改进的高斯-赛德尔迭代法的收敛性分析

黄湧辉
(华南师范大学 数学科学学院,广东 广州 510631)

【摘　要】本文讨论了改进的高斯-赛德尔迭代法的收敛性。在严格对角占优的L-矩阵条件下,该预条件加快了高斯-赛德尔迭代法的收敛速度,而且在该预条件下高斯-赛德尔迭代法的谱半径是单调下降的。最后用数值例子说明本文得出的结论。

【关键词】严格对角占优L-矩阵;预条件迭代法;谱半径;弱正则分裂;收敛速度
【中图分类号】O241.6　【文献标识码】A　【文章编号】1673-1891(2011)01-0015-03

引言

本文考虑实线性方程组

$$Ax=b \tag{1}$$

其中 $A=(a_{ij})_{n\times n}\in R^{n\times n}$ 为n阶方阵,$x\in R^n$ 和 $b\in R^n$ 是n维向量。对系数矩阵A作A=M-N的分裂,M为非奇异矩阵,则对应方程组(1)的基本迭代形式为:

$$x^{(k+1)}=M^{-1}Nx^{(k)}+M^{-1}b,k=0,1,2,\cdots n \tag{2}$$

其中称 M^{-1} 为方程(1)的迭代矩阵,迭代形式(2)是否收敛取决于迭代矩阵 $M^{-1}N$。当 $k\to\infty$,$M^{-1}N\to 0$,即 $k\to\infty$ 时,谱半径 $\rho(M^{-1}N)<1$,并且其收敛速度随谱半径 $\rho(M^{-1}N)$ 的减小而加快。

一般地,对线性方程组(1)的系数矩阵A做如下的分裂

$$A=D-L-U \tag{3}$$

其中D为非奇异对角矩阵,L为严格下三角矩阵,U为严格上三角矩阵。

为讨论方便,当A可逆时,总可以通过初等变换把A的对角元都化简为1,因此

$$A=I-L-U$$

对于形如(3)的系数矩阵,Gauss-Seidel迭代法的迭代矩阵为:$G=(I-L)^{-1}U$。

虽然迭代法是求解线性方程组的主要方法,但是当系数矩阵的条件数较大时,系数矩阵对于求解线性方程组是病态的,此时迭代法会出现不收敛和收敛速度慢的情况。对于迭代法收敛速度慢的问题,采用适当的预处理技术,可以使系数矩阵的特征值分布更加集中,降低矩阵的条件数,改良矩阵的病态特性。

现对线性方程组(1),两端左乘可逆矩阵 $P\in R^{n\times n}$,使方程组等价地转换为一个预条件的方程组

$$PAx=Pb \tag{4}$$

通常称可逆矩阵P为预处理矩阵(或预处理因子),则与(2)对应的基本迭代形式为:

$$x^{(k+1)}=(MP)^{-1}NPx^{(k)}+(MP)^{-1}Pb,k=0,1,2,\cdots n$$

其中PA=PM-PN是PA的分裂,且PM是可逆的。

本文主要考虑,对于线性方程组(1),如何选取预处理矩阵P,使得用Gauss-Seidel迭代法解经过预处理后的方程组(4)有更快的收敛速度。本文在文[1]的基础上,考虑用一类新的预处理方法解线性方程组(1),取

$$P_\alpha=I+S_\alpha=\begin{pmatrix} 1 & 0 & \Lambda & 0 & 0 \\ -\alpha_2 a_{21} & 1 & \Lambda & 0 & 0 \\ M & M & O & M & M \\ -\alpha_{n-1}a_{n-1,1} & 0 & \Lambda & 1 & 0 \\ -\alpha_n a_{n,1} & 0 & \Lambda & 0 & 1 \end{pmatrix}$$

