Displacement Structure of Weighted Pseudoinverses *

Jianfeng Cai¹ Yimin Wei²

¹ Institute of Mathematics, Fudan University, Shanghai, 200433, PR China.

² Department of Mathematics, Fudan University, Shanghai, 200433,PR China. E-mail: ymwei@fudan.edu.cn

Abstract

Estimates for the rank of $A_{MN}^{\dagger}V - UA_{MN}^{\dagger}$ and more general displacement of A_{MN}^{\dagger} are presented, where A_{MN}^{\dagger} is the weighted pseudoinverse of a matrix A. The results are applied to the close-to-Toeplitz, close-to-Vandermonde and close-to-Cauchy matrices. We extend the results due to G. Heinig and F. Hellinger in 1994.

Keywords:displacement, weighted pseudoinverse, structured matrix AMS subject classification:15A09,65F20

1 Introduction

The modern study of structured matrices was largely motivated by [6, 7], in which a basic concept of the displacement rank was introduced. If the rank of a matrix's displacement is small, fast algorithms for the matrix are available. If the UV-displacement of matrix A fulfills a Sylvester equation

AU - VA = E,

we call it Sylvester UV-displacement. If it fulfills a Stein equation

$$A - VAU = E,$$

we call it *Stein UV-displacement*. For example, a Toeplitz matrix T have Stein displacement $T - Z_m^* T Z_n = E$ and the rank of the displacement is 2, so fast algorithms have been constructed. For Sylvester and Stein displacement, if A is invertible, it is easy to show that the

 $^{^{*}{\}rm This}$ project is supported by National Natural Science Foundation of China and Doctoral Point Foundation of China.

rank of VU-displacement of A^{-1} is same to the rank of UV-displacement of A. In [2, 3, 18] the VU-displacement of pseudoinverse and group inverse were discussed if A is not invertible. In our paper, we are interested in generalized inverses with as small as possible rank.

We present the representation of the weighted pseudoinverse A_{MN}^{\dagger} of the matrix A. This means we study the so-called displacement $A_{MN}^{\dagger}V - UA_{MN}^{\dagger}$ or $A_{MN}^{\dagger} - UA_{MN}^{\dagger}V$ for structured matrix A. The detail of the weighted pseudoinverse and weighted linear least squares solution can be found in [4, 5, 9, 11-17].

Let $M \in C^{m \times m}, N \in C^{n \times n}$ be Hermitian positive definite matrices. We define the weighted inner product in C^m and C^n :

$$(x,y)_M = y^* M x, \quad x, y \in C^m,$$

 $(x,y)_N = y^* N x, \quad x, y \in C^n.$

So, the weighted conjugate transpose matrix $A^{\sharp} = N^{-1}A^*M$ of $A \in C^{m \times n}$ can be defined by

$$(Ax, y)_M = (x, A^{\sharp}y)_N \quad \forall x \in C^m, y \in C^n.$$

The weighted pseudoinverse of $A \in C^{m \times n}$ is the unique solution $A_{MN}^{\dagger}[1]$ of the following four equations:

$$AA_{MN}^{\dagger}A = A, \quad A_{MN}^{\dagger}AA_{MN}^{\dagger} = A_{MN}^{\dagger}, \quad (AA_{MN}^{\dagger})^{\sharp} = AA_{MN}^{\dagger}, \quad (A_{MN}^{\dagger}A)^{\sharp} = A_{MN}^{\dagger}A.$$

From the weighted singular value decomposition (SVD)[10], we know that for any $m \times n$ matrix A with rank r, there exist two weighted unitary matrices $R \in C^{m \times m}, S \in C^{n \times n}$ such that

$$A = R \begin{pmatrix} \Sigma & 0\\ 0 & 0 \end{pmatrix} S^*, \tag{1}$$

and $R^*MR = I_m, S^*N^{-1}S = I_n$, where $\Sigma = diag(\sigma_1, \sigma_2, \dots, \sigma_r)$ and $\sigma_1 \ge \sigma_2 \ge \dots \sigma_r > 0$ are nonzero eigenvalues of $A^{\sharp}A$. The weighted pseudoinverse A_{MN}^{\dagger} has an explicit expression given by:

$$A_{MN}^{\dagger} = N^{-1} S \begin{pmatrix} \Sigma^{-1} & 0\\ 0 & 0 \end{pmatrix} R^* M.$$
 (2)

We introduce

$$Q_{(MN)} \equiv A_{MN}^{\dagger}A, \quad Q_{(MN)*} \equiv AA_{MN}^{\dagger}, \quad P_{(MN)} \equiv I_n - Q_{(MN)}, \quad P_{(MN)*} \equiv I_m - Q_{(MN)*}.$$

It is easy to show that

$$Im(Q_{(MN)}) = Im(A^{\sharp}) = Im(A_{MN}^{\dagger}), \quad Im(Q_{(MN)*}) = Im(A);$$
$$Im(P_{(MN)}) = Ker(A), \quad Im(P_{(MN)*}) = Ker(A_{MN}^{\dagger}) = Ker(A^{\sharp}).$$

2 Displacement structure for Sylvester displacement

At first, we discuss the Sylvester displacement structure of the weighted pseudoinverse A_{MN}^{\dagger} of $A \in C^{m \times n}$.

Proposition 2.1. Let $A \in C^{m \times n}, U \in C^{n \times n}$ and $V \in {}^{m \times m}$, then

$$A_{MN}^{\dagger}V - UA_{MN}^{\dagger} = A_{MN}^{\dagger}VP_{(MN)*} - P_{(MN)}UA_{MN}^{\dagger} - A_{MN}^{\dagger}(AU - VA)A_{MN}^{\dagger}.$$
 (3)

Proof. Since

$$A_{MN}^{\dagger}(AU - VA)A_{MN}^{\dagger} = (I - P_{(MN)})UA_{MN}^{\dagger} - A_{MN}^{\dagger}V(I - P_{(MN)*}),$$

we can immediately draw the conclusion.

