

Nested splitting conjugate gradient method for matrix equation $AXB = C$ and preconditioning



Mohammad Khorsand Zak^{a,*}, Faezeh Toutounian^{a,b}

^a Department of Applied Mathematics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

^b The Center of Excellence on Modelling and Control Systems, Ferdowsi University of Mashhad, Mashhad, Iran

ARTICLE INFO

Article history:

Received 29 December 2012

Received in revised form 30 April 2013

Accepted 8 May 2013

Keywords:

Conjugate gradient

NSCG method

Matrix equations

Preconditioning

Nested iterations

ABSTRACT

In this paper, we present a nested splitting conjugate gradient (NSCG) iteration method for solving a class of matrix equations with nonsymmetric coefficient matrices. This method is actually inner/outer iterations, which employs a CG-like method as inner iteration to approximate each outer iterate, while each outer iteration is induced by a convergent and symmetric positive definite splitting of the coefficient matrices. Convergence conditions of this method are studied in depth and numerical experiments show the efficiency of this method. Moreover, we show that the use of the quasi-Hermitian splitting as a preconditioner can induce an accurate, robust and effective preconditioned Krylov subspace method.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Matrix equations appear frequently in many areas of applied mathematics and play vital roles in a number of applications such as control theory, model reduction and image processing; see [1–4] and their references. Hence, many researches are performed on the various type of matrix equations; for example see [5,2,6–12].

In this paper, we consider the matrix equation

$$AXB = C, \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$, C and $X \in \mathbb{R}^{n \times m}$. The linear matrix equation (1) can be written as the following $nm \times nm$ linear system:

$$(B^T \otimes A) \text{vec}(X) = \text{vec}(C), \quad (2)$$

where $\text{vec}(X)$ is the vector of \mathbb{R}^{nm} obtained by stacking the columns of the $n \times m$ matrix X and \otimes denotes the Kronecker product: $(A \otimes B) = [a_{ij}B]$. Of course, this is a numerically poor way to determine the solution X of the matrix equation (1), as the linear system of equations (2) is costly to solve and can be ill-conditioned.

In this paper, we present an iterative method for solving the matrix equation (1) by using the symmetric and skew-symmetric splitting of the matrices A and B in a matrix variant of the nested splitting conjugate gradient (NSCG) method. A class of nested splitting conjugate gradient methods was first proposed in [13] for solving large sparse systems of linear equations. Via the NSCG method, the solution of the system of linear equations (2) is first approximated successively by a sequence of fixed-point matrix equations, which is induced by the splitting $\mathcal{A} = \mathcal{H} - \mathcal{S}$, where $\mathcal{A} = B^T \otimes A$, $\mathcal{H} = \frac{1}{2}(\mathcal{A}^T + \mathcal{A})$ and $\mathcal{S} = \frac{1}{2}(\mathcal{A}^T - \mathcal{A})$. Then, at each step of iteration a matrix equation with coefficient matrix $\mathcal{H} = H_B \otimes H_A - S_A \otimes S_B$ (where

* Corresponding author. Tel.: +98 9125119690.

E-mail addresses: mo.khorsand@stu-mail.um.ac.ir, khorsandzak@yahoo.com (M. Khorsand Zak), toutouni@math.um.ac.ir (F. Toutounian).

H_A, S_A and H_B, S_B are the symmetric and skew-symmetric parts of A and B , respectively) is solved iteratively by a CG-like method.

In the remainder of this paper, we use $\lambda(M), \|M\|_2, \|M\|_F$ and I_n to denote the eigenvalue, the spectral norm, the Frobenius norm of a matrix $M \in \mathbb{R}^{n \times n}$, and the identity matrix with dimension n , respectively. Note that $\|\cdot\|_2$ is also used to represent the 2-norm of a vector. For a nonsingular matrix \mathcal{B} , we denote by $\kappa(\mathcal{B}) = \|\mathcal{B}\|_2 \|\mathcal{B}^{-1}\|_2$ its Euclidean condition number, and for a symmetric positive definite matrix \mathcal{B} , we define the $\|\cdot\|_{\mathcal{B}}$ norm of a vector $x \in \mathbb{R}^n$ as $\|x\|_{\mathcal{B}} = \sqrt{x^T \mathcal{B} x}$. Then the induced $\|\cdot\|_{\mathcal{B}}$ norm of a matrix $H \in \mathbb{R}^{n \times n}$ is defined as $\|H\|_{\mathcal{B}} = \|\mathcal{B}^{\frac{1}{2}} H \mathcal{B}^{-\frac{1}{2}}\|_2$. In addition, it holds that $\|Hx\|_{\mathcal{B}} \leq \|H\|_{\mathcal{B}} \|x\|_{\mathcal{B}}, \|H\|_{\mathcal{B}} \leq \sqrt{\kappa(\mathcal{B})} \|H\|_2$ and $\|I\|_{\mathcal{B}} = 1$, where I is the identity matrix. Furthermore, we have the following relationships between the Frobenius norm of a matrix R and the 2-norm of a vector r :

$$r = \text{vec}(R) \Rightarrow \|r\|_2 = \|R\|_F.$$

The organization of this paper is as follows. Section 2 contains a brief description of the NSCG method. Section 3 presents our own contribution, i.e., the NSCG method for matrix equation $AXB = C$ and its convergence properties. Section 4 is devoted to numerical experiments. Finally, we present our conclusions in Section 5.

