Relationships among Various Kinds of Eigenvalue and Singular Value Decompositions

Yoshio Takane

McGill University, 1205 Dr. Penfield Avenue Montreal QC H3A 1B1
takane@takane2.psych.mcgill.ca

Summary. Eigenvalue decomposition (EVD) and/or singular value decomposition (SVD) play important roles in many multivariate data analysis techniques as computational tools for dimension reduction. A variety of EVD and SVD have been developed to deal with specific kinds of dimension reduction problems. This paper explicates various relationships among those decompositions with the prospect of exploiting them in practical applications of multivariate analysis.

Key words: Ordinary and Generalized Eigenvalue Decompositions (EVD and GEVD), Ordinary, Product, Quotient, Restricted, and Generalized Singular Value Decompositions (SVD, PSVD, QSVD, RSVD, and GSVD), Canonical Correlation Analysis (CANO), Constrained Principal Component Analysis (CPCA).

1 Introduction

In many multivariate data analysis (MVA) techniques, we look for best reduced-rank approximations of (some functions of) data matrices. In principal component analysis (PCA), for example, we obtain a reduced-rank approximation of a (standardized) data matrix or that of a covariance (or correlation) matrix. In canonical correlation analysis (CANO), we obtain a reduced-rank approximation of $P_G P_H$, where $P_G = G(G'G)^- G'$ and $P_H = H(H'H)^- H'$ (with $^-$ indicating a generalized inverse (g-inverse) of a matrix) are orthogonal projectors formed from the two sets of data matrices, G and H. Eigenvalue (or spectral) decomposition (EVD) and/or singular value decomposition (SVD) play important roles in this rank-reduction process.

Standard PCA has recently been extended to accommodate external information on both rows and columns of data matrices (CPCA; Takane and Shibayama 1991, Takane and Hunter 2001). Similar extensions have also been made to CANO (GCCANO; Takane and Hwang 2002, Takane et al. 2002, Yanai and Takane 1992). These techniques first decompose data matrices or projectors formed from the data matrices according to the external information. Decomposed matrices are then subjected to a rank-reduction process. Again, EVD or SVD plays important roles in the rank-reduction process.

Ordinary EVD and SVD have been extended in various ways: Generalized EVD (GEVD), Product SVD (PSVD), Quotient SVD (QSVD), Restricted SVD (RSVD), Generalized SVD (GSVD), etc. GEVD obtains EVD of a square (often symmetric and/or nonnegative definite (*nnd*)) matrix with respect to another *nnd* matrix. PSVD obtains SVD of a product of matrices without explicitly forming the product. QSVD obtains SVD of a rectangular matrix with respect to another rectangular matrix. RSVD extends QSVD to a triplet of matrices. GSVD obtains SVD of a matrix under non-identity *nnd* metric matrices. There are interesting relationships among these decompositions. For example, it is well known that SVD($P_G P_H$) mentioned above is simply related to GSVD($(G'G)^- G'H(H'H)^-, G'G, H'H$) (i.e., GSVD of $(G'G)^- G'H(H'H)^-$ with row metric $G'G$ and column metric $H'H$) (Takane and Shibayama 1991, Takane and Hunter 2001; see also the application section below). In this paper we systematically investigate these relationships and suggest ways to exploit them in practical applications of multivariate analysis.

In what follows, U and V denote (full) orthogonal matrices of order m and n, respectively, and X, X_1, and X_2 denote square nonsingular matrices of appropriate orders. Diag$(\cdots)$ indicates a block diagonal matrix with matrices in parentheses constituting its diagonal blocks. Sp(A) and Ker(A) indicate, respectively, the range and the null spaces of A. Let C^- denote a g-inverse of C. The following conditions called Moore-Penrose Conditions characterize various kinds of g-inverse matrices.

$$CC^-C = C, \tag{1}$$

$$C^-CC^- = C^- \qquad \text{(reflexivity)}, \tag{2}$$

$$(CC^-)' = CC^- \qquad \text{(least squares)}, \tag{3}$$

and

$$(C^-C)' = C^-C \qquad \text{(minimum norm)}. \tag{4}$$

Conditions (3) and (4) may, respectively, be generalized into:

$$(SCC^-)' = SCC^- \qquad \text{(S-least squares)}, \tag{5}$$

and

$$(TC^-C)' = TC^-C, \qquad \text{(T-minimum norm)}, \tag{6}$$

where S and T are *nnd* matrices (called metric matrices). Matrices C_l^- and C_m^- denote a least squares and a minimum norm g-inverse of C, respectively, satisfying (1) and (3), and (1) and (4), respectively.

