

CHAPTER II

THEORY

Polymers are susceptible to attack by a wide variety of naturally occurring and man-made agents. An exposure of polymers to certain environments over a period of time can lead to deterioration of physical properties and hence degradation, which limits the lifetime of the polymers. Failure induced by degradation of polymers could cause catastrophic fracture. Therefore, a stabilizer is an essential part for polymers, especially for commodity plastics such as polyethylene.

2.1 High Density Polyethylene

Polyethylene is a thermoplastic polymer produced by polymerization of ethylene. HDPE is a linear polymer that can be produced via a low pressure process by Ziegler-Natta polymerization, by the Phillips process and by the Standard Oil (Indiana) process. The coordination polymerization utilizes a catalyst of an aluminum alkyl and $TiCl_4$ in a solvent such as heptane. Ethylene is added to the reaction vessel under slight pressure, at a temperature of 50-70°C. HDPE forms as white powder; it is insoluble in the reaction mixture. The HDPE is then filtered or centrifuged off, washed and dried at the completion of the reaction.

Typical linear polyethylenes are highly crystalline polymers with over 90% of crystallinity. They contain less than one side chain per 200 carbon atoms in the main chain. Their melting point is above 127°C (typically about 135°C) and their density is in the range of 0.941 to 0.965 g/cm³.

Most of the differences in the properties between low density and high density polyethylenes can be attributed to the higher crystallinity of the latter polymers. Linear polyethylenes are decidedly stiffer than the branched ones. This is evident in the modulus of elasticity of 100,000 psi (689.5 MPa) for HDPE as compared with 20,000 psi (137.9 MPa) for LDPE. Linear PE also has a higher crystalline melting point, greater tensile strength and hardness. The good chemical resistance of branched polyethylene is still retained or enhanced in linear PE, and such properties as low-temperature brittleness and low permeability to gases and vapors are improved in the linear material.

The linear polyethylene has enhanced properties superior to the branched structure as shown in Table 2.1.

Table 2.1 : Properties of high density polyethylene and low density polyethylene.

Commercial properties	HDPE	LDPE
Density, g/cm ³	0.941-0.965	0.910-0.925
Approximate % crystallinity	65-90	42-53
Branching equivalent CH ₃ group/1000 carbon atoms	1-5	15-30
Crystalline melting temperature, °C	120-140	95-130
Tensile strength, psi	3100-5490	590-2300
Tensile modulus, psi	60000-181200	14000-38000
Impact strength (ft-lb/in of notch)	0.5-2.0	>16
Elongation to break, %	20-1300	90-800
Hardness, Shore D	50-70	41-50

Bottles and other containers processed by blow molding accounts for about 40% of the use of linear polyethylene. The adjustment of structure variables to obtain high resistance to environmental stress cracking, allowing the HDPE to be used in detergent bottles, produced a large expansion in the packaging field. About 25% of the linear polyethylene produced is used in the injection molding of crates, pails, tubs, caps and closures and housewares. The higher stiffness and heat resistance of linear PE have led to the replacement of branched polyethylene in applications where these properties are important. Other major uses of HDPE include film and sheet, rotational molding, wire and cable insulation, extrusion coating, and pipe and conduit.

2.2 Polymer Degradation

Degradation is defined as any undesirable changes in polymer properties that occur after the polymer has been exposed to a degradation factor either during processing or during service life.

Degradations may be classified by their mechanisms or factors causing degradation as shown in Table 2.2. The main mechanisms taking place as dominating events during degradation are depolymerization, random chain scission, cross-linking, bonds changes and side group elimination. The factors which initiate degradation are heat (thermal degradation), heat plus oxygen (thermo-oxidative degradation), high energy radiation, ultraviolet light (photodegradation), mechanical stress (mechano-chemical degradation), chemical attack and biological organisms (biodegradation). In practically important situations, several factors may be involved. Thus, in weathering, not

only sunlight and air (which leads to photo-oxidation) but also water vapour, heat, ozone and pollutants may participate in deteriorating polymers.

Table 2.2 : Polymer degradative factors.

Degradative factor	Degradative Occurrence
Heat	During production (Polymerization, Degassing, Drying, Extrusion, Pelletization) During processing (Mixing, Molding, etc.) During use
Ultraviolet	During use (Indoors, Outdoors)
Oxygen	During production (Air drying) During mixing (Rolls, Banbury, Kneader, Extruder) During use
Heavy metal ions	During production (Catalyst residue) During processing (Equipment) During use (Contact with metals, Composites with inorganic fillers)
Mechanical shear	During production (Extrusion, Pelletization) During processing (Mixing, Molding, etc.)
Ozone	During use

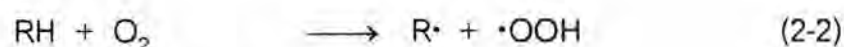
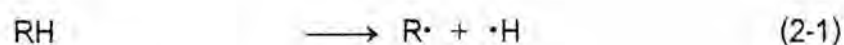
Most polymers are more or less susceptible to various forms of degradation. It generally occurs in every polymer's life cycle. Thermoplastic polymers, like other polymers, are mainly subjected to two different degradation stages during their lifetime. The first stage occurs in the enclosed processing machine, where the polymer is subjected simultaneously to a high temperature and mechanical shear stress with little of the presence of atmospheric oxygen. Degradation may occur via various mechanisms such as chain scission, crosslinking and also the occurrence of oxygen-containing chemical groups, which initiate other degradation form in the final material. The second stage of degradation occurs in the finished material, when the product is subjected to many external effects. The most predominantly

dangerous effects are UV radiation and high temperature. From these two degradation stages, it is necessary to protect the polymer against various external effects by means of incorporation of appropriate stabilizers.

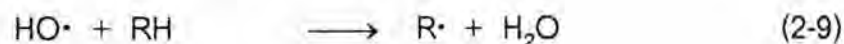
2.2.1 General Mechanisms of Polymer Degradation

Oxidation of polyethylene occurs in air upon exposure to ultraviolet light and/or elevated temperature. The reaction of polyethylene and other hydrocarbon polymers with oxygen proceeds by a free radical mechanism. The reactions may be represented by three schemes as follow [1]:

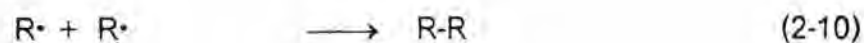
(i) Initiation:



(ii) Propagation:



(iii) Termination:





where RH = Polymer R \cdot = alkyl radical
 RO \cdot = alkoxy radical ROO \cdot = peroxy radical
 ROOH = hydroperoxide HO \cdot = hydroxy radical

