Orthogonal similarity transformation into semiseparable matrices of semiseparability rank k

Marc Van Barel
Ellen Van Camp
Nicola Mastronardi

Report TW 380, January 2004

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

Orthogonal similarity transformation into semiseparable matrices of semiseparability rank k

Marc Van Barel
Ellen Van Camp
Nicola Mastronardi

Report TW 380, January 2004

Department of Computer Science, K.U.Leuven

Abstract

Very recently, an algorithm, which reduces any symmetric matrix into a semiseparable one of semiseparability rank 1 by similar orthogonality transformations, has been proposed by Vandebril, Van Barel and Mastronardi. Partial execution of this algorithm computes a semiseparable matrix whose eigenvalues are the Ritz-values obtained by the Lanczos' process applied to the original matrix. Also a kind of nested subspace iteration is performed at each step.

In this paper, we generalize the above results and propose an algorithm to reduce any symmetric matrix into a similar block-semiseparable one of semiseparability rank k, with $k \in \mathbb{N}$, by orthogonal similarity transformations. Also in this case partial execution of the algorithm computes a block-semiseparable matrix whose eigenvalues are the Ritz-values obtained by the block-Lanczos' process with k starting vectors, applied to the original matrix. Subspace iteration is performed at each step as well.

Keywords : Similarity transformation, block-semiseparable matrix, semiseparability rank k, block-Lanczos' algorithm, Ritz-values, subspace iteration.
AMS(MOS) Classification : Primary : 65F15, Secondary : 15A21, 15A23.

NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2003; **999**:1–15 *Prepared using* nlaauth.cls *[Version: 2002/09/18 v1.02]*

Orthogonal similarity transformation into semiseparable matrices of semiseparability rank k

M. Van Barel[1*], E. Van Camp[1] and N. Mastronardi[2]

[1] *Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Leuven (Heverlee), Belgium.*
[2] *Istituto per le Applicazioni del Calcolo "M. Picone", sez. Bari, Consiglio Nazionale delle Ricerche, Via G. Amendola, 122/I, I-70126 Bari, Italy.*

SUMMARY

Very recently, an algorithm, which reduces any symmetric matrix into a semiseparable one of semiseparability rank 1 by similar orthogonality transformations, has been proposed by Vandebril, Van Barel and Mastronardi. Partial execution of this algorithm computes a semiseparable matrix whose eigenvalues are the Ritz-values obtained by the Lanczos' process applied to the original matrix. Also a kind of nested subspace iteration is performed at each step.

In this paper, we generalize the above results and propose an algorithm to reduce any symmetric matrix into a similar block-semiseparable one of semiseparability rank k, with $k \in \mathbb{N}$, by orthogonal similarity transformations. Also in this case partial execution of the algorithm computes a block-semiseparable matrix whose eigenvalues are the Ritz-values obtained by the block-Lanczos' process with k starting vectors, applied to the original matrix. Subspace iteration is performed at each step as well. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: Similarity transformation, block-semiseparable matrix, semiseparability rank k, block-Lanczos' algorithm, Ritz-values, subspace iteration.

*Correspondence to: Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Leuven (Heverlee), Belgium. email: Marc.VanBarel@cs.kuleuven.ac.be, phone: +32/016 327563, fax: +32/016 327996.

Contract/grant sponsor: The research of the first two authors was 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), and G.0184.02 (CORFU: Constructive study of Orthogonal Rational Functions), by the K.U.Leuven (Bijzonder Onderzoeksfonds), 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. The research of the third author was partially supported by MIUR, grant number 2002014121.

Received 10 December 2003

1. Introduction

In [9] Van Barel, Vandebril and Mastronardi developed an algorithm which transforms any symmetric matrix $A \in \mathbb{R}^{n \times n}$ into a semiseparable one of semiseparability rank 1 by means of orthogonal similarity transformations. Analogously to the Lanczos' algorithm, while executing the latter transformation, at the ith step, $i = 1, \ldots, n-1$, a semiseparable matrix of order $i+1$ is constructed whose eigenvalues are the Lanczos-Ritz values. Also a kind of nested subspace iteration is performed. Hence for several eigenvalue distributions, this algorithm has the important feature that, after few steps, a diagonal block is created whose eigenvalues are very accurate approximations of the extreme eigenvalues of the original matrix. These extreme eigenvalues are very important in a lot of applications, especially in large scale problems appearing in, for example, principal component analysis, data mining, computing the eigensolutions of the Schrödinger equation on a grid, magnetic resonance spectroscopy, gene expression data, microarray data analysis, ... (see, for instance, [2, 6, 10, 11, 12, 13] and the references therein).

When dealing with eigenvalues of higher multiplicity, however, or with clusters of eigenvalues, the block-Lanczos behavior is more preferable (see, for instance, [1, 3, 4, 5, 7] and the references therein). Therefore, an algorithm with the block-Lanczos behavior is constructed in this paper. This algorithm is a block-version of the one presented in [9] and transforms any symmetric matrix into a block-semiseparable one of semiseparability rank k, with $k \in \mathbb{N}$, by means of orthogonal similarity transformations. Also subspace iteration is performed at each step. Hence the block-algorithm, generically, approximates the extreme eigenvalues of the original matrix first with multiplicity k if their multiplicity is greater or equal to k and with their proper multiplicity otherwise. Also clusters of eigenvalues are approximated accurately.

The paper is structured as follows. In §2, after repeating the definition of a semiseparable matrix of semiseparability rank 1, two extensions to rank k and suitable representations are presented. Section 3 generalizes the transformation of a symmetric matrix into a semiseparable matrix of semiseparability rank 1 to a transformation into a block-semiseparable one of semiseparability rank k. The properties of block-Lanczos behavior and subspace iteration are discussed in §4 and Section 5 briefly treats the computational complexity. In §6 numerical experiments confirm the properties mentioned in §4 and finally the conclusion is stated in §7.

2. Definitions and representations

Slightly different definitions of semiseparable matrices can be found in the literature. In [8], an overview of different definitions and representations of symmetric semiseparable matrices of semiseparability rank 1, as well as their properties, is given. From these definitions, one is chosen which preserves the properties dual to the class of tridiagonal matrices and a suitable new representation is built.

In this section, we repeat this definition and representation of a semiseparable matrix of semiseparability rank 1 and propose two generalizations of the latter definition and representation to semiseparable matrices of semiseparability rank k, with $k \in \mathbb{N}$. The first one is an obvious extension of the rank 1 case. The second one is a block version.

