



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Computational Statistics &amp; Data Analysis ■■■ (■■■■) ■■■–■■■

**COMPUTATIONAL  
STATISTICS  
& DATA ANALYSIS**[www.elsevier.com/locate/csda](http://www.elsevier.com/locate/csda)

# An extended redundancy analysis and its applications to two practical examples

Yoshio Takane<sup>a,1</sup>, Heungsun Hwang<sup>b,\*</sup><sup>a</sup>*Department of Psychology, McGill University, 1205, Dr. Penfield Avenue, Montréal, Québec H3A 1B1, Canada*<sup>b</sup>*Department of Marketing, HEC Montréal, 3000 Chemin de la Côte-Ste-Catherine, Montréal, Québec H3T 2A7, Canada*

Received 24 September 2003; received in revised form 6 May 2004; accepted 5 June 2004

---

## Abstract

An extension of redundancy analysis is proposed that allows analyzing a variety of directional relationships among multiple sets of variables. The proposed method subsumes an existing redundancy analysis method as a special case. It is also extended further to analyze more complex relationships among variables such as direct effects of observed exogenous variables, higher-order components and multi-sample comparisons. An alternating least-squares algorithm is developed for parameter estimation. A small simulation study is conducted to investigate the performance of the proposed method. Two real examples are given to illustrate the empirical use of the proposed method.

© 2004 Elsevier B.V. All rights reserved.

**Keywords:** Redundancy analysis; Alternating least squares

---

## 1. Introduction

Redundancy analysis (RA; van den Wollenberg, 1977), also called reduced rank regression (Anderson, 1951) or principal components of instrumental variables (Rao, 1964), is a useful technique for analyzing a directional relationship between two sets of multivariate data (Lambert et al., 1988). RA aims to extract a series of components from a set of

---

\* Corresponding author. Tel.: +1-5143407008; fax: +1-5143405631.

E-mail addresses: [takane@takane2.psych.mcgill.ca](mailto:takane@takane2.psych.mcgill.ca) (Y. Takane), [heungsun.hwang@hec.ca](mailto:heungsun.hwang@hec.ca) (H. Hwang).

<sup>1</sup> Requests for reprints should be addressed to: Yoshio Takane, Department of Psychology, 1205 Dr. Penfield Avenue, Montréal, Québec H3A 1B1, Canada.

exogenous variables in such a way that they are mutually orthogonal and successively account for the maximum variance of a set of endogenous variables.

A few attempts have been made to extend RA to more than two sets of variables (Velu, 1991; Reinsel and Velu, 1998). However, they are limited to model and fit a particular type of relationship among three sets of variables. In this paper, we extend RA in such a way that it enables to specify and fit a variety of relationships among multiple sets of variables. The method proposed herein is called an extended redundancy analysis (ERA).

The remaining sections of this paper are organized as follows. In Section 2, we discuss ERA in full detail. We present the basic ERA model and parameter estimation procedures, and discuss how to examine structural hypotheses regarding model parameters by imposing linear constraints. We also evaluate the performance of ERA in terms of recovery of parameters by conducting a small simulation study. We then discuss the relationships of ERA with other methods, which can also handle similar models. We then analyze a practical example to illustrate the empirical feasibility of ERA. In Section 3, we provide useful extensions of the basic ERA model, including higher-order components, direct effects of observed endogenous variables, and multi-sample comparisons. We empirically illustrate each of the extensions. In the final section, we briefly summarize the previous sections and discuss further prospects of ERA.

## 2. An extended redundancy analysis

ERA, as the name suggests, is an extension of RA. Thus, we begin with a brief description of RA. We then discuss the proposed extension of RA in detail.

### 2.1. Redundancy analysis

Let  $\mathbf{Z}_1$  denote an  $n \times r$  matrix of dependent or endogenous variables. Let  $\mathbf{Z}_2$  denote an  $n \times t$  matrix of independent or exogenous variables. Then, the model for RA may be written as

$$\begin{aligned}\mathbf{Z}_1 &= \mathbf{Z}_2 \mathbf{W} \mathbf{A} + \mathbf{E} \\ &= \mathbf{F} \mathbf{A} + \mathbf{E},\end{aligned}\tag{1}$$

with

$$\text{rank}(\mathbf{W} \mathbf{A}) \leq \min(r, t),\tag{2}$$

where  $\mathbf{W}$  is a  $t \times d$  matrix of component weights,  $\mathbf{A}$  is a  $d \times r$  matrix of component loadings,  $\mathbf{E}$  is an  $n \times r$  matrix of residuals, and  $\mathbf{F} (= \mathbf{Z}_2 \mathbf{W})$  is an  $n \times d$  matrix of component scores. Due to the rank restriction in (2), this model is also called the reduced-rank regression model (Anderson, 1951; Izenman, 1975; Davies and Tso, 1982; Reinsel and Velu, 1998).

To illustrate further, the RA model may be expressed in terms of a path diagram, a pictorial representation of a system of equations, as given in Fig. 1.

In the figure, square boxes are used to indicate observed variables, circles are used to represent components, and arrows are used to signify that the variable at the base of an arrow affects the variable at the head of the arrow. The path diagram implies that the RA

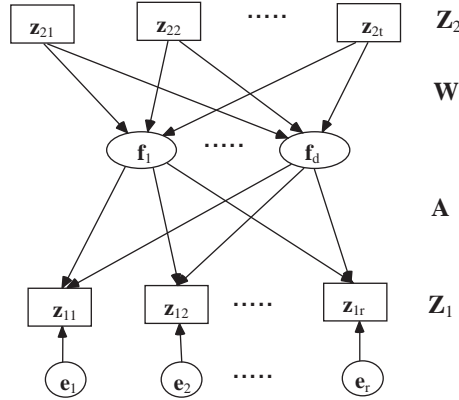


Fig. 1. A path diagram for the RA model.

model may be viewed as a simple type of structural equation model (SEM) between two sets of variables, where latent variables are specified as components of  $\mathbf{Z}_2$  that influence  $\mathbf{Z}_1$ . This RA model is comparable to the multiple indicators/multiple causes (MIMIC) model (Jöreskog and Goldberger, 1975). Let  $\mathbf{F}^* = \mathbf{Z}_2\mathbf{W} + \mathbf{D}$ , where  $\mathbf{D}$  is an  $n \times d$  matrix of residuals. The MIMIC model is given by  $\mathbf{Z}_1 = \mathbf{F}^*\mathbf{A} + \mathbf{E}$ , or equivalently  $\mathbf{Z}_1 = (\mathbf{Z}_2\mathbf{W} + \mathbf{D})\mathbf{A} + \mathbf{E} = \mathbf{Z}_2\mathbf{WA} + (\mathbf{DA} + \mathbf{E}) = \mathbf{Z}_2\mathbf{WA} + \mathbf{E}^*$ , where  $\mathbf{E}^* = \mathbf{DA} + \mathbf{E}$ . Thus, the RA model can be viewed as a special case of the MIMIC model, where  $\mathbf{D} = \mathbf{0}$  (also see Reinsel and Velu, 1998, pp. 17–18).

To estimate parameters in RA, we seek to minimize the following least-squares (LS) criterion

$$f = \text{SS}(\mathbf{Z}_1 - \mathbf{Z}_2\mathbf{WA}), \quad (3)$$

with respect to  $\mathbf{W}$  and  $\mathbf{A}$ , where  $\text{SS}(\mathbf{X}) = \text{trace}(\mathbf{X}'\mathbf{X})$ . In (3), it is assumed that  $\mathbf{W}'\mathbf{Z}_2'\mathbf{Z}_2\mathbf{W} = \mathbf{F}'\mathbf{F} = \mathbf{I}$  for identification. Minimizing (3) computationally boils down to calculating the generalized singular value decomposition (GSVD) of  $(\mathbf{Z}_2'\mathbf{Z}_2)^{-1}\mathbf{Z}_2'\mathbf{Z}_1$  with metric matrices  $\mathbf{Z}_2'\mathbf{Z}_2$  and  $\mathbf{I}$  (Takane and Shibayama, 1991). (For the computation of GSVD, refer to Greenacre, 1984, Appendix A.) This indicates that ordinary RA between two data sets has a closed-form solution.

## 2.2. The ERA model

We now extend RA to accommodate more diverse relationships among multiple sets of variables. Let  $\mathbf{Z}^{(1)}$  denote an  $n \times p$  matrix consisting of endogenous variables. Let  $\mathbf{Z}^{(2)}$  denote an  $n \times q$  matrix consisting of exogenous variables. When a variable is exogenous as well as endogenous, it is included in both  $\mathbf{Z}^{(1)}$  and  $\mathbf{Z}^{(2)}$ . Assume that both  $\mathbf{Z}^{(1)}$  and  $\mathbf{Z}^{(2)}$  are (columnwise) standardized. Then, the model of extended redundancy analysis may be

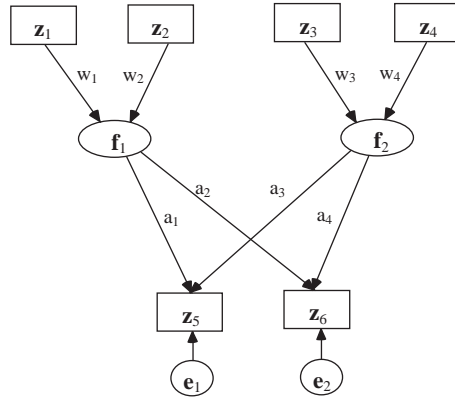


Fig. 2. An exemplary ERA model among three data sets.

generally expressed as

$$\begin{aligned}\mathbf{Z}^{(1)} &= \mathbf{Z}^{(2)}\mathbf{W}\mathbf{A} + \mathbf{E} \\ &= \mathbf{F}\mathbf{A} + \mathbf{E},\end{aligned}\quad (4)$$

with

$$\text{rank}(\mathbf{W}\mathbf{A}) \leq \min(q, p), \quad (5)$$

where  $\mathbf{W}$  denotes a  $q \times d$  matrix of component weights,  $\mathbf{A}$  denotes a  $d \times p$  matrix of component loadings,  $\mathbf{E}$  denotes an  $n \times p$  matrix of residuals,  $\mathbf{F}(=\mathbf{Z}^{(2)}\mathbf{W})$  denotes an  $n \times d$  matrix of component scores. To remove scale indeterminacy between  $\mathbf{F}$  and  $\mathbf{A}$ ,  $\mathbf{F}$  is restricted to be  $\text{diag}(\mathbf{F}'\mathbf{F}) = \mathbf{I}$ . In (4),  $\mathbf{W}$  and/or  $\mathbf{A}$  are structured according to the specified model.