则用 P_α 左乘线性方程组(1)的两边,

$$\begin{aligned} A_\alpha &= P_\alpha A=(I+S_\alpha)A=(I+S_\alpha)(I-L-U) \\ &= I-L-U+S_\alpha-S_\alpha L-S_\alpha U=I-L-U+S_\alpha-(D_S+L_S+U_S) \\ &= (I-D_S)-(L+L_S-S_\alpha)-(U+U_S) \\ &= \begin{pmatrix} 1 & a_{12} & \Lambda & a_{1,n} \\ a_{21}-\alpha_2 a_{21} & 1-\alpha_2 a_{12}a_{21} & \Lambda & a_{2,n}-\alpha_2 a_{1,n}a_{21} \\ M & M & O & M \\ a_{n,1}-\alpha_n a_{n,1} & a_{n,2}-\alpha_n a_{12}a_{n,1} & \Lambda & 1-\alpha_n a_{1,n}a_{n,1} \end{pmatrix} \\ &= D_\alpha-L_\alpha-U_\alpha \end{aligned}$$

其中 $S_\alpha L=0$,$D_\alpha=diag(1,1-\alpha_2 a_{12}a_{21},\cdots,1-\alpha_n a_{1,n}a_{n,1})=I-D_S>0$

$$L_\alpha=\begin{pmatrix} 0 & 0 & \Lambda & 0 \\ (\alpha_2-1)a_{21} & 0 & \Lambda & 0 \\ M & M & O & M \\ (\alpha_n-1)a_{n,1} & \alpha_n a_{12}a_{n,1}-a_{n,2} & \Lambda & 0 \end{pmatrix}=L+L_S-S_\alpha$$

$$U_\alpha=\begin{pmatrix} 0 & -a_{12} & \Lambda & -a_{1,n} \\ 0 & 0 & \Lambda & \alpha_2 a_{1,n}a_{21}-a_{2,n} \\ M & M & O & M \\ 0 & 0 & \Lambda & 0 \end{pmatrix}=U+U_S$$

收稿日期:2011-02-10
作者简介:黄湧辉(1989-　),男,广东揭阳人,本科,研究方向:计算数学。

此时，预条件Gauss-Seidel迭代法的迭代矩阵为

$G_\alpha=(D_\alpha-L_\alpha)^{-1}U_\alpha=[(I-G_S)-(L+L_S-S_\alpha)]^{-1}(U+U_S)$

1 预备知识

定义1[2] 设$Z^{n\times n}=\{A=(a_{ij})_{n\times n}\mid A\in R^{n\times n},a_{ij}\leqslant 0,\forall i,j\in N,i\neq j\}$，则称$Z^{n\times n}$的矩阵A为Z矩阵；设矩阵$A=(a_{ij})_{n\times n}\in R^{n\times n}$，如果对$\forall i=j,a_{ij}\geqslant 0$且$\forall i\neq j,a_{ij}\leqslant 0$，则称矩阵A是L矩阵；如果A是L矩阵，满足$A=sI-B,B\geqslant 0$且谱半径$\rho(B)\leqslant s$，称A是M矩阵。特别地，当$\rho(B)<s$时，称A为非奇异M矩阵；当$\rho(B)=s$时，称A是奇异M矩阵，其中$\rho(B)$为矩阵B的谱半径。

定义2[3] 设矩阵$A=(a_{ij})_{n\times n}\in R^{n\times n}$，若对于所有的$i(1\leqslant i\leqslant n)$都有$|a_{ii}|>\sum_{j=1,j\neq i}^{n}|a_{ij}|$成立，则称矩阵A为严格对角占优的。

定义3[2] 设$A\in R^{n\times n}$，如果M是非奇异n阶矩阵，则称$A=M-N$为矩阵A的分裂。如果$M^{-1}N$是收敛的，则称分裂$A=M-N$是收敛的；如果$M^{-1}\geqslant 0$且$N\geqslant 0$，则称分裂$A=M-N$是正则的；如果$M^{-1}\geqslant 0$且$M^{-1}N\geqslant 0$，则称分裂$A=M-N$是弱正则的。

定义4[3] 设A为Z矩阵且$A^{-1}\geqslant 0$，则称A为非奇异M矩阵。

引理1[4] 设A是非奇异M矩阵，$A=M_1-N_1=M_2-N_2$为A的两个弱正则分裂，如果$M_2^{-1}\leqslant M_1^{-1}$且$N_2\geqslant 0$，则$\rho(M_1^{-1}N_1)\leqslant\rho(M_2^{-1}N_2)$。

