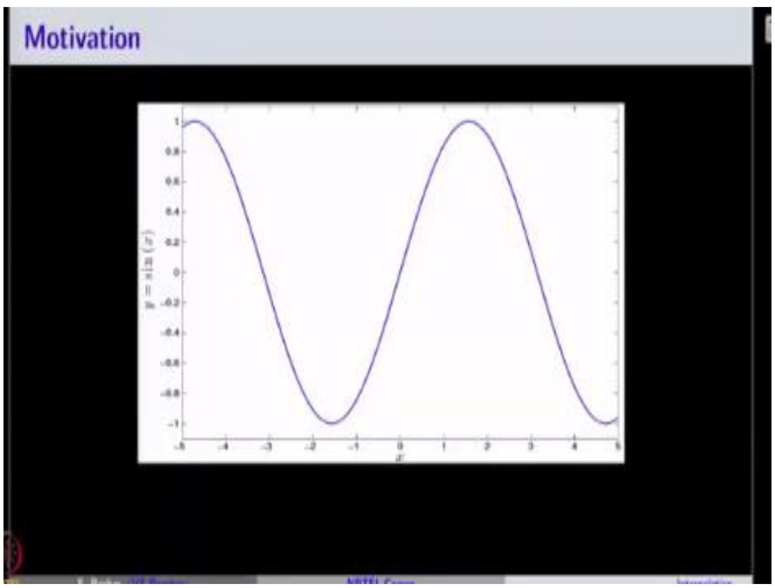
Numerical Analysis
Prof. S. Baskar
Department of Mathematics
Indian Institute of Technology, Bombay

Lecture - 38
Polynomial Interpolation: Existence and Uniqueness

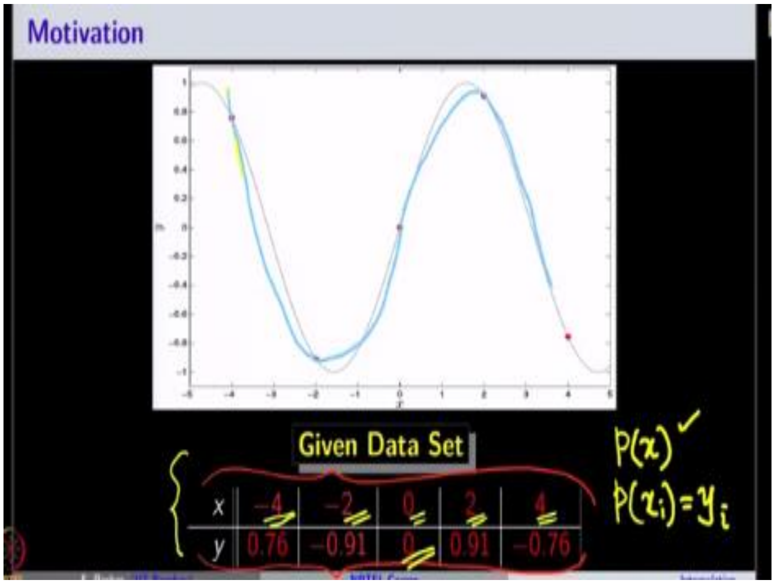
Hi, in this lecture we will start a new chapter on interpolation. We will be studying polynomial interpolations; in this lecture we will prove that there exists a unique polynomial interpolation for a given set of data. We will start our discussion with a small introduction to polynomial interpolation and also, we will see where it is useful.

(Refer Slide Time: 00:46)



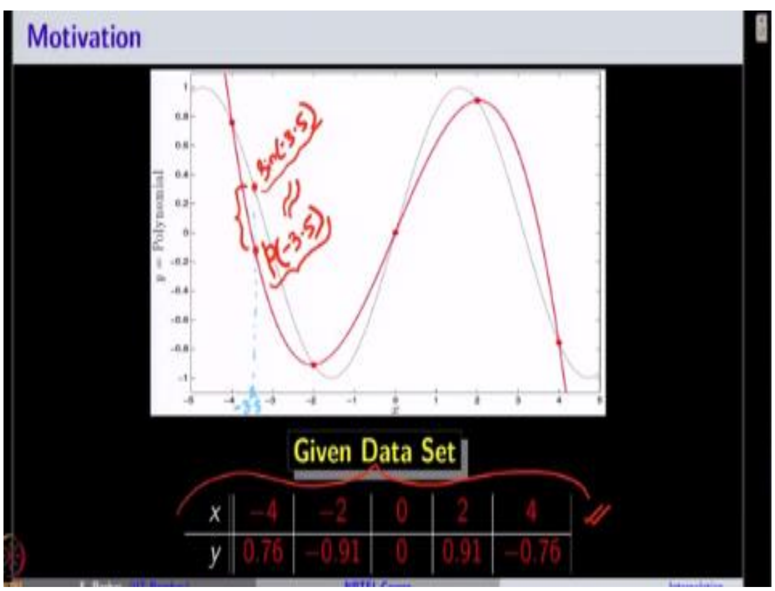
Let us consider the function $f(x) = sin\, x$. The graph of the function sin x is given in the solid blue line, the x-axis denotes the angle in radians and y-axis gives the sin value at x. Just imagine that we do not know what is the function given to us.

(Refer Slide Time: 01:17)



But we only know a set of x values along with the value of the function at these points. Now the question is, can I find a function say more precisely we are interested in a polynomial. The question is, can we find a polynomial such that $p(x_i) = y_i$ where x_i's or these points and y_i's are the corresponding values given in this table. We will call this table as data set and we are interested in finding a polynomial such that it passes through all these points that is something like this, that is our interest.

