BIAS AND VARIANCE CRITERIA FOR ESTIMATORS AND DESIGNS FOR FITTING POLYNOMIAL RESPONSES

by

M. J. Karson, R. J. Hader and A. R. Manson

This research was supported by a Shell Foundation Fellowship.

Institute of Statistics
Mimeograph Series No. 510
Raleigh, N. C. 1967

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

1. INTRODUCTION

It is generally recognized that response surface analysis is a statistical technique which uses the methodology of the design and analysis of experiments, to detect and describe unknown functional relationships between a dependent response variable and several independent controllable variables. Commonly the relationship is expressed as a surface built by response contours in the space of the controllable factors, properly scaled to dimensionless variables. The manner in which the underlying response surface is explored relies on the experimental levels of the controlled variables and the estimation procedure for fitting the response surface. These two aspects are obviously intimately related.

At some stage of a response surface investigation it often is necessary and convenient to assume a true underlying functional relationship,

$$(1) \qquad \eta = f(\xi_1, \xi_2, \ldots, \xi_p; \alpha_1, \alpha_2, \ldots, \alpha_r)$$

exists between the response η and p independent variables ξ_i, over some large region of operability in the space of the controllable factors, where the α_j represent unknown parameters. Over some smaller region of interest lying in the region of operability, N observations of the response η are taken at a set of N corresponding levels of each ξ_i; these levels constituting an experimental design. The data are then used to fit an estimating function $\hat{y}$ to η over the region of interest.

This thesis is concerned with

(1) criteria for choosing the estimator and

(2) criteria for choosing an experimental design.

Clearly the application of criteria to these problems must be consistent with the underlying problem. These criteria are used to derive estimating functions, experimental design moment requirements and some experimental designs.

More specifically, suppose the original variables ξ_i are standardized in a natural way to transformed variables x_i and denote by R the scaled region of interest. In most applications R is the p-dimensional cube defined by $-1 \leq x_i \leq 1$. Suppose that over the wider region of operability the true functional response η is a polynomial of degree d+k-1 and the estimating function $\hat{y}$, desired to fit η over the region R, is a polynomial of degree d-1. This is the practical or experimental way that the unknown form of the true model is taken into consideration. To account for the model inadequacy in fitting and from a result discussed by Box and Draper (1959, 1963), the primary criterion adopted for choosing an estimator is minimizing the squared bias, arising from the failure of $\hat{y}$ to represent the true model exactly, integrated over the region of interest. Sufficient conditions on the form of the estimator are derived and subject to the estimator satisfying these minimum bias requirements, estimators are found which satisfy the secondary criterion of minimizing the variance of $\hat{y}$ integrated over R. The bias criterion is

$$(1.2) \qquad B = c_1 \int_R \{E[\hat{y}(\underline{x})] - \eta(\underline{x})\}^2 d\underline{x}$$

and the variance criterion is

$$(1.3) \qquad V = c_2 \int_R \text{Var } \hat{y}(\underline{x})\, d\underline{x} \ ,$$

where c_1 and c_2 are constants.

Given that the estimator satisfying these criteria is to be used for fitting, conditions on the design moments are determined which make the experimental design satisfy the following protection criterion. Protection is against specified higher order effects being present in the true model. These effects are guarded against by selecting experimental designs which in turn yield estimators which are constant for specified values of k. This means, for example, that if the experimenter is fitting a linear estimator, the estimator depends on the degree of the true model assumed, and protection against different values of k, say against k=1 and k=2 for quadratic and cubic effects in the true model, is obtained by choosing designs which make the estimators identical for these values of k.

It should be noted that the plan followed for finding estimators and experimental designs is different in certain aspects from the more or less traditional approaches to these problems. In most cases variance type criteria have been preferred for choosing estimators (for example the least squares procedure assumes the fitted function is identical to the true model). However, when the underlying problem supposes a certain inadequacy in the fitted response model, satisfying variance criteria should not necessarily be the primary objective in choosing estimators. Both variance type criteria (e.g. rotatability) and bias type criteria (e.g. min max $| E[\hat{y}(\underline{x})] - \eta(\underline{x})|$) have been applied to the design problem, but essentially independently of the motivations for the criteria used in finding estimators.

