

**Scientific Computing Using Python**  
**Professor. Vivek Aggarwal and Professor. Mani Mehra**  
**Department of Mathematics**  
**Indian Institute of Technology, Delhi**  
**Lecture No. 12**

Welcome to Scientific Computing Using Python. Today, we are going to start a new topic, so we'll call it norms. In that, we will define what are vector norms and what are matrix norms. So, let's get started.

Today's topic is vector and matrix norms. Basically, first we will learn what are norms. So, norms are called how to measure the size and length of a vector and a matrix. Because, just like we have a vector, there can be a matrix as well. Vectors can also have a polynomial. So, in that case, we want to see their size or length. Like, when we used to take vectors at school level, we used to measure their size and length. So, can we measure the size of a matrix as well? Or the distance between two matrices — can we measure the distance between them? So, what will we say there? That this is the size. This is the size of this matrix. This is the size of this vector. So, for that, we will use a word called norms. So, norms basically — you can say that it is an extension of the magnitude that we used to take earlier of the vector. So, norms measure the size and the length of the vectors or matrices. So, where it is used, wherever we are dealing with vectors or matrices. Then all these things will be used there, like numerical energy, optimization, and machine learning.

So, definition — if we write what is the definition of norm, then norm is a function that assigns a non-negative length and size to vectors in the vector space. So, if we want to write it mathematically, then mathematically, you will see that if I want to write this, then we have defined it here that it is a mathematical definition.

(Refer slide time: 2:48)

- ▶ Norms measure the size or length of a vector or matrix.
- ▶ Commonly used in numerical analysis, optimization, and machine learning.

**A norm** is a function that assigns a non-negative length or size to vectors in a vector space.

**Mathematical Definition**

A norm is a function  $\|\cdot\| : V \rightarrow \mathbb{R}$  that satisfies:

- ▶  $\|x\| \geq 0$  and  $\|x\| = 0 \iff x = 0$  (Non-Negativity)
- ▶  $\|\alpha x\| = |\alpha| \|x\|$  for any scalar  $\alpha$  (Homogeneity)
- ▶  $\|x + y\| \leq \|x\| + \|y\|$  (Triangle Inequality)
- ▶  $\|x - y\| = 0 \implies x = y$  (Separation of vectors)

So now, if we see what is norm — basically, if we want to see, then it is a function which we will represent. We will represent it in this way, okay? So, there will be two lines, dot, two

lines. So, it gives us — like when we used to write modulus, we used to write the modulus value of any value — so by extending it a little, we added another line. So, we have given this — we will represent norm in this way, and it is defined from some vector space to the real line.

So, what is it now? Now we are defining it in this. So, what will we call norm? So, we will call norm only that who satisfies this condition. We have these three conditions. If it satisfies this, then we will call that norm. So, first we have this. In this case, it is called non-negativity. So, what does it mean? That our norm, which is the length — it is obvious that the length will always be positive, or it can be zero, but cannot be negative. So, we call this non-negativity. So, if  $x$  norm is greater than or equal to 0, it will always be. And if norm of any vector becomes zero, then what does it mean? That the vector is zero in itself. So, the norm of the zero vector itself can become zero. So, first, this is it.

Second, what is our homogeneity? What does it mean? That if we multiply any scalar by any vector, then both of them — if I want to take the norm of this alpha, then its modulus should come and the norm of  $x$  should come. Then the positive value of alpha will come. If alpha is scalar and real, then there is no problem in real. But if we have alpha as complex then we will multiply it by its modulus value, okay? So, this means that the norms will always be either zero or positive.

And the third one is triangle inequality. So, triangle inequality means that if we take two vectors and sum them, and after that take their norm, and if we have already taken the norm of the vectors, okay, then if it is like this, then it will satisfy it.

What does it mean? That the norm of  $x$  plus  $y$  is always less than or equal to the norm of  $x$  and the norm of  $y$ . So, this is triangle inequality, okay? Because in triangle, we know that the sum of the third side is always less than or equal to the sum of the other sides — that is always there. So, these three property— this is a property that will always define the name. The fourth one is just a property. The outcome is that if we take two vectors, and we say that the norm of the two vectors is zero and the difference between the two is zero, then it means that we will assume this value. Because we know that if the norm of a vector is zero, then it will represent us. This will show that the  $x-y$  should be zero. And from here, we will get  $x$  equal to  $y$ . So, now we can show from here that the separation of vectors — if their separation is zero, then it means that the two vectors will be the same. So, we will use this thing in error estimates, wherever we are dealing with vectors or matrices. Basically, if we go ahead and deal with matrices and vectors, we will use this thing a lot there.

Now we are defining vector norms. First of all, what is there in vector norm — that we have defined one vector. This one is okay. So, if we look at vectors in general, in school we used to write it like this —  $x_1, x_2, x_3$ . In this way, we will write it in the form of a column vector. So, even now we will write it in the form of a column vector.

So, we have written  $x$  transpose is a row vector. So, it means that  $x$  is a column vector. So, in this vector, we have  $n$  number of elements. So, I have to define its norm. So, first of all, I define norm — it is called one norm. So, what is one norm doing? One norm is basically the modulus of  $x_1, x_2$  — its modulus,  $x_n$  — its modulus. We have taken the modulus of all of them and added them. So, the norm that we got, we called it norm. Okay? So, we will call it one norm.

(Refer slide time: 8:22)

Suppose we have a vector  $x^T = (x_1, x_2, \dots, x_n)$  ✓

Various Norms

**1-Norm**

$$\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|$$

**2-Norm (Euclidean Norm)**

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = (x^T x)^{1/2}$$

**p-Norm (Holder's Norm)**

Now, if someone asks us that we are taking the modulus value in this — so we are taking the modulus value so that it always remains positive. If we do not take the modulus value, then we get to know that its resultant comes out to be negative. So, if it comes out to be negative, then it means that the definition of norm itself will change — because we know that norm will always be positive or non-negative. So, we have to take the modulus value.

