Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v13n1p47

Canonical correlation analysis of principal component scores for multiple-set random vectors

By Ogura, Murakami

Published: 02 May 2020

Electronic Journal of Applied Statistical Analysis
Vol. 13, Issue 01, May 2020, 47-74
DOI: 10.1285/i20705948v13n1p47

Canonical correlation analysis of principal component scores for multiple-set random vectors

Toru Ogura*[a] and Hidetoshi Murakami[b]

[a] *Clinical Research Support Center, Mie University Hospital, 2-174, Edobashi, Tsu-city, Mie, Japan*
[b] *Department of Applied Mathematics, Tokyo University of Science, 1-3 Kagurazaka, Shinjyuku-ku, Tokyo, Japan*

Published: 02 May 2020

Canonical correlation analysis (CCA) is often used to analyze correlations between the variables of two random vectors. As an extension of CCA, multiple-set canonical correlation analysis (MCCA) was proposed to analyze correlations between multiple-set random vectors. However, sometimes interpreting MCCA results may not be as straightforward as interpreting CCA results. Principal CCA (PCCA), which uses CCA between two sets of principal component (PC) scores, was proposed to address these difficulties in CCA. We propose multiple-set PCCA (MPCCA) by applying the idea to multiple-set of PC scores. PCs are ranked in descending order according to the amount of information they contain. Therefore, it is enough to use only a few PC scores from the top instead of using all PC scores. Decreasing the number of PC makes it easy to interpret the result. We confirmed the effectiveness of MPCCA using simulation studies and a practical example.

keywords: Canonical correlation analysis, Multiple-set, Multivariate analysis, Principal components analysis.

1 Introduction

Canonical correlation analysis (CCA) is often used to examine the relationship between the variables of two random vectors (Anderson, 1951, 2003; Hotelling, 1935, 1936). However, the canonical variables may not to be easy to interpret (Ter Braak, 1990), and it

*Corresponding author: t-ogura@clin.medic.mie-u.ac.jp

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

may be difficult to estimate the inverse matrix necessary for CCA if a high correlation coefficient is included in the correlation matrix. Principal CCA (PCCA) was proposed as one method to address these difficulties (Yamamoto et al., 2007). PCCA is CCA of two sets of principal component (PC) scores, where each set of PC scores (or components) is calculated from individual random vectors using principal component analysis (PCA) (Anderson, 1963, 2003; Hotelling, 1933). PC scores have two characteristics: (I) PCA transforms a given data set of correlated variables into a new data set of uncorrelated variables, or PC scores, and (II) PCs are ranked in descending order according to the amount of information they contain. Thus, when the researcher uses PCCA, only a few PCs with high information content are used in many cases. When comparing PCCA using only a few PCs from the top and PCCA using all PCs, there is only a small loss of information. It is easier to interpret PCCA results using only a few PCs from the top. The following methods were proposed for determining the number of PCs for PCCA: cumulative contribution rate (Palatella et al., 2010), Kaiser criterion (Skourkeas et al., 2013), information criterion (Ogura, 2010; Ogura et al., 2013). As a theoretical verification of PCCA, Sugiyama et al. (2007) showed that the limit distribution of the difference between principal canonical correlation coefficient (PCCC) in PCCA and canonical correlation coefficient (CCC) in CCA was a normal distribution.

By extending CCA, multiple-set CCA (MCCA) was proposed as an analysis method to determine relevance among three or more sets (Hwang et al., 2013; Kettenring, 1971). The canonical variables of CCA were determined to maximize CCC, but the number of multiple-set CCCs (MCCCs) in k sets were obtained ${}_k\mathrm{C}_2$. Because there are no canonical variables that maximize all MCCCs at the same time, several criteria to determine MCCCs were proposed by Kettenring (1971); Tenenhaus and Tenenhaus (2011) ((i) the maximum sum of MCCCs, (ii) the maximum sum of squares of MCCCs, (iii) the maximum sum of absolute values of MCCCs, etc.). MCCA results are complicated because they are interpreted using the canonical variables of each set and several MCCCs. Therefore, interpretation of MCCA results may not be as straightforward as interpretation of CCA results. As a method to address these difficulties, we extend the idea of PCCA and propose multiple-set PCCA (MPCCA). That is, PC scores are calculated for each set, and MCCA is performed between PC scores in each set. Similarly to PCCA, it is possible to select the number of PC scores used by MPCCA. PCA can be based on either the covariance matrix or the correlation matrix. The cumulative contribution rate criterion can be used for both of the covariance matrix and the correlation matrix. In this paper, the number of PC scores used for MPCCA is determined based on the cumulative contribution rate criterion that can be used with high versatility. Because the calculation of MPCCC is more complicated than that of PCCC, it is difficult to verify using the same theoretically techniques. Therefore, we verify the effectiveness of MPCCA with several population correlation matrices in simulation studies. We used a combination of the following conditions for verification: the correlation levels within blocks, scored as high or low, and the correlation levels between blocks, scored as high or low. Furthermore, using a practical example, we consider whether interpretation of MPCCA is simpler than interpretation of MCCA.

In Section 2, we derive MPCCA in population and sample. In Section 3, we verify the effectiveness of MPCCA by several simulation studies. In Section 4, we discusses interpretation of MPCCA results using a practical example. In Section 5, we examine the difference of MPCCCs by some cumulative contribution rate criteria. Conclusions are presented in Section 6.

2 Derivation of MPCCA

CCC between two sets is determined to be the maximum in CCA. However, the number of MCCCs is ${}_k\mathrm{C}_2$ in the case of k-set. Several methods for determining MCCCs have been proposed by Kettenring (1971); Tenenhaus and Tenenhaus (2011). We extend those methods to MPCCA.

2.1 Population of MPCCA

We summarize the notation based on Anderson (2003); Ogura et al. (2013). Suppose the random vector $\boldsymbol{X}$ of $(p_1 + \cdots + p_k)$ components has a covariance matrix Ψ. Because we are only interested in the variance and covariance in this section, we assume $E(\boldsymbol{X}) = \boldsymbol{0}$ without loss of generality. We partition $\boldsymbol{X}$ into $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(k)}$ of $p_1, \ldots, p_k$ components, respectively, as follows:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}^{(1)} \\ \vdots \\ \boldsymbol{X}^{(k)} \end{pmatrix}, \quad \boldsymbol{X}^{(1)} = \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_{p_1}^{(1)} \end{pmatrix}, \ldots, \boldsymbol{X}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ \vdots \\ x_{p_k}^{(k)} \end{pmatrix}.$$

Similarly, the $\{(p_1 + \cdots + p_k) \times (p_1 + \cdots + p_k)\}$ covariance matrix Ψ is partitioned into a $p_i \times p_j$ matrix $\Psi^{(ij)}$ $(i, j = 1, \ldots, k)$:

$$\Psi = \mathrm{Cov}(\boldsymbol{X}) = \begin{pmatrix} \Psi^{(11)} & \cdots & \Psi^{(1k)} \\ \vdots & \ddots & \vdots \\ \Psi^{(k1)} & \cdots & \Psi^{(kk)} \end{pmatrix}. \tag{1}$$

Let $\lambda_1^{(1)} \geq \cdots \geq \lambda_{p_1}^{(1)} > 0$ be the ordered latent roots of $\Psi^{(11)}$, and let $\boldsymbol{\gamma}_1^{(1)}, \ldots, \boldsymbol{\gamma}_{p_1}^{(1)}$ be the corresponding latent vectors with $\boldsymbol{\gamma}_a^{(1)\prime}\boldsymbol{\gamma}_{a'}^{(1)} = \delta_{aa'}$, where $\delta_{aa'}$ is the Kronecker delta, i.e., $\delta_{aa} = 1$ and $\delta_{aa'} = 0$ for $a \neq a'$. This definition is repeated. Let $\lambda_1^{(k)} \geq \cdots \geq \lambda_{p_k}^{(k)} > 0$ be the ordered latent roots of $\Psi^{(kk)}$, and let $\gamma_1^{(k)}, \ldots, \gamma_{p_k}^{(k)}$ be the corresponding latent vectors with $\boldsymbol{\gamma}_a^{(k)\prime}\boldsymbol{\gamma}_{a'}^{(k)} = \delta_{aa'}$. We can decompose $\Psi^{(11)}, \ldots, \Psi^{(kk)}$ as

$$\Gamma^{(1)\prime}\Psi^{(11)}\Gamma^{(1)} = \Sigma^{(11)} = \Lambda^{(1)}, \ldots, \Gamma^{(k)\prime}\Psi^{(kk)}\Gamma^{(k)} = \Sigma^{(kk)} = \Lambda^{(k)}, \tag{2}$$

where $\Lambda^{(1)} = \mathrm{diag}(\lambda_1^{(1)}, \ldots, \lambda_{p_1}^{(1)}), \ldots, \Lambda^{(k)} = \mathrm{diag}(\lambda_1^{(k)}, \ldots, \lambda_{p_k}^{(k)})$ are the diagonal matrices, and $\Gamma^{(1)} = (\boldsymbol{\gamma}_1^{(1)} \cdots \boldsymbol{\gamma}_{p_1}^{(1)}), \ldots, \Gamma^{(k)} = (\boldsymbol{\gamma}_1^{(k)} \cdots \boldsymbol{\gamma}_{p_k}^{(k)})$ are the orthonormal matrices.

