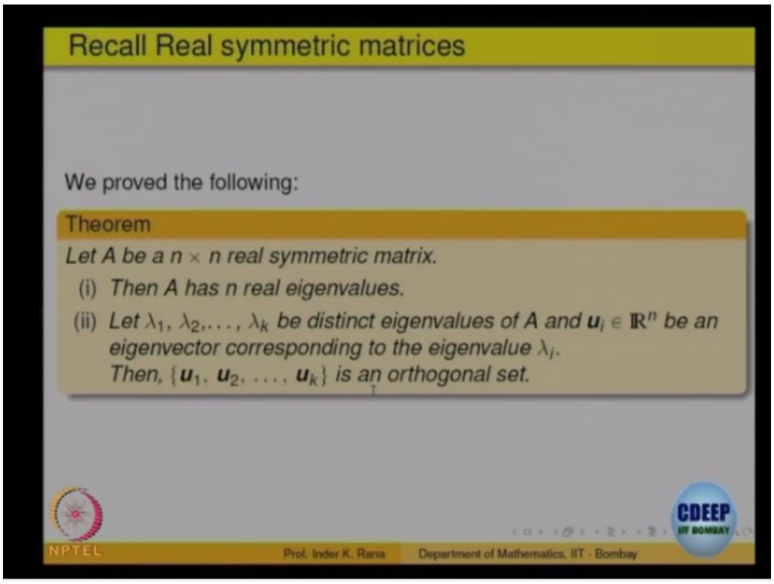
**Basic Linear Algebra**
**Prof. Inder K. Rana**
**Department of Mathematics**
**Indian Institute of Technology – Bombay**

**Lecture - 34**
**Diagonalization and its Applications - I**

Okay so let us begin today's lecture by recalling the last lecture we had proved the following theorem.

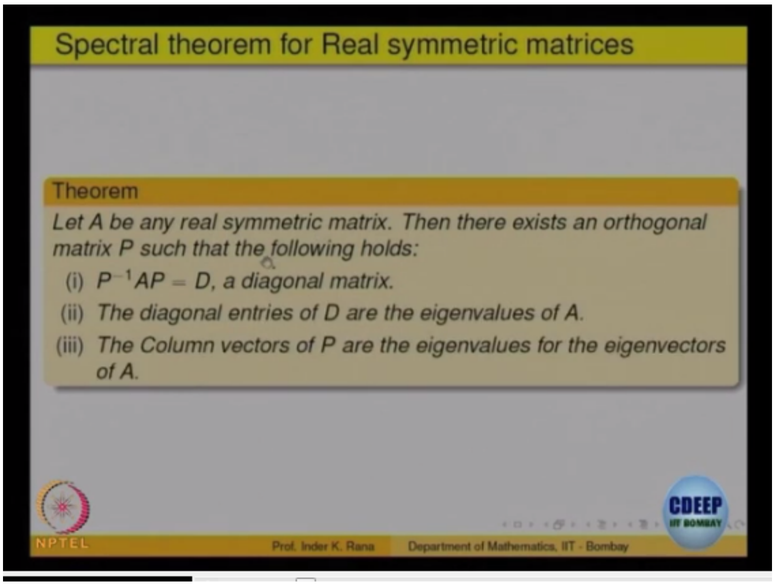
**(Refer Slide Time: 00:39)**




That if A is n x n real symmetric matrix, one it has n real eigenvalues and secondly if lambda 1, lambda 2, lambda k are distinct eigenvalues and ui's are corresponding eigenvectors then the set u1, u2, uk is an orthogonal set, so basically saying that for a general for a given matrix the eigenvectors are linearly independent for distinct eigenvalues. If it is a real symmetric matrix, then the eigenvectors corresponding to distinct eigenvalues are also orthogonal.

So using this because it has already got n eigenvalues and so there will be n eigenvectors right and we already know that because they are orthogonal, they are also linearly independent.

**(Refer Slide Time: 01:35)**




So improvement on the previous theorem which was for diagonalizable, diagonalization of matrices says the following. That if A is a real symmetric matrix, then there exists an orthogonal matrix P such that P inverse AP is a diagonal matrix okay and the diagonal entries are the diagonal eigenvalues of the matrix A. The column vector P are the eigenvectors of A, this will be the eigenvectors of A.