So, this norm will satisfy all the properties, right? So, that's why we can call it norm. If it is not satisfying all the properties, then it means it is not a norm.

So, now we see whether it is able to satisfy the properties or not. So, suppose I take a vector  $x$  — I have  $x_1, x_2$  — I have only two elements of it. So, we took one norm of it. So, we got  $x_1$  modulus and  $x_2$  modulus. So, this is its one norm.

Now, we know that it will always be positive. If this norm is one — if I assume that it is zero, then what does it mean?  $x_1$  modulus plus  $x_2$  modulus — both of them are zero. And we know that the sum of two is zero — if there are two positive elements and their addition is zero, then what does it mean? That both of them are zero separately.

(Refer slide time: 10:12)

Suppose we have a vector  $x^T = (x_1, x_2, \dots, x_n)$  ✓

Various Norms

**1-Norm**

$$\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|$$

**2-Norm (Euclidean Norm)**

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = (x^T x)^{1/2}$$

**p-Norm (Holder's Norm)**

$x = (x_1, x_2)$   
 $\|x\|_1 = |x_1| + |x_2|$   
 $\|x\|_1 = 0$   
 $\Rightarrow |x_1| + |x_2| = 0$   
 $\Rightarrow |x_1| = 0 \quad |x_2| = 0$

And from here, we came to know that  $x_1$  will also be zero and  $x_2$  will also be zero. So, from this we get — now we know that both the values have become zero. So, from here we get that if the value of our  $x$  is 0, then  $x$  is also zero in itself. So, our first property was satisfied.

The second property is that if we multiply it by  $\phi$ . What does multiplying by  $\phi$  mean? If I multiply this by something — then suppose I multiply by  $\alpha x$  by this, then what would  $\alpha x$  mean?  $\alpha$  would be multiplied by  $x_1$  plus, it would also be multiplied by  $\alpha x_2$ .

Now, we can take  $\alpha$  as common. So, if we take  $\alpha$  as common, we will get the positive value of  $\alpha$ , and inside we have  $x_1, x_2$ . We can write it like this because these are real numbers. So, we can take it as common. And from here we got to know that  $\alpha$  times — and this has become the norm of  $x$ . So, this will satisfy it.

And the third property would be  $x + y$ . Now we have taken  $x$ . I will also take  $y$ , a vector —  $y_1, y_2$ . So, we add them. So, if we look here, we will have  $x_1 + y_1$  — one element, and  $x_2, y_2$  is the second element. Vector — we have to take its norm. So, this will come —  $x_1 + y_1$ , plus  $x_2 + y_2$ .

So, if we look at this, it will satisfy — now the sum of the elements is their modulus value. So, we know that, by the number properties of the number, we can write it like this — plus like this — it has become ours —  $x_1 + x_2$ , okay, plus  $y_1, y_2$ . We can write from here that the  $x + y$  that is less than — what is this? Norm of  $x$  and this — norm of  $y$ . So, the third property that we had will be satisfied — this one, triangle inequality okay?

So, all these three properties of ours have been satisfied. So, we will say that this norm is it, and we have named it, one norm. Now, the next thing that will come is two norms. Two norms means Euclidean norms, euclidean norm — whatever we have done till now in schooling, we have taken this norm.

What are we doing? Elements — we will square their elements, we will sum them all, and take them under the root. So, what does it mean? That we are doing this work, and this norm is used the most — because if we have taken this vector —  $x$  is our vector — so if we define the vector  $x$  here like this, then if our  $x$ , as I have just taken, is  $x_1, x_2$ , then if we remember that if we used to take its modulus or we write it two norm, then we used to do  $x_1$ , square,  $x_2$ , square, under the root.

(Refer slide time: 14:48)

Suppose we have a vector  $x' = (x_1, x_2, \dots, x_n)$

**Various Norms**

**1-Norm**  
 $\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|$

**2-Norm (Euclidean Norm)**  
 $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$   
 $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$   
 $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = (x^T x)^{1/2}$

**p-Norm (Holder's Norm)**  
 $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \quad 1 \leq p < \infty$

Handwritten notes on the slide:  
 $x = (x_1, x_2)$   
 $\|x\|_1 = |x_1| + |x_2|$   
 $\|x\|_2 = 0 \Rightarrow |x_1| + |x_2| = 0 \Rightarrow |x_1| = 0 \wedge |x_2| = 0 \Rightarrow x_1 = 0 \wedge x_2 = 0 \Rightarrow x = 0$

So, this work is always done in schooling. So now, we can write it in this form, and I can call it — that this is the euclidean norm or two norm. Now, what does it mean? It means that we can also write it like this — the x transpose will become the row — it has become the row vector — and x is the column vector,  $x_1, x_2$ . Okay? And it's under the root. So, if we multiply it in the form of matrix, then  $x_1$  square plus  $x_2$  square — under the root — came. So, this became our euclidean norm. So, what does it mean? That we can define euclidean norm like this also. So, we can check that it will also satisfy all the three properties of the norm, and that is why we will call it a norm. So, its name is two norm or euclidean norms. We saw that we took one norm, we took two norm — in this way, we can take different norms. We can take three norms, can take five norms. So, we have defined a general norm. It is called p-norm, and we also name it Holder's norm.

So, what is Holder's norm? That here we will write p — summation i from one to n — modulus of  $x_i$  is raised to the power p — we will sum all of them, and we will take the root of 1 by p. And p is starting from 1 and going up to infinity, excluding the infinity. So, we call this norm — p-norm. So now, in this we can define many norms of p. If you take the value one, then it comes to this. If you take the value two, then it comes to 3, then 5, 6. We can take norms like this, and we will get to know that we can define different types of norms, and this satisfies all the three properties. That's why we call it norm. So, this is p-norm which we are taking in the form of vector — for the vectors. So now, after this, we see one, two, three, p here but did not take infinity. So, you can see that if we want to take infinity norm, what would it mean here?

