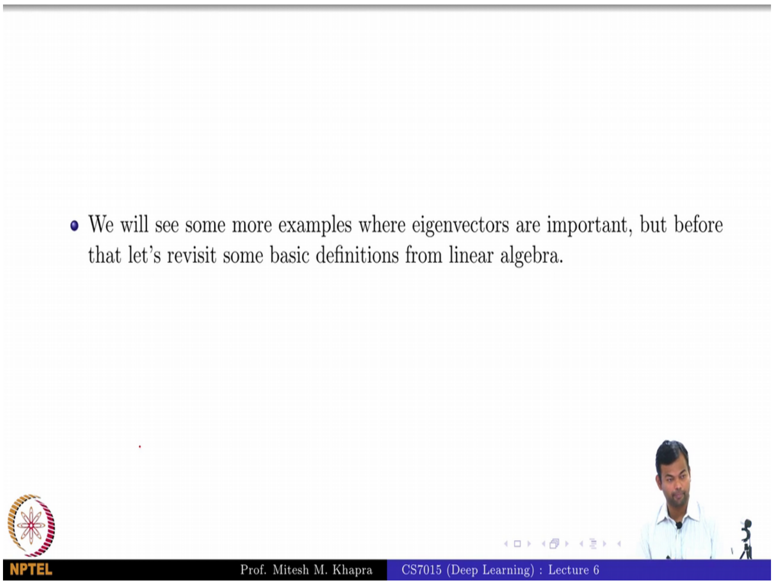
Deep Learning
Prof. Mithesh M. Khapra
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Module – 6.2
Lecture - 06
Linear Algebra - Basic Definitions

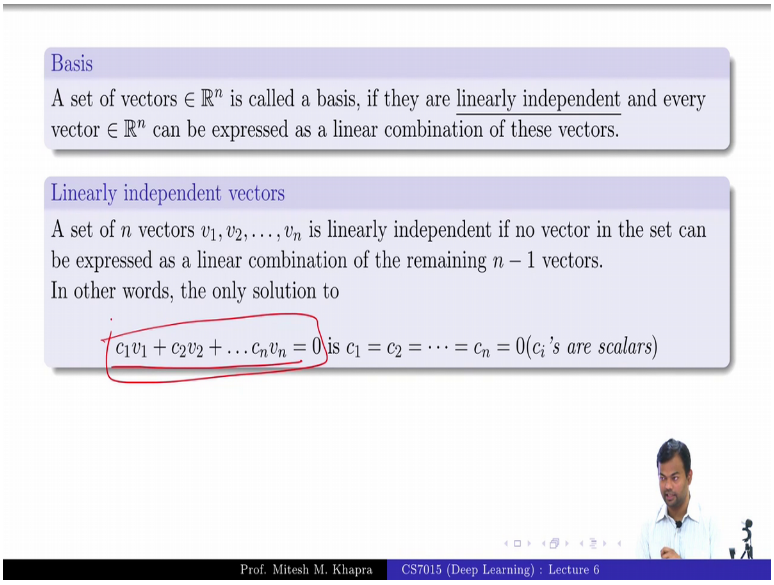
Now, from here on we will go on to something even more basic. We will start defining some basic definitions from linear algebra and these are again important for something that I need in the next lecture. So, let us start with this.

(Refer Slide Time: 00:27)



I mean, in the process we all just see, why the eigenvectors are important for us in this course? Ok.

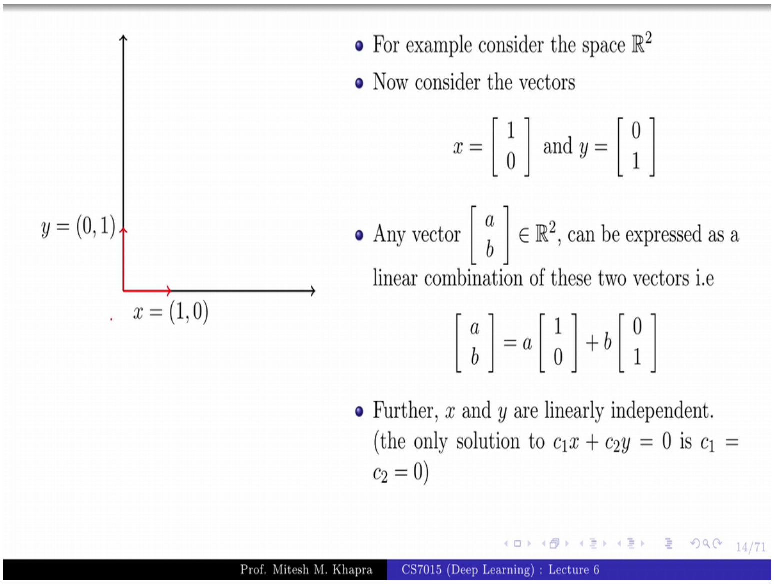
(Refer Slide Time: 00:33)



