

非奇异阵特征极值和条件数的近似计算*

雷 光 耀

(北京 中国科学院应用数学研究所, 1990年5月3日收到)

摘 要

本文从共轭梯度法的公式推导出对称正定阵 A 与三对角阵 B 的相似关系, B 的元素由共轭梯度法的迭代参数确定.因此,对称正定阵的条件数计算可以化成三对角阵条件数的计算,并且可以在共轭梯度法的计算中顺带完成.它只需增加 $O(s)$ 次的计算量, s 为迭代次数.这与共轭梯度法的计算量相比是可以忽略的.当 A 为非对称正定阵时,只要 A 非奇异,即可用共轭梯度法计算 $A^T A$ 的特征极值和条件数,从而得出 A 的条件数.对不同算例的计算表明,这是一种快速有效的简便方法.

关键词 对称正定矩阵 共轭梯度法 特征值 条件数

一、前 言

对称正定矩阵条件数 $\text{cond}(A)$ 的计算($\text{cond}(A) = M(A)/m(A)$)或最大特征值 $M(A)$ 和最小特征值 $m(A)$ (本文简称特征极值)的计算,是数值分析中的一个重要问题.有两种传统方法可以计算特征极值,一种是使用标准的特征值计算程序,算出全部特征值,从而得到特征极值.另一种是用幂法分别计算 A 和 A^{-1} 的最大特征值,这两种方法的主要缺点是要占用大量内存来存储 A 的所有元素,并需花费大量机时,在大量使用的中小型计算机上,这类方法最多只能计算数百阶矩阵的特征极值.以Micro-VAX为例,计算 $N=400$ 阶矩阵的特征值要占用几乎全部内存,且需计算数小时.这显然是不方便的.

在大量的科学与工程计算问题中,常常要求解 A 为对称正定阵的代数方程组 $Ax=b$.预处理共轭梯度法(preconditioned conjugate gradient, 简称为PCG)是近十余年来常用的方法之一^[1].其迭代矩阵为 AM^{-1} ,其中 M 为预处理阵^[2].当预处理阵为单位阵时,其迭代矩阵 $AM^{-1}=A$,PCG退化为共轭梯度法.本文从PCG的计算公式推导出 AM^{-1} 与另一个三对角阵 B 之间的相似关系,从而使计算大为简化.当PCG迭代达到给定精度时,对 B 的低阶子阵作特征值计算,即可得到 AM^{-1} 的条件数和特征极值的近似值(相对误差小于 10^{-4}).大量计算表明,迭代次数在8~15次左右即可得到四位有效数字的结果.本文应用这一方法对二维Poisson方程差分阵、高阶ICCG和MICCG的迭代矩阵以及块PCG的迭代矩阵,都得了令人满意的结果.

* 泰元勋推荐.

国家自然科学基金资助的课题.

$$(AM^{-1})R_s = R_s B_s \quad (2.12)$$

其中 R_s 是由 r_0, r_1, \dots, r_{s-1} 为列向量的 $N \times s$ 矩阵。当 $\|r_s\|$ 足够小时，可以用 B_s 的特征极值和条件数分别作为 AM^{-1} 的特征极值和条件数的近似值。在大量计算问题中， $\|r_s\|$ 随 s 增大而迅速减小。因而在 $s \ll N$ 时即能得到相当精确的结果。 B_s 的特征极值计算远比 AM^{-1} 的简单。它可以作为副产品在 PCG 计算中“捎带”完成。而仅仅增加 $O(s)$ 的计算量和存储量。这与 PCG 的计算量和存储量相比是微不足道的。PCG 的计算可以采用稀疏矩阵存储的方式。在中小型计算机上可以完成 $N=16384$ 阶矩阵问题的计算。

这种方法的另一个优点是，它不需要明显地算出 M^{-1} 和 AM^{-1} ，因而也可以应用于迭代矩阵未显式给出的块 PCG 计算以及类似的计算问题。

应该指出，文[3]曾在1986年给出了与(2.12)类似的公式和部分计算结果，但其计算公式和计算结果均有错误。

三、数值结果

本文采用上述方法对三种问题的特征极值和条件数进行了近似计算，所有结果都是在 Micro-VAX 机上用双精度 FORTRAN 程序得到的，其中三对角阵的特征值计算程序引自[4]。

例1 正规5点格式的差分阵。

设矩阵 A 为下述问题的正规5点差分阵，

$$\begin{cases} -u_{xx} - u_{yy} = f(x, y) & x, y \in (0, 1) \\ u = 0 & x = 0, 1 \text{ 或 } y = 0, 1 \end{cases}$$

