STRUCTURED MATRICES AND INVERSES *

P. COMON†

Abstract. A matrix (and any associated linear system) will be referred to as structured if it has a small displacement rank. It is known that the inverse of a structured matrix is structured, which allows fast inversion (or solution), and reduced storage requirements. According to two definitions of displacement structure of practical interest, it is shown here that several types of inverses are also structured, including the Moore-Penrose inverse of rank-deficient matrices.

Key Words. Displacement rank, Structured matrix, Töplitz, Hankel, Inverse, Schur, Moore-Penrose, Pseudo-inverse, Deconvolution.

AMS(MOS) subject classification. 15A03, 15A06, 15A09, 15A57, 65F20, 65F30.

1. Introduction. Close to Töplitz or close to Hankel matrices appear in various areas including signal processing and automatic control (e.g. prediction of second-order nearly stationary time series). In radar or sonar (or more generally antenna processing), Töplitz matrices are encountered when far-field sources impinge an array of regularly spaced sensors after propagating through an homogeneous medium. If 2-dimensional regular arrays are utilized, then block-Töplitz matrices can be found. Other applications include optics, image processing (when the spreading function is shift invariant), differential or integral equations under certain boundary conditions and for certain discretizations (e.g. oil prospecting), seismics, geophysics, transmission lines, and communications... In general, these applications correspond to the solution of some inverse problems. When shift invariance properties are satisfied, the linear operator to invert is Töplitz, or block-Töplitz, and it is dealt with a deconvolution problem.

However, Töplitz matrices in the strict sense are rarely encountered in the real word, because the abovementioned invariance properties are not satisfied. For instance, second-order stationarity of long time series, or homogeneity of propagation media, are idealized assumptions. In antenna array processing, the decalibration of the array is the main cause of many problems, among which the deviation from Töplitz is one of the mildest ones. For instance in sonar, decalibration occurs because of the effects of pressure, temperature, and usage, among others. Another major cause of problems is the distorsion of wave fronts impinging the array due to inhomogeneity of the medium or to local turbulences (note that improvements can be obtained by assuming that the celerity is random with a small variance, but this is out of the scope of the present discussion). Lastly, a simple

---

* This work has been carried out while the author was visiting the Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, in April 1992.

† Thomson-Sintra, Sophia Antipolis BP138, 06561 Valbonne Cedex France. The author can be reached at na.comon@na-net.ornl.edu.

deviation from Töplitz that has been already studied is the effect of limited length of the data. The proximity to Töplitz then depends on the way the matrix is calculated: its displacement rank ranges from 2 to 4 under ideal assumptions.

Since the set of Töplitz (or Hankel) matrices is a linear space, it is easy to compute the closest Töplitz approximate of any matrix by a simple projection. However, this operation should be avoided in general, since it would destroy other important structures (e.g. just the rank). On the other hand, finding the best approximate of a matrix by another of given rank and given displacement rank is still an open problem. It is true that some simple iterative algorithms have already been proposed in the literature for the Töplitz case, but the convergence issue has not been completely covered.

Since the early works by Schur (1917), Levinson (1947), Durbin (1960), Trench (1964), and Bareiss (1969), a lot of work has been done. In particular, Kailath and others introduced in the seventies the concept of displacement rank, which allows in some way to measure a distance to Töplitz [16]. By the way, the concept of displacement rank may be seen to have some connections with integral and differential equations [18]. An excellent survey of related works can be found in [17]. Other recent investigations are also reported in [6].

It is known that a linear system $Tx = b$ can be solved with $O(n^2)$ flops if T is a $n \times n$ Töplitz matrix. If T is just only close to Töplitz, it is useful to define a displacement rank, δ, measuring a distance to the Töplitz structure [11]. Then it has been shown that the solution requires only $O(\delta n^2)$ flops, to be compared to the $O(n^3)$ complexity necessary to solve a dense linear system of general form. More recently, superfast algorithms have been proposed to solve Töplitz systems, and their complexity ranges from $O(nlog^2 n)$ to $O(\alpha nlogn)$, $\alpha < n$, [1] [2] [3].

The displacement rank of a linear system is clearly related to the complexity of its solution. It has been shown in [4] [5] that this complexity reduction also holds for the calculation of various factorizations, provided the Schur algorithm is run on the appropriate block-matrix. In this paper, the displacement rank will be defined in a slightly more general framework, such that the structure of a wider class of matrices can be taken into account. In this framework, the first step in the quest of fast algorithms is to check whether the system considered has a displacement structure, and under what displacement operator its displacement rank is the smallest. Building explicitly fast algorithms taking advantage of this structure is the next question. However, our investigations are limited in this paper to the study of the displacement rank itself, and it will not be discussed how to build the corresponding fast algorithm.

The paper is organized as follows. Definitions and basic properties are given in section 2. In section 3, the structure of inverses and products of full-rank structured matrices is analyzed. Section 5 is devoted to the study of structured rank-deficient matrices, and utilizes preliminary results derived in section 4.

2. Definitions and first properties. The structure that will be considered in this paper is exclusively the *displacement structure* [16] [11]. Roughly speaking, a structured matrix is the sum of displaced versions of a unique generating matrix of small rank. For instance, sparse matrices may not have any interesting displacement structure. Displacement operators can be defined in different ways, and two definitions will be used subsequently.

