

一类矩阵方程谱范数下最小二乘解的交替方向算法

崔学莲

安徽新华学院, 安徽 合肥 230000

DOI:10.61369/ETI.2025050010

摘要： 基于交替方向法给出矩阵方程谱范数下的最小二乘解的表达式，给出计算最小二乘解的数值算法和收敛性证明，并用数值算例加以证明算法的有效性。

关键词： 矩阵方程；谱范数；最小二乘解；交替方向

Alternating Direction Algorithm for Least Squares Solutions of Quasi-Matrix Equations under Spectral Norm

Cui Xuelian

Anhui Xinhua University, Hefei, Anhui 230000

Abstract： Based on the alternating direction method, the expression of the least squares solution of the matrix equation under the spectral norm is presented. A numerical algorithm for calculating the least squares solution and a proof of convergence are provided. The effectiveness of the algorithm is demonstrated through numerical examples.

Keywords： matrix equation; spectral norm; least squares solution; alternating direction

现代金融理论、系统工程、信息与图像处理、统计分析、稳定性理论、时间序列分析、控制论和信息论、参数识别、结构设计、生物学、线性最优控制等领域中的诸多数学模型，通过有限元、有限差分、矩量法、矩形积分法等离散化方法进行离散处理后，往往对未知矩阵有诸多附加的约束条件，进而导致不同约束条件下的各类矩阵方程及其最小二乘问题。近年来，国内外诸多专家和学者对各类约束矩阵方程及其最小二乘解问题的求解进行了研究，并且取得了一系列很好的研究成果^[1-5]。但是仍然还有诸多有待继续研究的问题。

如线性矩阵方程 $AXB + CYD = E$ 是一类比较常见的矩阵方程，诸多专家和学者对其进行了研究。如2005年，Liao等^[6]利用一般内积空间的投影理论研究了该矩阵方程的最佳逼近解，同时利用矩阵的CCD分解和广义奇异值分解研究其极小范数的最小二乘问题；2007年，袁世芳和廖安平^[7]利用Kronecher积和广义逆也研究了极小范数的最小二乘问题；接着，蒋家尚^[8]也利用Kronecher积研究了该矩阵有对称解的充分必要条件和解的表达式；2008年，Peng^[9]给出了求解该矩阵方程对称和反对称、中心对称和双对称解的迭代解法；2009年，Li^[10]利用共轭梯度法研究了该矩阵方程的对称解；2014年，Li等^[11]利用Kron8echer积和广义逆研究了该矩阵方程的最小二乘解。2015年，Ke等^[12]利用交替方向法研究了该矩阵方程的最小二乘解、非负最小二乘解及其最佳逼近解的问题。一般情况下，约束矩阵方程 $AXB + CYD = E$ 问题都是在Frobenius范数下讨论的，对于谱范数下该问题的最小二乘解研究较少。

为此，本文研究谱范数下该约束矩阵方程的最小二乘问题。

一、符号说明

$R^{m \times n}$ ：全体 $m \times n$ 阶实矩阵组成的集合

$SR^{n \times n}$ ：全体 $n \times n$ 阶实对称矩阵组成的集合

$A \geq 0$ ：矩阵 A 的元素大于零

A^T ：矩阵 A 的转置

$\langle A, B \rangle$ ：矩阵 A 与 B 的内积，其定义为： $\langle A, B \rangle = \text{tr}(A^T B)$

二、问题的解及收敛性证明

求解矩阵方程 $AXB + CYD = E$ 谱范数下最小二乘解即等价于

问题1 给定列满秩矩阵 A ， $C \in R^{m \times n}$ ，行满秩矩阵 B ， $D \in R^{n \times p}$ 和任意矩阵 $E \in R^{m \times p}$ ，求矩阵 $X, Y \in R^{m \times n}$ 满足

$$\min \|AXB + CYD - E\|_2, \quad (1)$$

其中 $\|\cdot\|_2$ 表示谱范数，其定义为 $\|Q\|_2 = \sigma_1(Q)$ ， $\sigma_1(Q) \geq \sigma_2(Q) \geq \dots \geq \sigma_n(Q) \geq 0$ 为矩阵 $Q \in R^{m \times n}$ 的奇异值。

显然问题1等价于

项目来源：安徽省高等学校科学研究项目（自然科学类）

项目名称：基于深度学习的图像融合网络的探究

项目号：2024AH050611

作者简介：崔学莲（1993.03-），女，汉族，安徽长丰人，硕士研究生，助教，研究方向：数值代数。

$$\min \|Z\|_2 \quad (2)$$

$$s.t. AXB + CYD - E - Z = 0$$

$$X \in R^{n \times n}, Y \in R^{m \times n}$$

问题(2)对应的增广拉格朗日函数为

$$L_\rho(X, Y, Z, M) = \|Z\|_2 - \langle M, AXB + CYD - Z - E \rangle + \frac{\rho}{2} \|AXB + CYD - Z - E\|_F^2, \quad (3)$$