To illustrate further, we suppose there are three sets of variables:  $\mathbf{Z}_1 = [\mathbf{z}_1, \mathbf{z}_2]$ ,  $\mathbf{Z}_2 = [\mathbf{z}_3, \mathbf{z}_4]$ , and  $\mathbf{Z}_3 = [\mathbf{z}_5, \mathbf{z}_6]$ . We further suppose a certain relationship among the three sets of variables, as displayed in Fig. 2.

Fig. 2 shows that two components, one obtained from  $\mathbf{Z}_1$  (i.e.,  $\mathbf{f}_1$ ), and the other from  $\mathbf{Z}_2$  (i.e.,  $\mathbf{f}_2$ ), are combined and affect  $\mathbf{Z}_3$ . This relationship can be expressed as

$$\begin{aligned}\mathbf{Z}_3 &= [\mathbf{Z}_1, \mathbf{Z}_2] \begin{bmatrix} w_1 & 0 \\ w_2 & 0 \\ 0 & w_3 \\ 0 & w_4 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \mathbf{E} \\ &= \mathbf{Z}^{(2)}\mathbf{W}\mathbf{A} + \mathbf{E} \\ &= \mathbf{F}\mathbf{A} + \mathbf{E},\end{aligned}\quad (6)$$

where

$$\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2], \quad \mathbf{W} = \begin{bmatrix} w_1 & w_2 & 0 & 0 \\ 0 & 0 & w_3 & w_4 \end{bmatrix}', \quad \mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix},$$

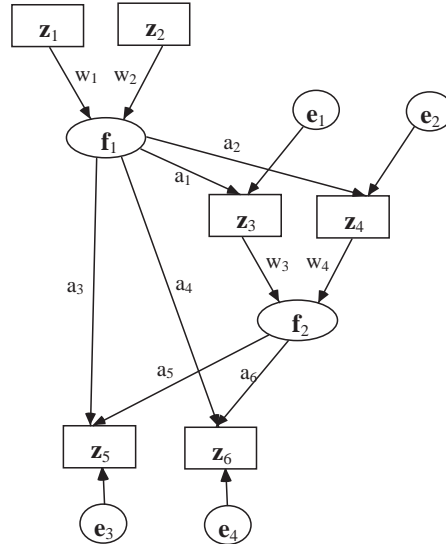


Fig. 3. An exemplary ERA model among three data sets.

and

$$\mathbf{F} = \mathbf{Z}^{(2)} \mathbf{W} = [\mathbf{f}_1, \mathbf{f}_2].$$

Let us consider another relationship among the three sets of variables, as presented in Fig 3.

In Fig. 3, we see that  $\mathbf{f}_1$  has an effect on  $\mathbf{Z}_2$  in addition to the effect on  $\mathbf{Z}_3$ . This relationship can be expressed as

$$\begin{aligned} [\mathbf{Z}_2, \mathbf{Z}_3] &= [\mathbf{Z}_1, \mathbf{Z}_2] \begin{bmatrix} w_1 & 0 \\ w_2 & 0 \\ 0 & w_3 \\ 0 & w_4 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & a_5 & a_6 \end{bmatrix} + \mathbf{E} \\ &= \mathbf{Z}^{(2)} \mathbf{W} \mathbf{A} + \mathbf{E} \\ &= \mathbf{F} \mathbf{A} + \mathbf{E}, \end{aligned} \quad (7)$$

where

$$\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4], \mathbf{W} = \begin{bmatrix} w_1 & w_2 & 0 & 0 \\ 0 & 0 & w_3 & w_4 \end{bmatrix}',$$

and

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & a_5 & a_6 \end{bmatrix}.$$

Model (4) reduces to the RA model when no variables are shared by  $\mathbf{Z}^{(1)}$  and  $\mathbf{Z}^{(2)}$  (i.e.,  $p = r$  and  $q = t$ ), and no constraints other than (5) are imposed on  $\mathbf{W}$  and  $\mathbf{A}$ . Velu (1991)

proposed an extended model of RA (also see [Reinsel and Velu, 1998](#), p. 73). His model is identical to (6), which is a special case of the proposed method. In (6) and (7), only a single component is extracted from  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  each, that is,  $\mathbf{f}_1$  and  $\mathbf{f}_2$ , respectively. If more than one component are obtained from each set, we simply replace the vectors of component weights, say  $\mathbf{w}_1 = [w_1, w_2]$ , and  $\mathbf{w}_2 = [w_3, w_4]$ , by matrices of weights, imposing mutual orthogonality of the components.

### 2.3. Parameter estimation

We estimate the unknown parameters in (4), i.e.,  $\mathbf{W}$  and  $\mathbf{A}$ , in such a way that the sum of squares of the residuals  $\mathbf{E} = \mathbf{Z}^{(1)} - \mathbf{Z}^{(2)}\mathbf{W}\mathbf{A}$  is as small as possible. This is equivalent to minimizing

$$\begin{aligned} f &= \text{SS}(\mathbf{Z}^{(1)} - \mathbf{Z}^{(2)}\mathbf{W}\mathbf{A}) \\ &= \text{SS}(\mathbf{Z}^{(1)} - \mathbf{F}\mathbf{A}), \end{aligned} \quad (8)$$

with respect to  $\mathbf{W}$  and  $\mathbf{A}$ , subject to  $\text{diag}(\mathbf{F}'\mathbf{F}) = \mathbf{I}$ .

As has been shown in (6) and (7),  $\mathbf{W}$  and/or  $\mathbf{A}$  may contain prescribed, fixed (zero) elements. Due to the fixed parameters, (8) cannot be solved in a closed form. Instead, it should be minimized by an iterative method. We use an alternating least squares (ALS) algorithm to minimize (8). Our algorithm is a simple adaptation of the ALS algorithm developed by [Kiers and ten Berge \(1989\)](#), which was used for simultaneous components analysis for two or more populations ([Millsap and Meredith, 1988](#)). In the algorithm, parameter matrices  $\mathbf{W}$  and  $\mathbf{A}$  are alternately updated until convergence. The updates of one parameter matrix are obtained in such a way that they minimize (8), with the other set fixed. The ALS algorithm is monotonically convergent ([de Leeuw et al., 1976](#)). Kiers and ten Berge's algorithm reduces to the algorithm developed here when  $\mathbf{Z}^{(1)}$  and  $\mathbf{Z}^{(2)}$  consist of distinct variables in a single population.

To use the ALS algorithm, we rewrite (8) as

$$f = \text{SS}(\text{vec}(\mathbf{Z}^{(1)}) - \text{vec}(\mathbf{Z}^{(2)}\mathbf{W}\mathbf{A})) \quad (9.a)$$

$$= \text{SS}(\text{vec}(\mathbf{Z}^{(1)}) - (\mathbf{A}' \otimes \mathbf{Z}^{(2)}) \text{vec}(\mathbf{W})) \quad (9.b)$$

$$= \text{SS}(\text{vec}(\mathbf{Z}^{(1)}) - (\mathbf{I} \otimes \mathbf{F}) \text{vec}(\mathbf{A})), \quad (9.c)$$

where  $\text{vec}(\mathbf{X})$  denotes a supervector consisting of all columns of  $\mathbf{X}$  one below another, and  $\otimes$  denotes a Kronecker product.

The algorithm consists of two main steps. In the first step, we update  $\mathbf{W}$  for fixed  $\mathbf{A}$ . (To compute an initial estimate of  $\mathbf{W}$ ,  $\mathbf{A}$  is initialized with arbitrary values.) There could be quite a few fixed elements in  $\text{vec}(\mathbf{W})$ , depending on the model to be fitted. Let  $\mathbf{w}$  denote the  $l \times 1$  vector formed by eliminating fixed elements from  $\text{vec}(\mathbf{W})$ , where  $l$  is the number of free parameters in  $\mathbf{W}$ . Let  $\mathbf{\Omega}$  denote the  $np \times l$  matrix formed by eliminating the columns of  $\mathbf{A}' \otimes \mathbf{Z}^{(2)}$  corresponding to the fixed elements in  $\text{vec}(\mathbf{W})$ . Then, we obtain the LS estimate of  $\mathbf{w}$  by

$$\hat{\mathbf{w}} = (\mathbf{\Omega}'\mathbf{\Omega})^{-1}\mathbf{\Omega}' \text{vec}(\mathbf{Z}^{(1)}), \quad (10)$$

assuming that  $\mathbf{\Omega}'\mathbf{\Omega}$  is nonsingular. (The regular inverse may be replaced by the Moore–Penrose inverse if  $\mathbf{\Omega}'\mathbf{\Omega}$  is singular.) We can simply reconstruct  $\text{vec}(\mathbf{W})$  from  $\mathbf{w}$  by putting back the fixed elements to their original positions, and the updated  $\mathbf{W}$  from  $\text{vec}(\mathbf{W})$ . We then obtain  $\mathbf{F} = \mathbf{Z}^{(2)}\mathbf{W}$  and normalize it so that  $\text{diag}(\mathbf{F}'\mathbf{F}) = \mathbf{I}$ . When more than one component are extracted from same sets of exogenous variables, the columns of  $\mathbf{F}$  are orthonormalized by the Gram–Schmidt orthonormalization method: Let  $\mathbf{F}'\mathbf{F} = \mathbf{M}\mathbf{M}'$ . The orthonormalized  $\mathbf{F}$  is then obtained by  $\mathbf{F}\mathbf{M}^{-1}$ .

