Adv Comput Math (2013) 40:865–880
DOI 10.1007/s10444-013-9330-3

Nested splitting CG-like iterative method for solving the continuous Sylvester equation and preconditioning

Mohammad Khorsand Zak · Faezeh Toutounian

Received: 17 February 2013 / Accepted: 8 October 2013 /
Published online: 23 October 2013

Abstract We present a nested splitting conjugate gradient iteration method for solving large sparse continuous Sylvester equation, in which both coefficient matrices are (non-Hermitian) positive semi-definite, and at least one of them is positive definite. This method is actually inner/outer iterations, which employs the Sylvester conjugate gradient method as inner iteration to approximate each outer iterate, while each outer iteration is induced by a convergent and Hermitian positive definite splitting of the coefficient matrices. Convergence conditions of this method are studied and numerical experiments show the efficiency of this method. In addition, we show that the quasi-Hermitian splitting can induce accurate, robust and effective preconditioned Krylov subspace methods.

Keywords Sylvester equation · Preconditioning · Conjugate gradient · Nested iterations

Mathematics Subject Classifications (2010) 15A24 · 15A30 · 15A69 · 65F10 · 65F30

Communicated by: J. M. Peña

M. Khorsand Zak (✉) · F. Toutounian
Department of Applied Mathematics, School of Mathematical Sciences,
Ferdowsi University of Mashhad, Mashhad, Iran
e-mail: mo.khorsand@stu-mail.um.ac.ir

F. Toutounian
The Center of Excellence on Modelling and Control Systems,
Ferdowsi University of Mashhad, Mashhad, Iran
e-mail: toutouni@math.um.ac.ir

1 Introduction

Consider the continuous Sylvester equation

$$AX + XB = C, \tag{1}$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{m\times m}$ and $C \in \mathbb{C}^{n\times m}$ are defined matrices and $X \in \mathbb{C}^{n\times m}$ is an unknown matrix. A Lyapunov equation is a special case with $m = n$, $B = A^H$, and $C = C^H$. Here and in the sequel, W^T (W^H) is used to denote the transpose (conjugate transpose) of the matrix $W \in \mathbb{C}^{m\times m}$. Equation (1) has a unique solution if and only if A and $-B$ have no common eigenvalues, which will be assumed throughout this paper. The Sylvester equation appears frequently in many areas of applied mathematics and plays vital roles in a number of applications such as control theory, model reduction and image processing, see [10–12] and their references.

The matrix Eq. (1) is mathematically equivalent to the linear system of equations

$$\mathcal{A}x = c, \tag{2}$$

where the matrix $\mathcal{A}$ is of dimension $nm \times nm$ and is given by

$$\mathcal{A} = I_m \otimes A + B^T \otimes I_n, \tag{3}$$

where $\otimes$ denotes the Kronecker product ($A \otimes B = [a_{ij}B]$) and

$$c = vec(C) = (c_{11}, c_{21}, \cdots, c_{n1}, c_{12}, c_{22}, \cdots, c_{n2}, \cdots, c_{nm})^T,$$
$$x = vec(X) = (x_{11}, x_{21}, \cdots, x_{n1}, x_{12}, x_{22}, \cdots, x_{n2}, \cdots, x_{nm})^T.$$

Of course, this is a numerically poor way to determine the solution X of the Sylvester Eq. (1), as the linear system of Eq. (2) is costly to solve and can be ill-conditioned.

Standard methods for numerical solution of the Sylvester Eq. (1) are the Bartels-Stewart [8] and the Hessenberg-Schur [17] methods, which consist of transforming coefficient matrices A and B into triangular or Hessenberg form by an orthogonal similarity transformation and then solving the resulting system directly by a back-substitution process. These methods are classified as direct methods and are applicable and effective for general Sylvester equations of reasonably small size. When the coefficient matrices A and B are large and sparse, iterative methods such as the alternating direction implicit (ADI) method [9], the Krylov subspace based algorithms [14, 20, 21, 26], the preconditioned conjugate gradient method [16], the Hermitian and skew-Hermitian splitting (HSS) method, and the inexact variants (IHSS) of HSS iteration method [3, 5] are often the methods of choice for efficiently and accurately solving the Sylvester Eq. (1).

In this paper, we present an iterative method for solving the Sylvester Eq. (1) by using the Hermitian and skew-Hermitian splitting of the matrices A and B in a matrix variant of the nested splitting conjugate gradient method. A class of nested splitting conjugate gradient (NSCG) method was first proposed in [1] for solving large sparse linear systems of equations and in [23] the NSCG method was investigated for a class of matrix equations. Via the NSCG method, the solution of the linear system of Eq. (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 $\mathcal{H} = \frac{1}{2}(\mathcal{A}^H + \mathcal{A})$

and $\mathcal{S} = \frac{1}{2}(\mathcal{A}^H - \mathcal{A})$. Then, at each step of iteration for matrix case, a Sylvester equation with coefficient matrices H_A and H_B (which are the Hermitian part of A and B, respectively) is solved iteratively by the Sylvester conjugate gradient method. When the matrices A and B are positive semi-definite, and at least one of them is positive definite, we prove that the iterations converge to the exact solution of the Sylvester Eq. (1). Numerical examples are given to further illustrate the effectiveness and robustness of the NSCG method over some well-known iterative methods when they are applied to the Sylvester equation. Moreover, we show that the quasi-Hermitian splitting can induce accurate, robust and effective preconditioned Krylov subspace methods.

