Journal of Mathematical Modeling
Vol. 2, No. 2, 2015, pp. 170-186

JMM

Global least squares solution of matrix equation $\sum_{j=1}^{s} A_j X_j B_j = E$

Saeed Karimi*

Department of Mathematics, Persian Gulf University, Bushehr, Iran
Emails: karimi@pgu.ac.ir, karimijafarbigloo@gmail.com

Abstract. In this paper, an iterative method is proposed for solving matrix equation $\sum_{j=1}^{s} A_j X_j B_j = E$. This method is based on the global least squares (GL-LSQR) method for solving the linear system of equations with the multiple right hand sides. For applying the GL-LSQR algorithm to solve the above matrix equation, a new linear operator, its adjoint and a new inner product are defined. It is proved that the new iterative method obtains the least norm solution of the mentioned matrix equation within finite iteration steps in the exact arithmetic, when the above matrix equation is consistent. Moreover, the optimal approximate solution $(X_1^*, X_2^*, \ldots, X_s^*)$ to a given multiple matrices $(\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_s)$ can be derived by finding the least norm solution of a new matrix equation. Finally, some numerical experiments are given to illustrate the efficiency of the new method.

Keywords: Matrix equation, GL-LSQR algorithm, iterative method, linear operator, matrix nearness problem.
AMS Subject Classification: 15A06, 65F08, 65F10.

1 Introduction

Matrix equations appear frequently in many areas of applied mathematics and play vital roles in a many applications such as control theory, model reduction and image processing; see [1, 3, 4, 5, 6] and their references. Hence,

*Corresponding author.
Received: 14 September 2014 / Revised: 5 October 2014 / Accepted: 19 October 2014.

many researches are performed on the various type of matrix equations; for example see [1, 4, 7, 8, 11, 12, 14, 17, 18]. In [13, 15] a simple type of (4) has been studied. The methods used in [15] and [13] include singular value decomposition (SVD) and an iterative method, respectively. Throughout this paper, $\mathbb{C}^{n\times m}$ represents the set of all $n \times m$ complex matrices. Also the notations $tr(A)$ and A^H use to denote the trace and the transpose conjugate of the matrix A, respectively. For two matrices X and Y in $\mathbb{C}^{n\times m}$, $\langle X, Y\rangle_F$ denotes the Frobenius inner product, e.g., $\langle X, Y\rangle_F = tr(Y^H X)$, and associated Frobenius norm is denoted by $||.||_F$.

The following inner product is defined on $\Re = \mathbb{C}^{p_1\times q_1} \times \mathbb{C}^{p_2\times q_2} \times \ldots \times \mathbb{C}^{p_s\times q_s}$.

Definition 1. Let $\mathcal{X} = (X_1, X_2, \ldots, X_s)$ and $\mathcal{Y} = (Y_1, Y_2, \ldots, Y_s)$ be in $\Re$. Then

$$\langle\langle \mathcal{X}, \mathcal{Y}\rangle\rangle = tr(\sum_{j=1}^{s} Y_j^H X_j),$$

and its associated norm is denoted by $\sharp\mathcal{X}\sharp = \sqrt{\langle\langle \mathcal{X}, \mathcal{X}\rangle\rangle}$.

It is easy to show that $\sharp\mathcal{X}\sharp = \sqrt{\sum_{j=1}^{s} ||X_j||_F^2}$.
The symbol $*$ stands for the following product:

$$\mathbb{V} * y := \sum_{j=1}^{s} y_j \mathcal{V}_j, \tag{1}$$

where $\mathbb{V} = [\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_s]$, $\mathcal{V}_j \in \Re$ for $1 \le j \le s$, and $y \in \mathbb{C}^s$. By the same way, it is defined

$$\mathbb{V} * T := [\mathbb{V} * T(:, 1), \mathbb{V} * T(:, 2), \ldots, \mathbb{V} * T(:, s)], \tag{2}$$

where T is a $s \times s$ matrix and $T(:, j)$ denotes the jth column of T. It is easy to show that the following relations are satisfied

$$\mathbb{V} * (y + z) = \mathbb{V} * (y) + \mathbb{V} * (z), \qquad (\mathbb{V} * T) * y = \mathbb{V} * (Ty), \tag{3}$$

where y and z are in $\mathbb{C}^s$.

In this paper, the solution of the following matrix equation is considered

$$\sum_{j=1}^{s} A_j X_j B_j = E, \tag{4}$$

where $A_j \in \mathbb{C}^{n\times p_j}, B_j \in \mathbb{C}^{q_j\times m},\ j = 1, 2, \ldots, s$ and $E \in \mathbb{C}^{n\times m}$. Also, the solution of the following matrix nearness problem is considered

$$\min_{\mathcal{X}\in\mathcal{S}_{\mathcal{X}}} \sharp\mathcal{X} - \bar{\mathcal{X}}\sharp^2, \tag{5}$$

where $\bar{\mathcal{X}}$ is a given multiple matrices of $\Re$ and $\mathcal{S}_{\mathcal{X}}$ is the solution set of the linear equation (4). Matrix nearness problem (5) occurs frequently in experimental design. For more details on the matrix nearness problem, one can refer to References [2, 10, 9].

Recently, Toutounian et al. [16] proposed the GL-LSQR algorithm for obtaining the approximate solution of matrix equation $AX = B$, where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{R}^{n\times s}$. The GL-LSQR method generates two set matrices $V_1, V_2, \ldots, V_k$ and $U_1, U_2, \ldots, U_k$ which form an F-orthonormal basis for block Krylov subspaces $\mathbb{K}_k(A^TA; V_1)$ and $\mathbb{K}_k(AA^T; U_1)$, respectively.

In this paper, a new iterative method is presented for solving the matrix equation (4). This method is based on the GL-LSQR algorithm. It is proved that the new method obtains the least norm solution with respect to the Frobenius norm of (4). In addition, using the new iterative method, the solution of the minimization problem (5) can be obtained by first finding the least norm solution of a new matrix equation. Finally the numerical results of this paper are compared to those of [13].