2 Various Decompositions

2.1 SVD

Let A be an m by n matrix of rank a, and let D be an m by n semi-diagonal matrix of the form $D = \begin{bmatrix} {}_a\Delta_a & {}_a0 \\ 0_a & 0 \end{bmatrix}$, where ${}_a\Delta_a$ is a positive definite (pd) diagonal matrix of order a. (We sometimes write ${}_a\Delta_a$ as Δ.) Then, matrix decomposition

$$A = UDV' \tag{7}$$

is called (complete) SVD of A and is denoted by SVD(A). Let U and V be partitioned according to the partition of D above. That is, $U = [U_1, U_2]$, and $V = [V_1, V_2]$, where U_1 is m by a, and V_1 is n by a. Then,

$$A = U_1 \Delta V_1'. \tag{8}$$

This is called incomplete SVD of A. In most data analysis applications only the incomplete version of SVD is of direct interest. In this paper, however, the word SVD refers to the complete SVD unless otherwise specified. Note that $\mathrm{Sp}(A) = \mathrm{Sp}(U_1)$, $\mathrm{Ker}(A') = \mathrm{Sp}(U_2)$, $\mathrm{Sp}(A') = \mathrm{Sp}(V_1)$, and $\mathrm{Ker}(A) = \mathrm{Sp}(V_2)$. (It should be understood that some blocks in partitioned matrices may be null. For example, in the above $D = [\Delta, 0]$ and $U = U_1$ if $m = a$, and $D' = [\Delta, 0]$ and $V = V_1$, if $n = a$.)

The diagonal elements in Δ are assumed to be all distinct and in decreasing order of magnitude, and columns of U and V are arranged accordingly. Let $U_{(r)}$ and $V_{(r)}$ denote the matrices with the first $r(\leq a)$ columns of U_1 and V_1, and let $\Delta_{(r)}$ denote the matrix with the leading diagonal block of order r in Δ. Then, $A_{(r)} = U_{(r)}\Delta_{(r)}V_{(r)}'$ gives the best rank r approximation to A in the least squares sense. That is, the above $A_{(r)}$ gives the minimum of $\mathrm{SS}(A - A_{(r)}) = \mathrm{tr}(A - A_{(r)})'(A - A_{(r)})$ over all matrices of rank r (Eckart and Young 1936; see ten Berge (1993) for an elegant proof of this property). It is this best reduced-rank approximation property that makes SVD extremely useful in multivariate analysis. Mirsky (1960) later showed that this optimality was not restricted to the SS norm. It holds generally for all unitarily (orthogonally) invariant norms.

The following property of SVD is important from a computational point of view (Takane and Hunter 2001, Theorem 1):

Property 1. Let B and C be columnwise orthogonal matrices, i.e., $B'B = I$ and $C'C = I$. Let $A = UDV'$ denote SVD(A), and let $BAC' = U^*D^*V^{*\prime}$ denote SVD(BAC'). Then, $U^* = BU$ (or $U = B'U^*$), $V^* = CV$ (or $V = C'V^*$), and $D^* = D$.

2.2 EVD

Let S be an *nnd* matrix of order n and of rank a. Let $\tilde{D}^2$ denote an *nnd* diagonal matrix of order n with Δ^2 as the leading diagonal block. Then,

$$S = V\tilde{D}^2V' \tag{9}$$

is called (complete) EVD of S and is denoted as EVD(S). Matrix S can also be expressed as

$$S = V_1\Delta^2V_1', \tag{10}$$

which is called incomplete EVD of S. When S is obtained by $A'A$, EVD(S) can be derived from SVD(A) by $S = A'A = V\tilde{D}^2V' = V_1\Delta^2V_1'$, where $\tilde{D}^2 = D'D$. Alternatively, if S is obtained by AA', then $S = AA' = UDD'U' = U_1\Delta^2U_1'$. As in the SVD of A, $S_{(r)} = V_{(r)}\tilde{\Delta}^2_{(r)}V'_{(r)}$ gives the best rank r approximation of S in the least squares sense.

2.3 PSVD

Let A be as defined earlier, and let C be a p by n matrix. The pair of decompositions,

$$A = UDX' \quad \text{and} \quad C' = (X')^{-1}JV', \tag{11}$$

is called PSVD of the matrix pair, A and C, and is denoted by PSVD(A, C) (Fernando and Hammarling 1988). Here,

$$D = \begin{bmatrix} {}_h\Delta_h & {}_h0_s & {}_h0_t & {}_h0 \\ {}_t0_h & {}_t0_s & {}_tI_t & {}_t0 \\ 0_h & 0_s & 0_t & 0 \end{bmatrix}, \quad \text{and} \quad J = \begin{bmatrix} {}_hI_h & {}_h0_s & {}_h0 \\ {}_s0_h & {}_sI_s & {}_s0 \\ {}_t0_h & {}_t0_s & {}_t0 \\ 0_h & 0_s & 0 \end{bmatrix},$$

where, $h = \text{rank}(AC')$, $s = \text{rank}(C) - h$, $t = \text{rank}(A) - h$, and ${}_h\Delta_h$ is diagonal and *pd*. It follows that

$$AC' = U(DJ)V' \tag{12}$$

gives SVD of AC'. Note that portions of X corresponding to the last column block of D (and the last row block of J) are not unique (although this usually does not cause much difficulties).

