

EQUIVALENT FORMULAE FOR THE SUPREMUM AND STABILITY OF WEIGHTED PSEUDOINVERSES

MUSHENG WEI

ABSTRACT. During recent decades, there have been a great number of research articles studying interior-point methods for solving problems in mathematical programming and constrained optimization. Stewart and O’Leary obtained an upper bound for scaled pseudoinverses $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|_2$ of a matrix X where \mathcal{P} is a set of diagonal positive definite matrices. We improved their results to obtain the supremum of scaled pseudoinverses and derived the stability property of scaled pseudoinverses. Forsgren further generalized these results to derive the supremum of weighted pseudoinverses $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|_2$ where \mathcal{P} is a set of diagonally dominant positive semidefinite matrices, by using a signature decomposition of weighting matrices W and by applying the Binet-Cauchy formula and Cramer’s rule for determinants. The results are also extended to equality constrained linear least squares problems.

In this paper we extend Forsgren’s results to a general complex matrix X to establish several equivalent formulae for $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|_2$, where \mathcal{P} is a set of diagonally dominant positive semidefinite matrices, or a set of weighting matrices arising from solving equality constrained least squares problems. We also discuss the stability property of these weighted pseudoinverses.

1. INTRODUCTION

In this paper we will use the following notation. $C^{m \times n}$ ($R^{m \times n}$) is the set of m by n matrices with complex (real) entries, $C_r^{m \times n}$ is a subset of $C^{m \times n}$ in which any matrix has rank r , $C^m(R^m) = C^{m \times 1}(R^{m \times 1})$. I_m denotes the identity matrix of order m , $0_{m \times n}$ is the m by n matrix with zero entries (if no confusion occurs, we will omit the subscript), and $\|\cdot\| \equiv \|\cdot\|_2$ is the Euclidean vector norm or the spectral matrix norm. For any matrix $X \in C_r^{m \times n}$, X^+ is the Moore-Penrose pseudoinverse of X , X^T is the transpose of X , X^H is the conjugate transpose of X , $\text{rank}(X)$ is the rank of X , $R(X)$ is the range of X . $X^{(i)} \prec X$ means that $X^{(i)}$ is a submatrix of X formed with a set of rows of X . $\mathcal{P}(X)$ is a set of real symmetric positive semidefinite matrices of order $m = \text{rows of } X$ such that for any $W \in \mathcal{P}(X)$, $\text{rank}(WX) = \text{rank}(X)$. $\text{inf}_+(X)$ denotes the smallest nonzero singular value of X . For a vector of sequence of increasing r integers $J = \{i_1, \dots, i_r\} \subset \{1, \dots, m\}$, $X \in C_r^{m \times n}$ and $D = \text{diag}(d_1, \dots, d_m)$, X_J denotes a submatrix of X formed with those rows $i_1, \dots, i_r \in J$ and $D_{JJ} = \text{diag}(d_{i_1}, \dots, d_{i_r})$. $J(X)$ is a set of indexes

Received by the editor March 27, 1996.

1991 *Mathematics Subject Classification*. Primary 15A09, 65F35.

Key words and phrases. Weighted pseudoinverse, supremum, stability.

This work was supported by the National Natural Science Foundation, P.R. China.

defined by

$$(1.1) \quad J(X) = \{J = \{i_1, \dots, i_r\} : 1 \leq i_1 < \dots < i_r \leq m, \text{rank}(X_J) = \text{rank}(X) = r\}.$$

Consider the following mathematical programming and nonlinear constrained optimization problem

$$(1.2) \quad \min_{y \in R^n} f(y), \quad \text{subject to } y \geq 0 \text{ and } Cy = d,$$

in which $f : R^n \rightarrow R$, $C \in R^{m \times n}$ and $d \in R^m$ are given. When solving (1.2) by an interior method [1], [5], [6], [17], [22]-[24], one will obtain the following weighted least squares (WLS) problem

$$(1.3) \quad \min_{x \in R^n} \|W^{\frac{1}{2}}(Xx - g)\|,$$

where $W = W(\tau) \in \mathcal{P}(X)$, $\tau > 0$ is a parameter. Similarly, when solving the equality constrained least squares problem (LSE) [9]

$$(1.4) \quad \min_{x \in R^n} \|W^{\frac{1}{2}}(Kx - g_2)\| \quad \text{subject to } Lx = g_1$$

by the weighting method, one will also obtain a WLS problem like (1.3). When $\tau \rightarrow +\infty$, the minimum 2-norm solution of (1.3) will tend to the minimum 2-norm solution of (1.2) or (1.4).

Let $X \in C^{m \times n}$, $g \in C^m$ and $W = W(\tau) \in \mathcal{P}(X)$ be given. Let $\bar{X} = X + \delta X$ and $\bar{g} = g + \delta g$ be the perturbed versions of X and g respectively satisfying $\text{rank}(W^{\frac{1}{2}}\bar{X}) = \text{rank}(W^{\frac{1}{2}}X)$. Then the minimum 2-norm least squares solutions π and $\bar{\pi}$ of the problems

$$(1.5) \quad \min_{x \in C^n} \|W^{\frac{1}{2}}(Xx - g)\| \quad \text{and} \quad \min_{x \in C^n} \|W^{\frac{1}{2}}(\bar{X}x - \bar{g})\|$$

are respectively

$$(1.6) \quad \pi = (W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}g \quad \text{and} \quad \bar{\pi} = (W^{\frac{1}{2}}\bar{X})^+W^{\frac{1}{2}}\bar{g}.$$

It is well known [7], [16] that when $\|\delta X\| \rightarrow 0$, $\|\delta g\| \rightarrow 0$ and $\text{rank}(W^{\frac{1}{2}}\bar{X}) = \text{rank}(W^{\frac{1}{2}}X)$, then

$$\bar{\pi} \rightarrow \pi.$$

In general, when the parameter $\tau \rightarrow +\infty$, the condition number of the matrix $W^{\frac{1}{2}}X$ would be unbounded, so in this case the WLS problems (1.5) would be “ill-conditioned”. On the other hand, the boundedness of π and $\bar{\pi}$ respectively depend upon the upper bounds of $\|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|$ and $\|(W^{\frac{1}{2}}\bar{X})^+W^{\frac{1}{2}}\|$ when W ranges over a subset $\mathcal{P} \subseteq \mathcal{P}(X)$. Also, the perturbation bound of $\|\bar{\pi} - \pi\|$ depends upon the upper bounds of $\|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|$ and $\|(W^{\frac{1}{2}}\bar{X})^+W^{\frac{1}{2}}\|$ when W ranges over a subset $\mathcal{P} \subseteq \mathcal{P}(X)$, as well as the bounds $\|\delta X\|$ and $\|\delta g\|$. So the problems of specifying sets $\mathcal{P} \subseteq \mathcal{P}(X)$ such that $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\| < +\infty$ and determining conditions such that $\sup_{\|\bar{X}-X\| \leq \eta} \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}\bar{X})^+W^{\frac{1}{2}}\| < +\infty$ for some constant η are important subjects.

By applying the Binet-Cauchy formula for determinant and Cramer’s rule for solving a system of linear equations, the authors of [2], [3], [4], [10], [13], [17] discussed the geometry of the minimum 2-norm solution of the WLS problem (1.3)

and linear equations. The authors of [5], [6], [8], [17], [18], [22]-[24] proposed several algorithms and discussed stability for solving the constrained optimization problem.

When $X \in R_n^{m \times n}$, Stewart [15] and O'Leary [12] showed that when \mathcal{P} is the set of positive diagonal matrices of order m , then $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\| < +\infty$. They

also provided an estimate for $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|$. We [21] generalized their result

to obtain $\frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|} = \min_{X^{(i)} \prec X} \inf_+(X^{(i)})$. Recently Forsgren [4] further

generalized this result to the case when $X \in R_n^{m \times n}$ and $\mathcal{P} \subseteq \mathcal{P}(X)$ is a set of diagonally dominant positive semidefinite matrices, or a set of matrices arising from handling the LSE problem (1.4) by the weighting method. Forsgren constructed a signature decomposition of any matrix $W \in \mathcal{P}$ and applied the Binet-Cauchy formula and Cramer's rule to derive these results.

In this paper we extend Forsgren's results to a general complex matrix X to establish several equivalent formulae for $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|$ with some $\mathcal{P} \subseteq \mathcal{P}(X)$.

In particular, we will discuss the case that $\mathcal{P} \subseteq \mathcal{P}(X)$ is a set of semi-positive diagonal matrices, a set of diagonally dominant positive semidefinite matrices, or a set of weighting matrices arising from solving the LSE problem (1.4) by the weighting method. We also discuss stability properties of these weighted pseudoinverses.

The paper is organized as follows. Section 2 will derive equivalent formulae for the supremum of weighted pseudoinverses; Section 3 will derive equivalent formulae for the supremum of weighted pseudoinverses arising from the LSE problem with the weighting method; Section 4 will discuss the stability of weighted pseudoinverses; Section 5 will conclude the paper with some remarks.

In §§3-4, we need the following fact.