(Refer Slide Time: 02:16)

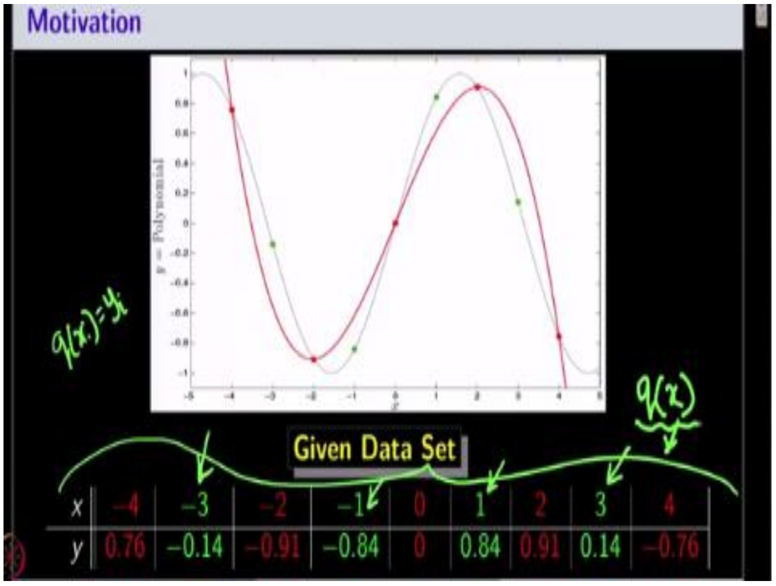


And what we are going to do with this is, suppose we want to find the value of the function at some point x say x is equal to say 3.5 we want to find the value of this function say it is this value. But since we do not know this function of course we will not be able to evaluate using the exact function rather if we find such a polynomial then we can actually get the value of the polynomial and take that as the value that we want that is you can evaluate the value at -3.5.

We are interested in the value at -3.5, then you can just evaluate this. Therefore, we are actually interested in finding the value, $sin(-3.5)$. Since we do not know this function therefore, we cannot find this value rather we constructed the polynomial that passes through this data points and then we will find the value of that polynomial and consider this as an approximation to what actually we want. In this process what we are losing as an error is the distance between these two points.

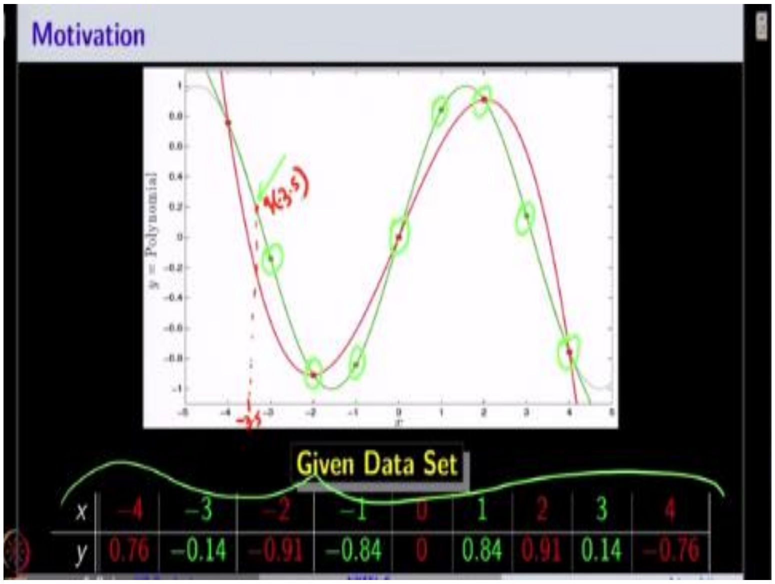
In this case you can see that the approximation is not that good because $p(-3.5)$ is quite far away from the exact value.

(Refer Slide Time: 04:08)



Rather, what I will do now is I will introduce little more points into my data set that is, I am inserting more points into my data set. And now I will ask the same question, can I find a polynomial $p(x)$ let us call this as now $q(x)$ that passes through all these points that is $q(x_i)$ will be equal to y_i and graphically, we are expecting this polynomial to pass through all the given data points.

(Refer Slide Time: 04:56)



So, you can see that such a polynomial is shown in the green solid line and you can see that this green solid line passes through all the points. So, it passes through all the points that we have taken as data in our table and this is actually polynomial that comes out of this data set. Now if I ask can I find the value of the function at the point -3.5. Of course, we cannot find this value from the exact function because we do not know what is that function.

Rather, we can now use $q(x)$ to find that value at least graphically you can see that $q(x)$ is giving us a good approximation to the exact value that we want. So, from here what we see that if you are given a data set then we would like to find a polynomial that satisfies the data set. Graphically the polynomials graph should pass through all the points that are given in the data set that is what we want. Why we want to construct such a polynomial?

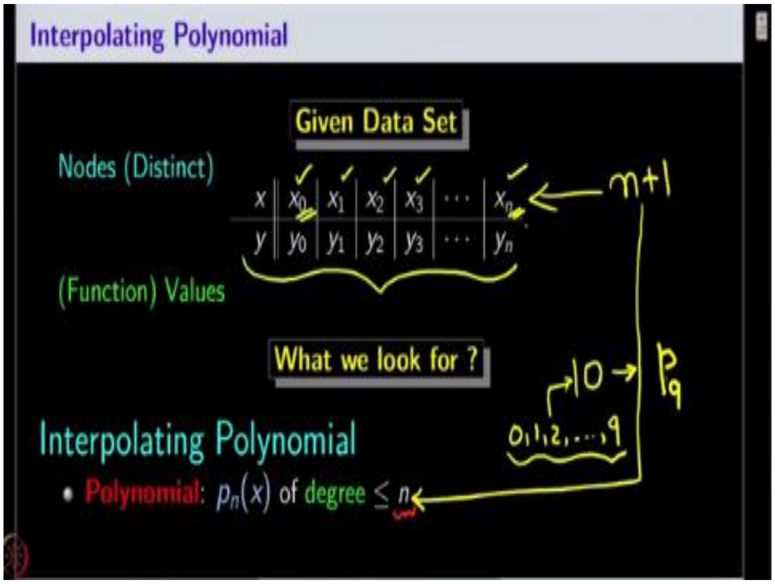
Often in practical applications we have some data but sometimes we miss certain data in the given data set due to some problems. For instance, suppose you have a model in which you have to give certain inputs that are generated from an experiment. For instance, you have to run an experiment and get the data for every half an hour in one day and that set of 48 data should be plugged into a model set.