PSVD of two matrices can be easily extended to the product of three matrices, A, $B(q \times m)$, and C. Let $A = X_1DX_2'$, $B = UEX_1^{-1}$, and $C' = (X_2')^{-1}JV'$. Then, $BAC' = U(EDJ)V'$ gives the SVD of BAC'. PSVD of the matrix triplet, A, B and C, is denoted by PSVD(A, B, C). Here, D, E, and J are of the following form

$$E = \begin{bmatrix} {}_hI_h & {}_h0_s & {}_h0_k & {}_h0_j & {}_h0_i & {}_h0 \\ {}_j0_h & {}_j0_s & {}_j0_k & {}_jI_j & {}_j0_i & {}_j0 \\ {}_i0_h & {}_i0_s & {}_i0_k & {}_i0_j & {}_iI_i & {}_i0 \\ 0_h & 0_s & 0_k & 0_j & 0_i & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} {}_h\Delta_h & {}_h0_s & {}_h0_t & {}_h0_k & {}_h0_j & {}_h0 \\ {}_s0_h & {}_sI_s & {}_s0_t & {}_s0_k & {}_s0_j & {}_s0 \\ {}_k0_h & {}_k0_s & {}_k0_t & {}_kI_k & {}_k0_j & {}_k0 \\ {}_j0_h & {}_j0_s & {}_j0_t & {}_j0_k & {}_jI_j & {}_j0 \\ {}_i0_h & {}_i0_s & {}_i0_t & {}_i0_k & {}_i0_j & {}_i0 \\ 0_h & 0_s & 0_t & 0_k & 0_j & 0 \end{bmatrix}, \quad \text{and} \quad J = \begin{bmatrix} {}_hI_h & {}_h0_s & {}_h0_t & {}_h0 \\ {}_s0_h & {}_sI_s & {}_h0_t & {}_s0 \\ {}_t0_h & {}_t0_s & {}_tI_t & {}_t0 \\ {}_k0_h & {}_k0_s & {}_k0_t & {}_k0 \\ {}_j0_h & {}_j0_s & {}_j0_t & {}_j0 \\ 0_h & 0_s & 0_t & 0 \end{bmatrix},$$

where $h = \text{rank}(BAC')$, $s = \text{rank}(AC') - h$, $t = \text{rank}(C) - \text{rank}(AC')$, $k = \text{rank}(A) - \text{rank}(BA) - \text{rank}(AC') + h$, $j = \text{rank}(BA) - h$, and $i = \text{rank}(B) - \text{rank}(BA)$.

In multivariate analysis, there are many situations in which SVD of a product of two or more matrices is obtained. One classical example is inter-battery factor analysis, where a best reduced-rank approximation of the product of two sets of variables, G and H, is obtained. As another example, Koene and Takane (1999) recently proposed a regularization technique for multi-layered feed-forward neural networks. This method obtains a best reduced-rank approximation of YW, where Y is the matrix of input variables, and W is the matrix of estimated weights associated with the connections from input variables to hidden units. Takane and Yanai (2002) used PSVD of a matrix triplet (Zha 1991b) to show what kind of g-inverse of BAC' is necessary for $\text{rank}(A - AC'(BAC')^- BA) = \text{rank}(A) - \text{rank}(BAC')$, to hold. The result generalizes the Wedderburn-Guttman theorem (Guttman 1944, 1957; Wedderburn 1934).

2.4 QSVD

Let A and C be as defined in PSVD. The pair of decompositions,

$$A = UDX' \quad \text{and} \quad C = VJX', \tag{13}$$

is called QSVD of A with respect to C and is denoted by QSVD(A, C) (Van Loan, 1976). Matrices D and J are of the form

$$D = \begin{bmatrix} {}_h\Delta_h & {}_h0_s & {}_h0_t & {}_h0 \\ {}_t0_h & {}_t0_s & {}_tI_t & {}_s0 \\ 0_h & 0_s & 0_t & 0 \end{bmatrix}, \quad \text{and} \quad J = \begin{bmatrix} {}_hI_h & {}_h0_s & {}_h0_t & {}_h0 \\ {}_s0_h & {}_sI_s & {}_s0_t & {}_s0 \\ 0_h & 0_s & 0_t & 0 \end{bmatrix},$$

where Δ is *pd* and diagonal, and $h = \dim(\text{Sp}(A') \cap \text{Sp}(C'))$, $s = \dim(\text{Ker}(A) \cap \text{Sp}(C'))$, and $t = \dim(\text{Sp}(A') \cap \text{Ker}(C))$. As in PSVD, the portions of X pertaining to the last block are not unique.

