Numerical Methods

Lecture 06: Curve Fitting and Interpolation

Kundan Kumar
Associate Professor
Department of ECE

Faculty of Engineering (ITER)
S'O'A Deemed to be University, Bhubaneswar, India-751030

Introduction

- In many scientific and engineering experients, observations of physical quantities are measured and recorded.
- For example, the strength of many metals depends on the size of the grains.
- Testing specimens with different grain sizes yields a discrete set of numbers (d- average grain diameter, σ_y - yield strength) as

Table 6-1: Strength-grain size data.

d (mm)	0.005	0.009	0.016	0.025	0.040	0.062	0.085	0.110
σ_y (MPa)	205	150	135	97	89	80	70	67

- The experimental records are typically referred to as data points.

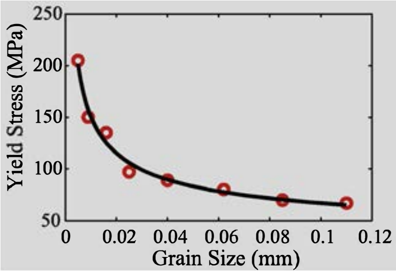
Introduction

- Often the data is used for developing, or evaluating, mathematical formulas (equations) that represent the data.
- This is done by curve fitting in which a specific form of an equation is assumed, or pro vided by a guiding theory, and then the parameters of the equation are determined such that the equation best fits the data points.
- Curve fitting can be carried out with many types of functions and with polynomials of various orders.
- Sometimes the data points are used for estimating the expected values between the known points, a procedure called interpolation,
- For predicting how the data might extend beyond the range over which it was measured, a procedure called extrapolation.

Curve Fitting

Curve Fitting

- Curve fitting is a procedure in which a mathematical formula (equation) is used to best fit a given set of data points.
- The objective is to find a function that fits the data points overall. This means that the function does not have to give the exact value at any single point, but fits the data well overall.



- A curve that shows the best fit of a power function ($\sigma = Cd^m$) to the data points.
- It can be observed that the curve fits the general trend of the data but does not match any of the data points exactly.
- Generally, all experimental measurements have built-in errors or uncertainties, and requiring a curve fit to go through every data point is not beneficial.

Curve Fitting with a linear equation

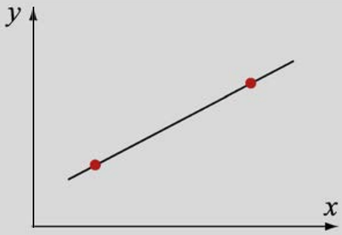
Curve Fitting with a linear equation

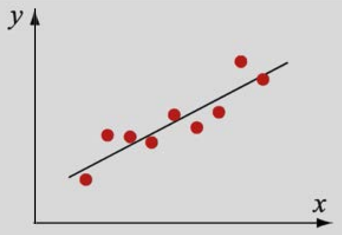
- Curve fitting using a linear equation (first degree polynomial) is the process by which an equation of the form:

$$y = a_1 x + a_0$$

is used to best fit given data points.

- This is done by determining the constants a1 and a0 that give the smallest error when the data points are substituted in the equation.



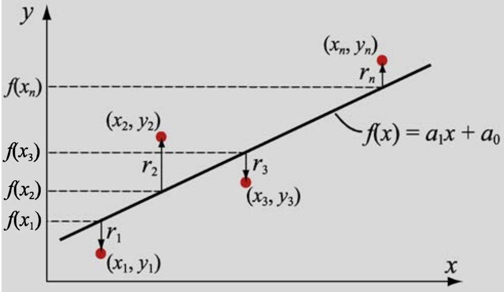


Measuring How Good Is a Fit

- The fit between given data points and an approximating linear function is determined by first calculating the error, also called the residual, which is the difference between a data point and the value of the approximating function, at each point.

- Subsequently, the residuals are used for calculating a total error for all the points.
- The residual r_i at a point, (x_i, y_i), is the difference between the value y_i of the data point and the value of the function $f(x_i)$ used to approximate the data points:

$$r_i = y_i - f(x_i)$$



Measuring How Good Is a Fit

- A criterion that measures how well the approximating function fits the given data can be obtained by calculating a total error E in terms of the residuals.

$$E = \sum_{i=1}^{n} r_i = \sum_{i=1}^{n} [y_i - (a_1 x_i + a_0)]$$

or $$E = \sum_{i=1}^{n} |r_i| = \sum_{i=1}^{n} |y_i - (a_1 x_i + a_0)|$$

or $$E = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} [y_i - (a_1 x_i + a_0)]^2$$

- A smaller E in indicates a better fit. This measure can be used to evaluate or compare proposed fits, and last eqation can be used to calculate the coefficients a_1 and a_0 in the linear function.