Section 2 constitutes a survey of the pertinent literature and includes additional remarks in the above vein. Section 3 derives the estimator for the single controllable variable case and Section 4 derives design moment conditions for fitting linear functions and guarding against quartics and less, fitting quadratic functions and guarding against quintics and less, and fitting cubic and protecting against quintics and less. Section 5 generates some experimental designs for these situations. Chapter 6 offers an extension of the structure and principles to higher dimensions, generalizing the estimator and giving some design conditions and designs.

2. REVIEW OF LITERATURE

This survey of pertinent literature is divided into three parts. The first reviews general response surface methodology and is mostly concerned with the underlying motivation for many response surface investigations, i.e. finding optimum operating conditions. The second part covers different criteria proposed both for fitting response surfaces and selecting experimental designs. These works have arisen in the traditional regression analysis context as well as in the response surface framework. The last section specifically reviews those articles concerned with the bias criterion given in (1.2) and discusses the motivation for the present work.

2.1 Response Surface Methodology

Hotelling (1941) considered the problem of arranging an experiment for determining the value of x which maximizes a function

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \dots$$

whose coefficients $\beta_0, \beta_1, \beta_2, \dots$ are unknown. An early exploratory study is initially performed to provide sufficient preliminary approximate knowledge of the neighborhood of the maximum. Advance knowledge of certain higher order parameters are also required and may be obtained from small experiments in the neighborhood of the maximum. When a fairly good knowledge of the response function is available the function is fitted with a quadratic equation

$$\hat{y} = b_0 + b_1 x + b_2 x^2$$

by least squares and the estimate, x_0, of the maximizing level is

$-b_1/2b_2$. The errors arising from the inadequacy of the quadratic approximation and from experimental error are examined and are used as criteria to determine an experimental design. Letting B_3 be the coefficient of β_3 and B_4 be the coefficient of β_4 in $E(b_1-\beta_1)$, i.e. $E(b_1-\beta_1) = B_3\beta_3 + B_4\beta_4 + \ldots$, and assuming the variance of b_1 is fixed, Hotelling shows how to allocate the experimental points to make B_3 vanish and to minimize B_4. Three distinct levels of x are required and they must satisfy

$$x_1x_2 + x_1x_3 + x_2x_3 = 0 .$$

He also briefly considered the two variable problem.

It is important to note that this early work was concerned directly with model inadequacy in experimental design.

Friedman and Savage (1947) also restrict themselves to problems where the primary object is to find that combination of the controllable factors at which the maximum occurs. For this purpose they reject the factorial design because 1) it devotes observations to exploring regions that may be of no interest because they are far from the maximum, 2) it either explores a small region or explores a large region very superficially and 3) it does not take advantage of the fact that some of the variables are continuous. The authors suggest a sequential, one factor-at-a-time scheme, as follows: Choose in advance the best starting combination of factors, using the best preliminary estimate of the optimum combination. Somehow order the independent variables and perform a series of experiments varying the first factor holding the others constant. Estimate the level of the first factor yielding maximum response. Holding this factor fixed at its maximizing value, find the optimal level of the second factor in the same manner. Continue for all variables, making a "round" of experiments. The process is repeated until convergence to a maximum is attained.

Box and Wilson (1951) laid the foundation for present response surface analysis using a strategy involving varying more than one factor at a time in order to find optimum operating conditions. In the entire k-dimensional factor space, there is a region, R, called the experimental region where experimentation is to be performed to find that combination of factors yielding maximum response. The general method is to begin with a full or fractional 2^k factorial design. These designs provide estimates of first order effects and of certain interaction effects. If the first order effects are large compared to second order interactions, then further experiments are performed along the path of steepest ascent determined from the first order effects. When first order effects are small it is inferred that a near-stationary region has been reached and a full second order model is fitted using a composite design, formed by adding points which allow for determination of quadratic effects to the factorial.

Box and Wilson were also concerned with the bias due to fitting an inadequate polynomial and introduced the alias matrix to reflect the bias in the least squares estimates of the unknown parameters.

A further discussion of the general considerations in response surface methodology is given by Box (1954) who included in the statement of the problem the experimenter's concern with elucidating certain aspects of the underlying functional relationship, and not simply finding maximum or minimum producing operating conditions. This paper represents a cursory review of the previous article, discussing factor dependence, experimental designs, ridge systems (see Draper, 1963) and further exposition of the method of steepest ascent.

