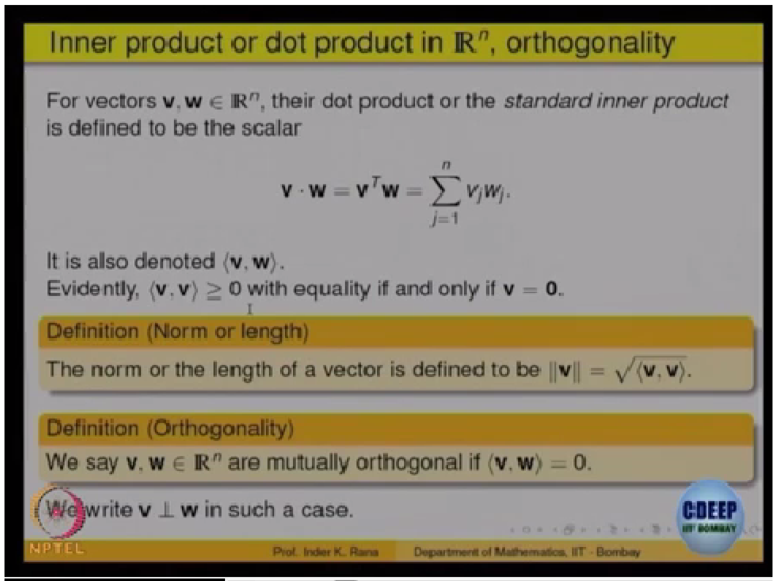
Basic Linear Algebra
Prof. Inder K. Rana
Department of Mathematics
Indian Institute of Technology - Bombay

Lecture – 27
Orthonormal Basis, Geometry in R^2 III

Right, so let us start doing it in Rn.

(Refer Slide Time: 00:27)



For Rn 2 vectors v and w, their dot product you can define it as, if you treat v as a column vector, with n*1, if you write it, transpose this 1*n multiplied by another, what your answer will be 1*1, the scalar. So the dot product of v and w is defined as v transpose w, the product of, as matrices. Or it is same as v1w1+v2w2+...vnwn, either way. So this is the matrix form. This is the algebraic form and this is the expanded form of the dot product, okay, in Rn.

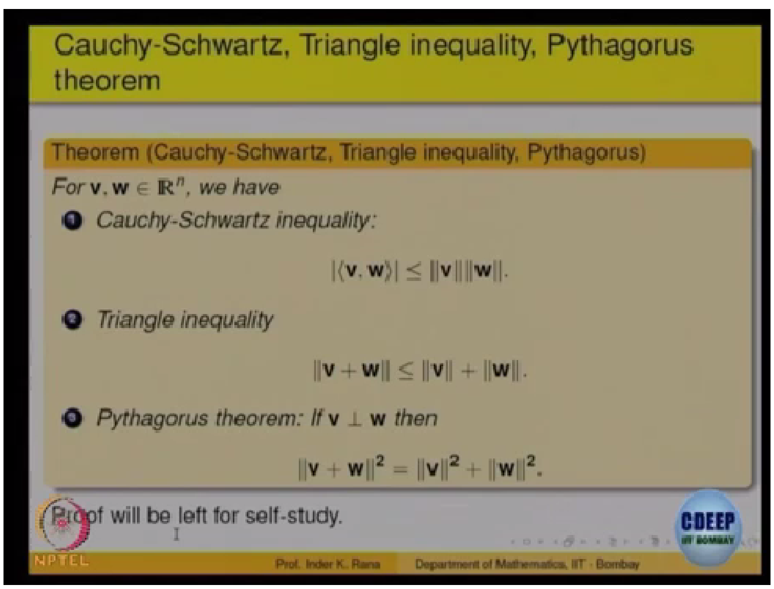
This dot product normally is also instead of, sometimes dot is not very visible, so one writes it as this bracket, angle bracket. So this is called the inner product between v and w. So dot product in general settings is called inner product. That same terminology goes over when you have some abstract spaces also, okay. So we are getting out of the standard terminology of R2R3n, taking a general terminology and general definitions.

So dot product in Rn is also called inner product. So obvious properties that inner product is always v,v, right, inner product of v with itself will always non-negative, right. And it is 0. When it will be 0? It is vi square, sum of squares=0, that means each vi must be equal to 0. So it is 0 if and only if the vector v is 0. Is it okay? Right. So this is also called the norm or the length of this vector, right.

Is the magnitude in R2 or R3, we start calling it as the norm of the vector or you can also call the length of the vector. So what is that? That v, inner product v with v, that gives you sigma vi square square root, right. So we say 2 vectors are orthogonal in Rn are mutually orthogonal if the dot product or the inner product between them is equal to 0.