So basically the same process that we do for general diagnozability. Given a matrix A, we find eigenvalues, we find the corresponding eigenvectors right. If it is real symmetry, they are going to be n eigenvalues right accounting multiplicity, so each eigenspace is going to have a basis consisting of eigenvector right. So but the only thing is that eigenvectors corresponding to distinct eigenvalues for real symmetric will be orthogonal but inside the eigenspace they may not be orthogonal.

They will be only linearly independent, so what we can do is we can make using Gram-Schmidt process, we can make those eigenvectors also mutually orthogonal. So for each eigenvalue will have the eigenvectors right, there is a basis consisting of eigenvectors because it is diagonalizable. So we can find orthogonal, orthonormal basis convert that basis to an orthonormal basis.

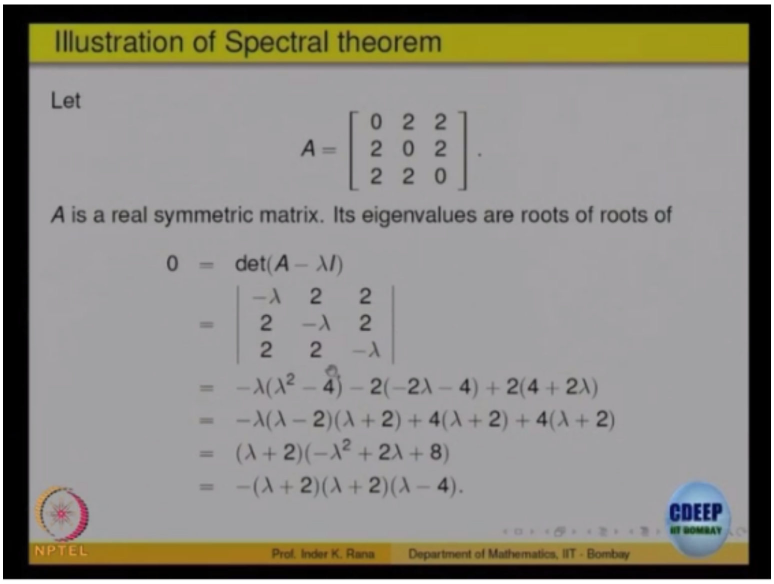
And this says that for a real symmetric matrix this will always happen right, so for a real symmetric matrix there will be n eigenvalues right. There is a basis consisting of eigenvectors right which form an orthonormal set right. So there is a basis consisting of orthonormal vectors, so that is what is called the spectral theorem for real symmetric matrices right. So keep in mind for a general matrix, it may not be diagonalizable at all.

Because it may not have eigenvalues, even if it has eigenvalues the algebraic multiplicities may not be equal to the geometric multiplicities but for a real symmetric matrix, always there exist eigenvalues, algebraic multiplicity of each eigenvalue is equal to the geometric multiplicity right and as a consequence it becomes diagonalizable right. Because it is diagonalizable, eigenvectors corresponding to distinct eigenvalues are mutually orthogonal.

But inside each eigen-subspace right, the basis exists but it may not be orthogonal. So you can use Gram-Schmidt process to make it orthogonal right. So that is the basic idea of this spectral theorem. The advantage is that this matrix P is an orthogonal matrix, so for an orthogonal matrix what is P inverse? You know orthogonal means P transpose P is=identity right, so that means the inverse is the transpose itself.

So there is advantage here that for a real symmetric you do not have to compute P inverse, it is just the transpose of the matrix P that you get okay.

