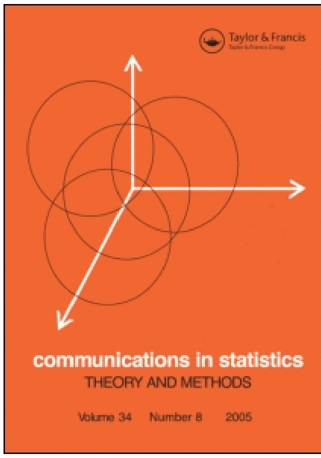

Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Geometry in Canonical Correlations

Maria Adam [a]; John Maroulas [a]
[a] Department of Mathematics, National Technical University, Athens, Greece

Online Publication Date: 01 December 2006

To cite this Article: Adam, Maria and Maroulas, John (2006) 'Geometry in Canonical Correlations', Communications in Statistics - Theory and Methods, 35:12, 2263 — 2273

To link to this article: DOI: 10.1080/03610920600853522
URL: http://dx.doi.org/10.1080/03610920600853522

Communications in Statistics—Theory and Methods, 35: 2263–2273, 2006

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920600853522

Multivariate Analysis

Geometry in Canonical Correlations

MARIA ADAM AND JOHN MAROULAS

Department of Mathematics, National Technical University,
Athens, Greece

In this article, a useful proposition relating the canonical correlations in multi-way layout to the singular values of a specific matrix is proved and also a geometrical explanation of canonical correlations as an angle between subspaces is given.

Keywords Canonical correlation; Projectors.

Mathematics Subject Classification 62H20; 15A18.

1. Introduction

This article continues our search of canonical correlations in multi-way layout, where some results have been published in Adam and Maroulas (2004). Our approach is to give a geometrical interpretation of canonical correlations of k-factors in a set of observations, by generalizing the results in Styan (1986) for two- and three-way layout. As we have noted in Adam and Maroulas (2004), the statistical model is written as

$$E(\mathbf{y}) = X_1\mathbf{b}_1 + X_2\mathbf{b}_2 + \cdots + X_k\mathbf{b}_k,$$

where $\mathbf{y}$ is a $n \times 1$ vector of all the observations, the vectors $\mathbf{b}_i$ $(i = 1, 2, \ldots, k)$ consist of the effects (row, column, treatment, etc.), and X_i are $n \times n_i$ *design matrices* identifying the correspondence between the elements of $\mathbf{y}$ and $\mathbf{b}_i$ of the k-way layout. Following, we consider as in Adam and Maroulas (2004) the types of canonical correlations:

(i) $\rho_h^{(1*2*\cdots*k)}$ between $X_1^T\mathbf{y}$, $X_2^T\mathbf{y}, \ldots,$ and $X_k^T\mathbf{y}$, (k factors total)
(ii) $\rho_h^{(1*2*\cdots*k\,|\,s)}$ between $X_1^TP_s\mathbf{y}, X_2^TP_s\mathbf{y}, \ldots,$ and $X_k^TP_s\mathbf{y}$, with $s \neq 1, 2, \ldots, k$ (several factors each adjusted for the sth factor)
(iii) $\rho_h^{(1*2*\cdots*k\|s)}$ between $X_1^T\mathbf{y}, X_2^T\mathbf{y}, \ldots, X_{k-1}^T\mathbf{y}$ and $X_k^TP_s\mathbf{y}$, with $s \neq 1, 2, \ldots, k$ (partial adjusted for the factor s)

Received September 23, 2005; Accepted March 17, 2006
Address correspondence to John Maroulas, Department of Mathematics, National Technical University, Athens 15780, Greece; E-mail: maroulas@math.ntua.gr.

where in (ii) and (iii) the matrix P_s is an orthogonal projection matrix onto the null space of X_s^T. The sign $*$ corresponds to the partition of $n \times (n_1 + \cdots + n_k)$ design matrix X in block submatrices, and $h = 1, \ldots, m$. The upper limit m of the index h denotes the number of positive canonical correlations and we write $h = 0$, when the associated vectors are uncorrelated. In Styan (1983), the inequalities $0 \leq \rho_m \leq \cdots \leq \rho_1 \leq 1$ are presented and the relationships among the numbers u, t, m have been referred, where u denotes the number of canonical correlations of unit value, t is the number of positive canonical correlations less than 1, and $m = u + t$ is equal to the number of positive canonical correlations.

In this article, we calculate the canonical correlations (i)–(iii) in k-way layout via the singular values of a suitable matrix depending on $X_1, X_2, \ldots, X_k$ and also it is proved that any type of canonical correlations is equal to the cosine of the minimum principal angle between suitable subspaces, where the columns of X_i are involved.

2. Results

Let $X = [X_1 \ldots X_k]$ be the design matrix for the k-way layout, where the submatrices X_i, $(i = 1, \ldots, k)$ are of dimensions $n \times n_i$, with $n \geq n_1 + \cdots + n_k$. Without loss of generality we assume that $rank(X_i) = dim\, R(X_i) = n_i$, where $R(X_i)$ stands for the range of X_i (i.e., the subspace spanned by the columns of X_i). Furthermore, we remind that

$$H_i = X_i(X_i^T X_i)^{-1} X_i^T, \quad (i = 1, 2, \ldots, k) \tag{1}$$

is an $n \times n$ orthogonal projector onto $R(X_i)$, assuming that $\det(X_i^T X_i) > 0$. A more general form of this projector is

$$H_i = X_i X_i^- \tag{2}$$

where the superscript "−" is referred to a least squares generalized inverse, in the sense that $XX^-X = X$ and $X^-XX^- = X^-$, (see Lancaster and Tismenetsky, 1984, p. 430). If $rank(X_i) < n_i$, clearly in (1) the $(X_i^T X_i)^{-1}$ has to be $(X_i^T X_i)^-$.

