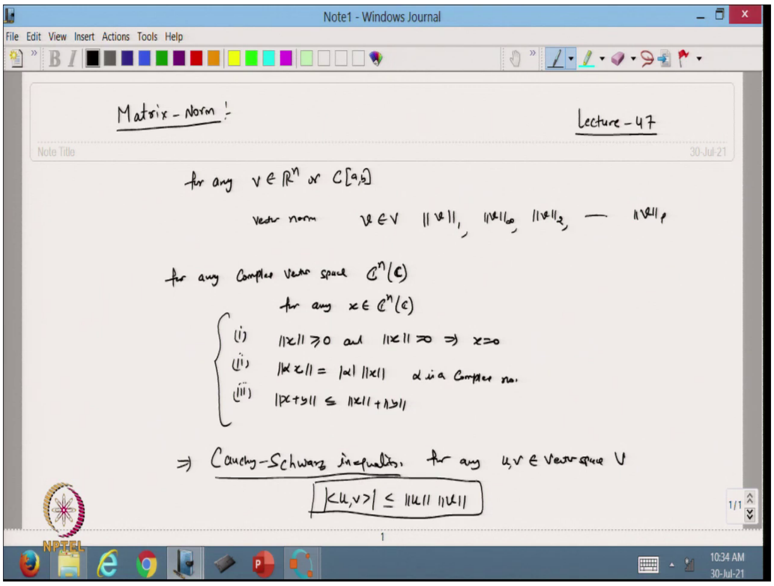
Matrix Computation and its applications
Dr. Vivek Aggarwal
Prof. Mani Mehra
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture - 47
Matrix norm

(Refer Slide Time: 00:17)



Hello viewers. So, welcome back to the course on matrix computation and its application. So, in the previous lecture, we discussed the vector norm or defined on a vector space. So, today in this lecture, we are going to discuss matrix norms. So, let us start with that. So, in the previous lecture, we have discussed that for any vector $V \in R^n$ or a set of polynomials or a continuous function; so, these are things we are defining for, then we have discussed that we can define the vector norm.

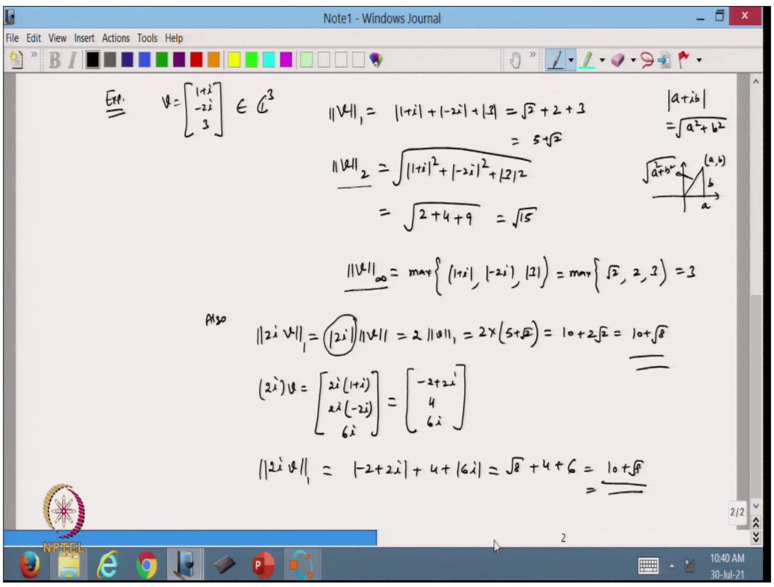
So, this is we have defined for any $v \in V$, this is my vector norm and we have discussed 1-norm or infinity norm or 2-norm and in the end, we have discussed that we can define p norm. Now, these things we have done there and now, we want to define that the same thing is applicable for any vector belonging to complex vector space. So, now, for any complex vector space, we call it C^n define over the complex number C, now the same things is going to happen there that for any x $\in C^n$.

So, this is my x a vector coming from this one. Now, in this case also we have to discussed first thing is that $\|x\| \geq 0, \|x\| = 0 => x = 0$. Second thing is also that if I take any scalar alpha that is $\|\alpha x\| = |\alpha|.\|x\|$. And the third one is that if I take any two vectors, then this is equal to $\|x + y\| \leq \|x\| + \|y\|$

So, this is also true in the case of a vector space that is defined in the complex plane. Then, we also discussed the famous Cauchy-Schwarz inequality. So, in this case, for any u and v that belongs to a vector space V, then we have we can generalize this one that $|< u, v >| \leq \|u\|.\|v\|$

So, this is also how we can generalize any of the vector space and inner product, we know that we can define it with the different different norms or maybe we have discussed the inner product, we have seen how we can define the inner product, when we change the vector space. So, these things we already saw.

(Refer Slide Time: 04:50)



So, now, I want to discuss what is going to happen, when we are taking a vector. So, suppose

I take a vector v that is $v = \begin{bmatrix} 1+i \\ -2i \\ 3 \end{bmatrix} \in C^3$. Now, I want to define its 1-norm. So, 1-norm in this case, it will be taking the modulus value or the magnitude of 1 +i; this one. Now, it is the magnitude of this.

So, a +ib and if I take the magnitude, then it is a square plus b square under the root. So, that we already know. So, it is in this case, it will be because this magnitude is basically supposed this is my complex number a. So, this is a b, then this is b. So, I know that this value is a square b square under the root. So, this is the magnitude of this one.

