Indian Institute of Technology Madras
Presents

NPTEL
NATIONAL PROGRAMME ON TECHNOLOGY ENHANCED LEARNING

Pattern Recognition

Module 01

Lecture 05

Relevant basics of linear algebra, Vector spaces

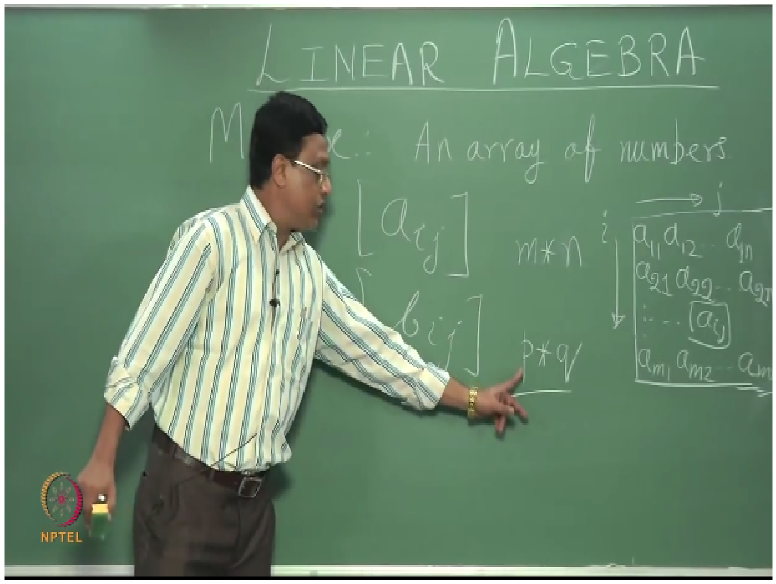
Prof. Sukhendu Das
Department of CS&E, IIT Madras

So in the last class we had a discussion on the difference between classification and clustering in feature space now to implement most of these algorithms or computational methods we will need a many mathematical tools and methods which are based on concepts of probability statics linear algebra and vector spaces of course there are other methods to do clustering and classification which are based on new networks and graph based methods or syntactic analysis as well through this course we will be mainly concentrating on statistical methods.

And of course we will touch upon if you neural based methods for classification as well and to understand these methods in a better way we expect the students to have some background on the mathematical principles of linear algebra vector space probability and statistics the concepts of probability and statistics will be covered later in the next class today we will look into basics of linear algebra and vector spaces.

For those in the field of mathematics and also computer science and electrical engineering this may be revision of the basics and if you are conversation with this topic then you can actually skip this lecture and go forward to the next lecture itself for those who are not used to these methods may be in the recent passed without to provide you one lecture on a very basics of linear algebra and vector spaces.

(Refer Slide Time: 02:40)



So we start off with the concept of what is a matrix. Matrix is a an array of numbers it as applications in may branches of mathematics computer science and electrical engineering using a matrix you can represent a graph you can solve a system of linear humongous equations you can use it to represent any set of numbers which have some order between them okay and of course we are now going to use matrix in our field of pattern recognition also.

What is the symbol which is used to represent a matrix you can represent it a matrix usually by a capital letter A or B and also it is rectangular array of numbers a matrix is a rectangular array of numbers a typical notion which you may have is something like this aij or bij where very simply these are considered to be the elements of the matrix A or B okay the question comes is what sis the size what dimension of this matrix okay.

So you have an array of numbers so we are talking of a 2 dimensional array of numbers and let us say the size of A is m * m these are 2 integers okay the 2 integers or it can be something like p * q as well essentially it consist of a set of rows and columns so the number of rows multiplied by the number of columns of a matrix give the size odd number of elements within a matrix there are 2 subscripts which are used here which we will be trying to use consistently inj correct.

So one of them will represent an index for the column the other for the row typically if I draw the elements of a matrix say aij are the elements matrix so I can write them as a11, a12 and a13 and so on up to the last element a1n this is the 1st row of the matrix A you can write the elements of the second row as similarly as a21, a22, a23 and so on up to a2n very simply and so on for 3^{rd} 4^{th} one so on until you reach the m^{th} row where you write them as am1,am2 and so on up to amn.

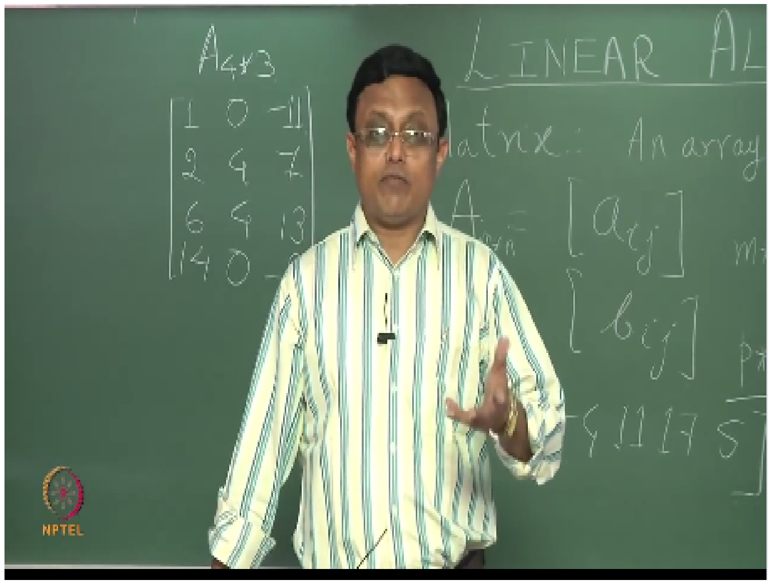
Okay the same thing will be applicable for the elements bij for the matrix B where we just simply replaced the charter a/ b and you will have b11, b12 up to b1n, b21, b22 up to b2n and similarly bm1 sorry you have to be careful because the size of the array B is p*q so it start from b11, b12 and so on up to b1q, b21, b22 up to b2q and then finally a bp1, bp2 up to bpq okay notion wise sometime the size of the array is return as a subscript here where you will find a cross or a star indicating.

