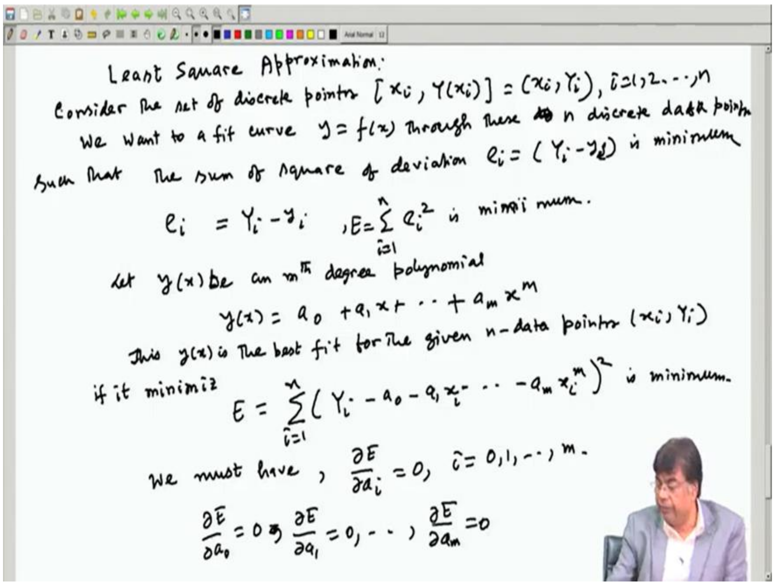
Advanced Computational Techniques
Professor Somnath Bhattacharyya
Department of Mathematics
Indian Institute of Technology Kharagpur
Lecture 08
Least Square Approximations

(Refer Slide Time: 01:16)



Okay. So, now, we talk about data points when we have a set of data points. Now, we have already discussed about the polynomial interpolation where the data points are fit by a polynomial. So, if we have n+1 data points a polynomial of degree n can be obtained without using the function derivative values if we use the that is why Lagrange or Newton's formula and all.

Now, similarly, if we use the Hermite interpolation polynomial where the function as well as its derivative is used for calculation of the polynomial. So, then using n data points we can have 2n-1 degree polynomial. We talk about a least square approximation for a set of data points now, that is, we would like to fit a curve through a given set of data points.

So, let us consider the set of discrete points say $[x_i, y(x_i)]$ which is not known, let us call this $[x_i, y(x_i)]$= $(x_i, y_i), i = 1,2, \dots, n$

and we would like to approximate and we want to fit a curve y=f(x) through this we are given, $i = 1,2, \dots, n$ through this discrete data points through these n discrete data points such that the sum of square of the deviation e_i let us call this, $e_i = (Y_i - y_i)$,

is minimum.

Sum of square of the deviation between these exact data points and the approximate value whatever we are obtaining through the curve fitting. So, that means, what we need this we are defining as $Y_i - y_i$ in such a way that the e_i^2 is minimum. So, let us call this E as

$$E = \sum_{i=1}^{n} e_i^2$$

minimum. So, this is our requirement.

Now, let us approximate y(x), m^{th} degree polynomial so that means

y(x)=$a_0 + a_1 x + \cdots + a_m x^m$

So, now what do we know is so this y(x) is the best fit for the given n data points (x_i, y_i) if it minimizes E,

E=$\sum_{i=1}^{n}(Y_i - a_0 - a_1 x_i - \cdots - a_m x_i^{m})^2$

Now, what is in our choice or what is the parameter with which we can choose in order to minimize this E in form of the polynomial. So, that means a_0, a_1 they are to be chosen in such a way this summation the square sum is minimum. So, that means, what we need is at those extreme point the derivative should be zero.

