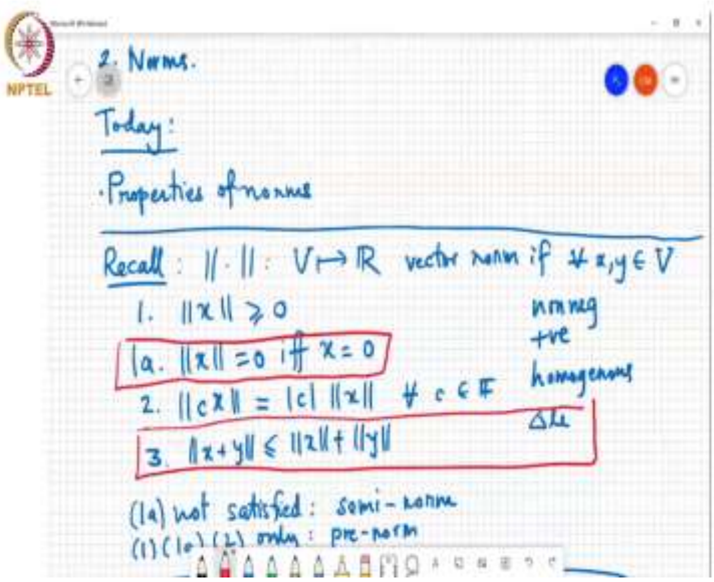
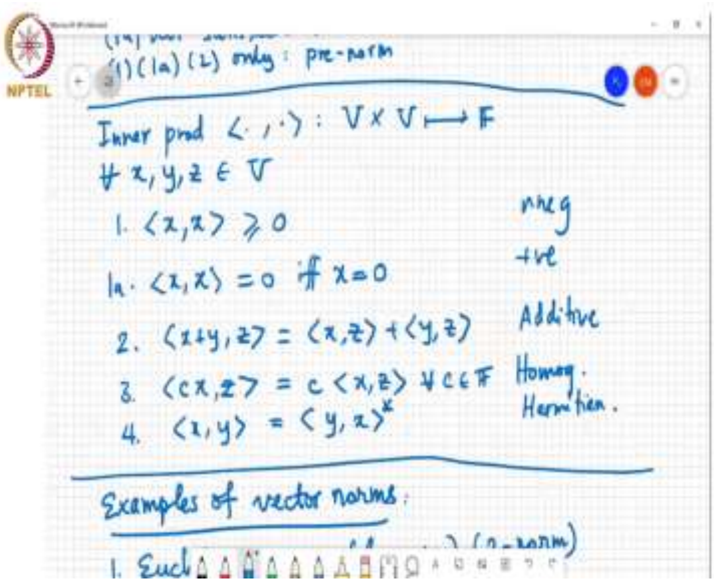
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Technology, Bangalore
Vector norms and their properties

Yeah, I mean, I think two weeks after the classes have begun, begun, you should know which meeting to join. So, please join the right meeting. So, we will begin. So, the last time we were discussing about, we discussed a bit about determinants. And then we started discussing about norms. And today we will discuss several properties of norms.

(Refer Slide Time: 00:45)





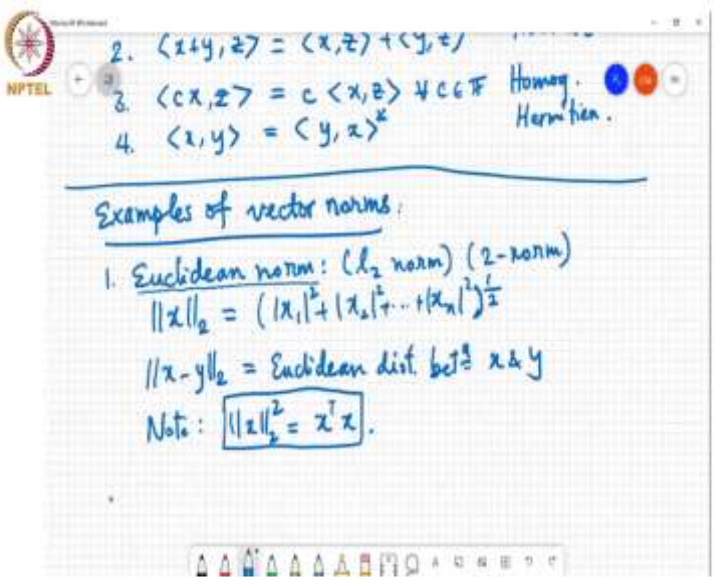
So, yeah, so, today we will discuss several properties of norms. Recall that we have this right starting point, recall that something is a norm provided for any x and y belonging to this vector space V over which the norm is defined, the norm of any vector x is greater than or equal to 0. And it is equal to 0 if and only if x is 0.

