STATISTICAL SOFTWARE SACS — SENSITIVITY ANALYSIS IN COVARIANCE STRUCTURE ANALYSIS

Shingo Watadani *and Yutaka Tanaka†

ABSTRACT

Statistical software SACS is developed for sensitivity analysis in covariance structure analysis. It consists of two parts, SACS(FA) and SACS(SEM), which deal with the problem of detecting influential observations in confirmatory as well as exploratory factor analysis models and in LISREL type linear structural equation models, respectively. Some computational problems are discussed.

1. Introduction

Covariance structure analysis (CSA) is a family of multivariate methods which have a common fundamental assumption that the covariance matrix Σ of observable variable is expressed as a function of a set of parameters $\underline{\theta} = (\theta_1, \dots, \theta_m)$, i.e., $\Sigma = \Sigma(\underline{\theta})$. This family contains confirmatory as well as exploratory factor analyses (FA), path analysis, LISREL type linear structural equation analysis among others.

Sensitivity analysis procedures have been proposed by Tanaka, Watadani and Moon (1991) and Tanaka and Watadani (1992) for CSA without/with equality constraints. In these papers the proposed procedures are illustrated in confirmatory and exploratory factor analyses. Tanaka, Watadani and Inoue (1992) discuss an application to LISREL type structural equation models. Several computer programs have been made during these studies, but they are not suitable for general uses, because they have been made with minimum requirement for specific purposes to analyze the numerical examples. Having modified them we have developed statistical software SACS which consists of SACS(FA) and SACS(SEM) for general uses of sensitivity analysis in FA models and LISREL type structural equation models, respectively. This software is written in Microsoft (Quick) BASIC and runs on MS-DOS. Some computational problems will be discussed in the final section.

2. Influence functions in CSA

In CSA an estimate $\hat{\underline{\theta}}$ for parameter vector $\underline{\theta}$ is obtained by minimizing a function called discrepancy function $G(S, \Sigma(\underline{\theta}))$ which measures the discrepancy between the sample

*Sanyo Gakuen Junior College, Hirai, Okayama 703, Japan.

†Department of Environmental and Mathematical Sciences, Okayama University, Tsushima, Okayama 700, Japan.

Key words: factor analysis; influence function; influential observations; LISREL; structural equation models.

covariance matrix $S = n^{-1} \sum (\mathbf{z}_i - \bar{\mathbf{z}})(\mathbf{z}_i - \bar{\mathbf{z}})^T$ and the reproduced covariance matrix $\Sigma(\underline{\theta})$ using the estimated parameters. The discrepancy function is given by

$$G_{\text{ULS}}(S, \Sigma) = \frac{1}{2} \text{tr}[(S - \Sigma)^2], \quad (2.1)$$

in unweighted least squares (ULS) estimation,

$$G_{\text{GLS}}(S, \Sigma) = \frac{1}{2} \text{tr}[(I - S^{-1}\Sigma)^2], \quad (2.2)$$

in generalized least squares (GLS) estimation or

$$G_{\text{ML}}(S, \Sigma) = \text{tr}[\Sigma^{-1}S] - \log|\Sigma^{-1}S| - p, \quad (2.3)$$

in maximum likelihood (ML) estimation (see, e.g., Jöreskog, 1978).

The major objective of sensitivity analysis is to detect influential observations. To evaluate the influence of each observation, we use the idea of so-called influence function proposed by Hampel (1974). Let a parameter vector $\underline{\theta}$ be expressed as a functional of the distribution function F , i.e., $\underline{\theta} = \underline{\theta}(F)$. The theoretical influence function (*TIF*) of $\underline{z}$ for $\underline{\theta}$ is defined by $TIF(\underline{z}; \underline{\theta}) = \lim_{\epsilon \rightarrow 0} [\{\underline{\theta}(F_\epsilon) - \underline{\theta}(F)\} / \epsilon]$, where $F_\epsilon = (1 - \epsilon)F + \epsilon\delta_{\underline{z}}$ is a perturbed distribution function by $\delta_{\underline{z}}$, the distribution function of a unit point mass at $\underline{z}$. In practice, we usually focus our interest on the influence of each observation $\underline{z} = \mathbf{z}_i$ ($i = 1, \dots, n$) on the estimate $\hat{\underline{\theta}} = \underline{\theta}(\hat{F})$ of the parameter vector. For this purpose the empirical influence function (*EIF*) of $\mathbf{z}_i$ for $\hat{\underline{\theta}}$ is defined by substituting the empirical distribution function $\hat{F}$ into F in the definition of the *TIF*.

