Hindawi Publishing Corporation
International Journal of Mathematics and Mathematical Sciences
Volume 2012, Article ID 634698, 13 pages
doi:10.1155/2012/634698

Research Article

Approximation of the *p*th Roots of a Matrix by Using Trapezoid Rule

Amir Sadeghi and Ahmad Izani Md. Ismail

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia

Correspondence should be addressed to Amir Sadeghi, sadeghi.usm@gmail.com

Received 23 August 2011; Accepted 18 October 2011

Academic Editor: Qing-Wen Wang

The computation of the roots of positive definite matrices arises in nuclear magnetic resonance, control theory, lattice quantum chromo-dynamics (QCD), and several other areas of applications. The Cauchy integral theorem which arises in complex analysis can be used for computing *f(A)*, in particular the roots of A, where A is a square matrix. The Cauchy integral can be approximated by using the trapezoid rule. In this paper, we aim to give a brief overview of the computation of roots of positive definite matrices by employing integral representation. Some numerical experiments are given to illustrate the theoretical results.

1. Introduction

It is well known that contour integrals which form a component of the Cauchy integral theorem have an important role in complex analysis. The trapezoid rule is popular for the approximation of integrals due to its exponential accuracy if particular conditions are satisfied. It has been established in [1] that the trapezoid rule can be used to compute contour integrals to give a powerful algorithm for the computation of matrix functions.

Kellems [1] has studied the case of computing a matrix square root and a matrix exponential function by utilizing the trapezoid rule. In particular, he focused on the matrix exponential e^A and its use in the heat equation. Only a few trapezoid rule points were required for very high accuracy. Davies and Higham [2] have investigated the computation of a matrix-vector product $f(A)b$ without explicitly computing $f(A)$. Their proposed methods were specific to the logarithm and fractional matrix powers which were based on quadrature and solution of an ordinary differential equation initial value problem, respectively. Hale et al. in [3] have presented new methods for the numerical computation of $f(A)$ and $f(A)b$, where $f(A)$ is a function such as $A^{1/2}$ or $\log(A)$ with singularities in $(-\infty, 0]$ and A is a matrix with eigenvalues on or near $(0, \infty)$. The methods in [3] were based on combining contour integrals evaluated using the periodic trapezoid rule with conformal maps involving elliptic functions.

In this paper, we investigate computation of the pth roots of positive definite matrices by utilizing integral representation. Our approach is based on the work of Kellems [1] who compute the square root of positive definite matrix using trapezoid rule. The integral identity will be computed by employing trapezoid rule. We also study some matrix factorization procedure and their application in computing matrix roots using trapezoid rule. The outline of this paper is as follows. In Section 2 we introduce some basic definitions and also integral representation of function of matrices and trapezoid rule. In Section 3 we will obtain some formulas to compute the integral representation of the matrix pth root. Numerical experiments will be discussed in Section 4, and the conclusions will be presented in Section 5.

2. Approximation of the Matrix pth Roots

The Cauchy integral theorem states that the value of $f(a)$ can be evaluated by an integral representation as follows:

$$f(a) = \frac{1}{2\pi i}\oint_{\Gamma} \frac{f(z)}{(z-a)} dz, \tag{2.1}$$

where Γ is a contour in $\mathbb{C}$ such that Γ enclose a and $f(z)$ is an analytic and inside Γ [1]. The generalization of this formula in the matrix case can be presented as

$$f(A) = \frac{1}{2\pi i}\oint_{\Gamma} f(z)(zI - A)^{-1} dz \tag{2.2}$$

and can be defined element by element as follows:

$$f(A) = f_{ij} \Longrightarrow f_{kj} = \frac{1}{2\pi i}\oint_{\Gamma} f(z) e_k^T (zI - A)^{-1} e_j dz, \tag{2.3}$$

where the entries of $(zI-A)^{-1}$ are analytic on Γ and also $f(z)$ is analytic function in the neighborhood of the spectrum of A [4]. Cauchy's integral formula can be simplified by considering Γ to be a circle of radius r centered at some point z_c, defined by $z = z_c + re^{i\theta}$. This substitution gives us the following identity [1]:

$$f(A) = \frac{1}{2\pi}\int_0^{2\pi} f(z)(zI - A)^{-1} re^{i\theta} d\theta. \tag{2.4}$$

Writing $z - z_c = re^{i\theta}$ and substituting into (2.4) gives

$$f(A) = \frac{1}{2\pi}\int_0^{2\pi} (z - z_c) f(z)(zI - A)^{-1} d\theta. \tag{2.5}$$

A primary pth root of a square matrix $A \in \mathbb{C}^{n\times n}$, with a p positive integer, is a solution of the matrix equation $X^p - A = 0$ that can be written as a polynomial of A. If A has ℓ distinct eigenvalues and none of which is zero, then A has exactly p^{ℓ} primary pth roots. This result by (2.2) was obtained where f is any of the p^{ℓ} analytic functions defined on the spectrum of A, such that $f(z)^p = z$ and Γ is a closed contour which encloses the spectrum of A. If A has no nonpositive real eigenvalues, then there exists only one primary pth root whose eigenvalues lie in the sector $S_p = \{z \in \mathbb{C} \setminus \{0\} : |\arg(z)| < \pi/p\}$ [5]. In this paper, we demonstrate the method for $f(A) = A^{1/p}$.

In order to calculate the integral (2.2) accurately, we first split the interval of integration $[a, b]$ into N smaller uniform subintervals and then apply the trapezoidal rule on each of them. The composite trapezoidal rule is as follows [6]:

$$\int_a^b f(x)dx \approx \frac{h}{2}\left(f(x_0) + 2\sum_{j=1}^{N-1} f(x_j) + f(x_N) \right), \tag{2.6}$$

where $h = (b-a)/N$ and $x_j = a + jh$, $j = 1, \ldots, N-1$. Since $f(a) = f(b)$, (2.6) can be reformulated as

$$\int_a^b f(x)dx \approx \frac{b-a}{N}\sum_{j=0}^{N-1} f(x_j). \tag{2.7}$$

For (2.5), let the integrand be the function $g(\theta)$. If we take N equally spaced points on Γ and consider that $f(0) = f(2\pi)$, then (2.7) can be written as

$$f(A) \approx \frac{1}{N}\sum_{j=0}^{N-1} g(\theta_j), \tag{2.8}$$

which is the general formula for the computation of a matrix function when Γ is a circle [1].