Same definition carried over now. See if v is perpendicular to w, then w also is perpendicular to v, because this multiplication is commutative, right. Whether you write inner product v with w, dot product with v and w, that is same as dot product of w with v. So it does not matter. So this is a symmetric. v perpendicular to w is same as w perpendicular to v.

(Refer Slide Time: 03:16)



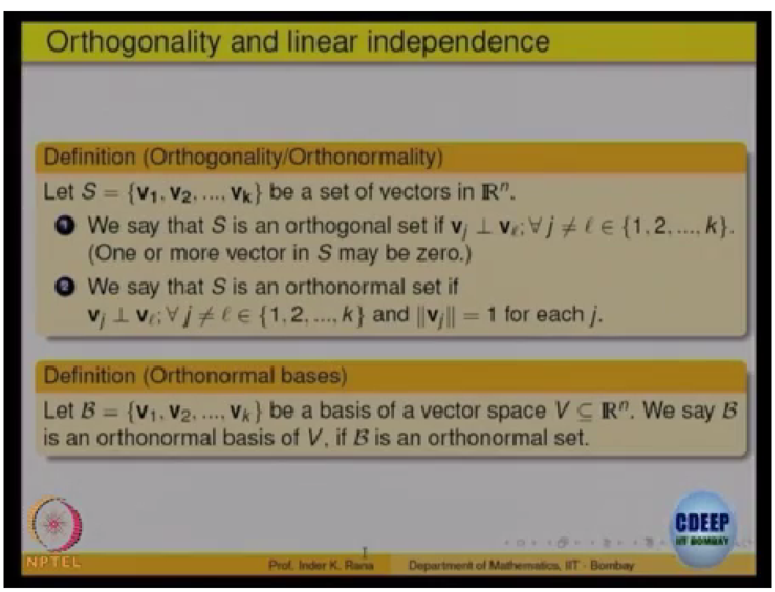
Here are somethings which I think we will not prove them. We will just assume. This is what is called Cauchy-Schwartz inequality, you must have proved it in R2 and R3 as a; same proof actually continues. If you go back and look at the same proof, it continues. Namely, the absolute value of the dot product of v and w is less than, because essentially the formula comes. This is equal to norm v, norm v cos theta, right and absolute value of cos; same relation actually is valid.

You can give. But here, in Rn, you first proof this and then define angle. See this is, in R2 and R3, this is not proved using cos theta, okay. Because we do not know what is theta. How to define the angle in Rn. You first proof this. So if it is less than or equal to, it is equal to; multiplied by something, right, if you want to write equality, that equality that thing you call as cos theta.

So that is how theta is defined. So this divided by this is called cos of the angle between those 2 vectors in Rn, okay. We will not going to do that. So what I am saying is, a proof of this without going to theta, can be given, without (()) (04:38) angle, it can be given. Triangle inequality, this follows from Cauchy-Schwartz actually, right. If you write this as a square open out, then this follows.

We will not, I am just indicating some proofs of that because some of you may be keen to know what is the proof but for examination point of view, we are not including the proofs of this, okay. You just going to use these things and Pythagoras theorem, right, if 2 perpendicular A+B square=A square+B square. So that essentially goes, right. So I have said proof is self-study, self-enrichment I would call it. If you want to know really how it works, then try to read a proof, okay, right.

(Refer Slide Time: 05:27)



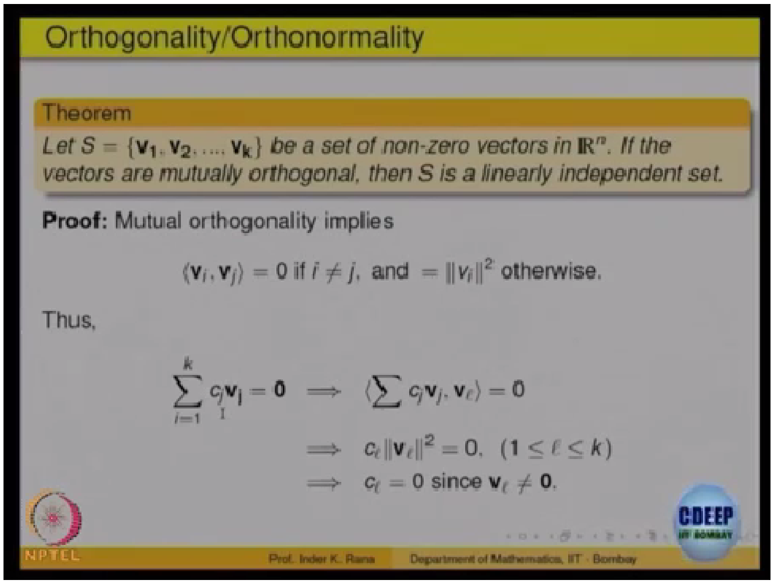
So let us now formally define what is orthonormal bases, okay. So given a set S, v1 v2, vk, a set of vectors in Rn, we say that this forms an orthogonal set if vij is perpendicular to vl whenever j is not equal to l. Any 2 of them, distinct ones are perpendicular to each other, right. One of the vectors may be 0, we are not saying all are non-0. The 0 vector will be a perpendicular to everything, you can say that way, right.