DEFINITION 2.1. *For any fixed pair of matrices (Z, N) of appropriate dimension, the displacement of a matrix A with respect to displacement operator $\nabla_{Z,N}$ is defined as*

$$\nabla_{Z,N}A = A - ZAN. \tag{1}$$

DEFINITION 2.2. *For any fixed pair of matrices (Z, N) of appropriate dimension, the displacement of a matrix A with respect to displacement operator $\Delta_{Z,N}$ is defined as*

$$\Delta_{Z,N}A = ZA - AN. \tag{2}$$

In the remaining, matrices Z and N will be referred to as *displacement matrices*, and the pair $\{, Z, N\}$ to as the *displacement pattern.* . Once the above definitions are assumed in the primal space, then it is convenient to use the definitions below in the dual space, denoting by $(^*)$ the transposition:

$$\begin{aligned} \nabla_{N,Z}(A^*) &= A^* - NA^*Z, \\ \Delta_{N,Z}(A^*) &= NA^* - A^*Z. \end{aligned} \tag{3}$$

DEFINITION 2.3. *Matrices for which any of the four displaced matrices (1), (2), (3) or (3) has a rank bounded by a value that does not depend on the size of A will be referred to as structured. This rank will be called the displacement rank of A with respect to the displacement operator considered, and will be denoted as $\delta^{\nabla}_{Z,N}\{A\}$, $\delta^{\nabla}_{N,Z}\{A^*\}$, $\delta^{\Delta}_{Z,N}\{A\}$, or $\delta^{\Delta}_{N,Z}\{A^*\}$.*

This definition is consistent with [6]. Displacement matrices Z and N are usually very simple (typically formed only of ones and zeros). Additionally, it can be seen that the displacement operator (1) is easily invertible as soon as *either Z or N* is nilpotent. To see this, assume that $Z^{k+1} = 0$

and explicit the displacement $\nabla_{Z,N}$ in the sum

$$\sum_{i=0}^{k} \nabla_{Z,N}\{Z^i A N^i\}.$$

Then this expression can be seen to be nothing else than A itself. For additional details, see [17] and references therein. Note that the results shown in this paper will not require a particular form for matrices Z and N (nilpotent for instance), unless otherwise specified. Other considerations on invertibility of displacement operators are also tackled in [4]. In [20], displacement operators are defined (in a manner very similar to [4]), but displacement ranks of products or pseudo-inverses are unfortunately not obtained explicitely. Lastly, other displacement structures, including (2), are being investigated by G.Heinig.

EXAMPLE 2.4. Denote S the so-called lower shift matrix:

$$S = \begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}. \tag{4}$$

For Hankel matrices, it is easy to check out that we have

$$\delta^{\nabla}_{Z,N}\{H\} \le 2, \quad \text{for } (Z,N) = (S,S), \tag{5}$$

$$\delta^{\Delta}_{Z,N}\{H\} \le 2, \quad \text{for } (Z,N) = (S,S^*), \tag{6}$$

whereas for Töplitz matrices, we have

$$\delta^{\nabla}_{Z,N}\{T\} \le 2, \quad \text{for } (Z,N) = (S,S^*), \tag{7}$$

$$\delta^{\Delta}_{Z,N}\{T\} \le 2, \quad \text{for } (Z,N) = (S,S). \tag{8}$$

In these particular cases, the non-zero entries of displaced matrices are indeed contained only in one row and one column. These four statements hold also true if matrices Z and N are permuted. In other words,

$$\delta^{\Delta}_{S^*,S}\{H\} \le 2, \text{ and } \delta^{\nabla}_{S^*,S}\{T\} \le 2.$$

It turns out that the definitions 2.1 and 2.2 yield displacement ranks that are not independent to each other. We have indeed the following

THEOREM 2.5. *For any given matrices Z, N, A, the two inequalities below hold*

$$\delta^{\nabla}_{Z,N}\{A\} \le \delta^{\Delta}_{Z^*,N}\{A\} + \delta^{\nabla}_{Z,Z^*}\{I\}, \tag{9}$$

$$\delta^{\Delta}_{Z,N}\{A\} \le \delta^{\nabla}_{Z^*,N}\{A\} + \delta^{\nabla}_{Z,Z^*}\{I\}, \tag{10}$$

where I denotes the identity matrix having same dimensions as A.

Proof. $\nabla_{Z,N}A = Z(Z^*A - AN) + (I - ZZ^*)A$ shows the first inequality, and $\Delta_{Z,N}A = Z(A - Z^*AN) - (I - ZZ^*)AN$ shows the second one. $\square$

EXAMPLE 2.6. If A is a circulant Töplitz matrix, e.g.,

$$A = \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \end{pmatrix},$$

then it admits a displacement rank $\delta^{\nabla}_{Z,N}\{A\} = 1$ provided the following displacement pattern is assumed: $Z = S_3$ as given by (4), and

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

In this example, we also have $\delta^{\Delta}_{Z^*,N}\{A\} = 1$, which is conform to theorem 2.5.

EXAMPLE 2.7. Let A be a $m \times n$ Töplitz matrix. Define $N = S_n$, and

$$Z = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then it can be seen that $\delta^{\Delta}_{Z^*,N}\{A\} = 2$, and $\delta^{\nabla}_{Z,N}\{A\} = 3$. This example shows that equality can occur in theorem 2.5.

EXAMPLE 2.8. If T is Töplitz $m \times n$ and H is Hankel $m \times p$, then the block matrix $(T\ H)$ has a displacement rank equal to 3 with respect to the displacement pattern $\{Z, N\} = \{S_m, S_n^* \oplus S_p\}$.

The notation $A \oplus B$ will be subsequently used when A and B are square to denote the block-diagonal matrix having A and B as diagonal blocks.