In the next step,  $\mathbf{A}$  is updated for fixed  $\mathbf{W}$ . Let  $\mathbf{a}$  denote the  $c \times 1$  vector formed by eliminating fixed elements from  $\text{vec}(\mathbf{A})$ , where  $c$  is the number of free parameters in  $\mathbf{A}$ . Let  $\mathbf{\Gamma}$  denote the  $np \times c$  matrix formed by eliminating the columns of  $\mathbf{I} \otimes \mathbf{F}$  corresponding to the fixed elements in  $\text{vec}(\mathbf{A})$ . We then obtain the LS estimate of  $\mathbf{a}$  by

$$\hat{\mathbf{a}} = (\mathbf{\Gamma}'\mathbf{\Gamma})^{-1}\mathbf{\Gamma}'\text{vec}(\mathbf{Z}^{(1)}), \quad (11)$$

assuming that  $\mathbf{\Gamma}'\mathbf{\Gamma}$  is nonsingular. (Again, the regular inverse may be replaced by the Moore–Penrose inverse if  $\mathbf{\Gamma}'\mathbf{\Gamma}$  is singular.) We can also easily recover  $\mathbf{A}$  from  $\mathbf{a}$ . The above two steps are alternated until convergence is reached.

A few remarks concerning the proposed algorithm are in order. First of all, criterion (8) or equivalently (9) is consistently minimized by this algorithm. However, the algorithm does not guarantee that the obtained minimum is a global minimum. This so-called convergence to non-global minimum problem may be avoided in two ways (e.g., [ten Berge, 1993](#)). When we choose good (or rational) initial values, the function value is likely to start near the global minimum, and it is more likely to obtain the global minimum. Here, we compute the PCA solutions of  $\mathbf{Z}^{(2)}$ , and use the principal coordinates as the rational start for  $\mathbf{W}$ . Then, initial values of  $\mathbf{A}$  are simply obtained by the LS estimate, given  $\mathbf{W}$ . The second possible remedy against the non-global minimum solution is to repeat the ALS optimization procedure with many random initial starts. The obtained function values after convergence are compared, and the solution associated the smallest function value is chosen as the best one.

When  $n$  is large relative to  $q$ , the above algorithm may be made more efficient by the following procedure. Let  $\mathbf{Z}^{(2)} = \mathbf{Q}\mathbf{R}'$  be portions of the QR decomposition of  $\mathbf{Z}^{(2)}$ , pertaining to the column space of  $\mathbf{Z}^{(2)}$ , where  $\mathbf{Q}$  is an  $n \times q$  semi-orthonormal matrix, so that  $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$ , and  $\mathbf{R}'$  is a  $q \times n$  upper-triangular matrix. Then, (8) can be rewritten as follows:

$$\begin{aligned} f &= \text{SS}(\mathbf{Z}^{(1)} - \mathbf{Q}\mathbf{R}\mathbf{B}) \\ &= \text{SS}(\mathbf{Z}^{(1)} - \mathbf{Q}\mathbf{Q}'\mathbf{Z}^{(1)} + \mathbf{Q}\mathbf{Q}'\mathbf{Z}^{(1)} - \mathbf{Q}\mathbf{R}\mathbf{B}) \\ &= \text{SS}(\mathbf{Z}^{(1)} - \mathbf{Q}\mathbf{Q}'\mathbf{Z}^{(1)}) + \text{SS}(\mathbf{Q}(\mathbf{Q}'\mathbf{Z}^{(1)} - \mathbf{R}\mathbf{B})), \end{aligned} \quad (12)$$

where  $\mathbf{B} = \mathbf{W}\mathbf{A}$ . The first term of the right-hand side in (12) does not depend on  $\mathbf{B}$ , and minimizing (12) reduces to minimizing

$$f^* = \text{SS}(\mathbf{Q}(\mathbf{Q}'\mathbf{Z}^{(1)} - \mathbf{R}'\mathbf{B})) = \text{SS}(\mathbf{Q}'\mathbf{Z}^{(1)} - \mathbf{R}'\mathbf{B}). \quad (13)$$

It is computationally more efficient to minimize (13) than (8) because the size of  $\mathbf{R}'$  is usually much smaller than that of  $\mathbf{Z}^{(2)}$ . Moreover, this procedure permits us to use covariance or correlation matrices instead of data matrices since  $\mathbf{Z}^{(2)'}\mathbf{Z}^{(2)} = \mathbf{R}\mathbf{R}'$  and  $\mathbf{Z}^{(2)'}\mathbf{Z}^{(1)} = \mathbf{R}\mathbf{Q}'\mathbf{Z}^{(1)}$ , so that  $\mathbf{Q}'\mathbf{Z}^{(1)}$  can be obtained by  $\mathbf{R}^{-1}\mathbf{Z}^{(2)'}\mathbf{Z}^{(1)}$ . This is beneficial when data are available only in the form of covariance or correlation matrices (which is often the case in the published

literature). When we analyze covariance or correlation matrices, however, there is no way to obtain  $\mathbf{F}$  (Takane et al., 1995).

In ERA, the total fit of a hypothesized model to data is measured by the total variance of the observed endogenous variables explained by the exogenous variables, or equivalently, by the total variance of the endogenous variables minus its unexplained variance. This is given by

$$\text{FIT} = 1 - \frac{SS(\mathbf{Z}^{(1)} - \mathbf{Z}^{(2)}\mathbf{W}\mathbf{A})}{SS(\mathbf{Z}^{(1)})}. \quad (14)$$

This fit index ranges from 0 to 1. The larger the fit value, the more the variance of the endogenous variables explained by the exogenous variables.

Resampling methods such as the jack-knife and the bootstrap methods may be employed to estimate the standard errors of parameter estimates. In ERA, the bootstrap method (Efron, 1982; Efron and Tibshirani, 1998) is used to estimate the standard errors. The bootstrapped standard errors can be used to assess the reliability of the parameter estimates. The critical ratio (CR), i.e., a parameter estimate divided by its standard error, can be used to test the significance of the parameter estimate (e.g., a parameter estimate having a CR greater than two in absolute value is considered significant at a .05 level).

#### 2.4. Imposing additional constraints on parameters

It is often of interest to examine structural hypotheses regarding model parameters. For instance, one may examine a hypothesis that some elements in  $\mathbf{W}$  are equal to each other, or some elements in  $\mathbf{A}$  are identical to each other, and so on. A variety of structural hypotheses on model parameters may be incorporated in the form of linear constraints (Böckenholt and Böckenholt, 1990; Böckenholt and Takane, 1994; Takane and Shibayama, 1991; Takane et al., 1991; ter Braak, 1986; Yanai, 1986).

The linear constraints may be specified by either the reparameterization or the null-space method (Böckenholt and Takane, 1994; Takane et al., 1991). The former method specifies the space spanned by column vectors of a constraint matrix, while the latter specifies its ortho-complement space. The choice of the reparameterization method or the null-space method depends solely on specific empirical interests of the investigator (Takane et al., 1991). The reparameterization method appears to be a more natural choice when parameters need to be directly constrained, while the null-space method seems to be a more natural choice when the effects of some variables need to be eliminated from other variables (Böckenholt and Takane, 1994). Nonetheless, the reparameterization method appears to have some computational advantage over the null-space method in that it may involve a fewer number of parameters to be estimated directly compared to the null-space method (Takane et al., 1991).

In ERA, all linear constraints are imposed by the reparameterization method. Let  $\mathbf{H}$  denote a matrix of linear constraints on  $\mathbf{a}$ . In the final step of the ALS algorithm, we incorporate  $\mathbf{H}$  into  $\mathbf{a}$  as follows:

$$\mathbf{a} = \mathbf{H}\boldsymbol{\alpha}, \quad (15)$$



for some  $\alpha$ . The LS estimate of  $\alpha$  is then given by

$$\hat{\alpha} = (\mathbf{H}'\Gamma'\Gamma\mathbf{H})^{-1}\mathbf{H}'\Gamma'\text{vec}(\mathbf{Z}^{(1)}), \quad (16)$$

which leads to

$$\hat{\mathbf{a}} = \mathbf{H}\hat{\alpha} = \mathbf{H}(\mathbf{H}'\Gamma'\Gamma\mathbf{H})^{-1}\mathbf{H}'\Gamma'\text{vec}(\mathbf{Z}^{(1)}). \quad (17)$$

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 reparameterization form. The transformation is straightforward. Let

$$\mathbf{P}'\mathbf{a} = \mathbf{0} \quad (18)$$

represent the constraints in the null-space form. Suppose that the first and the last elements of  $\mathbf{a}$  are equal; then,  $\mathbf{P}'$  becomes a vector whose first element is 1, the last element is  $-1$ , and the other elements are zeros. We may reparameterize (18) into the form of (15) by defining  $\mathbf{H} = \mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-}\mathbf{P}'$ , where  $(\mathbf{P}'\mathbf{P})^{-}$  is a g-inverse of  $\mathbf{P}'\mathbf{P}$ . This implies that  $\text{Ker}(\mathbf{P}') = \text{Sp}(\mathbf{H})$ , where  $\text{Ker}(\mathbf{P}')$  denotes the null space of  $\mathbf{P}'$ , and  $\text{Sp}(\mathbf{H})$  denotes the space spanned by the column vectors of  $\mathbf{H}$ . Linear constraints can be imposed on  $\mathbf{w}$  in a similar way. The validity of certain hypotheses may be empirically investigated by comparing fits of the constrained and unconstrained solutions. The standard errors and critical ratios of the obtained parameter estimates can also be used to evaluate the hypotheses.

