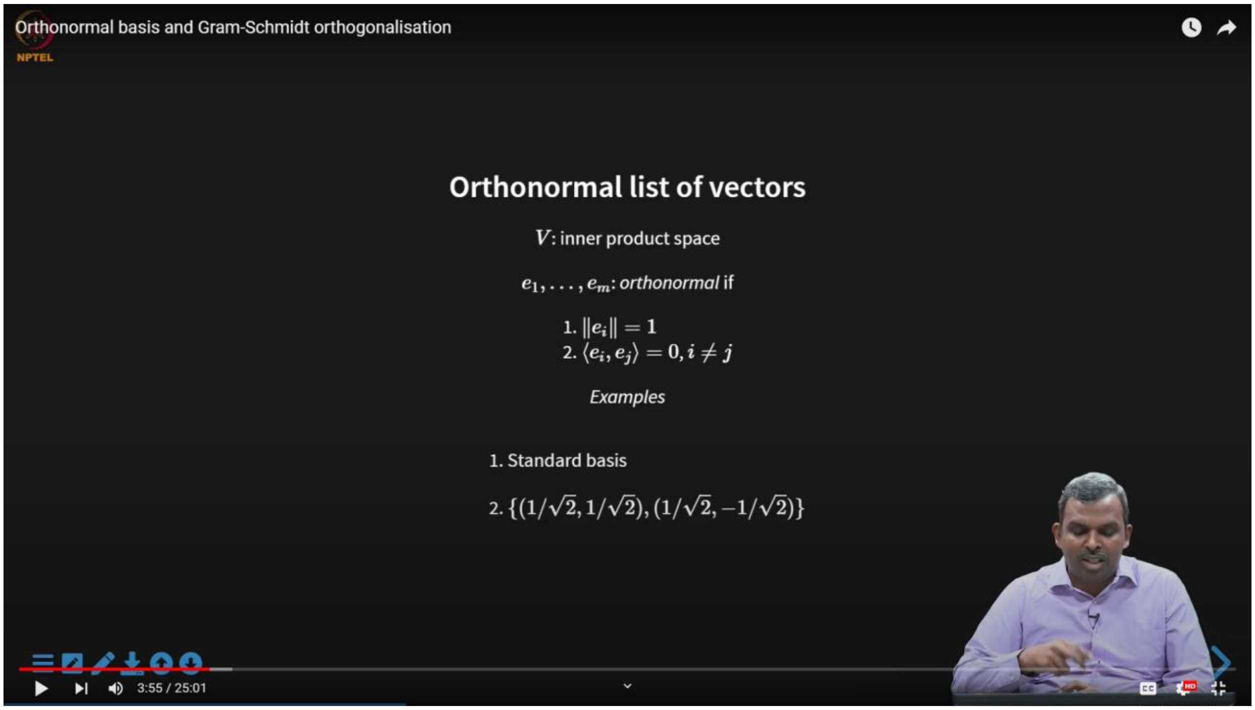
Applied Linear Algebra
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

Week 07
Orthonormal basis and Gram-Schmidt orthogonalisation

Hello and welcome to this lecture. In the previous lecture we introduced this notion of inner product and norm and saw how these natural notions of dot product and length you've been defining in euclidean spaces extends very naturally to abstract vector spaces as well. And we saw some basic properties of inner products and norms, how they are related etc. Now we will start using inner products and norms in the main problem we have, right? So how to understand operators. So we had vector spaces before. Now we have inner product spaces. How do you better understand operators in inner product spaces? Most of the applications we would deal with would be in an inner product space. So in an inner product space can we say more about operators? So if you remember before, we came up to eigenvalues, we saw operators had eigenvalues, eigenvectors and some of them could be diagonalized, all of them had an upper triangular matrix representation, all of that was true in any vector space, right? So now that you have inner product spaces, can we say something more special? Can we think of operators more clearly? Can we classify them? Can we understand a lot of these properties, okay? So that's what we're going to start doing. And the crucial starting point is this notion of an orthonormal basis and this notion of Gram-Schmidt orthogonalization. So these things help a lot, okay? So let's get started.

