文章编号：1001-2486（2002）02-0073-04

块三对角矩阵的修正型局部块分解预条件※

吴建平[1]，李晓梅[2]

（1. 国防科技大学计算机学院，湖南 长沙　410073；2. 指挥技术学院，北京　101416）

摘　要：利用块三对角阵分解因子构造了一类修正型不完全分解预条件子，分析了该预条件子的存在性及其若干性质。针对从二维 Laplace 算子离散得到的五点差分矩阵，给出了预条件后的实际条件数，结果表明，条件数与矩阵阶数的平方根成正比，并且比例因子随局部分解步长的增大而逐渐减小。具体实现时，考虑了其高效实现方案，并针对从二维 Laplace 算子与系数不连续的二维椭圆型算子离散得到的五点差分矩阵，在主频为 550MHz，内存为 256MB 的微机上作了大量实验，且与其他较有效的预条件方法进行了比较，结果表明该预条件方法效率优于其他测试预条件。

关键词：对称正定矩阵；不完全分解；预条件子

中图分类号：TP301　　　**文献标识码**：A

Modified Preconditioners to Block Tridiagonal Matrices Based on Local Factorization

WU Jian-ping[1]，LI Xiao-mei[2]

（1. College of Computer，National Univ. of Defense Technology，Changsha 410073，China；2. Institute of Command and Technology，Beijing 101416，China）

Abstract：A modified type of preconditioner is constructed with the help of local block factorization of block tridiagonal matrices. Then the existence and the properties are analyzed. For the standard 5 – point matrices，which are derived from the 2 – D Laplace operator，the actual condition numbers of the preconditioned matrices are computed. The result shows that the condition number is proportioned to the square root of the order of the matrix. What's more，the longer the step of the local factorization，the smaller the coefficient is. Then efficient implementations of the preconditioners are focused on and three of them provided. Finally lots of experiments are performed for the constructed preconditioners and the well – known effective ones on the personal computer with main frequency of 550MHz and memory of 256M. The matrices in these experiments include the standard five point ones，and the ones derived from a 2 – D elliptic operator with discontinuous coefficients. The results also show that the preconditioners are more efficient than the other tested ones.

Key words：symmetric positive definite matrix；incomplete factorization；preconditioner

线性方程组的迭代法面临不收敛与收敛速度慢两个主要问题，构造预条件子是解决这些问题的有效途径。不完全分解预条件是串行计算时最有效的方法之一，包括基于因子中非零元结构分析的 $ILU(k)$[1]及基于元素幅度大小的 $ILU(p, \tau)$[2]，对后者还可基于极小舍弃重排[3]进行计算，但开销大。这些预条件方法针对一般的稀疏矩阵，对具体的块三对角矩阵，雷光耀等对对角占优矩阵用阶矩阵的思想作不完全分解[4,5]得到了较有效的预条件子。

1　块三对角阵的修正型局部块分解预条件

考虑由 $m\times m$ 个块组成的三对角矩阵 $A=\mathrm{tridiag}(E(i), T(i), F(i+1))$，其中每个 $E(i)$，$T(i)$，$F(i)$ 都是 $n\times n$ 矩阵。现对 A 作块 LU 分解，分解因子 $L=\mathrm{tridiag}(L(i), D(i), 0)$，$U=\mathrm{tridiag}(0, I, U(i+1))$，其中 $D(i)$，$L(i)$，$U(i)$ 均为 n 阶矩阵，则知 $D(1)=T(1)$，且对 $k\geqslant 2$ 有：

$$\begin{cases} D(k)=T(k)-E(k)D(k-1)^{-1}F(k) \\ L(k)=E(k),\ U(k)=D(k-1)^{-1}F(k) \end{cases} \tag{1}$$

※ 收稿日期：2001-10-08
作者简介：吴建平（1974—），男，博士生。

设对 A 作不完全分解 $A \approx B_A = L_A U_A$，其中

$$L_A = \mathrm{tridiag}(E(i), D_A(i), 0), U_A = \mathrm{tridiag}(0, I, U_A(i+1)) \tag{2}$$