And say orthonormal, orthogonality is different from orthogonal essentially. Orthogonal is perpendicular to each other and orthonormal says not only it is orthogonal, something more, the norm of each vector is 1, magnitude is equal to 1. So it is something more than being orthogonal and as a consequence of this because norm of each vector is 1, 0 cannot appear in an orthonormal set.

In an orthonormal set, 0 cannot be a member. Because magnitude is 1. If there is a vector 0, then its magnitude is 0. So if a set of vectors is orthonormal, then 0 cannot be a member of that. Be sure of that, right. So definition of an orthonormal bases, bases v1 v2 vk of a vector space v in Rn is called orthonormal bases if this forms an orthonormal set.

That means it is a bases, right like every vector is a linear combination, +added property, length of each vector is equal to 1. And any 2 of them are perpendicular to each other when they are distinct. Then we say it is an orthonormal bases. And advantage I have already showed you what is the advantage of an orthonormal bases?

(Refer Slide Time: 07:26)



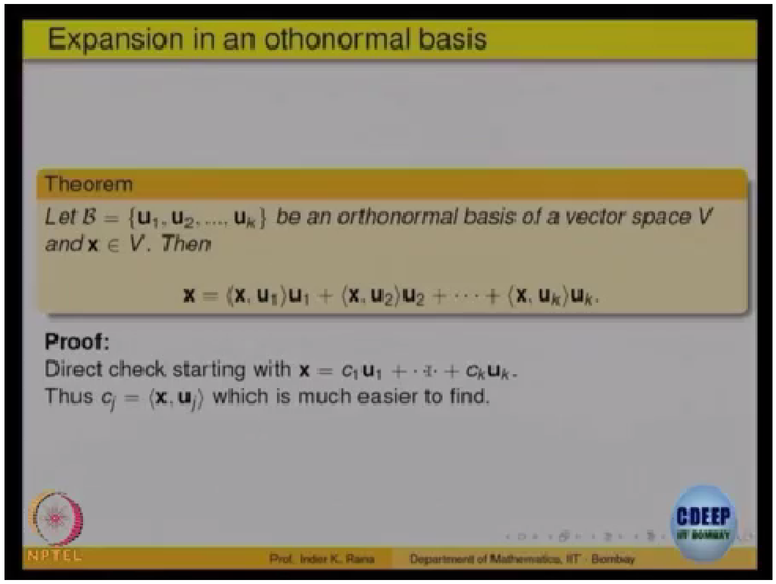
v1 v2 vk, okay, if they are orthonormal, right, then what happens? I can compute the coordinates very easily, right, okay. Okay, this is a different, I am going somewhere else. So let us first proof if we are given a orthonormal set, then it is automatically linearly independent. So it gives you is it is something more than that. Why it is linearly independent? Take a linear combination, cjvj=0, right.

I can take inner product of this on both sides with any vector vj, right. Take the inner product on both sides. On the right hand side, inner product will be 0, that will be 0. On the left hand side, what you will get? cjvjvl. Summation is of vl I have taken, running index is j. I have taken with vl, okay, any particular one. Then what you will get? What happens to this inner product? This is 0 when j is not equal to l. When j=l, it is norm of vj square.

So only 1 term survives here, right when j=l. So this is equal to cl norm vl square. And that is equal to 0 means what? This is orthonormal vector, so this is 1. That means cl should be equal to 0, right. So what you are seeing is if a linear combination of an orthonormal set is 0, then each one of them must be 0 simply by taking dot product on both sides for each vl.

Only one term in the sum will survive when vl, j=l and that gives you, this is 0 and this norm being 1, cl=0. So an orthonormal set automatically is linearly independent. The question comes, how do I make a linearly independent set orthonormal, right? So we will do that. But before that, let us just go through what I was saying.

