THE UNIVERSITY OF TEXAS AT SAN ANTONIO, COLLEGE OF BUSINESS

Working Paper SERIES

Date June 23, 2009

WP # 0090MSS-253-2009

Classification Rules for Multivariate Repeated Measures Data with Equicorrelated Correlation Structure on both Time and Spatial Repeated Measurements

Anuradha Roy
Department of Management Science and Statistics
University of Texas at San Antonio

Ricardo Leiva
Universidad Nacional de Cuyo

UTSA COLLEGE OF BUSINESS KNOWLEDGE FOR A NEW WORLD

ONE UTSA CIRCLE
SAN ANTONIO, TEXAS 78249-0631
210 458-4317 | BUSINESS.UTSA.EDU

Classification Rules for Multivariate Repeated Measures Data with Equicorrelated Correlation Structure on both Time and Spatial Repeated Measurements

Anuradha Roy
Department of Management Science and Statistics
The University of Texas at San Antonio
San Antonio, TX 78249, USA
E-mail: aroy@utsa.edu

Ricardo Leiva
Departamento de Matemática
F.C.E., Universidad Nacional de Cuyo
5500 Mendoza, Argentina

Abstract

We study the problem of classification for multivariate repeated measures data with structured correlations on both time and spatial repeated measurements. This is a very important problem in many biomedical as well as in engineering field. Classification rules as well as the algorithm to compute the maximum likelihood estimates of the required parameters are given.

Key Words: Kronecker product covariance structure; Repeated observations; Maximum Likelihood Estimates.

JEL Classification: C10, C13

1 Introduction

We develop classification rules for multivariate repeated measures data with structured correlations on repeated measures on both spatial as well as over time. The available classification rules for multivariate repeated measures data consider structured correlation only on repeated measures over time. Nevertheless, in many biomedical applications just one variable is measured on different parts of the body and repeatedly over time, where use of structured correlations on repeated measures on spatial as well as over time would be natural/ beneficial. These problems are computationally very challenging, as it is not possible to tract them analytically or find any closed form solution.

Classification problem on multivariate repeated measures data, where measurements on a number of variables are measured repeatedly over time, was first studied by Gupta (1980, 1986).

Roy and Khattree (2005 a, b; 2007) considered the problem in small sample situation by assuming Kronecker product structure on variance covariance matrix. They assumed equicorrelated or compound symmetry correlation structure on the repeated measures in their 2005 b paper, and an autoregressive of order one $\big(\text{AR}(1)\big)$ structure on repeated measures in their other two papers. In many clinical trial problems it is found that the measurements on a single variable is measured on different body positions and repeatedly over time. For example positron emission tomography (PET) imaging aids in diagnosing different types of dementia. A healthy brain shows normal metabolism levels (measurements) throughout the scan. Low metabolism in the temporal and parietal lobes (sides and back) on both sides (sites) of the brain is seen in Alzheimer's disease. Repeated measurements of PET scan may diagnosis a patient with Alzheimer's disease. In another example, for the classification of patients between two different osteoporosis drug treated populations in two clinical trials. Osteoporosis can be detected by a test of Bone Mineral Density (BMD), the assessments of which are obtained at different anatomic locations of the body, such as the spine, radius, femoral neck and the total hip and all the measurements were observed at repeatedly over time. In this article we develop classification rules for these kinds of data. Different time points as well as different sites may have different measurement variations for the variables, and we should take these variations into account while analyzing these kinds of data. It is well known (Hand, 1997) that the correlation structure on the repeated measurements follows a simple pattern such as compound symmetry or a first-order autoregressive $\big(\text{AR}(1)\big)$ structure as opposed to the unstructured variance-covariance matrix, where the mean vectors and the variances and covariances among the pu variables are arbitrary. Therefore, for both the data sets it is expected that measurement variation over sites as well as over time both will have patterned covariance structures. In other words, marginal variance-covariance matrices over different sites as well as over different time points will have patterned covariance structures. In this paper we will develop classification rules for multivariate repeated measures data where both the marginal variance-covariance matrices over different sites as well as over different time points have patterned covariance structures.

