文章编号：1674-8085（2023）02-0001-08

Sylvester 矩阵方程的改进 IO 迭代算法

康　靖，*马昌凤

（福建师范大学数学与统计学院，福建，福州 350117）

摘　要：通过对 Sylvester 矩阵方程的理论分析，可知 IO 迭代算法中迭代矩阵的谱半径随内迭代次数的增大而减小，更新了 IO 迭代算法中内迭代次数的选择方法，并证明了该算法收敛性与初始矩阵无关。Sylvester 矩阵在满足一些特定条件下，为了进一步提高收敛速度，可通过选择适当的相关参数，使得 IO 迭代算法有较好的收敛速度且比 Smith 算法的迭代次数明显减少。

关键词：Sylvester 矩阵方程；IO 迭代算法；Smith 算法；收敛性；迭代因子

中图分类号：O224.2　　**文献标识码：**A　　DOI：10.3969/j.issn.1674-8085.2023.02.001

IMPROVED IO ITERATIVE ALGORITHM FOR SYLVESTER MATRIX EQUATION

KANG Jing，*MA Chang-feng

（School of Mathematics and Statistics, Fujian Normal University, Fuzhou, Fujian 350007, China）

Abstract：Through the theoretical analysis of the Sylvester matrix equation, it is known that the spectral radius of the iterative matrix in the IO iteration algorithm decreases with the increase of the number of inner iterations. The selection method of the number of inner iterations in the IO iteration algorithm is updated, and the convergence of the algorithm is proved to be independent of the initial matrix. When the Sylvester matrix satisfies some specific conditions, in order to further improve the convergence speed, by selecting appropriate relevant parameters, the IO iteration algorithm has a better convergence speed and the number of iterations is significantly reduced than that of the Smith algorithm.

Key words：Sylvester matrix equations; IO iteration algorithm; Smith method; convergence; iteration factors

0 引言

将考虑下述的 Sylvester 矩阵方程

$$X - AXB = C \tag{0.1}$$

其中，$A \in \mathbb{R}^{m\times m}$，$B \in \mathbb{R}^{n\times n}$，$C \in \mathbb{R}^{m\times n}$ 是已知矩阵，$X \in \mathbb{R}^{m\times n}$ 为待求矩阵。关于 Sylvester 矩阵方程的求解。

Sylvester 矩阵方程在许多领域都有广泛应用，例如控制理论、数值代数、图像处理等[1-3]方面。近几十年来，取得了一定的研究成果，但还有一些需要改进的问题，寻找更高效的求解方法是当今很多人的研究方向。近几年，许多求解方程的数值实验得到很大发展，例如 Smith 算法[4]，ADI 算法[5]，HSS 型算法[6-7]。在文献[8]中，Tian 等人结合了 Smith 方法中的 IO 迭代算法，提出了 MSIO 算法求解 CLMEs。在文献[9]中，Ke 和 Ma 提出了求解特定 Sylvester 矩阵方程的交替方向法。

收稿日期：2022-04-07；修改日期：2022-05-25
基金项目：国家自然科学基金项目(11901098)；福建省自然科学基金项目(2020J05034).
作者简介：康　靖(1998-)，女，山西平遥人，硕士生，主要从事数值代数方面研究(E-mail:1339048414@qq.com)；
*马昌凤(1962-)，男，湖南隆回人，教授，博士，博士生导师，主要从事数值代数及其应用研究(E-mail:macf@fjnu.edu.cn).

笔者基于 IO 迭代算法的收敛原则，分析了方程的迭代矩阵相关性质，得到即使在迭代因子 $\gamma>1$ 的情况下，可通过选择合适的参数来限制内迭代，使 IO 迭代算法具有更快的收敛速度。

本研究结构如下：第一节提出 IO 迭代算法，并证明了 IO 迭代算法在某些特定条件下的收敛性（其中包含了 $\gamma>1$ 的情况），分析这种情况下，为了加快收敛速度选择迭代因子 l_k 和 γ 的策略；第二节比较在特定条件下，改进 IO 迭代算法和 Smith 算法的收敛速度；第三节给出两个数值实验，比较了改进 IO 迭代算法和 Smith 算法的收敛速度；最后对全文进行总结，并对今后的工作进行展望。

