Report Number: WUCS-87-1

1987-02-01

The QR Decomposition of Toeplitz Matrices

Adam W. Bojanczyk

We present a new algorithm for computing the QR factorization of an mxn Toeplitz matrix in O(mn) multiplications. The algorithm exploits the procedure for the rank-1 modification and the fact that successive columns of a Toeplitz matrix are related to each other. Both matrices Q and R are generated column by column, starting for their first columns. Each column is calculated from the previous column after rank-1 modification to the matrix R and a step of Gramm-Schmidt orthogonalization process applied to two auxiliary vectors.

THE QR DECOMPOSITION OF TOEPLITZ MATRICES

Adam W. Bojanczyk

WUCS-87-1

February 1987

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899

The QR Decomposition of Toeplitz Matrices

Adam W. Bojanczyk
Department of Computer Science
Washington University
St. Louis, MO 63130

1. Abstract

We present a new algorithm for computing the QR factorization of an mxn Toeplitz matrix in O(mn) multiplications. The algorithm exploits the procedure for the rank-1 modification and the fact that succesive columns of a Toeplitz matrix are related to each other. Both matrices Q and R are generated column by column, starting from their first columns. Each column is calculated from the previous column after rank-1 modification to the matrix R and a step of Gramm-Schmidt orthogonalization process applied to two auxiliary vectors.

2. Introduction

An mxn matrix T is Toeplitz if elements on any diagonal are all equal, i.e., $t_{i,j} = t_{i-j}$, i=1,..,m, j=1,..,n. Toeplitz matrices arise in many engineering applications. For most of these applications it is required to compute the QR decomposition of the matrix T,

$$T=QR$$

where Q is an orthonormal matrix and R is upper triangular.

Methods for calculating the QR decomposition of a general rectangular matrix are known and they require $O(mn^2)$ multiplications. Because Toeplitz matrices have a very special structure, it might be expected that the QR decomposition of a Toeplitz matrix could be calculated in less multiplications than in general case. This is indeed true.

Recently, Sweet [4] has proposed an O(nm) algorithm for the QR decomposition of a rectangular Toeplitz matrix. The algorithm exploits the procedures for rank-1 modification and the fact that both principal mxn submatrices of T are identical. When the fast Givens rotations are used, Sweet's algorithm requires 9mn+5n^2+O(m+n) multiplications to compute both Q and R.

Another approach has been proposed by Cybenko [2]. Instead of computing the QR decomposition of the matrix T, one computes an *inverse decomposition*, i.e., a matrix P having orthogonal columns and U upper triangular for which

$$TU=P$$

The method uses inner products as in Gramm-Schmidt orthogonalization process. An inverse QR factorization is computed in 10nm + O(n^2) multiplications.

Yet another algorithm for QR factorization of Toeplitz matrix have been developed in [1]. The approach is similar to that used in Sweet's algorithm but is logically less complex. The algorithm requires 7nm+4n^2+O(m+n) multiplication for computing Q and R when fast Givens rotations are used.

Computer tests [3] have shown that none of the three algorithms, at least in their present form, gives satisfactory numerical results. All three algorithms do not guarantee orthogonality of the computed matrix Q.

In this paper we propose an algorithm for calculating the QR decomposition of a rectangular mxn Toeplitz matrix in $12mn + 6n^2 + O(m)$ multiplications. The algorithm can be viewed as a combination of the Gramm-Schmidt procedure and results obtained in [1].

Although the operation count is slightly higher for the new algorithm, computer tests suggest that the algorithm may give more accurate results than the other methods. More test are necessary to verify this claim.

3. The Algorithm

Consider an mxn Toeplitz matrix **T**,

$$\mathbf{T} = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdot\ \cdot & t_{-n+1} \\ t_1 & t_0 & t_{-1} & \cdot\ \cdot & t_{-n+2} \\ t_2 & t_1 & t_0 & \cdot\ \cdot & t_{-n+3} \\ \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\ \cdot & \cdot \\ t_{m-1} & t_{m-2} & t_{m-3} & \cdot\ \cdot & t_{m-n} \end{bmatrix}$$

Let $\mathbf{x}_k$ denote the kth column of **T**. Define recursively matrices $\mathbf{X}_k$ and $\mathbf{Y}_k$ as

$$\mathbf{X}_1 := \mathbf{x}_1, \quad \mathbf{X}_k := [\mathbf{X}_{k-1}, \mathbf{x}_k]$$

$$\mathbf{Y}_1 := \mathbf{x}_2, \quad \mathbf{Y}_k := [\mathbf{Y}_{k-1}, \mathbf{x}_{k+1}]$$

