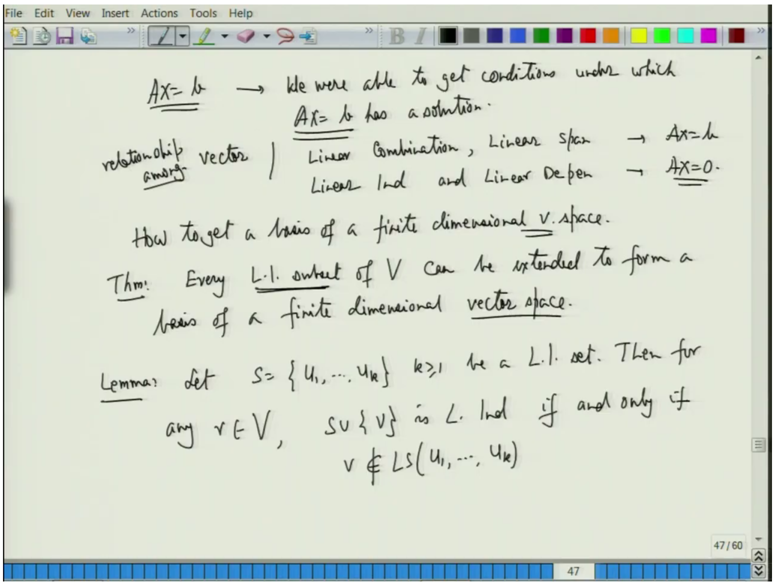
Linear Algebra
Prof. Arbind Kumar Lal
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture – 50
Recapitulate ideas on Inner Product Spaces

So, I want to recapitulate everything in this lecture as far as inner product space is concerned, this is a very crucial step, very crucial time for us to understand what exactly we have done and proceed fine.

(Refer Slide Time: 00:33)



So, what we had done was we have started with, if I look at the class where I started with trying to solve the system AX is equal to b that was our starting point fine; for that we define matrix multiplication invertibility and so on. And, then using whatever method we had elementary matrices, rref and so on; we were able to solve. So, we were able to get conditions under which AX is equal to b has a solution fine.

This is one thing we learnt using rref and then there are some equivalent conditions when A was a square matrix fine. We also learned that solving a system relates to the idea of coordinates or the vectors alright. So, solving also relates to relationship between vectors, relationship among vectors alright. And, this let us to a study what is called linear combination, linear span related with AX is equal to b, then there was this, what is called linear independence and linear dependence.

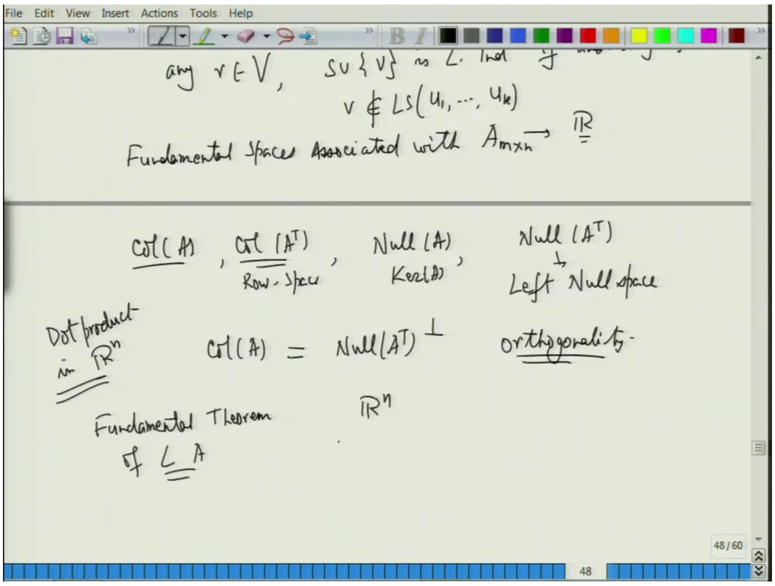
This was related with the unique solution of the system AX is equal to 0 alright fine. So, we went to vector spaces and the most important theorem that we had was related with basis alright. So, how to get a basis? How to get a basis of a linearly of a finite dimensional vector space; of a finite dimensional vector space alright?

So, that was the idea. So, what we did? We had a theorem which says theorem was there, that every linearly independent subset of V can be extended to form a basis of a finite dimensional vector space alright. So, this was important that we did this part fine.

Now, to prove this one of the crucial result was crucial result that is lemma we can put our theorem was that, let S be a linearly. So, every linearly independent; so, we are talking about. So, let start with u k, k greater than or equal to 1 be a linearly independent set. Then for any v belonging to V, S union V is linearly independent; if and only if V does not belong to linear span of u 1 to u k; this is what we learnt alright fine. This is a very crucial result fine.

And, at this stage from here we could go to the finite dimensional to get a basis. And, we are able to get finite basis because see that number of element in the generating set is say 100 or 1000 or million or trillion whatever it is, but it is a finite number. And, hence after some time the process has to end that was the idea alright. And, we related this with rref and so; from there we related different ideas.

