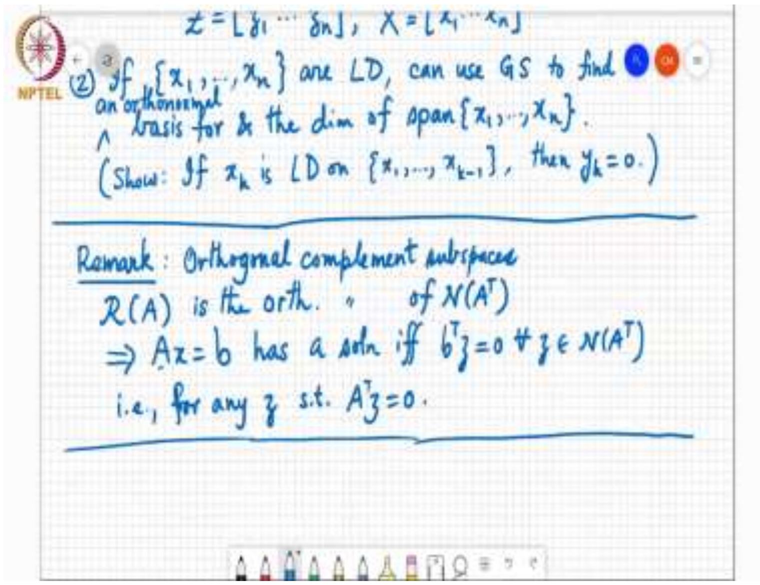
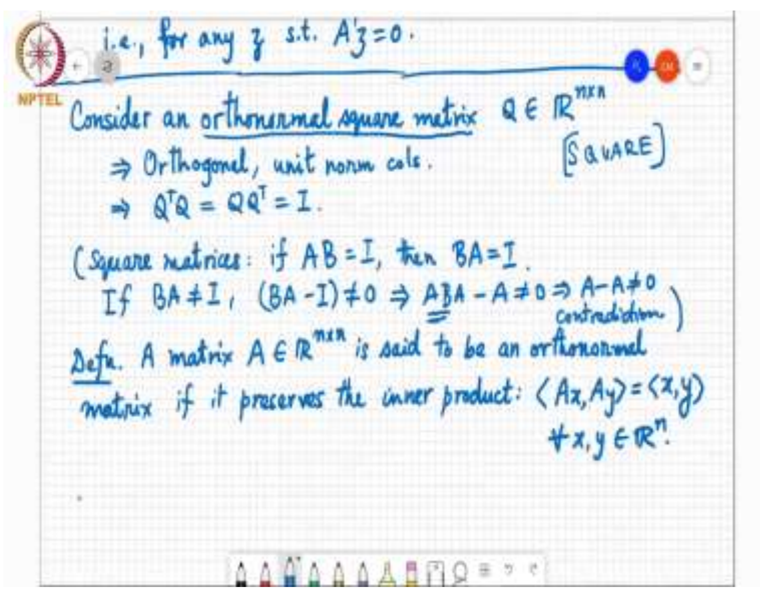
Matrix Theory
Professor. Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bengaluru
Orthonormal matrices definition

(Refer Slide Time: 0:14)





Now, there is a couple of points I want to make, and then I will state one proposition. So consider, did somebody have a question? Was it a question? So, I am defining an orthonormal square matrix. Go ahead. So, let us consider an orthonormal square matrix. This is a matrix such that it has orthogonal and unit norm columns. That is the definition of an orthonormal matrix.

And it has the property that another way to define it is matrix Q satisfies Q transpose Q, equals Q, Q transpose equals the identity matrix.

So by the way, in general, it is always true that for square matrices, if A B equals the identity matrix, then B A is also equal to the identity matrix, this is easy to see. Basically, if B A is not equal to the identity matrix, then B A minus I is non-zero, which implies now if I pre multiply by A, then A B A minus A is not equal to 0, but A B is equal to the identity matrix, which implies A minus A is not equal to 0, which is a contradiction.

