Singular value decomposition of multi-companion matrices

Boshnakov, Georgi

2006

MIMS EPrint: **2006.362**

ISSN 1749-9097

Singular value decomposition of multi-companion matrices

Georgi N. Boshnakov

School of Mathematics
The University of Manchester

Georgi N. Boshnakov

School of Mathematics
The University of Manchester
Sackville str.
Manchester M60 1QD
UK

E-mail: georgi.boshnakov@manchester.ac.uk
Tel: +44 161 306 3684

Abstract

We obtain the singular value decomposition of multi-companion matrices. We completely characterise the columns of the matrix U and give a simple formula for obtaining the columns of the other unitary matrix, V, from the columns of U. We also obtain necessary and sufficient conditions for the related matrix polynomial to be hyperbolic.

1 Introduction

An $n \times n$ matrix F is said to be multi-companion of order d or d-companion if it has the form

$$F = \begin{pmatrix} A & B \\ I & \mathbf{0} \end{pmatrix}, \tag{1}$$

where A, B, I and $\mathbf{0}$ are matrices of size $d \times (n-d)$, $d \times d$, $(n-d) \times (n-d)$, $(n-d) \times d$, respectively, for some integer d, $1 \leq d < n$. In the important particular case when n is a multiple of d, multi-companion matrices are often called block-companion (or companion) and are extensively studied, see [1]. For companion matrices $d = 1$.

Our interest in the class of multi-companion matrices was motivated by the fact that multi-companion matrices arise naturally in various models for time series in discrete and continuous time, see [2]. We have used the description of the eigenstructure of multi-companion matrices and their factorisation as products of companion matrices obtained in [2] to reparameterize some models for periodically correlated and multivariate time series. This parameterisation allows for generation of models with specified dynamic characteristics since these are determined by the eigenstructure.

Generation of models with specified eigen- or singular structure is of interest also in inverse eigenvalue problems (see [3]) and simulation.

In this paper we describe the singular value decomposition of multi-companion matrices. It turns out that for this problem the block structure that matters most is the one given in Equation (1). We also discuss some other decompositions.

There is a close relationship of this class of matrices to matrix polynomials $\lambda^2 I - \lambda B_0 + C_0$ with Hermitian coefficients. We explore this in some length. We obtain necessary and sufficient conditions for hyperbolicity of a polynomial, $T(\lambda)$, of this type in terms of the matrices A and B or the unit singular values of F.

The singular value decomposition of companion matrices ($d = 1$) has been given in [4] and the polar decomposition in [5]. The singular values (but not the vectors) of block-companion matrices are obtained in [6], who also give a sufficient condition for the hyperbolicity of the respective matrix polynomial. Our results strenghten the latter by providing conditions which are (i) necessary and sufficient, and (ii) hold in the more general case when n is not a multiple of d.

2 Notation

Let $F = U\Sigma V^*$ be the singular value decomposition of F with unitary matrices U and V and nonnegative diagonal matrix Σ. The singular values of F are the arithmetic square roots of the eigenvalues of FF^* (or F^*F). We have

$$H = FF^* = \begin{pmatrix} A & B \\ I & \mathbf{0} \end{pmatrix} \begin{pmatrix} A^* & I \\ B^* & \mathbf{0}^* \end{pmatrix} = \begin{pmatrix} AA^* + BB^* & A \\ A^* & I \end{pmatrix}$$
$$F^*F = \begin{pmatrix} A^* & I \\ B^* & \mathbf{0}^* \end{pmatrix} \begin{pmatrix} A & B \\ I & \mathbf{0} \end{pmatrix} = \begin{pmatrix} A^*A + I & A^*B \\ B^*A & B^*B \end{pmatrix}$$

The blocking of H and F^*F induces natural blocking of their eigenvectors $u = (u_t^* \, u_b^*)^*$ and $v = (v_t^* \, v_b^*)^*$, respectively, such that u_t and v_b are $d \times 1$, while u_b and v_t are $(n-d) \times 1$ (notice the way the sizes match). Subscript t here stands for "top part", b for "bottom part".

The singular values do not change if F is multiplied by unitary matrices. In particular, if columns are permuted. Let

$$F_1 = \begin{pmatrix} B & A \\ \mathbf{0} & I \end{pmatrix}, \qquad F_1^* = \begin{pmatrix} B^* & \mathbf{0}^* \\ A^* & I \end{pmatrix}.$$

Then F_1 has the same eigenvalues as F and, moreover, $F_1F_1^* = FF^*$ since

$$F_1F_1^* = \begin{pmatrix} B & A \\ \mathbf{0} & I \end{pmatrix} \begin{pmatrix} B^* & \mathbf{0}^* \\ A^* & I \end{pmatrix} = \begin{pmatrix} AA^* + BB^* & A \\ A^* & I \end{pmatrix} = FF^*.$$

We say that (λ, x) is an eigenpair of a matrix M if x is an eigenvector of M associated with the eigenvalue λ, i.e. $Mx = \lambda x$, $N(M)$ denotes the null space of M. $S_S^\perp$ is the complement of a subspace S, $S_{BB^*,\lambda}$ is the subspace spanned by all eigenvectors of BB^* associated with the eigenvalue λ.

