Approximating the PSVD and QSVD

Sanzheng Qiao
Department of Computer Science and Systems
McMaster University
Hamilton, Ontario
Canada L8S 4K1
qiao@maccs.dcss.mcmaster.ca

ABSTRACT. This paper presents an adaptive method for approximating the SVD of the product AB or the quotient AB^{-1} of two matrices A and B. Specifically, it computes an approximation of the complete orthogonal decomposition of AB or AB^{-1}. The algorithm is an extension of the ULV decomposition algorithm.

KEYWORDS. PSVD, QSVD, complete orthogonal decomposition, ULV decomposition.

1 Introduction

In this paper we consider the problem of the SVD of the product AB (PSVD) and the quotient AB^{-1} (QSVD) of two matrices A and B. We present an adaptive method for approximating the PSVD or the QSVD. Specifically, this method computes an approximation of the complete orthogonal decomposition [1] of the product AB or the quotient AB^{-1} without explicit matrix multiplication or inversion. The algorithm is a generalization of the ULV decomposition algorithm [2]. For example, it computes the decomposition

$$AB = U \begin{pmatrix} G & E \\ 0 & F \end{pmatrix} V^H \tag{1.1}$$

where U and V are orthonormal, G and F are upper triangular, and $\|E\|_F^2 + \|F\|_F^2$ is approximately the sum of the squares of the "small" singular values of AB.

This algorithm can be generalized to any finite number of matrices.

This paper is organized as follows. In section 2, we present a generalization of the estimator in [3]. It gives an estimated right singular vector associated with the smallest singular value of AB or AB^{-1}. In sections 3 and 4, we extend the deflation and refinement schemes in [2] to the cases of AB and AB^{-1}. Finally, in section 5, we describe the updating procedures.

In the following sections, we use 2-norm for all vector and matrix norms and adopt MATLAB notations, e.g., $A_{i:j,k:l}$ denotes the submatrix of A consisting the entries in rows from i to j and columns from k to l.

2 The Estimator

To begin with, we describe the estimator for the right singular vector corresponding to the smallest singular value of the product AB. Without the loss of generality, we assume both A and B are upper triangular and of order n. The estimator consists of two stages: first, it finds a vector d of unit norm such that the norm of the solution x to the system

$$B^H A^H x = d$$

is large; second, it solves for v in $ABv = x$. Then v is an approximation of the right singular vector associated with the smallest singular value of AB and $\|x\|/\|v\|$ is an approximation of the smallest singular value of AB. We only describe the first stage for the second one is straightforward.

The procedure of finding d is incremental in that d is determined one element at a time. Let $\hat{A}$ and $\hat{B}$ be the order $k-1$ leading principal submatrices of A and B respectively, then the kth order leading principal submatrices are:

$$A(k) \stackrel{\text{def}}{=} \begin{pmatrix} \hat{A} & a \\ 0 & A_{k,k} \end{pmatrix} \quad \text{and} \quad B(k) \stackrel{\text{def}}{=} \begin{pmatrix} \hat{B} & b \\ 0 & B_{k,k} \end{pmatrix} \tag{2.1}$$

where $a = A_{1:k-1,k}$ and $b = B_{1:k-1,k}$. Suppose that we have found a $(k-1)$-vector $\hat{d}$ of unit norm so that the norm of the solution $\hat{x}$ to the system $\hat{B}^H \hat{A}^H \hat{x} = \hat{d}$ is large. Our goal is to increment $\hat{d}$ to a k-vector of unit norm and, at the same time, make the norm of the solution x to $B(k)^H A(k)^H x = d$ large. Because of the unit norm of d, we can assume that

$$d = \begin{pmatrix} s\hat{d} \\ c \end{pmatrix} \qquad \text{where} \quad |c|^2 + |s|^2 = 1. \tag{2.2}$$