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 $M \in \mathbb{C}^{n\times n}$, and the identity matrix with dimension n, respectively. Note that $||.||_2$ is also used to represent the 2-norm of a vector. For a Hermitian positive definite matrix $\mathcal{B}$, we denote by $\kappa(\mathcal{B}) = ||\mathcal{B}||_2||\mathcal{B}^{-1}||_2$ its Euclidean condition number, and we define the $||\cdot||_{\mathcal{B}}$ norm of a vector $x \in \mathbb{C}^n$ as $||x||_{\mathcal{B}} = \sqrt{x^H\mathcal{B}x}$. Then the induced $||\cdot||_{\mathcal{B}}$ norm of a matrix $H \in \mathbb{C}^{n\times n}$ is define as $||H||_{\mathcal{B}} = \left\|\mathcal{B}^{\frac{1}{2}}H\mathcal{B}^{-\frac{1}{2}}\right\|_2$. 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 equivalent relationships between the Frobenius norm of a matrix R and the 2-norm of a vector $r = vec(R)$:

$$||r||_2 = \sqrt{\sum_{i=1}^{mn}|r_i|^2} \Leftrightarrow ||R||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|R_{ij}|^2}.$$

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 the continuous Sylvester equation and its convergence properties. Section 4 is devoted to numerical experiments. Finally, we present our conclusions in Section 5.

2 The NSCG method

A matrix $\mathcal{A} \in \mathbb{R}^{n\times n}$ is called positive definite if for all $x \in \mathbb{R}^n$ and $x \neq 0$ we have $x^T\mathcal{A}x > 0$. $\mathcal{A} = \mathcal{B} - \mathcal{C}$ is called a splitting of the matrix $\mathcal{A}$ if $\mathcal{B}$ is nonsingular. This splitting is a symmetric positive definite splitting if $\mathcal{B}$ is a symmetric positive definite matrix; a convergent splitting if $\rho(\mathcal{B}^{-1}\mathcal{C}) < 1$; and a contractive splitting if $\left\|\mathcal{B}^{-1}\mathcal{C}\right\| < 1$ for some matrix norm.

Let $\mathcal{A} = \mathcal{B} - \mathcal{C}$ be a symmetric positive definite splitting of the coefficient matrix $\mathcal{A}$. Then the linear system of Eq. (2) is equivalent to the fixed-point equation

$$\mathcal{B}x = \mathcal{C}x + c.$$

Assume that this splitting satisfies the condition $\rho(\mathcal{B}^{-1}\mathcal{C}) < 1$. Given an initial guess $x^{(0)} \in \mathbb{R}^n$, suppose that we have computed approximations $x^{(1)}, x^{(2)}, \cdots, x^{(l)}$ to the

solution $x^* \in \mathbb{R}^n$ of the system Eq. (2). Then the next approximation $x^{(l+1)}$ may be defined as either an exact or an inexact solution of the linear system of equations

$$\mathcal{B}x = \mathcal{C}x^{(l)} + c. \tag{4}$$

In [1], the authors proposed to solve the linear system of Eq. (4) by the CG method and they established the following theorem about the convergence properties of this method which called NSCG method.

Theorem 2.1 [1] *Let $\mathcal{A} \in \mathbb{R}^{n\times n}$ be a nonsingular and non-symmetric matrix, and $\mathcal{A} = \mathcal{B} - \mathcal{C}$ a contractive (with respect to the $||\cdot||_{\mathcal{B}}$-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 lth approximation to the solution $x^* \in \mathbb{R}^n$ of the system of linear* Eq. (2)*, obtained by solving the linear system $\mathcal{B}x = \mathcal{C}x^{(l)} + c$ with k_l steps of CG iterations. Then*

(a) $\left\|x^{(l)} - x^*\right\|_{\mathcal{B}} \leq \gamma^{(l)} \left\|x^{(l-1)} - x^*\right\|_{\mathcal{B}}, \quad l = 1, 2, 3, \cdots,$

(b) $\left\|c - \mathcal{A}x^{(l)}\right\|_{\mathcal{B}} \leq \widetilde{\gamma}^{(l)} \left\|c - \mathcal{A}x^{(l-1)}\right\|_{\mathcal{B}}, \quad l = 1, 2, 3, \cdots,$

where

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

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

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

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

we have $\gamma^{(l)} \leq \gamma$ $(l = 1, 2, 3, \cdots)$, and the sequence $\{x^{(l)}\}_{l=0}^{\infty}$ converges to the solution x^* of the system of linear Eq. (2)*. For $\varrho \in (0, \sqrt{2} - 1)$ and some $\widetilde{\gamma} \in ((1 + \varrho)\varrho/(1 - \varrho), 1)$, and*

$$k_l \geq \frac{\ln(((1 - \varrho)\widetilde{\gamma} - \varrho(1 + \varrho))/(2(1 + \varrho)^2))}{\ln((\sqrt{\kappa(\mathcal{B})} - 1)/(\sqrt{\kappa(\mathcal{B})} + 1))}, \quad l = 1, 2, 3, \cdots,$$

we have $\widetilde{\gamma}^{(l)} \leq \widetilde{\gamma}$ $(l = 1, 2, 3, \cdots)$, and the residual sequence $\{c - \mathcal{A}x^{(l)}\}_{l=0}^{\infty}$ converges to zero.

