

TWISTED FACTORIZATION OF A BANDED MATRIX

JASON SLEMONS[†] AND CHRISTOF VÖMEL[‡]

Friday 16th May, 2008, 0:13am

Abstract. The twisted factorization of a tridiagonal matrix plays an important role in inverse iteration as featured in the MRRR algorithm. The twisted structure simplifies the computation of the eigenvector approximation and can also improve the accuracy.

This paper gives a constructive proof for the existence of the twisted factorizations of a general banded matrix and investigates the implications on inverse iteration.

AMS subject classifications. 65F15, 65Y15.

Key words. Banded matrix, block-tridiagonal matrix, double factorization, twisted factorization.

1. Introduction. Let $A \in \mathcal{R}^{n \times n}$ denote a banded matrix with semi-bandwidth $b > 0$, for example a tridiagonal ($b := 1$) or a pentadiagonal ($b := 2$). We do not require A to be symmetric, but we do assume for now that it has the same number of upper and lower off-diagonal bands.

We further assume existence of the double factorization

$$A = L_+ D_+ U_+ = U_- D_- L_-, \quad (1.1)$$

with L_+, L_- and U_+, U_- respectively being lower and upper triangular with unit diagonal, and with D_+ and D_- containing the pivots of the ‘forward’ and ‘backward’ factorizations.

We are interested in the existence of a twisted factorization

$$A = M_k \Delta_k N_k, \quad (1.2)$$

where Δ_k is diagonal, and the k -th column of M_k and the k -th row of N_k correspond to the k -th column and row of the identity, that is

$$M_k e_k = e_k, \quad e_k^t N_k = e_k^t. \quad (1.3)$$

So far, a twisted factorization of this kind has mainly been considered for tridiagonal matrices, see for example [5, 6, 10] and also [6, Section 4] for historic references. One of its important applications lies in the computation of eigenvectors. Indeed, the symmetric tridiagonal twisted factorization is at the core of the MRRR algorithm [1, 2, 3, 4, 6, 7] because it allows us to compute a good approximation to the eigenvector of a relatively isolated eigenvalue. Consider for example inverse iteration to compute an approximate right eigenvector of a symmetric tridiagonal matrix. If $\gamma_k := e_k^t \Delta_k e_k$, then from (1.2) and (1.3), one has for the special right-hand side e_k that

$$Az = \gamma_k e_k \Leftrightarrow N_k z = e_k,$$

[†]Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA. slemons@amath.washington.edu

[‡]Institute of Computational Science, ETH Zürich, CAB, Universitätsstraße 6, 8092 Zürich, Switzerland, cvomel@inf.ethz.ch

THEOREM 2.1. *Let $A \in \mathcal{R}^{n \times n}$ denote a banded matrix with semi-bandwidth $b > 0$ and partitioned as in (2.2). Provided that the factorizations (2.3), (2.4), and (2.7) exist, both (2.8) and (2.9) constitute non-blocked twisted factorizations of A .*

Let us briefly compare to the symmetric tridiagonal case investigated by Parlett & Dhillon [6] where the dimension of A_{22} is one. There is only a single twisted factorization per index k , and the Schur complement is a scalar γ_k .

2.2. Connection to A^{-1} . As indicated in [6, Section 7], the block twisted factorization (2.5) of the block-partitioned matrix A from (2.2) allows several straightforward block-extensions of results on the tridiagonal twisted factorization from [6].

THEOREM 2.2. *Let A be as in (2.2).*

- *If $D_+(2, 2) := A_{22} - A_{21}A_{11}^{-1}A_{12}$ denotes the second block pivot from its forward block factorization, and $D_-(2, 2) := A_{22} - A_{23}A_{33}^{-1}A_{32}$ denotes the second block pivot from its backward block factorization, then*

$$A'_{22} = D_+(2, 2) + D_-(2, 2) - A_{22}. \quad (2.10)$$

- *If A is nonsingular, then*

$$[A'_{22}]^{-1} = (A^{-1})_{22}. \quad (2.11)$$

If the block double factorization of A exists, then

$$\text{blockdiag}(A) + [\text{blockdiag}(A^{-1})]^{-1} = \text{blockdiag}(D_+(\cdot, \cdot)) + \text{blockdiag}(D_-(\cdot, \cdot)). \quad (2.12)$$

It is interesting to compare this to the non-blocked factorizations of a band matrix, as for example those in (2.8) and (2.9).

