学校编码：10384　　　　分类号 ______ 密级 ______
学号：19020071152084　　　　UDC ______

厦门大学

硕　士　学　位　论　文

Sylvester方程的不平衡ADI因子迭代方法

Factor Imbalance ADI Iteration Method For Sylvester Equation

林　一　丁

指导教师姓名：卢 琳 璋 教授
专 业 名 称：计 算 数 学

论文提交日期：2010 年 5 月
论文答辩日期：2010 年 　 月
学位授予日期：2010 年 　 月

答辩委员会主席：____________
评　　阅　　人：____________

2010 年 5 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　　　　　　　　　　　）课题（组）的研究成果，获得（　　　　　　）课题（组）经费或实验室的资助，在（　　　　　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　）1. 经厦门大学保密委员会审查核定的保密学位论文，于　年　月　日解密，解密后适用上述授权。

（　　）2. 不保密，适用上述授权。

（请在以上相应括号内打"√"或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

目　录

Contents

摘　　要

本文讨论如下形式的低秩 Sylvester 方程

$$AX - XB = GF^* \tag{0.1}$$

和广义低秩 Sylvester 方程

$$A_1XB_1 + A_2XB_2 = GF^* \tag{0.2}$$

的迭代法，提出了一个不平衡交替方向因子迭代 (FIADI) 算法，它是 FADI [7] 的一种变形. 考虑到 Sylvester 方程 (0.1) 中 A 和 B 的阶数可能不同，但在交替方向迭代中的地位相同，尤其是在计算中对 A 和 B 需要进行相同数量的位移求逆运算. 本文提出了 FIADI, 它打破了这种平等，通过对阶数小的矩阵多求逆，而对阶数大的矩阵少求逆这种方式来减少迭代整体的计算量. 本文还总结了能利用 FADI 来求解的低秩 Sylvester 方程一般形式： $X + \tilde{A}X\tilde{B} = \tilde{G}\tilde{F}^*$, 就可以利用右端的低秩性，拆成两个方向来解. 通过分析，可知广义低秩 Sylvester 方程 (0.2) 实际具有这种形式，因此可以用 FADI 和本文提出的办法来解.

关键词： Sylvester 方程， 交替方向迭代 (ADI)，不平衡交替方向因子迭代 (FIADI) , Lyapunov 方程， 广义 Sylvester 方程.

Abstract

This paper is concerned with the numerical solution of Sylvester equations

$$AX - XB = GF^* \tag{0.1}$$

and generalized Sylvester equations

$$A_1XB_1 + A_2XB_2 = GF^*. \tag{0.2}$$

In this paper, we present Factor Imbalance Alternating-Directional-Implicit (FIADI) iteration for Sylvester equations(0.1), which is a variation of FADI [7]. Considering the different order of A and B but the same role in FADI, we present FIADI iteration which breaks the balance. The method reduces the overall computation by inverting the matrix of small order more and inverting the other less. Furthermore, we discuss the form of generalized Sylvester equations, which can take advantage of FADI and FIADI. They can usually be transformed into the form of $X - \tilde{A}X\tilde{B} = \tilde{G}\tilde{F}^*$. By analyzing, we find generalized Sylvester equations (0.2) has this form. Therefore, FADI and FIADI can be applied to generalized Sylvester equation (0.2).

Key words: Sylvester equation, ADI, FIADI , Lyapunov equation, generalized Sylvester equation.