So, its meaning is this — so, we have defined a new type of norm. It is called maximum norm or infinity norm. So, what happened in that — if I took p to infinity, then what would happen? So, I wrote x and here the value of p became infinity.

So, when we calculated, we found that it is equal to the maximum of  $x_i$ . What does it mean? That we took the maximum of the set —  $x_1, x_2$  modulus,  $x_3$  — taking absolute value — and in the last, absolute value  $x_n$ , this.

So, this value which came out — absolute value came out — we found out the maximum value in it. So, it became our infinity norm. So, we named it — that from now on, we will call it infinity norm, okay?

(Refer slide time: 18:21)

The image shows a presentation slide with the following content:

- At the top, a formula for the p-norm:  $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$  for  $1 \leq p < \infty$ . There is a red checkmark next to it.
- The slide title is "Vector and Matrix Norms".
- Under "Examples", there is a section for "Vector Norms" and "Infinity Norm (Max Norm)".
- The infinity norm is defined as  $\|x\|_\infty = \max |x_i| = \max\{|x_1|, |x_2|, \dots, |x_n|\}$ . The infinity symbol and the formula are circled in red.
- Handwritten notes in red ink show:
  - $\|x+y\|_1 = |x_1| + |x_2| = |x_1 + x_2| = \|x\|_1$
  - $\|x+y\|_\infty = \max\{|x_1+y_1|, |x_2+y_2|\}$

So, this is also satisfy all the properties, and that is why there is infinity norm. The advantage in this is that infinity norm — it is used a lot, and all these norms are easy to define. In this, you will see the easiest one is infinity norm — look at all the values, their modulus values, and from that, we took the maximum value. So, what does it mean? That is why we also call it max norm. And we can prove it as well — we can test it. Like if I have taken any vector, what can you do with that vector? Find one norm, then find two norm, then find three norm, then suppose  $p$  — take 10, then take  $p$  as 100. Then, it will be known how many values have come for 100, and ultimately, you will see that is infinity norm and the value of  $p$  for 100 will be the same value. So, this means we can say that  $p$  is going toward infinity. So, if we define it, then it will become infinity norm. So, now that we have defined this, we can define different types of examples in it. I can take different types of examples. So, let's take an example — that how can we calculate it.

So, like I took  $x$ , I took a vector  $x$ . I defined it as 1 minus 2, 3, 5. I took this value and transpose it, so it means column vector. So what did we do after the column vector? I defined it as its one norm. Let me see what will be its one norm. What will be one plus two? Because the minus value so modulus value plus, 3, plus, 5, so this comes out to be 3, plus, 3, 6, so 11 comes. So the one norm comes out to be 11. It means the size of this vector is 11. When we calculate it in this way, now we see two norm. What I did now? See the 2 norm of  $x$ , will come out to be 1 square plus the square of 2 plus the square of 3 plus the square of 5 under the root of this value. Okay, so 1, plus, 4 plus 9 plus 25 under the root of this. So how much will it come out to be? 5, 14, under the root 39. Okay, so under the root of 39. So the value of this will be approximations, and you saw that 6 point something came out, okay. Because if it would be 36 then 6 would come. So, 6 point something value comes out. So, this means that our two norms, as we saw come to be 6 point something. and one norm which came out was 11.

Now, what did we do? We can calculate it like this with the help of a calculator. So what will we do now? What did we do? That I can take the 10 norm of this. So I took the ten norm. So I saw that 1, plus 2 to the power of 10 plus 3 to the power of 10 plus 5 to the power 10, and after that, I wrote 1 by 10. Okay, so I calculated it like this. We can do this with the help of a calculator. We can do it in the computer as well. After that, I saw that let me take infinity norm. So what do I have to do in infinity norm? The maximum one, minus two norm under root minus two modulus value 3,5. So the maximum value that is coming in this is 5, so its infinity norm is 5. Okay, so you see that one norm is coming as 11. The two norm is coming as 6 point, something and 10 norm, we can calculate this. And the infinity norm that is coming is 5. So it was the easiest to calculate.

So, infinity norm is called for this because the value of  $p$  that we are taking is the maximum value, that is why we call it infinity norm. And if you see here, the one norm is the largest. It means that if we measure the size of this vector, then its size in one norm is 11. And in infinity norm, maximum norm, its size is 5. So it does not mean that this maximum is giving the maximum value. It means that the size of the norm that comes out will be the maximum. So we can calculate it in this way. In the same way, we can calculate complex vectors as well. What happens in complex is that if we have some vector in complex form, then suppose I take some vector in complex form. What will we do in vectors? We know that if we have any  $z$ ,  $x$  plus  $iy$ . If you remember we want to find its modulus. Okay. So, square of modulus then we write it, its conjugate multiplied by  $z$ , right? So from there, we can calculate it.

So its conjugate is  $x$  minus  $iy$ , and this will become  $x$  plus  $iy$ . So it will come out as  $x$  square plus  $y$  square. So, by using the calculation formula, we will simply apply the minus  $i$  square, which is minus. So this value came, and the  $z$  came out to as  $x$  square plus  $y$  square. So, the value of this is what we had, we are doing it in complex numbers. So, in the same manner if I have some vector, suppose I have only 2 quantity, first I take some  $5i$ , in second I have  $1+i$ , this then we got one vector. So, if we want to find its norm. So you will see that we had defined it. So the norm, two norm, if I want to define it like this, then what will come? So I wrote  $x$  transpose. So now we are not taking  $x$  transpose, we have to take the conjugate also. Okay, and I multiply it with  $x$  and write it here as square.

