An implicit Q-theorem for Hessenberg-like matrices.

Raf Vandebril
Marc Van Barel
Nicola Mastronardi

Report TW 394, June 2004

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

An implicit Q-theorem for Hessenberg-like matrices.

Raf Vandebril
Marc Van Barel
Nicola Mastronardi

Report TW 394, June 2004

Department of Computer Science, K.U.Leuven

Abstract

The implicit Q-theorem for Hessenberg matrices is a widespread and powerful theorem. It is used in the development of for example implicit QR-algorithms to compute the eigendecomposition of Hessenberg matrices. Moreover it can also be used to prove the essential uniqueness of orthogonal similarity transformations of matrices to Hessenberg form. The theorem is also valid for symmetric tridiagonal matrices, proving thereby also in the symmetric case its power.

Currently there is a growing interest to so-called semiseparable matrices. These matrices can be considered as the inverses of tridiagonal matrices. In a similar way, one can consider Hessenberg-like matrices as the inverses of Hessenberg matrices.

In this paper, we formulate and prove an implicit Q-theorem for Hessenberg-like matrices. Similarly, like in the Hessenberg case the notion of unreduced Hessenberg-like matrices is introduced and also a method for transforming matrices via orthogonal transformations to this form is proposed. Moreover, as the theorem is valid for Hessenberg-like matrices it is also valid for symmetric semiseparable matrices.

Keywords : inverse of Hessenberg matrices, semiseparable matrices, implicit Q-theorem, Hessenberg-like matrices.
AMS(MOS) Classification : Primary : 15A18, Secondary : 65F15.

An implicit Q-theorem for Hessenberg-like matrices. ⋆

Raf Vandebril [a,1] Marc Van Barel [a,2] Nicola Mastronardi [b,3]

[a] *Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, 3001 Heverlee, Belgium*

[b] *Istituto per le Applicazioni del Calcolo "M. Picone", sez. Bari*

Abstract

The implicit Q-theorem for Hessenberg matrices is a widespread and powerful theorem. It is used in the development of for example implicit QR-algorithms to compute the eigendecomposition of Hessenberg matrices. Moreover it can also be used to prove the essential uniqueness of orthogonal similarity transformations of matrices to Hessenberg form. The theorem is also valid for symmetric tridiagonal matrices, proving thereby also in the symmetric case its power.

Currently there is a growing interest to so-called semiseparable matrices. These matrices can be considered as the inverses of tridiagonal matrices. In a similar way, one can consider Hessenberg-like matrices as the inverses of Hessenberg matrices.

In this paper, we formulate and prove an implicit Q-theorem for Hessenberg-like matrices. Similarly, like in the Hessenberg case the notion of unreduced Hessenberg-like matrices is introduced and also a method for transforming matrices via orthogonal transformations to this form is proposed. Moreover, as the theorem is valid for Hessenberg-like matrices it is also valid for symmetric semiseparable matrices.

Key words: inverse of Hessenberg matrices, semiseparable matrices, implicit Q-theorem, Hessenberg-like matrices

1 Introduction

Nowadays there is a growing interest to the computation of the eigenvalues of semiseparable matrices. Several different approaches are possible. There exist divide and conquer methods [1,2], methods for transforming semiseparable matrices to tridiagonal or banded form and hence compute the eigenvalues via these resulting matrices [3,4], and also QR-methods applicable directly on semiseparable and related matrices exist [5,6,7,8]. For the development of implicit QR-algorithms suitable for semiseparable matrices an implicit Q-theorem, just like in the tridiagonal case, would be a valuable tool.

Recently also techniques were designed for transforming arbitrary symmetric matrices to semiseparable form via orthogonal similarity transformations. The same technique can also be used for transforming not necessarily symmetric matrices to Hessenberg-like form (see [9]) via orthogonal similarity transformations. As the matrices resulting from these orthogonal transformations are semiseparable or Hessenberg-like, an implicit Q-theorem would make it possible to prove the essential uniqueness of the upper left $k \times k$ block of the resulting semiseparable or Hessenberg-like matrix. Here k denotes the unreduced number, which will be defined in the paper. Essentially unique means that only the signs of the elements of the matrices can differ.

In [5] an implicit Q-theorem for symmetric semiseparable matrices was proposed. The implicit Q-theorem proposed in this paper differs significantly from the one in [5]. The authors of [5] restrict their implicit Q-theorem to the class of symmetric semiseparable matrices, while the theorem presented in this paper is also valid for nonsymmetric Hessenberg-like matrices and thereby also covering the class of symmetric semiseparable matrices as a special case. Moreover there are some other small differences, but with important consequences. We will restrict ourselves now to the class of semiseparable matrices to make clear

⋆ The research was partially supported by the Research Council K.U.Leuven, project OT/00/16 (SLAP: Structured Linear Algebra Package), by the Fund for Scientific Research–Flanders (Belgium), projects G.0078.01 (SMA: Structured Matrices and their Applications), G.0176.02 (ANCILA: Asymptotic aNalysis of the Convergence behavior of Iterative methods in numerical Linear Algebra), G.0184.02 (CORFU: Constructive study of Orthogonal Functions) and G.0455.0 (RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation), and by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, project IUAP V-22 (Dynamical Systems and Control: Computation, Identification & Modelling). The scientific responsibility rests with the authors.
[1] Email:raf.vandebril@cs.kuleuven.ac.be
[2] Email:marc.vanbarel@cs.kuleuven.ac.be
[3] Email:irmanm21@area.ba.cnr.it

also the other differences between this paper and the paper [5].