Let $y_{jr,ts}$ be the measurement on the r^{th} individual at the s^{th} site (location) and at the t^{th} time point in the j^{th} population; $r = 1, \dots, n$, $s = 1, \dots, u$, $t = 1, \dots, p, j = 1, \dots, k$. Let $\boldsymbol{y}_{jr,t}$ be the u-variate vector of all measurements corresponding to the r^{th} individual at the t^{th} time point, that is, for each r, and t, $\boldsymbol{y}_{jr,t}$ is obtained by stacking the response of the r^{th} individual at the t^{th} time point at the first site (location), then stacking the response at the second site and so on. Let $\boldsymbol{y}_{jr} = (\boldsymbol{y}'_{jr,1}, \boldsymbol{y}'_{jr,2}, \dots, \boldsymbol{y}'_{jr,p})'$ be the pu-variate vector of all measurements corresponding to the r^{th} individual. For two populations, $\boldsymbol{Y}_j = \left[\boldsymbol{y}_{\boldsymbol{j1}}, \boldsymbol{y}_{j2}, \dots, \boldsymbol{y}_{jn}\right]$ be n_j independent random samples from populations $N_{pu}\left(\boldsymbol{\mu}_j, \boldsymbol{\Omega}_j\right)$, where $\boldsymbol{\mu}_j \in \mathbb{R}^{pu}$ and the matrix $\boldsymbol{\Omega}_j$ is assumed to be $pu \times pu$ positive definite matrix. We assume the form of the covariance matrix $\boldsymbol{\Omega}_j$ as

$$\underset{pu\times pu}{\boldsymbol{\Omega}_j} = \underset{p\times p}{\boldsymbol{V}_j} \otimes \underset{u\times u}{\boldsymbol{\Delta}_j},$$

where both $\boldsymbol{V}_j$ and $\boldsymbol{\Delta}_j$ have equicorrelated structures. However, this may result in identifiability problem. We circumvent this problem by taking $\boldsymbol{V}_j$ as a equicorrelated correlation structure and $\boldsymbol{\Delta}_j$ as a equicorrelated covariance structure. The matrix $\boldsymbol{\Delta}_j$ has the form

$$\boldsymbol{\Delta}_j = \left(\sigma_{j,0}^2 - \sigma_{j,1}^2\right) \mathbf{I}_u + \sigma_{j,1}^2 \mathbf{J}_u,$$

where $\mathbf{I}_u$ is the $u \times u$ identity matrix, $\mathbf{1}_u$ is a $u \times 1$ vector containing all elements as unity, $\mathbf{J}_u = \mathbf{1}_u \mathbf{1}'_u$. It is well known that

$$\boldsymbol{\Delta}_j^{-1} \;=\; \left(\sigma_{j,0}^2 - \sigma_{j,1}^2\right)^{-1} \mathbf{I}_u + \frac{1}{u}\left[\left(\sigma_{j,0}^2 + (u-1)\sigma_{j,1}^2\right)^{-1} - \left(\sigma_{j,0}^2 - \sigma_{j,1}^2\right)^{-1}\right] \mathbf{J}_u.$$

That is, $\boldsymbol{\Delta}_j^{-1}$ also has the form

$$\boldsymbol{\Delta}_j^{-1} = h_j \mathbf{I}_u + k_j \mathbf{J}_u, \tag{1}$$

where

$$h_j = \left(\sigma_{j,0}^2 - \sigma_{j,1}^2\right)^{-1},$$

and

$$k_j \;=\; \frac{1}{u}\left[\left(\sigma_{j,0}^2 + (u-1)\sigma_{j,1}^2\right)^{-1} - \left(\sigma_{j,0}^2 - \sigma_{j,1}^2\right)^{-1}\right]$$

The determinant of $\boldsymbol{\Delta}_j$ is given by

$$|\boldsymbol{\Delta}_j| = \left|\sigma_{j,1}^2 - \sigma_{j,0}^2\right|^{u-1} \left|\sigma_{j,0}^2 + (u-1)\sigma_{j,1}^2\right|. \tag{2}$$

The correlation matrix $\boldsymbol{V}_j,\ j = 1, 2$ is given by

$$\boldsymbol{V}_j = (1-\rho_j)\mathbf{I}_p + \rho_j \mathbf{1}_p \mathbf{1}'_p.$$

The elements v_i^{lm} of $\boldsymbol{V}_j^{-1}$ is given by

$$v_j^{lm} = \begin{cases} \dfrac{1+(p-2)\rho_j}{(1-\rho_j)\{1+(p-1)\rho_j\}}, & \text{if} \quad l = m, \\ -\dfrac{\rho_j}{(1-\rho_j)\{1+(p-1)\rho_j\}}, & \text{if} \quad l \neq m. \end{cases} \tag{3}$$

The determinant of $\boldsymbol{V}_j$ is given by

$$|\boldsymbol{V}_j| = \big(1+(p-1)\rho_j\big)(1-\rho_j)^{p-1}, j = 1, 2. \tag{4}$$

Since $\boldsymbol{V}_j$ has to be positive definite, we should have $-\dfrac{1}{p-1} < \rho_j < 1$. However, we further assume that $0 < \rho_j < 1$.