第一章 绪 论

本文讨论如下形式的低秩 Sylvester 方程

$$\begin{aligned} AX - XB &= GF^*, \\ A_1XB_1 + A_2XB_2 &= GF^*, \end{aligned} \tag{1.1}$$

其中 A, A_1 和 A_2 是 $m\times m$ 的矩阵，B, B_1 和 B_2 是 $n\times n$ 的矩阵，G 是 $m\times r$ 的矩阵，F 是 $n\times r$ 的矩阵，$r = rank(GF^*), r \ll min(m, n)$, 解 X 是 $m\times n$ 的矩阵. 该方程的特殊情况是 Lyapunov 方程

$$AX + XA^T = -GG^T, \tag{1.2}$$

其中的 A 是 $n\times n$ 的矩阵，且该矩阵的特征值都位于左半平面，即 $re(\lambda_i(A)) < 0$, G 是 $n\times r$ 的矩阵，$r \ll n$,X 是 $n\times n$ 的矩阵.

Sylvester 方程 $AX - XB = C(X \in \mathbb{C}^{m\times n})$ 在应用数学中有广泛的应用，例如矩阵的特征问题 [15], 控制论 [12, 22, 29], 模型降阶 [1, 2, 6, 32], Riccati 方程的求解 [11, 13] 以及图像处理 [10]. 同时在理论上，关于 Sylvester 方程及其特殊情况 Lyapunov 方程已经有了很多结论 [9, 16].

求解 Sylvester 方程已经有了很多办法，较早的办法是 R. H. Bartels 和 G. W. Stewart [5] 和 G. H. Golub, S. Nash 和 C. F. Van Loan [14], 他们算法的关键是通过正交变换把 B 变成拟上三角的形式，对 A 进行的变换只需变成近似上三角即可，这时展开成 Kronecker 积的形式就可以发现他们的算法实际就是块的向前算法解了一个近似块下三角的方程，一列一列的得到解. 他们的算法对中小型稠密矩阵 A 和 B 很有效，而对大型稀疏的矩阵，要通过正交变换把 A 和 B 变成近似上三角阵会破坏稀疏性，因此不适合大规模稀疏问题.

对大型稀疏的 Sylvester 方程，Krylov 子空间的投影类办法 [3, 4, 17, 18, 19, 28, 27, 30, 31] 很有效，它们的基础是第二章定理 3: Sylvester 方程的解能写成式子：$X = \int_0^\infty e^{-tA}Ce^{tB}dt$, 用 $\mathbb{P}(A)$ 近似 e^{-tA} 和用 $\mathbb{Q}(B)$ 近似 e^{tB}(其中 $\mathbb{P}(x), \mathbb{Q}(x)$

可以是多项式，也可以是分式) 就可以得到近似解. 它们的好处是不用知道关于 A 和 B 的谱的任何信息，主要问题是对右端低秩的 Sylvester 方程没办法重启.

现在所提出的方程 [10, 22] 除了系数矩阵 A 和 B 大规模稀疏外，还有右端项 C 低秩的性质. 交替方向隐式迭代 (ADI)[19, 20, 21, 23, 24, 26] 很好的利用了右端的低秩性质，当 A 和 B 的谱的性质较好时 (指 A 和 B 的谱各自集中但整体距离较远)，比较容易取得最 (次) 优的 ADI 位移参数 [24, 26], 这时收敛速度比投影法快. 在用交替方向隐式迭代求解 Lyapunov 方程时，J.-R. Li 和 J. White [20] 通过倒序参数的办法，把计算量降了一个数量级，可谓经典算法. Peter Benner,Ren-Cang Li 和 Ninoslav Truhar [7] 在这之后不久便将该算法把它推广到 Sylvester 方程. 交替方向迭代的主要问题是需要选择参数使迭代矩阵的谱半径尽可能的小，而如何选择好的参数是个很复杂的问题，它需要知道 A 和 B 谱的信息，然后求一个极小极大问题，对一般的系数矩阵 A 和 B 还没有比较好的办法. 而且还有个很严重的问题，有些矩阵本身就很难找到参数使交替方向迭代的迭代矩阵谱半径很小.

