M-矩阵线性方程组的一类非定常迭代法*

关晋瑞，温瑞萍

(太原师范学院数学系，山西 晋中 030619)

摘要:为了计算M-矩阵线性方程组的解,本文基于M-矩阵的性质以及矩阵分析技巧,提出了一类非定常迭代法以计算方程组的解,并给出了相应的收敛性分析.理论分析表明,该方法具有二次收敛率.数值实验表明,新方法是可行的,而且在一定情况下也是较为有效的.

关键词:线性方程组;M-矩阵;非定常迭代法;收敛率

中图分类号:O241.6 **DOI**:10.19789/j.1004-9398.2021.06.003

0 引 言

在科学计算和工程应用中,许多问题最终都归结为线性方程组的求解,因此研究线性方程组解的理论以及求解方法具有重要的理论意义和应用价值[1-4].如何快速有效地求得线性方程组的解是计算数学的一个基本问题,也是目前仍在继续研究的重要课题之一[5-16].

本文考虑M-矩阵线性方程组的求解

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \tag{1}$$

式中$\boldsymbol{A}\in\mathbf{R}^{n\times n}$是非奇异的M-矩阵,$\boldsymbol{b}\in\mathbf{R}^{n}$是给定的向量.这种类型的线性方程组在微分方程数值解、线性互补问题、Markov链以及物理学、生物学和经济学中都有着广泛的应用[2,4,17].例如,对于Poisson方程第一边值问题,利用差分法离散化之后即得到一个系数矩阵为M-矩阵的线性方程组.在线性互补问题中,很多情况下的系数矩阵都是M-矩阵.在经济学中,投入产出分析对应的矩阵也是M-矩阵[17].

对于一般的线性方程组已有很多经典的解法,如Gauss消去法、平方根法、Jacobi迭代法、Gauss-Seidel迭代法、SOR、共轭梯度法、Krylov子空间方法[1-4]和HSS迭代法[18]等.而对于一些具有特殊性质和结构的方程组,也有很多学者进行了深入研究.本文利用M-矩阵的特点,针对M-矩阵线性方程组提出了一类迭代法.理论分析和数值实验表明,新方法是可行的,而且在一定情况下也是较为有效的.

1 预备知识

首先简要介绍本文的符号记法及将要用到的一些预备知识.

$\mathbf{R}^{m\times n}$表示实数域上全体$m\times n$的矩阵,$\mathbf{R}^{n}$表示实数域上全体n维列向量,$\rho(\boldsymbol{A})$表示矩阵$\boldsymbol{A}$的谱半径.n阶单位矩阵记为$\boldsymbol{I}_n$,在没有混淆的情况下写为$\boldsymbol{I}$.用$\|\cdot\|$来表示一个给定的矩阵范数.

设$\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{n\times n}$,若$\boldsymbol{A}$的每个元素都是非负的,则称$\boldsymbol{A}$为非负矩阵,记为$\boldsymbol{A}\geqslant 0$.设$\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{n\times n}$,若对于任意的$i\neq j$都有$a_{ij}\leqslant 0$成立,则称$\boldsymbol{A}$为Z-矩阵.一个Z-矩阵$\boldsymbol{A}$若可以表示为$\boldsymbol{A}=s\boldsymbol{I}-\boldsymbol{B}$,其中$\boldsymbol{B}$为非负矩阵且$s\geqslant\rho(\boldsymbol{B})$,则称该Z-矩阵为M-矩阵.特别地,当$s>\rho(\boldsymbol{B})$时,称$\boldsymbol{A}$为非奇异的M-矩阵,当$s=\rho(\boldsymbol{B})$时,称$\boldsymbol{A}$为奇异的M-矩阵.关于M-矩阵的详细内容可参考文献[17].

分裂迭代法是求解线性方程组(1)的最简单的一类迭代法.分裂迭代法的基本思路是,给定系数矩阵$\boldsymbol{A}$的一个分裂

$$\boldsymbol{A}=\boldsymbol{M}-\boldsymbol{N},$$

式中$\boldsymbol{M}$是可逆的,代入式(1)得到原方程组的等价形式$\boldsymbol{M}\boldsymbol{x}=\boldsymbol{N}\boldsymbol{x}+\boldsymbol{b}$,进而可以得到迭代法

$$\boldsymbol{x}_{k+1}=\boldsymbol{M}^{-1}\boldsymbol{N}\boldsymbol{x}_k+\boldsymbol{M}^{-1}\boldsymbol{b},\ k=0,1,2,\cdots, \tag{2}$$

式中$\boldsymbol{M}^{-1}\boldsymbol{N}$称为分裂迭代法(2)的迭代矩阵.

