Local influence analysis of multivariate probit latent variable models

Bin Lu\textsuperscript{a}, Xin-Yuan Song\textsuperscript{b,}\textsuperscript{*}

\textsuperscript{a}School of Finance, Nanjing University of Finance and Economics, Nanjing, Jiangsu 210003, PR China
\textsuperscript{b}Department of Statistics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Received 22 November 2004
Available online 9 December 2005

Abstract

The multivariate probit model is very useful for analyzing correlated multivariate dichotomous data. Recently, this model has been generalized with a confirmatory factor analysis structure for accommodating more general covariance structure, and it is called the MPCFA model. The main purpose of this paper is to consider local influence analysis, which is a well-recognized important step of data analysis beyond the maximum likelihood estimation, of the MPCFA model. As the observed-data likelihood associated with the MPCFA model is intractable, the famous Cook’s approach cannot be applied to achieve local influence measures. Hence, the local influence measures are developed via Zhu and Lee’s [Local influence for incomplete data model, J. Roy. Statist. Soc. Ser. B 63 (2001) 111–126.] approach that is closely related to the EM algorithm. The diagnostic measures are derived from the conformal normal curvature of an appropriate function. The building blocks are computed via a sufficiently large random sample of the latent response strengths and latent variables that are generated by the Gibbs sampler. Some useful perturbation schemes are discussed. Results that are obtained from analyses of an artificial example and a real example are presented to illustrate the newly developed methodology.

© 2005 Elsevier Inc. All rights reserved.


Keywords: Conformal normal curvature; Dichotomous variables; Gibbs sampler; Local influence; $Q$-displacement function

\textsuperscript{*} Corresponding author. Fax: +852 26035188.
E-mail address: xysong@sta.cuhk.edu.hk (X.-Y. Song).
1. Introduction

Correlated dichotomous data often arise in behavioral, medical and psychological researches, ranging from measurements of random cross-section subjects to repeated measurements in longitudinal studies. The multivariate probit (MP) model is a popular method for analyzing this kind of data. This model is described in terms of a correlated multivariate normal distribution of the underlying latent variables that are manifested as discrete variables through a threshold specification, and hence allows the flexible modeling of the correlation structure and easy interpretation of the parameters.

Since the pioneer work of Ashford and Sowden [1], numerous attempts have been made to solve the computational difficulty of evaluating the multivariate normal orthant probabilities that are involved in the observed-data likelihood function of a MP model. A common approach is to use much less restrictive covariance structures to reduce the computational burden of evaluating the probabilities; see Ochi and Prentice [21]. In particular, Bock and Aitkin [3] used the exploratory factor analysis (EFA) model for the covariance structure and applied an EM algorithm [8] to obtain the ML solution. Bock and Gibbons [4], and Gibbons and Wilcox-Gök [11] extended the Bock and Aitkin [3] model to an EFA model with fixed covariates. The approaches that were used by Bock and Aitkin [3], Bock and Gibbons [4], and Gibbons and Wilcox-Gök [11] applied the Gauss–Hermite quadrature to approximate the integrals in relation to the marginal probabilities. Meng and Schilling [18] reanalyzed the EFA model of Bock and Aitkin [3] and pointed out some deficiencies in using the Gauss–Hermite quadrature to approximate the integrals that are associated with the marginal probabilities. They then recommended a better approach that is based on the Monte Carlo EM (MCEM) algorithm [30]. Chib and Greenberg [5] developed a Markov chain Monte Carlo (MCMC) method for a MP model with a general covariance structure. As these methods involve simulation of observations from a multivariate truncated normal distribution, the computational burden is heavy. Recently, Song and Lee [27] generalized the MP model of Bock and Gibbons [4], and Gibbons and Wilcox-Gök [11] by incorporating a more useful confirmatory factor analysis (CFA) model rather than the EFA model, and called it the MPCFA model. For a psychological research point of view, the MPCFA model can be regarded as a CFA model with dichotomous manifest variables and fixed covariates. They applied the MCEM algorithm [30] to maximum likelihood (ML) estimation, and showed that their approach can analyze the MP model with a general covariance with less computation effect than Chib and Greenberg [5].

Local influence analysis is a general statistical technique to assess the stability of the estimation outputs with respect to the model inputs. Model inputs may include data, parameters or other characteristics. Outputs may include the parameters estimates, final objective function values, estimates of residuals, etc. Local influence may be regarded as an important step of data analysis after estimation for drawing statisticians’ attention to influential aspects of the inputs in relation to the underlying model and problem. The identified problems caused by influential aspects may give ideas for improving the model assumptions and/or input data in establishing a better model. For instance, appropriate treatment of potential outliers will give a better fitted model. Cook [6] proposed a unified approach for assessment of local influence in minor perturbations of a statistical model. This approach has been widely applied to local influence analysis of statistical models. Typical examples are the applications to nonlinear regression models [28], growth curve models [22], principal component analysis [26], factor analysis and structural equation models [29, 24, 14], among many others. However, as the observed-data likelihood of a MPCFA model involves intractable integrals, it is very difficult and inappropriate to directly apply Cook’s [6]
approach to obtain local influence measures. Hence, as far as we know, no local influence analysis for MPCFA model has been developed.

Recently, Zhu and Lee [31] developed an approach for achieving local influence analysis for general statistical models with missing data, by working with a $Q$-displacement function that is closely related to the conditional expectation of the complete-data log-likelihood at the E-step of an EM algorithm. Inspired by Lee and Xu [15], and Lee and Tang [13], in this article we apply Zhu and Lee’s [31] local influence approach to the MPCFA model by treating the latent continuous measurements and the latent factor scores in the model as missing data. The local influence measures are developed on the basis of the conformal normal curvature [23] and MCMC methods such as the Gibbs sampler [10] and the Metropolis–Hastings (MH) algorithm [19,12]. It will be shown that the proposed approach involves no intractable integrals, hence its computational burden is not heavy.