So, how many of you know what a basis is. So, a set of vectors belonging to R n is called a basis. If they are linearly independent right, and every vector in R n can be expressed as a linear combination of these vectors. So, a set of n vectors v1 to vn is linearly independent. If no vector in the set can be expressed as a linear combination of the remaining n minus 1 vectors ok. So, a more weird we are stating it that. So, that everyone get confuse is that if you take this linear combination ok.

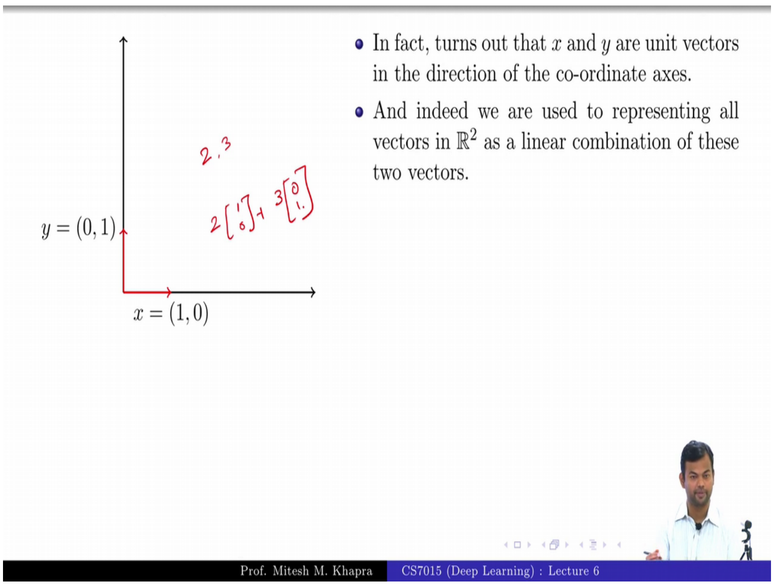
The only solution to this is all the ci's is equal to 0 and that make perfect sense right, that is that same as that linear combination, linear independence and all that thus, is make sense to everyone ok. So, what does linear independence mean that any vector from this set cannot be expressed as the linear combination of the other set; other vectors in the set and a more formal way of saying that is this ok. Everyone gets this what is linear independence.

(Refer Slide Time: 01:35)



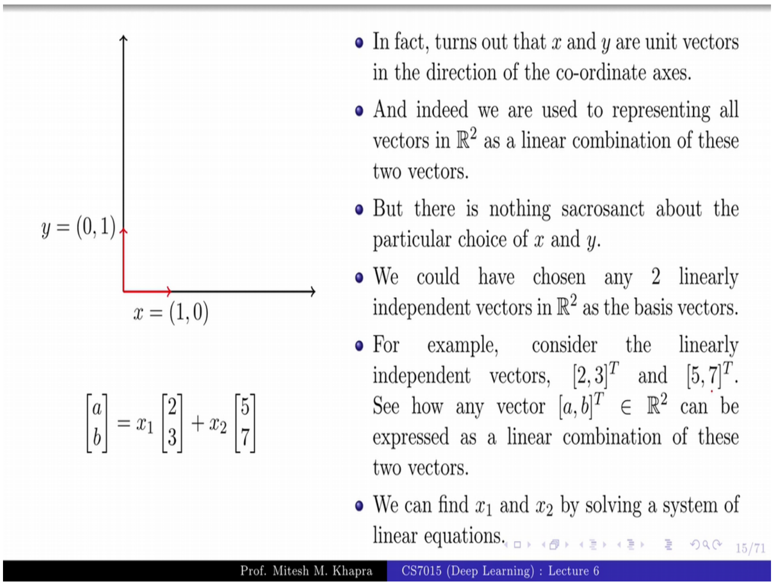
Now, let us consider some very stupid examples, again the space R 2 and we consider these 2 vectors 1 0 and 0 1. Are they linearly independent? Yes, they cannot be expressed as a multiple of each other right. Now, any vector ab belonging to R square can be expressed as a linear combination of these 2 vectors ok. And x and y are linearly independent the only solution is c 1 x plus oh sorry, c 1 x plus c 2 y is c 1 and y equal to 0; What about if I move to R 3 1 0 0 0 1 0 and 0 0 1. So, x y and z axis right are the unit vector.