Now you go to perform the experiment and every half an hour you will take the reading from the experiment and the next day you will collect that 48 set of data and give to your model, this is what you are doing. And suppose in between something happened and you were not able to take the reading for certain data. Like for instance at 1 o'clock you went for lunch and you forgot to take the reading from your experiment.

The experiment is running but you miss to take the experiment data at 1 o'clock say. Then at the end of the day you will be having 47 data where x coordinate is the time and y coordinate is the observation that you had from the experiment. Now in between at 1 o'clock you missed the reading and therefore this is not a complete data set that can be plugged into your model for further analysis. Now what you will do?

You can probably guess something randomly and feed that as the value at 1 o'clock. But if it is ok then you can go for generating a polynomial that satisfies the given 47 data sets and then take the value of that polynomial that will also be not exactly what you would have observed in the experiment. What you lost is definitely lost but constructing the polynomial and taking the value from that polynomial may be mathematically justified way of choosing or guessing. See that is the only advantage of going for such tools.

(Refer Slide Time: 08:48)



With that let us post some question. Suppose we are given a data set like this. In this the x coordinate is generally called nodes and we always take these nodes as distinct nodes that is each

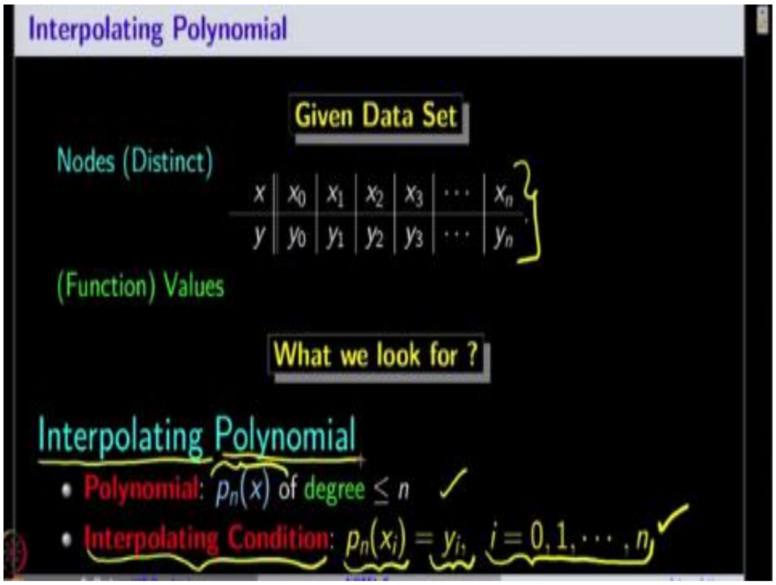
$x_i \neq x_j$ for every $0 \leq i, j \leq n$. At least for the construction of interpolating polynomial we will consider all the nodes as distinct nodes and we are also given corresponding values.

These values may be generated from a function like how we did in our previous example or these values may be coming from some other source. Something like these are coming as observations from some experiments. Now the question is, can we find a polynomial whose graph passes through these points. Such a polynomial is called the interpolating polynomial. Let us give the definition more precisely.

We are looking for a function which is a polynomial of degree less than or equal to n. What is this n? You are given a set of data where there are $n + 1$ nodes. You can see that the nodes index starts from 0 and goes up to n. Therefore, there are $n + 1$ nodes. If you are given $n + 1$ nodes then you want polynomial with degree n. Suppose you are given 10 nodes then you want a polynomial of degree one less that is 9. So, this you should remember.

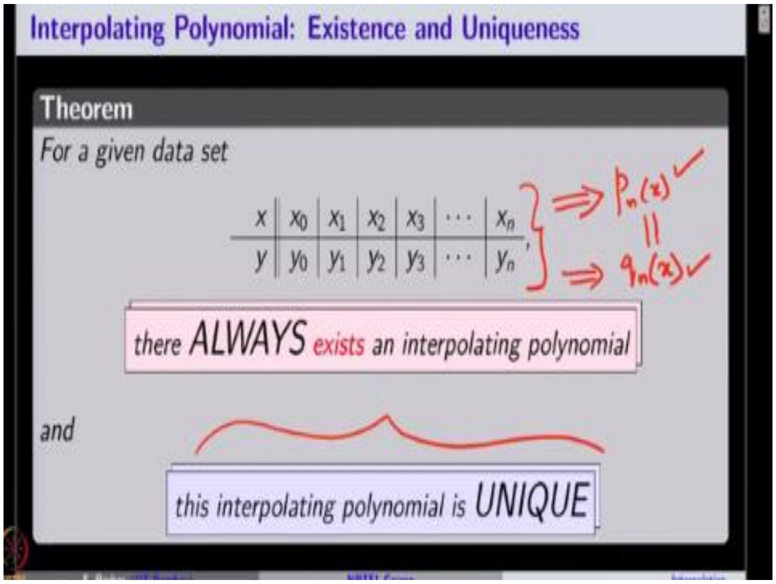
Always we will index the nodes from 0, 1, 2 and so on up to 9 therefore you are given 10th set of nodes and therefore its corresponding interpolating polynomial will be p_9. So, that you have to remember.

