [11] showed that including covariates did not necessarily lead to estimates that have smaller variance than when they are omitted.

#### Choice of Covariates

Rao examined the effect of assuming patterned dispersion matrices which arise in various mixed models. These patterned dispersion matrices may also be appropriate when the model is of the special case of (1) considered by Gleser and Olkin [5].

Suppose that  $\Sigma = B\Gamma B' + H\Theta H' + I\sigma^2$  where  $\Gamma$  and  $\Theta$  are unknown dispersion matrices,  $\sigma^2$  is an unknown variance and  $H$  is a known matrix of dimension  $(p \times q)$  of full rank such that  $H'B = 0$ . When we make our transformation we let  $B_1 = B(B'B)^{-1}$ . We then find that the covariance matrix between corresponding columns of  $Y_1$  and  $Y_2$  is

$$\begin{aligned} \text{Cov}(Y_{1i} Y_{2i}) &= (B'B)^{-1} B' (B\Gamma B' + H\Theta H' + I\sigma^2) B_2 \\ &= 0. \end{aligned}$$

Here  $Y_2$  provides no information about  $Y_1$  and should be disregarded.

Inferences about  $\xi$  should be made using conventional multivariate techniques applied to  $Y_1$ , which is equivalent to unweighted least squares, or choosing  $G = I$  in Potthoff's and Roy's development.

If  $\theta$  or  $H = 0$  this model for variance becomes identical to one discussed by Rao in [9] in which he showed that the best estimate is unweighted. These conditions will be met when the parameters describing each individual's response are considered to be random variables with dispersion matrix  $\Gamma$  and in addition to this source of variation there is an independent source with covariance matrix  $I\sigma^2$  and expected value zero. The first reflects the deviation of each individual's curve from the best description for its group, and the second source reflects the variation of each individual about its own curve. This is an assumption made and often reasonable in mixed model problems.

Suppose that  $\Sigma = T\Gamma T' + \sigma^2 I$  where  $\Gamma$  is an unknown dispersion matrix,  $\sigma^2$  is an unknown variance, and  $T$  is an arbitrary known matrix. When we make our transformation, we let  $B_1 = B(B'B)^{-1}$ ,  $B_2$  be a column basis  $(I - B(B'B)^{-1}B')T$ , and  $B_3$  be such that  $B_3'B = 0$ ,  $B_3'T = 0$ , and  $[B_1:B_2:B_3]$  is  $(p \times p)$  of full rank. The transformation is

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} B_1' \\ B_2' \\ B_3' \end{bmatrix} X.$$

We then find that the covariance matrix between corresponding columns of  $Y_1$  and  $Y_2$  is

$$\text{Cov}(Y_{1i} Y_{2i}) = (BB')^{-1} B' (T\Gamma T' + \sigma^2 I) B_2$$

between corresponding columns of  $Y_1$  and  $Y_3$  is

$$\text{Cov}(Y_{1i} Y_{3i}) = (BB')^{-1} B' (T\Gamma T' + \sigma^2 I) B_3$$

$$= 0$$



and the covariance matrix between corresponding columns of  $Y_2$  and  $Y_3$  is

$$\begin{aligned} \text{Cov}(Y_{2i} Y_{3i}) &= B_2'(TT' + \sigma^2 I)B_3 \\ &= 0. \end{aligned}$$

Here  $Y_3$  provides no information about  $Y_1$  or  $Y_2$  and may be disregarded. We then make inferences about  $\xi$  using the conditional distribution of  $Y_1$  given  $Y_2$  as described earlier. This partitions the within individual error space into parts which are useful as covariables and a part which can be discarded.

As a general principal only those vectors in the within individual error space that are correlated with the vectors in the within individual estimation space should be used as covariables. Due to the patterns assumed above the two groups of vectors are mutually independent. Other assumptions about the patterns in  $\Sigma$  and about the model could lead to the choice of only a few of the vectors in the error space being included as covariables.

