In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

7/25/68

MINIMUM BIAS ESTIMATION OF THE

SLOPE OF A RESPONSE SURFACE

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

by

Richard James Matteson

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in the School of Industrial Engineering

Georgia Institute of Technology

June 1972

ACKNOWLEDGMENTS

I wish to extend my heartfelt thanks to Dr. Douglas C. Montgomery both for suggesting the problem to me and for his advice and assistance during the conduct of the research and the writing and editing of this thesis.

I also gratefully acknowledge the interest and guidance of Drs. Donovan Young and Russell Heikes, members of my reading committee.

Credit is due also to Mrs. Peggy Weldon for her skillful assistance in the typing, composition and publication of this thesis.

Finally, to my family, most especially to my wife, Barbara, I offer my loving appreciation for their tolerance of a husband and father who was too often distracted and short of both temper and patience during the completion of this work.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

This thesis derives necessary and sufficient conditions for designs to estimate the slope of a response surface with minimal bias contribution to expected mean square error. It also investigates the traditional least squares estimation all-bias design for estimating slope and presents a necessary condition for the design which is easily calculated for a spherical region of interest.

The investigation demonstrates that in the case of a linear approximation to a quadratic response, the easily generated least squares estimate of slope is unbiased for an all-bias design and is independent of design parameters. Two examples illustrate this significant result. They also show the importance of considering the full true relationship in selecting designs to estimate slope by minimum bias estimation.

CHAPTER I

INTRODUCTION

1.1 Background

Polynomial models are widely used in the physical sciences and in engineering to approximate the response of a system. The procedure usually is to gather observational data relating a response, say y, to a set of k independent variables, x_1, $x_2, \ldots, x_k$, and then to fit an appropriate polynomial by least squares. One problem in this procedure is deciding on the degree of polynomial required. The desire for simple models often leads the experimenter to select a polynomial of too low a degree to adequately represent the mechanism producing the response. When the polynomial is of too low a degree, the error in the response includes not only random observational error, or variance, but also systematic error due to an inadequate model, or bias.

In 1958, Box and Draper (2) demonstrated that in certain situations the bias error resulting from the choice of an inadequate approximating polynomial is a more significant contributor to expected mean square error than is variance. That disclosure stimulated considerable research, because at that time the criteria for selecting a response surface design were based on minimization of variance and considered bias only as a secondary measure of effectiveness. The main result of Box and Draper's research was the class of all-bias designs which attempted to select design parameters such that the bias contribution to expected mean square error was minimized.

These all-bias or near all-bias designs had a smaller expected mean square error than designs resulting from classical criteria (2). Further research indicated that in some, if not most, cases a still smaller expected mean square error might be achieved. In 1969, Karson, et al., (3) introduced the concept of minimum bias estimation which minimized the bias component of expected mean square error by the choice of the estimator. By using the resulting design flexibility to minimize variance, they were able to achieve a smaller expected mean square error than with the all-bias design. The result is a more effective design.

1.2 Estimating the Slope of a Response Surface

The data collected in an experiment, i.e., the response variable, is not always the variable of interest to the experimenter. In meteorology, for example, researchers use balloons to measure wind velocity. The balloons are light enough that the velocity of the balloon may be taken as the actual wind velocity. It is not possible to measure the velocity of the balloon directly, but its exact position can be determined very accurately. In this example the first derivative of the data, i.e., the slope of the response surface, is the variable of interest.* Similar problems occur in rocketry where the observed variable is position and the variable of interest is the first derivative (velocity) or the second derivative (acceleration) of the data.

The development of effective designs to estimate slope has been considered from several viewpoints. Atkinson (1) proposed designs to estimate the slope of a response surface. He applied the Box-Draper effectiveness criterion and used least squares estimation. Lure and

---

*Example taken from Lure and Wenning (4).

Wenning (4) investigated the degree of polynomial required to provide a sufficiently accurate estimate of the first derivative of a data set. They also used least squares to estimate the unknown parameters.

1.3 Statement of the Problem

Previous efforts to develop designs to estimate the slope of a response surface have concentrated on least squares as the estimation technique. It would seem reasonable to expect that a smaller expected mean square error could be achieved when minimum bias estimation is applied to the problem. In fact, such a result has been shown when the response is represented by a polynomial containing only one independent variable (5). The main thrust of this investigation will be to apply the concept of minimum bias estimation to estimate the slope of a multifactor response surface.

CHAPTER II

LITERATURE SURVEY

Munsch (5) has provided an excellent and thorough survey of the literature pertinent to experimental design criteria in response surface methodology. He also organized much of the more important work and examined relationships among the various concepts and criteria. The interested reader is referred to that work. This chapter will present only those portions of the published works upon which this research is based.

2.1 Box and Draper

Box and Draper (2) proposed the criterion of minimization of expected mean square error over a specified region of interest as a measure of design effectiveness. They further showed that the expected mean square error, say J, may be expressed as the sum of two components: V, the variance and B, the squared bias.

Let $\eta(\underline{\xi})$ be a response which is a function of $\underline{\xi}$, the vector of independent variables, i.e., $\underline{\xi}' = (\xi_1, \xi_2, \ldots, \xi_p)$. We assume that $\eta(\underline{\xi})$ may be truly represented by a polynomial of degree d_2. We will approximate $\eta(\underline{\xi})$ with the quantity $\hat{y}\,(\underline{\xi})$ where $\hat{y}\,(\underline{\xi})$ is represented by a polynomial of degree $d_1 < d_2$. The expected squared difference $E\{\hat{y}(\underline{\xi} - \eta(\underline{\xi})\}^2$ is to be minimized on the average over the region R. We transform the vector $\underline{\xi}$ to a vector, $\underline{x}$, such that the center of the region R is at the origin of the x's. Thus $\underline{x}' = (x_1, x_2, \ldots, x_p)$, and the expected squared difference becomes

$$E\{\hat{y}(\underline{x}) - \eta(\underline{x})\}^2$$

To average the squared difference over R, we integrate over R and divide the result by the volume of the region R as follows:

$$K \int_R E\{\hat{y}(\underline{x}) - \eta(\underline{x})\}^2 \, d(\underline{x})$$

where

$$d(\underline{x}) = d(x_1), d(x_2), \ldots, d(x_p)$$

and

$$K^{-1} = \int_R d(\underline{x}) \; .$$

In order to compare designs with different numbers of points we write the design effectiveness criterion, J, as

$$J = \frac{NK}{\sigma^2} \int_R E\{\hat{y}(\underline{x}) - \eta(\underline{x})\}^2 \, d(\underline{x}) \; . \qquad (2.1.1)$$

One then attempts to find experimental designs which will minimize J for a specific situation and model.

Clearly, J may be more simply expressed. Squaring the quantity in brackets in (2.1.1), we obtain

$$[\hat{y}(\underline{x})]^2 - 2\hat{y}(\underline{x})\eta(\underline{x}) + [\eta(\underline{x})]^2 =$$

$$[\hat{y}(\underline{x})]^2 - E^2[\hat{y}(\underline{x})] + E^2[\hat{y}(\underline{x})] - 2\hat{y}(\underline{x})\eta(\underline{x}) + [\eta(\underline{x})]^2.$$

Taking expectation

$$E[\hat{y}(\underline{x})]^2 - E^2[\hat{y}(\underline{x})] + E^2[\hat{y}(\underline{x})] - 2E[\hat{y}(\underline{x})]\eta(\underline{x}) + [\eta(\underline{x})]^2 ,$$

thus equation (2.1.1) becomes

$$J = \frac{NK}{\sigma^2}\int_R \left\{E[\hat{y}(\underline{x})]^2 - E^2[\hat{y}(\underline{x})]\right\} d(\underline{x}) + \frac{NK}{\sigma^2}\int_R \left\{E[\hat{y}(\underline{x})] - \eta(\underline{x})\right\}^2 d(\underline{x}) \ . \quad (2.1.2)$$

The first term in the right hand side of equation (2.1.2) is recognized as the variance of $\hat{y}(\underline{x})$ averaged over the region R, which was called V. Likewise, the second term is the squared bias averaged over R, which was called B. Therefore

$$J = V + B \ .$$

2.2 Minimum Bias Estimation

Karson, et al., (3) proposed a different approach to the problem of minimizing J, the expected mean square error. They assumed the presence of bias due to higher order terms than those present in the model and adopted a method of estimation aimed directly at minimization of that bias.