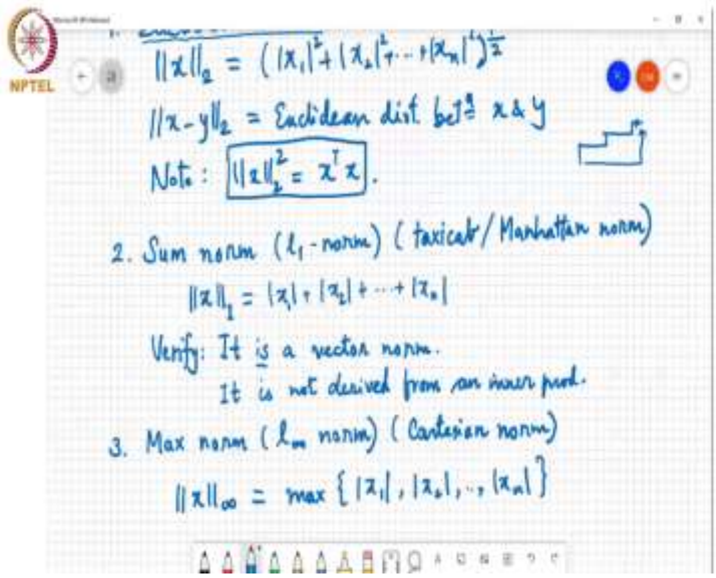
It has the homogeneity property naming namely that the norm of C times x is equal to the magnitude of C times the norm of x for every C in this field F. And finally, the triangle inequality norm of x plus y is less than or equal to norm x plus norm y. If property three is not satisfied, then we call it a pre-norm. And if property 1a is not satisfied, then we call it a semi-norm, also recall the definition of an inner product, the inner product is defined like this, if it is defined from two points, you have to choose two points in V.

And then it maps those two points to the field F. So given any x, y and z belonging to V, the inner product of x with itself is non negative and the inner product is equal to 0 if and only if x is equal to 0, the inner product is additive. So, if you have x plus y comma z, then that is x comma z plus y comma z. And it is homogenous in the first argument, namely, that Cx z inner product is the same as C times x, z for every C in the field F and it is also Hermitian.

If you exchange the order, you do inner product of y x, you get the complex conjugate of the inner product between x and y. So also, we saw one crucial property that if this some dot coma dot is, an inner product, then x comma x power half the inner product of x with itself power half is a vector norm on V. So, that is one crucial property that connects inner products to norms. So, using any inner product, you can define a norm.

(Refer Slide Time: 03:30)





Now, today, we will start by discussing some example norms, these are perhaps the most popular norms that are used in different applications. The most popular norm is the Euclidean norm, which is also known as the l2 norm or more simply as the two norm. It is just the sum of the squares of the entries of x raised to the power half. So, one important, simple formula is that norm x squared, this is called the l2 norm.

So, l2 norm of x equal to x transpose x. It is a very useful formula. And it gives you an algebraic way of writing this particular Euclidean norm, the Euclidean norm of x, the norm of x minus y l2 norm of x minus y measures the Euclidean distance between x and y meaning the our conventional notion of length between x and y. The second norm I want to discuss is what is known as the l1 norm or the sum norm.