(Refer Slide Time: 04:45)



And, then we came to what was called the fundamental spaces, fundamental spaces associated with a matrix A which is m cross m. We will be bother only about coming elements coming from real numbers fine.

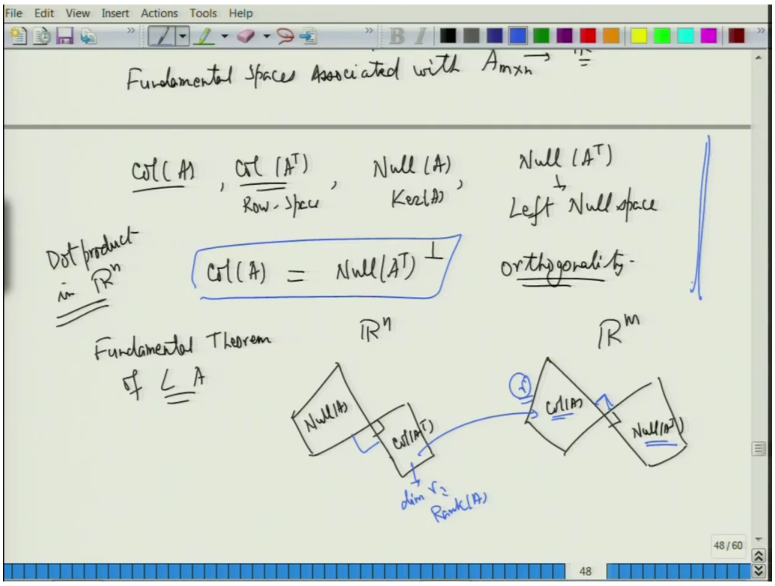
So, here we have notion of what is called the column space of A, column space of A transpose which was same as the row space, column space row space. And, then there was this null space or the kernel of A and then there was this null space of A transpose which was called the left null space fine; we had this.

When we were trying to study this, I did not define any inner product; we took just about we talked about what is called the dot product, dot product is R n; we looked at dot product in R

n. Just the generalization of component wise multiplication and then adding as it happens for R 2 and R 3.

And, there we showed that column space of A and null space of A transpose they are related. And what is the relationship? They are orthogonal to each other. So, the notion of orthogonality was there fine and then we related.

(Refer Slide Time: 06:31)



So, the fundamental theorem of linear algebra was there of linear algebra which said that, I can write R n as well as R m as orthogonal things alright. So, R n if I want to look at; so, null space has so, A is m cross n. So, null space is there A here, null space of A transpose will come here. Then there will be column space and there will be column space of A transpose, this is what it says here that column space of A is perpendicular to null space of A transpose.

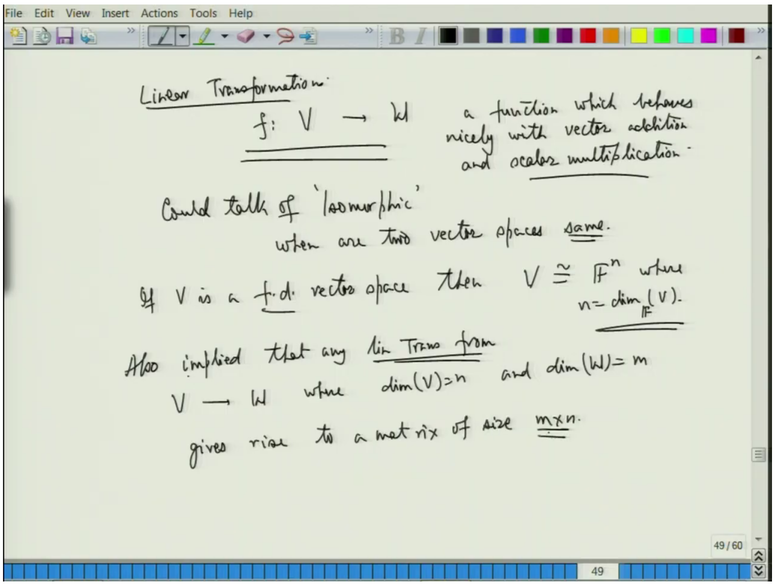
So, this is null space of A transpose column space is here, they are perpendicular to each other alright fine. There is a perpendicularity. Similarly, if I replace A by A transpose, this is what I will get; they are perpendicular here. I did not prove that there is a one to one relationship here because I did not had time and it is not and it is out of syllabus for us.

But, at the same time what we know is that dimension of this was r the rank of A, where r is rank of A. And, this is also dimension is r itself alright; so, we related these ideas and got something's fine.

But, what we realized at this stage was that at no stage other than this part when I look at the fundament spaces, we did not have any geometric coming into play. There was no notion of angle between two vectors alright.