The rest of this paper is organized as follows. In the next section, a short review of the GL-LSQR method is presented. Section 3 is devoted to the new iterative method. In Section 4, numerical examples are given to illustrate the efficiency of the new method. Finally, some concluding remarks are given in Section 5.

2 The GL-LSQR method

In this section, some fundamental properties of the GL-LSQR method [16] is reviewed for solving matrix equation $AX = B$. The GL-LSQR method uses a process, namely Global-Bidiag process, to reduce the coefficient matrix A to a global lower bidiagonal form. The Global-Bidiag process can be described as follows.

Global-Bidiag (starting matrix B; reducing to global lower bidiagonal form):

$$\begin{gathered}\beta_1 U_1 = B, \qquad \alpha_1 V_1 = A^H U_1,\\ \left.\begin{aligned}\beta_{i+1}U_{i+1} &= AV_i - \alpha_i U_i\\ \alpha_{i+1}V_{i+1} &= A^H U_{i+1} - \beta_{i+1}V_i,\end{aligned}\right\}, \quad i = 1, 2, \ldots,\end{gathered} \tag{6}$$

where $V_i \in \mathbb{C}^{m\times q}, U_i \in \mathbb{C}^{n\times q}, i = 1, 2, \ldots$. The scalars $\alpha_i \geq 0$ and $\beta_i \geq 0$ are chosen so that $\|U_i\|_F = \|V_i\|_F = 1$.

With the definitions

$$\begin{array}{l} \mathcal{U}_k \equiv [U_1, U_2, \ldots, U_k], \\ \mathcal{V}_k \equiv [V_1, V_2, \ldots, V_k], \end{array} \qquad T_k \equiv \begin{bmatrix} \alpha_1 & & & & \\ \beta_2 & \alpha_2 & & & \\ & \ddots & \ddots & & \\ & & \beta_k & \alpha_k & \\ & & & \beta_{k+1} & \end{bmatrix},$$

the recurrence relations (6) may be rewritten as:

$$\mathcal{U}_{k+1} * (\beta_1 e_1) = B, \tag{7}$$

$$A\mathcal{V}_k = \mathcal{U}_{k+1} * T_k, \tag{8}$$

$$A^T\mathcal{U}_{k+1} = \mathcal{V}_k * T_k^T + \alpha_{k+1} V_{k+1} * e_{k+1}^T. \tag{9}$$

Proposition 1. *([16]) Suppose that k steps of the Global-Bidiag process have been taken, then the $n \times s$ block vectors $V_1, V_2, \ldots, V_k$ and $U_1, U_2, \ldots, U_k, U_{k+1}$ are F-orthonormal basis of the Krylov subspaces $\mathbb{K}_k(A^HA, V_1)$ and $\mathbb{K}_{k+1}(AA^H, U_1)$, respectively.*

Proposition 2. *([16]) The Global-Bidiag process will be stopped at step m if and only if $m = \min\{\mu, \lambda\}$, where μ and λ are the grades of V_1 and U_1 with respect to A^HA and AA^H, respectively.*

Proposition 3. *([16]) Let $\mathcal{U}_k = [U_1, U_2, \ldots, U_k]$, where the $n \times s$ matrices U_i, $i = 1, \ldots, k$, are generated by the Global-Bidiag process. Then*

$$||\mathcal{U}_k * \eta||_F = ||\eta||_2,$$

where $\eta \in \mathbb{R}^k$ and $||.||_2$ is the ℓ^2-norm.

By using the Global-Bidiag process, the GL-LSQR algorithm constructs an approximate solution of the form $X_k = \mathcal{V}_k * y_k$, where $y_k \in \mathbb{R}^k$, which solves the least squares problem

$$\min_X ||B - AX||_F.$$

The main steps of the GL-LSQR algorithm can be summarized as follows.

Algorithm 1. *Gl-LSQR algorithm*

1. Set $X_0 = 0$
2. $\beta_1 = ||B||_F$, $U_1 = B/\beta_1$, $\alpha_1 = ||A^HU_1||_F$, $V_1 = A^HU_1/\alpha_1$.

3. Set $W_1 = V_1$, $\bar{\phi}_1 = \beta_1$, $\bar{\rho}_1 = \alpha_1$
4. For $i = 1, 2, \ldots$ until convergence, Do:
5. $\quad \bar{W}_i = AV_i - \alpha_i U_i$
6. $\quad \beta_{i+1} = ||\bar{W}_i||_F$
7. $\quad U_{i+1} = \bar{W}_i / \beta_{i+1}$
8. $\quad \bar{S}_i = A^H U_{i+1} - \beta_{i+1} V_i$
9. $\quad \alpha_{i+1} = ||\bar{S}_i||_F$
10. $\quad V_{i+1} = \bar{S}_i / \alpha_{i+1}$
11. $\quad \rho_i = (\bar{\rho}_i^2 + \beta_{i+1}^2)^{1/2}$
12. $\quad c_i = \bar{\rho}_i / \rho$
13. $\quad s_i = \beta_{i+1} / \rho_i$
14. $\quad \theta_{i+1} = s_i \alpha_{i+1}$
15. $\quad \bar{\rho}_{i+1} = c_i \alpha_{i+1}$
16. $\quad \phi_i = a_i \bar{\phi}_i$
17. $\quad \bar{\phi}_{i+1} = c_i \bar{\phi}_i$
18. $\quad \phi_i = c_i \bar{\phi}_i$
19. $\quad \bar{\phi}_{i+1} = -s_i \phi_i$
20. $\quad X_i = X_{i-1} + (\phi_i / \rho_i) W_i$
21. $\quad W_{i+1} = V_{i+1} - (\theta_{i+1} / \rho_i) W_i$
22. $\quad$ If $|\bar{\phi}_{i+1}|$ is small enough then stop
23. EndDo.

More details about the GL-LSQR algorithm can be found [16].

