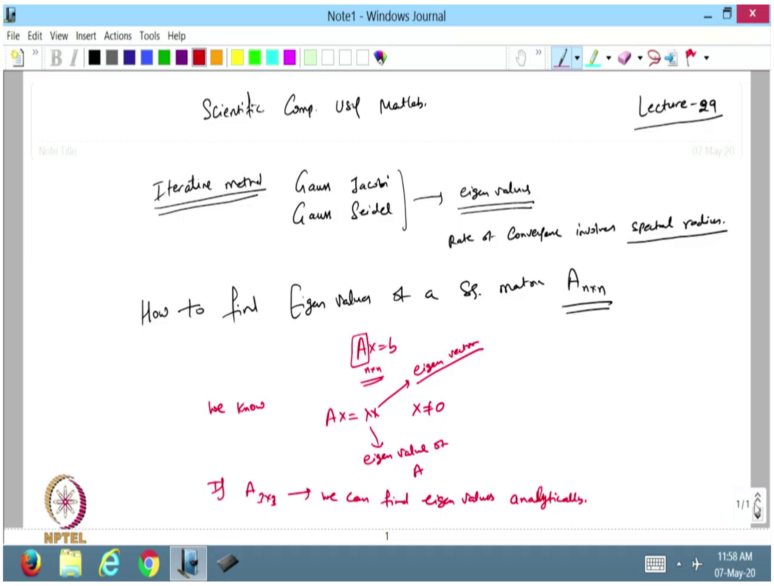
Scientific Computing Using Matlab
Professor. Vivek Aggarwal & Professor. Mani Mehra
Department of Mathematics
Indian Institute of Technology, Delhi
Lecture No. 29
Power Method for Solving Eigenvalues of a Matrix

Hello viewers. Welcome back to the course on scientific computing in Matlab. So, today we will discuss another method that is called how we can find the eigenvalues, because in the previous lecture we have seen that the eigenvalues play a very important role in the convergence of the iterative method. So, now the next question is how we can find the eigenvalue using the numerical computation?

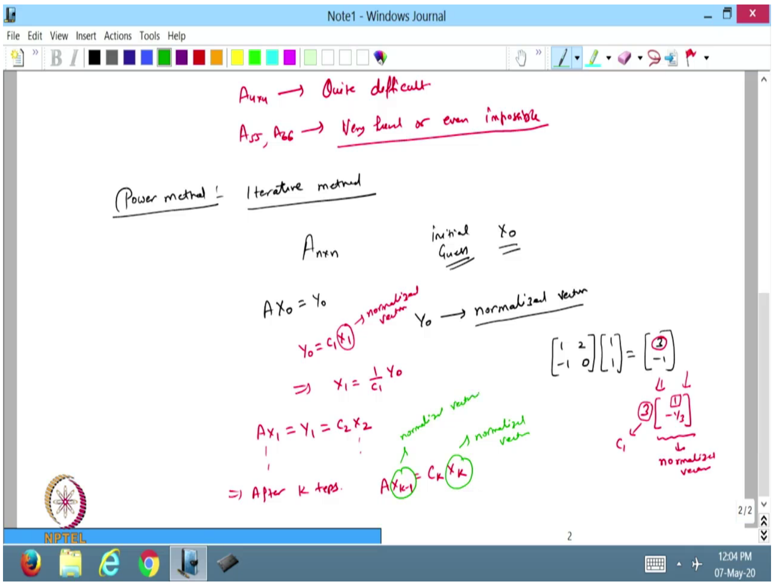
(Refer Slide Time: 0:49)



So, now from the previous one we found that the iterative method like Gauss Jacobi or Gauss Seidel or other method, they are dependent on the eigenvalues. The rate of convergence involves spectral radius. So, the question comes: how to find the eigenvalues of the matrix? So, let us the next topic is how to find the eigenvalues of a square matrix that is A, that is n cross n because we are involved with solving the system Ax is equal to b and then we have to find what is the eigenvalue of this n corrosion system.

