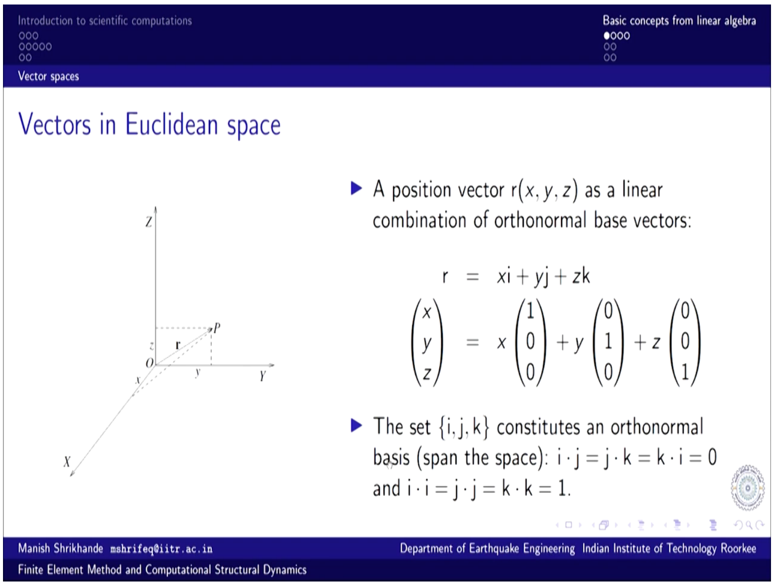
Finite Element Method and Computational Structural Dynamics
Prof. Manish Shrikhande
Department of Earthquake Engineering
Indian Institute of Technology, Roorkee

Lecture - 03
Basic Concepts of Linear Algebra

Hello. So, as we discussed in our last lecture about the Errors in Floating Point Operation, the next logical point of discussion is how the approximate solutions are constructed.

(Refer Slide Time: 00:44)



And, that is based on the concept of vector spaces, how approximate solution is created and then how do we analyze about the quality of approximation and that is what we will discuss in our lecture today.

Now, we move on to the next topic that is also essential, that also forms the basis of all approximations. So, vector spaces, so vectors in Euclidean space we all know about the vectors from high school levels that is the first time we probably started using vectors.

So, Cartesian coordinate system x, y, z and (Refer Time: 01:38) point P can be represented as a summation of 3 unit vectors. So, a position vector r that you can see, that designates the defines the point P in the any 3-dimensional geometry (Euclidean geometry) can be represented as a linear combination, I mean when I say linear combination it is single power.

So, linear combination i + j + k, some proportion of i plus some proportion of j plus some proportion of k, where i, j, k are the unit vectors that we so very well know. So, that represents the position vector r. So, the triad of the coordinates x, y, z can be represented as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

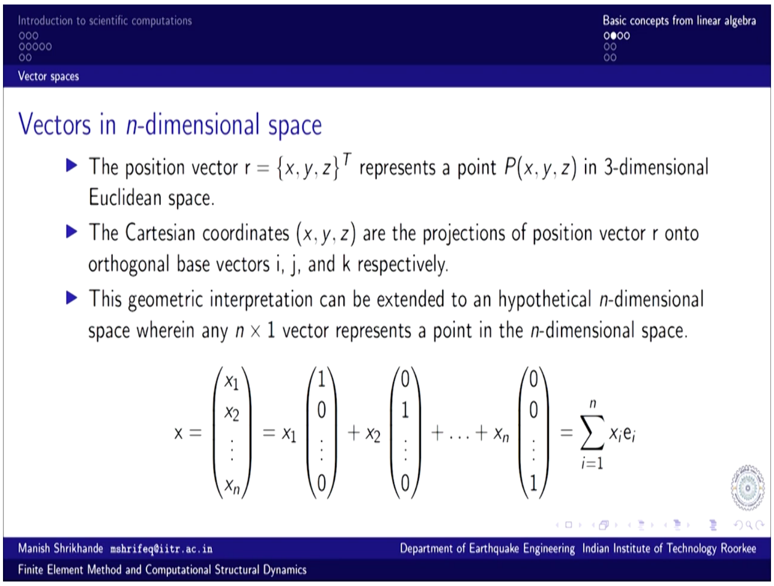
So, this vector x, y, z, so, this is position vector r that can be represented as a linear combination of these 3 base vectors. So, these are orthonormal base vectors we call them orthonormal because they are all of unit length, so they are unit vectors, so, unit length. And they are all perpendicular to each other that is when I say perpendicular to each other or orthogonal to each other, the projection of one vector on to another base vector is 0. So, if I try to compute the projection of i along the direction of j that is 0. And if I try to compute the projection of unit vector j along unit vector k that is again 0 and that projection is given by dot product between these vectors. So, $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ multiplied by $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, that is the dot product. So, that will be again if you calculate, so $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ transpose that is the row vector and then column vector $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ that will yield 0. So, that is the dot product that projection and that works out to 0. So, that is the orthogonality.

