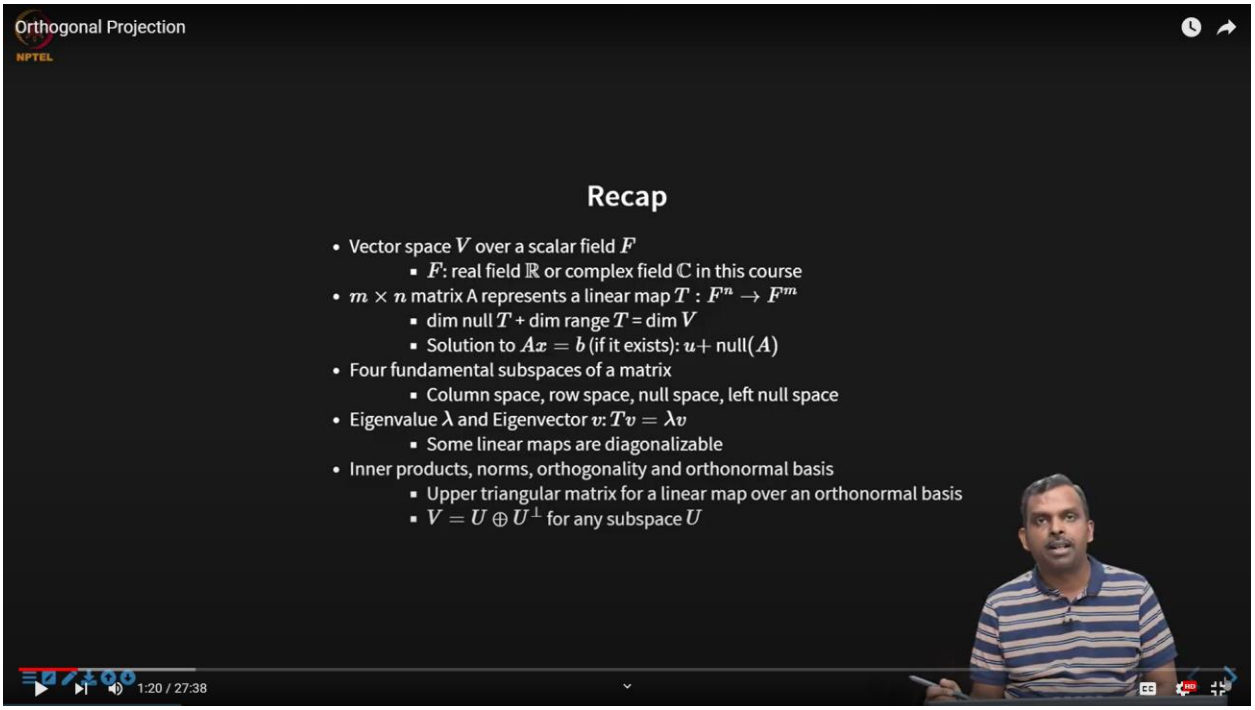
Applied Linear Algebra
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

Week 08
Orthogonal Projection

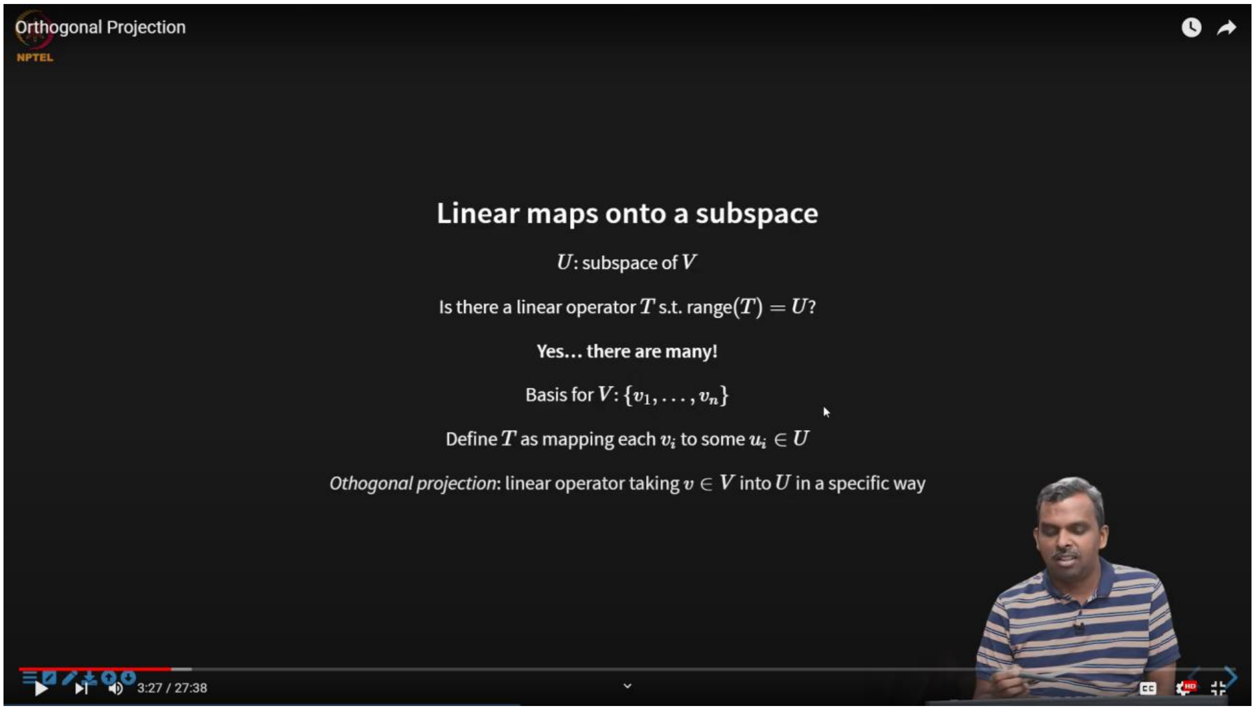
Hello and welcome to this lecture. We're going to talk about orthogonal projection in this short lecture. We will talk about how, you know, you can use an operator, a linear operator in a vector space to do something called orthogonal projection onto a subspace. In the subsequent lectures, we will see some very important applications of this. But the idea of this is actually quite simple. In small dimensions one can visualize this very easily. We already saw a little bit of it in a very simple case. As we go forward, we'll study more about this idea, okay? So let us get started.