PC scores of $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(k)}$ are then defined by $\boldsymbol{U}^{(1)} = \Gamma^{(1)\prime}\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{U}^{(k)} = \Gamma^{(k)\prime}\boldsymbol{X}^{(k)}$. We denote PC scores as

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{U}^{(1)} \\ \vdots \\ \boldsymbol{U}^{(k)} \end{pmatrix}, \quad \boldsymbol{U}^{(1)} = \begin{pmatrix} u_1^{(1)} \\ \vdots \\ u_{p_1}^{(1)} \end{pmatrix}, \ldots, \boldsymbol{U}^{(k)} = \begin{pmatrix} u_1^{(k)} \\ \vdots \\ u_{p_k}^{(k)} \end{pmatrix}.$$

Because PCs descend in the order of the amount of information that they contain, we use the first to q_1th PC scores of $\boldsymbol{U}^{(1)}$. This definition is repeated. The first to q_kth PC scores of $\boldsymbol{U}^{(k)}$ are defined as $(1 \le q_1 \le p_1, \ldots, 1 \le q_k \le p_k)$:

$$\boldsymbol{U}^{(1)*} = \begin{pmatrix} u_1^{(1)} \\ \vdots \\ u_{q_1}^{(1)} \end{pmatrix}, \ldots, \boldsymbol{U}^{(k)*} = \begin{pmatrix} u_1^{(k)} \\ \vdots \\ u_{q_k}^{(k)} \end{pmatrix}.$$

The covariance matrix of

$$\boldsymbol{U}^* = \begin{pmatrix} \boldsymbol{U}^{(1)*} \\ \vdots \\ \boldsymbol{U}^{(k)*} \end{pmatrix}$$

is defined by

$$\mathrm{Cov}(\boldsymbol{U}^*) = \Sigma^* = \begin{pmatrix} \Sigma^{(11)*} & \cdots & \Sigma^{(1k)*} \\ \vdots & & \vdots \\ \Sigma^{(k1)*} & \cdots & \Sigma^{(kk)*} \end{pmatrix}, \tag{3}$$

where $\Sigma^{(ij)*}$ is a $q_i \times q_j$ matrix. There is a relationship between Ψ and Σ^* such that

$$\Gamma^{*\prime}\Psi\Gamma^* = \Sigma^*, \tag{4}$$

which can be expressed as

$$\begin{aligned} \Sigma^* &= \begin{pmatrix} \Gamma^{(1)*\prime}\Psi^{(11)}\Gamma^{(1)*} & \cdots & \Gamma^{(1)*\prime}\Psi^{(1k)}\Gamma^{(k)*} \\ \vdots & & \vdots \\ \Gamma^{(k)*\prime}\Psi^{(k1)}\Gamma^{(1)*} & \cdots & \Gamma^{(k)*\prime}\Psi^{(kk)}\Gamma^{(k)*} \end{pmatrix} \\ &= \begin{pmatrix} \Lambda^{(1)*} & & \Gamma^{(1)*\prime}\Psi^{(1k)}\Gamma^{(k)*} \\ & \ddots & \\ \Gamma^{(k)*\prime}\Psi^{(k1)}\Gamma^{(1)*} & & \Lambda^{(k)*} \end{pmatrix}, \end{aligned}$$

where

$$\Gamma^* = \begin{pmatrix} \Gamma^{(1)*} & & 0 \\ & \ddots & \\ 0 & & \Gamma^{(k)*} \end{pmatrix}, \ \Gamma^{(1)*} = (\boldsymbol{\gamma}_1^{(1)} \cdots \boldsymbol{\gamma}_{q_1}^{(1)}), \ldots, \Gamma^{(k)*} = (\boldsymbol{\gamma}_1^{(k)} \cdots \boldsymbol{\gamma}_{q_k}^{(k)}),$$

and $\Lambda^{(1)*} = \mathrm{diag}(\lambda_1^{(1)}, \ldots, \lambda_{q_1}^{(1)})$, ..., $\Lambda^{(k)*} = \mathrm{diag}(\lambda_1^{(k)}, \ldots, \lambda_{q_k}^{(k)})$.

Consider an arbitrary linear combination, $v^{(i)*} = \boldsymbol{\alpha}^{(i)*\prime}\boldsymbol{U}^{(i)*}$, of $\boldsymbol{U}^{(i)*}$ in the same method as CCA, where $\boldsymbol{\alpha}^{(i)*} = (\alpha_1^{(i)*} \cdots \alpha_{q_i}^{(i)*})'$. Because the correlation of $c_1 v^{(i)*}$ and $c_2 v^{(j)*}$ is the same as the correlation of $v^{(i)*}$ and $v^{(j)*}$ (c_1, c_2: constant), we can apply arbitrary normalization to $\boldsymbol{\alpha}^{(i)*}$. We require $\boldsymbol{\alpha}^{(i)*}$ to be such that $v^{(i)*}$ has unit variance:

$$1 = E(v^{(i)*2}) = E(\boldsymbol{\alpha}^{(i)*\prime}\boldsymbol{U}^{(i)*}\boldsymbol{U}^{(i)*\prime}\boldsymbol{\alpha}^{(i)*}) = \boldsymbol{\alpha}^{(i)*\prime}\Sigma^{(ii)*}\boldsymbol{\alpha}^{(i)*}.$$

We note that $E(v^{(i)*}) = E(\boldsymbol{\alpha}^{(i)*\prime}\boldsymbol{U}^{(i)*}) = \boldsymbol{\alpha}^{(i)*\prime}E(\boldsymbol{U}^{(i)*}) = 0$. The correlation between $v^{(i)*}$ and $v^{(j)*}$ is

$$E(v^{(i)*}v^{(j)*}) = E(\boldsymbol{\alpha}^{(i)*\prime}\boldsymbol{U}^{(i)*}\boldsymbol{U}^{(j)*\prime}\boldsymbol{\alpha}^{(j)*}) = \boldsymbol{\alpha}^{(i)*\prime}\Sigma^{(ij)*}\boldsymbol{\alpha}^{(j)*}.$$

CCA defines the canonical variables such that CCC is maximized. MCCA requires a criterion for determining the canonical variables for MCCC calculated for each set. The following three methods are proposed as criteria focused on MCCCs (Kettenring, 1971; Tenenhaus and Tenenhaus, 2011): (i) the maximum sum of MCCCs for all combinations (SUMCOR), (ii) the maximum sum of squares of MCCCs for all combinations (SSQ-COR), and (iii) the maximum sum of absolute values of MCCCs for all combinations (SABSCOR).

First, we discuss using SUMCOR. To obtain the first MPCCCs of all combinations, we take $\boldsymbol{\alpha}^{(i)*} = \boldsymbol{\alpha}_1^{(i)*} = (\alpha_{11}^{(i)*} \cdots \alpha_{1q_i}^{(i)*})'$. $v_1^{(i)*} = \boldsymbol{\alpha}_1^{(i)*\prime}\boldsymbol{U}^{(i)*}$ is the normalized linear combination of $\boldsymbol{U}^{(i)*}$ with the maximum sum of the correlation coefficients for all combinations:

$$\max_{\boldsymbol{\alpha}_1^{(1)*},\ldots,\boldsymbol{\alpha}_1^{(k)*}} \sum_{i=1}^{k-1}\sum_{j=i+1}^{k} \boldsymbol{\alpha}_1^{(i)*\prime}\Sigma^{(ij)*}\boldsymbol{\alpha}_1^{(j)*}, \tag{5}$$

where the condition of $E(v_1^{(i)*2}) = 1$ is satisfied. The first MPCCCs are defined as $\rho_1^{(ij)*} = \mathrm{Cor}(v_1^{(i)*}, v_1^{(j)*})$. To obtain the bth MPCCC for all combinations, we take $\boldsymbol{\alpha}^{(i)*} = \boldsymbol{\alpha}_b^{(i)*} = (\alpha_{b1}^{(i)*} \cdots \alpha_{bq_i}^{(i)*})'$ ($b = 2, \ldots, \min(q_i)$). We consider finding the bth linear combination of $\boldsymbol{U}^{(i)*}$, say $v_b^{(i)*} = \boldsymbol{\alpha}_b^{(i)*\prime}\boldsymbol{U}^{(i)*}$, corresponding to all linear combinations that are uncorrelated with $v_1^{(i)*}$, ..., $v_{b-1}^{(i)*}$. These have the maximum sum of correlation coefficients for all combinations:

$$\max_{\boldsymbol{\alpha}_b^{(1)*},\ldots,\boldsymbol{\alpha}_b^{(k)*}} \sum_{i=1}^{k-1}\sum_{j=i+1}^{k} \boldsymbol{\alpha}_b^{(i)*\prime}\Sigma^{(ij)*}\boldsymbol{\alpha}_b^{(j)*}, \tag{6}$$

where the conditions of $E(v_b^{(i)*2}) = 1$ and $E(v_1^{(i)*}v_b^{(i)*}) = \cdots = E(v_{b-1}^{(i)*}v_b^{(i)*}) = 0$ are satisfied ($b = 1, \ldots, \min(q_i)$). The bth MPCCCs are defined as $\rho_b^{(ij)*} = \mathrm{Cor}(v_b^{(i)*}, v_b^{(j)*})$. This method of analysis includes CCA as a special case when $k = 2$.

Similarly, the first MPCCCs using SSQCOR and SABSCOR are calculated using $\boldsymbol{\alpha}_1^{(i)*}$ set as follows:

$$\text{SSQCOR} \quad : \quad \max_{\boldsymbol{\alpha}_1^{(1)*},\ldots,\boldsymbol{\alpha}_1^{(k)*}} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left(\boldsymbol{\alpha}_1^{(i)*\prime} \Sigma^{(ij)*} \boldsymbol{\alpha}_1^{(j)*} \right)^2, \tag{7}$$

$$\text{SABSCOR} \quad : \quad \max_{\boldsymbol{\alpha}_1^{(1)*},\ldots,\boldsymbol{\alpha}_1^{(k)*}} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left| \boldsymbol{\alpha}_1^{(i)*\prime} \Sigma^{(ij)*} \boldsymbol{\alpha}_1^{(j)*} \right|, \tag{8}$$

where the condition of $E(v_1^{(i)*2}) = 1$ is satisfied. The bth MPCCCs using SSQCOR and SABSCOR are calculated using $\boldsymbol{\alpha}_b^{(i)*}$ set as follows:

$$\text{SSQCOR} \quad : \quad \max_{\boldsymbol{\alpha}_b^{(1)*},\ldots,\boldsymbol{\alpha}_b^{(k)*}} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left(\boldsymbol{\alpha}_b^{(i)*\prime} \Sigma^{(ij)*} \boldsymbol{\alpha}_b^{(j)*} \right)^2, \tag{9}$$

$$\text{SABSCOR} \quad : \quad \max_{\boldsymbol{\alpha}_b^{(1)*},\ldots,\boldsymbol{\alpha}_b^{(k)*}} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left| \boldsymbol{\alpha}_b^{(i)*\prime} \Sigma^{(ij)*} \boldsymbol{\alpha}_b^{(j)*} \right|, \tag{10}$$

where the conditions of $E(v_b^{(i)*2}) = 1$ and $E(v_1^{(i)*} v_b^{(i)*}) = \cdots = E(v_{b-1}^{(i)*} v_b^{(i)*}) = 0$ are satisfied.