**(Refer Slide Time: 05:09)**

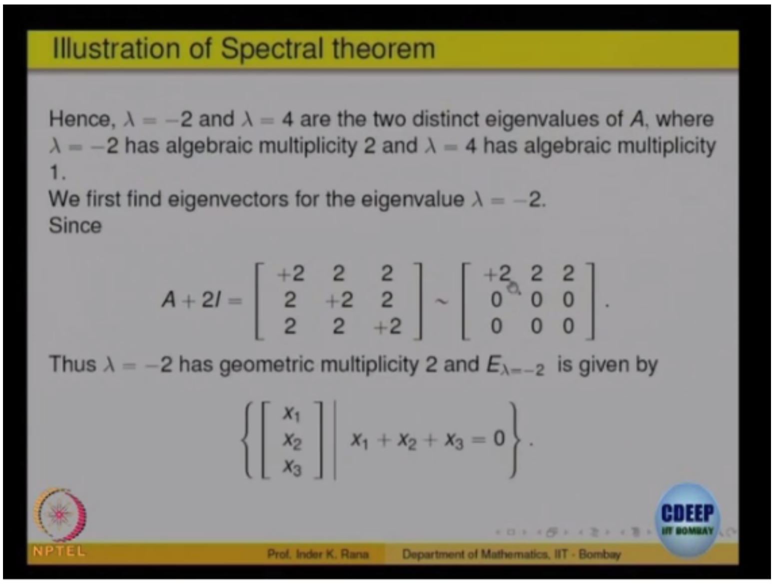



So we will look at one example completely illustrating this idea. So look at the matrix A with the entries 0 2 2 2 0 2 and 2 2 0. So first observation is this is a real matrix and it is symmetric. So the row here 0 2 2 is same as the column here. So interchange rows and columns, this is a symmetric matrix right. So as our theorem says it should have n eigenvalues right.

So that means the characteristic polynomial should be completely factorizable. So here the characteristic polynomial will be a cubic of degree 3 right. So how do you find that? So look at A-lambda I, so 0-lambda 0-lambda. So along the diagonal it will become –lambda, expand that determinant okay. So that determinant 3 x 3, you expand and simplify, it comes out to be –lambda+2 lambda +2 and lambda-4, so there are 3 factors right.

So that means what? Lambda=-2 is repeated, it has algebraic multiplicity 2 and lambda=4 has got algebraic multiplicity 1 right. So lambda=4 that is not going to be a problem because that is eigenvalue lambda=4 you can find an eigenvector corresponding to it right. The only work to be done is for lambda=-2 right.

**(Refer Slide Time: 06:46)**

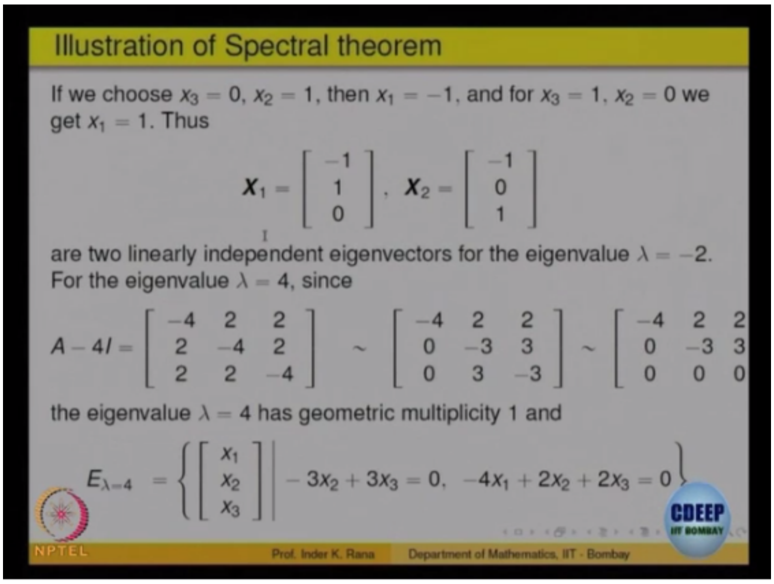



So let us compute that. So algebraic multiplicity of lambda=-2 is 2 and of lambda=4 is 1. So let us look at the eigenvectors for the eigenvalues lambda=-2. So look at A-lambda I, so that is A+2I right. So that it is easy to reduce it to row echelon form. So that gives 2 2 2 and 0 0 0, so that means what is the rank of this matrix? That is 1, that means nullity is=2 right. That means there is going to be 2 linearly independent solutions for this eigenvalue lambda=2.

