Block Tridiagonal Reduction of Perturbed Normal and Rank Structured Matrices

Roberto Bevilacqua[a], Gianna M. Del Corso[a,*], Luca Gemignani[a]

[a]Università di Pisa, Dipartimento di Informatica, Largo Pontecorvo, 3, 56127 Pisa, Italy

Abstract

It is well known that if a matrix $A \in \mathbb{C}^{n\times n}$ solves the matrix equation $f(A, A^H) = 0$, where $f(x, y)$ is a linear bivariate polynomial, then A is normal; A and A^H can be simultaneously reduced in a finite number of operations to tridiagonal form by a unitary congruence and, moreover, the spectrum of A is located on a straight line in the complex plane. In this paper we present some generalizations of these properties for almost normal matrices which satisfy certain quadratic matrix equations arising in the study of structured eigenvalue problems for perturbed Hermitian and unitary matrices.

Keywords: Block tridiagonal reduction, rank structured matrix, block Lanczos algorithm
2000 MSC: 65F15

1. Introduction

Normal matrices play an important theoretical role in the field of numerical linear algebra. A square complex matrix is called normal if

$$A^H A - AA^H = 0,$$

where A^H is the conjugate transpose of A. Polyanalytic polynomials [15] of degree at most k are functions of the form $p(z) = \sum_{j=0}^{k} h_{k-j}(z)\bar{z}^j$, where $h_j(z)$, $0 \le j \le k$, are complex polynomials of degree less than or equal to j. A polyanalytic polynomial of minimal total degree that annihilates A, i.e., such that $p(A) = 0$, is called a minimal polyanalytic polynomial of A [15]. Over the years many equivalent conditions have been found [12, 7], and it has been discovered that the class of normal matrices can be partitioned in accordance with a parameter $s \in \mathbb{N}$, $s \le n-1$, where s is the minimal degree of a particular polyanalytic polynomial $p_s(z) = \bar{z} - n_s(z)$ such that $p_s(A) = A^H - n_s(A) = 0$, and $n_s(z)$ is a polynomial of degree s.

For a normal matrix the assumption of being banded imposes strong constraints on the localization of the spectrum and the degree of minimal polyanalytic polynomials. It is well known that the minimal polyanalytic polynomial of an irreducible normal tridiagonal matrix has degree one and, moreover, the spectrum of the matrix is located on a straight line in the complex plane [9, 14]. Generalizations of these properties to normal matrices with symmetric band structure are provided in [17]. Nonsymmetric structures are considered in the papers [18, 8] where it is shown that the customary Hessenberg reduction procedure applied to a normal matrix always returns a banded matrix with upper bandwidth at most k if and only if $s \le k$. A way to arrive at the Hessenberg form is using the Arnoldi method which amounts to construct a sequence of nested Krylov subspaces. A symmetric variation of the Arnoldi method named generalized Lanczos procedure is devised in [6] and applied in [6, 15] and [11] for the block tridiagonal reduction of normal and perturbed normal matrices, respectively. The reduction is rational –up to square root calculations– and finite but not computationally appealing since it essentially reduces to the orthonormalization of the sequence of generalized powers $A^j {A^k}^H \mathbf{v}$, $j + k = m$, $m \ge 0$. In [16] it is shown that any normal matrix can be unitarily reduced to a band matrix whose bandwidth is proportional to the degree of the minimal

*Corresponding author
Email addresses: bevilacq@di.unipi.it (Roberto Bevilacqua), delcorso@di.unipi.it (Gianna M. Del Corso), gemignan@di.unipi.it (Luca Gemignani)

polyanalytic polynomial of the matrix, an an algorithmic procedure based on a generalized Krylov method is given.

In [1] the class of *almost normal matrices* is introduced, that is the class of matrices for which $[A, A^H] = A^H A - AA^H = CA - AC$ for a low rank matrix C. In the framework of operator theory conditions upon the commutator $[A, A^H]$ are widely used in the study of structural properties of hypernormal operators [19]. Our interest in the class of almost normal matrices stems from the analysis of fast eigenvalue algorithms for rank–structured matrices. If A is a rank–one correction of a Hermitian or unitary matrix than A satisfies $[A, A^H] = CA - AC$ for a matrix C of rank at most 2. Furthermore, this matrix C is involved in the description of the rank structure of the matrices generated starting from A under the QR process [2, 4, 20]. Thus the problem of simultaneously reducing both A and C to symmetric band structure is theoretically interesting but it also might be beneficial for the design of fast effective eigenvalue algorithms for these matrices. Furthermore, condensed representations expressed in terms of block matrices [5] or the product of simpler matrices [3, 21] tends to become inefficient as the length of the perturbation increases [10]. The exploitation of condensed representations in banded form can circumvent these difficulties.

In [1] it is shown that we can always find an almost normal block tridiagonal matrix with blocks of size 2 which fulfills the commutator equation for a certain C with $\text{rank}(C) = 1$. Although an algorithmic construction of a block tridiagonal solution is given, no efficient computational method is described in that paper for the block tridiagonal reduction of a prescribed solution of the equation. In this contribution, we first propose an algorithm based on the application of the block Lanczos method to the matrix $A + A^H$ starting from a suitable set of vectors associated with the range of the commutator $[A, A^H]$ for the block tridiagonal reduction of an eligible solution of the commutator equation. Then we generalize the approach to the case where $\text{rank}(C) = 2$ that is relevant for the applications to rank–structured eigenvalue problems. We also report experimental evidence that in these problems the proposed reduction effectively impacts the tracking of the rank structures under the customary QR process. Finally, we show that similar results still partially hold when A is a rank–one modification of particular normal matrices whose eigenvalues lie on a real algebraic curve of degree 2. In the latter case the matrix C of rank at most 2 could not exist and, therefore, the analysis of this configuration is useful to put in evidence the consequences of such a missing.

2. Simultaneous block tridiagonalization

In this section we discuss the reduction to block tridiagonal form of almost normal matrices.

Definition 1. *Let A be an $n \times n$ matrix. If there exists a rank-k matrix C such that*

$$[A, A^H] = A^H A - AA^H = CA - AC,$$

we say that A is a k-almost normal matrix.