We know that the eigenvalues can be found analytically as $Ax = \lambda x$ where $x \neq 0$. Then in that case we say that this λ is an eigenvalue of matrix A and this x is called corresponding Eigenvector. Now I want to find first what is the eigenvalue and then the Eigenvector and then we know how we can find this one. Now, if A is 3 cross 3, then we can find eigenvalues analytically. Analytically means that with the pen or paper.

(Refer Slide Time: 3:12)



I can find the Cartesian equation and then we can solve this one. Maybe for 4 cross 4 it is quite difficult and then A 5 cross 5 and so on. So, I can say that it is very hard or even impossible. So, now to find out this one we take the help of numerical computation. So, to deal with this one, how to find the eigenvalue of the matrix there is a method called power method. So, the power method is an iterative method. So, this is an iterative method. So, how can we deal with this one?

Suppose I have a matrix A that is n cross n matrix. Now, what I do because this is iterative methods, so I start with the initial x_0. So initial x_0 means this is my initial guess. So, that is my initial guess x_0. Now what do I do? I will take Ax_0. So, I will get another vector and that other vector I call it maybe y_0.

So, in this case what I do is that, this y_0 is a vector. So, now I reduce this vector into the normalized form, normalized vector. So, whatever the vector y_0 I am getting, I will call it a normalized vector. So, what is the meaning of a normalized vector? Suppose I have a matrix like 1, 2, -1, 0, suppose this is my matrix and I start with the process 1 1, then I will get the value, this will multiply and this will add.

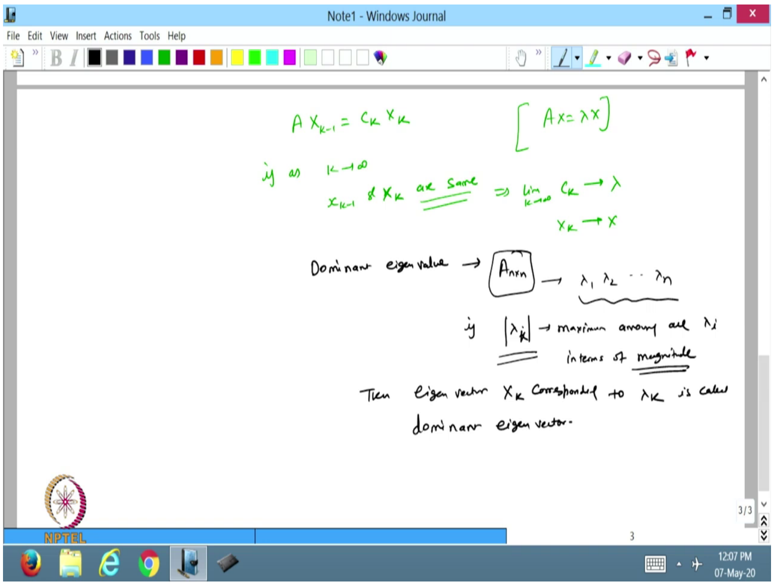
So, this is 3 and this will be 1 and this will be -1. So, now in this case this is the element with the highest magnitude. So, in this case what I will do, I will take 3 common from this. So, this will be 1 and this is -1 by 3. Now the highest value, highest component in this vector is 1. So, this vector is called a normalized vector. Does not matter what is the sign because we want the highest in terms of numerical value or in magnitude value. So, that this vector becomes the normalized vector.

So, what do you do? We make it x_0. So, this x_0, this y_0, what do I do? In the naught I will take this element common and then I will make this x1 and that I will call it c1. So, what am I doing here? Now from this vector I have taken this element, the highest element common. So, I call it c1 now and then the remaining vector becomes x1. So, x1 is my normalized vector.

So, from here I can say that my $x_1 = \frac{1}{c_1} y_0$. So, that is my normalized vector in this case. Now from here, what do I do? Now I will take the next step. I will put Ax_1. So in this case what will happen? I will put the Ax_1, from here I will get y_1. Now with the help of y1 I will reduce this one into c2 x2. So, I will keep doing that. Then from here I can say that after k steps because after 1 step, I will get $Ax_0 = c_1 x_1$.

So, from here I will get Ax_1 in the after 2 step I will get Ax_1. So, after k step I will get $Ax_{k-1} = c_k x_k$. So, from here I can write this one, where this is that you have to keep in mind that this is a normalized vector. This is also a normalized vector.

