*Communications in Statistics—Theory and Methods*, 41: 1411–1420, 2012

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610926.2010.530369



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

ANURADHA ROY[1] AND RICARDO LEIVA[2]

[1]Department of Management Science and Statistics,
The University of Texas at San Antonio, San Antonio, Texas, USA
[2]Departamento de Matemática, F.C.E., Universidad Nacional de Cuyo,
Mendoza, Argentina

*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.*

**Keywords** Covariance structure; Maximum likelihood estimates; Repeated observations.

**Mathematics Subject Classification** Primary 62H30; Secondary 62H12.



## 1. Introduction

We develop classification rules for multivariate repeated measures data with structured correlations on repeated measurements on both spatial as well as over time. The available classification rules (Roy and Khattree, 2007) for multivariate repeated measures data consider structured correlation only on repeated measurements 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 measurements on spatial as well as over time would be natural or beneficial. These problems are computationally very challenging, as it is not possible to tract them analytically or find any closed-form solution.

Received February 28, 2010; Accepted October 4, 2010
Address correspondence to 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

Classification problems 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 (2007) considered the problem in small sample situation by assuming Kronecker product structure on a variance–covariance matrix. They assumed equicorrelated or compound symmetry as well as autoregressive of order one (AR(1)) correlation structure on the repeated measurements on a single variable and an unstructured covariance matrix on the variables at a particular time point. As mentioned before in many clinical trial problems, it is found that the measurements on a single variable are 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 over time may diagnose a patient with Alzheimer's disease. Another example is 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 anatomical locations of the body, such as the spine, radius, femoral neck, and the total hip, and all the measurements are observed repeatedly over time. In this article, we develop classification rules for these kinds of data. Different time points ($p$) as well as different sites ($u$) may have different measurement variations for the variable, and we should take these variations into account while analyzing these kinds of data. It is well known (Harville, 1997) that the correlation structure on the repeated measurements follows a simple pattern such as compound symmetry or a first-order autoregressive AR(1) structure as opposed to the unstructured variance–covariance matrix, where the mean vectors, and the variances and covariances among the $pu$ measurements are arbitrary. Therefore, for both data sets it is expected that both measurement variations over time as well as over sites will have patterned covariance structures. In other words, marginal variance–covariance matrices over different time points as well as over different sites will have patterned covariance structures. In this article, we develop classification rules for multivariate repeated measures data for two classes where both the marginal variance–covariance matrices over different time points as well as over different sites 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, $r = 1, \ldots, n_j$, $s = 1, \ldots, u$, $t = 1, \ldots, p$, $j = 1, 2$. 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 $\boldsymbol{y}_{jr} = (\boldsymbol{y}'_{jr,1}, \boldsymbol{y}'_{jr,2}, \ldots, \boldsymbol{y}'_{jr,p})'$ be the $pu$-variate vector of all measurements corresponding to the $r$th individual in the $j$th population. For two populations let $\boldsymbol{Y}_j = [\boldsymbol{y}_{j1}, \boldsymbol{y}_{j2}, \ldots, \boldsymbol{y}_{jn_j}]$ be $n_j$ independent random samples from populations $N_{pu}(\boldsymbol{\mu}_j, \boldsymbol{\Omega}_j)$, where $\boldsymbol{\mu}_j \in \mathbb{R}^{pu}$ and the matrix $\boldsymbol{\Omega}_j$ is assumed to be $pu \times pu$ positive definite matrix. When $\boldsymbol{\Omega}_j$ is unknown and completely unspecified, a total of $pu(pu + 1)/2$ unknown parameters must be estimated for any statistical inference. This number increases very rapidly with $p$ and $u$. Estimation of so many parameters will require a very large sample ($n_j > pu$), which may not always be feasible for many biomedical researchers. We circumvent this problem by imposing some structure on $\mathbf{\Omega}_j$. We assume the form of the covariance matrix $\mathbf{\Omega}_j$ as

$$\underset{pu\times pu}{\mathbf{\Omega}_j} = \underset{p\times p}{\boldsymbol{V}_j} \otimes \underset{u\times u}{\mathbf{\Delta}_j}, \tag{1}$$

