Numer. Math. 49, 81-94 (1986)

Numerische Mathematik
© Springer-Verlag 1986

QR Factorization of Toeplitz Matrices

A.W. Bojanczyk, R.P. Brent and F.R. de Hoog

Computer Sciences Laboratory and Centre for Mathematical Analysis, Australian National University, Canberra, A.C.T. 2601, Australia, and Division of Mathematics and Statistics, GPO Box 1965, Canberra, A.C.T. 2601, Australia

Summary. This paper presents a new algorithm for computing the QR factorization of an $m \times n$ Toeplitz matrix in $O(mn)$ operations. The algorithm exploits the procedure for the rank-1 modification and the fact that both principal $(m-1) \times (n-1)$ submatrices of the Toeplitz matrix are identical. An efficient parallel implementation of the algorithm is possible.

Subject Classifications: AMS(MOS): 65F05; CR: G1.3.

1. Introduction

A matrix $\mathbf{T}$ is Toeplitz if the elements on each diagonal are constant. Toeplitz systems of linear equations and Toeplitz linear least squares problems arise from many sources, see [2], and it is important to be able to solve such systems in as few operations as possible. As a square Toeplitz matrix is determined by at most $2n-1$ numbers, it is not surprising that there are algorithms which solve Toeplitz systems of linear equations using only $O(n^2)$ operations or even $O(n \log^2 n)$ when fast techniques are applied.

Most algorithms for solving Toeplitz systems require that all principal submatrices of $\mathbf{T}$ are well conditioned. For example, well conditioned positive definite matrices possess this property. If some principal submatrix is singular, the algorithms may fail. This difficulty could be circumvented if the orthogonal decomposition of the matrix $\mathbf{T}$ were available. Recently, Sweet [6] has proposed an $O(n^2)$ algorithm which computes the QR decomposition of a square Toeplitz matrix. His algorithm essentially calculates the Cholesky factors of $\mathbf{T}^T\mathbf{T}$ without explicitly forming $\mathbf{T}^T\mathbf{T}$ and requires $10n^2+O(n)$ multiplications to compute $\mathbf{R}$ and $25n^2+O(n)$ multiplications to compute both $\mathbf{Q}$ and $\mathbf{R}$.

In this paper we present a simpler and slightly faster algorithm for the QR factorization of Toeplitz matrix. Our algorithm exploits principles similar to those used in Sweet's algorithm but is applicable to rectangular matrices and requires only $mn+6n^2+O(n)$ multiplications to compute $\mathbf{R}$ and $13mn+6n^2$

$+O(n)$ multiplications to compute both $\mathbf{Q}$ and $\mathbf{R}$. The matrix $\mathbf{R}$ is generated row by row and the matrix $\mathbf{Q}$ column by column, starting from the first row and column respectively. Each row of $\mathbf{R}$ (column of $\mathbf{Q}$) is calculated from the previous row (column) after three implicit modifications of rank-1 to the matrix $\mathbf{R}$, one updating and two downdatings. The procedure for rank-1 updating is as in Gill, Golub, Murray and Saunders [3], while that for rank-1 downdating can be regarded as the reverse of rank-1 updating. Both updating and downdating operate on rows of $\mathbf{R}$ (columns of $\mathbf{Q}$) from left to right (from top to bottom) which makes efficient parallel implementation possible. This is explained in detail in [1].

Section 2 contains a description of rank-1 updating and rank-1 downdating. In Sect. 3 an algorithm for the QR decomposition of a Toeplitz matrix is given. Some concluding remarks are made in Sect. 4.

2. Methods for Rank-1 Modification

Our method for the QR decomposition of a Toeplitz matrix uses procedures for the rank-1 modification of a Cholesky factorization. In this section we describe algorithms for rank-1 updating and rank-1 downdating. The derivation of the first algorithm is taken from Gill et al. [3] while the derivation of the latter follows the one given in Lawson and Hanson [4]. We start with the description of the updating procedure.

Let $\mathbf{R}$ be an $n \times n$ nonsingular upper triangular matrix with positive elements on the diagonal. For a given n-vector $\mathbf{x}$ consider the matrix $\mathbf{X}$,

$$\mathbf{X} = \mathbf{R}^T \mathbf{R} + \mathbf{x}\mathbf{x}^T. \tag{2.1}$$

As $\mathbf{X}$ is positive definite the matrix $\mathbf{X}$ can be uniquely factorized in the form

$$\mathbf{X} = \tilde{\mathbf{R}}^T \tilde{\mathbf{R}} \tag{2.2}$$

where $\tilde{\mathbf{R}}$ is upper triangular with positive diagonal elements. We want to find the matrix $\tilde{\mathbf{R}}$. This problem is usually described as that of updating the Cholesky factors following a rank-1 modification. Clearly, we should be able to calculate $\tilde{\mathbf{R}}$ from $\mathbf{R}$ and $\mathbf{x}$ using fewer operations than the Cholesky factorization of $\mathbf{X}$ would require. This is indeed true and the matrix $\tilde{\mathbf{R}}$ can be determined by the following procedure.

Consider the augmented matrix $\mathbf{Y}$ of dimension $(n+1) \times n$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{x}^T \\ \mathbf{R} \end{bmatrix} \tag{2.3}$$

and premultiply $\mathbf{Y}$ by an orthogonal $(n+1) \times (n+1)$ matrix $\mathbf{U}$ which transforms $\mathbf{Y}$ into upper triangular form $[\hat{\mathbf{R}}^T, \mathbf{O}]^T$ with positive diagonal elements. The matrix $\hat{\mathbf{R}}$ is the desired matrix $\tilde{\mathbf{R}}$. To see this consider the product $\mathbf{Y}^T\mathbf{Y}$ and the following calculation

$$\mathbf{Y}^T \mathbf{Y} = \mathbf{Y}^T \mathbf{U}^T \mathbf{U} \mathbf{Y} = [\hat{\mathbf{R}}^T, \mathbf{O}][\hat{\mathbf{R}}^T, \mathbf{O}]^T = \hat{\mathbf{R}}^T \hat{\mathbf{R}}. \tag{2.4}$$

On the other hand we have

$$\mathbf{Y}^T\mathbf{Y} = \mathbf{R}^T\mathbf{R} + \mathbf{x}\mathbf{x}^T = \mathbf{X} = \tilde{\mathbf{R}}^T\tilde{\mathbf{R}}. \tag{2.5}$$

From the uniqueness of the Cholesky factorization, the matrices $\tilde{\mathbf{R}}$ and $\hat{\mathbf{R}}$ are identical.