Let N be such that $\text{Sp}(N) = \text{Ker}(C)$. (This matrix N could be V_2 if $C'C = V\tilde{D}^2V' = V_1D_1^2V_1'$ represents the EVD of $C'C$.) Define further $C^-_{N/A'A} = Q_{N/A'A}C_l^-$, where $Q_{N/A'A} = I - N(N'A'AN)^- N'A'A$ (and C_l^- is a least squares g-inverse of C). Then, $\mathrm{PSVD}(A, C^-_{N/A'A}) = \mathrm{SVD}(AC^-_{N/A'A})$ is obtained from $\mathrm{QSVD}(A, C)$ by

$$\begin{aligned} AC^-_{N/A'A} &= AQ^-_{N/A'A}C^-_l \\ &= UDX'Q_{N/A'A}(X')^{-1}J^-_l V' \\ &= UDQ_{N^*/D'D}J^-_l V' \\ &= U(DJ^-_{N^*/D'D})V', \end{aligned} \tag{14}$$

where $N^* = X'N$, and $J^-_{N^*/D'D} = Q_{N^*/D'D}J^-_l$. Note that in going from the first equation to the second we used $(X')^{-1}J^-_l V' \in \{C^-_l\}$ if $C = VDX'$ (Rao and Mitra 1971, Complement 3.2 (iv)). Leading diagonals in $DJ^-_{N^*/D'D} = \mathrm{diag}(\Delta, 0, 0, 0)$ have nonzero and finite singular values of A with respect to C. (All improper (infinite and indeterminate) singular values become zero in $\mathrm{SVD}(AC^-_{N/A'A})$.) Note that $A'A = XD'DX'$, and that $X'Q_{N/A'A}(X')^{-1} = I - N^*(N^{*\prime}D'DN^*)^-N^{*\prime}D'D = Q_{N^*/D'D}$. Note also that

$$J^-_l = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ Z_1 & Z_2 & Z_3 \\ E_1 & E_2 & E_3 \end{bmatrix}, \quad \text{but} \quad J^-_{N^*/D'D} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \\ E^*_1 & E^*_2 & E^*_3 \end{bmatrix},$$

where Z_i's and E_i's, and E^*_i's $(i = 1, \cdots, 3)$ are arbitrary.

G-inverse, $C^-_{N/A'A} = Q_{N/A'A}C^-_l$, satisfies (1), (3), and (6) of the extended Moore-Penrose Conditions, namely

$$CC^-_{N/A'A}C = C, \tag{15}$$

$$(CC^-_{N/A'A})' = CC^-_{N/A'A} \qquad \text{(least squares)}, \tag{16}$$

and

$$(A'AC^-_{N/A'A}C)' = A'AC^-_{N/A'A}C \qquad (A'A\text{-minimum norm}). \tag{17}$$

The first two of these conditions follow immediately from $CQ_{N/A'A} = C$. The last condition follows from $A'AC^-_{N/A'A}C = A'AQ_{N/A'A}$, (the latter is clearly symmetric), and from

$$A'AQ_{N/A'A} = C'\Lambda C \tag{18}$$

for some Λ. That is, $A'AC^-_{N/A'A}C = A'AQ_{N/A'A} = A'AQ_{N/A'A}C^-_l C = C'\Lambda CC^-_l C = C'\Lambda C = A'AQ_{N/A'A}$. G-inverse, $J^-_{N^*/D'D} = Q_{N^*/D'D}J^-_l$, has a similar property. The fact that $C^-_{N/A'A}$ can be expressed as $Q_{N/A'A}C^-_l$ (and that $J^-_{N^*/D'D}$ can be expressed as $Q_{N^*/D'D}J^-_l$) is sufficient, but not necessary, for the above three conditions ((1), (3), and (6)) to hold. In terms of the expression of $J^-_{N^*/D'D}$ given above, the zero matrix in the third row and the third column block could be arbitrary in order to satisfy the three conditions.

The name "Quotient" SVD derives from the fact that it obtains SVD of $AC^-_{N/A'A}$, which is analogous to taking the quotient (a/c) of two numbers (De Moor 1991, De Moor and Golub 1991). QSVD used to be called GSVD by computational matrix algebraists (Golub and Van Loan 1989, Paige 1986, Paige and Saunders 1981, Van Loan 1976). This was quite confusing because the same terminology has been used by French data analysts for a different decomposition. See below. GSVD in the sense of computational matric algebraists has recently been renamed into QSVD (De Moor and Golub 1991).

2.5 RSVD

Let A, B, and C be as introduced earlier. The triplet of decompositions,

$$A = X_1 D X_2', \quad B = UEX_1', \quad \text{and} \quad C = VJX_2', \tag{19}$$

is called RSVD of A with respect to B and C and is denoted by RSVD(A, B, C) (De Moor and Golub 1991, Zha 1991a). Here,