And this is also known as the taxicab norm or the Manhattan norm. So, this the Manhattan area of New York is famous for having perfectly rectangular streets divided into perfectly rectangular grid. And so, this essentially measures if you are given a point A and point B, you have to go. It is like this grid, you see in the background here. So, if you have a point here and another point here, the way you can go from this point to this point, you can go like this, or you can take… and go like this.

But however, you go, the total distance you will traverse is actually the same, as long as you are going along the sides of this grid. And that is basically this sum norm. So, it is equal to it is written as norm xl1 and it is equal to mod x1 plus mod x2 plus etc, plus mod xn, the sum of the magnitudes of the entries in x. So, small exercise for you is to verify that this is in fact a vector norm, that means it must satisfy those four properties that we discussed just now.

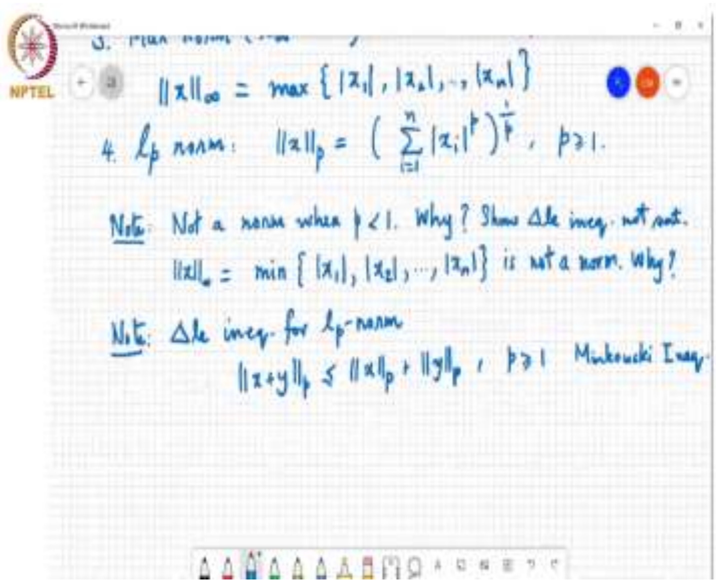
And another property is that it is not derived from an inner product. The third norm is what is known as the max norm or the l infinity norm. This is also called the Cartesian norm. And this is written as x infinity. And the reason for the subscripts will be obvious in a second, it is the max of, first the max magnitude entries of x, the largest magnitude entry next, that is the norm.

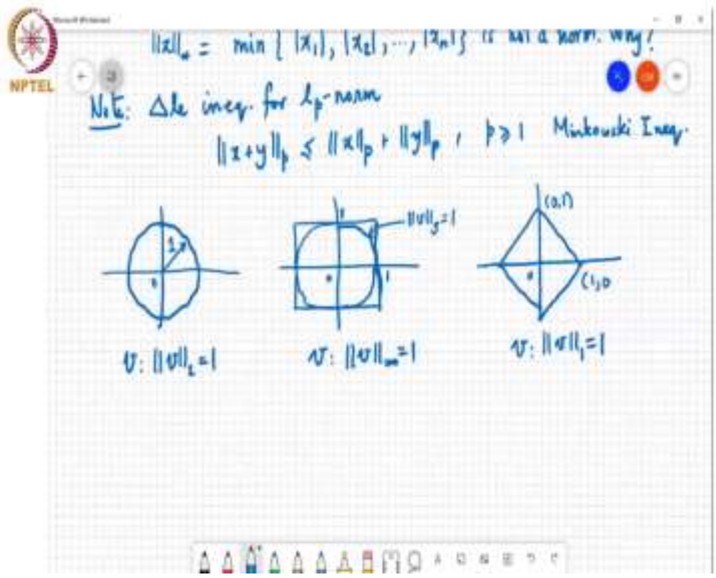
So, one thing is that if you I mean, if you think about it, these are all different ways of measuring the length of a vector. And taking the sum of the squares and taking to the power half is one way to measure the length of a vector. And in a two-dimensional space, if I draw a vector, say like this, then the length of the vector is actually this squared plus this squared power half, we are just using Pythagoras theorem to say that is the length of this vector.