where $\boldsymbol{V}_j$ and $\mathbf{\Delta}_j$, respectively, are $p \times p$ and $u \times u$ positive definite matrices and $\otimes$ represents the Kronecker product. Several authors (Boik, 1991; Chaganty and Naik, 2002; Galecki, 1994; Huizenga et al., 2002; Mardia and Goodall, 1993; Naik and Rao, 2001; Roy, 2006a,b; Shults, 2000) have observed many advantages of using this Kronecker product (separable) structure over usual unstructured variance–covariance matrix for analyzing multivariate repeated measures data to model the dependence among spatio-temporal repeated measurements.

Note that if $\mathbf{\Omega}_j = \boldsymbol{V}_j \otimes \mathbf{\Delta}_j$, then $\mathbf{\Omega}_j = \left(\frac{1}{\alpha}\boldsymbol{V}_j\right) \otimes \left(\alpha\mathbf{\Delta}_j\right)$, for any non zero real number $\alpha$, and therefore parameters in each matrix $\boldsymbol{V}_j$ and $\mathbf{\Delta}_j$ are not jointly identifiable unless we impose an appropriate condition. There are several possible ways to handle this. It is always possible to obtain an estimate of either of them by taking one of the diagonal elements of either of the component matrices $\boldsymbol{V}_j$ and $\mathbf{\Delta}_j$ to be one. Normally, the first or the last diagonal element of $\boldsymbol{V}_j$ is taken to be one. SAS *PROC MIXED* (SAS Institute, 2004) always takes the first diagonal element of $\boldsymbol{V}_j$ to be one when both $\boldsymbol{V}_j$ and $\mathbf{\Sigma}_j$ are unstructured covariance matrices.

As mentioned before, the repeated measurements on the same unit or subject are correlated on each other and it is often modeled through patterned covariance structures such as autoregressive of order one AR(1), antedependence, ARMA, or equicorrelated covariance structure. Among these structures equicorrelation is a commonly applied assumption in the analyses of many medical and biomedical univariate repeated measures data. We thus assume both $\boldsymbol{V}_j$ and $\mathbf{\Delta}_j$ in (1) have equicorrelated or compound symmetry structures. However, this may result in an identifiability problem as discussed above. We evade this problem by taking $\boldsymbol{V}_j$ as an equicorrelated correlation structure so that all diagonal elements of it are one, and $\mathbf{\Delta}_j$ as an equicorrelated covariance structure. The matrix $\mathbf{\Delta}_j$ has the form

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

where $\mathbf{I}_u$ is the $u \times u$ identity matrix, $\mathbf{1}_u$ is an $u \times 1$ vector containing all elements as unity, $\mathbf{J}_u = \mathbf{1}_u\mathbf{1}'_u$, and the correlation matrix $\boldsymbol{V}_j$, $j = 1, 2$ has the form

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

In this way, $\mathbf{\Omega}_j$ inherits the benefit of both structures (equicorrelation and separability) yields an enormous reduction in the number of parameters. Also note that the matrix $\mathbf{\Omega}_j$ has the form

$$\begin{aligned}\mathbf{\Omega}_j &= \boldsymbol{V}_j \otimes \mathbf{\Delta}_j \\ &= \mathbf{I}_p \otimes \left(\mathbf{\Gamma}_{j,0} - \mathbf{\Gamma}_{j,1}\right) + \mathbf{J}_p \otimes \mathbf{\Gamma}_{j,1}\end{aligned}$$

$$= \begin{bmatrix} \boldsymbol{\Gamma}_{j,0} & \boldsymbol{\Gamma}_{j,1} & \cdots & \boldsymbol{\Gamma}_{j,1} \\ \boldsymbol{\Gamma}_{j,1} & \boldsymbol{\Gamma}_{j,0} & \cdots & \boldsymbol{\Gamma}_{j,1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Gamma}_{j,1} & \boldsymbol{\Gamma}_{j,1} & \cdots & \boldsymbol{\Gamma}_{j,0} \end{bmatrix},$$

