SCIENTIFIC COMPUTING USING PYTHON

Professor Vivek Kumar Aggarwal & Professor Mani Mehra

Department of Mathematics

Indian Institute of Technology, Delhi

Lecture No. 26

Welcome, everybody, to Scientific Computing Using Python. So today, we have found out an approximation. So now let's see that the approximation that we have done, we have done it according to the interpolating polynomial, okay, and in that we saw that we defined the interpolating polynomial in different ways. So today we're going to do that if we are not doing the interpolation then how to do approximation. Okay, because if we have to do approximations, so for that we have a theorem and we call the theorem Wierstrass theorem then what does theorem say if a function, we define a square function as f x is continuous if it is continuous on a finite interval a to b then for given epsilon for given epsilon greater than zero there exist a number n, then this n is the function of epsilon, this will depend on this and a polynomial p x of degree n such that the function for which we have to find our approximate, I will take the difference between the two, that means I will take the difference between the two. Its magnitude will less than epsilon for all x belongs to a to b, this is our Wierstrass theorem. So, what does Wierstrass theorem say? What have we done till now? If we look at it, till now we have two ways. One is that we have a function given f x, okay, and x belongs to a to b, so that function should be continuous. So, we will use this theorem to find out the approximating polynomial. So, this is continuous form. Second, if we do not have a function given, we have such givens. We take some values, say x 0 is a, then suppose x 1, x 2, x3 and x n equal to b, suppose so what we did is we took n plus 1 points from this interval and there we also know the value of the function, f x 0, f of x1, we also know all these values. So, what we did is, on the basis of these values, we calculated its approximations from the polynomial, and we called that polynomial as interpolating polynomial but what will happen when we have a function given and that function is just a continuous function. So, if we don't want to find the interpolating polynomial. So how do we find approximation with the help of polynomial how much accuracy do we need? The epsilon is our accuracy. So, we know that we have studied Taylor series, but the function in Taylor series is the function and its derivative is well defined at a point. But what is there in this in Wierstrass is that the function we have is just a continuous function in the interval. So, for that we need an approximation polynomial whose degree will depend upon epsilon. So, what does it mean? The smaller the value of epsilon the higher is the degree of polynomial. So, this is the Wierstrass theorem. So, let's see with the help of this. So, we were told that we are not proving its proof. We just have this statement. So now the question is, now how to find p x a polynomial. So, how to find out a polynomial. And if I give it n, then n means the n th degree polynomial. Let's approximate the p n x. So, now see what do we do with this? With the help of approximations, we write it like this: c0 phi 0 x plus c0 phi 1 x thus, c n phi n x comes to us. So, I have written it with respect to the function as a linear combination of phi0, phi1, phi2, phi3, phi4, phi n x, where the functions phi0, phi1, phi2, phi3, phi4, phi n x that we have taken are all linearly independent. They are LI so LI, our phi i which we have calculated, the phi i which we have

calculated that zero,1 and not calculated that we have defined is called basis function. It is also called coordinate function. So, it is also called basis, it is also called coordinate.

(Refer Slide Time 08:05)

So, what does it mean that we will take the linear combination of our basis and from there we will get a polynomial of degree n, after that we will approximate it depending upon what degree of polynomial we have to take. We will calculate it according to the value of epsilon. Now the question is what are our phi i's so phi i's can be anything. It can be polynomial. It can be sine cosine function, trigonometry function, exponential, it can be anything. So, we have to proceed according to that. So, first of all we take functions which are phi I's, so generally we take phi I's, so phi is our x to the power one. And i is 0,1,2 up to n. We will take this. Now see, the phi0 that we will have will always be the same 1, phi1 is x, phi2 is x square and phi n is x to the power n. And our x belongs to zero to one. So, we know that this is the basis function or these are standard basis. We know that they are linearly independent. So, phi I here is linearly independent and these are standard basis. The standard basis of the vector space of the polynomial of degree less equal to n, these are its standard basis. So, we have taken the standard basis. We know that they are linearly independent. So, these became basis. So, we took this. After that, if we write it like this, then our If you see we have any function, now we have a function f x which we have to approximate. Okay, with the help of p x, now there are two values between them, so we have taken the mode, we can write it like this, basically it is an error, so I can take it like this f x minus p n x, now if this is written, it will come to c0 c1 x plus c2 x square c n x to the power of n. Okay, now the modulus value that we are taking, the distance that we are taking, we can take any of it, so I have made it a norm, we had defined vector norm matrix norm, so such functions also have norms, I will tell you that now, so here we have the difference between the two, so we have an error, okay and this error, if you see, if I write it with error, I will write it as error e and it depends on the function f and c0 c1 c2 up to c n, okay, what is that function and c1 c0 is unknown to us, what are their values Okay, so we got the error, so what is our purpose in this? Our main purpose is that we want to reduce our error, the E f c vector, its c, so the c is basically c1 up to c n, so our purpose is that this error should be the minimum least.