其中 $\rho > 0$ 为罚参数, $M \in R^{m \times p}$ 为拉格朗日乘子。

利用目标函数可分离特性, 采用交替方向法求解子问题, 有如下迭代形式

$$\begin{cases} X^{k+1} = \arg \min \{L_\rho(X, Y^k, Z^k, M^k) | X \in R^{n \times n}\} \\ Y^{k+1} = \arg \min \{L_\rho(X^{k+1}, Y, Z^k, M^k) | Y \in R^{m \times n}\} \\ Z^{k+1} = \arg \min \{L_\rho(X^{k+1}, Y^{k+1}, Z, M^k) | Z \in R^{m \times p}\} \\ M^{k+1} = M^k - \rho(AX^{k+1}B + CY^{k+1}D - Z^{k+1} - E) \end{cases},$$

而交替方向法对于求解三个变量问题的收敛性并不能保证,

为此引入一个修正矩阵变量 ξ , 进而设计如下迭代格式

$$\begin{cases} \tilde{X}^k = \arg \min \left\{ \|Z\|_2 - \langle M^k, AXB + CY^kD - \xi^k - E \rangle + \frac{\rho}{2} \|AXB + CY^kD + \xi^k - E\|_F^2 \right\} \\ \tilde{Y}^k = \arg \min \left\{ \|Z\|_2 - \langle M^k, A\tilde{X}^kB + CYD - \xi^k - E \rangle + \frac{\rho}{2} \|A\tilde{X}^kB + CYD + \xi^k - E\|_F^2 \right\} \\ \tilde{Z}^k = \arg \min \left\{ \|Z\|_2 - \langle M^k, A\tilde{X}^kB + C\tilde{Y}^kD - Z - E \rangle + \frac{\rho}{2} \|A\tilde{X}^kB + C\tilde{Y}^kD - Z - E\|_F^2 \right\} \\ \tilde{M}^k = M^k - \rho(A\tilde{X}^kB + C\tilde{Y}^kD - \tilde{Z}^k - E). \end{cases} \quad (4)$$

$$\begin{cases} X^{k+1} = \tilde{X}^k \\ Y^{k+1} = \tilde{Y}^k \\ Z^{k+1} = \tilde{Z}^k \\ M^{k+1} = \tilde{M}^k - \alpha(M^k - \tilde{M}^k). \end{cases} \quad (5)$$

问题(4)中 X 子问题的求解

$$\begin{aligned} \tilde{X}^k &= \arg \min \left\{ \|Z\|_2 - \langle M^k, AXB + CY^kD - \xi^k - E \rangle + \frac{\rho}{2} \|AXB + CY^kD - \xi^k - E\|_F^2 | X \in R^{n \times n} \right\} \\ &= \arg \min \left\{ \frac{\rho}{2} \|AXB + CY^kD - \xi^k - E + \frac{1}{\rho} M^k\|_F^2 | X \in R^{n \times n} \right\}. \end{aligned} \quad (6)$$

记问题(6)中的目标函数为 $f(X)$, 对其求梯度可得

$$\nabla f(X) = \rho A^T (AXB + CY^kD - \eta^k - E + \frac{1}{\rho} M^k) B^T,$$

令 $\nabla f(X) = 0$ 可得

$$\tilde{X}^k = (A^T A)^{-1} A^T (\xi^k + E + \frac{1}{\rho} M^k - CY^kD) B^T (BB^T)^{-1} \quad (7)$$

问题(6)中 Y 子问题的求解

$$\tilde{Y}^k = (C^T C)^{-1} C^T (\xi^k + E + \frac{1}{\rho} M^k - A\tilde{X}^kB) D^T (DD^T)^{-1}. \quad (8)$$

问题(6)中 Z 的子问题的求解

$$\begin{aligned} \tilde{Z}^k &= \arg \min \left\{ \|Z\|_2 - \langle M^k, A\tilde{X}^kB + C\tilde{Y}^kD - Z - E \rangle + \frac{\rho}{2} \|A\tilde{X}^kB + C\tilde{Y}^kD - Z - E\|_F^2 | Z \in R^{m \times p} \right\} \\ &= \arg \min \left\{ \|Z\|_2 + \frac{\rho}{2} \|A\tilde{X}^kB + C\tilde{Y}^kD - Z - E - \frac{1}{\rho} M^k\|_F^2 | Z \in R^{m \times p} \right\}. \end{aligned} \quad (9)$$