Because **T** is Toeplitz, $\mathbf{X}_k$ and $\mathbf{Y}_k$ can be partitioned as follows

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{T}_k \\ \mathbf{a}_k^T \end{bmatrix}, \qquad \mathbf{a}_k^T = [t_{m-1}, \ \cdots \ , t_{m-k}]$$

$$\mathbf{Y}_k = \begin{bmatrix} \mathbf{b}_k^T \\ \mathbf{T}_k \end{bmatrix}, \qquad \mathbf{b}_k^T = [t_{-1}, \ \cdots \ , t_{-k}]$$

$$\mathbf{T}_k = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & . & . & t_{-k+1} \\ t_1 & t_0 & t_{-1} & . & . & t_{-n+2} \\ t_2 & t_1 & t_0 & . & . & t_{-n+3} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ t_{m-2} & t_{m-3} & t_{m-4} & . & . & t_{m-k} \end{bmatrix}$$

We will also consider an augmented matrix $\mathbf{S}_{k-1}$ defined as

$$\mathbf{S}_{k-1} = \begin{bmatrix} \mathbf{b}_{k-1}^T \\ \mathbf{T}_{k-1} \\ \mathbf{a}_{k-1}^T \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{k-1}^T \\ \mathbf{X}_{k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{k-1} \\ \mathbf{a}_{k-1}^T \end{bmatrix} \tag{3.1}$$

Assume that the QR factorizations of $\mathbf{X}_{k-1}$ and $\mathbf{Y}_{k-1}$ are known,

$$\mathbf{X}_{k-1} = [\mathbf{x}_1, \mathbf{x}_2, \ . \ . \ . \ , \mathbf{x}_{k-1}] = \begin{bmatrix} \mathbf{T}_{k-1} \\ \mathbf{a}_{k-1}^T \end{bmatrix} = \mathbf{Q}_{a_{k-1}} \mathbf{R}_{a_{k-1}} =$$

$$[\mathbf{q}_1, \mathbf{q}_2, \ . \ . \ . \ , \mathbf{q}_{k-1}] \begin{bmatrix} r_{11}^a & r_{12}^a & . & . & r_{1k-1}^a \\ . & r_{22}^a & . & . & r_{2k-1}^a \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & r_{k-1k-1}^a \end{bmatrix} \tag{3.2}$$

and

$$\mathbf{Y}_{k-1} = [\mathbf{x}_2, \mathbf{x}_3, \ . \ . \ . \ , \mathbf{x}_k] = \begin{bmatrix} \mathbf{b}_{k-1}^t \\ \mathbf{T}_{k-1} \end{bmatrix} = \mathbf{Q}_{b_{k-1}} \mathbf{R}_{b_{k-1}} =$$

$$[\mathbf{p}_1,\mathbf{p}_2,\ \cdots,\mathbf{p}_{k-1}]\begin{bmatrix} r^b_{11} & r^b_{12} & \cdot & \cdot & r^b_{1k-1} \\ \cdot & r^b_{22} & \cdot & \cdot & r^b_{2k-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & r^b_{k-1k-1} \end{bmatrix} \tag{3.3}$$

From (3.1)-(3.3) we have

$$\mathbf{S}_{k-1} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_{a_{k-1}} \end{bmatrix}\begin{bmatrix} \mathbf{b}^T_{k-1} \\ \mathbf{R}_{a_{k-1}} \end{bmatrix} \tag{3.4}$$

and

$$\mathbf{S}_{k-1} = \begin{bmatrix} \mathbf{0} & \mathbf{Q}_{b_{k-1}} \\ 1 & \mathbf{0}^T \end{bmatrix}\begin{bmatrix} \mathbf{a}^T_{k-1} \\ \mathbf{R}_{b_{k-1}} \end{bmatrix} \tag{3.5}$$

Let $G(\theta_{k-1}),G(\theta_{k-2}),...,G(\theta_1)$ be a sequence of plane rotations which triangularizes the upper Hessenberg matrix on the right hand side of (3.4),

$$G(\theta_{k-1})G(\theta_{k-2})...G(\theta_1)\begin{bmatrix} \mathbf{b}^T_{k-1} \\ \mathbf{R}_{a_{k-1}} \end{bmatrix} = \begin{bmatrix} \mathbf{R}^{\theta}_{a_{k-1}b_{k-1}} \\ \mathbf{0}^T \end{bmatrix} \tag{3.6}$$

