Vol. 34 (2014)
No. 5

数 学 杂 志
J. of Math. (PRC)

矩阵方程 $A_1Z+ZB_1=C_1$ 的广义自反最佳逼近解的迭代算法

杨家稳 [1,2]，孙合明 [2]
(1. 滁州职业技术学院基础部, 安徽 滁州 239000)
(2. 河海大学理学院, 江苏 南京 210098)

摘 要: 本文研究了 Sylvester 复矩阵方程 $A_1Z+ZB_1=C_1$ 的广义自反最佳逼近解. 利用复合最速下降法，提出了一种的迭代算法. 不论矩阵方程 $A_1Z+ZB_1=C_1$ 是否相容, 对于任给初始广义自反矩阵 Z_0, 该算法都可以计算出其广义自反的最佳逼近解. 最后, 通过两个数值例子, 验证了该算法的可行性.

关键词: Sylvester 矩阵方程; Kronecker 积; 复合最速下降法; 最佳逼近; 广义自反矩阵
MR(2010) **主题分类号:** 65F30　　**中图分类号:** O241.6
文献标识码: A　　**文章编号:** 0255-7797(2014)05-0968-09

1 引言

我们知道关于反射矩阵的自反矩阵或反自反矩阵在系统与控制理论、工程、科学计算和其他领域都有广泛的应用 [1−3]. 近年来, Sylvester 复矩阵方程的自反和反自反解的研究非常活跃. Peng 和 Hu[4] 给出方程 $AX=B$ 的自反和反自反解. 文献 [5] 给出方程 $A^HXB=C$ 的自反和反自反解. 2009 年, Xu 和 Li[6] 研究了方程 $AX=B$ 的反问题的 Hermitian 自反解. 文献 [7] 给出反 Hermitian、广义反 Hamiltonian 矩阵的最佳逼近解, 等等.

上述文献都是考虑方程有解情况下的最佳逼近解, 方程无解情况下的最佳逼近解目前研究很少. 不论矩阵方程是否有解, 我们利用复合最速下降法 [8] (the hybrid steepest descent method, HSDM), 求 $A_1Z+ZB_1=C_1$ 广义自反最佳逼近解.

首先介绍本文中的符号. $C^{m\times n}$ 表示 $m\times n$ 阶复矩阵的集合, $R^{m\times n}$ 表示 $m\times n$ 阶实矩阵的集合, A^H 表示矩阵 A 的共轭转置. I 表示单位矩阵, 对于矩阵 A, $B\in C^{m\times n}$, $A\otimes B$ 表示 A 与 B 的 Kronecker 积, $\langle A,B\rangle=\mathrm{trace}(B^HA)$ 表示 A 与 B 的内积. $\|A\|$ 表示矩阵 A 的 Frobe-nius 范数, $\|A\|^2=\langle A,A\rangle$. $\|\alpha\|_2$ 表示向量 $\alpha=(x_1,x_2,\cdots x_n)^T$ 的 2- 范数, $\|\alpha\|_2=\sqrt{\alpha^T\alpha}$. 若 $P^2=I$ 且 $P^H=P$, 则称矩阵 P 是广义反射矩阵. 若 $A=PAQ$, 则称矩阵 A 为关于广义反射 P, Q 的广义自反矩阵. 记 ${C_r}^{n\times n}(P,Q)=\{A\in C^{m\times n}|PAQ=A\}$. 定义矩阵 $A=(a_{ij})\in C^{m\times n}$ 的拉伸算子为

$$\mathrm{vec}(A)=[a_{11},a_{21},\cdots,a_{m1},a_{12},a_{22},\cdots,a_{m2},\cdots,a_{1n},a_{2n},\cdots,a_{mn}]^T.$$

∇ 表示梯度算子, 即 $\nabla f(\alpha)=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n}\right)^T$, 其中 $\alpha=(x_1,x_2,\cdots,x_n)^T$, H 是希尔伯特空间, 对于映射 $T:H\to H$, 所有关于 T 的不动点的集合表示为 $\mathrm{Fix}(T):=\{x\in H|T(x)=x\}$, P_K 表示到非空凸集 K 上的投影算子.

*收稿日期: 2012-06-25　　接收日期: 2012-09-14
基金项目: 安徽省教育厅自然科学基金资助 (KJ2011B119).
作者简介: 杨家稳 (1972–), 男, 安徽滁州, 副教授, 主要研究方向: 优化算法.