(Refer Slide Time: 01:20)



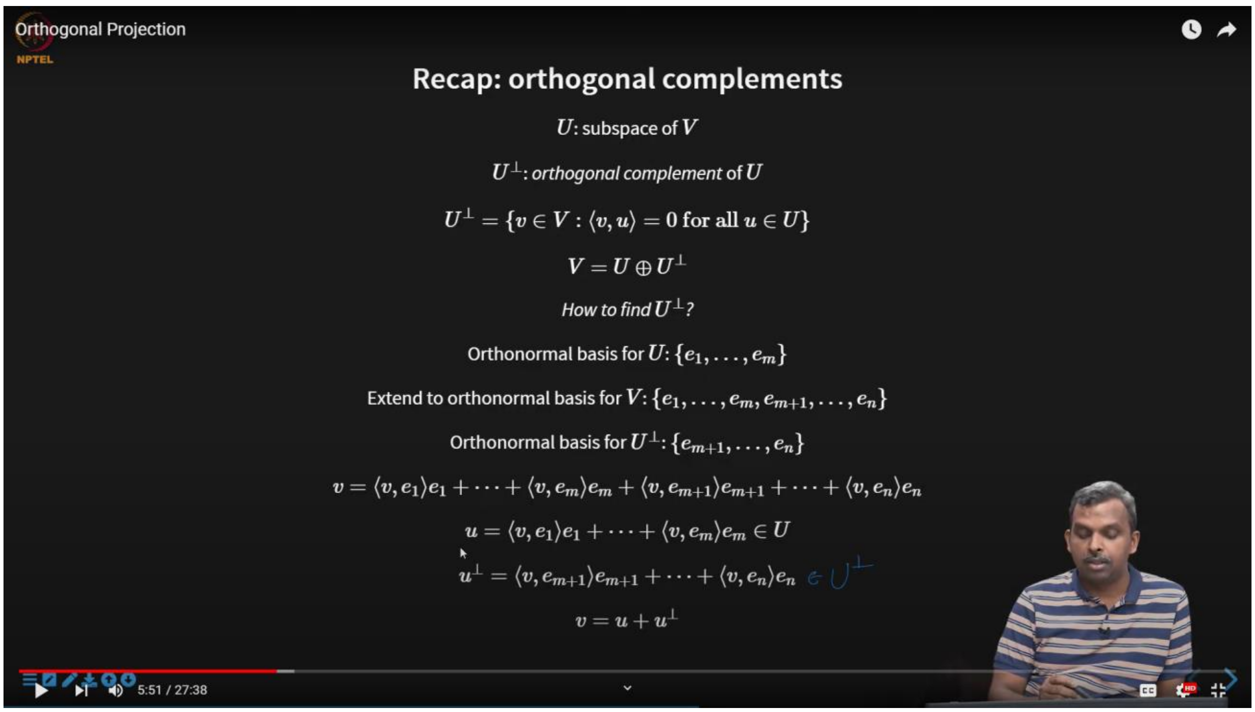
Okay. A quick recap. We have been studying vector spaces over real and complex fields. We know their association with matrices and linear maps and the four fundamental subspaces, solving linear equations, eigenvalues and eigenvectors, how they are defined and how the invariant subspaces simplify the matrix representation into a diagonalizable matrix in some cases. And then we studied inner products, norms, orthogonality, orthonormal basis. That's where we are right now. We looked at orthogonal complement and we'll put all that into some use in this, in defining orthogonal projection, okay? So that's where we are currently. Okay. So firstly let me begin with general linear maps. Let us say you have a subspace of V and you want to think of constructing an operator. Not just any operator in the vector space but an operator whose range will be, say, equal to U, okay? Or may be contained in U, okay? So I put equal to here. but range is contained in U or equal to U or something like that. So the range has got to do with the subspace that you have chosen, okay? So is that possible? Is that feasible? Or am I asking for something which is too much? Well, it turns out yes. In fact, there are quite a few, right? You can, you know, start with... So how do you specify a linear map T? Linear operator T? You have to look at the basis for V, any basis that you like. And for each v_i in the basis, I have to say what T maps it into. So as long as T maps each vector here into something in U, right, you will always have the range being inside U, okay? In fact you can make it equal to U by choosing, you know, enough independent vectors inside U in this for mapping, okay? So this seems like it is easy to do, right? So I can make a linear operator in this fashion very easily. So I can send a vector from my vector space into the subspace of choice, okay?

(Refer Slide Time: 03:28)



