PSYCHOMETRIKA—VOL. 69, NO. 1, 65–79
MARCH 2004

A MULTIVARIATE REDUCED-RANK GROWTH CURVE MODEL WITH UNBALANCED DATA

HEUNGSUN HWANG

HEC MONTREAL

YOSHIO TAKANE

MCGILL UNIVERSITY

A multivariate reduced-rank growth curve model is proposed that extends the univariate reduced-rank growth curve model to the multivariate case, in which several response variables are measured over multiple time points. The proposed model allows us to investigate the relationships among a number of response variables in a more parsimonious way than the traditional growth curve model. In addition, the method is more flexible than the traditional growth curve model. For example, response variables do not have to be measured at the same time points, nor the same number of time points. It is also possible to apply various kinds of basis function matrices with different ranks across response variables. It is not necessary to specify an entire set of basis functions in advance. Examples are given for illustration.

Key words: Multivariate growth curve models, reduced-rank restriction, longitudinal data.

1. Introduction

In a variety of areas, observations are taken over multiple time points on a particular characteristic, often called a response variable, to investigate temporal patterns of change on the characteristic. For instance, students may be asked to take a standardized test repeatedly over several months. Satisfaction of customers toward a particular brand may be tracked down every quarter. Effects of a certain drug on animals, blood sugar concentrations may be measured over time.

Data of this type are usually analyzed by the growth curve model, initiated by Potthoff and Roy (1964), and extensively studied by numerous authors, including Khatri (1966), Grizzle and Allen (1969), and Rao (1965). (Refer to von Rosen (1991) for a nice review on the growth curve model.) The basic idea of the growth curve model is to introduce some known functions, so-called basis functions (e.g., polynomial functions), so as to capture patterns of change for time-dependent measurements. The traditional growth curve model was designed for the situations where individuals are measured on a single response variable. Reinsel (1982) extended the univariate growth curve model to the multivariate case, where several response variables are measured over multiple time points. The multivariate growth curve model enables us to examine relationships between different response variables (also see Carter & Hubert, 1984; Lundbye-Christensen, 1996; Nummi & Möttönen, 2000). Another recent extension is to impose reduced-rank restrictions on the univariate growth curve model, motivated by the fact that the mathematical structure of the growth curve model is akin to that of the reduced-rank

The work reported in this paper was supported by Grant A6394 from the Natural Sciences and Engineering Research Council of Canada to the second author. We thank Jennifer Stephan for her helpful comments on an earlier version of this paper. We also thank Patrick Curran and Terry Duncan for kindly letting us use the NLSY and substance use data, respectively. The substance use data were provided by Grant DA09548 from the National Institute on Drug Abuse.
Requests for reprints should be sent to Heungsun Hwang, HEC Montreal, Department of Marketing, 3000 Chemin de la Cote Ste-Catherine, Montreal, Quebec, H3T 2A7, CANADA. Email: heungsun. hwang@hec. ca.

0033-3123/2004-1/2001-0910-A $00.75/0

regression model (Albert & Kshirsagar, 1993; Reinsel & Velu, 1998, pp. 171–176). This univariate reduced-rank growth curve model may provide more parsimonious results than the standard growth curve model, if the rank restrictions are reasonable.

In this paper, we propose a multivariate reduced-rank growth curve model, which extends the univariate reduced-rank growth curve model to the multivariate case. The proposed model may lead to simpler interpretations about relationships among a number of response variables than the traditional multivariate growth curve model. Moreover, the method is more flexible than the traditional (multivariate) growth curve models in various aspects. For example, response variables do not have to be measured at the same time points, nor the same number of time points. This can be an advantage compared with the existing multivariate growth curve models that assume identically time-structured or balanced response variables. It is also possible to apply diverse kinds of basis function matrices with different rank across response variables. Furthermore, it is not necessary to specify an entire set of basis functions in advance, and some of the functions can be left unknown to be freely estimated to obtain their more optimal forms. This is distinct from the traditional growth curve model in which basis functions are all fixed beforehand.

This paper is organized as follows: Section 2 discusses the proposed method in detail. It provides the proposed model and estimation of model parameters. Section 3 illustrates empirical validity of the proposed method with two examples. The final section briefly summarizes the previous sections and discusses further prospects of the proposed method.

2. The Method

Suppose that N individuals are measured on the jth response variable ($j = 1, \ldots, J$) at T_j different time points. Let $\mathbf{Y}_j$ denote an N by T_j matrix of complete repeated measurements on the jth response variable. Let $\mathbf{X}_j$ denote a known N by P_j matrix of time-invariant explanatory variables, where P_j corresponds with the number of the explanatory variables for the jth response variable. Let $\mathbf{A}_j$ denote a D_j by T_j matrix of basis functions that represent specific aspects of change of $\mathbf{Y}_j$ across T_j time points. Let $\mathbf{B}_j$ denote a P_j by D_j matrix of unknown coefficients. Let $\mathbf{E}_j$ denote an N by T_j matrix of error. Then, the reduced-rank growth curve model for the jth response variable is given by

$$\mathbf{Y}_j = \mathbf{X}_j\mathbf{B}_j\mathbf{A}_j + \mathbf{E}_j, \tag{1}$$

with

$$\text{rank}(\mathbf{B}_j\mathbf{A}_j) \leq D_j \leq \min(P_j, T_j). \tag{2}$$

Model (1) is the univariate reduced-rank growth curve model since there is only a single (jth) response variable involved. This model is distinct from the standard growth curve model that assumes $D_j \leq \min(P_j, T_j)$ (e.g., Reinsel & Velu, 1998, p. 155). In the model, $\mathbf{A}_j$ may be a priori known or partially known (i.e., some elements in $\mathbf{A}_j$ are left unknown to be estimated). When $\mathbf{A}_j$ is prescribed, the model amounts to the standard growth curve model with a reduced-rank restriction of $\text{rank}(\mathbf{B}_j) \leq D_j$. In this case, the maximum likelihood (ML) estimates of $\mathbf{B}_j$ can be analytically obtained in the same way as in the standard growth curve model. The constrained principal components analysis by Takane and Shibayama (1991) may also include the univariate reduced-rank univariate growth curve model as a special case, where the portion of $\mathbf{Y}_j$ explained by both fixed row ($\mathbf{X}_j$) and column constraints ($\mathbf{A}_j$) is analyzed. If $\mathbf{A}_j$ is totally unknown, on the other hand, model (1) reduces to the reduced-rank regression model (e.g., Anderson, 1951; Davies & Tso, 1982; Izenman, 1975; Rao, 1964), or equivalently the redundancy analysis model (van den Wollenberg, 1977), given some identification restriction. Van der Leeden (1990, p. 121) has applied the reduced-rank regression model for analysis of repeated measurements with emphasis on the rank restrictions as a distinguishing aspect from the standard growth curve model.