So how do you find that? You know the process, so what are the variables which are non-pivotal? X1 is pivotal, X2 is non-pivotal, X3 is non-pivotal. So non-pivotal variables will be given the arbitrary values and pivotal compute it in terms of them and because here the dimension is 2, so the methodology is you first give X2=1, X3=0 and find out X1 in terms of that, so that gives you 1.

And second is put X2=0 and X3=1, that gives you another solution for the same eigen-subspace right a null space and those two are going to be linearly independent automatically because we have chosen X2 and X3 suitably okay. So just writing this that means it is 2 times X1 X2 X3 right=0 right. All linear combinations so that in a null space.

**(Refer Slide Time: 08:35)**

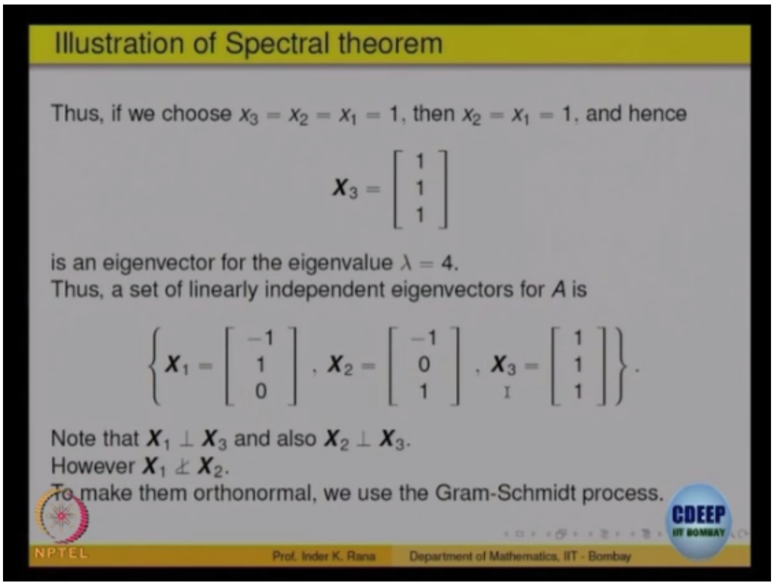



So you get two vectors, eigenvectors for this eigenvalue X1 X2=1 X3=0 and similarly for X2 it is X2 is 0, the component X3 is 1. You compute accordingly other one that is -1 and -1, so these are the two eigenvectors corresponding to the eigenvalues lambda=-2 right and these are linearly independent by our choice of X2 and X3. The only problem is that these they may not be see if you multiply what is the dot product of these two?

So that is -1 -1 that is 1 0 0, so dot product is not equal to 0, so these are only linearly independent but they are not orthogonal, so will orthogonalize them. Before that let us find for lambda=4, so A-lambda I so along the diagonal becomes -4 determinant, reduce it to row echelon form, it comes out this. So one row is 0, two rows are nonzero, so the rank of this matrix is=2.

So rank is=2, nullity is=1, so only one-dimensional null space. So you can find that, so how do you find that? What is the last equation? -3 so what will the last equation give you? -3 X2+3 X3=0. So what are the non-pivotal variables? Non-pivotal is only X3 right third one, so X3 is given arbitrary value, you find X2 from this equation, put those values back, you get the first variable right, only one solution, non-pivotal gets arbitrary values others are obtained in terms of the non-pivotal.

