ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 206 (2007) 1051 – 1069
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Vector least-squares solutions for coupled singular matrix equations

Adem Kılıçman*, Zeyad Abdel Aziz Al Zhour

Department of Mathematics and Institute for Mathematical Research, University Putra Malaysia (UPM), 43400, Serdang, Selangor, Malaysia

Received 28 August 2006; received in revised form 18 September 2006

Abstract

The weighted least-squares solutions of coupled singular matrix equations are too difficult to obtain by applying matrices decomposition. In this paper, a family of algorithms are applied to solve these problems based on the Kronecker structures. Subsequently, we construct a computationally efficient solutions of coupled restricted singular matrix equations. Furthermore, the need to compute the weighted Drazin and weighted Moore–Penrose inverses; and the use of Tian's work and Lev-Ari's results are due to appearance in the solutions of these problems. The several special cases of these problems are also considered which includes the well-known coupled Sylvester matrix equations. Finally, we recover the iterative methods to the weighted case in order to obtain the minimum D-norm G-vector least-squares solutions for the coupled Sylvester matrix equations and the results lead to the least-squares solutions and invertible solutions, as a special case.

MSC: 15A24; 15A69; 15A09

Keywords: Kronecker products; Matrix least-squares problems; Generalized inverses; Matrix norms; Iterative methods

1. Introduction and preliminary results

Let us recall some concepts that we will use in this study and throughout the study we consider matrices over the field of complex numbers $\mathbb{C}$ or real numbers $\mathbb{R}$. The set of $m \times n$ complex matrices is denoted by $M_{m,n}(\mathbb{C}) = \mathbb{C}^{m\times n}$. For simplicity we write $M_{m,n}$ instead of $M_{m,n}(\mathbb{C})$ and when $m = n$, we write M_n instead of $M_{n,n}$. The notations A^{T}, A^*, A^{-1}, $r(A)$, $R(A)$ and $\mathrm{tr}(A)$ stand for the transpose, conjugate transpose, inverse, rank, range and trace of matrix A, respectively.

Given two matrices $A = [a_{ij}] \in M_{m,n}$ and $B = [b_{kl}] \in M_{p,q}$, then the Kronecker product of A and B is defined by (see, e.g., [2,3,12,17,26,38])

$$A \otimes B = [a_{ij}B]_{ij} \in M_{mn,pq}. \tag{1.1}$$

While if $A = [a_1\ a_2\ \cdots\ a_n] \in M_{m,n}$ and $B = [b_1\ b_2\ \cdots\ b_n] \in M_{p,n}$ are matrices (where a_i and b_i are the ith columns of A and B, respectively, $i = 1, 2, \ldots, n$), then the columns of $A \otimes B$ are $\{a_i \otimes b_j\}$ for all i and j, that is

$$A \otimes B = [a_1 \otimes b_1\ \cdots\ a_1 \otimes b_n\ \cdots\ a_n \otimes b_1\ \cdots\ a_n \otimes b_n] \in M_{mp,n^2}. \tag{1.2}$$

* Corresponding author. Tel.: +60 3 89466813; fax: +60 3 89437958.
E-mail addresses: akilic@fsas.upm.edu.my (A. Kılıçman), zeyad1968@yahoo.com (Z.A.A.A. Zhour).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2006.09.009

Thus, the Khatri–Rao product of A and B is defined by (see, [5,18,26])

$$A \oslash B = [a_1 \otimes b_1 \ \ a_2 \otimes b_2 \ \ \cdots \ \ a_n \otimes b_n] \in M_{mp,n} \tag{1.3}$$

consists of a subset of the columns of $A \otimes B$. Additionally, if both matrices $A = [a_{ij}]$ and $B = [b_{ij}] \in M_{m,n}$ have the same size, then the Hadamard product of A and B is defined by (see, e.g., [2,3,17,26,38])

$$A \circ B = [a_{ij} b_{ij}]_{ij} \in M_{m,n}. \tag{1.4}$$

This product is much simpler than Kronecker and Khatri–Rao products and it can be connected with isomorphic diagonal matrix representations. The fundamental relationship between Kronecker, Khatri–Rao and Hadamard products can be expressed as follows, see [18]

$$P_n^{\mathrm{T}}(A \oslash B) = P_n^{\mathrm{T}}(A \otimes B)P_n = A \circ B, \tag{1.5}$$

where the selection matrix P_n is given by

$$P_n = [e_1 \ \ e_{n+2} \ \ e_{2n+3} \ \ \cdots \ \ e_{n^2}] \in M_{n^2,n}, \tag{1.6}$$

and e_k is an n^2-column vector with a unity element in the kth position and zeros elsewhere ($1 \leqslant k \leqslant n^2$). Moreover, the columns of a selection matrix P_n are mutually orthonormal, that is, $P_n^{\mathrm{T}} P_n = I_n$.

The mentioned three matrix products and some vector operators affirming their capability of solving some matrix equations. Such equations can be readily converted into the standard linear equation form by using the well-known identities (e.g., [10,12,14,18,39])

$$\mathrm{Vec}(AXB^{\mathrm{T}}) = (B \otimes A)\mathrm{Vec}(X), \tag{1.7}$$

$$\mathrm{Vec}(AXB^{\mathrm{T}}) = (B \oslash A)\mathrm{Vecd}(X) : X \text{ is diagonal}, \tag{1.8}$$

$$\mathrm{Vecd}(AXB^{\mathrm{T}}) = (B \circ A)\mathrm{Vecd}(X) : X \text{ is diagonal}. \tag{1.9}$$

where $\mathrm{Vec}(X) = [x_{11} \ \cdots \ x_{m1} \ \ x_{12} \ \cdots \ x_{m2} \ \cdots \ x_{1m} \ \cdots \ x_{mn}]^{\mathrm{T}} \in M_{mn,1}$ denotes vectorization by columns of arbitrary matrix $X \in M_{m,n}$, and $\mathrm{Vecd}(X) = [x_{11} \ \ x_{22} \ \cdots \ x_{nn}]^{\mathrm{T}} \in M_{n,1}$ denotes vectorization by diagonal elements of a square matrix $X \in M_n$.

The weighted generalized inverses of an arbitrary matrix (including singular and rectangular) are very useful in various applications such as control system analysis, statistics, singular differential and difference equations, Markov chains, iterative methods, generalized least-squares problem, weighted perturbation theory, neural networks problem and many other subjects that can be found in the literature e.g. [4,6,8,24,25,27,31–37,39]. Here we study the following weighted generalized inverses:

(a) The *weighted Moore–Penrose inverse* (WMPI) of a matrix $A \in M_{m,n}$ with respect to the two positive definite matrices $M \in M_m$ and $N \in M_n$ is defined to be the unique solution of the following four matrix equations:

$$AXA = A, \quad XAX = X, \quad (MAX)^* = MAX, \quad (NXA)^* = NXA, \tag{1.10}$$

and is often denoted by $X = A^+_{M,N}$, In particular, when $M = I_m$ and $N = I_n$, then $A^+_{M,N}$ is reduced to the Moore–Penrose inverse (MPI) A^+, while if A is square and non-singular matrix, then A^+ is reduce to A^{-1}. The important properties related to WMPI of matrix $A \in M_{m,n}$ might be given by (see, [30,31]):

(i)

$$(A^+_{M,N})^* = (A^*)^+_{N^{-1},M^{-1}}. \tag{1.11}$$

(ii) If A has full column-rank, then

$$A^+_{M,N} = (A^* M A)^{-1} A^* M. \tag{1.12}$$

(iii) If A has full row-rank, then

$$A^+_{M,N} = N^{-1} A^* (A N^{-1} A^*)^{-1}. \tag{1.13}$$

(b) The *weighted Drazin inverse* (WDI) of a matrix $A \in M_{m,n}$ with respect to the matrix $W \in M_{n,m}$ is defined to be the unique solution $X \in M_{m,n}$ of the following three matrix equations [8]:

$$(AW)^{k+1}XW = (AW)^k, \quad XWAWX = X, \quad AWX = XWA, \tag{1.14}$$

where

$$k = \max\{\mathrm{Ind}(AW), \mathrm{Ind}(WA)\}, \tag{1.15}$$

and is often denoted by $X = A_{d,W}$. In particular, when A is a square matrix of order $m \times m$ and $W = I_m$, then $A_{d,W}$ is reduced to the Drazin inverse (DI) A_d. Note that for a matrix $A \in M_m$, $\mathrm{Ind}(A) = k$ is the smallest positive integer such that

$$r(A^{k+1}) = r(A^k). \tag{1.16}$$

For any compatible matrices A, B, C and D, we shall frequently use the following properties of Kronecker product (see, e.g., [12,17]):

(i) If AC and BD are well defined, then

$$(A \otimes B)(C \otimes D) = AC \otimes BD. \tag{1.17}$$

(ii) For any natural number s,

$$(A \otimes B)^s = A^s \otimes B^s. \tag{1.18}$$

(iii)

$$r(A \otimes B) = r(A)r(B). \tag{1.19}$$

(iv) If A and B are positive definite matrices, then $A \otimes B$ is a positive definite.

(v) If $A \in M_{m,n}$ and $B \in M_{p,q}$ are matrices with respect to the positive definite matrices $M \in M_m$, $N \in M_n$, $L \in M_p$ and $Q \in M_q$, then [30]

$$(A \otimes B)^+_{F,G} = A^+_{M,N} \otimes B^+_{L,Q}, \tag{1.20}$$

where $F = M \otimes L$ and $G = N \otimes Q$.