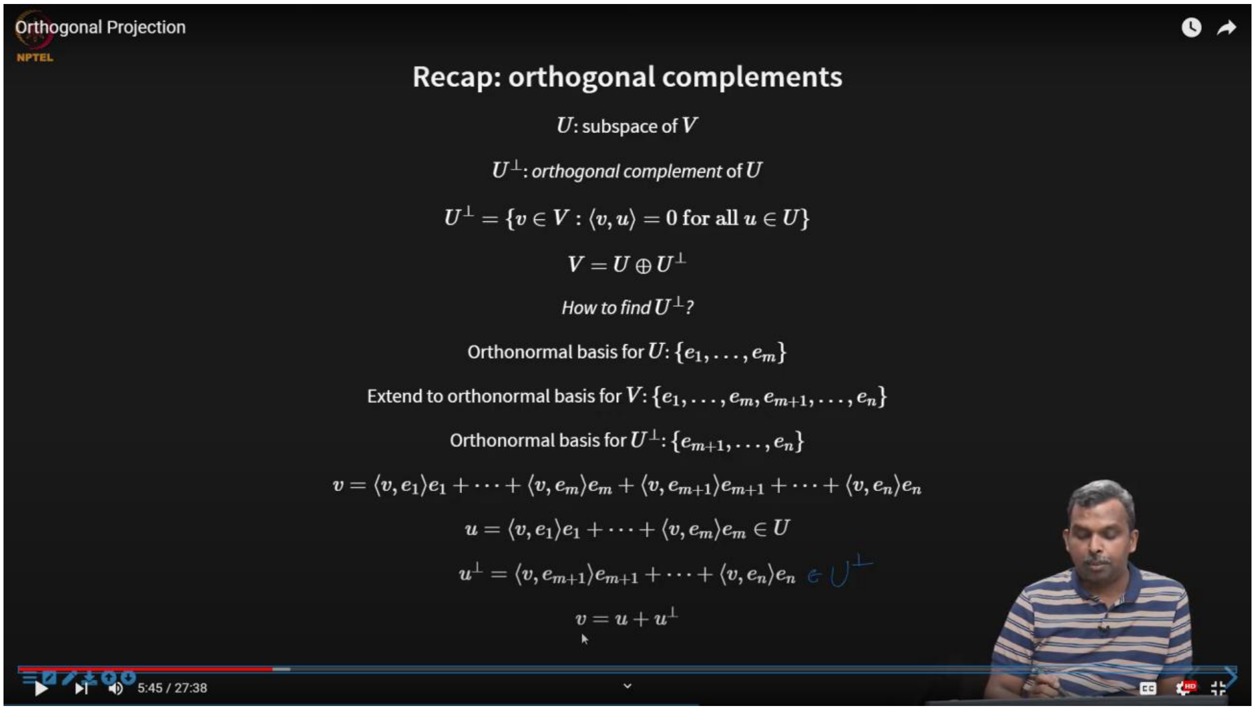
So out of all these kinds of choices, there is one special choice which is called orthogonal projection and it has a lot of powerful properties. And it's one specific linear operator which will take any vector into the subspace U or on to the subspace U and in a particular way. This orthogonal projection sort of way, I will describe what that is as we go along, but this is the context. And you can keep this in mind. Then also remember that there are many possible operators which can take you into a space. Out of them, this orthogonal projection is a little bit special. So we will see what is special about it as we go along. Okay. So a quick recap of orthogonal complements. We saw that in a previous lecture, supposing you have a subspace of a vector space. We define the orthogonal complement as the collection of all vectors that are orthogonal to every vector in U, right? So that is how we define the orthogonal complement. We saw this wonderful property that U and $U^{\perp}$, the orthogonal complement of U is denoted $U^{\perp}$, U and $U^{\perp}$ have a direct sum which is equal to the entire vector space V, okay? So we saw that. And how do you find $U^{\perp}$? I gave you this little method here. There are other methods also, but this method is: you start with an orthonormal basis for U, you extend it into an orthonormal basis for V and then $U^{\perp}$ is simply the extended vectors that you have, okay? So the orthonormality is important here. And that gives you all that you want for U. This is a very quick and easy way for doing, constructing $U^{\perp}$ and finding it out, okay?

(Refer Slide Time: 05:53)



And here is a little bit of a more of an explanation of what can happen here. So if you take an arbitrary vector v in the entire vector space, how do you find its coordinate representation in an orthonormal basis in this chosen orthonormal basis? You have this very simple formula, right? Because it is orthonormal, you simply have to take the inner product, okay? With respect to which it is orthonormal, okay? $< v, e_1 > e_1 +$ ... entirely, okay? So this gives you v, the entire vector v and you can quickly identify two vectors, right? One in U and one in $U^{\perp}$, right? This is not mentioned here, but this belongs to $U^{\perp}$, okay? And the sum of these two guys $u + u^{\perp} = v$, okay? That just directly comes from the way this, you know, orthonormal basis property behaves. Because it is orthonormal, you have this very simple expression for v. And because it is such a simple expression, it sort of splits into two halves. u and $u^{\perp}$. So you are able to identify here from the orthonormal basis extension v as the sum of two vectors u and $u^{\perp}$ and they add up to v. And we also know that this is unique, right? This u is unique. Irrespective of whatever orthonormal basis you pick, this u ends up being unique. You can prove that as well, right? So that property we know, okay?

(Refer Slide Time: 5:49)