And so, it is a reasonable way of measuring the length of a vector. Similarly, the elbow norm it is measuring the length you have to travel if you were restricted to go along the sides, and the l infinity norm, essentially picks off the biggest entry in x in magnitude. And that is also I cannot give you an example of how that will be the length but the, but you can imagine that maybe the cost is completely dominated by the largest segment in one of the dimensions that you need to traverse and therefore, that is the l infinity norm.

However, for example, if I took the min here, the min of the magnitude entries of x, that is not a norm. Can anybody think why?

(Refer Slide Time: 08:56)





Student: So, is this second one where we are taking the meaning is it due to the second property of positivity?

Professor: Yeah, that is one, you can also show that. Yeah, so certainly positivity does not hold, if any one entry is 0 then the star is going to be equal to 0. So, it would not be 0 only if the if all the entries are 0, you can also probably show that it does not satisfy the triangle inequality. That is also easy to show. So, similarly, this when p is less than 1, this definition of lp norm, it does not satisfy the triangle inequality. So, show this can show.

So, the triangle inequality for the lp norm for p greater than or equal to 1, it basically reads norm x plus y, and p is less than or equal to norm x p plus norm y, p. Again, it is something to think about how you show this, for any p greater than or equal to 1, if you define the lp norm like this, then it satisfies this triangle inequality. And this inequality is called me Minkouski inequality.

So, here in this inequality, if I substitute in this definition of lp norm, if I substitute p equals 1, then it is mod xi power 1 whole power 1. And so that reduces to the sum norm. And if I take p very, very, very large, then what happens is that, when I am taking mod xi to a very large power, and I am adding them up across all the excise, the largest magnitude of a magnitude entry in the vector x will completely dominate this sum.

And therefore, they this, the value of the sum is equal to Sp tends to infinity, the value of this sum will be equal to the magnitude of the largest entry of x raised to the power p, and then I am taking it to the power 1 over p. So, I this will lead to the largest entry in magnitude in x as the lp norm as p tends to infinity and that is the reason for this notation.

Norm x infinity equals the maximum of these entries, so this p norm, it reduces also when p equals 2, it is the sum of the squares of the magnitudes of x and the entries of x raised to the power half which is exactly the same as the Euclidean norm.

So, this is a generalization that includes the l1 norm l infinity norm and l2 norm as its special cases. So, now also to just get a feel for how these norms look like one can ask, what is the… So, you can look at a two-dimensional space and ask what is the locus of points that have a fixed norm? So, if I take the l2 norm, if I take the set of points V, such that norm V l2 equals 1 on the two-dimensional plane, what will it look like?

Student: (())(14:54)

Professor: It is a circle. So, assume that this is a circle, and its radius will be equal to 1. And if I take the set of points such that V such that norm V infinity equals 1. So, the norm is now the largest entry. And so, basically the, what will that look like? You are fleet footed you can think about it.

Student 1: Squares.

Student 2: Squares.

Professor: Exactly, so that will look like a square. So, for any point along this line here, the largest, so whatever the value of y the value of x is equal to 1. And so, so this is 1 and this is 1, and this is the origin. And for any point along this line, the x value is equal to 1 and so, the infinity norm of any point along this line is equal to one similarly, any point along this line, the l infinity norm is 1 like that. So, that is how you get the square. And finally, if I take all the V such that l1 norm of V equals 1 what shape will I get?