1 IO 迭代算法

在文献[8]中，Tian 等人提出了用 IO 迭代算法求解 Sylvester 矩阵方程，其中 IO 迭代算法可表述为如下：

$$X_{k+1}-\gamma AX_{k+1}B=(1-\gamma)AX_kB+C, k=0,1,2,\cdots \tag{1.1}$$

其中，$0<\gamma<1$ 方程是 IO 迭代算法的外迭代过程。令 $W=(1-\gamma)AX_kB+C$ 且 $Y_0=X_k, Y_{l_k}=X_{k+1}$，可以得到定义如下的内迭代过程：

$$Y_{i+1}=\gamma AY_iB+W, i=0,1,2,\cdots,l_k-1 \tag{1.2}$$

其中，l_k 为方程迭代到第 k 步时的内迭代次数。

算法 1: IO 迭代算法

输入: 将方程(1.1)向量化后输入 $B^T\otimes A=H$，$vec(C)=c$；
初始值 $X_0=O$，$RES=0$，终止条件 ϵ，且记 $vec(X_k)=x_k$；
输入γ的值.

输出: x_k

1: **If** $RES<\epsilon$ **then**

2: 迭代结束

3: **else**

4: $G_{l_k}=(\gamma H)^{l_k}+(1-\gamma)\sum_{i=1}^{l_k-1}(\gamma H)^iH$;

5: $F_{l_k}=\sum_{i=0}^{l_k-1}(\gamma H)^i$.

6: $x_{k+1}=G_{l_k}x_k+F_{l_k}c$;

7: $RES=\|x_{k+1}-x_k\|_2$;

8: $x_k=x_{k+1}$.

9: **end if**

引理 1[4,8] 若 $\rho(A)\rho(B)<1$ 且 $1<\gamma<1$，则由迭代算法产生的序列 $\{x_k\}_{k=0}^{\infty}$ 收敛于方程的精确解，且在文献[8]中比较 IO 迭代算法和 Smith 算法的收敛速度与 IO 迭代算法收敛速度更快。

引理 1 已经给出了 IO 迭代算法在 $0<\gamma<1$ 的情况下的收敛性。接下来，为了讨论 IO 迭代算法在 γ 更大的范围内的收敛性及收敛速度，需要进行一些恰当的假设。

假设 1 设 $\rho(A)\rho(B)<1$，

$$0<\gamma<\frac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)},$$

$$\tau<\frac{1-\rho(A)\rho(B)}{1+\rho(A)\rho(B)}。$$

在假设 1 下，可以通过简单计算得到下述结论：

（1）$\gamma\rho(A)\rho(B)<1$；

（2）$\dfrac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)}>1$；

（3）$\lim\limits_{n\to\infty}(\gamma\rho(A)\rho(B))^n=0$，即 $\exists N\in\mathbb{N}$，当 $\forall n>N$，有 $(\gamma\rho(A)\rho(B))^n<\tau<\dfrac{1-\rho(A)\rho(B)}{1+\rho(A)\rho(B)}$，

因此可取 $N=\left\lceil\dfrac{\ln\dfrac{1-\rho(A)\rho(B)}{1+\rho(A)\rho(B)}}{\ln(\gamma\rho(A)\rho(B))}\right\rceil\doteq\omega$，

其中 $\lceil x\rceil$ 代表对 x 向上取整。

证明 （1）

$$\gamma\rho(A)\rho(B)<\rho(A)\rho(B)\frac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)}<$$

$$\frac{(1-\tau)(1+\rho(A)\rho(B))}{2}<\frac{(1+\rho(A)\rho(B))}{2}<1 \tag{1.3}$$