So, for square matrices, it is always true that if A B equals I, then B A equals I therefore, if I say Q transpose Q equals I, then Q Q transpose must also be equal to the identity matrix. So, there is another definition for an orthonormal matrix, which is the definition I am going to use for stating this proposition which is that a matrix A in R.

Student: Excuse me sir.

Professor: Yeah?

Student: Sir in above proof B A we are taking not is equal to I.

Professor: Yeah.

Student: So, why writing A B A is equal to A.

Professor: A B is equal to I. If A B equals I, that is what we are trying to show here that if A B is I, then B A is not, it must also be equal to I. So, A B is equal to the identity matrix, identity times A is always equal to A whatever A is.

Student: Yes sir, I got it. Thank you.

Professor: Yeah?

Student: I have a doubt. So, on the top you have written Q transpose Q equals Q Q transpose equal to identity matrix, but dimensional Q transpose Q will not be equal to the dimension of Q Q transpose, right?

Professor: Square.

Student: It is for n cross n fine. If it is suppose if it is in rectangular matrix in how will this.

Professor: Then it is no longer true that if Q transpose Q is I, then Q Q transpose must be equal to I.

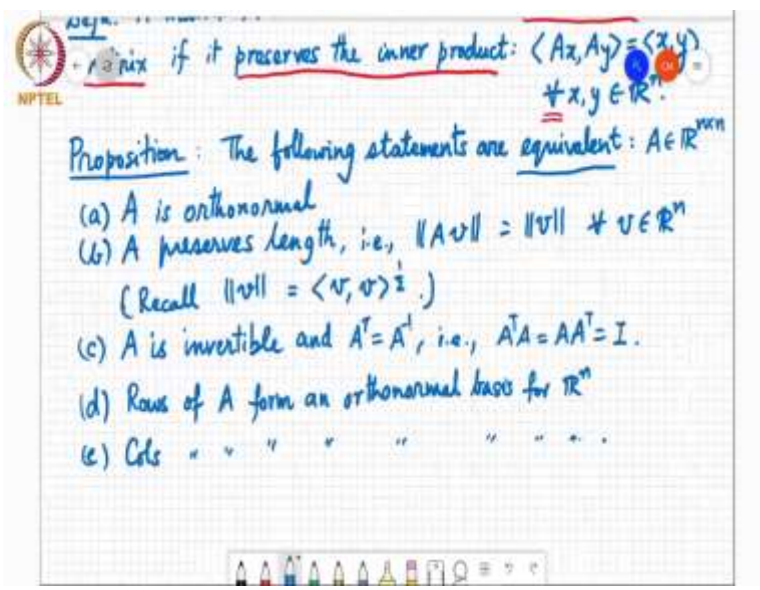
Student: But either of them will hold true, right? Like at least one in, if one does not hold true then other will definitely hold true, right? In case it is an orthogonal matrix.

Professor: So, if the columns are orthonormal then Q transpose Q will contain as it entries inner products between columns and so, that will be equal to the identity matrix but Q Q transpose may not be the identity matrix. Similarly, if the rows are orthonormal then Q Q transpose will be the identity matrix, but Q transpose Q need not be the identity matrix.

Student: Fine sir.

Professor: So, this is said to be an orthonormal matrix if it preserves the inner product so, what I mean by that is if I take the inner product between Ax and Ay that is equal to the inner product between X and Y for every x and y belonging to R to the n. So, that is also another definition of an orthonormal matrix .

(Refer Slide Time: 6:55)



So, here is the proposition. So, the following statements are equivalent a. A is so, here A is matrix in R to the n cross n, b is A preserves the length by that I mean Av for every v in R to the n. So, recall, that is the definition of the (())(8:49). In fact, for the moment, I would not write the subscript 2 because this result is actually more general, it does not require the inner product to be defined the way we have been defining it until now we have seen only one example of an inner product, there are other possible definitions for the inner product.