As mentioned in [1] and motivated by [4], 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 non-symmetric 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 [5] and [2]. When the original coefficient matrix $\mathcal{A}$ is symmetric, the NSCG iteration method naturally reduces to the RCG (regularized conjugate gradient) iteration method which first proposed and studied in [4]. More details about the NSCG method can be found in [1].

3 The NSCG method for the Sylvester equation

In this section, we establish the NSCG method for solving the Sylvester Eq. (1). Here and in the sequel, we call W a positive definite or positive semi-definite matrix if so is its Hermitian part $H_W = \frac{1}{2}(W + W^H)$; note that a positive definite or positive semi-definite matrix is not necessarily Hermitian. We suppose that both coefficient matrices $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$ are (non-Hermitian) positive semi-definite, and at least one of them is positive definite. It is easy to show that under these assumptions the coefficient matrix $\mathcal{A}$ defined in Eq. (3) is positive definite. Now, we consider the Hermitian and skew-Hermitian splitting $\mathcal{A} = \mathcal{H} - \mathcal{S}$, where

$$\mathcal{H} = \frac{\mathcal{A}^H + \mathcal{A}}{2}, \quad \mathcal{S} = \frac{\mathcal{A}^H - \mathcal{A}}{2}, \tag{5}$$

are the Hermitian and skew-Hermitian parts of matrix $\mathcal{A}$, respectively, see [5] and [2]. From Eqs. (3) and (5), by using the Kronecker product's properties [19], we have

$$\mathcal{H} = I_m \otimes H_A + H_{B^T} \otimes I_n, \tag{6}$$

$$\mathcal{S} = I_m \otimes S_A + S_{B^T} \otimes I_n, \tag{7}$$

where H_A, S_A, H_{B^T}, S_{B^T} are the Hermitian and skew-Hermitian parts of A and B^T, respectively. By choosing $\mathcal{B} = \mathcal{H}$ and $\mathcal{C} = \mathcal{S}$ and using the Eqs. (6) and (7), we obtain the following positive definite linear system of equations

$$(I_m \otimes H_A + H_{B^T} \otimes I_n)x = (I_m \otimes S_A + S_{B^T} \otimes I_n)x^{(l)} + c, \tag{8}$$

which can be arranged equivalently as

$$H_A X + X H_B = S_A X^{(l)} + X^{(l)} S_B + C. \tag{9}$$

Under the assumption A and B are positive semi-definite and at last one of them is positive definite, we can easily see that there is no common eigenvalue between the matrices H_A and $-H_B$, so the Sylvester Eq. (9) has unique solution for all given right hand side matrices. For obtaining $X^{(l+1)}$, we can solve the matrix Eq. (9) iteratively by the Sylvester CG method [16]. Now, based on the above observations, we can

establish the following algorithm for the NSCG method for solving the continuous Sylvester Eq. (1).

3.1 Implementation of the NSCG method

An implementation of the NSCG method is given by the following algorithm. In the following algorithm, $k_{\max}$ 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.

Algorithm 1 The NSCG algorithm for the continuous Sylvester equation

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

3.2 Convergence analysis

In the sequel, we need the following lemmas.

Lemma 3.1 [22] *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 $\left\|\mathcal{A}^{\frac{1}{2}}x\right\|_2 = ||x||_{\mathcal{A}}$ and*

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

Lemma 3.2 [25] *Suppose that $A, B \in \mathbb{C}^{n\times n}$ be two Hermitian matrices, and denote the minimum and the maximum eigenvalues of a matrix M with $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$, respectively. Then*

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

Lemma 3.3 [25] *Let $A, B \in \mathbb{C}^{n\times n}$, 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 $x \otimes y$.*

For the linear system of Eq. (2) when the coefficient matrix is non-Hermitian positive definite, Wang and Bai [27] presented sufficient conditions for the convergent splittings. Also, the convergence conditions for splitting iteration methods for non-Hermitian coefficient matrix are studied in [29]. By noting that Theorem 2.1and the Lemmas 3.1 hold also in the complex case for Hermitian positive definite matrices, in the following lemma, we prove a similar result in the case of matrix equations for the Sylvester Eq. (1).

Lemma 3.4 *Suppose that $\mathcal{H}$ and $\mathcal{S}$ are as in* Eqs. (6) *and* (7)*. If*

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

where $\theta = \sqrt{\frac{\lambda_{\max}(H_A)+\lambda_{\max}(H_B)}{\lambda_{\min}(H_A)+\lambda_{\min}(H_B)}}$, 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 From Eqs. (6) and (7), by Lemmas 3.2 and 3.3, we have

$$||\mathcal{H}||_2 = \lambda_{\max}(\mathcal{H}) \ge \lambda_{\min}(\mathcal{H}) \ge \lambda_{\min}(H_A)+\lambda_{\min}(H_B),$$

and

$$||\mathcal{S}||_2 = \max_{\lambda\in\Lambda(\mathcal{S})}|\lambda(\mathcal{S})| \le \max|\lambda(S_A)|+\max|\lambda(S_B)|,$$

where $\Lambda(\mathcal{S})$ is the set of all eigenvalues of $\mathcal{S}$. Therefore, it follows that