As discussed by Tanaka, Watadani and Moon (1991), the *EIF* of $\mathbf{z}_i$ for $\hat{\underline{\theta}}$ is obtained as

$$EIF(\mathbf{z}_i; \hat{\underline{\theta}}) = - \left[\frac{\partial^2 G(\underline{s}, \hat{\underline{\theta}})}{\partial \underline{\theta} \partial \underline{\theta}^T} \right]^{-1} \left[\frac{\partial^2 G(\underline{s}, \hat{\underline{\theta}})}{\partial \underline{\theta} \partial \underline{s}^T} \right] EIF(\mathbf{z}_i; \underline{s}), \quad (2.4)$$

in the case of CSA with no constraints, where $\underline{s}$ denotes a vector which consists of the elements of covariance matrix S , i.e., $\underline{s} = \text{vech}(S) = (s_{11}, \dots, s_{p1}; s_{22}, \dots, s_{p2}; \dots; s_{pp})^T$ and the *EIF* of $\mathbf{z}_i$ for $\underline{s}$ is given by $EIF(\mathbf{z}_i; \underline{s}) = \text{vech}[(\mathbf{z}_i - \bar{\mathbf{z}})(\mathbf{z}_i - \bar{\mathbf{z}})^T - S]$. Tanaka and Watadani (1992) discuss the case of CSA with r equality constraints expressed as

$$\underline{h}(\underline{\theta}) = \begin{pmatrix} h_1(\underline{\theta}) \\ \vdots \\ h_r(\underline{\theta}) \end{pmatrix} = \underline{0}. \quad (2.5)$$

Using the Lagrangian function

$$G^*(\underline{s}, \underline{\theta}^*) = G^*(\underline{s}, \underline{\theta}, \underline{\lambda}) = G(\underline{s}, \underline{\theta}) + \underline{\lambda}^T \underline{h}(\underline{\theta}), \quad (2.6)$$

where $\underline{\lambda} = (\lambda_1, \dots, \lambda_r)$ is an $r \times 1$ vector of Lagrangian multipliers and $\underline{\theta}^* = (\underline{\theta}^T, \underline{\lambda}^T)^T$, the *EIF* of $\mathbf{z}_i$ for $\hat{\underline{\theta}}$ is given by

$$EIF(\mathbf{z}_i; \hat{\underline{\theta}}) = -Q(\underline{s}, \hat{\underline{\theta}}, \hat{\underline{\lambda}}) \left[\frac{\partial^2 G(\underline{s}, \hat{\underline{\theta}})}{\partial \underline{\theta} \partial \underline{s}^T} \right] EIF(\mathbf{z}_i; \underline{s}), \quad (2.7)$$

where $Q(\underline{s}, \hat{\underline{\theta}}, \hat{\underline{\lambda}})$ is the upper left $m \times m$ submatrix of the $(m+r) \times (m+r)$ matrix

$$\left[\frac{\partial^2 G^*(\underline{s}, \hat{\underline{\theta}}^*)}{\partial \hat{\underline{\theta}}^* \partial \hat{\underline{\theta}}^{*T}} \right]^{-1} = \begin{bmatrix} \frac{\partial^2 G(\underline{s}, \hat{\underline{\theta}})}{\partial \underline{\theta} \partial \underline{\theta}^T} + \sum_{j=1}^r \lambda_j \left[\frac{\partial^2 h_j(\hat{\underline{\theta}})}{\partial \underline{\theta} \partial \underline{\theta}^T} \right] & \frac{\partial \underline{h}^T(\hat{\underline{\theta}})}{\partial \underline{\theta}} \\ \frac{\partial \underline{h}(\hat{\underline{\theta}})}{\partial \underline{\theta}^T} & 0 \end{bmatrix}^{-1}. \quad (2.8)$$

To calculate the *EIF* for $\hat{\underline{\theta}}$ the second derivatives of discrepancy function G are required. It is known that the first derivatives of G are given in the form as

$$\frac{\partial G(S, \Sigma)}{\partial \theta_i} = \text{tr} \left[A(\Sigma - S) A \frac{\partial \Sigma}{\partial \theta_i} \right], \quad (2.9)$$

where $\Sigma(\underline{\theta})$ is simply expressed as Σ and

$$A = \begin{cases} I, & \text{in ULS;} \\ S^{-1}, & \text{in GLS;} \\ \Sigma^{-1}, & \text{in ML.} \end{cases} \quad (2.10)$$

From (2.9) the second derivatives of G are obtained as

$$\begin{aligned} \frac{\partial^2 G(\underline{s}, \underline{\theta})}{\partial \theta_i \partial \theta_j} &= \text{tr} \left[A \frac{\partial \Sigma}{\partial \theta_i} A \frac{\partial \Sigma}{\partial \theta_j} \right] + 2 \text{tr} \left[A(\Sigma - S) \frac{\partial A}{\partial \theta_i} \frac{\partial \Sigma}{\partial \theta_j} \right] \\ &\quad + \text{tr} \left[A(\Sigma - S) A \frac{\partial^2 \Sigma}{\partial \theta_i \partial \theta_j} \right], \end{aligned} \quad (2.11)$$

and

$$\frac{\partial^2 G(\underline{s}, \underline{\theta})}{\partial \theta_i \partial s_{jk}} = -\text{tr} \left[A E_{jk}^* A \frac{\partial \Sigma}{\partial \theta_i} \right] + 2 \text{tr} \left[A(\Sigma - S) \frac{\partial A}{\partial s_{jk}} \frac{\partial \Sigma}{\partial \theta_i} \right], \quad (2.12)$$

where E_{jk}^* is a $p \times p$ matrix with 1's in the (j, k) and (k, j) elements and 0's in the other elements. The first derivatives of A are given by

$$\frac{\partial A}{\partial \theta_i} = \begin{cases} 0, & \text{in ULS and GLS;} \\ -\Sigma^{-1} [\partial \Sigma / \partial \theta_i] \Sigma^{-1}, & \text{in ML,} \end{cases} \quad (2.13)$$

and

$$\frac{\partial A}{\partial s_{ij}} = \begin{cases} 0, & \text{in ULS and ML;} \\ -S^{-1} E_{ij}^* S^{-1}, & \text{in GLS.} \end{cases} \quad (2.14)$$

The second derivatives (2.11) and (2.12) are often approximated by only the first terms of the right-hand sides neglecting the remaining terms which contain the factor $(\Sigma - S)$.