So, now, from here, I can write this as $\sqrt{2}+2+3=5+\sqrt{2}$. So, that is my 1-norm for the vector space or for a vector in C^3. Now, the same thing I can define for that 2-norm. So, 2-norm is $||1+ i||^2+|-2i|^2 +|3|^2$ = $\sqrt{2+4+9}=\sqrt{15}$ So, this is my 2-norm and I can define it similarly, I can define an infinity norm.

So, infinity norm will be the maximum {|1+i|, |-2i|,|3|}So, it is again I am taking the maximum{ $\sqrt{2}$,2,3}=3. So, I am able to define the infinity norm, the 2-norm, 1-norm of any vector that is coming from the C^3. Also, one more thing I want to discuss; I want to see what will happen when I take 2iv . So, this one I want to see.

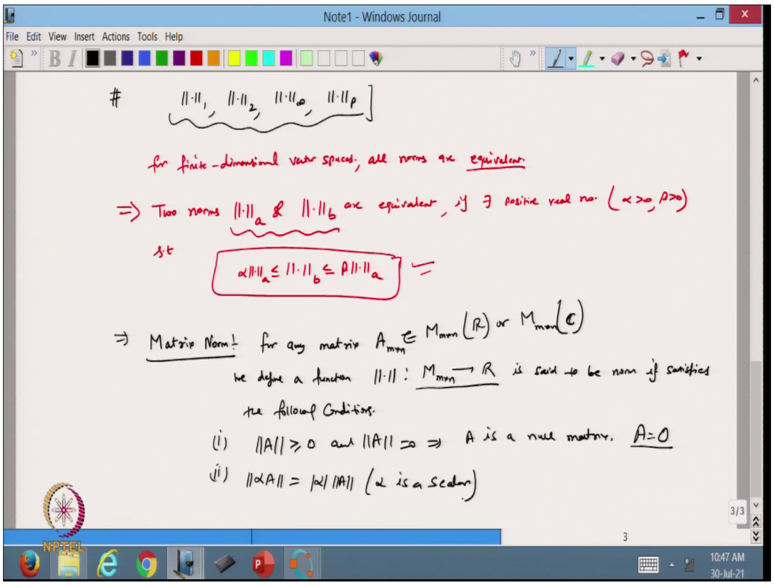
Then, by the property, this property, we are going to use it should be equal to |2i|||v||. Now, suppose, I take for one; so, in this case, I am taking 1-norm. So, here, it should be two and that should be 1-norm and it is 2 into; so 1-norm of this is this. So, it should be $2(5+\sqrt{2})$.

So, that should be the value. Now, if I take $(2iv) = \begin{bmatrix} 2i(1+i) \\ 2i(-2i) \\ 6i \end{bmatrix} = \begin{bmatrix} 2-2i \\ 4 \\ 6i \end{bmatrix}$

$$||2iv||_1 = |-2+2i| + 4 + |6i| = 10+\sqrt{8}$$

So, in this case, what we have seen is this modulus value, we have to take when we deal with the complex numbers. So, this is one of the examples, we can define in complex vector space.

(Refer Slide Time: 10:33)



Now, one thing we want to discuss is that now we are able to see lots of norms; the vector norm I have defined, I have defined 2-norm, I have defined the infinity norm or maybe I can define p norm. Now, it may happen because we have seen that for any vector, all norms are different. So, the question comes that it may happen that in 1-norm, if I am taking the convergence in of the vectors, then it may happen that in 1-norm it is very high and in another norm, it is very low or it in 1-norm, it may converge and in if I take the another norm, it may diverge.

So, these things generally do not happen. So, from here, we have a very important condition about that for finite dimensional vector spaces, all norms are equivalent. So, all norms are equivalent. So, what is the definition of this one? Now, from here, I can define those two norms. Suppose, I take a and b, two norms. So, 2-norms are equivalent, if they exist there exists a positive number.

The positive number means positive real number. Suppose, I take $\alpha > 0, \beta > 0$ such that $\alpha \,||\,.\,||_a \leq ||\,.\,||_b \leq \beta \,||\,.\,||_a$

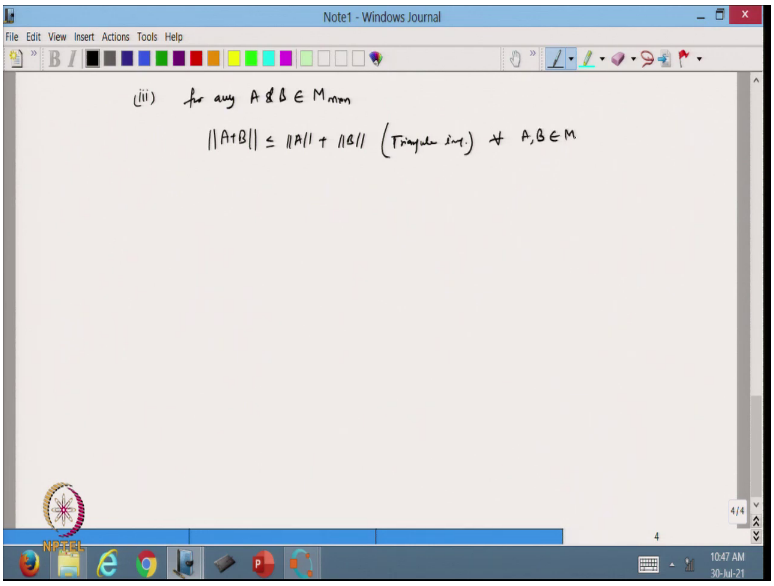
So, if we are able to write this one, then we say that these 2-norms a and b are equivalent. It means that the value of a vector in norm 1-norm is bounded by the values of the same vector in another norm. So, that is called the equivalence relation. Now, based on this one, now let us define. So, once we have defined the vector norms, then the next question is how we can define the matrix norm. So, we want to define a matrix norm.

