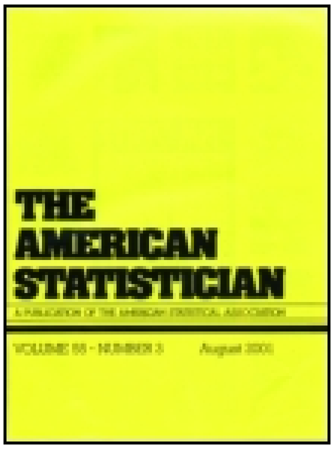


The American Statistician
Publication details, including instructions for authors and subscription information:
http://amstat.tandfonline.com/loi/utas20

Two Graphical Techniques Useful in Detecting Correlation Structure in Repeated Measures Data

Kathryn S. Dawson [a] , Chris Gennings [a] & Walter H. Carter [a]
[a] Department of Biostatistics Medical College of Virginia , Virginia Commonwealth University , Box 32, Richmond , VA , 23298-0032 , USA
Published online: 17 Feb 2012.

Statistical Computing and Graphics

Two Graphical Techniques Useful in Detecting Correlation Structure in Repeated Measures Data

Kathryn S. DAWSON, Chris GENNINGS, and Walter H. CARTER

Analysis of repeated measures data using a mixed model includes specifying a form for the covariance matrix of the within-subject observations. This reduction in the number of estimated parameters from the unspecified structure may improve the efficiency of inferences made. An implementation of this technique has been incorporated in the MIXED procedure of the SAS® statistical package, and includes a wide range of options for the structure of the covariance matrix. It is demonstrated that draftman's display plots and/or plots in a coordinate system with parallel axes can aid in visualizing the dispersion structure.

KEY WORDS: Draftman's display; Mixed modes; Parallel axes.

1. INTRODUCTION

Repeated measures data in which a response is measured on each observational unit on more than one occasion are often encountered in medical and social sciences applications. In particular, in longitudinal studies the response is observed over time. Multivariate modeling techniques have been used successfully in the analysis of such data. Here the p vector of observations for the ith subject y_i is commonly assumed to be normally distributed with $p \times 1$ mean vector and $p \times p$ dispersion matrix. If no specific structure for the dispersion matrix, such as compound symmetry, is assumed, and if the number of time points is large, the number of parameters in this dispersion matrix increases and may be poorly estimated (Laird and Ware 1982).

The mixed model approach to analyzing repeated measures data has been described by several authors, including Harville (1977) and Laird and Ware (1982). One advantage of this approach is the ability to model the covariance or correlation structure for the repeated measurements on a subject where positive correlation is expected. This can result in a reduction in the number of parameters in the dispersion matrix to be estimated, and can therefore improve the efficiency of inferences and estimates made. This is especially true when the data are unbalanced and the number of time points is large relative to the number of observations (Ware 1985).

Assuming no knowledge about the covariance matrix, the analyst could attempt to fit a mixed model with an unspecified structure, and use the resulting estimated covariance matrix to suggest a better defined structure. The appropriateness of the more specified model can be evaluated. If, however, the number of subjects and/or time points is large, algorithms used to attain the initial unspecified estimated structure may not converge. This paper demonstrates how, without first fitting a model, two different graphical techniques can aid the user in determining an initial form of the covariance matrix. It is not the purpose of this paper to compare the two graphical procedures, but rather to point out that they both provide useful insights into the structure of the data.

The two methods considered in the following are the draftman's display and parallel axis plots. The well-known draftman's display is a two-dimensional array of scatter plots $X_i \times X_j, i = 1, 2, \ldots, p, j = 1, 2, \ldots, p, i \neq j$. A p-dimensional data point $(x_1, x_2, \ldots, x_p)$ can be displayed as a series of points (x_i, x_j) $i \neq j$ each plotted in the appropriate $X_i \times X_j$ coordinate system. This technique has been shown to be helpful in detecting clustering and outliers (Chambers, Cleveland, Kleiner, and Tukey 1983).

Plotting in a parallel axis system as discussed by Inselberg (1985) involves using a set of connected line segments to plot a p-dimensional point. p parallel axes are drawn one unit apart corresponding to the variables $X_1, X_2, \ldots, X_p$. The point $(x_1, x_2, \ldots, x_p)$ is plotted by drawing lines from the values x_i on the X_i axis to the value x_{i+1} on the ad-

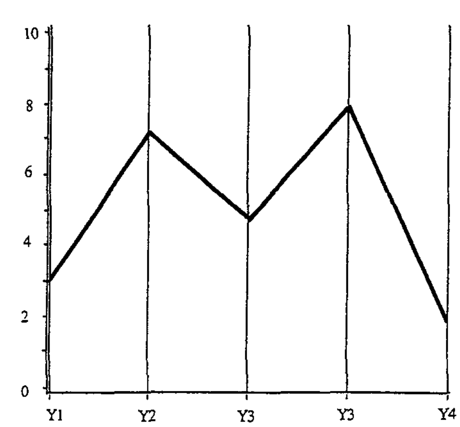


Figure 1. Five-Dimensional Point (3, 7, 5, 8, 2) Plotted in a Parallel Coordinate System.

Kathryn S. Dawson is Assistant Professor, Chris Gennings is Associate Professor, and Walter H. Carter is Professor, Department of Biostatistics, Medical College of Virginia, Virginia Commonwealth University, Box 32, Richmond, VA 23298-0032 (E-mail: dawson@gems.vcu.edu).

Table 1. Covariance Structures

Unequal Diagonal Elements

Unstructured

$$\begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_{22}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_{33}^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44}^2 \end{bmatrix}$$

Banded Main Diagonal

$$\begin{bmatrix} \sigma_{11}^2 & 0 & 0 & 0 \\ 0 & \sigma_{22}^2 & 0 & 0 \\ 0 & 0 & \sigma_{33}^2 & 0 \\ 0 & 0 & 0 & \sigma_{44}^2 \end{bmatrix}$$

Equal Diagonal Elements

Simple

$$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}$$

Compound Symmetry

$$\begin{bmatrix} \sigma^2+\sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma^2+\sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma^2+\sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma^2+\sigma_1^2 \end{bmatrix}$$

First-Order Autoregressive

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Toeplitz

$$\sigma^2 \begin{bmatrix} 1 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & 1 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & 1 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & 1 \end{bmatrix}$$

Spatial

$$\sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} \\ \rho^{d_{12}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} \\ \rho^{d_{13}} & \rho^{d_{23}} & 1 & \rho^{d_{34}} \\ \rho^{d_{14}} & \rho^{d_{24}} & \rho^{d_{34}} & 1 \end{bmatrix}$$

where $d_{ij} = |t_i - t_j|$ and t_i, t_j are the times associated with the ith and jth element of the correlation matrix

jacent $X_{i+1}, i = 1, 2, \ldots, p-1$ axis. The point (3, 7, 5, 8, 2) is shown in Figure 1. A complete set of data can be plotted by drawing all of the connected line segments on a common set of axes in a single graph (see Figure 4). Gennings, Dawson, Carter, and Myers (1990) developed analytical results useful in interpreting higher dimensional parallel axis plots of a polynomial model. Wegman (1990) demonstrated the usefulness of this plotting technique in observing structure and clustering in higher dimensional data. With respect to repeated measures data Weiss and Lazaro (1992) showed, by noting trends and patterns in a parallel axis plot of the observed residuals, that the quality of the fitted model can be assessed. In addition, unusual observations can be detected. Plots in this system can also be useful in noting distributional properties of the multivariate data (Wilkinson 1992).

