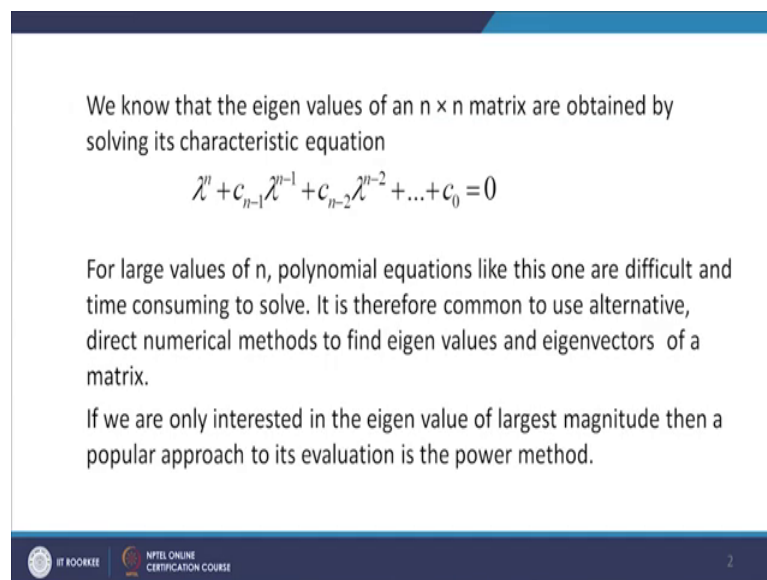


Numerical Linear Algebra
Dr. P. N. Agrawal
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture – 56
Power Method

Hello friends, welcome to my lecture on power method. We know that the eigen values of an n by n matrix are obtained by solving its characteristic equation determinant of $A - \lambda I = 0$. So, when we expand that determinant, we get the polynomial equation in λ ; $\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \dots + c_0 = 0$.

(Refer Slide Time: 00:43)



We know that the eigen values of an $n \times n$ matrix are obtained by solving its characteristic equation

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

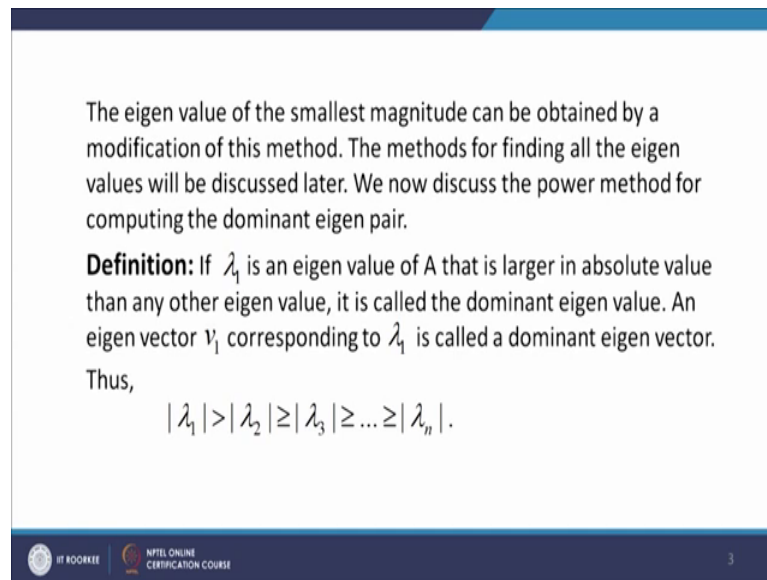
For large values of n , polynomial equations like this one are difficult and time consuming to solve. It is therefore common to use alternative, direct numerical methods to find eigen values and eigenvectors of a matrix.

If we are only interested in the eigen value of largest magnitude then a popular approach to its evaluation is the power method.

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 2

Now, for large values of n polynomial equations like this one are difficult and time consuming to solve. And it is therefore, common to use alternative direct numerical methods to find the eigen values and eigenvectors of a matrix. Now in many cases we need only the eigen value of largest magnitude. So, if we are only interested in the eigen value of largest magnitude, then a popular approach to its evaluation is the power method.

(Refer Slide Time: 01:26)



The eigen value of the smallest magnitude can be obtained by a modification of this method. The methods for finding all the eigen values will be discussed later. We now discuss the power method for computing the dominant eigen pair.

Definition: If λ_1 is an eigen value of A that is larger in absolute value than any other eigen value, it is called the dominant eigen value. An eigen vector v_1 corresponding to λ_1 is called a dominant eigen vector.

Thus,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|.$$

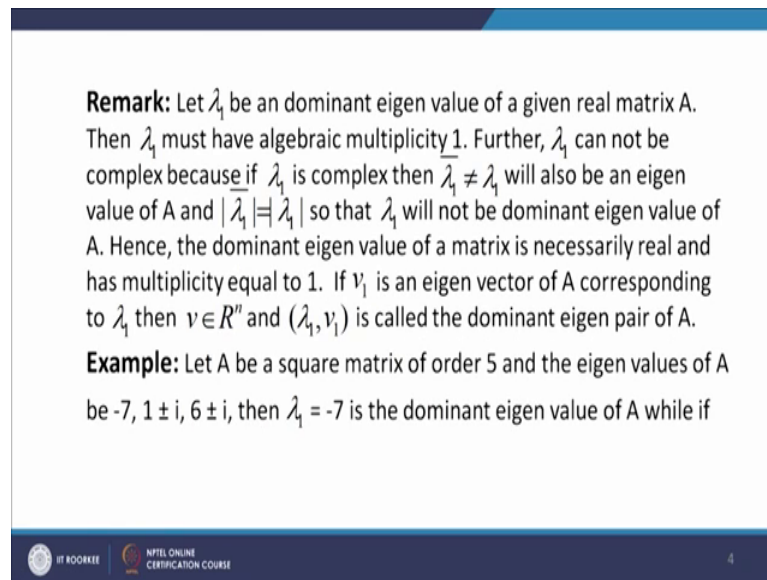
IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 3

The eigen value of the smallest magnitude can be obtained by a slight modification of this method, which we shall be discussing after this power method. So, and then the methods for finding all the eigen values we shall be discussing later. First, we shall discuss the power method for computing the dominant eigen value and the corresponding eigenvector, which we call as the dominant eigen value together with the corresponding eigenvector is called as the dominant eigen pair.

So, we are going to discuss how we can compute that. So, if λ_1 is an eigen value of A that is larger in absolute value, than any other eigen value it is called the dominant eigen value. And eigenvector v_1 corresponding to λ_1 is called a dominant eigenvector. So, if λ_1 is the dominant eigen value then $|\lambda_1|$ will be strictly greater than $|\lambda_2|$, and $|\lambda_2|$ may be equal to or greater than $|\lambda_3|$. Then λ_3 maybe such that $|\lambda_3|$ is equal to $|\lambda_4|$ are greater than $|\lambda_4|$ and so on greater than or equal to $|\lambda_n|$.

