Journal of Computational Mathematics
Vol.35, No.5, 2017, 620–641.
http://www.global-sci.org/jcm
doi:10.4208/jcm.1608-m2015-0430

THE ALTERNATING DIRECTION METHODS FOR SOLVING THE SYLVESTER-TYPE MATRIX EQUATION $AXB + CX^\top D = E$*

Yifen Ke and Changfeng Ma
School of Mathematics and Computer Science & FJKLMAA, Fujian Normal University, Fuzhou 350117, China
Email: keyifen2017@163.com, macf@fjnu.edu.cn

Abstract

In this paper, we present two alternating direction methods for the solution and best approximate solution of the Sylvester-type matrix equation $AXB + CX^\top D = E$ arising in the control theory, where A, B, C, D and E are given matrices of suitable sizes. If the matrix equation is consistent (inconsistent), then the solution (the least squares solution) can be obtained. Preliminary convergence properties of the proposed algorithms are presented. Numerical experiments show that the proposed algorithms tend to deliver higher quality solutions with less iteration steps and CPU time than some existing algorithms on the tested problems.

Mathematics subject classification: 65F10, 15A24.
Key words: Sylvester-type matrix equation, Alternating direction method, The least squares solution, Best approximate solution.

1. Introduction

Consider the following Sylvester-type matrix equation and its best approximate problem:

Problem 1. *Given $A \in \mathbb{R}^{s\times m}$, $B \in \mathbb{R}^{n\times t}$, $C \in \mathbb{R}^{s\times n}$, $D \in \mathbb{R}^{m\times t}$ and $E \in \mathbb{R}^{s\times t}$, find a matrix $X \in \mathbb{R}^{m\times n}$ such that*

$$AXB + CX^\top D = E. \tag{1.1}$$

Problem 2. *Let Problem 1 be consistent and its solution set be denoted by S_r. For given $X_f \in \mathbb{R}^{m\times n}$ and $X_f \notin S_r$, find a matrix $X^* \in S_r$ such that*

$$\frac{1}{2}\|X^* - X_f\|_F^2 = \min_{X\in S_r} \frac{1}{2}\|X - X_f\|_F^2.$$

According to Theorem 4.3.8 and Corollary 4.3.10 in [1], there exists a permutation matrix $P(m, n)$ such that the matrix equation (1.1) is equivalent to

$$H\mathrm{vec}(X) = \mathrm{vec}(E), \tag{1.2}$$

where $H = B^\top \otimes A + (D^\top \otimes C)P(m, n)$, the symbol $\otimes$ denotes the Kronecker product, i.e., $A \otimes B = (a_{ij}B)$ and

$$\mathrm{vec}(X) = (x_{11}, x_{21}, \dots, x_{m1}, x_{12}, x_{22}, \dots, x_{m2}, \dots, x_{mn})^\top \in \mathbb{R}^{mn}.$$

* Received October 21, 2015 / Revised version received May 14, 2016 / Accepted August 30, 2016 / Published online July 1, 2017 /

The matrix $P(m,n)$ has the following explicit form

$$P(m,n) = \begin{pmatrix} E_{11}^\top & E_{12}^\top & \cdots & E_{1n}^\top \\ E_{21}^\top & E_{22}^\top & \cdots & E_{2n}^\top \\ \vdots & \vdots & \ddots & \vdots \\ E_{m1}^\top & E_{m2}^\top & \cdots & E_{mn}^\top \end{pmatrix} \in \mathbb{R}^{mn\times mn},$$

where $E_{ij}(i=1,\ldots,m; j=1,\ldots,n)$ is an $m\times n$ matrix with the element at position (i,j) being 1 and the others being 0.

The system of linear equations (1.2) has a solution $\operatorname{vec}(X)\in\mathbb{R}^{mn}$ if and only if $HH^\dagger\operatorname{vec}(E)=\operatorname{vec}(E)$. In this case, the general solution can be described as $\operatorname{vec}(X)=H^\dagger\operatorname{vec}(E)+(I_{mn}-H^\dagger H)\operatorname{vec}(Y)$, where $Y\in\mathbb{R}^{m\times n}$ is an arbitrary matrix and $H^\dagger$ is the Moore-Penrose generalized inverse of the coefficient matrix H [2,3].

There are many algorithms for solving the system of linear equations (1.2), e.g., the Jacobi and Gauss-Seidel methods [4], the GMRES method [5] and the NSCG method [6]. For more details on the system of linear equations $Ax=b$, we refer to [4–13] and references therein. However, they are usual costly and impractical while the coefficient matrix H is large and sparse.

The matrix equation (1.1) plays an important role in system theory and control theory [14,15], for example, eigenstructure assignment [16], observer design [17], system control with constraint input [18] and fault detection [19].

Many results have been obtained about the matrix equation (1.1). In [20], by applying the conjugate gradient method, Wang et al. proposed two iterative algorithms to solve the matrix equation (1.1). In [21], Wang gave a new iterative algorithm for the matrix equation (1.1), which is based on the LSQR algorithm [22]. By using Kronecker product and the vectorization operator, Hajarian [23] developed the bi-conjugate gradients (Bi-CG) and bi-conjugate residual (Bi-CR) methods for the solution of the generalized Sylvester-transpose matrix equation

$$\sum_{i=1}^{p}(A_iXB_i+C_iX^\top D_i)=E.$$

In [24], by extending the idea of the conjugate gradient method, Song et al. constructed an iterative algorithm for the coupled Sylvester transpose matrix equations

$$\sum_{\eta=1}^{p}A_{i\eta}X_\eta B_{i\eta}+C_{i\eta}X_\eta^\top D_{i\eta}=E_i,\qquad i=1,\ldots,p.$$

For more details on the matrix equation (1.1), we refer to [3,25–32,52–60] and references therein.

Our purpose in the present paper is twofold. Firstly, we present the equivalent forms of Problems 1 and 2. Moreover, we consider to apply the classical alternating direction method of multipliers (ADMM) [33,34] for solving the equivalent linear constrained optimization problems.

Throughout this paper, we adopt the standard notation in matrix theory. The symbol I_n stands for the identity matrix of order n. Let $A^\top$ and $\|A\|_F$ be the transpose and the Frobenius norm of a real matrix $A\in\mathbb{R}^{m\times n}$, respectively. The inner product in space $\mathbb{R}^{m\times n}$ is defined as

$$\langle A,B\rangle := \operatorname{tr}(A^\top B)=\sum_{i,j}a_{ij}b_{ij},\qquad \forall\ A,B\in\mathbb{R}^{m\times n}.$$