$$D = \begin{bmatrix} {}_h\Delta_h & {}_h0_s & {}_h0_t & {}_h0_k & {}_h0_j & {}_h0 \\ {}_s0_h & {}_sI_s & {}_h0_t & {}_s0_k & {}_s0_j & {}_s0 \\ {}_k0_h & {}_k0_s & {}_k0_t & {}_kI_k & {}_k0_j & {}_k0 \\ {}_j0_h & {}_j0_s & {}_j0_t & {}_j0_k & {}_jI_j & {}_j0 \\ {}_i0_h & {}_i0_s & {}_i0_t & {}_i0_k & {}_i0_j & {}_i0 \\ 0_h & 0_s & 0_t & 0_k & 0_j & 0 \end{bmatrix}, \quad E = \begin{bmatrix} {}_hI_h & {}_h0_j & {}_h0_i & {}_h0 \\ {}_s0_h & {}_s0_j & {}_s0_i & {}_s0 \\ {}_k0_h & {}_k0_j & {}_k0_i & {}_k0 \\ {}_j0_h & {}_jI_j & {}_j0_i & {}_j0 \\ {}_i0_h & {}_i0_j & {}_iI_i & {}_i0 \\ 0_h & 0_j & 0_i & 0 \end{bmatrix},$$

and

$$J = \begin{bmatrix} {}_hI_h & {}_h0_s & {}_h0_t & {}_h0_k & {}_h0_j & {}_h0 \\ {}_s0_h & {}_sI_s & {}_s0_t & {}_s0_k & {}_s0_j & {}_s0 \\ {}_t0_h & {}_t0_s & {}_tI_t & {}_t0_k & {}_t0_j & {}_t0 \\ 0_h & 0_s & 0_t & 0_k & 0_j & 0 \end{bmatrix},$$

where $h = \rho_{abc} - \rho_{ab} - \rho_{ac} + \text{rank}(A)$, $s = \rho_{ab} - \rho_{abc} + \text{rank}(C)$, $t = \rho_{ac} - \text{rank}(A)$, $k = \rho_{abc} - \text{rank}(A) - \text{rank}(B)$, $j = \rho_{ab} - \rho_{abc} + \text{rank}(B)$, and $i = \rho_{ab} - \text{rank}(A)$ with $\rho_{abc} = \text{rank}\begin{bmatrix} A & B' \\ C & 0 \end{bmatrix}$, $\rho_{ab} = \text{rank}[A, B']$, $\rho_{ac} = \text{rank}\begin{bmatrix} A \\ C \end{bmatrix}$.

SVD$((B^-_{M/AA})'AC^-_{N/A'A})$ follows from RSVD(A, B, C), where M is such that Sp(M) = Ker(B) (which is analogous to N for C). Note that $(B^-_{M/AA'})'A C^-_{N/A'A} = U(K^-_{M^*/DD'})'DJ^-_{N^*/D'D}V'$, where U and V are, respectively, matrices of left and right singular vectors, and $(K^-_{M^*/DD'})'DJ^-_{N^*/D'D}$ a diagonal matrix of nonzero finite singular values of $(B^-_{M/AA'})'AC^-_{N/A'A}$. (All other singular values of A with respect to B and C become zero.) Matrix $B^-_{M/AA}$ has similar properties to those possessed by $C^-_{N/A'A}$.

2.6 GEVD (QEVD)

Let T be an *nnd* matrix of the same order as S. The pair of decompositions,

$$S = X\tilde{D}^2X' \quad \text{and} \quad T = X\tilde{J}^2X', \tag{20}$$

is called GEVD of S with respect to T and is denoted by GEVD(S, T). Here, $\tilde{D}^2 = \text{diag}(\Delta^2, 0, I, 0)$ (where Δ^2 is diagonal and *pd*), and $\tilde{J}^2 = \text{diag}(I, I, 0, 0)$. The four diagonal blocks in these matrices pertain to $\text{Sp}(S) \cap \text{Sp}(T)$, $\text{Ker}(S) \cap \text{Sp}(T)$, $\text{Sp}(S) \cap \text{Ker}(T)$, and $\text{Ker}(S) \cap \text{Sp}(T)$, and their sizes reflect the dimensionality of these subspaces. As in QSVD (and PSVD), portions of X pertaining to the last block are usually non-unique.

Let $W = (X')^{-1}$. In the more traditional approach to GSVD, this W may be of more direct interest. This W has the property of simultaneously diagonalizing both S and T. That is, $W'SW = \tilde{D}^2$, and $W'TW = \tilde{J}^2$. De Leeuw (1982) has given a complete solution to this problem. Let $T = FF'$ be a full rank square root decomposition of T, and let N be a columnwise orthogonal matrix such that $\text{Ker}(T) = \text{Sp}(N)$. Let $W = [W_1, W_2]$, where W_1 and W_2 correspond with $\text{Sp}(T) = \text{Sp}(F)$ and $\text{Ker}(T) = \text{Sp}(N)$, respectively. Define $Q_{N/S} = (I - N(N'SN)^- N'S)$. (This is analogous to $Q_{N/A'A}$ in QSVD.) Then, W can be obtained by

$$W_1 = Q_{N/S}(F')^+ W_1^* + Q_{N/S}NZ_1, \tag{21}$$

and

$$W_2 = NW_2^* D^*, \tag{22}$$

where W_1^* and W_2^* are complete sets of eigenvectors of $F^+Q'_{N/S}SQ_{N/S}(F')^+$ and $N'SN$, respectively, $^+$ indicates the Moore-Penrose inverse ($F^+ = (F'F)^{-1}F'$ and $(F')^+$ is its transpose), and Z_1 is arbitrary. D^* is such that $D^* = \begin{bmatrix} \Delta_2^{-1} & 0 \\ 0 & Z_2 \end{bmatrix}$, where Δ_2^2 is the diagonal matrix of nonzero eigenvalues of $N'SN$, and Z_2 is a nonnull (but otherwise arbitrary) square matrix. Note that $\text{Sp}(Q_{N/S}N)$ pertains to the joint null space of S and T, that is, $\text{Ker}(S) \cap \text{Ker}(T)$. The fact that W is usually non-unique corresponds with the fact that X is usually non-unique.