2. A brief review of the NSCG method

Axelsson et al. [13] proposed an efficient nested iterative method for solving the system of linear equations

$$Ax = b, \tag{3}$$

where $A \in \mathbb{R}^{n \times n}$ is a large sparse matrix. This method, which is called the nested splitting conjugate gradient (NSCG) method, is established as follows:

Let $A = B - C$ be a symmetric positive definite splitting of the coefficient matrix A , where B is a symmetric positive definite matrix [13]. Assume that this splitting satisfies the condition $\rho(B^{-1}C) < 1$. Then the system of linear equations (3) is equivalent to the fixed-point equation

$$Bx = Cx + b.$$

Given an initial guess $x^{(0)} \in \mathbb{R}^n$, suppose that we have computed approximations $x^{(1)}, x^{(2)}, \dots, x^{(l)}$ to the solution $x^* \in \mathbb{R}^n$ of the system (3). Then the next approximation $x^{(l+1)}$ may be defined as either an exact or an inexact solution of the system of linear equations

$$Bx = Cx^{(l)} + b. \tag{4}$$

Considering that the CG method is quite efficient for solving an SPD system of linear equations whose coefficient matrix has a tightly clustered spectrum [14–16], Axelsson et al. [13] proposed to solve the system of linear equations (4) by the CG method and established the following theorem about the convergence properties of this method.

Theorem 2.1 (Theorem 3.1 in [13]). *Let $A \in \mathbb{R}^{n \times n}$ be a nonsingular and nonsymmetric matrix, and $A = B - C$ a contractive (with respect to the $\|\cdot\|_B$ -norm) and symmetric positive definite splitting. Suppose that the NSCG method is started from an initial guess $x^{(0)} \in \mathbb{R}^n$ and produces an iterative sequence $\{x^{(l)}\}_{l=0}^{\infty}$, where $x^{(l)} \in \mathbb{R}^n$ is the l th approximation to the solution $x^* \in \mathbb{R}^n$ of the system of linear equations (3), obtained by solving the linear system $Bx = Cx^{(l)} + b$ with k_l steps of CG iterations. Then*

- (a) $\|x^{(l)} - x^*\|_B \leq \gamma^{(l)} \|x^{(l-1)} - x^*\|_B, \quad l = 1, 2, 3, \dots,$
- (b) $\|b - Ax^{(l)}\|_B \leq \tilde{\gamma}^{(l)} \|b - Ax^{(l-1)}\|_B, \quad l = 1, 2, 3, \dots,$

where

$$\gamma^{(l)} = 2 \left(\frac{\sqrt{\kappa(B)} - 1}{\sqrt{\kappa(B)} + 1} \right)^{k_l} (1 + \varrho) + \varrho, \quad \tilde{\gamma}^{(l)} = \gamma^{(l)} \frac{1 + \varrho}{1 - \varrho}, \quad l = 1, 2, 3, \dots$$

and $\varrho = \|B^{-1}C\|_B$.

Moreover, for some $\gamma \in (\varrho, 1)$, and

$$k_l \geq \frac{\ln((\gamma - \varrho)/(2(1 + \varrho)))}{\ln\left(\frac{(\sqrt{\kappa(B)} - 1)}{(\sqrt{\kappa(B)} + 1)}\right)}, \quad l = 1, 2, 3, \dots,$$

we have $\gamma^{(l)} \leq \gamma$ ($l = 1, 2, 3, \dots$), and the sequence $\{x^{(l)}\}_{l=0}^{\infty}$ converges to the solution x^* of the system of linear equations (3).

For $\varrho \in (0, \sqrt{2} - 1)$ and some $\tilde{\gamma} \in ((1 + \varrho)\varrho/(1 - \varrho), 1)$, and

$$k_l \geq \frac{\ln(((1 - \varrho)\tilde{\gamma} - \varrho(1 + \varrho))/(2(1 + \varrho)^2))}{\ln\left(\frac{(\sqrt{\kappa(B)} - 1)}{(\sqrt{\kappa(B)} + 1)}\right)}, \quad l = 1, 2, 3, \dots,$$

we have $\tilde{\gamma}^{(l)} \leq \tilde{\gamma}$ ($l = 1, 2, 3, \dots$), and the residual sequence $\{b - Ax^{(l)}\}_{l=0}^{\infty}$ converges to zero.

As mentioned in [13] and motivated by [17], we may adopt regularization techniques to obtain a proper splitting of the coefficient matrix \mathcal{A} . For example, we can choose the splitting to be a quasi-Hermitian splitting [18]

$$\mathcal{A} = \mathcal{B}(\nu) - \mathcal{C}(\nu),$$

where

$$\mathcal{B}(\nu) = \mathcal{B} + \nu I, \quad \mathcal{C}(\nu) = \mathcal{C} + \nu I.$$

For a nonsymmetric and positive definite matrix \mathcal{A} , we can choose the splitting $\mathcal{A} = \mathcal{B} - \mathcal{C}$ as

$$\mathcal{B} = \frac{\mathcal{A}^T + \mathcal{A}}{2}, \quad \mathcal{C} = \frac{\mathcal{A}^T - \mathcal{A}}{2},$$

which are symmetric and skew-symmetric parts of matrix \mathcal{A} , respectively; see [19,20]. When the original coefficient matrix \mathcal{A} is symmetric, the NSCG iteration method naturally reduces to the RCG (regularized conjugate gradient) iteration method which was first proposed and studied in [17]. More details about the NSCG method can be found in [13].

3. NSCG method for $AXB = C$

In the sequel, we need the following lemmas.

Lemma 3.1 ([21]). Suppose that $A, B \in \mathbb{R}^{n \times n}$ be two symmetric matrices and denote the minimum and the maximum eigenvalues of a matrix M with $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$, respectively. Then

$$\begin{aligned} \lambda_{\max}(A + B) &\leq \lambda_{\max}(A) + \lambda_{\max}(B), \\ \lambda_{\min}(A + B) &\geq \lambda_{\min}(A) + \lambda_{\min}(B). \end{aligned}$$

Lemma 3.2 ([22]). Let $A, B \in \mathbb{R}^{n \times n}$, and λ and μ be the eigenvalues of A and B , and x and y be the corresponding eigenvectors, respectively. Then $\lambda\mu$ is an eigenvalue of $A \otimes B$ corresponding to the eigenvector $x \otimes y$.