(Refer Slide Time: 11:08)



The next thing is what is called the interpolation condition. It means $p_n(x_i) = y_i$ for each i equal to 0 to n. So, if a polynomial is obtained for a given data set that satisfies these two conditions, then we will call that polynomial as the interpolating polynomial. I hope you understood what is mean by interpolating polynomial.

(Refer Slide Time: 11:41)



Let us go to ask the next question. Suppose I am given a data set like this can I always get a interpolating polynomial that is the first question and also it is often in applications we would like to have methods that finally gives us a unique output. This is because, if our method gives many outputs even though the method is very good, we will have the confusion of which output we should choose in order to fulfil our requirements.

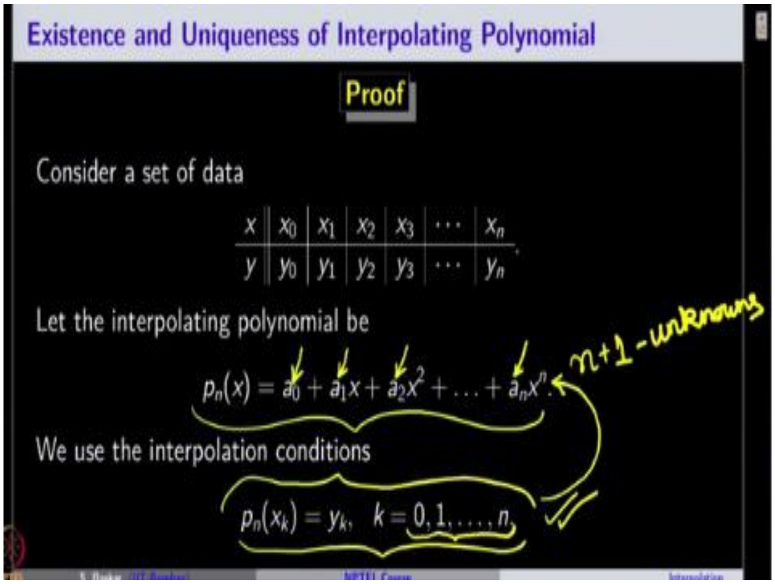
Therefore, it is always comfortable for us to work with those methods that give unique output. Therefore, often in numerical analysis we will also look for methods that give unique solution. In that way the interest for us also, to see whether the interpolating polynomial for a given set of data is unique. In fact, in this theorem, we have positive answer for both these questions. The theorem says that you give me a set of data remember you are giving a set of data and that data is fixed.

Now for that data, I can always find an interpolating polynomial and moreover that interpolating polynomial will be unique that is for the theorem says. It means for this particular data I cannot get two different polynomials say $p_n(x)$ and $q_n(x)$. Say for instance, I use a method and find

$p_n(x)$ and you use another method to construct the interpolating polynomial which you denoted by $q_n(x)$, the theorem says that we both will have the same interpolating polynomial.

We cannot have two different interpolating polynomials even if we use two different methods to construct such a polynomial that is what the theorem says. Let us try to prove this theorem.

(Refer Slide Time: 13:51)

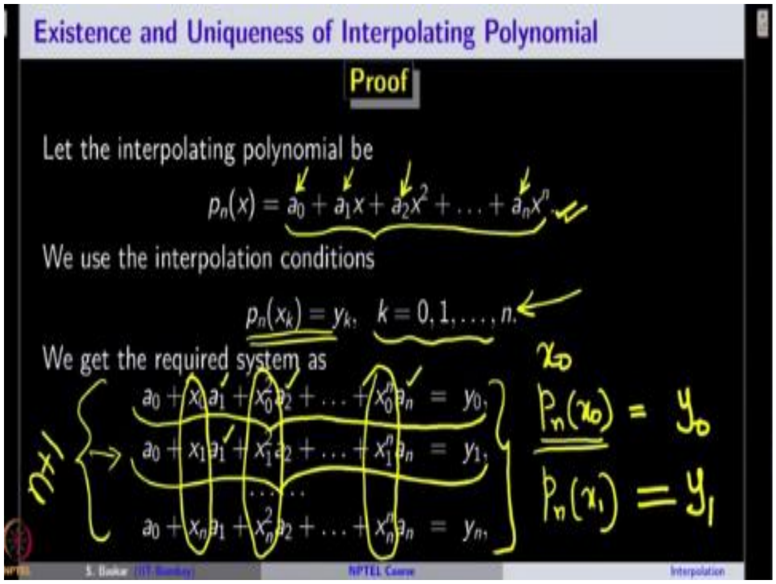


Consider the given set of data. Now what we want to do, is to see that their existing interpolating polynomial. Remember the interpolating polynomial is basically a polynomial with degree less than or equal to *n*. So, therefore we will assume that our interpolating polynomial is of this form. Remember we can write any polynomial of degree less than or equal to *n* in this form where we have to only find these coefficients a naught a_1, a_2, upto a_n.

If we manage to find these coefficients it means we do have such a polynomial that is the understanding. And remember that we have this interpolation conditions that is we want to find such a polynomial such that it satisfies these conditions. Now itself you can see how the proof goes on. You see you have some unknowns. How many unknowns are there? There are *n* + 1 unknowns are there and how many conditions are there?

There are again *n* + 1 conditions. These are the interpolation conditions. Now we will impose these *n* + 1, conditions into this general form and see what happens.