But you can notice that the λ_1 eigen value such that its modulus is strictly greater than the modulus of λ_2 .

(Refer Slide Time: 02:55)



Remark: Let λ_1 be an dominant eigen value of a given real matrix A. Then λ_1 must have algebraic multiplicity 1. Further, λ_1 can not be complex because if λ_1 is complex then $\bar{\lambda}_1 \neq \lambda_1$ will also be an eigen value of A and $|\bar{\lambda}_1| = |\lambda_1|$ so that $\bar{\lambda}_1$ will not be dominant eigen value of A. Hence, the dominant eigen value of a matrix is necessarily real and has multiplicity equal to 1. If v_1 is an eigen vector of A corresponding to λ_1 then $v \in \mathbb{R}^n$ and (λ_1, v_1) is called the dominant eigen pair of A.

Example: Let A be a square matrix of order 5 and the eigen values of A be $-7, 1 \pm i, 6 \pm i$, then $\lambda_1 = -7$ is the dominant eigen value of A while if

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE | 4

And because of this if λ_1 is an dominant eigen value of a given real matrix then λ_1 must have algebraic multiplicity equal to 1 because mod of λ_1 is strictly greater than mod of λ_2 . Now further λ_1 cannot be a complex number because if λ_1 is complex then we are dealing with a real matrix. So, eigen values so the a polynomial equation in λ will have real coefficients and therefore, if λ_1 is a root of that and λ_1 is complex then λ_1 conjugate will also be a root of that.

So, then if λ_1 is complex then λ_1 conjugate is not equal to λ_1 , but their moduli are same ok. So, modulus of λ_1 conjugate is equal to modulus of λ_1 and therefore, λ_1 will not be a dominant eigen value. So, λ_1 has algebraic multiplicity 1 and moreover; it is a real eigen value. So, hence the dominant eigen value of a matrix is necessarily real and has multiplicity equal to 1. If v_1 is an eigenvector of a corresponding to λ_1 then v belongs to \mathbb{R} to the power n and $\lambda_1 v_1$ is called the dominant eigen pair of A.

For example, let us say a be a square matrix of order 5 and the eigen values of a are suppose $-7, 1 \pm i, 6 \pm i$ then $\lambda_1 = -7$ is the dominant eigen value of A because modulus of λ_1 which is 7 is strictly greater than the modulus of all other eigen values.

(Refer Slide Time: 04:46)

the eigen values are $3, -3, 2 \pm i, -2$ then A does not have a dominant eigen value. Further, if the eigen values are $-3 \pm i, 2 \pm i, 1$ then A does not have dominant eigen value.

Definition: Let $x \in R^n$ then $\max(x)$ is defined $\max(x) = x_i$ where x_i is any element of x with the property that

$$|x_i| = \|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}.$$

If there are many entries of x having the same maximum modulus value then $\max(x) =$ first entry of x having the maximum modulus value. Thus, $\max(\cdot)$ is a well defined function on R^n .

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE | 5

Now if the eigen values are 3 minus 3 2 plus minus 1 minus 2 then a does not have a dominant eigen value because there are 2 values whose moduli are same 3 and minus 3 . Further if the eigen values are plus minus 3 plus minus 1 2 plus minus 1 then ok. Then since dominant eigen value must be real and there is a only 1 eigen value which is real and it is 1 and this eigen value is not strictly greater than the moduli of other eigen values therefore, a has does has a does not have a dominant eigen value.

Now, let us discuss how the maximum of a of a vector x belonging to R^n is defined here. So, maximum will be defined as a function here. So, let x belong to R to the power n , then maximum of x is defined maximum of x equal to x_i where x_i is any element of x with the property that modulus of x_i is equal to norm of x infinity and which is maximum of mod of x_1 mod of x_2 mod of x_n . So, the maximum of x is equal to x_i such that modulus of x_i equal to norm of x infinity. If there are many entries of of x having the same maximum modulus value then maximum at maximum of x is the first entry of x having the maximum modulus value.

Suppose there are 2 3 entries in x whose modulus is equal to norm of x infinity that is the maximum modulus modulus, then the first entry whose max whose modulus becomes norm of x infinity will be will defined the maximum of x . So, this way we can say that maximum is a well-defined function on R to the power n .

(Refer Slide Time: 06:37)

If $x \in \mathbb{R}^n$ then

$$\|x\|_{\infty} = \begin{cases} \max(x), & \text{if } \max(x) \geq 0 \\ -\max(x), & \text{if } \max(x) < 0 \end{cases}$$

Example: Let $x = [-3 \ 2 \ 3 \ -1 \ 3]^T$ then $\max(x) = -3$ and
 $\|x\|_{\infty} = 3.$

If $x = [-4 \ 4 \ -7]^T$ then $\max(x) = -7$ and
 $\|x\|_{\infty} = 7.$

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 6

Thus, if x belongs to \mathbb{R}^n we can say that norm of x infinity is equal to maximum of x , if maximum of x is greater than or equal to 0 and minus maximum of x if maximum of x is less than 0.

For example, suppose x is $[-3 \ 2 \ 3 \ -1 \ 3]^T$ then you can see that absolute value of -3 is 3 this 3 is there this 3 is there. So, norm of x infinity is 3, but maximum of x will be -3 , because the first entry in x whose modulus becomes norm of x infinity equal to norm of x infinity is -3 . So, maximum of x is -3 and here in x equal to $[-4 \ 4 \ -7]^T$, -7 is the value whose modulus becomes the maximum in all the 3 components.

(Refer Slide Time: 07:35)

Theorem: (i) If $x \in R^n$ and $c \in R$ then
$$\max(cx) = c \max(x)$$

(ii) $\max(\cdot)$ is a continuous function on R^n .

Proof: (i) Let $\max(x) = x_i$, where i is the first entry from 1 to n such that
$$|x_i| = \max\{|x_1|, |x_2|, \dots, |x_n|\}.$$

Clearly,
$$\max(cx) = cx_i = c \max(x)$$

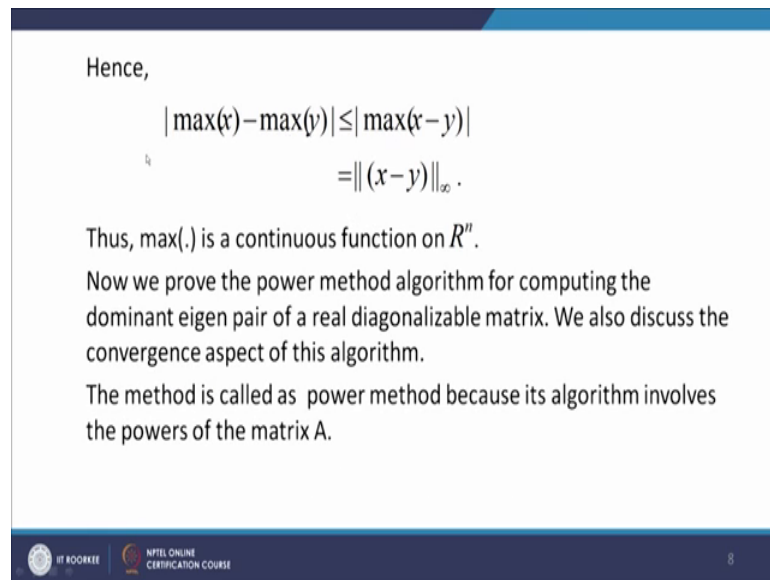
(ii) Let $x, y \in R^n$ then
$$\max(x) \leq \|x\|_\infty \text{ and } \max(y) \leq \|y\|_\infty$$

At the bottom of the slide, there are logos for IIT ROORKEE and NPTEL ONLINE CERTIFICATION COURSE, and a page number 7.