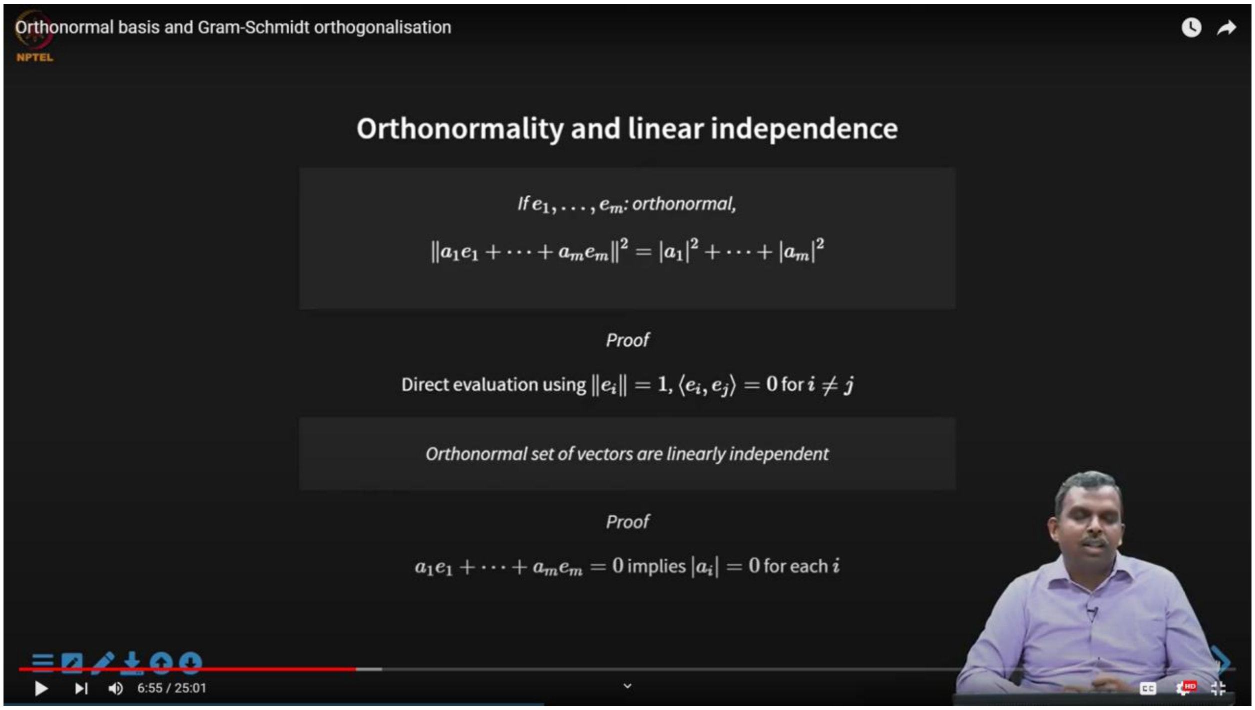
Okay. A quick recap. We've seen all of these before. What we saw in the very last lecture, the previous lecture was this notion of inner products and norms and orthogonality and related properties and very nice results based on inner products and norms. Particularly the Cauchy-Schwarz inequality. Very nice inequality which relates inner products and norms. Okay. So we'll make a very small definition which looks like a simple definition, but it will have very, a lot of interesting simplifications will happen in our understanding of operators when we do this. So throughout this lecture we will think of V as an inner product space. Maybe one or two times will not have that assumption, but mostly it will be an inner product space. I will call a list of vectors $\{e_1, e_2, \dots, e_m\}$ as orthonormal if the two conditions are satisfied. What are the two conditions? One is ortho, one is normal. Normal means the norm is 1, okay? So that's sort of like the way in which it's defined. And in this sense, orthonormal sort of implies norm is 1 and ortho means any two vectors, two distinct vectors have an inner product of 0, okay? So basically it just says $< e_i, e_j >= 1$ if $i = j$ and 0 of $i \neq j$, right? So that's what this also means, right? So that's the notion. Because we know that the, you know, the norm comes from the inner product, all of that is assumed here. Since that is true, this is the definition, okay? So this orthonormal list of vectors, do they exist? Can you come up with an orthonormal list of vectors? I mean naturally you might ask these questions. But what is the benefit of having an orthonormal list of vectors? What's the big deal with, you know, orthogonality, okay? So we will see they simplify a lot of things. We will start looking at initially some simplifications. As we go further, you will see the power of the orthogonality and simplifications entering the picture, okay?

(Refer Slide Time: 03:55)



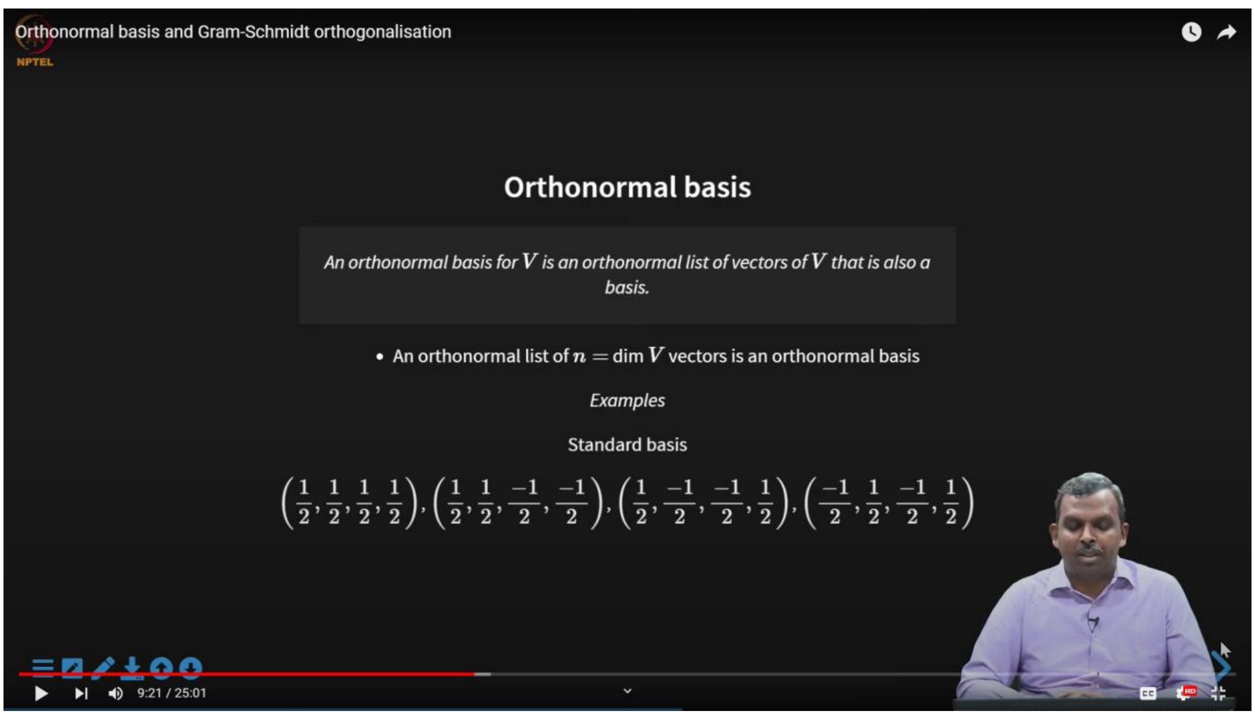
Here are a couple of basic examples just to motivate and convince you that orthonormal vectors do exist, okay? The standard basis is a standout example, okay? So an immediate example. You quickly see that orthonormality is satisfied for the standard basis. It's very easy to convince yourself that that's true. Here is another example which is maybe slightly more non-trivial, okay? So you see $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$. So this satisfies all the conditions that we have. Each vector is unit norm and, you know, I mean I'm thinking of the two norm here, okay? So in this thing, it's euclidean two norm, okay? So it's got unit norm and dot product is zero. So familiar dot product is what I'm using, okay? So this is orthonormality. Okay. So there is a connection between orthonormality and linear independence, okay? And also orthonormality simplifies things when we look at things, okay? So let's look at why that is so. Here is a very nice result, okay? If you have an orthonormal list of vectors and you do a linear combination of them, okay? I do a linear combination with the orthonormal set, okay? $a_1e_1 + \ldots + a_me_m$. So a_is are scalars, okay? So maybe I should mention that here. a_is are scalars, okay? So you do a linear combination with the orthonormal list of vectors and you look at the norm square after the linear combination, okay? So maybe you're thinking, you know, it's a linear combination, I have to use the properties of the inner product etc. But remember all of them are orthogonal. If you use those properties, and you start cancelling out etc. you will simply directly get sum of the absolute value square of the a_i, okay? So it's a very easy and simple result. So you see already that linear combinations are very, very simple to deal with when you think of orthonormal vectors and so that's at the heart of a lot of simplification in our understanding of linear operators, okay? So proof is very easy. I don't want to go into details. You just write it as an inner product of this vector, the linear combination vector with itself, then expand it out, a lot of cross terms will cancel, the same terms will simply give you value one and you get the answer, okay?

