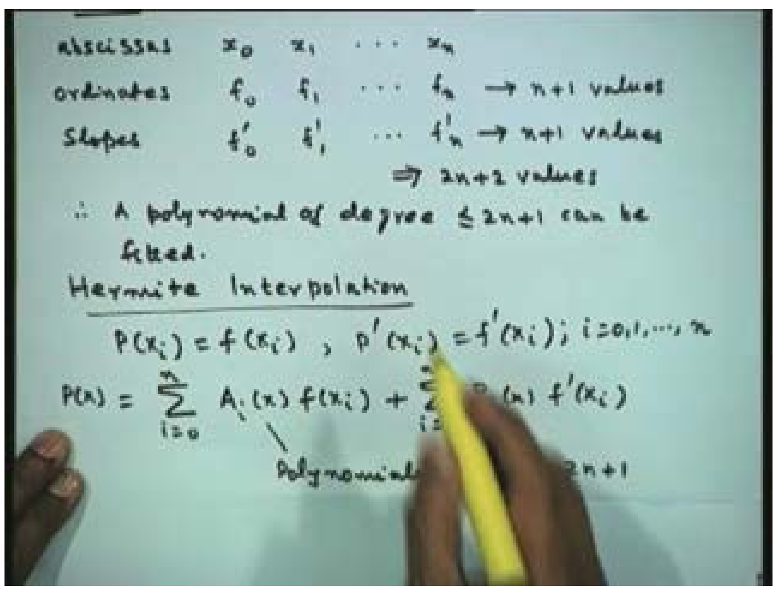
Numerical Methods and Computation
Prof. S.R.K. Iyengar
Department of Mathematics
Indian Institute of Technology Delhi

Lecture No - 30
Interpolation and Approximation (Continued)

In our previous lectures we have derived various forms of interpolating polynomials to fit a given data which consists of an abscissa and the corresponding ordinate at n plus 1 points. Now we derived the various forms in the sense we have Lagrange interpolation, divided difference interpolation and if it as a equispaced data we have the Newton's formulas of backward forward formulas and if now, if you add for this data one more item like the slope of the function that is representing the given data, then we can have a different type of polynomial all together, so let us now consider data of this particular form.

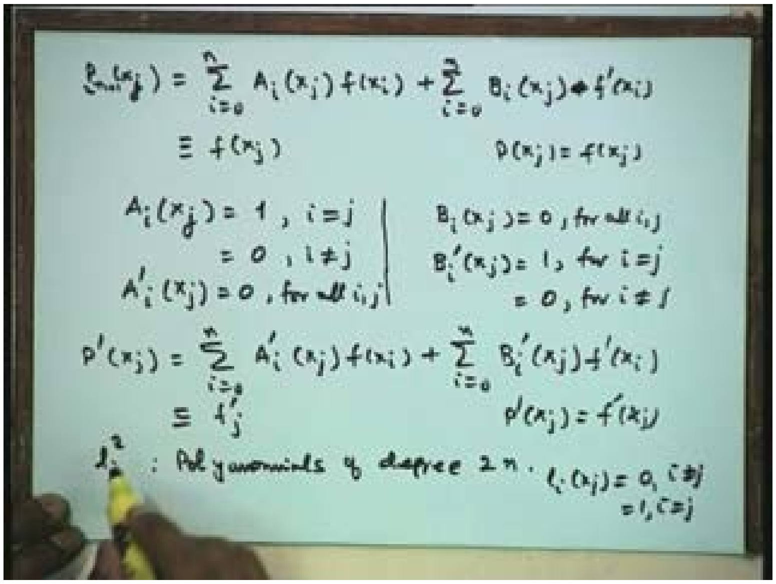
(Refer Slide Time: 01:42)



The data that is given to us shall be the abscissas that we take it as x_i, this this is your x_0, x_1, so on x_n, then we have the ordinates given to us as f_0, f_1, f_n. Now this was the data that we considered earlier, let us now add to this slope at these points also, so let us also consider the case when we have slopes also at this, which we shall write it as f prime at 0, f prime at 1, so on f prime at n. Now what type of interpolating polynomial can we derive for this particular data, now if you look at this number of ordinates that we have, we have n plus 1 values of this ordinates. Now if you consider simply the data x_i f_i that is this data, then we were able to write down a polynomial of degree less than is equal to n which fits this data exactly. Now we have added to it n plus 1 more values, these are again n plus more values. Therefore this implies that we have a total of 2 n plus 2 values that are available to us, therefore in this data we have 2 n plus 2 values available to us.

Hence this implies that we can fit a polynomial of degree less than is equal 2 n plus 1 for this data, therefore a polynomial of degree less than or equal to 2 n plus 1 can be fitted and the polynomial which does this job, we shall call it as Hermite interpolation, so we shall call it as Hermite interpolation. That means we want to construct a polynomial P, such that p at x_i is equal to f at x_i and P prime at x_i is f prime at x_i for i is equal to 0 1 so on n. Now we shall follow the procedure that we have adopted in deriving the Lagrange interpolation, now the interpolating polynomial must be, now a linear combination of these ordinates and these slopes also, that means I must be able to write down the polynomial P(x) is equal to summation of i is equal to 0 to n, some $A_i(x)$ f of x_i plus summation 0 to n $B_i(x)$ f prime at x_i, so the polynomial should be a linear combination of these ordinates, n plus 1 ordinates and n plus 1 slopes. Now since this polynomial is of degree than is equal to 2 n plus 1, $A_i(x)$ and B(x) must be polynomials of degree 2 n plus 1, these are polynomials of degree 2 n plus 1. Since this polynomial should fit this data exactly, we can find out the conditions under which this approximation is possible for us.

