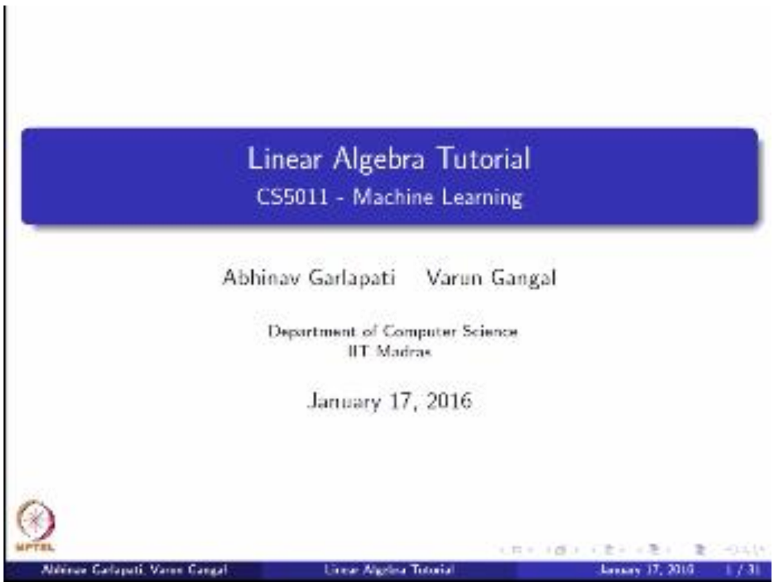
NPTEL
NPTEL ONLINE CERTIFICATION COURSE

Introduction to Machine Learning

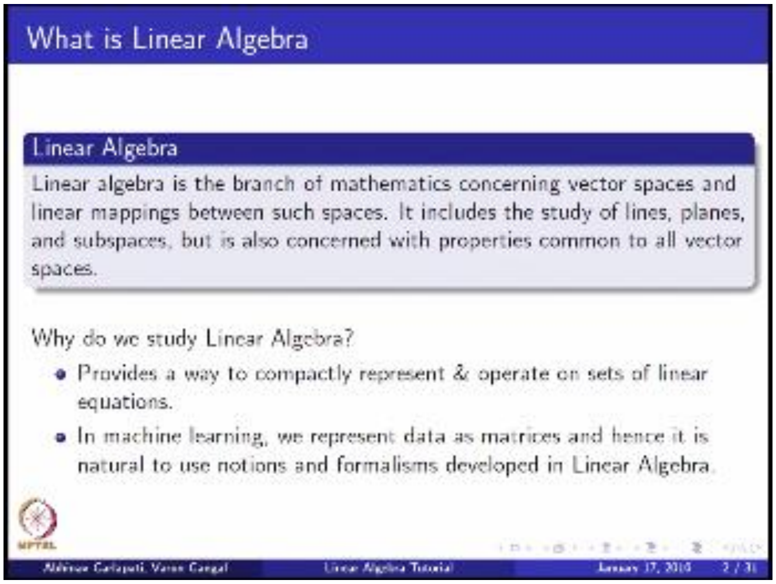
Linear Algebra-1

(Refer Slide Time: 00:15)



Hai, everyone welcome to the second tutorial of this, of the introduction to machine learning course, so in this tutorial we shall be taking a tool of the aspens of linear algebra which you would need for the course, we will cover a variety of concepts, subspaces, bases and decompositions Eigen values, Eigen vectors over the course of the tutorial.

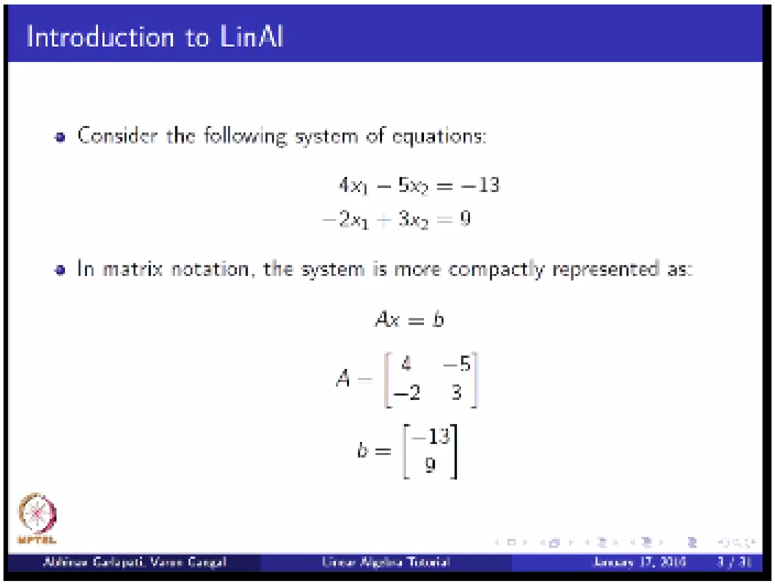
(Refer Slide Time: 00:47)