It is well known that the *weighted matrix Frobenius norm* (WMFN) of a matrix $A \in M_{m,n}$ with respect to positive definite matrices $M \in M_m$ and $N \in M_n$ is given by Wang [31] as follows:

$$\|A\|_{M,N} = \|M^{1/2}AN^{-1/2}\|_2, \tag{1.21}$$

where $\|P\|_2 = (\mathrm{tr}(P^*P))^{1/2}$ is called the *Frobenius norm* of matrix $P \in M_{m,n}$. Many scientific applications gave rise to the *weighted least-squares problem* (WLSP):

$$\min_x \|Ax - b\|_M, \tag{1.22}$$

where $\|y\|_M = \|M^{1/2}y\|_2$ is called the *weighted vector Frobenius norm* (WVFN) of $y \in \mathbb{C}^m$ with respect to positive definite matrix $M \in M_m$. Generally speaking, the WLSP has multiple solution. In such a case, e.g., Wang [31] considered that the unique *minimum* N-norm M-least-squares solution (or weighted solution) of (1.22) as follows:

$$x = A^+_{M,N}b. \tag{1.23}$$

In the next we need to compute the A^+, $A^+_{M,N}$, A_d, $A_{d,W}$ and A^{-1}, and use the following two lemmas which was given by Tian [27] that are due to appearance in the solutions of coupled least-squares problems (LSPs).

Lemma 1. *If the block matrix* $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in M_{m,n}$ *satisfies the following five conditions*:

$$r(A) = r(A_{11}) + r(A_{22}), \quad R(A_{12}) \subseteq R(A_{11}), \quad R(A_{21}) \subseteq R(A_{22}),$$

$$R(A^*_{21}) \subseteq R(A^*_{11}), R(A^*_{12}) \subseteq R(A^*_{22}).$$

Then the MPI of A can be expressed as

$$A^+ = \begin{bmatrix} S^+_{A_{22}} & -A^+_{11}A_{12}S^+_{A_{11}} \\ -S^+_{A_{11}}A_{21}A^+_{11} & S^+_{A_{11}} \end{bmatrix}, \tag{1.24}$$

where $S_{A_{22}} = A_{11} - A_{12}A^+_{22}A_{21}$ *and* $S_{A_{11}} = A_{22} - A_{21}A^+_{11}A_{12}$ *are the Schur complements of* A_{22} *and* A_{11}, *respectively, in A*.

Lemma 2. *If the block matrix* $A = \begin{bmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{bmatrix} \in M_{m,n}$ *satisfies the following five conditions*:

$$r(A) = r(A_{11}) + r(A_{22}), \quad R(A_{12}) = R(A_{11}), \quad R(A_{21}) = R(A_{22}),$$
$$R(A^*_{21}) = R(A^*_{11}), R(A^*_{12}) = R(A^*_{22}).$$

Then the MPI of A can be expressed as

$$A^+ = \begin{bmatrix} S^+_{A_{22}} & S^+_{A_{12}} \\ S^+_{A_{21}} & S^+_{A_{11}} \end{bmatrix} = \begin{bmatrix} (A_{11} - A_{12}A^+_{22}A_{21})^+ & (A_{21} - A_{22}A^+_{12}A^+_{11})^+ \\ (A_{12} - A_{11}A^+_{21}A^+_{22})^+ & (A_{22} - A_{21}A^+_{11}A_{12})^+ \end{bmatrix}. \tag{1.25}$$

The coupled matrix equations have been widely used in several areas such as stability theory, control theory, communication systems, perturbation analysis, economics and many other fields of pure and applied mathematics; and recently it is in the context of the analysis and numerical simulation of descriptor systems therefore many interesting problems can lead to coupled matrix equations. For example, the non-zero sum differential games, optimal control system, Sylvester matrix equations, matrix Schrödinger equations, axial *N*-index transportation problems and state-space equations which were discussed, respectively, by Cruz et al. [9], Zhang [39], Ding and Chen [10], Carlson [7], Al Zhour and Kılıçman [1] and Mouroutsos and Sparis [21].

Research on solving systems of linear matrix equations has also been active for past years, for example, the conditions for the existence of a solution and a representation of the general common solution to the matrix equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$ were provided in [20], a representation of the general common solution to the matrix equations $A_1XB_1 = C_1$; $A_2XB_2 = C_2$ were also studied by Navarra et al. [22], the existence of a common solution X to the matrix equations $A_iXB_j = C_{ij}$, $(i, j) \in \Gamma$ were obtained by van der Woude [28], the iterative method for symmetric solutions and optimal approximation of the system matrix equations $A_1XB_1 = C_1$; $A_2XB_2 = C_2$ were also presented in [23] the nearest Kronecker product problems are solved in [29], the perturbation for the constrained and weighted least-squares problems were derived by Gulliksson et al. [13], and the solutions of coupled matrix convolution and matrix differential equations were studied by Kılıçman and Al Zhour [14–16]. Finally, Fulton and Wu [11] described an implementation of the matrices decompositions such as *QR*, *SVD*, *LU*, and Van Loan [29] mentioned also that the linear system of the form $(A \otimes B)x = c$ can be solved fast by compared with matrices decompositions.

Depending on the problem in consideration, different coupling terms may appear. However, in several cases, it is difficult to find the weighted least-squares solutions by using matrix decomposition. In present paper, a family of coupled singular matrix equations are formulated and several algorithms for computing the weighted solutions of these coupled are also proposed by using the effective Kronecker structures. We construct a computationally efficient solution of coupled restricted singular matrix equations (RSME) and derive the representation for the WDI of the Kronecker product in order to find the weighted solution of RSME. We also consider several special cases which includes the well-known coupled Sylvester matrix equations.

For some applications such as stability analysis, it is often not necessary to compute exact solutions, approximate solutions are sufficient since sometimes computational efforts will rapidly increase with the size of matrix functions.

Here we recover the iterative methods given by the Ding and Chen [10] due to the weighted case in order to find the minimum D-norm G-vector least-squares solutions of the coupled Sylvester matrix equations.

2. Restricted singular matrix equations

First of all, we derive the representation for the WDI of the Kronecker product $A \otimes B$ as follows:

Theorem 3. *Let* $A \in M_{m,n}, W \in M_{n,m}, B \in M_{p,q}$ *and* $R \in M_{q,p}$ *be matrices, and let* $Z = W \otimes R$ *and* $k = \max\{k_1, k_2\}$ *such that*

$$k_1 = \max\{\text{Ind}(AW), \text{Ind}(WA)\}, \quad k_2 = \max\{\text{Ind}(BR), \text{Ind}(RB)\}.$$

Then

(i)

$$\text{Ind}\{(A \otimes B)Z\} = k, \tag{2.1}$$

(ii)

$$(A \otimes B)_{d,Z} = A_{d,W} \otimes B_{d,R}. \tag{2.2}$$

Proof. (i) By assumptions, we have

$$r(AW)^{k_1} = r(AW)^{k_1+1}, \quad r(BR)^{k_2} = r(BR)^{k_2+1}.$$

From properties of Kronecker products, we have

$$r((A \otimes B)Z)^s = r((A \otimes B)(W \otimes R))^s = r(AW \otimes BR)^s = r(AW)^r r(BR)^s.$$

Similarly,

$$r((A \otimes B)Z)^{s+1} = r(AW)^{s+1} r(BR)^{s+1}$$

It is obvious that the smallest non-negative integer that

$$r((A \otimes B)Z)^{s+1} = r((A \otimes B)Z)^s$$

is $k = \max\{k_1, k_2\}$. Hence (2.1) is true.

(ii) Let $X = A_{d,W} \otimes B_{d,R}$ and $Z = W \otimes R$. From properties of the Kronecker product and (1.14) we have

$$\begin{aligned}((A \otimes B)Z)^{k+1} XZ &= ((A \otimes B)(W \otimes R))^{k+1} (A_{d,W} \otimes B_{d,R})(W \otimes R) \\ &= ((AW)^{k+1} A_{d,W} W) \otimes ((BR)^{k+1} B_{d,R} R) \\ &= (AW)^k \otimes (BR)^k = (AW \otimes BR)^k = ((A \otimes B)(W \otimes R))^k \\ &= \{(A \otimes B)Z\}^k,\end{aligned} \tag{2.3}$$

$$\begin{aligned}XZ(A \otimes B)ZX &= (A_{d,W} \otimes B_{d,R})(W \otimes R)(A \otimes B)(W \otimes R)(A_{d,W} \otimes B_{d,R}) \\ &= (A_{d,W} WAWA_{d,W}) \otimes (B_{d,R} RBRB_{d,R}) = A_{d,W} \otimes B_{d,R} \\ &= X,\end{aligned} \tag{2.4}$$

$$\begin{aligned}(A \otimes B)ZX &= (A \otimes B)(W \otimes R)(A_{d,W} \otimes B_{d,R}) \\ &= AWA_{d,W} \otimes BRB_{d,R} = A_{d,W} WA \otimes B_{d,R} RB \\ &= (A_{d,W} \otimes B_{d,R})(W \otimes R)(A \otimes B) \\ &= XZ(A \otimes B).\end{aligned} \tag{2.5}$$

From (2.3)–(2.5) we can obtain (2.2) immediately. □

One of the important application of Theorem 3 is that the WDI of Kronecker product arise naturally in solving the so-called RSME as follows.

