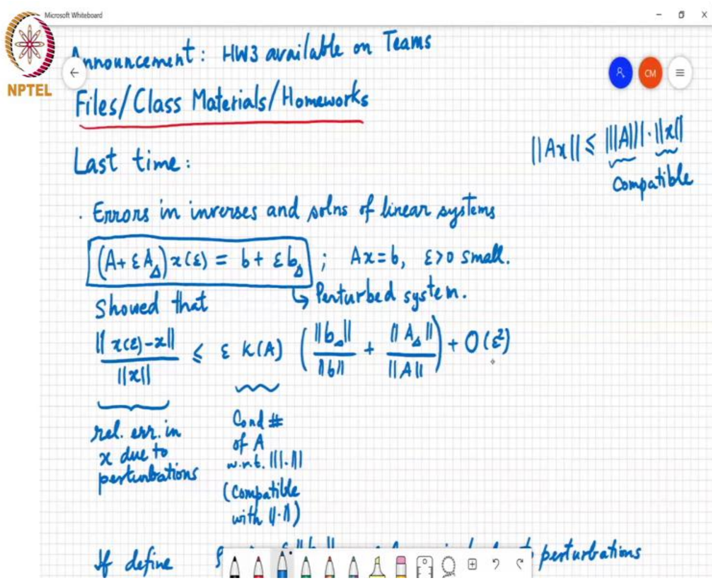
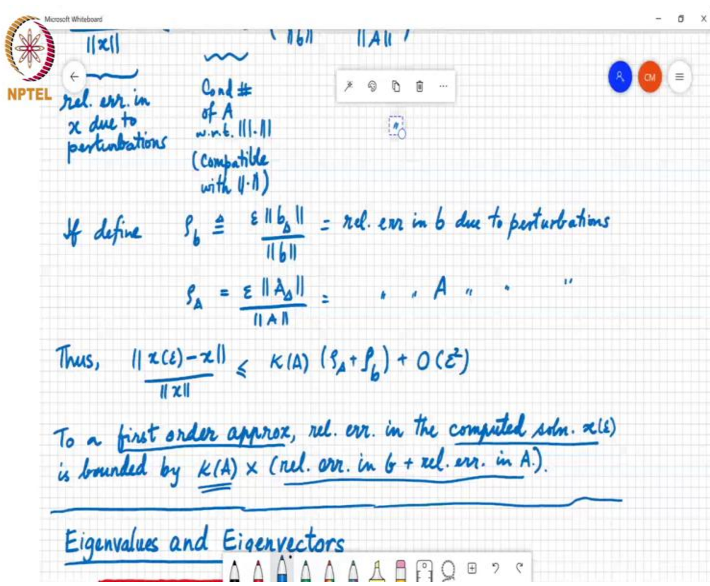
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Introduction to Eigenvalues and Eigenvectors

(Refer Slide Time: 00:14)





So, the last time we were looking at errors in inverses and solutions to linear systems. A related concept that we saw was the concept of compatible norms and compatible norms are norms such that they satisfy a sub multiplicativity type property, but with a combination of vector norm and a matrix norm. So, specifically there is a vector norm such that Ax is less than or equal to the norm of A times the vector norm of x then this vector norm and this matrix norm are said to be compatible.

So, yeah, so based on this we were looking at bounding the errors and computing solutions to linear systems or equations, we saw one formulation and then towards the end of the previous class, we saw this other formulation where we look at perturb system, this one here. So, where the matrix A and the right hand side B are both perturbed by the matrix A delta and a vector B delta multiplied by some small number epsilon.

And we are interested in understanding the first order behavior of how the solution x epsilon is related to the solution x as the (())(1:48) for very small values of epsilon. And so, this is what we call the perturbed system. And we showed that if you look at the relative error in x due to perturbations, that is the norm of x epsilon minus x divided by the norm of x that can be written as, that can be upper bounded by epsilon times the condition number of A with respect to some norm, which is actually a compatible norm with the norm used to evaluate this relative error in x.

So, this times the relative error in B, actually, epsilon times this is the relative error in B. And epsilon times norm of A delta over A is the relative error in A plus a term which is order of epsilon square. So, if epsilon is small enough, this term will dominate this term. And so you can drop this term. And essentially what, what we said is if yeah. So, as I mentioned, the relative error in B is epsilon times norm of B delta divided by norm of B.