Linear Least-Squares Regression

- An experiment produces a set of data points $(x_1, y_1), \ldots, (x_n, y_n)$, where are abscissas $\{x_k\}$ are distint.
- One goal of numerical methods is to determine a formula $y = f(x)$ that relates these variables.

$$y = f(x) = a_1 x + a_0$$

- Linear least-squares regression is a procedure in which the coefficients a_1 and a_0 of a linear function $y = a_1 x + a_0$ are determined such that the function has the best fit to a given set of data points.
- The best fit is defined as the smallest possible total error that is calculated by adding the squares of the residuals

$$E = \sum_{i=1}^{n} [y_i - (a_1 x_i + a_0)]^2$$

Linear Least-Squares Regression

- Take the partial derivative of of above equation, we get

$$\frac{\partial E}{\partial a_0} = -2\sum_{i=1}^{n}(y_i - a_1 x_i - a_0) = 0$$

$$\frac{\partial E}{\partial a_1} = -2\sum_{i=1}^{n}(y_i - a_1 x_i - a_0)x_i = 0$$

- Above two equations are a system of two linear equations for the unknows a_1 and a_0, and can be rewritten in the form as

$$na_0 + \left(\sum_{i=1}^{n} x_i\right)a_1 = \sum_{i=1}^{n} y_i$$

$$\left(\sum_{i=1}^{n} x_i\right)a_0 + \left(\sum_{i=1}^{n} x_i^2\right)a_1 = \sum_{i=1}^{n} x_i y_i$$

Linear Least-Squares Regression

- Solution can be written as

$$a_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad a_0 = \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i y_i\right)\left(\sum_{i=1}^{n} x_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

- The values of a_1 and a_0 in the equation $y = a_1 x + a_0$ that has the best fit to n data points (x_i, y_i)

$$a_1 = \frac{nS_{xy} - S_x S_y}{nS_{xx} - (S_x)^2} \qquad a_0 = \frac{S_{xx}S_y - S_{xy}S_x}{nS_{xx} - (S_x)^2}$$

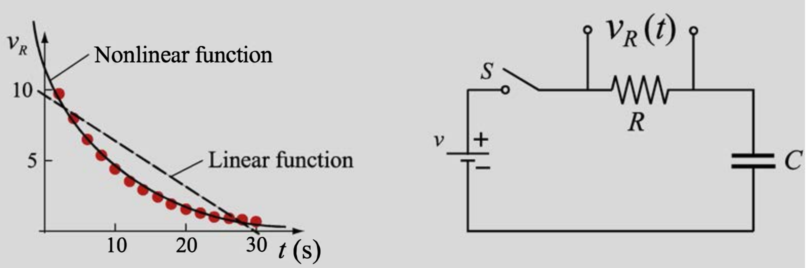
where,

$$S_x = \sum_{i=1}^{n} x_i, \quad S_y = \sum_{i=1}^{n} y_i, \quad S_{xy} = \sum_{i=1}^{n} x_i y_i, \quad S_{xx} = \sum_{i=1}^{n} x_i^2$$

Curve Fitting with Nonlinear Equation

Curve Fitting with Nonlinear Equation

- Many situations in science and engineering show that the relationship between the quantities that are being considered is not linear.
- For example, the data points meansured in RC circuit.



- It is obvious from the plot that curve fitting the data points with a nonlinear function gives a much better fit than curve fitting with a linear function.

- There are many kinds of nonlinear functions which can be used with linear-squares regression method to determine the coefficients that gives the best fit.
- Examples of nonlinear functions used for curve fitting in the present section are:

$$y = bx^m \quad \text{(power function)}$$

$$y = be^{mx} \text{ or } y = b10^{mx} \quad \text{(exponential function)}$$

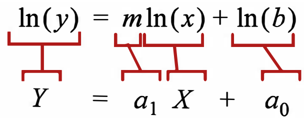
$$y = \frac{1}{mx + b} \quad \text{(reciprocal function)}$$