(Refer Slide Time: 06:58)



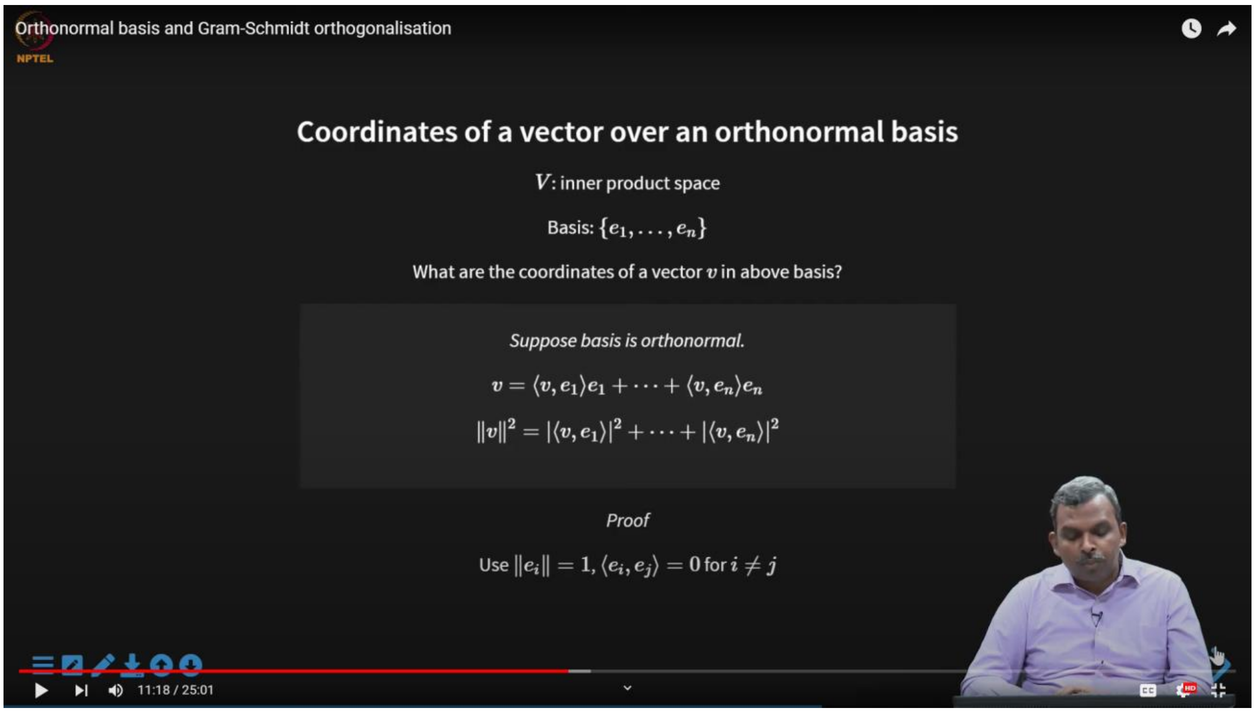
Here's another result, okay? An orthonormal list of vectors, orthonormal set of vectors, they are also linearly independent, okay? So it's quite easy to see. The proof can come directly from above, okay? If it turns out the linear combination is 0, then their norm is 0 which means each of the $|a_i|^2$ is equal to 0, each of the a_i is 0, okay? So that shows that a set of orthogonal vectors or orthonormal vectors, they are also linearly independent, okay? So you immediately see the connection there. It is not very difficult to imagine. You can have a linearly dependent set which is not orthonormal, that's possible, okay? So it's not, if and only for anything. There are very many linearly independent sets which are not orthonormal, but orthonormal sets are linearly independent, okay? So this immediately gives you a lot of restriction. So if you have a finite dimensional vector space, how many orthonormal vectors can you have? As many as the number of basis vectors, right? The dimension limits the number of orthonormal vectors that you can have as well, okay? So these are all nice results that come about once you identify the connection between linear independence and orthonormal set, okay? So you see that, you know, orthogonality is linear independence and more, okay? So it gives you linear independence. But also means the dot product is zero, right? So that gives you more, a stronger characterization of basis vectors, okay? Right? So that's interesting to have, okay? So this leads to the definition of an orthonormal basis, okay? An orthonormal basis for an inner product space V is simply what the name says. It should be a basis and it should also be an orthonormal set, okay? So now that we know that an orthonormal set is also linearly independent, it is very obvious that if you have a finite dimensional vector space V and the dimension is n and if you produce n orthonormal vectors, right, then that forms a basis. So we see that this orthonormal basis is going to be something that's really very interesting, right? So when you have an inner product space, this is what distinguishes the inner product space from a space without an inner product. You can have an orthonormal basis, okay? You can have the possibility of an orthonormal basis. So far we haven't shown that there exists an orthonormal basis. We are only shown one can even at least conceive of an orthonormal basis, okay? And we see that orthonormality can give you a lot of simplifications, particularly expressing vectors in terms of an orthonormal basis will end up being very nice and simple and elegant. And they are connected to the inner product and all that. We will see that soon enough in the next few slides.