且 $D_A(i)$，$U_A(i)$ 均为 n 阶矩阵，$U_A(i) = (D_A(i-1))^{-1}F(i)$，则 B_A 可作为 A 的预条件子。

局部块分解在计算 $U_A(k)$ 时，假设其前某个 $D(\max(k-l,1))$ 为 $T(\max(k-l,1))$，再利用(1)计算得到，下文称 l 为局部分解步长。本文修改 D_A 的对角元使 $B_A e = Ae$，称此时 B_A 为修正型局部块分解预条件，其中 e 是元素全为1的向量。记修改的 D_A 为 $D_A - \Lambda$，且 $\Lambda = \mathrm{diag}(\Lambda(1), \Lambda(2), \cdots, \Lambda(m))$，其中每个 $\Lambda(k)$ 为 n 阶对角阵。

定理1 设 B_l 是局部分解步长为 l 时的修正型预条件子($l < m$)并设其非奇，则 $(B_l)^{-1}A$ 以1为至少 $(l+1)n + m - l - 1$ 重特征值。

证明 设 $R_l = B_l - A$，则知 $R_l = \mathrm{diag}(R_l(1), \cdots, R_l(m))$，其中 $R_l(k)$ 为 $n \times n$ 矩阵，且当 $k \leqslant l+1$ 时 $R_l(k) = 0$。此外由 $B_l e = Ae$ 知对每个 $k > l+1$，均有 $R_l(k)e = 0$，故0亦为每个 $R_l(k)$ 的至少一重特征值。所以0为 R_l 的至少 $(l+1)n + m - l - 1$ 重特征值，从而由 $(B_l)^{-1}A = I - (B_l)^{-1}R_l$ 知定理成立。

对模型五点差分矩阵 $A = \mathrm{tridiag}(-I, T, -I)$，其中 $T = \mathrm{tridiag}(-1, 4, -1)$，可证明如上定义的预条件子 B_l 存在且 $(B_l)^{-1}A$ 具有最小特征值1，为此先证如下引理：

引理1 设 $t > 2$，$d = (t + (t^2 - 4)^{1/2})/2$ 且 $T_t = \mathrm{tridiag}(-1, t, -1)$，则 $(T_t)^{-1}e < de/(d-1)^2$。

证明 考虑 T_t 的 LDL^{T} 分解，设 $D = \mathrm{diag}(d_1, d_2, \cdots, d_n)$，则 d_i 严格单调下降收敛于 d，计算知当 $j < i$ 时 $L^{-1}(i,j) = (d_{i-1})^{-1}\cdots(d_j)^{-1}$，故 $L^{-1}e < de/(d-1)$，从而 $(T_t)^{-1}e < de/(d-1)^2$。

引理2 设 $Ce = 0$，且 C 的非对角元全不小于0，则 C 的特征值不大于0。

证明 由 $|\lambda - C_{ii}| \leqslant \sum_{j=1\sim n, j\neq i} C_{ij}$ 及 $Ce = 0$ 知 $C_{ii} = -\sum_{j=1\sim n, j\neq i} C_{ij}$，故 $\lambda \leqslant C_{ii} + \sum_{j=1\sim n, j\neq i} C_{ij} = 0$。

引理3 设按公式(1)作模型五点差分矩阵 A 的块 LU 分解，则对任意自然数 k，$D(k)$ 为对称 M 矩阵，$\lambda_{\max}(D(k)^{-1}) < k/(k+1)$ 且 $D(k)^{-1}e < ke/(k+1)$。

证明 (数学归纳法) 1) $D(1) = T$ 为对称 M 矩阵，且由引理1知 $D(k)^{-1}e < e/2$，由此知有 $\lambda_{\max}(D(1)^{-1}) < 1/2$。2) 假设引理对 k 成立，则由 $D(k+1) = T - D(k)^{-1}$ 知 $\lambda_{\max}(D(k+1)^{-1}) < (k+1)/(k+2)$，且 $D(k+1)$ 为对称 M 阵。此外 $D(k+1)^{-1}e = \sum_{j=0\sim+\infty}(T^{-1}D(k)^{-1})^j T^{-1}e < (k+1)e/(k+2)$，故引理3成立。

引理4 设按公式(1)对模型五点差分矩阵 A 作块 LU 分解，则对任意自然数 k，l，都成立不等式 $D(k+l)^{-1}e - D(l)^{-1}e < ke/((l+1)(k+l+1))$。