3 The new iterative method

Consider the linear operator

$$\begin{aligned} \mathcal{L} &: \Re \to \mathbb{C}^{n \times m}, \\ \mathcal{L}(\mathcal{X}) &= \sum_{j=1}^{s} A_j X_j B_j, \end{aligned} \tag{10}$$

where $A_j \in \mathbb{C}^{n \times p_j}$, $B_j \in \mathbb{C}^{q_j \times m}$, $j = 1, 2, \ldots, s$ and $\mathcal{X} = (X_1, \ldots, X_s)$. Therefore, the matrix equation (4) can be written as

$$\mathcal{L}(\mathcal{X}) = E. \tag{11}$$

Definition 2. Let $\mathcal{L}$ be the linear operator given in equation (11). Then linear operator

$$\mathcal{L}^* : \mathbb{C}^{n\times m} \to \Re,$$

that satisfies

$$\langle \mathcal{L}(\mathcal{X}), Y\rangle_F = \langle\langle \mathcal{X}, \mathcal{L}^*(Y)\rangle\rangle,$$

for all $\mathcal{X} \in \Re,\ Y \in \mathbb{R}^{n\times m}$, is called the adjoint of $\mathcal{L}$.

It is easy to prove the following remark.

Remark 1. Let $\mathcal{L}$ be linear operator (11). Then

$$\mathcal{L}^*(Y) = (A_1^H Y B_1^H, A_2^H Y B_2^H, \ldots, A_s^H Y B_s^H).$$

Similar to the block Krylov subspace $\mathbb{K}_k(A; R)$, the following block Krylov subspace is defined for $\mathcal{L}$.

Definition 3. Let $\mathcal{L}$ be linear operator (11). Then

$$\mathbb{K}_k(\mathcal{L}; R) = \text{span}\{R, \mathcal{L}(R), \mathcal{L}^2(R), \ldots, \mathcal{L}^{k-1}(R)\},$$

where $\mathcal{L}^i = \overbrace{\mathcal{L} \circ \mathcal{L} \circ \cdots \circ \mathcal{L}}^{i \text{ times}}$ and $\circ$ is the combination of two operators.

Definition 4. Let $\mathcal{L}$ be linear operator (11) and $\mathbb{V}_k = [\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k]$, where $\mathcal{V}_j \in \Re,\ j = 1, 2, \ldots, k$. Then

$$\mathcal{L}(\mathbb{V}_k) = [\mathcal{L}(\mathcal{V}_1), \mathcal{L}(\mathcal{V}_2), \ldots, \mathcal{L}(\mathcal{V}_k)] \in \mathbb{R}^{n\times km}.$$

To approximate the solution of the operator equation (11), a new algorithm is presented, will be referred by $\mathcal{L}$-GLS. Same as the GL-LSQR algorithm, the $\mathcal{L}$-GLS algorithm uses a global bidiagonal process, will be referred to $\mathcal{L}$-Bidiag. The $\mathcal{L}$-Bidiag process reduces the linear operator $\mathcal{L}$ to the lower bidiagonal matrix form. The $\mathcal{L}$-Bidiag process can be described as follows.

$\mathcal{L}$-Bidiag (starting matrix E; reduction to lower bidiagonal matrix form):

$$\beta_1 U_1 = E,\ \alpha_1 \mathcal{V}_1 = \mathcal{L}^*(U_1)$$

$$\left.\begin{array}{l} \beta_{i+1} U_{i+1} = \mathcal{L}(\mathcal{V}_i) - \alpha_i U_i \\ \alpha_{i+1} \mathcal{V}_{i+1} = \mathcal{L}^*(U_{i+1}) - \beta_{i+1} \mathcal{V}_i \end{array}\right\}, \quad i = 1, 2, ..., \qquad (12)$$

where $U_i \in \mathbb{C}^{n\times m}$ and $\mathcal{V}_i \in \Re$. The scalars $\alpha_i \geq 0$ and $\beta_i \geq 0$ are chosen so that $||U_i||_F = 1$ and $\sharp \mathcal{V}_i \sharp = 1$.

With the definitions

$$\mathcal{U}_k \equiv [U_1, U_2, ..., U_k], \quad \mathbb{V}_k \equiv [\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k], \qquad T_k \equiv \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \beta_k & \alpha_k \\ & & & \beta_{k+1} \end{bmatrix},$$

and using the notation $*$ and Definition 4, the recurrence relations (6) can be rewritten as:

$$\mathcal{U}_{k+1} * (\beta_1 e_1) = E, \tag{13}$$

$$\mathcal{L}(\mathbb{V}_k) = \mathcal{U}_{k+1} * T_k. \tag{14}$$

Proposition 4. *Suppose that k steps of the $\mathcal{L}$-Bidiag process have been taken, then $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k$ and $U_1, \ldots, U_{k+1}$ are the orthonormal basis of the Krylov subspaces $\mathbb{K}_k(\mathcal{L}^*\mathcal{L}; \mathcal{V}_1)$ and $\mathbb{K}_{k+1}(\mathcal{L}\mathcal{L}^*; U_1)$, respectively.*

Proof. The proof is similar to Proposition 1 and is omitted. □

Proposition 5. *Let $\mathbb{V}_k = [\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k]$, where $\mathcal{V}_i$, $i = 1, \ldots, k$, are generated by the $\mathcal{L}$-Bidiag process. Then*

$$\sharp \mathbb{V}_k * \eta \sharp = ||\eta||_2,$$

where $\eta \in \mathbb{R}^k$.

