

A PC PROGRAM FOR PERFORMING MULTIGROUP LONGITUDINAL COMPARISONS USING THE POTTHOFF-ROY ANALYSIS AND ORTHOGONAL POLYNOMIALS

THOMAS R. TEN HAVE^a, CHARLES J. KOWALSKI^b, EMET D. SCHNEIDERMAN^c and STEPHEN M. WILLIS^c

^a*Department of Biostatistics and* ^b*Department of Oral Biology, The University of Michigan, Ann Arbor, MI, 48109 and* ^c*Department of Oral and Maxillofacial Surgery, Baylor College of Dentistry, 3302 Gaston Avenue, Dallas, TX, 75246 (U.S.A.)*

(Received October 16th, 1991)

(Revision received November 26th, 1991)

(Accepted November 26th, 1991)

A PC-program performing the Potthoff-Roy (PR) multigroup (G -sample) analysis of longitudinal data is described and illustrated. This program and the underlying statistical model are useful in the comparison of several longitudinal samples. Applications include the study of growth, development, adaptation, aging, and treatment effects (in short, any phenomenon in which the passage of time is important) for which serial data are available. Specifically, this method fits polynomials to the average growth curves in the samples, and tests hypotheses concerning the curves themselves and the individual coefficients of the polynomials. The program features the utilization of orthogonal polynomial regression coefficients (OPRCs) and is written in *GAUSS*, a relatively inexpensive yet comprehensive matrix programming language. It is documented that using OPRCs to comprise the within-individual or time design matrix has several advantages over the more usual choice of the successive-powers-of- t form of this matrix and an example of one important such advantage is provided. *GAUSS* was employed to make the program readily-accessible (i.e., executable code) to biomedical investigators. The *GAUSS* compiler is not required to run this program. Information regarding the availability of the program is provided in the Appendix.

Keywords: Longitudinal studies; Group comparisons; Multivariate analysis; Growth curves; Orthogonal polynomials; PC program

Introduction

In a recent paper [1] we described the PR [2] method for comparing several groups of individuals on the basis of longitudinal data and implemented it using SAS. This program was limited to multivariate comparisons of polynomial regression coefficients, i.e., polynomials of degree D were fitted to each individual's growth pattern and the averages of the resulting $P (= D + 1)$ regression coefficients in the groups

Correspondence to: Emet D. Schneiderman, Department of Oral and Maxillofacial Surgery, Baylor College of Dentistry, 3302 Gaston Avenue, Dallas, TX, 75246, U.S.A.

were compared simultaneously (by MANOVA, the multivariate analysis of variance). No provision was made to allow the user to examine the coefficients one-at-a-time in order to isolate those coefficients responsible for any overall difference found. Nor was the user given any flexibility with regard to the choice of the within-individual (time) design matrix: given observations at times t_1, t_2, \dots, t_T the program computed and used the $T \times P$ (successive powers-of- t) matrix:

$$\mathbf{W} = \begin{bmatrix} 1 & t_1 & \dots & t_1^D \\ 1 & t_2 & \dots & t_2^D \\ \vdots & \vdots & \dots & \vdots \\ 1 & t_T & \dots & t_T^D \end{bmatrix} \quad (1)$$

for this purpose. The polynomial regression coefficients were then estimated by

$$\hat{\boldsymbol{\tau}}_{N \times P} = \mathbf{X} \mathbf{A}^{-1} \mathbf{W} (\mathbf{W}' \mathbf{A}^{-1} \mathbf{W})^{-1} \quad (2)$$

where \mathbf{X} is the $N \times T$ data matrix with elements x_{ij} , the value of the measurement for individual i ($i = 1, 2, \dots, N$) at time t_j ($j = 1, 2, \dots, T$) and \mathbf{A} is an arbitrary $T \times T$ matrix (often $\mathbf{A} = \mathbf{S}$, the sample covariance matrix, or $\mathbf{A} = \mathbf{I}$, the identity matrix, is used), corresponding to the weights used to obtain the weighted least squares estimate of $\boldsymbol{\tau}$.