Bijleveld and de Leeuw (1991) proposed a longitudinal reduced rank regression model that extended the reduced rank regression model to take into account autoregressive effects in $\mathbf{X}_j\mathbf{B}_j$. However, their method does not explicitly capture time-dependent changes in repeated measurements.

Albert and Kshirsagar (1993) and Reinsel and Velu (1998, pp. 171–176) proposed a special case of (1) under the same terminology. Their model can indeed be viewed as a second-order (reduced-rank) growth curve model (i.e., linear components of $\mathbf{X}_j$ are nested within second-order components) where $\mathbf{B}_j$ is the product of two unknown sub-matrices of coefficients, say $\mathbf{B}_{j1}$ and $\mathbf{B}_{j2}$ (i.e., $\mathbf{B}_j = \mathbf{B}_{j1}\mathbf{B}_{j2}$), so that $\text{rank}(\mathbf{B}_j) = \text{rank}(\mathbf{B}_j\mathbf{A}_j) \leq D = \min(D_j, P_j)$, and $\mathbf{A}_j$ is fixed.

Let $T = \sum_j T_j$, $D = \sum_j D_j$, and $P = \sum_j P_j$. The multivariate reduced-rank growth curve model for measurements of N individuals on J response variables over multiple time points may be expressed as

$$\begin{aligned}\left[\mathbf{Y}_1, \ldots, \mathbf{Y}_J\right] &= \left[\mathbf{X}_1, \ldots, \mathbf{X}_J\right]\begin{bmatrix}\mathbf{B}_1 & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{B}_J\end{bmatrix}\begin{bmatrix}\mathbf{A}_1 & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{A}_J\end{bmatrix} + \left[\mathbf{E}_1, \ldots, \mathbf{E}_J\right],\\ \mathbf{Y} &= \mathbf{XBA} + \mathbf{E},\end{aligned} \tag{3}$$

where $\mathbf{Y} = [\mathbf{Y}_1, \ldots, \mathbf{Y}_J]$ is an N by T matrix of observations on J response variables, $\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_J]$ is an N by P matrix of explanatory variables associated with J response variables, $\mathbf{B} = \text{diag}[\mathbf{B}_1, \ldots, \mathbf{B}_J]$ is a P by D matrix of unknown coefficients, $\mathbf{A} = \text{diag}[\mathbf{A}_1, \ldots, \mathbf{A}_J]$ is a D by T matrix of basis functions linked to J response variables, and $\mathbf{E} = [\mathbf{E}_1, \ldots, \mathbf{E}_J]$ is an N by T error matrix. In this model, $\mathbf{X}_j$ and $\mathbf{A}_j$ are allowed to differ across J response variables. Different rank restrictions may be imposed on the response variables, for example, $\text{rank}(\mathbf{B}_j\mathbf{A}_j) = h$ while $\text{rank}(\mathbf{B}_{j+1}\mathbf{A}_{j+1}) = h'$, where $h \neq h'$. Moreover, $\mathbf{Y}_j$ is not necessarily identically time-structured, that is, all individuals do not need to be measured on J response variables at the same time points, nor the same number of time points.

We note that in its most general form model (3) is equivalent to J separate analyses of the univariate reduced-rank growth curve model. Nevertheless, it should be emphasized that various constraints can be imposed on parameters across J response variables, thus allowing evaluation of a variety of hypotheses on relationships among the response variables. For instance, we can examine whether certain elements of $\mathbf{B}_j$ are identical across the response variables by imposing equality constraints. Often, the same explanatory variables can be duplicated across response variables or some comparable explanatory variables measured on the same scale can be observed in different response variables. In such cases, equality constraints are of use to reduce the number of redundant parameters, and provide simpler interpretations of solutions than the unconstrained case. Moreover, if the constraints are consistent with the data, we can obtain more reliable parameter estimates.

We may consider numerous special cases of model (3). For example, often, all matrices of explanatory variables are identical across J response variables, that is, $\mathbf{X}_1 = \mathbf{X}_2 = \cdots = \mathbf{X}_J$, so that $\mathbf{X}$ comes down to a single matrix. Then, (3) can be expressed as $\mathbf{Y} = \mathbf{X}[\mathbf{B}_1, \ldots, \mathbf{B}_J]\mathbf{A} + \mathbf{E}$. If $\mathbf{A}_j$'s are all known and no rank restrictions are involved, this case is essentially the same as the standard growth curve model, but deals with more than one response variable (Reinsel, 1982). This special case is further simplified when all known basis function matrices are also equal across J response variables, that is, $\mathbf{A}_1 = \mathbf{A}_2 = \cdots = \mathbf{A}_J$. In either case, the ML estimates of $\mathbf{B}$ are obtained in a similar way to the standard growth curve model. Model (3) also handles the second-order structures in $\mathbf{B}_j$ across J response variables in a simple way. To accommodate the second-order structure of coefficients for each response variable, the model can be expressed as

$$\begin{aligned}\mathbf{Y} &= \mathbf{X}\begin{bmatrix}\mathbf{B}_{11} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{B}_{J1}\end{bmatrix}\begin{bmatrix}\mathbf{B}_{12} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{B}_{J2}\end{bmatrix}\mathbf{A}+\mathbf{E},\\ &= \mathbf{X}\mathbf{B}_1\mathbf{B}_2+\mathbf{E}\\ &= \mathbf{X}\mathbf{B}\mathbf{A}+\mathbf{E},\end{aligned} \tag{4}$$

where

$$\mathbf{B}=\mathbf{B}_1\mathbf{B}_2=\begin{bmatrix}\mathbf{B}_{11} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{B}_{J1}\end{bmatrix}\begin{bmatrix}\mathbf{B}_{12} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{B}_{J2}\end{bmatrix}.$$

This model includes a multivariate extension of the univariate second-order growth curve model (Albert & Kshirsagar, 1993; Reinsel & Velu, 1998, pp. 171–176) as a special case.

Under the assumption that each row of $\mathbf{E}$ is *iid* multivariate normal, we seek to maximize the log likelihood function to derive the ML estimates of model parameters:

$$f(\mathbf{B},\mathbf{A},\mathbf{\Sigma})=\rho+\frac{N}{2}\log|\mathbf{\Sigma}|-\frac{1}{2}\mathrm{tr}[(\mathbf{Y}-\mathbf{XBA})\mathbf{\Sigma}^{-1}(\mathbf{Y}-\mathbf{XBA})'], \tag{5}$$

where $\rho=-\frac{1}{2}NT\log(2\pi)$, and $\mathbf{\Sigma}$ is the T by T unknown population covariance matrix. In (5), $\mathbf{A}$ is assumed at most partially known.

To maximize (5), we may use an optimization procedure similar to an alternating maximum likelihood (AML) procedure (de Leeuw, 1989; van der Leeden, 1990). The procedure consists of two global steps: In the first step, (5) is optimized over $\mathbf{B}$ and $\mathbf{A}$ for fixed $\mathbf{\Sigma}$. In the second step, (5) is optimized over $\mathbf{\Sigma}$, for fixed $\mathbf{B}$ and $\mathbf{A}$. These steps are alternated until convergence is obtained.

