International Journal of Mathematical Archive-5(6), 2014, 1-5
Available online through www.ijma.info ISSN 2229 – 5046

POWER METHOD AND ITS CONVERGENCE FOR APPROXIMATING DOMINANT EIGENVALUE AND ITS CORRESPONDING EIGENVECTOR

Alanur Hussain Laskar and Samira Behera*

Department of Mathematics, Assam University, Silchar-788011, India.

(Received on: 19-05-14; Revised & Accepted on: 03-06-14)

ABSTRACT

In this study, we examine power method for computing the dominant eigenvalue and its corresponding eigenvector of real square matrices. We produced an improvement in the convergence of power method. Our work is based on choosing of initial vector in power method for acceleration purpose. Finally, some examples are presented to illustrate the method and results discussed.

Keywords: *Dominant eigenvalue, power method, convergence of power method.*

1. INTRODUCTION

Eigenvalues and eigenvectors play an important part in the applications of linear algebra. The naive method of finding the eigenvalues of a matrix involves finding the roots of the characteristic polynomial of the matrix. In industrial sized matrices, however, this method is not feasible, and the eigenvalues must be obtained by other means. Fortunately, there exist several other techniques for finding eigenvalues and eigenvectors of a matrix, some of which fall under the realm of iterative methods. These methods work by repeatedly refining approximations to the eigenvectors or eigenvalues, and can be terminated whenever the approximations reach a suitable degree of accuracy. Iterative methods form the basis of much of modern day eigenvalue computation. In this paper, we outline power method, and summarize convergence of power method, derivations, procedures, and advantages. The method to be examined is the power method.

Section 2 of this paper provides a brief review of some of the linear algebra background required to understand the concepts that are discussed. In section 3, we have presented power method and are studied in brief detail with example. In section 4, we produced convergence of power method with numerical examples. Finally, in section 5, we summarized some concluding remarks that are used in practice.

For the purpose of this paper, we restrict our attention to real-valued, square matrices with a full set of real eigenvalues.

2. LINEAR ALGEBRA REVIEW

We begin by reviewing some basic definitions from linear algebra. It is assumed that the reader is comfortable with the notions of matrix and vector multiplication.

Definition 2.1: Let $A \in R^{n\times n}$. A non zero vector $x \in R^n$ is called an eigenvector of A with corresponding eigenvalue $\lambda \in C$ if $Ax = \lambda x.$

Note that eigenvectors of a matrix are precisely the vectors in R^n whose direction is preserved when multiplied with the matrix. Although eigenvalues may not be real in general, we will focus on matrices whose eigenvalues are all real numbers. This is true in particular if the matrix is symmetric.

It is often necessary to compute the eigenvalues of a matrix. The most immediate method for doing so involves finding the roots of characteristic polynomials.

*Corresponding author: Samira Behera**
Department of Mathematics, Assam University, Silchar-788011, India.

Definition 2.2: Let $\lambda_1, \lambda_2, ..., \lambda_n$ be the eigenvalues of an $n \times n$ matrix A. λ_1 is called the dominant eigenvalue of A if

$$|\lambda_1| > |\lambda_i|, i = 2, ..., n.$$

The eigenvectors corresponding to λ_1 are called dominant eigenvectors of A.

Definition 2.3: Eigenvectors corresponding to distinct eigenvalues are linearly independent. However, two or more linearly independent eigenvectors may correspond to the same eigenvalue.

Definition 2.4: An $n \times n$ matrix is diagonalizable if and only if it possesses n linearly independent eigenvectors.

Definition 2.5: Let $A^{(1)}, A^{(2)}, A^{(3)}, ...$ be a sequence of matrices in $R^{m \times n}$. We say that the sequence of matrices converges to a matrix $A \in R^{m \times n}$ if the sequence $A_{i,j}^{(k)}$ of real numbers converges to $A_{i,j}$ for every pair $1 \le i \le m, 1 \le j \le n$, as k approaches infinity. That is, a sequence of matrices converges if the sequences given by each entry of the matrix all converge.

3. DESCRIPTION OF THE POWER METHOD

We saw that the eigenvalues of an $n \times n$ matrix A are obtained by solving its characteristic equation

$$\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + ... + c_0 = 0$$

For large value of n, polynomial equations like this one are difficult and time-consuming to solve. Moreover, numerical techniques for approximating roots of polynomial equations of high degree are sensitive to rounding errors. We look at an alternative method for approximating eigenvalues. Here, the method can be used only to find the eigenvalue of A that is largest in absolute value – we call this eigenvalue the dominant eigenvalue of A. The dominant neigenvalues are of primary interest in many physical applications.

Example 3.1: Let us consider the matrix $A = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}$ for finding dominant eigenvalue and corresponding dominant eigenvectors.

