Higher-Order Bi-CGSTAB and Bi-CRSTAB Algorithms To Solve Some Tensor Equations

Conference Paper · May 2021

Eisa Khosravi Dehdezi
Persian Gulf University

Higher-Order Bi-CGSTAB and Bi-CRSTAB Algorithms To Solve Some Tensor Equations

Eisa Khosravi Dehdezi*
Department of Mathematics, Persian Gulf University, Bushehr, Iran
and Saeed Karimi
Department of Mathematics, Persian Gulf University, Bushehr, Iran

Abstract. This paper investigates the tensor form of the Bi-CGSTAB and Bi-CRSTAB methods, by employing Kronecker product and vectorization, to solve the generalized coupled Sylvester tensor equations with no matricization. Some numerical examples are provided to compare the efficiency of the proposed methods.
Keywords: Tensor equations, HOBi-CGSTAB, HOBi-CRSTAB, Iterative methods, k-mode product.
AMS Mathematical Subject Classification [2010]: 15A69, 65F10, 65W05.

1. Introduction

In this paper, we are concerned with the generalized coupled Sylvester tensor equations

$$\sum_{j=1}^{n} \mathcal{X}_j \times_1 A_{ij1} \times_2 A_{ij2} \times \cdots \times_d A_{ijd} = \mathcal{E}_i, \qquad i = 1, 2, \ldots, n, \tag{1}$$

where the matrices $A_{ijl} \in \mathbb{C}^{n_{ijl} \times n_{ijl}}$ $(i, j = 1, 2, \ldots, n$ and $l = 1, 2, \ldots, d)$, tensors $\mathcal{E}_i \in \mathbb{C}^{n_{i1} \times \cdots \times n_{id}}(i = 1, 2, \ldots, n)$ are known and $\mathcal{X}_j \in \mathbb{C}^{n_{j1} \times \cdots \times n_{jd}}(j = 1, 2, \ldots, n)$ are unknown tensors and the j-mode product $\times_j$ will be defined later. The (coupled) Sylvester tensor equations often arise from the finite element, finite difference or spectral methods [3]. In [1], Khosravi Dehdezi and Karimi proposed the extended conjugate gradient squared and conjugate residual squared methods for solving (1). In this paper, tensors are written as calligraphic capital letters such as $\mathcal{A}, \mathcal{B}, \ldots$. Let N be a positive integer, an order N real tensor $\mathcal{A} \in \mathbb{R}^{\times I_1 \times \cdots \times I_N}$ is the following multidimensional array

$$\mathcal{A} = (a_{i_1 i_2 \cdots N})(1 \leq i_j \leq n_j,\ j = 1, 2, \ldots, N),\ a_{i_1 i_2 \cdots N} \in \mathbb{R},$$

with $H(H = n_1 n_2 \ldots n_N)$ entries [2]. Each entry of $\mathcal{A}$ is denoted by $a_{i_1 i_2 \cdots N}$. $\mathcal{O}$ with all entries zero denote the zero tensor. With this definition of tensor, matrices are tensors of order two, where signified by capital letters, e.g., A. As usual, $\mathbb{C}$ denotes the complex number field.

Definition 1.1. [2] The operators $\times_k (k = 1, 2, \ldots, n)$ represent the k-mode product of a tensor $\mathcal{X}$ with a matrix $A \in \mathbb{C}^{m \times n_k}$ defined as follows

*Speaker

$$(\mathcal{X} \times_k A)_{i_1 i_2 \cdots_{k-1} j i_{k+1} \cdots_d} = \sum_{i_k=1}^{n_k} x_{i_1 i_2 \cdots_{k-1} i_k i_{k+1} \cdots_d} a_{j i_k}.$$

DEFINITION 1.2. [2] Let N, M be positive integers. The inner product of two tensors $\mathcal{X}, \mathcal{Y} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$ is defined by

$$< \mathcal{X}, \mathcal{Y} >= \sum_{j_M=1}^{J_M} \cdots \sum_{j_1=1}^{J_1} \sum_{i_N=1}^{I_N} \cdots \sum_{i_1=1}^{I_1} x_{i_1 \ldots i_N j_1 \ldots j_M} \bar{y}_{j_1 \ldots j_M i_1 \ldots i_N},$$

so the tensor norm that generated by this inner product is

$$\|\mathcal{X}\| = \sqrt{< \mathcal{X}, \mathcal{X} >} = \sqrt{\sum_{j_M=1}^{J_M} \cdots \sum_{j_1=1}^{J_1} \sum_{i_N=1}^{I_N} \cdots \sum_{i_1=1}^{I_1} |x_{i_1 \ldots i_N j_1 \ldots j_M}|^2},$$

which is the tensor Frobenius norm. We say that $\mathcal{X}, \mathcal{Y}$ are orthogonal if $< \mathcal{X}, \mathcal{Y} >= 0$.