Denote by $\Delta(A) := [A, A^H] = A^H A - AA^H$ the commutator of A and by $\mathcal{S}$ the range of $\Delta(A)$. It is clear that if, for a given C, any solution of the nonlinear matrix equation

$$[X, X^H] = X^H X - XX^H = CX - XC, \quad C, X \in \mathbb{C}^n, \tag{1}$$

exists, then it is not unique. Indeed, if A is an almost-normal matrix such that $\Delta(A) = CA - AC$ then $B = A + \gamma I$, with a complex constant γ, is almost normal as well and $\Delta(B) = \Delta(A) = CB - BC$. In [1] the structure of almost normal matrices with rank–one perturbation is studied by showing that a block-tridiagonal matrix with 2×2 blocks can be determined to satisfy (1) for a certain C of rank one. Here we take a different look at the problem by asking whether a solution of (1) for a given C can be reduced to block tridiagonal form.

The block Lanczos algorithm is a technique for reducing a Hermitian matrix $H \in \mathbb{C}^{N \times n}$ to block tridiagonal form. There are many variants of the basic block Lanczos procedure. The method stated below is in the spirit of the block Lanczos algorithm described in [13].

```
Procedure Block_Lanczos
Input: H, Z ∈ ℂ^{n×ℓ} nonzero, ℓ ≤ n;
[Q,Σ,V] = svd(Z); s = rank(Σ);
U(:,1:s) = Q(:,1:s); s_0 = 1, s_1 = s;
while s_1 < n
  W = A_H · U(:,1:s); T(s_0:s_1, s_0:s_1) = (U(:,1:s))^H · W;
  if s_0 = 1
    W = W − U(:,s_0:s_1) · T(s_0:s_1, s_0:s_1);
  else
    W = W − U(:,s_0:s_1) · T(s_0:s_1, s_0:s_1);
    W = W − U(:,ŝ_0:ŝ_1) · T(ŝ_0:ŝ_1, s_0:s_1);
  end
  [Q,Σ,V] = svd(W); s_new = rank(Σ);
  if s_new = 0
    disp('premature stop'); return;
  else
    Σ = Σ(1:s_new, 1:s_new) · (V(:,1:s))^H;
    ŝ_0 = s_0, ŝ_1 = s_1, s_0 = s_1 + 1, s_1 = s_1 + s_new;
    U(:,s_0:s_1) = Q(:,1:s_new), T(s_0:s_1, ŝ_0:ŝ_1) = Σ(1:s_new, 1:s);
    T(ŝ_0:ŝ_1, s_0:s_1) = (T(s_0:s_1, ŝ_0:ŝ_1))^H, s = s_new;
  end
end
T(s_0:s_1, s_0:s_1) = (U(:,s_0:s_1))^H · A_H · U(:,s_0:s_1);
```

The procedure, when terminates without a premature stop, produces a block-tridiagonal matrix and a unitary matrix U such that

$$U^H H U = T = \begin{bmatrix} A_1 & B_1^H & & \\ B_1 & A_2 & \ddots & \\ & \ddots & \ddots & B_{p-1}^H \\ & & B_{p-1} & A_p \end{bmatrix},$$

where $A_k \in \mathbb{C}^{i_k \times i_k}$, $B_k \in \mathbb{C}^{i_{k+1} \times i_k}$, and $\ell \geq i_k \geq i_{k+1}$, $i_1 + i_2 + \cdots + i_p = n$. In fact, the size of the blocks can possibly shrink when the rank of the matrices W is less than ℓ.

Let $Z \in \mathbb{C}^{n\times\ell}$, and denote by $\mathcal{K}_j(H,Z)$ the block Krylov subspace generated by the column vectors in Z, that is the space spanned by the columns of the matrices $Z, HZ, H^2Z, \ldots, H^{j-1}Z$. It is well known that the classical Lanczos process builds an orthonormal basis for the Krylov subspace $\mathcal{K}_{n-1}(H,\mathbf{z})$, for $\mathbf{z} = \alpha U(:,1)$. Similarly, when the block Lanczos process does not break down, $\operatorname{span}\{U(:,1), U(:,2) \ldots, U(:,n)\} = \mathcal{K}_j(H,Z)$ for j such that $\dim(\mathcal{K}_j(H,Z)) = n$. When the block-Lanczos procedure terminates before completion it means that $\mathcal{K}_j(H,Z)$ is an invariant subspace and a null matrix W has been found in the above procedure. In this case the procedure has to be restarted and the final matrix T is block diagonal. As an example, we can consider the matrix $H = U + U^H$, where U is the Fourier matrix $U := \mathcal{F}_n = \frac{1}{\sqrt{n}}\Omega_n$ of order $n = 2m$. Due to the relation $\Omega_n^2 = n\Pi$, where Π is a suitable symmetric permutation matrix, it is found that for any starting vector $\mathbf{z}$ the **Block_Lanczos** procedure applied with $Z = [\mathbf{z}|U\mathbf{z}]$ breaks down within the first three steps. In this case the reduction scheme has to be restarted and the initial matrix can be converted into the direct sum of diagonal blocks.

2.1. *Case of rank one*

In this section we consider the case where the matrix $A \in \mathbb{C}^{n\times n}$ solves (1) for a prescribed nonzero matrix C of rank one, that is $C = \mathbf{u}\mathbf{v}^H$, $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$. We show that A can be unitarily converted to a block tridiagonal form with blocks of size at most 2 by applying the block-Lanczos procedure starting from a basis of $\mathcal{S}$ the column space of $\Delta(A)$.

Let us introduce the Hermitian and antihermitian part of A denoted as

$$A_H :\, = \frac{A+A^H}{2}, \quad A_{AH} :\, = \frac{A-A^H}{2}.$$

Observe that

$$\Delta(A) = A^H A - AA^H = 2(A_H A_{AH} - A_{AH} A_H).$$

In the next theorems we prove that the Krylov subspace of A_H obtained starting from a basis of $\mathcal{S}$, the column space of $\Delta(A)$, coincides with the Krylov space of A_{AH}, and hence with that of A. We first need some technical lemmas.

