Block Householder Transformation for Parallel QR Factorization

F. ROTELLA
Laboratoire de Génie de Production
Ecole Nationale d'Ingénieurs de Tarbes
47 avenue d'Azereix, 65016 Tarbes Cedex, France

I. ZAMBETTAKIS
Institut Universitaire de Technologie Paul Sabatier
1 rue Lautréamont, 65016 Tarbes Cedex, France
and
Laboratoire d'Analyse et d'Architecture des Systèmes, CNRS
7 avenue du Colonel Roche, 31077 Toulouse Cedex, France

Abstract—A new form of the QR factorization procedure is presented which is based on a generalization of the Householder transformation. This extension is a block matrical form of the usual Householder procedure which leads to a dichotomic algorithm which allows parallel implementation. © 1999 Elsevier Science Ltd. All rights reserved.

1. INTRODUCTION

The well-known QR factorization of a matrix, namely writing a $(m \times n)$ matrix A on the form

$$A = QR,$$

where Q is an orthogonal matrix, $Q^{-1} = Q^{\mathsf{T}}$, and R is an upper triangular matrix, is obtained with the use of a sequence of Householder transformations [1,2].

This type of transformation is based on the following orthogonal and symmetric matrix [3]

$$H(v) = I - 2\frac{vv^{\mathsf{T}}}{v^{\mathsf{T}}v},$$

where v is a nonnull vector and I is the identity matrix of order size (v) and on the following result [1].

THEOREM 1.1. If $e_1 = [1 \ 0 \cdots 0]^{\mathsf{T}}$, then for every $x \neq 0$, if we take $v_x = x + \sigma e_1$, with $\sigma = \text{sign}(x^{\mathsf{T}} e_1)\sqrt{x^{\mathsf{T}} x}$, then

$$H(v_x)x = \sigma e_1.$$

REMARK 1. In this theorem, the sign function is introduced for a numerical robustness property and is defined by

$$\text{sign}(z) = \begin{cases} -1, & \text{if } z < 0, \\ +1, & \text{if } z \geq 0. \end{cases}$$

REMARK 2. For $x = 0$, this result is always true by considering $H(v_x) = I$.

This property is used to construct the QR form of a matrix with n columns by applying $n - 1$ successive Householder transformations. It is well known that this so-called reduction of matrix is an efficient and stable tool for numerical calculations and analysis [4,5]. It is used for instance to obtain the Schur (complex or real) form of a matrix which is a step for eigenvalues numerical determination.

In this paper, we will propose a matrix form of the Householder transformation which will allow to reduce in one step the last components of several vectors. This will lead to a parallelized version of the QR algorithm where only $\log_2(n)$ steps will be necessary. Consequently, the paper is organized as follows: in the following section is described our extension of the Householder transformation and the application to the reduction of a matrix, and in a final part is presented the application to the QR factorization which allows a parallelization of the implied calculations. A simple example illustrates the presented ideas.

2. EXTENSION OF THE HOUSEHOLDER TRANSFORMATION

Let us consider a full column rank matrix V and if we introduce the matrix defined by

$$H(V) = I_n - 2V \left(V^\top V\right)^{-1} V^\top, \tag{2.1}$$

which appears as a matrix extension of the usual Householder transform, we have the following result.

THEOREM 2.1. For every $(n \times r)$ matrix V, such that $\operatorname{rank}(V) = r$, then $H(V)$ is symmetric and orthogonal.

PROOF. As $\operatorname{rank}(V) = r$, $V^\top V$ is nonsingular and $H(V)$ is well defined. We can notice here that $(V^\top V)^{-1}V^\top = V^\dagger$, the Moore-Penrose pseudo-inverse of the matrix V [6,7], thus we can write $H(V) = I_n - 2VV^\dagger$. Consequently, as by definition of the pseudo-inverse, we have $VV^\dagger = (VV^\dagger)^\top$ and $VV^\dagger V = V$, it is a trivial trick to verify that $[H(V)]^\top = H(V)$ and $H(V)H(V) = I_n$. ∎

After this preliminary step, let us state now the main result of our paper.

