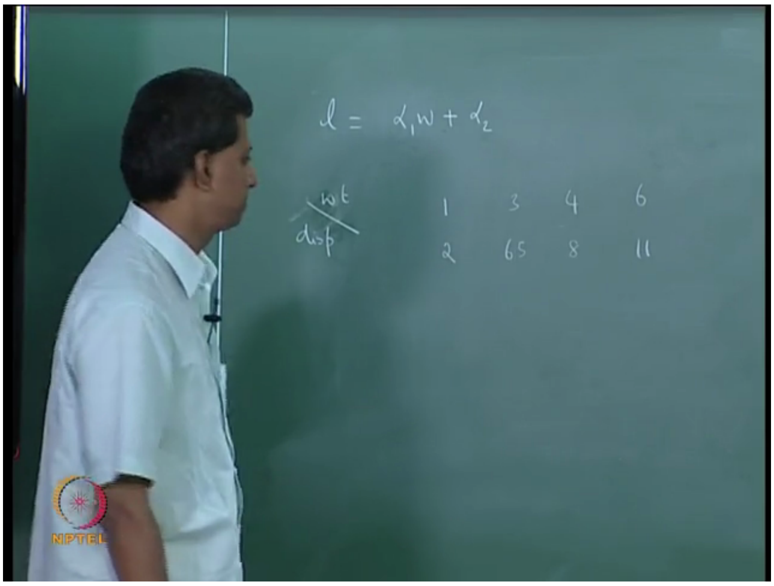
Linear Algebra
Professor K.C Shivakumar
Department of Mathematics
Indian Institute of Technology, Madras
Module 12-Best Approximation
Lecture 44
Orthogonal Complementary Subspaces, Orthogonal Projections

See I have not given an example coming from real life for this least squares solution, so let me give one example the model comes from physics Hooks law which states that the displacement of a spring is proportional to the force that's apply ok, the displacement of a spring is proportional to force that is being applied to the spring, proportional means directly proportional, this is a linear relationship ok that is if L is the length that the spring moves and if W is the weight that is being applied ok,

(Refer Slide Time: 01:09)



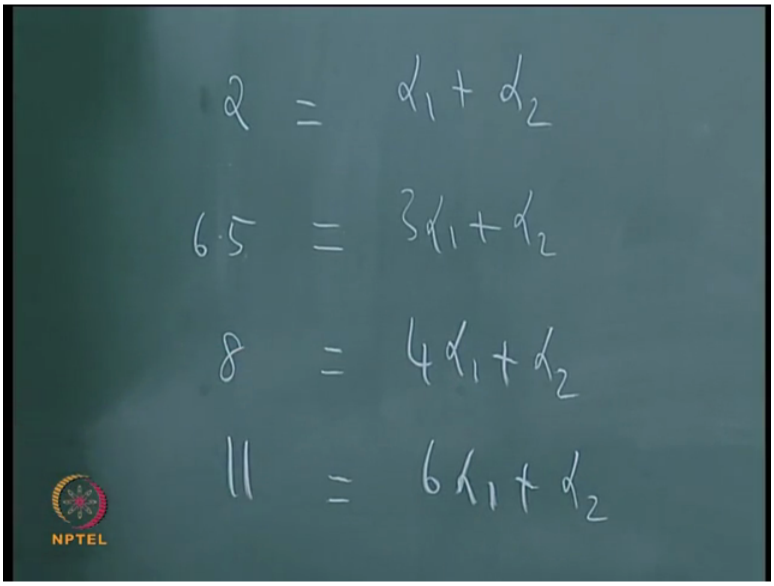
If L is the length that the spring moves if W is the weight that is applied then Hooks law states that L and W are related by this formula, L is a constant times I will call it alpha 1 W plus another constant alpha 2. L is a length W is the weight, so this is Hooks law.

Now in order to determine the numbers alpha 1 alpha 2, see this relationship is known the unknowns are alpha 1 alpha 2, in order to determine the numbers alpha 1 alpha 2 you do experiments in the lab actually attached physical weight to the spring measure the displacement substitute into this and then you can determine alpha 1 alpha 2 for that spring but what usually happens in lab experiments is that we make errors. So in order to avoid errors we make more number of experiments do more number of experiments and then try to see whether we can determine this relationship ok.

Now when you do more number of experiments typically you will get an over determine system of equations, over determine means the number of equations is more than the number of unknowns, typically over determine systems will be typically inconsistent, that is where you have least square solution coming ok. I will give an example, suppose we have the suppose experiment has been done lets say 3 times there are two unknowns it is 3 times and I have the numbers in this tabular column, let us say the displacement is I have weight here displacement here ok.

So lets say weight and here it is the displacement lets say I have 1 gram then 3 grams, 4 grams and lets say we have one more experiment so I have four, what is a displacement? Displacement is lets say 2, 6.5 something like 8 and lets say 11 here, this are the numbers lets say this are in inches whatever. So this is the table that I have from this table for this particular spring I must determine this equation ok. So you substitute this into this, see this is W this is L and so I have the following equations, this gives rise to the following four equations into unknowns alpha 1 alpha 2 right, 6.5 inches or millimeter.