$$\text{So, } \frac{\partial E}{\partial a_i} = 0$$

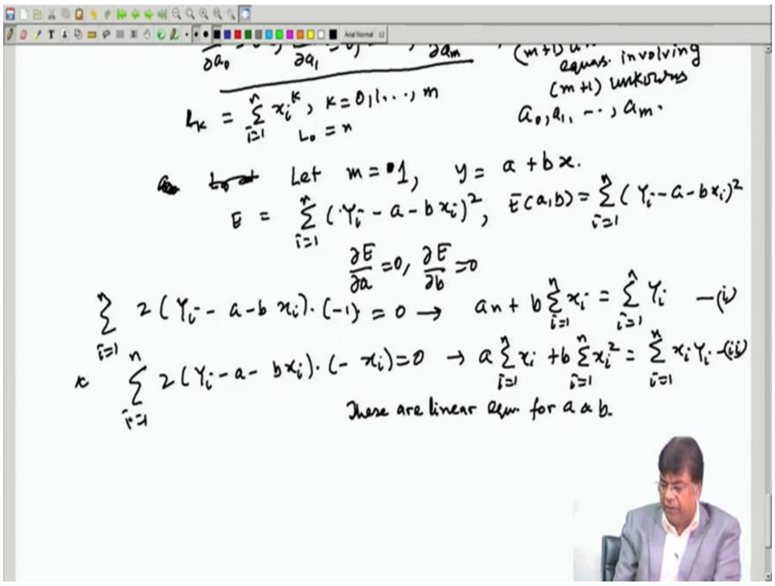
I=1, 2,…, m.

So, that means, what we should need is

$\frac{\partial E}{\partial a_0} = \frac{\partial E}{\partial a_1} = 0, \dots, \frac{\partial E}{\partial a_m} = 0$

because unknowns are $a_0, a_1, \dots, a_m$ and they are appearing in a linear fashion.

(Refer Slide Time: 07:29)



So, this will lead to how many equations m+1 linear and they will be algebraic equations involving m+1 unknowns $a_0, a_1, \ldots, a_m$. So, basically what we have given a condition that you have a set of data points and you choose the best fit for those data points a curve or that which is a polynomial of degree m to this curve such that the square of the deviation between the function value and the data point value is minimum.

So, this is the restriction or constraint we have imposed. So, this will lead to a set of n +1 number of equations which need to be solved so, if I take the derivative, so, what we get is they take the derivative and summation. The first one is

$$l_k = \sum_{i=1}^{n} {x_i}^k, k = 0,1, \ldots, m$$

First let us consider, so, that way we can proceed now, let us consider a linear fit so m= 1 so, that means, we are fitting a polynomial by this manner

$$y = a + bx`$$

So, in that case what we have is

$E = \sum_{i=1}^{n}(Y_i - a - bx_i)^2$Y

Okay, let me write in this way.