（2）将 $\tau<\dfrac{1-\rho(A)\rho(B)}{1+\rho(A)\rho(B)}$ 代入可得

$$\frac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)}>$$

$$(1-\frac{1-\rho(A)\rho(B)}{1+\rho(A)\rho(B)})\frac{1+\rho(A)\rho(B)}{2\rho(A)\rho(B)}=$$

$$\frac{2\rho(A)\rho(B)}{1+\rho(A)\rho(B)}\frac{1+\rho(A)\rho(B)}{2\rho(A)\rho(B)}=1 \qquad (1.4)$$

（3）由结论(1)以及结合收敛的定义易得，在此不做赘述。

1.1 收敛性

下述定理给出 IO 迭代算法在假设 1（其中包含了 $\gamma>1$ 的情况)下的收敛性的详细证明。

定理 1 若方程满足假设 1 且 $l_k>\omega$，则 $\rho(G_{l_k})<1$。

证明 由算法 1 的第 4 行可知，

$$G_{l_k}=(\gamma H)^{l_k}+(1-\gamma)\sum_{i=1}^{l_k-1}(\gamma H)^i H$$

设 λ_i 为 H 的特征值，则有

$$\phi_i^{(l_k)}=(\gamma\lambda_i)^{l_k}+(1-\gamma)\sum_{j=0}^{l_k-1}(\gamma\lambda_i)^{l_k}\lambda_i=$$

$$1-\frac{(1-\lambda_i)(1-(\gamma\lambda_i)^{l_k})}{1-\gamma\lambda_i} \qquad (1.5)$$

又因为 $\rho(H)=\rho(B^T\otimes A)=\rho(A)\rho(B)$，则

$$|\phi_i^{(l_k)}|=|1-\frac{(1-\lambda_i)(1-(\gamma\lambda_i)^{l_k})}{1-\gamma\lambda_i}|=$$

$$|\frac{\lambda_i(1-\gamma+\gamma^{l_k}\lambda_i^{l_k-1}-(\gamma\lambda_i)^{l_k})}{1-\gamma\lambda_i}|<$$

$$\frac{|1-\gamma|\ |\lambda_i|+|\gamma\lambda_i|^{l_k}|1-\lambda_i|}{|1-\gamma\lambda_i|}\le$$

$$\frac{|1-\gamma|\ \rho(H)+(\gamma\rho(H))^{l_k}(1+\rho(H))}{1-\gamma\rho(H)}=$$

$$\frac{|1-\gamma|\ \rho(A)\rho(B)+(\gamma\rho(A)\rho(B))^{l_k}(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)} \qquad (1.6)$$

由假设 1 的分析可得，$\gamma\rho(A)\rho(B)<1$，且当 $l_k>\omega$ 时，有 $(\gamma\rho(A)\rho(B))^{l_k}<\tau$，则根据 γ 与 1 的大小关系分类考虑有下述两种情况：

当 $0<\gamma<1$ 时有

$$|\phi_i^{(l_k)}|\le$$

$$\frac{|1-\gamma|\ \rho(A)\rho(B)+(\gamma\rho(A)\rho(B))^{l_k}(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}=$$

$$\frac{(1-\gamma)\rho(A)\rho(B)+(\gamma\rho(A)\rho(B))^{l_k}(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}<$$

$$\frac{(1-\gamma)\rho(A)\rho(B)+\tau(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}<$$

$$\frac{(1-\gamma)\rho(A)\rho(B)+1-\rho(A)\rho(B)}{1-\gamma\rho(A)\rho(B)}=$$

$$\frac{1-\gamma\rho(A)\rho(B)}{1-\gamma\rho(A)\rho(B)}=1. \qquad (1.7)$$

当 $1<\gamma<\frac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)}$ 时有

$$|\phi_i^{(l_k)}|\le$$

$$\frac{|1-\gamma|\ \rho(A)\rho(B)+(\gamma\rho(A)\rho(B))^{l_k}(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}=$$

$$\frac{(\gamma-1)\rho(A)\rho(B)+(\gamma\rho(A)\rho(B))^{l_k}(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}<$$

$$\frac{(\frac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)}-1)\rho(A)\rho(B)+\tau(1+\rho(A)\rho(B))}{1-\frac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)}\rho(A)\rho(B)}=1 \qquad (1.8)$$

