Block Householder transformation for parallel QR factorisation

F. Rotella(*), I. Zambettakis(**,***)
(*) Laboratoire de Génie de Production
Ecole Nationale d'Ingénieurs de Tarbes
47 avenue d'Azereix, 65016 Tarbes CEDEX, France.
(**) Institut Universitaire de Technologie Paul Sabatier
1 rue Lautréamont, 65016 Tarbes CEDEX, France.
(***) Laboratoire d'Analyse et d'Architecture des Systèmes – CNRS
7 avenue du Colonel Roche, 31077 Toulouse CEDEX, France.

February 22, 2011

Abstract

A new form of the QR factorisation procedure is presented which is based on a generalisation 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.

1. Introduction

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

$$A = QR,$$

where Q is an orthogonal matrix, $Q^{-1} = Q^T$, and R is an upper triangular matrix, is obtained with the use of a sequence of Householder transformations [2, 4].

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

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

where v is a non null vector and I is the identity matrix of order size(v) and on the following result [2]:

Theorem 1.1. *If* $e_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$*, then for every* $x \neq 0$*, if we take* $v_x = x + \sigma e_1$*, with* $\sigma = sign(x^T e_1)\sqrt{x^T x}$*, 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 :*

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

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 [8, 9]. 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 matric 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 factorisation 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 :

$$H(V) = I_n - 2V(V^TV)^{-1}V^T, \tag{2.1}$$

which appears as a matric extension of the usual Householder transform, we have the following result :

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

Proof : As rank$(V) = r$, V^TV is non singular and $H(V)$ is well defined. We can notice here that $(V^TV)^{-1}V^T = V^\dagger$, the Moore-Penrose pseudo-inverse of the matrix V [1, 7], thus we can write $H(V) = I_n - 2VV^\dagger$. Consequently, as by definition of the pseudo-inverse, we have $VV^\dagger = (VV^\dagger)^T$ and $VV^\dagger V = V$, it is a trivial trick to verify that $[H(V)]^T = H(V)$ 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 $(r \times r)$ non singular matrix, if we choose :

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

where X is given by :

$$X = P^T\sqrt{D}PA_1,$$

and $\sqrt{D} = \mathrm{diag}_{i=1}^r\left\{\sqrt{d_i}\right\}$, where the non negative scalar d_i and the orthogonal matrix P, are defined by :

$$I_r + (A_2A_1^{-1})^T(A_2A_1^{-1}) = P^T\mathrm{diag}_{i=1}^r\{d_i\}\,P,$$

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 rank$(A) = r$, written as :

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

where A_1 is a $(r \times r)$ non singular 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 :

$$H(V)A = A - 2V(V^TV)^{-1}V^TA,$$

with :

$$\begin{array}{l} V^TV = A^TA + X^TA_1 + A_1^TX + X^TX, \\ V^TA = A^TA + X^TA_1. \end{array}$$

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 :

$$\begin{array}{rcl} E_1 & = & A_1(V^TV)^{-1}[V^TV - 2V^TA] - 2X(V^TV)^{-1}V^TA, \\ E_2 & = & A_2(V^TV)^{-1}[V^TV - 2V^TA]. \end{array}$$

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

$$V^TV = 2V^TA,$$

or equivalently :

$$A^TA = X^TX + A_1^TX - X^TA_1. \tag{2.2}$$

This last relationship can be written :

$$(X + A_1)^T(X - A_1) = A_2^TA_2,$$

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

$$A_1^TX = X^TA_1.$$

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

$$X^T X = A^T A,$$

where X is such that $A_1^T X$ is symmetric.

In order to solve this constrained equation, let us introduce the matrix $Z = XA_1^{-1}$ which must also be a symmetric one. As A_1 is non singular, we can write the previous identity as :

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

where $\Lambda = A_2 A_1^{-1}$, and with the remark that $I_r + \Lambda^T \Lambda$ is positive definite, we can obtain its square root by considering its diagonalisation through an orthogonal transform [7, 6] :

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

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

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

where $\sqrt{D} = \mathrm{diag}_{i=1}^r \left\{\sqrt{d_i}\right\}$. Following the meanning of Z, we obtain, in the one hand, the solution :

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

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 rank$(A) = r$, there exists a matric Householder 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 rank$(A) = r$, we can consider a maximal rank factorisation of A, namely [1] :

$$A = FG,$$

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 factorisation

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

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

where A_1 is a $(r \times r)$ non singular matrix and the matrices A_1, A_2, B_1 and B_2 are well dimensionned. Then by choosing $H_1 = H(V_A)$, as in the previous theorem, with respect to $A = \begin{pmatrix} A_1^T & A_2^T \end{pmatrix}^T$, 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 factorisation of a $(2^N \times 2^N)$ matrix M can be obtained, in N steps with the following scheme :

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

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

$$H_3H_2H_1M =$$

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

$$\vdots$$

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 factorisation, as described in the introduction can be performed in 3 steps and leads to :

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

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 2 steps to the QR factorisation :

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

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 organisation 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 independant processors. This can reduce the computational time to obtain this factorisation and more generally in every numerical calculation where it is employed, as for instance the QR algorithm of Francis[2] or in other applications [5].

References

[1] Ben-Israel, A., Greville, T.N.E., *Generalized inverses : theory and applications*, John Wiley & Sons, 1974.

[2] Golub, G.H., Van Loan, C.F., *Matrix computations*, North Oxford Academic, 1986.

[3] Householder, A.S., *The theory of matrices in numerical analysis,* Blaisdell Publishing Co., 1965.

[4] Kailath, T., "Matrix triangularization : Householder transformation", *Concise Encyclopedia of Modelling and Simulation*, Eds Atherton, D.K., Borne, P., pp. 256–7, Pergamon Press, 1992.

[5] Kailath, T., "State-space modelling : square root algorithms", *Concise Encyclopedia of Modelling and Simulation*, Eds Atherton, D.K., Borne, P., pp. 455–61, Pergamon Press, 1992.

[6] Lancaster, P., *Theory of matrices*, Academic Press, 1969.

[7] Rotella, F., Borne, P., *Théorie et pratique du calcul matriciel*, Technip, 1995.

[8] Stewart, G.W., *Introduction to matrix computations*, Academic Press, 1973.

[9] Strang, G., *Linear algebra and its applications*, Academic Press, 1980.