The first global step minimizes

$$f(\mathbf{B},\mathbf{A}\mid\mathbf{\Sigma})=\tfrac{1}{2}\mathrm{tr}[(\mathbf{Y}-\mathbf{XBA})\mathbf{\Sigma}^{-1}(\mathbf{Y}-\mathbf{XBA})'], \tag{6}$$

for fixed $\mathbf{\Sigma}$. Let $\mathbf{\Sigma}^{-1}=\mathbf{RR}'$, and this is equivalent to minimizing

$$\begin{aligned}f^* &= \mathrm{SS}((\mathbf{Y}-\mathbf{XBA})\mathbf{R})\\ &= \mathrm{SS}(\tilde{\mathbf{Y}}-\mathbf{XB}\tilde{\mathbf{A}})\end{aligned} \tag{7}$$

where $\mathrm{SS}(\mathbf{M})=\mathrm{trace}(\mathbf{M}'\mathbf{M})$, $\tilde{\mathbf{Y}}=\mathbf{YR}$, and $\tilde{\mathbf{A}}=\mathbf{AR}$ (e.g., Rao, 1980). Due to the zero structure of $\mathbf{B}$ and $\mathbf{A}$, (7) is not solved in a closed-form. Instead, (7) should be minimized by an iterative method. We use an alternating least squares (ALS) algorithm to minimize (7). Our algorithm is a simple adaptation of the ALS algorithm developed by Kiers and ten Berge (1989). In the algorithm, $\mathbf{B}$ and $\mathbf{A}$ are updated alternately until convergence is reached. The updates of one parameter matrix are optimally obtained such that they minimize (7) in the least squares sense, while the other is fixed.

To employ the ALS algorithm, specifically, we may rewrite (7) as

$$f^*=\mathrm{SS}(\mathrm{vec}(\tilde{\mathbf{Y}})-\mathrm{vec}(\mathbf{XB}\tilde{\mathbf{A}})) \tag{8a}$$

$$=\mathrm{SS}(\mathrm{vec}(\tilde{\mathbf{Y}})-(\tilde{\mathbf{A}}'\otimes\mathbf{X})\mathrm{vec}(\mathbf{B})) \tag{8b}$$

$$=\mathrm{SS}(\mathrm{vec}(\tilde{\mathbf{Y}})-(\mathbf{R}'\otimes\mathbf{XB})\mathrm{vec}(\mathbf{A})) \tag{8c}$$

where vec($\mathbf{M}$) denotes a supervector formed by stacking all columns of $\mathbf{M}$ one below another, and $\otimes$ denotes a Kronecker product. The algorithm can then repeat the following local steps until convergence is reached:

1. Update $\mathbf{B}$ for fixed $\mathbf{A}$ as follows: Let $\mathbf{b}$ denote the vector formed by eliminating zero elements from vec($\mathbf{B}$) in 8b. Let $\boldsymbol{\Omega}$ denote the matrix formed by eliminating the columns of $\tilde{\mathbf{A}} \otimes \mathbf{B} \rightarrow \tilde{\mathbf{A}}' \otimes \mathbf{X}$ in 8b corresponding to the zero elements in vec($\mathbf{B}$). Then, the least squares estimate of $\mathbf{b}$ is obtained by

$$\hat{\mathbf{b}} = (\boldsymbol{\Omega}'\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}'\text{vec}(\tilde{\mathbf{Y}}). \tag{9}$$

   The updated $\mathbf{B}$ is reconstructed from $\hat{\mathbf{b}}$.
2. Update $\mathbf{A}$ for fixed $\mathbf{B}$ as follows: Let $\mathbf{a}$ denote the vector formed by eliminating any fixed (or known) elements from vec($\mathbf{A}$) in 8c. Let $\boldsymbol{\Gamma}$ denote the matrix formed by eliminating the columns of $\mathbf{R}' \otimes \mathbf{XB}$ in 8c corresponding to the fixed elements in vec($\mathbf{A}$). Then, the least squares estimate of $\mathbf{a}$ is obtained by

$$\hat{\mathbf{a}} = (\boldsymbol{\Gamma}'\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}'\text{vec}(\tilde{\mathbf{Y}}). \tag{10}$$

   The updated $\mathbf{A}$ is recovered from $\hat{\mathbf{a}}$.

In the next global step, we update $\boldsymbol{\Sigma}$ for fixed $\mathbf{B}$ and $\mathbf{A}$. This amounts to maximizing

$$f(\boldsymbol{\Sigma} \mid \mathbf{B}, \mathbf{A}) = \rho + \frac{N}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\text{tr}[\mathbf{V}\boldsymbol{\Sigma}^{-1}], \tag{11}$$

where $\mathbf{V} = (\mathbf{Y} - \mathbf{XBA})'(\mathbf{Y} - \mathbf{XBA})$. Given $\mathbf{B}$ and $\mathbf{A}$, it is well known that the ML estimate of $\boldsymbol{\Sigma}$ is $\hat{\boldsymbol{\Sigma}} = N^{-1}\mathbf{V}$ (e.g., Anderson, 1984, p. 62).

We alternate the two global steps until convergence of (5) is reached. The AML algorithm seems to be fairly efficient thus far according to our experience with a number of examples. It converges fast with random starts (usually within less than 10 iterations) and seems to be hardly affected by the nonglobal minimum problem when being run with a number of different starts. The optimization procedure can be extended to fit various special cases of model (3). For instance, when the elements of $\mathbf{A}$ are all known, the algorithm becomes simpler since Step (10) is not required. In addition, the algorithm can be readily extended to fit model (4) that accommodates the second-order structure in $\mathbf{B}$. In this case, the first global step of ALS repeatedly updates each of the parameter matrices, $\mathbf{B}_1$, $\mathbf{B}_2$, and $\mathbf{A}$, with the others fixed.

We may be interested in testing various structural hypotheses regarding $\mathbf{B}$ and/or $\mathbf{A}$. A variety of structural hypotheses on parameters can be incorporated in the form of linear constraints. The linear constraints may be specified by either the reparametrization or the null-space method (Böckenholt & Takane, 1994; Takane, Yanai, & Mayekawa, 1991). The former method specifies the space spanned by column vectors of a constraint matrix, whereas the latter specifies its ortho-complement space. In the proposed method, all linear constraints are imposed by the reparametrization method. For example, let $\mathbf{H}$ denote a matrix of linear constraints on $\mathbf{b}$. In the step of (9), we incorporate $\mathbf{H}$ into $\mathbf{b}$ as follows:

$$\mathbf{b} = \mathbf{H}\boldsymbol{\beta}, \tag{12}$$

for some $\boldsymbol{\beta}$. A least squares estimate of $\boldsymbol{\beta}$ is then given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{H}'\boldsymbol{\Omega}'\boldsymbol{\Omega}\mathbf{H})^{-1}\mathbf{H}'\boldsymbol{\Omega}'\text{vec}(\tilde{\mathbf{Y}}), \tag{13}$$

which leads to

$$\hat{\mathbf{b}} = \mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{H}(\mathbf{H}'\boldsymbol{\Omega}'\boldsymbol{\Omega}\mathbf{H})^{-1}\mathbf{H}'\boldsymbol{\Omega}'\mathrm{vec}(\tilde{\mathbf{Y}}). \tag{14}$$