### 2.5. Relationships with two methods for SEM

The extended RA models considered herein may be viewed as special kinds of structural equation models where (1) endogenous variables are always observed ones; (2) unobserved or latent variables are approximated by components of observed exogenous variables; and (3) a formative relationship is always assumed between observed and latent variables (i.e., observed variables determine latent variables). In principle, therefore, two existing methods for structural equation models can also be employed to fit the extended RA models: analysis of covariance structures (ACOVs) and partial least squares (PLS).

Nonetheless, PLS (Wold, 1966, 1973, 1982; Lohmöller, 1989) appears to be more suitable to fit extended RA models than ACOVS (Bock and Bargmann, 1966; Jöreskog, 1970, 1973, 1977) for several reasons. First of all, PLS treats latent variables as components of observed variables while ACOVS as common factors (Velicer and Jackson, 1990). Also, PLS directly analyzes a data matrix whereas ACOVS analyzes a covariance or correlation matrix. Furthermore, PLS is capable of handling both formative and reflective (i.e., observed variables are caused by latent variables) relationships between observed and latent variables. In ACOVS, on the other hand, the reflective relationship is typically assumed because it is consistent with the ACOVS statistical algorithm based on all the covariances among observed variables (Chin, 1998). Modeling a formative relationship in ACOVS requires certain conditions (e.g., a latent variable defined by linear combinations of observed variables needs to emit at least two paths to different latent variables) in order to avoid some identification problems (MacCallum and Browne, 1993). Yet, it seems difficult to satisfy those conditions in extended RA models where endogenous variables are always observed ones.

For fitting extended RA models, ERA and PLS share the same benefits. For example, both methods can avoid the problem of improper solutions (e.g., latent variable correlation estimates greater than  $\pm 1$ , negative variance estimates, etc.) due to replacing latent variables by components. They can also be free from stringent distributional assumptions, such as multivariate normality of observed variables, since they employ LS procedures for parameter estimation: ERA uses an ALS algorithm and PLS uses a fixed point algorithm (Lyttekens, 1968, 1973; Wold, 1965, 1981), where a set of parameters are divided into subsets, and each subset is partially estimated by ordinary multiple regression analysis with other subsets fixed.

Despite such benefits, PLS has a substantial limitation when fitting extended RA models. PLS does not solve a global optimization problem for parameter estimation (Jöreskog and Wold, 1982; Fornell and Bookstein, 1982), although there were a few attempts to alleviate this problem (e.g., Bookstein, 1982; Hanafi and Qannari, 2002). This indicates that there exists no criterion consistently minimized or maximized to determine parameter estimates. The lack of a global optimization criterion makes it difficult to evaluate the PLS estimation procedures (McDonald, 1996). More seriously, PLS has no mechanism to evaluate the overall fit of the model. Given no overall goodness of fit measures, it is difficult to examine how well the hypothesized model fits to the data and also to compare it with alternative models. On the other hand, ERA employs a well-defined global LS criterion for parameter estimation, thereby allowing an overall model fit measure. Due to the global optimization criterion, ERA can also deal with multi-sample comparisons with optional impositions of across-sample constraints, which PLS cannot accommodate (this multi-sample analysis will be discussed in a later section). Therefore, we conclude that the proposed method appears to be superior to PLS in handling extended RA models.

## 2.6. A simulation study

A simulation study was conducted to investigate the small sample behavior of ERA. In particular, the study focused on how well ERA performed in terms of recovery of the original underlying structure of the data in small samples. For this purpose, the model of (6) was chosen with the following parameter values:

$$\mathbf{W} = \begin{bmatrix} .6 & 0 \\ .6 & 0 \\ 0 & .6 \\ 0 & .6 \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} .2 & .2 \\ .2 & .2 \end{bmatrix}. \quad (19)$$

To generate  $\mathbf{Z}^{(1)}$  and  $\mathbf{Z}^{(2)}$ , it was further assumed that  $\mathbf{E} \sim N(\mathbf{0}, \mathbf{\Sigma})$  and  $\mathbf{Z}^{(2)} \sim N(\mathbf{0}, \mathbf{\Pi})$ , where  $\mathbf{\Sigma}$  and  $\mathbf{\Pi}$  were pre-specified as

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & .1 \\ .1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{\Pi} = \begin{bmatrix} 1 & .3 & .1 & .1 \\ .3 & 1 & .1 & .1 \\ .1 & .1 & 1 & .3 \\ .1 & .1 & .3 & 1 \end{bmatrix}. \quad (20)$$

The specification of  $\mathbf{\Pi}$  indicates that the within-set correlation between observed exogenous variables is equal to .3, while the between-set correlation is equal to 0.1.

In the simulation study, sample sizes varied as follows:  $n = 50, 100, 200$ , and  $400$ . For each sample size, 1000 Monte Carlo samples were obtained. Matrix  $\mathbf{Z}^{(2)}$  was created only once for each sample size, and considered fixed for the entire samples. The true parameter values were used as the initial values for the ALS algorithm so as to increase the speed of convergence as well as the possibility of convergence to the global minimum.

The proposed algorithm was found to converge in all cases. Table 1 shows the mean parameter estimates, estimated standard errors, and relative biases ( $100 \times [\text{parameter estimate} - \text{population parameter}] / \text{population parameter}$ ) obtained for different sample sizes. Table 1 also provides the mean congruence coefficient (Tucker, 1951) between parameters and estimates for each sample size as a recovery measure, which is defined as follows: Let  $\boldsymbol{\theta}$  and  $\mathbf{p}$  denote the vectors of the population parameters and the parameter estimates obtained from a single Monte Carlo sample, respectively. Then, the congruence coefficient is  $(\boldsymbol{\theta}'\mathbf{p})/(\sqrt{\boldsymbol{\theta}'\boldsymbol{\theta}}\sqrt{\mathbf{p}'\mathbf{p}})$ . The congruence coefficient indicates the degree of similarity between the population parameters and the parameter estimates.

In Table 1, when  $n = 50$ , the estimates of  $\mathbf{W}$  show somewhat large relative biases (greater than 10%) and large standard errors. On the other hand, those of  $\mathbf{A}$  appear to show acceptably small relative biases (less than 10%) even though their standard errors are large. The mean congruence coefficient (.71) is smaller than .90, which is a conventional rule of thumb criterion as an acceptable degree of congruence (Mulaik, 1972). On the other hand, when  $n \geq 100$ , the estimates of both  $\mathbf{W}$  and  $\mathbf{A}$  seem to show small relative biases and small standard errors. Moreover, the mean congruence coefficients appear to be close to .90 (.87 when  $n = 100$ ) or greater than .90 for  $n \geq 200$ . These simulation results suggest that ERA seems to recover population parameters sufficiently well unless a sample size is too small ( $n \leq 50$ ).

This simulation study provides useful information on the performance of ERA in small samples. Nonetheless, the scope of the study is admittedly limited since it focuses mainly on the capability of recovering the original data structure on the basis of a single, simple model. To evaluate the performance of ERA from more diverse perspectives, further simulation studies seem to be needed, considering a more variety of experimental conditions such as model complexity, distributional assumptions, model misspecification, and so forth.

## 2.7. An illustrative example: the basic health indicator data

In this section, we present an example to demonstrate the empirical use of ERA. This example is part of the so-called basic health indicator data collected by the World Health Organization in the United Nations. They appear in the 1999 World Health Report and are also available through the internet (<http://www.who.int>). From the entire data set, we only used six variables measured in different countries. The six observed variables were as follows: (1) infant mortality rate (IMR), defined as the number of deaths per 1000 live births between birth and exact age of one year in 1998, (2) maternal mortality ratio (MMR), defined as the number of maternal deaths per 100,000 live births in 1990, (3) real gross domestic product (GDP) per capita adjusted for purchasing power parity, expressed in 1985 US dollars, (4) the average number of years of education given for females aged 25 years and above (FEUD), (5) the percentage of children immunized against measles in 1997 (Measles), and (6) total health expenditures as a percentage of GDP in 1995 (Healthexp).

Table 1

Mean parameter estimates, mean standard errors, and mean relative biases for an ERA model in different sample sizes (1000 Monte Carlo samples per sample size)

par	<i>n</i> = 50			<i>n</i> = 100			<i>n</i> = 200			<i>n</i> = 400		
	est	s.e.	rbias	est	s.e.	rbias	est	s.e.	rbias	est	s.e.	rbias
$w_1 = .6$	.420	.519	−29.951	.592	.335	−1.285	.629	.255	4.882	.596	.169	−.727
$w_2 = .6$	.518	.486	−13.657	.558	.341	−7.038	.561	.265	−6.527	.605	.170	.901
$w_3 = .6$	.513	.476	−14.431	.586	.319	−2.297	.592	.255	−1.403	.630	.162	5.073
$w_4 = .6$	.456	.502	−24.032	.550	.334	−8.382	.587	.265	−2.246	.566	.169	−5.734
$a_1 = .2$	.197	.147	−1.616	.200	.102	0.022	.181	.072	−9.467	.189	.047	−5.752
$a_2 = .2$	.211	.150	5.655	.214	.101	7.216	.184	.069	−7.843	.194	.047	−2.931
$a_3 = .2$	.203	.147	1.309	.206	.100	2.964	.183	.069	−8.434	.193	.048	−3.638
$a_4 = .2$	.207	.149	3.711	.219	.099	9.366	.192	.069	−3.937	.200	.047	.198
Congruence = .71			Congruence = .87			Congruence = .92			Congruence = .96			

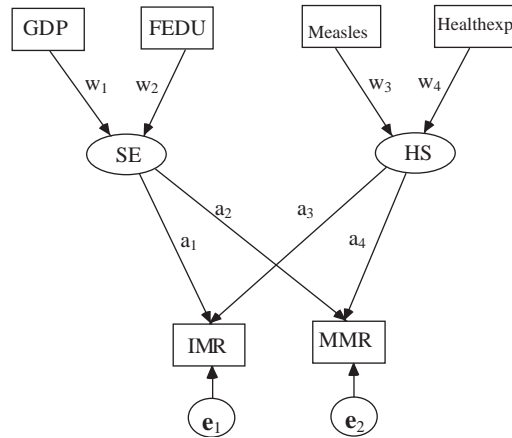


Fig. 4. The ERA model specified for the WHO data.