We will call this rho B and similarly, rho A is epsilon times norm of A delta over norm of A, then we can upper bound, this relative error in x hence the condition number times rho A plus, plus order epsilon square. So, this is the punch line that to a first order approximation, the relative error in the computed solution x of epsilon is bounded by the condition number of A times the sum of the two relative errors.

So, that is what, that is where we stopped in the previous class and again, it brings out the fact that if you have a well conditioned matrix K of A will be close to 1 and so, the error in the solution is going to be of the same order as the error in A or B. Whereas if A is a poorly conditioned matrix K of A will be a large number. So, the error in the solution will be much larger than the error in either A or B.

It is not will be, it is this is the bound we have on this so again, it is possible that for specific A's and specific right hand sides b the error relative error and the solution may not be as big as this, it does not mean that it will always be as big as this but this is the bound we are able to get. So let us continue. So now that concludes this chapter on norms. So now, I am moving.

Student: Sir, one question.

Professor: Yeah.

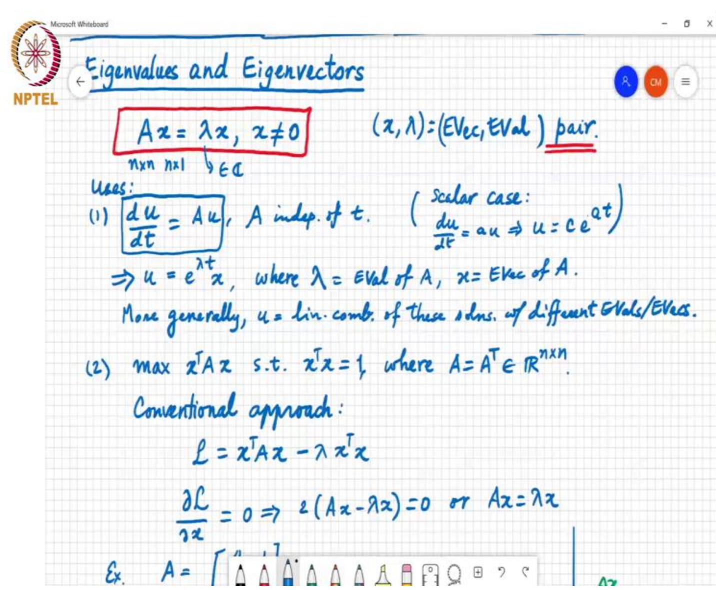
Student: Sir, now that we derive this thing, for, the same Epsilon for matrix A and B, that is same perturbation parameter. So, if the epsilon is different for A and B will the final formula that relative error plus B plus relative error in A times condition number will hold.

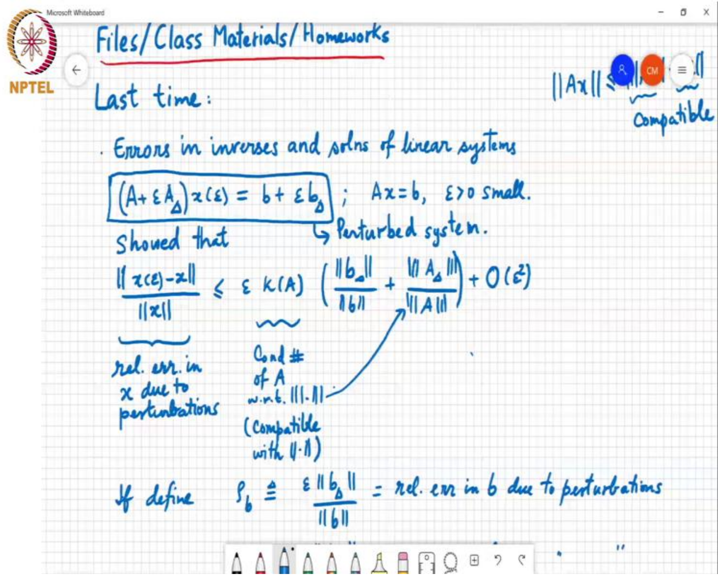
Professor: So, these are all just upper bounds. So, if you want to perturb them by different amounts, a simple fix to this is to take epsilon to be the max of the two epsilons. And everything we are saying here is valid. Here, we are just looking at the first order behavior, the sensitivity to small perturbations in A or B. And the punch line is again that whatever is the relative error in B plus the relative error in A, that get amplified by this coefficient K of A.

