CHAPTER III

CHAPTER III

THE METHODS FOR COMPUITING EIGENVALUES AND EIGENVECTORS

(3.1) Introduction

There are two types of methods for computing eigenvalues and eigenvectors namely (i) Iterative methods, (ii) Direct or transformation methods.

The iterative methods are most useful when the matrix is large and sparse and having good estimate of the eigenvector.

The power method, Deflation method, Inverse iteration technique, Simultaneous iteration for real symmetric matrices by Clint and Jennings are some known iterative methods for computing eigenvalues and eigenvectors. Here we discuss the Power method in detail.

(3.2) THE POWER METHOD

The matrix power method is the iterative method which is used to obtain the largest eigenvalue and the corresponding eigenvector of an n×n matrix.

The method is as follows,

Let A be an n×n matrix with linear elementary divisors whose eigenvalues satisfy,

$$|\lambda_1| = |\lambda_2| = |\lambda_3| = \ldots\ldots = |\lambda_r| > |\lambda_{r+1}| \geq \ldots \geq |\lambda_n| \qquad (1)$$

The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ will be refered as the dominant eigenvalues. By assumption there exists n linearly independent eigenvectors $x_1, x_2, \ldots, x_n$ and an arbitrary vector $Z^{(0)}$ can be

expressed in the form

$$Z^{(0)} = \Sigma \alpha_i x_i \qquad (2)$$

where α_i are scalars, not all zero.

Let us define the iterative scheme

$$Z^{(k)} = A\, Z^{(k-1)} \qquad (3)$$

$k = 1, 2, 3, \ldots$

where $Z^{(0)}$ is arbitrary, then

$$Z^{(k)} = AZ^{(k-1)} = A^2 Z^{(k-2)} = \ldots\ldots = A^k Z^{(0)}$$

$$Z^{(k)} = \Sigma \alpha_i \lambda_i^k x_i \qquad (4)$$

Here provided that $\alpha_1, \alpha_2, \alpha_3, \ldots\ldots, \alpha_r$ are not all zero. The right hand side of equation (4) is altimately dominated by the term $\Sigma \alpha_i \lambda_i^k x_i$. In particular if $r = 0$ and we assume that $\alpha_1 \neq 0$, we have

$$Z^{(k)} = \lambda_1^k \left\{ \alpha_1 x_1 + \Sigma \alpha_i (\lambda_i/\lambda_1)\, x_i \right\}$$

$$= \lambda_1^k \left\{ \alpha_1 x_1 + \varepsilon_k \right\} \qquad (5)$$

for k sufficiently large, where ε_k is a vector with very small elements. The vector $Z^{(k)}$ is an approximation to the unnormalized eigenvector x_1 and is accurate if $\|\varepsilon_k\|$ is sufficiently small.

Since $Z^{(k+1)} = \lambda_1^{k+1} \{\alpha_1 x_1 + \varepsilon_{k+1}\}$

It follows that for any i

$$\frac{(Z^{(k+1)})_i}{(Z^{(k)})_i} = \lambda_1 \left[\frac{\alpha_1 (x_1)_i + (\in_{k+1})_i}{\alpha_1 (x_1)_i + (\in_k)_i} \right]$$

$$\longrightarrow \lambda_1 \text{ as } k \longrightarrow \infty$$

The rate of convergence depends on the ratios $|\lambda_2/\lambda_1|$, $|\lambda_3/\lambda_1|, \ldots, |\lambda_n/\lambda_1|$ and if these ratios are smaller the convergence is faster.

Computer Implementation:

The Power method can be implemented in a computer as follows:

Step 1 : Input the matrix A.

Step 2 : Choose the initial vector $Z(0)$ as

$$Z^{(0)} = (1\ 1\ 1\ \ldots.)^T$$

Step 3 : Form $Z^{(p+1)}$ and $U^{(p+1)}$ as

$$U^{(p+1)} = A\, Z^{(p)},$$

$$\beta_{p+1} = \max| U^{(p+1)} |,$$

$$Z^{(p+1)} = 1/\beta_{p+1}\ U^{(p+1)},$$

where $p = 1, 2, 3, \ldots.$

Step 4 : The largest magnitude eigenvalue is the largest magnitude component of $U^{(p+1)}$ and the corresponding standardized eigenvector is $Z^{(p+1)}$ (when p is large).

(3.3) JACOBI METHOD :

Jacobi method is the one of the most stable direct method for computing the complete eigensystem of a real symmetric matrix.