Now How to find $P_n(x) \rightarrow$ nth degree poly

$P_n(x) = c_0 \phi_0(x) + c_1 \phi_1(x) + \cdots + c_n \phi_n(x)$

where $\phi_0, \phi_1, \phi_2, \cdots \phi_n \rightarrow$ linearly independent (LI)

$\left\{ \phi_i \right\}_{i=0,1 \cdots n} \Rightarrow$ Basis functions / coordinate fn.

$\phi_i(x) = x^i \quad i = 0,1,2 \cdots n$

$\phi_0(x) = 1, \quad \phi_1(x) = x, \quad \phi_2(x) = x^2, \quad \cdots \quad \phi_n(x) = x^n \quad x \in [a,b]$

$\phi_i \rightarrow$ LI standard basis of $P_n(x)$

$\|f(x) - P_n(x)\| = \|f(x) - (c_0 + c_1 x + c_2 x^2 + \cdots c_n x^n)\| = E(f, c_0, c_1, c_2, \cdots c_n) \qquad c = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}$

We want that $E(f,c)$ should be minimum

$f(x) \quad x \in [a,b] \rightarrow$ Continuous form

$x_0 = a, x_1, x_2, x_3 \cdots x_n = b$
$f(x_0), f(x_1) \cdots f(x_n) \Rightarrow$

(Refer Slide Time 12:28)

So, if the error is the lowest between the two, then we can say from here that what would it mean, the best approximation means p n x would be our best approximation, right? So, we call it that we can make best approximations, so we have to make it minimum, so first of all we define the norms, so we know discrete norms, so see we have two types of things, now look we have two ways, as I also told above that if there is discrete data, then we have defined vectors norms here, right? Vector norms, we have defined one norm, okay, we did 2 norms, what did we name it p norms, infinity norms or all the norms we have defined. Okay, so the norms we have defined for this, like we have defined if there is a vector, we have to take its 2 norms, so we know what we have done, we have done summation, we have taken the square of each x I and taken it under the root. So, we have defined its norm, this is the 2 norm, so we have done this for discrete. Now we have to see how we can do this for continuous, for continuous data, continuous data, if I write what does it mean, how can we define a function, okay, so what are we doing for this, for function, now we will define its norm for the function. Now see, we have a function f, x is belonging to a,b continuous function, suppose, then what will we do, I define its norm, suppose I define p norm, so what will happen in p norm, see what was coming in it, summation was coming, okay, if I have any x i's here, okay, I want to write anything for x I's. We have some values here, okay, write its p norm, so summation x to the power p and after that p root this comes to us, okay, because x is a vector, the x we have will be x1, x2 to x n, like this we have a function, so what to do in summation, see, summation, we can take for finite discrete value for it, but as this becomes continuous, then this summation will become our integration, from where to where a to b now when we define norm, like we used to do in that also with respect to define the sum weight function, this w x is the weight function, if we take one, it will become simple, what will we do after that, f x, okay, so and we know that norm is always positive, so we will take the modulus value, we will take the power p d x, we will take its 1 by p value, so we call it P norm or LP norm, which are also called L spaces, and we are not discussing it here, so we defined norm in it, so we call it LP norm, okay, so in this Integration comes in this way.