Solution: We know that the characteristic polynomial of A is $\lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2)$. Therefore the eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = -2$, of which the dominant one is $\lambda_2 = -2$. From the same example we know that the dominant eigenvectors of A (those corresponding to $\lambda_2 = -2$) are of the form $X = t\begin{bmatrix} 3 \\ 1 \end{bmatrix}, t \neq 0.$

3.2. THE POWER METHOD

Power method is normally used to determine the largest eigenvalue (in magnitude) and the corresponding eigenvector of the system $AX = \lambda X$. The power method for approximating eigenvalues is iterative. First we assume that the matrix A has a dominant eigenvalue with corresponding dominant eigenvectors. Then we choose an initial approximation X_0 of one of the dominant eigenvectors of A. This initial approximation must be a nonzero vector in R^n.

Finally we form the sequence given by

$$X_1 = AX_0$$
$$X_2 = AX_1 = A(AX_0) = A^2 X_0$$
$$X_3 = AX_2 = A(A^2 X_0) = A^3 X_0$$
$$...$$
$$...$$
$$X_k = AX_{k-1} = A(A^{k-1} X_0) = A^k X_0$$

For large powers of k, and by properly scaling this sequence, we will see that we obtain a good approximation of the dominant eigenvector of A. This procedure is illustrated in example 3.3.

Example 3.3: Let us now consider the matrix $A = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}$ to approximate a dominant eigenvector.

Solution: We begin with an initial approximation $X_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

We then obtain the following approximations.

$$\begin{array}{ll} \quad\quad\quad\quad \mathit{Itereation} & \quad\quad \mathit{Approximation} \\ X_1 = AX_0 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -10 \\ -4 \end{pmatrix} & \rightarrow \; -4\begin{pmatrix} 2.50 \\ 1.00 \end{pmatrix} \\ X_2 = AX_1 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}\begin{pmatrix} -10 \\ -4 \end{pmatrix} = \begin{pmatrix} 28 \\ 10 \end{pmatrix} & \rightarrow \; 10\begin{pmatrix} 2.80 \\ 1.00 \end{pmatrix} \\ X_3 = AX_2 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}\begin{pmatrix} 28 \\ 10 \end{pmatrix} = \begin{pmatrix} -64 \\ -22 \end{pmatrix} & \rightarrow \; -22\begin{pmatrix} 2.91 \\ 1.00 \end{pmatrix} \\ X_4 = AX_3 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}\begin{pmatrix} -64 \\ -22 \end{pmatrix} = \begin{pmatrix} 136 \\ 46 \end{pmatrix} & \rightarrow \; 46\begin{pmatrix} 2.96 \\ 1.00 \end{pmatrix} \\ X_5 = AX_4 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}\begin{pmatrix} 136 \\ 46 \end{pmatrix} = \begin{pmatrix} -280 \\ -94 \end{pmatrix} & \rightarrow \; -94\begin{pmatrix} 2.98 \\ 1.00 \end{pmatrix} \\ X_6 = AX_5 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}\begin{pmatrix} -280 \\ -94 \end{pmatrix} = \begin{pmatrix} 568 \\ 190 \end{pmatrix} & \rightarrow \; 190\begin{pmatrix} 2.99 \\ 1.00 \end{pmatrix} \end{array}$$

Note that the approximation in example 3.3 appear to be approaching scalar multiples of $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$, which we know from example 3.1 is a dominant eigenvector of the matrix $A = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}$.

In example 3.3, the power method was used to approximate a dominant eigenvector of the matrix A. In that example we already knew that the dominant eigenvalue of A was $\lambda = -2$.

4. CONVERGENCE OF POWER METHOD

Theorem 4.1: If A is an $n \times n$ diagonalizable matrix with a dominant eigenvalue, then there exists a nonzero vector X_0 such that the sequence of vectors given by $AX_0, A^2X_0, A^3X_0, \ldots\ldots\ldots, A^kX_0, \ldots\ldots\ldots$ approaches a multiple of the dominant eigenvector of A.

Proof: Since A is diagonalizable, then we know that it has n linearly independent eigenvectors $X_1, X_2, X_3, \ldots\ldots, X_n$ with corresponding eigenvalues of $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots, \lambda_n$. We assume that these eigenvalues are ordered so that λ_1 is the dominant eigenvalue (with a corresponding eigenvector of X_1). Because the n eigenvectors $X_1, X_2, X_3, \ldots\ldots, X_n$ are linearly independent, they must form a basis for R^n. For the initial approximation X_0, we choose a nonzero vector such that the linear combination

$$X_0 = c_1X_1 + c_2X_2 + c_3X_3 + \ldots\ldots\ldots + c_nX_n$$

has nonzero leading coefficients. If $c_1 = 0$, the power method may not converge, and a different X_0 must be used as the initial approximation. Now, multiplying both sides of this equation by A produces

$$\begin{aligned} AX_0 &= A(c_1X_1 + c_2X_2 + ... + c_nX_n) \\ &= c_1(AX_1) + c_2(AX_2) + ... + c_n(AX_n) \\ &= c_1(\lambda_1X_1) + c_2(\lambda_2X_2) + ... + c_n(\lambda_nX_n) \end{aligned}$$

