Journal of Computational Mathematics, Vol.11, No.1, 1993, 50–62.

SOME PROPERTIES OF THE QUOTIENT SINGULAR VALUE DECOMPOSITION*1)

Hong-yuan Zha
(Scientific Computing and Computational Mathematics, Stanford University, Stanford, USA)

Abstract

A new derivation is given for the Quotient Singular value Decomposition (QSVD) of matrix pair (A, B) having the same number of columns. Certain properties of the quotient singular values are proved. The relation between the QSVD and the SVD is analyzed in some detail.

§1. Introduction

In this paper, we will discuss some properties of the Quotient Singular Value Decomposition (QSVD) of a matrix pair (A, B). The QSVD was first proposed by Van loan[18], who used the name the B-singular value decomposition, and further generalized by Paige and Saunders[12], where the name the generalized singular value decomposition was used. We adopt in this paper the name QSVD in accordance with the standardized nomenclature proposed in [5]. Numerical algorithms for computing the QSVD were developed in [10], [15], [19]. Parallel implementations can be found in [1]. There are quite a few papers discussing applications of the QSVD, for example [3], [9], [11], [14], [17]. As pointed out by Speiser, the QSVD together with matrix-vector multiplication, orthogonal triangular decomposition (QR decomposition) and the SVD forms the core linear algebra operations required in most signal processing problems [13]. Despite all those efforts, there are still some questions associated with the QSVD that deserve further investigation. This paper will analyze some theoretical problems concerning the QSVD: in Section 2, we give a new constructive proof of the QSVD, which, when properly adapted, forms a basis of a numerical algorithm for computing the QSVD2); In Section 3, we propose an algorithm for computing the orthonormal basis of the maximal common row space of two matrices having the same number of columns, and we will show how this problem is intimately connected with the QSVD; we also touch on the problem of computing the orthonormal basis of the maximal common row space of

---

* Received September 4, 1990.

1) Part of the work was supported by NSF grant DRC-8412314.

2) We will not elaborate on the algorithmic aspect of the QSVD in this paper; the reader is referred to [1], [2], [21] for more details.

an arbitrary number od matrices; in Section 4, we generalize the Eckart-Young-Mirsky matrix approximation theorem to handle the case of the quotient singular values; in Section 5, we analyze the relation between the QSVD and the SVD in some detail. A certain form of generalized inverse of matrices generated by the QSVD will also be discussed.

Notation. We also use the following abbreviations in this paper:

$$r_A = \text{rank}\,(A), \quad r_B = \text{rank}\,(B), \quad r_{AB} = \text{rank}\begin{pmatrix} A \\ B \end{pmatrix}.$$

Throughout the paper, matrices are denoted by capitals, vectors by lower case letters. The symbol $R^{m\times n}$ represents the set of $m\times n$ real matrices. $\|\cdot\|$ is the spectrum norm and $\|\cdot\|_F$ the Frobenius norm. The identity matrix of order j is denoted by I_j; we will omit the subscript when the dimension is clear from the context. A zero matrix is denoted by O with various dimensions. We also adopt the following convention for block matrices: whenever a dimension indicating integer in a block matrix is zero, the corresponding block row or column should be omitted, and all expressions and equations in which a block matrix of that block row or block column appears, can be discarded.

§2. A New Constructive Proof of the QSVD

In this section, we will give a constructive proof of the QSVD using the SVD and the Gaussian elimination technique. The presentation of the theorem is a dual and slightly generalized version of Theorem 2.3 in [11], where the case of two matrices having the same number of rows is discussed. The techniques used in our proof are quite different from those in [12] and [18]. Extension of the techniques to handle the case of matrix triplets can be found in [20]. As a further generalization in [6] we have provided a systematic and unified treatment for a tree of generalizations of the SVD for any number of matrices with compatible dimensions.

Theorem 2.1. *Let $A \in R^{m\times n}$ and $B \in R^{p\times n}$ have the same number of columns. Then there exist orthogonal matrices U, V and Q such that*

$$U^T AQ = \Sigma_A(L, O), \quad V^T BQ = \Sigma_B(L, O), \tag{2.1}$$

with

$$\Sigma_A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} I & O & O \\ O & C & O \\ O & O & O \end{pmatrix} \end{array}, \quad \Sigma_B = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} O & O & O \\ O & S & O \\ O & O & I \end{pmatrix} \end{array} \tag{2.2}$$

where

$$C = \text{diag}\,(\alpha_{r+1}, \cdots, \alpha_{r+s}), \quad S = \text{diag}\,(\beta_{r+1}, \cdots, \beta_{r+s}),$$

