# Canonical Correlation Analysis of Longitudinal Data

Jayesh Srivastava* Dayanand N. Naik†

**Abstract**

Studying the relationship between two sets of variables is an important multivariate statistical analysis problem in statistics. Canonical correlation coefficients are used to study these relationships. Canonical correlation analysis (CCA) is a general multivariate method that is mainly used to study relationships when both sets of variables are quantitative. In this paper, we have generalized CCA to analyze the relationships between two sets of repeatedly or longitudinally observed data using a block Kronecker product matrix to model dependency of the variables over time. We then apply canonical correlation analysis on this matrix to obtain canonical correlations and canonical variables.

**Key Words:** Canonical correlation analysis, Kronecker product, longitudinal data.

## 1. Introduction

Canonical correlation analysis (CCA) is a well known statistical technique used to identify and measure the association between two sets of random vectors using specific matrix functions of variance-covariance matrices of these variables. This is also one of the most general methods for data reduction in multivariate analysis. CCA was introduced by Hotelling (1936) while studying the relationship between two sets of variables in instructional research. Now CCA has found many applications in different fields and it is routinely discussed in many multivariate statistical analysis textbook. For example, see Johnson and Wichern (2002), Khattree and Naik (2000), or Mardia, Kent and Bibby (1979).

Suppose $\mathbf{\Sigma}_{xx}$ is the variance-covariance matrix the $p \times 1$ random vector $\mathbf{x}$ and $\mathbf{\Sigma}_{yy}$ is that of $q \times 1$ random vector $\mathbf{y}$. Also suppose the covariance between the vectors $\mathbf{x}$ and $\mathbf{y}$ is given by $\mathbf{\Sigma}_{xy} = \text{cov}(\mathbf{x}, \mathbf{y})$. Then we can write the variance covariance matrix of $(\mathbf{y},\ \mathbf{x})$ as

$$\begin{bmatrix} \mathbf{\Sigma}_{yy} & \mathbf{\Sigma}_{yx} \\ \mathbf{\Sigma}_{xy} & \mathbf{\Sigma}_{xx} \end{bmatrix}.$$

The main idea behind canonical correlation analysis is to find a $q \times 1$ vector $\mathbf{a}$ and a $p \times 1$ vector $\mathbf{b}$, given $\mathbf{\Sigma}_{xx}$, $\mathbf{\Sigma}_{yy}$ and $\mathbf{\Sigma}_{xy}$, so that the correlation between $\mathbf{a}'\mathbf{y}$ and $\mathbf{b}'\mathbf{x}$ is maximized.

The $i^{th}$ pair of canonical variables $(\mathbf{a}_i'\mathbf{y}, \mathbf{b}_i'\mathbf{x})$ is obtained by solving

$$\begin{aligned} \mathbf{\Sigma}_{yy}^{-1}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}\mathbf{a}_i &= \rho_i^2\mathbf{a}_i & (1) \\ \text{and } \mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1}\mathbf{\Sigma}_{yx}\mathbf{b}_i &= \rho_i^2\mathbf{b}_i, & (2) \end{aligned}$$

where $\rho_i$ is a canonical correlation and $\rho_i^2$ is eigenvalue of

$$\mathbf{\Sigma}_{xx}^{-1/2}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1/2} \text{ or } \mathbf{\Sigma}_{yy}^{-1/2}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1/2}.$$

Kettenring (1971) has generalized CCA to more than two sets of variables; other generalizations of the method can also be found in the literature. Beaghen (1997) has used canonical variate method to analyze the means of longitudinal data. However, no methods have been found in the literature to perform CCA on longitudinally observed data. Focus in this paper is to generalize canonical correlation analysis of repeatedly observed $\mathbf{x} = (x_1, ..., x_p)'$ and $\mathbf{y} = (y_1, ..., y_q)'$.

## 2. Repeated Canonical Correlation Analysis

Suppose we have observed $\mathbf{x}$ and $\mathbf{y}$ repeatedly over $t$ time periods. Let $\mathbf{x}_i$ and $\mathbf{y}_i$ be the vectors $\mathbf{y}$ and $\mathbf{x}$ observed at the $i^{th}$ occasion. Define $\mathcal{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_t')'$ and $\mathcal{X} = (\mathbf{x}_1', \ldots, \mathbf{x}_t')'$.