- In order to be able to use linear regression, the form of a nonlinear equation of two variables is changed such that the new form is linear with terms that contain the original variables.
- For example, the power function $y = bx^m$ can be put into linear form by taking the natural logarithm (ln) of both sides:

$$\ln(y) = \ln(bx^m) = m\ln(x) + \ln(b)$$

Writing a nonlinear equation in linear form

- This equation is linear for $ln(y)$ in terms $ln(x)$. The equation is in the form $Y = a1X + a0$ where $Y = ln(y)$, $a1 = m$, $X = ln(x)$, and $a0 = ln(b)$:



- This means that linear least-squares regression can be used for curve fitting an equation of the form $y = bx^m$ to a set of data points x_i, y_i.
- Once a_1 and a_0 are known, the constants b and m in the exponential equation are calculated by:

$$m = a_1 \qquad \text{and} \qquad b = e^{a_0}$$

Transforming nonlinear equations to linear form

Nonlinear equation	Linear form	Relationship to $Y = a_1 X + a_0$	Values for linear least-squares regression	Plot where data points appear to fit a straight line
$y = bx^m$	$\ln(y) = m\ln(x) + \ln(b)$	$Y = \ln(y),\ X = \ln(x)$ $a_1 = m,\ a_0 = \ln(b)$	$\ln(x_i)$ and $\ln(y_i)$	y vs. x plot on logarithmic y and x axes. $\ln(y)$ vs. $\ln(x)$ plot on linear x and y axes.
$y = be^{mx}$	$\ln(y) = mx + \ln(b)$	$Y = \ln(y),\ X = x$ $a_1 = m,\ a_0 = \ln(b)$	x_i and $\ln(y_i)$	y vs. x plot on logarithmic y and linear x axes. $\ln(y)$ vs. x plot on linear x and y axes.
$y = b10^{mx}$	$\log(y) = mx + \log(b)$	$Y = \log(y),\ X = x$ $a_1 = m,\ a_0 = \log(b)$	x_i and $\log(y_i)$	y vs. x plot on logarithmic y and linear x axes. $\log(y)$ vs. x plot on linear x and y axes.
$y = \frac{1}{mx+b}$	$\frac{1}{y} = mx + b$	$Y = \frac{1}{y},\ X = x$ $a_1 = m,\ a_0 = b$	x_i and $1/y_i$	$1/y$ vs. x plot on linear x and y axes.
$y = \frac{mx}{b+x}$	$\frac{1}{y} = \frac{b}{m}\frac{1}{x} + \frac{1}{m}$	$Y = \frac{1}{y},\ X = \frac{1}{x}$ $a_1 = \frac{b}{m},\ a_0 = \frac{1}{m}$	$1/x_i$ and $1/y_i$	$1/y$ vs. $1/x$ plot on linear x and y axes.

How to choose an appropriate nonlinear function for curve fitting

Curve fitting with quadratic and higher order polynomials

Curve fitting with quadratic and higher order polynomials

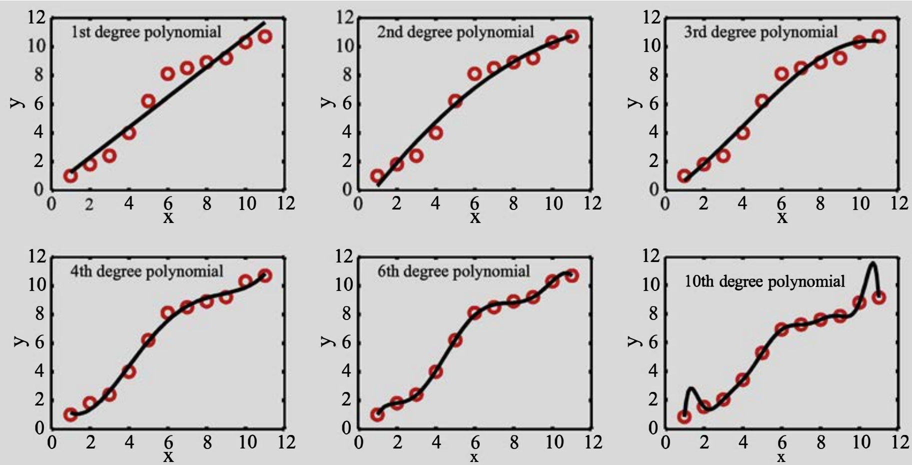
Background:

- Polynomials are functions that have the form:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

  The coefficients $a_n, a_{n-1}, \ldots, a_1, a_0$ are real numbers, and n, which is a nonnegative integer, is the degree, or order, of the polynomial.
- A plot of the polynomial is a curve. A first-order polynomial is a linear function, and its plot is a straight line. Higher-order polynomials are nonlinear functions, and their plots are curves.
- A quadratic (second-order) poly nomial is a curve that is either concave up or down (parabola).
- A third-order polynomial has an inflection point such that the curve can be con cave up (or down) in one region, and concave down (or up) in another.