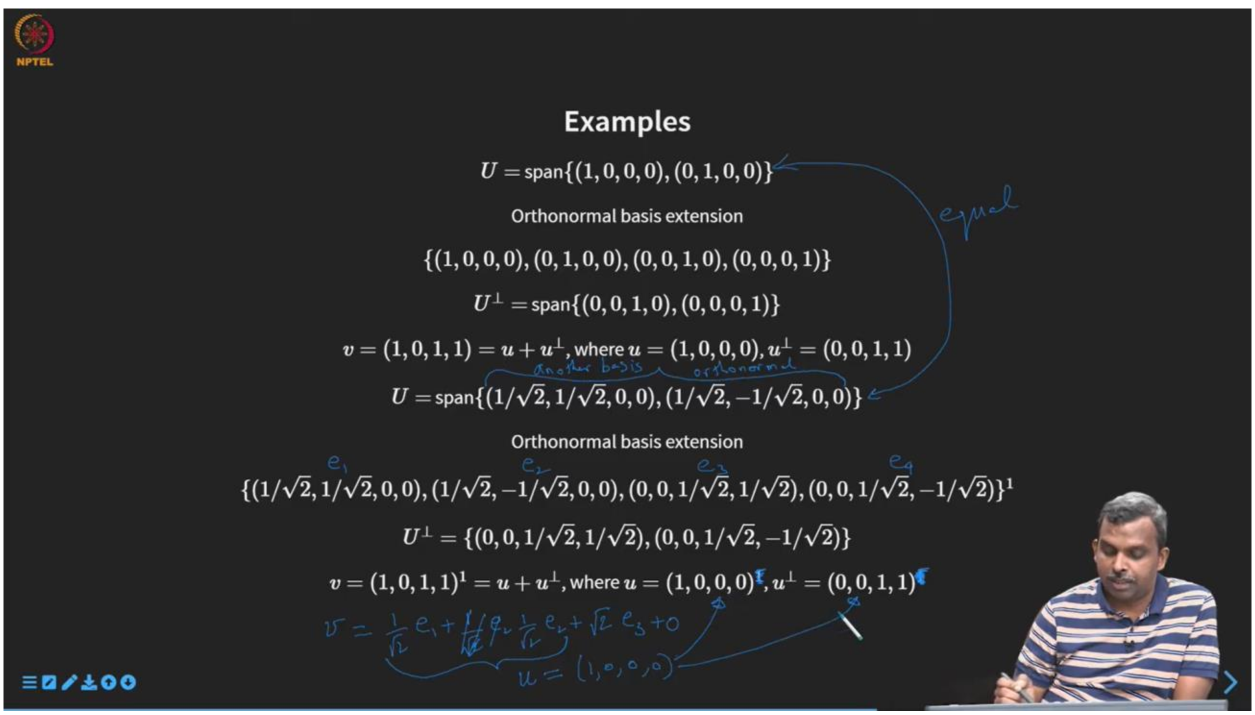


So let's see an example of this and sort of see how this works in practice. So here's a very simple example. In class I tend to do simple examples to give you an idea of how this whole thing works. Here is a U, right? I am doing this example in $\mathbb{R}^4$, right? You can evidently see that this example is in $\mathbb{R}^4$, okay? We'll start with a very simple subspace. It's spanned by these two basis vectors. $(1, 0, 0, 0)$ and $(0, 1, 0, 0)$. This is clearly a basis. I'll do an orthonormal basis extension. It's very easy to do, okay? It's the standard basis, right? So I extend it in this fashion. So I quickly identify $U^{\perp}$ as the span of the remaining two guys. That's easy. And any v, I have taken an example v here, $(1, 0, 1, 1)$, and it can easily quite easily be seen as $u + u^{\perp}$. And that is just $(1, 0, 0, 0) + (0, 0, 1, 1)$. You can see that this one belongs to $U^{\perp}$. This one belongs to U. And, I mean, if you want, you can write it down in the clear way as before. The inner product, right? So how do you identify the coefficient? You take the inner product of these guys. So you see this is 1, this is all 0, okay? So you get $(1, 0, 0, 0)$ as the u part. For the $u^{\perp}$ part, you see that the first one is 0, second one is 0, third and fourth, it's quite trivial to see that that formula works, okay?

So I am going to do the same thing but with a different basis. So you can see here these two are equal, they are equal. But the basis is different, okay? So different basis, okay? It's another basis.

It is also orthonormal, okay? So this is also orthonormal. You can check that it is an orthonormal basis, okay? But it is not the same as the standard basis, it is a different basis. So that hopefully gives you another example. I just want to show what happens if you take another orthonormal basis, right? So if you take this orthonormal basis, things worked out like this. What if you take another orthonormal basis and do something else with it? Here is an orthonormal basis extension, okay? So I've taken this guy and I've extended it to a basis for the entire vector space V. I am putting this one here simply to indicate that this is a different basis, okay? From the previous one, from the standard basis that we had. So $U^{\perp}$ now you can again identify is... So you can see $U^{\perp}$ is also the same as this, right? So it's the same thing. But the basis is different, okay? So different orthonormal basis. Now if I take v in this basis, right, in this new basis, I am taking a vector v. You can write it as $u + u^{\perp}$ and that would again be $(1, 0, 0, 0) + (0, 0, 1, 1)$. So if you want to see why that is, you can see, you know... So this v maybe I should write it down in more detail. So this v, so let us call it, I will call this e_1, this one e_2, this is e_3 and this is e_4 just for convenience. So what is going to be $< v, e_1 >$? So that is going to be $\left(\frac{1}{\sqrt{2}}\right) e_1 + \left(\frac{1}{\sqrt{2}}\right) e_2$, right? So I am taking the inner product here with this.

(Refer Slide Time: 12:57)



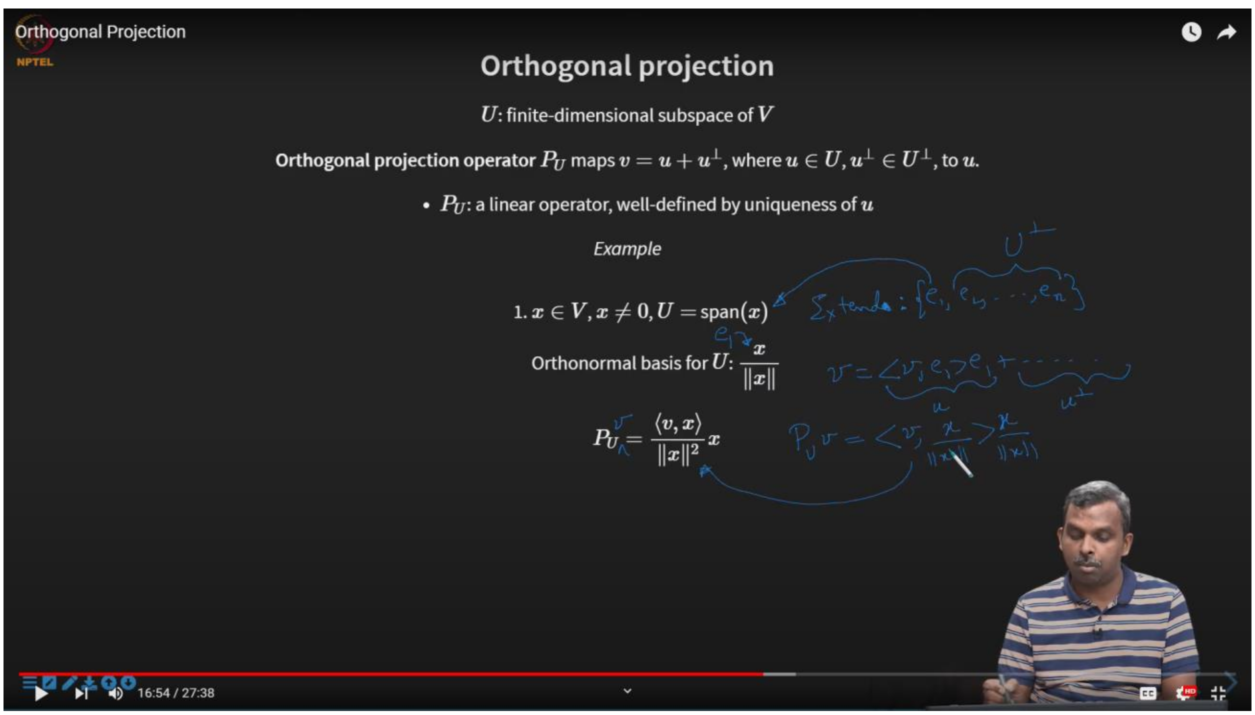
No, I think, did I make a mistake? Let me just strike it out. No, that is correct, I think that is correct. I do not know why I wrote that out. Plus these two would again work out to something, right? So if you do $< v, e_3 >$, that's going to be $\sqrt{2} e_3$ plus, if I do an inner product with this guy I am going to get zero, right? So this, the last one just works out as zero, okay? So you do not have an e_4 component in this, right? So this is how it works out, okay? And that is v in, written in this basis.