THEOREM 2.3. *Let $A \in \mathcal{R}^{n \times n}$ as in (2.2), with $n = p + b + q$, $\dim(A_{11}) = p$, $\dim(A_{22}) = b$, $\dim(A_{33}) = q$. As in (2.7), let the double factorization*

$$A'_{22} = L_{2+}D_{2+}U_{2+} = U_{2-}D_{2-}L_{2-}. \quad (2.13)$$

Then

$$\mu_{p+1}^{-1} := e_{p+1}^t A^{-1} e_{p+1} = e_1^t D_{2-}^{-1} e_1 \quad (2.14)$$

and

$$\nu_{p+b}^{-1} := e_{p+b}^t A^{-1} e_{p+b} = e_b^t D_{2+}^{-1} e_b, \quad (2.15)$$

that is the respective last pivots of the back- and forward factorizations are reciprocals of the diagonal elements of A^{-1} . (The other entries of D_{2+} and D_{2-} are not as directly related to the diagonal of A^{-1} .)

In the tridiagonal case [6, Corollary 7], one could explicitly construct a product representation of the scalar γ_k at the twist index in terms of the pivots from the double factorization. Since μ_{p+1} and ν_{p+b} do not arise directly but come from a second-level triangular factorization of the Schur complement, such a product representation is no longer possible for general banded matrices. However, the twisted factorizations can still be used advantageously for eigenvector computations as we will discuss in Section 2.3.

2.3. Connection to eigenvectors. The twisted factorizations (2.8) and (2.9) enable us to stably and efficiently compute eigenvector approximations where the right-hand side is a row or column of the identity.

For example, let k denote a twist index and let $A = M_k \Delta_k N_k$ as in (2.8). Further, let ν_k denote the last entry of D_{2+} . Then the approximate right eigenvector z_r

$$Az_r = \nu_k e_k \Leftrightarrow \begin{bmatrix} U_+ & D_+^{-1} L_+^{-1} A_{12} \\ & U_{2+} \\ & D_-^{-1} U_-^{-1} A_{32} & L_- \end{bmatrix} z_r = e_k. \quad (2.16)$$

The k -th row of the matrix on the right-hand side is a row of the identity, hence $z_r(k) = 1$. Moreover, the $(1 : 2, 1 : 2)$ block-submatrix is right upper triangular, so that components $z_r(1 : k-1)$ can be computed using backward substitution. The semi-bandwidth of A determines the complexity of this computation. In the tridiagonal case, $z_r(1 : k-1)$ can be obtained using simply multiplications, in the more general case it is a difference of b terms. Once $z_r(k-b+1 : k)$ have been obtained, one can compute $z_r(k+1 : n)$ via forward substitution.

For an approximate left eigenvector z_l^* , using analogous notation, one finds

$$z_l^* A = \nu_k e_k^* \Leftrightarrow z_l^* \begin{bmatrix} L_+ \\ A_{21} U_+^{-1} D_+^{-1} & L_{2+} & A_{23} L_-^{-1} D_-^{-1} \\ & & U_- \end{bmatrix} = e_k^*. \quad (2.17)$$

Thus again $y(k) = 1$. One can compute $y(1 : k-1)$ using backward substitution. At last, $y(k+1 : n)$ follows from $y(k-b+1 : k)$ via forward substitution.

If A is normal, its unitary eigen-decomposition $A = V \Lambda V^*$ allows us to find us via connections to the inverse (Section 2.2) useful expressions for the pivots at the twist index: let λ_j denote the simple eigenvalue of A closest to zero, then from (2.14)

$$\frac{1}{\nu_k} = e_k^* V \Lambda^{-1} V^* e_k = \frac{|v_j(k)|^2}{\lambda_j} + \sum_{i \neq j} \frac{|v_i(k)|^2}{\lambda_i}. \quad (2.18)$$

As a consequence, one has for shifted $A - \sigma I$ the following

THEOREM 2.4. [6, Lemma 13] *Let $A - \sigma I$ be a normal, invertible matrix, and let (2.8) exist. Suppose that σ is closer to the simple eigenvalue λ_j than to any other eigenvalue of A . Then if $v_j(k) \neq 0$, one has for z_r from (2.16) that*

$$\frac{|\nu_k|}{\|z_r\|_2} = \frac{|\lambda_j - \sigma|}{|v_j(k)|} [1 + (|v_j(k)|^{-2} - 1) \mathcal{A}]^{-1/2} \leq \frac{|\lambda_j - \sigma|}{|v_j(k)|}. \quad (2.19)$$

Here \mathcal{A} is a weighted arithmetic mean of $\left\{ \left| \frac{\lambda_i - \sigma}{\lambda_i - \sigma} \right|^2, i \neq j \right\}$, $0 \leq |\mathcal{A}| < \frac{|\lambda_j - \sigma|^2}{gap^2(\sigma)}$, where $gap(\sigma) = \min_{i \neq j} |\lambda_i - \sigma|$.