这篇文章所提出的办法是 FADI [7] 的一种变形，主要想法是 Sylvester 方程的两个系数矩阵 A 和 B 阶数可能不同，但在交替方向迭代中处于相同的地位，要求相同数量的逆. 本文提出的算法打破了这种平衡，通过对阶数小的矩阵多求逆，而对阶数大的矩阵少求逆来减少整体的计算量，并通过数值实验验证这种办法的可能性和有效性. 文章还总结了能用 FADI 来求解的广义低秩 Sylvester 方程形式，它们一般都能简化成 $X + \tilde{A}X\tilde{B} = \tilde{G}\tilde{F}^*$ 的形式，再利用了右端的低秩性，用两个循环就可以算出解的因子. 广义低秩 Sylvester 方程 $A_1XB_1 + A_2XB_2 = GF^*$ 实际能简化成这种形式，可以用 FADI 和本文提出的办法来解.

全文共分五个部分：第一章简要的介绍了 Sylvester 方程及其背景；第二章介绍几类 Sylvester 方程解的存在唯一性和解的形式；第三章给出了 Peter Benner 的因子交替方向 (FADI) 迭代算法 [7] 和适用 FADI 的广义低秩 Sylvester 方程的形式；第四章仿造 Peter Benner 的过程推导出了本文的算法，并用例子说明；第五章把算

法应用在广义低秩 Sylvester 方程 $A_1XB_1 + A_2XB_2 = GF^*$, 并用例子说明.

记号说明：$A \in \mathbb{C}^{m\times m}$ 指 A 是复数域上的 $m \times m$ 矩阵. F^* 指 F 的共轭转置. G^T 指 G 的转置. $r = rank(C)$ 指 C 的秩是 r. $\otimes$ 指 Kronecker 积. 若 $C \in \mathbb{C}^{m\times n}$ 按列分块为 $[c_1, c_2, \cdots, c_n]$, $vecC$ 指 C 按列拉直, 即 $(vecC)^T = ({c_1}^T, {c_2}^T, \cdots, {c_n}^T)$. $eig(A)$ 指 A 的所有特征值的集合. $\lambda_i(A), i = 1, 2, \ldots, n$ 指 A 的按某种顺序排列的特征值. $re(\lambda(A))$ 指 A 的特征值实部. $\rho(A)$ 指 A 的谱半径. $\|A\|_\infty$ 指 A 的无穷范数. $\|C\|_F$ 指 C 的 F- 范数.

第二章 预备知识

§2.1 解的存在唯一性

关于 Sylvester 方程的解的存在唯一性有如下结论：

定理 1. *[16]* 矩阵方程 $AX - XB = C$,$A \in \mathbb{C}^{m\times m}$,$B \in \mathbb{C}^{n\times n}$,$C \in \mathbb{C}^{m\times n}$ 有唯一解 $X \in \mathbb{C}^{m\times n}$ 的充分必要条件是 A 和 B 没有相同的特征值，即

$$\lambda_i - \mu_j \neq 0,\ i = 1, \ldots, m,\ j = 1, \ldots, n.$$

证明 将矩阵方程两端用矩阵的拉直运算操作后得到：

$$(I_n \otimes A - B^T \otimes I_m)vecX = vecC.$$

该方程有唯一解的充分必要条件是系数矩阵 $(I_n \otimes A - B^T \otimes I_m)$ 没有零特征值.

如果设 A 的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_m$, B(或 B^T) 的特征值为 $\mu_1, \mu_2, \ldots, \mu_n$, 那么 $(I_n \otimes A - B^T \otimes I_m)$ 的特征值为 $\lambda_i - \mu_j (i = 1, \ldots, m,\ j = 1, \ldots, n)$, 于是原矩阵方程有唯一解的充分必要条件是 $\lambda_i - \mu_j \neq 0$, 即 A 与 B 没有相同的特征值. □

对于 Lyapunov 方程 $AX + XA^T = -GG^T$, 当 A 的特征值实部均小于零时，A 与 $-A^T$ 必没有相同的特征值，此时这个低秩 Lyapunov 方程必有唯一解.

