A Note on the Recursive Calculation of Dominant Singular Subspaces

Nicola Mastronardi
Marc Van Barel
Raf Vandebril

Report TW 393, June 2004

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

A Note on the Recursive Calculation of Dominant Singular Subspaces

Nicola Mastronardi
Marc Van Barel
Raf Vandebril

Report TW 393, June 2004

Department of Computer Science, K.U.Leuven

In [Chahlaoui, Gallivan and Van Dooren, 2004] a recursive procedure is designed for computing an approximation of the left and right dominant singular subspaces of a matrix, whose columns are produced incrementally. The method is particularly suited for matrices with many more rows than columns. The procedure consists of a few steps. In one of these steps a Householder transformation is multiplied to an upper triangular matrix. The following step consists in recomputing an upper triangular matrix from the latter product. In [Chahlaoui, Gallivan and Van Dooren, 2004] it is said that the latter step is accomplished in $O(k^3)$ operations, where k is the order of the triangular matrix. In this short note we show that this step can be accomplished in $O(k^2)$ operations. The note is organized as follows. In Section 2 the algorithm proposed in [Chahlaoui, Gallivan and Van Dooren, 2004] is described and the proposed modification is described in Section 3.

Abstract

Keywords : Householder matrix, Givens matrix, QR factorization, URV factorization, updating, singular value decomposition.
AMS(MOS) Classification : Primary : 15A15, Secondary : 15A09, 15A23.

A NOTE ON THE RECURSIVE CALCULATION OF DOMINANT SINGULAR SUBSPACES

NICOLA MASTRONARDI*, MARC VAN BAREL†, AND RAF VANDEBRIL

Abstract. A recursive procedure for computing an approximation of the left and right dominant singular subspaces of a given matrix is proposed in [1]. The method is particularly suited for matrices with many more rows than columns. The procedure consists of a few steps. In one of these steps a Householder transformation is multiplied to an upper triangular matrix. The following step consists in recomputing an upper triangular matrix from the latter product. In [1] it is said that the latter step is accomplished in $O(k^3)$ operations, where k is the order of the triangular matrix. In this short note we show that this step can be accomplished in $O(k^2)$ operations.

Key words. Householder matrix, Givens matrix, QR factorization, URV factorization, updating, singular value decomposition.

AMS subject classifications. 15A15, 15A09, 15A23

1. Introduction. In [1], a recursive procedure is designed for computing an approximation of the left and right dominant singular subspaces of a matrix, whose columns are produced incrementally. The method is particularly suited for matrices with many more rows than columns. The procedure consists of a few steps. In one of these steps a Householder transformation is multiplied to an upper triangular matrix. The following step consists in recomputing an upper triangular matrix from the latter product. In [1] it is said that the latter step is accomplished in $O(k^3)$ operations, where k is the order of the triangular matrix. In this short note we show that this step can be accomplished in $O(k^2)$ operations. The note is organized as follows. In Section 2 the algorithm proposed in [1] is described and the proposed modification is described in Section 3.

2. Recursive Calculation of Dominant Singular Subspaces. In this section we shortly describe the algorithm proposed in [1] to recursively compute dominant singular subspaces, whose columns are known incrementally, i.e., at each step one more column is added to the matrix. The algorithm is particularly suited for matrices $A \in \mathbb{R}^{m\times n}, m \gg n$.

At each recursion, the following steps are performed:

Step 1. Known the orthogonal matrix $U \in \mathbb{R}^{m\times k}$, and the upper triangular matrix $R \in \mathbb{R}^{k\times k}$ of the URV factorization of the matrix $\tilde{A} \in \mathbb{R}^{m\times k}$, $k < n \ll m$, $\tilde{A} = URV$, with $V \in \mathbb{R}^{k\times k}$ orthogonal (V does not need to be known), a new column b is added to $\tilde{A}$,

$$\hat{A} = [\tilde{A}, b].$$

Step 2. Update the URV factorization of the extended matrix $\hat{A}$ via Gram–Schmidt orthogonalization,

$$\begin{aligned} r &= U^T b \\ \hat{b} &= b - Ur \\ \rho &= \|\hat{b}\|_2 \\ \hat{u} &= \hat{b}/\rho. \end{aligned}$$

Then

$$[\tilde{A}|b] = [UR|b]\left[\begin{array}{c|c} V & \\ \hline & 1 \end{array}\right] = [U|\hat{u}]\left[\begin{array}{c|c} R & r \\ \hline & \rho \end{array}\right]\left[\begin{array}{c|c} V & \\ \hline & 1 \end{array}\right] = [U|\hat{u}]\hat{R}\left[\begin{array}{c|c} V & \\ \hline & 1 \end{array}\right].$$

