INTRODUCTION

we review some of the fundamental ideas required by least squares. No attempt is made to completely cover the topic, since many excellent treatments are already available [6, 29a, 45].

I. The general problem.

We begin by assuming that we have a function

$$y = f(x_1, x_2, \ldots, x_m; \alpha_1, \alpha_2, \ldots \alpha_p). \qquad \ldots \qquad (1)$$

we are also given a set of n observations $(y_i, x_{1i}, x_{2i}, \ldots x_{mi})$, where $i = 1, 2, \ldots, n$ and $n > p$. The variable y is called the dependent variable, while the x's are the independent variables. Problem is to to determine estimates of the p parameters $\alpha_K (k=1, 2, \ldots, p)$. Of the many possible estimates of the α_K, we decide to choose the set of numbers that will minimize the sum of squares of the deviations of the observed y_i from the function. Thus, we wish to minimize

$$Q = \sum_{i=1}^{n} w_i [y_i - f(x_{1i}, \ldots, x_{mi}; \alpha_1, \alpha_2, \ldots, \alpha_p)]^2, \qquad \ldots (2)$$

where w_i are the weights associated with each of the y_i. Let us assume that each y_i has come from some distribution with mean

$$E(y_i) = f(x_{1i}, \ldots, x_{mi}; \alpha_1, \alpha_2, \ldots, \alpha_p) \qquad (3)$$

and variance proportional to a given function of $(x_1, \ldots, x_m)$, i.e.,

$$\sigma^2(y_i) = \sigma^2 h^2(x_{1i}, \ldots, x_{mi}) \qquad (4)$$

Here we have added a new parameter, σ^2, which (unless otherwise

specified) must be estimated. Setting

$$w_i = I/\, h^2(x_{Ii},\ldots,x_{mi}) = \sigma^2/\sigma^2(y_i) \qquad (5)$$

and assuming the y_i are each from a normal (gaussian) distribution with mean and variance given by (3) and (4), we are able to get an unbiased estimate of σ^2 with

$$s^2 = \sum_{i=1}^{n} w_i[y_i - f(x_{Ii},\ldots,x_{mi};\, a_I,a_2,\ldots,a_p)]^2/(n-p) \qquad (6)$$

where a_k's are the estimates of α_K's .

When the function $h^2(x_I,\ldots,x_m)$ is constant for all values of the x_i. we say that we have equally weighted data. We may assume these weights are constant and equal to I. From a purely standpoint, we can minimize the sum of squares Q in (2) with respect to the α_K by differentiating Q with respect to α_K, setting the derivatives equal to zero. and solving the resulting set of p simultaneous equations. these equations can be written

$$\frac{\partial Q}{\partial \alpha_K} = -2\sum_{i=1}^{n}\left(\frac{\partial f}{\partial \alpha_K}\right)_i [y_i - f(x_{Ii},\ldots,x_{mi};\alpha_1,\ldots,\alpha_p)] = 0 \qquad (7)$$

for k=I 2...,p and where $\left(\frac{\partial f}{\partial \alpha_K}\right)_i$ denotes the value of the kth partial derivative for the ith data point. Transposing and setting $f_i = f(x_{Ii},\ldots,x_{mi};\alpha_1,\ldots,\alpha_p)$, we obtain the more usual form

$$\begin{aligned} \sum_i^n f_i(\partial f/\partial \alpha_1)_i &= \sum_i^n y_i(\partial f/\partial \alpha_1)_i \,, \\ \sum_i^n f_i(\partial f/\partial \alpha_2)_i &= \sum_i^n y_i(\partial f/\partial \alpha_2)_i \,. \end{aligned} \qquad (8)$$

$$\sum_{i}^{n} f_i (\partial f/\partial \alpha_p)_i = \sum_{i}^{n} y_i (\partial f/\partial \alpha_p)_i \quad .$$

These are called the "normal equations" of the problem. The form (8) is, in general, a system of p nonlinear equations; and there is no guarantee that a solution to the system exists or that, if a solution does exist, it is unique.

2. The linear problem.

There is a large class of functions for which solutions do exist. An important subset of these functions is that in which tha parameters α_K appear as linear coefficients of the independent variables. The estimation of the parameters can be done by the well known linear multiple regression. Here, the general model can be written

$$f(x_I, \ldots, x_m; \alpha_1, \ldots, \alpha_p) = \alpha_1 g_I(x_I, \ldots, x_m) + \alpha_2 g_2(x_I, \ldots, x_m) + \ldots + \alpha_p g_p(x_I, \ldots, x_m) \qquad (9)$$

where g_k are any well behaved functions of the independent variables Hald [9a]. Special cases for example:

$$f(x_I, \ldots, x_m; \alpha_1, \ldots, \alpha_p) = \alpha_1 x_I + \alpha_2 x_2 + \ldots + \alpha_p x_p \quad , \qquad (10)$$

and $$f(x_I \ldots, x_m; \alpha_1, \ldots, \alpha_p) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_q x^q \qquad (11)$$

are very important and they are dealt nicely in many standard text books in regression analysis.

