Canonical Correlation Analysis With Linear Constraints

Haruo Yanai
The National Center for University Entrance Examination
2-19-23 Komaba, Meguro-ku
Tokyo 153, Japan

and

Yoshio Takane
Department of Psychology
McGill University
Montréal Québec, Canada

Submitted by J. K. Baksalary

ABSTRACT

We develop canonical correlation analysis by imposing linear constraints upon parameters corresponding to two sets of variables. The results of our method, which we call CANOLC, are shown in terms of projection operators both orthogonal and oblique. Further, CALC (correspondence analysis with linear constraints) turns out to be a special case of CANOLC.

I. INTRODUCTION

Over the past twenty years, Hotelling's canonical correlation analysis has received much attention. This may be due to the fact that canonical correlation analysis subsumes a number of multivariate techniques, including multiple regression analysis, canonical discriminant analysis, correspondence analysis, etc. Using the theory of generalized inverse (g-inverse) matrices, Khatri [5] has shown that canonical correlation analysis can be extended to the case in which the covariance matrix of two sets of variables may be singular.

Further, in the case of a linear regression problem (i.e., $y = X\beta + \varepsilon$), estimation of β may be done subject to a linear constraint $A\beta = c$. (For

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

example, see [8, pp. 84–88].) In view of this, the problem may duly be applied so as to estimate the unknown parameters involved in canonical correlation analysis, taking some constraints into consideration. More often, natural forms of constraints may follow from specific empirical questions posed by the investigators concerned.

With such a formulation, one specifies the space in which the original parameter vector should lie, and then proceeds to test the constructed hypothesis by finding an appropriate test statistic. However, little work has been done on canonical correlation analysis following this kind of approach, except for the work by Bockenholt and Bockenholt [2], who derived a correspondence analysis by incorporating linear constraints (CALC) on row and column scores of contingency tables.

In this paper, we extend the earlier results and derive general solutions for canonical correlation analysis with linear constraints by employing projection operators (called projectors for simplicity), and show that the method of Lagrange multipliers and the method of orthogonal projectors for finding constrained least squares estimates of unknown paramters in linear regression and CALC turn out to be special cases of our solution. Further, we develop some obtained results so as to express canonical correlation coefficients with linear constraints in terms of oblique projectors.

First, we shall briefly review the algebra of projection operators, and establish some necessary lemmas and theorems.

2. SOME RESULTS ON PROJECTION OPERATORS

Let X and Y be $n \times p$ and $n \times q$ matrices, respectively, where $X'X$ and $Y'Y$ may be singular. The symbols $S(X)$ and $S(Y)$ stand for the subspaces spanned by the column vectors of X and Y, respectively. Further, let P_X and P_Y be orthogonal projectors onto $S(X)$ and $S(Y)$. They are more explicitly written as

$$P_X = X(X'X)^{-}X' \quad \text{and} \quad P_Y = Y(Y'Y)^{-}Y'. \tag{2.1}$$

which are unique for any choices of g-inverse matrices of $X'X$ and $Y'Y$.

Further, it can be shown that

$$Q_X = I_n - P_X \quad \text{and} \quad Q_Y = I_n - P_Y \tag{2.2}$$

are orthogonal projectors onto $S(X)^{\perp}$ and $S(Y)^{\perp}$, i.e., orthocomplement subspaces of $S(X)$ and $S(Y)$. Let Ker X' be the kernel of X', or equivalently the null space of X'. Then Ker $X' = S(Q_X)$.

With regard to the orthogonal projectors, the following relationships hold.

LEMMA 2.1. *Let $P_{(X:Y)}$ be the orthogonal projector onto $S(X:Y)$. Then*

$$P_{(X:Y)} = P_X + P_Y \quad \textit{if and only if} \quad X'Y = 0; \tag{2.3a}$$

$$P_{(X:Y)} = P_X + P_Y - P_X P_Y \quad \textit{if and only if} \quad P_X P_Y = P_Y P_X. \tag{2.3b}$$

Equation (2.3a) is easy to prove. For a proof of (2.3b), see [7].

Next, we consider projectors which are not symmetric.

