

# Structures preserved by matrix inversion

*Steven Delvaux      Marc Van Barel*

*Report TW 414, December 2004*



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

# Structures preserved by matrix inversion

*Steven Delvaux*      *Marc Van Barel*

*Report TW414, December 2004*

Department of Computer Science, K.U.Leuven

## Abstract

In this paper we investigate some matrix structures on  $\mathbb{C}^{n \times n}$  that have a good behaviour under matrix inversion. The first type of structure is closely related to low displacement rank matrices. Next, we show that for a matrix having a low rank submatrix, also the inverse matrix must have a low rank submatrix, which we can explicitly determine. This allows us to generalize a theorem due to Fiedler. The generalization consists in the fact that our rank structures may have their own shift matrix  $\Lambda_k \in \mathbb{C}^{m \times m}$ , for suitable  $m$ , with Fiedler's theorem corresponding to the limiting cases  $\Lambda_k \rightarrow 0$  and  $\Lambda_k \rightarrow \infty I$ .

**Keywords :** displacement structures, Hermitian plus low rank, rank structures, lower semiseparable (plus diagonal) matrices, matrix inversion.

**AMS(MOS) Classification :** Primary : 15A09, Secondary : 15A03,65F05.

# Structures preserved by matrix inversion

Steven Delvaux \*, Marc Van Barel \*

24th December 2004

## Abstract

In this paper we investigate some matrix structures on  $\mathbb{C}^{n \times n}$  that have a good behaviour under matrix inversion. The first type of structure is closely related to low displacement rank matrices. Next, we show that for a matrix having a low rank submatrix, also the inverse matrix must have a low rank submatrix, which we can explicitly determine. This allows us to generalize a theorem due to Fiedler. The generalization consists in the fact that our rank structures may have their own shift matrix  $\Lambda_k \in \mathbb{C}^{m \times m}$ , for suitable  $m$ , with Fiedler's theorem corresponding to the limiting cases  $\Lambda_k \rightarrow 0$  and  $\Lambda_k \rightarrow \infty I$ .

**Keywords:** displacement structures, Hermitian plus low rank, rank structures, lower semiseparable (plus diagonal) matrices, matrix inversion

## 1 Introduction

The aim of this paper is to handle structures of a matrix  $A \in \mathbb{C}^{n \times n}$  that carry over to its inverse.

Section 2 deals with the inversion of displacement structures. The idea is to generalize the classical examples of displacement structures (for example Toeplitz-like, Cauchy-like, Vandermonde-like, circulant matrices etc., see [11]), by ‘decoupling’ the displacement equation. This means that the displacement equation is allowed to involve two variables  $A$  and  $B$  rather than only one variable  $A$ . We will then illustrate these decoupled displacement structures by

---

\*Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Leuven (Heverlee), Belgium. email: {Steven.Delvaux,Marc.VanBarel}@cs.kuleuven.ac.be.

The research was partially supported by the Research Council K.U.Leuven, project OT/00/16 (SLAP: Structured Linear Algebra Package), by the Fund for Scientific Research–Flanders (Belgium), projects G.0078.01 (SMA: Structured Matrices and their Applications), G.0176.02 (ANCILA: Asymptotic aNalysis of the Convergence behavior of Iterative methods in numerical Linear Algebra), G.0184.02 (CORFU: Constructive study of Orthogonal Functions) and G.0455.0 (RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation), and by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, project IUAP V-22 (Dynamical Systems and Control: Computation, Identification & Modelling). The scientific responsibility rests with the authors.

some examples; one of these examples involves Hermitian plus low rank matrices, for which we provide an alternative characterization.

Section 3 handles the inversion of *rank structures*. A rank structure on  $\mathbb{C}^{n \times n}$  will be defined as a collection of structure blocks  $\{\mathcal{B}_k\}_k$ : these are low rank submatrices which lie in the bottom left corner of a given matrix  $A \in \mathbb{C}^{n \times n}$ , together with a certain correction term for the block diagonal positions of  $A$ , called the shift matrix. Suppose then that  $\mathcal{B}$  is a structure block according to the above definition. It turns out that there can be defined in a natural way an inverse structure block  $\mathcal{B}^{-1}$ , by just replacing the shift matrix  $\Lambda$  by  $\Lambda^{-1}$  (assuming  $\Lambda$  is nonsingular).

Section 4 handles some generalizations of Section 3. A first generalization is the inversion of structure blocks whose shift matrix  $\Lambda$  is singular. To solve this problem, we make a reduction to the case where  $\Lambda = \Lambda_{\text{ns}} \oplus 0$ , with  $\Lambda_{\text{ns}}$  nonsingular. Then it could be expected that  $\Lambda^{-1} = \Lambda_{\text{ns}}^{-1} \oplus \infty I$ , where  $\infty I$  is the diagonal matrix with diagonal entries ‘equal to  $\infty$ ’; we will show that there can be given an exact meaning to this statement.

A second generalization is to absorb permutation matrices into the structure. This allows us to move the structure blocks to an arbitrary matrix position, not necessarily situated in the bottom left matrix corner anymore. Moreover, in the case of shift matrix  $\Lambda \rightarrow 0$  or  $\Lambda \rightarrow \infty I$ , we obtain in this way an alternative derivation of a theorem due to Fiedler [5] (see Corollary 20 further on).

It should be emphasized that Fiedler’s theorem will be just a special case of the theory. Moreover, there are several connections between Section 2, 3 and 4, in the sense that the main results for rank structures can be derived as a special case of the ‘decoupled’ displacement structures of Section 2, as will become clear soon.

## 2 Displacement structures

In this section we handle displacement structure. As a general reference, we can refer to [11, 10] for an overview of the many applications of displacement theory in numerical linear algebra. Some references of historical interest are [6, 9]. For our purposes, however, we will only be interested in matrix inversion.

Let us start with some classical examples of displacement structure. Let  $A$  be a Toeplitz matrix, i.e.  $A = [a_{i-j}]_{i,j=1}^n$ . Putting  $Z := [\mathbf{e}_2 \dots \mathbf{e}_n \mathbf{0}]$  with  $\mathbf{e}_k$  the  $k$ th column of the identity matrix, it is easy to check that

$$A - ZAZ^T = \text{Rk } 2, \tag{1}$$

with  $\text{Rk } 2$  denoting a matrix of rank at most 2. Therefore, Toeplitz matrices can be embedded in the class of *Toeplitz-like* matrices, i.e. the class of matrices  $A$  satisfying (1).

Let  $A$  be a Cauchy matrix, i.e.  $A = [\frac{1}{x_i - y_j}]_{i,j=1}^n$ . It is easy to check that

$$D_{\mathbf{x}}A - AD_{\mathbf{y}} = \text{Rk } 1, \tag{2}$$

with  $\text{Rk } 1$  denoting a matrix of rank at most 1. (Here we are using the notation  $D_{\mathbf{x}} = \text{diag}(x_i)_i$  for any vector  $\mathbf{x}$ ). Therefore, Cauchy matrices can be embedded in the class of *Cauchy-like* matrices, i.e. the class of matrices  $A$  satisfying (2).

Now we come to some formal definitions. The main difference with (1) and (2) is that the variable  $A$  is ‘decoupled’ into two variables  $A$  and  $B$ .

**Definition 1** *Let  $G, H \in \mathbb{C}^{m \times n}$  and let  $r \in \mathbb{N}$ . Then for  $A \in \mathbb{C}^{m \times m}$  and  $B \in \mathbb{C}^{n \times n}$ ,*

1. *we say  $A$  and  $B$  to satisfy the Stein type displacement equation induced by  $(G, H, r)$  if*

$$A - GBH^T = \text{Rk } r, \quad (3)$$

*where  $\text{Rk } r$  denotes a matrix of rank at most  $r$ .*

2. *We say  $A$  and  $B$  to satisfy the Sylvester type displacement equation induced by  $(G, H, r)$  if*

$$AG - HB = \text{Rk } r, \quad (4)$$

*where  $\text{Rk } r$  denotes a matrix of rank at most  $r$ .*

Note that the dimension requirements in Definition 1 are equivalent to saying that  $A$  and  $B$  are square and the block matrix

$$\begin{bmatrix} B^{-1} & H^T \\ G & A \end{bmatrix} \quad (5)$$

has compatible matrix dimensions. (Here we assumed that  $B^{-1}$  exists). Moreover, this block representation is useful in the sense that the left hand side of (3) can be realized as a Schur complement in (5). This is the basis of the following inversion result.