Lemma 1.1. *Suppose that $L \in C^{m_1 \times n}$ and $K \in C^{m_2 \times n}$. Let*

$$(1.7) \quad X = \begin{pmatrix} L \\ K \end{pmatrix}, \quad Y = \begin{pmatrix} L & 0 \\ KL^+L & K(I_n - L^+L) \end{pmatrix}, \quad Z = \begin{pmatrix} L^+L \\ I_n - L^+L \end{pmatrix}.$$

Then

$$(1.8) \quad \text{rank}(X) = \text{rank}(Y) = \text{rank}(L) + \text{rank}(K(I_n - L^+L)).$$

Proof. It can be shown that $X = YZ$ and $Y = XZ^H$. From this we have

$$\text{rank}(X) \geq \text{rank}(XZ^H) = \text{rank}(Y) \geq \text{rank}(YZ) = \text{rank}(X).$$

Combining above inequalities and (1.7), the equalities in (1.8) follow. \square

2. EQUIVALENT FORMULAE FOR THE SUPREMUM OF WEIGHTED PSEUDOINVERSES

In this section, we will derive several equivalent formulae for the supremum of weighted pseudoinverses of a matrix $X \in C_r^{m \times n}$. We first derive the Moore-Penrose pseudoinverse of $(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}$ for any $W \in \mathcal{P}(X)$. Let the unitary decomposition of X^H be

$$(2.1) \quad X^H = QA \quad \text{so} \quad X = A^H Q^H,$$

in which $Q^H Q = I_r$ and A has full row rank r . Then we have

Lemma 2.1. For a given matrix $X \in C_r^{m \times n}$ and any weighting matrix $W \in \mathcal{P}(X)$, define

$$(2.2a) \quad B = (W^{\frac{1}{2}}X)^+W^{\frac{1}{2}} \quad \text{and} \quad C = WX(WX)^+X.$$

Then

$$\text{rank}(B) = \text{rank}(C) \quad \text{and} \quad B^+ = C.$$

Proof. Let the unitary decomposition of X be as in (2.1). It can be shown that if $\text{rank}(WX) = \text{rank}(X) = r$, then

$$\text{rank}(AWA^H) = \text{rank}(AW^2A^H) = \text{rank}(X) = r,$$

so both AWA^H and AW^2A^H are nonsingular. Then

$$(2.2b) \quad \begin{aligned} B &= (W^{\frac{1}{2}}A^HQ^H)^+W^{\frac{1}{2}} = Q(AWA^H)^{-1}AW, \\ C &= WA^HQ^H(WA^HQ^H)^+A^HQ^H = WA^H(AW^2A^H)^{-1}AWA^HQ^H, \end{aligned}$$

and so

$$\begin{aligned} BC &= (Q(AWA^H)^{-1}AW)(WA^H(AW^2A^H)^{-1}AWA^HQ^H) = QQ^H = (BC)^H, \\ CB &= (WA^H(AW^2A^H)^{-1}AWA^HQ^H)(Q(AWA^H)^{-1}AW) \\ &= WA^H(AW^2A^H)^{-1}AW = (CB)^H. \end{aligned}$$

From this it is obvious that

$$BCB = (BC)B = QQ^HB = B \quad \text{and} \quad CBC = C(BC) = CQQ^H = C;$$

that is, C satisfies all four conditions as the unique Moore-Penrose pseudoinverse of B [1]. So $C = B^+$. \square

For given $X \in C_r^{m \times n}$ and a set $\mathcal{P} \subseteq \mathcal{P}(X)$, we define the following two sets of vectors. Let

$$(2.3) \quad \begin{aligned} \mathcal{X} &= \{x \in R(X) : \|X^+x\| = 1\}, \\ \mathcal{Y}(\mathcal{P}) &= \{y \in C^m : \text{there exists } W \in \mathcal{P} \text{ such that } (WX)^Hy = 0\}, \\ \rho(\mathcal{P}) &= \inf_{x \in \mathcal{X}, y \in \mathcal{Y}(\mathcal{P})} \|x - y\|. \end{aligned}$$

We now present the first main result of this section.

Theorem 2.1. Suppose that $X \in C_r^{m \times n}$ with $r > 0$, and a set of matrices $\mathcal{P} \subseteq \mathcal{P}(X)$ such that

$$(2.4) \quad \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\| < +\infty.$$

Then

$$(2.5) \quad \frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|} = \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X) = \rho(\mathcal{P}).$$

Proof. (1) We first prove that

$$\frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|} = \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X).$$

Let $W \in \mathcal{P}$ be given. From Lemma 2.1, $((W^{\frac{1}{2}}X)^+W^{\frac{1}{2}})^+ = WX(WX)^+X$. Let the nonzero singular values of $(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}$ and $WX(WX)^+X$ be arranged respectively in decreasing order,

$$(2.6) \quad \begin{aligned} \sigma_1((W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}) &\geq \dots \geq \sigma_r((W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}) > 0, \\ \sigma_1(WX(WX)^+X) &\geq \dots \geq \sigma_r(WX(WX)^+X) > 0. \end{aligned}$$

Then we have

$$(2.7) \quad \sigma_j(WX(WX)^+X) = \frac{1}{\sigma_{r+1-j}((W^{\frac{1}{2}}X)^+W^{\frac{1}{2}})} \text{ for } j = 1, \dots, r.$$

Therefore

$$\begin{aligned} \inf_+(WX(WX)^+X) &= \sigma_r(WX(WX)^+X) = \frac{1}{\sigma_1((W^{\frac{1}{2}}X)^+W^{\frac{1}{2}})} = \frac{1}{\|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|} \\ &\geq \frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|}. \end{aligned}$$

Because the above inequality holds for any matrix $W \in \mathcal{P}$,

$$\inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X) \geq \frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|}.$$

By reversing the above procedure, we obtain that for any matrix $W \in \mathcal{P}$,

$$\frac{1}{\|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|} = \inf_+(WX(WX)^+X) \geq \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X),$$

so

$$\frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|} \geq \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X).$$

(2) We then prove that

$$\inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X) = \rho(\mathcal{P}).$$

Let the unitary decomposition of X be as in (2.1). Then from Lemma 2.1, for given $W \in \mathcal{P}$,

$$WX(WX)^+X = WA^H(AW^2A^H)^{-1}AWA^HQ^H.$$

Let z be the right singular vector of $WA^H(AW^2A^H)^{-1}AWA^H$ associated with $\inf_+(WX(WX)^+X)$, where $\|z\| = 1$. Then

$$\inf_+(WX(WX)^+X) = \|WA^H(AW^2A^H)^{-1}AWA^Hz\| = \|WX(WX)^+X(Qz)\|$$

and it can be shown that $x = X(Qz) \in \mathcal{X}$. Let $y = (I - WX(WX)^+)x$. Then $(WX)^Hy = 0$ so $y \in \mathcal{Y}(\mathcal{P})$ and

$$\inf_+(WX(WX)^+X) = \|WX(WX)^+x\| = \|x - y\| \geq \rho(\mathcal{P}).$$

Because the above inequality holds for any $W \in \mathcal{P}$,

$$\inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X) \geq \rho(\mathcal{P}).$$

On the contrary, for any $\epsilon > 0$, there exist $x_\epsilon \in \mathcal{X}$, $W_\epsilon \in \mathcal{P}$, $y_\epsilon \in \mathcal{Y}(\mathcal{P})$ such that $(W_\epsilon X)^H y_\epsilon = 0$ and

$$\rho(\mathcal{P}) > \|x_\epsilon - y_\epsilon\| - \epsilon.$$

Because $(W_\epsilon X)^H y_\epsilon = 0$, y_ϵ should be of the form $y_\epsilon = (I - W_\epsilon X(W_\epsilon X)^+)z$ for some vector $z \in C^m$. From the knowledge of the least squares problem [7] and noting that $x_\epsilon = XX^+x_\epsilon$, we obtain

$$\begin{aligned} \|x_\epsilon - y_\epsilon\| &\geq \min_{v \in C^m} \|x_\epsilon - (I - W_\epsilon X(W_\epsilon X)^+)v\| = \|x_\epsilon - (I - W_\epsilon X(W_\epsilon X)^+)x_\epsilon\| \\ &= \|W_\epsilon X(W_\epsilon X)^+X(X^+x_\epsilon)\| \geq \inf_+(W_\epsilon X(W_\epsilon X)^+X), \end{aligned}$$

so

$$\rho(\mathcal{P}) > \inf_+(W_\epsilon X(W_\epsilon X)^+X) - \epsilon \geq \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X) - \epsilon.$$

Because $\epsilon > 0$ is arbitrary, letting $\epsilon \rightarrow 0^+$ we obtain

$$\rho(\mathcal{P}) \geq \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X),$$

proving the assertions of the theorem. □

Notice that Theorem 2.1 is a generalization of [15], [12], [21].

Next we will present the second main result of this section by using an idea in [4, Theorem 3.1].

Theorem 2.2. *Suppose that $X \in C_r^{m \times n}$ with $r > 0$, integer $l \geq r$. Let $\mathcal{U} \subseteq R^{m \times l}$ and $\mathcal{P} \subseteq \mathcal{P}(X)$ be two given sets of matrices which respectively satisfy*

$$(2.8) \quad \begin{aligned} &\forall U \in \mathcal{U}, \text{rank}(U^T X) = r, \\ &\sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T\| < +\infty, \end{aligned}$$

where $U_J^T \equiv (U^T)_J$, and

$$(2.9) \quad \begin{aligned} \mathcal{P} &= \{W = UDU^T : U \in \mathcal{U}, D = \text{diag}(d_1, \dots, d_l) \geq 0 \\ &\text{such that } \text{rank}(WX) = \text{rank}(X)\}. \end{aligned}$$

Then

$$(2.10) \quad \begin{aligned} \frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}\|} &= \frac{1}{\sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T\|} = \rho(\mathcal{P}) \\ &= \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X) \\ &= \inf_{U \in \mathcal{U}} \min_{J \in J(U^T X)} \inf_+((U_J^T)^+ U_J^T X). \end{aligned}$$

Proof. Notice that $\mathcal{P} \subseteq \mathcal{P}(X)$.