Similarly, let $G(\omega_{k-1}),G(\omega_{k-2}),...,G(\omega_1)$ be a sequence of plane rotations which triangularizes the upper Hessenberg matrix on the right hand side of (3.5),

$$G(\omega_{k-1})G(\omega_{k-2})...G(\omega_1)\begin{bmatrix} \mathbf{a}^T_{k-1} \\ \mathbf{R}_{b_{k-1}} \end{bmatrix} = \begin{bmatrix} \mathbf{R}^{\omega}_{a_{k-1}b_{k-1}} \\ \mathbf{0}^T \end{bmatrix} \tag{3.7}$$

From the uniqness of the QR factorization of $\mathbf{S}_{k-1}$ the following relations hold,

$$\mathbf{R}^{\omega}_{a_{k-1}b_{k-1}} = \mathbf{R}^{\theta}_{a_{k-1}b_{k-1}} =: \mathbf{R}_{a_{k-1}b_{k-1}}$$

and

$$\begin{bmatrix} 1 & 0 & . & . & . & 0 \\ \mathbf{0} & \mathbf{q}_1 & . & . & . & \mathbf{q}_{k-1} \end{bmatrix} G^T(\theta_1)G^T(\theta_2)\ \cdots\ G^T(\theta_{k-1})\begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} = \tag{3.8}$$

$$\begin{bmatrix} \mathbf{0} & \mathbf{p}_1 & . & . & . & \mathbf{p}_{k-1} \\ 1 & 0 & . & . & . & 0 \end{bmatrix} G^T(\omega_1)G^T(\omega_2)\ \cdots\ G^T(\omega_{k-1})\begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}$$

where $\mathbf{I}_{k-1}$ is the (k-1)-dimensional identity matrix. The relation (3.8) states that only the first k-1 columns on both sides of (3.8) are equal. This is because $\mathbf{R}_{a_{k-1}b_{k-1}}$ has rank k-1.

Define an (m+1)x(k-1) matrix $\mathbf{V}_{k-1}$ and a vector $\overline{\mathbf{v}}_k$ as

$$\begin{bmatrix} 1 & 0 & . & . & . & 0 \\ 0 & \mathbf{q}_1 & . & . & . & \mathbf{q}_{k-1} \end{bmatrix} G^T(\theta_1)G^T(\theta_2) \cdots G^T(\theta_{k-1}) = \tag{3.9}$$

$$\left[\mathbf{v}_1,\mathbf{v}_2, \cdots ,\mathbf{v}_{k-1},\overline{\mathbf{v}}_k\right] = \left[\mathbf{V}_{k-1},\overline{\mathbf{v}}_k\right]$$

Note that

$$\begin{bmatrix} \mathbf{0} & \mathbf{p}_1 & . & . & . & \mathbf{p}_{k-1} \\ 1 & 0 & . & . & . & 0 \end{bmatrix} G^T(\omega_1)G^T(\omega_2) \cdots G^T(\omega_{k-1}) = \tag{3.10}$$

$$\left[\mathbf{v}_1,\mathbf{v}_2, \cdots ,\mathbf{v}_{k-1},\overline{\mathbf{z}}_k\right] = \left[\mathbf{V}_{k-1},\overline{\mathbf{z}}_k\right]$$

where $\overline{\mathbf{z}}_k$, in general, is not equal to $\overline{\mathbf{v}}_k$.

The rest of this Section is divided into two parts. In the first part we show how to obtain the QR decomposition of $\mathbf{X}_k$ from the QR decomposition of $\mathbf{X}_{k-1}$ and the QR decomposition of $\mathbf{Y}_{k-1}$. In the second part we show how to obtain the QR decomposition of $\mathbf{Y}_k$ from the QR decompositions of $\mathbf{X}_k$,$\mathbf{S}_{k-1}$ and $\mathbf{Y}_{k-1}$. All quantities computed in step k-1 are assumed to be known before the execution of step k starts.