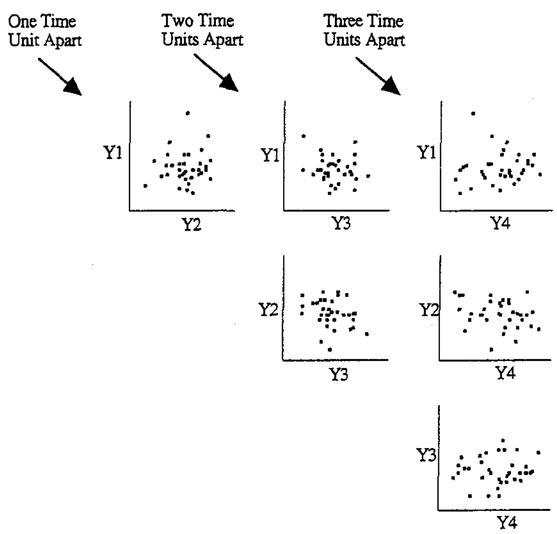


Figure 2. Ordering of Scatterplots in a Draftman's Display to View Correlation Structure.

In Section 2 a repeated measures model is defined. A centering and scaling technique is described in Section 3 that retains certain distributional properties assumed in the statement of the model. By plotting these transformed data a visualization of the dispersion structure can be made without first fitting the model. This information can then be used to specify an actual form for the covariance matrix.

2. A LINEAR REPEATED MEASURES MODEL WITH FIXED EFFECTS BASED ON THE MIXED MODEL

Assume there are t treatment groups with n_i subjects in each. Although it is also assumed that the experiment was designed to take measurements at p well-defined and fixed time points, the actual observed number of observations for each individual is $p_{ij}, 0 \le p_{ij} \le p, j = 1, 2, \ldots, n_i; i = 1, 2, \ldots, t$. A mixed linear model is given as follows:

$$\mathbf{Y}_{ij} = \mathbf{X}_{ij}\boldsymbol{\beta} + \mathbf{Z}_{ij}\nu_{ij} + \zeta_{ij}$$

where

$\mathbf{Y}_{ij}$ = $p_{ij} \times 1$ vector of responses over time for the jth subject in the ith treatment group,

$\mathbf{X}_{ij}$ = $p_{ij} \times q$ design matrix associated with the fixed effects for the jth subject in the ith treatment group,

$\boldsymbol{\beta}$ = $q \times 1$ vector of unknown fixed effect parameters,

$\mathbf{Z}_{ij}$ = $p_{ij} \times 1$ vector of 1s associated with the random effects for the jth subject in the ith treatment group,

γ^2 = unknown variance,

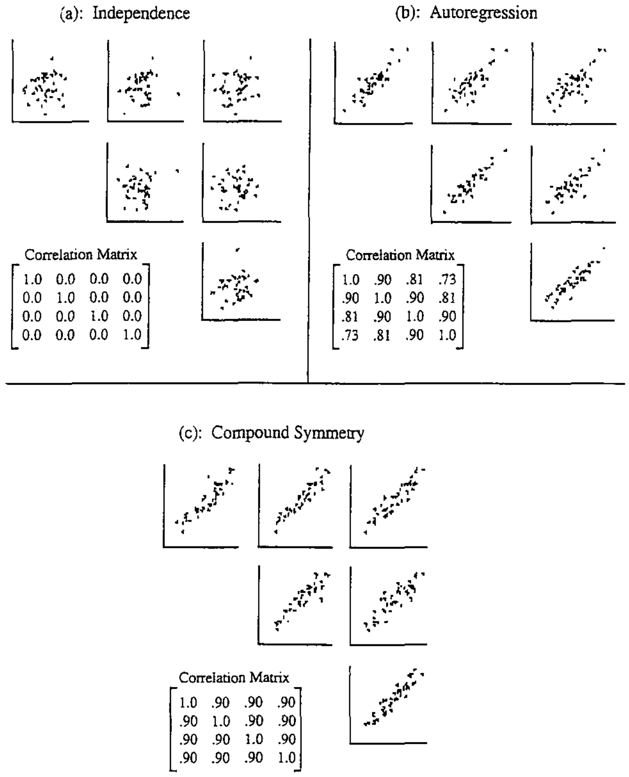


Figure 3. Simulated Four-Dimensional Normal Data with Various Correlation Structure Plotted in a Draftman's Display.

ν_{ij} = scalar random effect associated with the jth subject in the ith treatment group assumed to be independent and identically distributed (iid) $N(0, \gamma^2)$,
$\mathbf{H}$ = $p \times p$ unknown covariance matrix,
$\mathbf{H}_{ij}$ = $p_{ij} \times p_{ij}$ submatrix of $\mathbf{H}$ corresponding to the nonmissing times for the jth subject in the ith treatment group,
ζ_{ij} = $p_{ij} \times 1$ vector of unobserved random errors assumed to be iid $N_{p_{ij}}(\mathbf{0}, \mathbf{H}_{ij})$ and cov $(\zeta_{ijk}, \nu_{ij}) = 0, k = 1, 2, \ldots, p_{ij}$.

It follows that $\mathbf{Y}_{ij}$ is $N_{p_{ij}}(\mathbf{X}_{ij}\boldsymbol{\beta}, \gamma^2 \mathbf{Z}_{ij}\mathbf{Z}'_{ij} + \mathbf{H}_{ij})$.

