Computers and Mathematics with Applications 68 (2014) 1409–1420

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

# A preconditioned nested splitting conjugate gradient iterative method for the large sparse generalized Sylvester equation

Yi-Fen Ke, Chang-Feng Ma *

*School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, PR China*

ARTICLE INFO

*Article history:*
Received 26 April 2014
Received in revised form 8 September 2014
Accepted 13 September 2014
Available online 11 October 2014

*Keywords:*
Generalized Sylvester equation
Conjugate gradient method
Preconditioning
Nested iteration method
Convergence conditions

ABSTRACT

A nested splitting conjugate gradient (NSCG) iterative method and a preconditioned NSCG (PNSCG) iterative method are presented for solving the generalized Sylvester equation with large sparse coefficient matrices, respectively. Both methods are actually inner/outer iterations, which employ the CG-like method as inner iteration to approximate each outer iteration, while each outer iteration is induced by a convergent and symmetric positive definite splitting of the coefficient matrices. Convergence conditions of both methods are studied in depth and numerical experiments demonstrate the efficiency of the proposed methods. Moreover, experimental results show that the PNSCG method is more accurate, robust and effective than the NSCG method.



## 1. Introduction

Matrix equations are one of the most interesting and intensively studied classes of mathematical problems and play vital roles in applications, such as image processing, control theory and model reduction; see [1–3] and their references. There are many results about various type of matrix equations, for example see [4–27].

In this paper, we consider the following generalized Sylvester equation

$$AXB + CXD = E, \tag{1.1}$$

where $A, C \in \mathbb{R}^{n\times n}$, $B, D \in \mathbb{R}^{m\times m}$ and $E \in \mathbb{R}^{n\times m}$ are given, large sparse and real matrices. Based on the Kronecker product, (1.1) can be written as

$$(B^\top \otimes A + D^\top \otimes C)\mathrm{vec}(X) = \mathrm{vec}(E), \tag{1.2}$$

where $\otimes$ denotes the Kronecker product, i.e., $A \otimes B = (a_{ij}B)$ and

$$\mathrm{vec}(X) = (x_{11}, x_{21}, \ldots, x_{n1}, x_{12}, x_{22}, \ldots, x_{n2}, \ldots, x_{nm})^\top \in \mathbb{R}^{nm}.$$

Here, the superscript '$\top$' denotes the transpose of a vector or a matrix. As is well known, the matrix equation (1.1) has a unique solution if and only if $B^\top \otimes A$ and $-D^\top \otimes C$ possess no common eigenvalues and the unique solution can be obtained by means of the inversion of the $nm \times nm$ matrix $\mathcal{A} = B^\top \otimes A + D^\top \otimes C$.

Many results have been obtained about the matrix equation (1.1). For example, Ding, Liu and Ding [12] presented a gradient based and a least-squares based iterative algorithms for solving the generalized Sylvester equation by using the

* Corresponding author. Tel.: +86 13763827962.
*E-mail addresses:* macf88@163.com, macf@fjnu.edu.cn (C.-F. Ma).

http://dx.doi.org/10.1016/j.camwa.2014.09.009
0898-1221/

hierarchical identification principle. Zhang and Cai [13] constructed a parameter iterative method for solving large generalized Sylvester equation. Bao, Lin and Wei [14] proposed Galerkin and minimal residual methods for iteratively solving generalized Sylvester equation of the form $AXB - X = C$, which used the Krylov subspace methods. In [19], Bai gave a Hermitian and skew-Hermitian splitting (HSS) iteration method for solving large sparse continuous Sylvester equations with non-Hermitian and positive definite/semi-definite matrices. A software package had been developed by Gardiner et al. [15] to solve efficiently the Sylvester-type matrix equation $AXB^\top + CXD^\top = E$. And a transformation method is used, which employed the QZ algorithm to structure the equation in such a way that it can be solved columnwise by a back substitution technique. This method is an extension of the Bartels–Stewart method and the Hessenberg–Schur method.

Our purpose in the present paper is twofold. Firstly, we present the nested splitting conjugate gradient (NSCG) iterative method for solving the generalized Sylvester equation (1.1) by using the symmetric and skew-symmetric splitting of the matrices $A$, $B$, $C$ and $D$. This method was first proposed in [16] for solving system of linear equations $Ax = b$, then was extended to solve the matrix equation $AXB = C$ and $AX + XB = C$, respectively, see [17,18]. For more detail on this method, we refer to [28,29] and references therein. Moreover, based on the NSCG method, we propose a preconditioned nested splitting conjugate gradient (PNSCG) iterative method.

The organization of this paper is as follows. In Section 2, we present some necessary notations and useful lemmas as well as a brief description of the NSCG method. In Section 3, we give the NSCG method for the generalized Sylvester equation $AXB + CXD = E$ and analyze its convergence properties. In Section 4, we propose a preconditioned NSCG method and give its convergence conditions. The numerical results about the proposed methods are shown and discussed in Section 5. Finally, in Section 6, we end the paper by some concluding remarks.