(1) We first prove that

$$\inf_{U \in \mathcal{U}} \min_{J \in J(U^T X)} \inf_+((U_J^T)^+ U_J^T X) = \frac{1}{\sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T\|}.$$

Let the unitary decomposition of X be as in (2.1). For any $U \in \mathcal{U}$ and $J \in J(U^T X)$, we have $(U^T X)_J = U_J^T X$ and (see (1.1))

$$\text{rank}((U^T X)_J) = \text{rank}(U_J^T X) = \text{rank}(U^T X) = \text{rank}(X) = \text{rank}(U_J^T A^H) = r,$$

so $U_J^T A^H$ is nonsingular and U_J^T has full row rank r . Then we have the following identities

$$(2.11) \quad ((U_J^T X)^+ U_J^T)^+ = (Q(U_J^T A^H)^{-1} U_J^T)^+ = (U_J^T)^+ (U_J^T A^H) Q^H = (U_J^T)^+ U_J^T X,$$

from which we obtain by using the same argument as in step (1) of the proof of Theorem 2.1,

$$\inf_{U \in \mathcal{U}} \min_{J \in J(U^T X)} \inf_+((U_J^T)^+ U_J^T X) = \frac{1}{\sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T\|} > 0.$$

(2) We then prove that

$$\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}\| = \sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T\|$$

by applying the Binet-Cauchy formula and Cramer’s rule. For given $W = UDU^T \in \mathcal{P}$ with $U \in \mathcal{U}$ we have $\text{rank}(WX) = \text{rank}(X) = r$. Then

$$\text{rank}(X) \geq \text{rank}(U^T X) \geq \text{rank}(DU^T X) \geq \text{rank}(WX) = \text{rank}(X);$$

that is,

$$(2.12) \quad \text{rank}(U^T X) = \text{rank}(DU^T X) = \text{rank}(WX) = \text{rank}(X),$$

so AWA^H is nonsingular. For any vector $g \in C^m$, let

$$(2.13) \quad \pi = (W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}} g = Q(AWA^H)^{-1} AWg = Q(\tilde{A}^H D \tilde{A})^{-1} \tilde{A}^H D \tilde{g},$$

in which $\tilde{A} = U^T A^H$ and $\tilde{g} = U^T g$. Notice that for any $J \in J(U^T X)$, $\tilde{A}_J = (U^T A^H)_J = U_J^T A^H$ is nonsingular. Then by applying the Binet-Cauchy formula and Cramer’s rule [11],

$$\begin{aligned} (Q^H \pi)_i &= [(\tilde{A}^H D \tilde{A})^{-1} \tilde{A}^H D \tilde{g}]_i \\ &= \frac{\det(\tilde{A}^H D(\tilde{A} + (\tilde{g} - \tilde{A}e_i)e_i^T))}{\det(\tilde{A}^H D \tilde{A})} \\ &= \sum_{J \in J(U^T X)} \frac{\det((\tilde{A}_J)^H) \cdot \det(D_{JJ}) \cdot \det(\tilde{A}_J)}{\sum_{N \in J(U^T X)} \det((\tilde{A}_N)^H) \cdot \det(D_{NN}) \cdot \det(\tilde{A}_N)} ((\tilde{A}_J)^{-1} \tilde{g}_J)_i. \end{aligned}$$

Notice that $\tilde{A}_J Q^H = U_J^T X$ and $\tilde{g}_J = U_J^T g$. Then we deduce that

$$(2.14) \quad \pi = Q(Q^H \pi) = \sum_{J \in J(U^T X)} \frac{|\det(\tilde{A}_J)|^2 \cdot \det(D_{JJ})}{\sum_{N \in J(U^T X)} |\det(\tilde{A}_N)|^2 \cdot \det(D_{NN})} (U_J^T X)^+ U_J^T g.$$

From (2.14), π is the convex combination of the basic solutions. Then in exactly the same way as in the proof of [4, Theorem 3.1, Corollary 5.2], we obtain

$$(2.15) \quad \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}} g\| = \sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T g\|$$

and

$$(2.16) \quad \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}\| = \sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T\|.$$

The remaining equalities of (2.10) follow from Theorem 2.1. □

From Theorem 2.2, with \mathcal{U} satisfying (2.8) and \mathcal{P} defined in (2.9), one of the following inequalities implies the other:

$$\sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+ U_J^T\| < +\infty,$$

$$\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}\| < +\infty.$$

Especially, if \mathcal{U} only contains a finite number of matrices, then the above inequalities are always true.

In the case that $l = m$ and $\mathcal{U} = \{I_m\}$ so that \mathcal{P} is a set of semi-positive diagonal matrices, we then obtain essentially the same results as in [21].

Corollary 2.1. *Under the notation and the conditions in Theorem 2.2, if*

$$(2.17) \quad \mathcal{P} = \{W = \text{diag}(d_1, \dots, d_m) : d_j \geq 0 \text{ for } j = 1, \dots, m, \text{ such that } \text{rank}(WX) = \text{rank}(X)\},$$

then

$$(2.18) \quad \frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}\|} = \frac{1}{\max_i \|X^{(i)}\|} = \rho(\mathcal{P}) = \min_i \inf_+ (X^{(i)}),$$

in which the maximum or minimum is taken over all $X^{(i)} \prec X$ such that $\text{rank}(X^{(i)}) = r = \text{number of rows of } X^{(i)}$. □

In the case that \mathcal{P} is a set of $m \times m$ real symmetric diagonally dominant positive semidefinite matrices, Forsgren [4] has constructed an interesting signature decomposition of $W = UDU^T \in \mathcal{P}$, where $U \equiv U(s)$ has the following form, with $l = \frac{m(m+1)}{2}$:

$$(2.19a) \quad \begin{aligned} U(s) &\equiv (u(s)_1, \dots, u(s)_l) \in R^{m \times l} \text{ with} \\ u(s)_i &= e_i \text{ for } i = 1, \dots, m, \\ u(s)_{m+t(i,j)} &= e_i + s_{t(i,j)} e_j \text{ for } 1 \leq i < j \leq m, \end{aligned}$$

in which e_j is the j -th column of the identity matrix I_m , and for $1 \leq i < j \leq m$,

$$t(i, j) = m(i - 1) - \frac{i(i + 1)}{2} + j \quad \text{and} \quad s_{t(i,j)} = \begin{cases} 1 & \text{if } w_{ij} \geq 0, \\ -1 & \text{if } w_{ij} < 0, \end{cases}$$

and

$$(2.19b) \quad \begin{aligned} D &= \text{diag}(d_1, \dots, d_l) \text{ with } d_i = w_{ii} - \sum_{j=1, j \neq i}^m |w_{ij}| \text{ for } i = 1, \dots, m, \\ d_{m+t(i,j)} &= |w_{ij}| \text{ for } 1 \leq i < j \leq m. \end{aligned}$$

For a detailed description of the signature decomposition of W , we refer to [4]. From (2.19a) and (2.19b) we see that \mathcal{U} only contains a finite number of $m \times l$ matrices. Then we obtain the following results by applying Theorem 2.2 and [4, §3-§5].

Corollary 2.2. *Under the notation and the conditions of Theorem 2.2, if furthermore, \mathcal{P} is the set of real $m \times m$ symmetric positive semidefinite diagonally dominant matrices, then with $U \in \mathcal{U}$ defined in (2.19a), we have*

$$\begin{aligned}
 \frac{1}{\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|} &= \frac{1}{\sup_{U \in \mathcal{U}} \max_{J \in J(U^T X)} \|(U_J^T X)^+U_J^T\|} = \rho(\mathcal{P}) \\
 (2.20) \qquad \qquad \qquad &= \inf_{W \in \mathcal{P}} \inf_+(WX(WX)^+X) \\
 &= \inf_{U \in \mathcal{U}} \min_{J \in J(U^T X)} \inf_+((U_J^T)^+U_J^T X). \quad \square
 \end{aligned}$$

Remarks. 1. For given $X \in C_r^{m \times n}$, \mathcal{U} satisfying (2.8), \mathcal{P} defined in (2.9), (2.14) provides the geometric structure of the minimum 2-norm WLS solution π of (1.3) for given $W \in \mathcal{P}$, which is a generalization of well known formulae in [2], [3], [4], [10], [13], [17].

2. Theorems 2.1, 2.2 and Corollaries 2.1, 2.2 provide several equivalent formulae for $\sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\|$, which extend the results in [15], [12], [21], [4].

3. EQUIVALENT FORMULAE FOR THE SUPREMUM OF WEIGHTED PSEUDOINVERSES ARISING FROM SOLVING EQUALITY CONSTRAINED LINEAR LEAST SQUARES PROBLEMS WITH THE WEIGHTING METHOD

In [4] Forsgren also used the Binet-Cauchy formula and Cramer’s rule to obtain the supremum of weighted pseudoinverses arising from solving the LSE problem [9]

$$(3.1a) \qquad \min_x \|W^{\frac{1}{2}}(Kx - g_2)\|, \text{ subject to } Lx = g_1,$$

where $L \in R^{m_1 \times n}$, $K \in R^{m_2 \times n}$ are given matrices, and $g_1 \in R^{m_1}$, $g_2 \in R^{m_2}$ are given vectors. In this section we will generalize Forsgren’s results to derive equivalent formulae for the supremum of such weighted pseudoinverses for general complex matrices L, K and some set $\mathcal{P}_2 \subseteq \mathcal{P}(KP)$ with $P = I_n - L^+L$.

Let $L \in C_p^{m_1 \times n}$, $K \in C^{m_2 \times n}$, $g_1 \in C^{m_1}$ and $g_2 \in C^{m_2}$ be given. One method of solving (3.1a) is the weighting method. Let $m = m_1 + m_2$ and define

$$(3.2) \qquad g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}, \quad X = \begin{pmatrix} L \\ K \end{pmatrix} \text{ and } W(\tau) = \begin{pmatrix} \tau I_{m_1} & \\ & W \end{pmatrix}.$$

Suppose that $\text{rank}(X) = r$. Let $\pi(\tau)$ be the minimum 2-norm least squares solution of

$$(3.1b) \qquad \min_{x \in C^n} \|W(\tau)^{\frac{1}{2}}(Xx - g)\|,$$

so

$$\pi(\tau) = (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}g.$$

Then with

$$(3.3) \qquad \pi = \lim_{\tau \rightarrow +\infty} \pi(\tau) = \lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}g,$$

π is the minimum 2-norm solution of the LSE problem (3.1a).