This approach is called the projection method (see Seber, 1984, pp. 403–405; Takane, Yanai, & Mayekawa, 1991). However, it is sometimes easier to specify constraints in the null-space form (e.g., equality or zero constraints). In such cases, the constraints are first expressed in the null-space form, and then transformed into the reparametrization form. The transformation is straightforward. Let

$$\mathbf{L}'\mathbf{b} = \mathbf{0} \tag{15}$$

represent the constraints in the null space form. Suppose that the first and the last elements of $\mathbf{b}$ are equal, then, $\mathbf{L}'$ reduces to a vector whose first element is 1, last element is -1, and other elements are zeros. We may reparametrize (15) into the form of (12) by defining $\mathbf{H} = \mathbf{I} - \mathbf{L}(\mathbf{L}'\mathbf{L})^{-}\mathbf{L}'$. This implies that $\mathrm{Ker}(\mathbf{L}') = \mathrm{Sp}(\mathbf{H})$. Linear constraints on $\mathbf{a}$ can be imposed in a similar way.

The proposed method makes it possible to fit a wide range of models to the data in hand. To assess the goodness of fit of fitted models, we use the Akaike information criterion (AIC). AIC is defined as follows: Let Γ denote the -2 log maximum likelihood value for a particular model, and γ denote the number of parameters estimated in the model. Then $\mathrm{AIC} = \Gamma + 2\gamma$. As shown in the formula, AIC adds some penalty to the log maximum likelihood value for increasing the number of parameters. The criterion aims to balance model fit (represented by Γ) and model parsimony (implied by γ). AIC favors simpler models over complex models if a similar fit sustains. It also applies to comparison across both nested and nonnested models. A model that minimizes AIC is regarded as the most appropriate one among fitted models. AIC is valid when γ is smaller than $2\sqrt{N}$ or $N/2$ at most (Sakamoto, Ishiguro, & Kitagawa, 1986, p. 83). We also note that the actual value of AIC for a specified model provides little information on the goodness of fit of the model itself. It is only beneficial when being compared to the AIC values of other models, that is, only the differences of AIC matter, not the actual values. Thus, we employ another fit index that furnishes certain information on the goodness of fit of a particular model itself. We call the index EV (which stands for Explained Variance) since it is proportional to the total variance of $\mathbf{Y}$ explained by the assumed model. The EV is given by

$$\mathrm{EV} = 1 - \frac{\mathrm{SS}(\mathbf{Y} - \mathbf{XBA})}{\mathrm{SS}(\mathbf{Y})}. \tag{16}$$

It ranges from 0 to 1. The larger the value, the more variance of $\mathbf{Y}$ is explained. Thus, a model that maximizes EV is preferable. Note, however, that EV is obviously affected by model complexity, that is, the more parameters, the larger EV. We thus consider both fit measures at the same time for more elaborate model selection.

In actual model selection, we may specify a class of candidate models for the data, and then compare the goodness of fit of the models. However, it may be implausible to fit all possible models for which the statistical fit measures are evaluated. Thus, additional knowledge from prior investigation about the data structure or any theoretical information can be utilized to eliminate less sensible models from the list of alternative models. Furthermore, it is often difficult to choose a final model solely in terms of statistical measures of model fit, since different models may have identical model fit. In such a case, nonstatistical considerations such as interpretability of the fitted model play a crucial role in making a final selection, although they are usually more complicated to justify since they are largely subjective.

3. Examples

3.1. The NLSY Data

The first example is part of the National Longitudinal Survey of Youth (NLSY), conducted by the U. S. Department of Labor. Starting in 1986, a large sample of children and their mothers were administered a set of assessment instruments every other year until 1992. Interviews were conducted on each child and her/his mother about the child. From the original sample of children and mothers, we analyzed a smaller sample of child-mother pairs, provided in Curran (1998). (Note that Curran (1998) considered only one biological child from each mother.) The sample consisted of 221 pairs of children and mothers, who completed interviews at four time points.

Two variables were repeatedly measured over the four time points: Antisocial behavior and reading recognition. Antisocial behavior of children was measured as a sum of the mother's responses to six items from the Behavior Problems Index antisocial behavior subtest. The child's reading recognition skill was computed by summing the total number of correct items by children out of 84 items of the Peabody Individual Achievement Reading Recognition subtest. Besides the two repeatedly measured variables, three variables were assessed once at the initial time point: cognitive stimulation for children at home, emotional support for children at home, and gender. A measure of cognitive stimulation was obtained as a sum of the mother's responses to 14 items in the cognitive stimulation subscale of the Home Observation for Measurement of the Environment-Short Form (HOME-SF). Emotional support was measured by summing the mother's responses to 13 items from the HOME-SF. Female child was coded as -1 and male child as 1.

The proposed method was applied to investigate optimal patterns of change in antisocial behavior and reading recognition over time and also to examine the effects of cognitive stimulation, emotional support, and gender on the temporal patterns. For our analysis, antisocial behavior and reading recognition were used as two response variables in $\mathbf{Y}$, so that $J = 2$, and $T_1 = T_2 = 4$. The former response variable was regarded as $\mathbf{Y}_1$ and the latter as $\mathbf{Y}_2$. Cognitive stimulation, emotional support, and gender were considered as explanatory variables in $\mathbf{X}_j$. In addition, $\mathbf{X}_j$ contained an intercept term. Matrix $\mathbf{X}_j$ was common to the two response variables, that is, $\mathbf{X}_1 = \mathbf{X}_2$.