(Refer slide time: 26:26)

$$\text{Complex form} \left\{ \begin{array}{l} z = x + iy \\ |z|^2 = \bar{z} z = (x - iy)(x + iy) = x^2 + y^2 \\ |z| = \sqrt{x^2 + y^2} \end{array} \right.$$

$$x = \begin{bmatrix} 5i \\ 1+i \end{bmatrix} \quad \|x\|^2 = \bar{x}^T x$$

So now we will write it here. Now we will take its conjugate. Okay, if we take the conjugate, then minus  $5i$  became, and  $1$  minus  $i$  became. This one became our vector, and here it came  $5i$ , and this came  $1$  plus  $i$ . Okay, so we will multiply these with each other, and whatever value comes after multiplying this, we will take its under root. So our  $x$  is the under root that we have, and whatever values come after multiplying it—because  $5i$  square will come  $25$ —so whatever we add, its value will come, and that will come to the norm of this vector. So we have to keep in the complex that we also take conjugate along with transpose. So mostly, if we see in this course, we are dealing only with real numbers, but if you have to use it further, so we can do it like this. So this is a vector norm. Now we want to see the vector norm. So now we know how to calculate it, if we have a vector.

Now we have to see the matrix norm. So if we see, the matrix is also arranging the vectors in column form or row form. Then we get a matrix. So it means that the matrix norm is somewhere closely related with the vector norms.

So let's see what is its definition. So how do we define the matrix norm? So the matrix norm is a function that assigns a non-negative scalar to the matrix and satisfies the following properties. So what are the following properties? The first property is non-negativity. So if the norm of a matrix will always be greater than equal zero. And if I say that the norm of a matrix is zero, then it means that the matrix itself is zero. It is a zero matrix. A zero matrix whose all elements are zero. So we call it the zero matrix. So we have defined it as the zero matrix. Zero made in the form of matrix now we have absolute. Absolute means if we multiplying a scalar then the norm of the matrix will be equal to this value, as we have seen that the same thing happens in it as well, after that if we see the third, then triangle equality should be in the norm  $a$  plus  $b$ , their norm must be less equal to the norm of  $a$  plus the norm of  $b$ , so we had these three things in vector but what is in matrix norm in this one more thing comes because we know that if we multiply the matrix, then we know the way in which

matrix multiplication is done, so one additional thing comes in it and we call it Schwarz inequality, so one more thing comes in the norms in the matrix that if we multiply two matrices and take their norm, then the norm will be less than equal to the norm of a into the norm of b, so we will have to take care of this thing a little bit, will we also need the additional condition required, so if it is satisfied then we will say that the size of the matrix is this, so we can calculate its size from here, so now we have the matrix norm, now we have to see how we will define the matrix norms, so here we have the induced norm, so how do we define the induced norm or operator norm.

(Refer slide time: 30:31)

## Induced (Operator) Norms

For an  $m \times n$  matrix  $\mathbf{A}$ , the induced norm is defined as:

$$\|\mathbf{A}\|_p = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p}$$

if  $\mathbf{x}$  is replaced by  $\mathbf{x}/\|\mathbf{x}\|$ , then we can write

$$\|\mathbf{A}\|_p = \max_{\mathbf{x}} \|\mathbf{Ax}\|_p \|\mathbf{x}\| = 1,$$

### Compatibility of matrix and vector norms

The norm of a matrix  $\|\mathbf{A}\|$  is said to be compatible (consistent) with the norm of a vector  $\|\mathbf{x}\|$ , if  $\|\mathbf{Ax}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|$

so here we have for an m cross n matrix A, the induced norm is defined as this one. why we are saying induced norm because it is being induced from the vector norm, if we have a matrix, I have a matrix A where suppose 2 cross 2, if I multiply it by any vector 2 cross 1, then we know that the vector that will come on the right side, if I write that vector, then the 2 cross 1 will be a vector, so it means that Ax is a vector and we have p, taken its norm, we have divided it, the norm of the x is also a vector, this is also a vector norm, okay and I wrote that the maximum of all the x such that x cannot be equal to zero because we know If x is zero then its norm will also be zero, so division will not be allowed here, so we did it from here and after calculating from here, the norm that we got, we said that this norm is an induced norm; it is an induced norm because it is being induced from the vector norm. So now if we change this norm and replace x by x divided by its magnitude, it means that I can write the same thing like this that the induced norm of the matrix A can be written as the maximum of all the x such that the value of these names is one. So if we have any matrix, like I have a matrix A, I can take it like 2,1,1,2 , and I have to find its norm, so I can take any value of p, so I have done this What do I take as the value of suppose I am calculating in this? So suppose I take p as 1. So, if I want to find the value of one norm of A, if I want to induce from here, then what will I do? Maximum of x, A, I have defined x as x1, x2, and from here, if you see, we will get a vector, and we will get 2 x1 plus x2, and x1 plus 2 x2 is a vector. I will take the one norm of that vector. Okay, so if we see, we have an Ax that comes out this, 2 x1 plus x2 and x1 plus 2 x2. Now, if I want to take its one norm, then the one norm, if you

see, its one norm will be the modulus of  $2x_1$  plus  $x_2$  modulus, plus the modulus of  $x_1$  plus  $x_2$ . This is what we have.

(Refer slide time: 33:58)

### Induced (Operator) Norms

For an  $m \times n$  matrix  $\mathbf{A}$ , the induced norm is defined as:

$$\|\mathbf{A}\|_p = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p}$$

$$\|\mathbf{A}\|_1 = \max_x \begin{bmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix}$$

if  $\mathbf{x}$  is replaced by  $\mathbf{x}/\|\mathbf{x}\|$ , then we can write

$$\|\mathbf{A}\|_p = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|_p$$

Compatibility of matrix and vector norms

The norm of a matrix  $\|\mathbf{A}\|$  is said to be compatible (consistent) with the norm of a vector  $\|\mathbf{x}\|$ , if  $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|$

Okay, now how can we do this? Now, I can write it a little more. So now, the values that we have, we will calculate in such a way that I will take all those  $\mathbf{x}$  whose norm is one. So suppose I have taken  $\mathbf{x}$ , some  $(1,0)$  its norm is one, and one norm, we know that. Okay, like this, I have taken an element 0, and 1 then I have to take all those elements whose one norm is one, and keep applying it, and see that the maximum that comes out will become our one norm. So this is called induced norm. Okay, so we will use this to find out how we can define different norms.