$d_i \to \underline{L\cdot I}$ standard basis of $P_n(x)$

$$\|f(x) - P_n(x)\| = \|f(x) - (c_0 + c_1 x + c_2 x^2 + \cdots - c_n x^n)\| = E(f, c_0, c_1, c_2, \cdots c_n) \qquad c = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}$$

We want that $E(f, c)$ should be minimum/least

$\Rightarrow P_n(x) \to$ Best approximation

Discrete data $\to$ Vector norms :— $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_p, \|\cdot\|_\infty$

$$\|x\|_2 = \sqrt{\sum x_i^2} \qquad \|x\|_p = \left(\sum x_i^p\right)^{1/p} \qquad x = (x_1, x_2, \cdots x_n)$$

$\Rightarrow$ Continuous data $\to$ function $f(x) \quad x \in [a, b]$

norm $\quad \|f\|_p = \left(\displaystyle\int_a^b w(x)\, |f(x)|^p \, dx\right)^{1/p} \longrightarrow L_p \to$ norm

(Refer Slide Time 16:36)

If I take any function and want to define its 2 norm, then what would be the 2 norm, if we take the whole square of f x and take the square root, then this would be our L2 norm, okay, like this we can define infinity norm, now what is infinity norm because the function f x is our continuous function in the interval a to b, then its maximum will be found, okay, if the function continuous in a close interval, then its maximum it attain its maximum, we know, so what will we do in this, we will take the maximum of x between a to b and take the magnitude of the function, then we have infinity norm, okay, so we call it uniform, so we have three numbers, so this third one we call uniform norm, okay this is 2 norm, p L p norm, so what should we do now, now see what we have to do in this, we have to reduce the error, okay, we have to make minimum error and under some number, all this work will have to be done under some norm. Okay, so now we can take different numbers, so in this we are told that if I use Euclidean norm, okay, I mean which we also call 2 norm, if we use Euclidean norm for finding the minimum error, which we have to find, this process is called least square approximation, okay, and if we take uniform, if we use uniform, okay, then this process is called uniform approximation.

$$\text{norm} \begin{cases} \|f\|_p = \left(\int_a^b w(x) |f(x)|^p dx \right)^{1/r} \to L_p \to \text{norm} \\[2mm] \|f\|_2 = \left(\int_a^b w(x) (f(x))^2 \right)^{1/2} \Rightarrow L_2 \text{ norm} \\[2mm] \|f\|_\infty = \max_{a \leq x \leq b} |f(x)| \to \text{Uniform norm} \end{cases}$$

If we use Euclidean norm (2-norm) for finding the minimum error ⇒ least square app
Uniform n

NPTEL

(Refer Slide Time 19:20)

So now it depends which norm we take, okay, so what will we do in this, we will see with both norm but our main purpose in this case is to use Euclidean norm, 2 norm because this is least square approximation is widely used to find out approximate function, whichever function is the best approximation, okay, so for example, like I want to see how to find the norm for example, I take any example, I took the function f x is equal to 1 plus x and the x which is, I took it from suppose zero to 1, okay and let also I assume that the weight function is one, then this is a function, suppose I define 1 norm, what would it be from zero to 1. Now, 1 plus x, okay, will come as positive dx is positive between zero to 2. So, we have solved this. So, it will come as x plus x square by and I put the limit on it. So, this 1 plus 1 by 2 means 3 by 2. So, we take the norm of this function. Okay, it means its length. Basically, length is the measurement. So, its value is 3 by 2. Similarly, if I take its 2 norm, then I have taken the zero to 1, weight function is 1, whole square of 1 plus x d x, now we will take its under root. So, we can take the under root or I can square it. Now it comes to us from here. Okay, now if I do this, it becomes zero to 1 plus x square plus 2 x d x. Okay, and under the root. Now, see, if we solve this, we will get it here x plus x x will come to x cube by 3 plus 2x square by 2 zero to one and all 1 by 2 How many values do we have from here So this will come to 1 plus 1 by 3 plus 1. Okay, because from here two will cancel out one 1 by 3 and 1 and its under root so how much will it come to 2 plus 1 by3 under root so 7 by 3 its under root so here we have its two norm, Euclidean norm, okay like this I can take the 3 norm, 4 norm, 5 norm any function, if I want to take infinity norm so infinity what is mine, I have to take the maximum which x brings in our interval and this function which is magnitude of 1 plus x so 1 plus x is positive now we have to take its maximum x belongs to 0 1 x one so if we look at this then it will come to two only, okay so we have this 2, this Euclidean norm, this 1 norm, we have this one so like this we can define different norms.