The model can be simplified by first noting that $\gamma^2 \mathbf{Z}_{ij}\mathbf{Z}'_{ij} = \gamma^2 \mathbf{J}_{p_{ij}}$ where $\mathbf{J}_{p_{ij}}$ is a $p_{ij} \times p_{ij}$ matrix of 1s. Then, by letting $\varepsilon_{ij} = \nu_{ij}\mathbf{Z}_{ij} + \zeta_{ij}$, a repeated measures model with fixed effects can be written as

$$\mathbf{Y}_{ij} = \mathbf{X}_{ij}\boldsymbol{\beta} + \varepsilon_{ij} \qquad j = 1, 2, \ldots, n_i; i = 1, 2, \ldots, t$$

where

Σ = $p \times p$ unknown within subject covariance matrix,
$\boldsymbol{\Sigma}_{ij}$ = $p_{ij} \times p_{ij}$ submatrix of Σ corresponding to the nonmissing times for the ith subject in the jth treatment group,
ε_{ij} = $p_{ij} \times 1$ vector of unobserved random errors assumed to be iid $N_{p_{ij}}(\mathbf{0}, \boldsymbol{\Sigma}_{ij})$.

Furthermore, $\mathbf{Y}_{ij}$ is distributed as $N_{p_{ij}}(\mathbf{X}_{ij}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_{ij})$ where $\boldsymbol{\Sigma}_{ij} = \mathbf{H}_{ij} + \gamma^2 \mathbf{J}_{p_{ij}}$.

An implementation of this modeling technique is included in the MIXED procedure of the SAS® statistical package (SAS Institute Inc. 1992). Maximum likelihood or restricted maximum likelihood (REML) estimators for $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ are found based on algorithms developed by several authors including Harville (1977), Laird, Lange, and Stram (1987), and Jennrich and Schluchter (1986). In order to fit the model, however, the user must first choose a structure for $\boldsymbol{\Sigma}$ from a wide range of choices provided in the procedure (examples are provided in Table 1). Plots of certain centered and scaled observations can aid the user in determining an appropriate form for this dispersion structure. Draftman's plots and parallel axis plots are discussed in Sections 4 and 5. Those plots are constructed without first fitting the model, thereby avoiding the numerical problems associated with fitting a possibly misspecified or overparameterized model. The next section describes properties of the centered and scaled observations that are useful in a graphical examination of the data.

3. PROPERTIES OF THE CENTERED AND SCALED OBSERVATIONS

For the ith treatment group let $\boldsymbol{\Sigma}_i$ denote the $p \times p$ covariance matrix with kth diagonal element σ^2_{ik}. The associated $p \times p$ correlation matrix is given $\rho_i = \mathbf{D}_{\Sigma_i}\boldsymbol{\Sigma}_i\mathbf{D}_{\Sigma_i}$ where $\mathbf{D}_{\Sigma_i}$ denotes the $p \times p$ diagonal matrix with the jth diagonal element $(\sigma^2_{ik})^{-(1/2)}$. This correlation matrix is estimated by the $p \times p$ sample correlation matrix $\mathbf{r}_i$. The (st) element of $\mathbf{r}_i$ is given by

$$r_{ist} = \frac{\sum_{j=1}^{n_i} (Y_{ijs} - \bar{Y}_{i.s})(Y_{ijt} - \bar{Y}_{i.t})}{\left[\sum_{j=1}^{n_i} (Y_{ijs} - \bar{Y}_{i.s})^2 \sum_{j=1}^{n_i} (Y_{ijt} - \bar{Y}_{i.t})^2\right]^{1/2}}$$

where n_i is the number of subjects in the ith treatment group, summation is over the nonmissing pairs of observa-

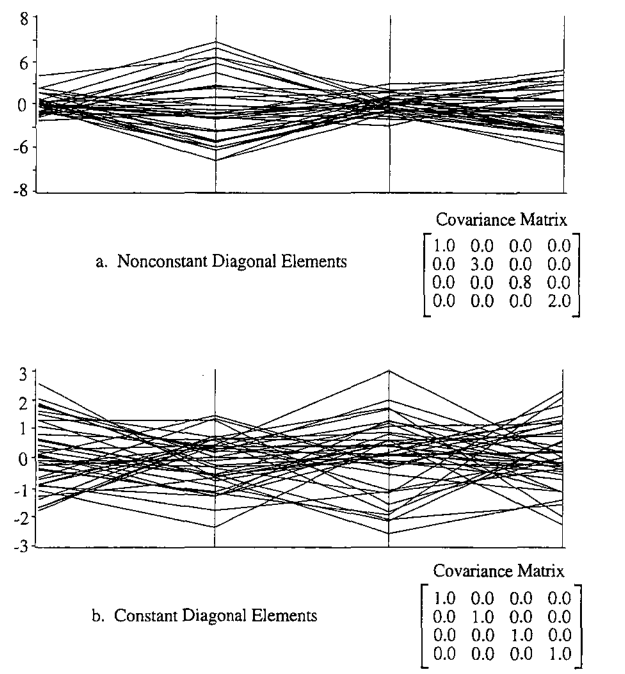


Figure 4. Comparing the Diagonal Elements of the Covariance Matrix in a Parallel Axis System Using Simulated Four-Dimensional Normal Data with Different Correlation Structures. (a) Nonconstant diagonal elements. (b) Constant diagonal elements.

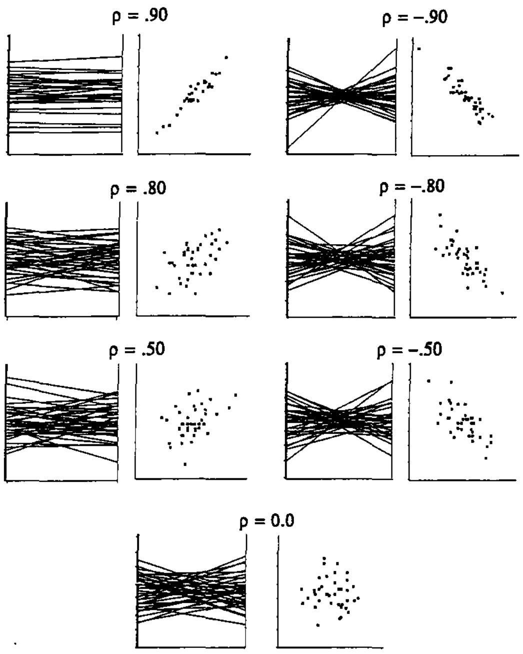


Figure 5. Simulated Bivariate Normal Data with Various Correlation Coefficients Plotted in Both a Parallel Axis System and a Scatterplot.

tions, and $\bar{Y}_{i.k} = (\sum_{j=1}^{n_{ik}} Y_{ijk}/n_{ik})$. Patterns observed in $\mathbf{r}_i$ can suggest the structure of the corresponding correlation matrix.