By (2.16), the quotient $|\nu_k|/\|z_r\|_2$ on the left-hand side of (2.19) is exactly the scaled residual norm of z_r with respect to σ . The bound on the right-hand side of (2.19) is tightest when the twist index k corresponds to the largest entry of the eigenvector v . By (2.18), this corresponds to the ν_k smallest in magnitude.

For a non-normal matrix A , let λ denote a simple, generally complex, eigenvalue of A closest to zero and y and x its associated left and right eigenvectors. With the bi-orthonormality condition $(y, Y_2)^*(x, X_2) = (x, X_2)^*(y, Y_2) = I$, one has the spectral

decomposition [8, pp. 244-245]

$$A = (x, X_2) \begin{pmatrix} \lambda & \\ & L_2 \end{pmatrix} \begin{pmatrix} y^* \\ Y_2^* \end{pmatrix} = \lambda xy^* + X_2 L_2 Y_2^*.$$

Again, (2.14) yields

$$\frac{1}{\nu_k} = \frac{x(k)\bar{y}(k)}{\lambda} + e_k^* X_2 L_2^{-1} Y_2^* e_k. \quad (2.20)$$

If λ is complex, then ν_k will be complex as well. Provided that λ is well separated from the spectrum of L_2 and $x(k)\bar{y}(k) \neq 0$, a small $|\nu_k|$ corresponds to a large product $|x(k)\bar{y}(k)|$.

3. A remark on the twisted factorization of a general block matrix. Extending the procedure from Section 2.1, we show the existence of twisted factorizations of a matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}. \quad (3.1)$$

The difference from (2.2) are the generally non-vanishing blocks A_{31} and A_{13} . The blocks on the diagonal are assumed to have square shape. If A is a banded matrix, the dimension of A_{22} is the minimum of the lower and upper band-width of A .

If the forward factorization

$$A_{11} = L_+ D_+ U_+ \quad (3.2)$$

exists and is nonsingular, we can write A from (3.1) as

$$\begin{bmatrix} L_+ & & \\ A_{21} U_+^{-1} D_+^{-1} & I & \\ A_{31} U_+^{-1} D_+^{-1} & & I \end{bmatrix} \cdot \begin{bmatrix} D_+ & & \\ & A'_{22} & A'_{23} \\ & A'_{32} & A'_{33} \end{bmatrix} \cdot \begin{bmatrix} U_+ & D_+^{-1} L_+^{-1} A_{12} & D_+^{-1} L_+^{-1} A_{13} \\ & I & \\ & & I \end{bmatrix}, \quad (3.3)$$

with

$$A'_{ij} = A_{ij} - A_{i1} U_+^{-1} D_+^{-1} L_+^{-1} A_{1j}, \quad 2 \leq i, j \leq 3. \quad (3.4)$$

Continuing, let

$$A'_{33} = U'_- D'_- L'_- \quad (3.5)$$

be nonsingular, then we have from (3.3) that $A =$

$$\begin{bmatrix} L_+ & & & \\ A_{21} U_+^{-1} D_+^{-1} & I & A'_{23} L_-^{-1} D_-^{-1} & \\ A_{31} U_+^{-1} D_+^{-1} & & U'_- & \end{bmatrix} \cdot \begin{bmatrix} D_+ & & \\ & A''_{22} & \\ & & D'_- \end{bmatrix} \cdot \begin{bmatrix} U_+ & D_+^{-1} L_+^{-1} A_{12} & D_+^{-1} L_+^{-1} A_{13} \\ & I & \\ D_-^{-1} U_-^{-1} A'_{32} & & L'_- \end{bmatrix}, \quad (3.6)$$

with

$$A''_{22} = A'_{22} - A'_{23} U_-^{-1} D_-^{-1} L_-^{-1} A'_{32}. \quad (3.7)$$

At last, let exist the double factorization

$$A''_{22} = L''_+ D''_+ U''_+ = U''_- D''_- L''_-. \quad (3.8)$$

Then, we find the following two twisted factorizations of the block matrix (3.1).

The one with a twist in the *lower right* corner of the A_{22} is

$$\begin{bmatrix} L_+ & & & \\ A_{21}U_+^{-1}D_+^{-1} & L''_+ & A'_{23}L_-^{-1}D_-^{-1} & \\ A_{31}U_+^{-1}D_+^{-1} & & U'_- & \end{bmatrix} \cdot \begin{bmatrix} D_+ & & \\ & D''_+ & \\ & & D'_- \end{bmatrix} \cdot \begin{bmatrix} U_+ & D_+^{-1}L_+^{-1}A_{12} & D_+^{-1}L_+^{-1}A_{13} \\ & U''_+ & \\ & D_-^{-1}U_-^{-1}A'_{32} & L'_- \end{bmatrix}, \quad (3.9)$$