We show a conceptual diagram of MPCCA in Figure 1, when $k = 3$, $p_1 = p_2 = p_3 = 5$ and $q_1 = q_2 = q_3 = 2$.

2.2 Estimation of MPCCA

Let T be the sample covariance matrix based on the sample

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}^{(1)} \\ \vdots \\ \boldsymbol{X}^{(k)} \end{pmatrix}, \quad \boldsymbol{X}^{(1)} = \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_{p_1}^{(1)} \end{pmatrix}, \ldots, \boldsymbol{X}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ \vdots \\ x_{p_k}^{(k)} \end{pmatrix}$$

of size $N = n + 1$, and partition T as

$$T = \begin{pmatrix} T^{(11)} & \cdots & T^{(1k)} \\ \vdots & & \vdots \\ T^{(k1)} & \cdots & T^{(kk)} \end{pmatrix}$$

in accordance with the partition of $\boldsymbol{X}$. The latent roots $\lambda_1^{(i)} \geq \cdots \geq \lambda_{p_i}^{(i)}$ of $\Psi^{(ii)}$ and the corresponding latent vectors $\boldsymbol{\gamma}_1^{(i)}, \ldots, \boldsymbol{\gamma}_{p_i}^{(i)}$ are then estimated using the latent roots $\ell_1^{(i)} \geq \cdots \geq \ell_{p_i}^{(i)}$ of $T^{(ii)}$ and the corresponding latent vectors $\boldsymbol{h}_1^{(i)}, \ldots \boldsymbol{h}_{p_i}^{(i)}$. We can decompose $T^{(11)}, \ldots, T^{(kk)}$ as

$$H^{(i)\prime} T^{(ii)} H^{(i)} = S^{(ii)} = D^{(i)},$$

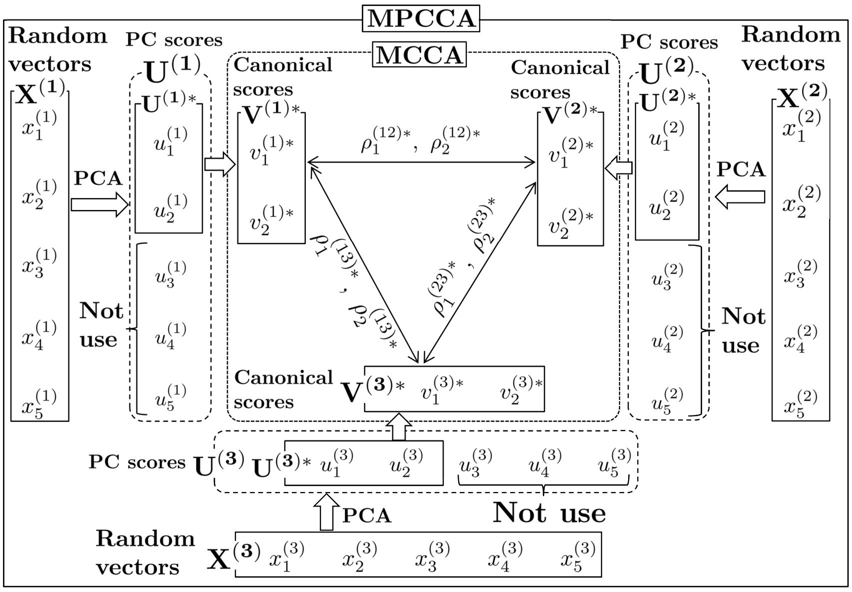

Figure 1: Conceptual diagram of MPCCA when $k = 3$, $p_1 = p_2 = p_3 = 5$ and $q_1 = q_2 = q_3 = 2$. First, each set of PC scores is calculated from the individual random vectors using PCA. Second, MCCA is applied using selected PC scores.

where $D^{(i)} = \mathrm{diag}(\ell_1^{(i)}, \ldots, \ell_{p_i}^{(i)})$ is the diagonal matrix, and $H^{(i)} = (\boldsymbol{h}_1^{(i)} \cdots \boldsymbol{h}_{p_i}^{(i)})$ is the orthonormal matrix. PC scores of $\boldsymbol{X}^{(i)}$ are defined by $\boldsymbol{U}^{(i)} = H^{(i)}\boldsymbol{X}^{(i)}$. We use the first to q_ith PC score of $\boldsymbol{U}^{(i)}$ $(1 \le q_i \le p_i)$:

$$\boldsymbol{U}^{(1)*} = \begin{pmatrix} u_1^{(1)} \\ \vdots \\ u_{q_1}^{(1)} \end{pmatrix}, \cdots, \boldsymbol{U}^{(k)*} = \begin{pmatrix} u_1^{(k)} \\ \vdots \\ u_{q_k}^{(k)} \end{pmatrix}.$$

The covariance matrix of

$$\boldsymbol{U}^{*} = \begin{pmatrix} \boldsymbol{U}^{(1)*} \\ \vdots \\ \boldsymbol{U}^{(k)*} \end{pmatrix}$$

is then estimated by Σ^*:

$$\begin{aligned} S^* &= H^{*\prime} T H^* \\ &= \begin{pmatrix} S^{(11)*} & \cdots & S^{(1k)*} \\ \vdots & & \vdots \\ S^{(k1)*} & \cdots & S^{(kk)*} \end{pmatrix} = \begin{pmatrix} D^{(1)*} & & H^{(1)*\prime} T^{(1k)} H^{(k)*} \\ & \ddots & \\ H^{(k)*\prime} T^{(k1)} H^{(1)*} & & D^{(k)*} \end{pmatrix}, \end{aligned}$$

where

$$H^* = \begin{pmatrix} H^{(1)*} & & 0 \\ & \ddots & \\ 0 & & H^{(k)*} \end{pmatrix}, \ H^{(i)*} = (\boldsymbol{h}_1^{(i)} \cdots \boldsymbol{h}_{q_i}^{(i)}),$$

and the diagonal matrix is given as $D^{(i)*} = \mathrm{diag}(\ell_1^{(i)}, \ldots, \ell_{q_i}^{(i)})$. Using the same procedure as Subsection 2.1, the bth MPCCCs are estimated to be $r_b^{(ij)*} = \hat{\rho}_b^{(ij)*}$ ($b = 1, \ldots, \min(q_i)$).

3 Simulation studies

It was difficult to verify the effectiveness of MPCCA by a logical method because the calculation of MPCCC was complicated. We verified the effectiveness of MPCCA by simulation studies in five cases. We used Mathematica version 11.0 (Wolfram Research, 2016) for the calculation of MPCCA and MCCA. The population correlation levels within and between blocks were set for the following four cases ($k = 3$):

- Case 1: The correlation levels within blocks and between blocks were high ($p_1 = p_2 = p_3 = 5$).
- Case 2: The correlation level within blocks was high, and the correlation level between blocks was low ($p_1 = 9$, $p_2 = 10$, $p_3 = 11$).
- Case 3: The correlation level within blocks was low, and the correlation level between blocks was high ($p_1 = p_2 = p_3 = 5$).
- Case 4: The correlation levels within blocks and between blocks were low ($p_1 = 9$, $p_2 = 10$, $p_3 = 11$).

In this paper, the correlation level was determined as follows. For each correlation matrix $\Phi^{(ii)}$ or $\Phi^{(ij)}$, the number of correlation coefficient absolute values at or above 0.4 was counted. If the mean of count was greater than or equal to 2 within blocks $\Phi^{(ii)}$, the correlation level within blocks was defined as high. Otherwise, the correlation level within blocks was defined as low. Similarly, if the mean of count was greater than or equal to 2 between blocks $\Phi^{(ij)}$, the correlation level between blocks was defined as high. Otherwise, the correlation level between blocks was defined as low. Furthermore, Case 5 was set to verify validity when $k = 4$. The correlation levels within blocks and between blocks were high ($p_1 = 5$, $p_2 = 6$, $p_3 = 7$, $p_4 = 8$). Palatella et al. (2010) selected 3 PC scores from 15 PC scores using the 95% cumulative contribution rate criterion. However, using the 95% cumulative contribution rate criterion for most datasets would hardly reduce the number of PC scores. If the 95% cumulative contribution rate criterion was used in practical example of Section 4, all PC scores would be used without any decrease. In this paper, the 70% cumulative contribution rate criterion was used to reduce the number of PC scores while retaining a lot of information from the original dataset. We examined the difference in results obtained using the 60%, 70%, and 80% cumulative contribution rate criteria in Section 5.

The simulation study was conducted using the following procedure:

1. Generate $\boldsymbol{X}$ of sample size N, where $\boldsymbol{X} \sim N(\boldsymbol{0}, \Psi)$.
2. Calculate the covariance matrix $T = \mathrm{Cov}(\boldsymbol{X})$, and partition T for each set: $T^{(ii)}$ $(i = 1, \ldots, k)$.
3. Applying PCA for each $T^{(ii)}$, transform $\boldsymbol{X}$ into $\boldsymbol{U}$.
4. Set $\boldsymbol{U}^*$ using the 70% cumulative contribution rate criterion from $\boldsymbol{U}$.
5. Calculate MPCCCs from $\boldsymbol{U}^*$, and MCCCs from $\boldsymbol{X}$.
6. Independently, repeat steps 1–5 10,000 times.
7. Calculate the mean and the standard deviation (SD) of MPCCC and MCCC.