在求解问题(9)之前, 首先给出引理1, 引理2和引理3。

引理1^[13] 给定 $\tau > 0$ 和矩阵 $\bar{Y} \in R^{m \times n}$, 令 $\bar{Y} \in R^{m \times n} (m \leq n)$ 的奇异值分解为

$$\bar{Y} = U [\Sigma(\bar{Y}), 0] V^T,$$

其中 $\Sigma(\bar{Y}) = \text{diag}(\sigma(\bar{Y}))$, $\sigma(\bar{Y}) = [\sigma_1, \sigma_2, \dots, \sigma_m] \in R^m$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m$. 则

列极小化问题

$$\min \{\|y - \sigma(\bar{Y})\|_1, \|y\|_1 \leq \tau\}, \text{ 其中 } \|y\|_1 = \sum_{i=1}^m |y_i|,$$

有最优解 $y^* \in R^m$. 若令 k 是使不等式 $\sigma_k > z_k \geq \sigma_{k+1}$ 成立的最小下标, 定义 z_k 为 $z_k = \frac{1}{k} \sum_{i=1}^k \sigma_i$, y^* 的分量可通过如下方式计算:

$$(y^*)_i = \begin{cases} \sigma_i - z_k, & i \leq k, \\ 0, & i > k. \end{cases}$$

引理2^[13] 给定矩阵 $\bar{Y} \in R^{m \times n} (m \leq n)$, 定义核范数下半径为 τ 的闭球为

$$B_\tau^* = \{X \in R^{m \times n} | \|X\|_* \leq \tau\}.$$

利用 John von Neumann 迹不等式和 Frobenius 范数和谱范数的酉不变性, 则 $\bar{Y}$ 在闭球 B_τ^* 上的投影 $P_{B_\tau^*}(\bar{Y})$, 即矩阵极小化问题

$$\min \{\|Y - \bar{Y}\|_F | \|Y\|_* \leq \tau\}$$

有最优解

$$Y^* = U [\text{diag}(y^*), 0] V^T,$$

其中 $\|Y\|_*$ 表示 Y 的核范数, 即 $\|Y\|_* = \sum_{i=1}^n \sigma_i(Y)$, y^* 可通过引理1计算。

引理3^[13] 给定 $\tau > 0$ 和矩阵 $\bar{Y} \in R^{m \times n}$, 利用闭凸函数的 Moreau-Yosida 正则化, 则问题

$$\min_{X \in R^{m \times n}} \tau \|X\|_2 + \frac{1}{2} \|X - \bar{Y}\|_F^2$$

有最优解

$$X^* = \bar{Y} - P_{B_\tau^*}(\bar{Y}),$$

其中 $P_{B_\tau^*}(\cdot)$ 表示到闭球 $B_\tau^* = \{X \in R^{m \times n} | \|X\|_* \leq \tau\}$ 内的投影算子。

令 $\bar{Y} = A\tilde{X}^kB + C\tilde{Y}^kD - E - \frac{1}{\rho} M^k$, $\bar{Y} \in R^{m \times p} (m \leq p)$ 的奇异值分解为

$$\bar{Y} = U [\Sigma(\bar{Y}), 0] V^T,$$

其中 $\Sigma(\bar{Y}) = \text{diag}(\sigma(\bar{Y}))$, $\sigma(\bar{Y}) = [\sigma_1, \sigma_2, \dots, \sigma_m] \in R^m$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m$. 基于以上引理和定理可得式(9)有解

$$\begin{aligned} \tilde{Z}^k &= \bar{Y} - P_{B_\tau^*}(\bar{Y}) \\ &= \bar{Y} - U [\text{diag}(y^*), 0] V^T, \end{aligned} \quad (10)$$

其中 y^* 是列极小化问题 $\min \{\|y - \sigma(\bar{Y})\|_1, \|y\|_1 \leq \tau/\rho\}$, $\|y\|_1 = \sum_{i=1}^m |y_i|$ 的最优解, 令 k 是使不等式 $\sigma_k > z_k \geq \sigma_{k+1}$ 成立的最小下标, 定义 z_k 为 $z_k = \frac{1}{k} \sum_{i=1}^k \sigma_i$, y^* 分量可通过如下方式计算:

$$(y^*)_i = \begin{cases} \sigma_i - z_k, & i \leq k, \\ 0, & i > k. \end{cases}$$

下面给出求解问题(2)的迭代格式

算法1

1) 给定参数 $\rho > 0$, $\varepsilon > 0$, $\alpha \in (0, \frac{3-\sqrt{5}}{2})$, 设置初始矩阵 (Y^0, Z^0, M^0) , $\xi^0 = -Z^0$ 和 $k = 0$;

2) 按照式(7), 式(8), 式(9)计算 $\tilde{X}^k$, $\tilde{Y}^k$, $\tilde{Z}^k$;

3) 计算 $\tilde{M}^k = M^k - \rho(A\tilde{X}^kB + C\tilde{Y}^kD - \tilde{Z}^k - E)$, 通过式(5)产生新的迭代 $(X^{k+1}, Y^{k+1}, \xi^{k+1}, M^{k+1})$;

