FITTING LONGITUDINAL REDUCED-RANK REGRESSION MODELS BY
ALTERNATING LEAST SQUARES

CATRIEN C. J. H. BIJLEVELD
DEPARTMENT OF PSYCHOMETRICS
LEIDEN UNIVERSITY

JAN DE LEEUW
DEPARTMENTS OF PSYCHOLOGY AND MATHEMATICS
UNIVERSITY OF CALIFORNIA AT LOS ANGELES

An alternating least squares method for iteratively fitting the longitudinal reduced-rank regression model is proposed. The method uses ordinary least squares and majorization substeps to estimate the unknown parameters in the system and measurement equations of the model. In an example with cross-sectional data, it is shown how the results conform closely to results from eigenanalysis. Optimal scaling of nominal and ordinal variables is added in a third substep, and illustrated with two examples involving cross-sectional and longitudinal data.

Key words: reduced-rank regression, state space analysis, optimal scaling.

Introduction

The linear model and the reduced-rank regression model are generalized to situations where a dependence exists between observations on different occasions. Thus, the techniques developed are intended for describing dynamic or longitudinal situations, in contrast to purely cross-sectional ones. The general framework also includes spatial multivariate data, in which the input of a given region influences not only the output of the region, but also the output of adjacent regions, or regions in the immediate neighborhood. In general, our models are appropriate for observations ordered in some clearly defined way (time and space are merely the most obvious examples), and when there is reason to suppose that close observations influence each other.

Consider the following empirical situation. At a number of occasions \( t = 1, \ldots, T \), we observe two vector variables \( x_t \) and \( y_t \). When \( y_t \) is influenced by \( x_t \) (i.e., \( x_t \) is the cause of \( y_t \)), \( x_t \) can be thought of as an input variable, and \( y_t \) as an output variable. In econometrics, \( x_t \) is called exogenous, and \( y_t \) endogenous. In psychometrics, and in various other areas of applied statistics, \( x_t \) is called the independent variable, and \( y_t \) dependent. Thus, we have two sets of variables, and the two sets play a somewhat different, asymmetric, role in our thinking.

In multivariate analysis the occasions are often replications of the same basic structure, and the index \( t \) denotes individuals or objects, considered as a sample from some well-defined population. It is assumed that there is no causal connection between variables with different indices. Thus \( x_1 \) influences \( y_1 \), \( x_2 \) influences \( y_2 \), and so on, but there is no influence of \( x_1 \) on \( x_2 \) or on \( y_2 \). This is called the independence assumption. Another important aspect of this type of model is stationarity, where the influence of \( x_1 \)...
on $y_1$ is supposed to be the same as that of $x_2$ on $y_2$, and so on. Such models are at the basis of regression analysis, and of linear models generally.

There is also a slightly more complicated class of independent and stationary models, which goes under various names: reduced-rank regression models, growth curve models, MIMIC models, or errors-in-variables models. Here, the influence of $x$ on $y$ is mediated by an unobserved latent vector variable $z$, with $x$ determining $z$, and $z$ determining $y$. In general, the number of latent variables in $z$ is smaller than the number of variables in $x$ or $y$, and in this sense, $z$ filters the relationships between the two sets of variables. We call the space of the variables in $z$ the latent or state space, and we use $p$ for its dimensionality. For various versions and applications of reduced-rank regression, we refer to Anderson (1951, 1984) and Jöreskog and Goldberger (1975). Alternating least squares algorithms for fitting reduced-rank regression models have been discussed by de Leeuw and Bijleveld (1987).

If the independence assumption is dropped in the reduced-rank regression models, the dynamic generalization we discuss is obtained, which as pointed out below is identical to the state space models studied in mathematical systems theory (Kalman, Falb, & Arbib, 1969). State space models have been discussed recently in the context of covariance structure models by MacCallum and Ashby (1986) and Otter (1986). Because we have social and behavioral science applications in mind, however, a quite different algorithm is developed in this paper that does not rely on the assumption of multivariate normal errors and allows for optimal scaling of the input and output variables.

State Space Models

To simplify the discussion, we shall use several concepts borrowed from factor analysis. Explicitly, in factor analysis, $m$ variables in the vector $y = (y_1, \ldots, y_m)'$ are observed that are correlated. It is assumed that there exist $p$ unobserved variables or factors in the vector $z = (z_1, \ldots, z_p)'$ that “explain” the association between the observed variables, in the sense that the observed variables are independent given the factors. In the reduced-rank regression model, the dependence of the output $y$ on the $k$ variables in the input vector $x = (x_1, \ldots, x_k)'$ is decomposed into dependence of the output $y$ on the latent factor $z$, and dependence of the latent factor $z$ on the input $x$. In the dynamic case, there is the unobserved state variable $z$ to mediate the influence of the input $x$ on the time-dependent $y$. This dependence of the output variables is accommodated by assuming that all influence of the past on the present is mediated by the present state variables. This first and basic assumption renders the model Markovian.

The state space model can be written as follows:

$$z_t = Fz_{t-1} + Gx_t + \epsilon_t,$$
$$y_t = Hz_t + \delta_t,$$

with $F$ the $p$ by $p$ matrix containing parameters specifying the influence of the past $p$ states $z_{t-1}$ on the present $p$ states $z_t$, $G$ the $p$ by $k$ matrix with parameters specifying the influence of the $k$ input variables $x$ on the $p$ state variables $z$, and $H$ the $m$ by $p$ matrix with parameters specifying the influence of the $p$ state variables $z$ on the $m$ output variables $y$. The errors terms $\epsilon_t$ and $\delta_t$ are needed since we do not expect a perfect fit to real data.

In the multivariate normal situation the random variables $\epsilon_t$ and $\delta_t$ each have independent centered multivariate normal distributions, where $\epsilon_t$ has covariance matrix $Q$ and $\delta_t$ has covariance matrix $W$. The maximum likelihood method can be used to estimate the structural parameters of the system (e.g., Hannan & Deistler, 1988; and