Serdica

Bulgariacae mathematicae publicationes

Сердика

Българско математическо списание

SERDICA — Bulgaricae
mathematicae publicationes
18 (1992) 293-298

TWO-SIDED METHODS FOR COMPUTING THE ROOTS OF MATRICES

J. A. ZARNAN, M. G. PETKOV

ABSTRACT. The present paper deals with four two-sided methods for calculating the roots of matrices, namely: Chords-Newton, Secants-Newton, Parallel Chords-Newton and Method of successive approximations.

1. Introduction. One of the essential problems of linear algebra is the problem of computing the roots of matrices, i.e. finding $\sqrt[m]{A}$, where A is a given $n \times n$ matrix, and $m \geq 2$ is a given natural number. Let us point out that usually the problem is treated for $m = 2$ and for A being a real symmetric positive definite matrix. Our paper ends with some references on the problem of computing the roots of matrices.

The problems of computing the roots of matrices are connected with iteration methods for solving systems of equations, for treating the problem of stability, etc.

In the present paper, the two-sided methods for calculating the roots of the symmetric positive definite matrices are described. These methods except for their greater capacities for parallelizing have the advantage of allowing the application of easier criteria of stopping and estimating the error of successive approximations.

2. First method (Chords-Newton). Let A be an $n \times n$ real symmetric positive definite matrix and let $m \geq 2$ be a natural number. As known there exists a unique symmetric positive definite matrix $X = \sqrt[m]{A}$ such that

$$X^m = A.$$

Let us calculate X by the following two-sided iteration process (Chords-Newton):

$$B_{k+1} = B_k - (B_k^{m-1} + B_k^{m-2}C_0 + \ldots + B_kC_0^{m-2} + C_0^{m-1})^{-1}(B_k^m - A) \tag{1}$$

$$\frac{C_{k+1} = (p-1)C_k + AC_k^{-m+1}}{p} \tag{2}$$

where $p \geq m$, $k = 0, 1, 2, \dots$ and $B_0 = B_0^T > 0$ and $C_0 = C_0^T > 0$, are chosen so as to commutate with each other and A. Moreover $B_0^m < A$ and $C_0^m > A$ (the inequality $P > 0$ denotes that P is symmetric and positive definite, and $P > Q$ is $P - Q > 0$). Thus with the assumptions made, we have

$$X < C_{k+1} < C_k \Rightarrow X \quad (k = 0, 1, 2, \dots) \tag{3}$$

$$X > B_{k+1} > B_k \Rightarrow X \quad (k = 0, 1, 2, \dots) \tag{4}$$

We are going to prove only (4) since in [8] the convergence of the iteration process (2) is proved.

Let us prove (4) by induction.

Let $B_k A = A B_k$, $B_k C_0 = C_0 B_k$, $B_k = B_k^T > 0$, $B_k^m < A$, i.e. $B_k < X$. Hence we obtain that $B_k < B_{k+1} = B_{k+1}^T$. For $B_{k+1} - X$ we have

$$\begin{aligned} B_{k+1} - X &= B_k - X - (\sum_{s=0}^{m-1} B_k^s C_0^{m-1-s})^{-1} (B_k^m - A) \\ &= B_k - X - (\sum_{s=0}^{m-1} B_k^s C_0^{m-1-s})^{-1} (B_k^m - X^m) \\ &= B_k - X + (X - B_k)(\sum_{s=0}^{m-1} B_k^s C_0^{m-1-s})^{-1} \sum_{s=0}^{m-1} B_k^s X_0^{m-1-s} \\ &< B_k - X + (X - B_k)(\sum_{s=0}^{m-1} B_k^s C_0^{m-1-s})^{-1} \sum_{s=0}^{m-1} B_k^s C_0^{m-1-s} = 0 \end{aligned}$$

i.e. $B_{k+1} < X$ or $B_{k+1}^m < A$. Thus we obtain that B_k is an increasing sequence of symmetric and positive definite matrices which is bounded above by the positive definite matrix X. Therefore B_k is convergent and we obtain X as its limit.

The following inequality may be used as a criterion for stopping

$$\|C_k - B_k\| < \varepsilon$$

where $\varepsilon > 0$ is an arbitrary small number chosen in advance.

The method described has been experimented with $n \times n$ matrices of type

$$A = (I - \alpha w w^T)^m \tag{5}$$

where $w = (1/n^{1/2}, 1/n^{1/2}, \dots, 1/n^{1/2})^T$ and $\alpha \in (0, 1)$, with the following initial approximations $B_0 = (I + A^{-1})^{-1}$, $C_0 = I + A$. Experiments are made with different m, n and α (see item 6).