LEMMA 2.2 [13]. *For a $p \times r$ matrix A and an n.n.d. matrix M of order p, the following three statements are equivalent.*

$$\operatorname{rank}(A'M) = \operatorname{rank} A, \tag{2.4a}$$

$$A(A'MA)^{-}A'MA = A, \tag{2.4b}$$

$$S(A) \oplus \operatorname{Ker} A'M = E^n, \tag{2.4c}$$

where E^n is the n-dimensional Euclidean space.

Let $P_{A(M)} = A(A'MA)^{-}A'M$. Then if rank $A'M$ = rank A, $P_{A(M)}$ is the projector onto $S(A)$ along Ker $A'M$.

We give a lemma which generalizes (2.3a and 2.3b).

LEMMA 2.3. *Let K be a positive definite matrix of order n. Further, let*

$$P_{(X:Y)(K)} = (X:Y)[(X:Y)'K(X:Y)]^{-}(X:Y)'K.$$

Then

$$P_{(X:Y)(K)} = P_{X(K)} + P_{Y(K)} \quad \textit{if and only if} \quad X'KY = 0, \tag{2.5a}$$

$$P_{(X:Y)(K)} = P_{X(K)} + P_{Y(K)} - P_{X(K)}P_{Y(K)} \quad \textit{if and only if}$$

$$P_{X(K)}P_{Y(K)} = P_{Y(K)}P_{X(K)}. \tag{2.5b}$$

Proof. Equations (2.5a) and (2.5b) follow immediately from (2.3a) and (2.3b) on noting that

$$K^{1/2}P_{X(K)}K^{-1/2} = P_{K^{1/2}X} \tag{2.6}$$

and $P_{K^{1/2}X}$ is an orthogonal projector. Similarly for $P_{Y(K)}$ and $P_{(X:Y)(K)}$. ∎

We now give one more lemma, which generalizes Lemma 2.2.

LEMMA 2.4. *Let A and B be $p \times r_x$ and $p \times r_y$ matrices such that $A'B = 0$ and* rank A + rank $B = p$. *Further, let M and N be n.n.d. matrices of order p satisfying the following conditions*:

(a) rank A = rank $A'M$,
(b) rank B = rank $B'N$, *and*
(c) $A'MNB = 0$.

Then:

(i) *One has*

$$I_p = P_{A(M)} + (P_{B(N)})'. \tag{2.7}$$

where $P_{A(M)} = A(A'MA)^- A'M$, $P_{B(N)} = B(B'NB)^- B'N$ are projectors onto $S(A)$ along Ker $A'M$ *and onto $S(B)$ along* Ker $B'N$, *respectively*.

(ii) *If $S(M) \supset S(B)$, and we choose $N = M^-$, then*

$$M = MA(A'MA)^- A'M + B(B'M^- B)^- B'. \tag{2.8}$$

Part (i) was given by Takane, Yanai, and Mayekawa [11, pp. 681–682], and part (ii) was given by Khatri [6].

NOTE 1. Suppose that N is nonsingular and put $M = N^{-1}$ in (2.7). Then

$$N^{-1}\left[I_p - A(A'N^{-1}A)^- A'N^{-1}\right] = B(B'NB)^- B', \tag{2.9}$$

which is sometimes called Khatri's lemma [4]. This also follows from (2.3a) on setting $X = N^{1/2}B$, $Y = N^{-1/2}A$.

Using Lemma 2.4, the following theorem is established.

THEOREM 2.1. *Let A and B be matrices as defined in Lemma 2.4, and let X be an* $n \times p$ *matrix. Then*

$$P_X = P_{X_* A} + P_{XB}, \tag{2.10}$$

where P_X, X_*, $P_{X_* A}$, *and* P_{XB} *depend on whether* (i) $X'X$ *is nonsingular or not, and/or* (ii) $A = X'W$ *or not.*

Case 1. $X'X$ *is nonsingular*:

$$P_x = X(X'X)^{-1}X', \; X_* = X(X'X)^{-1},$$

$$P_{X_* A} = X(X'X)^{-1}A\left[A'(X'X)^{-1}A\right]^{-}A'(X'X)^{-1}X', \tag{2.11}$$

$$P_{XB} = XB(B'X'XB)^{-}B'X'.$$

Case 2. $X'X$ *is singular and* $A = X'W$ *for some* W:

$$P_X = X(X'X)^{-}X', \qquad X_* = X(X'X)^{-},$$

$$P_{X_* A} = X(X'X)^{-}A\left[A'(X'X)^{-}A\right]^{-}A'(X'X)^{-}X', \tag{2.12}$$