If we use Euclidean norm (2-norm) for finding the minimum error $\Rightarrow$ **Least Square app**

" " " Uniform norm " $\Rightarrow$ **Uniform app**

$\underline{Exp}$ $f(x) = 1+x$, $x \in [0,1]$ & $w(x) = 1$

$$\|f\|_1 = \int_0^1 (1+x)\,dx = \left[x + \frac{x^2}{2}\right]_0^1 = 1 + \frac{1}{2} = \boxed{\frac{3}{2}}$$

$$\|f\|_2 = \left(\int_0^1 (1+x)^2\,dx\right)^{1/2} = \left(\int_0^1 (1+x^2+2x)\,dx\right)^{1/2} = \left(\left[x + \frac{x^3}{3} + 2\frac{x^2}{2}\right]_0^1\right)^{1/2}$$

$$= \sqrt{1 + \frac{1}{3} + 1} = \sqrt{2 + \frac{1}{3}} = \sqrt{\frac{7}{3}}$$

$$\|f\|_\infty = \max_{x \in [a,b]} |1+x| = \max_{x \in [0,1]} 1+x = 2$$

NPTEL

(Refer Slide Time 23:30)

So, we can define norms so our main purpose is that now we use only least square approximation. So, our purpose is to do least square approximation. What is least square approximation? So, we will do least square approximation for both. For discrete and for, we can take continuous as well as discrete. So now we have to see how to do this. So first we take for discrete data. Now see, if we have discrete data, then we have values. Suppose now see, we have x0, x1, x2 x n, so we can take these n points as well or we can take n plus 1 points as well. Now we don't have any problem. Because we are not defining an interpolating polynomial. We have points. So now onward, we can define it with x1 as well and start. Because there is no data in this that depends on the number of points. We will not do it now we have done this okay so what do I have to do we have a function so this value is given and here the values of the function are given what is f x1, f x2, f x n these values okay so what did we do the function f x I have used approximations c0 plus c1 x plus c2 x square c n x n with the help of this okay what is this my this is a polynomial of degree n now what did I do I took the 2 norm between these two okay so we will see this will become a function which will be dependent now we have to see on the x I's okay so this is basically because we will have this which will be true what will it mean first let's write it down so what are we going to do in this how are we going to define it see it will be defined like this a for x1 so this will be fx1 will be minus c0 c1 x1 plus c2 x1 square c n x1 n so this value has come okay so we will take the difference of these two one this is just one value okay so this is the value we have the difference of both We can take it in the same way, we will have fx2, this will be the value c0 c1 x2 c2 x2 square c n x n to the power n of c2, so what does it mean, f x1 is the same as this, we have approximated it from p x1, okay, fx2 is approximated from p x2, f x n is approximated from p x n and what are p, these are the values, okay, so what did we do, the difference between these two, now it all depends on this, these values are also known, these values are also known, so if we give, then this function of ours will be just c0, c1, c2, c n

function of these, okay, now if we have to write its function, now I am taking a two norm, then how will we define the two norm, so we have to take a two norm, so what do we have to do to define the two norm, see on one, on two and sum, so what do we have, weight function will come first at all the x I's, after that we will write f x i, this minus summation, I can write it that, I can write this k from we can take k from zero to n. This is C k x to the power k. Okay, I will square both of them. And after that, the under root. So, let us leave the under root because we will take the under root later. If we take the under root of this, it will become 2 norm. So, if we want to find the minimum of the function, then its square root will also be minimum. So, there is no problem in that. So, we have defined it like this. Okay, and this is true for all I 1 to n. Okay, so this value is now this value, small n and capital n can also be different. Okay, so in this, N and n can also be different. So, let me write it separately. It depends on what we need. We will also understand when we take an example. Okay, so this capital N, this difference is ours. This is it. So, we have done it like this. Now, if we want to see the same thing in it, then this is for discrete.

