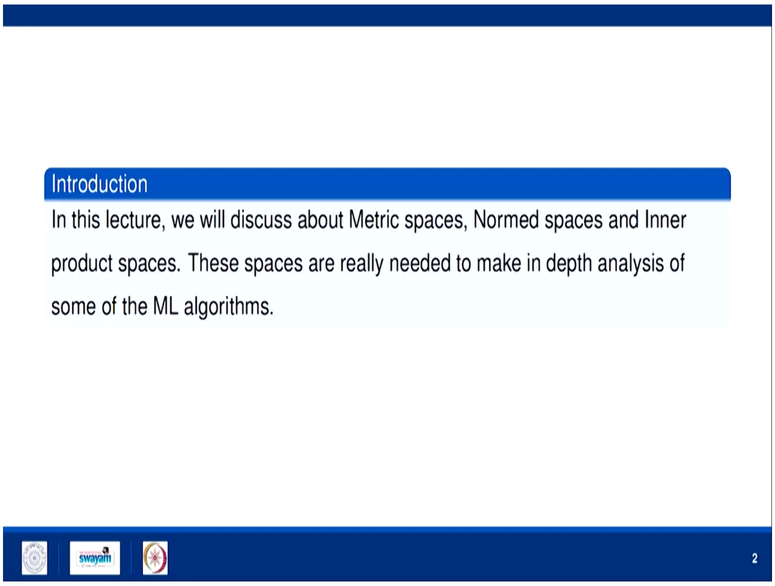
Essential Mathematics for Machine Learning
Prof. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture - 06
Norms and Spaces

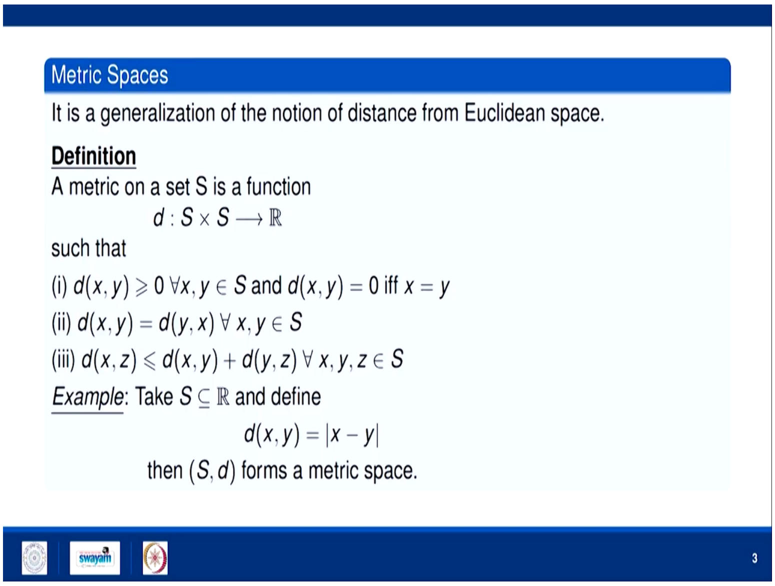
Hello friends. So, welcome to the 6th lecture of this course: Essential Mathematics for Machine Learning.

(Refer Slide Time: 00:29)



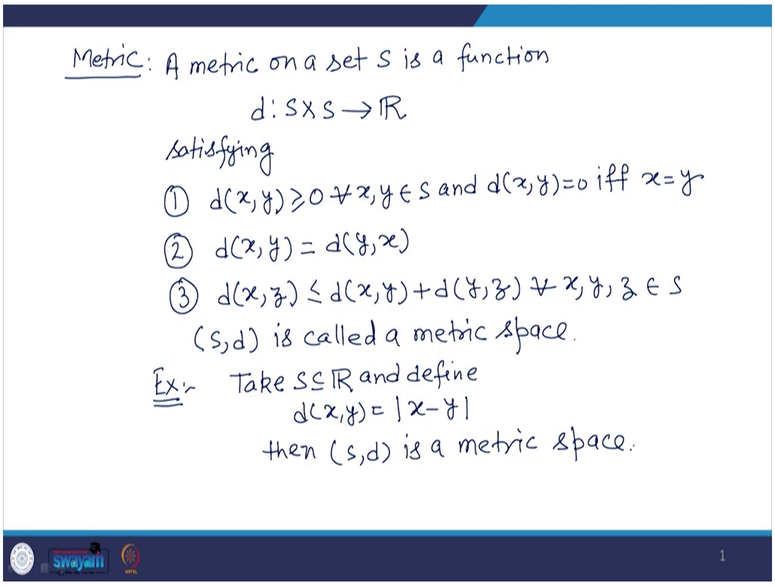
In this lecture, we will discuss about metric spaces, normed spaces and inner product spaces. These spaces are really important to make in depth analysis of some of the machine learning algorithm.