$$P_{XB} = XB(B'X'XB)^{-}B'X'.$$

Case 3. $X'X$ *is singular, and* $A \neq X'W$ *for any* W: *Let* $N = X'X + AA'$. *Put* $M = N_r^-$, *where* N_r^- *is the symmetric reflexive g-inverse of* N. *Further, put*

$$P_X = X'MX, \qquad X_* = XM, \qquad (2.13)$$

$$P_{X_* A} = X_* A(A'MA)^{-}A'(X_*)', \qquad P_{XB} = XB(B'X'XB)^{-}B'X'.$$

The results (2.10), (2.11) and (2.12) are new in that they are written in the form of projectors, although they can easily be obtained from Lemma 2.4.

It is to be noted, here, that $P_x = XMX'$, $P_{X_* A}$, and P_{XB} as given in (2.13) are not projectors themselves except for the case in which $S(X')$ and $S(A)$ are disjoint. In that case, $XMX' = X(X'X)^{-}X'$ and $X_* A = XMA = X(X'X$

$+ AA')^{-}A = 0$, thus establishing

$$X(X'X)^{-}X = XB(B'X'XB)^{-}B'X = P_{XB}.$$

Then A is called an identification restriction [8, p.74].

NOTE 2. Let $P_{X(K)} = X(X'KX)^{-1}X'K$, where K is a positive definite matrix. Further, let A and B be matrices as defined in Theorem 2.1. Then

$$P_{X(K)} = P_{X_{**}(K)} + P_{XB(K)} \tag{2.14}$$

where

$$X_{**} = X(X'KX)^{-1},$$

$$P_{X_{**}A(K)} = X(X'KX)^{-1}A\left(A'(X'KX)^{-1}A\right)^{-}A'(X'KX)^{-1}X'K,$$

$$P_{XB(K)} = XB(B'X'KXB)^{-}B'X'K.$$

A proof of (2.14) follows from (2.5a), (2.6), and the relationship

$$(X_{**}A)'K(XB) = A'B = 0.$$

3. CANONICAL CORRELATION ANALYSIS WITH LINEAR CONSTRAINTS

Let z be an $n \times 1$ random vector with covariance matrix proportional to the identity matrix, and let X and Y be centered $n \times p$ and $n \times q$ matrices, respectively. Then the joint covariance matrix of $X'z$ and $Y'z$ is proportional to the matrix

$$V = \begin{pmatrix} X'X & X'Y \\ Y'X & Y'Y \end{pmatrix}. \tag{3.1}$$

We first consider representing the canonical correlation coefficient between $X'z$ and $Y'z$ without any constraints. We may maximize the correlation coefficient between the composite variables $f'z = (Xa)'z$ and $g'z =$

$(Yb)'z$, i.e.,

$$\rho(f,g) = \frac{\mathrm{Cov}(f'z, g'z)}{\sqrt{\mathrm{Var}(f'z)\,\mathrm{Var}(g'z)}}$$

$$= \frac{a'X'Yb}{\sqrt{(a'X'Xa)(b'Y'Yb)}}. \tag{3.2}$$

It is well known that the maximum value of (3.2), which is called the canonical correlation coefficient between $X'z$ and $Y'z$, can be obtained as the square root of the eigenvalue as shown in the following lemma.

LEMMA 3.1 [13]. *The solutions a and b maximizing* (3.2) *can be obtained by any of the following three statements*:

$$P_X Yb = \mu Xa \quad \text{and} \quad P_Y Xa = \mu Yb, \tag{3.3a}$$

$$(P_X P_Y)Xa = \mu^2 Xa \quad \text{and} \quad P_Y Xa = \mu Yb, \tag{3.3b}$$

$$(P_Y P_X)Yb = \mu^2 Yb \quad \text{and} \quad P_X Yb = \mu Xa. \tag{3.3c}$$

First, observe that the ith largest, canonical correlation coefficient between two random vectors $X'z$ and $Y'z$ is denoted as $\mathrm{cc}_i(X'z, Y'z)$, and the set of all corresponding positive canonical correlations as $\mathrm{cc}(X'z, Y'z)$. Further, the set of all nonzero eigenvalues of a square matrix A is denoted as $\mathrm{nzch}(A)$. Then we have

$$\mathrm{cc}^2(X'z, Y'z) = \mathrm{nzch}(P_X P_Y) = \mathrm{nzch}(P_Y P_X), \tag{3.4}$$

which imply that canonical correlation coefficients between $X'z$ and $Y'z$ are unique for any choice of g-inverses of $X'X$ and $Y'Y$.