So, the (i, j, k) constitutes an orthonormal basis. When I say basis it means that it spans the space; that means, any vector in 3-dimensional geometry can be represented as a linear combination of these 3 orthonormal vectors, base vectors. Now, these are by no means the only set of base vectors, these are only the convenient base vectors that we deal with.

But as I said the orthogonality properties, dot product between i and j is equal to dot product between j and k and that is equal to the dot product between k and i, and that is equal to 0 because orthogonal vectors will not have any projection on other direction. And also they are all unit length, so the dot product with themselves is equal to unity.

So, that is only a matter of definition. But what is important is they constitute orthonormal basis, the significance of basis is that it spans the space. When I say spanning the space it means essentially what I am trying to say here is any vector in this 3-dimensional geometry can be represented as a linear combination of these 3 base vectors. So, as you can see in the context of this position vector, position vector can be anywhere in the 3-dimensional geometry and this can be represented as a base vector.

(Refer Slide Time: 06:18)



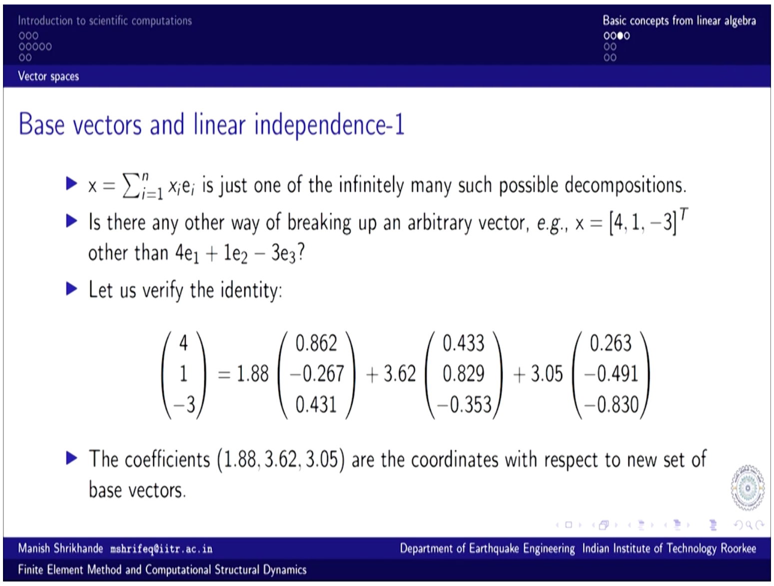
So, as I said, now we can extend this from 3-dimensions, we can extend it to n-dimension just a simple mathematical extension. So, we can talk of instead of 3-dimension, we can talk of n-dimensional space. So, just as a position vector r = $\begin{matrix} x\ x \\ y\ y \\ z\ z \end{matrix}$ represents a point P (x, y, z) in 3-dimensional Euclidean space. Any vector n-dimensional and the Cartesian coordinates x, y, z they are the projections of position vector r on to the, orthogonal vectors.

So, if I have this position vector r, I take its projection along x axis along the unit vector i that gives us the coordinate x coordinate. I take the projection along unit vector j that gives me the coordinate y and I take the projection along the unit vector k along the z direction and that gives me the coordinate z. So, those are the Cartesian coordinates because those are the projections of the position vector along the base vectors i, j and k respectively. So, Cartesian coordinates are the projection of position vector r onto orthogonal base vectors i, j and k respectively.