Furthermore, we denote

$$H_{j\,|\,s} = P_s X_j\,(X_j^T P_s X_j)^- X_j^T P_s, \quad s \neq j \tag{3}$$

the orthogonal projection onto $R(P_s X_j)$, where $P_s = I - H_s$ is a symmetric, idempotent projector matrix. If $P_s X_j$ has full column rank, then $(X_j^T P_s X_j)^-$ has to be $(X_j^T P_s X_j)^{-1}$.

Denoting by $ch_h(\cdot)$ the hth largest real eigenvalue, in two-way layout, it is known in Khatri (1976) and Bérubé et al. (1993) that the canonical correlation is given by

$$\rho_h^{(i*j)} = ch_h^{1/2}(H_i H_j).$$

In multi-way layout the canonical correlation of type (i) and (ii) have been defined in Adam and Maroulas (2004) as

$$\rho_h^{(1*2*\cdots*k)} = ch_h^{1/2}(H_1 H_2 \ldots H_k) \tag{4}$$

and

$$\rho_h^{(1*2*\cdots*k\,|\,s)} = ch_h^{1/2}(H_{1\,|\,s}H_{2\,|\,s}\dots H_{k\,|\,s}). \tag{5}$$

If we apply the QR factorization to matrices $X_1, X_2, \dots, X_k$ we have:

$$X_i = Q_i T_i \tag{6}$$

where Q_i are $n \times n_i$ orthogonal matrices ($Q_i^T Q_i = I_{n_i}$), and T_i are $n_i \times n_i$ non singular upper triangular matrices, $i = 1, 2, \dots, k$. Clearly, the n_i columns of Q_i form an orthonormal basis for $R(X_i)$.

Lemma 2.1. *The orthogonal projector H_i in* (1) *or* (2). *is equal to*

$$H_i = Q_i Q_i^T. \tag{7}$$

Proof. By the QR factorization of matrix X_i in (6), the H_i is written in the form

$$\begin{aligned} H_i &= X_i(X_i^T X_i)^{-1}X_i^T = Q_i T_i (T_i^T Q_i^T Q_i T_i)^{-1} T_i^T Q_i^T \\ &= Q_i T_i T_i^{-1} (T_i^T)^{-1} T_i^T Q_i^T = Q_i Q_i^T. \end{aligned}$$

□

Björck and Golub have proved that the canonical correlations of type (i), only for two-way layout, are connected by the eigenvalues of an eigenvalue problem. Generalizing this result, we say the following.

Proposition 2.1. *Let $\mathcal{H} = H_2 H_3 \dots H_{k-1}$. The eigenvalues of the generalized eigenvalue problem*

$$\begin{bmatrix} O & X_1^T \mathcal{H} X_k \\ X_k^T X_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} X_1^T X_1 & O \\ O & X_k^T X_k \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \tag{8}$$

are equal to the canonical correlations $\rho_h^{(1*2*\cdots*k)}$.

Proof. Since $diag(X_1^T X_1, X_k^T X_k)$ is non singular, Eq. (8) is equivalent to

$$\begin{bmatrix} (X_1^T X_1)^{-1} & O \\ O & (X_k^T X_k)^{-1} \end{bmatrix} \begin{bmatrix} O & X_1^T \mathcal{H} X_k \\ X_k^T X_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix},$$

and then

$$\begin{bmatrix} \lambda I_{n_1} & -(X_1^T X_1)^{-1} X_1^T \mathcal{H} X_k \\ -(X_k^T X_k)^{-1} X_k^T X_1 & \lambda I_{n_k} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

Moreover, we have,

$$\begin{aligned} \det &\begin{bmatrix} \lambda I_{n_1} & -(X_1^T X_1)^{-1} X_1^T \mathcal{H} X_k \\ -(X_k^T X_k)^{-1} X_k^T X_1 & \lambda I_{n_k} \end{bmatrix} \\ &= \det(\lambda I_{n_1}) \cdot \det\big[\lambda I_{n_k} - \lambda^{-1}(X_k^T X_k)^{-1} X_k^T X_1 (X_1^T X_1)^{-1} X_1^T \mathcal{H} X_k\big] \end{aligned}$$

$$= \lambda^{n_1-n_k} \cdot \det\big[\lambda^2 I_{n_k} - (X_k^T X_k)^{-1} X_k^T X_1 (X_1^T X_1)^{-1} X_1^T \mathscr{H} X_k\big]$$
$$= \lambda^{n_1-n_k} \cdot \det\big[\lambda^2 I_{n_k} - (X_k^T X_k)^{-1} X_k^T H_1 \mathscr{H} X_k\big],$$

i.e., the non zero eigenvalues of the eigenvalue problem (8) are roots of the equation

$$\det\big[\lambda^2 I_{n_k} - (X_k^T X_k)^{-1} X_k^T H_1 \mathscr{H} X_k\big] = 0.$$

On the other hand, using the relationship $\det(\lambda I_n - AB) = \lambda^{n-m} \cdot \det(\lambda I_m - BA)$ for the $n \times m$ matrix A and the $m \times n$ matrix B, we obtain

$$\det(\lambda^2 I_n - H_1 \mathscr{H} H_k) = \det\big[\lambda^2 I_n - H_1 \mathscr{H} X_k (X_k^T X_k)^{-1} X_k^T\big]$$
$$= \lambda^{2(n-n_k)} \cdot \det\big[\lambda^2 I_{n_k} - (X_k^T X_k)^{-1} X_k^T H_1 \mathscr{H} X_k\big] = 0.$$

Clearly by (4), the canonical correlations $\rho_h^{(1*2*\cdots*k)}$ are identified by the eigenvalues of the generalized eigenvalue problem (8). □

Using the simplified symbol ρ_h for the $\rho_h^{(1*2*\cdots*k)}$, the ρ_h^+ denotes the positive canonical correlations and by Proposition 2.1, for k-way layout we have

$$\rho_h^+ = \{\lambda : \det(\lambda^2 I - H_1 \mathscr{H} H_k) = 0\} \cap \mathbb{R}^+.$$

The quantities $1 - \rho_h^+$ are called *canonical efficiency factors* (see James and Wilkinson, 1971).