Let $\hat{y}$ be the solution to the triangular system $\hat{B}^H \hat{y} = \hat{d}$, then, from (2.1) and (2.2), the solution y to $B(k)^H y = d$ is given by

$$y = \begin{pmatrix} s\hat{y} \\ \eta \end{pmatrix}, \qquad \eta = (c - s\pi)/B^*_{k,k}$$

where c^* denotes the complex conjugate of a complex number c and

$$\pi \stackrel{\text{def}}{=} b^H \hat{y}.$$

Similarly, if $\hat{x}$ is the solution to the triangular system $\hat{A}^H \hat{x} = \hat{y}$, then the solution to $A(k)^H x = y$ is

$$x = \begin{pmatrix} s\hat{x} \\ \xi \end{pmatrix}, \qquad \xi = (\eta - s\delta)/A^*_{k,k}$$

where

$$\delta \stackrel{\text{def}}{=} a^H \hat{x}.$$

In solving a triangular system, we often keep a partial sum vector, for example, in this case

$$\hat{p} \stackrel{\text{def}}{=} {A_{1:k-1,k:n}}^H \hat{x}.$$

Consequently, the first component of $\hat{p}$ is δ and we have the partition

$$\hat{p} \stackrel{\text{def}}{=} \begin{pmatrix} \delta \\ \tilde{p} \end{pmatrix}$$

where $\tilde{p} = {A_{1:k-1,k+1:n}}^H \hat{x}$.

Having solved for x, we can update the partial sum vector using

$$p \stackrel{\text{def}}{=} {A_{1:k,k+1:n}}^H x = s\tilde{p} + \xi {A_{k,k+1:n}}^H.$$

Since the partial sum p will affect the solutions in the subsequent steps, we choose c and s so that the norm

$$\left\| \begin{pmatrix} x \\ p \end{pmatrix} \right\|$$

is maximized.

After some algebraic manipulation, we obtain

$$x = W_1 \begin{pmatrix} c \\ s \end{pmatrix} \quad \text{and} \quad p = W_2 \begin{pmatrix} c \\ s \end{pmatrix}$$

where

$$W_1 = (A^*_{k,k} B^*_{k,k})^{-1} \begin{pmatrix} 0 & B^*_{k,k} A^*_{k,k} \hat{x} \\ 1 & -\pi - B^*_{k,k}\delta \end{pmatrix},$$

and

$$W_2 = (A^*_{k,k} B^*_{k,k})^{-1} ({A_{k,k+1:n}}^H \quad A^*_{k,k} B^*_{k,k} \tilde{p} - (\pi + B^*_{k,k}\delta){A_{k,k+1:n}}^H).$$

Thus $(c, s)^T$ is the normalized right singular vector associated with the larger singular value of the $n \times 2$ matrix

$$W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}.$$

Now we derive the estimator for the right singular vector corresponding to the smallest singular value of the quotient AB^{-1} where, again, both A and B are upper triangular. Similarly, the crucial step is to incremently find a vector d of unit norm such that the norm of the solution x to the system

$$B^{-H} A^H x = d$$

is large.

Using the partitions in (2.1) and (2.2), let $\hat{y} = \hat{B}^H \hat{d}$, then

$$y = \begin{pmatrix} s\hat{y} \\ sb^H \hat{d} + cB^*_{k,k} \end{pmatrix}.$$

The solution x and its corresponding partial sum p are

$$x = \begin{pmatrix} s\hat{x} \\ (sb^H\hat{d} + cB^*_{k,k} - s\delta)/A^*_{k,k} \end{pmatrix}$$

and

$$p = s\tilde{p} + (sb^H\hat{d} + cB^*_{k,k} - s\delta){A_{k,k+1:n}}^H/A^*_{k,k}$$

recalling that

$$\hat{p} \stackrel{\text{def}}{=} \begin{pmatrix} \delta \\ \tilde{p} \end{pmatrix} \quad \text{where} \quad \delta = a^H\hat{x}.$$

Denote

$$x = W_1 \begin{pmatrix} c \\ s \end{pmatrix} \quad \text{and} \quad p = W_2 \begin{pmatrix} c \\ s \end{pmatrix}$$

we have

$$W_1 = (A^*_{k,k})^{-1} \begin{pmatrix} 0 & A^*_{k,k}\hat{x} \\ B^*_{k,k} & b^H\hat{d} - \delta \end{pmatrix}$$

and

$$W_2 = (A^*_{k,k})^{-1}(B^*_{k,k}{A_{k,k+1:n}}^H \quad A^*_{k,k}\tilde{p} + (b^H\hat{d} - \delta){A_{k,k+1:n}}^H).$$