(Refer Slide Time 30:02)

Now for continuous data, continuous data, what does it mean that the function given f So, in that case, see what will it become now I c0,c1, c n what will it become what will it become in this f x will be minus p x okay n degree polynomial so n degree polynomial is any we can take k degree polynomial okay so let me write this p it can be of any degree so I will take this f x minus p x and take its 2 norm under root that's why I am not taking it if I take the under root then if it is minimum then the under root will also be minimum okay so we don't use the under root as unnecessary so we remove it so I have defined this means I took two norms when we take the norm then we will not take its under root so what will it become this will become integration a to b weight function any weight function after that f x minus p x whole square and under the root and I am not taking it okay so here comes from here so what does it mean this comes we have a to b weight function and from here comes f x minus c0, c1 x c2 x

square c n x to the power n and their square with respect to x integration we have passed so what happens now this function is c0, c1, c2 now if we want to find the minimum then we will have to find such c 0's so that this difference which is this norm should be minimum because this is an error so this thing which we are doing we call it least square approximation so in this the number of elements which are number of points can be anything so now if I want to define this then I can take the values in phi x so if you see I have taken phi i as x i so now see what to do with this.

$$\text{for Continuous data} :- \qquad f(x) \qquad b$$

$$I(c_0, c_1, \cdots c_n) = \| f(x) - p(x) \|_2 = \int_a^b w(x)\, (f(x) - p(x))^2\, dx$$

$$= \int_a^b w(x) \left[f(x) - (c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n) \right]^2 dx$$

(Refer Slide Time 32:43)

I do this that I take an example then we will get more clarity from that then what should I do now suppose I have a function f x is equal to x cube belongs between zero to 1 okay it is a continuous function question I find its linear and quadratic least square Approximation, I have least square approximation or best fit approximation with weight function equal to 1. So, if we solve this, the function is continuous function then first I will do it for linear. What is going to happen in linear? Let us do it in linear. Let us assume that our polynomial is linear. Suppose c0 c1 x is a polynomial, because it is linear then it will be like this. So, we have accepted this 1. Now, what do we have to calculate, f x minus c0 plus c1 x. What will happen to this? We will take its 2 norm, whose function, is it? The function is c0, c1. And how will we calculate this? I will calculate it from zero to 1 because zero to 1 is given. Weight function is one given. So, it will become a f x x square minus c0 c1 x. Its whole square d x. Okay, under the root, I have removed it. For the time being, we have to use this. I will assume it to be I c. Okay, because if this function is minimum, then its square root will also be minimum. So, what did we do? Let me write it as square. So, This will come, so what will we do now, we will remove the values of this and I will go ahead and write it like this, what will be my C0 c1, what will it be zero to 1 x cube minus c0 plus c1 x's whole square d x this comes, so we have accepted this function, so if it is minimum then obvious its square or its under root will also be minimum, so that's why we write this thing like this, okay, so generally we want

to write it, so what do we do, this 2, we remove it because we know that if there is a function, if there is some value, if it is minimum, then its square will also be minimum, okay, so it depends on us that with its help we have to find c0, c1 and now see, I will have to calculate it, I will have to find c0, c1, so what is this, I is a function of two variables c0, c1, okay and I have to find its minimum, so we know how to find the minimum, which is the necessary conditions to find out the minimum maximum. What happens is that del i by del c0 should be zero first. So, if we calculate this and del i by del c1 should also be zero. So, from here we will get two equations. So, if I calculate this, then it will come as 2 zero to 1 x cube minus c0 plus c1 x. And the square comes ahead and the coefficient of c0 is minus 1, so minus will come and this will become d x. So, from here if I write it, I will get zero to 1 x cube minus c0 minus c1 x d x is equal to 0. So, we have this one equation. We can solve this equation easily. So, we will solve it. So, what will come out to be x to the power of 4 minus c0 x minus c1 x square by 2, zero to 1 that is equal to 0, this equation that we have is a necessary condition that gives us a critical point which will give us the value of c0, c1. We say normal equation, okay, so normal equation, with the help of this, I have calculated it.