(Refer Slide Time: 9:11)



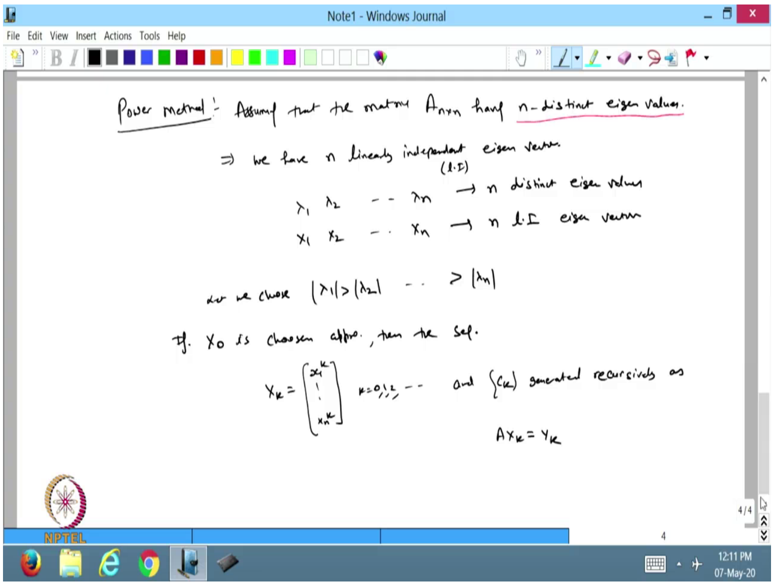
So, now from here you can see that I am getting this type of algorithm that is ck xk. Now, if you just compare with this $Ax = \lambda x$, now you can see from here that if as $k \to \infty$ that after many iterations if x_{k-1} and x_k are the same, same means they will take that difference between these L2 norm and that norm is less than the required tolerance is almost same which implies that the limit k tends to infinity ck will tends to λ.

So, this will tend to λ and then my xk tends to the corresponding Eigenvector. So, that is the process of the power series method. Now before that one so there is one more term we want to define that dominant Eigen. So, what is the dominant eigenvalue? Dominant eigenvalue means suppose I have a matrix, n cross n matrix and now from the linear algebra we know that if we have a n cross n matrix, then we have a n number of eigenvalues.

So, suppose I write $\lambda_1, \lambda_2, \dots, \lambda_n$, this is the n eigenvalues corresponding to this matrix. These eigenvalues may be complex and also do not matter. Now, if we choose λ, any λ ith value taking the magnitude, that is maximum maximum among all $\lambda_i' s$, I can call it k among all i's then this is called the dominant eigenvalue. So, we call it in terms of magnitude maximum all lambda i's in terms of magnitude.

Then the eigenvector that is xk corresponding to λ_k is called the dominant Eigenvector. So, that is called the dominant Eigenvector. So, now we are ready to apply that power method. So, this is the power method we are going to define now.

(Refer Slide Time: 12:32)