In this section we will derive equivalent formulae for the supremum

$$(3.4) \qquad \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}\| (< +\infty)$$

for some set $\mathcal{P}_2 \subseteq \mathcal{P}(KP)$. We first need the following result:

Lemma 3.1. *Let $L \in C^{m_1 \times n}$, $K \in C^{m_2 \times n}$ and $W \in R^{m_2 \times m_2}$ be given matrices. Suppose that $\text{rank}(X) = r$ and $\text{rank}(L) = p$, where $X = \begin{pmatrix} L \\ K \end{pmatrix}$. Then with $W(\tau)$ defined in (3.2) such that $\text{rank}(W(\tau)X) = \text{rank}(X)$,*

$$(3.5) \quad \begin{aligned} \lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}} &= (L^+_{W^{\frac{1}{2}}K}, (W^{\frac{1}{2}}KP)^+W^{\frac{1}{2}}), \\ \lim_{\tau \rightarrow +\infty} W(\tau)X(W(\tau)X)^+X &= \begin{pmatrix} L \\ WKP(WKP)^+K \end{pmatrix}, \end{aligned}$$

in which

$$(3.6) \quad P = I - L^+L, \quad L^+_{W^{\frac{1}{2}}K} = (I - (W^{\frac{1}{2}}KP)^+W^{\frac{1}{2}}K)L^+.$$

So

$$(3.7) \quad \begin{pmatrix} L \\ WKP(WKP)^+K \end{pmatrix}^+ = (L^+_{W^{\frac{1}{2}}K}, (W^{\frac{1}{2}}KP)^+W^{\frac{1}{2}}).$$

Proof. Let $\tilde{K} = W^qK$ ($q = \frac{1}{2}, 1$) and let the SVD of $\begin{pmatrix} L \\ \tilde{K} \end{pmatrix}$ be [7]

$$(3.8) \quad \tilde{X} \equiv \begin{pmatrix} L \\ \tilde{K} \end{pmatrix} = ZTH^H = Z_1T_1H_1^H,$$

where Z, H are unitary matrices of appropriate sizes,

$$T = \text{diag}(T_1, 0), \quad T_1 = \text{diag}(t_1, \dots, t_r)$$

with $t_1 \geq \dots \geq t_r > 0$ the nonzero singular values of \tilde{X} , Z_1, H_1 are respectively the first r columns of Z, H . Let the CSD of Z be [14]

$$(3.9) \quad Z = \begin{pmatrix} Z_{11} & Z_{21} \\ Z_{21} & Z_{22} \end{pmatrix} \begin{matrix} m_1 \\ m_2 \\ r, & m-r \end{matrix} = \begin{pmatrix} U_1 & \\ & U_2 \end{pmatrix} \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} V_1^H \\ V_2^H \end{pmatrix},$$

where U_1, U_2, V_1 and V_2 are unitary matrices,

$$(3.10) \quad \begin{aligned} D_{11} &= \text{diag}(I_j, C, 0), & D_{12} &= \text{diag}(0, S, I_{m_1-p}) \\ D_{21} &= \text{diag}(0, S, I_{r-p}), & D_{22} &= \text{diag}(I_{m_2+j-r}, -C, 0), \end{aligned}$$

in which C, S are positive definite diagonal matrices of order $p - j$ with $C^2 + S^2 = I_{p-j}$ (when $p - j = 0$, both C and S are naught). Then from (3.8)-(3.10),

$$(3.11) \quad \begin{aligned} \lim_{\tau \rightarrow +\infty} (W(\tau)^qX)^+W(\tau)^q &= \lim_{\tau \rightarrow +\infty} (H_1T_1(\tau^{2q}Z_{11}^H Z_{11} + Z_{21}^H Z_{21})T_1H_1^H)^+H_1T_1(\tau^{2q}Z_{11}^H, Z_{21}^H W^q) \\ &= \lim_{\tau \rightarrow +\infty} H_1T_1^{-1}(\tau^{2q}Z_{11}^H Z_{11} + Z_{21}^H Z_{21})^{-1}(\tau^{2q}Z_{11}^H, Z_{21}^H W^q) \\ &= H_1T^{-1}(V_1 \text{diag}(I_j, C^{-1}, 0)U_1^H, V_1 \text{diag}(0, 0, I_{r-p})U_2^H W^q) \\ &= H_1T_1^{-1}(Z_{11}^+, (I_r - Z_{11}^+ Z_{11})Z_{21}^H W^q). \end{aligned}$$

Also from (3.8)-(3.10),

$$\begin{aligned}
 (3.12) \quad & \lim_{\tau \rightarrow +\infty} W(\tau)^q X(W(\tau)^q X)^+ X = \lim_{\tau \rightarrow +\infty} \begin{pmatrix} \tau^q Z_{11} \\ Z_{21} \end{pmatrix} T_1 H_1^H \left(\begin{pmatrix} \tau^q Z_{11} \\ Z_{21} \end{pmatrix} T_1 H_1^H \right)^+ X \\
 & = \lim_{\tau \rightarrow +\infty} \begin{pmatrix} \tau^q Z_{11} \\ Z_{21} \end{pmatrix} (\tau^{2q} Z_{11}^H Z_{11} + Z_{21}^H Z_{21})^{-1} (\tau^q Z_{11}^H, Z_{21}^H) X \\
 & = \begin{pmatrix} Z_{11} Z_{11}^+ & \\ & I - Z_{22} Z_{22}^+ \end{pmatrix} X.
 \end{aligned}$$

So with $q = \frac{1}{2}$ we have from (3.11) and [19, (3.1)-(3.3), (3.7), (3.8), (3.11)] that

$$\begin{aligned}
 (3.13) \quad & \lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}} X)^+ W(\tau)^{\frac{1}{2}} = (H_1 T_1^{-1} Z_{11}^+, H_1 T_1^{-1} (I - Z_{11}^+ Z_{11}) Z_{21}^H W^{\frac{1}{2}}) \\
 & = (L_{W^{\frac{1}{2}} K}^+, (W^{\frac{1}{2}} K P)^+ W^{\frac{1}{2}}),
 \end{aligned}$$

where $L_{W^{\frac{1}{2}} K}^+$ defined in (3.6) is the weighted pseudoinverse for the LSE problem (3.1a).

With $q = 1$ we also have from (3.12) and [19, (3.1)-(3.3), (3.7), (3.8), (3.11)] that

$$(3.14) \quad \lim_{\tau \rightarrow +\infty} W(\tau) X(W(\tau) X)^+ X = \begin{pmatrix} Z_{11} Z_{11}^+ L \\ (I - Z_{22} Z_{22}^+) K \end{pmatrix} = \begin{pmatrix} L \\ W K P (W K P)^+ K \end{pmatrix}.$$

Then (3.7) follows from Lemma 2.1 and (3.5). □

Now we can present the main result of this section by applying Theorem 2.1, Lemma 3.1 and by using the idea in [4, §6].

Theorem 3.1. *Let $L \in C^{m_1 \times n}$, $K \in C^{m_2 \times n}$ be given and $X = \begin{pmatrix} L \\ K \end{pmatrix}$. Suppose that $\text{rank}(X) = r$ and $\text{rank}(L) = p$. Let $\mathcal{U}_2 \subseteq R^{m_2 \times l}$ be a given set of matrices for some integer $l \geq m_2$, such that*

$$\begin{aligned}
 (3.15) \quad & \forall U \in \mathcal{U}_2, \text{rank}(U^T K P) = \text{rank}(K P) = r - p, \\
 & \sup_{U \in \mathcal{U}_2} \max_{J \in J(U^T K P)} \left\| \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I \\ U_J^T \end{pmatrix} \right\| < +\infty,
 \end{aligned}$$

where $J(U^T K P)$ is a set of vectors of indexes,

$$(3.16) \quad J(U^T K P) = \{J = \{i_1, \dots, i_{r-p}\} : \text{rank}(U_J^T K P) = \text{rank}(K P) = r - p\},$$

and let \mathcal{P}_2 be a set of $m_2 \times m_2$ real matrices defined by

$$\begin{aligned}
 (3.17) \quad & \mathcal{P}_2 = \{W = U D U^T : D \text{ is a semi-positive diagonal matrix of order } l, \\
 & U \in \mathcal{U}_2 \text{ such that } \text{rank}(W K P) = \text{rank}(K P)\}.
 \end{aligned}$$

Then with $W(\tau)$ defined in (3.2),

$$\begin{aligned}
 (3.18) \quad & \frac{1}{\sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}X) + W(\tau)^{\frac{1}{2}}\|} = \frac{1}{\sup_{U \in \mathcal{U}_2} \max_{J \in J(U^T K P)} \left\| \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I \\ U_J^T \end{pmatrix} \right\|} \\
 & = \inf_{W \in \mathcal{P}_2} \inf_+ \left(\begin{matrix} L \\ W K P (W K P)^+ K \end{matrix} \right) = \inf_{U \in \mathcal{U}_2} \min_{J \in J(U^T K P)} \inf_+ \left(\begin{matrix} L \\ (U_J^T)^+ U_J^T K \end{matrix} \right) \\
 & = \frac{1}{\sup_{W \in \mathcal{P}_2} \|(L^+_{W^{\frac{1}{2}} K}, (W^{\frac{1}{2}} K P)^+ W^{\frac{1}{2}})\|}.
 \end{aligned}$$

Proof. (1) We first prove that

$$\sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}X)^+ W(\tau)^{\frac{1}{2}}\| = \sup_{U \in \mathcal{U}_2} \max_{J \in J(U^T K P)} \left\| \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I \\ U_J^T \end{pmatrix} \right\|$$

by applying the Binet-Cauchy formula and Cramer’s rule.

Let the unitary decomposition of X be as in (2.1),

$$(3.19) \quad X = \begin{pmatrix} L \\ K \end{pmatrix} = A^H Q^H = \begin{matrix} m_1 \\ m_2 \end{matrix} \begin{pmatrix} L_1 \\ \tilde{K} \end{pmatrix} Q^H.$$

Then $\text{rank}(L_1) = \text{rank}(L) = p$. Let the unitary decomposition of L_1 be

$$(3.20) \quad L_1 = M \tilde{L},$$

where $M^H M = I_p$ and \tilde{L} has full row rank p . For any $W = U D U^T \in \mathcal{P}_2$, $\tau > 0$ and $g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \in C^m$, let

$$(3.21) \quad \pi(\tau) = (W(\tau)^{\frac{1}{2}}X)^+ W(\tau)^{\frac{1}{2}}g = Q(\tilde{A}^H \tilde{D}(\tau) \tilde{A})^{-1} \tilde{A} \tilde{D}(\tau) \tilde{g},$$

in which

$$(3.22) \quad \tilde{A} = \begin{pmatrix} \tilde{L} \\ U^T \tilde{K} \end{pmatrix}, \quad \tilde{D}(\tau) = \begin{pmatrix} \tau I_p & \\ & D \end{pmatrix} \quad \text{and} \quad \tilde{g} = \begin{pmatrix} M^H g_1 \\ U^T g_2 \end{pmatrix}.$$

Then with exactly the same procedure as deriving (2.14),

$$(3.23) \quad \pi(\tau) = \sum_{J \in J(\tilde{A})} \frac{|\det(\tilde{A}_J)|^2 \cdot \det(\tilde{D}(\tau)_{JJ})}{\sum_{N \in J(\tilde{A})} |\det(\tilde{A}_N)|^2 \cdot \det(\tilde{D}(\tau)_{NN})} Q(\tilde{A}_J)^+ \tilde{g}_J.$$

Notice that $\det(\tilde{D}(\tau)_{JJ})$ has the maximum power τ^p of τ if and only if \tilde{A}_J contains entire \tilde{L} . Therefore when $\tau \rightarrow +\infty$, in both denominators and numerators of (3.23) only those terms with \tilde{A}_J and \tilde{A}_N containing entire \tilde{L} remain. So when $\tau \rightarrow +\infty$, (3.23) only contains those $J, N \in J(\tilde{A})$ such that $\{1, \dots, p\} \subset J$ and $\{1, \dots, p\} \subset N$.