Lemma 3.3 ([23]). Let $\mathcal{A} \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix. Then for all $x \in \mathbb{R}^n$, we have $\|\mathcal{A}^{\frac{1}{2}}x\|_2 = \|x\|_{\mathcal{A}}$ and

$$\sqrt{\lambda_{\min}(\mathcal{A})}\|x\|_{\mathcal{A}} \leq \|\mathcal{A}x\|_2 \leq \sqrt{\lambda_{\max}(\mathcal{A})}\|x\|_{\mathcal{A}}.$$

Consider the matrix equation

$$AXB = C, \tag{5}$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$ are nonsymmetric matrices, and $C, X \in \mathbb{R}^{n \times m}$. By using the symmetric and skew-symmetric splitting for coefficient matrices A and B , i.e., $A = H_A - S_A$ and $B = H_B - S_B$, the matrix equation (5) can be rewritten as

$$H_A X H_B + S_A X S_B = H_A X S_B + S_A X H_B + C. \tag{6}$$

By applying the vec operator [22], the matrix equation (6) is converted to

$$(H_B \otimes H_A - S_B \otimes S_A)x = (H_B \otimes S_A - S_B \otimes H_A)x + c, \tag{7}$$

where $x = vec(X)$ and $c = vec(C)$. Easily, we can see that

$$\begin{aligned} (H_B \otimes H_A - S_B \otimes S_A)^T &= H_B \otimes H_A - S_B \otimes S_A \\ (H_B \otimes S_A - S_B \otimes H_A)^T &= -(H_B \otimes S_A - S_B \otimes H_A). \end{aligned}$$

Therefore, $\mathcal{H} = H_B \otimes H_A - S_B \otimes S_A$ is symmetric and $\mathcal{S} = H_B \otimes S_A - S_B \otimes H_A$ is skew-symmetric. By using Lemmas 3.1 and 3.2, we can easily show that if

$$\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A)) > 0, \tag{8}$$

then the matrix \mathcal{H} is a symmetric positive definite matrix. By choosing $\mathcal{B} = \mathcal{H}$ and $\mathcal{C} = \mathcal{S}$, we obtain the following positive definite system of linear equations for the NSCG method,

$$(H_B \otimes H_A - S_B \otimes S_A)x = (H_B \otimes S_A - S_B \otimes H_A)x^{(l)} + c, \tag{9}$$

which can be arranged equivalently as

$$H_A X H_B + S_A X S_B = H_A X^{(l)} S_B + S_A X^{(l)} H_B + C. \tag{10}$$

For obtaining $X^{(l+1)}$, we can solve the matrix equation (10) iteratively by the CG-like method. Now, based on the above observations, we can establish the following algorithm for the NSCG method for solving the matrix equation (1).

3.1. The NSCG algorithm

An implementation of the NSCG method for the matrix equation (1) is given by the following algorithm. In the following algorithm, l_{\max} and j_{\max} are the largest admissible number of the outer and the inner iteration steps, respectively. $X^{(0)}$ is an initial guess for the solution, and the outer and the inner stopping tolerances are denoted by ε and η , respectively.

The NSCG algorithm for matrix equation $AXB = C$

1. $X^{(0,0)} = X^{(0)}$
2. $R^{(0)} = C - AX^{(0)}B$
3. For $l = 0, 1, \dots, l_{\max}$ Do:
4. $\hat{C} = C + H_A X^{(l,0)} S_B + S_A X^{(l,0)} H_B$
5. $\hat{R}^{(k,0)} = \hat{C} - H_A X^{(l,0)} H_B - S_A X^{(l,0)} S_B$
6. $P^{(0)} = \hat{R}^{(l,0)}$
7. For $j = 0, 1, \dots, j_{\max}$ Do:
8. $W^{(j)} = H_A P^{(j)} H_B + S_A P^{(j)} S_B$
9. $\alpha_j = \frac{(\hat{R}^{(l,j)}, \hat{R}^{(l,j)})_F}{(W^{(j)}, \hat{R}^{(l,j)})_F}$
10. $X^{(l,j+1)} = X^{(l,j)} + \alpha_j P^{(j)}$
11. $\hat{R}^{(l,j+1)} = \hat{R}^{(l,j)} - \alpha_j W^{(j)}$
12. If $\|\hat{R}^{(l,j+1)}\|_F \leq \eta \|\hat{R}^{(l,0)}\|_F$ GoTo 16
13. $\beta_j = \frac{(\hat{R}^{(l,j+1)}, \hat{R}^{(l,j+1)})_F}{(\hat{R}^{(l,j)}, \hat{R}^{(l,j)})_F}$
14. $P^{(j+1)} = \hat{R}^{(l,j+1)} + \beta_j P^{(j)}$
15. End Do
16. $X^{(l+1)} = X^{(l,j+1)}$
17. $R^{(l+1)} = C - AX^{(l+1)}B$
18. If $\|R^{(l+1)}\|_F \leq \varepsilon \|R^{(0)}\|_F$ Stop
19. $X^{(l+1,0)} = X^{(l+1)}$
20. End Do

We mention that Lines 7–15 in the above algorithm are similar to the global CG algorithm in [7]. For more details about the global CG iterative method, see [7,10].

3.2. Convergence analysis

In the sequel, we suppose that the condition (8) holds for the matrix equation (5). For the system of linear equations (3) when the coefficient matrix is non-Hermitian positive definite, Wang and Bai [24] presented sufficient conditions for the convergent splittings. Also, the convergence conditions for splitting iteration methods for the non-Hermitian coefficient matrix are studied in [25]. In the following lemma, we prove a similar result in the case of matrix equations for the Eq. (5).

Lemma 3.4. Let $\mathcal{H} = H_B \otimes H_A - S_B \otimes S_A$ and $\mathcal{S} = H_B \otimes S_A - S_B \otimes H_A$. If

$$\frac{\tau}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))} < \frac{1}{\theta^3},$$

where $\theta = \sqrt{\frac{\lambda_{\max}(H_B)\lambda_{\max}(H_A) + \max(-\lambda(S_B)\lambda(S_A))}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))}}$ and

$$\tau = (\lambda_{\max}^2(H_B) \max |\lambda(S_A)|^2 + \lambda_{\max}(H_B S_B \otimes S_A H_A + S_B H_B \otimes H_A S_A) + \lambda_{\max}^2(H_A) \max |\lambda(S_B)|^2)^{\frac{1}{2}},$$

then $\mathcal{A} = \mathcal{H} - \mathcal{S}$ is a contractive splitting (with respect to the $\|\cdot\|_{\mathcal{H}}$ -norm), i.e., $\|\mathcal{H}^{-1}\mathcal{S}\|_{\mathcal{H}} < 1$.