Let $m = d - \text{rank}(A)$. Usually n is significantly larger than d and all typical features are available for $n > 2d$ but the results below hold for any $n > d$.

3 The eigenvector decomposition of $H = FF^*$

The matrix U can be obtained from the eigenvector decomposition of H since $H = FF^* = U\Sigma VV^*\Sigma U^* = U\Sigma^2U^*$, and so $HU = U\Sigma^2$.

H is Hermitian semidefinite with non-negative eigenvalues $\alpha_1, \ldots, \alpha_n$. The eigenvectors corresponding to unequal eigenvalues are orthogonal and those corresponding to repeated eigenvalues can be chosen to be orthogonal.

Let (λ, u) be an eigenpair of H, i.e. $Hu = \lambda u$. Then using the partitioning of u introduced in section 2 we have

$$Hu = \begin{pmatrix} AA^* + BB^* & A \\ A^* & I \end{pmatrix} \begin{pmatrix} u_t \\ u_b \end{pmatrix} = \begin{pmatrix} \lambda u_t \\ \lambda u_b \end{pmatrix},$$

i.e.,

$$(AA^* + BB^*)u_t + Au_b = \lambda u_t \tag{2}$$

$$A^*u_t + u_b = \lambda u_b \tag{3}$$

We formulate a rearranged version of this as a lemma for easier reference.

Lemma 1. *For any eigenpair (λ, u) of H the following relations hold,*

$$Au_b = (\lambda I - AA^* - BB^*)u_t \tag{4}$$

$$(\lambda - 1)u_b = A^*u_t. \tag{5}$$

It is immediately obvious that the bottom part, u_b, of any eigenvector corresponding to an eigenvalue $\alpha \neq 1$ is determined uniquely by its top part by

$$u_b = (1 - \alpha)^{-1}A^*u_t. \tag{6}$$

Putting expression (6) into Equation (4) gives, after some simplification,

$$T(\lambda)u_t = \mathbf{0}_d,$$

where

$$T(\lambda) = \lambda^2 I - \lambda(I + AA^* + BB^*) + BB^*.$$

Equation (6) shows also that $u_t = \mathbf{0}_d$ and $\lambda \neq 1$ imply $u_b = \mathbf{0}_{n-d}$, i.e. $u = (\mathbf{0}_d^* \, \mathbf{0}_{n-d}^*)^*$. Since the zero vector is not an eigenvector we conclude that the top part of an eigenvector associated with $\lambda \neq 1$ is not zero.

A direct check shows that $T(\lambda)u_t = \mathbf{0}_d$ for eigenvectors associated with $\lambda = 1$, as well. Also, the top part, u_t, of any eigenvector associated with $\alpha = 1$ is in the null space of A^* (i.e. $A^*u_t = \mathbf{0}_{n-d}$). Moreover, setting $u_t = \mathbf{0}_d$ we get $n - 2d + m$ eigenvectors associated with $\alpha = 1$ that have the form $\binom{\mathbf{0}_d}{u_b}$ and are such that u_b's form a basis of the null space of (the rows of) A.

Eigenvectors whose top parts are non-zero and belong to the null space of A^* have special structure.

Lemma 2. *If (λ, u) is an eigenpair of H such that $A^*u_t = \mathbf{0}_{n-d}$, then*

$$(\lambda I - BB^*)u_t = Au_b \tag{7}$$

$$u_t^*(\lambda I - BB^*)u_t = 0 \tag{8}$$

$$\lambda = u_t^* BB^* u_t / u_t^* u_t \qquad (\textit{if } u \neq \mathbf{0}_n). \tag{9}$$

Proof. Equation (7) follows directly from $A^*u_t = \mathbf{0}_{n-d}$ and Lemma 1. To get Equation (8) we left-multiply Equation (7) by u_t^* and notice that $u_t^* A u_b = (A^*u_t)^* u_b = 0$. Solving Equation (8) for λ we get Equation (9). □

The matrix polynomial $T(\lambda)$ appears also as a factor of the determinant of $H - \lambda I$. Indeed, using the formula $\det(M) = \det(M_{22})\det(M_{11} - M_{12}M_{22}^{-1}M_{21})$, (see e.g. [7]) for the determinant of a partitioned matrix, we get

$$\begin{aligned}\det(H - \lambda I) &= (1-\lambda)^{n-d}\det(AA^* + BB^* - \lambda I - \frac{1}{1-\lambda}AA^*) \\ &= \frac{(1-\lambda)^{n-d}}{(1-\lambda)^d}\det((1-\lambda)(AA^* + BB^* - \lambda I) - AA^*) \\ &= (1-\lambda)^{n-2d}\det(\lambda^2 I - \lambda(I + AA^* + BB^*) + BB^*). \end{aligned} \tag{10}$$

Formula (10) agrees with the one obtained in [6] for the case when n is a multiple of d.

More detailed analysis leads to the following theorem whose main features are summarised in Table 1.