That this is the total number of elements in the array or there are so many rows multiplied by so many columns so the index 1 which you see here is actually indicating the number of rows is like the y axis in an x, y graph or plot and the j indicates the column under consideration so obituary element aij here will be talking about the i^{th} row and the j^{th} column I repeat 1^{st} 2^{nd} and so up to m rows in the matrix A and 1^{st}, 2^{nd} up to m columns for the matrix A as well.

So the element aij will correspond to an element at the i^{th} row and the ^{jth} column similarly you may have also a bij as the i^{th} and the j^{th} column element of the matrix b of course in this case you must ensure that I, n. j as not exceed the value p and q respectively of course there are some notions which might actually start from 00 as here as the first element and stop at m – 1, n – 1here okay or p – 1, q – 1 here it is only a matter of a convention okay.

So this is of size p * q so the lower case let us are representing here the position of the element i^{th} row j^{th} column here okay the small let us indicate the element of a matrix capital letters is the convention which we will used represent the entire matrix let us take a few examples of matrices a very simple examples let us say a matrix has.

(Refer Slide Time: 09:35)



A matrix A of size 4 x 3 what does it indicate basically there are four rows and three columns and let us write some elements okay that it is a typical example of a matrix with consist of 4 rows and just three columns a matrix can be a special time which can just have let us say I write a matrix B as something like this very simply how many columns just on column how many rows three rows this is also possible because what we are talking of is that the value of M is 3 here the three rows and the value of n is 1 here just 1 column.

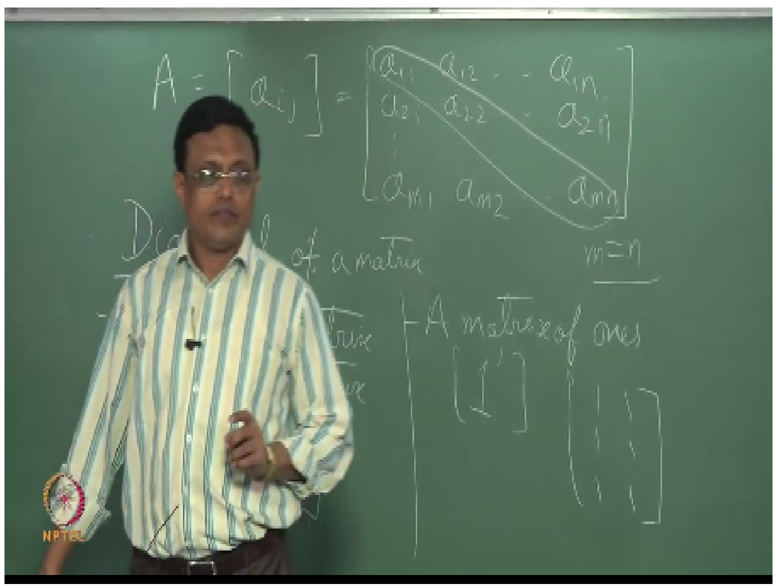
In this special case of a matrix this is also can be considered to be vector will introduced the word vector little bit later on and for the time way we will see that this can form the elements of a particular vector as well of course you can have an array which is having many columns but juts on row a typical example could be another matrix A which is something like this 4 11 17 and 5 basically it contains 4 columns and 1 row, so this is actually sometimes called a row vector okay because the difference.

This as one column this is 1 row okay sometimes it is called column vector and row vector especially case of matrices if you just introduced what is a matrix and what it is the element could be we are talking about real number in this case but member there are some mathematical analysis where you can have evaluates to b complex number is well they can exist you can have complex matrices, but we will keep them out of the scope of our discussion today we may not need that much.

In the field of pattern recognition that is the main reason we restrict our scope of elements aij and bij to be real numbers well not about applications of matrices okay our focus will be w us them in the field of pattern recognition for the purpose of classification clustering and called suitable analysis which are required for classification and clustering analysis for that we need to manipulate the elements of a matrix there are various operations which are possible with matrices we will first find.

A few properties which are associated with matrices we all find some special type of matrices which exist and then go into various manipulations and operations which can exist with matrices before we start looking into different properties and operations on matrices will get use to a few simple terminologies associated with matrices okay.