The function $f(z) = z^{1/p}$ is analytic (i.e., $f_x = -i f_y$) everywhere in $\mathbb{C}$ except at $z = 0$. Consider a matrix A which has eigenvalues in the unit disk centered at z_c. The contour is a disk of radius r, parameterized as $z = z_c + re^{i\theta}$, and from (2.5) we have

$$A^{1/p} \approx \frac{1}{N}\sum_{j=0}^{N-1} (z_j - z_c) z_j^{1/p} (z_j I - A)^{-1}. \tag{2.9}$$

An important property of the trapezoidal rule approximation is that it has better accuracy than the standard matrix pth root algorithms [1].

Now we suppose the Random matrix and use trapezoid rule with $z_c = 3$ and $r = 2$. Figure 1 shows us the convergence of this method for matrices of dimension 4, 8, 16, 32, and 64. In each case exponential accuracy was found. Since the eigenvalues are well clustered, a few points need to be used. For matrices with larger spectral radius or more scattered eigenvalues, the convergence will be slower. This is consistent with the finding in [1] for the case of square root ($p = 2$).

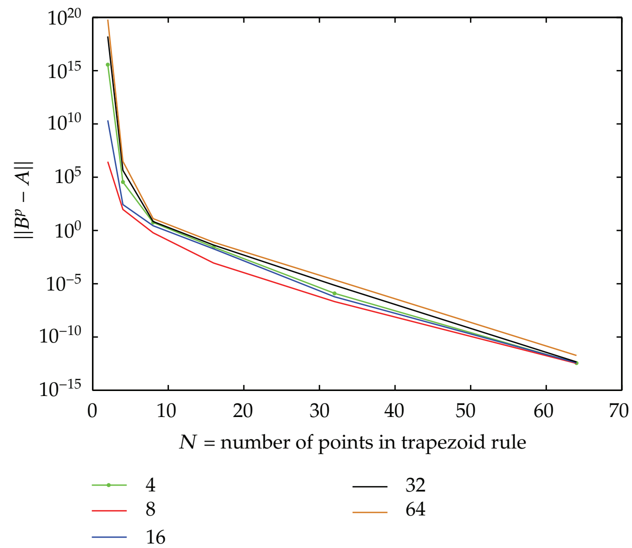


Figure 1: Error in trapezoid rule for computing $A^{1/p}$ as matrix size increases.

3. Employing Matrix Decompositions

Matrix factorizations are utilized to compute the pth root of a matrix using trapezoid rule. Given a square matrix A, we are interested as to the simplest form of matrix B in C or R under unitary similarity transform $A = QBQ^*$ or similarity transform $A = XBX^{-1}$. Matrix B presents some information on A because many features and structure of matrices are invariant under similarity transform. In this part, three factorizations: Schur, Eigenvalue, and Hessenberg are investigated in relation to the use of the trapezoidal rule.

3.1. Schur Decomposition

One of the most applicable factorization of matrices is Schur decomposition which is presented in the following theorem [4].

Theorem 3.1 (Schur decomposition theorem). *Let $A \in \mathbb{C}^{n\times n}$; then there exists a unitary $Q \in \mathbb{C}^{n\times n}$ such that*

$$Q^* AQ = T = D + N, \tag{3.1}$$

where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and N is strictly upper triangular. Further, $Q = \operatorname{diag}(q_1, \ldots, q_n)$ is a column partitioning of the unitary matrix Q where q_i is referred to as Schur vectors, and from $AQ = QT$ Schur vectors satisfy

$$Aq_k = \lambda_k q_k + \sum_{i=1}^{k-1} n_{ik} q_i, \qquad k = 1, \ldots, n. \tag{3.2}$$

One of the most famous algorithms to compute matrix roots is Smith's algorithm proposed in [7]. Generally, this algorithm is presented as follows.

(i) Compute the Schur factorization $A = QTQ^T$.

(ii) Matrix T is upper triangular and so we then set $R_{j,j} = T_{j,j}^{1/p}$.

(iii) Else operate column-by-column on T to produce the upper triangular pth root matrix R.

This algorithm uses $(28 + (p-1)/3)n^3$ flops in total. The matrix pth root is given as $B = QRQ^T$. It can be verified that

$$\begin{aligned} T = R^p &\Longrightarrow QTQ^T = QR^pQ^T \\ &\Longrightarrow A = \left(QRQ^T\right)\dots\left(QRQ^T\right) \\ &\Longrightarrow A = B^p. \end{aligned} \tag{3.3}$$

This can be used to speed up the trapezoid rule method: implement a preliminary factorization of A, operate on the factored matrix, and then combine the factors at the end of the computation [1]. Using the Schur factorization and the unitary of Q, we then have

$$z_j I - A = z_j I - QTQ^* = Q(z_j I - T)Q^*. \tag{3.4}$$

Using this in (2.9) gives

$$A^{1/p} \approx \frac{1}{N} Q \left(\sum_{j=0}^{N-1} (z_j - z_c) z_j^{1/p} (z_j I - T)^{-1} \right) Q^*. \tag{3.5}$$

3.2. Eigenvalue Decomposition

This factorization is also called spectral decomposition and is presented as follows [4].

Theorem 3.2 (Eigenvalue decomposition theorem). *Let $A \in \mathbb{C}^{n\times n}$; there exists a nonsingular $\mathrm{X} \in \mathbb{C}^{n\times n}$ which can diagonalize A*

$$\mathrm{X}^{-1} A X = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \tag{3.6}$$

if and only if the geometric multiplicities of all eigenvalue λ_i are equal to their algebraic multiplicities.