So, now, what we are going to do is the matrix norm. Now, because the matrix is made up of vectors, we can define the matrix norm as. So, I am defining for any matrix A. So, this matrix I am taking. So, whether it can be a rectangular or a square matrix does not matter. So, for any matrix A, we define the norm. We define a function that is defined from the space of order m cross n and it can be real and complex.

So, it does not matter. So, we are defining a function over the space vector space of matrix m cross n to a real number. So, this is basically, I can say that positive real numbers or real numbers R. So, we define a function that is said to be the norm, if it satisfies the following conditions. So, we are defining on the set of all the matrices and this can be real or it can be complex.

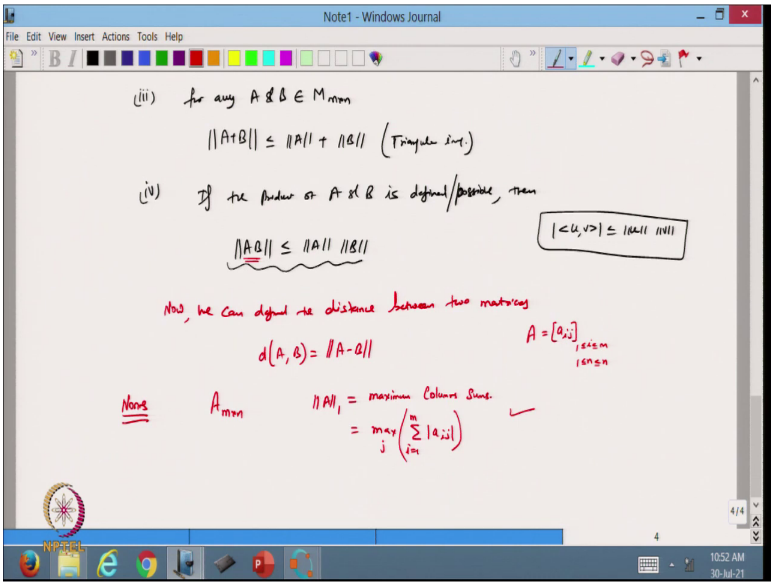
So, these are the following conditions. First one is that again if I define the norm of a matrix and it is $\| A \| \geq, \| A \| = 0$ which imply that A is a null matrix, it means that A= 0, 0 matrix. Second one is that again if I take some scalar $\| \alpha A \| = | \alpha | . \| A \|$ this value where alpha is a scalar; so, it can be real or complex.

(Refer Slide Time: 17:02)



Third one is if I take two matrix for any A and B belongs to m*n. So, of course, if I take the sum, then it should be of the same dimension. So, if I take the norm, then it is $| A + B |\leq\| A \| . \| B \|$. So, this is the triangular inequality and this is true for all A and B belongs to well this is already I have written that for any this one.

(Refer Slide Time: 17:47)



Now fourth one; so, the fourth one is I take A into B. So, I am writing that A into B only it can be written if it is so if I can write that if the product of A and B is possible is defined or possible, then I can write AB, the product of AB taking the norm, then this is $\| AB \| \leq \| A \| . \| B \|$.

So, this condition if you see this is analogous to the Cauchy Schwarz inequality because in the Cauchy-Schwarz inequality, we are also taking the dot product of the two vectors and then, this gives you this condition. So, now, this condition is happening in the case of matrices; but this is only possible when A and B are defined.

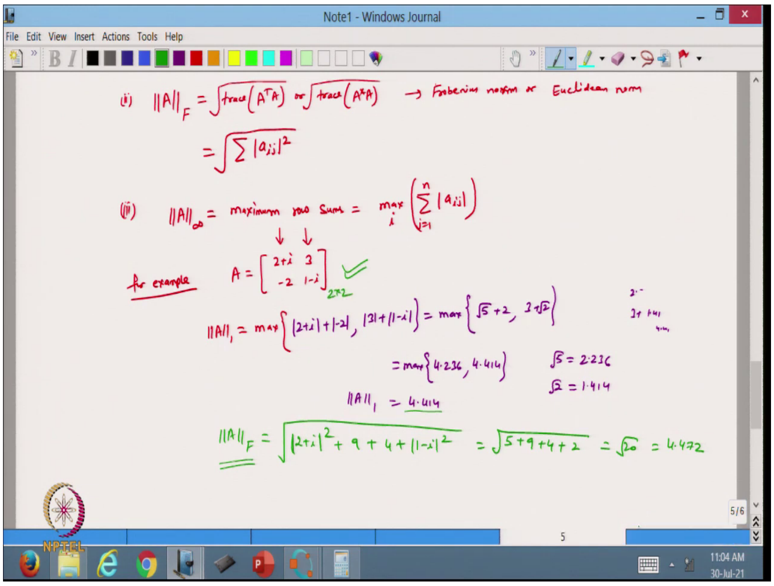
So, if I take all the sets of all the square matrices, then it is ok, no problem; but if the matrix is a rectangular matrix, then we have to take care about this product. So, these are the four conditions we have to define. So, after defining this one, now we can define the distance between two matrices. So, I can define the distance between suppose A and B, then it becomes A minus B norm. So, then I can define the distance between two matrices.