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

Again, the use of Lemmas 3.2 and 3.3 implies that

$$\sqrt{\kappa(\mathcal{H})} = \sqrt{\frac{\lambda_{\max}(\mathcal{H})}{\lambda_{\min}(\mathcal{H})}} \le \sqrt{\frac{\lambda_{\max}(H_A)+\lambda_{\max}(H_B)}{\lambda_{\min}(H_A)+\lambda_{\min}(H_B)}} = \theta. \tag{10}$$

So, we can write

$$\left\|\mathcal{H}^{-1}\mathcal{S}\right\|_{\mathcal{H}} \leq \theta^3 \frac{\max|\lambda(S_A)| + \max|\lambda(S_B)|}{\lambda_{\min}(H_A) + \lambda_{\min}(H_B)}. \tag{11}$$

This clearly proves the lemma. □

Let $\eta = \theta^3 \frac{\max|\lambda(S_A)|+\max|\lambda(S_B)|}{\lambda_{\min}(H_A)+\lambda_{\min}(H_B)}$, and $\varrho = ||\mathcal{H}^{-1}\mathcal{S}||_{\mathcal{H}}$. When $\eta < 1$, from the proof of Lemma 3.4 (Eq. (11)), we have $\varrho \leq \eta < 1$. Therefore, $\mathcal{A} = \mathcal{H} - \mathcal{S}$ is a contractive (in the $||\cdot||_{\mathcal{H}}$-norm) and Hermitian positive definite splitting. In this case, Lemma 3.1 and part (a) of Theorem 2.1 imply that

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

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

$$||\mathcal{H}(x^{(l)} - x^*)||_2 \leq \gamma^{(l)} \frac{\sqrt{\lambda_{\max}(\mathcal{H})}}{\sqrt{\lambda_{\min}(\mathcal{H})}}||\mathcal{H}(x^{(l-1)} - x^*)||_2. \tag{12}$$

Furthermore,

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

Now, by using

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

where $E^{(l)} = X^{(l)} - X^*$, the Eq. (12) can be arranged equivalently as

$$||H_A E^{(l)} + E^{(l)}H_B||_F \leq \omega^{(l)}||H_A E^{(l-1)} + E^{(l-1)}H_B||_F, \tag{13}$$

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

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

Under this restriction, from Eq. (13), we have

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

Therefore, the sequence $\{X^{(l)}\}_{l=0}^{\infty}$ converges to the solution X^* of the Sylvester Eq. (1).

Similarly, by using part (b) of Theorem 2.1, for residual $r^{(l)} = c - \mathcal{A}x^{(l)}$, we obtain

$$||\mathcal{H}r^{(l)}||_2 \leq \widetilde{\gamma}^{(l)} \frac{\sqrt{\lambda_{\max}(\mathcal{H})}}{\sqrt{\lambda_{\min}(\mathcal{H})}} ||\mathcal{H}r^{(l-1)}||_2, \tag{14}$$

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

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

The Eq. (14) can also be arranged equivalently as

$$\left\|H_A R^{(l)} + R^{(l)} H_B\right\|_F \leq \widetilde{\omega}^{(l)} \left\|H_A R^{(l-1)} + R^{(l-1)} H_B\right\|_F, \tag{15}$$

where $R^{(l)} = C - AX^{(l)} - X^{(l)}B$ and

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

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 $\widetilde{\omega} \in ((1+\eta)\eta\theta/(1-\eta), 1)$, we have $\widetilde{\omega}^{(l)} \leq \widetilde{\omega}$ if

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

Under this restriction, from Eq. (15), we have

$$\begin{aligned}||H_A R^{(l)} + R^{(l)} H_B||_F &\leq \widetilde{\omega}^{(l)} \left\|H_A R^{(l-1)} + R^{(l-1)} H_B\right\|_F \\ &\leq \Pi_{k=0}^{l} \widetilde{\omega}^{(k)} \left\|H_A R^{(0)} + R^{(0)} H_B\right\|_F \\ &\leq \widetilde{\omega}^{l+1} \left\|H_A R^{(0)} + R^{(0)} H_B\right\|_F.\end{aligned}$$

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

The above analysis is summarized in the following theorem.

Theorem 3.5 *Consider the Sylvester* Eq. (1) *which both coefficient matrices A and B are (non-Hermitian) positive semi-definite, and at least one of them is positive definite. Let* H_A, S_A, H_B, *and* S_B *be the Hermitian and skew-Hermitian parts of*

A and B, respectively. Suppose that $\eta < 1$*, and the NSCG method is started from an initial guess* $X^{(0)} \in \mathbb{C}^{n\times m}$*, and produces an iterative sequence* $\{X^{(l)}\}_{l=0}^{\infty}$*, where* $X^{(l)} \in \mathbb{C}^{n\times m}$ *is the lth approximation to the solution* $X^* \in \mathbb{C}^{n\times m}$ *of the Sylvester* Eq. (1)*, by solving the Sylvester* Eq. (9) *with* k_l *steps of the Sylvester CG iterations. Then*

$$(a) \quad \left\|H_A E^{(l)} + E^{(l)} H_B\right\|_F \le \omega^{(l)} \left\|H_A E^{(l-1)} + E^{(l-1)} H_B\right\|_F, \quad l = 1, 2, 3, \cdots,$$
$$(b) \quad \left\|H_A R^{(l)} + R^{(l)} H_B\right\|_F \le \widetilde{\omega}^{(l)} \left\|H_A R^{(l-1)} + R^{(l-1)} H_B\right\|_F, \quad l = 1, 2, 3, \cdots,$$

where

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

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

Moreover, if $\eta \in \left(0, \frac{1}{\theta}\right)$ *then for some* $\omega \in (\eta\theta, 1)$*, and*