where

$$\boldsymbol{\Gamma}_{j,0} = \boldsymbol{\Delta}_j,$$

and

$$\boldsymbol{\Gamma}_{j,1} = \rho_j \boldsymbol{\Delta}_j.$$

That is, $\boldsymbol{\Omega}_j = \boldsymbol{V}_j \otimes \boldsymbol{\Delta}_j$ is a particular case of the blocked compound symmetry (BCS) structure. Hypothesis testing on this BCS structure is discussed in Roy and Leiva (2011). It is clear that $\boldsymbol{\Gamma}_{j,0}$ models the dependence among the components in each vector $\boldsymbol{y}_{jr,t}$, and $\boldsymbol{\Gamma}_{j,1}$ models the dependence among the components of different vectors $\boldsymbol{y}_{jr,t}$ and $\boldsymbol{y}_{jr,t^*}$ with $t \neq t^*$.

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

$$\boldsymbol{\Delta}_j^{-1} = h_j \mathbf{I}_u + k_j \mathbf{J}_u, \tag{2}$$

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

$$\left|\boldsymbol{\Delta}_j\right| = \left|\sigma_{j,1}^2 - \sigma_{j,0}^2\right|^{u-1} \left|\sigma_{j,0}^2 + (u-1)\sigma_{j,1}^2\right|. \tag{3}$$

Now, the elements $v_i^{lm}$ of $\boldsymbol{V}_j^{-1}$ are given by

$$v_j^{lm} = \begin{cases} \dfrac{1+(p-2)\rho_j}{(1-\rho_j)\{1+(p-1)\rho_j\}}, & \text{if } l = m, \\ -\dfrac{\rho_j}{(1-\rho_j)\{1+(p-1)\rho_j\}}, & \text{if } l \neq m, \end{cases} \tag{4}$$

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

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

Since $V_j$ has to be positive, we should have $-\frac{1}{p-1} < \rho_j < 1$. However, we further assume that $0 < \rho_j < 1$.

## 2. Classification Rules

*Case* 1. $\boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_2$ ($V_1 = V_2, \boldsymbol{\Delta}_1 = \boldsymbol{\Delta}_2$).

Sample classification rule is given by:

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

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

and to Population 2 otherwise.

*Maximum likelihood estimation (MLE) of* $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{V}$, and $\boldsymbol{\Delta}$. Let $n = n_1 + n_2$ be the total number of random samples $\boldsymbol{Y}_j = [\boldsymbol{y}_{j1}, \boldsymbol{y}_{j2}, \ldots, \boldsymbol{y}_{jn_j}]$ from Population $j$, $j = 1, 2$. Here we assume $\boldsymbol{\mu}_j = (\mu_{j,ts})'_{t=1,\ldots,p;s=1,\ldots,u}$. Using (2) and(3) the log likelihood function $\ln L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{V}, \boldsymbol{\Delta}; \boldsymbol{Y}_1, \boldsymbol{Y}_2)$ is given by

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

An alternative expression for $\ln L$ is

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

where

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

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

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

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

$$\ln L = -\frac{npu}{2}\ln(2\pi) - \frac{n(p-1)u}{2}\ln(1-\rho) - \frac{nu}{2}\ln\{1 + (p-1)\rho\}$$

$$-\frac{np}{2}\ln|\boldsymbol{\Delta}| - \frac{1}{2(1-\rho)}c_1^* + \frac{\rho}{2(1-\rho)\{1+(p-1)\rho\}}d_1^*, \tag{6}$$