本文考虑如下问题:

问题 Ⅰ 给定矩阵 A_1、$P_1 \in C^{m\times m}$, B_1、$Q_1 \in C^{n\times n}$, $C \in C^{m\times n}$, 求 $Z \in C_r^{m\times n}(P_1, Q_1)$ 使得

$$\|A_1Z + ZB_1 - C_1\| = \min. \tag{1.1}$$

问题 Ⅱ 设问题Ⅰ的解集合为 S_E, 给定 $Z^* \in C_r^{m\times n}(P_1, Q_1)$, 求 $\hat{Z} \in S_E$ 使得

$$\left\|\hat{Z} - Z^*\right\| = \min_{Z\in S_E} \|Z - Z^*\|. \tag{1.2}$$

问题Ⅱ是在 S_E 里找一个与矩阵 $Z^* \in C_r^{m\times n}(P_1, Q_1)$ 最接近的矩阵 $\hat{Z}$.

本文的结构如下: 第 2 部分介绍问题Ⅰ和问题Ⅱ的一个迭代算法并证明该算法是收敛的; 第 3 部分用两个数值例子来验证该算法的可行性; 第 4 部分给出结论.

2 迭代算法

本节首先给出解决问题Ⅰ、Ⅱ的迭代算法, 然后证明该算法收敛, 最后研究如何计算到非空凸集 K 上的投影.

算法:

步骤 1 输入矩阵 $C_1 \in C^{m\times n}$, $Z_0 \in C^{m\times n}$ 和 $Z^* \in C_r^{m\times n}(P_1, Q_1)$.

步骤 2 计算

$$A = \mathrm{real}(A_1), B = \mathrm{imag}(A_1), C = \mathrm{real}(B_1), D = \mathrm{imag}(B_1), E = \mathrm{real}(C_1), F = \mathrm{imag}(C_1),$$
$$W = I\otimes A + C^T\otimes I, N = I\otimes B + D^T\otimes I, M = 2\begin{pmatrix} W^TW + N^TN & N^TW - W^TN \\ W^TN - N^TW & W^TW + N^TN \end{pmatrix},$$
$$\nu = \frac{2}{\|M\|}, X^* = \mathrm{real}(Z^*),\quad Y^* = \mathrm{imag}(Z^*), k := 1.$$

步骤 3

$$\begin{aligned}
&X_{k-1} = \mathrm{real}(Z_{k-1}), Y_{k-1} = \mathrm{imag}(Z_{k-1}),\\
&\nabla\psi(\mathrm{vec}(Z_{k-1})) = M\begin{pmatrix} \mathrm{vec}(X_{k-1}) \\ \mathrm{vec}(Y_{k-1}) \end{pmatrix} - 2\begin{pmatrix} W^T & N^T \\ -N^T & W^T \end{pmatrix}\begin{pmatrix} \mathrm{vec}(E) \\ \mathrm{vec}(F) \end{pmatrix},\\
&T\begin{pmatrix} \mathrm{vec}(X_{k-1}) \\ \mathrm{vec}(Y_{k-1}) \end{pmatrix} = P_K\left\{\begin{pmatrix} \mathrm{vec}(X_{k-1}) \\ \mathrm{vec}(Y_{k-1}) \end{pmatrix} - \nu\nabla\psi(\mathrm{vec}(Z_{k-1}))\right\},
\end{aligned}$$

其中

$$K = \{Z|Z = P_1ZQ_1\}, \mathrm{vec}(Z^*) = \begin{pmatrix} \mathrm{vec}(X^*) \\ \mathrm{vec}(Y^*) \end{pmatrix},$$
$$\mathrm{vec}(Z_k) = (1 - \frac{2}{k+2})T(\mathrm{vec}(Z_{k-1})) + \frac{2}{k+2}(\mathrm{vec}(Z^*)).$$

步骤 4 若 $\|Z_k - Z_{k-1}\| < \varepsilon$, 则停止; 否则, $k := k + 1$, 转向步骤 3.

为了证明该算法的收敛性, 首先给出下面的定义和定理.