证明 $D(k+1)^{-1}e - D(1)^{-1}e = D(k+1)^{-1}D(k)^{-1}T^{-1}e$，由引理3有 $D(k+1)^{-1}e - D(1)^{-1}e < ke/(2(k+2))$。又由 $D(k+l)^{-1}e - D(l)^{-1}e = D(k+l)^{-1}(D(k+l-1)^{-1} - D(l-1)^{-1})D(l)^{-1}e$ 知：$D(k+l)^{-1}e - D(l)^{-1}e = D(k+l)^{-1}\cdots D(k+2)^{-1}(D(k+1)^{-1} - D(1)^{-1})D(2)^{-1}\cdots D(l)^{-1}e$

从而 $D(k+l)^{-1}e - D(l)^{-1}e < ke/((l+1)(k+l+1))$。

定理2 对模型五点差分矩阵，当 $l = 0$ 时，对任意自然数 k，$D_A(k)$ 为对称 M 矩阵且 $\Lambda(k)$ 的元素介于0到1之间。从而 B_0 存在且为对称正定矩阵，进而 $\lambda_{\max}((B_0)^{-1}A) = 1$。

证明 (数学归纳法证明前一部分) 1) 由 $D_A(1) = T$ 知其为对称 M 阵，由 $\Lambda(1) = 0$ 知结论成立。2) 设对 k 成立，记 $M = \mathrm{tridiag}(-1, 3, -1)$，则 $\Lambda(k+1)e = (D_A(k))^{-1}e < M^{-1}e$。由引理1知 $M^{-1}e < e$，所以 $\Lambda(k+1) < I$，故 $D_A(k+1) = T - \Lambda(k+1) \geqslant M$ 且其非对角元不大于0，故为对称 M 阵。所以对所有自然数 k，$D_A(k)$ 为对称 M 阵。故 B_0 存在且对称正定。

对任意自然数 k，$R_0(k)e = 0$ 且 $R_0(k) = (D_A(k))^{-1} - \Lambda(k)$，则由引理2知其特征值不大于0。由于 B_0 对称正定，则由 $(B_0)^{-1}R_0$ 相似于 $(B_0)^{-1/2}R_0(B_0)^{-1/2}$ 且后者合同于 R_0 知 $(B_0)^{-1}R_0$ 的特征值不大于为0，从而 $(B_0)^{-1}A = I + (B_0)^{-1}R_0$ 的特征值不大于1，故由定理1知 $\lambda_{\max}((B_0)^{-1}A) = 1$。

定理3 对模型五点差分矩阵 A，当 $l \geqslant 1$ 时，对任意 $k \geqslant l+1$，$D_A(k)$ 为对称 M 矩阵且 $\Lambda(k)$ 的元素介于0到 $1/(l+1)$ 之间。从而 B_l 存在且为对称正定矩阵，进而 $\lambda_{\max}((B_l)^{-1}A) = 1$。

证明 （数学归纳法证明前一部分）1）对 $k=l+1$，由预条件子的构造知显然成立。2）当 $k\geqslant l+1$ 时，设对某个 k 成立，由 $D_A(k)=T-D^{-1}(l)-\Lambda(k)$ 知

$$\Lambda(k+1)e=((D_A(k))^{-1}-D^{-1}(l))e=(D_A(k))^{-1}(D^{-1}(l)-D^{-1}(l-1)+\Lambda(k))D^{-1}(l)e$$

故 $\Lambda(k+1)e=(I-T^{-1}D^{-1}(l)-T^{-1}\Lambda(k))^{-1}T^{-1}(D^{-1}(l)-D^{-1}(l-1)+\Lambda(k))D^{-1}(l)e$。由非负矩阵的性质知矩阵 $T^{-1}D^{-1}(l)+T^{-1}\Lambda(k)$ 的谱半径不大于 $T^{-1}D^{-1}(l)+T^{-1}/(l+1)$，从而小于 $1/2$，所以 $(I-T^{-1}D^{-1}(l)-T^{-1}\Lambda(k))^{-1}e=\sum\limits_{j=0\sim+\infty}(T^{-1}D^{-1}(l)+T^{-1}\Lambda(k))^je<2e$