Repeated multiplication of both sides of this equation by A produces

$$A^kX_0 = c_1(\lambda_1^kX_1) + c_2(\lambda_2^kX_2) + ... + c_n(\lambda_n^kX_n),$$

which implies that

$$A^kX_0 = \lambda_1^k[c_1X_1 + c_2(\lambda_2/\lambda_1)^kX_2 + ... + c_n(\lambda_n/\lambda_1)^kX_n]$$

Now, from our original assumption that λ_1 is larger in absolute value than the other eigenvalues it follows that each of the fractions $\lambda_2/\lambda_1, \lambda_3/\lambda_1, ..., \lambda_n/\lambda_1$ is less than 1 in absolute value. Therefore each of the factors $(\lambda_2/\lambda_1)^k, (\lambda_3/\lambda_1)^k, ..., (\lambda_n/\lambda_1)^k$ must approach 0 as k approaches infinity. This implies that the approximation $A^kX_0 \approx \lambda_1^kc_1X_1, c_1 \neq 0$, improves as k increases. Since X_1 is a dominant eigenvector, it follows that any scalar multiple of X_1 is also a dominant eigenvector. Thus we have shown that A^kX_0 approaches a multiple of the dominant eigenvector of A.

N.B.: The proof of the above theorem provides some insight into the rate of convergence of the power method. That is, if the eigenvalues of A are ordered so that

$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq .... \geq |\lambda_n|$, then the power method will converge quickly if $|\lambda_2|/|\lambda_1|$ is small, and slowly if $|\lambda_2|/|\lambda_1|$ is close to 1. This fact is illustrated in example 4.2.

Example 4.2: The rate of convergence of the power method

(a)The matrix $A = \begin{pmatrix} 4 & 5 \\ 6 & 5 \end{pmatrix}$, has eigenvalues of $\lambda_1 = 10$ and $\lambda_2 = -1$. Thus the ratio $|\lambda_2|/|\lambda_1|$ is 0.1. For this matrix, we have seen that only four iterations are required to obtain successive approximations that agree when rounded to three significant digits, as shown in table 1.1.

Table - 1.1.

X_0	X_1	X_2	X_3	X_4
$\begin{bmatrix} 1.000 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.818 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.835 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.833 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.833 \\ 1.000 \end{bmatrix}$

(b) The matrix $A = \begin{pmatrix} -4 & 10 \\ 7 & 5 \end{pmatrix}$, has eigenvalues of $\lambda_1 = 10$ and $\lambda_2 = -9$. For this matrix, the ratio $|\lambda_2|/|\lambda_1|$ is 0.9, and the power method does not produce successive approximations that agree to three significant digits until sixty-eight iterations have been performed, as shown in table 1.2.

Table - 1.2.

X_0	X_1	X_2	……………………..	X_{66}	X_{67}	X_{68}
$\begin{bmatrix} 1.000 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.500 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.941 \\ 1.000 \end{bmatrix}$	…………………….		$\begin{bmatrix} 0.715 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.714 \\ 1.000 \end{bmatrix}$

$$\begin{bmatrix} 0.714 \\ 1.000 \end{bmatrix}$$

From above two tables we conclude that if the ratio $|\lambda_2|/|\lambda_1|$ is small, then the power method will converge quickly and if the ratio $|\lambda_2|/|\lambda_1|$ is close to 1, then the power method will converge slowly.

5. RESULT AND DISCUSSION

We have discussed the use of the power method to approximate the dominant eigenvalue of a matrix. This method can be modified to approximate other eigenvalues through use of a procedure called deflation. Moreover, the power method is only one of several techniques that can be used to approximate the dominant eigenvalue and its corresponding eigenvector of a matrix.

CONCLUSION

The purpose of this paper was to provide an overview of the power method and its rate of convergence used to compute the dominant eigenvalue and its corresponding eigenvector of real-valued square matrices. Here, we used the new initial vector for the power method for the approximation of dominant eigenvalue and its corresponding eigenvector.

Mainly, in this paper we have seen that with examples 4.2(a) and 4.2(b), the rate of convergence of power method is faster if $|\lambda_2|/|\lambda_1|$ is small and slow rate of convergence if $|\lambda_2|/|\lambda_1|$ is close to 1.

REFERENCES

[1] G.H.Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins University Press, 1996.

[2] Kendall E. Atkinson, An Introduction to Numerical Analysis, Second Edition: 1988, John Wiley and Sons Publishers.

[3] C. Moler, G. Stewart, an Algorithm for Generalised Matrix Eigenvalue Problems, SIAM J. Number. Anal., Vol 10, No 2, 1973.

[4] G.W. Stewart, Introduction to Matrix Computations, Academic Press, New York, 1973.

[5] J.B. Fraleigh, A.R.Beauregard, Linear Algebra, Addison-Wesley Publishing Company, 1995.

[6] G.H. Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, 1996.

[7] M.K. Jain, S.R.K. Iyengar, R.K. Jain, Numerical Methods for Scientific and Engineering Computation, Fourth Edition: 2003, New Age International (P) Limited, Publishers.

Source of support: Nil, Conflict of interest: None Declared