But right now, we are working with the dot product as the notion of the inner product, but any other valid notion of the inner product can be used to define a norm defined like this and that is called a norm that is induced by an inner product and this is true for any definition, any valid definition of an inner product.

Student: Can you explain the previous definition Ax Ay equal to x, y? What is Ax? This is the column?

Professor: x is a vector. So A x is a vector. So, (())(9:54) inner product between A x and A y.

Student: So what is x and y, it is numbers or vectors?

Professor: It is right here, x, y belongs to R to the n.

Student: So, it is the column one and column two it will be 1 and 2.

Professor: What do you mean? x is a vector in R to the n, y is a vector and R to the n, x and y are vectors, x is a vector, and y is a vector, which will compute the inner product between x and y, whatever number you get, is the same number as what you would get if you computed the inner product between Ax and Ay. And that is true for any pair of vectors I choose. I think this is important that everybody understands the definition. Otherwise, this entire proposition makes no sense.

So, if you have other questions about the definition of an orthonormal matrix? Please ask. The definition is that it should preserve the inner product. That is you take two vectors in R to the n, if you find the inner product that is x, y. And then you take Ax and Ay, those are also two vectors in R to the n, and you find their inner product. And those two numbers must be equal for all pairs of vectors in R to the n. Not for just one particular pair. Go ahead.

Student: Sir actually in that inner product line is a matrix into a vector is actually a transformation, right?

Professor: Yes.

Student: So what is the meaning, I mean inner product is actually the length, I mean, the component of X along Y or something, in that sense, right? I mean inner product between X and Y means the component of x along y.

Professor: Correct, it is preserving the relative orientations of x and y. That is what it means.

Student: In whatever frame we transform this vector.

Professor: Yes, yes. So, if, for example, x and y were orthogonal, they were (perpendi) transformation by A, they will remain perpendicular to each other, if co-linear with each other, then even after transformation by A, they will remain collinear with each other. That is what it means.

Student: okay.

Professor: And mathematically, this is the precise meaning of a matrix being an orthonormal matrix. It preserves all inner products, not just when x and y are orthogonal, or not just when they are co-linear.

Student: Hello sir, sir just a question like it is not, when multiplied a vector with a it is not changing the length. So is it, it will rotate the vector?

Professor: That is a good way to think about it, it will rotate the vector, but it will not change its length. So, in some say, so another way to put it is that every rotation matrix is actually an orthonormal matrix. And every orthonormal matrix can be thought of as a rotation matrix.

Student: Okay sir.

Student: Sir is this sometimes called orthogonal matrix as well?

Professor: So that is actually some, that is actually a good point, depending on the textbook. Some textbooks refer to this, this kind of matrix as an orthogonal matrix, some textbooks refer to it as an orthonormal matrix. I prefer the terminology of orthonormal matrix, because it also clearly tells you that every column is unit norm, an orthogonal matrix, if you really wanted a more general definition, it is a matrix where the columns need not be unit norm, but they are still orthogonal to each other. That could be a more general definition of an orthogonal matrix.

But yeah, just to be clear, I want to call it an orthonormal matrix in this course, and it is a matrix whose columns are unit norm, and are orthogonal to each other.

Student: Sure sir.