Proof. We have

$$\sharp \mathbb{V}_k * \eta \sharp^2 = \langle\langle \sum_{i=1}^{s} \eta_i \mathcal{V}_i, \sum_{i=1}^{s} \eta_i \mathcal{V}_i \rangle\rangle,$$

since $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k$ are orthonormal with respect to norm $\sharp.\sharp$, it results that

$$\sharp \mathbb{V}_k * \eta \sharp^2 = \sum_{i=1}^{s} \eta_i^2 \langle\langle \mathcal{V}_i, \mathcal{V}_i \rangle\rangle = \sum_{i=1}^{s} \eta_i^2,$$

and this proves the proposition. □

The quantities generated from the linear operator $\mathcal{L}$ and E by the $\mathcal{L}$-Bidiag process will now be used to solve the least squares problem,

$$\min_{\mathcal{X} \in \Re} ||E - \mathcal{L}(\mathcal{X})||_F.$$

Let the quantities

$$\mathcal{X}_k = \mathbb{V}_k * y_k, \tag{15}$$

$$R_k = E - \mathcal{L}(\mathcal{X}_k), \tag{16}$$

be defined, where $y_k \in \mathbb{R}^k$. According to linearity of the operator $\mathcal{L}$, it is easy to show that

$$\mathcal{L}(\mathcal{X}_k) = \mathcal{L}(\mathbb{V}_k) * y_k.$$

Also, it readily follows from (15), (16) and properties of product $*$ that the equation

$$R_k = E - \mathcal{L}(\mathbb{V}_k) * y_k = \mathcal{U}_{k+1} * (\beta_1 e_1) - (\mathcal{U}_{k+1} * T_k) * y_k = \mathcal{U}_{k+1} * (\beta_1 e_1 - T_k y_k),$$

holds to working accuracy.

To minimize the kth residual $||R_k||_F$, since $\mathcal{U}_{k+1}$ is F-orthonormal and by using the Proposition 5, we choose y_k so that

$$||R_k||_F = ||\beta_1 e_1 - T_k y_k||_2, \tag{17}$$

is minimum. This minimization problem is carried out by applying the QR decomposition [16], where a unitary matrix Q_k is determined so that

$$Q_k \begin{bmatrix} T_k & \beta_1 e_1 \end{bmatrix} = \begin{bmatrix} \mathcal{R}_k & f_k \\ 0 & \bar{\phi}_{k+1} \end{bmatrix} = \begin{bmatrix} \rho_1 & \theta_1 & & & & \phi_1 \\ & \rho_2 & \theta_3 & & & \phi_2 \\ & & \ddots & \ddots & & \vdots \\ & & & \rho_{k-1} & \theta_k & \phi_{k-1} \\ & & & & \rho_k & \phi_k \\ & & & & 0 & \bar{\phi}_{k+1} \end{bmatrix},$$

where ρ_l, θ_l and ϕ_l are scalars. The above QR factorization is determined by constructing the kth plane rotation $Q_{k,k+1}$ to operate on rows k and $k+1$ of the transformed $\begin{bmatrix} T_k & \beta_1 e_1 \end{bmatrix}$ to annihilate β_{k+1}. This gives the following simple recurrence relation:

$$\begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix} \begin{bmatrix} \bar{\rho}_k & 0 & \bar{\phi}_k \\ \beta_{k+1} & \alpha_{k+1} & 0 \end{bmatrix} = \begin{bmatrix} \rho_k & \theta_{k+1} & \phi_k \\ 0 & \bar{\rho}_{k+1} & \bar{\phi}_{k+1} \end{bmatrix},$$

where $\bar{\rho}_1 \equiv \alpha_1$, $\bar{\phi}_1 \equiv \beta_1$ and the scalars c_k and s_k are the nontrivial elements of $Q_{k,k+1}$. The quantity $\bar{\rho}_k$ and $\bar{\phi}_k$ are intermediate scalars that are subsequently replaced by ρ_k and ϕ_k.

By setting

$$y_k = \mathcal{R}_k^{-1} * f_k,$$

the approximate solution is given by

$$\mathcal{X}_k = \mathbb{V}_k * (\mathcal{R}_k^{-1} * y_k), \tag{18}$$

$$= (\mathbb{V}_k * \mathcal{R}_k^{-1}) * f_k. \tag{19}$$

Letting

$$\mathbb{P}_k \equiv \mathbb{V}_k * \mathcal{R}_k^{-1} \equiv [\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_k],$$

then

$$\mathcal{X}_k = \mathbb{P}_k * f_k.$$

The last column of $\mathbb{P}_k$, $\mathcal{P}_k$, can be computed from the previous $\mathcal{P}_{k-1}$ and $\mathcal{V}_k$, by the simple update

$$\mathcal{P}_k = (\mathcal{V}_k - \theta_k \mathcal{P}_{k-1})/\rho_k, \tag{20}$$

also note that,

$$f_k = \begin{bmatrix} f_{k-1} \\ \phi_k \end{bmatrix},$$

in which

$$\phi_k = c_k \bar{\phi}_k.$$

Thus, $\mathcal{X}_k$ can be updated at each step, via the relation

$$\mathcal{X}_k = \mathcal{X}_{k-1} + \phi_k \mathcal{P}_k.$$

The residual norm $||R_k||_F$ is computed directly from the quantity $\bar{\phi}_{k+1}$ as

$$||R_k||_F = |\bar{\phi}_{k+1}|.$$

Some of the work in (20) can be eliminated by using matrices $\mathcal{W}_k = \rho_k \mathcal{P}_k$ in place of $\mathcal{P}_k$. The main steps of the $\mathcal{L}$-GLS algorithm can be summarized as follows.

Algorithm 2. *$\mathcal{L}$-GLS algorithm*

1. Set $\mathcal{X}_0 = 0 \in \Re$
2. $\beta_1 = ||E||_F$, $U_1 = E/\beta_1$, $\alpha_1 = \sharp \mathcal{L}^*(U_1) \sharp$, $\mathcal{V}_1 = \mathcal{L}^*(U_1)/\alpha_1$.
3. Set $\mathcal{W}_1 = \mathcal{V}_1$, $\bar{\phi}_1 = \beta_1$, $\bar{\rho}_1 = \alpha_1$
4. For $i = 1, 2, \ldots,$ until convergence, Do:
5. $\quad \hat{W}_i = \mathcal{L}(\mathcal{V}_i) - \alpha_i U_i$
6. $\quad \beta_{i+1} = ||\hat{W}_i||_F$
7. $\quad U_{i+1} = \hat{W}_i/\beta_{i+1}$
8. $\quad \hat{S}_i = \mathcal{L}^*(U_{i+1}) - \beta_{i+1}\mathcal{V}_i$
9. $\quad \alpha_{i+1} = \sharp \hat{S}_i \sharp$
10. $\quad \mathcal{V}_{i+1} = \hat{S}_i/\alpha_{i+1}$
11. $\quad \rho_i = (\bar{\rho}_i^2 + \beta_{i+1}^2)^{1/2}$
12. $\quad c_i = \bar{\rho}_i/\rho$