So, norm of x infinity is 7 and maximum of x is minus 7 ok. So, if x belongs to R to the power n let us discuss some properties of the maximum function if x belongs to R to the power n and c is the real number in R then maximum of cx is equal to c times maximum of x and the second observation is that maximum is a continuous function on R to the power n . Let us say maximum of the vector x is equal to x_i , where i is the first entry from 1 to n ok, such that; modulus of x_i is equal to modulus of x_1 modulus of x_2 modulus of x_n . Then when you multiply the vector x by c maximum of cx will be equal to cx_i and which is equal to c times maximum of x so this is very simple.

Now let us say x, y belong to R to the power n ok. Then maximum of x is always less than or equal to norm of x infinity, by definition of norm of x infinity and the maximum function and my maximum of y is less than or equal to norm of y infinity.

(Refer Slide Time: 08:40)



Hence,

$$|\max(x) - \max(y)| \leq |\max(x - y)|$$
$$= \|x - y\|_{\infty}.$$

Thus, $\max(\cdot)$ is a continuous function on \mathbb{R}^n .

Now we prove the power method algorithm for computing the dominant eigen pair of a real diagonalizable matrix. We also discuss the convergence aspect of this algorithm.

The method is called as power method because its algorithm involves the powers of the matrix A .

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE

Now, it is clear that the modulus of the maximum of x minus the maximum of y by definition of the maximum of x minus the maximum of y is less than or equal to the modulus of the maximum of x minus y , but this modulus of the maximum of x minus y is nothing but the infinity norm of x minus y .

So, when the infinity norm of x minus y is less than δ , the modulus of the maximum of x minus the maximum of y can be made less than ϵ . And therefore, we can say that the maximum of x is a continuous function in \mathbb{R}^n . Now let us discuss the power method algorithm for computing the dominant eigen pair of a real diagonalizable matrix. So, we will also discuss the convergence aspect of this algorithm. This method is called as the power method because when we will see the algorithm we will see that it involves the powers of the matrix A . So, we call it as power method.

(Refer Slide Time: 09:44)

Theorem: Let A be a real, diagonalizable $n \times n$ matrix with a dominant eigen value. Suppose that A has the eigen values $\lambda_1, \lambda_2, \dots, \lambda_n$ ordered so that



$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq |\lambda_4| \geq \dots \geq |\lambda_n|.$$

Let x_1, x_2, \dots, x_n be n linearly independent eigen vectors of A , corresponding to the eigen values $\lambda_1, \lambda_2, \dots, \lambda_n$ respectively. Suppose $y \in \mathbb{R}^n$ be any vector, such that if we write

$$y_0 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

then $\alpha_1 \neq 0$. For $k = 1, 2, 3, \dots$, define vectors y_k and z_k by $z_k = Ay_{k-1}$,

$$y_k = \frac{z_k}{\max(z_k)}.$$



9

So, let us say A be a real diagonalizable n by n matrix with a dominant eigen value and suppose that the values eigen values of A are ordered. So, so that modulus of λ_1 is greater than modulus of λ_2 modulus of λ_2 is greater than or equal to modulus of λ_3 and so on greater than or equal to mod of λ_n . Let us say x_1, x_2, \dots, x_n be n linearly independent eigenvectors of A we know that the matrix A is diagonalizable if and only if it has an linearly independent eigenvectors.

So, here we are assuming A is to be a diagonalizable matrix therefore, we can say that there exists n linearly independent eigenvectors. Let us take them as x_1, x_2, \dots, x_n corresponding to the eigen values $\lambda_1, \lambda_2, \dots, \lambda_n$ of A . Now suppose that y is any vector \mathbb{R}^n belonging to \mathbb{R}^n such that; if we write y naught equal to $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$ we can consider we can take α_1 to be nonzero. Now for k equal to $1, 2, 3, \dots$ we define the vectors y_k and z_k as z_k is defined as $z_k = Ay_{k-1}$ and y_k is defined as $y_k = \frac{z_k}{\max(z_k)}$. Then we shall see that as k goes to infinity z_k goes to λ_1 the dominant eigen value and y_k goes to $\alpha_1 x_1$ is the eigenvector corresponding to λ_1 .

(Refer Slide Time: 11:18)



Then, as $k \rightarrow \infty$, $\{\max(z_k)\} \rightarrow \lambda_1$ and $\{y_k\} \rightarrow cx_1$, where c is a real constant.

Proof: First, we prove that

$$y_k = \frac{A^k y_0}{\max(A^k y_0)} \text{ for } k = 1, 2, 3, \dots$$

Let us show it by induction on k . For $k = 1$, we have

$$y_1 = \frac{A y_0}{\max(A y_0)}.$$



10

So, as k goes to infinity maximum of Z_k goes to λ_1 and y_k goes to cx_1 . So, where c is a real constant. Now first we prove that y_k is equal to $A^k y_0$ over maximum of $A^k y_0$ for k equal to 1 2 3 and so on let us show it by induction on k .

(Refer Slide Time: 11:53)

Since, $z_1 = A y_0$ and



$$y_1 = \frac{z_1}{\max(z_1)}, \text{ we have } y_1 = \frac{z_1}{\max(A y_0)} = \frac{A y_0}{\max(A y_0)}.$$

Thus the result holds for $k = 1$. Let us assume that it holds for $k = m$ then

$$y_m = \frac{A^m y_0}{\max(A^m y_0)}.$$

Now,

$$y_{m+1} = \frac{z_{m+1}}{\max(z_{m+1})} \text{ where } z_{m+1} = A y_m.$$



11

So, for k equal to 1 we have from this equation y_1 equal to $A y_0$ over maximum of $A y_0$ ok. Now we have z_1 equal to $A y_0$ we can see by definition of Z_k equal to $A y_{k-1}$ z_1 equal to $A y_0$. So, z_1 equal to $A y_0$ and y_1 equal to z_1 over maximum of z_1 y_1 equal to z_1 over maximum of z_1 .