To model the dependency of repeated measures we assume that the variance covariance matrix $\mathbf{D}$ of $\mathbf{u} = (\mathcal{Y}',\ \mathcal{X}')'$ has a Kronecker product matrix structure. That is,

$$\mathbf{D} = \begin{bmatrix} \mathbf{\Omega}_{yy} \otimes \mathbf{\Sigma}_{yy} & \mathbf{\Omega}_{yx} \otimes \mathbf{\Sigma}_{yx} \\ \mathbf{\Omega}_{xy} \otimes \mathbf{\Sigma}_{xy} & \mathbf{\Omega}_{xx} \otimes \mathbf{\Sigma}_{xx} \end{bmatrix}. \quad (3)$$

[*] Statistician, Product Management, High Point Insurance, New Jersey

[†] Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA 23529

The matrices $\mathbf{\Omega}_{yy}$, $\mathbf{\Omega}_{xx}$ and $\mathbf{\Omega}_{yx}$ are used to model the dependency over $t$ time periods of repeated measurements on $\mathbf{y}$, on $\mathbf{x}$ and of the covariance matrix between repeated measurements of $\mathbf{y}$ and $\mathbf{x}$ respectively. Kronecker product structures have been successfully utilized to analyze multivariate normal repeated measures data in Naik and Rao (2001), Roy and Khattree (2005), Srivastava, Nahtman and von Rosen (2007), and Srivastava, Nahtman and von Rosen (2008).

The problem here is to determine linear functions $U = \mathbf{a}'\mathcal{Y}$ and $V = \mathbf{b}'\mathcal{X}$ such that the correlation between them is maximum. Here $\mathbf{a}$ is $qt \times 1$ and $\mathbf{b}$ is $pt \times 1$ vectors. Assuming $E(\mathcal{Y}) = 0$, $E(\mathcal{X}) = 0$ and restricting $U$ and $V$ to have unit variances, i.e

$$E(U^2) = 1 \Rightarrow \mathbf{a}'E(\mathcal{Y}'\mathcal{Y})\mathbf{a} = \mathbf{a}'\mathbf{\Omega}_{yy} \otimes \mathbf{\Sigma}_{yy}\mathbf{a} = 1 \tag{4}$$

$$E(V^2) = 1 \Rightarrow \mathbf{b}'E(\mathcal{X}'\mathcal{X})\mathbf{b} = \mathbf{b}'\mathbf{\Omega}_{xx} \otimes \mathbf{\Sigma}_{xx}\mathbf{b} = 1, \tag{5}$$

the correlation between U and V is given by

$$E(UV) = E(\mathbf{a}'\mathcal{Y}\mathcal{X}'\mathbf{b}) = \mathbf{a}'E(\mathcal{Y}\mathcal{X}')\mathbf{b} = \mathbf{a}'\mathbf{\Omega}_{yx} \otimes \mathbf{\Sigma}_{yx}\mathbf{b}. \tag{6}$$

Thus the algebraic problem is to find $\mathbf{a}$ and $\mathbf{b}$ to maximize (6) subject to the conditions (4) and (5). Solution to this problem can be easily obtained using Lagrangian multipliers method and vectors $\mathbf{a}$ and $\mathbf{b}$ are obtained by solving the equations:

$$\Big((\mathbf{\Omega}_{yx} \otimes \mathbf{\Sigma}_{yx})(\mathbf{\Omega}_{xx} \otimes \mathbf{\Sigma}_{xx})^{-1}(\mathbf{\Omega}_{xy} \otimes \mathbf{\Sigma}_{xy}) - \lambda^2(\mathbf{\Omega}_{yy} \otimes \mathbf{\Sigma}_{yy})\Big)\mathbf{a} = 0 \tag{7}$$

and

$$\Big((\mathbf{\Omega}_{xy} \otimes \mathbf{\Sigma}_{xy})(\mathbf{\Omega}_{yy} \otimes \mathbf{\Sigma}_{yy})^{-1}(\mathbf{\Omega}_{yx} \otimes \mathbf{\Sigma}_{yx}) - \lambda^2(\mathbf{\Omega}_{xx} \otimes \mathbf{\Sigma}_{xx})\Big)\mathbf{b} = 0. \tag{8}$$

From equations (7) and (8) it is clear that $\lambda^2$ is an eigenvalue of

$$(\mathbf{\Omega}_{yy}^{-1/2}\mathbf{\Omega}_{yx}\mathbf{\Omega}_{xx}^{-1}\mathbf{\Omega}_{xy}\mathbf{\Omega}_{yy}^{-1/2}) \otimes (\mathbf{\Sigma}_{yy}^{-1/2}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1/2}).$$

and of

$$(\mathbf{\Omega}_{xx}^{-1/2}\mathbf{\Omega}_{xy}\mathbf{\Omega}_{yy}^{-1}\mathbf{\Omega}_{yx}\mathbf{\Omega}_{xx}^{-1/2}) \otimes (\mathbf{\Sigma}_{xx}^{-1/2}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1/2}).$$