综上所述，证得 $\rho(G_{l_k})<1$，IO 迭代算法是收敛的。

1.2 分析参数 l_k 和γ的选择策略

定理 2 若方程满足假设 1 且 $l_k>\omega$，则 l_k 越大 IO 迭代算法收敛速度越快。

证明 由定理 1 中的(1.5)可知，

$$\phi_i^{(l_k)}=1-\frac{(1-\lambda_i)(1-(\gamma\lambda_i)^{l_k})}{1-\gamma\lambda_i}$$

而 H 的特征值 $\lambda_i<1$，且由(1.3)可知 $\gamma\lambda_i<\gamma\rho(A)\cdot\rho(B)<1$。

令 $h(l_k)=1-\frac{(1-\lambda_i)(1-(\gamma\lambda_i)^{l_k})}{1-\gamma\lambda_i}$，对 $h(l_k)$ 关于 l_k 求导后有

$$h'(l_k)=\frac{\ln(\gamma\lambda_i)(1-\lambda_i)(\gamma\lambda_i)^{l_k}}{1-\gamma\lambda_i}. \tag{1.9}$$

显然，$h'(l_k)<0$。所以，$h(l_k)$是关于l_k的递减函数，即$\rho(G_{l_k})$随着内迭代次数l_k的增大而减小，收敛速度更快。

定理 3 若方程满足假设 1 且 $l_k>\omega$，则$\gamma\to 1$时 IO 迭代算法收敛效果比较好。

证明 由定理 1 中的(1.6)可知，

$$|\phi_i^{(l_k)}|<\frac{|1-\gamma|\ \rho(A)\rho(B)+(\gamma\rho(A)\rho(B))^{l_k}(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)},$$

又因为$l_k>\omega$，有$(\gamma\rho(A)\rho(B))^{l_k}<\tau$，代入上式可以得到

$$|\phi_i^{(l_k)}|<\frac{|1-\gamma|\ \rho(A)\rho(B)+\tau(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)} \tag{1.10}$$

记$\gamma\rho(A)\rho(B)=\sigma$，令$f(\gamma)=\dfrac{|1-\gamma|\ \rho(A)\rho(B)+\tau(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}=$

$$\begin{cases}\dfrac{(1-\gamma)\rho(A)\rho(B)+\tau(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}, & 0<\gamma<1,\\ \dfrac{(\gamma-1)\rho(A)\rho(B)+\tau(1+\rho(A)\rho(B))}{1-\gamma\rho(A)\rho(B)}, & 1<\gamma<\dfrac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)},\end{cases}$$

$$\begin{cases}\dfrac{(1-\gamma)\sigma+\tau(1+\sigma)}{1-\gamma\sigma}, & 0<\gamma<1,\\ \dfrac{(\gamma-1)\sigma+\tau(1+\sigma)}{1-\gamma\sigma}, & 1<\gamma<\dfrac{(1-\tau)(1+\sigma)}{2\sigma}.\end{cases} \tag{1.11}$$

对(1.11)式求导进行化简后有

$$f(\gamma)=\begin{cases}\dfrac{-\sigma(1-\gamma\sigma)-[(1-\gamma)\sigma+\tau(1+\sigma)](-\sigma)}{(1-\gamma\sigma)^2}, & 0<\gamma<1,\\ \dfrac{\sigma(1-\gamma\sigma)-[(\gamma-1)\sigma+\tau(1+\sigma)](-\sigma)}{(1-\gamma\sigma)^2}, & 1<\gamma<\dfrac{(1-\tau)(1+\sigma)}{2\sigma},\end{cases}=$$

$$\begin{cases}\dfrac{\sigma^2-\sigma+\tau\sigma(1+\sigma)}{(1-\gamma\sigma)^2}, & 0<\gamma<1,\\ \dfrac{\sigma-\sigma^2+\tau\sigma(1+\sigma)}{(1-\gamma\sigma)^2}, & 1<\gamma<\dfrac{(1-\tau)(1+\sigma)}{2\sigma}.\end{cases} \tag{1.12}$$

