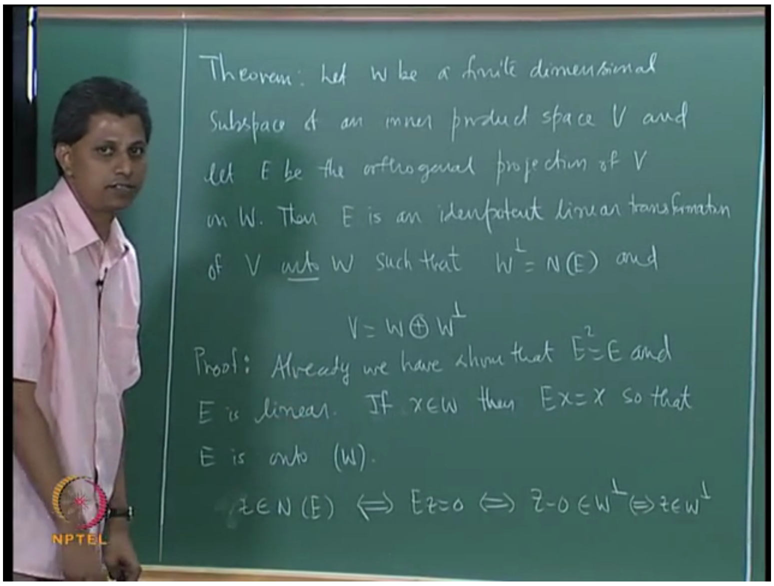
Linear Algebra
Professor K.C Shivakumar
Department of Mathematics
Indian Institute of Technology, Madras
Module 12-Best Approximation
Lecture 45
Projection Theorem Linear Functionals

Let me recall we are proving this result.

(Refer Slide Time: 00:21)



Let W be a finite dimensional subspace of an inner product space V and let E be the orthogonal projection of V on W ok then E is an idempotent linear transformation, an idempotent linear transformation of V onto W so we must show that this am onto map, such that the following hold. First condition is that W perpendicular is the null space of E and see remember in order to write this we must know that the E is linear, see what we knows is that the null space or range space of a linear transformation is a subspace.

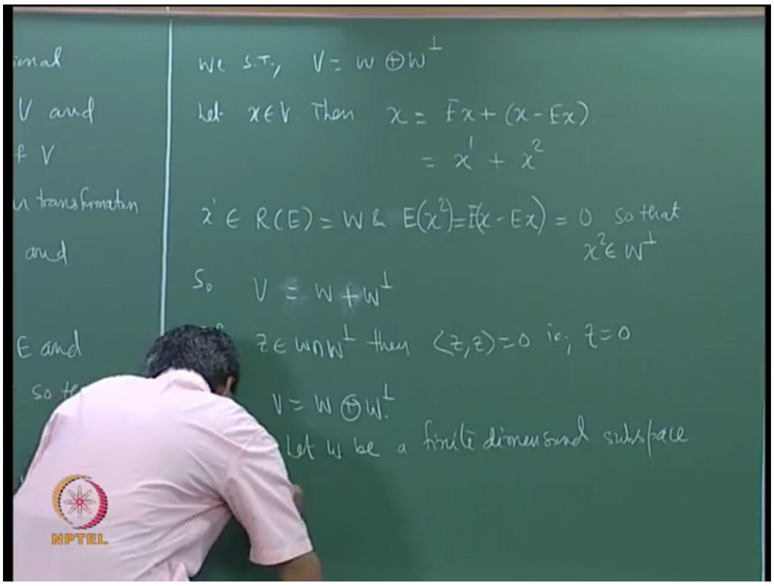
So in order to write this we must know that E is linear ok, this is one condition that E must satisfy and more importantly V is the orthogonal direct sum of this two subspaces ok. For the proof I remember that we have done E is idempotent and V is linear ok. So let me just write this already we have shown that E square equal to E that is E is idempotent and E is linear, we must show that V is onto in W perpendicular is null space of V and this ok. Lets first dispose this off, this is straight forward V is onto W that is given any x in W can I find ok, given any U in W can i find an x in V such that E x equals U? It is the same thing.

So let me write like this, if x belongs to W then what we know is that E x equal to x so there are (())(03:28) x itself. So that E is onto let me just emphasize W, see E is a linear operator on V so from V to V it is not an onto map but if you look at the subspace W think of E as a mapping from V to W then it is onto ok, that is what this means. So it is the orthogonal projection of V onto W we must show that W perpendicular is null space of E ok.

So lets take we need to show this and then this. Let's take ok lets say Y or u or z ok, Z belongs null space of E implies E z is zero by definition infact if and only if isn't it by definition? But E z is equal to zero if and only if remember that to any fixed vector z in V, E assigns the best approximation the unique best approximation to Z from W. So and we also know that condition, what is that condition? If V x equal to U, then x minus U is perpendicular to W this happens if and only if x minus U is perpendicular to W that is I am writing it like this, Z minus zero belongs to W perpendicular, that is the same as saying, z belongs to W perpendicular.