Recall equation (2.1.2). Define e to be a random variable such that $E(e) = 0$ and $Var(e) = \sigma^2$. The bias term is

$$B = \frac{NK}{\sigma^2}\int_R \left\{E[\hat{y}(\underline{x})] - \eta(\underline{x})\right\}^2 d(\underline{x}) \quad (2.2.1)$$

where

$$\eta(\underline{x}) = \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2 + e \ , \quad (2.2.2)$$

is now a polynomial of degree $d + k - 1$ and

$$\hat{y}(\underline{x}) = \underline{x}_1'\underline{b}_1 \tag{2.2.3}$$

is a fitted polynomial of degree d - 1. Then $E[\hat{y}(\underline{x})]$ is a polynomial of degree d - 1, say

$$E[\hat{y}(\underline{x})] = \underline{x}_1'\underline{\alpha}_1 \ .$$

So equation (2.2.1) may be written as

$$\begin{aligned} B &= \frac{NK}{2}\int_R [\underline{x}_1'(\underline{\alpha}_1 - \underline{\beta}_1) - \underline{x}_2'\underline{\beta}_2]^2 d(\underline{x}) \\ &= \frac{NK}{2}\int_R [(\underline{\alpha}_1 - \underline{\beta}_1)'\underline{x}_1\underline{x}_1'(\underline{\alpha}_1 - \underline{\beta}_1) - 2(\underline{\alpha}_1 - \underline{\beta}_1)'\underline{x}_1\underline{x}_2'\underline{\beta}_2 + \underline{\beta}_2'\underline{x}_2\underline{x}_2'\underline{\beta}_2]d(\underline{x}) \\ &= \frac{N}{\sigma^2}[(\underline{\alpha}_1 - \underline{\beta}_1)'\mu_{11}(\underline{\alpha}_1 - \underline{\beta}_1) - 2(\underline{\alpha}_1 - \underline{\beta}_1)'\mu_{12}\beta_2 + \underline{\beta}_2'\mu_{22}\underline{\beta}_2] \ , \end{aligned} \tag{2.2.4}$$

where

$$\mu_{11} = K\int_R \underline{x}_1\underline{x}_1' d(\underline{x}) \tag{2.2.5a}$$

$$\mu_{12} = K\int_R \underline{x}_1\underline{x}_2' d(\underline{x}) \tag{2.2.5b}$$

$$\mu_{22} = K\int_R \underline{x}_2\underline{x}_2' d(\underline{x}) \ . \tag{2.2.5c}$$

Differentiating (2.2.4) with respect to $\underline{\alpha}_1$ and equating the results to zero yields

$$\frac{\partial B}{\partial \underline{\alpha}_1} = 2\mu_{11}(\underline{\alpha}_1 - \underline{\beta}_1) - 2\mu_{12}\underline{\beta}_2 = 0 \ ,$$

which implies

$$\underline{\alpha}_1 = \underline{\beta}_1 + \mu_{11}^{-1}\mu_{12}\underline{\beta}_2 \qquad (2.2.6)$$

$$= A\underline{\beta} ,$$

where

$$A = [I|\mu_{11}^{-1}\mu_{12}]$$

$$\underline{\beta}' = [\underline{\beta}_1'|\underline{\beta}_2'] .$$

The condition for minimum B shown in (2.2.7) is subject only to the quantity $A\underline{\beta}$ being estimable.

Now, if we let $\underline{y}$ be the vector of N observations (data vector), then $\underline{b}$ may be expressed as a linear transformation of $\underline{y}$, say

$$\underline{b}_1 = T'\underline{y} .$$

Referring to equations (2.2.2) and (2.2.3) we define X_1 as the matrix of values taken by the terms of $\underline{x}_1$ for the N experimental combinations, X_2 is a similar matrix for the values taken by the terms of $\underline{x}_2$, and $X = [X_1|X_2]$. Then, $E(\underline{y})$ may be expressed as

$$E(\underline{y}) = X\underline{\beta} \qquad (2.2.8)$$

which implies

$$T'X = A .$$

We may solve for the minimum value of B by substituting from (2.2.6) into (2.2.4) to obtain

$$\text{Min } B = \frac{N}{\sigma^2}\,\underline{\beta}_2'(\mu_{22} - \mu_{12}'\mu_{11}^{-1}\mu_{12})\underline{\beta}_2 .$$

Karson also points out that for the case where $T' = (X_1'X_1)^{-1}X_1'$ the Box-Draper condition is satisfied, i.e.,

$$(X_1'X_1)^{-1}X_1'X_2 = \mu_{11}^{-1}\mu_{12} .$$

The all-bias design is therefore shown to be a special case of the class of designs which minimize B.

2.3 Estimating the Slope of a Response Surface

The work of interest on this topic is that of Atkinson (1). He proposed designs to estimate the slope of a response surface; that is, the first derivative of the data. His designs are based on the Box-Draper effectiveness criterion, and he uses substantially the same approach to the problem. Of particular interest is his general discussion of the case where the data is equated to a polynomial in several independent variables,

$$\hat{y} = \underline{x}_1'\underline{b}_1 ,$$

where

$$\underline{x}_1 = (1, x_1, x_2, \ldots, x_p) .$$

The least squares estimates of the parameters are

$$\underline{b}_1 = (X_1'X_1)^{-1}X_1'\underline{y} , \qquad (2.3.1)$$

where X_1 and $\underline{y}$ are defined as in section 2.2 above.

The true response is

$$\eta = \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2 + e$$

The estimated slopes at a point $\underline{x}_0$ are

$$\underline{g} = \{g_i(\underline{x}_0)\} = \left\{\frac{\partial(y)}{\partial x_i}\bigg|_{\underline{x}=\underline{x}_0}\right\}$$

The true slopes at $\underline{x}_0$ are

$$\underline{\gamma} = \{\gamma_i(\underline{x}_0)\} = \left\{\frac{\partial(\eta)}{\partial x_i}\bigg|_{\underline{x}=\underline{x}_0}\right\}$$

For example, if the model is a first order approximation to a quadratic response which is a function of three independent variables ($p = 3$), then

$$\begin{aligned}
\hat{y} &= \underline{x}_1\underline{b}_1 \\
\underline{g} &= \underline{b}_1 \qquad\qquad (2.3.2) \\
\eta &= \underline{x}_1\underline{\beta}_1 + \underline{x}_2\underline{\beta}_2 \\
\underline{\gamma} &= \underline{\beta}_1 + D(\underline{x}_0)\underline{\beta}_2 ,
\end{aligned}$$

where

$$\begin{aligned}
\underline{x}_1 &= (1,\ x_1,\ x_2,\ x_3) \\
\underline{b}_1 &= (b_0,\ b_1,\ b_2,\ b_3) \\
\underline{\beta}_1 &= (\beta_0,\ \beta_1,\ \beta_2,\ \beta_3) \\
\underline{x}_2 &= (x_1^2,\ x_2^2,\ x_3^2,\ x_1x_2,\ x_1x_3,\ x_2x_3) \\
\underline{\beta}_2 &= (\beta_{11},\ \beta_{22},\ \beta_{33},\ \beta_{12},\ \beta_{13},\ \beta_{23}) \\
D(\underline{x}_0) &= \begin{bmatrix} 2x_{10} & 0 & 0 & x_{20} & x_{30} & 0 \\ 0 & 2x_{20} & 0 & x_{10} & 0 & x_{10} \\ 0 & 0 & 2x_{30} & 0 & x_{10} & x_{20} \end{bmatrix} .
\end{aligned}$$

The vector of biases may be calculated from

$$E(\underline{g}) - \underline{\gamma} \, .$$

Substituting from (2.3.1) into (2.3.2)

$$E(\underline{g}) = E\left\{(X_1'X_1)^{-1}X_1'\underline{y}\right\}$$
$$= (X_1'X_1)^{-1}X_1'E[\underline{y}] \, ,$$

and substituting from (2.2.8) yields

$$E(\underline{g}) = \underline{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\underline{\beta}_2 \, , \qquad (2.3.3)$$

where X_2 is defined as in section 2.2 above. We now use the results of (2.3.3) to write the vector of biases at $\underline{x}_0$ as

$$E(\underline{g}) - \underline{\gamma} = [(X_1'X_1)^{-1}X_1'X_2 - D(\underline{x}_0)]\underline{\beta}_2 \, . \qquad (2.3.4)$$

Atkinson carries the expression forward one more step placing it in terms of adjusted x variables, but the above equation is sufficient for our purposes. It is worth emphasizing that (2.3.4) is the vector of biases at the point $\underline{x}_0$. The next step by Atkinson was to derive a design criterion based on minimizing expected mean square error. He proceeded by first specifying a direction in which slope is to be estimated and then minimizing error after averaging over all such directions.