同样用拉直运算可以得到：

定理 2. 矩阵方程 $X - AXB = C$,$A \in \mathbb{C}^{m\times m}$,$B \in \mathbb{C}^{n\times n}$,$C \in \mathbb{C}^{m\times n}$ 有唯一解 $X \in \mathbb{C}^{m\times n}$ 的充分必要条件是

$$\lambda_i \mu_j \neq 1,\ i = 1, \ldots, m,\ j = 1, \ldots, n.$$

其中 λ_i, μ_j 分别是 A 和 B 的特征值.

关于 Sylvester 方程 $AX - XB = C$ 精确解的形式有：

定理 3. *[16]* 如果 $A \in \mathbb{C}^{m\times m}$,$B \in \mathbb{C}^{n\times n}$,$C \in \mathbb{C}^{m\times n}$ 满足

$$\lim_{t\to\infty} e^{-tA}Ce^{tB}dt = 0, \text{且} \int_0^\infty e^{-tA}Ce^{tB}dt\text{收敛}, \tag{2.1}$$

那么矩阵方程 $AX - XB = C$ 的精确解为

$$X = \int_0^\infty e^{-tA}Ce^{tB}dt. \tag{2.2}$$

证明 考虑常微分方程

$$\begin{aligned} \frac{d}{dt}Y(t) &= -AY(t) + Y(t)B, \\ Y(0) &= C, \end{aligned} \tag{2.3}$$

其中 $Y(t)$ 是 $m\times n$ 的函数矩阵， A 是 $m\times m$ 矩阵， B 是 $n\times n$ 的矩阵， C 是 $m\times n$ 的矩阵. 利用 Kronecker 积拉直可以得到：

$$\begin{aligned} \frac{d}{dt}vecY(t) &= (-I_n \otimes A + B^T \otimes I_m)vecY(t), \\ vecY(0) &= vecC. \end{aligned} \tag{2.4}$$

容易得到这个线性常微分方程组的解：

$$vecY(t) = e^{t(-I_n\otimes A+B^T\otimes I_m)}vecY(0) = (e^{tB^T} \otimes e^{-tA})vecY(0) = (e^{tB^T} \otimes e^{-tA})vecC$$

把解还原为矩阵形式得到：

$$Y(t) = e^{-tA}Y(0)e^{tB} = e^{-tA}Ce^{tB}.$$

另一方面，如果：

$$\begin{gathered} \lim_{t\to\infty} e^{-tA}Ce^{tB} = 0(\text{即} \lim_{t\to\infty} Y(t) = 0), \\ \text{且} \int_0^\infty e^{-tA}Ce^{tB}dt\text{收敛} \end{gathered} \tag{2.5}$$

那么对原来的常微分方程组两边从 0 到 ∞ 积分可得：

$$-C = Y(\infty) - Y(0) = -A\int_0^\infty e^{-tA}Ce^{tB}dt + \int_0^\infty e^{-tA}Ce^{tB}dtB,$$

由此可知 $X=\int_0^\infty e^{-tA}Ce^{tB}dt$ 是 $AX-XB=C$ 的精确解. □

这个定理是投影法的基础，由上面可以看出 X 的列空间是由 A 和 C 生成的 Krylov 子空间，而 X 的行空间是由 B 和 C 生成的 Krylov 子空间. 投影法实际相当于用 $\mathbb{P}(A)$ 近似 e^{-tA} 和用 $\mathbb{Q}(B)$ 近似 e^{tB}(其中 $\mathbb{P}(x),\mathbb{Q}(x)$ 可以是多项式，也可以是分式).

对于 Lyapunov 方程 $AX+XA^T=-GG^T$ 当 $re(\lambda_i(A))<0, i=1,2,\ldots,n$ 时，容易验证:

$$\lim_{t\to\infty} e^{tA}GG^Te^{tA^T}dt=0, \text{且} \int_0^\infty e^{tA}GG^Te^{tA^T}dt \text{收敛} \tag{2.6}$$

由上面的定理， Lyapunov 方程的精确解为 $X=\int_0^\infty e^{tA}GG^Te^{tA^T}dt$.