On the other hand, for each $J \in J(\tilde{A})$ such that $\{1, \dots, p\} \subset J$, denote $J = \{1, \dots, p, i_{p+1}, \dots, i_r\}$ and $J_2 = \{i_{p+1} - p, \dots, i_r - p\}$. Notice that from (3.19)-(3.20), $\tilde{K}(I_r - \tilde{L}^+ \tilde{L})Q^H = K P$, so $\text{rank}(U_{J_2}^T \tilde{K}(I_r - \tilde{L}^+ \tilde{L})) = \text{rank}(U_{J_2}^T K P)$. Then from (3.19)-(3.22) and Lemma 1.1 we have $\tilde{A}_J = \begin{pmatrix} \tilde{L} \\ U_{J_2}^T \tilde{K} \end{pmatrix}$ and

$$(3.24) \quad r = \text{rank}(\tilde{A}_J) = \text{rank}(\tilde{L}) + \text{rank}(U_{J_2}^T \tilde{K}(I_r - \tilde{L}^+ \tilde{L})) = p + \text{rank}(U_{J_2}^T K P),$$

from which we see that $J \in J(\tilde{A})$ with $\{1, \dots, p\} \subset J$, if and only if $J_2 \in J(U^T KP)$. With the above observation, we obtain by letting $\tau \rightarrow +\infty$ in (3.23),

$$(3.25)$$

$$\begin{aligned} \pi &= \lim_{\tau \rightarrow +\infty} \pi(\tau) \\ &= \sum_{J \in J(U^T KP)} \frac{|\det \begin{pmatrix} \tilde{L} \\ U_J^T \tilde{K} \end{pmatrix}|^2 \cdot \det(D_{JJ})}{\sum_{N \in J(U^T KP)} |\det \begin{pmatrix} \tilde{L} \\ U_N^T \tilde{K} \end{pmatrix}|^2 \cdot \det(D_{NN})} Q \begin{pmatrix} \tilde{L} \\ U_J^T \tilde{K} \end{pmatrix}^{-1} \begin{pmatrix} M^H g_1 \\ U_J^T g_2 \end{pmatrix}, \end{aligned}$$

in which for simplicity we have used $J, N \in J(U^T KP)$ instead of $J_2, N_2 \in J(U^T KP)$. Notice that for any $J \in J(U^T KP)$, $\begin{pmatrix} \tilde{L} \\ U_J^T \tilde{K} \end{pmatrix}$ is nonsingular. By using the notation in (2.1), (3.19) and (3.20) we have

$$\begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} g_1 \\ U_J^T g_2 \end{pmatrix} = \begin{pmatrix} M \tilde{L} Q^H \\ U_J^T \tilde{K} Q^H \end{pmatrix}^+ \begin{pmatrix} g_1 \\ U_J^T g_2 \end{pmatrix} = Q \begin{pmatrix} \tilde{L} \\ U_J^T \tilde{K} \end{pmatrix}^{-1} \begin{pmatrix} M^H g_1 \\ U_J^T g_2 \end{pmatrix}.$$

Substituting this into (3.25) we get

$$(3.26) \quad \begin{aligned} \pi &= \lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}} X)^+ W(\tau)^{\frac{1}{2}} g \\ &= \sum_{J \in J(U^T KP)} \frac{|\det \begin{pmatrix} \tilde{L} \\ U_J^T \tilde{K} \end{pmatrix}|^2 \cdot \det(D_{JJ})}{\sum_{N \in J(U^T KP)} |\det \begin{pmatrix} \tilde{L} \\ U_N^T \tilde{K} \end{pmatrix}|^2 \cdot \det(D_{NN})} \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} g_1 \\ U_J^T g_2 \end{pmatrix}. \end{aligned}$$

Then by applying the same argument as in [4, §6],

$$(3.27)$$

$$\begin{aligned} \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}} X)^+ W(\tau)^{\frac{1}{2}} g\| &= \sup_{U \in \mathcal{U}_2} \max_{J \in J(U^T KP)} \left\| \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} g_1 \\ U_J^T g_2 \end{pmatrix} \right\|, \\ \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}} X)^+ W(\tau)^{\frac{1}{2}}\| &= \sup_{U \in \mathcal{U}_2} \max_{J \in J(U^T KP)} \left\| \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I_{m_1} & \\ & U_J^T \end{pmatrix} \right\|. \end{aligned}$$

(2) We then prove that

$$\inf_{U \in \mathcal{U}_2} \min_{J \in J(U^T KP)} \inf_+ \left(\begin{pmatrix} L \\ (U_J^T)^+ U_J^T \tilde{K} \end{pmatrix} \right) = \frac{1}{\sup_{U \in \mathcal{U}_2} \max_{J \in J(U^T KP)} \left\| \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I_{m_1} & \\ & U_J^T \end{pmatrix} \right\|}.$$

By applying (2.1), (3.19) and (3.20) we have

$$\begin{aligned} \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I_{m_1} & \\ & U_J^T \end{pmatrix} &= \begin{pmatrix} M \tilde{L} Q^H \\ U_J^T \tilde{K} Q^H \end{pmatrix}^+ \begin{pmatrix} I_{m_1} & \\ & U_J^T \end{pmatrix} \\ &= Q \begin{pmatrix} \tilde{L} \\ U_J^T \tilde{K} \end{pmatrix}^{-1} \begin{pmatrix} M^H & \\ & U_J^T \end{pmatrix}. \end{aligned}$$

Because $Q^H Q = I_r$ and $\begin{pmatrix} M^H & \\ & U_J^T \end{pmatrix}$ has full row rank r , from the above identity we obtain

$$(3.28) \quad \left(\begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I_{m_1} & \\ & U_J^T \end{pmatrix} \right)^+ = \begin{pmatrix} M & \\ & (U_J^T)^+ \end{pmatrix} \begin{pmatrix} \tilde{L} \\ U_J^T \tilde{K} \end{pmatrix} Q^H = \begin{pmatrix} L \\ (U_J^T)^+ U_J^T K \end{pmatrix}.$$

Notice that the above equality holds for any $U \in \mathcal{U}_2$ and $J \in J(U^T K P)$. We obtain by using the same argument as in step (1) of the proof of Theorem 2.1,

$$\inf_{U \in \mathcal{U}_2} \min_{J \in J(U^T K P)} \inf_+ \left(\begin{pmatrix} L \\ (U_J^T)^+ U_J^T K \end{pmatrix} \right) = \frac{1}{\sup_{U \in \mathcal{U}_2} \max_{J \in J(U^T K P)} \left\| \begin{pmatrix} L \\ U_J^T K \end{pmatrix}^+ \begin{pmatrix} I_{m_1} & \\ & U_J^T \end{pmatrix} \right\|}.$$

The remaining equalities of the theorem can be proved by applying Theorems 2.1, 2.2 and Lemma 3.1. \square

In the case that $l = m_2$ and $\mathcal{U}_2 = \{I_{m_2}\}$ so that \mathcal{P}_2 is a set of semi-positive diagonal matrices, we have the following result:

Corollary 3.1. *Under the notation and the conditions in Theorem 3.1, furthermore suppose that \mathcal{P}_2 is defined as*

$$(3.29) \quad \mathcal{P}_2 = \{W = \text{diag}(d_1, \dots, d_{m_2}) \geq 0 : \text{rank}(W K P) = \text{rank}(K P) = r - p\}.$$

Then

$$(3.30) \quad \frac{1}{\sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}} X)^+ W(\tau)^{\frac{1}{2}}\|} = \frac{1}{\max_i \left\| \begin{pmatrix} L \\ K^{(i)} \end{pmatrix}^+ \right\|} = \min_i \inf_+ \left(\begin{pmatrix} L \\ K^{(i)} \end{pmatrix} \right),$$

in which the maximum or minimum is taken over all $K^{(i)} \prec K$ such that $\text{rank}(K^{(i)} P) = r - p =$ number of rows of $K^{(i)}$.

Proof. In Theorem 3.1, take $l = m_2$ and $\mathcal{U}_2 = \{I_{m_2}\}$. \square

In the case that \mathcal{P}_2 is a set of $m_2 \times m_2$ diagonally dominant positive semidefinite matrices, we can use a similar signature decomposition of any matrix $W = U D U^T \in \mathcal{P}_2$, as in [4], to obtain the following result:

Corollary 3.2. *Under the notation and the conditions in Theorem 3.1, furthermore suppose that \mathcal{P}_2 is defined as*

$$(3.31) \quad \mathcal{P}_2 = \{W \in \mathcal{P}(K P) : w_{ii} \geq \sum_{j=1, j \neq i}^{m_2} |w_{ij}|\}.$$

Let $l = \frac{m_2(m_2+1)}{2}$. For any $W \in \mathcal{P}_2$, define matrices $D = \text{diag}(d_1, \dots, d_l)$ and $U \in \mathcal{U}_2 \subset R^{m_2 \times l}$ as

$$(3.32) \quad \begin{aligned} d_i &= w_{ii} - \sum_{j=1, j \neq i}^{m_2} |w_{ij}| \text{ for } 1 \leq i \leq m_2, \\ d_{m_2+t(i,j)} &= |w_{ij}| \text{ for } 1 \leq i < j \leq m_2, \end{aligned}$$

$$\begin{aligned}
 (3.33) \quad U &\equiv U(s) = (u(s)_1, \dots, u(s)_l) \in R^{m_2 \times l}, \\
 u(s)_i &= e_i \quad \text{for } i = 1, \dots, m_2, \\
 u(s)_{m_2+t(i,j)} &= e_i + s_{t(i,j)}e_j \quad \text{for } 1 \leq i < j \leq m_2,
 \end{aligned}$$

in which e_j is the j -th column of the identity matrix I_{m_2} and for $1 \leq i < j \leq m_2$,

$$t(i, j) = m_2(i - 1) - \frac{i(i + 1)}{2} + j \quad \text{and} \quad s_{t(i,j)} = \begin{cases} 1 & \text{if } w_{ij} \geq 0, \\ -1 & \text{if } w_{ij} < 0. \end{cases}$$

Then

$$\begin{aligned}
 (3.34) \quad &\frac{1}{\sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}\|} = \inf_{W \in \mathcal{P}_2} \inf_+ \left(\begin{matrix} L \\ WKP(WKP)^+K \end{matrix} \right) \\
 &= \inf_{U \in \mathcal{U}_2} \min_{J \in J(U^T X)} \inf_+ \left(\begin{matrix} L \\ (U_J^T)^+U_J^T K \end{matrix} \right) = \frac{1}{\sup_{W \in \mathcal{P}_2} \|(L^+_{W^{\frac{1}{2}}K}, (W^{\frac{1}{2}}KP)^+W^{\frac{1}{2}})\|}.
 \end{aligned}$$

Proof. From [4, §3], any matrix $W \in \mathcal{P}_2$ has the form $W = UDU^T$ with D, U defined in (3.32) and (3.33). □

Remarks. 1. For given $L \in C_p^{m_1 \times n}, K \in C^{m_2 \times n}$ such that $X = \begin{pmatrix} L \\ K \end{pmatrix} \in C_r^{m \times n}, \mathcal{U}_2$ satisfying (3.15), \mathcal{P}_2 defined in (3.17), (3.26) provides the geometric structure of the minimum 2-norm solution π of (3.1a) for given $W \in \mathcal{P}_2$, which is a generalization of the formula in [4].