(Refer Slide Time: 04:15)

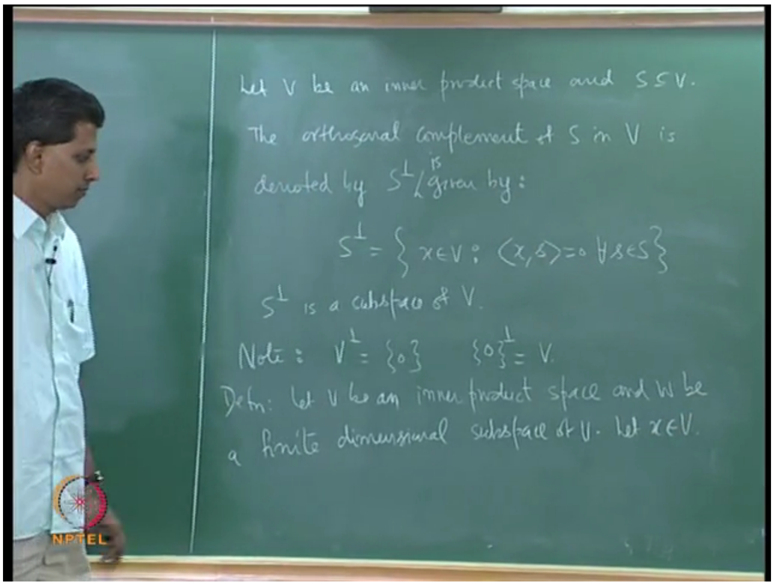


No no it is 6.5, ok then we have the following equations, L on the left I have 2 equals W is 1 alpha 1 plus alpha 2, 6.5 is 3 alpha 1 plus alpha 2, 8 is 4 alpha 1 plus alpha 2 and finally 11 is 6 alpha 1 plus alpha 2 ok I have these equations. Now you can verify that this will be inconsistent for example lets take equation 1 and 4 subtract 1 from the other alpha 2 is gone 3 alpha 1 is 6 so alpha 1 is 2 and alpha 2 is zero I am just taking equations 1 and 3 alpha 1 is 2 alpha 2 is zero but alpha 1 is 2 alpha 2 is 0 does not satisfy the fourth equation so this is inconsistent ok.

You can actually verify by the rank condition, rank of A must be equal to rank of A, the right hand side column vector that will not happen here, this ranks will be different. So we need to but we need to solve this you can't discard any equation here, each equation is as important as the other equation because this are experiments done under the same conditions in the laboratory. So you can't discard any of those so this have to be taken as they are but then they cannot be solved so what one does is to look at the least square solution, that is the solution which minimizes norm A x minus b where the norm is the 2 norm ok.

So I am going to leave the rest of the calculations for you to complete ok, determine this is like A x equal to b pre-multiply by A transpose in this case can you tell me just by inspection whether the least square solution is unique we have given a condition last time. The columns of A are ok, so A transpose A is invertible and so there is a uniquely square solution for this problem ok, you please complete the rest of the problem ok lets get back to this problem of best approximation and look at this notion a little deeper that it deserves but before that lets look at the concept of this orthogonality a little into a little more detail.

(Refer Slide Time: 07:04)



Lets take V as an inner product, so this is my frame work let V be an inner product space and S is a subset of W not necessarily a subspace just a subset then the orthogonal compliment of S in V I am going to define this it is denoted by S perpendicular and this S perpendicular is given by S perpendicular is the set of all vectors x in V such that inner product of x with let us say S this is zero for all S and S, S is not necessarily a subspace just a subset, I collect all vectors that are orthogonal to each vector in S that is S perpendicular.

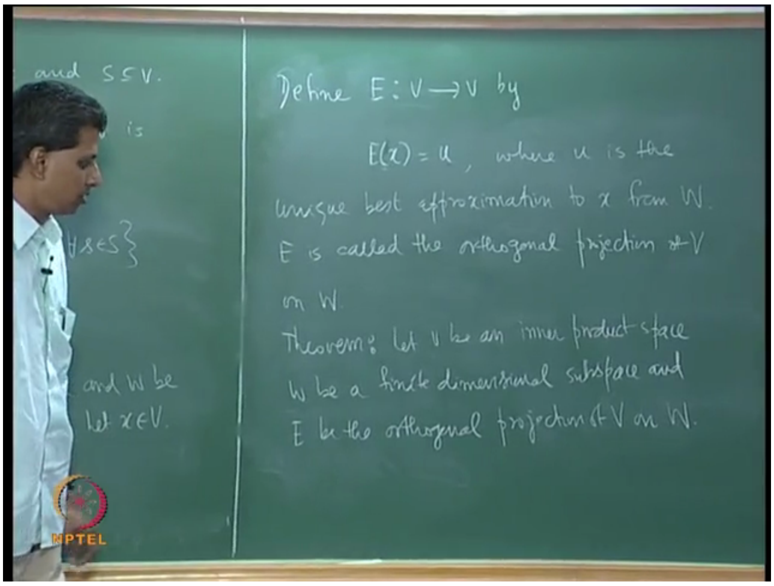
The name orthogonal is clear, why is it orthogonal compliment, that is not clear it should be clear probably by the end of today's lecture, S perpendicular is this whatever be S, S perpendicular is always a subspace ok. The reason is that see the second coordinate second argument is coming from S you can think of this as a linear map null space of a linear map, there is a null space of a linear map ok. So this S perpendicular is a subspace, I am going to leave that as an exercise, S perpendicular is a subspace for one thing zero is perpendicular to any set of vectors so zero belongs to S perpendicular the subspace has to have atleast a zero vector ok.