定义 2.1 设 $S\subset H$ 是非空凸集, $\lambda\in(0,1)$, $f(x)$ 是定义在 S 上的函数, 如果对任意 $x_1,x_2\in S$, 有

$$f(\lambda x_1+(1-\lambda)x_2)\le \lambda f(x_1)+(1-\lambda)f(x_2),$$

则称函数 $f(x)$ 是 S 上的凸函数.

定义 2.2 若 $\forall x_1,x_2\in S\subset H$, $\exists\kappa>0$, 使 $\|T(x)-T(y)\|\le\kappa\|x-y\|$ 成立, 则映射 $T:H\to H$ 称为 κ-Lipschitzian (或 κ-Lipschitzian 连续), 特别地, 若对于一切 $x,y\in H$, 有 $\|T(x)-T(y)\|\le\|x-y\|$ 成立, 则映射 $T:H\to H$ 称为非扩展映射.

定义 2.3 对于给定的集合 $S\subset H$, 若 $\forall x,y\in S$, $\exists\eta>0$ 使

$$\langle F(x)-F(y),x-y\rangle\ge\eta\|x-y\|^2$$

成立, 则称映射 $F:H\to H$ 在集合 S 上 η 强单调.

定理 2.1[8] (HSDM) 设 $T:H\to H$ 是非扩展映射且 $\mathrm{Fix}(T)\ne\varnothing$. 假设函数 $\theta:H\to R\bigcup\infty$ 存在 Gateaux 导数 $\theta':H\to H$, 且 θ' 在 $T(H)$ 上 κ-Lipschitzian 且 η 强单调. 对于任意的 $u_0\in H$ 和正实数序列 $(\lambda_n)_{n\ge1}$ 满足:

(L1) $\lim\limits_{n\to\infty}\lambda_n=0$; (L2) $\sum\limits_{n\ge1}\lambda_n=\infty$; (L3) $\lim\limits_{n\to\infty}(\lambda_n-\lambda_{n+1})\lambda_{n+1}^{-2}=0$;

(或者 (W1) $\lim\limits_{n\to\infty}\lambda_n=0$; (W2) $\sum\limits_{n\ge1}\lambda_n=\infty$; (W3) $\sum\limits_{n\ge1}|\lambda_n-\lambda_{n+1}|<\infty$),

则由 $u_{n+1}:=T(u_n)-\lambda_{n+1}\theta'(T(u_n))$ 生成的序列 $(u_n)_{n\ge0}$ 强收敛于唯一的 u^*, 其中 u^* 满足 $\theta(u^*)=\inf\{\theta(\mathrm{Fix}(T))\}$ (例如: $\lambda_n=1/n$ 是满足 (W1)–(W3) 的序列 $(\lambda_n)_{n\ge1}$).

定理 2.2 [9] 设 $K\subset H$ 是闭凸子集, 假设:

(i) $\psi:H\to R\bigcup\{\infty\}$ 是 Gateaux 可微的凸函数且它的导数 $\psi:'H\to H$ 在 H 上满足 γ-Lipschitzian;

(ii) $\theta:H\to R\bigcup\{\infty\}$ 是 Gateaux 可微的凸函数且它的导数 $\theta':H\to H$ 在 $T(H)$ 上满足 κ-Lipschitzian 和 η 强单调, 其中 $T:=P_k(I-v\psi')$, $v\in(0,2/\gamma]$;

则 $\forall u_0\in H$ 由 $u_n:=T(u_{n-1})-\lambda_n\theta'(T(u_{n-1}))$ 产生的序列 $(u_n)_{n\ge0}$ 强收敛于唯一的

$$x^*=\arg\inf_{x\in K_\psi}\theta(x),$$

其中 $K_\psi:=\arg\inf\limits_{x\in K}\psi(x)\ne\varnothing$.

在集合 K 上定义函数 ψ 和 θ, 并证明它们满足定理 2.2 的条件.