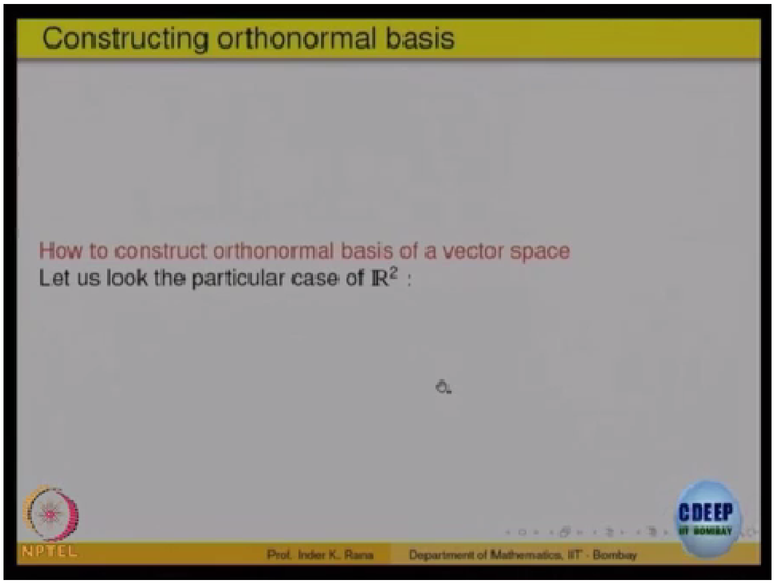
(Refer Slide Time: 09:39)



If B is an orthonormal basis, then what is that vector x. It is a linear combination of u1 u2 uk, right, because that is the basis. It should be linear combination. But it says what are those scalars? They are precisely xu1 xu2 dot product xuk dot product. So coordinates which I said earlier, if you are given an orthonormal basis, its coordinates are immediately known, right. Those unique scalars alpha 1 to alpha n which give a linear combination, is immediately known.

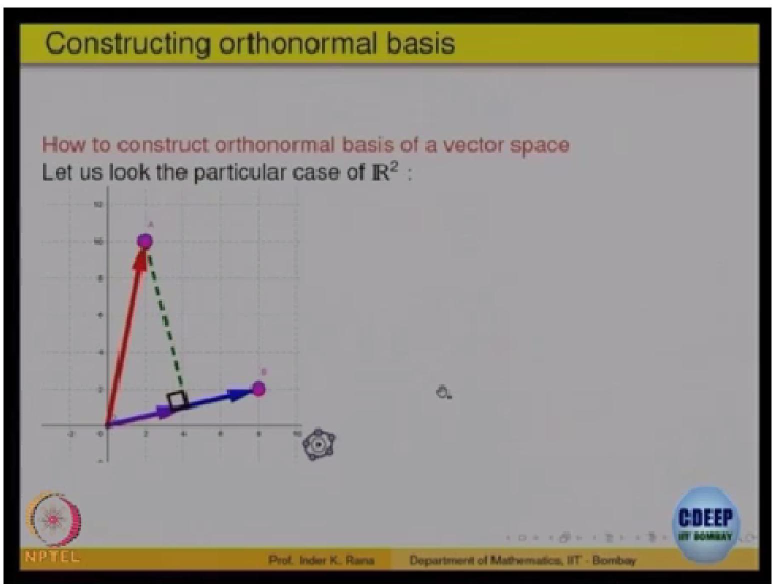
Take the vector, form its dot product with that corresponding basis elements in the orthonormal basis, you get the corresponding coefficients. So that is the advantage of orthonormal basis. So one would like whenever you are given a vector space, I would like to convert it into a or I would like to construct out of it a orthonormal basis, right. So because of this advantage.

(Refer Slide Time: 10:39)



So to do that, there is the question, how to do that. And the process is very simple.

(Refer Slide Time: 10:46)



It is very geometric. Let us look at this, okay. Supposing this is a vector, red one is 1 vector, this is another vector. What is called the projection of 1 vector on the other, ordinarily? You will say drop a perpendicular, right. Is the shadow, right. Projection is just the shadow of something. If you are standing on the ground, what is your projection? Your shadow, right. So if I drop a perpendicular, then this is a projection, right, up to here.

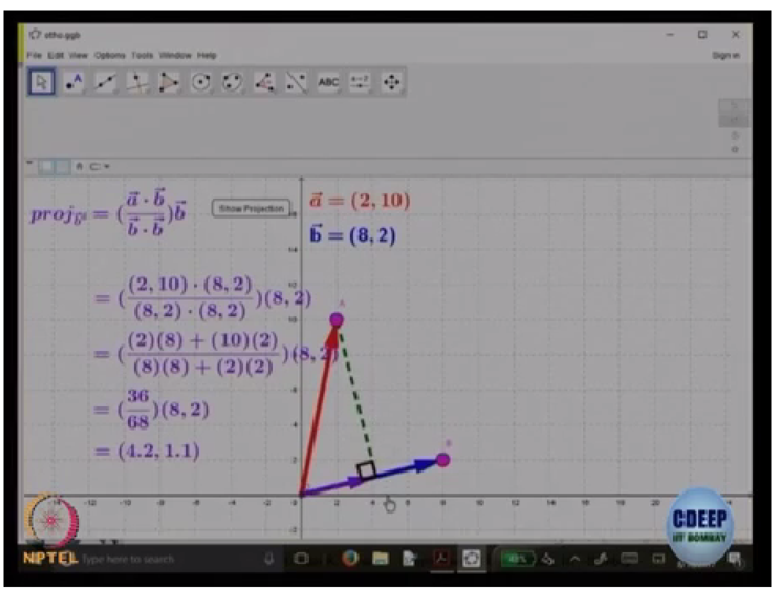
Now from this vector, if I subtract that, what I will get? This is a vector, this is another vector, I subtract, I will get this vector, right. Vector addition. This plus this should be, right, equal to, but