(Refer Slide Time: 02:12)



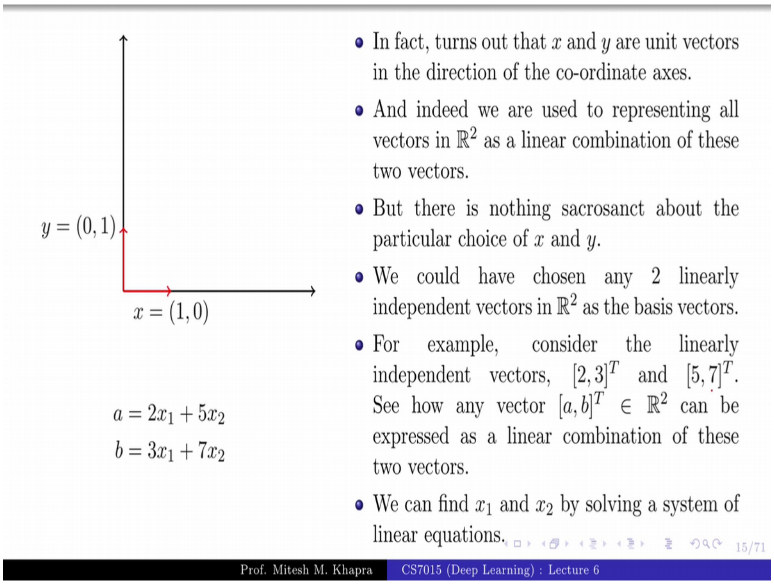
So, in that x and y turns to be unit vectors in the direction of the coordinate axis. And we are used to representing every point in R 2 as a linear combination of these 2 vector is that exactly what I what we do. So, when we say that I have a point 2 coma 3, I am actually telling you that the point is 2 1 0 plus 3 0 1 right. I am expressing at are the linear combination of the coordinate axis.

(Refer Slide Time: 02:38)



But now this nothing sacrosanct about x and y right; I could have chosen just about any other axis ok. So, in particular we could have chosen this as our basis are these 2 vectors linearly independent can any vector and R 2 be expressed as a linear combination of these 2 vectors sure. So, I give you a vector a b, how do you going to express it as a linear combination of these 2 vectors? So, you will do it this way right. How will you find that values are the x 1 and x 2; so, other linear system of linear equations right.

(Refer Slide Time: 03:09)

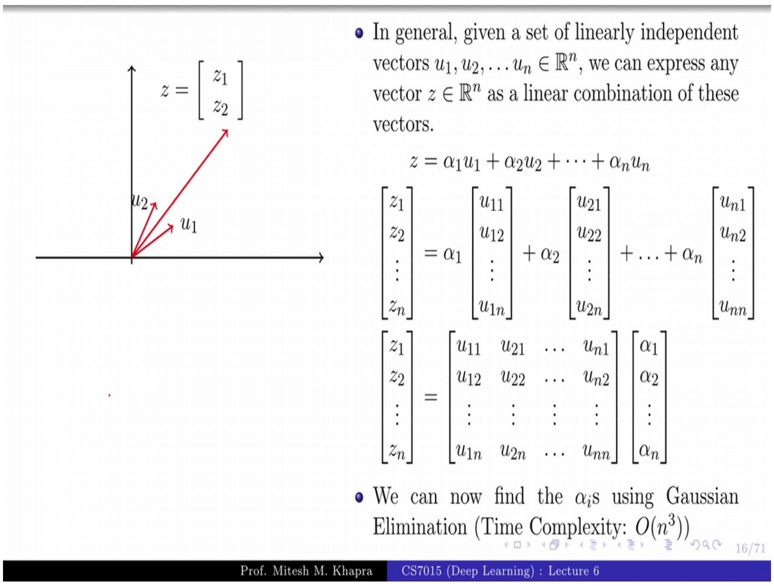


So, this is what you will do. I know all are good in doing this and what do we actually do when we do this? What is the algorithm that we use? How do we solve this? What is the algorithm that you use solving this?