13. $s_i = \beta_{i+1}/\rho_i$
14. $\theta_{i+1} = s_i\alpha_{i+1}$
15. $\bar{\rho}_{i+1} = c_i\alpha_{i+1}$
16. $\phi_i = a_i\bar{\phi}_i$
17. $\bar{\phi}_{i+1} = c_i\bar{\phi}_i$
18. $\phi_i = c_i\bar{\phi}_i$
19. $\bar{\phi}_{i+1} = -s_i\phi_i$
20. $\mathcal{X}_i = \mathcal{X}_{i-1} + (\phi_i/\rho_i)\mathcal{W}_i$
21. $\mathcal{W}_{i+1} = \mathcal{V}_{i+1} - (\theta_{i+1}/\rho_i)\mathcal{W}_i$
22. If $|\bar{\phi}_{i+1}|$ is small enough then stop
23. EndDo.

In the following, it is proved that $\mathcal{L}$-GLS gives the minimum norm solution of (4). That is, it solves the optimization problem $\min \sharp\mathcal{X}\sharp$ such that satisfies (4). Let $N(\mathcal{L})$ and $R(\mathcal{L})$ denote the null space and range of an operator $\mathcal{L}$, respectively.

Lemma 1. *Let $\mathcal{L}$ be the linear operator (11) and $\tilde{\mathcal{X}} = (\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_s) \in \Re$ satisfies $\mathcal{L}^*(\mathcal{L}(\tilde{\mathcal{X}})) = 0$, then $\tilde{\mathcal{X}} \in N(\mathcal{L})$.*

Proof. According to the definition of $\mathcal{L}^*$ and since $\mathcal{L}^*(\mathcal{L}(\tilde{\mathcal{X}})) = 0$, we have

$$\langle\langle\tilde{\mathcal{X}}, \mathcal{L}^*(\mathcal{L}(\tilde{\mathcal{X}}))\rangle\rangle = 0 \Rightarrow \langle\mathcal{L}(\tilde{\mathcal{X}}), \mathcal{L}(\tilde{\mathcal{X}})\rangle_F = 0.$$

So $\mathcal{L}(\tilde{\mathcal{X}}) = 0$ and this completes the proof. □

Let $\mathcal{L}$ be the linear operator (10) and $\mathcal{L}$-GLS applies to solve (11), then the following theorem is presented.

Theorem 3. *$\mathcal{L}$-GLS returns the minimum norm solution.*

Proof. The final $\mathcal{L}$-GLS solution satisfies $\mathcal{L}^*(\mathcal{L}(\mathcal{X}_k^{GLS})) = \mathcal{L}^*(E)$, and any other solution $\hat{\mathcal{X}}$ satisfies $\mathcal{L}^*(\mathcal{L}(\hat{\mathcal{X}})) = \mathcal{L}^*(E)$. With $\tilde{\mathcal{X}} = \hat{\mathcal{X}} - \mathcal{X}_k^{GLS}$, the difference between the two normal equations gives $\mathcal{L}^*(\mathcal{L}(\tilde{\mathcal{X}})) = 0$, so that $\mathcal{L}(\tilde{\mathcal{X}}) = 0$ by Lemma 1. From $\alpha_1\mathcal{V}_1 = \mathcal{L}^*(U_1)$ and $\alpha_{k+1}\mathcal{V}_{k+1} = \mathcal{L}^*(U_{k+1}) - \beta_{k+1}\mathcal{V}_k$, we see that $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k \in R(\mathcal{L}^*)$. From $\mathcal{L}(\tilde{\mathcal{X}}) = 0$, it follows that $\tilde{\mathcal{X}}$ is orthogonal to $\mathbb{V}_k$, that is, $\tilde{\mathcal{X}} \perp \mathbb{V}_k$. Therefore,

$$\begin{aligned}
\sharp\hat{\mathcal{X}}\sharp^2 - \sharp\mathcal{X}_k^{GLS}\sharp^2 &= \sharp\mathcal{X}_k^{GLS} + \tilde{\mathcal{X}}\sharp^2 - \sharp\mathcal{X}_k^{GLS}\sharp^2, \\
&= \langle\langle\tilde{\mathcal{X}}, \tilde{\mathcal{X}}\rangle\rangle + 2\langle\langle\mathcal{X}_k^{GLS}, \tilde{\mathcal{X}}\rangle\rangle, \\
&= \sharp\tilde{\mathcal{X}}\sharp^2 + 2tr(\tilde{\mathcal{X}}^H\mathbb{V}_k y_k^{GLS}) \\
&= \sharp\tilde{\mathcal{X}}\sharp^2 \geq 0,
\end{aligned}$$

and this completes the proof. □

To solve the matrix nearness problem (5), let the linear operator equation (11) be consistent, then

$$\mathcal{L}(\mathcal{X}) = E \;\Leftrightarrow\; \mathcal{L}(\mathcal{X} - \bar{\mathcal{X}}) = E - \mathcal{L}(\bar{\mathcal{X}}), \quad \forall \mathcal{X} \in \mathcal{S}_{\mathcal{X}}, \bar{\mathcal{X}} \in \Re.$$

Let $\tilde{\mathcal{X}} = \mathcal{X} - \bar{\mathcal{X}}$ and $\tilde{E} = E - \mathcal{L}(\bar{\mathcal{X}})$, then the matrix nearness problem (5) is equivalent to first finding the least norm solution(with respect to norm $\sharp.\sharp$) of the linear operator equation $\mathcal{L}(\tilde{\mathcal{X}}) = \tilde{E}$. Using $\mathcal{L}$-GLS algorithm, the unique least norm solution $\tilde{\mathcal{X}}^*$ of the linear operator equation $\mathcal{L}(\tilde{\mathcal{X}}) = \tilde{E}$ can be obtained, then the unique solution $\hat{\mathcal{X}}$ of the matrix nearness problem (5) can be computed as $\hat{\mathcal{X}} = \tilde{\mathcal{X}}^* + \bar{\mathcal{X}}$.