4) 若满足终止条件 $\max \{\|CY^kD - C\tilde{Y}^kD\|_F, \|\xi^k + \tilde{Z}^k\|_F, \|M^k - \tilde{M}^k\|_F\} < \varepsilon$, 则运行结束; 否则令 $k = k + 1$, 转2)。

给定算法1的当前迭代步 $(\tilde{X}^k, \tilde{Y}^k, \tilde{Z}^k, \tilde{M}^k)$, 可得如下变分不等式

$$\begin{cases} \langle X - \tilde{X}^k, -A^T \tilde{M}^k B^T + \rho A^T C(Y^k - \tilde{Y}^k) D B^T + \rho A^T (\xi^k + \tilde{Z}^k) B^T \rangle \geq 0, \\ \langle Y - \tilde{Y}^k, -C^T \tilde{M}^k D^T + \rho C^T (\xi^T + \tilde{Z}^k) D^T \rangle \geq 0, \\ \langle Z - \tilde{Z}^k, \partial \|\tilde{Z}^k\|_2 + \tilde{M}^k \rangle \geq 0, \\ \left\langle M - \tilde{M}^k, A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E - \frac{1}{\rho} (M^k - \tilde{M}^k) \right\rangle \geq 0. \end{cases} \quad (11)$$

显然, 当 $CY^k D - C\tilde{Y}^k D = 0$, $\xi^k + \tilde{Z}^k = 0$, $M^k - \tilde{M}^k = 0$ 时, $(\tilde{X}^k, \tilde{Y}^k, \tilde{Z}^k, \tilde{M}^k)$ 是问题 (2) 的解, 故可得算法 1 的终止准则为

$$\max \left\{ \|CY^k D - C\tilde{Y}^k D\|_F, \|\xi^k + \tilde{Z}^k\|_F, \|M^k - \tilde{M}^k\|_F \right\} < \varepsilon.$$

3 算法的收敛性证明及数值例子

定义 1^[14] 设 $f(X)$ 是 $R^{m \times n}$ 上的凸函数, 那么 $f(X)$ 在 X 点的次微分 $\partial f(X)$ 的定义如下:

$$\partial f(X) = \left\{ \xi \in R^{m \times n} \mid \langle \xi, Y - X \rangle + f(X) \leq f(Y), \forall Y \in R^{m \times n} \right\},$$

其中矩阵 $\xi \in \partial f(X)$ 称为次微分中的元素, 也简称为次微分或次梯度。

根据定义 1, 令 $f(X) = 0$, $g(Y) = 0$, $h(Z) = \|Z\|_2$, 则有

$$F(X) = \partial f(X) = 0, \quad G(Y) = \partial g(Y) = 0, \quad H(Z) = \partial h(Z).$$

因而求解问题 (2) 的最优解等价于寻求

$$W^* = (X^*, Y^*, Z^*, M^*) \in w(R^{n \times n}, R^{n \times n}, R^{m \times p}, R^{m \times p}),$$

使其满足变分不等式:

$$\langle W - W^*, Q(W) \rangle \geq 0, \quad (12)$$

$$\text{其中 } W = \begin{pmatrix} X \\ Y \\ Z \\ M \end{pmatrix}, \quad Q(W) = \begin{pmatrix} -A^T M B^T \\ -C^T M D^T \\ \partial \|Z\|_2 + M \\ A X B + C Y D - Z - E \end{pmatrix},$$

即找到 (X^*, Y^*, Z^*, M^*)

使其满足下式:

$$\begin{cases} \langle X - X^*, -A^T M^* B^T \rangle \geq 0, \\ \langle Y - Y^*, -C^T M^* D^T \rangle \geq 0, \\ \langle Z - Z^*, \partial \|Z^*\|_2 + M^* \rangle \geq 0, \\ \langle M - M^*, A X^* B + C Y^* D - Z^* - E \rangle \geq 0. \end{cases} \quad (13)$$

引理 4^[15] 令 $W^* = (X^*, Y^*, Z^*)$ 是问题 (2) 的任意一个解, M^* 是相应的拉格朗日乘子, 则算法产生的迭代序列 $(X^k, Y^k, Z^k, \xi^k, M^k)$ 满足

$$\begin{aligned} & \frac{1}{\rho} \langle M^k - M^*, A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E \rangle \\ & \geq \|A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E\|_F^2 + \langle A \tilde{X}^k B - A X^* B, C \tilde{Y}^k D - C \tilde{Y}^k D + \tilde{Z}^k - \xi^k \rangle \\ & \quad + \langle C \tilde{Y}^k D - C Y^* D, \xi^k + \tilde{Z}^k \rangle. \end{aligned} \quad (14)$$