Theorem 1. *If $(\alpha_i, u^{(i)})$, $i = 1, \ldots, n$, are the eigenvalue-eigenvector pairs of H, then the top parts of the eigenvectors satisfy $T(\alpha_i)u_t^{(i)} = \mathbf{0}_d$, and the following classification holds.*

1. *eigenvectors with top parts in the nullspace, $N(A^*)$, of A^*.*

   (a) *$n_a = n - 2d + m$ eigenvectors associated with the eigenvalue 1 and determined solely by the matrix A. They have the form $\binom{\mathbf{0}_d}{u_b}$ where the u_b's span the complement, $S_A^\perp$, of the row space, S_A, of A. We call the eigenvectors in this set* structural eigenvectors.

   This is the only group of eigenvectors whose top part may be zero. The top parts of the eigenvectors in the remaining groups cannot be $\mathbf{0}_d$.

   (b) *$n_b = \dim(N(A^*) \cap S_{BB^*,1})$ eigenvectors associated with the eigenvalue 1 and having the form $\binom{u_t}{\mathbf{0}_{n-d}}$, where u_t form a basis for $N(A^*) \cap S_{BB^*,1}$. Here $u_b = \mathbf{0}_{n-d}$ in the sense that it can be chosen to be $\mathbf{0}_{n-d}$.*

Table 1: Eigenstructure of $H = FF^*$. Each eigenvector, $u = (u_t^* \, u_b^*)^*$, is split into a top part, u_t, and a bottom part, u_b. The number of elements in a group, N, is equal to the dimension of the space (where specified) spanned by the top part. The bulk of the elements are in group 1a with eigenvalue 1. The "generic" elements are in group 2 with eigenvalues $\neq 1$. The remaining groups account for specific relations between A and B.

	N	val	vec	top in $N(A^*)$?	top part*	bottom part*
1a	n_a	1	$\binom{\mathbf{0}_d}{u_b}$	Yes	$\mathbf{0}_d$	$S_A^\perp$
1b	n_b	1	$\binom{u_t}{\mathbf{0}_{n-d}}$	Yes	$N(A^*) \cap S_{BB^*,1}$	$\mathbf{0}_{n-d}$
1c	n_c	1	$\binom{u_t}{u_b}$	Yes	$\mathrm{Im}(A) \cap S_K$ †	$Au_b = (I - BB^*)u_t$ ‡
1d	n_{d_1}	$\lambda_1 \neq 1$	$\binom{u_t}{\mathbf{0}_{n-d}}$	Yes	$N(A^*) \cap S_{BB^*,\lambda_1}$	$\mathbf{0}_{n-d}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1d	n_{d_p}	$\lambda_p \neq 1$	$\binom{u_t}{\mathbf{0}_{n-d}}$	Yes	$N(A^*) \cap S_{BB^*,\lambda_p}$	$\mathbf{0}_{n-d}$
2	n_2	$\neq 1$	$\binom{u_t}{u_b}$	No	$\notin N(A^*)$	$\frac{1}{1-\alpha} A^* u_t$

* Vectors are (orthonormal) basis of the designated space or equal the specified value.

† $S_K = \{y : y = (I - BB^*)x, x \in N(A^*) \cap S^\perp_{BB^*,1}\}$.

‡ one u_b for each u_t.

(c) $n_c = \dim(Im(A) \cap \{y : y = (I - BB^*)x, x \in N(A^*) \cap S^{\perp}_{BB^*,1}\})$ *eigenvectors associated with the eigenvalue 1 and having the form* $\binom{u_t}{u_b}$*, where* u_t *are linearly independent vectors such that the vectors* $(I - BB^*)u_t$ *form a basis of the space* $Im(A) \cap \{y : y = (I - BB^*)x, x \in N(A^*) \cap S^{\perp}_{BB^*,1}\}$*,* u_b *is a solution of* $Au_b = (I - BB^*)u_t$ *and can be selected to be the unique solution which is orthogonal to* $N(A)$*.*

(d) $n_{d_1}, \dots, n_{d_p}$*, eigenvectors associated with* $\alpha_1, \dots, \alpha_p$*, respectively, where* $\alpha_i \neq 1$*,* $\alpha_i \neq \alpha_j$ *for* $i \neq j$*, and for each* i*,* $i = 1, \dots, p$*,* α_i *is an eigenvalue of* BB^**,* $n_{d_i} = \dim(N(A^*) \cap S_{BB^*,\lambda_i})$*, and the eigenvectors are of the form* $\binom{u_t}{\mathbf{0}_{n-d}}$*, with tops,* u_t*, forming a basis for the space* $N(A^*) \cap S_{BB^*,\lambda_i}$*.*

(In this case u_b *must be* $\mathbf{0}_{n-d}$*, compare with the above.)*

2. $n - n_a - n_b - n_c - \sum n_{d_i}$ *eigenvectors with top parts not in* $N(A^*)$*.*

Each of these has the form $\binom{u_t}{u_b}$*, where* $\alpha \neq 1$*,* $u_t \notin N(A^*)$*,* (α, u_t) *is a latent pair of* $T(\lambda)$*, i.e.* $T(\alpha)u_t = \mathbf{0}_d$*, and* $u_b = (1 - \alpha)^{-1}A^*u_t$*.*

Remark 1. If $(1, x)$ is a pair, such that $T(1)x = \mathbf{0}_d$, then x may or may not be the top part of an eigenvector of H. And if it is, it is not necessarilly associated with eigenvalue 1. Note that such an $x \in N(A^*)$ and compare to part 2 of the theorem.