So the first question one would ask is why we need linear algebra at all and what is linear algebra, so you may have come across this in school or your +1 or +2 there will be but just to recap I mean linear algebra is the branch of mathematics which deals the vectors and vector spaces and the linear mappings between these spaces. So why do we study linear algebra here which is in the context of machine learning is first to be it is this is the way to if you let represented operate sets of linear equations.

But by do these linear equations come up in machine learning in the first quiz. So the reason was that is in machine learning we presented data is an entropies matrix so in this number of data point entropies the number of features. So it is natural that we have to use notions that formalisms developed in linear algebra. What is the data even I mean the parameter where user represented as vector so as a result linear algebra has an important role to today in machine learning.

(Refer Slide Time: 02:05)

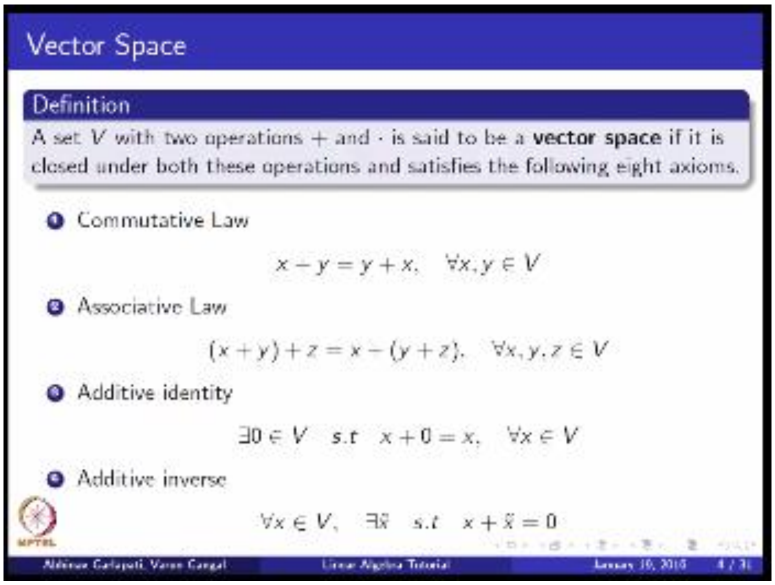


So we can see here a system of linear equations two equations and two variables 4x1-5x2=-13 and -2x1+3x2=9 so we can write away see here that advantages of a matrix notation so if you see below you can represent the same system of two equations directly as one equation the form of Ax=b, when A is the set of coefficient and x is the 2x1 matrix, 2 rows and 1 column or you may also call it a two dimensional vector x1,x2 so you can see that when you multiply the matrix A with x1, x2 the 2x1 matrix you will get back the same elegies or the set of two elegies and b the matrix b represents the RHS.

So we can it is very easy to verify that if you represent does not matrices you can get back the original representation and from that representation you can get this, to solving this in general way them matrices how you would solve this is you would first solve the one variable and then substitute to get the other. Now using matrices you can even solve this directly so just multiply both the sides by inverse, so you would get x and a inverse v of course when you and see you want to care about is that all matrices do not have an inverse but in most cases they do.

So in that case you can directly get the solution of x in the form of x=a inverse b whereas all being if this so one equation. So as we said only inverse and many other data gives in this favor to manipulates several equations at once in multiple variables.