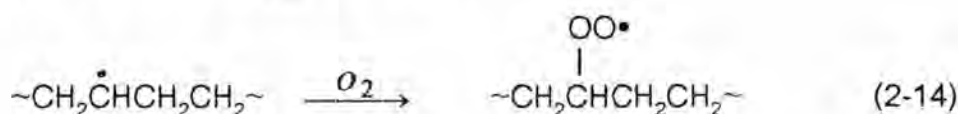
2.2.2 Degradation of Polyethylene

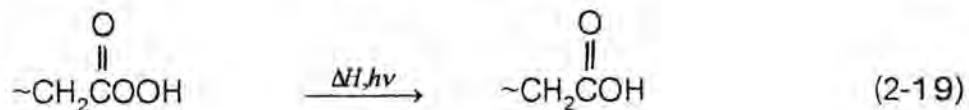
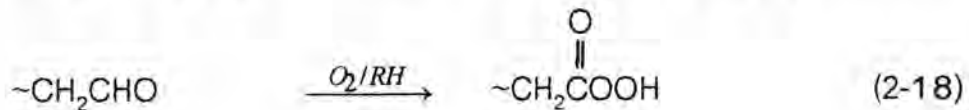
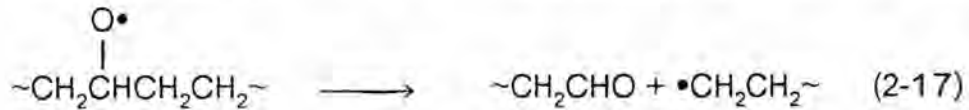
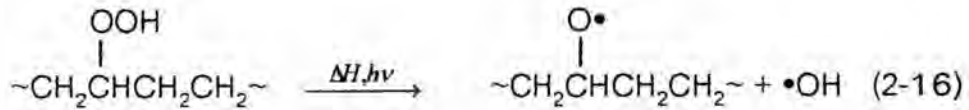
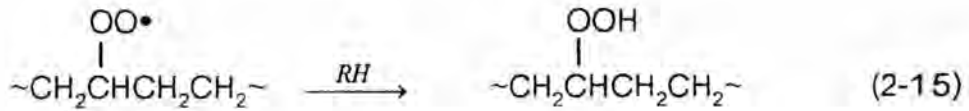
In polyolefins and polymers with hydrocarbon fragments, degradation proceeds as a radical reactions with branched kinetic chains. The chain reaction mechanism of polyethylene generally follows a mechanism quite similar to the general mechanism established for simple hydrocarbon degradation. The reaction involves free radical initiation, and a chain reaction in which the chain length may be comprised of several steps. Reactions (2-13) through (2-21) summarize the alternative reactions of alkyl radicals during high temperature processing of polyethylene.

(i) Initiation :

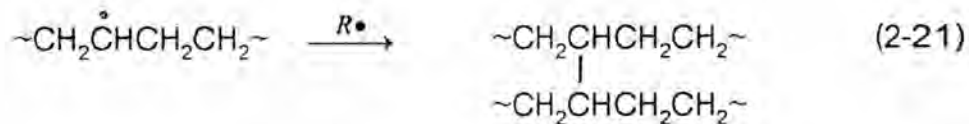
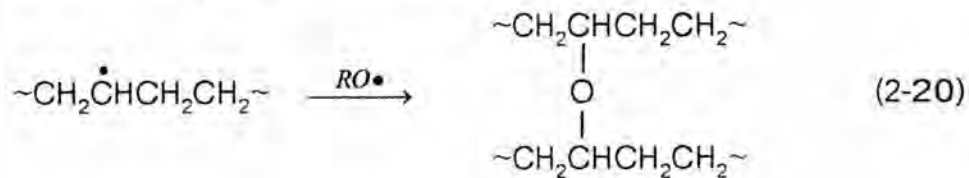


(ii) Chain scission (oxygen excess) :





(iii) Cross-linking (oxygen deficiency) :



In the initiation step of thermo-oxidative degradation of polyethylene, as shown in Reaction (2-13), hydrogen is removed from the secondary carbon atom in polyethylene leaving free polymeric radical. Initial formation of a strongly initiating primary polymeric radical may be induced by impurity such as polymerization catalyst residue, traces of hydroperoxide and so on.

After the start of the initiation step, the subsequent chain reaction of polyethylene may undergo two primary reactions which are chain scission and

cross-linking, depending on the oxygen concentration. In an abundant oxygen atmosphere, chain scission, which dominantly occurs during processing, produces hydroperoxide units on the polyethylene backbone. These hydroperoxides, which are shown in Reaction (2-15), are effective initiators of degradation to create the strongly abstracting hydroxyl radical and another radicals. Each of the radicals formed by hydroperoxide homolysis could react with additional polyethylene molecule to form the initiation of new reactive chains. Hence, the rate of degradation is accelerated while hydroperoxides accumulate and decompose into radicals. The result of chain scission is a reduction in the chain length and/or the molecular weight.

In a limited oxygen concentration environment and particularly at very high temperature, polyethylene is predominately subjected to cross-linking as presented in Reactions (2-20) and (2-21). Recombination of chain radicals can produce crosslinks in polyethylene structure resulting in molecular enlargement and free radical chain termination.

2.3 Antioxidants

Although all polymers degrade at high temperature in the absence of air, degradation is almost always faster in the presence of oxygen. The role of antioxidant in preventing or inhibiting the oxidation of polymers is usually filled by a substance which itself is readily oxidized, although in some cases the antioxidant may act by combining with the oxidizing polymer to form a stable product [2]. Antioxidants increase the stability of the polymer at high temperature and also extend the service life of such polymer at ambient temperature.

Antioxidants interrupt the degradation process in different ways, depending on their structures. The two major classifications of antioxidants are primary and secondary antioxidants.

2.3.1 Primary antioxidants

A majority of primary antioxidants are sterically hindered phenols or secondary aromatic amines. Primary antioxidants act or participate mainly in Cycle I of Figure 2.1. They are capable of undergoing fast reactions with peroxy radicals, therefore, they are frequently called "*radical scavengers*".

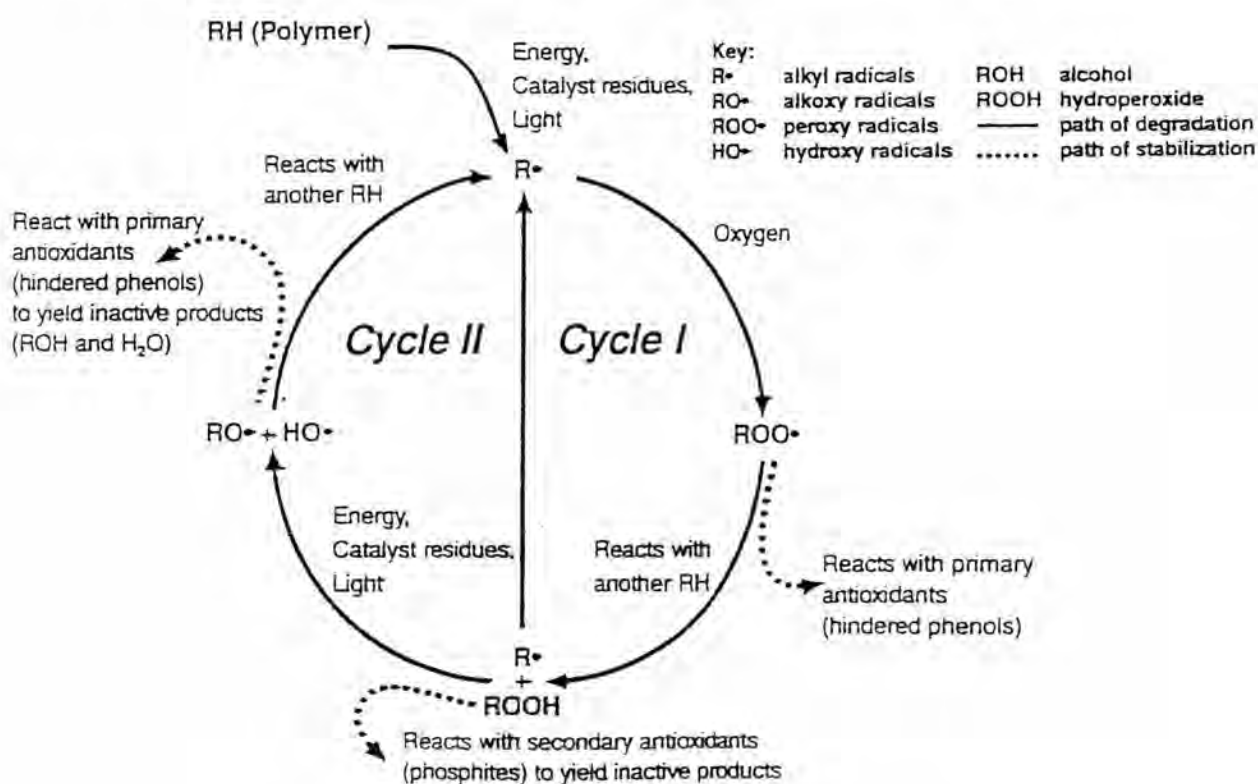
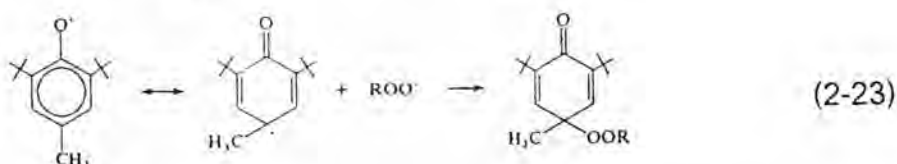
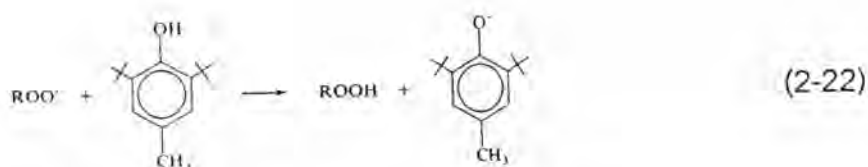