如果对于任意给定的初始向量$\boldsymbol{x}_0\in\mathbf{R}^{n}$,由式(2)生成的序列$\{\boldsymbol{x}_k\}$都收敛于方程组(1)的解$\boldsymbol{x}^{*}$,则

收稿日期:2020-11-09

* 山西省高等学校科技创新项目(2019L0783);山西省应用基础研究计划项目(201901D211423)

称迭代法(2)是收敛的.

引理 1.1[19]　分裂迭代法(2)收敛当且仅当迭代矩阵$\boldsymbol{M}^{-1}\boldsymbol{N}$的谱半径$\rho(\boldsymbol{M}^{-1}\boldsymbol{N})<1$.

引理 1.2[20]　设$\|\cdot\|$是任意一个矩阵范数,则对任意的$\boldsymbol{A}\in\mathbf{R}^{n\times n}$有$\lim\limits_{k\to\infty}\left\|\boldsymbol{A}^k\right\|^{\frac{1}{k}}=\rho(\boldsymbol{A})$.

2　一类非定常迭代法

本节提出一类非定常迭代法,以求解M-矩阵线性方程组,并给出该方法的收敛性分析.

设$\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{n\times n}$为非奇异的M-矩阵,则$\boldsymbol{A}$存在自然的分裂

$$\boldsymbol{A}=s\boldsymbol{I}-\boldsymbol{B},\tag{3}$$

式中$s>0$,而$\boldsymbol{B}\geqslant 0$是非负矩阵.事实上,不难验证只要选取$s\geqslant\max\limits_{1\leqslant i\leqslant n}\{a_{ii}\}$,令$\boldsymbol{B}=s\boldsymbol{I}-\boldsymbol{A}$即可.于是基于上述分裂便可以得到如下的一类迭代法

$$s\boldsymbol{x}_{k+1}=\boldsymbol{B}\boldsymbol{x}_k+\boldsymbol{b},\ k=0,1,2,\cdots,$$

或者

$$\boldsymbol{x}_{k+1}=\frac{1}{s}\boldsymbol{B}\boldsymbol{x}_k+\frac{1}{s}\boldsymbol{b},\ k=0,1,2,\cdots,\tag{4}$$

式中$\boldsymbol{x}_0\in\mathbf{R}^n$是任意给定的初始向量.

由于$\boldsymbol{A}$是非奇异的M-矩阵,可知$s>\rho(\boldsymbol{B})$.于是利用引理1.1,可得到如下的收敛性.

定理 2.1　设线性方程组(1)的系数矩阵是非奇异的M-矩阵,则对于任意给定的初始向量,迭代法(4)都是收敛的.

在迭代法(4)中存在一个参数s.下面讨论参数s对迭代的影响,并给出最优参数的值.

定理 2.2　迭代法(4)中的最优参数为$s=\max\limits_{1\leqslant i\leqslant n}\{a_{ii}\}$.

证明　设$\boldsymbol{A}=s_1\boldsymbol{I}-\boldsymbol{B}_1=s_2\boldsymbol{I}-\boldsymbol{B}_2$为矩阵$\boldsymbol{A}$的2个分裂,其中$s_1,s_2\geqslant\max\limits_{1\leqslant i\leqslant n}\{a_{ii}\}$都是正数,且$s_1>s_2$,而$\boldsymbol{B}_1,\boldsymbol{B}_2$是非负矩阵.相应的迭代矩阵谱半径分别为$\dfrac{\rho(\boldsymbol{B}_1)}{s_1}$和$\dfrac{\rho(\boldsymbol{B}_2)}{s_2}$.