The paper is organized as follows. Section 2 defines the MPCFA model and briefly depicts the MCEM algorithm for obtaining the ML estimates. In Section 3, we give a brief sketch of the local influence approach for models with incomplete data and develop the methodology required for MPCFA models. Some perturbation schemes are discussed. To illustrate the developed methodology, an artificial example and a real example are analyzed in Section 4. Concluding remarks are given in Section 5. Some technical details are given in the appendices.

2. The MPCFA model

For introducing some notation and motivating the proposed local influence approach, we define the MPCFA model [27], and roughly describe the related procedure for sampling observations at the E-step of the EM algorithm for ML estimation in this section.

In the MPCFA model, it is assumed that each subject has a covariate vector that can be any mixture of discrete and continuous variables, and each subject produces $J$ distinct quantal responses or is classified with respect to $J$ dichotomous categories. More specifically, let $u_i = (u_{i1}, \ldots, u_{iJ})'$ denote the collection of observed dichotomous 0/1 responses in $J$ variables on the $i$th subject, $i = 1, \ldots, n$, $x_{ij}$ be a $kj \times 1$ vector of covariates, $k = \sum_{j=1}^{J} k_j$, and

$$X_i = \begin{bmatrix}
 x_{i1}' & 0 & \cdots & 0 \\
 0 & x_{i2}' & \cdots & 0 \\
 \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \cdots & x_{iJ}'
\end{bmatrix},$$

be a $J \times k$ matrix. Let $z_i = (z_{i1}, \ldots, z_{iJ})'$ denote a $J$-variate normal vector of “response strengths” such that for $j = 1, \ldots, J$,

$$u_{ij} = 1 \quad \text{if } z_{ij} > 0; \quad u_{ij} = 0 \quad \text{if } z_{ij} \leq 0. \quad (1)$$

Here, the exact measurement of “response strengths” $z_i$ is not observed, and its information is given by an observed dichotomous vector $u_i = (u_{i1}, \ldots, u_{ij})'$ with $u_{ij}$ defined by Eq. (1). In the MPCFA model, $z_i$ is modeled by incorporating the covariates $X_i$ and a CFA model:

$$z_i = X_i \beta + \Lambda \zeta_i + \epsilon_i, \quad i = 1, \ldots, n, \quad (2)$$

where $\beta' = (\beta'_1, \ldots, \beta'_J)$, $\beta_j$ is a $k_j \times 1$ unknown parameter vector, $\epsilon_i$ is a $J \times 1$ vector of residuals. Here, $\beta$ is the $k \times 1$ vector of regression coefficients of $z_i$ on $X_i$. We also assume that $\zeta_i$
is independently distributed as $\mathcal{N}_q(0, \Phi)$, $\epsilon_i$ is independently distributed as $\mathcal{N}_J[0, \Psi]$, where $\Phi$ is an arbitrary covariance or correlation matrix, $\Psi$ is a diagonal covariance matrix, and $\zeta_i$ and $\epsilon_i$ are uncorrelated. Note that the MP model in Bock and Gibbons [4], and Gibbons and Wilcox-Gök [11] assumed that $\zeta_i \overset{D}{=} \mathcal{N}_q(0, I)$. It has been shown by Song and Lee [27] that the model can be identified by fixing $\Psi$ to be a diagonal matrix with preassigned diagonal elements, and some elements in $\Lambda$ and/or $\Phi$ at preassigned values.

Let $U = (u_1, \ldots, u_n)$ be the observed data matrix of the dichotomous outcomes, $Z = (z_1, \ldots, z_n)$ be the matrix of latent continuous measurements underlying $U$, $\Omega = (\zeta_1, \ldots, \zeta_n)$ be the matrix of latent variables, and $\theta$ be the parameter vector which contains all distinct unknown parameters in $\beta$, $\Lambda$ and $\Phi$. As the observed-data log-likelihood function, $L_0(U; \theta)$, involves complex high-dimensional integrals, it is very difficult to directly work with it. In their estimation procedure, Lee [27], is not presented. In the following section, we derive the local influence measures with a given ML estimate.

3. Local influence of the MPCFA model

3.1. Local influence based on a $Q$-displacement function

The general approach that was developed in Zhu and Lee [31] for local influence analysis of incomplete data models will be applied to obtain the local influence measures for the MPCFA model.
Consider a perturbation vector \( \omega = (\omega_1, \ldots, \omega_n)' \) which is in some open set \( \Theta \subseteq \mathbb{R}^n \). Let \( L_c(\mathbf{D}; \theta, \omega) \) be the complete-data log-likelihood of the perturbed model. As usual, we assume that there is a null point \( \omega^0 \) such that \( L_c(\mathbf{D}; \theta, \omega^0) = L_c(\mathbf{D}; \theta) \) for all \( \theta \). Let \( \hat{\theta}(\omega) \) be the ML estimate of \( \theta \) for the perturbed model which maximizes

\[
Q(\theta, \omega|\hat{\theta}) = E[L_c(\mathbf{D}; \theta, \omega)|\mathbf{U}, \hat{\theta}].
\]

(4)

Obviously, \( \hat{\theta}(\omega^0) = \hat{\theta} \). Inspired by Zhu and Lee [31], we consider the following \( Q \)-displacement function:

\[
f_Q(\omega) = 2\{Q(\hat{\theta}|\hat{\theta}) - Q(\hat{\theta}(\omega)|\hat{\theta})\}.
\]

(5)