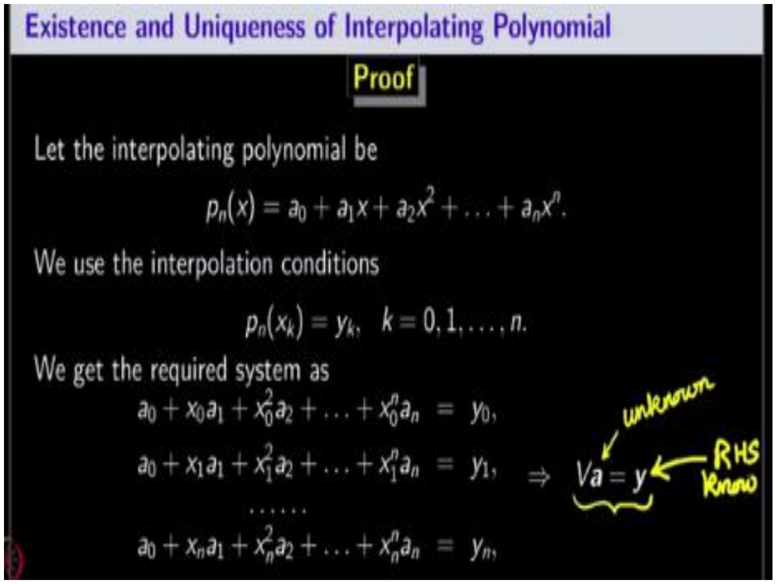
(Refer Slide Time: 15:29)



We will get a system of equations. You take x_0 that is the first point in our data set and plug into this expression what you will get $p_n(x_0)$ should be equal to y_0. What is $p_n(x_0)$? You put $x = x_0$ in this expression that is what gives us the first equation. You can see $a_0 + x_0a_1 + x_0^2a_2$ and so on up to $x_0^na_n = y_0$. Similarly, you put $p_n(x_1)$ that is you put x_1 on the right-hand side of this equation and that should be equal to y_1 and that gives you the second equation.

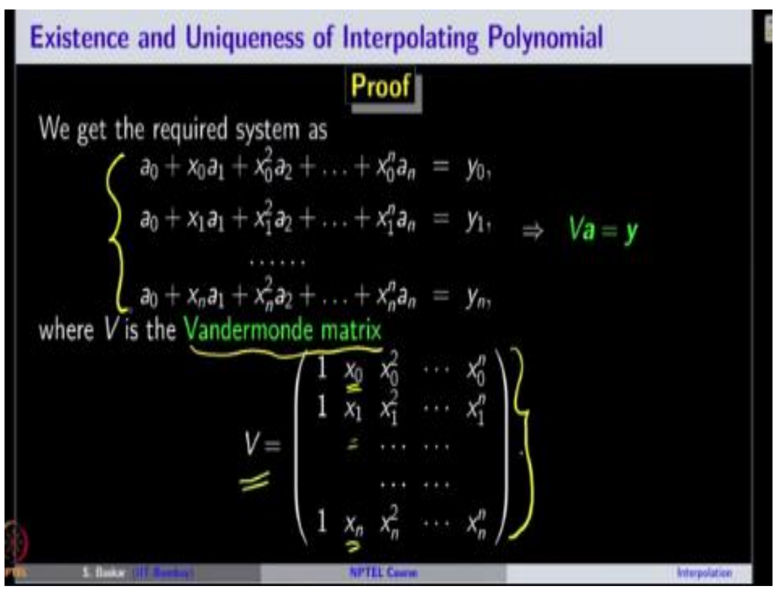
Similarly, you use all the interpolation conditions. Each condition gives you one equation in this system and therefore you have n + 1 equations remember x_0, x_1 up to x_n are known to us. Therefore, all these quantities are known quantities to us. What are all unknowns here? Unknowns are precisely the coefficients of this polynomial a_0, a_1 up to a_n and they are appearing here with power 1.

(Refer Slide Time: 17:14)



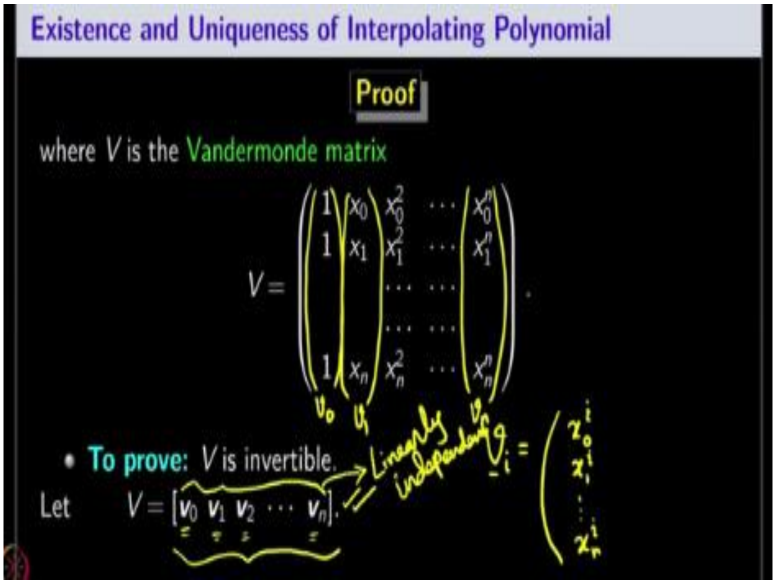
Therefore, this is in fact a linear system of equations and therefore we can write this linear system in the form $Va = y$ where a is unknown, y is the right-hand side vector which is known to as through the data set that is given to us. And what is the coefficient matrix V?

(Refer Slide Time: 17:47)



Coefficient matrix V is given like this where x_0, x_1, x_2 up to x_n are also known to us from the given data set and this matrix V is generally called the Vandermonde matrix. Now our problem of proving the existence and uniqueness of the interpolating polynomial of a given data set is reduced to seeing whether this system has a unique solution or not. Now from your linear algebra course you know that a linear system has a unique solution if and only if the coefficient matrix is invertible.

