

REPUBLIC OF THE PHILIPPINES
POLYTECHNIC UNIVERSITY OF THE PHILIPPINES
STA. MESA, MANILA

COLLEGE OF ENGINEERING
COMPUTER ENGINEERING DEPARTMENT



CMPE30063

NUMERICAL METHODS

INSTRUCTIONAL MATERIAL

ENGR. JOSHUA BENJAMIN RODRIGUEZ

Table of Contents

| | |
|-----------------------------------------------|----|
| 1. Approximation and Rounding Errors..... | 1 |
| 2. Roots and Optimization..... | 5 |
| a. Bracketing Method..... | 7 |
| i. Bisection Method..... | 7 |
| ii. False Position Method..... | 8 |
| b. Open Method..... | 9 |
| i. Fixed Point Iteration..... | 9 |
| ii. Newton Rhapsion Method..... | 9 |
| iii. Secant Method..... | 10 |
| c. Optimization..... | 11 |
| i. Golden Section Method..... | 13 |
| ii. Newton's method..... | 15 |
| iii. Multimodal functions..... | 16 |
| 3. Linear System..... | 18 |
| a. Introduction to Matrices..... | 19 |
| b. Non iterative method..... | 22 |
| i. Graphical method..... | 22 |
| ii. Determinant and Cramer's rule..... | 22 |
| iii. Naïve Gauss elimination..... | 23 |
| iv. Pivoting..... | 25 |
| v. Tridiagonal System..... | 26 |
| vi. Gauss Jordan method..... | 28 |
| vii. LU decomposition..... | 29 |
| viii. Inverse of matrix..... | 30 |
| ix. Cholesky Factorization/Decomposition..... | 31 |
| c. Iterative method | 32 |
| i. Gauss-Seidel method..... | 32 |
| ii. Jacobi iteration..... | 35 |
| 4. Curve Fitting | 37 |
| a. Introduction to statistic..... | 38 |
| b. Linear Least squares Regression..... | 39 |
| c. Polynomial Regression..... | 44 |
| d. Polynomial Interpolation..... | 47 |
| i. Newton Interpolating Polynomial..... | 48 |
| ii. Lagrange Interpolating Polynomial..... | 51 |
| Activities..... | 54 |

Module 1:
APPROXIMATION AND ROUNDING ERRORS

This module is all about the approximations and rounding errors that is being used for numerical analysis of equations and formulas.

Overview:

Numerical methods must be sufficiently exact and precise to satisfy the requirements of electrical engineering problems may it be theoretical (Ohm's law, Kirchoff's Voltage and Current Law, Fault Analysis) and practical (sizing and length of wire estimation, selecting of optimal brand of devices based on cost). Thus, it is important to differentiate the terminologies and its application first before going into the numerical analysis.

Module Objectives:

After successful completion of this module, you should be able to:

- Differentiate between accuracy, precision and bias
- Define the significant figures
- Determine different error types
- Conduct an error analysis for solutions and answers

Course Materials:

The term accuracy, precision and bias are terminologies use to describe the veracity or authenticity of a certain thing. For example, one must determine the accuracy of a certain news, by verifying to a trusted source to determine its degree of precision to the real issue without bias in either side. In the same way, that these terms are being used to determine the integrity of a certain calculation, on how it deviates from the real or actual numbers. This is crucial since engineering especially electrical engineering deals with such meticulous calculation.

These terms are often defined interchangeably, as this have very similar meaning in terms of its usage. To illustrate its difference, below is a four shooting board (Figure 1.1):

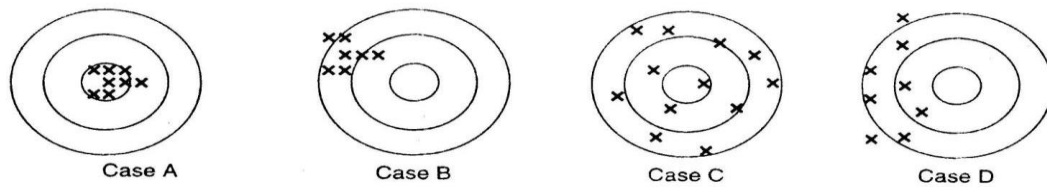


Figure 1.1 *Shooting boards showing Accuracy, Precision and Bias*

As shown, the following observation can be observed and can be summarize by Table 1:

- Case A is successful in terms of its accuracy (close to the target) and precision (how holes are close to each other).
- In Case B, the holes agree with each other (consistency or precision) but they deviate considerably from where the shooter was aiming (no correctness) and hence, it lacks exactness or inaccurate.
- Case C lacks both correctness (accuracy) and consistency (precision).
- Case D lacks consistency (precision).
- The shooters of targets C and D were imprecise since both of their holes are inconsistent with each other.

| Target | Bias | Precision | Accuracy |
|--------|-----------------|-----------|----------|
| A | None (unbiased) | High | High |
| B | High | High | Low |
| C | None (unbiased) | Low | Low |
| D | Moderate | Low | Low |

Table 1: Summary of Shooting Boards

From this example, **Accuracy** can be define as the degree to which the measurements deviate from the true value (in this case, the exactness of the hole to the target), **Precision** as the ability to give multiple estimates that are near to each other (how the holes agree with each other) and **Bias** is the difference between the center of the holes and the center of the target or a systematic deviation of values from the true value.

In regard to this, significant figures are being used to determine how accurate or precise is one solution to the real answer. For example, if 46.23 is exact to the four digits shown, it has four significant digits which means that the last digit is imprecise in which the error can be estimated to be no more than 0.005. The digits from 1 to 9 are always significant, with zero being significant where it is not being used to set the position of the decimal point. For example, numbers 2410, 2.41, 0.00241 has both three significant digits where 0 in 2410 is used only to set the decimal place. Thus, rounding off of answers should be made at the end of computation, not at intermediate calculation to avoid inaccuracy of the solution.

Scientific notation can then be used to avoid confusion in significant figures. For example, 2.41×10^3 has three significant digits while 2.410×10^3 has four significant digits. Thus, a mathematical operation using an imprecise digit is imprecise since it will affect the accuracy of the final answer.

In general, errors can be classified based on their sources such as non-numerical and numerical errors:

Non-Numerical errors can be classified as follows:

- a. Modelling errors which is generated by assumptions and limitations such as the exact value of pi.
- b. Blunders and mistakes which are attributed to human errors such as misreading of weighing scale.
- c. Uncertainty in information and data gathered.

Numerical errors can be classified as follows:

- a. Round off errors is due to a limited number of significant digits.
- b. Truncation errors is due to the shortened terms of a certain calculation such as the infinite Taylor series.
- c. Propagation errors is due to a sequence of operations which can reduced with a good computation order. In example, in summarizing several values, one can rank the values in ascending order before performing the summation.
- d. Mathematical approximation errors are due to estimating a certain model. For example, using a linear model for representing a nonlinear expression.

Error is defined as the difference between the computed and true values of number x. This can be defined by the equation below:

$$e = x_c - x_t$$

$$e_r = \frac{x_c - x_t}{x_t} = \frac{e}{x_t}$$

which: e- error x_c - computed value x_t - true value e_r =relative error

Error can be defined by the following terms:

- True value= approximation + absolute error
- Absolute error= true value-approximation
- Relative error=absolute error/true value

This can be summarized by the following formulas:

$$\mathcal{E}_t = \frac{\text{absolute error}}{\text{true value}} * 100$$

$$\mathcal{E}_a = \frac{\text{approximate error}}{\text{approximate value}} * 100 = \frac{(\text{actual} - \text{previous}) \text{ approximation}}{\text{actual approximation}} * 100$$

where: \mathcal{E}_t =true error \mathcal{E}_a =absolute error

To learn more about Approximation and Rounding errors, stick with these following resources:

Watch:

- Floating Point Representation and Rounding error | Chad Higdon-Topaz
(<https://www.youtube.com/watch?v=wbxSTxhTmrs>)

Read:

- Applied Numerical Methods with MatLAB for Engineers and Scientists by S.T. Chapra
Chapter 4 – Roundoff and Truncation errors

Module 2:

ROOTS AND OPTIMIZATION

This module is all about the roots and optimization that is being used to determine the solutions and answer to a particular objective function.

Overview:

Finding the roots of a certain equation is to determine its approximate value or solutions where it will converge to zero ($f(x)=0$). One of the conventional way to obtain approximate solution is to plot the function to determine where it crosses the x axis but such approach may inappropriate for some functions because of its lack of consistency to the true answer. Or one may try the gruesome trial and error method where one will guess the value of x and repeatedly evaluating $f(x)$ is zero to determine whether the new value provides a better estimate of the root. Such methods are inefficient for the requirements of electrical engineering. Thus, numerical methods provide means to approximate but employ strategic to determine the feasible solution to $f(x)=0$. Besides from finding the root, another point of interest is finding the optimal maximum and minimum values for each function. Back in calculus, such solutions can be obtained analytically by determining the value at which the function is flat that is when its derivative is zero. From a numerical analysis standpoint, this is very similar to the root location methods which both involve guessing and searching for a location on a function. The fundamental difference between the two types can be summarized by the figure below:

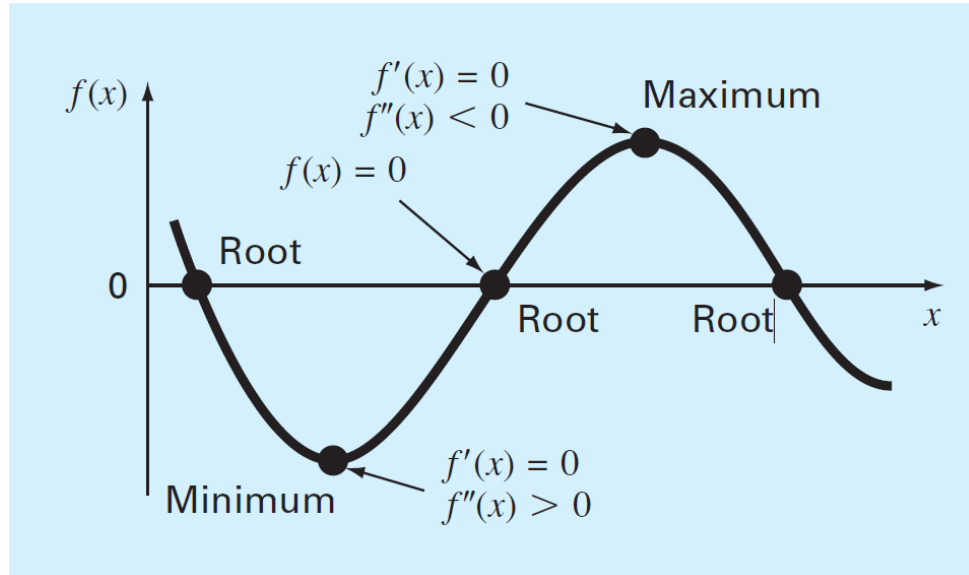


Figure 2.1: Difference between roots and optima using a single variable

Module Objectives:

After successful completion of this module, you should be able to:

- Utilize bracketing methods such as false position and bisection to determine the root of an equation.
- Utilize open methods such as Newton Rhapson, secant and fixed point to determine the root of an equation.
- Compare the difference between bracketing and open methods in terms of its accuracy and precision in find the root and convergence to the true value.
- Utilize optimization methods such as golden section search method to determine the optima of an equation.
- Utilize computer software such as Microsoft Excel for both root location and optimization.