3.1 Computation of $\mathbf{Q}_{a_k}$ and $\mathbf{R}_{a_k}$

From (3.2) and (3.3) we have

$$\mathbf{X}_k = [\mathbf{q}_1,\mathbf{p}_1, \ldots ,\mathbf{p}_{k-1}] \begin{bmatrix} r_{11}^a & \mathbf{0}^T \\ \mathbf{0} & \mathbf{R}_{b_{k-1}} \end{bmatrix} \tag{3.11}$$

We can make $\mathbf{q}_1$ orthogonal to the space $Lin(\mathbf{p}_1,\mathbf{p}_2, \ldots ,\mathbf{p}_{k-1})$ using, for example, the Gramm-Schmidt procedure,

$$\mathbf{u}_0 := \mathbf{q}_1$$

for i=1,2,...,k-1:

$$\mathbf{u}_i := \mathbf{u}_{i-1} + \mu_i\mathbf{p}_i \qquad \mu_i = -(\mathbf{u}_{i-1},\mathbf{p}_i) \tag{3.12}$$

By construction $\mathbf{u}_i$ is orthogonal to $Lin(\mathbf{p}_1, \ldots, \mathbf{p}_{i-1})$. Clearly $\mathbf{u}_{k-1}$ and μ_{k-1} can be generated recursively from $\mathbf{u}_{k-2}$ and $\mathbf{p}_{k-1}$.

Based on (3.12) we can express $\mathbf{q}_1$ in terms of $\mathbf{p}'s$ and $\mathbf{u}_{k-1}$,

$$\mathbf{q}_1 = \mathbf{u}_{k-1} - (\mu_1\mathbf{p}_1 + \ldots + \mu_{k-1}\mathbf{p}_{k-1}) \tag{3.13}$$

Moreover, from orthogonality of $\mathbf{p}'s$ and $\mathbf{u}_{k-1}$,

$$||\mathbf{q}_1||^2 = ||\mathbf{u}_{k-1}||^2 + \mu_1^2 + \ldots + \mu_{k-1}^2 \tag{3.14}$$

Let $\mathbf{m}_{k-1} = [\mu_1, \cdots, \mu_{k-1}]^T$. Combining (3.11) and (3.13) we obtain

$$\mathbf{X}_k = [\mathbf{p}_1, \ldots, \mathbf{p}_{k-1}, \mathbf{u}_{k-1}] \begin{bmatrix} -\mathbf{m}_{k-1} & \mathbf{I} \\ 1 & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} r_{11}^a & \mathbf{0}^T \\ \mathbf{0} & \mathbf{R}_{b_{k-1}} \end{bmatrix} \tag{3.15}$$

Define

$$\bar{\mathbf{u}}_{k-1} = \frac{\mathbf{u}_{k-1}}{||\mathbf{u}_{k-1}||} \tag{3.16}$$

then (3.15) can be rewritten as

$$\mathbf{X}_k = [\mathbf{p}_1, \ldots, \mathbf{p}_{k-1}, \bar{\mathbf{u}}_{k-1}] \begin{bmatrix} -r_{11}^a \mathbf{m}_{k-1} & \mathbf{R}_{b_{k-1}} \\ r_{11}^a ||\mathbf{u}_{k-1}|| & \mathbf{0}^T \end{bmatrix} \tag{3.17}$$

Note that the matrix $[\mathbf{p}_1, \ldots, \mathbf{p}_{k-1}, \bar{\mathbf{u}}_{k-1}]$ has orthonormal columns.

Let $G(\gamma_1)G(\gamma_k) \cdots G(\gamma_{k-1})$ be a sequence of plane rotations such that

$$G(\gamma_1)G(\gamma_2) \cdots G(\gamma_{k-1}) \begin{bmatrix} -r_{11}^a \mathbf{m}_{k-1} & \mathbf{R}_{b_{k-1}} \\ r_{11}^a ||\mathbf{u}_{k-1}|| & \mathbf{0}^T \end{bmatrix} \tag{3.18}$$

is upper triangular.

From the uniqueness of the QR decomposition of $\mathbf{X}_k$ and (3.17)-(3.18) we conclude that

$$[\mathbf{p}_1, \ . \ . \ . \ , \mathbf{p}_{k-1}, \bar{\mathbf{u}}_{k-1}] G^T(\gamma_{k-1}) G^T(\gamma_{k-2}) \cdot \cdot \cdot \ G^T(\gamma_1) = [\mathbf{q}_1, \ . \ . \ . \ , \mathbf{q}_k] = \mathbf{Q}_{a_k} \qquad \mathbf{(3.20)}$$

and

$$G(\gamma_1) G(\gamma_2) \cdot \cdot \cdot \ G(\gamma_{k-1}) \begin{bmatrix} -r_{11}^a \mathbf{m}_{k-1} & \mathbf{R}_{b_{k-1}} \\ r_{11}^a \, ||\mathbf{u}_{k-1}|| & \mathbf{0}^T \end{bmatrix} = \mathbf{R}_{a_k} \qquad (3.21)$$

Note that by (3.14)

$$||\mathbf{u}_k||^2 + \mu_k = ||\mathbf{u}_{k-1}||^2$$

Thus $G(\gamma_j)$, $j \leq k-2$, generated in step k-1 are the same as those generated in step k. Hence, in each step k we have to compute only one additional rotation $G(\gamma_k)$.