Remark 2. Eigenvectors of H associated with unequal eigenvalues are automatically orthogonal. Eigenvectors of repeated eigenvalues may need to be orthogonalised (among themselves) in order to get orthogonal system. If the bases in 1(a)–1(c) are chosen orthogonal, then the eigenvectors associated with 1 will form an orthogonal system as well.

Remark 3. $BB^*u = u$ in (1b) implies $u^*BB^*u = u^*u$.

Remark 4. $u^*BB^*u \neq u^*u$ in (1d).

Remark 5. $A^*u \neq \mathbf{0}_{n-d}$ implies $\lambda \neq 1$ in (2).

With the help of the Cauchy interlace theorem [8, p. 186] it can be shown as in [6] that the smallest d eigenvalues of H are ≤ 1 and the largest d are ≥ 1. Together with parts 1a–1c of Theorem 1 this gives the following corollary.

Corollary 1. *The multiplicity of the eigenvalue 1 is at least* $n - 2d + m$ *and at most* $n - 2d + 2m = n - 2\,\mathrm{rank}(A^*)$*. The difference between the number of eigenvalues greater than 1 and the number of those smaller than 1 is at most* $2m$*.*

Proof of Theorem 1. We have already shown (1a) and (2).

Suppose now that (λ, u) is an eigenpair of H such that $A^* u_t = \mathbf{0}_{n-d}$ and $u \neq 0$. Let $\beta = u_t^* BB^* u_t / u_t^* u_t$. From Lemma 2 it follows that $\beta = \lambda$.

If $\beta \neq 1$ then from $A^* u_t = \mathbf{0}_{n-d}$ and Equation (5) we get $u_b = \mathbf{0}_{n-d}$, which used together with Equation (7) leads to $BB^* u_t = \lambda u_t$. This shows that (λ, u_t) is an eigenpair of BB^* and $\sqrt{\lambda}$ is a singular value of B. Hence, $u = \binom{u_t}{\mathbf{0}_{n-d}}$, where u_t is an eigenvector of BB^* associated with eigenvalue λ. The number of linearly independent vectors of this type for β is $\dim(N(A^*) \cap S_{BB^*,\lambda_i})$. Therefore (1d) holds.

If $\beta = 1$ then the equations (5) and $A^* u_t = \mathbf{0}_{n-d}$ give $\mathbf{0}_{n-d} = 0 u_b$, which does not provide a unique expression for u_b. But Equation (7) gives (since $\beta = \lambda = 1$) $Au_b = (I - BB^*)u_t$. There are two subcases here.

- $(I - BB^*)u_t = \mathbf{0}_d$, i.e. $(1, u_t)$ is an eigenpair of BB^* and $Au_b = \mathbf{0}_d$.

  This is similar to (1d) except that the bottom part of u is not necessarily $\mathbf{0}_{n-d}$. However, it must be zero if we require u to be orthogonal to the structural eigenvectors. Indeed, $Au_b = \mathbf{0}_d$ shows that u_b is in the null space of A which is spanned by the bottom parts of the structural eigenvectors. Hence, the vector $\binom{u_t}{u_b}$ will be orthogonal to the structural eigenvectors if and only if $u_b = \mathbf{0}_{n-d}$. So, (1b) holds.

- $(I - BB^*)u_t \neq \mathbf{0}_d$.

  Here, $Au_b \neq \mathbf{0}_d$. For $\binom{u_t}{u_b}$ to be orthogonal to the structural eigenvectors, u_b should be the unique orthogonal to $N(A)$ solution of $Au_b = (I - BB^*)u_t$. It is now not difficult to get (1c).

From the above it follows that if (λ, u) is an eigenpair of H and (λ, u_t) is an eigenpair of BB^*, then u_b, the bottom part of u, is $\mathbf{0}_{n-d}$. The opposite is also true. Suppose that u is any eigenvector of H with $u_b = \mathbf{0}_{n-d}$. Then from Lemma 1 we get $A^* u_t = \mathbf{0}_{n-d}$ and $(\lambda I - AA^* - BB^*)u_t = \mathbf{0}_d$. Hence, $(\lambda I - BB^*)u_t = \mathbf{0}_d$ or $BB^* u_t = \lambda u_t$. Hence (λ, u_t) is an eigenpair of BB^*. □

4 Singular value decomposition of F

We turn our attention now to the matrix V. If F is non-singular we may obtain V from the formula $V = F^{-1} U \Sigma$. This does not work however if F is singular. It turns out that there is a neat relation between V and U which does not involve any inverse matrices.