and the one with a twist in the *left upper* corner of the A_{22} is

$$\begin{bmatrix} L_+ & & & \\ A_{21}U_+^{-1}D_+^{-1} & U''_- & A'_{23}L_-^{-1}D_-^{-1} & \\ A_{31}U_+^{-1}D_+^{-1} & & U'_- & \end{bmatrix} \cdot \begin{bmatrix} D_+ & & \\ & D''_- & \\ & & D'_- \end{bmatrix} \cdot \begin{bmatrix} U_+ & D_+^{-1}L_+^{-1}A_{12} & D_+^{-1}L_+^{-1}A_{13} \\ & L''_- & \\ & D_-^{-1}U_-^{-1}A'_{32} & L'_- \end{bmatrix}. \quad (3.10)$$

Similar to Section 2.2, one can derive connections to the inverse A^{-1} .

However, exploiting the structure of the factors to efficiently compute an eigenvector approximation as in Section 2.3 requires a bit more care. For example, for the approximation of the right eigenvector, one has

$$Az_r = \nu_k e_k \Leftrightarrow \begin{bmatrix} U_+ & D_+^{-1}L_+^{-1}A_{12} & D_+^{-1}L_+^{-1}A_{13} \\ & U''_+ & \\ & D_-^{-1}U_-^{-1}A'_{32} & L'_- \end{bmatrix} z_r = e_k. \quad (3.11)$$

In this case, one first has to compute $z_r(k-b+1:k)$ from equations $k-b+1:k$, afterwards one can obtain $z_r(k+1:n)$ from the last $n-k+1$ equations, and finally one can compute $z_r(1:k-1)$. Computing the approximation of the left eigenvector is similar.

The task may partially simplify if at least one of the blocks A_{31} or A_{13} vanishes. For example, $A_{31} = 0$ if A is an upper Hessenberg matrix. In this case, the computation of a left eigenvector approximation reduces to the standard banded case, but the right eigenvector approximation remains as described here.

We also remark that the factorization technique could be applied recursively if A''_{22} is large enough. One can block-partition A''_{22} again as in (3.1) instead of directly performing the double factorization (3.8). If the bandwidth is large enough, this can also be done for A'_{22} as an alternative to (2.7).

4. Summary and conclusions. This paper gives a constructive proof for the existence of twisted factorizations of a band matrix. One of their important applications is inverse iteration and we showed how to exploit the twisted form.

We close with a question to the reader regarding choice of nomenclature: because of their 2-level construction, one may argue that ‘twisted’ is not the most descriptive adjective to describe the banded factorizations (2.8) and (2.9). This is even truer for the general factorizations (3.10) and (3.9). Maybe ‘interwoven’ should be used? Only in the simplest, tridiagonal, case, the factorization is simply twisted.

Acknowledgment. We thank B. Parlett for inspiring and encouraging us to pursue this topic.

REFERENCES

- [1] I. S. Dhillon. *A New $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/Eigenvector Problem*. PhD thesis, University of California, Berkeley, California, 1997.
- [2] I. S. Dhillon and B. N. Parlett. Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices. *Linear Algebra and Appl.*, 387:1–28, 2004.
- [3] I. S. Dhillon and B. N. Parlett. Orthogonal eigenvectors and relative gaps. *SIAM J. Matrix Anal. Appl.*, 25(3):858–899, 2004.
- [4] I. S. Dhillon, B. N. Parlett, and C. Vömel. The design and implementation of the MRRR algorithm. *ACM Trans. Math. Software*, 32(4):533–560, 2006.
- [5] K. V. Fernando. On Computing an Eigenvector of a Tridiagonal Matrix. Part I: Basic Results. *SIAM J. Matrix Anal. Appl.*, 18(4):1013–1034, 1997.
- [6] B. N. Parlett and I. S. Dhillon. Fernando’s solution to Wilkinson’s problem: an application of double factorization. *Linear Algebra and Appl.*, 267:247–279, 1997.
- [7] B. N. Parlett and I. S. Dhillon. Relatively robust representations of symmetric tridiagonals. *Linear Algebra and Appl.*, 309(1-3):121–151, 2000.
- [8] G. W. Stewart. *Matrix Algorithms II: Eigensystems*. SIAM, Philadelphia, 2001.
- [9] W. Xu and S. Qiao. A Twisted Factorization Method for Symmetric SVD of a Complex Symmetric Tridiagonal Matrix. Technical Report CAS 06-01-SQ, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada, 2006.
- [10] Z.-Y. Zhang and T.-W. Ouyang. Computing Eigenvectors of Normal Matrices with Simple Inverse Iteration. *Journal of Computational Mathematics*, 21(5):657–670, 2003.