Table 1 shows descriptive statistics for the two response variables measured at four different time points. From the descriptive statistics it seemed that there was a linear trend in both response variables because the mean levels of the response variables measured at each time point increased over time. In addition, the response variables were merely measured over four time

TABLE 1.
Descriptive statistics for two response variables measured at four time points in the NLSY data

Response variable	Mean	S. D.	Minimum	Maximum
Antisocial Behavior				
Time 1	1.49	1.54	0	7.00
Time 2	1.84	1.79	0	9.00
Time 3	1.88	1.80	0	10.00
Time 4	2.07	2.08	0	9.00
Reading Recognition				
Time 1	2.52	0.88	0.70	7.20
Time 2	4.04	1.00	1.60	6.20
Time 3	5.02	1.10	2.20	8.40
Time 4	5.80	1.22	2.50	8.30

TABLE 2.
Summary of fit for various multivariate growth curve models using the NLSY data

Type of Model	Rank of $\mathbf{B}_j\mathbf{A}_j$	γ	EV	AIC
Model 1: $\mathbf{A}_j$ is known	4 ($j = 1, 2$)	32	.86	5231.2
Model 2: $\mathbf{A}_j$ is known	3 ($j = 1, 2$)	24	.86	5230.4
Model 3: $\mathbf{A}_j$ is partially known	3 ($j = 1, 2$)	32	.86	5317.0
Model 4: $\mathbf{A}_j$ is known	2 ($j = 1, 2$)	16	.85	5304.9
Model 5: $\mathbf{A}_j$ is known and $\mathbf{B}_j$ is constrained	2 ($j = 1, 2$)	14	.85	5302.4
Model 6: $\mathbf{A}_j$ is known	1 ($j = 1, 2$)	8	.75	5732.9

points. This relatively small number of time points may not be sufficient to provide a nonlinear temporal change of the response variables. According to these prior investigations about the data, we initially assumed a multivariate growth curve model with a prescribed $\mathbf{A}_j$ of dimension 2×4, so that the rank of $\mathbf{B}_j\mathbf{A}_j$ is equal to 2, for both response variables. In this model, $\mathbf{A}_j$ was defined as a known matrix of orthogonal polynomials of order 1 to represent a linear trend of temporal change in both antisocial behavior and reading recognition (the exact form of the orthogonal polynomials are given below). This model is denoted by model 4 in the sequel. Based on the model of rank 2, a number of models were contemplated as possible competing models. The competing models were created in a hierarchical fashion that successively increased/decreased the rank of the assumed model and also imposed/released constraints on parameters within the same model. For final model selection, then, the assumed model was compared with the competing models.

A summary of the goodness of fit of the fitted models is presented in Table 2. In Table 2, model 1 corresponds with a multivariate growth curve model, in which the rank of $\mathbf{B}_j\mathbf{A}_j$ was equal to 4. In model 1, $\mathbf{A}_j$ was pre-specified as a matrix of orthogonal polynomials of order 3, that is,

$$\mathbf{A}_j = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 1 & -1 & -1 & 1 \\ -1 & 3 & -3 & 1 \end{bmatrix}. \tag{17}$$

Model 2 is a model with rank of $\mathbf{B}_j\mathbf{A}_j$ equal to 3, in which $\mathbf{A}_j$ was given as a matrix of orthogonal polynomials of order 2 for both response variables. Model 2 hypothesized that both antisocial behavior and reading recognition changed in a quadratic fashion over time. Model 3 is also a model with rank of $\mathbf{B}_j\mathbf{A}_j$ equal to 3. In model 3, however, $\mathbf{A}_j$ were assumed only partially known in such a way that the first two rows of $\mathbf{A}_j$ were given as a matrix of orthogonal polynomials of order 1, while the last row of $\mathbf{A}_j$ was left unknown to be estimated. That is,

$$\mathbf{A}_j = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ a_{j1} & a_{j2} & a_{j3} & a_{j4} \end{bmatrix}, \tag{18}$$

where a_{jt} is unknown ($t = 1, \ldots, 4$). Model 5 is a constrained version of model 4, in which the coefficients corresponding to the mean intercepts and mean growth rates were assumed to be equivalent across the response variables. We contemplated these equality constraints since the response variables were measured on a similar scale. Although the same explanatory variables are replicated over the two response variables, on the other hand, we did not presume that their effects on the response variables were also identical because the response variables were quite different types of characteristics. The last model (model 6) specified $\mathbf{A}_j$ as a matrix of orthogonal polynomials of zero order, assuming that there was stability in both response variables over time. For each of the fitted models, $\mathbf{A}_j$ was equal across $\mathbf{Y}_j$, that is, $\mathbf{A}_1 = \mathbf{A}_2$. Therefore, each of them is a special case of (3) in which both explanatory variables and basis functions are replicated over two response variables.

The EV values of models 1–5 were quite similar, in each case approximately 85% or 86% of the variance of $\mathbf{Y}$ was accounted for. Model 2, on the other hand, had the smallest AIC value. Model 2 presented that the measurements on both antisocial behavior and reading recognition varied in a quadratic manner over time. However, it was found that the estimated mean quadratic trends for antisocial behavior and reading recognition were equal to −.26 (s.e. = .11) and −.13 (s.e. = .05), respectively. It suggests that the levels of antisocial behavior and reading recognition tend to first increase linearly and then gradually decrease toward the end of the time points. The inverted U shaped quadratic trend seemed to be rather inconsistent with the descriptive statistics that suggested monotonically increasing patterns of change in both response variables over time. Thus, we did not select model 2 as a final model due to the inconsistent results, even though it was found to be the best-fitting model in terms of AIC. Such conflicting solutions were also found in the full-rank model (model 1), which had the second smallest AIC.

Instead, we chose model 5 as a final model because it had the third smallest AIC. More crucial was that model 5 provided sensible and simpler interpretations of obtained results than other fitted models. Model 5 was favored over the initially assumed model (model 4) because its AIC was smaller than model 4's AIC, providing essentially the same but simpler interpretations. The final model posits that both antisocial behavior and reading recognition vary in a linear trend or growth rate over time. The linear pattern of change in antisocial behavior of the same children was also reported by Curran and Bollen (1999). Moreover, it shows that the mean intercept and mean growth rate of the response variables are identical across the response variables.

Table 3 provides the ML estimates of $\mathbf{B}_j$ in the final model with their standard errors in the parentheses. The first column of $\mathbf{B}_j$ under the label of Initial provides the effects of the explanatory variables on the response variables at the initial status, and the second column under the label of Linear represents the effects of the explanatory variables on a growth rate of the response variables over time. The estimated mean intercept is equal to 3.19 (s.e. = .14) and the mean growth rate is equal to .27 (s.e. = .04), which are constrained to be the same for both response variables. The constrained mean intercept and growth rate estimates indicate that antisocial behavior and reading recognition seem to increase in the same linear pattern during the study. The effect of gender on the initial status of antisocial behavior is .64 (s.e. = .09). It suggests that boys show a higher level of antisocial behavior than girls through the study. The effect of gender on the growth rate is .02 (s.e. = .02), indicating that boys tend to increase antisocial behavior at a higher rate compared to girls. Yet, this effect appears less reliable. The effects of cognitive stimulation on the initial status and the growth rate of antisocial behavior are equal to −.11 (s.e. = .02) and −.03 (s.e. = .00), respectively. It suggests that children

TABLE 3.
The ML estimates of $\mathbf{B}_j$ in the final model for the NLSY data (standard errors in parentheses). I = Intercept, G = Gender, C = Cognitive stimulation, E = Emotional support

	Antisocial Behavior		Reading Recognition	
	Initial	Linear	Initial	Linear
I	3.19 (.14)	.27 (.04)	3.19 (.14)	.27 (.04)
G	.64 (.09)	.02 (.02)	−.07(.06)	.01 (.01)
C	−.11(.02)	−.03(.00)	.06 (.01)	.00 (.00)
E	−.15(.02)	.00 (.00)	.06 (.01)	.02 (.00)

receiving higher levels of cognitive stimulation at home show lower levels of antisocial behavior, and also increase antisocial behavior at a lower rate than those receiving lower levels of cognitive stimulation. The effect of emotional support on the initial status of antisocial behavior is $-.15$ (s.e. $= .02$), indicating that the higher levels of emotional support at home, the lower levels of antisocial behavior. On the other hand, the effect of emotional support on the growth rate of antisocial behavior is .00 (s.e. $= .00$). This estimate, however, looks less sizable to interpret.