(Refer Slide Time: 04:09)

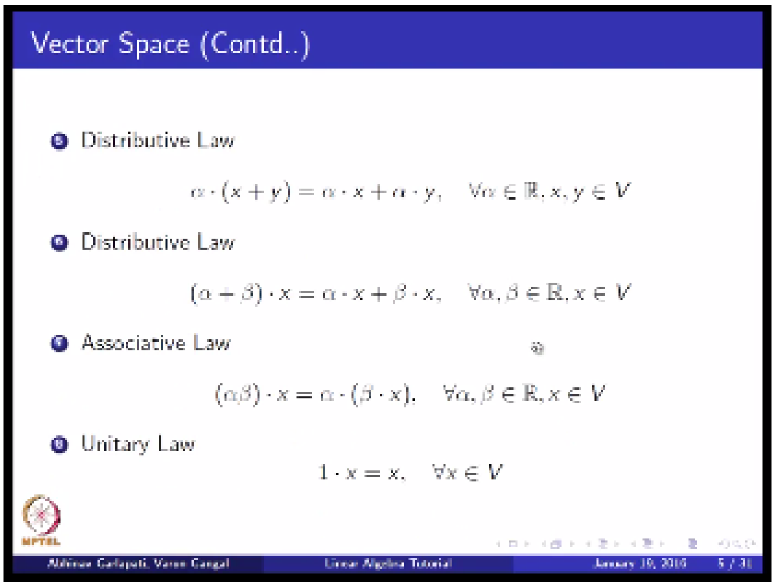


A fundamental definition in linear algebra is that pass of vector space, a set of vectors V is said to be a vector space if it is closed under the operations of vector addition and scalar multiplication and in addition satisfies the axioms where distributed by knowing to being that we have two elements from this set x and y where x+y will also lies in the set V. in addition to this if we take a scalar α or real number that is and multiply a vector from this set by let it then αy also becomes 2V. In both these properties satisfied then the set of vector is said to be closed with respect to vector addition and scalar multiplication.

Now let us have a look at the axioms the first one is the commutative law, the commutative law states that if you pick any two elements from the set V x and y then x+y=y+x. The Associative law says that if you pick any three element from this set x,y,z then x+y added together +z is equate to x+y+z added together. The additive identity losses that an exist an additive identity or 0 so as to say in this set.

Since that if you pick any element from the set and add the 0 to it you get back those same element. The additive inverse law says that there exists for every element there exists another for every element x there exists a corresponding x̄ such that x+x̄=0.

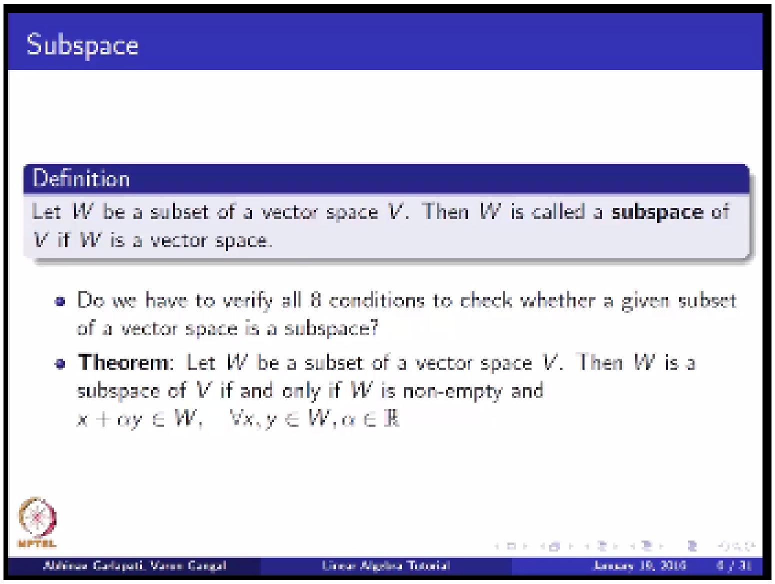
(Refer Slide Time: 06:15)



The five law is the distributive law, this law says that if you have a inverse scalar α with which you multiply the sum of two vectors x+y then that should be equal to αx+αy. The second distributive law says that if you have the sum of two scalars multiplying the vector x from the set α+βx x then that should be equal to αtimes x+β times x. The associative law says that if you first multiply two scalars and then multiply the vector will them that should be equal to multiplying first with second scalar β and then with α.

The unitary law says that on multiplication minus scalar real number 1 you get back the same vector, so this is important because we would not want multiplication because any unexpected scaling, if you multiply by the scalar k has a vector should be exactly k times it should not be say √k.