Now, let's generalize Lemma 3.1 by assuming that $V(z)$ is proportional to a positive definite matrix K of order n. Then the covariance matrix between $X'z$ and $Y'z$ turns out to be

$$V_{(K)} = \begin{pmatrix} X'KX & X'KY \\ Y'KX & Y'KY \end{pmatrix}. \tag{3.5}$$

LEMMA 3.2. *Under the conditions stated above, solutions a and b maximizing*

$$r(f,g) = \frac{\mathrm{Cov}(f'z, g'z)}{\sqrt{\mathrm{Var}(f'z)\,\mathrm{Var}(g'z)}}$$

$$= \frac{a'X'KYb}{\sqrt{(a'X'KXa)(b'Y'KYb)}} \tag{3.6}$$

can be obtained from any of the following three statements:

$$P_{X(K)}Yb = \mu Xa \quad \text{and} \quad P_{Y(K)}Xa = \mu Yb, \tag{3.7a}$$

$$\left(P_{X(K)}P_{Y(K)}\right)Xa = \mu^2 Xa \quad \text{and} \quad P_{Y(K)}Xa = \mu Yb, \tag{3.7b}$$

$$\left(P_{Y(K)}P_{X(K)}\right)Yb = \mu^2 Yb \quad \text{and} \quad P_{X(K)}Yb = \mu Xa, \tag{3.7c}$$

where $P_{X(K)} = X(X'KX)^{-}X'K$ and $P_{Y(K)} = Y(Y'KY)^{-}Y'K$ are projectors onto $S(X)$ along Ker $X'K$ *and onto $S(Y)$ along* Ker $Y'K$, *respectively.*

Proof of Lemma 3.2. A straightforward proof follows from Lemma 3.1 on replacing X and Y in Lemma 3.1 with $K^{1/2}X$ and $K^{1/2}Y$, and using the relationship (2.5). ∎

Let $\mathrm{cc}(X'z, Y'z)_K$ be the set of all positive canonical correlations between $X'z$ and $Y'z$ which are obtained from the covariance matrix $V_{(K)}$ in (3.5). Then we have

$$\mathrm{cc}^2(X'z, Y'z)_K = \mathrm{nzch}\left(P_{X(K)}P_{Y(K)}\right) = \mathrm{nzch}\left(P_{Y(K)}P_{X(K)}\right).$$

Next, we consider canonical correlation analysis with some constraints on a and/or b.

For given matrices A ($p \times r_x$, $r_x \leqq p$) and C ($q \times r_y$, $r_y \leqq q$), we consider linear constraints of the following forms:

$$A'a = 0 \quad \text{and} \quad C'b = 0, \tag{3.8}$$

which imply $a \in \operatorname{Ker} A'$ and $b \in \operatorname{Ker} C'$. Further, let B and D be $p \times (p - r_x)$ and $q \times (q - r_y)$ matrices such that

$$A'B = 0 \quad \text{and} \quad C'D = 0. \tag{3.9}$$

with

$$p = \operatorname{rank}(A : B) = \operatorname{rank} A + \operatorname{rank} B,$$

$$q = \operatorname{rank}(C : D) = \operatorname{rank} C + \operatorname{rank} D. \tag{3.10}$$

The result is summarized in the following theorem.

THEOREM 3.1. *The solutions a and b maximizing* (3.2) *subject to the linear constraints* (3.8) *are given by either of the following two statements*:

$$P_{XB}Yb = \mu Xa \quad \textit{and} \quad P_{YD}Xa = \mu Yb, \tag{3.11a}$$

$$\left(P_X - P_{X_*A}\right)Yb = \mu Xa \quad \textit{and} \quad \left(P_Y - P_{Y_*C}\right)Xa = \mu Yb. \tag{3.11b}$$

where if $X'X$ is nonsingular, P_X and P_{X_*A} are given by (2.11); *if $X'X$ is singular and $A = X'W$ for some X, then P_X and P_{X_*A} are given by* (2.12); *and if $X'X$ is singular and $A \neq XW$ for any W, then P_X and P_{X_*A} are given by* (2.13).