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

we have $\omega^{(l)} \le \omega$ $(l = 1, 2, 3, \cdots)$*, and the sequence* $\{X^{(l)}\}_{l=0}^{\infty}$ *converges to the solution* X^* *of the Sylvester* Eq. (1)*. For* $\eta \in \left(0, \frac{\sqrt{(\theta+1)^2+4\theta}-(\theta+1)}{2\theta}\right)$ *and some* $\widetilde{\omega} \in ((1+\eta)\eta\theta/(1-\eta), 1)$*, and*

$$k_l \ge \frac{\ln((\widetilde{\omega}(1-\eta) - \theta\eta(1+\eta))/(2\theta(1+\eta)^2))}{\ln((\theta-1)/(\theta+1))}, \ l = 1, 2, 3, \cdots,$$

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

3.3 Regularized NSCG

In the regularized case, by setting $\nu = \alpha + \beta$, where $\alpha > 0$ and $\beta > 0$, the Eq. (9) transforms to the following system

$$(H_A + \alpha I_n)X + X(H_B + \beta I_m) = (S_A + \alpha I_n)X^{(l)} + X^{(l)}(S_B + \beta I_m) + C. \quad (16)$$

Therefore, $X^{(l+1)}$ can be obtained by solving the matrix Eq. (16) with the Sylvester CG method, where $\alpha > 0$ and $\beta > 0$ are the regularizing parameters.

The Eq. (16) can be arranged in the Kronecker product form equivalently as

$$\mathcal{H}(\nu)x = \mathcal{S}(\nu)x^{(l)} + c, \quad (17)$$

where

$$\mathcal{H}(\nu) = \mathcal{H} + \nu\mathcal{I} = I_m \otimes (H_A + \alpha I_n) + (H_B + \beta I_m)^T \otimes I_n,$$

and

$$\mathcal{S}(\nu) = \mathcal{S} + \nu\mathcal{I} = I_m \otimes (S_A + \alpha I_n) + (S_B + \beta I_m)^T \otimes I_n.$$

Now, it is easy to show that when the splitting $\mathcal{A} = \mathcal{H} - \mathcal{S}$ is a contractive splitting (with respect to the $||\cdot||_{\mathcal{H}}$-norm) the splitting $\mathcal{A} = \mathcal{H}(\nu) - \mathcal{S}(\nu)$ is also a contractive

splitting (with respect to the $||\cdot||_{\mathcal{H}(\nu)}$-norm). In addition a theorem similar to Theorem 3.5 can be stated for this splitting.

The following discussion describes how to choose the best possible parameter ν in order to improve the convergence rate of the NSCG method. Since $\mathcal{H}(\nu)$ is Hermitian positive definite and $\mathcal{S}(\nu)$ is skew-Hermitian, one can easily obtain that

$$\rho(\mathcal{H}^{-1}(\nu)\mathcal{S}(\nu)) \leq ||\mathcal{H}^{-1}(\nu)\mathcal{S}(\nu)||_2 \leq ||\mathcal{H}^{-1}(\nu)||_2||\mathcal{S}(\nu)||_2 = \frac{\sqrt{\sigma_{\max}^2(\mathcal{S}) + \nu^2}}{\lambda_{\min}(\mathcal{H}) + \nu} = g(\nu),$$

where, $\sigma_{\max}(\mathcal{S})$ denotes the maximum singular value of the matrix $\mathcal{S}$. We easily see that $g(\nu)$ admits the minimum

$$g(\nu^*) = \frac{\sigma_{\max}(\mathcal{S})}{\sqrt{\lambda_{\min}^2(\mathcal{H}) + \sigma_{\max}^2(\mathcal{S})}}$$

at

$$\nu^* = \frac{\sigma_{\max}^2(\mathcal{S})}{\lambda_{\min}(\mathcal{H})}. \tag{18}$$

We emphasize that the optimal parameter ν^* minimizes only the upper bound $g(\nu)$ of the spectral radius of the iteration matrix but not does the spectral radius itself.

3.4 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 methods (see [7]), in Section 4, by numerical computation, we show that the quasi-Hermitian splitting

$$\mathcal{A} = \mathcal{H}(\nu) - \mathcal{S}(\nu), \tag{19}$$

with $\mathcal{H}(\nu)$ and $\mathcal{S}(\nu)$ defined in the Section 3.3, can be used as a splitting preconditioner and induce accurate, robust and effective preconditioned Krylov subspace iteration methods for solving the continuous Sylvester equation.