And what we are seeing here is that, when I looked at the fundamental spaces somehow this orthogonality has played a role in trying to understand my column space, row space and so on alright. So, in the next chapter in place of going to the inner product we went first to alright. So, we went first to what I called linear transformations.

(Refer Slide Time: 08:21)

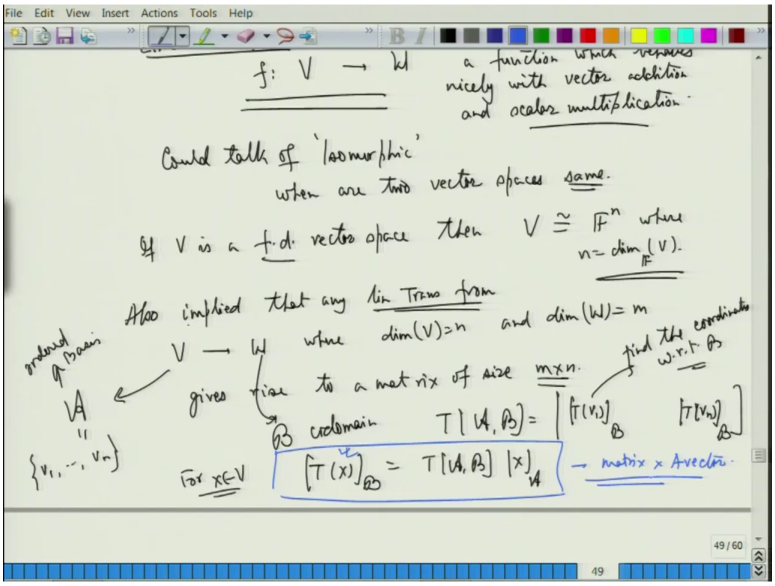


So, we went to linear transformations. And what was the idea of linear transformations? It was just a function f from a vector space V to W, it is just a function f from V to be a function which behaves nicely with vector addition and a scalar multiplication fine. So, what we use the word was that they commute, the operation of taking addition and then the image and first image and then the addition and things like that they commute alright.

So, and we saw that these were very very important in the sense that we could define alright, that when our two vector spaces isomorphic. So, could talk of could talk of isomorphic word or talk of when are two vector spaces same alright, fine. And, then obtain a lot of results there, we also saw that if V is a finite dimensional vector space then V is isomorphic to F n, where n was equal to dimension of V fine over F. So, dimension of V over F alright; this is what we saw fine.

This also helped us, also implied that any linear transformation any linear transformation from V to W, where dimension of V is n and dimension of W is m implied that any linear transformation gives rise to a matrix of size m cross n alright. And how do you do it?

(Refer Slide Time: 11:13)

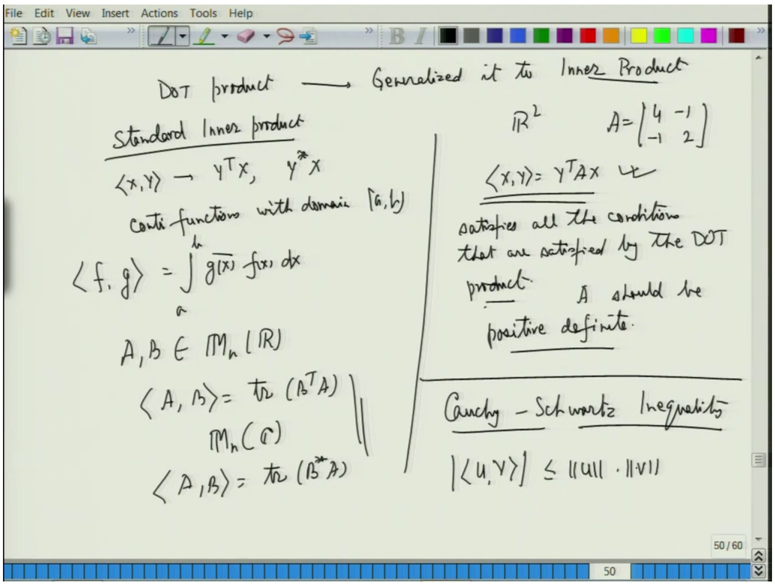


So, there was a notion of what is called an ordered basis here in the domain, similarly there was an ordered basis here in the co-domain fine. And, then we have the definition of T of A, B fine and the idea of T A, B was so, A has vectors say V 1 to V n. So, what you do? Compute T of V 1 till T of V n. So, compute this maps then evaluate them with respect to B, evaluate them means find the coordinates of these vectors with respect to W alright.

So, find the coordinates coordinate with respect to B alright and that is the one that gave us the answer for us and we saw that any T of X for X belonging to V, T of X is an element of W. So, I can write it with respect to B. So, this is nothing, but T of A B times X which is an element of A alright. So, we could do this part. So, what we are saying is that any linear transformation is a matrix times the vector alright.

So, matrix into a vector this is what we have alright. So, we saw that part. Now, once we have understood that everything can be in terms of F n itself we went back to our dot product or the inner product whatever we want say.