If you want to write it out like that, okay? And what is the u part? So these two together will make the u part of it, okay? So the u part is now $\left(\frac{1}{\sqrt{2}}\right)e_1 + \left(\frac{1}{\sqrt{2}}\right)e_2$. And that will work out to, you know, you can see that that will work out to $(1,0,0,0)$, right? That is not in this basis though. So one needs to be a bit more careful here. So it's just $\left(\frac{1}{\sqrt{2}}\right)e_1 + \left(\frac{1}{\sqrt{2}}\right)e_2$. So this is in the... I'm sorry for that, I think I messed this up a little bit. So hopefully, let me just clean this up here. So this 1 is wrong, this should not be there, it should also not be there, okay? So this is in the original basis. So even if I pick $(1,0,1,1)$ in the new basis, you see it decomposes into $u + u^{\perp}$ and it goes into the same u in the original basis, okay? So of course if you write u in the new basis, it will be something else. But it goes back to the same u in the original basis, so that tells you the uniqueness of u and $u^{\perp}$. Even if you change the basis with respect to which I'm doing the orthonormal basis calculation for figuring out how we decompose this into u and $u^{\perp}$, I will get the same vector u and $u^{\perp}$, okay? So that's clear enough to see from that point of view, okay?

So hopefully this was a clear enough example. So basically I am showing an example of a subspace U and two different basis choices for the orthonormal basis and two different orthonormal expansions. One ends up being the standard basis, the other ends up being some other basis. And then I try to write a vector, okay? Which is $(1,0,1,1)$ in the original basis. I am writing it in terms of u plus u perp and I get $(1,0,0,0)$ and $(0,0,1,1)$ and then I take a vector which is $(1,0,1,1)$ in the new basis, okay? And that will end up being something else in the original basis. But because I picked $(1,0,1,1)$ and then when I wrote $u + u^{\perp}$ and when I saw what happens here, the u ends up being the same, right? So you get the same thing there. It's just that it has to work out like that, okay? Yeah, so that works out correctly, okay? So u is $(1,0,1,1)$ in the old basis and $u^{\perp}$ is again $\sqrt{2}$ times this. That will work out as $(0,0,1,1)$ in the old basis, okay? So you pick $(1,0,1,1)$ but this $(1,0,1,1)$ is in the new basis. It works out in this fashion, okay? So hopefully this example was interesting enough for you to see how the same... You can pick whatever basis expansion you want and you would get the result that you want here, okay? So writing in terms of u and u perp is illustrated in this fashion, okay?

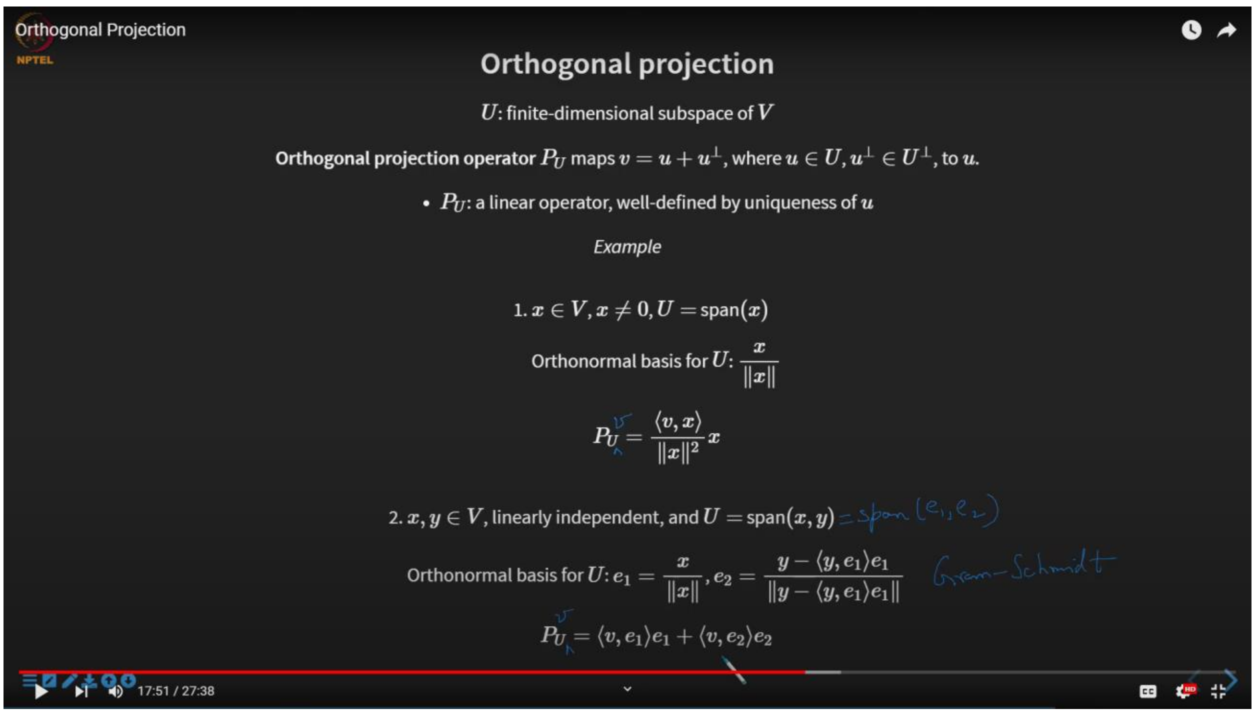
Okay. So let us move on to define orthogonal projection. So hopefully by now you are convinced that if you have a subspace U, finite dimensional subspace of a vector space V, you can always write it as $U + U^{\perp}$. And how do you do that? You pick your favorite orthogonal bases for U, extend it to an orthonormal basis for V and then pick up only the U part of it, you get your U, okay? So that seems like a clear enough way. And it’s unique as per the definition here, okay? So the orthogonal projection operator P_U basically works in this fashion, okay? So it maps a vector v to u, okay? And what is this u? u is obtained by the decomposition of v into $u + u^{\perp}$. And where u comes from the vector space U and $u^{\perp}$ comes from the orthogonal complement $U^{\perp}$, okay? So clearly this is a linear operator. You do very clearly linear operations. And it’s well defined because u is unique. Whatever, I mean just because you change something, u is not going to change. u is going to be the same thing, okay? All right. So here's an example, okay? So let's take a simple definition for U. U is a one dimensional subspace, okay? So how do you project, how do you do an orthogonal projection onto a one dimensional subspace U, okay? So let us take one non-zero vector and define the one dimensional subspace as a span of that non-zero vector x, okay? You see quickly that the orthonormal basis for u is simply $x/||x||$, right? So this is the definition for the orthonormal basis. And P_U will work out as, P_U acting on any v, sorry I think I am missing that here, P_U acting on any v will simply be $< v, x > \frac{x}{||x||^2}$. So how did I get this? You can see that. So any v can be written as, so you, how would you do this? So let us say we call this e_1, okay? And you extend e_1. So you extend to get a basis $\{e_1, e_2, \ldots, e_n\}$ right? So this guy gives you U, all of these give you $U^{\perp}$, okay? So when you write v as $< v, e_1 > e_1 + \cdots$ this is the projection u, this part is the $u^{\perp}$, okay? So when you project, what do you do? You write v as something that belongs to U plus something that belongs to $U^{\perp}$, okay? All that belongs to $U^{\perp}$. You can throw away, only this you keep, okay? And that you can see is the same as this, right? So what is this? This will work out as, so $P_U(v)$ it'll work out as $< v, x/||x| >$ times e_1 again which is $x/||x||$. And you can see that is the same as this, right? So $\frac{1}{||x||}$ will come out. So you will get $< v, x > \frac{x}{||x||^2}$, okay? So this is how you do projection onto a one dimensional subspace, okay? So it just works exactly by that basis extension argument.