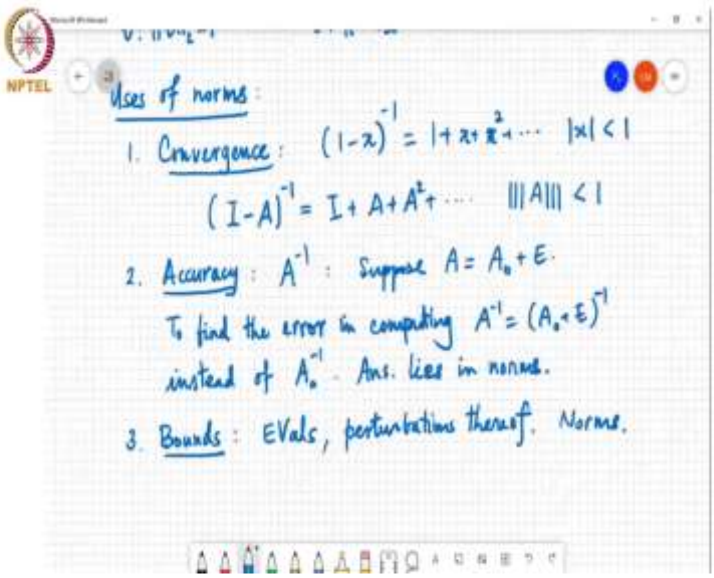
Student: A 9.

Professor: Yeah, I will get a…

Student: 4th point.

Professor: I will get a diamond; I just call it a diamond simplicity. And these are points. This is point 1 comma 0 and this is 0 comma 1, this is the origin. So, for any point along this line, the sum of these two coordinates is always equal to 1. And so that is how you get this diamond shape. So, that is kind of the shape of these norms. And if I take the l3 norm or l4 norm that will be like a circle that is further bulged out, it will end up looking a bit like this. It is not quite a circle is bulged out compared to a circle. So, I am trying to draw it a little more bulged out. So, this could be like the. So, that is how these will look.

(Refer Slide Time: 17:46)



Now, what are norms good for? There are several things that they are good for. So, I will just give some examples. So, one very important use is for showing convergence. So basically, for example, we know this formula that if I take 1 minus x inverse, or 1 over 1 minus x, and I can write this as 1 plus x plus x squared, plus etc, it is an infinite series. Now, when is this true?

Student 1: x less than 1.

Student 2: x infinity, x to the power infinity tends to 0.

Professor: So, x mod x should be less than 1 and that is the magnitude of x should be less than one. So, this is true for a scalar. But suppose I wanted to find identity minus of matrix A inverse and so when can I write it as i plus a plus a squared plus etc. Now, obviously, this condition here suggests that maybe we should we need a condition on somehow the size of this matrix A? And the answer is that, this is true if a matrix norm on A which I am going to write with three lines.

So going forward, I will use three lines to denote matrix norms. And I need to tell you in what which norm I will use here, and it turns out that any matrix norm will do and if you can find a matrix nom, under which the norm of A is less than 1 then a formula like this can be used to compute the inverse of i minus A.

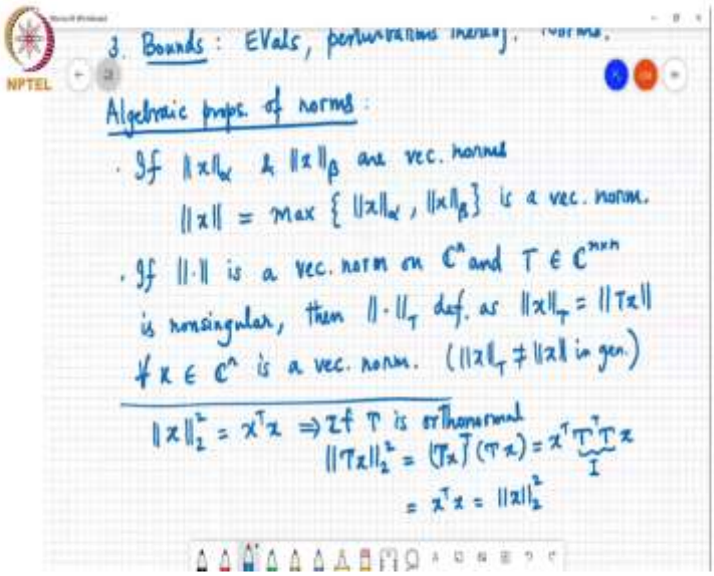
And the other use is if you know that mod x is less than 1, you can actually bound how much how big the rest of the series will be. And in turn, you can determine how many terms you need to use in the summation in order to get a sufficiently accurate estimate of 1 minus x inverse and similarly the norm of A will tell us how many terms I need to include in the series in order to get a sufficiently accurate representation of i minus A inverse.