(Refer Slide Time: 12:52)



So, we have a notion of dot product and then we generalized it to; generalized it to the word what is called inner product fine. For us most of the times we are looking at what are called the standard inner product inner product. And what were they? They were basically coming as if I want to look at inner product of X and Y, it was either Y transpose X, if it was real Y star X if it is complex fine.

Similarly, if I am coming from say continuous functions, continuous functions with domain a, b; then I can look at a to b g X bar f X d X, here I am looking at f comma g. So, bar for the complex fine, if I have got two matrices A and B belonging to M n of R. Then, I have the notion of inner product of A, B as trace of B transpose A.

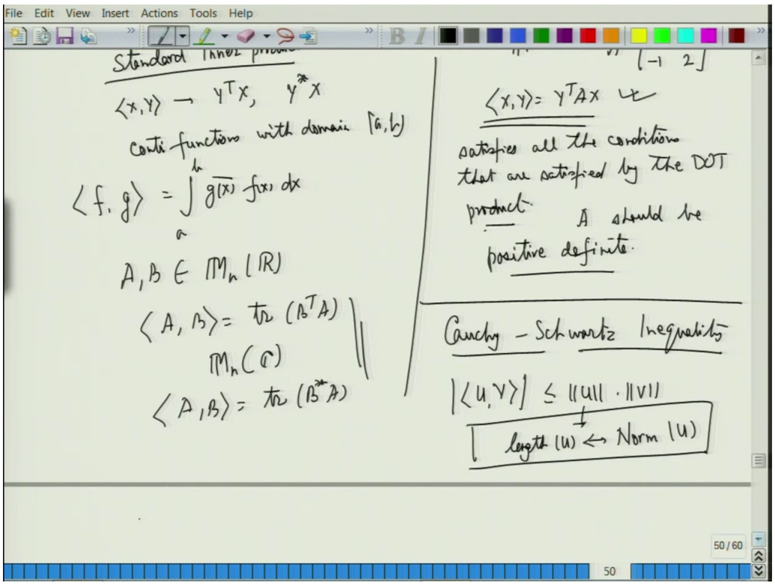
If they are coming from M n of complex number then a standard inner product was trace of B a star A. And basically they were nothing, but the usual standard inner products and so on alright. So, this is what the study inner products where. The generalization was with the idea that something more can be done and in that for R 2, we gave an example of a matrix A which was something like 4 minus 1 minus 1 2; I do not remember exactly.

But, this that I can define my inner product as Y transpose AX everything is nice everything is satisfied, but the angles and other thing they change alright, but it satisfies. So, this definition alright satisfies all the conditions that are satisfied by the dot product alright ok. So, the only thing we wanted was that this matrix A should be positive definite.

So, that a few exercises also with this with respect to this and we calculated certain things here fine. The next idea that we had was once we have this dot product and so on, we have the notion of what is called Cauchy Schwartz inequality alright.

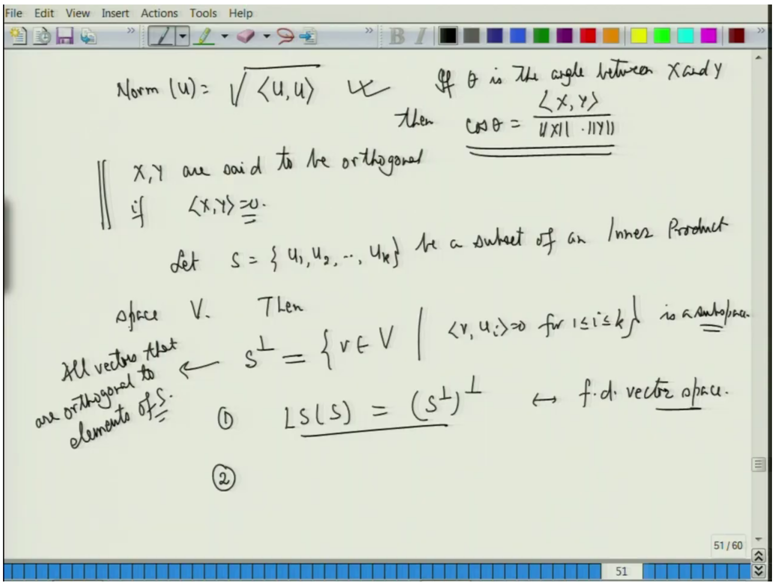
So, what we saw was that it is not only the a standard inner product, but any inner product that I am looking at satisfies this Cauchy Schwartz inequality. And what is that inequality? That look at u and u, two vectors, look at the absolute value of that. So, this is less than equal to length of u into length of v alright.

(Refer Slide Time: 16:28)



So, I am using the word length, but basically we talk of; so, length of u, but we talk in terms of what is called norm of u; this is what we talk of alright. At every stage it is called the norm and what is the definition of norm?