Student: Gaussian elimination

Gaussian elimination right, in 2 variables of course we do not call it an algorithm. That is what we did in 8 standard or something, but when we come to engineering we call it Gaussian elimination right; So, the same algorithm ok.

(Refer Slide Time: 03:29)



So, in general given a set of linearly independent vectors, we can express any vector that belonging to Rn as a linear combination of these vectors right. I can say z is equal to alpha 1, u 1 plus alpha 2 u 2 and so on given alpha 1 to alpha n are linearly independent ok. So; that means, any vector in Rn can be expressed using these vectors which form the basis of Rn does that make sense ok. That is why call the basis vector because, anything else these are the fundamental vectors using these anything else can be expressed in that space it is that clear ok.

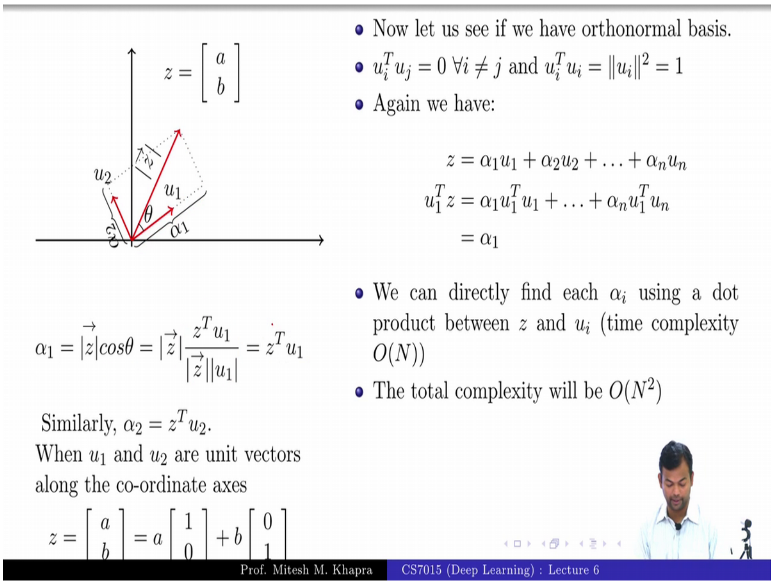
So, this is how it will be. How do I write this in matrix notation A? There are lot of these and these thing I do not really understand what you mean by that ok, yeah good. So, this is what you mean. So, that we writing same in matrix notation and now this is again a dash a system of linear equation there was a lot of space to fill and one dash good. So, system of linear equation and again you can solve them using.

Student: Gaussian elimination

Gaussian elimination, what is the complexity of Gaussian elimination. Let us see options right n n square n cube [FL], n cube right the Gaussian elimination the complexity is O N cube right. And I am not doing all this just to the sake of time pass I have a point of make which I will make on the next class right. So now, this was for any basis; that means, if you have any n linear independent vectors.

Now, I will consider a special basis where instead of n linearly independent vectors. In addition, these vectors are also orthogonal ok. Orthogonal vectors are linearly independent ok. So, a set of orthogonal vectors are linearly independent, but the converse is not all this right ok.

(Refer Slide Time: 05:12)



So now, let us see what if we have an orthonormal basis; that means, a basis consisting of orthonormal vectors. So, orthonormal is combination of 2 words ortho means, the 2 vectors are orthogonal and normal means all the vectors are unit vectors; that means, I am normalized them by their magnitude fine.

So, what is the condition that holds ui transpose uj is equal to 0? If I is not equal to j and ui transpose ui is equal to 1. Ok now, what happens in this special case, so we have this again we can express any vectors that as a linear combination of these. Ok now, let me try to do this I am just pre multiplying this equation by u 1 transpose. What happens on the right hand side everything disappears all of the this terms will disappear because, they are of the form ui transpose uj where is not equal to j and the first term is.