**(Refer Slide Time: 10:37)**

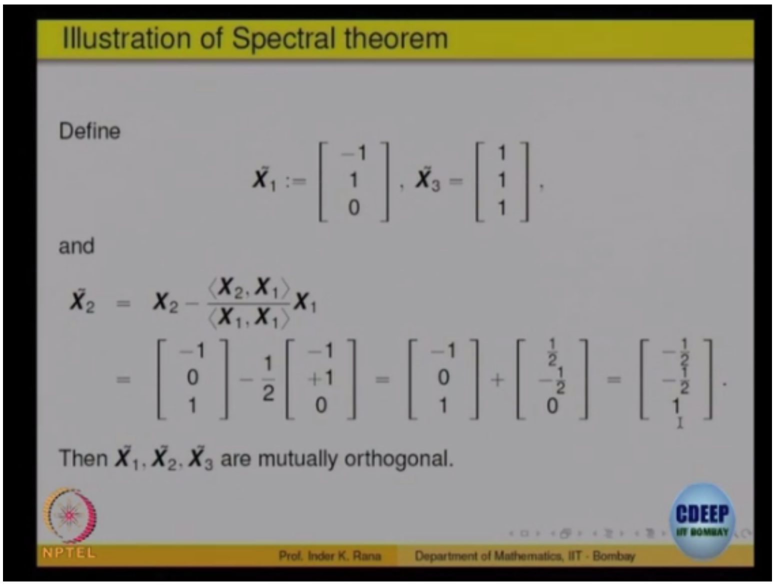



So once you write that you get right you put X2=1, X3 and you get 1 1 1, so these are eigenvectors for lambda=4. Basically, it is solving a system of linear equations right, finding rank, finding nullity. So you get 3 vectors X1, X2, and X3 right and that will always happen for a real symmetric matrix will get 3 linearly independent eigenvectors right if the order is 3 x 3, n x n that we will get n linearly independent eigenvectors for a real symmetric matrix.

Now in this observe, these X1 and X2 were the eigenvectors for the eigenvalue lambda=-2 right and X3 is for lambda=4. So X3 is perpendicular to x2 because they are for distinct eigenvalues right and X1 is also perpendicular to X3, you can just dot product and see that is=0. So eigenvectors corresponding to distinct eigenvalues will be orthogonal but among themselves X1 and X2 are not orthogonal to each other right.

So what we do is we know that X3 is orthogonal to these two, so that is not a problem. So these two we make them perpendicular to each other by using Gram-Schmidt orthonormalization process right. So what will be my first step? So X1 is not perpendicular, so we use so if we are given these 3, how do you orthonormalize them? Gram-Schmidt process from successively you are going to remove the projections right. So let us do that.

**(Refer Slide Time: 12:24)**

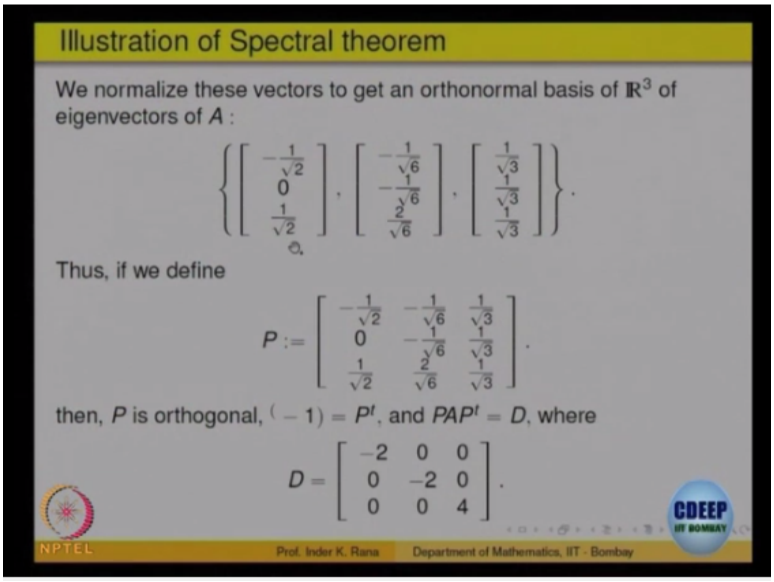



So define X1 and X3, tilde to be as it is okay, the original ones because they are perpendicular to X1 right. They are perpendicular to, X2 is the one which is not perpendicular, X1 and X3 are perpendicular anyway. So let us define that. So X2 is the one which is to be defined now right, either X1 or X2 one of them you can modify so that becomes perpendicular right.

So you define X2 so that will be perpendicular to right X1. X1 and X2 are not orthogonal right; from X2 remove the projection of X1, so that they become orthogonal right. So once you do that, so that is the way you X2-this projection on to X1 right. They are not orthogonal, so remove the projection. So once you compute, it comes out to be this. So Gram-Schmidt process right that is being used to orthonormalize okay.

So you get 3 vectors which are mutually orthogonal. Gram-Schmidt only gives you orthogonal right, will remove the projections, you get only orthogonal but you have to orthonormalize them if you want right. So then you divide each one of them by the norm of it.

