# **GSVD Progress**

Ji Wang

# 1 Week of July 1, 2019

# 1.1 GSVD Pre-processing Routine

Given that  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times n}$ , rank(B) = l and rank([A; B]) = k + l = r, here's a detailed derivation of pre-processing procedure.

Step 1: QR decomposition with column pivoting of B, and determine the effective rank of B = l.

$$BP_1 = Q_1 \cdot \begin{bmatrix} l & n-l \\ B_{11}^{(1)} & B_{12}^{(1)} \\ 0 & 0 \end{bmatrix}$$

**Step** 2: RQ decomposition of  $\begin{bmatrix} B_{11}^{(1)} & B_{12}^{(1)} \end{bmatrix}$  when  $p \ge l$  and  $n \ne l$ .

$$BP_{1} = Q_{1} \cdot \frac{l}{p-l} \begin{bmatrix} 0 & B_{12}^{(2)} \\ 0 & 0 \end{bmatrix} \cdot Q_{2}$$

$$BP_{1}Q_{2}^{T} = Q_{1} \cdot \frac{l}{p-l} \begin{bmatrix} 0 & B_{12}^{(2)} \\ 0 & 0 \end{bmatrix}$$

$$Q_{1}^{T}BP_{1}Q_{2}^{T} = \frac{l}{p-l} \begin{bmatrix} 0 & B_{12}^{(2)} \\ 0 & 0 \end{bmatrix}$$

**Step** 3: View  $AP_1Q_2^T$  as block matrix and apply QR with column pivoting of  $A_{11}^{(1)}$  when  $A_{11}^{(1)}$  is not empty, and determine the effective rank of  $A_{11}^{(1)} = k$ .

$$AP_{1}Q_{2}^{T} = m \begin{bmatrix} n-l & l \\ A_{11}^{(1)} & A_{12}^{(1)} \end{bmatrix}$$

$$A_{11}^{(1)}P_{2} = Q_{3} \cdot \begin{bmatrix} k & n-l-k \\ A_{11}^{(2)} & A_{12}^{(2)} \\ m-k \end{bmatrix}$$

**Step** 4: RQ decomposition of  $\begin{bmatrix} A_{11}^{(2)} & A_{12}^{(2)} \end{bmatrix}$  when  $n-l \ge k$ .

$$A_{11}^{(1)}P_2 = Q_3 \cdot k \begin{bmatrix} 0 & A_{12}^{(3)} \\ 0 & 0 \end{bmatrix} \cdot Q_4$$

View  $A_{12}^{(1)}$  as block matrix, we have:

$$Q_3^T A P_1 Q_2^T P_2 Q_4^T = \begin{bmatrix} k & l & l \\ 0 & A_{12}^{(3)} & A_{13}^{(3)} \\ 0 & 0 & A_{23}^{(3)} \end{bmatrix}$$

$$Q_1^T B P_1 Q_2^T P_2 Q_4^T = \begin{bmatrix} l & l & l \\ l & l & l \\ l & l & l \end{bmatrix}$$

**Step** 5: QR decomposition of  $A_{23}^{(3)}$  when  $m \ge k$ .

Let  $U = Q_3 Q_5$ ,  $Q = P_1 Q_2^T P_2 Q_4^T$ , and  $V = Q_1$ , we have:

# 2 Week of June 21, 2019

## 2.1 Quick exit from GSVD main routine

Given that  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times n}$ , rank(B) = l and rank([A; B]) = k + l = r, if m = k, we can exit from main routine once preprocessing is done. Justification is the following:

In 3.1, after preprocessing, we obtain  $R_1$  and  $R_2$ . Let's take a close look at how we compute  $R_1$  and  $R_2$  progressively.

Initially,  $A^{(0)} = A$ ,  $B^{(0)} = B$ .

**Step** 1: URV decomposition of B, and determine l.

$$\begin{array}{c}
m \\
p \\
\end{array} \begin{bmatrix}
A^{(1)} \\
B^{(1)}
\end{bmatrix} = 
\begin{array}{c}
m \\
\iota \\
p-l
\end{array} \begin{bmatrix}
A^{(1)}_{11} & A^{(1)}_{12} \\
0 & B^{(1)}_{12} \\
0 & 0
\end{bmatrix}$$

**Step** 2: URV decomposition of  $A_{11}^{(1)}$ , if  $A_{11}^{(1)}$  is not empty (0 col).

where s is the rank of  $A_{11}^{(1)}$ .