Course Materials:

A simple approach of finding the estimate of the root of an equation $f(x)=0$ is to plot the said function in a graph and observe where it crosses the x axis and at this point which provides the rough approximation of the root.

Graphical techniques however have a limited practical value because they are imprecise especially in obtaining the actual values of a function. This however can be used to obtain rough estimate of roots. This is where methods with initial guesses comes in. It is where one must repeatedly make guesses until the function was sufficiently close to zero. The two major classes of methods available by the type of initial guess are:

- a. **Bracketing Method**- are based on two initial guesses that “bracket” or border the root.
- b. **Open methods** are methods which can involve one or more initial guesses, but there is no need for them to bracket the root.

Bracketing method as start with guesses, or contain, the root and then systematically reduce the width of the bracket. The two most common variation of this method are **Bisection Method** and **False Position Method**.

The **Bisection method** is a variation of the incremental search method in which the interval is always divided in half. If a function changes sign over an interval, the function value at the midpoint is evaluated (Figure 2.2). The location of the root is then determined as lying within the subinterval where the sign change occurs. The subinterval then becomes the interval for the next iteration. The process is repeated until the root is known to the required precision. It is given by equation 1 and condition:

$$x_r = \frac{x_l + x_u}{2} \quad (1) \quad \text{where: } x_l = \text{lower limit} \quad x_r = \text{median} \quad x_u = \text{upper limit}$$

The following conditions will be used for the succeeding iteration:

- If $f(x_l)f(x_r) > 0$, no sign change occurs between the lower bound and the midpoint. Thus, the root must be located between x_r and x_u .
- If $f(x_l)f(x_r) < 0$, the root is now lower than x_l and x_r .

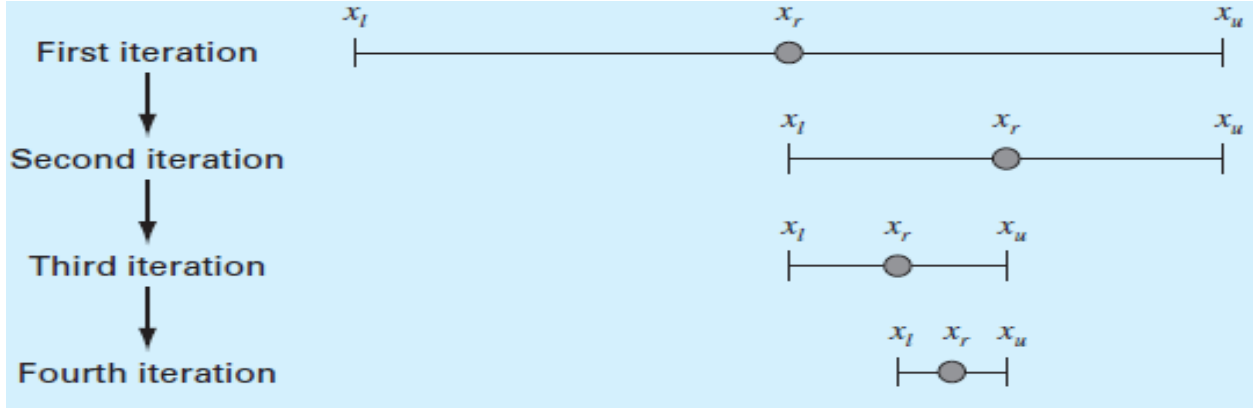


Figure 2.2: Graphical depiction of bisection method by successive iteration

False position (also called the linear interpolation method) is another well-known bracketing method. It is very similar to bisection with the exception that it uses a different strategy to come up with its new root estimate. Rather than bisecting the interval, it locates the root by joining $f(x_l)$ and $f(x_u)$ with a straight line in the figure. The intersection of this line with the x axis represents an improved estimate of the root. Thus, the shape of the function influences the new root estimate. Using similar triangles (Figure 2.3), the intersection of the straight line with the x axis can be estimated as for equation 2.

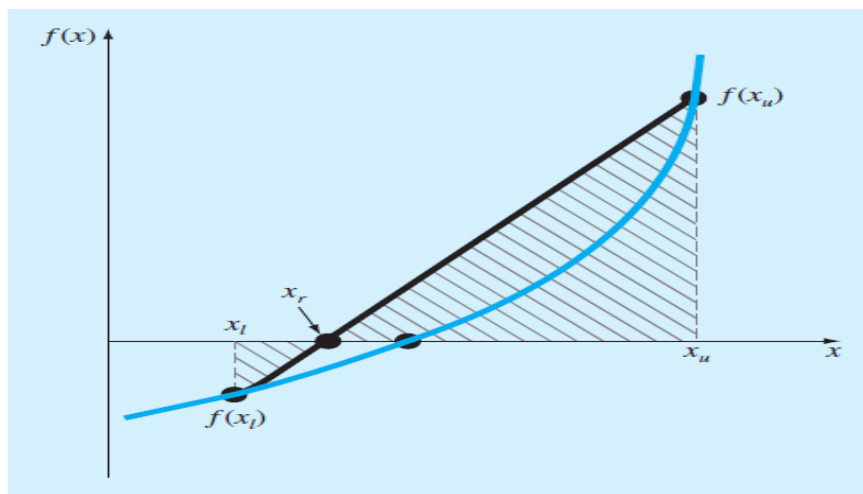


Figure 2.3: Graphical representation of False position method

by similar triangle, $\frac{x_u - x_r}{f(x_u)} = \frac{x_l - x_u}{f(x_l) - f(x_u)}$

thus, $x_r = x_u - \frac{f(x_u)(x_l - x_u)}{f(x_l) - f(x_u)}$ (2)

Open method will start with also with a guesses, but unlike **Bracketing Method**, it will only require single starting value or two starting values that do not necessarily bracket the root. As such, they sometimes diverge or move away from the true root as the computation progresses. The three most common variation of this method are **Fixed Point Iteration Method**, **Newton Rhapson Method** and **Secant Method**.

Fixed Point iteration (also called *one-point iteration* or *successive iteration*) is done by rearranging the function $f(x)=0$ so that x is on left-hand side of the equation such $x=g(x)$. This transformation can be accomplished either by algebraic manipulation or by simply adding x to both sides of the original equation such $x_{i+1}=g(x_i)$. Figure 2.4 illustrates the graphical illustration of fixed point iteration.

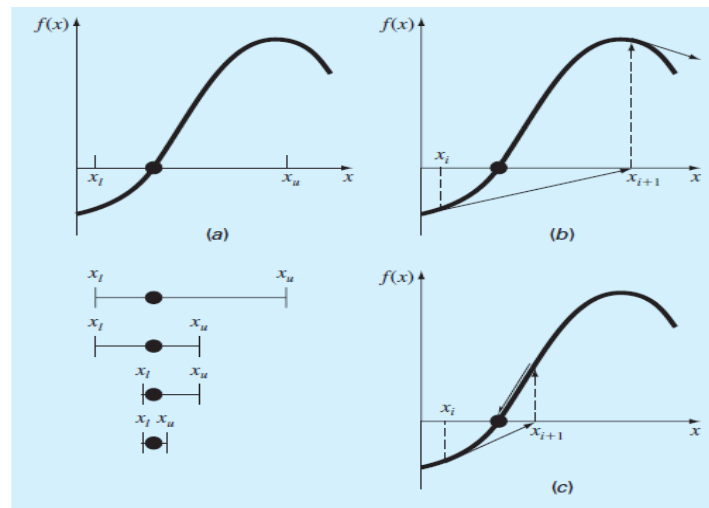


Figure 2.4: Graphical illustration of Fixed point iteration

Newton Rhapson method is perhaps the most widely used of all root-locating formulas. This can be illustrated that if initial guess at the root is x_i , a tangent can be extended from the

point $[x_i, f(x_i)]$ as shown in Figure 2.5. The point where this tangent crosses the x axis usually represents an improved estimate of the root. This method can be derived on the basis of geometrical interpretation of similar triangles. The first derivative at x is equivalent to the slope and is shown by equation 3:

$$f'(x_i) = \frac{f(x_i) - 0}{x_i - x_{i+1}} \quad \text{hence, } x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \quad (3)$$

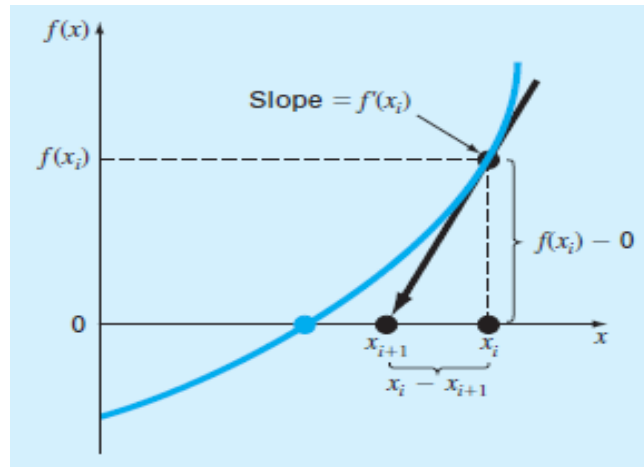


Figure 2.5: Graphical representation of Newton Raphson method

Secant method addresses the problem in implementing the Newton Raphson method which is the evaluation of the derivative. Although, this is not inconvenient for polynomials and many other functions, there are certain functions whose derivatives may be difficult or inconvenient to evaluate. For these cases the derivative can be approximated by a backward finite divided difference which is shown by equation 4.

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \quad (4)$$

The approximation in Newton Raphson can be substituted to yield the following iterative formula shown in equation 5:

$$x_{i+1} = x_i - \frac{f(x_i)(x_{i-1}-x_i)}{f(x_{i-1})-f(x_i)} \quad (5) \quad \text{which is the secant method}$$

An alternative approach involves a fractional perturbation of the independent variable to estimate $f'(x)$.

$$f'(x_i) = \frac{f(x_i + \delta x_i) - f(x_i)}{\delta x_i}$$

where δ = small perturbation fraction and it can be substituted to yield the equation:

$$x_{i+1} = x_i - \frac{\delta x_i f(x_i)}{f(x_i + \delta x_i) - f(x_i)} \quad \text{which is the **modified secant method**.}$$

Optimization or mathematical programming and root finding whether it is open or bracket method, both involve guessing and searching for a point on a function such as $f(x)$. However, the fundamental difference is that root finding focuses on searching for zeros of a function/s while optimization focuses on finding the minimum or the maximum of a function of several variables. It can be stated as:

Find x , which minimizes or maximizes $f(x)$ that is subject to,

$$d_i(x) \leq a_i \quad i = 1, 2, \dots, m^*$$

$$d_i(x) \leq a_i \quad i = 1, 2, \dots, m^*$$

Where x is an n -dimensional design vector, $f(x)$ is the objective function, $d_i(x)$ are inequality

Optimization problems can be classified on the basis of the form of $f(x)$:

- If $f(x)$ and the constraints are linear, we have linear programming
- If $f(x)$ is quadratic and the constraints are linear, we have quadratic programming.

- If $f(x)$ is not linear or quadratic and/or the constraints are non linear, we have nonlinear programming.

When equations(*) are included, we have a *constrained optimization* problem; otherwise, it is *unconstrained optimization* problem.

A **unimodal** function has a single maximum or a minimum in the a given interval as shown in Figure 2.6. For a *unimodal* function:

- ▶ First pick two points that will bracket your extremum $[x_l, x_u]$.
- ▶ Pick an additional third point within this interval to determine whether a maximum occurred.
- ▶ Then pick a fourth point to determine whether the maximum has occurred within the first three or last three points
- ▶ The key is making this approach efficient by choosing intermediate points wisely thus minimizing the function evaluations by replacing the old values with new values.

