Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Applied Mathematics and Computation 187 (2007) 1434–1441
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

# A new iteration method for the matrix equation $AX = B$ ☆

Guo Konghua *, Xiyan Hu, Lei Zhang

*College of Mathematics and Econometrics, Hunan University, Changsha 410082, PR China*

**Abstract**

An iteration method for the matrix equation $AX = B$ is constructed. By this iteration method, the least-norm solution for the matrix equation can be obtained when the matrix equation is consistent and the least-norm least-squares solutions can be obtained when the matrix equation is not consistent. The related optimal approximation solution is also obtained by this iteration method. A preconditioned method for improving the iteration rate is put forward. Finally, some numerical examples are given.


*Keywords:* Matrix equation; Iteration method; Optimal approximation

## 1. Introduction

Matrix equation problem is one of the topics of very active research in the computational mathematics, and has been widely applied in various areas, such as structural design, system identification, principal component analysis, exploration and remote sensing, biology, electricity, solid mechanics, molecular spectroscopy, structural dynamics, automatics control theory, vibration theory, and so on.

We use $R^n$ to denote the set of all real vectors of $n$ dimensions, $I_n$ the identity matrix of order $n$, and $R^{n\times m}$ all $n\times m$ real matrices. Let $\|A\|_{\mathrm{F}}$, $A^+$, $A^{\mathrm{T}}$ denote especially the Frobenius norm, the Moore–Penrose generalized inverse, and the transpose of a matrix $A$. $\mathrm{tr}(A)$ means the trace of matrix $A$, $R(A)$ the column space of matrix $A$, $R^{\perp}(A)$ the orthogonal complement space of $R(A)$, and for any $A\in R^{m\times n}$, $B\in R^{n\times p}$, $A\otimes B$ means the Kronecker product of the matrices $A$ and $B$.

The following problems are considered in this paper.

**Problem 1.1.** Given $A, B\in R^{m\times n}$, find $X\in R^{n\times n}$, such that

$$AX = B. \tag{1}$$

☆ Research supported by National Natural Science Foundation of China (10571047).
* Corresponding author.
*E-mail address:* khguo@126.com (G. Konghua).

0096-3003/$ - see front matter 
doi:10.1016/j.amc.2006.09.059

**Problem 1.2.** Let Problem 1.1 is consistent, and its solution set is $S_E$, for $X_0 \in R^{n\times n}$, find $\widehat{X} \in S_E$ such that

$$\|\widehat{X} - X_0\|_{\mathrm{F}} = \min_{X\in S_E} \|X - X_0\|_{\mathrm{F}}. \tag{2}$$

In fact, Problem 1.2 is to find the optimal approximation solution to the given matrix $X_0$ in the solution set of Problem 1.1. Since Bjerhammar considered Problem 1.1 and obtained the general expression of its solution by using matrix generalized inverse in 1951 [1], Problems 1.1 and 1.2 have been studied and have been solved for different classes of structured matrices by many authors, and for more details we can refer to literatures [2–9]. In these literatures, the problem was solved by using matrices decomposition such as the singular value decomposition (SVD), the generalized SVD (GSVD), the quotient SVD (QSVD) and the canonical correlation decomposition (CCD). However, it is difficult to apply these methods for solving problems of matrix equations in subspace such as solving symmetric solution of the matrix equation $AXB = C$. In 2005, Peng put forward an iteration method for symmetric solution of the matrix equation $AXB = C$ [10]. The advantage of the iteration method is that when the problem is consistent, its solution can be obtained theoretically within finite iteration, and the disadvantage of the method is that the convergence rate cannot be analyzed, so the iteration method cannot be improved.

In this paper, we construct a new iterative method for the matrix equation $AX = B$, by which we can obtain the least-norm solution of Problem 1.1 when the problem is consistent and obtain the least-norm least-squares solution of Problem 1.1 when the problem is not consistent, and furthermore, we show that the convergence rate of the method is related to the singular value of the matrix $A$. In the case the solution set of problem is not empty, Problem 1.2 has a unique solution and we can obtain it by the iteration method.

The paper is organized as follows: In Section 2 we first introduce a new iterative method for the matrix equation $AX = B$ and prove the convergence of the method. In Section 3 we solve Problem 1.2 by using this iteration method. In Section 4 we put forward a kind of improvement of the iteration method in order to increase the convergence rate. In the last section, we will give some numerical examples to verify the method and compare the convergence rate between the original method and the improved method.

## 2. The iteration method for solving Problem 1.1

In this section, we will introduce a new iteration method for solving Problem 1.1, and then we will prove the convergence of the iteration method.

**Iteration method 2.1**

Step 1: Select $B_0 = B$, $X_0 = O$.
Step 2: Let $\alpha_k = \frac{\|A^{\mathrm{T}}B_k\|_{\mathrm{F}}^2}{\|AA^{\mathrm{T}}B_k\|_{\mathrm{F}}^2}$ $(k = 0, 1, 2, \ldots)$.
Step 3: Let $\Delta X_k = \alpha_k A^{\mathrm{T}} B_k$ $(k = 0, 1, 2, \ldots)$.
Step 4: If $\Delta X_k = 0$, stop, otherwise, let $X_{k+1} = X_k + \Delta X_k$ $(k = 0, 1, 2, \ldots)$.
Step 5: Let $B_{k+1} = B_k - A\Delta X_k = B_0 - AX_{k+1}$ $(k = 0, 1, 2, \ldots)$, goto step 2.

**Definition 2.1.** Suppose $A, B \in R^{m\times n}$, then $\mathrm{tr}(A^{\mathrm{T}}B)$ is called the inner product of the matrices $A$, $B$, denoted by $\langle A, B\rangle$.

**Definition 2.2.** Assume $A, B \in R^{m\times n}$, if $\langle A, B\rangle = 0$ i.e. $\mathrm{tr}(A^{\mathrm{T}}B) = 0$, then matrices $A$, $B$ are called orthogonal each other.

**Lemma 2.1.** *In Iteration method 2.1, the selection of $\alpha_k$ will make $\|B_{k+1}\|_F$ the minimal and will make $B_{k+1}$ and $A\Delta X_k$ orthogonal.*

**Proof.** From Iteration method 2.1, we have that

$$\|B_{k+1}\|_{\mathrm{F}}^2 = \langle B_k - \alpha_k AA^{\mathrm{T}}B_k, B_k - \alpha_k AA^{\mathrm{T}}B_k\rangle = \|B_k\|_{\mathrm{F}}^2 - 2\alpha_k\langle B_k, AA^{\mathrm{T}}B_k\rangle + \alpha_k^2\|AA^{\mathrm{T}}B_k\|_{\mathrm{F}}^2.$$