Lemma 1. *Let A be a 1-almost normal matrix, and let $C = \mathbf{u}\mathbf{v}^H$ be a rank–one matrix such that $\Delta(A) = CA - AC$ is nonzero. Then $\Delta(A)$ has rank two and $(\mathbf{u}, A\mathbf{u})$ and $(\mathbf{v}, A^H\mathbf{v})$ are two bases of $\mathcal{S}$. Moreover, if $\mathbf{u}$ and $\mathbf{v}$ are linearly independent then $(\mathbf{u}, \mathbf{v})$ is a basis for $\mathcal{S}$ as well.*

PROOF. Note that $CA - AC = \mathbf{u}\mathbf{v}^H A - A\mathbf{u}\mathbf{v}^H$, $\Delta(A)$ is Hermitian and, therefore, $\Delta(A)$ has rank two and $\mathcal{S} = \text{span}\{\mathbf{u}, A\mathbf{u}\}$. Because of the symmetry of $\Delta(A)$, $(\mathbf{v}, A^H\mathbf{v})$ is a basis of $\mathcal{S}$ as well. Moreover, if $\mathbf{u}$ and $\mathbf{v}$ are linearly independent they form a basis for $\mathcal{S}$ since both vectors belong to $\mathcal{S}$.

Lemma 2. *Let A be a 1-almost normal matrix with $C = \mathbf{u}\mathbf{v}^H$, where $\mathbf{u}$ and $\mathbf{v}$ are linearly independent. Then we have*

$$AA_H^k\mathbf{u} = \sum_{j=0}^{k} \lambda_j^{(k)} A_H^j \mathbf{u} + \sum_{j=0}^{k} \mu_j^{(k)} A_H^j \mathbf{v}, \qquad k = 1, 2, \ldots;$$

and similarly

$$A^H A_H^k\mathbf{v} = \sum_{j=0}^{k} \hat{\lambda}_j^{(k)} A_H^j \mathbf{u} + \sum_{j=0}^{k} \hat{\mu}_j^{(k)} A_H^j \mathbf{v}, \qquad k = 1, 2, \ldots.$$

PROOF. We prove the first case by induction on k. Observe that it holds

$$A_H A - AA_H = \frac{\Delta(A)}{2},$$

which gives

$$AA_H\mathbf{u} = A_H A\mathbf{u} - \frac{\Delta(A)}{2}\mathbf{u}.$$

From $\mathcal{S} = \text{span}\{\mathbf{u}, \mathbf{v}\}$, $\Delta(A)\mathbf{u} \in \mathcal{S}$ and $A\mathbf{u} \in \mathcal{S}$ we deduce the relation for $k = 1$. Then assume the thesis is true for k and let prove it for $k+1$. Denote by $\mathbf{x} = A_H^k\mathbf{u}$, we have

$$AA_H^{k+1}\mathbf{u} = AA_H\mathbf{x} = A_H A\mathbf{x} - \frac{\Delta(A)}{2}\mathbf{x}.$$

Since $\Delta(A)\mathbf{x} \in \mathcal{S}$, applying the inductive hypothesis we get the thesis. The proof of the second case proceeds analogously by using

$$A^H A_H - A_H A^H = \frac{\Delta(A)}{2}.$$

Lemma 3. *Let $Z = [\mathbf{z}_1|\ldots|\mathbf{z}_\ell] \in \mathbb{C}^{n\times\ell}$ and $X \in \mathbb{C}^{n\times\ell}$ such that span$\{Z\} :\, =$ span$\{\mathbf{z}_1, \ldots, \mathbf{z}_\ell\} =$ span$\{X\}$, then $\mathcal{K}_j(A, Z) = \mathcal{K}_j(A, X)$.*

PROOF. If $\text{span}\{Z\} = \text{span}\{X\}$, then there exists a square nonsingular matrix B such that $Z = XB$. Let $\mathbf{u} \in \mathcal{K}_j(A, Z)$, then we can write $\mathbf{u}$ as a linear combination of the vectors of $\mathcal{K}_j(A, Z)$, that is there exists a $(j+1)s$ vector $\mathbf{a}$ such that

$$\begin{aligned} \mathbf{u} &= [Z, AZ, \ldots, A^j Z]\mathbf{a} = [XB, AXB, \ldots, A^j XB]\mathbf{a} = \\ &= [X, AX, \ldots, A^j X]\begin{bmatrix} B & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{bmatrix}\mathbf{a} = [X, AX, \ldots, A^j X]\mathbf{b} \in \mathcal{K}_j(A, X), \end{aligned}$$

where $\mathbf{b} = (I_{j+1} \otimes B)\mathbf{a}$.

We denote as $\mathcal{K}_j(A, <\mathbf{z}_1,\ldots,\mathbf{z}_l>)$ the Krylov subspace of A generated starting from any initial matrix $X \in \mathbb{C}^{n\times\ell}$ satisfying $\operatorname{span}\{\mathbf{z}_1,\ldots,\mathbf{z}_\ell\} = \operatorname{span}\{X\}$.

The main result of this section is the following.

Theorem 4. *Let A be a 1-almost normal matrix with $C = \mathbf{u}\mathbf{v}^H$. If $\mathbf{u}$ and $\mathbf{v}$ are linearly independent then $\mathcal{K}_j(A_{AH}, <\mathbf{u},\mathbf{v}>) \subseteq \mathcal{K}_j(A_H, <\mathbf{u},\mathbf{v}>)$ for each j. Thus, if the block Lanczos process does not break down prematurely, $U^H A_H U$ and $U^H A_{AH} U$ are block-tridiagonal and, hence, $U^H AU$ is block tridiagonal as well. The size of the blocks is at most two.*