(Refer Slide Time: 13:41)

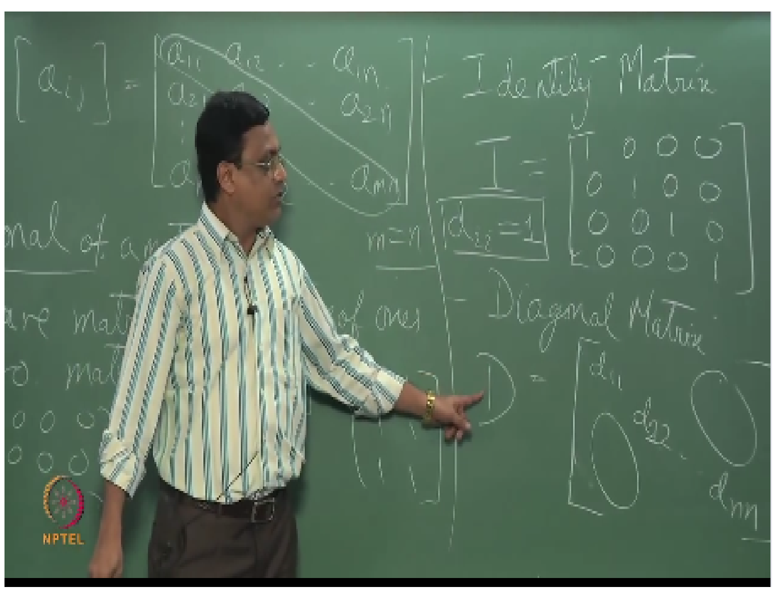


So we had a matrix a_{ij} and a_{11} a_{12} a1n second row third row and writing the last row here I will write the second row also diagonal of a matrix of an also called the main diagonal of the matrix or the elements which in the diagonal, so if you ask me what of the main diagonal elements here that is a typical example a11 a22 a33 and so on up to of course m where we will assume here m = n we will assume here that m is equal to n that means the matrices is a square matrix that is another term.

A matrix can be either square or non square typically that it is a simple thing if m is not equal to n the matrix is rectangular or non square if it is if the case is m = n matrix is a square matrix we will also define a very simple matrix called the 0 matrix this is diagonal of a matrix manure diagonal is not a matrix the set of elements which are following on the diagonal going from top left to bottom right not the other diagonal line is the diagonal of a matrix these are special matrices, so we can have a square matrix or a 0 matrix elements are all 0 typical example this can be a set of elements.

On a single row or on a single column or even few rows a set of a rows in columns forming well all the elements are 0 okay you can have the matrix of one's similarly which is sometimes well casually written as the matrix of ones the typical example could be elements are all 1 okay very simple elements.

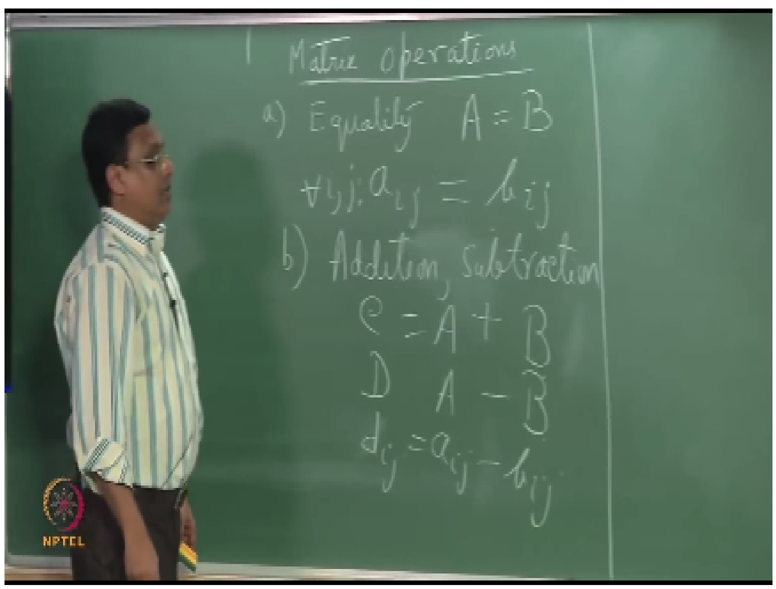
(Refer Slide Time: 16:55)



A next matrix is an identity matrix which is given by the symbol I this is a special case of what is called as a diagonal matrix where the elements are all at the diagonal the elements are equal to 1 and the non diagonal elements are equal to 0 okay whereas in the case of a diagonal matrix which is often referred to by the symbol D it is often written as d11 d12 up to sorry d22 up to d_{nn} if the matrix is of course square and all the other elements are 0 okay and instead of writing all th elements as 0 as here.

Is often represented by a big 0 means all the elements which are of diagonal that means they are not sitting on the diagonal are 0 forms diagonal matrix this is the very important matrix is w used and a special case of diagonal matrix is in annuity matrix where the diagonal elements dii = 1 this is case which happens when diagonal matrix becomes and identify matrix let us look at a few operations.

(Refer Slide Time: 18:48)

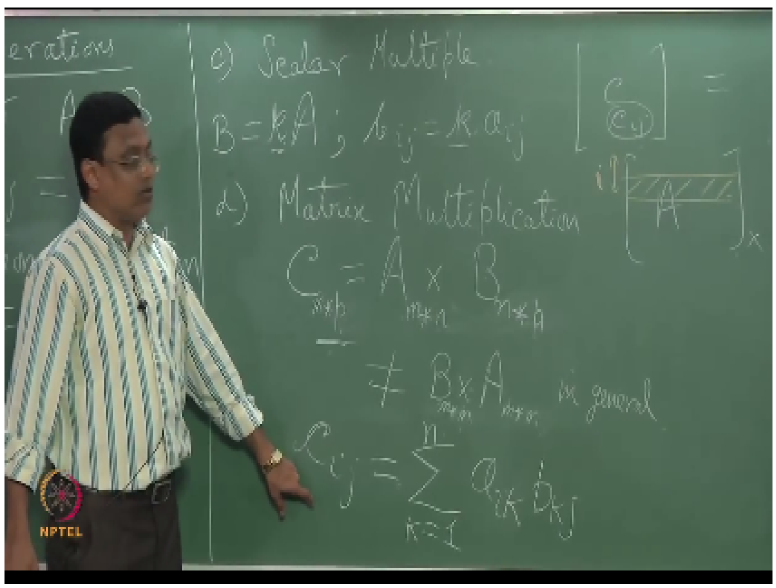


Operations are not matrix very simple preliminary operations on matrices the first of them will with the equality of a matrix, when do you consider two matrices to be equal so if we say two matrices A and B are equal only if first of all two matrices should be of equal size that means if this is mxn this is also mxn and the elements aij and bij for the corresponding matrices A as element aij, B has bij should be equal.