(Refer Slide Time: 00:46)



So, first we are going to define metric spaces. So, metric spaces is a generalization of the notion of distance from Euclidean space. So, like you have seen in earlier mathematics, school mathematics that how to find out distance between two points in R 2 R 3 or R n. So, if you generalise this concept, we define metric. So, my first definition is about metric.

(Refer Slide Time: 01:21)



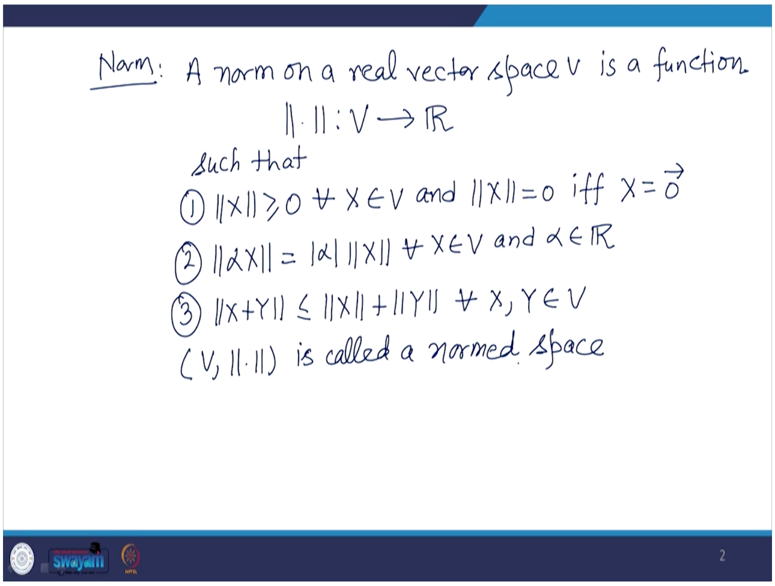
So, a metric on a set S is a function d, defined from S cross S to set of real numbers satisfying certain properties. And what are those properties?

The 1st property is that d of x y is always non negative. For all x y belongs to S and it is 0 if and only if x equals to y. The 2nd property is symmetry that is d x y equals to d of y x. And the 3rd property is triangle inequality that if you are having three elements from the set is x y and z. So, the d of x z is always less than equals to d of x y plus d of y z. For all x y z belongs to S.

So, if a function from S cross S to R satisfy these three properties, then we say that the function d is a metric. Now, the set S together with this metric is called a metric space. If you want to see an example of a metric space. So, take the set S a subset of set of real numbers and define the metric d of x y equals to absolute value of the difference between x and y. Then, S d is a metric space.

So, as I told you d is a generalisation of the notion of distance. Similarly, we are having definition of norm. So, norm is a generalisation of notion of length.

(Refer Slide Time: 04:44)



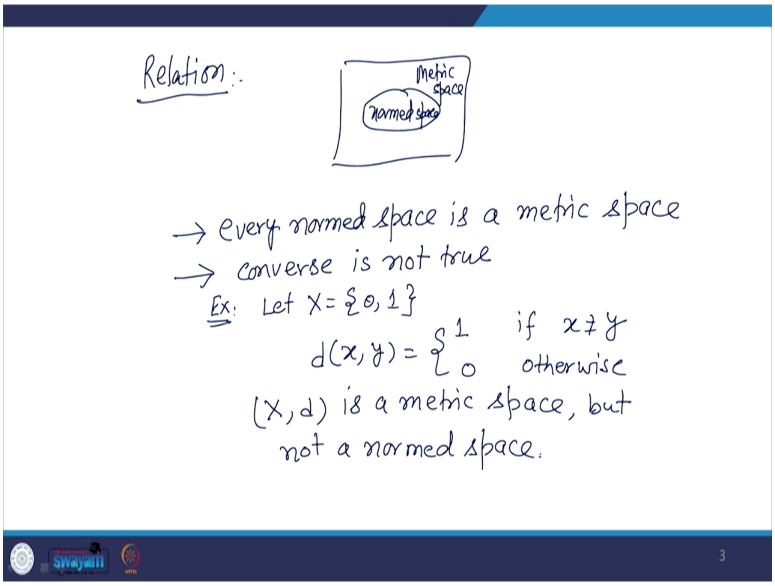
So, let us come to norm. So, a norm on a real vector space. Why we are taking real vector space? Because we are talking in a sense of machine learning. So, space V is a function denoted by this symbol and it is a function from vector space V to the field R; such that this function satisfy certain properties.