2.1. A semiseparable matrix of semiseparability rank 1

Definition 1. *A matrix S of order n is called* a (strictly) lower (upper) semiseparable matrix of semiseparability rank 1 *if all submatrices which can be taken out of the (strictly) lower (upper) triangular part of the matrix S have rank ≤ 1.*

A lower and strictly upper semiseparable matrix of semiseparability rank 1, is called a semiseparable matrix of semiseparability rank 1.

For a semiseparable matrix of semiseparability rank 1, the representation proposed in [8] consists of a vector $d^{(l)} = \left[d_1^{(l)}, d_2^{(l)}, \ldots, d_n^{(l)}\right]$ of length n and $n-1$ Givens rotations $G_1^{(l)}, G_2^{(l)}, \ldots, G_{n-1}^{(l)}$, in order to construct the lower triangular part, and a vector $d^{(u)} = \left[d_1^{(u)}, d_2^{(u)}, \ldots, d_{n-1}^{(u)}\right]$ of length $n-1$ and $n-2$ Givens rotations $G_1^{(u)}, G_2^{(u)}, \ldots, G_{n-2}^{(u)}$ in order to construct the upper triangular part.

With this vector-Givens representation, the lower triangular part is built as follows.

Let S be an $n \times n$ matrix of zeros, except for the upper-left corner element $d_1^{(l)}$. First we apply the Givens rotation $G_1^{(l)}$ to the first and the second row and add $d_2^{(l)}$ at the second diagonal element. We denote by $\boxtimes$ the entries of the part of the matrix having already the semiseparable structure. Only the first 2 rows and columns are shown:

$$\begin{pmatrix} d_1^{(l)} & 0 \\ 0 & 0 \end{pmatrix} \rightarrow G_1^{(l)} \begin{pmatrix} d_1^{(l)} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & d_2^{(l)} \end{pmatrix} \rightarrow \begin{pmatrix} \boxtimes & 0 \\ \boxtimes & d_2^{(l)} \end{pmatrix}.$$

At the second step, we apply $G_2^{(l)}$ to the second and third row and add $d_3^{(l)}$ at the third diagonal element. Only the involved rows and the first 3 columns are shown:

$$\begin{pmatrix} \boxtimes & d_2^{(l)} & 0 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow G_2^{(l)} \begin{pmatrix} \boxtimes & d_2^{(l)} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & d_3^{(l)} \end{pmatrix} \rightarrow \begin{pmatrix} \boxtimes & \boxtimes & 0 \\ \boxtimes & \boxtimes & d_3^{(l)} \end{pmatrix}.$$

Repeating this process $n-1$ times, the lower triangular part of a semiseparable matrix of semiseparability rank 1 is constructed.

The strictly upper block-triangular part can be constructed in a similar way, starting from $n-1$ elements $d_i^{(u)}$ and $n-2$ Givens rotations $G_i^{(u)}$. Instead of applying the Givens rotations $G_i^{(u)}$ to the rows, their transpose must be applied to the columns. Hence the complete semiseparable matrix can be reconstructed in this way.

Moreover, if the semiseparable matrix is symmetric, only the knowledge of the vector-Givens representation $d^{(l)}$ and $G_i^{(l)}$, for $i = 1, \ldots, n-1$, of the lower triangular part is needed to reconstruct the whole matrix. In the remainder of this paper we will work with symmetric matrices, so we can omit the superscript (l). Denoting the Givens transformation G_i by

$$G_i = \begin{pmatrix} c_i & -s_i \\ s_i & c_i \end{pmatrix},$$

the constructed symmetric semiseparable matrix has the following form:

$$\begin{pmatrix} c_1 d_1 & c_2 s_1 d_1 & \cdots & c_{n-1} s_{n-2:1} d_1 & s_{n-1:1} d_1 \\ c_2 s_1 d_1 & c_2 d_2 & \cdots & c_{n-1} s_{n-2:2} d_2 & s_{n-1:2} d_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{n-1} s_{n_2:1} d_1 & c_{n-1} s_{n-2:2} d_2 & \cdots & c_{n-1} d_{n-1} & s_{n-1} d_{n-1} \\ s_{n-1:1} d_1 & s_{n-1:2} d_2 & \cdots & s_{n-1} d_{n-1} & d_n \end{pmatrix},$$

with $s_{a:b} \equiv s_a s_{a-1} \dots s_b$.

2.2. A semiseparable matrix of semiseparability rank k

In this subsection, we propose a first extension of Definition 1 to the rank k case, with $k \in \mathbb{N}$, and present a suitable representation.

Definition 2. *A matrix S is called* a (strictly) lower (upper) semiseparable matrix of semiseparability rank k *if all submatrices which can be taken out of the (strictly) lower (upper) triangular part of the matrix S have rank $\leq k$.*

A lower and strictly upper semiseparable matrix of semiseparability rank k, is called a semiseparable matrix of semiseparability rank k.

The representation of the lower triangular part of an n-dimensional semiseparable matrix of semiseparability rank k, consists of k vectors $d^{(1)}, d^{(2)}, \dots, d^{(k)}$ with $d^{(i)} = \left[d_1^{(i)}, d_2^{(i)}, \dots, d_{n-i+1}^{(i)}\right]$ and $nk - \frac{3}{2}k^2 + \frac{1}{2}k$ Givens rotations G_i.

The construction of this lower triangular part is explained by means of an example because this illustrates the idea clearly and a general explanation would be too technical.

Let us construct a 9×9 semiseparable matrix of semiseparability rank 3. The representation of the lower triangular part consists of 3 vectors $d^{(1)} = \left[d_1^{(1)}, d_2^{(1)}, \dots, d_9^{(1)}\right], d^{(2)} = \left[d_1^{(2)}, d_2^{(2)}, \dots, d_8^{(2)}\right]$ and $d^{(3)} = \left[d_1^{(3)}, d_2^{(3)}, \dots, d_7^{(3)}\right]$, and 15 Givens rotations $G_1, \dots, G_{15}$.