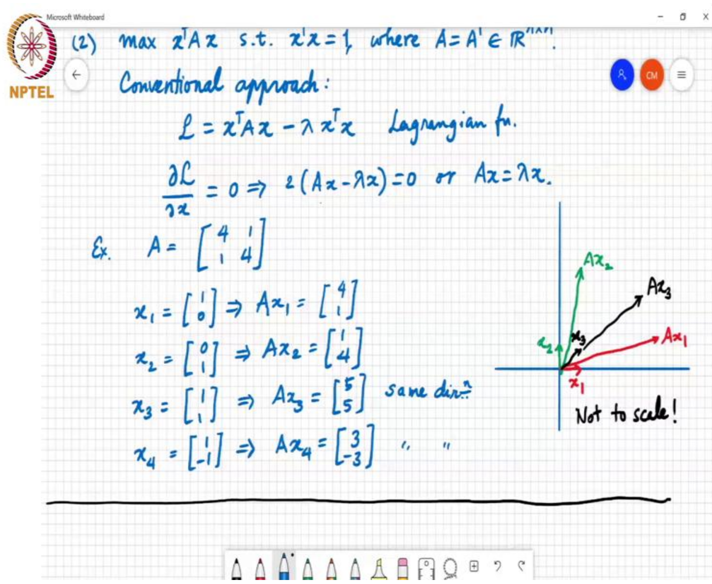
Student: Yes sir.

Professor: No, do that, and delete. So, we will continue.

(Refer Slide Time: 06:04)







So, now I come to actually chapter one of the Holland Johnson textbook, which is on eigenvalues and eigenvectors, so.

Student: Sir?

Professor: Yeah?

Student: Can you scroll up a little bit? Yeah. No scroll up in the first page. Yeah. Sir, again, somewhere you have written. Sir is this see, is this normal matrix norm per A what you have written?

Professor: Ha good question. This should have been the matrix norm, Yeah.

Student: Yeah. Sorry,(())(06:44).

Professor: The typo, this is the same matrix norm that is used here to compute the condition number of A. So, now we come to, now you sort of how to rewire your brain a little bit and this is a different topic, its eigenvalues and eigenvectors again, a super important topic from the point of view of matrix theory. So, the fundamental equation of eigenvalues and eigenvectors is very simple. It is stated right here, Ax equals lambda x and here A is an n cross n matrix.

And x is an n by 1 vector, lambda is a scalar and so lambda belongs to C. And x can be any vector, as long as it is not the all 0 vector, of course, the all 0 vector will satisfy this equation, but we do not care about that that is a trivial solution. So, for non 0 x, if you can find a vector x and a scalar lambda such that Ax equals lambda x, we call this pair x and lambda as an Eigen value, this is not correct. To be sort of consistent, this is an eigenvector eigenvalue pair.

And the key thing, reason I underlined this here is because these always occur in pairs. So, you associate an Eigen value and associated with any Eigen value, I mean you cannot define an Eigen value without saying that there is an x naught equal to 0, such that x Ax equals lambda x. And similarly, you cannot say that x is an Eigenvector, without saying that, for some complex valued lambda, Ax equals lambda x holds. So, they are always in pairs.

So, just to motivate there are two quick simple examples where eigenvalues and eigenvectors matter. So, suppose you want to find a solution to this differential equation here, du over dt equals au. Here a is some constant matrix, which is independent of t. But u is a function of t and you are trying to solve for u of t, u of t is a vector, and it is evolving with time. And the way it evolves is such that it satisfies this differential equation, du by dt equals au.

Of course, in the scalar case, you have certainly seen this in your undergraduate program, if I give you this equation du by dt equals au, you will take u to down there and dt up here and then you will integrate both sides, you will get log u is equal to at. And from that, you get u is equal to some constant times e power at. Where the value of the constant depends on the initial condition that is the value of u at t equals 0.

So, if somebody tells you what the value of u at t equals 0 is you can find what this constant is, and you know that this is the solution. So, if I had this and say what happens in the matrix case, I could potentially think about writing capital A in the exponential here. But for now, just consider u is equal to e power lambda t times x, where x is an Eigen value of A. And sorry, lambda is an Eigen value of A. And x is an eigenvector of A.

This if I make the substitution u is equal to e power lambda t times x, if I substitute that, if yeah, so how do I, how do I explain this to you. So, A times u will be A times e power lambda t times x, which is equal to e power, lambda t is a scalar. So, I can take that out of the multiplication. So it is e power lambda t times Ax, and Ax is the same as lambda x, because lambda and x are an eigenvalue eigenvector pair. So, A times u is equal to lambda times e power lambda t times x. So, basically, au will be equal to lambda times U.