Figure 2.1 : Polymer degradation and stabilization.

An example of the reaction sequence of primary antioxidant is shown in Reactions (2-22) to (2-23).



Stabilization of polymers is achieved by the fact that Reaction (2-22) competes with Reaction (2-5), transforming the reactive peroxy radical into the much less reactive phenoxyl radical, which, in turn, is capable to react with a second peroxy radical as shown in Reaction (2-23). There is abundant literature where the isolation of the generated peroxycyclohexadienones in Reaction (2-23) is described [3], together with numerous other oxidation products from various antioxidants.

The thermal stability of the peroxycyclohexadienones in Reaction (2-23) is limited, their decomposition leads to new reaction chains also at temperatures below 150°C. The phenoxyl radicals in Reaction (2-22) do not initiate new radical chains at the normal temperatures of use and testing, however, this propagation reaction becomes possible at high temperatures. This means that the effectiveness of sterically hindered phenols is decreasing with increasing temperature [3].

The phenolic antioxidants are generally nonstaining and non-discoloring although upon degradative process they may form highly colored quinoid structure. As an example on the conversion of the phenolics during stabilization, the loss of the active form of the antioxidant are resulted in

quinone methides, peroxy-cyclohexadienones and benzoquinones. Their structure are demonstrated in Figure 2.2

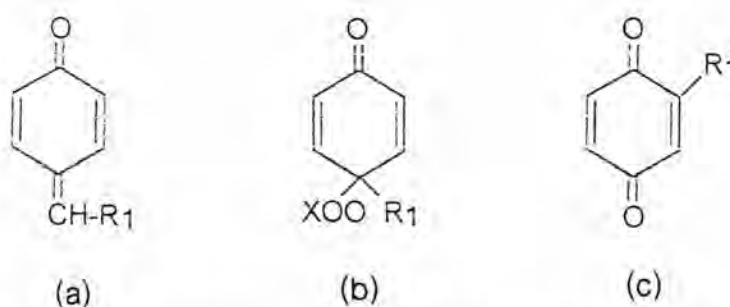


Figure 2.2 : The quinoid structure : (a) quinone methide, (b) peroxy-cyclohexadienones ($R_2 = \text{H}$ or alkyl) and (c) benzoquinones.

2.3.2 Secondary antioxidant

Secondary antioxidants are also termed "*preventive stabilizers*" because they can prevent the proliferation of oxidative deterioration by counteracting the formation of free radicals. Their function is to decompose the hydroperoxy group ($\text{HOO}\sim$) present in the polymer, acting in Cycle II of Figure 2.1. Secondary antioxidants are classified by two chemical groups namely phosphites and thioesters.

2.3.2.1 Phosphites

The phosphites primarily act as hydroperoxide decomposers by reducing hydroperoxides to alcohols, while they themselves are converted to phosphates. The reaction is shown in Reaction (2-24).

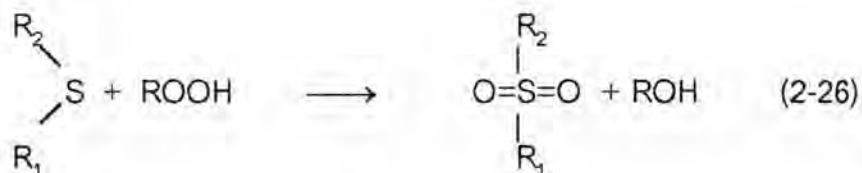
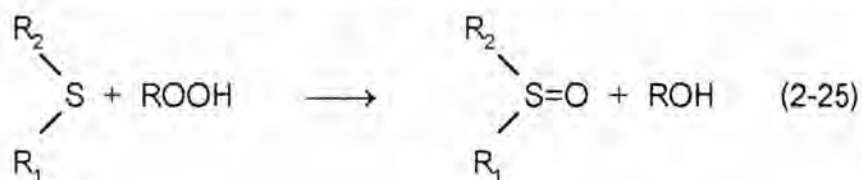


The performance of phosphites is not only nondiscoloring but also color stability. They are extensively used with phenolics to inhibit the formation of the discoloring quinoid reaction products formed upon oxidation of the phenolic antioxidant. They can also improve effectively the melt index.

Due to their major drawback on hydrolysis, the utilization of phosphites is limited in some products. The formation of phosphoric acid resulted from hydrolysis of phosphites can cause corrosion of processing equipments.

2.3.2.2 Thioesters

Most notable thioesters are aliphatic ester of β -thiodipropionic acid. They are highly effective peroxide decomposer for long-term stability to polymers. Sulfoxides, the active forms of sulfides, are capable of decomposing many hydroperoxide to form stable products. The interaction of thioester with hydroperoxide is represented in Reactions (2-25) and (2-26).



Being very good synergistic companions for primary antioxidant, the thioester are commonly used with phenolic antioxidant.

2.3.3 Synergism

Commercially, antioxidants are often used in combination. Two antioxidants may exhibit a co-operative effect and be more powerful than either antioxidant used alone, even when the total antioxidant concentration is the same. This is known as *synergism*.

In commercially practical terms, it may be that although synergism occurs, one additive may be sufficiently less expensive for more to be used to achieve the required stability and still gives the lowest overall cost.

In antioxidants, synergism appears to arise either from one antioxidant effectively regenerating another so that the latter does not become consumed or from the two antioxidants functioning via different mechanisms. The latter is more important and it is easy to see how effective a combination of peroxide decomposer and chain breaking antioxidant can be [4].

The peroxide decomposer will drastically reduce the number of radicals, which can then be more effectively 'mopped-up' by the chain breaking materials [4]. It is possible to envisage most powerful combinations where a chain breaking antioxidant, a regenerating agent, a peroxide decomposer, a metal deactivator and an ultra-violet absorber are all employed together.

Two points should be noted while considering antioxidant synergism. Firstly, some antioxidants which function by more than one mechanism are in all probability autosynergistic. Secondly, certain additives which appear to act

synergistically with one antioxidant may be antagonistic to the second one. The most well-known and important example of this is carbon black, which has an adverse effect with many phenols and amines but carbon black is synergistic with some phenolic sulphides [4].

2.4 Optical Brightener

For a long time fluorescent whitening agents or optical brighteners have been used to improve the appearance of finished articles in the paper, soap and textile industries. Along with the development of thermoplastics, using fluorescent whitening agents have penetrated the field of plastics too.