(Refer Slide Time: 16:56)



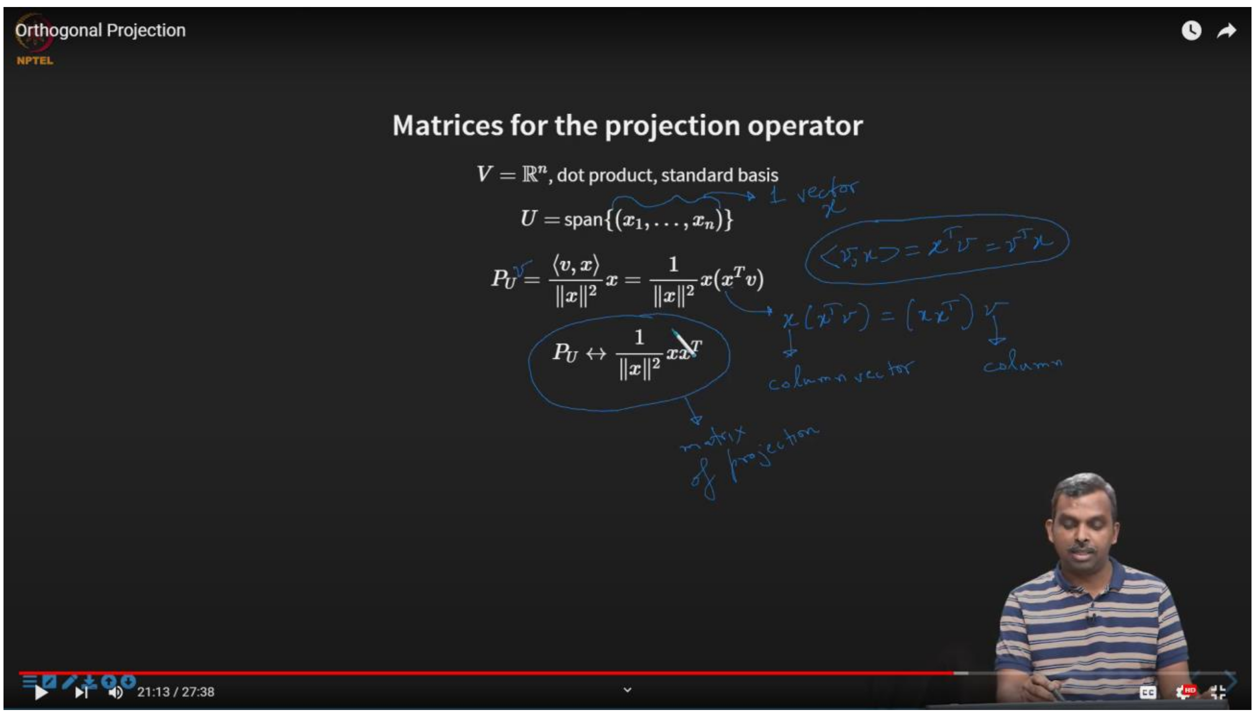
So you can also do, for a two dimensional subspace you take linearly independent x and y and then define u as the span(x, y). You do orthonormal basis. This is just Gram-Schmidt, right? So this is just the Gram-Schmidt followed on two, Gram-Schmidt done on two vectors. You know how to do this. So $P_U(v)$. I am missing the v again, apologies for that, okay? Okay? So $P_U(v)$ is simply $< v, e_1 > e_1 + < v, e_2 > e_2$, right? So that's it. So in this case, this is the same as the span of $\{e_1, e_2\}$ and simply the projection on v is simply $< v, e_1 > e_1 + < v, e_2 > e_2$, okay? So it is easy enough to see. So you find an orthonormal basis, any orthonormal basis that you want for your subspace U and simply do this formula v comma each of these things times the basis. That gives you the orthogonal projection of a vector v onto the subspace U, okay? And it does not matter what subspace you pick, what basis you pick here. Any orthonormal basis you pick, you will get the same projection operator P_U, okay? So that is something that we saw. So let us look at matrices for the projection operator, okay? So we see that projection is a linear operation and we have an operator for it. We have sort of a description for it in terms of inner product, that's not too bad, it's pretty good.