Lemma 2.1. The VU-displacement rank of A_{MN}^{\dagger} satisfies the following estimate: $rank(A_{MN}^{\dagger}V - UA_{MN}^{\dagger}) \leq rank(Q_{(MN)*}VP_{(MN)*}) + rank(P_{(MN)}UQ_{(MN)}) + rank(AU - VA).$ (4)

Proof.

$$rank(A_{MN}^{\dagger}VP_{(MN)*}) = rank(A_{MN}^{\dagger}VP_{(MN)*})^{\sharp} = dim[P_{(MN)*}V^{\sharp}Im(A_{MN}^{\dagger})^{\sharp}]$$

$$= dim[P_{(MN)*}V^{\sharp}Im(Q_{(MN)*})] = rank(Q_{(MN)*}VP_{(MN)*})^{\sharp}$$

$$= rank(Q_{(MN)*}VP_{(MN)*})$$

$$rank(P_{(MN)}UA_{MN}^{\dagger}) = dim[P_{(MN)}UIm(A_{MN}^{\dagger})] = dim[P_{(MN)}UIm(Q_{(MN)})]$$
$$= rank(P_{(MN)}UQ_{(MN)})$$

Taking these two into account, we obtain (4).

Proposition 2.2. Let $A \in C^{m \times n}, U \in C^{n \times n}$ and $V \in {}^{m \times m}$, then

$$rank(P_{(MN)}UQ_{(MN)}) + rank(Q_{(MN)*}VP_{(MN)*}) \le rank(AU^{\sharp} - V^{\sharp}A),$$
(5)

where $U^{\sharp} = N^{-1}U^*N$ and $V^{\sharp} = M^{-1}V^*M$.

Proof. We set $F \equiv AU^{\sharp} - V^{\sharp}A$.Let

$$A = R \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} S^*$$

be the weighted SVD of A.Partition

$$R^* M V^{\sharp} R = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}, \quad S^* U^{\sharp} N^{-1} S = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

where $U_{11}, V_{11} \in C^{r \times r}$ and r = rank(A). Therefore,

$$R^*MFN^{-1}S = \begin{pmatrix} \Sigma U_{11} - V_{11}\Sigma & \Sigma U_{12} \\ V_{21}\Sigma & 0 \end{pmatrix}.$$

Since

$$Q_{(MN)}U^{\sharp}P_{(MN)} = N^{-1}S\begin{pmatrix} 0 & U_{12} \\ 0 & 0 \end{pmatrix}S^{*}$$

and

$$P_{(MN)*}V^{\sharp}Q_{(MN)*} = R\begin{pmatrix} 0 & 0\\ V_{21} & 0 \end{pmatrix} R^*M_{21}$$

it follows from [8] that

$$rank(F) = rank(R^*MFN^{-1}S)$$

$$\geq rank(\Sigma U_{12}) + rank(V_{21}\Sigma)$$

$$= rank(U_{12}) + rank(V_{21})$$

$$= rank(Q_{(MN)}U^{\sharp}P_{(MN)}) + rank(P_{(MN)*}V^{\sharp}Q_{(MN)*})$$

$$= rank(P_{(MN)}UQ_{(MN)}) + rank(Q_{(MN)*}VP_{(MN)*}).$$

L		

From Proposition 2.1,2.2 and Lemma 2.1, we conclude that

Theorem 2.1. Let $A \in C^{m \times n}$ and A_{MN}^{\dagger} its weighted pseudoinverse. Then

$$rank(A_{MN}^{\dagger}V - UA_{MN}^{\dagger}) \le rank(AU - VA) + rank(AU^{\sharp} - V^{\sharp}A).$$
⁽⁶⁾

Corollary 2.1. If U, V are both weighted self-adjoint $U = U^{\sharp}, V = V^{\sharp}$ or weighted unitary $U^{-1} = U^{\sharp}, V^{-1} = V^{\sharp}$, then

$$rank(A_{MN}^{\dagger}V - UA_{MN}^{\dagger}) \le 2 \ rank(AU - VA).$$
⁽⁷⁾

3 Displacement structure for generalized displacement

In order to generalize Theorem 2.1 we introduce a generalized displacement concept[2].Let $a = [a_{ij}]_0^1$ denote a nonsingular 2 × 2 matrix. We associate a with the polynomial in two variables

$$a(\lambda,\mu) = \sum_{i,j=0}^{1} a_{ij}\lambda^{i}\mu^{j}$$

and the linear fractional function

$$f_a(\lambda) = \frac{a_{10} + a_{11}\lambda}{a_{00} + a_{01}\lambda}.$$

For any fixed $U \in C^{n \times n}$ and $V \in C^{m \times m}$, the generalized (a, U, V) displacement of $A \in C^{m \times n}$ generated by $a(\lambda, \mu)$ is defined by

$$a(V,U)A = \sum_{i,j=0}^{1} a_{ij}V^{i}AU^{j}.$$

 \mathbf{If}

$$a = d \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

we just get Sylvester displacement that we have discussed in Section 2.If

$$a = d \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

we get Stein displacement.

Lemma 3.1.[2] Let $a = [a_{ij}]_{0}^{1}, b = [b_{ij}]_{0}^{1}, c = [c_{ij}]_{0}^{1}, d = [d_{ij}]_{0}^{1}$ be nonsingular 2 × 2 matrices such that

$$a = b^T dc, (8)$$

then

$$(b_{00} + b_{01}\lambda)^{-1}a(\lambda,\mu)(c_{00} + c_{01}\mu)^{-1} = d(f_b(\lambda), f_c(\mu)).$$
(9)

Lemma 3.2.[2] Let $d = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, then there exist 2×2 matrices b, c such that (8) holds and $b_{00} + b_{01}V$ and $c_{00} + c_{01}U$ are invertible.

Taken Lemma 3.1 and Lemma 3.2 together, we obtain the following

Proposition 3.1.[2] Let b and c be matrices satisfying the conditions in Lemma 3.2, then for $A \in C^{m \times n}$,

$$a(V,U)A = (b_{00} + b_{01}V)[Af_c(U) - f_b(V)A](c_{00} + c_{01}U).$$

The following is very important to generalize Theorem 2.1 for general (a, U, V) displacement.

Proposition 3.2.