Step 3: URV decomposition of  $A_{23}^{(2)}$ , such that  $QA_{23}^{(3)}=A_{23}^{(2)}$ .

$$\begin{bmatrix}
n & s & l \\
n & A^{(3)} \\
p & B^{(3)}
\end{bmatrix} = \begin{bmatrix}
m-s \\
l \\
p-l
\end{bmatrix} \begin{bmatrix}
0 & A_{12}^{(3)} & A_{13}^{(3)} \\
0 & 0 & A_{23}^{(3)} \\
0 & 0 & B_{12}^{(3)} \\
0 & 0 & 0
\end{bmatrix} = \begin{bmatrix}
m \\
R_1 \\
R_2
\end{bmatrix}$$

As is shown at **Step** 2, if  $A_{11}^{(1)}$  is full (row) rank, namely, its rank is m, then  $A_{23}^{(2)}$  will no longer exist. Thus, **Step** 3 is not needed anymore. Further, we have:

$${}^{m} \left[ \begin{array}{c} R_{1} \\ P_{1} \end{array} \right] = {}^{m} \left[ \begin{array}{c} 0 & A_{12}^{(2)} & A_{13}^{(2)} \\ \hline 0 & 0 & B_{13}^{(2)} \\ 0 & 0 & 0 \end{array} \right] = {}^{m+l} \left[ \begin{array}{c} 0 & R \\ 0 & 0 \end{array} \right],$$

which itself is an upper triangular matrix.

Therefore, at this step, we find the exact GSVD of A, B with rank determination,

$$\begin{array}{c} m \\ p \\ \end{array} \begin{bmatrix} A \\ B \\ \end{bmatrix} = \begin{array}{c} m \\ p \\ \end{array} \begin{bmatrix} U_1 & 0 \\ 0 & V_1 \\ \end{bmatrix} \begin{array}{c} m+l \\ p-l \\ \end{array} \begin{bmatrix} I \\ 0 \\ \end{bmatrix} \begin{array}{c} m+l \\ m+l \\ \end{array} \begin{bmatrix} n-m-l & m+l \\ 0 & R \\ \end{bmatrix} n \begin{array}{c} n \\ Q_1^T \\ \end{bmatrix}$$

## 2.2 QR: Givens rotations VS Householder's method

Test matrices A and B:

```
julia> A = [1.0 2 1 0; 2 3 1 1; 3 4 1 2; 4 5 1 3; 5 6 1 4]
5×4 Array{Float64,2}:
    1.0 2.0 1.0 0.0
    2.0 3.0 1.0 1.0
    3.0 4.0 1.0 2.0
    4.0 5.0 1.0 3.0
    5.0 6.0 1.0 4.0

julia> B = [6.0 7 1 5; 7 1 -6 13; -4 8 9 -2]
3×4 Array{Float64,2}:
    6.0 7.0 1.0 5.0
    7.0 1.0 -6.0 13.0
    -4.0 8.0 9.0 -2.0
```

Output of Givens rotations:

```
julia> test1()
```

0.000124 seconds (345 allocations: 39.219 KiB)

Output of Householder's method:

```
julia> test1()
     0.000163 seconds (159 allocations: 20.094 KiB)
```

# 3 Week of May 31, 2019

## 3.1 Quick exit from GSVD main routine

Given that  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times n}$ , rank(B) = l and rank([A; B]) = k + l = r, if  $m \le k$ , we can exit from main routine once preprocessing is done. Justification is the following:

In 3.1, after preprocessing, we obtain  $R_1$  and  $R_2$ . Let's take a close look at the structure of  $R_1$  and  $R_2$ , respectively.

$$R_{1} = \begin{bmatrix} k & l & l \\ 0 & R_{12}^{(a)} & R_{13}^{(a)} \\ 0 & 0 & R_{23}^{(a)} \\ 0 & 0 & 0 \end{bmatrix}$$

$$R_2 = \begin{bmatrix} l & 0 & 0 & R_{13}^{(b)} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As is shown, if  $m \leq k$ ,  $R_{23}^{(a)}$  will no longer exist. Further, we have: (The following is wrong)