Utilizing the property of X, one has

$$z_j I - A = z_j I - XDX^{-1} = X(z_j I - D)X^{-1}. \tag{3.7}$$

Replacing this into (2.9) *yields*

$$A^{1/p} \approx \frac{1}{N} X \left(\sum_{j=0}^{N-1} (z_j - z_c) z_j^{1/p} (z_j I - D)^{-1} \right) X^{-1}. \tag{3.8}$$

3.3. *Hessenberg Decomposition*

This factorization is also called spectral decomposition and is presented as follows [4].

Theorem 3.3 (Hessenberg decomposition theorem). *Let $A \in \mathbb{R}^{n\times n}$; then there exists a orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ such that*

$$Q^T A Q = H, \tag{3.9}$$

where H is a Hessenberg matrix which means that the elements below the subdiagonal are zero.

Applying the Hessenberg factorization and the orthogonality of Q, one can write

$$z_j I - A = z_j I - QHQ^T = Q(z_j I - H)Q^T. \tag{3.10}$$

Substituting this into (2.9) *will give*

$$A^{1/p} \approx \frac{1}{N} Q \left(\sum_{j=0}^{N-1} (z_j - z_c) z_j^{1/p} (z_j I - H)^{-1} \right) Q^T. \tag{3.11}$$

4. Numerical Experiments

In this section we present some numerical experiment to illustrate the theory which is developed. All the computations have been carried out using MATLAB 7.10(Ra). We assume positive definite matrices with positive nonzero eigenvalues. These matrices, which are given in MATLAB gallery, are used to compute roots of matrices. Recall that if $\widetilde{X}$ is approximated value of $A^{1/p}$ by using different methods, then the absolute error and residual errors can be considered as follows:

$$\begin{aligned} e\left(\widetilde{X}\right) &= \left\|\widetilde{X}^p - A\right\|_F, \\ \operatorname{Res}\left(\widetilde{X}\right) &= \frac{\left\|\widetilde{X}^p - A\right\|_F}{\|A\|_F}, \end{aligned} \tag{4.1}$$

respectively, where $\|\cdot\|_F$ is Frobenius norm.

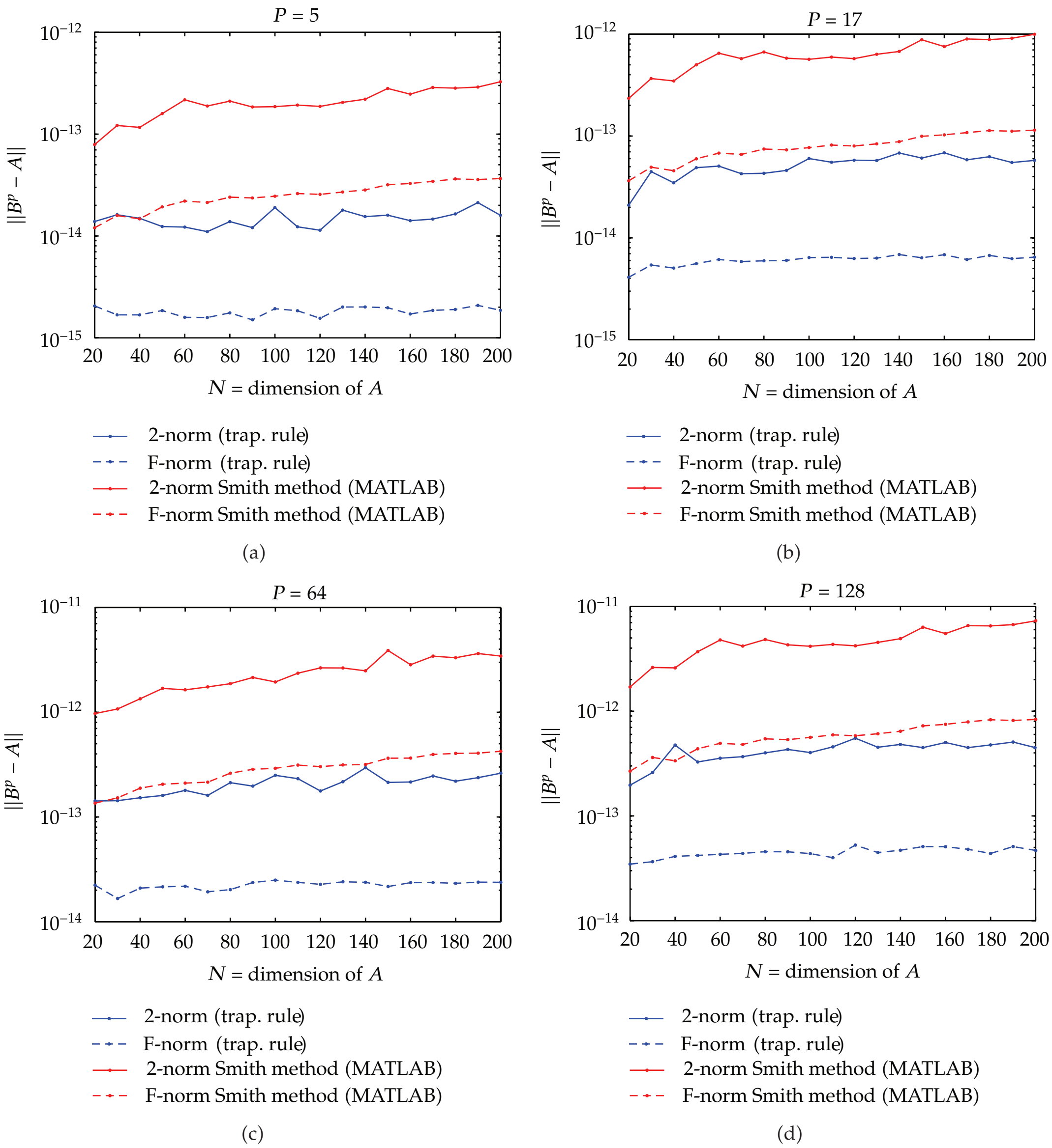


Figure 2: Residual error for computing $A^{1/p}$ for Random matrix.