So, this geometric interpretation can be extended to an hypothetical n-dimensional space, vector space, where we have a vector of n number of elements and I can represent it as a linear combination of base vectors. So,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^{n} x_i e_i$$

(Refer Slide Time: 08:31)



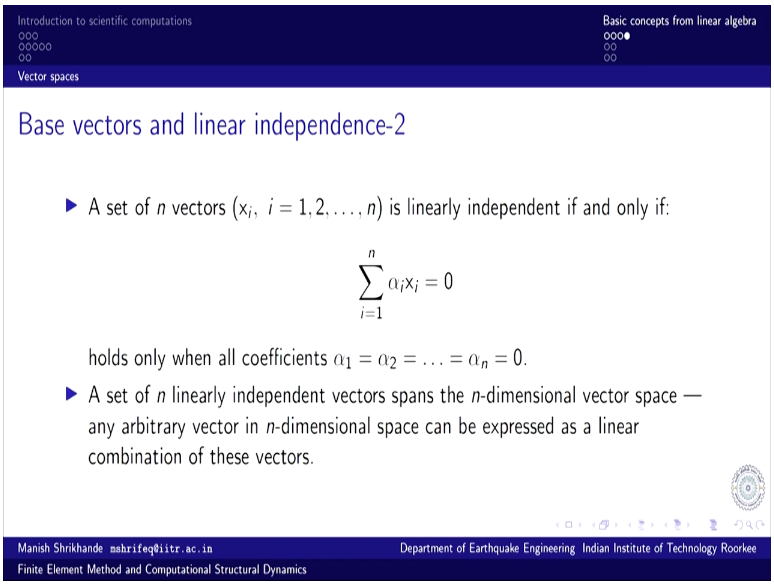
So, now x is equal to x= $\sum_{i=1}^{n} x_i\, e_i$ $\sum_{i=1}^{n} x_i\, e_i$ is just one of the infinitely many such possible decompositions. It is a convenient decomposition very easy to construct, but it is not the only possibility. There are in principle infinite such possibilities, I can find out infinite set of base vectors which mutually orthogonal and of unit length, so orthonormal set of vectors which can be used in similar fashion for the expansion of the any arbitrary vector, right.

So, just as an numerical example to explain the concept, so this arbitrary vector again we will work with vector of 3-dimension just for it is easier to understand, it is easier to visualize, but the results are very general. It can go up to any value of n even extending to infinity. So, any vector let us say x= $[4,1,-3]^T$ $[4,1,-3]^T$ can be represented in any way other than 4 e_1 + 1 e_2 - 3 e_3 that is the Cartesian representation using e_1, e_2, e_3 are synonyms for i, j and k. So, let us look at it from the other way around. I just have these numbers. Does this add up to this left-hand side is equal to right hand side? So, 1.88 again I am using 3 significant digits here.

$$1.88 \begin{pmatrix} 0.862 \\ -0.267 \\ 0.431 \end{pmatrix} + 3.62 \begin{pmatrix} 0.433 \\ 0.829 \\ -0.353 \end{pmatrix} + 3.05 \begin{pmatrix} 0.263 \\ -0.491 \\ -0.830 \end{pmatrix}$$

And you will find that within the finite precision arithmetic more or less you will get the left hand side is equal to right hand side. So, what does this mean? vectors that we see here this vector here and this vector here and this vector here, together these 3 also constitute a base vector, a set of base vectors. So, that is also is basis in this 3-dimensional space. And these the numbers they are the coordinates. And I can write this as a linear transformation, so coefficients 1.88, 3.62 and 3.05 they are the coordinates with respect to new set of base vectors. So, I just effected a transformation.