(Refer Slide Time: 17:51)



But let's see if we can get a matrix, right? So for that I'll specialize to this simple example of V being $\mathbb{R}^n$. I look at dot product as the inner product and the standard basis as the basis of choice, okay? So let's do all this. And here is a U, I have taken the vector space U to be a one dimensional subspace, okay? So this whole thing is a vector. Do not think of this as n vectors, this whole thing is one vector, okay? Just one vector x, okay? So think of it as one vector x, okay? And I know my projection operator acting on v, once again sorry about that... I am missing this $P_U(v)$ here. Is simply $< v, x > \frac{x}{||x||^2}$, okay? So that would be, I can rewrite this, right? See, notice this. See in $\mathbb{R}^n$ for instance, $< v, x >$ you can write in matrix form. If you have the, you know, coordinate expansions, you can write it as $x^T v$. You can also write it as $v^T x$. Both of these are the same, okay? So this is something that I didn't maybe emphasize too much when we discuss inner product. In $\mathbb{R}^n$ where the inner product is simply, you know, the dot product, $< v, x >$ inner product is the same as $x^T v$. You can also write it as $v^T x$, both of these are the same, okay? Important to understand this. So once I know this, this inner product $< v, x >$ I can write it as $x^T v$, okay? So there is a reason why I am writing it as $x^T v$. If you write it as $v^T x$, x it will not work out quite so nicely. So this is $x(x^T v)$. I can write this as, right, so this is all scalars now, so I can push it to this side if I want. So I simply get that $x(x^T v)$ is the same as $(xx^T)v$, okay? See, remember, x is a column vector. I am thinking of x as a column vector. So x^T will become a row vector and v also is a column vector, okay? So (xx^T) will be a matrix, $n \times n$ matrix. And then v will multiply on the right, okay? So this P_U which is the operator can be identified now with $\frac{1}{\|x\|^2}$ (xx^T) isn't it? So this is the crucial thing. This is a matrix of a projection operator, okay? So it will look like this, okay? You can see it is quite simple. It's (xx^T), it is a rank one $n \times n$ matrix, okay? Then multiplied by $\frac{1}{||x||^2}$. So this becomes the projection operator.

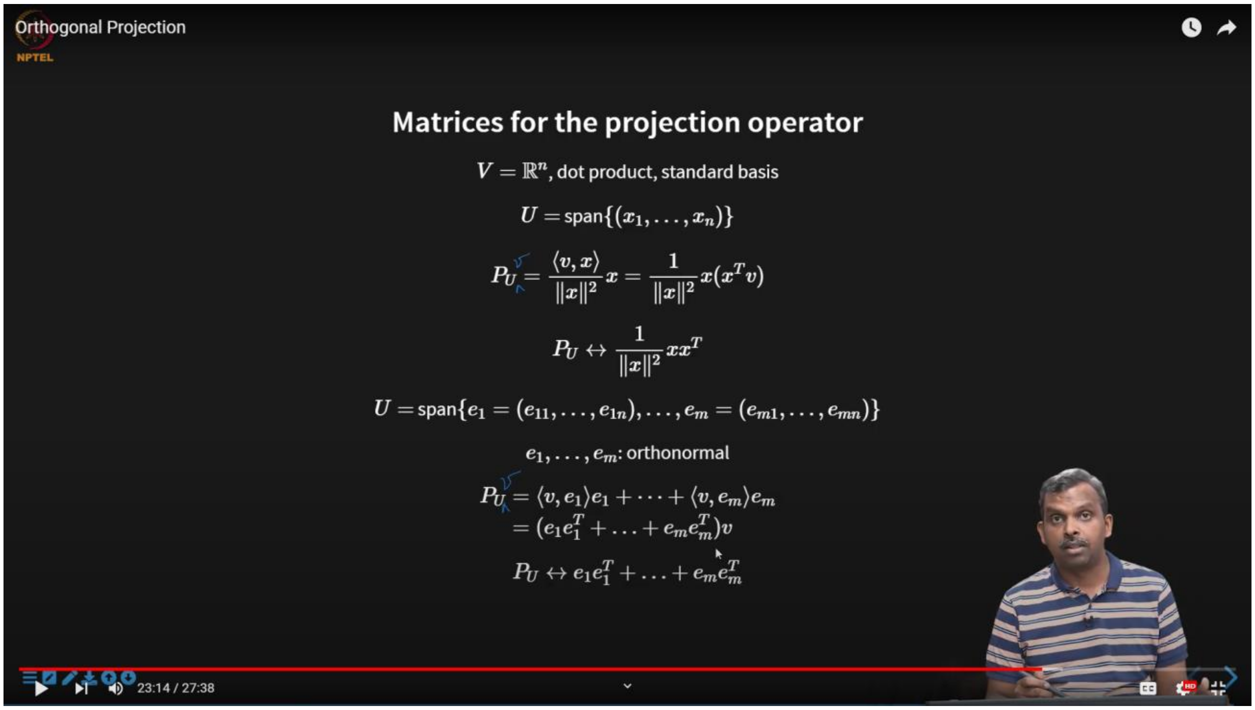
(Refer Slide Time: 21:15)



So you can generalize this. It's not very difficult to generalize this. So you find an orthonormal basis for your subspace U onto which you want to project, okay? And then if you want to do $P_U(v)$, I am getting this in every single slide here, if you want to do $P_U(v)$, should do it here also, okay, that will be simply $< v, e_1 > e_1$, so on, okay? So this we know. You can pick any orthonormal basis that you like for U and $P_U(v)$ simply becomes this. So now every $< v, e_1 >$ I can write as $e_1^T v$, okay? And then if you do that, you can pull the v common outside and you will simply get