The class of semiseparable matrices considered in [5] equals the class of generator representable semiseparable matrices as defined in this paper. As will be mentioned in the paper, the class of generator representable semiseparable matrices is a subset of the class of semiseparable matrices as defined here. Briefly one can say that (for invertible matrices) the inverse of a generator representable semiseparable matrix is an irreducible tridiagonal matrix, whereas the inverse of a semiseparable matrix is also tridiagonal but not necessarily irreducible. Hence, the implicit Q-theorem as presented here is not restricted to the class of generator representable semiseparable matrices, which is the case in [5]. Another point of difference between the two papers, is the connection with the tridiagonal matrices. The proof as presented in [5] is based on partial inversion of the involved semiseparable matrices. As these inverses are irreducible tridiagonal matrices, the authors use the implicit Q-theorem for tridiagonal matrices to base their proof on. In the proof as presented here, no use is made of invertibility. The proof is constructed by induction on the columns of the orthogonal matrices. It is constructed in a similar way as the proof of the implicit Q-theorem for Hessenberg matrices [10]. In this way the proof stands on its own and no initial knowledge on tridiagonal matrices is needed. The results however as presented in [5] were interesting and helped us in developing the version of the implicit Q-theorem as presented in this paper.

The paper is organized as follows. In Section 2 we define Hessenberg-like matrices. As unreduced matrices play an important role in developing for example implicit QR-algorithms, we define them in Section 3 investigate some of its properties and propose an orthogonal similarity transformation to reduce matrices to this unreduced form. In Section 4 we finally propose and prove the implicit Q-theorem.

2 Hessenberg-like matrices

Let us firstly define what is meant with a Hessenberg-like matrix. We use Matlab [4] style notation. With $Z(i:n;1:i)$ the submatrix of the matrix Z is denoted consisting of the rows $i, i+1, \ldots, n$ and the columns $1, 2, \ldots i$.

Definition 1 *A matrix Z of dimension $n \times n$ is called a Hessenberg-like matrix if it satisfies the following rank conditions:*

$$\operatorname{rank}(Z(i:n;1:i))) \leq 1 \qquad \forall i \in \{1, \ldots, n\}.$$

[4] Matlab is a registered trademark of the Mathworks inc.

Briefly spoken this means that all the subblocks taken out of the lower triangular part of this matrix (including the diagonal) have rank at most 1. Note however, that the lower triangular part is not necessarily coming from a rank 1 matrix as the following example illustrates.

Example 2 *The matrix A is a Hessenberg-like matrix, but the lower triangular part is not coming from a rank* 1 *matrix:*

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

We remark that this is a common mistake. More information about frequent mistakes concerning semiseparable matrices can be found in [11].

Hessenberg-like matrices, for which the lower triangular part is coming from a rank 1 matrix are called generator representable Hessenberg-like matrices throughout the remainder of the paper. It is clear that these matrices can be written as the sum of a rank 1 matrix plus a strictly upper triangular matrix. Moreover the two vectors defining the rank one matrix are called the generators of the Hessenberg-like matrix. Important is the connection between generator representable Hessenberg-like matrices, Hessenberg-like matrices and their inverses.

Theorem 3 *The inverse of an invertible*

- *generator representable Hessenberg-like matrix is an irreducible Hessenberg matrix;*
- *Hessenberg-like matrix is a Hessenberg matrix.*

PROOF. These are well-known facts and therefore we do not include the proof. Proofs can be found for example in [12,13,14].

As we want to prove a general implicit Q-theorem it is natural that we do not restrict ourselves to the class of generator representable Hessenberg-like matrices, but that we formulate the theorem in terms of Hessenberg-like matrices. An implicit Q-theorem for symmetric generator representable Hessenberg-like matrices, these are in fact generator representable symmetric semiseparable matrices, is already available in the literature [5]. An important reason, why one should not restrict the implicit Q-theorem to the class of generator representable Hessenberg-like matrices is the following one: If one applies several steps of the QR-algorithm to a Hessenberg-like matrix, one desires that this matrix converges to a matrix with a zero block in the lower left position. This allows us to do deflation and continue working with smaller Hessenberg-like matrices. As however the Hessenberg-like matrix with a zero block in the lower left position is not a generator representable Hessenberg-like matrix anymore, the theorem as provided in [5] becomes inapplicable. An implicit Q-theorem, valid for general Hessenberg-like matrices however, remains valid.

In the remainder of the section we briefly mention two theorems, relating Hessenberg-like and generator representable Hessenberg-like matrices. These theorems and proofs can be found in [11].