And these properties are; the 1st property is, the norm of a vector is non negative for all X belongs to V and the norm is 0 if and only if the vector itself is a 0 vector.

2nd property is, if you multiply the vector X by a scalar alpha then, it is equals to absolute value of alpha times the norm of X. For all X belongs to the vector space V and alpha belongs to the field R. The 3rd property we are having that, the norm of the sum of two vectors X and Y is less than equals to the norm of X plus norm of Y, for all X Y belongs to V. So, it is again triangle inequality.

So, the vector space V together with the norm is called a normed space. In some of the reference is, we write it normed vector spaces or normed linear spaces. Now, what is the relation between metric space and normed space?

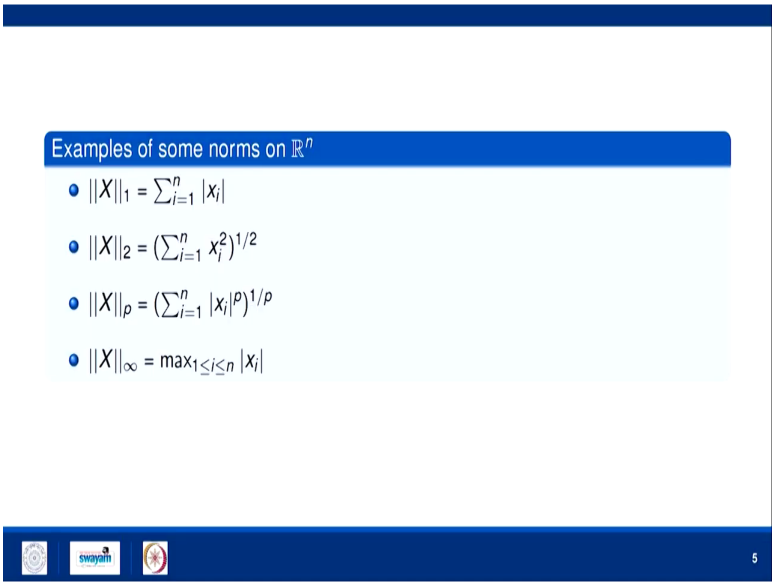
(Refer Slide Time: 07:32)



So, let us try to make a relation between these two definition. So, what I want to say that? Every normed space is a metric space, but reverse is not true. So, every normed space is a metric space, converse is not always true. So, if you take a example of this converse where we are defining a metric space, but that is not a normed space. So, let X be a set containing two element 0 and 1. So, it is a subset of real numbers.

Now, define the metric d x y as 1 if x naught equals to y and 0 otherwise. Then X together with d is a metric space, but not a normed space. You can verify it by using the definition.

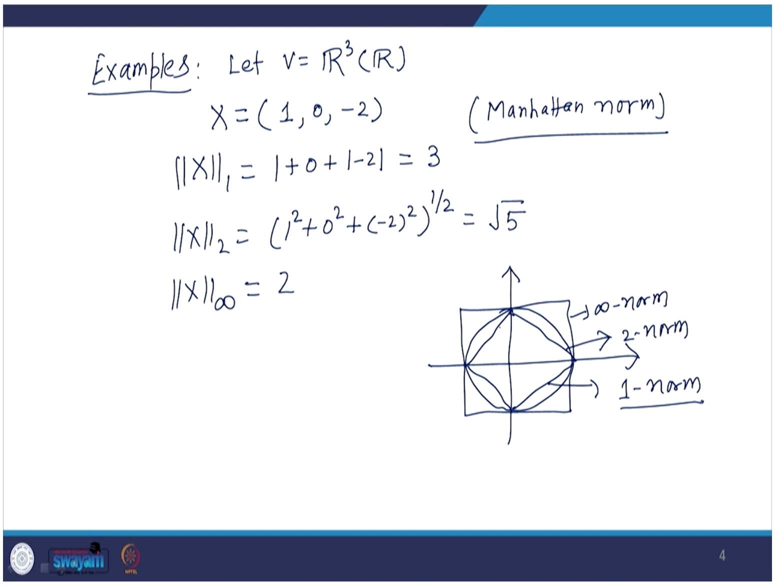
(Refer Slide Time: 09:24)