**(Refer Slide Time: 13:47)**




So you can divide, so you get these 3 vectors right which are going to be which form a orthonormal set. They are 3 of them, so they will form a basis for R3 and each one of them is an eigenvector right because you have taken eigenvector to only modify them, so these are eigenvectors. So what is the corresponding matrix that you will get? So P is the matrix which is going to be, columns of P are going to be these vectors.

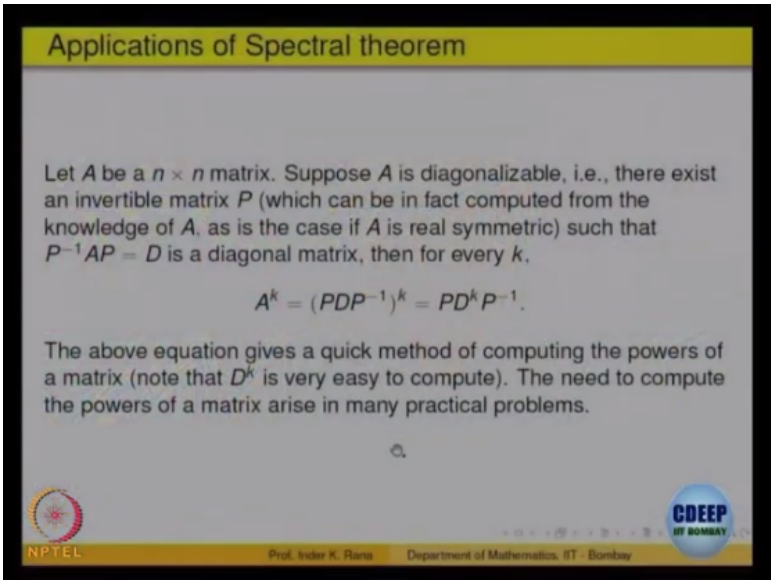
So first column is the first eigenvector, second column second eigenvector, third the third eigenvector, process is same whether it is ordinary diagonalizable matrix or it is real symmetric. You find as many linearly independent eigenvectors as is the order of the matrix. If you are able to find, it is diagonalizable. If not, you are which is not diagonalizable. For real symmetric, you will be able to find 3 not only linearly independent, mutually orthogonal and orthonormal set of vectors which form a basis, so you get this right.

So first column, second column and that is a third column and this matrix is automatically orthogonal because the columns are mutually orthogonal and unit length is 1 right. PP transpose is=identity okay. So there is this is a typo here, orthogonal so the inverse right, so P inverse is=P transpose. So you can write PAP transpose=diagonal, you can verify that okay. So diagnozability of a real symmetric matrix means first find the eigenvalues, you will be able to find as many as the order right.

Unlike the ordinary matrix where it may not have any solution at all right, there may not be any eigenvalue for an ordinary real matrix but here for a real symmetric you will be able to factorize right and you will get n eigenvalues, some of them maybe repeated right but each will have algebraic multiplicity same as geometric multiplicity, for a real symmetric algebraic will be always equal to geometric multiplicity.

Distinct eigenvectors corresponding to distinct eigenvalues will be mutually orthogonal, inside right the eigenvectors corresponding to the same eigenvalue you are able to find as many as algebraic multiplicity, they may not be mutually orthogonal. So use Gram-Schmidt process so that you get orthonormal basis right. So that is the process. So what are the applications?

**(Refer Slide Time: 16:49)**




I think in the beginning I had said once you are able to diagonalize the matrix, there is lot of advantages in that. One is supposing a matrix is diagonalizable, that means there is invertible matrix so that P inverse AP is=D right. So if you want to compute powers of A that becomes very easy because if this is=D then what is A is equal to that is PDP inverse right. From this equation, if P inverse AP is D then what is A?

Multiply on the left side by P, on the right side by P inverse, so you get A is=PDP inverse right. So what is A to the power k? So it is PDP inverse to the power k. Now when you expand it what happens? When you expand say 2, when you multiply it will be P inverse A and the next one will be coming PDP inverse, so PP inverse will cancel right. So it is for any k this power is P the diagonal matrix raised to power k into P inverse.