2 Classification Rules

Case 1: $\mathbf{\Omega}_1 = \mathbf{\Omega}_2$.

Sample classification rule is given by:

Classify an individual with response $\boldsymbol{y}$ to Population 1 if

$$(\widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\mu}}_2)' \left(\widehat{\boldsymbol{V}}^{-1} \otimes \widehat{\mathbf{\Delta}}^{-1}\right) \boldsymbol{y} \geq \frac{1}{2} (\widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\mu}}_2)' \left(\widehat{\boldsymbol{V}}^{-1} \otimes \widehat{\mathbf{\Delta}}^{-1}\right) (\widehat{\boldsymbol{\mu}}_1 + \widehat{\boldsymbol{\mu}}_2)',$$

and to Population 2 otherwise.

Maximum likelihood estimation of $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{V}$, *and* $\mathbf{\Delta}$: Let $n = n_1 + n_2$ be the total number of random samples $\boldsymbol{Y}_j = \left[\boldsymbol{y}_{j1}, \boldsymbol{y}_{j2}, \ldots, \boldsymbol{y}_{jn_j}\right]$ from Population $j, j = 1, 2$. Here we assume $\boldsymbol{\mu}_j = (\mu_{j,ts})'_{t=1,\ldots,p;s=1,\ldots,u}$. Using (1) and(2) the log likelihood function $\ln L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{V}, \mathbf{\Delta}; \boldsymbol{Y}_1, \boldsymbol{Y}_2)$ is given by

$$\begin{aligned} \ln L \;=\; & -\frac{npu}{2} \ln(2\pi) - \frac{nu}{2} \ln|\mathbf{V}| - \frac{np(u-1)}{2} \ln\left|\sigma_0^2 - \sigma_1^2\right| - \frac{np}{2} \ln\left|\sigma_0^2 + (u-1)\sigma_1^2\right| \\ & -\frac{1}{2} \sum_{j=1}^{2} \sum_{r=1}^{n_j} \sum_{m=1}^{p} \sum_{l=m-1}^{m+1} \sum_{s=1}^{u} hv^{lm} (y_{jr,ls} - \mu_{j,ls})(y_{jr,ms} - \mu_{j,ms}) \\ & -\frac{1}{2} \sum_{j=1}^{2} \sum_{r=1}^{n_j} \sum_{m=1}^{p} \sum_{l=m-1}^{m+1} \sum_{s=1}^{u} \sum_{s^*=1}^{u} kv^{lm} (y_{jr,ls} - \mu_{j,ls})(y_{jr,ms^*} - \mu_{j,ls^*}). \end{aligned} \tag{5}$$

An alternative expression for $\ln L$ is

$$\begin{aligned} \ln L \;=\; & -\frac{npu}{2} \ln(2\pi) - \frac{n}{2} \ln|\boldsymbol{V} \otimes \mathbf{\Delta}| - \frac{1}{2} \mathrm{tr}(\boldsymbol{V} \otimes \mathbf{\Delta})^{-1} (\boldsymbol{S}_1 + \boldsymbol{S}_2) \\ & -\frac{1}{2} \mathrm{tr}(\boldsymbol{V} \otimes \mathbf{\Delta})^{-1} \sum_{j=1}^{2} n_j (\overline{\boldsymbol{y}}_j - \boldsymbol{\mu}_j)(\overline{\boldsymbol{y}}_j - \boldsymbol{\mu}_j)'. \end{aligned}$$

where

$$\boldsymbol{S}_j = \sum_{r=1}^{n_j} (\boldsymbol{y}_{jr} - \overline{\boldsymbol{y}}_j)(\boldsymbol{y}_{jr} - \overline{\boldsymbol{y}}_j)', \text{ for } j = 1, 2,$$