Theorem 4 *Suppose a Hessenberg-like matrix Z is given. Then this matrix can be written as the sum of a strictly upper triangular matrix and a block diagonal matrix, for which all the blocks are generator representable Hessenberg-like matrices.*

Definition 5 *The pointwise limit of a collection of matrices $A_\epsilon \in \mathbb{R}^{n\times n}$ for $\epsilon \in I$ (if it exists), with the matrices A_ϵ as*

$$A_\epsilon = \begin{pmatrix} (a_{1,1})_\epsilon & \cdots & (a_{1,n})_\epsilon \\ \vdots & \ddots & \vdots \\ (a_{n,1})_\epsilon & \cdots & (a_{n,n})_\epsilon \end{pmatrix}$$

is defined as:

$$\lim_{\epsilon\to\epsilon_0} A_\epsilon = \begin{pmatrix} \lim_{\epsilon\to\epsilon_0}(a_{1,1})_\epsilon & \cdots & \lim_{\epsilon\to\epsilon_0}(a_{1,n})_\epsilon \\ \vdots & \ddots & \vdots \\ \lim_{\epsilon\to\epsilon_0}(a_{n,1})_\epsilon & \cdots & \lim_{\epsilon\to\epsilon_0}(a_{n,n})_\epsilon \end{pmatrix}.$$

Theorem 6 *The pointwise closure of the class of generator representable Hessenberg-like matrices is the class of Hessenberg-like matrices.*

Before formulating and proving the implicit Q-theorem for Hessenberg-like matrices, we will define unreduced Hessenberg-like matrices.

3 Unreduced Hessenberg-like matrices, definition and properties

This section is divided in three subsections. A first section briefly defines unreduced Hessenberg-like matrices. As however the definition of unreducedness is closely related to implicit QR-algorithms, the second section focusses on some properties of the unreduced matrices, w.r.t. the uniqueness of QR-decompositions. Essential for the implicit QR-algorithm for Hessenberg matrices is that the initial Hessenberg matrix is unreduced, therefore the third section proposes a method for transforming Hessenberg-like matrices via orthogonal similarity transformations to unreduced Hessenberg-like form.

3.1 The definition

Let us define now an unreduced Hessenberg-like matrix, in the following way:

Definition 7 *An Hessenberg-like matrix Z is said to be unreduced if*

(1) all the blocks $Z(i+1:n, 1:i)$ (for $i = 1, \ldots, n-1$) have rank equal to 1; this corresponds to the fact that there are no zero blocks below the diagonal;
(2) all the blocks $Z(i:n, 1:i+1)$ (for $i = 1, \ldots, n-1$) have rank strictly higher than 1, this means that on the superdiagonal, no elements are includable in the semiseparable structure.

Note 1 *If an Hessenberg-like matrix Z is unreduced it is also nonsingular. This can be seen by calculating the QR-factorization of Z as described in [15]. Because of the rank 1 assumptions below the diagonal, a sequence of $n-1$ Givens transformations from bottom to top applied on the left of Z will make the matrix upper triangular. Because none of the elements above the diagonal is includable in the semiseparable structure below the diagonal, all the diagonal elements of the upper triangular matrix R in the QR-factorization of Z will be different from zero. This implies the nonsingularity of the Hessenberg-like matrix Z.*

The first demand in the definition is quite natural and corresponds to the demands placed on a Hessenberg matrix to be unreduced. The second demand seems a little awkward but is essential for the implicit Q-theorem, as will become clear in the proof.

3.2 Some properties

Unreduced Hessenberg matrices are very important for the development of an implicit QR-algorithm for Hessenberg matrices. It is essential for the application of an implicit QR-algorithm that the corresponding Hessenberg matrix is unreduced, otherwise the algorithm would break down.

Another feature of an unreduced Hessenberg matrix H is that it has an essentially unique QR-factorization. Moreover also $H - \kappa I$, with κ a shift has an essentially unique QR-factorization. This is straightforward because of the zero structure below the subdiagonal. Only the last column can be dependent of the first $n-1$ columns. Essentially unique means, only the sign of the columns of the orthogonal matrices Q can differ, and the signs of the elements in R. Because the previous $n-1$ columns are linearly independent, the first $n-1$ columns of the matrix Q in the QR-factorization of $H = QR$, are linearly independent. Q is an orthogonal matrix, and the first $n-1$ columns are essentially unique. As the dimension is n and we already have $n-1$ orthogonal columns, the last column is uniquely defined, orthogonal to the first $n-1$ columns. This means that the QR-factorization is essentially unique. Summarizing, this means that the QR-factorization of an unreduced Hessenberg matrix H has the following form:

$$H = Q \left(\begin{array}{c|c} R & w \\ \hline 0 & \alpha \end{array} \right),$$

for which R is a nonsingular upper triangular matrix of dimension $(n-1) \times (n-1)$, w is a column vector of length $(n-1)$, and α is an element in $\mathbb{R}$. We have that $\alpha = 0$ if and only if H is singular. This means that the last column of H depends on the previous $n-1$.

When we have an unredecued Hessenberg-like matrix, we have the following theorem connected to the QR-factorization, similar like in the Hessenberg case:

Theorem 8 *Suppose Z to be an unreduced Hessenberg-like matrix. Then the matrix $Z - \kappa I$, with κ as a shift has an essentially unique QR-factorization.*

PROOF. If $\kappa = 0$ the theorem is true because Z is nonsingular. Suppose $\kappa \neq 0$. We can apply the $n-1$ Givens transformations from bottom to top on the left of the Hessenberg-like matrix to make the Hessenberg-like matrix upper triangular. Because the Hessenberg-like matrix Z is unreduced, these Givens transformations are nontrivial. Applying them to the left of the matrix $Z - \kappa I$ results therefore in an unreducible Hessenberg matrix. As this Hessenberg matrix has the first $n-1$ columns independent of each other also the matrix $Z - \kappa I$ has the first $n-1$ columns independent. Hence, the QR-factorization of $Z - \kappa I$ is essentially unique.