$E(a, b) = \sum_{i=1}^{n}(Y_i - a - bx_i)^2$

So, what we need is

$$\frac{\partial E}{\partial a} = 0, \frac{\partial E}{\partial b} = 0$$

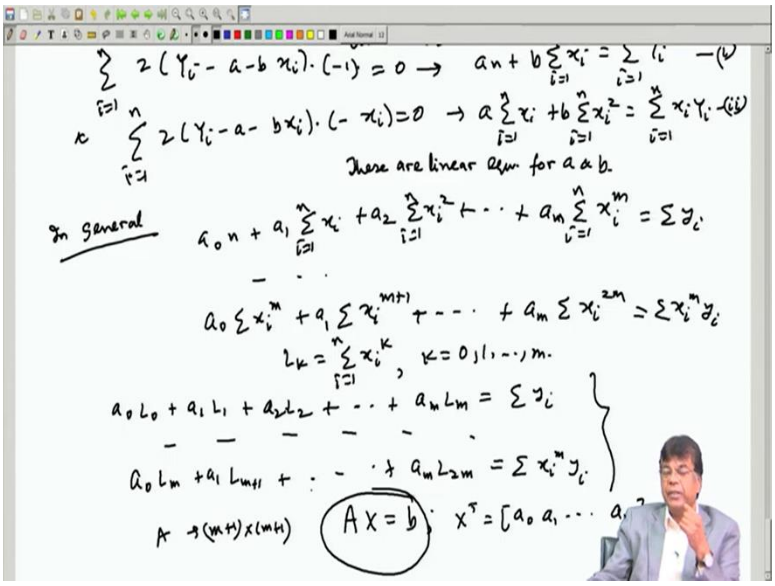
So, if we take the derivative

$$\sum_{i=1}^{n} 2(Y_i - a - bx_i)(-1) = 0 \rightarrow an + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} 2(Y_i - a - bx_i)(-x_i) = 0 \ \rightarrow a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i{}^2 = \sum_{i=1}^{n} x_i y_i$$

So, these two set of equations to be solved so obviously this is forming linear equations for a and b which can be solved to get the values of a,b so that this minimize the squared deviation from the given data points.

(Refer Slide Time: 14:25)



So, same thing with generalized for any m^{th} degree polynomial. So, in general what we get

$a_0 n + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} {x_i}^2 + \cdots + a_m \sum_{i=1}^{n} {x_i}^m = \sum y_i$

Similarly, we get like this way.

$$a_0 \sum {x_i}^m + a_1 \sum {x_i}^{m+1} + \cdots + a_m \sum {x_i}^{2m} = \sum {x_i}^m y_i$$

$l_k = \sum_{i=1}^{n} {x_i}^k, k = 0,1,\ldots,m$

So, if I define this way so what I get

$a_0 L_0 + a_1 L_1 + a_2 L_2 + \cdots + a_m L_m = \sum y_i$

…

$$a_0 L_m + a_1 L_{m+1} + \cdots + a_m L_{2m} = \sum {x_i}^m y_i$$

gives m +1 equations in m +1 variables which is a complete compact system.

So, one can solve this and get it. So, this leads to a formula like this way

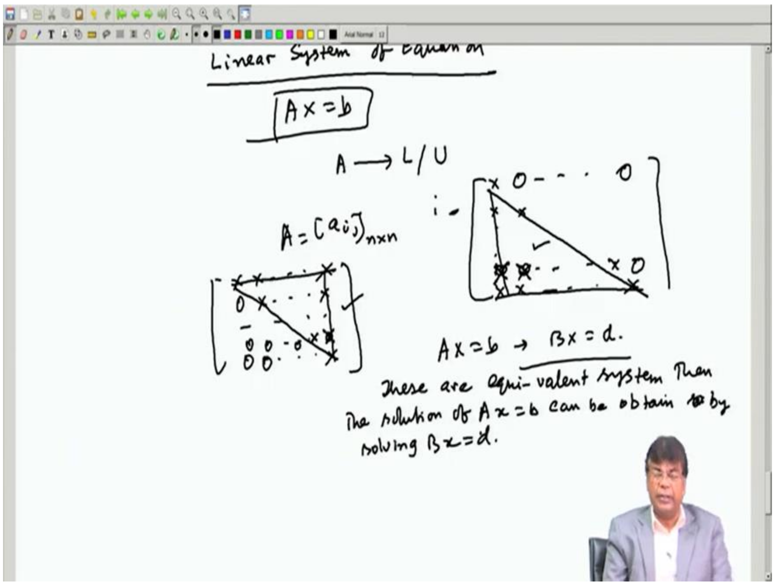
Ax=b, $x^T = [a_0 a_1 \ldots a_m]$, A→ $(m+1) \times (m+1)$

So, this is how the least squares approximation for a given set of data points are determined. So, that means you have a set of data points and through that is a curve fitting is obtained.

Now, obviously, as it appears that this curve fitting is much simpler than in general form of polynomial and other representations. Now, representations of a function can be by several ways, one is Taylor series expansion that is if we know the function values and from there one can obtain the Taylor series expansion. Similarly, one can express in a Fourier series. Fourier series representation or any orthogonal function if we have then one can obtain an expression by the inner product of those.