Proposition 2.2. *Let rank*$(\mathscr{H} X_k) = n_k$, *and let* $\mathscr{H} X_k = WZ$ *be its QR factorization. The absolute values of non zero eigenvalues of the generalized eigenvalue problem*

$$\begin{bmatrix} O & X_1^T \mathscr{H} X_k \\ X_k^T \mathscr{H}^T X_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} X_1^T X_1 & O \\ O & X_k^T \mathscr{H}^T \mathscr{H} X_k \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \tag{9}$$

are identified by the singular values of $n_1 \times n_k$ *matrix* $Q_1^T W$.

Proof. Using the QR factorizations of matrices X_1, and $\mathscr{H} X_k$, the generalized eigenvalue problem in (9) is written as

$$\begin{bmatrix} O & T_1^T Q_1^T W Z \\ Z^T W^T Q_1 T_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} T_1^T Q_1^T Q_1 T_1 & O \\ O & Z^T W^T W Z \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}. \tag{10}$$

Since $Q_1^T Q_1 = I_{n_1}$, $W^T W = I_{n_k}$, and the matrix $diag(T_1, Z)$ is non singular, by Eq. (10) we take

$$\begin{bmatrix} O & Q_1^T W \\ W^T Q_1 & O \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{a}} \\ \tilde{\mathbf{b}} \end{bmatrix} = \lambda \begin{bmatrix} \tilde{\mathbf{a}} \\ \tilde{\mathbf{b}} \end{bmatrix},$$

where $\tilde{\mathbf{a}} = T_1 \mathbf{a}$, $\tilde{\mathbf{b}} = Z \mathbf{b}$. Thus,

$$Q_1^T W \tilde{\mathbf{b}} = \lambda \tilde{\mathbf{a}}, \quad W^T Q_1 \tilde{\mathbf{a}} = \lambda \tilde{\mathbf{b}},$$

and it is evident that $|\lambda|$ is singular value of $Q_1^T W$, since these relationships lead to

$$(Q_1^T WW^T Q_1)\tilde{\mathbf{a}} = \lambda(Q_1^T W)\tilde{\mathbf{b}} = \lambda^2\tilde{\mathbf{a}}, \quad (W^T Q_1 Q_1^T W)\tilde{\mathbf{b}} = \lambda(W^T Q_1)\tilde{\mathbf{a}} = \lambda^2\tilde{\mathbf{b}}. \qquad \square$$

Proposition 2.3. *Let the matrices H_i and H_j commute for $i, j = 1, 2, \ldots, k-1$, and $\mathscr{H}X_k = WZ$, $X_k = Q_k T_k$ are their QR factorizations. If the non singular matrices Z and T_k are related by $Z^T Z = T_k^T T_k$, the singular values of $Q_1^T W$ are equal to the positive canonical correlations $\rho_h^{(1*2*\cdots*k)}$.*

Proof. Since $diag(X_1^T X_1, X_k^T\mathscr{H}^T\mathscr{H}X_k)$ is non singular, by (9) we have

$$\begin{bmatrix} O & (X_1^T X_1)^{-1}X_1^T\mathscr{H}X_k \\ (X_k^T\mathscr{H}^T\mathscr{H}X_k)^{-1}X_k^T\mathscr{H}^T X_1 & O \end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \lambda\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix},$$

i.e., the eigenvalues of problem (9) are computed by the equation

$$\det\begin{bmatrix} \lambda I_{n_1} & -(X_1^T X_1)^{-1}X_1^T\mathscr{H}X_k \\ -(X_k^T\mathscr{H}^T\mathscr{H}X_k)^{-1}X_k^T\mathscr{H}^T X_1 & \lambda I_{n_k} \end{bmatrix} = 0.$$

Hence, the non zero eigenvalues, as in Proposition 2.1, are given by

$$\det\left[\lambda^2 I_{n_k} - (X_k^T\mathscr{H}^T\mathscr{H}X_k)^{-1}X_k^T\mathscr{H}^T H_1\mathscr{H}X_k\right] = 0$$

or by

$$\det\left[\lambda^2 I_n - \mathscr{H}X_k(X_k^T\mathscr{H}^T\mathscr{H}X_k)^{-1}X_k^T\mathscr{H}^T H_1\right] = 0.$$

The commutativity of symmetric matrices H_i and H_j implies that $\mathscr{H} = \mathscr{H}^T$, and $\mathscr{H}^2 = \mathscr{H}$. Then, by Lemma 2.1 and the relationship $ZT_k^{-1} = Z^{-T}T_k^T$ we have:

$$\begin{aligned}
\det\left[\lambda^2 I_n - H_1\mathscr{H}H_k\right] &= \det\left[\lambda^2 I_n - H_1\mathscr{H}^2 H_k\right] = \det\left[\lambda^2 I_n - \mathscr{H}H_k\mathscr{H}H_1\right] \\
&= \det\left[\lambda^2 I_n - \mathscr{H}Q_k Q_k^T\mathscr{H}H_1\right] = \det\left[\lambda^2 I_n - \mathscr{H}Q_k T_k T_k^{-1}Q_k^T\mathscr{H}^T H_1\right] \\
&= \det\left[\lambda^2 I_n - \mathscr{H}X_k Z^{-1}ZT_k^{-1}Q_k^T\mathscr{H}^T H_1\right] \\
&= \det\left[\lambda^2 I_n - WZ^{-T}T_k^T Q_k^T\mathscr{H}^T H_1\right] \\
&= \det\left[\lambda^2 I_n - WW^T H_1\right] = \det\left[\lambda^2 I_n - WW^T Q_1 Q_1^T\right] \\
&= \lambda^{2(n-n_1)}\cdot\det\left[\lambda^2 I_{n_1} - Q_1^T WW^T Q_1\right].
\end{aligned}$$