So, once I am able to define a norm which satisfies all these conditions, then we say that this norm is well-defined. So, let us discuss the different types of norms. Now, first of all, so I will define norms, if I take a matrix A and suppose its dimension is m*n. Now, I define its 1-norm. So, 1-norm also we have defined for the vectors. So, similarly, I am defining the norm one and here the 1-norm is a maximum column sum. It is $\| A \|_1 = \max_j \left(\sum_{i=1}^{m} \| a_{ij} \| \right)$

So, I am taking the column sum. So, I am taking the first column and then, taking the magnitude and i moving from 1 to m whatever the value is there and then, so first column taking the sum, second column and then taking the maximum of this one.

So, this is the maximum column sum. Similarly, this is one of the norms that we can verify that it satisfies all these conditions. So, this is the 1-norm.

(Refer Slide Time: 22:05)



Now, similarly, I can define another norm and that we have already seen is the Frobenius norm. So, Frobenius norm, we have seen in the inner product of the matrix. So, this is you have you know that this we have taken from finding $\| A \|_F = \sqrt{trace(A^T A)} or \sqrt{trace(A^T * A)} = \sqrt{\sum |a_{ij}|^2}$

So, this one we take from the inner product, so this norm is called. We know that this norm is called the Frobenius norm or Euclidean norm. So, this is another norm we have defined. So, this norm actually we know that this is equal to taking the magnitude square and then, under the root. So, that is equal to this one. Now, the third norm, we are going to define the infinity norm. So, the infinity norm is again, it is the maximum row sum. So, in this case, I am taking the maximum over the rows and taking the summation of $\| A \|_\infty = \max_i \left(\sum_{j=1}^{n} \| a_{ij} \| \right)$.

So, it is called the maximum row sum. Now, this sums we can define. So, for example, now I take a matrix A. So, I am just taking a 2 by 2 matrix maybe. So, I just writing $A = \begin{bmatrix} 2+i & 3 \\ -2 & 1-i \end{bmatrix}$. So, this is the complex matrix. So, I want to take 1-norm. So, 1-norm, I have just defined taking the maximum column sum.

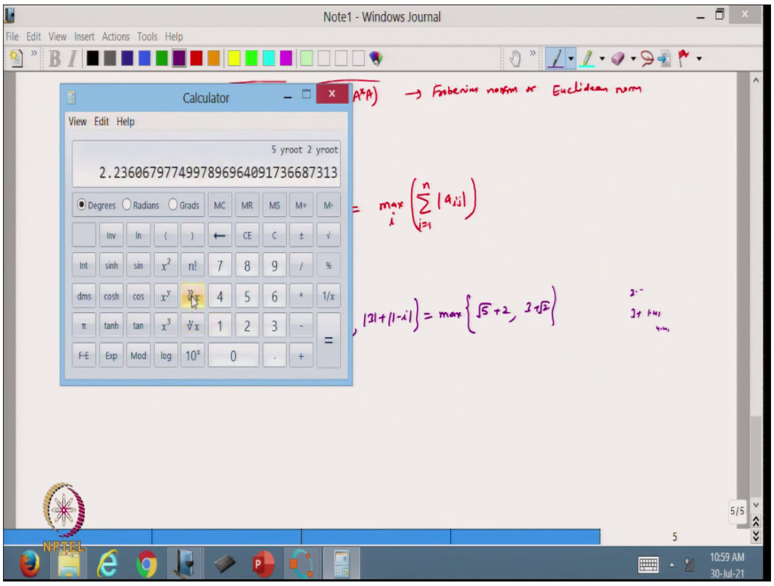
So, what I am going to do is that I am taking the maximum of the column sum. So, first, I am taking this column and then, this is my column. So, this value is I am taking

$$\|A\|_1 = \max[|2+i|+|-2|,|3|+|1-i|]$$

$$= \max\left\{\sqrt{5}+2, 3+\sqrt{2}\right\}$$

So, it is this one. So, in this case, the maximum value if I take, so this one we can find.

(Refer Slide Time: 26:04)

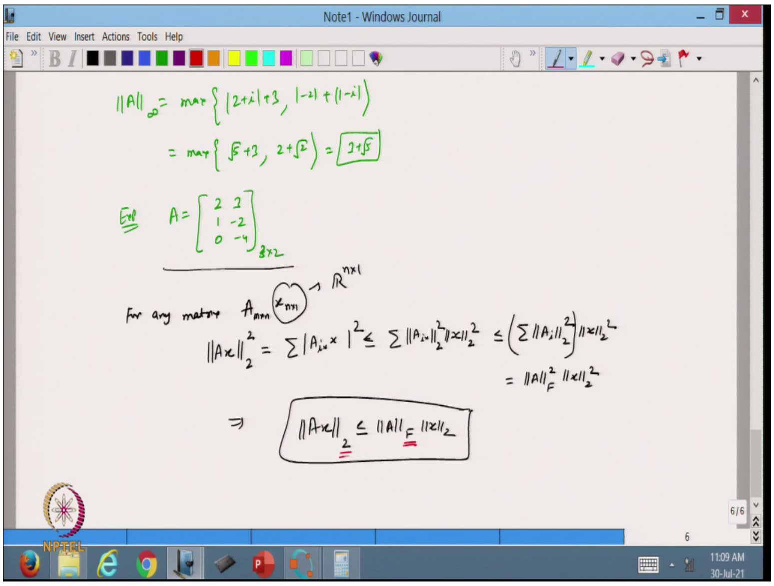


$$= \max\left\{2.236+2, 3+1.414\right\}$$
$$= \max\left\{4.236, 4.414\right\} = 4.414$$

So, that is my 1-norm. Similarly, I can define another norm; maybe I can define, Frobenius norm. So, in this case, what I am going to do is take the modulus value of each square.

