Computing Orthogonal Decompositions of Block Tridiagonal or Banded Matrices

Wilfried N. Gansterer

Institute of Distributed and Multimedia Systems,
University of Vienna, Austria

Abstract. A method for computing orthogonal URV/ULV decompositions of block tridiagonal (or banded) matrices is presented. The method discussed transforms the matrix into structured triangular form and has several attractive properties: The block tridiagonal structure is fully exploited; high data locality is achieved, which is important for high efficiency on modern computer systems; very little fill-in occurs, which leads to no or very low memory overhead; and in most practical situations observed the transformed matrix has very favorable numerical properties. Two variants of this method are introduced and compared.

1 Introduction

In this paper, we propose a method for computing an orthogonal decomposition of an irreducible symmetric block tridiagonal matrix

$$M_p := \begin{pmatrix} B_1 & C_1 & & & \\ A_1 & B_2 & C_2 & & \\ & A_2 & B_3 & \ddots & \\ & & \ddots & \ddots & C_{p-1} \\ & & & A_{p-1} & B_p \end{pmatrix} \in \mathbb{R}^{n\times n} \tag{1}$$

with $p > 1$. The blocks $B_i \in \mathbb{R}^{k_i\times k_i}$ $(i = 1, 2, \ldots, p)$ along the diagonal are quadratic (but not necessarily symmetric), the off-diagonal blocks $A_i \in \mathbb{R}^{k_{i+1}\times k_i}$ and $C_i \in \mathbb{R}^{k_i\times k_{i+1}}$ $(i = 1, 2, \ldots, p-1)$ are arbitrary. The block sizes k_i satisfy $1 \le k_i < n$ and $\sum_{i=1}^{p} k_i = n$, but are otherwise arbitrary.

We emphasize that the class of matrices of the form (1) comprises *banded* symmetric matrices, in which case the C_i are upper triangular and $C_i = A_i^\top$. Alternatively, given a banded matrix with upper and lower bandwidth b, a block tridiagonal structure is determined by properly selecting the block sizes k_i.

Motivation. Banded matrices arise in numerous applications. We also want to highlight a few situations where block tridiagonal matrices occur. One example from acoustics is the modelling of vibrations in soils and liquids with different

V.S. Sunderam et al. (Eds.): ICCS 2005, LNCS 3514, pp. 25–32, 2005.

layers using finite elements [1]. Every finite element is only linked to two other elements (a deformation at the bottom of one element influences the deformation at the top of the next element). Consequently, when arranging the local finite element matrices into a global system matrix, there is an overlap between these local matrices resulting in a block tridiagonal global matrix. Another example is the block tridiagonalization procedure introduced by Bai et al. [2]. Given a symmetric matrix, this procedure determines a symmetric block tridiagonal matrix whose eigenvalues differ at most by a user defined accuracy tolerance from the ones of the original matrix.

In this paper, we present two orthogonal factorizations of M_p—a URV and a ULV factorization, where U and V are orthogonal matrices, and R and L have special upper and lower triangular structure, respectively. These decompositions are important tools when block tridiagonal or banded linear systems have to be solved. This aspect is discussed in more detail in Section 3.

Although many research activities on ULV/URV decompositions have been documented in the literature (see, for example, [3, 4, 5]), we are not aware of work specifically targeted towards block tridiagonal matrices. A distinctive feature of the approach described in this paper is that it fully exploits the special structure defined in Eqn. (1).

Synopsis. The algorithm for computing the two orthogonal decompositions and their properties are described in Section 2, important applications are summarized in Section 3, and concluding remarks are given in Section 4.

2 Factorization of Block Tridiagonal Matrices

In this section, we summarize two closely related methods for computing an orthogonal factorization of the block tridiagonal matrix M_p defined by Eqn. (1).

The basic idea behind the algorithms is to eliminate off-diagonal blocks one after the other by computing singular value decompositions of submatrices and performing the corresponding update. Depending on how the submatrices are chosen, there are two basic variants of this factorization, a URV and a ULV factorization. In the first one, M_p is transformed into upper triangular structure, whereas in the second one, M_p is transformed into lower triangular structure. A comparison between the two variants for specific applications is given in Section 3.2.