4 Numerical experiments

All numerical experiments presented in this section were computed in double precision with a number of MATLAB codes. Our test problems were chosen such that the restriction of Lemma 3.4 is satisfied for them. All iterations are started from the zero matrix for initial $X^{(0)}$ and terminated when the current iterate satisfies $\frac{\|R^{(k)}\|_F}{\|R^{(0)}\|_F} \leq 10^{-10}$, where $R^{(k)} = C - AX^{(k)} - X^{(k)}B$ is the residual of the kth iterate. Also we use the tolerance $\varepsilon = 0.01$ for inner iterations in NSCG method. 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 NSCG method with some iterative methods. The iterative methods which used in this section are presented in Table 1.

Dagger (†) shows that no convergence has been obtained after 5000 iterations. For regularized iterations and regularized preconditioned Krylov subspace methods, the optimal value of regularization parameter ν^* in Eq. (18) and $\alpha = \beta = \nu^*/2$ were used.

Table 1 Description the iterative methods which used

Method	Description
NSCG	Nested splitting conjugate gradient method described in Section 3
RNSCG	Regularized NSCG method described in Section 3
HSS	Hermitian and skew-Hermitian splitting method described in [3]
IHSS	Inexact HSS method described in [3]
BiCGSTAB	BiCGSTAB method for the Sylvester equation, see [15]
NSCG-BiCGSTAB	BiCGSTAB preconditioned by NSCG
RNSCG-BiCGSTAB	BiCGSTAB preconditioned by RNSCG
GMRES(m)	GMRES method for the Sylvester equation with $m = 10$, see [21, 26]
NSCG-FGMRES	FGMRES preconditioned by NSCG
RNSCG-FGMRES	FGMRES preconditioned by RNSCG

Example 1 For the first set of our examples, we use the matrices

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

where $M = \text{tridiag}(-1, 2, -1)$, $N = \text{tridiag}(0.5, 0, -0.5)$ and $r = 0.01$ [3]. We apply the iteration methods to this problem with different dimensions. The results are given in Table 2. The pair (n, m) in the first row of the table, represents the dimension of matrices A and B, respectively. For these problems, the optimal value of the regularization parameter obtained by Eq. (18) is equal to zero. Therefore, the RNSCG method reduces to the NSCG method.

From the results presented in Table 2, we observe that the NSCG method is more efficient than the other methods (except the NSCG-BiCGSTAB method for $(m, n) = (256, 256)$) in terms of CPU time. In addition, using the quasi-Hermitian splitting as a preconditioner for the BiCGSTAB and GMRES methods leads to achieve more efficient results in terms of the number of iterations and the CPU time.

Table 2 Results of the Example 1

Method	(8, 8)	(16, 16)	(32, 32)	(64, 64)	(128, 128)	(256, 256)
NSCG	5(0.002)	5(0.003)	6(0.006)	6(0.123)	8(1.010)	10(12.074)
HSS	19(0.010)	34(0.040)	63(0.276)	118(2.060)	220(16.407)	407(191.796)
IHSS	16(0.003)	30(0.009)	59(0.073)	104(0.489)	203(3.823)	379(43.251)
BiCGSTAB	10(0.002)	20(0.006)	42(0.031)	81(0.276)	157(2.812)	310(35.084)
NSCG-BiCGSTAB	6(0.002)	5(0.005)	10(0.026)	8(0.191)	4(1.029)	3(10.514)
GMRES(10)	15(0.004)	31(0.009)	70(0.046)	140(0.530)	339(3.512)	960(62.353)
NSCG-FGMRES	3(0.002)	3(0.004)	4(0.031)	9(0.319)	5(3.117)	9(59.511)

Example 2 For the second experiment, consider $A = \text{tridiag}(-2, 4, -1)$ and $B = \text{tridiag}(-1, 4, -2)$ with dimensions 2048×2048 and 128×128, respectively. This is a problem of strong Hermitian part [24, 28]. The numerical results for this problem are listed in Table 3.

Respect to the Table 3, it is obvious that the RNSCG method is more effective in terms of CPU time versus the other methods. Moreover, using the quasi-Hermitian splitting as a preconditioner for the BiCGSTAB method, improves the results obtained by the BiCGSTAB method. For the GMRES method, we observe that the use of NSCG or RNSCG methods as preconditioner decreases the number of iterations and residual Frobenius norm, but it increases the CPU time. Figure 1 represents the convergence history of the NSCG method versus the other iterative methods for this problem.

Example 3 Finally, for the last experiment, we used the nonsymmetric sparse matrix SHERMAN3 of dimension 5005×5005 with 20033 nonzero entries from the Harwell-Boeing collection [13] as the coefficient matrix A. For the coefficient matrix B, we used $B = \text{tridiag}(-1, 4, -2)$ of dimension 8×8. This is a problem of strong skew-Hermitian part [24, 28]. The CPU-time, residual Frobenus-norm and the number of iterations required for convergence are presented in Table 4.