(a) If $\phi = [\phi_{ij}]_0^1$ is nonsingular and $\phi_{00} + \phi_{01}U$ is invertible, then

$$rank(P_{(MN)}UQ_{(MN)}) = rank(P_{(MN)}UQ_{(MN)}),$$

where $\tilde{U} \equiv f_{\phi}(U)$.

(b) If $\psi = [\psi_{ij}]_0^1$ is nonsingular and $\psi_{00} + \psi_{01}V$, then

$$rank(Q_{(MN)*}VP_{(MN)*}) = rank(Q_{(MN)*}VP_{(MN)*}),$$

where $\tilde{V} \equiv f_{\psi}(V)$.

Proof. We define

$$\mathcal{S} = Ker(A) \cap Ker(AU^{\sharp})$$
, $\mathcal{S}_1 = Ker(A) \ominus \mathcal{S}$.

We show that $Q_{(MN)}U^{\sharp}$ is one-to-one on S_1 . If $Q_{(MN)}U^{\sharp}x = 0$ and $x \in S_1$, then $U^{\sharp}x \in Ker(Q_{(MN)}) = Ker(A)$. That means $AU^{\sharp}x = 0$. Noting that $x \in Ker(A)$, we conclude $x \in S$. Thus x = 0.

Furthermore, $Q_{(MN)}U^{\sharp}x = 0$ for all $x \in \mathcal{S}$. Hence

$$rank(P_{(MN)}UQ_{(MN)}) = rank(Q_{(MN)}U^{\sharp}P_{(MN)}) = dim(\mathcal{S}_1).$$

$$(10)$$

Analogously we define

$$\tilde{\mathcal{S}} = Ker(A) \cap Ker(A\tilde{U}^{\sharp}) \ , \ \tilde{\mathcal{S}}_1 = Ker(A) \ominus \tilde{\mathcal{S}},$$

and we will get

$$rank(P_{(MN)}\tilde{U}Q_{(MN)}) = rank(Q_{(MN)}\tilde{U}^{\sharp}P_{(MN)}) = dim(\tilde{\mathcal{S}}_1).$$
(11)

Now we show that the invertible matrix $\overline{\phi}_{00} + \overline{\phi}_{01}U^{\sharp}$ bijectively maps \mathcal{S} onto $\tilde{\mathcal{S}}$. Suppose that $x \in \mathcal{S}$. Then $x, U^{\sharp}x \in Ker(A)$. Hence $y \equiv (\overline{\phi}_{10} + \overline{\phi}_{11}U^{\sharp})x$ and $z \equiv (\overline{\phi}_{00} + \overline{\phi}_{01}U^{\sharp})x$ are all contained in Ker(A). Thus $y = \tilde{U}^{\sharp}z$ and we conclude that $z, \tilde{U}^{\sharp}z \in Ker(A)$, which implies $z \in \tilde{\mathcal{S}}$. Convesely, with the same arguments we get $(\overline{\phi}_{00} + \overline{\phi}_{01}U^{\sharp})^{-1}z \in \mathcal{S}$ for all $z \in \tilde{\mathcal{S}}$.

This implies

$$dim(\mathcal{S}_1) = dim[Ker(A)] - dim(\mathcal{S}) = dim[Ker(A)] - dim(\tilde{\mathcal{S}}) = dim(\tilde{\mathcal{S}}_1).$$

According to (10) and (11), we get assertion (a).

Assertion (b) is proved analogously.

Now we can generalize Theorem 2.1 for general (a, U, V) displacement.

Theorem 3.1. Let a, b be 2×2 nonsingular matrices, then

$$rank[a(U,V)A_{MN}^{\dagger})] \le rank[a^{T}(V,U)A] + rank[b(V^{\sharp},U^{\sharp})A].$$
(12)

Proof. According to Lemma 3.2 there exist 2×2 matrices w, x, y, z such that $w_{00} + w_{01}U$, $x_{00} + x_{01}V$, $y_{00} + y_{01}U$, $z_{00} + z_{01}V$ are invertible and

$$a = w^T dz,$$
$$b = x^* d\overline{y}.$$

Hence,

$$\begin{aligned} \operatorname{rank}[a(U,V)A_{MN}^{\dagger}] - \operatorname{rank}[a^{T}(V,U)A] \\ &= \operatorname{rank}[f_{w}(U)A_{MN}^{\dagger} - A_{MN}^{\dagger}f_{z}(V)] - \operatorname{rank}[f_{z}(V)A - Af_{w}(U)] \\ &\leq \operatorname{rank}[P_{(MN)}f_{w}(U)Q_{(MN)}] + \operatorname{rank}[Q_{(MN)*}f_{z}(V)P_{(MN)*}] \\ &= \operatorname{rank}[P_{(MN)}f_{y}(U)Q_{(MN)}] + \operatorname{rank}[Q_{(MN)*}f_{x}(V)P_{(MN)*}] \\ &\leq \operatorname{rank}[f_{\overline{x}}(V^{\sharp})A - Af_{\overline{y}}(U^{\sharp})] \\ &= \operatorname{rank}[b(V^{\sharp}, U^{\sharp})A]. \end{aligned}$$

Corollary 3.1. If U, V are weighted unitary or weighted self-adjoint matrix, then

$$rank[a(U,V)A_{MN}^{\dagger}] \le 2 \ rank[a^{T}(V,U)A].$$

$$\tag{13}$$

Proof. Let

$$b = \begin{cases} a^T & \text{if } U^{\sharp} = U, V^{\sharp} = V, \\ ia^T & \text{if } U^{\sharp} = U, V^{\sharp} = V^{-1}, \\ a^T i & \text{if } U^{\sharp} = U^{-1}, V^{\sharp} = V, \\ ia^T i & \text{if } U^{\sharp} = U^{-1}, V^{\sharp} = V^{-1}, \end{cases}$$

where *i* denote the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. We immediately obtain (13) from Theorem 3.1. \Box

4 Computation of the displacement

For practical purpose it is important to know not only the displacement rank of A_{MN}^{\dagger} but the explicit form of the displacement. For simplicity we restrict our explanation to the case of Sylvester displacement and to the case of a matrix A.

Our starting point is (3).