The effects of gender on the initial status and the growth rate of reading recognition are $-.07$ (s.e. $= .06$) and .01 (s.e. $= .01$), respectively. It suggests that girls show a higher level of reading recognition than boys through the study, while boys seem to show a higher rate of increase in reading recognition. However, both effects appear less reliable. The effect of cognitive stimulation on the initial status of reading recognition is equal to .06 (s.e. $= .01$). It indicates that the higher levels of cognitive stimulation at home, the higher levels of reading recognition skill acquired. The effect of cognitive stimulation on the growth rate is .00 (s.e. $= .00$), which appears less sizable to interpret. The effects of emotional support on the initial status and the growth rate of reading recognition are .06 (s.e. $= .01$) and .02 (s.e. $= .00$), respectively. It suggests that children with higher levels of emotional support gain higher levels of reading recognition at the initial time point and also increase their reading recognition skills at a higher rate than those with lower levels of emotional support.

The chosen model seems to fit well, providing useful information on the relationship between time-dependent response variables and time-invariant predictor variables. However, it does not necessarily mean that the model fits perfectly and thus reflects the actual state of affairs.

3.2. The Substance Use Data

The second example comes from a longitudinal study on the predictors and consequences of substance use among adolescents from American northwestern urban areas (Duncan, Duncan, Alpert, Hops, Stoolmiller, & Muthén, 1994). For the present analysis, 632 adolescents were measured on their use of three drugs such as marijuana, cigarettes, and alcohol over four time points. The three measures of substance use were assessed based on a self-reported 5-point item: (1) life time abstainers, (2) 6-month abstainers, (3) current use of less than four times a month, (4) current use of between 4 and 29 times a month, and (5) current use of 30 or more times a month. Five additional variables were measured once at the initial time point, including parental marital status, family status, socio-economic status (SES), age, and gender. Marital status was classified as follows: $0 =$ single and $1 =$ married or living in a committed relationship. Family status was categorized as follows: $0 =$ step or foster families and $1 =$ others. SES was calculated as the average of parental annual income and education level. Parental annual income was assessed based on a 16-point scale ranging from "6,000 dollars and below" to "50,000 dollars or more. " Education levels range from "Grade level 6 or less" to "Graduate level. " Male and female were coded as 0 and 1, respectively.

The three measures of substance use were employed as response variables in $\mathbf{Y}$. That is, $J = 3$, and $T_1 = T_2 = T_3 = 4$. Use of marijuana, cigarettes, and alcohol were taken as $\mathbf{Y}_1$, $\mathbf{Y}_2$, and $\mathbf{Y}_3$, respectively. The five nonrepeated measures were used as explanatory variables in $\mathbf{X}_j$. Also, $\mathbf{X}_j$ consisted of an intercept term. Matrix $\mathbf{X}_j$ was identical across the three response variables, that is, $\mathbf{X}_1 = \mathbf{X}_2 = \mathbf{X}_3$. Table 4 shows descriptive statistics for the three response variables measured at 4 time points.

From Table 4 it appeared that there was a linear trend in all three response variables because the mean levels of the response variables measured at each time point were monotonically increasing over time. Moreover, the number of repeated measurements seemed to be insufficient to observe a nonlinear temporal change of the response variables. Therefore, we initially assumed a multivariate growth curve model of a prescribed $\mathbf{A}_j$ of dimension 2×4 as in Example 1. In

TABLE 4.
Descriptive statistics for three response variables measured at four time points in the substance use data

Response variable	Mean	S. D.	Minimum	Maximum
Marijuana				
Time 1	1.52	0.95	1	5
Time 2	1.70	1.00	1	5
Time 3	1.84	1.07	1	5
Time 4	2.01	1.09	1	5
Cigarettes				
Time 1	1.85	1.14	1	5
Time 2	2.07	1.27	1	5
Time 3	2.20	1.26	1	5
Time 4	2.50	1.35	1	5
Alcohol				
Time 1	2.23	1.04	1	5
Time 2	2.46	1.00	1	5
Time 3	2.65	1.00	1	5
Time 4	2.94	0.94	1	5

this model, $\mathbf{A}_j$ was specified as a matrix of orthogonal polynomials of order 1 to capture a linear trend of temporal change in the three response variables. Starting from the model of rank 2, various models were specified as alternative models in a similar way to the previous example. They were fitted and compared with the assumed model.

Table 5 provides the goodness of fit of the fitted models. In Table 5, model 1 is a multivariate growth curve model, where the rank of $\mathbf{B}_j\mathbf{A}_j$ was equal to 4. In model 1, $\mathbf{A}_j$ was given as a matrix of the same orthogonal polynomials as (17). In model 2, $\mathbf{A}_j$ was given as a matrix of orthogonal polynomials of order 2, representing a quadratic pattern of change in the response variables over time. Model 3 is the model we originally assumed. The rank of model 4 is equal to model 3. In model 4, however, $\mathbf{A}_j$ was assumed to be partially known. The first row of $\mathbf{A}_j$ was given as a matrix of orthogonal polynomials of zero order, while the second row of $\mathbf{A}_j$ was left unknown to be estimated. Model 5 is a constrained version of model 3 that imposed equality constraints on all four explanatory variables except gender (i.e., parental marital status, family status, SES, and age) across the three response variables. We expected that the effects of the four explanatory variables could be interpreted in the same direction for all the response variables. For example, adolescents living with a single parent might be expected to show higher levels of use of all three substances than those living with both parents. Adolescents living with other types of families were expected to show more frequent use of all the substances than those with single or

TABLE 5.
Summary of fit for various multivariate growth curve models using the substance use data

