求解奇异线性方程组的双逐次投影法

孟宪亮

（温州大学数学与信息科学学院，浙江温州　325035）

摘　要：用双逐次投影迭代法来求解奇异线性方程组，当线性方程组的系数矩阵是对称半正定时，给出了不同情形时有关参量的选取以及相应的算法，并就收敛结果分别与雅可比迭代法和 Gauss-Seidel 迭代法进行了比较，数值结果表明，该方法对求解奇异线性方程组是很有效的．

关键词：双逐次投影法；奇异线性方程组；内积；对称半正定矩阵

中图分类号：O24　　**文献标志码**：A　　**文章编号**：1674-3563(2009)05-0027-07

DOI：10.3875/j.issn.1674-3563.2009.05.005　　本文的 PDF 文件可以从 xuebao.wzu.edu.cn 获得

最近，文献[1]和[2]讨论了用双逐次投影法求解对称正定线性方程组 $Ax=b$ 的问题，其中 $A\in R^{n\times n}$ 是对称正定矩阵，$x\in R^n$，$b\in R^n$．

双逐次投影法的主要思想是分别构建搜寻空间 κ 和约束空间 φ，κ 和 φ 可取相同的空间，选取初始解 $x_0\in R^n$，以及线性无关的向量 $v_1,v_2\in R^n$，使得 κ,φ 由 v_1,v_2 张成，即 $\kappa=\varphi=span\{v_1,v_2\}$，并按如下的要求来寻找近似解 x，x 属于仿射空间 $x_0+\kappa$，残向量 $b-Ax$ 与 φ 正交，即

$$x\in x_0+\kappa,\ b-Ax\perp\varphi. \tag{1}$$

本文将用双逐次投影法来求解奇异线性方程组：

$$Ax=b, \tag{2}$$

其中，$A\in R^{n\times n}$ 是对称半正定矩阵，$b\in R(A)$．在本文，$R(A)$ 和 $N(A)$ 分别表示 A 的值域和零空间．

1 算法分析

双逐次投影迭代法的迭代格式为：

$$x_{k+1}=x_k+\alpha v_1+\beta v_2, \tag{3}$$

其中，α,β 为常数，$v_1\in R^n$ 和 $v_2\in R^n$ 为线性无关的向量．

对于求解奇异线性方程组（2），向量 v_1,v_2 的选取非常重要，它关系着上述迭代过程的敛散性以及收敛的速度．由于 A 是对称半正定矩阵，定义内积和半范数分别为：

$$\langle x,y\rangle_A=\langle Ax,y\rangle=y^TAx,\ \forall x,y\in R^n.$$

收稿日期：2009-02-02

作者简介：孟宪亮（1981-），男，山东枣庄人，硕士研究生，研究方向：计算数学

$$\|x\|_A^{\ 2}=\langle Ax,x\rangle=x^T Ax\ .$$

为了方便，本文引入下面的记号：

$$a=\langle Av_1,v_1\rangle\ ,\quad c=\langle Av_1,v_2\rangle=\langle Av_2,v_1\rangle\ ,\quad d=\langle Av_2,v_2\rangle\ ,$$

$$p_1=\langle Ax_k-b,v_1\rangle\ ,\quad p_2=\langle Ax_k-b,v_2\rangle\ .$$

因为矩阵 A 是对称半正定的，由内积和半范数的性质可知：$a\ge 0$，$d\ge 0$.

根据双逐次投影法的思想和（3）式来寻找 x_{k+1}，使得

$$x_{k+1}\in x_k+\kappa, b-Ax_{k+1}\perp\varphi\ ,$$

即

$$\begin{cases}\langle b-A_{k+1},v_1\rangle=0\\ \langle b-A_{k+1},v_2\rangle=0\end{cases}\tag{4}$$

即

$$\begin{aligned}\langle b-Ax_k-\alpha Av_1-\beta Av_2,v_1\rangle&=\langle b-Ax_k,v_1\rangle-\alpha\langle Av_1,v_1\rangle-\beta\langle Av_2,v_1\rangle\\&=-p_1-a\alpha-c\beta=0\end{aligned}\ ,\tag{5}$$

$$\begin{aligned}\langle b-Ax_k-\alpha Av_1-\beta Av_2,v_2\rangle&=\langle b-Ax_k,v_2\rangle-\alpha\langle Av_1,v_2\rangle-\beta\langle Av_2,v_2\rangle\\&=-p_2-c\alpha-d\beta=0\end{aligned}\ .\tag{6}$$