In general the vectors $\mathbf{a}_i$ and $\mathbf{b}_i$, such that $(\mathbf{a}_i'\mathcal{Y}, \mathbf{b}_i'\mathcal{X})$ is the $i^{th}$ pair of canonical variables, are obtained as the solutions to

$$(\mathbf{\Omega}_{yy}^{-1/2}\mathbf{\Omega}_{yx}\mathbf{\Omega}_{xx}^{-1}\mathbf{\Omega}_{xy}\mathbf{\Omega}_{yy}^{-1/2}) \otimes (\mathbf{\Sigma}_{yy}^{-1/2}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1/2})\mathbf{a}_i = \lambda_i^2\mathbf{a}_i$$

and

$$(\mathbf{\Omega}_{xx}^{-1/2}\mathbf{\Omega}_{xy}\mathbf{\Omega}_{yy}^{-1}\mathbf{\Omega}_{yx}\mathbf{\Omega}_{xx}^{-1/2}) \otimes (\mathbf{\Sigma}_{xx}^{-1/2}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1/2})\mathbf{b}_i = \lambda_i^2\mathbf{b}_i.$$

Suppose $\boldsymbol{\lambda}^{\mathbf{2}}$ is the vector of eigenvalues $\lambda_i$. Then using the properties of Kronecker product we have the following: $\boldsymbol{\lambda}^{\mathbf{2}} = \boldsymbol{\lambda}_{\mathbf{\Omega}}^2 \otimes \boldsymbol{\lambda}_{\mathbf{\Sigma}}^2$, where $\boldsymbol{\lambda}_{\mathbf{\Omega}}^2$ and $\boldsymbol{\lambda}_{\mathbf{\Sigma}}^2$ are the vectors of eigenvalues of $(\mathbf{\Omega}_{xx}^{-1/2}\mathbf{\Omega}_{xy}\mathbf{\Omega}_{yy}^{-1}\mathbf{\Omega}_{yx}\mathbf{\Omega}_{xx}^{-1/2})$ and $(\mathbf{\Sigma}_{yy}^{-1/2}\mathbf{\Sigma}_{yx}\mathbf{\Sigma}_{xx}^{-1}\mathbf{\Sigma}_{xy}\mathbf{\Sigma}_{yy}^{-1/2})$ respectively. It is interesting to note that the canonical correlations for repeated measures data are the scaled versions of the canonical correlations in the usual case. However, the scaling is by the square root of the eigenvalues of the repeated effect matrix $(\mathbf{\Omega}_{yy}^{-1/2}\mathbf{\Omega}_{yx}\mathbf{\Omega}_{xx}^{-1}\mathbf{\Omega}_{xy}\mathbf{\Omega}_{yy}^{-1/2})$.

Notice that if there is no repeated effect (that is, $\mathbf{\Omega}_{ij} = I$, for $i, j = x, y$) or all the repeated effect is same (that is, $\mathbf{\Omega}_{ij} = \mathbf{\Omega}$) then

$$(\mathbf{\Omega}_{xx}^{-1/2}\mathbf{\Omega}_{xy}\mathbf{\Omega}_{yy}^{-1}\mathbf{\Omega}_{yx}\mathbf{\Omega}_{xx}^{-1/2}) = \omega\mathbf{I}_{tt}$$

and $\boldsymbol{\lambda}_{\mathbf{\Omega}}^2 = \omega\mathbf{1_t}$, where $\omega$ is a positive constant.

## 3. Estimation and Hypothesis Testing

Usually the matrices $\mathbf{\Sigma}_{yy}$, $\mathbf{\Sigma}_{yx}$, $\mathbf{\Sigma}_{xx}$, $\mathbf{\Omega}_{yy}$, $\mathbf{\Omega}_{yx}$ and $\mathbf{\Omega}_{xx}$ are not known and need to be estimated from the data. The population canonical correlation will be estimated by the sample canonical correlations. Let us assume that $\mathbf{u} = (\mathcal{Y}',\ \mathcal{X}')'$ is distributed as multivariate normal with mean vector $\boldsymbol{\mu}$ and variance covariance matrix $\mathbf{D}$.

Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be the random sample from the $N(\boldsymbol{\mu},\ \mathbf{D})$ where variance-covariance matrix $\mathbf{D}$ is given by equation 3.