Thus we choose c and s so that $(c, s)^T$ is the normalized right singular vector associated with the larger singular value of the matrix

$$W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}.$$

3 The Deflation

Suppose v is the approximation of the singular vector computed by the estimator described in the previous section. To deflate AB, we first find a sequence of row rotations transforming v into the nth unit vector $e_n = (0, ..., 0, 1)^T$. We then apply the rotations to the columns of B and, at the same time, restore the triangular structure of B by applying a sequence of rotations to the rows of B. These row rotations are propagated to A in order to maintain the product AB. Finally, we apply another sequence of row rotations to restore the triangular structure of A. We can show that

$$AB = URV^H$$

where both U and V are orthonormal and R is nearly block diagonal:

$$R = \begin{pmatrix} \hat{R} & r \\ 0 & \sigma \end{pmatrix} \tag{3.1}$$

where $||r||^2 + |\sigma|^2$ approximates the square of the smallest singular value of AB.

If we denote Q^H as the sequence of rotations applied to the rows of B, then we can write

$$A = UR_AQ^H \quad \text{and} \quad B = QR_BV^H \tag{3.2}$$

where R_A and R_B are resultant triangular matrices and $R_A R_B = R$.

The deflation of the quotient AB^{-1} can be obtained by noticing that $B^{-1}V = (V^H B)^{-1}$. Specifically, assuming v is the estimated right singular vector associated with the smallest singular value of AB^{-1}. we first apply the sequence of rotations which transforms v into the nth unit vector to the rows of B and simultaneously restore the triangular structure of B by column rotations. Then, as in the case of the product AB, we propagate these rotations to A and restore the triangular structure of A.

Similarly, if we denote Q as the sequence of rotations applied to the columns of B (or the columns of A), then we have

$$A = U R_A Q^H \quad \text{and} \quad B = V R_B Q^H \tag{3.3}$$

where $R_A R_B^{-1} = R$.

The deflation procedure can be repeated until the decomposition (1.1) is achieved depending on a tolerance for the numerical rank.

If v is the exact singular vector associated with the smallest singular value $\sigma_{\min}$ of, say AB, then

$$|\sigma_{\min}|^2 = \|ABv\|^2 = \|r\|^2 + |\sigma|^2$$

which implies that

$$|\sigma| \leq |\sigma_{\min}|.$$

On the other hand, since R and AB have the same singular values and $\sigma = e_n^T R e_n$, it follows from the minimax theorem for singular values that

$$|\sigma| \geq |\sigma_{\min}|.$$

This shows that $|\sigma| = |\sigma_{\min}|$ and $r = 0$, when v is the exact singular vector associated with the smallest singular value. Thus an accurate estimate of the singular vector can bring R close to block diagonal. To improve the accuracy of the singular vector approximation v computed by the estimator described in the previous section, we can apply the inverse power method using the estimate v as an initial vector. Alternatively, we can bring R arbitrary close to block diagonal and σ arbitrary close to $\sigma_{\min}$ using a refinement scheme presented in the following section.

4 The Refinement

Let $C = AB$, we first find a rotation on the $(n-1,n)$-plane to eliminate $C_{n-1,n}$ using $C_{n-1,n-1}$. This rotation is determined by the five elements in A and B: $A_{n-1,n-1}$, $A_{n-1,n}$, $A_{n,n}$, $B_{n-1,n}$, and $B_{n,n}$, since $C_{n-1,n} = A_{n-1,n-1}B_{n-1,n} + A_{n-1,n}B_{n,n}$ and $C_{n-1,n-1} = A_{n-1,n-1}B_{n-1,n-1}$. Then, after applying this rotation to the nth and the $n-1$st columns of B, we annihilate the $(n-1,n)$-entry of B using a row rotation on the $(n-1,n)$-plane.