**Theorem 2** *(Stein type inversion:)* *With the notation of Definition 1, if  $A$  and  $B$  are nonsingular matrices satisfying*

$$A - GBH^T = \text{Rk } r, \quad (6)$$

*then the inverses satisfy*

$$B^{-1} - H^T A^{-1} G = \widetilde{\text{Rk}} \tilde{r}, \quad (7)$$

*where  $\widetilde{\text{Rk}} \tilde{r}$  denotes a matrix of rank at most  $\tilde{r} := r + n - m$ .*

PROOF. (Based on [11], Lemma 1.5.1). Let us denote  $A \sim B$  if these matrices can be obtained out of each other by elementary Gaussian row and/or column operations. Consider the embedded matrix (5). By the nonsingularity of  $B^{-1}$ , we can use it as pivot block for a Gaussian elimination process, and the Schur complement formula yields

$$\begin{bmatrix} B^{-1} & H^T \\ G & A \end{bmatrix} \sim \begin{bmatrix} B^{-1} & 0 \\ 0 & A - GBH^T \end{bmatrix}. \quad (8)$$

Similarly, by the nonsingularity of  $A$ , we can use it also as pivot block and obtain

$$\begin{bmatrix} B^{-1} & H^T \\ G & A \end{bmatrix} \sim \begin{bmatrix} B^{-1} - H^T A^{-1} G & 0 \\ 0 & A \end{bmatrix}. \quad (9)$$

The proof can then be finished by comparing (8) and (9), and by using the fact that  $\sim$  is a rank-preserving relation.  $\square$

We come to the second, more easy inversion result.

**Theorem 3** (*Sylvester type inversion:*) *With the notation of Definition 1, if  $A$  and  $B$  are nonsingular matrices satisfying*

$$AG - HB = \text{Rk } r, \quad (10)$$

*then the inverses satisfy*

$$GB^{-1} - A^{-1}H = \widetilde{\text{Rk}} r, \quad (11)$$

*where  $\widetilde{\text{Rk}} r := A^{-1}(\text{Rk } r)B^{-1}$  denotes a new matrix of rank at most  $r$ .*

PROOF. This follows immediately by multiplying (10) on the left with  $A^{-1}$  and on the right with  $B^{-1}$ .  $\square$

**Remark 4** *The above proof remains valid if we add a ‘quadratic’ and ‘constant’ term to the structure, i.e. if we replace (10) by*

$$AFB + AG - HB + J = \text{Rk } r,$$

*and (11) by*

$$F + GB^{-1} - A^{-1}H + A^{-1}JB^{-1} = \widetilde{\text{Rk}} r,$$

*for certain  $F, G, H, J \in \mathbb{C}^{m \times n}$ . Thus we see that the quadratic term  $AFB$  has been transformed into the constant term  $F$  under matrix inversion, and vice versa.*

Let us illustrate why it is useful to ‘decouple’ the variable  $A$  into two variables  $A$  and  $B$ . As an illustrative example, we will focus on the following displacement equation.

**Corollary 5** *Let  $A$  and  $B$  be nonsingular matrices satisfying*

$$A - B = \text{Rk } r, \quad (12)$$

*then the inverses satisfy*

$$A^{-1} - B^{-1} = \widetilde{\text{Rk}} r. \quad (13)$$

This corollary is a special case of both the Stein and Sylvester type inversion results. Hence the matrix  $\widetilde{\text{Rk}} r$  occurring in (13) is given explicitly by  $-A^{-1}(\text{Rk } r)B^{-1}$  or  $-B^{-1}(\text{Rk } r)A^{-1}$ , by the proof of Theorem 3. Yet another explicit formula for  $\widetilde{\text{Rk}} r$  is the so-called *Sherman-Morrisson-Woodbury formula* [7, Section 2.1], stating that

$$\widetilde{\text{Rk}} r = -B^{-1}U(V^T B^{-1}U + I)^{-1}V^T B^{-1},$$

where we assumed a decomposition  $\text{Rk } r = UV^T$  with  $U, V \in \mathbb{C}^{n \times r}$ . Note that this formula does not involve  $A^{-1}$  anymore.

Now we come to some illustrations of Corollary 5. First we can take  $B$  to be unitary. Note that the inverse matrix  $B^{-1}$  is again unitary; hence Corollary 5 reveals the following fact

**Corollary 6** *The property  $A = \text{Uni} + \text{Rk } r$ , i.e.  $A$  is unitary plus rank at most  $r$  is preserved by matrix inversion.*

We could state a similar property for the property  $A = \text{Herm} + \text{Rk } r$ , i.e.  $A$  is Hermitian plus rank at most  $r$ . But here we encounter the problem that for a nonsingular matrix  $A = \text{Herm} + \text{Rk } r$ , the Hermitian component  $\text{Herm}$  does not necessarily have to be nonsingular too. This means that the nonsingularity conditions in Corollary 5 will not always be satisfied. Nevertheless, it turns out that the preservation of Hermitian plus low rank structure under inversion will still be valid, even if this nonsingularity assumption is not satisfied.

To show this, we will give an alternative characterization of the property  $A = \text{Herm} + \text{Rk } r$ .

We recall the following definition.

**Definition 7** *Two Hermitian matrices  $A, B \in \mathbb{C}^{n \times n}$  are called congruent if there exists a nonsingular matrix  $T \in \mathbb{C}^{n \times n}$  such that  $A = TBT^H$ . Given a Hermitian matrix  $\text{Herm}$ , then  $\text{Inertia}(\text{Herm})$  is defined as the 3-tuple  $(\pi, \nu, \zeta)$  with  $\pi$  the number of positive eigenvalues,  $\nu$  the number of negative eigenvalues and  $\zeta$  the number of zero eigenvalues of  $\text{Herm}$ . (This definition makes sense since the eigenvalues of  $\text{Herm}$  are known to be real).*

Inertia and congruence are classical tools in linear algebra. For example, *Sylvester's law of inertia* states that  $\text{Inertia}(A) = \text{Inertia}(B)$  if and only if  $A$  and  $B$  are congruent. Moreover, inertia is *subadditive* in the sense that for all Hermitian matrices  $A$  and  $B$ , we have  $\pi(A + B) \leq \pi(A) + \pi(B)$  and  $\nu(A + B) \leq \nu(A) + \nu(B)$ : see for example [8, Lemma 2] for an easy proof of this property. (Note that by adding these two equations, we get the well-known property of subadditivity of *rank*).

Now we prove

**Theorem 8** *The following are equivalent:*

- (i)  $A = \text{Herm} + \text{Rk } r$ , i.e.  $A$  is Hermitian plus rank at most  $r$ ;

(ii)  $i(A - A^H) = \text{Rk } 2r$ , where  $i := \sqrt{-1}$  and where  $\text{Rk } 2r$  is a matrix of rank at most  $2r$  such that  $\text{Inertia}(\text{Rk } 2r) = (\pi, \nu, \zeta)$  with  $\max\{\pi, \nu\} \leq r$ .

PROOF.

a. First we prove the implication (i)  $\Rightarrow$  (ii). Thus let us assume that  $A = \text{Herm} + \text{Rk } r = \text{Herm} + \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^H$  for suitable column vectors  $\mathbf{u}_k, \mathbf{v}_k \in \mathbb{C}^n$ . The matrix  $\text{Herm}$  can be eliminated by considering

$$i(A - A^H) = \sum_{k=1}^r i(\mathbf{u}_k \mathbf{v}_k^H - \mathbf{v}_k \mathbf{u}_k^H). \quad (14)$$

We will first prove the theorem for rank upper bound  $r = 1$ . Thus we will prove that for any  $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ ,

$$\text{Inertia}(i(\mathbf{u}\mathbf{v}^H - \mathbf{v}\mathbf{u}^H)) = (\pi, \nu, \zeta) \quad \text{with } \max\{\pi, \nu\} \leq 1. \quad (15)$$

To prove this, let us write

$$i(\mathbf{u}\mathbf{v}^H - \mathbf{v}\mathbf{u}^H) = \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix} \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}^H. \quad (16)$$

If the columns  $\mathbf{u}, \mathbf{v}$  in (16) are linearly dependent, then  $i(\mathbf{u}\mathbf{v}^H - \mathbf{v}\mathbf{u}^H)$  has rank at most one, and hence (15) must obviously be true. If the columns  $\mathbf{u}, \mathbf{v}$  are independent, then from (16), Sylvester's law of inertia implies that  $\text{Inertia}(i(\mathbf{u}\mathbf{v}^H - \mathbf{v}\mathbf{u}^H)) = \text{Inertia}\left(\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}\right)$  (completed with  $n - 2$  zero eigenvalues). But the latter 2 by 2 matrix has determinant  $-1$ , and hence it has exactly one positive and one negative eigenvalue. This proves (15), i.e. the theorem has been proven now for rank upper bound  $r = 1$ .