Theorem 4. *Let* $A \in M_{m,n}$, $W \in M_{n,m}$, $B \in M_{p,q}$, $R \in M_{q,p}$ *and* $C \in M_{n,q}$ *be given constant matrices and* $X \in M_{m,p}$ *be an unknown matrix to be solved. Also, let*

$$L = R \otimes W, \quad k_1 = \mathrm{Ind}((B \otimes A)L), \quad k_2 = \mathrm{Ind}(L(B \otimes A))$$

such that

$$r((B \otimes A)L)^{k_1} = r(L(B \otimes A))^{k_2}, \quad \mathrm{Vec}\, C \in R(L(B \otimes A))^{k_2}, \quad \mathrm{Vec}\, X \in R((B \otimes A)L)^{k_1}. \tag{2.6}$$

Then the unique solution of the following RSME

$$(WAW)X(RBR)^{\mathrm{T}} = C \tag{2.7}$$

is given by

$$X = A_{d,W} C B_{d,R}^{\mathrm{T}}. \tag{2.8}$$

Proof. On using the identity (1.7) it is not difficult to transform (2.7) into the following vector form:

$$(L(B \otimes A)L)\mathrm{Vec}\, X = \mathrm{Vec}\, C. \tag{2.9}$$

It is easy to verify under (2.6) that the unique solution of (2.9) is

$$\begin{aligned} \mathrm{Vec}\, X &= (B \otimes A)_{d,L} \mathrm{Vec}\, C = (B_{d,R} \otimes A_{d,W})\mathrm{Vec}\, C \\ &= \mathrm{Vec}(A_{d,W} C B_{d,R}^{\mathrm{T}}), \end{aligned}$$

which is the required result. □

In particular case, if $m = n$, $p = q$, $W = I_m$ and $R = I_p$, we obtain the following corollary:

Corollary 5. *Let* $A \in M_m$, $B \in M_p$ *and* $C \in M_{m,p}$ *be given constant matrices and* $X \in M_{m,p}$ *be an unknown matrix to be solved. Then the unique solution of the following RSME*:

$$AXB^{\mathrm{T}} = C : \mathrm{Vec}\, C, \quad \mathrm{Vec}\, X \in R(B \otimes A)^k, \quad k = \mathrm{Ind}(B \otimes A) \tag{2.10}$$

is given by

$$X = A_d C B_d^{\mathrm{T}}. \tag{2.11}$$

Another important case can be obtained from (1.7), (1.20), (1.11), (1.23) and Magnus and Neudecker's idea [19, Theorems 12 and 13, p. 37] as in the following corollary:

Corollary 6. *Let* $A \in M_{m,n}$, $B \in M_{p,q}$, $C \in M_{m,p}$ *be given constant matrices and* $X \in M_{n,q}$ *be an unknown matrix to be solved. Also, let* $M \in M_m$, $N \in M_n$, $L \in M_p$ *and* $Q \in M_q$ *be positive definite matrices, and* $G = M \otimes L$ *and* $D = N \otimes Q$. *Then the minimum D-norm G-least squares solution of* $AXB^{\mathrm{T}} = C$ *is given by*

$$X = A^{+}_{M,N} C (B^{+}_{L,Q})^{\mathrm{T}} = A^{+}_{M,N} C (B^{\mathrm{T}})^{+}_{Q^{-1},L^{-1}}. \tag{2.12}$$

Furthermore, a necessary and sufficient condition for the matrix equation $AXB^{\mathrm{T}} = C$ *in order to have a weighted solution can be given by*

$$AA^{+}_{M,N} C (BB^{+}_{L,Q})^{\mathrm{T}} = C, \tag{2.13}$$

in which case, the general minimum D-norm G-least squares solution is given by

$$X = A^{+}_{M,N} C (B^{+}_{L,Q})^{\mathrm{T}} + F - A^{+}_{M,N} A F (B^{+}_{L,Q} B)^{\mathrm{T}}, \tag{2.14}$$

where F is an arbitrary matrix of order $n \times q$.

In particular, the unique least-squares solution of $AXB^{\mathrm{T}} = C$ *can be given by*

$$X = A^{+}C(B^{+})^{\mathrm{T}}. \tag{2.15}$$

Now we note that if $A \in M_m$, $B \in M_p$ are non-singular matrices, then $\mathrm{Ind}(A) = \mathrm{Ind}(B) = 0$, $A_d = A^{+} = A^{-1}$ and the unique solution of $AXB^{\mathrm{T}} = C$ is given by

$$X = A^{-1}C(B^{-1})^{\mathrm{T}}. \tag{2.16}$$

3. Coupled restricted singular matrix equations

In this section, we will study the vector least-squares solutions of the so-called *coupled restricted singular matrix equations* (CRSME) based on Tian's Lemmas and Kronecker structures and we consider some special cases. First of all, we have the following theorem:

Theorem 7. *Let* $A_1, B_1, C_1, D_1, A_2, B_2, C_2$ *and* $D_2 \in M_{m,n}$ *be given full column-rank matrices such that*

$$R(D_1 \otimes C_1) \subseteq R(B_1 \otimes A_1), \quad R(B_2 \otimes A_2) \subseteq R(D_2 \otimes C_2),$$

$$R(B_2^{*} \otimes A_2^{*}) \subseteq R(B_1^{*} \otimes A_1^{*}), \quad R(D_1^{*} \otimes C_1^{*}) \subseteq R(D_2^{*} \otimes C_2^{*}).$$

Also, let $E, F \in M_m$ *be given constant matrices, and* $X, Y \in M_n$ *be unknown matrices to be solved. Then the vector least-squares solution of the following CRSME*:

$$A_1XB_1^{\mathrm{T}} + C_1YD_1^{\mathrm{T}} = E, \quad A_2XB_2^{\mathrm{T}} + C_2YD_2^{\mathrm{T}} = F \tag{3.1}$$

is given by

$$\begin{aligned} \mathrm{Vec}\,X = &-(B_1^{+}D_1 \otimes A_1^{+}C_1)(D_2 \otimes C_2 - B_2B_1^{+}D_1 \otimes A_2A_1^{+}C_1)^{+}\mathrm{Vec}\,F \\ &+ (B_1 \otimes A_1 - D_1D_2^{+}B_2 \otimes C_1C_2^{+}A_2)^{+}\mathrm{Vec}\,E, \end{aligned} \tag{3.2}$$

$$\begin{aligned} \mathrm{Vec}\,Y = &-(D_2 \otimes C_2 - B_2B_1^{+}D_1 \otimes A_2A_1^{+}C_1)^{+}(B_2B_1^{+} \otimes A_2A_1^{+})\mathrm{Vec}\,E \\ &+ (D_2 \otimes C_2 - B_2B_1^{+}D_1 \otimes A_2A_1^{+}C_1)^{+}\mathrm{Vec}\,F. \end{aligned} \tag{3.3}$$

Proof. The CRSME in (3.1) can easily be transformed to the following vector form:

$$\begin{bmatrix} B_1 \otimes A_1 & D_1 \otimes C_1 \\ B_2 \otimes A_2 & D_2 \otimes C_2 \end{bmatrix} \begin{bmatrix} \mathrm{Vec}\,X \\ \mathrm{Vec}\,Y \end{bmatrix} = \begin{bmatrix} \mathrm{Vec}\,E \\ \mathrm{Vec}\,F \end{bmatrix}. \tag{3.4}$$

Since $\begin{bmatrix} B_1 \otimes A_1 & D_1 \otimes C_1 \\ B_2 \otimes A_2 & D_2 \otimes C_2 \end{bmatrix}$ is a full-column rank, then the least-squares solution of (3.4) is

$$\begin{bmatrix} \mathrm{Vec}\,X \\ \mathrm{Vec}\,Y \end{bmatrix} = \begin{bmatrix} B_1 \otimes A_1 & D_1 \otimes C_1 \\ B_2 \otimes A_2 & D_2 \otimes C_2 \end{bmatrix}^{+} \begin{bmatrix} \mathrm{Vec}\,E \\ \mathrm{Vec}\,F \end{bmatrix}. \tag{3.5}$$

Now applying (1.24) of Lemma 1 and (1.20) into (3.5), we establish (3.2)–(3.3). □

Similarly by using (1.25) of Lemma 2 we obtain the following corollary:

Corollary 8. *Let* A_1, B_1,C_1 , D_1 , A_2, B_2, C_2 *and* $D_2 \in M_{m,n}$ *be given full column-rank matrices such that*

$$R(D_1 \otimes C_1) = R(B_1 \otimes A_1), \quad R(B_2 \otimes A_2) = R(D_2 \otimes C_2),$$

$$R(B_2^{*} \otimes A_2^{*}) = R(B_1^{*} \otimes A_1^{*}), \quad R(D_1^{*} \otimes C_1^{*}) = R(D_2^{*} \otimes C_2^{*}).$$

Also, if we let E and $F \in M_m$ be given constant matrices, and X, $Y \in M_n$ be unknown matrices to be solved then the vector least-squares solution of the following CRSME:

$$A_1 X B_1^{\mathrm{T}} + C_1 Y D_1^{\mathrm{T}} = E, \quad A_2 X B_2^{\mathrm{T}} + C_2 Y D_2^{\mathrm{T}} = F \tag{3.6}$$

is given by

$$\begin{aligned}\mathrm{Vec}\, X = {} & (B_1 \otimes A_1 - D_1 D_2^{+} B_2 \otimes C_1 C_2^{+} A_2)^{+} \mathrm{Vec}\, E \\ & + (B_2 \otimes A_2 - D_2 D_1^{+} B_1 \otimes C_2 C_1^{+} A_1)^{+} \mathrm{Vec}\, F,\end{aligned} \tag{3.7}$$

$$\begin{aligned}\mathrm{Vec}\, Y = {} & (D_1 \otimes C_1 - B_1 B_2^{+} D_2 \otimes A_1 A_2^{+} C_2)^{+} \mathrm{Vec}\, E \\ & + (D_2 \otimes C_2 - B_2 B_1^{+} D_1 \otimes A_2 A_1^{+} C_1)^{+} \mathrm{Vec}\, F.\end{aligned} \tag{3.8}$$

Corollary 9. *Let* A_1, B_1,C_1 , D_1 , A_2, B_2, C_2 and $D_2 \in M_n$ and E, $F \in M_n$ *be given constant matrices, and X, $Y \in M_n$ be unknown matrices to be solved. Then the vector least-squares solution of the following coupled matrix equations*:

$$A_1 X B_1^{\mathrm{T}} + C_1 Y D_1^{\mathrm{T}} = E, \quad A_2 X B_2^{\mathrm{T}} + C_2 Y D_2^{\mathrm{T}} = F \tag{3.9}$$

is given by

$$\begin{aligned}\mathrm{Vec}\, X = {} & -(B_1^{-1} D_1 \otimes A_1^{-1} C_1)(D_2 \otimes C_2 - B_2 B_1^{-1} D_1 \otimes A_2 A_1^{-1} C_1)^{-1} \mathrm{Vec}\, F \\ & + (B_1 \otimes A_1 - D_1 D_2^{-1} B_2 \otimes C_1 C_2^{-1} A_2)^{-1} \mathrm{Vec}\, E,\end{aligned} \tag{3.10}$$

$$\begin{aligned}\mathrm{Vec}\, Y = {} & -(D_2 \otimes C_2 - B_2 B_1^{-1} D_1 \otimes A_2 A_1^{-1} C_1)^{-1} (B_2 B_1^{-1} \otimes A_2 A_1^{-1}) \mathrm{Vec}\, E \\ & + (D_2 \otimes C_2 - B_2 B_1^{-1} D_1 \otimes A_2 A_1^{-1} C_1)^{-1} \mathrm{Vec}\, F\end{aligned} \tag{3.11}$$

assuming that all relevant inverses exist.