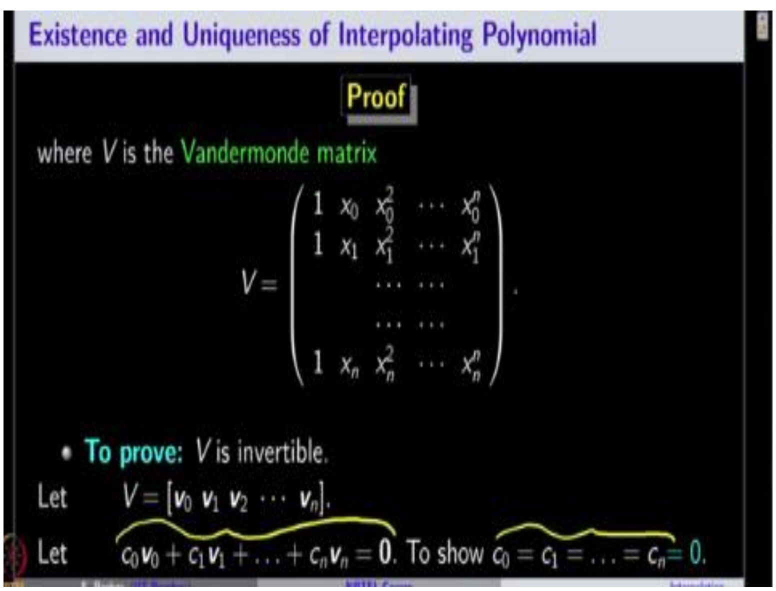
(Refer Slide Time: 18:41)



Therefore, our proof is complete if we prove that the Vandermonde matrix is invertible. Let us see how to prove that V is invertible for us we will just write the matrix V in the form $[\boldsymbol{v_0}\boldsymbol{v_1}, \boldsymbol{v_2} \cdots \boldsymbol{v_n}]$ where $\boldsymbol{v_i}$ is a column vector with coordinates as $x_0^i,\ x_1^i$ and so on up to x_n^i. That is this is $\boldsymbol{v_0}$ this is $\boldsymbol{v_1}$ and so on up to this is your $\boldsymbol{v_n}$. We are just writing this matrix with this notation, that is all we have not done anything.

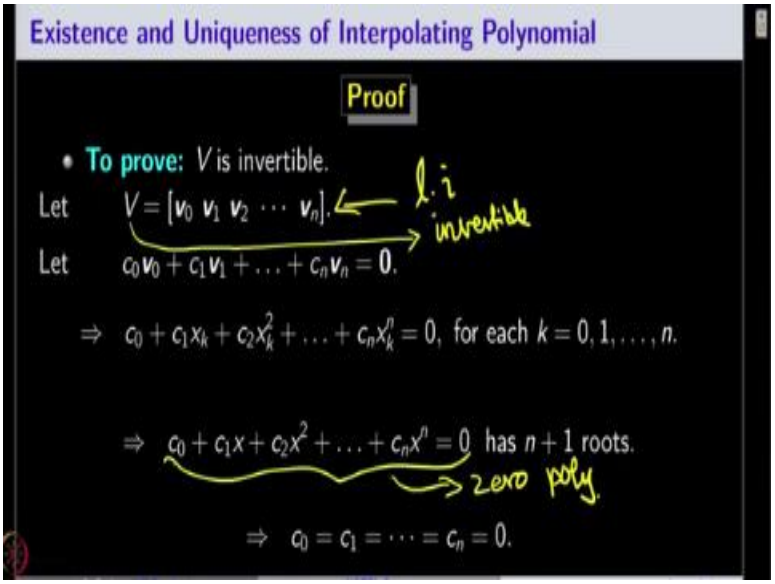
Now we want to prove that this matrix is invertible that is equivalent to proving that all this column vectors are linearly independent. How are we going to prove that all this vectors $\boldsymbol{v_i}$'s are linearly independent?

(Refer Slide Time: 20:02)



Let us write this equation, this is the vector equation which is obtained by writing the linear combinations of $\boldsymbol{v}$ is and you just equate that to 0 it gives you a linear system again. Homogeneous linear system with solutions as c_0, c_1 up to c_n. And in order to prove that $\boldsymbol{v_i}$'s are linearly independent we have to show that this system has the 0 solution that is $c_0 = c_1 = \cdots = c_n$ all these scalar should be equal to zero. So, let us see how to prove this.

(Refer Slide Time: 20:57)



Let us take this system and write the kth equation of this system. So, we are taking this system and we are writing the kth equation of this system. So, that is precisely given by this where k varies from 0 to n. So, for each k you have one such equation. You can see that the left-hand side is a polynomial in x evaluated at x_k that shows that this polynomial equation has $n + 1$ distinct roots. What are they?

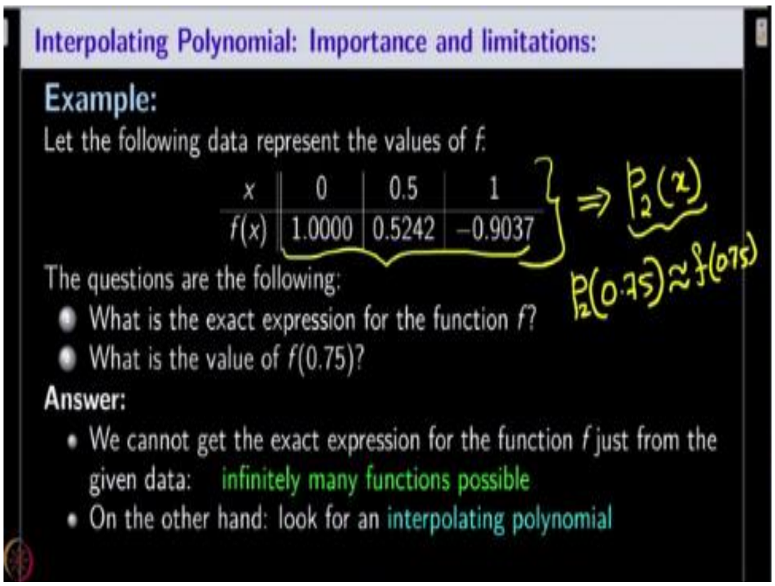
They are x_0, x_1, up to x_n because for each x_k this system says that this polynomial equation holds at $x = x_k$. So, that is what is meant by this system and that implies that this polynomial has $n + 1$ distinct roots. Remember we have considered our nodes in the data set as distinct nodes therefore all this x_i's are distinct. Now you see we have the polynomial with degree n and it has $n + 1$ roots. What it means? It means this is a zero polynomial.