So you do not have to do anything other than if it is invertible we find out what is the matrix D which is diagonal right, which is the eigenvalues of the matrix A. Once you have found that, powers are very easy to compute, you have to just find D to the power k and for a diagonal matrix what is the kth power? For the diagonal matrix, it is just diagonal entries raised to this appropriate power right. Is this clear?

D square is nothing but the diagonal entries raised to power 2, Dk is diagonal entries raised to power k. So this for a diagonal matrix, it is very easy to compute D to the power k. So this becomes very useful tool and this is useful when you are doing lot of applications. For example, there is something called (()) (19:00) where this comes out as a probability. This matrix A is the probability matrix of going from one state to another of some process right.

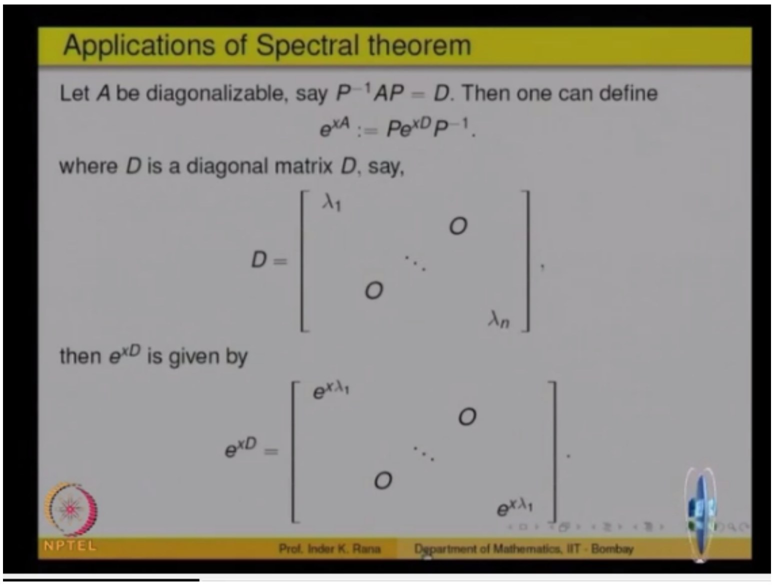
So if you are able to diagonalise that right and you want to apply that again and again and see what happens eventually to the system okay. Let me just give you an idea. If there is a system which is being observed, imagine and there are two possible outcomes of that system, system can write, can go in one state or in another state. There is a probability some probability actually go and stay in state 1, it is some probability it will go from state 1 to state 2 right.

And some probability that will stay in state 2, so you can compute that, you can find that matrix, first entry will be states 1 1 right, second entry in that row is the probability that from state 1 to state 2 bottom state 1 to state 2 and then state 2 to state 2 itself. So that is where there each row and column gives you total probability of this. Now what you want to do? This is the probability applied, X1 X2 will tell you where will X1 go where will X2 go right with this probabilities.

Now if you want to apply again and again right then you want to compute if after k applications where will the system be right. That means multiplying the X1 and X2 by kth power of it right and eventually you want to know whether a system will stabilize somewhere or not, so you want to compute higher powers of the matrix, probability matrix that comes there, so these are applications in many stochastic modeling and such things.

So will not go into that but idea is that once it is diagonalizable, powers are easy to compute, so that is the advantage.

**(Refer Slide Time: 21:10)**




And you will see in your course in differential equations, you may have to compute what is called the e to the power of matrix, exponential of a matrix. You know given a number you can compute e to the power x right but how do you define what is e raised to power a matrix A. So if it is diagonalizable, e raised to power xD is very easy to compute right. So e raised to power xA you can define if it is diagonalizable P e raised to power the diagonal matrix okay.

And e raised to power diagonal is very easy to compute. If the entry are lambda 1, lambda 2, lambda n then e raised to power D is just diagonal entries raised to power that e raised to power that right. It happens actually not only for exponential for any function, any function of D you correspondingly look at f of lambda 1 and f of this. So diagnozability has a lot of applications in computing functions right.

When the domain of a function is a matrix exponential is 1 which most probably will come across when you look at systems of differential equations and finding solutions, so eigenvalues and everything will come back there. So there is one advantage of diagnozability that you can compute powers and the functions of.