(Refer Slide Time: 16:42)



So, norm of u is basically a square root of inner product between u and u itself, is nothing to do with a standard inner product. This you can do it for anything and therefore, this makes sense fine. So, we had this norm with u and the Cauchy Schwartz inequality. So, using Cauchy Schwartz inequality you could define.

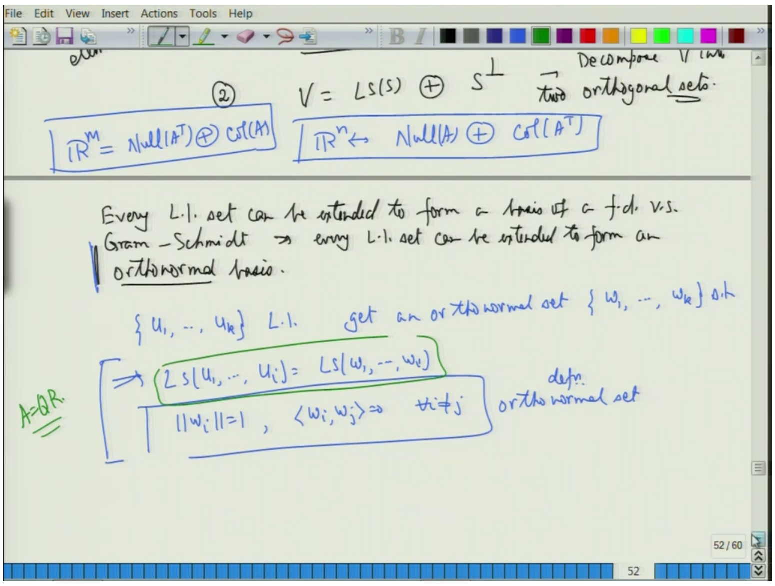
So, if theta is the angle between X and Y then we define cos theta is equal to u is equal to inner product of X Y divided by norm of X into norm of Y; I could define it alright. We did not do much here, what we did was basically said that when something is orthogonal and when something is not orthogonal.

So, we looked at X and Y are said to be orthogonal if inner product between X and Y is 0. We did only with real here, but in general you can talk of complex also. And, you can have a similar condition and things like that; I did not spend time on that part. I wanted you to understand this fine, then there was a notion of what was called orthogonal complement and orthogonal.

So, let me write that. So, let S be a subset of say you have u 1, u 2, u k be a subset of an inner product space, the space V. Then we talked of the notion of what is called S perp, S perp was all v belonging to V such that inner product of v with u i is 0 for 1 less than equal to i less than equal to k. So, we are looking at all vectors that are orthogonal to elements of S alright.

And what we saw was that this is a subspace fine, this is a subspace fine. Some things we concluded here that if I want to look at LS of S, then this is same as S perp whole perp. And, this was only for finite dimensional vector space, we did only for finite dimensional vector space. Otherwise there is some strict in containment in certain examples, that I did not do alright. But, this was important for me. 2, what we did was at later stage was that every linearly independent set can be extended to form a basis.

(Refer Slide Time: 19:57)



Let me write that, every linearly independent set can be extended to form a basis of a finite dimensional vector space. And, this and the Gram-Schmidt's process, Gram-Schmidt implies every linearly independent set can be extended to form an orthonormal basis. I will define orthonormal bases just after some time, but I would like to write here that. So, I can write because of this orthonormal that I can create it I can write V as LS of S direct sum S perp that I can decompose.

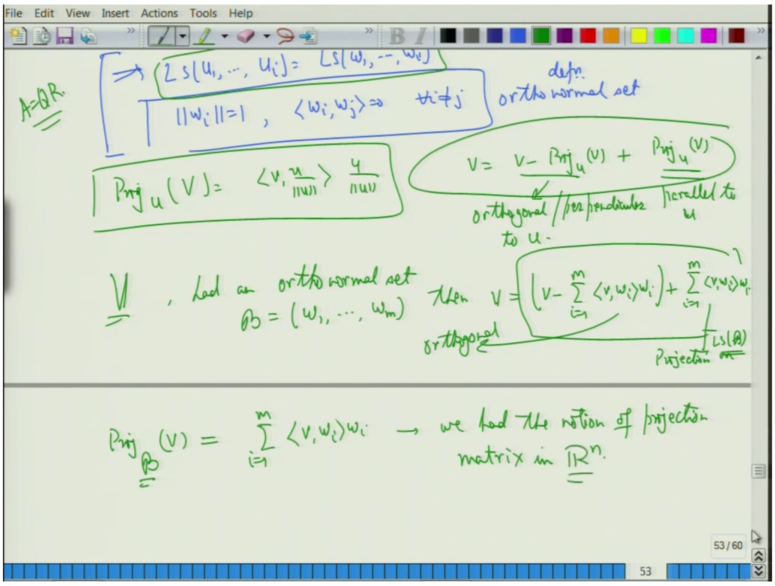
So, decompose V into two orthogonal sets alright. The way that we used to do for; so, that we had already done for R n was coming from; recall it was coming from null space of A and it was orthogonal to column space of A transpose. And, R m was orthogonal to null space of A transpose fine. So, and direct sum m components and column space of A fine.