3.1 Simulation studies for three sets

The population covariance matrix Ψ_1 in Case 1 was given as follows:

$$
\begin{aligned}
\Psi_1 &= \begin{pmatrix} \Psi_1^{(11)} & \Psi_1^{(12)} & \Psi_1^{(13)} \\ \Psi_1^{(21)} & \Psi_1^{(22)} & \Psi_1^{(23)} \\ \Psi_1^{(31)} & \Psi_1^{(32)} & \Psi_1^{(33)} \end{pmatrix}, \\
\Psi_1^{(11)} &= \begin{pmatrix} 58.8 & & & & \\ -65.1 & 248.6 & & & \\ -17.0 & 34.3 & 256.6 & & \\ -30.7 & 87.4 & 105.5 & 303.2 & \\ 65.7 & -136.9 & -53.1 & -103.8 & 249.0 \end{pmatrix}, \\
\Psi_1^{(21)} = \Psi_1^{(12)\prime} &= \begin{pmatrix} -25.7 & 70.9 & 13.5 & 14.4 & -49.2 \\ 2.3 & -25.1 & -157.4 & -97.1 & 14.6 \\ 32.2 & -85.3 & -18.1 & -1.3 & 67.9 \\ -0.4 & 1.7 & 101.8 & 93.4 & -36.3 \\ 37.0 & -81.7 & -9.8 & -43.7 & 102.5 \end{pmatrix}, \\
\Psi_1^{(31)} = \Psi_1^{(13)\prime} &= \begin{pmatrix} 17.8 & -60.1 & -44.3 & -32.9 & 46.7 \\ 35.6 & -92.0 & -30.0 & -53.6 & 70.2 \\ 3.7 & -4.7 & -65.3 & -47.6 & 16.9 \\ -16.3 & 21.1 & 74.4 & 88.1 & -48.0 \\ -55.0 & 166.9 & 59.0 & 87.9 & -127.0 \end{pmatrix},
\end{aligned}
$$

$$\Psi_1^{(22)} = \begin{pmatrix} 139.3 & & & & \\ -5.5 & 357.9 & & & \\ -38.5 & 38.1 & 113.8 & & \\ 14.5 & -134.8 & -14.6 & 204.1 & \\ -60.4 & 1.4 & 46.8 & -40.3 & 154.1 \end{pmatrix},$$

$$\Psi_1^{(32)} = \Psi_1^{(23)\prime} = \begin{pmatrix} -41.4 & 27.6 & 8.6 & -17.2 & 33.4 \\ -38.7 & 17.7 & 30.3 & -8.2 & 32.1 \\ -6.9 & 75.6 & 11.8 & -36.3 & 2.8 \\ 2.1 & -78.1 & -3.9 & 63.5 & -16.3 \\ 109.9 & -18.9 & -44.6 & 20.1 & -80.8 \end{pmatrix},$$

$$\Psi_1^{(33)} = \begin{pmatrix} 60.5 & & & & \\ 35.1 & 57.7 & & & \\ 26.2 & 12.2 & 56.0 & & \\ -15.5 & -16.8 & -28.6 & 68.8 & \\ -91.2 & -78.5 & -38.5 & 36.8 & 237.8 \end{pmatrix}.$$

The population correlation matrix Φ_1 transformed from Ψ_1 was shown in Supplementary Material 2; the correlation levels within blocks and between blocks were found to be high. We applied PCA to $\Psi_1^{(11)}$, $\Psi_1^{(22)}$, and $\Psi_1^{(33)}$. Table 1 showed the latent roots, contribution rate, cumulative contribution rate, and the latent vectors. Using the 70% cumulative contribution rate criterion, the variables for PC scores used in MPCCA were $q_1 = q_2 = q_3 = 2$. We calculated population MPCCCs using $q_1 = q_2 = q_3 = 2$ and population MCCCs using all variables ($p_1 = p_2 = p_3 = 5$). The results using the criteria of SUMCOR, SSQCOR, and SABSCOR were shown in Table 2. Population MPCCC and MCCC had little difference between the three criteria.

The simulation study results using SUMCOR were shown in Table 3. Since each the SD was small, we compared the mean. For $N = 100$, the difference between the first MPCCCs and MCCCs were less than 0.1: $(\bar{r}_1^{(12)} - \bar{r}_1^{(12)*}, \bar{r}_1^{(13)} - \bar{r}_1^{(13)*}, \bar{r}_1^{(23)} - \bar{r}_1^{(23)*}) = (0.081, 0.067, 0.044)$, where $\bar{r}_b^{(ij)*}$ and $\bar{r}_b^{(ij)}$ were the mean of the simulation study results in MPCCCs and MCCCs, respectively. The difference between the second MPCCCs and MCCCs were less than 0.15: $(\bar{r}_2^{(12)} - \bar{r}_2^{(12)*}, \bar{r}_2^{(13)} - \bar{r}_2^{(13)*}, \bar{r}_2^{(23)} - \bar{r}_2^{(23)*}) = (0.083, 0.134, 0.062)$. Thus, the overall interpretation was not changed by this difference. When N increased, the difference between MPCCCs and MCCCs tended to decrease. Thus, MPCCA was a reliable indicator for the analysis results, even for $N = 100$.

CCCs were known to have a positive bias (Lawley, 1959). Similarly, MPCCCs and MCCCs were considered to have the bias. Thus, we investigated the bias of MPCCCs and MCCCs. The differences between the population and the mean of the simulation study results in MPCCCs and MCCCs were given below:

$$
\begin{aligned}
&(\bar{r}_1^{(12)*} - \rho_1^{(12)*}, \bar{r}_1^{(13)*} - \rho_1^{(13)*}, \bar{r}_1^{(23)*} - \rho_1^{(23)*}) = (0.057, 0.021, 0.005),\\
&(\bar{r}_2^{(12)*} - \rho_2^{(12)*}, \bar{r}_2^{(13)*} - \rho_2^{(13)*}, \bar{r}_2^{(23)*} - \rho_2^{(23)*}) = (-0.081, -0.046, -0.019),\\
&(\bar{r}_1^{(12)} - \rho_1^{(12)}, \bar{r}_1^{(13)} - \rho_1^{(13)}, \bar{r}_1^{(23)} - \rho_1^{(23)}) = (0.036, 0.032, 0.042),\\
&(\bar{r}_2^{(12)} - \rho_2^{(12)}, \bar{r}_2^{(13)} - \rho_2^{(13)}, \bar{r}_2^{(23)} - \rho_2^{(23)}) = (0.014, -0.021, 0.002),
\end{aligned}
$$

when $N = 100$. As N increased to 150 and 200 in MPCCCs and MCCCs, the mean of the simulation study results approached the population, and the SD became smaller. The bias was included in MPCCCs and MCCCs, but its influence was minimal on the simulation study results in Case 1.

The simulation study results using SSQCOR and SABSCOR were shown in Supplementary Material 4. The simulation study results of SUMCOR and SSQCOR similar, but MPCCCs and MCCCs of SABSCOR were occasionally small. That is, in the simulation study results of SABSCOR, the mean of MPCCCs and MCCCs was small, and the SD of MPCCCs and MCCCs was large. Therefore, we recommend using SUMCOR and SSQCOR in MPCCA. Since SUMCOR and SSQCOR gave similar results, we only discussed SUMCOR in this paper.

Since the simulation study results were stable at $N = 100$ in Case 1, we ran the simulation study with $N = 100$ in Cases 2, 3, 4, and 5. The population covariance matrices in Cases 2, 3, and 4 were shown as Ψ_2, Ψ_3 and Ψ_4 in Supplementary Material 1, and the population correlation matrices were shown as Φ_2, Φ_3 and Φ_4 in Supplementary Material 2, respectively. The results of PCA for Cases 2, 3, and 4 were shown in Tables C2, C3, and C4 in Supplementary Material 3. Using the 70% cumulative contribution rate criterion, PC scores used in Case 2 were $q_1 = 3$, $q_2 = 4$, $q_3 = 5$, PC scores used in Case 3 were $q_1 = q_2 = q_3 = 2$, and PC scores used in Case 4 were $q_1 = 5$, $q_2 = 6$, $q_3 = 7$. The simulation study results of MPCCCs and MCCCs for Cases 2, 3, and 4 were shown in Tables 4, 5, and 6, respectively. Since our purpose was to compare MPCCC and MCCC, each table showed MPCCC from 1st to q_1th, MCCC also showed the same 1st to q_1th.

For Case 2, the difference between MCCCs and MPCCCs were given below:

$$
\begin{aligned}
&(\bar{r}_1^{(12)} - \bar{r}_1^{(12)*}, \bar{r}_1^{(13)} - \bar{r}_1^{(13)*}, \bar{r}_1^{(23)} - \bar{r}_1^{(23)*}) = (0.109, 0.104, 0.069),\\
&(\bar{r}_2^{(12)} - \bar{r}_2^{(12)*}, \bar{r}_2^{(13)} - \bar{r}_2^{(13)*}, \bar{r}_2^{(23)} - \bar{r}_2^{(23)*}) = (0.099, 0.130, 0.136),\\
&(\bar{r}_3^{(12)} - \bar{r}_3^{(12)*}, \bar{r}_3^{(13)} - \bar{r}_3^{(13)*}, \bar{r}_3^{(23)} - \bar{r}_3^{(23)*}) = (0.214, 0.297, 0.228).
\end{aligned}
$$

The first differences were about 0.1, and the second differences were less than 0.15, the third differences were greater than 0.2. Although MPCCCs loss seemed to be large, these differences were due to the large bias. The differences between the population and the mean of the simulation study results in MPCCCs and MCCCs were given below:

$$
\begin{aligned}
&(\bar{r}_1^{(12)*} - \rho_1^{(12)*}, \bar{r}_1^{(13)*} - \rho_1^{(13)*}, \bar{r}_1^{(23)*} - \rho_1^{(23)*}) = (0.057, 0.035, 0.018),\\
&(\bar{r}_2^{(12)*} - \rho_2^{(12)*}, \bar{r}_2^{(13)*} - \rho_2^{(13)*}, \bar{r}_2^{(23)*} - \rho_2^{(23)*}) = (-0.006, -0.011, 0.064),\\
&(\bar{r}_3^{(12)*} - \rho_3^{(12)*}, \bar{r}_3^{(13)*} - \rho_3^{(13)*}, \bar{r}_3^{(23)*} - \rho_3^{(23)*}) = (-0.028, 0.056, -0.005),\\
&(\bar{r}_1^{(12)} - \rho_1^{(12)}, \bar{r}_1^{(13)} - \rho_1^{(13)}, \bar{r}_1^{(23)} - \rho_1^{(23)}) = (0.134, 0.108, 0.056),\\
&(\bar{r}_2^{(12)} - \rho_2^{(12)}, \bar{r}_2^{(13)} - \rho_2^{(13)}, \bar{r}_2^{(23)} - \rho_2^{(23)}) = (0.100, 0.067, 0.054),\\
&(\bar{r}_3^{(12)} - \rho_3^{(12)}, \bar{r}_3^{(13)} - \rho_3^{(13)}, \bar{r}_3^{(23)} - \rho_3^{(23)}) = (0.147, 0.208, 0.130).
\end{aligned}
$$

MCCCs with many variables had to consider the influence of bias. It was an advantage that MPCCCs with only the selected PC score had less bias.

