

Elementary Numerical Analysis
Prof. Rekha P. Kulkarni
Department of Mathematics
Indian Institute of Technology, Bombay

Lecture No. # 21
Vector and Matrix Norms

We have considered some methods for solving a system of linear equations. So, the methods which we have considered are choas decomposition, then gauss elimination method, with and without partial pivoting.

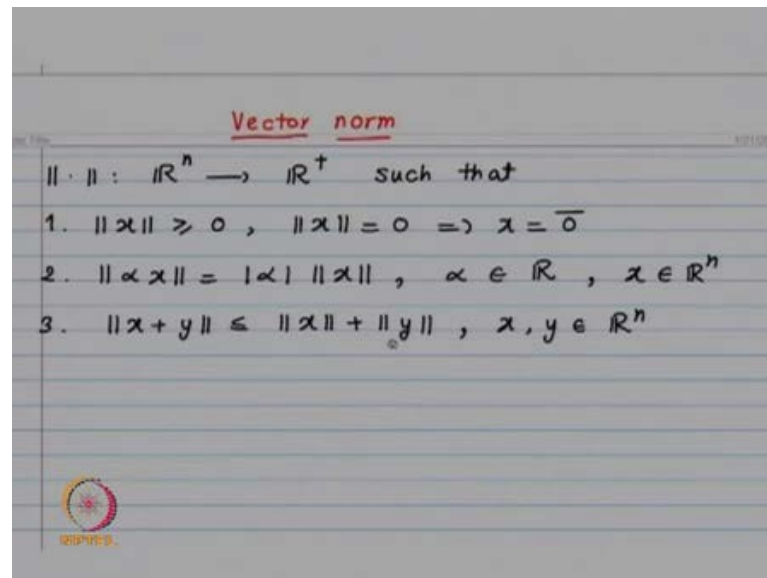
Now, these methods they are meant for solving big systems of linear equations using a computer. If you are doing hand computations, then you will be solving small system, say, n is equal to 3, n is equal to 4 and then, it does not matter which method you use, but the methods which we have considered and their relative merits, it is meant for a big system of linear equations.

So, you are necessarily going to use computer, now in computer no matter how powerful your computer is, you have finite precision; that means, after the decimal point you are going to have a fixed word length, so that is going to introduce a error. Also, the linear system which you are trying to solve, it may be coming from some experimental data. In that case, there is going to be experimental error also, we want to solve an exact equation $Ax = b$, but because of the finite precision of the computer, you are going to solve a perturb system; that means, there is going to be error in the elements of your matrices, matrix - coefficient matrix. Similarly, there will be error in the right hand side, so you get a computed solution.

Now, one wants to know, how near your computed solution is to the exact solution. So, for that we need to have a measure to decide, how near your computed solution is to the exact solution, so that is why we are going to consider norms. Now, the norm will give us a way to calculate distance between two vectors, you are familiar with Euclidean norm. So, in \mathbb{R}^3 the Euclidean length is going to be $x_1^2 + x_2^2 + x_3^2$ whole thing raise to half.

Then, we are going to define some other vector norm and then, we will be considering the matrix norm. So, today's topic is going to be vector and matrix norms, and as has been the case so far, our numbers, they are going to be all real numbers. So, matrices involved are real, then vectors involved they are real, so let us define a norm. So, x is going to be a vector in \mathbb{R}^n and we will define norm.

(Refer Slide Time: 03:46)



So, norm is a function from \mathbb{R}^n to \mathbb{R}^+ , so \mathbb{R}^+ is all non-negative real numbers, which should satisfy the property that norm x is bigger than or equal to 0. Actually, that is included, when I say that norm is from \mathbb{R}^n to \mathbb{R}^+ ; that means, norm x is bigger than or equal to 0, it should be equal to 0 should imply x is equal to 0 vector, and the converse also should be true that x is equal to 0 vector, that should imply that norm x is equal to 0.

The second condition is α is a real number; α times x ; that means, you multiply each component of x by α . So, norm of this new vector should be equal to modulus of α times norm x , this should be valid for all α belonging to \mathbb{R} and for all vectors x belonging to \mathbb{R}^n . And the third condition is known as triangle inequality that norm of x plus y is less than or equal to norm x plus norm y , two vectors in \mathbb{R}^n you add component y , so if you have a vector $x = x_1, x_2, x_n$; y is vector y_1, y_2, y_n . When you want to add x plus y , you add the corresponding components. So, it will be x_1 plus y_1 ,

x_2 plus y_2 , x_n plus y_n , and scalar multiplication is α times x will be vector with the components αx_1 , αx_2 , αx_n .

Now, let us look at some of the examples of norms, you can define norm in various manners, we are going to consider three norms. So, one is the Euclidean norm **and** then second is known as a one norm, and the third norm is the maximum norm or infinity norm. And the reason we consider 1-norm and infinity norm will become clear when we look at corresponding matrix norm.

(Refer Slide Time: 06:11)

Examples: $x = [x_1, x_2, \dots, x_n]^T$

$$\|x\|_2 = \left(\sum_{j=1}^n x_j^2 \right)^{1/2} : \text{Euclidean norm}$$

$$\|x\|_1 = \sum_{j=1}^n |x_j| : \text{1-norm}$$

$$\|x\|_\infty = \max_{1 \leq j \leq n} |x_j| : \text{\infty-norm}$$

So, here Euclidean norm is summation j goes from 1 to n x_j square whole thing raise to half, then norm x_1 is equal to summation j goes from 1 to n modulus of x_j and norm x_∞ is maximum mod x_j 1 less than or equal to j less than or equal to n . Now, it is easy to verify that all these three definitions, they will satisfy the three properties of the norm, so that norm x should be bigger than or equal to 0, so look at Euclidean norm. You are considering summation x_j square and then, you are taking positive root - positive square root, so it is going to be bigger than or equal to 0. If it is equal to 0 then, your sum of the squares is 0, so **it has to be** the each component has to be 0; that means, x is a 0 vector and if x is a 0 vector norm x will be equal to 0.