由式 (14) 可得如下引理。

引理 5^[15] 令 $W^* = (X^*, Y^*, Z^*)$ 是问题 (2) 的任意一个解, M^* 是相应的拉格朗日乘子, 则算法产生的迭代序列 $(X^k, Y^k, Z^k, \xi^k, M^k)$ 满足

$$\begin{aligned} & \frac{1}{\rho} \langle M^k - M^*, A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E \rangle + \\ & \langle C Y^k D - C \tilde{Y}^k D, C Y^k D - C \tilde{Y}^k D \rangle + \langle Z^* - \xi^k, C Y^k D - C \tilde{Y}^k D + \tilde{Z}^k - \xi^k \rangle \\ & \geq \|A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E\|_F^2 + \|\xi^k + \tilde{Z}^k\|_F^2 + \|C Y^k D - C \tilde{Y}^k D\|_F^2 + \\ & \langle A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E, C Y^k D - C \tilde{Y}^k D + \xi^k + \tilde{Z}^k \rangle + \langle C Y^k D - C \tilde{Y}^k D, \xi^k + \tilde{Z}^k \rangle. \end{aligned} \quad (15)$$

由式 (15) 可得如下引理。

引理 6^[15] 令 $W^* = (X^*, Y^*, Z^*)$ 是问题 (5) 的任意一个解, M^* 是相应的拉格朗日乘子, 则算法产生的迭代序列 $(X^k, Y^k, Z^k, \xi^k, M^k)$ 满足

$$\begin{aligned} & \|C Y^{k+1} D - C Y^* D\|_F^2 + \|M^{k+1} - M^*\|_{\frac{1}{\rho^2}}^2 + \|\xi^{k+1} + Z^*\|_F^2 \\ & \leq \|C Y^k D - C Y^* D\|_F^2 + \|M^k - M^*\|_{\frac{1}{\rho^2}}^2 + \|\xi^k + Z^*\|_F^2 - \\ & \quad \alpha \left(\frac{3 - \sqrt{5}}{2} - \alpha \right) \left\{ \|A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E\|_F^2 + \right. \\ & \quad \left. \|\xi^k + \tilde{Z}^k\|_F^2 + \|C Y^k D - C \tilde{Y}^k D + \xi^k + \tilde{Z}^k\|_F^2 \right\}. \end{aligned} \quad (16)$$

定理 1 由算法产生的迭代序列 $\{\tilde{X}^k, \tilde{Y}^k, \tilde{Z}^k, \tilde{M}^k\}$ 收敛于问题 (2) 的一个解。

证明 由算法 1 中 $\alpha \in (0, \frac{3 - \sqrt{5}}{2})$, 可知式 (16) 满足下式

$$\begin{aligned} & \|C Y^{k+1} D - C Y^* D\|_F^2 + \|M^{k+1} - M^*\|_{\frac{1}{\rho^2}}^2 + \|\xi^{k+1} + Z^*\|_F^2 \\ & \leq \|C Y^k D - C Y^* D\|_F^2 + \|M^k - M^*\|_{\frac{1}{\rho^2}}^2 + \|\xi^k + Z^*\|_F^2. \end{aligned} \quad (17)$$

故 $\{C Y^k D\}$, $\{\xi^k\}$, $\{M^k\}$ 均有界, 且式 (16) 可以写成

$$\begin{aligned} & \alpha \left(\frac{3 - \sqrt{5}}{2} - \alpha \right) \left\{ \|A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E\|_F^2 + \right. \\ & \quad \left. \|\xi^k + Z^*\|_F^2 + \|C Y^k D - C \tilde{Y}^k D + \xi^k + \tilde{Z}^k\|_F^2 \right\} \\ & \leq (\|C Y^k D - C Y^* D\|_F^2 + \|M^k - M^*\|_{\frac{1}{\rho^2}}^2 + \|\xi^k + Z^*\|_F^2) - \\ & \quad (\|C Y^{k+1} D - C Y^* D\|_F^2 + \|M^{k+1} - M^*\|_{\frac{1}{\rho^2}}^2 + \|\xi^{k+1} + Z^*\|_F^2). \end{aligned} \quad (18)$$

当 $k = 0, 1, \dots, \infty$ 时, 将其依次相加求和可得

$$\begin{aligned} & \alpha \left(\frac{3 - \sqrt{5}}{2} - \alpha \right) \sum_{k=0}^{\infty} \left\{ \|A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E\|_F^2 + \right. \\ & \quad \left. \|\xi^k + Z^*\|_F^2 + \|C Y^k D - C \tilde{Y}^k D + \xi^k + \tilde{Z}^k\|_F^2 \right\} \\ & \leq \|C Y^0 D - C Y^* D\|_F^2 + \|M^0 - M^*\|_{\frac{1}{\rho^2}}^2 + \|Z^* - Z^0\|_F^2 < \infty. \end{aligned}$$