If $S = A'A$ and $T = C'C$, where A and C are those used in QSVD, GEVD(S, T) is obtained from QSVD(A, C) by setting $S = X\tilde{D}^2X'$ and $T = X'\tilde{J}^2X$, where $\tilde{D}^2 = D'D$, and $\tilde{J}^2 = J'J$. In fact, QSVD was invented initially (Van Loan, 1976) to obtain GEVD($A'A, C'C$) without explicitly calculating $A'A$ and $C'C$.

2.7 GSVD

Let A be as defined earlier, and let K and L be *nnd* matrices of order m and n, respectively. Then, the triplet of matrix decompositions

$$A = X_1 D X_2', \quad K = (X_1')^{-1} \tilde{E}^2 X_1^{-1}, \text{ and } L = (X_2')^{-1} \tilde{J}^2 X_2^{-1}, \tag{23}$$

is called GSVD of A with metric matrices K and L. Matrix D is similar to that in PSVD, and $\tilde{E}^2 = \text{diag}(I, 0, 0, I, I, 0)$ and $\tilde{J}^2 = \text{diag}(I, I, I, 0, 0, 0)$ are themselves *nnd*, and are obtained by $\tilde{E}^2 = E'E$ and $\tilde{J}^2 = J'J$. Let $R_K' = \tilde{E} X_1^{-1}$, and $R_L = (X_2')^{-1} \tilde{J}$, and define $(R_K')_*^- = X_1 \tilde{E}^+$, and $(R_L)_*^- = \tilde{J}^+ X_2'$ (where $^+$ indicates the Moore-Penrose inverse). Then, $(R_K')_*^- R_K' A R_L (R_L)_*^- = X_1 (\tilde{E}^+ \tilde{E} D \tilde{J} \tilde{J}^+ X_2' = X_1 D^* X_2'$. (Matrix $(R_K')_*^-$ satisfies (1), (2), (3), and (6) of the extended Moore-Penrose Conditions with $T = (X_1 X_1')^{-1}$, while $(R_L)_*^-$ satisfies (1), (2), (3), and (5) with $S = X_2 X_2'$. Leading diagonals of D^* have nonzero generalized singular values of A under metric matrices K and L. GSVD of A under metric matrices K and L is denoted by $\text{GSVD}(A, K, L)$.

$\text{GSVD}(A, K, L)$ is typically calculated as follows: Let $K = F_K F_K'$ and $L = F_L F_L'$ denote full rank square root decompositions of K and L. Let $F_K' A F_L = U * D * V^{*\prime}$ be $\text{SVD}(F_K' A F_L)$. Then, $P_{F_K} A P_{F_L} = X_1 \tilde{D} X_2'$, where P_{F_K} and P_{F_L} are orthogonal projectors defined by F_K and F_L, $U = F_K (F_K' F_K)^{-1} U^*$, $\tilde{D} = D^*$, and $V = F_L (F_L' F_L)^{-1} V^*$. Note that $X_1' K X_1 = \tilde{E}^2$, and $X_2' L X_2 = \tilde{J}^2$.

There is an important property associated with GSVD of a matrix, which is analogous to Property 1 mentioned in the section entitled SVD (Takane and Hunter 2001, Theorem 2).

Property 2. Let B and C be such that BAC' can be formed. Let $BAC' = UDV'$ denote $\text{GSVD}(BAC', K, L)$, and let $A = U^* D^* V^{*\prime}$ denote $\text{GSVD}(A, B'KB, C'LC)$. Then, $U = K^- KBU^*$, $V = L^- LCV^*$ and $D = D^*$, or $U^* = (B'KB)^- B'KU$, $V^* = (C'LC)^- C'LV$ and $D^* = D$.

3 Applications

How can we effectively use these decompositions in representative methods of multivariate analysis? Let us take canonical correlation analysis (CANO) and constrained principal component analysis (CPCA; Takane and Shibayama 1991, Takane and Hunter 2001) as examples.