Similarly, when you consider 1-norm and infinity-norm, it is going to satisfy that it is bigger than or equal to 0 and is equal to 0, if and only if, x is equal to 0 vector, it is simple to verify.

Next is, norm alpha x should be equal to mod alpha times norm x, so this also for all the three definitions you substitute, look at the formula and then you can verify. The third property is the triangle inequality. So, this triangle inequality, it is easy to verify for 1-norm and infinity-norm, its straight forward. Like, look at the 1-norm, so norm of x plus y, it will be summation j goes from 1 to n modulus of x j plus y j, modulus of x j plus y j, x j and y j these are real numbers, so that will be less than or equal to mod x j plus mod y j and then separate the summation.

Similarly, for the infinity-norm it is straight forward, the triangle inequality for the 2-norm one needs to use what is known as Cauchy Schwarz inequality. So, I am not going to prove Cauchy Schwarz inequality, but I will state Cauchy Schwarz inequality and show you how using Cauchy Schwarz inequality, we get the triangle inequality for the 2-norm.

(Refer Slide Time: 09:06)

Inner Product on \mathbb{R}^n

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j$$

$$\langle x, x \rangle^{1/2} = \left(\sum_{j=1}^n x_j^2 \right)^{1/2} = \|x\|_2$$

Cauchy-Schwarz Inequality:

$$|\langle x, y \rangle| \leq \|x\|_2 \|y\|_2$$

$$\left| \sum_{j=1}^n x_j y_j \right| \leq \left(\sum_{j=1}^n x_j^2 \right)^{1/2} \left(\sum_{j=1}^n y_j^2 \right)^{1/2}$$

So, now this inner product on \mathbb{R}^n , so its inner product of x comma y, x and y are vectors in \mathbb{R}^n , it is defined as summation j goes from 1 to n x j y j. We will come across inner product little later and then, that time we will study the properties of the inner product. In fact, we did use inner product, when? We had talked about the gauss points. The gauss points they were 0's of the Legendre polynomials, that was need they were needed in gaussian quadrature and that is where we had talked about the inner product, so, you are

familiar with inner product. So now, look at the inner product on \mathbb{R}^n and then, you consider positive square root of $\langle x, x \rangle$, so that is nothing, but our 2-norm of x .

(Refer Slide Time: 10:08)

Inner Product on \mathbb{R}^n

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j$$
$$\langle x, x \rangle^{1/2} = \left(\sum_{j=1}^n x_j^2 \right)^{1/2} = \|x\|_2$$

Cauchy-Schwarz Inequality:

$$|\langle x, y \rangle| \leq \|x\|_2 \|y\|_2$$
$$\left| \sum_{j=1}^n x_j y_j \right| \leq \left(\sum_{j=1}^n x_j^2 \right)^{1/2} \left(\sum_{j=1}^n y_j^2 \right)^{1/2}$$

So, $\langle x, x \rangle$ raise to half is equal to norm x_2 , so the Cauchy Schwarz inequality is modulus of inner product of x with y is less than or equal to norm x_2 into norm y_2 . So, that is modulus of summation j goes from 1 to n $x_j y_j$, because that is our definition of inner product of x with y , is less than or equal to summation j goes from 1 to n x_j^2 raise to half which is norm x_2 and this quantity is norm y_2 , so this is the Cauchy Schwarz inequality.

(Refer Slide Time: 10:50)

The image shows a handwritten derivation on lined paper. At the top, it is titled "Cauchy-Schwarz Inequality". The first line is the inequality: $|\sum_{j=1}^n x_j y_j| \leq (\sum_{j=1}^n x_j^2)^{1/2} (\sum_{j=1}^n y_j^2)^{1/2}$. The second line shows the expansion of the squared 2-norm of the sum of two vectors: $(\|x+y\|_2)^2 = \sum_{j=1}^n (x_j + y_j)^2$. The third line expands this sum: $= \sum_{j=1}^n x_j^2 + \sum_{j=1}^n y_j^2 + 2 \sum_{j=1}^n x_j y_j$. The fourth line applies the Cauchy-Schwarz inequality to the cross term: $\leq (\|x\|_2)^2 + (\|y\|_2)^2 + 2 \|x\|_2 \|y\|_2 = (\|x\|_2 + \|y\|_2)^2$. A small logo is visible in the bottom left corner of the slide.

And now, look at norm of x plus y , it is 2 norm and then square of it, so that is summation j goes from 1 to n x_j plus y_j square; x_j plus y_j is going to be j th component of x plus y . So, I expand a square and I will get summation over j x_j square plus summation over j y_j square plus two times summation $x_j y_j$. This is nothing but 2 norm of x square summation over j y_j square is, norm y its 2 norm square plus, now I use the Cauchy Schwarz inequality to dominate this by two times norm x 2 norm y 2, so this is nothing but norm x plus norm y square. So, now you take positive square root and then that will prove the triangle inequality for 2 norm, so this is about the vector norms.

Now, I want to consider matrix norm, so I have got a to be n by n matrix, so I want to define the norm. So, it should satisfy our three properties, the positive definiteness, triangle inequality, and the third one was that, α times a should be equal to mod α times norm a . So, why not treat our a as a vector of length n square because, what is matrix? It is an arrangement of n square number, we write them as rows and then columns. So, as such the matrix A , it has got n square elements, so I can treat it as a **n square** a vector of length n square, I know how to define vector norm.