The starting matrix consists of all zero elements, except the first three elements of $d^{(1)}$ on the main diagonal, the first three elements of $d^{(2)}$ on the first subdiagonal and the first three elements of $d^{(3)}$ on the second subdiagonal. We only show the non-zero elements:

$$\begin{pmatrix} d_1^{(1)} & & \\ d_1^{(2)} & d_2^{(1)} & \\ d_1^{(3)} & d_2^{(2)} & d_3^{(1)} \\ & d_2^{(3)} & d_3^{(2)} \\ & & d_3^{(3)} \end{pmatrix}.$$

In the first step we apply the first Givens rotation G_1 to the 5th and 6th row, G_2 to rows 4 and 5, G_3 to the 3rd and 4th row and finally, add the fourth element of $d^{(1)}, d^{(2)}$ and $d^{(3)}$ to the corresponding diagonals. We denote by $\boxtimes$ the entries of the matrix belonging to the semiseparable structure. The next figure only shows the first 6 rows and 4 columns.

$$\begin{array}{c} \\ \\ \\ \\ \rightarrow \\ \rightarrow \end{array}\begin{pmatrix} d_1^{(1)} &&& \\ d_1^{(2)} & d_2^{(1)} && \\ d_1^{(3)} & d_2^{(2)} & d_3^{(1)} & \\ 0 & d_2^{(3)} & d_3^{(2)} & 0 \\ 0 & 0 & d_3^{(3)} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{G_1} \begin{array}{c} \\ \\ \\ \rightarrow \\ \rightarrow \\ \\ \end{array}\begin{pmatrix} d_1^{(1)} &&& \\ d_1^{(2)} & d_2^{(1)} && \\ d_1^{(3)} & d_2^{(2)} & d_3^{(1)} & \\ 0 & d_2^{(3)} & d_3^{(2)} & 0 \\ 0 & 0 & \boxtimes & 0 \\ 0 & 0 & \boxtimes & 0 \end{pmatrix}$$

$$\xrightarrow{G_2} \begin{array}{c} \\ \\ \rightarrow \\ \rightarrow \\ \\ \\ \end{array}\begin{pmatrix} d_1^{(1)} &&& \\ d_1^{(2)} & d_2^{(1)} && \\ d_1^{(3)} & d_2^{(2)} & d_3^{(1)} & \\ 0 & \boxtimes & \boxtimes & 0 \\ 0 & \boxtimes & \boxtimes & 0 \\ 0 & 0 & \boxtimes & 0 \end{pmatrix} \xrightarrow{G_3} \begin{pmatrix} d_1^{(1)} &&& \\ d_1^{(2)} & d_2^{(1)} && \\ \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & \boxtimes & \boxtimes & 0 \\ 0 & 0 & \boxtimes & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} d_1^{(1)} &&& \\ d_1^{(2)} & d_2^{(1)} && \\ \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & \boxtimes & \boxtimes & 0 \\ 0 & 0 & \boxtimes & 0 \end{pmatrix} + \begin{pmatrix} 0 &&& \\ 0 & 0 && \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & d_4^{(1)} \\ 0 & 0 & 0 & d_4^{(2)} \\ 0 & 0 & 0 & d_4^{(3)} \end{pmatrix} = \begin{pmatrix} d_1^{(1)} &&& \\ d_1^{(2)} & d_2^{(1)} && \\ \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & d_4^{(1)} \\ 0 & \boxtimes & \boxtimes & d_4^{(2)} \\ 0 & 0 & \boxtimes & d_4^{(3)} \end{pmatrix}.$$

Performing the second step, we apply Givens rotation G_4 to the 6th and 7th row, Givens rotation G_5 to rows 5 and 6 and rotation G_6 to the 4th and 5th row. Then we add the fifth elements of the three vectors to the corresponding diagonals. We only show the first 7 rows and 5 columns in the next figure.

$$\begin{array}{c} \\ \\ \\ \\ \\ \rightarrow \\ \rightarrow \end{array}\begin{pmatrix} \boxtimes &&&& \\ \boxtimes & \boxtimes &&& \\ \boxtimes & \boxtimes & \boxtimes && \\ \boxtimes & \boxtimes & \boxtimes & d_4^{(1)} & \\ 0 & \boxtimes & \boxtimes & d_4^{(2)} & 0 \\ 0 & 0 & \boxtimes & d_4^{(3)} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{G_4} \begin{array}{c} \\ \\ \\ \\ \rightarrow \\ \rightarrow \\ \\ \end{array}\begin{pmatrix} \boxtimes &&&& \\ \boxtimes & \boxtimes &&& \\ \boxtimes & \boxtimes & \boxtimes && \\ \boxtimes & \boxtimes & \boxtimes & d_4^{(1)} & \\ 0 & \boxtimes & \boxtimes & d_4^{(2)} & 0 \\ 0 & 0 & \boxtimes & \boxtimes & 0 \\ 0 & 0 & \boxtimes & \boxtimes & 0 \end{pmatrix}$$

$$\xrightarrow{G_5} \begin{array}{c} \\ \\ \\ \rightarrow \\ \rightarrow \\ \\ \\ \end{array}\begin{pmatrix} \boxtimes &&&& \\ \boxtimes & \boxtimes &&& \\ \boxtimes & \boxtimes & \boxtimes && \\ \boxtimes & \boxtimes & \boxtimes & d_4^{(1)} & \\ 0 & \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & 0 & \boxtimes & \boxtimes & 0 \end{pmatrix} \xrightarrow{G_6} \begin{pmatrix} \boxtimes &&&& \\ \boxtimes & \boxtimes &&& \\ \boxtimes & \boxtimes & \boxtimes && \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & 0 & \boxtimes & \boxtimes & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} \boxtimes &&&& \\ \boxtimes & \boxtimes &&& \\ \boxtimes & \boxtimes & \boxtimes && \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & \boxtimes & \boxtimes & \boxtimes & 0 \\ 0 & 0 & \boxtimes & \boxtimes & 0 \end{pmatrix} + \begin{pmatrix} 0 &&&& \\ 0 & 0 &&& \\ 0 & 0 & 0 && \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & d_5^{(1)} \\ 0 & 0 & 0 & 0 & d_5^{(2)} \\ 0 & 0 & 0 & 0 & d_5^{(3)} \end{pmatrix} = \begin{pmatrix} \boxtimes &&&& \\ \boxtimes & \boxtimes &&& \\ \boxtimes & \boxtimes & \boxtimes && \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & d_5^{(1)} \\ 0 & \boxtimes & \boxtimes & \boxtimes & d_5^{(2)} \\ 0 & 0 & \boxtimes & \boxtimes & d_5^{(3)} \end{pmatrix}.$$

Repeating the former steps another two times, leads to the following matrix:

$$\begin{pmatrix} \boxtimes & & & & & & \\ \boxtimes & \boxtimes & & & & & \\ \boxtimes & \boxtimes & \boxtimes & & & & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & & & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & d_7^{(1)} \\ 0 & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & d_7^{(2)} \\ 0 & 0 & \boxtimes & \boxtimes & \boxtimes & \boxtimes & d_7^{(3)} \end{pmatrix}$$

Hence nine rows are already partially constructed and the last two steps are slightly different from the former ones.