引理2[3] 若矩阵$A\in Z^{n\times n}$，那么有如下等价性质：

(1)矩阵A是一个非奇异的M-矩阵；(2)存在向量x，使得$Ax>0$。

引理3 若A是严格对角占优的L-矩阵，则A为非奇异M-矩阵并且$A^{-1}\geqslant 0$。

证明：因为A是严格对角占优的，则对所有的$i(1\leqslant i\leqslant n)$都有$|a_{ii}|>\sum_{j=1,j\neq i}^{n}|a_{ij}|$成立，即有$|a_{ii}|-\sum_{j=1,j\neq i}^{n}|a_{ij}|>0$。记$a=(1,1,\Lambda,1)$，则$Aa=(\sum_{j=1}^{n}a_{1j},\Lambda,\sum_{j=1}^{n}a_{nj})^T$。又因为A是L-矩阵，则对$\forall i=j,a_{ij}\geqslant 0$且$\forall i\neq j,a_{ij}\leqslant 0$因而有$\sum_{j=1}^{n}a_{ij}=a_{ii}-\sum_{j=1,j\neq i}^{n}a_{ij}=|a_{ii}|-\sum_{j=1,j\neq i}^{n}|a_{ij}|>0$。

所以$Aa>0$。因而，由引理2得A为非奇异的M-矩阵。由定义4知$A^{-1}\geqslant 0$。证毕

引理4 当系数矩阵A是严格对角占优的L-矩阵，则预处理后的矩阵A_α为严格对角占优的M-矩阵。

证明：设预处理后得到的矩阵$A_\alpha=(\tilde{a}_{ij})_{n\times n}$。由于矩阵A是严格对角占优的L-矩阵，则当$i\neq j$时，有$-1<a_{ij}\leqslant 0$，所以有$0\leqslant a_{1j}a_{j1}<1$，从而$1-a_{1j}a_{j1}>0$，$(j=2,\Lambda,n)$所以矩阵$A_\alpha$的对角元部分$D_\alpha=diag(1,1-\alpha_2a_{12}a_{21},\cdots,1-\alpha_na_{1,n}a_{n,1})>0$。

而对于矩阵A_α的非对角元部分有

$$\tilde{a}_{ij}=\begin{cases}a_{1j} & i=1,j=2,\Lambda,n\\ a_{i1}-\alpha_ia_{i1} & j=1,i=2,\Lambda,n\\ a_{ij}-\alpha_ia_{1j}a_{i1} & \text{其他}\end{cases}$$

由于$a_{1j}\leqslant 0$，$a_{i1}-\alpha_ia_{i1}\leqslant 0$，$a_{ij}-\alpha_ia_{1j}a_{i1}\leqslant a_{ij}\leqslant 0$，矩阵$A_\alpha$的非对角元部分为非正的。

所以矩阵A_α为L-矩阵。记$a=(1,1,\Lambda,1)$，则$Aa=(\sum_{j=1}^{n}a_{1j},\Lambda,\sum_{j=1}^{n}a_{nj})^T>0$。

由于$P_\alpha=I+S_\alpha\geqslant 0$，则$A_\alpha=P_\alpha Aa>0$。 证毕

引理5[4] 设$A=(a_{ij})\in Z^{n\times n}$为非奇异M矩阵且$B\in Z^{n\times n}$，满足$B\geqslant A$，那么有$A^{-1},B^{-1}$都存在且$A^{-1}\geqslant B^{-1}\geqslant 0$。

2 主要结论

定理1 若线性方程组(1)的系数矩阵A是严格对角占优的L-矩阵，用G表示矩阵A的Gauss-Seidel迭代矩阵，G_α表示在预条件$P_\alpha=I+S_\alpha$下系数矩阵A_α的Gauss-Seidel迭代矩阵，则对于$\alpha_i\in(0,1],i=2,3,\Lambda,n$，有$\rho(G_\alpha)\leqslant\rho(G)$成立。

证明 由于矩阵A是严格对角占优的L-矩阵，则A为非奇异M-矩阵并且矩阵A_α为严格对角占优的非奇异M-矩阵。令A_α有形如下分裂：

$A_\alpha=D_\alpha-L_\alpha-U_\alpha=(I-D_s)-(L+L_s-S_\alpha)-(U+U_s)=E_\alpha-F_\alpha$