Type of Model	Rank of $\mathbf{B}_j\mathbf{A}_j$	γ	EV	AIC
Model 1: $\mathbf{A}_j$ is known	4 ($j = 1, 2, 3$)	72	.82	16180
Model 2: $\mathbf{A}_j$ is known	3 ($j = 1, 2, 3$)	54	.82	16180
Model 3: $\mathbf{A}_j$ is known	2 ($j = 1, 2, 3$)	36	.82	16167
Model 4: $\mathbf{A}_j$ is partially known	2 ($j = 1, 2, 3$)	48	.82	16171
Model 5: $\mathbf{A}_j$ is known and $\mathbf{B}_j$ is constrained	2 ($j = 1, 2, 3$)	20	.82	16158
Model 6: $\mathbf{A}_j$ is known	1 ($j = 1, 2, 3$)	18	.81	16461

foster families. Adolescents with low SES were expected to use the three substance more than those with high SES. Older adolescents were expected to show more use in all the substances, compared to younger ones. In addition, the four explanatory variables were duplicated over the three response variables measured at the same scale. Thus, we assumed that the effects of the four explanatory variables were identical across the response variables. Although gender was also replicated across the response variables, on the other hand, we assumed that there would be some gender differences in use of the substances (e.g., male adolescents may use more marijuana while female adolescents more cigarettes). Moreover, we presumed that there would be mean differences in use of the three substances on the basis of the descriptive statistics. Thus, we did not apply equality constraints to intercepts and gender across the response variables. Model 6 specified $\mathbf{A}_j$ as a matrix of orthogonal polynomials of zero order, assuming that there was stability in the response variables over time. For each of the fitted models, $\mathbf{A}_j$ was equal across $\mathbf{Y}_j$, that is, $\mathbf{A}_1 = \mathbf{A}_2 = \mathbf{A}_3$. Therefore, each of these is a special case of (3) where both explanatory variables and basis functions are duplicated across response variables.

In this example, model 5 was chosen as a final model since its AIC is lowest. The EV value of the model was essentially equivalent to those of more complex models, indicating that it explained about 82% of the total variance of the response variables. Moreover, it provided much simpler and more plausible interpretations of the obtained results than other models. The final model posits that all three response variables change in a linear fashion over the four time points. This is consistent with the result of Duncan et al. (1997). It also indicates that the effects of four explanatory variables (i.e., parental marital status, family status, SES, and age) are equal while the effects of intercepts and gender are different across all the response variables.

Table 6 presents the ML estimates of $\mathbf{B}_j$ in the final model with their standard errors in parentheses. The estimated mean intercept and mean growth rate for use of marijuana are .31 (s.e. = .08) and .13 (s.e. = .02), respectively. The estimated mean intercept and mean growth rate for cigarette use are .67 (s.e. = .10) and .16 (s.e. = .02), respectively. The estimated mean intercept and mean growth rate for alcohol use are 1.08 (s.e. = .08) and .15 (s.e. = .02), respectively. These indicate linearly increasing trends of change in all response variables. Due to the equality constraints imposed, the effects of parental marital status, family status, SES, and age on the initial status and growth rate are identical across the three response variables. The effect of parental marital status on the initial status of the response variables is equal to −.29 (s.e. = .04) across the response variables. It indicates higher levels of substance use by adolescents living with single parents than those living with both parents. On the other hand, the effect of parental marital status on the growth rate of the response variables is .01 (s.e. = .01), indicating that adolescents living with nonsingle parents seem to increase their use of substance at a higher rate. Yet, this estimate looks less reliable. The effects of family status on the initial status and the growth rate

TABLE 6.
The ML estimates of $\mathbf{B}_j$ in the final model for the substance use data (standard errors in parentheses). I = Intercept, P = Parental marital status, F = Family status, S = SES, A = Age, G = Gender

	Marijuana		Cigarettes		Alcohol	
	Initial	Linear	Initial	Linear	Initial	Linear
I	.31 (.08)	.13 (.02)	67 (.10)	.16 (.02)	1.08 (.08)	.15 (.02)
P	−.29 (.04)	.01 (.01)	−.29 (.04)	.01 (.01)	−.29 (.04)	.01 (.01)
F	.12 (.01)	.01 (.00)	.12 (.01)	.01 (.00)	.12 (.01)	.01 (.00)
S	−.05 (.02)	−.03 (.00)	−.05 (.02)	−.03 (.00)	−.05 (.02)	−.03 (.00)
A	.12 (.01)	−.00 (.00)	.12 (.01)	−.00 (.00)	.12 (.01)	−.00 (.00)
G	−.02 (.02)	−.01 (.00)	.03 (.03)	−.02 (.01)	.04 (.02)	−.02 (.00)

of the response variables are .12 (s.e. = .01) and .01 (s.e. = .00), respectively. It suggests that adolescents living with other families rather than step or foster families show higher levels of substance use, and also their substance use increases at a higher rate over time.

The effects of SES on the initial status and the growth rate are equal to −.05 (s.e. = .02) and −.03 (s.e. = .00), respectively, for the response variables. It suggests that socially and economically more disadvantaged adolescents show higher levels of substance use, and also show a higher rate of increase in substance use than those less disadvantaged. The effect of age on the initial status of the response variables is equal to .12 (s.e. = .01), indicating that older adolescents show higher levels of substance use compared to younger adolescents over the study. The effect of age on the growth rate of the response variables is −.00 (s.e. = .00), which looks negligibly small. The effects of gender on the initial status for use of marijuana, cigarettes, and alcohol are −.02 (s.e. = .02), .03 (s.e. = .03), and .04 (s.e. = .02), respectively. It suggests that male adolescents seem to use higher levels of marijuana but lower levels of cigarettes and alcohol compared to female adolescents over the study. Yet, they seem to be less accurately measured. On the other hand, the effects of gender on the growth rate of use of marijuana, cigarettes, and alcohol are −.01 (s.e. = .00), −.02 (s.e. = .01), and −.02 (s.e. = .00), respectively. It indicates that male adolescents tend to increase use of all the drugs at a higher rate, compared to female adolescents.

4. Concluding Remarks

The multivariate reduced-rank growth curve model enables us to investigate diverse relationships among more than one response variable measured over multiple time points in a more parsimonious way than the traditional growth curve models. It helps us capture a more parsimonious pattern of change on time-dependent measurements. It also allows us to examine a variety of hypotheses among response variables. The method subsumes existing growth curve models as special cases. The method is quite flexible in that it allows for unbalanced response variables, and permits different rank restrictions for each response variable. In addition, some of the basis functions may be left unknown, and estimated to obtain their more optimal forms.

The proposed method may be extended in various ways. Inspired by Chinchilli and Elswick (1985), for example, it may be extended to include a MANOVA component as well as a growth curve component (also see Reinsel & Velu, 1998, p. 166). This extended model allows for the effect of additional explanatory variables on a response variable, which are not associated with basis functions. It may be attractive when the effects of some explanatory variables are examined in terms of time-dependent structure, whereas the effects of other variables are needed to be tested independently of such a structure. The proposed method may also be extended to a mixture of reduced-rank growth curve and full-rank growth curve components (e.g., Reinsel & Velu, 1998, p. 176).

In the proposed method, the population covariance matrix $\boldsymbol{\Sigma}$ is assumed to be unstructured or unconstrained. The unstructured covariance matrix becomes computationally less attractive when the number of repeated measurements becomes large, often leading to less reliable parameter estimates (Laird & Ware, 1982). Instead, we may assume that the population covariance matrix is structured in a certain way. The random effects growth curve model (e.g., Laird & Ware, 1982; Nummi, 1997; Rao, 1965; Reinsel, 1982) provides one type of special covariance structure, in which the covariance matrix is considered as a sum of the sub-covariance matrices of errors and random coefficients. Furthermore, we may apply the generalized estimating equation (GEE) approach (Liang & Zeger, 1986; Zeger & Liang, 1986) for the proposed method to accommodate a variety of covariance structures.