Note that in Cook’s [6] approach, the following likelihood-displacement function was considered:

\[
LD(\omega) = 2\{L_0(\mathbf{U}; \hat{\theta}) - L_0(\mathbf{U}; \hat{\theta}_w, \omega)\},
\]

where \( \hat{\theta}_w^* \) is the vector that maximizes \( L_0(\mathbf{U}; \theta, \omega) \), and the local behavior of \( LD(\omega) \) is studied by examining the normal curvature of the influence graph \( \gamma_L(\omega) = (\omega', LD(\omega))' \) for developing the local influence measures. But it is very difficult to obtain the local influence measures based on the \( LD(\omega) \) for some complicated model, for example, MPCFA model discussed in this paper, because the building-blocks in the associated diagnostic measures involve intractable integrals which are inherited from the observed-data likelihood for the incomplete-data models. But if we use the \( Q \)-displacement function \( f_Q(\omega) \) instead of the likelihood-displacement function, then we can avoid the difficulty mentioned above. It has been shown that (see [31]) \( f_Q(\omega) \) has similar properties as \( LD(\omega) \); for example, it is the measure between \( \hat{\theta} \) and \( \hat{\theta}(\omega) \), \( f_Q(\omega^0) = 0 \), and \( \hat{c} f_Q(\omega)/\hat{c} \omega = 0 \), etc. Following the key idea of Cook’s [6] approach, we study the local behavior of the \( Q \)-displacement function by examining the corresponding influence graph.

The influence graph of \( f_Q(\omega) \) is defined as

\[
\gamma(\omega) = (\omega', f_Q(\omega))' .
\]

(6)

In differential geometry, a surface of this form is frequently called Monge patch [20]. Based on the same reasoning as given in Cook [6], the normal curvature \( C_{f_Q, h} \) of \( \gamma(\omega) \) at \( \omega^0 \) in the direction of a unit vector \( h \) can be used to summarize the local behavior of the object function \( f_Q(\omega) \). It can be shown that the normal curvature \( C_{f_Q, h} \) of \( \gamma(\omega) \) at \( \omega^0 \) in the direction of a unit vector \( h \) is

\[
C_{f_Q, h} \triangleq -2h' \hat{Q}_{\omega^0} h = -2h' \Delta_{\omega^0} (\hat{Q}_{\omega}(\hat{\theta}))^{-1} \Delta_{\omega^0} h,
\]

(7)

where

\[
\hat{Q}_{\omega}(\hat{\theta}) \triangleq \left. \frac{\partial^2 Q(\hat{\theta})}{\partial \hat{\theta} \partial \hat{\theta}'} \right|_{\theta = \hat{\theta}} \quad \text{and} \quad \Delta_{\omega^0} \triangleq \left. \frac{\partial^2 \hat{Q}(\hat{\theta}(\omega)|\hat{\theta})}{\partial \hat{\theta} \partial \hat{\theta}'} \right|_{\theta = \hat{\theta}, \omega = \omega^0}.
\]

Under some mild regularity conditions, \( \hat{Q}_{\omega}(\hat{\theta}) \) and \( \hat{Q}_{\omega^0} \) are semi-positive definite. The normal curvature based on the likelihood-displacement function proposed by Cook [6] may take any value and is not invariant under a uniform change of scale. And the above normal curvature based on the \( Q \)-displacement function also possesses these drawbacks. Inspired by Poon and Poon [23], a conformal normal curvature is employed for our procedure. This kind of curvatures is a one-to-one function of the normal curvature and takes value in the closed interval \([0, 1]\). Moreover, this
curvature is invariant with respect to conformal reparameterization of \( \omega \). Based on the reasoning given in Poon and Poon [23], and Zhu and Lee [31], the conformal normal curvature \( B_{fQ, h} \) at \( \omega^0 \) in the direction of a unit vector \( h \) is given as follows:

\[
B_{fQ, h} = \frac{-2 h^i \ddot{Q}_{\omega^0} h^j}{tr[-2 \dot{Q}_{\omega^0}]}.
\]  

(8)

Let \( B = -2 \ddot{Q}_{\omega^0} / tr[-2 \dot{Q}_{\omega^0}] \), and \( \lambda_1 \geq \cdots \geq \lambda_r > 0 \) be the \( r \) non-zero eigenvalues of \( B \), and \( e_1, \ldots, e_r \) be the corresponding orthogonal eigenvectors. According to Lesaffre and Verbeke [17], Poon and Poon [23], and Zhu and Lee [31], the following aggregate contribution vector of all eigenvectors that are associated with all non-zero eigenvalues

\[
M(0) = \sum_{i=1}^{r} \lambda_i e_i^2,
\]

where \( e_i^2 = (e_i^1, \ldots, e_i^m)' \), is used for assessing local influence. For \( j = 1, \ldots, m \), it follows from Zhu and Lee [31] that the \( j \)th component of \( M(0) \), \( M(0)_j = b_{jj} \) for \( j = 1, \ldots, m \), where \( b_{jj} \) is the \( j \)th diagonal element of matrix \( B \). So, it is very simple to compute \( b_{jj} \) and largely reduce the computing burden, because no eigenfunctions and eigenvalues are involved. Therefore, our local influence measures are based on the conformal normal curvature rather than the classical normal curvature, because the conformal normal curvature possesses the above-mentioned nice properties.

To get the basic building blocks of local influence measures, we need to derive expressions for \( \Delta_{\omega^0} \) and \( \dot{Q}_\theta(\hat{\theta}) \) (see (7)). Assuming the legitimacy of interchange of integration and differentiation, we have

\[
\dot{Q}_\theta(\hat{\theta}) = E \left[ \frac{\partial^2 L_c(D; \theta)}{\partial \theta \partial \theta'} \bigg| U, \hat{\theta} \right] \bigg|_{\theta = \hat{\theta}},
\]  

(9)

\[
\Delta_{\omega^0} = E \left[ \frac{\partial^2 L_c(D; \theta, \omega)}{\partial \theta \partial \omega'} \bigg| U, \hat{\theta} \right] \bigg|_{\theta = \hat{\theta}, \omega = \omega^0}.
\]  

(10)

Let \( \ddot{L}_c(D; \theta) = \frac{\partial^2 L_c(D; \theta)}{\partial \theta \partial \theta'} \). It follows from (3) that

\[
L_c(D; \theta) = L_{c, 1}(D; \beta, \Lambda) + L_{c, 2}(\Omega; \Phi),
\]  