The columns of V satisfy

$$Fv = \lambda u, \tag{11}$$

i.e.

$$\begin{pmatrix} A & B \\ I & \mathbf{0} \end{pmatrix} \begin{pmatrix} v_t \\ v_b \end{pmatrix} = \begin{pmatrix} \lambda u_t \\ \lambda u_b \end{pmatrix}.$$

Note that the blocks of u and v have different dimensions. So,

$$Av_t + Bv_b = \lambda u_t \tag{12}$$

$$v_t = \lambda u_b \tag{13}$$

From these two equations we deduce that $\lambda Au_b + Bv_b = \lambda u_t$ and so,

$$Bv_b = \lambda(u_t - Au_b).$$

From equations (2)–(3) we get (left-multiply (3) by A and subtract from (2))

$$BB^* u_t = \lambda(u_t - Au_b). \tag{14}$$

Hence,

$$Bv_b = BB^* u_t,$$

which can be written also as

$$B(v_b - B^* u_t) = \mathbf{0}_d.$$

One solution is $v_b = B^* u_t$ (it is the only one if B is non-singular).

The above discussion suggests that a vector v corresponding to u and λ may be obtained by the formula

$$v = \begin{pmatrix} \lambda u_b \\ B^* u_t \end{pmatrix}. \tag{15}$$

A direct check shows that it indeed satisfies equation (11),

$$\begin{aligned} Fv = \begin{pmatrix} A & B \\ I & \mathbf{0} \end{pmatrix} \begin{pmatrix} \lambda_i u_b \\ B^* u_t \end{pmatrix} &= \begin{pmatrix} \lambda A u_b + BB^* u_t \\ \lambda u_b \end{pmatrix} = \begin{pmatrix} AA^* u_t + Au_b + BB^* u_t \\ \lambda u_b \end{pmatrix} \\ = \begin{pmatrix} \lambda u_t \\ \lambda u_b \end{pmatrix} = \lambda u \end{aligned}$$

To check if such vectors form an orthogonal system let $u^{(1)}$, $u^{(2)}$ be two eigenvectors of H and x_1, x_2 be related to them via Equation (15),

$$x_1 = \begin{pmatrix} \lambda_1 u_b^{(1)} \\ B^* u_t^{(1)} \end{pmatrix}, \; x_2 = \begin{pmatrix} \lambda_2 u_b^{(2)} \\ B^* u_t^{(2)} \end{pmatrix}, \quad \text{where} \quad u^{(1)} = \begin{pmatrix} u_t^{(1)} \\ u_b^{(1)} \end{pmatrix}, \; u^{(2)} = \begin{pmatrix} u_t^{(2)} \\ u_b^{(2)} \end{pmatrix}.$$

Then

$$
\begin{aligned}
x_1^* x_2 &= \lambda_1\lambda_2 u_b^{(1)*} u_b^{(2)} + u_t^{(1)*} BB^* u_t^{(2)} \\
&= \lambda_1\lambda_2 u_b^{(1)*} u_b^{(2)} + u_t^{(1)*} \lambda_2 (u_t^{(2)} - A u_b^{(2)}) \qquad \text{(by Equation (14))} \\
&= \lambda_2 \left(\lambda_1 u_b^{(1)*} u_b^{(2)} + u_t^{(1)*} (u_t^{(2)} - A u_b^{(2)}) \right) \\
&= \lambda_2 (\lambda_1 u_b^{(1)*} u_b^{(2)} + u_t^{(1)*} u_t^{(2)} - u_t^{(1)*} A u_b^{(2)}) \\
&= \lambda_2 (u_t^{(1)*} u_t^{(2)} + \lambda_1 u_b^{(1)*} u_b^{(2)} - u_t^{(1)*} A u_b^{(2)}) \\
&= \lambda_2 (u_t^{(1)*} u_t^{(2)} + (\lambda_1 u_b^{(1)*} - u_t^{(1)*} A) u_b^{(2)}) \\
&= \lambda_2 (u_t^{(1)*} u_t^{(2)} + u_b^{(1)*} u_b^{(2)}) \qquad \text{(by Equation (5))} \\
&= \lambda_2 (u^{(1)})^* u^{(2)}.
\end{aligned}
$$

The last expression is zero if $u^{(1)}$ and $u^{(2)}$ are orthogonal, as required.

This calculation shows also that $x_1^* x_1 = \lambda_1$. So we can normalise these vectors if $\lambda_1 \neq 0$. If $\lambda_1 = 0$, then $x_1 = 0$ (by Theorem 1,part 1d) and therefore equation (15) does not provide eigenvectors for $\lambda = 0$.

When $\lambda = 0$ the system (12)–(13) becomes

$$
\begin{aligned}
A v_t + B v_b &= \mathbf{0}_d \\
v_t &= \mathbf{0}_{n-d}.
\end{aligned}
$$

Hence, $v_t = \mathbf{0}_{n-d}$ and $B v_b = \mathbf{0}_d$. So, a system of orthonormal vectors such that $B v_b = \mathbf{0}_d$ will provide the required vectors corresponding to the zero singular value(s) via the formula

$$
v = \begin{pmatrix} \mathbf{0}_{n-d} \\ v_b \end{pmatrix}, \qquad \text{where } B v_b = \mathbf{0}_d.
$$

Since

$$
x_1^* v = u_t^{(1)*} B v_b = 0,
$$

these vectors are orthogonal to the vectors associated with $\lambda \neq 0$ and thus we get an orthogonal system of v's.