Student: One.

One, so what remains alpha 1. So, you can directly find alpha 1 using a dot product of 2 vectors. What is the complexity of this operation n the is just n products ok? Now how many such alphas do we need to find.

Student: (Refer Time: 6:09).

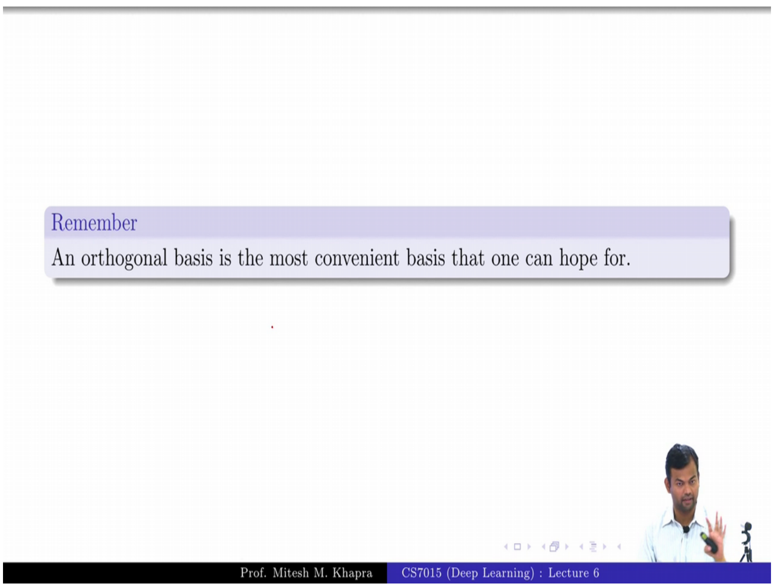
n of those. So, what is the complexity of d n square. So, that is now you see why an orthonormal basis is a very convenient basis, you can get all these coefficients just by doing a dot product between 2 vectors and later on I will show you that you might not need all of these. You might just need some subset k of these right so; that means, you just do k of these dot products and get these values. So, do you now understand the meaning of what is why, why do I say it is orthonormal basis is the most convenient basis that you can hope for right ok.

So, the another way of looking it right at it is again just to make few more comfortable with vectors and projections and so on right. So, this was your original point z 1 z which is a comma b right. And how do you actually draw that vector that, this is a and this is b ok. So, how do you find the coordinates actually you projects on to your basis vectors which were these x and y vectors. That is how you found the components along those the coefficient along those.

Now, instead of this x and y, I have any other set of vectors which is u 1 and u 2 and I will do the same thing I will project this on to u1 ok. I will project this on to u 2 and that projection will give me alpha 1 and alpha 2 right. So now, what is alpha 1 and that sense this is z, this is alpha 1 and this is theta right. So, alpha 1 is equal to z into cos theta ok. And what this cos theta? Ok. So, again you arrive at the same thing fine.

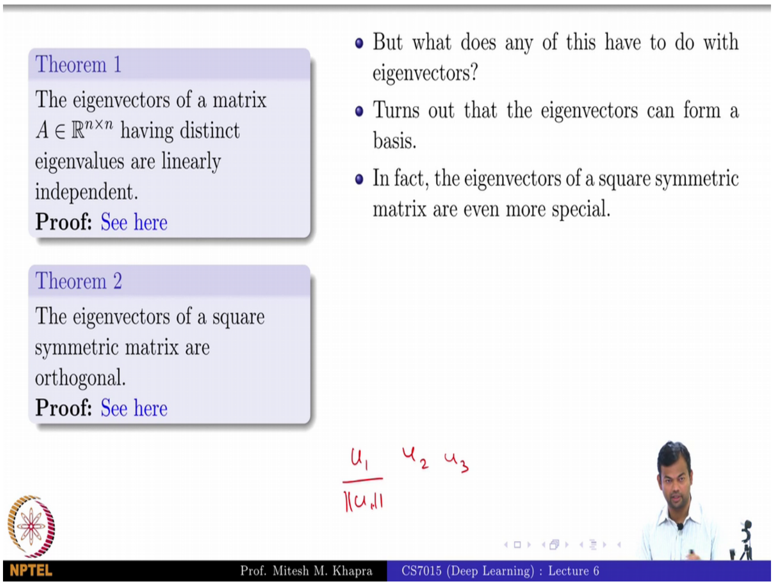
So, essentially taking a projection of a vector on to your basis is that. Fine to everyone, there is just to difference arriving at the same formula ok. That alpha is are given by a dot product between the basis vector and your original vector fine.