If we set $B_1 = C_1 = B_2 = C_2 = I$ in Corollary 8 we obtain the vector least-squares solution of coupled Sylvester matrix equations as follows:

Corollary 10. *Let A, B, C and $D \in M_{m,n}$ be given full column-rank matrices such that*

$$R(I \otimes A) = R(C \otimes I), \quad R(I \otimes B) = R(D \otimes I), \quad R(I \otimes A^*) = R(I \otimes B^*),$$

$$R(C^* \otimes I) = R(D^* \otimes I).$$

Also, if we let E and $F \in M_m$ be given constant matrices and X, $Y \in M_n$ be unknown matrices to be solved then the vector least-squares solutions of the coupled restricted singular Sylvester matrix equations

$$AX + YC^{\mathrm{T}} = E, \quad BX + YD^{\mathrm{T}} = F \tag{3.12}$$

can be given by

$$\mathrm{Vec}\, X = (I \otimes A - CD^{+} \otimes B)^{+} \mathrm{Vec}\, E + (I \otimes B - DC^{+} \otimes A)^{+} \mathrm{Vec}\, F, \tag{3.13}$$

$$\mathrm{Vec}\, Y = (C \otimes I - D \otimes AB^{+})^{+} \mathrm{Vec}\, E + (D \otimes I - C \otimes BA^{+})^{+} \mathrm{Vec}\, F. \tag{3.14}$$

Now, if we set $B_1 = D_1 = B_2 = A_2 = D_2$ in Corollary 8 we obtain the following corollary:

Corollary 11. *Let A, B and $C \in M_{m,n}$ be given full column-rank matrices such that*

$$R(C \otimes A) = R(C \otimes B), \quad R(C^* \otimes A^*) = R(C^* \otimes B^*).$$

Now, if E and $F \in M_m$ are constant matrices, and X, $Y \in M_n$ are unknown matrices to be solved then the vector least-squares solution of the CRSME

$$AXC^{\mathrm{T}} + BYC^{\mathrm{T}} = E, \quad BXC^{\mathrm{T}} + AYC^{\mathrm{T}} = F \tag{3.15}$$

can be given by

$$\mathrm{Vec}\, X = (C \otimes A - C \otimes BA^{+}B)^{+}\mathrm{Vec}\, E + (C \otimes B - C \otimes AB^{+}A)^{+}\mathrm{Vec}\, F, \tag{3.16}$$

$$\mathrm{Vec}\, Y = (C \otimes B - C \otimes AB^{+}A)^{+}\mathrm{Vec}\, E + (C \otimes A - C \otimes BA^{+}B)^{+}\mathrm{Vec}\, F. \tag{3.17}$$

As an example, we will discuss the efficient least-squares solution of the following coupled matrix equations:

$$AXC^{\mathrm{T}} + BYC^{\mathrm{T}} = E, \quad BXC^{\mathrm{T}} + AYC^{\mathrm{T}} = F, \tag{3.18}$$

where $A, B \in M_{m,n}, C \in M_{p,n}, E$ and $F \in M_{m,p}$ are given scalar matrices, and X and $Y \in M_n$ are unknown matrices to be solved. We also assume that $n < mp$, so that the coupled matrix equations (3.18) is over-determined, which suggests using a least-square approach. We consider the CLSP:

$$\min_{X,Y} \left\| \begin{bmatrix} E \\ F \end{bmatrix} - \begin{bmatrix} AXC^{\mathrm{T}} + BYC^{\mathrm{T}} \\ BXC^{\mathrm{T}} + AYC^{\mathrm{T}} \end{bmatrix} \right\|_2^2. \tag{3.19}$$

It is not difficult to transform (3.19) into the vector CLSP form:

$$\min_{X,Y} \left\| \begin{bmatrix} \mathrm{Vec}\, E \\ \mathrm{Vec}\, F \end{bmatrix} - \begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix} \begin{bmatrix} \mathrm{Vec}\, X \\ \mathrm{Vec}\, Y \end{bmatrix} \right\|_2^2 \tag{3.20}$$

which has the following solution:

$$\begin{bmatrix} \mathrm{Vec}\, X \\ \mathrm{Vec}\, Y \end{bmatrix} = \left\{ \begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix}^* \begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix} \right\}^{-1} \begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix}^* \begin{bmatrix} \mathrm{Vec}\, E \\ \mathrm{Vec}\, F \end{bmatrix}. \tag{3.21}$$

It is easily verified that the

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} I & -I \\ I & I \end{bmatrix}$$

is a unitary matrix and

$$\begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix} = U \begin{bmatrix} C \otimes (A+B) & 0 \\ 0 & C \otimes (A-B) \end{bmatrix} U^{\mathrm{T}}. \tag{3.22}$$

If $Q = A + B$, $T = A - B$, $H = C \otimes (A + B)$ and $W = C \otimes (A - B)$, then we have

$$\begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix}^* \begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix} = U \begin{bmatrix} (C \otimes Q)^*(C \otimes Q) & 0 \\ 0 & (C \otimes T)^*(C \otimes T) \end{bmatrix} U^{\mathrm{T}} \tag{3.23}$$

and

$$\begin{aligned} &\left(\begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix}^* \begin{bmatrix} C \otimes A & C \otimes B \\ C \otimes B & C \otimes A \end{bmatrix} \right)^{-1} \\ &\quad = U \begin{bmatrix} ((C \otimes Q)^*(C \otimes Q))^{-1} & 0 \\ 0 & ((C \otimes T)^*(C \otimes T))^{-1} \end{bmatrix} U^{\mathrm{T}} \\ &\quad = \frac{1}{2} \begin{bmatrix} I & -I \\ I & I \end{bmatrix} \begin{bmatrix} (H^*H)^{-1} & 0 \\ 0 & (W^*W)^{-1} \end{bmatrix} \begin{bmatrix} I & I \\ -I & I \end{bmatrix} \\ &\quad = \frac{1}{2} \begin{bmatrix} (H^*H)^{-1} + (W^*W)^{-1} & (H^*H)^{-1} - (W^*W)^{-1} \\ (H^*H)^{-1} - (W^*W)^{-1} & (H^*H)^{-1} + (W^*W)^{-1} \end{bmatrix}. \end{aligned} \tag{3.24}$$

Now, the least-squares solutions (3.21) can be written into the form

$$\begin{bmatrix}\mathrm{Vec}\,X\\ \mathrm{Vec}\,Y\end{bmatrix}=\frac{1}{2}\begin{bmatrix}(H^*H)^{-1}+(W^*W)^{-1} & (H^*H)^{-1}-(W^*W)^{-1}\\ (H^*H)^{-1}-(W^*W)^{-1} & (H^*H)^{-1}+(W^*W)^{-1}\end{bmatrix}\begin{bmatrix}(C\otimes A)^* & (C\otimes B)^*\\ (C\otimes B)^* & (C\otimes A)^*\end{bmatrix}\begin{bmatrix}\mathrm{Vec}\,E\\ \mathrm{Vec}\,F\end{bmatrix}. \tag{3.25}$$

This gives the following vector least-squares solutions:

$$\begin{aligned}\mathrm{Vec}\,X &= \tfrac{1}{2}(\{(H^*H)^{-1}+(W^*W)^{-1}\}(C\otimes A)^*+\{(H^*H)^{-1}-(W^*W)^{-1}\}(C\otimes B)^*)\mathrm{Vec}\,E\\ &= +\tfrac{1}{2}(\{(H^*H)^{-1}+(W^*W)^{-1}\}(C\otimes B)^*\\ &\quad +\{(H^*H)^{-1}-(W^*W)^{-1}\}(C\otimes A)^*)\mathrm{Vec}\,F, \end{aligned}\tag{3.26}$$

$$\begin{aligned}\mathrm{Vec}\,Y &= \tfrac{1}{2}(\{(H^*H)^{-1}-(W^*W)^{-1}\}(C\otimes A)^*+\{(H^*H)^{-1}+(W^*W)^{-1}\}(C\otimes B)^*)\mathrm{Vec}\,E,\\ &\quad +\tfrac{1}{2}(\{(H^*H)^{-1}-(W^*W)^{-1}\}(C\otimes B)^*\\ &\quad +\{(H^*H)^{-1}+(W^*W)^{-1}\}(C\otimes A)^*)\mathrm{Vec}\,F. \end{aligned}\tag{3.27}$$

In order to be able to use (3.21), (3.26) and (3.27) we must ascertain that the matrix

$$\begin{bmatrix}C\otimes A & C\otimes B\\ C\otimes B & C\otimes A\end{bmatrix}^*\begin{bmatrix}C\otimes A & C\otimes B\\ C\otimes B & C\otimes A\end{bmatrix} \tag{3.28}$$

is invertible if and only one

$$H^*H=(C\otimes(A+B))^*(C\otimes(A+B)),\quad W^*W=(C\otimes(A-B))^*(C\otimes(A-B)) \tag{3.29}$$

are invertible matrices.

