Sankhyā : The Indian Journal of Statistics
Special Issue on Biostatistics
2000, Volume 62, Series B, Pt. 1, pp. 5–24

A STEP-DOWN PROCEDURE FOR ANALYSIS OF TIME-VARYING COVARIATES IN MULTI-VISIT STUDIES

By SOLANGE ANDREONI*
Laboratórios Pfizer Ltda, São Paula

SUMMARY. In multi-visit studies where several subjects are followed for a certain prespecified period of time, the goal may be viewed as the description of patterns of change of the response variable(s) as a function of covariates, time, and or treatment. Often there are many covariates measured at each visit. In view of the cost involved, one question that may arise is whether it is necessary to have the collection of all covariates at each visit. The objective here is to study the necessity of using the covariates at each time point in the analysis, as opposed to using only the baseline measurements. Some hypotheses of interest are formulated in terms of the associated dispersion matrices adjusted for the covariates, and test statistics for comparisons between them based on the step-down procedure are presented. This study may help to design future investigations, or to stop data collection on covariates early in the study.

1. **Introduction**

It is a common practice in many multi-visit studies to collect data on response variables and potential covariates repeatedly over time for each subject. Traditionally at each visit t a set of covariates $\mathbf{X}_t$ and responses $\mathbf{Y}_t$ are observed for n subjects. Often precise recording of $\mathbf{X}_t$ may be costly and time consuming. In view of the issues of cost and precision involved, one question that may arise is whether the observation of the covariates at each visit provides more information than the observation of the covariates at the first visit alone.

One multivariate model that handles the problem of having the repeated covariates over time is the multiple design multivariate (MDM) model (Srivastava, 1966) which allows different dependent variables to have different design matrices.

AMS (1991) *subject classification.* Primary 62H05, 62H10, 62H15.
Key words and phrases. longitudinal data, repeated-measures, monotone multiple design multivariate model, likelihood ratio.

* Research supported in part by CAPES-Brazil while the author was a student at the University of North Carolina, Chapel Hill. This research does not necessarily reflect the views or practices of Pfizer Laboratories.

McDonald (1975) derived a hypotheses testing procedure for the MDM model with normally distributed errors. His method involves the application of the standard multivariate test statistics as long as the hypotheses are nested. The procedure assumes that the time-varying covariates are non-stochastic and does not take into account the fact that the model under the alternative hypothesis is restricted (nested).

A correlation model that allows a decreasing pattern of observation of both responses and covariates over time is considered in this work. It is assumed that the set of observations corresponding to a certain experimental unit or individual follows a multivariate normal distribution and the estimation process to obtain the maximum likelihood estimates of the parameters involved is described. The hypothesis of interest associated with the parameters of the dispersion matrix is decomposed into an intersection of a finite number of hypotheses. The use of the step-down procedure is advocated for testing the null hypotheses (Roy, Gnanadesikan and Srivastava, 1971).

The proposed methodology can help to design data collection schemes, and avoid the recording of other variables which might increase the costs of future investigations.

2. **Correlation Model**

Consider the situation where there are n subjects followed for p visits. At each visit t $(t = 1, \ldots, p)$ r random covariates and q responses are observed for n_t subjects, such that $n = n_1 \geq n_2 \geq \ldots \geq n_p$ and the n_t subjects have their covariates and responses measurements observed at the prior visits. At the beginning of the study f design variables that do not change over time are recorded for all individuals and are not considered to be random. Note that the r covariates change over time for each individual, and that n_p individuals have their covariate and response measurements observed at all p visits, $n_{p-1} - n_p$ individuals have their covariate and response measurements observed at the first $p - 1$ visits, and so on, such that $n_1 - n_2$ individuals have their covariate and response measurements observed only at the first visit. In order to apply the estimation procedure described later on, it is assumed that $n_p > f + 1 + p(r + q) - q$ and also that the drop-out mechanism is independent of the responses and covariates.

Let $\mathbf{x}_{i0\bullet}$ be the $((f + 1) \times 1)$ vector of fixed design covariates with first element being 1, $\mathbf{x}_{i1\bullet}$, $\mathbf{x}_{i2\bullet}$, ..., $\mathbf{x}_{it\bullet}$ be the $(r \times 1)$ vectors of random covariates, $\mathbf{y}_{i1\bullet}$, $\mathbf{y}_{i2\bullet}$, ..., $\mathbf{y}_{it\bullet}$ be the $(q \times 1)$ vectors of responses, and $\mathbf{e}_{i1\bullet}$, $\mathbf{e}_{i2\bullet}$, ..., $\mathbf{e}_{it\bullet}$ be the $(q \times 1)$ vectors of random errors associated with each individual i, for $i = 1, \ldots, n_t$ and $t = 1, \ldots, p$. Let

$$\mathbf{X}_0^{(t)} = \begin{bmatrix} 1 & x_{101} & \ldots & x_{10f} \\ 1 & x_{201} & \ldots & x_{20f} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n_t01} & \ldots & x_{n_t0f} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{10\bullet}^T \\ \mathbf{x}_{20\bullet}^T \\ \vdots \\ \mathbf{x}_{n_t0\bullet}^T \end{bmatrix} = [\ \mathbf{x}_{\bullet 01} \quad \ldots \quad \mathbf{x}_{\bullet 0f}\ ],$$

$$\mathbf{X}_u^{(t)} = \begin{bmatrix} x_{1u1} & \dots & x_{1ur} \\ x_{2u1} & \dots & x_{2ur} \\ \vdots & \ddots & \vdots \\ x_{n_t u1} & \dots & x_{n_t ur} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1u\bullet}^T \\ \mathbf{x}_{2u\bullet}^T \\ \vdots \\ \mathbf{x}_{n_t u\bullet}^T \end{bmatrix} = [\ \mathbf{x}_{\bullet u1} \quad \mathbf{x}_{\bullet u2} \quad \dots \quad \mathbf{x}_{\bullet ur}\ ],$$

$$\mathbf{Y}_u^{(t)} = \begin{bmatrix} y_{1u1} & \dots & y_{1uq} \\ y_{2u1} & \dots & y_{2uq} \\ \vdots & \ddots & \vdots \\ y_{n_t u1} & \dots & y_{n_t uq} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_{1u\bullet}^T \\ \mathbf{y}_{2u\bullet}^T \\ \vdots \\ \mathbf{y}_{n_t u\bullet}^T \end{bmatrix} = [\ \mathbf{y}_{\bullet u1} \quad \mathbf{y}_{\bullet u2} \quad \dots \quad \mathbf{y}_{\bullet uq}\ ],$$

$$\mathbf{E}_u^{(t)} = \begin{bmatrix} e_{1u1} & \dots & e_{1uq} \\ e_{2u1} & \dots & e_{2uq} \\ \vdots & \ddots & \vdots \\ e_{n_t u1} & \dots & e_{n_t uq} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{1u\bullet}^T \\ \mathbf{e}_{2u\bullet}^T \\ \vdots \\ \mathbf{e}_{n_t u\bullet}^T \end{bmatrix} = [\ \mathbf{e}_{\bullet u1} \quad \mathbf{e}_{\bullet u2} \quad \dots \quad \mathbf{e}_{\bullet uq}\ ],$$

$$\mathbf{X}^{(t)} = \left[\mathbf{X}_1^{(t)} \ \dots \ \mathbf{X}_t^{(t)}\right]$$

and

$$\mathbf{Y}^{(t)} = \left[\mathbf{Y}_1^{(t)} \ \dots \ \mathbf{Y}_{t-1}^{(t)}\right]$$

for $u = 1, \dots, t$ and $t = 1, \dots, p$. Suppose that $\mathbf{x}_{i\bullet\bullet}^{(t)} = (\mathbf{x}_{i1\bullet}^T \, \mathbf{x}_{i2\bullet}^T \dots \mathbf{x}_{it\bullet}^T)^T$ and $\mathbf{e}_{i\bullet\bullet}^{(t)} = (\mathbf{e}_{i1\bullet}^T \, \mathbf{e}_{i2\bullet}^T \dots \mathbf{e}_{it\bullet}^T)^T$ are uncorrelated random vectors distributed according to a multivariate normal distribution with mean

$$E\left[\begin{pmatrix} \mathbf{x}_{i\bullet\bullet}^{(t)} \\ \mathbf{e}_{i\bullet\bullet}^{(t)} \end{pmatrix}\right] = \begin{pmatrix} \mathbf{\Theta}_{x0}^{(t)T} \mathbf{x}_{i0\bullet} \\ \mathbf{0} \end{pmatrix}$$

and variance-covariance matrix

$$\mathbf{\Sigma}^{(t)} = \begin{pmatrix} \mathbf{\Sigma}_{xx.x_0}^{(t)} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{ee}^{(t)} \end{pmatrix}$$

where

$$\mathbf{\Theta}_{x0}^{(t)} = [\mathbf{\Theta}_{x10} \ \mathbf{\Theta}_{x20} \ \dots \ \mathbf{\Theta}_{xt0}],$$

$$\mathbf{\Sigma}_{xx.x_0}^{(t)} = \begin{pmatrix} \mathbf{\Sigma}_{x_1x_1.x_0} & \mathbf{\Sigma}_{x_1x_2.x_0} & \dots & \mathbf{\Sigma}_{x_1x_t.x_0} \\ \mathbf{\Sigma}_{x_2x_1.x_0} & \mathbf{\Sigma}_{x_2x_2.x_0} & \dots & \mathbf{\Sigma}_{x_2x_t.x_0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Sigma}_{x_tx_1.x_0} & \mathbf{\Sigma}_{x_tx_2.x_0} & \dots & \mathbf{\Sigma}_{x_tx_t.x_0} \end{pmatrix}$$

and

$$\mathbf{\Sigma}_{ee}^{(t)} = \begin{pmatrix} \mathbf{\Sigma}_{e_1e_1} & \mathbf{\Sigma}_{e_1e_2} & \dots & \mathbf{\Sigma}_{e_1e_t} \\ \mathbf{\Sigma}_{e_2e_1} & \mathbf{\Sigma}_{e_2e_2} & \dots & \mathbf{\Sigma}_{e_2e_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Sigma}_{e_te_1} & \mathbf{\Sigma}_{e_te_2} & \dots & \mathbf{\Sigma}_{e_te_t} \end{pmatrix}.$$

Note that $\mathbf{\Theta}_{x0}^{(t)}$ corresponds to the first $t(f+1)$ columns of $\mathbf{\Theta}_{x0}^{(p)} = \mathbf{\Theta}_{x0}$, $\mathbf{\Sigma}_{xx.x_0}^{(t)}$ corresponds to the first tr columns and rows of $\mathbf{\Sigma}_{xx.x_0}^{(p)} = \mathbf{\Sigma}_{xx.x_0}$, and $\mathbf{\Sigma}_{ee}^{(t)}$ corresponds to the first tq columns and rows of $\mathbf{\Sigma}_{ee}^{(p)} = \mathbf{\Sigma}_{ee}$. The fact that the sets of covariate observations and random errors associated with different individuals are uncorrelated implies that they are independent under the assumption of normality.