(Refer Slide Time: 05:47)



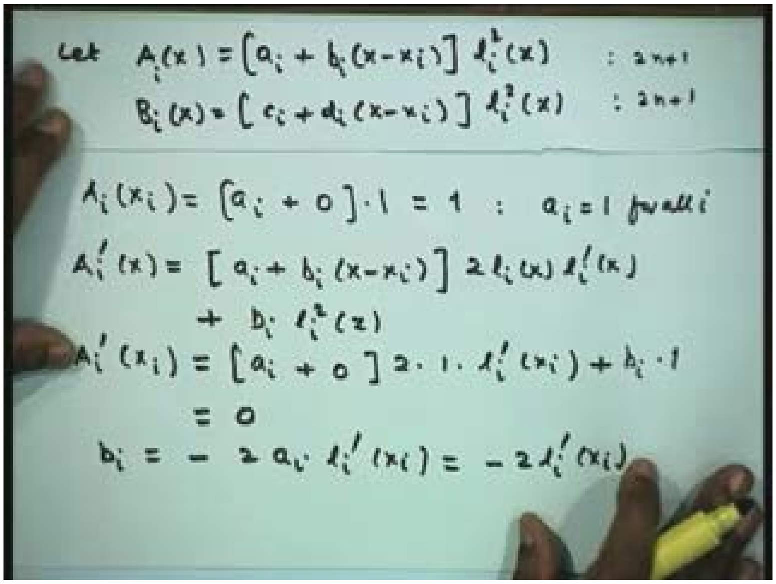
So let us substitute x_i here and see what we would get the conditions on a_i and b_i, some x_j, let us put x_j, i is equal to 0 to n, $A_i(x_j)$ into $f(x_i)$ plus summation i is 0 to n, $B_i(x_j)$ plus into f prime of x_i. Now this should be identically equal to, this is $P_{2n+1,}$ suffix $P_{2n+1,}$ this should be identically be equal to f of x_j, that is your P of x_j is equal to f of x_j, that is the data given to us. Now if this is to be true it should be identically, now all this $B_i(x_j)$ must be 0 because there is no derivative here, so let us put it here in two ways $B_i(x_j)$ is equal to 0 for all i j, $B_i(x_j)$ must be equal to 0 for all i j.

You write it in the 2 columns, so we shall fill up the remaining data over here. Now let us look at $A_i(x_j)$, this should be equal to $f(x_j)$ therefore all of them would be 0 except when the both suffixes are same, so you will have here $A_i(x_i)$ is equal to 1 $A_i(x_j)$ is equal to 1, for i is equal to j, is equal to 0 for i not equal to j. Then this will give us simply 1 into f of x_j so it will be identically equal to this.

Now we shall fill up here, I will just leave some space over here. Let us differentiate this and set x_j here, so I am differentiating this particular polynomial and then to it, so it will give us derivative of A_i and here it will give derivative of B_i, these are constants so only derivative of A_i and B_i will come, so that we can write it as summation i is equal to 0 to n, A_i prime x_j $f(x_i)$ plus summation i is 0 to n, B_i prime x_j f prime at x_i. Now you can see that this should be identically equal to f prime of x_j because the interpolation condition is P dash x_j is equal to f prime of x_j. Therefore this should be identically equal to this, therefore this implies that there cannot be any ordinates here that means this A_i prime x_j will be 0 for all i j, so that means I can now write here A prime $i(x_j)$ is equal to 0 for all i j. Whether it is equal or not equal, in all the cases the value of A prime x_j will be equal to 0 and B prime x_j will be equal to 1 when i is equal to j and it will be zero for i not equal to j because this should produce f prime at x_j. Therefore we will have here B_i prime x_j is equal to 1 for i is equal to j, is equal to 0 for i not equal to j.

Now we need to construct, now A_i and B_i just as we have done in the Lagrange interpolation by looking at the property of A_i and its derivative, B_i and its derivative. Now we remember that A_i and B_i are polynomials of degree 2 n plus 1, so what we will do it, we shall take advantage of the Lagrange fundamental polynomials, which are polynomials of degree n. So if I consider l_i square, this will be polynomial of degree 2 n, polynomials of degree 2 n and this satisfies our properties that $l_i(x_j)$ is equal to 1 for i is equal to j, $l_i(x_j)$ is equal to 0 for, we know this property that $l_i(x_j)$ is equal to 0 for i not equal to j, is equal to 1 for i equal to j. Because of this property which is inbuilt for A construction of this, I shall take, use this l_i square Lagrange fundamental polynomials in building this A_i and B_i.

(Refer Slide Time: 11:07)



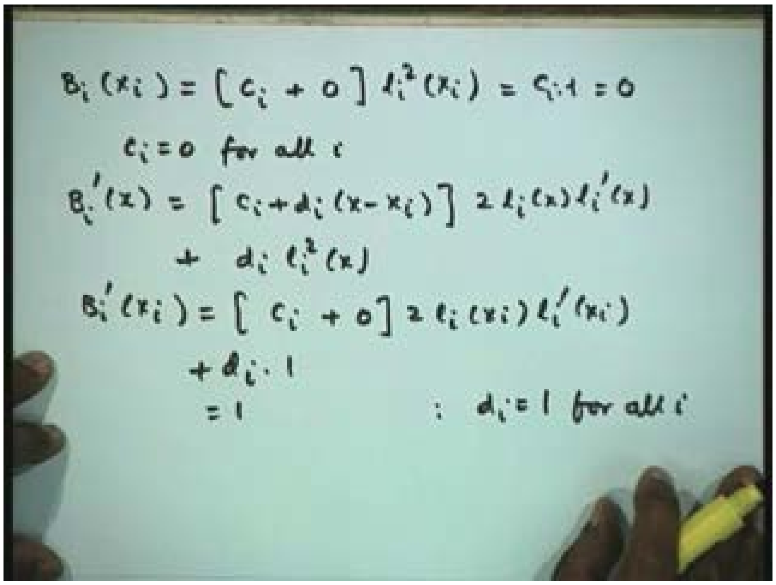
So what I would do since l_i square is a polynomial of degree 2 n and A_i B_i are only polynomials of degree 2 n plus 1, I need to multiply this only by linear polynomial that means we shall assume that let A(x) is equal to suffix I, some suffix i you put it, [a_i plus b_i into (x minus x_i)] l_i square x and B_i(x) is equal to some c_i d_i (x minus x_i) l_i square x. Now l_i square is a polynomial of degree 2 n, I am now multiplying this by linear polynomial of the special form which I have taken it in this particular form because it is easy for us to find the constants in that case. Now this is a polynomial of degree 2 n plus 1 and this is also a polynomial of degree 2 n plus 1.