(Refer Slide Time 38:20)

It will come out that 1 by 4 minus c0 minus c1 by 2 equals to 0 and from here it will come to us that c0 plus c1 by 2 equal 1 by 4, this equation what will come in the second one, we will do it like this, okay, just doing it with respect to c1 If we calculate this, then when we take its derivative, then the x will come along with it. Okay so if we see here, it will become minus 2 0 to 1 x cube minus c0 minus c1 x x d x equal to 0. Okay, and if we put it equal to zero, the condition that we get from there will be zero to 1, now if there is x, then this will come x4 minus c0 x minus c1 x square d x equal to 0. And if we calculate this, then this will come x 5 by 5 minus c0 x square by 2 minus c1 x cube by 3 zero to 1 is equal to 0. Okay, so we have calculated this, then the values that we get, if we put one one, then this will come to us a c 0 by 2 plus c1 by 3 equal 1 by 5. This is what we have. Okay, so we have the second equation,

so now I can write this as three and I can write this as four, so what will we have from 3 and four, it will give us c0 and c1, from there we will have the polynomial of first degree we will get it and this polynomial will be the best polynomial, linear polynomial for this function x cube. So, in this way we did approximations of this function.

(Refer Slide Time 40:50)

Similarly, we can do it in quadratic, so what will happen in quadratic, what will be our p x will be c0 c1 x c2 x square. We will go up to here. So, basis function we are taking the standard bases which we had taken, now we have to calculate it, so you see in this my i will be c0 c1 c2. This will be the function and if we see, this value remains, we have f x f x from zero, to 1 which is our x cube minus c0 minus c1 x minus c2 x square d x. So, this is what we have its square and under the root, so we know, so we took its square, we took the approximations, now you know that if I have to find its maxima minimum, then for this we have del i by del c0, del i by del c1, partial derivatives del I by del c1, del i by del c2. This should be zero. So now if we calculate this, we will get three equations. So, in this case, what we have is, we will call it normal equation. So, if we calculate this, this normal equation is now if you see, it will come like this, see zero to 1, you can directly write it, it will become x cube minus c0 minus c1x minus c2 x square d x equal 0. The first one is this, the second one will come 0 to 1 x cube minus c0 minus c1x minus c2x square and with respect to c1, what is its coefficient x is so x d x equals 0 and the third one is okay, the second and third one will be what will be the third one, 0 to 1 x cube minus c0 minus c1 x minus c2 x square and what is here, coefficient of c2 x square so x square d x equal to 0 okay, so from this we will get three equations and three of us are unknown c0, c1, c2. So, if we calculate this, from here the values that we have are from here we will get the values of c0, c2 from all the three equations.

from ③ & ④ $\Rightarrow$ $\boxed{c_0, c_1}$ $\Rightarrow$ $\boxed{P(x) = c_0 + c_1 x}$

<u>Similarly</u> :- Quadratic $P(x) = c_0 + c_1 x + c_2 x^2$

$$I(c_0, c_1, c_2) = \int_0^1 \left(x^2 - c_0 - c_1 x - c_2 x^2 \right)^2 dx$$

$$\frac{\partial I}{\partial c_0} = 0 \qquad \frac{\partial I}{\partial c_1} = 0 \qquad \frac{\partial I}{\partial c_2} = 0$$

normal Equations

$$\int_0^1 (x^2 - c_0 - c_1 x - c_2 x^2)\, dx = 0 \; ; \qquad \int_0^1 (x^2 - c_0 - c_1 x - c_2 x^2)\, x\, dx = 0 \; ;$$

$$\int_0^1 (x^2 - c_0 - c_1 x - c_2 x^2)\, x^2\, dx = 0$$

NPTEL

(Refer Slide Time 43:40)