So, what do we have? So we have y_1 equal to z_1 over maximum of Ay naught and z_1 is Ay naught. So, Ay naught over maximum of Ay naught and therefore, we can say that the result holds for k equal to 1. Now let us assume that the result holds for k equal to m then by hypothesis y_m equal to $a_m y$ naught over maximum of Ay_m naught now we have to show that y_{m+1} is equal to $a_{m+1} y$ naught over maximum of $a_{m+1} y$ naught we go to the definition of y_{m+1} . So, y_{m+1} by definition is Z_{m+1} divided by maximum of Z_{m+1} .

But by definition Z_{m+1} is equal to Ay_m . So, let us replace the value here ok.

(Refer Slide Time: 13:10)

Hence,

$$y_{m+1} = \frac{Ay_m}{\max(Ay_m)} = \frac{A^{m+1}y_0}{\max(A^{m+1}y_0)}.$$

Thus, the result holds for $k = m+1$. So, by mathematical induction it holds for all $k = 1, 2, 3, \dots$. The initial guess is chosen so that

$$y_0 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

$$\Rightarrow A^k y_0 = \alpha_1 (A^k x_1) + \alpha_2 (A^k x_2) + \dots + \alpha_n (A^k x_n)$$

$$\Rightarrow A^k y_0 = \lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right]$$

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE | 12

So, y_{m+1} equal to Ay_m over maximum of Ay_m now let us put the value of y_m , y_m equal to y_m is equal to $a_m y$ naught over maximum of Ay_m naught. So, we can prove this easily we have y_{m+1} equal to Ay_m divided by maximum of Ay_m .

(Refer Slide Time: 13:40)

we have $Ax_i = \lambda_i x_i$, for $i=1, 2, \dots, n$

$A(Ax_i) = \lambda_i(Ax_i)$
 $A^2 x_i = \lambda_i^2 x_i$

$y_{m+1} = \frac{A y_m}{\max(A y_m)}$

where $y_m = \frac{A^m y_0}{\max(A^m y_0)}$

Then, we have

$$y_{m+1} = \frac{A^{m+1} y_0}{\max(A^{m+1} y_0)} = \frac{A^{m+1} y_0}{\max(A^{m+1} y_0)} \cdot \frac{\max(A^m y_0)}{\max(A^m y_0)} = \frac{1}{\max(A^m y_0)} \max(A^{m+1} y_0)$$

$y_0 = \alpha_1 x_1 + \dots + \alpha_n x_n$
 $A y_0 = A(\alpha_1 x_1 + \dots + \alpha_n x_n)$
 $= \alpha_1 A x_1 + \alpha_2 A x_2 + \dots + \alpha_n A x_n$
 $A^2 y_0 = \alpha_1 A^2 x_1 + \alpha_2 A^2 x_2 + \dots + \alpha_n A^2 x_n$
 $A^k y_0 = \alpha_1 A^k x_1 + \alpha_2 A^k x_2 + \dots + \alpha_n A^k x_n$

And y_m and y_{m+1} are equal to where y_m equal to $A^m y_0$ over maximum of $A^m y_0$. So, let us replace. So, then replacing the value of y_m we have y_{m+1} equal to $A y_m$ ok. $A y_m$ then we operate by A on this equation $A y_m$ becomes $A^{m+1} y_0$ divided by this is this is scalar value ok.

So, we get maximum of $A^m y_0$ ok. This y_{m+1} now divided by maximum of $A^m y_0$ is $A^m y_0$. So, divided by maximum of; so, this is equal to $A^{m+1} y_0$ divided by maximum of $A^m y_0$ divided by, now this becomes $A^{m+1} y_0$ and this is a scalar values. So, it will come outside the maximum. So, 1 upon this m maximum of $A^m y_0$ ok.

So, this is what we get. So, this since with this and we get $A^{m+1} y_0$ divided by maximum of $A^m y_0$. So, this proves the result. So, the result holds for all k by induction ok. So, by induction it holds for all k . Now let us choose the initial guess y_0 to be such that $\alpha_1 \neq 0$. Then if you apply the operator A to the power k $A^k y_0$ will be equal to $\alpha_1 A^k x_1 + \dots + \alpha_n A^k x_n$ where α_1 is not equal to 0 . then if you apply the operator A to the power k $A^k y_0$ will be equal to $\alpha_1 A^k x_1 + \dots + \alpha_n A^k x_n$ where α_1 is nonzero ok.

So, $A y_0$ if you apply $A y_0$ will give you A of $\alpha_1 x_1 + \dots + \alpha_n x_n$ and so on $\alpha_1 A x_1 + \alpha_2 A x_2 + \dots + \alpha_n A x_n$ ok. Now $\alpha_1, \alpha_2, \dots, \alpha_n$ are scalars. So, $\alpha_1 A x_1 + \alpha_2 A x_2 + \dots + \alpha_n A x_n$ and so on $\alpha_1 A x_1 + \alpha_2 A x_2 + \dots + \alpha_n A x_n$ we can again operate by A . So, we will get a square y_0 we

will get $\alpha_1 A^2 x_1$, $\alpha_2 A^2 x_2$ and so on $\alpha_n A^2 x_n$ and so on ok. Similarly we will give you $\alpha_1 A^k x_1$ plus $\alpha_2 A^k x_2$ and so on $\alpha_n A^k x_n$ ok.

now $\lambda_1, \lambda_2, \dots, \lambda_n$ are eigen values of a $n \times n$ matrix A are the corresponding eigenvectors of A . So, we have $Ax_i = \lambda_i x_i$ for $i = 1, 2, \dots, n$ and so on up to n ok. So, when we operate again by A we get $A(Ax_i) = A(\lambda_i x_i) = \lambda_i (Ax_i) = \lambda_i^2 x_i$ because λ_i is a scalar I can write like this now this is our $A^2 x_i = \lambda_i^2 x_i$ and $Ax_i = \lambda_i x_i$. So, we get $\lambda_i^2 x_i$. So, if λ_i is an eigen value of A then λ_i^2 is the eigen value of A^2 and x_i is the corresponding eigenvector.

So, we can do it again k times if we do pre-multiplication by the matrix A we get $A^k x_i = \lambda_i^k x_i$ for $i = 1, 2, \dots, n$. So, this we can proceed similarly we get $A^k x_i = \lambda_i^k x_i$ for $i = 1, 2, \dots, n$.