As another sample version, the sample influence function (*SIF*) of $\underline{z}_i$ for $\hat{\underline{\theta}}$ is defined as

$$SIF(\underline{z}_i; \hat{\underline{\theta}}) = -(n-1)(\hat{\underline{\theta}}_{(i)} - \hat{\underline{\theta}}), \quad (2.15)$$

where the subscript (i) indicates that the quantity with it is based on the sample with the i -th observation $\underline{z}_i$ deleted. The *SIF* has clear “leave-one-out” interpretation, but requires high computing cost.

3. Influence measures

Changes (or influence functions) of the estimate $\hat{\underline{\theta}}$ are in general vector-valued. So we need to transform them into scalar measures. Here we introduce three kinds of influence measures which reflect the influence on the estimate $\hat{\underline{\theta}}$, on the precision of the estimate and on the goodness-of-fit. For the convenience of computation we use the linear approximation for the estimate $\hat{\underline{\theta}}_{(i)}$ based on the *EIF* such as

$$\hat{\underline{\theta}}_{(i)} \cong \tilde{\underline{\theta}}_{(i)} = \hat{\underline{\theta}} - (n-1)^{-1}EIF(\underline{z}_i; \hat{\underline{\theta}}), \quad (3.1)$$

in evaluating those measures. Symbol ($\sim$) indicates that the quantity with it is obtained by using this kind of linear approximation.

Influence measures on the estimate and on the precision can be defined for a part of the parameters as well as all the parameters.

1) Influence on the estimate

Two measures are defined based on the approximate change of the estimate $\Delta \tilde{\underline{\theta}} = \tilde{\underline{\theta}}_{(i)} - \hat{\underline{\theta}} = -(n-1)^{-1}EIF(\underline{z}_i; \hat{\underline{\theta}})$.

(a) Euclidean norm:

$$\|\Delta \tilde{\underline{\theta}}\| = (n-1)^{-1} \|EIF(\underline{z}_i; \hat{\underline{\theta}})\|. \quad (3.2)$$

(b) Generalized Cook's distance:

$$\begin{aligned} \widetilde{CD} &= \Delta \tilde{\underline{\theta}}^T [\widehat{\text{acov}}(\hat{\underline{\theta}})]^{-1} \Delta \tilde{\underline{\theta}} \\ &= (n-1)^{-2} EIF(\underline{z}_i; \hat{\underline{\theta}})^T [\widehat{\text{acov}}(\hat{\underline{\theta}})]^{-1} EIF(\underline{z}_i; \hat{\underline{\theta}}), \end{aligned} \quad (3.3)$$

where $\widehat{\text{acov}}(\hat{\underline{\theta}})$ is an estimate of the asymptotic covariance matrix of $\underline{\theta}$. It is known that $\widehat{\text{acov}}(\hat{\underline{\theta}})$ is given in the case of *ML* as

$$\widehat{\text{acov}}(\hat{\underline{\theta}}) = \frac{2}{n} \left[\frac{\partial^2 G(\underline{s}; \hat{\underline{\theta}})}{\partial \underline{\theta} \partial \underline{\theta}^T} \right]^{-1}, \quad (3.4)$$

in *CSA* without constraints and as

$$\widehat{\text{acov}}(\hat{\underline{\theta}}) = \frac{2}{n} Q(\underline{s}, \hat{\underline{\theta}}, \hat{\underline{\lambda}}), \quad (3.5)$$

in *CSA* with equality constraints.

2) Influence on the precision

To evaluate the change of the estimated asymptotic covariance matrix $\widehat{\text{acov}}(\hat{\underline{\theta}})$ the following measure is introduced by analogy with *COVRATIO* in regression diagnostics (see Belsley, Kuh and Welsch, 1980).

(c) *COVRATIO*-like measure:

$$\widetilde{CVR} = |\widehat{\text{acov}}(\tilde{\underline{\theta}}_{(i)})| / |\widehat{\text{acov}}(\hat{\underline{\theta}})|. \quad (3.6)$$

The fact that $\widehat{CVR} < 1$ for an observation suggests that the precision becomes better if the observation is deleted.

3) Influence on the goodness-of-fit

In LISREL model the likelihood ratio statistic X^2 , the goodness-of-fit index (GFI) and the root mean square residual (RMR) (see, e.g., Jöreskog and Sörbom, 1986) are known as goodness-of-fit measures. They are defined as

$$X^2 = nG(S, \Sigma), \quad (3.7)$$

in ML,

$$GFI = 1 - \text{tr}[\{\hat{A}(S - \hat{\Sigma})\}^2] / \text{tr}[(\hat{A}S)^2], \quad (3.8)$$

and

$$RMR = \sum_{i=1}^p \sum_{j=1}^i (s_{ij} - \hat{\sigma}_{ij}), \quad (3.9)$$

where $\Sigma(\hat{\theta})$ is simply expressed as $\hat{\Sigma} = (\hat{\sigma}_{ij})$. The first measure is used in ML . We define the change of each statistic as the influence measure on the goodness-of-fit.