So zero belongs to S perpendicular is clear, you can infact show that it is a subspace ok lets look at two extreme examples, what is V perpendicular? Set of all vectors orthogonal to every vector in V, single term zero it must be a subspace and what is zero perpendicular? It is V, it is kind of taking the compliment here and then intuitively we might think of taking the double compliment when we take the double compliment we intuitively feel that we should get back the same, that happens in a finite dimension spaces ok.

So these are two extreme examples, lets then look at this notion see this is important in the context of the fact that if you are in a product space and if you have a finite dimensional subspace W then there is unique vector U in W such that for every x in V, given in x in V there is a unique vector U in W such that x minus U is perpendicular to W ok. So one needs to understand W perpendicular, W in our case is a subspace ok, in general one can define S perpendicular in this manner without S being a subspace ok.

So lets then go back to this problem to make the following definition, I go back to the problem of the best approximation. Let V be an inner product space and W be a finite dimensional subspace, let us fix x in V then we know that there is unique U in W such that x minus U is perpendicular to W associated with this X is that U ok,

(Refer Slide Time: 11:54)



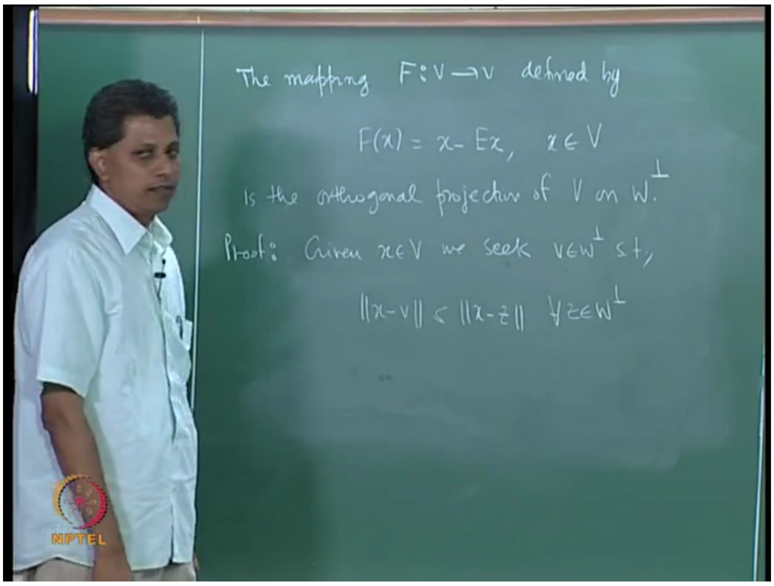
So lets define our mapping I will call it E, E is a mapping on V, E is defined by E of x equals U, what is U? U is to emphasize where U is the unique best approximation to x from W, whenever we define a map we must know that if you define a map if you want to define the image of an element x then you must know that the image in unique only then it is a function.

The fact that this right hand side is unique comes from the result that we proved earlier that if W is a finite dimensional subspace then there is a unique U, ok so U is a unique best approximation to x from W, this U is called the projection of x from V on W E is called see we denoted we call this U as a projection of x on W you are calling E as a orthogonal projection map infact is called the orthogonal projection map so orthogonal projection of V on W. In other words for every x there is a unique U I assign a map call that E then this E is called the orthogonal projection of V on W.

For one thing it is not clear that E is linear will prove a little later that E is linear ok. but before we look at the proof that E is linear and also derives some other properties there is also a relationship see we will show that E is linear then we will establish a relationship between the range of E and the subspace W ok, but before that we look at a complimentary notion, a notion complimentary to the orthogonal projection E and prove the following result.

See I will illustrate all this by means of a numerical example I will do it next after proving this theorem. Let V be an inner product space W be a finite dimensional subspace of V and E be the orthogonal projection of V on W, E is orthogonal projection of V on W.

(Refer Slide Time: 15:25)



The mapping I will call it F from V to V define by you will see that this F is related to E, F from V to V we define by F of x equals x minus E x, it is I minus E, F is I minus E, this is the orthogonal projection of V on what do you expect the subspace to be? W perpendicular, E is the orthogonal projection of V on W, I minus E you will show is the orthogonal projection of V on W perpendicular.