It should also be mentioned that Anderson (1953) reviews the first three articles of this section.

2.2 Criteria Useful in Response Surface Fitting

Since regression analysis is an important part of response surface methodology, criteria have been applied to regression problems for both designs and estimation of parameters. Prior to 1958 variance type criteria were employed, while after 1958 bias type criteria were introduced. This section is primarily concerned with criteria that may be useful in exploring a response surface.

Elfving (1952) considered the general linear model with two independent variables x_1 and x_2, r possible combinations of the two factors, a total of n observations with np_i replicates of the i^{th} experiment, $i = 1,2,\ldots,r$. The problem is to find the optimum weights $p_1,p_2,\ldots,p_r$. The sense of optimality depends on the estimation problem. Let the two unknown parameters be β_1 and β_2 in the model $y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \epsilon_i$, and let $\bar{y}_i$ be the means of the np_i responses. The errors are independent with mean zero and variance σ^2. The problem of estimating $\theta = a_1\beta_1 + a_2\beta_2$, by $\hat{\theta} = a_1\hat{\beta}_1 + a_2\hat{\beta}_2$, where a_1 and a_2 are known and $\hat{\beta}_1$ and $\hat{\beta}_2$ are the least squares estimates of β_1 and β_2 is considered. Since $\hat{\beta}_1$ and $\hat{\beta}_2$ minimize the weighted squared sum $\Sigma\, p_i(\bar{y}_i - x_{i1}\beta_1 - x_{i2}\beta_2)^2$, they depend on $p_1,p_2,\ldots,p_r$. Elfving introduces the criterion of choosing the p_i which minimize the variance of the estimate. For estimating θ he found that experiments were required at two and only two points. For estimating β_1 and β_2 experiments at three points, and in general for estimating s β's, at most $s(s+1)/2$ experiments are required for optimality. His approach was via a purely geometric argument.

Chernoff (1953) generalized the work of Elfving and introduced the criterion of minimizing the variance of the asymptotic distribution of the maximum likelihood estimate of the parameters.

De la Garza (1954) studied spacing of experiments for a single factor when the relation is a polynomial of degree m. Using the criterion of minimizing the maximum variance of least squares estimators, he showed that no more than m+1 distinct values of the factor were required over a certain interval.

Box and Hunter (1957) introduced the important concept of rotatable designs. A design in which the variance of an estimated response at a given point in the factor space is dependent only on the distance of the point from the origin of the design is called a rotatable design.

With a standardized set of k factors, the unknown relationship requiring exploration is assumed to be represented by a polynomial of degree d in the k factors. The estimated response $\hat{y}$ is fitted by least squares and the criterion of rotatability requires that $\text{var}\ (\hat{y}(\underline{x})) = \text{var}\ (\hat{y}(\underline{z}))$, where $\underline{x}$ and $\underline{z}$ are two points in the factor space at equal distances from the origin. By using design moment generating functions it is shown that the moments of a rotatable design are the moments up to order 2d of a spherical distribution. Let

$$[1^{\alpha_1}, 2^{\alpha_2}, \ldots, k^{\alpha_R}] = N^{-1} \sum_n x_{1n}^{\alpha_1} x_{2n}^{\alpha_2} \ldots x_{kn}^{\alpha_k} .$$

In particular the moment requirements for rotatability are

(a) all odd order moments are zero (i.e. if any α_i is odd)

(b) $[1^{\alpha_1}, 2^{\alpha_2}, \ldots, d^{\alpha_k}] = \dfrac{\lambda_\alpha \prod_{i=1}^{k} \alpha_i!}{2^{\alpha/2} \prod_{i=1}^{k} (\frac{1}{2}\alpha_2)!}$ for all α_i even,

where λ_α depends on d, the degree of the approximating polynomial, and generating function coefficients.

Hoel (1958), (1961) used the criterion of minimizing the generalized variance of least squares estimates in the general linear model with uncorrelated errors to choose experimental designs. He also applied this criterion to models with specific non-diagonal covariance matrices.

Folks (1958) considered the design problem for approximating true polynomial responses with a lower degree polynomial, using least squares estimates over an experimental region R. He considered eight different optimality criteria.