3. Displacement of various inverses and products. There is a number of situations where the displacement rank of a matrix can be quite easily shown to be small. Since our main concern is inverses, let us start with the simplest case.

THEOREM 3.1. *Let A be an invertible square matrix. Then*

$$(i)\ \delta^{\Delta}_{Z,N}\{A\} = \delta^{\Delta}_{N,Z}\{A^{-1}\},\ \textit{and}\ (ii)\ \delta^{\nabla}_{Z,N}\{A\} = \delta^{\nabla}_{N,Z}\{A^{-1}\}. \tag{11}$$

In other words, A and A^{-1} have the same displacement rank with respect to dual displacement patterns.

To prove the theorem, it it useful to recall the following lemma.

LEMMA 3.2. *Let f and g be two linear operators, and denote E_λ^h the eigenspace of operator h associated with the eigenvalue λ. If λ is an eigenvalue of fog, then it is also an eigenvalue of gof. In addition, the eigenspaces have the same dimension as soon as λ is non-zero:*

$$dim\{E_\lambda^{fog}\} = dim\{E_\lambda^{gof}\}.$$

Proof. Assume λ is an eigenvalue of fog. Then for some non-zero vector x, $fog(x) = \lambda x$. Composing both sides by operator g immediately shows that

$$gof(g(x)) = \lambda g(x). \tag{12}$$

Next there are two cases: (i) if $g(x) \neq 0$, then $g(x)$ is an eigenvector of gof associated with the same eigenvalue λ; (ii) if $g(x) = 0$, then $fog(x) = 0$ and necessarily $\lambda = 0$. Now assume without restricting the generality of the proof that $dim\{E_\lambda^{fog}\} > dim\{E_\lambda^{gof}\}$. Then there exists a vector x in E_λ^{fog} such that $g(x) = 0$ (since otherwise relation (12) would imply that g(x) is also in E_λ^{gof}). Yet composing by f yields $fog(x) = 0$ and consequently $\lambda = 0$. As a conclusion, if $\lambda \neq 0$, eigenspaces must have the same dimension. □

Proof. of theorem. We have by definition $\delta_{Z,N}^{\Delta}\{A\} = rank\{ZA - AN\} = rank\{Z - ANA^{-1}\}$, and $\delta_{N,Z}^{\Delta}\{A^{-1}\} = rank\{NA^{-1} - A^{-1}Z\} = rank\{ANA^{-1} - Z\}$. But these two matrices are opposite, and therefore have the same rank. This proves (i).

Similarly since the rank does not change by multiplication by a regular matrix, we have $\delta_{Z,N}^{\nabla}\{A\} = rank\{A - ZAN\} = rank\{I - ZANA^{-1}\}$. On the other hand $\delta_{N,Z}^{\nabla}\{A^{-1}\} = rank\{A^{-1} - NA^{-1}Z\} = rank\{I - NA^{-1}ZA\}$. Now from lemma 3.2 we know that $ker\{I - fog\}$ and $ker\{I - gof\}$ have the same dimension. If f and g are endomorphisms in the same space, this implies in particular that $rank\{I - fog\} = rank\{I - gof\}$. Now applying this result to $f = ZA$, $g = NA^{-1}$ eventually proves (ii). □

The proof that an invertible matrix and its inverse have the same displacement rank has been known for a long time, and proved for symmetric matrices [17]. However, the proof for general Töplitz matrices seems to have been given only recently in [6] for a displacement of type (1). Our theorem is slightly more general.

COROLLARY 3.3. *For any given square matrix A, let the regularized inverse be given by $R = (A + \eta I)^{-1}$, for some number η such that $A + \eta I$ is regular. Then the displacement ranks of A and R are linked by the inequality below*

$$\delta_{N,Z}\{R\} \leq \delta_{Z,N}\{A\} + \delta_{Z,N}\{I\}, \tag{13}$$

this inequality holding for both displacements ∇ and Δ.

Proof. Just write $\delta^{\nabla}_{N,Z}\{R\} = \delta^{\nabla}_{Z,N}\{R^{-1}\} = \delta^{\nabla}_{Z,N}\{A+\eta I\}$, and since the rank of a sum is smaller than the sum of the ranks, we eventually obtain the theorem. In order to prove the inequality for the displacement Δ, proceed exaclty the same way. □

When close to Töplitz or close to Hankel matrices are considered, the displacement matrices Z and N are essentially either the lower shift matrix S or its transposed. In such a case, it is useful to notice that

$$\delta^{\nabla}_{S,S^*}\{I\} = \delta^{\nabla}_{S^*,S}\{I\} = 1. \tag{14}$$

On the other hand for any matrix Z (and S or S^* in particular):

$$\delta^{\Delta}_{Z,Z}\{I\} = 0. \tag{15}$$

For a Töplitz matrix T, we have a stronger (and obvious) result, because T and $T+\eta I$ are both Töplitz.

$$\delta^{\nabla}_{S^*,S}\{R\} = \delta^{\nabla}_{S,S^*}\{T\}.$$

COROLLARY 3.4. *Let M be the 2×2 block matrix below*

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where A and D are square of dimension n_1 and n_2, respectively. Assume M and A are invertible. When the last $n_2\times n_2$ block of the matrix M^{-1} is invertible, it can be written as $\bar{A}^{-1}$, where $\bar{A}$ is the so-called Schur complement of A in M: $\bar{A} = D - CA^{-1}B$. If M has a displacement rank $\delta_{N,Z}\{M\}$ with respect to a displacement pattern $\{Z,N\} = \{Z_1\oplus Z_2, N_1\oplus N_2\}$, where Z_i and N_i are $n_i\times n_i$ matrices, then the displacement rank of $\bar{A}$ satisfies the inequality below for both displacements ∇ and Δ:

$$\delta_{Z_2,N_2}\{\bar{A}\} \le \delta_{Z,N}\{M\}. \tag{16}$$

Proof. Applying twice the theorem 3.1, and noting that the rank of M is always larger than the rank of any of its submatrices, yield $\delta_{Z_2,N_2}\{\bar{A}\} = \delta_{N_2,Z_2}\{\bar{A}^{-1}\} \le \delta_{N,Z}\{M^{-1}\} = \delta_{Z,N}\{M\}$. □

This kind of property has been noticed for several years by Chun and Kailath. See for instance [4] [6]. This corollary restates it in the appropriate framework.