由上式分析得，在$0<\gamma<1$时，$f'(\gamma)<0$，即$f(\gamma)$关于γ递减；在$1<\gamma<\dfrac{(1-\tau)(1+\sigma)}{2\sigma}$时，$f'(\gamma)>0$，即$f(\gamma)$关于$\gamma$递增；在$\gamma=1$处$f(\gamma)$取到最小值。又因为在$\gamma=1$时，只是$f(\gamma)$取到了上界的最小值，而不是直接关于谱半径$\rho(G_{l_k})$的，所以 IO 迭代算法在$\gamma\to 1$时有较好的收敛效果。

2 收敛速度

在假设 1 下，相关参数l_k遵循最优选择策略，

γ 满足假设 1，分析改进 IO 迭代算法收敛速度，进一步判断是否比 Smith 方法收敛更快。现将 Smith 算法做简单引入，具体可参考文献[4]。

对方程(0.1)两边向量化展开得

$$x=(I-B^T\otimes A)^{-1}c=$$

$$\sum_{k=1}^{\infty}((B^T)^{k-1}\otimes A^{k-1})c=\sum_{k=1}^{\infty}H^{k-1}c \tag{2.1}$$

算法 2: Smith 算法

输入: 将方程（1.1）向量化后输入 $B^T\otimes A=H$ ， $vec(C)=c$ ；初始值 $X_0=O$，$RES=0$，终止条件 ϵ ，且记 $vec(X_k)=x_k$ ；

输出: x_k

1: **If** $RES<\epsilon$ **then**

2: 迭代结束

3: **else**

4: $x_{k+1}=x_k+H^kc$;

5: $RES=\|x_{k+1}-x_k\|_2$;

6: $x_k=x_{k+1}$;

7: **end if**

由(2.1)可得 Smith 算法的迭代矩阵为 H，因此要证明 IO 迭代算法比 Smith 算法收敛速度更快，只需证明 $\rho(G_{l_k})<\rho(H)$ 即可。为了证明方便，先引入一个关于谱半径的不等式。

引理 2[7] 令 $\rho(A)$ 为矩阵 A 的谱半径，对于任意的 A 和任意的 $\mu>0$，存在算子范数 $\|\cdot\|_\zeta$ 满足

$$\rho(A)\le\|A\|_\zeta\le\rho(A)+\mu .$$

定理 4 若假设方程的系数矩阵 A,B 满足 $\rho(A)\rho(B)<1$， $0<\gamma<\dfrac{(1-\tau)(1+\rho(A)\rho(B))}{2\rho(A)\rho(B)}$ 且 $\left(\dfrac{1-\rho(A)\rho(B)}{1+\rho(A)\rho(B)}\right)^2<\tau<\dfrac{1-\rho(A)\rho(B)}{1+\rho(A)\rho(B)}$。则对于任意的 $\mu>0$，存在 $p_k\in\mathbb{N}$，当 $l_k>p_k$ 时，IO 迭代算法比 Smith 算法收敛速度更快。

证明 由算法 1 的第 4 行可得，

$$G_{l_k}=(\gamma H)^{l_k}+(1-\gamma)\sum_{i=1}^{l_k-1}(\gamma H)^iH .$$

令

$$M=1-\gamma H,N=(1-\gamma)H,R=\gamma H. \tag{2.2}$$

代入迭代矩阵 G_{lk} 中得到

$$G_{l_k}=R^{l_k}+(I-R^{l_k})M^{-1}N=$$

$$I-(I-R^{l_k})(I-M^{-1}N). \tag{2.3}$$

由假设条件可知，$\gamma\rho(A)\rho(B)<1$，可以得到 $\lim\limits_{l_k\to\infty}(\gamma H)^{l_k}=O$ 。则存在 $p_k\in\mathbb{N}$，满足当 $l_k>p_k$ 时，有