Let $\hat{y}(\underline{x})$ be the estimated response at the point $\underline{x}$. Let $\hat{\underline{\beta}}$ be the vector of least squares estimates for the coefficients, and let $B(\underline{x})$ be the bias at $\underline{x}$ due to the inadequacy of $E(\hat{y})$ to represent the true model. Also let D denote the design matrix. Then the criteris studied were

1) $\min_{D} \max_{\underline{x}} \text{Var } \hat{y}(\underline{x})$

2) $\min_{D} \int_{R} \text{Var } \hat{y}(\underline{x}) \, d\underline{x}$

3) $\min_{D} | \text{Var } \hat{\underline{\beta}} |$

4) $\min_{D} \max_{\underline{x}} B(\underline{x})$

5) $\min_{D} \int_{R} B(\underline{x}) d\underline{x}$

6) $\min_{D} \int_{R} B^2(\underline{x}) d\underline{x}$

7) $\min_{D} \int_{R} \text{MSE}(\hat{y}) d\underline{x}$

8) $\min_{D} \max_{\underline{x}} \text{MSE } (\hat{y})$.

Criteria (1)-(3) are variance type criteria, (4)-(6) are bias type criteria and the last two involve both types.

Folks showed the invariance of optimum designs using these criteria under simple linear transformations, enabling restricting attention to cubical regions $-1 \leq x_i \leq 1$ and spherical regions of radius unity and still generalizing the design to any rectangular or ellipsoidal regions. He developed design conditions for fitting linear approximations to quadratic models in p variables for the variance and bias criteria over cubical regions as well as for other specific situations and criteria.

David and Arens (1959) used mean square error type criteria to find an optimal spacing of a single controllable factor when fitting a straight line to a true quadratic response. Using n/2 observations at each of 2 locations x_1 and x_2, with corresponding mean responses $\bar{y}_1$ and $\bar{y}_2$, the fitted line is

$$\hat{y}(x) = \hat{c}_0 + \hat{c}_1(x-\bar{x}) ,$$

where $\hat{c}_0 = \bar{y}$, $\hat{c}_1 = (\bar{y}_2-\bar{y}_1)/(x_2-x_1)$. The observed response is assumed to be $y(x) = f(x) + \epsilon$, where $f(x)$ is quadratic and ϵ has mean zero and variance σ^2. As an extension of the non-random problem of fitting polynomials by Legendre and Tchebysheff spacing, the authors introduce the criteria

1) minimum expected squared error integrated over $[-1,1]$,

2) minimum maximum expected squared error.

The first leads to the symmetric spacing $x_1 = -x_2$ with x_2 being a root of $r^4(3r^2-1) = (2\sigma^2/9nc_2^{\,2})$, where c_2 is the coefficient of the Legendre polynomial $p_2(x)$ in the expression $f(x) = c_0 + c_1p_1(x) + c_2p_2(x)$. The second criterion leads to the symmetric spacing $x_1 = -x_2$ with $x_2 = \min\{\frac{1}{2}[1 + (1 + 32\sigma^2/9nc_2^{\,2})^{\frac{1}{2}}]^{\frac{1}{2}},1\}$. It is also shown that when fitting a least squares line there is no improvement by spacing at more than 2 distinct levels.

Kiefer (1959) offered various criteria for optimum experimental designs under the setting of inference about s given linearly independent parametric functions $\psi_j = \Sigma\, c_{ij}\theta_i$, in the parameters θ_i. Let V be the covariance matrix of the best linear unbiased estimates of the ψ_j for a given design. The problem is considered to be that of testing the hypothesis that the ψ_j are all zero, under the normality assumption, against alternatives for which $\Sigma\, \psi_j^2 = c\sigma^2$. Let ϕ represent a test and let $\bar{\beta}_\phi(c)$ be the infimum of the corresponding power function. Different kinds of optimum designs and their relationships are discussed depending on the following criteria. Designs are

1) M-optimum if $\sup_\phi \bar{\beta}_\phi(c)$ is a maximum,
2) L-optimum if they are locally M-optimum,
3) D-optimum if $|V|$ is a minimum,
4) E-optimum if the maximum characteristic root of V is a minimum,
5) A-optimum if trace (V) is a minimum,
6) G-optimum (for regression problems) if the supremum of the variance of the difference between the estimated regression function, assumed correct, and the true model, when fitted by least squares, is a minimum.