Rao [11] shows that the difference in variance between the weighted and unweighted estimates are

$$B_1' \Sigma B_1 - \frac{N-r-1}{N-r-(p-q)-1} (B' \Sigma^{-1} B)^{-1}.$$

$B_1' \Sigma B_1 - (B' \Sigma^{-1} B)^{-1}$  is non-negative definite. But since

$(N-r-1)/[N-r-(p-q)-1] \geq 1$ , the difference in variance is not necessarily in favor of the weighted estimate. By reducing the number of covariates, we make  $p-q$  small and hence the multiplier moves closer to its lower bound of unity, and by discarding the vectors that are independent of  $\hat{\xi}$  we decrease the effect of random variation in the weights actually used.

It was illustrated above that depending on the pattern in  $\Sigma$ , all of the vectors in  $Y_2$  may not carry useful concomitant information. In particular fields of inquiry, the most useful  $Y_2$  could be found, and then only those concomitants should be used in the analysis.

Consider what could happen if there were 20 observations made on each individual and a linear trend provided an adequate description for each individual. Then  $p-q = 18$  and if a large reduction in variance did not result from using all covariates, efficiency would be lost. However, it is entirely reasonable to suppose that 5 or 6 properly chosen covariates could effect a reduction in variance comparable to using all 18. Then instead of  $p-q = 18$  we should have 5 and efficiency should be gained. When the number of useful concomitant variables can be decreased without sacrificing an appreciable amount of information, more powerful tests will result by increasing the degrees of freedom and by decreasing the variance of  $\hat{\xi}$ .

An alternative approach is to use the same data to find the useful concomitants as was used in making the test. Of course this will result in the probability of a Type I error no longer being exactly  $\alpha$ , but it might be worth the price. Rao [11] points out that the dominant characteristic vectors of

$$[I - B(B'B)^{-1}B']S$$

can be determined. These constitute the vectors in  $B_2$  instead of the previously defined form. We require that  $B'B_2 = 0$ ; hence it is essential that the left handed characteristic vector be used. It is easily seen that if  $S$  approximates  $I\sigma^2$ , the roots will be  $\sigma^2$  or 0 with multiplicity  $p-q$  and  $q$  respectively.

Williams [14] discussion throws some light on allocating effort toward making  $p$  or  $N$  large when the amount of experimental effort and  $q$  are fixed. He conjectures that when the response model is exact, it is better to replicate a small design over many individuals than to sample the response at too many points. The pattern in  $B_1' \Sigma B_2$  would be the deciding factor if it were known. The fact that one can sometime do as good or better when using a sub-set of covariates on even none than by using all is evidence that effort may have been wasted.

#### Confidence Intervals

Confidence bounds on  $C \xi V$  were derived by Khatri [6]. He shows that  $(1 - \alpha)$  100% confidence bounds for  $a'(C \xi V)b$  are given by

$$a'(C \hat{\xi} V)b \pm [\lambda_\alpha (a'Q_2a)(b'V'R_1Vb)]^{\frac{1}{2}}$$

where  $Q_2 = C(B'S^{-1}B)^{-1}C'$ ,  $R_1 =$  the leading upper left  $r \times r$  minor of  $DD'$  and  $\lambda_\alpha = \frac{u_\alpha}{1-u_\alpha}$ ,  $u_\alpha$  being the  $\alpha$ th percentage point of the largest root test criteria with degrees of freedom  $s = \min(c, v)$ ,  $m = (|v-c| - 1)/2$ ,  $n = (N-r-(p-q)-c-1)/2$ . These confidence bounds hold for all vectors  $a$  and  $b$ . Notice that  $Q_2$  is the error sum of products matrix associated with the test of the hypothesis  $C \xi V = 0$  when the model given by equation (4a) is used. When  $s = 1$ ,  $\lambda_\alpha$  can be found from the F distribution. In the case  $v = 1$ ,  $\lambda_\alpha = \frac{c}{N-r-(p-q)-c+1} F_{[\alpha, c, N-r-(p-q)-c+1]}$  and when  $c = 1$ ,  $\lambda_\alpha = \frac{v}{N-r-(p-q)} F_{[\alpha, v, N-r-(p-q)]}$ .