(Refer Slide Time: 19:15)

we have $Ax_i = \lambda_i x_i$, for $i=1, 2, \dots, n$

$$A(Ax_i) = \lambda_i(Ax_i)$$

$$A^2 x_i = \lambda_i^2 x_i$$

$$\vdots$$

$$A^k x_i = \lambda_i^k x_i, i=1, 2, \dots, n$$

Then

$$A^k y_0 = \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \dots + \alpha_n \lambda_n^k x_n$$

$$= \lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k x_n \right]$$

$y_0 = \alpha_1 x_1 + \dots + \alpha_n x_n, \alpha_1 + \dots + \alpha_n = 1$
 $A y_0 = A(\alpha_1 x_1 + \dots + \alpha_n x_n)$
 $= \alpha_1 A x_1 + \alpha_2 A x_2 + \dots + \alpha_n A x_n$
 $A^2 y_0 = \alpha_1 A^2 x_1 + \alpha_2 A^2 x_2 + \dots + \alpha_n A^2 x_n$
 $A^k y_0 = \alpha_1 A^k x_1 + \alpha_2 A^k x_2 + \dots + \alpha_n A^k x_n$

$$\frac{A^{m+1} y_0}{\max(A^{m+1} y_0)} = \frac{A^{m+1} y_0}{\max(A^{m+1} y_0)}$$

$$= \frac{A^{m+1} y_0}{\frac{1}{\max(A^m y_0)} \max(A^{m+1} y_0)}$$

So, we can do this here so then the then $A^k y_0$ will be equal to $\alpha_1 \lambda_1^k x_1$ plus $\alpha_2 \lambda_2^k x_2$ and so on $\alpha_n \lambda_n^k x_n$ ok.

Now, what we do we can write it as $\lambda_1^k \alpha_1 x_1$ plus $\lambda_2^k \alpha_2 x_2$ plus $\lambda_n^k \alpha_n x_n$ and so on. We can write like this now, since

λ_1 is the dominant eigen value mod of λ_j over λ_1 will be strictly less than one for all j equal to 2 3 and so on up to n . And therefore, mod of λ_j over λ_1 Goes to 0 as k goes to infinity for j equal to 2 3 and so on up to n .

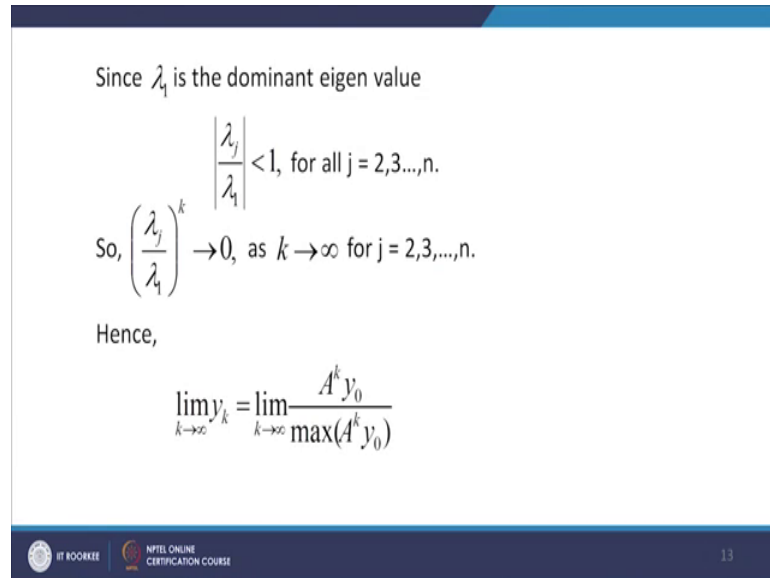
(Refer Slide Time: 20:22)

Since λ_1 is the dominant eigen value

$$\left| \frac{\lambda_j}{\lambda_1} \right| < 1, \text{ for all } j = 2, 3, \dots, n.$$

So, $\left(\frac{\lambda_j}{\lambda_1} \right)^k \rightarrow 0$, as $k \rightarrow \infty$ for $j = 2, 3, \dots, n$.

Hence,

$$\lim_{k \rightarrow \infty} y_k = \lim_{k \rightarrow \infty} \frac{A^k y_0}{\max(A^k y_0)}$$


IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 13

And therefore, we can say that limit k goes to infinity y_k is equal to we have shown earlier that y_k is equal to $A^k y_0$ naught over maximum of $A^k y_0$ ok.

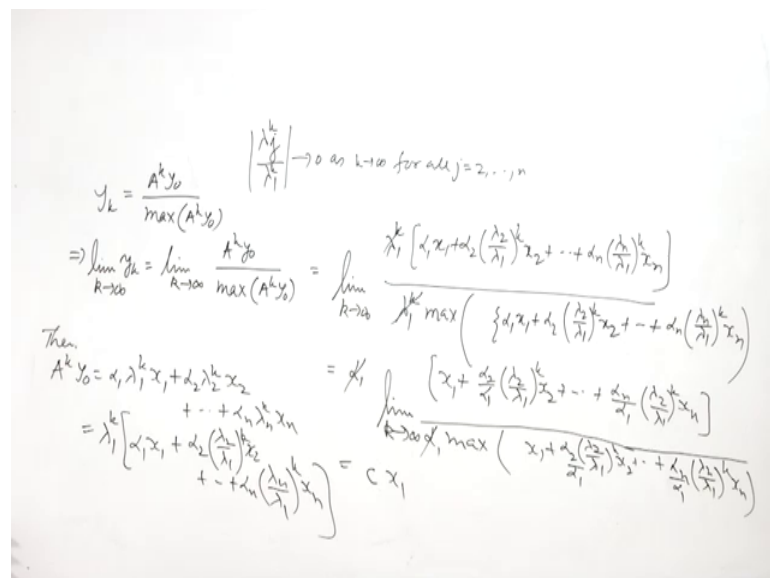
(Refer Slide Time: 20:53)

$$y_k = \frac{A^k y_0}{\max(A^k y_0)} \quad \left| \frac{\lambda_j}{\lambda_1} \right| \rightarrow 0 \text{ as } k \rightarrow \infty \text{ for all } j = 2, \dots, n$$

$$\Rightarrow \lim_{k \rightarrow \infty} y_k = \lim_{k \rightarrow \infty} \frac{A^k y_0}{\max(A^k y_0)} = \lim_{k \rightarrow \infty} \frac{\lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right]}{\lambda_1^k \max \left(\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right)}$$

Then

$$A^k y_0 = \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \dots + \alpha_n \lambda_n^k x_n$$

$$= \lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right] = c x_1$$


So, this implies limit k goes to infinity y_k is equal to limit k goes to infinity. $A^k y_0$ naught divided by maximum of $A^k y_0$ naught.

Now, so in the numerator, what we have? Aky naught we have seen just now this is Aky naught ok. Aky naught as k goes to infinity as k goes to infinity, what we have? we can say that this is equal to $\alpha_1 \times 1$ now when this this goes to 0 this goes to 0 λ_1 to the power k and then we have here maximum of I may write like this; limit k goes to infinity see this is λ_1 to the power k $\alpha_1 \times 1$ $\alpha_2 \times \lambda_2$ by λ_1 raised to the power k and then we have x_2 also here. So, x_2 and here we have $\alpha_n \lambda_n$ by λ_1 raised to the power k x_n this divided by maximum of maximum of Aky naught.