Now, a condition comes together — compatibility of the matrix and vector norms. So how will we know? Look here, the vector norm is being defined. In this, the matrix norm is being defined. So both are being linked. So we will say, how is the compatibility? So if we say that the norm of the matrix is said to be compatible, consistent with the norm of the vector  $\mathbf{x}$ , if the norm of  $\mathbf{Ax}$  becomes a vector, then that should be less than equal to the norm, or the vector becomes the matrix norm, and this vector. If it is satisfied, then only we will say that the vector norm and the matrix norm are compatible with each other because a matrix is a matrix, right, which has rows and columns, and we have defined the vector norm till now. So, from that, we will know whether both are compatible or not. On the basis of that, we will use it to find out the induced norm.

So now let us see what norms we can define in the matrix. So see, in the matrix, first of all we defined one norm. So one norm in vector — we knew what is one norm in vector — we have just defined it. So using that one norm, we induced matrix norm one. So we defined it. So we saw what is the maximum column size. Maximum column size means we have to sum over the columns. Now this is our matrix. In it, first column, second column, third column. So we have to sum column-wise, and the maximum column sum in it becomes our one norm. So you see,  $i$  is 1 to  $m$ , and  $j$  is 1 to  $n$ . Okay, so I maximum and this. So what will happen? If I calculate it, then its values are coming. See, maximum  $j$  between 1 to  $n$ , and the summation is  $a_{1j} + a_{2j} + \dots + a_{mj}$ . This has come.

Okay, so one  $j$  will move completely to  $j$  will move completely. So the maximum that comes in the column sum, we will call it one norm. Similarly, infinity norm came. We named infinity norm maximum row sum.

(Refer slide time: 37:50)

Induced (for  $p = 1$ ) 1-Norm (Maximum Column Sum)

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| = \max_{1 \leq j \leq n} (|a_{1j}| + |a_{2j}| + \dots + |a_{mj}|)$$

Induced (for  $p = \infty$ ) Infinity Norm (Maximum Row Sum)

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$$

Spectral Norm (for  $p = 2$ ) (Induced 2-Norm) (largest singular value  $\sigma$ )

$$\|A\|_2 = \sigma_{\max}(A)$$

Now our summation is over the row, that is, we did row-wise sum. Summed the first row, summed the second, summed the third row. Okay, here we have to do it by taking the absolute value. Summed them. Then after that, we saw the maximum value. So we got infinity norm.

Okay, so if we get any norm — let's go ahead and do an example — we have defined it with infinity norm. So the one norm is maximum column sum and infinity norm, that maximum row sum. Now one more norm comes which is important, which we call it is called spectral norm or two norm, or it is also called spectral norm. So let's take the spectral norm now. So the spectral norm is for  $p$  is equal to 2, and it means the largest singular value. Singular value means the eigenvalue. So we define it that if we have to take the two norm of any matrix, then take out all its eigenvalues — singular values of the matrix. Okay, we will take their under the root. So basically, if we can prove that we have to find the  $A$  transpose  $A$  — basically for the eigenvalue — so that's why we call it singular value. So if we take this singular value of a matrix  $A$ , then the maximum singular value in it, we will define it as two norm and name it spectral norm.

Now, till now we have taken out the euclidean norm in all. Okay, so in a similar way, there is a norm in a matrix. It is called the Frobenius norm, the Euclidean norm, or the Schur norm. So this norm is like that. So, as we have been doing it, what will happen in the Frobenius norm? What does it mean? Square the elements of the matrix. Square all the elements. Then add them all, and then take the under root of it. We can also represent it like this: transpose the matrix, then multiply it with  $A$  and then find its trace. Trace means the sum of the diagonal elements, and take its under root. So, if we do this, then the norm that comes out will be called Frobenius norm. So, what should we write for this? It measures over energy of the matrix. So, the values of norm — we say how much energy this matrix has — and it is calculated a lot for numerical stability. It is used a lot. So, from this, we can define the norms.

So now, if we take an example, then let us take a simple example so that we can calculate it here. So let us take a matrix A. I gave its value to it. Suppose I take it 2, 3, 1, 4, -2, 0, 0. Okay, and I gave it here. Let's take zero here as well. I take the upper triangular matrix — very means easily we can do this. This is an upper triangular matrix, upper triangular matrix.

(Refer slide time: 41:45)

Ex  $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix} \rightarrow$  upper triangular matrix



Now, what we have to do is define the one norm of this matrix. So before defining the one norm, we had told that this maximum will take the maximum column sum. Okay, so what does it mean to take from column sum? Now, I summed the first columns, so the maximum of all the columns' sum. So we summed the first column, so it came out to be maximum of 2, for second this come 3 plus 4, and if we did the sum of the third, it came out to be 1 plus minus 2. So we calculated from here. So this came out to be maximum of 2, 7, and 3 so from here, the value came out 7. So we can say that the size of matrix A is seven.