The linear parametric function  $a'(C \xi V)b$ , where  $a$  and  $b$  are arbitrary, is an expression of the general linear function belonging to

the vector space generated by  $C \xi V$ . Thus the Khatri bounds permit us to control simultaneously the error levels for all linear parametric functions in subspace  $C \xi V$  of the total parameter space  $\xi$ . The smaller we can keep  $C \xi V$ , while retaining the comparisons of interest, the shorter will be the confidence intervals or the more powerful will be the tests. This technique is an extension of Scheffé's method of multiple comparisons.

When only a few specific contrasts are tested, shorter intervals can usually be obtained by using the Bonferroni inequality which results in using  $F_{\alpha/k}$  where  $k$  is the number of confidence intervals, than from the intervals based on the largest root criteria which cover all linear functions in the same space as  $C \xi V$  with probability  $1 - \alpha$ . The lengths of the confidence intervals can be compared by computing

$$\frac{c}{N-r-(p-q)+1} F_{\frac{\alpha}{k}}[\frac{\alpha}{k}, c, N-r-(p-q)-c+1] \text{ or } \frac{v}{N-r-(p-q)} F_{\frac{\alpha}{k}}[\frac{\alpha}{k}, v, N-r-(p-q)] \text{ whichever}$$

is appropriate and comparing it to  $\frac{u_{\alpha}}{1-u_{\alpha}}$ .

Examples

Example 1. Data, taken from [4], presented in Table 2, are the ramus heights, measured in mm., of a cohort of boys at 8, 8 1/2, 9, and 9 1/2 years of age. The objective is to establish a normal growth curve for the use of orthodontists. It is apparent within the range covered, a straight line should fit the data.

Table 2

Ramus Height of 20 Boys

Individual	Age in Years			
	8	8 1/2	9	9 1/2
1	47.8	48.8	49.0	49.7
2	46.4	47.3	47.7	48.4
3	46.3	46.8	47.8	48.5
4	45.1	45.3	46.1	47.2
5	47.6	48.5	48.9	49.3
6	52.5	53.2	53.3	53.7
7	51.2	53.0	54.3	54.5
8	49.8	50.0	50.3	52.7
9	48.1	50.8	52.3	54.4
10	45.0	47.0	47.3	49.3
11	51.2	51.4	51.6	51.9
12	48.5	49.2	53.0	55.5
13	52.1	52.8	53.7	55.0
14	48.2	48.9	49.3	49.8
15	49.6	50.4	51.2	51.8
16	50.7	51.7	52.7	53.3
17	47.2	47.7	48.4	49.5
18	53.3	54.6	55.1	55.3
19	46.2	47.5	48.1	48.4
20	46.3	47.6	51.3	51.8
Mean	48.66	49.62	50.57	51.45
s.d.	2.52	2.54	2.63	2.73

The analysis fits into the general model (2) as follows: Since there is only a single group, the design matrix for individuals is

$$A = (1, \dots, \dots, 1),$$

$1 \times 20$

and the within individual design can be expressed in terms of orthogonal polynomial coefficients as

$$B = \begin{bmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 1 \\ 1 & 3 \end{bmatrix} .$$

The parameter matrix is  $\xi' = [\xi_0, \xi_1]$ .

To test the fit of the linear model, we can choose

$$B_2 = \begin{bmatrix} 1 & -1 \\ -1 & 3 \\ -1 & -3 \\ 1 & 1 \end{bmatrix}$$

which are the orthogonal polynomial coefficients for quadratic and cubic terms.

The fit of this model is tested by starting with a model in which the fact that 4 measurements made on each individual are functionally related is ignored, i.e.,

$$E(X) = \begin{matrix} \xi^* & A \\ 4 \times 20 & 4 \times 1 & 1 \times 20 \end{matrix} ,$$

and testing the hypothesis  $C\xi^*V = 0$ , where  $C = B_2'$  and  $V = I$ . Then if this hypothesis is accepted, the vectors  $B_2'X$  are assigned to the error space.

