INTRODUCTION

Growth curve analysis usually refers to the fitting and testing hypotheses about a series of measurements in time of some variable of interest.

Growth curves and repeated measure experiments are properly analyzed using multivariate techniques, since there are correlations between observations on the same experimental unit over time. Wishart (1938) introduced a class of models which has had a great influence on the statistical analyses of growth studies. This approach was to reduce a growth curve to set a least squares estimates of orthogonal polynomial coefficients. These coefficients are available at each point of an experimental design grid and may be subjected to univariate and multivariate analysis to test the differences between the treatments. This approach has been extended by several authors for equally spaced time points and to test for overall significance between treatments. Dempster (1963) proposed the stepwise method of analysis. Patel (1919) has described the stepwise method for both equally and unequally spaced time points. In this paper, two additional methods for stepwise analysis of growth curves are developed:

(I) Stepwise MANOVA with Linear Trend in Dose

When the data are measured at a series of time points for several doses, then one might wish to test the overall significance between doses but also to test for a linear dose effect.

(II) Stepwise MANOVA for a 2 x 2 Design

For the 2 x 2 factorial experiment with repeated measurements, we are interested in testing main effects and interactions.

Using PROC MATRIX of SAS, two MACROS were written for these methods. These techniques are illustrated with numerical examples.

METHOD ONE: TEST FOR LINEAR DOSE EFFECT

Suppose there are I treatments (doses) with Ni subjects assigned to the i-th dose group, each observed at K equally or unequally spaced time points. We wish to determine the significance of a linear trend in doses whose effects are measured in time. Using orthogonal polynomials, the responses can be modeled in terms of suitable degree polynomials which adequately characterize the curves. By this approach we reduce the dimensionality of the data without sacrificing essential information.

The polynomial is fitted to growth curve data with the aid of orthogonal polynomials. Each succeeding step gives a "better" fit to the curve than the previous curve. The purpose of using higher order polynomials is to differentiate treatments when lower order comparisons do not show any significant differences. At present there is no test for lack of fit or any technique for handling missing data.

(1.1) MODEL

The model for the multivariate analysis in one-way classification is

$$
Y = P \beta + \varepsilon
$$

where \( Y \) is an \( N \times K \) data matrix, \( P \) is an \( N \times (I + 1) \) design matrix, \( \beta \) is an \( (I + 1) \times K \) matrix of unknown parameters, \( \varepsilon \) is an \( N \times K \) error matrix, each row of which is assumed to be independently distributed as multivariate \( N(0, \Sigma) \), \( N = E_i N_i = \) number of subjects in all treatments, \( I = \) number of treatments (doses), and \( K = \) number of repeated measurements (time points).

The data are transformed into polynomials to compare different meaningful aspects of the curves. Therefore, it is necessary to generate the general orthogonal polynomial coefficients corresponding to equally or unequally spaced time points. This is done as follows.

(1.2) GENERATION OF GENERAL ORTHOGONAL POLYNOMIAL COEFFICIENTS

Let \( \tilde{X} = (X_1, \ldots, X_K) \) be a \( 1 \times K \) row vector of times at which observations are taken. Then the general orthogonal polynomial coefficients corresponding to \( \tilde{X} \) are obtained by the relation

$$
A_k = B_k - \sum_{i=0}^{K-1} \left( \frac{A_i}{A_k} \right) A_k
$$

where

$$
B_k = [(X_1 - \tilde{X})^T, \ldots, (X_K - \tilde{X})^T]^T, \quad \tilde{X} = \sum_i X_i / K,
$$

for \( k = 1, \ldots, p \), the degree of the polynomial (\( p < K \)).

When the range of the time points is very large, these polynomial coefficients are usually very small at one end and extremely

224
Thus the matrix \( Q = (\mathbf{q}_i) \) is a \( K \times (p+1) \) standardized orthogonal polynomial matrix. This can also be obtained by using \( Q = \text{ORPOL}(X, P) \).

Using matrices \( Y \) and \( U \), the score matrix is obtained as \( Y = X U \). Let \( Q_{nk1} = T \), where \( T \) is the matrix whose partitions consist of the score matrix \( X_{nk(p+1)} \), the design matrix \( \mathbf{D}_{nk1} \), and the linear dose matrix \( \mathbf{D}_{nk1} \), where \( J = p + I + 3 \), \( L \) is column vector of actual dose (or log-dose).

The design matrix \( D \) is that standardly used in analysis of variance situations for the one-way classification model, which can be generated by setting the first column of \( D \), \( D_0 = (1, \ldots, 1)^T \), and the remaining \( D_i \), \( i \) columns are such that
\[
D_{i, j} = \begin{cases} 
1, & \text{if subject } i \text{ receives treatment } j, \\
0, & \text{otherwise}.
\end{cases}
\]

Using this matrix, the raw sum of squares and cross-products matrix is computed as \( \mathbf{SSR} = \mathbf{Y}^T \mathbf{C} \mathbf{Y} \).