The details of Atkinson's method are not critical to this investigation, but his conclusions are vital. He found that the best design for estimating the slope over any region centered at that point. Further, he demonstrated that designing experiments to estimate the slope of a surface in any specific direction is the same as designing experiments to obtain precise estimates of the slopes along the factor axes.

These conclusions are used as starting points to develop a simpler criterion statement using least squares estimation (section 3.1 below). Atkinson's second conclusion is interpreted to mean that the slope at any point in the region may be expressed as a vector of components of slope parallel to the factor axes. The bias vector is similarly expressed and is averaged over the entire region of interest. This technique includes of necessity averaging over all directions. The transformations made on independent variables heretofore and carried through in the following chapters are precisely those which center the region of interest at the point where the slope is to be estimated.

2.4 Minimum Bias Estimation of the Slope of a Response Surface

In a 1971 master's thesis Munsch (5) applied the concept of minimum bias estimation to the problem of estimating the slope of a response surface. He limited his development to the case where the data is a function of one independent variable only. By restricting the analysis to the case of one independent variable, he was able to express the conditions for minimum bias estimation in terms of the matrices μ_{11} and μ_{12} defined in equation (2.2.5a-c). Thus Munsch's results closely resemble those of Karson.

When one attempts to extend Munsch's development to the case where the data is a function of k independent variables, $k > 1$, the matrices, μ_{ij}, are no longer convenient. Munsch points out that the μ_{ij}, matrices may be defined as moment matrices of a uniform probability distribution over the region R. In the following chapter a different set of matrices will be defined which may be thought of as moment matrices of the slope of the same distribution over the same region.

CHAPTER III

MINIMUM BIAS ESTIMATION OF THE SLOPE OF A RESPONSE SURFACE

3.1 Estimation by Least Squares, the All-bias Design

This section displays a development of the all-bias design for estimating slope that is closer in concept to the Box and Draper (2) method than is Atkinson's (1) approach. Definitions and concepts of use in later sections are introduced. The derivation depends on summing and averaging the components (parallel to the factor axes) of variance and squared bias. Clearly that technique cannot be used without assuming Atkinson's conclusions (see section 2.3).

Assume a response, say η, which is accurately represented by a polynomial of degree d in k ($k > 1$) independent variables, i.e.,

$$\eta = \beta_0 + \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2 + e ,$$

where e is a random variable such that $E(e) = 0$, $\mathrm{Var}(e) = \sigma^2$ and $\underline{x}' = (x_1, x_2, \ldots, x_k)$ is a vector of transformed factors as defined in section 2.1. Further, let $\underline{x}_2$ contain only terms of order d, and $\underline{x}_1$ contain all remaining terms. For example, if $k = 2$ and $d = 3$, then

$$\underline{x}_1' = (x_1,\ x_2,\ x_1^2,\ x_2^2,\ x_1x_2)$$

$$\underline{x}_2' = (x_1^3,\ x_2^3,\ x_1^2x_2,\ x_1x_2^2)$$

The true response, η, is to be approximated by a fitted polynomial of degree d -1, i.e.,

$$\hat{y} = b_0 + \underline{x}_1'\underline{b}_1 \ . \tag{3.1.1}$$

Suppose, however, that the variable of interest is the slope of the response η, say $\underline{\gamma} = \frac{d\eta}{d\underline{x}}$. Then the intercept term, β_0, is of no interest, and it is convenient to define a new response variable, say η^*, where

$$\eta^* = (\eta - \beta_0) = \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2 + e \ .$$

Clearly $\underline{\gamma}$ is unchanged, i.e.,

$$\underline{\gamma} = \frac{d\eta}{d\underline{x}} = \frac{d\eta^*}{d\underline{x}} = D_1\underline{\beta}_1 + D_2\underline{\beta}_2$$

where D_j is a matrix, the i^{th} row of which is obtained by the partial differentiation of $\underline{x}_j'$ with respect to x_i. The true slope, $\underline{\gamma}$, may be approximated by the slope of y, say $\underline{g}$. Here again the intercept, b_0, is of no interest. We may define

$$\hat{y}^* = (y - b_0) = \underline{x}_1'\underline{b}_1 \tag{3.1.2}$$

without changing $\hat{\underline{g}}$, i.e.,

$$\hat{\underline{g}} = \frac{d\hat{y}}{d\underline{x}} = \frac{d\hat{y}^*}{d\underline{x}} = D_1\underline{b}_1 \ .$$

The expected squared difference between the true slope $\underline{\gamma}$ and its approximation $\underline{g}$ is

$$E(\hat{\underline{g}} - \underline{\gamma})^2 \ .$$

It can be easily shown, just as in section 2.1, that the expected mean square error over the region R is

$$J = \frac{NK}{\sigma^2}\int_R E(\underline{\hat{g}} - \underline{\gamma})^2 d\underline{x}$$

where

$$K^{-1} = \int_R d\underline{x} \quad \text{and} \quad d\underline{x} = dx_1, dx_2, \ldots, dx_k$$

The argument of the expectation in this expression is a vector squared. While there are possibly other definitions of the square of a vector, here it is taken to be a scalar, the product of the transpose of the vector and the vector itself. Thus,

$$\begin{aligned}(\underline{\hat{g}} - \underline{\gamma})^2 &= (\underline{\hat{g}} - \underline{\gamma})'(\underline{\hat{g}} - \gamma) \\ &= \underline{\hat{g}}'\underline{\hat{g}} - \underline{\hat{g}}'\gamma - \underline{\gamma}'\underline{\hat{g}} + \underline{\gamma}'\underline{\gamma} \\ &= \underline{\hat{g}}'\underline{\hat{g}} - [E(\underline{\hat{g}})]^2 + [E(\hat{g})]^2 - 2\underline{\hat{g}}'\underline{\gamma} + \underline{\gamma}'\underline{\gamma}\ .\end{aligned}$$

Thence, since $E(\underline{\hat{g}})$ is itself a vector, we obtain

$$\begin{aligned}(\underline{\hat{g}} - \underline{\gamma})^2 = \underline{\hat{g}}'\underline{\hat{g}} - [E(\hat{g})]'[E(\hat{g})] + [E(\underline{\hat{g}})]'[E(\underline{\hat{g}})] \\ - 2\underline{\hat{g}}'\underline{\gamma} + \underline{\gamma}'\underline{\gamma}\ .\end{aligned}$$

Taking the expectation yields

$$\begin{aligned}E(\underline{\hat{g}} - \underline{\gamma})^2 &= E(\underline{\hat{g}}'\underline{\hat{g}}) - [E(\hat{g})]'[E(\hat{g})] + [E(\underline{\hat{g}})]'[E(\underline{\hat{g}})] \\ &\qquad - 2[E(\underline{\hat{g}})]'\underline{\gamma} + \underline{\gamma}'\underline{\gamma} \\ &= E(\underline{\hat{g}}^2) - [E(\underline{\hat{g}})]^2 + [E(\underline{\hat{g}}) - \underline{\gamma}]^2\end{aligned}$$

which implies

$$J = \frac{NK}{\sigma^2}\int_R \left\{E(\underline{\hat{g}}^2) - [E(\underline{\hat{g}})]^2\right\}d\underline{x} + \frac{NK}{\sigma^2}\int_R [E(\underline{\hat{g}}) - \underline{\gamma}]^2\, d\underline{x}\ . \qquad (3.1.3)$$

The first term in the right hand side of the equation is plainly the sum of the variances of the components of $\hat{\underline{g}}$ integrated and averaged over the region R, which we shall call Var $(\hat{\underline{g}})$. The second term is the sum of the squared biases of the components of $\hat{\underline{g}}$ integrated and averaged over R.

Atkinson (1) has shown that designing experiments to estimate slope in any direction is the same as designing experiments to estimate slope parallel to the factor axes. The components of $\hat{\underline{g}}$ are precisely the slopes parallel to the factor axes. Then designing an experiment to minimize the sum of the squared biases of the components of $\hat{\underline{g}}$, which sum is averaged over R, will be the same as designing an experiment to minimize the bias of the estimate of slope in any specific direction.