注意到，方程组（2）的解等价于如下极小化问题的解：

$$f(x)=\frac{1}{2}\langle Ax,x\rangle-\langle b,x\rangle\to\inf\ .\tag{7}$$

由内积的定义可知：

$$\begin{aligned}\|x_{k+1}-x_*\|_A-\|x_k-x_*\|_A&=\langle Ax_{k+1}-Ax_*,x_{K+1}-x_*\rangle-\langle Ax_k-Ax_*,x_k-x_*\rangle\\&=\langle Ax_{k+1},x_{k+1}\rangle-2\langle b,x_{k+1}\rangle-\big(\langle Ax_k,x_k\rangle-2\langle b,x_k\rangle\big)\ ,\\&=2f(x_{k+1})-2f(x_k)\end{aligned}\tag{8}$$

其中，x_* 是（2）的某个精确解．当 $2f(x_{k+1})-2f(x_k)<0$ 时，求解奇异线性方程组（2）的双逐次投影迭代法就收敛．

由 $x_{k+1}=x_k+\alpha v_1+\beta v_2$ 和（7），知：

$$\begin{aligned}f(x_{k+1})&=f(x_k+\alpha v_1+\beta v_2)\\&=f(x_k)+\alpha\langle Ax_k-b,v_1\rangle+\beta\langle Ax_k-b,v_2\rangle+\frac{1}{2}\alpha^2\langle Av_1,v_1\rangle\ ,\\&=+\alpha\beta\langle Av_1,v_2\rangle+\frac{1}{2}\beta^2\langle Av_2,v_2\rangle\end{aligned}$$

即　　$$f(x_{k+1})-f(x_k)=p_1\alpha+p_2\beta+\frac{1}{2}a\alpha^2+c\alpha\beta+\frac{1}{2}d\beta^2\ .$$

令　　$$g(\alpha,\beta)=p_1\alpha+p_2\beta+\frac{1}{2}a\alpha^2+c\alpha\beta+\frac{1}{2}d\beta^2\ ,\tag{9}$$

易知，$g(\alpha,\beta)$要取得最小值，则α,β应满足如下条件：

$$\begin{cases}\dfrac{\partial g(\alpha,\beta)}{\partial\alpha}=p_1+a\alpha+c\beta=0\\ \dfrac{\partial g(\alpha,\beta)}{\partial\beta}=p_2+c\alpha+d\beta=0\end{cases},\tag{10}$$

从而（10）式与（5）和（6）式是等效的.

由于A是对称半正定的，对（10）式的求解情况比A是对称正定矩阵时复杂多了，特别是当$c=0$时迭代的收敛性得不到保证，因此这里我们假定$c\neq 0$.

事实上关于$c\neq 0$的v_1,v_2的选取很简单，因为对任何$A\in R^{n\times n}$，若A对称，则其值域$R(A)$和零空间$N(A)$构成R^n的正交投影分解，即$R(A)\oplus N(A)=R^n$，从而只需在$R(A)$中选取两个线性无关的向量便可充当v_1和v_2. 因为$a\geq 0,d\geq 0$，所以我们对（10）求解可得如表 1 的四种情形.

表 1　四种情形的各数据结果

Table 1　The Data of Four Cases

序列情形	参变量			
	a,d	α	β	$g(\alpha,\beta)$
(i)	$a=0,d=0$	$-\frac{p_2}{c}$	$-\frac{p_1}{c}$	$-\frac{p_1p_2}{c}$
(ii)	$a>0,d=0$	$-\frac{p_2}{c}$	$\frac{ap_2}{c^2}-\frac{p_1}{c}$	$-\frac{p_1p_2}{c}+\frac{ap_2{}^2}{2c^2}$
(iii)	$a=0,d>0$	$\frac{-p_2c-dp_1}{c^2}$	$-\frac{p_1}{c}$	$-\frac{p_1p_2}{c}+\frac{dp_1{}^2}{2c^2}$
(iv)	$a>0,d>0$	$\frac{cp_2-dp_1}{ad-c^2}$	$\frac{ap_2-cp_1}{ad-c^2}$	$-\frac{dp_1{}^2+ap_2{}^2-2cp_1p_2}{2(ad-c^2)}$