and

$$1 > \alpha_{r+1} \geq \cdots \geq \alpha_{r+s} > 0, \quad 0 < \beta_{r+1} \leq \cdots \leq \beta_{r+s} < 1,$$

$$\alpha_i^2 + \beta_i^2 = 1, \quad i = r+1, \cdots, r+s. \tag{2.3}$$

The block dimensions of Σ_A and Σ_B are as follows:

	block columns of Σ_A and Σ_B	block rows of Σ_A	block rows of Σ_B
1	r	r	$p-k+r$
2	s	s	s
3	$k-r-s$	$m-r-s$	$k-r-s$

where

$$r = r_{AB} - r_B, \quad s = r_A + r_B - r_{AB}, \quad k = r_{AB}.$$

The structure of L is the following: $L = (L_{ij})_{i,j=1}^3$ is lower triangular with L_{11} and L_{33} diagonal matrices of order r and $k-r-s$ respectively. The diagonal elements of L_{11} and L_{33} are positive and are ordered in nonincreasing magnitude. Specifically, let $\sigma_i = \alpha_{r+i}/\beta_{r+i}$ $(i = 1, \cdots, s)$ and $CS^{-1} = \mathrm{diag}(\sigma_{i_1} I_{s_1}, \cdots, \sigma_{i_l} I_{s_l})$, where $\sigma_{i_1} > \cdots > \sigma_{i_l}$ with $\sum_{i=1}^l s_i = s$, represent the distinct elements of the diagonal entries of CS^{-1}. Then the (i,i) diagonal block $L_{ii}^{(2)} \in R^{s_i \times s_i}$ of $L_{22}(i = 1, \cdots, l)$ are diagonal matrices with positive diagonal entries arranged in nonincreasing magnitude.

Proof. The proof consists of four recursive steps, each of which brings us closer to the desired form in (2.1). In each of the first three steps, the transformation to the next step is of the following form:

$$A_{k+1} = U_k^T A_k P_k, \quad B_{k+1} = V_k^T B_k P_k, \tag{2.4}$$

where U_k and V_k are orthogonal and P_k is nonsingular. The matrices A_k and B_k are the transformed results of A and B at step k, which are initially set to

$$A_1 = A, \quad B_1 = B.$$

Step 1. Applying the SVD of B, we can decompose B as

$$(U^{(1)})^T B V^{(1)} = \begin{pmatrix} O & O \\ O & \Sigma_1 \end{pmatrix},$$

where $\Sigma_1 = \mathrm{diag}(s_1, \cdots, s_t)$ and $s_1 \geq \cdots \geq s_t > 0$, and we have $t = \mathrm{rank}(B)$. Set

$$U_1 = I, \quad V_1 = U^{(1)}, \quad P_1 = V^{(1)} \mathrm{diag}(I, \Sigma_1^{-1}),$$

and recall equation (2.4) with $k = 1$. The pair (A, B) is transformed to

$$A_2 = (\overset{n-t}{A_1^{(2)}}, \overset{t}{A_2^{(2)}}) \text{ and } B_2 = \begin{pmatrix} O & O \\ O & I_t \end{pmatrix}.$$

Step 2. Let the SVD of $A_1^{(2)}$ be

$$(U^{(2)})^T A_1^{(2)} V^{(2)} = \begin{pmatrix} \Sigma_2 & O \\ O & O \end{pmatrix},$$

where $\Sigma_2 = \mathrm{diag}(t_1, \cdots, t_r)$ and $t_1 \geq \cdots \geq t_r > 0$, i.e., $r = \mathrm{rank}(A_2^{(2)})$. Setting

$$U_2 = U^{(2)}, \quad V_2 = I, \quad P_2 = \mathrm{diag}(V^{(2)}, I)\, \mathrm{diag}(\Sigma_2^{-1}, I)$$

leads to

$$A_2 = \begin{array}{cc} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} I & O & A_{13}^{(3)} \\ O & O & A_{23}^{(3)} \end{pmatrix} \end{array},$$

	1	2	3
row	r	$m-r$	
column	r	$n-r-t$	t

The block dimensions of A_2 are shown in the tabular on the right. Moreover, B_3 remains the same $B_3 = B_2$.