(Refer Slide Time: 13:05)

The image shows handwritten mathematical definitions for two matrix norms on lined paper. The first definition is for the Frobenius norm: $A = [a_{ij}]$ is an $n \times n$ matrix, and $\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2}$, labeled as the Frobenius norm. The second definition is for the maximum norm: $\|A\|_{\max} = \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} |a_{ij}|$. A small logo is visible in the bottom left corner of the paper.

So, then I will consider the corresponding norm, norm which corresponds to the Euclidean norm will be a $i j$ square. Now, instead of single summation you will have summation i goes from 1 to n , summation j goes from 1 to n and then, whole thing raise to half, so this is known as Frobenius norm.

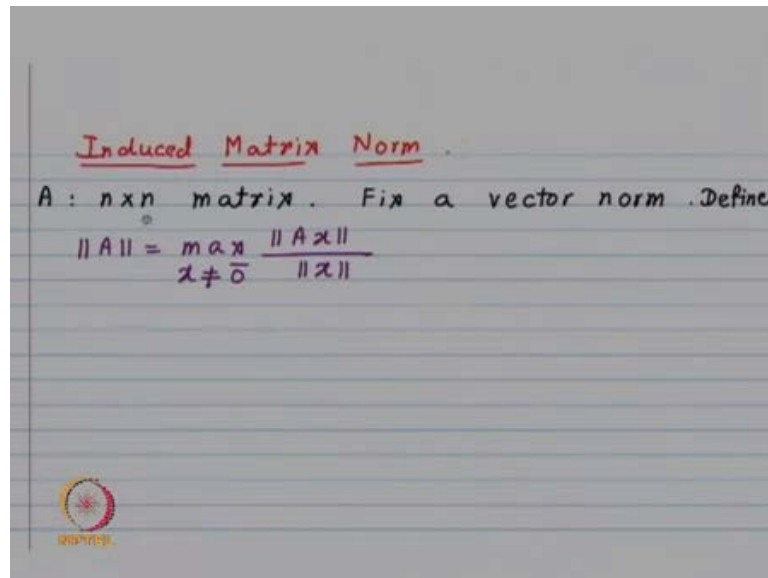
And then, norm A_{\max} , now I am not writing infinity, because norm A_{∞} I want to reserve that symbol for something else, so norm A_{\max} will be maximum of modulus of a_{ij} , $1 \leq i \leq n$, $1 \leq j \leq n$. And then, one can verify that these definitions, they will satisfy our three properties of norms. Now, what we want to do is, we want to relate our vector norm and the matrix norm, because look at our system of linear equation, it is $Ax = b$, so we are considering matrix into vector. So, I will like to have some relation between A - the matrix norm and a vector norm and another thing is I can multiply two matrices A and b , if they are I am considering square matrices, you cannot multiply two vectors, but you can multiply two matrices.

So, then again norm of A b and norm A norm b , I will like to have some relation between them. So, in order to do that what we are going to do is, we are going to consider induced matrix norm.

So, **we will have** we are going to define a general way of defining a matrix norm. So, we are going to start with a vector norm and from that we will give a general definition of

induced matrix norm. And then, since we are interested in our 1-norm, 2-norm, infinity-norm, we will see what are going to be the corresponding induced matrix norm, so now, first definition of Induced matrix norm.

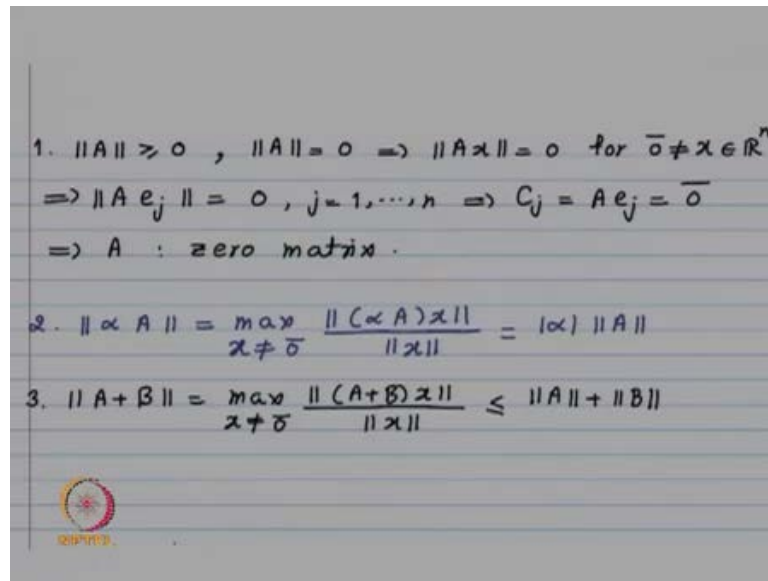
(Refer Slide Time: 15:35)



So, A is n by n matrix, you fix a vector norm, any norm which you want; that means, it should satisfy the three properties of norm. And now, I am going to look at maximum of norm Ax by norm x, x not equal to 0 vector that is my definition of norm of A. So, now, the first thing we want to show is that if I define norm A in this manner, then it satisfies those three properties of norm.

So, the first property is positive definiteness that norm A is bigger than or equal to 0, and it is equal to 0, if and only if, A is a now 0 matrix. The second property will be, norm alpha A should be equal to mod alpha times norm A, and the third property is the triangle inequality norm of A plus B is less than or equal to norm A plus norm B. So, three properties we are going to deduce using the fact that our vector norm satisfies these properties, you will see that the proofs are straight forward.