基于上面的理论本文给出如下算法，其中α,β的取值为情形（iv），v_1,v_2的取值属文后数值算例 3 的情况，其它的情形及算例的α,β，v_1,v_2只要对算法作出相应的修改就可.

1）取x_0以及精度Tol

2）$r_0=b-Ax_0$

$r_{01}=\|r_0\|_2$

k=0；

for $i=2,3,\cdots,n$；

$j=i-1$；

$v_1=0.99\times E_i+9\times f$；

$v_2=0.99\times E_j+9\times f$；

$a=\langle Av_1,v_1\rangle$；

$c=\langle Av_1,v_2\rangle$；

$d = \langle Av_2, v_2 \rangle$；

$p_1 = \langle Ax_k - b, v_1 \rangle$；

$p_2 = \langle Ax_k - b, v_2 \rangle$；

$\alpha = \dfrac{cp_2 - dp_1}{ad - c^2}$；

$\beta = \dfrac{ap_2 - cp_1}{ad - c^2}$；

$x_{k+1} = x_k + \alpha v_1 + \beta v_2$；

$r_1 = b - Ax$；

$r_{11} = \|r_1\|_2$；

k=k+1

if $r_{11} / r_{01} \le Tol$；

break；

end；

end.

2 数值实验

这里给出了三个试验，其中例 1 说明情形（i），例 2 说明情形（ii）和（iii），例 3 说明情形（iv）．所用计算机的主频为 1.90GHz，内存为 512M，以 MATLAB 7.0 为平台进行计算．

在下面的几个数值算例中取 $x_0 = (0, 0, \cdots, 0)^T$， $f = (1, 1, \cdots, 1)^T$， $b = Af$，矩阵 H 是对角矩阵，对角线上的元素为 $H_{i,i} = 0.0001 \times i^2$，$i = 1, 2, \cdots n$，其它的元素为 0，$H_i$ 是 H 的第 i 列．E 是单位矩阵，E_i 为 E 的第 i 列．Tol 取 10^{-6}．但在数值算例 3 中 b 取 A 与一随机向量的乘积，该随机向量由 Matlab 命令 rand 生成．

数值算例 1　矩阵 A 是如下对称半正定的带状矩阵（见文献[3]），阶数为奇数时矩阵 A 是奇异的，这里取 $n = 35$，即 A 的阶数为 1 225×1 225．选取 $v_1 = H_i + f$，$v_2 = H_j + f$，$\alpha = -\dfrac{p_2}{c}$，$\beta = -\dfrac{p_1}{c}$．

$$
A = \begin{bmatrix}
0.5 \times U_n & -M_n^T & & & & & \\
-M_n & U_n & -M_n^T & & & & \\
& \ddots & \ddots & \ddots & & & \\
& & -M_n & U_n & -M_n^T & & \\
& & & \ddots & \ddots & \ddots & \\
& & & & -M_n & U_n & -M_n^T \\
& & & & & -M_n & 0.5 \times U_n
\end{bmatrix}
$$

其中，

$$U_n=\begin{bmatrix} 0 & -1 & & & & & \\ -1 & 0 & -1 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & -1 & 0 & -1 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 0 & -1 \\ & & & & & -1 & 0 \end{bmatrix}，M_n=\begin{bmatrix} 1/2 & & & & & & 1 \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & & & & & 1/2 \end{bmatrix}.$$

收敛情况如图 1（a）. 由表 1 可看出（ii）和（iii）是一种类型，下面以算例 2 来说明这一情形.