Note 2 *An unreduced Hessenberg-like matrix Z is always nonsingular, and has therefore, always an essentially unique QR-factorization. If one however admits that the block $Z(n-1 : n, 1 : n)$ is of rank 1, the matrix Z also will have an essentially unique QR-factorization.*

Note 3 *The unreduced Hessenberg-like matrix Z is always nonsingular, but the matrix $Z - \kappa I$ can be singular.*

Note 4 *An unreduced Hessenberg-like matrix is always a generator representable Hessenberg-like matrix. This is a consquence of the definition of unreducedness and Theorem 4.*

In the following section we address how one can transform a Hessenberg-like matrix via orthogonal similarity transformations to unreduced form.

3.3 The reduction to unreduced Hessenberg-like form

The implicit QR-algorithm for Hessenberg matrices is based on the unreducedness of the corresponding Hessenberg matrix. For Hessenberg matrices it is straightforward how to split the matrix into two or more Hessenberg matrices which are unreduced (See Section 7.5.1 in [10]).

In this section we will reduce a Hessenberg-like matrix to unreduced form as presented in Definition 7. The first demand, that there are no zero blocks in the lower left corner of the matrix can be satisfied by dividing the matrix into different blocks, in a complete similar way as for the Hessenberg case.

The second demand, the fact that the Hessenberg-like structure does not extend above the diagonal is not solved in such an easy way. It can be seen that matrices having this special structure are singular. In the solution we propose, we will chase in fact the dependent rows completely to the bottom where they will form zero rows which can be removed. This chasing technique is achieved by performing a special QR-step without shift on the matrix. We will demonstrate this technique on a 5×5 matrix. The $\boxtimes$ denote the elements satisfying the Hessenberg-like structure. One can clearly see that the Hessenberg-like structure of the matrix extends above the diagonal:

$$Z = \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \end{pmatrix}.$$

We start the annihilation with the traditional orthogonal matrix $Q_1 = G_4 G_3 G_2 G_1$, which consists of 4 Givens transformations (see [15]). This results in the following matrix $\tilde{R} = G_4^T G_3^T G_2^T G_1^T Z$:

$$\tilde{R} = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

In normal circumstances one would apply the transformation Q_1 now on the right of the above matrix to complete one step of QR without shift. Instead of applying now this transformation we continue and annihilate the elements marked with $\otimes$ with the Givens transformations G_5 and G_6. G_5^T is performed on rows 3 and 4, while G_6^T is performed on rows 4 and 5:

$$\begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & \otimes & \times \\ 0 & 0 & 0 & 0 & \times \end{pmatrix} \xrightarrow{G_5^T \tilde{R}} \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & \otimes \end{pmatrix}$$
$$\xrightarrow{G_6^T G_5^T \tilde{R}} \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We have now finished performing the transformations on the left side of the matrix. Denote $R = G_6^T G_5^T \tilde{R}$. To complete the QR-step applied to the matrix Z, we have to perform the transformations on the right-side of the matrix:

$$R = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{RG_1} \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \boxtimes & \times \\ 0 & 0 & 0 & \boxtimes & \boxtimes \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\xrightarrow{RG_1G_2} \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \boxtimes & \boxtimes & \times \\ 0 & 0 & \boxtimes & \boxtimes & \boxtimes \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{RG_1G_2G_3} \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \boxtimes & \times & \times & \times \\ 0 & \boxtimes & \boxtimes & \boxtimes & \times \\ 0 & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{RQ_1} \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{RQ_1 G_5} \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{RQ_1 G_5} \left(\begin{array}{cccc|c} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

It is clear that the last matrix can be deflated such that we get an unreduced matrix and one eigenvalue which has already converged.

The same technique can be applied for matrices with larger blocks crossing the diagonal or more blocks crossing the diagonal.

4 The implicit Q-theorem

We will now prove an implicit Q-theorem for Hessenberg-like matrices. The theorem is proved in a similar way as the implicit Q-theorem for Hessenberg matrices as formulated in Section 7.4.5 from [10]. First, two definitions and some propositions are needed.

Definition 9 *Two matrices Z_1 and Z_2 are called essentially the same if there exists a matrix $W = \text{diag}(\pm 1, \pm 1, \ldots, \pm 1)$ such that the following equation holds:*

$$Z_1 = W Z_2 W^T.$$

Before proving the implicit Q-theorem, we define the so called unreduced number of a Hessenberg-like matrix.

Definition 10 *Suppose Z to be a Hessenberg-like matrix, the unreduced number k of Z is the smallest integer such that one of the following two conditions is satisfied:*

(1) The submatrix $S(k+1:n; 1:k) = 0$.
(2) The element $S(k+1, k+2)$ with $k < n-1$ is includable in the lower semiseparable structure.

If the matrix is unreduced, $k = n$.