and $\overline{\boldsymbol{y}}_j$ is the sample mean vector for the jth group. The vector $\overline{\boldsymbol{y}}_j = \left(\overline{\boldsymbol{y}}'_{j,1}, \overline{\boldsymbol{y}}'_{j,2}, \ldots, \overline{\boldsymbol{y}}'_{j,p}\right)'$, with $\overline{\boldsymbol{y}}_{j,t} = \frac{1}{n_j} \sum_{r=1}^{n_j} \boldsymbol{y}_{jr,t} = (\overline{y}_{j\cdot,t1}, \overline{y}_{j\cdot,t2}, \ldots, \overline{y}_{j\cdot,tu})'$, for $t = 1, \ldots, p$. It is obvious that the MLEs of $\boldsymbol{\mu}_j$ are $\widehat{\boldsymbol{\mu}}_j = \overline{\boldsymbol{y}}_j$ for $j = 1, 2$. Now, replacing $\boldsymbol{\mu}_j$ by $\widehat{\boldsymbol{\mu}}_j$ the log likelihood function reduces to

$$\ln L = -\frac{npu}{2} \ln(2\pi) - \frac{n}{2} \ln\left(|\boldsymbol{V}|^u |\mathbf{\Delta}|^p\right) - \frac{1}{2} \mathrm{tr}\left(\boldsymbol{V}^{-1} \otimes \mathbf{\Delta}^{-1}\right) \boldsymbol{S},$$

where $\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2$. By substituting the values of $|\boldsymbol{V}|$ and $\boldsymbol{V}^{-1}$ in the above equation we get

$$\begin{aligned} \ln L = & -\frac{npu}{2} \ln 2\pi - \frac{n(p-1)u}{2} \ln(1-\rho) - \frac{nu}{2} \ln\{1 + (p-1)\rho\} \\ & - \frac{np}{2} \ln|\mathbf{\Delta}| - \frac{1}{2(1-\rho)} c_1^* + \frac{\rho}{2(1-\rho)\{1 + (p-1)\rho\}} d_1^*, \end{aligned} \tag{6}$$

where $c_1^* = \text{tr}[(\mathbf{I}_p \otimes \boldsymbol{\Delta}^{-1})\boldsymbol{S}]$ and and $d_1^* = \text{tr}[(\mathbf{J}_p \otimes \boldsymbol{\Delta}^{-1})\boldsymbol{S}]$.
Differentiating (6) with respect to ρ, equating it to zero and simplifying we get,

$$\begin{aligned}(p-1)k_0\rho^3 + \{k_0 - (p-1)k_0 + (p-1)^2 c_1^* - (p-1)d_1^*\}\rho^2 \\ + \{2(p-1)c_1^* - k_0\}\rho + (c_1^* - d_1^*) = 0,\end{aligned} \tag{7}$$

where $k_0 = nu(p-1)p$. Alternatively, from (5) we get

$$\begin{aligned}\ln L &= -\frac{npu}{2}\ln(2\pi) - \frac{nu}{2}|\boldsymbol{V}| - \frac{np(u-1)}{2}\ln\left|h^{-1}\right| - \frac{np}{2}\ln\left|m^{-1}\right| \\ & -\frac{1}{2}h\left(b_{1,1}^* - \frac{1}{u}b_{1,2}^*\right) - \frac{1}{2u}mb_{1,2}^*,\end{aligned}$$

where

$$\begin{aligned}h &= \frac{1}{\sigma_0^2 - \sigma_1^2}, \\ m &= \frac{1}{\sigma_0^2 + (u-1)\sigma_1^2},\end{aligned}$$

$$b_{1,1}^* = \sum_{j=1}^{2}\sum_{r=1}^{n_j}\sum_{m=1}^{p}\sum_{l=1}^{p}\sum_{s=1}^{u} v^{lm}\left(y_{jr,ls} - \overline{y}_{j\cdot,ls}\right)\left(y_{jr,ms} - \overline{y}_{j\cdot,ms}\right)',$$

$$\text{and} \quad b_{1,2}^* = \sum_{j=1}^{2}\sum_{r=1}^{n_j}\sum_{m=1}^{p}\sum_{l=1}^{p}\sum_{s=1}^{u}\sum_{s^*=1}^{u} v^{lm}\left(y_{jr,ls} - \overline{y}_{j\cdot,ls}\right)\left(y_{jr,ms^*} - \overline{y}_{j\cdot,ms^*}\right).$$