In the general case  $r \geq 1$ , it follows from (14) that the matrix  $i(A - A^H)$  is the sum of all the  $i(\mathbf{u}_k \mathbf{v}_k^H - \mathbf{v}_k \mathbf{u}_k^H)$ ,  $k = 1, \dots, r$ . The theorem follows then by (15) and the subadditivity of inertia.

b. Now we prove the implication (ii)  $\Rightarrow$  (i). Thus let us assume that  $\text{Inertia}(i(A - A^H)) = (\pi, \nu, \zeta)$ , where  $\max\{\pi, \nu\} \leq r$ . By symmetry, we may suppose that  $\pi \geq \nu$ . We define the block diagonal matrix

$$D = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \oplus \dots \oplus \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \oplus 1 \oplus \dots \oplus 1, \quad (17)$$

where the first factor occurs precisely  $\nu$  times and the second factor  $\pi - \nu$  times. By construction,  $D$  has exactly the same inertia as  $i(A - A^H)$  (except for a loss of zero eigenvalues). By Sylvester's law of inertia, there exists a maximal column rank matrix  $T$  such that

$$i(A - A^H) = TDT^H. \quad (18)$$

Now we define vectors  $\mathbf{u}_k, \mathbf{v}_k$  by setting  $T =: [\mathbf{u}_1, \mathbf{v}_1; \dots; \mathbf{u}_\nu, \mathbf{v}_\nu; \mathbf{u}_{\nu+1}; \dots; \mathbf{u}_\pi]$  and  $\mathbf{v}_k := \frac{i}{2}\mathbf{u}_k$  for  $k = \nu + 1, \dots, \pi$ . It is easy to check that

$$\sum_{k=1}^{\pi} i(\mathbf{u}_k \mathbf{v}_k^H - \mathbf{v}_k \mathbf{u}_k^H) = TDT^H. \quad (19)$$

Together with (18), Equation (19) shows that the matrix  $A - \sum_{k=1}^{\pi} \mathbf{u}_k \mathbf{v}_k^H$  must be Hermitian. Since moreover  $\pi = \max\{\pi, \nu\} \leq r$ , this yields us a decomposition  $A = \text{Herm} + \text{Rk } r$ , hence finishing the proof.  $\square$

From this theorem, we can derive some further properties of Hermitian plus rank at most  $r$  matrices, including an alternative (and complete) proof of their preservation under inversion.

**Corollary 9** *Let  $r \in \mathbb{N}$  be a given number, then*

1. *the class  $\{A \in \mathbb{C}^{n \times n} \mid A = \text{Herm} + \text{Rk } r\}$  is topologically closed;*
2. *the class  $\{A \in \mathbb{C}^{n \times n} \mid A = \text{Herm} + \text{Rk } r\}$  is closed under matrix inversion;*
3. *in case  $A$  is real, the inertia condition in Theorem 8 can be dropped, i.e. from  $i(A - A^H) = \text{Rk } 2r$ , it automatically follows that  $\text{Inertia}(\text{Rk } 2r) = (\pi, \nu, \zeta)$  with  $\max\{\pi, \nu\} \leq r$ .*

PROOF.

1. Obviously, for a family of matrices  $A_\epsilon \in \mathbb{C}^{n \times n}$ ,  $\epsilon \in \mathbb{C} \setminus \{0\}$  with  $\lim_{\epsilon \rightarrow 0} A_\epsilon = A \in \mathbb{C}^{n \times n}$ , the property  $\text{Inertia}(i(A_\epsilon - A_\epsilon^H)) = (\pi, \nu, \zeta)$  with  $\max\{\pi, \nu\} \leq r$  can not be lost for the limiting matrix  $A = \lim_{\epsilon \rightarrow 0} A_\epsilon$ . We can then conclude by Theorem 8.
2. We use again Theorem 8. Thus let  $A$  be a matrix satisfying  $i(A - A^H) = \text{Rk } 2r$  and  $\text{Inertia}(\text{Rk } 2r) = (\pi, \nu, \zeta)$  with  $\max\{\pi, \nu\} \leq r$ . From Corollary 5 (or by direct verification), we obtain that

$$i(A^{-1} - A^{-H}) = -\widetilde{\text{Rk}} 2r, \quad (20)$$

where  $\widetilde{\text{Rk}} 2r := A^{-1}(\text{Rk } 2r)A^{-H}$ . Since  $\text{Rk } 2r$  and  $\widetilde{\text{Rk}} 2r$  are congruent, by Sylvester's law of inertia they have the same inertia. In particular, the property  $\max\{\pi, \nu\} \leq r$  must carry over to  $\widetilde{\text{Rk}} 2r$ , and hence to  $-\widetilde{\text{Rk}} 2r$  in (20).

3. It can be easily checked that for a *real* matrix  $A$ , the eigenvalues of the matrix  $i(A - A^H)$  always come in (real) pairs  $\lambda, -\lambda$ . Hence from  $i(A - A^H) = \text{Rk } 2r$  it automatically follows that  $\pi = \nu = \max\{\pi, \nu\} \leq r$ , leading to the desired simplification of Theorem 8.  $\square$

### 3 Inversion of rank structures

In this section we handle the inversion of rank structures. These rank structures are a generalization of the ones in [2]. We start with the following definition.

**Definition 10** We define a rank structure on  $\mathbb{C}^{n \times n}$  as a collection of so-called structure blocks  $\mathcal{R} = \{\mathcal{B}_k\}_k$ . Each structure block  $\mathcal{B}_k$  is characterized as a 4-tuple

$$\mathcal{B}_k = (i_k, j_k, r_k, \Lambda_k),$$

where  $i_k$  is the row index,  $j_k$  the column index,  $r_k$  the rank upper bound and  $\Lambda_k \in \mathbb{C}^{(j_k - i_k + 1) \times (j_k - i_k + 1)}$  is called the shift matrix of  $\mathcal{B}_k$ . We say a matrix  $A \in \mathbb{C}^{n \times n}$  to satisfy the rank structure if for each  $k$ , making a  $k$ -dependent partitioning

$$A =: \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix}_k, \quad (21)$$

where  $A_{2,2}$  is square and containing rows and columns  $i_k, \dots, j_k$ , then

$$\begin{bmatrix} A_{2,1} & A_{2,2} - \Lambda_k \\ A_{3,1} & A_{3,2} \end{bmatrix}_k = \text{Rk } r_k, \quad (22)$$

i.e. a matrix of rank at most  $r_k$ .

As a special case,  $\mathcal{B}_k$  is called a pure structure block if it has shift matrix  $\Lambda_k = 0$ . Pure structure blocks will be denoted as  $\mathcal{B}_{\text{pure},k}$ . Also if  $j_k - i_k + 1 \leq 0$ , then  $\mathcal{B}_k$  is considered to be a pure structure block, since it does not intersect the diagonal: see Figure 1. (Note that (22) is invalid for  $j_k - i_k + 1 < 0$ , since it would contain negative matrix dimensions. We should replace it here by  $[A_{3,1}]_k = \text{Rk } r_k$ , where  $A_{3,1}$  contains rows  $i_k, \dots, n$  and columns  $1, \dots, j_k$ ).

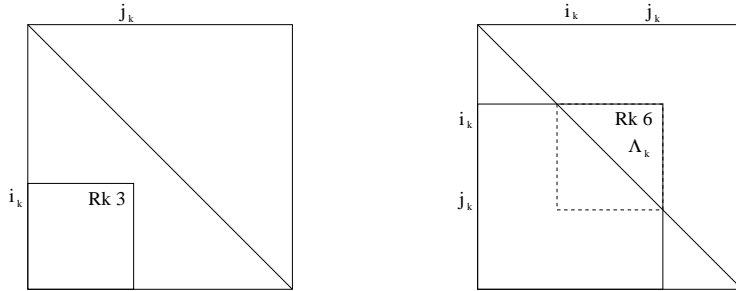


Figure 1: The structure block in the left figure has  $j_k - i_k + 1 \leq 0$  and is pure. The structure block in the right picture has the following meaning: after subtracting the shift matrix  $\Lambda_k \in \mathbb{C}^{4 \times 4}$  from the dashed square submatrix in the middle, the indicated bottom left submatrix must be of rank at most 6.

In case  $\Lambda_k = \lambda_k I$  for certain  $\lambda_k \in \mathbb{C}$ , Definition 10 leads to the structure blocks that were studied in [2]. We proved there that these structure blocks are preserved by the shifted QR-algorithm. For the present paper, we do not need to make this restriction on the shift matrices  $\Lambda_k$ .

Let us illustrate Definition 10. We can use it to describe Hessenberg matrices:  $\mathcal{B}_{\text{pure},k} = (k+2, k, 0, 0)$ ,  $k = 1, \dots, n-2$ ; upper triangular matrices; lower semiseparable matrices:  $\mathcal{B}_{\text{pure},k} = (k, k, 1, 0)$ ,  $k = 1, \dots, n$ ; lower semiseparable plus diagonal matrices:  $\mathcal{B}_k = (k, k, 1, \lambda_k)$ ,  $k = 1, \dots, n$ . Note that for this last example, the shift matrices  $\Lambda_k =: \lambda_k \in \mathbb{C}$  are scalar. Also for upper triangular matrices, we could absorb the diagonal elements  $\lambda_k$  into the structure, if we would want to. Of course, there are also more general or ‘chaotic’ rank structures than the ones that we just mentioned.

Concerning matrix inversion, let us first indicate why we should expect a positive result. Some well-known examples are the following: upper triangular structure is preserved under matrix inversion; lower semiseparable matrices and Hessenberg matrices are each others inverses; the *induced pure structure* of lower semiseparable plus diagonal matrices, i.e.  $\mathcal{B}_{\text{pure},k} = (k+1, k, 1, 0)$ ,  $k = 1, \dots, n-1$ , is preserved under matrix inversion.

Having a second look at these examples, we can note that for upper triangular matrices, the diagonal elements  $\lambda_k$  are precisely transformed into  $1/\lambda_k$  under matrix inversion. The same holds for the shift elements  $\lambda_k$  of a lower semiseparable plus diagonal matrix, as we will show (see also [13, Theorem 4.1]). Since it is sufficient to consider the behaviour of a single structure block  $\mathcal{B}_k$ , from now on we will put  $\mathcal{B}_k =: \mathcal{B}$  and drop the index  $k$ .