From (3.20) it follows that

$$[\mathbf{p}_{k-1}, \bar{\mathbf{u}}_{k-1}] \mathbf{G}^T(\gamma_{k-1}) = [\mathbf{z}, \mathbf{q}_k] \qquad (3.22)$$

Thus $\mathbf{q}_k$ can be generated from $\mathbf{p}_{k-1}, \bar{\mathbf{u}}_{k-1}$ and $G(\gamma_{k-1})$; the vector $\mathbf{z}$ is of no importance in the further considerations.

Comparison of the last columns on both sides of (3.21) yields

$$G(\gamma_1) G(\gamma_2) \cdot \cdot \cdot \ G(\gamma_{k-1}) \begin{bmatrix} r_{1k-1}^b \\ . \\ . \\ r_{k-1k-1}^b \\ 0 \end{bmatrix} = \begin{bmatrix} r_{1k}^a \\ . \\ . \\ r_{k-1k}^a \\ r_{kk}^a \end{bmatrix} \qquad (3.23)$$

The relations (3.22) and (3.23) give us a means for generating the QR decomposition of $\mathbf{X}_k$ from the QR decomposition of $\mathbf{X}_{k-1}$ and the QR decomposition of $\mathbf{Y}_{k-1}$.

3.2 Computation of $\mathbf{Q}_{b_k}$ and $\mathbf{R}_{b_k}$

In the sequel we develop a procedure for computing the QR decomposition of $\mathbf{Y}_k$ from the QR decomposition of $\mathbf{X}_k$ and the QR decompositions of $\mathbf{Y}_{k-1}$ and $\mathbf{S}_{k-1}$.

Recall that the augmented matrix $\mathbf{S}_{k-1}$ is defined as follows

$$\mathbf{S}_{k-1} = \begin{bmatrix} \mathbf{b}_{k-1}^T \\ \mathbf{T}_{k-1} \\ \mathbf{a}_{k-1}^T \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{k-1}^T \\ \mathbf{X}_{k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{k-1} \\ \mathbf{a}_{k-1}^T \end{bmatrix}$$

Assume by induction that $\mathbf{R}_{a_{k-1}b_{k-1}}$ and $[\mathbf{V}_{k-1},\bar{\mathbf{v}}_k]$ defined by (3.7) and (3.9) are known.

Consider the augmented matrix $\mathbf{S}_k$. We have

$$\mathbf{S}_k = \begin{bmatrix} \mathbf{b}_k^T \\ \mathbf{X}_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{Q}_{a_k} \end{bmatrix} \begin{bmatrix} \mathbf{b}_k^T \\ \mathbf{R}_{a_k} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{Q}_{a_k} \end{bmatrix} \begin{bmatrix} \mathbf{b}_{k-1}^T & t_{-k} \\ \mathbf{R}_{a_{k-1}} & \mathbf{r}_{.k}^a \end{bmatrix} \tag{3.26}$$

where $\mathbf{r}_{.k}^a = [r_{1k}^a, r_{2k}^a, \; . \; . \; . \; , r_{kk}^a]^T$ is the last column of the matrix $\mathbf{R}_{a_k}$.

From (3.6)

$$G(\theta_{k-1})G(\theta_{k-2})...G(\theta_1) \begin{bmatrix} \mathbf{b}_{k-1}^T & t_{-k} \\ \mathbf{R}_{a_{k-1}} & \mathbf{r}_{.k}^a \end{bmatrix} = \tag{3.27}$$

$$\begin{bmatrix} \mathbf{R}_{a_{k-1}b_{k-1}} & \mathbf{r}_{1:k-1,k}^{ab} \\ \mathbf{0}^T & \bar{r}_{kk}^{ab} \\ \mathbf{0}^T & r_{kk}^a \end{bmatrix}$$

where $\mathbf{r}_{1:k-1,k}^{ab}$ denotes elements 1 thru k-1 of the kth column of $\mathbf{R}_{a_k b_k}$. Rewriting the relation (3.27) for the last column gives

$$G(\theta_{k-1})G(\theta_{k-2})...G(\theta_1) \begin{bmatrix} \beta_{k+1} \\ r_{1k}^a \\ . \\ . \\ r_{kk}^a \end{bmatrix} = \begin{bmatrix} r_{1k}^{ab} \\ . \\ r_{k-1k}^{ab} \\ \bar{r}_{kk}^{ab} \\ r_{kk}^a \end{bmatrix} \tag{3.28}$$