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

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

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

$$\begin{aligned}h &= \frac{1}{\sigma_0^2 - \sigma_1^2}, \\ m &= \frac{1}{\sigma_0^2 + (u-1)\,\sigma_1^2}, \\ b_{1,1}^* &= \sum_{j=1}^{2}\sum_{r=1}^{n_j}\sum_{m=1}^{p}\sum_{l=1}^{p}\sum_{s=1}^{u} v^{lm}\left(y_{jr,ls} - \bar{y}_{j\cdot,ls}\right)\left(y_{jr,ms} - \bar{y}_{j\cdot,ms}\right)', \quad \text{and} \\ b_{1,2}^* &= \sum_{j=1}^{2}\sum_{r=1}^{n_j}\sum_{m=1}^{p}\sum_{l=1}^{p}\sum_{s=1}^{u}\sum_{s^*=1}^{u} v^{lm}\left(y_{jr,ls} - \bar{y}_{j\cdot,ls}\right)\left(y_{jr,ms^*} - \bar{y}_{j\cdot,ms^*}\right).\end{aligned}$$

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

$$\begin{aligned}\widehat{h^{-1}} &= \frac{1}{np\,(u-1)}\left(b_{1,1}^* - \frac{1}{u}b_{1,2}^*\right), \quad \text{and} \\ \widehat{m^{-1}} &= \frac{1}{np}b_{1,2}^*.\end{aligned}$$

After some simplifications, we get

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

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

The MLEs $\hat{\rho}$, $\hat{\sigma}_0^2$ and $\hat{\sigma}_1^2$ are obtained by simultaneously and iteratively solving (7)–(9), by substituting the values of $v^{lm}$; $l, m = 1, 2, \ldots, p$, from Eq. (4). The computations can be carried out by the following algorithm. The MLE of $\boldsymbol{V}$ is obtained from

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

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

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

Algorithm outline:

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

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

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

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

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

**Step 6.** Repeat Steps 2–5 until convergence is attained. This is ensured by verifying that the maximum of the absolute difference between two successive values of $\hat{\rho}$, $\hat{\sigma}_0^2$, and $\hat{\sigma}_1^2$ is less than $\epsilon$. Even though $\rho$ is always between $-\frac{1}{p-1}$ and 1, we have assumed $0 < \rho < 1$. Still, $\hat{\rho}$ 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. $\boldsymbol{\Omega}_1 \neq \boldsymbol{\Omega}_2$ ($\boldsymbol{V}_1 \neq \boldsymbol{V}_2, \boldsymbol{\Delta}_1 \neq \boldsymbol{\Delta}_2$).
Sample classification rule is given by:
Classify an individual with response $\boldsymbol{y}$ to Population 1 if

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

and to Population 2 otherwise.

*Maximum likelihood estimation (MLE) of* $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{V}_1$, $\boldsymbol{V}_2$, $\boldsymbol{\Delta}_1$, and $\boldsymbol{\Delta}_2$: As before let $n_j$ random samples $\boldsymbol{Y}_j = [\boldsymbol{y}_{j1}, \boldsymbol{y}_{j2}, \ldots, \boldsymbol{y}_{jn_j}]$ be drawn from Population $j$, $j = 1, 2$. Here, we assume $\boldsymbol{\mu}_j = (\mu_{j,ts})'_{t=1,\ldots,p;s=1,\ldots,u}$. Using (2) and (3) the the log likelihood function $\ln L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2; \boldsymbol{Y}_1, \boldsymbol{Y}_2)$ is given by:

$$\begin{aligned}\ln L = &-\frac{npu}{2}\ln 2\pi - \frac{n_1(p-1)u}{2}\ln(1-\rho_1)\\ &-\frac{n_2(p-1)u}{2}\ln(1-\rho_2) - \frac{n_1u}{2}\ln\{1+(p-1)\rho_1\}\\ &-\frac{n_2u}{2}\ln\{1+(p-1)\rho_2\} - \frac{n_1p}{2}\ln|\boldsymbol{\Delta}_1|\end{aligned}$$

$$
- \frac{n_2 p}{2} \ln|\boldsymbol{\Delta}_2| - \frac{1}{2(1-\rho_1)} c_1 - \frac{1}{2(1-\rho_2)} c_2
+ \frac{\rho_1}{2(1-\rho_1)\{1+(p-1)\rho_1\}} d_1 + \frac{\rho_2}{2(1-\rho_2)\{1+(p-1)\rho_2\}} d_2, \qquad (12)
$$

where $c_j = \mathrm{tr}[(\mathbf{I}_p \otimes \boldsymbol{\Delta}_j^{-1})\boldsymbol{S}_j]$ 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:

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

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

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

where

$$
h_j = \frac{1}{\sigma_{j,0}^2 - \sigma_{j,1}^2},
$$

$$
m_j = \frac{1}{\sigma_{j,0}^2 + (u-1)\sigma_{j,1}^2},
$$

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

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

After some algebraic simplification from (14), we get

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

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

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

After simplification, we get

$$\hat{\sigma}_{j,0}^2 = \frac{b_{j,1}}{n_j pu}, \quad j = 1, 2, \quad \text{and} \tag{15}$$

$$\hat{\sigma}_{j,1}^2 = \frac{b_{j,2} - b_{j,1}}{n_j pu\,(u-1)}, \quad j = 1, 2. \tag{16}$$

The maximum likelihood estimates $\hat{\rho}_1, \hat{\rho}_2, \hat{\sigma}_{10}^2, \hat{\sigma}_{11}^2, \hat{\sigma}_{20}^2$, and $\hat{\sigma}_{21}^2$ 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

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

## Acknowledgment

The first author would like to thank the College of Business at the University of Texas at San Antonio for their continued support with this publication.

## References

Boik, J. B. (1991). Scheff's mixed model for multivariate repeated measures: A relative efficiency evaluation. *Commun. Statist. Theor. Meth.* 20:1233–1255.

Chaganty, N. R., Naik, D. N. (2002). Analysis of multivariate longitudinal data using quasi-least squares. *J. Statist. Plann. Infer.* 103:421–436.

Galecki, A. T. (1994). General class of covariance structures for two or more repeated factors in longitudinal data analysis. *Commun. Statist. Theor. Meth.* 22:3105–3120.

Gupta, A. K. (1980). On a multivariate statistical classification model. In: Gupta, R. P., ed. *Multivariate Statistical Analysis*. Amsterdam: North-Holland, pp. 83–93.

Gupta, A. K. (1986). On a classification rule for multiple measurements. *Comput. Math. Applic.* 12A:301–308.

Hand, D. J. (1997). *Construction and Assessment of Classification Rules*. Chichester, England: Wiley.

Harville, D. A. (1997). *Matrix Algebra from Statistician's Perspective*. New York: Springer-Verlag.

Huizenga, H., Munck, J., Waldorp, L., Grasman, R. (2002). Spatio-temporal EEG/MEG source analysis based on a parametric noise covariance model. *IEEE Trans. Biomed. Eng.* 49:533–539.

Mardia, K., Goodall, C. (1993). Spatial-temporal analysis of multivariate environmental data. In: Patil, G., Rao, C., eds. *Multivariate Environmental Statistics*. Amsterdam: Elsevier, pp. 347–386.

Naik, D. N., Rao, S. S. (2001). Analysis of multivariate repeated measures data with a kronecker product structured covariance matrix. *J. Appl. Statist.* 28:91–105.

Roy, A. (2006a). Estimating correlation coefficient between two variables with repeated observations using mixed effects model. *Biom. J.* 48:286–301.

Roy, A. (2006b). A new classification rule for incomplete doubly multivariate data using mixed effects model with performance comparisons on the imputed data. *Statist. Med.* 25:1715–1728.

Roy, A., Khattree, R. (2007). Classification rules for repeated measures data from biomedical research. In: Khattree, R., Naik, D. N., eds. *Computational Methods in Biomedical Research*. Boca Raton, FL: Chapman & Hall/CRC, pp. 323–370.

Roy, A., Leiva, R. (2011). Estimating and testing a structured covariance matrix for three-level multivariate data. *Commun. Statist. Theor. Meth.* 40:1945–1963.

SAS Institute. (2004). *SAS/STAT User's Guide Version 9*. Cary, NC: SAS Institute Inc.

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.

Shults, J. (2000). Modeling the correlation structure of data that have multiple levels of association. *Commun. Statist. Theor. Meth.* 29:1005–1015.