THEOREM 3.5. *Let A_1 and A_2 be two full-rank matrices of size $n_1\times n_2$ and $n_2\times n_1$, respectively, with $n_1\le n_2$. Then the displacement rank of the* *matrix A_1A_2 is related to the displacement ranks of A_1 and A_2 for either displacement ∇ or Δ by*

$$\delta_{Z_1,Z_2}\{A_1A_2\} \leq \delta_{Z_1,N_1}\{A_1\} + \delta_{N_1,N_2}\{I_{n_2}\} + \delta_{N_2,Z_2}\{A_2\}, \tag{17}$$

Proof. To prove the theorem, form the square matrix M of size n_1+n_2:

$$M = \begin{pmatrix} I & A_2 \\ A_1 & 0 \end{pmatrix},$$

consider the displacement pattern $\{N_2 \oplus Z_1, N_1 \oplus Z_2\}$ and apply corollary 3.4. Again, since the displacement pattern is block-diagonal, the displaced block matrix is formed of the displaced blocks.

In the present case, the Schur complement is precisely the product $-A_1A_2$. This proof is identical to that already proposed in [6] for particular structured matrices. □

Note that if $N_1 = N_2$, (15) implies $\delta^{\Delta}_{N_1,N_2}\{I\} = 0$. On the other hand, if $N_1 = N_2^* = S$, then $\delta^{\nabla}_{N_1,N_2}\{I\} = 1$ from (14). For particular displacement matrices Z and N, the general bounds given by theorem 3.5 may be too loose. In particular for Töplitz or Hankel matrices, the corollary below is more accurate.

Corollary 3.6. *Let S be the lower shift matrix defined in (4), T_1 and T_2 be Töplitz matrices, and H_1 and H_2 be Hankel. Then under the conditions of theorem 3.5:*

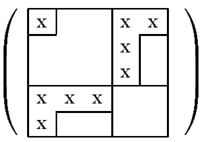
$$\textbf{(a)}\ \delta^{\Delta}_{S,S}\{T_1T_2\} \leq 4, \quad \textbf{(b)}\ \delta^{\Delta}_{S,S}\{H_1H_2\} \leq 4, \quad \textbf{(c)}\ \delta^{\Delta}_{S,S^*}\{T_1H_2\} \leq 4, \tag{18}$$

$$\textbf{(a)}\ \delta^{\nabla}_{S,S^*}\{T_1T_2\} \leq 4, \quad \textbf{(b)}\ \delta^{\nabla}_{S,S^*}\{H_1H_2\} \leq 4, \quad \textbf{(c)}\ \delta^{\nabla}_{S,S}\{T_1H_2\} \leq 4. \tag{19}$$

Proof. Equations (18) result from a combination of example 2.4 and theorem 3.5. In fact, take $Z_i = N_i = S$ for **(a)**, $Z_1 = Z_2 = N_1^* = N_2^* = S$ for **(b)**, and $Z_1 = N_1 = N_2 = Z_2^* = S$ for **(c)**.

On the other hand, if we try to apply theorem 3.5 to prove (19), we find a result weaker than desired, for we obtain $\delta^{\nabla} \leq 5$. A more careful analysis is therefore necessary. Restart the proof of theorem 3.5: if T_1 and T_2 are full rank Töplitz, the displaced block matrix $\nabla_{S\oplus S,S^*\oplus S^*}M$ has the following form:

$$\begin{pmatrix} \nabla I & \nabla T_2 \\ \nabla T_1 & 0 \end{pmatrix} =$$



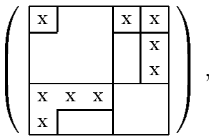
where crosses indicate the only locations where the matrix is allowed to have non-zero entries: only in two rows and two columns. Such a matrix is clearly of rank at most 4. Following the same lines as in theorem 3.5, it can be seen that the product T_1T_2 has a displacement rank bounded by 4.

A similar proof could be derived in the case of two Hankel matrices, and will not be detailed here. In order to prove (19**c**), let us consider finally the block matrix

$$M = \begin{pmatrix} I & H_2 \\ T_1 & 0 \end{pmatrix}.$$

Assuming the displacement pattern $\{S \oplus S, S^* \oplus S\}$, the displaced matrix ∇M is now of the form

$$\begin{pmatrix} \nabla I & \nabla H_2 \\ \nabla T_1 & 0 \end{pmatrix} =$$



which is again obviously of rank at most 4. The last result follows. □

The theorem 3.5 was valid only for full-rank matrices of transposed sizes. For further purposes it is useful to extend it to products of rectangular matrices of general form.