Curve fitting with polynomials of different order



Curve fitting with quadratic and higher order polynomials

- A given set of n data points can be curve-fit with polynomials of different order up to an order of $(n-1)$.
- The coefficients of a polynomial can be determined such that the polynomial best fits the data by minimizing the error in a least squares sense.
- For n points the polynomial that passes through all of the points is one of order $(n-1)$. Even though the high-order polynomial gives the exact values at all the data points, it cannot be used reliably for interpolation or extrapolation.

Polynomial regression

- Polynomial regression is a procedure for determining the coefficients of a polynomial of a second degree, or higher, such that the polynomial best fits (minimizing the total error) a given set of data points.
- If the polynomial, of order m, that is used for curve fitting is:

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \ldots + a_1 x + a_0$$

- Then, for a given set of n data points $\{(x_i, y_i)\}_{i=1}^{n}$ (m is smaller than $n-1$), the total error is given by:

$$E = \sum_{i=1}^{n} \left[y_i - (a_m x_i^m + a_{m-1} x_i^{m-1} + \ldots + a_1 x_i + a_0) \right]^2$$

- The function E has a minimum at the values of a_0 through a_m where the partial derivatives of E with respect to each of the variables is equal to zero.

Polynomial regression

- For the simplicity, let us consider the case of $m = 2$ (quadratic polynomial)

$$E = \sum_{i=1}^{n} \left[y_i - (a_2 x_i^2 + a_1 x_i + a_0) \right]^2$$

- Taking the partial derivatives with respect to a0, a1, and a2 , and setting them equal to zero gives:

$$\frac{\partial E}{\partial a_0} = -2 \sum_{i=1}^{n} (y_i - a_2 x_i^2 - a_1 x_i - a_0) = 0$$

$$\frac{\partial E}{\partial a_1} = -2 \sum_{i=1}^{n} (y_i - a_2 x_i^2 - a_1 x_i - a_0) x_i = 0$$

$$\frac{\partial E}{\partial a_2} = -2 \sum_{i=1}^{n} (y_i - a_2 x_i^2 - a_1 x_i - a_0) x_i^2 = 0$$

Polynomial regression

$$na_0 + \left(\sum_{i=1}^{n} x_i\right)a_1 + \left(\sum_{i=1}^{n} x_i^2\right)a_2 = \sum_{i=1}^{n} y_i$$

$$\left(\sum_{i=1}^{n} x_i\right)a_0 + \left(\sum_{i=1}^{n} x_i^2\right)a_1 + \left(\sum_{i=1}^{n} x_i^3\right)a_2 = \sum_{i=1}^{n} x_i y_i$$

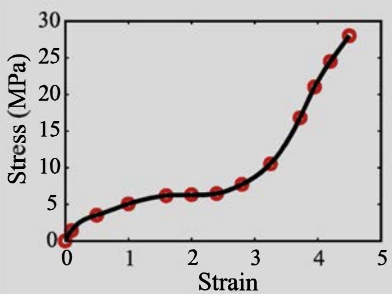
$$\left(\sum_{i=1}^{n} x_i^2\right)a_0 + \left(\sum_{i=1}^{n} x_i^3\right)a_1 + \left(\sum_{i=1}^{n} x_i^4\right)a_2 = \sum_{i=1}^{n} x_i^2 y_i$$

- The solution of the system of equations gives the values of the coefficients $a0$, $a1$, and $a2$ of the polynomial $y = a_2 x_i^2 + a_1 x_i + a_0$ that best fits then data points $\{(x_i, y_i)\}_{i=1}^{n}$.
- The coefficients for higher-order polynomials are derived in the same way.

Interpolation

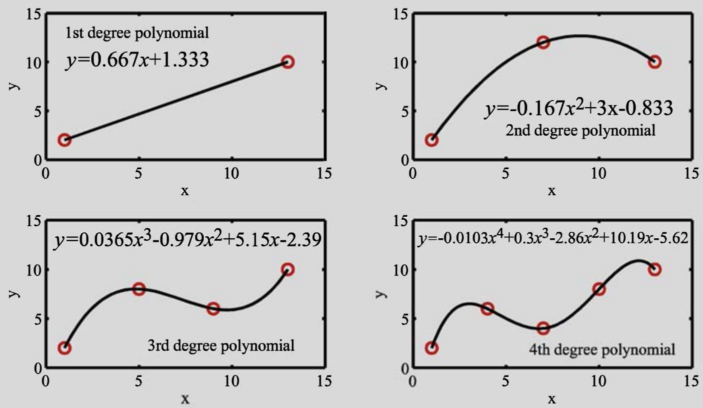
Interpolation

- Interpolation is a procedure for estimating a value between known val ues of data points.
- It is done by first determining a polynomial that gives the exact value at the data points, and then using the polynomial for calculating values between the points.