(d) Change of X^2 :

$$\Delta \tilde{X}^2 = \tilde{X}_{(i)}^2 - X^2. \quad (3.10)$$

(e) Change of GFI :

$$\Delta \widetilde{GFI} = \widetilde{GFI}_{(i)} - GFI. \quad (3.11)$$

(f) Change of RMR :

$$\Delta \widetilde{RMR} = \widetilde{RMR}_{(i)} - RMR. \quad (3.12)$$

The fact that $\Delta \tilde{X}^2 < 0$, $\Delta \widetilde{GFI} > 0$ or $\Delta \widetilde{RMR} < 0$ suggests that the data set fits the assumed model better if the observation is deleted.

4. Influence of subsets of observations

For evaluating the influence of a subset of k observations $Z_A = \{\mathbf{z}_i; i \in A\}$ a generalized influence function was introduced by Tanaka (1992) using $F_k(Z_A) = k^{-1} \sum_{i \in A} \delta_{\mathbf{z}_i}$ instead of $\delta_{\mathbf{z}}$ in the definition of the influence function. The generalized influence function of Z_A is given as the average of the ordinary influence function of each member $\mathbf{z}_i$ of the set. Then the estimate based on the sample with Z_A deleted is approximately evaluated as

$$\hat{\theta}_{(A)} \cong \tilde{\theta}_{(A)} = \hat{\theta} - \frac{1}{n-k} \sum_{i \in A} EIF(\mathbf{z}_i; \hat{\theta}). \quad (4.1)$$

Based on this relation we can say that we should search for the observations which are relatively influential and also have similar influence patterns with each other to detect influential subsets. For this purpose we can apply principal component analysis (PCA) or canonical variate analysis (CVA) using $\widehat{\text{acov}}(\hat{\theta})$ as its metric to the EIF vectors for $\hat{\theta}$.

5. Application to FA models

We consider the following basic model of FA:

$$\underline{x} = \Lambda \underline{f} + \underline{e}, \quad (5.1)$$

where $\underline{x}$ is a $p \times 1$ vector of observable variables, $\underline{f}$ is a $q \times 1$ ($q < p$) vector of common factor scores, $\underline{e}$ is a $p \times 1$ vector of unique factor scores and $\Lambda = (\lambda_{ir})$ is a $p \times q$ factor loading matrix. Means and covariances of common and unique factors are assumed as $E(\underline{f}) = \underline{0}$, $E(\underline{e}) = \underline{0}$, $E(\underline{f}\underline{f}^T) = \Phi = (\phi_{rs})$ (every diagonal element is unity), $E(\underline{e}\underline{e}^T) = \Psi$ (diagonal matrix) and $E(\underline{f}\underline{e}^T) = 0$. Then the covariance matrix Σ of observable variables $\underline{x}$ is expressed as

$$\Sigma(\Lambda, \Phi, \Psi) = \Lambda \Phi \Lambda^T + \Psi. \quad (5.2)$$

The first and second derivatives of the covariance matrix (5.2) with respect to parameters are given as

$$\begin{cases} \frac{\partial \Sigma(\Lambda, \Phi, \Psi)}{\partial \lambda_{is}} = E_{is} \Phi \Lambda^T + \Lambda \Phi E_{si}, \\ \frac{\partial \Sigma(\Lambda, \Phi, \Psi)}{\partial \phi_{st}} = \Lambda E_{st}^* \Lambda^T, \quad \frac{\partial \Sigma(\Lambda, \Phi, \Psi)}{\partial \psi_{ii}} = E_{ii}^*, \end{cases} \quad (5.3)$$

and

$$\begin{cases} \frac{\partial^2 \Sigma(\Lambda, \Phi, \Psi)}{\partial \lambda_{is} \partial \lambda_{jt}} = \phi_{st} (E_{is} + E_{jt}), \\ \frac{\partial^2 \Sigma(\Lambda, \Phi, \Psi)}{\partial \lambda_{is} \partial \phi_{tu}} = \delta_{st} [E_{iu} \Lambda^T + \Lambda E_{ui}] + \delta_{su} [E_{it} \Lambda^T + \Lambda E_{ti}], \end{cases} \quad (5.4)$$

where the symbol δ represents Kronecker's delta and E_{ab} is a matrix unit with 1 in the (a, b) element and 0's in the other elements.

1) Confirmatory FA

In confirmatory FA no equality constraints are imposed, but a unique solution is usually obtained, because some elements of Λ and Φ are assumed to have values specified in advance. Therefore we may identify confirmatory FA as a special case of CSA without constraints in which the parameters to be estimated are the unspecified elements of Λ , Φ and Ψ .