So, λ_1 to the power k λ_1 to the power k then we have $\alpha_1 \times 1$ plus $\alpha_2 \lambda_2$ by λ_1 to the power k x_2 and so on. $\alpha_n \lambda_n$ by λ_1 raised to the power k x_n . And then we have limit k goes to infinity ok. So, this can be written as we can take λ_1 to the power k outside here because it is a scalar. I can write it here and then this λ_1 to the power k λ_1 to the power k can be cancelled α_1 is nonzero.

So, I can take α_1 outside and then we have α_1 times limit k goes to infinity. I can take outside so we get x_1 plus α_2 over $\alpha_1 \lambda_2$ by λ_1 raised to the power k x_2 and so on. α_n by $\alpha_1 \lambda_2$ by λ_1 raised to the power k and we have x_n here, and divided by divided by maximum of $\alpha_1 \times 1$ plus $\alpha_2 \lambda_2$ by λ_1 to the power k x_2 and so on α_n by sorry α and λ_n by λ_1 to the power k x_n .

So, when k goes to infinity λ_j by λ_1 , this goes to 0 as k goes to infinity. This to be power k goes to 0 as k goes to infinity for all j from 2 to n ok. So, these terms go to 0 and we get x_1 here ok. And then this quantity this quantity gives you a maximum 1 upon maximum of this when k goes to infinity this defines a constant. So, we get α_1 here we will have α_1 be connected α_1 outside here also then also α_1 is a constant ok. So, this α_1 can be cancelled with this α_1 and we will have some c times x_1 ok. This x_1 this quantity is tend to 0 and here we have these quantity tend to 0. So, maximum of this quantity we can consider.

So, 1 upon maximum of that can define a constant c . So, we have $c x_1$. So, by k as k goes to infinity gives us c times x_1 and where α_1 is not equal to 0.

(Refer Slide Time: 26:44)

$$= x_1 \lim_{k \rightarrow \infty} \left[\frac{\lambda_1^k}{\max \left(\lambda_1^k x_1 + \frac{\alpha_2}{\alpha_1} \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \frac{\alpha_n}{\alpha_1} \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right)} \right]$$

$$= c x_1 \text{ where } \alpha_1 \neq 0.$$

To find c , we note that

$$y_k = \frac{A^k y_0}{\max(A^k y_0)} \Rightarrow \max y_k = 1, \text{ for all } k=1,2,\dots$$

Now we need to find the value of the constant c . So, for this we need we note that y_k is equal to $A^k y_0$ naught over maximum of $A^k y_0$ naught. So, when we will find maximum value of k , maximum value of k will be maximum of $A^k y_0$ naught divided by maximum of $A^k y_0$ naught y_k is equal to $A^k y_0$ naught divided by maximum of $A^k y_0$ naught.

(Refer Slide Time: 27:12)

$$y_k = \frac{A^k y_0}{\max(A^k y_0)}$$

$$\Rightarrow \lim_{k \rightarrow \infty} y_k = \lim_{k \rightarrow \infty} \frac{A^k y_0}{\max(A^k y_0)} = \lim_{k \rightarrow \infty} \frac{\lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right]}{\lambda_1^k \max \left(\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right)}$$

$$= x_1 \lim_{k \rightarrow \infty} \frac{\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n}{\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n}$$

$$= \frac{1}{\max(A^k y_0)} \max(A^k y_0) = 1$$

because $\max(cy) = c \max(y)$

So, when we find maximum here maximum of y_k . Then this is this one maximum of $A^k y_0$ naught is a constant ok .

So, we shall have this 1 upon maximum of $A_k y$ naught and then maximum of $A_k y$ naught this is equal to 1; this is because maximum of a constant c times y equal to c times maximum of y we have earlier scene where c is any constant. So, because of that, maximum of y_k is equal to 1 for all k equal to 1 2 3 and so on.

(Refer Slide Time: 28:23)

Since $\max(\cdot)$ is a continuous function

$$y_k \rightarrow c x_1, \text{ as } k \rightarrow \infty \Rightarrow \max(y_k) \rightarrow \max(c x_1), \text{ as } k \rightarrow \infty$$

but $\max(y_k) = 1$ for all $k = 1, 2, \dots, n$ hence

$$1 = \max(c x_1) = c \max(x_1) \Rightarrow c = \frac{1}{\max(x_1)}.$$

Further by definition $z_k = A y_{k-1}$ and so $\lim_{k \rightarrow \infty} z_k = A(c x_1) = c A(x_1) = c \lambda_1 x_1$
because $y_k \rightarrow c x_1$, as $k \rightarrow \infty$. Hence

$$\lim_{k \rightarrow \infty} \max(z_k) = \max(c \lambda_1 x_1) = c \lambda_1 \max(x_1) = \lambda_1$$

as

$$c = \frac{1}{\max(x_1)}. \text{ This proves the theorem.}$$

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE | 15

Now y_k max we have seen that maximum is a continuous function. So, then y_k goes to $c x_1$ as k goes to infinity maximum of y_k will go to maximum of $c x_1$ and maximum of $c x_1$ will be equal to c times maximum of x_1 ok.

So, we will have c equal to 1 upon maximum of x_1 let us notice that limit k goes to infinity y_k equal to c of x_1 ok.

(Refer Slide Time: 28:56)

$$y_k = \frac{A^k y_0}{\max(A^k y_0)}$$

$$\Rightarrow \lim_{k \rightarrow \infty} y_k = \lim_{k \rightarrow \infty} \frac{A^k y_0}{\max(A^k y_0)}$$

$$y_k = \frac{A^k y_0}{\max(A^k y_0)}$$

$$\max(y_k) = \max\left(\frac{A^k y_0}{\max(A^k y_0)}\right)$$

$$= \frac{1}{\max(A^k y_0)} \max(A^k y_0) \quad \text{because } \max(cy) = c \max(y)$$

$$= 1$$

$$\lim_{k \rightarrow \infty} y_k = c x_1$$

Since $\max(\cdot)$ is a continuous function on \mathbb{R}^n

$$\lim_{k \rightarrow \infty} \max(y_k) = \max(c x_1)$$

$$1 = \max(c x_1)$$

$$= c \max(x_1)$$

$$\Rightarrow c = \frac{1}{\max(x_1)}$$

Since, maximum of maximum is a continuous function in \mathbb{R}^n it is a continuous function on \mathbb{R}^n . Maximum of y_k as k goes to infinity will be equal to maximum of $c x_1$. So now, maximum of y_k we have seen maximum of y_k equal to 1 for all k ok. So, 1 will be equal to maximum of $c x_1$ which is equal to c times maximum of x_1 and this gives us c equal to 1 upon maximum of x_1 ok.