So, we already had seen these decompositions when we looked at fine fundamental spaces or the fundamental theorem of linear algebra. We have generalized this idea to any vector space alright, the generalization had been done and that was done because of this Gram-Schmidt process.

So, let us go to that part and see what exactly we learned in Gram-Schmidt process fine. So, there the idea was that I have a set u 1 to u k alright, linearly independent get a an orthonormal set normal set w 1 to w k; such that linear span of u 1 to u i is same as linear span of w 1 to w i. Length of w i is 1 and dot product between w i and w j is 0; for all i not equal to j.

This is what the definition was as far as orthonormality is concerned, orthonormals see there are the definition here fine. And, that was the Gram-Schmidt process that gave me this alright. So, this process also gave me because of this restriction that at each stage linear span is maintained, I had A equal to QR; that is the one that gave me QR, that I can sequentially create my columns of Q, is that ok? So, that was one important idea that we got fine.

(Refer Slide Time: 23:26)



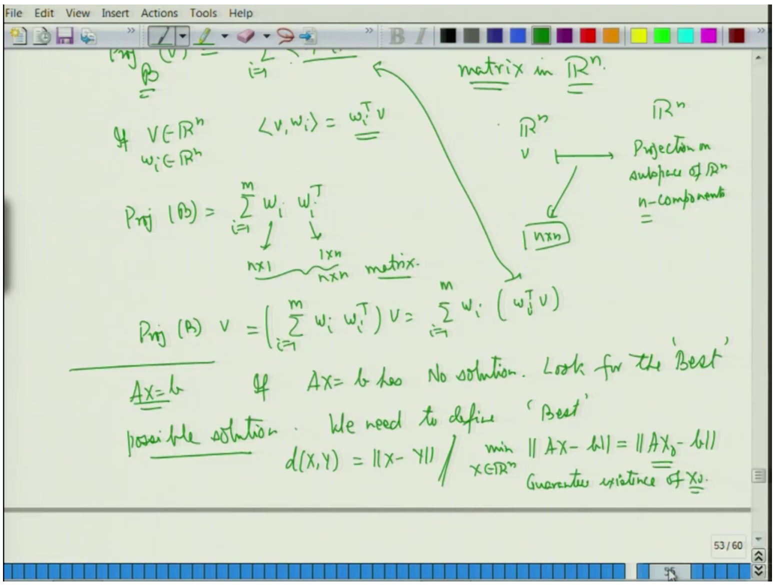
We also had because of this part orthogonality, that I am talking about that given any vector V I have the notion of projection of V on u as V u divided by length of u into u upon length of u; this is what I had projection of V on u I could talk of this. And therefore, decompose V as alright I am looking at V minus projection of V on u plus projection of V on u.

So, this was composition parallel to parallel to u and this was orthogonal or perpendicular to u fine. So, that was for one vector, then we generalize this idea to any set. So, then what we did was the next thing we did was that next was that given any vector space V, I had orthonormal set B ordered bases you can say in some sense as w 1 till w m alright; orthonormal set this.

Then we decomposed any vector V as V minus summation v w i w i, i equal to 1 to m plus i is equal to 1 to m v w i w i alright. This decomposition is similar to this decomposition, they are same in some sense you can see here. There is V minus something, there is V minus projection. So, these are the projection that I am looking at here. So, these were projection, projection on LS of this B whatever, I have fine.

These are projection on this and this is the orthogonal part. So, this is what you have to understand that I had all these notions with me and from here from the projection part. So, from the projection of V on B in some sense alright the notation is not proper, it should be linear span and so on; which is same as v comma w i w i; i going from 1 to m. We had the notion of projection matrix in R n. So, what was the idea of projection matrix in R n? So, let us go back again and try to understand.

(Refer Slide Time: 26:20)



So, when I am looking at a projection; so, if I am taking. So, if V belongs to R n then their standard one it says that this is equal to w i transpose v alright fine and I want a matrix fine, using this and this idea I want to get a matrix. So, since it is a matrix and I have to pick an element from R n; so, I want a matrix, I want to pick up an element from R n which is V. I want to send it to the projection on subspace of R n alright.

Since, this is a projection on a subspace of R n, it will also have n components fine. So, I have looking at R n itself it as n components. So, this matrix that I need is n cross n matrix and w i has also w i also has n components. So, I use this idea to look at the matrix A I defined projection matrix on B as summation i going from 1 to m w i w i transpose.