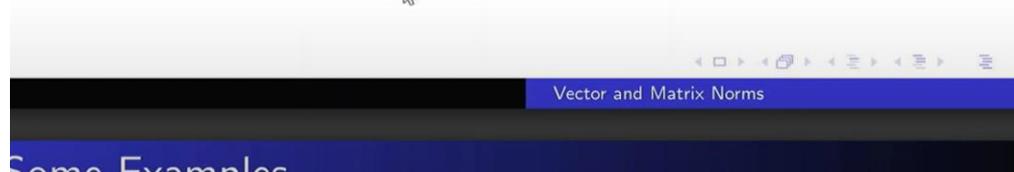
Now, what I do is take the infinity norm of this matrix, and we know the infinity norm — it's the maximum row sum. So what did we do? Maximum and take in rows. So the first row is 2,3,1 then 4 and then minus 2 is here. We took its maximum value. So what is its value? 3 plus two — 5 and 6. So its value is 6, and the value is one norm is 7 and the infinity norm is 6.

Now, what will we do with this norm? Now we don't know what is its eigenvalues of matrix. So that's why I took simple. So A, if I write its eigenvalues, singular values only, then we say that the matrix is not necessarily a square matrix. So that's why we talk about singular values. So if we find out the eigenvalue of A, then in this we can give 2, 4, -2 are eigenvalues.

(Refer slide time: 44:20)

Ex  $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix} \rightarrow$  upper triangular matrix

$$\|A\|_1 = \max\{2, |3|+|1|, |1|+|-2|\} = \max\{2, 7, 3\} = \underline{7}$$
$$\|A\|_\infty = \max\{|2|+|3|+|1|, 4, |-2|\} = \max\{6, 4, 2\} = \underline{6}$$
$$\text{eigenvalue}(A) = \underline{2, 4, -2}$$



So if we see in this, there is real eigenvalue. It can happen that we have complex eigenvalues, so in complex eigenvalues we have to do  $A$  transpose  $A$ . After that, we find out its value. We will tell that from here. If I want to find the two norm of this matrix — the two norm or whatever we have — then we have said that the norm of this matrix, the Frobenius norm, is approximately 5, meaning it is less than 6. I found the one norm that was 7, and the infinity norm was 6. The two norm, the spectral norm, came out to be four, and the Frobenius norm also came out to be approximately close to six.

So, in this way, we can define norms of different types, depending on our requirement. How can we find out? So if we have to find out the energy norm, the energy of the matrix, then we will use Frobenius norms.

And otherwise, generally, what we can do is that we use spectral norms. If we have stability or numerical schemes, if the eigenvalues are involved somewhere in them, then what we can do is we can use spectral norms. From there, we can find out the largest eigenvalue — what we will have — and on its basis, we will discuss its stability and convergence later. So, in this way, we can define one norm, infinity norm, all the norms.

(Refer slide time: 46:09)

Some Examples

$$\|A\|_F = \sqrt{4+9+1+16+4} = \sqrt{34} \approx 5.9$$
$$A^T A = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 4 & 0 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 2 \\ 6 & 25 & 3 \\ 2 & 3 & 5 \end{bmatrix}$$
$$\text{trace} = 4+25+5 = 34$$
$$\|A\|_F = (\text{trace}(A^T A))^{1/2} = \sqrt{34}$$

Okay, so now if we look at the spectral norm, we can say that — how did the norm come about? How did the two norm become related to the eigenvalue? So if we look at it on the basis of eigenvalue, then there is a derivation of it. We will discuss the derivation — how do we link these two norms with each other with the eigenvalue? Okay, so it will have to be linked with the eigenvalue. So now we do this — look at the spectral norm — we just have to give it a little idea that how is it being defined in this form. So see, we have to find out the two norm. So we can write it like this: if we had just defined induced norm, then on its basis I can write that — how will it be the maximum of all the  $x$  such that  $Ax$  — which is over the matrix, over the vectors whose norm is one. Okay, and this two norm.

So what are we doing? We will pick all those vectors whose two norm is one, and we will define the maximum value that comes over it. So here we will get the two norm. We can write it like this — see, the two norm square — so if I want to write it as square, what will it mean? Maximum over all the  $x$ , its square will also come. Now if I take it as square, then it becomes this — we have maximum. Okay, so  $Ax$  square. Okay, so we are taking the two norm, everywhere we are taking two norm. So if we know that there is an  $x$ , if its norm is square, then you know that we can take it like this. Okay, so what does it mean? If you see, it is called inner product. It is the inner product — just like we used to take the dot product — so I can write it here like this: maximum over  $x$ , and the inner product is defined like this —  $\langle x, x \rangle$  this is inner product — so I can write this,  $\langle Ax, Ax \rangle$  this is the inner product of  $Ax$ .

Okay, we have these values. Now this I can write — maximum of over the  $x$  — so we know we can write it like this. So I can define here as transpose of  $Ax$  into  $Ax$ . We will take transpose and we know that the definition of transpose is this. Now we have this — we have maximum over the  $x$ . It is going along with us, right? It is not like it was till here — it is always with it. So we can write it like this —  $x$  transpose,  $A$  transpose  $A$ ,  $x$ .

Okay, so if we see, then we have  $A$  transpose  $A$  is a new matrix. So I named it — so this became ours matrix  $B$ . Now going forward — maximum over the  $x$ ,  $x$  transpose  $Bx$  came. Okay, so now from here we get to know that our  $B$  is  $A$  transpose  $A$  and is always symmetric — was shown in the examples. If you see, then I had calculated it. So look, we have a matrix that is a symmetric matrix. Okay, so this matrix is a symmetric matrix, and going forward we will also check that this matrix is also positive definite.