We have since [3] documented the general superiority of orthogonal polynomial regression coefficients (OPRCs) for use in longitudinal data-analytic contexts, especially with regard to comparing individual regression coefficients among G groups. These are obtained when the within-individual design matrix

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & \phi_1(t_1) & \dots & \phi_D(t_1) \\ 1 & \phi_1(t_2) & \dots & \phi_D(t_2) \\ \vdots & \vdots & \dots & \vdots \\ 1 & \phi_1(t_T) & \dots & \phi_D(t_T) \end{bmatrix} \quad (3)$$

is used in place of \mathbf{W} , where ϕ_d is a polynomial of degree d with the properties

$$\sum_{j=1}^T \phi_d(t_j) = 0 \quad (4)$$

and

$$\sum_{j=1}^T \phi_d(t_j) \phi_k(t_j) = 0 \quad (5)$$

Polynomials satisfying (4) and (5) are said to be orthogonal; if they also satisfy

$$\sum_{j=1}^T \phi_d^2(t_j) = 1 \quad (6)$$

they are called orthonormal polynomials. The OPRCs for each of the N individuals are then estimated by

$$\hat{\alpha} = \mathbf{X}\mathbf{A}^{-1} \Phi(\Phi' \mathbf{A}^{-1} \Phi)^{-1} \quad (7)$$

and these, as will be illustrated below, accommodate the comparison of individual coefficients to an extent apparently not matched by $\hat{\tau}$. While the *multivariate* tests based on $\hat{\tau}$ and $\hat{\alpha}$ agree (the multivariate tests are *invariant*), *univariate* tests on *individual* coefficients do not in general agree; and we claim, largely on an empirical basis [3], that the use of $\hat{\alpha}$ often leads to more parsimonious conclusions. In any event, the program documented in this paper will permit the use of $\hat{\alpha}$ in the PR analysis. For any given run of the program the user is asked to choose between \mathbf{W} and (three forms of) Φ ; and to select the arbitrary matrix \mathbf{A} from the possibilities $\mathbf{A} = \mathbf{I}$, $\mathbf{A} = \mathbf{S}$ or a user-defined matrix. For a more detailed discussion of the choice between \mathbf{W} and Φ , see Ref. 3. Timm [4] provides some useful information concerning the choice of the arbitrary matrix, \mathbf{A} .

The primary purpose of this paper, then, is to implement the PR method for longitudinal data analysis using OPRCs in *GAUSS*, a matrix programming language that runs on PC-compatible microcomputers. The use of OPRCs adds to the usefulness of the PR method; the use of *GAUSS* extends accessibility to investigators who may not have access to a mainframe computer supporting SAS. Indeed, our program does not even require that *GAUSS* be installed by potential users. An important secondary aim is to illustrate the claimed advantage of $\hat{\alpha}$ over $\hat{\tau}$ when, as is often the case, the investigator wishes to follow a significant MANOVA by an examination of the individual regression coefficients in an attempt to discover 'where the significance lies', i.e., what aspects of the growth curves differ.

We begin by describing the program, we then consider an example of its use on a data set involving three groups of Guatemalan schoolchildren. Finally, we discuss the output and its implications for effective and consistent data analysis.

The Program

The *GAUSS* program, PR, tests for group differences in polynomial growth curves estimated by the PR approach. The first step involves determining the degree, D , of the polynomial to be used to model the growth processes in the groups. The user is given the choice of either specifying D or allowing the program to perform a series of significance tests to determine the smallest D adequate to fit the data. This involves what is referred to in the program as the 'reduced' and 'full' polynomial growth curve models (this terminology is fairly standard [5]). Let D_R denote the degree of the reduced model and D_F the degree of the full model. If the user wishes to specify the degree, D , to be used in the analysis (perhaps based on plots of the data) it is necessary only to take $D_R = D_F = D$. If the user wishes to determine D

on the basis of the data using a series of preliminary significance tests, then choosing numerical values for D_R and D_F ($D_R < D_F$) produces $D_F - D_R$ goodness-of-fit tests, these testing respectively the adequacy of the fit of $D_R, D_R + 1, \dots, D_F - 1$ degree polynomials to the data. D is said to be 'adequate' if the hypothesis that the higher-order regression coefficients (namely, the coefficients of t^{D+1}, t^{D_F}) in each of the groups is zero is not rejected by the test provided. Often the user will wish to start with $D_R = 1$ (linearity) and choose $D_F \leq T - 1$ as a degree which should be more than adequate to describe the growth curves in each of the groups.