定理 2.3 设 $\psi(\mathrm{vec}(Z))=\|A_1Z+ZB_1-C_1\|^2$, 其中 $Z,C_1\in C^{m\times n}$, $A_1\in C^{m\times m}$, $B_1\in C^{n\times n}$. 令

$$A=\mathrm{real}(A_1),B=\mathrm{imag}(A_1),C=\mathrm{real}(B_1),D=\mathrm{imag}(B_1),E=\mathrm{real}(C_1),F=\mathrm{imag}(C_1),$$
$$X=\mathrm{real}(Z),Y=\mathrm{imag}(Z),W=I\otimes A+C^T\otimes I,N=I\otimes B+D^T\otimes I,$$

则

$$\nabla\psi(vec(Z))=\nabla\psi\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}=M\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}-2\begin{pmatrix}W^T & N^T\\ -N^T & W^T\end{pmatrix}\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}. \tag{2.1}$$

证

$$
\begin{aligned}
&\psi(\mathrm{vec}(Z))=\psi\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}\\
=\ &\|A_1Z+ZB_1-C_1\|^2=\|AX-BY+XC-YD-E\|^2\\
&+\|BX+AY+XD+YC-F\|^2\\
=\ &\|\mathrm{vec}(AX)-\mathrm{vec}(BY)+\mathrm{vec}(XC)-\mathrm{vec}(YD)-\mathrm{vec}(E)\|_2^2\\
&+\|\mathrm{vec}(BX)+\mathrm{vec}(AY)+\mathrm{vec}(XD)+\mathrm{vec}(YC)-\mathrm{vec}(F)\|_2^2\\
=\ &\left\|(I\otimes A)\mathrm{vec}(X)-(I\otimes B)\mathrm{vec}(Y)+(C^T\otimes I)\mathrm{vec}(X)-(D^T\otimes I)\mathrm{vec}(Y)-\mathrm{vec}(E)\right\|_2^2\\
&+\left\|(I\otimes B)\mathrm{vec}(X)+(I\otimes A)\mathrm{vec}(Y)+(D^T\otimes I)\mathrm{vec}(X)+(C^T\otimes I)\mathrm{vec}(Y)-\mathrm{vec}(F)\right\|_2^2\\
=\ &\left\|\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}\right\|_2^2\\
=\ &\left\langle\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix},\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}\right\rangle\\
=\ &\left\langle\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix},\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}\right\rangle\\
&-2\left\langle\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix},\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}\right\rangle+\left\langle\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix},\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}\right\rangle\\
=\ &\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}^T\begin{pmatrix}W & -N\\ N & W\end{pmatrix}^T\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}-2\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}^T\begin{pmatrix}W & -N\\ N & W\end{pmatrix}^T\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}\\
&+\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}^T\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix}.
\end{aligned}
$$

故

$$
\begin{aligned}
&\nabla\psi(\mathrm{vec}(Z))=\nabla\psi\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}\\
=\ &2\begin{pmatrix}W & -N\\ N & W\end{pmatrix}^T\begin{pmatrix}W & -N\\ N & W\end{pmatrix}\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}-2\begin{pmatrix}W & -N\\ N & W\end{pmatrix}^T\begin{pmatrix}\mathrm{vec}(E)\\ \mathrm{vec}(F)\end{pmatrix},
\end{aligned}
$$

则 (2.1) 式得证.

定理 2.4 设 $\theta(\mathrm{vec}(z))=\|z-z^*\|^2$, $Z=X+i*Y$, $Z^*=X^*+i*Y^*$, 其中 Z^*, $Z\in C^{m\times n}$, $X,Y\in R^{m\times n}$, 则

$$
\nabla\theta(\mathrm{vec}(z))=\nabla\theta\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}=\begin{pmatrix}2\mathrm{vec}(X-X^*)\\ 2\mathrm{vec}(Y-Y^*)\end{pmatrix}. \tag{2.2}
$$

证

$$\begin{aligned}
& \theta(\mathrm{vec}(z)) = \|z-z^*\|^2 = \|(X-X^*)+i*(Y-Y^*)\|^2 \\
= & \ \langle X-X^*, X-X^*\rangle = \langle \mathrm{vec}(X-X^*), \mathrm{vec}(X-X^*)\rangle \\
= & \ \|X-X^*\|^2 + \|Y-Y^*\|^2 = \|\mathrm{vec}(X-X^*)\|_2^2 + \|\mathrm{vec}(Y-Y^*)\|_2^2 \\
= & \ \langle \mathrm{vec}(X-X^*), \mathrm{vec}(X-X^*)\rangle + \langle \mathrm{vec}(Y-Y^*), \mathrm{vec}(Y-Y^*)\rangle \\
= & \ [\mathrm{vec}(X-X^*)]^T \mathrm{vec}(X-X^*) + [\mathrm{vec}(Y-Y^*)]^T \mathrm{vec}(Y-Y^*),
\end{aligned}$$

则 (2.2) 式得证.