The log-likelihood function of the parameters given the observed data is

$$L(\boldsymbol{\mu}, \mathbf{D}) = -0.5(n\ log(|\mathbf{D}|) + \sum_{i=1}^{n}(\mathbf{u}_i - \boldsymbol{\mu})'\mathbf{D}^{-1}(\mathbf{u}_i - \boldsymbol{\mu})). \tag{9}$$

The estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\mathbf{D}}$ can be obtained by maximizing the above log-likelihood function. In the context of analysis of multivariate repeated measures data, Naik and Rao (2001) have provided the maximum likelihood estimates. Srivastava, Nahtman and von Rosen (2007), and Srivastava, Nahtman and von Rosen (2008) have provided proofs that the maximum likelihood estimates exist and are unique. We used SAS' non linear optimization routine for maximizing the log-likelihood function. Suppose $\hat{\boldsymbol{\Omega}}_{yy}$, $\hat{\boldsymbol{\Omega}}_{yx}$, and $\hat{\boldsymbol{\Omega}}_{xx}$ are the maximum likelihood estimates of $\boldsymbol{\Omega}_{yy}$, $\boldsymbol{\Omega}_{yx}$, and $\boldsymbol{\Omega}_{xx}$ respectively and $\hat{\boldsymbol{\Sigma}}_{yy}$, $\hat{\boldsymbol{\Sigma}}_{yx}$, and $\hat{\boldsymbol{\Sigma}}_{xx}$ are the maximum likelihood estimates of $\boldsymbol{\Sigma}_{yy}$, $\boldsymbol{\Sigma}_{yx}$, and $\boldsymbol{\Sigma}_{xx}$ respectively.

Then the sample canonical correlations $\hat{\lambda}_i^2$ are obtained as the positive square roots of the nonzero eigenvalues of

$$(\hat{\boldsymbol{\Omega}}_{yy}^{-1/2}\hat{\boldsymbol{\Omega}}_{yx}\hat{\boldsymbol{\Omega}}_{xx}^{-1}\hat{\boldsymbol{\Omega}}_{xy}\hat{\boldsymbol{\Omega}}_{yy}^{-1/2}) \otimes (\hat{\boldsymbol{\Sigma}}_{yy}^{-1/2}\hat{\boldsymbol{\Sigma}}_{yx}\hat{\boldsymbol{\Sigma}}_{xx}^{-1}\hat{\boldsymbol{\Sigma}}_{xy}\hat{\boldsymbol{\Sigma}}_{yy}^{-1/2}).$$

The vectors $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{b}_i}$ corresponding to $i^{th}$ pair of canonical variables are obtained as the solution of

$$(\hat{\boldsymbol{\Omega}}_{yy}^{-1/2}\hat{\boldsymbol{\Omega}}_{yx}\hat{\boldsymbol{\Omega}}_{xx}^{-1}\hat{\boldsymbol{\Omega}}_{xy}\hat{\boldsymbol{\Omega}}_{yy}^{-1/2}) \otimes (\hat{\boldsymbol{\Sigma}}_{yy}^{-1/2}\hat{\boldsymbol{\Sigma}}_{yx}\hat{\boldsymbol{\Sigma}}_{xx}^{-1}\hat{\boldsymbol{\Sigma}}_{xy}\hat{\boldsymbol{\Sigma}}_{yy}^{-1/2})\hat{\mathbf{a}}_i = \hat{\lambda}_i^2\hat{\mathbf{a}}_i$$

and

$$(\hat{\boldsymbol{\Omega}}_{xx}^{-1/2}\hat{\boldsymbol{\Omega}}_{xy}\hat{\boldsymbol{\Omega}}_{yy}^{-1}\hat{\boldsymbol{\Omega}}_{yx}\hat{\boldsymbol{\Omega}}_{xx}^{-1/2}) \otimes (\hat{\boldsymbol{\Sigma}}_{xx}^{-1/2}\hat{\boldsymbol{\Sigma}}_{xy}\hat{\boldsymbol{\Sigma}}_{yy}^{-1}\hat{\boldsymbol{\Sigma}}_{yx}\hat{\boldsymbol{\Sigma}}_{xx}^{-1/2})\hat{\mathbf{b}}_i = \hat{\lambda}_i^2\hat{\mathbf{b}}_i.$$

Before performing any canonical correlation analysis using the samples $\mathbf{u}_1, \ldots, \mathbf{u}_n$, the following hypotheses may be tested.