Table 2

Parameter estimates and their standard errors (in the parentheses) obtained from the unconstrained and constrained ERA models for the WHO data

	Unconstrained	Constrained
FIT	.6512	.6491
GDP $\rightarrow$ SE ( $w_1$ )	-.50 (.17)	-.49 (.16)
FEDU $\rightarrow$ SE ( $w_2$ )	-.57 (.16)	-.57 (.16)
Measles $\rightarrow$ HS ( $w_3$ )	-.96 (.12)	-1.00 (.00)
Healthexp $\rightarrow$ HS ( $w_4$ )	-.13 (.24)	0
SE $\rightarrow$ IMR ( $a_1$ )	.58 (.10)	.61 (.11)
SE $\rightarrow$ MMR ( $a_2$ )	.43 (.10)	.47 (.13)
HS $\rightarrow$ IMR ( $a_3$ )	.41 (.09)	.40 (.11)
HS $\rightarrow$ MMR ( $a_4$ )	.45 (.11)	.43 (.15)

The sample size was 51, which represented the number of countries for which the data were available.

Two components were assumed for the last four observed variables. One component called ‘social and economic (SE) factor’ was defined as a linear combination of GDP and FEUD, and the other called ‘health services (HS) factor’ as that of Measles and Healthexp. The two components were in turn deemed to influence two observed endogenous variables, IMR and MMR. The specified two-component model is presented in Fig. 4.

Using ERA, this model was fitted to the data. The parameter estimates and their bootstrapped standard errors obtained with 100 bootstrap samples are given in Table 2.

The goodness of fit of the model is equal to .6512, indicating that about two-thirds of the total variance of the endogenous variables is accounted for by the two-component model. The component weights associated with SE are all significant and negative ( $w_1 = -.50$ ,

s.e. = .17, CR = -2.94;  $w_2 = -.57$ , s.e. = .16, CR = -3.56). This suggests that SE is characterized as social and economic underdevelopment. Similarly, the component weights of Measles and Healthexp are negative ( $w_3 = -.96$ , s.e. = .12, CR = -8.0;  $w_4 = -.13$ , s.e. = .24, CR = -.54), indicating that HS is likely to represent a low level of health service. However, only one variable, Measles, seems to be significantly associated with HS. Both components are found to have significant and positive effects on IMR ( $a_1 = .58$ , s.e. = .10, CR = 5.8;  $a_2 = .43$ , s.e. = .10, CR = 4.3) and MMR ( $a_3 = .41$ , s.e. = .09, CR = 4.56;  $a_4 = .45$ , s.e. = .11, CR = 4.09). This suggests that social and economic underdevelopment and the low level of health services are likely to increase infant mortality rate and maternal mortality ratio. The correlation between the two components is equal to .47, which is significant (s.e. = .12, CR = 3.91).

Given the solutions in the two-component model, we further assumed that the component weight for Healthexp was equal to zero (i.e.,  $w_4 = 0$ ). This additional assumption could be incorporated by imposing zero constraints on  $\mathbf{w}$ , as explained in Section 2.4. Let  $\mathbf{p}'$  denote the vector of the zero constraints. The  $\mathbf{p}'$  was defined as [0 0 0 1]. It was required that  $\mathbf{p}'\mathbf{w} = 0$ . Then, the  $\mathbf{p}'$  was transformed into a reparameterization form by the procedure described in Section 2.4.

The results of fitting the constrained model are also presented in Table 2. Due to the zero constraint, the component weight for Healthexp is shown as zero. This indicates that HS is solely defined by Measles. It is thus analogous to eliminating HS from the analysis, and hypothesizing direct effects of Measles on the endogenous variables. Measles is found to have negative direct effects on IMR (-.40) and MMR (-.43). The fit of the constrained model is .6491. This is almost the same as that of the unconstrained model. The correlation between the components becomes somewhat smaller (.41) than that from the unconstrained case. This is due to the elimination of Healthexp from HS. Yet, the correlation turns out to be significant (s.e. = .13, CR = 3.15). Despite the zero constraint, however, the interpretations of the solutions from the constrained model seem to be essentially the same as those from the unconstrained model. Therefore, it may be safe to say that the additional assumption on the component weight for Healthexp is acceptable. This allows for simpler interpretations of the obtained solutions by reducing the number of parameters.

### 3. Some extensions

#### 3.1. Three possible extensions

The ERA model in (4) can be further extended in various ways. In particular, we discuss how to handle higher-order components (i.e., components nested within other components), direct effects of observed variables on other observed variables, and multi-sample comparisons. It is also shown that the extended models can be expressed in essentially the same form as (4), and essentially the same estimation procedure can be applied to fit them.

To include the  $K$ th-order components, the ERA model may be expressed as follows:

$$\mathbf{Z}^{(1)} = \mathbf{Z}^{(2)}\tilde{\mathbf{W}}\mathbf{A} + \mathbf{E}, \quad (21)$$

where

$$\tilde{\mathbf{W}} = \prod_{k=1}^K \mathbf{W}^{(k)}. \quad (22)$$

In (22),  $\mathbf{W}^{(k)}$  denotes the matrix of component weights for the  $k$ th-order components ( $k = 1, \dots, K$ ), and each  $k$ th-order component is restricted to be of unit length for identification. Model (21) is essentially the same as (4). Hence, a similar ALS algorithm can be used to fit the model. In this case, however, a set of parameter matrices are split into  $K + 1$  matrices, that is,  $\mathbf{W}^{(k)}$ 's and  $\mathbf{A}$ , and we update them alternately until convergence is obtained. For instance, an ERA model with second-order components may be written as

$$\begin{aligned} \mathbf{Z}^{(1)} &= \mathbf{Z}^{(2)} \mathbf{W}^{(1)} \mathbf{W}^{(2)} \mathbf{A} + \mathbf{E} \\ &= \mathbf{Z}^{(2)} \tilde{\mathbf{W}} \mathbf{A} + \mathbf{E}, \end{aligned} \quad (23)$$

where  $\tilde{\mathbf{W}} = \mathbf{W}^{(1)} \mathbf{W}^{(2)}$ . To estimate  $\tilde{\mathbf{W}}$  and  $\mathbf{A}$ , in (23), we may update  $\mathbf{W}^{(1)}$  for fixed  $\mathbf{A}$  and  $\mathbf{W}^{(2)}$ , normalize  $\mathbf{Z}^{(2)} \mathbf{W}^{(1)}$ , update  $\mathbf{W}^{(2)}$  for fixed  $\mathbf{A}$  and  $\mathbf{W}^{(1)}$ , normalize  $\mathbf{Z}^{(2)} \tilde{\mathbf{W}}$ , and update  $\mathbf{A}$  for fixed  $\mathbf{Z}^{(2)} \tilde{\mathbf{W}}$  in each iteration. An example of a higher-order component model will be given in the next section.

The ERA model including the direct effects may also be written in the same form as (4). For example, we suppose that in (6),  $\mathbf{z}_1$  of  $\mathbf{Z}^{(2)}$  has a direct effect on  $\mathbf{z}_5$  of  $\mathbf{Z}^{(1)}$  in addition to the effect on  $\mathbf{z}_3$  through  $\mathbf{f}_1$ . To include this effect, we may write the ERA model as follows.

$$\mathbf{Z}^{(1)} = \mathbf{Z}^{(2)} \mathbf{W} \mathbf{A} + \mathbf{E}, \quad (24)$$

where

$$\mathbf{W} = \begin{bmatrix} 1 & w_1 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \\ 0 & 0 & w_4 \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

Model (24) is a simple variant of (4), where the only distinction is in that  $\mathbf{W}$  and  $\mathbf{A}$  contain an extra (first) column and row, respectively, in order to represent a direct effect of  $\mathbf{z}_1$  on  $\mathbf{z}_3$ . In this way, we can readily incorporate the direct effects of any observed variables in  $\mathbf{Z}^{(2)}$ . The supplementary column of  $\mathbf{W}$  consists of all fixed elements, zeros or unities, whereas the corresponding row of  $\mathbf{A}$  has free parameters to be estimated as well as fixed elements. The same ALS estimation procedure is employed to fit (24). To demonstrate the addition of direct effects to the ERA model, an illustrative example will be presented in Section 3.2.