Proof. By Lemmas 3.1 and 3.2, we have

$$\|\mathcal{H}\|_2 = \lambda_{\max}(\mathcal{H}) \geq \lambda_{\min}(\mathcal{H}) \geq \lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A)),$$

and

$$\begin{aligned} \|\mathcal{S}\|_2^2 &= \rho(\mathcal{S}\mathcal{S}^T) \\ &= \max \lambda(\mathcal{S}\mathcal{S}^T) \\ &= \max \lambda(H_B H_B^T \otimes S_A S_A^T + H_B S_B \otimes S_A H_A + S_B H_B \otimes H_A S_A + S_B S_B^T \otimes H_A H_A^T) \\ &\leq \max \lambda(H_B H_B^T \otimes S_A S_A^T) + \max \lambda(H_B S_B \otimes S_A H_A + S_B H_B \otimes H_A S_A) + \max \lambda(S_B S_B^T \otimes H_A H_A^T) \end{aligned}$$

$$\begin{aligned}
 &= \max (\lambda^2(H_B)|\lambda(S_A)|^2) + \max \lambda(H_B S_B \otimes S_A H_A + S_B H_B \otimes H_A S_A) + \max (|\lambda(S_B)|^2 \lambda^2(H_A)) \\
 &\leq \lambda_{\max}^2(H_B) \max |\lambda(S_A)|^2 + \lambda_{\max}(H_B S_B \otimes S_A H_A + S_B H_B \otimes H_A S_A) + \lambda_{\max}^2(H_A) \max |\lambda(S_B)|^2.
 \end{aligned}$$

So we have

$$\begin{aligned}
 \|\mathcal{H}\|_2 &\leq (\lambda_{\max}^2(H_B) \max |\lambda(S_A)|^2 + \lambda_{\max}(H_B S_B \otimes S_A H_A + S_B H_B \otimes H_A S_A) + \lambda_{\max}^2(H_A) \max |\lambda(S_B)|^2)^{\frac{1}{2}} \\
 &= \tau.
 \end{aligned}$$

Therefore, it follows that

$$\begin{aligned}
 \|\mathcal{H}^{-1} \mathcal{H}\|_{\mathcal{H}} &\leq \sqrt{\kappa(\mathcal{H})} \|\mathcal{H}^{-1} \mathcal{H}\|_2 \\
 &\leq \sqrt{\kappa(\mathcal{H})} \|\mathcal{H}^{-1}\|_2 \|\mathcal{H}\|_2 \\
 &\leq (\kappa(\mathcal{H}))^{\frac{3}{2}} \frac{\|\mathcal{H}\|_2}{\|\mathcal{H}\|_2} \\
 &\leq (\kappa(\mathcal{H}))^{\frac{3}{2}} \frac{\tau}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))}.
 \end{aligned}$$

Again, the use of Lemmas 3.1 and 3.2 implies that

$$\sqrt{\kappa(\mathcal{H})} = \sqrt{\frac{\lambda_{\max}(\mathcal{H})}{\lambda_{\min}(\mathcal{H})}} \leq \sqrt{\frac{\lambda_{\max}(H_B)\lambda_{\max}(H_A) + \max(-\lambda(S_B)\lambda(S_A))}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))}} = \theta. \tag{11}$$

So, we can write

$$\|\mathcal{H}^{-1} \mathcal{H}\|_{\mathcal{H}} \leq \frac{\tau \theta^3}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))}. \tag{12}$$

This clearly proves the lemma. \square

Consider the following NSCG iterations for system (5)

$$H_A X^{(l+1)} H_B + S_A X^{(l+1)} S_B = H_A X^{(l)} S_B + S_A X^{(l)} H_B + C. \tag{13}$$

By using the Kronecker product, we can rewrite the above-described NSCG iteration in the following matrix–vector form:

$$(H_B \otimes H_A - S_B \otimes S_A) X^{(l+1)} = (H_B \otimes S_A - S_B \otimes H_A) X^{(l)} + c, \tag{14}$$

which can be arranged equivalently as

$$\mathcal{H} X^{(l+1)} = \mathcal{H} X^{(l)} + c. \tag{15}$$

Evidently, the iteration scheme (15) is a special case of the NSCG method for the system of linear equations $\mathcal{A}x = c$, with $\mathcal{A} = \mathcal{H} - \mathcal{H}$; see [13].

Let $\varrho = \|\mathcal{H}^{-1} \mathcal{H}\|_{\mathcal{H}}$ and

$$\eta = \frac{\tau \theta^3}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))}.$$

When $\eta < 1$, from the Proof of Lemma 3.4 (relation (12)), we have $\varrho \leq \eta < 1$. Therefore, $\mathcal{A} = \mathcal{H} - \mathcal{H}$ is a contractive (in the $\|\cdot\|_{\mathcal{H}}$ norm) and symmetric positive definite splitting. In this case, Lemma 3.3 and Part (a) of Theorem 2.1 imply that

$$\begin{aligned}
 \frac{1}{\sqrt{\lambda_{\max}(\mathcal{H})}} \|\mathcal{H}(x^{(l)} - x^*)\|_2 &\leq \|x^{(l)} - x^*\|_{\mathcal{H}} \\
 &\leq \gamma^{(l)} \|x^{(l-1)} - x^*\|_{\mathcal{H}} \\
 &\leq \frac{\gamma^{(l)}}{\sqrt{\lambda_{\min}(\mathcal{H})}} \|\mathcal{H}(x^{(l-1)} - x^*)\|_2,
 \end{aligned}$$

where

$$\gamma^{(l)} = 2 \left(\frac{\sqrt{\kappa(\mathcal{H})} - 1}{\sqrt{\kappa(\mathcal{H})} + 1} \right)^{k_l} (1 + \varrho) + \varrho.$$