(Refer Slide Time: 09:24)



So are there orthonormal bases? Yes definitely, right? So you can immediately come up with an example. The standard basis is an orthonormal basis, right? In any vector space that we take we have taken so far, $\mathbb{R}^n$, $\mathbb{C}^n$, you take the standard basis. It's an orthonormal basis. So orthonormal bases do exist. They are there. But are they there all the time in an inner product space? We don't know that yet. But they are there at least. There are at least one or two examples, okay? I'll give you one more example. Maybe it's not immediately obvious, but this is an $\mathbb{R}^4$ and you can see this is an orthonormal basis, right? So you can check for yourself that any two vectors are orthonormal here. And, you know, you can do the calculations if you like. And at the same time they are all unit norm. This is $\mathbb{R}^4$, right? And I am doing the dot product and usual two norm, okay? So this is the, this is an example. So there are, looks like there are many orthonormal bases. We will see soon enough that there are indeed orthonormal bases. But orthonormal basis is, looks like it exists.

(Refer Slide Time: 11:18)

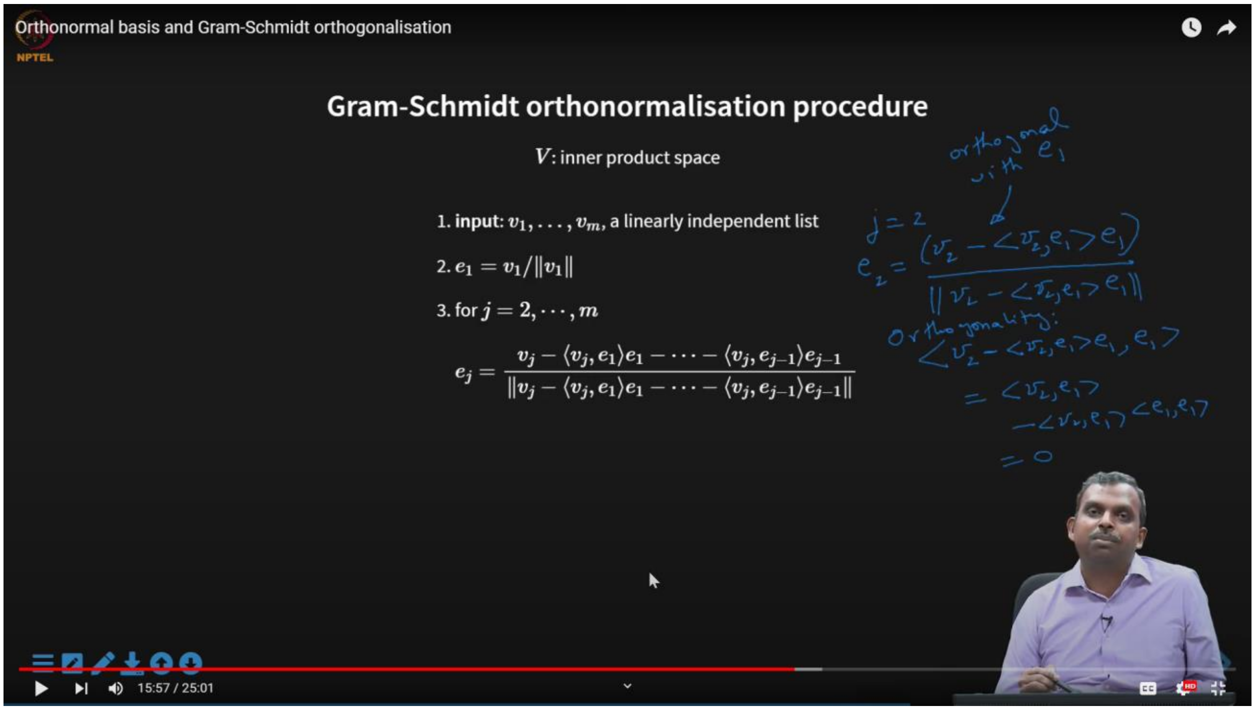


But what is the advantage? Is there, are there advantages? Are there simplifications in using an orthonormal basis? It turns out yes. Already if you just think in terms of coordinates, you'll see an advantage. And then when you think in terms of operators, you'll see there are more interesting advantages, okay? So first thing is, if you think of an inner product space and you have a basis $\{e_1, \ldots, e_n\}$, if you ask me what are the coordinates of a vector v in the above basis, I do not have an easy way to establish this, right? So you have a v and then I have to write down some equation, I have to see how to express this. Maybe a matrix equation, maybe some linear equation solving, all sorts of things are needed if I don't have orthogonality. If you have orthonormality, look at what happens. If this basis is orthonormal, all you have to do is evaluate inner products, okay? You take your vector v, you evaluate inner products with each of the basis vectors and do a linear combination with them, you will get v, okay? So it's very, it seems like a very intuitive and easy result. Maybe you learned this in so many ways in previous classes but one can prove this quite easily, okay? So v when expressed, okay, in the basis $\{e_1, \ldots, e_n\}$, okay? First of all it's a basis. So you know that v can be expressed as a linear combination of e_1 to e_n. And after that it's a very easy exercise to show $< v, e_1 >$ is simply the first coefficient. $< v, e_n >$ is the last coefficient. Not only that, we saw before that the $||v||^2$ is simply the norm square of, I mean mod square sum of each of these coefficients, okay? So all this is very nice. So expressing a vector in terms of an orthonormal basis is very, very trivial. You simply what's called project onto each of these, you know, basis vectors, take the inner product with each of these basis vectors and simply scale them and add them, you get your answer, okay? So it's an easy proof. I'll skip the details of that, okay? So this has become simplified.