Fluorescent whitening agents can be used in all plastics. The objectives of using fluorescent whitening agent in plastics are the same as in paper and textiles. The fluorescent whitening agent helps in improving the initial color of plastics, which are inherently slightly yellowish. It produces "brilliant white" effect on the end-use articles, such as fashion goods of imitation leather, packaging materials, etc. For colored-articles, it helps in increasing the brilliancy of colored and black-pigmented articles.

A suitable fluorescent whitening agent shows high whitening effectiveness at low concentrations. In practice, concentrations of 100 to 500 ppm are used in thermoplastics. Only special applications require concentrations exceeding 1000 ppm [3].

2.5 Free radical scavengers

Apart from the absorption of harmful radiation by UV absorbers and the decomposition of hydroperoxides by some compounds containing phosphorus and/or sulfur, the scavenging of free radical intermediates is another possibility for stabilization, analogous to that used in thermal degradation. However, these reactions were not taken into consideration for a long time, probably because the phenolic antioxidants do not inhibit photooxidation efficiently enough for most practical purposes. Nevertheless, it has been shown that some phenolic antioxidants can improve the light stability of polypropylene (PP) by an order of magnitude in comparison with the unstabilized PP [3].

Hindered Amine Light Stabilizers (HALS) is an example of free radical scavengers. Its general molecular structure is shown in Figure 2.3.

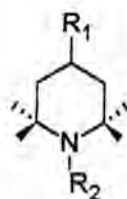
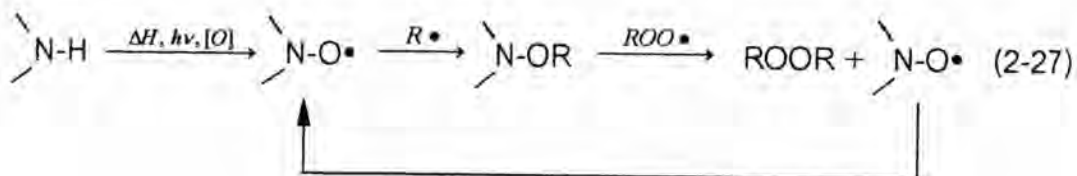


Figure 2.3 : The molecular structure of Hindered Amine Light Stabilizers (HALS).

During light exposure, stable nitroxyl radicals are generated. These are the radical scavengers, reacting according to Reaction (2-27) [1].



2.6 Statistical Approaches for Experimental Analysis

2.6.1 Response Surface Methodology (RSM)

RSM is a collection of mathematical and statistical techniques useful for analyzing problems in which several independent variables influence a dependent variable or response. If the independent variables are denoted as x_1, x_2, \dots, x_k . If these variables are continuous and controllable by the experimenter with minimal error, the response (y) is assumed to be a random variable. The response (y) can be written as a function of the levels of the independent variables as shown in Equation (2-28)

$$y = f(x_1, x_2, \dots, x_k) + e \quad \dots(2-28)$$

where e is a random error component. If the expected response is represented by $E(y)$ which is η , is called a *response surface* then the surface represented by a relationship shown in Equation (2-29).

$$\eta = f(x_1, x_2, \dots, x_k) \quad \dots(2-29)$$

Equation (2-29) may represent a two-dimensional response surface graphically by drawing the x_1 and x_2 axes in the plane of the paper and visualizing the $E(y)$ axis perpendicular to the plane of the paper. Then plotting response surface curve and contours of constant expected response yields the response surface as in Figure 2.4 and 2.5.

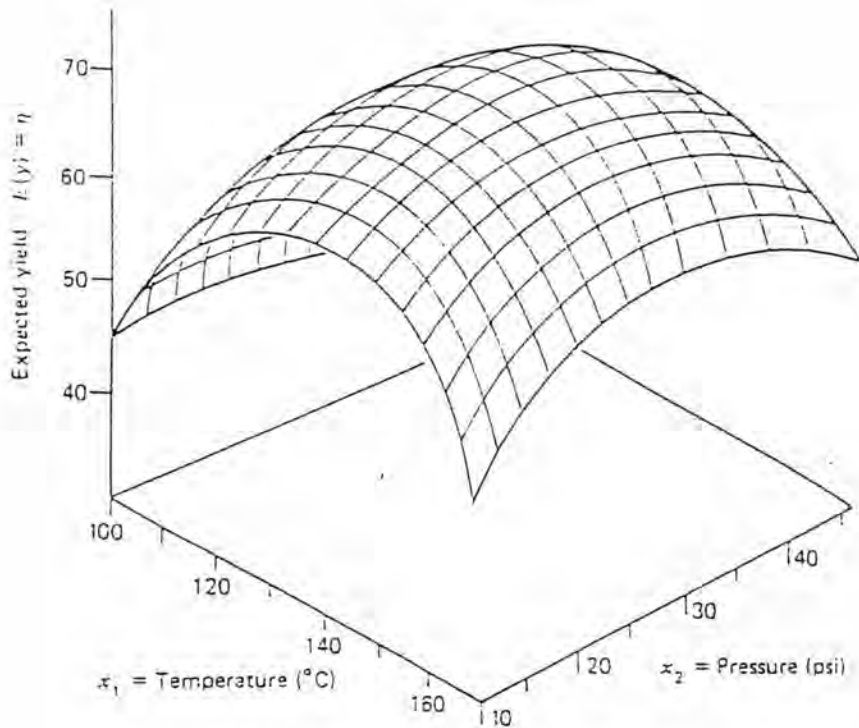


Figure 2.4 : Example of a response surface show the relationship between the response (η) with independent variables x_1 and x_2 .

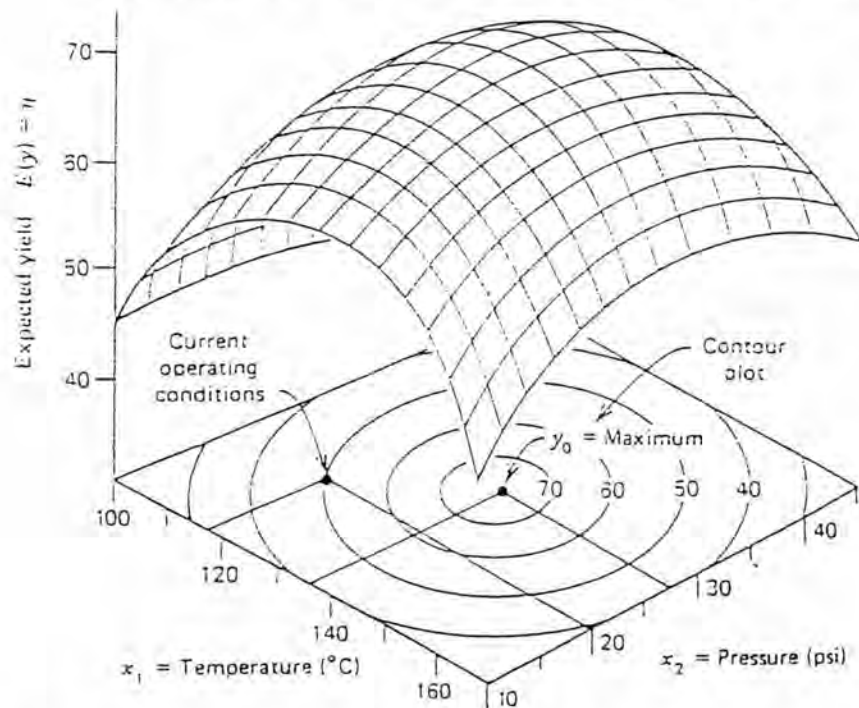


Figure 2.5 : Example of a contour curve show the relationship between the response (η) with independent variables x_1 and x_2 .