Differentiating (Harville, 1997) the above equation with respect to h^{-1} and m^{-1} separately and then equating them to zero we get

$$\widehat{h^{-1}} = \frac{1}{np(u-1)}\left(b_{1,1}^* - \frac{1}{u}b_{1,2}^*\right),$$

$$\text{and} \quad \widehat{m^{-1}} = \frac{1}{np}b_{1,2}^*.$$

After some simplifications we get

$$\widehat{\sigma}_0^2 = \frac{b_{1,1}^*}{npu}, \tag{8}$$

$$\text{and} \quad \widehat{\sigma}_1^2 = \frac{b_{1,2}^* - b_{1,1}^*}{npu(u-1)}. \tag{9}$$

The MLEs $\widehat{\rho}, \widehat{\sigma}_0^2$ and $\widehat{\sigma}_1^2$ are obtained by simultaneously and iteratively solving (7), (8) and (9) by substituting the values of v^{lm}; $l, m = 1, 2, \ldots, p$, from equation (3). The computations can be carried out by the following algorithm. The MLE of $\boldsymbol{V}$ is obtained from

$$\widehat{\boldsymbol{V}} = (1-\widehat{\rho})\mathbf{I}_p + \widehat{\rho}\mathbf{1}_p\mathbf{1}_p', \tag{10}$$

and the MLE of $\mathbf{\Delta}$ is obtained from

$$\widehat{\mathbf{\Delta}} = \mathbf{I}_u \left(\widehat{\sigma}_0^2 - \widehat{\sigma}_1^2\right) + \mathbf{J}_u \widehat{\sigma}_1^2. \tag{11}$$

Algorithm Outline:

Step 1: Get the pooled sample variance covariance matrix $\boldsymbol{G}$ for the repeated measures. Then obtain an initial estimate of ρ as $\widehat{\rho}_o = (\mathbf{1}_p' \boldsymbol{G} \mathbf{1}_p - \operatorname{tr} \boldsymbol{G})/p(p-1)$.

Step 2: Compute $\widehat{\sigma}_0^2$ and $\widehat{\sigma}_1^2$ from (8) and (9), and then compute $\widehat{\mathbf{\Delta}}$ from (11).

Step 3: Compute c_1^* and d_1^* using $\widehat{\mathbf{\Delta}}$ obtained in Step 2.

Step 4: Compute $\widehat{\rho}$ by solving the cubic equation (7). Ensure that $0 < \widehat{\rho} < 1$. Truncate $\widehat{\rho}$ to 0 or 1, if it is outside this range.

Step 5: Compute the revised estimate $\widehat{\boldsymbol{V}}$ from $\widehat{\rho}$ by using (10).

Step 6: Repeat Steps 2 to 5 until convergence is attained. This is ensured by verifying that the maximum of the absolute difference between two successive values of $\widehat{\rho}$, $\widehat{\sigma}_0^2$ and $\widehat{\sigma}_1^2$ is less than ϵ.

Even though ρ is always between $-\frac{1}{p-1}$ and 1, we have assumed $0 < \rho < 1$. Still, $\widehat{\rho}$ may fall at the boundary $\rho = 1$, in which case the standard asymptotic theory may not be directly applicable. See, Self and Liang, (1987) for more details.

Case 2: $\mathbf{\Omega}_1 \neq \mathbf{\Omega}_2$ ($\boldsymbol{V}_1 \neq \boldsymbol{V}_2, \mathbf{\Delta}_1 \neq \mathbf{\Delta}_2$).

Sample classification rule is given by:

Classify an individual with response $\boldsymbol{y}$ to Population 1 if

$$\begin{aligned} &\sum_{j=1}^{2} (-1)^{j-1} \left[\overline{\boldsymbol{y}}_j' \left(\widehat{\boldsymbol{V}}_j^{-1} \otimes \widehat{\mathbf{\Delta}}_j^{-1} \right) \boldsymbol{y} - \frac{1}{2} \boldsymbol{y}' \left(\widehat{\boldsymbol{V}}_j^{-1} \otimes \widehat{\mathbf{\Delta}}_j^{-1} \right) \boldsymbol{y} \right] \\ \geq\ & \frac{1}{2} \sum_{j=1}^{2} (-1)^{j-1} \left[\ln \left| \widehat{\boldsymbol{V}}_j \right|^u \left| \widehat{\mathbf{\Delta}}_j \right|^p + \overline{\boldsymbol{y}}_j' \left(\widehat{\boldsymbol{V}}_j^{-1} \otimes \widehat{\mathbf{\Delta}}_j^{-1} \right) \overline{\boldsymbol{y}}_j \right], \end{aligned}$$

and to Population 2 otherwise.