**1.** First test for the repeated effect on the variance covariance matrices of $\mathbf{y}$, $\mathbf{x}$ and on $cov(\mathbf{x}, \mathbf{y})$, i.e. test

$$H_0 : D = \begin{bmatrix} I_{yy} \otimes \boldsymbol{\Sigma}_{yy} & I_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ I_{xy} \otimes \boldsymbol{\Sigma}_{xy} & I_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix} \text{ vs } H_a : D = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix}$$

Note that the null hypothesis here specifies that the variance and covariance matrices do not change with the time factor. Here as well as in the cases that follow, the alternative hypothesis is assumed to be as in our assumed model, that it is unstructured Kronecker product block matrix. Testing can be performed using the likelihood ratio test (LRT). Maximizing the log-likelihood function $L(\boldsymbol{\mu}, \mathbf{D}) = -0.5(n\ log(|\mathbf{D}|) + \sum_{i=1}^{n}(\mathbf{u}_i - \boldsymbol{\mu})'\mathbf{D}^{-1}(\mathbf{u}_i - \boldsymbol{\mu}))$ under $H_0$ and $H_a$ will produce the maximum likelihood estimates. The likelihood ratio test statistic is then

$$-2log\Lambda = -2log(\ell_0/\ell_a),$$

where $\ell_0$ and $\ell_a$ denote the maximized likelihood functions under the null and alternative hypothesis. Under $H_0$, $-2log\Lambda$ has a chi-squared distribution, as $n \to \infty$. The degrees of freedom of the chi-square is the difference between the dimensions of the parameter spaces under $H_0 \cup H_a$ and under $H_0$.

**2.** If we accept $H_0$ then we can do the usual canonical correlation analysis by merging all the data. Otherwise we will test whether the effect of time (or the repeated effect) is on the covariances between ($\mathbf{x}$ and $\mathbf{y}$) only. This amounts to testing

$$H_{01} : D = \begin{bmatrix} I_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & I_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix} \text{ vs } H_a : D = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix}.$$

To test this hypothesis, $\ell_a$ is as in the previous case, i.e. as in (**1**) above. The maximum likelihood estimates and the maximum value of the likelihood function under $H_{01}$ can be obtained by maximizing (9) under $H_{01}$.

**3.** If we accept $H_{01}$, then we can perform canonical correlation analysis (CCA) using the estimated variance covariance matrix given under $H_{01}$ in (**2**) above. Otherwise we will test for repeated effect on variance covariance matrices of $\mathbf{y}$, $\mathbf{x}$, by testing,

$$H_{ox} : D = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & I_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix} \text{ vs } H_a : D = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix},$$

$$H_{oy} : D = \begin{bmatrix} I_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix} \text{ vs } H_a : D = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix}$$

As before the MLE and the value of the maximum likelihood function under $H_{ox}$ (and $H_{oy}$) can be obtained by maximizing (9) under the null hypothesis. Under $H_a$, the value $\ell_a$ remains the same.

**4.** If we accept $H_{ox}$ or $H_{oy}$ then we can perform canonical correlation analysis (CCA) using the corresponding estimated variance covariance matrix as in (**3**). Otherwise we will test for the same repeated effect, that is, test
$H_{tt} : D = \begin{bmatrix} \boldsymbol{\Omega}_{tt} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{tt} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{tt} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{tt} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix}$ vs $H_a : D = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix}$
The MLE of the common $\boldsymbol{\Omega}_{tt}$ and the other parameters can be obtained by maximizing (9) under $H_{tt}$ and in the same way as before the LRT can be constructed.

**5.** If we accept $H_{tt}$ then it suggest that change in variance covariance matrices over time is same and we should perform canonical correlation analysis (CCA) using the estimated structured variance covariance matrix as discussed in (**4**) above. Otherwise we should proceed with the general structured variance covariance matrix

$$\mathbf{D} = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix}.$$

## 4. A Simulation Example

In order to illustrate the analysis discussed here, we will work with simulated data. First we use the Helmert matrix to generate the positive definite matrices. The general form of a Helmert matrix $\mathbf{H}_k$ of order $k$ has $k^{-1/2}\mathbf{1}'_k$ for its first row, and each of its other $k-1$ rows for $i = 1, \ldots, k-1$ has the partitioned form

$$\left[ \begin{array}{c|c|c} \mathbf{1}'_i & -i & 0 \end{array} \right] / \sqrt{a_i}$$

with $a_i = i(i+1)$. A Helmert matrix is an orthogonal matrix, that is, $\mathbf{H}'\mathbf{H} = \mathbf{H}\mathbf{H}' = \mathbf{I}_k$. For example, the $4^{th}$ order Helmert matrix is given by

$$\mathbf{H}_4 = \begin{bmatrix} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{-3}{\sqrt{12}} \end{bmatrix}.$$