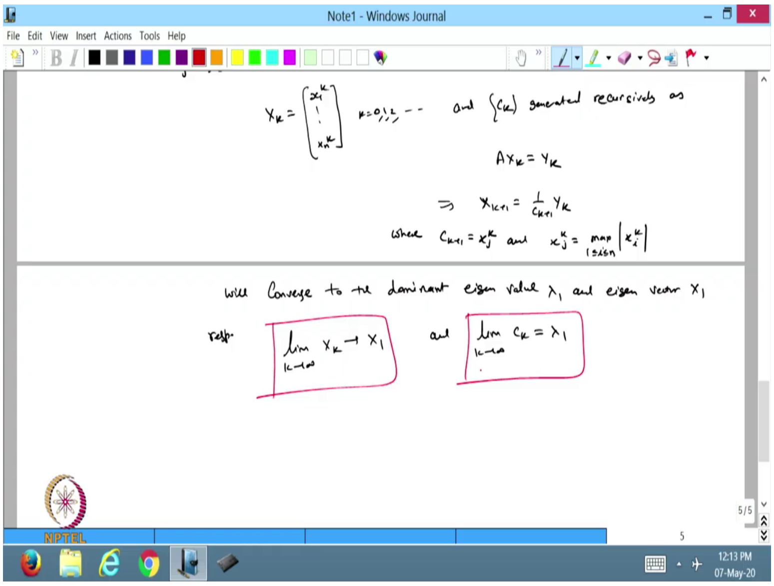


Power method, so assuming that the matrix A that is n cross n matrix having n distinct eigenvalues. So, that is the condition. We are dealing with n distinct eigenvalues because we know that the eigenvalue may be repeating also. So, that we are not keeping in mind we are having the n distinct eigenvalues. Now we know that if we have n distinct eigenvalues then which implies that we have n linearly independent Eigenvectors.

So, we have now $\lambda_1, \lambda_2, \ldots, \lambda_n$. These are n distinct eigenvalues and I take it v1, v2, vn or we can take this as x1, x2, xn. So, these are n linearly independent. So, this is LI. I can write it as LI Eigenvectors. So, these are there. Now, I know that this is the linearly independent Eigenvector. So, let us choose without loss of generality that λ_1 is the highest one.

So, I can call it $|\lambda_1| > |\lambda_2|, \ldots > |\lambda_n|$ because they are n distinct eigenvalues. So, we could call it. Now if, x_0 is chosen appropriately, then the sequence xk, so this is x_k I am writing in terms of a vector because this will be a vector definitely. So, this is x_1^k, x_n^k where k is 0 1 2 and so on and ck generated recursively as $Ax_k = y_k$

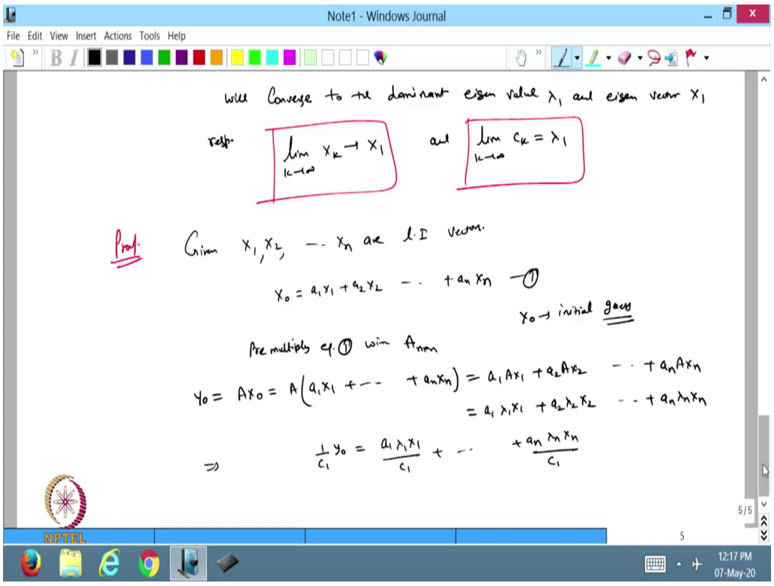
(Refer Slide Time: 16:04)



Which implies my $x_{k+1} = \frac{1}{c_{k+1}} y_k$ where where the $c_{k+1} = x_j^k$. Suppose I take that $x_j^k = \max_{1 \le i \le n} |x_i^k|$. I already told you that it should be the maximum value and that maximum I will take k common that is my ck as we have discussed here. So, this is basically my ck and that is the maximum value.

Then the sequence will converge to the dominant eigenvalue that is λ_1 and eigen vector that is x1 respectively. So, that is the $\lim_{k \to \infty} x_k \to x_1$ and $\lim_{k \to \infty} c_k \to \lambda_1$. So, this is my statement. I have already explained for a given problem how we can implement this power method. So, that is the statement.