In order to graphically visualize correlation properties of the observed data it will be useful to first remove the variability in the data associated with differences in the means and variances over time. Let n_{ik} be the number of nonmissing observations for the ith treatment group at time k. For $i = 1, 2, \ldots, t; j = 1, 2, \ldots, n_{ik}$; and $k = 1, 2, \ldots, p$, let

$$\Delta_{ijk} = Y_{ijk} = \bar{Y}_{i.k}$$

and

$$\Delta^*_{ijk} = \frac{\Delta_{ijk}}{s_{i.k}}$$

where

$$s^2_{i.k} = \sum_{j=1}^{n_{ik}} \frac{\Delta^2_{ijk}}{n_{ik} - 1}.$$

Because $\bar{\Delta}^*_{i.k} = 0$ the sample correlation matrix based on the centered and scaled values is given by

$$r^*_{ist} = \frac{\sum_{j=1}^{n_i} (\Delta^*_{ijs})(\Delta^*_{ijt})}{\left[\sum_{j=1}^{n_i} (\Delta^*_{ijs})^2 \sum_{j=1}^{n_i} (\Delta^*_{ijt})^2\right]^{1/2}}$$

$$= \frac{\sum_{j=1}^{n_i} \left(\frac{Y_{ijs}-\bar{Y}_{i.s}}{s_{i.s}}\right)\left(\frac{Y_{ijt}-\bar{Y}_{i.t}}{s_{i.t}}\right)}{\left[\sum_{j=1}^{n_i} \left(\frac{Y_{ijs}-\bar{Y}_{i.s}}{s_{i.s}}\right)^2 \sum_{j=1}^{n_i} \left(\frac{Y_{ijt}-\bar{Y}_{i.t}}{s_{i.t}}\right)^2\right]^{1/2}} = r_{ist}.$$

Therefore the sample correlation structure is retained by the centered and scaled observations.

Other distributional properties associated with the univariate marginal distributions at each time point will be shown to be useful in the interpretation of the parallel axis plots of these centered and scaled observations. For ease of notation consider the case of no missing values so that $p_{ij} = p$ for all i, j. Let $\boldsymbol{\mu}_i$ denote the $p \times 1$ vector of mean responses for the ith treatment group. When it has been assumed for $j = 1, 2, \ldots, n_{ik}$ that the $p \times 1$ vectors $\mathbf{Y}_{ij}$ are iid $N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, it follows that at each time point, $k = 1, 2, \ldots, p, Y_{ijk}$ is iid $N(\mu_{ik}, \sigma^2_{ik})$. This implies $E(\Delta_{ijk}) = 0$ and $E(\Delta^*_{ijk}) = 0$. Furthermore, var $(\Delta_{ijk}) = (n_{ik} - 1/n_{ik})\sigma^2_{ik}$. Let $\xrightarrow{p}$ denote converge in probability and $\xrightarrow{d}$ convergence in distribution. Because $s^2_{i.k} \xrightarrow{p} \sigma^2_{ik}$ and $\bar{Y}_{i.k} \xrightarrow{p} \mu_{ik}$, it follows by Slutsky's Theorem that $\Delta_{ijk} = Y_{ijk} - \bar{Y}_{i.k} \xrightarrow{p} Y_{ijk} - \mu_{ik}$ and $\Delta^*_{ijk} = (Y_{ijk}/s_{i.k}) - (\bar{Y}_{i.k}/s_{i.k}) \xrightarrow{d} (Y_{ijk} - \mu_{ik}/\sigma_{ik})$. Hence when the number of observations in the ith treatment group at the kth time is large, Δ_{ijk} converges to $N(0, \sigma^2_{ik})$ and Δ^*_{ijk} converges to $N(0, 1)$. Although, in general, the number of these observations at each time point within each treatment group may be limited, these properties may still be useful in the examination of the plots of these centered and scaled data.

The following describes two graphical techniques that can be used to display distributional patterns in the centered and scaled data.

4. DRAFTMAN'S DISPLAY

Assuming there are p time points, $(p(p-1)/2)$ pairwise

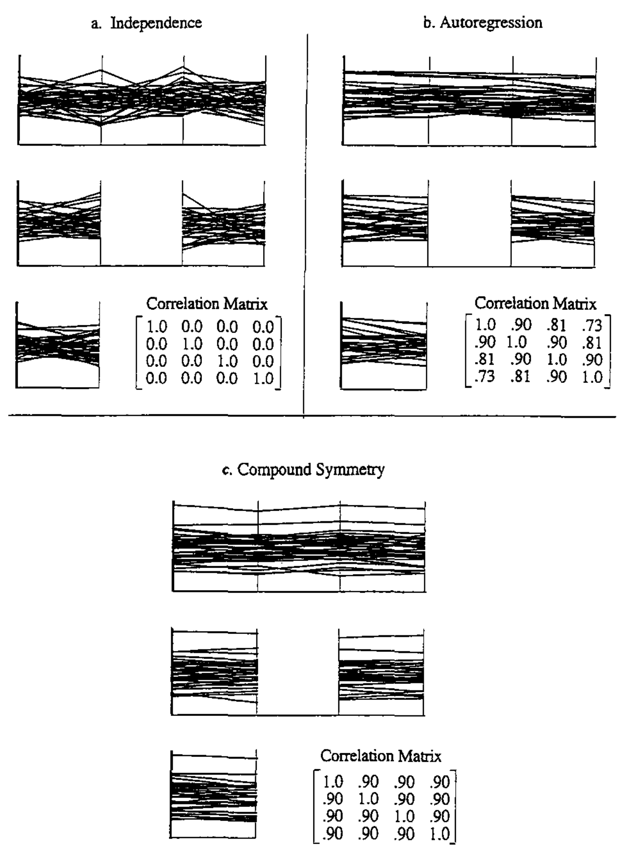


Figure 6. Simulated Four-Dimensional Normal Data with Various Correlation Structure Plotted in a Parallel Axis System.