2.1 URV Decomposition

Based on the SVD of the block comprising the first diagonal and the first subdiagonal block of M_p,

$$\begin{pmatrix} B_1 \\ A_1 \end{pmatrix} = U_1 \begin{pmatrix} \Sigma_1 \\ \mathbf{0} \end{pmatrix} V_1^\top = \begin{pmatrix} U_1^1 & U_1^3 \\ U_1^2 & U_1^4 \end{pmatrix} \begin{pmatrix} \Sigma_1 \\ \mathbf{0} \end{pmatrix} V_1^\top$$

we can transform M_p into

$$\begin{pmatrix} \Sigma_1 & \tilde{C}_1 & \tilde{F}_1 & & & \\ \mathbf{0} & \tilde{B}_2 & \tilde{C}_2 & & & \\ & A_2 & B_3 & \ddots & & \\ & & \ddots & \ddots & C_{p-1} \\ & & & A_{p-1} & B_p \end{pmatrix}$$

by updating with $U_1^\top$ from the left and with V_1 from the right. The existing blocks C_1, B_2, and C_2 are modified in this process (indicated by a tilde),

$$\tilde{C}_1 = U_1^{1\top} C_1 + U_1^{2\top} B_2 \tag{2}$$

$$\tilde{B}_2 = U_1^{3\top} C_1 + U_1^{4\top} B_2 \tag{3}$$

$$\tilde{C}_2 = U_1^{4\top} C_2, \tag{4}$$

and one block in the second upper diagonal is filled in:

$$\tilde{F}_1 = U_1^{2\top} C_2. \tag{5}$$

Next, we continue with the SVD of the subblock comprising the updated diagonal block and the subdiagonal block in the second block column:

$$\begin{pmatrix} \tilde{B}_2 \\ A_2 \end{pmatrix} = U_2 \begin{pmatrix} \Sigma_2 \\ \mathbf{0} \end{pmatrix} V_2^\top .$$

Again, we perform the corresponding updates on $\tilde{C}_2$, B_3, C_3, the fill-in of $\tilde{F}_2$ with $U_2^\top$ from the left, and the update of the second block column with V_2 from the right. Then, we compute the SVD of

$$\begin{pmatrix} \tilde{B}_3 \\ A_3 \end{pmatrix},$$

update with $U_3^\top$ and with V_3, and so on.

After $p-1$ such steps consisting of SVD and update operations, followed by the SVD of the diagonal block in the bottom right corner

$$\tilde{B}_p = U_p \Sigma V_p^\top$$

we get a factorization

$$M_p = URV^\top \tag{6}$$

with upper triangular R (see Fig. 1). More specifically,

$$R = \begin{pmatrix} \Sigma_1 & \tilde{C}_1 & \tilde{F}_1 & & & \\ & \Sigma_2 & \tilde{C}_2 & \tilde{F}_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \Sigma_{p-2} & \tilde{C}_{p-2} & \tilde{F}_{p-2} \\ & & & & \Sigma_{p-1} & \tilde{C}_{p-1} \\ & & & & & \Sigma_p \end{pmatrix}, \tag{7}$$

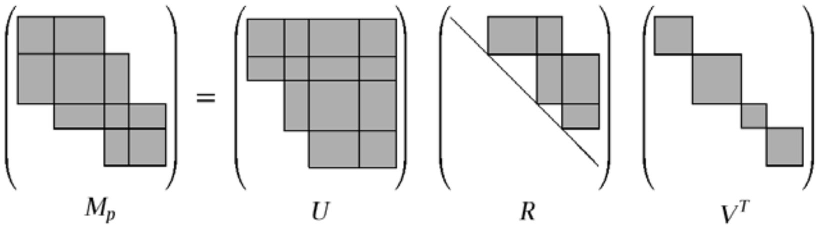

Fig. 1. URV decomposition of M_p

U is orthogonal and represented as the product

$$U = \begin{pmatrix} U_1^1 & U_1^3 & & & & \\ U_1^2 & U_1^4 & & & & \\ & & I & & & \\ & & & I & & \\ & & & & \ddots & \\ & & & & & I \end{pmatrix} \times \begin{pmatrix} I & & & & & \\ & U_2^1 & U_2^3 & & & \\ & U_2^2 & U_2^4 & & & \\ & & & I & & \\ & & & & \ddots & \\ & & & & & I \end{pmatrix} \times \cdots$$

$$\cdots \times \begin{pmatrix} I & & & & & \\ & I & & & & \\ & & \ddots & & & \\ & & & I & & \\ & & & & U_{p-1}^1 & U_{p-1}^3 \\ & & & & U_{p-1}^2 & U_{p-1}^4 \end{pmatrix} \times \begin{pmatrix} I & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & I & & \\ & & & & I & \\ & & & & & U_p \end{pmatrix},$$

and V is also orthogonal and block diagonal:

$$V = \text{block-diag}(V_1, V_2, \ldots, V_p).$$

The correctness of this decomposition algorithm can be verified directly by multiplying out (6) and taking into account the respective relationships corresponding to Eqns. (2, 3, 4, 5).