And then if I consider du by dt, so if I differentiate this with respect to t, the only thing that depends on T is just this e power lambda t and its derivative is lambda e power lambda t. And so du by dt is also equal to lambda times e power lambda t times x, which is equal to lambda times u. So, au is lambda times u and du by dt is also equal to lambda times u. So, it satisfies this differential equation.

And more generally, u can be written as a linear combination of solutions of this form corresponding to different eigenvalues and eigenvectors. Now, now another problem is, suppose you want to solve a constrained optimization problem, such as maximize x transpose Ax, subject to the constraint x transpose x equals 1, where A is a real valued matrix, which is also symmetric.

So, A equals A transpose, then the conventional approach is to use the method of Lagrange multipliers, where you define this Lagrangian function, L which is Ax transpose Ax minus lambda times x transpose x. Then if we differentiate this with respect to x, now this is a vector derivative. So, you will have to take this on faith. But the simple explanation is that the way to differentiate with respect to a vector is to differentiate with respect to each of the components of the vector partially and then stack them together as a vector.

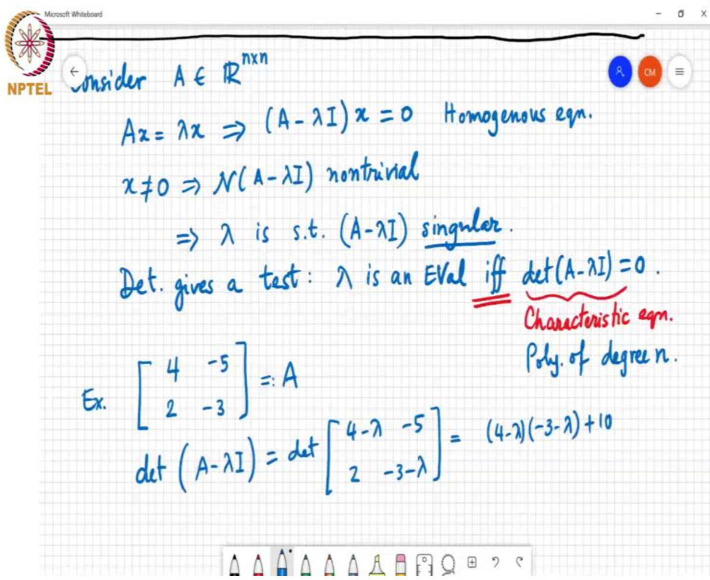
The derivative of a scalar with respect to a vector is a vector whose dimension equals the dimension of the vector. And the entries are equal to the derivatives of the scalar with respect to each of the components stacked one above the other. And if you do that, for this particular Lagrangian function, it is not difficult to show that the derivative is two times Ax minus lambda x. And so if you set the derivative equal to 0, you get this equation here, two times Ax minus lambda x equals 0.

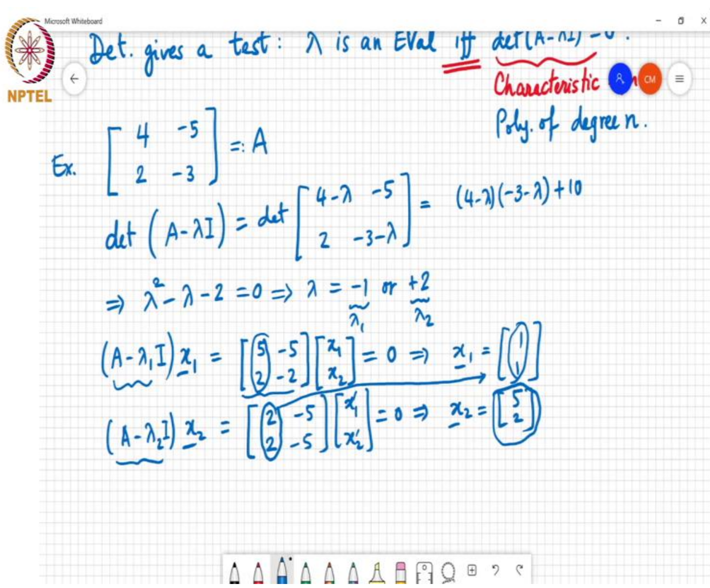
Or, in other words, Ax equals lambda x, which is the eigenvalue eigenvector equation. So, these are two simple examples where eigenvalues and eigenvectors that arise naturally, and you are trying to solve some problems. And here is an example to just visualize eigenvalues and eigenvectors. So, suppose is A simple 2 cross 2 matrix with entries 4 1 1 4. If I take x1 to be this vector, 1 0 then Ax 1 will be the first column of this matrix which is 4 1.