As is well known, CANO between G and H amounts to $\text{SVD}(P_G P_H)$ or equivalently to $\text{GSVD}((G'G)^- G'H(H'H)^-, G'G, H'H)$. The former obtains canonical scores directly, while the latter obtains weights applied to G and H to obtain the canonical scores. Property 2 mentioned above indicates there is a simple relationship between the two decompositions. Let $P_G P_H = UDV'$ represent $\text{SVD}(P_G P_H)$, and let $(G'G)^- G'H(H'H)^- = X_1 D^* X_2'$ represent $\text{GSVD}((G'G)^- G'H(H'H)^-, G'G, H'H)$. Then, $U = GX_1$ (or $X_1 = (G'G)^- G'U$), $V = HX_2$ (or $X_2 = (H'H)^- H'V$), and $D = D^*$. It may look as if the dimensionality of the former problem is usually much larger than the latter. Thus, it is wise to obtain the latter first and then derive the former from it. However, the dimensionality of the former can also be directly reduced using Property 1 in SVD. Let Y_G and Y_H denote orthonormal bases for $\mathrm{Sp}(G)$ and $\mathrm{Sp}(H)$, respectively. Then, P_G and P_H can be written as $P_G = Y_G Y_G'$ and $P_H = Y_H Y_H'$, respectively. Let $Y_G' Y_H = UDV'$ denote $\mathrm{SVD}(Y_G' Y_H)$. Then, $\mathrm{SVD}(P_G P_H) = \mathrm{SVD}(Y_G(Y_G' Y_H) Y_H')$ is obtained by $Y_G UDV' Y_H'$.

In GCCANO (Takane and Hwang 2002, Takane et al. 2002), the situation is only slightly more complicated. In GCCANO, matrices from which projectors are formed (that is, matrices analogous to G and H in standard CANO) are obtained by products of two or more matrices. However, the structure of the computational problem remains essentially the same.

CPCA, on the other hand, involves five matrices. CPCA of a data matrix Z with two external information matrices, G and H, and two metric matrices, K and L, is denoted by $\mathrm{CPCA}(Z, G, H, K, L)$, and it subsumes a number of existing MVA techniques as special cases. For example, CANO between G and H follows when $Z = I$, $K = I$ and $L = I$. CPCA amounts to $\mathrm{GSVD}(P_{G/K} Z P_{H/L}', K, L)$ or equivalently to $\mathrm{GSVD}((G'KG)^- G'KZLH(H'LH)^-, G'KG, H'LH)$. Again by Property 2, there is a simple relationship between the two decompositions. Let $G^* = R_K' G$, $H^* = R_L' H$, and $A^* = R_K' A R_L$. Then, $R_K' P_{G/K} Z P_{H/L}' R_L = P_{G^*} A^* P_{H^*}$, where P_{G^*} and P_{H^*} are orthogonal projectors defined by G^* and H^*, respectively. Let Y_{G^*} and Y_{H^*} be matrices of orthonormal bases spanning $\mathrm{Sp}(G^*)$ and $\mathrm{Sp}(H^*)$, respectively. Then, $P_{G^*} A^* P_{H^*} = Y_{G^*} Y_{G^*}' A^* Y_{H^*} Y_{H^*}'$. Let $Y_{G^*}' A^* Y_{H^*} = UDV'$ be $\mathrm{SVD}(Y_{G^*}' A^* Y_{H^*})$. Then, $\mathrm{SVD}(Y_{G^*}(Y_{G^*}' A^* Y_{H^*}) Y_{H^*}')$ is obtained by $Y_{G^*} UDV' Y_{H^*}'$. From this, $\mathrm{GSVD}(P_{G/K} Z P_{H/L}', K, L)$ is obtained by $(R_K')^- Y_{G^*} UDV' Y_{H^*}' R_L^-$ and $\mathrm{GSVD}((G'KG)^- G'KZLH(H'LH)^-, G'KG, H'LH)$ is obtained by $(G^{*\prime} G^*)^- G^{*\prime} UDV' H^* (H^{*\prime} H^*)^-$.

The above procedures represent more conventional procedures for computing CANO (and GCCANO), and CPCA. In the light of the new breed of SVD's discussed in this paper, we have other options, which may yield numerically more stable solutions. Virtually any one of the EVD's and SVD's discussed in this paper can be used to obtain solutions for CANO and CPCA by preprocessing matrices appropriately. This is summarized in the following table. Which one to use in which situation depends on the size of matrices involved (e.g., EVD of $A'A$ is much faster than SVD of A when A is very tall), how crucial the numerical accuracy is, etc. In both CANO and CPCA, $Q_{N/S} = I$ (i.e., $\mathrm{Sp}(S) \cap \mathrm{Ker}(T) = \{0\}$) which simplifies the formula considerably.

A little more elegant way of computing $\mathrm{CANO}(G, H)$ is to combine more than one kind of decomposition. $\mathrm{CANO}(G, H)$ may be solved by obtaining $\mathrm{QSVD}(\mathrm{PSVD}(I, G', H'), G', H')$, which De Moor (1991) calls QPPQ-SVD, or by $\mathrm{PSVD}(\mathrm{QSVD}(I, G, H), G, H)$, which De Moor calls PQQP-SVD. Similarly, $\mathrm{CPCA}(A, G, H, K, L)$ may be solved by $\mathrm{QSVD}(\mathrm{PSVD}(A^*, G^{*\prime}, H^{*\prime}), G^{*\prime},$

Table 1. CANO and CPCA Solutions by Various EVD's and SVD's.