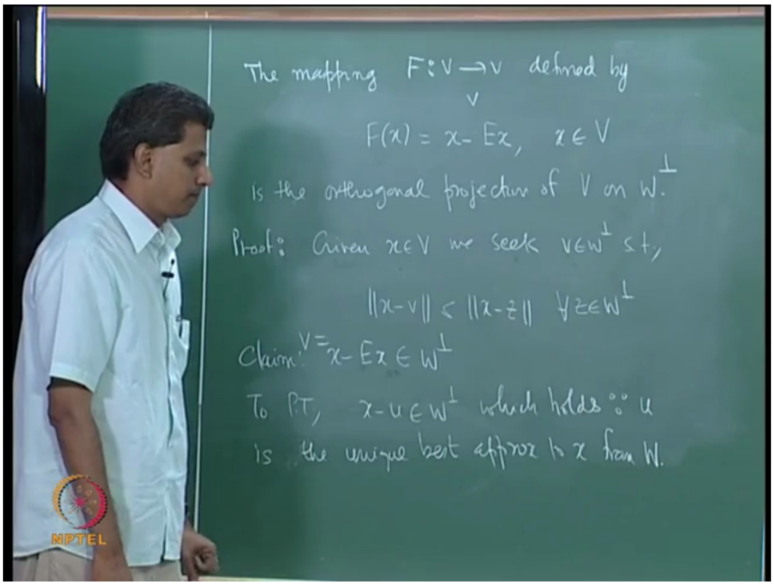
Remember the name orthogonal projection comes because you have projected the vector x orthogonally onto the subspace W ok it is actually an onto map we will prove it to be an onto map but right now we will simply say on (V) E is a projection of V on W I minus E is a projection of V on W perpendicular this are onto maps infact ok lets prove this first, the function F we want to show is orthogonal projection of V on W perpendicular which means I must consider the best approximation problem instead of that is given x in V I must look at lets call it V given x in V I must look at small v.

Now coming from W perpendicular which approximates x, ok so let me write given x in V we are seeking V instead of V (ok) lets call ok V is not a bad idea, seek V in W perpendicular such that see this time we must understand this problem but this is similar to the previous problem. Given x we seek V in W perpendicular such that norm x minus V is less or equal to norm x minus I will use Z if you don't mind Z belongs to W perpendicular. I seek a V in W perpendicular which approximates x from among all those vectors in W perpendicular. So this Z belongs to W perpendicular arbitrary.

Ok now this is the this is the problem for best approximation from W perpendicular I must show that this x minus E x is a best approximation to x coming from W perpendicular isn't it? In order to show that this is the orthogonal projection of V on W perpendicular I must show that this right hand side, in order to show that this E is (ortho) why how do you get the orthogonal projection of V on W? The right hand side vector U is the best approximation to the x that we started with, U belongs to W the problem now is I must show that x minus E x that is why I called that as U I am calling this as V, I am using V for this ok.

So let us call those as V then I am I must show that V given x in V this small v is a best approximation to x from among vectors in W perpendicular, is that clear, then it follows that this F is the orthogonal projection of V on W perpendicular ok, first of all I must verify that this what is the problem there? E x equal to U, U belongs to W, I must verify that x minus E x belongs to W perpendicular ok, that is the first thing.

(Refer Slide Time: 20:21)

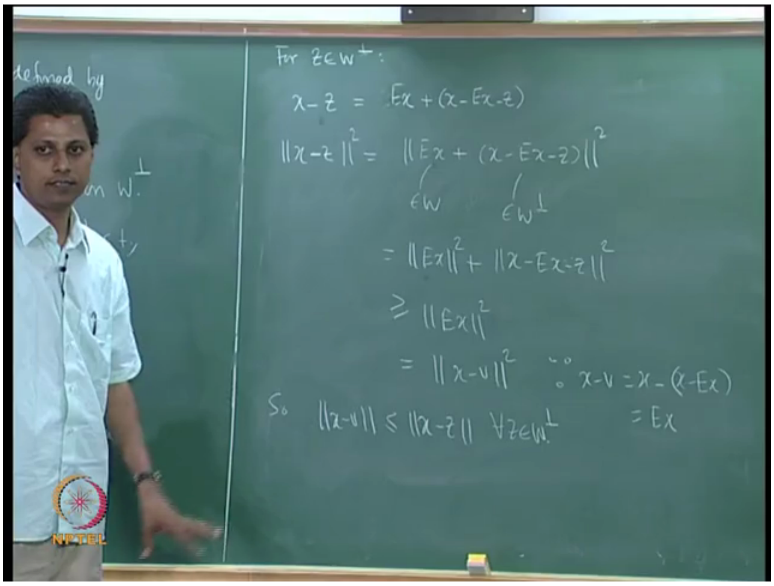


X minus ok I will call it V, V equal to x minus E x belongs to W perpendicular that is we need to prove that x minus what is E x? For x there is a U, once I fix the x there is a U.

I must show that x minus U belongs to W perpendicular where U is the best approximation to x from W, but this is something we have proved before. Infcat U is a best approximation to x if and only if x minus U is perpendicular to W that is x minus U belongs to W perpendicular this holds, which holds since U is the unique best approximation to x from W because W is a finite dimensional subspace this comes from the map E but we know that this happens if and only if x minus U is perpendicular to W that is the same as saying x minus U belongs to W perpendicular.