I am taking, opposite direction is there. So that gives me this vector. So from this vector if I subtract the projection, I get a vector which is perpendicular to it. That is the simple idea. And if I look at the plane, plane can be generated by this vector and this vector, original.

Or it can be generated by this and this perpendicular vectors. These 2 perpendicular, one is the original one, other is obtained from the second one by removing the projection, I get a perpendicular one and these 2 generate the same thing which the earlier one was generating. So that is the basic idea. So what we do, given some vectors, first one take it as it is, okay. The next one, remove from it the projection of this in that, right, remove that, you get a second vector.

Now you have got 2 vectors which are perpendicular to each other and generate a subspace. Now you are given a third vector. So what you will do? How you will get to the third one, perpendicular to this and giving the same thing? For the third one, project it on the plane now and remove that projection, subtract that projection, you will get a perpendicular one, right. So that one along with that plane will give you all, whole of 3 thing, 3-dimension. So that process you go on doing it, okay. I just wanted to show you something, right.

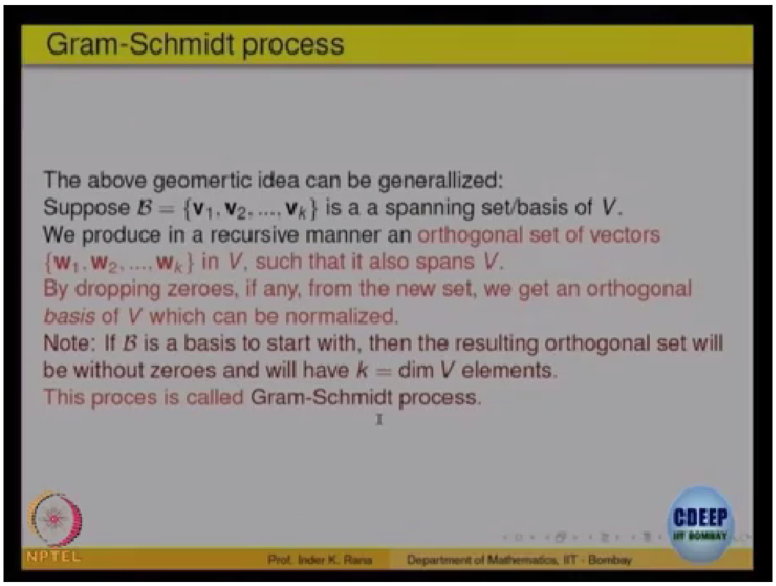
(Refer Slide Time: 13:06)



So here not only it gives you, what is this vector, is 2,n. So this gives you 1=2,10, right. b=8,2. So it gives you b. And it tells you what is the projection? The formula is, you must have done it calculating projection in terms of dot product, right. What is that, if you know the angle theta, projection is cos theta of that and from there you can easily calculate that, right. So it says it is a.b/b.b*b, that is the projection, that is this projection that you will get here and you subtract.

So it gives you everything. It gives you the computation of, so this gives you, right. This vector is, right. This is the projection vector is computing only. This is a projection. Now if you subtract from the original one, you will get this perpendicular one, right. You can do it in R3 also. This has a possibility of doing in R3 also. Nice to play with this and that is how you will learn also what is the formula, right.

(Refer Slide Time: 14:23)



So here is what we do. Suppose you are given what we start with is, given a vector space, start with a basis of it. Because eventually when you want to convert something into orthonormal, that is going to be linearly independent anyway, right. And if you know how to construct basis of a vector space. So start with a basis, right. If not basis, at least it should be a spanning set, meaning every vector is a linear combination of these vectors.

At least that much you should start doing, okay. What we want to do? We want a process by which I can construct w1 w2 wk which is orthonormal and spans the same space as this one, right. That is what we want to do. And the process that we had done, will be, first of all what you can do is? You can drop 0's if anything in this, the original one, right. Because the basis, it want nothing will be 0.