PROOF. The proof is by induction on j. For $j = 1$, let $\mathbf{x} \in \mathcal{K}_1(A_{AH}, <\mathbf{u},\mathbf{v}>)$, we need to prove that $\mathbf{x} \in \mathcal{K}_1(A_H, <\mathbf{u},\mathbf{v}>)$. Since $\mathbf{x} \in \mathcal{K}_1(A_{AH}, <\mathbf{u},\mathbf{v}>)$, then $\mathbf{x} \in \operatorname{span}\{\mathbf{u},\mathbf{v},A_{AH}\mathbf{u},A_{AH}\mathbf{v}\}$. It is enough to prove that $A_{AH}\mathbf{u} \in \mathcal{K}_1(A_H, <\mathbf{u},\mathbf{v}>)$ and $A_{AH}\mathbf{v} \in \mathcal{K}_1(A_H, <\mathbf{u},\mathbf{v}>)$. From

$$A_{AH} + A_H = A, \quad A_{AH} - A_H = -A^H \tag{2}$$

we obtain that

$$A_{AH}\mathbf{u} = -A_H\mathbf{u} + A\,\mathbf{u}.$$

Since from Lemma 1 $A\mathbf{u} \subseteq \operatorname{span}\{\mathbf{u},\mathbf{v}\} \in \mathcal{K}_1(A_H, <\mathbf{u},\mathbf{v}>)$, we conclude that $A_{AH}\mathbf{u} \in \mathcal{K}_1(A_H, <\mathbf{u},\mathbf{v}>)$. Similarly we find that

$$A_{AH}\mathbf{v} = A_H\mathbf{v} - A^H\,\mathbf{v} \in \mathcal{K}_1(A_H, <\mathbf{u},\mathbf{v}>).$$

Assume now that the thesis holds for j and prove it for $j+1$. For the linearity of the Krylov subspaces we can prove the thesis on the monomials and for each of the starting vectors $\mathbf{u}$ and $\mathbf{v}$. Let $\mathbf{x} = A_{AH}^j\,\mathbf{u}$. We have

$$A_{AH}^{j+1}\,\mathbf{u} = A_{AH}\mathbf{x}$$

Since by inductive hypothesis $\mathbf{x} \in \mathcal{K}_j(A_H, <\mathbf{u},\mathbf{v}>)$, $\mathbf{x} = \sum_{k=0}^{j} \alpha_k A_H^k\mathbf{u} + \sum_{k=0}^{j} \beta_k A_H^k\mathbf{v}$, and using (2) we obtain that

$$\begin{aligned} A_{AH}^{j+1}\mathbf{u} &= A_{AH}\mathbf{x} = (-A_H + A)\sum_{k=0}^{j} \alpha_k A_H^k\mathbf{u} + (A_H - A^H)\sum_{k=0}^{j} \beta_k A_H^k\mathbf{v} \\ &= -\sum_{k=1}^{j+1} \alpha_k A_H^k\mathbf{u} + \sum_{k=1}^{j+1} \beta_k A_H^k\mathbf{v} + \sum_{k=0}^{j} \alpha_k A A_H^k\mathbf{u} - \sum_{k=0}^{j} \beta_k A^H A_H^k\mathbf{v}. \end{aligned}$$

By applying Lemma 2 to each term of the form $AA_H^k\mathbf{u}$ and $A^H A_H^k\mathbf{v}$ in the previous relation we obtain that $A_{AH}^{j+1}\mathbf{u} \in \mathcal{K}_{j+1}(A_H,, \mathbf{u},\mathbf{v}>)$. With a similar technique we prove that $A_{AH}^{j+1}\,\mathbf{v} \in \mathcal{K}_{j+1}(A_H, <\mathbf{u},\mathbf{v}>)$.

Lemma 3, provided $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, proves that we can apply the block Lanczos procedure to any pair of linearly independent vectors in $\mathcal{S}$. The remaining case where $\mathbf{u}$ and $\mathbf{v}$ are not linearly independent, that is the rank–one correction has the form $C = \alpha\mathbf{u}\mathbf{u}^H$, can treated as follows.

Theorem 5. *Let A be a 1-almost normal matrix with $C = \alpha\mathbf{u}\mathbf{u}^H$. Then $\mathcal{K}_j(A_{AH},\mathbf{u}) \subseteq \mathcal{K}_j(A_H,\mathbf{u})$. Hence, if breakdown does not occur, the classic Lanczos process applied to A_H with starting vector $\mathbf{u}$ returns a unitary matrix U which reduces A_{AH} to tridiagonal form and therefore also $U^H AU$ is tridiagonal.*

PROOF. Set $B = -\frac{i}{\alpha}A$ obtaining for B the following relation

$$B^H B - BB^H = \hat{C}B - B\hat{C},$$

with $\hat{C} = i\mathbf{u}\mathbf{u}^H$, that is with an antihermitian correction. Since $\Delta(B)$ is hermitian, we obtain

$$\hat{C}B - B\hat{C} = B^H\hat{C}^H - \hat{C}^H B^H = -B^H\hat{C} + \hat{C}B^H,$$

and hence

$$\hat{C}B_{AH} = B_{AH}\hat{C},$$

meaning that $\operatorname{span}\{B_{AH}\mathbf{u}\} \subseteq \operatorname{span}\{\mathbf{u}\}$. This proves that $\mathcal{K}_j(B_{AH},\mathbf{u}) \subseteq \operatorname{span}\{\mathbf{u}\} \subseteq \mathcal{K}_j(B_H,\mathbf{u})$, and hence that B_{AH} is brought in tridiagonal form by means of the same unitary matrix which tridiagonalizes B_H. Then B and A are brought to tridiagonal form by the same U.

Note that Theorem 5 states that any 1-almost normal matrix with $C = \alpha\mathbf{u}\mathbf{u}^H$ can be unitarily transformed into tridiagonal form.

2.2. Case of rank two

The class of 2-almost normal matrices is a richer and more interesting class. For example, rank–one perturbations of unitary matrices, such as the companion matrix, belong to this class. Also generalized companion matrices for polynomials expressed in the Chebyshev basis [4, 20] can be viewed as rank–one perturbation of Hermitian matrices and are 2-almost normal.

Assume $A^H A - AA^H = CA - AC$, where $C = \mathbf{u}\mathbf{v}^H + \mathbf{x}\mathbf{y}^H$. Note that $\dim(\mathcal{S}) \leq 4$. If the column space of $\Delta(A)$ has dimension exactly 4 then possible bases for $\mathcal{S}$ are $< \mathbf{u}, \mathbf{x}, A\mathbf{u}, A\mathbf{x} >$, $< \mathbf{v}, \mathbf{y}, A^H\mathbf{v}, A^H\mathbf{y} >$ and $< \mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y} >$ when the four vectors are linearly independent.