So, this is how we obtain the constant c c is equal to 1 upon maximum of x_1 . Now by definition Z_k equal to $A y_k$ minus 1 . So, as k goes to infinity Z_k goes to limit of $A y_k$ minus 1 y_k goes to $c x_1$. So, y_k minus 1 also goes to $c x_1$. So, A of $c x_1$ we have A of $c x_1$ is equal to c times $A x_1$ we can see here Z_k equal to $A y_k$ minus 1 . So, limit k goes to infinity Z_k equal to limit by limit k goes to infinity A of y_k minus 1 . So, y_k minus 1 will also go to $c x_1$. So, $A c x_1$ which is equal to c times $A x_1$ and $A x_1$ is equal to λx_1 because λ is an eigen value of A and x_1 is the corresponding eigenvector.


So, this is $c \lambda x_1$. Now hence limit k goes to infinity maximum of Z_k ok. Maximum is again let us recall that maximum is a continuous function. So, when Z_k goes to $c \lambda x_1$ maximum of Z_k will go to maximum of $c \lambda x_1$ $c \lambda$ is a constant. So, $c \lambda$ times maximum of x_1 , but maximum of x_1 into c is equal to 1 because c is equal to 1 upon maximum of x_1 . So, this is equal to λ and therefore, as k goes to infinity maximum of Z_k goes to λ . So, this proves the theorem.

(Refer Slide Time: 31:42)

Remark 1: In the statement of the theorem, we assumed that the initial guess y_0 is such that if

$$y_0 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

then $\alpha_1 \neq 0$ i.e. y_0 has a non zero projection on the eigen space corresponding to the eigen vector x_1 and so y_k converges in the direction of the dominant vector x_1 . Here one can make an objection that we do not know the dominant eigen vector a priori then it can be said that without any loss of generality this assumption can be made because even if $\alpha_1 = 0$, the round off errors in the subsequent iterations enforce that y_k has a non zero projection on the eigen space spanned by the dominant eigen vector x_1 .



Now let us look at some on the remark one in a this statement of the theorem we assumed that the initial guess y_0 is such that; y_0 is equal to $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$ and where α_1 is not equal to 0 that is y_0 has a nonzero projection on the eigen space corresponding to the eigenvector x_1 . And so y_k converges in the direction of the dominant vector x_1 .

Now, here one can raise an objection that we do not know the dominator eigenvector a priori then it can be said that without any loss of generality this assumption can be made because even if α_1 is equal to 0 the round of errors in the subsequent iterations will enforce that y_k has a nonzero projection on the eigen space spanned by the a dominant eigenvector x_1 .

(Refer Slide Time: 32:40)

Remark 2: It is customary to take the initial guess

$$y_0 = [1 \ 1 \ \dots \ 1]^T$$

in the power method algorithm.

Example: Let

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -5 & 2 \\ 5 & 2 & -10 \end{bmatrix}$$

Then the eigen values of A are 0.0887, -5.5638 and -11.3249. Hence, the dominant eigen value is -11.3249. Now, let us find the dominant eigen value by using power method.



Now it is called when we solve a any matrix for the dominant eigen value and the corresponding eigenvector it is, conventional to take the initial guess as a 1111 ok.

So, we take y naught equal to 1111 in the power method algorithm. Now let us take an example suppose we take a 3 by 3 matrix. So, a equal to minus 1 1 2 3 minus 5 2 5 2 minus 10 ok. Then the eigen values of a are 0.0887 minus 5.5638 and minus 11.3249. Here is the dominant eigen value is minus 11.3249 ok. You can see among all these 3 eigen values the greatest one the whose absolute value is the greatest is minus 11.3249.

Now, let us find the dominant eigen value. So, this is the dominant eigen value let us find this dominant eigen value by the power method let y naught equal to 1 1 1 transpose.

(Refer Slide Time: 32:50)

Let $y_0 = [1 \ 1 \ 1]^T$ then $z_1 = Ay_0 = [2 \ 0 \ -3]^T$
 and $\max(z_1) = -3$. Hence

$$y_1 = \frac{z_1}{\max(z_1)} = \frac{[2 \ 0 \ -3]^T}{-3}$$

$$= [-0.6667 \ 0 \ 1.0000]^T.$$

Now, $z_2 = Ay_1 = [2.6667 \ 0 \ -13.3333]^T$ and
 $\max(z_2) = -13.3333$ hence

$$y_2 = \frac{z_2}{\max(z_2)} = [-0.2000 \ 0 \ 1.0000]^T$$

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE | 18

As we said in it is conventional to take y naught as 1 1 1 be here it is 3 by 3 matrix should we take y naught equal to 1 1 1 a maybe 3 components. Then z_1 we find Z_k equal to Ay_k minus 1 ok. So, for k equal to 1 k equal to 1 means; we are doing the first iteration ok. So, Z_k equal to Ay_k minus 1 becomes z_1 equal to Ay naught. And when you multiply the matrix A the given matrix A by the y naught vector 1 1 1 we will get a column vector 2 0 minus 3 ok.

And from 2 0 minus 3 we can say that maximum of z_1 is minus 3 by definition of maximum, we know that maximum of z_1 is the value the value the value the component which has the maximum absolute value ok. So, maximum absolute value in 2 0 minus 3 is 3. So, we have maximum of z_1 equal to minus 3 hence, y_1 is equal to by definition y_1 equal to y_k equal to Z_k over maximum of Z_k by that definition. So, y_1 equal to z_1 over maximum of z_1 z_1 is 2 0 minus 3 transpose divided by maximum of z_1 minus 3.

So, we get column vector minus 0.6667 0 1.0000. Now we do the second iteration. So, z_2 equal to Ay_1 ok. A matrix is multiplied by the y_1 vector this 1 minus 0.6667 0 1.0000 transpose. So, by this vector when we multiply the matrix A we get 2.6667 0 minus 13.3333. Now minus 13.3333 is the a component which has which has the maximum absolute value.

So, we maximum of z_2 is equal to minus 13.3333 hence y_2 equal to z_2 over maximum of z_2 . So, z_2 vector this vector z_2 vector is divided by a maximum of maximum of z_2 value and what we get is? Minus 0.20000 1.0000 transpose.

(Refer Slide Time: 36:07)

Next $z_3 = Ay_2 = [2.2000 \quad 1.4000 \quad -11.0000]^T$ and
 $\max(z_3) = -11.000$ so
 $y_3 = \frac{z_3}{\max(z_3)} = [-0.2000 \quad -0.1273 \quad 1.0000]^T$.
 Proceeding like this, we obtain
 $z_{11} = Ay_{10} = [1.9368 \quad 2.6585 \quad -11.3243]^T$
 and hence $\max(z_{11}) = -11.3243$ and
 $y_{11} = \frac{z_{11}}{\max(z_{11})} = [-0.1710 \quad -0.2348 \quad 1.0000]^T$

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 19

Now we for k equal to 3 we get z_3 equal to Ay_2 . So, z_3 comes out to be 2.2000 1.4000 minus 11.0000 transpose again maximum z_3 is minus 11.0000.