We define a new inner product which is needed in the following.

DEFINITION 1.3. Let $H_j, j = 1, 2, \ldots, n$ be the linear space $\mathbb{C}^{n_{j1} \times \cdots \times n_{jd}}, j = 1, 2, \ldots, n$. Define

$$\begin{cases} \mathcal{L} : H_1 \times H_2 \times \cdots \times H_n \to H_1 \times H_2 \times \cdots \times H_n \\ \mathcal{L}(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n) = \begin{pmatrix} \mathcal{L}_1(\mathcal{X}_j) \\ \mathcal{L}_2(\mathcal{X}_j) \\ \cdots \\ \mathcal{L}_n(\mathcal{X}_j) \end{pmatrix}, \end{cases}$$

where

$$\mathcal{L}_i(\mathcal{X}_j) = \sum_{j=1}^{n} \mathcal{X}_j \times_1 A_{ij1} \times_2 A_{ij2} \times \cdots \times_d A_{ijd}, \; i = 1, 2, \ldots, n.$$

According to this definition, the linear system (1) can be rewritten as

$$\mathcal{L}(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n) = \mathcal{E}, \; \mathcal{E} = \left(\mathcal{E}_1{}^T, \mathcal{E}_2{}^T, \ldots, \mathcal{E}_n{}^T\right)^T. \tag{2}$$

The remainder of this paper is organized as follows. In Section 2,the higher-order Bi-CGSTAB and Bi-CRSTAB methods are obtained according to tensor form for solving the tensor equations (1). Finally in Section 3, we show comparative results.

2. Higher Order Bi-CGSTAB and Bi-CRSTAB Methods to Solve (1)

Two of the important iterative methods for solving large sparse non-Hermitian linear systems of equations

$$Ax = b, \qquad\qquad A \in \mathbb{R}^{n \times n}, x \in \mathbb{R}^n.$$

are the bi-conjugate gradient (Bi-CG) and bi-conjugate residual (Bi-CR) methods based on the non-symmetric Lanczos procedure. Van der Vorst in [4] introduced one of the most successful improvements, a fast and smoothly convergent variant of Bi-CG that avoids calculating the matrix A^*, known as the Bi-CGSTAB algorithm. In exact arithmetic, the Bi-CGSTAB algorithm terminates with a true solution after $j \leq n$ steps [4]. In the following, we present the higher order Bi-CGSTAB

(HOBi-CGSTAB) and higher order Bi-CRSTAB (HOBi-CRSTAB) algorithms to solve the generalized coupled Sylvester tensor Eq. (1). The mode-k matricization of a tensor $\mathcal{X} \in \mathbb{C}^{n_1 \times n_2 \times \cdots \times n_M}$ is denoted by $\mathcal{X}(k)$ and the mode-k fibres are arranged to be the columns of the resulting matrix. The operator "vec" denotes the columns of a matrix or tensor to form a vector. For a matrix $A = (a_1, a_2, \ldots, a_n) = (a_{ij}) \in \mathbb{C}^{m \times n}$ and a matrix B, $A \otimes B = (a_{ij}B)$ is a Kronecker product and $vec(A)$ is a vector defined by $vec(A) = (a_1^T, a_2^T, ..., a_n^T)^T$, where $a_i, 1 \leq i \leq n$ is the i-th column of A and for tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_M}$, $\mathcal{X}(1)$ is the mode-1 matricization of the tensor $\mathcal{X}$ [**2**]. By using the property of the Kronecker product it can be shown that tensor Eq. (1) are equivalent to the following equations

$$\sum_{j=1}^{n} G_{ij} vec(\mathcal{X}_j) = vec(\mathcal{E}_i),\ G_{ij} = A_{ijd} \otimes \cdots \otimes A_{ij2} \otimes A_{ij1},\ i, j = 1, 2, \ldots, n,$$

and $\otimes$ stands Kronecker product.