Table 2. *Tumor Size (mm^3) Over Course of the Experiment (Koziol)*

	Day										
Mouse	7	11	12	13	14	15	17	18	19	20	21
1	35.3	157.1	122.5	217.6	340.3	379.0	556.6	661.3	634.8		
2	19.6	152.2	129.6	176.6	213.9	317.9	356.4	580.0	415.2	460.0	520.1
3	27.0	122.4	196.1	196.1	332.2	388.9	469.3	397.1	505.4	541.5	
4	55.0	95.0	205.9	205.9	270.0	307.3	405.1	726.0	950.4	661.5	798.6
5	24.6	68.8	135.3	196.0	340.2	340.4	507.3	767.2	820.0	937.5	
6	12.6	85.0	70.1	225.1	225.1	289.0	317.9	529.1	653.4	687.7	750.2
7	35.2	129.8	180.0	274.7	420.1	340.3	507.2	634.8	714.3	777.6	912.6
8	29.8	157.0	126.8	202.5	225.0	307.2	320.1				
9	70.0	129.7	196.0	205.8	375.7	419.1	421.2	573.4	701.8		
10	29.5	156.9	176.7	225.0	289.0	372.6	379.2	529.2	573.3	560.1	520.0
11	48.6	115.3	90.8	176.5	317.9	421.2	529.2	388.8	629.0		
12	66.7	289.0	215.6	268.8	388.8	487.4	551.3	767.1	677.6	846.4	634.9
13	24.5	143.7	115.0	90.7	194.3	559.6	629.3	573.3	540.0		
14	14.4	84.7	135.2	191.2	176.4	356.4	397.1	551.4	605.0	480.0	634.8
15	10.8	70.0	80.0	118.3	156.8	215.6	268.8	346.8	551.3	946.4	440.0
16	11.3	15.0	205.8	289.0	346.8	529.2	629.2	551.3	714.2	772.6	806.4
17	18.0	56.7	115.3	96.8	177.5	268.8	320.0	372.6	487.4	573.3	683.6
18	60.0	166.6	166.7	324.0	420.0	440.0	634.8	500.0	289.0	560.0	748.8
19	29.4	152.1	122.4	186.3	186.3	274.7	485.1	397.0			
20	41.1	186.2	176.6	274.6	361.0	379.1	440.0	415.2			
21	12.5	108.0	96.8	186.2	202.5	213.8	379.1	379.0	433.2	379.0	500.0
22	23.4	129.6	176.5	196.6	320.0	397.1	500.0	687.7	767.1	806.4	937.5
23	22.2	65.0	176.4	191.3	213.8	274.6	405.0	520.0	796.6	978.7	864.0
24	11.2	52.9	70.0	129.6	152.1	303.5	415.0	440.0	556.7	812.5	1014.0
25	66.6	147.0	260.1	420.0	460.0	653.4	806.4				
26	11.4	115.2	65.1	32.0	10.8	3.2	1.4	.0	.0	.0	.0
27	22.1	55.0	115.2	55.0	93.6	118.8	118.3	230.4	217.6	243.2	217.6
28	40.5	156.8	65.0	84.7	191.2	291.5	400.0				
29	32.0	44.6	108.9	258.8	247.5	405.0	372.6	388.0	451.3	580.0	573.3
30	10.0	118.3	166.6	176.4	186.2	340.2	361.0	556.6	556.6	268.8	346.8

plots of the $(\Delta^*_{ijr}, \Delta^*_{ijs}), r, s = 1, 2, \ldots, p, r \neq s$ values at all nonmissing distinct combinations of time points can be used to view the sample correlation structure. This array of plots given in Figure 2 can be laid out in an order similar to the correlation matrix. The plots along a diagonal, that is, $Y1 \times Y2, Y2 \times Y3$, and so on, correspond to the scaled data a fixed time unit apart. The upper left-hand plot corresponds to the first and second time points, and the upper right-hand corner plot corresponds to the first and last time points. It is possible to gain an indication of the correlation structure between the corresponding pairs of variables through the use of these plots.

Figure 3 shows simulated four-dimensional normally distributed datasets with differing correlation structures. In Figure 3a the within-subject observations are independent, which is indicated by the absence of a linear trend in any of the pairwise plots. Autoregression is indicated in Figure 3b by the decrease in correlation as the time interval between measurements increases, as well as by the consistency in the correlation for plots associated with fixed differences in time. The similarity of the pairwise plots of Figure 3c suggests compound symmetry.

5. PARALLEL AXIS DISPLAY

In a parallel axis display properties of the p univariate distributions can be visualized in addition to the correlation structure. In plots of Δ_{ijk} and Δ^*_{ijk} the patterns formed by the intersections of the line segments with each axis allow graphical assessment of the observed univariate sample distributions at each time point. In the plot of the centered and scaled Δ^*_{ijk} values the patterns of the line segments between the axis allow visualization of the observed correlation between the corresponding pairs of variables.

If the normality assumption is correct $E(\Delta_{ijk}) = 0$ and $var(\Delta_{ijk}) = (n_{ik} - 1/n_{ik})\sigma^2_{ik}$. This implies that when n_{ik} is constant, $k = 1, 2, \ldots, p$, a visual comparison of the $\sigma^2_{ik}, k = 1, 2, \ldots, p$ values can be made in a parallel axis plot of the Δ_{ijk} values by examining the distribution of the intersection points of the line segments with the associated parallel axis. In Figure 4a data simulated from a normal distribution with a nonconstant $\sigma^2_{ik}, k = 1, 2, 3, 4$ are plotted. The larger variability associated with the second and fourth measurements is indicated by the wider spread in the distribution of the intersection points with these axes. In contrast, in Figure 4b where data were simulated from a normal distribution with $\sigma^2_{ik} = 1, k = 1, 2, 3, 4$, constant variability is suggested. Of course, in cases when n_{ik} is not constant for all k the variability in the plots associated with the nonconstant sample sizes must be taken into account.

Similarly, because, when n_{ik} is large, the distribution of the centered and scaled Δ^*_{ijk} approximates $N(0, 1)$ in a plot of these values, the intersection points along each of the axes should be consistently centered at 0 and virtually all between -3 and 3 (Fig. 4b). Deviations in this pattern when n_{ik} is not large may not necessarily indicate nonnormality, but may warrant further examination.

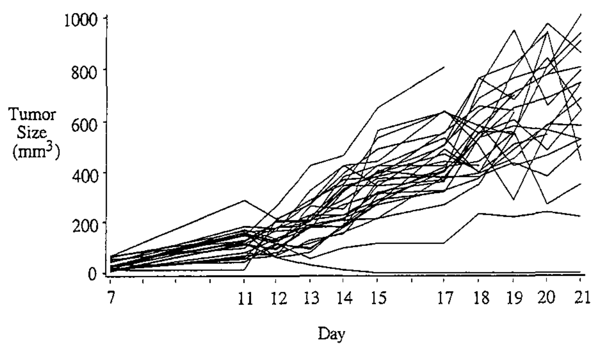


Figure 7. Observed Koziol Data Plotted Over Time.