The rotation $G(\theta_k)$ which zeroes the bottom element r_{kk}^a and produces r_{kk}^{ab} can now be easily determined from $\bar{r}_{kk}^{ab}$ and r_{kk}^a. The relation (3.28) gives us a means for generating $\mathbf{r}_{.k}^{ab}$, the last column of the matrix $\mathbf{R}_{a_k b_k}$.

We also have

$$\begin{bmatrix} 1 & 0 & . & . & . & 0 \\ \mathbf{0} & \mathbf{q}_1 & . & . & . & \mathbf{q}_k \end{bmatrix} G^T(\theta_1)G^T(\theta_2) \cdots G^T(\theta_k) =$$

$$\begin{bmatrix} \mathbf{v}_1,\mathbf{v}_2, \cdots ,\mathbf{v}_{k-1},\overline{\mathbf{v}}_k,\hat{\mathbf{q}}_k \end{bmatrix} G^T(\theta_k) = \begin{bmatrix} \mathbf{V}_k,\overline{\mathbf{v}}_{k+1} \end{bmatrix}$$

where $\hat{\mathbf{q}}_k^T = [0,\mathbf{q}_k^T]$. For the last two columns in the above relation we have

$$[\overline{\mathbf{v}}_k,\hat{\mathbf{q}}_k]G^T(\theta_k) = [\mathbf{v}_k,\overline{\mathbf{v}}_{k+1}] \tag{3.29}$$

We will now derive similar relations for $\mathbf{Q}_{b_k}$ and $\mathbf{R}_{b_k}$.

From (3.5)

$$\mathbf{S}_k = \begin{bmatrix} \mathbf{0} & \mathbf{p}_1 & . & . & \mathbf{p}_k \\ 1 & 0 & . & . & 0 \end{bmatrix} \begin{bmatrix} \mathbf{a}_k^T \\ \mathbf{R}_{b_k} \end{bmatrix}$$

The sequence of plane rotations $G(\omega_k),G(\omega_{k-1}), \cdots ,G(\omega_1)$ transforms the upper Hessenberg matrix on the right hand side to the upper triangular matrix $\mathbf{R}_{a_k b_k}$,

$$G(\omega_k)G(\omega_{k-1}) \cdots G(\omega_1) \begin{bmatrix} \mathbf{a}_k^T \\ \mathbf{R}_{b_k} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{a_k b_k} \\ \mathbf{0}^T \end{bmatrix} \tag{3.30}$$

$G(\omega_k)G(\omega_{k-1}) \cdots G(\omega_1)$ are known from step k-1. We want to find $G(\omega_k)$ and $\mathbf{r}_{.k}^b$, the last column of $\mathbf{R}_{b_k}$.

Recall from (3.8) that

$$\begin{bmatrix} \mathbf{0} & \mathbf{p}_1 & . & . & \mathbf{p}_k \\ 1 & 0 & . & . & 0 \end{bmatrix} G^T(\omega_1)G^T(\omega_2)...G^T(\omega_k) \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & . & . & . & 0 \\ \mathbf{0} & \mathbf{q}_1 & . & . & . & \mathbf{q}_k \end{bmatrix} G^T(\theta_1)G^T(\theta_2)...G^T(\theta_k) \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} =$$

or using notation (3.25) and (3.29),

$$\begin{bmatrix} \mathbf{v}_1, \; . \; . \; . \; ,\mathbf{v}_k,\overline{\mathbf{v}}_{k+1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0}^t & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1, \; . \; . \; . \; ,\mathbf{v}_{k-1},\overline{\mathbf{z}}_k,\hat{\mathbf{p}}_k \end{bmatrix} G^T(\omega_k) \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0}^t & 0 \end{bmatrix} \tag{3.31}$$

where $\hat{\mathbf{p}}_k^T = [\mathbf{p}_k,0]$ is being sought as well.