Thus, the general coupled Sylvester tensor Eq. (1) can be transformed into the following linear system

$$\underbrace{\begin{pmatrix} G_{11} & G_{12} & \ldots & G_{1n} \\ G_{21} & G_{22} & \ldots & G_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ G_{n1} & G_{n2} & \ldots & G_{nn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} vec(\mathcal{X}_1) \\ vec(\mathcal{X}_2) \\ \ldots \\ vec(\mathcal{X}_n) \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} vec(\mathcal{E}_1) \\ vec(\mathcal{E}_2) \\ \ldots \\ vec(\mathcal{E}_n) \end{pmatrix}}_{b}.$$

It is obvious that the size of this linear system is very large even for small values of N and thus using of the Bi-CGSTAB and Bi-CRSTAB algorithms to solve the linear system $Ax = b$ instead of corresponding tensor Eq. (1) will consume much more computer time and memory space as the dimension increases. To overcome this problem, we propose the HOBi-CGSTAB and HOBi-CRSTAB algorithms for solving the tensor Eq. (1). For this purpose, we first provide the common Bi-CGSTAB and Bi-CRSTAB algorithms for solving $Ax = b$ as follows. Let $\mathcal{X}_{i,k}$ and $\mathcal{P}_{i,k} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, $i, j = 1, \ldots, n,\ k = 0, 1, 2, \ldots,$ be the k-th equation solution tensor and k-th search direction tensor, respectively. By taking

$$\mathcal{R}_{i,k} = \mathcal{E}_i - \mathcal{L}_i(\mathcal{X}_{i,k}),\ \mathcal{Q}_{i,k} = \mathcal{L}_i(\mathcal{P}_{j,k}),\ \mathcal{S}_{i,k} = \mathcal{R}_{i,k} - \alpha_k \mathcal{Q}_{i,k},$$

$$\mathcal{W}_{i,k} = \mathcal{L}_i(\mathcal{S}_{j,k}),\ \mathcal{R}_{i,k+1} = \mathcal{S}_{i,k} - \omega_k \mathcal{W}_{i,k},$$

in step k and using variables of Bi-CGSTAB and Bi-CRSTAB algorithms and choosing $\mathcal{R}^*_{i,0} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, $i = 1, 2, \ldots, n$, arbitrary, the following vectors can be rearranged with the corresponding tensors. Set

$$b = \begin{pmatrix} vec(\mathcal{E}_1) \\ vec(\mathcal{E}_2) \\ \ldots \\ vec(\mathcal{E}_n) \end{pmatrix}, x_k = \begin{pmatrix} vec(\mathcal{X}_{1,k}) \\ vec(\mathcal{X}_{2,k}) \\ \ldots \\ vec(\mathcal{X}_{n,k}) \end{pmatrix}, \tag{3}$$

$$p_k = \begin{pmatrix} vec(\mathcal{P}_{1,k}) \\ vec(\mathcal{P}_{2,k}) \\ \ldots \\ vec(\mathcal{P}_{n,k}) \end{pmatrix}, w_k = \begin{pmatrix} vec(\mathcal{W}_{1,k}) \\ vec(\mathcal{W}_{2,k}) \\ \ldots \\ vec(\mathcal{W}_{n,k}) \end{pmatrix},$$

(4) $$q_k = \begin{pmatrix} vec(\mathcal{Q}_{1,k}) \\ vec(\mathcal{Q}_{2,k}) \\ \dots \\ vec(\mathcal{Q}_{n,k}) \end{pmatrix}, r_k = \begin{pmatrix} vec(\mathcal{R}_{1,k}) \\ vec(\mathcal{R}_{2,k}) \\ \dots \\ vec(\mathcal{R}_{n,k}) \end{pmatrix}, r_0^* = \begin{pmatrix} vec(\mathcal{R}_{1,0}^*) \\ vec(\mathcal{R}_{2,0}^*) \\ \dots \\ vec(\mathcal{R}_{n,0}^*) \end{pmatrix},$$

where $\mathcal{C}_i, \mathcal{X}_{i,k}, \mathcal{P}_{i,k}, \mathcal{Q}_{i,k}, \mathcal{V}_{i,k}, \mathcal{W}_{i,k}, \mathcal{R}_{i,k}, \mathcal{R}_{i,0}^* \in \mathbb{C}^{n_{i1}\times\cdots\times n_{id}}$ for $i = 1, 2, \dots, n$ and $k = 0, 1, 2, \dots$. By using Eqs. (3) and (4), we have the following relation

$$\begin{aligned} <r_{k+1}, r_0^*> &= \left\langle \begin{pmatrix} vec(\mathcal{R}_{1,k+1}) \\ vec(\mathcal{R}_{2,k+1}) \\ \dots \\ vec(\mathcal{R}_{n,k+1}) \end{pmatrix}, \begin{pmatrix} vec(\mathcal{R}_{1,0}^*) \\ vec(\mathcal{R}_{2,0}^*) \\ \dots \\ vec(\mathcal{R}_{n,0}^*) \end{pmatrix} \right\rangle \\ &= \sum_{i=1}^{n} < vec(\mathcal{R}_{i,k+1}), vec(\mathcal{R}_{i,0}^*) > \\ &= \sum_{i=1}^{n} < \mathcal{R}_{i,k+1}, \mathcal{R}_{i,0}^* > . \end{aligned}$$