$$\|R^{l_k}\|_\zeta=\|(\gamma H)^{l_k}\|_\zeta<\mu \tag{2.4}$$

结合 τ 的范围，

$$\rho(M^{-1}N)=\rho[(1-\gamma H)^{-1}(1-\gamma)H]=$$

$$\frac{|1-\gamma|\ \rho(A)\rho(B)}{1-\gamma\rho(A)\rho(B)}<$$

$$\rho(A)\rho(B)<1. \tag{2.5}$$

由引理 2 得，存在算子范数 $\|\cdot\|_\zeta$ 和 $\mu>0$，满足

$$\|M^{-1}N\|_\zeta\le\rho(M^{-1}N)+\mu<\rho(A)\rho(B)+\mu \tag{2.6}$$

将不等式和结合关于迭代矩阵 G_{l_k} 的(2.3)关系式中，则有：

$$\|G_{l_k}\|_\zeta=\|R^{l_k}+(I-R^{l_k})(M^{-1}N)\|_\zeta\le$$

$$\|R^{l_k}\|_\zeta+(1+\|R^{l_k}\|_\zeta)\|M^{-1}N\|_\zeta<$$

$$\mu+(1+\mu)(\rho(A)\rho(B)+\mu)\triangleq g(\mu) \tag{2.7}$$

由于不等式(2.7)是对于任意的 $\mu>0$ 成立，则由 $\lim\limits_{\mu\to0}g(\mu)=\rho(A)\rho(B)=\rho(H)$ ，可以证得 IO 迭代算法比 Smith 算法收敛速度更快。

3 数值实验

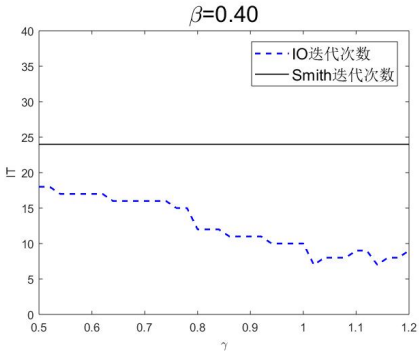
下面用两个数值例子来说明所提出的改进 IO 迭代算法的收敛性能，以及对相关参数 l_k 和 γ 选择策略的可视化。两种算法终止条件为 $RES<10^{-7}$，数值实验是在 MATLAB-R2017a 软件运行。

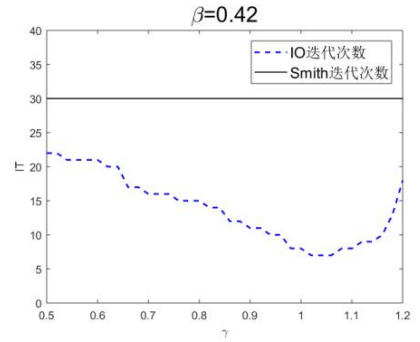
例 1 针对求解 Lyapunov 矩阵方程 $X-AXA^T=C$，其中

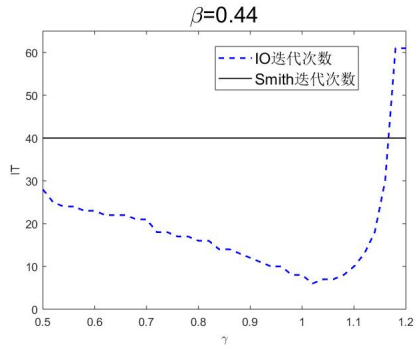
$$A=\begin{pmatrix} 0 & \beta & & & & \\ -\beta & 0 & \beta & & & \\ & -\beta & 0 & \beta & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \beta \\ & & & & -\beta & 0 \end{pmatrix},$$

$$C=\begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

在这个例子中，我们取 $n=15$。由图 1 分析得到当 $\beta\to 0.5$，$\gamma\in(0.5,1.2)$ 有 $\rho(A)\to 1$。对于 β 取一些不同值的情况时，分析所需迭代次数，可知 $\gamma>1$ 时也是收敛的，且 $\gamma\to 1$ 有较好的收敛速度，IO 迭代算法比 Smith 方法收敛速度整体要快一些。