Test 1

For the first experiment, consider 20×20 Random matrix in the form

$$A = \mathrm{randn}(N)/\mathrm{sqrt}(N) + 3 * \mathrm{eye}(N), \tag{4.2}$$

which has positive eigenvalues. We estimate the pth root of A for different values of p using trapezoid rule. Furthermore, the Schur, eigenvalue, and Hessenberg factorizations are utilized for speeding up the computations. Moreover, the residual error $\mathrm{Res}(\widetilde{X})$ and used time

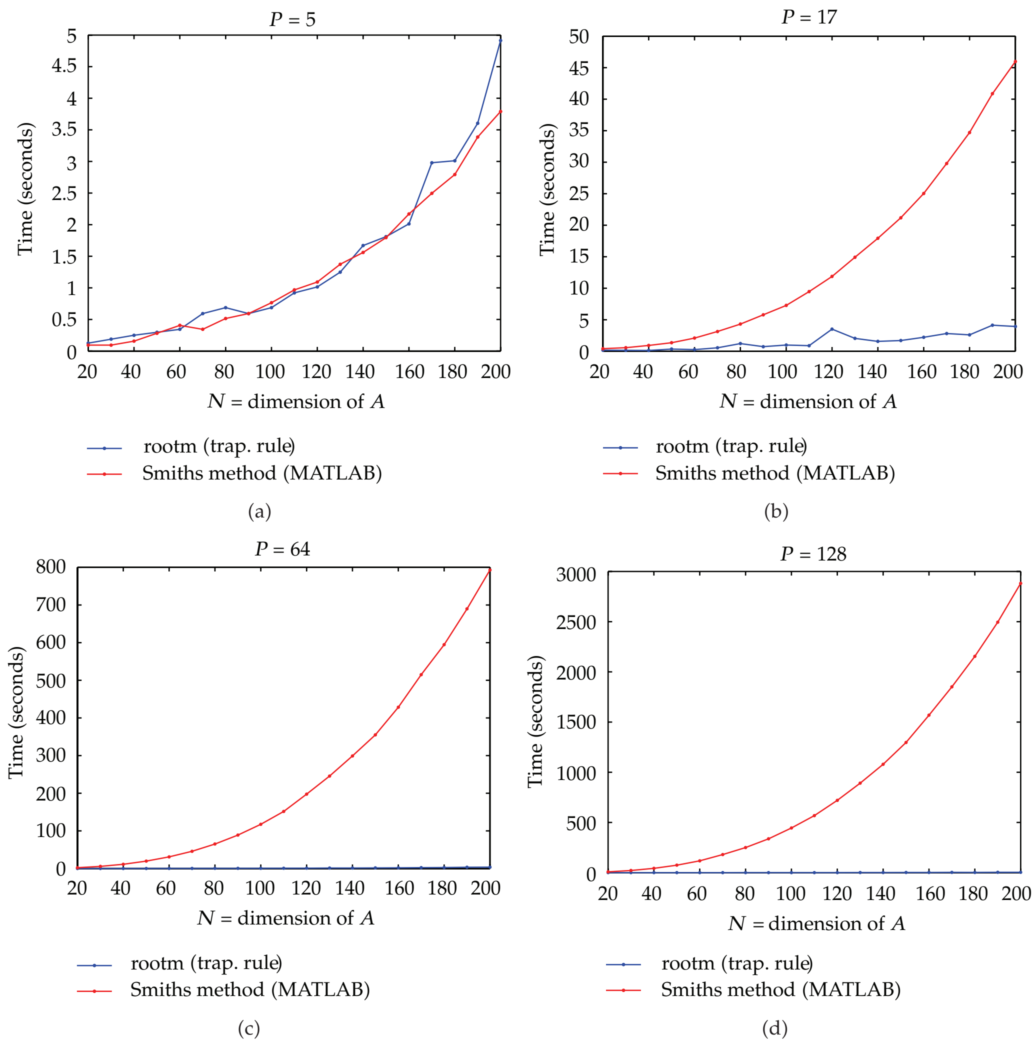


Figure 3: Time in seconds for computing $A^{1/p}$ for Random matrix.

in proposed method for the computation of matrix pth root, for $p = 2, 16, 52, 128$, and 2012, are compared. The results are observed in Tables 1 and 2. It should be mentioned that the number of the point in trapezoid rule is considered as $N = 128$. As can be shown in the result, the Schur factorization has almost exactly the same error as MATLAB's algorithm. The Hessenberg factorization gives the best accuracy among the three methods. In fact, the Hessenberg factorization is quicker than some algorithms in MATLAB. The trapezoid rule for the computation of the matrix pth root may be more effective than several other algorithms but this is dependent on the spectrum of A. This is consistent with the finding in [1] for the case of square root.

Test 2

In this example consider 10×10 Random, Hilbert, and Lehmer matrices. We first fix the value of p and then use trapezoid rule to compute matrix pth root. The relation between the num-