The basis of this algorithm is the similarity transformation, which aims to reduce the original matrix to diagonal form by carrrying out a sequence of plane rotations.

The method is based on following theorems and definition

Theorem (1): All the eigenvalues of a Hermitian matrix are real.

Teorem(2): If an n×n matrix is real symmetric then there exists a real orthogonal matrix R such that $R^{-1}A\,R = D$. Where D is a diagonal matrix.

Unitary matrix: A matrix R is said to be unitary if and only if

$$R^{*} = R^{-1}.$$

The orthogonal matrix is a particular case of unitary matrix.

The aim of Jacobi's algorithm is to reduce to zero the off diagonal elements of a matrix A by means of the transformations.

$$\text{Define} \quad E_{(s)} = \sum_{\substack{j,i=1 \\ i \neq j}}^{n} |a_{ij}^{(s)}|^2 \qquad (1)$$

$$D_{(s)} = \sum_{i=1}^{n} |a_{ii}^{(s)}|^2 \qquad (2)$$

and if we choose the parameteres so that

$$E_{(s+1)} \leq E_{(s)}$$

we have that

$$D_{(s+1)} \geq D_{(s)}.$$

Let A be a n×n real symmetric matrix. Then construct successively the sequence of matrices $A_{(s)}$ such that

$$A_{(s+1)} = R(p,q)A_{(s)}R^T(p,q)$$

Where R(p,q) is a rotation through an angle θ in the plane (p,q). The plane and the rotation angle θ is chosen to ensure that $a_{pq}^{(s+1)}$ is identically zero. Where $a_{pq}^{(s)}$ is taken through the elements of $A_{(s)}$ lying above the diagonal and of maximum modulus.

We know that a similarity transformation using the plane rotation R(p,q), affected the entries in rows and columns p and q only. The modified elements are given by

$$a_{ip}^{(s+1)} = a_{ip}^{(s)} \cos\theta + a_{iq}^{(s)} \sin\theta = a_{pi}^{(s+1)}$$

$$a_{iq}^{(s+1)} = - a_{ip}^{(s)} \sin\theta + a_{iq}^{(s)} \cos\theta = a_{qi}^{(s+1)}$$

$$a_{pp}^{(s+1)} = a_{pp}^{(s)} \cos^2\theta + 2a_{pq}^{(s)} \cos\theta \sin\theta + a_{qq}^{(s)} \sin^2\theta \qquad (3)$$

$$a_{qq}^{(s+1)} = a_{pp}^{(s)} \sin^2\theta - 2a_{pq}^{(s)} \cos\theta \sin\theta + a_{qq}^{(s)} \cos^2\theta$$

$$a_{pq}^{(s+1)} = (a_{qq}^{(s)} - a_{pp}^{(s)}) \cos\theta \sin\theta + a_{pq}^{(s)}(\cos^2\theta - \sin^2\theta)$$

$$= a_{qp}^{(s+1)}.$$

If the rotation angle θ is chosen to annihilate $a_{pq}^{(s+1)}$ we require

$$-(a_{pp}^{(s)} - a_{qq}^{(s)}) \sin 2\theta + 2a_{pq}^{(s)} \cos 2\theta = 0$$

i.e

$$\tan 2\theta = \frac{2a_{pq}^{(s)}}{(a_{pp}^{(s)} - a_{qq}^{(s)})} \qquad (4)$$

the angle θ is chosen so that

$$-\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4}$$

if $a_{pp}^{(s)} - a_{qq}^{(s)} = 0$ then θ is chosen to be

$$\left(a_{pq}^{(s)} / |a_{pq}^{(s)}| \right) \frac{\pi}{4}$$

from equation (3)

$$E_{(s+1)} = E_{(s)} - 2 (a_{pq}^{(s)})^2 \qquad (5)$$

giving

$$D_{(s+1)} = D_{(s)} + 2(a_{pq}^{(s)})^2$$

Then the matrix $A_{(s)}$ will be a diagonal matrix.

that is as $s \rightarrow \infty$,

$$\left| a_{pp}^{(s+1)} - a_{pp}^{(s)} \right| \rightarrow 0$$

From equation (3) we have that

$$a_{pp}^{(s+1)} - a_{pp}^{(s)}$$

$$= -a_{pp}^{(s)}(1 - \cos^2\theta) + 2a_{pq}^{(s)}\cos\theta \sin\theta + a_{qq}^{(s)}\sin^2\theta$$

$$=(a_{qq}^{(s)} - a_{pp}^{(s)})\sin^2\theta + 2\, a_{pq}^{(s)}\cos\theta \sin\theta$$