§2.2 几类 Sylvester 方程的级数形式的精确解

定理 4. *[8] 如果存在正数 r 使 $A\in\mathbb{C}^{m\times m}$,$B\in\mathbb{C}^{n\times n}$ 满足*

$$eig(B)\subseteq\{z\in\mathbb{C}:|z|<r\}, \text{且} eig(A)\subseteq\{z\in\mathbb{C}:|z|>r\}, \tag{2.7}$$

则关于 X 的方程 $AX-XB=C$ 的解为

$$X=\sum_{k=0}^{\infty}A^{-k-1}CB^k.$$

证明 只需证明这个级数收敛，然后容易验证 X 是方程的解.

选取 $r_1<r<r_2$ 使得

$$eig(B)\subseteq\{z\in\mathbb{C}:|z|<r_1\}, eig(A)\cap\{z\in\mathbb{C}:|z|\le r_2\}=\emptyset.$$

则 $eig(A^{-1})\subseteq\{z\in\mathbb{C}:|z|<r_2^{-1}\}$.

根据 Gelfand 谱半径公式 $\rho(A)=\lim_{k\to\infty}\|A^k\|_\infty^{\frac{1}{k}}$, 存在正整数 N 使得当 $k\ge N$ 时

$$\|B^k\|_\infty\le r_1^k, \|A^{-k}\|_\infty\le r_2^{-k}.$$

因次，对于 $k \geq N$,

$$\|A^{-k-1}CB^k\|_\infty \leq (\frac{r_1}{r_2})^k \|A^{-1}C\|_\infty.$$

所以解的级数收敛. □

类似的有如下定理：

定理 5. 如果存在正数 r 使 $A \in \mathbb{C}^{m\times m}$,$B \in \mathbb{C}^{n\times n}$ 满足

$$eig(B) \subseteq \{z \in \mathbb{C} : |z| > r\}, \text{且} eig(A) \subseteq \{z \in \mathbb{C} : |z| < r\}, \tag{2.8}$$

则关于 X 的方程 $AX - XB = C$ 的解为：

$$X = -\sum_{k=0}^{\infty} A^k C B^{-k-1}.$$

在上述两个定理中解都表达成级数形式，但有个非常重要的区别是它们乘法和求逆的对象不一样. 只要条件允许，在计算解时我们自然希望是对阶数低的矩阵求逆，而对阶数高做乘法.

类似还有如下定理：

定理 6. 如果 $A \in \mathbb{C}^{m\times m}$,$B \in \mathbb{C}^{n\times n}$ 满足

$$\rho(A)\rho(B) < 1, \tag{2.9}$$

则关于 X 的方程 $X - AXB = C$ 的解为：

$$X = \sum_{k=0}^{\infty} A^k C B^k. \tag{2.10}$$

实际上，如果能找到某种算子范数 $\|\cdot\|$, 使得 $\|A\| \cdot \|B\| < 1$, 那么可以得到 $\rho(A)\rho(B) < 1$.

注意解的这种求和形式使得从不同方向做乘法再累加成为可能.

观察 Lyapunov 方程 $AX + XA^T = -GG^T$(其中 A 的所有的特征值实部小于零，即 $re(\lambda_i(A)) < 0, i = 1, 2, \ldots, n$), 对于任取的 $\alpha \in \mathbb{C}$, 该方程均可表示成：

$$(A + \alpha I)X(A^T + \alpha I) - (A - \alpha I)X(A^T - \alpha I) = -2\alpha GG^T.$$

若 α 也取在左半平面，即 $re(\alpha) < 0$, 易知 $(A + \alpha I)$ 与 $(A^T + \alpha I)$ 可逆.