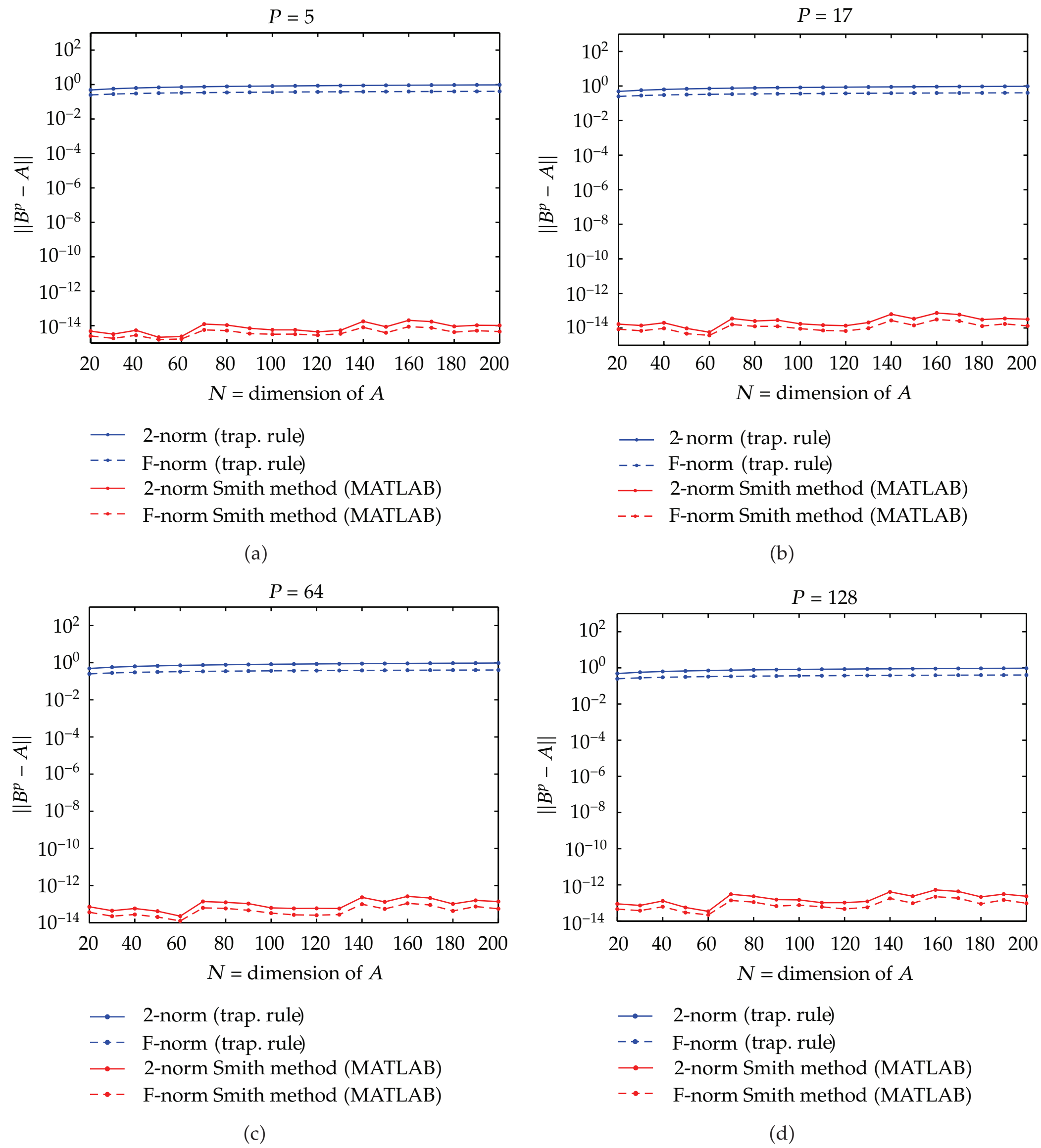


Figure 4: Residual error for computing $A^{1/p}$ for Hilbert matrix.

Table 1: Comparison residual error for computing the pth root of Random matrix.

p	Schur decomposition	Eigenvalue decomposition	Hessenberg decomposition
2	$4.660633743656308e-015$	$5.030320198311745e-015$	$1.554377109452457e-015$
16	$2.311964104013483e-014$	$3.524734357427803e-014$	$7.749850409855744e-015$
52	$8.706212825269891e-014$	$1.121352227197161e-013$	$2.904341268646203e-014$
128	$2.358332191022634e-013$	$2.524139574310145e-013$	$7.105511057327197e-014$
2012	$3.448827113054843e-012$	$3.472268532385117e-012$	$1.170919968312386e-012$