This work was continued by Kiefer (1961) and applied to rotatable designs and systematic designs (which arise when a time variable affects the experimental procedure).

Kiefer and Wolfowitz (1959) consider optimum designs in regression analysis for the regression functions $\eta(x) = \sum_{i=1}^{k} a_i f_i(x)$. They allow for three inference problems in choosing designs. These are (1) estimation of a_k assuming the other a_i are known, (2) estimation of $s < k$ of the a_i and (3) estimation of the whole regression function. For the first situation designs are sought which minimize the variance of the best linear unbiased estimator of a_k. These designs depend on the best Tchebycheff approximation to $f_k(x)$. For the second inference situation justification is given for using as a design criterion D-optimality as previously defined, i.e. minimizing the generalized variance of the best linear unbiased estimates of the s a_i. The criterion used for the third situation is to choose designs which minimize the maximum expected squared difference between $\eta(x)$ and the best linear unbiased estimator of $\eta(x)$. For polynomial regression on the interval $-1 \leq x \leq 1$, the best design places 1/k of the observations at -1,+1 and at the zeros of the derivative of the Legendre polynomials. They also show that if a design putting mass 1/k on each of k points is optimum for set-up (3) then the design is D-optimum for s=k.

The final criterion mentioned in this section was discussed by Hoel and Levine (1964). They employed the criterion of minimizing the variance of the predicted value of a polynomial response at a specified point beyond the interval of observations, defined by $-1 \leq x \leq 1$.

2.3 Minimum Bias Designs

Box and Draper (1959) motivated and set the framework for the present study. It is required to fit a response function $\eta = g(\underline{x})$ by a polynomial $\hat{y}(\underline{x})$ of degree d_1 over some standardized region of interest R. In R it is assumed that η can be expressed exactly as a polynomial of degree $d_2 > d_1$, but that the lower degree polynomial $\hat{y}$ will be adequate for practical purposes. They assumed that $\hat{y}(\underline{x})$ is fitted by least squares and recognized both variance error (due to sampling error) and bias error (due to the inadequacy of the polynomial $\hat{y}$ to exactly fit η). As a result of their primary objective of requiring the design to allow $\hat{y}$ to represent η as well as possible within R, they considered the problem of choosing designs which minimized J, the expected mean square error averaged over R. Thus,

$$J = \frac{N\Omega}{\sigma^2} \int_R E\{\hat{y}(\underline{x}) - \eta(\underline{x})\}^2 d\underline{x}$$

where $\Omega^{-1} = \int_R d\underline{x}$ and σ^2 is the experimental error variance. This can be partitioned into two meaningful components, the integrated variance and the integrated bias yielding

$$J = V + B,$$

where

$$V = \frac{N\Omega}{\sigma^2} \int_R \text{Var}\ \hat{y}(\underline{x}) d\underline{x}$$

$$S = \frac{N\Omega}{\sigma^2} \int_R \{E\hat{y}(\underline{x}) - \eta(\underline{x})\}^2 d\underline{x}\ .$$

They showed, even though designs minimizing J depended on certain unknown parameters, that the bias contribution generally far outweighed the variance contribution. Hence it is useful to consider designs for the "all-bias" case.

Designs which minimize B can be found by equating the moments of the design up to order d_1+d_2 to the corresponding moments of a uniform distribution over R. Thus, if $\hat{y}(\underline{x}) = \underline{x}_1'\underline{b}_1$ and $\eta(\underline{x}) = \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2$ where $\underline{b}_1$ is the vector of least squares estimates, and if X_1 is the design matrix and X_2 is the matrix representing the bias type terms, then the necessary and sufficient conditions for a design to minimize B are

$A = \mu_1^{-1}\mu_2$ where $A = (X_1'X_1)^{-1}X_1'X_2$ and $\mu_1 = \Omega \int_R \underline{x}_1\underline{x}_1' d\underline{x}$ and $\mu_2 = \Omega \int_R \underline{x}_1\underline{x}_2' d\underline{x}$. Particular examples of designs are given for fitting linear functions to true quadratic surfaces within a spherical region.