and similarly

$$\begin{aligned} < w_k, w_k > &= \sum_{i=1}^{n} < \mathcal{W}_{i,k}, \mathcal{W}_{i,k} >, < q_k, r_0^* > \\ &= \sum_{i=1}^{n} < \mathcal{Q}_{i,k}, \mathcal{R}_{i,0}^* >, < w_k, s_k > \\ &= \sum_{i=1}^{n} < \mathcal{W}_{i,k}, \mathcal{S}_{i,k} > . \end{aligned}$$

Inspired by common Bi-CGSTAB and Bi-CRSTAB, the tensors $\mathcal{P}_{i,k}, \mathcal{Q}_{i,k}$ are auxiliary tensors and $\mathcal{R}_{i,k}, i = 1, 2, \dots, n$ are k-th residual of i-th equation, i.e. $\mathcal{R}_{i,k} = \mathcal{E}_i - \mathcal{L}_i(\mathcal{X}_{j,k}),\ i = 1, 2, \dots, n,\ k = 0, 1, 2, \dots$.

In regard to (2), the k-th residual is as $\mathcal{R}_k = \mathcal{E} - \mathcal{L}(\mathcal{X}_{1,k}, \mathcal{X}_{2,k}, \dots, \mathcal{X}_{n,k})$. Therefore, the residual norm is $||\mathcal{R}_k||_* = \sqrt{\sum_{i=1}^{n} ||\mathcal{R}_{i,k}||^2}$. According to the above discussions the HOBi-CGSTAB and HOBi-CRSTAB algorithms for solving the generalized coupled Sylvester tensor Eq. (1), can be presented as follows.

Algorithm (HOBi-CGSTAB)

Input matrices A_{ijl} and tensors $\mathcal{X}_{j,0}, \mathcal{E}_i$ for $i, j = 1, 2, \dots, n$ and $l = 1, 2, \dots, d$.

(1) Set $\mathcal{P}_{i,0} = \mathcal{R}_{i,0} = \mathcal{E}_i - \mathcal{L}_i(\mathcal{X}_{j,0})$.
(2) Choose arbitrary tensors $\mathcal{R}_{i,0}^*$ such that $\sum_{i=1}^{n} < \mathcal{R}_{i,0}, \mathcal{R}_{i,0}^* > \neq 0$.
(3) For $i = 1, 2, \dots, n$ and $k = 0, 1, \dots,$ until $||\mathcal{R}_k||_*$ small enough Do
(4) $\mathcal{Q}_{i,k} = \mathcal{L}_i(\mathcal{P}_{j,k})$, $\alpha_k = (\sum_{i=1}^{n} < \mathcal{R}_{i,k}, \mathcal{R}_{i,0}^* >)/(\sum_{i=1}^{n} < \mathcal{Q}_{i,k}, \mathcal{R}_{i,0}^* >)$.
(5) $\mathcal{S}_{i,k} = \mathcal{R}_{i,k} - \alpha_k \mathcal{Q}_{i,k}$, $\mathcal{W}_{i,k} = \mathcal{L}_i(\mathcal{S}_{j,k})$,
$\omega_k = (\sum_{i=1}^{n} < \mathcal{W}_{i,k}, \mathcal{S}_{i,k} >)/\sum_{i=1}^{n} < \mathcal{W}_{i,k}, \mathcal{W}_{i,k} >$.
(6) $\mathcal{X}_{i,k+1} = \mathcal{X}_{i,k} + \alpha_k \mathcal{P}_{i,k} + \omega_k S_{i,k}$, $\mathcal{R}_{i,k+1} = \mathcal{S}_{i,k} - \omega_k \mathcal{W}_{i,k}$.
(7) $\beta_k = (\frac{\sum_{i=1}^{n} <\mathcal{R}_{i,k+1}, \mathcal{R}_{i,0}^*>}{\sum_{i=1}^{n} <\mathcal{R}_{i,k}, \mathcal{R}_{i,0}^*>})(\frac{\alpha_k}{\omega_k})$, $\mathcal{P}_{i,k+1} = \mathcal{R}_{i,k+1} + \beta_k(\mathcal{P}_{i,k} - \omega_k \mathcal{Q}_{i,k})$.
(8) End Do

Due to the similarity of HOBi-CRSTAB with the HOBi-CGSTAB algorithm and the limited number of pages of the submitted article, writing the HOBi-CRSTAB algorithm is avoided.