In some problems (for example the multistatic antenna processing problem) the unknown matrices might be diagonal. As we observed earlier, when the unknown matrices X and $Y\in M_n$ are diagonal in CLSP in (3.19), the solution for Vec X and Vec Y are highly inefficient since most of the elements of X and Y vanish. Instead we can use the more compact vectorization identity (1.8) to rewrite the CLSP in (3.19) in the reduced-order vector form:

$$\min_{X,Y}\left\|\begin{bmatrix}\mathrm{Vec}\,E\\ \mathrm{Vec}\,F\end{bmatrix}-\begin{bmatrix}C\oslash A & C\oslash B\\ C\oslash B & C\oslash A\end{bmatrix}\begin{bmatrix}\mathrm{Vecd}\{X\}\\ \mathrm{Vecd}\{Y\}\end{bmatrix}\right\|_2^2. \tag{3.30}$$

Note that Vecd$\{X\}$ and Vecd$\{Y\}$ consists of only the nontrivial (i.e., diagonal) elements of matrices X and Y. The explicit efficient solution of (3.30) is can easily show that

$$\begin{aligned}\mathrm{Vecd}\{X\} &= \tfrac{1}{2}(\{(R^*R)^{-1}+(S^*S)^{-1}\}(C\oslash A)^*+\{(R^*R)^{-1}-(S^*S)^{-1}(C\oslash B)^*\})\mathrm{Vec}\,E\\ &\quad +\tfrac{1}{2}(\{(R^*R)^{-1}+(S^*S)^{-1}\}(C\oslash B)^*\\ &\quad +\{(R^*R)^{-1}-(S^*S)^{-1}\}(C\oslash A)^*)\mathrm{Vec}\,F, \end{aligned}\tag{3.31}$$

$$\begin{aligned}\mathrm{Vecd}\{Y\} &= \tfrac{1}{2}(\{(R^*R)^{-1}-(S^*S)^{-1}\}(C\oslash A)^*+\{(R^*R)^{-1}+(S^*S)^{-1}\}(C\oslash B)^*)\mathrm{Vec}\,E\\ &\quad +\tfrac{1}{2}(\{(R^*R)^{-1}-(S^*S)^{-1}(C\oslash B)^*\\ &\quad +\{(R^*R)^{-1}+(S^*S)^{-1}\}(C\oslash A)^*)\mathrm{Vec}\,F, \end{aligned}\tag{3.32}$$

where $R=C\oslash(A+B)$ and $S=C\oslash(A-B)$. In order to be able to use (3.31) and (3.32), we must ascertain also that

$$R^*R=(C\oslash(A+B))^*(C\oslash(A+B)),\quad S^*S=(C\oslash(A-B))^*(C\oslash(A-B)) \tag{3.33}$$

are invertible matrices.

It turns out that the expressions (3.31) and (3.32) can be also implemented involving Hadamard product by applying identities (1.5) and (1.9).

Corollary 12. *Let A, B, C, E and $F \in M_n$ be given constant matrices. Then the coupled matrix equations*:

$$AXC + BYC = E, \quad BXC + AYC = F \tag{3.34}$$

has a unique solution if and only if C , $Q = A + B$ and $T = A - B$ are invertible matrices, in this case, the unique solution is given by

$$X = \tfrac{1}{2}\{(A + B)^{-1}(E + F) + (A - B)^{-1}(E - F)\}C^{-1}, \tag{3.35}$$

$$Y = \tfrac{1}{2}\{(A + B)^{-1}(E + F) - (A - B)^{-1}(E - F)\}C^{-1}. \tag{3.36}$$

4. Weighted least-squares iterative solutions

In this section, we study the weighted least-squares iterative algorithms to solve the coupled Sylvester matrix equations

$$AX + YB = C, \quad DX + YE = F, \tag{4.1}$$

where $A, D \in M_m$ and $B, E \in M_n$ and $C , F \in M_{m,n}$ are given constant matrices, $X, Y \in M_{m,n}$ are the unknown matrices to be solved.

First, let us introduce a large family of iterative methods to solve the linear equation

$$Ax = b, \tag{4.2}$$

where $A = [a_{ij}] \in M_n$ is a given full-rank matrix with non-zero diagonal elements, $b \in \mathbb{C}^n$ is a constant vector, and $x \in \mathbb{C}^n$ is an unknown vector to be solved. Let $F \in M_n$ be a full-rank matrix to be determined and $\dot{\alpha} > 0$ be the step-size or convergence factor. Ding and Chen [10] presented a large family of iterative methods as follows:

$$x_k = x_{k-1} + \alpha F(b - Ax_{k-1}), \quad k = 1, 2, 3, \ldots \tag{4.3}$$

which includes the Jacobi and Gauss–Seidel iterations as special cases (where x_k is the iterative solution of x). For example, when $F = D^{-1}$and $\alpha = 1$, we get the Jacobi method; when $F = (L + D)^{-1}$ and $\alpha = 1$, we obtain the Gauss–Seidel method (where D and L are diagonal and strictly lower triangular parts of A). Unfortunately, the Jacobi and Gauss–Seidel iterations cannot guarantee that x_k converges to the exact solution $x = A^{-1}b$, and are not suitable for solving the non-square system: $Tx = g$ with $T \in M_{m,n}$. Ding and Chen [10] mentioned also that

(i) if we take $F = A^*$, then the gradient iterative algorithm,

$$x_k = x_{k-1} + \alpha A^*(b - Ax_{k-1}), \quad 0 < \alpha < \frac{2}{\lambda_{\max}[A^*A]} \text{ or } 0 < \alpha < \frac{2}{\|A\|_2^2} \tag{4.4}$$

yields $\lim_{k\to\infty} x_k = x$.

(ii) if we take $F = A^{-1}$, then the following iterative algorithm converges to x:

$$x_k = x_{k-1} + \alpha A^{-1}(b - Ax_{k-1}), \quad 0 < \alpha < 2. \tag{4.5}$$

This motivates us to study the so-called weighted least-squares iterative method in the following lemma. This lemma is straightforward and its proof is omitted.

Lemma 13. *If A be a non-square $m \times n$ full column-rank matrix with respect to positive definite matrices $M \in M_m$ and $N \in M_n$. Then we have $\lim_{k\to\infty} x_k = x$ in the following weighted least-squares iterative algorithm,*

$$x_k = x_{k-1} + \alpha(A^*MA)^{-1}A^*M(b - Ax_{k-1}), \quad 0 < \alpha < 2, \tag{4.6}$$

for $k = 1, 2, 3, \ldots$.

It is also easy to prove that the weighted iterative solution in (4.6) converges to the minimum N-norm M-least-squares solution $x=(A^*MA)^{-1}A^*Mb$ at a fast exponential rate, or it is linearly convergent. When $\alpha=1$, the iteration in (4.6) gives $x_1=(A^*MA)^{-1}A^*Mb$.

The weighted iterative algorithm in (4.6) is also suitable for solving non-square systems and is very useful for finding the weighted iterative solutions of coupled matrix equations to be studied later; the convergence factor α do not rely on the matrix A and is easy to choose, although the algorithm in (4.6) require computing weighted matrix inversion only at the first step.

In order to derive the weighted iterative solutions to (4.1), we need to introduce the intermediate b_1 and b_2 as follows:

$$b_1=\begin{bmatrix} C-YB \\ F-YE \end{bmatrix}, \quad b_2=[C-AX \;\; F-DX]. \tag{4.7}$$

Then from (4.1), we obtain two fictitious subsystems

$$S_1: GX=b_1, \quad S_2: YH=b_2, \tag{4.8}$$

where $G=[\begin{smallmatrix} A \\ D \end{smallmatrix}]\in M_{2m,m}$ and $H=[B \;\; E]\in M_{n,2n}$ are full column and full row-rank matrices, respectively.

Let X_k and Y_k be a weighted iterative solutions of X and Y, respectively. By using Lemma 13 and (1.12)–(1.13), it is easy to find the weighted iterative solutions to S_1 and S_2 with respect to positive definite matrices $M\in M_{2m}$ and $N\in M_{2n}$ as follows:

$$X_k=X_{k-1}+\alpha(G^*MG)^{-1}\begin{bmatrix} A \\ D \end{bmatrix}^* M\left\{b_1-\begin{bmatrix} A \\ D \end{bmatrix}X_{k-1}\right\}, \tag{4.9}$$

$$Y_k=Y_{k-1}+\alpha\{b_2-Y_{k-1}[B \;\; E]\}N^{-1}\;[B \;\; E]^*(HN^{-1}H^*)^{-1}. \tag{4.10}$$

Substituting (4.7) into (4.9) and (4.10) gives

$$X_k=X_{k-1}+\alpha(G^*MG)^{-1}\begin{bmatrix} A \\ D \end{bmatrix}^* M\begin{bmatrix} C-YB-AX_{k-1} \\ F-YE-DX_{k-1} \end{bmatrix}, \tag{4.11}$$

$$Y_k=Y_{k-1}+\alpha[C-AX-Y_{k-1}B \;\; F-DX-Y_{k-1}E]N^{-1}\;[B \;\; E]^*(HN^{-1}H^*)^{-1}. \tag{4.12}$$

We note that here, a difficulty arises in the expressions on the right-hand sides of (4.11) and (4.12) contain the unknown parameter matrices X and Y, respectively, so it is impossible to realize the algorithm in (4.11) and (4.12). The solution is based on the hierarchical identification principle; the unknown variables Y in (4.11) and X in (4.12) are replaced by their estimates $Y_{k-1}$and X_{k-1}. Thus, we obtain the weighted least-squares iterative solutions X_k and Y_k of the coupled Sylvester equations in (4.1) as

$$X_k=X_{k-1}+\alpha(G^*MG)^{-1}\begin{bmatrix} A \\ D \end{bmatrix}^* M\begin{bmatrix} C-AX_{k-1}-Y_{k-1}B \\ F-DX_{k-1}-Y_{k-1}E \end{bmatrix}, \tag{4.13}$$

$$Y_k=Y_{k-1}+\alpha[C-AX_{k-1}-Y_{k-1}B \;\; F-DX_{k-1}-Y_{k-1}E]N^{-1}\;[B \;\; E]^*(HN^{-1}H^*)^{-1}, \tag{4.14}$$

where

$$\alpha=\frac{1}{m+n} \quad \text{or} \quad \alpha=\frac{1}{\lambda_{\max}[G(G^*MG)^{-1}G^*M]+\lambda_{\max}[N^{-1}H^*(HN^{-1}H^*)^{-1}H^*]}. \tag{4.15}$$

The weighted least-squares iterative algorithm in (4.13)–(4.15) requires to compute the weighted matrix inversions $(G^*MG)^{-1}$ and $(HN^{-1}H^*)^{-1}$ only once at the first step. To initialize the algorithm, we take $X(0)=Y(0)=0$ or some small real matrix, e.g., $X(0)=Y(0)=10^{-6}\mathbf{1}_{m\times n}$ with $\mathbf{1}_{m\times n}$ being an $m\times n$ matrix whose elements are 1.