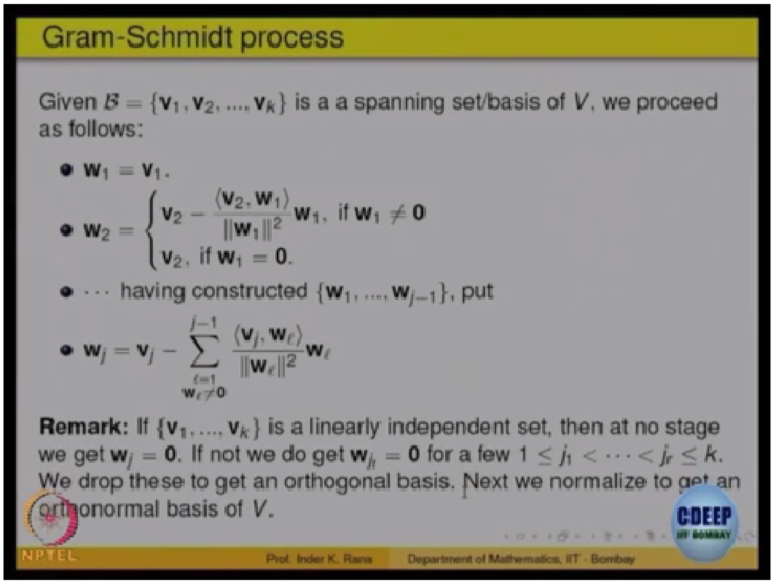
If it is spanning set, there may be some 0. So drop that, okay. They are not going to be useful, okay. And so what we do is, first we, from these v1 v2 vk, first we get those perpendicular ones, right. From that vector, subtract the projection, you get a vector which is perpendicular to this. But the perpendicular one, the 2 new ones that you have gotten, they are perpendicular to each other but they may not be of magnitude 1.

So first what you do is? The given set of vector, you convert them into an orthogonal set where each is perpendicular to each other and then divide by each one by its norm, you will get a normal one also, right. So that is the 2 step. First orthogonalizes it, then normalize it also, okay. So that is what the process. And this process is what is called Gram-Schmidt process. This process of orthonormalization is called Gram-Schmidt process of given a spanning set or a basis.

If the basis, none of the vectors that you construct successfully will be 0. But if the 2 are dependent on each other; you can see, if 2 vectors in R2 which are linearly dependent on each other, then what is the projection? It is not going to give you anything new, right. So you will be getting when you subtract that, you will get 0 only, right.

You will get; so v from v, you will get 0 only, right. So if 2 vectors are linearly dependent in the given set, when you do orthonormalization, this Gram-Schmidt process, you may get a 0 if one of this is linearly dependent on the previous ones. We will drop them and go to the next step, okay. So that is the idea. So here is the process.

(Refer Slide Time: 17:18)



The first one, v1 is given to you. So define w1 as v1, okay? What is w2? If this w1 is not 0, right, you could have dropped them. You need not right this case at all. v2-the projection, this is the projection of w2 on w1. So this is what you are subtracting and the remaining one we said should be a perpendicular. From v2, from the second one, remove the projection of first one on the second one, right.

So that is the first stage. You have gotten w2. And you go on doing it inductively. Up to j-1 you have done it. You want the next one. So what is wj? From vj, right, remember that plane thing of 2 of them you have gotten the plane perpendicular one. The third one, drop the projection on to that plane now and remove that. So that is the projection on the remaining ones, j-1. So on j-1, take wl, take sorry vj, take its projection on the linear combination of the remaining ones, that plane, essentially.

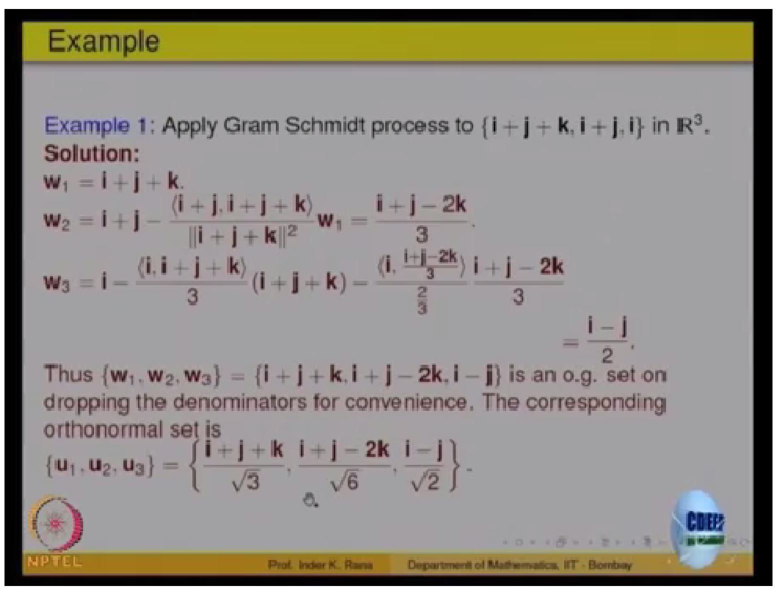
So how to get from second to the third one and remove it from vj. Remove the projection. So this is the projection on the earlier ones put together on that space generated by that. Remove that and that you should be getting a perpendicular to the original one, right. So this is an algorithm which you can compute successively. First one, as it is. Second one, remove the projection. What is the projection?