But, you know, what else can happen with operators? That's also an important question, okay? So so far we have been sort of skirting the question of: do orthonormal bases exist, do they have, are there orthonormal bases for a finite dimensional vector space, can you find them, are there easy ways to find them. And that's what's given by this Gram-Schmidt orthonormalization procedure. It's a very simple procedure. It sort of converts a linearly independent set into an orthonormal set with the same span, same incremental span. So I will talk about how that, how it works and why that is so shortly. But take that as an important definition. So basically if you have a linearly independent set, this Gram-Schmidt process will output an orthonormal set but with spans being the same. So at every level spans will be the same. So it's a very powerful and very simple and elegant procedure, okay? So how does that go? The input like I said is a set of linearly independent vectors in your inner product space. So the first thing you do is you take the first vector and normalize it, okay? So this process of dividing a vector by its norm, it is called normalizing. So what happens after normalizing? The norm of e_1 will simply become 1, isn't it? So that's a nice thing to do. So you can normalize. Norm of e_1 becomes 1. So this is called normalization, okay? So notice this normalization. So dividing a vector by its norm makes the overall norm as 1, okay? So you can quickly prove it. It's not very hard, just use the homogeneity property and you will get just 1, okay? So this normalization will be one property. But normalization alone is not needed, right? So just because you do e_1, you cannot say e_2 is simply norm, you know, $v_2/||v_2||$. But then v_1 and v_2 will not be orthogonal. I also want the orthogonality, right? I want the output from this process to be orthogonal. So for that we will use some sort of an orthonormal decomposition idea and this is what will happen. So this is what the, this is what the process does, okay? So in general for a particular j, this is the step. But let me just show you what happens for $j = 2$, okay? So that will help you. Let's go to $j = 2$.

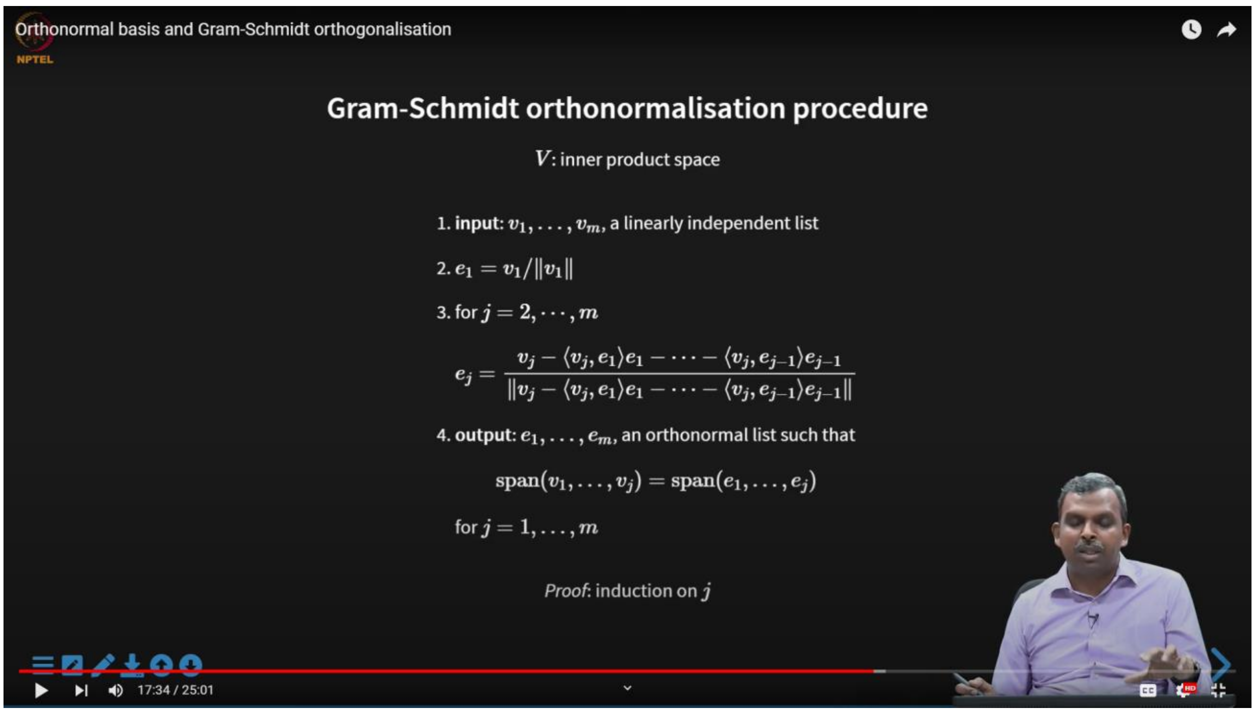
You will define e_2 as $(v_2 - < v_2, e_1 > e_1)$ divided by the norm of the whole thing, okay? Forget about the norm coming in the denominator, so that is just to normalize, right? So this $(v_2 - < v_2, e_1 > e_1)$, that is going to be... This guy, the guy in the numerator is orthogonal with e_1, okay? So that's the crucial part, okay? And that's very easy to show, right? So for instance you can show, so, you know, $(v_2 - < v_2, e_1 > e_1)$, right? What will this work out as? $< v_2, e_1 >$, okay, just use the additivity property, minus $< v_2, e_1 >< e_1, e_1 >$. And $< e_1, e_1 >$ is 1. So that ends up being true, okay? So how do you prove the orthogonality? Simply by evaluation. So what's being done here? It's not very hard to imagine, right? So you are taking this, the same orthogonal decomposition approach, okay? So you decompose v_2 into a part that lies along e_1 and then you subtract it out, you will get a part that is orthogonal to e_1, right? So that's the same orthogonal decomposition. I am just writing this out once again to show you that that's true. And then once I get the orthogonal thing, I simply divide by its norm to get e_2. So norm of e_2 will be 1. And e_2 and e_1 will be orthogonal, okay? So that is the condition here.

