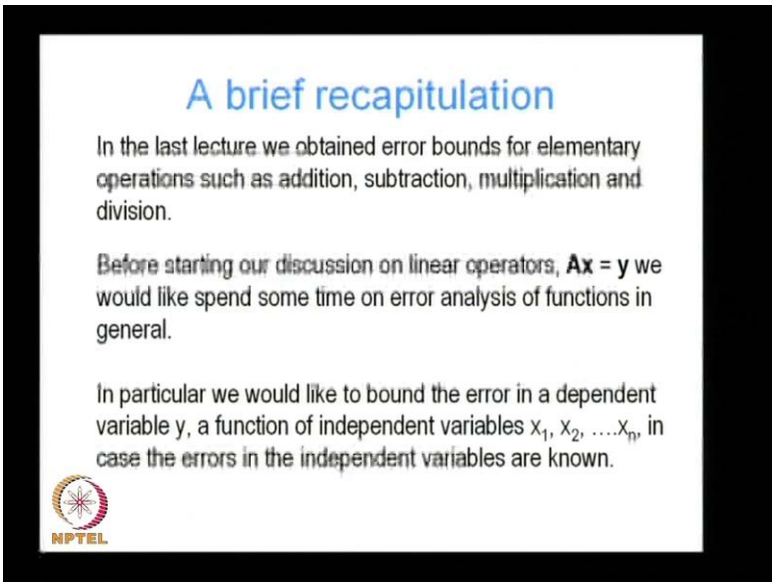
Numerical Methods in Civil Engineering
Prof. Arghya Deb
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

Lecture - 3
Introduction to Linear Systems

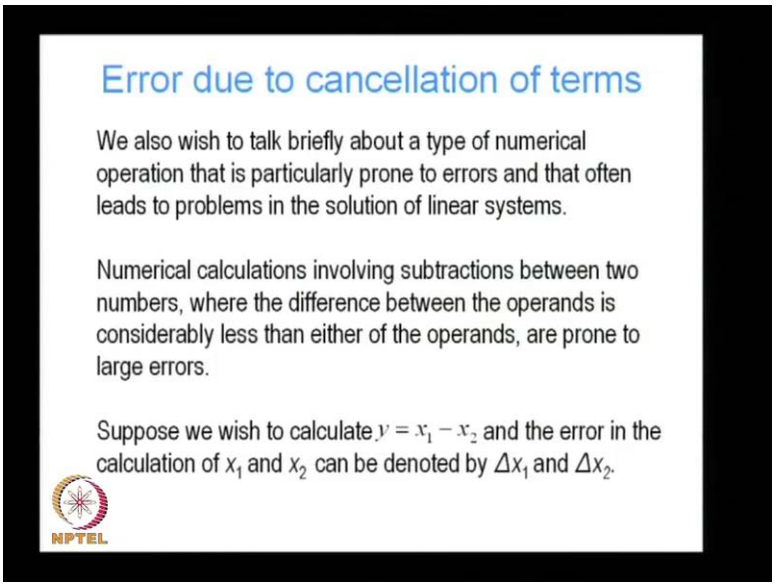
(Refer Slide Time: 00:20)



In civil engineering, we are going to talk about Linear Systems. In the last lecture we obtained error bounds for elementary operations, such as addition subtraction, multiplication and division. Before starting our discussion on linear operators, which are typically written as A times x is equal to y, we would like to spend some time on error analysis of functions in general. In particular, we would like to bound the error in a dependent variable y, a function of in a dependent variables x 1, x 2 up to x n.

In case the errors in the independent variables are known, for instance if our function is a linear function, if we know the errors in the depend in the independent variables A x, x then we want to find out what would be the error in y, after we operate on x operator with the linear operator A.

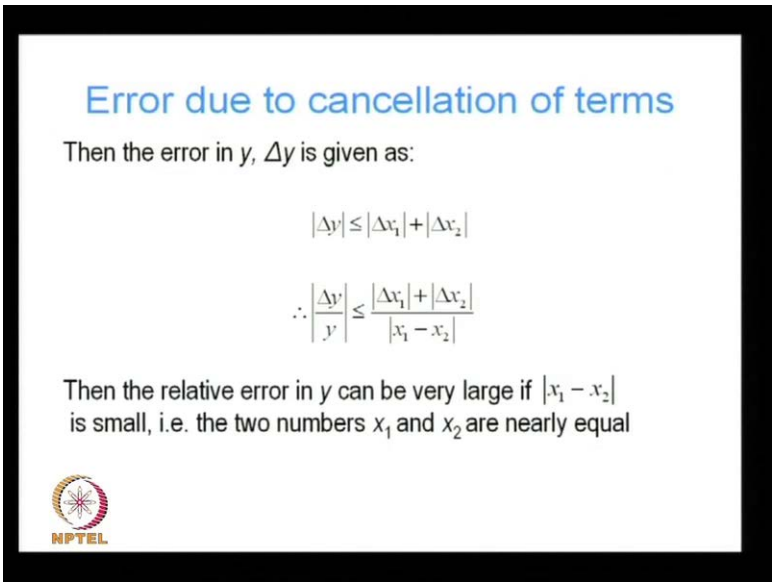
(Refer Slide Time: 01:32)



We would also like to wish, we would also like to talk briefly about a type of numerical operation, that is particularly prone to errors, and that often leads to problems in the solution of linear systems. Numerical calculations involving subtractions between two numbers, where the difference between the operands is considerably less than either of the operands, are prone to large errors.

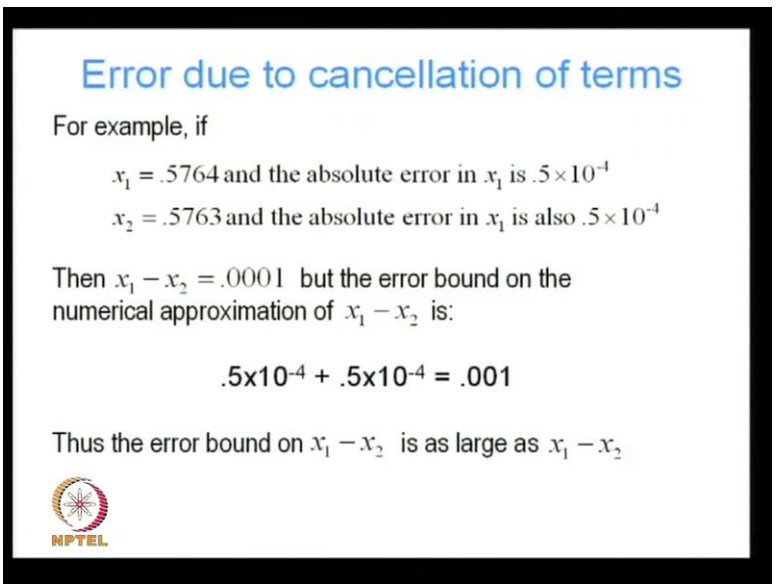
Suppose, we wish to calculate y is equal to x 1 minus x 2 and the error in the calculation of x 1 and x 2 can be denoted by delta x 1 and delta x 2, that is x 1 and x 2 are my independent variables; and I know the error in x 1 and x 2 which is given by delta x 1 and delta x 2.

(Refer Slide Time: 02:34)



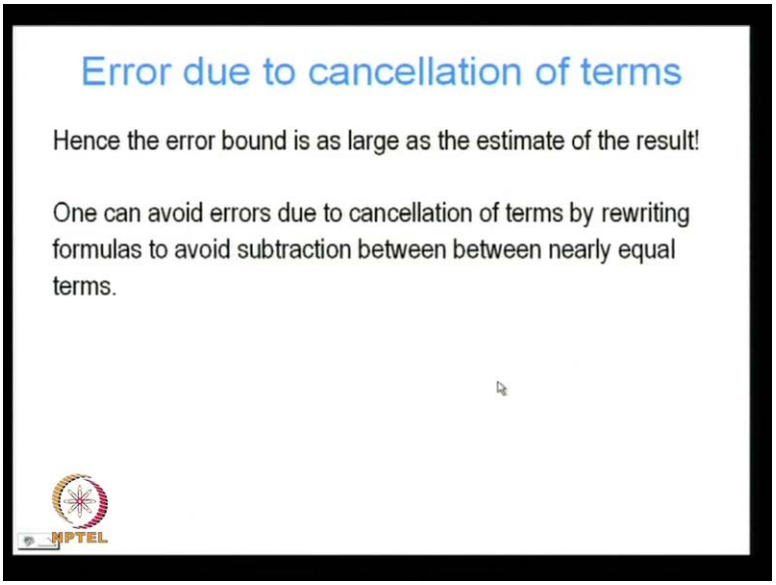
Then the error in y, delta y is given as bound on the error in delta y, mod of delta y is lesser than or equal to mod of delta x 1 plus mod of delta x 2, we have seen this from our last lecture, where we found the bounds on simple operations likes subtraction. So, from that equation dividing both sides by mod of y, we get delta y by y mod of that is lesser than or equal to mod of delta x 1 plus mod of delta x 2 divided by mod of delta x 1 minus x 2. It is clear from the last expression, that the relative error in y will be very large, if mod of x 1 minus x 2 is small that is the two numbers x 1 and x 2 and nearly equal.