Instead of applying three Givens rotations, only two are needed in this step and also only two elements must be added at the end: apply G_{13} to rows 8 and 9 and G_{14} to rows 7 and 8 and add the last but one elements of $d^{(1)}$ and $d^{(2)}$ at the diagonals. The figure only shows the last three rows.

$$\begin{array}{c} \\ \rightarrow \\ \rightarrow \end{array} \begin{pmatrix} \boxtimes & \boxtimes & \boxtimes & \dots & \boxtimes & d_7^{(1)} & \\ 0 & \boxtimes & \boxtimes & \dots & \boxtimes & d_7^{(2)} & 0 \\ 0 & 0 & \boxtimes & \dots & \boxtimes & d_7^{(3)} & 0 \end{pmatrix} \xrightarrow{G_{13}}$$

$$\begin{array}{c} \rightarrow \\ \rightarrow \\ \\ \end{array} \begin{pmatrix} \boxtimes & \boxtimes & \boxtimes & \dots & \boxtimes & d_7^{(1)} & \\ 0 & \boxtimes & \boxtimes & \dots & \boxtimes & \boxtimes & 0 \\ 0 & \boxtimes & \boxtimes & \dots & \boxtimes & \boxtimes & 0 \end{pmatrix} \xrightarrow{G_{14}} \begin{pmatrix} \boxtimes & \boxtimes & \dots & \boxtimes & \\ \boxtimes & \boxtimes & \dots & \boxtimes & 0 \\ 0 & \boxtimes & \dots & \boxtimes & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} \boxtimes & \boxtimes & \dots & \boxtimes & \\ \boxtimes & \boxtimes & \dots & \boxtimes & 0 \\ 0 & \boxtimes & \dots & \boxtimes & 0 \end{pmatrix} + \begin{pmatrix} 0 & \dots & 0 & \\ 0 & \dots & 0 & d_8^{(1)} \\ 0 & \dots & 0 & d_8^{(2)} \end{pmatrix} = \begin{pmatrix} \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & d_8^{(1)} \\ 0 & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & d_8^{(2)} \end{pmatrix}$$

The last step consists of applying the last Givens rotation G_{15} to the last 2 rows and adding $d_9^{(1)}$ to the main diagonal. Only the last two rows are shown.

$$\begin{array}{c} \rightarrow \\ \rightarrow \end{array} \begin{pmatrix} \boxtimes & \boxtimes & \dots & \boxtimes & d_8^{(1)} & \\ 0 & \boxtimes & \dots & \boxtimes & d_8^{(2)} & 0 \end{pmatrix} \xrightarrow{G_{15}} \begin{pmatrix} \boxtimes & \boxtimes & \dots & \boxtimes & \boxtimes & \\ \boxtimes & \boxtimes & \dots & \boxtimes & \boxtimes & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} \boxtimes & \dots & \boxtimes & \boxtimes & \\ \boxtimes & \dots & \boxtimes & \boxtimes & 0 \end{pmatrix} + \begin{pmatrix} 0 & \dots & 0 & \\ 0 & \dots & 0 & d_9^{(1)} \end{pmatrix} = \begin{pmatrix} \boxtimes & \dots & \boxtimes & \boxtimes & \\ \boxtimes & \dots & \boxtimes & \boxtimes & d_9^{(1)} \end{pmatrix}$$

This explains how the lower triangular part of a semiseparable matrix of semiseparability rank k can be constructed by a vector-Givens representation.

The construction of the upper triangular part is similar. As we already noted, for a symmetric matrix the construction of the lower triangular part is sufficient. Given a semiseparable matrix of semiseparability rank k, the vector-Givens representation can always be constructed because there always exist Givens rotations which annihilate the elements of the lower triangular matrix up to the k diagonals in reversed order of the above construction. In the same way the vector-Givens representation of the strictly upper semiseparable matrix can be constructed and hence the representation of the whole semiseparable matrix is known.

2.3. A block-semiseparable matrix of semiseparability rank k

Here we propose a second alternative extension of semiseparable matrices of semiseparability rank 1, namely a block-version. Also for this definition a suitable representation is proposed.

Definition 3. *Starting from an $n \times n$ block-matrix S, whose blocks are of size $k \times k$, a matrix S is called* a (strictly) lower (upper) block-semiseparable matrix of semiseparability rank k, *with $k \in \mathbb{N}$, if all subblock-matrices which can be taken out of the (strictly) lower (upper) block-triangular part of the matrix S have rank $\leq k$.*

If n is not a multiple of k, but the blocks of size less than $k \times k$ can be extended to full $k \times k$ blocks such that all subblock-matrices which can be taken out of the (strictly) lower (upper) block-triangular part of this larger block-matrix have rank $\leq k$, the matrix S is also called a (strictly) lower (upper) block-semiseparable matrix of semiseparability rank k.

A lower and strictly upper block-semiseparable matrix of semiseparability rank k, is called a block-semiseparable matrix of semiseparability rank k.

During the remainder of this paper, we assume that $n = kl$. If this would be not the case, some technical details change, but we will omit them.

The representation we use for the lower block-triangular part of an $n \times n$ block-semiseparable matrix of semiseparability rank k consists of a block-vector of l $k \times k$ blocks $[D_1, D_2, \ldots, D_l]$ and $l-1$ "*generalized Givens transformations*" $G_1, G_2, \ldots, G_{l-1}$ where a generalized Givens transformation is a $2k \times 2k$ orthogonal matrix.

Starting from a block-vector and generalized Givens transformations, a block-semiseparable matrix is built completely analogously to the semiseparability rank 1 case; each time a Givens transformation was used in Subsection 2.1, a generalized Givens transformation should be used and instead of adding 1 element on the diagonal, a $k \times k$ block must be added on the block-diagonal.