For Case 3, the difference between MCCCs and MPCCCs were given below:

$$
\begin{aligned}
&(\bar{r}_1^{(12)} - \bar{r}_1^{(12)*}, \bar{r}_1^{(13)} - \bar{r}_1^{(13)*}, \bar{r}_1^{(23)} - \bar{r}_1^{(23)*}) = (0.380, 0.192, 0.343),\\
&(\bar{r}_2^{(12)} - \bar{r}_2^{(12)*}, \bar{r}_2^{(13)} - \bar{r}_2^{(13)*}, \bar{r}_2^{(23)} - \bar{r}_2^{(23)*}) = (0.278, 0.477, 0.423).
\end{aligned}
$$

The first MCCCs were highly correlated, but the correlation between the first MPCCCs was small compared with that for MCCCs. The second MCCCs showed the high correlation, but the second MPCCCs showed low correlation. Therefore, reasonable simulation study results were obtained by MCCA, but not by MPCCA.

In Case 4, because the correlation level within blocks was low, the deviation of the latent root was small. Therefore, there was no advantage using PC score. When population MCCCs were small, the population and the mean of the simulation study results in MPCCCs were also small. Therefore, we could not determine the relationship between the three sets in this case.

From these simulation study results, it was found that whether the application of MPCCA was effective could be judged by the correlation matrix. MPCCA with a reduced number of PC scores showed the simulation study results similar to MCCA using all variables when the correlation levels within blocks and between blocks were high. On the other hand, MPCCA has the following disadvantages: (i) MPCCCs were small when MCCCs were small; (ii) many PC scores were required for the 70% cumulative contribution rate criterion when the differences between sets of latent roots were small; and (iii) MPCCCs were small when the correlation level between blocks was small, even if MCCCs were large.

3.2 Simulation studies of four sets

This subsection examined the effectiveness of MPCCA for four sets in Case 5. The population covariance matrix Ψ_5 was shown in Supplementary Material 1, and the population correlation matrix Φ_5 was shown in Supplementary Material 2. The results of PCA were shown Table C5 in Supplementary Material 3. Using the 70% cumulative contribution rate criterion, PC scores used in Case 5 were $q_5 = q_5 = 2$, $q_5 = q_5 = 3$. Table 4 showed the simulation study results from MPCCC and MCCC.

For Case 5, the difference between MCCCs and MPCCCs were shown below:

$$\begin{aligned}&(\bar{r}_1^{(12)} - \bar{r}_1^{(12)*}, \bar{r}_1^{(13)} - \bar{r}_1^{(13)*}, \bar{r}_1^{(14)} - \bar{r}_1^{(14)*}, \bar{r}_1^{(23)} - \bar{r}_1^{(23)*}, \bar{r}_1^{(24)} - \bar{r}_1^{(24)*}, \bar{r}_1^{(34)} - \bar{r}_1^{(34)*})\\&= (-0.014, 0.046, 0.070, 0.120, 0.088, 0.026),\\&(\bar{r}_2^{(12)} - \bar{r}_2^{(12)*}, \bar{r}_2^{(13)} - \bar{r}_2^{(13)*}, \bar{r}_2^{(14)} - \bar{r}_2^{(14)*}, \bar{r}_2^{(23)} - \bar{r}_2^{(23)*}, \bar{r}_2^{(24)} - \bar{r}_2^{(24)*}, \bar{r}_2^{(34)} - \bar{r}_2^{(34)*})\\&= (0.047, 0.071, 0.097, 0.197, 0.173, 0.046).\end{aligned}$$

The first differences were less than 0.1, except for $\bar{r}_1^{(23)} - \bar{r}_1^{(23)*}$, and the second differences were less than 0.1, except for $\bar{r}_2^{(23)} - \bar{r}_2^{(23)*}$ and $\bar{r}_2^{(24)} - \bar{r}_2^{(24)*}$. The reason for these larger differences was that the biases were large for MCCCs: $\bar{r}_1^{(23)} - \rho_1^{(23)} = 0.082$, $\bar{r}_2^{(23)} - \rho_2^{(23)} = 0.096$, and $\bar{r}_2^{(24)} - \rho_2^{(24)} = 0.111$. In contrast, the biases of the same combination of MPCCCs were small: $\bar{r}_1^{(23)*} - \rho_1^{(23)*} = 0.024$, $\bar{r}_2^{(23)*} - \rho_2^{(23)*} = -0.003$, and $\bar{r}_2^{(24)*} - \rho_2^{(24)*} = -0.013$. We also found that the bias decreased as the variable decreased in Case 5.

3.3 Computation time

The computation time required for MPCCA and MCCA was longer than that for the CCA. The condition in (6) became complicated when $\min(q_i)$ and k increase, so the calculation time of MPCCA (or MCCA) became long. We measured the computation time of MPCCA 10,000 times for $N = 100, 150, 200$, $k = 3, 4$, and $q = 2, 3, 4, 5$, where $q = q_1 = \cdots = q_k$. Table 8 showed the mean of the computation time for MPCCA. When $k = 3$, the mean of the computation time in $q = 5$ was about 24 times as long as that of $q = 2$ for every N. When $k = 4$, the mean of the computation time in $q = 5$ was about 14 times as long as that of $q = 2$ for every N. The computation time also increased with N or k, but could be shortened by decreasing PC score. Thus, to reduce the variables in terms of the computation time, MPCCA had merit. Note that, although MPCCA and MCCA differed with regard to whether PC scores were used, the time required to convert random vectors to PC scores was very short.

4 Practical example

We used the educational data to illustrate the performance of three sets in MPCCA. The three sets were obtained from the academic records of 147 high-school seniors, specifically, their first- and second-grade academic records, and the scores from the Joint First-Stage Achievement Test (these data were obtained by personal communication with Prof. Takakazu Sugiyama). The Joint First-Stage Achievement Test was a general, basic content exam for first- and second-grades in Japanese high schools. This test was used to select university applicants in Japan. The three sets were adjusted to obtain data within the range of 0-100. We arrived at the following covariance matrix T for $\boldsymbol{X}^{(1)} = (x_1^{(1)}, \ldots, x_5^{(1)})'$ (representing the academic records of high-school seniors in the first-grade), $\boldsymbol{X}^{(2)} = (x_1^{(2)}, \ldots, x_5^{(2)})'$ (representing the academic records of high-school seniors in the second-grade), and $\boldsymbol{X}^{(3)} = (x_1^{(3)}, \ldots, x_5^{(3)})'$ (representing the Joint First-Stage Achievement Test scores), where $x_1^{(1)}$, $x_1^{(2)}$, and $x_1^{(3)}$ were English, $x_2^{(1)}$, $x_2^{(2)}$, and $x_2^{(3)}$ were Japanese, $x_3^{(1)}$, $x_3^{(2)}$, and $x_3^{(3)}$ were Mathematics, $x_4^{(1)}$, $x_4^{(2)}$, and $x_4^{(3)}$ were Science, and $x_5^{(1)}$, $x_5^{(2)}$, and $x_5^{(3)}$ were Social Studies:

$$
\begin{aligned}
T &= \begin{pmatrix} T^{(11)} & T^{(12)} & T^{(13)} \\ T^{(21)} & T^{(22)} & T^{(23)} \\ T^{(31)} & T^{(32)} & T^{(33)} \end{pmatrix}, \\
T^{(11)} &= \begin{pmatrix} 366.5 & & & & \\ 209.4 & 358.2 & & & \\ 159.2 & 90.9 & 352.9 & & \\ 88.8 & 85.7 & 174.4 & 406.8 & \\ 113.7 & 196.0 & 35.5 & 91.1 & 359.6 \end{pmatrix}, \\
T^{(21)} = T^{(12)'} &= \begin{pmatrix} 265.2 & 236.9 & 128.9 & 110.6 & 157.2 \\ 168.3 & 274.0 & 28.1 & 59.6 & 192.0 \\ 138.6 & 94.2 & 200.7 & 231.9 & 44.0 \\ 96.2 & 153.5 & 135.3 & 223.4 & 88.1 \\ 129.6 & 186.5 & 60.8 & 145.8 & 162.1 \end{pmatrix}, \\
T^{(31)} = T^{(13)'} &= \begin{pmatrix} 189.0 & 180.7 & 68.9 & 32.8 & 109.3 \\ 48.1 & 113.4 & -65.2 & -17.9 & 106.6 \\ 91.5 & 64.1 & 332.8 & 335.7 & -30.3 \\ 56.4 & 54.8 & 140.0 & 294.2 & 1.8 \\ 121.6 & 169.2 & 25.6 & 83.0 & 127.8 \end{pmatrix}, \\
T^{(22)} &= \begin{pmatrix} 339.5 & & & & \\ 229.7 & 363.8 & & & \\ 105.7 & 70.7 & 368.0 & & \\ 117.9 & 131.0 & 161.8 & 416.6 & \\ 164.9 & 187.3 & 109.1 & 152.2 & 390.3 \end{pmatrix}, \\
T^{(32)} = T^{(23)'} &= \begin{pmatrix} 246.7 & 156.3 & 56.0 & 61.0 & 105.1 \\ 83.1 & 173.2 & -28.9 & -0.3 & 157.7 \\ 28.8 & -34.8 & 364.9 & 216.4 & 36.6 \\ 36.0 & 14.7 & 243.0 & 222.1 & 102.7 \\ 174.3 & 190.3 & 37.2 & 118.5 & 235.7 \end{pmatrix},
\end{aligned}
$$

$$T^{(33)} = \begin{pmatrix} 546.7 & & & & \\ 130.5 & 473.8 & & & \\ 26.5 & -162.2 & 1392.6 & & \\ 67.4 & -50.7 & 494.3 & 782.3 & \\ 229.1 & 175.8 & 6.2 & 181.7 & 574.5 \end{pmatrix}.$$

We then applied PCA to each of $T^{(11)}$, $T^{(22)}$, and $T^{(33)}$; the results obtained were shown in Table 9.