The spectral decomposition of a symmetric matrix, $\mathbf{A}$ is $\mathbf{A} = \sum \lambda_i \mathbf{u}_i \mathbf{u}'_i$, where the $\mathbf{u}_i$'s are the eigenvectors of $\mathbf{A}$. Now to generate a $k \times k$ positive definite matrix we take the $k^{th}$ order Helmert matrix, whose columns will give us the eigenvector of the desired matrix. Then choosing $k$ positive eigenvalues and using the spectral decomposition property we can construct the desired $k \times k$ positive definite matrix. We will use thus constructed positive definite matrix as $\boldsymbol{\Sigma}$.

Partitioning $\boldsymbol{\Sigma}$ will give

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{pmatrix},$$

and $\boldsymbol{\Sigma}_{yy}, \boldsymbol{\Sigma}_{xx}$ and $\boldsymbol{\Sigma}_{yx}$ can be used as variance-covariance matrix for $\mathbf{y}$, $\mathbf{x}$ and covariance matrix between $\mathbf{y}$ and $\mathbf{x}$ respectively. Then by choosing $t \times t$ modeling matrix $\boldsymbol{\Omega}_{yy}$ to associate with $\boldsymbol{\Sigma}_{yy}$, $\boldsymbol{\Omega}_{xx}$ with $\boldsymbol{\Sigma}_{xx}$ and $\boldsymbol{\Omega}_{yx}$ with $\boldsymbol{\Sigma}_{yx}$ we can construct the desired matrix

$$\mathbf{D} = \begin{bmatrix} \boldsymbol{\Omega}_{yy} \otimes \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Omega}_{yx} \otimes \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Omega}_{xy} \otimes \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Omega}_{xx} \otimes \boldsymbol{\Sigma}_{xx} \end{bmatrix}.$$

We can simulate any desired number of observations from the multivariate Normal $N(\mathbf{0}, \mathbf{D})$ and do repeated canonical correlation analysis on them as discussed earlier.

For our simulation example, we chose three $\mathbf{y}$ components, two $\mathbf{x}$ components and three repeated measurements, that is, $q = 3$, $p = 2$, and $t = 3$. A Helmert matrix of order 5 is chosen and used to determine a $5 \times 5$ positive definite variance covariance matrix $\boldsymbol{\Sigma}$. In addition, we picked 9.5262261, 8.7983733, 4.3901993, 2.2263795 and 1.9919697 as the eigenvalues which yields the following positive definite matrix $\boldsymbol{\Sigma}$ by the method described earlier.

$$\boldsymbol{\Sigma} = \begin{bmatrix} 5.3866296 & 2.1523738 & 0.3669639 & -1.729692 & 1.580125 \\ 2.1523738 & 5.3951715 & 1.9648395 & 1.3893513 & 1.0761869 \\ 0.3669639 & 1.9648395 & 4.7251901 & 0.2368742 & 0.183482 \\ -1.729692 & 1.3893513 & 0.2368742 & 8.4097148 & -0.864846 \\ 1.580125 & 1.0761869 & 0.183482 & -0.864846 & 3.016442 \end{bmatrix}.$$

**Table 1**: Hypothesis Testing

| **Hypothesis** | **Chi Square Test Statistics** | **Dof** | **p-value** |
|---|---|---|---|
| $H_0$ | 108.0483 | 3 | 0 |
| $H_{01}$ | 85.490536 | 2 | 0 |
| $H_{ox}$ | 53.496484 | 1 | 2.59E-13 |
| $H_{oy}$ | 44.505918 | 1 | 2.54E-11 |
| $H_{tt}$ | 4.3754035 | 2 | 0.1121743 |

**Table 2**: Canonical Correlation Estimates

| **Can. Corr.** | **Parameter** | **Estimate** | **Root MSE** | **Bias** |
|---|---|---|---|---|
| | $\rho$ | $\hat{\rho}$ | $\sqrt{E((\rho-\hat{\rho})^2)}$ | $\lvert(\rho-\hat{\rho})\rvert$ |
| $\rho_1$ | 0.237273404 | 0.2418999 | 0.025882426 | 0.004626013 |
| $\rho_2$ | 0.208651984 | 0.2124004 | 0.018033303 | 0.003754997 |
| $\rho_3$ | 0.177922968 | 0.1889365 | 0.020921281 | 0.011013628 |
| $\rho_4$ | 0.1591106 | 0.159546475 | 0.018627936 | 0.009110434 |
| $\rho_5$ | 0.14793029 | 0.1447519 | 0.016281892 | 0.00317805 |
| $\rho_6$ | 0.126144001 | 0.1247618 | 0.018398369 | 0.001382172 |