In this context the appropriate correlation model that incorporates a monotone MDM linear model for the vector of responses at the tth visit of the ith individual is given by

$$\mathbf{y}_{it\bullet} = \mathbf{B}_{t0}^{T}\mathbf{x}_{i0\bullet} + \mathbf{B}_{t}^{T}\mathbf{x}_{i\bullet\bullet}^{(t)} + \mathbf{e}_{it\bullet}, \tag{2.1}$$

where $\mathbf{B}_{t0}$ is the $((f+1)\times q)$ matrix of unknown regression coefficients associated with the fixed covariates, $\mathbf{B}_t = (\mathbf{B}_{t1}^T\,\mathbf{B}_{t2}^T\,\ldots\,\mathbf{B}_{tt}^T)^T$ is the $(tr\times q)$ matrix with $\mathbf{B}_{tu}$ being the $(r\times q)$ matrix of unknown regression coefficients associated with the random covariates observed at visit u for $u = 1,\ldots,t$, $i = 1,\ldots,n_t$, and $t = 1,\ldots,p$. Then $\mathbf{z}_{i\bullet\bullet}^{(t)} = (\mathbf{x}_{i\bullet\bullet}^{(t)T}\ \mathbf{y}_{i\bullet\bullet}^{(t)T})^T$ $\ i = 1,\ldots,n_t$ are normally distributed with mean

$$\boldsymbol{\mu}_{z_i}^{(t)} = \begin{pmatrix} \mathbf{\Theta}_{x0}^{(t)T}\mathbf{x}_{i0\bullet} \\ \mathbf{B}_0^{(t)T}\mathbf{x}_{i0\bullet} + \mathbf{B}^{(t)T}\mathbf{\Theta}_{x0}^{(t)T}\mathbf{x}_{i0\bullet} \end{pmatrix}$$

and variance-covariance matrix

$$\mathbf{\Sigma}_{zz}^{(t)} = \begin{pmatrix} \mathbf{\Sigma}_{xx.x_0}^{(t)} & \mathbf{\Sigma}_{xx.x_0}^{(t)}\mathbf{B}^{(t)} \\ \mathbf{B}^{(t)T}\mathbf{\Sigma}_{xx.x_0}^{(t)} & \mathbf{\Sigma}_{yy.x_0x}^{(t)} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma}_{xx.x_0}^{(t)} & \mathbf{\Sigma}_{xx.x_0}^{(t)}\mathbf{B}^{(t)} \\ \mathbf{B}^{(t)T}\mathbf{\Sigma}_{xx.x_0}^{(t)} & \mathbf{\Sigma}_{ee}^{(t)} + \mathbf{B}^{(t)T}\mathbf{\Sigma}_{xx.x_0}^{(t)}\mathbf{B}^{(t)} \end{pmatrix}$$

where

$$\mathbf{B}_0^{(t)} = (\mathbf{B}_{10}\,\mathbf{B}_{20}\ldots\mathbf{B}_{t0})$$

and

$$\mathbf{B}^{(t)} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{21} & \ldots & \mathbf{B}_{t1} \\ \mathbf{0} & \mathbf{B}_{22} & \ldots & \mathbf{B}_{t2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{B}_{tt} \end{pmatrix}.$$

Note that $\mathbf{B}_0^{(t)}$ corresponds to the matrix formed by the first $t(f+1)$ columns of $\mathbf{B}_0^{(p)} = \mathbf{B}_0$, $\mathbf{B}^{(t)}$ corresponds to the matrix formed by first tq columns and tr rows of $\mathbf{B}^{(p)} = \mathbf{B}$, and that $\mathbf{\Sigma}_{yy.x_0x}^{(t)}$ corresponds to the first tq columns and rows of $\mathbf{\Sigma}_{yy.x_0x}^{(p)} = \mathbf{\Sigma}_{yy.x_0x} = \mathbf{\Sigma}_{ee} + \mathbf{B}^T\mathbf{\Sigma}_{xx.x_0}\mathbf{B}$.

3. **Estimation**

The usual method of maximum likelihood involves the maximization of the joint distribution of $(\mathbf{x}_{i\bullet\bullet}^{(t)T}\ \mathbf{e}_{i\bullet\bullet}^{(t)T})^T$ for $i = n_{t+1}+1,\ldots,n_t$ and $t = 1,\ldots,p$ with respect to $\mathbf{B}$ and $\mathbf{\Sigma}$. Since $\mathbf{x}_{i\bullet\bullet}^{(t)}$ and $\mathbf{e}_{i\bullet\bullet}^{(t)}$ for $i = n_{t+1}+1,\ldots,n_t$ and $t = 1,\ldots,p$ are uncorrelated random vectors, thus independent under the normality assumption, the likelihood under the model (2.1) is given by

$$L(\mathbf{x}_{\bullet\bullet\bullet}, \mathbf{e}_{\bullet\bullet\bullet}, \mathbf{B}, \boldsymbol{\Sigma})$$
$$= (2\pi)^{-\frac{N}{2}} \prod_{t=1}^{p} |\boldsymbol{\Sigma}_{ee}^{(t)}|^{-\frac{(n_t-n_{t+1})}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{p} \mathrm{tr}\boldsymbol{\Sigma}_{ee}^{(t)-1} \sum_{i=n_{t+1}+1}^{n_t} \mathbf{e}_{i\bullet\bullet}^{(t)}\mathbf{e}_{i\bullet\bullet}^{(t)T}\right]$$
$$\prod_{t=1}^{p}|\boldsymbol{\Sigma}_{xx.x_0}^{(t)}|^{-\frac{(n_t-n_{t+1})}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{p} \mathrm{tr}\boldsymbol{\Sigma}_{xx.x_0}^{(t)-1} \sum_{i=n_{t+1}+1}^{n_t}(\mathbf{x}_{i\bullet\bullet}^{(t)}-\boldsymbol{\Theta}_{x0}^{(t)T}\mathbf{x}_{i0\bullet})(\mathbf{x}_{i\bullet\bullet}^{(t)}-\boldsymbol{\Theta}_{x0}^{(t)T}\mathbf{x}_{i0\bullet})^T\right] \tag{3.1}$$

where $n_{p+1} = 0$, $N = (q+r)\sum_{t=1}^{p} n_t$, and $\boldsymbol{\Sigma} = \mathrm{diag}(\mathbf{I}_{n_p} \otimes \boldsymbol{\Sigma}_{xx.x_0}^{(p)}, \mathbf{I}_{n_{p-1}-n_p} \otimes \boldsymbol{\Sigma}_{xx.x_0}^{(p-1)}, \ldots, \mathbf{I}_{n_1-n_2} \otimes \boldsymbol{\Sigma}_{xx.x_0}^{(1)}, \mathbf{I}_{n_p} \otimes \boldsymbol{\Sigma}_{ee}^{(p)}, \mathbf{I}_{n_{p-1}-n_p} \otimes \boldsymbol{\Sigma}_{ee}^{(p-1)}, \ldots, \mathbf{I}_{n_1-n_2} \otimes \boldsymbol{\Sigma}_{ee}^{(1)})$. Direct calculation of the maximum likelihood estimators of the parameters in (3.1) often requires iterative procedures. An alternative approach is to apply an orthogonal transformation on $\mathbf{x}_{\bullet\bullet\bullet}$ and $\mathbf{e}_{\bullet\bullet\bullet}$ (see Fujisawa (1995)) to obtain explicit solutions of the maximum likelihood estimators of functions of the parameters in $\mathbf{B}$ and $\boldsymbol{\Sigma}$ and then transform back to obtain the maximum likelihood estimators of the original parameters. For that purpose observe that if $\mathbf{x}_{i1\bullet}$, $\mathbf{y}_{i1\bullet}$, $\mathbf{x}_{i2\bullet}$, $\mathbf{y}_{i2\bullet}$, …, $\mathbf{x}_{ip\bullet}$, $\mathbf{y}_{ip\bullet}$ for each individual i are distributed according to a multivariate normal distribution, then the variables

$$\begin{array}{l}
\mathbf{x}_{i1\bullet} - \mathrm{E}(\mathbf{x}_{i1\bullet} \mid \mathbf{x}_{i0\bullet})\\
\mathbf{y}_{i1\bullet} - \mathrm{E}(\mathbf{y}_{i1\bullet} \mid \mathbf{x}_{i0\bullet}, \mathbf{x}_{i1\bullet})\\
\mathbf{x}_{i2\bullet} - \mathrm{E}(\mathbf{x}_{i2\bullet} \mid \mathbf{y}_{i1\bullet}, \mathbf{x}_{i0\bullet}, \mathbf{x}_{i1\bullet})\\
\mathbf{y}_{i2\bullet} - \mathrm{E}(\mathbf{y}_{i2\bullet} \mid \mathbf{y}_{i1\bullet}, \mathbf{x}_{i0\bullet}, \mathbf{x}_{i1\bullet}, \mathbf{x}_{i2\bullet})\\
\vdots\\
\mathbf{x}_{ip\bullet} - \mathrm{E}(\mathbf{x}_{ip\bullet} \mid \mathbf{y}_{i1\bullet}, \ldots, \mathbf{y}_{ip-1\bullet}, \mathbf{x}_{i0\bullet}, \mathbf{x}_{i1\bullet}, \ldots, \mathbf{x}_{ip-1\bullet})\\
\mathbf{y}_{ip\bullet} - \mathrm{E}(\mathbf{y}_{ip\bullet} \mid \mathbf{y}_{i1\bullet}, \ldots, \mathbf{y}_{ip-1\bullet}, \mathbf{x}_{i0\bullet}, \mathbf{x}_{i1\bullet}, \ldots, \mathbf{x}_{ip\bullet})
\end{array} \tag{3.2}$$

are noncorrelated and follow a multivariate normal distribution with mean $\mathbf{0}$ and variance-covariance matrix

$$\boldsymbol{\Sigma}^C = \mathrm{diag}(\boldsymbol{\Sigma}_{x_1x_1.x_0}, \boldsymbol{\Sigma}_{e_1e_1}, \ldots, \boldsymbol{\Sigma}_{x_px_p.x_0x_1\ldots x_{p-1}}, \boldsymbol{\Sigma}_{e_pe_p.e_1e_2\ldots e_{p-1}}).$$

The likelihood function based on the transformed variables (3.2) is given by

$$L_c =$$
$$(2\pi)^{-\frac{N}{2}}|\boldsymbol{\Sigma}_{x_1x_1.x_0}|^{-\frac{n_1}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Sigma}_{x_1x_1.x_0}^{-1}\sum_{i=1}^{n_1}(\mathbf{x}_{i1\bullet} - \boldsymbol{\Theta}_{x10}^T\mathbf{x}_{i0\bullet})(\mathbf{x}_{i1\bullet} - \boldsymbol{\Theta}_{x10}^T\mathbf{x}_{i0\bullet})^T\right]$$
$$\prod_{t=2}^{p}|\boldsymbol{\Sigma}_{x_tx_t.x_0x_1\ldots x_{t-1}}|^{-\frac{n_t}{2}} \exp\left\{-\frac{1}{2}\sum_{t=2}^{p}\mathrm{tr}\,\boldsymbol{\Sigma}_{x_tx_t.x_0x_1\ldots x_{t-1}}^{-1}\sum_{i=1}^{n_t}[\mathbf{x}_{it\bullet}-\boldsymbol{\Theta}_{xt0}^T\mathbf{x}_{i0\bullet}-\right.$$
$$\left.\boldsymbol{\Theta}_{xt}^T(\mathbf{x}_{i\bullet\bullet}^{(t-1)}-\boldsymbol{\Theta}_{x0}^{(t-1)T}\mathbf{x}_{i0\bullet})][\mathbf{x}_{it\bullet}-\boldsymbol{\Theta}_{xt0}^T\mathbf{x}_{i0\bullet}-\boldsymbol{\Theta}_{xt}^T(\mathbf{x}_{i\bullet\bullet}^{(t-1)}-\boldsymbol{\Theta}_{x0}^{(t-1)T}\mathbf{x}_{i0\bullet})]^T\right\}$$
$$|\boldsymbol{\Sigma}_{e_1e_1}|^{-\frac{n_1}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Sigma}_{e_1e_1}^{-1}\sum_{i=1}^{n_1}\mathbf{e}_{i1\bullet}\mathbf{e}_{i1\bullet}^T\right]\prod_{t=2}^{p}|\boldsymbol{\Sigma}_{e_te_t.e_1\ldots e_{t-1}}|^{-\frac{n_t}{2}}$$

$$\times \exp\left[-\frac{1}{2}\sum_{t=2}^{p}\mathrm{tr}\,\boldsymbol{\Sigma}_{e_te_t.e_1\ldots e_{t-1}}^{-1}\sum_{i=1}^{n_t}(\mathbf{e}_{it\bullet}-\boldsymbol{\Theta}_{et}^{T}\mathbf{e}_{i\bullet\bullet}^{(t-1)})(\mathbf{e}_{it\bullet}-\boldsymbol{\Theta}_{et}^{T}\mathbf{e}_{i\bullet\bullet}^{(t-1)})^{T}\right] \tag{3.3}$$

where $\mathbf{x}_{i\bullet\bullet}^{(t-1)} = (\mathbf{x}_{i1\bullet}^{T}\,\mathbf{x}_{i2\bullet}^{T}\ldots\mathbf{x}_{it-1\bullet}^{T})^{T}$, $\mathbf{e}_{i\bullet\bullet}^{(t-1)} = (\mathbf{e}_{i1\bullet}^{T}\,\mathbf{e}_{i2\bullet}^{T}\ldots\mathbf{e}_{it-1\bullet}^{T})^{T}$,