(Refer slide time: 51:42)

**Derivation of Spectral Norm**

$$\|A\|_2 = \max_x \|Ax\|_2 \quad (\|x\|_2=1)$$

$$\|A\|_2^2 = \left( \max_x \|Ax\|_2 \right)^2 = \max_x \|Ax\|^2 = \max_x \langle Ax, Ax \rangle$$

$$\|x\|^2 = x^T x \rightarrow \text{inner product } \langle x, x \rangle$$

$$= \max_x (Ax)^T (Ax)$$

$$= \max_x (x^T A^T A x)$$

$$= \max_x x^T \underbrace{(A^T A)}_B x$$

$$= \max_x x^T B x$$

$B = A^T A \rightarrow$  Symmetric

Okay, so it is a symmetric matrix. So the eigenvalues are always real, and if we take positive definite, then its eigenvalues will always be positive. So what have we done? Now look — when we used to find the eigenvalue of any matrix, then we used to write it like this —  $Ax$  is equal to  $\lambda x$ . So what am I doing from there? If we look here, then what I do — I multiply it on both sides by  $x^T$ , and from here I write  $\lambda$  as  $x^T Ax / x^T x$ . So what does it become?  $x^T Ax / x^T x$  — and this is below — so we can write it like this — square. It will always be positive. So the values of  $\lambda$  will depend on this. If this value is positive, then  $\lambda$  will be positive, and if it is negative, then it will be negative. So everything depends on this eigenvalue. So we said that our matrix  $B$  is symmetric and is positive definite in this case. So we accepted it.

So what does it mean? That now we are taking these values in it — we are calculating divide by this norm, meaning the vector  $x$  — but the  $x$  vector that we have and its norm is automatically one. So it means that from here we can say that we have maximum  $x^T Bx / x^T x$ . I can write it like this — because it is one, right?

So what does it mean? This is the  $\lambda$ , okay. So if our  $\lambda$  in this case is our eigenvalue, then  $x$  becomes the eigenvector, corresponding, right? So what does it mean that we are taking our  $x$  is the eigenvector corresponding to the eigenvalue  $\lambda$ , right? So what does it mean that the our norm, two norm, suppose we take the maximum eigenvalue, okay. So we will take the maximum eigenvalue of  $B$ , meaning  $A^T A$  which means we will get the maximum eigenvalue of  $A^T A$ , under the root. And now you will see that we have to find the eigenvalue,  $A^T A$ . So if we have the matrix  $A$ , so that matrix is a symmetric matrix, then if we are in a symmetric matrix, or we have a symmetric matrix, then we know that the eigenvalue of  $A^T A$  is same so the that comes here is  $\lambda^2$ , if its is symmetric. If we take its under root, then  $\lambda$  will come.

So basically, with the help of this expression, we can find the spectral norm. Where the matrix is given to us, we will multiply its transpose by this. So the matrix that we get here is symmetric and is positively definite. Positively definite means that only then we can take its under root. If we take its negative under root, then it will come in the form of complex. So this is our  $A^T A$  is positive definite. So, from there we can find all its eigenvalues and the maximum on value in it, if we take its under root, it will become our spectral norm. So, in this way, our norm is spectral norm. So, we have just given an idea that how can we take the largest eigenvalue in it, and we can write it in the form of spectral norm. So this spectral norm is with us. So we can write it in the form of spectral norm. We call it spectral norms.

We have given its examples earlier. Now we have a summary sheet. So what is a summary sheet?

(Refer slide time: 56:38)

## Summary Table

Norm Type	Formula	Interpretation
1-Norm	$\ x\ _1 = \sum  x_i $	Sum of absolute values
2-Norm	$\ x\ _2 = \sqrt{\sum x_i^2}$	Euclidean distance
Infinity Norm	$\ x\ _\infty = \max  x_i $	Largest abs. component
Induced 1-Norm	$\ A\ _1 = \max \sum  a_{ij} $	Max column sum
Induced Infinity Norm	$\ A\ _\infty = \max \sum  a_{ij} $	Max row sum
Spectral Norm	$\ A\ _2 = \sigma_{\max}(A)$	Largest eigen value
Frobenius Norm	$\ A\ _F = \sqrt{\sum a_{ij}^2}$	Overall energy of matrix

One norm, this is done. Two norm, this is done, infinity norm, largest value done, induced one norm of the matrix is this, okay. Induced infinity norm, so maximum column, maximum row, spectral norm, this is. Frobenius norm, this is, okay.

So overall, all our norms are the same. We will use, if we are dealing with vectors or matrix, then we have done this. So we can also show a little code of it, that if we, then we can write this in the form of code also. We can calculate this in Python also, matrix.

(Refer slide time: 57:27)

## Vector and Matrix Norms

```
[5]: # Vector Norm
import numpy as np

#vec = np.arange(5) # initialize vector
vec=np.array([-3,5,9,2,6,10])

# compute norm of vector
vec_norm = np.linalg.norm(vec)
print('Given vector is:', vec)
print("Vector norm:", vec_norm)
```

```
Given vector is: [-3  5  9  2  6 10]
Vector norm: 15.968719422671311
```

```
[14]: # Matrix norm
import numpy as np

# initialize matrix
mat = np.array([[ 1, 3, 1],
                [4, 5, 2]])

# compute norm of matrix
mat_norm = np.linalg.norm(mat)
```



NPTEL

So what did we do? I wrote a code in which we defined. First we did import NumPy. Then I defined a vector, its value is -3,5, 9,2, 6, 10. So suppose there are six values inside it, okay. So now I have to find its norm. So we will do a, the inbuilt library or function that is in NP is called linear algebra or see, L-I-N-A-L-G. Linear algebra has all the linear algebra-related work inside it. So if we are in linear algebra, then there is defined norm. So what did we do? I defined a vector norm and took the norm of this vector. Here I printed it and pressed enter. So we have this vector and the norm that came out, and it came out. The vector norm is basically considered in this case. We are taking the Euclidean norm. We can define it. If I do not want to take that, suppose I have just taken some value. What should I do? I take the three and take the four. So we know 3 square 9 plus 16 under root 5. So this vector norm came out, okay. Like this, I can define a matrix norm. What did we do in matrix norm? We initialized a matrix. So import NumPy as np. I made it a matrix, and this is the matrix that I have created. I have this 2 cross 3 matrix, okay.