The bias component of J is

$$B = \frac{NK}{\sigma^2} \int_R [E(\hat{\underline{g}}) - \underline{\gamma}]^2 \, d\underline{x} \,. \tag{3.1.4}$$

We need $E(\hat{\underline{g}})$. The quantity

$$\hat{\underline{g}} = D_1 \underline{b}_1$$

implies

$$\begin{aligned} E(\hat{\underline{g}}) &= E(D_1 \underline{b}_1) \\ &= D_1 E(\underline{b}_1) \,. \end{aligned}$$

The least squares estimate of $\underline{b}_1$ is well known to be

$$\underline{b}_1 = (X_1'X_1)^{-1} X_1' \underline{y}^*$$

where X_1 is a matrix the elements of the i^{th} row of which are the values taken by the terms of $\underline{x}_1$ for the i^{th} experimental combination, and $\underline{y}^*$ is the vector of the N observations of the transformed response variable y^*. So

$$E(\underline{b}_1) = E[(X_1'X_1)^{-1}X_1'\,\underline{y}^*]$$

$$= (X_1'X_1)^{-1}X_1'E(\underline{y}^*) \ .$$

Since it can be shown that

$$E(y^*) = X_1\underline{\beta}_1 + X_2\underline{\beta}_2$$

where X_2 is a matrix similar to X for the terms of $\underline{x}_2$, we have

$$E(\underline{b}_1) = (X_1'X_1)^{-1}\ X_1'X_1\underline{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\underline{\beta}_2$$

$$= \underline{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\underline{\beta}_2$$

and finally

$$E(\hat{\underline{q}}) = D_1[\underline{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\underline{\beta}_2\,].$$

Then, substituting in (3.1.4) yields

$$B = \frac{NK}{\sigma^2}\int_R \left\{D_1[\underline{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\underline{\beta}_2] - \underline{\gamma}\right\}^2 d\underline{x} \ .$$

Since

$$\underline{\gamma} = D_1\underline{\beta}_1 + D_2\underline{\beta}_2 \ ,$$

we obtain

$$B = \frac{NK}{\sigma^2}\int_R \left\{[D_1(X_1'X_1)^{-1}X_1'X_2 - D_2]\underline{\beta}_2\right\}^2 d\underline{x}$$

$$= \frac{NK}{\sigma^2}\int_R \underline{\beta}_2'(D_1A - D_2)'(D_1A - D_2)\underline{\beta}_2 d\underline{x} ,$$

where

$$A = (X_1'X_1)^{-1}X_1'X_2 .$$

Then

$$B = \frac{NK}{\sigma^2}\int_R \underline{\beta}_2'(A'D_1'D_1A - A'D_1'D_2 - D_2'D_1A + D_2'D_2)\underline{\beta}_2 dx . \qquad (3.1.5)$$

Define

$$\Delta_{11} = K\int_R D_1'D_1 d\underline{x} \qquad (3.1.6a)$$

$$\Delta_{12} = K\int_R D_1'D_2 \, d\underline{x} \qquad (3.1.6b)$$

$$\Delta_{22} = K\int_R D_2'D_2 d\underline{x} \qquad (3.1.6c)$$

Substitute the Δ_{ij}'s into (3.1.5) to obtain

$$B = \frac{N}{\sigma^2}[\underline{\beta}_2'(A'\Delta_{11}A - A'\Delta_{12} - \Delta_{12}'A + \Delta_{22})\underline{\beta}_2] .$$

This can be rewritten to isolate A as follows:

$$B = \frac{N}{\sigma^2}\left\{\underline{\beta}_2'[\Delta_{22} - \Delta_{12}'\Delta_{11}^{-1}\Delta_{12} + (A - \Delta_{11}^{-1}\Delta_{12})'\Delta_{11}(A - \Delta_{11}^{-1}\Delta_{12})]\underline{\beta}_2\right\} . \quad (3.1.7)$$

The minimum value of B cannot be calculated from (3.1.7), but it can be said that a necessary condition for minimum B is that $(A - \Delta_{11}\Delta_{12})$ be the null matrix, because Δ_{11} is positive definite (See Appendix B).

Therefore, for minimum B

$$A - \Delta_{11}^{-1}\Delta_{12} = [0] \tag{3.1.8}$$

implies

$$A = \Delta_{11}^{-1}\Delta_{12} .$$

Since

$$A = (X_1'X_1)^{-1}X_1'X_2 ,$$

a sufficient condition for (3.1.8) is

$$X_1'X_1 = \Delta_{11} \quad \text{and} \quad X_1'X_2 = \Delta_{12} \tag{3.1.9}$$

To present this condition a more familiar form, rewrite A as

$$A = N(X_1'X_1)^{-1}N^{-1}(X_1'X_2) ,$$

which implies that $N^{-1}(X_1'X_1) = \Delta_{11}$ and $N^{-1}(X_1'X_2) = \Delta_{12}$. The quantities $N^{-1}(X_1'X_1)$ and $N^{-1}(X_1'X_2)$ clearly are moment matrices of the form of M_{11} and M_{12} for the model of (3.1.2), say,

$$M_{11}^* = N^{-1}(X_1'X_1)$$

$$M_{12}^* = N^{-1}(X_1'X_2)$$

The observed response, however, is y; not y^*. Therefore, for the condition (3.1.9) to be meaningful, it must be related to the model of (3.1.1). The moment matrices, M_{11} and M_{12}, of this model are

$$M_{11} = N^{-1}(\tilde{X}_1'\tilde{X}_1)$$

$$M_{12} = N^{-1}(\tilde{X}_1'\tilde{X}_2)$$

where $\tilde{X}_2 = X_2$ and $\tilde{X}_1 = [\underline{1} \vdots X_1]$, $\underline{1}$ being a $(N \times 1)$ vector of 1's. Then clearly, condition (3.1.9) specifies all elements of M_{11} except those in the first row and first column and all elements of M_{12} except those in the first row, i.e.,

$$M_{11} = \begin{bmatrix} \cdot & \vdots & \cdots\cdots \\ & \vdots & \Delta_{11} \end{bmatrix}, \quad M_{12} = \begin{bmatrix} \cdots\cdots \\ \Delta_{12} \end{bmatrix}.$$

3.2 Minimum Bias Estimation

In this approach, suggested initially by Karson, et al., (K, 1969), we search for necessary and sufficient conditions for the estimator $\underline{b}$ which insure that the value of B is minimal. As in the preceding section we lean heavily on Atkinson's conclusions. Starting with an equation for the value of B we derive conditions for $\underline{b}$ to minimize B. Recall (3.1.4),

$$B = \frac{NK}{\sigma^2}\int_R [E(\hat{\underline{g}}) - \underline{\gamma}]^2 d\underline{x} .$$

Note that $\hat{\underline{g}}$ is a vector each component of which is a polynomial of degree d - 2. Then $E(\hat{\underline{g}})$ is also a vector of polynomials of degree d - 2,

$$E(\hat{\underline{g}}) = D_1\underline{\alpha} . \tag{3.2.1}$$

Then $B = \frac{NK}{\sigma^2}\int_R [D_1\underline{\alpha} - \underline{\gamma}]^2 d\underline{x}$. Expand and substitute the Δ_{ij} matrices to obtain

$$B = \frac{N}{\sigma^2}\,[(\underline{\alpha} - \underline{\beta}_1)'\Delta_{11}(\underline{\alpha} - \underline{\beta}_1) - 2(\underline{\alpha} - \underline{\beta}_1)'\Delta_{12}\underline{\beta}_2 + \underline{\beta}_2'\Delta_{22}\underline{\beta}_2]. \tag{3.2.2}$$

setting $\frac{\partial B}{\partial \underline{\alpha}} = \underline{0}$ yields

$$2\Delta_{11}(\underline{\alpha} - \underline{\beta}_1) - 2\Delta_{12}\underline{\beta}_2 = \underline{0}$$

implies

$$\underline{\alpha} = \underline{\beta}_1 + \Delta_{11}^{-1}\Delta_{12}\underline{\beta}_2 \tag{3.2.3}$$

$$= R\underline{\beta}$$

where

$$R = [I : \Delta_{11}^{-1}\Delta_{12}] \quad \text{and} \quad \underline{\beta}' = [\underline{\beta}_1' : \underline{\beta}_2'] \ .$$

Substituting in (3.2.1) yields

$$E(\underline{a}) = D_1 R\underline{\beta} \ ,$$

and, since $E(\underline{a}) = D_1 E(\underline{b}_1)$, we obtain

$$E(\underline{b}_1) = R\,\underline{\beta},$$

a necessary and sufficient condition for minimum B is that the estimator $\underline{b}_1$ be such that $E(\underline{b}_1) = R\,\underline{\beta}$. The condition is subject only to the restriction that $R\,\underline{\beta}$ be estimable in $X = [X_1 : X_2]$.