由引理 3 知 $D(l)^{-1}e<le/(l+1)$，由引理 4 知 $D(l+1)^{-1}e-D(l)^{-1}e<e/(l(l+1))$ 且由假设知 $\Lambda(k)e<e/(l+1)$，从而有 $\Lambda(k+1)e<e/(l+1)$。

另方面 $(D_A(k+1))^{-1}=\sum\limits_{j=0\sim+\infty}(T^{-1}D^{-1}(l)+T^{-1}\Lambda(k+1))^jT^{-1}\geqslant 0$。又由 $D_A(k+1)$ 的特征值全大于 1 知其为对称正定矩阵，则对角元全大于 0，显然非对角元不大于 0 且 $D_A(k+1)$ 非奇异，从而 $D_A(k+1)$ 为对称 M 矩阵，所以定理的前部分对所有 $k\geqslant l+1$ 成立。由此可知 B_l 存在且对称正定。

对所有 $k\geqslant l+2$ 有 $R_l(k)e=0$，又 $R_l(k)=D_A(k)+(D_A(k-1))^{-1}-T=(D_A(k-1))^{-1}-D^{-1}(l)-\Lambda(k)$，从而 $R_l(k)+\Lambda(k)=(D_A(k-1))^{-1}(D^{-1}(l)-D^{-1}(l-1)+\Lambda(k-1))D^{-1}(l)\geqslant 0$，所以 $R_l(k)$ 的非对角元非负，由引理 2 知其特征值不大于 0。类似于定理 2 的证明知 $\lambda_{\max}((B_l)^{-1}A)=1$。

图 1 给出了 $l=0\sim5$ 时模型五点差分矩阵预处理后的条件数，结果表明条件数与矩阵阶数的平方根成正比，并且比例因子随局部分解步长的增大而逐渐减小。

2 修正型局部块 LU 分解预条件的实现

预处理时要解形如 $L_AU_AZ=W$ 的方程组，这可通过解两类方程组得到：1）$L_AZ=W$，2）$U_AZ=W$，所以关键是解 $D_A(k)Z(k)=W(k)$ 形式的方程组。设所有 $F(k)$ 非奇，考虑

$$D_l(k)=T_l(k)-E(k)D_{l-1}^{-1}(k-1)F(k)$$

则知可以利用 F_k 的逆进行求解。对 F_k 为非对角矩阵时，求逆要用近似值以提高性能，例如当 F_k 为三对角矩阵时，可用三对角或五对角近似逆矩阵。

当 $l=0$ 时可利用矩阵分解直接求解 $D_0(k)Z(k)=W(k)$，其中 $D_0(k)=T(k)-\Lambda(k)$ 且对角阵 $\Lambda(k)$ 利用 $\Lambda(k)e=(D_0(k-1))^{-1}e$ 计算得到。

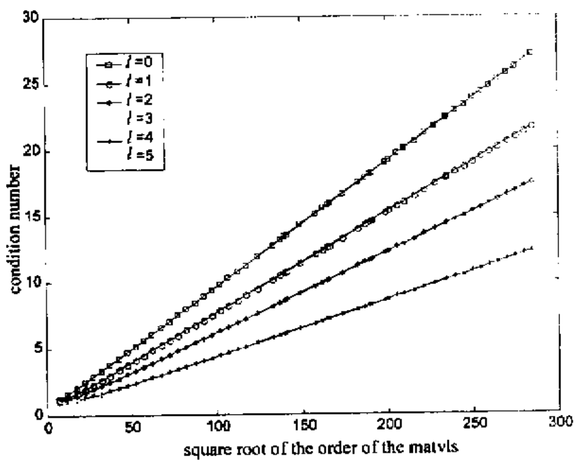

图 1 模型五点差分矩阵预条件后的条件数

Fig.1 Condition numbers

当 $l=1$ 时，对 $k=1$ 时的 $D_1(k)Z(k)=W(k)$ 可利用 $l=0$ 时的方法求解，对 $k\geqslant2$ 可如下求解（其中 $\Lambda(k)$ 利用 $\Lambda(k)e=(D_1(k-1))^{-1}e-D^{-1}(k-l,k-1)e$ 计算得到，$D(k-l,k-1)$）是指假设 $D(k-l)=T(k-l)$ 时再用(1)递推得到的 $D(k-1)$）：

$$\begin{cases}((T(k)-\Lambda(k))F^{-1}(k)T(k-1)-E(k))V(k)=W(k)\\ Z(k)=F^{-1}(k)T(k-1)V(k)\end{cases}$$