(11)

where

\[
L_{c, 1}(D; \beta, \Lambda) = -\frac{Jn}{2} \log(2\pi) - \frac{n}{2} \log |\Psi| - \frac{1}{2} \sum_{i=1}^{n} (z_i - X_i \beta - \Lambda \zeta_i)' \Psi^{-1} (z_i - X_i \beta - \Lambda \zeta_i),
\]

\[
L_{c, 2}(\Omega; \Phi) = -\frac{qn}{2} \log(2\pi) - \frac{n}{2} \log |\Phi| - \frac{1}{2} \sum_{i=1}^{n} \zeta_i' \Phi^{-1} \zeta_i.
\]

Because \( L_{c, 1} \) and \( L_{c, 2} \) are functions with separable parameters, \( \ddot{L}_c(D; \theta) \) is a diagonal block matrix. For completeness, expressions for \( \ddot{L}_c(D; \theta) \) are listed in Appendix B. The conditional expectations of the second derivatives in (9) and (10) cannot be evaluated in closed forms. We
overcome this difficulty via the Monte Carlo approximation. Let \((Z^{(m)}, \Omega^{(m)}); m = 1, \ldots, M\) be a sample randomly drawn from the joint conditional distribution \([Z, \Omega] \mid U, \theta\), the building blocks can be approximated by

\[
\tilde{Q}_\theta(\hat{\theta}) \approx \frac{1}{M} \sum_{m=1}^{M} \frac{\partial^2 L_c(U, Z^{(m)}, \Omega^{(m)}; \theta)}{\partial \theta \partial \theta'} \bigg|_{\theta=\hat{\theta}},
\]

\[
\Delta_{\omega^0} \approx \frac{1}{M} \sum_{m=1}^{M} \frac{\partial^2 L_c(U, Z^{(m)}, \Omega^{(m)}; \theta, \omega)}{\partial \theta \partial \omega'} \bigg|_{\theta=\hat{\theta}, \omega=\omega^0}.
\]

This random sample can usually be obtained by the sampling-based procedure that is developed for simulating observations at the E-step of the MCEM algorithm for the ML estimation. Hence, the additional programming effort is light.

### 3.2. Perturbation schemes

It follows from (7) and (8) that the local influence measures \(M(0)_j\) depend on \(\tilde{Q}_\theta(\hat{\theta})\) and \(\Delta_{\omega^0}\) with respect to each perturbation scheme \(\omega\). We note from (11) that \(L_c(D; \theta)\) contains two separate terms that involve different separable functions. This sample form of \(L_c(D; \theta)\) has the following two advantages. First, Hessian matrix is a diagonal block matrix. This saves a lot of programming and computational efforts in evaluating \(\Delta_{\omega^0}, \tilde{Q}_\theta(\hat{\theta}), \tilde{Q}_{\omega^0}, B_{f_0, h}, \) and hence \(M(0)_j\). Second, it is more flexible to consider perturbations for assessing sensitivity of various aspects of model inputs, see the perturbations below. Moreover, interpretation of the perturbations is more apparent. Some interesting perturbations of the MPCFA model are given as below:

**Scheme 1: Perturbation on cell frequencies.** The observations in a MPCFA model are the cell frequencies of \(n_0\) distinct cells (or \(n_0\) response patterns). A common interest is to identify the influential cells via minor perturbation on the cell frequencies. Let \(\omega = (\omega_1, \ldots, 1)'\). The cell frequency \(\pi_i\) is perturbed to \(\omega_i \pi_i\), for \(i = 1, \ldots, n_0\). Apart from a constant, the perturbed complete-data log-likelihood function \(L_c(D; \theta, \omega)\) with respect to this kind of data structure is given by

\[
-\frac{1}{2} \sum_{i=1}^{n_0} \omega_i \pi_i \left\{ (z_i - X_i \beta - \Lambda \varsigma_i)' \Psi^{-1} (z_i - X_i \beta - \Lambda \varsigma_i) + \log |\Phi| + \varsigma_i' \Phi^{-1} \varsigma_i \right\}.
\]

For \(i = 1, \ldots, n_0; j = 1, \ldots, J\), we have

\[
\frac{\partial^2 L_{c,1}}{\partial \beta \partial \omega_j} = \pi_i X_i' \Psi^{-1} (z_i - X_i \beta - \Lambda \varsigma_i),
\]

\[
\frac{\partial^2 L_{c,1}}{\partial \Lambda_j \partial \omega_i} = \psi_{ij}^{-1} \pi_i (z_{ij} - X_{ij} \beta_j - \Lambda_j \varsigma_i \varsigma'_i),
\]

\[
\frac{\partial^2 L_{c,2}}{\partial \Omega \partial \omega_i} = \frac{\pi_i}{2} \Phi^{-1} [\varsigma'_i \varsigma_i - \Phi] \Phi^{-1}.
\]

**Scheme 2: Perturbation on explanatory variables.** Similar to Cook’s [6] method for perturbing the design matrix in regression, we define the following scale matrix to account for the different
measurement units associated with each element of \( X_j \):

\[
S = \begin{bmatrix}
s'_1 & 0 & \cdots & 0 \\
0 & s'_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & s'_j 
\end{bmatrix},
\]

where \( s_j \) is a \( k_j \times 1 \) constant vector, \( j = 1, \ldots, J \). In this scheme, the perturbation vector is \( \omega = (\omega_1, \ldots, \omega_n)' \), and \( X_j \) is replaced by \( X_j(\omega) = X_j + \omega_j S \). The null point is \( \omega^0 = (0, \ldots, 0)' \). The perturbed complete-data log-likelihood is given by

\[
L_c(\mathbf{D}; \theta, \omega) = L_{c,1}(\mathbf{D}; \beta, \Lambda, \omega) + L_{c,2}(\Omega; \Phi),
\]

where \( L_{c,2}(\Omega; \Phi) \) is the same as one in (11), and \( L_{c,1}(\mathbf{D}; \beta, \Lambda, \omega) \) is given by

\[
\frac{-Jn}{2} \log(2\pi) - \frac{n}{2} \log |\Psi| - \frac{1}{2} \sum_{i=1}^{n} (z_i - (X_i + \omega_i S)\beta - \Lambda z_i)' \times \Psi^{-1}(z_i - (X_i + \omega_i S)\beta - \Lambda z_i).
\]