By the last equality and Propositions 2.1 and 2.2, it is implied that the non zero singular values of $Q_1^T W$ are equal to the positive real numbers $\rho_h^{(1*2*\cdots*k)}$. $\square$

Furthermore, we present a geometric approach for $\rho_h^{(1*2*\cdots*k)}$. We remember that if F and G are subspaces of $\mathbb{R}^n$, then the angle $\theta \in [0, \pi/2]$ between the subspaces is defined by the relationship

$$\cos\theta = \max_{\mathbf{u}\in F, \mathbf{v}\in G} \mathbf{u}\cdot\mathbf{v},$$

where $\mathbf{u}, \mathbf{v}$ are unit vectors and $\cdot$ denotes the scalar product. Additionally, if $q = min\{dim\, F, dim\, G\} \geq 1$, the *principal angles* $0 \leq \theta_1 \leq \theta_2 \leq \cdots \leq \theta_q \leq \frac{\pi}{2}$ between F and G are defined, as in Björck and Golub (1973), by

$$\begin{aligned}
\cos\theta_1 &= \max_{\mathbf{u}\in F, \mathbf{v}\in G} \mathbf{u}\cdot\mathbf{v} = \mathbf{u}_1\cdot\mathbf{v}_1,\\
\cos\theta_2 &= \max_{\mathbf{u}\in F_1, \mathbf{v}\in G_1} \mathbf{u}\cdot\mathbf{v} = \mathbf{u}_2\cdot\mathbf{v}_2, \quad \text{where} \quad F_1 = F\cap[span\{\mathbf{u}_1\}]^{\perp},\\
&\qquad\qquad G_1 = G\cap[span\{\mathbf{v}_1\}]^{\perp},\\
\cos\theta_3 &= \max_{\mathbf{u}\in F_2, \mathbf{v}\in G_2} \mathbf{u}\cdot\mathbf{v} = \mathbf{u}_3\cdot\mathbf{v}_3, \quad \text{where} \quad F_2 = F\cap[span\{\mathbf{u}_1,\mathbf{u}_2\}]^{\perp},\\
&\qquad\qquad G_2 = G\cap[span\{\mathbf{v}_1,\mathbf{v}_2\}]^{\perp},\\
&\ \vdots \\
\cos\theta_q &= \max_{\mathbf{u}\in F_{q-1}, \mathbf{v}\in G_{q-1}} \mathbf{u}\cdot\mathbf{v} = \mathbf{u}_q\cdot\mathbf{v}_q, \quad \text{where} \quad F_{q-1} = F\cap[span\{\mathbf{u}_1,\ldots,\mathbf{u}_{q-1}\}]^{\perp},\\
&\qquad\qquad G_{q-1} = G\cap[span\{\mathbf{v}_1,\ldots,\mathbf{v}_{q-1}\}]^{\perp}.
\end{aligned} \tag{11}$$

Proposition 2.4. *Let the matrices H_i, H_j, for $i, j = 1, 2, \ldots, k-1$ commute and $Z^T Z = T_k^T T_k$, where Z, T_k are the non singular matrices in the QR factorization of matrices $\mathscr{H}X_k$ and X_k. The angle between the subspaces $R(X_1)$ and $R(\mathscr{H}X_k)$ is equal to $min\{\arccos\theta_p : \theta_p$ principal angles, $p = 1, \ldots, \ell\}$ and $\rho_h^+ = \{\cos\theta_p : p = 1, \ldots, \ell\}$.*

Proof. By the QR factorizations of matrices $X_1 = Q_1 T_1$ and $\mathscr{H}X_k = WZ$, it is evident that $R(X_1) = R(Q_1)$ and $R(\mathscr{H}X_k) = R(W)$. Moreover, by the singular value decomposition of $Q_1^T W$ we have

$$Y^T Q_1^T W U = S,$$

where the $n_1 \times n_k$ matrix S has diagonal elements $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_\ell$, with $\ell = rank(Q_1^T W) \leq min\{n_1, n_k\}$ and zeros elsewhere. We consider

$$Q_1 Y = [\mathbf{y}_1 \ldots \mathbf{y}_{n_1}] \quad WU = [\mathbf{z}_1 \ldots \mathbf{z}_{n_k}]. \tag{12}$$

The columns of these matrices are orthogonal basis of subspaces $R(X_1)$ and $R(\mathscr{H}X_k)$, and it is implied that

$$R(X_1) = R(Q_1) = span\{\mathbf{y}_1, \ldots, \mathbf{y}_{n_1}\}, \quad R(\mathscr{H}X_k) = R(W) = span\{\mathbf{z}_1, \ldots, \mathbf{z}_{n_k}\}.$$

By Theorem 1 in Björck and Golub (1973), it is implied that the singular values of $Q_1^T W$ are related by the principal angles $0 \leq \theta_1 \leq \theta_2 \leq \cdots \leq \theta_\ell < \frac{\pi}{2}$, which are given by the equalities

$$\cos\theta_p = \sigma_p(Q_1^T W) = \mathbf{y}_p\cdot\mathbf{z}_p, \quad p = 1, \ldots, \ell,$$

with $\mathbf{y}_p, \mathbf{z}_p$ in (12) and $\cos\theta_1 \geq \cos\theta_2 \geq \cdots \geq \cos\theta_\ell$.