当 $l=2$ 时，对 $k=1,2$ 时的 $D_2(k)Z(k)=W(k)$ 可分别利用 $l=0$，$l=1$ 时的方法求解，对 $k\geqslant3$ 可如下求解（其中 $\Lambda(k)$ 利用 $\Lambda(k)e=(D_2(k-1))^{-1}e-D^{-1}(k-l,k-1)e$ 计算得到）：

$$\begin{cases}S_2(k-1)=T(k-1)F^{-1}(k-1)T(k-2)-E(k-1)\\ ((T(k)-\Lambda(k))F^{-1}(k)S_2(k-1)-E(k)F^{-1}(k-1)T(k-2))V(k)=W(k)\\ Z(k)=F^{-1}(k)S_2(k-1)V(k)\end{cases}$$

当 $l\geqslant3$ 时 $D_l(k)Z(k)=W(k)$ 的求解算法可类似得到，不再赘述。

3 数值实验

本节实验如前所述的预条件方法，并与已知的较有效的方法进行比较，其中 CG 是指无预条件 CG

法,DCG为对角预条件CG法,BCG为块预条件CG法,IC(1,1),IC(1,2),IC(1,3)为文[5]中预条件的修正型ICCG法,SSOR1是指采用的SSOR预条件CG法,SSORO是指采用 $\omega=2/(1+\sin(\pi/(n+1)))$ 的SSOR预条件CG法[6,7],其他方法为采用文中前述预条件CG方法,所有实验结果均在主频为550兆赫的微机上得到。

实验1针对系数矩阵为 $n\times n$ 阶模型五点差分矩阵的线性方程组,其中涉及次对角块与向量乘时直接赋值为该向量。实验2中的矩阵由算子

$$-\frac{\partial}{\partial x}A(x,y)\frac{\partial u(x,y)}{\partial x}-\frac{\partial}{\partial y}A(x,y)\frac{\partial u(x,y)}{\partial y}+C(x,y)u(x,y)$$

在区域 $[0,2.1]\times[0,2.1]$ 上用五点差分格式离散得到,其中

$$A(x,y)=\begin{cases}1.0, & x\in[0,1)\text{且 }y\in[0,2.1]\\ 1.0, & y\in[0,1)\text{且 }x\in[0,2.1]\\ 2.0, & x\in[1,2]\text{且 }y\in[1,2]\\ 3.0, & x\in(2,2.1]\text{且 }y\in[1,2.1]\\ 3.0, & y\in[2,2.1]\text{且 }x\in[1,2.1]\end{cases},\qquad C(x,y)=\begin{cases}0.02, & x\in[0,1)\text{且 }y\in[0,2.1]\\ 0.02, & y\in[0,1)\text{且 }x\in[0,2.1]\\ 0.03, & x\in[1,2]\text{且 }y\in[1,2]\\ 0.05, & x\in(2,2.1]\text{且 }y\in[1,2.1]\\ 0.05, & y\in(2,2.1]\text{且 }x\in[1,2.1]\end{cases},$$

表1,2分别列出了这两个实验的部分计算结果,其中iters为迭代次数,time为以秒为单位的包括预条件生成在内的总执行时间。所有迭代中终止条件为残量范数减少到小于初始时的 ε 倍,此处范数由内积 $(.,.)_B$ 导出,其中 B 为预条件子。右端项利用假定的真实解计算得到,该真实解中取第 i 个分量为 $x(i)=i^2/n^2$。

表1 在主频为550兆赫的微机上的实验1($\varepsilon=1E-10$,$n=800$)

Tab.1 Experiment No.1 on a PC with main frequency of 550MHz ($\varepsilon=1E-10$, $n=800$)

	CG	DCG	BCG	SSOR1	SSORO	IC(1,1)	IC(1,2)	IC(1,3)	$l=0$	$l=1$	$l=2$	$l=3$
iters	2544	2544	1796	911	136	146	102	84	52	47	42	39
time	791.0	928.1	758.8	488.3	73.76	77.23	55.32	51.02	30.22	36.82	40.15	47.60