THEOREM 2.2. For any full column rank $(m \times r)$ matrix A

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix},$$

where A_1 is a $(r \times r)$ nonsingular matrix, if we choose

$$V_A = \begin{pmatrix} A_1 + X \\ A_2 \end{pmatrix},$$

where X is given by

$$X = P^\top \sqrt{D} P A_1,$$

and $\sqrt{D} = \operatorname{diag}_{i=1}^r \{\sqrt{d_i}\}$, where the nonnegative scalar d_i and the orthogonal matrix P, are defined by

$$I_r + \left(A_2 A_1^{-1}\right)^\top \left(A_2 A_1^{-1}\right) = P^\top \operatorname{diag}_{i=1}^r \{d_i\} P,$$

then

$$H(V_A)A = \begin{pmatrix} -X \\ O_{((m-r) \times r)} \end{pmatrix},$$

where I_r is the $(r \times r)$ identity matrix and $O_{((m-r) \times r)}$ is the $((m - r) \times r)$ null matrix.

PROOF. Let us consider a $(m \times r)$ matrix A, such that $\operatorname{rank}(A) = r$, written as

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix},$$

where A_1 is a $(r \times r)$ nonsingular matrix. Then if we choose

$$V = \begin{pmatrix} A_1 + X \\ A_2 \end{pmatrix},$$

we can look for X such that we have

$$H(V)A = \begin{pmatrix} E_1 \\ O_{((m-r) \times r)} \end{pmatrix},$$

where $O_{((m-r) \times r)}$ is the $((m-r) \times r)$ null matrix.

To solve this problem, following (2.1), we have:

$$H(V)A = A - 2V \left(V^\top V\right)^{-1} V^\top A,$$

with

$$V^\top V = A^\top A + X^\top A_1 + A_1^\top X + X^\top X,$$
$$V^\top A = A^\top A + X^\top A_1.$$

Then

$$H(V)A = \begin{pmatrix} E_1 \\ E_2 \end{pmatrix},$$

where E_1 is a $(r \times r)$ matrix, E_2 is a $((m-r) \times r)$ one, given by

$$E_1 = A_1 \left(V^\top V\right)^{-1} \left[V^\top V - 2V^\top A\right] - 2X \left(V^\top V\right)^{-1} V^\top A,$$
$$E_2 = A_2 \left(V^\top V\right)^{-1} \left[V^\top V - 2V^\top A\right].$$

As we want $E_2 = 0$, this can be obtained by letting

$$V^\top V = 2V^\top A,$$

or equivalently

$$A^\top A = X^\top X + A_1^\top X - X^\top A_1. \tag{2.2}$$

This last relationship can be written

$$(X + A_1)^\top (X - A_1) = A_2^\top A_2,$$

which indicates that $(X + A_1)^\top (X - A_1)$ is a symmetric matrix, thus we must have

$$A_1^\top X = X^\top A_1.$$

From (2.2), this remark leads to the following equation which must be verified by X

$$X^\top X = A^\top A,$$

where X is such that $A_1^\top X$ is symmetric.

In order to solve this constrained equation, let us introduce the matrix $Z = X A_1^{-1}$ which must also be a symmetric one. As A_1 is nonsingular, we can write the previous identity as

$$Z^2 = I_r + \Lambda^\top \Lambda,$$

where $\Lambda = A_2 A_1^{-1}$, and with the remark that $I_r + \Lambda^\top \Lambda$ is positive definite, we can obtain its square root by considering its diagonalization through an orthogonal transform [7,8]

$$I_r + \Lambda^\top \Lambda = P^\top D P,$$

where P is orthogonal, $P^\top P = I_r$ and $D = \operatorname{diag}_{i=1}^r \{d_i\}$, $d_i > 0$. It comes out

$$Z = P^\top \sqrt{D} P,$$

where $\sqrt{D} = \operatorname{diag}_{i=1}^r \{\sqrt{d_i}\}$. Following the meaning of Z, we obtain, in the one hand, the solution

$$X = P^\top \sqrt{D} P A_1,$$

which leads to $E_2 = 0$, and in the other hand, $E_1 = -X$.   ∎

More generally, we can extend the previous result to matrices which are not full column rank by the following.