(16)

It can be shown that for \( i = 1, \ldots, n; j = 1, \ldots, J \),

\[
\frac{\partial^2 L_{c,1}(\mathbf{D}; \beta, \Lambda, \omega)}{\partial \beta \partial \omega_i} = -X_i' \Psi^{-1} S \beta,
\]

\[
\frac{\partial^2 L_{c,1}(\mathbf{D}; \beta, \Lambda, \omega)}{\partial \Lambda_j \partial \omega_i} = -\psi_{jj}^{-1} s'_j \beta_j z_i.'
\]

Clearly, \( \frac{\partial^2 L_{c,2}(\Omega; \Phi)}{\partial \Omega \partial \omega'} = 0 \).

Scheme 3: Perturbation on latent variables. Consider an additive perturbation scheme on \( \zeta \) via a vector \( \omega = (\omega_1, \ldots, \omega_n)' \), such that \( \zeta_i(\omega) = \zeta_i + \omega_i 1_q \), where \( 1_q = (1, \ldots, 1)' \). In this case, \( \omega^0 = (0, \ldots, 0)' \). Ignoring a constant, the complete-data log-likelihood function for the perturbed model \( L_c(\mathbf{D}; \theta, \omega) \) is given by

\[
\frac{-1}{2} \sum_{i=1}^{n} \left \{ (z_i - X_i \beta - \Lambda (\zeta_i + \omega_i 1_q))' \Psi^{-1} (z_i - X_i \beta - \Lambda (\zeta_i + \omega_i 1_q)) + \log |\Phi| + (\zeta_i + \omega_i 1_q)' \Phi^{-1} (\zeta_i + \omega_i 1_q) \right \}.
\]

(17)

It can be shown that for \( i = 1, \ldots, n; j = 1, \ldots, J \),

\[
\frac{\partial^2 L_{c,1}(\mathbf{D}; \beta, \Lambda, \omega)}{\partial \beta \partial \omega_i} = -X_i' \Psi^{-1} \Lambda 1_q,
\]

\[
\frac{\partial^2 L_{c,1}(\mathbf{D}; \beta, \Lambda, \omega)}{\partial \Lambda_j \partial \omega_i} = \psi_{jj}^{-1} (\Lambda_j 1_q (\omega_i 1_q' + \zeta_i') + [z_{ij} - x_{ij} \beta_j - \Lambda_j (\omega_i 1_q + \zeta_i)] 1_q'),
\]

\[
\frac{\partial^2 L_{c,2}(\Omega; \Phi, \omega)}{\partial \Omega \partial \omega_i} = \frac{1}{2} \Phi^{-1} (\omega_i 1_q' + \zeta_i') + (\omega_i 1_q + \zeta_i) 1_q') \Phi^{-1}.
\]
Scheme 4: Perturbation on all unknown parameters. Let \( \omega_1 \) and \( \omega_2 \) the perturbation vectors corresponding to \( \theta_1 = (\beta, \Lambda) \) and \( \theta_2 = \Phi \), respectively. The perturbed parameter vectors are defined as follows:

\[
\theta_1(\omega) = \theta_1 + U_1 \omega_1, \quad \theta_2(\omega) = \theta_2 + U_2 \omega_2,
\]

where \( U_1 \) and \( U_2 \) are diagonal matrices of appropriate orders that can be chosen according to the investigators special concerns. In this case, \( \omega_1^0 = (0, \ldots, 0)' \) and \( \omega_2^0 = (0, \ldots, 0)' \). The complete-data log-likelihood function of the perturbed model is

\[
L_c(D; \theta(\omega)) = L_{c,1}(D; \theta_1(\omega)) + L_{c,2}(D; \theta_2(\omega)).
\]

Hence,

\[
\tilde{Q}_{\omega^0} = \begin{pmatrix} \frac{\partial \theta(\omega)}{\partial \omega'} & \frac{\partial^2 L_c(D; \theta(\omega))}{\partial \theta(\omega) \partial \theta(\omega)'} & \frac{\partial \theta(\omega)}{\partial \omega'} \end{pmatrix}.
\]

where

\[
\frac{\partial \theta(\omega)}{\partial \omega'} = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix},
\]

with \( U_1 = \partial \theta_1(\omega) / \partial \omega'_1 \) and \( U_2 = \partial \theta_2(\omega) / \partial \omega'_2 \).

Influential parameters should receive more attention in the statistical analysis of the model. For example, when drawing statistical conclusions using an interval of an influential parameter, an interval with higher confidence is more desirable. In model modification, we should avoid fixing an influential parameter because two slightly different preassigned values will give quite different statistical results.

4. Numerical illustrations

4.1. An artificial example

Results obtained from analysis of an artificial example are presented here to illustrate the performance of the diagnostic measures. A MPCFA model defined in (1) and (2) with nine item response variables, three fixed coefficient parameters \((\beta_0, \beta_1, \beta_2)\), and three latent variables \((\xi^1, \xi^2, \xi^3)\) is considered. The loading matrix \( \Lambda \) is specified as

\[
\Lambda^T = \begin{bmatrix} \lambda_{11} & \lambda_{21} & \lambda_{31} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{42} & \lambda_{52} & \lambda_{62} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{73} & \lambda_{83} & \lambda_{93} \end{bmatrix},
\]

where the \( \lambda_{ij} \)'s are the unknown factor loading parameters. To identify the model, the 0's in \( \Lambda \) are fixed, \( \Psi \) is fixed to be an identity matrix, and \( \Phi = (\varphi_{ij}) \) is taken to be a correlation matrix. True population values of unknown parameters are given by: \( \lambda_{11} = \lambda_{21} = \lambda_{31} = \lambda_{42} = \lambda_{52} = \lambda_{62} = \lambda_{73} = \lambda_{83} = \lambda_{93} = 0.8, (\varphi_{12}, \varphi_{13}, \varphi_{23}) = (0.3, 0.3, 0.3) \), and

\[
\beta' = \begin{bmatrix} -0.5 & -0.5 & -0.5 & -0.5 & -0.5 & -0.5 & -0.5 & -0.5 & -0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \end{bmatrix}.
\]
Fig. 1. Index plots of $M(0)_j$ for perturbation on explanatory variables: artificial data.