So, if you are want to see some of the examples of normed space. So, take the vector spaces R n, defined on over the field of real numbers. And then define the norm h l 1 norm. So, this is called l 1 norm. So, norm of X equals to some of the absolute values of the components of the vector. So, it is your l 1 norm. So, this l 1 norm together with the vector space R n forms a normed space.

Similarly, we are having the l 2 norm which is also called Euclidean norm. So, it is define like: i equals to 1 2 n x i square and then, whatever sum of square you are having the square root of that. Then we are having p norm. So, again you are taking the power p of the absolute value of the component of vector X and sum of all those. And then we are taking the p th root of the sum.

(Refer Slide Time: 10:39)

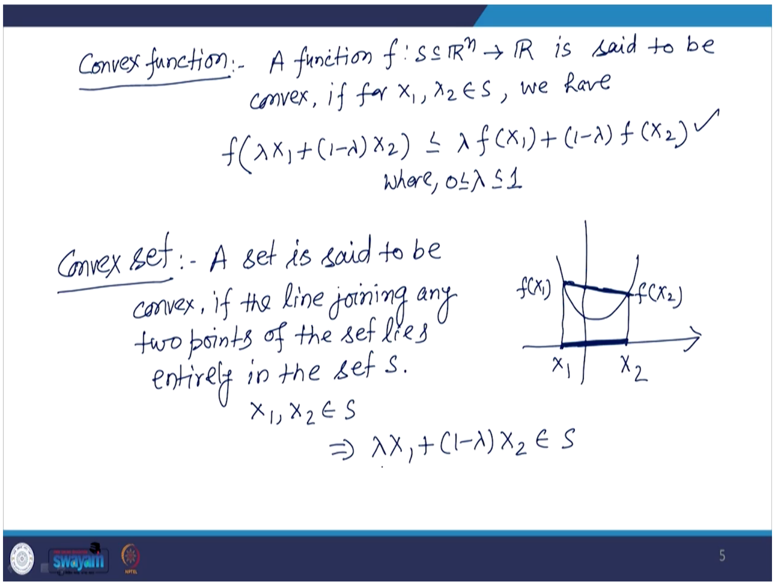


So, if you want to see example. So, some of the examples. So, let the vector space is R 3 defined on R. Now, take a vector X in R 3 form by the 1 0 and minus 2. Now, what is X 1 norm. So, X 1 norm is means. What is one norm? It is 1 plus 0 plus absolute value of minus 2. So, the value comes out to be 3. This is also called l 1 norm as Manhattan norm.

If you want to find the Euclidean norm here of the vector X, then it will become 1 square plus 0 square plus minus 2 square and then square root of this. So, it becomes square root of 5. If you want to find out the infinity norm here or max norm, then it is maximum of the absolute values of the component. So, absolute value of 1 is 1 0 and 2. So, it comes out to be 2. So, in that way you can calculate norms.

Furthermore, if you want to see geometrically. So, this square is a max norm infinity norm. If you talk about Euclidean norm. So, it will be this ball here. And then your l 1 norm will be like this. So, geometrically we can see these three norms in this way. Now, one of the important property of the norms are all the norms are convex. How?