可以证得 $\psi'(\mathrm{vec}(z)) = \nabla\psi(\mathrm{vec}(z))$ 和 $\theta'(\mathrm{vec}(z)) = \nabla\theta(\mathrm{vec}(z))$. 因此分别用 $\nabla\psi(\mathrm{vec}(z))$, $\nabla\theta(\mathrm{vec}(z))$ 来代替 $\psi'(\mathrm{vec}(z))$ 和 $\theta'(\mathrm{vec}(z))$.

定理 2.5 设 $\psi(\mathrm{vec}(z))$ 和 $\theta(\mathrm{vec}(z))$ 如定理 2.3, 2.4 中所示, 则有如下的结论:

(a) $\psi(\mathrm{vec}(z))$ 是凸函数;

(b) $\theta(\mathrm{vec}(z))$ 是凸函数;

(c) $\nabla\psi(\mathrm{vec}(z))$ 是 γ-Lipschitizian;

(d) $\nabla\theta(\mathrm{vec}(z))$ 是 κ-Lipschitzian;

(e) $\nabla\theta(\mathrm{vec}(z))$ 是 η 强单调的.

证 (a) 设 $\psi(\mathrm{vec}(Z)) = \|A_1Z + ZB_1 - C_1\|$, 得

$$\begin{aligned}
& \psi(\lambda\mathrm{vec}(Z_1) + (1-\lambda)\mathrm{vec}(Z_2)) \\
= & \ \|A_1[\lambda Z_1 + (1-\lambda)Z_2] + [\lambda Z_1 + (1-\lambda)Z_2]B_1 - C_1\| \\
= & \ \|\lambda(A_1Z_1 + Z_1B_1 - C_1) + (1-\lambda)(A_1Z_2 + Z_2B_1 - C_1)\| \\
\le & \ \lambda\|A_1Z_1 + Z_1B_1 - C_1\| + (1-\lambda)(\|A_1Z_2 + Z_2B_1 - C_1\| \\
= & \ \lambda\psi(\mathrm{vec}(Z_1)) + (1-\lambda)\psi(\mathrm{vec}(Z_2)), \lambda \in (0,\ 1).
\end{aligned}$$

所以由凸集定义可得, $\psi(\mathrm{vec}(Z))$ 是凸函数.

(b) 设 $f(\mathrm{vec}(Z)) = \|Z - Z^*\|$, 则

$$\begin{aligned}
& f[\lambda\mathrm{vec}(Z_1) + (1-\lambda)\mathrm{vec}(Z_2)] = \|\lambda Z_1 + (1-\lambda)Z_2 - Z^*\| \\
= & \ \|\lambda Z_1 - \lambda Z^* + \lambda Z^* + (1-\lambda)Z_2 - Z^*\| \\
= & \ \|\lambda(Z_1 - Z^*) + (1-\lambda)(Z_2 - Z^*)\| \le \|\lambda(Z_1 - Z^*)\| + \|(1-\lambda)(Z_2 - Z^*)\| \\
= & \ \lambda\|Z_1 - Z^*\| + (1-\lambda)\|Z_2 - Z^*\| = \lambda f(\mathrm{vec}(Z_1)) + (1-\lambda)f(\mathrm{vec}(Z_2)),
\end{aligned}$$

其中 $\lambda \in (0,\ 1)$. 所以, 由凸集定义可得 (b) 成立.

(c) 令

$$\begin{aligned}
W &= I\otimes A + C^T\otimes I, N = I\otimes B + D^T\otimes I, \\
M &= 2\begin{pmatrix} W^TW + N^TN & N^TW - W^TN \\ W^TN - N^TW & W^TW + N^TN \end{pmatrix},
\end{aligned}$$

根据 (2.1) 式有

$$\begin{aligned} &\|\nabla\psi(vec(Z_1))-\nabla\psi(\mathrm{vec}(Z_2))\| = \left\|M\begin{pmatrix}\mathrm{vec}(X_1)\\ \mathrm{vec}(Y_1)\end{pmatrix}-M\begin{pmatrix}\mathrm{vec}(X_2)\\ \mathrm{vec}(Y_2)\end{pmatrix}\right\|_2 \\ \le\ & \|M\|\left\|\begin{pmatrix}\mathrm{vec}(X_1)\\ \mathrm{vec}(Y_1)\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(X_2)\\ \mathrm{vec}(Y_2)\end{pmatrix}\right\|_2 = \|M\|\,\|\mathrm{vec}(Z_1)-\mathrm{vec}(Z_2)\|_2. \end{aligned}$$

因此, $\nabla\psi(\mathrm{vec}(Z))$ 是 γ-Lipschitizian, 其中 $\gamma=\|M\|$.