2) Exploratory FA

In exploratory FA, on the other hand, no elements are specified in advance. Instead to obtain a unique solution it is usually assumed that common factors are not correlated, i.e., $E(\underline{f}\underline{f}^T) = \Phi = I$, and $\Lambda^T \Lambda$ is diagonal in ULS or $\Lambda^T \Psi^{-1} \Lambda$ is diagonal in ML. Thus we may identify exploratory FA as a special case of CSA with $r = q(q-1)/2$ equality constraints

$$h_{st}(\Lambda) = [\Lambda^T \Lambda]_{st} = 0 \quad (1 \leq s < t \leq q), \quad (5.5)$$

in ULS, or

$$h_{st}(\Lambda, \Psi) = [\Lambda^T \Psi^{-1} \Lambda]_{st} = 0 \quad (1 \leq s < t \leq q), \quad (5.6)$$

in ML. The parameters to be estimated are the elements of Λ and the diagonal elements of Ψ . The first and second derivatives of h_{st} are given as

$$\frac{\partial h_{st}(\Lambda)}{\partial \lambda_{iu}} = \delta_{su} \lambda_{it} + \delta_{tu} \lambda_{is} \quad (5.7)$$

and

$$\frac{\partial^2 \mathbf{h}_{st}(\Lambda)}{\partial \lambda_{iu} \partial \lambda_{jv}} = \delta_{ij}(\delta_{su} \delta_{tv} + \delta_{sv} \delta_{tu}) \quad (5.8)$$

in ULS, or

$$\frac{\partial \mathbf{h}_{st}(\Lambda, \Psi)}{\lambda_{iu}} = \frac{\delta_{su} \lambda_{it} + \delta_{tu} \lambda_{is}}{\psi_{ii}}, \quad \frac{\partial \mathbf{h}_{st}(\Lambda, \Psi)}{\psi_{ii}} = -\frac{\lambda_{is} \lambda_{it}}{\psi_{ii}^2}, \quad (5.9)$$

and

$$\begin{cases} \frac{\partial^2 \mathbf{h}_{st}}{\partial \lambda_{iu} \partial \lambda_{jv}} = \frac{\delta_{ij}(\delta_{su} \delta_{tv} + \delta_{sv} \delta_{tu})}{\psi_{ii}}, \\ \frac{\partial^2 \mathbf{h}_{st}}{\partial \lambda_{iu} \partial \psi_{jj}} = -\frac{\delta_{ij}(\delta_{su} \lambda_{it} + \delta_{tu} \lambda_{is})}{\psi_{ii}^2}, \quad \frac{\partial^2 \mathbf{h}_{st}}{\partial \psi_{ii} \partial \psi_{jj}} = \frac{2\delta_{ij} \lambda_{is} \lambda_{it}}{\psi_{ii}^3}, \end{cases} \quad (5.10)$$

in ML.

6. Application to LISREL type structural equation models

The general LISREL model is defined by three equations,

$$\begin{cases} \underline{\eta} = \mathbf{B}\underline{\eta} + \Gamma\underline{\xi} + \underline{\zeta}, \\ \underline{y} = \Lambda_y \underline{\eta} + \underline{\varepsilon}, \\ \underline{x} = \Lambda_x \underline{\xi} + \underline{\delta}, \end{cases} \quad (6.1)$$

where $\underline{y}$ ($p_1 \times 1$) and $\underline{x}$ ($p_2 \times 1$) are vectors of dependent and independent observable variables, respectively, $\underline{\eta}$ ($q_1 \times 1$) and $\underline{\xi}$ ($q_2 \times 1$) are dependent and independent latent variables, respectively, $\mathbf{B}$, Γ , Λ_y and Λ_x are coefficient matrices with appropriate sizes, and $\underline{\zeta}$, $\underline{\varepsilon}$ and $\underline{\delta}$ are random vectors of residuals. In this model it is assumed that all variables, observable as well as latent, are measured in deviation from their means and $\underline{\zeta}$, $\underline{\varepsilon}$ and $\underline{\delta}$ are uncorrelated with other random vectors of latent variables. Let Φ , Ψ , Θ_ε and Θ_δ be the covariance matrices of $\underline{\xi}$, $\underline{\zeta}$, $\underline{\varepsilon}$ and $\underline{\delta}$, respectively, then the covariance matrix Σ ($((p+q) \times (p+q))$) of $\underline{z} = (\underline{y}^T, \underline{x}^T)^T$ is expressed as

$$\Sigma = \begin{bmatrix} \Lambda_y(I - \mathbf{B})^{-1}(\Gamma\Phi\Gamma^T + \Psi)(I - \mathbf{B}^T)^{-1}\Lambda_y^T + \Theta_\varepsilon & \text{(symmetric)} \\ \Lambda_x\Phi\Gamma^T(I - \mathbf{B}^T)^{-1}\Lambda_y^T & \Lambda_x\Phi\Lambda_x^T + \Theta_\delta \end{bmatrix}. \quad (6.2)$$