A theorem analogous to Theorem 4 for 2-almost normal matrices uses a generalization of Lemmas 1 and 2.

Lemma 6. *Let A be a 2-almost normal matrix, and let $C = UV^H$, with $U, V \in \mathbb{C}^{n\times 2}$ be a rank-2 matrix such that $\Delta(A) = CA - AC$ has rank 4. Then, the columns of the matrices $[U, AU]$ and of $[V, A^H V]$ span the space $\mathcal{S}$. Moreover, if rank$([U,V]) = 4$ the columns of the matrix $[U,V]$ form a basis for $\mathcal{S}$ as well.*

Similarly Lemma 2 can be generalized replacing the vectors $\mathbf{u}$ and $\mathbf{v}$ with two $n \times 2$ matrices.

Lemma 7. *Let A be a 2-almost normal matrix with $C = UV^H$, with $U, V \in \mathbb{C}^{n\times 2}$, with rank$(\Delta(A)) = 4$ and rank$([U,V]) = 4$. Then we have*

$$AA_H^j U \in \mathcal{K}_j(A_H, [U,V]) \qquad j = 1, 2, \ldots$$

and similarly

$$A^H A_H^j V \in \mathcal{K}_j(A_H, [U,V]) \qquad j = 1, 2, \ldots.$$

We are now ready for the desired generalization of the main result. The proof is similar to that of Theorem 4 and it is omitted here.

Theorem 8. *Let A be a 2-almost normal matrix with $C = UV^H$, with $U, V \in \mathbb{C}^{n\times 2}$. If rank$([U,V]) = 4$ and rank$(\Delta(A)) = 4$, then we have $\mathcal{K}_j(A_{AH}, [U,V]) \subseteq \mathcal{K}_j(A_H, [U,V])$ for each j. Hence, if the block Lanczos process does not break down, the unitary matrix which transforms A_H to block-tridiagonal form brings also A_{AH} to block-tridiagonal form, and hence also A is brought to block tridiagonal form with blocks of size at most 4.*

Generalizations of these results to generic k-almost normal matrices with $k \geq 2$ are straightforward.

3. Almost Hermitian or unitary matrices

In this section we specialize the previous results for the remarkable cases where A is a rank–one perturbation of a Hermitian or a unitary matrix. The case of perturbed Hermitian matrices is not directly covered by Theorem 8. In fact, it assumes that rank$(\Delta(A)) =$ rank$([U,V])$ or, equivalently, that there exists a set of $2k$ linearly independent vectors spanning the column space of $\Delta(A)$ whenever A is $k-$almost normal. If $A = H + \mathbf{x}\mathbf{y}^H$, where H is a Hermitian matrix, then it is easily seen that A is 2-almost normal and $C = \mathbf{y}\mathbf{x}^H - \mathbf{x}\mathbf{y}^H$. Generically, $\Delta(A)$ has rank 4 but $U = V$ and, therefore, rank$([U,V]) = 2$. However, it is worth noticing that in this case $C = UV^H$ is antihermitian, i.e., $C^H = -C$. By exploiting this additional property of C we can prove that

$$\mathcal{K}_j(A_{AH}, U) \subseteq \mathcal{K}_1(A_H, U), \quad j \geq 1, \tag{3}$$

meaning that the same unitary matrix which transforms the Hermitian part of A toi block tridiagonal form, also transforms the antihermitian part of A to block tridiagonal form and, therefore, A. In order to deduce (3), let us observe that $\Delta(A)$ is Hermitian and

$$CA - AC = A^H C^H - C^H A^H.$$

Replacing $C^H = -C$, we obtain that

$$C(A - A^H) = (A - A^H)C, \tag{4}$$

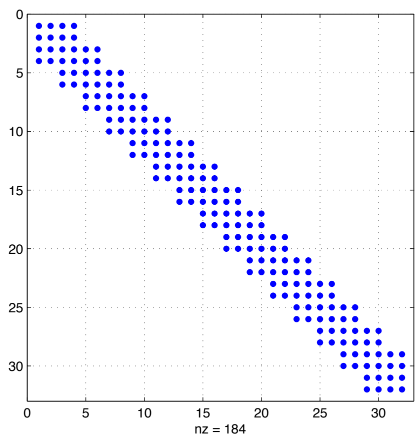

Figure 1: Shape of the block tridiagonal matrix obtained from the block-Lanczos procedure applied to a arrow matrix with starting vectors in the column space of C

meaning that the antihermitian part of A commutes with C. Multiplying both sides of (4) by the matrix V we have

$$(A - A^H)U = U(V^H(A - A^H)V)(V^H V)^{-1}.$$

which gives (3).

Summing up, in the case of a rank–one modification of a Hermitian matrix we can apply the block-Lanczos procedure to A_H starting with only two linearly independent vectors in the column space of C, for example $\mathbf{x}$ and $\mathbf{y}$, if known, thus computing a unitary matrix which transforms A to a block-tridiagonal matrix with block size two. Differently, we can also employ a basis of $\mathcal{S}$ by obtaining a first block of size 4 which immediately shrinks to size 2 in the subsequent steps. In Figure 1 and 2 we illustrate the shapes of the block tridiagonal matrices determined from the block-Lanczos procedure applied to an arrow matrix with starting vectors in the column space of C and $\mathcal{S}$, respectively.

It is worth noticing that the matrix C plays an important role for the design of fast structured variants of the QR iteration applied to perturbed Hermitian matrices. Specifically, in [4, 21] it is shown that the sequence of perturbations $C_k :\ = Q_k^H C_{k-1} Q_k$ yields a description of the upper rank structure of the matrices $A_k :\ = Q_k^H A_{k-1} Q_k$ generated under the QR process applied to $A_0 = A$.

The case where $A = U + \mathbf{x}\mathbf{y}^H$ is a rank–one correction of a unitary matrix U is particularly interesting for applications to polynomial root-finding. If A is invertible, then it is easily seen that

$$C = \mathbf{y}\mathbf{x}^H + \frac{U\mathbf{x}\mathbf{y}^H U^H}{1 + \mathbf{y}^H U^H \mathbf{x}}$$

is such that

$$A^H A - CA = I_n, \quad AA^H - AC = I_n,$$

which implies

$$A^H A - AA^H = CA - AC.$$

Thus, by applying Theorem 8 to the unitary plus rank–one matrix A we find that the block-Lanczos procedure applied to A_H starting with four linearly independent vectors in $\mathcal{S}$ reduces A to a block tridiagonal form as depicted in Figure 3.