(Refer Slide Time: 13:09)



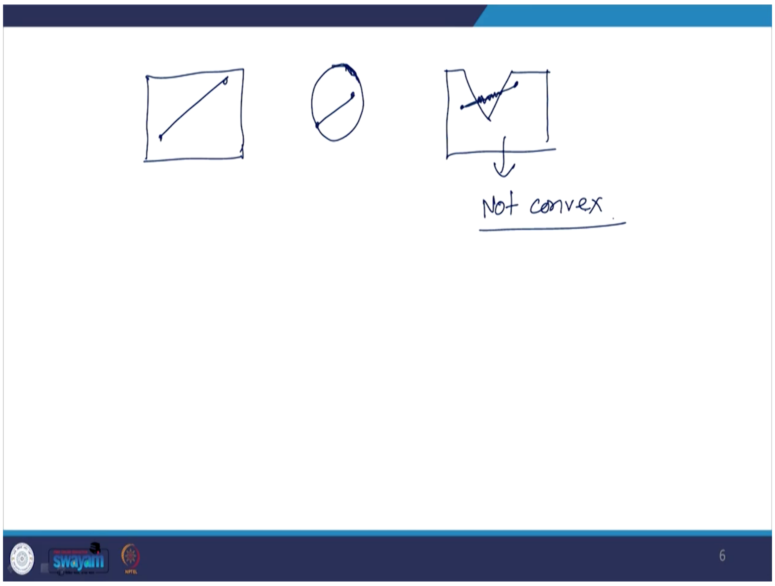
Let us first define a convex function. So, a function f S which is a convex of set of R n to R is said to be convex, if for X 1 X 2 belongs to S, we have the value of lambda time X 1 plus 1 minus lambda times X 2 is less than equals to lambda times f of X 1 plus 1 minus lambda times f of X 2.

So, mathematically we can say a function we satisfy this property is called a convex function, where lambda is between 0 to 1. So, geometrically, what it is saying? It is saying that a function if you take let us say like this. So, let me see this as X 1 this point as X2. Now, what is right hand side? So, this point becomes f X 1 and this point is f X 2.

So, what is the right hand side? What right hand side is a code joining f X 1 and f of X 2. Now, what this particular inequality is saying that the this line is; this line is lambda X 1 plus 1 minus lambda times X 2. So, what this inequality is saying that, the functional image of this code joining X 1 and X 2 is always lie under the code joining f of X 1 and f of X 2.

So, if you are having a function like we satisfy this property for all X 1 X 2 belongs to S, then we say that function f is a convex function. Similarly, a convex set is. A set is said to be convex, if the line joining any two points of the set lies entirely in the set S. It means if I am having X 1 X 2 belongs to S, then lambda of X 1 plus 1 minus lambda times of X 2 is also in S.

(Refer Slide Time: 16:56)

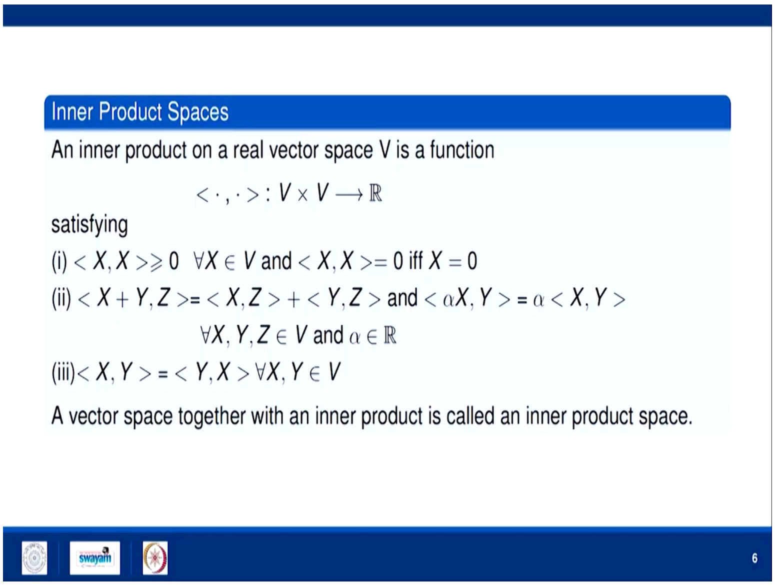


So, if the point of a set is satisfied this property for all arbitrary point X 1 and X 2 then, the set is said to be convex set. Geometrically you can see, if I take a rectangle like this. You take any two points in this rectangle. The line joining these two points will be entirely in this rectangle. So, this rectangle is a convex set.