In plots of Δ^*_{ijk} the patterns of the line segments between the axes are associated with the correlation structure. Positive correlation in a parallel axis plot is indicated by a pattern of parallel line segments between the axes. In addition, if the positively correlated data are jointly centered about 0, the slopes of the line segments will be close to 0. As the degree of positive correlation diminishes the lines will tend to intersect more often, and the variability in slopes will increase. In contrast, as negative correlation approaches -1.0 the line segments intersect in a smaller and smaller region. In fact, if there is perfect negative correlation, the lines all intersect at a single point (Inselberg 1985). As the degree of negative correlation diminishes the common intersection region widens. Figure 5a and b shows simulated data for various values of both positive and negative ρ, respectively.

In order to observe correlation patterns between all possible pairs of variables a set of parallel axis plots will be needed. Rather than considering $(p(p-1)/2)$ pairwise parallel plots, plotting several dimensions on a single set of axes is convenient. Wegman (1990) has shown that for p variables a minimal number of parallel plots to show all possible pairwise combinations is $(p+1/2)$. The permutation of the axes using this method results in an ordering of the axes that does not reflect the natural order of the measurements by time. A more convenient ordering of the axes for this type of data using $(p-1)$ plots is suggested. This is illustrated with the three plots shown in Figure 6a. Here the first parallel plot orders the variables naturally so that correlation between measurements one time unit apart can be observed. In addition, the user can simultaneously observe the centered and scaled observations through time. The second parallel plot orders the axes so that correlations between measurement taken two time units apart can be observed. Note in this plot that the line segments between the axes representing times 3 and 2 are not connected because this comparison is not considered in this plot. The third graph considers values taken three time units apart.

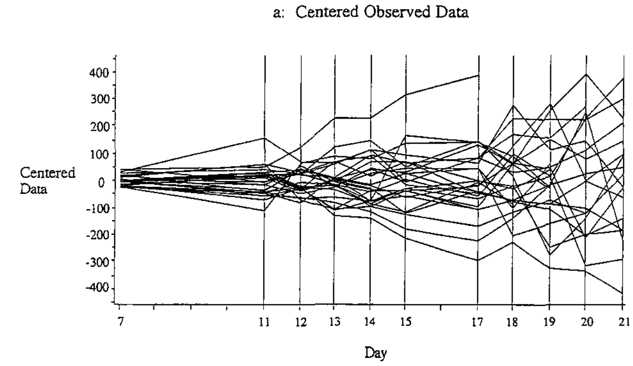
Figure 6 illustrates this technique for the datasets previously plotted in the draftman's display (Fig. 3). In Figure 6a there is no indication of positive correlation due to the absence of parallel lines between any of the axes. In addition, because the regions formed by the intersection of lines between the axes are wide, negative correlation is not indicated. Independence is therefore suggested in Figure 6a. A decrease in positive correlation associated with autoregression is indicated in Figure 6b by the weakening in the pattern of parallel lines as the differences in time between measurements increases. Compound symmetry is suggested in Figure 6c because the line patterns between all pairs of variables seem constant.

The following example illustrates the usefulness of both the parallel axis plots and the draftman's display at certain stages of a repeated measures analysis. In particular, these plots will be used, without first fitting a model, to suggest the structure of the dispersion matrix.

6. EXAMPLE

Koziol and Maxwell (1981) report a study that was conducted to test the efficacy of three therapies against colon carcinomas in mice. Thirty mice injected with mouse colon carcinoma cells were randomly divided into the three treatment groups. The size of the resulting tumor was recorded on 11 different occasions, days 7, 11, 12, 13, 14, 15, 17, 18, 19, 20, and 21. Nine mice died before the end of the experiment so that measurements at later time points for these mice are missing. The data appear in Table 2. A plot of the observed data (Fig. 7) indicates an increase in tumor size over time. The sixth mouse in group 3 is clearly unusual, and was eliminated from the remainder of this analysis.

At this point the analyst could examine the estimated dispersion matrix after fitting a linear repeated measures model with an unstructured covariance matrix. An attempt to perform such an analysis using PROC MIXED in SAS® failed to converge. The graphical techniques previously discussed can now be used to provide valuable information that will permit the analyst to proceed.



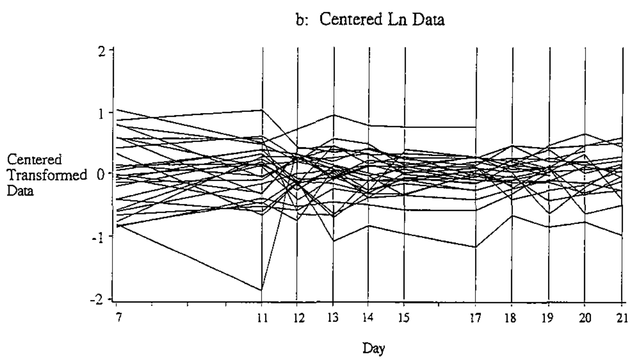


Figure 8. Centered Observed and Transformed Koziol Data Plotted in a Parallel Axis System. (a) Centered data. (b) Centered ln transformed data.

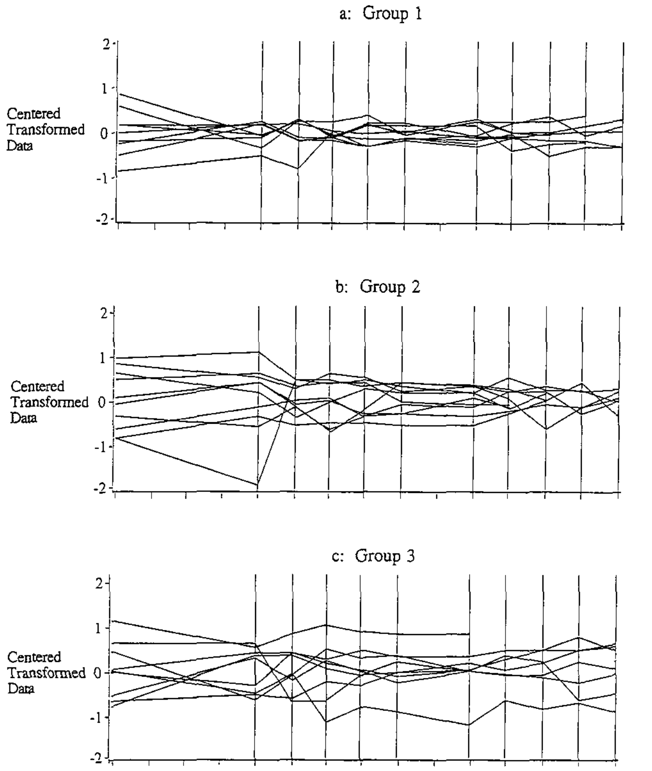


Figure 9. Center ln Transformed Koziol Data by Group. (a) Group 1. (b) Group 2. (c) Group 3.