Theorem 2. *The singular value decomposition of the multi-companion matrix F is $F = U\Sigma V^*$, where*

1. *the singular values of F are $\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_n}$, the square roots of the zeroes of $\det(T(\lambda))$. At least $n - 2d + m$ of them are equal to 1.*

2. *$\Sigma = diag(\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_n})$.*

3. *the columns,* $\begin{pmatrix} u_t^{(i)} \\ u_b^{(i)} \end{pmatrix}$, $i = 1, \ldots, n$, *of* U *are an orthonormal set of eigenvectors of* $H = FF^*$ *having the structure described in Theorem 1. The* i*th column,* $u^{(i)}$, *corresponds to* α_i.

4. *for* $i = 1, \ldots, n$, *the* i*th column,* $v^{(i)}$, *of* V *is given by*

$$v^{(i)} = \begin{cases} \frac{1}{\sqrt{\lambda_i}} \begin{pmatrix} \lambda_i u_b^{(i)} \\ B^* u_t^{(i)} \end{pmatrix} & \lambda_i \neq 0, \\ \begin{pmatrix} \mathbf{0}_{n-d} \\ v_b^{(i)} \end{pmatrix} & \lambda_i = 0, \end{cases}$$

where the components $v_b^{(i)}$ *corresponding to singular values equal to zero are chosen to form an orthonormal basis of the null space of* B.

5 Reducing the dimension of the problem

Let $A = U_A D_A V_A^*$ be the singular value decomposition of A. Then the matrix F_1 can be factored as

$$\begin{pmatrix} B & A \\ \mathbf{0} & I \end{pmatrix} = \begin{pmatrix} U_A & 0 \\ 0 & V_A \end{pmatrix} \begin{pmatrix} U_A^* B U_A & D_A \\ 0 & I \end{pmatrix} \begin{pmatrix} U_A^* & 0 \\ 0 & V_A^* \end{pmatrix},$$

where the left and right factors on the right are orthogonal. Hence F has the same singular values as

$$F_2 \equiv \begin{pmatrix} U_A^* B U_A & D_A \\ 0 & I \end{pmatrix}.$$

D_A contains the non-zero singular values of A in the upper left corner and is zero elsewhere. Therefore $F_2 = \left(\begin{smallmatrix} X & 0 \\ 0^* & I \end{smallmatrix} \right)$, where X is of size $(2d - m) \times (2d - m)$. So, we have reduced the original problem of finding the singular value decomposition of the $n \times n$ matrix F to that for the smaller matrix F_2.

5.1 Diagonal case

Of some interest is the case when $U_A^* B U_A = D_1 Q_1^*$ with D_1 diagonal and Q_1 unitary (in particular, identity). In this case F_2 can be further simplified to have diagonal blocks,

$$F_2 = \begin{pmatrix} D_1 & D_A \\ \mathbf{0} & I \end{pmatrix} \begin{pmatrix} Q_1^* & 0 \\ 0 & I \end{pmatrix}.$$

Note that the diagonal elements of D_1 are the singular values of B but not necessarily ordered. Then

$$F_3 \equiv \begin{pmatrix} D_1 & D_A \\ \mathbf{0} & I \end{pmatrix} \begin{pmatrix} D_1^* & 0^* \\ D_A^* & I \end{pmatrix} = \begin{pmatrix} D_1 D_1^* + D_A D_A^* & D_A \\ D_A^* & I \end{pmatrix},$$

and

$$\begin{aligned} \det(F_3 - \lambda I) &= \det((1-\lambda)I) \det(D_1 D_1^* + D_A D_A^* - \lambda I - \frac{1}{1-\lambda} D_A D_A^*) \\ &= (1-\lambda)^d \det(\operatorname{diag}(d_{1i}^2 + d_{2i}^2 - \lambda - \frac{1}{1-\lambda} d_{2i}^2)) \\ &= (1-\lambda)^{d-d} \prod((1-\lambda)(d_{1i}^2 + d_{2i}^2 - \lambda) - d_{2i}^2) \\ &= \prod(\lambda^2 - \lambda(1 + d_{1i}^2 + d_{2i}^2) + d_{1i}^2 + d_{2i}^2 - d_{2i}^2). \end{aligned}$$

Hence,

$$\det(F_3 - \lambda I) = \prod_{i=1}^{d} (\lambda^2 - \lambda(1 + d_{1i}^2 + d_{2i}^2) + d_{1i}^2). \tag{16}$$

Kittaneh's result [4] about companion matrices (i.e. $d = 1$) is obtained as a particular case. Let $d = 1$, $B = -a_1$, and $A = -(a_n, \dots, a_2)$. With this choice of A and B the matrix F is companion. The singular value of B is $d_{11} = |a_1|$, the non-zero singular value of A is $d_{2i} = \sqrt{\sum_{i=2}^{n} |a_i|^2}$. Formula (16) takes the form

$$(\lambda^2 - \lambda(1 + |a_1|^2 + \sum_{i=2}^{n} |a_i|^2) + |a_1|^2) = (\lambda^2 - \lambda(1 + \sum_{i=1}^{n} |a_i|^2) + |a_1|^2), \tag{17}$$

If A is not zero it is of full rank 1 and $m = 0$. Then the two non-trivial singular values of a companion matrix are the zeroes of this polynomial which agrees with the result of Kittaneh.