(1) If we know the full-rank decomposition (For many structured matrices, it is easy to get the decomposition.) r

$$AU - VA = GF^* = \sum_{i=1}^{r} g_i f_i^*,$$

we only need to compute the weighted least squares solutions [1, 15] $A_{MN}^{\dagger}g_i$ and $f_i^*A_{MN}^{\dagger}$ to get

$$A_{MN}^{\dagger}(AU - VA)A_{MN}^{\dagger}.$$

(2) For the purpose to get $A_{MN}^{\dagger} V P_{(MN)*}$, we start with another full-rank decomposition

$$A^{\sharp}V - UA^{\sharp} = KL^*.$$

We denote $C \equiv \begin{bmatrix} A^* M^{1/2} \\ L^* M^{-1/2} \end{bmatrix}$. It is obvious that $Ker(C) \subseteq M^{-1/2}Ker(A^*)$, so we can find an orthonormal system of vectors w_1, \dots, w_p forming a basis of the orthogonal complement of Ker(C) in $M^{-1/2}Ker(A^*)$. If we introduce the matrix $W = [w_1, \dots, w_p]$, then we have

$$M^{-1/2}Ker(A^*) = Ker(C) \oplus Im(W).$$

Proposition 4.1. Let $R \equiv M^{-1/2}WW^*M^{1/2}$, then

$$RP_{(MN)*} = R. (14)$$

Proof. It is obvious that

$$Im(RQ_{(MN)*}) = RIm(Q_{(MN)*}) = RIm(A) = Im(RA) = Im(M^{-1/2}WW^*M^{1/2}A).$$

Noting that $Im(W) \subseteq M^{-1/2}Ker(A^*)$, we get

$$A^*M^{1/2}W = 0$$

Hence, $W^*M^{1/2}A = 0$. So, we conclude $RQ_{(MN)*} = 0$. Taking $P_{(MN)*} = I_m - Q_{(MN)*}$ into account, we obtain (14).

Proposition 4.2. Let *R* be defined as Proposition 4.1, then

$$A_{MN}^{\dagger} V(I_m - R) P_{(MN)*} = 0.$$
(15)

Proof. In view of $Im(P_{(MN)*}) = Ker(A^{\sharp})$, one can easily check

$$Im[A_{MN}^{\dagger}V(I_m - R)P_{(MN)*}] = A_{MN}^{\dagger}VM^{-1/2}(I_m - WW^*)M^{1/2}Ker(A^{\sharp})$$
$$= A_{MN}^{\dagger}VM^{-1/2}(I_m - WW^*)M^{-1/2}Ker(A^*).$$

For all $x \in M^{-1/2}Ker(A^*)$, there exists a unique decomposition x = y + z such that $y \in Ker(C)$ and $z \in Im(W)$. Noting WW^* is an orthogonal projection onto Im(W), we get $(I_m - WW^*)(y + z) = y \in Ker(C)$. Hence

$$Im[A_{MN}^{\dagger}V(I_m - R)P_{(MN)*}] \subseteq A_{MN}^{\dagger}VM^{-1/2}Ker(C).$$

$$(16)$$

Now we show $A_{MN}^{\dagger}VM^{1/2}Ker(C) = 0$. Surpose $x \in Ker(C)$. Then

$$A^*M^{1/2}x = 0$$
 and $L^*M^{-1/2}x = 0$.

Hence,

$$A^{\sharp}VM^{-1/2}x - UA^{\sharp}M^{-1/2}x = KL^*M^{-1/2}x = 0.$$

Thus,

$$A^{\sharp}VM^{-1/2}x = UA^{\sharp}M^{-1/2}x = UN^{-1}A^{*}M^{1/2}x = 0.$$

So, we have $VM^{-1/2}x \in Ker(A^{\sharp}) = Ker(A_{MN}^{\dagger})$. This means $A_{MN}^{\dagger}VM^{-1/2}Ker(C) = 0$.

Noting (16), we obtain (15).

According to Proposition 4.1 and 4.2, we have

$$A_{MN}^{\dagger} V P_{(MN)*} = A_{MN}^{\dagger} V M^{-1/2} W W^* M^{1/2}.$$

(3) We proceed analogously for $P_{(MN)}UA_{MN}^{\dagger}$. Let $C_* \equiv \begin{bmatrix} AN^{-1/2} \\ K^*N^{1/2} \end{bmatrix}$ and $S \equiv N^{-1/2}ZZ^*N^{1/2}$, where $Z = [z_1, \dots, z_q]$ and z_1, \dots, z_q is an orthonormal basis of the orthogonal complement of $Ker(C_*)$ in $Ker(AN^{-1/2})$. The result obtained is

$$P_{(MN)}UA_{MN}^{\dagger} = SUA_{MN}^{\dagger} = N^{-1/2}ZZ^*N^{1/2}UA_{MN}^{\dagger}.$$

In order to compute the displacement, one have to find 2r weighted least squares solutions $A_{MN}^{\dagger}g_i$ and $f_i^*A_{MN}^{\dagger}$, where r = rank(AU - VA), and p + q weighted least squares solutions $A_{MN}^{\dagger}VM^{-1/2}w_i$ $(i = 1, \dots, p)$ and $z_j^*N^{1/2}UA_{MN}^{\dagger}$ $(j = 1, \dots, q)$, where $p + q \leq rank(A^{\sharp}V - UA^{\sharp})$.

5 Full rank matrices

In this section we consider a special case that A has full rank. If this condition is fulfilled, then A_{MN}^{\dagger} has an explicit form given by $A_{MN}^{\dagger} = A^{\sharp} (AA^{\sharp})^{-1}$ or $A_{MN}^{\dagger} = (A^{\sharp}A)^{-1}A^{\sharp}$.

We show that under the assumption made above one can find a more general estimate for the displacement rank. In fact, in the case under consideration the $U^{\sharp}V^{\sharp}$ -displacement can be displaced by the $U^{\sharp}W^{\sharp}$ -displacement for arbitrary $W \in C^{m \times m}$, or by the $W^{\sharp}V^{\sharp}$ -displacement for arbitrary $W \in C^{n \times n}$. This is important for a series of applications.

We will start with Sylvester displacement as we have done in Section 2. Then we generalize it to the generalized displacement.