By partitioning $\mathbf{\Sigma}$ we get $\mathbf{\Sigma}_{yy}$, $\mathbf{\Sigma}_{xx}$ and $\mathbf{\Sigma}_{yx}$ as:

$$\mathbf{\Sigma}_{yy} = \begin{bmatrix} 5.3866296 & 2.1523738 & 0.3669639 \\ 2.1523738 & 5.3951715 & 1.9648395 \\ 0.3669639 & 1.9648395 & 4.7251901 \end{bmatrix},$$

$$\mathbf{\Sigma}_{xx} = \begin{bmatrix} 8.4097148 & -0.864846 \\ -0.864846 & 3.016442 \end{bmatrix} \text{ and } \mathbf{\Sigma}_{yx} = \begin{bmatrix} -1.729692 & 1.580125 \\ 1.3893513 & 1.0761869 \\ 0.2368742 & 0.183482 \end{bmatrix}.$$

We assume $AR(1)$ structure for repeated modeling matrices $\mathbf{\Omega}_{yy}$, $\mathbf{\Omega}_{xx}$, and $\mathbf{\Omega}_{yx}$ with correlation parameter $\rho_y = 0.1$, $\rho_x = 0.2$, and $\rho_{yx} = 0.1$ respectively. Arranging all the matrices together we have

$$\mathbf{D} = \begin{bmatrix} \mathbf{\Omega}_{yy} \otimes \mathbf{\Sigma}_{yy} & \mathbf{\Omega}_{yx} \otimes \mathbf{\Sigma}_{yx} \\ \mathbf{\Omega}_{xy} \otimes \mathbf{\Sigma}_{xy} & \mathbf{\Omega}_{xx} \otimes \mathbf{\Sigma}_{xx} \end{bmatrix}.$$

We simulated 500 observations from the multivariate normal ($N(\mathbf{0}, \mathbf{D})$) distribution and estimated the population parameters $\mathbf{\Sigma}_{yy}$, $\mathbf{\Sigma}_{xx}$, $\mathbf{\Sigma}_{yx}$, $\rho_y$, $\rho_x$, and $\rho_{yx}$. The estimates were found by maximizing the log-likelihood function using SAS *NLPQN* optimization routine.

## 5. Results and Discussion

To illustrate testing of various hypothesis discussed above, we used a set of data generated in the simulation. The chi-square test statistics and the asymptotic P-values for testing different hypothesis are shown in Table 1. As can be seen from the Table 1, all of the p-values are quite small except for the $H_{tt}$ hypothesis ($p - val = 0.1121743$). Thus all hypotheses except the $H_{tt}$ are rejected. In hypothesis $H_{tt}$ we are testing that the repeated effect is same on all components. In our simulation we have used the $AR(1)$ structure for the repeated correlation matrix with correlation parameters $\rho_y = 0.1$, $\rho_x = 0.2$, and $\rho_{yx} = 0.1$. Apparently these values are not very different to reject $H_{tt}$ using likelihood ratio test and this sample data. Although those values are not provided here to save space, when we chose quite different values for AR(1) parameters, the LRT did reject $H_{tt}$.

Next, in order to estimate the bias in estimating canonical correlations and other parameters, we repeated the simulation 5000 times and calculated the average values of all estimates. Table 3 shows the average of the parameter estimates based on these simulations. Table 2 presents the means the estimated canonical correlations. In the table, at the left of estimates we have provided true canonical correlation values calculated for the matrix used for simulation. In Table 2, minimum and maximum bias values are 0.001382172 and 0.011013628 respectively. Similarly in Table 3 biases ranges from $4.23009E - 05$ to 0.00409767. From both the tables it can be said that the estimates are very close to the true values.