THEOREM 2.3. *For any* $(m \times n)$ *matrix* A, *such that* $\mathrm{rank}\,(A) = r$, *there exists a matrix House-holder transformation* H_A *(2.1), such that*

$$H_A A = \begin{pmatrix} \tilde{A} \\ O_{((m-r) \times n)} \end{pmatrix},$$

where $\tilde{A}$ *is* $(r \times n)$.

PROOF. When $\mathrm{rank}\,(A) = r$, we can consider a maximal rank factorization of A, namely [6]

$$A = FG,$$

where F is $(m \times r)$ and G is $(r \times n)$. Then by the application of the previous result, we obtain

$$H(V_F)F = \begin{pmatrix} -X_F \\ O_{((m-r) \times r)} \end{pmatrix},$$

where X_F is a $(r \times r)$ matrix. Then we have

$$H(V_F)A = \begin{pmatrix} -X_F G \\ O_{((m-r) \times n)} \end{pmatrix},$$

which states the result.

If $A = 0$, then we can choose obviously $H_A = I$.   ∎

3. APPLICATION TO A BLOCK QR FACTORIZATION

Let us consider now a $(m \times n)$ partitioned matrix

$$M = \begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \end{pmatrix},$$

where A_1 is a $(r \times r)$ nonsingular matrix and the matrices A_1, A_2, B_1, and B_2 are well dimensioned. Then by choosing $H_1 = H(V_A)$, as in the previous theorem, with respect to $A = (\, A_1^\mathsf{T} \ A_2^\mathsf{T} \,)^\mathsf{T}$, we obtain

$$H_1 M = \begin{pmatrix} -X & B_1^* \\ O_{((m-r) \times r)} & B_2^* \end{pmatrix},$$

which is an upper block triangular matrix. This matrix can be reduced by a block diagonal Householder transformation

$$H_2 = \begin{pmatrix} H_2^1 & O_{(r \times (n-r))} \\ O_{((m-r) \times r)} & H_2^2 \end{pmatrix},$$
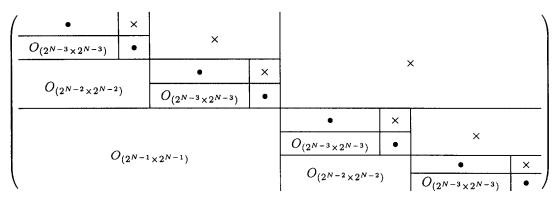
where H_2^1, and H_2^2 are extended Householder transformation which act, respectively, on X and B_2^*.

Then the QR factorization of a $(2^N \times 2^N)$ matrix M can be obtained, in N steps with the following scheme

$$H_1 M = \begin{pmatrix} \bullet & \times \\ O_{(2^{N-1} \times 2^{N-1})} & \bullet \end{pmatrix},$$

$$H_2 H_1 M = \begin{pmatrix} \bullet & \times & \times & \\ O_{(2^{N-2} \times 2^{N-2})} & \bullet & \bullet & \times \\ O_{(2^{N-1} \times 2^{N-1})} & & O_{(2^{N-2} \times 2^{N-2})} & \bullet \end{pmatrix},$$

$$H_3 H_2 H_1 M =$$

$$
\begin{pmatrix}
\bullet & \times \\
O_{(2^{N-3} \times 2^{N-3})} & \bullet & & \times & & \times \\
& & \bullet & \times \\
O_{(2^{N-2} \times 2^{N-2})} & O_{(2^{N-3} \times 2^{N-3})} & \bullet \\
& & & & \bullet & \times \\
O_{(2^{N-1} \times 2^{N-1})} & & O_{(2^{N-3} \times 2^{N-3})} & \bullet & & \times \\
& & O_{(2^{N-2} \times 2^{N-2})} & & \bullet & \times \\
& & & O_{(2^{N-3} \times 2^{N-3})} & \bullet
\end{pmatrix}
$$