* Istituto per le Applicazioni del Calcolo ”M. Picone”, Consiglio Nazionale delle Ricerche, via G. Amendola 122/D, I–70126 Bari, Italy (n.mastronardi@area.ba.cnr.it).,

†Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Leuven (Heverlee), Belgium ({Raf.Vandebril,Marc.VanBarel}@cs.kuleuven.ac.be)

Step 3. Let σ_{k+1} be the smallest singular value of $\hat{R}$ and u_{k+1} the corresponding left singular vector. Let G_u be an orthogonal matrix such that

$$G_u^T u_{k+1} = e_{k+1} = [0,\ldots,0,1]^T.$$

and compute

$$G_u^T \hat{R} = \hat{S},$$

with $\hat{S} \in \mathbb{R}^{(k+1)\times(k+1)}$ no more upper triangular.

Step 4. Compute the RQ factorization of $\hat{S}$,

$$\hat{S} = \tilde{R}\tilde{H},$$

where $\tilde{R}$ and $\tilde{H}$ are upper triangular and orthogonal matrices of order $(k+1)$, respectively. It turns out that

$$\tilde{R} = \left[\begin{array}{c|c} R_n & \\ \hline & \sigma_{k+1} \end{array}\right].$$

Step 5. Update the orthogonal matrix U

$$U \leftarrow UG_u, \ \ V \leftarrow \tilde{H}^T V,$$

and set

$$\tilde{A} = U(:,1:k)R_n V(1:k,:).$$

REMARK 1. *It is not necessary to update he matrix V, as stated in [1], if only the left dominant singular subspaces are needed.*

The matrix G_u in Step 3 could be either a product of k Givens rotations or a Householder matrix. If the product of k Givens rotations is chosen, Step 4 can be accomplished in $O(k^2)$ operations in a way similar to the one described in [2], and the matrix $\tilde{H}$ is given by the product of k different Givens rotations.

However, the costly part of the algorithm is the update of the matrix U, and hence it is preferable to choose G_u to be a Householder transformation. In fact, the cost of the latter product UG_u is $4m(k+1)$ if G_u is a Householder transformation, whereas the cost is doubled if the orthogonal matrix G_u is given by the product of k Givens rotations. In [1] is said that the price to pay choosing G_u a Householder transformation is that the RQ factorization in Step 4 is computed in $O(k^3)$ operations.

In the next section we show that the latter factorization can be accomplished in $O(k^2)$ operations.

3. Computation of the *RQ* factorization of the matrix $\hat{S}$. Let $G_u = I - \beta zz^T$ be the Householder transformation such that $G_u^T u_{k+1} = e_{k+1}$. Let us denote by $\hat{s}_i,\ i = 1,\ldots,k+1$, the ith column of $\hat{S}$. Similarly, we use the notation $\hat{r}_i$ for the ith column of the matrix $\hat{R}$.

The product $G_u^T \hat{R}$ can be written as

$$G_u^T \hat{R} = [G_u^T \hat{r}_1, G_u^T \hat{r}_2, \ldots, G_u^T \hat{r}_{k+1}].$$

We observe that

$$G_u^T \hat{r}_i = (I - \beta zz^T)\hat{r}_i = \hat{r}_i - (\beta z^T \hat{r}_i)z.$$

Since $\hat{R}$ is an upper triangular matrix, the ith column of $\hat{S}$ below the main diagonal is a multiple of z, i.e.,

$$\hat{s}_i(i+1:k+1) = (G_u^T \hat{r}_i)(i+1:k+1) = -(\beta z^T \hat{r}_i) z(i+1:k+1).$$

In other words, the strictly lower triangular part of $\hat{S}$ is the strictly lower triangular part of a rank–one matrix. Therefore, the retriangularization of $\hat{S}$ can be accomplished applying only $2k$ Givens rotations to the right, with $O(k^2)$ operations as described below.

We observe that he strictly lower triangular part is not explicitly computed, since the entries of the ith column below the main diagonal are given by $-\beta z^T \hat{r}_i z(k),\ k > i$, where $z(k)$ is the kth entry of the vector z. Therefore, for any column it is only needed to store the values $\gamma(i) = \beta z^T \hat{r}_i,\ i = 1, \ldots, k$.