The all-bias design of section 3.1 may be shown to be a special case of minimum bias estimation. Define $\underline{y}^*$ as before; then $\underline{b}_1$ may be expressed as a linear transformation of $\underline{y}^*$, say

$$\underline{b}_1 = T'\underline{y}^* \ .$$

This implies that

$$E(\underline{b}_1) = T'E(\underline{y}^*) = R\,\underline{\beta} \, .$$

Since $E(\underline{y}^*) = X\underline{\beta}$, where X and $\underline{\beta}$ are as previously defined, we have

$$T'X\underline{\beta} = R\underline{\beta}$$

which implies

$$T'X = R .$$

If $T' = (X_1'X_1)^{-1}X_1'$, then $[(X_1'X_1)^{-1}X_1'X_1 : (X_1'X_1)^{-1}X_1'X_2] = [I : \Delta_{11}^{-1}\Delta_{12}]$, or $X_1'X_1 = \Delta_{11}$ and $X_1'X_2 = \Delta_{12}$,

which are just the conditions of (3.1.9) for an all bias design.

The minimum value of B may be calculated quite simply. Substitute the value for $\underline{\alpha}$ from (3.2.3) into the expression for B in (3.2.2). This yields

$$\text{Min } B = \frac{N}{\sigma^2}\underline{\beta}_2'(\Delta_{22} - \Delta_{12}'\Delta_{11}^{-1}\Delta_{12})\underline{\beta}_2 .$$

3.3 Minimization of Variance

A class of designs has been demonstrated which minimizes B, the bias contribution to expected mean square error. From that class the particular design which yields minimum variance within the class is now to be selected. Specifically we wish to minimize Var $(\hat{\underline{g}})$ subject to $E(\underline{b}_1) = R\underline{\beta}$. Recall that $\hat{\underline{g}}$ is a vector, and that the quantity Var $(\hat{\underline{g}})$ was defined in section 3.1 as the sum of the variances of the components of $\underline{g}$.

A component of $E(\hat{\underline{g}})$, say $E(\hat{g}_i)$, may be written,

$$E(\hat{g}_i) = \underline{d}_{1i} R\underline{\beta}$$

where $\underline{d}_{1i}$ is a row vector, the i^{th} row of D_1. The minimum variance point estimate of a linear combination, $\underline{d}_{1i} R\underline{\beta}$, is

$$\hat{g}_i = \underline{d}_{1i} R(X'X)^{-1}X'\underline{y}^* \qquad (3.3.1)$$

Note that the entire X matrix, $X = [X_1 : X_2]$ is involved. Further, $X'X$ must be nonsingular (or the generalized inverse of $X'X$, say G, used in place of $(X'X)^{-1}$, and $R\underline{\beta}$ must be estimable in X. The variance portion of J in (3.1.3) is

$$V = \frac{NK}{\sigma^2}\int_R \left\{E(\hat{\underline{g}}^2) - [E(\hat{\underline{g}})]^2\right\} d\underline{x} \ ,$$

and the argument is recognized as the sum of the variances of the components of $\hat{\underline{g}}$. Then V will be minimized if each of the terms in the argument satisfies the requirements of (3.3.1). Specifically, since $\hat{\underline{g}} = D_1 T'\underline{y}^*$, then for each i,

$$\underline{d}_{1i} T'\underline{y}^* = \underline{d}_{1i} R(X'X)^{-1}X'\underline{y}^*$$

implies

$$T = R(X'X)^{-1}X' \ ,$$

and the minimum variance of $\hat{g}_i$ is

$$\text{Min Var } (\hat{g}_i) = \sigma^2 \underline{d}_{1i} R(X'X)^{-1}R'\underline{d}'_{1i}$$

which implies

$$\text{Min } V = NK \sum_{i=1}^{k} \left\{\int_R [\underline{d}_{1i} R(X'X)^{-1}R'\underline{d}'_{1i}] d\underline{x}\right\} .$$

The derivations presented in this chapter have shown a new, simple way to specify the all-bias design for estimating the slope of a response surface. Further, a necessary and sufficient condition has been demonstrated for the class of bias minimizing designs which estimate slope. The all-bias design is shown to be a special case of the class of bias minimizing designs. A second special case which minimizes variance within the class has been developed. In the next chapter we shall compare these two cases for specific design problems.

CHAPTER IV

DESIGNS FOR USE IN PRACTICE

Two examples of the case $k = 2$, $d = 2$ are examined. The minimum variance design with minimum bias estimation is compared to the all-bias design.

4.1 The Case $k = 2$, $d = 2$, $x_1 = \pm a_1$, $x_2 = \pm a_2$, a 4 Point Design with no Interaction

The true quadratic relationship is

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + e$$

$$\eta = (\eta - \beta_0) = \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + e$$

and the fitted model

$$y = b_0 + b_1 x_1 + b_2 x_2$$

$$y^* = (y - b_0) = b_1 x_1 + b_2 x_2$$

The X, $\underline{y}^*$, and $(X'X)^{-1}$ matrices are,

$$X = [X_1 \vdots X_2] = \begin{array}{c} \begin{array}{cccc} \underline{x_1} & \underline{x_2} & \underline{x_1^2} & \underline{x_2^2} \end{array} \\ \begin{bmatrix} a_1 & 0 & a_1^2 & 0 \\ -a_1 & 0 & a_2^2 & 0 \\ 0 & a_2 & 0 & a_2^2 \\ 0 & -a_2 & 0 & a_2^2 \end{bmatrix} \end{array}, \quad \underline{y}^* = \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix},$$

$$(X'X)^{-1} = \frac{1}{2}\begin{bmatrix} a_1^{-2} & 0 & 0 & 0 \\ & a_2^{-2} & 0 & 0 \\ & & a_1^{-4} & 0 \\ & & & a_2^{-4} \end{bmatrix}.$$

<u>Part A: Least Squares Estimation, All-bias Design</u>

The least squares estimator of $\underline{b}_1$ is

$$\underline{b}_1 = (X_1'X_1)^{-1}X_1'\underline{y}^* ,$$

and the bias and variance components of expected mean square error can be shown to be

$$V = \frac{NK}{\sigma^2}\int_R \left\{E(\hat{\underline{g}}^2) - [E(\hat{\underline{g}})]^2\right\} d\underline{x}$$

$$= \frac{NK}{\sigma^2}\int_R \mathrm{Var}(\hat{\underline{g}})\, d\underline{x}$$

where

$$\mathrm{Var}(\hat{\underline{g}}) = \sum_{i=1}^{2} \mathrm{Var}\,(g_i)$$

$$= \frac{1}{2}\,\sigma^2(a_1^{-2} + a_2^{-2})$$

which implies $V = \frac{NK}{\sigma^2}\left[\frac{\sigma^2(a_1^{2} + a_2^{2})}{2K}\right] = \frac{N}{2}\,(a_1^{-2} + a_2^{-2})$,

and

$$B = \frac{NK}{\sigma^2}\int_R [E(\hat{\underline{g}}) - \underline{Y}]^2\, d(\underline{x})$$

$$= \frac{N}{\sigma^2}[\beta_2'\Delta_{22}\underline{\beta}_2 - \underline{\beta}_2'\Delta_{12}'\Delta_{11}^{-1}\Delta_{12}\underline{\beta}_2 + \underline{\beta}_2\Delta'\Delta_{11}\Delta\,\underline{\beta}_2]$$

where

$$\Delta_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Delta_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Delta_{22} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Therefore, $B = \frac{N}{\sigma^2} (\beta_{11}^2 + \beta_{22}^2)$.

The bias is independent of the choice of a_1 and a_2. Consider, however, that the model is first order in the presence of a second order system. Further, $\underline{b}_1$ does not include the intercept. A similar result is well known in another situation. A 2^k factorial design used to fit a first order model in the presence of a second order system without interaction introduces bias only to the intercept term (6, 114).