Similarly, if you take a ball circular ball, then it is again a convex set. Because you take any two points in this ball, the line joining these two points will lie entirely in the circle. If you are having a set like this, then this set is not a convex set. Because if I take a point here and a point here, then the line joining these two points are not entirely inside the set.

So, this portion of the line is outside the set. So, it is not a convex set. So, this is the definition of convex set and convex function. Now, if you come back. You can see from the geometry of all these l 1 norm, l 2 norm, and l infinity norm that, all these three are making the convex set. So, hence all norms are convex.

(Refer Slide Time: 18:20)

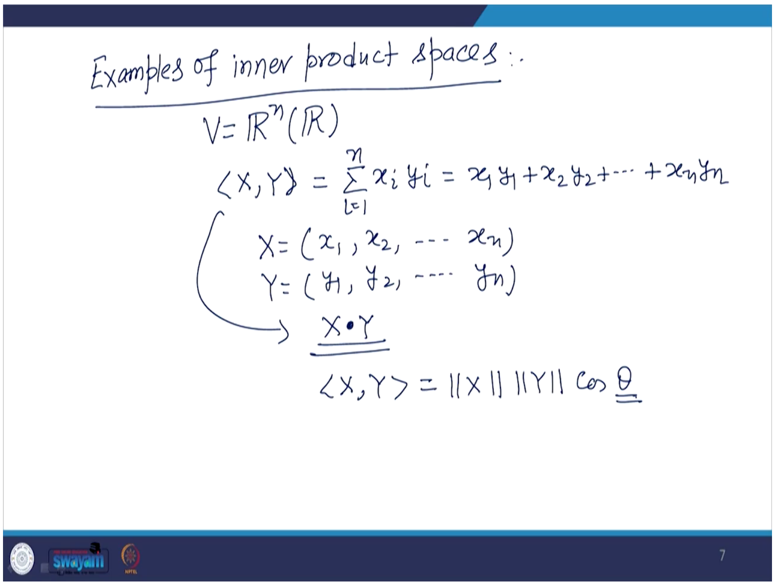


My next definition is inner product spaces. Inner product spaces are really important in terms of machine learning and analysis of any classifier. So, an inner product on a real vector space V is a function denoted by this symbol. So, it is acting on two elements of V.

So, it is from V cross V to R, satisfying that X belongs to V inner product of X X is always greater than equals to 0. For all X belongs to V and it is 0 when X is 0. If you take the inner product of vectors X plus Y and Z, then it is equals to the inner product of X and Z plus inner product of Y and Z. And if you take the inner product of alpha X and Y it is equals to alpha times X comma Y, for all X Y Z belongs to V and alpha belongs to R.

The third property symmetry that the inner product of X and Y equals to inner product of Y and X, for all X Y belongs to V. A vector space together with an inner product is called an inner product space. So, if you want to see the example of inner product, then we can have like this.

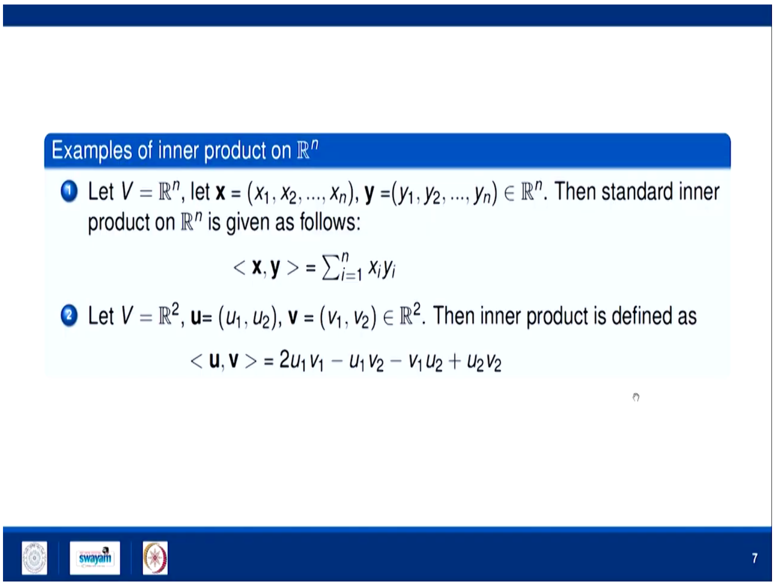
(Refer Slide Time: 19:59)



So, take the vector space as R n define over the field R. So, R n real vector space of n tuples. Now, you define inner product of X Y belongs to R n as summation i equals to 1 to n and x i y i. Here X is the vector having component x 1 x 2 x n. And the vector Y is having components y 1 y 2 up to y n. So, then this will become x 1 y 1 plus x 2 y 2 plus x n y n. And what is this? It is nothing just usual dot product of the vector X and Y.