So, it will be $\|A\|_F = \sqrt{|2+i|^2+9+4+|1-i|} = \sqrt{5+9+4+2} = \sqrt{20} = 4.472$. So, this is the Frobenius norm. So, 1-norm is 4.14 and Frobenius is 4.472.

(Refer Slide Time: 28:47)



Similarly, we can define infinity norm. So, now, so 1-norm is there, then now I am taking the infinity norm. So, it is the maximum of the row sum. So, $\|A\|_{\infty} = \max\{|2+i|+3, |-2|+|1-i|\} = \max\{\sqrt{5}+3, 2+\sqrt{2}\} = 3+\sqrt{5}$

So, that is my infinity norm. So, this way, we can define the different types of norms for a given matrix and I have taken this matrix as a square matrix with complex eigenvalues or complex entries. But these things can be done for any matrix, maybe I can take example of a matrix A, which is like $A = \begin{bmatrix} 2 & 3 \\ 1 & -2 \\ 0 & -4 \end{bmatrix}$. So, this is 3* 2 matrix; the same way I can define these values.

Now, after defining these, I wanted to see this because we have seen here that the matrix norm and the vector norms look similar. Then, we want to define one condition here because we have seen the Frobenius norm in a vector form also and the matrix form also. So, I just want to define that now since, now if I want to define a matrix A, any matrix.

Now, for any matrix A, I define Ax. So, Ax I know that for any matrix A, Ax is a vector because I am taking x as a vector. Suppose, x A is from m cross n and then, definitely x will be n *1. So, in this case, I can say that this belongs to R^n and I am taking this column vector. So, I can write like this one. Now, suppose I take its 2-norm. So, I am just taking the 2-norm.

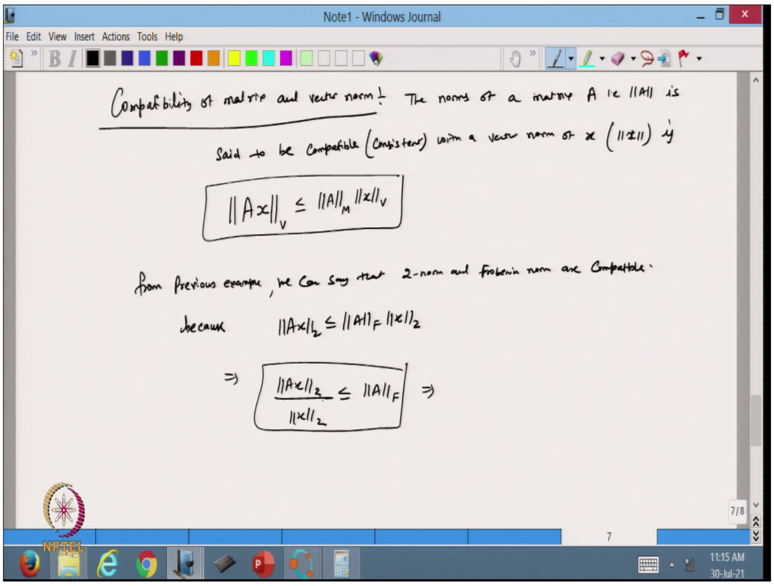
So, I know that this is equal to summation taking the multiplication and then taking the square of that. So, this is basically I am writing $\| Ax \|_2^2 = \sum | A_{i*}x |^2 \leq \sum \| A_{i*} \|_2^2 \| x \|_2^2$ Now, I can apply because I know the Cauchy-Schwarz inequality.

So, from here, I can write that this is always less than equal to summation. So, I have to take and then, this value because this x is complete x, I am taking. So, this can be less than equal to this value by the Cauchy-Schwarz inequality and this one I can write again as because x is so this sum is going on the matrix now and this x can be taken out.

So, this will be equal to summation square and now I am defining it here. So, this is basically the 2-norm basically I am taking; this is a 2-norm. So, this one this value and what is this? It is just that we have taken that this is equal to the Frobenius norm. So, it is not the two norm; it is just we have written Frobenius norm $\| Ax \|_2^2 \leq \left(\sum \| A_i \|_2^2 \right) \| x \|_2^2 = \| A \|_2^2 \| x \|_2^2$

So, from here, I can say that 2-norm is less than equal to Frobenius norm and these 2-norm, it means that if I take the A x taking the 2-norm that is always less than equal to the Frobenius norm of A into the 2-norm of x. It means that the 2-norm of the vector and the Frobenius norm of the matrix are connected to each other. So, such a type of connection, we call it compatibility.

(Refer Slide Time: 34:47)



So, let us define one more term here; the compatibility of matrix of matrix and vector norm. Now, the norms of a matrix suppose I take the matrix A. So, that is I write this one is said to be compatible or I can say consistent with a vector norm if, so compatible with a vector norm of suppose I take x that is if. So, I take the matrix A, then Ax should be defined.

So, then I take its vector norm because it is a vector, so I am taking the vector norm. $\| A \|_v \leq \| A \|_M \| x \|_v$ So, this is the matrix norm I am taking and this is the vector norm I am taking. So, the norms of the matrix, so this is a matrix norm M is said to be compatible with the vector norm of x. So, this is a vector norm I am taking, if this is the condition; this condition is satisfied ok.