Because the remainder of this paper only involves symmetric matrices, we will only have a closer look at the symmetric block-semiseparable ones.

Let us denote the generalized Givens transformations involved as

$$G_i = \begin{pmatrix} T_i & U_i \\ V_i & W_i \end{pmatrix},$$

with T_i, U_i, V_i and W_i $k \times k$ matrices. Then the corresponding symmetric block-semiseparable matrix is of the form:

$$\begin{pmatrix} T_1 D_1 & D_1^T V_1^T T_2^T & \ldots & D_1^T V_{l-2:1}^T T_{l-1}^T & D_1^T V_{l-1:1}^T \\ T_2 V_1 D_1 & T_2 D_2 & \ldots & D_2^T V_{l-2:2}^T T_{l-1}^T & D_2^T V_{l-1:2}^T \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ T_{l-1} V_{l-2:1} D_1 & T_{l-1} V_{l-2:2} D_2 & \ldots & T_{l-1} D_{l-1} & D_{l-1}^T V_{l-1}^T \\ V_{l-1:1} D_1 & V_{l-1:2} D_2 & \ldots & V_{l-1} D_{l-1} & D_l \end{pmatrix}$$

where $V_{a:b} = V_a V_{a-1} \ldots V_b$. The condition of being a symmetric matrix also leads to some extra constraints on the building elements, namely $T_i D_i$ must be equal to $D_i^T T_i^T$ for $i = 1, \ldots, l-1$ and $D_l = D_l^T$. When we start from a symmetric matrix, as we will do in Section 3, in order to construct a block-semiseparable matrix, these conditions are clearly fulfilled.

Given a block-semiseparable matrix of semiseparability rank k, a representation by means of block-vectors and generalized Givens transformations can always be found because, for the lower block-triangular part, there always exist generalized Givens transformations which annihilate the blocks up to the main block-diagonal in reversed order of the above construction. In the same way the representation of the strictly upper block-semiseparable part can be constructed and hence the representation of the whole block-semiseparable matrix is known.

In the remainder of this paper, we will construct an extension of the algorithm proposed in [9] which transforms any symmetric matrix into a semiseparable one of semiseparability rank 1 by means of orthogonal similarity transformations.

Also this algorithm can be extended to one that transforms a symmetric matrix into a semiseparable matrix of semiseparability rank k or to one that builds a block-semiseparable matrix of semiseparability rank k.

In order to keep the analogy as much as possible with the rank 1 case during the implementation, the block-semiseparable matrix suits best, so during the remainder of this paper we will focus on the block-version, but the results also hold for Definition 2 and its representation.

3. Transformation of a symmetric matrix into a block-semiseparable one of semiseparability rank k

In this section, we extend Theorem 3.1 of [9], which states that any symmetric matrix can be transformed into a semiseparable matrix of semiseparability rank 1 by means of orthogonal similarity transformations, to block-semiseparable matrices of semiseparability rank k with $k \in \mathbb{N}$. The algorithm which performs this transformation, is a block-version of the one proposed in [9].

Theorem 1. *Let A be a symmetric matrix. Then there exists an orthogonal matrix U such that*

$$U^T A U = S, \tag{1}$$

where S is a symmetric block-semiseparable matrix of semiseparability rank k.

Proof. The proof is analogous to the one of [9, Theorem 3.1]: replace matrix elements by $k \times k$ blocks and Givens transformations by generalized Givens transformations. Hence the algorithm constructed during the proof to reduce a symmetric matrix into a block-semiseparable one of semiseparability rank k, is a block-version of the algorithm proposed in [9].

At each step k rows and columns are added to the block-semiseparable structure (instead of 1) and the dimension of the block-semiseparable matrix appearing in the lower-right corner increases with k.

For completeness the mth step of the algorithm is pictured in Figure 1. Here 0 denotes a k dimensional zero matrix, $\times$ a $k \times k$ block, $\boxtimes$ a $k \times k$ block that is part of the block-semiseparable structure of semiseparability rank k and $\otimes$ a block that will be annihilated in the next step.

∎

The implementation of the algorithm constructed in the proof of Theorem 1 and of the one presented in [9] are analogous. One must only be careful with the transpose of matrices and the fact that in the rank 1 case a Givens transformation only involves c and s and in the block rank k case a generalized Givens transformation is determined by T, U, V and W. So the formula's change a bit, but the idea stays the same.

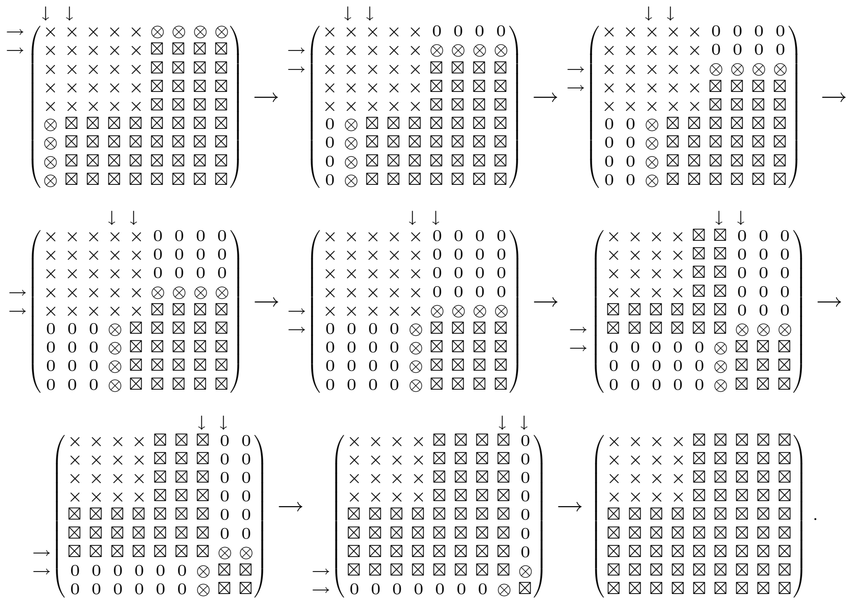

Figure 1. Description of one step of similarity transformations used to retrieve the block-semiseparable structure.