**Theorem 11** *Let  $A \in \mathbb{C}^{n \times n}$  be a nonsingular matrix satisfying the structure block  $\mathcal{B} = (i, j, r, \Lambda)$ , with  $\Lambda$  nonsingular. (We assume here implicitly that  $j - i + 1 \geq 0$ , i.e. that  $\mathcal{B}$  is intersecting or lying just below the main diagonal of  $A$ .) Then the inverse matrix will satisfy the structure block  $\mathcal{B}^{-1} := (i, j, r, \Lambda^{-1})$ : see Figure 2.*

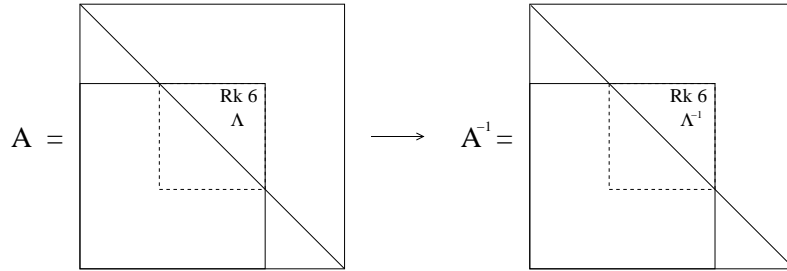


Figure 2: Under matrix inversion, the shift matrix  $\Lambda$  of the Rk 6 structure block will transform into the new shift matrix  $\Lambda^{-1}$ . The rank number itself is preserved.

PROOF. Making a partition of  $A$  as in (21), it follows that  $A - B = \text{Rk } r$  where

$$B = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ 0 & \Lambda & A_{2,3} \\ 0 & 0 & A_{3,3} \end{bmatrix}.$$

Then by (13), or by direct verification, it follows that  $A^{-1} - B^{-1} = \widetilde{\text{Rk}} r$ . By the form of  $B$ , this implies that  $A^{-1}$  will satisfy the inverse structure block  $\mathcal{B}^{-1}$ , and the theorem is proved.

Note that we assumed here implicitly that the matrix  $B$  is nonsingular, i.e. that the square blocks  $A_{1,1}$  and  $A_{3,3}$  are nonsingular. The case where this nonsingularity is not satisfied can be handled by a continuity argument. (We add an infinitesimally small correction to the  $A_{1,1}$  and  $A_{3,3}$  blocks, which will not influence the given structure block).  $\square$

**Remark 12** 1. Note that the above theorem has been stated in terms of rank upper bounds  $r$ . We proved in fact that the structure block  $\mathcal{B}^{-1}$  can have at most the same rank as the original structure block  $\mathcal{B}$ , i.e.  $r_{\mathcal{B}^{-1}} \leq r_{\mathcal{B}}$ . By applying the same argument to  $\mathcal{B}^{-1}$ , we obtain also the inequality  $r_{\mathcal{B}^{-1}} \geq r_{\mathcal{B}}$ , and it follows that the ranks must be exactly the same. A similar remark could have been made also for the theorems in Section 2.

2. As stated in the formulation of the theorem, the above proof works also for the case where  $\Lambda$  is lying just below the main diagonal of  $A$ , i.e. for the case where  $j - i + 1 = 0$ . One should then consider  $\Lambda$  and  $\Lambda^{-1}$  as empty matrices. Structure blocks with  $j - i + 1 = 0$  have already received a lot of attention in the literature, from the point of view of efficient inversion algorithms: see for example [3, 4].

## 4 Inversion of rank structures: some extensions

In this section we will extend Theorem 11 in several directions.

### 4.1 Singular shift matrices

In the statement of Theorem 11, it was assumed that the shift matrix  $\Lambda$  is nonsingular. We will now remove this condition.

The first step is to construct unitary matrices  $U$  and  $V$  such that

$$U^H \Lambda V = \Lambda_{\text{ns}} \oplus 0, \tag{23}$$

with  $\Lambda_{\text{ns}}$  nonsingular. This matrix decomposition can be considered as an ‘incomplete singular value decomposition’. The word *incomplete* means that we are only concerned with transforming the dependent rows and columns of  $\Lambda$  into zeros, which is a relatively easy operation.

Now by (23), we have that  $(I \oplus V^H \oplus I)A^{-1}(I \oplus U \oplus I)$  satisfies the structure block

$$\mathcal{B}^{-1}, \quad \text{where } \mathcal{B} = (i, j, r, \Lambda_{\text{ns}} \oplus 0).$$

So the structure of  $A^{-1}$  is known as soon as we know the structure block  $\mathcal{B}^{-1}$ . Hence from now on, we will suppose that  $\Lambda = \Lambda_{\text{ns}} \oplus 0$  with  $\Lambda_{\text{ns}}$  nonsingular.

By continuity reasons, we could then expect that  $\Lambda^{-1} = \Lambda_{\text{ns}}^{-1} \oplus \infty I$ .

**Definition 13** We define a structure block  $\mathcal{B} = (i, j, r, \Lambda)$  on  $\mathbb{C}^{n \times n}$ , where  $\Lambda = \Lambda_{\text{fin}} \oplus \infty I$  with  $\Lambda_{\text{fin}} \in \mathbb{C}^{m \times m}$  for certain  $m$ , as follows. We identify  $\mathcal{B}$  with the ‘structure block’ obtained by dropping all rows and columns involving  $\infty$ , and with the rank upper bound  $r$  decreased by the number of these dropped rows: see Figure 3.

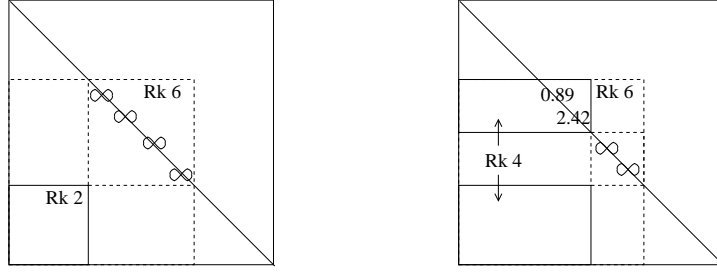


Figure 3: Given the Rk 6 structure block  $\mathcal{B}$  in the left picture. We have here  $\Lambda = \infty I_4$ , and  $\mathcal{B}$  can be identified with the smaller Rk 2 structure block. Next, consider the Rk 6 structure block  $\mathcal{B}$  in the right picture. We have here  $\Lambda = \text{diag}(0.89, 2.42, \infty, \infty)$ , and  $\mathcal{B}$  can be identified with the smaller Rk 4 structure block, consisting of two pieces. Note that the shift submatrix  $\text{diag}(0.89, 2.42)$  is inherited.

Formally, we will use  $\mathbb{C} \cup \{\infty\}$  to denote the one-point-compactification of  $\mathbb{C}$ . Thus by definition, we have  $x_\epsilon \rightarrow \infty$  if and only if the moduli  $|x_\epsilon| \rightarrow \infty$ .

Of course we have to motivate Definition 13. Thus let us show that it is indeed the correct, i.e. continuous definition for shift matrix  $\infty I$ .

**Theorem 14** Given fixed integers  $i, j, r, m_1$  and  $m_2$  such that  $j - i + 1 = m_1 + m_2$ . Consider a family of matrices  $A_\epsilon, \epsilon \in \mathbb{C} \setminus \{0\}$  satisfying the conditions

F1 each  $A_\epsilon$  satisfies a structure block  $\mathcal{B}_\epsilon = (i, j, r, \Lambda_\epsilon)$ , where  $\Lambda_\epsilon = \Lambda_{\epsilon,1} \oplus \Lambda_{\epsilon,2}$  with  $\Lambda_{\epsilon,k} \in \mathbb{C}^{m_k \times m_k}, k = 1, 2$ ;

F2a  $\lim_{\epsilon \rightarrow 0} \Lambda_{\epsilon,1} =: \Lambda_{\text{fin}} \in \mathbb{C}^{m_1 \times m_1}$ , and  $\lim_{\epsilon \rightarrow 0} \Lambda_{\epsilon,2} = \infty I$ ;

F2b  $\lim_{\epsilon \rightarrow 0} A_\epsilon =: A \in \mathbb{C}^{n \times n}$ .

Then

1. (Continuity:)  $A$  satisfies the limiting structure block  $\mathcal{B} = (i, j, r, \Lambda)$ , with  $\Lambda = \Lambda_{\text{fin}} \oplus \infty I$ ;
2. (Approximation by finite shift matrices:) conversely, every matrix  $A$  satisfying  $\mathcal{B} = (i, j, r, \Lambda)$  can be realized as the limit of a family  $A_\epsilon$  satisfying the above conditions F1 and F2.

PROOF. For simplicity, we restrict the proof to the ‘pure’ case where

Assumption:  $\Lambda = \infty I$ , i.e.  $m_1 = 0$  and  $m_2 = j - i + 1$ .