Comparing the kths columns on both sides of (3.31) we obtain

$$c(\omega_k)\bar{\mathbf{z}}_k + s(\omega_k)\begin{bmatrix}\mathbf{p}_k \\ 0\end{bmatrix} = \mathbf{v}_k \tag{3.32}$$

where $c(\omega_k)$ and $s(\omega_k)$ are cosine and sine defining $G(\omega_k)$. From the relation above it is clear that

$$c(\omega_k) = \frac{v_{mk}}{\bar{z}_{mk}} \tag{3.33}$$

where v_{mk} and $\bar{z}_{mk}$ are the last components of $\mathbf{v}_k$ and $\bar{\mathbf{z}}_k$ respectively.

In order to compute $s(\omega_k)$ note that $\hat{\mathbf{p}}_k$ is in $Lin(\bar{\mathbf{z}}_k,\mathbf{v}_k)$ and is orthogonal to $\bar{\mathbf{z}}_k$. By Gramm-Schmidt procedure, the vector $\bar{\mathbf{p}}_k$ defined by

$$\bar{\mathbf{p}}_k = \mathbf{v}_k + \nu_k\bar{\mathbf{z}}_k \ , \qquad \nu_k = -(\mathbf{v}_k,\bar{\mathbf{z}}_k) \tag{3.34}$$

has the same direction as $\hat{\mathbf{p}}_k = [\mathbf{p}_k^T,0]^T$. Thus

$$\begin{bmatrix}\mathbf{p}_k \\ 0\end{bmatrix} = \frac{\bar{\mathbf{p}}_k}{||\bar{\mathbf{p}}_k||}$$

Comparing (3.32) and (3.34) we see that

$$c(\omega_k) = -\nu_k, \qquad s(\omega_k) = \frac{1}{||\bar{\mathbf{p}}_k||} \tag{3.35}$$

Note that (3.33) is another alternative for computing $c(\omega_k)$.

Knowing $G(\omega_k)$ it is now straightforward to compute $\mathbf{r}^b_{.k}$ from the relation (3.30) as

$$\begin{bmatrix} t_{m-k} \\ r^b_{1k} \\ \cdot \\ \cdot \\ r^b_{kk} \end{bmatrix} = G^T(\omega_1)\cdots G^T(\omega_k)\begin{bmatrix} r^{ab}_{1k} \\ \cdot \\ \cdot \\ r^{ab}_{kk} \\ 0 \end{bmatrix} \tag{3.36}$$

Also we can compute $\bar{\mathbf{z}}_{k+1}$ from (3.31) as

$$[\bar{\mathbf{z}}_k,\hat{\mathbf{p}}_k]G^T(\omega_k) = [\mathbf{v}_k,\bar{\mathbf{z}}_{k+1}] \tag{3.37}$$

which completes the derivation of the kth step of the recursive procedure. An outline of the algorithm together with costs of individual operations are given in Appendix.

Appendix

For completeness we give a brief outline of the kth step of the recursive algorithm.

Algorithm

Assume that all relevant quantities computed in step k-1 are known. In step k do:

1) compute $\mathbf{u}_{k-1}$ and μ_{k-1} (see (3.12) and (3.16))

$$\mathbf{u}_{k-1} = \mathbf{u}_{k-2} + \mu_{k-1}\mathbf{p}_{k-1} \quad \text{where} \quad \mu_{k-1} = -(\mathbf{u}_{k-2},\mathbf{p}_{k-1})$$

$$\bar{\mathbf{u}}_{k-1} = \frac{\mathbf{u}_{k-1}}{||\mathbf{u}_{k-1}||}$$

cost: 2m + O(1) multiplications

2) determine $c(\gamma_{k-1})$ and $s(\gamma_{k-1})$, parameters of the rotation $G(\gamma_{k-1})$, from the following relation (see (3.21))

$$\begin{bmatrix} c(\gamma_{k-1}) & s(\gamma_{k-1}) \\ -s(\gamma_{k-1}) & c(\gamma_{k-1}) \end{bmatrix} \begin{bmatrix} -\mu_{k-1} \\ ||\mathbf{u}_{k-1}|| \end{bmatrix} = \begin{bmatrix} ||\mathbf{u}_{k-1}|| \\ 0 \end{bmatrix}$$

cost: O(1) multiplications

3) compute the vector $\mathbf{q}_k$ (see (3.22))

$$[\mathbf{z},\mathbf{q}_k] = [\mathbf{p}_{k-1},\bar{\mathbf{u}}_{k-1}]G^T(\gamma_{k-1})$$