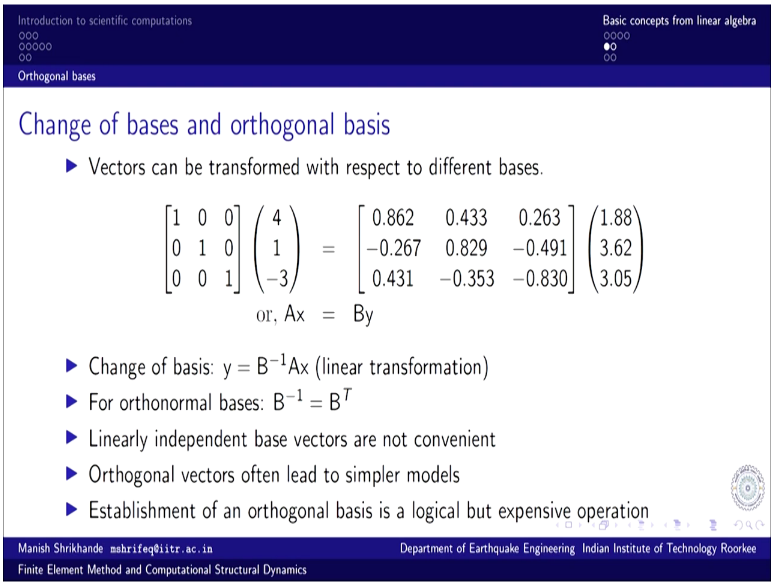
(Refer Slide Time: 11:24)



So, these are the basic definition. Just to emphasize any set of vectors is linearly independent as long as any linear combination is equal to null vector only if all the coefficients alpha 1, alpha 2, alpha 3 that are the coefficients of the vectors associated with it. Only when these coefficients are 0, only then this linear summation is equal to 0, then that set of vectors is called linear independent linearly independent. And it is important for any set of vectors to be constituted as a base vector that they have to be linearly independent first. Orthogonality is a very stringent criterion. All orthogonal vectors will of course be linearly independent, but not all linearly independent vectors need be orthogonal, right. So, orthogonality is a much more stringent criterion.

So, a set of n linearly independent vectors they will they are sufficient to span the n-dimensional vector space and any arbitrary vector in that n-dimensional space can be represented as a linear combination of these vectors. And this is very fundamental statement which allows us to construct approximate solutions. We will come to this in the next few lectures when we start constructing finite element approximations, how do we derive finite element approximations. So, the basis is essentially this. So, if I have a set of linearly independent vectors then I can construct any arbitrary vector can be represented in terms of independent, n number of linearly independent vectors.

(Refer Slide Time: 13:25)



And just as an example, vectors can be transformed with respect to different bases.

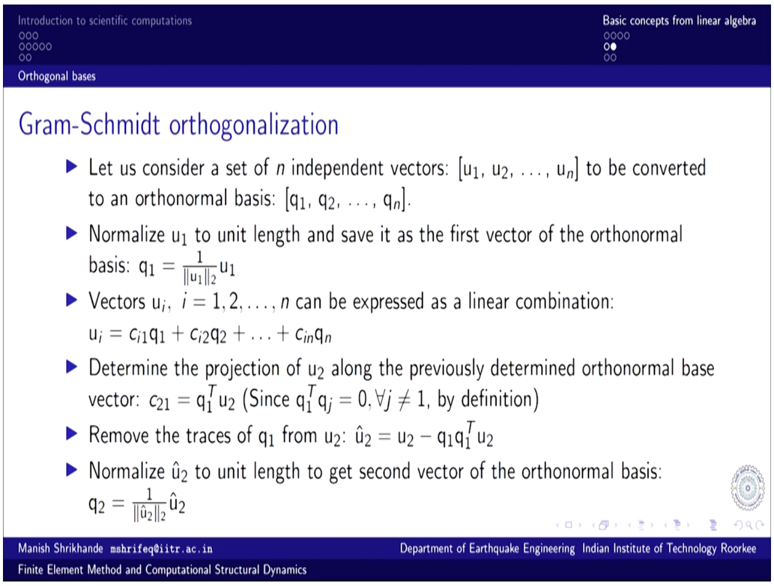
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 4 \\ 1 \\ -3 \end{pmatrix} = \begin{bmatrix} 0.862 & 0.433 & 0.263 \\ -0.267 & 0.829 & -0.491 \\ 0.431 & -0.353 & -0.830 \end{bmatrix} \begin{pmatrix} 1.88 \\ 3.62 \\ 3.05 \end{pmatrix}$$

So, both of them are identical. They are referring to the same point, but with respect to different set of base vectors. So, essentially the point representation that I have is I am referring to the same position, same point in 3-dimensional space, but it is with reference to different frames of references.