(Refer Slide Time: 16:01)



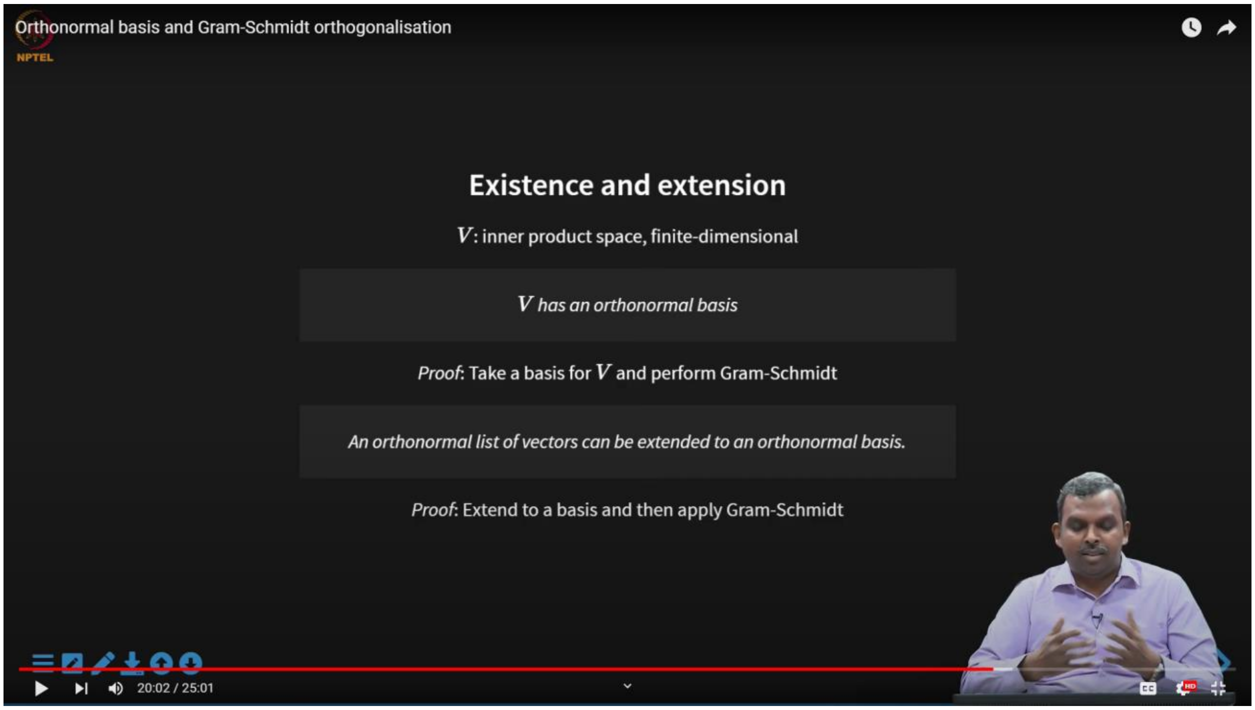
Now what do you do for arbitrary j? You take v_j and simply subtract $< v_j, e_1 > e_1$ minus $< v_j, e_2 > e_2$ so on till $< v_j, e_{j-1} > e_{j-1}$. And then divide by its norm. So this process you can see every step will keep giving you normalized vectors, norm 1. And at the same time all of them being orthogonal to each other at every step, they are orthogonal, orthogonal, okay? So that is the nice little result. I mean there is nothing more to prove here. It is a very simple process and you can take any set of vectors and try it out, it is very easy to prove this, okay? So the output e_1 through e_m is an orthonormal list. Not only that, okay, the orthonormal list part is easy enough to show, but notice what has happened here. This is very, very crucial. The $span\{v_1, \ldots, v_j\}$ for every j is the same as $span\{e_1, \ldots, e_j\}$, okay? This maybe takes a little bit of proving and you can prove it by induction, it is very easy, okay? So but the span is the same and it's not too difficult to imagine why. Because e_1 was proportional to v_1 and e_2 was v_2 minus something, right? So it was still, the span will remain the same, okay? So this linear independence and the span does not change, okay? So this is crucial. And for every j the span remains the same. Not just the overall span, okay? So it's almost like the Gaussian Elimination sort of process, right? So there are these processes which are very important. First was that row, elementary row operations which helped you do rank and all that. Next was this upper triangularization with eigenvalues which helped you, you know, do something. And here is an orthonormalization process. From linear independence, you can go to orthonormal case also, okay? So the proof you can do by induction on j, particularly the fourth part, the last part, the span being the same is easy to do with induction on j. I'm skipping the details. You can see in the book also. There are details of the proof. But you can see intuitively why this should be true, okay? So this process which is called the Gram-Schmidt orthonormalization procedure is very, very important, okay? So this shows you that there are orthonormal sets of vectors that you can easily generate. Given linearly independent sets, we know there are linearly independent sets. So take linearly independent sets. From there you can generate orthonormal sets, okay? So that's very nice.