进而可得

$$\begin{aligned} & \lim_{k \rightarrow \infty} \|A \tilde{X}^k B + C \tilde{Y}^k D - \tilde{Z}^k - E\|_F^2 = 0, \\ & \lim_{k \rightarrow \infty} \|C Y^k D - C \tilde{Y}^k D + \xi^k + \tilde{Z}^k\|_F^2 = 0, \\ & \lim_{k \rightarrow \infty} \|\xi^k + \tilde{Z}^k\|_F^2 = 0. \end{aligned} \quad (19)$$

故有 $\lim_{k \rightarrow \infty} \|C Y^k D - C \tilde{Y}^k D\|_F^2 = 0$, 且序列 $\{A \tilde{X}^k B\}$, $\{C \tilde{Y}^k D\}$, $\{\tilde{Z}^k\}$ 和 $\{M^k\}$ 是有界的, 又由 A , C 是列满秩矩阵, B , D 是行满秩矩阵, 可知序列 $\{\tilde{X}^k\}$, $\{\tilde{Y}^k\}$ 和 $\{\tilde{Z}^k\}$ 也是有界的。

假设序列 $\{\tilde{X}^k\}$, $\{\tilde{Y}^k\}$, $\{\tilde{Z}^k\}$, $\{\xi^k\}$, $\{M^k\}$, 分别收敛于 $\bar{X}, \bar{Y}, \bar{Z}, \bar{U}, \bar{M}$. 令其子序列 $\{\tilde{X}^{k_j}\}$, $\{\tilde{Y}^{k_j}\}$, $\{\tilde{Z}^{k_j}\}$, $\{\xi^{k_j}\}$, $\{M^{k_j}\}$ 也分别收敛于 $\bar{X}, \bar{Y}, \bar{Z}, \bar{U}, \bar{M}$. 对式 (11) 求极限可得

$$\begin{cases} \lim_{j \rightarrow \infty} \langle X - \tilde{X}^{k_j}, -A^T \tilde{M}^{k_j} B^T + \rho A^T C(Y^k - \tilde{Y}^{k_j}) D B^T + \rho A^T (\xi^k + \tilde{Z}^{k_j}) B^T \rangle \geq 0, \\ \lim_{j \rightarrow \infty} \langle Y - \tilde{Y}^{k_j}, -C^T \tilde{M}^{k_j} D^T + \rho C^T (\xi^T + \tilde{Z}^{k_j}) D^T \rangle \geq 0, \\ \lim_{j \rightarrow \infty} \langle Z - \tilde{Z}^{k_j}, \partial \|\tilde{Z}^{k_j}\|_2 + \tilde{M}^{k_j} \rangle \geq 0, \\ \lim_{j \rightarrow \infty} \left\langle M - \tilde{M}^{k_j}, A \tilde{X}^{k_j} B + C \tilde{Y}^{k_j} D - \tilde{Z}^{k_j} - E - \frac{1}{\rho} (M^k - \tilde{M}^{k_j}) \right\rangle \geq 0. \end{cases}$$

进而由式 (19) 可得

$$\begin{cases} \langle X - \bar{X}, -A^T \bar{M} B^T \rangle \geq 0, \\ \langle Y - \bar{Y}, -C^T \bar{M} D^T \rangle \geq 0, \\ \langle Z - \bar{Z}, \partial \|Z\|_2 + \bar{M} \rangle \geq 0, \\ \langle M - \bar{M}, A\bar{X}B + C\bar{Y}D - \bar{Z} - E \rangle \geq 0. \end{cases}$$

因此, $(\bar{X}, \bar{Y}, \bar{Z}, \bar{M})$ 是问题 (2) 的一个解。

易知 (X^*, Y^*, Z^*, M^*) 是问题 (2) 的任意一个解, 用 $(\bar{X}, \bar{Y}, \bar{Z}, \bar{M})$ 替代 (X^*, Y^*, Z^*, M^*) 可得式 (17) 为

$$\begin{aligned} & \|CY^{k+1}D - C\bar{Y}D\|_{\alpha}^2 + \|M^{k+1} - \bar{M}\|_{\frac{1}{\rho^2}}^2 + \|\xi^{k+1} + \bar{Z}\| \\ & \leq \|CY^kD - C\bar{Y}D\|_{\alpha}^2 + \|M^k - \bar{M}\|_{\frac{1}{\rho^2}}^2 + \|\xi^k + \bar{Z}\|. \end{aligned}$$

因为 C, D 分别为列满秩矩阵和行满秩矩阵, 故序列 $\{Y^k, \xi^k, Z^k, M^k\}$ 有唯一的收敛点, 又 A, B 分别为列满秩和行满秩矩阵, 故序列 $\{X^k\}$ 也有唯一的收敛点。因此, 由算法产生的序列 $\{\bar{X}^k, \bar{Y}^k, \bar{Z}^k, \bar{M}^k\}$ 收敛于问题 (2) 的唯一最优解 $(\bar{X}, \bar{Y}, \bar{Z}, \bar{M})$ 。