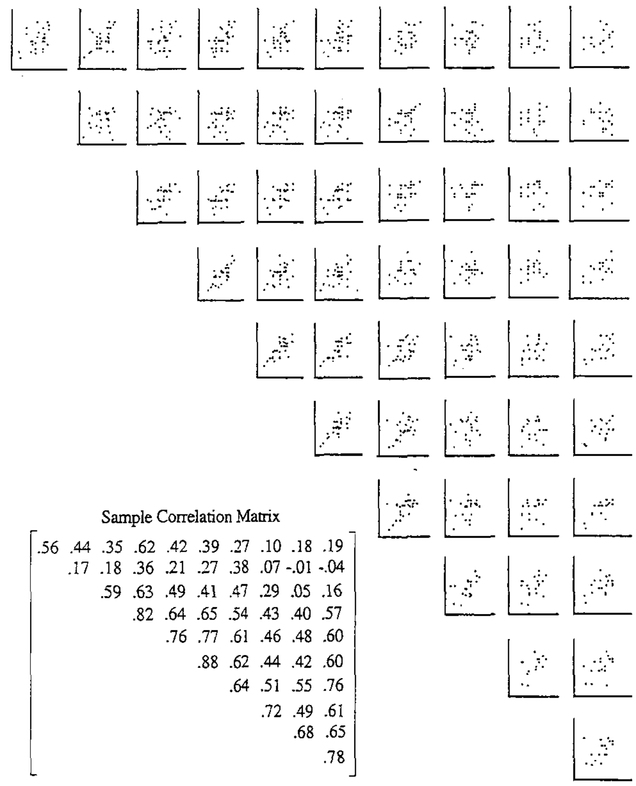


Figure 11. Centered and Scaled ln Transformed Koziol Data in a Draftman's Display with Sample Correlation Matrix.

In order to compare the diagonal elements of the covariance matrix graphically the centered data using the observed means for each of the three treatment groups at each of the 11 time points are plotted in Figure 8a. An increase in variability is indicated by the widening distribution of the intersections of the line segments with each axis. Because covariance structures associated with constant diagonal elements have fewer parameters to estimate, a transformation to stabilize this variability is suggested and was considered. A plot of the centered ln transformed data (Fig. 8b) suggests constant diagonal elements of the covariance matrix associated with these transformed data. Consequently, the ln transformation $Y_{ij}^* = \ln(Y_{ij} + 1), i = 1, 2, 3; j = 1, 2, \ldots, n_i$ was employed for the remainder of this analysis.

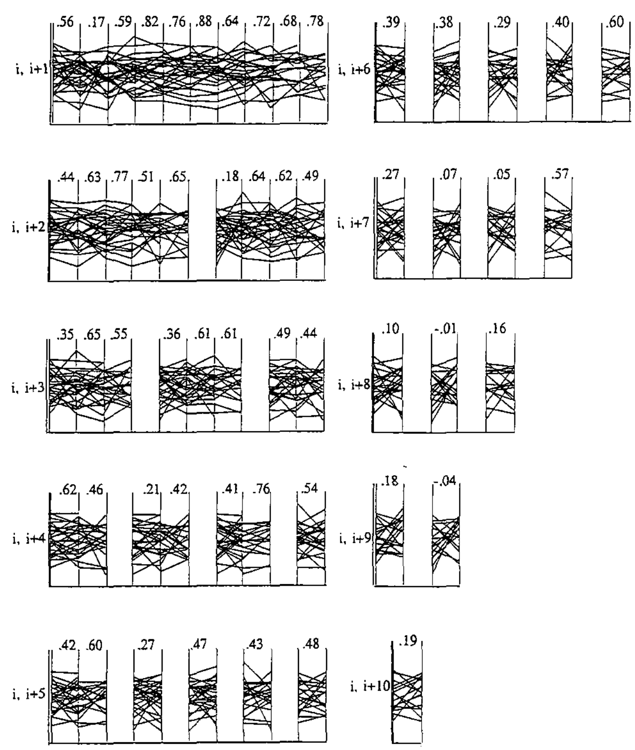


Figure 10. Centered and Scaled ln Transformed Koziol Data in a Parallel Axis System. The i,jth position of the correlation matrix is represented in each plot. The sample correlation is shown between each parallel axis.

To assess the consistency of the form of the covariance structure by treatment group the centered ln data by treatment group (Fig. 9) are plotted. Group 3 appears to exhibit a slightly larger degree of variability at each time point, and groups 1 and 2 show a possible decrease in variability

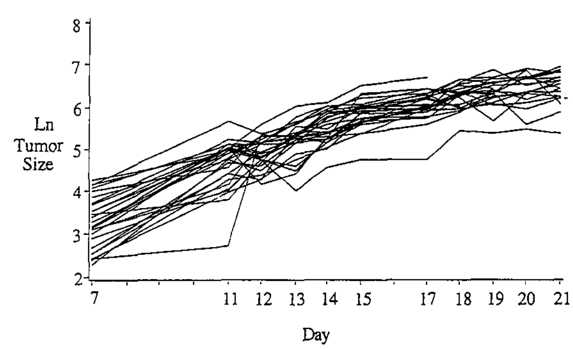


Figure 12. ln Transformed Koziol Data Plotted Over Time.

Table 3. Preliminary Analysis Results Using Proc Mixed in SAS®

Covariance parameters (Table 1)	Simple Parameter estimate	SE	Compound symmetry Parameter estimate	SE	Spatial Parameter estimate	SE
			(a) Covariance Structure			
σ^2	.1601	.0135	.0997	.0088	.1967	.0261
σ_1^2			.0680	.0219		
ρ					.7148	.0257
Criterion						
Akaike statistic*	−176.08		−136.92		−124.71	
Schwartz statistic*	−177.90		−140.56		−128.36	

Model parameters (Table 1)	Simple Parameter estimate	SE	Compound symmetry Parameter estimate	SE	Spatial Parameter estimate	SE
			(b) Model			
β_{01}	3.4501	.1195	3.4549	.1253	3.459	.1385
β_{02}	3.2761	.1193	3.2801	.1252	3.304	.1385
β_{03}	3.1217	.1256	3.1250	.1318	3.140	.1459
β_{11}	.3946	.0348	.3914	.0275	.3933	.0426
β_{12}	.4097	.0345	.4061	.0274	.4077	.0424
β_{13}	.4128	.0359	.4057	.0284	.4077	.0440
β_{21}	−.0123	.0023	−.0120	.0019	−.0123	.0030
β_{22}	−.0127	.0023	−.0124	.0018	−.0129	.0030
β_{23}	−.0130	.0023	−.0120	.0019	−.0126	.0030

	Simple	Compound symmetry	Spatial
		(c) Test Group of Differences	
p value	.0739	.1467	.1993

*Higher values are associated with models deemed best.

over time. Bartlett's test for the homogeneity of the covariance matrices, across groups (Morrison 1976, p. 252) can be applied once it has been determined that there is no indication of nonnormality. Maridia's tests for multivariate normality (Mardia 1980, p. 310) fail to detect a deviation from normality (skewness: $p = .1108$, kurtosis $p = .5959$), and significant heterogeneity is not indicated ($p = 1.000$). Homogeneity of the covariance matrix by group was therefore concluded.