And so we have proved in one single step that the null space of E is W perpendicular ok, this is just the definition how is E defined, to any vectors Z it assigns the unique best approximation and that best approximation U satisfies x minus U perpendicular to W that is what I have used ok.

(Refer Slide Time: 05:48)



Finally this equation we need to show that but that straight forward we show that V is W direct sum W perpendicular. Lets take x in V then x can be written as I wanted to be in this plus this so I can write E x plus x minus E x.

Let's call this as x1 and this as x2, x1 plus x2 then x1 belongs to range of E but range of V is by definition W and x2 is x minus E x so I want you to look at E x2 is E of x minus E x that is zero E x minus E square x but E square is E, so E x2 is zero just now we have shown that null space of E is W perpendicular so x2 belongs to W perpendicular. So for one thing we have shown that V can be written as a sum of W1 sorry W and W perpendicular, we must show that it is a direct sum that is we must show that the intersection is the subspace consisting of the zero vector alone but that is straight forward.

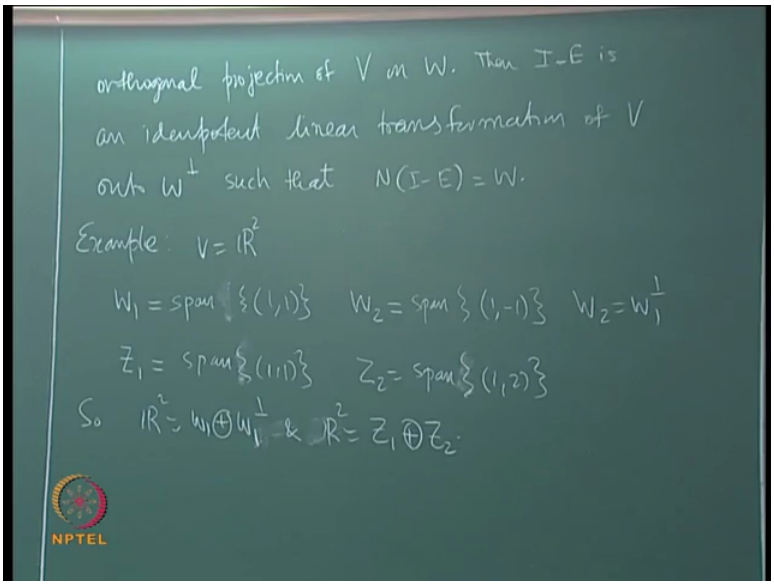
So what we have shown is that, E is contained in W direct sum W perpendicular so I can write equality because this is a subspace direct sum is also a subspace so this is equal, I am sorry, W equals just W plus ok finally if lets say Z, Z belongs to W intersection W perpendicular then by definition the inner product of Z with itself must be zero because Z is orthogonal to every vector in W and Z belongs to W perpendicular also but this means Z is zero. So the intersection is single term and so it follows that V is that direct sum of these two subspaces.

So remember that there is no restriction on V, the restriction is only on W, it must be a finite dimensional subspace V could be an infinite dimensional inner product space but if W is not a finite dimensional subspace then this is not true ok. If time permits I will provide an example later. Ok so this is finite dimensional result. Ok corresponding to E, we have this result the version for F which kind of compliments E that version is similar, so let me give you that version for the record. So the following is a consequence of the previous result Corollary.

Remember the mapping F, F is I minus E, ok and we have seen that F is a projection of V onto of V on W perpendicular but what follows from this theorem is that it is an idempotent linear transformation of V onto W perpendicular this time and what is a null space of F? What do you expect the null space of F to be? What is F? F is I minus E, null space of is W perpendicular I am sorry just W ok. Null space of F is range of I minus E ok so lets null space of F is range of E yes that is W so lets quickly dispose this one.

Corresponding to E we define an S do.

(Refer Slide Time: 10:37)



Let W be a finite dimensional subspace of an inner product space V and E denote the orthogonal projection of V on W as before then I minus E is an idempotent linear transformation of V onto W perpendicular such that the null space of I minus E is W. So previously we had W perpendicular to be the null space of E that is I am looking at the null space of the orthogonal projection at any stage. So a null space of I minus E is W, the proof is almost there so I will not write down.

The fact that E is idempotent and linear implies that I minus E is linear in the first place and I minus E we have seen this last time I minus E square is I minus E so it is an idempotent linear transformation. I minus E that is the mapping that takes x to x minus E x we have seen last time that it is the orthogonal projection of V on W perpendicular that was proved, we called it F the mapping F from V to V defined by F of x equals x minus E x is the orthogonal projection of V on W perpendicular but from what we have seen just now it must be onto also similar to E, I minus E behaves very similar to E but its complimentary to E.