$$\begin{aligned}
\boldsymbol{\Theta}_{xt} &= \boldsymbol{\Sigma}_{xx.x_0}^{(t-1)-1}\boldsymbol{\Sigma}_{x_1\ldots x_{t-1}\;x_t.x_0} = \left[\boldsymbol{\Theta}_{xt1}^{T}\;\ldots\;\boldsymbol{\Theta}_{xtt-1}^{T}\right]^{T},\\
\boldsymbol{\Theta}_{et} &= \boldsymbol{\Sigma}_{ee}^{(t-1)-1}\boldsymbol{\Sigma}_{e_1\ldots e_{t-1}\;e_t} = \left[\boldsymbol{\Theta}_{et1}^{T}\;\ldots\;\boldsymbol{\Theta}_{ett-1}^{T}\right]^{T},\\
\boldsymbol{\Sigma}_{x_tx_t.x_0x_1\ldots x_{t-1}} &= \boldsymbol{\Sigma}_{x_tx_t.x_0}-\boldsymbol{\Sigma}_{x_t\;x_1\ldots x_{t-1}.x_0}\boldsymbol{\Sigma}_{xx.x_0}^{(t-1)-1}\boldsymbol{\Sigma}_{x_1\ldots x_{t-1}\;x_t.x_0},\\
\boldsymbol{\Sigma}_{e_te_t.e_1\ldots e_{t-1}} &= \boldsymbol{\Sigma}_{e_te_t}-\boldsymbol{\Sigma}_{e_t\;e_1\ldots e_{t-1}}\boldsymbol{\Sigma}_{ee}^{(t-1)-1}\boldsymbol{\Sigma}_{e_1\ldots e_{t-1}\;e_t},\\
\boldsymbol{\Sigma}_{x_t\;x_1\ldots x_{t-1}.x_0} &= [\boldsymbol{\Sigma}_{x_tx_1.x_0}\,\boldsymbol{\Sigma}_{x_tx_2.x_0}\ldots\boldsymbol{\Sigma}_{x_tx_{t-1}.x_0}]
\end{aligned} \tag{3.4}$$

and

$$\boldsymbol{\Sigma}_{e_t\;e_1\ldots e_{t-1}} = [\boldsymbol{\Sigma}_{e_te_1}\,\boldsymbol{\Sigma}_{e_te_2}\ldots\boldsymbol{\Sigma}_{e_te_{t-1}}]$$

for $t = 2,\ldots,p$. The parameters of each of the $2p$ components of the likelihood can then be estimated by maximizing $2p$ multivariate regression models separately. The maximum likelihood estimators of $\mathbf{B}$ and $\boldsymbol{\Sigma}$ of the original correlation model can be obtained by transformation.

To find the maximum likelihood estimators of the parameters involved consider the repeated use of the following two well known theorems.

THEOREM 3.1. *Let $\boldsymbol{\Phi}$ and $\mathbf{T}$ be symmetric positive definite matrices and $f(\boldsymbol{\Phi}) = |\boldsymbol{\Phi}|^{-n}\exp[-n\mathrm{tr}\,(\boldsymbol{\Phi}^{-1}\mathbf{T})]$. Then $f(\boldsymbol{\Phi})$ is maximized with respect to $\boldsymbol{\Phi}$ when $\boldsymbol{\Phi} = \mathbf{T}$.*

THEOREM 3.2. *Let $\boldsymbol{\Phi}$ be a $(u\times v)$ matrix, $\mathbf{T}_1$ be a $(v\times v)$ symmetric positive definite matrix, $\mathbf{T}_2$ be a $(u\times u)$ symmetric positive definite matrix, and and $f(\boldsymbol{\Phi}) = -n\mathrm{tr}\,(\mathbf{T}_1\boldsymbol{\Phi}^{T}\mathbf{T}_2\boldsymbol{\Phi})$. Then $f(\boldsymbol{\Phi})$ is maximized with respect to $\boldsymbol{\Phi}$ when $\boldsymbol{\Phi} = \mathbf{0}$.*

Consider the terms of the likelihood (3.3) that correspond to the distribution of $\mathbf{x}_{\bullet\bullet\bullet}$, that is,

$$\begin{aligned}
L_x &= (2\pi)^{-\frac{N_x}{2}}|\boldsymbol{\Sigma}_{x_1x_1.x_0}|^{-\frac{n_1}{2}}\prod_{t=2}^{p}|\boldsymbol{\Sigma}_{x_tx_t.x_0x_1\ldots x_{t-1}}|^{-\frac{n_t}{2}}\\
&\quad\exp\left[-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_{x_1x_1.x_0}^{-1}\sum_{i=1}^{n_1}(\mathbf{x}_{i1\bullet}-\boldsymbol{\Theta}_{x10}^{T}\mathbf{x}_{i0\bullet})(\mathbf{x}_{i1\bullet}-\boldsymbol{\Theta}_{x10}^{T}\mathbf{x}_{i0\bullet})^{T}\right]\\
&\quad\exp\Bigg[-\frac{1}{2}\sum_{t=2}^{p}\mathrm{tr}\,\boldsymbol{\Sigma}_{x_tx_t.x_0x_1\ldots x_{t-1}}^{-1}\\
&\quad\sum_{i=1}^{n_t}[\mathbf{x}_{it\bullet}-\boldsymbol{\Theta}_{xt0}^{T}\mathbf{x}_{i0\bullet}-\boldsymbol{\Theta}_{xt}^{T}(\mathbf{x}_{i\bullet\bullet}^{(t-1)}-\boldsymbol{\Theta}_{x0}^{(t-1)T}\mathbf{x}_{i0\bullet})]\\
&\qquad[\mathbf{x}_{it\bullet}-\boldsymbol{\Theta}_{xt0}^{T}\mathbf{x}_{i0\bullet}-\boldsymbol{\Theta}_{xt}^{T}(\mathbf{x}_{i\bullet\bullet}^{(t-1)}-\boldsymbol{\Theta}_{x0}^{(t-1)T}\mathbf{x}_{i0\bullet})]^{T}\Bigg]
\end{aligned}$$

where $N_x = r\sum_{t=1}^{p} n_t/2$.

First take the likelihood corresponding to the distribution of $\mathbf{x}_{\bullet 1\bullet}$

$$L_{x_1} = (2\pi)^{-\frac{rn_1}{2}}|\mathbf{\Sigma}_{x_1x_1.x_0}|^{-\frac{n_1}{2}}\exp\left[-\frac{1}{2}\mathrm{tr}\mathbf{\Sigma}^{-1}_{x_1x_1.x_0}\sum_{i=1}^{n_1}(\mathbf{x}_{i1\bullet}-\mathbf{\Theta}^T_{x10}\mathbf{x}_{i0\bullet})(\mathbf{x}_{i1\bullet}-\mathbf{\Theta}^T_{x10}\mathbf{x}_{i0\bullet})^T\right]. \tag{3.5}$$

Theorem 3.3. *For the likelihood function* (3.5) *the maximum likelihood estimators of* $\mathbf{\Theta}_{x10}$ *and* $\mathbf{\Sigma}_{x_1x_1.x_0}$ *are*

$$\hat{\mathbf{\Theta}}_{x10} = \left\{\sum_{i=1}^{n_1}\mathbf{x}_{i0\bullet}\mathbf{x}^T_{i0\bullet}\right\}^{-1}\left\{\sum_{i=1}^{n_1}\mathbf{x}_{i0\bullet}\mathbf{x}^T_{i1\bullet}\right\} = \left\{\mathbf{X}_0^{(1)T}\mathbf{X}_0^{(1)}\right\}^{-1}\left\{\mathbf{X}_0^{(1)T}\mathbf{X}_1^{(1)}\right\}$$

and

$$\begin{aligned}\hat{\mathbf{\Sigma}}_{x_1x_1.x_0} &= \frac{1}{n_1}\sum_{i=1}^{n_1}(\mathbf{x}_{i1\bullet}-\hat{\mathbf{\Theta}}^T_{x10}\mathbf{x}_{i0\bullet})(\mathbf{x}_{i1\bullet}-\hat{\mathbf{\Theta}}^T_{x10}\mathbf{x}_{i0\bullet})^T \\ &= \frac{1}{n_1}\left(\mathbf{X}_1^{(1)}-\mathbf{X}_0^{(1)}\hat{\mathbf{\Theta}}_{x10}\right)^T\left(\mathbf{X}_1^{(1)}-\mathbf{X}_0^{(1)}\hat{\mathbf{\Theta}}_{x10}\right).\end{aligned} \tag{3.6}$$

The proofs to Theorem 3.3 and the following theorems can be found in Andreoni (1996).

Now take the likelihood of each $\mathbf{x}_{\bullet t\bullet}$ regressed on $\mathbf{x}_{\bullet 1\bullet},\ldots,\mathbf{x}_{\bullet t-1\bullet}$ for $t=2,\ldots,p$, that is

$$\begin{aligned}L_{x_t.x_0x_1\ldots x_{t-1}} = {} & (2\pi)^{-\frac{rn_t}{2}}|\mathbf{\Sigma}_{x_tx_t.x_0x_1\ldots x_{t-1}}|^{-\frac{n_t}{2}}\exp\left\{-\frac{1}{2}\mathrm{tr}\,\mathbf{\Sigma}^{-1}_{x_tx_t.x_0x_1\ldots x_{t-1}}\right. \\ & \sum_{i=1}^{n_t}[\mathbf{x}_{it\bullet}-\mathbf{\Theta}^T_{xt0}\mathbf{x}_{i0\bullet}-\mathbf{\Theta}^T_{xt}(\mathbf{x}^{(t-1)}_{i\bullet\bullet}-\mathbf{\Theta}^{(t-1)T}_{x0}\mathbf{x}_{i0\bullet})] \\ & \left.[\mathbf{x}_{it\bullet}-\mathbf{\Theta}^T_{xt0}\mathbf{x}_{i0\bullet}-\mathbf{\Theta}^T_{xt}(\mathbf{x}^{(t-1)}_{i\bullet\bullet}-\mathbf{\Theta}^{(t-1)T}_{x0}\mathbf{x}_{i0\bullet})]^T\right\}.\end{aligned} \tag{3.7}$$

Define

$$[\tilde{\mathbf{\Theta}}^{(t-1)}_{x0}\ \ \tilde{\mathbf{\Theta}}_{xt0}] = \left(\mathbf{X}_0^{(t)T}\mathbf{X}_0^{(t)}\right)^{-1}\mathbf{X}_0^{(t)T}[\mathbf{X}^{(t)}\ \ \mathbf{X}_t^{(t)}] \tag{3.8}$$

and

$$\mathbf{Q}_{tt} = \frac{1}{n_t}[\mathbf{X}^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}^{(t-1)}_{x0}\ \ \mathbf{X}_t^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}_{xt0}]^T[\mathbf{X}^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}^{(t-1)}_{x0}\ \ \mathbf{X}_t^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}_{xt0}]. \tag{3.9}$$

with submatrices

$$\mathbf{Q}_{tt11} = \frac{1}{n_t}(\mathbf{X}^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}^{(t-1)}_{x0})^T(\mathbf{X}^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}^{(t-1)}_{x0}),$$

$$\mathbf{Q}_{tt22} = \frac{1}{n_t}(\mathbf{X}_t^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}_{xt0})^T(\mathbf{X}_t^{(t)}-\mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}_{xt0})$$

and

$$\mathbf{Q}_{tt12} = \mathbf{Q}^{T}_{tt21} = \frac{1}{n_t}(\mathbf{X}^{(t)} - \mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}_{x0}^{(t-1)})^{T}(\mathbf{X}_t^{(t)} - \mathbf{X}_0^{(t)}\tilde{\mathbf{\Theta}}_{xt0}).$$

THEOREM 3.4. *For the likelihood function* (3.7) *the maximum likelihood estimators of* $\mathbf{\Theta}_{xt0}$, $\mathbf{\Theta}_{xt}$ *and* $\mathbf{\Sigma}_{x_t x_t . x_0 x_1 \ldots x_{t-1}}$ *are given by*

$$\hat{\mathbf{\Theta}}_{xt0} = \tilde{\mathbf{\Theta}}_{xt0} - [\tilde{\mathbf{\Theta}}_{x0}^{(t-1)} - \hat{\mathbf{\Theta}}_{x0}^{(t-1)}]\hat{\mathbf{\Theta}}_{xt},$$

$$\hat{\mathbf{\Theta}}_{xt} = \mathbf{Q}_{tt11}^{-1}\mathbf{Q}_{tt12}$$

and

$$\hat{\mathbf{\Sigma}}_{x_t x_t . x_0 x_1 \ldots x_{t-1}} = \mathbf{Q}_{tt22} - \mathbf{Q}_{tt21}\mathbf{Q}_{tt11}^{-1}\mathbf{Q}_{tt12}.$$

where $\tilde{\mathbf{\Theta}}_{xt0}$ *and* $\tilde{\mathbf{\Theta}}_{x0}^{(t-1)}$ *are given by* (3.8) *and* $\mathbf{Q}_{tt}$ *by* (3.9).