For instance, if we consider the matrix

$$
M = \begin{pmatrix}
1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 \\
1 & -1 & -1 & -1
\end{pmatrix},
$$

its usual QR factorization, as indicated in the introduction can be performed in three steps and leads to

$$
R = \begin{pmatrix}
-2 & 0 & 0 & 1 \\
0 & 2 & 0 & 1 \\
0 & 0 & 2 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}, \qquad
Q = \begin{pmatrix}
-.5 & .5 & .5 & .5 \\
-.5 & -.5 & .5 & -.5 \\
-.5 & .5 & -.5 & -.5 \\
-.5 & -.5 & -.5 & .5
\end{pmatrix}.
$$

If we apply the previous scheme, we obtain

$$
H_1 = \begin{pmatrix}
-.7071 & 0 & -.7071 & 0 \\
0 & -.7071 & 0 & -.7071 \\
-.7071 & 0 & .7071 & 0 \\
0 & -.7071 & 0 & -.7071
\end{pmatrix},
$$

$$
H_2 = \begin{pmatrix}
-.7071 & -.7071 & 0 & 0 \\
-.7071 & .7071 & 0 & 0 \\
0 & 0 & -.7071 & -.7071 \\
0 & 0 & -.7071 & .7071
\end{pmatrix},
$$

which leads, in two steps to the QR factorization

$$
R = \begin{pmatrix}
2 & 0 & 0 & -1 \\
0 & 2 & 0 & 1 \\
0 & 0 & 2 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}, \qquad
Q = \begin{pmatrix}
.5 & .5 & .5 & .5 \\
.5 & -.5 & .5 & -.5 \\
.5 & .5 & -.5 & -.5 \\
.5 & -.5 & -.5 & .5
\end{pmatrix}.
$$

4. CONCLUSION

We have proposed, in this paper an extension of the Householder transformation which can be applied on a full column rank matrix. The application of this result leads to parallelize the QR factorization of a matrix. Indeed the parallelization is ensured by the the block organization of the algorithm which allows to calculate the necessary matrices at a given step, for example H_2^1 and H_2^2, in two completely independent processors. This can reduce the computational time to obtain this factorization and more generally in every numerical calculation where it is employed, as for instance the QR algorithm of Francis [1] or in other applications [9].

REFERENCES

1. G.H. Golub and C.F. Van Loan, *Matrix Computations*, North Oxford Academic, (1986).
2. T. Kailath, Matrix triangularization: Householder transformation, In *Concise Encyclopedia of Modelling and Simulation*, (Edited by D.K. Atherton and P. Borne), pp. 256–257, Pergamon Press, (1992).
3. A.S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, (1965).
4. G.W. Stewart, *Introduction to Matrix Computations*, Academic Press, (1973).
5. G. Strang, *Linear Algebra and Its Applications*, Academic Press, (1980).
6. A. Ben-Israel and T.N.E. Greville, *Generalized Inverses: Theory and Applications*, John Wiley & Sons, (1974).
7. F. Rotella and P. Borne, *Théorie et pratique du calcul matriciel*, Technip, (1995).
8. P. Lancaster, *Theory of Matrices*, Academic Press, (1969).
9. T. Kailath, State-space modelling: Square root algorithms, In *Concise Encyclopedia of Modelling and Simulation*, (Edited by D.K. Atherton and P. Borne), pp. 455–61, Pergamon Press, (1992).