(Refer Slide Time: 16:54)



1. $\|A\| \geq 0$, $\|A\| = 0 \Rightarrow \|Ax\| = 0$ for $\bar{0} \neq x \in \mathbb{R}^n$
 $\Rightarrow \|Ae_j\| = 0, j=1, \dots, n \Rightarrow C_j = Ae_j = \bar{0}$
 $\Rightarrow A$: zero matrix.

2. $\|\alpha A\| = \max_{x \neq \bar{0}} \frac{\|(\alpha A)x\|}{\|x\|} = |\alpha| \|A\|$

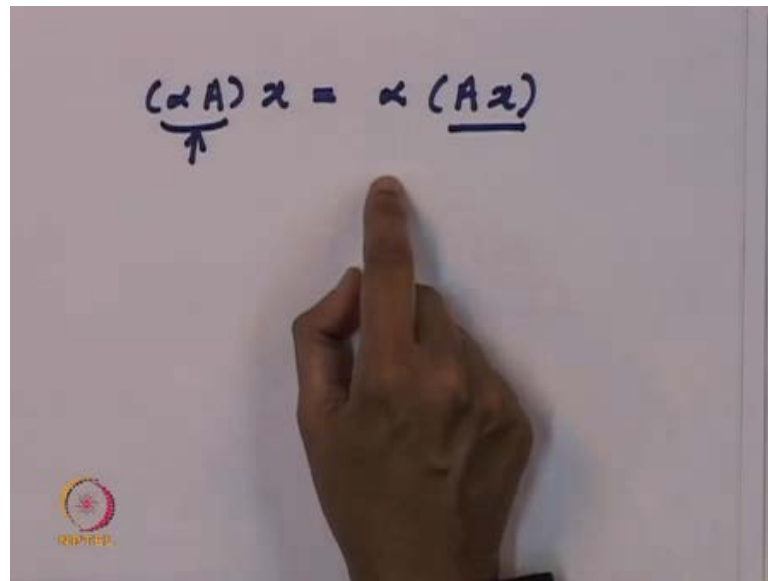
3. $\|A+B\| = \max_{x \neq \bar{0}} \frac{\|(A+B)x\|}{\|x\|} \leq \|A\| + \|B\|$



So, norm A , A is bigger than or equal to 0 that is clear, norm A is equal to 0, it will imply that norm Ax is equal to 0. Remember, our norm A is maximum - I have written here - so norm of A will be maximum of norm Ax divided by norm x , x not equal to 0 vector. So, if norm A is 0 then norm Ax is going to be equal to 0 for any non-zero vector in \mathbb{R}^n . So, now, look at our canonical vectors, e_j is going to be vector with 1 at j th place and 0 elsewhere.

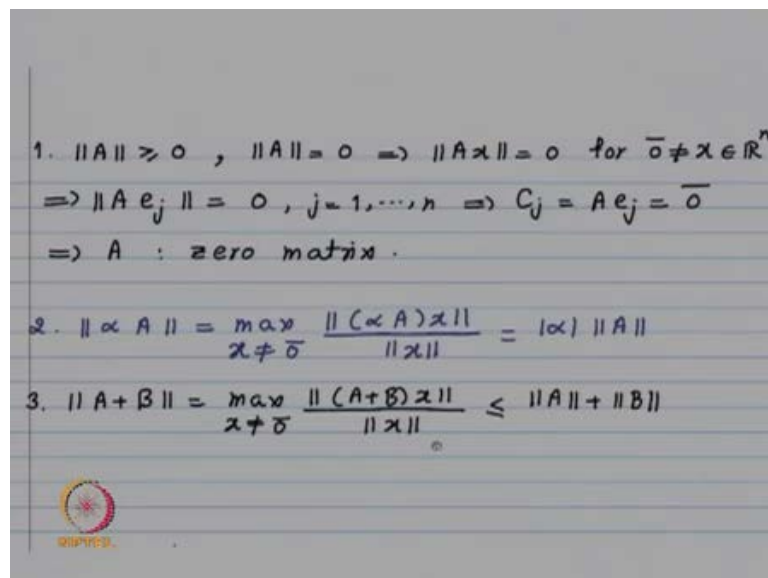
So, I will have norm Ae_j to be equal to 0, but what is our Ae_j ? Ae_j is nothing but the j th column. Now, using the property of vector norm, you will get C_j to be a 0 vector and you get A to be a 0 matrix. So, we are using the property of vector norm, the second property is, look at norm αA , this will be by definition maximum x not equal to 0 vector norm of αAx by norm x , but then our α times A into x will be nothing but α times Ax .

(Refer Slide Time: 18:13)



What is the difference? Here, you are multiplying matrix A by alpha, here you are first calculating A into x and then, multiplying by scalar alpha.

(Refer Slide Time: 18:35)



You have norm alpha A is maximum x not equal to 0 vector alpha times A x divided by norm x, and using the property of vector norm you get mod alpha times norm A because, mod alpha comes out it does not depend on maximum, so it will come out of the maximum, and maximum of norm A x by norm x, x not equal to 0 vector is norm A. And similarly, triangle inequality norm of A plus B is maximum x not equal to 0, norm of A

plus Bx divided by norm x use the fact that norm of $Ax + Bx$ is less than or equal to norm Ax plus norm Bx .