数值算例 2　选取文献[1]中例 1 的矩阵 A，$a_{i,i}=4n$，$a_{i,i+1}=a_{i+1,i}=n$，$a_{i,j}=0.5$，$i=1,\cdots,n$，$j\neq i,i+1$. 这里把矩阵 A 最后一行的元素和最后一列的元素全变为 0. 矩阵 A 是 900×900 阶对称半正定矩阵. $\alpha=-\frac{p_j}{c}$，$\beta=\frac{(ap_2-cp_1)}{c^2}$，取 $v_1=0.001\times E_i+9.9\times f$，$v_2=0.001\times E_j+9.9\times f$.

$$A=\begin{bmatrix} 4n & n & 0.5 & \cdots & 0.5 & 0 \\ n & 4n & n & \ddots & \vdots & 0 \\ 0.5 & n & \ddots & \ddots & 0.5 & \vdots \\ \vdots & \ddots & \ddots & 4n & n & 0 \\ 0.5 & \cdots & 0.5 & n & 4n & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}.$$

收敛情况见图 1（b）. 下面的数值算例 3 属于情形（iv）.

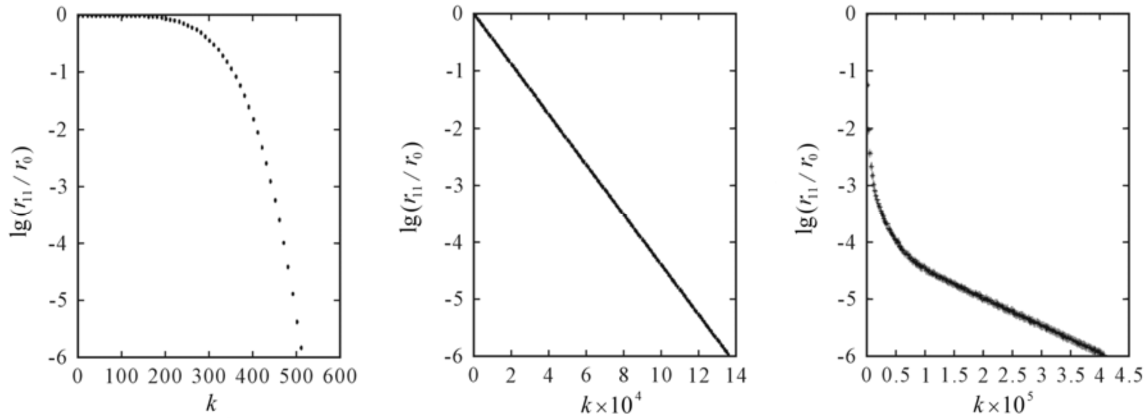
数值算例 3　该矩阵 A 为文献[4]中的例 1，即用五点中心差分离散泊松方程所得到的系数矩阵，矩阵形式如下：

$$A=\begin{bmatrix} 0.5\times D_n & -N_n & & & & & \\ -N_n & D_n & -N_n & & & & \\ & \cdots & \cdots & \cdots & & & \\ & & -N_n & D_n & -N_n & & \\ & & & \cdots & \cdots & \cdots & \\ & & & & -N_n & D_n & -N_n \\ & & & & & -N_n & 0.5\times D_n \end{bmatrix},$$

其中，

$$D_n=\begin{bmatrix} 2 & -1 & & & & & \\ -1 & 4 & -1 & & & & \\ & \cdots & \cdots & \cdots & & & \\ & & -1 & 4 & -1 & & \\ & & & \cdots & \cdots & \cdots & \\ & & & & -1 & 4 & -1 \\ & & & & & -1 & 2 \end{bmatrix}，N_n=\begin{bmatrix} 1/2 & & & & & & \\ & 1 & & & & & \\ & & \cdots & & & & \\ & & & 1 & & & \\ & & & & \cdots & & \\ & & & & & 1 & \\ & & & & & & 1/2 \end{bmatrix},$$

取 $v_1 = 0.99 \times E_i + 9 \times f$ ，$v_2 = 0.99 \times E_j + 9 \times f$ ，$\alpha = \dfrac{cp_2 - dp_1}{ad - c^2}$ ，$\beta = \dfrac{ap_2 - cp_1}{ad - c^2}$ ．A 是 900×900 的．收敛情况见图 1（c）．