The variances and covariances are shown above the diagonal and the correlations below in the following matrix.

6.33	6.19	5.78	5.55
0.97	6.45	6.16	5.92
0.87	0.92	6.93	6.95
0.81	0.85	0.97	7.47

The test for the fit of the model yields

$$C \hat{\xi}^* V = B_2' \hat{\xi}^* 1 = [-0.09, -0.04].$$

$2 \times 4$     $4 \times 1$     $1 \times 1$

The likelihood ratio criterion yields  $\chi^2 = .18$  with 2 degrees of freedom. This is far below the critical value. Therefore, we assign  $B_2'X$  to the error space.

Proceeding to the next stage of the analysis we compute  $B_1'$  and  $B_2'X$  where

$$B_1' = \begin{bmatrix} 1/4 & -3/20 \\ 1/4 & -1/20 \\ 1/4 & 1/20 \\ 1/4 & 3/20 \end{bmatrix},$$

and  $B_2$  is as defined above. It is easy to check that  $B_1'B = I$  and  $B_2'B = 0$ . Since  $S$  has been computed, the correlations between the vectors in  $B_1'X$  and  $B_2'X$  are easily found to be

	$B_1'X$	$B_2'X$
$B_1'X$	1.00	0.12
$B_2'X$	-0.08	-0.06

  

	$B_1'X$	$B_2'X$
$B_1'X$	1.00	0.15
$B_2'X$	0.15	1.00

This shows that only the vector expressing the cubic term in  $B_2'X$  is correlated with the vectors in  $B_1'X$ . Therefore, it seems reasonable to expect a smaller variance when only this term is used as a covariate.

This covariate is easily incorporated into the analysis by augmenting the original  $A$  by  $B_2'X$  where  $B_2' = [-1, 3, -3, 1]$ . If we define

$$B_2' = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 3 & -3 & 1 \end{bmatrix},$$

the quadratic and cubic terms will be incorporated as covariates which is equivalent to weighting by  $S^{-1}$ .

The results of the three analyses are shown below.

	Covariables Used		
	None	Cubic	Quadratic and Cubic
$\hat{\xi}_0$	50.07	50.07	50.05
$\hat{\xi}_1$	.4665	.4629	.4654
$20(\hat{\text{Var}}(\hat{\xi}))$	$\begin{bmatrix} 6.27 & .090 \\ .090 & .085 \end{bmatrix}$	$\begin{bmatrix} 7.00 & .070 \\ .070 & .062 \end{bmatrix}$	$\begin{bmatrix} 7.85 & .087 \\ .087 & .069 \end{bmatrix}$
400(Gen. var.)*	.525	.430	.535

\*The generalized variance is defined as the determinant of the  $\hat{\text{Var}}(\hat{\xi})$  in accordance with [2], page 166.

This shows that the unweighted analysis is preferable to weighting by  $S^{-1}$ . The unweighted analysis and the analysis using two covariates correspond to Methods I and III in [4]. The reason that Method I gave slightly narrower confidence bands in [4] is due to the smaller variance of the estimated intercept in unweighted analysis. The analysis in which the cubic term is used as a covariable is intermediate between Methods I



and III in regard to the variance of the intercept and its slope has the smallest estimated variance. The analysis using only the cubic term has the smallest generalized variance.

Example II. The data shown in Table 1 are used in this example.

They are typical of the type of data obtained from nondescript dogs often used in medical experimentation.

The covariance and correlation matrices are

0.2261	0.1721	0.1724	0.2054	0.1705	0.1958	0.1817
0.88	0.1696	0.1840	0.1919	0.1628	0.1700	0.1644
0.58	0.71	0.3917	0.3473	0.2370	0.1876	0.2194
0.65	0.70	0.84	0.4407	0.3689	0.2870	0.2582
0.54	0.60	0.57	0.84	0.4337	0.3733	0.3178
0.57	0.57	0.41	0.60	0.78	0.5235	0.4606
0.53	0.56	0.50	0.54	0.67	0.89	0.5131

The variances and covariances are shown above the diagonal and correlations below.