Theorem 14. *If the coupled Sylvester matrix equation determined by* (4.1) *has a unique solution X and Y, then the weighted iterative solution X_{n+1} and Y_{n+1} given by the algorithms in* (4.13)–(4.15) *converges to X and Y for any finite initial values $X(0)$ and $Y(0)$, that is*

$$\lim_{k\to\infty} X_k=X, \quad \lim_{k\to\infty} Y_k=Y. \tag{4.16}$$

Proof. Though the proof of Theorem 14 is quite similar to the proof of Theorem 1 in [10] for a special case when $M = I_{2m}$ and $N = I_{2n}$. We give proof of the general case for the sake of convenience. Define two error matrices

$$\widetilde{X}_k = X_k - X \quad \text{and} \quad \widetilde{Y}_k = Y_k - Y. \tag{4.17}$$

By using (4.1) and (4.13)–(4.14), it is easy to get

$$\widetilde{X}_k = \widetilde{X}_{k-1} + \alpha(G^*MG)^{-1}\begin{bmatrix} A \\ D \end{bmatrix}^* M \begin{bmatrix} -A\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}B \\ -D\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}E \end{bmatrix}, \tag{4.18}$$

$$\widetilde{Y}_k = \widetilde{Y}_{k-1} + \alpha[-A\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}B \quad -D\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}E]N^{-1}[B \quad E]^*(HN^{-1}H^*)^{-1}. \tag{4.19}$$

Taking the WMFN of (4.18) with respect to positive definite matrices $M \in M_{2m}$, $Z \in M_n$ and $R \in M_m$ using the following formula:

$$\begin{aligned}
&\|G(X + (G^*MG)^{-1}Y)\|^2_{M,Z} \\
&\quad = \|M^{1/2}G(X + (G^*MG)^{-1}Y)Z^{-1/2}\|^2_2 \\
&\quad = \operatorname{tr}[Z^{-1/2}(X^* + Y^*((G^*MG)^{-1})^*)G^*M^{1/2}M^{1/2}G(X + (G^*MG)^{-1}Y)Z^{-1/2}] \\
&\quad = \operatorname{tr}[Z^{-1/2}X^*G^*M^{1/2}M^{1/2}GXZ^{-1/2}] + \operatorname{tr}[Z^{-1/2}X^*G^*M^{1/2}M^{1/2}G(G^*MG)^{-1}YZ^{-1/2}] \\
&\qquad + \operatorname{tr}[Z^{-1/2}Y^*((G^*MG)^{-1})^*G^*M^{1/2}M^{1/2}GXZ^{-1/2}] \\
&\qquad + \operatorname{tr}[Z^{-1/2}Y^*((G^*MG)^{-1})^*G^*M^{1/2}M^{1/2}G(G^*MG)^{-1}YZ^{-1/2}] \\
&\quad = \operatorname{tr}[(M^{1/2}GXZ^{-1/2})^*(M^{1/2}GXZ^{-1/2})] + 2\operatorname{tr}[Z^{-1/2}X^*YZ^{-1/2}] \\
&\qquad + \operatorname{tr}[(M^{1/2}G(G^*MG)^{-1}YZ^{-1/2})^*M^{1/2}G(G^*MG)^{-1}YZ^{-1/2}] \\
&\quad = \|M^{1/2}GXZ^{-1/2}\|^2_2 + 2\operatorname{tr}[Z^{-1/2}X^*YZ^{-1/2}] + \|M^{1/2}G(G^*MG)^{-1}YZ^{-1/2}\|^2_2 \\
&\quad = \|GX\|^2_{M,Z} + 2\operatorname{tr}[Z^{-1/2}X^*YZ^{-1/2}] + \|G(G^*MG)^{-1}Y\|^2_{M,Z}
\end{aligned} \tag{4.20}$$

gives

$$\begin{aligned}
&\|G\widetilde{X}_k\|^2_{M,Z} \\
&\quad = \|G\widetilde{X}_{k-1}\|^2_{M,Z} + 2\alpha \operatorname{tr}\left\{Z^{-1/2}\widetilde{X}^*_{k-1}\begin{bmatrix} A \\ D \end{bmatrix}^* M \begin{bmatrix} -A\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}B \\ -D\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}E \end{bmatrix} Z^{-1/2}\right\} \\
&\qquad + \alpha^2 \left\| G(G^*MG)^{-1}\begin{bmatrix} A \\ D \end{bmatrix}^* M \begin{bmatrix} -A\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}B \\ -D\widetilde{X}_{k-1} - \widetilde{Y}_{k-1}E \end{bmatrix}\right\|^2_{M,Z} \\
&\quad \leqslant \|G\widetilde{X}_{k-1}\|^2_{M,Z} - 2\alpha \operatorname{tr}\{(A\widetilde{X}_{k-1})^*(A\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}B) + (D\widetilde{X}_{k-1})^{\mathrm{T}}(D\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}E)\} \\
&\qquad + \alpha^2 m\{\|A\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}B\|^2_{R,Z} + \|D\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}E\|^2_{R,Z}\}.
\end{aligned} \tag{4.21}$$

Similarly, taking the WMFN of (4.19) with respect to positive definite matrices $N \in M_{2n}$, $Z \in M_n$ and $R \in M_m$ we have

$$\begin{aligned}
&\|\widetilde{Y}_k H\|^2_{R,N} \\
&\quad \leqslant \|\widetilde{Y}_{k-1}H\|^2_{R,N} - 2\alpha \operatorname{tr}(\widetilde{Y}_{k-1}B)^*(A\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}B) + (\widetilde{Y}_{k-1}E)^*(D\widetilde{X}_1 + \widetilde{Y}_{k-1}E) \\
&\qquad + \alpha^2 n\{\|A\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}B\|^2_{R,Z} + \|D\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}E\|^2_{R,Z}\}.
\end{aligned} \tag{4.22}$$

Defining a non-negative definite function:

$$W_k = \|G\widetilde{X}_k\|^2_{M,Z} + \|\widetilde{Y}_k H\|^2_{R,N}, \tag{4.23}$$

and using (4.21) and (4.22), we have

$$\begin{aligned} W_k &\leqslant W_{k-1} - 2\alpha \left\{ \|A\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}B\|^2_{R,Z} + \|D\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}E\|^2_{R,Z} \right\} \\ &\quad + \alpha^2(m+n)\{\|A\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}B\|^2_{R,Z} + \|D\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}E\|^2_{R,Z}\} \\ &\leqslant W_{k-1} - \alpha\{2 - \alpha(m+n)\}\{\|A\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}B\|^2_{R,Z} + \|D\widetilde{X}_{k-1} + \widetilde{Y}_{k-1}E\|^2_{R,Z}\} \\ &\leqslant W_0 - \alpha\{2 - \alpha(m+n)\} \sum_{i=1}^{k-1} \{\|A\widetilde{X}_i + \widetilde{Y}_i B\|^2_{R,Z} + \|D\widetilde{X}_i + \widetilde{Y}_i E\|^2_{R,Z}\}. \end{aligned} \tag{4.24}$$

If the convergence factor is chosen to satisfy: $0 < \mu < 2/(m+n)$, then

$$\sum_{k=1}^{\infty} \{\|A\widetilde{X}_k + \widetilde{Y}_k B\|^2_{R,Z} + \|D\widetilde{X}_k + \widetilde{Y}_k E\|^2_{R,Z}\} < \infty. \tag{4.25}$$

It follows that as $k \to \infty$,

$$\|A\widetilde{X}_k + \widetilde{Y}_k B\|^2_{R,Z} + \|D\widetilde{X}_k + \widetilde{Y}_k E\|^2_{R,Z} = 0,$$

or

$$A\widetilde{X}_k + \widetilde{Y}_k B = 0, \quad D\widetilde{X}_k + \widetilde{Y}_k E = 0. \tag{4.26}$$

According to (1.7) we have

$$\begin{bmatrix} I_n \otimes A & B^{\mathrm{T}} \otimes I_m \\ I_n \otimes D & E^{\mathrm{T}} \otimes I_m \end{bmatrix} \begin{bmatrix} \operatorname{Vec} \widetilde{X}_k \\ \operatorname{Vec} \widetilde{Y}_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{4.27}$$

and the unique solution of (4.27) is $\widetilde{X}_k \to 0$ and $\widetilde{Y}_k \to 0$ as $k \to \infty$. This completes the proof of Theorem 14. □

The convergence factor in (4.15) may not the best and may be conservative. In fact, there exist a best α such that the fast convergence rate of X_k to X and Y_k to Y can be obtained as in the following numerical example which is given in [10].

Example 15. Suppose that the coupled matrix are

$$AX + YB = C, \quad DX + YE = F$$

with

$$A = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -0.2 \\ 0.2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} -2 & -0.5 \\ 0.5 & 2 \end{bmatrix},$$

$$E = \begin{bmatrix} -1 & -3 \\ 2 & -4 \end{bmatrix}, \quad C = \begin{bmatrix} 13.2 & 10.6 \\ 0.6 & 8.4 \end{bmatrix}, \quad F = \begin{bmatrix} -9.5 & -18 \\ 16 & 3.5 \end{bmatrix}.$$

Then the exact solutions of X and Y are

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}, \quad Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -2 & 3 \end{bmatrix}.$$

Taking $X(0) = Y(0) = 10^{-6}\mathbf{1}_{2\times 2}$, $M = N = I_4$, and applying the algorithms in (4.13) and (4.14) to compute X_k and Y_k. The iterative solutions X_k and Y_k is shown for $\alpha = 1/1.1$ as in Table 1 where

$$\delta = \sqrt{\frac{\|X_k - X\|^2_{R,Z} + \|Y_k - Y\|^2_{R,Z}}{\|X\|^2_{R,Z} + \|Y\|^2_{R,Z}}} \tag{4.28}$$

is the relative error.