# 3.2 Reduce space consumption in Givens rotations still working on it, stay tuned...

# 4 Week of May 6, 2019

# 4.1 Derivation of GSVD from ([A;B]) without rank revealing

$$\begin{array}{c} n \\ m \\ p \end{array} \begin{bmatrix} A \\ B \end{bmatrix} \text{ preprocessing } {m \atop = \ \ \ \, p} \begin{bmatrix} W_1 & 0 \\ 0 & V_1 \end{bmatrix} {m \atop = \ \, p} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} n \\ Q_1^T \end{bmatrix} \\ \text{split } QR {m \atop = \ \, p} \begin{bmatrix} W_1 & 0 \\ 0 & V_1 \end{bmatrix} {m \atop = \ \, p} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} n \\ R_3 \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} n \\ Q_1^T \end{bmatrix} \\ \text{Split } QR {m \atop = \ \, p} \begin{bmatrix} W_1 & 0 \\ 0 & V_1 \end{bmatrix} {m \atop = \ \, p} \begin{bmatrix} Q_1 \\ 0 \\ 0 & V_1 \end{bmatrix} {m \atop = \ \, p} \begin{bmatrix} W_2 & 0 \\ 0 & V_2 \end{bmatrix} {m \atop = \ \, p} \begin{bmatrix} C \\ S \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} n \\ Q_2^T \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} n \\ R_3 \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} n \\ U_1 U_2 & 0 \\ 0 & V_1 V_2 \end{bmatrix} {m \atop = \ \, p} \begin{bmatrix} C \\ S \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} N \\ 0 & V_1 V_2 \end{bmatrix} {m \atop = \ \, p} \begin{bmatrix} N \\ S \end{bmatrix} {n \atop = \ \, p} \begin{bmatrix} N$$

# 4.2 Golub & Van Loan's definition, algorithm of GSVD (partially adopted by MatLab)

# 5 Week of April 29, 2019

## 5.1 Ordering of cos and sin

#### 5.1.1 LAPACK 3.6.0

The generalized (or quotient) singular value decomposition of an m-by-n matrix A and a p-by-n matrix B is given by the pair of factorizations:

$$A = U\Sigma_1[0, R]Q^T$$
 and  $B = V\Sigma_2[0, R]Q^T$ 

 $\Sigma_1$  is m-by-r,  $\Sigma_2$  is p-by-r. (The integer r is the rank of  $\begin{pmatrix} A \\ B \end{pmatrix}$ , and satisfies  $r \leq n$ .) Both are real, non-

negative and diagonal, and  $\Sigma_1^T \Sigma_1 + \Sigma_2^T \Sigma_2 = I$ . Write  $\Sigma_1^T \Sigma_1 = diag(\alpha_1^2, ..., \alpha_r^2)$  and  $\Sigma_2^T \Sigma_2 = diag(\beta_1^2, ..., \beta_r^2)$ , where  $\alpha_i$  and  $\beta_i$  lie in the interval from 0 to 1. The ratios  $\alpha_1/\beta_1, ..., \alpha_r/\beta_r$  are called the generalized singular values of the pair A, B. If  $\beta_i = 0$ , then the generalized singular value  $\alpha_i/\beta_i$  is infinite. There is **no finite ordering** of  $\alpha$  and  $\beta$  in GSVD in LAPACK.

### 5.2 Features of my Julia version

#### 5.2.1 Mathematical Definition

The generalized singular value decomposition (GSVD) of an m-by-n matrix A and p-by-n matrix B is the following, where  $m + p \ge n$ :

$$A = UD_1RQ^T$$
 and  $B = VD_2RQ^T$ 

where U, V and Q are orthogonal matrices. Let k + l be the effective numerical rank of the matrix  $\begin{pmatrix} A \\ B \end{pmatrix}$ ,

then R is a (k+l)-by-n matrix of structure  $[0 R_0]$  where  $R_0$  is (k+l)-by-(k+l) and is nonsingular upper triangular matrix, D1 and D2 are m-by-(k+l) and p-by-(k+l) non-negative "diagonal" matrices and satisfy  $D_1^T D_1 + D_2^T D_2 = I$ . The nonzero elements of  $D_1$  are in **non-increasing** order while the nonzero elements of  $D_2$  are in **non-decreasing** order.

(TO BE REVEALED) Detailed structure of  $D_1$  and  $D_2$ :

#### 5.2.2 API design

- 1. U, V, Q, alpha, beta, R, k, l = gsvd(A, B, 0)
- 2. U, V, Q, D1, D2, R, k, l = gsvd(A, B, 1) or F = gsvd(A, B)

Notice that both A and B are overwritten.