The first derivatives of the covariance matrix (6.2) are obtained as

$$\begin{cases} \frac{\partial \Sigma}{\partial \lambda_{yij}} = I_y E_{ij} J_y^T + J_y E_{ji} I_y^T, & \frac{\partial \Sigma}{\partial \lambda_{xij}} = I_x E_{ij} J_x^T + J_x E_{ji} I_x^T, \\ \frac{\partial \Sigma}{\partial \beta_{ij}} = J_\eta E_{ij} J_y^T + J_y E_{ji} J_\eta^T, & \frac{\partial \Sigma}{\partial \gamma_{ij}} = J_\eta E_{ij} J_x^T + J_x E_{ji} J_\eta^T, \\ \frac{\partial \Sigma}{\partial \phi_{ij}} = J_\xi E_{ij}^* J_\xi^T, & \frac{\partial \Sigma}{\partial \psi_{ij}} = J_\eta E_{ij}^* J_\eta^T, \\ \frac{\partial \Sigma}{\partial \theta_{\varepsilon ij}} = I_y E_{ij}^* I_y^T, & \frac{\partial \Sigma}{\partial \theta_{\delta ij}} = I_x E_{ij}^* I_x^T, \end{cases} \quad (6.3)$$

where

$$\begin{cases} I_y = \begin{bmatrix} I \\ 0 \end{bmatrix}, & I_x = \begin{bmatrix} 0 \\ I \end{bmatrix}, \\ J_\eta = \begin{bmatrix} \Lambda_y(I - B)^{-1} \\ 0 \end{bmatrix}, & J_\xi = \begin{bmatrix} \Lambda_y(I - B)^{-1}\Gamma \\ \Lambda_x \end{bmatrix}, \\ J_x = \begin{bmatrix} \Lambda_y(I - B)^{-1}\Gamma\Phi \\ \Lambda_x\Phi \end{bmatrix}, & J_y = \begin{bmatrix} \Lambda_y(I - B)^{-1}(\Gamma\Phi\Gamma^T + \Psi)(I - B^T)^{-1} \\ \Lambda_x\Phi\Gamma^T(I - B^T)^{-1} \end{bmatrix}. \end{cases} \quad (6.4)$$

As explained in section 2, the second derivatives of discrepancy function G , (2.11) and (2.12), can be approximated only with the first terms of the right-hand sides. In our software this kind of approximation is used and, therefore, the second derivatives of Σ with respect to parameters are not required.

7. Statistical software SACS

Statistical software SACS(FA) and SACS(SEM) are developed for sensitivity analysis in CSA for the cases of FA models and LISREL type structural equation models, respectively. They are written in Microsoft (Quick) BASIC on MS-DOS and have been tested on NEC PC-9800 series personal computers. The present version of SACS(FA) can deal with confirmatory FA models using the ML, ULS or GLS method and exploratory FA models using the ML or ULS method. That of SACS(SEM) can deal with LISREL type structural equation models using the ML, ULS or GLS method. Each of SACS(FA) and SACS(SEM) consists of 5 parts, (1) data entry and preliminary analysis, (2) parameter estimation, (3) single-case diagnostics, (4) multiple-case diagnostics, and (5) posterior analysis.

1) Data entry and preliminary analysis

In this part SACS reads input files, the parameter file and the data file, and computes basic statistics.

2) Parameter estimation

In this part SACS provides estimated parameters in the specified model using the modified Newton-Raphson method. The iterative process is terminated if the difference of two successive approximations are sufficiently small, i.e., $\max_i |\theta_i^{(s+1)} - \theta_i^{(s)}| < \varepsilon$ for a given ε , where the superscript for θ indicates the iteration number. The user can assign this constant ε and the maximum iteration number to terminate.

3) Single-case diagnostics

This part computes the *EIF* of each observation for the estimated parameters and summarize them into influence measures on the estimate, either Euclidean norm or generalized Cook's distance, and also influence measures on the precision and on the goodness-of-fit. The *EIF* and computed influence measures can be displayed in index plots on the screen. If the user wants, SACS computes the *SIF* and influence measures based on the *SIF*, and displays scatter diagrams of influence measures based on the *SIF* versus the *EIF*.

4) Multiple-case diagnostics

In this part PCA or CVA is applied to the *EIF* for all parameters or a part of the parameters to detect observations which are relatively influential individually and have similar influence patterns with each other.

Table 1. Parameter file and data file for SACS(FA)

nc=88 nv=5 nf=2	77,82,67,67,81
std	63,78,80,70,81
ml	75,73,71,66,81
confirmatory	55,72,63,70,68
lambda	63,63,65,70,63
1,0	.
1,0	.
0,1	.
0,1	5,30,44,36,18
0,1	12,30,32,35,21
oblique	5,26,15,20,20
data=open.dat	0,40,21, 9,14

5) *Posterior analysis*

This part provides the results of CSA without specified subsets of observations. The user can specify a subset of observations as a candidate of influential subset and confirm whether they are really influential or not, comparing the results for the sample with and without the specified set of observations.

Example. For illustration we shall show the applications of SACS(FA) and SACS(SEM) to a numerical example of a set of open/closed book data (Mardia, Kent and Bibby, 1979, Table 1.2.1). This data set was analyzed by Tanaka, Watadani and Moon (1991) and Tanaka and Watadani (1992) as well as by Mardia et al. The set of data consists of $n = 88$ observations (students) and $p = 5$ variables (examinations), in which the first two examinations were performed with the books closed and the remaining three examinations were performed with the books opened. From this property and the results of previous works we assume a two factor model, in which the first factor is related to the first two variables and the second factor is related to the remaining three variables.