The issue concerning the relationship between the matrices C_k and A_k generated under the QR iteration is more puzzling and, indeed, actually much of the work of fast companion eigensolvers is spent for the updating of the rank structure in the upper triangular portion of A_k. Moreover, if $A = A_0$ is initially transformed to upper Hessenberg form by a unitary congruence then the rank of the off–diagonal blocks in the

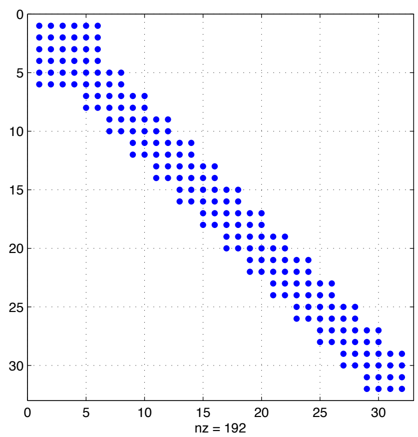

Figure 2: Shape of the block tridiagonal matrix obtained from the block-Lanczos procedure applied to a arrow matrix with starting vectors in the column space of $\mathcal{S}$

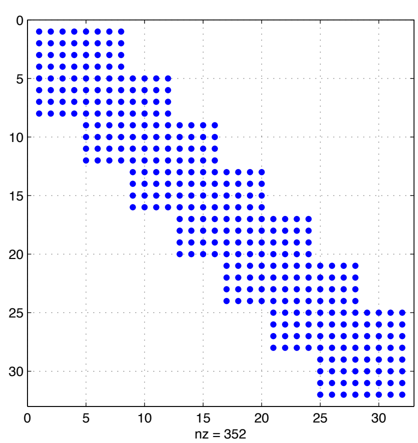

Figure 3: Shape of the block tridiagonal matrix obtained from the block-Lanczos procedure applied to a companion matrix

upper triangular part of A_k is generally three whereas the rank of C_k is two. Notwithstanding, the numerical behavior of the QR iteration seems to be different if we apply the iterative process directly to the block tridiagonal form of the matrix. In this case, under some mild assumption, it is verified that the rank of the blocks in the upper triangular portion of A_k located out of the block tridiagonal profile is at most 2 and, in addition, the rank structure of these blocks is completely specified by the matrix C_k.

4. Eigenvalues on an algebraic curve

The property of block tridiagonalization is inherited by a larger class of perturbed normal matrices [11]. It is interesting to consider such extension even in simple cases in order to enlighten the specific features of almost normal matrices with respect to the band reduction and to the QR process. In this section we show that the block-Lanczos procedure can be employed for the block tridiagonalization of rank–one perturbations of certain normal matrices whose eigenvalues lie on an algebraic curve of degree at most two. These matrices are not in general almost normal, but the particular distribution of the eigenvalues of the normal part, guarantees the existence of a polyanalytic polynomial of small degree relating the antihermitian part of A with the Hermitian part of A.

Let $A \in \mathbb{C}^{n\times n}$ be a matrix which can be decomposed as

$$A = N + \mathbf{u}\mathbf{v}^H, \mathbf{u}, \mathbf{v} \in \mathbb{C}^n, \quad NN^H - N^H N = 0. \tag{5}$$

Also suppose that the eigenvalues of N, $\lambda_j = \Re(\lambda_j) + \mathrm{i}\Im(\lambda_j)$, $1 \le j \le n$, lie on a real algebraic curve of degree 2, i.e., $f(\Re(\lambda_j), \Im(\lambda_j)) = 0$, where $f(x,y) = ax^2 + by^2 + cxy + dx + ey + f = 0$. From

$$\Re(\lambda) = \frac{\lambda + \bar\lambda}{2}, \quad \Im(\lambda) = \frac{\lambda - \bar\lambda}{2\mathrm{i}},$$

by setting

$$x = \frac{z + \bar z}{2}, \quad y = \frac{z - \bar z}{2\mathrm{i}},$$

it follows that λ_j, $1 \le j \le n$, belong to an algebraic variety $\Gamma = \{z \in \mathbb{C} \colon p(z) = 0\}$ defined by

$$p(z) = a_{2,0}z^2 + a_{1,1}z\bar z + a_{0,2}\bar z^2 + a_{1,0}z + a_{0,1}\bar z + a_{0,0} = 0,$$

with $a_{k,j} = \bar a_{j,k}$. This also means that the polyanalytic polynomial $p(z)$ annihilates N in the sense that

$$p(N) = a_{2,0}N^2 + a_{1,1}NN^H + a_{0,2}{N^H}^2 + a_{1,0}N + a_{0,1}N^H + a_{0,0}I_n = 0. \tag{6}$$

If $a_{2,0} = \bar a_{0,2} = 0$ and $a_{1,1} = 0$ then Γ reduces to

$$a_{0,1}\bar z = -(a_{1,0}z + a_{0,0}).$$

that is the case of a shifted Hermitian matrix $N = H + \gamma I$. As observed in the previous section rank–one corrections of Hermitian matrices are almost-normal, and shifted almost normal matrices are almost-normal as well. Thus we can always suppose that the following condition named (*Hypothesis 1*) is fulfilled

$$a_{2,0} + a_{0,2} - a_{1,1} \neq 0. \tag{7}$$

In fact when Hypothesis 1 is violated, but not all the terms above are zero, then we can consider the modified matrix $A' = e^{\mathrm{i}\theta}A = e^{\mathrm{i}\theta}N + \mathbf{u}'\mathbf{v}'^H$ and observe that the eigenvalues of $e^{\mathrm{i}\theta}N$ belongs to the algebraic variety

$$a'_{2,0}z^2 + a'_{1,1}z\bar z + a'_{0,2}\bar z^2 + a'_{1,0}z + a'_{0,1}\bar z + a'_{0,0} = 0,$$

where

$$a'_{2,0} = a_{2,0}/e^{2\mathrm{i}\theta}, \quad a'_{0,2} = a_{0,2}/e^{-2\mathrm{i}\theta}, \quad a'_{1,1} = a_{1,1}.$$