So, $[4,1,-3]^T$ $[4,1,-3]^T$ is with respect to some set of unit vectors, Euclidean base vectors, and the same point can have coordinates $[1.88, 3.62, 3.05]^T$ if I use the base vectors which are oriented along these directions as indicated. And I can have Ax. So, A is this matrix and this is the column vector, so that represents Ax. And Ax is again another column vector that is the position vector that is equal to B is this arrangement of base vectors and with associated coordinates, so, By. So, it also refers to the same position vector albeit with different basis. So, this is the change of basis. I can change coordinates, I can find coordinates with respect to one from the other and if I know the basis I can find out the coordinates with respect to new basis by just taking this linear transformation B inverse A and that would be the linear transformation. Now, inverse computation is a very tedious process and we do not really recommend. I mean inverse of a matrix should never be computed that is the thumb rule of numerical computation, and we can find out several ways of doing the operation. I mean most of the time we do not really need to find the inverse.

And interestingly for orthonormal basis inverse of a matrix, inverse of the matrix is same as its transpose. So, the problem is solved. If I choose an orthonormal basis then I do not really need to compute the inverse, I just have to find the transpose of the matrix and the problem is solved, that transpose is its inverse and subsequent computations are also facilitated very easily. So, that is why linearly independent vectors although they are very general they are not convenient because of the need to more elaborate computation for finding the coordinates. Orthogonal vectors are much more simpler and they will often lead to simpler models in computation. And establishment of an orthogonal basis is a logical process, very logical process, but it requires little bit of expensive calculations, expensive operations that is called through the Gram-Schmidt orthogonalization, right.

(Refer Slide Time: 17:09)



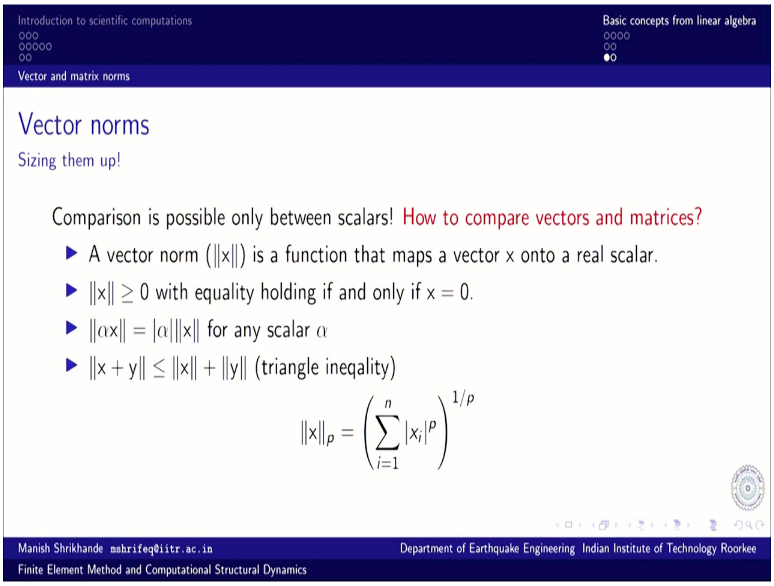
We will not get into this Gram-Schmidt orthogonalization at the moment, that is not very crucial for these subsequent discussions, but it is very important when we come to the process of numerical computations. So, while we are at the vector space discussion let us also briefly discuss this Gram-Schmidt orthogonalization that will help in understanding the operations that we do for orthogonalization and what are the benefits.

So, the point is any linearly independent victors let us say [u_1 ……… u_n] they can be converted into an orthonormal basis that is [q_1, q_2 ……. q_n]. And the first process starts with the first vector anything any arbitrary vector u_1, it can be normalized to unit length, so that is just sum of squares and square root and divided each element by that number and that is take that takes care of the orthonormal basis. And then we go by the definition of orthogonalization that the projection has to be 0, projection on the resultant vector has to be 0 on the other components. So, if I want to make any vector orthogonal with respect to let us say one direction then I just find out its projection on that direction and then subtract that component. Once I compute the projection I know exactly by what amount that direction is contained in this vector and then I sweep out that contribution. So, any vector u_i can be expressed as a linear combination. Remember we have already computed q_1 here, so theoretically u_i can be represented as, any arbitrary vector u_i can be represented as $c_{i1} q_1+c_{i2} q_2$ ……. $c_{in} q_n$.