4 Numerical experiments

In this paper, some numerical examples are given to illustrate the feasibility and effectiveness of the new method. The numerical results of the new method are compared to the one that's in [13]. In all examples, the initial guess was taken the zero element of $\Re$ and stopping criterion is $\|R_i\|_F < 10^{-10}$, where R_i is the ith residual at the ith iteration. All the numerical experiments were computed in double precision with some MATLAB codes.

Example 1. ([13]) Consider the linear equation

$$AXB + CYD = E,$$

where the matrices A, B, C, D and E are as follows:

$$A = \begin{bmatrix} 1 & 3 & 1 & 3 & 1 \\ 3 & -7 & 3 & -7 & 3 \\ 3 & -2 & 3 & -2 & 3 \\ 11 & 6 & 11 & 6 & 11 \\ -5 & 5 & -5 & 5 & -5 \\ 9 & 4 & 9 & 4 & 9 \end{bmatrix}, \; B = \begin{bmatrix} -1 & 4 & -1 & 4 & -1 \\ 5 & -1 & 5 & -1 & 5 \\ -1 & -2 & -1 & -2 & -1 \\ 3 & 9 & 3 & 9 & 3 \\ 7 & -8 & 7 & -8 & 7 \end{bmatrix},$$

$$C = \begin{bmatrix} 3 & -4 & 3 & -4 & 1 & 6 \\ -1 & 3 & -1 & 3 & -3 & -1 \\ 3 & -5 & 3 & -5 & 2 & 5 \\ 3 & -4 & 3 & -4 & 1 & 6 \\ -1 & 3 & -1 & 3 & -3 & -1 \\ 3 & -5 & 3 & -5 & 2 & 5 \end{bmatrix}, \; D = \begin{bmatrix} -5 & 4 & -1 & -5 & 4 \\ -2 & 3 & 5 & -2 & 3 \\ 3 & 5 & -1 & 3 & 5 \\ 2 & -6 & 3 & 2 & -6 \\ 1 & 11 & 7 & 1 & 11 \\ 4 & -1 & 4 & -5 & 4 \end{bmatrix},$$

$$E = \begin{bmatrix} -79 & 613 & -172 & 126 & 424 \\ 245 & 479 & 317 & 612 & 128 \\ 124 & 975 & -7 & 519 & 584 \\ 14 & 2099 & -79 & 1612 & 517 \\ -247 & -969 & -175 & -836 & -364 \\ 58 & 1791 & -73 & 1335 & 518 \end{bmatrix}.$$

By applying the $\mathcal{L}$-GLS algorithm to this example, after 34 iterations while in [13] after 40 iterations, the unique least norm solution is obtained as follows:

$$X_{34} = \begin{bmatrix} 1.2075 & 0.7524 & -0.9367 & 3.8822 & -1.3053 \\ -0.1886 & -0.9652 & 0.4140 & -1.5433 & -0.6884 \\ 1.2075 & 0.7524 & -0.9367 & 3.8822 & -1.3053 \\ -0.1886 & -0.9652 & 0.4140 & -1.5433 & -0.6884 \\ 1.2075 & 0.7524 & -0.9367 & 3.8822 & -1.3053 \end{bmatrix},$$

$$Y_{34} = \begin{bmatrix} 0.1461 & -0.6742 & 1.5150 & -1.3108 & 0.8278 & -0.2846 \\ 0.2668 & 1.4287 & -2.1160 & 1.5454 & -0.3976 & -0.4103 \\ 0.1461 & -0.6742 & 1.5150 & -1.3108 & 0.8278 & -0.2846 \\ 0.2668 & 1.4287 & -2.1160 & 1.5454 & -0.3976 & -0.4103 \\ 1.2104 & 1.0492 & -2.5987 & 0.8949 & -1.7203 & 1.0718 \\ 1.8359 & 0.3841 & 0.8009 & -2.0708 & 1.5019 & -1.1077 \end{bmatrix},$$

with a corresponding residual norm

$$||R_{34}||_F = ||E - AX_{34}B - CY_{34}D||_F = 1.1079e - 011.$$

Let

$$\bar{X} = \begin{bmatrix} -6 & 2 & -3 & 2.5 & -6 \\ 2 & -1 & 3 & -5.5 & 2 \\ -3 & 3 & -3 & 3 & -2 \\ 2.5 & -5 & 3 & 2 & 2 \\ -6 & 2 & -2 & 2 & -1 \end{bmatrix},$$

$$\bar{Y} = \begin{bmatrix} -1 & 1.5 & 1 & 0 & 0.5 & 3 \\ 1.5 & 3 & -1 & 2 & 0 & 0 \\ 1 & -1 & 2 & -2 & 0.5 & -1 \\ 0 & 2 & -2 & 2 & 1.5 & -1.5 \\ 0.5 & 0 & 0.5 & 1.5 & -3 & 1.5 \\ 3 & 0 & -1 & -1.5 & 1.5 & -1 \end{bmatrix},$$

then using $\mathcal{L}$-GLS algorithm and iterating 33 steps, while [13] iterating 39 steps, the unique least norm solution on the new equation is obtained as follows:

$$\tilde{X}^*_{33} = \begin{bmatrix} 0.5177 & 0.1722 & -0.3541 & 1.4982 & -0.7179 \\ 0.4025 & -0.0617 & -0.2136 & 0.9487 & -0.7641 \\ 0.5177 & 0.1722 & -0.3541 & 1.4982 & -0.7179 \\ 0.4025 & -0.0617 & -0.2136 & 0.9487 & -0.7641 \\ 0.5177 & 0.1722 & -0.3541 & 1.4982 & -0.7179 \end{bmatrix},$$