So this right hand side vector belongs to W perpendicular so it makes sense to talk about this function being orthogonal projection of V on W perpendicular offcourse we need to prove further that this is what do we have to prove? Yeah we must show that among all vectors in W perpendicular this holds ok this needs to be proved offcourse but this belongs to W perpendicular has been established ok. We need to show something like this precisely this I have denoted this by V so we need to show precisely this ok but lets look at that so this is the first part I must show that this holds.

(Refer Slide Time: 22:38)

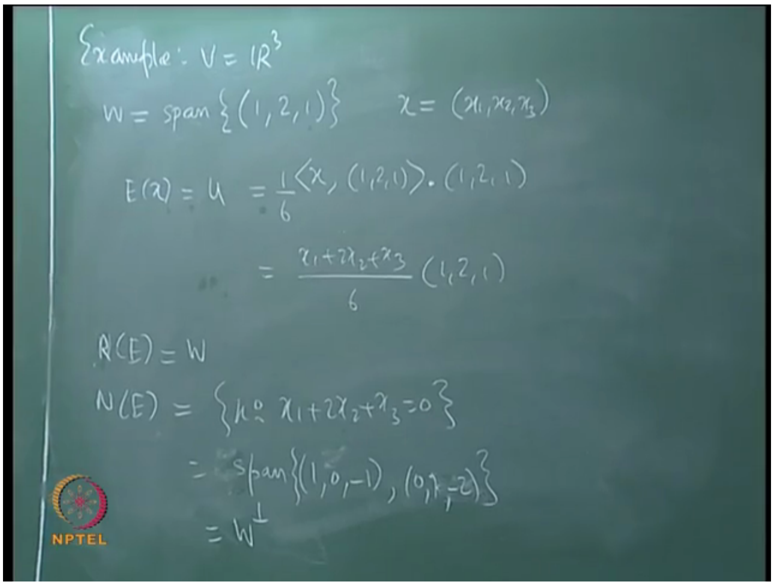


See I will show see we want to show this, what I will do is show that norm x minus Z square is greater than or equal to norm x minus V square ok. I want to show norm x minus Z whole square but look at x minus Z I can write this as x I can write this as E x so I add and subtract E x, E x plus x minus E x minus Z, so I have added and subtracted E x. See this z comes from W perpendicular ok, so for Z in W perpendicular we have this so norm x minus Z square is norm E x plus x minus E x minus Z square look at E x now the first term, see I am going to apply Pythagoras theorem.

Look at the first term first term is E x, E x by definition belongs to W so this belongs to W. look at x minus E x minus Z we have just now shown that x minus E x belongs to W perpendicular Z is coming from W perpendicular, W perpendicular, is a subspace, so this vector belongs to W perpendicular so they are orthogonal so I can apply Pythagoras theorem, so this is norm E x square plus norm of x minus E x minus Z square this is non-negative so this is grater than or equal to norm E x square ok but we want to show that it is greater than or equal to norm x minus V but what is V? V is x minus E x so can I write this as norm x minus V?

Because since x minus V is x minus V is x minus E x which is E x so this greater than or equal to norm x minus E square so it follows that norm x minus V is less and or equal to norm x minus Z for all Z and W perpendicular ok, so this is a first result then if V is a projection of V on W then I minus E is a projection of V on W perpendicular we have not yet proved that E is a linear map but before we prove that E is a linear map I want to look at an example ok and then this example will kind of act as a sandwich between this result and the next result so let us look at the following.

(Refer Slide Time: 25:36)



V has R3 with the usual inner product let me take W to be the subspace span by this vector and for convenience I will take row vectors, what I want to do is determine E ok, W is given to me I want to determine E and derive properties for I minus E ok, W is span of this lets say I want to look at the vector x as 1 minus 1 1 I can determine a general x itself I want to determine the projection so I will take a general x1 x2 x3 I want to determine the projection E for this problem W is given ok, what is the definition of V? E of x is U, where U is the best approximation where U is the projection of x on W and I have a formula for that W id finite dimension W is one dimensional I will simply take an orthonormal vector just divide this vector by the norm ok.

So can you tell me what E x is? E x is U I want a formula for U, you remember the formula for U if U1 U2 etc U n is an orthonormal basis for W then U is summation j equals 1 to n inner product x with U j into U j orthonormal basis I have just an ordinary basis but I there is only one vector and divide by the norm so can you tell me what U is? It is a inner product of x with U j the norm of this is root 6 right so 1 by root 6 it will go with another 1 by root 6 so it is 1 by 6 times x with 1, 2, 1 into 1, 2, 1 ok which is x1 plus 2x 2 plus x3 by 6 times 1, 2, 1 so this is E of x obviously linear ok E is linear.