Denoting, as in (11), $\cos\theta_1 = \max_{\mathbf{y}\in R(X_1)\mathbf{z}\in R(\mathscr{H}X_k)} \mathbf{y}\cdot\mathbf{z} = \mathbf{y}_1\cdot\mathbf{z}_1$, and for $s = 1,\ldots,\ell-1$,

$$F_s = R(X_1)\cap[span\{\mathbf{y}_1,\ldots,\mathbf{y}_s\}]^{\perp},\quad G_s = R(\mathscr{H}X_k)\cap[span\{\mathbf{z}_1,\ldots,\mathbf{z}_s\}]^{\perp},$$

it is evident that θ_{s+1} is the smallest angle between F_s and G_s, and by Proposition 2.3, clearly the positive canonical correlations $\rho_h^+ \equiv \{\cos\theta_1,\ldots,\cos\theta_\ell\}$. □

3. Special Cases

Now, we study in an analog way the canonical correlations of type (ii), $\rho_h^{(1*2*\cdots*k\,|\,s)}$ with $s \neq 1,2,\ldots,k$, and we denote

$$\mathscr{H}_s = H_{2|s}H_{3|s}\cdots H_{k-1|s}.$$

Proposition 3.1. *The canonical correlations* $\rho_h^{(1*2*\cdots*k\,|\,s)}$ *are eigenvalues of the generalized eigenvalue problem*

$$\begin{bmatrix} O & X_1^TP_s\mathscr{H}_sX_k \\ X_k^TP_sX_1 & O\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \lambda\begin{bmatrix} X_1^TP_sX_1 & O \\ O & X_k^TP_sX_k\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}, \tag{13}$$

where $P_s = I - H_s$.

Proof. Let $rank(X_1^TP_sX_1) = r_1 \le n_1$, and $X_1^TP_sX_1 = Y_1Y_2$, where the factors Y_1, Y_2 are $n_1\times r_1$ and $r_1\times n_1$ matrices, with $rank(Y_1) = rank(Y_2) = r_1$. Similarly, let $rank(X_k^TP_sX_k) = r_k \le n_k$ and $X_k^TP_sX_k = \Psi_1\Psi_2$, where Ψ_1,Ψ_2 are $n_k\times r_k$ and $r_k\times n_k$ full rank matrices. Then, we have

$$D = \begin{bmatrix} X_1^TP_sX_1 & O \\ O & X_k^TP_sX_k\end{bmatrix} = \begin{bmatrix} Y_1 & O\\ O & \Psi_1\end{bmatrix}\begin{bmatrix} Y_2 & O\\ O & \Psi_2\end{bmatrix} = D_1D_2$$

and even $D_1^- = diag(Y_1^-,\Psi_1^-)$, and $D_2^- = diag(Y_2^-,\Psi_2^-)$, where

$$Y_1^- = (Y_1^TY_1)^{-1}Y_1^T,\quad \Psi_1^- = (\Psi_1^T\Psi_1)^{-1}\Psi_1^T,\quad Y_2^- = Y_2^T(Y_2Y_2^T)^{-1},\quad \Psi_2^- = \Psi_2^T(\Psi_2\Psi_2^T)^{-1}.$$

Then, by (13) we have:

$$\begin{bmatrix} O & X_1^TP_s\mathscr{H}_sX_k \\ X_k^TP_sX_1 & O\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \lambda\begin{bmatrix} Y_1 & O\\ O & \Psi_1\end{bmatrix}\begin{bmatrix} Y_2 & O\\ O & \Psi_2\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}$$

and consequently,

$$\begin{bmatrix} O & Y_1^-X_1^TP_s\mathscr{H}_sX_k \\ \Psi_1^-X_k^TP_sX_1 & O\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \lambda\begin{bmatrix} Y_2 & O\\ O & \Psi_2\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}. \tag{14}$$

Setting

$$\begin{bmatrix} Y_2 & O\\ O & \Psi_2\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix}\hat{\mathbf{a}}\\ \hat{\mathbf{b}}\end{bmatrix} \Longrightarrow \begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix} Y_2^- & O\\ O & \Psi_2^-\end{bmatrix}\begin{bmatrix}\hat{\mathbf{a}}\\ \hat{\mathbf{b}}\end{bmatrix}$$

i.e., $\mathbf{a} = Y_2^- \hat{\mathbf{a}}$, and $\mathbf{b} = \Psi_2^- \hat{\mathbf{b}}$. Hence, Eq. (14) is written

$$\begin{bmatrix} O & Y_1^- X_1^T P_s \mathscr{H}_s X_k \Psi_2^- \\ \Psi_1^- X_k^T P_s X_1 Y_2^- & O \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{bmatrix} = \lambda \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} \lambda I_{r_1} & -Y_1^- X_1^T P_s \mathscr{H}_s X_k \Psi_2^- \\ -\Psi_1^- X_k^T P_s X_1 Y_2^- & \lambda I_{r_k} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{15}$$