In a second paper Box and Draper (1963) extend their earlier work by applying their development to fitting quadratic polynomials when the true model is cubic. They considered second order rotatable designs and found, for example, that with two factors and R being the unit sphere, that the design must satisfy $\Sigma x_1^2 = \Sigma x_2^2 = N(.2652) = Nc$ and $\lambda = 3\lambda_4/c^2$ and $\lambda_4 = \Sigma x_1^4/3N = \Sigma x_2^4/3N = \Sigma x_1^2x_2^2/N$.

Lawrence (1964), applied the Box and Draper all-bias results to different kinds of regions. He let R be a triangle in two dimensions, a tetrahedron in three dimensions and a k-dimensional parallelepiped.

A further contribution was added to the Box and Draper all-bias development by Manson (1966). He considered the problem of fitting $y = \Sigma b_j e^{jz}$ to a true exponential response $\eta = \alpha + \beta e^{\gamma z}$. When transformed to the problem of fitting $\Sigma b_j x^j$ to $\alpha + \beta x^\gamma$ the Box and Draper framework is apparent. Manson applied some aspects of numerical analysis pertinent to polynomial equations to actually deriving minimum B designs.

At this point it is necessary to consider an experimental outlook which ties together the design problem and the estimation problem. Clearly without an estimation goal (or more generally, an inference goal) there is no meaningful design criterion. This viewpoint is reflected in Kiefer (1959) and more explicitly in Elfving (1952) where it is recognized that the sense of design optimality depends on the estimation problem which the design is to help solve. It follows that it is also necessary to consider the estimation criteria along with the design criteria. These criteria must relate to the entire underlying problem. Optimization with respect to both design and estimation jointly can be achieved by first finding the optimum estimation procedure for any given design and then choosing the design(s) which, together with the optimum estimation procedure, gives the "best" results for some appropriate over-all criterion. It is, of course, necessary that criteria of optimality used in each of the two stages be consistent.

From the foregoing more general viewpoint the Box and Draper work leading to designs which minimize bias alone (or those which minimize some combination of variance and bias) is really an optimization in a somewhat restricted sense, the restriction being that only least squares estimation is allowed. Traditionally the justification of least squares estimation is that minimum variance unbiased estimators are obtained if the model fitted is correct. In the Box and Draper work, however, it is assumed that the model fitted may well not be correct. The least squares estimates are then biased, and it is not at all clear that the minimum variance property is at all relevant.

The Box-Draper results indicate rather conclusively that the bias contribution to mean square error exerts much more influence on the choice of design than does the variance contribution. Since least squares estimation essentially optimizes with respect to variance, it appears that its use may be an unnecessary restriction. In the work which follows this restriction is dropped. An estimation method is used which, for a fixed design, minimizes the integrated bias due to specified terms not included in the fitted equation, and subject to this, also minimizes the integrated variance of the estimator. It is found that using this method of estimation the same minimum bias is achieved for any design (subject to easily satisfied conditions). For designs satisfying certain special conditions the estimation method reduces to the familiar least squares method, but it is emphasized that this is only a special case of the more general method.

Since the primary criterion of minimizing bias and a secondary criterion of minimizing variance subject to the minimum bias are satisfied essentially by choice of the estimator alone, an additional criterion is introduced to select among designs.

3. DERIVATION OF THE ESTIMATOR

In this chapter the bias and variance criteria are developed and applied to deriving the estimating polynomial. Sufficient conditions are first found which minimize the bias criterion. This is followed by determining the exact representation of the estimator arising from minimizing the variance criterion subject to these conditions.

3.1 Set-Up for the One-Dimensional Case

Over an entire region of operability, the true model η at the point x is assumed to be a polynomial of degree d+k-1 given by

(3.1) $$\eta(x) = \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2$$

where

(3.2) $$\underline{x}_1' = (1, x, x^2, \ldots, x^{d-1})$$

(3.3) $$\underline{x}_2' = (x^d, x^{d+1}, \ldots, x^{d+(k-1)})$$

(3.4) $$\underline{\beta}_1' = (\beta_0, \beta_1, \beta_2, \ldots, \beta_{d-1})$$

(3.5) $$\underline{\beta}_2' = (\beta_d, \beta_{d+1}, \ldots, \beta_{d+k-1}) .$$

It is clear that $\underline{x}_1$ and $\underline{\beta}_1$ are (dxl) vectors and that $\underline{x}_2$ and $\underline{\beta}_2$ are (kxl) vectors, and that $\underline{\beta}_1$ and $\underline{\beta}_2$ represent the unknown parameters in the assumed true model.