(Refer Slide Time: 03:39)



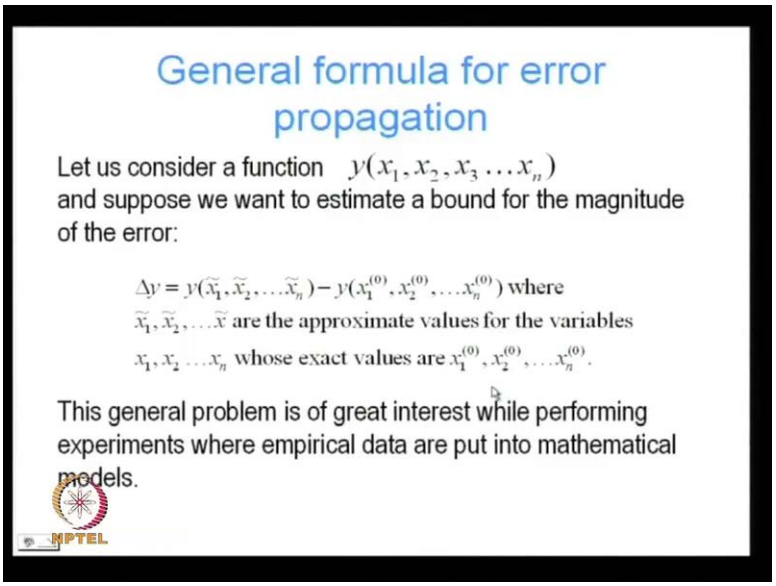
For example, if x 1 is equal to 0.5764 and the absolute error in x 1 that is delta x 1 is 0.5 into 10 to the power minus 4, and x 2 is equal to 0.5763, and the absolute error in x 2 is also 0.5 into 10 to the power minus 4. Then x 1 minus x 2 is equal to 0.0001, but the error bound on the numerical approximation of x 1 minus x 2 is the sum of this absolute errors in x 1 and x 2. So, that is equal to 0.5 into 10 to the power minus 4 plus 0.5 into 10 to the power minus 4 equal to 0.001, thus the error bound on x 1 minus x 2 is as large as x 1 minus x 2, recall x 1 minus x 2 is 0.0001. But, the error bound is actually 0.001, which is actually in order of magnitude greater than x 1 minus x 2.

(Refer Slide Time: 04:54)



Hence, the error bound is as large as the estimate of the result, one can avoid errors due to cancellation of terms by rewriting formulas to avoid subtraction between nearly equal terms.

(Refer Slide Time: 05:10)

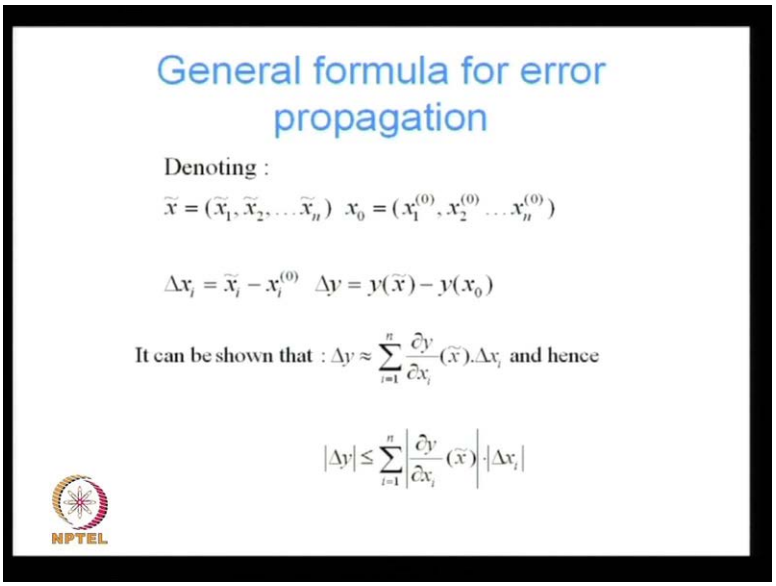


Let us consider in general a function y, which is in general function of x 1, x 2, x 3 up to x n and let us suppose we want to estimate a bound for the magnitude of the error, that is we want to find out a bound on delta y. Where y is actually function of x 1 tilde, x 2 tilde, up to x n tilde, minus y of x 1 0, x 2 0 up to x n where x 1 tilde, x 2 tilde, x n tilde are the approximate values for the variables, while x 1 0, x 2 0, x n 0 are the exact where values of the variables.

So, the variables are x 1, x 2, x n whose exact values are x 1 0, x 2 0 an x n 0 and x 1 tilde, x 2 tilde, x n tilde are the approximate values for those variables. We want to find given that we have the approximate values of x 1 through x n, which is x 1 tilde through x n tilde, we want to find what would be the error in y due to the error in x. This general problem is of great interest while performing experiments, where empirical data are put into mathematical models, for instance we do some experiments where we find the values of x 1, where the independent variables in my experiments are x 1 through x n.

I do that experiment and I find the values x 1 tilde, x 2 tilde, x n tilde, because of some experimental errors x 1 tilde, x 2 tilde, x n tilde are not the except values, which are actually x 1 0, x 2 0 and x n. But, given x 1 tilde, x 2 tilde an x n up to x n tilde I want to find what is the error in y due to the error in my experiment, but due to my experimental error.

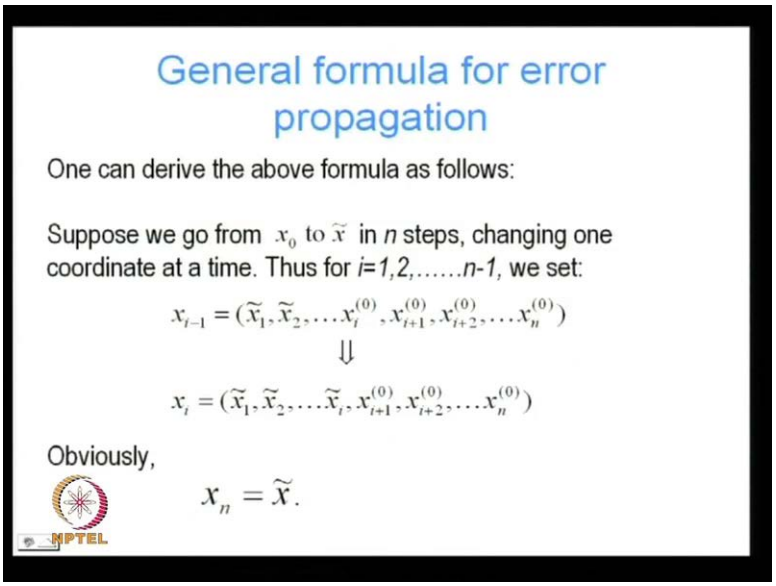
(Refer Slide Time: 07:18)



So, we denote x tilde is equal to x 1 tilde, x 2 tilde through x n tilde and x 0 is equal to x 1 0, x 2 0 through x n 0 and we denote delta x i is equal to x 1 tilde minus x i 0 and delta y is equal to y of x tilde minus y of x 0. That is the numerical solution which is y of x tilde minus the exact solution y of x 0, which gives me my error in y, it can be shown the delta y is approximately equal to the sum of the partial derivatives of y, with respect to the independent variables x i evaluated add the approximate solution x tilde times delta x i.

Since that is true the bound on y must be lesser than or equal to the sum of the individual partial derivatives, times the error in each of the individual variables, that is mod of delta y is lesser than or equal to sigma i is equal to 1 to n mod of del y del x i evaluated at x tilde times mod of delta x i. This result again we get from our previous lecture, were we talked about the error bounds on sums.

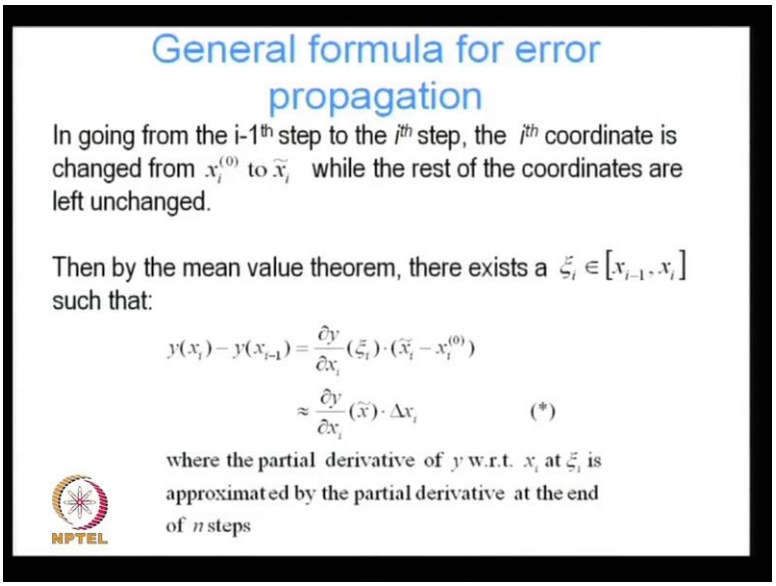
(Refer Slide Time: 09:03)



One can derive the previous formula, this formula as follows, suppose we go from x 0 to x tilde, x 0 being the exact solution, x tilde being the approximate solution in n steps. In each step we change we start with x 0, x 1 0, x 2 0, x 3 0 up to x n 0 at each step, we change one of the individual variables to it is approximate value. For instance, we when we go from step i minus 1 to step i, we change x i 0 to x i tilde, otherwise all the variables, all the independent variables have the same value between x i minus 1 and x i.