In most RSM problems, the form of the relationship between the response and the independent variables is unknown. Thus, the first step in RSM is to find a suitable approximation for the true functional relationship between y and the set of independent variables. Usually, a low-order polynomial in some region of the independent variables is employed. If the response is well-modeled by a linear function of the independent variables, then the approximating function is the first-order model as shown in Equation (2-30).

$$y = B_0 + B_1x_1 + B_2x_2 + \dots + B_kx_k + e \quad \dots(2-30)$$

If there is curvature in the system, then a polynomial of higher degree, such as the second-order model must be used.

$$y = B_0 + \sum_{i=1}^k B_i x_i + \sum_{i=1}^k B_{ii} x_i^2 + \sum_{i=1}^k \sum_{j=1}^k B_{ij} x_i x_j + e \quad \dots(2-31)$$

The method of least squares is used to estimate the parameters (B) in the approximating polynomials [5]. Figure 2.6 illustrates how, by suitable choices of the coefficients, the second order surface in x_1 and x_2 can take on a variety of useful shapes. Both the contour plots and the associated surfaces are shown in Figure 2.6.

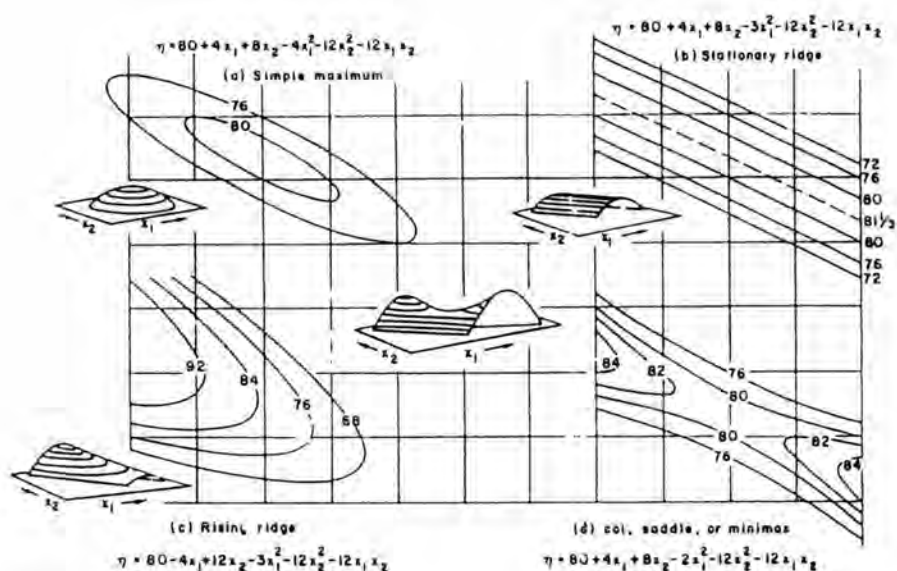


Figure 2.6 : Some examples of the types of surfaces defined by second-order polynomials in two predictor variables, x_1 and x_2 .

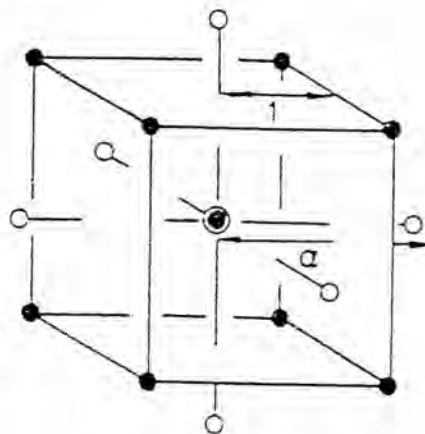
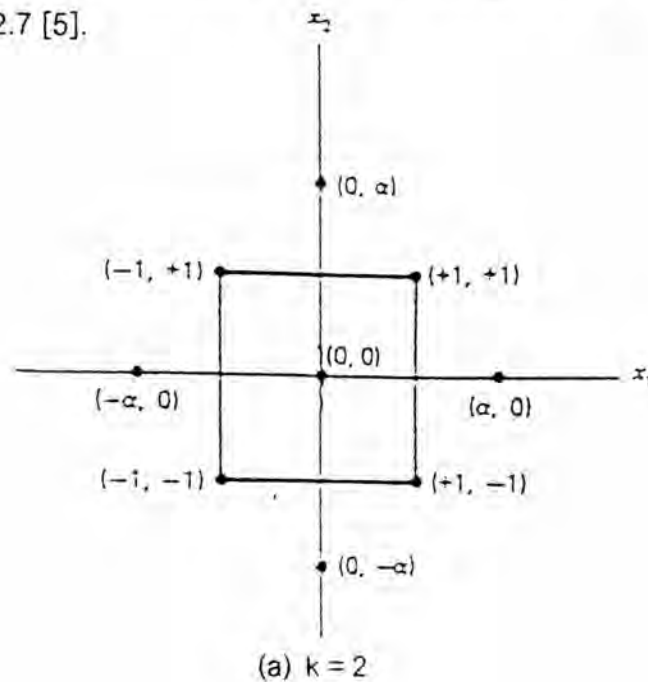
A simple maximum is shown in Figure 2.6(a), and a stationary or flat ridge in Figure 2.6(b). Figure 2.6(c) shows a rising ridge, and Figure 2.6(d) shows what is variously called a col, saddle, or minimax [6].

2.6.2 Experimental Designs for Fitting Response Surface

An experimental design or fitting a second-order model must have at least three levels of each factor so that the model parameters can be estimated. The preferred class of second-order response surface designs is the class of *rotatable* designs.

The most widely used design for fitting a second-order model is the central composite design. These designs consist of a 2^k factorial or fractional factorial (coded to the usual ± 1 notation) augmented by $2k$ axial points $(\pm\alpha, 0, 0, \dots, 0)$, $(0, \pm\alpha, 0, \dots, 0)$, $(0, 0, \pm\alpha, \dots, 0)$, \dots , $(0, 0, 0, \dots, \pm\alpha)$ and n_0 center

points $(0, 0, \dots, 0)$. Central composite rotatable designs for $k = 2$ and $k = 3$ are shown in Figure 2.7 [5].



- Factorial (or "cube") points, with coordinates $(\pm 1, \pm 1, \pm 1)$
- Axial (or "star") points, distance α from origin
- ⊙ Center points. If the design is blocked into cube and star portions, some center points would be in the cube block, some in the star block

Figure 2.7 : Central composite rotatable designs for $k = 2$ and $k = 3$.

Rotatable designs for any number k of x -variables can be built up from these three components. The value of α must be $2^{k/4}$ in order to make the design rotatable. Table 2.3 shows the components of the design for $k = 2, 3, 4, 5$ and 6 . Note that with 5 and 6 x -variables, the size of the experiment is reduced by using a half-replicate of the 2^k factorial. With a half-replicate, The value of α becomes $2^{(k-1)/4}$. Tables 2.4 and 2.5 show examples of the central composite rotatable design for 2 independent variables, x_1, x_2 , and 3 independent variables, x_1, x_2, x_3 respectively.

Table 2.3 : A comparison of the number of experiments designed by factorial design and by central composite rotatable design for k independent variables at 5 levels of variables.

Number of Independent Variables (k)	Number of experiments with 5 levels of variables					
	Factorial (2^k)	Central Composite Rotatable				
		Cube Point	Star Point	Center Point	Total	α value
2	$5^2 = 25$	4	4	5	13	1.414
3	$5^3 = 125$	8	6	6	30	1.682
4	$5^4 = 625$	16	8	7	31	2.000
5	$5^5 = 3,125$	16	10	6	32	2.000
6	$5^6 = 15,625$	32	12	9	53	2.378

It is convenient not to have to deal with the actual numerical measures of the variables X , but instead to work with coded or "standardized" variables x . For example, if at some stage of an investigation, the current region of interest for X are defined to be $X_0 \pm t$, where X_0 is the center of the region, then it would be convenient to define an equivalent working variable x where [6].

$$x = \frac{X - X_0}{t} \quad \dots(2-32)$$

Table 2.4 : Example of the central composite rotatable design for 2 independent variables, x_1 and x_2 .