If we are able to find uniquely a_i, b_i, c_i, d_i from the data that we have, then we have what we have written in the formula is correct. So let us try to find that one, let us look at, let us just put this condition over here. We shall now determine the parameters a_i, b_i, c_i and d_i using the conditions on a_i and b_i which we have obtain earlier as this, where $A_i(x_j)$ is equal to 1 for i is equal to j, it is 0 for i not equal to j. Similarly A_i prime x_j is equal to 0 for all i and j and similarly the conditions on B_i. We shall apply these conditions one after the other to determine the constants a_i, b_i, c_i and d_i. Let us first of all substitute x_i in $A_i(x)$, so if I substitute x is equal to x_i, I get here a_i plus 0, this is 0, then I get $l_i(x_i)$ is 1 therefore $l_i(x)$ square is equal to 1 and this should be equal to 1, a_i is equal to 1 for all i. Now we have determined one of the parameters a_i in this capital A_i. Now let us differentiate A_i therefore I will get A_i prime of x is equal to, it is a product of 2 functions, so let us write this product as [a_i plus b_i into (x minus x_i)] derivative of this is 2 times $l_i(x)$ l_i prime x plus the derivative of this gives us simply b_i into l_i square of x. Now we have the condition that A_i prime x_j is equal to 0 for all I j.

Let us now substitute and see here, if i put A_i prime x_i here, I would get here is a_i plus, x is equal to x_i that is 0, then I will have here 2 times $l_i(x_i)$ is equal to 1, I will have l_i prime of x_i plus $l_i(x_i)$

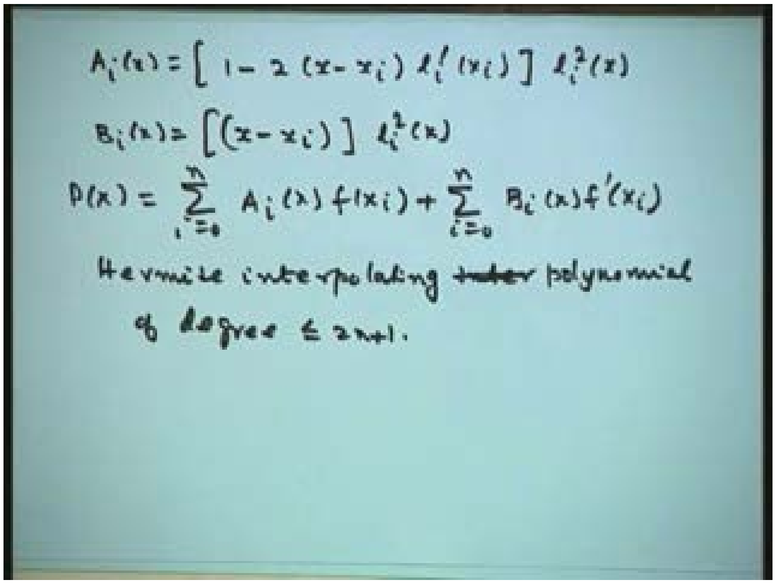
is equal to 1 therefore I will have here b_i into 1 and the value of A_i prime x_i is equal to 0 that is the condition that we have here, that A_i prime x_j is equal to zero for all i and j. Therefore I can determine b_i from here therefore I will have b_i is equal to minus 2 times a_i into l_i prime of x_i but a_i is equal to 1 therefore I will get 2 times l_i prime of x_i. Now similarly we shall apply the conditions on b_i to determine the constant c_i and d_i.

(Refer Slide Time: 15:22)



Now let us substitute x is equal to x_i in B_i of x, so if I have put $B_i(x_i)$ that gives me, I am now substituting here in this the, I am substituting x is equal to x_i in this, so let us keep this slide here, this is [c_i plus 0] l_i square of x_i and this is equal to 1 therefore this gives us c_i into 1 but this is equal to 0, this is equal to 0. Therefore we get c_i is equal to 0 for all i. Then let us differentiate $B_i(x)$, let us differentiate $B_i(x)$ from here, so I will have here B_i prime of x is equal to [c_i plus d_i into (x minus x_i)] derivative of l_i square is 2 times $l_i(x)$ l_i prime of x plus derivative of the first one gives us d_i into l_i square of x. Now use the condition that B_i prime x_i is equal to 1, so if I put B_i prime of x_i here, I get here c_i plus this is 0, that is 2 times $l_i(x_i)$ l_i prime of x_i plus d_i into 1. Now B_i prime of x_i is equal to 1 therefore this is equal to 1. Now c_i is equal to 0 therefore this bracket goes off and we have simply d_i therefore we obtain d_i is equal to 1 for all i. Now we have determined all the 4 constants. So we can now substitute in the expression for $A_i(x)$ and $B_i(x)$ that is this expression $A_i(x)$ and this.

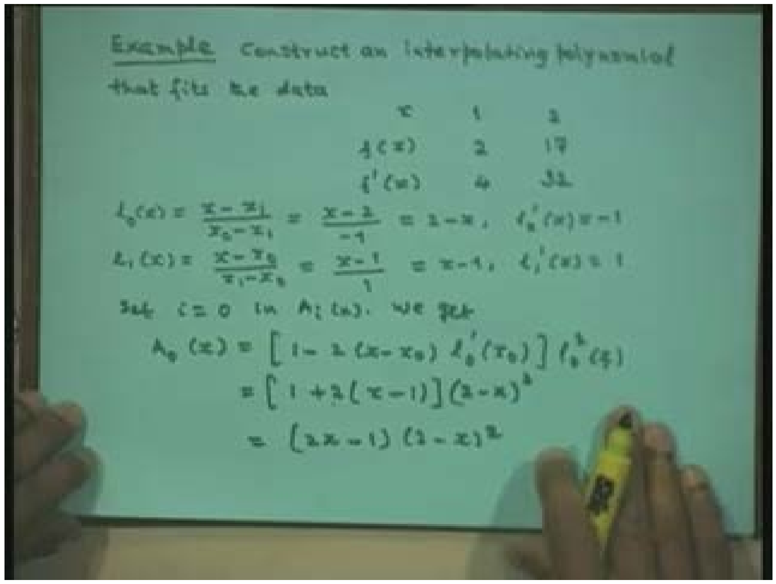
(Refer Slide Time: 17:43)