Even when the parameters do not naturally appear linearly. it is often possible to make a transformation of either the dependent or independent variables so that the parameters (or simple transformation of them) satisfy the form (8). For example, we can transform the function $Y=c_I \exp(c_2 x)$ by taking natural logarithms of both sides to obtain $\ln Y = \ln c_I + c_2 x$. Setting $y=\ln Y$, $\alpha_1 = \ln c_I$ and $\alpha_2 = c_2$, we have a function of the general linear form.

3. The non-linear problem.

When a function is such that 'linearization' of the parameters is not feasible, it is clear that another approach is needed. One such method is due to Gauss and it is known as Gauss method. The Gauss method [4Ia] consists essentially of linearizing the desired function with respect to each of the parameters by means of a truncated Taylor's series. Using initial estimates of the parameters to evaluate the coefficients of the expansion, new estimates are obtained. The process is repeated until some convergence criterion is satisfied. The method will be shown to amount to repeated application ofthe form given by equation (I0).

Suppose, now, that our function is of the form in equation (I) and that we wish to minimize Q in (2). Suppose further that we have initial estimates of the parameters, and let us denote the estimate of the kth parameter by $a_{k,o}$. Thus, our set of estimates can be thought of as a point $(a_{I,o}, a_{2,o}, \ldots, a_{p,o})$ in the p-dim-

ensional parameter space. If we expand Equation (I) in a Taylor's series about this point, we obtain, for each i (i=I,2,...,n)

$$\Delta y_{i,o} = y_i - f(x_{Ii}, x_{2i}, \ldots, x_{mi};\ a_{I,o}, a_{2,o}, \ldots, a_{p,o})$$

$$= \left(\frac{\partial f}{\partial \alpha_1}\Big|_{i,o}\right)\Delta a_{1,1} + \left(\frac{\partial f}{\partial \alpha_2}\Big|_{i,o}\right)\Delta a_{2,1} + \cdots + \left(\frac{\partial f}{\partial \alpha_p}\Big|_{i,o}\right)\Delta a_{p,1},$$

$$= \sum_{1}^{p}\left(\frac{\partial f}{\partial \alpha_k}\Big|_{i,o}\right)\Delta a_{k,I}, \qquad (I2)$$

where $\left(\frac{\partial f}{\partial \alpha_k}\Big|_{i,o}\right)$ means that the partial derivative of (I) with respect to α_k is evaluated at $(a_{I,o}, a_{2,o}, \ldots, a_{p,o})$ for the ith data point, and the $\Delta y_{i,o}$ are the differences between y_i and the value of the function for the ith set of independent variables and the estimates of the α_k. We have now reduced the problem to one to which linear methods can be applied. The $\Delta y_{i,o}$ are the dependent variables, the $\left(\frac{\partial f}{\partial \alpha_k}\Big|_{i,o}\right)$ are the independent variables, and the $\Delta a_{k,I}$ are the parameters to be estimated. The normal equations (8) become

$$\Delta a_{1,1}\left[\sum_{i=1}^{n}\left(\frac{\partial f}{\partial \alpha_j}\Big|_{i,o}\right)^2\right] + \cdots + \Delta a_{p,1}\left[\sum_{i=1}^{n}\left(\frac{\partial f}{\partial \alpha_j}\Big|_{i,o}\right)\left(\frac{\partial f}{\partial \alpha_p}\Big|_{i,o}\right)\right]$$

$$= \sum_{i}^{n}\Delta y_{i,o}\left(\frac{\partial f}{\partial \alpha_j}\Big|_{i,o}\right), \qquad (I3)$$

for j=I 2,...,p.

The system (I3) is a set of p linear equations in p unknown. If we get a solution to the system (I3), we have obtained a set of values $\Delta a_{k,I}$ with which to modify each of the $a_{k,o}$. We do this by applying

$$a_{k,I} = a_{k.o} + \Delta a_{k,I} \quad . \qquad (I4)$$

The "improved" estimates of the a_k are then placed into equations (I3) and (I4), and the process is repeated until, after q iterations, the $\Delta a_{k.q}$ are all deemed "sufficiently small". When this occurs, we say that the process has converged and we take $a_{k,q}$ to be the least squares estimate of α_k.