Since $Y_2^- Y_1^-$ and $\Psi_2^- \Psi_1^-$ are the Moore Penrose generalized inverses of $Y_1 Y_2$ and $\Psi_1 \Psi_2$, clearly $Y_2^- Y_1^- = (Y_1 Y_2)^-$, and $\Psi_2^- \Psi_1^- = (\Psi_1 \Psi_2)^-$. Hence, we have

$$\begin{aligned} &\det\begin{bmatrix} \lambda I_{r_1} & -Y_1^- X_1^T P_s \mathscr{H}_s X_k \Psi_2^- \\ -\Psi_1^- X_k^T P_s X_1 Y_2^- & \lambda I_{r_k} \end{bmatrix} \\ &= \lambda^{r_1 - r_k} \cdot \det\left[\lambda^2 I_{r_k} - \Psi_1^- X_k^T P_s X_1 Y_2^- Y_1^- X_1^T P_s \mathscr{H}_s X_k \Psi_2^-\right] \\ &= \lambda^{r_1 - r_k} \cdot \det\left[\lambda^2 I_{r_k} - \Psi_1^- X_k^T P_s X_1 \,(Y_1 Y_2)^- X_1^T P_s \mathscr{H}_s X_k \Psi_2^-\right] \\ &= \lambda^{r_1 - r_k} \cdot \det\left[\lambda^2 I_{r_k} - \Psi_1^- X_k^T P_s X_1 (X_1^T P_s X_1)^- X_1^T P_s \mathscr{H}_s X_k \Psi_2^-\right] \\ &= \lambda^{r_1 - r_k} \lambda^{2(r_k - n_k)} \cdot \det\left[\lambda^2 I_{n_k} - X_k^T P_s X_1 (X_1^T P_s X_1)^- X_1^T P_s \mathscr{H}_s X_k \Psi_2^- \Psi_1^-\right] \\ &= \lambda^{r_1 + r_k - 2n_k} \cdot \det\left[\lambda^2 I_{n_k} - X_k^T P_s X_1 (X_1^T P_s X_1)^- X_1^T P_s \mathscr{H}_s X_k (\Psi_1 \Psi_2)^-\right] \\ &= \lambda^{r_1 + r_k - 2n_k} \cdot \det\left[\lambda^2 I_{n_k} - X_k^T P_s X_1 (X_1^T P_s X_1)^- X_1^T P_s \mathscr{H}_s X_k (X_k^T P_s X_k)^-\right] \\ &= \lambda^{r_1 + r_k - 2n_k} \cdot \det\left[\lambda^2 I_{n_k} - (X_k^T P_s X_k)^- X_k^T P_s X_1 (X_1^T P_s X_1)^- X_1^T P_s \mathscr{H}_s X_k\right]. \end{aligned}$$

Clearly, the non zero eigenvalues of the eigenvalue problem (15) are roots of the equation

$$\begin{aligned} &\det\left[\lambda^2 I_{n_k} - (X_k^T P_s X_k)^- X_k^T P_s X_1 (X_1^T P_s X_1)^- X_1^T P_s \mathscr{H}_s X_k\right] \\ &\quad = \det\left[\lambda^2 I_{n_k} - (X_k^T P_s X_k)^- X_k^T P_s \underbrace{P_s X_1 (X_1^T P_s X_1)^- X_1^T P_s}\mathscr{H}_s P_s X_k\right] \\ &\quad = \det\left[(\lambda^2 I_{n_k} - (X_k^T P_s X_k)^- X_k^T P_s H_{1|s} \mathscr{H}_s P_s X_k\right] = 0. \end{aligned}$$

Thus by (5) and

$$\begin{aligned} \det\left[\lambda^2 I_n - H_{1|s} \mathscr{H}_s H_{k|s}\right] &= \det\left[\lambda^2 I_n - H_{1|s} \mathscr{H}_s (P_s X_k)(X_k^T P_s X_k)^- X_k^T P_s\right] \\ &= \lambda^{2(n - n_k)} \cdot \det\left[\lambda^2 I_{n_k} - (X_k^T P_s X_k)^- X_k^T P_s H_{1|s} \mathscr{H}_s P_s X_k\right], \end{aligned}$$

it is implied that the canonical correlations $\rho_h^{(1*2*\cdots*k|s)}$ are identified by the eigenvalues of the generalized eigenvalue problem (13). □

Similarly, if $\operatorname{rank}(X_k^T P_s X_k) = r_k \leq n_k$, there exist full rank matrices Z_1 and Z_2 such that $X_k^T P_s X_k = Z_1 Z_2$, and by (6) using $X_1 = Q_1 T_1$, the following is proved.

Corollary 3.1. *The canonical correlations* $\rho_h^{(1*2*\cdots*k\|s)}$ *of the type* (iii) *are eigenvalues of the generalized eigenvalue problem*

$$\begin{bmatrix} O & X_1^T \mathscr{H} P_s X_k \\ X_k^T P_s X_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} X_1^T X_1 & O \\ O & X_k^T P_s X_k \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \tag{16}$$

where $P_s = I - H_s$ *and* $\mathscr{H} = H_2 H_3 \ldots H_{k-1}$.

If $rank(P_s X_i) = n_i$, $i \neq s, i = 1, 2, \ldots, k$, we consider the QR factorization of matrices $P_s X_i$, and let

$$P_s X_i = \widetilde{Q}_i \widetilde{T}_i \tag{17}$$

with $\widetilde{Q}_i$ being $n \times n_i$ orthogonal matrices ($\widetilde{Q}_i^T \widetilde{Q}_i = I_{n_i}$), $\widetilde{T}_i$ be $n_i \times n_i$ non singularing upper triangular matrices, and $P_s = I - H_s$. Since $P_s^T = P_s = P_s^2$, by (3), and (17) we have

$$H_{i|s} = P_s X_i (X_i^T P_s^T P_s X_i)^{-1} (P_s X_i)^T = \widetilde{Q}_i \widetilde{T}_i (\widetilde{T}_i^T \widetilde{Q}_i^T \widetilde{Q}_i \widetilde{T}_i)^{-1} \widetilde{T}_i^T \widetilde{Q}_i^T$$

i.e.,

$$H_{i|s} = \widetilde{Q}_i \widetilde{Q}_i^T, \quad i \neq s. \tag{18}$$

Analogous conclusions of Propositions 2.2 and 2.3, connecting the singular values of a matrix with the canonical correlations of types (ii), (iii), we present in the following.