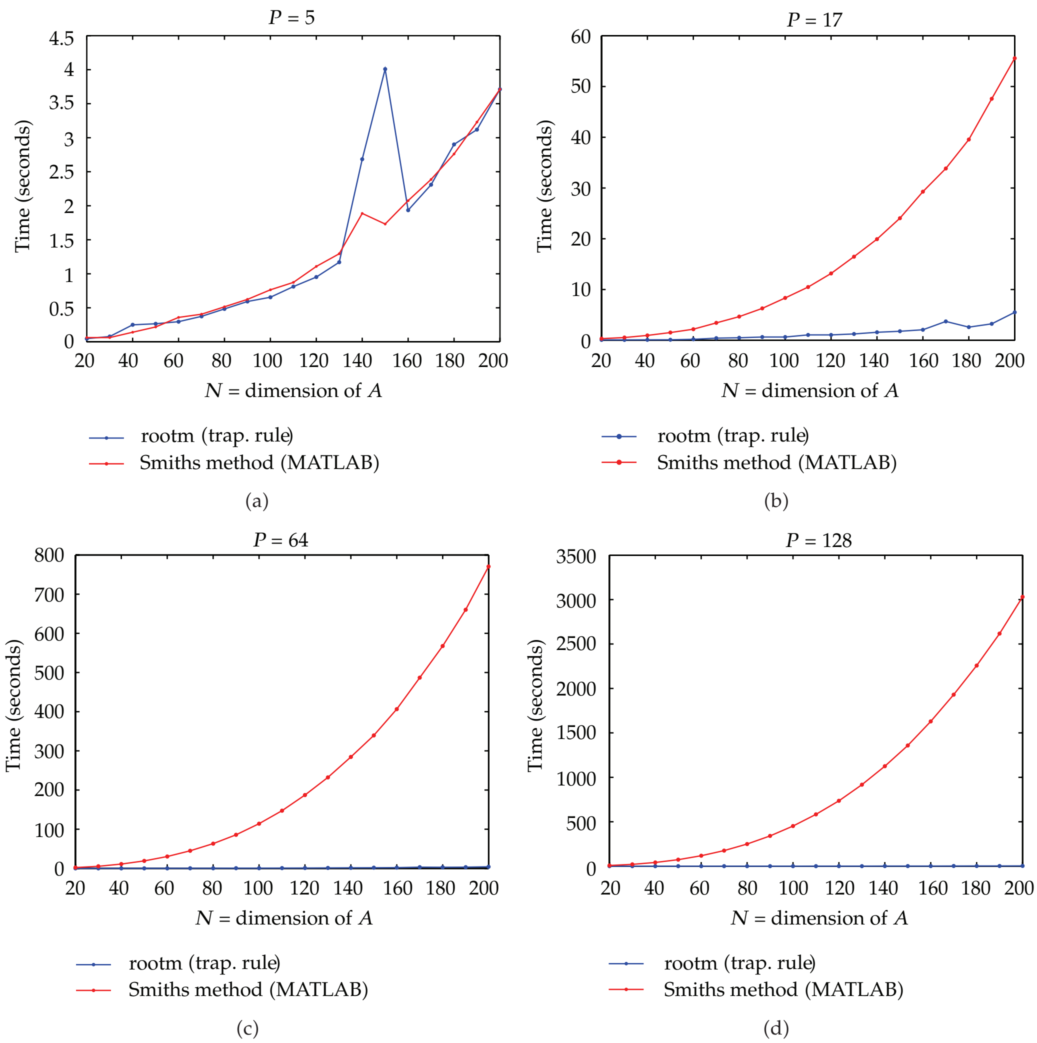


Figure 5: Time in seconds for computing $A^{1/p}$ for Hilbert matrix.

ber of points in trapezoid rule and obtained absolute errors $e(\widetilde{X})$ is investigated. In the implementation, $p = 100$ is supposed and the number of points is increased as $N = 2^{\ell}$, for $\ell = 1, \ldots, 10$. Comparison between errors for various matrices is illustrated in Table 3. It can be seen that errors by increasing the number of points in trapezoid rule will be reduced.

Test 3

In the last test Random matrix, Lehmer matrix, and Hilbert matrix are supposed. We have computed the 5th, 17th, 64th, and 128th root of these matrices using trapezoidal rule and Smith's algorithm. Furthermore, in this example the absolute errors are estimated. As shown in the figures, the accuracy is measured in either the Frobenius norm or the 2-norm for different matrices. The relations between dimension of A and absolute errors and also time in

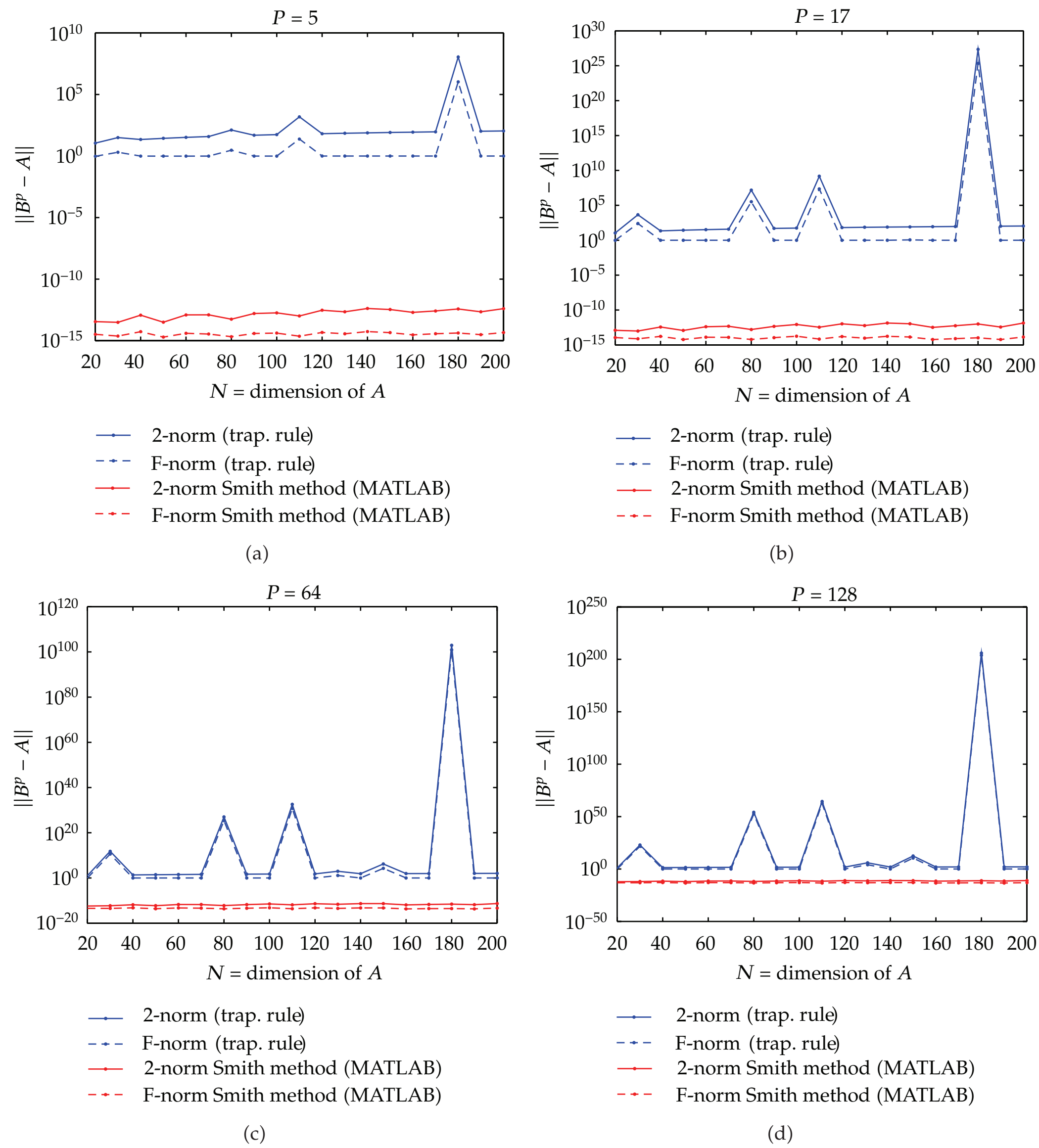


Figure 6: Residual error for computing $A^{1/p}$ for Lehmer matrix.