Let us substitute it and see what we get, we get $A_i(x)$ is equal to, is equal to A [1 minus 2 times (x minus x_i) into l_i prime of x_i] into l_i square of x. Now this we have obtained it because we have got here a_i is equal to 1 and we have obtained b_i is equal to minus 2 times a_i l_i prime of this x_i, so we have substituted for b_i and this is (x minus x_i) over here, therefore this is the expression in the brackets for $A_i(x)$ and outside the bracket we have l_i square x. Similarly we get for $B_i(x)$, this is equal to, now c_0 is 0 and d_i is equal to 1 therefore I simply get (x minus x_i) into l_i square of x and therefore for the required polynomial, P(x) is equal to summation of i is equal to 0 to n $A_i(x)$ f of x_i plus summation i is equal to 0 to n $B_i(x)$ of f prime of x_i.

We call this as the Hermite interpolating polynomial, we call this as the Hermite interpolating polynomial, interpolating polynomial, which is of degree less than or equal to 2 n plus 1. Now to compute this interpolating polynomial we need to determine $l_i(x)$, I need to determine l_i prime then substitute the values over here, determine my A(x) $B_i(x)$ from here, sum them up, simplify it to finally arrive at the polynomial of degree 2 n plus 1 or less than 2 n plus 1 and that represents the interpolating polynomial which fits exactly that given data.

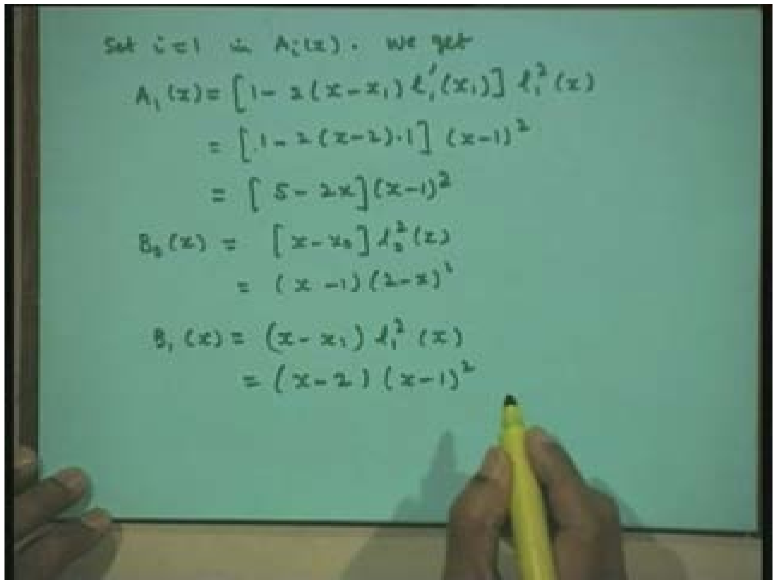
(Refer Slide Time: 20:23)



Now let us take an example on this, let me write the example. Now construct an interpolating polynomial that fits the data x, let us take only 2 points, f(x) is 2 values 2 ordinates and the slopes f prime x is 4 and 32. Now we need to first of all write down our Lagrange fundamental polynomials $l_i(x)$, there only 2 data points therefore we will have $l_0(x)$ is equal to (x minus x_1) upon (x_0 minus x_1) that is (x minus 2) divided by 1 minus 2 that is minus 1, which is your 2 minus x and we need the derivative also, let us differentiate it l_0 of x also, lets write down x also, that is equal to minus 1. Here it is a linear polynomial therefore derivative is a constant otherwise this would not be a constant, if we take more data points it will be a function of x.

Now let us write down $l_1(x)$, this is (x minus x_0) divided by (x_1 minus x_0) that is (x minus 1) divided by 1 that is your (x minus 1). We need its derivative also, let us differentiate it, this is equal to 1. Now we need to find the quantities A_0, A_1, B_0, B_1 to use this particular expression. Now we set i is equal to 0 in this to get $A_0(x)$, now we shall say, set i is equal to 0 in $A_i(x)$ then we get $A_0(x)$ is equal to, now we are setting i is equal to 0, therefore I will get here x_0, this is l_0 prime x_0, this is l_0 square x, so I will get here [1 minus 2 (x minus x_0) l_0 prime x_0] into l_0 square of x. Now I will substitute the values of l_0 prime x_0 is minus 1, $l_0(x)$ is equal to (2 minus x), therefore I will get here 1, this is negative sign so I will write it as plus 2 (x minus x_0) is 1, l_0 square is (2 minus x) whole square. We can simplify and write this as (2 x minus 1) into (2 minus x) whole square.

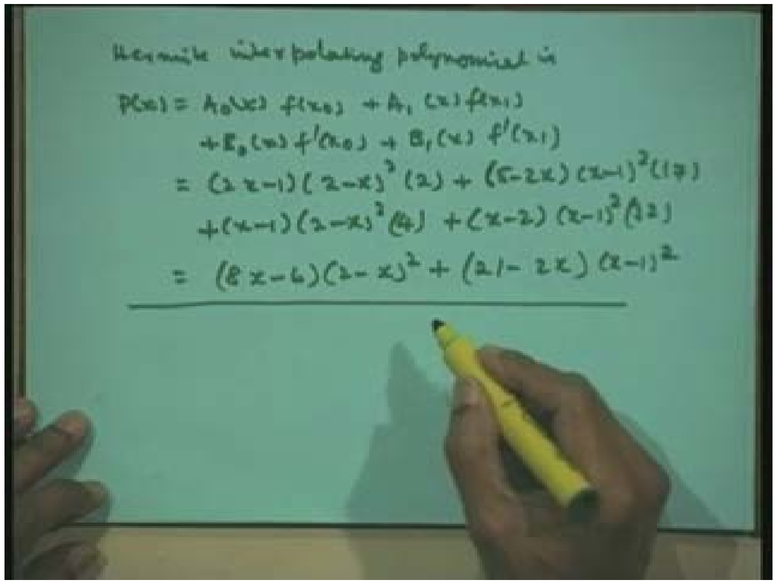
(Refer Slide Time: 23:51)