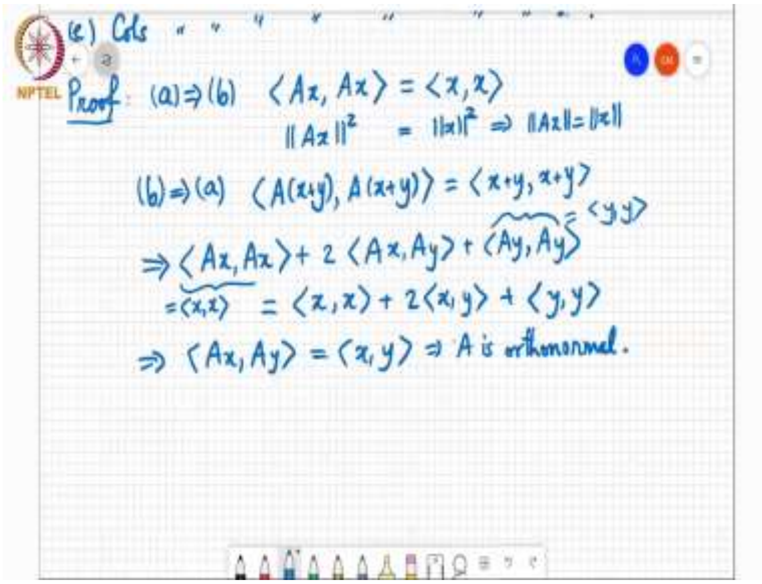
Professor: These are two definitions, but for the purposes of showing this proposition, I will go with this definition, meaning that it is an orthonormal matrix if it preserves the inner product. So the third property is that A transpose is equal to A inverse. A is an invertible matrix, which means that A transpose A equals A A transpose, which equals the identity matrix. And the fourth property is that the rows of A form an orthonormal basis for R to the n. And the last one is that the columns also form an orthonormal basis for R to the n.

So how do you prove such a proposition? So it says that these statements are equivalent, meaning that if I tell you that A is orthonormal, it is the same as me telling you that A preserves length. So, if you have 5 statements like this, the way to show that all of these are equivalent is to say, for example, you take the first two statements, you should show that a implies b and b implies a, that means a and b are equivalent, then you take the say, the third statement the c, then what you have to show is that either a or b, one of those two statements implies c, and c implies either a or b one of those two statements, then it means that a, b and c are equivalent.

Then similarly, you take d and you show that d implies one of these three statements, a, b, or c. And you should show also that a, b or C, one of these three statements implies d, and so on. If you show all of that, then then you have shown that these statements are all equivalent. Of course, there are many ways to do it.

For example, you show a implies b and b implies a, then you show b implies c and c implies b, and you show c implies d and d implies c and d implies e and e implies d. So that also means that these statements are all equivalent. So, let us see how to show this.

(Refer Slide Time: 17:17)



So first, we will tackle a implies b. So we want to show that if a is an orthonormal matrix, then it preserves the length meaning that for any vector if I look at the length of V, it is equal to the length of the transformed version of V which is Av. So, and then vice versa, if a preserves length of every vector in R to the n, then a must be an orthonormal matrix.

So, to show first a implies b, an orthonormal matrix implies that it preserves the length, then what I need to do is I just need to consider the inner product between A x and A x and this, the inner product between Ax and Ax because A is an orthonormal matrix, by definition, it preserves inner products. So basically, this is equal to the inner product between x and x.

So, that means that if I, so this is nothing but Ax square, and this is nothing but x square, which means that Ax equals x. So this means that so this is what we wanted to show that a preserves length, so a implies b. Similarly, b implies a, what I need to do is I need to show that if a preserves the length, then a is an orthonormal matrix, that means that all inner products are unchanged by multiplication by a. So, what I will do is I will consider.

Student: Sir.

Professor: Yeah?

Student: Sir could you explain how inner product of Ax with itself is equals to inner product of x with itself?

Professor: That is from the definition. So, what I want to show here is that if a is orthonormal, then it preserves length. So if a is orthonormal, by definition, because it is an orthonormal matrix, it preserves the inner product, the inner product between any two vectors, in fact, y not be equal to x. So, if a is an orthonormal matrix, then Ax Ay is equal to the inner product between x and y. This is true for all pairs of vectors, x and y.

And so all I am doing is to take a special case of this, where y equals x. So, the inner product between Ax and Ax is equal to the inner product between x and x. So, you can see that although the statements look quite different from each other, saying that a is an orthonormal matrix, meaning that it preserves in our products and saying that a preserves length they seem different from each other in terms of statements, but in fact, it follows trivially from the definition, is it clear?