Part B: Minimum Bias Estimation, Minimum Variance Design

The estimator is

$$\underline{b}_1 = T'\underline{Y}^*$$

where

$$T' = R(X'X)^{-1}X' .$$

Thus

$$\underline{b}_1 = \begin{bmatrix} a_1^{-1} & -a_1^{-1} & 0 & 0 \\ 0 & 0 & a_2^{-1} & -a_2^{-1} \end{bmatrix} \begin{bmatrix} Y_1^* \\ Y_2^* \\ Y_3^* \\ Y_4^* \end{bmatrix}$$

Further, note that

$$E(\underline{b}_1) = R\,\underline{\beta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_{11} \\ \beta_{12} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} .$$

Again, $\underline{b}_1$ is unbiased. The variance V has been shown to be

$$\text{Var}\ (\hat{g}_i) = \sigma^2\, \underline{d}_{1i} R(X'X)^{-1} R' \underline{d}_{1i}'$$

$$= \frac{\sigma^2}{2}\,(a_i^{-2})$$

implies

$$V = \frac{NK}{\sigma^2} \int_R \frac{\sigma^2}{2}\,(a_1^{-2} + a_2^{-2})\, d\,\underline{x}$$

$$= \frac{N}{2}\,(a_1^{-2} + a_2^{-2})$$

We need go no further. The two designs are the same for this case.

4.2 The Case $k = 2$, $d = 2$, with Interaction, a 6 Point Design

The true quadratic relationship is

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + e\ ,$$

and the fitted model is

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2\ .$$

Proceeding exactly as in the previous example we obtain for X, $\underline{y}^*$, and $(X'X)^{-1}$

$$
\begin{array}{c}
\begin{array}{ccccc}
\underline{x_1} & \underline{x_2} & \underline{x_1^2} & \underline{x_2^2} & \underline{x_1x_2}
\end{array} \\
X = \begin{bmatrix}
\frac{\sqrt{3}\,a}{2} & -\frac{a}{2} & \frac{3a^2}{4} & \frac{a^2}{4} & -\frac{\sqrt{3}\,a^2}{4} \\
0 & -a & 0 & a^2 & 0 \\
-\frac{\sqrt{3}\,a}{2} & -\frac{a}{2} & \frac{3a^2}{4} & \frac{a^2}{4} & \frac{\sqrt{3}\,a^2}{4} \\
-\frac{\sqrt{3}\,a}{2} & \frac{a}{2} & \frac{3a^2}{4} & \frac{a^2}{4} & -\frac{\sqrt{3}\,a^2}{4} \\
0 & a & 0 & a^2 & 0 \\
\frac{\sqrt{3}\,a}{2} & \frac{a}{2} & \frac{3a^2}{4} & \frac{a^2}{4} & \frac{\sqrt{3}\,a^2}{4}
\end{bmatrix}, \quad \underline{y}^* = \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \end{bmatrix},
\end{array}
$$

$$
(X'X)^{-1} = \begin{bmatrix}
\frac{1}{3}a^{-2} & 0 & 0 & 0 & 0 \\
0 & \frac{1}{3}a^{-2} & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}a^{-4} & -\frac{1}{6}a^{-4} & 0 \\
0 & 0 & -\frac{1}{6}a^{-4} & \frac{1}{2}a^{-4} & 0 \\
0 & 0 & 0 & 0 & \frac{4}{3}a^{-4}
\end{bmatrix}.
$$

<u>Part A. Least Squares Estimation, All-bias Design</u>

The variance and bias components of expected mean square error are

$$
V = \frac{NK}{\sigma^2} \int_R \sum_{i=1}^{2} \mathrm{Var}(g_i)\, d\underline{x}
$$

$$
= \frac{2}{3} N a^{-2}
$$

and

$$B = \frac{NK}{\sigma^2} \int_R [E(\hat{y}) - Y]^2 \, d\underline{x}$$

$$= \frac{N}{\sigma^2} (\beta_{11}^2 + \beta_{22}^2 + \beta_{12}) \; .$$

It should be noted that the above design is by no means the only six point all-bias design available to the experimenter in this situation. In fact, if he were interested only in an all-bias design, the experimenter would more likely choose the 2^k design of the preceding section with two center points added. But, as is pointed out in section 4.3, that design would yield a singular X'X matrix. The design which is presented allows development of a minimum bias estimation, minimum variance design in the following paragraphs with simple numerical manifestations.

Part B. Minimum Bias Estimation, Minimum Variance Design

As in the previous example,

$$\underline{b}_1 = T'y^*$$

$$\frac{1}{a} \begin{bmatrix} \frac{\sqrt{3}}{6} & 0 & \frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{6} & 0 & \frac{\sqrt{3}}{6} \\ -\frac{1}{6} & -\frac{1}{3} & -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \end{bmatrix}$$

and

$$E(\underline{b}_1) = R\,\underline{\beta}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_{11} \\ \beta_{22} \\ \beta_{12} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}.$$

Further, $\mathrm{Var}(g_i) = \sigma^2 \underline{d}_{1i} R(X'X)^{-1} R' \underline{d}_{1i}'$

$$= \frac{1}{3}\,\sigma^2\, a^{-2}$$

implies

$$V = \frac{2}{3}\, N\, a^{-2}$$

4.3 Interpretation and Comments

In both examples the bias component of expected mean square error was independent of design. In fact the estimates of slope at the center of the region of interest were unbiased. This result indicates that, in these cases of linear approximation to quadratic responses, the all-bias design is equivalent to a minimum bias estimation design in that bias is independent of design. Further, as we would expect from the general result shown in section 3.2, i.e., that the all-bias design is a special case of minimum bias estimation, the value of the bias component of the all bias design is minimal.

The implications are that an all-bias design, which is relatively easy to specify, may be used to minimize variance as well as bias. For example, in the case described in section 4.1 the variance component of expected mean square error is $V = \frac{N}{2}(a_1^{-2} + a_2^{-2})$. To minimize V, make a_1 and a_2 as large as possible, i.e., $a_1 = a_2 = 1$. This assignment of values to a_1 and a_2 has no effect on the bias component, $B = \frac{N}{\sigma^2}(\beta_{11}^2 + \beta_{22}^2)$, since it is independent of a_1 and a_2.

Both examples are the same except for assumptions concerning interaction. In the second example interaction is permitted, adding the term β_{12} to the vector $\underline{\beta}$. The number of terms in the whole vector $\underline{\beta}$ has significance when a minimum bias estimation design is contemplated. The estimator for $\underline{b}_1$ involves $(X'X)^{-1}$. If that inverse is to exist, then the rank of the full matrix, X, must be equal to the number of columns in X, that is, to the number of terms in the full vector $\underline{\beta}$. Thus, for the first example a design of at least four points is necessary. For the second example at least five design points are required. Further, adding a center point to the first design would not suffice to produce a design for the second example since that would only enlarge the X matrix but not change its rank. The conclusion is that the design to provide a linear approximation to a quadratic response must be carefully chosen considering the full quadratic equation. Otherwise, generating the minimum bias estimation design will require finding a generalized inverse.

An examination of the matrices, Δ_{11} and Δ_{12}, for the case of linear approximation to a quadratic response in k variables shows that Δ_{11} is always the identity matrix, $I_{k\times k}$, and Δ_{12} is always the null matrix, $[0]_{k\times p}$, where p is the number of terms in $\underline{\beta}_2$. Therefore, for this case, in general, the all-bias design is always orthogonal, the estimates of slope are unbiased, and the all-bias design is the same as the minimum bias estimation design.

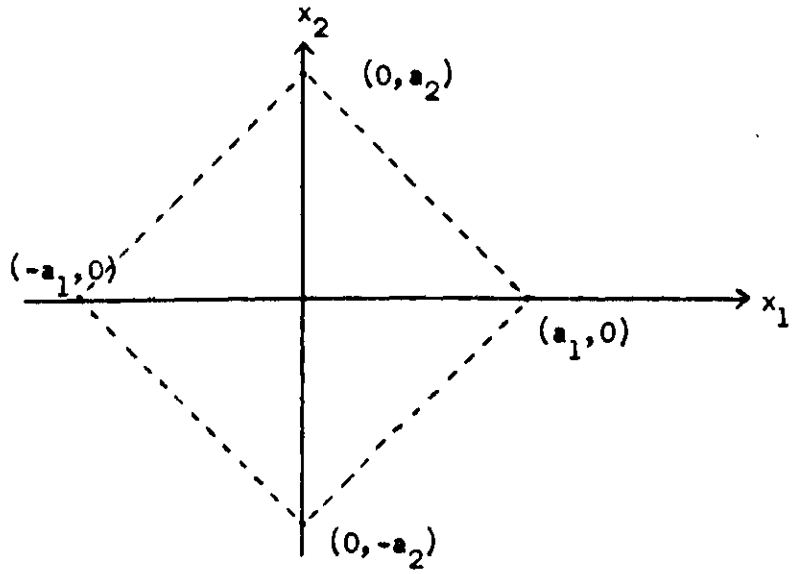

Figure 1. Design for Example 4.1.