Proposition 11 *Suppose we have a Hessenberg-like matrix Z which is not unreduced, and does not have any zero blocks below the diagonal, i.e., whose bottom left element is nonzero. Then this matrix can be transformed via similarity transformations to a Hessenberg-like matrix, for which the upper left $(n-l)\times(n-l)$ matrix is of unreduced form and the last l rows equal to zero, where l equals the nullity of the matrix. Moreover if k is the unreduced number of the matrix Z then the orthogonal transformation can be chosen in such a way that the upper left $k \times k$ block of this orthogonal transformation equals the identity matrix.*

PROOF. We will explicitely construct the transformation matrix. We illustrate this technique on a matrix of dimension 5×5 of the following form

$$Z = \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \end{pmatrix}.$$

We annihilate first the elements of the fifth and fourth row, with Givens transformations G_1 and G_2. This results in the matrix

$$G_2^T G_1^T Z = \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ 0 & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & \otimes \end{pmatrix}.$$

The Givens transformation G_3 is constructed to annihilate the element marked with $\otimes$:

$$G_3^T G_2^T G_1^T Z = \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ 0 & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Completing the similarity transformation by applying the transformations G_1, G_2 and G_3 on the right of $G_3^T G_2^T G_1^T Z$ gives us the following matrix:

$$\tilde{Z} = G_3^T G_2^T G_1^T Z G_1 G_2 G_3 = \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix does not yet satisfy the desired structure, as we want the upper left 4×4 block to be of Hessenberg-like form. To do so, we remake the matrix semiseparable (techniques for making a complete matrix semiseparable can be found in [9]). Let us perform the transformation G_4 on the right of the matrix $\tilde{Z}$ to annihilate the marked element. Completing this similarity transformation gives us:

$$\tilde{Z} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ 0 & 0 & \otimes & \times & \times \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\tilde{Z}G_4} \begin{pmatrix} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{G_4^T \tilde{Z} G_4} \left(\begin{array}{cccc|c} \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The resulting matrix has the upper left 4×4 block of unreduced Hessenberg-like form, while the last row is zero. The transformations involved did not change the upper left 2×2 block as desired.

Note 5 *The resulting matrices from Proposition 11 are sometimes called zero-tailed matrices (see [5]).*

Note 6 *Proposition 11 can be generalized to Hessenberg-like matrices with zero blocks below the diagonal. These matrices can be transformed to a Hessenberg-like matrix with zero rows on the bottom and all the Hessenberg-like blocks on the diagonal of unreduced form.*

Proposition 12 *Suppose Z is an unreduced Hessenberg-like matrix. Then Z can be written as the sum of a rank 1 matrix and a strictly upper triangular matrix, where the superdiagonal elements of this matrix are different from zero.*

We have that

$$Z = uv^T + R,$$

the vectors u_n *and* v_1 *are different from zero and* R *is a strictly upper triangular matrix having nonzero supdiagonal elements.*

PROOF. Straightforward, because of the fact that an unreduced Hessenberg-like matrix is generator representable.

Proposition 13 *Suppose we have the following equality:*

$$WZ = XW, \tag{1}$$

where W *is an orthogonal matrix with the first column equal to* e_1 *and* Z *and* X *are two* $n \times n$ *Hessenberg-like matrices of the following form:*

$$Z = \begin{pmatrix} Z_1 & \times & \times \\ 0 & Z_2 & \times \\ 0 & 0 & 0 \end{pmatrix} \quad X = \begin{pmatrix} X_1 & \times & \times \\ 0 & X_2 & \times \\ 0 & 0 & 0 \end{pmatrix}$$

both having l *zero rows at the bottom and the matrices* Z_1*,* Z_2*,* X_1 *and* X_2 *of unreduced form. If we denote the dimension of the upper left block of* Z *with* n_{Z_1} *and the dimension of the upper left block of* X *with* n_{X_1}*. Then we have that* $n_{X_1} = n_{Z_1}$ *and* W *has the lower left* $(n - n_{X_1}) \times n_{X_1}$ *block equal to zero.*

PROOF. Assume $n_{Z_1} \geq n_{X_1}$. When considering the first n_{Z_1} columns of equation (1) we have

$$W(1:n, 1:n_{Z_1})\, Z_1 = V, \tag{2}$$

where V is of dimension $n \times n_{Z_1}$, with the last l rows of V equal to zero. Because Z_1 is invertible we know that $W(1:n, 1:n_{Z_1}) = VZ_1^{-1}$ has the last l rows equal to zero. We will prove by induction that all the columns w_k with $1 \leq k \leq n_{X_1}$ have the components with index higher than n_{X_1} equal to zero.

Step 1: $k = 1$. Because $We_1 = e_1$, we know already that this is true for $k = 1$. Let us write the matrices Z_1 and X_1 as (Using Proposition 12):

$$Z_1 = u^{(1)} {v^{(1)}}^T + R^{(1)},$$
$$X_1 = u^{(2)} {v^{(2)}}^T + R^{(2)},$$

with $v_1^{(1)} = v_1^{(2)} = 1$. Multiplying equation (1) to the right with e_1 gives us

$$W \begin{pmatrix} u^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} u^{(2)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Note that $u^{(1)}$ is of length n_{Z_1} and $u^{(2)}$ is of length n_{X_1}.