The matrix $\mathbf{U}$ can be chosen as a product of n plane rotations $\mathbf{U}_{k,k+1}$, $k=1,\dots,n$, where $\mathbf{U}_{k,k+1}$ is a rotation in the $(k,k+1)$ plane such that

$$\mathbf{U}_{n,n+1}\mathbf{U}_{n-1,n}\cdots\mathbf{U}_{1,2}\mathbf{Y} = \begin{bmatrix}\hat{\mathbf{R}}\\ \mathbf{O}^T\end{bmatrix}.$$

The matrix $\mathbf{U}_{k,k+1}$ has the following form

$$\mathbf{U}_{k,k+1} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c_k & s_k & \\ & & -s_k & c_k & \\ & & & & \ddots & \\ & & & & & 1\end{pmatrix} \leftarrow k\text{th row} \tag{2.6}$$

where scalars c_k and s_k are chosen so that the $(k+1,k)$ element in $\mathbf{U}_{k-1,k}\cdots\mathbf{U}_{1,2}\mathbf{Y}$ is reduced to zero. Note that the matrix $\mathbf{U}_{k,k+1}$ operates only on the kth and the $(k+1)$st row of the product $\mathbf{U}_{k-1,k}\cdots\mathbf{U}_{1,2}\mathbf{Y}$, so the $(k+1)$st row of the product $\mathbf{U}_{k-1,k}\cdots\mathbf{U}_{1,2}\mathbf{Y}$ is the same as the kth row of the matrix $\mathbf{R}$. If we denote the kth row of $\mathbf{U}_{k-1,k}\cdots\mathbf{U}_{1,2}\mathbf{Y}$ by $[0,\dots,0,\hat{a}_{k,k},\dots,\hat{a}_{k,n}]$ and the kth row of $\mathbf{R}$ by $[0,\dots,0,r_{k,k},\dots,r_{k,n}]$, then for rows k and $k+1$ we have

$$\begin{pmatrix} c_k & s_k\\ -s_k & c_k\end{pmatrix}\begin{pmatrix}0,\dots,0, & \hat{a}_{k,k},\dots,\hat{a}_{k,n}\\ 0,\dots,0, & r_{k,k},\dots,r_{k,n}\end{pmatrix} = \begin{pmatrix}0,\dots,0,\tilde{r}_{k,k},\tilde{r}_{k,k+1},\dots,\tilde{r}_{k,n}\\ 0,\dots,0,\ 0,\ \hat{a}_{k+1,k+1},\dots,\hat{a}_{k+1,n}\end{pmatrix}, \tag{2.7}$$

where

$$\begin{aligned}\tilde{r}_{k,k} &= ((\hat{a}_{k,k})^2 + (r_{k,k})^2)^{\frac{1}{2}},\\ c_k &= \hat{a}_{k,k}/\tilde{r}_{k,k},\\ s_k &= r_{k,k}/\tilde{r}_{k,k},\end{aligned} \tag{2.8}$$

and for $j=k+1,\dots,n$

$$\begin{aligned}\tilde{r}_{k,j} &= c_k\hat{a}_{k,j} + s_k r_{k,j},\\ \hat{a}_{k+1,j} &= -s_k\hat{a}_{k,j} + c_k r_{k,j}.\end{aligned}$$

Here $[0,\dots,0,\tilde{r}_{k,k},\tilde{r}_{k,k+1},\dots,\tilde{r}_{k,n}]$ denotes the kth row of the desired matrix $\tilde{\mathbf{R}}$.

The above calculations are repeated for $k=1,\dots,n$. Thus the overall updating procedure requires $2n^2+O(n)$ multiplications.

Now we proceed to the description of the downdating procedure. Two approaches will be presented, both leading to the same algorithm.

As before, let $\mathbf{R}$ be an $n\times n$ nonsingular upper triangular matrix with positive diagonal elements. For a given n-vector $\mathbf{x}$ we want to find the Cholesky factorization of the matrix

$$\mathbf{X} = \mathbf{R}^T \mathbf{R} - \mathbf{x}\,\mathbf{x}^T, \tag{2.9}$$

where it is assumed that $\mathbf{x}$ is such that $\mathbf{X}$ is positive definite. The Cholesky factor $\tilde{\mathbf{R}}$ is unique if we impose the condition that $\tilde{\mathbf{R}}$ has positive diagonal elements. We shall refer to the problem of finding the Cholesky factor $\tilde{\mathbf{R}}$ for the matrix $\mathbf{X}$ as the downdating problem.

Once again it is possible to calculate $\tilde{\mathbf{R}}$ directly from $\mathbf{R}$ and $\mathbf{x}$ in a more efficient way than by explicitly forming the Cholesky factorization of $\mathbf{X}$. Let $\mathbf{i}$ denote the imaginary unit ($\mathbf{i}^2 = -1$). Then (2.9) can be rewritten as

$$\mathbf{X} = \mathbf{Y}^T \mathbf{Y}, \tag{2.10}$$

where

$$\mathbf{Y} = \begin{bmatrix} \mathbf{i}\,\mathbf{x}^T \\ \mathbf{R} \end{bmatrix}$$

is an $(n+1) \times n$ matrix.

Now consider an $(n+1) \times (n+1)$ matrix $\tilde{\mathbf{V}}_{k,k+1}$ of the form

$$\tilde{\mathbf{V}}_{k,k+1} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \mathbf{i}\,\alpha_k & \beta_k & \\ & & -\beta_k & \mathbf{i}\,\alpha_k & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \leftarrow k\text{th row} \tag{2.11}$$

where α_k and β_k are real and $\tilde{\mathbf{V}}_{k,k+1}$ has the property that

$$\tilde{\mathbf{V}}_{k,k+1}^T \tilde{\mathbf{V}}_{k,k+1} = \mathbf{I}. \tag{2.12}$$

Note that $\tilde{\mathbf{V}}_{k,k+1}$ operates exclusively in the plane $(k, k+1)$. Thus premultiplication by $\tilde{\mathbf{V}}_{k,k+1}$ alters elements only in rows k and $k+1$.

Consider any $(n+1)$-vector which in rows k and $k+1$ has elements $\mathbf{i}\,\gamma$ and δ respectively, γ and δ real. If

$$\delta^2 > \gamma^2$$

then it is possible to choose α_k and β_k in such a way that in rows k and $k+1$ we have

$$\begin{pmatrix} \mathbf{i}\,\alpha_k & \beta_k \\ -\beta_k & \mathbf{i}\,\alpha_k \end{pmatrix} \begin{pmatrix} \mathbf{i}\,\gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \omega \\ 0 \end{pmatrix}. \tag{2.13}$$

Indeed if

$$\omega = (\delta^2 - \gamma^2)^{\frac{1}{2}}$$

then

$$\alpha_k = \gamma/\omega$$
$$\beta_k = \delta/\omega$$

satisfies both (2.12) and (2.13).