And this should be true for all i and j, for all values of i and j this should be true. Similar to operations on any elements in mathematics you can also have addition and subtraction which involves basically that if there are two matrices A and B you can either do an addition or you can also do a subtraction operation maybe I will write it below giving rise to a result in matrix C or D depending up on whether you add the elements and what does it involve typically the elements cij of the matrix C will be aij+bij for all elements i and j again the two matrices have to be all the same size are order as A and B.

And if you replace addition by subtraction you get another matrix, D with corresponding elements dij.

(Refer Slide Time: 21:27)



Let us look at a few more examples, scalar multiple of a matrix so if you have a matrix A and you want to multiply that with a constant k which is called as scalar value you get another matrix say B, what does it involve the elements of the matrix B result in matrix B all of this elements are very simply k.aih, what is this k any number actually it can be real or complex but of course in this case we are looking at only real numbers.

The fourth most important operation matrix multiplication, important means among those which we have discussed so far given two matrices A and B I want to multiply A and B to get another matrix C correct, and the orders are very important here. You cannot multiply all the time the order has to be such that or the size of the matrices have to be such that, that if the size of this multiplication is talking about the product or a matrix multiplication between A and B.

And that is why is the reason why I have used different symbol to indicate the number of elements in A, then B should have something like this nxp. And if this is so this has to be there that means the number of columns in a has to be equal to the number of rows in B correct, I repeat it is possible only to multiply two matrices when the number the number of columns in A is equal to the number of rows in B.

It is possible to multiply two matrices and get a result in matrix C for with the size will be very simply mxp. mxn is the size of A and nxp the size of B result in size will be mxp, the other thing which you must remember of course what is the size, what are the elements of C. Before I get into the element Cij of the matrix C as a function of the element Aij and Bij here I would also like to state here that A are these two same.

That means if I reverse the order m, I will not be able to do that in general if the sizes are given like that like the way they are given. I can reverse these two provided p=m because what will happen here that means the number of columns in B must be the same as the number of rows in A okay. So let us say for the time being m=p so that means I am talking of nxp, the pxm here let us say okay, or let us say I talk of a situation where the sizes of both of them are same.

So let us say I am talking about and the same thing holds good here, now both of them are square matrices of same size it is not necessary for both to be same to get a matrix multiplication then we talked about this that this value must be the same as this value, what this indicates number of columns in A must be the same as the number of rows in B correct, that must be the same.

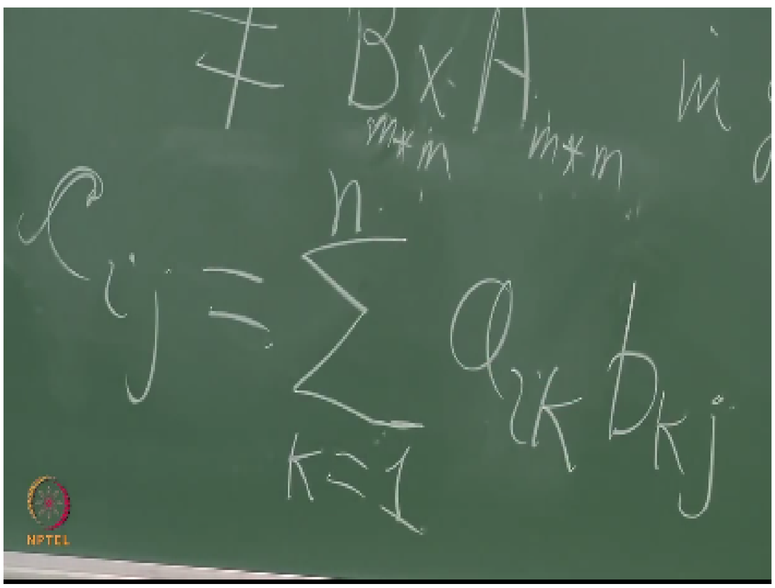
Then I should be able to multiply I have changed all this numbers to represent that AB are square matrices both are of same size because I want to reverse and in general this value which you get now you will get a mxm here, but this is not equal to this in general this is not true. Under special cases this is possible but we will not discuss that now and this operation of matrix multiplication is extremely importance and use in many braches of science and engineering not only this field of pattern recognition.

We need to multiple lot of matrices but only thing you know it is a very common operation that is what I mean and the property that the number of columns equal to the number of rows is very,

very important what are the elements of the matrix here, as I said before that let us go back and restore this that the way we talked about. M x n x value n x n get this thing to come, so let us take this and this can be remain it is fine, so cij that means there are m x p x elements in c so one such element cij when cij can be represented as a function of some of the element of a and b which basically means I am talking of the I^{th} row of a and the j^{th} column of b and it can be represent as the summation like this.