Moreover, we define index sets  $R = \{1, \dots, i - 1\}$ ,  $S = \{i, \dots, j\}$ ,  $T = \{j + 1, \dots, n\}$ : this is the situation in the left picture of Figure 4.

1. First, suppose that the family  $A_\epsilon$  satisfies conditions F1 and F2. Define a family  $\tilde{A}_\epsilon$  from  $A_\epsilon$  by setting

$$\tilde{A}_\epsilon(S, S) := A_\epsilon(S, S) - \Lambda_\epsilon.$$

Thus by F2, for  $\epsilon \rightarrow 0$  the diagonal entries of  $\tilde{A}_\epsilon(S, S)$  converge to  $\infty$  while the off-diagonal entries converge to finite values. Hence it is easy to see that  $\tilde{A}_\epsilon(S, S)^{-1}$  exists for all  $\epsilon$  sufficiently small and satisfies

$$\lim_{\epsilon \rightarrow 0} \tilde{A}_\epsilon(S, S)^{-1} = 0. \quad (24)$$

Moreover, the nonsingularity of  $\tilde{A}_\epsilon(S, S)$  allows us to use it as pivot block for a Gaussian elimination process. Then the Schur complement formula yields (we suppress the tilde whenever allowed)

$$\begin{aligned} \text{Rank } \tilde{A}_\epsilon(S \cup T, R \cup S) = \\ |S| + \text{Rank} \left( A_\epsilon(T, R) - A_\epsilon(T, S) \tilde{A}_\epsilon(S, S)^{-1} A_\epsilon(S, R) \right). \end{aligned} \quad (25)$$

But by F1, the left hand side of (25) can not exceed  $r$ , i.e.

$$\begin{aligned} r - |S| &\geq \limsup_{\epsilon \rightarrow 0} \text{Rank} \left( A_\epsilon(T, R) - A_\epsilon(T, S) \tilde{A}_\epsilon(S, S)^{-1} A_\epsilon(S, R) \right) \\ &\geq \text{Rank} \lim_{\epsilon \rightarrow 0} \left( A_\epsilon(T, R) - A_\epsilon(T, S) \tilde{A}_\epsilon(S, S)^{-1} A_\epsilon(S, R) \right) \\ &= \text{Rank } A(T, R), \end{aligned}$$

where the last transition follows from (24). This shows that the limiting matrix  $A$  satisfies the structure block  $\mathcal{B} = (i, j, r, \infty I)$ .

2. Conversely, suppose that  $A$  is a matrix satisfying  $\mathcal{B} = (i, j, r, \infty I)$ . Define a matrix  $\tilde{A}$  from  $A$  by setting  $\tilde{A}(S, S) = A(S, S) - \frac{1}{\epsilon} I$  (considering  $\epsilon$  as a symbol). Define a family  $A_\epsilon$ ,  $\epsilon \in \mathbb{C} \setminus \{0\}$  from  $A$  by

$$A_\epsilon(T, R) := A(T, R) + A(T, S) \tilde{A}(S, S)^{-1} A(S, R), \quad (26)$$

as in the right picture of Figure 4. Finally, define a family  $\tilde{A}_\epsilon$  from  $A_\epsilon$  by setting  $\tilde{A}_\epsilon(S, S) := A_\epsilon(S, S) - \frac{1}{\epsilon}I$ . Then for  $\epsilon \rightarrow 0$  the diagonal entries of  $\tilde{A}_\epsilon(S, S)$  converge to  $\infty$  while the off-diagonal entries are just constant values. Hence the derivation of (24) and (25) remains valid. But now by the above definition of  $A_\epsilon$ , the Schur complement in the right hand side of (25) simplifies to  $A_\epsilon(T, R) - A(T, S)\tilde{A}_\epsilon(S, S)^{-1}A(S, R) = A(T, R)$ . Hence (25) collapses to

$$\begin{aligned} \text{Rank}(\tilde{A}_\epsilon(S \cup T, R \cup S)) &= |S| + \text{Rank } A(T, R) \\ &\leq |S| + (r - |S|) \\ &= r, \end{aligned}$$

where the second transition follows from our assumptions on  $A$ . Thus we established that the family  $A_\epsilon$  satisfies F1 and F2a, with  $\Lambda_\epsilon := \frac{1}{\epsilon}I$ . Finally, the fact that  $A = \lim_{\epsilon \rightarrow 0} A_\epsilon$  follows from the above definition of  $A_\epsilon$ , together with (24). □

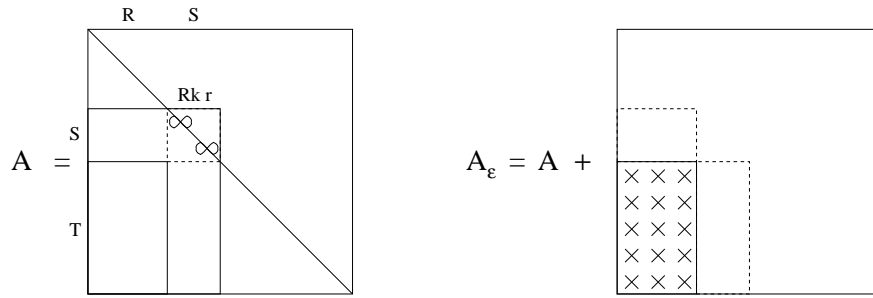


Figure 4: Let  $A$  satisfy a structure block  $\mathcal{B}$  with shift matrix  $\Lambda = \infty I$ . We can approximate  $\Lambda$  by finite shift matrices by adding an infinitesimally small correction term to  $A(T, R)$ . Note that, moreover, the row space of this correction term is contained in the row space of  $A(S, R)$  and the column space is contained in the column space of  $A(T, S)$  (see (26)).

**Remark 15** *The above proof was made under the assumption  $\Lambda = \infty I$ . For the general case, we make an additional partition  $S = S_1 \cup S_2$  with  $S_1$  the indices corresponding to  $\Lambda_{\text{fin}}$  and  $S_2$  the indices corresponding to  $\infty I$ . Then to prove Theorem 14.2, say, we proceed as follows: (i) we observe that the problem reduces to the matrix  $A_{\text{pure}} := A - 0 \oplus \Lambda_{\text{fin}} \oplus 0$ ; (ii) we realize  $A_{\text{pure}} = \lim_{\epsilon \rightarrow 0} A_{\text{pure}, \epsilon}$  by just applying the result which was already proved for Theorem 14.2, but now with index sets  $R := R \cup S_1$ ,  $S := S_2$  and  $T := T \cup S_1$ .*

Theorem 14 allows a restatement in terms of topological closure. Let us illustrate this for a particular example. Let  $\mathcal{M}$  be the set of ‘partially lower semiseparable’ matrices satisfying  $\mathcal{R} = \{\mathcal{B}_k\}_{k \in K}$ , where  $K$  is a certain index set, and with each structure block  $\mathcal{B}_k = (k, k, 1, \lambda_k)$ , where the shift element  $\lambda_k$  is allowed to take any value in  $\mathbb{C}$ . We claim that the topological closure  $\bar{\mathcal{M}}$  can be obtained in exactly the same way, but now allowing  $\lambda_k \in \mathbb{C} \cup \{\infty\}$ . This means that also *Hessenberg structure blocks* should be allowed in the structure: see Figure 5.

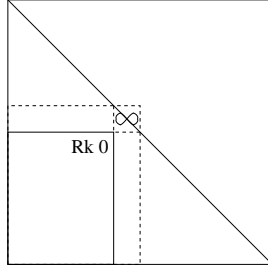


Figure 5: Each structure block  $\mathcal{B}_k = (k, k, 1, \infty)$ ,  $k = 1, \dots, n$  can be seen as a Hessenberg structure block. In particular, each Hessenberg matrix can be approximated by a sequence of lower semiseparable plus diagonal matrices.

Indeed: let us show that to each  $A \in \bar{\mathcal{M}}$  there corresponds a set of shift elements  $\lambda_k \in \mathbb{C} \cup \{\infty\}$ . Thus let  $A_\epsilon \in \mathcal{M}$  be a converging sequence of matrices, with corresponding shift elements  $\lambda_{\epsilon, k}$ . By the fact that  $\mathbb{C} \cup \{\infty\}$  is compact, for each  $k$  there must exist a convergent subsequence  $\lambda_{k, \epsilon} \rightarrow \lambda_k \in \mathbb{C} \cup \{\infty\}$ . The result follows then by Theorem 14.1 (‘Continuity’).

Conversely, let us show that if  $A$  satisfies a set of shift elements  $\lambda_k \in \mathbb{C} \cup \{\infty\}$ , we have  $A \in \bar{\mathcal{M}}$ . But this follows just by Theorem 14.2 (‘Approximation by finite shift matrices’). The only subtlety is that the infinitesimal correction term, constructed in Theorem 14.2 for one structure block  $\mathcal{B}_k$  with  $\lambda_k = \infty$ , should not destroy the other structure blocks  $\mathcal{B}_{\tilde{k}}$ ,  $\tilde{k} \neq k$ . The fact that this is satisfied, follows since the row and column spaces of these correction terms are well-behaved; see the explanation in Figure 4.

We will not go further into this.