Hence, for a suitable choice of θ it follows

$$a'_{2,0} + a'_{0,2} - a'_{1,1} \neq 0.$$

Under Hypothesis 1 it is easily seen that the leading part of $p(z)$ can be represented in some useful diverse ways. In particular, the 3×3 linear system in the variables α, β and γ determined to satisfy

$$\alpha(z-\bar{z})z+\beta(z+\bar{z})z+\gamma(z+\bar{z})\bar{z}=a_{2,0}z^2+a_{1,1}z\bar{z}+a_{0,2}\bar{z}^2, \tag{8}$$

is given by

$$\begin{cases} \gamma=a_{0,2}; \\ \alpha+\beta=a_{2,0}; \\ \beta+\gamma-\alpha=a_{1,1}. \end{cases}$$

This system is solvable and, moreover, we have $\alpha=\dfrac{a_{2,0}+a_{0,2}-a_{1,1}}{2}\neq 0$. Analogously, the 3×3 linear system in the variables α, β and γ determined to satisfy

$$\alpha(z-\bar{z})\bar{z}+\beta(z+\bar{z})z+\gamma(z+\bar{z})\bar{z}=a_{2,0}z^2+a_{1,1}z\bar{z}+a_{0,2}\bar{z}^2, \tag{9}$$

is given by

$$\begin{cases} \beta=a_{2,0}; \\ \gamma-\alpha=a_{0,2}; \\ \beta+\gamma+\alpha=a_{1,1}. \end{cases}$$

Again the system is solvable and $\alpha=-\dfrac{a_{2,0}+a_{0,2}-a_{1,1}}{2}\neq 0$.

For $A=N+\mathbf{u}\mathbf{v}^H$, with N normal matrix, the matrix $\Delta(A)=A^HA-AA^H$ is a matrix of rank four at most. Specifically, we find that

$$\Delta(A)=A^H\mathbf{u}\mathbf{v}^H+\mathbf{v}\mathbf{u}^HN-A\mathbf{v}\mathbf{u}^H-\mathbf{u}\mathbf{v}^HN^H,$$

and, hence, the space $\mathcal{S}$ is included in the subspace

$$\mathcal{D}:=\mathrm{span}\{\mathbf{u},\mathbf{v},A^H\mathbf{u},A\mathbf{v}\}\subseteq\mathcal{D}_s:=\mathrm{span}\{\mathbf{u},\mathbf{v},A^H\mathbf{u},A^H\mathbf{v},A\mathbf{u},A\mathbf{v}\}.$$

Also, recall that

$$A_{AH}\cdot A_H-A_H\cdot A_{AH}=\frac{1}{2}\Delta(A).$$

From this by induction it is easy to prove the following result, analogous to Lemma 1.

Lemma 9. *For any positive integer j we have*

$$A_{AH}\cdot A_H^j=A_H^j\cdot A_{AH}+\frac{1}{2}\sum_{k=0}^{j-1}A_H^k\cdot\Delta(A)\cdot A_H^{j-1-k}.$$

If the procedure block-Lanczos applied to A_H with initial matrix $Z\in\mathbb{C}^{n\times\ell}$, $\ell\leq 6$ such that $\mathrm{span}\{Z\}=\mathcal{D}_s$ terminates without premature stop then at the very end the unitary matrix U transforms A_H into the Hermitian block tridiagonal matrix $T=U^H\cdot A_H\cdot U$ with blocks of size at most 6. The following result says that $H:=U^H\cdot A_{AH}\cdot U$ is also block-tridiagonal with blocks of size at most 6.

Theorem 10. *Let $A\in\mathbb{C}^{n\times n}$ be as in (5), (6) and(7). Then we have $\mathcal{K}_j(A_{AH},Z)\subseteq\mathcal{K}_j(A_H,Z)$ for each $j\geq 0$, whenever span$\{Z\}=\mathcal{D}_s=$ span$\{\mathbf{u},\mathbf{v},A^H\mathbf{u},A^H\mathbf{v},A\mathbf{u},A\mathbf{v}\}$. Hence, if the block Lanczos process does not break down, the unitary matrix which transform A_H to block-tridiagonal form brings also A_{AH} to block-tridiagonal form, and hence also A is brought to block tridiagonal form with blocks of size at most 6.*

PROOF. Let $U(:,1:i_1)$ be the first block of columns of U spanning the subspace $\mathcal{D}_s$. The proof follows by induction on j. Consider the initial step $j=1$. We have

$$A_{AH}\mathbf{u}=-A_H\mathbf{u}+A\mathbf{u}\in\mathcal{K}_1(A_H,U(:,1:i_1))$$

and a similar relation holds for $A_{AH}\mathbf{v}$. Concerning $A_{AH}A\mathbf{u}$ from (8) we obtain that

$$\alpha(N-N^H)N+\beta(N+N^H)N+\gamma(N+N^H)N^H=a_{2,0}N^2+a_{1,1}NN^H+a_{0,2}{N^H}^2,$$

and, hence, by using (6) we find that

$$-\alpha(N-N^H)N = \beta(N+N^H)N + \gamma(N+N^H)N^H + (a_{1,0}N + a_{0,1}N^H + a_{0,0}I). \quad (10)$$

By plugging $N = A - \mathbf{u}\mathbf{v}^H$ into (10) we conclude that

$$\frac{(A-A^H)}{2} A\mathbf{u} \in \mathcal{K}_1(A_H, U(:,1:i_1)).$$

We can proceed similarly to establish the same property for the remaining vectors $A_{AH}A\mathbf{v}$ and $A_{AH}A^H\mathbf{u}$, $A_{AH}A^H\mathbf{v}$ by using (9).

To complete the proof, assume that

$$A_{AH}^{j}U(:,1:i_1) \in \mathcal{K}_j(A_H, U(:,1:i_1)),$$

and prove the same relation for $j+1$. We have

$$A_{AH}^{j+1}U(:,1:i_1) = A_{AH}\,(A_{AH}^{j}U(:,1:i_1) = A_{AH}X.$$

By induction X belongs to $\mathcal{K}_j(A_H, U(:,1:i_1))$ and, therefore, the thesis is proven by applying Lemma 9.