Note 1. Let us point out that (5) is a symmetric positive definite matrix whose Todd's condition number is

$$\rho(A) = \frac{1}{(1-\alpha)^m}.$$

Note 2. It is preferable to use the following best approximations as initial ones:

$$B_0 = (I + \frac{A^{-1}}{m})^{-1}, \quad C_0 = I + \frac{A}{m}.$$

3. Second method (Secants-Newton). Now the same problem will be solved by the formulas:

$$B_{k+1} = B_k - (\sum_{s=0}^{m-1} B_k^s C_k^{m-1-s})^{-1}(B_k^m - A), \quad k = 1,2,3,\dots \tag{6}$$

$$C_{k+1} = \frac{(p-1)C_k + AC_k^{-m+1}}{p}, \quad k = 1,2,3,\dots \tag{7}$$

with the same assumptions made for the initial approximations.

Again (3) and (4) hold for (6) and (7). We are not going to prove (6). Its proof does not differ essentially from that of the iteration process described in 2.

This method has also been algorithmisized, programmed and experimented with a number of matrices of type (5) with the same initial approximations (see item 6).

4. Third method (parallel Chords-Newton). This method is characterized by the following two iteration processes:

$$B_{k+1} = C_k - (mA)^{-1}(C_k^m - A)$$

$$C_{k+1} = \frac{(p-1)C_k + AC_k^{-m+1}}{p} = C_k - (pC_k^{m-1})^{-1}(C_k^m - A)$$

where it is assumed that $\|A\| \le 1$. Thus the other assumptions for the initial approximations are the same as above and we have $mA \le mA^{\frac{m-1}{m}}$ and

$$X > B_{k+1} > B_k \Rightarrow X;$$

$$X < C_{k+1} < C_k \Rightarrow X.$$

This method has been experimented with matrices of type

$$B = \frac{1}{2}A \tag{8}$$

where A is a matrix of type (5).

Note. If $\|A\| > 1$, we use the method with respect to $B = cA$ where constant c is chosen so that $\|B\| \leq 1$.

5. Fourth method (successive approximations). This method is used for calculating $\sqrt{A}$, $(A = A^T > 0)$ by the method of the successive approximations. It is obtained as follows. We put $X = Y - I$ in the following equation

$$X^2 - A = 0$$

where $\|A\| \leq 1$. Thus we have

$$Y^2 - 2Y + I - A = \mathbf{0}$$

Next we have to find the solution $Y = I + \sqrt{A}$. To find Y we construct the following two sided iteration process:

$$B_{k+1} = (2I - B_k)^{-1}(I - A) \tag{9}$$

$$C_{k+1} = (2I - C_k)^{-1}(I - A) \tag{10}$$

where $k = 0, 1, 2, \ldots$, $B_0 = 0$, $C_0 = I$. For (9) and (10) we show that

$$0 \leq B_k < B_{k+1} \Rightarrow Y = I - \sqrt{A}$$

$$0 \leq C_{k+1} < C_k \Rightarrow Y = I - \sqrt{A}.$$

These experiments have also been carried out with matrices of type (8) (see item 6).

6. Experiments and comparisons. The experiments have been accomplished with matrices of various n,m and condition numbers (see the following table).

We can see from the table that each of the four two-sided methods comprises some estimates for the well conditioned matrices. For $m = 2$ the last method is the slowest and for $m > 2$ (when the last method is excluded), the method of Secants-Newton is the fastest and the method of Chords-Newton is the slowest.