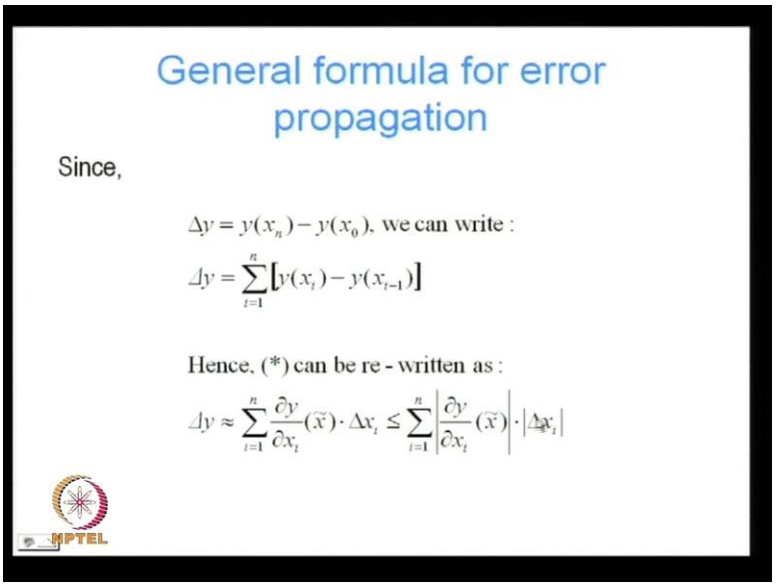
It is only the independent variable x i, we changes value from x y 0 to x i tilde obviously, if you perform this through n steps, at each step changing the value of a one particular independent variable, at the end of the n steps are x n will be equal to our approximate solution x tilde.

(Refer Slide Time: 10:15)



In going from the i minus 1 th step to i th step, the i th coordinate is changed from x i 0 to x i tilde, while the rest of the coordinates are left unchanged. Because of this, due to the mean value theorem, there exists a xi i which belongs to the interval x i minus 1 x xi such that, y of x i minus y of x i minus 1 is equal to del y del x i evaluated at xi i times the interval which is equal to x i tilde minus x i 0, which is approximately equal to del y del x i evaluated at x tilde dotted with delta x i. The approximation is because, we have evaluated the partial derivative of y with respect to x i at xi by the partial derivative at the end of my n steps, that is an evaluating del y del x i at the final value which is x tilde.

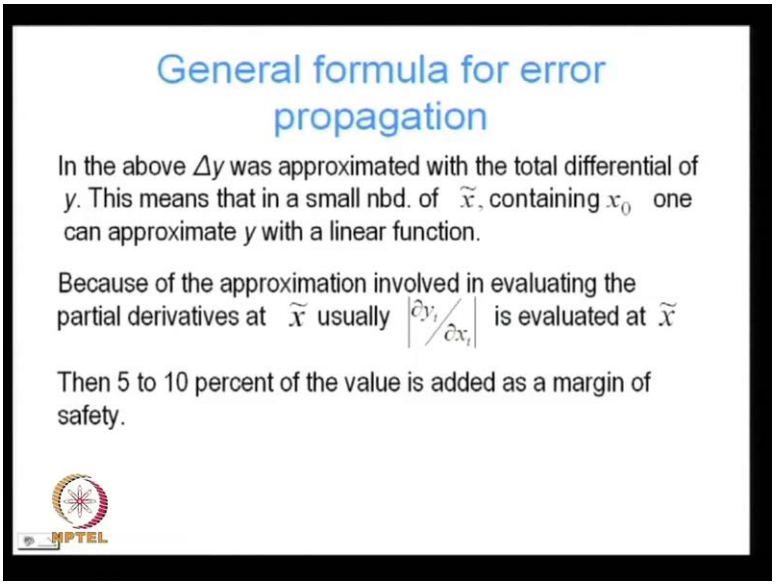
(Refer Slide Time: 11:29)



Since delta y is equal to y of x n minus y of x 0, recall y of x 0 is the solution where with the exact values for the independent variables. And y of x n is the solution where we have replace the independent variables, the exact values of the independent variables with the approximate values. So, delta y is equal to y of x n minus y of x 0, we can actually write delta y as the sum over i to n y of x i minus y of x i minus 1, it turns out that if you do this summation all the intermediate quantities are going to cancel out. And we would be left with y of x n with i equal to n minus y of x 0 were i is equal to 1.

Because of this, we can rewrite this expression as delta y is approximately equal to summation, basically we are taking the sum ((Refer Time: 12:44)), if we take the sum on both sides of this expression. Then we get delta y from this expression, delta y is approximately equal to sigma i equal to 1 to n del y del x i x tilde dotted with delta x i, which again from our bounds an additions is lesser than or equal to sum of sigma i equal to 1 to n mod of delta y del xi x tilde dotted with mod of delta x i.

(Refer Slide Time: 13:20)

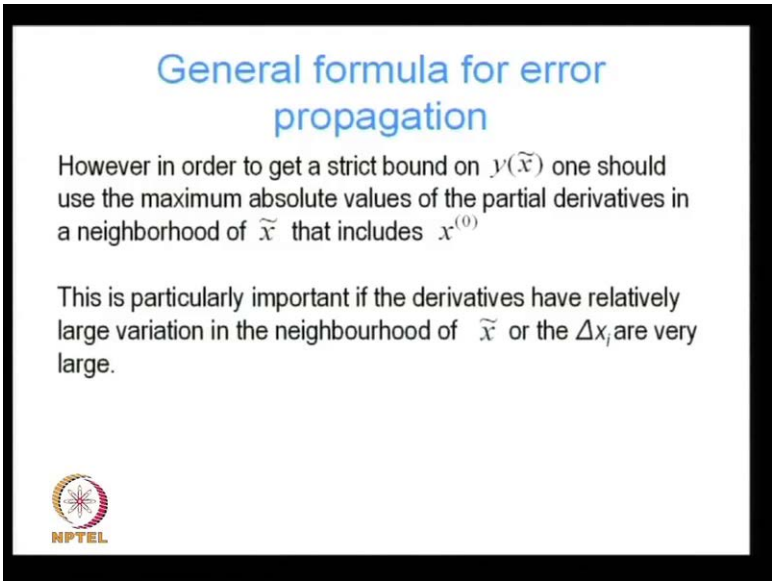


In the above if you go back to our previous equation and look, we are actually approximating delta y by the total differential of y, this means that in a small neighborhood of x tilde containing x 0, we are actually approximating y with a linear function. However, because of the approximation involved in evaluating the partial derivatives of at x tilde, recall that instead of evaluating the partial derivatives at xi, we

are evaluating the partial derivative at x tilde, there is some approximation involved there.

So, what is actually done is that all though we evaluate del y del x i mod of del y del x i at x tilde, 5 to 10 percent of the value is added as a margin of safety. Because, we are not evaluating del y del x i at the exact point xi i,, but at x tilde, which is the approximate value.

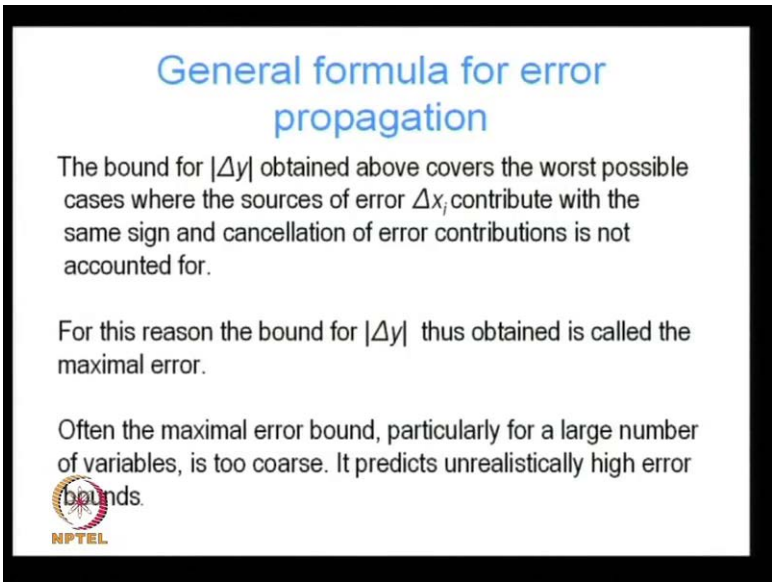
(Refer Slide Time: 14:31)



In order to get a strict bound of y of x tilde, one should use the maximum absolute values of the partial derivatives in a neighborhood of x tilde, that includes x 0. This becomes particularly important, if the derivatives have relatively large variation in the neighborhood of x tilde, or the delta x i are very large. So, we are saying that in order to get a strict bound, we should not really evaluate the partial derivatives at x tilde, we should actually look for the maximum value of the partial derivatives in a neighborhood of x tilde, that includes the true solution x 0.

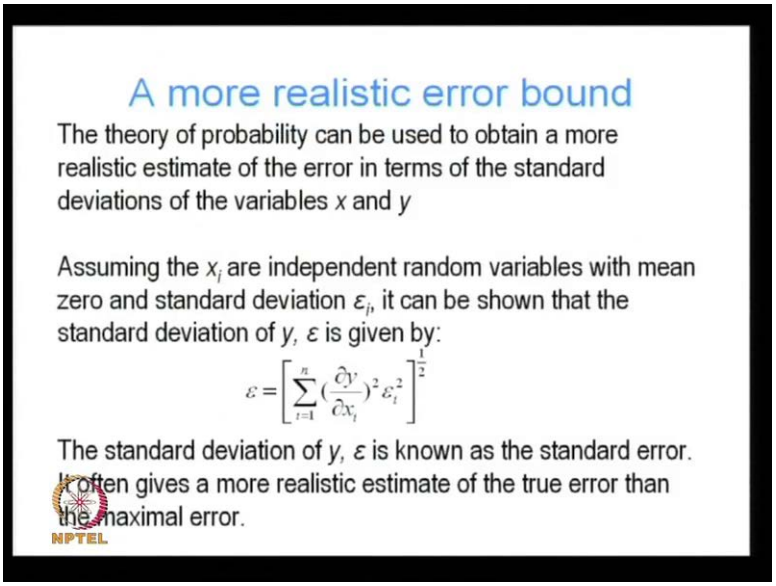
This is important particularly when the derivatives are varying significantly, there is significant variation in the derivatives. So, it is important that way in order to obtain a proper upper bound, that we take the maximum value of the partial derivative in a neighborhood of x tilde.