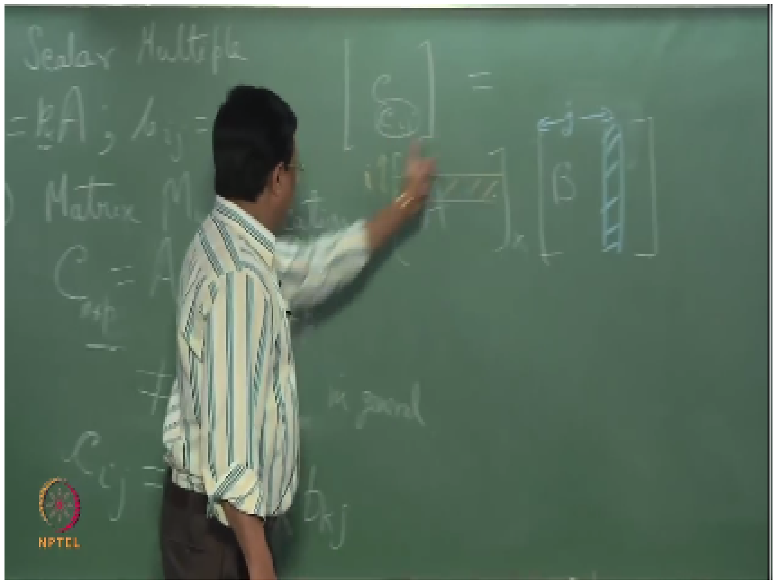
You are able to see what I have written I have written this as a function of the element of a and b and it is basically a some of products okay, look at how the k is varying when k is varying it is the second subscript it basically means that I am talking about the elements in a particular I^{th} row all the elements of the I^{th} row can be obtain by varying a and the j^{th} column here can be obtain.

(Refer Slide Time: 29:01)



So what you are talking basically is if I can write c as a matrix.

(Refer Slide Time: 29:08)

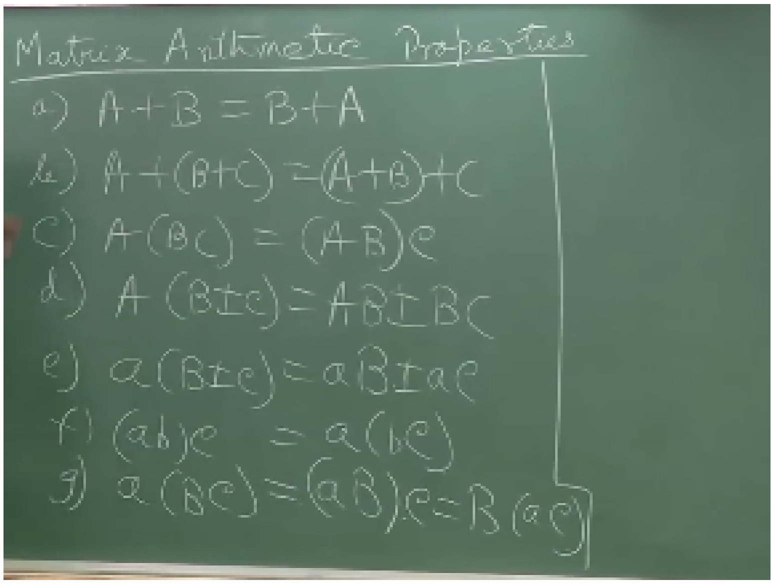


A x B then what I can tell you is that for an arbitrary element c I rake here arbitrary element c I take here I am talking about the I^{th} row let us say this is I, and j^{th} column of b, and I did say sometime earlier that the number of columns sorry the number of columns in a equal to the number of rows in b's so that means if you are doing an element wise multiplication correspondingly this is what this function is talking what, this function says that I take the first element of an i^{th} row of a first element of j^{th} column of b multiplied m get the first round in this summation.

Second term that the second and so on how many terms n number of terms because that is what it is the n^{th} column versus the n^{th} row of the and this is what we need at the matrix what we will get. Can you on me with discussion the next operation which we can do is call the transpose of a matrix, transpose of a matrix a again if I write as a ij m x n transpose of a matrix a if a is a matrix then the symbol typically use to represent the transpose of a matrix is this or some books might also use a notation like this a very short notation okay.

Basically a ji and x size will be n x m, basically the rows have become the columns and the columns have become the rows of a matrix. Trace of a matrix the trace of a matrix a is define as is equal to Σ aij assuming that there are it is a square matrix assuming it is a square matrix of size n x m this is a trace of a matrix, after learning a few operations on matrix will trying out to see we will see a few properties of matrix operations.

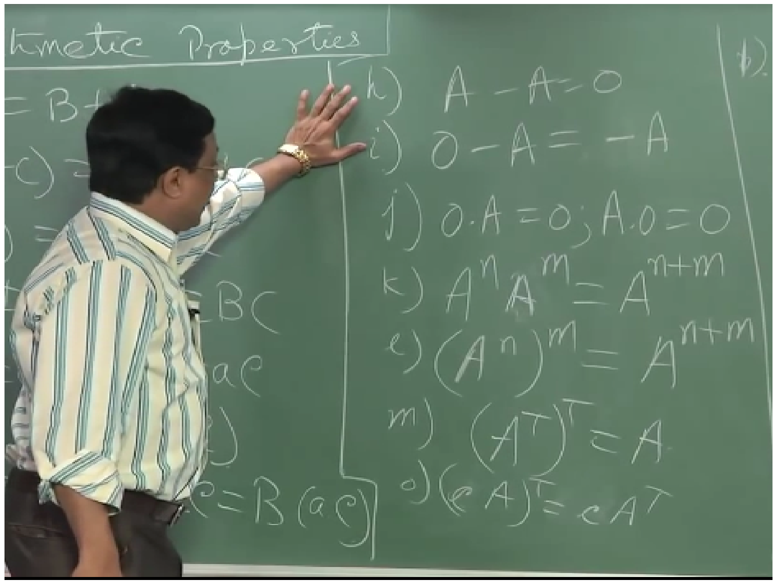
(Refer Slide Time: 32:55)