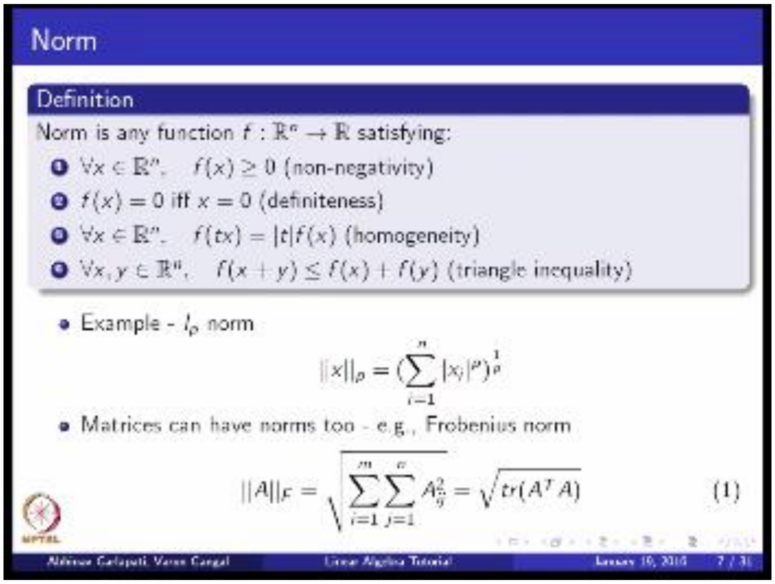
(Refer Slide Time: 07:34)



A second related definition is that of a subspace, sunset W of a vector space V is said to be a sunspace in W is a vector space, now this means that W should be closed under vector addition and scalar multiplication, it should also satisfy the eight axioms be stated earlier. Now the question that I raises means to being to verify all these eight condition given that we know that W is already a sunset of a vector space.

It is enough the check jut for the following two conditions, firstly the W is non empty that in other words it has at least to single element and secondly that if I pick any two elements x and y from this set and any real number α then x+αy should belong to W.

(Refer Slide Time: 08:39)



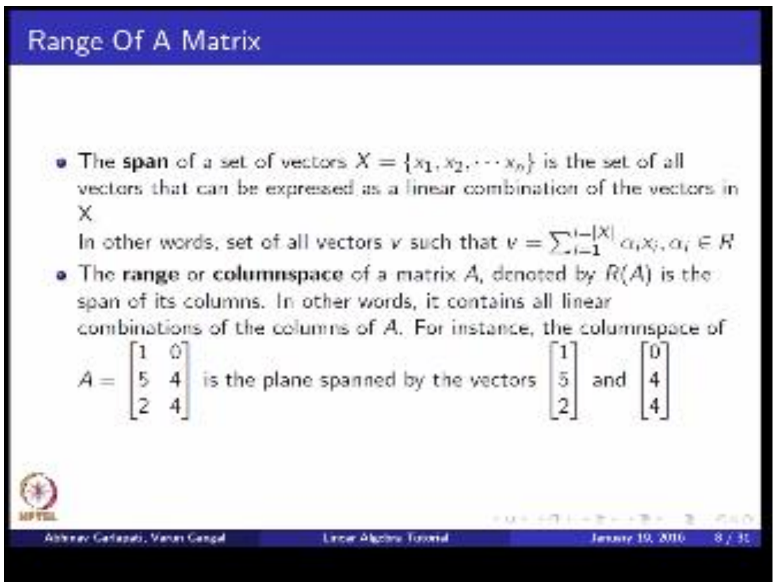
Now let us have look at a definition of a norm, so intuitively a norm is a measure of the length of a vector or it is magnitude, it is a function from this vector space which mostly happens to be R^n where n is the dimension of a vector to the space of real numbers R, so for a function to be a norm it should satisfy the whole conditions which we have given here. Firstly it should be it should always be norm negative.

Secondly, it should be 0 if and only if the vector is 0, thirdly for every vector if you multiplied it by a scalar it is norm should get multiplied by the module of the scalar by modeling here we mean the absolute value and fourth being that if we take any pair on vectors in our vector space which is R^n the norm of the sum of these two vectors should be less than sum of their norms this is also known as triangle inequality.

So this is in where related to the fact that the third side of a triangle should always be less than sum of the other two sides. Now an example of an norm is the l_p norm there you sum up the absolute value is along h dimension that is xp and then take the $1/p^{th}$ root of this, so when p=2µ get their two norm which is the magnitude of a vector whereas we have learned in our earlier studies $\sqrt{x^2+y^2}$ if you are looking at just at the space r^2.