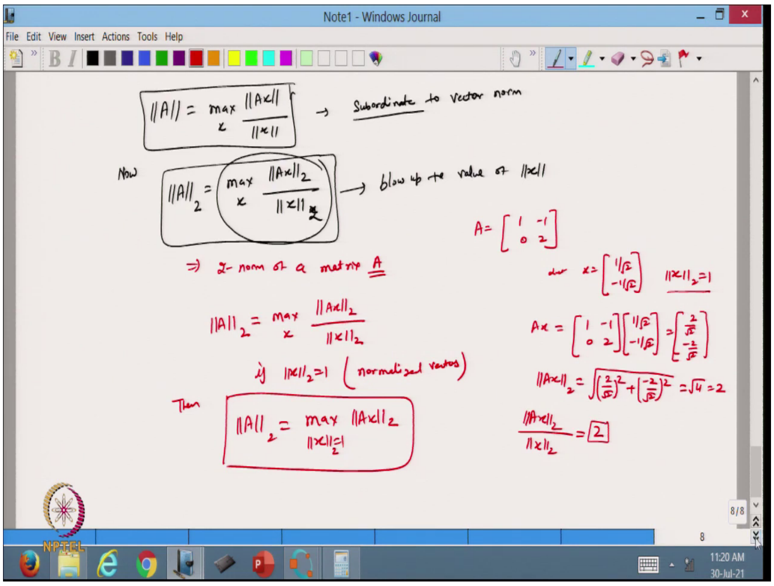
So, in this case, both the norms are coming, M is also coming and V is also coming. Now, from the previous one, we have seen that the 2-norm of the vector is compatible with the Frobenius norm of the matrix. So, from the previous example, we can say that 2-norm and Frobenius norm are compatible because I can write A x 2-norm that is less than equal to Frobenius norm and 2-norm.

So, this is the way we can define. Now, we have defined the Frobenius norm and we have defined the 2-norm, now the next question is how we can define the 2-norm of a matrix. So, let us define from here, we define the 2-norms because I want to see what will happen if I instead of Frobenius, I just take the 2-norm. Now, from here, I just define that from here I can write $\| Ax \|_2 \leq \| A \|_F \| x \|_2$

$$\frac{\| Ax \|_2}{\| x \|_2} \leq \| A \|_F$$

and from here, using this condition, I can define a norm.

(Refer Slide Time: 39:11)



So, from here, I define a norm as maximum because in this case, I am taking that this norm is always greater than equal to this value. So, if I take the maximum of this value, then this will be equal to this. So, I am taking A x over x and I am taking the overall x. So, from here, I can define this 1-norm and this norm, we are defining. $\| A \| = \max_x \frac{\| Ax \|}{\| x \|}$

So, this norm is called the subordinate norm and it is coming from the vector norms by taking this one. So, this is a subordinate to vector norm; subordinate to vector norm. So, I have taken the idea from here and then, we define it 2-norm, new norm that is repented by this one

$$\| A \|_2 = \max_x \frac{\| Ax \|_2}{\| x \|_2}$$

So, Frobenius norm, we have taken with respect to 2. Now, from that idea, I want to see what will be the 2-norm of a given matrix with respect to the 2-norm of a vector. So, this is a norm definition; definition of a norm. So, let us see how we can find this one. So, for example, now if you see from here, I am taking x taking its norm multiplying A x; then taking its norm and taking the maximum value.

So, if you see from here, then we are blowing up the value of x. Because I am taking the maximum values. For example, let us define what is going to be a 2-norm here. So, this is called the 2-norm of a matrix A, let us do. So, suppose I take a matrix A, $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ and I want to define its 2-norm. So, this value, I am going to take. It will be a maximum of x. Now, if I take the normalized vector, so that its norm is 1; then, so it is normalized vectors. Then, this becomes maximum or all the x whose norm is 1, 2. So, this is 2.

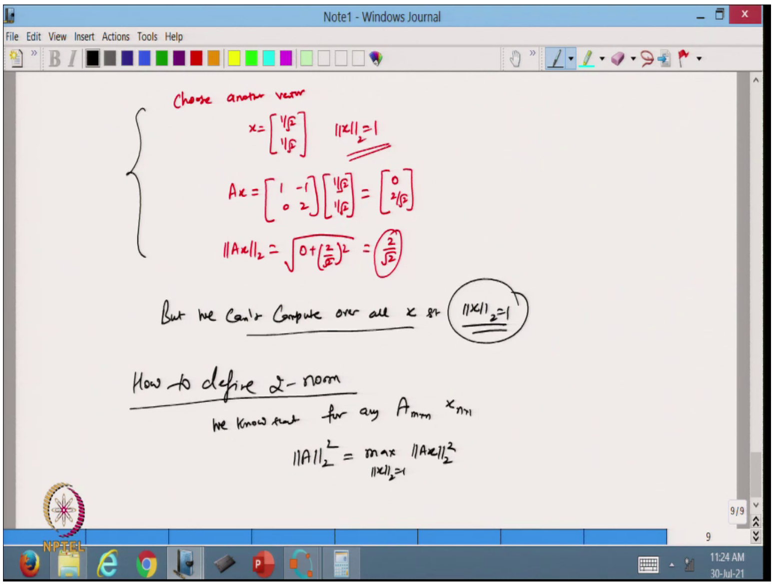
So, this is another definition, this one. Now, I will use this one. So, I have taken it here. So, let me take $x = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$. Why have I taken this one? Because its 2-norm is 1. Now, I multiply by $Ax = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{2} \\ -2/\sqrt{2} \end{bmatrix}$

$$\| A \|_2 = \sqrt{\left(\frac{2}{\sqrt{2}}\right)^2 + \left(\frac{-2}{\sqrt{2}}\right)^2} = \sqrt{\frac{4}{2} + \frac{4}{2}} = 2$$

So, now, my A x = 2. So, from here, I can define; so, I can define from here

$\frac{\| A \|_2}{\| x \|_2} = 2$ So, I started with a vector, whose norm was 1 and now, I defined this one and its value became 2.