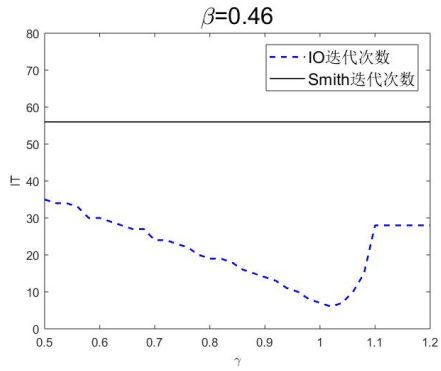


图 1 对于不同的 γ 和 β 的 IO 迭代算法和 Smith 算法迭代次数

Fig.1 The number of iterations of IO iterative method and Smith method for different γ and β

由定理 1 可知，l_k 越大，收敛速度越快，不同的 γ 会导致迭代矩阵的谱半径不同，从而 l_k 的最佳选择不同，详细数据在表 1 中体现。随着 γ 越大，为了更快的收敛，可以取更合适的 l_k，从而使得迭代次数最少。

表 1 不同的 β 下的 l_k 和迭代次数 IT

Table 1 l_k and iteration times IT for different β

IO 迭代算法	$\beta=0.40$		$\beta=0.42$		$\beta=0.44$		$\beta=0.46$	
γ	l_k	IT	l_k	IT	l_k	IT	l_k	IT
0.5	2	18	2	22	2	28	3	35
0.6	2	17	2	21	3	23	4	30
0.7	2	16	3	16	3	21	5	24
0.8	3	12	3	15	4	16	6	19
0.9	3	11	4	11	5	12	8	14
1.0	3	10	5	8	7	8	12	7
1.1	4	9	6	8	10	10	21	28
1.2	5	9	9	18	18	61	99	28
Smith IT	24		30		40		56	

由表 1 分析可得，l_k 的适当选取会明显降低总的迭代次数。

例 2 针对求解方程 $X-AXB=C$，其中

$$A=\begin{pmatrix} D & -\beta I & & & & \\ -\beta I & D & -\beta I & & & \\ & -\beta I & D & -\beta I & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & -\beta I \\ & & & & -\beta I & D \end{pmatrix}_{pq\times pq},$$

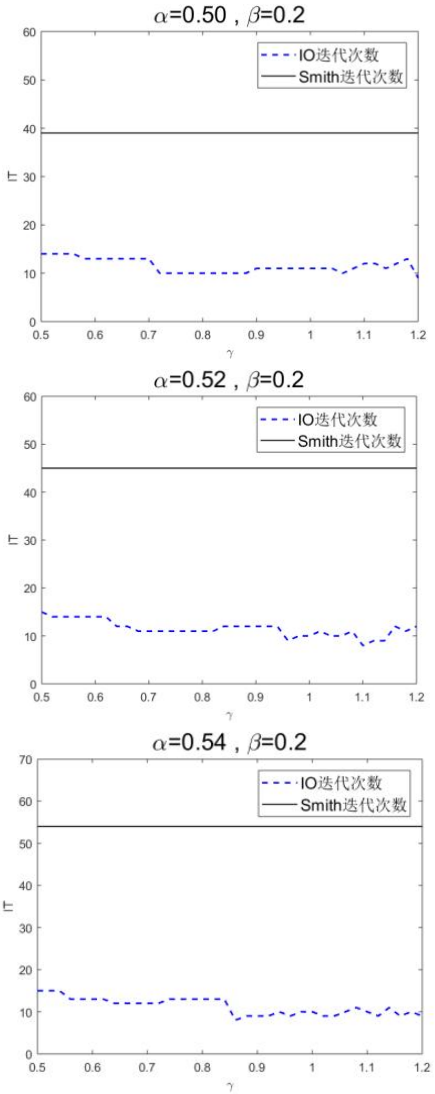
$$D=\begin{pmatrix} 0 & \alpha & & & & \\ \alpha & 0 & \alpha & & & \\ & \alpha & 0 & \alpha & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \alpha \\ & & & & \alpha & 0 \end{pmatrix}_{p\times p}$$