下面我们将给出数值算例来说明算法 1 求解问题 1 的高效可行性。在算法 1 中选取算法的终止准则为 $\|X_{k+1} - X_k\| \leq \varepsilon = 10^{-8}$, 参数 $\alpha = 0.2$, 罚参数 $\rho = 0.5$, 最大迭代步为 5000。

例 1 给定矩阵 A, B, C, D 和 E 如下

$$\begin{aligned} A &= \begin{pmatrix} -0.4848 & -0.1393 & 1.2724 & 1.0843 & -0.0410 & 1.5581 \\ 0.9136 & 0.2739 & 0.9193 & 0.6330 & 0.0957 & -1.0303 \\ 0.9962 & -0.9908 & 0.3843 & 0.4592 & -1.3617 & 0.9366 \\ 2.1895 & 0.1835 & 0.4090 & -0.5971 & -0.5808 & -0.5080 \\ -0.9084 & 3.2833 & -0.1765 & 0.1496 & 0.5860 & -1.8030 \\ -1.3050 & -0.8001 & 0.8636 & -0.1991 & 0.3364 & -0.3045 \\ -0.1579 & -1.1011 & 0.3229 & -0.3945 & 1.1707 & 0.9635 \end{pmatrix}, \\ B &= \begin{pmatrix} -0.2820 & 1.3940 & 0.3256 & -0.6764 & 0.2515 & -0.0172 & 1.8866 \\ -0.9764 & 0.1343 & -0.6998 & 1.0333 & 0.2361 & -1.2238 & 1.2770 \\ 0.1336 & -0.6636 & 1.1153 & -2.2744 & -0.3956 & 0.6331 & -1.8291 \\ -0.3501 & -0.0038 & -0.0630 & 0.2836 & 0.6932 & -0.3672 & -0.1284 \\ -0.5483 & 0.5312 & -0.0716 & -0.7797 & 1.9544 & 0.5210 & -1.5708 \\ 0.5263 & 1.1642 & -1.9361 & 1.9967 & -0.4026 & 0.8704 & 0.2513 \end{pmatrix}, \\ C &= \begin{pmatrix} 0.8506 & -0.2466 & -0.7847 & 0.0223 & 1.3915 & -0.7903 \\ 0.4410 & -0.2343 & -1.1099 & 1.8614 & -0.4857 & 1.5095 \\ 0.5836 & 1.9730 & 0.7163 & -0.2871 & -0.7235 & 1.0624 \\ 1.4605 & -0.4929 & 0.7858 & -0.2833 & 0.0990 & -0.5722 \\ -0.1287 & 0.0248 & -0.2472 & 0.9468 & -0.3820 & 0.2677 \\ 0.9521 & -1.5321 & -1.0843 & 0.2992 & 0.6247 & -0.7946 \\ -0.3574 & -0.1511 & -0.0477 & 0.5269 & -1.3248 & -1.6097 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} D &= \begin{pmatrix} 0.9215 & -0.3705 & -1.9203 & 0.5140 & 0.6255 & -0.2374 & 0.6127 \\ 0.8781 & 1.2236 & -1.7230 & 0.0620 & 0.9066 & -1.2843 & -1.2393 \\ 1.1203 & -0.5277 & -1.3462 & 0.0739 & 0.6267 & 0.1525 & 2.2334 \\ -0.0508 & -0.3522 & 0.3727 & -1.8480 & 0.8108 & -2.6492 & 0.3422 \\ 1.0223 & -0.8194 & -0.0869 & 0.0022 & 1.3302 & -0.4514 & 0.2322 \\ -0.5544 & -2.0849 & 1.0659 & -0.3397 & -0.3211 & 0.5151 & 0.6454 \end{pmatrix}, \\ E &= \begin{pmatrix} 0.8472 & 0.5602 & -0.4694 & 0.7649 & -0.2169 & 0.5952 & -1.3638 \\ 0.5349 & 0.0541 & 0.8049 & 0.2427 & 0.0061 & 0.0561 & 1.6418 \\ 0.9165 & 0.3012 & 0.0862 & -0.2829 & 0.0279 & -0.3093 & -0.8191 \\ -1.3078 & 0.4950 & -0.6174 & -1.3021 & -0.4408 & 0.1198 & 0.2964 \\ -0.7647 & -0.5535 & -1.2159 & 0.7198 & -2.1699 & -1.0138 & -0.3381 \\ -0.1525 & 0.0715 & 0.3037 & -0.5736 & -0.0140 & 1.0372 & 2.8760 \\ 0.9750 & 0.1449 & 0.2396 & 0.9415 & -1.3960 & 0.7195 & 0.0376 \end{pmatrix} \end{aligned}$$