So, this is possible only with zero polynomial and that means that all this coefficient c_i's are equal to 0. So, this shows that all this $\boldsymbol{v_i}$'s are linearly independent and that shows that the matrix V is invertible. That means that your linear system has unique solution. What is the solution? The solution is precisely the vector a which is equal to $(a_0, a_1, \cdots, a_n)^T$. They are precisely the coefficients of this polynomial.

And since this system is obtained by imposing the interpolation, you see that such a polynomial of degree less than or equal to n do exist and it satisfies all this n + 1 interpolation conditions. It means the interpolation polynomial exists and unique for the given data set. So, this completes the proof of our existence and uniqueness theorem for interpolating polynomials. Let us give an example to understand the importance and the limitation of the polynomial interpolation.

The interpolation polynomial is a nice tool for approximating a given data set but it has its own limitations. Let us try to understand what is the importance and limitations. You may be understanding the importance from our previous discussion.

(Refer Slide Time: 24:11)



Let us just try to put forth once again. Suppose we are given this data set and we ask this question. Give me the exact expression of the function from where I have generated this data that is my first question and the second question is you also give me the value of the function f at the point 0.75. That is in between these two node points, I want to pick up one point, I want to find the value of the function at that point.

What is the answer for the first question? It is simply impossible for us to trace back what was that function f from where we got this data. Why, because we can get infinitely many such functions. Therefore, how do we know from which function we have generated this data unless we come forward and reveal that secret. So, this is simply impossible and therefore the answer for the first question is just not possible.

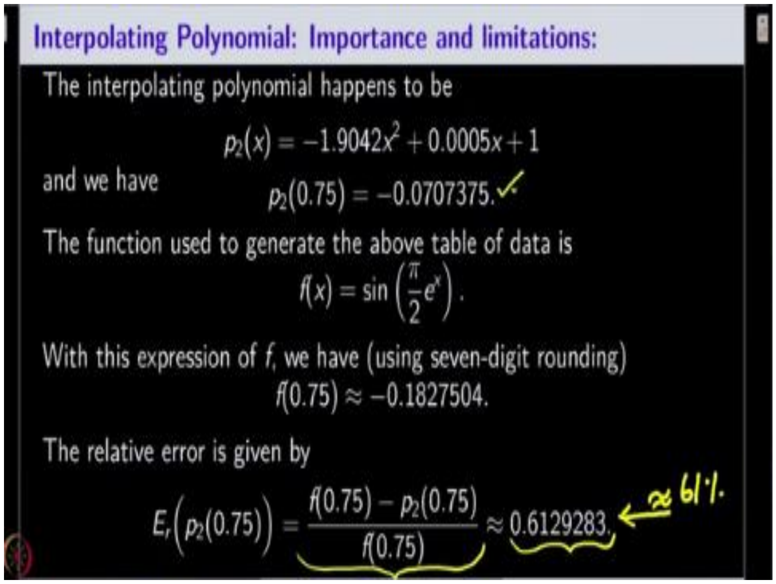
Therefore, if you have only certain data sets there is no possibility that you can retrieve the function or it is not possible for us to exactly find the value of the function at some point where the data is not given to us. It is simply impossible. Suppose you have a situation like what we have just considered in our initial example. Suppose you have a situation where you missed certain data but you very badly need it. Then what you can do? One thing is you can simply guess it.

For instance, you may simply have a wild guess of this as say 0.15 or something like that. But what is the guarantee that there is any logic that this function is going to take this value. So, at least mathematically we have a rigorous way of guessing this value. Again, I am telling you this is again as good as guessing only but only better situation is that there is a mathematical justification of guessing this value that is through the interpolating polynomial.

What you can do is? You can go for interpolating polynomial, construct the polynomial for this data and then take that polynomial say you have three node points. Therefore, from here you can get the degree one less that is a quadratic polynomial you can get. And now once you get the quadratic polynomial you can go to find the value of that quadratic polynomial and be happy with it by believing that it is approximately equal to $f(0.75)$.

So, that is life and this is what we can do at the best as long as this tool is concerned that is the interpolating polynomial. Otherwise, you have only one choice that is to just widely guess what is the value of this function at 0.75. Let us do this.

(Refer Slide Time: 27:34)

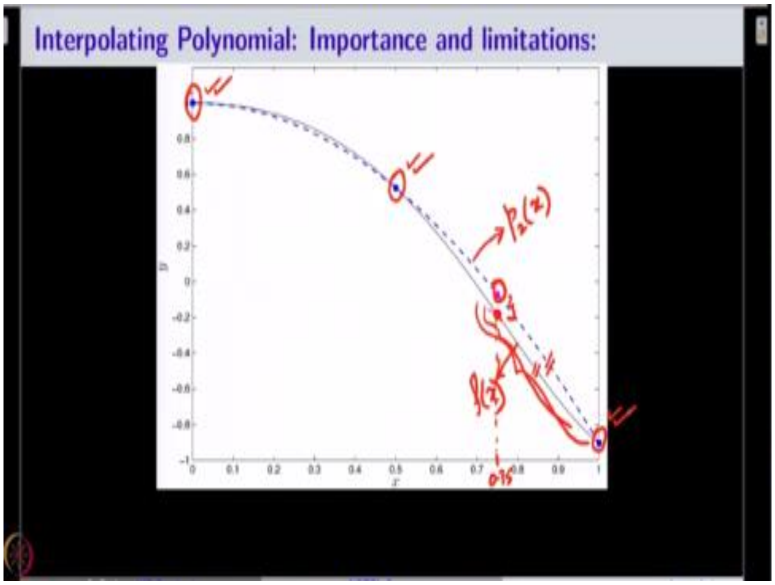


The interpolating polynomial for the given data set is given like this. I am not going to tell you how I constructed it. Later in the next lecture we will see two ways for constructing this polynomial. In this example we are only interested in understanding the importance and the limitation of this tool. The construction of such polynomials we will discuss in the next class. Therefore, I will just pause what is this quadratic polynomial that interpolates the given data and it is given like this.