**Table 3**: Maximum Likelihood Estimates, root MSE, and Bias

| **Pop. Para.** | **Para.** | **Estimate** | **Root MSE** | **Bias** |
|---|---|---|---|---|
| | $\theta$ | $\hat{\theta}$ | $\sqrt{E((\theta-\hat{\theta})^2)}$ | $\lvert(\theta-\hat{\theta})\rvert$ |
| $\Sigma_{yy}(1,1)$ | 5.38663 | 5.386232775 | 0.196693162 | 0.000397225 |
| $\Sigma_{yy}(2,2)$ | 5.39517 | 5.394642743 | 0.201315673 | 0.000527257 |
| $\Sigma_{yy}(3,3)$ | 4.72519 | 4.724343576 | 0.175820078 | 0.000846424 |
| $\Sigma_{yy}(1,2)$ | 2.15237 | 2.149807789 | 0.151351577 | 0.00256221 |
| $\Sigma_{yy}(1,3)$ | 0.36696 | 0.367644263 | 0.131625226 | 0.000684262 |
| $\Sigma_{yy}(2,3)$ | 1.96484 | 1.963606497 | 0.139320494 | 0.001233503 |
| $\Sigma_{xx}(1,1)$ | 8.40971 | 8.406462064 | 0.312531758 | 0.003247938 |
| $\Sigma_{xx}(2,2)$ | 3.01644 | 3.014373036 | 0.111293755 | 0.002066964 |
| $\Sigma_{xx}1,2$ | -0.86485 | -0.864892301 | 0.131922326 | 4.23009E-05 |
| $\Sigma_{yx}(1,1)$ | -1.72969 | -1.727133373 | 0.177219073 | 0.002556627 |
| $\Sigma_{yx}(1,2)$ | 1.58013 | 1.57732949 | 0.111941503 | 0.002800511 |
| $\Sigma_{yx}(2,1)$ | 1.38935 | 1.387468714 | 0.175657622 | 0.001881287 |
| $\Sigma_{yx}(2,2)$ | 1.07619 | 1.073755928 | 0.109225913 | 0.002434073 |
| $\Sigma_{yx}(3,1)$ | 0.23687 | 0.232772331 | 0.159496395 | 0.00409767 |
| $\Sigma_{yx}(3,2)$ | 0.18348 | 0.182754843 | 0.097614036 | 0.000725157 |
| $\rho_y$ | 0.1 | 0.100341533 | 0.024503061 | 0.000341533 |
| $\rho_x$ | 0.2 | 0.199774535 | 0.034666987 | 0.000225466 |
| $\rho_{yx}$ | 0.1 | 0.100673171 | 0.041535527 | 0.000673171 |

## 6. Concluding Remarks

In this paper, we have provided an easy to implement procedure to perform canonical correlation analysis of repeatedly observed data sets. To accommodate the effects of repeated measure we have adopted a Kronecker product structure to the variance covariance matrices. To account for the existence of repeated measure effects on different blocks of the variance covariance matrix, we have provided testing of different hypothesis. All of the procedures have been implemented on simulated data sets.

We have also proposed methods for performing correspondence analysis (CA) and canonical correspondence analysis (CCPA) of longitudinally observed data and those results will be reported elsewhere.

## References

Beaghen, M. (1997), "Canonical Variate Analysis and Related Methods with Longitudinal Data", *Ph.D. thesis, Department of Statistics, Virginia Tech.*

Hotelling, H. (1936), "Relations Between Two Sets of Variates", *Biometrika*, 28, 321-377.

Johnson, R. A., and Wichern D. W. (2002), *Applied Multivariate Statistical Analysis*, New Jersey: Prentice-Hall.

Kettenring, J. R. (1971), "Canonical Analysis of Several Sets of Variables", *Biometrika*, 58, 433-451.

Khattree, R., and Naik, D. N. (2000), *Multivariate Data Reduction and Discrimination with SAS Software.*, North Carolina: Wiley-SAS.

Mardia, K. V., Kent, J. J., & Bibby, J. M. (1979), *Multivariate Analysis*, New York: Academic Press.

Naik, D. N., and Rao, S. (2001), "Analysis of Multivariate Repeated Measures Data with a Kronecker Product Structured Covariance Matrix", *Journal of Applied Statistics*, 28, 91-105.

Roy, A., and Khattree, R. (2005), "On implementation of a test for Kronecker product covariance structure for multivariate repeated measures data", *Statistical Methodology*, 2, 297-306.

Srivastava, M. S., Nahtman, T., and von Rosen, D. (2007), "Models with a Kronecker product covariance structure: Estimation and testing", *Research Report, Swedish University of Agricultural Sciences*, 26 pages.

Srivastava, M. S., Nahtman, T., and von Rosen, D. (2008), "Estimation in general multivariate linear models with Kronecker product covariance structure", *Research Report, Swedish University of Agricultural Sciences*, 21 pages.