Having determined D , the program then performs MANOVAs and ANOVAs to test for overall and pairwise group differences with respect to the coefficients of the selected polynomial model. The program also plots the fitted average curve for each group. It estimates the polynomial growth curve coefficients from either orthogonal, orthonormal or original time scores. Moreover, the program can read the response data (the X matrix) from either an ASCII file or a *GAUSS* data set.

The program is invoked by the command **GAUSSRUN PR**; The program will first ask the user if the response data are in an ASCII file or a *GAUSS* data set. If the data are in an ASCII file, the program will prompt the user for the name of the file, the number of subjects (rows), and the number of variables or columns in the file. If the data are in a *GAUSS* data set, the program will ask for the name of the *GAUSS* file.

The program will then prompt the user for the following information regardless of the type of data set input: (1) the numbers of the columns containing (a) the group variable, (b) the first response variable, and (c) the last response variable; (2) the degree of the reduced polynomial model (D_R); (3) the degree of the full polynomial model (D_F); (4) the numerical values of the time points (one per line); the default is the integers 1, 2, ..., T ; (5) a value indicating which of the four forms of the time design matrix, W , the user has selected [1, 3], non-integer orthogonal scores, integer orthogonal scores for equally spaced time points, orthonormal scores, or original time scores; (6) a value indicating which form of the arbitrary matrix, A , the user has selected, identity matrix, the covariance matrix of the response variables, or user provided matrix (if this selection is made the program then prompts the user for either an ASCII or *GAUSS* data set containing this matrix); and (7) the p -value above which models are considered to fit the data adequately.

The program prints the time design matrix (W), the arbitrary matrix (A), and the matrix (Y) of coefficients corresponding to the full polynomial model for each individual in the sample. Note that the program uses Y for what we have called $\hat{\gamma}$.

The program then tests the adequacy of the reduced model for the average curve and each higher degree polynomial until one satisfies the user-specified p -value criterion. The program selects the corresponding degree as the degree of the average model. If no polynomial with degree less than the full model fits the average curve then the full model is analyzed.

The coefficients of the individual and average curves based on the final model are displayed, as well as the covariance matrix of the average coefficients.

The program then provides the results of multivariate and univariate tests of group differences with respect to the coefficients of the final model. First, MANOVA statistics, including the Hotelling-Lawley trace, Wilks's lambda and

Pillai's trace [5], are displayed along with the associated F -tests of overall group differences in the coefficients. Second, F -tests are performed for pairwise group comparisons of all the coefficients together. Third, univariate F -tests are performed on each coefficient for overall and pairwise group differences.

Finally, the program generates a plot of the group average fitted curves and a plot of the average observed curves. Details about how one may obtain a copy of this program are provided in the Appendix.

An Example

In order to illustrate our program we consider three samples of children living in Guatemala which were studied in depth in [6]. The children comprising these samples differ in socioeconomic status (SES) and ethnicity: One is of high SES Ladino children (G_1); the second is of low SES Ladino children (G_2); and the third is of low SES Mayan children (G_3). There are 20 individuals in each group and we analyze their growth in stature, this being measured $T = 6$ times at ages 7, 8, 9, 10, 11 and 12 years. This same data set was used in [7] to study the tracking behavior of these individuals.

We found that, using the 5% level of significance, a linear equation was not quite adequate to fit the data in the three groups ($p \approx 0.03$ for the three multivariate test statistics described earlier), but a polynomial of degree $D = 2$ was adequate ($p \approx 0.20$). The estimated AGCs in the three groups are plotted in Fig. 1.