COROLLARY 3.1. *Consider the following four statements*:

$$(P_{XB}P_{YD})Xa = \mu^2 Xa \quad \textit{and} \quad P_{YD}Xa = \mu Yb, \tag{3.12a}$$

$$\left(P_X - P_{X_*A}\right)\left(P_Y - P_{Y_*C}\right)Xa = \mu^2 Xa \quad \textit{and} \quad \left(P_Y - P_{Y_*C}\right)Xa = \mu Yb, \tag{3.12b}$$

$$(P_{YD}P_{XB})Yb = \mu^2 Yb \quad \textit{and} \quad P_{XB}Yb = \mu Xa, \tag{3.12c}$$

$$\left(P_Y - P_{Y_*C}\right)\left(P_X - P_{X_*A}\right)Yb = \mu^2 Yb \quad \textit{and} \quad \left(P_X - P_{X_*A}\right)Yb = \mu Xa. \tag{3.12d}$$

The six statements given in (3.11) *and* (3.12) *are all equivalent.*

Proof of Theorem 3.1. We differentiate

$$a'X'Yb - \mu_1(a'X'Xa - 1) - \mu_2(b'Y'Yb - 1) - a'A\lambda_1 - b'C\lambda_2 \quad (3.13)$$

(where λ_1 and λ_2 are vectors of Lagrangian multipliers of orders p and q, respectively) with respect to a and b, and set the result equal to zero. We obtain

$$X'Yb - \mu X'Xa = A\lambda_1, \quad (3.14)$$

$$Y'Xa - \omega Y'Yb = C\lambda_2. \quad (3.15)$$

By multiplying the first and second equation by a' and b', respectively, we get $\mu = \mu_1 = \mu_2$. Premultiply (3.14) by $XB(B'X'XB)^{-}B'$. We obtain

$$XB(B'X'XB)^{-}B'X'Yb = \mu XB(B'X'XB)^{-}B'X'Xa = \mu Xa, \quad (3.16)$$

using (3.9). Observe that $A'a = 0$ and $A'B = 0$ imply a $\in S(B)$. This shows that the right side of the above equation (3.16) is equal to μXa. Thus (3.15) reduces to $P_{XB}Yb = \mu Xa$. Similarly, $P_{YD}Xa = \mu Yb$ follows immediately from (3.15) on noting that $C'b = 0$ and $C'D = 0$ implies $b \in S(D)$. The proof of (3.11b) follows immediately, using Theorem 2.1. ∎

We call this analyses *canonical correlation analysis with linear constraints* (CANOLC).

NOTE 3. The equations (3.12a) can be written in terms of the matrices A and C as

$$(P_{XQ_{A'}}P_{YQ_{C'}})Xa = \mu Xa \quad \text{and} \quad P_{YQ_{C'}}Xa = \mu Yb, \quad (3.17)$$

where $Q_{A'}$ and $Q_{C'}$ are defined similarly to (2.2). Equation (3.17) implies that canonical correlation analysis between X and Y subject to constraints of the form (3.8) or (3.9) is equivalent to canonical correlation analysis between $XQ_{A'}$ and $YQ_{C'}$. With regard to $XQ_{A'}$, observe that minimization of $\mathrm{tr}(X - WA')'(X - WA')$ yields $X - \hat{W}A' = XQ_{A'}$, where $\hat{W}$ is a least squares estimate of W. In the context of Takane and Shibayama [10], $XQ_{A'}$ can be interpreted as the residual data matrix eliminating the effects of A', i.e., external information on column variables from X.

Let us denote by $\mathrm{cc}_i(X'z, Y'z)$ the ith largest canonical correlation between two sets of random variables $X'z$ and $Y'z$ obtained from (3.5). Then the results stated above are summarized in the following corollary.

COROLLARY 3.2. *Let A, B, C, and D be matrices as defined in Theorem 3.1. Then the following four sets are identical*:

(i) $\mathrm{cc}^2((XB)'z, (YD)'z)$,
(ii) $\mathrm{cc}^2((XQ_{A'})'z, (YQ_{C'})'z)$,
(iii) $\mathrm{nzch}(P_{XB}P_{YD})$,
(iv) $\mathrm{nzch}((P_X - P_{X_*A})(P_Y - P_{Y_*C}))$.