Fig. 2. Index plots of $M(0)_j$ for perturbation on latent variables $\zeta_i$: artificial data.

For the fixed covariates, we generated $x_{ij}$ from $\text{Gamma}(j, 1)$ distribution for $i = 1, \ldots, 300; j = 1, 2, 3$. Hence, the fixed covariates are continuous. The ML estimates of parameters are obtained via the MCEM algorithm as described in Song and Lee [27].

We generate a random sample with size 300. Based on the data set generated, two data sets with artificial outliers are created. First, to consider the perturbation on explanatory variables, we add 10 to each entry of the original $x_i$ for $i = 10, 20, 50, 290$. Plots of $M(0)_j$ for this perturbation scheme are shown in Fig. 1. Only the 10th, 20th, 50th, and 290th cases have relatively larger $M(0)_j$’s, so they are identified as influential. Second, to illustrate the perturbation on latent variables, we keep the original $x_i$, but subtract 10 from each entry of the original $\zeta_i$ for $i = 46, 125, 203, 274, 277$. Plots of the local influence measures $M(0)_j$ corresponding to the perturbation on latent variables are presented in Fig. 2. As we expected, only the 46th, 125th, 203th, 274th, 277th observations are detected as influential cases.

The above empirical findings indicate that the local influence measures corresponding to the perturbations have detected what they supposed to detect, and give no false influential cases.

4.2. A real example on compliance study of patients

It has been pointed out that patient adherence to prescribed medication is crucial to the success of medical treatment [7] and that non-adherence leads to misjudgment of the effectiveness of medication [25]. To enrich existing knowledge about patient non-adherence, the Department of Medicine and Therapeutics, Community and Family Medicine, and Pharmacy at the Chinese University of Hong Kong conducted a survey of ethnic Chinese patients who had been diagnosed
as suffering from hypertension. One objective was to measure and examine correlations among latent variables such as physician advice and concern, patient knowledge and belief, social cognition, and social influence, and the subsequent study reported non-adherence with reference to a factor analysis model. Because the study involved many dichotomous variables and the manifest indicators for the factors are influenced by covariates, it has been used in Song and Lee [27] as an illustrative example of the MPCFA model. In this article, we also use this real data set and the same MPCFA model that have been used in Song and Lee [27] to illustrate the local influence analysis after ML estimation.

Nine dichotomous manifest variables are selected as indicators of the latent variables who are patient “non-adherence”, “knowledge of medication”, and “health condition”. The questions are listed in Table 1, together with their frequencies. Two fixed covariates about patient education (coded by 0, 1, 2, 3), $x_1$ and the existence of “side-effects” (coded by 0 and 1), $x_2$. For brevity, we omit a small number of observations with missing entries, and the remaining sample size is 837. The loading matrix $\Lambda'$ of the MPCFA model is specified as (see [27]):

$$\Lambda' = \begin{bmatrix} \hat{\lambda}_{11} & \hat{\lambda}_{21} & \hat{\lambda}_{31} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{\lambda}_{42} & \hat{\lambda}_{52} & \hat{\lambda}_{62} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \hat{\lambda}_{73} & \hat{\lambda}_{83} & \hat{\lambda}_{93} \end{bmatrix},$$

where the $\hat{\lambda}_{ij}$'s are the unknown factor loading parameters, while the 0's are fixed for identifying the model. To identify the model, we also fix $\Psi$ to be an identity matrix and $\Phi = (\phi_{ij})$ to be a correlation matrix. For completeness, the ML estimates of parameters are reported in Table 2.

We first consider the perturbation on cell frequency (see Scheme 1). For this data set, there are a total of $2^9 \times 4 \times 2 = 4096$ cells, and the number of non-empty cells, $n_0$, is equal to 407. These cells are indexed by $i = 1, \ldots, n_0$; and each corresponding to a 11-dimensional cell. Plots of $M(0)_j$ are displayed in Fig. 3. From this figure, we identify the 16th, 70th, 104th, and 153th cells as influential. The covariates and response patterns of the dichotomous variables corresponding to these influential cells are presented in Table 3. From Fig. 3, the cell frequency corresponding to (00010000020) is most influential. Therefore, the doctors should pay more attention to the patients with this pattern.

Table 1

<table>
<thead>
<tr>
<th>Questions associated with the manifest variables</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$: Did you have any surplus in the previous prescribed drugs?</td>
<td>(175/662)</td>
</tr>
<tr>
<td>$u_2$: Did you stop/reduce/increase the dosage?</td>
<td>(69/768)</td>
</tr>
<tr>
<td>$u_3$: Did you forget to take medications?</td>
<td>(391/446)</td>
</tr>
<tr>
<td>$u_4$: Do you feel you have hypertension?</td>
<td>(363/474)</td>
</tr>
<tr>
<td>$u_5$: Do you know the reasons for taking drugs?</td>
<td>(650/187)</td>
</tr>
<tr>
<td>$u_6$: Do you know the reasons for taking drugs for a long term?</td>
<td>(605/232)</td>
</tr>
<tr>
<td>$u_7$: In the past 2 weeks, did you have emotional problems?</td>
<td>(387/450)</td>
</tr>
<tr>
<td>$u_8$: In the past 2 weeks, did your health cause any difficulties in daily activities?</td>
<td>(181/656)</td>
</tr>
<tr>
<td>$u_9$: In the past 2 weeks, did your health cause any difficulties in social activities?</td>
<td>(177/660)</td>
</tr>
</tbody>
</table>