4. Properties of the reduction of a symmetric matrix into a block-semiseparable one of semiseparability rank k

When transforming a symmetric matrix into a semiseparable one of semiseparability rank 1 by orthogonal similarity transformations as proposed in [9], at each step a semiseparable matrix is constructed in the lower-right corner. The eigenvalues of this semiseparable matrix are the Ritz-values of the original symmetric matrix A obtained by the Lanczos' algorithm.

Also a kind of nested subspace iteration is performed at each step.

Both properties provide important information about the spectrum, or part of it, of the original matrix. Hence, in this section we will investigate if the reduction to block-semiseparable matrices also has these properties.

4.1. Block-Lanczos-Ritz values

During the transformation of a symmetric matrix into a block-semiseparable one of semiseparability rank k by means of orthogonal similarity transformations as proposed in the algorithm of the proof of Theorem 1, in the lower-right corner a block-semiseparable matrix is constructed whose dimension increases with k (so with 1 block-row and 1 block-column) at each step.

Also in this case at each mth step with $m = 1, \ldots, l-1$, the eigenvalues of the latter block-matrix are the Ritz-values of the starting matrix A with respect to the Krylov subspace[†]:

$$\mathcal{K}_m = < e_{n-k+1}, \ldots, e_{n-1}, e_n, Ae_{n-k+1}, \ldots, Ae_{n-1}, Ae_n, \ldots, A^m e_{n-k+1}, \ldots, A^m e_{n-1}, A^m e_n >,$$

with e_i the ith vector of the standard basis in $\mathbb{R}^n$. Hence the convergence behavior of these eigenvalues is the same as the convergence behavior of the block-Lanczos' algorithm with the initial vectors $e_{n-k+1}, \ldots, e_{n-1}, e_n$. This implies that the extreme eigenvalues are approximated first.

The difference with the rank 1 case where eigenvalues with higher multiplicity are approximated only once, is that generically every eigenvalue with multiplicity p is approximated p times if $p \leq k$ and k times if $p \geq k$.

The proof of the block-Lanczos-Ritz behavior is a trivial extension of the Lanczos-Ritz behavior in the rank 1 case because the shape of the orthogonal matrices appearing in the block-version is similar to the shape of the ones in the proof of [9, Theorem 6.1].

4.2. Subspace iteration

In order to maintain the same property of subspace iteration as for the rank 1 reduction, some extra attention must be paid because the involved orthogonal matrices don't have the proper form according to [9, Section 7.1].

Looking at the block-version of the algorithm, it can be split into two parts, first we introduce zero elements and in a second part the block-semiseparable structure is created. For example in Figure 1, the three first transformations introduce zeros and from the 4th transformation, the block-semiseparable structure is created. The way of creating zero in the first part, doesn't influence the subspace iteration property, but in each transformation of the second part not only the block before the main diagonal must be annihilated, but also the elements up to the main diagonal in order to obtain subspace iteration. So in the second part, we should use generalized Givens transformations G_i such that

$$\underbrace{\left(\begin{array}{cccc|cccc} \times & \times & \ldots & \times & \times & \times & \ldots & \times \\ \times & \times & \ldots & \times & \times & \times & \ldots & \times \\ \times & \times & \ldots & \times & \times & \times & \ldots & \times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \times & \times & \ldots & \times & \times & \times & \ldots & \times \end{array}\right)}_{k \qquad\qquad k} G_i = \underbrace{\left(\begin{array}{cccc|cccc} 0 & 0 & \ldots & 0 & \times & \times & \ldots & \times \\ 0 & 0 & \ldots & 0 & 0 & \times & \ldots & \times \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & \times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \cdots & 0 & \times \end{array}\right)}_{k \qquad\qquad k} \qquad (2)$$

Using those generalized Givens transformations, at each step of the block-algorithm, the size of the vector subspace increases with k and, as in the rank 1 case, a change of coordinate system is performed. Theorem 7.2 of [9] about the convergence speed is still valid.

When combining the influence of the block-Lanczos-Ritz convergence behavior and subspace iteration for several distributions of the eigenvalues, the eigenvalues with the largest absolute values will be first approximated with multiplicity $\leq k$. Once these approximations are sufficiently close to the exact eigenvalues of the original matrix, the subspace iteration creates blocks and the problem will be split into two parts. This behavior is illustrated in the numerical experiments of Section 6.

[†] $< x, y, z, \ldots >$ we denote the subspace spanned by the vectors $x, y, z, \ldots$.

5. Computational complexity

As mentioned in Section 4.2, at each step the algorithm of transforming a symmetric matrix into a block-semiseparable one of semiseparability rank k can be split into two parts, one of creating zeros and one of constructing a block-semiseparable structure. The generalized Givens transformations used, can be constructed with both Householder transformations as well as Givens transformations. In our implementation, we used Givens rotations with a computational cost of 6 floating point operations. For fixed k, when using during the whole algorithm the specific kind of generalized Givens transformations (2) as mentioned in Subsection 4.2 in order to maintain the subspace iteration property, the computational complexity of all first steps, which introduce zeros, is given by

$$\frac{16}{3}n^3 + O(n^2)$$

and the computational complexity of all second steps, which create the semiseparable structure, is of order

$$O\left(n^2 k\right).$$

Remark that when one executes the first part of all steps, a block-tridiagonal matrix is constructed. So an alternative reduction of a symmetric matrix into a block-semiseparable one, first transforms a symmetric matrix into a block-tridiagonal one and then reduces this block-tridiagonal matrix into a block-semiseparable matrix by means of the block-algorithm proposed in the proof of Theorem 1. During the execution of the block-algorithm, the first part of each step does not need to be applied in this case. But because the reduction of a symmetric matrix into a block-tridiagonal one uses the same orthogonal similarity transformations as the first part of the block-algorithm, both algorithms are essentially the same. The block-algorithm however has the important properties mentioned in Section 4, while first reducing into a block-tridiagonal matrix misses this advantage.

6. Numerical experiments

As stated previously, the properties of Section 4 will be illustrated by means of numerical experiments performed using Matlab 6.5 [‡].

6.1. Experiment 1

In a first experiment we illustrate the block-Lanczos-Ritz behavior of the reduction of a symmetric matrix into a block-semiseparable one of semiseparability rank k with $k = 2$.