其中$E_\alpha=(I-D_s)-(L+L_s-S_\alpha)$，$F_\alpha=U+U_s$，由引理4的证明过程知，矩阵$A_\alpha$的对角元为正的，而非对角元部分为非正的，所以$E_\alpha$为严格对角占优的L-矩阵，由引理3得：$E_\alpha^{-1}\geqslant 0$，又$F_\alpha\geqslant 0$，所以有$E_\alpha^{-1}F_\alpha\geqslant 0$。

令$M_\alpha=(I+S_\alpha)^{-1}E_\alpha$，$N_\alpha=(I+S_\alpha)^{-1}F_\alpha$，由于$M_\alpha^{-1}=E_\alpha^{-1}(I+S_\alpha)\geqslant 0$，且$M_\alpha^{-1}N_\alpha=E_\alpha^{-1}(I+S_\alpha)(I+S_\alpha)^{-1}F_\alpha=E_\alpha^{-1}F_\alpha\geqslant 0$。

又$A=(I+S_\alpha)^{-1}A_\alpha=(I+S_\alpha)^{-1}(E_\alpha-F_\alpha)=M_\alpha-N_\alpha$，则$A=M_\alpha-N_\alpha$为A的一个弱正则分裂。

令$M=I-L$，$N=U$，则有$A=I-L-U=M-N$。由于矩阵A是严格对角占优的L-矩阵，则M为严格对角占优的L-矩阵，由引理3得：$M^{-1}\geqslant 0$。又$N\geqslant 0$，所以有$M^{-1}N\geqslant 0$。则$A=M-N$也为A的一个弱正则分裂。

由于$M-M_\alpha=(I-L)-(I+S_\alpha)^{-1}E_\alpha=(I-L)-(I-S_\alpha)[(I-D_s)-(L+L_s-S_\alpha)]=(D_s+L_s+S_\alpha^2-S_\alpha D_s-S_\alpha L-S_\alpha L_s)$，又$S_\alpha L=0,S_\alpha D_s=0,S_\alpha^2=0,S_\alpha L_s=0$，因而$M-M_\alpha=D_s+L_s\geqslant 0$，所以有$M\geqslant M_\alpha$，由引理5得$M^{-1}\leqslant M_\alpha^{-1}$，而$N\geqslant 0$。

由引理1得：$\rho(M_\alpha^{-1}N_\alpha)\leqslant\rho(M^{-1}N)<1$。

由于$M^{-1}N=G$,$M_\alpha^{-1}N_\alpha=E_\alpha^{-1}(I+S_\alpha)(I+S_\alpha)^{-1}F_\alpha=E_\alpha^{-1}F_\alpha=G_\alpha$,因而$\rho(G_\alpha)\leqslant\rho(G)<1$。 证毕

定理2 若线性方程组(1)的系数矩阵A是严格对角占优的L-矩阵,设$\gamma\leqslant\beta$,即$\forall i=2,3,\Lambda,n$,$\gamma_i\leqslant\beta_i$其中$\gamma=(\gamma_1,\gamma_2,\Lambda,\gamma_n)$,$\beta=(\beta_1,\beta_2,\Lambda,\beta_n)$且$\gamma_i\in[0,1]$,$\beta_i\in[0,1]$,$\forall i=2,3,\Lambda,n$,分别用$P_\gamma=I+S_\gamma$和$P_\beta=I+S_\beta$预处理矩阵A得到的Gauss-Seidel迭代矩阵表示为G_γ和G_β,则对于任意$\alpha_i\in(0,1]$,$i=2,3,\Lambda,n$,有$\rho(G_\beta)\leqslant\rho(G_\gamma)<1$。

证明 由于矩阵A是严格对角占优的L-矩阵,则A为非奇异M-矩阵并且矩阵A_β,A_γ也是严格对角占优的非奇异M-矩阵。令A_β,A_γ有如下分裂:

$A_\beta=D_\beta-L_\beta-U_\beta=(I-D_{\beta s})-(L+L_{\beta s}-S_\beta)-(U+U_{\beta s})=E_\beta-F_\beta$,

$A_\gamma=D_\gamma-L_\gamma-U_\gamma=(I-D_{\gamma s})-(L+L_{\gamma s}-S_\gamma)-(U+U_{\gamma s})=E_\gamma-F_\gamma$,