Let us suppose, for simplicity, that $k = 5$, and denote with $\boxtimes$ the entries of $\hat{S}$ belonging to the strictly lower triangular part:

$$\hat{S} = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \end{bmatrix}.$$

Applying a Givens rotation G_1 to the right in order to annihilate the entry $(3,1)$, the whole column below this entry is annihilated, since both the columns are a multiple of the vector z below the main diagonal.

$$\hat{S}G_1 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \boxtimes & \times & \times & \times & \times & \times \\ \otimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \end{bmatrix} G_1 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ & \boxtimes & \times & \times & \times & \times \\ & \boxtimes & \boxtimes & \times & \times & \times \\ & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \end{bmatrix}.$$

Let c_1 and s_1 be the Givens coefficients of G_1, i.e.,

$$G_1 = \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix},\ c_1 = \frac{z^T \hat{r}_2}{\sqrt{(z^T \hat{r}_1)^2 + (z^T \hat{r}_2)^2}},\ s_1 = \frac{z^T \hat{r}_1}{\sqrt{(z^T \hat{r}_1)^2 + (z^T \hat{r}_2)^2}}.$$

Of course, it is not necessary to compute $\hat{S}G_1$ below the upper Hessenberg part of the matrix. In fact the entries of the first column with row indexes greater than 2 are zero. Moreover, the entries of the second column with row indexes greater than 2 are given by

$$\beta z(i) \sqrt{(z^T \hat{r}_1)^2 + (z^T \hat{r}_2)^2},\ i > 2.$$

Therefore it is only needed to update the value $\gamma(2) = \beta \sqrt{(z^T \hat{r}_1)^2 + (z^T \hat{r}_2)^2}$.

In the same manner, applying a Givens rotation G_2 to the right in order to annihilate the entry $(4,2)$, the whole row below this entry is annihilated.

$$\hat{S}G_1G_2 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ & \boxtimes & \times & \times & \times & \times \\ & \otimes & \boxtimes & \times & \times & \times \\ & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \end{bmatrix} G_2 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times \\ & & \boxtimes & \times & \times & \times \\ & & \boxtimes & \boxtimes & \times & \times \\ & & \boxtimes & \boxtimes & \boxtimes & \times \end{bmatrix}.$$

Also in this case, only the upper Hessenberg part of columns 2 and 3 and the coeffcient $\gamma(3)$ need to be updated.

Proceeding in this way, after $k-1$ multiplications to the right by Givens rotations, the original matrix is transformed into an upper Hessenberg one with $O(k^2)$ operations,

$$\hat{S}G_1G_2\cdots G_{k-1} = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times \\ & & \times & \times & \times & \times \\ & & & \times & \times & \times \\ & & & & \times & \times \end{bmatrix}.$$

To reduce the latter Hessenberg matrix to upper triangular form, apply k Givens rotations to the right in order to annihilate the entries $(k+1,k)$, $(k,k-1)$, $\ldots$, $(2,1)$. The whole transformation is depicted in Fig. 3.1. Therefore the reduction of the matrix $\hat{S}$ into triangular form can be made in $O(k^2)$ operations.

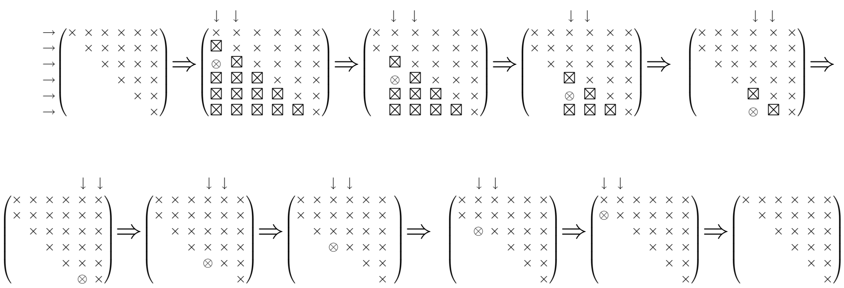

FIGURE 3.1. *Multiplication of $\hat{R}$ to the left by a Householder transformation followed by a sequence of $2k-1$ multiplications by Givens rotations to the right in order to retrieve the upper triangular structure. With $\otimes$ is denoted the entry of the matrix to be annihilated by a Givens transformation.*

REFERENCES

[1] Y. CHAHLAOUI, K. GALLIVAN, AND P. VAN DOOREN, *Recursive calculation of dominant singular subspaces,* SIAM J. Matrix Anal. Appl., 25:2 (2004), pp. 445–463.

[2] T.F. CHAN, *An Improved Algorithm for Computing the Singular Value Decomposition,* ACM Trans. Math. Software (TOMS), 8:1 (1982), pp. 72–83.