(Refer Slide Time: 19:14)

$$\begin{aligned} & \frac{(\alpha A)x}{\uparrow} = \alpha (Ax) \\ & \|(A+B)x\| \leq \|Ax\| + \|Bx\| \\ & \max_{x \neq 0} \frac{\|(A+B)x\|}{\|x\|} \leq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \\ & \quad + \max_{x \neq 0} \frac{\|Bx\|}{\|x\|} \\ & \|A+B\| \leq \|A\| + \|B\| \end{aligned}$$

So, maximum of norm of $Ax + Bx$, x not equal to 0 vector divided by norm x will be less than or equal to maximum of norm Ax by norm x , x not equal to 0 vector, plus maximum of norm Bx , x not equal to 0 vector, divided by norm x . So, this proves that norm of $A + B$ is less than or equal to norm A plus norm B .

We defined induced matrix norm and then, we showed that it satisfies all the properties of our vector norm. Now, we are going to prove one of the basic inequality which we will be using frequently later on. So, that inequality is going to relate the matrix norm to vector norm and we are going to show that norm of AB is less than or equal to norm A into norm B . Now, these two properties they will follow from our definition of induced matrix norm.

(Refer Slide Time: 21:01)

The image shows handwritten mathematical work on lined paper. At the top, the definition of the induced matrix norm is given: $\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$. Below this, two implications are shown. The first is $\frac{\|Ax\|}{\|x\|} \leq \|A\|$ for all $x \neq 0 \in \mathbb{R}^n$. The second is the boxed fundamental inequality: $\|Ax\| \leq \|A\| \|x\|$ for all $x \in \mathbb{R}^n$. A small logo is visible in the bottom left corner of the paper.

So, our norm A is maximum norm Ax by norm x , x not equal to 0 vector; that means, for all non-zero vector norm Ax by norm x is less than or equal to norm A , which will give us norm Ax to be less than or equal to norm A into norm x for all non-zero vector, but if x is a 0 vector, norm x will be 0 Ax being a 0 vector again this will be 0, so this fundamental inequality will be true for all x belonging to \mathbb{R}^n .

Now, I want you to note one thing that here norm Ax - that means vector norm - norm A means, induced matrix norm and norm x ; that means, again vector norm. So, it is clear from the context about which norm we are talking about, A being a matrix this will be a matrix norm; x being a vector, this will be a vector norm.

So, this is the first inequality and now, let us show the consistency condition; that means, if you have got A and B to be two matrices take their product, so AB is going to be again a n by n matrix, so norm AB is less than or equal to norm A into norm B . So, this fact will be proved using our basic inequality that norm Ax is less than or equal to norm A into norm x .

(Refer Slide Time: 22:54)

Consider $\|ABx\| \leq \|A\| \|Bx\| \leq \|A\| \|B\| \|x\|$
 $\Rightarrow \frac{\|ABx\|}{\|x\|} \leq \|A\| \|B\|$ for $0 \neq x \in \mathbb{R}^n$
 $\Rightarrow \|AB\| \leq \|A\| \|B\|$: Consistency condition

So, you have got norm ABx to be less than or equal to, by fundamental inequality, it will be norm A into norm Bx , treat Bx as one vector and A as our matrix. So, you have got norm A into norm Bx , this is less than or equal to... Now, use fundamental inequality for Bx , so norm Bx will be less than or equal to norm B into norm A . Now, if x is not equal to 0 vector, norm x will not be 0 , so I can divide and you will get norm ABx upon norm x to be less than or equal to norm A into norm B for all non-zero vectors in \mathbb{R}^n , so now take maximum of this over x not equal to 0 .

The right hand side is independent of x , so you get norm AB to be less than or equal to norm A into norm B , so this is consistency condition. So, when one talks about vector norm, it should satisfy those three properties. When one talks about matrix norm, it should satisfy the three properties of vector norm and in addition norm AB to be less than or equal to norm A into norm B . This is a very elegant definition, start with any vector norm you define the induced matrix norm showing that it satisfies the three properties of vector norm and then the consistency condition it was straight forward, so very nice situation.

But then, given a matrix A , I will like to calculate its norm, like given a vector x , if I want to calculate its Euclidean norm - I know- that take the squares of their components, add them up, take the positive square root. If I want 1-norm then I will take the modulus of each component add it up, if I want infinity-norm look at the moduli of its components

and look at the maximum. So now, suppose, I take one of these three norms and then, I want to know what the induced matrix norm? Now, in the definition of induced matrix norm, you have got maximum over all non-zero vector of certain quotient, that quotient is norm $A \times$ by norm A .

So, how am I going to calculate this maximum, I have got infinitely many vectors, so then I cannot calculate this induced matrix norm. So, now, we are going to derive a formula for calculating norm A_1 ; that means, when you fix vector norm to be 1, the corresponding induced matrix norm. So, we will have a formula in terms of the components of the matrix, so it is going to be something one can calculate.

Similarly, for infinity-norm of the matrix, we will have a formula whereas, the nice case of Euclidean norm, well, one has to be satisfied only with an upper bound. So, the Euclidean norm of matrix is something which you cannot calculate, so now, let us derive a formula for norm of A_1 . We are fixing our vector norm to be 1-norm, and then, I am going to look at norm $A \times 1$ divided by norm x_1 and look at their maximum.

(Refer Slide Time: 27:16)

$$\begin{aligned}
 &= \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \\
 \|Ax\|_1 &\leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| |x_j| \\
 &= \sum_{j=1}^n \sum_{i=1}^n |a_{ij}| |x_j| \\
 \|Ax\|_1 &= \sum_{j=1}^n |x_j| \sum_{i=1}^n |a_{ij}| \\
 &\leq \left(\max_j \sum_{i=1}^n |a_{ij}| \right) \left(\sum_{j=1}^n |x_j| \right) = \|x\|_1
 \end{aligned}$$

So, I have got norm A_1 is maximum of norm of $A \times 1$ divided by norm x_1 , x not equal to 0 vector, I want to find a formula for this. Now, the components of $A \times$, let me write as $A \times i$, this will be summation $a_{ij} x_j$ - j going from 1 to n that is, the matrix into vector multiplication. Norm of $A \times 1$, this will be summation i goes from 1 to n modulus of $A \times i$

that is our definition of 1-norm. Now, let me substitute for Ax , so it is going to be summation i goes from 1 to n modulus summation j goes from 1 to n $A_{ij} x_j$.