v2w1/the norm of w1 square w1, right. And this is the jth one and so on. So you will finish the induction when you have reached k, right. It is the inductive step. First one, using the first one, you construct the second one; to third one, you use first one and the second one which you have already constructed, right. Say this wl's up to j-1, these have already been constructed, right. So use them to construct the next one, plus removing the projection, that is the idea.

And that should give you an orthogonal set. This w1 w2 wk will be an orthogonal set whose span will be same as span of v1 v2 vk. If none of them is 0, right, as many as the original one, right. You can normalize them now. Each one ew, divide each w/its norm to get the normal vector so that it is an orthonormal basis. So from the given basis, you get an orthonormal basis, okay. So that is the process of; and the remark is just saying that the corresponding ones, if they are coming 0, just drop them, okay.

They will come if they are linearly dependent. So let us do one example for this.

(Refer Slide Time: 20:26)



Let us take these 3 vectors. What are I, j and k? They are the centered vectors in R3, right. i is 1 0 0, j is 0 1 0 and k is 0 0 1. I have taken v1. What is v1? i+j+k. So what will be this? 1 1 1. Right. In the ordinary notation, right, it will be 1 1 1, not in the vector, in the coordinate fashion. What will be i+j? 1 1, third component is 0 and i, that is 1 0 0, right. Check they form a basis of the given vector space, base that is R3.

They also form a basis, we can check, okay. So what is w1, the first one, i+j+k, right? What is w2? Remove the projection. So what is w2? i+j, that is the original one, right. Second one is i+j, w2, minus the projection. So that is i+j, that is the second one. Dot product with a.b, right. So dot with i+j+k, that is already done, w1, divided by the norm of that*w1. So that is the part of w1 you are removing, right from v2.

See essentially saying, you have got 2 components of w2. One in that direction, one in the perpendicular direction, right. Remove that component, you are left with this, the perpendicular one. You can look at that way also, right. So from i+j, that is the second one, I am removing the projection of w1. W1 is same as v1, removing the projection of that on to the second one, right. Then remove the projection, you get a vector this one.

And once you do your computation yourself, you will do it, right, okay. What is the third one? How do you get the third one now? W1 and w2 have been constructed. Coordinate w3. Look at the third one, that is my I, from that add to remove the project, some of the projections on the earlier ones. What are the earlier ones? w1 and w2. So let us remove the projection. So i, that is the last one, minus, this is the projection on i+j+k, this is the projection on i+j-, on i and i+jk, right.

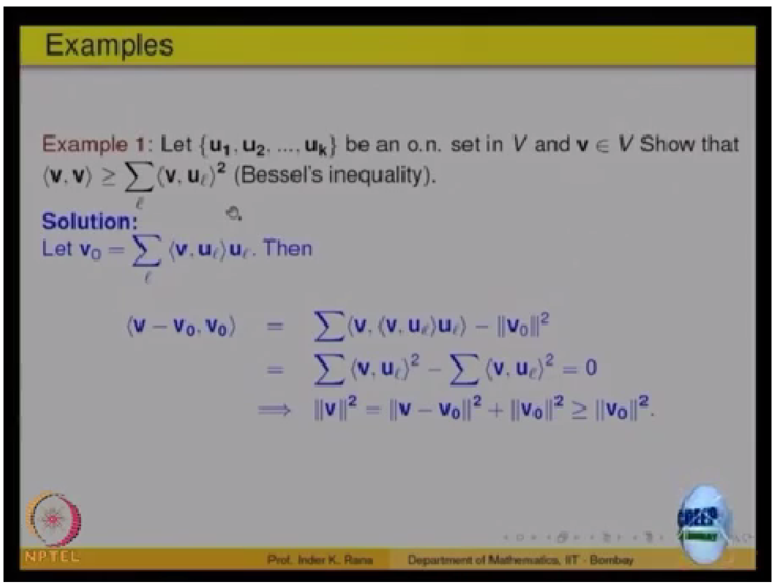
So this is w2, right. i+j-2k. So w1 and w2, for w1 and w2, remove those projections, the 2 terms and what you get is, if everything is written correctly, i-j/2. So from v1 v2 v3, we got w1 w2 w3. Claim is they are mutually orthogonal, right. What is the dot product of this with this? You can check, that should come out 0 if everything is okay, okay. And now normalize them, w1 w2 and w3, normalize them, you get an orthonormal basis.