(Refer Slide Time: 15:39)



The bound for mod of delta y obtained above, covers the worst possible cases where the sources of error delta x i contribute with the same sign, and cancellation of error contributions is not accounted for. For this reason the bound for mod of delta y thus obtained is called the maximal error, but the maximal error bound is often not very useful, particularly for a large number of variables, because it is true codes, it predicts an unrealistically high error estimate.

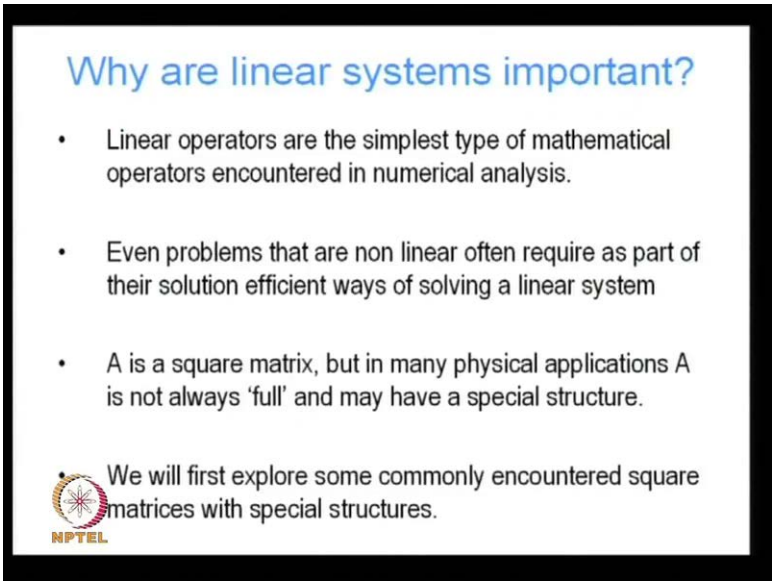
(Refer Slide Time: 16:24)



So, instead of using the maximal error, sometimes we use something which is known as the standard error, how do we get the standard error, we use the theory of probability to obtain the standard error. What do we do, we assume that the x i is are independent random variables with mean 0 and standard deviation epsilon i, if we assumed that it can be shown. That the standard deviation in the dependent variable y is given by a summation of the partial derivatives of y with respect to x i square times the standard deviation square and the square root of that.

If that is epsilon, which is by standard deviation in y and this standard deviation is known as the standard error, it often gives a more realistic estimate of the true error than the maximal error. Because, it gives the narrower bound, if instead of the previous error bound, this error bound this much more realistic.

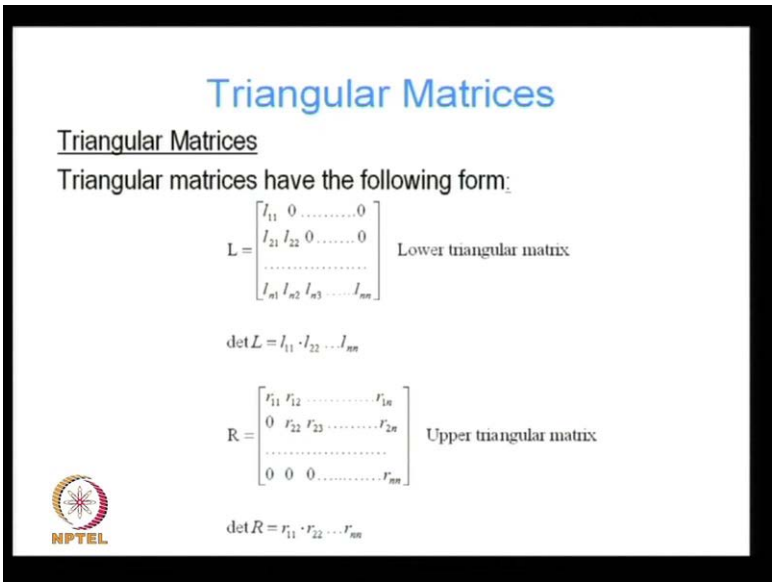
(Refer Slide Time: 17:39)



So, with that discussion we are ready to talk about linear systems, why are linear systems important, because linear operators are basically the simplest type of mathematical operators encountered in numerical analysis. Even problems that are non-linear often require as to obtain solutions of linear systems, for instance if A is a square matrix, and many physical applications A is not always full, and may have a special structure. Recall our definition of linear systems which we obtained earlier, which says that we want to solve the system A x equal to y. Typically in those linear systems A is a square matrix, but in many physical applications A is not always full, and may have a special structure.

We will first explore some commonly encountered square matrices with special structures.

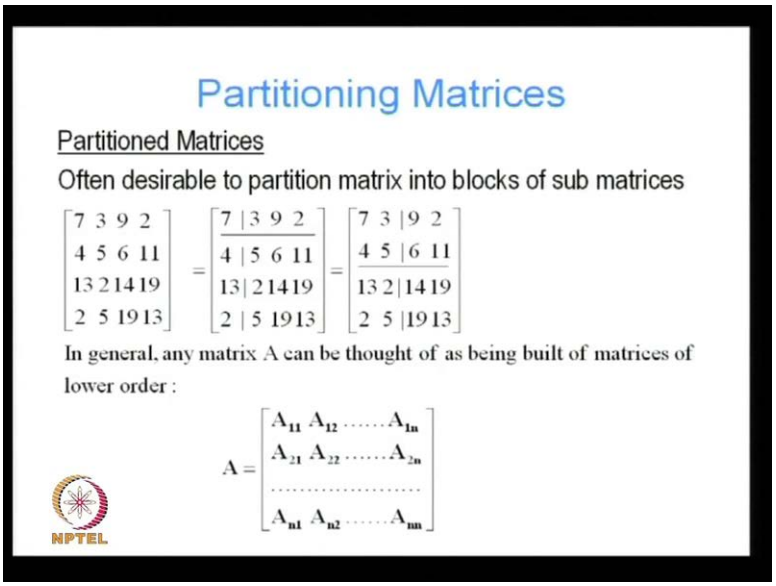
(Refer Slide Time: 18:55)



We will first talk about triangular matrices, triangular matrices have the following form, for instance we have some matrices known as lower triangular matrices, which contained only non-zero terms below or at the diagonal. Either the terms are non zero below the diagonal or along the diagonal, all terms above the diagonal are zero, that is a lower triangular matrix. And upper triangular matrix is a matrix which contains zeros along the diagonal, which contains zeros.

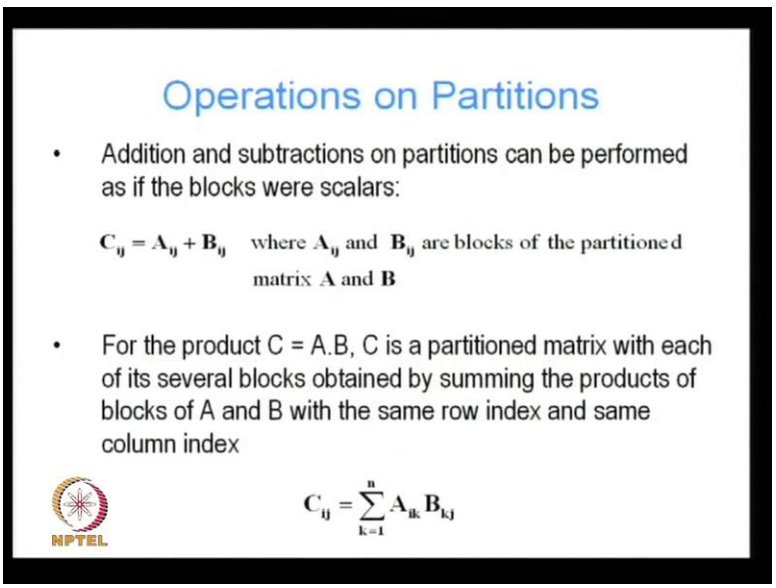
And all terms which are below the diagonal, which contain non zero terms along the diagonal and above the diagonal, that is an upper triangular matrix, but both upper and lower triangular matrix share a common property, which is that the determinant of such matrices are given by the products of the diagonal terms. So, for the lower triangular matrix, the determinant of L is equal to l 1 1 times l 2 2 times l n n, similarly for the upper triangular matrix the determinant of R is given by r 1 1 times r 2 2 times r n n.

(Refer Slide Time: 20:18)



Next the other type of typical matrix that we are going to encounter in the solution of linear systems, are partitioned matrices, it is often desirable to partition a matrix into blocks of sub matrices. In general any matrix can be thought of as being built of matrices of lower order, for instance if A is partition matrix comprising of A 1 1 A 1 2 through A 1 n, A 2 1 A 2 2 through A 2 n, A n 1 A n 2 through A n n. So, these A 1 1 A 1 2 A 1 n these are all actually matrices, but they are partitions of the bigger matrix A, so A 1 1 A 1 2 are all matrices of size smaller than the full matrix A.

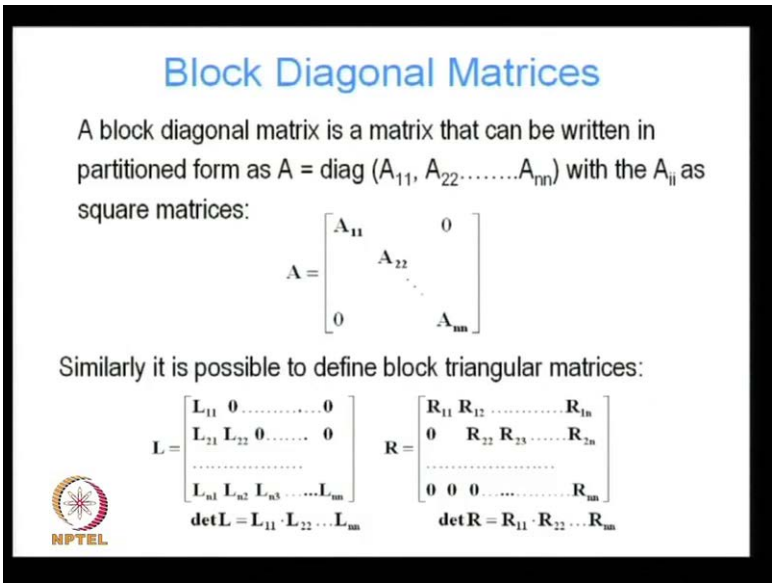
(Refer Slide Time: 21:18)



Operations on partitions addition and subtraction and partitions can be performed as if the blocks were scalars, that is if C i j is the summation of two partition matrices A and B. Then C i j is basically the sum of the partitions A i j plus B i j, A i j and B i j being blocks of the partition matrix A and B. So, basically let me go with that again if A and B are partition matrices with blocks A i j and B i j, then if we want to some those matrices A and B to obtained a third matrix C. Then the blocks of C, C i j a just the sum of the individual blocks A i j and B i j of A and B.