If $a_{qq}^{(s)} - a_{pp}^{(s)} = 0$

we have $| a_{pp}^{(s+1)} - a_{pp}^{(s)}| = |a_{pq}^{(s)}|$

since $|\theta| = \pi/4$

If we have reached the stage when $|a_{pq}^{(s)}| \leq \varepsilon$

it follows that $| a_{pp}^{(s+1)} - a_{pp}^{(s)}| \leq \varepsilon$

Thus this algorithm generates a sequence of matrices which tend to a fixed diagonal matrix, which is similar to our initial matrix. The diagonal elements are the eigenvalues of the initial matrix.

Calculation of the eigenvectors

An n×n matrix A has exactly n eigenvalues $\lambda_1, \lambda_2, \ldots\ldots \lambda_n$ and corresponding n eigenvectors $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$.

Let $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and

$X = (x^{(1)}, x^{(2)}, \ldots x^{(n)})$ be respectively the diagonal matrix with $\lambda_1, \lambda_2, \ldots, \lambda_n$ as diagonal elements and an n×n matrix with $x^{(i)}$ as its i^{th} column vector. Then from $Ax = \lambda x$ we can write

$$AX = XD$$

Premultiplying both sides by X^{-1} we get

$$X^{-1}AX = D$$

If A is an n×n real symmetric matrix and if we employ the Jacobi method, then $X = R_1R_2R_3 \ldots . R_k$. Thus while computing the eigenvalues we post multiply R_{k-1} by currently generated R_k. The resultant matrix is X, each of whose columns is an eigenvector; in k-th column is the k-th eigenvector of A corrseponding to the k-th eigenvalue i.e λ_k.

Note : Here we have to remember that at each stage in the reduction of a real symmetric matrix to diagonal form one must work with complete matrix. Again the elements reduced to zero in on e step may become nonzero in later step and the process is an infinite, particularly for larger matrix.

(3.4) GIVEN'S METHOD:

This method is applicable to an n×n real symmetric matrix A. It consists in carrying out a sequence of orthogonal transformations on A to produce a tridiagonal form.

In this process an element reduced to zero by one rotation never becomes nonzero in later rotation.

Let

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & c & -s & 0 & \cdots & 0 \\ 0 & s & c & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & & 1 \end{bmatrix}$$

be an orthogonal matrix where c and s are to be determined when A is subjected to the transformation $A_1 = S_1^T A S_1$ then (1,3)th element of A_1 = (3,1)th element of A_1 = $-sa_{12} + ca_{13}$ these vanish if we choose $\tan\theta = s/c = a_{13}/a_{12}$ or equivalently

$$s = a_{13} / (a_{12}^2 + a_{13}^2)^{1/2}$$

$$c = a_{12} / (a_{12}^2 + a_{13}^2)^{1/2}$$

such a transformation can be treated as a rotation in the (2,3) plane. Then the matrix A_1 has the form

$$A_1 = \begin{bmatrix} x & x & 0 & x & \ldots & x & x \\ x & x & x & x & \ldots & x & x \\ 0 & x & x & x & \ldots & x & x \\ x & x & x & x & \ldots & x & x \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ x & x & x & x & \ldots & x & x \end{bmatrix}$$

where x is any general value possibly not equal to zero. Successive rotations, thus carried out in the planes (2,3),(2,4),...,(2,n) where the θ's are so chosen that the new (1,3)rd, (1,4)th,.....(1,n)th elements vanish. After (n-2) such rotations we obtain A_{n-2} which has the form

$$A_{n-2} = \begin{bmatrix} x & x & 0 & x & \ldots & 0 & 0 \\ x & x & x & x & \ldots & x & x \\ 0 & x & x & x & \ldots & x & x \\ 0 & x & x & x & \ldots & x & x \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & x & x & x & \ldots & x & x \end{bmatrix}$$

The second row of A_{n-2} is treated in the same way as the first row. The rotations here are made in the planes (3,4),(3,5),.....(3,n).Thus after (n-1)(n-2)/ 2 rotations ,a symmetric tridiagonal matrix T of the form

$$T = \begin{bmatrix} x & x & 0 & 0 & \ldots & 0 & 0 & 0 \\ x & x & x & 0 & \ldots & 0 & 0 & 0 \\ 0 & x & x & x & \ldots & 0 & 0 & 0 \\ 0 & 0 & x & x & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & x & x & x \\ 0 & 0 & 0 & 0 & \ldots & 0 & x & x \end{bmatrix}$$

is obtained. Here no successive transformation affects the previous zeros. The eigenvalues of T and A are identical since the two matrices are similar.