So, Part (a) in Theorem 2.1 can be written as

$$\|\mathcal{H}(x^{(l)} - x^*)\|_2 \leq \gamma^{(l)} \frac{\sqrt{\lambda_{\max}(\mathcal{H})}}{\sqrt{\lambda_{\min}(\mathcal{H})}} \|\mathcal{H}(x^{(l-1)} - x^*)\|_2. \tag{16}$$

Furthermore,

$$\begin{aligned} \gamma^{(l)} &\leq 2 \left(\frac{\sqrt{\frac{\lambda_{\max}(H_B)\lambda_{\max}(H_A) + \max(-\lambda(S_B)\lambda(S_A))}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))} - 1}}{\sqrt{\frac{\lambda_{\max}(H_B)\lambda_{\max}(H_A) + \max(-\lambda(S_B)\lambda(S_A))}{\lambda_{\min}(H_B)\lambda_{\min}(H_A) + \min(-\lambda(S_B)\lambda(S_A))} + 1}} - 1 \right)^{k_l} (1 + \eta) + \eta \\ &= 2 \left(\frac{\theta - 1}{\theta + 1} \right)^{k_l} (1 + \eta) + \eta. \end{aligned}$$

Now, by using

$$H_A(X^{(l)} - X^*)H_B + S_A(X^{(l)} - X^*)S_B = H_A E^{(l)}H_B + S_A E^{(l)}S_B,$$

where $E^{(l)} = X^{(l)} - X^*$, the relation (16) can be arranged equivalently as

$$\|H_A E^{(l)}H_B + S_A E^{(l)}S_B\|_F \leq \omega^{(l)} \|H_A E^{(l-1)}H_B + S_A E^{(l-1)}S_B\|_F, \tag{17}$$

where

$$\omega^{(l)} = \left(2 \left(\frac{\theta - 1}{\theta + 1} \right)^{k_l} (1 + \eta) + \eta \right) \theta.$$

It is obvious that, for $\eta \in (0, \frac{1}{\theta})$ and $\omega \in (\eta\theta, 1)$, we will have $\omega^{(l)} \leq \omega$ if

$$k_l \geq \frac{\ln \left(\frac{\omega - \eta\theta}{2\theta(1+\eta)} \right)}{\ln \left(\frac{\theta-1}{\theta+1} \right)}, \quad l = 1, 2, 3, \dots$$

Under this restriction, from (17), we have

$$\begin{aligned} \|H_A E^{(l)}H_B + S_A E^{(l)}S_B\|_F &\leq \omega^{(l)} \|H_A E^{(l-1)}H_B + S_A E^{(l-1)}S_B\|_F \\ &\leq \prod_{k=0}^l \omega^{(k)} \|H_A E^{(0)}H_B + S_A E^{(0)}S_B\|_F \\ &\leq \omega^{l+1} \|H_A E^{(0)}H_B + S_A E^{(0)}S_B\|_F. \end{aligned}$$

Therefore, the sequence $\{X^{(l)}\}_{l=0}^\infty$ converges to the solution X^* of the system of matrix equations (5). Similarly, by using Part (b) of Theorem 2.1, for residual $r^{(l)} = c - \mathcal{A}X^{(l)}$, we obtain

$$\|\mathcal{H}r^{(l)}\|_2 \leq \tilde{\gamma}^{(l)} \frac{\sqrt{\lambda_{\max}(\mathcal{H})}}{\sqrt{\lambda_{\min}(\mathcal{H})}} \|\mathcal{H}r^{(l-1)}\|_2, \tag{18}$$

where $\tilde{\gamma}^{(l)} = \gamma^{(l)} \frac{1+\theta}{1-\theta}$, and

$$\tilde{\gamma}^{(l)} \leq \left(2 \left(\frac{\theta - 1}{\theta + 1} \right)^{k_l} (1 + \eta) + \eta \right) \frac{1 + \eta}{1 - \eta}.$$

The relation (18) can also be arranged equivalently as

$$\|H_A R^{(l)}H_B + S_A R^{(l)}S_B\|_F \leq \tilde{\omega}^{(l)} \|H_A R^{(l-1)}H_B + S_A R^{(l-1)}S_B\|_F, \tag{19}$$

where $R^{(l)} = C - \mathcal{A}X^{(l)}B$ and

$$\tilde{\omega}^{(l)} = \omega^{(l)} \frac{1 + \eta}{1 - \eta}.$$

As we observe, for $\eta \in \left(0, \frac{\sqrt{(\theta+1)^2 + 4\theta} - (\theta+1)}{2\theta} \right)$ we have $0 < (1 + \eta)\eta\theta / (1 - \eta) < 1$. So, for $\tilde{\omega} \in ((1 + \eta)\eta\theta / (1 - \eta), 1)$, we have $\tilde{\omega}^{(l)} \leq \tilde{\omega}$ if

$$k_l \geq \frac{\ln \left(\frac{\tilde{\omega}(1-\eta) - \eta\theta(1+\eta)}{2\theta(1+\eta)^2} \right)}{\ln \left(\frac{\theta-1}{\theta+1} \right)}, \quad l = 1, 2, 3, \dots,$$

Under this restriction, from (19), we have

$$\begin{aligned} \|H_A R^{(l)}H_B + S_A R^{(l)}S_B\|_F &\leq \tilde{\omega}^{(l)} \|H_A R^{(l-1)}H_B + S_A R^{(l-1)}S_B\|_F \\ &\leq \prod_{k=0}^l \tilde{\omega}^{(k)} \|H_A R^{(0)}H_B + S_A R^{(0)}S_B\|_F \\ &\leq \tilde{\omega}^{l+1} \|H_A R^{(0)}H_B + S_A R^{(0)}S_B\|_F. \end{aligned}$$

Therefore, the residual sequence $\{R^{(l)}\}_{l=0}^\infty$ converges to zero.

The above analysis is summarized in the following theorem.