Experimental	Coded variables of independent variables		Response (y)
	x_1	x_2	
1	-1	-1	y_1
2	1	-1	y_2
3	-1	1	y_3
4	1	1	y_4
5	-1.414	0	y_5
6	1.414	0	y_6
7	0	-1.414	y_7
8	0	1.414	y_8
9	0	0	y_9
10	0	0	y_{10}
11	0	0	y_{11}
12	0	0	y_{12}
13	0	0	y_{13}

Table 2.5 : Example of the central composite rotatable design for 3 independent variables, x_1 , x_2 and x_3 .

Experimental	Coded variables of independent variables			Responses (y)
	x_1	x_2	x_3	
1	-1	-1	-1	y_1
2	1	-1	-1	y_2
3	-1	1	-1	y_3
4	1	1	-1	y_4
5	-1	-1	1	y_5
6	1	-1	1	y_6
7	-1	1	1	y_7
8	1	1	1	y_8
9	-1.682	0	0	y_9
10	1.682	0	0	y_{10}
11	0	-1.682	0	y_{11}
12	0	1.682	0	y_{12}
13	0	0	-1.682	y_{13}
14	0	0	1.682	y_{14}
15	0	0	0	y_{15}
16	0	0	0	y_{16}
17	0	0	0	y_{17}
18	0	0	0	y_{18}
19	0	0	0	y_{19}
20	0	0	0	y_{20}

2.7. Regression Analysis

Regression analysis is a statistical tool used in analyzing the relationship between two or more quantitative variables so that one variable can be predicted from the other or others. After the determination of independent variables, dependent variable can be studied through the experimentation. Dependent variable can also be called response variable [7].

2.7.1 Linear Regression

A response y is uncontrollable in experimentation. Thus x_1, x_2, \dots, x_k are regarded as known constants since they are assumed to be measured without error and controllable in an experiment. The relationship between the independent variables x_1, x_2, \dots, x_k and a response y shown in a mathematical model is called "Linear regression model". Therefore, simple linear regression is derived as shown in Equation (2-33) and multiple linear regression as shown in Equation (2-34).

$$y = B_0 + B_1x + e \quad \dots(2-33)$$

$$y = B_0 + B_1x_1 + B_2x_2 + \dots + B_kx_k + e \quad \dots(2-34)$$

The unknown parameters (B) are usually called regression coefficients. Moreover, multiple linear regression can be written in other forms. The second order polynomial equation can be written as follows

$$y = B_0 + \sum_{i=1}^k B_i x_i + \sum_{i=1}^k B_{ii} x_i^2 + \sum_{i=1}^k \sum_{j=1}^k B_{ij} x_i x_j + e \quad (i < j) \quad \dots(2-35)$$

where e is a random error term with mean $E(e) = 0$

2.7.2 Estimation of Multiple Linear Regression Coefficients

Experimentation generally relates with many independent variables. In order to study the impact of independent variable on interested response, the number of experimentation (n) in regression coefficients has to more than the number of independent variables (k). As seen in Table 2.8 shown an example of the information derived from regression coefficient.

Table 2.6 : Data for Multiple Linear Regression

Independent Variables				Response Variables
x_1	x_2	...	x_k	y
x_{11}	x_{21}	...	x_{k1}	y_1
x_{12}	x_{22}		x_{k2}	y_2
\vdots	\vdots		\vdots	\vdots
x_{1u}	x_{2u}	...	x_{ku}	y_u
\vdots	\vdots		\vdots	\vdots
x_{1n}	x_{2n}	...	x_{kn}	y_n

The information above can be used to find the relationship between the response variable (y) and the independent variables (x_1, x_2, \dots, x_k) by forming the following multiple regression coefficient.

$$y = B_0 + B_1 x_{1u} + B_2 x_{2u} + \dots + B_k x_{ku} + e_u$$

$$= B_0 + \sum_{j=1}^k B_j x_{ij} + e_i \quad i = 1, 2, \dots, n \quad \dots(2-36)$$

The estimation of y value in Equation (2-36) will determine a new equation used for estimation as shown in Equation (2-37).

$$\hat{y}_u = b_0 x_{0u} + b_1 x_{1u} + b_2 x_{2u} + \dots + b_k x_{ku} \quad \dots(2-37)$$

when $x_{0u} = 1$, $b_0 = B_0$, $b_1 = B_1, \dots$, $b_k = B_k$ and the error of the observed value of y_u from the \hat{y}_u is shown in Equation (2-38).

$$e_u = y_u - \hat{y}_u \quad \dots(2-38)$$

The problem is now to obtain the estimates $b_0, b_1, b_2, \dots, b_k$ from the sample for the unknown parameters $B_0, B_1, B_2, \dots, B_k$. This can be done by the method of "Least square". This method minimizes the sum of square, SS_E , of the differences between the predicted values and the experimental values for the dependent variable. The sum of square can be obtained by using Equation (2-39).

$$\begin{aligned} SS_E = L &= \sum_{u=1}^n e_u^2 = \sum_{u=1}^n (y_u - \hat{y}_u)^2 \\ &= \sum_{u=1}^n (y_u - b_0 - b_1 x_{1u} - b_2 x_{2u} - \dots - b_k x_{ku})^2 \quad \dots(2-39) \end{aligned}$$

To determine the minimum, partial derivatives of L with respect to b_0 , b_1 , b_2, \dots, b_k are taken and then they are set to equal zero, as shown in Equation (2-40) to (2-43).

With respect to b_0 :

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial b_0} &= \frac{\partial}{\partial b_0} \left[\sum_{u=1}^n (y_u - b_0 x_{0u} - b_1 x_{1u} - b_2 x_{2u} - \dots - b_k x_{ku})^2 \right] \\ &= -2 \sum_{u=1}^n (y_u - b_0 x_{0u} - b_1 x_{1u} - b_2 x_{2u} - \dots - b_k x_{ku}) \\ &= 0\end{aligned}$$

then

$$\begin{aligned}-2 \sum_{u=1}^n y_u + 2nb_0 + 2b_1 \sum_{u=1}^n x_{1u} + 2b_2 \sum_{u=1}^n x_{2u} + \dots + 2b_k \sum_{u=1}^n x_{ku} &= 0 \\ nb_0 + b_1 \sum_{u=1}^n x_{1u} + b_2 \sum_{u=1}^n x_{2u} + \dots + b_k \sum_{u=1}^n x_{ku} &= \sum_{u=1}^n y_u \quad \dots(2-40)\end{aligned}$$

where $\sum_{u=1}^n x_{0u} = n$, so $b_0 \sum_{u=1}^n x_{0u} = nb_0$

With respect to b_1 :