Now set i is equal to 1 in $A_i(x)$, we get, I am setting i is equal to 1 so I will have here x_1 l_1 prime x_1 l_1 square x, therefore I will get here $A_1(x)$ is [1 minus 2 times (x minus x_1) l_1 prime x_1] into l_1 square of x. Now x_1 is equal to 2, l_1 prime of x_1 is 1 and $l_1(x)$ is (x minus 1), therefore I get here 1 minus 2 into (x minus 2) l_1 prime of x_1 is 1, so I will have this as 1 and l square of x is (x minus 1) whole square, that is (x minus 1) whole square. Therefore this I will get it as 4 plus 1, [5 minus 2 x] into(x minus 1) whole square.

Now set i is equal to 0 in $B_i(x)$, therefore I will get $B_0(x)$ is equal to [x minus x_0] into l_0 square x. Again our l_0 is (2 minus x) therefore this will simply give us (x minus 1) into (2 minus x) whole square. Now set again i is equal to 1 in $B_i(x)$ therefore I will get B_1 is (x minus x_1) l_1 square x, therefore I get $B_1(x)$ is (x minus x_1) l_1 square of x, x_1 is 2 therefore I get, here I get x_1 is 2 and we have got here $l_1(x)$ is (x minus 1), therefore this is equal to (x minus 1) whole square. Now I got this 4 quantities and which we can substitute now in P(x) to get our polynomial.

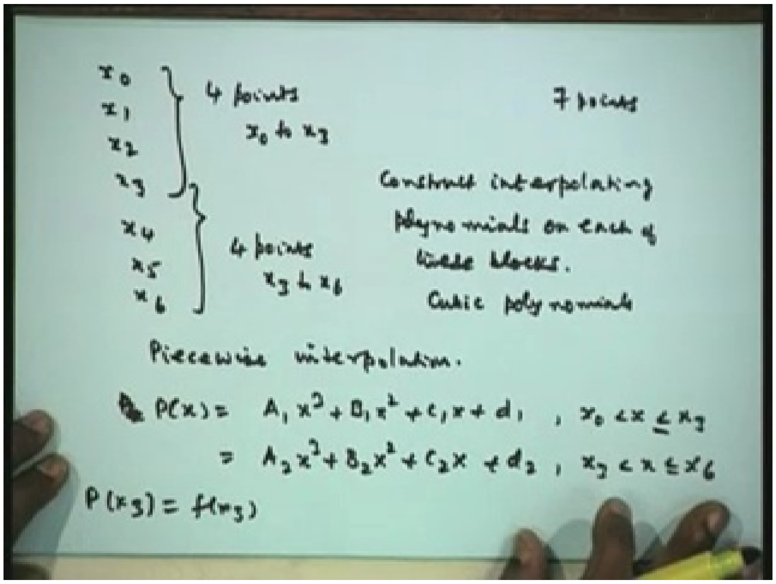
(Refer Slide Time: 26:25)



Therefore the required Hermite interpolating polynomial is, P(x) is equal to $A_0(x)$ f of x_0 plus $A_1(x)$ f of x_1 plus $B_0(x)$ f prime of x_0 plus $B_1(x)$ f prime of x_1. Now we can substitute the values that we have obtained earlier that gives us (2 x minus 1) into (2 minus x) whole square into 2 plus (5 minus 2 x) into (x minus 1) whole square into 17 plus (x minus 1) into (2 minus x) whole square into 4 plus (x minus 2) into (x minus 1) whole square into 32. Now we can simplify it and we get the result as (8 x minus 6) into (2 minus x) whole square plus (21 minus 2 x) into (x minus 1) whole square. Now you can easily verify that this fits our data exactly.

Now before we proceed further we should make some comments or a little word of caution on using the interpolating polynomials. If a large data is given say n plus 1 points, we can construct a polynomial of degree less than or equal to n to fit the data but the data that we have got is usually from an experiments or from some observations, therefore if you are given a, say 4 plus accuracy table, the last digit of this is always due to round off errors because you are rounding it off is there. Now if you are constructing a polynomial of degree 99 say for the data given as 100 data is given, if you are constructing polynomial of degree 99, we are now using the multiplications of these ordinates by numbers which are the coefficients in the polynomial. Therefore the round off error that is there in each data item gets multiplied and cumulatively the total round off error will be enormous and no experiment you can give the result exactly, therefore it is not advised and it is not used to construct higher order degree interpolating polynomial even though the data is very very large. The only alternative would be, to at the most go up to cubic or the forth degree polynomial beyond that we do not go and hence it is possible for us to break the data into blocks, say for example if you have got, suppose we have got a data of say at the 6 points, well let us write down this data.

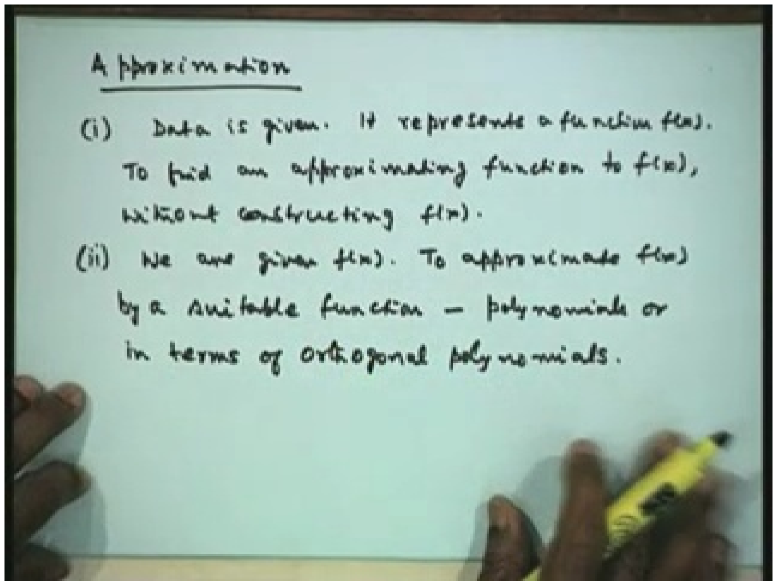
(Refer Slide Time: 29:54)