We can formally define a sequence of matrices $\tilde{\mathbf{V}}_{n,n+1}, \tilde{\mathbf{V}}_{n-1,n}, \ldots, \tilde{\mathbf{V}}_{1,2}$ which triangularize the matrix $\mathbf{Y}=[\mathbf{i}\,\mathbf{x}, \mathbf{R}^T]^T$. Let $[0, \ldots, 0, \mathbf{i}\,\hat{a}_{k,k}, \ldots, \mathbf{i}\,\hat{a}_{k,n}]$ denote the kth row of the matrix $\tilde{\mathbf{V}}_{k-1,k} \cdots \tilde{\mathbf{V}}_{1,2}\mathbf{Y}$. The $(k+1)$st row $\tilde{\mathbf{V}}_{k-1,k} \cdots \tilde{\mathbf{V}}_{1,2}\mathbf{Y}$ is identical to the kth row of $\mathbf{R}$ and has the representation $[0, \ldots, 0, r_{k,k}, \ldots, r_{k,n}]$. The transformation $\tilde{\mathbf{V}}_{k,k+1}$ is chosen so as to annihilate the $(k+1, k)$ element in the matrix $\tilde{\mathbf{V}}_{k-1,k} \cdots \tilde{\mathbf{V}}_{1,2}\mathbf{Y}$. The rows k and $k+1$ of the product $\tilde{\mathbf{V}}_{k,k+1}, \ldots, \tilde{\mathbf{V}}_{1,2}\mathbf{Y}$ are obtained from the formula

$$\begin{pmatrix} \mathbf{i}\,\alpha_k & \beta_k \\ -\beta_k & \mathbf{i}\,\alpha_k \end{pmatrix} \begin{pmatrix} 0, \ldots, 0, \mathbf{i}\,\hat{a}_{k,k}, \ldots, \mathbf{i}\,\hat{a}_{k,n} \\ 0, \ldots, 0, r_{k,k}, \ldots, r_{k,n} \end{pmatrix} = \begin{pmatrix} 0, \ldots, 0, \tilde{r}_{k,k}, \tilde{r}_{k,k+1}, \ldots, \tilde{r}_{k,n} & \\ 0, \ldots, 0,\ 0,\ \mathbf{i}\hat{a}_{k+1,k+1}, \ldots, \mathbf{i}\hat{a}_{k+1,n} \end{pmatrix}, \tag{2.14}$$

where

$$\tilde{r}_{k,k} = (r_{k,k}^2 - \hat{a}_{k,k}^2)^{\frac{1}{2}},$$
$$\alpha_k = \hat{a}_{k,k}/\tilde{r}_{k,k},$$
$$\beta_k = r_{k,k}/\tilde{r}_{k,k}, \tag{2.15}$$

and for $j = k+1, \ldots, n$

$$\tilde{r}_{k,j} = -\alpha_k\,\hat{a}_{k,j} + \beta_k\,r_{k,j}$$
$$\hat{a}_{k+1,j} = -\beta_k\,\hat{a}_{k,j} + \alpha_k\,r_{k,j}.$$

Here $[0, \ldots, 0, \tilde{r}_{k,k}, \ldots, \tilde{r}_{k,n}]$ denotes the kth row of the matrix $\tilde{\mathbf{R}}$. It can be shown that our assumption that $\mathbf{X} = \mathbf{R}^T\mathbf{R} - \mathbf{x}\,\mathbf{x}^T$ is positive definite implies that $\tilde{r}_{k,k}$ is real. Thus (2.14) can be performed entirely in real arithmetic. The matrix $\tilde{\mathbf{V}}_{n,n+1}\tilde{\mathbf{V}}_{n-1,n} \cdots \tilde{\mathbf{V}}_{1,2}\mathbf{Y} = [\hat{\mathbf{R}}^T, \mathbf{O}]^T$ is real and upper triangular.

Let $\tilde{\mathbf{V}} = \tilde{\mathbf{V}}_{n,n+1}\tilde{\mathbf{V}}_{n-1,n} \cdots \tilde{\mathbf{V}}_{1,2}$. Then using (2.11) we see that

$$\mathbf{X} = \mathbf{Y}^T\,\mathbf{Y} = \mathbf{Y}^T\,\tilde{\mathbf{V}}^T\,\tilde{\mathbf{V}}\,\mathbf{Y} = \tilde{\mathbf{R}}^T\,\tilde{\mathbf{R}},$$

i.e., $\tilde{\mathbf{R}}$ is the desired Cholesky factor of $\mathbf{X}$. As Calculations (2.14) and (2.15) are to be repeated n times, the cost of obtaining $\tilde{\mathbf{R}}$ is $2n^2 + O(n)$ multiplications.

The structure of the downdating procedure is analogous to the structure of the updating procedure. The significant point is that in both cases the matrix $\tilde{\mathbf{R}}$ can be generated from the matrix $\mathbf{R}$ and vector $\mathbf{x}$, row by row starting from the top row, and that each consecutive row of $\tilde{\mathbf{R}}$ is computed using knowledge of only two other rows: see (2.7) and (2.14).

We now describe a conceptually different but algebraically equivalent approach to the downdating problem. Recall that we want to find the Cholesky factor $\tilde{\mathbf{R}}$ of the positive definite matrix $\mathbf{X}$ defined by (2.9), where $\mathbf{R}$ and $\mathbf{x}$ are given. For the sake of uniqueness we assume that $\tilde{\mathbf{R}}$ has positive diagonal elements.

Suppose for a moment that $\tilde{\mathbf{R}}$ is known and that we want to find $\mathbf{R}$. The Relation (2.9) is equivalent to

$$\mathbf{R}^T\,\mathbf{R} = \tilde{\mathbf{R}}^T\,\tilde{\mathbf{R}} + \mathbf{x}\,\mathbf{x}^T. \tag{2.16}$$

Thus we have the updating problem.

Let $\mathbf{U} = \mathbf{U}_{n,n+1} \ldots \mathbf{U}_{1,2}$ be a product of a sequence of plane rotations which transform $\mathbf{Y} = [\mathbf{x}, \tilde{\mathbf{R}}^T]^T$ into upper triangular form $[\mathbf{R}^T, \mathbf{O}]^T$. Denoting by $[0, \ldots, 0, \hat{a}_{k,k}, \ldots, \hat{a}_{k,n}]$ the kth row of $\mathbf{U}_{k-1,k} \cdots \mathbf{U}_{1,2}\mathbf{Y}$ and by

$[0, \ldots, 0, \tilde{r}_{k,k}, \ldots, \tilde{r}_{k,n}]$ and $[0, \ldots, 0, r_{k,k}, \ldots, r_{k,n}]$ the kth rows of $\tilde{\mathbf{R}}$ and $\mathbf{R}$ respectively, for rows k and $k+1$ we have relations similar to (2.7) and (2.8), i.e.,

$$\begin{pmatrix} c_k & s_k \\ -s_k & c_k \end{pmatrix} \begin{pmatrix} 0, \ldots, 0, \hat{a}_{k,k}, \ldots, \hat{a}_{k,n} \\ 0, \ldots, 0, \tilde{r}_{k,k}, \ldots, \tilde{r}_{k,n} \end{pmatrix} = \begin{pmatrix} 0, \ldots, 0, r_{k,k}, r_{k,k+1}, \ldots, r_{k,n} \\ 0, \ldots, 0, \ 0, \ \hat{a}_{k+1,k+1}, \ldots, \hat{a}_{k+1,n} \end{pmatrix}, \tag{2.17}$$

where

$$r_{k,k} = (\hat{a}_{k,k}^2 + \tilde{r}_{k,k}^2)^{\frac{1}{2}}, \tag{2.18a}$$

$$c_k = \hat{a}_{k,k}/r_{k,k}, \tag{2.18b}$$

$$s_k = \tilde{r}_{k,k}/r_{k,k}, \tag{2.18c}$$

and for $j = k+1, \ldots, n$

$$r_{k,j} = c_k \hat{a}_{k,j} + s_k \tilde{r}_{k,j}, \tag{2.18d}$$

$$\hat{a}_{k+1,j} = -s_k \hat{a}_{k,j} + c_k \tilde{r}_{k,j}. \tag{2.18e}$$

Now return to the downdating problem. Assume that $[0, \ldots, 0, \hat{a}_{k,k}, \ldots, \hat{a}_{k,n}]$ and $[0, \ldots, 0, r_{k,k}, \ldots, r_{k,n}]$ are known. By solving (2.18a) and (2.18d) for $[0, \ldots, 0, \tilde{r}_{k,k}, \ldots, \tilde{r}_{k,n}]$ we obtain

$$\tilde{r}_{k,k} = (r_{k,k}^2 - \hat{a}_{k,k}^2)^{\frac{1}{2}} \tag{2.19a}$$

and for $j = k+1, \ldots, n$

$$\tilde{r}_{k,j} = (r_{k,j} - c_k \hat{a}_{k,j})/s_k. \tag{2.19b}$$

Relations (2.19a), (2.19b) and (2.18e) give us an alternative way of computing the matrix $\tilde{\mathbf{R}}$ from $\mathbf{R}$ and $\mathbf{x}$.