Using (1) and (2) the the log likelihood function $\ln L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{V}_1, \boldsymbol{V}_2, \mathbf{\Delta}_1, \mathbf{\Delta}_2; \boldsymbol{Y}_1, \boldsymbol{Y}_2)$ is given by

$$\begin{aligned} \ln L = & -\frac{npu}{2} \ln 2\pi - \frac{n_1(p-1)u}{2} \ln(1-\rho_1) \\ & -\frac{n_2(p-1)u}{2} \ln(1-\rho_2) - \frac{n_1 u}{2} \ln\{1+(p-1)\rho_1\} \\ & -\frac{n_2 u}{2} \ln\{1+(p-1)\rho_2\} - \frac{n_1 p}{2} \ln|\mathbf{\Delta}_1| \\ & -\frac{n_2 p}{2} \ln|\mathbf{\Delta}_2| - \frac{1}{2(1-\rho_1)} c_1 - \frac{1}{2(1-\rho_2)} c_2 \\ & + \frac{\rho_1}{2(1-\rho_1)\{1+(p-1)\rho_1\}} d_1 + \frac{\rho_2}{2(1-\rho_2)\{1+(p-1)\rho_2\}} d_2, \end{aligned} \tag{12}$$

where $c_j = \mathrm{tr}[(\mathbf{I}_p \otimes \boldsymbol{\Delta}_j^{-1})\boldsymbol{S}_j]$ and and $d_j = \mathrm{tr}[(\mathbf{J}_p \otimes \boldsymbol{\Delta}_j^{-1})\boldsymbol{S}_j]$. Differentiating (12) with respect to $\rho_j, j = 1, 2$, equating it to zero and simplifying, results in the following equation

$$\begin{aligned}(p-1)k_{j0}\rho_j^3 + \{k_{j0} - (p-1)k_{j0} + (p-1)^2 c_j - (p-1)d_j\}\rho_j^2 \\ + \{2(p-1)c_j - k_{j0}\}\rho_j + (c_j - d_j) = 0,\end{aligned} \tag{13}$$

where $k_{j0} = n_j u(p-1)p$. Alternatively from (12) we get

$$\begin{aligned}\ln L &= -\frac{npu}{2}\ln(2\pi) - \frac{n_1 u}{2}\ln|\boldsymbol{V}_1| - \frac{n_2 u}{2}\ln|\boldsymbol{V}_2| - \frac{n_1 p(u-1)}{2}\ln\left|h_1^{-1}\right| \\ &\quad -\frac{n_2 p(u-1)}{2}\ln\left|h_2^{-1}\right| - \frac{n_1 p}{2}\ln\left|m_1^{-1}\right| - \frac{n_2 p}{2}\ln\left|m_2^{-1}\right| \\ &\quad -\frac{1}{2}h_1 b_{1,1} - \frac{1}{2}k_1 b_{1,2} - \frac{1}{2}h_2 b_{2,1} - \frac{1}{2}k_2 b_{2,2},\end{aligned} \tag{14}$$

where

$$\begin{aligned}h_j &= \frac{1}{\sigma_{j,0}^2 - \sigma_{j,1}^2}, \\ m_j &= \frac{1}{\sigma_{j,0}^2 + (u-1)\sigma_{j,1}^2}\end{aligned}$$

$$\begin{aligned}b_{j,1} &= \sum_{r=1}^{n_j}\sum_{m=1}^{p}\sum_{l=1}^{p}\sum_{s=1}^{u} v_j^{lm}\left(y_{jr,ls} - \overline{y}_{j\cdot,ls}\right)\left(y_{jr,ms} - \overline{y}_{j\cdot,ms}\right) \\ \text{and} \quad b_{j,2} &= \sum_{r=1}^{n_j}\sum_{m=1}^{p}\sum_{l=1}^{p}\sum_{s=1}^{u}\sum_{s^*=1}^{u} v_j^{lm}\left(y_{jr,ls} - \overline{y}_{j\cdot,ls}\right)\left(y_{jr,ms^*} - \overline{y}_{j\cdot,ms^*}\right).\end{aligned}$$