A plot of the data suggest that a third degree polynomial should fit the data. A test that the 4th, 5th, and 6th degree polynomial coefficients have expected value zero simultaneously yield  $\chi^2 = 8.259$  with 12 degrees of freedom. This does not approach statistical significance.

This test is made by defining the model  $E(X) = \xi * A'$  and testing  $C\xi * V$ , where A is as defined when the example was originally presented,  $V = I$  and

$$C = B_2' = \begin{bmatrix} 3 & -7 & 1 & 6 & 1 & -7 & 3 \\ -1 & 4 & -5 & 0 & 5 & -4 & 1 \\ 1 & -6 & 15 & -20 & 15 & -6 & 1 \end{bmatrix}$$

The correlations among the vectors in the estimation space and in the error space are shown in the following table.

	$B_1'X$				$B_2'X$		
	1.00						
$B_1'X$	0.42	1.00					
	-0.17	0.25	1.00				
	-0.16	-0.37	-0.05	1.00			
	0.22	-0.30	-0.70	-0.31	1.00		
$B_2'X$	-0.19	0.02	-0.04	-0.57	-0.12	1.00	
	-0.11	0.02	-0.20	0.43	0.18	-0.08	1.00

Notice that there are relatively high correlations between the vectors in the error space with the quadratic and cubic terms in the estimation space. Therefore, all the vectors in the error space are used as co-variables. The estimates of  $\xi$ ,  $\hat{\text{Var}}(\hat{\xi})$ , and tests of homogeneity are shown in Table 3. The estimated parameters have not been scaled by dividing by the sums of squares of the orthogonal polynomial coefficients.

Table 3

Comparison of Weighted and Unweighted Estimates									
		<u>No Covariates</u> (Unweighted)				<u>3 Covariates</u> (Weighted by $S^{-1}$ )			
	Group	Inter- cept	Linear	Quad.	Cubic	Inter- cept	Linear	Quad.	Cubic
$\langle \xi \rangle$	1	32.57	4.48	-3.63	-0.98	32.70	4.14	-4.68	-1.14
	2	24.87	-0.56	-0.78	0.11	25.07	-0.69	-1.00	0.11
	3	29.05	3.49	-4.54	0.23	28.59	4.10	-2.90	-0.48
	4	27.58	1.72	-1.79	0.16	27.12	1.91	-1.55	0.04
$\widehat{\text{Var}}(\hat{\xi})$	$\frac{1}{N_j}$	[ 12.75 4.67 -2.60 -0.51 ]				[ 14.10 6.66 -0.96 -1.00 ]			
		[            9.83    3.31 -1.02 ]				[            10.84    0.66 -1.04 ]			
		[                    18.47 -0.20 ]				[                    10.86    0.59 ]			
		[                                    0.76 ]				[                                    0.45 ]			
Generalized Variance		$\frac{1}{N_j^4}$ 1061				$\frac{1}{N_j^4}$ 349			
Test of Homogeneity of Groups		$\chi^2 = 27.27$ p = .008				$\chi^2 = 27.61$ p = .007			

These results show that the generalized variance was decreased weighting by  $S^{-1}$ . However when considered individually the decrease was not uniform for all parameters estimated. Only the quadratic and cubic terms show an important decrease in variance; the intercept and slopes actually showed increases.

The test for homogeneity shows that there are significant differences among the response patterns of the four groups. Comparisons among the four groups are shown in Table 4. The comparisons are made by setting  $C = I$  and

$V' = (100-1)$  to compare the first and fourth groups. Other comparisons can be made by changing  $V$  appropriately in the general form  $C \xi V = 0$ .

In spite of the large reduction in the generalized variance there is little change in the results of specific tests. However, either method has the ability to pinpoint fairly precisely where the differences lie.