So it is an orthogonal projection of V onto W perpendicular so this is done and null space of I minus E equals W is straight forward very similar to what we have done before ok, so I am going to leave this last part also ok, before so this is kind of the summary of the notion of best approximation ok how it helps in decomposing an inner product space in terms of a finite dimensional subspace and its orthogonal compliment ok. So what is important is the notion of best approximation ok for but remember that for finite dimensional subspaces the best approximation exists and it is unique that is I mean this two are important for us to do al this things.

Using again this notion we will go to the concept of adjoint of a linear transformation then the notion of unitary operators, normal operators, finally spectral finite dimensional spectral theorem ok but before that lets look at two examples ok, two numerical examples where I will show how ok where I will try to give you another clue as to what you must expect of the linear transformation E when it is the orthogonal projection ok I asked you this question, what is a difference between an ordinary decomposition direct sum decomposition and an orthogonal direct sum decomposition ok.

You may be able to answer by looking at this example so this is kind of reinforcing what we have done till now. So lets look at this example really two examples, lets take R2 ok, for R2 I have the following two decompositions, let me call W1 as a subspace that is span of 1, 1 W1 is span of 1, 1 ok W2 is span of 1, minus 1 and I will write this are row vectors then what is the relationship between this two? Can I say W2 is W1 perpendicular? Ok observe that W2 is W1 perpendicular lets look at another pair, I will call it Z1, Z1 is span of I will take the same vector as before, so Z1 is really W1, Z2 lets say is span of the vector 1 2.

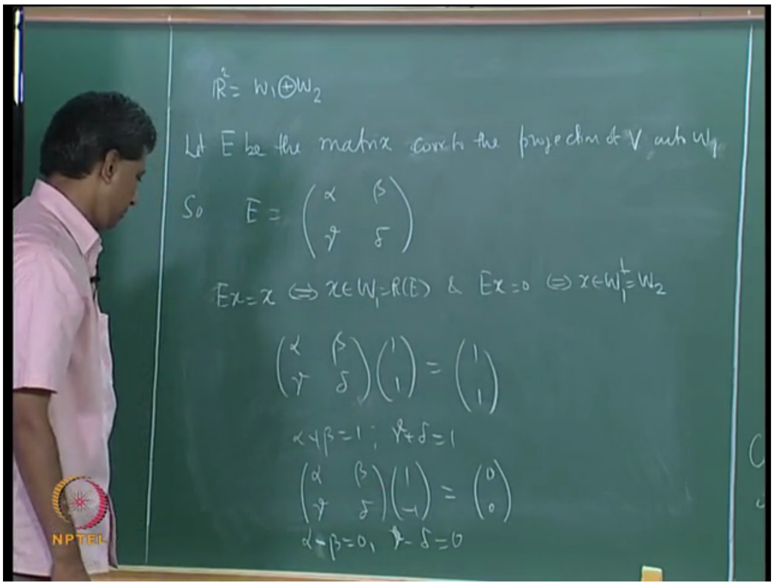
In the first instance obviously W1 direct sum W perpendicular is R2 in the second case can I say Z1 direct sum Z2 is R2? (First) the answer must be instant, answer is yes, why? This are independent vectors together they form a basis and this vector does not belong to the span of this, this doesn’t belong to the span of this, so the intersection is single term zero ok. So for the record R2 can be written as W1 direct sum W2 perpendicular and equal to ok let me write like this and R2 equals Z1 direct sum Z2, the intersection is single term zero and this two vectors together form a basis so any vector can be written as a linear combination of this two.

So sum Z1 plus Z2 is R2 is clear, Z1 intersection Z2 single term zero is also clear because this are independent vectors just by inspection. W1 direct sum W2 or I will write W1 perpendicular yes, ok lets find the maps E in both this cases ok. What is the property of the map E that we are going to use? So subspaces are given we want to find E, what I will do is not find E, I will just take E to be a matrix itself, remember there is this 1 to 1 correspondence between the matrix of a linear transformation and the transformation itself.

I will write down the matrix of E relative to the standard basis and see how it looks like by the way can you geometrically differentiate between this two direct sum decomposition, (geometric) for the first one it is 1, 1 you can think of the line y equals x ok that is a line which makes 45 degrees with the positive x axis the line 1 minus 1 ok makes 135 degrees with positive real axis and observe these two lines are perpendicular.

In the second case this is the same subspace but look at Z2, Z2 is not perpendicular the line passing through 1, 2 is not perpendicular to the line passing through 1, 1 ok but then you know that since they are independent the span of these two vectors must be the entire space but remember that the angle between this two lines, one passing through 1, 1 the other one passing through 1, 2 are not perpendicular this is crucial.