So, several ways a function can be represented. One of the simple one is the least square approximation so that means with a given data point say it is the linear or quadratic or cubic we fit those data points by a polynomial form. Now, obviously, what it appears that in even in the previous case that is the situation of the spline interpolation So, that means when we are representing a function by a cubic set of cubic polynomial in each sub interval so there also we came across a set of tridiagonal system of equations.

(Refer Slide Time: 20:01)



So, in several occasions even here we find a system of equations to be solved. So, linear system of equations. So, some techniques we should learn how to solve a linear system of equations. Now, first of all as we know that $Ax = b$ is a system of equations. Now, what we do is we make a transformation from $Ax=b$ by some manipulation.

So, our one of the intention is to transform to either a upper triangular or a lower triangular matrix. So, if we have A is any matrix say

A=$[a_{ij}]_{nXn}$

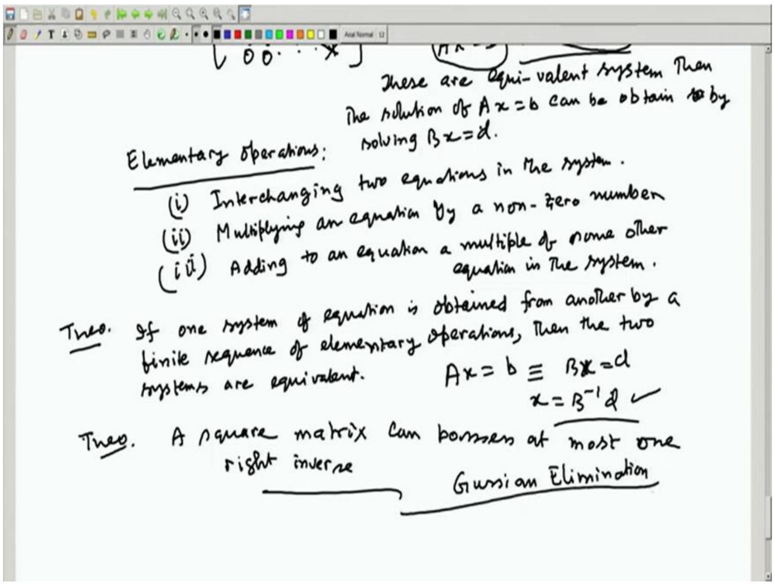
instead if we have a situation say a matrix where you have the only some places you have data point, x and zero.

So, if this is the situation or if it is upper triangular, say you have a situation like this that means, you have nonzero entries only up to diagonal and then you have all the edges as zero so like that way.

So, whenever you have reduced form then a system of equation can be solved much easily. So, now, we should have an equivalence between two systems.

So, if you have a stream you have to solve Ax = b $\rightarrow$ **B**x=d. Now, these two system are equivalent so then the solution of Ax = b can be obtained by solving **B**x=d whichever the simpler way.

(Refer Slide Time: 23:49)



Now, there are some restrictions for that any system which is equivalent to other if the second one that is **B**x=d are obtained by elementary operations. So, under these elementary operations if the reduced system equivalent to the original one. So, what are the elementary operations interchanging two equations in the system, in this case these remain equivalent. Multiplying an equation by a nonzero number, adding to an equation, a multiple of some other equation in the system. So, if we do these elementary operations to system should remain the equivalent. So, if B is obtained by those elementary this **B**x=d is determined by those elementary operations.

Under those elementary operations $Ax = b$ and **B**x $= d$ remains equivalent.

Now, we state one theory if one system of equation is obtained from another by a finite sequence of elementary operations then the two systems are equivalent. Now, we can obtain the solution as B inverse if inverse exist. Now, one other theory which proved the uniqueness of this is square matrix can possess at most one right inverse.

i.e. $x = B^{-1}d$

So, this gives uniqueness of the of the solution finding. Now, I will talk about the simple procedure like Gaussian elimination this is what we do from our, from the very beginning without any training in algorithm form one can write the Gaussian elimination procedure where the matrix is reduced to a triangular form. So, we will talk in the next class some of the iterative procedure. Thank you.