So, here I show that in red, so x1 is the vector 1 comma 0, and Ax1 is going along this direction, the x component is 4 and the y component is 1. And similarly, if I take x2 to be 0 1, then Ax 2 is the second column of this matrix, which is 1 4. And that is shown in green here, Ax2 is in this. So, you will see that x1 and x2, sorry, x1 and Ax1 point in different directions, x2 is like this Ax2 is pointing in a different direction.

Whereas if I choose x3, to be 1 1, then when I do Ax3, I get 5 5, which is five times this vector x3, so it points in the same direction as x3. So, that is shown in black here. So, similarly, if I take x4 as 1 minus 1, then Ax4 will be 3 minus 3, which is also pointing in the same direction, which is three times 1 minus 1. So, these Eigenvectors are very special vectors, where when you multiply the matrix by the Eigenvector, you get a vector that is pointing in the same direction as the original vector. So, how do we find these Eigenvalues? I guess, you guys know this already.

(Refer Slide Time: 16:11)





So, just quickly, for the sake of completeness, discuss this. So, consider A to be an n cross n matrix. It could be complex also the same thing, whatever I am going to say holds for complex also. So, if we consider the equation Ax equals lambda x, this implies that A minus lambda times the identity matrix times this vector x equals 0. And this kind of an equation where you have a matrix times a vector equals 0, this is called a homogeneous equation, the right hand side is 0 that is when it is called the homogenous equation.

So, one thing is that we wanted these Eigenvectors to be non 0 vectors. Of course, if I said x equals 0, this will satisfy this equation. But we do not want that solution. So, if x should be non

0, that means that this A minus lambda I must have a non-trivial null space. Or in other words, lambda is such that A minus lambda I is singular. So, at this point, it should look little magical to you. So, see lambda I is a highly structured matrix, it is just a scaled version of the identity matrix.

By subtracting lambda times I from A, I am able to arrive no matter what A is, if I take the right scaling lambda here and do A minus lambda I, the columns of A, A minus lambda I should become linearly dependent, then this matrix should become singular. So, I am looking for such kinds of lambdas. And if I want this matrix to be singular, one way to test it is to find its determinant.

And whenever the determinant of this matrix goes to 0, we know that this A minus lambda I is this going to be rank deficient and this matrix will be singular. So, the determinant gives us a test. So, lambda is an Eigen value if and only if determinant of A minus lambda I equals 0. So this is very, very crucial observation that lambda is an Eigen value if and only if this determinant of A minus lambda I goes to 0.

Obviously, if determinant of A minus lambda I is 0, then it means that the A minus lambda I is singular. And therefore, you will be able to find a non 0 vector x such that A minus lambda I times x is equal to 0. Contrary wise, if lambda is indeed an Eigenvalue of A, it implies from the definition that there is a non 0 x such that Ax equals lambda x, or A minus lambda I times x equals 0 for some x naught equal to 0, which means that the matrix A must be singular.

So, and therefore its determinant must be equal to 0. So, these two statements, these two points are actually an if and only if statement. A lambda is an eigenvalue of A and determinant of A minus lambda I equals 0 are if and only if conditions. And this equation determinant of lambda A minus lambda I equals 0 is called what?

Student: (())(20:11).

Professor: Yes, it is called the characteristic equation, it is a polynomial of degree of n, that comes about if you simply expand this definition from the definition of the determinant, you will see that this will be a polynomial of degree n. And so, and also corresponding to any eigenvalue lambda, there will always exist, at least one non 0 Eigenvector, by definition, they always occur in pairs I am repeating my point.

So, that is it. So, this is how we find the Eigenvalues, we have to set the characteristic equation or find the solutions or roots of the characteristic equation. And that gives us all the eigenvalues of the matrix. So, again, for example, if I consider the matrix 4 2 minus 5 and minus 3, then if I consider determinant of A minus lambda I equals, so if I so let us see determinant of. So, this is my A so let us say, A minus lambda I is the determinant of 4 minus lambda minus 5 2 and minus 3 minus lambda, which is equal to 4 minus lambda minus 3 minus lambda plus 10.