利用算法 1 可得问题 1 的解为

$$\begin{aligned} X &= \begin{pmatrix} 0.5178 & -0.5652 & 1.2428 & 4.5099 & -1.2296 & 0.8178 \\ 0.3103 & 0.8179 & 2.4280 & 3.9315 & -1.1909 & 1.5057 \\ -0.4535 & 1.3951 & -1.0949 & -7.0906 & 1.7088 & -0.7161 \\ 0.7646 & -3.0424 & 0.3728 & 9.3093 & -2.4219 & 0.5742 \\ 0.3560 & -1.2486 & 0.4201 & 4.1358 & -1.3309 & 0.5122 \\ 0.6494 & -0.2245 & 2.8771 & 7.2865 & -1.9660 & 1.9844 \end{pmatrix}, \\ Y &= \begin{pmatrix} 0.1738 & -0.1438 & -0.1068 & 0.0243 & -0.0082 & -0.0860 \\ -0.0879 & 0.2441 & 0.1239 & -0.0041 & -0.0989 & 0.1595 \\ -0.4761 & 0.4182 & 0.3025 & -0.0654 & 0.0079 & 0.2521 \\ -0.2414 & 0.2226 & 0.1577 & -0.0327 & -0.0024 & 0.1350 \\ -0.7511 & 0.7220 & 0.5026 & -0.1002 & -0.0249 & 0.4398 \\ -0.1149 & 0.1076 & 0.0757 & -0.0155 & -0.0021 & 0.0654 \end{pmatrix}. \end{aligned}$$

三、结束语

为求解矩阵方程 $AXB + CYD = E$ 在谱范数下的最小二乘解。本文利用交替方向法, 给出了该问题有解的条件和解的表达式, 并用数值例子加以说明结论的正确性。对于上述研究工作, 还有诸多问题就有待于进一步发展和改进, 如由于其该算法的收敛速度和罚参数 ρ 的选取有关, 所以对于罚参数的选取还需进一步的研究。而当矩阵规模较大时, 算法的运算量也会很大, 所以对其迭代算法还需进一步探究。最后可以进一步探索其在一定约束条件下的解。

参考文献

- [1] 戴华. 用振动实验最优校正刚度, 柔度和质量矩阵 [J]. 振动工程学报, 1998, 2: 168-768.
- [2] 朱德懋, 孙久厚. 动态有限元模型建立的逆方法 [J]. 南京航空学院学报, 1986, 4: 17-24.
- [3] 蒋正新, 陆超超. 谱约束下的矩阵最佳逼近问题 [J]. 计算数学, 1986, 1: 168-768.
- [4] 须田信英. 自动控制中的矩阵理论 [M]. 北京: 科学出版社, 1979, 1-443.
- [5] CHENEY E W. Introduction to Approximation Theory [J]. Chelsea, New York, 1996.
- [6] LIAO A P, BAI Z Z, YUAN L. Best approximate solution of matrix equation $AXB + CYD = E$ [J]. SIAM Journal on Matrix Analysis and Application, 2005, 27: 675-688.
- [7] 袁世芳, 廖安平, 雷渊. 矩阵方程 $AXB + CYD = E$ 对称最小范数最小二乘解 [J]. 计算数学, 2007, 2: 203-216.
- [8] 蒋家尚, 袁永新. 矩阵方程 $AXB + CYD = E$ 的对称解 (英文) [J]. 南京大学学报半年刊, 2008, 2: 141-148.
- [9] PENG Z Y. Solution of symmetry-constrained least-squares problems [J]. Numerical Linear Algebra Applications, 2008, 15: 373-389.
- [10] LI J F, HU X Y, ZHANG L. The submatrix constraint problem of matrix equation $AXB + CYD = E$ [J]. Applied Mathematics and Computation, 2009, 215(7): 2578-2590.
- [11] LI H Y, GAO Z S, ZHAO D. Least squares solutions of the matrix equation $AXB + CYD = E$ with the least norm for symmetric arrowhead matrices [J]. Applied Mathematics and Computation, 2014, 226: 719-724.
- [12] Ke Y F, Ma C F. Alternating direction method for generalized Sylvester matrix equation $AXB + CYD = E$ [J]. Applied Mathematics and Computation, 2015, 260: 565-576.
- [13] 李姣芬. 两类矩阵逆问题和几类约束矩阵方程问题的理论和算法 [D]. 长沙: 湖南大学, 2010, 25-39.
- [14] 高岩. 非光滑优化 [M]. 北京: 科学出版社, 2008.
- [15] HAN D, YUAN X, ZHANG W, et al. An ADM-based splitting method for separable convex programming [J]. Computation Optimization and Applications, 2003, 54: 343-369.