2. Lemma 3.1 and Theorem 3.1 provide algebraic relations between the matrix $\lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ and its pseudoinverse, which can be used to study properties of the constrained least squares problem.

3. Theorem 3.1 and Corollaries 3.1, 3.2 provide several equivalent formulae for

$$\sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}\|,$$

which extend the results in [4].

4. STABILITY OF WEIGHTED PSEUDOINVERSES

In the theory of pseudoinverses, one important problem is the boundedness of pseudoinverses under small perturbations. Let $X \in C_r^{m \times n}$ and $\hat{X} = X + \delta X \in C^{m \times n}$. Then the following results are well known (e.g., [16, pp. 136-152]).

1. If $\|\delta X\| \cdot \|X^+\| < 1$, then $\text{rank}(\hat{X}) \geq \text{rank}(X)$.
2. If $\|\delta X\| \cdot \|X^+\| < 1$ and $\text{rank}(\hat{X}) > \text{rank}(X)$, then $\|\hat{X}^+\| \geq \frac{1}{\|\delta X\|}$.
3. If $\|\delta X\| \cdot \|X^+\| < 1$ and $\text{rank}(\hat{X}) = \text{rank}(X)$, then $\frac{\|X^+\|}{1 + \|\delta X\| \cdot \|X^+\|} \leq \|\hat{X}^+\| \leq \frac{\|X^+\|}{1 - \|\delta X\| \cdot \|X^+\|}$.

So $\|\hat{X}^+\|$ are bounded for all small perturbations δX with $\|\delta X\| \cdot \|X^+\| \leq \eta < 1$ iff $\text{rank}(\hat{X}) = \text{rank}(X)$, where $0 \leq \eta < 1$ is a constant. For stability of weighted pseudoinverses, the situation is more complicated.

In this section we will discuss the stability property of $(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}$ over a set $\mathcal{P} \subseteq \mathcal{P}(X)$. We use the following definition.

Definition 4.1. For given $X \in C_r^{m \times n}$ and a set $\mathcal{P} \subseteq \mathcal{P}(X)$, suppose that

$$(4.1a) \quad \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}\| < +\infty.$$

We say that $(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}$ is stable over \mathcal{P} , if there exists a constant $\eta > 0$ which does not depend on the choice of $W \in \mathcal{P}$, such that

$$(4.1b) \quad \sup_{X+\delta X \in C_r^{m \times n}, \|\delta X\| \leq \eta} \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}(X + \delta X))^+W^{\frac{1}{2}}\| < +\infty.$$

We now have the following theorem.

Theorem 4.1. Suppose that $X \in C_r^{m \times n}$ with $0 < r \leq m$, an integer $l \geq r$. Let $\mathcal{U} \subseteq R^{m \times l}$ and $\mathcal{P} \subseteq \mathcal{P}(X)$ be two given sets of matrices which respectively satisfy

$$(4.2) \quad \begin{aligned} &\forall U \in \mathcal{U}, \text{rank}(U^T X) = \text{rank}(X), \\ &\inf_{U \in \mathcal{U}} \min_{J \in J(U^T X)} \inf_+((U_J^T)^+U_J^T X) > 0, \end{aligned}$$

$$(4.3) \quad \begin{aligned} \mathcal{P} = \{W = UDU^T : D = \text{diag}(d_1, \dots, d_l) \geq 0, U \in \mathcal{U}, \\ \text{such that } \text{rank}(WX) = \text{rank}(X)\}. \end{aligned}$$

Then $(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}$ is stable over \mathcal{P} , if and only if for any $U \in \mathcal{U}$ and any vector of sequence of increasing r integers $J = \{i_1, \dots, i_r\}$ with $1 \leq i_1 < \dots < i_r \leq m$,

$$(4.4) \quad \text{rank}(U_J^T) = r \quad \text{implies} \quad \text{rank}(U_J^T X) = r.$$

Furthermore, if the condition in (4.4) holds, then for any $\hat{X} = X + \delta X \in C_r^{m \times n}$ such that $\|\delta X\| \leq a \cdot \rho(\mathcal{P}) \equiv \eta$ with $0 < a < 1$ some constant, we have the following estimate:

$$(4.5) \quad \frac{1}{\rho(\mathcal{P}) + \|\delta X\|} \leq \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}\hat{X})^+W^{\frac{1}{2}}\| \leq \frac{1}{\rho(\mathcal{P}) - \|\delta X\|}.$$

Proof. Necessity. Suppose that there exists a matrix $\hat{U} \in \mathcal{U}$ and a vector of sequence of increasing r integers $N = \{i_1, \dots, i_r\}$ with $1 \leq i_1 < \dots < i_r \leq m$, such that $\text{rank}(\hat{U}_N^T) = r$ but $\text{rank}(\hat{U}_N^T X) = q < r$. Then with the unitary decomposition of $X = A^H Q^H$, see (2.1), for any $\epsilon > 0$, there always exists a matrix $E_\epsilon \in C^{m \times r}$, such that $\|E_\epsilon\| < \epsilon \cdot \rho(\mathcal{P})$ and with $A_\epsilon^H = A^H + E_\epsilon$, $\hat{U}_N^T A_\epsilon^H$ is nonsingular. Let $X_\epsilon = A_\epsilon^H Q^H$. Then $X_\epsilon \in C_r^{m \times n}$, $N \in J(\hat{U}^T X_\epsilon)$ and

$$\begin{aligned} &\inf_{U \in \mathcal{U}} \min_{J \in J(U^T X_\epsilon)} \inf_+((U_J^T)^+U_J^T X_\epsilon) \leq \inf_+((\hat{U}_N^T)^+\hat{U}_N^T X_\epsilon) = \sigma_r((\hat{U}_N^T)^+\hat{U}_N^T X_\epsilon) \\ &\leq \sigma_r((\hat{U}_N^T)^+\hat{U}_N^T X) + \|(\hat{U}_N^T)^+\hat{U}_N^T E_\epsilon Q^H\| \leq \|E_\epsilon\| < \epsilon \rho(\mathcal{P}). \end{aligned}$$

So according to Theorem 2.2, for any constant η with $0 < \eta \ll 1$,

$$\begin{aligned} &\sup_{X+\delta X \in C_r^{m \times n}, \|\delta X\| \leq \eta} \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}(X + \delta X))^+W^{\frac{1}{2}}\| \geq \lim_{\epsilon \rightarrow 0^+} \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}}X_\epsilon)^+W^{\frac{1}{2}}\| \\ &\geq \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon \rho(\mathcal{P})} = +\infty. \end{aligned}$$

So when the condition in (4.4) does not hold, $(W^{\frac{1}{2}}X)^+W^{\frac{1}{2}}$ is not stable over \mathcal{P} . *Sufficiency.* Suppose that the condition in (4.4) holds. Let $0 < \alpha < 1$ be a constant and let $\eta \equiv \alpha \cdot \rho(\mathcal{P})$. Suppose that any matrix $\delta X \in C^{m \times n}$ satisfies $\|\delta X\| \leq \alpha \cdot \rho(\mathcal{P})$ and $\hat{X} = X + \delta X \in C_r^{m \times n}$. Then for any $U \in \mathcal{U}$ and any vector $N \in J(U^T \hat{X})$,

$\text{rank}(U_N^T) = r$ and so $N \in J(U^T X)$. We then have from the perturbation of the singular values [7], [16],

$$\begin{aligned}
 \sigma_r((U_N^T)^+ U_N^T \hat{X}) &\geq \sigma_r((U_N^T)^+ U_N^T X) - \|(U_N^T)^+ U_N^T \delta X\| \\
 &\geq \rho(\mathcal{P}) - \|\delta X\| \geq (1 - \alpha)\rho(\mathcal{P}) > 0, \\
 \sigma_r((U_N^T)^+ U_N^T \hat{X}) &\leq \sigma_r((U_N^T)^+ U_N^T X) + \|(U_N^T)^+ U_N^T \delta X\| \\
 &\leq \sigma_r((U_N^T)^+ U_N^T X) + \|\delta X\|.
 \end{aligned}
 \tag{4.6}$$

Notice that (4.6) holds for any $U \in \mathcal{P}$ and any $N \in J(U^T \hat{X})$. We conclude by applying Theorem 2.2 that

$$\frac{1}{\rho(\mathcal{P}) + \|\delta X\|} \leq \sup_{W \in \mathcal{P}} \|(W^{\frac{1}{2}} \hat{X})^+ W^{\frac{1}{2}}\| \leq \frac{1}{\rho(\mathcal{P}) - \|\delta X\|} \leq \frac{1}{(1 - \alpha) \cdot \rho(\mathcal{P})},$$

proving (4.5). Because the above estimate holds for any $\hat{X} = X + \delta X \in C_r^{m \times n}$ with $\|\delta X\| < \alpha \cdot \rho(\mathcal{P})$, $(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}$ is stable over \mathcal{P} . \square

In the following we specify two special cases.