The next algorithm explains how to obtain the maximum likelihood estimators of $\mathbf{\Theta}_{x0}$ and $\mathbf{\Sigma}_{xx.x_0}$. It is a direct application of Theorems 3.3 and 3.4 and the relationships given in (3.4).

ALGORITHM 3.1. *The maximum likelihood estimators of* $\mathbf{\Theta}_{x0}$ *and* $\mathbf{\Sigma}_{xx.x_0}$ *can be obtained using the p steps described below.*

Step 1 *Obtain* $\hat{\mathbf{\Sigma}}_{x_1x_1.x_0}$ *and* $\hat{\mathbf{\Theta}}_{x10}$. *Then*

$$\hat{\mathbf{\Sigma}}^{(1)}_{xx.x_0} = \hat{\mathbf{\Sigma}}_{x_1x_1.x_0}$$

and

$$\hat{\mathbf{\Theta}}^{(1)}_{x0} = \hat{\mathbf{\Theta}}_{x10}.$$

Step 2 *Obtain* $\hat{\mathbf{\Sigma}}_{x_2x_2.x_0x_1}$, $\hat{\mathbf{\Theta}}_{x20}$ *and* $\hat{\mathbf{\Theta}}_{x2}$. *Since* $\mathbf{\Theta}_{x2} = \mathbf{\Sigma}^{-1}_{x_1x_1.x_0}\mathbf{\Sigma}_{x_1x_2.x_0}$ *and* $\mathbf{\Sigma}_{x_2x_2.x_0x_1} = \mathbf{\Sigma}_{x_2x_2.x_0} - \mathbf{\Sigma}_{x_2x_1.x_0}\mathbf{\Sigma}^{(1)-1}_{xx.x_0}\mathbf{\Sigma}_{x_1x_2.x_0}$, *we have that*

$$\hat{\mathbf{\Sigma}}_{x_1x_2.x_0} = \hat{\mathbf{\Sigma}}^{(1)}_{xx.x_0}\hat{\mathbf{\Theta}}_{x2},$$

$$\begin{aligned}\hat{\mathbf{\Sigma}}_{x_2x_2.x_0} &= \hat{\mathbf{\Sigma}}_{x_2x_2.x_0x_1} + \hat{\mathbf{\Sigma}}_{x_2x_1.x_0}\hat{\mathbf{\Sigma}}^{(1)-1}_{xx.x_0}\hat{\mathbf{\Sigma}}_{x_1x_2.x_0} \\ &= \hat{\mathbf{\Sigma}}_{x_2x_2.x_0x_1} + \hat{\mathbf{\Sigma}}_{x_2x_1.x_0}\hat{\mathbf{\Theta}}_{x2},\end{aligned}$$

$$\hat{\mathbf{\Sigma}}^{(2)}_{xx.x_0} = \begin{pmatrix} \hat{\mathbf{\Sigma}}_{x_1x_1.x_0} & \hat{\mathbf{\Sigma}}_{x_1x_2.x_0} \\ \hat{\mathbf{\Sigma}}_{x_2x_1.x_0} & \hat{\mathbf{\Sigma}}_{x_2x_2.x_0} \end{pmatrix}$$

and

$$\hat{\mathbf{\Theta}}^{(2)}_{x0} = [\hat{\mathbf{\Theta}}_{x10}\ \hat{\mathbf{\Theta}}_{x20}].$$

And so on, so that at the last step we have the following.

Step *p* *Obtain* $\hat{\mathbf{\Sigma}}_{x_px_p.x_0x_1\ldots x_{p-1}}$, $\hat{\mathbf{\Theta}}_{xp0}$ *and* $\hat{\mathbf{\Theta}}_{xp}$. *Since* $\mathbf{\Theta}_{xp} = \mathbf{\Sigma}^{(p-1)-1}_{xx.x_0}\mathbf{\Sigma}_{x_1\ldots x_{p-1}\,x_p.x_0}$ *and* $\mathbf{\Sigma}_{x_px_p.x_0x_1\ldots x_{p-1}} = \mathbf{\Sigma}_{x_px_p.x_0} - \mathbf{\Sigma}_{x_p\,x_1\ldots x_{p-1}.x_0}\mathbf{\Sigma}^{(p-1)-1}_{xx.x_0}\mathbf{\Sigma}_{x_1\ldots x_{p-1}\,x_p.x_0}$ *we have that*

$$\hat{\mathbf{\Sigma}}_{x_1\ldots x_{p-1}\,x_p.x_0} = \hat{\mathbf{\Sigma}}^{(p-1)}_{xx.x_0}\hat{\mathbf{\Theta}}_{xt},$$

$$\begin{aligned}\hat{\mathbf{\Sigma}}_{x_px_p.x_0} &= \hat{\mathbf{\Sigma}}_{x_px_p.x_0x_1\ldots x_{p-1}} + \hat{\mathbf{\Sigma}}_{x_p\ x_1\ldots x_{p-1}.x_0}\hat{\mathbf{\Sigma}}^{(p-1)-1}_{xx.x_0}\hat{\mathbf{\Sigma}}_{x_1\ldots x_{p-1}\ x_p.x_0}\\ &= \hat{\mathbf{\Sigma}}_{x_px_p.x_0x_1\ldots x_{p-1}} + \hat{\mathbf{\Sigma}}_{x_p\ x_1\ldots x_{p-1}.x_0}\hat{\mathbf{\Theta}}_{xp},\end{aligned}$$

$$\hat{\mathbf{\Sigma}}_{xx.x_0} = \hat{\mathbf{\Sigma}}^{(p)}_{xx.x_0} = \begin{pmatrix} \hat{\mathbf{\Sigma}}_{x_1x_1.x_0} & \cdots & \hat{\mathbf{\Sigma}}_{x_1x_p.x_0}\\ \vdots & \ddots & \vdots \\ \hat{\mathbf{\Sigma}}_{x_px_1.x_0} & \cdots & \hat{\mathbf{\Sigma}}_{x_px_p.x_0}\end{pmatrix}$$

and

$$\hat{\mathbf{\Theta}}_{x0} = \hat{\mathbf{\Theta}}^{(p)}_{x0} = [\hat{\mathbf{\Theta}}_{x10}\ \hat{\mathbf{\Theta}}_{x20}\ \ldots\ \hat{\mathbf{\Theta}}_{xp0}].$$

Now consider the terms of the likelihood (3.3) that correspond to the distribution of $\mathbf{e}_{\bullet\bullet\bullet}$, that is,

$$\begin{aligned}L_e &= (2\pi)^{-\frac{N_e}{2}}|\mathbf{\Sigma}_{e_1e_1}|^{-\frac{n_1}{2}}\exp\left(-\frac{1}{2}\mathrm{tr}\,\mathbf{\Sigma}^{-1}_{e_1e_1}\sum_{i=1}^{n_1}\mathbf{e}_{i1\bullet}\mathbf{e}^T_{i1\bullet}\right)\prod_{t=2}^{p}|\mathbf{\Sigma}_{e_te_t.e_1\ldots e_{t-1}}|^{-\frac{n_t}{2}}\\ &\exp\left[-\frac{1}{2}\sum_{t=2}^{p}\mathrm{tr}\,\mathbf{\Sigma}^{-1}_{e_te_t.e_1\ldots e_{t-1}}\sum_{i=1}^{n_t}(\mathbf{e}_{it\bullet}-\mathbf{\Theta}^T_{et}\mathbf{e}^{(t-1)}_{i\bullet\bullet})(\mathbf{e}_{it\bullet}-\mathbf{\Theta}^T_{et}\mathbf{e}^{(t-1)}_{i\bullet\bullet})^T\right].\end{aligned}$$

where $N_e = q\sum_{t=1}^{p} n_t/2$. First take the likelihood corresponding to the distribution of $\mathbf{e}_{i1\bullet}$

$$L_{e_1} = (2\pi)^{-\frac{qn_1}{2}}|\mathbf{\Sigma}_{e_1e_1}|^{-\frac{n_1}{2}}\exp\left(-\frac{1}{2}\mathrm{tr}\,\mathbf{\Sigma}^{-1}_{e_1e_1}\sum_{i=1}^{n_1}\mathbf{e}_{i1\bullet}\mathbf{e}^T_{i1\bullet}\right). \tag{3.10}$$

THEOREM 3.5. *For the likelihood function* (3.10) *the maximum likelihood estimators of* $\mathbf{B}_{10}$, $\mathbf{B}_{11}$ *and* $\mathbf{\Sigma}_{e_1e_1}$ *are*

$$\begin{aligned}\begin{bmatrix}\hat{\mathbf{B}}_{10}\\ \hat{\mathbf{B}}_{11}\end{bmatrix} &= \left\{\sum_{i=1}^{n_1}\begin{bmatrix}\mathbf{x}_{i0\bullet}\\ \mathbf{x}_{i1\bullet}\end{bmatrix}[\mathbf{x}^T_{i0\bullet}\ \mathbf{x}^T_{i1\bullet}]\right\}^{-1}\left\{\sum_{i=1}^{n_1}\begin{bmatrix}\mathbf{x}_{i0\bullet}\\ \mathbf{x}_{i1\bullet}\end{bmatrix}\mathbf{y}^T_{i1\bullet}\right\}\\ &= \left\{[\mathbf{X}^{(1)}_0\ \ \mathbf{X}^{(1)}_1]^T[\mathbf{X}^{(1)}_0\ \ \mathbf{X}^{(1)}_1]\right\}^{-1}[\mathbf{X}^{(1)}_0\ \ \mathbf{X}^{(1)}_1]^T\mathbf{Y}^{(1)}_1\end{aligned}$$

and

$$\begin{aligned}\hat{\mathbf{\Sigma}}_{e_1e_1} &= \frac{1}{n_1}\sum_{i=1}^{n_1}[\mathbf{y}_{i1\bullet}-\hat{\mathbf{B}}^T_{10}\mathbf{x}_{i0\bullet}-\hat{\mathbf{B}}^T_{11}\mathbf{x}_{i1\bullet}][\mathbf{y}_{i1\bullet}-\hat{\mathbf{B}}^T_{10}\mathbf{x}_{i0\bullet}-\hat{\mathbf{B}}^T_{11}\mathbf{x}_{i1\bullet}]^T\\ &= \frac{1}{n_1}[\mathbf{Y}^{(1)}_1-\mathbf{X}^{(1)}_0\hat{\mathbf{B}}_{10}-\mathbf{X}^{(1)}_1\hat{\mathbf{B}}_{11}]^T[\mathbf{Y}^{(1)}_1-\mathbf{X}^{(1)}_0\hat{\mathbf{B}}_{10}-\mathbf{X}^{(1)}_1\hat{\mathbf{B}}_{11}].\end{aligned}$$

Now take the likelihood corresponding to the distribution of $\mathbf{e}_{\bullet t\bullet}$ regressed on $\mathbf{e}_{\bullet 1\bullet},\ldots,\mathbf{e}_{\bullet t-1\bullet}$ for each $t = 2,\ldots,p$, that is,

$$\begin{aligned}L_{e_te_t.e_1\ldots e_{t-1}} &= (2\pi)^{-\frac{qn_t}{2}}|\mathbf{\Sigma}_{e_te_t.e_1\ldots e_{t-1}}|^{-\frac{n_t}{2}}\\ exp &\left[-\frac{1}{2}\mathrm{tr}\,\mathbf{\Sigma}^{-1}_{e_te_t.e_1\ldots e_{t-1}}\sum_{i=1}^{n_t}(\mathbf{e}_{it\bullet}-\mathbf{\Theta}^T_{et}\mathbf{e}^{(t-1)}_{i\bullet\bullet})(\mathbf{e}_{it\bullet}-\mathbf{\Theta}^T_{et}\mathbf{e}^{(t-1)}_{i\bullet\bullet})^T\right].\end{aligned} \tag{3.11}$$