令 $\tilde{A} = (A + \alpha I)^{-1}(A - \alpha I), \tilde{G} = (A + \alpha I)^{-1}G$. 方程可以写成

$$X - \tilde{A}X\tilde{A}^T = -2\alpha\tilde{G}\tilde{G}^T.$$

因为 A 的所有的特征值实部小于零，当 $re(\alpha) < 0$ 时，容易验证 $\rho(\tilde{A}) = \rho(\tilde{A}^T) < 1 \Rightarrow \rho(\tilde{A})\rho(\tilde{A}^T) < 1$,

于是解写成：

$$X = -2\alpha \sum_{k=0}^{\infty} \tilde{A}^k\tilde{G}\tilde{G}^T\tilde{A}^{Tk}. \tag{2.11}$$

令

$$M = (\tilde{G}, \tilde{A}\tilde{G}, \tilde{A}^2\tilde{G}, \tilde{A}^3\tilde{G}, \cdots), \tag{2.12}$$

则 $X = -2\alpha MM^T$.

要计算 X, 只需要用一个循环计算 M 即可.

第三章　Sylvester 的 ADI 因子迭代解法

§3.1　Sylvester 的 ADI 因子迭代解法

解矩阵方程 $AX - XB = C, A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}, C \in \mathbb{C}^{m\times n}$ 的 Alternating-Directional-Implicit(ADI) 迭代 [7] 如下：

$$\begin{aligned}
&\text{选取两组参数 } \{\alpha_i\}, \{\beta_i\} \text{ 和初始向量 } X_0 \\
&For \quad i = 0, 1, \cdots \\
&\qquad \text{I .解}(A - \beta_i I)X_{i+\frac{1}{2}} = X_i(B - \beta_i I) + C\text{得到}X_{i+\frac{1}{2}} \\
&\qquad \text{II .解}X_{i+1}(B - \alpha_i I) = (A - \alpha_i I)X_{i+\frac{1}{2}} - C\text{得到}X_{i+1}
\end{aligned} \tag{3.1}$$

把 II 代入 I 后整理可得：

$$\begin{aligned}
X_{i+1} =& (\beta_i - \alpha_i)(A - \beta_i I)^{-1}C(B - \alpha_i I)^{-1} \\
&+ (A - \alpha_i I)(A - \beta_i I)^{-1}X_i(B - \beta_i I)(B - \alpha_i I)^{-1},
\end{aligned} \tag{3.2}$$

记 X 为方程的精确解，则 X_{i+1} 的误差为：

$$\begin{aligned}
&X_{i+1} - X = (A - \alpha_i I)(A - \beta_i I)^{-1}(X_i - X)(B - \beta_i I)(B - \alpha_i I)^{-1} \\
&= \left[\prod_{j=0}^{i}(A - \alpha_j I)(A - \beta_j I)^{-1}\right](X_0 - X)\left[\prod_{j=0}^{i}(B - \beta_j I)(B - \alpha_j I)^{-1}\right].
\end{aligned} \tag{3.3}$$

参数 $\{\alpha_i\}$ 和 $\{\beta_i\}$ 的选取准则是使误差趋于零. 即使得

$$\rho\left(\prod_{j=0}^{i}(A - \alpha_j I)(A - \beta_j I)^{-1}\right) \cdot \rho\left(\prod_{j=0}^{i}(B - \beta_j I)(B - \alpha_j I)^{-1}\right) \to 0$$

这里 $\rho(A)$ 指 A 的谱半径.

这时实际存在某种算子范数 $\|\cdot\|$, 使得：

$$\left\|\prod_{j=0}^{i}(A - \alpha_j I)(A - \beta_j I)^{-1}\right\| \cdot \left\|\prod_{j=0}^{i}(B - \beta_j I)(B - \alpha_j I)^{-1}\right\| \to 0$$

这样 $X_{i+1} \to X$.