If we want to evaluate the product C of two matrices A and B, where A and B are partition matrices, then we obtained the product by summing the products of blocks of A and B with the same row index, and the same column index. That is if A i k and B k j are partitions of A and B, we obtained C i j a partition of C by taking the product of A i k B k j summing over k equal to 1 to n. So, by treating these individual blocks, as if there elements of a matrix we can obtain the partition C, corresponding to the product of the partition matrices A and B.

(Refer Slide Time: 23:24)



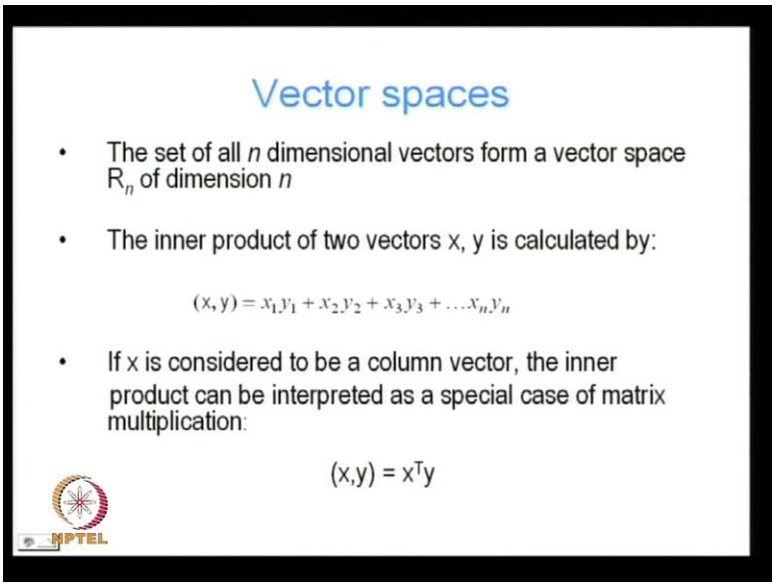
Next we want to talk about block diagonal matrices, a block diagonal matrix is a matrix that can be written in partition form as A is equal to diagonal A 1 1 A 2 2 through A n n that is a is a partition matrix, but A 1 1 A 2 2 A n n are actually matrices, they are not individual terms in the matrix A, they may be full matrices. But, the occupy diagonals on my full matrix A, on my matrix a on the diagonals I instead of having individual terms, I

have actually matrices. So, this type of matrix is known as a block diagonal matrix, it is as if the diagonal elements of A are actually blocks of matrices, blocks partitions.

Similarly, it is possible to define block triangular matrices, which have the same structure is triangular matrices, but instead of individual entities on the diagonal or below the diagonal we actually have blocks of matrices, at these locations. For instance L 1 1 L 2 1 L 2 2 are actually blocks of matrices, but they have with the same structure as a lower triangular matrix. Similarly, a block upper triangular matrix has the same structure, as an upper triangular matrix the only difference is this R 1 1 R 1 2 R 1 n, instead of being individual scalars are actually individual matrices.

However, when we tried to the evaluate the determinant of this block triangular matrix L, we can basically do the same, we can do the same thing that we did earlier. So, determinant of L is equal to L 1 1 times L 2 2 times L n n exactly like we did earlier, similarly determinant of R is equal to R 1 1 times R 2 2 through R n n. So, the same structure follows, the same structure with individual upper triangular, lower triangular matrices had passes over continues into block triangular matrices.

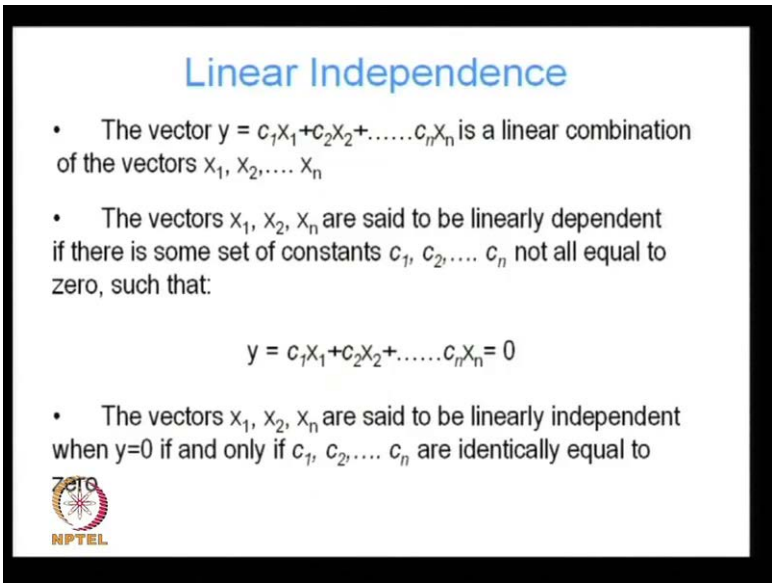
(Refer Slide Time: 25:51)



Next we want to talk about vector spaces, the set of all n dimensional vectors form a vector space R n of dimension n, for instance the set of all three dimensional vectors form a vector space R 3 of dimension three. Similarly, for two dimensions, the set of all two dimensional vectors form a vector space R 2 of dimension two. The inner product of

two vectors x and y is calculated as follows, how we are defining the inner product, we are saying that the inner product is defined by the product of the components summed together. So, the inner product of two vectors x y is equal to x 1 y 1 plus x 2 y 2 plus x 3 y 3 through x n y n. If x is considered to be a column vector, the inner product can be interpreted in matrix notation, the inner product between x and y can be interpreted in matrix notation, as the transpose of x times y is the product of the transpose of x and y.

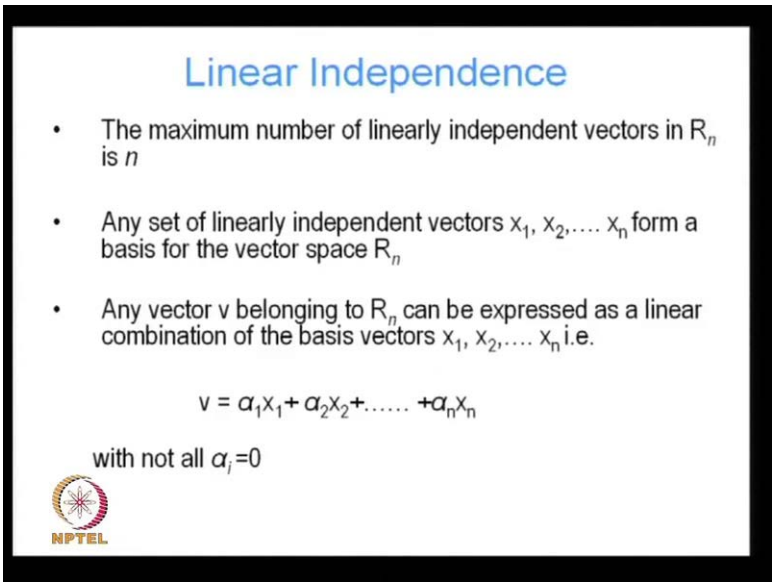
(Refer Slide Time: 27:17)



Next we want to talk about the notion of linear independence, a vector y which is obtained by taking a linear combination of vectors x 1, x 2 through x n by multiplying them with scalar C 1 through C n, is said to be in linearly independent. If there is some set of constant C 1 C 2 through C n not all equal to 0 such that, that sum C 1 X 1 plus C 2 X 2 through C n X n is equal to 0. On the other hand, the vectors x 1, x 2 through x n are said to be linearly independent, when this condition is only satisfied if and only if all the C 1 through C n are 0.

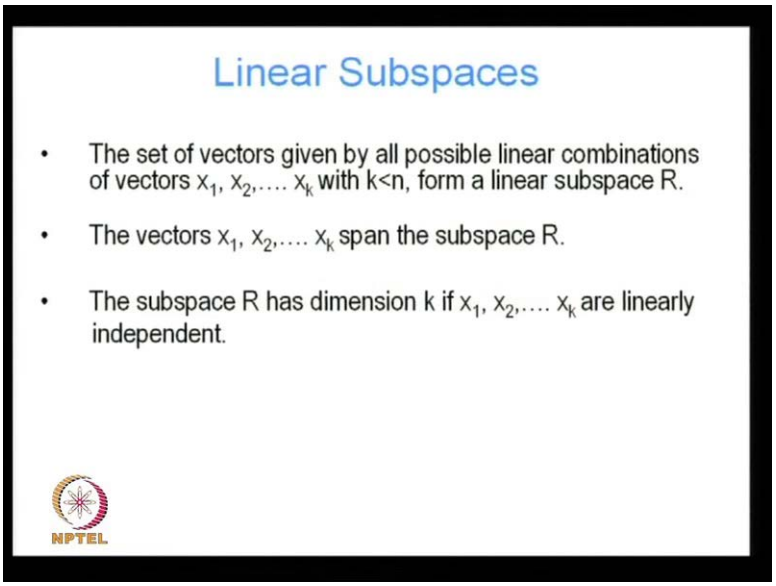
So, x 1 through x n are linearly dependent, if there is some set of constants C 1 through C n not all equal to 0 such that, C 1 X 1 plus C 2 X 2 plus through C n X n is equal to 0, X 1 through X n is linearly independent if that condition is only satisfied if all the C 1 are 0. Basically we are saying there is no set of non zero constants C 1 through C n such that, we can sum x 1, x 2 through x n after multiplying them with the scalars C 1 through C n such that, we get 0.