So, here we find y_3 by z_3 upon maximum of z_3 and we get minus 0.2000 minus 0.1273 and 1.0000 ok. Proceeding in this manner ok, proceeding in this manner we can find z_{11} ok. So, when we take the 11 in the 11th iteration we will get z_{11} z_{11} equal to Ay_{10} which will be 1.9368 2.655 minus 11.3243 to the transpose.

So, here maximum of z_{11} is minus 11.3243 ok. And y_{11} which is z_{11} z_{11} divided by maximum of z_{11} it will be minus 0.1710 minus 0.2348 and then 1.0000 transpose. Now let us see.

(Refer Slide Time: 37:25)

We observe that

$$\max(z_{11}) = -11.3243$$

is a good approximation for the dominant eigen value $\lambda_1 = -11.3249$

Also,

$$Ay_{11} - \lambda_1 y_{11} = [-0.0006 \quad 0.0021 \quad 0.0003]^T$$

which is small.

Hence, we can say that

$$\max(z_k) \rightarrow \lambda_1 \quad \text{and} \quad y_k \rightarrow c\lambda_1 \quad \text{as } k \text{ takes large values.}$$

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 20

So, here maximum of z_1 is minus 11.3243 and which you can see is a good approximation for the dominant eigen value λ_1 equal to minus 11.3249.

So, we can say that maximum of Z_k approaches to the dominant eigen value λ_1 and furthermore we also notice that $Ay_{11} - \lambda_1 y_{11}$ is very small. λ_1 we have found to be minus 11.3243. z_1 maximum of z_1 we are taking as minus 11.3243 as the dominant eigen value. So, this eigen value we are multiplying by y_{11} we have found earlier this is y_{11} . So, y_{11} is this and λ_1 is this ok.

So, we multiply here and subtract from Ay_{11} and when you do that what you get is this vector minus 0.0006 0.0021 0.0003 transpose, and which is very small you can see hence we can say that maximum of Z_k converges to λ_1 and y_k the y_k vector goes to c times λ_1 as k takes large values ok.

(Refer Slide Time: 38:50)

Example: Let

$$A = \begin{bmatrix} 9 & 5 & 3 & 7 \\ 8 & 3 & 8 & 4 \\ 7 & -6 & 6 & 2 \\ 4 & 6 & 7 & 9 \end{bmatrix}$$

Then the eigen values of A are 20.4530, $1.3114 \pm 6.0068i$ and 3.9242.

Hence the dominant eigen value of A is $\lambda_1 = 20.4530$.

Now, let us find λ_1 by power method algorithm. We choose

$$y_0 = [1 \ 1 \ 1 \ 1]^T. \text{ For } k=1,$$

$$z_1 = Ay_0 = [24 \ 23 \ 9 \ 26]^T.$$

So, let us now take the next example A equal to 9 5 3 7 8 3 8 4 7 minus 6 6 2 4 6 7 9.

So, here we are taking 4 by 4 matrix ok. The eigen value of the matrix A are 20.4530 1.3114 plus minus 6.008 into iota and then 3.9242. So, the absolute my maximum value here is 20.4345 20.4530 and therefore, lambda 1 the dominant eigen value of a is 20.4530. Now it is a 4 by 4 matrix we will choose initial guess y naught to be 1 1 1 1 ok.

(Refer Slide Time: 39:44)

Hence, $\max(z_1) = 26$.

Now,

$$y_1 = \frac{z_1}{\max(z_1)} = [0.9231 \ 0.8846 \ 0.3462 \ 1.0000]^T$$

for $k=2$,

$$z_2 = Ay_1 = [20.7692 \ 16.8077 \ 5.2308 \ 20.4231]^T.$$

$$\Rightarrow \max(z_2) = 20.7692$$

$$y_2 = \frac{z_2}{\max(z_2)} = [1.0000 \ 0.8093 \ 0.2519 \ 0.9833]^T.$$

For k equal to 1, we will find z_1 z_1 equal to Ay_0 now here you can see maximum of z_1 we can find ah. So now, y_1 equal to z_1 upon maximum of z_1 and we will get 0.9231 0.8846 0.3462 and 1.0000. Now for k equal to 2 z_2 equal to Ay_1 . So, we can

find Ay_1 into n it will come out to be 20.7692 16.8077 5.2308 20.4231 transpose and maximum of z_2 here you can see is 20.7692.

So, we can find y_2 here y_2 equal to z_2 upon maximum of z_2 it is 1.0000 0.8093 0.2519 0.9833 and the maximum absolute maximum value here is that you get is 0.999833.

(Refer Slide Time: 40:46)

Proceeding similarly,

$$z_{10} = [20.4530 \quad 16.4272 \quad 5.7523 \quad 19.2649]^T.$$

Note that $\max(z_{10})$ is the same as the dominant eigen value $\lambda_1 = 20.4530$ and $y_{10} = [1.0000 \quad 0.8032 \quad 0.2812 \quad 0.9419]^T$ is close to the direction of the dominant eigen vector x_1 of A as

$$Ay_{10} - \lambda_1 y_{10} = 10^{-4} [-0.0863 \quad -0.5301 \quad -0.5480 \quad -0.0964]^T$$

is small.

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 23

So, we we will get proceeding similarly z_{10} comes out to be 20.4530 16.4272 5.7523 19.2649 transpose. And we note that maximum of z_{10} maximum of z_{10} here is 20.4530 which is same as the dominant eigen value λ_1 equal to 20.4530, y_{10} you can see y_{10} it is y_{10} oh sorry y_{10} y_{10} comes out to be this 1.0000.

So, y_{10} is z_{10} divided by maximum of z_{10} ok. We have found z_{10} maximum of z_{10} is 20.453 0. So, when we divide z_{10} by 20.4530 that is that is 1.0000 0.8032 0.2812 0.9419. Now here you can see the actual dominant eigen value was a dominant eigen value was this 20.4530 and we could get this value 20 4 ah.4 5 3 0 after 10 iterations ok.

So, maximum of z_{10} is exactly same as the dominant eigen value 20.453 and the eigen vector y_1 y_{10} y_{10} is this one ok. So, when you multiply this y_{10} by the λ_1 20.4530 then Ay_{10} minus $\lambda_1 y_{10}$ you can see comes out to be 10^{-4} minus 0.0863 minus 0.5301 minus 0.5480 minus 0.0964 a transpose which is very small. So, we can say that as k goes to infinity maximum of Z_k converges to

$\lambda = 1$ and y_k convergence to $c \times x_1$ that is, y_k converges to x_1 in the direction of x_1 . So, the so the, this is because $\|A - \lambda I\|$ is small, with that I would like to conclude my lecture.

Thank you very much for your attention.