Proposition 5.1. Let $A, U, V, P_{(MN)}, P_{(MN)*}$ be defined as before and $W_1 \in C^{n \times n}, W_2 \in C^{m \times m}$ are arbitrary, then

$$\begin{aligned} A_{MN}^{\dagger}V - UA_{MN}^{\dagger} &= (A^{\sharp}A + P_{(MN)})^{-1}(A^{\sharp}V - W_{1}A^{\sharp})P_{(MN)*} \\ &+ P_{(MN)}(A^{\sharp}W_{2} - UA^{\sharp})(AA^{\sharp} + P_{(MN)*})^{-1} - A_{MN}^{\dagger}(AU - VA)A_{MN}^{\dagger}. \end{aligned}$$

Proof. By the weighted SVD, we have $P_{(MN)}A^{\sharp} = 0$ and $A^{\sharp}P_{(MN)*} = 0$. Hence,

$$P_{(MN)}A^{\sharp}W_2(AA^{\sharp} + P_{(MN)}^{*})^{-1} = 0$$
 and $(A^{\sharp}A + P_{(MN)})^{-1}W_1A^{\sharp}P_{(MN)^{*}} = 0.$

According to the weighted SVD, one can easily check that

$$A_{MN}^{\dagger} = (A^{\sharp}A + P_{(MN)})^{-1}A^{\sharp} = A^{\sharp}(AA^{\sharp} + P_{(MN)*})^{-1}.$$

Taking this into account and noting (3), we get the result.

We can immediately get the following two theorems through Proposition 5.1.

Theorem 5.1. Let $A \in C^{m \times n}$ be of row full rank and m < n, then for any arbitrary $W \in C^{m \times m}$,

$$rank(A_{MN}^{\dagger}V - UA_{MN}^{\dagger}) \le rank(AU - VA) + rank(AU^{\sharp} - W^{\sharp}A).$$

$$(17)$$

Theorem 5.2. Let $A \in C^{m \times n}$ be of column full rank and m > n, then for any arbitrary $W \in C^{n \times n}$,

$$rank(A_{MN}^{\dagger}V - UA_{MN}^{\dagger}) \le rank(AU - VA) + rank(AW^{\sharp} - V^{\sharp}A).$$
(18)

Now we turn to the generalized displacement.

Theorem 5.3. Let $A \in C^{m \times n}$ be of full rank and m < n, let a, b be nonsingular 2×2 matrices, then for any arbitrary $W \in C^{m \times m}$,

$$rank[a(U,V)A_{MN}^{\dagger}] \le rank[a^{T}(V,U)A] + rank[b(W^{\sharp},U^{\sharp})A].$$
⁽¹⁹⁾

Proof. Under the assumptions there exist 2×2 matrices w and z such that $a = w^T dz$ and the matrices $w_{00} + w_{01}U$ and $z_{00} + z_{01}V$ are invertible. Furthermore,

$$a(U,V)A_{MN}^{\dagger} = (w_{00} + w_{01}U)[A_{MN}^{\dagger}f_z(V) - f_w(U)A_{MN}^{\dagger}](z_{00} + z_{01}V)$$

together with

$$A_{MN}^{\dagger} f_z(V) - f_w(U) A_{MN}^{\dagger} = -P_{(MN)} f_w(U) Q_{(MN)} - A_{MN}^{\dagger} [A f_w(U) - f_z(V) A] A_{MN}^{\dagger}$$

implies

$$rank[a(U,V)A_{MN}^{\dagger}] \le rank[P_{(MN)}f_w(U)Q_{(MN)}] + rank[Af_w(U) - f_z(V)A].$$

Noting $a^T = -z^T dw$, we get

$$rank[Af_w(U) - f_z(V)A] = rank[a^T(V, U)A]$$

Under the assumptions there also exist 2×2 matrices x and y such that $a = x^T dy$ and the matrices $x_{00} + x_{01}W$ and $y_{00} + y_{01}U$ are invertible. Then

$$\begin{aligned} rank[P_{(MN)}f_w(U)Q_{(MN)}] &= rank[P_{(MN)}f_y(U)Q_{(MN)}] \\ &= rank[P_{(MN)}f_y(U)Q_{(MN)}] + rank[Q_{(MN)*}f_x(W)P_{(MN)*}] \\ &\leq rank[Af_{\bar{y}}(U^{\sharp}) - f_{\bar{x}}(W^{\sharp})A] \\ &= rank[b(W^{\sharp}, U^{\sharp})A]. \end{aligned}$$

The following theorem can be proved analogously.

Theorem 5.4. Let $A \in C^{m \times n}$ be of full column rank and m > n, let a, b be nonsingular 2×2 matrices, then for any arbitrary $W \in C^{n \times n}$,

$$rank[a(U,V)A_{MN}^{\dagger}] \le rank[a^{T}(V,U)A] + rank[b(V^{\sharp},W^{\sharp})A].$$
⁽²⁰⁾

6 Applications

In this section we apply the theorems proved before to many classical structured matrices, including Toeplitz, Hankel, Cauchy and Vandermonde matrices.

6.1 Close-to-Toeplitz matrices

Close-to-Toeplitz matrices are a class of matrices whose UV-displacement ranks are small compared with the sizes of the matrices for U and V being (forward or backward) (block) shifts, including Toeplitz, Hankel matrices, more general block matrices with Toeplitz or Hankel blocks, and sums, products, and inverses of these matrices.

We consider the case

$$U = Z_n \equiv \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \in C^{n \times n} , \quad V = Z_m^*$$

and let

$$r_+ \equiv rank(A - Z_m^* A Z_n)$$
, $r_- \equiv rank(A - Z_m A Z_n^*)$

Choosing

$$a = b = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

in Theorem 3.1 and noting

$$rank(A - (Z_m^*)^{\sharp}AZ_n^{\sharp}) = rank(MAN^{-1} - Z_mMAN^{-1}Z_n^*),$$

we obtain the following

Theorem 6.1.1. Let
$$r_{MAN^{-1}} = rank(MAN^{-1} - Z_mMAN^{-1}Z_n^*)$$
, then
 $rank(A_{MN}^{\dagger} - Z_nA_{MN}^{\dagger}Z_m^*) \le r_+ + r_{MAN^{-1}}$.