Table 1

k	x_{11}	x_{12}	x_{21}	x_{22}	y_{11}	y_{12}	y_{21}	y_{22}	$\delta(\%)$
5	3.6143	2.9900	2.9409	3.6971	3.3228	0.3895	−2.9754	3.2708	22.3326
10	3.5861	3.0545	2.9027	3.8764	2.3446	0.7818	−2.111	3.09467	7.8486
15	3.8223	3.0602	2.9532	3.9752	2.2117	0.8313	−2.1088	3.0717	4.3430
20	3.8947	3.0514	2.9703	3.9963	2.1074	0.9035	−2.0499	3.0407	2.4141
25	3.9404	3.0339	2.9826	4.0011	2.0624	0.9399	−2.0272	3.0252	1.4291
30	3.9645	3.0217	2.9894	4.0017	2.0364	0.9638	−2.0153	3.0151	0.85256
35	3.9788	3.0134	2.9936	4.0013	2.0217	0.9780	−20089	3.0091	0.51332
40	3.9872	3.0082	2.9961	4.0009	2.0130	0.9867	−2.0053	3.0055	0.30979
45	3.9923	3.0050	2.9977	4.0005	2.0079	0.9919	−2.0032	3.0034	0.18728
50	3.9953	3.0030	2.9986	4.0003	2.0047	0.9951	−2.0019	3.0020	0.11329
55	3.9972	3.0018	2.9991	4.0002	2.0029	0.9970	−2.0012	3.0012	0.06855
60	3.9983	3.0011	2.9994	4.0001	2.0017	0.9982	−2.0007	3.0007	0.04149
Solution	4	3	3	4	2	1	−2	3	

From Table 1, it is clear that δ are becoming smaller and smaller and goes to zero as k increases. This indicates the algorithm is effective.

How to extend the use of weighted iterative method in (4.6) by using the connection between Vec(.) and Kronecker products to find the weighted least-squares iterative solution of more general coupled matrix equations of the following form requires further research:

$$\begin{aligned} &A_{11}X_1B_{11} + A_{12}X_2B_{12} + \cdots + A_{1p}X_pB_{1p} = C_1,\\ &A_{21}X_1B_{21} + A_{22}X_2B_{22} + \cdots + A_{2p}X_pB_{2p} = C_2,\\ &\vdots\\ &A_{p1}X_1B_{p1} + A_{p2}X_2B_{p2} + \cdots + A_{pp}X_pB_{pp} = C_p, \end{aligned} \tag{4.29}$$

where $A_{ij} \in M_m$, $B_{ij} \in M_n$ and $C_i \in M_{m,n}$ are given constant matrices, $X_i \in M_{m,n}$ are the unknown matrix functions to be solved. This work requires further research. Although the weighted iterative algorithms are presented for coupled Sylvester matrix equations; the idea adopted may extend to find the weighted least-squares iterative solutions of the general coupled matrix equations determined by (4.29).

5. Concluding remarks

In general, there are two classes of methods to solve matrix equations:

- *Direct methods*, such as *LU*-factorization or *QR*-factorization or Kronecker structures. These methods *Theoretically* lead to an exact solution of the problem in finitely many steps.
- *Iterative methods*, such as Jacobi iteration or Gauss–Seidel iteration or SOR. These methods provide an approximate solution to the problem.

In fact, the Kronecker products and vector operators affirming their capability of solving matrix and matrix equations fast (more fast when the unknown matrices are diagonal). The way exists which transform the coupled matrix into the following simple form: $Ax = b$, and we can solve this system fast if A is a Kronecker product.

In order to demonstrate that application of Kronecker products method is effective, suppose we have to solve, for example, the following matrix equation:

$$BXA^{\mathrm{T}} = C, \tag{5.1}$$

where A, B and $C \in M_n$ are given scalar matrices and $X \in M_n$ is unknown matrix to be solved. Then it is not hard by using (1.7) to establish the following equivalence:

$$(A \otimes B)\operatorname{Vec} X = \operatorname{Vec} C. \tag{5.2}$$

If we ignore the Kronecker products structure, then we need to solve the following both matrix equations:

- $BY = C$,
  where Y can be obtained in $\mathrm{O}(n^3)$ arithmetic operations (flops) by using LU-factorization of matrix B (forward substitution).
- $XA^{\mathrm{T}} = Y$,
  where X can be obtained also in $\mathrm{O}(n^3)$ operations (flops) by using LU-factorization of matrix A (back substitution).

Without exploiting the Kronecker products structure, an $n^2 \times n^2$ system defined in (5.1) would normally (by Gaussian elimination) require $\mathrm{O}(n^6)$ operations to solve. But when we use Kronecker product structure in (5.2), the calculations show that the $\mathrm{Vec}\, X$ can be obtained only in $\mathrm{O}(n^3)$ operations by using LU-factorization of matrices A and B. Thus, we can say that the system in (5.2) can be solved fast since the Kronecker structure avoids the formation of $n^2 \times n^2$ matrices, and only the smaller lower and upper triangular matrices L_A, L_B, U_A, U_B are needed.

For example, if we consider matrices A and B are of order 3×3 and a vector C is of order 9×1. To demonstrate the usefulness of applying Kronecker product and Vec(.), we can return to the system problem in (5.2). If $A \otimes B$ is non-singular and regarding with LU-factorizations of $A = L_A U_A$ and $B = L_B U_B$, then a solution of system exists and can be written as

$$(U_A \otimes U_B)\mathrm{Vec}\, X = z, \quad (L_A \otimes L_B)z = \mathrm{Vec}\, C.$$

First, the lower triangular system: $(L_A \otimes L_B)z = \mathrm{Vec}\, C$, can be solved by forward substitution as the following:

$$\left(\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}\right) \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_9 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_9 \end{bmatrix},$$

that is

$$L_A \otimes L_B = \begin{bmatrix} a_{11}b_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{11}b_{21} & a_{11}b_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{11}b_{31} & a_{11}b_{32} & a_{11}b_{33} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21}b_{11} & 0 & 0 & a_{22}b_{11} & 0 & 0 & 0 & 0 & 0 \\ a_{21}b_{21} & a_{21}b_{22} & 0 & a_{22}b_{21} & a_{22}b_{22} & 0 & 0 & 0 & 0 \\ a_{21}b_{31} & a_{21}b_{32} & a_{21}b_{33} & a_{22}b_{31} & a_{22}b_{32} & a_{22}b_{33} & 0 & 0 & 0 \\ a_{31}b_{11} & 0 & 0 & a_{32}b_{11} & 0 & 0 & a_{33}b_{11} & 0 & 0 \\ a_{31}b_{21} & a_{31}b_{22} & 0 & a_{32}b_{21} & a_{32}b_{22} & 0 & a_{33}b_{21} & a_{33}b_{22} & 0 \\ a_{31}b_{31} & a_{31}b_{32} & a_{31}b_{33} & a_{32}b_{31} & a_{32}b_{32} & a_{32}b_{33} & a_{33}b_{31} & a_{33}b_{32} & a_{33}b_{33} \end{bmatrix}$$

which can be solved in $\mathrm{O}(n^2)$ operations.

Thus the first three equations are given as:

- $a_{11}b_{11}z_1 = c_1 \;\Rightarrow z_1 = \dfrac{c_1}{a_{11}b_{11}}$.
- $a_{11}b_{21}z_1 + a_{11}b_{22}z_2 = c_2 \;\Rightarrow z_2 = \dfrac{b_{11}c_2 - b_{21}c_1}{a_{11}b_{11}b_{22}}$.
- $a_{11}b_{31}z_1 + a_{11}b_{32}z_2 + a_{11}b_{33}z_3 = c_3 \;\Rightarrow z_3 = \dfrac{b_{11}b_{22}c_3 - b_{11}b_{32}c_2 - b_{22}b_{31}c_1 + b_{32}b_{21}c_1}{a_{11}b_{11}b_{22}b_{33}}$.

Now the next three equations are:

- $\mathbf{a}_{21}\mathbf{b}_{11}\mathbf{z}_1 + a_{22}b_{11}z_4 = c_4$.
- $\mathbf{a}_{21}\mathbf{b}_{21}\mathbf{z}_1 + \mathbf{a}_{21}\mathbf{b}_{12}\mathbf{z}_2 + a_{22}b_{21}z_4 + a_{22}b_{22}z_5 = c_5$.
- $\mathbf{a}_{21}\mathbf{b}_{31}\mathbf{z}_1 + \mathbf{a}_{21}\mathbf{b}_{32}\mathbf{z}_2 + \mathbf{a}_{21}\mathbf{b}_{33}\mathbf{z}_3 + a_{22}b_{31}z_4 + a_{22}b_{32}z_5 + a_{22}b_{33}z_6 = c_6$.

The first boldface expression $\mathbf{a}_{21}\mathbf{b}_{11}\mathbf{z}_1$ can be computed as $a_{21}c_1/a_{11}$. The second boldface expression

$$\mathbf{a}_{21}\mathbf{b}_{21}\mathbf{z}_1 + \mathbf{a}_{21}\mathbf{b}_{12}\mathbf{z}_2$$

can also be computed as $a_{21}c_2/a_{11}$. While the third boldface expression

$$\mathbf{a}_{21}\mathbf{b}_{31}\mathbf{z}_1 + \mathbf{a}_{21}\mathbf{b}_{32}\mathbf{z}_2 + \mathbf{a}_{21}\mathbf{b}_{33}\mathbf{z}_3$$

can be also computed as $a_{21}c_3/a_{11}$.

Now, we use the previous expressions for obtaining z_1, z_2 and z_3 in the first set of equations to simplify the second set of three equations. The simplified second set of equations becomes

$$\begin{aligned}
&a_{22}b_{11}z_4 = c_4 - \frac{a_{21}c_1}{a_{11}},\\
&a_{22}b_{21}z_4 + a_{22}b_{22}z_5 = c_5 - \frac{a_{21}c_2}{a_{11}},\\
&a_{22}b_{31}z_4 + a_{22}b_{32}z_5 + a_{22}b_{33}z_6 = c_6 - \frac{a_{21}c_3}{a_{11}}.
\end{aligned}$$

Solving the second set of equations takes $O(n)$ operations and the forward solve step takes $O(n^2)$ operations (flops), so obtaining z_4, z_5 and z_6 takes $O(n^2)$ time. This simplification and using the work from the previous solution step continuous so that solving each of n-sets of n-equations takes $O(n^2)$ time, resulting in an overall solution time of $O(n^2)$. Exploiting the Kronecker structure reduce the usual, expected $O(n^4)$ time to solve $(L_A \otimes L_B)z = \operatorname{Vec} C$ to $O(n^2)$.