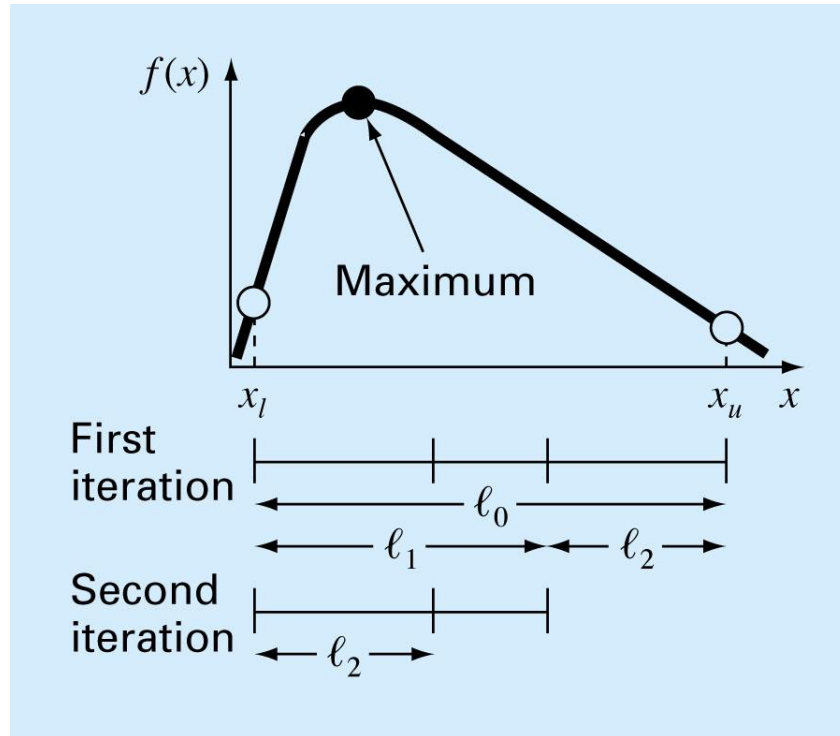


Figure 2.6: Graphical representation of unimodal function

$$l_0 = l_0 + l_1 \text{ thus, } \frac{l_1}{l_0} = \frac{l_2}{l_1}$$

- The first condition specifies that the sum of the two sub lengths l_1 and l_2 must equal the original interval length.
- The second say that the ratio of the length must be equal as shown in equation 3:

$$\frac{l_1}{l_1 + l_2} = \frac{l_2}{l_1} \quad \text{thus,} \quad R = \frac{l_2}{l_1} \quad (3)$$

$$1 + R = \frac{1}{R} \quad \text{thus,} \quad R^2 + R - 1 = 0$$

$$R = \frac{-1 + \sqrt{1 - 4(-1)}}{2} = \frac{\sqrt{5} - 1}{2} = 0.61803$$

The method starts with two initial guesses, x_l and x_u , that bracket one local extremum of $f(x)$:

- Next two interior points x_1 and x_2 are chosen according to the golden ratio as shown in equation 4.
- The function is then evaluated at these two interior points as shown in Figure 2.7.

$$d = \frac{\sqrt{5} - 1}{2}(x_u - x_l) \quad (4)$$

$$x_1 = x_l + d$$

$$x_2 = x_u - d$$

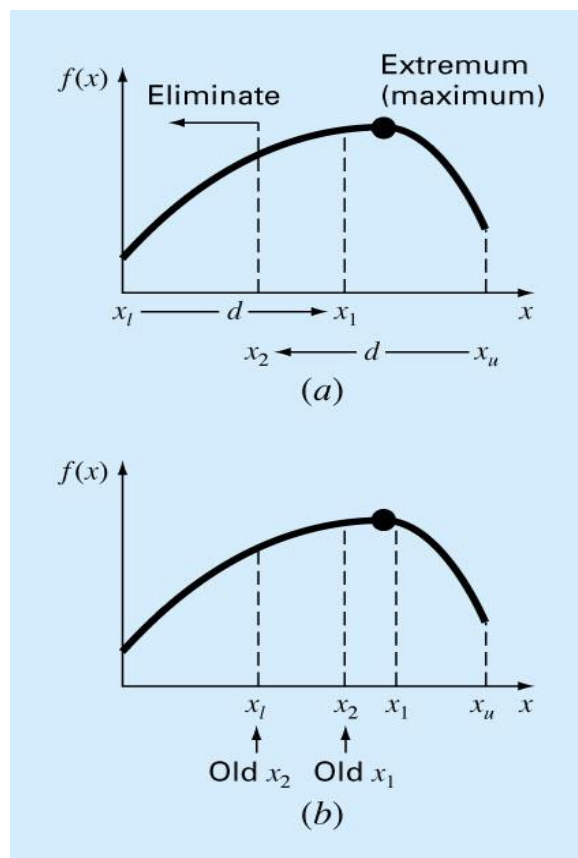


Figure 2.7: Graphical representation of an optimization using two initial guesses or golden search method

From this, two results can occur:

- ▶ If $f(x_1) > f(x_2)$ then the domain of x to the left of x_2 from x_l to x_2 , can be eliminated because it does not contain the maximum. Then, x_2 becomes the new x_l for the next round.
- ▶ If $f(x_2) > f(x_1)$, then the domain of x to the right of x_1 from x_l to x_2 , would have been eliminated. In this case, x_1 becomes the new x_u for the next round.

New x_i can now be determined as before by equation 4:

$$x_1 = x_l + \frac{\sqrt{5} - 1}{2} (x_u - x_l) \quad (4)$$

The real benefit from the use of golden ratio is because the original x_1 and x_2 were chosen using golden ratio, we do not need to recalculate all the function values for the next iteration.

Newton's Method is a similar approach to Newton- Raphson method can be used to find an optimum of $f(x)$ by defining a new function $g(x)=f'(x)$. Thus because the same optimal value x^* satisfies both:

$$f'(x^*) = g(x^*) = 0$$

We can use the following as a technique to the extremum of $f(x)$ as given by equation 5.

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)} \quad (5)$$

Newton's Method algorithm can be summarized by following step:

Initialization. Determine a reasonably good estimate for the maxima or the minima of the function $f(x)$.

Step 1. Determine $f'(x)$ and $f''(x)$

Step 2. Substitute x_i (initial estimate x_0 for the first iteration) $f'(x)$ and into $f''(x)$.

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

to determine x_{i+1} and the function value in iteration i .

Step 3. If the value of the first derivative of the function is zero then you have reached the optimum (maxima or minima). Otherwise, repeat Step 2 with the new value of x_i .

In **Multimodal functions**, both local and global optima can occur. In almost all cases, we are interest in finding the absolute highest or lowest value of a function as shown by cartesian and contour graph on Figure 2.8.

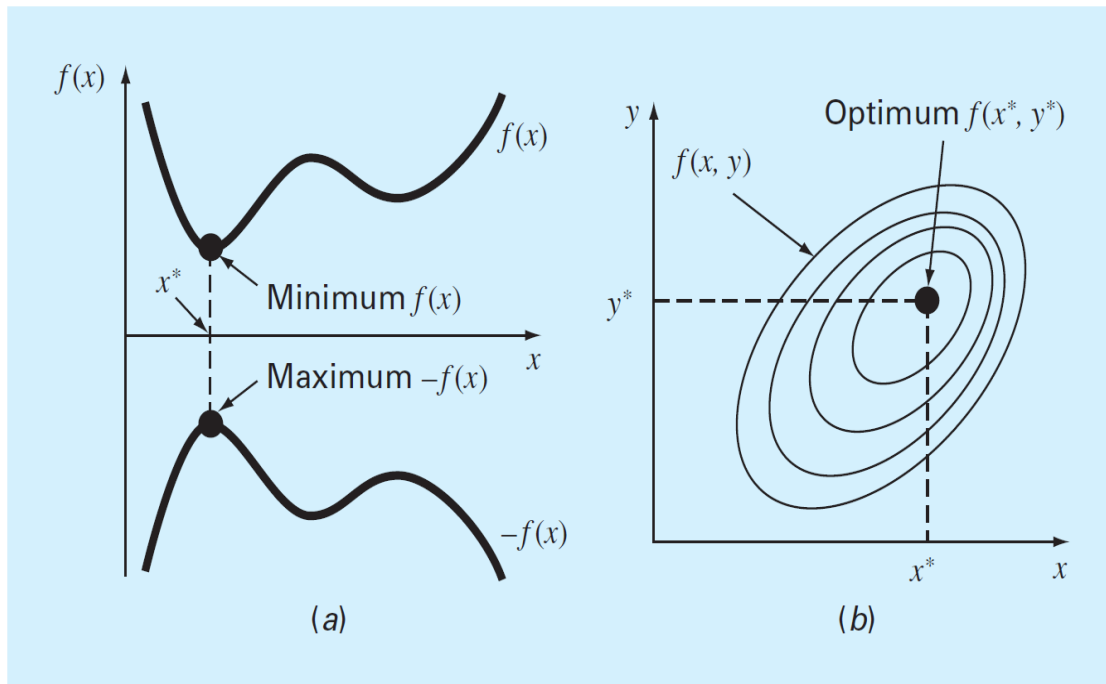


Figure 2.8: Representation of finding global maxima and minima using plot (a) and contour (b) graph

Global optimum can be distinguished from local optimum by:

- By graphing to gain insight into the behavior of the function (Figure 2.9).

- Using randomly generated starting guesses and picking the largest of the optima as global.
- Perturbating the starting point to see if the routing returns a better point or the same local minimum.

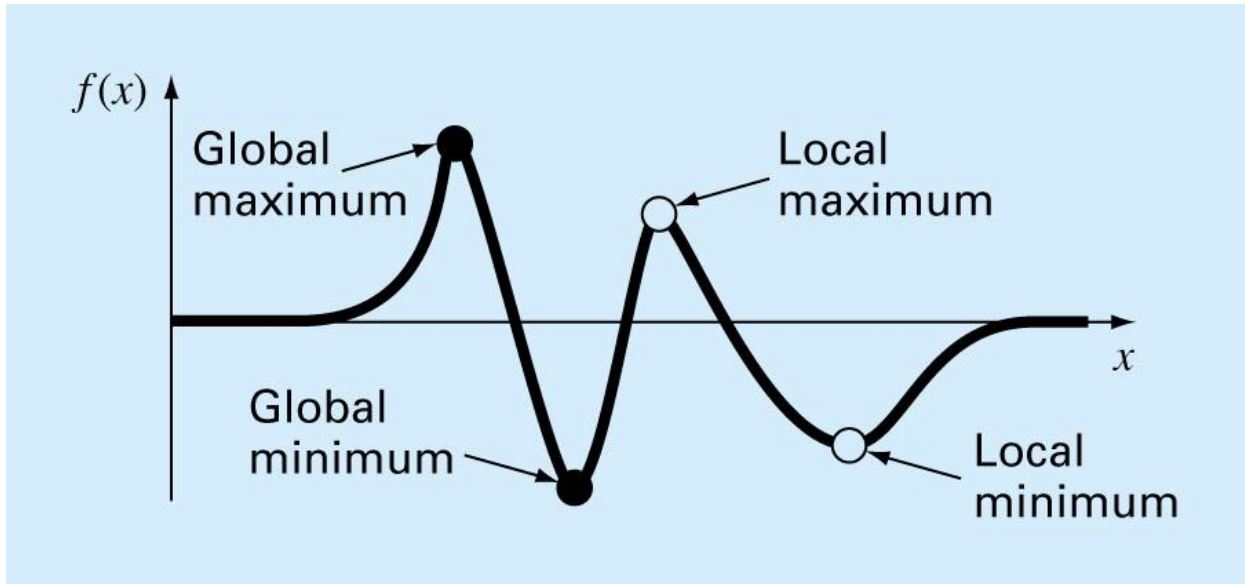


Figure 2.9: Graphical representation of Global and Local maximum

To learn more about Roots and Approximation, stick with these following resources:

Watch:

- MATLAB Introduction: Root Solving | Jonathan Curie
(<https://www.youtube.com/watch?v=4n4eXQLHGlg>)
- How to locate a root | ExamSolutions
(<https://www.youtube.com/watch?v=OzFuihxtbtA>)

Read:

- Applied Numerical Methods with MatLAB for Engineers and Scientists Chapter 5,6, and 7 by S.T. Chapra
- Applied Numerical Methods using MatLAB Chapter 4 by Won Young Yang, Wenwu Cao, Tae-Sang Chung and John Morris

Module 3:

LINEAR SYSTEM

This module is all about the system of linear algebraic equations which can be either linear or nonlinear in nature.