Now, this will be less than or equal to summation i goes from 1 to n , summation j goes from 1 to n modulus of a_{ij} modulus of x_j , so we are using triangle inequality. Now, this is finite summation, so I can interchange the order of the summation, **so, I am going to have**. So, this is our norm $\|Ax\|_1$ is less than or equal to this, so this is same as summation j goes from 1 to n , summation i goes from 1 to n modulus of A_{ij} and then, mod x_j .

Now, summation is over i , here you have x_j , so it will come out of the summation sign. So, it will be summation j goes from 1 to n mod x_j summation i goes from 1 to n modulus of A_{ij} . So here, what we are doing is, you fix j ; that means you are fixing the column, look at the entries in that column take their modulus and then add it up. So, I can say that this is less than or equal to maximum over j of the quantity, summation i goes from 1 to n modulus of A_{ij} into summation j goes from 1 to n mod x_j and this is nothing but norm x_1 .

Let me call this quantity to be alpha, so what we have is norm Ax_1 is less than or equal to alpha times norm x_1 . And hence, you will have norm A_1 to be less than or equal to this number.


(Refer Slide Time: 30:59)

Formula for calculating $\|A\|_1$

$$(Ax)(i) = \sum_{j=1}^n a_{ij} x_j \quad \|Ax\|_1 = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right|$$

$$\|Ax\|_1 \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| |x_j| = \sum_{j=1}^n |x_j| \left(\sum_{i=1}^n |a_{ij}| \right)$$

$$\leq \left(\max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \right) \sum_{j=1}^n |x_j|$$

$$= \alpha \|x\|_1$$


So, here, we have $\|Ax\|_1 \leq \alpha \|x\|_1$, where α is this number, summation i goes from 1 to n modulus of a_{ij} less than or equal to j less than or equal to n . So, I can very well calculate α , I will look at the columns of my matrix A , so look at the first column then take the modulus and add it up. So, that means I am considering 1-norm of the first column, then 1-norm of the second column and 1-norm of the n th column. So, I will get n numbers, the maximum among them that is going to be my α .

And my $\|Ax\|_1$ is going to be less than or equal to this α , but we will like to show that it is not only less than or equal to, but it is equal to. So, we have proved that $\|Ax\|_1$ is less than or equal to this number, and now, let us show that, in fact, it is equal to maximum.

(Refer Slide Time: 32:15)

$$\|Ax\|_1 \leq \alpha \|x\|_1, \text{ where}$$

$$\alpha = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

 For $x \neq \bar{0}$, $\frac{\|Ax\|_1}{\|x\|_1} \leq \alpha$

$$\Rightarrow \|A\|_1 = \max_{x \neq \bar{0}} \frac{\|Ax\|_1}{\|x\|_1} \leq \alpha.$$

So, you have $\|Ax\|_1 \leq \alpha \|x\|_1$, where α is this quantity and hence you have got $\|A\|_1$ to be less than or equal to α .

(Refer Slide Time: 32:28)

$$\|A\|_1 \leq \alpha = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| = \sum_{i=1}^n |a_{ij_0}|$$

for some j_0


Consider $A e_{j_0} = \begin{bmatrix} a_{1j_0} \\ a_{2j_0} \\ \vdots \\ a_{nj_0} \end{bmatrix}$

$\|A e_{j_0}\|_1 = \alpha$,

$\|e_{j_0}\|_1 = 1$.

$\|A\|_1 = \alpha$

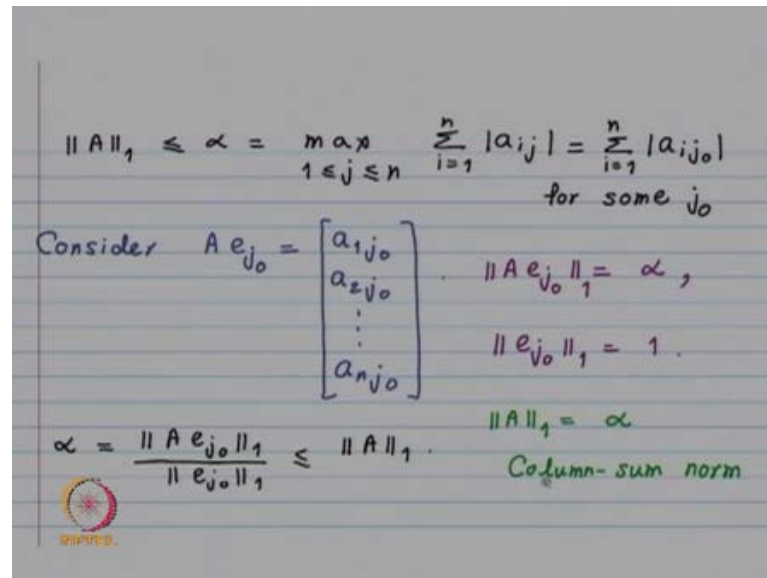
Column-sum norm

$$\alpha = \frac{\|A e_{j_0}\|_1}{\|e_{j_0}\|_1} \leq \|A\|_1$$


Now, look at this maximum, this will be equal to summation i goes from 1 to n modulus of a_{ij_0} , for some j . Such a j_0 need not be unique, it depends on your matrix, may be all the columns they are going to have the same 1-norm, but you are looking at maximum of n numbers, so look at least one number which is equal to maximum and let that column be equal to j_0 th column.