So there are lot to we will illustrate here, so I am just writing a few and will discuss them, the notations themselves are self sufficient here most of them are concerning addition subtraction of matrices or matrix multiplication in certain cases the two things which you must remember in mind you are keep in mind is that the small letters here represents a scalar number typically a real number but it can be a fraction also and the capital letter here represents a matrix so a and b are scalars here a is a scalar b and c are matrices.

So this is the same as this or it can be even written as this when I write a notation like this plus and minus that means you are either adding or subtracting does not matter the same results porch good here is well, so these are some of the loss of the operations which are possible arithmetic operations on matrices and these properties hold good continuing on that.

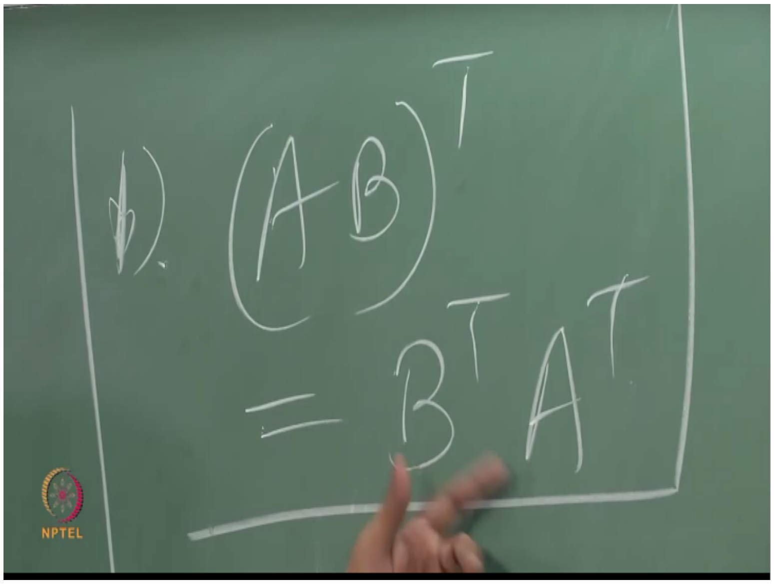
(Refer Slide Time: 36:21)



I will write just one more here this inverse this are again elementary operations so these re elementary operations on matrices again mostly to it subtraction, multiplication we are taking about epic matrix 0 so this is epic matrix of 0 matrices of 0 here this can be thought of 0 matrices or scalar matrix both is possible okay two matrices both are same here A and B are different matrices.

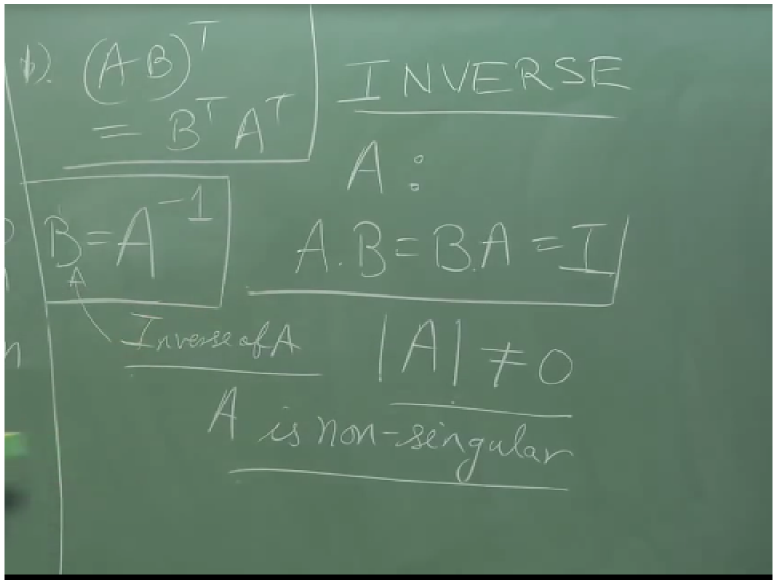
So both matrices are same then we should able to write here the same towards to the power and this is very simplify to prove we can take actually a small matrix A of any size which could even non square and transpose it twice basically it transpose operations involves interchanging of the rows and the columns so when you do it twice interchange of rows and columns throughout comes back to rows and the columns back to so will get the matrices A here.

(Refer Slide Time: 39:44)



This C is a scalar okay the last one is most important point which is a transpose over the product transpose over a product is equal to the product of the respective transpose in the reverse order I repeat again transpose of a product of two matrices will results is equal to the multiplication of two matrices A transposes in the reverse order whatever the order remember in journal I said A multiply B will not be equal to B multiply by A.

(Refer Slide Time: 40:21)



But then we able to transpose here we have to interchange the order of the multiplication then one of the most important operations of matrices is the inverse of a matrix so given a matrix A we were talking in general of a square matrix although the operations possible inverse of the matrix which is also non square but you take it up later on so we will talking about the inverse of a square matrix A is given by symbol usually like this A^{-1} indicating that inverses of a matrix.