Relations (2.19b) and (2.18e) are algebraically equivalent to the Relations (2.15). When we express c_k and s_k in (2.19b) and (2.18e) by (2.18b) and (2.18c) we get exactly the same formulae for $\tilde{r}_{k,j}$ and $\hat{a}_{k+1,j}$ as in (2.15). However, we prefer to use Formulae (2.19b) and (2.18e) for considerations related to their stability.

We conclude that one can treat the downdating problem as the reverse of the updating problem, and that there is a one-to-one correspondence between the matrix $\mathbf{U}_{k,k+1}$ defined by (2.17) and the matrix $\tilde{\mathbf{V}}_{k,k+1}$ defined by (2.14); each uniquely determines the other.

In the next section we make use of both downdating techniques described here.

3. An Algorithm for the QR Decomposition of a Toeplitz Matrix

Let $\mathbf{T}$ be a full rank $m \times n$ Toeplitz matrix, $m \geqq n$,

$$\mathbf{T} = \begin{pmatrix} t_0 & t_{-1} & \ldots t_{-n+1} \\ t_1 & t_0 & \ldots t_{-n+2} \\ \cdot & & \ddots \\ \cdot & & t_0 \\ \cdot & & t_1 \\ \cdot & & \vdots \\ t_{m-1} & t_{m-2} & t_{m-n} \end{pmatrix}.$$

Using the shift invariance property of Toeplitz matrices we can partition $\mathbf{T}$ in two ways:

$$\mathbf{T}=\begin{pmatrix} t_0 & \mathbf{y}^T \\ \mathbf{x} & \mathbf{T}_{-1} \end{pmatrix} \tag{3.1}$$

and

$$\mathbf{T}=\begin{pmatrix} \mathbf{T}_{-1} & \mathbf{y}^R \\ \mathbf{x}^{RT} & t_{m-n} \end{pmatrix}, \tag{3.2}$$

where $\mathbf{T}_{-1}$ is an $(m-1)\times(n-1)$ principal submatrix of $\mathbf{T}$,

$$\begin{aligned} \mathbf{x}^T &= [t_1, t_2, \ldots, t_{m-1}], \\ \mathbf{x}^{RT} &= [t_{m-1}, \ldots, t_{m-n+1}], \\ \mathbf{y}^T &= [t_{-1}, \ldots, t_{-n+1}], \\ \mathbf{y}^{RT}&[t_{-n+1}, \ldots, t_{m-n-1}]. \end{aligned}$$

3.1. Calculation of the Triangular Factor $\mathbf{R}$

Let $\mathbf{R}$ be an upper triangular factor of the Cholesky decomposition of $\mathbf{T}^T\mathbf{T}$, i.e.,

$$\mathbf{R}^T\mathbf{R}=\mathbf{T}^T\mathbf{T}. \tag{3.3}$$

It is well known that matrix $\mathbf{R}$ is the upper triangular factor from the *QR* decomposition of $\mathbf{T}$.

In a similar manner, we partition $\mathbf{R}$ in two different ways

$$\mathbf{R}=\begin{pmatrix} r_{1,1} & \mathbf{r}_{fr} \\ \mathbf{O} & \mathbf{R}_b \end{pmatrix} \tag{3.4}$$

where $\mathbf{r}_{fr}=[r_{1,2}, \ldots, r_{1,n}]$ is the first row of $\mathbf{R}$ except for its first element, $\mathbf{R}_b$ is an $(n-1)\times(n-1)$ principal bottom submatrix of $\mathbf{R}$, and

$$\mathbf{R}=\begin{pmatrix} \mathbf{R}_t & \mathbf{r}_{1c} \\ \mathbf{O} & r_{n,n} \end{pmatrix}, \tag{3.5}$$

where $\mathbf{r}_{1c}=[r_{1,n}, \ldots, r_{n-1,n}]^T$ is the last column of $\mathbf{R}$ except for its last element and $\mathbf{R}_t$ is an $(n-1)\times(n-1)$ principal top submatrix of $\mathbf{R}$.

From (3.1), (3.4) and (3.3) we get

$$\left(\begin{array}{c|c} r_{1,1}^2 & r_{1,1}\mathbf{r}_{fr} \\ \hline \mathbf{r}_{fr}^T r_{1,1} & \mathbf{r}_{fr}^T\mathbf{r}_{fr}+\mathbf{R}_b^T\mathbf{R}_b \end{array}\right)=\left(\begin{array}{c|c} t_0^2+\mathbf{x}^T\mathbf{x} & t_0\mathbf{y}^T+\mathbf{x}^T\mathbf{T}_{-1} \\ \hline t_0\mathbf{y}+\mathbf{T}_{-1}^T\mathbf{x} & \mathbf{y}\mathbf{y}^T+\mathbf{T}_{-1}^T\mathbf{T}_{-1} \end{array}\right). \tag{3.6}$$

Similarly, from (3.2), (3.5) and (3.3) we obtain

$$\left(\begin{array}{c|c} \mathbf{R}_t^T\mathbf{R}_t & \mathbf{R}_t^T\mathbf{r}_{1c} \\ \hline \mathbf{r}_{1c}^T\mathbf{R}_t & \mathbf{r}_{1c}^T\mathbf{r}_{1c}+r_{n,n}^2 \end{array}\right)=\left(\begin{array}{c|c} \mathbf{T}_{-1}^T\mathbf{T}_{-1}+\mathbf{x}^R\mathbf{x}^{RT} & \mathbf{T}^T\mathbf{y}^R+\mathbf{x}^R t_0 \\ \hline \mathbf{y}^{RT}\mathbf{T}_{-1}+t_{m-n}\mathbf{x}^{RT} & \mathbf{y}^{RT}\mathbf{y}^R+t_{m-n}^2 \end{array}\right). \tag{3.7}$$

Comparison of the upper left submatrices on both sides of (3.7) gives

$$\mathbf{R}_t^T \mathbf{R}_t = \mathbf{T}_{-1}^T \mathbf{T}_{-1} + \mathbf{x}^R \mathbf{x}^{RT} \tag{3.8}$$

while comparison of the lower right matrices on both sides of (3.6) gives

$$\mathbf{R}_b^T \mathbf{R}_b + \mathbf{r}_{fr}^T \mathbf{r}_{fr} = \mathbf{T}_{-1}^T \mathbf{T}_{-1} + \mathbf{y}\mathbf{y}^T. \tag{3.9}$$