4. Statement of the problem.

Earlier Stevens W.L. [37] has described a least-squares (Gauss method) technique for estimating the parameter ρ in the curve $E(y) = \alpha + \beta \rho^x$. It is expected that in many problems results may be improved by adding a linear term to $E(y) = \alpha + \beta \rho^x$ (Shah B.K. and Khatri C.G. [33]) . Thus, in this thesis the problem of estimating non-linear parameter in $E(y) = \alpha + \delta x + \beta \rho^x$ is considered. It is shown in chapter I, Shah B.K. and Patel I.R. [34] that only one initial estimate of the parameter ρ is necessary to obtained the least-squares solution using the Gauss method. However, this method requires a preliminary estimate of the non-linear parameter ρ and exactly how accurate must be the initial

estimate of the parameter ? Sometimes in general we have obtained convergence with estimates off by as much as a factor of 10^5; we have failed to obtain convergence when the estimates were within 20 % of what finally proved to be the correct values. At this time, it seems impossible to provide an all purpose answer to the general problem. Certainly it is obvious that some kind of estimates are needed to start the procedure. This decribed in subsequent chapters (chapters 2 and 3) in this thesis. Hartley [10,11] has described another method to find the estimates of the non-linear curves. This method is known as internal least squares. This is a very good method for obtaining initial estimates but is very complicated and not suitable on a desk calculators. Tootill [37,38,39], Richard [30], Cornell R.G.[3,4], Croxeton F.E.and Cowden D.J.[5], Linhart [18] and various other authors have described computational procedures for the linear combination of exponential family of curves.

In chapter 2, we have considered a very simple estimator Shah B.K. [31], based on ratio of two linear functions of y's. Efficiencies and biases in the estimators under the constant variance model are also given in Tables 2.1 and 2.2 respectively. When equation $E(y) = \alpha + \delta x + \beta \rho^x$ represents a biological growth, x is measure of time, and then it may be appropriate to incorporate this curve into increasing variance model that changes with time. In this situation theory of Brownian movement as discussed by S.Chandrasekar [2] is considered in detailed in this chapter.

In chapter 3, we have considered the estimate of ρ under two alternative methods,(Shah B.K. and Khatri C.G. [32]): (i) Patterson's [26] method of estimating ρ by considering a ratio of two quadtratic functions of y's and (ii) modified Hartley's method suggested by Khatri C.G. and Shah B.K. [I6]. These are known as Quadratic estimators. Various theoritical results are established in this chapter in more detail. In Quadratic estimators the efficiencies are increased. Out of the two method, modified Hartley's method is more efficient with zero bias than Patterson's Quadratic estimators. But the formula for estimating ρ, a ratio of quadratic functions of y's, is not suitable for practical purposes i.e , to say that it is not easy to calculate the estimator on a desk calculator.One can apply modified Hartley's estimator as an initial starting value in the Gauss method of iteration on a modern high speed computer, Shah B.K. [3Ia] .

In chapter 4, we have described how to obtain a quick estimate of ρ by solving a quadratic equation, which can be obtained by using the values of U_x and V_x , given in Table 4.I, in r. We can see that the efficienciesare increased than those of the quadratic estimators (Shah B.K. and Khatri C.G. [33]).

We have gereralised the method of chapter 4 in chapter 5. Here the estimate of the nonlinear parameter is

obtained by solving a cubic equation, which can be obtained by using the tabulated values of U_x, V_x and Z_x given in Table 5.I. (Shah B.K. and Khatri C.G. [33a]), in r. This method is found to be more efficient than all the previous existing method for large number of equally spaced observations. The efficiency by this method is very large nearly about 99.0% for all values of for n=I4. In this chapter procedure is described to obtain the values of U_x, V_x and Z_x for large n. Tables of U_x, V_x and Z_x can be prepared for large n>I4 using modern high speed computers. It is to be noted that the efficiences in the neighbourhood of ρ = 0.2, 0.5 and 0.8 are maximum and nearly 100.0 percent for all values of n.

In the last chapter, we have illustrated how a linear combination of exponential family of curves and even more complicated extensions of the exponential family of curves, can be solved using some modification in Harltey's [IO] method. Recently Lipton S. and Mcgilchrist [2I 2Ia] have studied the linear combination of exponential family of curves using Hartley's internal least squares method without giving reference to the earlier work made by Khatri C.G. and Shah B.K. [I6].