Student: Yes sir, thank you sir.

Professor: So, if I take the inner product between Ax plus y and Ax plus y, this because A preserves length, so now I want to show that b implies a, so A preserves length, so this is nothing but the length of A times x plus y square. And so this must be equal to the length of x plus y square.

But if I expand this out, by using the linearity of inner products, then this thing can be written like this, it is the inner product between Ax and Ax plus two times the inner product between Ax and Ay plus the inner product between Ay and Ay and the right hand side, if I expand this out, I get the inner product of x with itself plus two times the inner product between x and y plus the inner product between y and y.

Now, once again, I use the property that A preserves length. So this quantity, inner product between Ax and Ax, this is equal to the inner product between x and x and this term is equal to the inner product between y and y. So, this x x will cancel with this x x this y y cancels with this y y, and the 2 and 2 can cancel and so you are left with Ax Ay equals x y, so it preserves in the products and so then it is orthonormal. Then we consider c the statement c.

Student: Sir.

Professor: Yeah?

Student: Sir the moment you are considering inner product of Ax comma Ax is x comma x, sir that moment only you are considering it to be orthogonal, right? orthonormal.

Professor: That is what it is showing here actually.

Student: But in the second last step, you already considered Ax comma Ax is x comma x.

Professor: So, what we are trying to show here, I think it is important to pay attention a little bit to what we want to show. So, what we want to show is that b implies a, what is that? It is that if a preserves the length, that is if norm of Av equals norm of v for every v in R to the n, then A preserves the inner product, that is the inner product between Ax and Ay is equal to the inner product between x and y for all x, y in R to the n, that is what we want to show. That is what we are doing here.

It is not a difficult proof. It is very simple. But it is good to keep in mind exactly what it is that we are showing. So, in writing the first step, I am using the fact that b is true, I am saying if b is true, then I want to show that a will preserve an inner product between any pair of vectors. So this is true because a preserves the length. This is also true because a preserves the length Ax Ax equals x comma x, that is also true because a preserves the length and similarly Ay Ay equals y y is because a preserves the length.

And a consequence of this is that I am getting Ax Ay inner product is equal to the inner product between x and y, which means that a is orthonormal.

Student: Sir.

Professor: Yeah?

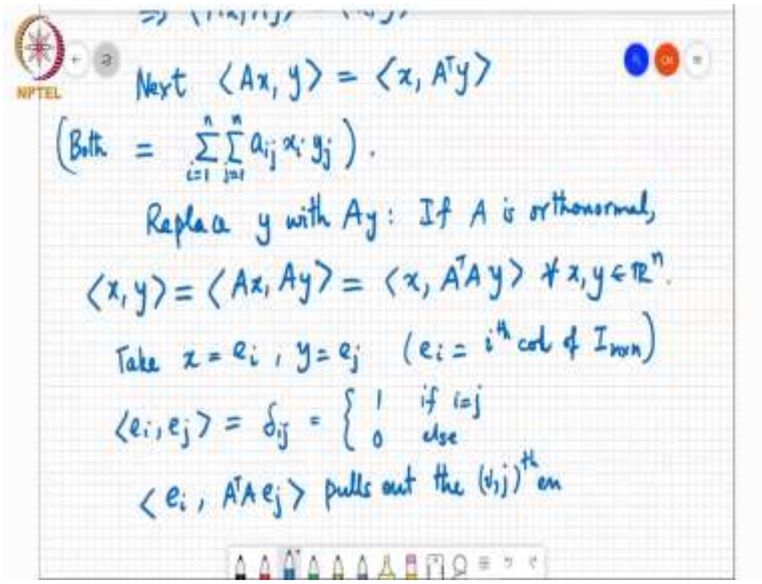
Student: Sir you give the definition that an orthonormal it is an, A is an orthonormal matrix, if it preserves the inner product, but that statement does not speak about the converse that is if it is, if it preserves the inner product then it should be orthonormal that part is not implied to the definition, right?