$$\tilde{Y}^*_{33} = \begin{bmatrix} -0.2792 & -0.1855 & 0.5667 & -0.1688 & 0.4475 & -0.4077 \\ -0.2792 & -0.7427 & 0.0062 & 0.0340 & -1.1861 & 0.7051 \\ -0.2792 & -0.1855 & 0.5667 & -0.1688 & 0.4475 & -0.4077 \\ -0.2792 & -0.7427 & 0.0062 & 0.0340 & -1.1861 & 0.7051 \\ 1.3686 & 1.8617 & -1.6553 & 0.3741 & 0.8100 & -0.1359 \\ -0.6697 & -0.7386 & 1.1736 & -0.3693 & 0.3534 & -0.5462 \end{bmatrix},$$

with a corresponding residual norm

$$||\tilde{R}^*_{33}||_F = ||\tilde{E} - A\tilde{X}^*_{33}B - C\tilde{Y}^*_{33}D||_F = 9.4622e - 011,$$

where $\tilde{E} = E - A\tilde{X}B - C\tilde{Y}D$. Hence, the optimal solution (5) is

$$\hat{X}^*_{33} = \tilde{X}^*_{33} + \bar{X}$$
$$= \begin{bmatrix} -5.4823 & 2.1722 & -3.3541 & 3.9982 & -6.7179 \\ 2.4025 & -1.0617 & 2.7864 & -4.5513 & 1.2359 \\ -2.4823 & 3.1722 & -3.3541 & 4.4982 & -2.7179 \\ 2.9025 & -5.0617 & 2.7864 & 2.9487 & 1.2359 \\ -5.4823 & 2.1722 & -2.3541 & 3.4982 & -1.7179 \end{bmatrix},$$
$$\hat{Y}^*_{33} = \tilde{Y}^*_{33} + \bar{Y}$$
$$= \begin{bmatrix} -1.2792 & 1.3145 & 1.5667 & -0.1688 & 0.9475 & 2.5923 \\ 1.2208 & 2.2573 & -0.9938 & 2.0340 & -1.1861 & 0.7051 \\ 0.7208 & -1.1855 & 2.5667 & -2.1688 & 0.9475 & -1.4077 \\ -0.2792 & 1.2573 & -1.9938 & 2.0340 & 0.3139 & -0.7949 \\ 1.8686 & 1.8617 & -1.1553 & 1.8741 & -2.1900 & 1.3641 \\ 2.3303 & -0.7386 & 0.1736 & -1.8693 & 1.8534 & -1.5462 \end{bmatrix},$$

and the minimum is

$$\min_{\mathcal{S}_{X,Y}} (||X - \bar{X}||_F^2 + ||Y - \bar{Y}||_F^2) \approx ||\bar{X}^*_{33} - \bar{X}||_F^2 + ||\bar{Y}^*_{33} - \bar{Y}||_F^2 = 31.4902.$$

Example 2. Let A, B, C and D be, respectively, a 100×100 complex symmetric Toeplitz matrix with the first row $(1, \frac{1}{2}, \ldots, \frac{1}{100}) + i(1, 1, \ldots, 1)$, a 100×100 complex Hankel matrix with the entry in position (k, l), $\frac{-1}{k+l-1} - i$ $(k, l = 1, 2, \ldots, 100)$, a 100×100 complex tridiagonal matrix with diagonal elements equal to $2 + 2i$ and off-diagonal elements to $(-1, -\frac{1}{2}, \ldots, -\frac{1}{99}) + i(1, 1, \ldots, 1)$ and $(1, \frac{1}{2}, \ldots, \frac{1}{99}) + i(1, 1, \ldots, 1)$, and a 100×100 upper triangular part matrix of the matrix A, where $i = \sqrt{-1}$. Let $E = A\check{X}B + C\check{Y}D$, where $\check{X}$ and $\check{Y}$ be 100×100 matrices with all elements 1. Then linear matrix equation $AXB + CYD = E$ is consistent and the $\mathcal{L}$-GLS algorithm and Algorithm 2.1 [13] can be applied to obtain its least norm solution. Figure 1 shows the convergence history of this example. As is shown from Figure 1, the $\mathcal{L}$-GLS algorithm monotonically converges faster than Algorithm 2.1. This fact causes a decreasing on run time when we use $\mathcal{L}$-GLS algorithm (17.98 in seconds) substitute of Algorithm 2.1 [13] (23.18 in seconds).

Example 3. Consider the following matrix linear equation

$$AXB + CYD + EZF = G,$$

where

$$A = \begin{bmatrix} \text{hilb}(4) & \text{zeros}(4,3) \\ \text{eye}(4) & \text{ones}(4,3) \end{bmatrix}, \quad B = \begin{bmatrix} \text{ones}(5,5) & \text{zeros}(5,4) \\ \text{zeros}(4,5) & \text{pasc}(4) \end{bmatrix},$$
$$C = \begin{bmatrix} \text{magic}(4) \\ \text{ones}(4,4) \end{bmatrix},$$

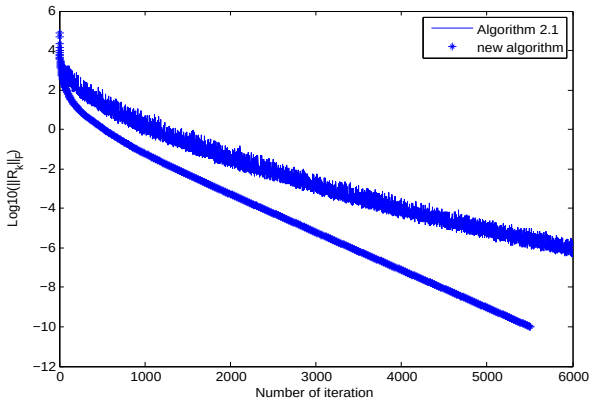

Figure 1: Convergence history of the Frobenius norm of the residual matrix for Example 2.

$$D = \begin{bmatrix} \text{hank}(1:4) & \text{zeros}(4,5) \\ \text{zeros}(5,4) & \text{zeros}(5,5) \end{bmatrix}, \quad E = \begin{bmatrix} \text{ones}(5,2) & \text{zeros}(5,3) \\ \text{zeros}(3,2) & \text{eye}(3) \end{bmatrix},$$

$$F = \begin{bmatrix} \text{toep}(1:5) & \text{ones}(5,4) \end{bmatrix},$$

where hilb(n), pasc(n) and magic(n) denote Hilbert matrix, Pascal matrix and Magic matrix of order n, respectively; and toep$(1:n)$ and hank$(1:n)$ denote Toeplitz matrix and Hankel matrix of order n, respectively, with their first rows being $(1, 2, \ldots, n)$, also ones(m, n) and zeros(m, n) denote $m \times n$ matrices with all components are 1 and 0, respectively. Let $G = A\hat{X}B + C\hat{Y}D + E\hat{Z}F$, where $\hat{X}, \hat{Y}$ and $\hat{Z}$ are ones$(7, 9)$, ones$(4, 9)$ and ones$(5, 5)$, respectively.