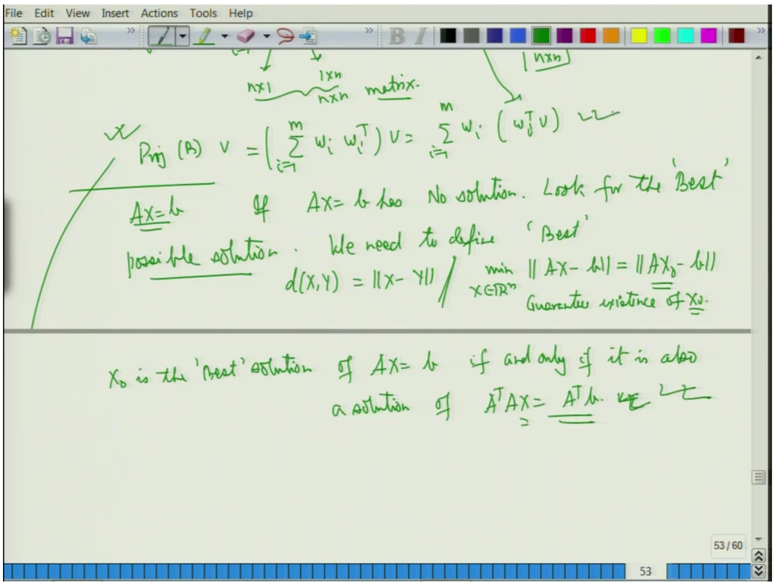
So, now if I look at w i is n cross 1 w i transpose is 1 cross n. So, I do get back an n cross n matrix fine. And if I look at projection of this times a vector V then this is nothing, but i is equal to 1 to m w i w i transpose of V which by linearity gives me i is equal to 1 to m w i w i transpose V which is what we have here alright fine. So, we could get that. So, this projection matrix was important.

So, now the idea was; so, let me sum these ideas that are small ideas that I had. So, the ideas were that I solve the system AX is equal to b, I had solution if AX is going to b has no solution, no solution look for the best possible solution. So, we have to define the difference; so, we had to we need to define this word best alright.

So, for us the best for us was looking at to define the best we looked at what is called the distance between X and Y, we define the word distance between X and Y as norm of X minus Y. And therefore, we looked at AX minus b minimum over X belonging to R n alright. So, we looked at this minimum of this. And, we said that suppose this is equal to AX naught minus b.

Now, why did I say that suppose; because I have to guarantee that there exists an X naught. So, we have to guarantee existence of X naught. So, here I could do it basically because I am an R n I have a standard in a product, this distance function is a continuous function; closed interval everything is nice, I could do it alright. In general setup I may not be able to do things, but at least in this setup I could do it and then what we showed was alright.

(Refer Slide Time: 29:53)



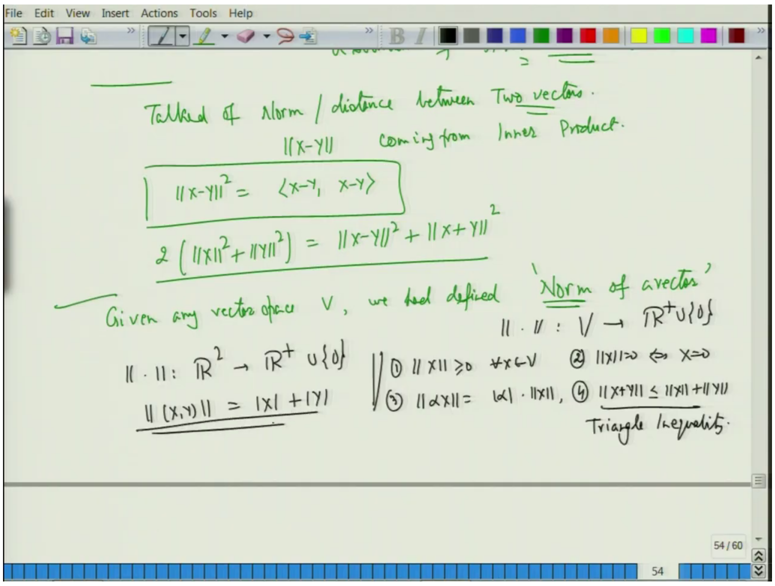
So, what we showed was from here that X naught is the best solution or the nearer solution of AX is equal to b; if and only if it is also a solution of A transpose A X is equal to A transpose b alright. So, we related the idea of best or the distance in terms of column space of A and null space of A transpose perp or null space you can say and then got this is all the proof of this alright.

So, this is one way of getting the best solution X naught. The other way is using this projection because projection is also nothing, but the best solution; because you are looking at the best nearest possible. In the sense that since it is a projection; so, it will be a perpendicular and perpendicular it is supposed to be the nearest as far as the plane is concerned alright. I have a plane I have any point there. So, the best the minimum distance from any point will be the perpendicular distance.

So, here I am looking at a perpendicular distance, here I am looking at the projection in two different languages, but the answers are going to be the same. So, that is what is more important for us; that we can do such a thing alright. So, there are different ways of finding projection this projection into different parts, three methods. One was using solving the system A transpose AX is equal to A transpose B. The other was using projection matrix.

And, the last one that I gave was alright was based on looking at the whole basis alright. I got a basis of the number of the null space or the column space and then its corresponding complement, solve the system because it is the full basis and then divided two parts for each one of them; I got the answers alright.