When the 70% cumulative contribution rate criterion was used, the variables for PC scores used in MPCCA were $q_1 = 3$, $q_2 = q_3 = 2$. From the result of the latent vectors, the first PCs of $\boldsymbol{X}^{(1)*}$ and $\boldsymbol{X}^{(2)*}$ were the sum of all subjects. The first PC of $\boldsymbol{X}^{(3)*}$ was to the Science course. The second PCs of $\boldsymbol{X}^{(1)*}$ and $\boldsymbol{X}^{(2)*}$ distinguished the liberal arts course from the science course. The second PC of $\boldsymbol{X}^{(3)*}$ was the Liberal Arts course. The third PC of $\boldsymbol{X}^{(1)*}$ was English and Mathematics versus Science and Social Studies.

MPCCCs using $q_1 = 3$ and $q_2 = q_3 = 2$, and MCCCs using all of the variables were summarized in Table 10. The difference between the first MPCCC and MCCC was small: $r_1^{(12)} - r_1^{(12)*} = 0.064$, $r_1^{(13)} - r_1^{(13)*} = 0.022$ and $r_1^{(23)} - r_1^{(23)*} = 0.015$. The difference between the second MPCCC and MCCC was also small: $r_2^{(12)} - r_2^{(12)*} = 0.065$, $r_2^{(13)} - r_2^{(13)*} = 0.117$ and $r_2^{(23)} - r_2^{(23)*} = 0.047$. Because the difference between the first MPCCCs and first MCCCs were less than 0.1, and the second MPCCCs and second MCCCs were less than 0.15, interpretation between the three sets could be considered as being almost the same in MPCCA and MCCA.

The coefficients of the canonical variables in MPCCA and MCCA were shown in Table 11. In the first MPCCCs, $\alpha_{11}^{(1)*}$, $\alpha_{11}^{(2)*}$, $\alpha_{11}^{(3)*}$, and $\alpha_{12}^{(3)*}$ were large and positive. Considering the characteristics of each PC and the first coefficients of the canonical variables in MPCCA, the first MPCCCs were the relevance of the sums of all subjects in the first- and second-grade academic records, and the Joint First-Stage Achievement Test. In the second MPCCCs, $\alpha_{22}^{(1)*}$, $\alpha_{22}^{(2)*}$, and $\alpha_{32}^{(3)*}$ were large and positive, and $\alpha_{31}^{(3)*}$ was large and negative. Considering the characteristics of each PC and the second coefficients of the canonical variables in MPCCA, the second MPCCCs were the relevance of the difference between the Science and Liberal Arts courses in the first- and second-grade academic records, and the Joint First-Stage Achievement Test.

In the first MCCCs, $\alpha_{12}^{(1)}$, $\alpha_{14}^{(1)}$, $\alpha_{11}^{(2)}$, $\alpha_{13}^{(2)}$, $\alpha_{11}^{(3)}$, $\alpha_{13}^{(3)}$, $\alpha_{14}^{(3)}$, and $\alpha_{15}^{(3)}$ were greater than 0.2. They were English, Japanese, and Science in the first-grade academic record, English and Mathematics in the second-grade academic record, and English, Mathematics, Science, and Social Studies in the Joint First-Stage Achievement Test. It was difficult to interpret the relationship between them. In the second MCCCs, $\alpha_{21}^{(i)}$, $\alpha_{22}^{(i)}$, and $\alpha_{25}^{(i)}$ were negative, and $\alpha_{23}^{(i)}$ and $\alpha_{24}^{(i)}$ were positive in all three sets. Considering the characteristics of each PC and the second coefficients of the canonical variables in MCCA, the second MCCCs were the relevance of the coefficients were distinguished by differences between Science and Liberal Arts courses.

Since the number of large coefficients of the canonical variables was a little in MPCCA, interpretations of the first and second MPCCC were easy. However, interpretation of the first MPCCC was difficult because there were many large coefficients in MCCA.

5 Optimal cumulative contribution rate criterion in MPCCCs

We compare MPCCCs when the number of PC scores used in MPCCA determined by the 60% or 80% cumulative contribution rate criterion and the 70% cumulative contribution rate criterion.

In simulation study of Case 1, using the 60% cumulative contribution rate criterion, the variables for PC scores used in MPCCA were $q_1 = q_2 = 2$, $q_3 = 1$ from Table 1. Then, the simulation study results of the first MPCCCs were $(\bar{r}_1^{(12)*}, \bar{r}_1^{(13)*}, \bar{r}_1^{(23)*}) = (0.585, 0.726, 0.568)$ for $N = 100$. These results were small compared to the first MPCCCs in Table 3. The second MCCCs using all variables were large values, but only the first MPCCCs were obtained. Using the 80% cumulative contribution rate criterion, the variables for PC scores used in MPCCA were $q_1 = q_2 = 3$, $q_3 = 2$ from Table 1. Then, the simulation study results of the first and the second MPCCCs were $(\bar{r}_1^{(12)*}, \bar{r}_1^{(13)*}, \bar{r}_1^{(23)*}) = (0.669, 0.753, 0.702)$ and $(\bar{r}_2^{(12)*}, \bar{r}_2^{(13)*}, \bar{r}_2^{(23)*}) = (0.573, 0.688, 0.584)$ for $N = 100$, respectively. The first and the second MPCCCs were larger than those in Table 3, but their differences were small.

In practical example, using the 60% cumulative contribution rate criterion, the variables for PC scores used in MPCCA were $q_1 = 2$, $q_2 = q_3 = 1$ from Table 9 and the first MPCCCs were obtained as $(r_1^{(12)*}, r_1^{(13)*}, r_1^{(23)*}) = (0.748, 0.496, 0.279)$. Since the first MPCCCs determined by the 70% cumulative contribution rate criterion were $(r_1^{(12)*}, r_1^{(13)*}, r_1^{(23)*}) = (0.800, 0.549, 0.571)$ from Table 10, $r_1^{(23)*}$ using the 60% cumulative contribution rate criterion was much smaller than the 70% cumulative contribution rate criterion. The second MCCCs using all variables were large values, but only the first MPCCCs were obtained. Using the 80% cumulative contribution rate criterion, the variables for PC scores used in MPCCA were $q_1 = q_2 = q_3 = 3$ from Table 9 and the first to third MPCCCs were obtained as $(r_1^{(12)*}, r_1^{(13)*}, r_1^{(23)*}) = (0.831, 0.547, 0.579)$, $(r_2^{(12)*}, r_2^{(13)*}, r_2^{(23)*}) = (0.689, 0.573, 0.664)$, $(r_3^{(12)*}, r_3^{(13)*}, r_3^{(23)*}) = (0.298, 0.198, 0.218)$, respectively. Since the second MPCCCs determined by the 70% cumulative contribution rate criterion were $(r_2^{(12)*}, r_2^{(13)*}, r_2^{(23)*}) = (0.646, 0.454, 0.615)$ from Table 10, the difference between the first and second MPCCCs using the 80% cumulative contribution rate criterion and the 70% cumulative contribution rate criterion were small. Comparing the third MCCCs using all variables in Table 10 and the third MPCCCs using the 80% cumulative contribution rate criterion, the third MPCCCs using the 80% cumulative contribution rate criterion were much smaller than the third MCCCs. MPCCCs using the 60% cumulative contribution rate criterion had large losses, and MPCCCs using the 80% cumulative contribution rate criterion had many PC scores but less advantages. Therefore, MPCCCs using the 70% cumulative contribution rate criterion were found to give reasonable results.

6 Conclusions

We proposed MPCCA for a relevance analysis method for multiple-set. Because PCs were ranked in order according to the amount of information they contain, we used a few PC scores such that $\boldsymbol{U}^{(1)*} = (u_1^{(1)} \ \cdots \ u_{q_1}^{(1)})'$, ..., $\boldsymbol{U}^{(k)*} = (u_1^{(k)} \ \cdots \ u_{q_k}^{(k)})'$. From simulation study results and numerical example, we found that selecting PC scores using the 70% cumulative contribution rate criterion yielded reasonable results. The criteria for determining MCCCs were proposed as SUMCOR, SSQCOR, and SABSCOR, and we extended them to the criteria for determining MPCCCs. The simulation study results showed that SUMCOR and SSQCOR gave similar results. However, the simulation study results using SABCOR were sometimes much smaller than population MPCCCs. Therefore, we discussed the results using SUMCOR in this paper. Furthermore, MPCCCs and MCCCs were compared using five simulation studies. From the simulation study results, when the correlation level within blocks was high, MPCCA analysis was effective. On the other hand, when the correlation level within blocks was low, MPCCA did not work well. Thus, whether it will be effective to use MPCCA can be judged from the correlation level within blocks. We compared MPCCA and MCCA using the practical example and also compared interpretation of those results. With regard to the canonical variables, because there were a few large coefficients in MPCCA, the result was easy to interpret. However, in MCCA, there were many large coefficients, making interpretation of the results more difficult.

Acknowledgement

The authors are grateful to the editor and the referee for their meaningful comments and suggestions, which led to great improvements in the revised manuscript.

References

Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *The Annals of Mathematical Statistics*, 22(3):327–351.

Anderson, T. W. (1963). Asymptotic theory for principal component analysis. *The Annals of Mathematical Statistics*, 34(1):122–148.

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis, 3rd Edition*. Wiley, New York.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441.

Hotelling, H. (1935). The most predictable criterion. *Journal of Educational Psychology*, 26(2):139–142.

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28(3):321–377.

Hwang, H., Jung, K., Takane, Y., and Woodward, T. S. (2013). A unified approach to multiple-set canonical correlation analysis and principal components analysis. *British Journal of Mathematical and Statistical Psychology*, 66(2):308–321.

Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika*, 58(3):433–451.

Lawley, D. N. (1959). Tests of significance in canonical analysis. *Biometrika*, 46(1–2):59–66.

Ogura, T. (2010). A variable selection method in principal canonical correlation analysis. *Computational Statistics & Data Analysis*, 54(4):1117–1123.

Ogura, T., Fujikoshi, Y., and Sugiyama, T. (2013). A variable selection criterion for two sets of principal component scores in principal canonical correlation analysis. *Communications in Statistics - Theory and Methods*, 42(12):2118–2135.