Now, if you consider A times e_{j_0} , e_{j_0} is a canonical vector with 1 at j_0 th place and 0 elsewhere. So, $A e_{j_0}$ gives you j_0 th column, so using this fact we will show that α is equal to $\|A\|_1$.

(Refer Slide Time: 33:34)



Handwritten mathematical derivation on lined paper:

$$\|A\|_1 \leq \alpha = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| = \sum_{i=1}^n |a_{ij_0}|$$

for some j_0

Consider $Ae_{j_0} = \begin{bmatrix} a_{1j_0} \\ a_{2j_0} \\ \vdots \\ a_{nj_0} \end{bmatrix}$

$\|Ae_{j_0}\|_1 = \alpha$,

$\|e_{j_0}\|_1 = 1$.

$\|A\|_1 = \alpha$

Column-sum norm

$$\alpha = \frac{\|Ae_{j_0}\|_1}{\|e_{j_0}\|_1} \leq \|A\|_1$$

A small logo is visible in the bottom left corner of the slide.

So, consider Ae_{j_0} that is the thing, but the j_0 th column, so $\|Ae_{j_0}\|_1$ is equal to α and if you consider $\|e_{j_0}\|_1$ that is going to be equal to 1. So, you have got α which is equal to $\|Ae_{j_0}\|_1$ divided by $\|e_{j_0}\|_1$, this is going to be less than or equal to $\|A\|_1$. So, thus we have proved that $\|A\|_1$ is equal to α and that is known as column sum norm, because what you do is, in each column you are adding the moduli of the entries and looking at their maxima, so here is a formula for $\|A\|_1$.

Now, we are going to look at $\|A\|_\infty$ - now $\|A\|_\infty$ we will see that. So, here, what was the crucial result? It was that you had finite summation, so you interchange the order of the summation. Now, in case of the infinity-norm, you do not have to do interchange, because there is going to be only one summation and one maximum. Here, in order to show that as such $\|A\|_\infty$, it is going to be row sum norm. That means, what you do is, look at each row and take the module of the entry and add it up, you will get n numbers, among this n numbers whichever is going to be the bigger one that is going to be $\|A\|_\infty$, so this fact now we are going to prove.

Now, in order to show that $\|A\|_\infty$ is equal to this showing less than or equal to will be simple, showing that it is equal to, that is, going to be a bit involved in case of 1-norm, it was easy you just looked at the corresponding canonical vector. Here, we will

have to construct a vector where, the infinity-norm is going to be at A, but then the proof is going to be something similar.

So, you fix infinity-norm and then, our aim is to calculate norm A infinity which is maximum of norm A x infinity divided by norm x infinity, x not equal to 0 vector, so let us start with norm A x infinity.

(Refer Slide Time: 36:46)

∞-norm of a matrix

$$\begin{aligned} \|Ax\|_{\infty} &= \max_{1 \leq i \leq n} |(Ax)(i)| \\ &= \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \\ &\leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| |x_j| \\ &\leq \|x\|_{\infty} \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad \|A\|_{\infty} \leq \beta \\ &\leq \beta \|x\|_{\infty} \end{aligned}$$

So, norm A x infinity will be maximum of modulus of A xi, 1 less than or equal to i less than or equal n, by definition of infinity-norm. Now, A xi as before will be given by summation j goes from 1 to n a ij x j then, for this modulus use triangle inequality. So, you will get less than or equal to, maximum one less than or equal to i less than or equal to n summation j goes from 1 to n modulus of a ij mod x j.

So, there are no two summations here now this mod x ji will dominate by norm x infinity. So, norm x infinity is going to be maximum of mod x jj going from 1 to n. So, if I dominate that it will come out of this summation and then maximum. So, you have norm x infinity maximum summation j goes from 1 to n mod a ij 1 less than or equal to i less than or equal to n. So, here you are fixing first i is equal to 1 then, you are looking at you are varying j; that means, you are looking at the elements in the first row taking their modulus summing it up.

Then put i is equal to 2, so you do it for the second row and put i is equal to n . So, like that you are going to get n numbers maximum among them that we are denoting by β . So, we get $\|Ax\|_\infty \leq \beta \|x\|_\infty$ to be less than or equal to β times $\|x\|_\infty$, from this conclude that $\|Ax\|_\infty$ divided by $\|x\|_\infty$ is less than or equal to β for x not equal to 0 vector take their maximum that is $\|A\|_\infty$, so you have got $\|A\|_\infty$ to be less than or equal to β .

And now, we want to show the other way inequality, that now we want to show that we have proved that $\|A\|_\infty$ is less than or equal to β . Now, let us show that β is less than or equal to $\|A\|_\infty$, so that combining these two inequalities we can show that $\|A\|_\infty$ is in fact equal to β .

So, now again β is maximum summation j goes from 1 to n modulus of a_{ij} maximum over i . So, this maximum will be attended for some i is equal to i_0 . Such a i_0 can be more than 1, so take one of them. So, get hold of i_0 such that β is equal to summation j goes from 1 to n modulus of a_{i_0j} , and now let us construct a vector.