Step k: Suppose $2 \leq k \leq n_{X_1}$. We prove that w_k has the components $n_{X_1}+1, \ldots, n$ equal to zero. We know by induction that this is true for the columns w_i with $i < k$. Multiply both sides of equation (1) with e_k, this gives us:

$$W \left(\begin{pmatrix} u^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} v_k^{(1)} + \begin{pmatrix} r_{1,k}^{(1)} \\ \vdots \\ r_{k-1,k}^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right) = X w_k.$$

This can be rewritten as:

$$u^{(2)} v_k^{(1)} + W \begin{pmatrix} r_{1,k}^{(1)} \\ \vdots \\ r_{k-1,k}^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = X w_k.$$

Because of the induction procedure, we know that the left-hand side of the former equation has the components below element n_{X_1} equal to zero. The vector w_k has the last l components equal to zero. Because the matrices X_1 and X_2 are nonsingular we know that w_k has only the first n_{X_1} components different from zero, because $X w_k$ can only have the first n_{X_1} components different from zero. This proves the induction step.

This means that the matrix W has in the lower left position an $(n - n_{X_1}) \times n_{X_1}$ block of zeros. A combination of equation (1) and the zero structure of W leads to the fact that WZ needs to have a zero block in the lower left position of dimension $(n - n_{X_1}) \times n_{X_1}$. Therefore the matrix Z_1 has to be of dimension n_{X_1} which proves the proposition.

Note 7 *The proposition above can be formulated in a similar way for more zero blocks below the diagonal.*

Using this property, we can prove the following implicit Q-theorem, for Hessenberg-like matrices.

Theorem 14 (implicit Q-theorem for Hessenberg-like matrices) *Suppose the following equations hold,*

$$Q_1^T A Q_1 = Z \tag{3}$$
$$Q_2^T A Q_2 = X \tag{4}$$

with $Q_1 e_1 = Q_2 e_1$, where Z and X are two $n \times n$ Hessenberg-like matrices, with unreduced numbers k_1 and k_2 respectively and Q_1 and Q_2 are orthogonal matrices. Let us denote $k = \min(k_1, k_2)$. Then we have that the first k columns of Q_1 and Q_2 are the same, up to the sign, and the upper left $k \times k$ submatrices of Z and of X are essentially the same. More precisely, there exists a matrix $V = \operatorname{diag}(1, \pm 1, \pm 1, \ldots, \pm 1)$, of size $k \times k$, such that we have the following two equations:

$$\begin{aligned} Q_1(1:n, 1:k) &= Q_2(1:n, 1:k)\, V, \\ Z(1:k, 1:k)\, V &= V\, X(1:k, 1:k). \end{aligned}$$

PROOF. Using Proposition 11 and the notes following the proposition we can assume that the matrices Z and X in (3) and (4) are of the following form (for simplicity, we assume the number of blocks to be equal to two):

$$Z = \begin{pmatrix} Z_1 & \times & \times \\ 0 & Z_2 & \times \\ 0 & 0 & 0 \end{pmatrix} \quad X = \begin{pmatrix} X_1 & \times & \times \\ 0 & X_2 & \times \\ 0 & 0 & 0 \end{pmatrix},$$

with Z_1, Z_2, X_1 and X_2 unreduced Hessenberg-like matrices. This does not affect our statements from the theorem, as the performed transformations from Proposition 11 do not affect the upper left $k \times k$ block of the matrices Z and/or X, or the first k columns of Q_1 and Q_2.

Denoting $W = Q_1^T Q_2$ and using the equations (3) and (4) the following equality holds:

$$ZW = WX, \text{ with } We_1 = e_1. \tag{5}$$

Moreover, we know by Proposition 13 that the matrix W has the lower left block of dimension $(n-n_1)\times n_1$ equal to zero, where n_1 is the dimension of the block Z_1.

If we can prove now that for the first n_1 columns of W the following equality holds: $w_k = \pm e_k$, then we have that (as $n_1 \geq k$ and by Proposition 11) the theorem holds.

According to Proposition 12 we can write the upper rows of the matrix Z in the following form:

$$Z(1:n_1;1:n) = u^{(1)}\left({v^{(1)}}^T, {\tilde{v}^{(1)}}^T\right) + R^{(1)},$$

where $u^{(1)}, v^{(1)}$ have length n_1. The vector $\tilde{v}^{(1)}$ is of length $n-n_1$ and is chosen in such a way that the matrix $R^{(1)} \in \mathbb{R}^{n_1\times n}$, has the last row equal to zero. Moreover the matrix $R^{(1)}$ has the left $n_1 \times n_1$ block strictly upper triangular. Also the left part of X can be written as

$$X(1:n;1:n_1) = \begin{pmatrix} u^{(2)} \\ \tilde{u}^{(2)} \end{pmatrix} {v^{(2)}}^T + R^{(2)}$$