Finally, (3.8) and (3.9) give the main relation

$$\mathbf{R}_b^T \mathbf{R}_b = \mathbf{R}_t^T \mathbf{R}_t + \mathbf{y}\mathbf{y}^T - \mathbf{x}^R \mathbf{x}^{RT} - \mathbf{r}_{fr}^T \mathbf{r}_{fr} \tag{3.10}$$

where, from (3.6), $\mathbf{r}_{fr}$ is given by

$$\mathbf{r}_{fr} = (t_0 \mathbf{y}^T + \mathbf{x}^T \mathbf{T}_{-1})/(t_0^2 + \mathbf{x}^T \mathbf{x})^{\frac{1}{2}}. \tag{3.11}$$

Relations (3.10) and (3.11) form a base for computing $\mathbf{R}$. Relation (3.10) says that matrix $\mathbf{R}_b$ can be obtained from matrix $\mathbf{R}_t$ following a rank three modification. When implemented as a rank-1 update followed by two rank-1 downdates, Relation (3.10) gives us a means of calculating the kth row of $\mathbf{R}_b$ from the first k rows of $\mathbf{R}_t$. However, the kth row of $\mathbf{R}_b$ is identical to the $(k+1)$st row of $\mathbf{R}_t$. As the first row of $\mathbf{R}_t$ is defined by (3.11), we have a recursion for calculating the rows of $\mathbf{R}$.

We write (3.10) as a sequence of three rank-1 mofifications

$$\mathbf{R}_1^T \mathbf{R}_1 = \mathbf{R}_t^T \mathbf{R}_t + \mathbf{y}\mathbf{y}^T, \tag{3.12a}$$

$$\mathbf{R}_2^T \mathbf{R}_2 = \mathbf{R}_1^T \mathbf{R}_1 - \mathbf{x}^R \mathbf{x}^{RT}, \tag{3.12b}$$

$$\mathbf{R}_b^T \mathbf{R}_b = \mathbf{R}_2^T \mathbf{R}_2 - \mathbf{r}_{fr}^T \mathbf{r}_{fr}. \tag{3.12c}$$

Define $\mathbf{U}$, $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{W}}$ by the following relations

$$\mathbf{U}\begin{bmatrix}\mathbf{y}^T\\ \mathbf{R}_t\end{bmatrix} = \begin{bmatrix}\mathbf{R}_1\\ \mathbf{O}^T\end{bmatrix}, \tag{3.13a}$$

$$\tilde{\mathbf{V}}\begin{bmatrix}\mathbf{i}\,\mathbf{x}^{RT}\\ \mathbf{R}_1\end{bmatrix} = \begin{bmatrix}\mathbf{R}_2\\ \mathbf{O}^T\end{bmatrix}, \tag{3.13b}$$

$$\tilde{\mathbf{W}}\begin{bmatrix}\mathbf{i}\,\mathbf{r}_{fr}\\ \mathbf{R}_2\end{bmatrix} = \begin{bmatrix}\mathbf{R}_b\\ \mathbf{O}^T\end{bmatrix}, \tag{3.13c}$$

where

$$\mathbf{U} = \mathbf{U}_{n-1,n} \cdots \mathbf{U}_{1,2}$$

is a product of plane rotations of the form (2.6), and

$$\tilde{\mathbf{V}} = \tilde{\mathbf{V}}_{n-1,n} \cdots \tilde{\mathbf{V}}_{1,2}$$

and

$$\tilde{\mathbf{W}} = \tilde{\mathbf{W}}_{n-1,n} \cdots \tilde{\mathbf{W}}_{1,2}$$

are products of transformations of the form (2.11).

In the Relation (3.13a) the vector $\mathbf{y}$ is known, and from (3.11) the first row of $\mathbf{R}_t$ is known as well. Thus we know the first two rows of the matrix $[\mathbf{y}, \mathbf{R}_t^T]^T$. This allows us to apply to $[\mathbf{y}, \mathbf{R}_t^T]^T$ the first rotation $\mathbf{U}_{1,2}$ from (3.13a). As $\mathbf{U}_{1,2}$ is the only rotation which operates on the first row of $[\mathbf{y}, \mathbf{R}_t^T]^T$, multiplication by $\mathbf{U}_{1,2}$ gives the first row of the matrix on the right hand side of (3.13a), i.e., the first row of $\mathbf{R}_1$. The second row of the product plays the role of $\mathbf{y}$ in the subsequent calculations. Now in the Relation (3.13b) the first two rows of $[\mathbf{i}\,\mathbf{x}^R, \mathbf{R}_1^T]^T$ are known: we know the vector $\mathbf{x}^R$ and, from the previous calculation, the first row of $\mathbf{R}_1$. This is sufficient to apply the first transformation $\tilde{\mathbf{V}}_{1,2}$ from (3.13b), which gives the first row of the matrix $\mathbf{R}_2$. Finally, in the relation (3.13c) we know the first two rows of the matrix $[\mathbf{i}\,\mathbf{r}_{fr}^T, \mathbf{R}_2^T]^T$; the vector $\mathbf{r}_{fr}$ is that given by (3.11) while the first row of $\mathbf{R}_2$ is known from the previous step. Thus we can apply the transformation $\tilde{\mathbf{W}}_{1,2}$ which gives us the first row of $\mathbf{R}_b$. The first row of $\mathbf{R}_b$ is identical to the second row of $\mathbf{R}_t$, so we can repeat the whole procedure to get the second row of $\mathbf{R}_1$, then the second row of $\mathbf{R}_2$, and finally the second row of $\mathbf{R}_b$ which is identical to the third row of $\mathbf{R}_t$. By applying this procedure to successive rows of $\mathbf{R}_t$, $\mathbf{R}_1$ and $\mathbf{R}_2$, we obtain all rows of $\mathbf{R}_b$ which, together with (3.11), gives us the matrix $\mathbf{R}$.

The kth step of the procedure for computing rows of the matrix $\mathbf{R}$ is described below.

Let us denote rows k and $k+1$ of $\mathbf{U}_{k+1,k} \cdots \mathbf{U}_{1,2}[\mathbf{y}, \mathbf{R}_t^T]^T$ by

$$[0, \ldots, 0, \tilde{r}_{k,k}, \ldots, \tilde{r}_{k,n-1}]$$

and

$$[0, \ldots, 0, r_{k,k}, \ldots, r_{k,n-1}],$$

rows k and $k+1$ of $\tilde{\mathbf{V}}_{k-1,k} \cdots \tilde{\mathbf{V}}_{1,2}[\mathbf{i}\,\mathbf{x}^R, \mathbf{R}_1^T]^T$ by

$$[0, \ldots, 0, \mathbf{i}\,\tilde{a}_{k,k}, \ldots, \mathbf{i}\,\tilde{a}_{k,n-1}]$$

and

$$[0, \ldots, 0, a_{k,k}, \ldots, a_{k,n-1}],$$

and rows k and $k+1$ of $\tilde{\mathbf{W}}_{k-1,k} \cdots \tilde{\mathbf{W}}_{1,2}[\mathbf{i}\,\mathbf{r}_{fr}^T, \mathbf{R}_2^T]^T$ by

$$[0, \ldots, 0\ \mathbf{i}\,\hat{a}_{k,k}, \ldots, \mathbf{i}\,\hat{a}_{k,n-1}]$$

and

$$[0, \ldots, 0, \hat{r}_{k,k}, \ldots, \hat{r}_{k,n-1}].$$

In the kth step of the recursion we perform the following calculations for rows k and $k+1$ (corresponding to multiplications by $\mathbf{U}_{k,k+1}$, $\tilde{\mathbf{V}}_{k,k+1}$ and $\tilde{\mathbf{W}}_{k,k+1}$):