- When a small number of points is involved, a single polynomial might be sufficient for interpola tion over the whole domain of the data points.
- Often, however, when a large number of points are involved, different polynomials are used in the intervals between the points in a process that is called spline inter polation.

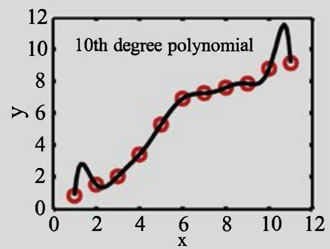
Interpolation



- As discussed in curve fitting, for any number of points n there is a polynomial of order $n-1$ that passes through all of the points.
- First, second, third, and fourth-order polynomials connect two, three, four, and five points, respectively.

Interpolation using a single polynomial

- Interpolation with a single polynomial gives good results for a small number of points.
- For a large number of points the order of the polynomial is high, and although the polynomial passes through all the points, it might deviate significantly between the points.



- Consequently, interpolation with a single polynomial might not be appropriate for a large number of points.
- For a large number of points, better interpolation can be done by using piecewise (spline) interpolation in which different lower-order polynomials are used for interpolation between different points of the same set of data points.

- For a given set of n points, only one (unique) polynomial of order m $(m = n - 1)$ passes exactly through all of the points.
- The polynomial, however, can be written in different mathematical forms.
- Three forms of polynomials are
  - Standard,
  - Lagrange, and
  - Newton's
- Standard form of an mth-order polynomial is:

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \ldots + a_1 x + a_0$$

- The coefficients in this form are determined by solving a system of $m + 1$ linear equations.
- The equations are obtained by writing the polynomial explicitly for each point (substituting each point in the polynomial). (Refer to curve fitting)

Lagrange Interplating Polynomials

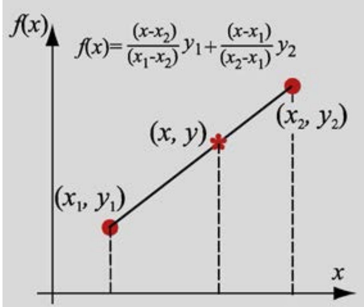
Lagrange Interplating Polynomials

- Lagrange interpolating polynomials are a particular form of polynomi als that can be written to fit a given set of data points by using the val ues at the points.
- The polynomials can be written right away and do not require any preliminary calculations for determining coefficients.
- Lagrange polynomials
  - First-order Lagrange polynomial
  - Second-order Lagrange polynomial
  - General form of an $n-1$ order Lagrange polynomial

First order Lagrange polynomial

- For two points, (x_1, y_1), and (x_2, y_2), the first-order Lagrange polynomial that passes through the points has the form:

$$f(x) = y = a_1(x - x_2) + a_2(x - x_1)$$



- Substitute the two point in the above equation

$$y_1 = a_1(x_1 - x_2) + a_2(x_1 - x_1) \quad \text{or} \quad a_1 = \frac{y_1}{(x_1 - x_2)}$$

$$y_2 = a_1(x_2 - x_2) + a_2(x_2 - x_1) \quad \text{or} \quad a_2 = \frac{y_2}{(x_2 - x_1)}$$

- Substitute a_1 and a_2 back

$$f(x) = \frac{(x - x_2)}{(x_1 - x_2)} y_1 + \frac{(x - x_1)}{(x_2 - x_1)} y_2$$

$$f(x) = \frac{(y_2 - y_1)}{(x_2 - x_1)} x + \frac{x_2 y_1 - x_1 y_2}{(x_2 - x_1)}$$

Second-order Lagrange polynomial

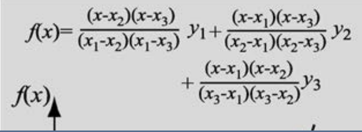
- For three points, (x_1, y_1), (x_2, y_2), and (x_3, y_3), the second-order Lagrange polynomial that passes through the points has the form:

$$f(x) = y = a_1(x - x_2)(x - x_3) + a_2(x - x_1)(x - x_3) + a_3(x - x_1)(x - x_2)$$