For this test problem, Table 4 shows that the HSS, the IHSS, the BiCGSTAB, and the GMRES methods did not converge after 5000 iterations. We observed that the HSS and the IHSS methods converge very slowly, the BiCGSTAB does not converge, and the GMRES stagnates. In addition, Table 4 shows that the NSCG and RNSCG methods are more effective in terms of CPU time than the other methods, and the use of NSCG method or RNSCG method as a preconditioner improves the results

Table 3 Results of the Example 2

Method	Iteration	CPU time	Residual-norm
NSCG	13	3.385	4.74e-7
RNSCG	12	2.979	9.06e-7
HSS	26	22.292	1.00e-6
IHSS	26	17.846	9.99e-7
BiCGSTAB	19	7.559	6.59e-7
NSCG-BiCGSTAB	5	6.115	2.44e-7
RNSCG-BiCGSTAB	4	4.533	1.51e-7
GMRES(10)	30	3.752	8.65e-7
NSCG-FGMRES	9	18.931	8.65e-8
RNSCG-FGMRES	6	13.218	7.61e-9

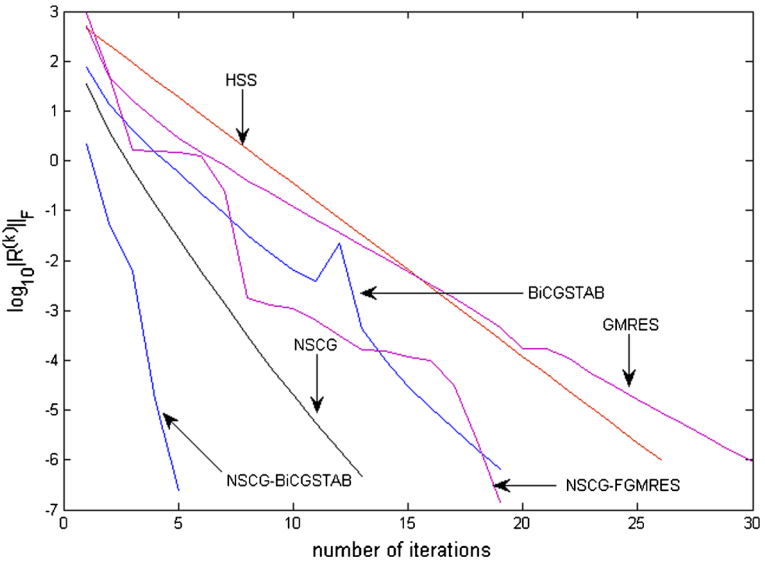

Fig. 1 Convergence history of NSCG versus the other iterative methods for Example 2

obtained by the corresponding methods. Finally, Fig. 2 represents the convergence history of the NSCG method versus the other methods.

In summary, by focusing on the results presented in Tables 1, 2, 3 and 4, one can observe that the NSCG method and its regularized variant are often superior to the other iterative methods. Moreover, the use of the quasi-Hermitian splitting as a preconditioner can induce accurate and effective preconditioned Krylov subspace methods.

Table 4 Results of the Example 3

Method	Iteration	CPU time	Residual-norm
NSCG	82	125.300	2.73e-6
RNSCG	78	119.371	3.64e-6
HSS	†	†	2.33
IHSS	†	†	1.89
BiCGSTAB	†	†	NaN
NSCG-BiCGSTAB	10	241.910	3.31e-6
RNSCG-BiCGSTAB	9	202.614	2.54e-7
GMRES(10)	†	†	5.03e+2
NSCG-FGMRES	9	289.147	5.08e-7
RNSCG-FGMRES	8	262.381	8.32e-8

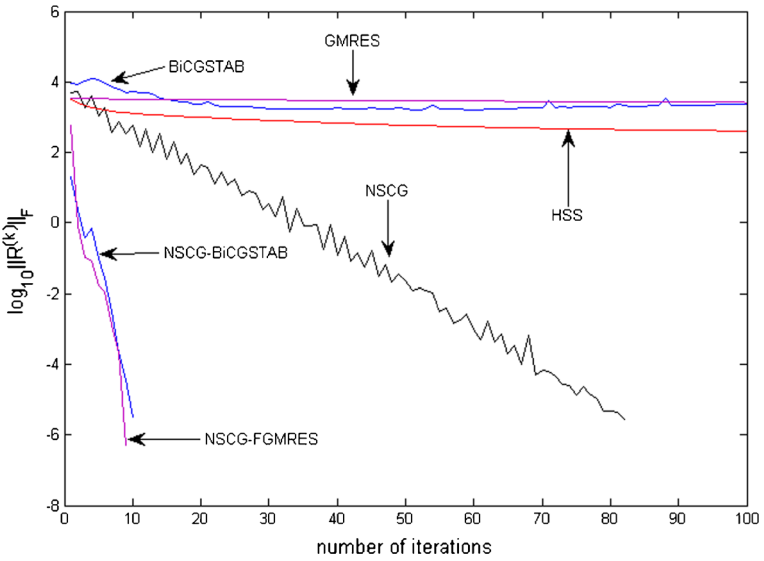

Fig. 2 Convergence history of NSCG versus the other iterative methods for Example 3

5 Conclusion

In this paper, we have proposed an efficient iterative method for solving the continuous Sylvester equation $AX+XB=C$. This method employs the Sylvester conjugate gradient 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 matrices A and B are positive semi-definite, and at least one of them is positive definite, we proved that the NSCG method converges to the exact solution of the Sylvester equation.

We have compared the NSCG method (and RNSCG method in regularized case), with well-known iterative methods such as the HSS and the IHSS methods, the BiCGSTAB method, and the GMRES method for several problems. We have observed that, for the problems of strong Hermitian parts, the NSCG and the RNSCG methods are superior to the other iterative methods in terms of CPU times. In addition, numerical computations showed that the quasi-Hermitian splitting can induce the accurate, robust and effective preconditioned Krylov subspace methods.