So process is clear. So I am just given you a simple example, try to, you can note it down if this and try to work out yourself what is the orthonormal basis? And check when I put the slides on the mural, check whether my computation is okay or not. If wrong, let me know if there is some typo in your thing, okay. It will give you also a practice of doing things, okay. So it is clear to everybody, how do you get the orthonormal basis?

Given a linearly independent set, let us assume that you are given an independent set which is a basis. If not, you can still do the process. At some stage, your w will become 0. So drop that as if it is not there and go to the next one. So let us assume you have got linearly independent set which is a basis, v1 v2 vk, first one, w1 is same as v1. Start with any one of them you like, okay.

w2, look at v2, from that remove the projection of v1 or w1, you have already constructed w1, so remove the projection of w1. What you get is w2. w3, from v3, remove the projection of v3 on w1 and also on w2, sum of the two, remove that, you will get another one and go on doing this process. That will give you orthogonal set, then normalize each one of them to get your orthonormal basis, okay.

(Refer Slide Time: 25:22)



So another one, another solution, okay. So here is something theory, is not. So let us go through that. What it says if u1 u2 uk is an orthonormal set, not necessarily a basis. It may not be a, it is just an orthonormal set, okay. Then it says if the magnitude of that is always bigger than or equal to the dot product of u with each one them and equality comes when this becomes an orthonormal basis, okay.

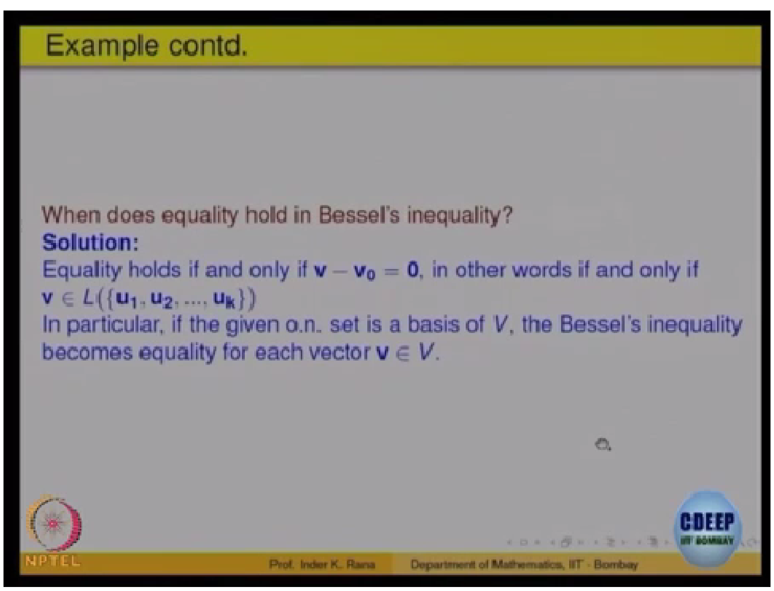
So okay, I do not want to say you something more which later on, it gets related with something called Fourier coefficients, those who have to do electrical engineering or something on Fourier series, you will see that these things come again, they will hit you again somewhere in some other form. There also it is called Bessel's inequality and Parseval's identity. Fourier coefficients will come, okay.

So this is related with something. So basically at present, what we are saying is, take a vector v0, I want to show this. So take any particular vector v0, okay. So that will be a linear combination. And what is the linear combination? So let us define v0 to be the just linear combination on the right hand side, right. And take the dot product. The dot product when you expand using linearity, will come out to be equal to this, right.

Here if you drop this, this is sum of 2 squares, this sign is a sum of 2 square terms, right. If I drop one of them, we will get bigger than or equal to, that is the other. Nothing more than that. Basically what we are saying is, this ul's are mutually orthogonal, right. So look at this and expand.

So what will happen in this thing. v, so these are sum. So it will be vul square-this square, that is equal to 0. And then you expand, okay, you will get norm v square=this and drop one of them. It is just writing using the fact that they are orthogonal, nothing more than that. You write and you will get it yourself, that is better, okay.

(Refer Slide Time: 27:40)



So equality holds when it becomes an orthonormal basis, right. So I think we will continue next time about this identity a bit. Basically the idea is given a basis on a vector space, how to make it an orthonormal basis and the importance of that lies in the fact that coordinates are immediately known and the process is a Gram-Schmidt orthonormalization process, right. So we will continue next time. Thank you.