PROPOSITION 2.1. *Let α_k, β_k and ω_k be the parameters obtained by HOBi-CGSTAB. The iterates in HOBi-CGSTAB satisfy the following properties*

i) $\sum_{j=1}^{n} \|\mathcal{R}_{j,k+1}\|^2 \leq \sum_{j=1}^{n} \|\mathcal{S}_{j,k}\|^2$,

ii) $\sum_{j=1}^{n} < \mathcal{S}_{j,k}, \mathcal{R}^*_{j,0} >= 0$,

iii) ω_k *minimizes* $\sum_{j=1}^{n} \|\mathcal{R}_{j,k+1}\|^2$,

iv) $\mathcal{R}_{i,k} = \mathcal{E}_i - \mathcal{L}_i(\mathcal{X}_{j,k})$,

where $i = 1, 2, \dots, n$ *and* $k = 0, 1, 2, \dots$.

3. Numerical Examples

In this section, due to existing restrictions, we give only a numerical example to show the efficiency of the HOBi-CGSTAB and HOBi-CRSTAB algorithms.

EXAMPLE 3.1. Consider the generalized coupled Sylvester tensor equation

$$\begin{cases} \mathcal{X} \times_1 A_1 \times_2 A_2 + \mathcal{Y} \times_1 B_1 \times_2 B_2 = \mathcal{E}_1, \\ \mathcal{X} \times_1 D_1 \times_2 D_2 + \mathcal{Y} \times_1 E_1 \times_2 E_2 = \mathcal{E}_2, \end{cases}$$

with

$$\begin{aligned} A_1 &= ones(m,m) + diag(3.5 + diag(rand(m))), \\ A_2 &= diag(1.5 + diag(rand(n))), \\ B_1 &= ones(m,m) - diag(1.5 + diag(rand(m))), \\ B_2 &= diag(2 + diag(rand(n))), \\ D_1 &= 1.5 \times ones(m,m) + diag(1 + diag(rand(m))), \\ D_2 &= diag(1.5 + diag(rand(n))), \\ E_1 &= ones(m,m) - diag(2.5 + diag(rand(m))), \\ E_2 &= diag(1.5 + diag(rand(n))). \end{aligned}$$

For $m = 50$ and $n = 40$, we apply the mentioned algorithms to compute the approximate solution $(\mathcal{X}_k, \mathcal{Y}_k)$. The numerical results are depicted in Figure 1, where $R_k = \log_{10} \sqrt{\|\mathcal{E}_1 - \mathcal{L}_1(\mathcal{X}_k, \mathcal{Y}_k)\|^2 + \|\mathcal{E}_2 - \mathcal{L}_2(\mathcal{X}_k, \mathcal{Y}_k)\|^2}$. As shown in Figure 1, the HOBi-CGSTAB and HOBi-CRSTAB algorithms have more superiority over the other algorithms.

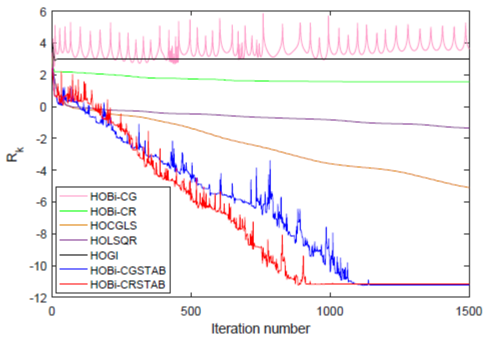

FIGURE 1. Comparison of residuals for Example 3.1.

References

1. E. Khosravi Dehdezi and S. Karimi, *Extended conjugate gradient squared and conjugate residual squared methods for solving the generalized coupled Sylvester tensor equations,* T. I. Meas. Control. (2020). DOI:10.1177/0142331220932385
2. T. G. Kolda, *Multilinear operators for higher-order decompositions,* Tech. Report SAND (2006) 2006–2081.
3. C. Lv and C. Ma, *A modified CG algorithm for solving generalized coupled Sylvester tensor equations,* Appl. Math. Comput. **365** (15) (2020) 124699.
4. H. A. Van der Vorst, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems,* SIAM J. Sci. Stat. Comput. **13** (2) (1992) 631–644.

E-mail: esakhosravidehdezi@gmail.com
E-mail: karimi@pgu.ac.ir