SCIENCE CHINA
Information Sciences

• LETTER •

August 2023, Vol. 66 189202:1–189202:2
https://doi.org/10.1007/s11432-021-3485-0

Effective distributed algorithm for solving linear matrix equations

Songsong CHENG[1], Jinlong LEI[2*], Xianlin ZENG[3], Yuan FAN[1] & Yiguang HONG[2]

[1]*Anhui Engineering Laboratory of Human-Robot Integration System and Intelligent Equipment, School of Electrical Engineering and Automation, Anhui University, Hefei* 230601*, China;*
[2]*Department of Control Science and Engineering, Tongji University, Shanghai* 200092*, China;*
[3]*Key Laboratory of Intelligent Control and Decision of Complex Systems, School of Automation, Beijing Institute of Technology, Beijing* 100081*, China*

Received 4 November 2021/Revised 28 January 2022/Accepted 12 April 2022/Published online 7 December 2022

Citation Cheng S S, Lei J L, Zeng X L, et al. Effective distributed algorithm for solving linear matrix equations. Sci China Inf Sci, 2023, 66(8): 189202, https://doi.org/10.1007/s11432-021-3485-0

Dear editor,

Solving linear matrix equations is a basic and important problem in many fields such as the computation of generalized inverses of matrices and (generalized) Sylvester equations. Also, the linear algebraic equation is a fundamental problem, which is a special form of linear matrix equations.

For linear algebraic equations of the form $A\boldsymbol{x} = \boldsymbol{c}$, which is a special case of $AX_oB = C$, a large number of distributed algorithms have been proposed to obtain the exact solution or least squares solution [1, 2]. However, distributed algorithms of solving linear matrix equations are far less investigated. Ref. [3] has developed a distributed algorithm to achieve the exact solution of the Sylvester equation. For solving the linear matrix equation of the form $AX_oB = C$, Ref. [4] has developed four distributed algorithms for all standard structures of A, B, and C. However, those distributed algorithms for solving $AX_oB = C$ are rather complex and the corresponding convergent rates are not clear.

In this study, we concentrate on a distributed least squares solution of linear matrix equations of the form $AX_oB = C$ with the following contributions: firstly, we establish a simpler optimization problem with two consensus constraints and propose a succinct algorithm with four variables; secondly, we show the proposed algorithm linearly converges to the optimum with an explicit rate. Before providing our main results, the following preliminaries on graph theory are necessary.

Graph theory. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote an undirected graph of a network, where $\mathcal{V} = \{1, \ldots, n\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ are sets of nodes and edges, respectively. We mark $j \in \mathcal{N}_i$ if $\{j, i\} \in \mathcal{E}$, namely, node j is the neighbor of node i and exchanges information with node i. $P = [p_{ij}] \in \mathbb{R}^{n\times n}$ is the adjacency matrix such that $p_{ij} = p_{ji} > 0$, if $\{j, i\} \in \mathcal{E}$ and $p_{ij} = 0$ otherwise. Moreover, $D = \mathrm{diag}\{\sum_{j=1}^n p_{1j}, \ldots, \sum_{j=1}^n p_{nj}\}$ and $L = [l_{ij}] = D - P$ is the Laplacian matrix of the graph $\mathcal{G}$. Specifically, if the undirected graph $\mathcal{G}$ is connected, we have that $L = L^{\mathrm{T}} \succeq 0$, $\mathrm{rank}(L) = n - 1$, and $\ker(L) = \{k\mathbf{1}_n | k \in R\}$.

Assumption 1. The graph $\mathcal{G}$ is connected and undirected.

Problem formulation. Consider the following linear matrix equation:

$$AX_oB = C, \tag{1}$$

where $A \in \mathbb{R}^{m\times p}$, $B \in \mathbb{R}^{q\times r}$, and $C \in \mathbb{R}^{m\times r}$ are known matrices, while $X_o \in \mathbb{R}^{p\times q}$ is an unknown matrix to be determined.

If $\mathrm{vec}(C) \in \mathrm{span}(B^{\mathrm{T}} \otimes A)$, then Eq. (1) has exact solutions; if $\mathrm{vec}(C) \notin \mathrm{span}(B^{\mathrm{T}} \otimes A)$, then Eq. (1) has least squares solutions. The objective of this study is to solve the least squares solution of (1) with rank conditions $\mathrm{rank}(A) = p$ and $\mathrm{rank}(B) = q$. $X_o^* = \min_{X_o} \frac{1}{2}\|AX_oB - C\|_{\mathrm{F}}^2$ is a least squares solution of (1), namely, $A^{\mathrm{T}}(AX_o^*B - C)B^{\mathrm{T}} = \mathbf{0}$.