### 1.3 Sweeping Method and Multivariate Analysis of Variance

The matrix \( \mathbf{SSR} \) is symmetric and singular because \( D_0 = \mathbf{1}_J \). We perform sweeping operations as described by Beaton (1964) and Dempster (1969). The purpose of sweeping is to invert submatrices indexed by the pivot, producing new matrices which are then used in a series of independent tests of significance on the scores obtained by orthogonal polynomials. The explanation of the elements obtained from sweeping is described in Dempster (1969) and Patel (1979).

To obtain the total sum of squares, within sum of squares, and sum of squares for the linear effect of dose, the sequence of sweeping is given in Table 1.

### 1.4 Test of Significance

Assuming that the observations constitute a random sample from a multivariate normal distribution with homogeneous dispersions, we test the null hypothesis for testing the linear effect in doses at each score.

At each score, the stepwise test statistic is given by
\[
F = \frac{\text{MS}_{LDE}}{\text{MS}_w}.
\]

Under \( H_0 \), this ratio has a \( F \) distribution with \( I \) and \( N-I+1 \) degrees of freedom. Thus a sequence of independent \( F \) tests \( F_1, \ldots, F_p \) are generated where the denominator degrees of freedom decrease by one at each step. In these tests, we have "adjusted" S.S. which means that we are looking at differences in scores after removing effects which can be explained in terms of previous variables.
METHOD TWO: BALANCED 2 x 2 DESIGN

Suppose we have two treatments each measured at two levels with n subjects assigned to each treatment combination, each observed at K equally or unequally spaced fixed time points. We assume that no observation is missing. We may also have some meaningful covariate(s) in addition to treatment effects.

We wish to determine the significance of the main effects and their interaction. Using Patel (1979), the test statistic can be obtained as follows.

(II.1) MODEL

The multivariate covariance model in (7) has been described by Patel (1979) as follows:

\[ \chi = D_0 + D_1 + D_2 + D_3 \]  

Let \( E_{MN} = \{ U | C | D \} \) be the matrix whose partitions consist of the score matrix \( U\), the covariates matrix \( C\), and the design matrix \( D\), where \( T = (p-1)\times W + 1\), \( W = \chi Q\).

The design matrix \( D\) is that standardly used in analysis of variance situations for a 2 x 2 factorial design model, which can be generated by setting the columns of \( D\) such that:

\[
\begin{array}{cccc}
\text{Factor} & D_0 & D_1 & D_2 & D_3 \\
A & B & C & D & E \\
1 & 1 & 1 & -1 & -1 & 1 \\
2 & 1 & 1 & 1 & -1 & 1 \\
2 & 1 & 1 & 1 & 1 & 1 \\
\end{array}
\]

Using this matrix, the raw sum of squares and cross-products matrix is computed as:

\[ SS_R = \chi^2 \cdot E \cdot E^T \]

Depending upon the number of variables \( K\) and number of subjects \( N\), the analysis could be performed on original data or on scores.

(II.2) SWEEPING METHOD AND MULTIVARIATE ANALYSIS OF COVARIANCE

The matrix \( SS_R\) is symmetric and non-singular. To obtain total S.S., within S.S., and S.S. for main effects and interaction, the sequence of sweeping is given in Table 2.

(II.3) TEST OF SIGNIFICANCE

To determine the significance of main effects and interaction at each \( T_i \) (covariate or score: degree of a polynomial), we test the null hypothesis

\[ H_0 : T_i = T_j \]

against the alternatives

\[ H_1 : T_i \neq T_j \] for \( i \neq j \)

The stepwise test statistic is given by

\[ F_a = MS_{T_a} / MS_{a} \]

This ratio has an F-distribution with \( (L,N-(n-W)) \) degrees of freedom.

These two methods were programmed using PROC MATRIX of SAS (Burr, et al. 1979). This also allowed us to use other capabilities in SAS such as PROC GLM for the overall univariate/multivariate analysis. Two examples were given to illustrate the method in SAS. The program listing is given in Appendix 1.

EXAMPLE I: STEPWISE MANOVA FOR TESTING LINEAR DOSE EFFECT

In this study there are 54 rats tested on three treatments. Body weights are measured on day 0, 6, 13, 16 and 20. The control group has 19 rats, the low dose group (300 mg/kg) has 18 rats, and the high dose group (600 mg/kg) has 17 rats. Cubic polynomials were fitted corresponding to the days. The significance between doses (treatments) and the linear dose effect (trend) is determined for each score. The results are given in Table 3.

It follows from the results given in Table 3 that treatments are not significant at any score,
but at score 3, the linear dose effect is significant (p = 0.06), suggesting that there is a linear dose effect.

**EXAMPLE II: STEPWISE MULTIVARIATE ANALYSIS OF THE 2 × 2 DESIGN**

In this study there are 40 rats tested on two treatments each at two levels. Body weights are measured on week 0, 1, 2, 3, and 4. Considering the initial body weight as a covariate, the cubic polynomials were fitted corresponding to weeks 1, 2, 3, and 4. The significance of main effects and interaction was determined for each covariate and score (degree of a polynomial). The results are given in Table 4.