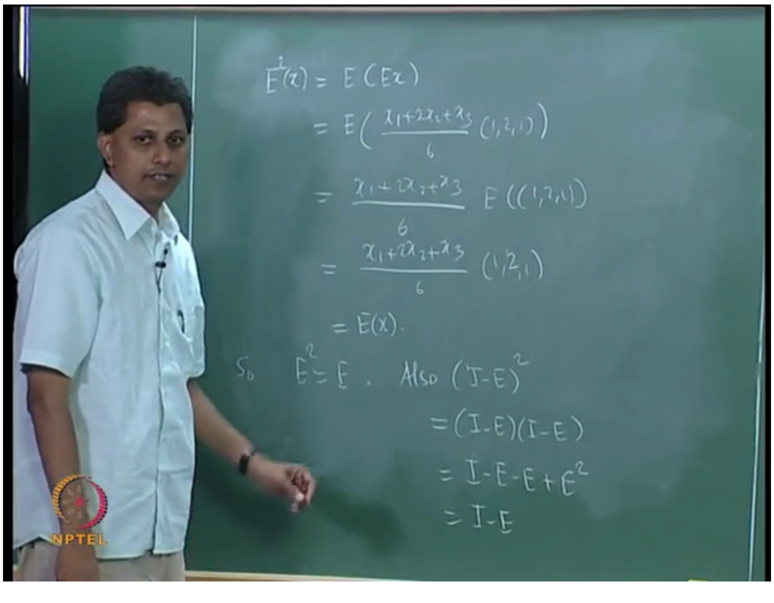
See what I have done is to take an orthonormal basis for W, W is one dimensional so I am just dividing by the norm which is 6 root 6 and this root 6 will come twice because x U j U j ok so this is my E x, what is the null space of E? What is the null space of E? What is the range of E? See in general we cannot talk about range of E because we have not yet shown E is linear but in this example, W, E the right hand side is in W so range of E is W.

What is null space of E? Null space of E is a set of all vectors x as T of x equals zero, set of all vectors x such that this numerator is zero, set of all x such that x1 plus 2x 2 plus x3 is zero this single equation two unknowns so there are two independent solutions so it is a two dimensional subspace, what is that subspace? It is ok, I will just draw it I will just write two independent vectors not necessarily orthogonal I can take 1 minus 1, 1 0 ok 1 0 minus 1 is 1 vector the other one is 0 2 minus sorry 0 1 minus 2 ok.

Is there relationship between these two vectors and this vector? They are orthogonal ok, this is orthogonal to this, this is orthogonal to this, these three together will form a basis for R3 but what is important is, this is equal to W perpendicular, this doesn't happen for a general linear transformation. What doesn't happen?

The range is the subspace W null space is the perpendicular in general this doesn't happen but for orthogonal projection this will be true ok. You can infact write down the formula for the orthogonal projection of R3 on W perpendicular it is just I minus E but I also want you to observe the following. E square is E, can we check that quickly.

(Refer Slide Time: 32:05)



E square is E, E square of x is E of E x that is E of E x is x1 plus 2x 2 plus x3 by 6 times 1, 2, 1 this is a scalar x1 plus 2x 2 plus x3 by 6 what is the image of V? Image of 1, 2, 1 under E?

x1 plus 2x 2 plus x3 by 6 but that 6, 6 by 6 is 1 so that scalar is 1 this is just x1 plus 2x 2 plus x3 by 6 into 1, 1, 1 which is E x that is yes so E square is equal to E, if E square equal to E then I minus E the whole square will be equal to I minus E, I minus E minus E plus E square but E square is E so this two terms get cancelled so that is I minus E. So if E such operators are called idempotent operators.

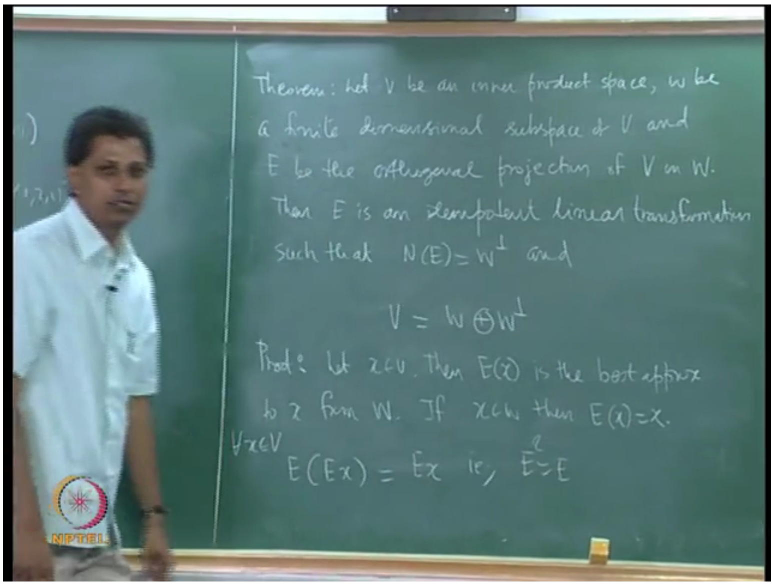
A matrix A is said to be a transformation T is said to be idempotent of T square equals T, so if T is idempotent then I minus T is also an idempotent I will leave the following problem for you to complete. What is a matrix of E under the standard basis? Is there a structure for that matrix? E square equal to E alright, so the matrix will also satisfy property but there is another structure for the matrix you can think of see E is a linear transformation you can think of associating a matrix given a basis.