Overview:

Many of the fundamental equations of electrical engineering are based on conservation law and some of the familiar quantities that conform to such laws are mass, energy and momentum. This can lead to balance or continuity equations that relate system behavior as represented by the levels or response of the quantity being modeled to the properties or characteristics of the system and the external stimuli or forcing functions acting on the system. When these dependencies are expressed mathematically, the resulting equations often of the linear algebraic form where x 's are usually measures of the magnitudes of the responses of the individual components. Aside from physical systems, simultaneous linear algebraic equations also arise in a variety of mathematical problem contexts. These result when mathematical functions are required to satisfy several conditions simultaneously. Each condition results in an equation that contains known coefficients and unknown variables

Module Objectives:

After successful completion of this module, you should be able to:

- Define matrix and its use in linear algebraic equations
- Discuss the importance of Gauss elimination and its application on solving simultaneous linear equation
- Formulate and discuss LU factorization from Gauss elimination.
- Describe how LU factorization can be employed to efficiently calculate the matrix inverse.
- Describe different iterative solution techniques in solving linear algebraic problem.

Course Materials:

Linear algebraic equations are of the general form as follows:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n$$

Where the a's are constant coefficients, the b's are constants, the x's are the unknowns, and n is the number of equation which can be simplified in the form of matrix.

A matrix consists of a rectangular array of elements represented by a single symbol. As depicted in Figure 3.1, [A] is the shorthand notation for the matrix and a_{ij} designates an individual element of the matrix. A horizontal set of elements is called a row and a vertical set is called a column. The first subscript i always designates the number of the row in which the element lies. The second subscript j designates the column. For example, element a_{23} is in row 2 and column 3. The matrix has m rows and n columns and is said to have a dimension of m by n (or $m \times n$). It is referred to as an m by n matrix.

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

Figure 3.1: Example of a matrix A

Matrices where $m = n$ are called *square matrices*. The diagonal consisting of the elements a_{11} , a_{22} , and a_{33} is termed the *principal* or *main diagonal* of the matrix. For example, a 3×3 matrix is:

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Asymmetric matrix is one where the rows equal the columns. that is, $a_{ij} = a_{ji}$ for all i 's and j 's.

An *identity matrix* is a diagonal matrix where all elements on the main diagonal are equal to 1, as in:

$$[I] = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

The identity matrix has properties similar to unity. That is, $[A][I] = [I][A] = [A]$

An *upper triangular matrix* is one where all the elements below the main diagonal are zero, as in:

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ & a_{22} & a_{23} \\ & & a_{33} \end{bmatrix}$$

A *lower triangular matrix* is one where all elements above the main diagonal are zero, as in:

$$[A] = \begin{bmatrix} a_{11} & & \\ a_{21} & a_{22} & \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Figure 3.2 shows the visual Depiction of how the rows and columns line up in matrix multiplication:

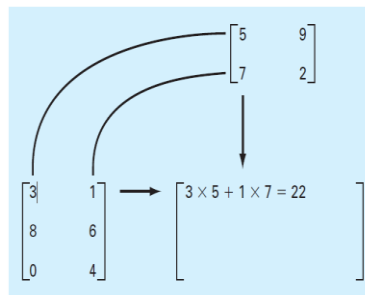
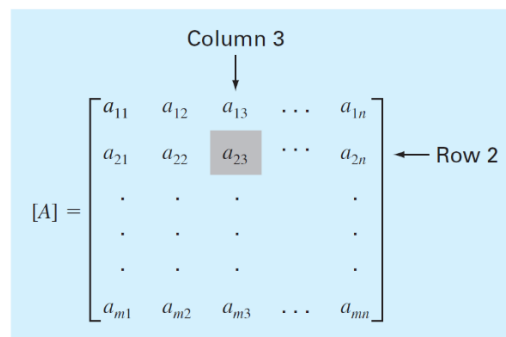


Figure 3.2: Row and columns in a matrix

Matrix Multiplication can be performed only if the inner dimensions are equal:



The product of two matrices is represented as $[C] = [A][B]$, where the elements of $[C]$ are defined using equation 1:

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \quad (1)$$

Although multiplication is possible, matrix division is not a defined operation. However, if a matrix $[A]$ is square and nonsingular, there is another matrix $[A]^{-1}$, called the *inverse* of $[A]$, for which

$$[A][A]^{-1} = [A]^{-1}[A] = [I]$$

Thus, the multiplication of a matrix by the inverse is analogous to division, in the sense that a number divided by itself is equal to 1. That is, multiplication of a matrix by its inverse leads to the identity matrix.

1. Graphical Method

A graphical solution is obtainable for two linear equations by plotting them on Cartesian coordinates with one axis corresponding to x_1 and the other to x_2 . Because the equations are linear, each equation will plot as a straight line. As shown in Figure 3.3

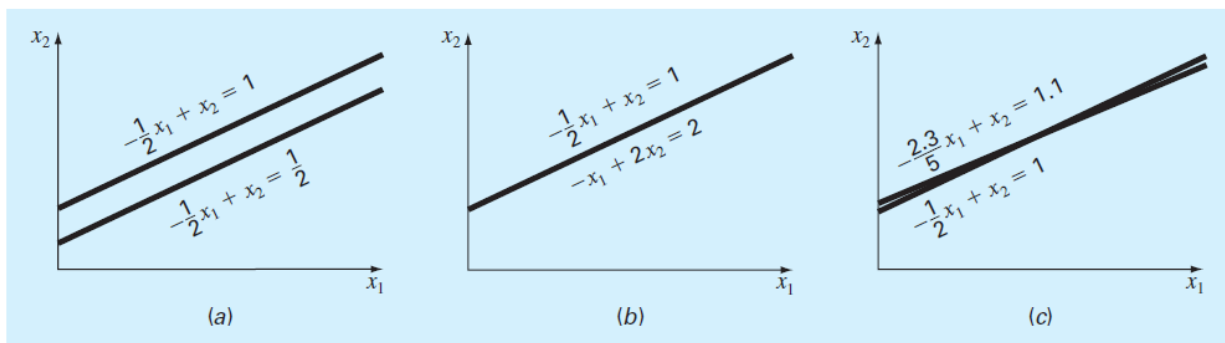


Figure 3.3: Graphical depiction of singular and ill conditioned system (a) no solution (b) infinite solutions (c) ill conditioned system where the slopes are close that the point of intersection is difficult to detect visually.

2. Determinants and Cramer Rule

The determinant can be illustrated for a set of three equations and the determinant of coefficient matrix A: $[A]\{x\}=\{b\}$

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

This rule states that each unknown in a system of linear algebraic equations may be expressed as a fraction of two determinants with denominator D and with the numerator obtained from D by replacing the column of coefficients of the unknown in question by the constants b_1, b_2, \dots, b_n . For example, for three equations, x_1 would be computed as:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{D}$$

3. Naïve Gauss elimination

The elimination of unknowns was used to solve a pair of simultaneous equations. The procedure consisted of two steps.

1. The equations were manipulated to eliminate one of the unknowns from the equations and reducing the entire system to upper triangular form. The result of this elimination step was that we had one equation with one unknown.
2. Consequently, this equation could be solved directly and the result back-substituted starting from the last variable into one of the original equations to solve for the remaining unknown.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \Rightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}' & a_{23}' \\ 0 & 0 & a_{33}' \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2' \\ b_3' \end{bmatrix}$$

For Forward elimination:

$$\begin{array}{l} \text{To eliminate } x_1 \\ \left. \begin{array}{l} a_{ij} \leftarrow a_{ij} - \left(\frac{a_{i1}}{a_{11}} \right) a_{1j} \quad (1 \leq j \leq n) \\ b_i \leftarrow b_i - \left(\frac{a_{i1}}{a_{11}} \right) b_1 \end{array} \right\} 2 \leq i \leq n \end{array}$$

$$\begin{array}{l} \text{To eliminate } x_2 \\ \left. \begin{array}{l} a_{ij} \leftarrow a_{ij} - \left(\frac{a_{i2}}{a_{22}} \right) a_{2j} \quad (2 \leq j \leq n) \\ b_i \leftarrow b_i - \left(\frac{a_{i2}}{a_{22}} \right) b_2 \end{array} \right\} 3 \leq i \leq n \end{array}$$

$$\begin{array}{l} \text{To eliminate } x_k \\ \left. \begin{array}{l} a_{ij} \leftarrow a_{ij} - \left(\frac{a_{ik}}{a_{kk}} \right) a_{kj} \quad (k \leq j \leq n) \\ b_i \leftarrow b_i - \left(\frac{a_{ik}}{a_{kk}} \right) b_k \end{array} \right\} k+1 \leq i \leq n \end{array}$$

Continue until x_{n-1} is eliminated.

For Backward substitution:

$$x_n = \frac{b_n}{a_{n,n}}$$

$$x_{n-1} = \frac{b_{n-1} - a_{n-1,n}x_n}{a_{n-1,n-1}}$$

$$x_{n-2} = \frac{b_{n-2} - a_{n-2,n}x_n - a_{n-2,n-1}x_{n-1}}{a_{n-2,n-2}}$$

$$x_i = \frac{b_i - \sum_{j=i+1}^n a_{i,j}x_j}{a_{i,i}}$$

4. Pivoting

- The primary reason that the foregoing technique is called 'naïve' is that during both the elimination and the back-substitution phases, it is possible that a division by zero can occur. The normalization of the first row would involve division by $a_{11} = 0$. Problems may also arise when the pivot element is close, rather than exactly equal, to zero because if the magnitude of the pivot element is small compared to the other elements, then round-off errors can be introduced.
- Therefore, before each row is normalized, it is advantageous to determine the coefficient with the largest absolute value in the column below the pivot element. The rows can then be switched so that the largest element is the pivot element. This is called *partial pivoting*.
- One of the problems of Naïve Gaussian Elimination is it may fail for very simple cases such as given by the matrix below in which the pivoting element is zero.

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

- Another thing is when the very small pivoting element may result in serious computation errors such as given by matrix below.

$$\begin{bmatrix} 10^{-10} & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

- Index vectors are used because it is much easier to exchange a single index element compared to exchanging the values of a complete row.
- In practical problems with very large N, exchanging the contents of rows may not be practical.