It follows from the results in Table 4 that factor A is significant for the covariate. Also the main effects (factors A and B) are significant for score (degree of a polynomial). The results are summarized in Table 3.

**SUMMARY**

The purpose of this paper is to show that the stepwise multivariate analysis can provide better insight into experimental results than can an overall test. Two methods of analysis for repeated measurements data have been described and illustrated. These methods perform stepwise tests for a linear effect of dose and the tests for main effects and interactions in the 2 × 2 design and they can be easily extended to summarize growth curves for other designs. The use of orthogonal polynomials has been explored as one approach to meet the needs of growth curve analysis. Further development of these procedures will be undertaken. Investigations directed toward improving procedures to analyze unbalanced data, and use splines as alternatives to polynomials are underway.

**ACKNOWLEDGMENTS:** The author is grateful to Professor A. P. Dempster and Dr. M. R. Selwyn for their helpful comments.

---

**TABLE 3**

**EXAMPLE I: STEPWISE MANOVA USING CUBIC POLYNOMIAL**

<table>
<thead>
<tr>
<th>Degree</th>
<th>F-Cal</th>
<th>D.F.</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (Mean)</td>
<td>2.22</td>
<td>2, 51</td>
<td>NS*</td>
</tr>
<tr>
<td>1 (Lin)</td>
<td>0.66</td>
<td>2, 51</td>
<td>NS</td>
</tr>
<tr>
<td>2 (Quad)</td>
<td>0.19</td>
<td>2, 51</td>
<td>NS</td>
</tr>
<tr>
<td>3 (Cubic)</td>
<td>0.26</td>
<td>2, 51</td>
<td>NS</td>
</tr>
</tbody>
</table>

**OVERALL WILK'S LAMBDA CRITERION:** 0.92 NS

* *(0.12)*

**TABLE 4**

**TEST FOR LINEAR DOSE EFFECT (LDE)**

<table>
<thead>
<tr>
<th>Degree</th>
<th>F-Cal</th>
<th>D.F.</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (Mean)</td>
<td>2.22</td>
<td>2, 51</td>
<td>NS*</td>
</tr>
<tr>
<td>1 (Lin)</td>
<td>0.66</td>
<td>2, 51</td>
<td>NS</td>
</tr>
<tr>
<td>2 (Quad)</td>
<td>0.19</td>
<td>2, 51</td>
<td>NS</td>
</tr>
<tr>
<td>3 (Cubic)</td>
<td>0.26</td>
<td>2, 51</td>
<td>NS</td>
</tr>
</tbody>
</table>

**EXAMPLE II: STEPWISE MANCOV USING CUBIC POLYNOMIAL (2 × 2 DESIGN)**

<table>
<thead>
<tr>
<th>Parameter Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F-Cal</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_1$</td>
<td>A</td>
<td>1</td>
<td>17139.6</td>
<td>17139.6</td>
<td>123.4 &lt;.01</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1</td>
<td>8.1</td>
<td>8.1</td>
<td>0.06 NS</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>1</td>
<td>22.5</td>
<td>22.5</td>
<td>0.2 NS</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>36</td>
<td>5002.2</td>
<td>136.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>39</td>
<td>22172.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_0$</td>
<td>A</td>
<td>1</td>
<td>15243.9</td>
<td>15243.9</td>
<td>34.1 &lt;.01</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1</td>
<td>2374.3</td>
<td>2374.3</td>
<td>3.3 0.03</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>1</td>
<td>99.5</td>
<td>99.5</td>
<td>0.2 NS</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>35</td>
<td>15526.8</td>
<td>446.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>38</td>
<td>32941.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_1$</td>
<td>A</td>
<td>1</td>
<td>19.9</td>
<td>19.9</td>
<td>0.7 NS</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1</td>
<td>1.9</td>
<td>1.9</td>
<td>0.07 NS</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>1</td>
<td>6.6</td>
<td>6.6</td>
<td>0.2 NS</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>34</td>
<td>918.7</td>
<td>27.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>37</td>
<td>945.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_2$</td>
<td>A</td>
<td>1</td>
<td>0.02</td>
<td>0.02</td>
<td>0.002 NS</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1</td>
<td>14.7</td>
<td>14.7</td>
<td>1.6 NS</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>1</td>
<td>15.8</td>
<td>15.8</td>
<td>1.7 NS</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>33</td>
<td>310.4</td>
<td>9.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>36</td>
<td>341.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_3$</td>
<td>A</td>
<td>1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.003 NS</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1</td>
<td>1.5</td>
<td>1.5</td>
<td>0.1 NS</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>1</td>
<td>0.07</td>
<td>0.07</td>
<td>0.008 NS</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>32</td>
<td>311.6</td>
<td>9.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>35</td>
<td>320.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NOTE: These MACROS can be generalized for p > 3 time points, l > 3rd degree and V > 1 covariate.

REFERENCES