(Refer Slide Time: 18:20)



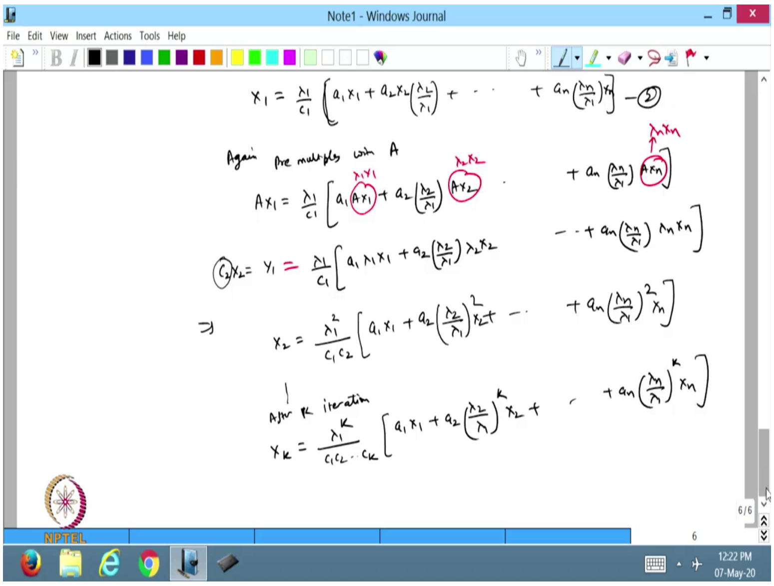
Now we can do the proof of this one. So, proof is quite easy. Now given that x1, x2 up to xn are linearly independent vectors. Now from here I can say that if I choose any vector x_0 that can be written as a linear combination so $x_0 = a_1 x1 + a_2 x_2 + \ldots + a_n x_n$. So, I can write these as a linear combination because these are the n vectors linearly in one vector.

So, any vector which is dimension n can be written as a linear combination of this one. So, that is my equation number 1. So, where x_0 is initial guess. So, that is the initial guess we are going to start with. Now I apply the matrix. So, pre-multiply equation 1 with the a matrix and the corrosion matrix whatever the matrix I am taken. So,

$$Ax_0 = A(a_1 x_1 + \ldots + a_n x_n) = a_1 Ax_1 + a_2 Ax_2 + \ldots + a_n Ax_n$$

Now the x1 is the eigenvector corresponding to the eigenvalue λ_1. So, from here I can write that this should be equal to $a_1 \lambda_1 x_1 + a_2 \lambda_2 x_2 + \ldots + a_n x_n$. So, this one we can write from here. So, from here I can write this one as now this is my ax2. Now from here I know that this $Ax_0 = y_0$.

Now I know that in the y0 I will take the factor ck common. So, what can I write from here? This one I can write as $\frac{1}{c_1} y_0 = \frac{a_1 \lambda_1 x_1}{c_1} + \ldots + \frac{a_n \lambda_n x_n}{c_n}$ (Refer Slide Time: 21:47)



From here this I know that becomes x1 because we already know that this becomes x 1. So, from here I can write that x1 can be written as now I can write this as so from what I am doing now, I will just take the common λ_1 over c1 from all. Now I will get from inside I will get $a_1 x_1 + a_2 x_2 \dfrac{\lambda_2}{\lambda_1}$. I will get this value and so on in the end I will get $a_n \dfrac{\lambda_n}{\lambda_1}$.

So, this one I so c1 I just take common and I will take λ_1 also common from all this. Now I will get from here, so this is again I can apply for this one. Now I will again so I can write it as a 2. Now again apply multiply so again multiply pre multiply with respect to A. So, I will get

$$Ax_1 = \frac{\lambda_1}{c_1}\left[a_1 Ax_1 + a_2 \frac{\lambda_2}{\lambda_1} Ax_2 + \ldots + a_n \frac{\lambda_n}{\lambda_1} Ax_n\right]$$

Now from here I know that this Ax_1 is again the λ_1. This is λ_2. This is $\lambda_1 x_1$. This is $\lambda_2 x_2$. This is $\lambda_n x_n$. So from here I can write this as

$$Ax_1 = \frac{\lambda_1}{c_1}\left[a_1 \lambda_1 x_1 + a_2 \frac{\lambda_2}{\lambda_1} \lambda_2 x_2 + \ldots + a_n \frac{\lambda_n}{\lambda_1} \lambda_n x_n\right]$$

Now again I can take this lambda 1 common and from here I can get a x1 will be what? That would be y1 and y1 I take the common factor. So, this will be equal to c2 x2. So, now from here I can write that my x2 will be λ_1 and again λ_1 I am taking the common square and it will be c1c2 because this factor is I have taken this as a common.