There are other norms Frobenius there are norm define for matrices it as well, even behind to define norms four vectors so the Frobenius norm is a matrix norm so what did does is essentially sums up this square of all the elements and then takes their root of that, so this also happens to be equal to that race of A transpose A, now the race of a matrix is simply the sum of it is diagonal elements.

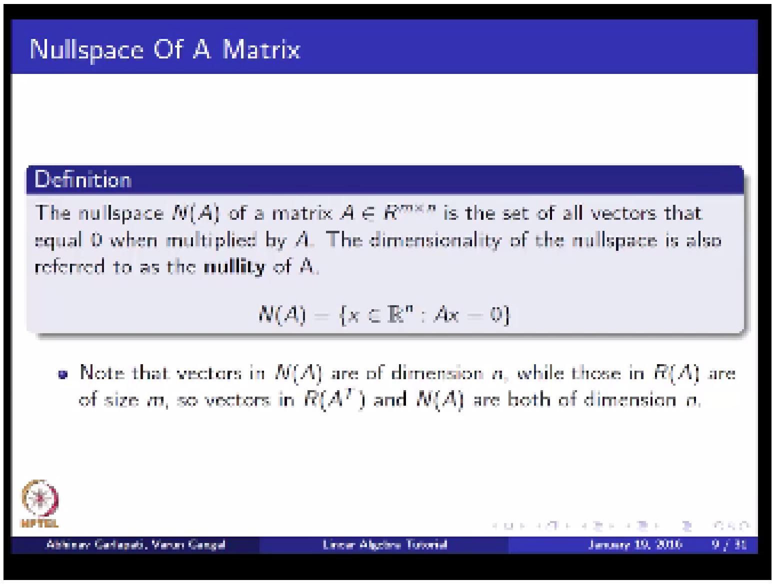
(Refer Slide Time: 11:12)



The span of a set of vectors is the set of all vectors which can be composed using these vectors by using the operations of vector addition and scalar multiplication so the name span consider the file type is set of vectors spans potentially larger set of vectors which is then call it as span so to define this more formally it is the set of all vectors v such that inverse someone i=1 to i equals mode X αiXi where αi is a real number now a related definition is that of range or column space so we think of a matrix each of it is columns is a vector so the set of all columns of a matrix is like a set of vectors.

Now the span of this set of vectors is called a range or column space of that matrix so if you consider the matrix given here the columns of this matrix A are ½ and 044 so what would be this columns space of this matrix it would be the span of the vectors 152 and 044 which is essentially the plane which is spanned by these two vectors.

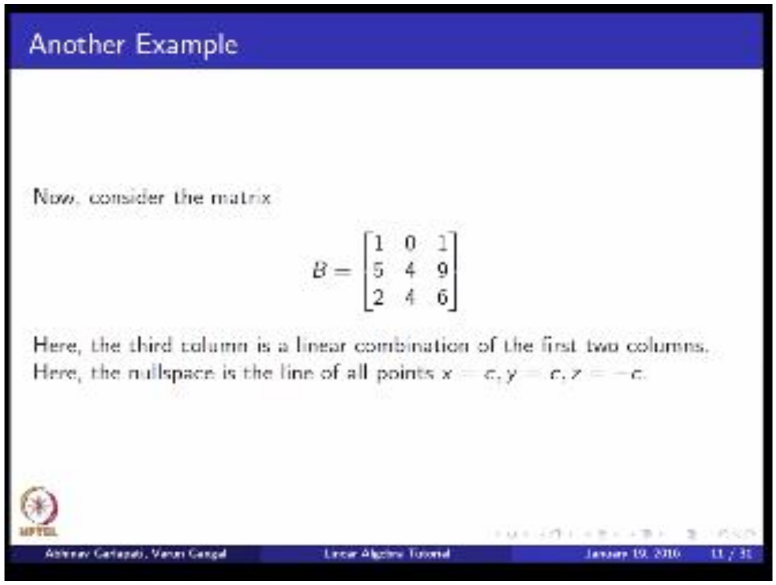
(Refer Slide Time: 12:56)



If we have a matrix A octavanges N x N then the null space is the sect if N x 1 vectors which given M x 1 0 vector on V multiplied by A in other words it is send of vectors X such that AX = 0 null is the ran for dimensionality of the null space we will replace the definition of nullity rate that once you defined rank more clearly another interesting fact above null spaces is that the null space of A N of A is dimension N by the range of area the columns space has be defined it earlier is a dimension M but M x 1 this means that vectors in R of A^T so note that A^T is N x M.