(Refer Slide Time: 17:34)



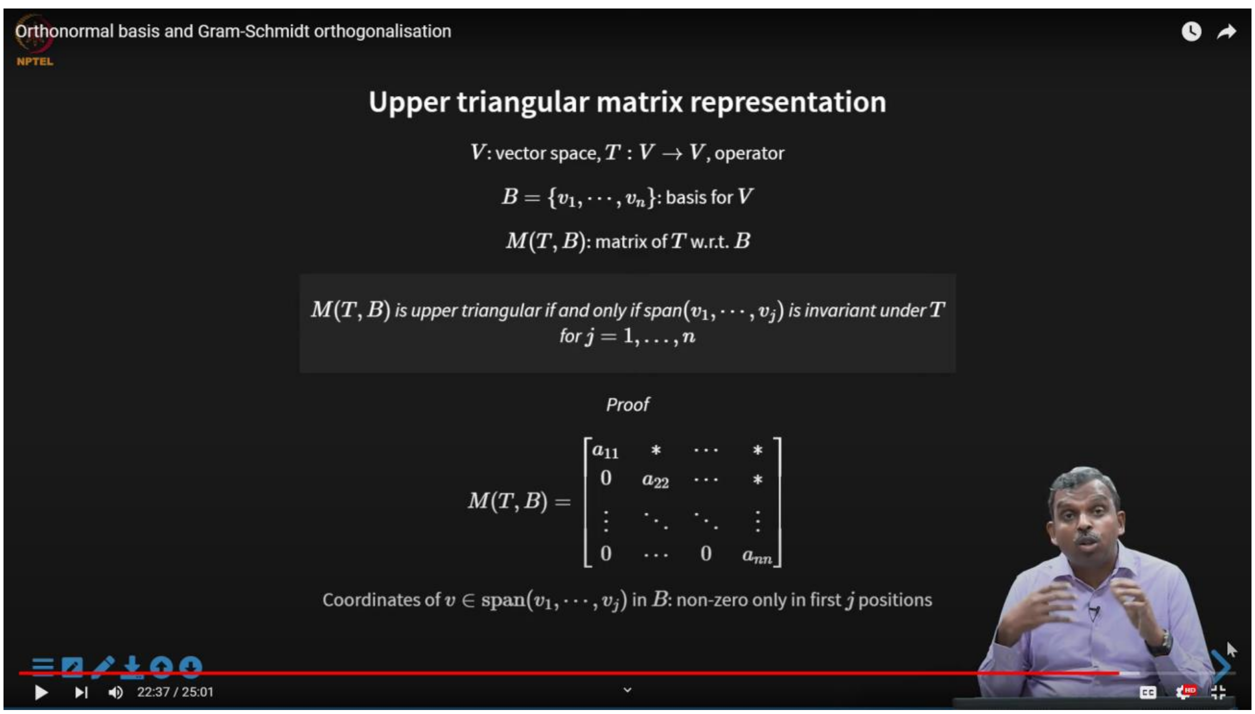
So this also gives you many more results, okay? So particularly about existence and extension possibilities, okay? So supposing you have an inner product space which is finite dimensional. Then that has an orthonormal basis. What do you do? How do you prove this? It's very easy. You simply take a basis for V, you know that there is a basis for V, and you perform Gram-Schmidt. If you do Gram-Schmidt, you will get an orthonormal basis, right? So you know that every inner product space now if it is finite dimensional it has an orthonormal basis. Next is: an orthonormal list of vectors can be extended to form an orthonormal basis. Supposing you have a large dimensional vector space and you have a few orthonormal vectors, a list of orthonormal vectors. You can extend them without affecting them, you can extend them to form an orthonormal basis overall, okay? So how do you do it? Again you do the same method, right? You take the original orthonormal list, you know it's a linearly independent list, you simply extend it to form a basis, a general basis without worrying about orthonormality. And then you apply Gram-Schmidt on this whole process, right? What will happen when you apply Gram-Schmidt? The initial set of orthonormal things will not change. The reason is Gram-Schmidt is going to keep doing these dot products and all of them will remain the same, okay? If you already have orthonormality, why will Gram-Schmidt do anything? It won't do anything, only afterwards it will start making it orthonormal, okay? So you will have an extension also possible. So this orthonormal basis in an inner product space is as powerful as a basis and as prevalent as a basis, okay? There are orthonormal bases and you can take small orthonormal sets and extend them to form a basis, okay? So in inner product spaces, you can happily have orthonormal bases. Of course you also have other bases which are not orthonormal, but orthonormal is particularly interesting. When you have orthonormal bases, so many computations become easy and you'll see later on also many more computations will become easy for you in an orthonormal basis. So as far as possible, you should try to work with an orthonormal basis. And only when orthonormality does not make sense you should move away from it, okay? So this is a very useful principle when you deal with vector spaces and operators, okay? So this is... And they do exist. Extension exists. Orthonormality also exists.

(Refer Slide Time: 20:04)