We may also be interested in fitting a single ERA model to more than one sample simultaneously. Such a simultaneous analysis enables us to test various hypotheses concerning the relationships among parameters across samples. Suppose that  $J$  samples are fitted by the same ERA model,

$$\mathbf{Z}_j^{(1)} = \mathbf{Z}_j^{(2)} \mathbf{W}_j \mathbf{A}_j + \mathbf{E}_j = \mathbf{F}_j \mathbf{A}_j + \mathbf{E}_j, \quad (25)$$

where  $\mathbf{F}_j = \mathbf{Z}_j^{(2)} \mathbf{W}_j$  ( $j = 1, \dots, J$ ). Here the structures of  $\mathbf{W}_j$  and  $\mathbf{A}_j$  are identical across  $J$  samples. Model (25) may be re-expressed as

$$\begin{bmatrix} \mathbf{Z}_1^{(1)} \\ \vdots \\ \mathbf{Z}_J^{(1)} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1^{(2)} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{Z}_J^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{W}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{W}_J \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_J \end{bmatrix} + \begin{bmatrix} \mathbf{E}_1 \\ \vdots \\ \mathbf{E}_J \end{bmatrix}. \quad (26)$$

It can thus be expressed as a single equation,

$$\dot{\mathbf{Z}}^{(1)} = \dot{\mathbf{Z}}^{(2)} \dot{\mathbf{W}} \dot{\mathbf{A}} + \dot{\mathbf{E}}, \quad (27)$$

where  $\dot{\mathbf{Z}}^{(1)} = [\mathbf{Z}_1^{(1)'} \dots \mathbf{Z}_J^{(1)'}]'$ ,  $\dot{\mathbf{Z}}^{(2)} = \text{diag}[\mathbf{Z}_1^{(2)}, \dots, \mathbf{Z}_J^{(2)}]$ ,  $\dot{\mathbf{W}} = \text{diag}[\mathbf{W}_1, \dots, \mathbf{W}_J]$ ,  $\dot{\mathbf{A}} = [\mathbf{A}_1' \dots \mathbf{A}_J']'$ , and  $\dot{\mathbf{E}} = [\mathbf{E}_1' \dots \mathbf{E}_J']'$ . Model (27) is essentially the same form as (4). To estimate parameters, the essentially same optimization criterion as (8) is minimized, that is,

$$f = \text{SS}(\dot{\mathbf{Z}}^{(1)} - \dot{\mathbf{Z}}^{(2)} \dot{\mathbf{W}} \dot{\mathbf{A}}), \quad (28)$$

with respect to  $\dot{\mathbf{W}}$  and  $\dot{\mathbf{A}}$ , subject to  $\text{diag}(\mathbf{F}_j' \mathbf{F}_j) = \mathbf{I}$ .

We note that in its most general form (25) or (26) is equivalent to  $J$  separate analyses of (4). Nevertheless, it should be emphasized that various constraints can be imposed on parameters across  $J$  samples, thus allowing evaluation of a variety of hypotheses on the parameters across the different samples. For instance, we can examine whether certain elements of  $\mathbf{W}_j$  and/or  $\mathbf{A}_j$  are identical across different samples by imposing equality constraints. These 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 (Böckenholt and Takane, 1994).

To test structural hypotheses concerning the parameters across  $J$  samples,  $J$  sets of parameters in  $\mathbf{W}_j$  and  $\mathbf{A}_j$  can be regarded as a single set of parameters in  $\dot{\mathbf{W}}$  and  $\dot{\mathbf{A}}$  as in a single sample, so that the same procedure in Section 2.4 can be used. More specifically, we update the vector, say  $\dot{\mathbf{w}}$ , which is formed by eliminating fixed elements from  $\text{vec}(\dot{\mathbf{W}})$ , by using (10), and recover the updated  $\dot{\mathbf{W}}$ . We normalize  $\mathbf{F}_j$  for each sample, such that  $\text{diag}(\mathbf{F}_j' \mathbf{F}_j) = \mathbf{I}$ . Then  $\dot{\mathbf{A}}$  is updated from  $\dot{\mathbf{a}}$  formed by eliminating fixed elements from  $\text{vec}(\dot{\mathbf{A}})$  based on (11). We may easily compare the means of the components,  $\mathbf{F}_j$ , across samples. In this case, however, the unstandardized data should be analyzed instead of the standardized data since the means are a priori eliminated in the standardized data. The set of exogenous variables should also include a constant term. An example of the simultaneous analysis of several samples will be provided in the next section.

### 3.2. Illustrative examples

The basic health indicator data are again used to illustrate the extensions discussed above. This data set is fitted by two extensions of the two-component model shown in Section 2.7: One adds a second-order component and the other a direct effect. A new data set, called social function data, is employed to show the feasibility of simultaneous analysis of several



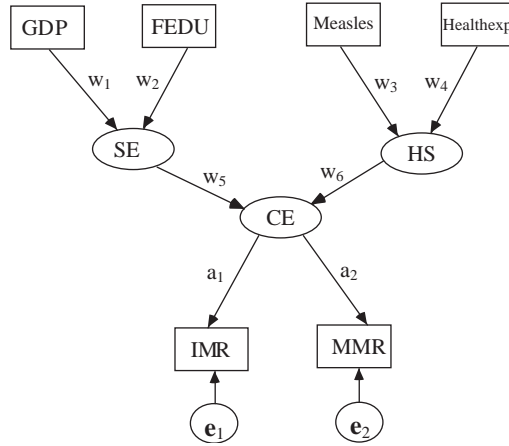


Fig. 5. The second-order component ERA model for the WHO data.

samples or multi-sample comparisons. Note that the two additional models for the basic health indicator data are specified without well-grounded theories or hypotheses to support model specification. They are simply improvised to illustrate the usefulness of the proposed method. On the other hand, an empirically well-motivated model is fitted in the analysis of the second data set.

### 3.2.1. The basic health indicator data

To demonstrate the feasibility of incorporating higher-order components, we supposed a second-order component nested within the two first-order components in the two-component model shown in Fig. 4. The second-order component was named ‘combined effect (CE)’. The specified model structure is given in Fig. 5.

In the model,  $\mathbf{W}^{(1)}$  was analogous to  $\mathbf{W}$  in (6), and  $\mathbf{W}^{(2)}$  and  $\mathbf{A}$  were specified as follows:

$$\mathbf{W}^{(2)} = \begin{bmatrix} w_5 \\ w_6 \end{bmatrix} \quad \text{and} \quad \mathbf{A} = [a_1 \quad a_2].$$

It is noteworthy that this second-order component model turns out to be equivalent to a simple RA model where only one component is assumed from  $\mathbf{Z}^{(2)}$ . Nonetheless, this model may allow additional analyses by imposing constraints on  $\mathbf{W}^{(2)}$  (e.g.,  $w_5 = w_6$ ).

The results of fitting the model are presented in Table 3.

This model shows a slightly worse fit (.6492) compared to that without the second-order component. Both component weights of SE and HS for CE are found to be significant and positive ( $w_5 = .67$ , s.e. = .13, CR = 5.15;  $w_6 = .50$ , s.e. = .13, CR = 3.85). The effects of CE on IMR and MMR also turn out to be significant and positive ( $a_1 = .87$ , s.e. = .02, CR = 43.5;  $a_2 = .77$ , s.e. = .04, CR = 19.25). This suggests that CE is an inverse measure of health supports, so that an increase of CE is likely to increase the possibility of infant and maternal deaths.

Table 3

Parameter estimates and their standard errors (in the parentheses) obtained from the second-order component ERA model for the WHO data

FIT	.6492
GDP $\rightarrow$ SE ( $w_1$ )	-.49 (.18)
FEDU $\rightarrow$ SE ( $w_2$ )	-.57 (.18)
Measles $\rightarrow$ HS ( $w_3$ )	-.96 (.13)
Healthexp $\rightarrow$ HS ( $w_4$ )	-.13 (.23)
SE $\rightarrow$ CE ( $w_5$ )	.67 (.13)
HS $\rightarrow$ CE ( $w_6$ )	.50 (.13)
CE $\rightarrow$ IMR ( $a_1$ )	.87 (.02)
CE $\rightarrow$ MMR ( $a_2$ )	.77 (.04)

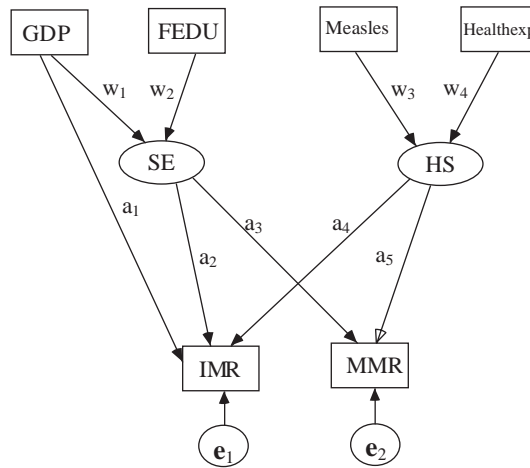


Fig. 6. The ERA model for the WHO data with a direct effect added.

To illustrate the addition of direct effects, we assumed that GDP had a direct effect on IMR in addition to its effect through SE. The specified two-component model with the direct effect is displayed in Fig. 6. The results of fitting the model are presented in Table 4.

The model shows almost the same fit (.6513) as that without the direct effect. This seems to be consistent with the fact that the direct effect of GDP is insignificant ( $a_1 = -.01$ , s.e. = .21, CR = -.05). The effect of SE on IMR also turns out to be insignificant ( $a_2 = .53$ , s.e. = 1.46, CR = .36). It may be due to adding the direct effect of GDP on IMR to the model, yielding a collinearity problem between GDP and SE.

### 3.2.2. The social function data

The second example was obtained from Park's (1996) social function data. Park (1996) studied cultural differences in the functional and structural aspects of social functions between South Korean and German adolescents. She assessed seven social functions, such

Table 4

Parameter estimates and their standard errors (in the parentheses) obtained from the ERA model for the WHO data with a direct effect added

FIT	.6513
GDP $\rightarrow$ SE ( $w_1$ )	-.46 (.25)
FEDU $\rightarrow$ SE ( $w_2$ )	-.60 (.24)
Measles $\rightarrow$ HS ( $w_3$ )	-.96 (.08)
Healthexp $\rightarrow$ HS ( $w_4$ )	-.13 (.21)
GDP $\rightarrow$ IMR ( $a_1$ )	-.01 (.21)
SE $\rightarrow$ IMR ( $a_2$ )	.53 (1.46)
SE $\rightarrow$ MMR ( $a_3$ )	.44 (.10)
HS $\rightarrow$ IMR ( $a_4$ )	.41 (.11)
HS $\rightarrow$ MMR ( $a_5$ )	.45 (.13)

as attachment, self-validation, intimacy, guidance, control, conflict, and comparison, which seemed to play important roles in the development of adolescents. Refer to [Park \(1996\)](#) for detailed explanations about these social functions.