其中

$$B=A^T, C=I_{pq\times pq}, \beta=0.2,$$

$$\alpha=0.50,0.52,0.54,0.56 。$$

显然 A,B 为非负矩阵。具体实验中我们取 $q=3,p=10,\beta=0.2$，α 不同值的情况。从图 2 和表 2 中可以看出，随着 γ 的变化，l_k 也在变化，对应的最佳 l_k 也随之不同，且 IO 迭代算法比 Smith 算法迭代次数少很多。



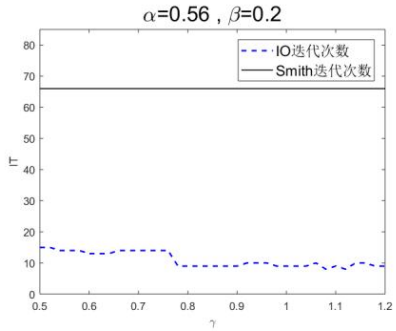


图 2 对于不同的 γ 和 α 的 IO 迭代算法和 Smith 算法迭代次数

Fig. 2 The number of iterations of IO iterative method and Smith method for different γ and α

表 2 不同的β下的 l_k 和迭代次数 IT

Table 2 l_k and iteration times IT for different β

IO 迭代算法	$B=0.40$		$B=0.42$		$B=0.44$		$B=0.46$	
γ	l_k	IT	l_k	IT	l_k	IT	l_k	IT
0.5	2	14	2	15	2	15	3	15
0.6	2	13	2	14	3	13	3	13
0.7	2	13	3	11	3	12	4	14
0.8	3	10	3	11	4	13	5	9
0.9	3	11	4	12	5	9	6	9
1.0	4	11	5	10	6	10	9	9
1.1	5	12	7	8	9	10	13	9
1.2	7	9	9	12	15	9	29	9
Smith IT	39		45		54		66	

4 结论

通过对 IO 迭代算法进行改进来求解 Sylvester 矩阵方程，结果表明：

1）IO 迭代算法比 Smith 方法迭代次数要少，收敛速度快，所改进的方法更加明显减少了迭代次数。

2）迭代因子 γ 的选择会影响所改进方法的迭代速度，在例 1 中有比较明显的体现。

3）虽然改进的方法加速了迭代速度，但对 Sylvester 矩阵方程有一定的特殊要求，具有一定的局限性。

4）未来希望在参数选取上有所突破，在两个参数选取之间的关系是否更密切。

5）可以为求解更一般的 Sylvester 矩阵方程提供了新的研究方向。

参考文献：

[1] Datta B N. Numerical methods for linear control systems design and analysis[M].Amsterdam:Elsevier Press,2003.

[2] Anderson B, Agathoklis P, Jury E. Stability and the matrix Lyapunov equation for discrete 2-dimensional systems[J]. EEE Transactions on Circuits and Systems, 1986, 33(3): 261-267.

[3] Liu B,Fu D,Qi M, et al. Noise-tolerant gradient-oriented neurodynamic model for solving the Sylvester equation[J]. Applied Soft Computing, 2021,109:107514.

[4] Smith R A. Matrix equation $XA+BX=C$.SIAM[J].Appl.Math., 1968,16(1):198-201.

[5] Calvetti D, Reichel L. Application of ADI iterative methods to the restoration of noisy images,SIAM [J]. Matrix Anal. Appl.,1996,17(1):165-186.

[6] Bai Z. On Hermitian and skew-Hermitian splitting iteration methods for continuous Sylvester equations[J]. Comput. Math.,2011,29(2):185-198.

[7] Zheng Q, Ma C. On normal and skew-Hermitian splitting iteration methods for large sparse continuous Sylvester equations[J].Comput. Appl. Math., 2014,268(1):145-154.

[8] Tian Z, Wang J, Dong Y, et al. A multi-step smith-inner-outer iteration algorithm for solving coupled continuous Markovian jump Lyapunov matrix equations[J]. Franklin. Inst., 2020,357:3656-3680.

[9] Ke Y, Ma C. Alternating direction method for a class of Sylvester matrix equations with linear matrix inequality constraint[J]. Numerical Functional Analysis and Optimization, 2018,39(3):257-275.