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial b_1} &= \frac{\partial}{\partial b_1} \left[\sum_{u=1}^n (y_u - b_0 x_{0u} - b_1 x_{1u} - b_2 x_{2u} - \dots - b_k x_{ku})^2 \right] \\ &= -2x_{1u} \sum_{u=1}^n (y_u - b_0 x_{0u} - b_1 x_{1u} - b_2 x_{2u} - \dots - b_k x_{ku}) \\ &= 0\end{aligned}$$

then

$$\begin{aligned}
 -2 \sum_{u=1}^n x_{1u} y_u + 2nb_0 \sum_{u=1}^n x_{1u} + 2b_1 \sum_{u=1}^n x_{1u}^2 + 2b_2 \sum_{u=1}^n x_{1u} x_{2u} + \dots + 2b_k \sum_{u=1}^n x_{1u} x_{ku} &= 0 \\
 nb_0 \sum_{u=1}^n x_{1u} + b_1 \sum_{u=1}^n x_{1u}^2 + b_2 \sum_{u=1}^n x_{1u} x_{2u} + \dots + b_k \sum_{u=1}^n x_{1u} x_{ku} &= \sum_{u=1}^n x_{1u} y_u
 \end{aligned}
 \tag{2-41}$$

and with respect to b_k :

$$\begin{aligned}
 \frac{\partial}{\partial b_k} &= \frac{\partial}{\partial b_k} \left[\sum_{u=1}^n (y_u - b_0 x_{0u} - b_1 x_{1u} - b_2 x_{2u} - \dots - b_k x_{ku})^2 \right] \\
 &= -2x_{ku} \sum_{u=1}^n (y_u - b_0 x_{0u} - b_1 x_{1u} - b_2 x_{2u} - \dots - b_k x_{ku}) \\
 &= 0
 \end{aligned}$$

then

$$\begin{aligned}
 -2 \sum_{u=1}^n x_{ku} y_u + 2nb_0 \sum_{u=1}^n x_{ku} + 2b_1 \sum_{u=1}^n x_{1u} x_{ku} + 2b_2 \sum_{u=1}^n x_{2u} x_{ku} + \dots + 2b_k \sum_{u=1}^n x_{ku}^2 &= 0 \\
 nb_0 \sum_{u=1}^n x_{ku} + b_1 \sum_{u=1}^n x_{1u} x_{ku} + b_2 \sum_{u=1}^n x_{2u} x_{ku} + \dots + b_k \sum_{u=1}^n x_{ku}^2 &= \sum_{u=1}^n x_{ku} y_u
 \end{aligned}
 \tag{2-42}$$

Equations (2-40), (2-41) and (2-42) are called "Normal equations" as

$$\begin{aligned}
 nb_0 + b_1 \sum_{u=1}^n x_{1u} + b_2 \sum_{u=1}^n x_{2u} + \dots + b_k \sum_{u=1}^n x_{ku} &= \sum_{u=1}^n y_u \\
 nb_0 \sum_{u=1}^n x_{1u} + b_1 \sum_{u=1}^n x_{1u}^2 + b_2 \sum_{u=1}^n x_{1u} x_{2u} + \dots + b_k \sum_{u=1}^n x_{1u} x_{ku} &= \sum_{u=1}^n x_{1u} y_u \\
 &\vdots \\
 nb_0 \sum_{u=1}^n x_{ku} + b_1 \sum_{u=1}^n x_{1u} x_{ku} + b_2 \sum_{u=1}^n x_{2u} x_{ku} + \dots + b_k \sum_{u=1}^n x_{ku}^2 &= \sum_{u=1}^n x_{ku} y_u
 \end{aligned}
 \tag{2-43}$$

It is simpler to solve the normal equations if they are expressed in matrix notation as

$$Y = XB + E \quad \dots(2-44)$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1} \quad X = \begin{bmatrix} x_{01} & x_{11} & x_{21} & \dots & x_{k1} \\ x_{02} & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ x_{0k} & x_{1k} & x_{2k} & \dots & x_{kn} \end{bmatrix}_{n \times (k+1)}$$

$$B = \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \\ B_k \end{bmatrix}_{(k+1) \times 1} \quad E = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}_{n \times 1}$$

In general, Y is an $(n \times 1)$ vector of the responses, X is an $n \times (k+1)$ matrix of the levels of the independent variables, B is a $(k+1) \times 1$ vector of the regression coefficients and E is an $(n \times 1)$ vector of random errors.

The solution to the normal equations will be the least square estimators, b . Equation (2-43) may be written in matrix form as shown in Equation (2-45).

$$\begin{bmatrix} n & \sum x_{1u} & \sum x_{2u} & \dots & \sum x_{ku} \\ n \sum x_{1u} & \sum x_{1u}^2 & \sum x_{1u} x_{2u} & \dots & \sum x_{1u} x_{ku} \\ n \sum x_{2u} & \sum x_{1u} x_{2u} & \sum x_{2u}^2 & \dots & \sum x_{2u} x_{ku} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ n \sum x_{ku} & \sum x_{1u} x_{ku} & \sum x_{2u} x_{ku} & \dots & \sum x_{ku}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} \sum y_u \\ \sum x_{1u} y_u \\ \sum x_{2u} y_u \\ \vdots \\ \sum x_{ku} y_u \end{bmatrix} \quad \dots(2-45)$$

which simplifies to

$$(X'X)b = X'Y \quad \dots(2-46)$$

Equations (2-46) are the least squares normal equations. To solve normal equations, multiply both sides of Equation (2-46) by an inverse of $X'X$. Thus, the least square estimator of b is

$$b = (X'X)^{-1} X'Y \quad \dots(2-47)$$

where

$$X'Y = (iy) = \begin{pmatrix} 0y \\ 1y \\ \vdots \\ ky \end{pmatrix} \quad \dots(2-48)$$

$$(X'X)^{-1} = c_{ij} = \begin{pmatrix} c_{00} & c_{01} & \dots & c_{0k} \\ c_{10} & c_{11} & \dots & c_{1k} \\ \vdots & \vdots & & \vdots \\ c_{k0} & c_{k1} & \dots & c_{kk} \end{pmatrix} \quad \dots(2-49)$$

so Equation (4.15) can be rewritten as

$$b_i = \sum_{j=0}^k C_{ij}(iy) \quad \dots(2-50)$$

2.8 Analysis of Variance; ANOVA

The analysis of variance is an attempt to analyze the variation of a response and to assign portions of this variation to each of a set of independent variables. The reason is that response variables vary only because of the variation in a set of unknown independent variables. The ANOVA partitions the sum of square of deviations, called "Total sum of squares of deviations", into parts, each of which is attributed to one of the independent variables in the experiment, plus a remainder that is associated with random error.

$$SS_T = SS_R + SS_E \quad \dots(2-51)$$

SS_T = Total sum of square is overall variability in the data

SS_R = Regression sum of square is variable of responses y due to effect of independent variables x_1, x_2, \dots, x_3

SS_E = Error sum of square is variable of responses due to the influence of uncontrollable factor

This sum of square decomposition is frequently recorded in a table. Since this table concerns various sums of squares or estimates of variances, it is called the analysis of variance table or ANOVA table which is shown in Table 2.9.

Table 2.7 : ANOVA Table for Multiple Linear Regression.

Source of Variation	Sum of Square	Degree of Freedom	Mean Square	F ₀
Regression	SS _R	k(k+3)/2	MS _R	MS _R /MS _E
- First order terms	SS _{R1}	k	MS ₁	MS ₁ /MS _E
- Second order terms	SS _{R2}	k(k+1)/2	MS ₂	MS ₂ /MS _E
Error	SS _E	n-1-[k(k+3)/2]	MS _E	
- Lack of Fit	SS _{LOF}	n ₂ -[k(k+3)/2]	MS _{LOF}	MS _{LOF} /MS _{PE}
- Pure Error	SS _{PE}	n ₁ -1	MS _{PE}	
Total	SS _T	n-1		

From the ANOVA table derived total sum of square (SS_T) from the following

$$\begin{aligned}
 SS_T &= \sum_{i=1}^n (y_i - \bar{y})^2 \\
 &= \sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n} \quad \dots(2-52) \\
 &= \sum_{i=1}^n y_i^2 - \frac{G^2}{n}
 \end{aligned}$$

where \bar{y} = the mean of all responses y

$\sum_{i=1}^n y_i^2$ = Grand total of square

G = Grand total of responses y

$\frac{G^2}{n}$ = Correction factor (CF)

where n are total observations, thus SS_T has n-1 degrees of freedom.