Starting from a 200×200 matrix with 50 equidistant eigenvalues $1, 2, \ldots, 50$ each with multiplicity 4, we compare at each step of the block-algorithm the Ritz-values of the block-semiseparable matrix occurring in the lower-right corner with the eigenvalues of the original matrix. On the x-axis the number of steps are indicated, on the y-axis the Ritz-values. When a Ritz-value differs less than 10^{-3} from an original eigenvalue, it is indicated with a cross. The four graphes of Figure 2 show when the eigenvalues $1, 2 \ldots, 50$ are approximated for the first,

[‡]Matlab is a registered trademark of the Mathworks Inc.

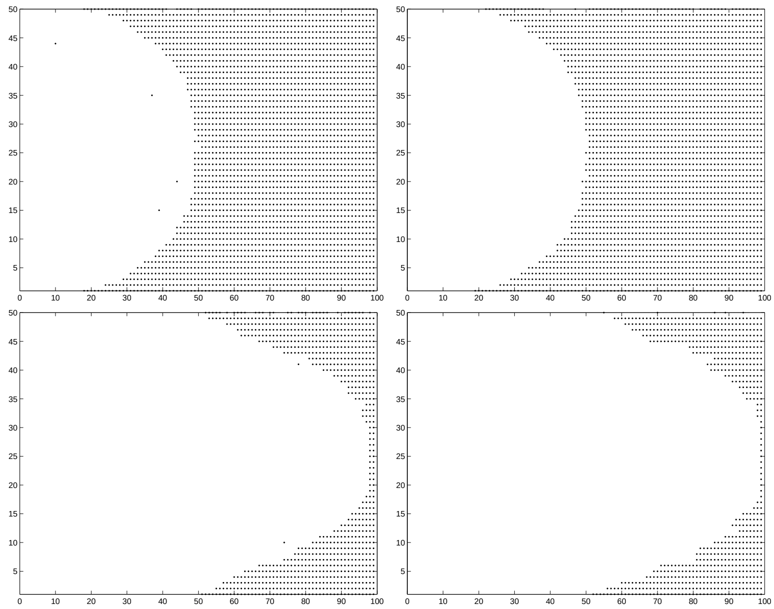

Figure 2. Experiment 1: The block-Lanczos-Ritz behavior

second, third and fourth time. Because $k = 2$ all eigenvalues are approximated twice after 50 steps. Immediately from step 51 the block-Lanczos-Ritz behavior is visible again for the approximation of all eigenvalues for the third and fourth time.

6.2. Experiment 2

In a second experiment, the convergence to the eigenvalues of largest absolute value with multiplicity $\leq k$ is illustrated and the cut off according to the subspace iteration.

The 50×50 symmetric matrix we start from, has eigenvalues $-100, -100, 100$, 100 and random ones between 0 and 1. Figure 3 indicates the elements of the similar matrix after 6 steps, when we can split the matrix into two parts. The more white the elements, the larger in absolute value, the darker the elements, the smaller they are. The cut off of a 4×4 block is visible and justified because the norm of the last 4 rows up to this 4×4 block is of order $O(10^{-8})$. The eigenvalues of the lower-right corner block are $-100, -100, 100, 100$ up to 15 exact digits.

6.3. Experiment 3

Next we illustrate the influence of k.

The starting matrix is a symmetric 20×20 matrix with eigenvalues $1, 2, 3, 4, 5$, all with multiplicity 4. We transform this matrix into a block-semiseparable one of semiseparability rank k for $k = 1, 2, 3, 4$. At each step we calculate the norm of the matrix formed by the rows left from the block-semiseparable matrix in the lower-right corner. If the latter norm is small, a cut off will occur. For $k = 1$, this cut off will occur when each eigenvalue is approximated

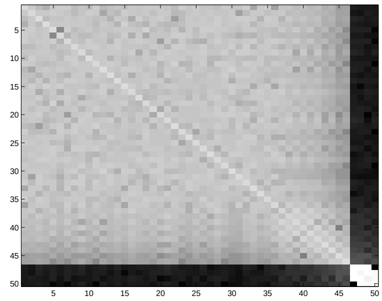

Figure 3. Experiment 2

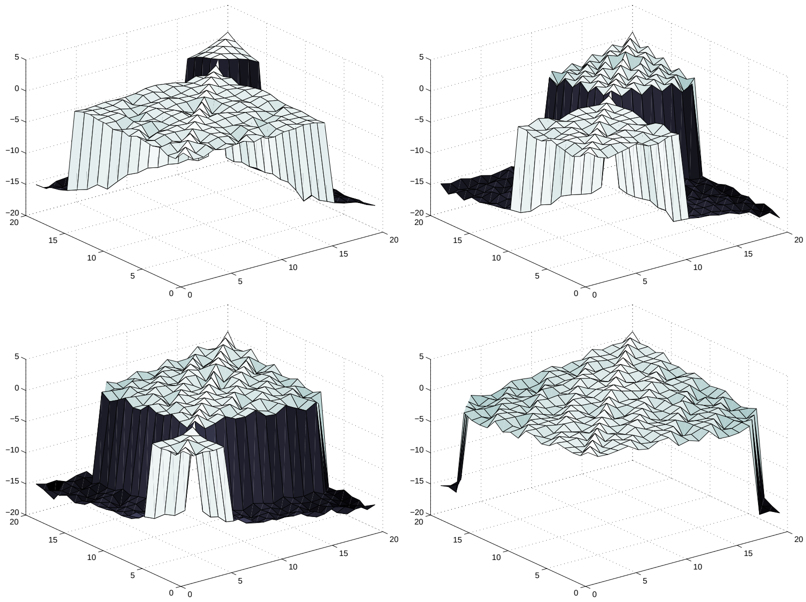

Figure 4. Experiment 3: Similar matrix when cut off occurs for $k = 1, 2, 3, 4$

once, for $k = 2$ when each eigenvalue is approximated with multiplicity 2 and so on. This is clearly shown in Figure 4 where for each k the elements of the transformed matrix are shown in a logarithmic scale at the moment of the cut off.

For $k = 2$, at each step the norm of the matrix formed by the rows left from the block-semiseparable lower-right corner is calculated, as well as the block-Lanczos-Ritz values $\hat{\lambda}_i$ of the block-semiseparable matrix and their relative error $\frac{\|\lambda_i - \hat{\lambda}_i\|_2}{\|\lambda_i\|_2}$ with λ_i the exact eigenvalues of the starting matrix. Table I shows this results, obtained by Matlab, until the cut off occurs.