Table 2: Comparison time in seconds for computing the pth root of Random matrix.

p	Schur decomposition	Eigenvalue decomposition	Hessenberg decomposition
2	0.054855	0.028401	0.017409
16	0.161744	0.226524	0.141455
52	0.404749	0.761385	0.663878
128	0.968970	1.914147	1.803490
2012	0.013331	0.025015	0.018379

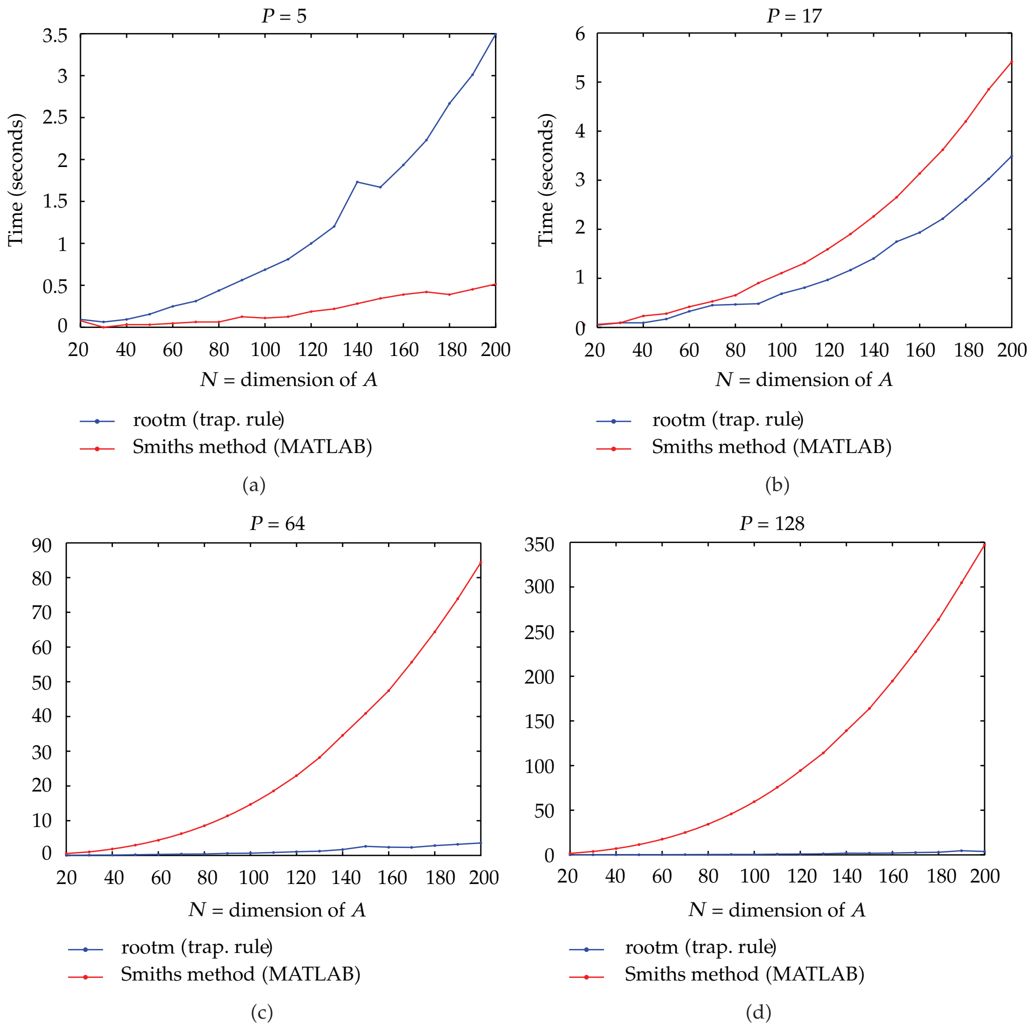


Figure 7: Time in seconds for computing $A^{1/p}$ for Lehmer matrix.

Table 3: Comparison absolute error for different numbers of points for different matrices.

No. of points	Schur decomposition	Eigenvalue decomposition	Hessenberg decomposition
2	4.235476891477593	3.617696711330309	9.018271218363070
4	7.642822555078070	2.268548808462573	85.569768110351674
8	6.509731621793472	16.059530670734521	0.951463218971532
16	0.035910610643974	0.342929548483509	0.001435059307873
32	0.000002770887073	0.342929548483509	0.000000210760699
64	0.000000000000375	0.003429295484835	0.000000000000349
128	0.000000000000002	0.003429295484835	0.000000000000002
256	0.000000000000002	0.003429295484835	0.000000000000002
512	0.000000000000004	0.003429295484835	0.000000000000003
1024	0.000000000000001	0.000342929548484	0.000000000000000

seconds are demonstrated in these figures. Figures 2–7 show the comparison of the residual errors and timings for these methods.

According to Figure 2 which has four parts, the absolute error in all cases ($p = 5, 17, 64$, and 128) in trapezoid rule are smaller than Smith's method. In addition, in Figure 3 it is illustrated that except the case of $p = 5$, the time of computation of the pth root using Smith's method is longer than trapezoidal rule. Also, Figures 4 and 5 are given difference between error and also time in trapezoid rule and Smith's method for Hilbert matrix. The residual error for trapezoid rule is large while for Smith's algorithm is small. For Hilbert matrix except case ($p = 5$), in all cases Smith's method is more time consuming than trapezoid rule. Finally, Figures 6 and 7 show the computation of the pth root of Lehmer matrix which in the most cases reveal that trapezoid rule is more expensive in time and also error than Smith's method by using $N = 128$ points. It must be mentioned that by increasing the number of points, considerable accurate solution can be obtained. For example, using 2^{20} points in trapezoid rule can achieve error of 10^{-6} in the last experiment.