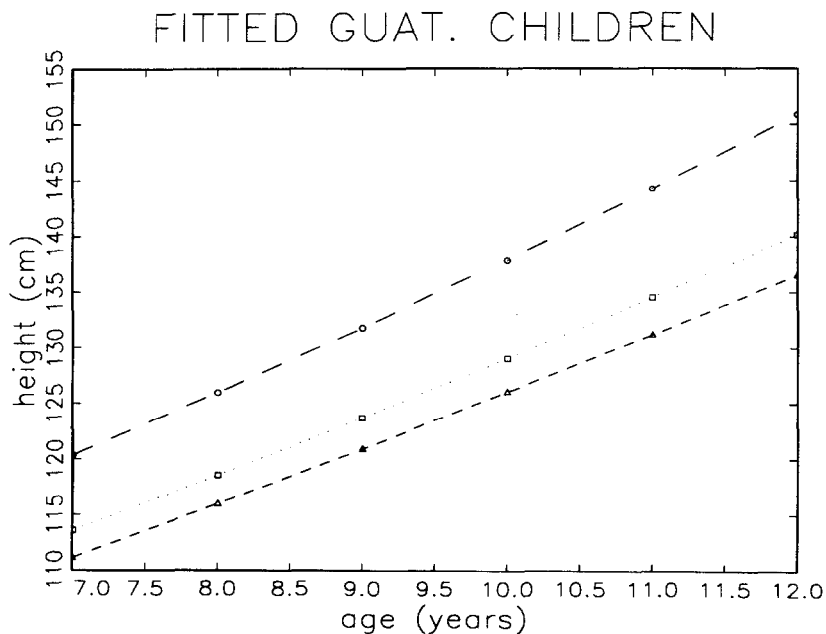


Fig. 1. Fitted polynomial growth curves for three groups of Guatemalan schoolchildren.

TABLE I

p -VALUES FOR THE TESTS OF PARALLELISM, COINCIDENCE AND THE INDIVIDUAL REGRESSION COEFFICIENTS USING W (τ) AND Φ (α) WITH $A = I$ AND $A = S$

	W		Φ	
	$A = I$	$A = S$	$A = I$	$A = S$
Parallelism	0.0035	0.0039	0.0035	0.0039
Coincidence	0.0001	0.0001	0.0001	0.0001
τ_1 (α_0)	0.2053	0.2831	0.0001	0.0001
τ_2 (α_1)	0.6767	0.9588	0.0006	0.0007
τ_3 (α_2)	0.4615	0.5630	0.4615	0.5630

Table I summarizes the results of the tests for parallelism (the coefficients, except possibly for the intercept, are equal in the three groups), coincidence (all the coefficients, including the intercept, are equal) and for the three coefficients considered one-at-a-time using W and Φ for $A = I$ and $A = S$. (There are reasons for the apparent notational quirk: we have, for convenience, consistently [1,3,7-10] subscripted the τ values starting with 1, but it is more convenient — and traditional — to start the α values with 0. The differences between the α values and τ values will become clear below when we consider a specific example.)

Not shown are the corresponding results for pairwise comparisons between the groups, but for parallelism, coincidence, α_0 and α_1 , G_1 differed from both G_2 and G_3 , but G_2 and G_3 were not significantly different. The values of the α values in the three groups when integer orthogonal scores are used in the time design matrix and $A = S$ are shown in Table II.

Discussion and Conclusion

It is seen from Table I that the multivariate tests (those for parallelism and coincidence) are, as noted earlier, invariant with respect to the choice of W or Φ : the p -values are exactly the same for a given choice of A . It is also clear that the particular choice of A makes but very little difference in this example. One can only con-

TABLE II

ORTHOGONAL POLYNOMIAL REGRESSION COEFFICIENTS IN THE THREE GROUPS WHEN INTEGER ORTHOGONAL SCORES ARE USED IN THE TIME DESIGN MATRIX

	α_0	α_1	α_2
G_1	135.180	3.045	0.084
G_2	126.594	2.664	0.057
G_3	123.627	2.541	0.039

clude that the growth curves in the three groups are not the same. However, when the τ values are considered individually no significant differences are found. This counter-intuitive set of results may be avoided by using the α values instead of the τ values: it is seen that α_0 and α_1 differ among the groups. This same phenomenon was noted in [3] for an entirely different data set. It has been our experience that the use of OPRCs generally leads to consistent results in such contexts whereas using the τ values can produce paradoxical conclusions.

In addition, a number of practical advantages of OPRCs were enumerated in [3]. One that has only recently come to light involves the use of high degree polynomials as recommended by [11] and used in [12]. Even with 80 bit precision (about 19 significant digits), as is employed by *GAUSS* [13], computational accuracy can be a problem when \mathbf{W} is used. Rounding errors can be substantial in the inversion of $\mathbf{W}'\mathbf{W}$ when \mathbf{W} is of greater dimension than 7×7 . There is a two-pronged reason for this problem [5]: multicollinearity (the columns of \mathbf{W} are highly correlated) and the fact that the elements of \mathbf{W} will often differ substantially in magnitude. Draper and Smith [14] give an explicit example of how the latter situation can cause serious problems in regression computations. The use of orthogonal polynomials represents one convenient and effective way to circumvent these problems.