(Refer Slide Time: 19:27)



In the first case I want to determine E, such that ok, let E be the matrix I am using the transformation notation itself if to denote the matrix let E be the matrix corresponding to ok, the projection of W1 onto W2 and so E is lets say alpha beta gamma delta E must be a 2 by 2 matrix, it is a linear transformation on R2, E is this to determine E completely I must use a property that E acts like identity on the range of E W1 and it must act like the zero operator on the perpendicular, do you agree? Range of E is W null space of E is W perpendicular ok and over range of V that is if x belongs to range of E then E x equal to x.

So E acts like identity on its range E acts like identity on W, it acts like a zero operator on W perpendicular, two conditions but you will get four equation in four unknowns alpha beta gamma delta so will determine E completely, so lets do that in both the cases, in the second case also this time we cannot talk about the orthogonal projection we will just talk about the usual projection, it is called an obliged projection, obliged projection so I will determine E in the second case such that E acts like identity on W1 that is Z1 and E acts like zero on Z2 then I will compare this two case, look at the structure of E how does E look like.

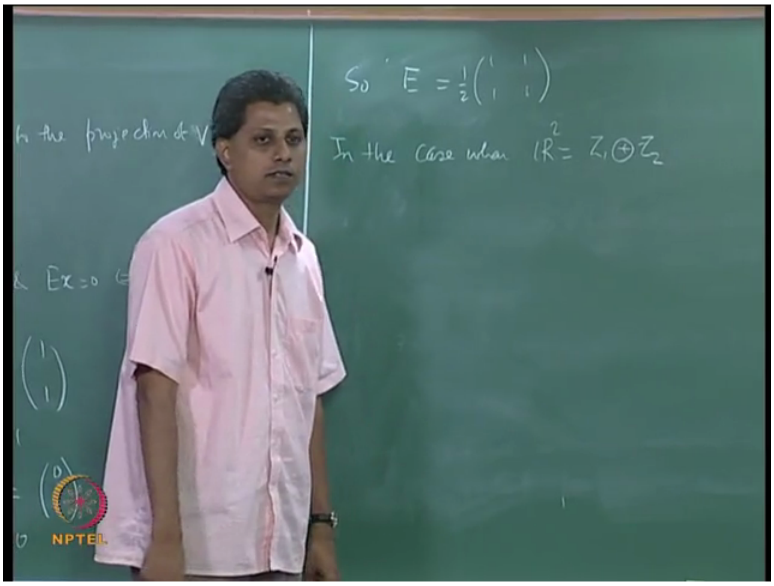
So first E x equals x if and only if x belongs to W that is range of E and so I am using this condition W1 right and E x equal to zero if and only if x belongs to W1 perpendicular that is W2in this example range of E is one dimensional W1 is one dimensional and span by the vectors 1 1, see I have not taken an orthonormal basis that does not matter, span by W1 is span by 1, 1 so E of that vector must be itself. So for one thing alpha beta gamma delta operating on 1 1 must be 1 1, 1, 1 belongs to W1 infact that is only vector in this only independent vector in W

Student: (())(22:48)

Professor: Yes it is V, V onto yes, projection of V onto W1 yes, I have this in the back of mind that this will correspond to W1, W2 yes it is a projection of V onto W1 ok. So I have this so this equation gives us rise to two equations ok, this matrix equation tells me alpha plus beta equals 1 gamma plus delta equals 1 ok two equations in four unknowns, two more equations will come from the second set of equations E x equal to zero so all the four unknowns can be determine. So look at this E x equal to zero if and only if x belongs to W2 I will take that vector 1, minus 1 equate that to zero alpha beta gamma delta into 1, minus 1 that must be this time the zero vector 0, 0 ok.

That gives rise to this two equations alpha plus beta is zero gamma sorry, alpha minus beta is zero gamma minus delta is zero, so alpha equal to beta equals half gamma equals delta equals half ok.

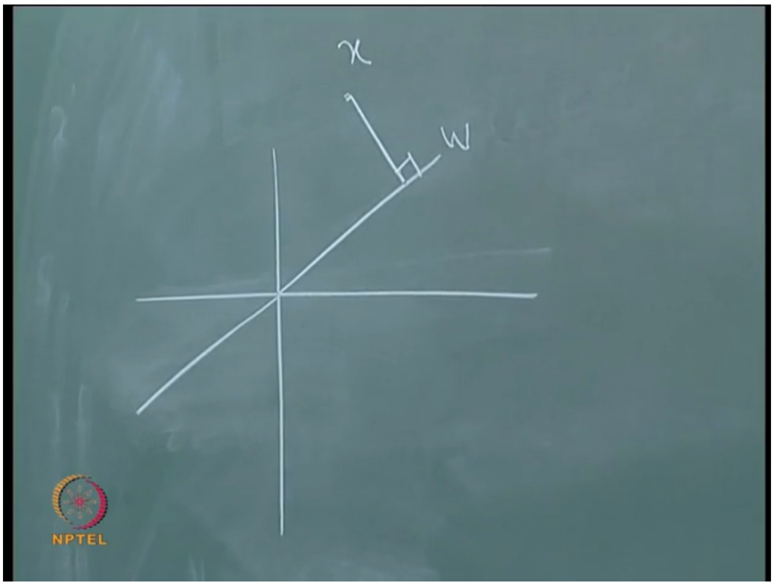
(Refer Slide Time: 24:21)