Palatella, L., Miglietta, M. M., Paradisi, P., and Lionello, P. (2010). Climate change assessment for mediterranean agricultural areas by statistical downscaling. *Natural Hazards and Earth System Sciences*, 10(4):1647–1661.

Skourkeas, A., Kolyva-Machera, F., and Maheras, P. (2013). Improved statistical downscaling models based on canonical correlation analysis, for generating temperature scenarios over greece. *Environmental and Ecological Statistics*, 20(3):445–465.

Sugiyama, T., Ogura, T., Sakaori, F., and Yamada, T. (2007). Distributions and bootstrap method of some statistics in principal canonical correlation analysis. *Journal of the Japan Statistical Society*, 37(2):239–251.

Tenenhaus, A. and Tenenhaus, M. (2011). Regularized generalized canonical correlation analysis. *Psychometrika*, 76(2):257–284.

Ter Braak, C. J. (1990). Interpreting canonical correlation analysis through biplots of structure correlations and weights. *Psychometrika*, 55(3):519–531.

Wolfram Research, I. (2016). *Mathematica*. Wolfram Research, Inc., Champaign, Illinois, version 11.0 edition.

Yamamoto, M., Sugiyama, T., Murakami, H., and Sakaori, F. (2007). Correlation analysis of principal components from two populations. *Computational Statistics & Data Analysis*, 51(9):4707–4716.

Table 1: PCA for each population covariance matrix in Case 1. The underline showed the 70% cumulative contribution rate.

			1st	2nd	3rd	4th	5th
$\Psi_1^{(11)}$	Latent roots		549.8	258.7	162.9	110.4	34.3
	Contribution rate		0.493	0.232	0.146	0.099	0.031
	Cumulative contribution rate		0.493	$\underline{0.724}$	0.870	0.969	1.000
	Latent vectors	$x_1^{(1)}$	0.182	−0.168	−0.089	0.016	0.965
		$x_2^{(1)}$	−0.484	0.497	0.151	0.681	0.181
		$x_3^{(1)}$	−0.367	−0.661	0.650	0.074	0.013
		$x_4^{(1)}$	−0.571	−0.388	−0.722	0.045	−0.027
		$x_5^{(1)}$	0.522	−0.371	−0.161	0.727	−0.190
$\Psi_1^{(22)}$	Latent roots		445.6	236.3	131.3	86.0	70.0
	Contribution rate		0.460	0.244	0.135	0.089	0.072
	Cumulative contribution rate		0.460	$\underline{0.704}$	0.839	0.928	1.000
	Latent vectors	$x_1^{(2)}$	−0.079	0.567	0.254	0.771	0.118
		$x_2^{(2)}$	0.842	0.275	−0.391	−0.025	0.250
		$x_3^{(2)}$	0.144	−0.365	−0.398	0.513	−0.652
		$x_4^{(2)}$	−0.502	0.115	−0.779	0.066	0.352
		$x_5^{(2)}$	0.113	−0.676	0.136	0.371	0.612
$\Psi_1^{(33)}$	Latent roots		323.7	75.1	37.5	27.9	16.7
	Contribution rate		0.673	0.156	0.078	0.058	0.035
	Cumulative contribution rate		0.673	$\underline{0.829}$	0.907	0.965	1.000
	Latent vectors	$x_1^{(3)}$	−0.362	−0.011	−0.366	0.275	0.812
		$x_2^{(3)}$	−0.315	−0.077	0.331	0.842	−0.278
		$x_3^{(3)}$	−0.190	0.564	−0.677	0.108	−0.419
		$x_4^{(3)}$	0.185	−0.778	−0.542	0.139	−0.219
		$x_5^{(3)}$	0.837	0.266	−0.068	0.430	0.200

Table 2: Population MPCCC using $q_1 = q_2 = q_3 = 2$ and MCCC using all variables ($p_1 = p_2 = p_3 = 5$) in Case 1.

MPCCC			MCCC					
$q_1 = q_2 = q_3 = 2$			$p_1 = p_2 = p_3 = 5$					
	$b = 1$	$b = 2$		$b = 1$	$b = 2$	$b = 3$	$b = 4$	$b = 5$
SUMCOR								
$\rho_b^{(12)*}$	0.566	0.586	$\rho_b^{(12)}$	0.668	0.574	0.350	0.015	0.010
$\rho_b^{(13)*}$	0.731	0.697	$\rho_b^{(13)}$	0.787	0.806	0.181	0.174	0.047
$\rho_b^{(23)*}$	0.671	0.577	$\rho_b^{(23)}$	0.663	0.617	0.328	0.344	0.106
SSQCOR								
$\rho_b^{(12)*}$	0.563	0.589	$\rho_b^{(12)}$	0.664	0.573	0.313	0.112	0.035
$\rho_b^{(13)*}$	0.739	0.689	$\rho_b^{(13)}$	0.804	0.795	0.019	0.174	0.041
$\rho_b^{(23)*}$	0.666	0.582	$\rho_b^{(23)}$	0.663	0.629	0.435	0.250	0.114
SABSCOR								
$\rho_b^{(12)*}$	0.566	0.586	$\rho_b^{(12)}$	0.668	0.574	0.252	0.210	0.022
$\rho_b^{(13)*}$	0.731	0.697	$\rho_b^{(13)}$	0.787	0.806	0.224	0.035	0.062
$\rho_b^{(23)*}$	0.671	0.577	$\rho_b^{(23)}$	0.677	0.618	0.393	0.185	0.120

Table 3: Simulation study results by SUMCOR for MPCCC using $q_1 = q_2 = q_3 = 2$ and MCCC using all the variables ($p_1 = p_2 = p_3 = 5$) in Case 1 (10,000 times).

N			MPCCC $q_1 = q_2 = q_3 = 2$			MCCC $p_1 = p_2 = p_3 = 5$				
			$b = 1$	$b = 2$		$b = 1$	$b = 2$	$b = 3$	$b = 4$	$b = 5$
100	Mean	$r_b^{(12)*}$	0.623	0.505	$r_b^{(12)}$	0.704	0.588	0.346	0.156	0.079
	SD		0.063	0.094		0.048	0.063	0.103	0.095	0.060
	Mean	$r_b^{(13)*}$	0.752	0.651	$r_b^{(13)}$	0.819	0.785	0.290	0.185	0.084
	SD		0.053	0.102		0.045	0.058	0.100	0.100	0.062
	Mean	$r_b^{(23)*}$	0.676	0.558	$r_b^{(23)}$	0.719	0.620	0.400	0.290	0.111
	SD		0.062	0.087		0.053	0.071	0.095	0.112	0.079
150	Mean	$r_b^{(12)*}$	0.613	0.525	$r_b^{(12)}$	0.690	0.585	0.336	0.128	0.067
	SD		0.053	0.070		0.041	0.053	0.095	0.085	0.051
	Mean	$r_b^{(13)*}$	0.747	0.665	$r_b^{(13)}$	0.813	0.787	0.259	0.175	0.074
	SD		0.047	0.077		0.043	0.053	0.092	0.096	0.055
	Mean	$r_b^{(23)*}$	0.670	0.567	$r_b^{(23)}$	0.704	0.619	0.380	0.299	0.107
	SD		0.054	0.071		0.049	0.064	0.086	0.104	0.070
200	Mean	$r_b^{(12)*}$	0.606	0.536	$r_b^{(12)}$	0.684	0.583	0.337	0.107	0.060
	SD		0.049	0.059		0.037	0.047	0.084	0.078	0.045
	Mean	$r_b^{(13)*}$	0.743	0.674	$r_b^{(13)}$	0.807	0.791	0.243	0.173	0.066
	SD		0.044	0.067		0.041	0.049	0.085	0.092	0.050
	Mean	$r_b^{(23)*}$	0.668	0.573	$r_b^{(23)}$	0.698	0.618	0.368	0.305	0.105
	SD		0.049	0.062		0.046	0.058	0.081	0.097	0.065

Table 4: MPCCC using $q_1 = 3$, $q_2 = 4$, $q_3 = 5$, and MCCC using all the variables ($p_1 = 9$, $p_2 = 10$, $p_3 = 11$) in the population and the simulation study results (10,000 times) in Case 2.

			MPCCC $q_1 = 3$, $q_2 = 4$, $q_3 = 5$				MCCC $p_1 = 9$, $p_2 = 10$, $p_3 = 11$		
			$b = 1$	$b = 2$	$b = 3$		$b = 1$	$b = 2$	$b = 3$
Population		$\rho_b^{(12)*}$	0.467	0.448	0.328	$\rho_b^{(12)}$	0.499	0.440	0.366
		$\rho_b^{(13)*}$	0.510	0.382	0.127	$\rho_b^{(13)}$	0.540	0.456	0.273
		$\rho_b^{(23)*}$	0.619	0.380	0.312	$\rho_b^{(23)}$	0.650	0.526	0.405
Simulation	Mean	$r_b^{(12)*}$	0.524	0.441	0.300	$r_b^{(12)}$	0.633	0.540	0.514
$N = 100$	SD		0.069	0.079	0.104		0.053	0.078	0.066
	Mean	$r_b^{(13)*}$	0.545	0.393	0.184	$r_b^{(13)}$	0.648	0.523	0.480
	SD		0.070	0.089	0.102		0.052	0.089	0.068
	Mean	$r_b^{(23)*}$	0.637	0.444	0.307	$r_b^{(23)}$	0.706	0.580	0.535
	SD		0.065	0.090	0.106		0.051	0.089	0.072

Table 5: MPCCC using $q_1 = q_2 = q_3 = 2$, and MCCC using all the variables ($p_1 = p_2 = p_3 = 5$) in the population and the simulation study results (10,000 times) in Case 3.