(a) $a=0,\ d=0$ (b) $a>0,\ d=0$ (c) $a>0,\ d>0$

(横坐标 k 表示迭代步数，每迭代 1000 次画出一个点．下同)

图 1 不同情形时的收敛图像

Fig 1 The Convergence Chart in Different Cases

3 双逐次投影法与其它迭代方法的比较

下面对双逐次投影法、雅可比迭代法、Gauss-Seidel 迭代法进行比较．

对上面数值算例 3 中的矩阵 A，双逐次投影法的有关参量的取值和数值算例 3 一样，矩阵 A 的大小为 225×225 阶的矩阵．采用双逐次投影法、雅可比法和 Gauss-Seidel 迭代法的收敛情况见图 2．

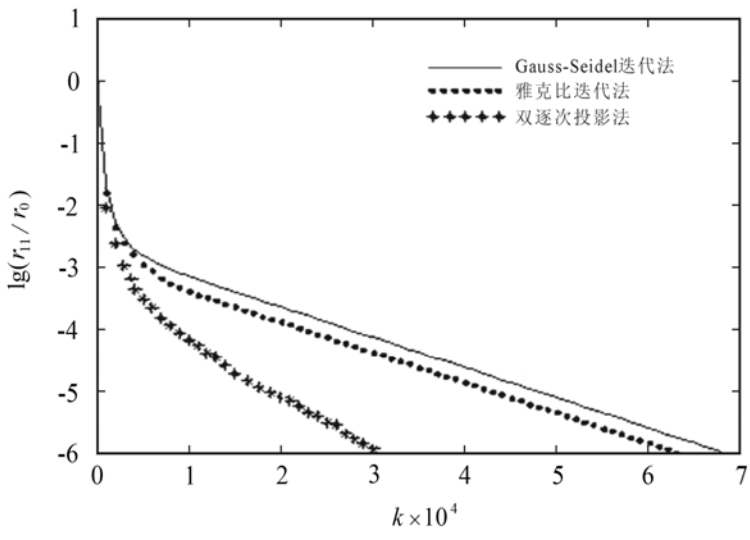


图 2 三种迭代法收敛曲线

Fig 2 The Convergence Curves of Three Iterative Methods

三种方法的迭代步数和 CPU 时间的比较见表 2．

表 2　三种迭代法收敛的比较

Table 2　The Comparison of Three Iterative Methods

迭代法	迭代步数	CPU 所用的时间 / 秒
双逐次投影法	3.7×10^4	23
雅可比法	6.3×10^4	175
Gauss-Seidel 法	5.3×10^4	90

由图 2 和表 2 可知，通过双逐次投影法求解线性方程组（2），使线性方程组（2）收敛且优于雅可比法和 Gauss-Seidel 法.

参考文献

[1] Jing Y F, Huang T Z. On a new Iterative Method for Solving Linear Systems and Comparion Results [J]. Journal of Computational and Applied Mathematics, 2008, 220(1-2), 74-84.

[2] Ujevic N. A New Iterative Method for Solving Linear Systems [J]. Applied Mathematics and Computation, 2006, 179(2), 725-730.

[3] James W D. 应用数值线性代数[M]. 王国荣, 译. 北京: 人民邮电出版社, 2007: 69-72.

[4] Zhang S L, Yoshio O, Augihara M. Necessary and Sufficient Conditions for the Convergence of Orthomin (k) on Singular and Inconsistent Linear Systems [J]. Numerische Mathematik, 2000, (87): 391-405.

Two Dimensional Double Successive Projection Method for Solving Singular Linear Equations

MENG Xianliang

(College of Mathematics and Information Science, Wenzhou University, Wenzhou, China　325035)

Abstract: In the discussion of issue of using two dimensional double successive projection method to solve singular linear equations, parameters of the different circumstances as well as the appropriate algorithm were given when the coefficient matrix was symmetric semidefinite positive definite. And the results of convergence were compared with the Jacobian iterative method and Gauss-Seidel iteration method. The results showed that the method for solving singular linear systems is effective.

Key words: Two Dimensional Double Successive Projection Method; Singular Linear Equations; Inner Product; Symmetric Positive Semidefinite Matrix

（编辑：王一芳）