Now this matrix has a special property this case, in this example I want you to see what it is ok, any guesses?

Student: sir what is the question actually?

What is the question? E is a linear transformation on R3 I can write down the matrix of E relative to let us say that standard basis relative to any basis, lets say the standard basis. There is a structure for E, there is a structure for matrix of E I am asking you to explore, you see what it turns out to be, but is there any guess? Never mind ok so lets now prove, so we have proved certain things for the numerical examples that is offcourse E is linear we have discover that E is idempotent then we have also observe that the null space of E is W perpendicular ok so lets prove this in the general case in a general inner product space.

(Refer Slide Time: 35:45)



This is happening for R3 with a usual inner product, let V be an inner product space W be a finite dimensional subspace and E denote the orthogonal projection of V on W then E is an idempotent linear transformation. Idempotent means E square equal to E what we proved in this example, E is an idempotent linear transformation such that null space of E is W perpendicular range of E is W is obvious by the definition null space of E is W perpendicular and what is more important is that, remember V need not be finite dimensional W is finite dimensional ok, more important is V can be written as W plus W perpendicular where this plus is the direct sum which means that W intersection W perpendicular is single term zero.

Intersection of two subspaces offcourse W plus W perpendicular is equal to V that is any element x in V can be written as a sum of two vectors one from W one from W perpendicular this will be done in a unique way because intersection is single term zero. Now look at go back to this definition, this perpendicular was called the orthogonal compliment of W, this is the reason, if W is finite dimensional then there is always an orthogonal compliment of W, orthogonal complimentary subspace. V maybe be an infinite dimensional inner product space but it can always be written as a direct sum of W and W perpendicular whenever W is a finite dimensional subspace ok.

If you demand W to be an infinite dimensional subspace the answer is no in general but whenever W is a finite dimensional subspace you can write V as a direct sum of this two subspaces ok. This should also remain this equation should also remind you of something that we have studied before not Rank Nullity dimension theorem, pardon

Student: (())(39:05)

Professor: No I am asking you this question, what does this equation remind you of? See dimensional be a problem, V is infinite dimensional possibly infinite see there is no, we have not restricted our attention to V being finite dimensional, so we don't know whether the dimensions will add up. If V is finite dimensional the dimensions will add up but if V is infinite dimensional W is infinite dimensional then this things don't make sense but I am asking you does this remind you of something we have studied before? Forget about the perpendicular, let us say I write V as W1 direct sum W2, have we done something?

If we are given such direct sum decomposition see if V is written as W1 plus W2 etc W k then we have shown that there exists k linear maps k idempotent k linear maps E1 E2 etc E k such that E1 plus E2 etc E k is identity then E i E j is zero if I is not equal to j that we have seen and we have also seen a connection between this and I gain values of a linear transformation ok, but forget about I gain values of a linear transformation what we have seen is a converse of what is happening now is we have defined a map and we have defined a map on a subspace W finite dimension subspace W and through this map we are getting direct sum decomposition.

What we have done earlier is a converse, given a direct sum decomposition we have constructed maps which are precisely idempotent ok they do not correspond to orthogonal direct sum decomposition but is a usual direct sum decomposition the difference between an orthogonal direct sum decomposition and usual decomposition will be clear a little later that is got to do with my question that I asked earlier, what is a structure of V, when you write down the matrix of V? Ok this are intimately connected but I just want to remind you that this converse question we have seen before it is not an orthogonal projection, it is just an ordinary projection.

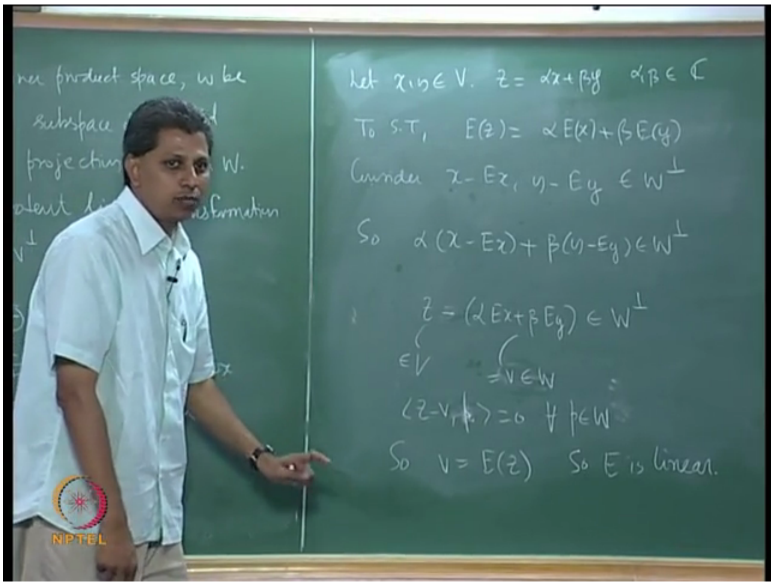
So the question is what is a difference between orthogonal projection and an ordinary projection? Just think it over, orthogonal projection see we have seen that ordinary (pro) I remember having used the word projection also earlier E square equal to E ok E square equal to E happens here also but there is something more to E again that question, what s special about the structure of E? The matrix of the linear transformation E that is the extra thing which connects you to the perpendicular ok.