So let me write down E in this case, are my calculations correct? Ok lets go to the second case now, in the case when R2 is Z1 direct sum Z2, you remember that there is I don't know if I have told you this before there is this terminology associated with perpendicular subspaces this E is called the projection of V onto W1 along W2, E is a projection of V onto W1 along W2 that is what you are (calcu) along W2

Along you can look at the geometry, ok, what is a geometry? Let's go back to the best approximation problem.

(Refer Slide Time: 25:41)

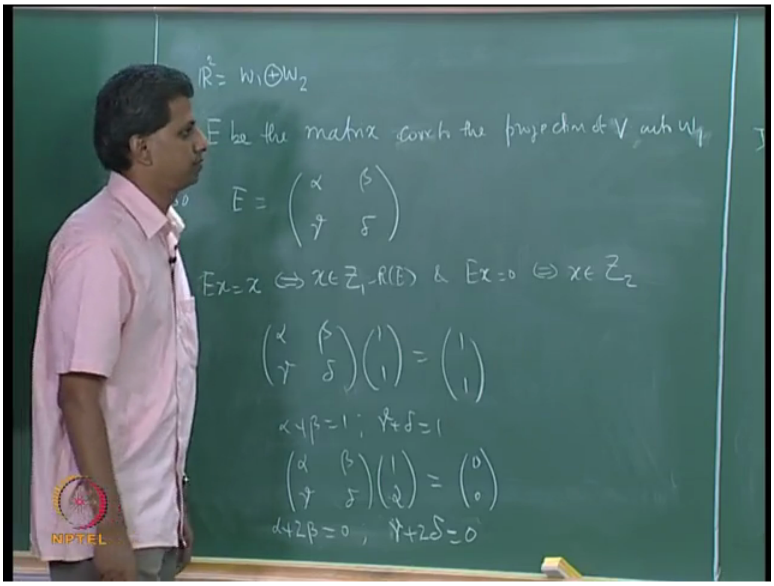


Say I have subspace pass through the origin, this is x ok, this is my subspace W and I want the projection of x onto W along W perpendicular, so I must go along perpendicular that is all. See this is perpendicular so I reach W along the perpendicular to W from x ok that is the reason it is called along W2. In this case it is the orthogonal projection but if you look at the case Z1 Z2 the second decomposition it is not orthogonal projection but we can still say that E is a linear transformation or the matrix is a linear transformation of R2 onto Z1 along Z2.

We can still say along Z2 ok, this time it is not an orthogonal projection, ok what I have tried to do is I have done the whole thing for orthogonal projections and I am trying to imitate this for the ordinary obliged projection I have taken the property of the orthogonal projection and trying to imitate to what it satisfies in the non-orthogonal case, what do I expect of a projection in the non-orthogonal case (what do I) essentially what do I expect? The property of a projection is that offcourse it is E square equal to E and it is a linear transformation ok.

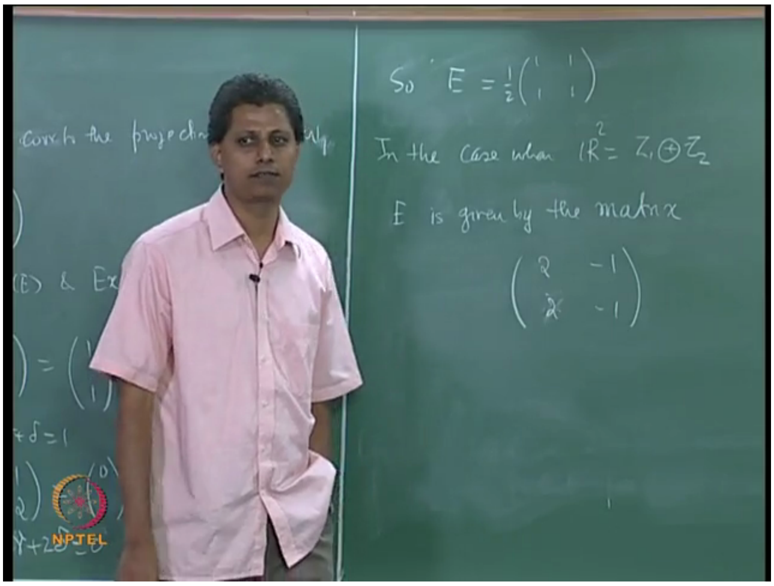
Apart from the what are the important properties, range of E is W1 null space of E is W2 this is essentially what I am trying to imitate in the non-orthogonal the obliged case and I am trying to see what is a difference between the matrixes that I get what is a difference between the transformations in this two cases? Ok, so lets go to this case when I write R2 in this manner again I will do a similar thing, so I will use this calculations.