This is the highest value we are taking to make this vector as a normalized vector and from here I can write this as a again

$$x_2 = \frac{\lambda_1^2}{c_1 c_2}\left[a_1 x_1 + a_2(\frac{\lambda_2}{\lambda_1})^2 x_2 + \ldots + a_n(\frac{\lambda_n}{\lambda_1})^2 x_n\right]$$

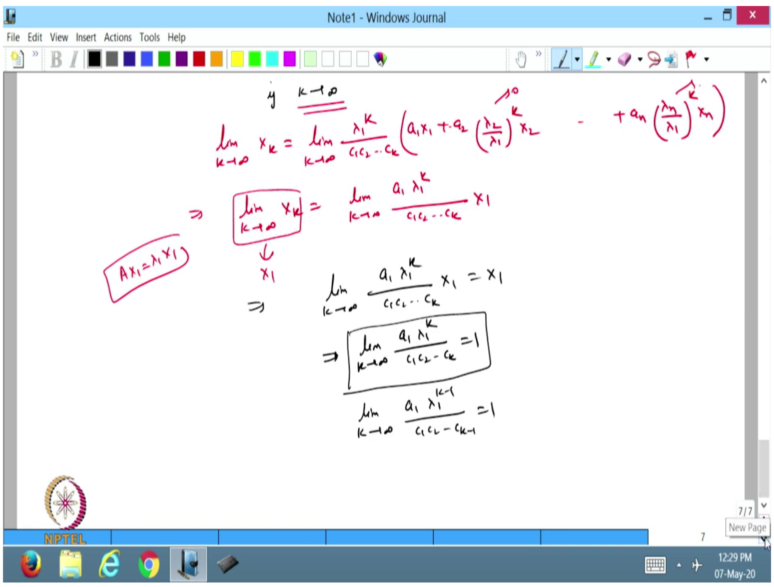
and so on. So, if we keep doing after k iterations, so the first iteration I will get x1, in the second iteration I will get x2. So after k iteration I will get x k.

So, it will be

$$x_2 = \frac{\lambda_1^k}{c_1 c_2 \ldots c_k}\left[a_1 x_1 + a_2(\frac{\lambda_2}{\lambda_1})^k x_2 + \ldots + a_n(\frac{\lambda_n}{\lambda_1})^k x_n\right]$$

. So, this is what we are finding.

(Refer Slide Time: 26:54)



Now from here now if $k \to \infty$ then let us see what will happen. Then in this case you know that so all these factors if you see then this factor if you see $\lambda_2 \lambda_1$, if you see this is very less than 1 in magnitude. So, I can take the value of this one.

Similarly, $(\frac{\lambda_n}{\lambda_1})^k < 1$. From here I can see that now if I chose λ_1, this is the condition we have taken that in magnitude it is value is 1. It may happen. So, let us see what will happen. Suppose I take $\lambda_1 = -8$ and λ_2 is suppose 4 and λ_3 is suppose -2. Then what will happen?

So, in magnitude this is the highest value. So, I will take λ_2 over λ_1. So, it is 4 over -8. This is -1 by 2. Now lambda 3 by λ_1, so it is -2 by -8, so it is 1 by 4. So, that is a value less than 1. This is also less than 1 in magnitude. So, this value is in magnitude less than 1 and this value is in magnitude less than 1.