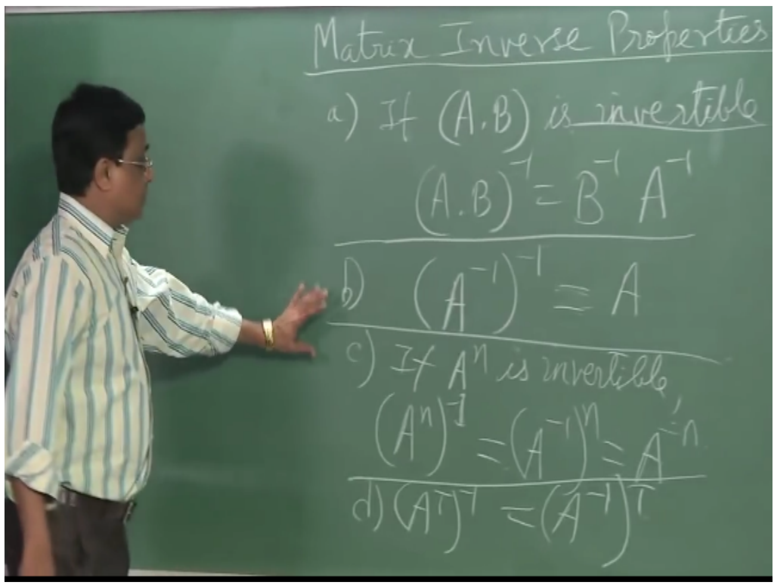
The elements of this matrix A can be represented as a elements of a matrix A itself but we think of this expression right now so if matrix B is the inverse of a matrix A that means you are trying to inverse some matrix A and you get a matrix B then this two matrices A and B they follow this property the product of the matrix and it is inverse in any order.

We can thought of as an inverse of matrix A or A also can be thought of them inverse of matrix B and they have multiplying any order it results in an identity matrix I identity matrix is a diagonal matrix which is the elements are all equal to 1 of that elements as simply equal to 0 this operations of trying to inverse a matrix there are many different algorithms talked about the field of mathematics and numerical analysis.

I will leave it to you for self study to find out methods this on numerical computation which can actually inverse the matrix I m not giving you that or I m saying is first of all this matrix must be square and then the matrix A must satisfy another property for its inverse matrix B to exist so will say here if this is a matrix A then this matrix B is the inverse of matrix A and it is possible to get this value or the elements of the matrix here instep the only if the following property holds good.

To determinant of the matrix A is not equal to 0 nd in such case is called A is non singular only when A is non singular you can obtain is inverse if A is singular then this is not good a determinant will be equal to 0 and it is possible only in the case of square matrix okay so that is there so we will now look at some properties which are to do with the inverse of a matrices.

(Refer Slide Time: 43:49)

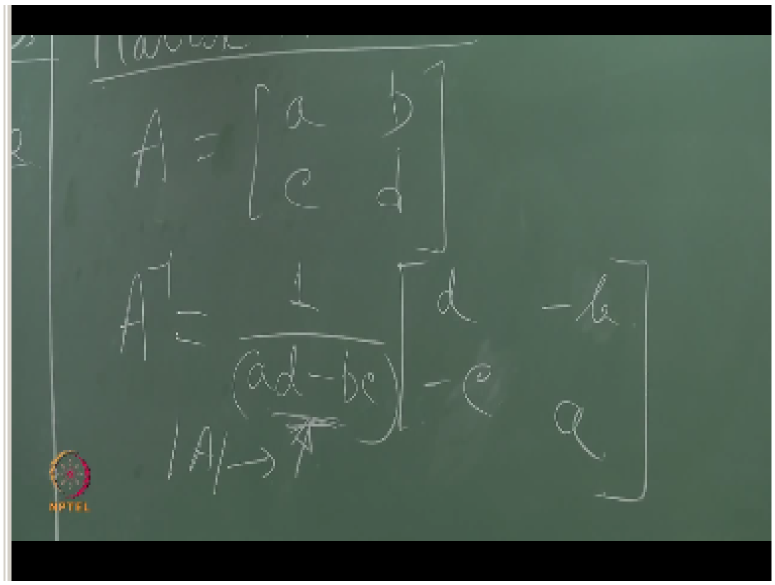


We will look at now into some properties of matrices inverse okay or properties will get to the inverse of a matrix if the product of two matrices A and B is invertible and you know the condition of invariability that means when you are able to invert that is the case when the matrices non singular if the product of the matrix A and B is non singular determinant is not equal to 0 then I can write this as inverse is, is the same thing which we had some time back with respect to transpose of the matrix transpose of the product okay.

And of course like I talked about the sequence of two transpose you can have invert the matrix A follows by inverse again you will get back the matrix