Proposition 3.2. *Let* $\mathscr{H}_s X_k$ *be full rank matrix, then*

(a) *The absolute values of non zero eigenvalues of the generalized eigenvalue problem*

$$\begin{bmatrix} O & X_1^T \mathscr{H}_s X_k \\ X_k^T \mathscr{H}_s^T X_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} X_1^T P_s X_1 & O \\ O & X_k^T \mathscr{H}_s^T \mathscr{H}_s X_k \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \tag{19}$$

are singular values of matrix $\widetilde{Q}_1^T \widetilde{W}$, *where* $\widetilde{Q}_1$ *is the* $n \times n_1$ *orthogonal matrix in* (17), *and* $\widetilde{W}$ *is defined by the QR factorization of* $\mathscr{H}_s X_k = \widetilde{W}\widetilde{Z}$, $(s \neq 1, k)$.

(b) *Let the matrices* $H_{i|s} H_{j|s} = H_{j|s} H_{i|s}$, *for* $i, j = 1, 2, \ldots, k-1$, $s \neq i, j, k$ *and* $\widetilde{Z}^T \widetilde{Z} = \widetilde{T}_k^T \widetilde{T}_k$, *where* $\widetilde{T}_k$ *has presented in* (17). *The singular values of* $\widetilde{Q}_1^T \widetilde{W}$ *are equal to the canonical correlations* $\rho_h^{(1*2*\cdots*k\,|\,s)}$ *of type* (ii).

Proof. (a) By $H_{j|s} P_s = P_s H_{j|s} = H_{j|s}$, for $j = 1, 2, \ldots, k-1$, it is evident that $P_s^T \mathscr{H}_s P_s = \mathscr{H}_s$, and Eq. (19) may be written as

$$\begin{bmatrix} O & X_1^T P_s^T \mathscr{H}_s P_s X_k \\ X_k^T P_s^T \mathscr{H}_s^T P_s X_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} X_1^T P_s^T P_s X_1 & O \\ O & X_k^T P_s^T \mathscr{H}_s^T \mathscr{H}_s P_s X_k \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}. \tag{20}$$

Applying the QR factorization to matrices $\mathscr{H}_s X_k = \mathscr{H}_s P_s X_k = \widetilde{W}\widetilde{Z}$ and $P_s X_1 = \widetilde{Q}_1 \widetilde{T}_1$, and following the same statements as in the proof of Proposition 2.2, we conclude that the absolute values of the non zero eigenvalues of problem (19) are also singular values of $\widetilde{Q}_1^T \widetilde{W}$.

(b) Moreover, the eigenvalues of problem (20) are given by the equation

$$\det\begin{bmatrix} \lambda I_{n_1} & -(X_1^T P_s^T P_s X_1)^{-1} X_1^T P_s^T \mathcal{H}_s P_s X_k \\ -(X_k^T P_s^T \mathcal{H}_s^T \mathcal{H}_s P_s X_k)^{-1} X_k^T P_s^T \mathcal{H}_s^T P_s X_1 & \lambda I_{n_k} \end{bmatrix} = 0$$

and then the non zero eigenvalues are computed by

$$\begin{aligned} 0 &= \det\left[\lambda^2 I_{n_k} - (X_k^T P_s^T \mathcal{H}_s^T \mathcal{H}_s P_s X_k)^{-1} X_k^T P_s^T \mathcal{H}_s^T P_s X_1 (X_1^T P_s^T P_s X_1)^{-1} X_1^T P_s^T \mathcal{H}_s P_s X_k\right] \\ &= \det\left[\lambda^2 I_{n_k} - (X_k^T P_s^T \mathcal{H}_s^T \mathcal{H}_s P_s X_k)^{-1} X_k^T P_s^T \mathcal{H}_s^T H_{1|s} \mathcal{H}_s P_s X_k\right] \end{aligned}$$

or by

$$\det\left[\lambda^2 I_n - H_{1|s} \mathcal{H}_s P_s X_k (X_k^T P_s^T \mathcal{H}_s^T \mathcal{H}_s P_s X_k)^{-1} X_k^T P_s^T \mathcal{H}_s^T\right] = 0.$$

Using $\mathcal{H}_s P_s X_k = \widetilde{W}\widetilde{Z}$, the last equation leads to $\det[\lambda^2 I_n - H_{1|s}\widetilde{W}\widetilde{W}^T] = 0$. On the other hand for $\widetilde{W} = \mathcal{H}_s P_s X_k \widetilde{Z}^{-1}$, $\widetilde{Z}^{-1}\widetilde{Z}^{-T} = \widetilde{T}_k^{-1}\widetilde{T}_k^{-T}$, and $P_s X_k = \widetilde{Q}_k \widetilde{T}_k$ we have

$$\begin{aligned} 0 &= \det\left[\lambda^2 I_n - H_{1|s}\widetilde{W}\widetilde{W}^T\right] \\ &= \det\left[\lambda^2 I_n - H_{1|s}\mathcal{H}_s P_s X_k \widetilde{Z}^{-1}\widetilde{Z}^{-T} X_k^T P_s^T \mathcal{H}_s^T\right] \\ &= \det\left[\lambda^2 I_n - H_{1|s}\mathcal{H}_s P_s X_k \widetilde{T}_k^{-1}\widetilde{T}_k^{-T} X_k^T P_s^T \mathcal{H}_s^T\right] \\ &= \det\left[\lambda^2 I_n - H_{1|s}\mathcal{H}_s Q_k Q_k^T \mathcal{H}_s^T\right] \\ &= \det\left[\lambda^2 I_n - \mathcal{H}_s^T H_{1|s}\mathcal{H}_s Q_k Q_k^T\right]. \end{aligned}$$

Equations (5), (18), and the commutativity of the matrices $H_{j|s}$ complete the proof of the statement (b). □

In the next corollary recalling $\mathcal{H} = H_2 H_3 \ldots H_{k-1}$, we study the relationship of canonical correlations of type (iii) with the singular values of suitable matrix.