If the estimate of Theorem 6.1.1 is small for a close-to-matrices A, then it lead to the the famous representation formula of Gohberg-Semencul type of A_{MN}^{\dagger} :

$$A_{MN}^{\dagger} = \sum_{k=1}^{r} L_k U_k$$

where r is the displacement rank of A_{MN}^{\dagger} , i.e., $r_{+} + r_{MAN^{-1}}$ here.

The importance of the representations consists the fact that with their help weighted least squares solutions A_{MN}^{\dagger} for a close-to-Toeplitz matrix A can be computed with the complexity O((m+n)log(m+n)) if the FFT is applied.

Corollary 6.1.1.

$$rank(A^{\dagger} - Z_n A^{\dagger} Z_m^*) \le r_+ + r_-.$$

Corollary 6.1.2. Let $r_M = rank(M - Z_m M Z_m^*), r_N = rank(N - Z_n^* N Z_n)$, then

$$rank(A_{MN}^{\dagger} - Z_n A_{MN}^{\dagger} Z_m^*) \le 2r_+ + r_M + r_N.$$

Proof. By [6], we have

$$rank(N - Z_n^*NZ_n) = rank(N^{-1} - Z_nN^{-1}Z_n^*)$$

We immediately obtain the corollary by the following

$$MAN^{-1} - Z_m MAN^{-1}Z_n^*$$

= $(M - Z_m MZ_m^*)AN^{-1} + Z_m MZ_m^*A(N^{-1} - Z_n N^{-1}Z_n^*) - Z_m M(A - Z_m^*AZ_n)N^{-1}Z_n^*$

However, the estimate in Corollary 6.1.2 is not always small, but if we choose the weight matrices M,N such that r_M,r_N is very small compared to the size of A, then the rank of the VU-displacement of A^{\dagger}_{MN} is also very small. For example, let M,N be Hermite positive definite Toeplitz matrices, then the displacement rank of the weighted pseudoinverse for a Toeplitz matrix is less than or equal to 8 through Corollary 6.1.2.

6.2 Close-to-Vandermonde matrices

Let $D(c) = diag(c_1, \dots, c_m)$. It is well known that the displacement rank $r = rank[Van_n(c)Z_n - D(c)Van_n(c)]$ of a Vandermonde matrix $Van_n(c) = [c_i^{j-1}]_{i=1,j=1}^{m,n}$ is equal to one, except for the trivial case r = 0. Hence, an $m \times n$ matrix is said to be close-to-Vandermonde if, for certain $c \in C^m$, the displacement rank $rank[AZ_n - D(c)A]$ is small compared with m and n.

We denote

$$r \equiv rank[AZ_n - D(c)A], \quad r' \equiv rank[A - D(c)AZ_n^*]$$

It is easily to check that a close-to-Vandermonde matrix admits a representation

$$A = \sum_{i=1}^{r} D_i Van_n(c) T_i + D_0 Van_n(c),$$
(21)

where D_i are diagonal matrices and T_i are upper triangular Toeplitz matrices with zeros at the main diagonal.Note that the matrices D_i and T_i can be found via the full-rank decomposition of $AZ_n - D(c)A$, and D_0 is related to the first column of A.

With the representation (21), we can show $r' \leq r+1$ and in particular, if A is Vandermonde matrix, r' = r = 1.

Theorem 6.2.1. Let $r_{MAN^{-1}} = MAN^{-1} - D(c)MAN^{-1}Z_n^*$. Suppose that $c_i \in R$ or $|c_i| = 1$ for all $i = 1, \dots, m$, then

$$rank[A_{MN}^{\dagger}D(c) - Z_n A_{MN}^{\dagger}] \le r + r_{MAN^{-1}}.$$

Proof. 3.1.Since

$$rank[AZ_n^{\sharp} - D(c)^{\sharp}A] = rank[MAN^{-1}Z_n^* - D(c)^*MAN^{-1}]$$

Set $b = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ if $c_i \in R$ and set $b = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ if $|c_i| = 1$ in Theorem

and

$$rank[A - D(c)^{\sharp}AZ_{n}^{\sharp}] = rank[MAN^{-1} - D(c)^{*}MAN^{-1}Z_{n}^{*}],$$

we immediately get the result.

We give the following corollary.

Corollary 6.2.1. Suppose that $c_i \in R$ or $|c_i| = 1$ for all $i = 1, \dots, m$, then

$$rank[A^{\dagger}D(c) - Z_nA^{\dagger}] \le r + r' \le 2r + 1.$$

Theorem 6.2.2. Let $r_N = rank(N - Z_n^*NZ_n), r_M = rank[M - D(c)MD(c)^*]$. Suppose that $|c_i| = 1$ for all $i = 1, \dots, m$, then

$$rank[A_{MN}^{\dagger}D(c) - Z_n A_{MN}^{\dagger}] \le 2r + r_M + r_N.$$

Proof. Since

$$MAN^{-1} - D(c)MAN^{-1}Z_n^*$$

= $(M - D(c)MD(c)^*)AN^{-1} + D(c)MD(c)^*A(N^{-1} - Z_nN^{-1}Z_n^*)$
 $-D(c)M(A - D(c)^*AZ_n)N^{-1}Z_n^*,$

noting $D(c)^* = D(c)^{-1}$ and by [6], we have

$$rank(N - Z_n^*NZ_n) = rank(N^{-1} - Z_nN^{-1}Z_n^*),$$

we obtain the theorem.

Theorem 6.2.3. Let $r'_N = rank(NZ_n - Z_n^*N), r'_M = rank[MD(c) - D(c)M]$. Suppose that $c_i \in R$ for all $i = 1, \dots, m$, then

$$rank[A_{MN}^{\dagger}D(c) - Z_{n}A_{MN}^{\dagger}] \le 2r + r_{M}^{'} + r_{N}^{'}.$$

Proof. Set $b = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ in Theorem 3.1.Therefore, $rank[A_{MN}^{\dagger}D(c) - Z_n A_{MN}^{\dagger}] \leq r + rank[AZ_n^{\sharp} - D(c)^{\sharp}A]$ $= r + rank[MAN^{-1}Z_n^* - D(c)^*MAN^{-1}].$

Since

$$MAN^{-1}Z_n^* - D(c)^*MAN^{-1}$$

= $(MD(c) - D(c)^*M)AN^{-1} + MA(N^{-1}Z_n^* - Z_nN^{-1}) + M(AZ_n - D(c)A)N^{-1},$

noting $D(c)^* = D(c)$ and

$$rank(NZ_n - Z_n^*N) = rank(N^{-1}Z_n^* - Z_nN^{-1}),$$

we obtain the theorem.