Suppose I have given the data is the seven points are given, I know that I can construct the polynomial of degree 6; if slopes are given we can construct the polynomial of degree 13 but what we are stating is in practice we should not use that. What we use it, for example I can break this into two blocks of these points, so this contains 4 points and this contains 4 points, continuously x_0 to x_3, this is your x_0 to x_3, this is x_3 to x_6. Then I construct interpolating polynomial on each of these blocks, so construct interpolating polynomials on each of these blocks. Now in this case they are cubic polynomials, in this case they are cubic polynomials. Then we shall call such interpolation as piecewise interpolation, so we shall call this as piecewise interpolation.

It has all the properties of the interpolating polynomial that we have discussed here except that the entire data is being made into small blocks and on each of them we are constructing a cubic polynomial, for example four data points, if we are taking three data points we will have a quadratic, quadratic, quadratic. Therefore we are going to write down the interpolating polynomial as something like, for example here I would write this as P(x), P(x) is equal to some A_1(x) cubed plus B_1(x) squared plus c_1(x) plus d_1, for x_0 less than x less than x_3 and I can put equal to because we have taken the point in both it is, both of them are fitting exactly at x_3 therefore continuity at x_3 is available for us. This is some A_2(x) cubed plus B_2(x) square plus c_2(x) plus d_2, x_3 is less than x less than or equal to x_6, either put equal to here, equal to here, because both of them are going to be the same thing. That of course that P at x_3 is equal to $f(x_3)$, so both of them would satisfy this. Therefore in this way we can give for the entire data, we can say that for this first block this is the polynomial, second block this is the polynomial, third block this is the polynomial and then use this piecewise interpolating polynomials for predicting the data values at any of the intermediate points. This is what we normally do in practice if you have a very huge data. Even though theoretically we construct very high degree polynomial but in practice we shall not be using it. Now let us look at the second problem in this that is your approximation.

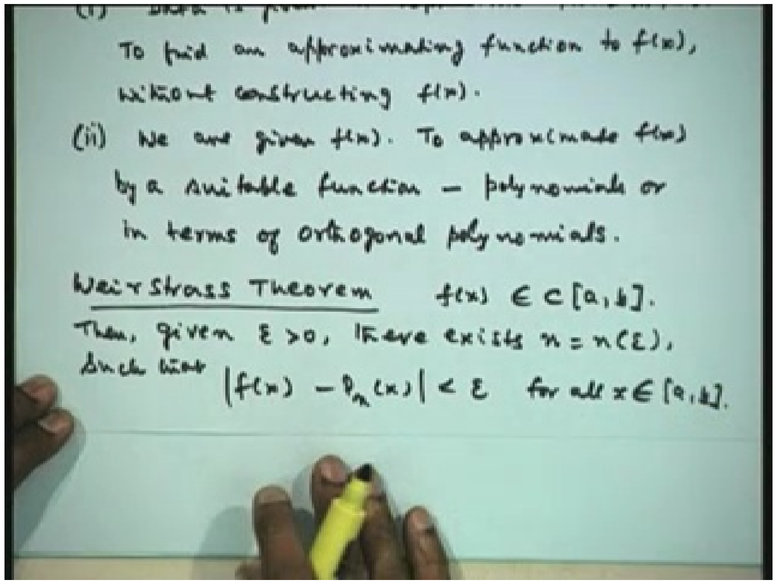
(Refer Slide Time: 33:23)



Now we mentioned earlier that the problem of approximation is of two type, one type is that data is given to us, data is given, we know that it represents a function f(x), it represents a function f(x) but without actually constructing this function f(x), we would like to write down an approximating polynomial or a function which approximate f(x) for this given data. Now the problem is therefore to find an approximating function, approximating function to f(x) without constructing f(x) that means we would like to construct the function without actually going through the process of interpolation.

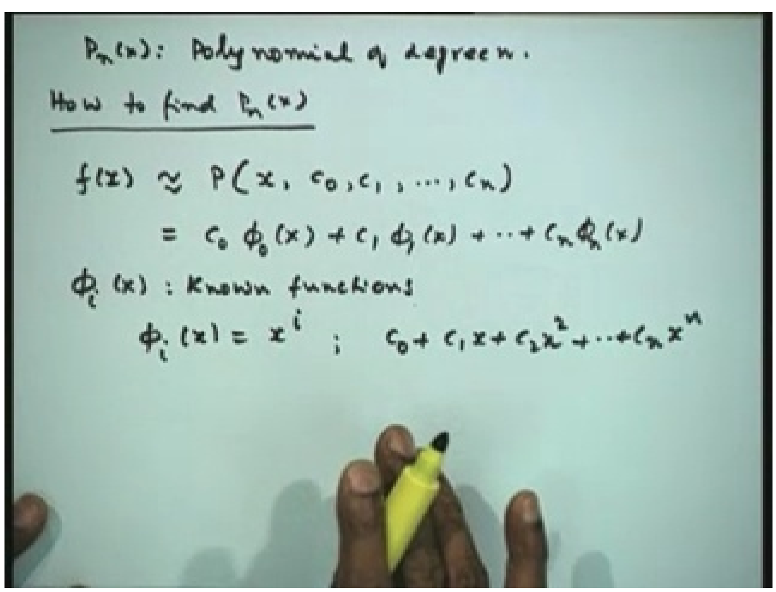
Now the second is that we are given a function f(x), here we are given f(x), if the problem is to approximate f(x) by a suitable function, what is this function? It could be polynomials or orthogonal polynomials; that is they are functions of polynomials or in terms of orthogonal polynomials. So that the properties of the function given to us, which was a complicated function can be studied through these polynomials because when once we write it in terms of an orthogonal polynomial, we know all the properties of the orthogonal polynomial implied there and we can use those properties to say about the behavior or any other property that we need of the given function f(x). Now if I want to construct these two, first of all we must guarantee ourselves that such a representation in terms of polynomials or in terms of orthogonal polynomials is guaranteed for us otherwise we are not sure whether what we are obtaining is correct or not, the answer for this is, yes we have a theorem called Weir Strass theorem, which states that if we have a continuous function over internal [a, b] then we can always approximate it by a polynomial. So that is known as Weir Strass theorem, which guarantees that what we are doing is correct and we will be able to get a unique polynomial from there.