- Once the coefficients are determined such that the polynomial passes through the three points, the polynomial (quadratic form) is:

$$f(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} y_1 + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} y_3$$

- Above equation can also be rewritten in the standard form $f(x) = a_2 x^2 + a_1 x + a_0$. (Home work to derive the standard form)



General form of Lagrange polynomial

- the general formula of an $n-1$ order Lagrange polynomial that passes through n points (x_1, y_1), (x_2, y_2), ... , (x_n, y_n) is:

$$f(x) = \frac{(x-x_2)(x-x_3)\ldots(x-x_n)}{(x_1-x_2)(x_1-x_3)\ldots(x_1-x_n)} y_1 + \frac{(x-x_1)(x-x_3)\ldots(x-x_n)}{(x_2-x_1)(x_2-x_3)\ldots(x_2-x_n)} y_2 +$$

$$\ldots + \frac{(x-x_1)(x-x_2)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_1)(x_i-x_2)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)} y_i + \ldots +$$

$$\frac{(x-x_1)(x-x_2)\ldots(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)\ldots(x_n-x_{n-1})} y_n$$

$$f(x) = \sum_{i=1}^{n} y_i L_i(x) = \sum_{i=1}^{n} y_i \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{(x-x_j)}{(x_i-x_j)}$$

where $L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{(x-x_j)}{(x_i-x_j)}$ are called the Lagrange functions

Conclusive remarks

- The spacing between the data points does not have to be equal.
- For a given set of points, the whole expression of the interpolation polynomial has to be calculated for every value of x. In other words, the interpolation calculations for each value of x are independent of others.
- Different from other forms where once the coefficients of the polynomial are determined, they can be used for calculating different values of x.
- If an interpolated value is calculated for a given set of data points, and then the data set is enlarged to include additional points, all the terms of the Lagrange polynomial have to be calculated again.
- As discussed in next topic, this is different from Newton's polynomials where only the new terms have to be calculated if more data points are added.

Newton's Interpolating Polynomials

Newton's Interpolating Polynomials

- Newton's interpolating polynomials are a popular means of exactly fitting a given set of data points.
- The general form of an $n-1$ order Newton's polynomial that passes through n points is:

$$f(x) = a_1 + a_2(x - x_1) + a_3(x - x_1)(x - x_2) + \ldots + a_n(x - x_1)(x - x_2)\ldots(x - x_{n-1})$$

- The special feature of this form of the polynomial is that the coeffi cients a1 through an can be determined using a simple mathematical procedure.
- (Determination of the coefficients does not require a solu tion of a system of n equations.)
- Once the coefficients are known, the polynomial can be used for calculating an interpolated value at any x.

Newton's Interpolating Polynomials

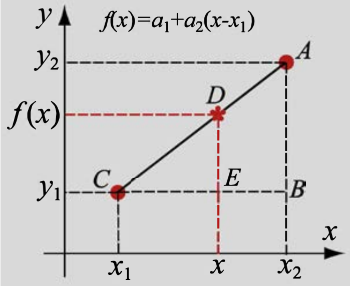
- Newton's interpolating polynomials have additional desirable features that make them a popular choice.
- The data points do not have to be in descending or ascending order, or in any order.
- Moreover, after the n coefficients of an $n-1$ order Newton's interpolating polynomial are determined for n given points, more points can be added to the data set and only the new additional coefficients have to be determined.

First-order Newton's polynomial

- For two given points, (x_1, y_1) and (x_2, y_2), the first-order Newton's polynomial has the form:

$$f(x) = a_1 + a_2(x - x_1)$$

It is an equation of a straight line that passes through the points.



- The coefficients a_1 and a_2 can be calculated by considering the similar triangles

$$\frac{DE}{CE} = \frac{AB}{CB}, \quad \text{or} \quad \frac{f(x) - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

$$f(x) = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

$$a_1 = y_1, \quad \text{and} \quad a_2 = \frac{y_2 - y_1}{x_2 - x_1}$$

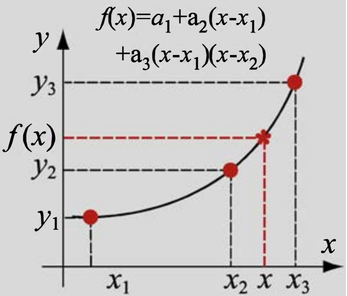
- The coefficient a_2 is the slope of the line that connects the two points.