with $u^{(2)}$ and $v^{(2)}$ of length n_1, $\tilde{u}^{(2)}$ is of length $n-n_1$ and $R^{(2)}$ a strictly upper triangular matrix of dimension $n\times n_1$. Both of the strictly upper triangular parts $R^{(1)}$ and $R^{(2)}$ have nonzero elements on the supdiagonals, because of the unreducedness assumption. The couples $u^{(1)}, v^{(1)}$ and $u^{(2)}, v^{(2)}$ are the generators of the semiseparable matrices Z_1 and X_1, respectively. They are normalized in such a way that $v_1^{(1)} = v_1^{(2)} = 1$. Denoting with P the projection operator $P = (I_{n_1},\ 0)$ we can calculate the upper left $n_1 \times n_1$ block of the matrices in equation (5):

$$\begin{aligned} PZWP^T &= PWXP^T \\ \left(u^{(1)}\left({v^{(1)}}^T, {\tilde{v}^{(1)}}^T\right) + R^{(1)}\right)WP^T &= PW\left(\begin{pmatrix} u^{(2)} \\ \tilde{u}^{(2)} \end{pmatrix}{v^{(2)}}^T + R^{(2)}\right). \end{aligned} \tag{6}$$

Denoting the columns of W as $(w_1, w_2, \ldots, w_n)$, the fact that $Q_1e_1 = Q_2e_1$ leads to the fact that $w_1 = e_1$. Hence, also the first row of W equals e_1^T. Multiplying (6) to the right by e_1, gives

$$u^{(1)} = PW\begin{pmatrix} u^{(2)} \\ \tilde{u}^{(2)} \end{pmatrix}. \tag{7}$$

Because of the structure of P and W, multiplying (7) to the left by e_1^T gives us that $u_1^{(1)} = u_1^{(2)}$. Using equation (7), we will prove now by induction that for $i \leq n_1 : w_i = \pm e_i$ and

$$\left(v^{(1)^T}, \tilde{v}^{(1)^T}\right) WP^T = \left(v^{(1)^T}, \tilde{v}^{(1)^T}\right)(w_1,\ w_2, \ldots, w_{n_1}) = v^{(2)^T}, \qquad (8)$$

which will prove the theorem.

Step 1 $l = 1$. This is a trivial step, because $v_1^{(1)} = v_1^{(2)} = 1$ and $w_1 = e_1$.

Step l $1 < l \leq n_1$. By induction we have that $w_i = \pm e_i \quad \forall 1 \leq i \leq l-1$ and (8) holds for the first $(l-1)$ columns. This means that:

$$\left(v^{(1)^T}, \tilde{v}^{(1)^T}\right)(w_1, w_2, \ldots, w_{l-1}) = \left(v_1^{(2)}, v_2^{(2)}, \ldots, v_{l-1}^{(2)}\right).$$

Taking (7) into account, (6) becomes

$$u^{(1)}\left(\left(v^{(1)^T}, \tilde{v}^{(1)^T}\right) WP^T - v^{(2)^T}\right) = \left(PWR^{(2)} - R^{(1)}WP^T\right). \qquad (9)$$

Multiplying (9) to the right by e_l, we have

$$u^{(1)}\left(\left(v^{(1)^T}, \tilde{v}^{(1)^T}\right) WP^T - v^{(2)^T}\right) e_l = \left(PWR^{(2)} - R^{(1)}WP^T\right) e_l. \qquad (10)$$

Because of the special structure of $P, W, R^{(1)}$ and $R^{(2)}$ the element in the n_1th position in the vector on the right-hand side of (10) is equal to zero.

We know that $u_{n_1}^{(1)}$ is different from zero, because of the unreducedness assumption. Therefore the following equation holds:

$$u_{n_1}^{(1)}\left(\left(v^{(1)^T}, \tilde{v}^{(1)^T}\right) WP^T - v^{(2)^T}\right) e_l = 0.$$

This means that

$$\left(v^{(1)^T}, \tilde{v}^{(1)^T}\right) w_l - v_l^{(2)} = 0.$$

Therefore equation (8) is already satisfied up to element l:

$$\left(v^{(1)^T}, \tilde{v}^{(1)^T}\right)(w_1, w_2, \ldots, w_l) = \left(v_1^{(2)}, v_2^{(2)}, \ldots, v_l^{(2)}\right). \qquad (11)$$

Using equation (11) together with equation (10) leads to the fact that the complete right-hand side of equation (10) has to be zero. This gives the following equation:

$$R^{(1)}WP^T e_l = PWR^{(2)} e_l$$

leading to

$$R^{(1)}We_l = \sum_{j=1}^{l-1} r_{j,l}^{(2)} w_j \qquad (12)$$

with the $r_{j,l}^{(2)}$ as the elements of column l of matrix $R^{(2)}$. Hence, the right-hand side can only have the first $l-1$ components different from zero.

Because the superdiagonal elements of the left square block of $R^{(1)}$ are nonzero and because the first n_1 columns of W have the last $(n - n_1)$ elements equal to zero, only the first l elements of w_l can be different from zero. This together with the fact that W is orthogonal and $w_i = \pm e_i$ for $i < l$, means that $w_l = \pm e_l$, which proves the induction step.