2.2 ULV Decomposition

We may as well start the process described in the previous section with the SVD of the block comprising the first diagonal and the first *super*diagonal block of M_p,

$$\begin{pmatrix} B_1 & C_1 \end{pmatrix} = U_1 \begin{pmatrix} \Sigma_1 & \mathbf{0} \end{pmatrix} V_1^\top = U_1 \begin{pmatrix} \Sigma_1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} V_1^1 & V_1^3 \\ V_1^2 & V_1^4 \end{pmatrix}^\top .$$

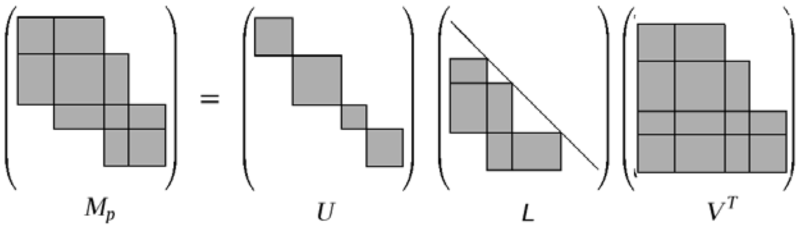

Fig. 2. ULV decomposition

The corresponding updates with $U_1^\top$ from the left and with V_1 from the right will transform M_p into

$$\begin{pmatrix} \Sigma_1 & \mathbf{0} & & & \\ \tilde{A}_1 & \tilde{B}_2 & C_2 & & \\ \tilde{F}_1 & A_2 & B_3 & \ddots & \\ & & \ddots & \ddots & C_{p-1} \\ & & & A_{p-1} & B_p \end{pmatrix}.$$

Next, we compute the SVD

$$\left(\tilde{B}_2 \; C_2 \right) = U_2 \left(\Sigma_2 \; \mathbf{0} \right) V_2^\top,$$

update again, compute the next SVD in the next block row, etc.

After $p-1$ such steps, again consisting of SVD and update operations, followed by the SVD of the diagonal block in the bottom right corner

$$\tilde{B}_p = U_p \Sigma V_p^\top,$$

this time we get a factorization

$$M_p = ULV^\top \tag{8}$$

with *lower* triangular L and orthogonal U and V (see Fig. 2).

2.3 Properties

The decompositions produced by the algorithms described in the previous section have the following important properties.

- $p-1$ lower or upper off-diagonal blocks are “pushed” onto the other side of the diagonal blocks, filling in the $p-2$ blocks F_i in the second upper or lower block diagonal. This implies that there is no increase in memory requirements, except when the transformation matrices U or V need to be stored.

- The blocks along the diagonal of R and L are diagonal matrices, whose entries are non-increasing in each block (because they are computed as the singular values of another matrix).
- In the URV decomposition, explicit representation of U requires considerable computational effort because of the overlap between consecutive block rows, whereas explicit representation of V is available directly because of its simple block diagonal structure.
- In the ULV decomposition, we have the opposite situation: Analogously to the previous argument, explicit representation of V requires considerable computational effort, whereas explicit representation of U is available directly.
- The factorizations (6) and (8) are not guaranteed to be rank-revealing (cf. [3]). However, our experience and current work shows that in practice it turns out to be rank-revealing in most cases and that it is very useful even if it turns out *not* to be rank-revealing [6].
- The arithmetic complexity of the algorithms described in Sections 2.1 and 2.2 is roughly $c_1 n \max_i k_i^2 + c_2 \max_i k_i^3$ with moderate constants c_1 and c_2. Consequently, the factorizations (6) and (8) can be computed highly efficiently, especially so for small block sizes k_i ("narrow" M_p).

3 Applications

The investigation of various applications of the orthogonal decompositions described in this paper is work in progress. In the following, we summarize the basic directions of this approach.

Clearly, one idea is to utilize the decompositions described here in the context of solving linear systems. Compared to standard methods for solving block tridiagonal (in particular, banded) linear systems, the approaches proposed in this paper do not involve pivoting. Consequently, they are expected to have important advantages in terms of efficiency in most practical situations. Some rare numerically "pathological" cases, which may require slightly more expensive techniques, are currently subject of intense research.

Our motivation comes from very specific linear systems arising in the context of the *block divide-and-conquer* eigensolver.

3.1 Block Divide-and-Conquer Eigensolver

The block divide-and-conquer (*BD&C*) eigensolver [7, 8] was developed in recent years for approximating the eigenpairs of *symmetric* block tridiagonal matrices M_p (B_i symmetric and $C_i = A_i^\top$ in Eqn. (1)).