#### 5.3 Issues to be resolved

# 6 Week of April 22, 2019

### 6.1 Ordering of cos and sin

#### 6.1.1 LAPACK 3.6.0

The generalized (or quotient) singular value decomposition of an m-by-n matrix A and a p-by-n matrix B is given by the pair of factorizations:

$$A = U\Sigma_1[0, R]Q^T$$
 and  $B = V\Sigma_2[0, R]Q^T$ 

 $\Sigma_1$  is m-by-r,  $\Sigma_2$  is p-by-r. (The integer r is the rank of  $\begin{pmatrix} A \\ B \end{pmatrix}$ , and satisfies  $r \leq n$ .) Both are real, non-

negative and diagonal, and  $\Sigma_1^T \Sigma_1 + \Sigma_2^T \Sigma_2 = I$ . Write  $\Sigma_1^T \Sigma_1 = diag(\alpha_1^2, ..., \alpha_r^2)$  and  $\Sigma_2^T \Sigma_2 = diag(\beta_1^2, ..., \beta_r^2)$ , where  $\alpha_i$  and  $\beta_i$  lie in the interval from 0 to 1. The ratios  $\alpha_1/\beta_1, ..., \alpha_r/\beta_r$  are called the generalized singular values of the pair A, B. If  $\beta_i = 0$ , then the generalized singular value  $\alpha_i/\beta_i$  is infinite.

#### 6.1.2 MATLAB 2019b

[U, V, X, C, S] = gsvd(A, B) returns unitary matrices U and V, a (usually) square matrix X, and non-negative diagonal matrices C and S so that

$$A = U * C * X'$$

$$B = V * S * X'$$

$$C' * C + S' * S = I$$

A and B must have the same number of columns, but may have different numbers of rows. If A is m-by-p and B is n-by-p, then U is m-by-m, V is n-by-n, X is p-by-q, C is m-by-q and S is n-by-q, where q = min(m+n,p).

The nonzero elements of S are always on its main diagonal. The nonzero elements of C are on the diagonal diag(C, max(0, q - m)). If  $m \ge q$ , this is the main diagonal of C.

sigma = gsvd(A, B) returns the vector of generalized singular values, sqrt(diag(C'\*C)./diag(S'\*S)). The vector sigma has length q and is in **non-decreasing** order, where  $\mathbf{q} = \min(\mathbf{m} + \mathbf{n}, \mathbf{p})$ . In other words, The nonzero elements of S are in **non-increasing** order while the nonzero elements of C are in **non-decreasing** order

It's interesting to notice that MATLAB has a compact form of products for gsvd.

#### 6.1.3 Julia 1.10

Compute the generalized SVD of A and B, returning a Generalized SVD factorization object F, such that  $A = F.U * F.D_1 * F.R_0 * F.Q'$  and  $B = F.V * F.D_2 * F.R_0 * F.Q'$ .

The entries of F.D1 and F.D2 are related, as explained in the LAPACK documentation for the generalized SVD and the xGGSVD3 routine which is called underneath (in LAPACK 3.6.0 and newer).

#### 6.2 Features of my Julia version

#### 6.2.1 Definition

The generalized singular value decomposition (GSVD) of an m-by-n matrix A and p-by-n matrix B is the following, where  $m + p \ge n$ :

$$A = UD_1RQ^T$$
 and  $B = VD_2RQ^T$ 

where U, V and Q are orthogonal matrices. Let k + l be the effective numerical rank of the matrix  $\begin{pmatrix} A \\ B \end{pmatrix}$ 

then R is a (k+l)-by-n matrix of structure  $[0 R_0]$  where  $R_0$  is (k+l)-by-(k+l) and is nonsingular upper triangular matrix, D1 and D2 are m-by-(k+l) and p-by-(k+l) non-negative "diagonal" matrices and satisfy  $D_1^T D_1 + D_2^T D_2 = I$ . The nonzero elements of  $D_1$  are in **non-increasing** order while the nonzero elements of  $D_2$  are in **non-decreasing** order.

#### 6.2.2 2 APIs

- 1. LAPACK-style: U, V, Q, alpha, beta, R, k, l = gsvd(A, B, 0)
- 2. Julia-style: U, V, Q, D1, D2, R, k, l = gsvd(A, B, 1)

## 6.3 Testing

#### 6.3.1 Alan testing

Still suffers some subtlety in the dimension of  $D_1$  and  $D_2$ .

#### 6.3.2 Spatial matrices testing

Tested example matrices pair from documentation, and publication.

#### 6.3.3 General testing

Tested function correctness, time performance and stability.