The proposed method may also be extended to deal with discrete (nominal or ordinal) variables through certain data transformations. In particular, the optimal scaling approach (e.g., Young, 1981) seems to be attractive since it may be readily suited to the AML estimation procedure. It is also important to handle missing observations, which frequently appear in large data sets. A simple way is to delete any cases having at least one missing observation. However, this is unsatisfactory if missing values are numerous and scattered throughout the data set, as deletion of the cases may incur substantial loss of information. A more effective way may be to impute missing observations by their estimates. The estimates for missing observations may be obtained in an iterative way (e.g., Gabriel & Zamir, 1979; Gifi, 1990): We start by completing the data with some initial estimates for missing observations, obtain model estimates by fitting the model to the complete data, update estimates of the missing observations based on the model estimates, fit the model to the updated data, and so on. These procedures are repeated until no significant changes take place in the estimates. This iterative imputation approach may be compatible with the estimation procedure of the proposed method.

Polynomials are typically chosen as basis functions for the growth curve model. These classic basis functions may not be sufficient for describing the shape of complex time-varying data (Ramsay, in press). Ramsay and Silverman (1997) suggested more diverse kinds of basis functions: For instance, the B-spline basis functions may be considered as good candidates for non-periodic data. Fourier series seems to be appropriate for periodic data. Different choices of basis functions may make the growth curve model more in line with functional linear models (Ramsay & Silverman, 1997).

References

Albert, J. M., & Kshirsagar, A. M. (1993). The reduced-rank growth curve model for discriminant analysis of longitudinal data. *Australian Journal of Statistics*, *35*, 345–357.

Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics*, *22*, 327–351.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley and Sons.

Bijleveld, C. C. J. H., & de Leeuw, J. (1991). Fitting longitudinal reduced-rank regression models by alternating least squares. *Psychometrika*, *56*, 433–447.

Böckenholt, U., & Takane, Y. (1994). Linear constraints in correspondence analysis. In M. J. Greenacre & J. Blasius (Eds.), *Correspondence Analysis in Social Sciences* (pp. 112–127). London: Academic Press.

Carter, E. M., & Hubert, J. J. (1984). A growth-curve model approach to multivariate quantal bioassay. *Biometrics*, *40*, 699–706.

Chinchilli, V. M., & Elswick, R. K. (1985). A mixture of the MANOVA and GMANOVA models. *Communications in Statistics: Theory and Methods*, *14*, 3075–3089.

Curran, P. J. (1998). Introduction to hierarchical linear models of individual growth: An applied example using the SAS data system. Paper presented at the First International Institute on Developmental Science, University of North Carolina, Chapel Hill.

Curran, P. J., & Bollen, K. A. (1999). A hybrid latent trajectory model of stability and change: Applications in developmental psychology. Paper presented at the biennial meeting of the Society for Research on Child Development, Albuquerque, New Mexico.

Davies, P. T., & Tso, M. K.-S. (1982). Procedures for reduced-rank regression. *Applied Statistics*, *31*, 244–255.

Duncan, T. E., Duncan, S. C., Alpert, A., Hops, H., Stoolmiller, M., & Muthén, B. (1997). Latent variable modeling of longitudinal and multilevel substance use data. *Multivariate Behavioral Research*, *32*, 275–318.

de Leeuw, J. (1989). Fitting reduced rank regression models by alternating maximum likelihood. UCLA Statistics Series 35. Department of Statistics, University of California at Los Angeles.

Gabriel, K. R., & Zamir, S. (1979). Low rank approximation of matrices by least squares with any choice of weights. *Technometrics*, *21*, 489–498.

Gifi, A. (1990). *Nonlinear Multivariate Analysis*. Chichester: John Wiley and Sons.

Grizzle, J. E., & Allen, D. M. (1969). Analysis of growth and dose response curves. *Biometrics*, *25*, 357–381.

Izenman, A. J. (1975). Reduced-rank regression for the multivariate linear model. *Journal of Multivariate Analysis*, *5*, 248–264.

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.

Kiers, H. A. L., & ten Berge, J. M. F. (1989). Alternating least squares algorithms for simultaneous components analysis with equal component weight matrices in two or more populations. *Psychometrika*, *54*, 467–473.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*, 963–974.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*, 13–22.

Lundbye-Christensen, S. (1991). A multivariate growth curve model for pregnancy. *Biometrics*, *47*, 637–657.

Nummi, T. (1997). Estimation in a random effects growth curve model. *Journal of Applied Statistics*, *24*, 157–168.

Nummi, T., & Möttönen, J. (2000). On the analysis of multivariate growth curves. *Metrika*, *52*, 77–89.

Potthoff, R. F., & Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, *51*, 313–326.

Ramsay, J. O. (in press). Multilevel modeling of longitudinal and functional data. To appear in D. Moskowitz & S. Hershberger (Eds.), *Modeling Intraindividual Variability with Repeated Measures Data: Methods and Applications.* New York: Erlbaum Associates.

Ramsay, J. O., & Silverman, B. W. (1997). *Functional Data Analysis*. New York: Springer-Verlag.

Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhyã* A, *26*, 329–358.

Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, *52*, 447–458.

Rao, C. R. (1980). Matrix approximations and reduction of dimensionality in multivariate statistical analysis. In P. R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 3–22). Amsterdam: North-Holland.

Reinsel, G. C. (1982). Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure. *Journal of the American Statistical Association*, *77*, 190–195.

Reinsel, G. C., & Velu, R. P. (1998). *Multivariate Reduced-Rank Regression: Theory and Applications*. New York: Springer-Verlag.

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike Information Criterion Statistics*. Boston: D. Reidel.

Seber, G. A. F. (1984). *Multivariate Observations*. New York: John Wiley and Sons.

Takane, Y., & Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika*, *56*, 97–120.

Takane, Y., Yanai, H., & Mayekawa, S. (1991). Relationships among several methods of linearly constrained correspondence analysis. *Psychometrika*, *56*, 667–684.

van den Wollenberg, A. L. (1977). Redundancy analysis: An alternative for canonical analysis. *Psychometrika*, *42*, 207–219.

van der Leeden, R. (1990). *Reduced Rank Regression with Structured Residuals*. Leiden: DSWO Press.

von Rosen, D. (1991). The growth curve model: A review. *Communications in Statistics: Theory and Methods*, *9*, 2791–2822.

Young, F. W. (1981). Quantitative analysis of qualitative data. *Psychometrika*, *46*, 357–388.

Zeger, S. L, & Liang, K. Y. (1986). The analysis of discrete and continuous longitudinal data. *Biometrics*, *42*, 121–130.

Manuscript received 29 JAN 2002
Final version received 24 JAN 2003