Otherwise, if $A = 0$, then $m = 1$, $\sum_{i=2}^{n} |a_i|^2 = 0$ and we get from the above

$$(\lambda^2 - \lambda(1 + |a_1|^2) + |a_1|^2) = (1 - \lambda)(|a_1|^2 - \lambda),$$

giving singular values 1 and $|a_1|$.

6 Hyperbolicity of $T(\lambda)$

The matrix polynomial

$$T(\lambda) = \lambda^2 I - \lambda(I + AA^* + BB^*) + BB^*$$

plays an important role in the singular value decomposition of F. In this section we exploit this relationship to study some properties of $T(\lambda)$.

The description of the properties of a general quadratic $\lambda^2 I - \lambda B_0 + C_0$ with Hermitian positive definite B_0 and positive semidefinite C_0 is particularly efficient when it is hyperbolic (or overdamped), i.e. $D_x > 0$ where D_x is the discriminant of the quadratic polynomial $x^*T(\lambda)x$, see [9].

For $T(\lambda)$ we have $B_0 - C_0 \geq I$. More importantly, the relation of $T(\lambda)$ to H provides some properties automatically. For example, from the analysis in section 3 it follows that all latent roots of this polynomial are real and non-negative and that there exist d linearly independent latent vectors of $T(\lambda)$.

The following observations are also useful.

$$T(\lambda) = \begin{cases} \lambda^2 I + |\lambda|(I + AA^* + BB^*) + BB^* > 0 & \text{for } \lambda < 0 \\ BB^* \geq 0 & \text{for } \lambda = 0 \\ -AA^* \leq 0 & \text{for } \lambda = 1 \end{cases} \tag{18}$$

We see that $x^*T(\lambda)x$ cannot be zero for negative λ. It can be zero at $\lambda = 1$ only when AA^* is singular (equivalently, when A^* is not of full column rank). Similarly, $x^*T(0)x$ can be zero only when BB^* is singular (equivalently, when B^* is not of full column rank).

Consider the quadratic form

$$\begin{aligned} x^*T(\lambda)x &= x^*(\lambda^2 I - \lambda(I + AA^* + BB^*) + BB^*)x \\ &= \lambda^2 x^*x - \lambda x^*(I + AA^* + BB^*)x + x^*BB^*x \\ &= a_x\lambda^2 - b_x\lambda + c_x \end{aligned}$$

with obvious definition of a_x, etc. Let $D_x = b_x^2 - 4a_xc_c$ be the discriminant of the quadratic. The roots of $x^*T(\lambda)x = 0$ are all real for all $x \neq \mathbf{0}_d$ if and only if $D_x \geq 0$ for all $x \neq \mathbf{0}_d$. Factorising the discriminant we get

$$\begin{aligned} D_x &= (x^*(I + AA^* + BB^*)x)^2 - 4x^*xx^*BB^*x \\ &= (x^*x + x^*AA^*x + x^*BB^*x)^2 - 4x^*xx^*BB^*x \\ &= \left(x^*x + x^*AA^*x + x^*BB^*x - 2\sqrt{x^*x}\sqrt{x^*BB^*x}\right) \\ &\quad \times \left(x^*x + x^*AA^*x + x^*BB^*x + 2\sqrt{x^*x}\sqrt{x^*BB^*x}\right) \\ &= \left((\sqrt{x^*x} - \sqrt{x^*BB^*x})^2 + x^*AA^*x\right)\left((\sqrt{x^*x} + \sqrt{x^*BB^*x})^2 + x^*AA^*x\right). \end{aligned}$$

The second factor in the last expression is always positive. The first factor is non-negative and is zero only if both $x^*AA^*x = \mathbf{0}_d$ and $x^*BB^*x = x^*x$.

The former is equivalent to $A^*x = \mathbf{0}_{n-d}$, the latter can be written equivalently as $x^*(I - BB^*)x = 0$. This case is treated by parts 1b–1c of Theorem 1. Notice also that $b_x \geq \sqrt{D_x}$.

We summarise the above in the following lemma.

Lemma 3. *The discriminant of the equation $x^*T(\lambda)x = 0$, $x \neq \mathbf{0}_d$, is*

$$D_x = \left((\sqrt{x^*x} - \sqrt{x^*BB^*x})^2 + x^*AA^*x\right)\left((\sqrt{x^*x} + \sqrt{x^*BB^*x})^2 + x^*AA^*x\right). \tag{19}$$

D_x is always non-negative and all roots of $x^*T(\lambda)x = 0$ are real and nonnegative. D_x is zero if and only if both $x^*AA^*x = 0$ and $x^*BB^*x = x^*x$.