(d) 设 $\theta(\mathrm{vec}(Z))=\|Z-Z^*\|^2$, 由 (2.2) 式得

$$\begin{aligned} &\|\nabla\theta(\mathrm{vec}(Z_1)-\nabla\theta(\mathrm{vec}(Z_2)\| = \|2\mathrm{vec}(Z_1-Z^*)-2\mathrm{vec}(Z_2-Z^*)\|_2 \\ =\ & \|2[\mathrm{vec}(X)-\mathrm{vec}(X^*)]-2[\mathrm{vec}(Y)-\mathrm{vec}(X^*)]\|_2 \\ =\ & 2\,\|\mathrm{vec}(X)-\mathrm{vec}(Y)\|_2 = 2\|\mathrm{vec}(Z_1-Z_2)\|_2 = 2\,\|Z_1-Z_2\|\,, \end{aligned}$$

因此 $\nabla\theta(\mathrm{vec}(z))$ 是 κ-Lipschitzian, 其中 $\kappa=2$.

(e) 再由 (2.2) 式可以得到

$$\begin{aligned} &\left\langle\nabla\theta\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_1))\\ \mathrm{vec}(\mathrm{imag}(Z_1))\end{pmatrix}-\nabla\theta\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_2))\\ \mathrm{vec}(\mathrm{imag}(Z_2))\end{pmatrix},\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_1))\\ \mathrm{vec}(\mathrm{imag}(Z_1))\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_2))\\ \mathrm{vec}(\mathrm{imag}(Z_2))\end{pmatrix}\right\rangle \\ =\ & 2\left\langle\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_1))\\ \mathrm{vec}(\mathrm{imag}(Z_1))\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_2))\\ \mathrm{vec}(\mathrm{imag}(Z_2))\end{pmatrix},\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_1))\\ \mathrm{vec}(\mathrm{imag}(Z_1))\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_2))\\ \mathrm{vec}(\mathrm{imag}(Z_2))\end{pmatrix}\right\rangle \\ =\ & 2\left\|\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_1))\\ \mathrm{vec}(\mathrm{imag}(Z_1))\end{pmatrix}-\begin{pmatrix}\mathrm{vec}(\mathrm{real}(Z_2))\\ \mathrm{vec}(\mathrm{imag}(Z_2))\end{pmatrix}\right\|_2^2. \end{aligned}$$

根据定义 2.2, 可以推出 $\nabla\theta(\mathrm{vec}(z))$ 是 η 强单调的, 其中 $\eta=2$.

由定理 2.5 可知, $\psi(\mathrm{vec}(z))$ 和 $\theta(\mathrm{vec}(z))$ 满足定理 2.2 的条件, 由定理 2.2 可以推出该算法是收敛的.

在算法中, 需要计算到凸集 K 上的投影. 为此我们给出如下定理.

定理 2.6 设 $P_1\in C^{m\times m},Q_1\in C^{n\times n}$ 是广义自反矩阵,$Z\in C^{m\times n}$ 是关于矩阵 P_1 和 Q_1 的自反矩阵, $X=\mathrm{real}(Z),Y=\mathrm{imag}(Z),K=\{Z|Z=P_1ZQ_1\},\alpha_1,\alpha_2,\cdots,\alpha_s$ 是矩阵 $\begin{pmatrix}P & -Q\\ Q & P\end{pmatrix}-I$ 零空间的标准正交基, 其中 $P=\mathrm{real}({Q_1}^T\otimes P_1),Q=\mathrm{imag}({Q_1}^T\otimes P_1)$, 则到凸集 K 上的投影为

$$P_K\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}=\left\langle\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix},\alpha_1\right\rangle\alpha_1+\left\langle\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix},\alpha_2\right\rangle\alpha_2\ +\cdots+\left\langle\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix},\alpha_s\right\rangle\alpha_s. \tag{2.3}$$