- multiplication by $\mathbf{U}_{k,k+1}$

$$\begin{pmatrix} c_k & s_k \\ -s_k & c_k \end{pmatrix} \begin{pmatrix} 0, \ldots, 0, \tilde{r}_{k,k}, \ldots, \tilde{r}_{k,n-1} \\ 0, \ldots, 0, r_{k,k}, \ldots, r_{k,n-1} \end{pmatrix} = \begin{pmatrix} 0, \ldots, 0, a_{k,k}, a_{k,k+1}, \ldots, a_{k,n-1} \\ 0, \ldots, 0,\ 0,\ \tilde{r}_{k+1,k+1}, \ldots, \tilde{r}_{k+1,n-1} \end{pmatrix} \quad (3.14\text{a})$$

- multiplication by $\tilde{\mathbf{V}}_{k,k+1}$

$$\begin{pmatrix} \mathbf{i}\alpha_k & \alpha_k \\ -\beta_k & \mathbf{i}\beta_k \end{pmatrix} \begin{pmatrix} 0,\dots,0,\mathbf{i}\tilde{a}_{k,k},\dots,\mathbf{i}\tilde{a}_{k,n-1} \\ 0,\dots,0,a_{k,k},\dots,a_{k,n-1} \end{pmatrix} = \begin{pmatrix} 0,\dots,0,\hat{r}_{k,k},\hat{r}_{k,k+1},\dots,\hat{r}_{k,n-1} \\ 0,\dots,0,\ 0,\ \mathbf{i}\tilde{a}_{k+1,k+1},\dots,\mathbf{i}\tilde{a}_{k+1,n-1} \end{pmatrix} \tag{3.14b}$$

- multiplication by $\tilde{\mathbf{W}}_{k,k+1}$

$$\begin{pmatrix} \mathbf{i}\gamma_k & \delta_k \\ -\delta_k & \mathbf{i}\gamma_k \end{pmatrix} \begin{pmatrix} 0,\dots,0,\mathbf{i}\hat{a}_{k,k},\dots,\mathbf{i}\hat{a}_{k,n-1} \\ 0,\dots,0,\hat{r}_{k,k},\dots,\hat{r}_{k,n-1} \end{pmatrix} = \begin{pmatrix} 0,\dots,0,\ r_{k+1,k+1},\ r_{k+1,k+2},\dots,r_{k,n-1} \\ 0,\dots,0,\ 0,\ \mathbf{i}\hat{a}_{k+1,k+1},\dots,\mathbf{i}\hat{a}_{k+1,n} \end{pmatrix}. \tag{3.14c}$$

As Steps (3.14a)–(3.14c) are to be repeated for $k=1,2,\dots,n-1$, the cost of obtaining $\mathbf{R}$ from $\mathbf{y}$, $\mathbf{x}^R$ and $\mathbf{r}_{fr}$ is $6n^2+O(n)$ multiplications. Because the calculation of the vector $\mathbf{r}_{fr}$ requires an additional $mn+O(n)$ multiplications, the cost of obtaining the upper triangular factor $\mathbf{R}$ in the QR decomposition of the matrix $\mathbf{T}$ is $mn+6n^2+O(n)$ multiplications.

3.2. *Calculation of the Orthogonal Factor Q*

Next we describe a procedure for generating an $m\times n$ matrix $\mathbf{Q}$ with orthonormal columns such that $\mathbf{T}=\mathbf{QR}$. We assume that the transformations $\mathbf{U}$, $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{W}}$ defined by Relations (3.13a)–(3.13c) are known from the first algorithm which computes the matrix $\mathbf{R}$. This in turn means that the sequence $\mathbf{U}_{n-1,n},\dots,\mathbf{U}_{1,2}$ of plane rotations which define the transformation $\mathbf{U}$, and the sequences $\mathbf{V}_{n-1,n},\dots,\mathbf{V}_{1,2}$ and $\mathbf{W}_{n-1,n},\dots,\mathbf{W}_{1,2}$ of plane rotations which define transformations $\mathbf{V}$ and $\mathbf{W}$ corresponding to transformations $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{W}}$, are also known.

Let the matrix $\mathbf{T}=[\mathbf{t}_1,\mathbf{t}_2,\dots,\mathbf{t}_n]$ be partitioned as follows:

$$\mathbf{T}=[\mathbf{C},\mathbf{t}_n]=[\mathbf{t}_1,\mathbf{D}]. \tag{3.15}$$

Using the Definitions (3.1) and (3.2), matrices $\mathbf{C}$ and $\mathbf{D}$ can be partitioned in the following way:

$$\mathbf{C}=[\mathbf{T}_{-1}^T,\mathbf{x}^R]^T, \quad \mathbf{D}=[\mathbf{y},\mathbf{T}_{-1}^T]^T. \tag{3.16}$$

We want to find an $m\times n$ matrix $\mathbf{Q}$, with orthonormal columns, such that

$$\mathbf{Q}^T\mathbf{T}=\mathbf{Q}^T[\mathbf{C},\mathbf{t}_n]=\begin{pmatrix} \mathbf{R}_t & \mathbf{z} \\ \mathbf{O} & r_{n,n} \end{pmatrix} =\mathbf{Q}^T[\mathbf{t}_1,\mathbf{D}]=\begin{pmatrix} r_{1,1} & \mathbf{r}_{fr} \\ \mathbf{O} & \mathbf{R}_b \end{pmatrix}=\mathbf{R}, \tag{3.17}$$

where $\mathbf{r}_{fr}$, $\mathbf{R}_t$ and $\mathbf{R}_b$ are defined by (3.4) and (3.5), $\mathbf{z}$ is the last column of $\mathbf{R}$ (except for the last element $r_{n,n}$), and $r_{1,1}$ and $r_{n,n}$ are elements of $\mathbf{R}$.

Assume for a moment that we know an $n \times m$ matrix $\mathbf{Q}_c^T$ with orthonormal rows such that

$$\mathbf{Q}_c^T \mathbf{C} = \begin{bmatrix} \mathbf{R}_t \\ \mathbf{O}^T \end{bmatrix} \tag{3.18}$$

and

$$\mathbf{Q}_c^T \mathbf{D} = \begin{bmatrix} \mathbf{r}_{fr} \\ \mathbf{R}_b \end{bmatrix}. \tag{3.19}$$

Comparing (3.17) with (3.18) and (3.19) we see that the matrix $\mathbf{Q}_c$ is the required matrix $\mathbf{Q}$.