On the other hand, it is often argued [15], that the τ values have a 'more convenient' interpretation than the α values. This view, however prevalent, is not universal. Our opinions were detailed previously [3]; and we are not completely alone. The claim that, 'interpretation is simpler and clearer with orthogonal polynomials', is forwarded by Kerlinger et al. [16]. To see what is involved here, note that in the context of the example considered in this paper the integer-valued orthogonal time design matrix for $T = 6$ time points and $D = 2$ is, either from tables [17] or our program [3],

$$\Phi = \begin{bmatrix} 1 & -5 & 5 \\ 1 & -3 & -1 \\ 1 & -1 & -4 \\ 1 & 1 & -4 \\ 1 & 3 & -1 \\ 1 & 5 & 5 \end{bmatrix} \quad (8)$$

the numbers in the second and third columns being the values of the orthogonal polynomials

$$\phi_1(t) = 2t \text{ and } \phi_2(t) = \frac{3}{2}t^2 - \frac{35}{8} \quad (9)$$

at the centered time points $t = -5/2, -3/2, -1/2, 1/2, 3/2$ and $5/2$, respectively. The AGC in a particular group is represented in terms of these quantities by

$$x(t) = \alpha_0 + \alpha_1\phi_1(t) + \alpha_2\phi_2(t) \quad (10)$$

i.e., by

$$x(t) = (\alpha_0 - \frac{35}{8} \alpha_2) + 2\alpha_1 t + \frac{3}{2} \alpha_2 t^2 \quad (11)$$

We wish to compare (10) and/or (11) with the form of the AGC involving the τ values, viz.,

$$x(t) = \tau_1 + \tau_2 t + \tau_3 t^2 \quad (12)$$

Begin by contrasting α_0 and τ_1 . The intercept, τ_1 , is the value of $x(t)$ when $t = 0$. In our example, the range of observations is from $t = 7$ to $t = 12$ years of age. We have no observations near $t = 0$ and do not even know whether or not a quadratic equation is appropriate in that region. On the other hand, α_0 is the average of the values of the fitted curve at the times of observation. It can be used to answer the question of whether or not the groups have the same average values over time.

Interpretation of the higher order coefficients in this case involves the important aspects of growth velocity and acceleration [8]. Using the dot notation for differentiation, for velocity the choice is between

$$\dot{x}(t) = 2\alpha_1 + 3\alpha_2 t \text{ and } \dot{x}(t) = \tau_2 + 2\tau_3 t \quad (13)$$

while for acceleration it is between

$$\ddot{x}(t) = 3\alpha_2 \text{ and } \ddot{x}(t) = 2\tau_3 \quad (14)$$

The expressions for these quantities involving the τ values may be somewhat simpler than those involving the α values, but this would seem a small price to pay for avoiding the possibility that the conclusions 'are based entirely on the vagaries of roundoff error' [14]. In any case, as pointed out in Ref. 3, nothing is lost by the use of OPRCs: one can always recover the τ values from the α values by using $\tau = \mathbf{W}^- \phi \alpha$ where \mathbf{W}^- denotes a generalized inverse [18] of \mathbf{W} .

Continuing with our example, the growth behavior in G_1 , the high SES Ladino children, is clearly different from that of the other two lower SES groups. Examination of the program output for pairwise comparisons (i.e., two groups compared at a time) reveals that α_0 differs significantly between G_1 and each of the other groups. That is, the most privileged group is, when averaged over the entire time range of observations, largest. The lack of parallelism and the differences in α_1 between G_1 and the others demonstrates that the individuals in this group are also growing at a faster rate (velocity). Finally, the program demonstrates that a quadratic equation is a good fit for the total data set. This indicates that, overall, the growth of these children is accelerating; the magnitude of acceleration is small (about 0.18 cm/year²), but significant. Thus, all three groups show a slight upturn in the rate of growth over the period 7–12 years, perhaps in anticipation of the pubertal growth spurt. It is interesting to note that the high SES children have the highest α_2 value, though it is not significantly greater than that of the other two groups. This does, however, suggest that the high SES children, who are taller and growing faster than

those in the other groups, may also be initiating the pubertal spurt earlier. The possibility of such a trend is consistent with the general observation that well-nourished and healthy children enter into puberty earlier than others.