Note that when we have a coefficient $a_{ij} = 0$ we may need less vectors in the approximating initial subspace. For example in the case N is unitary, the polynomial becomes

$$p(z) = z\bar{z} - 1,$$

meaning $a_{2,0} = a_{0,2} = a_{1,0} = a_{0,1} = 0$, $a_{1,1} = 1$ and $a_{0,0} = -1$. From (8) we have $\alpha = -1/2$, $\beta = 1/2$ and $\gamma = 0$, and hence we can see that everything works starting from the vectors $[\mathbf{u}, \mathbf{v}, A\mathbf{u}, A\mathbf{v}]$ independently of the invertibility of A as required in the previous section to establish the existence of a suitable matrix C.

In general, however, four initial vectors are not sufficient to start with the block tridiagonal reduction supporting the claim that for the given A there exist no matrix C of rank two satisfying $\Delta(A) = CA - AC$. However, due to the relations induced by the minimal polyanalytic polynomial of degree two it is seen that the construction immediately shrinks to size 4 after the first step. In figure 4 we show the shape of the matrix generated from the block-Lanczos procedure applied for the block tridiagonalization of a normal-plus-rank–one matrix where the normal component has eigenvalues located on some arc of parabola in the complex plane.

5. Conclusions

In this paper we have addressed the problem of computing a block tridiagonal matrix unitarily similar to a given almost normal or perturbed normal matrix. A computationally appealing procedure relying upon the block Lanczos method is proposed for this task. The application of the banded reduction for the acceleration of rank-structured matrix computations is an ongoing research topic.

References

[1] R. Bevilacqua, G.M. Del Corso, A condensed representation of almost normal matrices, Linear Algebra and Its Applications 438 (2013) 4408–4425.

[2] D.A. Bini, F. Daddi, L. Gemignani, On the shifted QR iteration applied to companion matrices, Electron. Trans. Numer. Anal. 18 (2004) 137–152 (electronic).

[3] S. Delvaux, M. Van Barel, A Givens-weight representation for rank structured matrices, SIAM J. Matrix Anal. Appl. 29 (2007) 1147–1170.

[4] Y. Eidelman, L. Gemignani, I. Gohberg, Efficient eigenvalue computation for quasiseparable Hermitian matrices under low rank perturbations, Numer. Algorithms 47 (2008) 253–273.

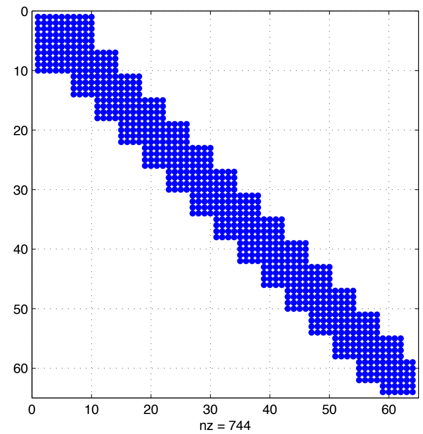

Figure 4: Shape of the block tridiagonal matrix obtained from the block-Lanczos procedure applied to a rank–one correction of a normal matrix whose eigenvalues lie on some arc of parabola

[5] Y. Eidelman, I. Gohberg, On generators of quasiseparable finite block matrices, Calcolo 42 (2005) 187–214.

[6] L. Elsner, K.D. Ikramov, On a condensed form for normal matrices under finite sequences of elementary unitary similarities, Lin. Alg. its Appl. 254 (1997) 79–98.

[7] L. Elsner, K.D. Ikramov, Normal matrices: an update, Lin. Alg. and its Appl. 285 (1998) 291–303.

[8] V. Faber, J. Liesen, P. Tichý, On orthogonal reduction to hessenberg form with small bandwidth, Numerical Algorithms 51 (2009) 133–142.

[9] V. Faber, T. Manteuffel, Necessary and sufficient conditions for the existence of a conjugate gradient method, SIAM J. Numer. Anal. 21 (1984) 352–362.

[10] K. Frederix, S. Delvaux, M. Van Barel, An algorithm for computing the eigenvalues of block companion matrices, Numer. Algorithms 62 (2013) 261–287.

[11] M. Gasemi Kamalvand, K.D. Ikramov, Low-rank perturbations of normal and conjugate-normal matrices and their condensed forms with respect to unitary similarities and congruences, Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet. (2009) 5–11, 56.

[12] R. Grone, C.R. Johnosn, E.M. Sa, H. Wolkowicz, Normal matrices, Lin. Alg. and its Appl. 87 (1987) 213–225.

[13] M.H. Gutknecht, T. Schmelzer, Updating the QR decomposition of block tridiagonal and block Hessenberg matrices, Appl. Numer. Math. 58 (2008) 871–883.

[14] T. Huckle, The Arnoldi method for normal matrices, SIAM J. Matrix Anal. Appl. 15 (1994) 479–489.

[15] M. Huhtanen, Orthogonal polyanalytic polynomials and normal matrices, Math. Comp. 72 (2003) 355–373 (electronic).

[16] K. Ikramov, L. Elsner, On matrices that admit unitary reduction to band form, Mathematical Notes 64 (1998) 753–760.

[17] K.D. Ikramov, On normal band matrices, Zh. Vychisl. Mat. Mat. Fiz. 37 (1997) 3–6.

[18] J. Liesen, P.E. Saylor, Orthogonal hessenberg reuction and orthogonal krylov subspace bases, SIAM J. Numer. Anal. 42 (2005) 2148–2158.

[19] M. Putinar, Linear analysis of quadrature domains. III, J. Math. Anal. Appl. 239 (1999) 101–117.

[20] R. Vandebril, G.M. Del Corso, An implicit multishift *QR*-algorithm for Hermitian plus low rank matrices, SIAM J. Sci. Comp. 16 (2010) 2190–2212.

[21] R. Vandebril, G.M. Del Corso, A unification of unitary similarity transforms to compressed representations, Numerische Mathematik 119 (2011) 641–665.