证

$$
\begin{aligned}
& Z=P_1ZQ_1 \Leftrightarrow \mathrm{vec}(Z)=\mathrm{vec}(P_1ZQ_1)\\
\Leftrightarrow\ & \mathrm{vec}(Z)={Q_1}^T\otimes P_1\mathrm{vec}(Z)\\
\Leftrightarrow\ & \mathrm{vec}(X)+i*\mathrm{vec}(Y)=(P+i*Q)(\mathrm{vec}(X)+i*\mathrm{vec}(Y))\\
\Leftrightarrow\ & \mathrm{vec}(X)+i*\mathrm{vec}(Y)=(P\mathrm{vec}(X)-Q\mathrm{vec}(Y))+i*(Q\mathrm{vec}(X)+P\mathrm{vec}(Y))\\
\Leftrightarrow\ & \begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}=\begin{pmatrix}P\mathrm{vec}(X)-Q\mathrm{vec}(Y)\\ Q\mathrm{vec}(X)+P\mathrm{vec}(Y)\end{pmatrix}\\
\Leftrightarrow\ & \left[\begin{pmatrix}P & -Q\\ Q & P\end{pmatrix}-I\right]\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}=0\Leftrightarrow\left[\begin{pmatrix}P & -Q\\ Q & P\end{pmatrix}-I\right]\begin{pmatrix}\mathrm{vec}(X)\\ \mathrm{vec}(Y)\end{pmatrix}=0. \qquad (2.4)
\end{aligned}
$$

凸集 K 中的元素就是线性方程 (2.4) 的解. 因此只要给定 (2.4) 式解集合的标准正交基, 就可以利用 (2.3) 式计算出 $\mathrm{vec}(Z_0)$ 在 K 上的投影.

3 数值例子

下面用两个数值例子来验证上述算法的可行性, 所有的实验都是在 MATLAB2007R 中进行的.

例 1 考虑矩阵方程 $A_1Z+ZB_1=C_1$, 其中

$$
A_1=\begin{pmatrix}2 & 1-i & 6\\ 5 & 4+2i & -3\\ -1+i & 4 & 8\end{pmatrix}, B_1=\begin{pmatrix}5-3i & 2 & -6\\ 6 & -7 & 0\\ 2+4i & 4 & -3\end{pmatrix},
$$

$$
C_1=\begin{pmatrix}-26-2i & -52+37i & 35+24i\\ -21+75i & 28+31i & -55+33i\\ 80+76i & 12+29i & -7-i\end{pmatrix},
$$

$$
P_1=\begin{pmatrix}0 & -i & 0\\ i & 0 & 0\\ 0 & 0 & 1\end{pmatrix}, Q_1=\begin{pmatrix}0 & i & 0\\ -i & 0 & 0\\ 0 & 0 & -1\end{pmatrix}, Z^*=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}, Z_0=\begin{pmatrix}10 & 10 & 10\\ 10 & 10 & 10\\ 10 & 10 & 10\end{pmatrix}.
$$

可以验证该矩阵方程是相容的, 且有唯一如下广义自反解:

$$
Z=\begin{pmatrix}-2-5i & 1+4i & -8+7i\\ 1+4i & 2+5i & 7+8i\\ 6+3i & -3+6i & 0\end{pmatrix}\in C_r^{3\times 3}(P_1,Q_1).
$$

利用上述算法, 令 $\varepsilon=1.0\text{e-}11$, 可得

$$
Z_{5140501}=\begin{pmatrix}-2.0000-5.0000i & 1.0000+4.0000i & -8.0000+7.0000i\\ 1.0000+4.0000i & 2.0000+5.0000i & 7.0000+8.0000i\\ 6.0000+3.0000i & -3.0000+6.0000i & 0\end{pmatrix}\in C_r^{3\times 3}(P_1,Q_1),
$$

相应的余项

$$R_{5140501}=\|A_1Z+ZB_1-C_1\|=2.8055\text{e-}4,$$

其中下标是迭代步数.