## 2. Preliminaries

In this section, we recall some necessary notations and useful results, which will be used in the following section.

A matrix $A \in \mathbb{R}^{n\times n}$ is said to be symmetric if $A^\top = A$, and symmetric positive definite if it is symmetric and satisfies $x^\top Ax > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. Let $A \in \mathbb{R}^{n\times n}$ be a given matrix and $B, C \in \mathbb{R}^{n\times n}$ satisfy $A = B - C$. Then $A = B - C$ is called a splitting of the matrix $A$ if $B$ is nonsingular. The splitting $A = B - C$ is called a convergent splitting if $\rho(B^{-1}C) < 1$; a symmetric splitting if $B$ is a symmetric matrix; a symmetric positive definite splitting if $B$ is a symmetric positive definite matrix; and a contractive splitting if $\|B^{-1}C\| < 1$ for some matrix norm.

In the remainder of this paper, we use the following notations. For $X = (x_{ij}), Y = (y_{ij}) \in \mathbb{R}^{n\times m}$, we define the following inner product $\langle X, Y\rangle = \mathrm{tr}(X^\top Y) = \sum_{i,j} x_{ij}y_{ij}$, where $\mathrm{tr}(\cdot)$ denotes the trace. We use $\lambda(A)$, $\lambda_{\max}(A)$, $\lambda_{\min}(A)$, $\|A\|_2$, $\|A\|_F$ and $I_n$ to denote the eigenvalue, the maximum eigenvalue, the minimum eigenvalue, the spectral norm, the Frobenius norm of a matrix $A \in \mathbb{R}^{n\times n}$ 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 $\mathcal{H}$, we denote by $\kappa(H) = \|\mathcal{H}\|_2\|\mathcal{H}^{-1}\|_2$ its Euclidean condition number. And for a symmetric positive definite matrix $\mathcal{H}$, we define the $\|\cdot\|_{\mathcal{H}}$ norm of a vector $x \in \mathbb{R}^n$ as $\|x\|_{\mathcal{H}} = \sqrt{x^\top \mathcal{H}x}$. Then the induced $\|\cdot\|_{\mathcal{H}}$ norm of a matrix $A \in \mathbb{R}^{n\times n}$ is defined as $\|A\|_{\mathcal{H}} = \|\mathcal{H}^{\frac{1}{2}}A\mathcal{H}^{-\frac{1}{2}}\|_2$. In addition, it holds that $\|Ax\|_{\mathcal{H}} \leqslant \|A\|_{\mathcal{H}}\|x\|_{\mathcal{H}}$, $\|A\|_{\mathcal{H}} \leqslant \sqrt{\kappa(\mathcal{H})}\|A\|_2$ and $\|I\|_{\mathcal{H}} = 1$. Furthermore, for a matrix $X \in \mathbb{R}^{n\times m}$ and the vector $\mathrm{vec}(X) \in \mathbb{R}^{nm}$, we have the following relationship, that is, $\|X\|_F = \|\mathrm{vec}(X)\|_2$.

**Lemma 2.1** (*[30]*). *Let $A, B \in \mathbb{R}^{n\times n}$ be two symmetric matrices. Then*

$$\lambda_{\max}(A + B) \leqslant \lambda_{\max}(A) + \lambda_{\max}(B),$$
$$\lambda_{\min}(A + B) \geqslant \lambda_{\min}(A) + \lambda_{\min}(B).$$

**Lemma 2.2** (*[31]*). *Let $A, B \in \mathbb{R}^{n\times n}$, $\lambda$ and $\mu$ be the eigenvalues of $A$ and $B$, $x$ 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 2.3** (*[32]*). *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. Then for all $x \in \mathbb{R}^n$, we have $\|A^{\frac{1}{2}}x\|_2 = \|x\|_A$ and*

$$\sqrt{\lambda_{\min}(A)}\|x\|_A \leqslant \|Ax\|_2 \leqslant \sqrt{\lambda_{\max}(A)}\|x\|_A.$$

In [16], Axelsson, Bai and Qiu proposed an efficient iteration method for solving the system of linear equations

$$Ax = b, \tag{2.1}$$

where $A \in \mathbb{R}^{n\times n}$ is a large sparse nonsingular matrix, $x, b \in \mathbb{R}^n$. Let $A = B - C$ be a symmetric positive definite splitting of the coefficient matrix $A$ and assume that the splitting satisfies the condition $\rho(B^{-1}C) < 1$. Then the system of linear equations (2.1) is equivalent to the fixed-point equation

$$Bx = Cx + b.$$

Given a starting vector $x^{(0)} \in \mathbb{R}^n$, suppose that we have computed approximations $x^{(1)}, x^{(2)}, \ldots, x^{(l-1)}$ to the solution $x^* \in \mathbb{R}^n$ of the system (2.1). Then the next approximation $x^{(l)}$ may be defined as either an exact or an inexact solution of the system