Step 3. Let the SVD of $A_{23}^{(3)}$ be

$$(U^{(3)})^T A_{23}^{(3)} V^{(3)} = \begin{pmatrix} \Sigma_3 & O \\ O & O \end{pmatrix},$$

where $\Sigma_3 = \operatorname{diag}(w_1, \cdots, w_s)$ and $w_1 \geq \cdots \geq w_s > 0$, with $s = \operatorname{rank}(A_{23}^{(3)})$. Furthermore, let $\alpha_i = w_i(1+w_i^2)^{-1/2}$ and $\beta_i = (1+w_i^2)^{-1/2}$, $i = r+1, \cdots, r+s$, and

$$C = \operatorname{diag}(\alpha_{r+1}, \cdots, \alpha_{r+s}), \quad S = \operatorname{diag}(\beta_{r+1}, \cdots, \beta_{r+s}).$$

It is easy to check that $\alpha_i, \beta_i (i = r+1, \cdots, r+s)$ satisfy (2.3). Setting

$$U_3 = \operatorname{diag}(I, U^{(3)}), \quad V_3 = \operatorname{diag}(I, (V^{(3)})^T),$$

$$P_3 = \begin{pmatrix} I & -A_{13}^{(3)} \\ O & I \end{pmatrix} \operatorname{diag}(I, V^{(3)}) \operatorname{diag}(I, S, I),$$

gives rise to

$$A_4 = \begin{array}{cc} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} I & O & O & O \\ O & O & C & O \\ O & O & O & O \end{pmatrix} \end{array},$$

	1	2	3	4
row	r	s	$m-r-s$	
column	r	$n-r-t$	s	$t-s$

and similarly

$$B_4 = \begin{array}{cc} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} O & O & O \\ O & S & O \\ O & O & I \end{pmatrix} \end{array},$$

	1	2	3
row	$p-k+r$	s	$k-r-s$
column	$n-t$	s	$t-s$

After applying certain suitable permutations Π_1 and Π_2, which of cause are orthogonal transformations, and setting $k = t + r$, we obtain

$$A_5 = A_4\Pi_1 = (\Sigma_A, O), \quad B_5 = \Pi_2 A_4 \Pi_1 = (\Sigma_B, O),$$

where Σ_A and Σ_B are of the form in (2.2).

Step 4. We first accumulate all the transforms applied to the right-hand side of (A, B) in the previous steps to a single matrix: $P := P_1 P_2 P_3 \Pi_1$. Using a variant of the QR decomposition, we can factorize P as

$$P = Q_1 \begin{pmatrix} L_{11}^{(1)} & O \\ L_{21}^{(1)} & L_{22}^{(1)} \end{pmatrix}^{-1},$$

where $L_{11}^{(1)} \in R^{k\times k}$ is lower triangular, and Q_1 is orthogonal. Let $L_{11}^{(1)} = (\tilde{L}_{ij}^{(1)})_{i,j=1}^3$ be partitioned compatibly with the block column partitioning of Σ_A and Σ_B, and let the SVD of $\tilde{L}_{11}$ and $\tilde{L}_{33}$ be

$$\tilde{L}_{11} = U_1^{(4)} L_{11} (V_1^{(4)})^T, \quad \tilde{L}_{33} = U_2^{(4)} L_{33} (V_2^{(4)})^T$$

where L_{11} and L_{33} are diagonal matrices with positive diagonal entries arranged in nonincreasing order; similarly let the SVD of $\hat{L}_{ii}^{(2)} \in R^{s_i\times s_i}$, the (i,i) diagonal block of $\tilde{L}_{22}$ $(i = 1,\cdots,l)$, be of the following form

$$(\hat{U}_i)^T \hat{L}_{ii}^{(2)} \hat{V}_i = L_{ii}^{(2)}, \quad i = 1,\cdots,l$$

where $L_{ii}^{(2)}$ is diagonal with positive diagonal entries arranged in nonincreasing order. Let

$$\tilde{U} = \operatorname{diag}(\hat{U}_1,\cdots,\hat{U}_l), \quad \tilde{V} = \operatorname{diag}(\hat{V}_1,\cdots,\hat{V}_l),$$

and accumulate orthogonal transformations:

$$U^T = \operatorname{diag}((U_1^{(4)})^T, \tilde{U}, (U_2^{(4)})^T)\Pi_2 U_3 U_2 U_1,$$

$$V^T = \operatorname{diag}((U_1^{(4)})^T, \tilde{U}, (U_2^{(4)})^T) V_3 V_2 V_1,$$

and the nonsingular transformations:

$$Q = \operatorname{diag}(V_1^{(4)}, \tilde{V}, V_2^{(4)}) Q_1.$$

Then, we obtain the desired decomposition. The expressions for the integer indices can be derived from

$$r_A = r + s, \quad r_B = k - r, \quad r_{AB} = k.$$

With a little more elementary calculation, we can also obtain the block dimensions for Σ_A and Σ_B.

Remark 1. Let $P = Q \operatorname{diag}(L^{-1}, I)$. Then P is nonsingular and

$$U^T AP = (\Sigma_A, O), \quad V^T BP = (\Sigma_B, O).$$

We will use this variant of Theorem 2.1 to simplify the presentations in the proofs of some of the later results in this paper.