Theorem 3.5. Suppose that for matrix equation (5), the symmetric and skew-symmetric parts of the coefficient matrices A and B satisfy the condition (8). Suppose that $\eta < 1$, and the NSCG method is started from an initial guess $X^{(0)} \in \mathbb{R}^{n \times m}$ and produces an iterative sequence $\{X^{(l)}\}_{l=0}^{\infty}$, where $X^{(l)} \in \mathbb{R}^{n \times m}$ is the l th approximation to the solution $X^* \in \mathbb{R}^{n \times m}$ of the matrix equations (5), by solving the linear system (13) with k_l steps of a CG-like iterations. Then

- (a) $\|H_A E^{(l)} H_B + S_A E^{(l)} S_B\|_F \leq \omega^{(l)} \|H_A E^{(l-1)} H_B + S_A E^{(l-1)} S_B\|_F, \quad l = 1, 2, \dots,$
- (b) $\|H_A R^{(l)} H_B + S_A R^{(l)} S_B\|_F \leq \tilde{\omega}^{(l)} \|H_A R^{(l-1)} H_B + S_A R^{(l-1)} S_B\|_F, \quad l = 1, 2, \dots,$

where

$$\omega^{(l)} = \left(2 \left(\frac{\theta - 1}{\theta + 1} \right)^{k_l} (1 + \eta) + \eta \right) \theta, \quad \tilde{\omega}^{(l)} = \omega^{(l)} \frac{1 + \eta}{1 - \eta}, \quad l = 1, 2, 3, \dots,$$

$$\theta = \sqrt{\frac{\lambda_{\max}(H_B) \lambda_{\max}(H_A) + \max(-\lambda(S_B) \lambda(S_A))}{\lambda_{\min}(H_B) \lambda_{\min}(H_A) + \min(-\lambda(S_B) \lambda(S_A))}}, \quad \eta = \frac{\tau \theta^3}{\lambda_{\min}(H_B) \lambda_{\min}(H_A) + \min(-\lambda(S_B) \lambda(S_A))}, \quad \text{and}$$

$$\tau = (\lambda_{\max}^2(H_B) \max |\lambda(S_A)|^2 + \lambda_{\max}(H_B S_B \otimes S_A H_A + S_B H_B \otimes H_A S_A) + \lambda_{\max}^2(H_A) \max |\lambda(S_B)|^2)^{\frac{1}{2}}.$$

Moreover, if $\eta \in (0, \frac{1}{\theta})$ then for any $\omega \in (\eta\theta, 1)$, and

$$k_l \geq \frac{\ln((\omega - \eta\theta)/(2\theta(1 + \eta)))}{\ln((\theta - 1)/(\theta + 1))}, \quad l = 1, 2, 3, \dots,$$

we have $\omega^{(l)} \leq \omega$ ($l = 1, 2, 3, \dots$), and the sequence $\{X^{(l)}\}_{l=0}^{\infty}$ converges to the solution X^* of the matrix equation (5). For $\eta \in (0, \frac{\sqrt{(\theta+1)^2 + 4\theta} - (\theta+1)}{2\theta})$ and any $\tilde{\omega} \in ((1 + \eta)\eta\theta/(1 - \eta), 1)$, and

$$k_l \geq \frac{\ln((\tilde{\omega}(1 - \eta) - \theta\eta(1 + \eta))/(2\theta(1 + \eta)^2))}{\ln((\theta - 1)/(\theta + 1))}, \quad l = 1, 2, 3, \dots,$$

we have $\tilde{\omega}^{(l)} \leq \tilde{\omega}$ ($l = 1, 2, 3, \dots$), and the residual sequence $\{R^{(l)}\}_{l=0}^{\infty}$ converges to zero.

3.3. Using the quasi-Hermitian splitting as a preconditioner

From the fact that any matrix splitting can naturally induce a splitting preconditioner for the Krylov subspace method (see [26]), in Section 4, by choosing the quasi-Hermitian splitting

$$\mathcal{A} = \mathcal{H}(v) - \mathcal{S}(v), \tag{20}$$

where

$$\mathcal{H}(v) = H_B(v) \otimes H_A(v) - S_B(v) \otimes S_A(v), \quad \mathcal{S}(v) = H_B(v) \otimes S_A(v) - S_B(v) \otimes H_A(v),$$

with regularization parameter v , and numerical computations, we show that the use of the quasi-Hermitian splitting (20) as a splitting preconditioner can induce accurate, robust, and effective preconditioned Krylov subspace iteration methods for solving matrix equation (1).

4. Numerical experiments

All numerical experiments presented in this section were computed in double precision with a number of MATLAB codes. All iterations are started from the zero matrix for initial $X^{(0)}$ and terminated when the current iterate satisfies $\frac{\|R^{(l)}\|_F}{\|R^{(0)}\|_F} \leq \varepsilon = 10^{-8}$, where $R^{(l)}$ is the residual of the l th iterate. Also we use the tolerance $\eta = 0.01$ for inner iterations. For each experiment we report the number of iterations or the number of total outer iteration steps and CPU time (in parentheses) and compare the new method with the BiCGSTAB and preconditioned BiCGSTAB (PBiCGSTAB) methods for the system (1). In the preconditioned BiCGSTAB method we used the NSCG method with a stopping criterion $\epsilon = 0.01$ as a preconditioner. Dagger (\dagger) and notation “> 1000” show that no solution has been obtained after 5000 iteration or CPU time is more than 1000 s respectively. For a regularized NSCG method (RNSCG) and a regularized preconditioned BiCGSTAB (PBiCGSTAB(v)) method, the experimentally computed optimal value of regularization parameter v was used.

For the first example, we used the matrices

$$A = M + 2rN + \frac{100}{(n + 1)^2}I \quad \text{and} \quad B = M + 2rN + \frac{100}{(m + 1)^2}I,$$

where $M = \text{tridiag}(-1, 2, -1)$, $N = \text{tridiag}(0.5, 0, -0.5)$, and $r = 0.01$ [5]. We apply the iteration methods to this problem with different dimensions, and the results are given in Table 1. The pair (n, m) in the first row of Table 1 represents the

Table 1
Results of the first set of examples.