6.3 Generalized Cauchy matrices

Let U, V be diagonal matrices,

$$U \equiv D(d) = diag(d_1, \cdots, d_n)$$
, $V \equiv D(c) = diag(c_1, \cdots, c_m)$.

A matrix A is said to be a generalized Cauchy matrix if for certain c and d, rank[AD(d) - D(c)A] is small compared with m and n.In case $c_i \neq d_j$ for all i and j, it has an explicit form

$$A = \left[\frac{f_i^* g_j}{c_i - d_j}\right]_{i=1,j=1}^{m,n} ,$$
 (22)

where $f_i, g_j \in C^r$ and r = rank[AD(d) - D(c)A].

In particular, if $r = 1, f_1 = g_1 = 1, A$ is classical Cauchy Matrix. If $f_1 = a = (a_1, \dots, a_m)^*, f_2 = (-1, \dots, -1)^*, g_1 = (1, \dots, 1)^*, g_2 = b = (b_1, \dots, b_n)^*, A$ is Loewner matrix, which has the form

$$A = \left[\frac{a_i - b_j}{c_i - d_j}\right]_{i=1,j=1}^{m,n} .$$

We assume that $c_i \in R$ or $|c_i| = 1$ for all *i*, and the same for d_j . In case $c \in R^m$, we have $D(c)^* = D(c)$; in case $|c_i| = 1$, we have $D(c)^* = D(c)^{-1}$.

Theorem 6.3.1. Let $r_{MAN^{-1}} = rank[MAN^{-1}D(d) - D(c)MAN^{-1}]$. Suppose that $c_i \in R$ or $|c_i| = 1$ for all $i = 1, \dots, m$ and $d_j \in R$ or $|d_j| = 1$ for all $j = 1, \dots, n$. Then

$$rank[A_{MN}^{\dagger}D(c) - D(d)A_{MN}^{\dagger}] \le r + r_{MAN^{-1}} ,$$

where r = rank[AD(d) - D(c)A].

Proof. Set

$$b = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} & \text{if } c_i \in R, d_i \in R \text{ or } |c_i| = 1, |d_i| = 1 \\ \text{if } c_i \in R, |d_i| = 1 \text{ or } |c_i| = 1, d_i \in R \end{cases}$$

in Theorem 3.1, we immediately obtain the theorem.

Corollary 6.3.1. If $c_i \in R$ or $|c_i| = 1$ for all *i*, and the same for d_j . Then

$$rank[A^{\dagger}D(c) - D(d)A^{\dagger}] \le 2r$$

where r = rank[AD(d) - D(c)A].

Corollary 6.3.2. Let $r_M = rank[MD(c) - D(c)M], r_N = rank[ND(d) - D(d)N]$. If $c_i \in R$ or $|c_i| = 1$ for all *i*, and the same for d_j , then

$$rank[A_{MN}^{\dagger}D(c) - D(d)A_{MN}^{\dagger}] \le 2r + r_M + r_N,$$

where r = rank[AD(d) - D(c)A].

Proof. Since

$$MAN^{-1}D(d) - D(c)MAN^{-1}$$

= $MA(N^{-1}D(d) - D(d)N^{-1}) + M(AD(d) - D(c)A)N^{-1} - (D(c)M - MD(c))AN^{-1}$

and

$$rank[N^{-1}D(d) - D(d)N^{-1}] = rank[ND(d) - D(d)N],$$

we obtain

$$r_{MAN^{-1}} \le r + r_N + r_M.$$

With Theorem 6.3.1, we conclude that A_{MN}^{\dagger} has the form (22). It indicates that one can construct a fast algorithm to compute the weighted least squares solution $x = A_{MN}^{\dagger} y$ for a generalized Cauchy system if $r + r_{MAN^{-1}}$ is small.

6.4 Upper bound for full rank matrices

In the above subsections, we give displacement estimates of A by the displacement of weight matrices M and N. In this subsection, we give an upper bound of the displacement ranks independent of the displacement of weight matrices. At first, we consider the case that A has full column rank and the Sylvester displacement.

Our start point is (3). Assume m > n. Because A has full column rank,
so (3) can be written into

$$A_{MN}^{\dagger}V - UA_{MN}^{\dagger} = A_{MN}^{\dagger}VP_{(MN)*} - A_{MN}^{\dagger}(AU - VA)A_{MN}^{\dagger}.$$

 Set

$$K \equiv A^{\dagger}_{MN} V P_{(MN)*}$$

We consider the system $K^{\sharp}x = 0$. This system can be written into the form

$$P_{(MN)*}V^{\sharp}(A^{\dagger}_{MN})^{\sharp}x = 0.$$

Since $Im[(A_{MN}^{\dagger})^{\sharp}] = Im(A)$, we get

$$rank(P_{(MN)*}V^{\sharp}A) = rank[P_{(MN)*}V^{\sharp}(A_{MN}^{\dagger})^{\sharp}].$$

Hence, the dimension of the solution space of $K^{\sharp}x = 0$ is equivalent to the dimension of the solution space of the following

$$P_{(MN)*}V^{\sharp}Ax = 0. \tag{23}$$

Now we consider another equation

$$(V^{\sharp}A - AW)x = 0 \quad , \quad W \in C^{n \times n}.$$

If x is a solution of (24), then

$$V^{\sharp}Ax = AWx \in Im(A) = Ker(P_{(MN)*}).$$

It means the solution space of (24) is a subset of solution space of (23). Therefore, if we give a lower bound of the solution space's dimension of (24), let it be d, then n - d is the upper bound of rank(K). In fact, if we set $W = A^{\dagger}V^{\sharp}A$, (24) is changed into

$$(I_m - AA^{\dagger})V^{\sharp}Ax = R \begin{bmatrix} 0 & 0\\ 0 & I_{m-n} \end{bmatrix} R^*V^{\sharp}Ax = 0.$$