Observe that

$$\begin{aligned}
\mathbf{e}_{it\bullet} - \mathbf{\Theta}_{et}^{T}\mathbf{e}_{i\bullet\bullet}^{(t-1)} &= \mathbf{y}_{it\bullet} - \mathbf{B}_{t0}^{T}\mathbf{x}_{i0\bullet} - \mathbf{B}_{t}^{T}\mathbf{x}_{i\bullet\bullet}^{(t-1)} - \mathbf{\Theta}_{et}^{T}\mathbf{e}_{i\bullet\bullet}^{(t-1)} \\
&= \mathbf{y}_{it\bullet} - \mathbf{B}_{t0}^{T}\mathbf{x}_{i0\bullet} - \sum_{u=1}^{t}\mathbf{B}_{tu}^{T}\mathbf{x}_{iu\bullet} - \sum_{u=1}^{t-1}\mathbf{\Theta}_{etu}^{T}\mathbf{e}_{iu\bullet} \\
&= \mathbf{y}_{it\bullet} - \mathbf{B}_{t0}^{T}\mathbf{x}_{i0\bullet} - \sum_{u=1}^{t}\mathbf{B}_{tu}^{T}\mathbf{x}_{iu\bullet} \\
&\quad - \sum_{u=1}^{t-1}\mathbf{\Theta}_{etu}^{T}\left[\mathbf{y}_{iu\bullet} - \mathbf{B}_{u0}^{T}\mathbf{x}_{i0\bullet} - \sum_{v=1}^{u}\mathbf{B}_{uv}^{T}\mathbf{x}_{iv\bullet}\right] \\
&= \mathbf{y}_{it\bullet} - \left(\mathbf{B}_{t0}^{T} - \sum_{v=1}^{t-1}\mathbf{\Theta}_{etv}^{T}\mathbf{B}_{v0}^{T}\right)\mathbf{x}_{i0\bullet} \\
&\quad - \sum_{u=1}^{t-1}\left(\mathbf{B}_{tu}^{T} - \sum_{v=u}^{t-1}\mathbf{\Theta}_{etv}^{T}\mathbf{B}_{vu}^{T}\right)\mathbf{x}_{iu\bullet} \\
&\quad - \mathbf{B}_{tt}^{T}\mathbf{x}_{it\bullet} \\
&\quad - \sum_{u=1}^{t-1}\mathbf{\Theta}_{etu}^{T}\mathbf{y}_{iu\bullet} \\
&= \mathbf{y}_{it\bullet} - \mathbf{\Gamma}_{t0}^{T}\mathbf{x}_{i0\bullet} - \sum_{u=1}^{t}\mathbf{\Gamma}_{tu}^{T}\mathbf{x}_{iu\bullet} - \sum_{u=1}^{t-1}\mathbf{\Theta}_{etu}^{T}\mathbf{y}_{iu\bullet} \\
&= \mathbf{y}_{it\bullet} - \hat{\mathbf{\Gamma}}_{t0}^{T}\mathbf{x}_{i0\bullet} - \hat{\mathbf{\Gamma}}_{t}^{T}\mathbf{x}_{i\bullet\bullet}^{(t-1)} - \hat{\mathbf{\Theta}}_{et}^{T}\mathbf{y}_{i\bullet\bullet}^{(t-1)} \\
&\quad + [(\hat{\mathbf{\Gamma}}_{t0}^{T} - \mathbf{\Gamma}_{t0}^{T})\ (\hat{\mathbf{\Gamma}}_{t}^{T} - \mathbf{\Gamma}_{t}^{T})\ (\hat{\mathbf{\Theta}}_{et}^{T} - \mathbf{\Theta}_{et}^{T})]\begin{bmatrix}\mathbf{x}_{i0\bullet}\\ \mathbf{x}_{i\bullet\bullet}^{(t)}\\ \mathbf{y}_{i\bullet\bullet}^{(t-1)}\end{bmatrix}
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{\Gamma}_{t0} &= \mathbf{B}_{t0} - \sum_{v=1}^{t-1}\mathbf{B}_{v0}\mathbf{\Theta}_{etv}, \\
\mathbf{\Gamma}_{tu} &= \mathbf{B}_{tu} - \sum_{v=u}^{t-1}\mathbf{B}_{vu}\mathbf{\Theta}_{etv}, \\
\mathbf{\Gamma}_{tt} &= \mathbf{B}_{tt}, \\
\mathbf{\Gamma}_{t} &= [\mathbf{\Gamma}_{t1}^{T}\ \ \mathbf{\Gamma}_{t2}^{T}\ \ \ldots\ \ \mathbf{\Gamma}_{tt}^{T}]^{T}
\end{aligned} \tag{3.12}$$

and

$$\begin{aligned}
\begin{bmatrix}\hat{\mathbf{\Gamma}}_{t0}\\ \hat{\mathbf{\Gamma}}_{t}\\ \hat{\mathbf{\Theta}}_{et}\end{bmatrix} &= \left\{\sum_{i=1}^{n_t}\begin{bmatrix}\mathbf{x}_{i0\bullet}\\ \mathbf{x}_{i\bullet\bullet}^{(t)}\\ \mathbf{y}_{i\bullet\bullet}^{(t-1)}\end{bmatrix}[\mathbf{x}_{i0\bullet}^{T}\ \mathbf{x}_{i\bullet\bullet}^{(t)T}\ \mathbf{y}_{i\bullet\bullet}^{(t-1)T}]\right\}^{-1}\left\{\sum_{i=1}^{n_t}\begin{bmatrix}\mathbf{x}_{i0\bullet}\\ \mathbf{x}_{i\bullet\bullet}^{(t)}\\ \mathbf{y}_{i\bullet\bullet}^{(t-1)}\end{bmatrix}\mathbf{y}_{it\bullet}^{T}\right\} \\
&= \left\{[\mathbf{X}_{0}^{(t)}\ \mathbf{X}^{(t)}\ \mathbf{Y}^{(t)}]^{T}[\mathbf{X}_{0}^{(t)}\ \mathbf{X}^{(t)}\ \mathbf{Y}^{(t)}]\right\}^{-1}[\mathbf{X}_{0}^{(t)}\ \mathbf{X}^{(t)}\ \mathbf{Y}^{(t)}]^{T}\mathbf{Y}_{t}^{(t)}
\end{aligned}$$

for $t = 2, \ldots, p$ and $u = 1, \ldots, t-1$.

THEOREM 3.6. *For the likelihood function* (3.11) *the maximum likelihood estimators of* $\mathbf{\Gamma}_{t0}$, $\mathbf{\Gamma}_t$, $\mathbf{\Theta}_{et}$ *and* $\mathbf{\Sigma}_{e_te_t.e_1\ldots e_{t-1}}$ *are given by*

$$\begin{bmatrix} \hat{\mathbf{\Gamma}}_{t0} \\ \hat{\mathbf{\Gamma}}_t \\ \hat{\mathbf{\Theta}}_{et} \end{bmatrix} = \left\{ \sum_{i=1}^{n_t} \begin{bmatrix} \mathbf{x}_{i0\bullet} \\ \mathbf{x}_{i\bullet\bullet}^{(t)} \\ \mathbf{y}_{i\bullet\bullet}^{(t-1)} \end{bmatrix} [\mathbf{x}_{i0\bullet}^T \ \mathbf{x}_{i\bullet\bullet}^{(t)T} \ \mathbf{y}_{i\bullet\bullet}^{(t-1)T}] \right\}^{-1} \left\{ \sum_{i=1}^{n_t} \begin{bmatrix} \mathbf{x}_{i0\bullet} \\ \mathbf{x}_{i\bullet\bullet}^{(t)} \\ \mathbf{y}_{i\bullet\bullet}^{(t-1)} \end{bmatrix} \mathbf{y}_{it\bullet}^T \right\}$$
$$= \left\{ [\mathbf{X}_0^{(t)} \ \mathbf{X}^{(t)} \ \mathbf{Y}^{(t)}]^T [\mathbf{X}_0^{(t)} \ \mathbf{X}^{(t)} \ \mathbf{Y}^{(t)}] \right\}^{-1} [\mathbf{X}_0^{(t)} \ \mathbf{X}^{(t)} \ \mathbf{Y}^{(t)}]^T \mathbf{Y}_t^{(t)} \tag{3.13}$$

and

$$\hat{\mathbf{\Sigma}}_{e_te_t.e_1\ldots e_{t-1}} = \frac{1}{n_t} \sum_{i=1}^{n_t} [\mathbf{y}_{it\bullet} - \hat{\mathbf{\Gamma}}_{t0}^T \mathbf{x}_{i0\bullet} - \hat{\mathbf{\Gamma}}_t^T \mathbf{x}_{i\bullet\bullet}^{(t)} - \hat{\mathbf{\Theta}}_{et}^T \mathbf{y}_{i\bullet\bullet}^{(t-1)}][\mathbf{y}_{it\bullet} - \hat{\mathbf{\Gamma}}_{t0}^T \mathbf{x}_{i0\bullet} - \hat{\mathbf{\Gamma}}_t^T \mathbf{x}_{i\bullet\bullet}^{(t)} - \hat{\mathbf{\Theta}}_{et}^T \mathbf{y}_{i\bullet\bullet}^{(t-1)}]^T$$
$$= \frac{1}{n_t} [\mathbf{Y}_t^{(t)} - \mathbf{X}_0^{(t)} \hat{\mathbf{\Gamma}}_{t0} - \mathbf{X}^{(t)} \hat{\mathbf{\Gamma}}_t - \mathbf{Y}^{(t)} \hat{\mathbf{\Theta}}_{et}]^T [\mathbf{Y}_t^{(t)} - \mathbf{X}_0^{(t)} \hat{\mathbf{\Gamma}}_{t0} - \mathbf{X}^{(t)} \hat{\mathbf{\Gamma}}_t - \mathbf{Y}^{(t)} \hat{\mathbf{\Theta}}_{et}]. \tag{3.14}$$

The next algorithm summarizes how to obtain the maximum likelihood estimators of $\mathbf{B}_{t0}$, $\mathbf{B}_t$ and $\mathbf{\Sigma}_{e_te_t}$ for $t = 1, \ldots, p$. It is a direct application of Theorems 3.1, 3.5 and 3.6, and the relationships (3.4) and (3.12).

ALGORITHM 3.2. *The maximum likelihood estimators of* $\mathbf{B}_{t0}$, $\mathbf{B}_t$ *and* $\mathbf{\Sigma}_{e_te_t}$ *for* $t = 1, \ldots, p$ *can be obtained using the* p *steps described below.*

Step 1 *Obtain* $\hat{\mathbf{\Sigma}}_{e_1e_1}$, $\hat{\mathbf{B}}_{10}$ *and* $\hat{\mathbf{B}}_1$. *Then*

$$\hat{\mathbf{\Sigma}}_{ee}^{(1)} = \hat{\mathbf{\Sigma}}_{e_1e_1},$$

and

$$\hat{\mathbf{B}}_{11} = \hat{\mathbf{\Gamma}}_{11} = \hat{\mathbf{B}}_1.$$

Step t ($t = 2, \ldots, p$) *Obtain* $\hat{\mathbf{\Sigma}}_{e_te_t.e_1\ldots e_{t-1}}$, $\hat{\mathbf{\Gamma}}_{t0}$, $\hat{\mathbf{\Gamma}}_t$ *and* $\hat{\mathbf{\Theta}}_{et}$. *Since* $\mathbf{\Theta}_{et} = \mathbf{\Sigma}_{ee}^{(t-1)-1} \mathbf{\Sigma}_{e_1\ldots e_{t-1}\,e_t}$ *and* $\mathbf{\Sigma}_{e_te_t.e_1\ldots e_{t-1}} = \mathbf{\Sigma}_{e_te_t} - \mathbf{\Sigma}_{e_t\,e_1\ldots e_{t-1}} \mathbf{\Sigma}_{ee}^{(t-1)-1} \mathbf{\Sigma}_{e_1\ldots e_{t-1}\,e_t}$ *we have that*

$$\hat{\mathbf{\Sigma}}_{e_1\ldots e_{t-1}\,e_t} = \hat{\mathbf{\Sigma}}_{ee}^{(t-1)-1} \hat{\mathbf{\Theta}}_{et},$$

$$\hat{\mathbf{\Sigma}}_{e_te_t} = \hat{\mathbf{\Sigma}}_{e_te_t.e_1\ldots e_{t-1}} + \hat{\mathbf{\Sigma}}_{e_t\,e_1\ldots e_{t-1}} \hat{\mathbf{\Sigma}}_{ee}^{(t-1)-1} \hat{\mathbf{\Sigma}}_{e_1\ldots e_{t-1}\,e_t}$$
$$= \hat{\mathbf{\Sigma}}_{e_te_t.e_1\ldots e_{t-1}} + \hat{\mathbf{\Sigma}}_{e_t\,e_1\ldots e_{t-1}} \hat{\mathbf{\Theta}}_{et},$$

and

$$\hat{\mathbf{\Sigma}}_{ee}^{(t)} = \begin{pmatrix} \hat{\mathbf{\Sigma}}_{e_1e_1} & \ldots & \hat{\mathbf{\Sigma}}_{e_1e_t} \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{\Sigma}}_{e_te_1} & \ldots & \hat{\mathbf{\Sigma}}_{e_te_t} \end{pmatrix}.$$