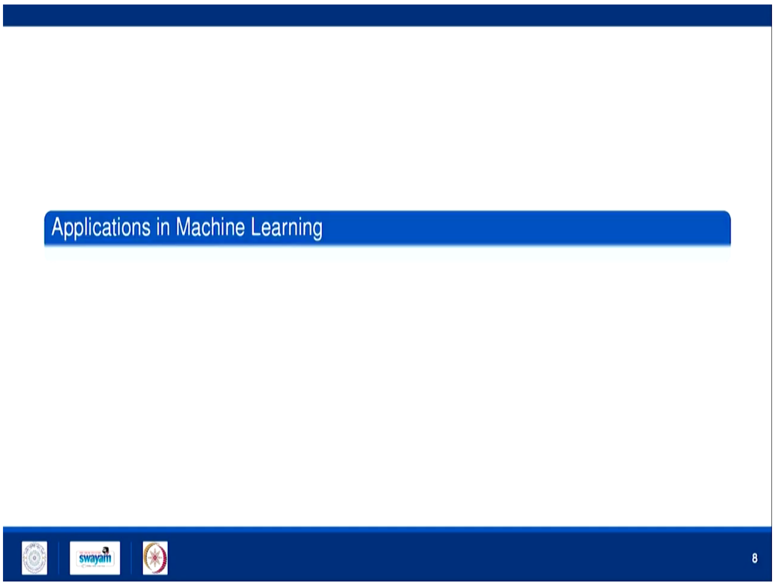
So, usual dot products of vector is an inner product. So, R n together with this inner product forms a vector inner product space. Here, what we can have? The angle between X and Y is defined as norm of X into norm of Y into cos theta, where theta is the angle between the vector X and Y.

(Refer Slide Time: 22:07)



My another example of inner product spaces from R 2. So, take the vector spaces R 2 and take two element two vectors in R 2 as u 1 u 2 and V as v 1 v 2. Then define the inner product like the inner product between u and v is defined like this 2 u 1 v 1 minus u 1 v 2 minus v 1 u 2 plus u 2 v 2. You can verify all those three properties of inner product using this definition of the inner product. R 2 together with this inner product forms a inner product space.

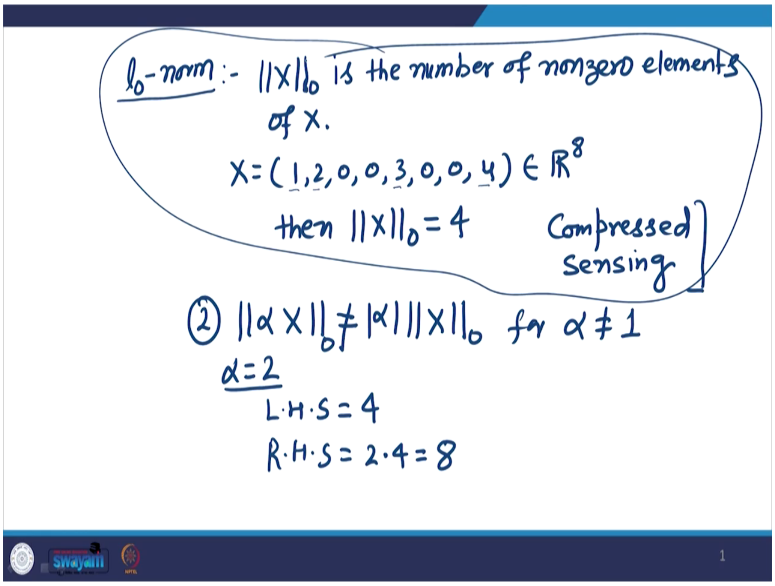
(Refer Slide Time: 22:49)



So, as I told you the metric are generalisation of the notion of distance. Norms are the generalisation of the notion of the length. And similarly inner products are generalisation of the dot product of the vectors which you have seen in your plus two. In that way, when you are having your feature space in machine learning and you are having feature vectors. So, these three concepts are really important to make the analysis of any particular machine learning algorithm or to develop any machine learning algorithm ok.