A good solution is when given $Ax=B$ in which x is a solution if $Ax-B=0$. Residual vector can be computed by equation (2):

$$R=Ax-B \quad (2)$$

However, due to rounding error, R may not be always zero. Thus, the solution is acceptable if

$$\text{Max } |r_i| \leq \epsilon_i.$$

5. Tridiagonal Systems

- The non-zero elements are in the main diagonal (red), super diagonal (blue) and subdiagonal (brown) as shown in the figure below.

$$a_{ij} = 0 \quad \text{if } |i - j| > 1$$

$$\begin{bmatrix} 5 & 1 & 0 & 0 & 0 \\ 3 & 4 & 1 & 0 & 0 \\ 0 & 2 & 2 & 2 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

- Tridiagonal systems occur in many applications.
- Selection of pivoting rows is unnecessary (under some conditions)
- Efficiently solved by Gaussian elimination.

Algorithms to solve Tridiagonal Systems

- Based on Naive Gaussian elimination.
- As in previous Gaussian elimination algorithms
 - Forward elimination step
 - Backward substitution step
- Elements in the super diagonal are not affected.
- Elements in the main diagonal, and B need updating

All the a elements will be zeros, need to update the d and b elements

The c elements are not updated

$$\begin{bmatrix} d_1 & c_1 & & & \\ a_1 & d_2 & c_2 & & \\ & a_2 & d_3 & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ & & & a_{n-1} & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} \Rightarrow \begin{bmatrix} d_1 & c_1 & & & \\ & d'_2 & c_2 & & \\ & & d'_3 & \ddots & \\ & & & \ddots & c_{n-1} \\ & & & & d'_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b'_2 \\ b'_3 \\ \vdots \\ b'_n \end{bmatrix}$$

Matrix A is diagonally dominant if:

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}| \quad \text{for } (1 \leq i \leq n)$$

The magnitude of each diagonal element is larger than the sum of elements in the corresponding row.

Examples :

$$\begin{bmatrix} 3 & 0 & 1 \\ 1 & 6 & 1 \\ 1 & 2 & -5 \end{bmatrix}$$

Diagonally dominant

$$\begin{bmatrix} -3 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Not Diagonally dominant

A tridiagonal system is diagonally dominant if:

$$|d_i| > |c_i| + |a_{i-1}| \quad (1 \leq i \leq n)$$

Forward elimination preserves diagonal dominance.

Forward Elimination

$$d_i \leftarrow d_i - \left(\frac{a_{i-1}}{d_{i-1}} \right) c_{i-1}$$

$$b_i \leftarrow b_i - \left(\frac{a_{i-1}}{d_{i-1}} \right) b_{i-1} \quad 2 \leq i \leq n$$

Backward Substitution

$$x_n = \frac{b_n}{d_n}$$

$$x_i = \frac{1}{d_i} (b_i - c_i x_{i+1}) \quad \text{for } i = n-1, n-2, \dots, 1$$

6. Gauss Jordan Method

- The method reduces the general system of equations $AX=B$ to $IX=B$ where I is an identity matrix.
- Only Forward elimination is done and no backward substitution is needed.
- It has the same problems as Naive Gaussian elimination and can be modified to do partial scaled pivoting.
- It takes 50% more time than Naive Gaussian method.

LU Decomposition

- Another way of solving a system of equations is by using a factorization technique for matrices called LU decomposition.
- This factorization involves two matrices, one lower triangular matrix and one upper triangular matrix.
- LU factorization methods separate the time-consuming elimination of the matrix $[A]$ from the manipulations of the right-hand-side $[b]$.
- Once $[A]$ has been factored (or decomposed), multiple right-hand-side vectors can be evaluated in an efficient manner.

$$L = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mm} \end{bmatrix} \quad U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ 0 & u_{22} & \cdots & u_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{mm} \end{bmatrix}$$

Where, $A=LU$

$$[A] = [L][U] = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

- LU factorization involves two steps as shown in Figure 3.4:
 - Factorization to decompose the $[A]$ matrix into a product of a lower triangular matrix $[L]$ and an upper triangular matrix $[U]$. $[L]$ has 1 for each entry on the diagonal.
 - Substitution to solve for $\{x\}$
 - Gauss elimination can be implemented using LU factorization

$$[A][B]=[I]=[B][A]$$

LU Decomposition be used to find the inverse of a matrix. Assume the first column of $[B]$ to be $[b_{11} \ b_{12} \ \dots \ b_{n1}]^T$. Using this and the definition of matrix multiplication

First column of $[B]$

$$[A] \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Second column of $[B]$

$$[A] \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

The remaining columns in $[B]$ can be found in the same manner and this can be verified by the following operation,

$$[A][A]^{-1}=[I]=[A]^{-1}[A]=1$$

Cholesky Factorization/Decomposition

- If A is a real, symmetric and positive definite matrix then there exists a unique lower triangular matrix L with positive diagonal element such that $A=LL^T=[U]^T[U]$.
- Symmetric systems occur commonly in both mathematical and engineering/science problem contexts, and there are special solution techniques available for such systems.
- The Cholesky factorization is one of the most popular of these techniques, and is based on the fact that a symmetric matrix can be decomposed as $[A]=[U]^T[U]$, where T stands for transpose.
- The rest of the process is similar to LU decomposition and Gauss elimination, except only one matrix, $[U]$, needs to be stored.
- The terms can be multiplied out and set equal to each other. The factorization can be generated efficiently by recurrence relations as by the equation 3 below for i th row.

$$u_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} u_{ki}^2} \quad (3a)$$

$$u_{ij} = \frac{a_{ij} - \sum_{k=1}^{i-1} u_{ki} u_{kj}}{u_{ii}} \quad \text{for } j = 1 + 1, \dots, n \quad (3b)$$

- After obtaining the factorization, it can be used to determine a solution for a right hand side vector (b) in a manner similar to LU factorization. First, an intermediate vector {d} is created by solving

$$[U]^T[d]=[b]$$

Then, the final solution can be obtained by solving,

$$[U][x]=[d]$$

Gauss-Seidel Method

Gauss-Seidel Method is a method being used to solve linear equations using iteration or successive substitution. Its basic procedure is as follows:

- Algebraically solve each linear equation for x_i
- Assume an initial guess solution array
- Solve for each x_i and repeat
- Use absolute relative approximate error after each iteration to check if error is within a pre-specified tolerance.

A set of n equations and n unknowns:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2$$

$$a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n = b_n$$

If: the diagonal elements are non-zero Rewrite each equation solving for the corresponding unknown

ex: First equation, solve for x_1 Second equation, solve for x_2

Rewriting each equation

$$x_1 = \frac{c_1 - a_{12}x_2 - a_{13}x_3 \dots - a_{1n}x_n}{a_{11}} \quad \leftarrow \text{From Equation 1}$$

$$x_2 = \frac{c_2 - a_{21}x_1 - a_{23}x_3 \dots - a_{2n}x_n}{a_{22}} \quad \leftarrow \text{From equation 2}$$

$$\vdots \quad \vdots \quad \vdots$$

$$x_{n-1} = \frac{c_{n-1} - a_{n-1,1}x_1 - a_{n-1,2}x_2 \dots - a_{n-1,n-2}x_{n-2} - a_{n-1,n}x_n}{a_{n-1,n-1}} \quad \leftarrow \text{From equation n-1}$$

$$x_n = \frac{c_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{nn}x_{n-1}}{a_{nn}} \quad \leftarrow \text{From equation n}$$

The general form of each equation can be written and for any row 'i',

$$x_1 = \frac{c_1 - \sum_{\substack{j=1 \\ j \neq 1}}^n a_{1j}x_j}{a_{11}} \quad x_{n-1} = \frac{c_{n-1} - \sum_{\substack{j=1 \\ j \neq n-1}}^n a_{n-1,j}x_j}{a_{n-1,n-1}} \quad x_n = \frac{c_n - \sum_{\substack{j=1 \\ j \neq n}}^n a_{nj}x_j}{a_{nn}}$$

$$x_i = \frac{c_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j}{a_{ii}}, i = 1, 2, \dots, n.$$

$$x_2 = \frac{c_2 - \sum_{\substack{j=1 \\ j \neq 2}}^n a_{2j}x_j}{a_{22}}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}$$

Solving the unknowns will involve assuming initial guess such as,

Use rewritten equations to solve for each value of x_i .

It is important to remember to use the most recent value of x_i . Which means to apply values calculated to the calculations remaining in the **current** iteration.

Absolute Relative approximate error can be computed by equation 4:

$$|\epsilon_a|_i = \left| \frac{x_i^{new} - x_i^{old}}{x_i^{new}} \right| \times 100 \quad (4)$$

The iterations are stopped when the absolute relative approximate error is less than a prespecified tolerance for all unknowns.

A pitfall of the Gauss-Seidel method that not all systems of equations will converge. One class of system of equations always converges: One with a *diagonally dominant* coefficient matrix.

[A] in [A] [X] = [C] is diagonally dominant if the following conditions are met:

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}| \text{ for all 'i' and } |a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}| \text{ for at least one 'i'}$$

Jacobi Iteration

A somewhat different approach in solving linear equations using iterative methods aside from Gauss Seidel method. Rather than using the latest available x 's, this technique uses successive substitution per iteration to compute a set of new x 's on the bases of a set of old x 's. Thus, as new values are generated, they are not immediately used but rather are retained for the next iteration.

The difference between the Gauss Seidel method and Jacobi iteration can be summarized in Figure 3.5.

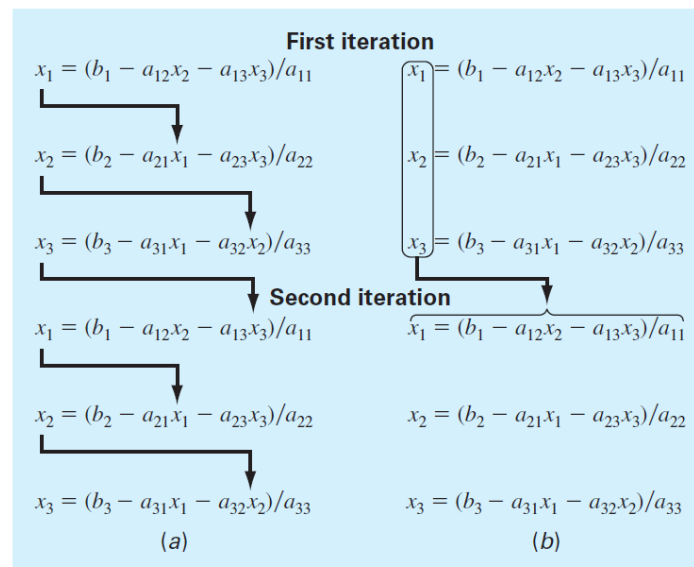


Figure 3.5: Fundamental difference between (a) Gauss Seidel method and (b) Jacobi iteration

To learn more about Linear System, stick with these following resources:

Watch:

Solving linear systems by substitution | Algebra Basics | Khan Academy
<https://www.youtube.com/watch?v=V7H1oUHXPkg>

- Gaussian elimination | Lecture 10 | Matrix Algebra for Engineers | Jeffrey Chasnov
<https://www.youtube.com/watch?v=RgnWMBpQPXk>

- Numerical Solutions of Linear Systems - Cholesky Decomposition/Factorisation | The Math Guy
(<https://www.youtube.com/watch?v=CzuWrFVZeIY>)
- Newton's method for solving nonlinear systems of Algebraic equations | The Math Guy
(https://www.youtube.com/watch?v=zPDp_ewoyhM)

Read:

- Applied Numerical Methods with MatLAB for Engineers and Scientists Chapter 8,9,10,11 and 12 by S.T. Chapra
- Applied Numerical Methods using MatLAB Chapter 2 by Won Young Yang, Wenwu Cao, Tae-Sang Chung and John Morris

Module 4:

CURVE FITTING

This module is all about the curve fitting methods that is being used for discrete values along a certain continuum and its application in engineering and science.