Then calculate

$$\hat{\mathbf{B}}_{tu} = \hat{\mathbf{\Gamma}}_{tu} + \sum_{v=u}^{t-1} \hat{\mathbf{B}}_{vu} \hat{\mathbf{\Theta}}_{tv}, \quad u = 1, \ldots, t-1,$$

$$\hat{\mathbf{B}}_{tt} = \hat{\mathbf{\Gamma}}_{tt}$$

and

$$\hat{\mathbf{B}}_{t0} = \hat{\mathbf{\Gamma}}_{t0} + \sum_{v=1}^{t-1} \hat{\mathbf{B}}_{v0} \hat{\mathbf{\Theta}}_{tv}.$$

4. **Hypotheses of Interest**

The objective is to test whether the inclusion of the time-varying covariates in the model provides a significant reduction in the dispersion matrix of the residual error of the responses when compared to the inclusion of the fixed covariates and the first visit measurements of the time-varying covariates. For that purpose call

$$\mathbf{\Sigma}^1_{e_t e_t} = \mathbf{\Sigma}_{y_t y_t . x_0 x_1 \ldots x_t},$$

$$\mathbf{\Sigma}^0_{e_t e_t} = \mathbf{\Sigma}_{y_t y_t . x_0 x_1},$$

$$\mathbf{\Sigma}^1_{e_t e_t . e_1 \ldots e_{t-1}} = \mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t}$$

and

$$\mathbf{\Sigma}^0_{e_t e_t . e_1 \ldots e_{t-1}} = \mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1}$$

where $y_t \,.\, y_1 \,\ldots\, y_{t-1}\, x_0\, x_1 \,\ldots\, x_t = y_t - E(y_t \mid y_1, \ldots, y_{t-1}, x_0, x_1 \ldots x_t)$ for $t = 1, \ldots, p$. The hypothesis being tested can be expressed as

$$H_0 : \bigcap_{t=2}^{p} \{H_{0t} :\ \mathbf{\Sigma}_{y_t y_t . x_0 x_1 \ldots x_t} = \mathbf{\Sigma}_{y_t y_t . x_0 x_1}\} \tag{4.1}$$

versus

$$H_1 : \bigcup_{t=2}^{p} \{H_{1t} :\ \mathbf{\Sigma}_{y_t y_t . x_0 x_1 \ldots x_t} < \mathbf{\Sigma}_{y_t y_t . x_0 x_1}\}. \tag{4.2}$$

We can rewrite (4.1) and (4.2) as

$$H^*_0 : \bigcap_{t=2}^{p} \{H^*_{0t} : \mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t} = \mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1} \text{ given } H^*_{0s} \text{ is true for } s = 1, \ldots, t-1\}$$

and

$$H^*_1 : \bigcup_{t=2}^{p} \{H^*_{1t} :\ \mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t} < \mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1}\}.$$

Observe that $H^*_{0t} = H_{0t}$ for $t = 1, \ldots, p$ if and only if H^*_{0s} is true for $s = 1, \ldots, t-1$. Observe that $\mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1} - \mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t}$ is positive semi-definite, since each element $\sigma_{y_{tj} y_{tj} . y_1 \ldots y_{t-1} x_0 x_1}$ in the diagonal of $\mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1}$ is greater than or equal to the corresponding element $\sigma_{y_{tj} y_{tj} . y_1 \ldots y_{p-1} x_0 x_1 \ldots x_p}$ in $\mathbf{\Sigma}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t}$, for $j = 1, \ldots, q$ and $t = 1, \ldots, p$; and if $\sigma_{y_{tj} y_{tj} . y_1 \ldots y_{t-1} x_0 x_1} = \sigma_{y_{tj} y_{tj} . y_1 \ldots y_{p-1} x_0 x_1 \ldots x_p}$

for some $j = 1, \ldots, q$, then $\sigma_{y_{tj}y_{tj'}.y_1\ldots y_{t-1}x_0x_1} - \sigma_{y_{tj}y_{tj'}.y_1\ldots y_{p-1}x_0x_1\ldots x_p} = \sigma_{y_{tj'}y_{tj}.y_1\ldots y_{t-1}x_0x_1} - \sigma_{y_{tj'}y_{tj}.y_1\ldots y_{p-1}x_0x_1\ldots x_p} = 0$ for $j' = 1, \ldots, q$.

The problem can be compared to the problem of testing independence between two sets of variables of the form

$$\begin{aligned} \mathbf{Z}_{t1}^{(t)} &= \mathbf{Y}_t^{(t)} \,.\, \mathbf{Y}_1^{(t)} \ldots \mathbf{Y}_{t-1}^{(t)}\, \mathbf{X}_0^{(t)}\, \mathbf{X}_1^{(t)} \\ &= \mathbf{Y}_t^{(t)} - E\left(\mathbf{Y}_t^{(t)} \mid \mathbf{Y}_1^{(t)}, \ldots, \mathbf{Y}_{t-1}^{(t)}, \mathbf{X}_0^{(t)}, \mathbf{X}_1^{(t)}\right) \end{aligned}$$

and

$$\begin{aligned} \mathbf{Z}_{t2}^{(t)} &= \left[\mathbf{X}_2^{(t)} \ldots \mathbf{X}_t^{(t)}\right] . \mathbf{Y}_1^{(t)} \ldots \mathbf{Y}_{t-1}^{(t)}\, \mathbf{X}_0^{(t)}\, \mathbf{X}_1^{(t)} \\ &= \left[\mathbf{X}_2^{(t)} \ldots \mathbf{X}_t^{(t)}\right] - E\left\{\left[\mathbf{X}_2^{(t)} \ldots \mathbf{X}_t^{(t)}\right] \mid \mathbf{Y}_1^{(t)}, \ldots, \mathbf{Y}_{t-1}^{(t)}, \mathbf{X}_0^{(t)}, \mathbf{X}_1^{(t)}\right\} \end{aligned}$$

satisfying the model

$$\mathbf{Z}_t^{(t)} = \left[\mathbf{Z}_{t1}^{(t)}\ \mathbf{Z}_{t2}^{(t)}\right] \sim N_{n_t m_t}(\mathbf{0}, \mathbf{I}_{n_t} \otimes \boldsymbol{\Sigma}_{z_t z_t})$$

with

$$\boldsymbol{\Sigma}_{z_t z_t} = \begin{pmatrix} \boldsymbol{\Sigma}_{z_{t1} z_{t1}} & \boldsymbol{\Sigma}_{z_{t1} z_{t2}} \\ \boldsymbol{\Sigma}_{z_{t2} z_{t1}} & \boldsymbol{\Sigma}_{z_{t2} z_{t2}} \end{pmatrix},$$

$\boldsymbol{\Sigma}_{z_t z_t}$ is $(q+r(t-1) \times q+r(t-1))$, $\boldsymbol{\Sigma}_{z_{t1} z_{t1}}$ is $(q \times q)$ and $\boldsymbol{\Sigma}_{z_{t2} z_{t2}}$ is $(r(t-1) \times r(t-1))$, that is, to test

$$H_{0t} : \boldsymbol{\Sigma}_{z_{t1} z_{t2}} = \mathbf{0} \quad \text{versus} \quad H_{1t} : \boldsymbol{\Sigma}_{z_{t1} z_{t2}} \neq \mathbf{0} \quad \text{for } t = 2, \ldots, p. \tag{4.3}$$

Let

$$\begin{aligned} \mathbf{G}_t &= n_t\, \hat{\boldsymbol{\Sigma}}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t} \\ &= \mathbf{Y}_t^{(t)T} [\mathbf{I}_{n_t} - \mathbf{X}_t^{*(t)} (\mathbf{X}_t^{*(t)T} \mathbf{X}_t^{*(t)})^{-1} \mathbf{X}_t^{*(t)T}] \mathbf{Y}_t^{(t)} \end{aligned}$$

and $\mathbf{H}_t = n_t\, \hat{\boldsymbol{\Sigma}}_{y_t\, x_2 \ldots x_t . y_1 \ldots y_{t-1} x_0 x_1} \hat{\boldsymbol{\Sigma}}^{-1}_{x_2 \ldots x_t\, x_2 \ldots x_t . y_1 \ldots y_{t-1} x_0 x_1} \hat{\boldsymbol{\Sigma}}_{x_2 \ldots x_t\, y_t . y_1 \ldots y_{t-1} x_0 x_1} =$

$$\mathbf{Y}_t^{(t)T} [\mathbf{X}_t^{*(t)} (\mathbf{X}_t^{*(t)T} \mathbf{X}_t^{*(t)})^{-1} \mathbf{C}_t^T [\mathbf{C}_t (\mathbf{X}_t^{*(t)T} \mathbf{X}_t^{*(t)})^{-1} \mathbf{C}_t^T]^{-1} \mathbf{C}_t (\mathbf{X}_t^{*(t)T} \mathbf{X}_t^{*(t)})^{-1} \mathbf{X}_t^{*(t)}] \mathbf{Y}_t^{(t)}$$

where

$$\mathbf{X}_t^{*(t)} = [\mathbf{X}_0^{(t)}\ \mathbf{X}_1^{(t)}\ \mathbf{X}_2^{(t)}\ \ldots\ \mathbf{X}_t^{(t)}\ \mathbf{Y}_1^{(t)}\ \ldots\ \mathbf{Y}_{t-1}^{(t)}]$$

and

$$\mathbf{C}_t = [\mathbf{0}_{r(t-1) \times (f+1+r)}\ \ \mathbf{I}_{r(t-1)}\ \ \mathbf{0}_{r(t-1) \times q(t-1)}].$$

The maximal invariant for each step is the set of ordered nonzero characteristic roots of $W_t = \mathbf{S}_{t11}^{-1} \mathbf{S}_{t12} \mathbf{S}_{t22}^{-1} \mathbf{S}_{t21}$ where $\mathbf{S}_{tij} = \mathbf{Z}_{ti}^T \mathbf{Z}_{tj}$ $(i, j = 1, 2)$.

Thus the hypotheses are now expressed in terms of canonical correlations

$$\mathbf{P}_t = \begin{pmatrix} \mathbf{I}_{m_{t1}} & \mathbf{R}_t \\ \mathbf{R}_t^T & \mathbf{I}_{m_{t2}} \end{pmatrix}, \quad \mathbf{R}_t = [\boldsymbol{\Delta}_t\ \mathbf{0}] : q \times r(t-1)$$

and $\mathbf{\Delta}_t = \text{diag}\ (\rho_{t1}, \ldots, \rho_{ts})$ with $s = \min[q, r(t-1)]$. In terms of Δ_t the hypothesis in (4.3) are written as

$$H_{0t}:\ \mathbf{\Delta}_t = \mathbf{0} \quad \text{versus} \quad H_{1t}:\ \mathbf{\Delta}_t \neq \mathbf{0} \quad \text{for } t = 2, \ldots, p.$$

At each step t $(t = 2, \ldots, p)$ the problem is invariant under the group of transformations $\mathcal{G}_t = \mathcal{O}(n_t) \times \text{GL}(q) \times \text{GL}[r(t-1)]$ acting on $\mathbf{Z}_t^{(t)}$ by

$$g_t(\mathbf{Z}_t^{(t)}) = (\mathbf{Q}_t \mathbf{Z}_{t1}^{(t)} \mathbf{C}_{t1}^{T}, \mathbf{Q}_t \mathbf{Z}_{t2}^{(t)} \mathbf{C}_{t2}^{T}) \quad \text{for } g_t = (\mathbf{Q}_t, \mathbf{C}_{t1}, \mathbf{C}_{t2}) \in \mathcal{G}_t$$

where $\mathcal{O}(n_t)$ is the set of all $(n_t \times n_t)$ orthogonal matrices, $\text{GL}(q)$ is the set of all $(q \times q)$ nonsingular matrices and $\text{GL}[r(t-1)]$ is the set of all $(r(t-1) \times r(t-1))$ nonsingular matrices.