由$\boldsymbol{B}_1=(s_1-s_2)\boldsymbol{I}+\boldsymbol{B}_2$,且$\rho(\boldsymbol{B}_1)=s_1-s_2+\rho(\boldsymbol{B}_2)$,可得

$$\frac{\rho(\boldsymbol{B}_2)}{s_2}<\frac{\rho(\boldsymbol{B}_2)+(s_1-s_2)}{s_2+(s_1-s_2)}=\frac{\rho(\boldsymbol{B}_1)}{s_1}.$$

因此,参数s越小,收敛越快.从而当$s=\max\limits_{1\leqslant i\leqslant n}\{a_{ii}\}$时,迭代法(4)收敛最快.证毕.

由于迭代法(4)的收敛速度与$\dfrac{1}{s}\rho(\boldsymbol{B})$相关,当$\rho(\boldsymbol{B})<<s$时,即矩阵远离奇异时,迭代收敛较快,而当$\rho(\boldsymbol{B})$接近$s$时,即矩阵接近奇异时,迭代收敛很慢.为了加速收敛,对其迭代格式进行改变,得到如下的非定常迭代

$$\begin{cases}\boldsymbol{x}_{k+1}=\boldsymbol{B}_k\boldsymbol{x}_k+\boldsymbol{x}_k,\\ \boldsymbol{B}_{k+1}=\boldsymbol{B}_k^2,\end{cases}\tag{5}$$

式中$\boldsymbol{B}_0=\dfrac{1}{s}\boldsymbol{B}$,$\boldsymbol{x}_0=\dfrac{1}{s}\boldsymbol{B}\bar{\boldsymbol{x}}_0+\dfrac{1}{s}\boldsymbol{b}$,而$\bar{x}_0$为(4)的初始向量.不难验证上述非定常迭代法(5)第k步计算的结果相当于迭代法(4)第2^k步计算的结果,因此(5)与(4)具有相同的收敛性.

定理 2.3　设线性方程组(1)的系数矩阵为非奇异的M-矩阵,则迭代法(5)是收敛的,且具有二次收敛率.

证明　由迭代法(4)可得方程组的精确解为

$$x^*=\sum_{j=0}^{\infty}\left(\frac{\boldsymbol{B}}{s}\right)^j\left(\frac{\boldsymbol{B}}{s}\bar{\boldsymbol{x}}_0+\frac{\boldsymbol{b}}{s}\right),$$

式中$\bar{\boldsymbol{x}}_0$为(4)的初始向量.直接验证可知非定常迭代法(5)得到的序列$\{x_k\}$满足

$$x_k=\sum_{j=0}^{2^k-1}\left(\frac{\boldsymbol{B}}{s}\right)^j\left(\frac{\boldsymbol{B}}{s}\bar{x}_0+\frac{\boldsymbol{b}}{s}\right).$$

于是有

$$x^*-x_k=\sum_{j=2^k}^{\infty}\left(\frac{\boldsymbol{B}}{s}\right)^j\left(\frac{\boldsymbol{B}}{s}\bar{x}_0+\frac{\boldsymbol{b}}{s}\right)=$$

$$\left(\frac{\boldsymbol{B}}{s}\right)^{2^k}\sum_{j=0}^{\infty}\left(\frac{\boldsymbol{B}}{s}\right)^j\left(\frac{\boldsymbol{B}}{s}\bar{x}_0+\frac{\boldsymbol{b}}{s}\right)=\left(\frac{\boldsymbol{B}}{s}\right)^{2^k}x^*,$$

两边取范数,并利用引理1.2,可得

$$\lim_{k\to\infty}\|\boldsymbol{x}^*-\boldsymbol{x}_k\|^{\frac{1}{2^k}}\leqslant\lim_{k\to\infty}\left\|\left(\frac{\boldsymbol{B}}{s}\right)^{2^k}\right\|^{\frac{1}{2^k}}\cdot\|\boldsymbol{x}^*\|^{\frac{1}{2^k}}=\frac{\rho(\boldsymbol{B})}{s}.$$

因此,序列$\{\boldsymbol{x}_k\}$是收敛的且具有二次收敛率.证毕.

尽管非定常迭代(5)相对于(4)运算量较大,但是上面的定理表明迭代收敛很快.因此是可行的.