Over some region of interest R, scaled to the closed interval $-1 \leq x \leq 1$, it is desired to fit η by taking N observations on η, where the observations on η are represented by $y = \eta$ + error in the usual manner, and using the approximating function, called the estimator

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

where the (dx1) vector $\underline{b}$ is to be determined.

Thus, (d-1) gives the degree of the fitted equation (e.g., d=2 implies fitting with a first order or linear equation and (d+k-1) gives the degree of the true model; so that k represents the "number of degrees" omitted in fitting.

The observational structure is

$$\underline{y}' = \text{the N observations on } \eta$$

(3.7) $$= (y_1, y_2, \ldots, y_N),$$

where

(3.8) $$\underline{y} = \underline{\eta} + \underline{\epsilon} ,$$

with the usual assumptions on $\underline{\epsilon}$, i.e. $E\underline{\epsilon} = \underline{0}$, Var $\underline{\epsilon} = I\,\sigma^2$, but not necessarily assuming normality. The "X" matrices are

(3.9) $$X_1 = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{d-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{d-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^{d-1} \end{pmatrix}$$

and

(3.10) $$X_2 = \begin{pmatrix} x_1^{d} & x_1^{d+1} & \cdots & x_1^{d+k-1} \\ x_2^{d} & x_2^{d+1} & \cdots & x_2^{d+k-1} \\ \vdots & \vdots & & \vdots \\ x_N^{d} & x_N^{d+1} & \cdots & x_N^{d+k-1} \end{pmatrix} .$$

Restricting attention to those $\underline{b}$'s which are linear functions of the observations, set

$$(3.11) \qquad \underline{b} = T'\underline{y},$$

where the (Nxd) matrix T is to be determined. The problem is now to determine T so that the desired optimality criteria are satisfied. Thus, choice of T defines the estimator, which can now be written as

$$(3.12) \qquad \hat{y}(x) = \underline{x}_1' \, T'\underline{y} \; .$$

Obviously, if the familiar least squares estimator is required then $T' = (X_1'X_1)^{-1}X_1'$, where X_1 is as in (3.9).

3.2 Bias Criterion

At the point x, the "bias contribution is defined to be

$$(3.13) \qquad \text{Bias } (x) = E[\hat{y}(x)] - \eta(x) \; ,$$

where

$$(3.14) \qquad E[\hat{y}(x)] = \underline{x}_1' \; E(\underline{b}) \; .$$

Similarly

$$(3.15) \qquad \text{Squared Bias} = \{E[\hat{y}(x)] - \eta(x)\}^2 \; .$$

In terms of the model (3.1), and the fitted equation (3.6) and using (3.11) and (3.14),

$$(3.16) \qquad E[\hat{y}(x)] = \underline{x}_1'T'E(\underline{y})$$

$$= \underline{x}_1'T'(X_1\underline{\beta}_1 + X_2\underline{\beta}_2)$$

$$= \underline{x}_1'T'\underline{\eta} \; .$$

Hence, (3.15) becomes

$$(3.17) \qquad \{E[\hat{y}(x)] - \eta(x)\}^2 = [\underline{x}_1' T'(X_1\underline{\beta}_1 + X_2\underline{\beta}_2) - (\underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\underline{\beta}_2)]^2$$

$$(3.18) \qquad = [\underline{x}_1'(T'X_1\underline{\beta}_1 + T'X_2\underline{\beta}_2 - \underline{\beta}_1) - \underline{x}_2'\underline{\beta}_2]^2 .$$

Define the (dx1) vector $\underline{\alpha}$ as

$$(3.19) \qquad \underline{\alpha} = \underline{\alpha}(T) = T'X_1\underline{\beta}_1 + T'X_2\underline{\beta}_2 - \underline{\beta}_1 .$$

Then (3.18) becomes, using (3.19), and remembering that $\underline{\alpha}$ is a function of T,

$$(3.20) \qquad \{E[\hat{y}(x)] - \eta(x)\}^2 = (\underline{x}_1'\underline{\alpha} - \underline{x}_2'\underline{\beta}_2)' \; (\underline{x}_1'\underline{\alpha} - \underline{x}_2'\underline{\beta}_2)$$

$$(3.21) \qquad = \underline{\alpha}'\underline{x}_1\underline{x}_1'\underline{\alpha} - 2\underline{\alpha}'\underline{x}_1\underline{x}_2'\underline{\beta}_2 + \underline{\beta}_2'\underline{x}_2\underline{x}_2'\underline{\beta}_2 .$$