To apply SACS(FA) the user prepares two input files, the parameter file and the data file as shown in Table 1. In the parameter file each item is separated by blank, comma or line feed. At first, the number of cases or observations (**nc**), number of variables (**nv**) and number of common factors (**nf**) must be assigned. Then the user selects standardized data or raw data (**std/raw**) to be analyzed, a method of estimation (**uls/gls/ml**), and confirmatory or exploratory FA model (**confirmatory/exploratory**). In the case of confirmatory FA models the patterns of the loading matrix Λ and of the correlation matrix Φ of the common factors are assigned. To specify the pattern of Λ it is required to write in the parameter file the word **lambda** followed by $p \times q$ numbers, which consist of 0's (indicating the elements fixed to be 0) and 1's (indicating the free elements). Similarly the pattern of Φ is specified by writing in the parameter file the word **phi** followed by the 0/1 pattern of off-diagonal lower half of Φ . It can also be specified by the word **orthogonal** (indicating orthogonal factors) or **oblique** (indicating oblique factors). The user also assigns the filename of the data file in the parameter file. In the data file a set of $n \times p$ data is written successively

Table 2. Parameter file for SACS(SEM)

nc=88 ny=0 nx=5 ne=0 nk=2	ini.Phi
std ml	1
ptn.Lx	0,1
1,0	ptn.Tdel
1,0	1
0,1	0,1
0,1	0,0,1
0,1	0,0,0,1
ini.Lx	0,0,0,0,1
0.5,0	ini.Tdel
0.5,0	0.75
0,0.5	0,0.75
0,0.5	0,0,0.75
0,0.5	0,0,0,0.75
ptn.Phi	0,0,0,0,0.75
0	data=open.dat
1,0	

separated by comma or line feed.

The parameter file for SACS(SEM) is shown in Table 2. At first, the number of cases or observations (nc), numbers of dependent and independent observable variables (ny and nx) and numbers of dependent and independent latent variables (ne and nk) must be assigned. The patterns of parameter matrices Λ_y , Λ_x , B, Γ , Φ , Ψ , Θ_ϵ and Θ_δ are assigned in the parameter file by the words ptn.Ly, ptn.Lx, ptn.beta, ptn.gamm, ptn.phi, ptn.psi, ptn.Teps and ptn.Tdel followed by 0/1 patterns, where 0's correspond to fixed parameters and 1's correspond to free parameters. Their initial estimates are specified by the the words ini.Ly, ini.Lx, ini.beta, ini.gamm, ini.phi, ini.psi, ini.Teps and ini.Tdel followed by assigned values. For symmetric matrices Φ , Ψ , Θ_ϵ and Θ_δ , 0/1 patterns and assigned values are required for only the lower half elements.

Tables 3 and 4 show the output of SACS(FA), the result of parameter estimation and generalized Cook's distance. Figure 1 shows the index plot of generalized Cook's distance for all parameters which is displayed on the screen. From this index plot it is found that case number 81 is the most influential on the estimates of all parameters. Similar output is obtained with SACS(SEM). Comparing the results of SACS(FA) and SACS(SEM) the estimated parameters are exactly the same, but the influence measures are slightly different. A major reason for these differences may be due to the fact that the second derivatives (2.11) and (2.12) of discrepancy function are approximated with the first terms of the right-hand sides in SACS(SEM).

Table 3. Output of SACS(FA): Result of parameter estimation

***** PARAMETER ESTIMATION *****		
	estimates (	S.E.)
1: lambda[1, 1]	0.700723 (	0.104795)
2: lambda[2, 1]	0.789764 (	0.104302)
3: lambda[3, 2]	0.925535 (	0.087315)
4: lambda[4, 2]	0.774010 (	0.094674)
5: lambda[5, 2]	0.725402 (	0.096771)
6: phi[2, 1]	0.817638 (	0.071112)
7: psi[1, 1]	0.508988 (	0.103043)
8: psi[2, 2]	0.376273 (	0.104165)
9: psi[3, 3]	0.143385 (	0.065805)
10: psi[4, 4]	0.400908 (	0.077831)
11: psi[5, 5]	0.473792 (	0.084778)
Chi-sq. =	2.09659 (with 4 d. f.)	
GFI =	0.99017	
AGFI =	0.96314	
RMR =	0.01946	

Table 4. Output of SACS(FA): generalized Cook's distance

***** INFLUENCE MEASURES BASED ON THE EIF *****				
===== Generalized Cook's Distance =====				
case#	theta	lambda	phi	psi
1	0.31468	0.29339	0.00713	0.02417
2	0.38066	0.33915	0.03605	0.01202
3	0.20471	0.17722	0.02639	0.00623
4	0.07547	0.05883	0.00784	0.01453
5	0.04862	0.03049	0.00726	0.01141
:				
86	0.09071	0.07022	0.01038	0.01480
87	0.92025	0.53117	0.06719	0.04738
88	0.65717	0.62739	0.04726	0.12981

Fig. 1. Index plot displayed on the screen

8. Discussion

Here we discuss some computational problems. As mentioned before a modified Newton-Raphson method is used for parameter estimation. The word “modified” means that in the process of each iteration

$$\hat{\underline{\theta}}^{(s+1)} = \hat{\underline{\theta}}^{(s)} - \left[\frac{\partial^2 G(\underline{s}, \hat{\underline{\theta}}^{(s)})}{\partial \underline{\theta} \partial \underline{\theta}^T} \right]^{-1} \left[\frac{\partial G(\underline{s}, \hat{\underline{\theta}}^{(s)})}{\partial \underline{\theta}} \right], \quad (8.1)$$

the second derivatives of the discrepancy function are approximated only by the first terms of right-hand side in (2.11), and further, if the difference of the last two previous approximations, i.e., $\max_i |\hat{\theta}_i^{(s)} - \hat{\theta}_i^{(s-1)}|$, is small, the inverse of the matrix of these second derivatives is evaluated by an approximation method used in Fletcher-Powell method.