(3.5) EIGENSYSTEM OF A SYMMETRIC TRIDIAGONAL MATRIX

The Given's transformation gives us a symmetric matrix to a symmetric tridiagonal form. The eigenvalues of such matrices are computed using Strum sequence property and bisection. The method is as follows,

Let T denote a symmetric tridiagonal matrix with elements

$$t_{ii} = d_i$$

$$t_{i,i+1} = t_{i+1,i} = e_i$$

assuming that none of the e_i is zero.

Thus

$$T = \begin{bmatrix} d_1 & e_1 & 0 & 0 & 0 & 0 & 0 \\ e_1 & d_2 & e_2 & 0 & 0 & 0 & 0 \\ 0 & e_2 & d_3 & e_3 & 0 & 0 & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & e_{n-2} & d_{n-1} & e_{n-1} \\ 0 & 0 & 0 & 0 & 0 & e_{n-1} & d_n \end{bmatrix}$$

Let $p_r(\lambda)$ denote the determinant of the leading principal minor of $T - \lambda I$.

Thus

$$p_r(\lambda) = \det \begin{vmatrix} d_1-\lambda & e_1 & 0 & 0 & 0 & 0 \\ e_1 & d_2-\lambda & e_2 & 0 & 0 & 0 \\ . & . & . & . & . & \\ . & . & . & . & . & . \\ 0 & 0 & 0 & e_{r-2} & d_{r-1}-\lambda & e_{r-1} \\ 0 & 0 & 0 & 0 & e_{r-1} & d_r-\lambda \end{vmatrix}$$

We define $p_0(\lambda) = 1$

$$p_1(\lambda) = d_1-\lambda.$$

expanding $p_r(\lambda)$ by the final row the result

$$p_r(\lambda) = (d_r-\lambda)p_{r-1}(\lambda) - e_{r-1}^2 p_{r-2}(\lambda)$$

Since $p_n(\lambda) = \det(T-\lambda I)$

we may compute it by means of the relations

$$p_0(\lambda) = 1$$

$$p_1(\lambda) = d_1-\lambda$$

$$p_r(\lambda) = (d_r-\lambda)p_{r-1}(\lambda) - e_{r-1}^2 p_{r-2}(\lambda), \quad r = 2,3,....n \quad ..(1)$$

Here the zeros of $p_r(\lambda)$ strictly separate those of $p_{r-1}(\lambda)$.

Using the formulas (1) we can evaluate the numbers $p_0(\lambda), p_1(\lambda), ... p_n(\lambda)$ for some value of λ. The number of agreements in sign between successive members of the sequence $\{p_r(\lambda)\}$ is denoted by $s(\lambda)$.

Then using the sturm sequence property we can locate approximate intervals in which each of the eigenvalue lie.

The sturm sequence property : The number of agreements in sign $s(\lambda)$ of successive members of the sequence $\{p_r(\lambda)\}$ is equal to the number of eigenvalues of T which are strictly greater than λ.

Using Sturm sequence property together with the method of bisection allows us to determine any eigenvalue to prescribed accuracy. The first step is to locate an interval $[\lambda_1^{(0)}, \lambda_2^{(0)}]$ in which only the eigenvalue λ lies, by means of sturm sequence procedure.

For some $k<n$ we have

$$s(\lambda_2^{(0)}) = k, \quad s(\lambda_1^{(0)}) = k+1.$$

Let $u^{(0)} = 1/2[\lambda_1^{(0)}+\lambda_2^{(0)}]$ and $s(u^{(0)})$ which must either have the value k or k+1. If $s(u^{(0)}) = k$, then λ lies in the interval $[\lambda_1^{(0)}, u^{(0)}]$; otherwise it is in $[u^{(0)}, \lambda_2^{(0)}]$. This process may be repetitively applied to determine a tight interval for λ.

The above technique is used to find the eigenvalues in aparticular interval, or the first few, or the last few eigenvalues of an $n \times n$ symmetric tridiagonal matrix.