One final note regarding the exploitation of the Kronecker structure of the system remains. Suppose the matrices A and B are different sizes. Then, the time required to solve the system: $(A \otimes B)\operatorname{Vec} X = \operatorname{Vec} C$, is $O(mn^2)$, where m and n are the sizes of A and B, respectively. In our work, the modeler has some choice for the size of the A and B matrices. Thus, a wise choice would make n small, reducing the effect of the n^2 term in the $O(mn^2)$ computation time.

While if $A \in M_{m,n}$, $B \in M_{p,n}$, $C \in M_{m,p}$ are given matrices and $X \in M_n$ is a diagonal matrix, then also by using (1.8), it is not difficult to transform (5.1) into the following equivalence:

$$(A \oslash B)\operatorname{Vecd}(X) = \operatorname{Vec} C. \tag{5.3}$$

Subsequently, to construct a computationally efficient solution of the least-squares problem

$$\min_X \|(A \oslash B)\operatorname{Vecd}(X) - \operatorname{Vec} C\|_2^2 \tag{5.4}$$

requires $O(n^3) + O((m+p)n^2)$ (multiply and add) operations. Then the explicit solution of (5.4) follows as

$$\operatorname{Vecd}(X) = [(A \oslash B)^*(A \oslash B)]^{-1}(A \oslash B)^*\operatorname{Vec} C. \tag{5.5}$$

In contrast, the most efficient known alternative (i.e., (5.5) requires $O(n^3)+O((mp)n^2)$ operations, which is significantly higher when $n \ll \min(m, p)$).

Similarly, when X and $C \in M_n$ are diagonal matrix, we return to the system problem:

$$(A \circ B)\operatorname{Vecd}(X) = \operatorname{Vecd}(C). \tag{5.6}$$

If $A \circ B$ is non-singular and regarding with LU-factorizations of $A \circ B = L_{A\circ B}U_{A\circ B}$, then a solution of system exists and can be written as:

$$U_{A\circ B}\operatorname{Vecd}(X) = y, \quad L_{A\circ B}\, y = \operatorname{Vecd}(C).$$

First, the lower triangular system $L_{A\circ B}\ y = \operatorname{Vecd}(C)$ can be solved by forward substitution as the following:

$$\begin{bmatrix} a_{11}b_{11} & 0 & 0 \\ a_{21}b_{21} & a_{22}b_{22} & 0 \\ a_{31}b_{31} & a_{32}b_{32} & a_{33}b_{33} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{22} \\ c_{33} \end{bmatrix}$$

which can be solved in O(n) operations as follows:

- $a_{11}b_{11}y_1 = c_{11} \Rightarrow y_1 = \dfrac{c_{11}}{a_{11}b_{11}}$.
- $a_{21}b_{21}y_1 + a_{22}b_{22}y_2 = c_{22} \Rightarrow y_2 = \dfrac{a_{11}b_{11}c_{22} - a_{21}b_{21}c_{11}}{a_{11}b_{11}a_{22}b_{22}}$.
- $a_{31}b_{31}y_1 + a_{32}b_{32}y_2 + a_{33}b_{33}y_3 = c_{33} \Rightarrow$

$$y_3 = \frac{a_{11}b_{11}a_{22}b_{22}c_{33} - [a_{31}b_{31}a_{22}b_{22} - a_{32}b_{32}a_{21}b21]c_{11} - a_{32}b_{32}a_{11}b_{11}c_{22}}{a_{11}b_{11}a_{22}b_{22}a_{33}b_{33}}.$$

Thus we can say that the system in (5.6) can be solved faster since the calculations of Vecd(X) can be obtained only in O(n) operations by using LU-factorization of $A \circ B$ since then we only need lower and upper triangular matrices $L_{A\circ B}$ and $U_{A\circ B}$.

Acknowledgement

The authors gratefully acknowledge that this research partially supported by MOSTI under the IRPA Project Grant no: 09-02-04-0898-EA001.

References

[1] Z. Al Zhour, A. Kılıçman, New algebraic method for solving the axial N-index transportation problem based on the Kronecker product, Matematika 21 (2) (2005) 113–123.
[2] Z. Al Zhour, A. Kılıçman, Extensions and generalization inequalities involving the Khatri–Rao product of several positive matrices, J. Inequalities Appl. 80878 (2006) 1–21.
[3] Z. Al Zhour, A. Kılıçman, Matrix equalities and inequalities involving Khatri–Rao and Tracy–Singh sums, J. Inequalities Pure Appl. Math. 7 (1)(Article 34) (2006), 1–17.
[4] A. Ben-Israel, T.N.E. Greville, Generalized Inverses: Theory and Applications, Wiley, New York, 1974.
[5] J.W. Brewer, Kronecker products and matrix calculus in system theory, IEEE Trans. Circuits Systems CAS-25 (9) (1978) 772–781.
[6] S.L. Campbell, C.D. Meyer, Generalized Inverses of Linear Transformations, Dover Publ. Inc, New York, 1999.
[7] R. Carlson, An inverse problem for the matrix Schrödinger equation, J. Math. Anal. Appl. 267 (2002) 564–575.
[8] R.E. Cline, T.N.E. Greville, A Drazin inverse for rectangular matrices, Linear Algebra Appl. 29 (1980) 53–62.
[9] J.B. Cruz, C.I. Chen, Series Nash solution of two-person non-zero sum linear differential games, J. Optimal Theory 7 (4) (1971) 240–257.
[10] F. Ding, T. Chen, Iterative least-squares solutions of coupled Sylvester matrix equations, Systems Control Lett. 54 (2005) 95–107.
[11] C.T. Fulton, L. Wu, Parallel algorithms for large least squares problems involving Kronecker products, Nonlinear Anal. Theory Methods Appl. 30 (8) (1997) 5033–5040.
[12] A. Graham, Kronecker Products and Matrix Calculus with Applications, first ed., Ellis Horwood Ltd., UK, 1981.
[13] M. Gulliksson, X.-O. Jin, Y. Wei, Perturbation bounds for constrained and weighted least squares problems, Linear Algebra Appl. 349 (2002) 221–232.
[14] A. Kılıçman, Z. Al Zhour, The general common exact solutions of coupled linear matrix and matrix differential equations, J. Anal. Comput. 1 (1) (2005) 15–29.
[15] A. Kılıçman, Z. Al Zhour, Iterative solutions of coupled matrix convolution equations, Soochow J. Math., accepted for publication.
[16] A. Kılıçman, Z. Al Zhour, Connection between the Kronecker and Hadamard convolution products of matrices and some applications, Internat. J. Pure Appl. Math., accepted for publication.
[17] S. Liu, Several inequalities involving Khatri–Rao products of positive semi-definite matrices, Linear Algebra Appl. 354 (2002) 175–186.
[18] H. Lev-Ari, Efficient solution of linear matrix equations with applications to multistatic antenna array processing, Comm. Inform. Systems 5 (1) (2005) 123–130.
[19] J.R. Magnus, H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrices, Revised ed., Wiley, New York, 1999.
[20] S.K. Mitra, A pair of simultaneous linear matrix equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$ and a matrix programming problem, Linear Algebra Appl. 131 (1990) 107–123.
[21] G. Mouroutsos, P.D. Sparis, Taylor series approach to system identification, analysis and optimal control, J. Franklin Institute Pergamon. 319 (3) (1985) 359–371.
[22] A. Navarra, P.L. Odell, D.M. Young, A representation of the general common solution to the matrix equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$ with applications, Comput. Math. Appl. 41 (2001) 929–935.
[23] Y.-X. Peng, X.-Y. Hu, L. Zhang, An iterative method for symmetric solutions and optimal approximation solution of the system of matrix equations $A_1XB_1 = C_1$, $A_2XB_2 = C_2$, Appl. Math. Comput., in press. Corrected proof available online 9 August 2006.
[24] V. Rakočević, Y. Wei, The representation theory for the Drazin inverse and its applications II, J. Austral. Math. Soc. 70 (2001) 189–197.

[25] C.R. Rao, S.K. Mitra, Generalized Inverses of Matrices and its Applications, Wiley, New York, 1971.
[26] C.R. Rao, M.B. Rao, Matrix Algebra and its Applications to Statistics and Econometrics, World Scientific Publishing, Singapore, 1998.
[27] Y. Tian, The Moore–Penrose inverses of $m \times n$ block matrices and their applications, Linear Algebra Appl. 283 (1998) 35–60.
[28] J.W. van der Woude, On the existence of a common solution X to the matrix $A_i X B_j = C_{ij}$, $(i, j) \in \Gamma$, Linear Algebra Appl. 375 (2003) 135–145.
[29] C.V. Van Loan, Generalizing the singular value decomposition, SIAM J. Numer. Anal. 13 (1976) 76–83.
[30] G. Wang, Weighted Moore–Penrose, Drazin and group inverses of the Kronecker product $A \otimes B$ and some applications, Linear Algebra Appl. 250 (1997) 39–50.
[31] D. Wang, Some topics on weighted Moore–Penrose inverse, weighted least squares and weighted regularized Tikhonov problems, Appl. Math. Comput. 157 (2004) 243–267.
[32] Y. Wei, Perturbation bound of singular linear system, Appl. Math. Comput. 105 (1999) 211–220.
[33] Y. Wei, Recurrent neural networks for computing weighted Moore–Penrose inverse, Appl. Math. Comput. 16 (2000) 279–287.
[34] Y. Wei, A characterization for the W-weighted Drazin inverse solution, Appl. Math. Comput. 125 (2002) 303–310.
[35] Y. Wei, Integral representation of the W-weighted Drazin inverse, Appl. Math. Comput. 144 (2003) 3–10.
[36] Y. Wei, C.-W. Woo, T. Lei, A note on the perturbation of the W-weighted Drazin inverse, Appl. Math. Comput. 149 (2004) 423–430.
[37] Y. Wei, H. Wu, The representation and approximation for the Drazin inverse, J. Comput. Appl. Math. 126 (2000) 417–432.
[38] F. Zhang, Matrix Theory: Basic Results and Techniques, Springer, New York, 1999.
[39] F. Zhang, Optimal control for mechanical vibration systems based on second order matrix equation, Mecha. System and Signal Process 16 (2002) 61–67.