In exploratory FA models parameters are often estimated using spectral decomposition

$$W^{-1/2} (S - \hat{\Psi}^{(s)}) W^{-1/2} = \Upsilon_1 A_1 \Upsilon_1^T + \Upsilon_2 A_2 \Upsilon_2^T, \quad (8.2)$$

where W is a $p \times p$ diagonal matrix given by

$$W = \text{diag}(w_{11}, \dots, w_{pp}) = \begin{cases} I, & \text{in ULS;} \\ \hat{\Psi}^{(s)}, & \text{in ML,} \end{cases} \quad (8.3)$$

$A_1 = \text{diag}(\alpha_1, \dots, \alpha_q)$ and $A_2 = \text{diag}(\alpha_{q+1}, \dots, \alpha_p)$ consist of the largest q eigenvalues and the remaining $(p - q)$ eigenvalues, respectively, and $\Upsilon_1 = (\underline{v}_1, \dots, \underline{v}_q)$ and $\Upsilon_2 = (\underline{v}_{q+1}, \dots, \underline{v}_p)$

are the matrices whose column vectors are their associated eigenvectors. Define

$$G_u(\Psi) = \min_{\Lambda} G(S, \Sigma(\Lambda, \Psi)). \quad (8.4)$$

Then it is known that the first and second derivatives of $G_u(\Psi)$ at $\Psi = \hat{\Psi}^{(s)}$ are evaluated by

$$\frac{\partial G_u(\hat{\Psi}^{(s)})}{\partial \psi_i} = -w_{ii}^{-1} [\Upsilon_2 A_2 \Upsilon_2^T]_{ii}, \quad (8.5)$$

and

$$\frac{\partial^2 G_u(\hat{\Psi}^{(s)})}{\partial \psi_i \partial \psi_j} \cong w_{ii}^{-1} w_{jj}^{-1} \left\{ [\Upsilon_2 \Upsilon_2^T]_{ij} \right\}^2, \quad (8.6)$$

using spectral decomposition (8.2) (see, e.g., Jöreskog, 1967). Thus $\hat{\Psi}$ is obtained by minimizing $G_u(\Psi)$ with the Newton-Raphson method and $\hat{\Lambda}$ is given by

$$\hat{\Lambda} = W^{1/2} \Upsilon_1 A_1^{1/2}. \quad (8.7)$$

In FA models, covariance matrix Σ is decomposed as (5.2), where Ψ is positive definite and $\Sigma - \Psi$ is of rank q and positive semi-definite. But, the estimate obtained by the iterative procedure not necessarily satisfies these conditions. The estimate not satisfying these conditions is called “improper solution.” For example, if $\hat{\psi}_i^{(s)}$ is close to zero or takes a negative value in ML, the algorithm described above cannot continue. In LISREL type structural equation models similar conditions exist on covariance matrices Φ , Ψ , Θ_ϵ and Θ_δ of latent variables. To avoid an improper solution it may be appropriate to consider covariance structure analysis with inequality constraints, though, in the present version of SACS, inequality constraints are not imposed. Instead, in our software, if some $\hat{\psi}_i^{(s)}$ becomes smaller than a preassigned small value c , say 0.0001, then $\hat{\psi}_i$ is assigned as $\hat{\psi}_i^{(t)} = c$ for $t \geq s$ and the iterative process continues with $\hat{\psi}_i$ fixed. The *SIF* is computed based on the estimates obtained in this manner.

Acknowledgement

The authors express deep thanks to the editor and the reviewers for their careful readings of the manuscript.

REFERENCES

- Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). *Regression diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons.
- Hampel, F. R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc.*, **69**, 383–393.
- Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, **32**, 443–482.
- Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, **43**, 443–477.
- Jöreskog, K. G. and Sörbom, D. (1986). *LISREL VI—Analysis of Linear Structural Relationship by the Method of Maximum Likelihood*. Scientific Software, Inc.
- Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*, Academic Press.
- Tanaka, Y. (1992). Sensitivity analysis in multivariate methods: Principles methods and software. *Technical Report of Okayama Statistical Association* **52**.
- Tanaka, Y. and Watadani, S. (1992). Sensitivity analysis in covariance structure analysis with equality constraints. *Comm. Statist.*, **A21**, 1501–1515.
- Tanaka, Y., Watadani, S. and Inoue, K. (1992). Sensitivity analysis in structural equation models. *Computational Statistics* (Edited by Dodge, Y. and Whittaker, J.), Physica-Verlag. **1**, 493–498.
- Tanaka, Y., Watadani, S. and Moon, S. H. (1991). Influence in covariance structure analysis: with an application to confirmatory factor analysis. *Comm. Statist.*, **A20**, 3805–3821.

(Received January 1994; Revised October 1994)