The $\mathcal{L}$-GLS algorithm was applied to this example and its convergence history has been shown in Figure 2.

Example 4. Consider the Example 2 in [13]. We applied $\mathcal{L}$-GLS algorithm and Algorithm 2.1 [13] for solving this example and the numerical results are given in the Table 1. In this table, notations "It" and "RN" denote the iteration number and residual norm. As is seen from this table, the iteration number and CPU time (in seconds) of two algorithms are close together but $\mathcal{L}$-GLS algorithm shows the relative superiority.

Table 1: Iteration number and CPU time of the $\mathcal{L}$-GLS algorithm and Algorithm 2.1 for Example 4.

	It	CPU-time	RN
$\mathcal{L}$-GLS	101	0.12	8.10e-11
Algorithm 2.1	107	0.17	8.56e-11

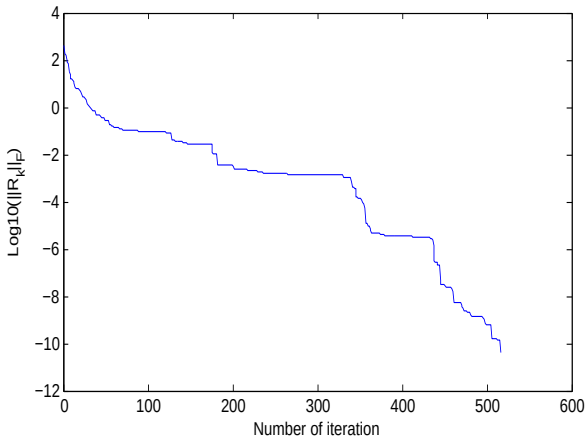

Figure 2: Convergence history of the Frobenius norm of the residual matrix for Example 3.

5 Conclusion

In this paper, an iterative method was proposed for obtaining the approximate solution of matrix equation (4). It was proved that the new iterative method obtains the least-norm solution of this equation. As an application, the new method was applied for solving the matrix nearness problem (5). Numerical results showed that the new method is more efficient than the one that's in [13].

Acknowledgements

The author would like to thank to the editor in chief and anonymous referee for their comments.

References

[1] Z.Z. Bai, *On Hermitian and skew-Hermitian splitting iteration methods for continuous Sylvester equations*, J. Comput. Math. **29** (2011) 185–198.

[2] M. Baruch, *Optimization procedure to correct stiffness and flexibility matrices using vibration tests*, Amer. Inst. Aeronaut. Astronaut. J. **16** (1978) 1208–1210.

[3] A. Björck, *Numerical methods for least squares problems*, SIAM, Phildelphia, PA, 1996.

[4] A. Bouhamidi, K. Jbilou, L. Reichel and H. Sadok, *An exterapolated TSVD method for linear discrete ill-posed problems with Kronecker structure*, Linear Algebra Appl. **434** (2011) 1677–1688.

[5] D. Calvetti and L. Reichel, *Application of ADI iterative methods to the restoration of noisy images*, SIAM J. Matrix Anal. Appl. **17** (1996) 165–186.

[6] B. Datta, *Numerical Methods for Linear Control Systems*, Elsevier Academic Press, 2004.

[7] Y.B. Deng, Z.Z. Bai and Y.H. Gao, *Iterative orthogonal direction methods for Hermitian minimum norm solution of two consistent matrix equations*, Numer. Linear Algebra Appl. **13** (2006) 801–823.

[8] A.El. Guennouni, K. Jbilou and J. Riquet, *Block Krylov subspace methods for solving large Sylvester equation*, Numer. Algor. **29** (2002) 75–96.

[9] N.J. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra Appl. **103** (1988) 103–118.

[10] Z. Jiang and Q. Lu, *Optimal application of a matrix under spectral restriction*, Math. Numer. Sin. **1** (1988) 47–52.

[11] M. Khorsand Zak and F. Toutounian, *Nested splitting conjugate gradient method for matrix equation $AXB = C$ and preconditioning*, Comput. Math. Appl. **66** (2013) 269–278.

[12] A.P. Liao, Z.Z. Bai and Y. Lei, *Best approximate solution of matrix equation AXB+CYD=E*, SIAM J. Matrix Anal. Appl. **27** (2005) 675–688.

[13] Z. Peng and Y. Peng, *An efficient iterative method for solving the matrix equation AXB+CYD=E*, Numer. Linear Algebra Appl. **13** (2006) 473–485.

[14] D.K. Salkuyeh, *CG-type algorithms to solve symmetric matrix equations*, Appl. Math. Comput. **172** (2006) 985–999.

[15] S.Y. Shim and Y. Chen, *Least squares solution of matrix equation* $AXB^* + CYD^* = E^*$, SIAM J. Matrix Anal. Appl. **24** (2003) 802–808.

[16] F. Toutounian and S. Karimi, *Global Least squares method (GL-LSQR) for solving general linear systems with several right-hand sides*, Appl. Math. Comput. **178** (2006) 452–460.

[17] M. Wang and Y. Feng, *An iterative algorithm for solving a class of matrix equations*, J. Control Theory Appl. **7** (2009) 68–72.

[18] L. Xie, J. Ding and F. Ding, *Gradient based iterative solutions for general linear matrix equations*, Comput. Math. Appl. **58** (2009) 1441–1448.