表2 在主频为550兆赫的微机上的实验2($\varepsilon=1E-10$,$n=800$)

Tab.2 Experiment No. 2 on a PC with main frequency of 550MHz ($\varepsilon=1E-10$, $n=800$)

	CG	DCG	BCG	SSOR1	SSORO	IC(1,1)	IC(1,2)	IC(1,3)	$l=0$	$l=1$	$l=2$	$l=3$
iters	8928	5598	3962	1976	1074	237	186	175	133	118	103	92
time	3148.	2312.	1872.	1179.	634.1	135.8	108.4	113.7	93.04	102.5	105.6	115.9

从以上实验可以看出,文中预条件方法优于已有方法,且就计算效率而言,通常取 $l=0$ 为最佳。

4 小 结

利用块三对角阵分解因子构造了一类修正型不完全分解预条件子,分析了该预条件子的存在性及其特征性质。针对模型五点差分矩阵,给出了预条件后的实际条件数,结果表明,条件数与矩阵阶数的平方根成正比,并且比例因子随局部分解步长的增加而逐渐减小。最后考虑了其高效实现并作了大量实验,结果表明该预条件方法优于其他测试方法。

(下转第100页)

一般地，关于 D 的 Bayes 估计的递推公式为

$$\hat{D}_k = \frac{\alpha_k}{\beta_k - 1}$$

式中 α_k，β_k 由递推公式计算：

$\alpha_k = \alpha_{k-1} + \Delta\alpha_k, \beta_k = \beta_{k-1} + \Delta\beta_k \quad k = 1$，…，$N$，$N$ 为截尾试验批数。$\Delta\alpha_k = \frac{n_k}{2}(S^{(k)})^2 + \frac{n_k(\bar{X}^{(k)} - \bar{X}^{(k-1)})^2}{n_k\eta_{k-1} + 1}, \Delta\beta_k = \frac{n_k}{2}$；

式中

$$\eta_k = \frac{1}{n_k}, (S^{(k)})^2 = \frac{1}{n_k}\sum_{i=1}^{n_k}(X_i^{(k)} - \bar{X}^{(k)})^2, \bar{X}^{(k)} = \frac{1}{n_k}\sum_{i=1}^{n_k}X_i^{(k)}$$

可知 $\hat{D}_k$ 的验后方差为

$$\mathrm{E}[(D - \hat{D}_k)^2 | X^{(k)}] = \frac{\alpha_k^2}{(\beta_k - 1)^2(\beta_k - 2)}$$

上述递推公式就是配合 Bayes 序贯截尾检验所作的 Bayes 递推估计方法。

示例具有特殊性，但上述思想方法具有普遍性，可以运用于其他分布参数的 Bayes 序贯检验和估计。

参考文献：

[1] Anscombe F J. Sequential estimation [J]. Jour. Roy. Stat. Soc., Series B，15. 1-29. 1953.

[2] 王正向. 导弹单发命中概率小子样问题研究 [J]. 系统工程与电子技术，1993 (3)：27 – 44.

[3] 张金槐. 特小样本下落点密集度鉴定——SPOT 方法的运用 [J]. 飞行器测控技术，1989，(1)：1-6

(上接第 76 页)

参考文献：

[1] Saad Y. Iterative methods for sparse linear systems [M]. Boston：PWS Publication Corporation，1996.

[2] Saad Y. ILUT：A dual threshold incomplete LU factorization [R]. Minnesota：Computer Science Department，University of Minnesota，Unsi – 92 – 38，1992.

[3] D'azevedo E F，Forsyth P A，Tang Wei – pai. Towards a cost-effective ILU preconditioner with higher level fills [R]. Oak Ridge：Mathematical Sciences Section，Oak Ridge National Laboratory，CS-92，1992.

[4] 雷光耀，张石峰. 阶矩阵及其在传统预处理方法中的应用 [J]. 计算物理，1991，8 (2)：196 – 202.

[5] 雷光耀，黄朝晖. ICCG 法误差阵模与条件数的估计 [J]. 计算物理，1996，13 (4)：489 – 495.

[6] 徐树方. 矩阵计算的理论与方法 [M]. 北京：北京大学出版社，1999.

[7] 李庆扬，易大义，王能超. 现代数值分析 [M]. 北京：高等教育出版社，1995.