According to Paige [12], corresponding to each column in (2.1) is ascribed a quotient singular value pair (α_i, β_i). Referring to (2.1), we take for the first k of those as

$$\begin{cases} \alpha_i = 1, \quad \beta_i = 0, & i = 1,\cdots,r, \\ \alpha_i, \beta_i, \text{ as in } C \text{ and } S, & i = r+1,\cdots,r+s, \\ \alpha_i = 0, \quad \beta_i = 1, & i = r+s+1,\cdots,k, \end{cases} \tag{2.5}$$

and call them the nontrivial quotient singular value pairs of (A,B). The quotients α_i/β_i $(i = 1,\cdots,k)$ are called the quotient singular values of (A,B). The other $n-k$ pairs corresponding to the zero columns in (2.1) are called trivial quotient singular pairs of (A,B). They also correspond to the common column null space of A and B.

§3. Computing the Intersection of the Row Spaces of Two Matrices

It is readily seen from Remark 1 that if we partition the rows of P^{-1} compatibly with the block column partitioning of (Σ_A, O) such that

$$P^{-1} = \begin{array}{c} r \\ s \\ k-r-s \\ n-k \end{array} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{pmatrix},$$

then the rows of P_2 form the basis of the maximal common row space of A and B, or the intersection of the row spaces of A and B. However, the QSVD provides a richer structure than a basis of the intersection. Basically what the QSVD does is to regroup the rows of A and B by applying orthogonal transformations to the left of A and B, which results in three groups: (i) those row vectors that are in the row space of A, but not in the row space of B; (ii) those row vectors that are in the row spaces of both A and B; (iii) those row vectors that are in the row space of B, but not in the row space of A. The quotient singular values or the quotient singular values pairs come from employing a special arrangement of the vectors in group (ii), i.e., each of the row vectors of $V^T B$ is a positive scalar multiple of (or parallel to) the corresponding row vectors of $U^T A$ in this group; and the scalar multipliers are exactly the quotient singular values of (A, B). As a matter of fact, it is not necessary to compute the full QSVD of (A, B) in order to obtain a basis of the intersection. Based on the idea in the constructive proof of the QSVD in Section 2, we will present an algorithm for computing the maximal common row space of two matrices having the same number of columns. In contrast to the algorithm in [9], which is based on the constructive proof of the QSVD due to Paige and Saunders [12], and the algorithm proposed in [4], the following algorithm can be implemented using only the QR decomposition. On the other hand, whenever doubts arise as to the rank decision of certain submatrices, the QR decomposition with column pivoting, the rank revealing QR decomposition, or even the SVD can be used to enhance the ability to detect the rank deficiency in the following algorithm.

Given $A \in R^{m\times n}$ and $B \in R^{p\times n}$, let $R_{\text{row}}(\cdot)$ denote the row space of a matrix. Then the following algorithm computes an orthonormal basis of $R_{\text{row}}(A) \cap R_{\text{row}}(B)$. Compress the columns of B such that

$$B = (O, B_2)V_1,$$

where B_2 has full column rank and V_1 is orthogonal. Partition

$$AV_1^T := (A_1, A_2)$$

compatibly with the column block partitioning of B, and compress the rows of A_1 such that

$$A_1 = U^T \begin{pmatrix} A_{11} \\ O \end{pmatrix}$$

where U is orthogonal and A_{11} has full row rank; write

$$A = U^T \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix} V_1.$$

Compress the columns of A_{22} such that

$$A_{22} = (A_{22}^{(1)}, O)V_2,$$

where $A_{22}^{(1)}$ has full column rank and V_2 is orthogonal. Let $V = \operatorname{diag}(I, V_2)V_1$ and write the final form as

$$A = U^T \begin{pmatrix} A_{11} & A_{12}^{(1)} & A_{13}^{(1)} \\ O & A_{22}^{(1)} & O \end{pmatrix} V, \quad B = (O, B_2^{(1)}, B_3^{(1)})V.$$

Partition the rows of V compatibly with the block column partitioning of A such that

$$V = \begin{pmatrix} V^{(1)} \\ V^{(2)} \\ V^{(3)} \end{pmatrix}.$$

Then the rows of $V^{(2)}$ form the orthonormal basis of the maximal row space of A and B.

The correctness of the above algorithm is proved in the following theorem. It can also be extended to handle the case of an arbitrary number of matrices.