Now, I can find out the projection along the previously determined orthogonal basis let us say q_1. So, that projection is given by $q_1^T u_2$ and that is c_{21}, and then remove the traces of q_1 from u_2 just as this operation and then this represents resultant vector is orthogonal, is rendered orthogonal to q_1. And that is how we continue the orthogonalization. We can normalize it to unit length and that is how we generate second set of the orthogonal orthonormal basis q_2.

Again, u_3 will be made orthogonal with respect to q_1 and q_2 by the similar process and then normalize it to get q_3. Then, u_4 will be made orthogonal with respect to q_1, q_2 and q_3 and then normalize subsequently normalized to unit length to get q_4, and so on it continues. So, it is a sequential process, and this process is called Gram-Schmidt orthogonalization.

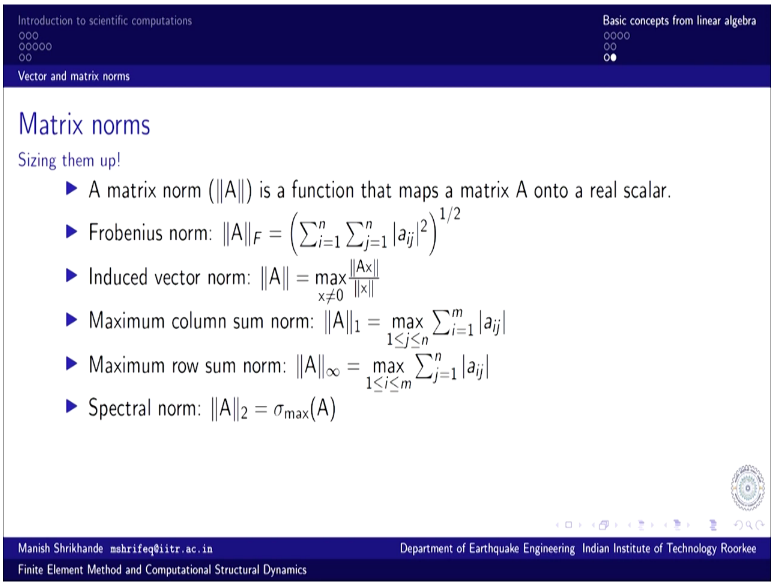
(Refer Slide Time: 20:36)



Now, another thing that we need to discuss is the how to we size it up. If you look at it this particular thing we are using, so this is what we call as l_2 norm or the length of the vector, length norm of the vector because when we are talking about approximate solutions we also have something called error. I mean the moment we are talking of approximate solution we are talking of errors. So, we need to find out what to, how to minimize the errors or when the error is small enough. Now, small or large can only be discussed between two scalars, two numbers, comparison is possible only between two scalars. If I have a vector of numbers and another vector of numbers how do I compare between them? How do I say that this vector is smaller than the other vector? So, during the process of approximation I have some error from the one set of approximation, I have another measure of another error in solution that is available as a vector from another set of approximation how do I judge which approximation is better? And that is possible if I define a vector norm is a function that maps a vector onto a real scalar. And this is a generic definition of a vector norm, any n-dimensional vector absolute value of each element raised to the power p and summed over all the elements and then taken the pth root. So, that is called pth norm. And we can generalize it norm of a vector is a real number greater than or equal to 0 with it is equal to 0 only if x is a null vector and norm of alpha times x is same as alpha times norm of x for any scalar number alpha.

And then, we have this triangle inequality. The norm of two vectors x plus y is less than or equal to the sum of individual norms.

(Refer Slide Time: 23:03)



And similarly, just as we have these vector norms we also have matrix norms that you can look it up and there are different norms, Frobenius norms, and then we can transform a matrix into a vector and that is called, then use the vector norm. So, that is called induced vector norm. And then we can always set convert it into a column vector or a row vector and then take the appropriate norm of the vectors or we can also have the spectral norm that is the singular values of the matrix, larger singular value or the eigen values of the matrix that can also be used as a measure of the matrix.

So, with this we stop our discussion. And we will continue with basic approximate solution using finite element method in our next lecture.

Thank you.