(Refer Slide Time: 08:08)



So, an orthogonal basis is the more convenient basis that you can hope for. That is the point, which I wanted to have you are convinced about that ok.

(Refer Slide Time: 08:16)



Now, but what does any of this have to do with eigenvectors. I started off with eigenvectors, I proved one property there and then I came to this linear algebra basic definitions and what a basis is set of linear independent vectors. And I eventually showed you that, an orthonormal basis is the most convenient basis that you can hope. So, what does any of this have to do with eigenvector.

Student: (Refer Time: 8:50).

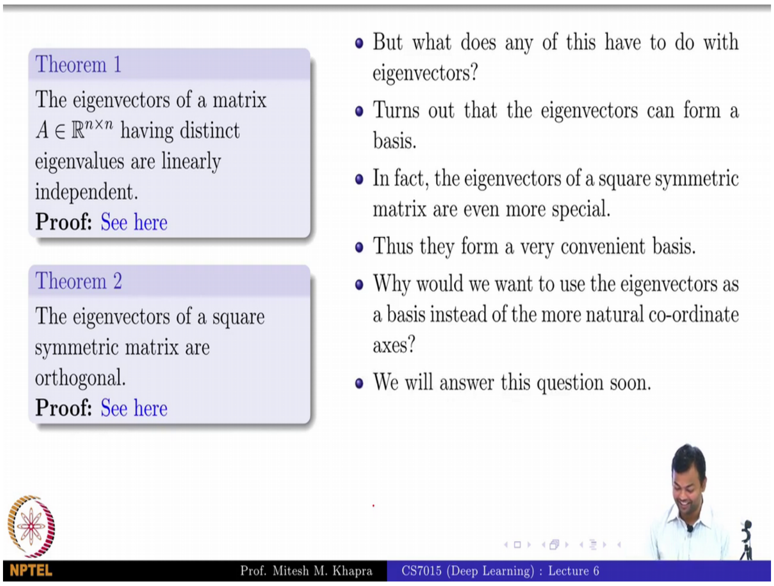
Always for us, square symmetric matrix right, why do you care about square symmetric matrix? Not sure yet. So, we get to that ok. So first of all, it is turns out the eigenvectors can form a basis and this is for any matrix. So, the eigenvectors of a matrix having distinct eigenvalues are linearly independent ok.

So, does every matrix, if I have an n cross n matrix will it have n eigenvectors? No, it can have less than or equal to eigenvector depending on the (Refer Time: 9:15). So, what is this saying is that if these eigenvectors are having distinct eigenvalues ok. Then these eigenvectors would be linearly independent, fine and turns out that for a square symmetric matrix that is the even more special, the eigenvector of a square symmetric matrix are.

Student: Orthogonal.

Orthogonal right and we already know that orthogonal is good right. So, remember, when we have orthogonal we do not really care about orthogonal. Because, that is it is a simple operation, if you have a set of vectors u 1 u 2 u 3 which are orthogonal, you can just divide them by the magnitudes and just get a set of orthonormal vectors right. So, orthogonal and orthonormal, I will use it interchangeably ok. And, whatever I done thus they form a very convenient basis. So, the eigenvectors of a square symmetric matrix form a very convenient basis.

(Refer Slide Time: 09:51)



So, that is how I connect the parts which was about the eigenvectors to the second part, which was about basis. And why would we want to do this and we already we had a coordinate axis that is the very good basis 1 0 0 0 0 1 0 0 1 and n dimension similarly.

So, why should I want to use the different basis? I have said that eigenvectors is a very convenient basis, but why do I care about it. I already have a very, very convenient basis which is just these 1 or 2 vectors are along these directions right. So, why do I care about a different basis? Ok, I understand that I that is there somewhere, but something more than that that is one advantage which I will talk about. What else more interesting? Ok, in what sense I love the power which comes with my job right. That you give a right answer and still I can embarrass you know. So, that is correct actually.