Then we obtain d = n - (m - n) = 2n - m, therefore,

$$rank(K) \le n - (2n - m) = m - n.$$

We generalize it to the generalized displacement $a(U, V)A_{MN}^{\dagger}$. By Lemma 3.2, there exist nonsingular 2×2 matrices w and z such that $a = w^T dz$ and $w_{00} + w_{01}U$ and $z_{00} + z_{01}V$ are invertible. Then

$$a(U,V)A_{MN}^{\dagger} = (w_{00} + w_{01}U)[A_{MN}^{\dagger}f_z(V) - f_w(U)A_{MN}^{\dagger}](z_{00} + z_{01}V).$$

Noting (3) and $P_{(MN)} = 0$, we obtain

$$A_{MN}^{\dagger}f_{z}(V) - f_{w}(U)A_{MN}^{\dagger} = A_{MN}^{\dagger}f_{z}(V)P_{(MN)*} - A_{MN}^{\dagger}[Af_{w}(U) - f_{z}(V)A]A_{MN}^{\dagger}.$$

Because $rank[Af_w(U) - f_z(V)A] = rank[a^T(V, U)A]$,so

$$rank[a(U,V)A_{MN}^{\dagger}] \leq rank[A_{MN}^{\dagger}f_{z}(V)P_{(MN)*}] + rank[a^{T}(V,U)A].$$

Now we estimate the upper bound of $rank[A^{\dagger}_{MN}f_z(V)P_{(MN)*}]$. In fact,

$$rank[A_{MN}^{\dagger}f_{z}(V)P_{(MN)*}] = rank[Q_{(MN)*}f_{z}(V)P_{(MN)*}] = rank[Q_{(MN)*}VP_{(MN)*}] = rank(K).$$

The same as the case for A has full row rank and m < n, we obtain

$$rank(P_{(MN)}UA_{MN}^{\dagger}) \le n-m.$$

Theorem 6.4.1. Let $A \in C^{m \times n}$ be of full rank, then

$$rank[a(U,V)A_{MN}^{\dagger}] \le min\{m,n,rank[a^{T}(V,U)A] + |m-n|\}.$$

The upper bound can be attained. For example,let T be a 20×12 full column rank Toeplitz matrix,

$$a = (3, 2, 3, 4, 5, 1, 2, 3, 5, 3, 2, 7)$$

be the first row of T and

$$b = (3, 4, 2, 3, 4, 6, 2, 5, 3, 4, 5, 6, 1, 2, 3, 6, 7, 8, 3, 4)^T$$

be the first column of T. The weight matrices

$$M = diag(1, 2, 7, 5, 6, 2, 4, 3, 4, 6, 4, 8, 4, 2, 2, 5, 6, 2, 5, 4),$$

and

$$N = diag(4, 1, 1, 5, 3, 6, 4, 8, 7, 8, 4, 5).$$

We obtain

$$rank(T_{MN}^{\dagger} - Z_n T_{MN}^{\dagger} Z_m^*) = 10 = rank(T - Z_m^* T Z_n) + m - n$$

References

- A. Ben-Israel and T.N.E. Greville, Generalized Inverses: Theory and Applications, Wiley, New York, 1974.
- [2] G. Heinig and F. Hellinger, Displacement structure of pseudoinverses, *Linear Algebra Appl.*, 197/198 (1994) 623-649.
- [3] G. Heinig and F. Hellinger, Displacement structure of generalized inverse matrices, *Linear Algebra Appl.*, 211 (1994) 67-83.
- [4] M.E. Gulliksson, Xiaoqing Jin and Yimin Wei, Perturbation bound for constrained and weighted least squares problem, *Linear Algebra Appl.*, 349 (2002) 221-232.
- [5] M.E. Gulliksson, P.A. Wedin and Yimin Wei, Perturbation identities for regularized Tikhonov inverses and weighted pseudoinverse, *BIT*, 40 (2000) 513-523.
- [6] T. Kailath, S.Y. Kung and M. Morf, Displacement rank of matrices and linear equations, J. Math. Anal. Appl., 68 (1979) 395-407.
- [7] T. Kailath and A. Sayed, Displacement structure: Theory and Applications, SIAM Review, 37 (1995) 297-386.
- [8] G. Marsaglia and G.P.H. Styan, Equalities and inequalities for ranks of matrices, *Linear and Multilinear Algebra*, 2 (1974) 269-292.

- [9] V. Pan and R. Schreiber, An improved Newton iteration for the generalized inverse of a matrix with applications, SIAM J. on Sci. and Stat. Comput., 12 (1991) 1109-1131.
- [10] C.F. Van Loan, Generalizing the singular value decomposition, SIAM J. Numer. Anal., 13 (1976) 76-83.
- [11] Yimin Wei and Hebing Wu, Expression for the perturbation of the weighted Moore-Penrose inverse, Comput. Math. Appl., 39 (2000) 13-18.
- [12] Yimin Wei, Recurrent neural networks for computing weighted Moore-Penrose inverse, Appl. Math. Comput., 116 (2000) 279-287.
- [13] Yimin Wei, Hebing Wu and Junyin Wei, Successive matrix squaring algorithm for parallel computing the weighted generalized inverse A_{MN}^+ , Appl. Math Comput., 116 (2000) 289-296.
- [14] Yimin Wei and Hebing Wu, The representation and approximation for the weighted Moore-Penrose inverse, Appl. Math. Comput., 121 (2001) 17-28.
- [15] Yimin Wei and Guorong Wang, PCR algorithm for parallel computing minimumnorm(T) least-squares(S) solution of inconsistent linear linear equations, Appl. Math. Comput., 133 (2002) 547-557.
- [16] Yimin Wei and Dingkun Wang, Condition numbers and perturbation of the weighted Moore-Penrose inverse and weighted linear least squares problems, *Appl. Math. Comput.*, to appear.
- [17] Yimin Wei and Michael Ng, Weighted Tikhonov filter matrices for ill-posed problems, submitted.
- [18] Yimin Wei and Michael Ng, Displacement structure of group inverses, submitted.