Theorem 3.1. *Using the notation as in the above algorithm, we have*

$$R_{\text{row}}(V^{(2)}) = R_{\text{row}}(A) \cap R_{\text{row}}(B).$$

Proof. Since both matrices $A_{22}^{(1)}$ and $B_2^{(1)}$ have full column rank, it is not difficult to verify that

$$R_{\text{row}}(V^{(2)}) \subseteq R_{\text{row}}(A) \cap R_{\text{row}}(B). \tag{3.6}$$

Now let $z \in R_{\text{row}}(A) \cap R_{\text{row}}(B)$. Then there exist row vectors x and y such that

$$z = xA =: (x_1, x_2) \begin{pmatrix} A_{11} & A_{12}^{(1)} & A_{13}^{(1)} \\ O & A_{22}^{(1)} & O \end{pmatrix} V$$

and similarly

$$z = yB = y(O, B_2^{(1)}, B_3^{(1)})V.$$

Comparing the right-hand side of the above two equations leads to

$$x_1 A_{11} = 0, \quad x_2 A_{22}^{(1)} = yB_2, \quad yB_3 = x_1 A_{13}^{(1)}.$$

Therefore

$$x_1 = 0, \quad yB_3 = 0,$$

and

$$z = yB = (yB_2)V^{(2)} \in R_{\text{row}}(V^{(2)}).$$

Hence we conclude that

$$R_{\text{row}}(A) \cap R_{\text{row}}(B) \subseteq R_{\text{row}}(V^{(2)}).$$

Combine this with (3.6), and the theorem is proved.

Remark 2. The above algorithm can be readily extended to the case of computing the intersection of several linear subspaces. As an illustration we consider the case of three matrices: A, B, C having the same number of columns. Similarly to the above algorithm, we can construct orthogonal matrices U_A, U_B, U_C and V such that

$$\begin{pmatrix} U_A A V^T \\ \hline U_B B V^T \\ \hline U_C C V^T \end{pmatrix} = \left(\begin{array}{cccc} A_{11} & A_{12} & A_{13} & A_{14} \\ O & A_{22} & A_{23} & A_{24} \\ O & O & O & A_{34} \\ \hline B_{11} & B_{12} & B_{13} & B_{14} \\ O & O & B_{23} & B_{24} \\ \hline O & C_2 & C_3 & C_4 \end{array}\right)$$

where $(C_2, C_3, C_4), (B_{23}, B_{24})$ and A_{34} are of full column rank, while $(B_{11}, B_{12}), A_{11}$ and A_{22} are of full row rank. Let A_{34} have column dimension r. We can similarly prove that the last r rows of V form the orthonormal basis of

$$R_{\text{row}}(A) \cap R_{\text{row}}(B) \cap R_{\text{row}}(C).$$

Extensions to the general case is straightforward.

§4. Rank Inequality Characterization of the Quotient Singular Values

In this section, we prove a generalization of the Eckart-Young-Mirsky theorem. We only consider the case of spectrum norm, although generalization of the case of orthogonally invariant norms is straightforward. We first cite the well known result:

Lemma 4.1 (Eckart-Young-Mirsky). *Let the singular values of A be*

$$\sigma_1 \geq \cdots \geq \sigma_n \geq 0.$$

Then

$$\sigma_i = \min_{E \in R^{m\times n}} \{\|E\| \,|\, \text{rank}\,(A+E) \leq i-1\}, \quad i = 1, \cdots, n,$$

and there exists a matrix E_i (not necessarily unique) satisfying $\|E_i\| = \sigma_i$ such that

$$\text{rank}\,(A + E_i) = i - 1, \quad i = 1, \cdots n.$$

The above result is the theoretical basis of the total least squares methods and the truncated SVD method. The following theorem generalizes the above result to the case of quotient singular values.

Theorem 4.1. *Let the QSVD of A and B be given as in Theorem 2.1. Then*

(a) *The quotient singular values can be characterized as*

$$\alpha_i/\beta_i = \min_{E \in R^{m\times p}} \{\|E\| \,|\, \text{rank}\,(A + EB) \leq i-1\}, \quad i = 1, \cdots, k.$$

(b) *Let $l = r_{AB} - r_B$ and $u = \min(m, r_{AB})$. Then for any $E \in R^{m\times p}$ we have*

$$l \leq \text{rank}\,(A + EB) \leq u,$$

and for any integer i satisfying $l \le i \le u$, there exists $E_i \in R^{m\times p}$ such that

$$\operatorname{rank}(A + E_i B) = i.$$

Proof. Using the notation in Theorem 2.1 and Remark 1, for arbitrary $E \in R^{m\times p}$, let $U^T E V = (E_{ij})_{i,j=1}^3$ be partitioned compatibly with the partitionings of Σ_A and Σ_B. Then

$$\begin{aligned}\operatorname{rank}(A+EB) &= \operatorname{rank}(U^T AP + U^T EVV^T BP)\\ &= \operatorname{rank}\left\{\begin{pmatrix} I & E_{12}S & E_{13} & O\\ O & C+E_{22}S & E_{23} & O\\ O & E_{32}S & E_{33} & O\end{pmatrix}\right\}\\ &= r + \operatorname{rank}\left\{\begin{pmatrix} CS^{-1} & O\\ O & O\end{pmatrix} + \begin{pmatrix} E_{22} & E_{23}\\ E_{32} & E_{33}\end{pmatrix}\right\}.\end{aligned}$$