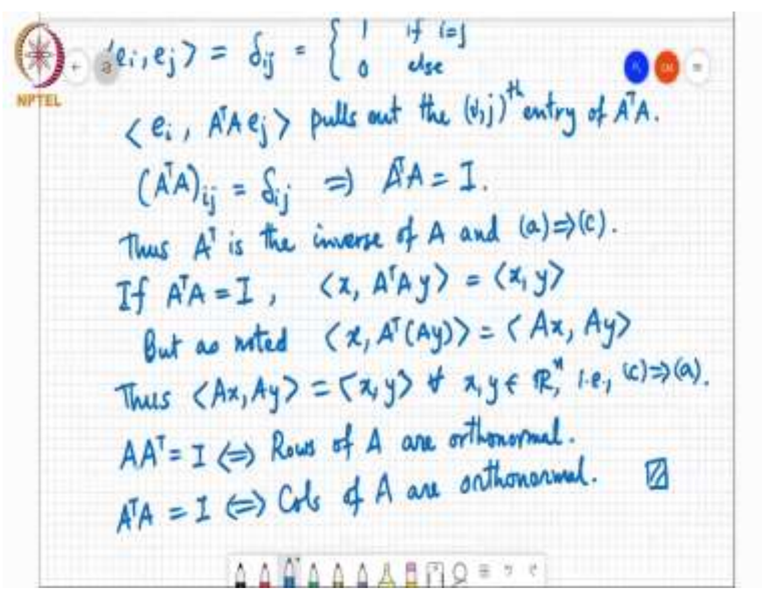
Professor: So, this is also a very good point. In mathematics a definition is always an if and only if statement. So, we write it like this, the matrix is said to be an orthonormal matrix if it preserves the inner product, but, when we are saying that it is said to be or it is defined as an orthonormal matrix, we mean that this is an if and only if condition.

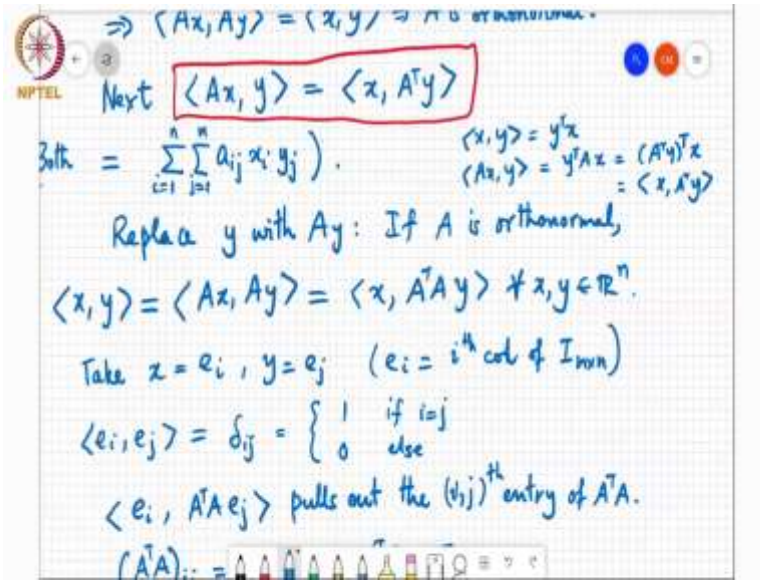
Student: Okay.

Professor: By definition an orthonormal matrix is one which preserves the inner product. All definitions in mathematics are like this, we define something to be in a particular way, it means that you know a is equal to b if I define a to be equal to b or a to b or if I say a matrix A is defined to have a property x if it satisfies y it means that x and y are equal to each other. Now, let us do the next step.