So vectors in the range of A^T will be of dimension N similar to the vectors in the null space of A so this means that the vectors in the range of A transpose and null space of A would both be of the dimension X cross 1.

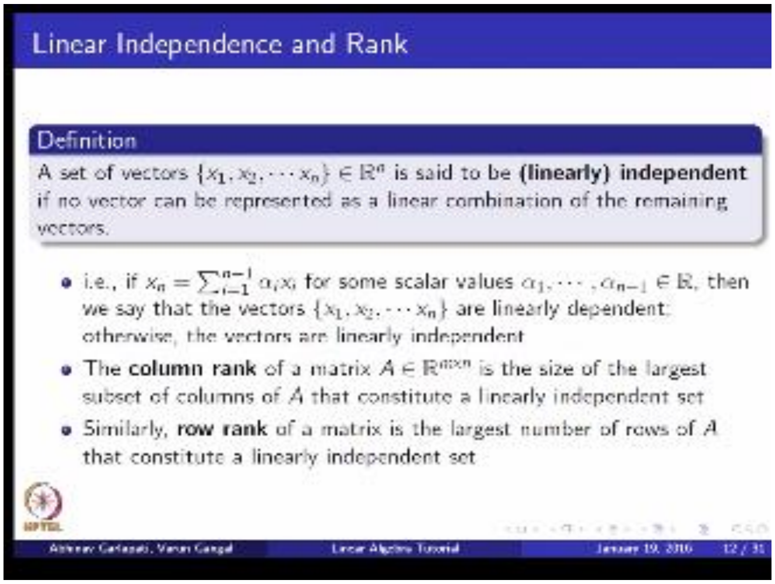
(Refer Slide Time: 14:15)



us consider an example to get a street the concept of a null space consider the matrix A given in the A is a 3 cross two matrix since the null space of A then may be vectors of dimension 2 what 2 cross 1 now we see that the and solving we get U = 0 V = 0 this means that a null space only contains the 0 vector the two dimensional 0 vectors 0 , 0 let us consider an another example to get illustrate null spaces better take the matrix being which is a 3 cross 3 matrix the null space would consist of 3 cross 1 vectors we leave the finding of the null space to the audience as an exercise.

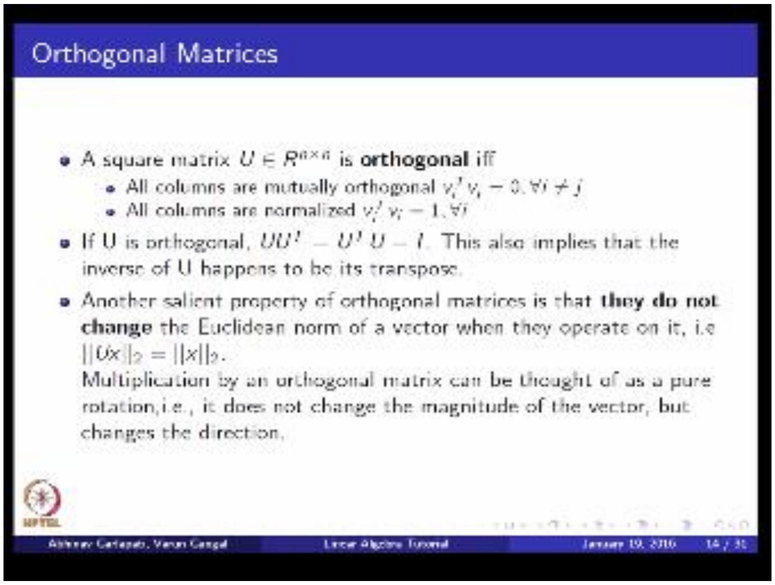
However then on solving we get the null space to be the set of all vectors of the form X = C by equal C and Z = - C then C is any real number and xy c refereed to the first second and third dimension respectively.