This approach proved to be highly attractive for several reasons. For example, it does not require tridiagonalization of M_p, it allows to approximate eigenpairs to arbitrary accuracy, and it tends to be very efficient if big parts of the spectrum need to be approximated to medium or low accuracy. However, in some situations eigenvector accumulation may become the bottleneck of BD&C—when accuracy requirements are not low and/or when not the entire spectrum, but only small parts of it have to be approximated.

With this in mind, we are currently developing a new approach for computing approximate eigenvectors of a symmetric block tridiagonal matrix as an alternative to eigenvector accumulation [6]. This new approach has the potential of significantly improving arithmetic complexity, proportionally reducing the computational effort if only a (small) part of the spectrum but not the full spectrum needs to be computed, and being highly parallelizable and scalable for large processor numbers.

3.2 Eigenvector Computation

We know that the eigenvector x corresponding to a eigenvalue λ of M_p is the solution of

$$(M_p - \lambda I)\, x = \mathbf{0}. \tag{9}$$

Since the shift operation preserves block tridiagonal structure, we can use one of the algorithms presented in this paper compute an orthogonal decomposition of the matrix $S := M_p - \lambda I$ and exploit it for solving the linear system (9) efficiently.

URV vs. ULV. In this specific context, there is a clear preference for one of the variants discussed in Section 2, as outlined in the following.

If we want to solve Eqn. (9) based on one of the decompositions described in this paper, we need to first solve $Ry = \mathbf{0}$ or $Ly = \mathbf{0}$ for the URV or the ULV decomposition, respectively. Then, we need to form $x = Vy$ to find a solution of Eqn. (9).

As discussed in Section 2.3, in the URV decomposition explicit representation of V is available directly and it has simple block diagonal structure, whereas multiplication with V from the ULV decomposition requires considerably higher computational effort. This clearly shows that the URV decomposition is preferable for this application context.

Backsubstitution. The rich structure of the matrix R resulting from the orthogonal decomposition described in this paper obviously allows for a very efficient backsubstitution process when solving a system $Ry = b$.

In the special context of eigenvector computations, many specific numerical aspects have to be taken into account in order to achieve certain requirements in terms of residuals and orthogonality properties (cf. [9] and several publications resulting from this work).

4 Conclusion

We have presented two related algorithms for computing orthogonal URV or ULV decompositions of general block tridiagonal matrices. We have shown that the matrices R and L can not only be computed efficiently, but that they also have very attractive structural properties.

Because of these properties the methods introduced in this paper are expected to have numerous applications. In particular, we expect them to be integral part of a new approach for computing approximate eigenvectors of symmetric block tridiagonal matrices.

Current and Future Work

Our current work focusses on reliable and efficient methods exploiting the decompositions presented in this paper for approximating eigenvectors in the context of the block divide-and-conquer eigensolver for symmetric block tridiagonal matrices. This goal requires careful investigation of several important numerical aspects of the decompositions described here, especially in the backsubstitution process. Our findings will be summarized in [6].

Moreover, we are also investigating other application contexts for the decompositions presented here.

Acknowlegdements

We would like to thank Robert C. Ward for many inspiring discussions when developing the concepts presented here and the anonymous referees for their comments.

References

1. Acoustics Research Institute of the Austrian Academy of Sciences: Vibrations in soils and liquids – transform method for layers with random properties. FWF-Project P16224-N07 (2004)
2. Bai, Y., Gansterer, W.N., Ward, R.C.: Block tridiagonalization of "effectively" sparse symmetric matrices. ACM Trans. Math. Softw. **30** (2004) 326–352
3. Stewart, G.W.: Updating a rank-revealing ULV decomposition. SIAM J. Matrix Anal. Appl. **14** (1993) 494–499
4. Park, H., Elden, L.: Downdating the rank-revealing URV decomposition. SIAM J. Matrix Anal. Appl. **16** (1995) 138–155
5. Stewart, M., Van Dooren, P.: Updating a generalized URV decomposition. SIAM J. Matrix Anal. Appl. **22** (2000) 479–500
6. Gansterer, W.N., Kreuzer, W., Ward, R.C.: Computing eigenvectors of block tridiagonal matrices. (2005) *in preparation.*
7. Gansterer, W.N., Ward, R.C., Muller, R.P.: An extension of the divide-and-conquer method for a class of symmetric block-tridiagonal eigenproblems. ACM Trans. Math. Softw. **28** (2002) 45–58
8. Gansterer, W.N., Ward, R.C., Muller, R.P., Goddard, III, W.A.: Computing approximate eigenpairs of symmetric block tridiagonal matrices. SIAM J. Sci. Comput. **25** (2003) 65–85
9. Dhillon, I.S.: A New $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/-Eigenvector Problem. PhD thesis, Computer Science Division (EECS), University of California at Berkeley (1997)