(Refer Slide Time: 40:06)

The slide contains the following handwritten mathematical content:

$$\|Ax\|_\infty \leq \beta \|x\|_\infty \quad \beta = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{i_0j}|$$

$$\text{Define } y_j = \begin{cases} \frac{|a_{i_0j}|}{a_{i_0j}}, & a_{i_0j} \neq 0 \\ 0, & a_{i_0j} = 0 \end{cases} \quad \|y\|_\infty = 1$$

$$(Ay)_{i_0} = \sum_{j=1}^n a_{i_0j} y_j = \sum_{j=1}^n |a_{i_0j}| = \beta$$

$$\beta = |(Ay)_{i_0}| \leq \|Ay\|_\infty \leq \|A\|_\infty \|y\|_\infty = \|A\|_\infty$$

So, we are looking at elements of the i_0 'th row, if a particular element is non-zero then you define y_j to be $|a_{i_0j}|$ divided by a_{i_0j} . If that number is equal to 0 then, you define it to be equal to 0. Now, if I define such a matrix modulus of y_j for any j is going to be either 1 or 0. And that will mean that $\|y\|_\infty$ should be equal to 1 at least 1

of the y_j should be not 0, because if y_j is equal to 0; that means, each entry in the i 'th row will be 0.

So, you will get beta to be 0 and that will be in that a is a 0 matrix. So, now, I have constructed a vector y_j , now I look at $A y$ of i 0, so i 'th component of my vector $A y$ that will be given by summation j goes from 1 to n $a_{ij} y_j$. Now, look at the where we have constructed y_j , if corresponding y_j is not 0 then, a_{ij} multiplied by y_j will be mod a_{ij} divided by a_{ij} , so a_{ij} will get canceled and you will get modulus of a_{ij} .

If the y_j is equal to 0, it will not contribute to your summation, so you will get this is equal to summation j goes from 1 to n modulus of a_{ij} and this is going to be equal to beta. So, we have got beta is equal to i 'th component of our vector $a y$, i 'th component will be less than or equal to norm of $A y$ infinity-norm; norm $A y$ infinity, we know that it is less than or equal to norm A infinity into norm y infinity and then, you get our norm $A y$ infinity will be less than or equal to norm A infinity into norm y infinity - norm y infinity will be 1, so you will get norm A infinity.

So, we have got beta to be less than or equal to norm A infinity and what was beta? It was this maximum, so this is going to be row sum norm, so now we have obtained a formula for norm A_1 and norm A infinity. Now, you know why I called norm A max for the term maximum over i n j of modulus of a_{ij} , because I want to reserve norm A infinity to the definition of induced matrix norm when the vector norm is infinity-norm.

So, we have got a formula for norm A_1 , we have got a formula for norm A infinity, but unfortunately for norm A_2 we have got only an upper bound. So, let us now calculate that upper bound, that upper bound we will be again using Cauchy Schwarz inequality. So, again the method is the same, we will look at norm of $A x_2$ norm obtain it to be less than or equal to say a gamma times norm x_2 , so that gamma will give us an upper bound.


(Refer Slide Time: 44:22)

2-norm of the matrix (upper bound)

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}, \quad \|x\|_2 = \left(\sum_{j=1}^n x_j^2\right)^{1/2}$$

$$\|Ax\|_2^2 = \sum_{i=1}^n (Ax)_i^2 = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j\right)^2$$

$$\leq \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij}^2\right) \left(\sum_{j=1}^n x_j^2\right) \quad (\text{using Cauchy-Schwarz inequality})$$

$$\leq \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2\right) \|x\|_2^2 = \|A\|_F^2 \|x\|_2^2$$


So, we have norm A_2 to be maximum of norm Ax by norm x , x not equal to 0 vector, recall that Euclidean norm 2 of x is summation j goes from 1 to n , x_j square or raise to half.

So, look at norm Ax , it is 2 norm square, it will be summation i goes from 1 to n $a_{ij} x_j$ square - by definition. Now i th component of Ax is given by summation j goes from 1 to n $a_{ij} x_j$, so this is Ax_i and its square. Now, apply Cauchy Schwarz inequality here, so square of this will be less than or equal to summation j goes from 1 to n a_{ij}^2 and summation x_j^2 .

So, using the Cauchy Schwarz inequality you get this quantity to be less than or equal to summation i goes from 1 to n , summation j goes from 1 to n a_{ij}^2 , summation j goes from 1 to n x_j^2 . This is nothing but Euclidean norm of x , its square and then you have got this number. Now, if you recall this is what is known as Frobenius norm, so you have got norm A_F square into norm x square, so that gives you 2 norm of A to be less than or equal to norm Ax . Now, let us recapitulate our norm, so we have got our norm A_1 that is the column sum norm, norm A_∞ that is the row sum norm, and norm A_2 is less than or equal to norm a Frobenius.

So, this is the best one could do for the 2 norm, but then for the 2 norm, we have got the basic inequality norm of Ax , its 2 norm it is going to be less than or equal to norm A Frobenius into norm of x 2 norm, so that inequality is available. If your matrix A is

symmetric, then 1-norm and infinity-norm they are going to be the same because for column sum norm that gives us 1-norm, if you do corresponding thing for row you get infinity-norm.

So, for the symmetric matrix both 1-norm and infinity-norm they are going to be the same, for a general matrix they can be different. So, now, we have proved or we have defined vector norm and matrix norm, using these we are going to analyze the behavior of the perturbed system $Ax = b$ is the original system or the system which we want to solve, because of the use of computers we are going to solve system $A + \delta A$ x_{cap} is equal to $B + \delta B$. So, x is the exact solution x_{cap} is the computed solution and we will like to say something about norm of $x - x_{cap}$.

So, this perturbation of the linear system and sensitivity of the computed solution to the change in the right hand side or to the perturbation that is going to be topic of our next lecture, so thank you.