THEOREM 3.7. *Let A and B be $m \times n$ and $n \times p$ matrices. Then the product AB is also structured, and the inequality below holds:*

$$\delta^{\Delta}_{Z_A,N_B}\{AB\} \leq \delta^{\Delta}_{Z_A,N_A}\{A\} + \delta^{\Delta}_{N_A,Z_B}\{I_n\} + \delta^{\Delta}_{Z_B,N_B}\{B\}, \tag{20}$$

where I_n denotes the $n \times n$ identity matrix.

Proof. Write first the displaced matrix as

$$\Delta_{Z_A,N_B}AB = (Z_AA - AN_A)B + A(N_AB - BN_B).$$

Then splitting the second term into $A(N_A - Z_B)B + A(Z_BB - BN_B)$ gives

$$\Delta_{Z_A,N_B}(AB) = \Delta_{Z_A,N_A}A \ . \ B + A \ . \ \Delta_{N_A,Z_B}I \ . \ B + A \ . \ \Delta_{Z_B,N_B}B, \tag{21}$$

which eventually proves the theorem, since the rank of a product is always smaller than the rank of each of its terms. □

A similar result holds for displacement ∇. A direct consequence of equation (21) is the following corollary, that looks like a differentiation rule.

COROLLARY 3.8. *When A and B have dual displacement patterns, we obtain the following simple result:*

$$\Delta_{Z,Z}(AB) = \Delta_{Z,N}A \ . \ B + A \ . \ \Delta_{N,Z}B.$$

In particular, if $AB = I$, this displaced matrix is null, because of (15).

COROLLARY 3.9. *Let A be a full-rank $m \times n$ matrix, with $m \geq n$. Then its pseudo-inverse $B = (A^*A)^{-1}A^*$ has a reduced displacement rank, as show the two bounds below:*

$$\delta^{\Delta}_{N,Z}\{B\} \leq \delta^{\Delta}_{Z,N}\{A\} + 2\delta^{\Delta}_{N,Z}\{A^*\}, \tag{22}$$

$$\delta^{\Delta}_{N,Z^*}\{B\} \leq 3\delta^{\Delta}_{Z,N}\{A\} + \delta^{\Delta}_{Z^*,Z}\{I_m\}. \tag{23}$$

Proof. Apply corollary 3.8 to A^*A, next theorem 3.1, and lastly theorem 3.7. □

EXAMPLE 3.10. If A is Töplitz, equation (22) claims that $\delta^{\Delta}_{S,S}\{B\} \leq 6$. In practice, it seems that no Töplitz matrix could yield a displacement rank larger than $\delta^{\Delta}_{S,S}\{B\} = 4$, which suggests that the bound is much too large.

DEFINITION 3.11. *Given any matrix A, if a matrix A^- satisfies*

$$\begin{aligned} &(i)\ AA^-A = A, \quad (ii)\ A^-AA^- = A^-, \\ &(iii)\ (AA^-)^* = AA^-, \quad (iv)\ (A^-A)^* = A^-A, \end{aligned} \tag{24}$$

then it will be called the Moore-Penrose (MP) pseudo-inverse of A. A so-called generalized inverse need only to satisfy conditions (i) and (ii).

It is well known that A^- is unique, and that A^- and A^* have the same range and the same null space [12]. On the other hand, a generalized inverse is not unique. When a matrix A is rank deficient, it is in general not possible to construct a MP pseudo-inverse having the same displacement rank, as will be demonstrated in section 5.

4. The space of P-symmetric matrices. In this section, more specific properties shared by matrices in a wide class will be investigated. The property of P-symmetry will be necessary in section 5 to transform a matrix into its transposed just by a congruent transformation.

DEFINITION 4.1. *Let P be a fixed orthogonal n by n matrix. The set of P-symmetric matrices is defined as follows:*

$$\mathcal{S}_P = \{M \in \mathbb{R}^{n\times n} / PMP^* = M^*\}, \tag{25}$$

where $(^*)$ denotes transposition and $\mathbb{R}$ the set of real numbers.

It will be assumed in this section that the matrix to invert (or the system to solve) belongs to $\mathcal{S}_P$, for some given known orthogonal matrix P. For instance, if a matrix A is square and Töplitz, then it is centro-symmetric and satisfies

$$JAJ^* = A^*,$$

which shows that $A \in \mathcal{S}_J$, where J denotes the reverse identity:

$$J = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix}. \tag{26}$$

If A is Hankel, then $A \in \mathcal{S}_I$ because A is symmetric. The property of P-symmetry is interesting for it is preserved under many transformations. For instance, singular vectors of a P-symmetric matrix are P-symmetric in the sense that if $\{u, v, \sigma\}$ is a singular triplet, then so is $\{Pv, Pu, \sigma\}$. A sum or a product of P-symmetric matrices is P-symmetric.

EXAMPLE 4.2. Define the 'alternate Töplitz matrix' below

$$A = \begin{pmatrix} 2 & -2 & -2 & 1 & 1 \\ 1 & -2 & -2 & 2 & 1 \\ 1 & 1 & 2 & -2 & -2 \\ 4 & -1 & 1 & -2 & -2 \\ -8 & 4 & 1 & 1 & 2 \end{pmatrix},$$

and assume the displacement pattern

$$Z = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}, \text{ and } N = -Z^*.$$

Then we have $PAP^* = A^*$ as requested in the definition above, with $P = J$. This matrix has displacement ranks $\delta^{\nabla}_{Z,N}\{A\} = 2$ and $\delta^{\Delta}_{Z^*,N}\{A\} = 2$, and is singular. The displacement rank of its MP pseudo inverse will be calculated in example 5.4.