3　数值实验

本节将通过几个数值检验迭代法(4)与(5)的数值效果.实验中,所有的程序都用Matlab(R2012a)编写,并在个人笔记本电脑(2G CPU和8G内存)上运行.实验结果给出2种方法所需的迭代次数(IT)、CPU运行时间(CPU,s)以及残差(RES),其

中残差定义为 $\mathrm{RES}=\dfrac{\|\boldsymbol{Ax}-\boldsymbol{b}\|_2}{\|\boldsymbol{b}\|_2}$.实验中,初始向量取为0向量,当残差满足 $\mathrm{RES}<10^{-8}$ 时迭代终止.

例 3.1[19] 考虑M-矩阵线性方程组(1),其中

$$\boldsymbol{A}=\begin{pmatrix} T & -I & & & \\ -I & T & -I & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & -I & T & -I \\ & & & -I & T \end{pmatrix}\in\mathbf{R}^{m\times m},$$

$$\boldsymbol{T}=\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}\in\mathbf{R}^{m\times m},$$

而 $n=m^2$, $b=(1,1,\cdots,1)^{\mathrm{T}}\in\mathbf{R}^n$.当 m 取不同值,2个迭代法得出的结果如表1所示.与迭代法(4)相比,迭代法(5)不仅迭代次数和运行时间都明显要少,而且求得解的精度也高.特别地,当系数矩阵的阶数增加时,迭代法(4)的迭代次数急剧增加,但迭代法(5)的迭代次数增加不明显.

表1 m 取不同值时不同迭代公式对应的结果

m	式(4)			式(5)		
	IT	CPU/s	RES	IT	CPU/s	RES
8	295	0.022 4	$9.591\,3\times10^{-9}$	9	0.028 6	$1.332\,3\times10^{-14}$
16	1 064	0.049 0	$9.929\,8\times10^{-9}$	11	0.048 2	$1.032\,6\times10^{-14}$
32	4 020	4.437 7	$9.957\,2\times10^{-9}$	12	1.469 6	$7.052\,5\times10^{-9}$

注:IT表示迭代次数;CPU表示运行时间;RES表示残差.

例 3.2 考虑M-矩阵线性方程组(1),其中 A 由下列命令生成

a=rand(n,n);A=diag(a*ones(n,1))−a+eye(n);

而 $\boldsymbol{b}=(1,1,\cdots,1)^{\mathrm{T}}\in\mathbf{R}^n$.对于阶数 n 的不同取值,表2中给出了实验结果.从表2中可见,随着系数矩阵阶数的增加,迭代法(4)的迭代次数与运行时间都急剧增加,但迭代法(5)的迭代次数和运行时间并没有明显的增加.

例 3.3 考虑M-矩阵线性方程组(1),其中

表2 n 取不同值时不同迭代公式对应的结果

n	式(4)			式(5)		
	IT	CPU/s	RES	IT	CPU/s	RES
100	1 095	0.027 2	$9.966\,6\times10^{-9}$	11	0.021 3	$1.240\,2\times10^{-14}$
200	2 034	0.050 6	$9.919\,8\times10^{-9}$	11	0.032 9	$5.622\,4\times10^{-13}$
500	4 933	1.039 7	$9.965\,7\times10^{-9}$	13	0.251 7	$1.172\,6\times10^{-13}$
1 000	9 667	10.505 5	$9.988\,0\times10^{-9}$	14	1.613 4	$3.139\,1\times10^{-13}$

$$A=\begin{pmatrix} 2+\varepsilon & -1 & -1 \\ -1 & 2+\varepsilon & -1 \\ -1 & -1 & 2+\varepsilon \end{pmatrix},b=\begin{pmatrix}1\\1\\1\end{pmatrix},$$

$\varepsilon>0$ 是正数.当 $\varepsilon\to0$ 时,矩阵 $\boldsymbol{A}$ 接近奇异.对于参数 ε 的不同取值,实验结果见表3.当参数越来越小时,迭代法(4)的迭代次数急剧增加,但迭代法(5)的迭代次数并没有明显的增加.

表3 参数取不同值时不同迭代公式对应的结果

ε	式(4)		式(5)	
	IT	RES	IT	RES
1	46	$7.939\,7\times10^{-9}$	6	$5.372\,4\times10^{-12}$
0.1	378	$9.782\,4\times10^{-9}$	9	$1.416\,1\times10^{-11}$
0.01	3 694	$9.967\,1\times10^{-9}$	12	$1.342\,2\times10^{-9}$
0.001	36 851	$9.997\,8\times10^{-9}$	16	0

由以上3个数值例子可见,迭代法(4)收敛较慢,特别当系数矩阵接近奇异时,收敛非常缓慢,而迭代法(5)始终收敛很快,且求得解的精度也比较高,因此迭代法(5)是较为有效的.