The following figure illustrates the procedure:

$$
B: \quad
\begin{array}{cccc}
 & & \downarrow & \downarrow \\
\times & \times & \times & \times \\
 & \times & \times & \times \\
 & & \times & \times \\
 & & & \times
\end{array}
\quad \longrightarrow \quad
\begin{array}{c}
\\ \\ \\ \rightarrow \\ \rightarrow
\end{array}
\begin{array}{cccc}
\\
\times & \times & \times & \times \\
 & \times & \times & \times \\
 & & \times & \otimes \\
 & & \times & \times
\end{array}
\quad \longrightarrow \quad
\begin{array}{cccc}
\\
\times & \times & \times & \times \\
 & \times & \times & \times \\
 & & \times & \\
 & & \times & \times
\end{array}
$$

The above row rotation must be propagated to A in order to maintain the product AB as shown in the following figure:

$$
A: \quad
\begin{array}{cccc}
 & & \downarrow & \downarrow \\
\times & \times & \times & \times \\
 & \times & \times & \times \\
 & & \times & \times \\
 & & & \times
\end{array}
\quad \longrightarrow \quad
\begin{array}{cccc}
\\
\times & \times & \times & \times \\
 & \times & \times & \times \\
 & & \times & \times \\
 & & \times & \times
\end{array}
$$

Now that both the resultant product AB and the resultant B have the structure:

$$
\begin{array}{cccc}
\times & \times & \times & \times \\
 & \times & \times & \times \\
 & & \times & \\
 & & \times & \times
\end{array}
\tag{4.1}
$$

it can be verified that the resultant A must have the same structure. This procedure can be repeated until A and B and their product have the following structure:

$$
\begin{array}{cccc}
\times & \times & \times & \\
 & \times & \times & \\
 & & \times & \\
\times & \times & \times & \times
\end{array}
$$

To restore the triangular structure, we start with applying a sequence of row rotations to triangularize B. These rotations are propagated to A to maintain the product AB. Finally we apply another sequence of left rotations to A to restore its triangular structure.

We can prove that if we iterate the above procedure, the off diagonal vector r in (3.1) converges to zero and σ to the smallest singular value of AB.

Similarly, in the case of the quotient AB^{-1}, let $C = AB^{-1}$, then the rotation eliminating $C_{n-1,n}$ is determined by $C_{n-1,n} = -A_{n-1,n-1}B_{n-1,n}/(B_{n-1,n-1}B_{n,n}) + A_{n-1,n}/B_{n,n}$ and $C_{n-1,n-1} = A_{n-1,n-1}/B_{n-1,n-1}$. Then we apply this rotation to the rows of B and annihilate the $(n-1,n)$-entry of B using a column rotation. Note that if B has the structure (4.1), then so does B^{-1}. The rest of the refinement procedure for the quotient AB^{-1} is similar to that for the product AB.

Note that all the rotations applied to the columns of A can be accumulated in the matrix Q in (3.2) or (3.3).

5 Updating

In updating the decomposition, we assume the form (3.2) or (3.3). In other words, the matrix Q is the product of all the rotations applied to the columns of A.

5.1 ADDING A ROW TO A

When a new row a^T is added to A, updating the decomposition can be achieved by writing

$$\begin{pmatrix} A \\ a^T \end{pmatrix} = \begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R_A \\ a^T Q \end{pmatrix} Q^H.$$

Then R_A is updated by a sequence of rotations eliminating $a^T Q$. The rotations are propagated to the left and absorbed by U.

5.2 ADDING A ROW OR COLUMN TO B

When a row is added to B in the case of the quotient AB^{-1}, the updating is symmetric to the procedure described above. In the case of the product AB, when a column b is added to B, we write

$$(B\ b) = Q(R_B\ Q^H b) \begin{pmatrix} V^H & 0 \\ 0 & 1 \end{pmatrix}.$$

Then R_B is updated by a sequence of column rotations which is absorbed by V.

After updating, we need to apply deflation and possible refinement since the numerical rank may change.

Acknowledgements

This work is supported in part by an NSERC grant OGP0046301. The author is grateful to C.T. Pan for valuable discussions.

References

[1] G.H. Golub and C.F. Van Loan. *Matrix computations.* Second Edition, Johns Hopkins Press, 1989.

[2] G.W. Stewart. Updating a rank-revealing ULV decomposition. *SIAM J. Matrix Anal. Appl.* **46**, pp 479-492, 1993.

[3] C. Van Loan. On estimating the condition of eigenvalues and eigenvectors. *Linear Algebra Appl.* **88/89**, pp 715-732, 1987.