And so, in in in a more general sense, it is useful for determining how many iterations of an iterative algorithm you need to use to solve a certain problem to a desired level of accuracy. And in fact, the second use is more about quantifying the accuracy of matrix computations. And these are again things that we are going to look at later in the course when we look at stability of matrix computations.

So, suppose we want to find A inverse, but instead the entries of A are noisy and so, what we get to see is suppose, so suppose A was equal to A naught plus E. And so what we have done is we have gone and completed A inverse, but what we really want is A0 inverse, then we want to know what is the potential error that we have incurred by computing A inverse.

So, to find the error in computing A inverse, which is A0 plus E inverse instead of A0 inverse, and again the answer lies in the norms. And the third use is in bounding Eigenvalues or perturbations of Eigenvalues. So, if you perturb a matrix by adding a small error matrix to it, how much will the eigenvalues get perturbed and all of these the answers lie in norms. So, this is also something that we are going to see later in the course.

(Refer Slide Time: 23:46)



Now, another thing is that we have seen a few kinds of norms, but the question is, can we can we come up with new norms based on existing norms that we know and so, for example, so these are you can do that and these exploit some properties, which are known as algebraic properties of norms. And I will give two examples here.

So, the first is that if say x alpha and x beta are vector norms. Then, if I define norm x to be equal to the max of these two numbers, then this is a vector norm again a property that is easy to verify, but you have to check that it satisfies those four properties of norms. Similarly, if this is vector norm to say on C to the n and T, and C to the n cross n is non-singular. It is a non-singular n cross n matrix.

Then, if I define this, so we will call it the T norm, this is equal to the norm of Tx, then for any x belonging to C to the n is a vector. So, you can produce lots of different norms, for example, you have a set of norms, you can take the max of any pair of them or any number of them, then you get another vector norm, take any matrix t in C to the n cross n that is non-singular then the norm of Tx gives you yet another norm, obviously the length of Tx is going to be the different from the norm of x.

So, in general, so in particular, for example, if I am taking the Euclidean norm, and if T is a unitary matrix, then xT will be equal to norm of x, but otherwise it may not be equal. So, yeah, so we can produce lots of different norms like this, but the question, now the question is, where do we use these different norms? In particular the…

Student: Sir?

Professor: Yeah, go ahead.

Student: Do orthogonal matrices only preserve Euclidean norms or every norm?

Professor: What do you think?

Student: Every norm.

Professor: No, orthogonal matrices preserve ortho-normal matrices?

Student: Yeah, orthonormal matrices, will they present Euclidean norm?

Professor: It preserve Euclidean norm only. And the reason is very simple.

Student: Because it is (())(27:25) to be inner product, right?

Professor: Inner product, in fact, what we call the usual inner product, so you can write this where it is equal to x transpose x. So, which implies if T is orthonormal then if I compute T x, this is going to be Tx transpose. So, this is bad notation. So, this is Tx transpose Tx, which is equal to x transpose T transpose Tx. And T transpose T is the identity matrix for orthonormal matrix. And so, this is equal to x transpose x, which is equal to x 2 square.

But for other norms, you cannot write it like this. And so, it is not true that it preserves other norms. So, you can in fact, ask, are there classes of matrices that preserve for example, the l1 norm, or that preserve the l infinity norm. And in general, it is hard to find matrices that will mean you can always find a matrix that will preserve the l1 norm or l infinity norm for a particular vector. But for any x, you cannot preserve its l1 norm or l infinity norm by multiplying it by an n cross n matrix. So, that is again something that you can try to show.