Second-order Newton's polynomial

- For three given points, (x_1, y_1), (x_2, y_2), and (x_3, y_3), the second-order Newton's polynomial has the form:

$$f(x) = a_1 + a_2(x - x_1) + a_3(x - x_1)(x - x_2)$$

It is an equation of a parabola that passes through the three points.



- The coefficients a_1, a_2, and a_3 can be determined by substituting the three points in above equation.
- Substituting $x = x_1$ and $f(x1) = y1$ gives: $a_1 = y_1$
- Substituting the second point, $x = x_2$ and $f(x_2) = y_2$, (and $a_1 = y_1$) in above eq. gives:

$$y_2 = y_1 + a_2(x_2 - x_1) \quad \text{or} \quad a_2 = \frac{y_2 - y_1}{x_2 - x_1}$$

Second-order Newton's polynomial

- Substituting the third point, $x = x_3$ and $f(x_3) = y_3$ (as well as $a_1 = y_1$ and $a_2 = \frac{y_2 - y_1}{x_2 - x_1}$) in $f(x)$ that gives:

$$y_3 = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_3 - x_1) + a_3(x_3 - x_1)(x_3 - x_2)$$

Above equation can be solved for a_3 and rearranged to give (after some algebra):

$$a_3 = \frac{\frac{y_3 - y_2}{x_3 - x_2} - \frac{y_2 - y_1}{x_2 - x_1}}{(x_3 - x_1)}$$

- The coefficients a1, and a2 are the same in the first-order and sec ond-order polynomials. This means that if two points are given and a first-order Newton's polynomial is fit to pass through those points, and then a third point is added, the polynomial can be changed to be of sec ond-order and pass through the three points by only determining the value of one additional coefficient.

Third-order Newton's polynomial

- For four given points, (x_1, y_1), (x_2, y_2), (x_3, y_3) and (x_4, y_4), the third-order Newton's polynomial that passes through the four points has the form:

$$f(x) = y = a_1 + a_2(x - x_1) + a_3(x - x_1)(x - x_2) + a_4(x - x_1)(x - x_2)(x - x_3)$$

- The formulas for the coefficients a1, a2, and a3 are the same as for the second order polynomial. The formula for the coefficient a4 can be obtained by substituting (x_4, y_4), in Eq and solving for a_4, which gives:

$$a_4 = \frac{\dfrac{\left(\dfrac{y_4 - y_3}{x_4 - x_3} - \dfrac{y_3 - y_2}{x_3 - x_2}\right)}{(x_4 - x_2)} - \dfrac{\left(\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}\right)}{(x_3 - x_1)}}{(x_4 - x_1)}$$

A general form of Newton's polynomial and its coefficients

- There is common pattern in all equations that can be clarified by defining so called divided differences.
- For two points, (x_1, y_1), and (x_2, y_2), the first divided difference, written as $f[x2, xi]$, is defined as the slope of the line connecting the two points:

$$f[x_2, x_1] = \frac{y_2 - y_1}{x_2 - x_1} = a_2$$

- The first divided difference is equal to the coefficient a_2.

A general form of Newton's polynomial and its coefficients

- For three points (x_1, y_1), (x_2, y_2), and (x_3, y_3) the second divided difference, written as $f[x_3, x_2, x_1]$, is defined as the difference between the first divided differences of points (x_3, y_3), and (x_2, y_2), and points (x_2, y_2), and (x_1, y_1) divided by $(x_3 - x_1)$:

$$f[x_3, x_2, x_1] = \frac{f[x_3, x_2] - f[x_2, x_1]}{x_3 - x_1} = \frac{\frac{y_3 - y_2}{x_3 - x_2} - \frac{y_2 - y_1}{x_2 - x_1}}{(x_3 - x_1)} = a_3$$

- The second divided difference is thus equal to the coefficient a_3

A general form of Newton's polynomial and its coefficients

- For four points (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4) the third divided difference, written as $f[x_4, x_3, x_2, x_1]$, is defined as the differ ence between the second divided differences of points (x_2, y_2), (x_3, y_3) and (x_4, y_4), and points (x_1, y_1), (x_2, y_2), and (x_3, y_3) divided by $(x_4 - x_1)$:

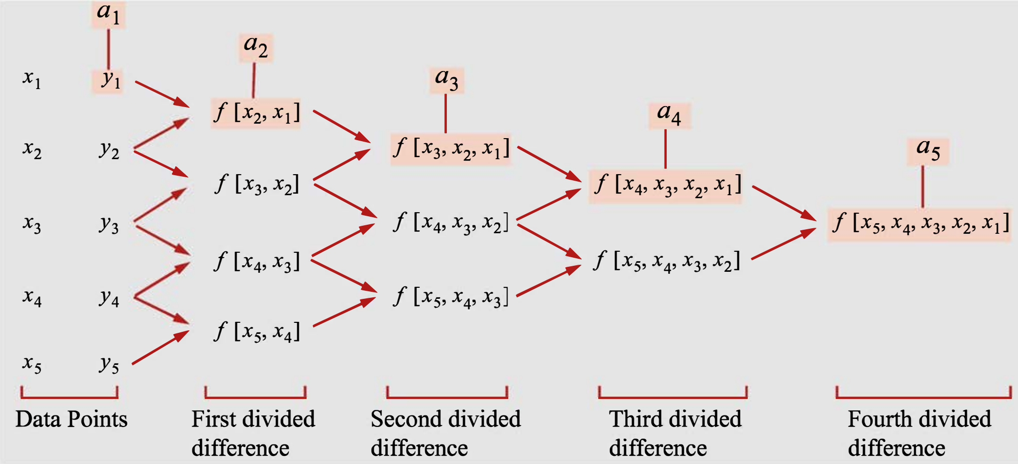
$$f[x_4, x_3, x_2, x_1] = \frac{f[x_4, x_3, x_2] - f[x_3, x_2, x_1]}{x_4 - x_1}$$

$$= \frac{\dfrac{f[x_4, x_3] - f[x_3, x_2]}{x_4 - x_2} - \dfrac{f[x_3, x_2] - f[x_2, x_1]}{x_3 - x_1}}{(x_4 - x_1)}$$

$$= \frac{\dfrac{\dfrac{y_4 - y_3}{x_4 - x_3} - \dfrac{y_3 - y_2}{x_3 - x_2}}{(x_4 - x_2)} - \dfrac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}}{(x_3 - x_1)}}{(x_4 - x_1)} = a_4$$

- The third divided difference is thus equal to the coefficient a_4.
- If more data points are given, the procedure for calculating higher dif ferences continues in the same manner.

A general form of Newton's polynomial and its coefficients



A general form of Newton's polynomial and its coefficients

- In general, when n data points are given, the procedure starts by calculating $(n-1)$ first divided differ ences. Then, $(n-2)$ second divided differences are calculated from the first divided differences. This is followed by calculating $(n-3)$ third divided differences from the second divided differences. The process ends when one nth divided difference is calculated from two $(n-1)$ divided differences to give the coefficient a_n.
- In general terms, for n given data points, (x_1, y_1), (x_2, y_2), ... , (x_n, y_n) , the first divided differences between two points (x_i, y_i), and (x_j, y_j) are given by:

$$f[x_j, x_i] = \frac{y_j - y_i}{x_j - x_i}$$

A general form of Newton's polynomial and its coefficients

- The kth divided difference for second and higher divided differences up to the $(n-1)$ difference is given by:

$$f[x_k, x_{k-1}, \ldots, x_2, x_1] = \frac{f[x_k, x_{k-1}, \ldots, x_3, x_2] - f[x_{k-1}, x_{k-2}, \ldots, x_2, x_1]}{x_k - x_1}$$

- With these definitions, the (n -1) order Newton's polynomial, Eq. (6. 46) is given by:

$$f(x) = y = \underbrace{y_1}_{a_1} + \underbrace{f[x_2, x_1]}_{a_2}(x-x_1) + \underbrace{f[x_3, x_2, x_1]}_{a_3}(x-x_1)(x-x_2) + \ldots + \underbrace{f[x_n, x_{n-1}, \ldots, x_2, x_1]}_{a_n}(x-x_1)(x-x_2)\ldots(x-x_{n-1})$$

Notes about Newton's polynomials

- The spacings between the data points do not have to be the same.
- For a given set of n points, once the coefficients a_1 through a_n are determined, they can be used for interpolation at any point between the data points.
- After the coefficients a_1 through a_n are determined (for a given set of n points), additional data points can be added (they do not have to be in order), and only the additional coefficients have to be determined.

Piecewise (Spline Interpolation)

References

[1] Amos Gilat and Vish Subramaniam. *Numerical Methods for Engineers and Scientists: An Introduction with Applications using MATLAB*. John Wiley & Sons, 2014.

Thank you!