Observe that the problem at each step remains invariant under the group of transformations $\mathcal{G}_t$ because

$$U_t^* = U_t^*[g_t(\mathbf{Z}_t^{(t)})] = \frac{|\mathbf{G}_t^*|}{|\mathbf{G}_t^* + \mathbf{H}_t^*|} = \frac{|\mathbf{G}_t|}{|\mathbf{G}_t + \mathbf{H}_t|} = U_t.$$

5. Step-Down Test Procedure

Let Λ be the likelihood ratio statistic for testing H_0 above when the observations of each individual come from a multivariate normal distribution. Observe that the likelihood ratio statistic can be written in several different ways. The following form is more convenient to evaluate the distribution under H_0

$$\begin{aligned}
\Lambda &= \frac{\max L^0(\mathbf{\Sigma})}{\max L^1(\mathbf{\Sigma})} \\
&= \prod_{t=2}^{p} \Lambda_t \\
&= \prod_{t=2}^{p} \left[\frac{|\hat{\mathbf{\Sigma}}^1_{e_t e_t . e_1 \ldots e_{t-1}}|}{|\hat{\mathbf{\Sigma}}^0_{e_t e_t . e_1 \ldots e_{t-1}}|} \right]^{n_t/2} \\
&= \prod_{t=2}^{p} \left[\frac{|\hat{\mathbf{\Sigma}}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t}|}{|\hat{\mathbf{\Sigma}}_{y_t y_t . y_1 \ldots y_{t-1} x_0 x_1}|} \right]^{n_t/2} \\
&= \prod_{t=2}^{p} \left[\frac{|\mathbf{G}_t|}{|\mathbf{G}_t + \mathbf{H}_t|} \right]^{n_t/2}, \qquad (5.1)
\end{aligned}$$

and the next one for the distribution under H_1

$$\Lambda = \prod_{t=2}^{p} \left[\frac{|\hat{\mathbf{\Sigma}}_{z_t z_t}|}{|\hat{\mathbf{\Sigma}}_{z_{t1} z_{t1}}||\hat{\mathbf{\Sigma}}_{z_{t2} z_{t2}}|} \right]^{n_t/2}$$

$$= \prod_{t=2}^{p} |\mathbf{I}_q - \hat{\mathbf{\Sigma}}_{z_{t1}z_{t1}}^{-1} \hat{\mathbf{\Sigma}}_{z_{t1}z_{t2}} \hat{\mathbf{\Sigma}}_{z_{t2}z_{t2}}^{-1} \hat{\mathbf{\Sigma}}_{z_{t2}z_{t1}}|^{n_t/2}$$
$$= \prod_{t=2}^{p} |\mathbf{I}_q - \hat{\mathbf{\Delta}}_t|^{n_t/2}$$

where $\hat{\mathbf{\Sigma}}_{y_t y_t \cdot y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t}$ and $\hat{\mathbf{\Sigma}}_{y_t y_t \cdot y_1 \ldots y_{t-1} x_0 x_1}$ denote the maximum likelihood estimator of the indicated variance based on the normal distribution given by

$$\hat{\mathbf{\Sigma}}_{y_t y_t \cdot y_1 \ldots y_{t-1} x_0 x_1} = \frac{1}{n_t}(\mathbf{Y}_t^{(t)} - \mathbf{X}_0^{(t)}\hat{\mathbf{\Gamma}}_{t0} - \mathbf{X}_1^{(t)}\hat{\mathbf{\Gamma}}_{t1} - \mathbf{Y}^{(t)}\hat{\mathbf{\Theta}}_t)^T(\mathbf{Y}_t^{(t)} - \mathbf{X}_0^{(t)}\hat{\mathbf{\Gamma}}_{t0} - \mathbf{X}_1^{(t)}\hat{\mathbf{\Gamma}}_{t1} - \mathbf{Y}^{(t)}\hat{\mathbf{\Theta}}_t)$$

and

$$\hat{\mathbf{\Sigma}}_{y_t y_t \cdot y_1 \ldots y_{t-1} x_0 x_1 \ldots x_t} = \frac{1}{n_t}(\mathbf{Y}_t^{(t)} - \mathbf{X}_0^{(t)}\hat{\mathbf{\Gamma}}_{t0} - \mathbf{X}^{(t)}\hat{\mathbf{\Gamma}}_{t} - \mathbf{Y}^{(t)}\hat{\mathbf{\Theta}}_t)^T(\mathbf{Y}_t^{(t)} - \mathbf{X}_0^{(t)}\hat{\mathbf{\Gamma}}_{t0} - \mathbf{X}^{(t)}\hat{\mathbf{\Gamma}}_{t} - \mathbf{Y}^{(t)}\hat{\mathbf{\Theta}}_t).$$

5.1 *Distribution under the null hypothesis.* It is not difficult to show that under the assumption of multivariate normality and H_0, the statistics $U_t = -2\log\Lambda_t$ are independently distributed, with each statistic converging to a chi-squared distribution with $qr(t-1)$ degrees of freedom as $n_t \to \infty$ (see Andreoni (1996), Chapter 2). The distribution of $-2\log\Lambda_t$ does not depend on $\mathbf{Y}_t^{(t)}$, $\mathbf{X}_0^{(t)}$, $\mathbf{X}^{(t)}$ and $\mathbf{Y}^{(t)}$ being held fixed or not, given that H_{0s} is true for $s = 1, \ldots, t$.

This fact suggests the use of a step-down testing procedure (Roy, Gnanadesikan and Srivastava, 1971) to test H_0. The procedure is to compare U_2 with the significance point u_2 for its respective degrees of freedom. If the observed value is larger than u_2, then reject H_0. If it is accepted, then compare U_3 with u_3. In sequence, the components are tested. If one is rejected, the sequence is stopped and the hypothesis H_0 is rejected. If all component null hypotheses are accepted, the composite hypothesis is accepted.

The proposed test has acceptance region A given by

$$A : \bigcap_{t=2}^{p} [U_t \le u_t],$$

where u_t's are to be chosen such that

$$P[U_t \le u_t | H_{0s}, \quad s = 1, \ldots, t] = 1 - \alpha_t,$$

and the probability of accepting H_0

$$P[\bigcap_{t=2}^{p} (U_t \le u_t | H_{0t})] = \prod_{t=2}^{p} (1 - \alpha_t) = 1 - \alpha,$$

for a prespecified α.

We can choose the significance levels based on the investigators' considerations. In the absence of any other reason, the component significance levels can be taken equal. If the investigator is more interested in the final measurements (like in the q responses in the pth visit) and α_p is a very small number, then it will take a relatively large deviation from the pth null hypothesis to lead to rejection.

The step-down procedure is appropriate, since there is a meaningful basis for considering the responses in the specified order. Considering them in different orders may, in general, lead to different conclusions.

One can improve the asymptotic distribution by using an asymptotic expansion for the distribution of the likelihood ratio criterion as explained in Anderson (1984) and Andreoni (1996).

5.2 *Power considerations.* In order to understand the complexity of the power function, consider the simplest situation when $q = 1$ and $f = 0$. Recall that in that case

$$\Lambda = \prod_{t=2}^{p} (1 - r_t^2)^{n_t/2}$$

where $r_t^2 = r^2_{z_{t1} z_{t2}}$ is the square of the multiple correlation coefficient between

$$\mathbf{Z}_{t1} = \mathbf{y}_t.\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}, \mathbf{X}_1$$

and

$$\mathbf{Z}_{t2} = \mathbf{X}_2, \ldots, \mathbf{X}_t.\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}, \mathbf{X}_1$$

for $t = 2, \ldots, p$. Let

$$\rho_t^2 = \frac{\boldsymbol{\sigma}^T_{z_{t2} z_{t1}} \boldsymbol{\Sigma}^{-1}_{z_{t2} z_{t2}} \boldsymbol{\sigma}_{z_{t2} z_{t1}}}{\sigma_{z_{t1} z_{t1}}}$$

be the square of the population multiple correlation coefficient between $\mathbf{Z}_{t1}$ and $\mathbf{Z}_{t2}$. When $q = 1$, it is well known (Anderson, 1984) that each $F_t = d_{t2} r_t^2 / [d_{t1}(1 - r_t^2)]$ is distributed according to a non-central F-distribution with $d_{t1} = r(t-1)$ and $d_{t2} = n_t - tr - 1$ degrees of freedom and non-centrality parameter given by

$$\delta_t = \frac{\boldsymbol{\sigma}^T_{z_{t2} z_{t1}} \boldsymbol{\Sigma}^{-1}_{z_{t2} z_{t2}} \mathbf{S}_{t22} \boldsymbol{\Sigma}^{-1}_{z_{t2} z_{t2}} \boldsymbol{\sigma}_{z_{t2} z_{t1}}}{\sigma_{z_{t1} z_{t1}.z_{t2}}}$$

for $t = 2, \ldots, p$ conditioning on the observed values of $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}, \mathbf{X}_1, \ldots, \mathbf{X}_t$ at each step t. The unconditional density function of r_t^2 is given by

$$\frac{\Gamma((d_{t1} + d_{t2})/2)}{\Gamma(d_{t1}/2)\,\Gamma(d_{t2}/2)} (r_t^2)^{(d_{t1}-2)/2} \, (1 - r_t^2)^{(d_{t2}-2)/2}$$

$$(1 - \rho_t^2)^{(d_{t1}+d_{t2})/2} \, {}_2F_1((d_{t1} + d_{t2})/2, (d_{t1} + d_{t2})/2; d_{t1}/2; \rho_t^2 \, r_t^2) \quad (0 < r_t^2 < 1).$$

where

$${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{z^k}{k!}$$

is the generalized hypergeometric function with $(a)_k = a(a+1)\cdots(a+k-1)$.

The distribution function of r_t^2 can be expressed in the form

$$P(r_t^2 \leq v_t) = \sum_{k=0}^{\infty} c_k P\left(F_{d_{t1}+2k, d_{t2}} \leq \frac{d_{t2}}{d_{t1}+2k}\frac{v_t}{1-v_t}\right),$$

where c_k is the negative binomial probability

$$c_k = (-1)^k \begin{pmatrix} -(d_{t1}+d_{t2})/2 \\ k \end{pmatrix} (1-\rho_t^2)^{(d_{t1}+d_{t2})/2}(\rho_t^2)^k.$$

The calculation of the distribution function is more complicated when $q \geq 2$.

The power of the proposed step-down procedure depends on the direction of the alternative hypothesis H_1. Let π denote the power function, that is,

$$\begin{aligned} \pi &= P(\text{Reject } H_0 \mid H_1) \\ &= 1 - P(\text{Accept } H_0 \mid H_1) \\ &= 1 - P(U_t \leq u_t, \text{ for } t = 2, \ldots, p \mid H_1) \\ &= 1 - P(U_2 \leq u_2 \mid H_1) \cdots P(U_p \leq u_p \mid H_1, U_s \leq u_s \text{ for } s = 2, \ldots, p-1). \end{aligned}$$

Let $\beta_t(\delta_t)$ denote the probability of accepting H_{0t} given that H_{0t} is false conditioned on the observed values of $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}, \mathbf{X}_1, \ldots, \mathbf{X}_t$. In that case, there are $2^{p-1} - 1$ situations for the alternative hypotheses as described below.