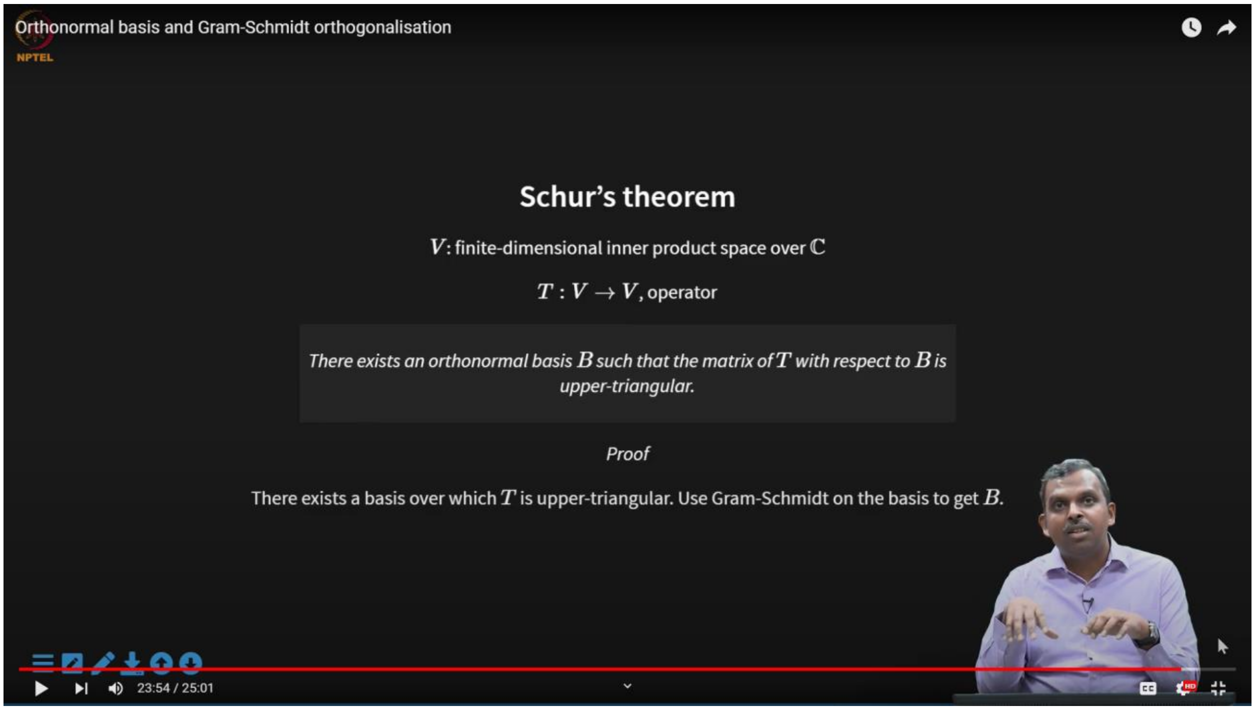
Okay. So let us come back to this upper triangular matrix representation and see a very interesting result about orthonormal bases in inner product spaces. Supposing you have an arbitrary vector space. At this point I am not saying inner product space. Vector space. And you think of an operator T, okay? And you have a basis, we always think of a matrix for that operator in that basis B, right? So we know how to do this. We know this very nice result about upper triangular matrix representations. But what do upper triangular matrix representations really mean? This matrix will be upper triangular if and only if the $span\{v_1, \dots, v_j\}$ for every j, the $span\{v_1, \dots, v_j\}$ should be invariant under T, okay? So this is an interesting way of characterizing the upper triangular matrix representation. So far we've just said upper triangular. But what does it really mean in terms of, you know, what the operator is? It turns out this is what's important. If, over a particular basis, the operator has an upper triangular matrix representation, this is an if and only if condition, okay, v_1 to v_j it has to be invariant under T, okay? So the proof is not very hard. I'll just write down the matrix for you and you'll see the proof will sort of follow. If your matrix is upper triangular like this, you can see if you take the first j coordinates and you do linear combinations with them, only the first j are going to be non-zero, right? So there is invariance sort of with respect to that. Bottom part is zero. And the other way also is true. If you have a basis in which the first j alone are going to be involved, if you express the operator in that basis, you will naturally get an upper triangular representation, okay? It's a very simple characterization of what it means to have an upper triangular matrix representation, okay? So this sounds very simple but it's very important.

(Refer Slide Time: 22:37)



So now you will see a lot of things will come together because the $span\{v_1, \dots, v_j\}$ has to be invariant under T. So this is the condition that's important, okay? We have seen before that, you know, every, if you have over the complex field, when you have eigenvalues, when you can have eigenvectors over the complex field, any operator has, over a vector space over complex numbers, has an upper triangular matrix representation, okay? So that we have seen before, right? So we saw one of the results. How to get an upper triangular matrix representation for a, for an operator in a complex vector space. What about orthonormal basis, okay? Can you have an upper triangular matrix representation with an orthonormal basis? That's a great question, right? And the answer is yes. And that's what's called Schur's theorem, okay? So it's a very popular idea, very important theorem. If you have a finite dimensional inner product space over $\mathbb{C}$, over the complexes, where you have eigenvalues and eigenvectors and all that and you have an operator T, it turns out there exists an orthonormal basis such that the matrix of T with respect to B is upper triangular, okay? So that's a powerful result. It's called Schur's theorem. And the proof is actually, we have already seen the proof, it is not very hard to see, okay? We know that there exists a basis in which T is upper triangular because it is over complexes, right? So eigenvalues, eigenvectors exist. So you can use that and make an upper triangular matrix. Now you take that basis in which its upper triangular and simply run Gram-Schmidt on it. When you run Gram-Schmidt on it, you get an orthonormal basis.

(Refer Slide Time: 25:34)



Not only that, the, you know, the $span\{e_1, \ldots, e_j\}$ is the span of, same as $span\{v_1, \ldots, v_j\}$, okay? So $span\{v_1, \ldots, v_j\}$ is invariant under T. $Span\{e_1, \ldots, e_j\}$ is also invariant under T. So the matrix with respect to E is also upper triangular, okay? And that's the end of story, okay? So look at this very nice result that we have now already. If you have an inner product space, we know that there are these orthonormal bases. And orthonormal basis simplifies the description of the coordinate system, simplifies operating with vectors etc. We will see more and more later on. But as it is, you can easily be convinced that that's true, okay? So we saw before that in a vector space where eigenvectors, eigenvalues exist, you can have an upper triangular matrix representation, vector space over complex numbers. Now because we have an inner product space, right, you can have an orthonormal basis in which the operator will have upper triangular matrix representation, okay? So that is very nice. It’s very convenient and a good thing to have, okay? So we can work with orthonormal basis in an inner product space and not really lose anything that is significant, okay? So that's a nice result to have. And I'll stop here for this lecture and pick up from here in the next one. Thank you very much.