Method	(256, 16)	(256, 32)	(256, 64)	(256, 128)
NSCG	8 (4.830)	8 (9.417)	14 (33.451)	41 (229.664)
RNSCG	8 (4.613)	8 (9.282)	14 (32.891)	41 (227.361)
BiCGSTAB	783 (2.012)	1715 (9.869)	3325 (49.873)	8977 (397.630)
PBiCGSTAB	4 (2.574)	4 (9.063)	5 (29.312)	9 (189.868)

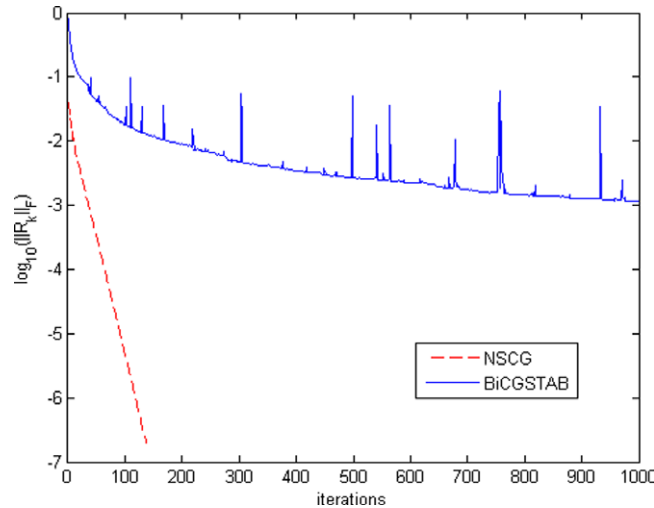


Fig. 1. Convergence history of NSCG versus BiCGSTAB for the first experiment with $m = n = 256$.

Table 2
Results of the second set of examples.

Method	$m = 100$	$m = 200$	$m = 300$	$m = 400$	$m = 500$
NSCG	4 (10.632)	4 (34.164)	4 (79.139)	4 (123.927)	4 (200.211)
RNSCG	4 (8.133)	4 (28.109)	4 (67.311)	4 (116.350)	4 (162.513)
BiCGSTAB	265 (15.475)	301 (54.225)	401 (153.816)	461 (352.349)	561 (750.286)
PBiCGSTAB	7 (6.146)	8 (24.008)	6 (56.534)	7 (113.256)	6 (137.920)

dimension of matrices A and B , respectively. From Table 1, we observe that the NSCG and RNSCG methods are more efficient than the BiCGSTAB method. In addition, the use of the quasi-Hermitian splitting as a preconditioner for the BiCGSTAB method leads to more efficient results compared to the other methods. Moreover, the comparison between the CPU time of the NSCG and RNSCG methods shows that the use of a good regularizing parameter can reduce the number of inner iterations and improve the CPU time. The convergence history of the NSCG method together with that of the BiCGSTAB method, for this experiment with $m = n = 256$, is illustrated in Fig. 1. From Fig. 1, one can observe that the NSCG method is more efficient than the BiCGSTAB method.

For the second example, we used the matrices

$$A = \text{tridiag} \left(1 - \frac{10}{(n+1)}, -2, 1 - \frac{10}{(n+1)} \right) \quad \text{and} \quad B = \text{tridiag} \left(-1 + \frac{10}{(m+1)}, 2, -1 + \frac{10}{(m+1)} \right).$$

We set $n = 500$ and apply the iteration methods to this problem with different dimensions m for the matrix B . For this example, the results are listed in Table 2. Again, it is obvious that the NSCG and RNSCG methods have much better convergence property than the BiCGSTAB method in terms of the CPU-time and the number of iterations required for convergence. In addition, the use of the quasi-Hermitian splitting as a preconditioner for the BiCGSTAB method improves the efficiency of the BiCGSTAB method. We also observe that the CPU time of the RNSCG method is less than that of the NSCG method which is due to the inner iterations of these algorithms.

Finally, in the following experiment, we test the numerical behavior of the quasi-Hermitian splitting (20), as a splitting preconditioner, by solving the matrix equation (1) with the preconditioned BiCGSTAB method and its regularized variant with the optimal value of ν .

For this experiment, for the coefficient matrix A , we used the nonsymmetric matrix SHERMAN3 of dimension 5005×5005 with 20033 nonzero entries from the Harwell–Boeing collection [27]. For the coefficient matrix B , we used the identity matrix of dimension $m = 4$. The CPU-time and the number of iterations required for convergence are presented in Table 3. For this test problem, we observe that the BiCGSTAB iteration does not converge, and the preconditioned BiCGSTAB (PBiCGSTAB) and

Table 3
Results of the last experiment.

Method	Iteration	CPU time	Residual-norm
NSCG	36	> 1000	1.2105e–004
RNSCG	30	> 1000	1.6723e–004
BiCGSTAB	†	†	NaN
PBiCGSTAB	10	54.366	6.3434e–005
PBiCGSTAB(ν)	6	46.067	1.8942e–004

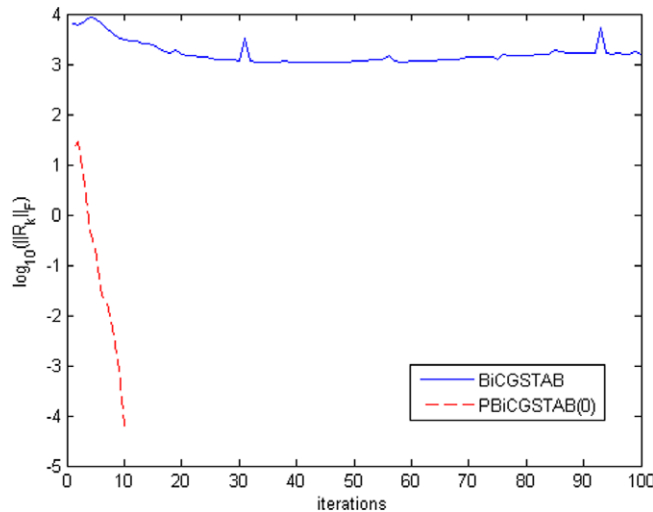


Fig. 2. Convergence history of BiCGSTAB versus PBiCGSTAB for $A = \text{SHERMAN3}$ and $B = I_{4 \times 4}$.

its regularized variant (PBiCGSTAB(ν)) based on the quasi-Hermitian splitting (20) are superior to the other methods. Fig. 2 shows the residuals obtained by the PBiCGSTAB and BiCGSTAB methods. From this figure, we can see that the PBiCGSTAB method is clearly superior.