PROPERTY 4.3. *The properties of P-symmetry and P^*-symmetry are equivalent.*

Proof. Let A be P-symmetric. Then transposing (25) gives $M = PM^*P^*$. Next pre- and post-multiplication by P^* and P, respectively, yields $P^*MP = M^*$. □

THEOREM 4.4. *If A is P-symmetric, then so is A^{-1} whenever A is invertible. If A is singular, then its Moore-Penrose inverse, A^-, is also P-symmetric.*

Proof. Inversion of both sides of the relation $PAP^* = A^*$ yields immediately $PA^{-1}P^* = A^{-1*}$. Now to insure that when A is singular, A^- is P-symmetric, it suffices to prove that the matrix $B = PA^{-*}P^*$ indeed satisfies the four conditions of definition (24). First, $ABA = APA^{-*}P^*A$ yields $ABA = PA^*A^{-*}A^*P^* = PA^*P = A$, which shows (i) of (24). Second, $BAB = PA^{-*}P^*APA^{-*}P^*$ yields similarly $BAB = PA^{-*}A^*A^{-*}P^* =$

$PA^{-*}P^*$, which equals B by definition. Next to prove (iii), consider $AB = APA^{-*}P^*$, which gives after premultiplication by PP^*: $AB = PA^*A^{-*}P^*$. But since $(A^-A)^* = A^-A$, we have $AB = PA^-AP^*$. Then insertion of P^*P yields finally $AB = PA^-P^*A^*$, which is nothing else then B^*A^*. The proof of (iv) can be derived in a similar manner. □

It may be seen that in the last proof, A does not need to be a normal matrix, which was requested in a similar statement in [15]. On the other hand, it is true that if A is P-symmetric, AA^* is in general not P-symmetric.

5. Displacement of MP pseudo-inverses. In section 3, it has been shown among other things that the pseudo-inverse of a full-rank matrix is structured. It will be now analyzed how the rank deficience weakens the structure of the MP pseudo-inverse.

THEOREM 5.1. *Let A be a P-symmetric square matrix, and let Z and N be two displacement matrices linked by the relation*

$$PZP = N. \tag{27}$$

Then the displacement ranks of A and A^- are related by

$$\delta^{\nabla}_{N,Z}\{A^-\} \leq 2\delta^{\nabla}_{Z,N}\{A\}. \tag{28}$$

In this theorem, the condition (27) is satisfied in particular for both close to Töplitz and close to Hankel matrices, with $(P, Z, N) = (J, S, S^*)$ and $(P, Z, N) = (I, S, S)$, respectively.

Proof. For conciseness, denote in short δ the displacement rank $\delta^{\nabla}_{Z,N}\{A\}$, and assume A is $n \times n$. In order to prove the theorem, it is sufficient to find two full-rank $n \times n - \delta$ matrices E_1 and E_2 such that

$$E_2^*\nabla_{N,Z}\{A^-\}E_1 = 0. \tag{29}$$

For this purpose, define the following full-rank matrices with n rows:

$$\begin{array}{rl} G_1 = & matrix\ whose\ columns\ span\quad Ker\nabla A \\ G_2 = & matrix\ whose\ columns\ span\quad Ker(\nabla A)^* \\ K_1 = & matrix\ whose\ columns\ span\quad KerAN \cap Ker\nabla A \\ K_2 = & matrix\ whose\ columns\ span\quad Ker(ZA)^* \cap Ker(\nabla A)^* \\ V_1 = & matrix\ whose\ columns\ span\quad ANG_1 \\ V_2 = & matrix\ whose\ columns\ span\quad A^*Z^*G_2. \end{array} \tag{30}$$

Then define the two matrices E_i as:

$$E_i = [V_i,\ W_i], with\ W_i = PK_i.$$

Let us prove first that E_i are indeed of rank $n-\delta$, and then that (29) is satisfied.

From (30), we have by construction $AK_1 = 0$. Then inserting a factor P^*P and premultiplying by P gives $PAP^*PK_1 = 0$, which shows that $A^*W_1 = 0$. Yet, V_1 is in the range of A by definition, and thus V_1 and W_1 are necessarily orthogonal as members of the orthogonal subspaces $KerA^*$ and ImA. In addition, P is bijective so that W_1 has the same dimension as K_1. As a consequence, $dimE_1 = dimV_1 + dimK_1$, which is nothing else but $dimG_1$ if we look at the definitions (30). Similarly, one can show that W_2 and V_2 are orthogonal because $W_2 \subset KerA$. This yields after the same argumentation that $dimE_2 = dimG_2 = n-\delta$.

Now it remains to prove (29). To do this, it is shown that the four blocks of $E_2^*\nabla A^- E_1$ are zero. The quantity $\mu = V_2^*\nabla A^- V_1$ is null since $\mu = G_2^*ZA(A^- - NA^-Z)ANG_1$ can be written $\mu = G_2^*ZANG_1 - G_2^*ZANA^-ZANG_1$, which is the difference of two identical terms by construction of matrices G_i. In fact from (30), $ZANG_1 = AG_1$ and $G_2^*ZAN = G_2^*A$. Next $W_2^*\nabla A^-$ is null because $W_2^*\nabla A^*$ is null (remember that A^- and A^* have the same null space). In fact, $W_2^*\nabla A^* = K_2^*P^*A^* - K_2^*P^*NA^*Z$ by definition of W_2 and ∇. Now using the relation (27) and P-symmetry of A yield $W_2^*\nabla A^*P = K_2^*A - K_2^*ZAN$. These two terms are eventually null by construction of K_2. It can be proved in a similar manner that $\nabla A^*W_1 = 0$. In fact, $\nabla A^*W_1 = A^*PK_1 - NA^*ZPK_1$ implies $P^*\nabla A^*W_1 = AK_1 - P^*NP^*PA^*P^*PZPK_1$. Again these two terms can be seen to be zero utilizing (27), P-symmetry of A, and the definition (30) of K_1. □

This theorem is an extension of a result first proved in [7]. As pointed out in [8], when the displacement rank of A is larger than its rank, the theorem above gives too weak results as is next shown.