With regard to the magnitude of the canonical correlation coefficients obtained above, the following properties hold.

COROLLARY 3.3.

(i) $\mathrm{cc}_i(X, Y) \geqq \mathrm{cc}_i(XB, YD)$ *for* $i = 1, \dots, r$, *where* $r = \mathrm{rank}(XB, YD)$, *with equality if* $\mathrm{rank}(XB) = \mathrm{rank}\, X$ *and* $\mathrm{rank}\, YD = \mathrm{rank}\, Y$ *hold simultaneously*.
(ii) *If* $P_{XB}P_{YD} = P_{YD}P_{XB}$, *then* $\mathrm{cc}_i(XB, YD) = 1$ *or* 0.

Proof. Property (i) is established by noting Lemma 4 of [1], which leads to

$$\mathrm{cc}_i(X, Y) \geqq \mathrm{cc}_i(XB, Y) \geqq \mathrm{cc}_i(XB, YB).$$

Property (ii) follows directly from (2.4). ■

NOTE 4. $\mathrm{rank}\, XB = \mathrm{rank}\, X$ is equivalent to $\mathrm{rank}(A : X') = \mathrm{rank}\, A + \mathrm{rank}\, X$, which implies that $S(A)$ and $S(X')$ are disjoint.

Finally, we generalize Theorem 3.1, although only in the case when both $X'KX$ and $Y'KY$ are nonsingular.

NOTE 5. The solutions a and b maximizing (3.2) subject to the linear constraints (3.8) or (3.9) are given by either of the following two statements, provided that both $X'KX$ and $Y'KY$ are nonsingular, where K is a p.d. matrix:

$$P_{XB(K)}Yb = \mu Xa \quad \text{and} \quad P_{YD(K)}Xa = \mu Yb. \tag{3.18a}$$

$$\left(P_{X(K)} - P_{X_{**}A(K)}\right)Yb = \mu Xa \quad \text{and} \quad \left(P_{Y(K)} - P_{Y_{**}C(K)}\right)Xa = \mu Yb, \tag{3.18b}$$

where $X_{**} = X(X'KX)^{-1}$ and $Y_{**} = Y(Y'KY)^{-1}$.

Further, it follows that the following four sets are identical:

(i) $\mathrm{cc}^2(((XB)'z, (YD)'z)_K$,
(ii) $\mathrm{cc}^2((XQ_{A'})'z, (YQ_{C'})'z)_K$.
(iii) $\mathrm{nzch}(P_{XB(K)}P_{YD(K)})$,
(iv) $\mathrm{nzch}((P_{X(K)} - P_{X_{**}A(K)})(P_{Y(K)} - P_{Y_{**}C(K)}))$.

4. RELATION TO SOME OTHER METHODS

In the previous section, we presented a general solution for canonical correlation analysis with some linear constraints. In this section, we consider some relationships that hold between our theorem (Theorem 3.1) and the previously established results.

We first derive a corollary from Theorem 3.1.

COROLLARY 4.1. *When $X'X$ and $Y'Y$ are nonsingular*, (3.11a) *and* (3.11b) *can be written as*

$$(I_p - P_{A(M)})X'Yb = \mu X'Xa, \tag{4.1a}$$

$$(I_q - P_{C(N)})'(Y'Y)^{-1}Y'Xa = \mu b, \tag{4.1b}$$

respectively, where

$$P_{A(M)} = A\left[A'(X'X)^{-1}A\right]^{-}A'(X'X)^{-1}, \tag{4.2a}$$

$$P_{C(N)} = C\left[C'(Y'Y)^{-1}C\right]^{-}C'(Y'Y)^{-1} \tag{4.2b}$$

are oblique projectors.

The proof follows immediately from Theorem 3.1, by observing that

$$(X'X)^{-1}P_{A(M)} = P'_{A(M)}(X'X)^{-1},$$

$$(Y'Y)^{-1}P_{C(N)} = P'_{C(N)}(Y'Y)^{-1}.$$

Thus, it follows that the generalized singular value decomposition [3] of the matrix

$$H = (X'X)^{-1}(I - P_{A(M)})X'Y(I - P_{C(N)})'(Y'Y)^{-1} \tag{4.3}$$

with the row metric $X'X$ and the column metric $Y'Y$ is identical to solving (3.11a) and (3.11b).