In conclusion, this program that performs the PR analysis, using OPRCs, is a straightforward and appropriate tool for the analysis and comparison of longitudinal samples. In that all of the major statistical software packages lack procedures specifically addressing this problem, this PC program should be a useful tool to many biomedical investigators.

Acknowledgement

This research was supported by grant DE08730 from the National Institute of Dental Research.

Appendix. Computer Implementation

This program can be obtained on a 5.25" double density floppy disk by sending \$10 to defray the cost of handling and licensing fees. The program will run on an IBM-PC/XT or AT compatible computer. The computer must be equipped with a numerical coprocessor from the 8087 family and 640K of memory. The computer must be configured so that at least 430K of memory is available, i.e., not tied up with memory resident programs such as Windows. EGA or VGA graphics capability is required to display the color graphics. No additional software is required (other than what one would normally use to enter a data set); run-time modules are supplied with the program so that no compiler or interpreter is necessary. The program, written in GAUSS, version 2.0, revision 20, requires no additional installation or modification, and is run with a single command. When requesting the program, address inquiries to the corresponding author and make checks payable to Baylor College of Dentistry.

References

- 1 Ten Have TR, Kowalski CJ and Schneiderman ED: Two SAS programs for performing multigroup longitudinal analyses, *Am J Phys Anthropol* (1992) in press.
- 2 Potthoff RF and Roy SN: A generalized multivariate analysis of variance model useful especially for growth curve problems, *Biometrika*, 51 (1964) 313–326.
- 3 Ten Have TR, Kowalski CJ and Schneiderman ED: A PC program for obtaining orthogonal polynomial regression coefficients for use in longitudinal data analysis, *Am J Hum Biol* (1992) in press.
- 4 Timm NH: *Multivariate Analysis with Applications in Education and Psychology*, Wadsworth, Belmont, CA, 1975.
- 5 Neter J, Wasserman W and Kutner MH: *Applied Linear Statistical Models*, 2nd Edn., Irwin, Homewood, IL, 1985.
- 6 Bogin B, Sullivan T, Hauspie R and MacVean RB: Longitudinal growth in height, weight, and bone age of Guatemalan Ladino and Indian School Children, *Am J Hum Biol*, 1 (1989) 103–113.
- 7 Schneiderman ED, Kowalski CJ and Ten Have TR: A GAUSS program for computing an index of tracking from longitudinal observations, *Am J Hum Biol*, 2 (1990) 475–490.
- 8 Schneiderman ED and Kowalski CJ: Implementation of Hills' growth curve analysis for unequal-time intervals using GAUSS, *Am J Hum Biol*, 1 (1989) 31–42.

- 9 Schneiderman ED and Kowalski CJ: Implementation of Rao's one-sample polynomial growth curve model using SAS, *Am J Phys Anthropol*, 67 (1985) 323–333.
- 10 Ten Have TR, Kowalski CJ and Schneiderman ED: PC program for analyzing one sample longitudinal data sets which satisfy the two-stage polynomial growth curve model, *Am J Hum Biol*, 3 (1991) 269–279.
- 11 Joossens JV and Brems-Heynes E: High power polynomial regression for the study of distance, velocity and acceleration of growth, *Growth*, 39 (1975) 535–551.
- 12 Zerbe GO: A new nonparametric technique for constructing percentiles and normal ranges for growth curves determined from longitudinal data, *Growth*, 43 (1979) 263–272.
- 13 Aptech: *The GAUSS System Version 2.0*, Aptech Systems, Kent, WA, 1988.
- 14 Draper N and Smith H: *Applied Regression Analysis*, Wiley, New York, 1966.
- 15 Goldstein H: *The Design and Analysis of Longitudinal Studies*, Academic Press, New York, 1979.
- 16 Kerlinger EN and Pedhazur EJ: *Multiple Regression in Behavioral Research*, Holt, Rinehart and Winston, New York, 1973.
- 17 Pearson ES and Hartley HO: *Biometrika Tables for Statisticians*, Vol. 1, 3rd edn, Cambridge University Press, London, 1966.
- 18 Graybill FA: *Introduction to Matrices with Applications in Statistics*, Wadsworth, Belmont, CA, 1969.