Now what will happen if I $(\frac{-1}{2})^k$? So, in this case this will be $(-1)^k$ and then $(\frac{1}{2})^k$. Now if I $k \to \infty$ then I know that this is going to be 0. So, from here I can say that this factor in magnitude is also going to be 0. So, if I put this k tends to infinity from here I can say that limit k tends to infinity xk and then I put the limit again. I take the limit on both sides. So, this will be

$$\lim_{k\to\infty} x_k = \lim_{k\to\infty} \frac{\lambda_1^k}{c_1 c_2 \ldots c_k}\left[a_1 x_1 + a_2(\frac{\lambda_2}{\lambda_1})^k x_2 + \ldots + a_n(\frac{\lambda_n}{\lambda_1})^k x_n\right]$$

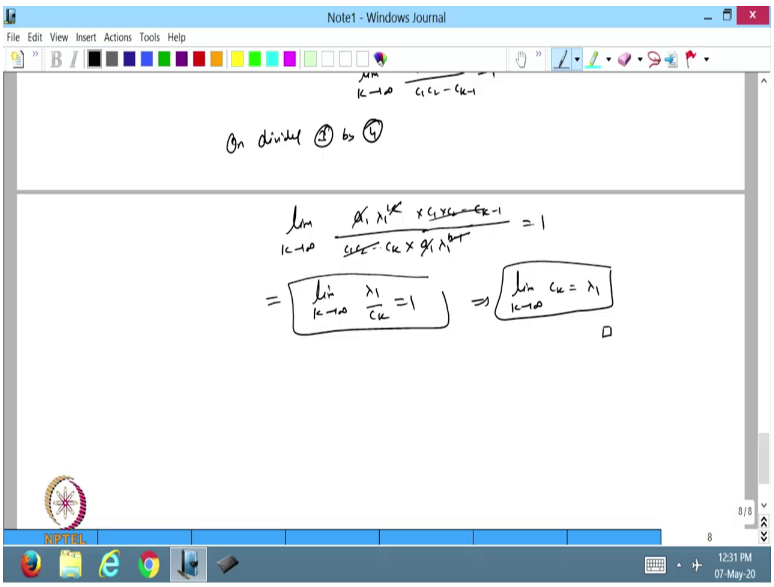
and now we also assumed that these matrices are distinct. These eigenvalues are distinct and real also but if it is not real then we have to take the magnitude. So, in that case we can take the magnitude of this one when the eigenvalues are complex.

Now if you see from here then I can write from here that the $\lim_{k\to\infty} x_k$ and on the right hand side this tends to 0. All these terms tend to 0. So from here I will get this will be equal to $\lim_{k\to\infty} x_k = a_1 \frac{\lambda_1^k}{c_1 c_2 \ldots c_k} x_1$

Now from here also I know that on the left hand side it is also converging to x1 because we know that whenever we will get $Ax_1 = \lambda_1 x_1$ only then the method will converge. So, this is the same eigenvector. So, from here I can say that this implies that as k tends to infinity, this factor $\lim_{k\to\infty} a_1 \frac{\lambda_1^k}{c_1 c_2 \ldots c_k} x_1$ should converge to x1.

So, that implies that $\lim_{k\to\infty} a_1 \frac{\lambda_1^k}{c_1 c_2 \ldots c_k} = 1$. Now from here this is my condition. Now the same I can write from here, $\lim_{k\to\infty} a_1 \frac{\lambda_1^{k-1}}{c_1 c_2 \ldots c_{k-1}} = 1$ because I am just in place of k I am putting k-1. So, that is it.

(Refer Slide Time: 32:35)



Now I will get 2. So, I call it 3. I call it 4. So, dividing 3 by 4 will get the limit k tends to infinity. So, I will get $\lim_{k\to\infty} a_1 \frac{\lambda_1^k c_1 c_2 \dots c_{k-1}}{c_1 c_2 \dots c_k \times a_1 \lambda_1^{k-1}} = 1$. So, this will cancel out with this and this will cancel out with this.

So, from here I will get the $\lim_{k\to\infty} \frac{\lambda_1}{c_k} = 1$ and from here I can say that the $\lim_{k\to\infty} c_k = \lambda_1$. So, that is my convergence of this one. So, if my vector is converging to x1, then that x1 is the dominant eigenvector then the eigenvalue ck will converge to the λ_1.

So, that is the proof of this power series theorem. So, let us stop today here. So, today we have started with the power series method. Then to find the dominant eigenvalues and the dominant eigenvector. So, I hope you have enjoyed this lecture. Will continue from the next lecture. So, thanks for watching this. Thanks very much.