例 2　仍然考虑矩阵方程 $A_1Z+ZB_1=C_1$, 其中 A_1, B_1, P_1, Q_1, Z_0 和 Z^* 同例 1 中的一样. 在本例中, 令

$$C_1=\begin{pmatrix} 8-i & -52+3i & 35+24i \\ -21+75i & 28+31i & -55+33i \\ 80+76i & 12+29i & -7-i \end{pmatrix}.$$

可以证明上述矩阵方程是不相容的. 令 $\varepsilon=1.0\text{e-}10$, 由上面算法得

$$Z_{5140501}=\begin{pmatrix} -1.0064-4.6660i & 1.8552+4.7525i & -7.4849+5.1686i \\ 1.8552+4.7525i & 1.0064+4.6660i & 5.1686+7.4849i \\ 5.6201+2.7592i & -2.7592+5.6201i & 0 \end{pmatrix}\in C_r^{3\times 3}(P_1,Q_1).$$

相应的余项

$$R_{5140501}=\|A_1Z+ZB_1-C_1\|=2.8055\text{e-}4.$$

4　结论

本文利用 HSDM, 给出求解矩阵方程 $A_1Z+ZB_1=C_1$ 广义自反最佳逼近解的迭代算法. 无论方程 $A_1Z+ZB_1=C_1$ 是否相容, 所给的算法都可以计算其广义自反的最佳逼近解. 数值例子也说明了算法的可行性, 但该算法的缺点是收敛速度较慢. 文中广义自反矩阵的集合是凸集, 若其它矩阵方程的未知约束矩阵是凸集, 则可以应用 HSDM 解决其最佳逼近问题.

参 考 文 献

[1] Chen Hsinchu. Generalized reflexive matrices: special properties and applications [J]. SIAM J. Matrix Anal Appl., 1998(19): 140–153.

[2] Chen Hsinchu, Ahmed H SAMEH. Numerical linear algebra algorithms on the cedar system [A]. Noor A K (Ed.). Parallel computations and their impact on mechanics[C]. The American Society of Mechanical Engineers, 1987, 86 : 101–125.

[3] Chen Hsinchu. The SAS domain decomposition method for structural analysis[R]. CSRD Teach, report 754, Center for Supercomputing Research and Development, Urbana, IL: University of Illinois, 1988.

[4] Peng Z Y, Hu X Y. The reflexive and anti-reflexive solutions of matrix equation $AX=B$ [J]. Linear Algebra Appl., 2003 (375): 147–155.

[5] Peng Xiangyang, Hua Xiyan, Zhang Lei. The reflexive and anti-reflexive solutions of the matrix equation [J]. Journal of Computational and Applied Mathematics, 2007 (200): 749–760.

[6] Xu Weiwei, Li Wen. The Hermitian reflexive solutions to the matrix inverse problem [J]. Applied Mathematics and Computation, 2009, (211): 142–147.

[7] 张忠志, 胡锡炎, 张磊. 线性约束下反 Hermitian 广义反 Hermitian 矩阵的最佳逼近问题 [J]. 工程数学学报, 2004, 21 (8): 88–92.

[8] Yamada I. The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings [A]. Butnariu D, Censor Y, Reich S (Eds.). Inherently parallel algorithm for feasibility and optimization and their applications[C]. Elsevier, 2001.

[9] Yamada I, Ogura N, Shirakawa N. A numerically robust hybrid steepest descent method for the convexly constrained generalized inverse problems [J]. Contemporary Mathematics, 2002 (313): 269–305.

AN ITERATIVE ALGORITHM FOR THE GENERALIZED REFLEXIVE OPTIMAL APPROXIMATION SOLUTIONS OF MATRIX EQUATIONS $A_1Z + ZB_1 = C_1$

YANG Jia-wen[1,2] , SUN He-ming[2]

(*1. Department of Basic Courses, Chuzhou Vocational and Technical College, Chuzhou 239000, China*)

(*2. College of Science, Hohai University, Nanjing 210098, China*)

Abstract: In this paper, we present an iterative algorithm to calculate the optimal approximation solutions of the Sylvester complex matrix equations $A_1Z + ZB_1 = C_1$ over generalized reflexive (anti-reflexive) matrices by using the hybrid steepest descent method. Whether matrix equations $A_1Z + ZB_1 = C_1$ are consistent or not, for arbitrary initial reflexive (anti-reflexive) matrix Z_0, the given algorithm can be used to compute the reflexive (anti-reflexive) optimal approximation solutions. The effectiveness of the proposed algorithm is verified by two numerical examples.

Keywords: Sylvester matrix equations; Kronecker product; hybrid steepest descent method; optimal approximation; reflexive matrix

2010 MR Subject Classification: 65F30