Note: Frequencies of (Yes ‘1’ / No ‘0’) are in parentheses.
Table 2
ML estimates in the compliance study of patients

<table>
<thead>
<tr>
<th>Para.</th>
<th>MLE</th>
<th>MLE</th>
<th>MLE</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{11}$</td>
<td>1.091</td>
<td>$\beta_{11}$</td>
<td>-0.722</td>
<td>$\beta_{12}$</td>
</tr>
<tr>
<td>$\lambda_{21}$</td>
<td>1.418</td>
<td>$\beta_{21}$</td>
<td>-1.207</td>
<td>$\beta_{22}$</td>
</tr>
<tr>
<td>$\lambda_{31}$</td>
<td>0.311</td>
<td>$\beta_{31}$</td>
<td>-0.022</td>
<td>$\beta_{32}$</td>
</tr>
<tr>
<td>$\lambda_{42}$</td>
<td>0.099</td>
<td>$\beta_{41}$</td>
<td>-0.121</td>
<td>$\beta_{42}$</td>
</tr>
<tr>
<td>$\lambda_{52}$</td>
<td>1.370</td>
<td>$\beta_{51}$</td>
<td>0.876</td>
<td>$\beta_{52}$</td>
</tr>
<tr>
<td>$\lambda_{62}$</td>
<td>1.471</td>
<td>$\beta_{61}$</td>
<td>0.729</td>
<td>$\beta_{62}$</td>
</tr>
<tr>
<td>$\lambda_{73}$</td>
<td>0.658</td>
<td>$\beta_{71}$</td>
<td>-0.081</td>
<td>$\beta_{72}$</td>
</tr>
<tr>
<td>$\lambda_{83}$</td>
<td>2.271</td>
<td>$\beta_{81}$</td>
<td>-1.179</td>
<td>$\beta_{82}$</td>
</tr>
<tr>
<td>$\lambda_{93}$</td>
<td>2.244</td>
<td>$\beta_{91}$</td>
<td>-1.163</td>
<td>$\beta_{92}$</td>
</tr>
</tbody>
</table>

Fig. 3. Index plots of $M(0)_j$ for perturbation on cell frequencies: compliance study of patients data.

Table 3
Influential cells in the compliance study of patients

<table>
<thead>
<tr>
<th>Cell no.</th>
<th>$u_1$</th>
<th>$u_2$</th>
<th>$u_3$</th>
<th>$u_4$</th>
<th>$u_5$</th>
<th>$u_6$</th>
<th>$u_7$</th>
<th>$u_8$</th>
<th>$u_9$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>MRC</th>
<th>TRC</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.256</td>
<td>2.077</td>
<td>1.032</td>
</tr>
<tr>
<td>70</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.287</td>
<td>2.340</td>
<td>1.047</td>
</tr>
<tr>
<td>104</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0.352</td>
<td>3.188</td>
<td>1.068</td>
</tr>
<tr>
<td>153</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.281</td>
<td>1.898</td>
<td>1.036</td>
</tr>
<tr>
<td>250</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.196</td>
<td>1.191</td>
</tr>
<tr>
<td>300</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.185</td>
<td>1.178</td>
<td>1.024</td>
</tr>
</tbody>
</table>

To study the sensitivity of the ML estimates with respect to an influential cell, we remove that influential cell and obtain the ML estimate $\hat{\theta}^p$ on the basis of the remaining data. The following two quantities are used to measure the difference between the original ML estimate, $\hat{\theta}$, and $\hat{\theta}^p$.

**Total Relative Changes:**

$$TRC = \sum_{i=1}^{n_p} |\hat{\theta}_i - \hat{\theta}^p_i|/\hat{\theta}_i,$$
and Maximum Relative Changes:

\[ MRC = \max_i |\hat{\theta}_i - \hat{\theta}_i^o|/\hat{\theta}_i, \]

where \( n_p \) is the number of parameters. Moreover, we compute the observed-data likelihood ratio, \( p_o(\hat{\theta}; U)/p_o(\hat{\theta}^o; U) \), to reveal the impact of the influential cells. For comparison sake, we randomly select two non-influential cells, and repeat the above analysis in obtaining the corresponding TRC, MRC, and observed-data likelihood ratio. The results are summarized in Table 3. From this table, we observe that the most influential cell (the 104th cell) has the largest TRC, MRC, and relative change in the observed-data likelihood. As expected, the influential cells have much stronger impact to the ML results than the non-influential cells. To seek for the possible reason for this phenomenon, we pay more attention to these influential cells and compare them with the non-influential cells. We find out that the frequencies of the influential cells (16th, 70th, 104th, and 153th) are 12, 22, 44, and 46, respectively, and which are significantly larger than the frequencies of the non-influential cells. Hence, influential cells associate with large frequencies. This reasonable finding agrees with the conclusion given in Poon et al. [24], and Lee and Xu [15].

We use the perturbation Scheme 3 to study the effect of an additive perturbation to latent variables \( \zeta_i \), for \( i = 1, \ldots, n \). Plots of \( M(0) \) are presented in Fig. 4. From this figure, we see that the 34th, 52th, 71st, 85th, 102th, 158th, 166th, 186th, 329th, 384th, 449th, and 557th observations are identified as the influential. The patterns of the covariates and the dichotomous variables of these 12 influential observations are presented in Table 4. We observe that they do not fall in the influential cells that are identified by the perturbation of cell frequency. Hence, these two schemes work complementary in local influence analysis. To compare the impact of these influential observations to the ML results, we repeat the above analysis without the influential observations. We obtain \( TRC = 2.084, MRC = 0.245, \) and the observed-data likelihood ratio is 1.021. In order to compare the impact of the non-influential observations, we repeat the analysis after removing 12 randomly selected non-influential observations. We get \( TRC = 2.047, MRC = 0.187, \) and the observed-data likelihood ratio is 1.006. Hence, the ML results are more sensitive to the influential observations.