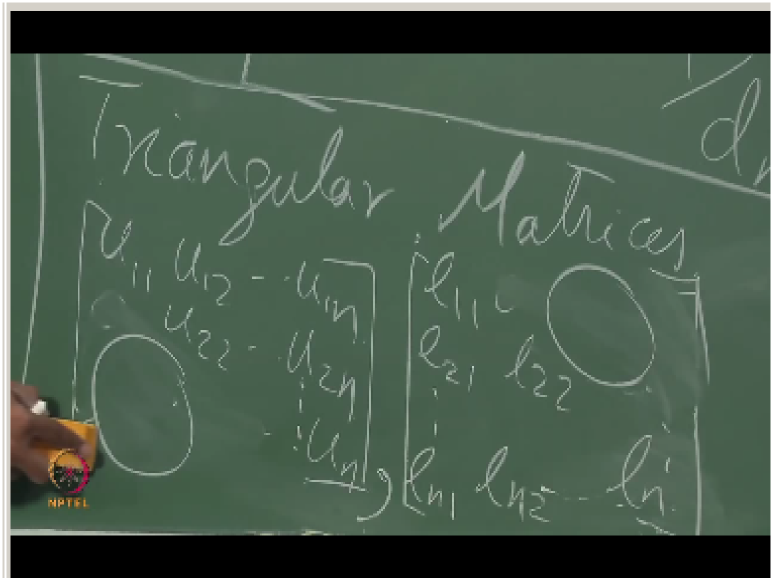
So for any integer n if a^n is invertible it basically means you are taking a product of a with itself n times then the corresponding inverse of this matrix a^n then also be written as this it is for –n or you take the inverse of a and then the power n raised to n is multiply n times the last of this list of properties is that we take a transpose of a matrix and take its inverse take an transpose of an matrix and take its inverse is equal to inverse also with transpose take the inverse of matrix a followed by its transpose .

(Refer Slide Time: 47:25)



The remaining algorithms to take the inverse of matrix but we will take simple example to understand the process of inversion let us say very simple 2.2 matrix a as elements ABCD then the inverse of a matrix a can be written as when this is the matrix then this is inverse this that what is this term which I put in the denominator at the beginning where it is to determine of it okay determined and if you look at the rest of the terms here they are basically called they are not defined another property with respect to matrix this elements which I have put here.

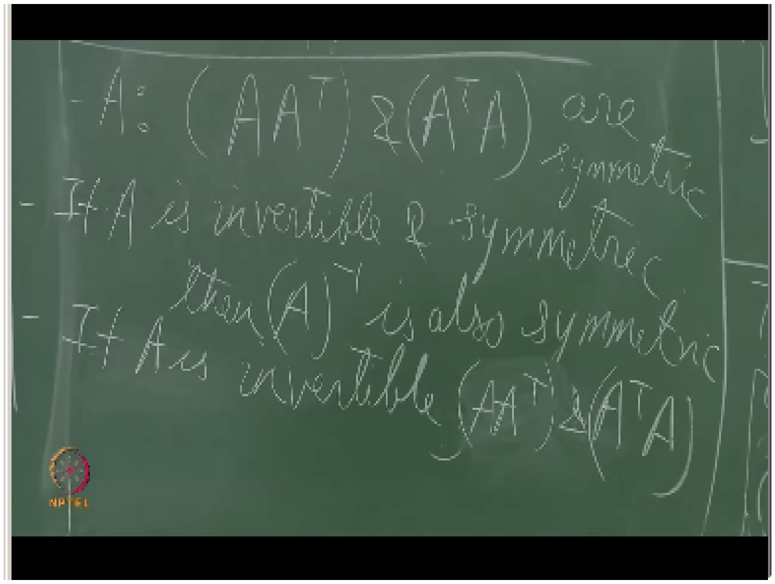
(Refer Slide Time: 48:54)



Triangular matrix if we take a diagonal matrix D I hope you remember the notation of this big zeroes which indicate that the oct diagonal terms are zero and the diagonal terms if you just want to and basically d11, d22 and up to dnn matrix cross M so for sake a simplicity the diagonal elements I just use to indicate that they are Di then the corresponding matrix D^{-1} is given by is what equal to this very simply taken.

The diagonal elements and put to reciprocal triangular matrices two special types of triangular matrices where you can have just the elements on the top upper right or you can have done on the upper right or lower left okay they are not the same l21 l22 up to l2n so can actually replace this part a big zero as we have done for the diagonal element or the same thing is applicable here that I have the diagonal terms and anything to the top right is a big 0

(Refer Slide Time: 52:04)



Symmetric matrices A is a symmetric matrix if aij=aji that means the element at the ith growth jth column is the same as the jth row ith column you want an example does not matter what is on the diagonal look at the symmetric part of the matrix this element should be same as this responding here so this is an example of symmetric matrix if you have an arbitrary matrix A a.

Some more properties for an arbitrary matrix A both AA^T and A^TA that means you can think of this two matrices they are symmetric if A is invertible and symmetric then A^{-1} is also symmetric or else if A is only right here A is only inverted if A is only invertible then A^{-1} and sorry A^TA IF A is invertible then the product take in a order of these two matrices are both inverted I think we will stop here.

Online Video Editing/Post Production

K.R.Mahendra Babu
Soju Francis
S.Pradeepa
S.Subash

Camera

Selvam
Robert Joseph
Karthikeyan
Ramkumar
Ramganesh
Sathiaraj

Studio Assistants

Krishnakumar
Linuselvan
Saranraj

Animations

Anushree Santhosh
Pradeep Valan.S.L

NPTEL Web&Faculty Assistance Team

Allen Jacob Dinesh
Bharathi Balaji
Deepa Venkatraman
Dianis Bertin
Gayathri
Gurumoorthi
Jason Prasad
Jayanthi
Kamala Ramakrishnan
Lakshmi Priya
Malarvizhi
Manikandasivam
Mohana Sundari
Muthu Kumaran
Naveen Kumar
Palani
Salomi
Senthil
Sridharan
Suriyakumari

Administrative Assistant

Janakiraman.K.S

Video Producers

K.R.Ravindranath
Kannan Krishnamurty

IIT Madras Production

Funded By

Department of Higher Education
Ministry of Human Resource Development
Government of India

www.nptel.ac.in