4 结束语

本文提出了一类非定常迭代法以求解M-矩阵线性方程组.理论分析表明该方法具有二次收敛率.数值实验表明,该方法是可行的,而且在一定情况下也是较为有效的.但是该方法在每步迭代中运算量较大,如何减少运算量并保持较快的收敛速度,还有待进一步的研究.另外,由于H-矩阵和M-矩阵具有很多相近的性质,本文的方法还可以推广到H-矩阵线性方程组.

参 考 文 献

[1] VARGA R S. Matrix iterative analysis[M]. Berlin:Springer-Verlag,2000.

[2] MEYER C D. Matrix analysis and applied linear algebra[M]. Philadelphia:SIAM,2000.

[3] DEMMEL J W. Applied numerical linear algebra[M]. Philadelphia:SIAM,1997.

[4] SAAD Y. Iterative methods for sparse linear systems[M].北京:科学出版社,2009.

[5] 孙蕾.求解非对称线性方程组的不完全最小联合向后扰动法[J].高等学校计算数学学报,2016,38(3):269−288.

[6] 孙蕾.求解大型非对称线性方程组的不完全广义最小向后扰动法[J].数学进展,2016,45(6):939−954.

[7] 温瑞萍,李苏丹.迭代求解非Hermitian正定线性方程组的衍生多分裂方法[J].应用数学,2018,31(1):1-11.
[8] 张迎春,吕全义,肖曼玉.复线性方程组的预处理MCG算法[J].工程数学学报,2018,35(3):308-318.
[9] 潘春平,王红玉,曹文方.非Hermitian正定线性方程组的外推的HSS迭代方法[J].计算数学,2019,41(1):52-65.
[10] 房喜明,令锋,傅守忠.求解线性方程组的一般共轭梯度法[J].数学理论与应用,2019,39(2):62-71.
[11] 蔡静.严格次对角占优线性方程组迭代法的收敛性分析[J].华东师范大学学报(自然科学版),2019(2):1-6.
[12] 王福胜,高娟,赵媛璐,等.混合约束Minimax问题的基于序列线性方程组的模松弛SQP算法[J].应用数学学报,2019,42(2):242-253.
[13] 杜亦疏,殷俊锋,张科.求解大型稀疏线性方程组的贪婪距离随机Kaczmarz方法[J].同济大学学报(自然科学版),2020,48(8):1224-1231.
[14] 温瑞萍,段辉.H-矩阵线性方程组的一类预条件并行多分裂SOR迭代法[J].应用数学,2020,33(4):814-825.
[15] 王燕荣,陈云兰.一种改进的求解大型线性方程组的Jacobi迭代法[J].山东大学学报(理学版),2020,55(6):122-126.
[16] 刘瑞华,邹洋杨,谢挺.求解Hilbert矩阵线性方程组的SOR迭代预处理方法[J].高等数学研究,2020,23(1):60-63.
[17] BERMAN A,PLEMMONS R J. Nonnegative matrices in the mathematical sciences[M]. New York:Academic Press,1994.
[18] BAI Z Z,GOLUB G H,NG M K. Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems[J]. SIAM Journal on Matrix Analysis and Applications,2003,24:603-626.
[19] 徐树方,高立,张平文.数值线性代数[M].2版.北京:北京大学出版社,2013.
[20] 徐树方.控制论中的矩阵计算[M].北京:高等教育出版社,2011.

A class of non-stationary iterative method for M-matrix linear systems

GUAN Jinrui, WEN Ruiping

(Department of Mathematics, Taiyuan Normal University, Jinzhong Shanxi 030619)

Abstract: This paper studies the numerical solution of M-matrix linear systems. Based on the properties of M-matrix and matrix analysis techniques, a class of non-stationary iterative methods is proposed to solve the linear systems, and the corresponding convergence analysis is given. Theoretical analysis shows that the method has quadratic convergence rate. Numerical experiments show that the new method is feasible and effective in some cases.

Keywords: linear systems; M-matrix; non-stationary iterative method; convergence rate

(责任编辑:马田田)