5. Conclusion

The CFA model has been extensively applied to behavioral, psychological, and social research for assessing the latent traits of manifest variables. Recently, this model has also received much
Table 4
Influential cases in the compliance study of patients

<table>
<thead>
<tr>
<th>Case no.</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>52</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>71</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>85</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>102</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>158</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>166</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>186</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>329</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>384</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>449</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>557</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Attention in biostatistics, and have been widely applied to medical research (see for example [2,16]). The MPCFA model proposed by Song and Lee [27] combines the MP model in biostatistics and the CFA model in psychometrics. It represents a more general model for analyzing multivariate dichotomous data.

Identification of influential observations is an important step in data analysis. As pointed out by Cook [6] and many others, it is important to study the sensitivity of the ML results in relation to the model and data inputs. Due to the complexity of the observed-data log-likelihood functions in MPCFA models, it is very difficult to obtain local influence measures by Cook’s [6] approach. To overcome this difficulty, we treat the latent variables as hypothetical missing data and develop the local influence measures based on the Q-displacement function, instead of the more complicated observed-data log-likelihood function. Several perturbation schemes are considered. Our empirical studies show that the proposed method is feasible, and is able to identify the influential aspects.

Acknowledgements

This work was fully supported by a grant from the Chinese University of Hong Kong (Project No. CUHK 2060279). The authors are grateful to Juliana C.N. Chan, Professor, Department of Medicine and Therapeutics, CUHK, for providing the data in the example.

Appendix A.

The required conditional distributions involved in the Gibbs sampler are briefly described as below.

\[ [\mathbf{Z} | \Omega, \mathbf{U}, \theta] \text{: Let } \Lambda_j \text{ be the } j\text{th row of } \Lambda, \text{ and } \psi_{jj} \text{ be the } j\text{th diagonal element of } \Psi. \text{ As } \mathbf{z}_i \text{ are mutually independent, it follows from the definition of the model that} \]

\[ p(\mathbf{Z} | \Omega^{(m)}, \mathbf{U}, \theta) = \prod_{i=1}^{n} p(z_i | \zeta_i, \mathbf{u}_i, \theta) = \prod_{i=1}^{n} \prod_{j=1}^{J} p(z_{ij} | \zeta_i, u_{ij}, \theta), \]  

(A.1)
where

\[
[z_{ij}|\xi_i, u_{ij}, \theta] \overset{D}{=} \begin{cases} 
N[x_{ij}^T \beta_j + A_j \xi_i, \psi_{jj}] \mathcal{I} \{z_{ij} \in (-\infty, 0)\}, & u_{ij} = 0, \\
N[x_{ij}^T \beta_j + A_j \xi_i, \psi_{jj}] \mathcal{I} \{z_{ij} \in (0, +\infty)\}, & u_{ij} = 1.
\end{cases}
\]

Note that (A.1) involves univariate rather than multivariate truncated normal distributions. The commonly used inverse distribution method [9] can be employed to simulate observations from this relatively simple distribution.

[\Omega|\mathbf{Z}, \mathbf{U}, \theta]: As \Omega is independent of \mathbf{U} with \mathbf{Z} given, and \xi_i, i = 1, \ldots, n are mutually independent, we have

\[
p(\Omega|\mathbf{Z}, \mathbf{U}, \theta) = p(\Omega|\mathbf{Z}, \theta) = \prod_{i=1}^{n} p(\xi_i|\mathbf{z}_i, \theta),
\]

where \( [\xi_i|\mathbf{z}_i, \theta] \overset{D}{=} \mathcal{N}_p(\Sigma^* \mathbf{A}' (\mathbf{z}_i - \mathbf{X}_i \beta), \Sigma^*) \), with \( \Sigma^* = (\Phi^{-1} + A \Psi^{-1} \mathbf{A}')^{-1} \). The simulation of observations from this multivariate normal distribution is fast and straightforward.

Appendix B.

Let \( \delta_{jk} \) be the Kronecker delta and \( \phi_s \) be the \( s \)th element of \( \Phi \), and \( n_\phi \) be the number of unknown distinct parameters in \( \Phi \). For \( j, \ell = 1, \ldots, J; s, t = 1, \ldots, n_\phi; i = 1, \ldots, n \).

\[
\frac{\partial^2 L_{c,1}(\mathbf{D}; \beta, \mathbf{A})}{\partial \beta \partial \beta'} = -\sum_{i=1}^{n} \mathbf{X}_i \Psi^{-1} \mathbf{X}_i,
\]

\[
\frac{\partial^2 L_{c,1}(\mathbf{D}; \beta, \mathbf{A})}{\partial \beta_j \partial \mathbf{A}'_{\ell}} = -\delta_{jj} \psi_{jj}^{-1} \sum_{i=1}^{n} \mathbf{x}_{ij} \xi_i',
\]

\[
\frac{\partial^2 L_{c,1}(\mathbf{D}; \beta, \mathbf{A})}{\partial \mathbf{A}_j \partial \mathbf{A}'_{\ell}} = -\delta_{j\ell} \psi_{jj}^{-1} \sum_{i=1}^{n} \xi_i' \xi_i',
\]

\[
\frac{\partial^2 L_{c,2}(\Omega; \Phi)}{\partial \phi_s \partial \phi_t} = -\frac{1}{2} tr \left( \Phi^{-1} \frac{\partial \Phi}{\partial \phi_s} \Phi^{-1} \left\{ \sum_{i=1}^{n} (2 \xi_i' \xi_i - \Phi) \right\} \Phi^{-1} \frac{\partial \Phi}{\partial \phi_t} \right).
\]

References