(Refer Slide Time: 27:57)



I want E to act like identity on W1 which is Z1 and E x equal to zero this time Z2 ok, observe Z2 is not perpendicular to Z1 ok, identity this equation will remain as it is ok, Z1 is the same as W1 but the other one will be 1 2 ok this is the only change the other one will be 1 2. So I give this equations alpha plus 2 beta equals 0 gamma plus 2 delta equals 0.

(Refer Slide Time: 28:51)



In the case when R2 equals this E is given by the matrix just to avoid the confusion I don't want to write two E's, so E is given by this matrix, so what is that this time?

Can someone make the quick calculations alpha is minus 2 beta, beta is minus 1, alpha is 2 is that correct? Beta is minus 1 alpha is 2, the next one I will write without looking at the equations, what is the next one? It is not the same is it the same or a minus of that? It's a same ok, (the other) see this will be a rank 1 matrix, so the second row will be a multiple of the first row in this case it is the same as the first row This will be rank 1 matrix, you know why? This will be a rank 1 matrix, this is also a rank 1 matrix I mean this is not invertible this will also not be invertible ok, in this case why is it not invertible?

Professor: No-no please that is not my question, you mean I will give you a 2 by 2 matrix ask you to tell me the determinant, I want a qualitative answer.

Student: (())(30:33)

Professor: E is linear transformation from R to 2R, what prevents E from being a bijective linear transformation? This is the question, pardon, why is it not an injective?

Student: (())(30:56)

Professor: Why should it not be? That is the question.

Student: (())(31:02)

Can we expect E to be invertible? That is the question. Can we expect E to be invertible?

Student: if W equal to V then we expect.

Professor: Yes that is the only situation, only this are the only two extremes you expect so the only injective see E square equal to E if E is invertible what happens? E is identity, the other extreme, E square equal to 0, the other extreme is that E is 0, now in one case V is equal to W so W perpendicular is trivial in the other case V equal to single term 0, so W perpendicular is the entire space V, this are the two extremes ok. Invertiblity cannot happen for a projection, if it happens it is a identity, identity there is no problem, the problem is not there in the first place ok.

So E can never be invertible, ok range of E equal to W so if E is invertible the W is the whole of V so W perpendicular is trivial ok, so this is my E in the second case, what is a difference between the structures of E? Is there any difference? You can offcourse verify that E square equal to E in this example as well as in this example ok but what is there a qualitative difference between the structures of E in this two examples? If yes what is it? Is there something that you can say for the first one, which you cannot say for the second one?

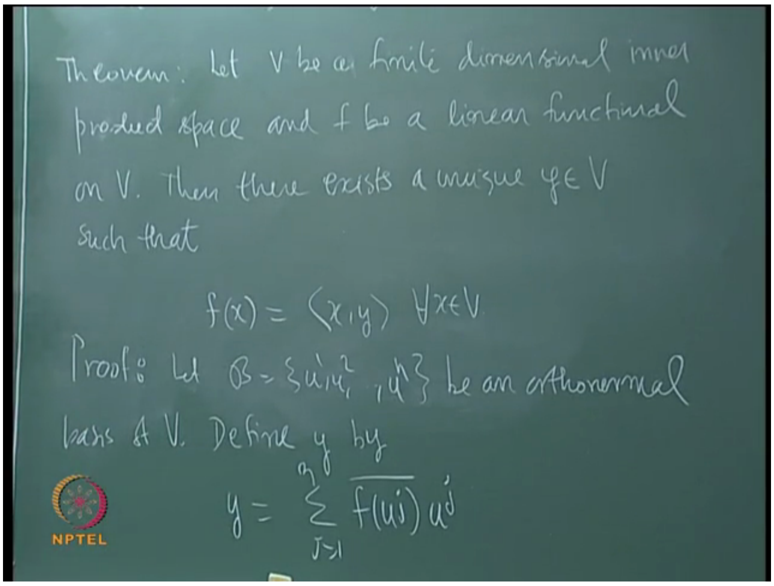
Structure, looking at the form E transpose is E, that is essential E transpose equal to E is infact another characterization for an orthogonal projection, E transpose not equal to E corresponds to the obliged projection. So E square equal to E, E transpose equal to E and range of E determine E completely E square equal to E, E transpose equal to E, range of E that is the subspace range of E must be known to me these three determine E uniquely in the orthogonal direct sum case, in the obliged direct sum case I must know range of E I must know I must know null space of E as well as the fact that E square equal to E.

In the obliged projection case, E square is not equal to E, sorry E transpose is not equal to E ok. So remember there is another important thing which I hope you have observed which is for the subspace W1 there is a unique sub space which together with the W1 forms a direct sum decomposition, this unique subspace is unique in the sense that it is perpendicular to W1 ok but for the second setup Z2 is just one complimentary subspace of Z1.