Overview:

The basic approach to fit a curve or a series of curves that pass directly through each of the points is when the data are known to be precise. Curve fitting is the process of constructing a curve or mathematical functions, which possess closest proximity to the series of data. Although many of the widely used engineering and scientific properties have been tabulated, there are great many more that are not available in tables. Special cases and new problem contexts often require that you measure your own data and develop your own predictive relationships. The attempts to find the best fit curve through a certain set of data can be seen on Figure 4.1 in which (a) and (b) falls under the category of Positive correlation while (c) has no correlation.

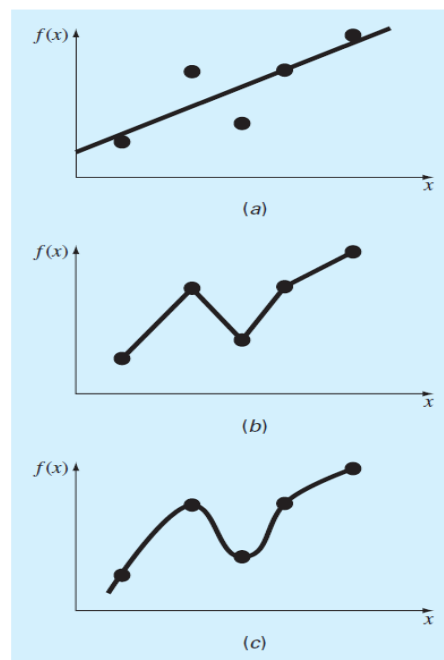


Figure 4.1: Three attempts to fit certain data through a curve (a) least square regression (b) linear interpolation (c) curvilinear interpolation

The type of application which are generally encountered when fitting a set of experimental data are:

- a. Trend analysis- represents the process of using the pattern to make predictions such as weather and load forecasting.
- b. Hypothesis testing—existing mathematical model is compared with measured data.

Module Objectives:

After successful completion of this module, you should be able to:

- Compute the slope and intercept of a best fit straight line with linear regression.
- Implement polynomial and multiple linear regression.
- Formulate general linear least-squares model.
- Evaluate polynomial coefficients with simultaneous equations in which is ill-conditioned problem.
- Perform an interpolation with Newton's and Lagrange polynomial.
- Solve an inverse interpolation problem.

Course Materials:

Before divulging into regression analysis, we will first review basic concepts on its mathematical background most especially in statistics.

Arithmetic mean is the sum of the individual data points (y_i) divided by the number of points (n) and is the measure of the central tendency as shown in equation 1.

$$\hat{y} = \frac{\sum y_i}{n}, \quad i = 1, \dots, n \quad (1)$$

Standard deviation is the most common measure of spread for a sample as shown in equation 2.

$$S_y = \sqrt{\frac{S_t}{n-1}}, \quad S_t = \sum (y_i - \hat{y})^2 \quad (2)$$

Variance represents the spread by the square of the standard deviation as shown in equation 3.

$$S_y^2 = \frac{\sum (y_i - \hat{y})^2}{n-1} = \frac{\frac{\sum y_i^2 - (\sum y_i)^2}{n}}{n-1} \quad (3)$$

Coefficient of variation has the utility to quantify the spread of data as shown in equation 4.

$$c.v. = \frac{S_y}{\hat{y}} \times 100 \quad (4)$$

Linear Least-Squares Regression

One approach to do best curve-fitting strategy is to derive an approximating function that fits the shape or general trend of the data without necessarily matching the individual points. This is to visually inspect the plotted data and then sketch a best line through the points. If points define a perfect straight line, different analysts would draw different lines and to remove this subjectivity, one must remove the discrepancy between the data points and the curve. The simplest way to fit a straight line to a set of paired observations such as: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. The mathematical expression for the straight line is given by equation 5:

$$y = a_0 + a_1x + e \quad (5)$$

Where: a_0 and a_1 =coefficients for intercept and slope respectively e = error or residual

Equation 5 can be rearranged to determine the residual given by equation 6:

$$e = y - a_0 - a_1x \quad (6)$$

Thus, in order to sought the “best” line through a set data, one must minimize the sum of the residual errors for all available data as given in equation 7.

$$\sum_{i=1}^n e_i = \sum_{i=1}^n (y_i - a_0 - a_1 x_i) \quad (7)$$

Where: n =total number of points.

To minimize the sum of the absolute values of the discrepancies, equation 3 can be modified to:

$$\sum_{i=1}^n |e_i| = \sum_{i=1}^n |y_i - a_0 - a_1 x_i| \quad (8)$$

Figure 4.2 shows some of the best criteria for “best fit” which are inadequate for regression.

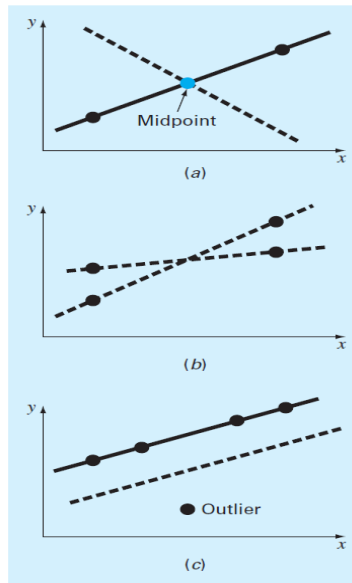


Figure 4.2: Criteria for “best fit” that are inadequate for regression (a) minimize the sum of the residuals (b) minimizes the sum of the absolute values of the residuals (c) minimizes the maximum error of any individual point.

Minimax criterion is a technique for fitting the best line in which the line is chosen that minimizes the maximum distance that an individual point falls from the line and is ill-suited for regression because it gives undue influence to an outlier (as shown in Figure 1.2 (c)) which is an single point in set of data with large error.

Which are called normal equations and can be solved simultaneously to determine slope as shown in equation 9:

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (9)$$

Thus, solving for intercept as shown by equation:

$$a_o = \hat{y} - a_1 \hat{x} \quad (10)$$

Where: \hat{y} and \hat{x} are the means of y and x respectively.

Recalling the formula for standard deviation and least square which has a striking similarity:

$$S_t = \sum (y_i - \hat{y})^2 \quad S_r = \sum_{i=1}^n (y_i - a_o - a_1 x_i)^2$$

In order to quantify the error of Linear Regression, a number of additional properties of this fit can be elucidated by examining more closely the way in which residuals were computed. Thus, for standard deviation, the square of the residual represented the square of the discrepancy between the data and a single estimate of the measure of the central tendency which is the mean while in least square, the square of the residual represents the square of the vertical distance between the data and another measure of central tendency which is the straight as shown in Figure 4.3.

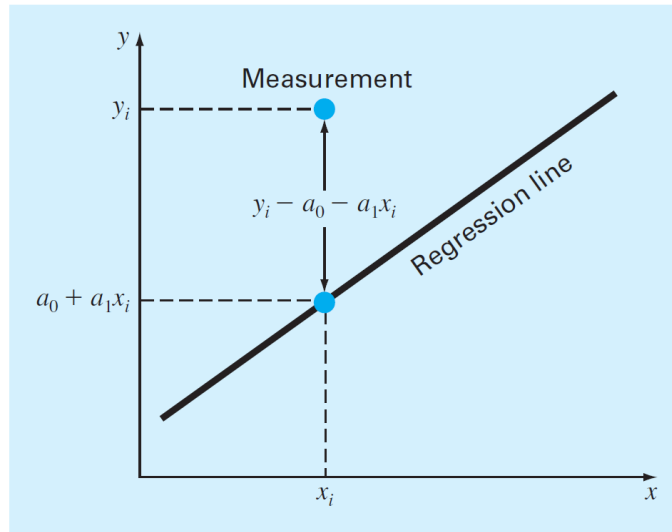


Figure 4.3: Graphic representation of residual in linear regression

These comparisons can be applied further for cases where:

1. Spread of the points around the line is of similar magnitude along the entire range of data
2. Distribution of these points about the line is normal.

If these criteria are met, least square regression will provide the most likely best estimates for a_0 and a_1 which are called maximum likelihood principle (Draper and Smith, 1981). An equation for “standard deviation” applied for regression line or standard error of the estimate is given by equation 11.

$$s_{y/x} = \sqrt{\frac{S_r}{n-2}} \quad (11)$$

Where: $s_{y/x}$ = standard error of the estimate in which the subscript notation “y/x” designates that the error is for a predicted value of y that corresponds to the value of x

n = number of points which must be greater than 2 to avoid “spread of data” around the straight connecting two points

The graphical representation of residual errors is shown in Figure 4.4 in which (a) has less errors or “best fit” than (b).

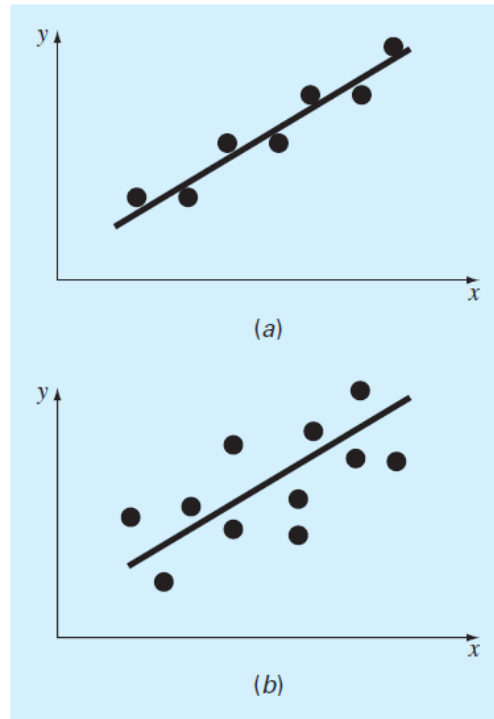


Figure 4.4: Graphical representation of linear regression

Unexplained sum of the squares happens when an error remains after the regression which quantifies the improvement or error reduction due to describing the data in terms of a straight line which is given by equation 12.

$$r^2 = \frac{S_t - S_r}{S_t} \quad (12)$$

Where: r^2 =coefficient of determination r =correlation coefficient

An alternative formula for r using the variables of standard deviation and least square will yield.