Fifty-nine South Korean adolescents and sixty German adolescents participated in her study. The mean age of the South Korean participants was 14 years and 1 month old and that of the German participants was 14 years and 2 months old. A self-report questionnaire was administered to measure the degree of the social functions to which South Korean and German adolescents would be exposed. The self-report questionnaire consisted of 18 items on adolescents' behaviors related to the seven social functions specified above. A measure of each social function was obtained as a sum of adolescents' responses to the following items: (1) attachment: "Did you feel happy when you did something with him/her together?", "Did you feel happy when he/she gave you something, or you could give him/her something?", "Did you worry about him/her?", (2) self-validation: "Did you gain recognition from him/her?", "Did you gain praise from him/her?", (3) intimacy: "Did you talk to him/her about what happened to you today?", "Did you speak to him/her about your innermost feelings?", (4) guidance: "Did he/she help you solve problems?", "Did the person advise you?", "Did you speak to the person about your problem?", (5) control: "Did the person say to you that you must do something (e.g., studying, making up the room, etc.)?", "Did the person punish you?", (6) conflict: "Did you quarrel with the person?", "Did you hate the person?", (7) comparison: "Did you feel superiority to the person?", "Did you feel inferiority to the person?", "Were you envious of the person?" Besides these items for the seven social functions, the following item was used to obtain an endogenous measure of 'trust': "How often will the person help you if you are in trouble?"

All participants were first asked to make a list of significant persons whom they have contacted at least once per week by various means, including by phone and by mail. The significant persons on the list were considered as social network members. The participants were then asked to answer the 18 items for each person on the list. The number of social network members was 741 for South Korean adolescents and 760 for German adolescents. The number of social network members corresponded to the size of each sample.

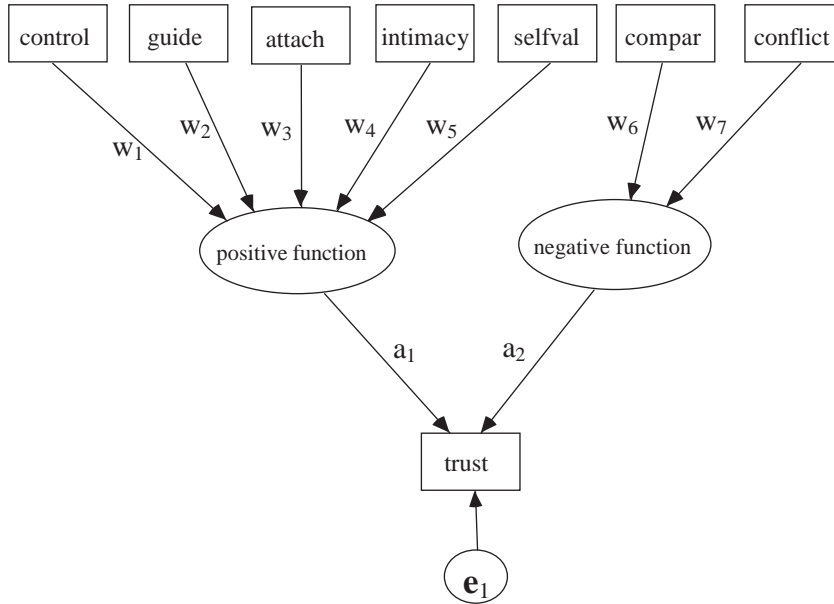


Fig. 7. The ERA model specified for the social function data.

On the basis of her preliminary analyses on the combined sample of South Korean and German adolescents, Park (personal communication, May 1998) has suggested two potential components underlying the seven social functions. One component, called ‘positive function’, was associated with five social functions such as attachment, control, guidance, intimacy, and self-validation, and the other, called ‘negative function’ was associated with two social functions such as comparison and conflict. Hence we assumed a single two-component model for both samples. In the model, the components were specified in the same way as above, and they were assumed to have direct influences on trust of adolescents toward other social network members. The specified model is displayed in Fig. 7.

In Fig. 7,  $\mathbf{Z}_j^{(2)}$  was an  $n_j \times 7$  matrix ( $j = 1, 2$ ), whose first five columns corresponded with the five social functions associated with positive function, whereas the last two corresponded with the two social functions related to negative function. On the other hand,  $\mathbf{Z}_j^{(1)}$  was a column vector of order  $n_j$ , corresponding to an endogenous variable, trust. In the model,  $\mathbf{W}_j$  and  $\mathbf{A}_j$  were specified equally across samples as follows:

$$\mathbf{W}_j = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 & w_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w_6 & w_7 \end{bmatrix}' \quad \text{and} \quad \mathbf{A}_j = [a_1 \quad a_2].$$

Using the multi-sample comparison feature in ERA, we simultaneously fit the two-component model to South Korean and German adolescent samples in order to examine any differences in two sets of parameter estimates. The results are presented in Table 5.

Table 5

Parameter estimates and their standard errors (in the parentheses) obtained from the unconstrained and constrained multi-sample ERA analyses for the social function data (PF = positive function, NF = negative function)

		Unconstrained	Constrained
FIT		.1660	.1658
Korean	Control $\rightarrow$ PF ( $w_1$ )	.40 (.08)	.40 (.06)
	Guide $\rightarrow$ PF ( $w_2$ )	.01 (.08)	.03 (.11)
	Attach $\rightarrow$ PF ( $w_3$ )	.42 (.12)	.45 (.07)
	Intimacy $\rightarrow$ PF ( $w_4$ )	.31 (.08)	.32 (.08)
	Selfval $\rightarrow$ PF ( $w_5$ )	.12 (.09)	.10 (.10)
	Compar $\rightarrow$ NF ( $w_6$ )	.04 (.48)	.00 (.47)
	Conflict $\rightarrow$ NF ( $w_7$ )	.85 (.35)	.94 (.28)
	PF $\rightarrow$ trust ( $a_1$ )	.46 (.03)	.44 (.02)
German	NF $\rightarrow$ trust ( $a_2$ )	-.09 (.03)	-.08 (.03)
	Control $\rightarrow$ PF ( $w_1$ )	.42 (.10)	.40 (.06)
	Guide $\rightarrow$ PF ( $w_2$ )	.14 (.09)	.14 (.08)
	Attach $\rightarrow$ PF ( $w_3$ )	.45 (.11)	.45 (.07)
	Intimacy $\rightarrow$ PF ( $w_4$ )	.26 (.13)	.24 (.09)
	Selfval $\rightarrow$ PF ( $w_5$ )	-.06 (.09)	-.05 (.09)
	Compar $\rightarrow$ NF ( $w_6$ )	.60 (.15)	.60 (.15)
	Conflict $\rightarrow$ NF ( $w_7$ )	.58 (.14)	.54 (.16)
	PF $\rightarrow$ trust ( $a_1$ )	.42 (.03)	.44 (.02)
	NF $\rightarrow$ trust ( $a_2$ )	-.19 (.04)	-.20 (.04)

The model fit obtained from this multi-sample analysis is .1660. It is found that in both South Korean and German samples, control, attachment, and intimacy seem to be significantly associated with positive function. In the South Korean sample, on the other hand, conflict is likely to be significantly related to negative function, whereas in the German sample both comparison and conflict are found to be significantly associated with negative function. In both samples of adolescents, positive function shows a positive and significant influence on trust ( $a_1 = .46$ , s.e. = .03, CR = 15.33 in South Korea;  $a_1 = .42$ , s.e. = .03, CR = 14.0 in Germany), while negative function had a negative and significant influence on trust ( $a_1 = -.09$ , s.e. = .03, CR = -3.0 in South Korea;  $a_1 = -.19$ , s.e. = .04, CR = -4.75 in Germany). Thus, positive function appears to have an effect of similar strength on trust in both samples. On the other hand, negative function in the German sample seems to show a stronger effect on trust than that in the South Korean sample (i.e., negative function affects trust more negatively). The correlations between the two components are .26 (s.e. = .12, CR = 2.17) and .45 (s.e. = .05, CR = 9.0) in South Korean and German adolescents, respectively.

On the basis of the obtained solutions in multi-sample analysis, we further assumed that the effects of control and attachment on positive function were identical across two samples. We also hypothesized that positive function had the same effect on trust across two samples. These across-sample assumptions were incorporated by imposing equality constraints on the portions of free parameters in  $\hat{\mathbf{W}}$  and  $\hat{\mathbf{A}}$ , that is,  $\hat{\mathbf{w}}$  and  $\hat{\mathbf{a}}$ , as described in the previous

section. Let  $\mathbf{P}'$  denote a matrix of equality constraints on  $\hat{\mathbf{w}}$ . Let  $\mathbf{p}'$  denote a vector of equality constraints on  $\hat{\mathbf{a}}$ . The  $\mathbf{P}'$  and  $\mathbf{p}'$  were easily specified as

$$\mathbf{P}' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$\mathbf{p}' = [1 \quad 0 \quad -1 \quad 0].$$

We require  $\mathbf{P}'\hat{\mathbf{w}} = \mathbf{0}$  and  $\mathbf{p}'\hat{\mathbf{a}} = 0$ . In the above, the first row of  $\mathbf{P}'$  indicated that the component weight for control (corresponding to  $w_1$ ) was identical across two samples. The second row of  $\mathbf{P}'$  represented that the component weight for attachment was equivalent across two samples. Likewise,  $\mathbf{p}'$  meant that the loadings for positive function were equal across the two samples. Then the  $\mathbf{P}'$  and  $\mathbf{p}'$  were transformed into reparameterization forms by the procedure described in Section 2.4.