Note 8 *In the definition of the unreduced number we assumed that also the element $Z(1,2)$ was not includable in the semiseparable structure. The reader can verify, that one can admit this element to be includable in the semiseparable structure, and the theorem and the proof still remain true for this newly defined unreduced number.*

Note 9 *This theorem can also be applied to the reduction algorithms which transform matrices to Hessenberg-like, symmetric semiseparable and upper triangular semiseparable form, thereby stating the uniqueness of these reduction algorithms in case the outcome is in unreduced form and $Q_1 e_1 = Q_2 e_2$, where Q_1 and Q_2 are the two different orthogonal matrices involved in the two different similarity transformations. Moreover, if the resulting matrix is not unreduced, uniqueness of the upper left $k \times k$ block can be stated, where k is the unreduced number of the resulting semiseparable matrix.*

Finally we give some examples, connected with this theorem. (See also [5].) The first condition in Definition 10 is quite logical and we will not give any examples connected with this condition.

Example 15 *Suppose we have the matrices:*

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

and Q_2=I. Then we have that $X = A$ and

$$Z = \begin{pmatrix} 1 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

One can see that the unreduced number of X and Z equals 1 because the element $X(2,3)$ is includable in the lower semiseparable structure of the matrix X. Thus we know by the theorem that equality is only garantueed for the upper left element of the matrices X and Z.

Example 16 *Suppose we have the matrices:*

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

and Q_2=I. *Then we have that* $X = A$ *and*

$$Z = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

One can see that the unreduced number of X *equals* 0 *because the element* $X(1,2)$ *is includable in the lower semiseparable structure of the matrix* X*. But using the definition of the unreduced number connected to the note following Theorem 14, we know that the unreduced number equals* 1 *and the equality holds for the upper left element of* Z *and* X*.*

5 Conclusions

In this paper we formulated and proved a general implicit Q-theorem for Hessenberg-like matrices. The theorem can be adapted in a simple way to suit also the symmetric case: the symmetric semiseparable matrices. Connected with this theorem we defined the notion of unreduced Hessenberg-like matrices and moreover some properties and a reduction towards this form were deduced.

The theorem can be used for example for proving the uniqueness of the reduction to semiseparable form as described in [9], when the first column of the two different orthogonal transformation matrices Q_1 and Q_2 are the same. Moreover the theorem is useful for proving the correctness of implicit QR-algorithms for semiseparable matrices.

References

[1] S. Chandrasekaran, M. Gu, A divide and conquer algorithm for the eigendecomposition of symmetric block-diagonal plus semi-separable matrices, Numerische Mathematik 96 (4) (2004) 723–731.

[2] N. Mastronardi, E. Van Camp, M. Van Barel, Divide and conquer type algorithms for computing the eigendecomposition of diagonal plus semiseparable matrices., Tech. Rep. 7 (5/2003), Istituto per le Applicazioni del Calcolo "M. Picone", Consiglio Nazionale delle Ricerche, Rome, Italy (2003).

[3] S. Chandrasekaran, M. Gu, Fast and stable eigendecomposition of symmetric banded plus semi-separable matrices, Linear Algebra and Its Applications 313 (2000) 107–114.

[4] D. Fasino, N. Mastronardi, M. Van Barel, Fast and stable algorithms for reducing diagonal plus semiseparable matrices to tridiagonal and bidiagonal form, Contemporary Mathematics 323 (2003) 105–118.

[5] R. Bevilacqua, G. M. Del Corso, Existence, uniqueness and algorithms for matrix unitary reduction to semiseparable form, Tech. Report TR-03-12, University of Pisa (July 2003).

[6] D. Bini, F. Daddi, L. Gemignani, On the shifted QR-iteration applied to Frobenius matrices, preprint (October 2003).

[7] D. Bini, L. Gemignani, V. Pan, QR-like algorithms for generalized semiseparable matrices, Tech. Report 1470, Department of Mathematics, University of Pisa (2004).

[8] R. Vandebril, M. Van Barel, N. Mastronardi, An implicit QR algorithm for semiseparable matrices to compute the eigendecomposition of symmetric matrices, Report TW 367, Department of Computer Science, K.U.Leuven, Leuven, Belgium (August 2003).

[9] M. Van Barel, R. Vandebril, N. Mastronardi, The Lanczos-Ritz values appearing in an orthogonal similarity reduction of a matrix into semiseparable form, Report TW 360, Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Leuven (Heverlee) (May 2003).

[10] G. H. Golub, C. F. Van Loan, Matrix Computations, 3rd Edition, The Johns Hopkins university Press, 1996.

[11] R. Vandebril, N. Van Barel, M. end Mastronardi, A note on the representation and definition of semiseparable matrices, Numerical Linear Algebra with Applications Accepted for publication (Report TW393).

[12] E. Asplund, Inverses of matrices a_{ij} which satisfy $a_{ij} = 0$ for $j > i + p$, Mathematica Scandinavica 7 (1959) 57–60.

[13] S. Asplund, Finite boundary value problems solved by Green's matrix, Mathematica Scandinavica 7 (1959) 49–56.

[14] M. Fiedler, T. Markham, Completing a matrix when certain entries of its inverse are specified, Linear Algebra and Its Applications 74 (1986) 225–237.

[15] E. Van Camp, N. Mastronardi, M. Van Barel, Two fast algorithms for solving diagonal-plus-semiseparable linear systems, Journal of Computational and Applied Mathematics 164-165 (2004) 731–747.
URL `file:/home/marc/OZOEK/ARTICLES/ma035/info.html`