其中 $f(x, y)$ 由精确解 $x(1-x)y(1-y)e^{xy}$ 决定。若 $h=1/(n+1)$ ，则 A 为块三对角形式，

$$A = \text{tridiag}(A_{i, i-1}, A_{ii}, A_{i, i+1}) \quad (i=1, 2, \dots, n)$$

其中

$$A_{ii} = \text{tridiag}(-1, 4, -1) \quad (i=1, 2, \dots, n)$$

$$A_{i, i+1} = A_{i+1, i} = -I \quad (i=1, 2, \dots, n-1)$$

矩阵 A 的特征值为

$$\lambda_{pq} = 4 \left(1 - \frac{1}{2} \cos p\pi h - \frac{1}{2} \cos q\pi h \right) \quad (p, q=1, 2, \dots, n)$$

从而有

$$\lambda_{\max}(A) = 4(1 + \cos(\pi/(n+1))) \quad (3.1)$$

$$\lambda_{\min}(A) = 4(1 - \cos(\pi/(n+1))) \quad (3.2)$$

用本文方法所得的计算结果以及由(3.1)和(3.2)算出的理论值在表1中给出。其中的迭代次数为 $|c_s - c_{s-1}|/c_s < 2 \times 10^{-5}$ 时的 s 值， c_s 为第 s 次迭代所得的条件数近似值（相对误差在 10^{-4} 到 10^{-6} 内）。由表1可知，只需作数十次 PCG 迭代（取 $M=I$ ）即可得到4至6位有效数字的近似值。2500阶矩阵的 PCG 计算在 Micro-VAX 机上约需半分钟。

例2 ICCG 与 MICCG 的迭代矩阵。

文[1]给出了 ICCG 算法，文[5]在 A 为 M 阵的条件下将 ICCG 算法推广并改进为 MICCG 算法。而文[6]则应用矩阵元素分阶的概念进一步将这两种方法系统化，提出了高阶 ICCG 算法和高阶 MICCG 算法。表2给出了0阶至6阶的 ICCG 和 MICCG 的特征极值

表1

例1的计算值与理论值的比较

N		10×10	20×20	30×30	40×40	50×50
计算值	s	23	41	54	67	77
	$\lambda_{\max}(A)$	7.83788	7.95517	7.97892	7.98759	7.99142
	$\lambda_{\min}(A)$	0.162028	0.0446767	0.0205227	0.0117368	0.00758672
	cond(A)	48.3736	178.061	388.785	680.559	1053.34
理论值	$\lambda_{\max}(A)$	7.83797	7.95532	7.97948	7.98826	7.99241
	$\lambda_{\min}(A)$	0.162028	0.0446767	0.0205227	0.0117368	0.00758668
	cond(A)	48.3742	178.064	388.812	680.617	1053.48

* s为迭代次数, cond(A)为A的条件数.

表2

ICCG和MICCG迭代矩阵的特征极值和条件数

阶p		0	1	2	3	4	5	6
ICCG	s	49	35	29	22	17	14	13
	$\lambda_{\max}(AM^{-1})$	1.204	1.171	1.131	1.122	1.132	1.112	1.083
	$\lambda_{\min}(AM^{-1})$	0.008178	0.02122	0.03135	0.05636	0.09549	0.1326	0.1628
	cond(AM^{-1})	147.2	55.21	36.07	19.91	11.86	8.383	6.650
MICCG	s	29	23	20	17	14	13	12
	$\lambda_{\max}(AM^{-1})$	19.58	10.61	7.778	5.590	4.307	3.500	2.965
	$\lambda_{\min}(AM^{-1})$	1.000	1.001	1.002	1.002	1.003	1.003	1.002
	cond(AM^{-1})	19.58	10.59	7.763	5.578	4.294	3.491	2.958

* s为迭代次数, cond(AM^{-1})为 AM^{-1} 条件数.

和条件数的计算结果. 所解的问题与例1相同, 但PCG迭代收敛标准为

$$\|r_s\|_2 / \|r_0\|_2 < 10^{-6}$$

初值为零向量, $u_0=0, h=1/64, N=3969$. 在MICCG计算中取 $\alpha=0$ 和 $\beta=1$, 这时MICCG为考虑行和一致的ICCG.

上述结果表明, 当阶p增高时ICCG和MICCG的迭代矩阵条件数 cond(AM^{-1})单调下降. 除0阶ICCG外, 衰减率在0.53至0.84之间. 值得注意的是, 各阶MICCG的最小特征值 $\lambda_{\min}(AM^{-1})$ 都十分接近于1, 从而使条件数和迭代次数与ICCG相比有较大的减少.