(Refer Slide Time: 36:49)



Let us define what our Weir Strass theorem is. So we have a function which is continuous, belongs to the class of continuous function over the interval [a, b]. Then the theorem states, then given an epsilon greater than 0, there exists a number n which is a function of epsilon such that f(x) minus $P_n(x)$ is less than epsilon for all x contained in [a, b], where $P_n(x)$ is a polynomial of degree n, where $P_n(x)$ is a polynomial of degree n.

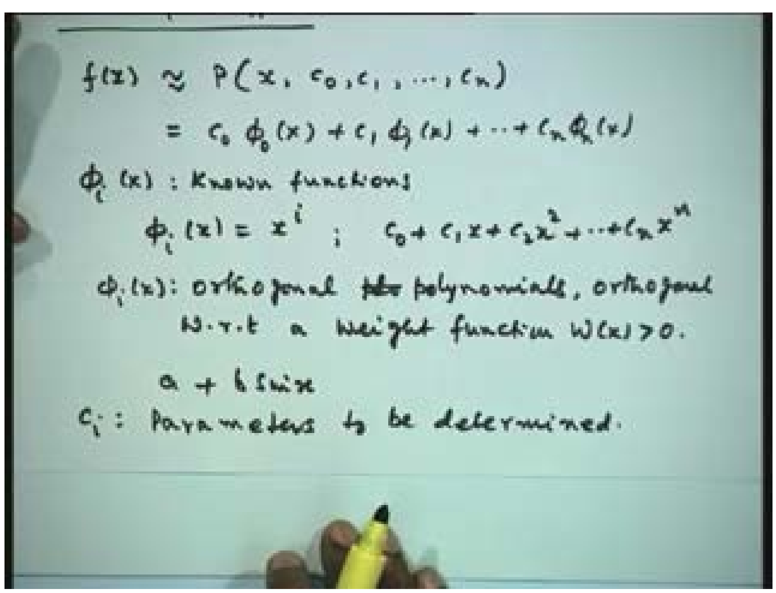
(Refer Slide Time: 37:47)



Therefore the Weir Strass theorem guarantees that we can approximate a continuous function by a polynomial of degree n, so we will take this assumption that it is possible. Then how to construct this particular function using this particular, the definition that f(x) minus $P_n(x)$ should be less than epsilon. Now the problem is therefore is to how to find $P_n(x)$. What we do is, we take $P_n(x)$ in terms of some function x that is polynomial or orthogonal polynomials whatever that we have given here, therefore f(x) will be approximating polynomial in terms of the variable x and we introduce some constants c_0, c_1, c_2, c_n, c_0 to c_n that means what we are essentially writing here is some c_0 phi_0 of x plus c_1 phi_1 of x plus so on c_n $phi_n(x)$. phi_i's are the known functions, these $phi_i(x)$ are the known functions, are the known functions, $phi_i(x)$ are the known functions.

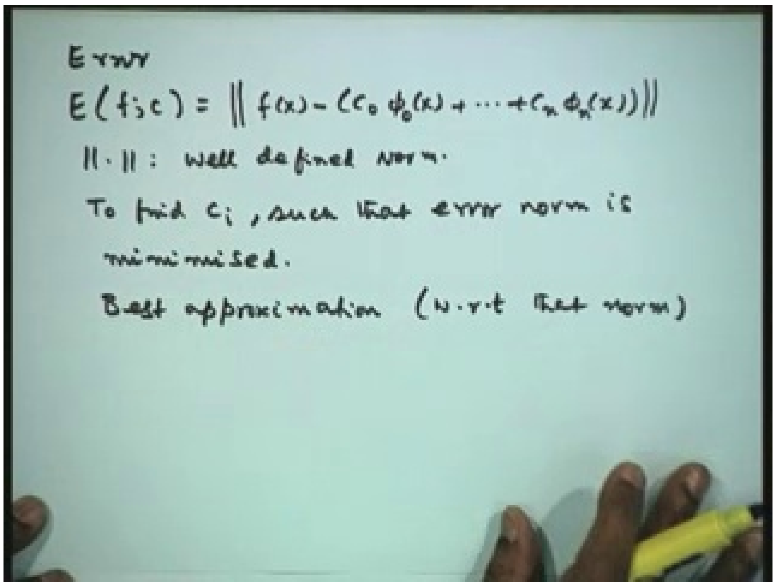
For example they can be taken as polynomials, I can simply take it as $phi_i(x)$ is equal to x to the power of i, I can take $phi_i(x)$ is equal to x to the power of i, that means phi_0 is 1, phi_1 is x, phi_2 is x square that means what I am really writing here is c_0 plus $c_1(x)$ plus $c_2(x)$ square plus so on $c_n(x)$ to the power of n that is simply a polynomial, that is what we have this or I can take $phi_i(x)$ as orthogonal polynomial, orthogonal polynomials, orthogonal polynomials which are orthogonal with respect to a weight function, orthogonal with respect to, with respect to a weight function. Now what we are really talking of is the Legendre polynomials and the Chebyshev polynomials. The Legendre polynomials are orthogonal with respect to 1; weight function is 1, whereas the Chebyshev polynomials are orthogonal with respect to weight function 1 upon under root 1 minus x square, that will come later on but we are just saying why we are introducing a w(x) here, because these orthogonal polynomials have a weight function over which they are orthogonal. Now further the phi(x) could be any other function which represents your experiment properly, for example if in an experiment the solution of the variable is in the form of sine or a cosine wave, we can write down this as, for example I could simply write the as some a plus b sin x.