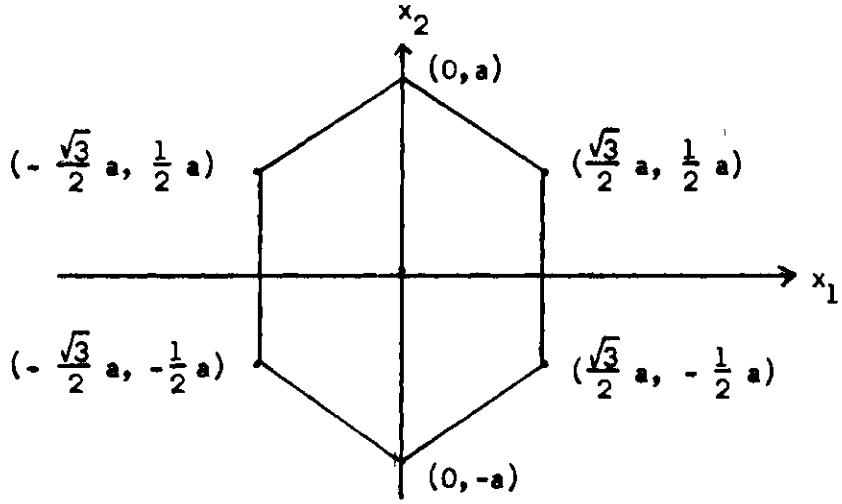

Figure 2. Design for Example 4.2.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

5.1 Results

The investigation into minimum bias estimation of the slope of a response surface has led to a new, simple way of specifying an all-bias design for making that estimation. Further, the examination of two examples of linear approximation of a quadratic response indicates that, in this important case, the easily generated all bias design is the same as the more complicated minimum bias estimation design (for a spherical region of interest). In fact, this conclusion has been generalized to all examples of the case.

All design conditions are in terms of the Δ_{ij} matrices which may be thought of as moment matrices of the region of interest for the slope of the $\underline{x}'$ vectors connected with the response. A simple method, well adapted to the use of computers, is demonstrated in Appendix B to calculate the elements of these matrices for a spherical region of interest. The method is general for this region. Only minor extensions are required to calculate the required matrices for the case of protection $\bar{\underline{g}}$ against more than one level of bias.

An approach to the problem of estimating a response which is a derivative of observed data has been outlined for the case where the variable of interest is the first derivative of the data. The pattern will require only slight modifications to extend it to consideration of the case where the desired variable is a second or higher derivative of the data.

5.2 Conclusions

1. The all-bias design to estimate the slope of a response surface has a minimal bias contribution to expected mean square error. Further, for the case of a linear approximation to a quadratic response, the bias is independent of design parameters.

2. The region moment matrices of slope developed herein provide an excellent method for characterizing designs to estimate slope whether the designs use least squares or minimum bias estimation.

3. Necessary and sufficient conditions for minimum bias estimators of the slope of a response surface have been derived.

5.3 Recommendations for Further Study

1. The results of this study suggest a regular structure of the moment matrices of design for designs of a given order which estimate derivatives while protecting against specified levels of bias. The investigation of the structure of designs to estimate the k^{th} $(k \geq 1)$ derivative while protecting against ℓ $(\ell \geq 1)$ levels of bias seems a natural extension of this study.

2. The designs developed here are specifically to estimate the slope at the center of the region of interest. A subject for further investigation is the effects of using the same designs to estimate slope at any point in the region of interest.

APPENDICES

APPENDIX A

NOTATION

Notation conventions used throughout this thesis are those thought to be standard in the experimental design literature. They are:

(1) Upper case letters denote matrices, e.g., X, M_{11}, Δ_{11};

(2) Lower case letters denote scalars or vectors as follows:

(a) Scalars - a lower case letter, e.g., x, β, etc.;

(b) Column vector - an underlined lower case letter, e.g., $\underline{x}$, $\underline{\beta}$;

(c) Row vector - transpose of a column vector denoted by a "prime," e.g., $\underline{x}'$, $\underline{\beta}'$.

(3) The above symbols are often subscripted with one or two indices. The meanings in such cases are clear by context.

Departures from these conventions are defined in text.

APPENDIX B

THE Δ_{rs} MATRIX

B.1 Introduction

The Δ_{rs} matrix is defined in Chapter III as

$$\Delta_{rs} = K \int_R D_r' D_s \, d\,\underline{x} \, , \tag{B.1.1}$$

where $K^{-1} = \int_R d\,\underline{x}$ and D_j is a matrix the i^{th} row of which is formed by partial differentiation of the row vector $\underline{x}_j'$ with respect to x_i. The variable x_i is the i^{th} component of the vector of independent variables, $\underline{x}' = (x_1, x_2, \ldots, x_k)$.

The terms of a subscripted row vector, say $\underline{x}_r'$, are the arguments of a polynomial. Each term is of the form $\prod_{u=1}^{k} x_u^{\alpha_u}$, where

$$\prod_{u=1}^{k} x_u^{\alpha_u} = (x_1^{\alpha_1})(x_2^{\alpha_2}) \ldots (x_k^{\alpha_k}) \, .$$

To uniquely designate each term of $\underline{x}_r'$, subscript the exponents of the j^{th} term. Thus

$$x_j = \prod_{u=1}^{k} x_u^{\alpha_{uj}} \tag{B.1.2}$$

In (B.1.2) all the exponents are non-negative integers. The right hand side of (B.1.2) may be partially differentiated with respect to x_i to obtain the ij^{th} term of D_r, say d_{ij}^r. The superscript r denotes only that d^r is an element of D_r. So,

$$d^r_{ij} = \alpha_{ij} x_{ij}^{\alpha_{ij}-1} \prod_{\substack{u=1 \\ u \neq i}}^{k} x_u^{\alpha_{uj}}$$

Consider a second row vector $\underline{x}$'s like $\underline{\psi}_r'$, r may or may not equal s. The components of $\underline{x}'$ remain the same. The D_s matrix associated with $\underline{x}_s'$ will have terms exactly like those of D_r. The two are differentiated in the following development by an asterisk, (*), on the exponents of elements of D_s. Thus

$$d^s_{ij} = \alpha^*_{ij} x_i^{\alpha^*_{ij}-1} \prod_{\substack{u=1 \\ u \neq i}}^{k} x_u^{\alpha^*_{uj}}$$

The matrix product $D_r' D_s$ of (B.1.1) is now easily formed.

The mnth element of $D_r' D_s$, say d^{rs}_{mn}, is the sum of the products of the corresponding terms in the mth column of D_r and the nth column of D_s, i.e.,

$$d^{rs}_{mn} = \sum_{i=1}^{k} \left[\alpha_{im} \alpha^*_{in} x_i^{(\alpha_{im} + \alpha^*_{in} - 2)} \prod_{\substack{u=1 \\ u \neq i}}^{k} x_u^{(\alpha_{um} + \alpha^*_{un})} \right]$$

Since an integral of a matrix is the matrix of the integrals of the elements of that matrix, we write the mnth element of Δ_{rs}, say δ^{rs}_{mn}, as

$$\begin{aligned} \delta^{rs}_{mn} &= K \int_R d^{rs}_{mn} \, d\,\underline{x} \\ &= K \int_R \left\{ \sum_{i=1}^{k} [\alpha_{im} \alpha^*_{in} x_i^{(\alpha_{im}+\alpha^*_{in}-2)}][\prod_{\substack{u=1 \\ u \neq i}}^{k} x_u^{(\alpha_{um}+\alpha^*_{un})}] \right\} d\,\underline{x} \end{aligned}$$

$$= K \sum_{i=1}^{k} \left\{ \alpha_{im}\alpha^*_{in} \int_R [x_i^{(\alpha_{im}+\alpha^*_{in}-2)}][\prod_{\substack{u=1\\u\neq i}}^{k} x_u^{(\alpha_{um}+\alpha^*_{un})}]\, d\underline{x} \right\} \qquad (B.1.3)$$

In one special, and important, case the δ^{rs}_{mn} of (B.1.3) is quite easily evaluated.