Regression sum of square SS_R may be obtained as shown in Equation (2-53).

$$SS_R = b_0(0y) + \sum_{i=1}^k b_i(iy) + \sum_{i=1}^k \sum_{j=1}^k b_{ij}(ijy) - \frac{G^2}{n} \quad \dots(2-53)$$

It can also written in the matrix form as shown in Equation (2-54).

$$SS_R = b'X'Y - \frac{G^2}{n} \quad \dots(2-54)$$

Regression sum of square estimated with second order polynomial equation has $k(k+3)/2$ degree of freedom and can be categorized into

- a. First order terms regression sum of square (SS_{R1}) with degree of freedom k derived from Equation (2-55).

$$SS_{R1} = \sum_{i=1}^k b_i(iy) \quad \dots(2-55)$$

- b. Second order terms regression sum of square (SS_{R2}) with degree of freedom $k(k+1)/2$ derived from Equation (2-56).

$$SS_{R2} = b_0(0y) + \sum_{i=1}^k \sum_{j=1}^k b_{ij}(ijy) - \frac{G^2}{n} \quad \dots(2-56)$$

2.8.1 Lack of Fit

Regression models are often fit to data when the true functional relationship is unknown. Naturally, one would like to know whether the order of the model tentatively assumed is correct.

The danger of using a regression model is a poor approximation of the true functional relationship. Obviously, a polynomial of degree two or greater should have been used for this hypothetical situation. The result is that a very poor model would be obtained [5].

The error sum of square consists of two components as shown in Equation (2-57).

$$SS_E = SS_{PE} + SS_{LOF} \quad \dots(2-57)$$

where SS_{PE} is the sum of squares attributable to "pure" experimental error, and SS_{LOF} is the sum of squares attributable to the lack of fit of the model.

when

$$SS_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \dots(2-58)$$

Equation (2-58) may be written in matrix form as shown in Equation (2-59).

$$\begin{aligned} SS_E &= (Y - Xb)'(Y - Xb) \\ &= Y'Y - b'X'Y \end{aligned} \quad \dots(2-59)$$

A derivative of Equation (2-51) yields the SS_E as shown in Equation (2-60).

$$SS_E = SS_T - SS_R \quad \dots(2-60)$$

The pure error sum of squares, SS_{PE} , is computed from the responses obtained by repeated observations at the same level of x at the center point

(0, 0, ..., 0) of the central composite rotatable experimental design. The pure error sum of squares, SS_{PE} , could be obtained by

$$SS_{PE} = \sum_{u=1}^n (y_{1u} - \bar{y}_1)^2 \quad \dots(2-61)$$

when y_{1u} is the response of each experiment at the center point. The sum of square for lack of fit is as shown in Equation (2-62).

$$SS_{LOF} = SS_E - SS_{PE} \quad \dots(2-62)$$

2.9 Hypothesis Testing in Multiple Linear Regression

The test for significance of regression is a test to determine if there is a linear relationship between the response y and a subset of the independent variables, x_1, x_2, \dots, x_k . The appropriate hypotheses are

$$H_0 : B_1 = B_2 = \dots = B_k = 0$$

$$H_1 : B_i \neq 0 \text{ for at least one } i \quad \dots(2-63)$$

The total sum of squares, SS_T , is partitioned into regression and error sums of squares, SS_E and SS_R are independent. Therefore, the hypothesis is tested by calculating F_0 :

$$F_0 = \frac{\frac{SS_R}{k(k+3)}}{\frac{SS_E}{n-1-\frac{k(k+3)}{2}}} = \frac{MS_R}{MS_E} \quad \dots(2-64)$$

where MS_R and MS_E are called mean square of regression and error respectively

Rejection of H_0 in Equation (2-62) if $F_0 > F_{\alpha, v_1, v_2}$ implies that at least one variable in the model contributes significantly to the fit. On the other hand, in case H_0 is accepted, if $F_0 < F_{\alpha, v_1, v_2}$; H implies that response has no relationship with independent variables x_1, x_2, \dots, x_k .

F_{α, v_1, v_2} could be obtained from table in index D in which α means the level of significance at a degree of confidence $1 - \alpha$. The term v_1 is the degree of freedom of the numerator in Equation (2-63) and v_2 is the degree of freedom of the denominator.

The hypothesis can also be tested by using F_0 shown in Equation (2-64).

$$F_0 = \frac{\frac{SS_{LOF}}{k(k+3)}}{\frac{SS_{PE}}{n_1 - 1}} = \frac{MS_{LOF}}{MS_{PE}} \quad \dots(2-65)$$

From the hypotheses

H_0 : The model adequately fits the data

H_1 : The model does not fit the data ...(2-66)

If $F_0 > F_{\alpha, v_1, v_2}$; rejecting H_0 , this indicates that the model does not fit the data. If $F_0 < F_{\alpha, v_1, v_2}$ accepting H_0 , this indicates that the model fits the data.

There is frequent interest in testing hypotheses on the individual regression coefficients. Such test would be useful in determining the value of each of the independent variables in the regression model [5].

The hypotheses for testing the significance of any individual regression coefficient, b_i , are

$$\begin{aligned} H_0 : b_i &= 0 \\ H_1 : b_i &\neq 0 \end{aligned} \quad \dots(2-67)$$

If $H_0: b_i = 0$ is not rejected, then this indicates that x_i can be deleted from the model. The appropriate test statistic for this hypothesis is

$$t_0 = \frac{b_i}{\sqrt{MSE C_{ii}}} \quad \dots(2-68)$$

The hypothesis $H_0: b_i = 0$ is rejected if $|t_0| > t_{\alpha/2, v_1}$. Where C_{ii} is the diagonal element of $(X'X)^{-1}$ corresponding to b_i and v is the degree of freedom of error. It should be noted that this is really a partial or marginal test, because the regression coefficient, b_i , depends on all the other regressor variables x_j ($i \neq j$) that are in the model [8].

2.10 Coefficient of Determination: R^2

The coefficient of determination is defined as follows:

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \quad \dots(2-69)$$

It measures the proportionate reduction of total variation in Y associated with the use of the set of X variables x_1, x_2, \dots, x_k . Since $0 \leq SS_E \leq SS_T$, it follows that

$$0 \leq R^2 \leq 1 \quad \dots(2-70)$$

R^2 assumes the value 0 when all $b_i = 0$ ($i = 0, 1, \dots, k$). R^2 takes on the value 1 when all observations fall directly on the fitted response surface, that is, when $y_u = \hat{y}_u$ for all u [7].

A large R^2 does necessarily imply that the fitted model is a useful one, for instance, observations may have been taken at only a few levels of the independent variables. Despite a high R^2 in this case, the fitted model may not be useful because most predictions would require extrapolations outside the region of observations. Again, even though R^2 is large, MS_E may still be too large for inferences to be useful in a case where high precision is required [7].

Adding more independent variables to the model can only increase R^2 and never reduce it, because SS_E can never become larger with more independent variables and SS_T is always the same for a given set of responses. Since R^2 often can be made large by including a large number of independent variables, it is sometimes suggested that a modified measure be used which recognizes the number of independent variables in the model.

This "Adjusted coefficient of determination", denoted R_a^2 , is defined as shown in Equation (2-70)

$$R_a^2 = 1 - \frac{\frac{SSE}{(n-k-1)}}{\frac{SST}{(n-1)}} = 1 + \frac{(n+1)}{(n-k-1)}(R^2 - 1) \quad \dots(2-71)$$