(Refer Slide Time: 15:34)



If for mean defining linear independents the recollect how we defined a linear combination a set of vectors is linearly independent is no vector in the set can be produced using a linear combination of the other vectors in this side now let us have a look at related concept of rank so the column rank of a N cross N matrix A is the size of the largest linearly independents of set of columns not at our columns you are a M cross one vectors the row rank is define in a similar ray for rows.

(Refer Slide Time: 16:22)



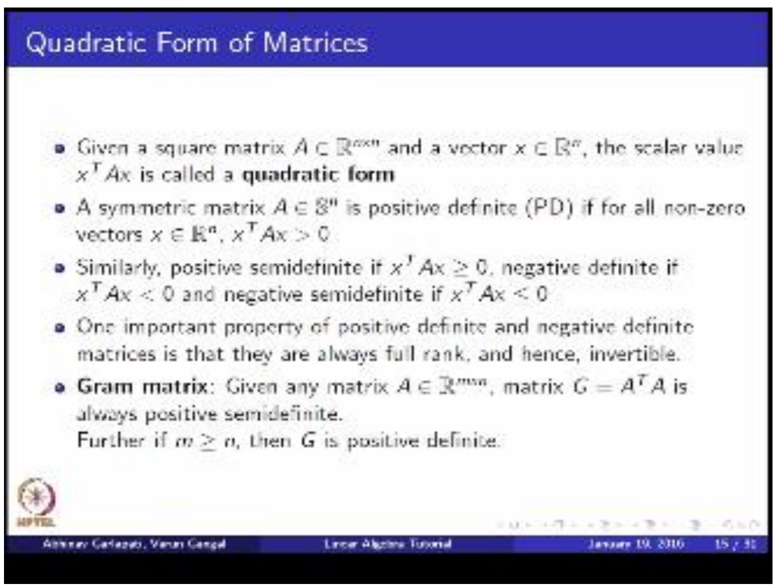
Let us work through some interesting properties of ranks for any M cross N matrix of real numbers the column rank is equal to the row rank we referred to this quantity as the rank of the matrix earlier we had looked at a quantity equal nullity, nullity is the rank of the null space of A some other interesting properties of ranks are also listed here for instance the rank of a matrix is utmost the minimum of it is two dimensions the row dimension in the column dimension secondly the rank of a matrix is the same of the rank of it is transpose thirdly if you multiply two matrixes A and P the rank of the resulted matrix is utmost the minimum.

Of the ranks of A and B if you add up to matrixes the rank of the result in matrix is utmost the some of the ranks of A and B square matrix of dimension N cross N is defined to be what problem in different only is the following two conditions with firstly all parts of distinct columns should be or probable by columns being are problem in that it dot prudent of any pair of distinct column vectors is 0 in other words VI transpose VJ = 0 for all I not equal to J is second condition is that a dot prudent of any column with itself or VI transpose VI equal 1.

In other words all the column vectors should be normalized and interesting implication of a matrix being also of nil is that U U transpose and U transpose U both end up equal to the N cross n identity matrix i this also means that U^T equals U inverse or the transpose of such here also matrix is also means equals and additional investing property is seen when we multiply a N cross 1 method x by a N cross n orthogonal matrix, the Euclidean are n to norm of such a vector X remains the same or multiplication by U.

It mutually we can understand this as orthogonal matrixes u performing good pure rotation or multiplying the vector X in other words they only change the direction of a vector by do not change it is magnitude.

(Refer Slide Time: 19:27)



We often uncounted the quadratic form which is the vector equivalent of a quadratic function the quadratic form with respect to the matrix A of a vector X where the matrix A is N cross N and the vector X is N cross 1 is given by the real number X^T AX based on the quadratic found the matrixes we can classify them as a positive definite negative definite positive semi definite and negative semi definite a matrix a is send to be positive definite in it is quadratic form is greater than 0 for any vector X similarly we can define it to be negative definite a matrix is positive semi definite.

If the quadratic form is greater than equal to 0 for any vector X note that equality 0 also may bold here but important property of positive and negative definite matrixes is that they are always full rack and implication of this is that A inverse always exist for a matrix A which is of dimension M cross N when that define and special matrix called a gram matrix the gram matrix is given by A^T A the save your property of the gram matrix is that it is always positive semi definite moreover if the number of rows exceeds the number of columns in another words if M is greater than equal to N this means this immunize that G is positive definite.

IIT Madras Production

Funded by

Department of Higher Education

Ministry of Human Resource Development

Government of India

www.nptel.ac.in