(Refer Slide Time: 29:00)



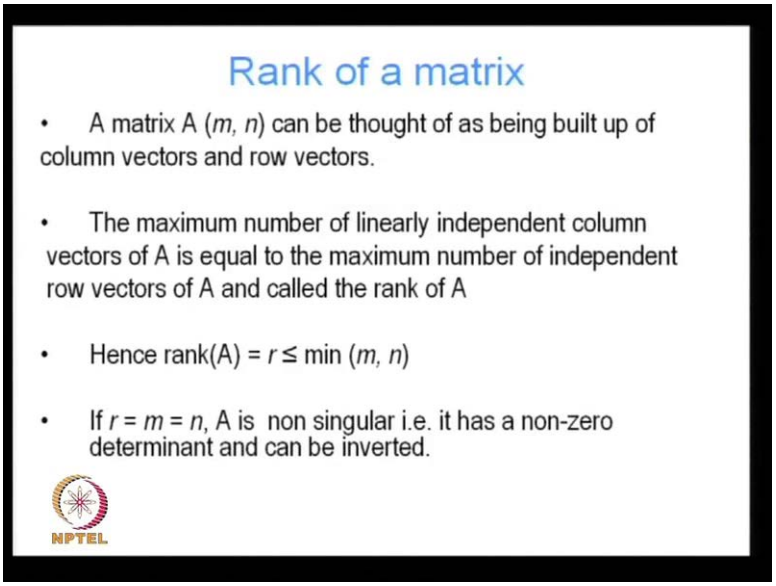
The maximum number of linearly independent vectors in R n is n, any set of linearly independent vectors x 1 through x n form a basis for the vector space R n. Any vector v belonging to R n can be expressed as a linear combination of the basis vectors x 1 through x n, that is v if x 1 through x n comprise a basis for this space R n. Any vector v can be represented as a linear combination of x 1 through x n, what do you mean by a linear combination, we can right be as a scalar multiple alpha 1 times x 1 plus alpha 2 times x 2 plus alpha 3 times x 3 through alpha n times x n, were not all the alpha are equal to 0.

(Refer Slide Time: 30:07)



Next we want to talk about linear subspaces, the set of vectors given by all possible linear combinations of the vectors x 1, x 2 through x k with k less than n form a linear subspace R. And the vectors x 1, x 2 through x k are set to span the subspace R, the subspace R has dimension k, if x 1, x 2 through x k are linearly independent, that is the subspace R k is has got basis vectors x 1, x 2 through x k.

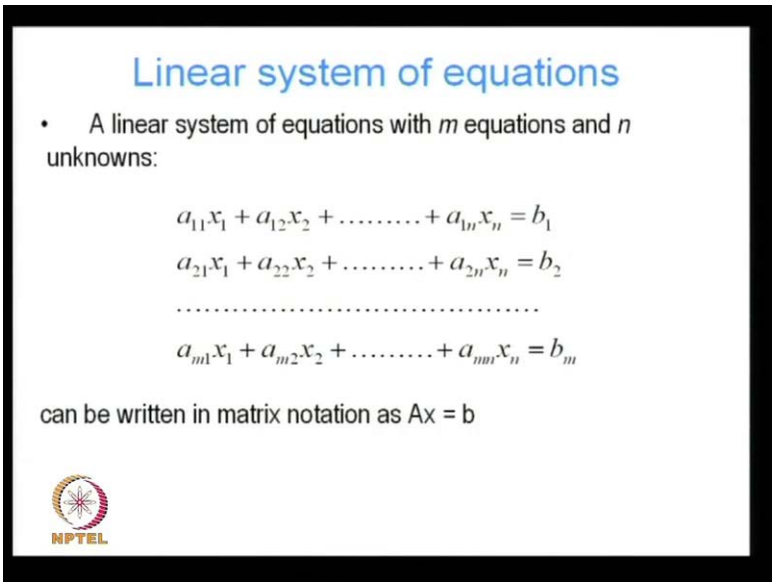
(Refer Slide Time: 30:59)



Next we want to talk about the rank of A matrix, a rank of a matrix A is set to be r, if r is the maximum number of linearly independent column vectors or row vectors of A. So, basically the rank is the maximum number of linearly independent column vectors or row vectors of A; and those the maximum number of linearly independent column vectors of A is always equal to the maximum number of independent row vectors of A. Rank of A therefore, must be lesser than or equal to the smaller of m and n, where m is the number of rows in A, and n is the number of columns in A.

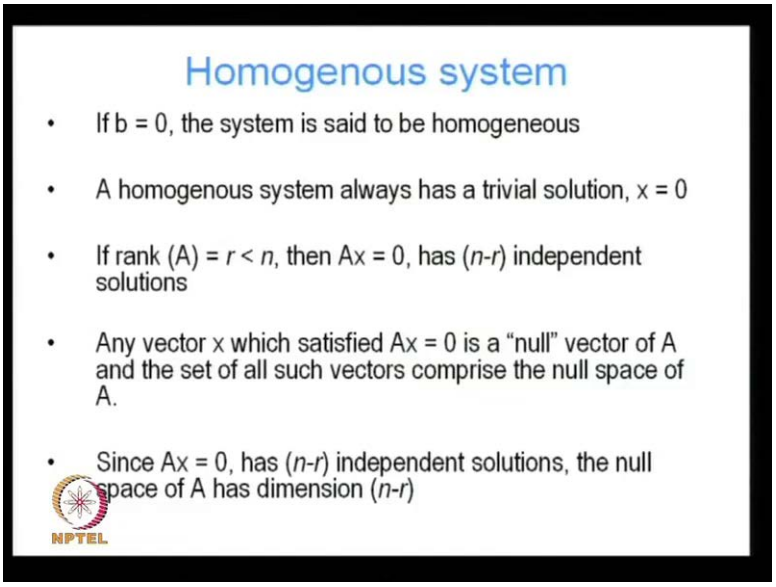
Since, the number of linearly independent columns must be equal to the number of linearly independent rows, it is obvious that the rank which is equal to the number of linearly independent rows, and number of linearly independent columns must be lesser than or equal to the minimum of m and n. If the rank is equal to m is equal to n that is if m is equal to n means A is a square matrix, and if the rank is equal to m and equal to n, then the matrix A is said to be nonsingular that is it has got a non zero determinant, that is it can be invertible.

(Refer Slide Time: 32:43)



A linear system of equations with m equations and n unknowns, we will be actually write it out, which we have in done up till now can be written like this a 1 1 a 1 2 a 1 n are the components of the co-efficient matrix A. x 1, x 2 through x n are my independent variables x, I can think of x as a vector with of dimension n and b is my right hand side, and we want to solve for x 1, x 2 through x n with satisfies this condition A x equal to b.

(Refer Slide Time: 33:34)

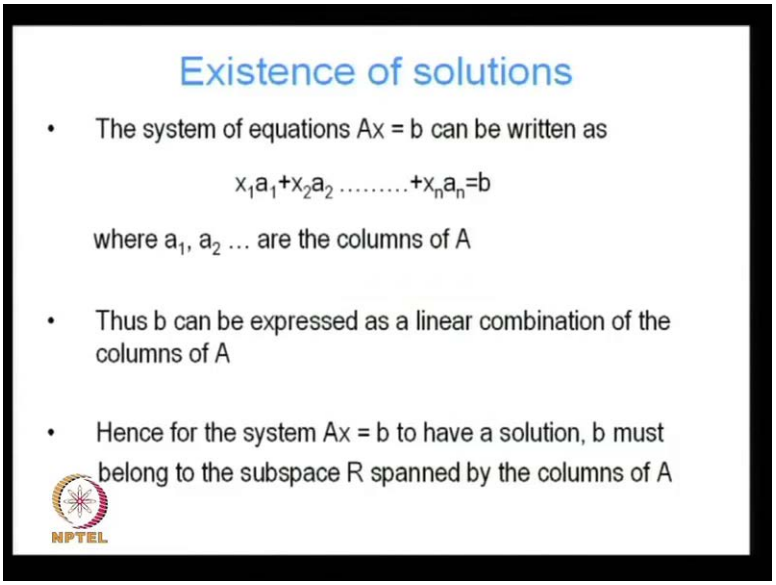


If b is equal to 0, the system is said to be homogeneous, so A x is equal to 0 is said to be a homogeneous system, a homogeneous system always has a trivial solution, what do we

mean by trivial solution, that is a homogeneous system will always be satisfied by x 1, x 2, x 3 to through x n equal to 0. If rank of A is equal to r is less than n, then A x equal to 0 has n minus r independent solutions, any vector x which satisfies A x is equal to 0 is a null vector of A.

And the set of all such vectors comprise the null space of A, the null space of A has dimension n minus r, because as we recall the set that if rank of A is equal to r less than n, then A x is equal to 0 has n minus r independent solutions, so the null space of A has dimension n minus r.