(Refer Slide Time: 45:29)



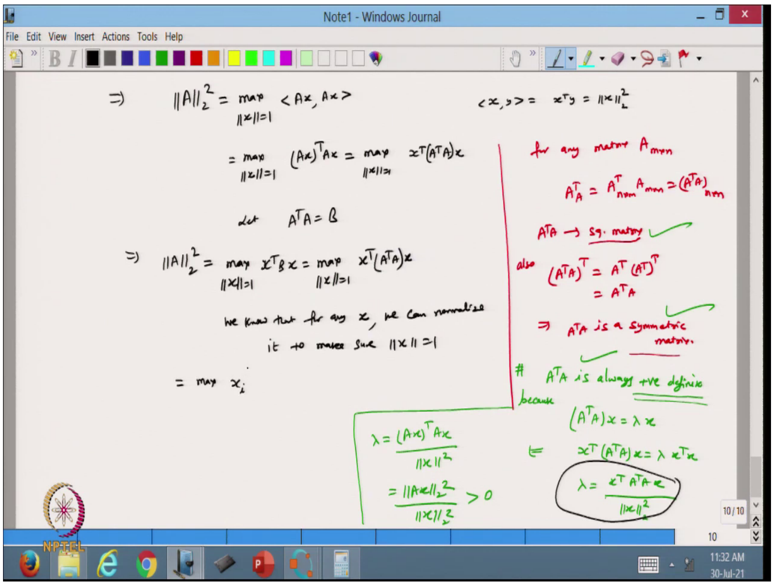
So, now, maybe I can choose another vector; choose another vector, suppose I take now x is equal to instead of 1, I just take $x = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$; I just take this vector norm. Now, I know that this 2-norm is 1 here. Now, A x will be again; so, it is $Ax = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 2/\sqrt{2} \end{bmatrix}$. So, this is in this case now, I am taking this value here. I take the 2-norm of this vector. So, if I take the 2-norm of this vector, then this is equal to $\| A \|_2 = \sqrt{0 + \left(\frac{2}{\sqrt{2}} \right)^2} = \frac{2}{\sqrt{2}}$

It is less than 2, but greater than 1. So, we have started with a vector and its value slightly increased to this one. Here we have started with a vector, whose norm is 1 and its value slightly changed and became 2. So, I will take all these vectors and then, take the maxima. So, this way, we can find out the norm; but we cannot compute all x such that the norm of x is 1 because there are an infinite number of vectors there.

So, these things, we cannot compute for all x to find out; but ultimately, we are reaching toward its maximum value and what is the maximum value we are going to take that will be the 2-norm in this case. So, from here, let us see how we can define the 2-norm here. Because this concept is ok, but it is not possible for us to find all these vectors x, whose norm is 1 and then, finding these values.

So, let us define how we can take the 2-norms in this case. Now, let us take this one value. So, the question is how to define 2-norm? This is what we are going to do ok. So, let us see we know that for any matrix A, it is supposed m cross n and my suppose n is n cross 1, let us define A x and take its 2-norm, then, so let us write in this way; I define 2-norm this one. So, that will be $\| Ax \|_2^2 = \max_{\|x\|_2=1} \| Ax \|_2^2$. This I can define by just taking the square.

(Refer Slide Time: 49:37)



Now, from here, this one I can write from that this is can be written as $\| Ax \|_2^2 = \max_{\|x\|_2=1} \| Ax \|_2^2$, then I can take the definition of inner product. Because this is coming from the Euclidean norm, from the inner product is a induced norm that I know, then I can write here that this is <Ax, Ax>, this value because we know that $<x,y> x^T y = \|x\|^2$ and it is coming as a 2-norm or maybe I can define it like this one.

So, this value, I am taking, is equal to maximum 1 and this is the inner product I am defining. So, I can write here $(Ax)^T$ Ax. This is again taking the maximum and I can write here $x^T(A^T A)$ x. Now, this matrix, I got that is A^T A. Now, from here, I know that now for any matrix that is A of order m cross n does not matter; A^T A will be always A^T will be n cross m and this will be m cross n.

So, A^T A is always n cross n. So, it is always a square matrix. So, A^T A is square matrix one thing is that; also A^T A, if I take the transpose, then this becomes $(A^TA)^T$ and that is A^T A and from here, I can say that A^T A is a symmetric matrix, this one.

So, now, it is always a square matrix; it is always a symmetric matrix. One more thing is there now I know that this is a symmetry matrix, I also say one more important property I am saying about this one that A transpose S is always positive definite. Positive definite means all its eigenvalues are positive. So, this is always true. How is it true?