	CANO(G, H)	CPCA(Z, G, H, K, L)
SVD	$A = P_G P_H$	$A = R_K' P_{G/K} Z P_{H/L}' R_L$
EVD	$S = P_H P_G P_H$	$S = R_L' P_{H/L} Z' K P_{H/L}' R_L$
PSVD	$A = G'H$, $B = G(G'G)^-$, $C = H(H'H)^-$	$A = G'ZH$, $B = G(G'G)^-$, $C = H(H'H)^-$
QSVD	$A = P_G H$, $C = H$	$A = R_K' P_{G/K} ZLH$, $C = R_L' H$
RSVD	$A = G'H$, $B = G$, $C = H$	$A = G'KZLH$, $B = R_K' G$, $C = R_L' H$
GEVD	$S = H' P_G H$, $T = H'H$	$S = H'LZ'KP_{G/K} ZLH$, $T = H'LH$
GSVD	$A = (G'G)^- G'H(H'H)^-$, $K = G'G$, $L = H'H$	$A = (G'KG)^- G'KZLH(H'LH)^-$, $K = GKG$, $L = H'LH$

$H^{*\prime}$) or PSVD(QSVD$(A^*, G^*, H^*), G^*, H^*$). Since A^*, G^*, and H^* are themselves products of two matrices, an even more sophisticated approach is to eliminate the multiplications to form these products altogether.

References

1. de Leeuw J (1982) Generalized eigenvalue problems with positive semi-definite matrices. Psyhometrika 47: 87–93
2. De Moor B (1991) Generalizations of the OSVD: Structure, properties and applications. In: Vaccaro RJ (ed) SVD and signal processing, II: Algorithms, analysis and applications. Elsevier, Amsterdam, pp 83–98
3. De Moor BLR, Golub GH (1991) The restricted singular value decomposition: Properties and applications. SIAM Journal: Matrix Analysis and Applications 12: 401–425
4. Eckart C, Young G (1936) The approximation of one matrix by another of lower rank. Psychometrika 1: 211–218
5. Fernando KV, Hammarling SJ (1988) A product induced singular value decomposition for two matrices and balanced realization. In: Datta BN, Johnson CR, Kaashoek MA, Plemmons R, Sontag E (eds) Linear algebra in signal systems and control. SIAM, Philadelphia, pp 128–140
6. Golub GH, Van Loan CF (1989) Matrix computations (2nd Ed). Johns Hopkins University Press, Baltimore
7. Guttman L (1944) General theory and methods of matric factoring. Psychometrika 9: 1–16
8. Guttman L (1957) A necessary and sufficient formula for matric factoring. Psychometrika 22: 79–81

9. Koene R, Takane Y (1999) Discriminant component pruning: Regularization and interpretation of multi-layered back-propagation networks. Neural Computation 11: 783–802
10. Mirsky L (1960) Symmetric gage functions and unitarily invariant norms. Quarterly Journal of Mathematics 11: 50–59
11. Paige CC (1986) Computing the generalized singular value decomposition. SIAM Journal: Scientific Computing 7: 1126–1146
12. Paige CC, Saunders MA (1981) Towards a generalized singular value decomposition. SIAM Journal: Numerical Analysis 18: 398–405
13. Rao CR, Mitra SK (1971) Generalized inverse of matrices and its applications. Wiley, New York
14. Takane Y, Shibayama T (1991) Principal component analysis with external information on both subjects and variables. Psychometrika 56: 97–120
15. Takane Y, Hunter MA (2001) Constrained principal component analysis: A comprehensive theory. Applicable Algebra in Engineering, Communication and Computing 12: 391–419
16. Takane Y, Hwang, H (2002) Generalized constrained canonical correlation analysis. Multivariate Behavioral Research 37: 163–195
17. Takane Y, Yanai H, Hwang H (2002) An improved method for generalized constrained canonical correlation analysis. Submitted to Computational Statistics and Data Analysis.
18. ten Berge JMF (1993) Least squares optimization in multivariate analysis. DSWO Press, Leiden
19. Van Loan CF (1976) Generalizing the singular value decomposition. SIAM Journal: Numerical Analysis 13: 76–83
20. Wedderburn JHM (1934) Lectures on matrices. Colloquium Publications 17, Providence, American Mathematical Society
21. Yanai H, Takane Y (1992) Canonical correlation analysis with linear constraints. Linear Algebra and Its Applications 176: 75–89
22. Zha H (1991a) The restricted singular value decomposition of matrix triplets. SIAM Journal: Matrix Analysis and Applications 12: 172–194
23. Zha H (1991b) The product singular value decomposition of matrix trplets. BIT 31: 711–726

In Yanai, H., Okada, A., Shigemasu, K., Kano, Y., and Meulman, J. J. (Eds.), *New developments in psychometrics* (pp. 45-56). Tokyo: Springer, 2002.