Acknowledgments The authors are very much indebted to the referees for their constructive comments and suggestions which greatly improved the original manuscript of this paper.

References

1. Axelsson, O., Bai, Z.-Z., Qiu, S.-X.: A class of nested iterative schemes for linear systems with a coefficient matrix with a dominant positive definite symmetric part. Numer. Algoritm. **35**, 351–372 (2004)

2. Bai, Z.-Z.: Splitting iteration methods for non-Hermitian positive definite systems of linear equations. Hokkaido Math. J. **36**, 801–814 (2007)
3. Bai, Z.-Z.: On Hermitian and skew-Hermitian splitting iteration methods for continuous Sylvester equations. J. Comput. Math. **29**(2), 185–198 (2011)
4. Bai, Z.-Z., Zhang, S.-L.: A regularized conjugate gradient method for symmetric positive definite system of linear equations. J. Comput. Math. **20**, 437–448 (2002)
5. Bai, Z.-Z., Golub, G.H., Ng, M.K.: Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. SIAM J. Matrix Anal. Appl. **24**, 603–626 (2003)
6. Bai, Z.-Z., Golub, G.H., Ng, M.K.: On inexact Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. Linear Algebra Appl. **428**, 413–440 (2008)
7. Bai, Z.-Z., Yin, J.-F., Su, Y.-F.: A shift-splitting preconditioner for non-Hermitian positive definite matrices. J. Comput. Math. **24**, 539–552 (2006)
8. Bartels, R.H., Stewart, G.W.: Algorithm 432: solution of the matrix equation $AX + XB = C$. Circ. Syst. Signal Proc. **13**, 820–826 (1994)
9. Benner, P., Li, R.-C., Truhar, N.: On the ADI method for Sylvester equations. J. Comput. Appl. Math. **233**, 1035–1045 (2009)
10. Bouhamidi, A., Jbilou, K.: Sylvester Tikhonov-regularization methods in image restoration. J. Comput. Appl. Math. **206**, 86–98 (2007)
11. Calvetti, D., Reichel, L.: Application of ADI iterative methods to the restoration of noisy images. SIAM J. Matrix Anal. Appl. **17**, 165–186 (1996)
12. Datta, B.: Numerical Methods for Linear Control Systems. Elsevier Academic Press, New York (2004)
13. Duff, I.S., Grimes, R.G., Lewis, J.G.: User's Guide for the Harwell-Boeing Sparse Matrix Collection, Technical Report RAL-92-086. Rutherford Applton Laboratory, Chilton (1992)
14. El Guennouni, A., Jbilou, K., Riquet, J.: Block Krylov subspace methods for solving large Sylvester equation. Numer. Algoritm. **29**, 75–96 (2002)
15. El Guennouni, A., Jbilou, K., Sadok, H.: A block version of BiCGSTAB for linear systems with multiple right-hand sides. Electron. Trans. Numer. Anal. **16**, 243–256 (2004)
16. Evans, D.J., Wan, C.R.: A preconditioned conjugate gradient method for $AX + XB = C$. Int. J. Comput. Math. **49**, 207–219 (1993)
17. Golub, G.H., Nash, S., Van Loan, C.: A Hessenberg-Schur method for the problem $AX + XB = C$. IEEE Trans. Control **AC-24**, 909–913 (1979)
18. Golub, G.H., Overton, M.L.: The convergence of inexact Chebyshev and Richardson iterative methods for solving linear systems. Numer. Math. **53**, 571–593 (1988)
19. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)
20. Hu, D.Y., Reichel, L.: Krylov-subspace methods for the Sylvester equation. Linear Algebra Appl. **172**, 283–313 (1992)
21. Jbilou, K., Messaoudi, A., Sadok, H.: Global FOM and GMRES algorithms for matrix equations. Appl. Numer. Math. **31**, 97–109 (1999)
22. Kelley, C.T.: Iterative Methods for Linear and Nonlinear Equations, no. 16. Frontiers in Applied Mathematics. SIAM, Philadelphia (1995)
23. Khorsand Zak, M., Toutounian, F.: Nested splitting conjugate gradient method for matrix equation $AXB = C$ and preconditioning. Comput. Math. Appl. **66**, 269–278 (2013)
24. Krukier, L.A., Martynova, T.S., Bai, Z.-Z.: Product-type skew-Hermitian triangular splitting iteration methods for strongly non-Hermitian positive definite linear systems. J. Comput. Appl. Math. **232**, 3–16 (2009)
25. Lütkepohl, H.: Handbook of Matrices. Wiley. England (1996)
26. Salkuyeh, D.K., Toutounian, F.: New approaches for solving large Sylvester equations. Appl. Math. Comput. **173**, 9–18 (2006)
27. Wang, C.-L., Bai, Z.-Z.: Sufficient conditions for the convergent splittings of non-Hermitian positive definite matrices. Linear Algebra Appl. **330**, 215–218 (2001)
28. Wang, L., Bai, Z.-Z.: Skew-Hermitian triangular splitting iteration methods for non-Hermitian positive definite linear systems of strong skew-Hermitian parts. BIT Numer. Math. **44**, 363–386 (2004)
29. Wang, L., Bai, Z.-Z.: Convergence conditions for splitting iteration methods for non-Hermitian linear systems. Linear Algebra Appl. **428**, 453–468 (2008)