In place of Z2 I could have taken span 1, 3 I could have taken span 1, 4 all this are complimentary subspaces to Z1 ok. So in the case of orthogonal projection there is a unique decomposition of the vectors space V into W plus W perpendicular. In the case of ordinary projection there are infinitely many direct sum decomposition that is given a subspace W there are infinitely many complimentary subspaces that those can be obtained by an obliged projection. The unique subspace is obtained by the orthogonal projection ok.

So this is really the essential difference between an orthogonal projection and an obliged projection ok. So lets then move on to the next topic, ok the topic is linear functions and adjoints. Adjoints of linear transformations, linear transformations and adjoints of linear transformations, linear functionals is what I want to discuss first.

(Refer Slide Time: 35:49)



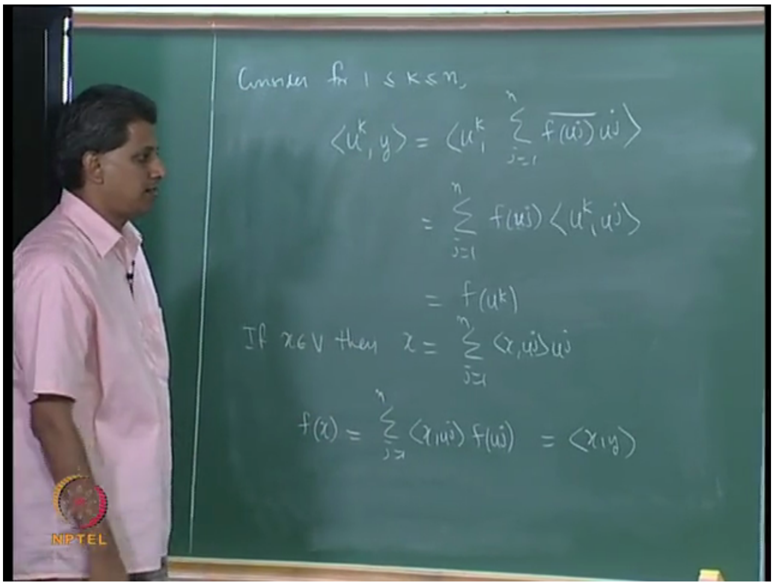
Linear Functionals and Adjoints, Adjoints of linear transformations, the notation of adjoint will generalize a notion of symmetry the operator, the operation of taking given A taking the operation A transpose.

Given A doing the operation A star that will be generalized in a inner product space will call that as a adjoint operation but before that linear functionals ok. So what is a linear functional? Definition, a linear functional definition, underline (())(36:39) ok that is a linear functional ok. What we will do is first give a representation for linear functions, this result is easy and this also should immediately remind you of a theorem in functional analysis which you will do a little later. Linear Functionals and Adjoints, so first I have the following, let V be a finite dimensional inner product space and F be a linear functional on V.

So the underline field is real or complex, so F is our linear transformation from V to R or C then the representation theorem says, there exists a unique vector Y in V such that the operation of this linear functional is like taking inner products with the vector Y such that the action of F on nay vector x is given by the inner products x, y for all x and V, for a fixed Y if it is a linear functional then it arises precisely in this manner ok the proof is just by producing the vector Y ok and will show that Y is unique ok this is called a representation theorem that is any linear functional is represented by this inner product .

The proof is we start with an orthonormal basis and give the formula for Y, lets say this is U1 U2 etc U n let B be an orthonormal basis of V, V is assumed to be a finite dimensional so it has an orthonormal finite orthonormal basis I know F I want to find Y which I must show is unique and satisfies this equation. Define Y by Y equals summation j equals 1 to n F of U j bar U j define Y to be this vector F is known so I can determine Y. Look at the image is of the basis vectors under F take the conjugates and then from this particular linear combination ok so will show that this Y satisfies this equation and show that this Y is unique.

(Refer Slide Time: 40:36)