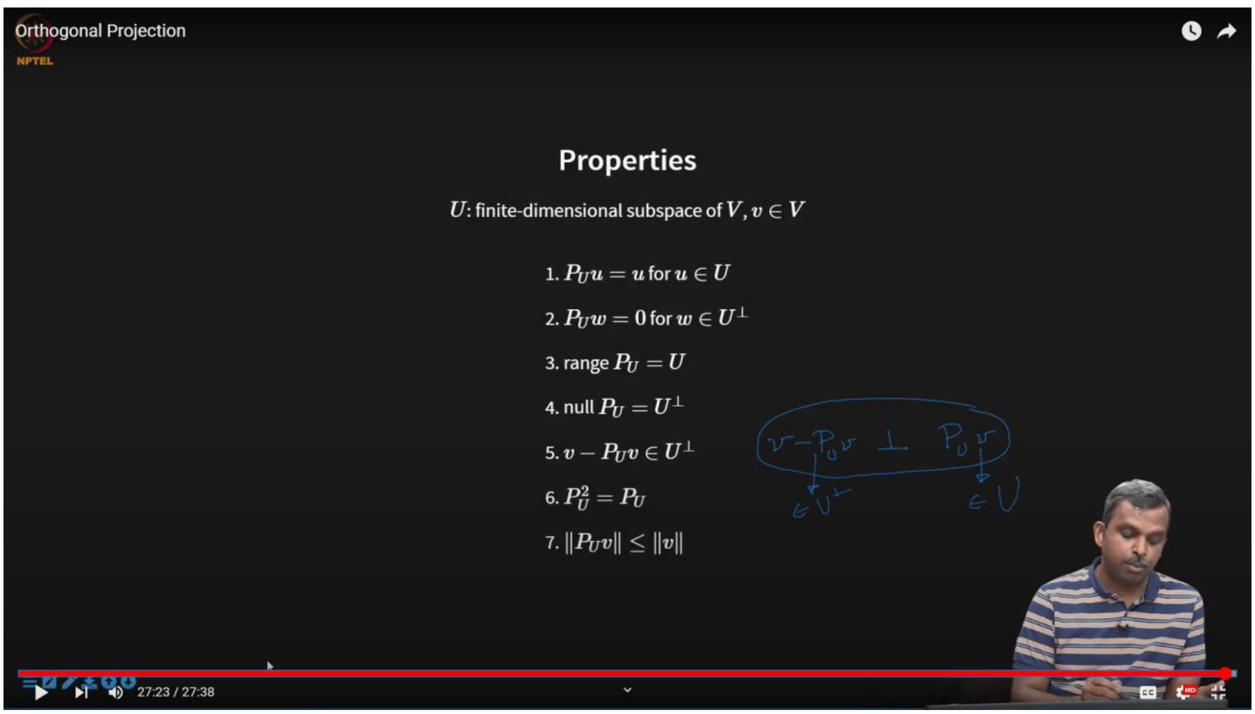
$(e_1 e_1^T + \cdots + e_m e_m^T)v$, okay? And you identify this P_U with this matrix that you got. $(e_1 e_1^T + \cdots + e_m e_m^T)$, okay? So this is your $n \times n$ matrix which represents projection onto the subspace U which is spanned by an orthonormal basis e_1 through e_m. So you see how this is coming through. So this probably gives you a good idea. You don't need the $\frac{1}{||x||^2}$ because, you know, all these e_is are already orthonormal, so they don't enter the picture. So it just becomes something like... So projection onto a subspace, when you have an orthonormal basis, finding the matrix is very easy, okay? You can repeat it quite easily for the complex case also, right? So instead of, you know, being $x^T v$, it will be, you know, $\bar{x}^T v$ or something like that. And then that would, so instead of $e_1 e_1^T$ it will be, you know. $\bar{e_1}^T e_1^T$ or something like that. So you should take the conjugate in one of these things and you will get what you want. Is that okay? So finding the matrix for the projection operator is also very easy, okay? So you have to get familiar with projection. It seems like a very easy thing to think of. It's not very complicated. But we will use some of these properties in some non-trivial way going forward, okay?

(Refer Slide Time: 23:14)



So here are some properties, quick properties one can very easily derive from the definitions I made so far. Supposing you start with a finite dimensional subspace of V and v is some vector in the vector space, okay? Here are a few properties, quick properties you can easily write down. If the projection operator operates on a vector that's already inside U, you will not get anything different. In fact you will get exactly u. So inside the subspace U, the projection operator is identity, okay? So that's nice to know. Outside what does it do? It makes it zero, okay? As an outside, meaning in the complement it makes it zero. And there are also many various vectors which are outside U and which are not in the complement also, right? So you will have $u + u^{\perp}$, then something else interesting would happen, okay? So, but this is the story as far as how the projection operator, if you restrict it to either U or $U^{\perp}$, how does it behave, okay? On u it becomes identity, on $u^{\perp}$ it becomes zero, okay? So that's, it sort of correlates with this description I have for the matrix, okay? If you think about what this matrix is, you will see that correlates very well, okay? Range of P_U is U itself. Null space of P_U is $U^{\perp}$, okay? These are all easy consequences. And if you do $v - P_U(v)$, you get something which is in $U^{\perp}$, right? So that is also very clear from the way we defined how orthogonal projection works. If you do $v - P_U(v)$, that belongs to $U^{\perp}$, okay? And if you do P_U^2, you get P_U itself, okay? So whatever vector you have, you know, if you square it you get the same thing as P_U, okay? So just think about it. If you repeatedly apply projection, you apply projection once, you got to some point that's already inside U. If you apply projection again, it's not going to change, it remains the same. So it's enough if you apply it once, right? So it's sort of, P_U^2 is a strange matrix or strange operator. You can't repeat it. The second time you repeat, it doesn't do anything new. It just gives you the same thing, okay? And here's an interesting result that norm of $P_U(v)$, the norm of the projection is lesser than the original vector. So you lose some magnitude when you project to a smaller subspace.

(Refer Slide Time: 27:23)



And there are proofs for all of this. And this one is a direct consequence of the Pythagorean identity, right? So v splits into its projection on u plus something which is orthogonal to that, okay? So when you have two orthogonal things adding to give you a vector, the norm is the sum of those two. So since you are losing something definite, you will get the norm for the projection being lesser than this, okay? So these are all properties. So projection is an interesting operation that you do on a vector. You can project onto a subspace of your choice. Orthonormal basis plays a big role in the projection and projection is a very interesting operator in the sense you do it, once you go to the subspace and after that it doesn't much happen to it, nothing much happens to it after that. And you lose something when you project, right? So the difference between the original vector and the projection is orthogonal to the projection, projected guy, okay? So this is something interesting. So in fact you can rephrase this a little bit also. See, $v - P_U(v)$ is orthogonal to $P_U(v)$. Is that okay? So this is a good result to have. Because this belongs to $U^{\perp}$. This belongs to U, okay? So there's a part of v which sort of, you're splitting v into two. One which is in U, one which is in $U^{\perp}$. And the difference is orthogonal, okay? So this is something important. We will take advantage of this later, okay?

Projection is important because it solves a very important and interesting minimization problem involving subspaces and norms and distances and all that. So we will take this up in the next lecture, okay? Thank you.