5. Conclusion

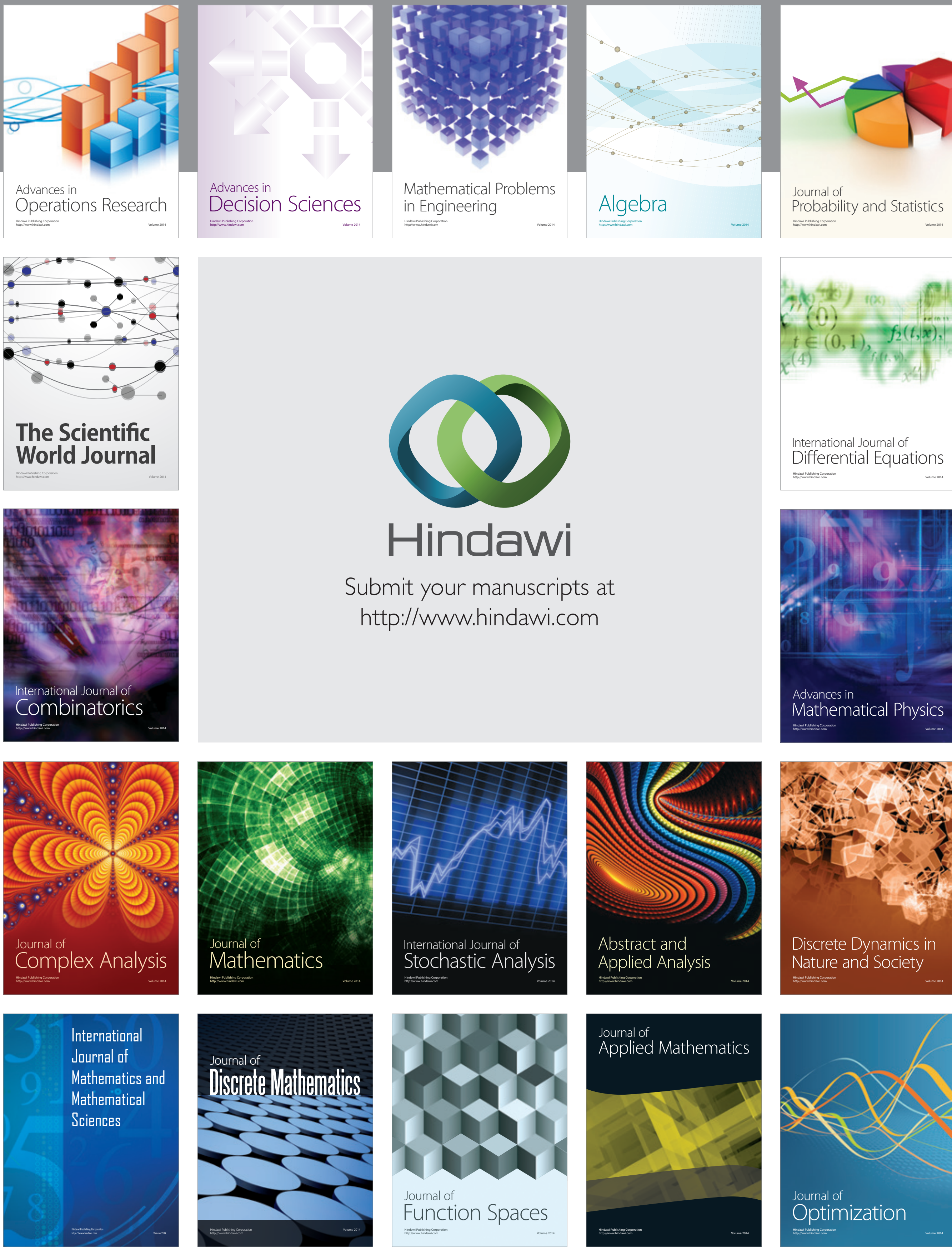
In this paper, we have studied the use of trapezoidal rule in conjunction with the Cauchy integral theorem to compute the pth roots of matrices. It was found that the technique is feasible and accurate.

Acknowledgments

The authors would like to thank Dr. Anthony Kellems and the anonymous referees for their helpful comments which improved the presentation.

References

[1] A. Kellems, "Computing functions of matrices via contour integrals and the Trapezoid rule," *Work Journal*, 2005.
[2] P. I. Davies and N. J. Higham, "Computing *f(A)b* for matrix functions *f*," Tech. Rep. 436, School of Mathematics, University of Manchester, 2004.
[3] N. Hale, N. J. Higham, and L. N. Trefethen, "Computing A^{α}, *log(A)*, and related matrix functions by contour integrals," *MIMS EPrint*, 2007.
[4] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, London, UK, 1996.
[5] F. Greco and B. Iannazzo, "A binary powering Schur algorithm for computing primary matrix roots," *Numerical Algorithms*, vol. 55, no. 1, pp. 59–78, 2010.
[6] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, vol. 12, Springer, New York, NY, USA, 3rd edition, 2002.
[7] M. I. Smith, "A Schur algorithm for computing matrix *p*th roots," *SIAM Journal on Matrix Analysis and Applications*, vol. 24, no. 4, pp. 971–989, 2003.



Advances in Operations Research
Advances in Decision Sciences
Mathematical Problems in Engineering
Algebra
Journal of Probability and Statistics
The Scientific World Journal
Hindawi
Submit your manuscripts at http://www.hindawi.com
International Journal of Differential Equations
International Journal of Combinatorics
Advances in Mathematical Physics
Journal of Complex Analysis
Journal of Mathematics
International Journal of Stochastic Analysis
Abstract and Applied Analysis
Discrete Dynamics in Nature and Society
International Journal of Mathematics and Mathematical Sciences
Journal of Discrete Mathematics
Journal of Function Spaces
Journal of Applied Mathematics
Journal of Optimization