The results of incorporating the equality constraints into the model are also provided in Table 5. Due to the equality constraints imposed, the component weight of control is identical across two samples and turns out to be significant ( $w_1 = .40$ , s.e. = .06, CR = 6.67). This is the case for the component weight of attachment ( $w_3 = .45$ , s.e. = .07, CR = 6.43) and the loading for positive function ( $a_1 = .44$ , s.e. = .02, CR = 22.0). The goodness of fit of the sample-wise constrained model is .1658. It is almost the same fit as that obtained from the unconstrained multi-sample analysis. Moreover, the constrained multi-sample analyses provide quite similar parameter estimates to the unconstrained case. Therefore, it might be safe to say that our hypotheses regarding the relationships among the parameters across samples are reasonable. This constrained analysis yields simpler interpretations of the solutions, reducing the number of parameters to be interpreted.

#### 4. Concluding remarks

In this paper, we proposed an extension of RA that could analyze a variety of hypothesized relationships among multiple sets of variables. The proposed method subsumes existing extended RA methods (Velu, 1991; Takane et al., 1995) as special cases. It is also versatile enough to analyze more complex relationships among variables such as direct effects of observed exogenous variables, higher-order components and multi-sample comparisons. A simple alternating least-squares algorithm was developed to estimate model parameters. According to our experiences, the algorithm seems to be efficient. It converges fast and seems to hardly suffer from the problem of converging to local minima.

A number of topics may be considered that further enhance the capabilities of the proposed method. For example, data transformation may be a significant consideration, which allows analyzing categorical variables within the method. It seems to be important to deal with missing observations, which frequently appear in large data sets. Furthermore, it is necessary to study robust estimation because the proposed method may not be robust against outliers as far as it is based on solving a simple least-squares criterion, which amounts to minimizing the sum of the ‘squared’ residuals. The assumption of normality is not essential for the method due to the least-squares fitting. If such an assumption is made, nonetheless,

the method enables to offer efficient estimators and perform statistical significance tests without recourse to resampling methods. Metric matrices (Takane and Shibayama, 1991) may be incorporated into the optimization criterion, leading to more diverse/effective analyses (Takane and Hunter, 2001). Future studies are needed to investigate the feasibility of these additional extensions.

## Acknowledgements

The work reported in this paper was supported by Grant A6394 from the Natural Sciences and Engineering Research Council of Canada to the first author. We wish to thank Yun Joo Park for kindly providing her social function data, Jamshid Etezadi-Amoli for his insightful comments on the second-order component model, and Jim Ramsay for his valuable comments on an earlier version of this paper.

## References

- Anderson, T.W., 1951. Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Ann. of Math. Statist.* 22, 327–351.
- Bock, R.D., Bargmann, R.E., 1966. Analysis of covariance structures. *Psychometrika* 31, 507–534.
- Böckenholt, U., Böckenholt, I., 1990. Canonical analysis of contingency tables with linear constraints. *Psychometrika* 55, 633–639.
- Böckenholt, U., Takane, Y., 1994. Linear constraints in correspondence analysis. In: Greenacre, M.J., Blasius, J. (Eds.), *Correspondence Analysis in Social Sciences*. Academic Press, London, pp. 112–127.
- Bookstein, F.L., 1982. Soft modeling: the basic design and some extensions. In: Jöreskog, K.G., Wold, H. (Eds.), *Systems under Indirect Observations II*. North-Holland, Amsterdam, pp. 55–74.
- Chin, W.W., 1998. Issues and opinion on structural equation modeling. *Management Inform. Systems Quart.* 22, 7–16.
- Davies, P.T., Tso, M.K.-S., 1982. Procedures for reduced-rank regression. *Appl. Statist.* 31, 244–255.
- de Leeuw, J., Young, F.W., Takane, Y., 1976. Additive structure in qualitative data: an alternating least squares method with optimal scaling features. *Psychometrika* 41, 471–503.
- Efron, B., 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*, SIAM, Philadelphia.
- Efron, B., Tibshirani, R.J., 1998. *An Introduction to the Bootstrap*, CRC Press, Boca Raton, FL.
- Fornell, C., Bookstein, F.L., 1982. Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *J. Marketing Res.* 19, 440–452.
- Greenacre, M.J., 1984. *Theory and Applications of Correspondence Analysis*, Academic Press, London.
- Hanafi, M., Qannari, E.M., 2002. An alternative algorithm to the PLS B problem, submitted for publication.
- Izenman, A.J., 1975. Reduced-rank regression for the multivariate linear model. *J. Multivariate Anal.* 5, 248–264.
- Jöreskog, K.G., 1970. A general method for analysis of covariance structures. *Biometrika* 57, 409–426.
- Jöreskog, K.G., 1973. A generating method for estimating a linear structural equation system. In: Goldberger, A.S., Duncan, O.D. (Eds.), *Structural Equation Models in the Social Sciences*. Academic Press, New York, pp. 85–112.
- Jöreskog, K.G., 1977. Structural equation models in the social sciences: specification, estimation, and testing. In: Krishnaiah, P.R. (Ed.), *Applications of Statistics*. North-Holland, Amsterdam, pp. 265–287.
- Jöreskog, K.G., Goldberger, A.S., 1975. Estimation of a model with multiple indicators and multiple causes of a single latent variable. *J. Amer. Statist. Association* 10, 631–639.
- Jöreskog, K.G., Wold, H., 1982. The ML and PLS techniques for modeling with latent variables: historical and comparative aspects. In: Wold, H., Jöreskog, K.G. (Eds.), *Systems Under Indirect Observation: Causality, Structure, Prediction, Part I*. North-Holland, Amsterdam, pp. 263–270.
- 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.

- Lambert, Z.V., Wildt, A.R., Durand, R.M., 1988. Redundancy analysis: an alternative to canonical correlation and multivariate multiple regression in exploring interset associations. *Psychol. Bull.* 104, 282–289.
- Lohmöller, J.B., 1989. *Latent Variable Path Modeling with Partial Least Squares*, Springer, New York.
- Lyttkens, E., 1968. On the fixed-point property of Wold's iterative estimation method for principal components. In: Krishnaiah, P.R. (Ed.), *Multivariate Analysis*. Academic Press, New York, pp. 335–350.
- Lyttkens, E., 1973. The fixed-point method for estimating interdependent systems with the underlying model specification. *J. Roy. Statist. Soc. A* 136, 353–394.
- MacCallum, R.C., Browne, M.W., 1993. The use of causal indicators in covariance structure models: some practical issues. *Psychol. Bull.* 114, 533–541.
- McDonald, R.P., 1996. Path analysis with composite variables. *Multivar. Behav. Res.* 31, 239–270.
- Millsap, R.E., Meredith, W., 1988. Component analysis in cross-sectional and longitudinal data. *Psychometrika* 53, 123–134.
- Mulaik, S.A., 1972. *The Foundations of Factor Analysis*, McGraw-Hill, New York.
- Park, Y., 1996. *Das soziale Netzwerk in der frühen Adoleszenz: eine vergleichende Untersuchung zweier Generationen aus Korea und Deutschland*. Doctoral Dissertation, University of Köln, Köln, Unpublished.
- Rao, C.R., 1964. The use and interpretation of principal component analysis in applied research. *Sankhya, A* 26, 329–358.
- Reinsel, G.C., Velu, R.P., 1998. *Multivariate Reduced-Rank Regression*, Springer, New York.
- Takane, Y., Hunter, M.A., 2001. Constrained principal component analysis: a comprehensive theory. *Appl. Algebra Eng. Comm. Comput.* 12, 391–419.
- 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.
- Takane, Y., Kiers, H., de Leeuw, J., 1995. Component analysis with different sets of constraints on different dimensions. *Psychometrika* 60, 259–280.
- ten Berge, J.M.F., 1993. *Least Squares Optimization in Multivariate Analysis*, DSWO Press, Leiden.
- ter Braak, C.J.F., 1986. Canonical correspondence analysis: a new eigenvalue technique for multivariate direct gradient analysis. *Ecology* 67, 1167–1179.
- Tucker, L.R., 1951. A method for synthesis of factor analysis studies. *Personnel Research Section Report No. 984*. US Department of the Army, Washington, DC.
- van den Wollenberg, A.L., 1977. Redundancy analysis: an alternative for canonical analysis. *Psychometrika* 42, 207–219.
- Velicer, W.F., Jackson, D.N., 1990. Component analysis versus common factor analysis: some issues in selecting and appropriate procedure. *Multivar. Behav. Res.* 25, 1–28.
- Velu, R.P., 1991. Reduced rank models with two sets of regressors. *Appl. Statist.* 40, 159–170.
- Wold, H., 1965. A fixed-point theorem with econometric background, I–II. *Ark. Mat.* 6, 209–240.
- Wold, H., 1966. Estimation of principal components and related methods by iterative least squares. In: Krishnaiah, P.R. (Ed.), *Multivariate Analysis*. Academic Press, New York, pp. 391–420.
- Wold, H., 1973. Nonlinear iterative partial least squares (NIPALS) modeling: some current developments. In: Krishnaiah, P.R. (Ed.), *Multivariate Analysis*. Academic Press, New York, pp. 383–487.
- Wold, H., 1981. *The Fixed Point Approach to Interdependent Systems*, North-Holland, Amsterdam.
- Wold, H., 1982. Soft modeling: the basic design and some extensions. In: Jöreskog, K.G., Wold, H. (Eds.), *Systems under Indirect Observations II*. North-Holland, Amsterdam, pp. 1–54.
- Yanai, H., 1986. Some generalizations of correspondence analysis in terms of projection operators. In: Diday, E., Escoufier, Y., Lebart, L., Pagès, J.P., Schektman, Y., Thomassone, R. (Eds.), *Data Analysis and Informatics IV*. North-Holland, Amsterdam, pp. 193–207.