B.2 Spherical Region of Interest

Let R in (B.1.3) be a unit hyper-sphere centered at the origin of the x's, i.e., R is such that $\sum_{i=1}^{k} x_i^2 \leq 1$. Then the integral of (B.1.3) is of the type

$$\int_{-1}^{1}\int_{-1}^{1} \cdots \int_{-1}^{1} (x_1^{\rho_1})(x_2^{\rho_2})\ldots(x_k^{\rho_k})dx_1,dx_2,\ldots,dx_k$$

where $\rho_1 = \alpha_{1m}+\alpha^*_{1n}$, $\rho_2 = \alpha_{2m}+\alpha^*_{2n}$, etc.

It is well-known that the value of an integral, say L, of this type is

$$L = \frac{\Gamma(\frac{\rho_1+1}{2})\Gamma(\frac{\rho_2+1}{2})\ldots\Gamma(\frac{\rho_k+1}{2})}{\Gamma\left[\frac{\sum_{u=1}^{k}(\rho_u+1)}{2}+1\right]} \qquad (B.2.1a)$$

if all ρ_u are even, and

$$L = 0 \qquad (B.2.1b)$$

if any ρ_u is odd.

The integral, K^{-1}, may be written in the same form with all $\rho_u = 0$. So,

$$K^{-1} = \frac{[\Gamma(\frac{1}{2})]^k}{\Gamma(\frac{k}{2}+1)}$$

since

$$\Gamma(p+1) = p\Gamma(p) \quad \text{and} \quad \Gamma(\tfrac{1}{2}) = \pi^{1/2}$$

we have

$$K^{-1} = \frac{(\pi)^{k/2}}{\Gamma(\frac{k+2}{2})} \tag{B.2.2}$$

Substituting for ρ_u in (B.2.1a and B) we get for the value of an integral, $L(m,n)_i$,

$$L(m,n)_i = \frac{\Gamma\left(\frac{\alpha_{im}+\alpha_{in}^{*}-2+1}{2}\right)\prod\limits_{\substack{u=1\\u\neq i}}^{k-}\Gamma\left(\frac{\alpha_{um}+\alpha_{un}^{*}+1}{2}\right)}{\Gamma\left[\frac{\alpha_{im}+\alpha_{in}^{*}-2+1+\sum\limits_{\substack{u=1\\u\neq i}}^{k}(\alpha_{um}+\alpha_{un}^{*}+1)}{2}+1\right]} \tag{B.1.3a}$$

if all sums $\alpha_{um} + \alpha_{un}^{*}$, $u = (1,2,\ldots,k)$, are even, and

$$L(m,n)_i = 0 \tag{B.1.3b}$$

if any sum $\alpha_{um} + \alpha_{un}^{*}$ is odd.

Note that $i \in \{u\}$, because, if $\alpha_{im}+\alpha_{in}^{*}$ is even, so is $\alpha_{im}+\alpha_{in}^{*}-2$. Substituting from (B.2.3a,b) into (B.1.3) yields

$$\delta_{mn}^{rs} = K\sum_{i=1}^{k}[\alpha_{im}\alpha_{in}^{*}L(m,n)_i] \tag{B.2.4}$$

Before doing an example, consider (B.2.3a) again. In the numerator there appears the expression, $\Gamma(\frac{\alpha_{im}+\alpha^*_{in}-2+1}{2})$. Conceivably the argument of that gamma function could be negative or zero since a_{ij} may equal zero or 1. Such an argument would violate the definition

$$\Gamma(p) = \int_0^\infty v^{p-1}e^{-v}\,dv\,; \quad p > 0$$

Resolution is simple. There are three possibilities:

(1) $\alpha_{im} = \alpha^*_{in} = 0$;

(2) α_{im} or $\alpha^*_{in} = 1$ and the other is zero;

(3) $\alpha_{im} + \alpha^*_{in} \geq 2$.

The third cases results in a valid gamme function in (B.1.3a). The other two cases do not arise because to obtain the i^{th} term of the sum that is δ^{rs}_{mn} the integral, $L(m,n)_i$, must be multiplied by $\alpha_{im}\alpha^*_{in}$ (see (B.2.4)).

B.3 Calculation of Δ_{11} for R a Unit Hypersphere

Suppose η as defined in Chapter III is a function of $\underline{x}$ where $k = 2$ and the polynomial of degree $d_3 = 2$, i.e.,

$$\eta = \beta_0 + \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2 + e$$

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2 + \beta_{12}x_1 x_2 + e$$

and

$$\hat{y} = b_0 + \underline{x}_1'\underline{b}_1$$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2$$

then

$$\underline{x}_1' = (x_1, x_2); \ \underline{x}_2' = (x_1^2, x_2^2, x_1 x_2)$$

or, in the form of (B.1.2),

$$\underline{x}_1' = (x_1^1 x_2^0, \ x_1^0 x_2^1) \qquad \text{(B.3.1)}$$

and

$$x_2' = (x_1^2 x_2^0, \ x_1^0 x_2^2, \ x_1^1 x_2^1)$$

Let D_1 and D_2 be the "D" matrices defined in section B.1 corresponding to $\underline{x}_1'$ and $\underline{x}_1'$ respectively. No further information is required to calculate any of the Δ_{ij} matrices associated with this model. An algorithm for the calculation of Δ_{11} is shown in tabular form at Table B.1. From column 10 of the table we construct,

$$\Delta_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{(B.3.2)}$$

Table 1. Calculation of Δ_{11} for $k = 2$, $d = 2$

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
m	n	i	α_{im}	α^*_{in}	$(\alpha_{im})\cdot(\alpha^*_{in})$	$\alpha_{im}+\alpha^*_{in}$	$u \neq i$	α_{um}	α^*_{un}	$\alpha_{um}+\alpha^*_{un}$	$L(m,n)_i$	6 x 12	$\sum_i 13$	K x 13 $= \delta_{mn}$
1	1	1	1	1	1	2	2	0	0	0	$\frac{[\Gamma(\frac{1}{2})]^2}{\Gamma(2)} =$	K^{-1}	K^{-1}	1
1	1	2	0	0	0							0		
1	2	1	1	0	0							0	0	0
1	2	2	0	1	0							0		
2	2	1	0	0	0							0	K^{-1}	1
2	2	2	1	1	1	2	1	0	0	0	K^{-1}	K^{-1}		

BIBLIOGRAPHY

Literature Cited

1. Atkinson, A. C. 1970. "The Design of Experiments to Estimate the Slope of a Response Surface," Biometrika, 57, 319-328.

2. Box, G. E. P. and Draper, N. R. 1959. "A Basis for the Selection of a Response Surface Design," Journal of the American Statistical Association, 54, 622-654.

3. Karson, M. J., Hader, R. J., and Manson, A. R. 1969. "Minimum Bias Estimation and Experimental Design for Response Surfaces," Technometrics, 11, 461-475.

4. Luers, J. K. and Wenning, R. H. 1971. "Polynomial Smoothing - Linear vs Cubic," Technometrics, 13, 589-600.

5. Munsch, R. C. 1971. Minimum Bias Designs for Response Surfaces.

6. Myers, R. H. 1971. Response Surface Methodology, Allen and Bacon, Rockleigh, N. J.

Other References

Box, G. E. P. and Wilson, K. B. 1951. "On the Experimental Attainment of Optimum Conditions," Journal of the Royal Statistical Society (Series B), 13, 1-41.

Draper, N. R. and Lawrence, W. E. 1965. "Designs which Minimize Model Inadequacies: Cuboidal Regions of Interest," Biometrika, 52, 111-118.

Graybill, F. A. 1961. An Introduction to Linear Statistical Models, Vol. 1, McGraw-Hill, New York.

Hotelling, H. 1941. "Experimental Determination of the Maximum of a Function," Journal of the American Statistical Association, 65, 1565-1572.

Kussmaul, K. 1969. "Protection Against Assuming the Wrong Degree in Polynomial Regression," Technometrics, 11, 677-682.