$$r = \frac{n \sum(x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (13)$$

Polynomial regression

The least-squares procedure can be used to fit the data to a higher polynomial or even quadratic such as:

$$y = a_0 + a_1x + a_2x^2 + e$$

In this case, the sum of the squares of the residuals as shown in equation 14 is:

$$S_r = \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2)^2 \quad (14)$$

The two dimensional case can be easily extended to an mth-order polynomial to give its general form as in equation 15:

$$y = a_0 + a_1x + a_2x^2 + \dots + a_mx^m + e \quad (15)$$

Recognizing that determining the coefficients of an mth order polynomial is equivalent to solving a system of m+1, the standard error is given by equation 16:

$$s_{\frac{y}{x}} = \sqrt{\frac{S_r}{n - (m + 1)}} \quad (17)$$

Where: n-number of points m-order of the polynomial

Setting these partial derivatives to be minimized and rearranging to develop a simultaneous linear equations with three unknowns a_0 , a_1 and a_2 as:

$$\begin{aligned}(n)a_0 + \left(\sum x_{1,i}\right)a_1 + \left(\sum x_{2,i}\right)a_2 &= \sum y_i \\ \left(\sum x_{1,i}\right)a_0 + \left(\sum x_{1,i}^2\right)a_1 + \left(\sum x_{1,i}x_{2,i}\right)a_2 &= \sum x_{1,i}y_i \\ \left(\sum x_{2,i}\right)a_0 + \left(\sum x_{1,i}x_{2,i}\right)a_1 + \left(\sum x_{2,i}^2\right)a_2 &= \sum x_{2,i}y_i\end{aligned}$$

From the previous discussion, simple linear, polynomial and multiple belong to the general linear least squares model as shown by equation 18.

$$y = a_0z_0 + a_1z_1 + a_2z_2 + \cdots + a_nz_n \quad (18)$$

Where z_0, z_1, \dots, z_m are $m+1$ basis function. Thus, it can be said it is a “linear” in a sense that it only refers to the model’s dependence on its function- a ’s. Equation 12 can be expressed then in matrix notation as in equation 19.

$$\{y\} = [Z]\{a\} + \{e\} \quad (19)$$

Where Z is the matrix of the calculated values of the basis functions such

$$[Z] = \begin{bmatrix} z_{01} & z_{11} & \cdots & z_{m1} \\ z_{02} & z_{12} & \cdots & z_{m2} \\ \vdots & \vdots & & \vdots \\ z_{0n} & z_{1n} & \cdots & z_{mn} \end{bmatrix}$$

m =number of variables n =number of data points

column vector $\{y\}$ =observed values of the dependent variable such

$$\{y\}^T = [y_1 \ y_2 \ \dots \ y_n]$$

column vector $\{a\}$ =unknown coefficients such

$$\{a\}^T = [a_0 \ a_1 \ \dots \ a_n]$$

column vector $\{e\}$ =residuals such

$$\{e\}^T = [e_1 \ e_2 \ \dots \ e_n]$$

The sum of squares of the residuals can be defined in equation 20.

$$S_r = \sum_{i=1}^n \left(y_i - \sum_{j=0}^m a_j z_{ij} \right)^2 \quad (20)$$

The outcome of this process can be generalized in its matrix form as given by equation 21:

$$[Z]^T [Z] \{a\} = \{[Z]^T \{y\}\} \quad (22)$$

The coefficient of the determination and the standard error can also be formulated in terms of matrix algebra such as given in equation 16. From equation 7,

$$r^2 = \frac{S_t - S_r}{S_t} = 1 - \frac{S_r}{S_t}$$

Thus, substituting the definitions of S_r and S_t yields,

$$r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

where \hat{y} =prediction of the least squares fit. Thus, it can be expressed in vector form as:

$$\{y\} - [Z]\{a\}$$

Polynomial Interpolation

Polynomial Interpolation is simply a technique to estimate intermediate values between precise data points. The general formula for an (n-1)th order polynomial can be written as:

$$f(x) = a_1 + a_2x + a_3x^2 + \dots + a_nx^{n-1} \quad (23)$$

For n data points, there is one and only one polynomial of order (n-1) that passes through all the points. This polynomial then provides a formula to compute intermediate values such as shown in figure 4.5.

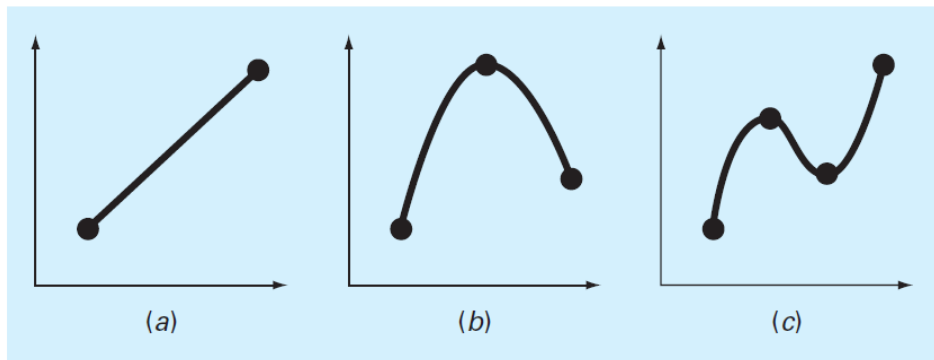


Figure 4.5: Examples of interpolating polynomials (a) first order (linear) (b) second order (quadratic) (c) third order

Polynomial interpolation consists the unique (n-1)th order polynomial that fits n data points. In order to be consistent with respect to x, equation 24 will use decreasing powers as in,

$$f(x) = p_1x^{n-1} + p_2x^{n-2} + \dots + p_1x + p_n \quad (24)$$

Although the above approach provides an easy way to solve interpolation it has serious deficiency as the last column [1,1,1] and it is prone to round off errors. This type of matrices is referred to as Vandermonde matrices and thus, it is inaccurate if applied to a much larger

simultaneous. Newton and Lagrange polynomials mitigates this shortcomings by plotting its objective function.

Newton Interpolating Polynomial

The simplest form of interpolation is to connect data points with a straight line as shown in Figure 4.6 which is called linear interpolation.

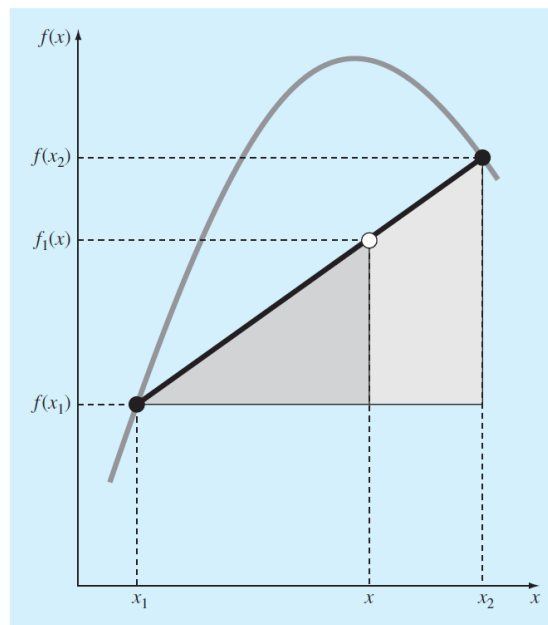


Figure 4.6: Graphical representation of linear interpolation

From Figure 4.6, using similar triangles,

$$\frac{f_1(x) - f(x_1)}{x - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Rearranging, to yield

$$f_1(x) = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \quad (25)$$

In which equation 25 is the Newton linear-interpolation formula. The notation $f_1(x)$ indicates that this is a first order interpolation polynomial and the term $\frac{f(x_2) - f(x_1)}{x_2 - x_1}$ is a finite difference of the first derivative. Thus, the smaller the interval between the data points, the better the approximation.

If three data points are available, this can be accomplished with a second-order polynomial (also called a quadratic polynomial or a parabola). A particularly convenient form for this purpose can be shown in the sample equation below:

$$f_2(x) = b_1 + b_2(x - x_1) + b_3(x - x_1)(x - x_2)$$

To determine the coefficients, substituting $x=x_1$, $x=x_2$ and $x=x_3$ by successive substitution to the original equation, yields.

$$\begin{aligned} b_1 &= f(x_1) \\ b_2 &= \frac{f(x_2) - f(x_1)}{x_2 - x_1} \\ b_3 &= \frac{\frac{f(x_3) - f(x_2)}{x_3 - x_2} - \frac{f(x_2) - f(x_1)}{x_2 - x_1}}{x_3 - x_1} \end{aligned}$$

These equations is beginning to manifest to a structure similar to Taylor series expansion in which terms are added sequentially to capture increasingly higher order curvature.

The foregoing analysis of (n-1)th polynomials can be written into its general form as given in equation 26,

$$f_{n-1}(x) = b_1 + b_2(x - x_1) + \cdots + b_n(x - x_1)(x - x_2) \cdots (x - x_{n-1}) \quad (27)$$

Using the data points to determine the coefficients for (n-1)th polynomial into its general form given by equation 28,

$$b_1 = f(x_1)$$

$$b_2 = f[x_2, x_1]$$

$$b_3 = f[x_3, x_2, x_1]$$

.

.

.

$$b_n = f[x_n, x_{n-1}, \dots, x_2, x_1] \quad (28)$$

The bracketed function evaluations are finite divided differences (graphically represented in Figure 4.8) which can be generalized as in equation 29,30 and 31 for first,second and nth divided difference respectively,

$$f[x_i, x_j] = \frac{f(x_i) - f(x_j)}{x_i - x_j} \quad (29)$$

$$f[x_i, x_j, x_k] = \frac{f[x_i, x_j] - f[x_j, x_k]}{x_i - x_k} \quad (30)$$

$$f[x_n, x_{n-1}, \dots, x_2, x_1] = \frac{f[x_n, x_{n-1}, \dots, x_2] - f[x_n, x_{n-1}, \dots, x_2]}{x_i - x_k} \quad (31)$$

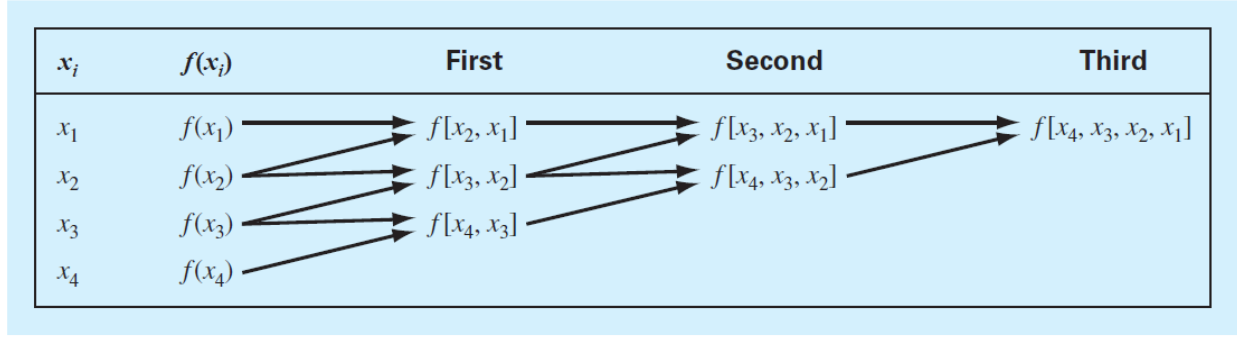


Figure 4.7: Graphical representation of finite divided difference which shows its recursive nature

Using the general forms of determining the coefficients and the bracketed function evaluations (equation 29 to 31) to the general form of (n-1)th polynomial (equation 28) will yield the general form of Newton's interpolating polynomials as shown in equation 32,

$$f_{n-1}(x) = f_1(x) + (x - x_1)f[x_2, x_1] + (x - x_1)(x - x_2)f[x_3, x_2, x_1] + \dots + (x - x_1)(x - x_2) \dots (x - x_{n-1})f[x_n, x_{n-1}, \dots, x_2, x_1] \quad (32)$$

Lagrange Interpolating Polynomial

Linear interpolating polynomial can be written as the weighted average of the two values that are connected by a straight line such as shown in equation 33.