 Here, this means that it is a normal equation, its unknowns will come and from there we will find out its value. So, if we see here, we are doing interpolation, we are not caring about whether the polynomial that we are taking and passing is not being done through a point, our only purpose is that our error should be the least, okay. So, this continuous function that we had, with the help of this we converted it into an approximation. Now if we want to do the same thing, now look, we have taken this example, it was in the form of continuous. Now, suppose someone told us that approximate are of x, but now what we have to do is, we are not going to get the value of the function, or we are not going to get the value of it, so x1 is our zero, x2 is our point 2, let us see what value I will take, okay. So, let's take x3 .5, let's take x4 .8, let's take x5 this value 1, what is the value of the function in this, f x 0 .2 gives .008, so this means f at x1, this is f at x2, f at x3 is the value .125 f at x 4 we have, so we get .512, at x4 comes to .512 and fx5 is one, so now we have these five points given and beyond that the value of the function is given, okay, so now what we have to do is to find the best approximate fit line, we have to find the best fit quadratic, okay, what does best fit mean that the error should be minimum, that is what we will call the best fit line or best fit quadratic curve, okay, so what will we do in this case, now suppose I want to find a linear, okay linear means best fit line and I am taking the weight function as one in this case also, so what did we do, our polynomial p x is the same, c0 plus c1 x, okay? So now we have this given, which means to find it out. So now see what to do in this. Now we have a function, but we do not know the function, but we know its value, so what would p at x1 be, on zero, what would be c0 plus c1 into 0 p at x2, what would be c0 plus .2 c1, p at x3, what would be c0 .5c1, okay, like this my p at x4 is c0 .8 c1 and p x5 is c0 into c1. Now the values that we find out are not necessary that they are equal to the function, so the interpolating polynomial used to do that, right, we do not know about it, okay, so its values, so we can say from here that if we want to find the total error, then what would it be, f at x1 minus p x1, I will square it plus f at x2

minus p x2, I will square it and in the end the whole square of f at x 5 minus p at x 5, okay and if I take its under root, then we have this if we see then this will become f the function we have minus p x, okay its Euclidean norm, 2 norm to least depends on the data value of this x i, so it will be like this, okay so if we see from here then the values that are being formed with us, if I write p x1 like this, if I write p x2 like this, then if we write it directly then we will get a function, there will be a function of c0 and c1 and what will be the summation this is f x I so x i minus c0 plus c1 x i okay because we are taking x i how much is it from one to 5 here okay and this is square so we have squared it so we have a function what is this function basically this is error if you see then this is total error okay we have seen at every point so whatever difference comes between them we will call it error definitely so this is an error we have got so if We got an error, so now we have to find it out.

(Refer Slide Time 50:30)

So, to find it out we will do the same thing, del i c0 c1, we will take its partial derivative zero and del i by del c1, we will take its partial derivative zero. So, these two equations that will come, we will call them normal equations. Right, this is a normal equation. So, if we see in this, what will happen in respect of c0, in the same way two will come ahead and the coefficient will come minus. So, we can directly write in it from here that equal to two will be f x i minus c0 plus c1 x I is equal to 0. Okay, summation of all. So, what will we get from here, we will get summation f i f i. I am writing it, okay, so whatever we have written, I am writing f i, so this is minus. Now see, if I take it inside the summation, then this will come, summation c0 minus summation c1 x i obviously, there are as many points as there are, so there are points up to five in it, so we have i which is one to five and I 1 to 5 now see in this we took the number of points, five points and we would have taken its approximations by Linear, if we would have done by interpolating polynomial, we would have got a polynomial of 4 degree, we would have done Newton, Lagrange's so what is happening here, here we are doing its linear approximation, now see, we are not doing it piece wise, okay, like we had

taken piece wise polynomials or splines, it is not happening like that, we have tried to approximate the whole data with one line, so the equation that we will get from here, if I see, it will be c0 summation and 5 times summation, so I would write it as c1 summation all x i and that becomes equal to summation f i so given