Now we come back to matrix inversion. We can use the result of Theorem 14 to remove the nonsingularity condition for  $\Lambda$  from the statement of Theorem 11.

**Corollary 16** *Let  $A \in \mathbb{C}^{n \times n}$  be a nonsingular matrix satisfying the structure block  $\mathcal{B} = (i, j, r, \Lambda)$ , where  $\Lambda = \Lambda_{\text{ns}} \oplus 0 \oplus \infty I$ , with  $\Lambda_{\text{ns}}$  nonsingular. Then the inverse matrix  $A^{-1}$  will satisfy the structure block  $\mathcal{B}^{-1} := (i, j, r, \Lambda^{-1})$ , with  $\Lambda^{-1} := \Lambda_{\text{ns}}^{-1} \oplus \infty I \oplus 0$  (hence using the rules  $\frac{1}{0} = \infty$  and  $\frac{1}{\infty} = 0$ ).*

PROOF. The case where  $\Lambda = \Lambda_{\text{ns}}$  has been proved in Theorem 11. The general case follows by continuity. Indeed: by Theorem 14.2, we can construct

a family  $A_\epsilon$ ,  $\epsilon \in \mathbb{C} \setminus \{0\}$  with  $\lim_{\epsilon \rightarrow 0} A_\epsilon = A$ , and such that the block  $\infty I$  is approximated by finite shift matrices  $\frac{1}{\epsilon}I$ . Next we add the correction term  $\epsilon I$  to the zero block. Thus each  $A_\epsilon$  satisfies the structure block  $\mathcal{B} = (i, j, r, \Lambda_\epsilon)$ , with

$$\Lambda_\epsilon = \Lambda_{\text{ns}} \oplus \epsilon I \oplus \frac{1}{\epsilon}I.$$

Hence

$$\Lambda_\epsilon^{-1} = \Lambda_{\text{ns}}^{-1} \oplus \frac{1}{\epsilon}I \oplus \epsilon I.$$

The theorem follows then by applying Theorem 14.1 ('Continuity').  $\square$

An interesting special case of Corollary 16 is when  $\Lambda = 0$  or  $\Lambda = \infty I$ . Then the theorem can be interpreted in terms of *pure* structure blocks: see figure 6.

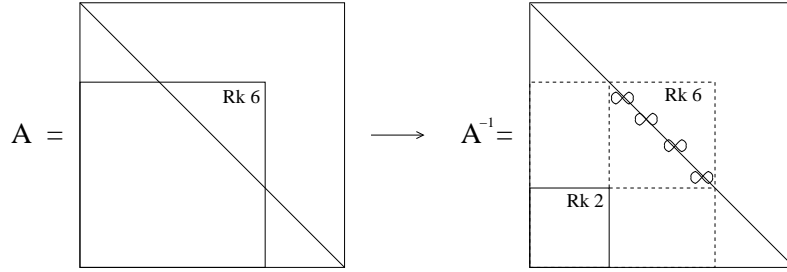


Figure 6: Given the matrix  $A$  in the left picture, satisfying the pure structure block  $\mathcal{B}_{\text{pure}} = (i, j, 6, 0)$ . Then the inverse matrix  $A^{-1}$  satisfies the pure structure block  $\mathcal{B}_{\text{pure}}^{-1} = (i, j, 6, \infty I)$  in the right picture.

In particular, we see that lower semiseparable and Hessenberg matrices are each others inverses. This follows since these matrices are both lower semiseparable plus diagonal, with structure blocks  $\mathcal{B}_k$  having shift elements  $\lambda_k = 0$  for lower semiseparable and  $\lambda_k = \infty$  for Hessenberg matrices.

For another example, suppose that  $A$  is a diagonal plus rank one matrix with diagonal correction  $\Lambda$ . Then there are two possibilities. If  $\Lambda$  is nonsingular, the inverse matrix  $A^{-1}$  is again diagonal plus rank one, with diagonal correction  $\Lambda^{-1}$ : this follows immediately from Theorem 11. Suppose now that  $\Lambda$  is singular, say  $\lambda_n = 0$ . Then the corresponding diagonal element of  $\Lambda^{-1}$  is  $\frac{1}{\lambda_n} = \infty$ , and hence  $A^{-1}$  will be an *arrowhead matrix*: see Figure 7.

Next let us assume that a certain square submatrix  $\Lambda$  of a matrix  $A$  is known. For example, let us suppose that

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & \Lambda \end{bmatrix},$$

with  $\Lambda \in \mathbb{C}^{k \times k}$ , for certain  $k$ . Then we claim that

$$A^{-1} - \begin{bmatrix} 0 & 0 \\ 0 & \Lambda^{-1} \end{bmatrix} = \text{Rk } n - k. \quad (27)$$

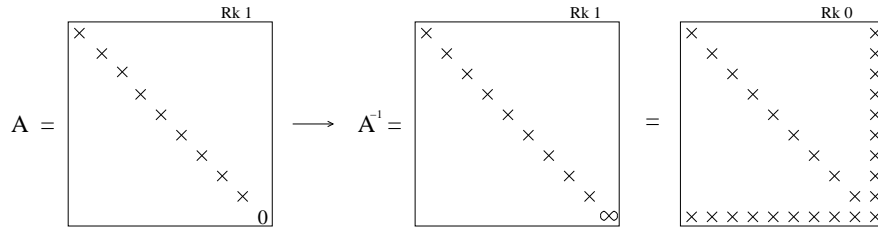


Figure 7: Given the matrix  $A$  in the left picture, which is diagonal plus rank one with last diagonal element equal to zero. Then the inverse matrix is arrowhead, and conversely.

Indeed: this can be seen by the argument in Figure 8.

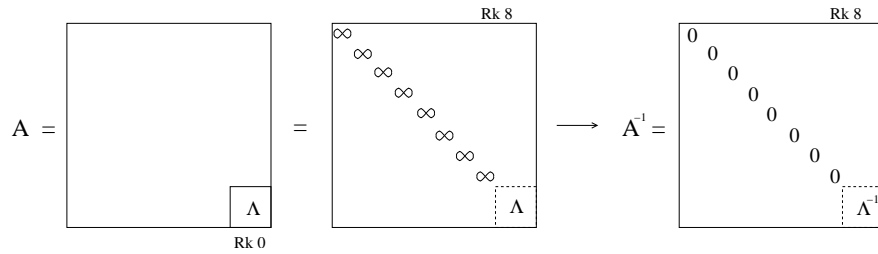


Figure 8: Given the matrix  $A \in \mathbb{C}^{10 \times 10}$  in the leftmost picture, with given  $(2, 2)$  square submatrix  $\Lambda$ . By suitably adding shift elements  $\infty$ , we can express these data in terms of a structure block  $\mathcal{B}$ . Hence the inverse matrix will satisfy the inverse structure block  $\mathcal{B}^{-1}$  of the rightmost picture, and this is precisely what (27) states.

We should mention here that also other proofs for (27) are possible. One possibility is to use an argument based on Schur complements. Another possibility is to translate the data to the decoupled displacement equations of Section 2 by writing  $PAP^T - \Lambda = \text{Rk } 0$ , where  $P$  is the projection matrix onto the last  $k$  columns. Hence Theorem 2 implies  $A^{-1} - P^T \Lambda P = \text{Rk } n - k$ , which is precisely (27).

Nevertheless, the argument based on shift elements  $\infty$  is maybe the ‘best’ in the sense that it can be used to handle also the case where  $\Lambda$  is rectangular; see further.

As a final example, let us assume that  $A = \text{Uni} + \text{Rk } r$  is a unitary plus low rank matrix. Then it is straightforward to check that  $A^{-1} - A^H = \text{Rk } 2r$  is a matrix of rank at most  $2r$ . Hence by Corollary 16, the structure blocks of  $A$  always come in *pairs*, as illustrated in Figure 9.

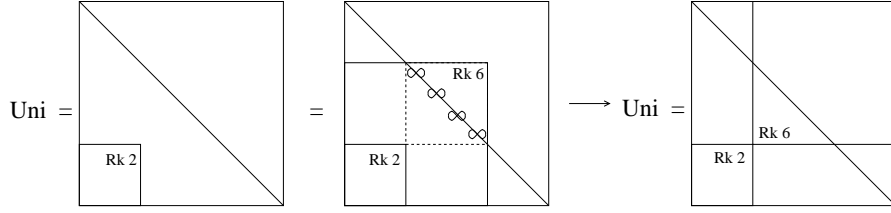


Figure 9: For a unitary matrix  $\text{Uni}$ , the structure blocks always come in pairs. The picture illustrates this for a given *pure*  $\text{Rk } 2$  structure block, but the result holds also for non-pure structure blocks.

## 4.2 Incorporating permutations

In this subsection we derive a generalization of Theorem 11 by absorbing permutation matrices into the structure.

We recall that for a permutation map  $P$ , the corresponding *permutation matrix*  $P$  (with same notation) is defined by its  $j$ th column being the basis vector  $\mathbf{e}_{P(j)}$ . Moreover, such matrices are unitary since  $P^{-1} = P^T = P^H$ .