$$f(x) = L_1f(x_1) + L_2f(x_2) \quad (33)$$

Where L's=weighting coefficient in which the first weighting coefficient is the straight line that is equal to 1 at x_1 and 0 at x_2 such $L_1 = \frac{x-x_2}{x_1-x_2}$ and the coefficient is the straight line that is equal to 1 at x_2 and 0 at x_1 such $L_2 = \frac{x-x_1}{x_2-x_1}$. Substituting these coefficients to equation 14 will yield:

$$f_1(x) = \frac{x - x_2}{x_1 - x_2} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2) \quad (34)$$

Equation 25 is referred to as linear Lagrange interpolating polynomial in which this particular equation indicates that is a first order polynomial as graphically represented in Figure 4.9. This is the same strategy as that of Newton's interpolating polynomial. For example, three parabolas will be fit through three points with each one passing through one of the points and equaling zero at the other two. Their sum would then represents the unique parabola that connects the three points and in which a second order Lagrange interpolating polynomial can be written as:

$$f_2(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)}f(x_1) + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)}f(x_2) + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}f(x_3)$$

Higher order Lagrange polynomials can be written in its general form as given in equation 35.

$$f_{n-1}(x) = \sum_{i=1}^n L_i(x)f(x_i) \quad (35)$$

Where

$$L = \prod_{j=1, j \neq i}^n \frac{x - x_j}{x_i - x_j}$$

Where: n=number of data points

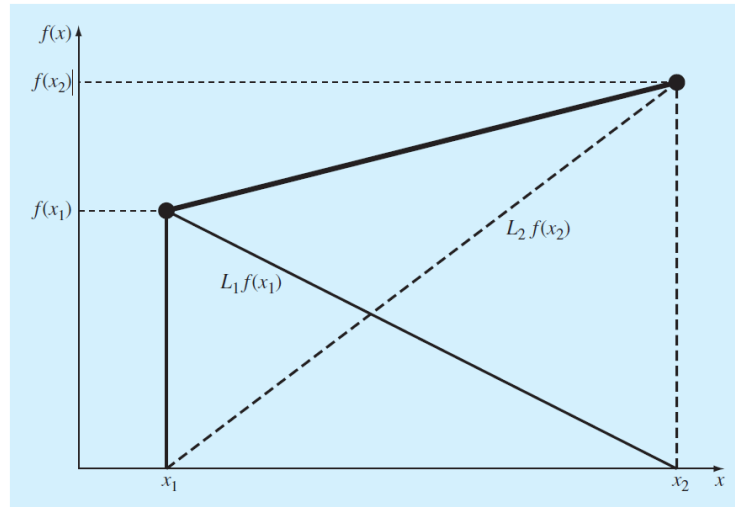


Figure 4.8: Graphical representation of Lagrange interpolating polynomials for first order case

To learn more about Curve fitting, stick with these following resources:

Watch:

- Part 5: Numerical Methods: Curve Fitting | Jacob Bishop
(<https://www.youtube.com/watch?v=yQsDxOdn1hk&list=PLYdroRCLMg5NTT00m-7ituVGdtY1X680M>)
- Regression: Crash Course Statistics #32 | CrashCourse
(<https://www.youtube.com/watch?v=WWqE7YHR4Jc>)
- Statistics 101: Linear Regression, The Very Basics | Brandon Foltz
(<https://www.youtube.com/watch?v=ZkjP5RJLQF4>)

Read:

- Applied Numerical Methods with MatLAB for Engineers and Scientists Chapter 14,15 and 17 by S.T. Chapra
- Applied Numerical Methods using MatLAB Chapter 3 by Won Young Yang, Wenwu Cao, Tae-Sang Chung and John Morris

Activities:

Instruction: Please accomplish this following problem with your corresponding solutions and final answer in a yellow pad with your Name, Section, Subject for each module/topic and send it to the respective courier (Air21/LBC) who sent you this module.

If you have problems and concerns regarding this activity, contact me at my cellphone number 09085212086. Please write eligibly.

Approximation and Rounding Errors:

- A. Determine how many significant digits for each figure:
 - a) 31.20002
 - b) 2014.2300
 - c) 2003×10^4
 - d) 201.003×10^3
 - e) 20.003×10^5
- A. Given an area of circle of 96 inches squared, determine its given radius by using the initial estimate of 4.5 inches.
- B. Essay: Based on your own learning from your previous subjects, How can you apply the concept of accuracy, bias and precision in the field of electrical engineering?

Roots and Optimization:

- A. Problem Solving (Bracketing Method): Solve the following problem using Bisection Method and False Position Method. Using the format below for your answer tabulation:

| N | X_l | X_u | X_r | Ea | Et | $f(x_l)f(x_r)$ |
|---|-------|-------|-------|----|----|----------------|
|---|-------|-------|-------|----|----|----------------|

- a. Determine the roots of $12 + 21x = 18x^2 - 2.75x^3$ with initial guess of -1 and 0 and stopping criterion of 1%.
- b. Locate the first non trivial root of $\sin(x)=x^2$ where x is in radians with the initial interval from 0.5 to 1 and perform the computation until the error is less than 0.02 %.

- c. Solve the roots of $\ln(x^2) + \ln(x^3) + \ln(x^4) = 1$ using an initial value of $x_i = 0.2$ and $x_u = 4$ at an error criterion of 0.005 %.

B. Problem Solving (Open Method): Solve the following problem using Newton Raphson Method and Modified Secant method. Using the format below for your answer tabulation for Newton Raphson and Modified Secant respectively:

| | | | | | | | |
|---|-------|----------|--------------------|-----------------------|----|----|-----------|
| n | | X_i | $f(x_i)$ | $f'(x_i)$ | Ea | Et | X_{i+1} |
| n | X_i | $f(x_i)$ | $X_i + \delta X_i$ | $f(x_i + \delta x_i)$ | Ea | Et | X_{i+1} |

- a. Employ the Newton-Raphson method to determine a real root for $f(x) = -2 + 6x - 4x^2 + 0.5x^3$, using an initial guess of **(a)** 4.5, and **(b)** 4.43
- b. A Parallel Circuit consisting of R, L and C has an impedance of:

$$\frac{1}{Z} = \sqrt{\frac{1}{R^2} + (wC - \frac{1}{wL})^2}$$

- c. Where Z =impedance (ohm), and w =angular frequency. Find the w that results in an impedance of 100 ohm using an initial value of 1 rad/s and perturbation factor of 10^{-6} at error criterion of 0.05 % using the modified secant method for the following parameters: $R=225$ ohm, $C=0.6$ microfarad, $L=0.5$ H.
- d. Use Newton Raphson Method to locate the root of $e^{-x} + \sin(x) = 2$ using an initial value of 0.1 at an error criterion of 0.02%. Use the following format for your solution.

C. Use the Golden Section method to solve the minimum of $f(x) = \ln(x^2) + \ln(x^3) + \ln(x^4)$ using an initial value of $x_i = 0.2$ and $x_u = 4$. Use the following format for your solution.

| N | X_1 | $f(x_1)$ | X_2 | $f(x_2)$ | X_i | $f(x_i)$ | X_u | $f(x_u)$ | d |
|---|-------|----------|-------|----------|-------|----------|-------|----------|---|
|---|-------|----------|-------|----------|-------|----------|-------|----------|---|

D. Use Newton Method to locate the minimum of:

$$f(x) = x^2 + e^{-x} + (4x + x^2)^2 - 2$$

Using an initial value of 0.05. Use the following formats for your answer.

| I | X | $f'(x)$ | $f''(x)$ | x estimate | F(x) |
|---|---|---------|----------|------------|------|
|---|---|---------|----------|------------|------|

Linear System:

A. Given the set of equation below,

$$17x+6y+4z+5t=17$$

$$8x+22y+2z+4t=20$$

$$6x+3y+17z+7t=11$$

$$3x+4y+7z+19t=22$$

- Determine the value of the unknowns using Naïve Gaussian elimination.
- Determine the value of the unknowns using Gauss Jordan elimination.
- Determine the value of the unknowns using LU factorization.
- Compute the Cholesky factorization for the coefficient matrix [A].

B. Given the set of equation below,

$$17x+6y+4z=17$$

$$8x+22y+2z=20$$

$$6x+3y+17z=11$$

- Determine the determinant of the coefficient.
- Determine the value of the unknowns using Cramer's Rule
- Determine the inverse of the coefficient of matrix

C. The following system of equations is designed to determine concentrations (the c.s in g/m³), c₁, c₂ and c₃, in a series of coupled reactors as a function of the amount of mass input to each reactor (the right-hand sides in g/day):

$$15c_1 - 3c_2 - c_3 = 3800$$

$$-3c_1 + 18c_2 - 6c_3 = 1200$$

$$-4c_1 - c_2 + 12c_3 = 2350$$

- a. Determine the value of the concentrations using Gauss Seidel method and using 0 as initial values until the absolute relative error reaches less than 1% or it reaches 10th iteration whichever comes first.
- b. Determine the value of the concentrations using Jacobi iteration and using 0 as initial values until the absolute relative error reaches less than 1% or it reaches 10th iteration whichever comes first.

Curve fitting:

1. Given the data

29.65 28.55 28.65 30.15 29.35 29.75 29.25
 30.65 28.15 29.85 29.05 30.25 30.85 28.75
 29.65 30.45 29.15 30.45 33.65 29.35 29.75
 31.25 29.45 30.15 29.65 30.55 29.65 29.25

Determine (a) mean (b) median (c) mode (d) range (e) standard deviation (f) variance (g) coefficient of variation

2. Use the least squares regression to fit a straight line to

| | | | | | | | | | | |
|---|---|---|---|---|---|----|----|----|----|----|
| x | 0 | 2 | 4 | 6 | 9 | 11 | 12 | 15 | 17 | 19 |
| y | 5 | 6 | 7 | 6 | 9 | 8 | 8 | 10 | 12 | 11 |

Along with the slope and intercept, compute the standard error of the estimate and the correlation coefficient. Plot the data and the regression line.

3. The following data were gathered to determine the relationship between pressure and temperature of a fixed volume of 1 kg nitrogen. The volume is 10 m³.

| | | | | | | |
|---------------------|------|------|------|--------|--------|--------|
| T, C ⁰ | -40 | 0 | 40 | 80 | 120 | 160 |
| P, N/m ² | 6900 | 8100 | 9350 | 10,500 | 11,700 | 12,800 |

Employ the ideal gas law $pV=nRT$ to determine R on the basis of data. Note that for the law, T must be expressed in kelvins.

4. Fit a cubic polynomial to the following data:

| | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|
| x | 3 | 4 | 5 | 7 | 8 | 9 | 11 | 12 |
| y | 1.6 | 3.6 | 4.4 | 3.4 | 2.2 | 2.8 | 3.8 | 4.6 |

Determine the coefficients, r^2 and $s_{y/x}$.

5. Three disease-carrying organisms decay exponentially in seawater according to the following model:

$$p(t) = Ae^{-1.5t} + Be^{-0.3t} + Ce^{-0.05t}$$

Estimate the initial concentration (at $t=0$) of each organism (A,B and C) using multiple linear regression given the following measurements (Hint: Linearize the said equation first before proceeding using the transformation at Figure 4.5)

| | | | | | | | | | |
|------|-----|-----|-----|-----|---|-----|-----|-----|-----|
| t | 0.5 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 9 |
| p(t) | 6 | 4.4 | 3.2 | 2.7 | 2 | 1.9 | 1.7 | 1.4 | 1.1 |