If I want to find the second equation, it will come as c1, okay, zero, so if we see here, it will be multiplied by x i, so in this case, if we see, the second normal equation will come as f x i, okay, minus c0 minus c1 I x i in x i equals 0, it will become the same thing that we did above, in integration. We did this when we were working with continuity and when our discrete data arrives, it will come to us like this, so if we do it from here, you will see what will come out, c0 will be multiplied by x i, so this will come out c0 summation x i plus c1 now x i will become so summation x I square and x i on the right side, so this will become summation x i f i, so see from here we have two equations, okay 2 cross two, so here we have a cross two system, what will I do, I will solve it, so this summation one which is 5 times any time, however many there are, summation x i summation x i summation x i square c0 c1 equal to summation f i summation x i f i this comes, so this system, this value is known, we know this, we also know this, we also know this, we also know this, so from here our system is formed, I will name it suppose, I will name it S, okay, so the system we have, we will solve it. Okay, so the matrix that we have, suppose I name it X, I name it C and I name it f. Let's name it f. So, if we look at the matrix X, it is a symmetric matrix. So, we have a symmetric matrix X and the rank of this matrix will always be 2, so it means it will be an invertible matrix. So, from here we can find out the inverse of C, because this is a two by two matrix. So obviously these rows are linearly independent, so its rank is two. So, we will get the solution c, so from here the value of c0 c1 will come out. So similarly, we have done this for discrete data. Similarly, if I have a second-degree polynomial, then if I approximate it to c0 c1 x c2 x

square, then we will have it like this. So, what will happen in that case, the error function that we have will be the function c0 c1 c2 on top of these. So if I am taking c0 c1 c2, then if we see from here, it will become summation over i i so from one to n, okay, f i minus c0 c1 c2 c1 x i c2 x i's square, its square becomes our error, total error, now on its base we have to find the normal equation, so del i by del c0 equals 0 del i by del c1 equals 0 del i by del c2 equals 0, so it will come out like this.

So now look, from here we can write the normal equation directly, okay, now we know how to find out, so if we want to find out the normal equation from these three, then we will get the first normal equation summation one c0 plus c1 summation x i, okay plus c2 summation x i's square, that will become equal to summation f i, so the first equation becomes this, the second equation will come to by differentiating with respect to c1, and multiply by x I so this becomes. See what I do to remember, now I multiply the 1 st equation by x, c0 summation x i plus c1 summation x i square plus c2 summation x I square and this becomes our x I f i. Now again we will divide by c2 so multiply by x square so it means divide it once more by x then we will get c0 summation x I square c1 summation x i cube okay and c2 summation x i to the power 4 and this comes summation x I square f i this comes to us now we have three equations and so if we see then the system that we have, if I want to write it then it will become summation one, summation x I, summation x i square, summation x I, summation x I square, summation x i cube, summation x i square, summation x i cube, summation x to the power 4 and this will become c0 c1 c2 and this will come to us on the side summation f i summation x I f i summation x I square f i so this matrix that we have formed now see this is also a symmetric matrix okay now it is a symmetric matrix now this is If we solve it, then I will also call it X C equal F ok so now it depends on what is its rank ok so if we see then we will solve it its rank will come 3 and then we have the unique best approximation is unique so from here we will find the value of C so that our C comes so from here our normal equation

comes this normal equation ok so we have normal equation so we solved it so see if you pay a little attention to this work if we do regression in statistics then the regression line is this which we have defined here linear ok and the second degree regression line is y so basically regression is also done in this way so now from here we will get best fit line or best fit quadratic equation okay.

(Refer Slide Time 01:01:37)

So, in this case the best fit line we got we can take cubic best fit We can define it to any degree, okay, and in this case, we have defined it in such a way that the error we have is minimal. So, if the error is minimal, then what does it mean that the approximation will be the best approximation, and it is independent of the number of our data points. Even if you have 100 data points, you would still get an approximation from a linear best fit line. So, thanks for looking at this. Hello h