After some algebraic simplification from (14) we get

$$\begin{aligned}\ln L &= -\frac{npu}{2}\ln(2\pi) - \frac{n_1 u}{2}\ln|\boldsymbol{V}_1| - \frac{n_2 u}{2}\ln|\boldsymbol{V}_2| - \frac{n_1 p(u-1)}{2}\ln\left|h_1^{-1}\right| \\ &\quad -\frac{n_2 p(u-1)}{2}\ln\left|h_2^{-1}\right| - \frac{n_1 p}{2}\ln\left|m_1^{-1}\right| - \frac{n_2 p}{2}\ln\left|m_2^{-1}\right| \\ &\quad -\frac{1}{2}h_1\left(b_{1,1} - \frac{1}{u}b_{1,2}\right) - \frac{1}{2}h_2\left(b_{2,1} - \frac{1}{u}b_{2,2}\right) - \frac{1}{2u}m_1 b_{1,2} - \frac{1}{2u}m_2 b_{2,2}.\end{aligned}$$

Differentiating (Harville, 1997) the above equation with respect to h_j^{-1} and m_j^{-1} separately and then equating them to zero we get

$$\widehat{h_j^{-1}} = \frac{1}{n_j p(u-1)}\left(b_{j,1} - \frac{1}{u}b_{j,2}\right),$$

and

$$\widehat{m_j^{-1}} = \frac{1}{n_j p u}b_{j,2}.$$

After simplification we get

$$\widehat{\sigma}_{j,0}^2 = \frac{b_{j,1}}{n_j p u},\ j = 1, 2, \tag{15}$$

$$\text{and}\quad \widehat{\sigma}_{j,1}^2 = \frac{b_{j,2} - b_{j,1}}{n_j p u\,(u-1)},\ j = 1, 2. \tag{16}$$

The maximum likelihood estimates $\widehat{\rho}_1, \widehat{\rho}_2, \widehat{\sigma}_{10}^2, \widehat{\sigma}_{11}^2$, $\widehat{\sigma}_{20}^2$ and $\widehat{\sigma}_{21}^2$ are obtained by simultaneously and iteratively solving (13), (15) and (16). The computations can be carried out by a similar algorithm presented in Case 1. The MLEs of $\boldsymbol{V}_j$ and $\boldsymbol{\Delta}_j$ are obtained as

$$\widehat{\boldsymbol{V}}_j = (1 - \widehat{\rho}_j)\mathbf{I}_p + \widehat{\rho}_j \mathbf{1}_p \mathbf{1}_p'.$$

$$\text{and}\quad \widehat{\boldsymbol{\Delta}}_j = \mathbf{I}_u\left(\widehat{\sigma}_{j,0}^2 - \widehat{\sigma}_{j,1}^2\right) + \mathbf{J}_u \widehat{\sigma}_{j,1}^2.$$

Acknowledgements: The first author was partially supported by the College of Business Summer Research Grant at the University of Texas at San Antonio.

References

[1] Gupta, A.K., 1986. On a Classification Rule for Multiple Measurements. Comput. Math. Applic. 12A, 301-308.

[2] Gupta, A.K., 1980. On a Multivariate Statistical Classification Model. Multivariate Statistical Analysis, R. P. Gupta (Ed.). Amsterdam: North-Holland, 83-93.

[3] Hand, D.J., 1997. Construction and Assessment of Classification Rules. Wiley, England.

[4] Harville, D.A., 1997. Matrix Algebra from Statistician's Perspective, Springer-Verlag, New York.

[5] Self, S.G., Liang, K., 1987. Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions. J. Amer. Statis. Assoc. 82(398), 605-610.

[6] Roy, A., Khattree, R., 2007. Classification of Multivariate Repeated Measures Data with Temporal Autocorrelation, Journal of Applied Statistical Science 15(3), 283-294.

[7] Roy, A., Khattree, R., 2005 a. Classification Based on Multivariate Repeated Measures with Time Effect on Mean Vector and an AR(1) Correlation Structure on the Repeated Measures, Calcutta Statistical Assoc. Bulletin 57, 49-65.

[8] Roy, A., Khattree, R., 2005 b. On Discrimination and Classification with Multivariate repeated Measures Data, Journal of Statistical Planning and Inference 134(2), 462-485.