THEOREM 5.2. *Let A be a square matrix, and denote by $r\{A\}$ its rank. Then there exist two other bounds for the displacement rank of its MP pseudo-inverse:*

$$\delta^{\nabla}_{N,Z}\{A^-\} < 2r\{A\} \quad if \ \delta^{\nabla}_{Z,N}\{A\} < 2r\{A\}, \text{ and} \tag{31}$$

$$\delta^{\nabla}_{N,Z}\{A^-\} \le 2r\{A\} \text{ otherwise.} \tag{32}$$

Proof. The proof of (32) is easy. In fact, it holds true for any matrix M since

$$rank\{M - ZMN\} \le rank\{M\} + rank\{ZMN\} \le 2rank\{M\}$$

is always true. So let us prove (31). Since A has rank $r\{A\}$, it may be written as

$$A = U\Sigma V^*,$$

where Σ is invertible and of size $r\{A\}$. Define the matrices $\mathcal{A} = [U\ ZU]$, $\mathcal{B} = [V\ N^*V]$, and $\Lambda = Diag(\Sigma, -\Sigma)$. Then it may be seen that

$$\nabla A = \mathcal{A}\Lambda\mathcal{B}^*.$$

Since Λ is of full rank, either $\mathcal{A}$ or $\mathcal{B}$ must be rank defficient, otherwise ∇A would be of rank $2rank\{A\}$ which is contrary to the hypothesis. Thus assume without restricting the generality of the proof that $rank\{\mathcal{A}\} < 2r\{A\}$. Then $rank\{\nabla A^-\} < 2r\{A\}$ because:

$$\nabla A^{-*} = \mathcal{A}\Lambda^{-1}\mathcal{B}^*.$$

This completes the proof. □

COROLLARY 5.3. *Let T and H be close to Töplitz and close to Hankel square matrices, respectively. Then*

$$\delta^{\nabla}_{S^*,S}\{T^-\} \leq 2\delta^{\nabla}_{S,S^*}\{T\},\quad \delta^{\nabla}_{S,S}\{H^-\} \leq 2\delta^{\nabla}_{S,S}\{H\}, \tag{33}$$

$$\delta^{\Delta}_{S,S}\{T^-\} \leq 2\delta^{\Delta}_{S,S}\{T\} + 1,\quad \delta^{\Delta}_{S^*,S}\{H^-\} \leq 2\delta^{\Delta}_{S,S^*}\{H\} + 1. \tag{34}$$

Proof. To prove (33), simply use theorem 5.1 and relation (14). In order to prove equations (34), utilize theorem 2.5 and relation (15). □

Note that the bounds are tight enough to be reached, as now shown in examples.

EXAMPLE 5.4. Take again the matrix defined in example 4.2. This matrix is of rank 4 and displacement rank 2. In addition, the displacement pattern satisfies $PZP = N$ as required in the theorem 5.1. With the notations of example 4.2, the MP pseudo inverse of A has displacement ranks $\delta^{\Delta}_{N,Z^*}\{A^-\} = 4$ and $\delta^{\nabla}_{N,Z}\{A^-\} = 4$. This is consistent with theorem 5.1.

EXAMPLE 5.5. Define the 5×5 Töplitz matrix of rank 3 :

$$A = \begin{pmatrix} 2 & 4 & 3 & 1 & 2 \\ 1 & 2 & 4 & 3 & 1 \\ 3 & 1 & 2 & 4 & 3 \\ 4 & 3 & 1 & 2 & 4 \\ 2 & 4 & 3 & 1 & 2 \end{pmatrix},$$

and assume as displacement pattern $Z = S_5$ and $N = S_5^*$. Then A has a displacement rank $\delta^{\nabla}_{Z,N} = 2$, and its MP pseudo-inverse has a displacement rank $\delta^{\nabla}_{N,Z}$ equal to 4. This result was expected, according to corollary 5.3.

EXAMPLE 5.6. If H is Hankel, then the displacement rank of H^- with respect to the displacement operator $\Delta_{S^*,S}$ is bounded by 5.

Other particular examples can be found in [7] and [8]. Let us now switch to the case of rectangular and rank-deficient structured matrices. In order to extend corollary 3.9, we need a variant of the inversion lemma:

LEMMA 5.7. *Let M be the block matrix:*

$$M = \begin{pmatrix} P & A_2 \\ A_1 & 0 \end{pmatrix},$$

where P is square invertible, and where A_1 and A_2 have the same rank. Then the MP-pseudo-inverse of M is:

$$M^- = \begin{pmatrix} Y & -P^{-1}A_2X \\ -XA_1P^{-1} & X \end{pmatrix},$$

where $X = -(A_1P^{-1}A_2)^-$, and $Y = P^{-1} + P^{-1}A_2XA_1P^{-1}$.