cost: 2m multiplications

4) compute the kth column of $\mathbf{R}_{a_k}$ (see (3.23))

$$G(\gamma_1)G(\gamma_2)\cdots G(\gamma_{k-1})\begin{bmatrix} r^b_{1k-1} \\ \cdot \\ r^b_{k-1k-1} \\ 0 \end{bmatrix} = \begin{bmatrix} r^a_{1k} \\ \cdot \\ r^a_{k-1k} \\ r^a_{kk} \end{bmatrix}$$

cost: 4k + O(1) multiplications

5) determine the parameters of the rotation $G(\theta_k)$ and the kth column of $\mathbf{R}_{a_k b_k}$ (see (3.28))

$$G(\theta_{k-1})G(\theta_{k-2})...G(\theta_1)\begin{bmatrix} t_{-k} \\ r^a_{1k} \\ \cdot \\ \cdot \\ r^a_{kk} \end{bmatrix} = \begin{bmatrix} r^{ab}_{1k} \\ \cdot \\ r^{ab}_{k-1k} \\ \bar{r}^{ab}_{kk} \\ r^a_{kk} \end{bmatrix}$$

$$\begin{bmatrix} c(\theta_k) & s(\theta_k) \\ -s(\theta_k) & c(\theta_k) \end{bmatrix}\begin{bmatrix} \bar{r}_{kk} \\ r^a_{kk} \end{bmatrix} = \begin{bmatrix} r^{ab}_{kk} \\ 0 \end{bmatrix}$$

cost: 4k + O(1) multiplications

6) find the vectors $\mathbf{v}_k$ and $\bar{\mathbf{v}}_{k+1}$ (see (3.29))

$$[\mathbf{v}_k,\bar{\mathbf{v}}_{k+1}] = [\bar{\mathbf{v}}_k,\hat{\mathbf{q}}_k]G^T(\theta_k)$$

cost: 4m multiplications

7) compute the vector $\mathbf{p}_k$ and $c(\omega_k)$,$s(\omega_k)$, the parameters of the rotation $G(\omega_k)$ (see (3.34) and (3.35)

$$\bar{\mathbf{p}}_k = \mathbf{v}_k + \nu_k \bar{\mathbf{z}}_k \qquad \text{where} \quad \nu_k = -(\mathbf{v}_k,\bar{\mathbf{z}}_k)$$

$$\begin{bmatrix} \mathbf{p}_k \\ 0 \end{bmatrix} = \frac{\bar{\mathbf{p}}_k}{||\bar{\mathbf{p}}_k||}, \qquad c(\omega_k) = -\nu_k, \quad s(\omega_k) = \frac{1}{||\bar{\mathbf{p}}_k||}$$

cost: 2m + O(1) multiplications

8) compute $\mathbf{r}^b_{.k}$, the kth column of $\mathbf{R}_{b_k}$ (see (3.36))

$$\begin{bmatrix} t_{m-k} \\ r^b_{1k} \\ \cdot \\ \cdot \\ r^b_{kk} \end{bmatrix} = G^T(\omega_1) \cdots G^T(\omega_k) \begin{bmatrix} r^{ab}_{1k} \\ \cdot \\ \cdot \\ r^{ab}_{kk} \\ 0 \end{bmatrix}$$

cost: 4k multiplications

9) find the vector $\overline{\mathbf{z}}_{k+1}$ (see (3.37))

$$[\mathbf{v}_k, \overline{\mathbf{z}}_{k+1}] = [\overline{\mathbf{z}}_k, \hat{\mathbf{p}}_k] G^T(\omega_k)$$

cost: 2m multiplications

The cost of step k is $12m + 12k$ multiplications. Thus the overall cost of the algorithm is $12mn + 6n^2$ multiplications. Some additional saving would be possible if the fast rotation were used. However, for simplicity of exposition we did not include this variant of the algorithm.

References

[1] Bojanczyk, A.W., Brent, R.P. and de Hoog, F.R., Linearly Connected Arrays for Toeplitz Least Squares Problems, *to appear in Journal of Parallel and Distributed Computing.*

[2] Cybenko, G., Fast Toeplitz Orthogonalization Using Inner Products, Tech. Report, Computer Science Dept., Tufts University, 1985.

[3] Bojanczyk, A.W., Evaluation of Fast Orthogonal Techniques for Triangularization of Toeplitz Matrices, *in preparation.*

[4] Sweet, D.R., Fast Toeplitz Orthogonalization, *Numer. Math.*, 43, pp. 1-21, 1984.