#### Summary

A method of analysis of "growth curves" is developed and illustrated which yields identical results to weighting inversely by the sample variance. This method has the additional feature of allowing flexibility in weighting by choosing subsets of covariates that have some special property. This permits exploitation of Rao's observation that especially selected subsets of the covariate may yield better estimates than the complete set in some cases. However, application of the results to actual data makes it clear that it is not easy to determine when it will be advantageous to use all the vectors in the error space as covariates (weighting by  $S^{-1}$ ), or a subset, or how to choose the "best" subset from examination of sample covariance or correlation matrices.

Table 4  
Tests of Differences Among Groups

Group Comparison	<u>Unweighted</u>				<u>Weighted by S<sup>-1</sup></u>				
	Intercept	Linear	Quad.	Cubic	Intercept	Linear	Quad.	Cubic	Multi- variate p
1 vs 4 100-1	5.00**	2.76	-1.84	-1.14**	5.59***	2.23	-3.13*	-1.15**	.007
2 vs 4 010-1	-2.71	-2.28	1.01	-0.05	-2.04	-2.59	0.55	-.11	.517
3 vs 4 001-1	1.47	1.77	-2.75	-0.39	1.47	2.19	-0.85	-0.48	.619
1 vs 3 10-10	3.53*	0.99	0.90	-0.75	4.12*	0.03	-2.28	-0.66	.084
2 vs 3 01-10	-4.18*	-4.05**	3.76	0.35	-3.51	-4.79***	1.41	0.59	.074
1 vs 2 1-100	7.71***	5.04**	-2.85	-1.10**	7.63***	4.82**	-3.68*	-1.25**	.001

Probabilities levels of univariate tests on each coefficient are indicated as

\*p < .05 ,

\*\*p < .01 ,

\*\*\*p < .005.

## REFERENCES

1. Allen, David. (1967). Multivariate Analysis of Nonlinear Models, North Carolina Institute of Statistics Mimeo Series No. 558.
2. Anderson, T. W. (1958). Introduction to Multivariate Statistical Analysis, John Wiley and Sons, Inc., New York.
3. Box, G. E. P. (1950). "Problems in the Analysis of Growth and Wear Curves." Biometrics, 6, 362-89.
4. Elston, R. C. and Grizzle, J. E. (1962). "Estimation of Time-Response Curves and Their Confidence Bands." Biometrics, 18, 148-59.
5. Gleser, Leon and Olkin, Ingram. (1966). "A k-Sample Regression Model with Covariance." Multivariate Analysis Proceedings of an International Symposium held in Dayton, Ohio June 14-19, 1965. Academic Press, 59-72.
6. Khatri, C. G. (1966). "A Note on a MANOVA Model Applied to Problems in Growth Curves." Annals of the Institute of Statistical Mathematics, 18, 75-86.
7. Morrison, Donald F. (1967). Multivariate Statistical Methods, McGraw-Hill Book Company, New York.
8. Potthoff, Richard R. and Roy, S. N. (1964). "A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems." Biometrika, 51, 313-26.
9. 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, 447-58.
10. Rao, C. R. (1966). "Covariance Adjustment and Related Problems in Multivariate Analysis." Multivariate Analysis Proceedings of an International Symposium held in Dayton, Ohio, June 14-19, 1965. Academic Press, 87-103.
11. Rao, C. R. (1967). "Least Squares Theory Using An Estimated Dispersion Matrix and Its Application to Measurement of Signals." Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Volume 1, 355-72.
12. Schatzoff, Martin. (1966). "Exact Distributions of Wilks' Likelihood Ratio Criterion." Biometrika, 53, 347-58.
13. Watson, Geoffrey S. (1967). "Linear Least Squares Regression." The Annals of Mathematical Statistics, 38, 1679-99.
14. Williams, James S. (1967). "The Variance of Weighted Regression Estimators." Journal of the American Statistical Association, 62, 1290-1301.