Hence, we have the following proposition.

Theorem 3. *$T(\lambda)$ is hyperbolic if and only if $x^*(I - BB^*)x \neq 0$ for each nonzero vector in the null space of A^*.*

From Theorem 1 we can see that the above condition is equivalent to the statement that FF^* has more than $n - 2d + m$ eigenvalues equal to 1. So, we can determine hyperbolicity by simply counting the unit roots of $\det(T(\lambda)) = 0$.

Theorem 4. *$T(\lambda)$ is hyperbolic if and only if $\lambda = 1$ is a root of $\det(T(\lambda)) = 0$ with multiplicity equal to $m = d - \operatorname{rank}(A)$.*

The following two useful sufficient conditions follow directly from the above results.

Corollary 2. *If A is of full rank then $T(\lambda)$ is hyperbolic.*

Corollary 3. *If $x^*(I - BB^*)x$ is definite (positive or negative), then $T(\lambda)$ is hyperbolic.*

In the particular case when n is a multiple of d the matrix A can be written in block form as $A = [A_1\, A_2\, \ldots\, A_{n/d}]$, where A_i are square. A sufficient, but not necessary, condition for A to be of full rank is that at least one of these blocks is non-singular. Thus, Corollary 2 strengthens Lemma 2.8 of [6]. (Note that in [6] the authors obtain their criterion from an inequality, rather than equality, for the discriminant, which cannot be used directly for an only if statement.)

For an eigenvector, say x, from Part 1b or 1c of Theorem 1 $x^*T(\lambda)x = (\lambda - 1)^2 x^*x$ with a double root, 1. This is yet another manifestation that if such eigenvectors exist, then the problem is not hyperbolic since for hyperbolic problems repeated eigenvalues never come in pairs.

6.1 Further remarks on hyperbolicity

The two corollaries may be obtained alternatively by using the fact that $T(\lambda)$ is hyperbolic if and only if the matrix $T(\lambda_0)$ is negative definite for some real λ_0 (see [1],[9]).

Indeed, $T(1) = -AA^*$ is negative definite if A is of full rank. Further,

$$T(\lambda) = \lambda^2 I - \lambda(I + AA^* + BB^*) + BB^* = -\left((1-\lambda)(\lambda I - BB^*) + \lambda AA^*\right)$$

If all eigenvalues of BB^* are smaller than one, then $(\lambda I - BB^*)$ is positive definite for values of λ close to 1 and then $T(\lambda)$ is negative definite. Similarly, if all eigenvalues of BB^* are greater than 1, then $(\lambda I - BB^*)$ is negative definite for values larger than 1 and smaller than the minimal eigenvalue of BB^*. In this case $(1-\lambda)(\lambda I - BB^*)$ is positive definite and $T(\lambda)$ is negative definite. A particularly efficient way to see all this at a glance is to note that since $BB^* = HDH^*$ for some diagonal D and unitary H, we have

$$\lambda I - BB^* = \lambda HH^* - HDH^* = H(\lambda I - D)H^* = H\mathrm{diag}(\lambda - d_1, \dots, \lambda - d_d)H^*.$$

7 Related decompositions

Once the decomposition $F = U\Sigma V^*$ is available, other decompositions are easy to determine. From $U\Sigma V^* = U\Sigma U^* UV^*$ we obtain the (left) polar decomposition $F = PQ$ with $P = U\Sigma U^*$ positive semidefinite and $Q = UV^*$ unitary.

Decompositions for the other possible multi-companion matrices where the non-trivial rows are at the bottom, or are columns rather than rows, are also straightforward.

References

[1] I. Gohberg, P. Lancaster, L. Rodman, Matrix polynomials., Computer Science and Applied Mathematics. New York etc.: Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers, 1982.

[2] G. N. Boshnakov, Multi-companion matrices, Linear Algebra and Its Applications 354 (1–3) (2002) 53–83.

[3] M. T. Chu, G. H. Golub, Inverse eigenvalue problems: theory, algorithms, and applications., Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press., 2005.

[4] F. Kittaneh, Singular values of companion matrices and bounds on zeros of polynomials., SIAM J. Matrix Anal. Appl. 16 (1) (1995) 333–340.

[5] P. van den Driessche, H. K. Wimmer, Explicit polar decomposition of companion matrices., Electron. J. Linear Algebra 1 (1996) 64–69.

[6] N. J. Higham, F. Tisseur, Bounds for eigenvalues of matrix polynomials., Linear Algebra Appl. 358 (1-3) (2003) 5–22.

[7] J. R. Magnus, H. Neudecker, Matrix differential calculus with applications in statistics and econometrics. Revised edition., Wiley Series in Probability and Statistics. Chichester: Wiley., 1999.

[8] B. N. Parlett, The symmetric eigenvalue problem., Prentice-Hall Series in Computational Mathematics. Englewood Cliffs, New Jersey: Prentice-Hall, 1980.

[9] P. Lancaster, Quadratic eigenvalue problems., Linear Algebra Appl. 150 (1991) 499–506.