(Refer Slide Time: 41:37)



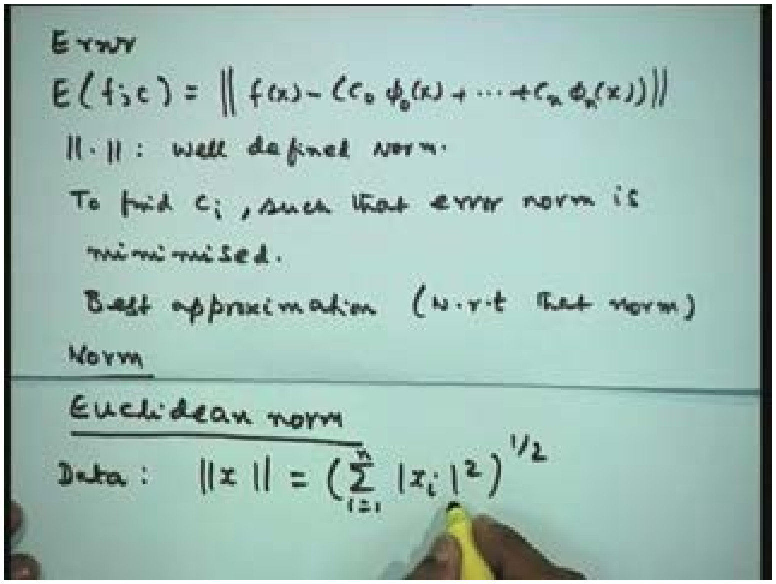
I can take an approximation for this as, a plus b sinx but I must be able to determine the parameters a and b. Now therefore these c_1, c_2, c_i, these c_i are parameters to be determined, parameters to be determined. Now we must give the rule under which these parameters can be determined, for that we shall consider the error that is coming from here.

(Refer Slide Time: 42:20)



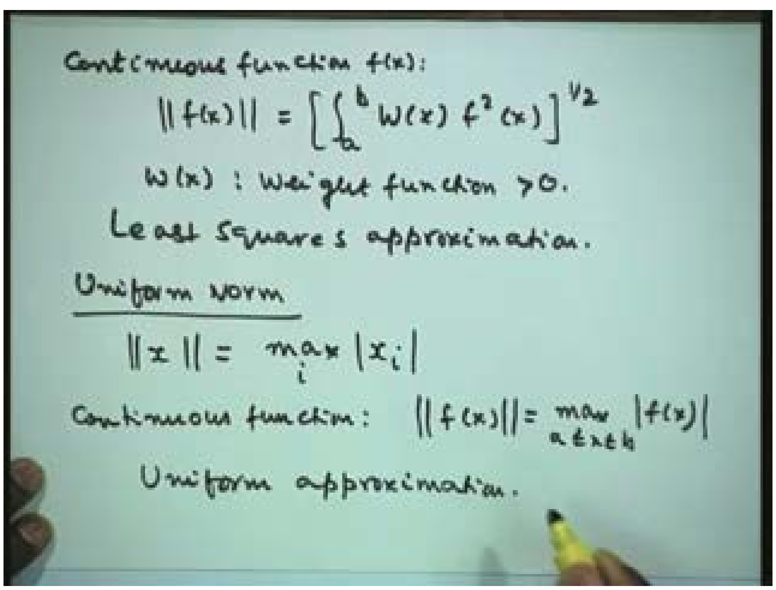
So we will take the error here, therefore the error, let us write down E of (f; c), c is the vector of this c_0, c_1, c_2, c_3 and we will take this as norm of f(x) minus (c_0 phi_0(x) c_n phi_n(x)), where this is a well defined norm, it is a one of the norms that we know, it is a well defined norm. Now therefore the problem reduces to how do you find this c_i, therefore the problem is to find c_i, to find c_i such that the error is minimized, the error should be smallest, to find c_i such that error norm is minimized. That particular approximation for which error is minimized that means that particular approximation, this approximation for which this error is minimized shall be called as the best approximation, so this shall be called as the best approximations. Of course we should qualify it as with respect to that norm which we have used, with respect to that norm. Now we have specified what we mean by this error, what are the things that we have to do, we have to find c_i such that the error norm is minimized. Now the next step we shall define is; what is norm? We have earlier given a number of definitions of norms of which we shall use two of the norms.

(Refer Slide Time: 44:32)



One norm that we shall use is the Euclidean norm, Euclidean norm. Now let us define what is this Euclidean norm, let us suppose we are given a data then we shall define this norm as, norm of x is equal to summation of magnitude of x_i square to the power of half. So we are taking the summation over all the elements, magnitude of x_i square whole to the power of half.

(Refer Slide Time: 45:13)



However if we are given a continuous function f(x), then we define the norm of f(x) is equal to integral of a to b w(x) f square of x to the power of half, where w(x) is the weight function that we are talking of earlier, is the weight function and which is greater than 0. So we can use this definition of Euclidean norm to obtain the values of the constants that we have just now defined in the error. In this case we would get what is known as the least square approximation; we obtain the least squares approximation, if I use this particular norm and determine the constants c_i such that this norm is minimized. Now we use another norm which we shall call it as uniform norm, uniform norm. Again if you are given a data, I would define norm of x is equal to the largest element in magnitude, maximum of x_i, of x_i and if you are given a continuous function again, if you are given a continuous function then we define this as, norm of f(x) is equal to maximum in the interval a to b of magnitude of f(x). Now we use this, either for the discrete data that is given or if you are given a continuous function, now if I use this particular norm to determine the constants c_i then what I would get is known as the uniform approximation, we get uniform approximation in this case. Now in our next lecture we shall see how we have actually apply this minimization problem that is the minimization of the norm to get least square of approximation or minimize these norms to get uniform approximation and thereby construct a polynomial or a function in terms of the orthogonal polynomials which gives us the best approximation gives us the best approximation. Okay, thank you.