Premultiplying $\mathbf{Q}_c^T \mathbf{D}$ by a sequence of plane rotations $\mathbf{W}_{1,2}, \ldots, \mathbf{W}_{n-1,n}$, the matrix $\mathbf{Q}_c^T \mathbf{D}$ can be transformed into upper triangular form $\mathbf{R}_2$ (see (3.13c)),

$$\mathbf{W}_{n-1,n} \cdots \mathbf{W}_{1,2} \mathbf{Q}_c^T \mathbf{D} = \begin{bmatrix} \mathbf{R}_2 \\ \mathbf{O}^T \end{bmatrix}. \tag{3.20}$$

If we denote by $\mathbf{W}$ the product of plane rotations $\mathbf{W}_{n-1,n}, \ldots, \mathbf{W}_{1,2}$, and define

$$\mathbf{Q}_d^T = \mathbf{W} \mathbf{Q}_c^T, \tag{3.21}$$

then we see that we have the following relation:

$$\mathbf{Q}_d^T \mathbf{D} = \begin{bmatrix} \mathbf{R}_2 \\ \mathbf{O}^T \end{bmatrix}. \tag{3.22}$$

Note that the matrix $\mathbf{W}$ corresponds to the matrix $\tilde{\mathbf{W}}$ of (3.13c).

To get the second relation, which will give us a means of calculating the matrix $\mathbf{Q}$, consider an $(m+1) \times (n-1)$ matrix $\mathbf{B}$:

$$\mathbf{B} \overset{df}{=} \begin{pmatrix} \mathbf{y}^T \\ \mathbf{T}_{-1} \\ \mathbf{x}^{RT} \end{pmatrix} = \begin{pmatrix} \mathbf{y}^T \\ \mathbf{C} \end{pmatrix} = \begin{pmatrix} \mathbf{D} \\ \mathbf{x}^{RT} \end{pmatrix}. \tag{3.23}$$

Define two $(n+1) \times (n+1)$ orthogonal matrices $\hat{\mathbf{Q}}$ and $\tilde{\mathbf{Q}}$ by

$$\hat{\mathbf{Q}} = \begin{pmatrix} 1 & \mathbf{O}^T \\ \mathbf{O} & \mathbf{Q}_c^T \end{pmatrix} \tag{3.24}$$

and

$$\tilde{\mathbf{Q}} = \begin{pmatrix} \mathbf{O}^T & 1 \\ \mathbf{Q}_d^T & \mathbf{O} \end{pmatrix}, \tag{3.25}$$

and form two products $\tilde{\mathbf{Q}} \mathbf{B}$ and $\hat{\mathbf{Q}} \mathbf{B}$. From (3.18) and (3.21) we have

$$\hat{\mathbf{Q}} \mathbf{B} = \begin{pmatrix} \mathbf{y}^T \\ \mathbf{R}_t \\ \mathbf{O}^T \end{pmatrix} \tag{3.26}$$

and

$$\tilde{\mathbf{Q}}\mathbf{B}=\begin{pmatrix}\mathbf{x}^{RT}\\ \mathbf{R}_2\\ \mathbf{O}^T\end{pmatrix}. \tag{3.27}$$

Let $\mathbf{U}_{-1}=\mathbf{U}_{n-1,n}\dots\mathbf{U}_{1,2}$ denote the product of the sequence of plane rotations which transform the matrix $[\mathbf{y},\mathbf{R}_t^T]^T$ into upper triangular form $\mathbf{R}_1$ (see (3.13a)). Then for the matrix

$$\mathbf{U}=\begin{pmatrix}\mathbf{U}_{-1} & \mathbf{O}\\ \mathbf{O}^T & 1\end{pmatrix}$$

we have

$$\mathbf{U}\hat{\mathbf{Q}}\mathbf{B}=\begin{bmatrix}\mathbf{R}_1\\ \mathbf{O}^T\\ \mathbf{O}^T\end{bmatrix} \tag{3.28}$$

The matrix $\mathbf{U}$ corresponds to the matrix $\mathbf{U}$ in (3.13a).

Similarly, let $\mathbf{V}_{-1}=\mathbf{V}_{n-1,n}\dots\mathbf{V}_{1,2}$ denote the product of a sequence of plane rotations which transforms the matrix $[\mathbf{x}^R,\mathbf{R}_2^T]^T$ into upper triangular form $[\mathbf{R}_1^T,\mathbf{O}]^T$. Then for the matrix

$$\mathbf{V}=\begin{pmatrix}\mathbf{V}_{-1} & \mathbf{O}\\ \mathbf{O}^T & 1\end{pmatrix}$$

we have

$$\mathbf{V}\tilde{\mathbf{Q}}\mathbf{B}=\begin{bmatrix}\mathbf{R}_1\\ \mathbf{O}^T\\ \mathbf{O}^T\end{bmatrix}. \tag{3.29}$$

The matrix $\mathbf{V}$ corresponds to the matrix $\tilde{\mathbf{V}}$ in (3.13b).

Comparing (3.28) with (3.29) we obtain a relationship between the first $n-1$ rows of $\mathbf{U}\hat{\mathbf{Q}}$ and $\mathbf{V}\tilde{\mathbf{Q}}$:

$$[\mathbf{I}_n,\mathbf{O}]\,\mathbf{U}\begin{pmatrix}1 & \mathbf{O}^T\\ \mathbf{O} & \mathbf{Q}_c^T\end{pmatrix}=[\mathbf{I}_n,\mathbf{O}]\,\mathbf{V}\begin{pmatrix}\mathbf{O}^T & 1\\ \mathbf{Q}_d^T & \mathbf{O}\end{pmatrix}. \tag{3.30}$$

Relations (3.22) and (3.30) form a base for the algorithm which calculates the orthogonal matrix $\mathbf{Q}$.

We describe the procedure for computing $\mathbf{Q}=\mathbf{Q}_c$ by considering the kth step of the algorithm. Recall that the sequences $\mathbf{U}_{n-1,n},\dots,\mathbf{U}_{1,2}$, $\mathbf{V}_{n-1,n},\dots,\mathbf{V}_{1,2}$ and $\mathbf{W}_{n-1,n},\dots,\mathbf{W}_{1,2}$ of plane rotations that define transformations $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$ are known from the algorithm for computing the matrix $\mathbf{R}$. Assume that prior to the kth step the first k rows of $\mathbf{Q}_c^T$ and the first $(k-1)$ rows of $\mathbf{Q}_d^T$ have been computed. In the kth step the kth row of $\mathbf{Q}_d^T$ is determined from the Relation (3.30) and knowledge of $\mathbf{U}_{k,k+1}$ and $\mathbf{V}_{k,k+1}$, and the $(k+1)$st row of $\mathbf{Q}_c^T$ is determined from the Relation (3.22) and knowledge of $\mathbf{W}_{k,k+1}$.