例3 高阶块PCG的迭代矩阵.

对高阶块PCG迭代矩阵的计算结果列于表3. 计算条件与例2相同, $u_0=0$, 但迭代收敛标准为

$$\|r_s\|_2 / \|r_0\|_2 < 10^{-8}$$

取 $h=1/65, N=4096$. p阶块PCG算法中的求逆计算采用p阶近似逆矩阵的算法和公式, 详见文[7]. 块PCG算法并未明显给出迭代矩阵 AM^{-1} . 对其特征极值和条件数的直接计算

表3

块PCG迭代矩阵 AM^{-1} 条件数和特征极值的计算结果

阶p	1	2	3	4	5	6	7	8	9	10
s	125	44	26	19	15	12	10	9	9	8
$\lambda_{\max}(AM^{-1})$	3.547	2.000	1.624	1.460	1.361	1.284	1.220	1.175	1.139	1.104
$\lambda_{\min}(AM^{-1})$	0.004139	0.02149	0.05451	0.09984	0.1511	0.2022	0.2492	0.2894	0.3224	0.3479
cond(AM^{-1})	856.8	93.10	29.80	14.63	9.007	6.353	4.896	4.059	3.532	3.172

* s为迭代次数, cond(AM^{-1})为 AM^{-1} 条件数.

难以进行。但本文的方法可以很方便地应用。由表3可知, $p=1, 2, \dots, 5$ 阶的块 PCG 条件数均明显大于同阶的 ICCG 和 MICCG, 因而迭代次数也较多。但三者每次迭代的计算量大体相同。这表明, 当 $p=1, 2, \dots, 5$ 时, p 阶块 PCG 的计算效率低于 p 阶 ICCG 和 p 阶 MICCG。

四、结 论

若矩阵 A 非奇异但非对称正定, 则可以用关于非对称阵的 CG 法, 即对 $A^T A$ 作 CG 法求解代数方程组 $A^T A x = A^T b$ 。得出 $A^T A$ 的特征极值和条件数, 从而得到 A 的条件数。因此, 上述方法可以用来计算任意非奇异阵的条件数。大量计算结果表明, 当 CG 法或 PCG 法迭代终止标准取 $\|r_k\|_2 / \|r_0\|_2 < 10^{-6}$ 时, 作为副产品的条件数近似值即可达到 4 至 6 位有效数字。因此, 这是一种相当简便而实用的算法。

参 考 文 献

- [1] Meijerink, J. A. and H. A. Van der Vorst, An iterative solution method for linear systems of which the coefficient matrix is a symmetric M -matrix, *Math. Comput.*, 31(137) (1977), 148—162.
- [2] Concus, P., G. H. Golub and D. P. O'Leary, *Sparse Matrix Computations*, Ed. by J. R. Bunch and D. J. Rose, Academic Press, New York (1976), 309—332.
- [3] Wong, Y. S., Preconditioned conjugate gradient methods applied to certain symmetric linear systems, *Intern. J. Computer Math.*, 19 (1986), 177—200.
- [4] 郭富印等,《FORTRAN 算法汇编》, 第三分册, 国防工业出版社, 北京 (1982), 37—38.
- [5] Gustafsson, I., A class of first order factorization methods, *BIT*, 18 (1978), 142—156.
- [6] 雷光耀, ICCG 及 MICCG 的讨论与改进, 应用数学学报 (待发表)。
- [7] Lei, G. Y., Block preconditioned conjugate gradient method for large sparse systems, Technical Report 050, Institute of Applied Math., Academia Sinica, Beijing (1989), 1—18.

On the Approximate Computation of Extreme Eigenvalues and the Condition Number of Nonsingular Matrices

Lei Guang-yao

(Institute of Applied Mathematics, Academia Sinica, Beijing)

Abstract

From the formulas of the conjugate gradient, a similarity between a symmetric positive definite (SPD) matrix A and a tridiagonal matrix B is obtained. The elements of the matrix B are determined by the parameters of the conjugate gradient. The computation of eigenvalues of A is then reduced to the case of the tridiagonal matrix B . The approximation of extreme eigenvalues of A can be obtained as a "by-product" in the computation of the conjugate gradient if a com-

putational cost of $O(s)$ arithmetic operations is added, where s is the number of iterations. This computational cost is negligible compared with the conjugate gradient. If the matrix A is not SPD, the approximation of the condition number of A can be obtained from the computation of the conjugate gradient on $A^T A$. Numerical results show that this is a convenient and highly efficient method for computing extreme eigenvalues and the condition number of nonsingular matrices.

Key words symmetric positive definite matrix, conjugate gradient, eigenvalues, condition number