Applying Lemma 4.1 to the two matrices in the above leads to the assertion in (a). Part (b) is a consequence of the following expressions:

$$k = r_{AB}, \quad r = r_{AB} - r_B,$$

which are established in Theorem 2.1.

Remark 3. The above theorem is also a generalization of the main theorem in [8], where application to the total least squares problem with partial exact columns is discussed; see also the algorithm in [17].

§5. Relation Between the QSVD and the SVD

It is easy to check that the quotient singular values of (A, B) are just the singular values of AB^{-1}, if B is nonsingular. In this section, we will further discuss the case that B is a general matrix.

Theorem 5.1. *Use the notation in Theorem 2.1 and assume that $r_{AB} = n$. Then*

$$B_A^+ := P\begin{pmatrix} O & O & O\\ O & S^{-1} & O\\ O & O & I_{n-s-r}\end{pmatrix} V^T$$

is uniquely defined and the singular values of AB_A^+ contain the noninfinite quotient singular values of (A, B).

Proof. Since $r_{AB} = n$, we can verify that any two sets of transformations in QSVD satisfy the following equations:

$$P_1 = P_2 \operatorname{diag}(U_{11}, U_{22}, V_{33}),\ U_1 = U_2^T \operatorname{diag}(U_{11}, U_{22}, U_{33}),\ V_1 = V_2^T \operatorname{diag}(V_{11}, V_{22}, V_{33}).$$

Hence

$$P_1 \begin{pmatrix} O & O & O \\ O & S^{-1} & O \\ O & O & I_{n-s-r} \end{pmatrix} V_1^T$$

$$= P_2 \begin{pmatrix} U_{11} & O & O \\ O & U_{22} & O \\ O & O & U_{33} \end{pmatrix} \begin{pmatrix} O & O & O \\ O & S^{-1} & O \\ O & O & I_{n-s-r} \end{pmatrix} \begin{pmatrix} V_{11}^T & O & O \\ O & U_{22}^T & O \\ O & O & V_{33}^T \end{pmatrix} V_2^T$$

$$= P_2 \begin{pmatrix} O & O & O \\ O & S^{-1} & O \\ O & O & I_{n-s-r} \end{pmatrix} V_2^T.$$

Therefore B_A^+ is uniquely determined. We also observe that

$$U^T A B_A^+ V = \mathrm{diag}(O, CS^{-1}, O),$$

and only the infinite quotient singular values of (A, B) are changed to zero ordinary singular values of AB_A^+; the others are preserved.

Corollary 5.1. *Let B^+ be the Moore-Penrose inverse of B. If B has full column rank, then the ordinary singular values of AB^+ contain the noninfinite quotient singular values of (A, B).*

Proof. The corollary is a consequence of Lemma 5.1 and the fact that $B_A^+ = B^+$, if B has full column rank.

Corollary 5.2[16]**.** *Let X be of full column rank, and x_j (x_j^+) be the j-th column (row) of X. Then the smallest perturbation e_j in x_j that will make X collinear (i.e. rank deficient) satisfies*

$$\frac{\|e_j\|}{\|x_j\|} = \kappa_j^{-1},$$

where $\kappa_j = \|x_j\| \, \|x_j^+\|$ is called the j-th collinear index of X.

Proof. The result follows from the above corollary and Theorem 4.1.

The following theorem compares the quotient singular values of (A, B) and the singular values of AB^+ and BA^+.

Theorem 5.2. *Let the QSVD of the matrix pair (A, B) be given in Theorem 2.1, and write $\sigma_i = \alpha_{r+i}/\beta_{r+i}$ ($i = 1, \cdots, s$); let the ordinary singular values of AB^+ be ordered in nonincreasing magnitude, while the singular values of BA^+ are ordered in nondecreasing magnitude. Then*

$$1/\sigma_i(BA^+) \le \sigma_i \le \sigma_i(AB^+), \quad i = 1, \cdots, s.$$

Proof. Using Lemma 4.1, for any integer i satisfying $1 \le i \le s$, we can find an $E_i \in R^{m\times p}$ satisfying $\|E_i\| = \sigma_i$ such that

$$\mathrm{rank}\,(AB^+ + E_i) = i - 1.$$

Then

$$\mathrm{rank}\,(A + E_i B) = \mathrm{rank}\,((AB^+ + E_i)B + A(I - B^+B)) \le i - 1 + r.$$

Using Theorem 4.1, we obtain

$$\sigma_i(AB^+) = \|E_i\| \geq \sigma_i.$$

Interchange the roles of A and B we can prove the other part of the inequality.