So, I have defined Manhattan norm, Euclidean norm, p norm, and maximum norm. There is one more very interesting norm which is quite useful in machine learning and that is called l 0 norm.

(Refer Slide Time: 23:42)



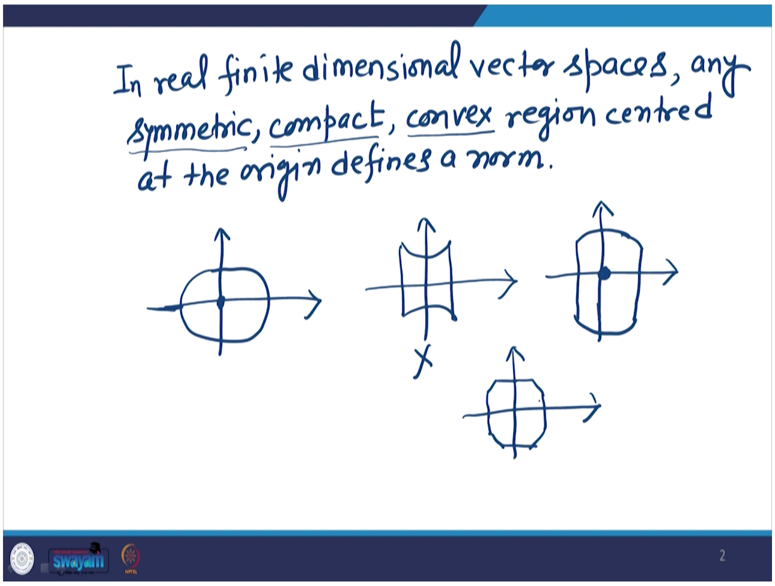
So, strictly it is an it is not a norm I will tell you, why? But we defined this norm as the. If you take a vector X then 0 norm of vector X is the number of nonzero elements of X. That is if you are having a vector X like this, 1 2 0 0 3 0 0 4 in belongs to R 8. Then, what is the 0 norm of this vector? It is number of nonzero elements in this vector. So, 1 2 3 and 4. So, this value is 4.

So, this norm is very much important in compressed sensing, where you are looking for a sparse solution of your data. Why I am saying strictly it is not a norm? Because, if you remember the second property of a norm we have told it, if you are having a scalar alpha and a vector X then, it is equals to the absolute value of alpha times norm of X.

Now, if you see here. If I take alpha equals to 2, then what is left hand side? Left hand side is having the same value 4, because if you multiply each of the element of this vector by 2, then it will become 2 4 0 0 6 and so on. So, it still it will be having only four nonzero elements.

However, right hand side will become 2 into 4 that is 8. So, l 0 norm is not satisfying this property for alpha not equals to 1. So, it will satisfy this property only when alpha is 1 or minus 1, otherwise it will not satisfy. However, we are using this particular 0 norm quite frequently in nowadays in, compressed sensing and machine learning.

(Refer Slide Time: 26:53)



Further, you can define your own norm. So, in real finite dimensional vector spaces, any symmetric, compact, convex, region centred at the origin defines a norm. So, here if you see this I have written symmetric. So, you know symmetricity very well. We have defined convex also convex region also in this lecture. The term compact is a bit more mathematical and if anyone of you are interested you can see it you can check it from any mathematics books.

So, when one a set is a compact set. So, for example, if you take this region this region is symmetric, centre is at origin ok. Compact and convex.

So, it defines a norm. Furthermore, if you take something like this. So, it is this region is not a convex region. So, it will not give you a norm, but if you are having region like this, then it is a convex region, just centre at origin and symmetric about all the axis.

Furthermore, if you take one more geometry something like this. So, this is again defines a norm. So, in that way you can define your own norms on real finite dimensional vector spaces. So, thank you for this lecture. In the next lecture, we will talk about a very important property of in inner product spaces and in general for vectors spaces, that is orthogonal vectors and projection transformation.

Thank you.