4.1. Correspondence Analysis with Linear Constraints (CALC)
Let X and Y be $n \times p$ and $n \times q$ dummy coded matrices. Then

$$K = X'X, \qquad L = Y'Y, \quad \text{and} \quad F = X'Y$$

are diagonal matrices of orders p and q. Further, it can be seen that $F = X'Y$ is a contingency table. Thus, (4.3) can be rewritten as

$$H = K^{-1}Q_{A(K-1)}FQ'_{C(L-1)}L^{-1}$$

where $Q_{A(K-1)} = I_p - A(A'K^{-1}A)^{-}A'K^{-1}$ and $Q_{C(L-1)} = I_q - C(C'L^{-1}C)^{-}C'L^{-1}$.

This was derived by Bockenholt and Bockenholt [2] in the context of correspondence analysis with linear constraints (CALC) on both rows and columns of a contingency table. Further, note that canonical correspondence analysis [9] can be made equivalent to CALC by judicious choice of the constraint matrices [11].

4.2. Multiple Regression Analysis with Linear Constraints
Choose $Y = y$ in (3.11b). We obtain

$$(P_x - P_{X_*A})y = \mu Xa. \tag{4.4}$$

If we put $y = X\beta + \varepsilon$ and choose $a = \beta$, then the least squares estimate of β with the constraint $A'\beta = 0$ is given by

$$X\hat{\beta} = X\left\{(X'X)^{-1}X' - (X'X)^{-1}A\left[A'(X'X)^{-1}A\right]^{-}A'(X'X)^{-1}X'\right\}y, \tag{4.5}$$

which implies

$$\hat{\beta} = \left\{(X'X)^{-1}X' - (X'X)^{-1}A\left[A'(X'X)^{-1}A\right]^{-}A'(X'X)^{-1}X'\right\}y \tag{4.6}$$

if $X'X$ is nonsingular. Equation (4.6) is the solution obtained by means of the Lagrange multiplier method. It is to be noted here that (4.6) generalizes an earlier result (for example, see [8, (3.59), p. 85]) in the sense that the term $(A'(X'X)^{-1}A)^{-1}$ is replaced by $(A'(X'X)^{-1}A)^{-}$. Further, from (3.11a), we have $P_{XB}y = \mu Xa$, which leads to

$$\hat{\beta} = B(B'X'XB)^{-}B'X'y. \tag{4.7}$$

It can be shown that (4.6) is identical to (4.7).

REFERENCES

1 J. K. Baksalary, S. Puntanen, and H. Yanai, Canonical correlations associated with symmetric reflexive generalized inverses of the dispersion matrix, *Linear Algebra Appl.* 176:61–74 (1992).

2 U. Bockenholt and I. Bockenholt, Canonical analysis of contingency tables with linear constraints, *Psychometrika* 55:633–639 (1990).

3 M. Greenacre, *Theory and Application of Correspondence Analysis*, Academic, 1984.

4 C. G. Khatri, A note on a MANOVA model applied to problems in growth curves. *Ann. Inst. Statist. Math.* 18:75–86 (1966).

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

6 C. G. Khatri, Some properties on BLUE in a linear model and canonical correlations associated with linear transformation, *J. Multivariate Anal.* 34:211–226 (1990).

7 C. G. Rao, and H. Yanai, General definition of a projector, its decomposition and application to statistical problems, *J. Statist. Plann. Inference* 3(1):1–17 (1979).
8 G. A. F. Seber, *Linear Regression Analysis*, Wiley-Interscience, (1977).
9 C. F. G. Ter Braak, Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis, *Ecology* 67:1167–1179 (1986).
10 Y. Takane and T. Shibayama, Principal component analysis with external information on both subjects and variables, *Psychometrika* 56:97–120 (1991).
11 Y. Takane, H. Yanai, and S. Mayekawa. Relationships among several methods of linearly constrained correspondence analysis, *Psychometrika* 56:667–684 (1991).
12 H. Yanai, Explicit expressions of projectors on canonical variables and distances between centroids of groups, *J. Japan Statist. Soc.* 11:43–53 (1981).
13 H. Yanai, Some generalized forms of least squares g-inverse, minimum norm g-inverse and Moore-Penrose inverse matrix, *Comput. Statist. Data Anal.* 10:251–260 (1990).

Received 16 October 1990; final manuscript accepted 18 May 1992