6.4. Experiment 4

The last example illustrates the connection between k and the multiplicity of the eigenvalues.

Table I. Numerical results for reduction into a block-semiseparable matrix of semiseparability rank 2

Step	Norm	Block-Lanczos-Ritz values	Relative error
1	1.277933643059884e+000	4.616801541416155e+000	7.663969171676896e-002
		4.024862703517305e+000	6.215675879326144e-003
		1.357878124560564e+000	3.578781245605636e-001
		1.648933314952031e+000	1.755333425239846e-001
2	1.045549327739640e+000	4.910846663115344e+000	1.783066737693115e-002
		4.722369379378176e+000	5.552612412436488e-002
		2.976377210260876e+000	7.874263246374635e-003
		2.744928200595344e+000	8.502393313488543e-002
		1.087490150518819e+000	8.749015051881903e-002
		1.184248032069537e+000	1.842480320695372e-001
3	8.219369256484070e-001	4.944268596276749e+000	1.114628074465021e-002
		4.987978306370172e+000	2.404338725965616e-003
		3.748786638252343e+000	6.280334043691438e-002
		3.558250112432622e+000	1.860833708108740e-001
		1.010480802847837e+000	1.048080284783648e-002
		1.031178421531595e+000	3.117842153159511e-002
		2.167536895263825e+000	8.376844763191249e-002
		2.258631957363070e+000	2.471226808789765e-001
4	3.345343300927901e-014	1.000000000000001e+000	6.661338147750939e-016
		1.000000000000003e+000	2.664535259100376e-015
		5.000000000000002e+000	3.552713678800501e-016
		4.999999999999996e+000	7.105427357601002e-016
		1.999999999999998e+000	1.110223024625157e-015
		2.000000000000001e+000	4.440892098500626e-016
		3.000000000000005e+000	1.776356839400251e-015
		3.000000000000002e+000	7.401486830834377e-016
		4.000000000000002e+000	4.440892098500626e-016
		3.999999999999999e+000	3.330669073875470e-016

When transforming a symmetric matrix with eigenvalues $1, 1, 1, 2, 3, 3, 3, 3, 4, 5, 5, 6, 6, 6, 7, 8, 8, 9, 9, 10$ into a block-semiseparable matrix with $k = 2$, a first block will be found with all eigenvalues with the exact multiplicity if their multiplicity is ≤ 2 and with multiplicity 2 otherwise. Figure 5 shows the block-Lanczos-Ritz values of the block-semiseparable matrix at each step until cut off occurs after 7 steps.

7. Conclusion

Any symmetric matrix can be transformed into a block-semiseparable matrix of semiseparability rank k by means of orthogonal similarity transformations. During this transformation block-semiseparable matrices are constructed whose eigenvalues have the block-Lanczos-Ritz convergence behavior and a kind of nested subspace iteration is performed.

REFERENCES

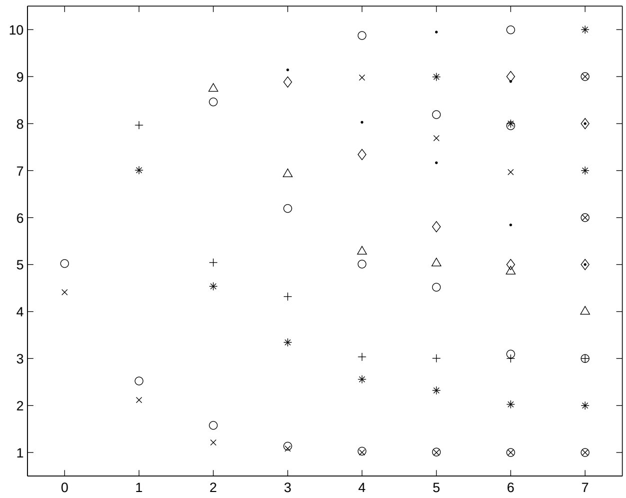

Figure 5. Experiment 4

1. Aliaga JI, Boley DL, Freund RW, Hernández V. A Lanczos-type method for multiple starting vectors. *Math. Comp.* 2000; **69**:1577–1601.
2. Braun M, Sofianos SA, Papageorgiou DG, Lagaris IE. An efficient Chebyshev Lanczos method for obtaining eigensolutions of the Schrödinger equation on a grid. *Journal of Computational Physics* 1996; **126**:315–327.
3. Cullum JK, Donath WE. A block Lanczos algorithm for computing the q algebraically largest eigenvalues and a corresponding eigenspace for large, sparse symmetric matrices. In *Proc. 1974 IEEE Conference on Decision and Control,* IEEE Press: New York, 1974; 505–509.
4. Cullum JK, Willoughby RA. *Lanczos algorithms for large symmetric eigenvalue computations, Volume 1, Theory.* Birkhäuser: Basel, 1995.
5. Golub GH, Underwood R. The block Lanczos method for computing eigenvalues. In *Mathematical Software III,* Rice JR (eds). Academic Press: New York, 1977; 361–377.
6. Joliffe IT. *Principal component analysis* Springer Verlag, 1986.
7. Ruhe A. Implementation aspects of band Lanczos algorithms for computation of eigenvalues of large sparse symmetric matrices. *Math. Comp.* 1979; **33**:680–687.
8. Vandebril R, Van Barel M, Mastronardi N. A note on the representation and definition of semiseparable matrices. Report TW 368, Dept. of Computer Science, Katholieke Universiteit Leuven, Belgium, 2003.
9. Van Barel M, Vandebril R, Mastronardi N. The Lanczos-Ritz values appearing in an orthogonal similarity reduction of a matrix into semiseparable form. Report TW 360, Dept. of Computer Science, Katholieke Universiteit Leuven, Belgium, 2003.
10. Wall ME, Dyck PA, Brettin TS. Svdman - singular value decomposition analysis of microarray data. *Bioinformatics* 2001; **17**(6): 566–568.
11. Wall ME, Rechtsteiner A, Rocha LM. Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis,* Berrar DP, Dubitzky W, Granzow M (eds). Kluwer, 2003.
12. Yeung MKS, Ruzzo WL. Principal component analysis for clustering gene expression data. *Bioinformatics* 2001; **17**(9): 763–774.
13. Yeung MKS, Tegnér J, Collins JJ. *Reverse engineering gene networks using singular value decomposition and robust regression,* PNAS 2002; 99: 6163–6168.