First we may consider permutations acting on the shift matrix  $\Lambda$ . Suppose  $P$  and  $\tilde{P}$  are permutations such that

$$P^{-1}\Lambda\tilde{P} = \Lambda_{\text{ns}} \oplus 0 \oplus \infty I.$$

Then by construction,  $(I \oplus \tilde{P}^{-1} \oplus I)A^{-1}(I \oplus P \oplus I)$  satisfies the structure block

$$\mathcal{B} = (i, j, r, \Lambda_{\text{ns}}^{-1} \oplus \infty I \oplus 0).$$

This allows us to choose arbitrarily the row and column indices of  $\Lambda$  where  $\infty$  is standing, a fact we already used implicitly in Figure 8.

Now we consider permutations acting on the *complete* matrix  $A$ .

**Definition 17** We define a generally positioned rank structure on  $\mathbb{C}^{n \times n}$  as a collection of so-called generally positioned structure blocks  $\mathcal{R}_{\text{genPo}} = \{\mathcal{B}_{\text{genPo},k}\}_k$ . Each  $\mathcal{B}_{\text{genPo}}$  (we drop the index  $k$ ) is characterized as a six-tuple

$$\mathcal{B}_{\text{genPo}} = (S, T, \tilde{R}, \tilde{S}, r, \Lambda), \quad (28)$$

where  $S, T, \tilde{R}$  and  $\tilde{S}$  are the index sets,  $r$  the rank upper bound and  $\Lambda \in \mathbb{C}^{|\tilde{S}| \times |S|}$  is called the shift matrix of  $\mathcal{B}_{\text{genPo}}$ . Here we impose that

*Condition:*  $S$  and  $T$  are disjoint,  $\tilde{R}$  and  $\tilde{S}$  are disjoint,  $|S| = |\tilde{S}|$  and  $|\tilde{R}| + |S| + |T| = n$ .

We say a matrix  $A \in \mathbb{C}^{n \times n}$  to satisfy  $\mathcal{B}_{\text{genPo}}$  if

$$\tilde{A}(S \cup T, \tilde{R} \cup \tilde{S}) = \text{Rk } r, \quad (29)$$

where  $\tilde{A}$  has been defined from  $A$  by  $\tilde{A}(S, \tilde{S}) = A(S, \tilde{S}) - \Lambda$ .

Let us show how the structure blocks of the above definition can be reduced to the structure blocks of Section 3.

By the condition occurring in Definition 17, there can be found index sets  $R$  and  $\tilde{T}$  such that  $R \cup S \cup T = \tilde{R} \cup \tilde{S} \cup \tilde{T}$  are both partitions of the set  $N := \{1, 2, \dots, n\}$ . Furthermore, we must have  $|R| = |\tilde{R}|$  and  $|T| = |\tilde{T}|$ .

Now consider the permutation

$$P : \begin{cases} \{1, \dots, |R|\} & \mapsto R \\ \{|R| + 1, \dots, n - |T|\} & \mapsto S \\ \{n - |T| + 1, \dots, n\} & \mapsto T. \end{cases}$$

Similarly, we define a permutation  $\tilde{P}$ . Then by construction, a matrix  $A$  will satisfy the generally positioned structure block  $\mathcal{B}_{\text{genPo}}$  of (28) if and only if the matrix  $P^{-1}A\tilde{P}$  satisfies the structure block  $\mathcal{B} = (i, j, r, \Lambda)$  in the sense of Section 3, with  $i := |R| + 1$  and  $j := |R| + |S|$ : see Figure 10.

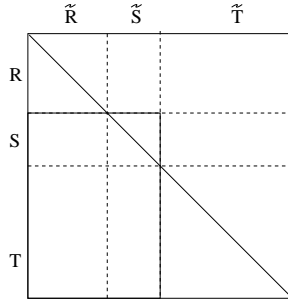


Figure 10: By using a transition  $A \mapsto P^{-1}A\tilde{P}$ , each structure block  $\mathcal{B}_{\text{genPo}}$  can be reduced to a usual structure block. The figure shows the position of the sets  $R, S, T$ ;  $\tilde{R}, \tilde{S}$  and  $\tilde{T}$  after permutation.

Concerning matrix inversion, we can use this correspondence  $A \mapsto P^{-1}A\tilde{P}$  to extend the inversion theorem of the previous subsection to the case of generally positioned structure blocks.

**Theorem 18** *Let  $\Lambda = \Lambda_{\text{ns}} \oplus 0 \oplus \infty I$ , and let  $A \in \mathbb{C}^{n \times n}$  be nonsingular and satisfying a structure block*

$$\mathcal{B}_{\text{genPo}} = (S, T, \tilde{R}, \tilde{S}, r, \Lambda). \quad (30)$$

*Then  $A^{-1}$  will satisfy the structure block*

$$\mathcal{B}_{\text{genPo}}^{-1} := (\tilde{S}, \tilde{T}, R, S, r, \Lambda^{-1}), \quad (31)$$

*see Figure 11.*

PROOF. We already remarked that the correspondence  $A \mapsto B := P^{-1}A\tilde{P}$  transforms the generally positioned structure block  $\mathcal{B}_{\text{genPo}}$  in (30) into the usual

structure block  $\mathcal{B} = (i, j, r, \Lambda)$  with  $i := |R| + 1$  and  $j := |R| + |S|$ . By Corollary 16, the inverse matrix  $B^{-1} = \tilde{P}^{-1}A^{-1}P$  will then satisfy the usual structure block  $\mathcal{B}^{-1} = (i, j, r, \Lambda^{-1})$ . Using again the definition of  $P$  and  $\tilde{P}$ , we see that  $A^{-1} = \tilde{P}B^{-1}P^{-1}$  satisfies precisely the inverse structure block  $\mathcal{B}_{\text{genPo}}^{-1}$  of (31), hence proving the theorem.  $\square$

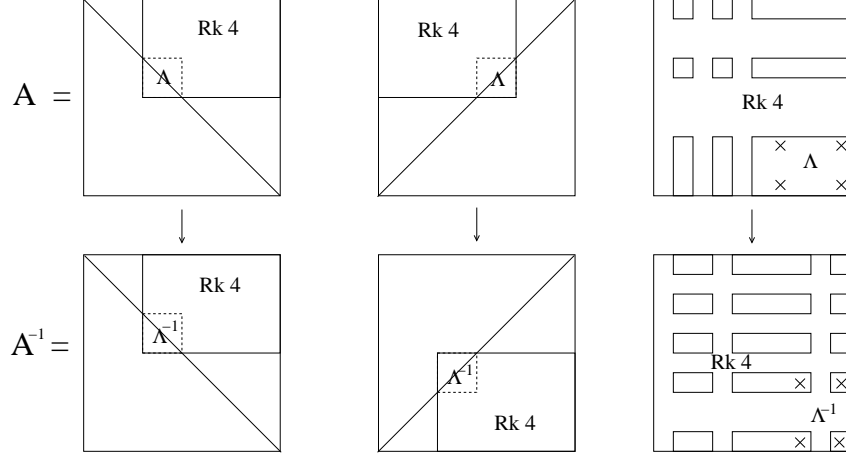


Figure 11: Given the structure blocks  $\mathcal{B}_{\text{genPo}}$  in the top pictures, with nonsingular shift matrix  $\Lambda \in \mathbb{C}^{2 \times 2}$ , the figure shows the corresponding structure blocks  $\mathcal{B}_{\text{genPo}}^{-1}$ . Note that the left and middle structure blocks are more or less preserved by matrix inversion, at least in shape, while the rightmost picture indicates that this is not always true in the general case.

### 4.3 Rectangular shift matrices

In this subsection we handle the case of arbitrarily shaped structure blocks, with possibly non-square shift matrix. More precisely, we will handle structure blocks of the form  $\mathcal{B}_{\text{genPo}} = (I, J, r, \Lambda)$  where  $\Lambda \in \mathbb{C}^{|I| \times |J|}$ . We say a matrix  $A$  to satisfy such a structure block if  $A(I, J) - \Lambda = \text{Rk } r$ .

The first step is to construct unitary matrices  $U$  and  $V$  (possibly of different sizes) such that

$$U^H \Lambda V = \begin{bmatrix} \Lambda_{\text{ns}} & 0 \\ 0 & 0_R \end{bmatrix}, \quad (32)$$

where  $\Lambda_{\text{ns}}$  is nonsingular and  $0_R$  possibly rectangular. This means that we transform the dependent rows and columns of  $\Lambda$  into zeros.

Now by (32), we have that  $(I \oplus V^H \oplus I)A^{-1}(I \oplus U \oplus I)$  satisfies the structure block

$$\mathcal{B}_{\text{genPo}}^{-1}, \quad \text{where } \mathcal{B}_{\text{genPo}} = (I, J, r, \begin{bmatrix} \Lambda_{\text{ns}} & 0 \\ 0 & 0_R \end{bmatrix}).$$

So the structure of  $A^{-1}$  is known as soon as we know the structure block  $\mathcal{B}_{\text{genPo}}^{-1}$ .