(Refer Slide Time: 26:33)







Next, if I consider the inner product between Ax and y, I will make a statement this is equal to the inner product between x and A transpose y. This can be seen by just writing out the expansion of what this will be. So, basically this A x y is nothing but summation i equals 1 to n sigma j equals 1 to n a i j x i and y j and y g and if you expand this out in terms of the entries of a you will find that this is also exactly the summation.

So, I should say. So, now, what I can do is instead of y I will replace y with a y then what happens is that I get so, I should write that out, so, do not have confusion later. Then what I get is the inner product between Ax and Ay now, because, so, suppose A is an orthonormal matrix then it preserves inner products. So, that means that this is equal to x y and this is equal to the inner product between x and A transpose A times y and this is true for every x y.

So, if a is orthonormal, then we have that this is equal to this for every x y in R to the n. So, then that means that what I can do is I can take example vectors for x and y, x equal to e I, the ith column of the identity matrix and y equal to e j where e I was the ith column of I n cross n. Then what I have is, if I consider the left hand side this ei, ej.

So, ei, ej is equal to, I can write it as delta ij which is this equal to 1 if i equals j and 0 otherwise, then the right hand side is the inner product between ei and A transpose A ej, which basically, if you expand this out, you will see that this exactly pulls out the i jth entry of A transpose A. So that means that the i jth entry of A transpose A is delta i j. Which means that A transpose A is the identity matrix.

So, then it means that is A transpose is the inverse of A and a in place c, so we have shown that now, conversely, if A transpose A is equal to the identity matrix, then if I consider the inner product between x and A transpose Ay, this is equal to, this is the identity matrix, so, this is always equal to x and y, the inner product, because a transpose it does not change y at all.

But then the left hand side this, I can move this A transpose over here. And I can so, more generally, so, let me just write it like this x, A transpose times Ay. I will consider this to be some vector, A transpose times something is equal to the inner product between Ax and Ay. And so thus Ax Ay equals x, y for all x, y in R to the n, and thus c implies a. And finally, the last step is just that, if A transpose A equals I, it means that the rows of A are orthonormal.

So, we have seen that. So, these two are equivalent statements, because this is just computing the inner product between rows of A. And if this is equal to I, it means if I take any pair of any distinct pair of rows, they are orthogonal to each other. And if I take any row, it has unit inner product with itself. So, the rows of A are orthonormal. They are both equivalent statements. And so c implies d and d implies c. And similarly, A transpose A equals i is the same as saying that the columns of A are orthonormal. Yes.

Student: Sir in order to show that inner product of A x and y, same as the inner product of x and A transpose y, can we use this formulation that inner product of x comma y is same as (())(34:13).

Professor: It is the same point, it is exactly the same point.

Student: Then we can (())(34:20).

Professor: This is it, this is the point.

Student: Sir because, you know, the submission sometimes does not click how to write in form of this or two summations, but I mean directly doing the transpose we can easily separate it out.

Professor: So, but that will be valid. So. you can do it that way also.

Student: Okay ir. Thank you.

Professor: So the thing is that, it follows from the definition of the inner product that we say so, it is you have to take the transpose of the first vector and then multiply it with the second vector.

So, if you think of it that way, it follows immediately. That is also a valid way to write it out. Any other questions?

Student: Sir is that true for all A? No, right?

Professor: Of course, this is always true. A x comma y inner product is always equal to x comma A transpose y what the previous student just said is that if I think of x, y to be equal to y transpose x, if I think of it this way, which is how I define the usual inner product earlier, then if I were to take Ax and y, this is equal to y transpose Ax, which is also equal to A transpose y whole transpose times x.

Student: I got it.

Professor: And which I can write as x transpose times A transpose y. No, I do not need to do that.

Student: Sir here to prove the equivalence of 5 statements, we are actually going through 10 proofs, but actually 6 proofs would suffice because a implies b, b implies c, c implies d, d implies e, then e implies a then all of them will imply each other.

Professor: So that is also a valid way to show it. There are many ways to write out this proof. I have shown you one way here.