Because I know that $(A^T A)x = \lambda$ x, this one I can write as an eigenvalue. From here, I can write as $x^T(A^T A)x = x^T \lambda$ x and from here, I can write that $\lambda = \frac{(Ax)^T Ax}{\|x\|^2} = \frac{\|Ax\|_2^2}{\|x\|_2^2}$ >0

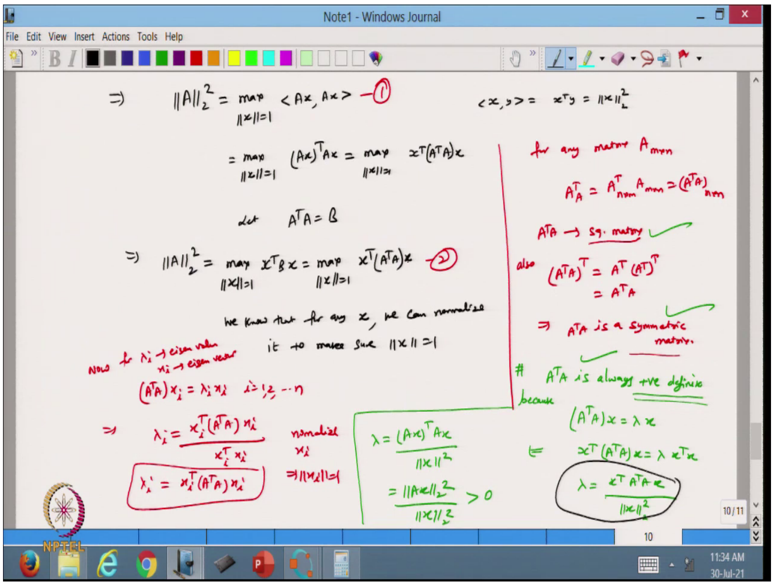
So, it is always positive and from here, if all the eigenvalues are positive, then we know that this is positive definite.

So, this matrix satisfies all this condition that it is a square matrix, it is a symmetric matrix and this is always positive definite. So, these things are defined here. Now, let I take A^T A=B. So, from here, I can write that A can be written as $\|Ax\|_2^2 = \max_{\|x\|=1} x^T Bx = \max_{\|x\|=1} x^T (A^T A)x$

this one I can write. Now, this matrix is a positive definite and everything is there, now from here I can write.

Now, from here, we know that for any x, we can normalize it to make sure that its norm is 1 that we can do. Now, from here, what I am going to write is that this thing I am going to write maximum. Now, this is my taking the same way as we have done here that any matrix can be written like this way. So, I can define this as maximum over all x i, suppose I am taking x i now.

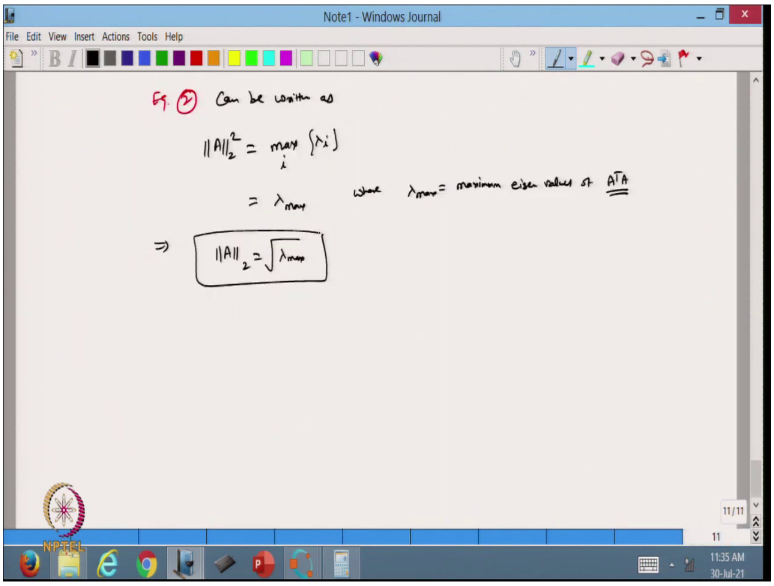
(Refer Slide Time: 57:09)



Now, A^T A, I can write as x_i for any eigenvector, I can define as $(A^T A)x_i = \lambda_i x_i$ for λ_i is the eigenvalue and x_i is the eigenvector. So, this concept we already know and it is a n cross n matrix. So, it will have a n number of eigenvalues and then, we can have the eigenvector; corresponding eigenvectors. So, this is true; i is equal to 1, 2, 3 up to n.

Now, from here, working with the same concept, I can write $\lambda_i = \dfrac{x_i^T (A^T A) x_i}{x_i^T x_i}$ Now, we can normalize; normalize x i that becomes 1. So, from here, I can write that this $\lambda_i = x_i^T (A^T A) x_i$ So, from here, these things can be written. So, I can write this as 1, 2.

(Refer Slide Time: 58:41)



So, I can write equation 2 can be written as now I am taking the maximum and this can be written as a lambda i's because I have taken shown that $\lambda_i = x_i^T(A^T A)x_i$ And that I took the maximum of all the x i's and that is this one lambda i's. So, this I can write the lambda i's over all i's because I am taking the maximum or all x i and that is the eigenvectors and from here, I can write that this is equal to lambda maximum, where lambda maximum is maximum eigenvalues of A transpose A and from here, I can define 2-norm as square root of maximum.

So, this way, we can define the 2-norm of a given matrix; this way. So, we will stop here now. So, in the today lecture, we have started with the matrix norm that how we can define different type of matrix norm, then we have discussed that how matrix norm and the vector norms are compatible to each other and then, in the end, we have discussed that how we can define the 2-norm of a given matrix in the terms of the eigenvalues. So, in the next lecture we also continue with this one.

So, thanks for watching.

Thanks very much.