(Refer Slide Time: 31:47)



So, I could do these things basically because of inner product. So, I could do because of; so, talked of norm or distance between two vectors alright, norm of X minus Y and it was coming from inner product alright. So, if you recall we at each stage we looked at some sort of a square like this which was X minus Y comma X minus Y and did lot of things alright.

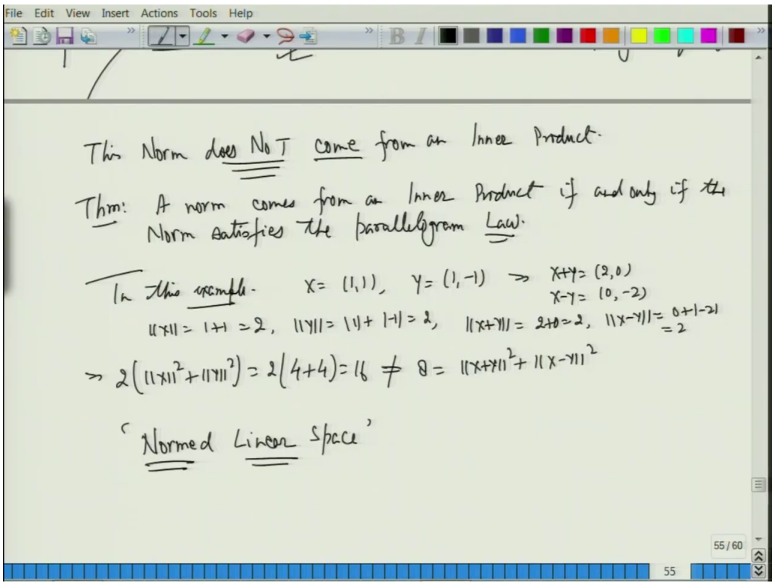
In this process we also proved generalization of Pythagoras theorem, there are other theorems also I did not prove. But, one thing that I proved was important was what is called the parallelogram law. We proved that parallelogram law rules that is if I look at the square of the diagonals add them up, then it is twice the square of the sizes different size sides alright; this is what it was.

So, let me try to use this idea to give you something more. So, I had defined the notion of; so, given any vector space V, given any vector space V we had defined norm of a vector. If you remember I defined the word norm of a vector fine and it has the property that. So, this is a function from the vector space. So, this is a function from vector space to R plus union 0 and it has certain property that I have done.

So, let me take an example. So, I want to define a function here from R 2 to R plus union 0 alright by any vector X, Y here length of this or the norm of this I define it to be equal to this plus this alright, I define it this way fine. Now, what I would like you to see that this has all the properties of that norm of X was supposed to be greater than equal to 0 for all X belonging to V that was 1.

2 norm of X was 0 if and only if X was 0, 3 norm of alpha X was equal to absolute value of alpha into norm of X. And, the fourth-one which was norm of X minus Y was less than equal to or norm of X plus Y, I think wrote as norm of X plus norm of Y; the triangle inequality. So, I had all these things with me fine.

(Refer Slide Time: 34:49)



So, I would like you to check that all these things are satisfied. It turns out that this inner this norm, this norm does not come from an inner product. So, it does not come from an inner product and the proof of this is that. So, there is this theorem which is out of syllabus for us, even this is out of syllabus for us that we are not bothered.

So, there is a theorem which says that a norm comes from an inner product if and only if; if and only if the norm satisfies the parallelogram law alright. So, we need that this norm that you have defined need to satisfy this parallelogram law alright, otherwise there is a problem. So, let us take an example and see here that why does it fail in this example that, I have taken here alright.

So, let so, in this example take X as the vector 1, 1; Y as 1 minus 1. This implies that X plus Y is 2, 0; X minus Y is 0 comma minus 2 fine. I hope X minus Y is 0 minus 2. So, what is norm of X? So, norm of X is 1 plus 1 which is 2, norm of Y is again 1 plus absolute value of this which is 2, norm of X plus Y is 2 plus 0 which is 2, norm of X minus Y is 0 plus absolute value of minus 2 which is 2.

So, therefore, 2 times norm of X square plus norm of Y square is 2 times norm of X is 2; so, it is 4 plus 4 which is 16 which is not equal to look at this part, norm of this is 2 squares of 4 plus 4 is 8. So, this is not equal to 8 which is norm of X plus Y whole square plus norm of X minus Y square alright because that is 4. So, we have got a set here or got a norm here which does not come from an inner product.

So, the whole theory related to things coming from what is called a normed linear space; normed linear space. So, what do I mean? Normal linear space means there has to be a norm and linear space means vector space. So, the whole theory related with non-linear space which is not an inner product which does not have an inner product is a thing that needs to be looked at for a higher mathematics and has to be understood well alright.

So, I end the this lecture here itself and also end the chapter on inner product spaces. In the next class, we looked at what are called eigen values and eigenvectors and then look at things fine. So, that is all.

Thank you.