Let $\mathbf{Q}_c^T$ have kth row $[q_{k,1},\dots,q_{k,m}]$ and $\mathbf{Q}_d^T$ have kth row $[\tilde{q}_{k,1},\dots,\tilde{q}_{k,m}]$. Denote rows k and $k+1$ of $\mathbf{U}_{k-1,k}\dots\mathbf{U}_{1,2}\mathbf{Q}$ by

$$\begin{pmatrix}\hat{a}_{k,1},\dots,\hat{a}_{k,m},\hat{a}_{k,m+1}\\ 0,q_{k,1},\dots,q_{k,m}\end{pmatrix}, \tag{3.31}$$

rows k and $k+1$ of $\mathbf{V}_{k-1,k}\cdots\mathbf{V}_{1,2}\tilde{\mathbf{Q}}$ by

$$\begin{pmatrix}\tilde{a}_{k,1},\ldots,\tilde{a}_{k,m},\tilde{a}_{k,m+1}\\ \tilde{q}_{k,1},\ldots,\tilde{q}_{k,m},0\end{pmatrix},\tag{3.32}$$

and rows k and $k+1$ of $\mathbf{W}_{k-1,k}\cdots\mathbf{W}_{1,2}\mathbf{Q}_c^T$ by

$$\begin{pmatrix}p_{k,1},\ldots,p_{k,m}\\ q_{k+1,1},\ldots,q_{k+1,m}\end{pmatrix}.\tag{3.33}$$

Premultiplying (3.31) by $\mathbf{U}_{k,k+1}$ and (3.32) by $\mathbf{V}_{k,k+1}$, and using (3.30), we have

$$\mathbf{U}_{k,k+1}\begin{pmatrix}\hat{a}_{k,1},\ldots,\hat{a}_{k,n},\hat{a}_{k,m+1}\\ 0,q_{k,1},\ldots,q_{k,m}\end{pmatrix}=\mathbf{V}_{k,k+1}\begin{pmatrix}\tilde{a}_{k,1},\ldots,\tilde{a}_{k,m},\tilde{a}_{k,m+1}\\ \tilde{q}_{k,1},\ldots,\tilde{q}_{k,m},0\end{pmatrix}.\tag{3.34}$$

Comparing kth rows on both sides of (3.34) we get an equation which can be solved for the vector $[\tilde{q}_{k,1},\ldots,\tilde{q}_{k,m}]$ as the other quantities involved are known from step $k-1$. Indeed, if

$$\mathbf{U}_{k,k+1}=\begin{pmatrix}\cos\alpha_k & \sin\alpha_k\\ -\sin\alpha_k & \cos\alpha_k\end{pmatrix},$$

$$\mathbf{V}_{k,k+1}=\begin{pmatrix}\cos\beta_k & \sin\beta_k\\ -\sin\beta_k & \cos\beta_k\end{pmatrix},$$

then $\tilde{q}_{k,j}$, $j=1,\ldots,m$, satisfies the relation

$$\hat{a}_{k,j}\cos\alpha_k+q_{k,j-1}\sin\alpha_k=\tilde{a}_{k,j}\cos\beta_k+\tilde{q}_{k,j}\sin\beta_k,\tag{3.35}$$

from which $\tilde{q}_{k,j}$ is easily determined.

Now, premultiplying (3.31) by $\mathbf{U}_{k,k+1}$ and (3.32) by $\mathbf{V}_{k,k+1}$ we obtain $[\hat{a}_{k+1,1},\ldots,\hat{a}_{k+1,m+1}]$ and $[\tilde{a}_{k+1,1},\ldots,\tilde{a}_{k+1,m+1}]$, i.e., the $(k+1)$st row of $\mathbf{U}_{k,k+1}\cdots\mathbf{U}_{1,2}\hat{\mathbf{Q}}$ and the $(k+1)$st row of $\mathbf{V}_{k,k+1}\cdots\mathbf{V}_{1,2}\tilde{\mathbf{Q}}$.

Knowing the kth row of $\mathbf{Q}_c^T$, the kth row of $\mathbf{Q}_d^T$ and the rotation $\mathbf{W}_{k,k+1}$, the $(k+1)$st row of $\mathbf{Q}_c^T$ can be determined from (3.21). Premultiplying (3.33) by $\mathbf{W}_{k,k+1}$ and comparing kth rows on both sides of (3.21), we obtain an equation similar to (3.35), from which $[q_{k+1,1},\ldots,q_{k+1,m}]$ can be easily computed. Indeed, if

$$\mathbf{W}_{k,k+1}=\begin{pmatrix}\cos\gamma_k & \sin\gamma_k\\ -\sin\gamma_k & \cos\gamma_k\end{pmatrix},$$

then $q_{k+1,j}$, $j=1,\ldots,m$, satisfies the relation

$$q_{k,j}\cos\gamma_k+q_{k+1,j}\sin\gamma_k=\tilde{q}_{k,j}\tag{3.36}$$

which uniquely determines $q_{k+1,j}$. Now knowing $[q_{k+1,1},\ldots,q_{k+1,m}]$ we can compute $[p_{k+1,1},\ldots,p_{k+1,m}]$ which will be required in the next step.

The Relation (3.34) gives us the kth row of $\mathbf{Q}_d^T$ and vectors $[\hat{a}_{k+1,1},\ldots,\hat{a}_{k+1,m+1}]$ and $[\tilde{a}_{k+1,1},\ldots,\tilde{a}_{k+1,m+1}]$, while the relation (3.36) gives us the $(k+1)$st row of $\mathbf{Q}_c^T$ and the vector $[p_{k+1,1}, \ldots, p_{k+1,m}]$. Thus we have all the data needed in step $(k+1)$. As the first row of $\mathbf{Q}_c^T$ is equal to $\mathbf{t}_1/r_{11}$, the description of the algorithm for computing $\mathbf{Q}$ is complete.

It is easy to see that computations implied by (3.30) require $8mn$ multiplications while those implied by (3.21) require $4mn$ multiplications. We conclude that our algorithms for computing $\mathbf{R}$ and $\mathbf{Q}$ require $13mn+6n^2+O(n)$ multiplications. This could be reduced by the use of "fast" Givens transformations.

4. Concluding Remarks

We have presented an algorithm for the QR factorization which is somewhat more efficient than Sweet's algorithm [6] and is able to deal with rectangular matrices. Furthermore, the implementation of the algorithm has indicated that it performs at least as well as Sweet's algorithm and is substantially more accurate on some examples. However, when the condition number of the Toeplitz matrix is large, we have observed that the decomposition obtained by our algorithm does not compare well with the decomposition obtained using standard $O(n^3)$ algorithms. In retrospect this is not surprising since the downdating of Cholesky factors (which is a key part of the algorithm) has been shown by Stewart [5] to be a poorly conditioned problem. Nevertheless, the proposed algorithm appears to compare very favourably with the other $O(n^2)$ algorithms available.

References

1. Bojanczyk, A.W., Brent, R.P., de Hoog, F.R.: Linearly Connected Arrays for Toeplitz Least Squares Problems. Technical Report CMA-R06-85, Centre for Mathematical Analysis, Australian National University 1985
2. Bunch, J.R.: Stability of Methods for Solving Toeplitz Systems of Equations. SIAM J. Sci. Stat. Comput. **6**, 349–364 (1985)
3. Gill, P.E., Golub, G.H., Murray, W., Saunders, M.A.: Methods for Modifying Matrix Factorizations. Math. Comput. **28**, 505–535 (1974)
4. Lawson, C.L., Hanson, R.J.: Solving Least Squares Problems. Englewood Cliffs, N.J.: Prentice-Hall 1974
5. Stewart, G.W.: The Effect of Rounding Error on an Algorithm for Downdating a Cholesky Factorization. J. Inst. Math. Appl. **23**, 203–213 (1979)
6. Sweet, D.R.: Fast Toeplitz Orthogonalization. Numer. Math. **43**, 1–21 (1984)

Received July 30, 1985/December 16, 1985