Proof. Let $A_i = U_iD_iV_i^*$ denote the SVD of A_i. Then define the matrix

$$\bar{M} = \begin{pmatrix} U_2^* & 0 \\ 0 & U_1^* \end{pmatrix} M \begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix},$$

and apply the usual inversion lemma to the invertible square portion of $\bar{M}$, denoted B. In other words we have:

$$\bar{M} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \text{ and } \bar{M}^- = \begin{pmatrix} B^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

The last lines of the proof are then just obvious manipulations. □

COROLLARY 5.8. *Let A be an $m \times n$ rectangular matrix with $m > n$. Then the displacement rank of its MP pseudo inverse verifies*

$$\delta^{\Delta}_{N,Z}\{A^-\} \leq 3\delta^{\Delta}_{N,Z}\{A^*\} + 2\delta^{\Delta}_{Z,N}\{A\} + 2\delta^{\nabla}_{N,N^*}\{I_m\}.$$

Proof. Write A^- as $(A^*A)^-A^*$, apply theorem 5.1 to the square matrix (A^*A), and then apply the product rule given in corollary 3.9. □

6. Concluding remarks. In this paper various aspects of the displacement rank concept were addressed in a rather general framework. In particular, displacement properties of rank-deficient matrices were investigated. However the bounds given in corollaries 3.9 and 5.8 are obviously too large. It is suspected that corollary 5.8 could be improved to $\delta\{B\} \leq 2\delta\{A\} + \delta\{I\}$ in most cases. On the other hand, particular examples have been found showing that the bounds given in other theorems are indeed reached (in particular theorems 5.1 and 5.2).

Another major limitation of this work lies in the fact that our proofs are in general not constructive, in the sense that they do not define suitable algorithms having the expected complexity. This is now the next question to answer. First ideas in this direction can be found in [4] and [14] and could be used for this purpose.

The author thanks Georg Heinig for his proofreading of the paper.

REFERENCES

[1] G. Ammar and W.B. Gragg, Superfast solution of real positive definite Töplitz systems, *SIAM Journal Matrix Analysis.,* vol.9, jan 1988, 61-76.

[2] A.W. Bojanczyk, R.P. Brent, and F.R. DeHoog, QR factorization of Töplitz matrices, *Numerische Mathematik,* vol.49, 1986, 81-94.

[3] R. Chan and G. Strang, Töplitz equations by conjugate gradients with circulant preconditioner, *SIAM Jour. Sc. Stat. Comput.* vol.10, jan 1989, 104-119.

[4] J. Chun, Fast array algorithms for structured matrices, *PhD thesis,* Stanford University, June 1989.

[5] J. Chun, T. Kailath, and H. Lev-Ari, Fast parallel algorithms for QR and triangular factorization, *SIAM Jour. Sci. Stat. Comput.,* vol.8, nov 1987, 899-913.

[6] J. Chun and T. Kailath, Displacement structure for Hankel- and Vandermonde-like matrices, *Signal Processing Part I: Signal Processing Theory, IMA Volumes in Mathematics and its Applications,* vol. 22, Springer Verlag, 1990 pp. 37–58.

[7] P. Comon and P. Laurent-Gengoux, Displacement rank of generalized inverses of persymmetric matrices, *Thomson Sintra report,* 90-C570-191, October 1990, to appear in *SIAM Journal Matrix Analysis.*

[8] P. Comon, Displacement rank of pseudo-inverses, *IEEE Int. Conf. ICASSP,* march 1992, San Francisco, vol.V, 49-52.

[9] P. Delsarte, Y.V. Genin, and Y.G. Kamp, A generalization of the Levinson algorithm for hermitian Töplitz matrices with any rank profile, *IEEE Trans ASSP,* vol.33, aug 1985, 964-971.

[10] K. Diepold and R. Pauli, Schur parametrization of symmetric matrices with any rank profile, *IEEE Int. Conf. ICASSP,* march 1992, San Francisco, vol.V, 269-272.

[11] B. Friedlander, M. Morf, T. Kailath, and L. Ljung, New inversion formulas for matrices classified in terms of their distance from Töplitz matrices, *Linear Algebra Appl.,* vol.27, 1979, 31-60.

[12] G.H. Golub and C.F. Van Loan, *Matrix computations,* Hopkins, 1983.

[13] C. Gueguen, An introduction to displacement ranks, *Signal processing XLV,* Lacoume, Durrani, Stora editors, Elsevier, 1987, 705-780.

[14] G. Heinig and K. Rost, Algebraic methods for Töplitz-like matrices and Operators, *Birkhäuser,* 1984.

[15] R.D. Hill, R.G. Bates, and S.R. Waters, On perhermitian matrices, *SIAM Journal Matrix Analysis,* April 1990, pp. 173–179.

[16] T. Kailath, A. Viera, and M. Morf, Inverses of Töplitz operators, innovations, and orthogonal polynomials, *SIAM Review,* 20, 1978, pp. 106–119.

[17] T. Kailath, Signal processing applications of some moment problems, *Proceedings of Symposia in Applied Mathematics, American Mathematical Society,* vol.37, 1987, pp. 71–109.

[18] T. Kailath, Remarks on the origin of the displacement-rank concept, *Applied Math. Comp.,* 45, 1991, pp. 193–206.

[19] S. Pombra, H. Lev-Ari, and T. Kailath, Levinson and Schur algorithms for Töplitz matrices with singular minors, *Int. Conf. ICASSP,* april 1988, New York, 1643-1646.

[20] D. Wood, Extending four displacement principles to solve matrix equations, submitted to *Math. Comp.,* preprint April 1992.

Published in: *Linear Algebra for Signal Processing*, A. Bojanczyk and G. Cybenko editors, vol.69, *IMA volumes in Mathematics and its Applications*, pp.1-16, Springer Verlag, 1995.