Corollary 4.1 [21, Theorem 3.1]. *If in Theorem 4.1,*

$$\mathcal{P} = \{W = \text{diag}(d_1, \dots, d_m) : 0 \leq W \in \mathcal{P}(X)\},
 \tag{4.7}$$

then $(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}$ is stable over \mathcal{P} , if and only if

$$\text{any } r \text{ rows of } X \text{ are linearly independent.} \tag{4.8} \quad \square$$

Corollary 4.2. *If in Theorem 4.1,*

$$\mathcal{P} = \{W \in \mathcal{P}(X) : w_{ii} \geq \sum_{j=1, j \neq i}^m |w_{ij}|\},
 \tag{4.9}$$

then $(W^{\frac{1}{2}} X)^+ W^{\frac{1}{2}}$ is stable over \mathcal{P} , if and only if the following condition holds:

Let e_i^T and x_i respectively be the i -th row of I_m and X for $i = 1, \dots, m$. Let U_J^T and $U_J^T X$ have the forms

$$U_J^T = \begin{pmatrix} e_{i_1}^T \\ \vdots \\ e_{i_q}^T \\ e_{j_1}^T \pm e_{k_1}^T \\ \vdots \\ e_{j_{r-q}}^T \pm e_{k_{r-q}}^T \end{pmatrix}, \quad U_J^T X = \begin{pmatrix} x_{i_1} \\ \vdots \\ x_{i_q} \\ x_{j_1} \pm x_{k_1} \\ \vdots \\ x_{j_{r-q}} \pm x_{k_{r-q}} \end{pmatrix},
 \tag{4.10}$$

in which q is any integer satisfying $q \leq r$, $1 \leq i_1 < \dots < i_q \leq m$ and $1 \leq j_t < k_t \leq m$ for $t = 1, \dots, r - q$. Then

$$\text{rank}(U_J^T) = r \text{ implies } \text{rank}(U_J^T X) = r.
 \tag{4.11}$$

Proof. Under the condition of the corollary, any $W \in \mathcal{P}$ has the signature decomposition $W = UDU^T$ [4], where $U \in R^{m \times l}$ ($l = \frac{m(m+1)}{2}$) has the form in (2.19a) and D has the form in (2.19b). Let a vector of indexes J be of the form $J = \{i_1, \dots, i_r\}$ with $1 \leq i_1 < \dots < i_r \leq l$. Then U_J^T and $U_J^T X$ respectively have the forms in (4.10). So by applying Theorem 4.1 we obtain the condition in (4.11). \square

Now we consider the stability condition for $\lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ arising from solving the LSE problem (1.4) by the weighting method.

Suppose that

(4.12a)

$$L \in C_p^{m_1 \times n}, K \in C^{m_2 \times n}, X = \begin{pmatrix} L \\ K \end{pmatrix} \in C_r^{m \times n} \text{ and } W(\tau) = \begin{pmatrix} \tau I_{m_1} & \\ & W \end{pmatrix},$$

where $m = m_1 + m_2$, $W \in \mathcal{P}(KP)$, and

$$(4.12b) \quad \hat{L} = L + E_L, \hat{K} = K + E_K \text{ and } \hat{X} = \begin{pmatrix} \hat{L} \\ \hat{K} \end{pmatrix} \equiv X + E_X$$

are the perturbed versions of L , K and X , respectively. From (3.5) and (3.7) we see that for given $W \in \mathcal{P}(KP)$, $\lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}\hat{X})^+W(\tau)^{\frac{1}{2}}$ is bounded for all small perturbations E_L , E_K and E_X ; we need conditions

$$(4.13) \quad \text{rank}(\hat{L}) = \text{rank}(L) = p \text{ and } \text{rank}(\hat{X}) = \text{rank}(X) = r.$$

Now we define the stability of $\lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ over a set $\mathcal{P}_2 \subset \mathcal{P}(KP)$.

Definition 4.2. Let L , K , X be given as in (4.12a) and $\mathcal{P}_2 \subset \mathcal{P}(KP)$ be a given set of matrices. For any $W \in \mathcal{P}_2$, $W(\tau)$ is defined in (4.12a). Suppose that

$$(4.14a) \quad \frac{1}{\rho(\mathcal{P}_2)} \equiv \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}\| < +\infty.$$

We say that $\lim_{\tau \rightarrow 0+} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ is stable over \mathcal{P}_2 , if there exists a constant $\eta > 0$ which does not depend on the choice of $W \in \mathcal{P}_2$, such that

$$(4.14b) \quad \sup_{\substack{X+E_X \in C_r^{m \times n}, \|E_X\| \leq \eta \\ L+E_L \in C_p^{m_1 \times n}}} \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}(X + E_X))^+W(\tau)^{\frac{1}{2}}\| < +\infty.$$

We now state the following theorem.

Theorem 4.2. Suppose that L , K , X are as in (4.12a), \mathcal{U}_2 satisfies (3.15) and \mathcal{P}_2 is defined in (3.17). Then $\lim_{\tau \rightarrow 0+} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ is stable over \mathcal{P}_2 , if and only if for any $U \in \mathcal{U}_2$ and any set of indexes $J = \{i_1, \dots, i_{r-p}\}$ with $1 \leq i_1 < \dots < i_{r-p} \leq m_2$,

$$(4.15a) \quad \text{rank}(U_J^T) = r - p \text{ implies } \text{rank}(U_J^T KP) = r - p,$$

which is equivalent to

$$(4.15b) \quad \text{rank}(U_J^T) = r - p \text{ implies } \text{rank} \begin{pmatrix} L \\ U_J^T K \end{pmatrix} = r.$$

Furthermore, if the condition in (4.15a) holds, then for any \hat{L} , \hat{K} and \hat{X} which are defined in (4.12b) and satisfy the conditions in (4.13), such that $\|E_X\| \leq a \cdot \rho(\mathcal{P}_2) \equiv \eta$ with $0 < a < 1$ some constant, we have the following estimate:

$$(4.16) \quad \frac{1}{\rho(\mathcal{P}_2) + \|E_X\|} \leq \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}\hat{X})^+W(\tau)^{\frac{1}{2}}\| \leq \frac{1}{\rho(\mathcal{P}_2) - \|E_X\|}.$$

Proof. Necessity. If there exists a matrix $U \in \mathcal{U}_2$ and a set of indexes $N = \{i_1, \dots, i_{r-p}\}$ such that $\text{rank}(U_N^T) = r - p$ but $\text{rank}(U_N^T K P) = q < r - p$, then for any $\epsilon > 0$, we can add into K a small perturbation E_K such that $\|E_K\| < \epsilon \cdot \rho(\mathcal{P}_2)$ and $\text{rank}(U_N^T \hat{K} P) = \text{rank}(\hat{K} P) = r - p$. Let $X_\epsilon = \begin{pmatrix} L \\ \hat{K} \end{pmatrix}$, then $\text{rank}(X_\epsilon) = \text{rank}(X) = r$ and $N \in J(U^T \hat{K} P)$. So by applying Theorem 3.1 we have

$$\begin{aligned} \inf_{U \in \mathcal{U}_2} \min_{J \in J(U^T \hat{K} P)} \inf_+ \left(\begin{matrix} L \\ (U_J^T)^+ U_J^T \hat{K} \end{matrix} \right) &\leq \sigma_r \left(\begin{matrix} L \\ (U_N^T)^+ U_N^T \hat{K} \end{matrix} \right) \\ &\leq \sigma_r \left(\begin{matrix} L \\ (U_N^T)^+ U_N^T K \end{matrix} \right) + \|(U_N^T)^+ U_N^T E_K\| \leq \|E_K\| < \epsilon \rho(\mathcal{P}_2), \end{aligned}$$

because in this case $\text{rank} \left(\begin{matrix} L \\ (U_N^T)^+ U_N^T K \end{matrix} \right) = \text{rank}(L) + \text{rank}((U_N^T)^+ U_N^T K P) < r$ by applying Lemma 1.1. So according to Theorem 3.1, for any $0 < \eta \ll 1$,

$$\begin{aligned} &\sup_{\substack{X + E_X \in C_r^{m \times n}, \|E_X\| \leq \eta \\ L + E_L \in C_p^{m_1 \times n}}} \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}}(X + E_X))^+ W(\tau)^{\frac{1}{2}}\| \\ &\geq \lim_{\epsilon \rightarrow 0^+} \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}} X_\epsilon)^+ W(\tau)^{\frac{1}{2}}\| \geq \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon \rho(\mathcal{P}_2)} = +\infty. \end{aligned}$$

So when the condition of the theorem does not hold, then $\lim_{\tau \rightarrow 0^+} (W(\tau)^{\frac{1}{2}} X)^+ W(\tau)^{\frac{1}{2}}$ is not stable over \mathcal{P}_2 . *Sufficiency.* Suppose that the condition in (4.15a) holds. Let a be a constant satisfying $0 < a < 1$ and let $\eta \equiv a \cdot \rho(\mathcal{P}_2)$. Let $\hat{L} = L + E_L$, $\hat{K} = K + E_K$, $\hat{X} = L + E_X$ be any perturbed matrices satisfying conditions in (4.13) with $\|E_X\| \leq \eta$. Then for any $U \in \mathcal{U}_2$ and $N \in J(U^T \hat{K} \hat{P})$, $\text{rank}(U_N^T) = \text{rank}(U_N^T \hat{K} \hat{P}) = r - p$ and so $\text{rank}(U_N^T K P) = r - p$. We then have from the perturbation of the singular values [7], [16],

$$\begin{aligned} (4.17) \quad \sigma_r \left(\begin{matrix} \hat{L} \\ (U_N^T)^+ U_N^T \hat{K} \end{matrix} \right) &\geq \sigma_r \left(\begin{matrix} L \\ (U_N^T)^+ U_N^T K \end{matrix} \right) - \left\| \begin{matrix} E_L \\ (U_N^T)^+ U_N^T E_K \end{matrix} \right\| \\ &\geq \rho(\mathcal{P}_2) - \|E_X\| \geq (1 - a)\rho(\mathcal{P}_2) > 0, \\ \sigma_r \left(\begin{matrix} \hat{L} \\ (U_N^T)^+ U_N^T \hat{K} \end{matrix} \right) &\leq \sigma_r \left(\begin{matrix} L \\ (U_N^T)^+ U_N^T K \end{matrix} \right) + \|E_X\|. \end{aligned}$$

Notice that (4.17) holds for any $U \in \mathcal{P}_2$ and any $N \in J(U^T \hat{K} \hat{P})$. We conclude by applying Theorem 3.1 that

$$\begin{aligned} \frac{1}{\rho(\mathcal{P}_2) + \|E_X\|} &\leq \sup_{W \in \mathcal{P}_2} \lim_{\tau \rightarrow +\infty} \|(W(\tau)^{\frac{1}{2}} \hat{X})^+ W(\tau)^{\frac{1}{2}}\| \\ &\leq \frac{1}{\rho(\mathcal{P}_2) - \|E_X\|} \leq \frac{1}{(1 - a)\rho(\mathcal{P}_2)}, \end{aligned}$$

proving (4.16). Because the above estimate holds for any \hat{L} , \hat{K} , \hat{X} satisfying (4.13) and $\|E_X\| \leq a \cdot \rho(\mathcal{P}_2)$, $\lim_{\tau \rightarrow 0^+} (W(\tau)^{\frac{1}{2}} X)^+ W(\tau)^{\frac{1}{2}}$ is stable over \mathcal{P}_2 . The equivalence of (4.15a) and (4.15b) is supported by Lemma 1.1. \square

In the following we specify two special cases.