Once you have this you can go to find the value of the quadratic polynomial at the point 0.75 and it is given like this. Once you get this let me reveal the function from where I have generated this data set. The function that I have considered is $sin\left(\frac{\pi}{2}e^x\right)$. Now let us take the value of this function at 0.75 and see how good our p_2's approximation is when compared to the exact value will the function value is given by - 0.1827504.

And what is the relative error in $p_2(0.75)$ that is the approximation that we obtained when compared to the exact value. The relative error if you recall it is computed using this formula and that is given by this that is quite upsetting because we have around approximately 61 percent error in our approximate value when compared to the exact value. Does it mean that this polynomial interpolation tool is a bad tool? Let us see graphically how this function f and p_2 are.

(Refer Slide Time: 29:39)

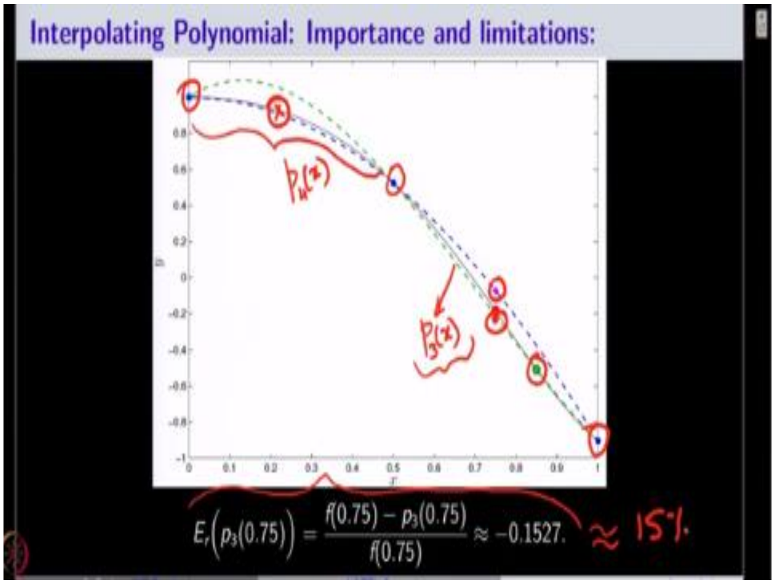


The graph of the function $f(x)$ is given in the grey colour solid line and the graph of the polynomial $p_2(x)$ that we constructed using interpolation conditions is given in the dashed blue line and I have used these points as the nodes and the corresponding value of the function f. Now you can see that we wanted the value at 0.75. It means we wanted this value but we got this value as the approximation. What is the error involved?

The error involved is the difference between these two points. You can see that the function graph is coming very nicely but there is a point of inflection starting from somewhere here and because of that the graph of the function f start a deviating much from the polynomial. Of course, it is unfair for us to expect our interpolating polynomial to sense this sudden change of the curvature in the curve and adapt to that because we are only giving these three nodes and the corresponding values.

Therefore, it is unfair for us to expect the interpolating polynomial to sense all this kind of properties of the function and try to get a better approximation. Now how will you get the better approximation? The natural thing is, give more information.

(Refer Slide Time: 31:31)



For instance, you give one more node and the corresponding value of the function then you can see that the green line now is the cubic polynomial that interpolates the points this and now the new point this and this point. So, now I am giving four points therefore I can get a cubic polynomial and now you see the cubic polynomial will give you this as the approximation that looks far better than what we got from the quadratic polynomial.

Of course, naturally this side we got a bad approximation. But suppose if you give one more node here and get p_4 out of that that will be obviously better than what we got from p_3 and so on. In fact, you can see the relative error involved in $p_3(0.75)$ that is approximately 15 percent error we had. This is also not good but when compared to the quadratic polynomial the approximation is little better in the case of the cubic polynomial interpolation.

Like this from this example we can see the limitations of the interpolating polynomial that is if you want better approximation you have to give more information about the function, that is quite natural. So, this is not something bad that is expected. If you give more information, it means you are going for higher degree polynomials. So, higher degree polynomials, at least by the first look seems to be getting better approximation to our function than the lower degree polynomials.

But this is not always true, you will see later in our discussion that arithmetic error can spoil the approximation if you keep on increasing the degree of the polynomial. In the coming class we will give two ways of constructing interpolating polynomial. Remember in this class we only discuss the existence and uniqueness of the polynomial interpolation; we never constructed these polynomials explicitly.

You remember, it is very difficult for us to solve the Vandermonde system to get the polynomial because that will be quite costly to do also. Computationally it is not a preferred way of doing. We have some better way of constructing the interpolating polynomial. There are in fact two ways one is Lagrangian form of interpolating polynomial and another is Newton's way of interpolating polynomial. In the next class we will introduce these two forms of interpolating polynomials.

They give a nice formula for these polynomials for a given data set and then in the coming class we will also learn how to do the error analysis for the interpolating polynomial. With this let us stop this lecture, thank you for your attention.