And if I set this equal to 0, then I will get, you have to simplify this, so that will give you lambda minus lambda times lambda is lambda square, and 3 lambda minus 4 lambda gives me minus lambda. And then I have minus 12 over here, but there is a plus 10. So, I am left with minus 2 equals 0. And the solutions to this are lambda equals minus 1 or plus 2. So, so if I now compute, so these are the two eigenvalues of this matrix, and if I now compute A minus lambda one times the identity, so let us call this lambda 1. And let us call this lambda 2.

Then, this times, if I take x1 to be a corresponding eigenvector, this will be equal to 5 2 minus 5 minus 2 times, now this is a slight abuse of notation, I will call this vector x1, so I will call this x1 x2. And I set this equal to 0 and I want to solve for what x1 x2 satisfies this it is easy to verify that x1 is the vector 1 1, if I just take 1 1 here, this could become 0, this becomes 0. And similarly, if I take A minus lambda 2 I times x 2, that becomes this minus, so 4 minus 2 is 2.

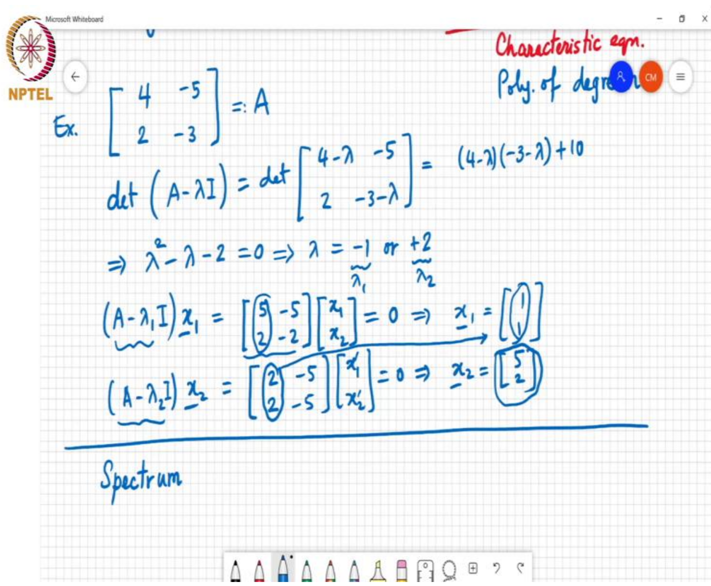
And this is 2 and this is minus 5 and minus 3 minus 2 is minus 5 again, this times say x1 dash x2 dash equals 0, that implies the vector x2 is equal to I can just take it to be 5 comma 2. So, notice that basically, if I take A minus lambda 1 I, that is this matrix, the column of this is 5 2, and that gives me x2. And similarly, if I do A minus lambda 2 I that gives me this column and that is going to be equal to x1.

So, just maybe just for I will just indicate it like this. So, this is something that is an interesting observation that the columns of A minus lambda 2I actually give you x 1, the Eigenvector corresponding to the first Eigen value, and vice versa. So, this only works for 2 by 2 matrices, it does not work for larger than dimensional matrices. But nonetheless, it is an interesting observation.

So, basically, when I multiply A with vector x, most vectors will not satisfy Ax equals lambda x only special numbers are Eigenvalues and special vectors are Eigenvectors. Normally, if I take A x, it will scale the different components of x by different amounts, and it will rotate the vector x and so it will not point in the same direction the ones that point in the same direction are called Eigenvectors and there is a corresponding scaling factor which is denoted by which is defined to be the Eigen value.

So, this is the basic notion of Eigen value and eigenvectors and how to find Eigenvalues and once you found the Eigenvalues, you compute A minus lambda I for each Eigen value and you find one vector in the non trivial null space of A minus lambda I and that gives you an idea or rather you find a basis for the span of the null space of this matrix and that gives you the Eigenvectors corresponding to that Eigen value.

(Refer Slide Time: 26:48)



So, now, couple of more definitions.

Student: Sir?

Professor: Yes?

Student: There is not any restriction on A, I mean A only should be a square matrix, it can be similar as well. Still, Eigenvalue and Eigenvector will adjust?

Professor: Yes. So, basically if A is singular then there is an x which is non 0 such that Ax equals 0. But of course, I can write 0 as 0 times x. So, then it satisfies Ax equals lambda x where lambda equals 0. So, if A is singular then certainly you can say that lambda is equal to 0 is one of the eigenvalues.

Student: Yeah, yeah Sir, thank you.