Thus we will assume now that  $\Lambda = \begin{bmatrix} \Lambda_{\text{ns}} & 0 \\ 0 & 0_R \end{bmatrix}$ . Then define index sets  $S$  and  $\tilde{S}$  which contain the rows and columns of  $\Lambda_{\text{ns}}$ , respectively, let  $T := I \setminus S$  and let  $\tilde{R} := J \setminus \tilde{S}$ . By construction,  $A$  satisfies the structure block  $\mathcal{B}_{\text{genPo}} = (S, T, \tilde{R}, \tilde{S}, r, \Lambda_{\text{ns}})$  and so the structure of  $A^{-1}$  can be obtained immediately by Theorem 18.

The only problem that could occur in this way is that the condition on the size restrictions of  $S, T$ , etc. occurring in Definition 17 is not satisfied yet. This can occur if

$$|\tilde{R}| + |S| + |T| \neq n. \quad (33)$$

But then we proceed as follows: if  $|\tilde{R}| + |S| + |T| > n$ , we just remove some indices from  $T, \tilde{R}$  and add them to  $S, \tilde{S}$  until the equality in (33) is restored. The involved indices can be considered as indices on which  $\Lambda$  ‘acts as the zero matrix’.

If  $|\tilde{R}| + |S| + |T| < n$ , we just leave  $T$  and  $\tilde{R}$  invariant and extend  $S, \tilde{S}$  by an extra index (which was not present yet in any of these index sets). Now we stipulate that  $\Lambda$  behaves as  $\infty$  on the added row and column, and update  $r := r + 1$ . This process can then be iterated until the equality in (33) is restored: see Figure 12.

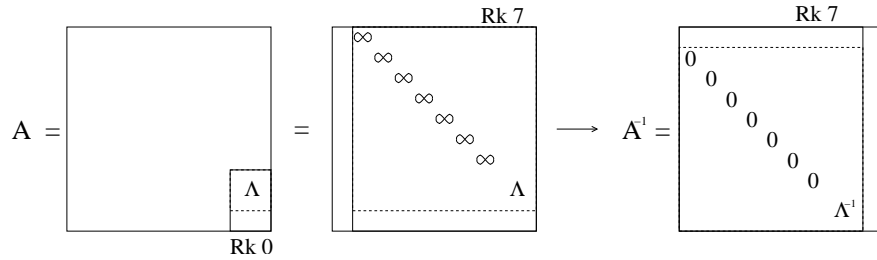


Figure 12: Given the matrix  $A \in \mathbb{C}^{10 \times 10}$  with specified 3 by 2 bottom right submatrix. We suppose here that the transformations  $U$  and  $V$  of (32) have already been applied, and that the shift matrix  $\Lambda$  in the figure is square and nonsingular. We can then add shift elements  $\infty$  until we have a structure block  $\mathcal{B}_{\text{genPo}}$  satisfying the conditions in Definition 17.  $\mathcal{B}_{\text{genPo}}^{-1}$  can then be obtained by Theorem 18.

We close with a theorem due to Fiedler [5]. We refer to [12, 14] and the references therein for more information about this theorem. First we give a definition.

**Definition 19** *The right nullity or shortly nullity  $\text{Null } A$  of a rectangular matrix  $A \in \mathbb{C}^{m \times n}$  is defined as the dimension of the right null space of  $A$ . The left nullity of  $A$  is defined in a similar way, but we do not reserve a special notation*

for it. The rank defect of  $A$  is defined as the minimum of the left and right nullity.

**Corollary 20** (See [5]:) *Let  $A \in \mathbb{C}^{n \times n}$  be nonsingular. Then for any two index sets  $I$  and  $J$ , we have that*

$$\text{Null } A^{-1}(I, J) = \text{Null } A(N \setminus J, N \setminus I), \quad (34)$$

where  $N := \{1, 2, \dots, n\}$ . Moreover, (34) remains valid if we replace everywhere Null by the left nullity or by the rank defect: see Figure 13.

PROOF. This is basically a restatement of Theorem 18 in the limiting cases  $\Lambda = 0$  and  $\Lambda = \infty I$ . Indeed: suppose for example that  $|I| + |J| \geq n$ . (The case where  $|I| + |J| \leq n$  can be handled in a similar way). As in the paragraphs preceding this corollary, we can translate the data in terms of a structure block  $\mathcal{B}_{\text{genPo}} = (S, T, \tilde{R}, \tilde{S}, r, 0)$ . Then the inverse matrix satisfies  $\mathcal{B}_{\text{genPo}}^{-1} = (\tilde{S}, \tilde{T}, R, S, r, \infty I)$ ; the  $\infty I$  means that all rows and columns in  $\tilde{S}, S$  should be dropped, and that the rank upper bound  $r$  should decrease by exactly this same value  $|S|$ , i.e.

$$\text{Null } A^{-1}(\tilde{T}, R) = \text{Null } A(S \cup T, \tilde{R} \cup \tilde{S}).$$

The proof is finished since by construction,  $I = S \cup T$  and  $J = \tilde{R} \cup \tilde{S}$ .  $\square$

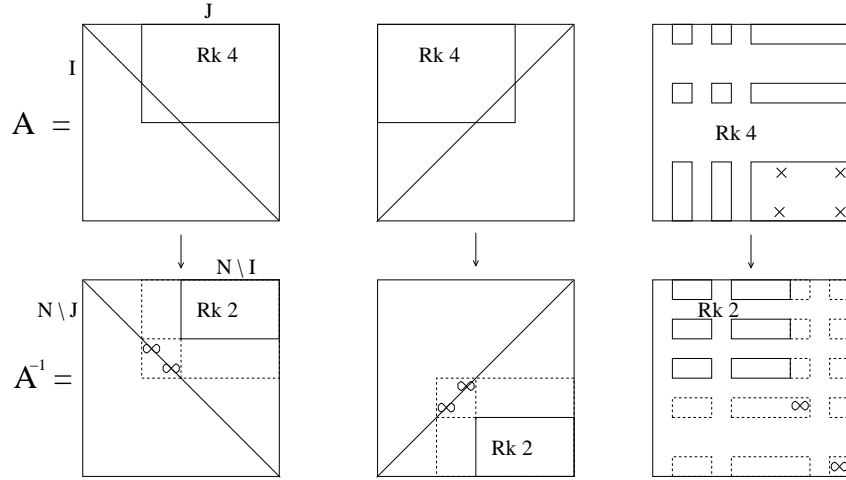


Figure 13: In case  $\Lambda = 0$ , we get the following specification of Figure 11.

## 5 Conclusion

In this paper we investigated some structures on  $\mathbb{C}^{n \times n}$  that have a good behaviour under matrix inversion. We handled two classes of them, namely ‘de-

coupled' displacement structures and rank structures. For the case of rank structures, we provided some generalizations to deal with singular shift matrices and generally positioned structure blocks, leading amongst others to a different interpretation of a theorem due to Fiedler [5], which corresponds to the limiting cases  $\Lambda = 0$  and  $\Lambda = \infty I$ . In [1], we show that these structures have also a good behaviour under Schur complementation.

## References

- [1] S. Delvaux and M. Van Barel. Structures preserved by Schur complementation. Report TW 415, Department of Computer Science, K.U.Leuven, Leuven, Belgium, December 2004.
- [2] S. Delvaux and M. Van Barel. Structures preserved by the QR-algorithm. Report TW 399, Department of Computer Science, K.U.Leuven, Leuven, Belgium, August 2004.
- [3] P. Dewilde and A.-J. van der Veen. *Time-varying systems and computations*. Kluwer academic publishers, Boston, June 1998.
- [4] Y. Eidelman and I. Gohberg. Fast inversion algorithms for diagonal plus semiseparable matrices. *Integral Equations and Operator Theory*, 27(2):165–183, 1997.
- [5] M. Fiedler and T. L. Markham. Completing a matrix when certain entries of its inverse are specified. *Linear Algebra and Its Applications*, 74:225–237, 1986.
- [6] G. Heinig and K. Rost. *Algebraic Methods for Toeplitz-like Matrices and Operators*. Mathematical Research. Akademie Verlag-Berlin, 1984.
- [7] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins university Press, third edition, 1996.
- [8] D. A. Gregory, B. Heyink, and K. N. Vander Meulen. Inertia and biclique decompositions of joins of graphs. *J. Combin. Theory*, (Series B 88):135–151, 2003.
- [9] T. Kailath, S. Kung, and M. Morf. Displacement ranks of matrices and linear equations. *Journal of Mathematical Analysis and its Applications*, 68:395–407, 1979.
- [10] T. Kailath and A. H. Sayed. Displacement structure: theory and applications. *SIAM Review*, 37:297–386, 1995.
- [11] T. Kailath and A. H. Sayed, editors. *Fast reliable algorithms for matrices with structure*. SIAM, Philadelphia, PA, USA, May 1999. ISBN0-89871-431-1.

- [12] G. Strang and T. Nguyen. The interplay of ranks of submatrices. *SIAM Review*, 46(4):637–646, 2004.
- [13] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi. Orthogonal rational functions and structured matrices. *SIAM Journal on Matrix Analysis and its Applications*, 2004. Accepted for publication.
- [14] R. Vandebril and M. Van Barel. A short note on the nullity theorem. Report TW 397, Department of Computer Science, K.U.Leuven, Leuven, Belgium, 2004.