Because it has been concluded that the diagonal elements of the covariance matrix are equal, further refinement of the dispersion structure can be made by considering plots of the centered and scaled data. Note, first, that the time intervals between consecutive measurements are not constant, that is, $|t_1 - t_2|$ =four days and $|t_2 - t_3|$ =one day. In order to view and compare correlations between successive measurements in the parallel axis plots, however, the distance between the axes should remain constant. Similarly, in the draftman's display the pairwise plots correspond to elements in the covariance structure. Hence this display need not reflect irregularly spaced time intervals. Using the parallel axis system (Fig. 10) some positive correlation is indicated, especially in the top left plot associated with the $(i, i+1), i = 1, 2, \ldots, 10$ positions of the sample correlation matrix. This indication of positive correlation is not as apparent in the plots associated with pairs of responses taken several days apart, that is, those in the $(i, i+k), k = 7, 8, 9$ positions of the sample correlation matrix. In the draftman's display (Fig. 11) the similarity of the scatter plots in the top rows suggests compound symmetry. Along each of the remaining rows a weakening degree of positive correlation is suggested. By examining the scatter plots along the diagonals the correlation between measurements at constant differences in time appears equal. Both graphical presentations and the sample correlation matrix shown in Figure 11 indicate that compound symmetry or spatial can be considered as possible forms for the covariance structure in the model fitting procedure. Note that because the time interval between successive measurements was not constant, first-order autoregression is not appropriate.

A plot of the observed ln transformed data indicates a possible curvilinear increase over time (Fig. 12). This implies that a linear repeated measures model with a quadratic term may be appropriate. The following defines the initial model considered for the ln data:

$$Y_{ij}^* = \beta_{0i} + \beta_{1i}t + \beta_{2i}t^2 + \varepsilon_{ij} \qquad i = 1, 2, 3$$

$$j = 1, 2, \ldots, n_i \quad t = 0, 1, \ldots, 10.$$

Table 3 summarizes the results of fitting this model using SAS® Proc Mixed and assuming compound symmetry and spatial correlation structures. A simple model assuming independence of the observations within a subject was also run for comparison. In all three cases the iterative procedure successfully converged to parameter estimates. A test comparing the treatment groups, however, results in differing p values, with the simple case actually suggesting a possible difference between the groups. In order to compare the models the Akaike Information Criterion and Schwartz's Bayesian Criteria can be examined. These statistics can be used as guidelines when comparing models with the same fixed effects, but with varying covariance structure (SAS Institute Inc. 1992, p. 326). Larger values of these statistics suggest more appropriate covariance structure. Thus at this stage of the analysis the spatial structure appears to be the most appropriate for this dataset.

7. SUMMARY

Although the mixed modeling approach to analyzing repeated measures data has often been discussed, its implementation in PROC MIXED of the SAS® statistical package has made it readily available for use. This procedure allows the analyst to specify a well-defined form for the covariance structure. In some cases, especially if the data are unbalanced, the procedure will fail to converge if a specific form for this matrix is not chosen. This paper has demonstrated that graphical procedures can be useful procedure in suggesting an appropriate form for the dispersion matrix for a repeated measure model with fixed effects. In particular, a centering and scaling technique was chosen so that properties of the covariance structure were retained in the scaled data. By plotting these scaled data properties of the covariance matrix can be visualized.

[Received May 1993. Revised January 1997.]

REFERENCES

Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A. (1983), *Graphical Methods for Data Analysis*, Belmont, CA: Wadsworth International Group.

Gennings, C., Dawson, K. S., Carter, W. H., Jr., and Myers, R. H. (1990), "Interpreting Plots of a Multidimensional Dose-Response Surface in a Parallel Coordinate System," *Biometrics*, 46, 719–735.

Harville, D. A. (1977), "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," *Journal of the American Statistical Association*, 72, 320–340.

Inselberg, A. (1985), "The Plane with Parallel Coordinates," *The Visual Computer*, 1, 69–91.

Jennrich, R. I., and Schluchter, M. D. (1986), "Unbalanced Repeated-Measures Models with Structured Covariance Matrices," *Biometrics*, 42, 805–820.

Kotz, S., and Johnson, N. L. (eds.) (1982), *Encyclopedia of Statistical Sciences*, New York: Wiley.

Koziol, J. A., Maxwell, D. A., Fukushima, M., Colmerauer, M. E. M., and Pilch, Y. H. (1981), "A Distribution-Free Test for Tumor-Growth Curve Analyses with Application to an Animal Tumor Immunotherapy Experiment," *Biometrics*, 37, 383–390.

Laird, N., Lange, N., and Stram, D. (1987), "Maximum Likelihood Computations with Repeated Measures: Application of the EM Algorithm," *Journal of the American Statistical Association*, 82, 97–105.

Laird, N. M., and Ware, J. H. (1982), "Random-Effects Models for Longitudinal Data," *Biometrics*, 38, 963–974.

Mardia, K. V. (1980), "Test of Univariate and Multivariate Normality," in *Handbook of Statistics, Vol. 1*, ed. P. R. Krishnaiah, New York: North-Holland, pp. 279–320.

Morrison, D. F. (1976), *Multivariate Statistical Methods*, New York: McGraw-Hill.

SAS Institute Inc. (1992), SAS® Technical Reports, p. 229, *SAS/STAT® Software: Changes and Enhancements, Release 6.07*, Cary, NC: SAS Institute Inc., 620 pp.

Ware, J. H. (1985), "Linear Models for the Analysis of Longitudinal Studies," *The American Statistician*, 39, 95–101.

Wegman, E. J. (1990), "Hyperdimensional Data Analysis Using Parallel Coordinates," *Journal of the American Statistical Association*, 85, 664–675.

Weiss, R. E., and Lazaro, C. G. (1992), "Residual Plots for Repeated Measures," *Statistics in Medicine*, 11, 115–124.

Wilkinson, L. (1992), "Graphical Displays," *Statistical Methods in Medical Research*, 1, 3–25.