其中

$E_\beta=(I-D_{\beta s})-(L+L_{\beta s}-S_\beta)$,$F_\beta=U+U_{\beta s}$,

$E_\gamma=(I-D_{\gamma s})-(L+L_{\gamma s}-S_\gamma)$,$F_\gamma=U+U_{\gamma s}$,

由定理1的证明过程知,$A=E_\beta-F_\beta=E_\gamma-F_\gamma$是矩阵A的两个弱正则分裂。

由于$\gamma\leqslant\beta$,$E_\beta-E_\gamma=[(I-D_{\beta s})-(L-S_\beta-L_{\beta s})]-[(I-D_{\gamma s})-(L-S_\gamma-L_{\gamma s})]=(D_{\gamma s}-D_{\beta s})+(S_\beta-S_\gamma+L_{\gamma s}-L_{\beta s})\leqslant0$,所以有$E_\gamma\geqslant E_\beta$,由引理5得$E_\gamma^{-1}\leqslant E_\beta^{-1}$,而$F_\gamma\geqslant0$。由引理1得:$\rho(E_\beta^{-1}F_\beta)\leqslant\rho(E_\gamma^{-1}F_\gamma)<1$。由于$E_\gamma^{-1}F_\gamma=G_\gamma$,$E_\beta^{-1}F_\beta=G_\beta$,所以有$\rho(G_\beta)\leqslant\rho(G_\gamma)<1$。 证毕

3 数值例子

下面给出一个数值例子来验证本文所得的结论。

设 $$A=\begin{pmatrix}1 & -0.3 & -0.4 & -0.1\\ -0.2 & 1 & -0.3 & -0.2\\ -0.4 & -0.2 & 1 & -0.2\\ -0.1 & -0.3 & -0.5 & 1\end{pmatrix}$$

则A为严格对角占优的L-矩阵。

为了方便计算,接下来我们取$\alpha_2=\alpha_3=\alpha_4$。对$\alpha_i(i=2,3,4)$取不同的值时,$\rho(G)$与$\rho(G_\alpha)$的比较如下表:

α_i	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\rho(G_\alpha)$	0.6347	0.6257	0.6162	0.6061	0.5954	0.5839	0.5716	0.5585	0.5442	0.5287

当$\alpha_2=\alpha_3=\alpha_4=0$时,$\rho(G_\alpha)$即为$\rho(G)$,则$\rho(G)=0.6347$。由上表得:$\rho(G)>\rho(G_\alpha)$。因而,预条件$P_\alpha=I+S_\alpha$能够加快Gauss-Seidel迭代法的收敛速度。

此外,由上表得:当$\alpha_i<\alpha_j(i\neq j)$时,有$\rho(G_{\alpha j})\leqslant\rho(G_{\alpha i})<1$。因而,预条件下Gauss-Seidel迭代法的谱半径是单调下降的。

注释及参考文献:

[1]Kohno T, Kotakemori H, Niki H, Usui M.Improving modified iterative methods for Z-matrices[J].Liner Algebra and Its Application, 1997, 267: 113-123.

[2]张谋成,黎稳.非负矩阵论[M].广州:广东高等教育出版社,1995:43-76.

[3]黄廷祝,杨传胜.特殊矩阵分析及应用[M].北京:科学出版社,2007:33-77.

[4]LI W, SUN W, Modified Gauss-Seidel type methods and Jacobi type methods for Z-matrices[J].Liner Algebra and its Applications, April 4, 2000, 317: 227-240.

The Convergence Analysis for the Preconditioned Gauss-Seidel Iterative Method

HUANG Yong-hui

(*South China Normal University, Guangzhou, Guangdong 510631*)

Abstract: In this paper, the convergence analysis for a new preconditioned Gauss-Seidel iterative method was discussed. If the matrix is the strictly dominant L-matrix, the convergence rate of the preconditioned Gauss-Seidel iterative method is faster than one of the original one. Furthermore, the spectral radius of the reconditioned Gauss-Seidel iterative method is monotonically decreasing. At last, an example was given to confirm the fact that the preconditioned Gauss-Seidel iterative method is better than the original one.

Key words: Strictly dominant L-matrix; Preconditioned iterative method; Spectral radius; Weak regular splitting; Convergence rate