(Refer Slide Time: 35:00)

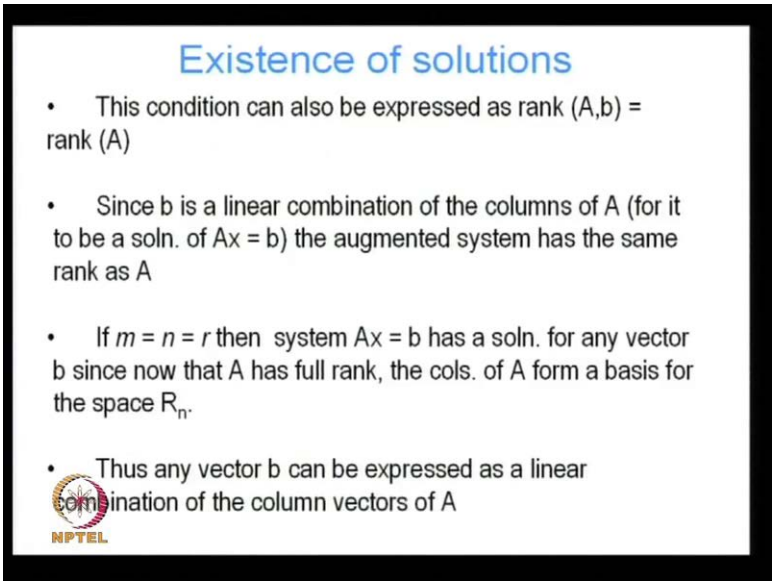


The system of equations A x is equal to b can be written as x 1 a 1 plus x 2 a 2 through x n a n equal to b, basically we are taking the rows of A. And multiplying scaling each row by x 1 through x n scaling the first row with the x 1, the second row with the x 2, the third row with the x 3 and so on, and so forth. And the last row a n by x n and setting that equal to b, this is just another way of writing the system A x equal to b, but this way of writing makes clear that b is actually a linear combination of the columns of A, which again makes it obvious that if the system A x is equal to b is to have a solution.

b must belong to the subspace r spanned by the columns of A, because we call that x 1 a 1 plus x 2 a 2 plus x n a n is equal to b, which basically means that b is a linear combination of a 1 a 2 through a n, x 1 and x 2 x through x n are just the scalar multiples. So, b is a linear combination of the columns of A, a 1 a 2 through a n thus b must belong

to the subspace R which is spanned by the columns of A, so in this case columns of A are the basis of that subspace. And b is obtained by taking a linear combination of those basis vectors a 1 a 2 through a n, after scaling them with the scalars x 1, x 2 through x n.

(Refer Slide Time: 37:05)

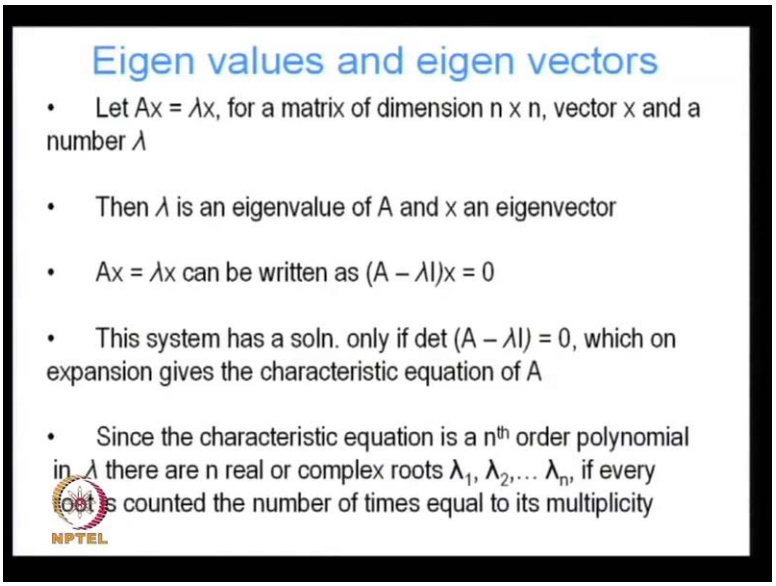


This condition can also be expressed by the requirement that rank of A, b is equal to rank of A, for this means since b is a linear combination of the columns of A that means, if we add we if we augment A by adding an additional column b we do not change the rank of A. Because, b is a linear combination of the columns of A, so we add an additional column to a which is b, it does not change the number of linearly independent columns of A. So, since b is a linear combination of the columns of A, the augmented system has the same rank as A.

If m is equal to n is equal to r that is if A is square matrix, an as got full rank, then A x equal to b has a solution for any vector b, since now that a has full rank, the columns of a form a basis for the space R n. A has full rank means, A has rank n that means, each of the columns of A are linearly independent vectors and since, the space has dimension n and since I have n linearly independent vectors, which are the columns of A, which means that this implies that the columns of A form a basis for the space R n.

And since the form a basis for the space of R n, any vector b can be represented by a linear combination of the column vectors of A, thus any vector b can be solved for, even any vector b we can find x such that, A x is equal to b.

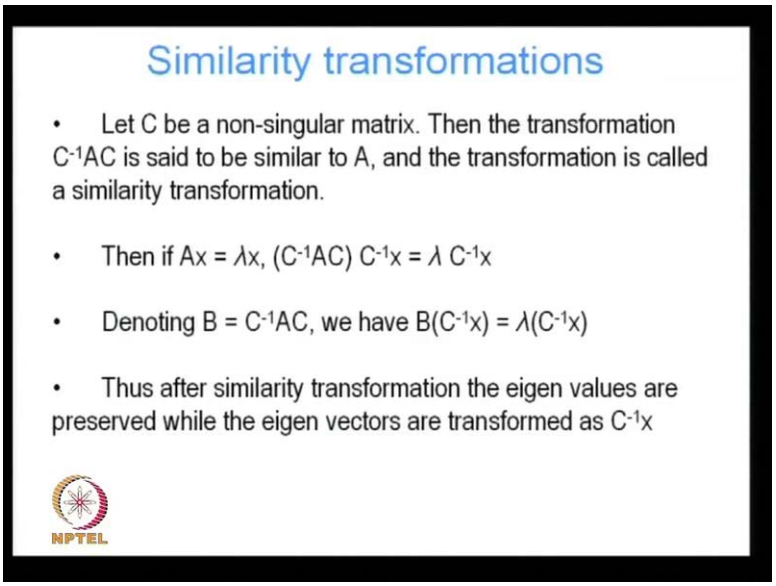
(Refer Slide Time: 39:12)



Next we wish to talk about Eigen values and eigen vectors, let A x is equal to lambda x, for a matrix of dimension n by n, a vector x and a scalar lambda. Then lambda is an Eigen value of A and x is an eigen vector of A, we can rewrite A x is equal to lambda x as a minus lambda i, where i is the identity matrix. A minus lambda i operating on x is equal to 0, notice that A minus lambda i x equal to 0 is a homogeneous system, from what we have looked at previously, this homogeneous system is only going to have a solution, if determinant of a minus lambda i is equal to 0.

If we expand this determinant out, if A write out the determinant, determinant A minus lambda i we get a polynomial in lambda, and that polynomial in lambda is called the characteristic equation of A. And depending on the dimension of A, the dimension of the polynomial also varies for instance, if A is a 3 by 3 matrix the characteristics equation is a polynomial is a cubic, is a polynomial of dimension three. If A is the 4 by 4 matrix the characteristic equation is a polynomial of dimension four, and depending on the roots of the polynomial we get the Eigen values of A, a lambda 1 lambda 2 lambda n which may be real or complex numbers.

(Refer Slide Time: 41:25)



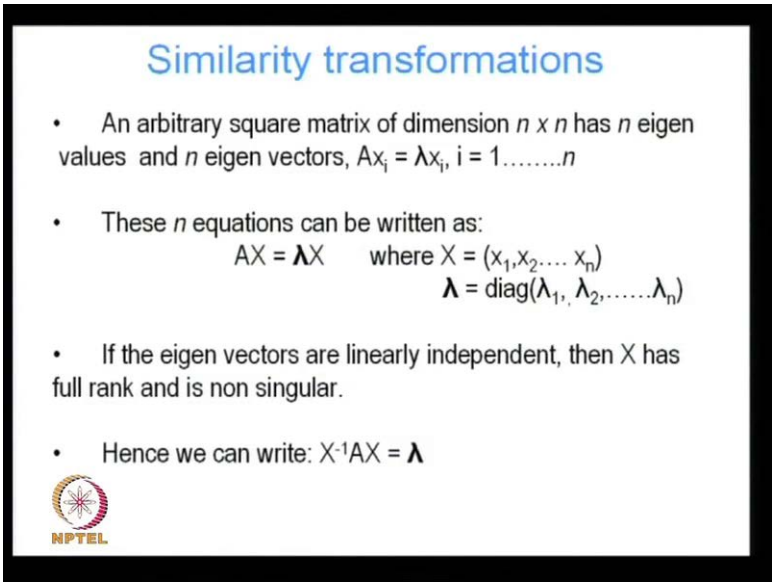
But, if we count each root equal to the number of times equal to it is multiplicity, we get n real or complex roots of that polynomial which are the Eigen values of A. Next, we want to talk about similarity transformations, let C be a non singular matrix that is C has determinant which is non zero, that is determinant of C is greater than 0. Then the transformation C inverse A C is said to be similar to A, that is if we start with the original matrix A.

And when we transform it according to the rules C inverse A C, then the matrix we get C inverse A C is said to be similar to A, and the transformation is called a similarity transformation. Let us go back to the Eigen value problem A x is equal to lambda x and suppose, we perform a similarity transformation on A to get C inverse A C, we can see that if we pre multiply both the left hand and right hand side with the matrix C inverse, we get C inverse x A X is equal to lambda times C inverse of x C inverse of x.

I can rewrite as C inverse A C operating on C inverse times x, because C and C inverse on taking the product they give me the identity matrix. So, I can rewrite A x is equal to lambda x as C inverse A C operating on C inverse x, which is equal to lambda operating on C inverse x. But, in this last equation, this has the exact structure of an Eigen value problem which becomes evident, if we replace C inverse A C with B, now so that we can write B operating on C inverse x is equal to lambda times C inverse of x.