Corollary 3.2. *If the matrix $\mathcal{H} P_s X_k$ has full rank, then*

(a) *The absolute values of nonzero eigenvalues of the generalized eigenvalue problem*

$$\begin{bmatrix} O & X_1^T \mathcal{H} P_s X_k \\ X_k^T P_s \mathcal{H}^T X_1 & O \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} X_1^T X_1 & O \\ O & X_k^T P_s \mathcal{H}^T \mathcal{H} P_s X_k \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \qquad (21)$$

are singular values of the matrix $Q_1^T \mathcal{W}$, where Q_1 is $n \times n_1$ orthogonal matrix in (6), *and $\mathcal{H} P_s X_k = \mathcal{W}\mathcal{Z}$, with $\mathcal{W}$ is $n \times n_k$ orthogonal matrix and $\mathcal{Z}$ is $n_k \times n_k$ non singular matrix, $(s \neq k)$.*

(b) *Let the matrices $H_i H_j = H_j H_i$, for every $i, j = 1, 2, \ldots, k-1$, and $\mathcal{Z}^T\mathcal{Z} = \widetilde{T}_k^T \widetilde{T}_k$, where $P_s X_k = \widetilde{Q}_k \widetilde{T}_k$ as in* (17). *The singular values of $Q_1^T \mathcal{W}$ are equal to the canonical correlations $\rho_h^{(1*2*\cdots*k\|s)}$.*

Similarly, by the proof of Proposition 2.4 we conclude the next corollaries.

Corollary 3.3. *Let* $H_{i|s}H_{j|s} = H_{j|s}H_{i|s}$, *for* $i, j = 1, 2, \ldots, k-1$, $s \neq i, j, k$, *and* $\widetilde{Z}^T\widetilde{Z} = \widetilde{T}_k^T\widetilde{T}_k$, *where* $\widetilde{Z}$, $\widetilde{T}_k$ *are the non singular matrices in the QR factorizations of matrices* $\mathscr{H}_sX_k$, *and* P_sX_k, *respectively. The angle between the subspaces* $R(P_sX_1)$ *and* $R(\mathscr{H}_sX_k)$ *is equal to* $\arccos(\rho_h^{(1*2*\cdots*k\,|\,s)})$.

Proof. By (17), and $\mathscr{H}_sX_k = \mathscr{H}_sP_sX_k = \widetilde{W}\widetilde{Z}$, it is evident that $R(P_sX_1) = R(\widetilde{Q}_1)$, and $R(\mathscr{H}_sX_k) = R(\widetilde{W})$. Hence, by Proposition 3.2, and in the same process as in Proposition 2.4, we obtain the result. □

Corollary 3.4. *Let* $H_iH_j = H_jH_i$, *for* $i, j = 1, 2, \ldots, k-1$, *and* $\mathscr{Z}^T\mathscr{Z} = \widetilde{T}_k^T\widetilde{T}_k$, *where* $\mathscr{Z}$, $\widetilde{T}_k$ *are the non singular matrices in the QR factorizations of matrices* $P_sX_k = \widetilde{Q}_k\widetilde{T}_k$, $s \neq k$, *as in* (17), *and* $\mathscr{H}P_sX_k = \mathscr{W}\mathscr{Z}$, *with* $\mathscr{W}$ *is* $n \times n_k$ *orthogonal matrix. The angle between the subspaces* $R(X_1)$ *and* $R(\mathscr{H}P_sX_k)$ *is equal to* $\arccos(\rho_h^{(1*2*\cdots*k\|s)})$.

Proof. By the equations $X_1 = Q_1T_1$ as in (6), and $\mathscr{H}P_sX_k = \mathscr{W}\mathscr{Z}$, it is obvious that $R(X_1) = R(Q_1)$, and $R(\mathscr{H}P_sX_k) = R(\mathscr{W})$. Thus, by Corollary 3.2 and in the same way as in Proposition 2.4, we obtain the result. □

Acknowledgment

Research supported by RC-NTUA under grant "C. Caratheodory."

References

Adam, M., Maroulas, J. (2004). Canonical correlations in multi-way layout. *Ann. Instit. Statist. Math.* 56:656–666.

Bérubé, J., Hartwig, R., Styan, G. P. H. (1993). On canonical correlations and the degrees of non-orthogonality in the three-way layout. In: Matusita, K., Puri, M. L., Hayakawa, T., eds. *Statist. Sci. Data Anal. Proc. Third Pacific Area Statist. Conf.* VSP, Utrecht, The Netherlands: International Science Publishers, pp. 247–252.

Björck, A., Golub, G. H. (1973). Numerical methods for computing angles between linear subspaces. *Math. Comp.* 27:579–594.

James, A. T., Wilkinson, G. N. (1971). Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance. *Biometrika* 58:279–294.

Khatri, C. G. (1976). A note on multiple and canonical correlations for a singular covariance matrix. *Psychometrika* 41:465–470.

Lancaster, P., Tismenetsky, M. (1984). *The Theory of Matrices*. Orlando: Academic Press.

Styan, G. P. H. (1983). Schur complements and linear statistical models. *Proc. First Tampere Sem. Linear Models*. Department of Mathematical Sciences, University of Tampere, Tampere pp. 37–75.

Styan, G. P. H. (1986). Canonical correlations in the three-way layout. In: Francis, I. S., Manly, B. F. J., Lam, F. C., eds. *Pacific Statist. Congr.* Amsterdam: North-Holland, pp. 433–438.