Table

(n,m,ρ)	ε	I	t
(10,2,4)	$\omega.10^{-6}$	12	0.84″
	0	7	0.57″
	0	10	0.63″
	$\omega.10^{-7}$	16	0.84″
(10,3,8)	0	30	1.95″
	$\omega.10^{-6}$	10	1.06″
	$\omega.10^{-7}$	25	1.48″
	—	-	—
(10,5,32)	$\omega.10^{-7}$	35	2.95″
	$\omega.10^{-6}$	6	0.98″
	$\omega.10^{-7}$	30	2.45″
	—	-	—
(20,2,4)	$\omega.10^{-7}$	12	4.67″
	$\omega.10^{-7}$	7	2.48″
	$\omega.10^{-7}$	8	2.82″
	$\omega.10^{-7}$	16	4.80″
(20,3,8)	$\omega.10^{-7}$	30	10.15″
	$\omega.10^{-6}$	10	6.62″
	$\omega.10^{-7}$	25	9.37″
	—	-	—
(20,5,32)	0	35	16.10″
	$\omega.10^{-7}$	6	5.89″
	$\omega.10^{-7}$	30	15.48″
	—	-	—
(30,2,4)	$\omega.10^{-6}$	11	13.49″
	$\omega.10^{-6}$	7	8.62″
	$\omega.10^{-7}$	5	5.67″
	$\omega.10^{-6}$	16	14.80″
(30,3,8)	$\omega.10^{-6}$	30	31.77″
	$\omega.10^{-7}$	10	21.37″
	$\omega.10^{-6}$	25	30.99″
	—	-	—
(30,5,32)	$\omega.10^{-6}$	35	52.25″
	$\omega.10^{-6}$	6	18.79″
	$\omega.10^{-6}$	30	51.13″
	—	-	—
(40,2,4)	$\omega.10^{-7}$	11	31.29″
	$\omega.10^{-7}$	7	19.67″
	$\omega.10^{-7}$	7	17.21″
	$\omega.10^{-7}$	16	33.49″
(40,3,8)	$\omega.10^{-7}$	30	2′36.3″
	$\omega.10^{-7}$	10	51.25″
	$\omega.10^{-7}$	25	1′50.30″
	—	-	—
(40,5,32)	$\omega.10^{-7}$	35	4′2.35″
	$\omega.10^{-7}$	6	44.00″
	$\omega.10^{-7}$	30	3′1.99″
	—	-	—
(50,2,4)	$\omega.10^{-7}$	11	58.52″
	$\omega.10^{-7}$	7	38.11″
	$\omega.10^{-7}$	7	32.66″
	$\omega.10^{-6}$	19	1′14.44″
(50,3,8)	$\omega.10^{-6}$	30	4′1.13″
	$\omega.10^{-6}$	10	1′37.96″
	$\omega.10^{-6}$	25	3′45.77″
	—	-	—
(50,5,32)	$\omega.10^{-6}$	35	7′13.66″
	$\omega.10^{-7}$	6	1′25.88″
	$\omega.10^{-6}$	30	6′12.52″
	—	-	—

There:

– The first column comprises triples (n,m,ρ);

– The second column shows the accuracy ε for the elements of $X = \sqrt[m]{A}$ in positions (1,1) and (1,2), $\omega \in (0,1)$;

– Column I comprises the number of iterations needed for reaching the accuracy ε;

– The last fourth column comprises the time t needed for obtaining the respective result.

Acknowledgement. This paper is partially supported by Contract MM-208/92.

REFERENCES

[1] ALBRECHT, J. Quadratically convergent iteration processes for calculating $\sqrt{A}$ and an inverse. *Numer. Math.*, **32** (1976) 9–15.

[2] ALBRECHT, J. Bemerkungen zu iteration verfahren zur berechnung von A^T und A^{-1}. *Z. angew. math. und mech.*, **57** (1977) T262–T263.

[3] APRILE, G., T. RAIMONDI. Iterative numerical calculation of roots of matrices. *Atti. accad. Sci. Lett. Arti. Palermo*, **36**, 1976-1977, 67–71 (1978).

[4] CROSS, G.W., P. LANCASTER Square roots of complex matrices. Linear and multilinear algebra, 1974, I, 289–293.

[5] HOSKINS, W.D., D.J. WALTON. A faster more stable method for computing the P th roots of positive definite matrices. *Linear Algebra and Appl.*, **26** (1979) 159-163.

[6] LAASONEN, P. On the iterative solution of matrix equation $AX^2 - I = 0$. *Math. Tables and Other Aids Comput.*, **12** (1958) 109–116.

[7] PETKOV, M. Iteration method for finding $\sqrt[m]{A}$. Proceedings of the Conference on Numerical methods and Applications, 1984, Sofia, 1985, 476–482.

[8] PETKOV, M., S. BORSHUKOVA. Modified Newton's method for calculating the roots of matrices. *Annual of the University of Sofia*, 1971–1972 **66** (1974) 341-347.

[9] PULAY, P. An iterative method for the determination of the square root of a positive definite matrix. *Z. angew. math. und mech.*, **46** (1966) 151.

Institute of Mathematics
Acad. G. Bonchev str., bl. 8
1113 Sofia
BULGARIA

Received 7.08.1992