Thus we see that after similarity transformation of A, the similar matrix C inverse A C has the same eigen values of as A, it has a same Eigen values lambda, but it is eigenvectors are no longer x. They are now C inverse x, thus after similarity transformation the eigen values are preserved, where the eigenvectors are transformed as C inverse of x.

(Refer Slide Time: 44:30)

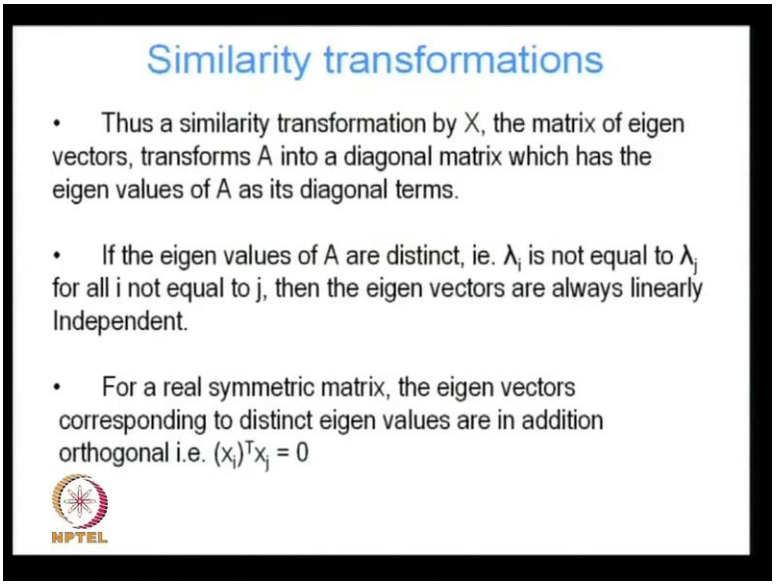


An arbitrary square matrix of dimension n by n has n eigen values and n eigen vectors, each of which satisfied the relation A x i is equal to lambda x i, these n equations. If we can combined all these equations together, we can write it in matrix form as A x is equal to lambda x. Where now A x is no longer of vector x is a combination of all this eigen vectors each eigen vector occupies a column of the capital matrix x. So, x 1, x 2 through x n are individual eigen vectors and have combined them together, to form the matrix big x and lambda is another matrix, whose which is basically a diagonal matrix.

And the diagonal elements of lambda are the individual eigen values lambda 1, lambda 2 through lambda n corresponding to the eigen vectors x 1, x 2 through x n. If the eigen vectors are linearly independent X is going to have full rank, and is going to be non-singular why, because if the eigenvectors are linearly independent each of the columns of A are linearly independent. And each of the columns of x are linearly independent, x has full rank and that is x can be invertible and x is non-singular. Since, x is invertible we can write x inverse A x is equal to lambda, basically we scale, we pre multiply A x is

equal to lambda x with x inverse both on the left hand and the right hand side, when we get x inverse A x is equal to lambda. And recall lambda is a diagonal matrix that is it has got 0 is everywhere, and lambda 1 lambda 2 through lambda n on the diagonals.

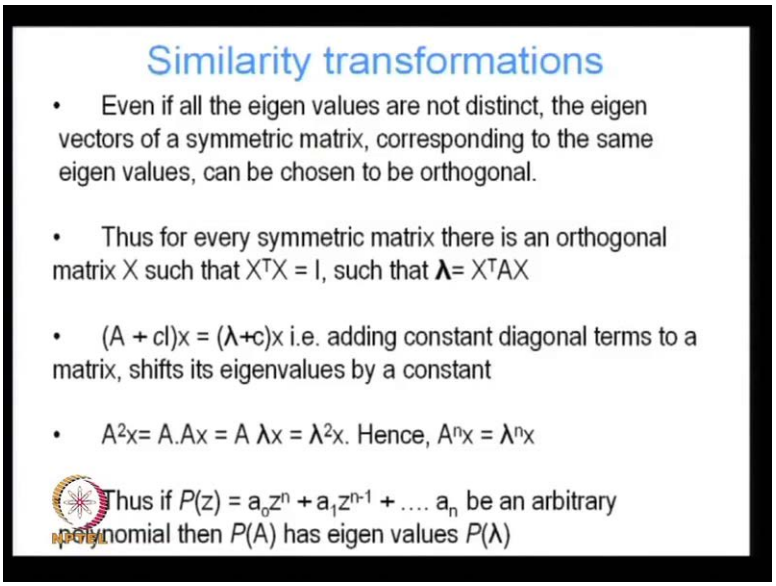
(Refer Slide Time: 46:48)



Thus a similarity transformation by X, the matrix of eigenvectors recall that earlier the transformation that we look that, this relationship this is basically a similarity transformation on A. So, what we are saying is that, if we do a similarity transformation on A using the matrix X, we get a diagonal matrix we transform a into a diagonal matrix which has the Eigen values of A as it is diagonal terms. If the Eigen values of A are distinct that is lambda i is not equal to lambda j for any e i or j, then the eigen vectors are always going to be linearly independent.

So, if a matrix has got distinct eigen values that each of it is Eigen values are different from each other, then the eigenvectors are also going to be linearly independent. And we know that for a real symmetric matrix, the eigen vectors corresponding to distinct Eigen values, in addition to being linearly independent they are also orthogonal. So, that is not really true for always, but it is only true for real symmetric matrices, when the eigenvectors corresponding to distinct eigen values are mutually orthogonal. What it means, what does orthogonal mean it means that, if we take the dot product of two of those eigen vectors x i dotted with x j or been matrix notation x i transpose x j, we are going to get 0.

(Refer Slide Time: 48:50)



Even if all the eigen values are not distinct, the eigenvectors of a symmetric matrix corresponding to the same eigen values can be chosen to be orthogonal, there are procedures for orthogonalization, which we are going to talk about later in this course. For basically for a symmetric matrix corresponding to distinct eigen values, corresponding to an eigen value, if there is more than one eigen vector we can make those eigenvectors orthogonal to each other.

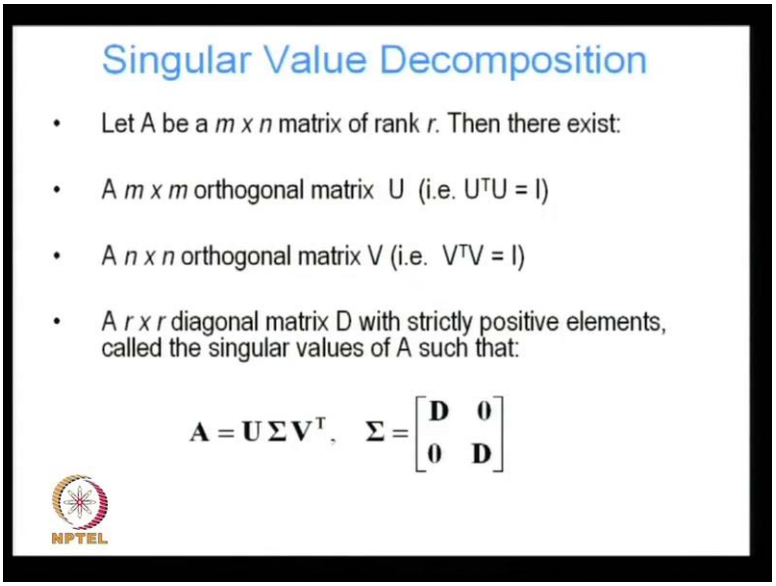
We can define the eigen vectors, so that dot product to those eigenvectors is equal to 0, thus for every symmetric matrix, there is an orthogonal matrix x such that, X transpose x is equal to i that is such that, lambda is equal to X transpose A x. Next, we want to see how the eigen values change if we perform certain simple operations on a matrix, for instance we have an initial matrix A. And suppose we add the same constant scalar C to each of the diagonal terms of A, that is we form the matrix A plus C i, i being the identity matrix.

In that case what happens is that the eigen values of a just get shifted, that is instead of the eigen values remaining as lambda for the eigen values become lambda plus C. So, adding a constant to the diagonal term just shifts the eigen value by a constant C, so that is one property of eigen values the second property, we want to talk about is. If we take powers of a matrix for instance, if we take a product if a matrix with itself, for instance A square x I can write it as A product with A x, but A x I know I can write as lambda x.

Thus A operating on A x I can write as a operating on lambda x, but we can since lambda is a scalar I can move it out and so we have lambda times A lambda x is equal to lambda times A x, but again A x is again equal to lambda x. So, we are going to get lambda square x, thus if we take a power of matrix, if we take n th power of A, then the eigen value of the n th power of A, is basically the eigen value of A raised to the n th power.

So, the eigen value of A to the power n is equal to lambda to the power n, because of this property if P is a polynomial of the form A 0 z to the power n plus A 1 z to the power n minus 1 through A n, an instead of z if we replace that by the matrix A, then we can see that P A is going to have eigen values P of lambda.

(Refer Slide Time: 52:31)



Finally, for the today's lecture I want to briefly introduce something, which is known as singular value decomposition. In our next lecture we are going to talk in much greater detail about this, but I just want to introduce this, so suppose A is an m by n matrix of rank R, then there will always be m by m orthogonal matrix recall what is an orthogonal matrix. An orthogonal matrix is a matrix all of whose columns are orthogonal to each other, so if I take the dot product of a column i, with a dot product of another column j I am going to get 0.

So, what I am saying is that, if we can always do a singular value decomposition of a matrix A into orthogonal matrices U and V and an r by r diagonal matrix D such that, I

can represent a as a product of U sigma and V transpose. A sigma is such the D occupies the diagonal elements of sigma.

Thank you, we will continue with this discussion in our next lecture.