Corollary 4.3. *If in Theorem 4.2,*

$$(4.18) \quad \mathcal{P}_2 = \{W = \text{diag}(d_1, \dots, d_{m_2}) : 0 \leq W \in \mathcal{P}(KP)\},$$

then $\lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ is stable over \mathcal{P}_2 , if and only if for any vector of $r - p$ indexes $J = \{i_1, \dots, i_{r-p}\}$ with $1 \leq i_1 < \dots < i_{r-p} \leq m_2$,

$$(4.19a) \quad \text{rank}(K_J P) = r - p,$$

which is equivalent to

$$(4.19b) \quad \text{rank} \begin{pmatrix} L \\ K_J \end{pmatrix} = r.$$

Proof. In this case $\mathcal{U} = \{I_{m_2}\}$. So for any vector of indexes $J = \{i_1, \dots, i_{r-p}\}$ with $1 \leq i_1 < \dots < i_{r-p} \leq m_2$, U_J^T is of full row rank, and $U_J^T K P = K_J P$. Then by applying Theorem 4.2 we immediately reach the assertion in (4.19a). The equivalence of (4.19a) and (4.19b) is supported by Lemma 1.1. \square

Corollary 4.4. *If in Theorem 4.2,*

$$(4.20) \quad \mathcal{P}_2 = \{W \in \mathcal{P}(KP) : w_{ii} \geq \sum_{j=1, j \neq i}^{m_2} |w_{ij}|\},$$

then $\lim_{\tau \rightarrow +\infty} (W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ is stable over \mathcal{P}_2 , if and only if the following condition holds:

Let e_i^T and k_i respectively be the i -th row of I_{m_2} and K for $i = 1, \dots, m_2$. Let U_J^T and $U_J^T K$ have the forms

$$(4.21) \quad U_J^T = \begin{pmatrix} e_{i_1}^T \\ \vdots \\ e_{i_q}^T \\ e_{j_1}^T \pm e_{l_1}^T \\ \vdots \\ e_{j_{r-p-q}}^T \pm e_{l_{r-p-q}}^T \end{pmatrix}, \quad U_J^T K = \begin{pmatrix} k_{i_1} \\ \vdots \\ k_{i_q} \\ k_{j_1} \pm k_{l_1} \\ \vdots \\ k_{j_{r-p-q}} \pm k_{l_{r-p-q}} \end{pmatrix},$$

in which q is any integer satisfying $q \leq r - p$, $1 \leq i_1 < \dots < i_q \leq m_2$ and $1 \leq j_t < l_t \leq m_2$ for $t = 1, \dots, r - p - q$. Then

$$(4.22a) \quad \text{rank}(U_J^T) = r - p \quad \text{implies} \quad \text{rank}(U_J^T K P) = r - p,$$

which is equivalent to

$$(4.22b) \quad \text{rank}(U_J^T) = r - p \quad \text{implies} \quad \text{rank} \begin{pmatrix} L \\ U_J^T K \end{pmatrix} = r.$$

Proof. Under the condition of the corollary, any $W \in \mathcal{P}_2$ has the signature decomposition $W = UDU^T$ [4], where $U \in R^{m_2 \times l}$ ($l = \frac{m_2(m_2+1)}{2}$) has the form in (3.33) and D has the form in (3.32). Let any vector of indexes J be of the form $J = \{i_1, \dots, i_{r-p}\}$ with $1 \leq i_1 < \dots < i_{r-p} \leq l$. Then U_J^T and $U_J^T K$ respectively have the forms in (4.21). So by applying Theorem 4.2 we obtain the equivalent conditions in (4.22a) and (4.22b). \square

Remark. As mentioned at the beginning of this section, for a given matrix X , rank preserving of perturbed matrices $\hat{X} = X + \delta X$ guarantees boundedness of $\|\hat{X}^+\|$ for all small perturbations δX satisfying $\|\delta X\| \cdot \|X^+\| \leq \eta < 1$ for any constant $0 < \eta < 1$.

From the discussion of this section, we see that when analyzing stability of weighted pseudoinverses of X , rank preserving of perturbed matrices \hat{X} is not adequate to guarantee the stability of weighted pseudoinverses of X over a set \mathcal{P} or \mathcal{P}_2 . We need some extra structural conditions on the matrix X . So when analyzing stability properties of interior methods, one should take special care.

5. CONCLUDING REMARKS

In this paper we have discussed equivalent formulae and the stability properties of weighted pseudoinverses $(W(\tau)^{\frac{1}{2}}X)^+W(\tau)^{\frac{1}{2}}$ over a set $\mathcal{P} \subseteq \mathcal{P}(X)$ when the supremum is finite. We have also considered several specific cases when the sets of weighting matrices are semi-positive diagonal, real symmetric diagonally dominant positive semidefinite matrices, and weighting matrices arising from solving the LSE problem. The results generalize those in [15], [12], [21], [4].

We have also studied algebraic properties of the weighted pseudoinverses arising from solving the LSE problem, which can be used to further study the LSE problem.

ACKNOWLEDGEMENT

The author is grateful to the editor and referees for useful suggestions and comments.

REFERENCES

1. A. Ben-Israel and T.N.E. Greville, *Generalized Inverses: Theory and Applications*, John Wiley, New York, 1974. MR **81h**:15005
2. A. Ben-Tal and M. Teboulle, *A geometric property of the least squares solution of linear equations*, Linear Algebra Appl. **139** (1990), 165-170. MR **91h**:15004
3. I.I. Dikin, *On the speed of an iterative process*, Upravlyaemye Sistemi **12** (1974), 54-60.
4. A. Forsgren, *On linear least-squares problems with diagonally dominant weight matrices*, SIAM J. Matrix Anal. Appl. **17** (1996), 763-788. CMP 97:01
5. A. Forsgren, P.E. Gill and J.R. Shinnerl, *Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization*, SIAM J. Matrix Anal. Appl. **17** (1996), 187-211. MR **96m**:90084
6. P.E. Gill, W. Murray and M.H. Wright, *Numerical Linear Algebra and Optimization, Vol. 1*, Addison-Wesley Publishing Company, Redwood City, 1991. MR **92b**:65001
7. G.H. Golub and C.F. Van Loan, *Matrix Computations*, 2nd Edit., The Johns Hopkins University Press, Baltimore, MD, 1989. MR **90d**:65055
8. C.C. Gonzaga, *Path-following methods for linear programming*, SIAM Review **34** (1992), 167-224. MR **93j**:90050
9. M. Gullikson and P.-Å. Wedin, *Modifying the QR decomposition to constrained and weighted linear least squares*, SIAM J. Matrix Anal. Appl. **13** (1992), 1298-1313. MR **93e**:65066
10. M.Hanke and M. Neumann, *The geometry of the set of scaled projections*, Linear Algebra Appl. **190** (1993), 137-148. MR **94g**:15002
11. R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985. MR **87e**:15001
12. D. P. O'Leary, *On bounds for scaled projections and pseudoinverses*, Linear Algebra Appl. **132** (1990), 115-117. MR **91f**:15056
13. J. Miao and A. Ben-Israel, *The geometry of basic, approximate, and minimum-norm solution of linear equations*, Linear Algebra Appl. **216** (1995), 25-41. MR **96e**:15003
14. C.C. Paige and M.A. Saunders, *Towards a generalized singular value decomposition*, SIAM J. Numer. Anal. **18** (1981), 398-405. MR **83c**:65082

15. G.W. Stewart, *On scaled projections and pseudoinverses*, Linear Algebra Appl. **112** (1989), 189-193. MR **90d**:15005
16. G.W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990. MR **92a**:65017
17. R.J. Vanderbei and J.C. Lagarias, *Dikin's convergence result for the affine-scaling algorithm*, Contemporary Mathematics **114** (1990), 109-119. MR **92d**:90046
18. S.A. Vavasis, *Stable numerical algorithms for equilibrium system*, SIAM J. Matrix Anal. Appl. **15** (1994), 1108-1131. MR **95m**:65080
19. M. Wei, *Algebraic properties of the rank- deficient equality constrained and weighted least squares problems*, Linear Algebra Appl. **161** (1992), 27-43. MR **92j**:65059
20. M. Wei, *Perturbation theory for the rank- deficient equality constrained least squares problem*, SIAM J. Numer. Anal. **29** (1992), 1462-1481. MR **93i**:65046
21. M. Wei, *Upper bound and stability of scaled pseudoinverses*, Numer. Math. **72** (1995), 285-293. MR **96k**:65031
22. M.H. Wright, *Interior methods for constrained optimization*, In A. Iserles, edit., Acta Numerica, pp.341-407, Cambridge University Press, Cambridge, UK, 1992. MR **93d**:90037
23. S. Wright, *Stability of linear equations solvers in interior-point methods*, SIAM J. Matrix Anal. Appl. **16** (1995), 1287-1307. MR **96f**:65055
24. S. Wright, *Stability of linear algebra computations in interior-point methods for linear programming*, Tech. Rep. MCS- P400- 1293, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, 1993.

DEPARTMENT OF MATHEMATICS, EAST CHINA NORMAL UNIVERSITY, SHANGHAI 200062, CHINA