So we have seen that we have studied the converse question, in a little more general sense, if V is W1 plus W2 etc W k then we have constructed idempotent maps E1 etc E k which have the property that their sum is equal to (idempotent) ok. Let just remember that see you learn a new concept you need to relate the new concept with what you have studied earlier. So I am just reminding you that the converse question has been studied not in this context of perpendicular but little more general sum of direct sum of several subspaces ok. I will just maybe prove that it is idempotent ok.

So proof, I want to show E is idempotent ok but ok so let me emphasize let x belongs to V then E x is the best approximation to x from W that is e x belongs to W ok, if x belongs to W what is E x? Yeah what is E x? If x belongs to W then x is equal to E, what is E x? E x is U, U is x so E x is x if x belongs to W then E x is x that itself is the best approximation. So E acts like identity on W ok. What is E of E of x then? For every x in W, E of E of x that is the E square x ok but look at E x, E x belongs to W, E of something in W must be itself. T x that is we have shown this for each x in V, so E square is E ok, so E is idempotent that is really straight forward E is idempotent, is that clear ok.

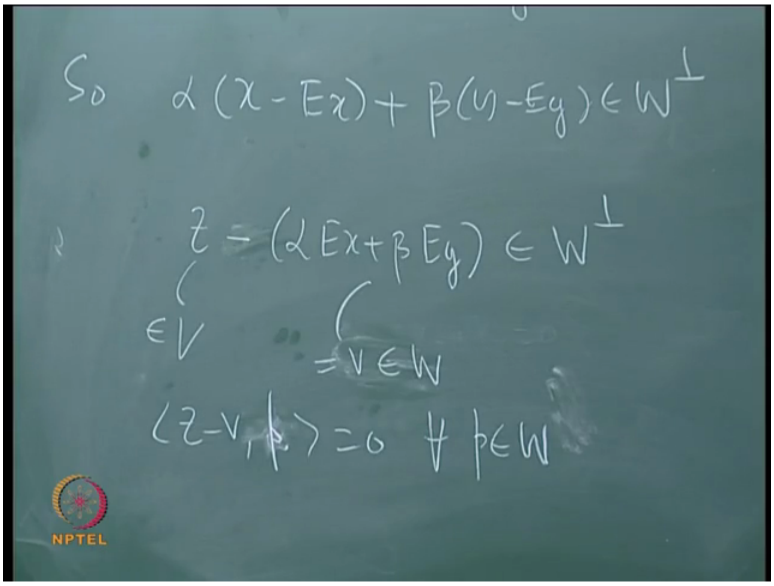
We need to show E is linear let me do.

(Refer Slide Time: 45:24)



We need to show E is linear let x y belong to V and lets call Z as alpha x plus beta y, alpha beta coming from the field I must show that E of Z equals alpha E x plus beta E y it would then follow that E is linear, this is a definition, E of alpha x plus beta y is alpha E x plus beta E y if I show this then E is linear ok. Lets consider x minus E x and y minus E y the function the mapping F from V to V F of x equals to x minus E x that is an orthogonal projection of V on W perpendicular these two elements are in W perpendicular, this two vectors belong to W perpendicular ok, x minus U really, x minus U, y minus V if you want, they must be orthogonal to W.

(Refer Slide Time: 46:45)



So this are vectors in W perpendicular, W perpendicular is a subspace so a linear combination, alpha times x minus E x plus beta times y minus E y that must also be in W perpendicular what is this? This is alpha x plus beta y that is Z, alpha plus beta y minus alpha E x plus beta E y, this belongs to W perpendicular ok. Now this Z comes from V, E x E y belong to W, W is a subspace this combination is in W, so this comes from W but this difference Z minus lets call it V ok equal to V that belongs to W ok.

Now the difference Z minus V belongs to W perpendicular that is ok, what does it mean? Z minus V W sorry I will use some other y maybe no not y, y is already there, p, Z minus V p equal to zero for every p in W perpendicular, Z minus V is orthogonal to every p in W, is this the same as saying that V is equal to E of Z, is this the same as saying? If x minus U is orthogonal to W then U is a image of x under E, if x minus U is perpendicular to W, U is the image of x under E, if x minus sorry, if Z minus V is perpendicular to W, it means V is the image of Z under E by definition ok we are through.

V is alpha E x plus beta E y on the one hand E of Z, Z is alpha x plus beta y, so E is linear, is that clear? What is E of x? E of x equal to U where U satisfy the property that x minus U is perpendicular to W I am writing E of Z equals V because Z minus U is, Z minus V is perpendicular to W ok, the second last part that V is a direct sum I will prove it in the next class.