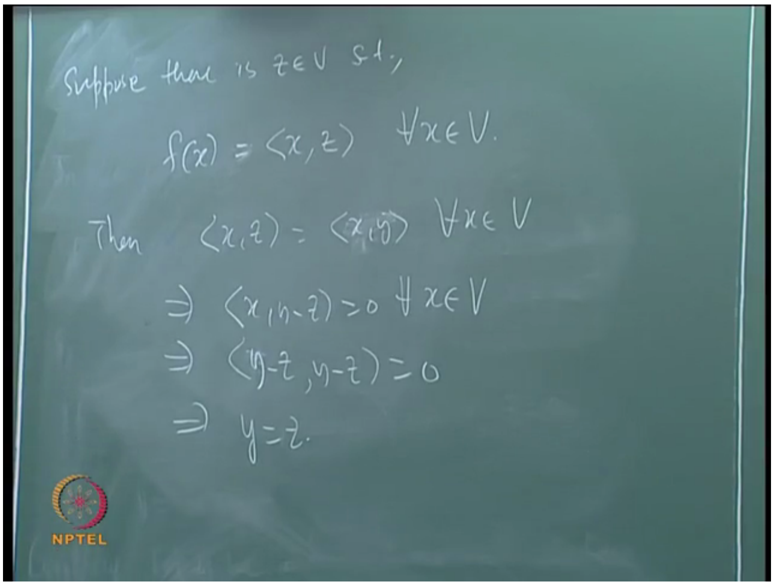


Consider for 1 less and or equal to k less and or equal to n the inner product of U k with Y this is inner product U k, Y is this summation j equals 1 to n F U j bar U j the inner product is conjugate linear with second one, so when this comes out it goes with the conjugate double conjugate summation F U j into the inner product U k with U j with j being the running index k has been fixed. So when j takes a value k this is 1 or all other terms are 0 this is an orthonormal basis so the only term that remains is when j is equal to k you substitute here it is F Uk.

So we have shown F Uk is Uk with Y this is an orthonormal basis so for any x this will be true because any x there is a linear combination right, is that clear that is if x belongs to V then x can be written as summation j equals 1 to n infact the coefficients are x U j with U j and so if you look at f x then f x is summation j equals 1 to n f is linear so it is x Uj F Uj but f U j is I want to show f x equals x, y ok x U j f U j this is ok could have been better to ok doesn’t matter can I just say this is inner product x with Y, you can verify, you could have started with x with y start with inner product x with y then it is x into that formula f U j goes out and then I must take the inner product of U j with x so I get this ok.

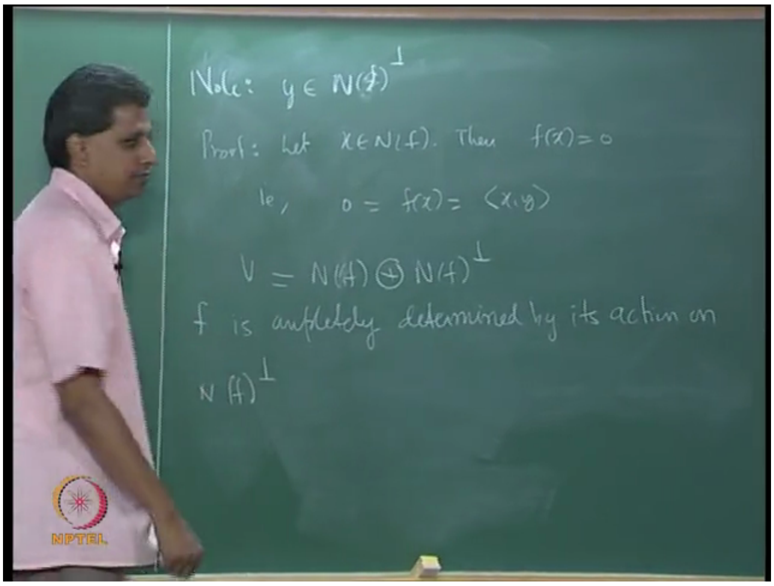
So you have this representation for any linear functional f, so there is a unique y uniqueness we have to show.

(Refer Slide Time: 43:22)



Suppose there is Z in V such that f x equals x, Z ok the you can compare this two it follows that inner product x with Z is inner product x with y this means that ok this means that inner product x with y minus Z equals zero for all x in V, this means since this is true for all x I can replace x by y minus Z which means y equal to Z, so uniqueness follows immediately. Ok so every linear functional on a finite dimensional inner product space arises from the inner product with a particular vector.

(Refer Slide Time: 44:32)



Observe that this y belongs to null space f perpendicular which is not coming from the proof but this can be shown, y belongs to null space of f perpendicular that is can you see why this is true? Proof, let x belong to null space of f I must show that inner product of x y is zero then f x is zero that is zero equals f x but f x I know is inner product of x with y. So this means y is perpendicular to null space of f if x is random vector from f then y with x is zero we have shown. So x belongs to null space of f V itself is finite dimensional null space of f is also finite dimensional so it has an orthogonal compliment.

So I can write V as null space of f direct sum null space of f perpendicular whenever you have a finite dimensional subspace it holds but V itself is finite dimensional so this holds. So what this means is that a linear functional on a finite dimensional inner product space is completely determine by its action on null space of f perpendicular, f is completely determined by its action on the subspace null space of f perpendicular. Offcourse the first term is null space of f so f takes a value zero there, f is completely determined by its action on null space of f perpendicular this is something which does not hold for a general linear transformation.

For a linear functional this result is true ok, using this representation theorem we will show how the conjugate transpose it is called the adjoint operator, will show how the adjoint operator for a linear transformation on a finite dimensional vector space exists and it is unique and also derives some of its properties yeah that will do next time.