In this study, we consider the case that A, B, and C have a row-column-column (RCC) structure, namely,

$$A = [A_{\mathrm{h}1}^{\mathrm{T}}, A_{\mathrm{h}2}^{\mathrm{T}}, \ldots, A_{\mathrm{h}n}^{\mathrm{T}}]^{\mathrm{T}},$$
$$B = [B_{\mathrm{v}1}, B_{\mathrm{v}2}, \ldots, B_{\mathrm{v}n}],$$
$$C = [C_{\mathrm{v}1}, C_{\mathrm{v}2}, \ldots, C_{\mathrm{v}n}],$$

where $A_{\mathrm{h}i} \in \mathbb{R}^{m_i\times p}$, $B_{\mathrm{v}i} \in \mathbb{R}^{q\times r_i}$, and $C_{\mathrm{v}i} \in \mathbb{R}^{m\times r_i}$ with $\sum_{i=1}^n m_i = m$ and $\sum_{i=1}^n r_i = r$. A subscript "h" ("v") means a "horizontal (vertical) partition" corresponding to a split of the matrix across the row (column) dimension.

In a distributed scheme, agent i only has access to subblocks $A_{\mathrm{h}i}$, $B_{\mathrm{v}i}$, $C_{\mathrm{v}i}$ and holds local variable $X_i \in \mathbb{R}^{p\times q}$. To decouple A and B, we introduce an intermediate matrix $Y_i = [Y_{i\mathrm{h}1}^{\mathrm{T}}, \ldots, Y_{i\mathrm{h}n}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{m\times q}$ with $Y_{i\mathrm{h}j} \in \mathbb{R}^{m_j\times q}$, $j \in \mathcal{V}$ for agent i, such that

$$Y_{i\mathrm{h}i} = A_{\mathrm{h}i}X_i,\ C_{\mathrm{v}i} = Y_iB_{\mathrm{v}i},\ X_i = X_j,\ Y_i = Y_j, \tag{2}$$

where $\forall i, j \in \mathcal{V}$. Furthermore, we transform (2) into the

* Corresponding author (email: leijinlong@tongji.edu.cn)

following distributed optimization problem:

$$\begin{aligned}&\min_{X,Y}\ \frac{1}{2}\sum_{i=1}^{n}\left(\|A_{\mathrm{h}i}X_i-Y_{i\mathrm{h}i}\|_{\mathrm{F}}^2+\|Y_iB_{\mathrm{v}i}-C_{\mathrm{v}i}\|_{\mathrm{F}}^2\right),\\&\text{s.t.}\quad X_i=X_j,\ Y_i=Y_j,\ \forall i,j\in\mathcal{V}.\end{aligned}\tag{3}$$

Remark 1. This study devotes to solving the linear matrix equation with the RCC structure of A, B, and C. The algorithm design and the convergence analysis of this study can be easily extended to other standard structures. We omit this discussion due to space limitations.

Different from the algorithm in [4] with 2 primal and 3 dual variables, we formulate the optimization problem by constructing two equalities $Y_iB_{\mathrm{v}i}=C_{\mathrm{v}i}$ and $A_{\mathrm{h}i}X_i=Y_{i\mathrm{h}i}$ in the objective function, simultaneously. Then the formulated optimization problem (3) only has consensus constraints of variables X and Y and is simpler than the counterpart of [4], which has three equality constraints.

For (3), we have an augmented Lagrange function:

$$\begin{aligned}\mathcal{L}=&\frac{1}{2}\sum_{i=1}^{n}\left(\|A_{\mathrm{h}i}X_i-Y_{i\mathrm{h}i}\|_{\mathrm{F}}^2+\|Y_iB_{\mathrm{v}i}-C_{\mathrm{v}i}\|_{\mathrm{F}}^2\right)\\&+\langle\Lambda_1,L_1X\rangle_{\mathrm{F}}+\langle\Lambda_2,L_2Y\rangle_{\mathrm{F}}+\frac{1}{2}\langle X,L_1X\rangle_{\mathrm{F}}\\&+\frac{1}{2}\langle Y,L_2Y\rangle_{\mathrm{F}},\end{aligned}\tag{4}$$

where $X=[X_1^{\mathrm{T}},X_2^{\mathrm{T}},\ldots,X_n^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{np\times q}$, $Y=[Y_1^{\mathrm{T}},Y_2^{\mathrm{T}},\ldots,Y_n^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{nm\times q}$, $\Lambda_1=[\Lambda_{11}^{\mathrm{T}},\Lambda_{12}^{\mathrm{T}},\ldots,\Lambda_{1n}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{np\times q}$, $\Lambda_2=[\Lambda_{21}^{\mathrm{T}},\Lambda_{22}^{\mathrm{T}},\ldots,\Lambda_{2n}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{nm\times q}$, $L_1=L\otimes I_q$, and $L_2=L\otimes I_m$. Based on (4), we propose the following algorithm from the primal-dual perspective:

$$\begin{cases}X_{i,t+1}=X_{i,t}-\alpha\Big[A_{\mathrm{h}i}^{\mathrm{T}}(A_{\mathrm{h}i}X_{i,t}-Y_{i\mathrm{h}i,t})+\sum_{j=1}^{n}l_{ij}\\\qquad\qquad\times(\Lambda_{1j,t}+X_{j,t})\Big],\\Y_{i,t+1}=Y_{i,t}-\alpha\Big[(Y_{i,t}B_{\mathrm{v}i}-C_{\mathrm{v}i})B_{\mathrm{v}i}^{\mathrm{T}}+\mathcal{I}_{m_i}(Y_{i\mathrm{h}i,t}\\\qquad\qquad-A_{\mathrm{h}i}X_{i,t})+\sum_{j=1}^{n}l_{ij}(\Lambda_{2j,t}+Y_{j,t})\Big],\\\Lambda_{1i,t+1}=\Lambda_{1i,t}+\alpha\sum_{j=1}^{n}l_{ij}X_{j,t},\\\Lambda_{2i,t+1}=\Lambda_{2i,t}+\alpha\sum_{j=1}^{n}l_{ij}Y_{j,t},\end{cases}\tag{5}$$

where $\mathcal{I}_{m_i}=[\mathbf{0}_{m_{[i-1]}\times m_i}^{\mathrm{T}},I_{m_i},\mathbf{0}_{(m-m_{[i]})\times m_i}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{m\times m_i}$, $\sum_{i=1}^{n}m_i=m$, $m_{[i]}=\sum_{j=1}^{i}m_j$, $X_{i,0}\in\mathbb{R}^{p\times q}$, $Y_{i,0}\in\mathbb{R}^{m\times q}$, $\Lambda_{1i,0}\in\mathbb{R}^{p\times q}$, and $\Lambda_{2i,0}\in\mathbb{R}^{m\times q}$.

We show that the least squares solution of (1) can be achieved by solving the distributed optimization problem in (3).

Theorem 1. Under Assumption 1, if $(\mathbf{1}_n\otimes X_o^*,\mathbf{1}_n\otimes Y_o^*)$ is an optimal solution of (3), then X_o^* is the least squares solution of (1) and (X_o^*,Y_o^*) satisfies $A^{\mathrm{T}}(AX_o^*-Y_o^*)=\mathbf{0}$.

Remark 2. The authors in [4] have established the optimization model with the necessary constraint $AX_o^*-Y_o^*=\mathbf{0}$ and introduced a dual variable for this constraint. In our optimization model, we relax the constraint $AX_o^*-Y_o^*=\mathbf{0}$ as $A^{\mathrm{T}}(AX_o^*-Y_o^*)=\mathbf{0}$ and still ensure that the least squares solution of (1) can be achieved by solving (3).

Define $H=\mathrm{blkdiag}\{H_1^{\mathrm{T}}H_1,H_2^{\mathrm{T}}H_2,\ldots,H_n^{\mathrm{T}}H_n\}$ with $H_i=[I_q\otimes(\mathbb{I}_{m_i}A),-I_q\otimes\mathbb{I}_{m_i};\mathbf{0},(\mathbb{I}_{r_i}B^{\mathrm{T}})\otimes I_m]$ and

$$0<\alpha<\frac{1}{h_m+s_1},\tag{6}$$

where h_m is the spectral norm of H and s_1 is the largest eigenvalue of L.

Next, we provide the explicit linear convergence rate of the algorithm in (5).

Theorem 2. Under Assumption 1, let X_t, Y_t, $\Lambda_{1,t}$, and $\Lambda_{2,t}$ be generated by (5). If Eq. (6) holds, then $X_{i,t}$ linearly converges to the least squares solution X_o^* with $\|X_{i,t}-X_o^*\|=\gamma^t\|X_{i,0}-X_o^*\|$, $\forall i\in\mathcal{V}$, where $\gamma=1-\alpha|\sigma_1|$ and $|\sigma_1|\in(0,h_m+s_1)$.

Remark 3. In comparison with the method for solving algebraic equations in [2], which is a special case of linear matrix equations, we propose a scheme to achieve the exact solution and least squares solution. Moreover, in comparison with the convergence results in [4], the proposed algorithm can achieve the optimum with an explicit linear rate of convergence. Furthermore, compared with the necessary computations on Moore-Penrose pseudoinverse in the network flow method in [3], the proposed method is more efficient.

Conclusion. This study designed an efficient distributed scheme for finding the least squares solution to linear matrix equations over undirected and connected graphs. Based on the simplified optimization model, we proposed a distributed optimization algorithm from the primal-dual perspective. Moreover, we investigated the effectiveness of the proposed optimization model and provided an explicit linear convergence rate for the proposed algorithm.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62103003, 62003240, 62073035, 61973002, 62173250, 61733018) and Anhui Provincial Natural Science Foundation (Grant No. 2008085J32).

Supporting information Appendixes A–D. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Wang P, Mou S, Lian J, et al. Solving a system of linear equations: from centralized to distributed algorithms. Annu Rev Control, 2019, 47: 306–322
2 Liu Y, Lageman C, Anderson B D O, et al. An Arrow-Hurwicz-Uzawa type flow as least squares solver for network linear equations. Automatica, 2019, 100: 187–193
3 Deng W, Hong Y, Anderson B D O, et al. Network flows that solve Sylvester matrix equations. IEEE Trans Automat Contr, 2022, 67: 6731–6738
4 Zeng X, Liang S, Hong Y, et al. Distributed computation of linear matrix equations: an optimization perspective. IEEE Trans Automat Contr, 2018, 64: 1858–1873