Corollary 5.3. *Let the nonzero finite quotient singular values of the matrix pairs* (A_1, B_1) *and* (A_2, B_2) *be*

$$\sigma_1^{(1)} \geq \cdots \geq \sigma_{s_1}^{(1)} > 0; \quad \sigma_1^{(2)} \geq \cdots \geq \sigma_{s_2}^{(2)} > 0.$$

And the singular values of $A_1B_1^+$ *and* $A_2B_2^+$ ($B_1A_1^+$ *and* $B_2A_2^+$) *are arranged in non-increasing (nondecreasing) order. Then*

$$\frac{\sigma_i(B_2A_2^+)}{\sigma_i(B_1A_1^+)} \leq \frac{\sigma_i^{(1)}}{\sigma_i^{(2)}} \leq \frac{\sigma_i(A_1B_1^+)}{\sigma_i(A_2B_2^+)}, \quad i = 1, \cdots, \min(s_1, s_2).$$

Proof. The inequalities follow directly from the above theorem.

We can similarly define A_B^+ as

$$A_B^+ = P \begin{pmatrix} I & O & O \\ O & C^{-1} & O \\ O & O & O \end{pmatrix} U^T.$$

The relation of A_B^+ and B_A^+ is given by the following interesting relations:

$$(A_B^+)_{B_A^+}^+ = A, \quad (B_A^+)_{A_B^+}^+ = B.$$

They follow directly from the definitions, and can be considered as a generalization of the relation $(A^{-1})^{-1} = A$. The QSVD of (A, B) is equivalent to the QSVD of (A_B^+, B_A^+) in the sense that we can derive one from the other.

To conclude the section, we discuss some properties of B_A^+. It is easy to check that B_A^+ satisfies the following equations:

$$BB_A^+B = B, \quad B_A^+BB_A^+ = B_A^+, \quad (BB_A^+)^T = BB_A^+.$$

Therefore B_A^+ is a $\{1,2,3\}$-inverse of B. We give the following two results which indicate how to uniquely characterize B_A^+ in the class of $\{1,2,3\}$-inverse of B.

Theorem 5.3. *If* $r_{AB} = n$, *then* B_A^+ *is the unique solution of the following constrained minimization problem:*

$$\min_{X \in R^{n\times p}} \|AX\|_F$$

subject to

$$BXB = B, \quad XBX = X, \quad (BX)^T = BX. \tag{5.7}$$

The minimum value is $\sqrt{\sum_{i=r+1}^{r+s}(\alpha_i/\beta_i)^2}$.

Proof. Since $r_{AB} = n$, we can write

$$B = V\Sigma_B P^{-1}.$$

Partition $P^{-1}XV := (X_{ij})_{i,j=1}^3$ compatibly with the partitionings of Σ_A and Σ_B. We can verify that X must be of the following form:

$$X = P \begin{pmatrix} O & X_{12} & X_{13} \\ O & S^{-1} & O \\ O & O & I_{n-r-s} \end{pmatrix} V$$

in order to satisfy the constraints (5.7). Since

$$\|AX\|_F^2 = \|U^T APP^{-1}XV\|_F^2 = \left\| \begin{pmatrix} I_r & O & O \\ O & C & O \\ O & O & O \end{pmatrix} \begin{pmatrix} O & X_{12} & X_{13} \\ O & S^{-1} & O \\ O & O & I_{n-r-s} \end{pmatrix} \right\|_F^2$$

$$= \|X_{12}, X_{13})\|_F^2 + \|CS^{-1}\|_F^2 \geq \|CS^{-1}\|_F^2 = \sum_{i=r+1}^{r+s} (\alpha_i/\beta_i)^2,$$

the equality is satisfied if and only if $X_{12} = O$ and $X_{13} = O$, namely $X = B_A^+$.

In plain English, this theorem tells us that B_A^+ is the unique {1, 2, 3}-inverse of B that minimizes $\|AX\|_F$ over all { 1, 2, 3}-inverse of B. In the following, we give another characterization of B_A^+. It is generalization of the Moor-Perose equations.