Case 1: If $H_1 = \{H_{02} \text{ true}, \ldots, H_{0p-1} \text{ true}, H_{0p} \text{ false}\}$ then the power given that $\mathbf{Y}_1$, $\mathbf{Y}_2$, …, $\mathbf{Y}_{p-1}$, $\mathbf{X}_1$, …, $\mathbf{X}_t$ are being held fixed is given by

$$\pi(\delta_p) = 1 - [\,\textstyle\prod_{t=2}^{p-1}(1-\alpha_t)\,]\,\beta_p(\delta_p).$$

Case 2: If $H_1 = \{H_{02} \text{ true}, \ldots, H_{0p-2} \text{ true}, H_{0p-1} \text{ false}, H_{0p} \text{ false}\}$ then

$$\begin{aligned} &\pi(\delta_{p-1}, \delta_p) = \\ &\quad 1 - [\,\textstyle\prod_{t=2}^{p-2}(1-\alpha_t)\,]\, P(U_{p-1} \leq u_{p-1}, U_p \leq u_p \mid H_1, U_t \leq u_t, \text{ for } t = 2, \ldots, p-2). \end{aligned}$$

Case 3: If $H_1 = \{H_{02} \text{ true}, \ldots, H_{0p-2} \text{ true}, H_{0p-1} \text{ false}, H_{0p} \text{ true}\}$ then

$$\begin{aligned} &\pi(\delta_{p-1}, \delta_p) = \\ &\quad 1 - [\,\textstyle\prod_{t=2}^{p-2}(1-\alpha_t)\,]\, P(U_{p-1} \leq u_{p-1}, U_p \leq u_p \mid H_1, U_t \leq u_t, \text{ for } t = 2, \ldots, p-2). \end{aligned}$$

Case $\mathbf{2^{p-1}}$–1: If $H_1 = \{H_{02} \text{ false}, \ldots, H_{0p} \text{ false}\}$ then

$$\pi(\delta_2, \ldots, \delta_p) \quad = \quad 1 - P(U_t \leq u_t, \text{ for } t = 2, \ldots, p \mid H_1).$$

The power function is given by $\pi = E[\pi(\boldsymbol{\delta}_c)]$ for each case c. Note that the calculation of the power function involves the knowledge of the joint cumulative distribution function of $(U_2, \ldots, U_p)$ because the test statistics involved are no longer independent under H_1.

Bootstrap method for power calculation. The empirical power can be calculated by means of bootstrap methods. Let $\hat{\mathbf{\Sigma}}_n = \text{diag}\,\{\hat{\mathbf{\Sigma}}_{xxn}, \hat{\mathbf{\Sigma}}_{een}\}$ be the maximum likelihood estimate of the variance-covariance matrix based on the original sample, $\hat{\mathbf{B}}_{t0}$ the maximum likelihood estimate of $\mathbf{B}_{t0}$ and $\hat{\mathbf{B}}_t$ the maximum likelihood estimate of $\mathbf{B}_t$.

Under the correlation model with independent errors (2.1) the covariates $\mathbf{x}_{it\bullet}$ and the residuals $\hat{\mathbf{e}}_{it\bullet}$ for $i = 1, \ldots, n_t$, $t = 1, \ldots, p$ are resampled independently.

Draw M samples from the original $n = n_1$ covariate measurements with replacement. For each sample draw n_p covariate observations with replacement from $\mathbf{x}_{i1\bullet}, \ldots, \mathbf{x}_{ip\bullet}$ for $i = 1, \ldots, n_p$; $n_{p-1} - n_p$ covariate observations with replacement from $\mathbf{x}_{i1\bullet}, \ldots, \mathbf{x}_{ip-1\bullet}$ for $i = 1, \ldots, n_{p-1} - n_p$; and so on, until $n_1 - n_2$ covariate observations with replacement from $\mathbf{x}_{i1\bullet}$ for $i = 1, \ldots, n_1 - n_2$ are obtained. Call these new covariate observations $\mathbf{x}^{*(p)}_{i\bullet\bullet} = [\mathbf{x}^*_{i1\bullet}, \ldots, \mathbf{x}^*_{ip\bullet}]^T$ for $i = 1, \ldots, n_p$; $\mathbf{x}^{*(p-1)}_{i\bullet\bullet} = \mathbf{x}^*_{i1\bullet}, \ldots, \mathbf{x}_{ip-1\bullet}$ for $i = 1, \ldots, n_{p-1} - n_p$; and so on $\mathbf{x}^{(1)}_{i\bullet\bullet} = \mathbf{x}_{i1\bullet}$ for $i = 1, \ldots, n_1 - n_2$.

Draw M samples from the $n = n_1$ residuals with replacement independently of the above sampling scheme for the covariate observations such that you draw n_p from $\hat{\mathbf{e}}_{i1\bullet}, \ldots, \hat{\mathbf{e}}_{ip\bullet}$ with replacement for $i = 1, \ldots, n_p$; $n_{p-1} - n_p$ from $\hat{\mathbf{e}}_{i1\bullet}, \ldots, \hat{\mathbf{e}}_{ip-1\bullet}$ with replacement for $i = 1, \ldots, n_{p-1} - n_p$; $\ldots$, $n_1 - n_2$ from $\hat{\mathbf{e}}_{i1\bullet}$ with replacement for $i = 1, \ldots, n_1 - n_2$.

For each sample put

$$\begin{aligned} \mathbf{y}^*_{it\bullet} &= \hat{\mathbf{B}}^T_{t0}\mathbf{x}_{i0\bullet} + \hat{\mathbf{B}}^T_{t1}\mathbf{x}^*_{i1\bullet} + \ldots + \hat{\mathbf{B}}^T_{tt}\mathbf{x}^*_{it\bullet} + \hat{\mathbf{e}}^*_{it\bullet} \\ &= \hat{\mathbf{B}}^T_{t0}\mathbf{x}_{i0\bullet} + \hat{\mathbf{B}}^T_{t}\mathbf{x}^{*(t)}_{i\bullet\bullet} + \hat{\mathbf{e}}^*_{it\bullet} \end{aligned}$$

The joint distribution of $\{(\mathbf{x}^{*(t)}_{i\bullet\bullet}, \mathbf{y}^{*(t)}_{i\bullet\bullet})\ i = 1, \ldots, n_t,\ t = 1, \ldots, p\}$, conditional on $\{(\mathbf{x}^{(t)}_{i\bullet\bullet}, \mathbf{y}^{(t)}_{i\bullet\bullet})\ i = 1, \ldots, n_t,\ t = 1, \ldots, p\}$, is the bootstrap estimate of the unconditional distribution of $\{(\mathbf{x}^{(t)}_{i\bullet\bullet}, \mathbf{y}^{(t)}_{i\bullet\bullet})\ i = 1, \ldots, n_t,\ t = 1, \ldots, p\}$ under the correlation model (2.1) with independent errors.

For each sample calculate $\hat{\mathbf{\Sigma}}^*_n$, the maximum likelihood estimate of the variance matrix for each bootstrap sample. For each $\hat{\mathbf{S}}^*_n$ calculate the corresponding $(T^*_2, \ldots, T^*_p)$ and count the number of times that the hypothesis is rejected, let us call this number m. So the empirical power of the procedure is given by

$$\pi_e = \frac{m}{M}.$$

Beran and Srivastava (1985) showed that the empirical distribution of the bootstrap test statistic converges to the distribution of the true test statistic.

6. Numerical Example

The data in this example correspond to a part of a randomized multi-visit clinical trial to test a cholesterol lowering drug against placebo. Initially there were 1835 subjects in the group on active treatment and 1843 taking placebo. The age of each subject was recorded at entrance in the trial. Two liver function measurements (alkaline phosphatase and SGOT) from two groups of men (on drug and placebo) were observed at eight regularly spaced visits and the alcohol consumption (grams/day) was recorded for each individual. The monotone missing data pattern of the data scheduled to be taken one year apart is shown in Table 1. Only 42.88% of the subjects have data for all eight measurements. Because of progressive recruitment and a fixed termination date, most of the subjects were followed for less than the full follow-up time.

Table 1. Pattern of missing data.

Visit 1	2	3	4	5	6	7	8	Number of Subjects
×	×	×	×	×	×	×	×	1577
×	×	×	×	×	×	×		884
×	×	×	×	×	×			167
×	×	×	×	×				149
×	×	×	×					305
×	×	×						323
×	×							160
×								113
3678	3565	3405	3082	2777	2628	2461	1577	3678

The following model was used:

$$\begin{pmatrix} y_{it1} \\ y_{it2} \end{pmatrix} = \mathbf{B}_{t0}^{T} \begin{pmatrix} x_{i01} \\ x_{i02} \\ x_{i03} \end{pmatrix} + \mathbf{B}_{t1}^{T}\, x_{i11} + \ldots + \mathbf{B}_{tt}^{T}\, x_{it1} + \begin{pmatrix} e_{it1} \\ e_{it2} \end{pmatrix}$$

where y_{it_1} is the transformed alkaline phosphatase, y_{it_2} is the transformed SGOT, x_{i01} is equal to 1, x_{i02} indicates the treatment: 1 if on drug and 0 if on placebo, x_{i03} is the age at randomization, and x_{it_1} is the transformed alcohol consumption for each individual i at visit t, for $t = 1, \ldots, 8$. The following transformations were used to reduce the skewness of the responses and random covariates: $y_{it1}^{*0.75}$ for alkaline phosphatase, $y_{it_2}^{*-0.50}$ for SGOT, and $x_{it_1}^{*0.25}$ for alcohol consumption.

The step-down procedure can be used to test whether the information on alcohol consumption recorded at visits 2 to 8 do not provide more information to explain the change in the liver functions than the alcohol consumption recorded at visit one alone. The step-down test statistics for that purpose are presented in the column labelled 2 to 8 in Table 2. For $\alpha = 0.05$, we can take $\alpha_s = 0.0073$ at each step $s = 1, \ldots, 7$, and observe that the results in Table 2 show that the step-down procedure rejected the null hypothesis in step 1, indicating that alcohol consumption should be collected at visit 2.

Table 2. STEP-DOWN STATISTICS.

			Inclusion of covariates from visit						
Step	df	Statistics	2 to 8	3 to 8	4 to 8	5 to 8	6 to 8	7 to 8	8
1	2	χ^2_{obs}	14.7645	2.5811	30.7800	3.8133	14.6130	10.1060	10.5139
		P-value	0.0006	0.2751	0.0000	0.1486	0.0007	0.0064	0.0052
2	4	χ^2_{obs}	10.3286	33.7152	4.1135	15.8978	10.6954	15.6930	
		P-value	0.0352	0.0000	0.3909	0.0032	0.0302	0.0035	
3	6	χ^2_{obs}	34.9135	7.2578	17.1625	11.1118	16.1670		
		P-value	0.0000	0.2977	0.0087	0.0850	0.0129		
4	8	χ^2_{obs}	12.4086	21.5340	18.3908	23.2426			
		P-value	0.1339	0.0059	0.0185	0.0030			
5	10	χ^2_{obs}	25.4326	20.8141	24.7437				
		P-value	0.0046	0.0224	0.0059				
6	12	χ^2_{obs}	22.5018	25.5302					
		P-value	0.0323	0.0125					
7	14	χ^2_{obs}	28.9578						
		P-value	0.0106						

Similar test statistics can be computed to test if the additional covariates observed at certain visits do not provide more information than the ones observed at the previous visits and are presented in Table 2. The results indicate that alcohol consumption should be recorded at every visit in order to explain the changes in the liver functions.

Acknowledgements. The author would like to thank the referees and editors for their helpful comments and suggestions which led to the improvement of this paper. The author is grateful to Dr P. K. Sen for enabling and supervising the completion of this work.

References

ANDERSON, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd edn, Wiley, New York.

ANDREONI, S. (1996). *A Step-Down Procedure for Analysis of Time-Varying Covariates in Multi-Visit Studies*, University of North Carolina, Institute of Statistics, Mimeo Series No. 2175T.

BERAN, R. AND SRIVASTAVA, M.S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix, *Ann. Statist.*, **13**, 95-115.

FUJISAWA, H. (1995). A note on the maximum likelihood estimators for multivariate normal distribution with monotone data, *Communications in Statistics - Theory and Methods*, **24**, 1377-1382.

MCDONALD, L.L. (1975). Tests for the general linear hypothesis under the multiple design multivariate linear model, *Ann. Statist.*, **3**, 461-466.

ROY, S.N., GNANADESIKAN, R. AND SRIVASTAVA, J.N. (1971). *Analysis and Design of Certain Quantitative Multi-Response Experiments*, Pergamon Press, New York.

SRIVASTAVA, J.N. (1966). Some generalizations of multivariate analysis of variance, In *Multivariate Analysis*, ed. P.R. Krishnaiah, Academic Press, New York, 129-145.

SOLANGE ANDREONI
LABORATÓRIOS PFIZER LTDA.
SÃO PAULO, SÃO PAULO
BRAZIL
Email: ANDRES1@pfizer.com