			MPCCC $q_1 = q_2 = q_3 = 2$			MCCC $p_1 = p_2 = p_3 = 5$	
			$b = 1$	$b = 2$		$b = 1$	$b = 2$
Population		$\rho_b^{(12)*}$	0.406	0.228	$\rho_b^{(12)}$	0.790	0.440
		$\rho_b^{(13)*}$	0.655	0.024	$\rho_b^{(13)}$	0.838	0.516
		$\rho_b^{(23)*}$	0.503	0.064	$\rho_b^{(23)}$	0.851	0.454
Simulation	Mean	$r_b^{(12)*}$	0.424	0.216	$r_b^{(12)}$	0.804	0.494
$N = 100$	SD		0.091	0.099		0.036	0.077
	Mean	$r_b^{(13)*}$	0.654	0.089	$r_b^{(13)}$	0.846	0.567
	SD		0.062	0.067		0.030	0.072
	Mean	$r_b^{(23)*}$	0.515	0.098	$r_b^{(23)}$	0.858	0.521
	SD		0.092	0.072		0.026	0.072

Table 6: MPCCC using $q_1 = 5$, $q_2 = 6$, $q_3 = 7$, and MCCC using all the variables ($p_1 = 9$, $p_2 = 10$, $p_3 = 11$) in the population and the simulation study results (10,000 times) in Case 4.

			MPCCC $q_1 = 5, q_2 = 6, q_3 = 7$				MCCC $p_1 = 9, p_2 = 10, p_3 = 11$		
			$b = 1$	$b = 2$	$b = 3$		$b = 1$	$b = 2$	$b = 3$
Population		$\rho_b^{(12)*}$	0.302	0.358	0.253	$\rho_b^{(12)}$	0.485	0.331	0.412
		$\rho_b^{(13)*}$	0.344	0.284	0.147	$\rho_b^{(13)}$	0.372	0.361	0.245
		$\rho_b^{(23)*}$	0.355	0.232	0.162	$\rho_b^{(23)}$	0.363	0.345	0.332
Simulation	Mean	$r_b^{(12)*}$	0.488	0.399	0.312	$r_b^{(12)}$	0.605	0.525	0.477
$N = 100$	SD		0.080	0.086	0.090		0.060	0.069	0.069
	Mean	$r_b^{(13)*}$	0.465	0.371	0.286	$r_b^{(13)}$	0.570	0.485	0.453
	SD		0.077	0.084	0.085		0.058	0.078	0.067
	Mean	$r_b^{(23)*}$	0.469	0.387	0.314	$r_b^{(23)}$	0.578	0.503	0.470
	SD		0.075	0.080	0.083		0.058	0.074	0.066

Table 7: MPCCC using $q_1 = q_2 = 2$, $q_3 = q_4 = 3$, and MCCC using all the variables ($p_1 = 5$, $p_2 = 6$, $p_3 = 7$, $p_4 = 8$) in the population and the simulation study results (10,000 times) in Case 5.

		MPCCC			MCCC		
		$q_1 = q_2 = 2$, $q_3 = q_4 = 3$			$p_1 = 5$, $p_2 = 6$, $p_3 = 7$, $p_4 = 8$		
			$b = 1$	$b = 2$		$b = 1$	$b = 2$
Population		$\rho_b^{(12)*}$	0.725	0.388	$\rho_b^{(12)}$	0.748	0.417
		$\rho_b^{(13)*}$	0.519	0.355	$\rho_b^{(13)}$	0.523	0.410
		$\rho_b^{(14)*}$	0.439	0.269	$\rho_b^{(14)}$	0.456	0.283
		$\rho_b^{(23)*}$	0.357	0.212	$\rho_b^{(23)}$	0.419	0.310
		$\rho_b^{(24)*}$	0.546	0.450	$\rho_b^{(24)}$	0.588	0.499
		$\rho_b^{(34)*}$	0.728	0.528	$\rho_b^{(34)}$	0.725	0.538
Simulation	Mean	$r_b^{(12)*}$	0.720	0.371	$r_b^{(12)}$	0.707	0.418
$N = 100$	SD		0.128	0.161		0.168	0.149
	Mean	$r_b^{(13)*}$	0.521	0.342	$r_b^{(13)}$	0.567	0.413
	SD		0.097	0.113		0.081	0.120
	Mean	$r_b^{(14)*}$	0.450	0.259	$r_b^{(14)}$	0.520	0.356
	SD		0.096	0.094		0.083	0.097
	Mean	$r_b^{(23)*}$	0.381	0.209	$r_b^{(23)}$	0.502	0.406
	SD		0.096	0.096		0.081	0.101
	Mean	$r_b^{(24)*}$	0.553	0.437	$r_b^{(24)}$	0.641	0.610
	SD		0.089	0.098		0.075	0.122
	Mean	$r_b^{(34)*}$	0.696	0.520	$r_b^{(34)}$	0.722	0.566
	SD		0.132	0.139		0.092	0.128

Table 8: Mean of the computation time to calculate MPCCA (Unit: Second).

k	$N \setminus q$	2	3	4	5
3	100	0.7	4.7	9.5	16.6
	150	0.8	5.6	11.4	19.0
	200	1.0	6.4	15.1	25.8
4	100	1.8	6.6	13.3	25.3
	150	2.1	7.4	14.6	30.2
	200	2.3	8.4	18.4	36.5

Table 9: Results of PCA in educational data. The underline showed the 70% cumulative contribution rate.

			1st	2nd	3rd	4th	5th
$T^{(11)}$	Latent roots		872.6	408.4	289.5	152.3	121.1
	Contribution rate		0.473	0.221	0.157	0.083	0.066
	Cumulative contribution rate		0.473	0.695	$\underline{0.852}$	0.934	1.000
	Latent vectors	$x_1^{(1)}$	0.498	−0.123	0.552	0.313	−0.577
		$x_2^{(1)}$	0.498	−0.408	0.064	0.258	0.717
		$x_3^{(1)}$	0.408	0.512	0.333	−0.645	0.211
		$x_4^{(1)}$	0.418	0.583	−0.551	0.422	−0.061
		$x_5^{(1)}$	0.403	−0.465	−0.526	−0.491	−0.322
$T^{(22)}$	Latent roots		954.7	363.1	233.0	212.5	115.0
	Contribution rate		0.508	0.193	0.124	0.113	0.061
	Cumulative contribution rate		0.508	$\underline{0.702}$	0.826	0.939	1.000
	Latent vectors	$x_1^{(2)}$	0.455	−0.339	−0.322	0.320	−0.687
		$x_2^{(2)}$	0.474	−0.456	−0.029	0.299	0.691
		$x_3^{(2)}$	0.355	0.611	−0.670	−0.129	0.188
		$x_4^{(2)}$	0.458	0.522	0.628	0.339	−0.090
		$x_5^{(2)}$	0.482	−0.178	0.229	−0.822	−0.083
$T^{(33)}$	Latent roots		1698.2	949.7	489.4	356.0	276.5
	Contribution rate		0.450	0.252	0.130	0.094	0.073
	Cumulative contribution rate		0.450	$\underline{0.702}$	0.832	0.927	1.000
	Latent vectors	$x_1^{(3)}$	0.050	0.540	−0.409	−0.658	0.325
		$x_2^{(3)}$	−0.118	0.423	−0.264	0.728	0.455
		$x_3^{(3)}$	0.859	−0.181	−0.445	0.144	−0.099
		$x_4^{(3)}$	0.489	0.264	0.747	−0.030	0.364
		$x_5^{(3)}$	0.076	0.653	0.085	0.125	−0.738

Table 10: MPCCC using $q_1 = 3$ and $q_2 = q_3 = 2$, and MCCC using all variables ($p_1 = p_2 = p_3 = 5$) in educational data.

MPCCC $q_1 = 3$, $q_2 = q_3 = 2$			MCCC $p_1 = p_2 = p_3 = 5$					
	$b = 1$	$b = 2$		$b = 1$	$b = 2$	$b = 3$	$b = 4$	$b = 5$
$r_b^{(12)*}$	0.800	0.646	$r_b^{(12)}$	0.865	0.711	0.496	0.132	0.088
$r_b^{(13)*}$	0.549	0.454	$r_b^{(13)}$	0.571	0.571	0.303	0.147	0.058
$r_b^{(23)*}$	0.571	0.615	$r_b^{(23)}$	0.586	0.662	0.496	0.168	0.013

Table 11: Coefficients of the canonical variables in MPCCA and MCCA.

MPCCA $q_1 = 3$, $q_2 = q_3 = 2$			MCCA $p_1 = p_2 = p_3 = 5$					
	$b = 1$	$b = 2$		$b = 1$	$b = 2$	$b = 3$	$b = 4$	$b = 5$
$\alpha_{b1}^{(1)*}$	0.033	−0.009	$\alpha_{b1}^{(1)}$	0.017	−0.014	−0.044	−0.047	−0.009
$\alpha_{b2}^{(1)*}$	0.012	0.041	$\alpha_{b2}^{(1)}$	0.021	−0.025	0.029	0.045	−0.039
$\alpha_{b3}^{(1)*}$	−0.007	−0.029	$\alpha_{b3}^{(1)}$	0.010	0.025	−0.024	0.048	0.027
			$\alpha_{b4}^{(1)}$	0.025	0.026	0.031	−0.030	−0.011
			$\alpha_{b5}^{(1)}$	−0.002	−0.019	0.006	−0.010	0.060
$\alpha_{b1}^{(2)*}$	0.031	−0.009	$\alpha_{b1}^{(2)}$	0.029	−0.014	−0.061	−0.027	−0.006
$\alpha_{b2}^{(2)*}$	0.014	0.050	$\alpha_{b2}^{(2)}$	−0.001	−0.033	0.032	0.056	−0.011
			$\alpha_{b3}^{(2)}$	0.023	0.033	−0.001	0.033	0.027
			$\alpha_{b4}^{(2)}$	0.013	0.016	0.016	−0.019	−0.047
			$\alpha_{b5}^{(2)}$	0.007	−0.010	0.030	0.037	0.037
$\alpha_{b1}^{(3)*}$	0.038	−0.031	$\alpha_{b1}^{(3)}$	0.042	−0.028	−0.071	−0.008	0.037
$\alpha_{b2}^{(3)*}$	0.041	0.050	$\alpha_{b2}^{(3)}$	0.008	−0.030	0.056	0.060	0.051
			$\alpha_{b3}^{(3)}$	0.026	0.020	−0.005	0.047	−0.024
			$\alpha_{b4}^{(3)}$	0.021	0.030	0.031	−0.047	0.053
			$\alpha_{b5}^{(3)}$	0.024	−0.030	0.031	−0.026	−0.082