So I want to get its matrix norm. So I defined it here as np dot linear algebra dot norm, and I printed it here, okay. So as soon as we print it, we get the answer, okay. So we can do it axis-wise as well. We can do it row-wise, column-wise. So from there, we can define how we can find the matrix norm.

So this norm, we can define different norms, okay. We will do that later. So just this, initially I have told, we wanted to check that if we see the vector norm and matrix norm with each other, then if I take the norm of the same vector, suppose I have defined an array. Its vector is 9,1,2,0,3, -2, -5,8,9. I took a vector which has 9 elements, okay.

After this, I reshaped it. We know that we reshaped it into a 3 by 3 matrix. So now this vector has become a matrix after being reshaped. Now I have defined the norm of that vector. So here I have taken the vector norm and printed it. I have taken the matrix norm of this and printed it. I have calculated it and seen that both the norms are coming out to be 16.40.

So what does it mean? That the matrix norm and the vector norm, after that the norm of the matrix formed from it is the same. What does it mean? That the matrix norm that is being taken here is frobenius norm, square of all the value. So square all the values of the matrix and under-root them. What did we do in this? Otherwise, what did we do? Square the vector and under-root them.

So this is how it comes out. So it means that our vector norm frobenius and this two norm of the vector will both give the same value, and we have the same value.

Okay, so now we can define all these norms in this way: Frobenius norm, one norm, infinity norm, spectral norm. So we wrote a code in which import norm NumPy, and I defined this. A function computer compute matrix norm A, here we will give A, in return we will get this value.

So we have taken a matrix: 1,2,3, 4, 5, 6, 7,8,9 okay. So its matrix here, we will call the function, and it will print it here. Like we did this norm, see it came. The Frobenius norm of this matrix is 16, we have checked it. The one norm that is coming is 18. The infinity norm that is coming is 24, and the spectral norm that is coming is 16.8.

So see, 16.8, Frobenius and spectral are coming the same. It means the largest eigenvalue is coming, okay. And the one norm that is 18, infinity norm is 24. It means the infinity norm

that is coming, the maximum row sum that is coming in this is 24. In this, it means the maximum norm that is coming.

So like this, I should change the value. 1,2 ,I wrote zero in place of 2. I wrote minus 3 in place of 3, okay. I did 5 in place of 1.

(Refer slide time: 1:02:56)

```
inf_norm = np.linalg.norm(A, np.inf)
spectral_norm = np.linalg.norm(A, 2)

return frobenius_norm, one_norm, inf_norm, spectral_norm

# Example matrix
A = np.array([[1, 0, -3],
              [4, 1, 6],
              [7, 8, 9]], dtype=float)

# Compute matrix norms
frobenius_norm, one_norm, inf_norm, spectral_norm = compute_matrix_norms(A)

# Print results
print("Matrix Norms:")
print("Frobenius Norm:", frobenius_norm)
print("1-Norm (Max Column Sum):", one_norm)
print("Infinity Norm (Max Row Sum):", inf_norm)
print("Spectral Norm (Max Eigen/Singular Value):", spectral_norm)
```

Matrix Norms:  
Frobenius Norm: 16.881943016134134  
1-Norm (Max Column Sum): 18.0  
Infinity Norm (Max Row Sum): 24.0  
Spectral Norm (Max Eigen/Singular Value): 16.84810335261421

I wrote zero here. I wrote minus. Now let's see what happens. Now this comes. So the frobenius norm 14 comes. The one norm 18 comes. The infinity norm 17 comes, and the spectral norm is 13.03, okay.

So this is our norm. We can calculate it in different ways. I can change it and see what happened. After calculating this, it is not matching, okay. This comes. So now see, the matrix is 2 cross 3 matrix. So its frobenius norm comes out to be 7.93. The one norm comes out to be 9, the maximum column, see the maximum column here, it is 3 and 6 becomes 9.

How much is coming in the row? 4 plus 5, 6 11 so it comes 11. Now see, this matrix is its 2 by 3 matrix. So we cannot find its eigenvalue. So what will we do? The spectral norm that we have defined, and how did we do it? Let's  $A^T A$ . So, we take  $A^T A$ . Suppose we take it, then you will see that the matrix is 3 cross 3 matrix. It will go, okay, right?

So, as soon as the 3 cross 3 matrix is formed, after that we found its eigenvalue and its maximum value, which was its square root. So, this 7 which we got, we cannot say that this 7 is the eigenvalue. So, in this case, we would call it singular value. So we are talking about maximum singular value.

(Refer slide time: 1:05:17)

```
inf_norm = np.linalg.norm(A, np.inf)
spectral_norm = np.linalg.norm(A, 2)

return frobenius_norm, one_norm, inf_norm, spectral_norm

# Example matrix
A = np.array([[1, 0, -3],
              [4, 1, 6]], dtype=float)

# Compute matrix norms
frobenius_norm, one_norm, inf_norm, spectral_norm = compute_matrix_norms(A)

# Print results
print("Matrix Norms:")
print("Frobenius Norm:", frobenius_norm)
print("1-Norm (Max Column Sum):", one_norm)
print("Infinity Norm (Max Row Sum):", inf_norm)
print("Spectral Norm (Max Eigen/Singular Value):", spectral_norm)

Matrix Norms:
Frobenius Norm: 7.937253933193772
1-Norm (Max Column Sum): 9.0
Infinity Norm (Max Row Sum): 11.0
Spectral Norm (Max Eigen/Singular Value): 7.560184073244183
```

So what does it mean? The spectral norm in this case is 7.56. So we can define norms for different matrices. We can define norms for different vectors, and with its help, we will use it further when we find out the iterative process to solve the matrix.

So, this is our Python code. With this, we can define norms of all the matrices that we have used, or we can define the norm of the vector. So, in the lecture that we saw today, we discussed every type of norm: vector norm, different types of vector norm, different types of matrix norm, and we will use these in the future.

So, thank you for watching this lecture.