Theorem 5.4. *If $r_{AB} = n$, then B_A^+ is the unique solution of the following four equations:*

$$BXB = B, \quad XBX = X, \quad (BX)^T = BX. \quad (A^T AXB)^T = A^T AX. \tag{5.8}$$

Proof. As in the proof of the above theorem, X must be of the following form:

$$X = P \begin{pmatrix} O & X_{12} & X_{13} \\ O & S^{-1} & O \\ O & O & I_{n-r-s} \end{pmatrix} V$$

in order to satisfy the first three constraints in (5.8). Since

$$A^T AX = P^{-T} \begin{pmatrix} I_r & O & O \\ O & C & O \\ O & O & O \end{pmatrix} UU^T \begin{pmatrix} I_r & O & O \\ O & C & O \\ O & O & O \end{pmatrix} P^{-1}P$$

$$= \begin{pmatrix} O & X_{12} & X_{13} \\ O & S^{-1} & O \\ O & O & I_{n-r-s} \end{pmatrix} VV^T \begin{pmatrix} I_r & O & O \\ O & S^{-1} & O \\ O & O & O \end{pmatrix} Q^{-1}$$

$$= Q^{-T} \begin{pmatrix} O & X_{12} & X_{13} \\ O & S^{-1} & O \\ O & O & I_{n-r-s} \end{pmatrix} Q^{-1}$$

$(A^T AXB)^T = A^T AXB$ is and only if $X_{12} = O$ and $X_{13} = O$, i.e., $X = B_A^+$.

Remark 4. The four quartions in (5.8) are a special case of the four equations in [7]. So the aobve theorem answers the open question of the uniqueness of the solution of the four equations in [7] under the condition that $(K^T, L^T)^T$ has full column rank and $M = I$.

References

[1] Z. Bai, The direct GSVD algorithm and its parallel implementation, CS-TR-1901, 1987, University of Maryland.

[2] Z. Bai and J. Demmel, Computing the generalized singular value decomposition, Research Report 91-09, Dept. of Math., Univ. of Kentucky.

[3] Å. Björck, Least squares method, in Handbook of Numerical Analysis, Vol.2, Editors, P.G. Ciarlet and J.L. Lions.

[4] Å. Björck and G. Golub, Numerical methods for computing angles between linear subspaces, *Mathematics of Computation*, **27** : 123, (1973), 579–594.

[5] B. De Moor and G. Golub, Generalized singular value decompositions: a proposal for a standardized nomenclature, NA-89-05, 1989, Stanford University.

[6] B. De Moor and H. Zha, A tree of generalizations of the ordinary singular value decomposition, Linear Algebra and Its Applications, 147 (1991), 469–500.

[7] L. Eldén, A weighted pseudoinverse, generalized singular values, and constrained least squares problems, *BIT*, 22 (1982), 487–502.

[8] G. Golub, A. Hoffman and C. Van Loan, A generalization of the Eckart-Young-Mirsky matrix approximation theorem, *Linear Algebra and Its Applications*, 88/89 (1987), 322–327.

[9] M. Moonen, B. De Moor, L.Vandenberghe and J. Vandewalle, On-and off-line identification of linear state space models, *International Journal of Control*, **49** (1989), 219–232.

[10] C. Paige, Computing the generalized singular value decomposition, *SIAM J. Scientific and Statistic Computing*, **7** (1986), 1126–1146.

[11] C. Paige, The general linear model and generalized singular value decomposition, *Linear Algebra and Its Applications*, **70** (1985), 269–284.

[12] C. Paige and M. Saunders, Towards a generalized singular value decomposition, *SIAM J. Numerical Analysis*, **18** (1981), 398–405.

[13] J. Speiser, Linear algebra problems arising in signal processing, Invited presentation 3, SIAM Annual Meeting 1989.

[14] J. Speiser and C. Van Loan, Signal processing using the generalized singular value decomposition, SPIE Vol. 495 Real Time Signal Processing VII, 1984, 47–55.

[15] G. Stewart, A method for computing the generalized singular value decomposition, in Matrix Pencils, Lecture Notes in Mathematics 973.

[16] G. Stewart, Collinearity and the least squares regression, Statistical Science, 2, (1987), 68–100.

[17] S. Van Huffel and J. Vandewalle, Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns of A are subject to errors, *SIAM J. Matrix Analysis and Applications*, **10** (1989), 294–315.

[18] C. Van Loan, Generalizing the singular value decomposition, *SIAM J. Numerical Analysis*, **13** (1976), 76–83.

[19] C. Van Loan, Computing the CS and the generalized singular value decompositions, *Numerische Mathematik*, **46** (1985), 479–492.

[20] H. Zha, Restricted singular value decomposition for matrix triplets, *SIAM J. Matrix Analysis and Applications*, **12** (1991), 172–194.

[21] H. Zha, Computing the restricted singular value decomposition of matrix triplets, to appear in *Linear Algebra and Its Applications*.