In summary, by focusing on the results presented in Tables 1–3, one can observe that the NSCG method and its regularized variant are often superior to the BiCGSTAB method. However, the use of the quasi-Hermitian splitting as a preconditioner can induce an accurate and effective preconditioned BiCGSTAB method.

5. Conclusion

In this paper, we have proposed an efficient iterative method for solving the matrix equation (1). This method, which is based on the NSCG method, employs a CG-like iteration method as the inner iteration to approximate each outer iterate, while each outer iteration is induced by a convergent and symmetric positive definite splitting of coefficient matrices. When the symmetric and skew-symmetric parts of the coefficient matrices of the matrix equation satisfy the condition (8), we proved that the NSCG method for the matrix equation (1) converges to the exact solution.

In addition, numerical computations showed that the quasi-Hermitian splitting can induce accurate, robust and effective preconditioned Krylov subspace methods.

Acknowledgments

The authors are grateful to the anonymous referees for valuable comments and suggestions which substantially improved the quality of this paper.

References

- [1] U. Baur, P. Benner, Cross-gramian based model reduction for data-sparse systems, *Electron. Trans. Numer. Anal.* 31 (2008) 256–270.
- [2] A. Bouhamidi, K. Jbilou, L. Reichel, H. Sadok, An extrapolated TSVD method for linear discrete ill-posed problems with Kronecker structure, *Linear Algebra Appl.* 434 (2011) 1677–1688.
- [3] D. Calvetti, L. Reichel, Application of ADI iterative methods to the restoration of noisy images, *SIAM J. Matrix Anal. Appl.* 17 (1996) 165–186.
- [4] B. Datta, *Numerical Methods for Linear Control Systems*, Elsevier Academic Press, 2004.
- [5] Z.-Z. Bai, On Hermitian and skew-Hermitian splitting iteration methods for continuous Sylvester equations, *J. Comput. Math.* 29 (2011) 185–198.
- [6] Y.-B. Deng, Z.-Z. Bai, Y.-H. Gao, Iterative orthogonal direction methods for Hermitian minimum norm solutions of two consistent matrix equations, *Numer. Linear Algebra Appl.* 13 (2006) 801–823.

- [7] A. El Guennoui, K. Jbilou, J. Riquet, Block Krylov subspace methods for solving large Sylvester equation, *Numer. Algorithms* 29 (2002) 75–96.
- [8] A.-P. Liao, Z.-Z. Bai, Y. Lei, Best approximate solution of matrix equation $AXB + CYD = E$, *SIAM J. Matrix Anal. Appl.* 27 (2005) 675–688.
- [9] D.K. Salkuyeh, F. Toutounian, New approaches for solving large Sylvester equations, *Appl. Math. Comput.* 173 (2006) 9–18.
- [10] D.K. Salkuyeh, CG-type algorithms to solve symmetric matrix equations, *Appl. Math. Comput.* 172 (2006) 985–999.
- [11] M. Wang, Y. Feng, An iterative algorithm for solving a class of matrix equations, *J. Control Theory Appl.* 7 (2009) 68–72.
- [12] L. Xie, J. Ding, F. Ding, Gradient based iterative solutions for general linear matrix equations, *Comput. Math. Appl.* 58 (2009) 1441–1448.
- [13] O. Axelsson, Z.-Z. Bai, S.-X. Qiu, A class of nested iterative schemes for linear systems with a coefficient matrix with a dominant positive definite symmetric part, *Numer. Algorithms* 35 (2004) 351–372.
- [14] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, New York, 1994.
- [15] G.H. Golub, C. Van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore, 1996.
- [16] Y. Saad, *Iterative Methods for Large Sparse Linear Systems*, SIAM, Philadelphia, 2002.
- [17] Z.-Z. Bai, S.-L. Zhang, A regularized conjugate gradient method for symmetric positive definite system of linear equations, *J. Comput. Math.* 20 (2002) 437–448.
- [18] G.H. Golub, M.L. Overton, The convergence of inexact Chebyshev and Richardson iterative methods for solving linear systems, *Numer. Math.* 53 (1988) 571–593.
- [19] Z.-Z. Bai, G.H. Golub, M.K. Ng, Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *SIAM J. Matrix Anal. Appl.* 24 (2003) 603–626.
- [20] Z.-Z. Bai, Splitting iteration methods for non-Hermitian positive definite systems of linear equations, *Hokkaido Math. J.* 36 (2007) 801–814.
- [21] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
- [22] H. Lütkepohl, *Handbook of Matrices*, John Wiley & Sons Press, England, 1996.
- [23] C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, in: *Frontiers in Applied Mathematics*, vol. 16, SIAM, Philadelphia, 1995.
- [24] C.-L. Wang, Z.-Z. Bai, Sufficient conditions for the convergent splittings of non-Hermitian positive definite matrices, *Linear Algebra Appl.* 330 (2001) 215–218.
- [25] L. Wang, Z.-Z. Bai, Convergence conditions for splitting iteration methods for non-Hermitian linear systems, *Linear Algebra Appl.* 428 (2008) 453–468.
- [26] Z.-Z. Bai, J.-F. Yin, Y.-F. Su, A shift-splitting preconditioner for non-Hermitian positive definite matrices, *J. Comput. Math.* 24 (2006) 539–552.
- [27] I.S. Duff, R.G. Grimes, J.G. Lewis, *User's guide for the Harwell–Boeing sparse matrix collection*, Technical Report RAL-92-086, Rutherford Applton Laboratory, Chilton, UK, 1992.