The bias criterion for choice of T is (3.21) "averaged" over the region of interest, and therefore, it is required to choose T, independently of $\underline{\beta}_1$ and $\underline{\beta}_2$ to minimize

$$(3.22) \qquad B = B(T) = (\tfrac{1}{2}) \int_{-1}^{1} \{E[\hat{y}(x)] - \eta(x)\}^2 dx ,$$

where the integrand is given by (3.21) and $\frac{1}{2} = (\int_{-1}^{1} dx)^{-1}$.

Thus, from (3.22) and (3.21)

$$(3.23) \qquad B = (\tfrac{1}{2}) \int_{-1}^{1} (\underline{\alpha}'\underline{x}_1\underline{x}_1'\underline{\alpha} - 2\underline{\alpha}'\underline{x}_1\underline{x}_2'\underline{\beta}_2 + \underline{\beta}_2'\underline{x}_2\underline{x}_2'\underline{\beta}_2)dx$$

$$(3.24) \qquad = \underline{\alpha}'W_1\underline{\alpha} - 2\underline{\alpha}'W_2\underline{\beta}_2 + \underline{\beta}_2'W_3\underline{\beta}_2 ,$$

where

$$(3.25) \qquad W_1 = (\tfrac{1}{2}) \int_{-1}^{1} \underline{x}_1 \underline{x}_1' dx$$

$$(3.26) \qquad W_2 = (\tfrac{1}{2}) \int_{-1}^{1} \underline{x}_1 \underline{x}_2' dx$$

$$(3.27) \qquad W_3 = (\tfrac{1}{2}) \int_{-1}^{1} \underline{x}_2 \underline{x}_2' dx \ .$$

It is understood that where the integrand is a matrix in (3.25), (3.26) and (3.27), the integration is performed element by element. The three matrices W_1, W_2, W_3 are constants independent of T.

The (dxd) matrix W_1 is symmetric, positive definite (Manson, 1966) and therefore, B can be written

$$(3.28) \qquad B = (\underline{\alpha} - W_1^{-1} W_2 \underline{\beta}_2)' W_1 (\underline{\alpha} - W_1^{-1} W_2 \underline{\beta}_2) + \underline{\beta}_2' (W_3 - W_2' W_1^{-1} W_2) \underline{\beta}_2 \ .$$

Since T appears in B only through $\underline{\alpha}$, B = B(T) will be a minimum with respect to T when

$$(3.29) \qquad \underline{\alpha} = W_1^{-1} W_2 \underline{\beta}_2 \ .$$

Using (3.19), the definition of $\underline{\alpha}$, (3.29) becomes

$$(3.30) \qquad T' X_1 \underline{\beta}_1 + T' X_2 \underline{\beta}_2 = W_1^{-1} W_2 \underline{\beta}_2 + \underline{\beta}_1 \ .$$

Hence, sufficient conditions on T for satisfying (3.30), independent of $\underline{\beta}_1$ and $\underline{\beta}_2$, and hence for minimizing B are

$$(3.31) \qquad T' X_1 = I_d$$

$$(3.32) \qquad T' X_2 = W_1^{-1} W_2 \ .$$

Conditions (3.31) and (3.32), which can be written as

$$(3.33) \qquad T'X = A ,$$

where

$$(3.34) \qquad X = (X_1 \,\vdots\, X_2)$$

$$(3.35) \qquad A = (I_d \,\vdots\, W_1^{-1}W_2)$$

are the minimum B (min B) conditions.

The conditions (3.33) on the elements of the T matrix mean that if a T matrix satisfies (3.33) for a given X matrix then using the estimator $\hat{y}(x) = \underline{x}_1' T' \underline{y}$ will give min B, when η is the true model. For example, if least squares is required then T' is automatically fixed at $(X_1'X_1)^{-1}X_1'$ and then the conditions (3.32) which yield $(X_1'X_1)^{-1}X_1'X_2 = W_1^{-1}W_2$ are <u>forced on the design</u>.

Furthermore, the value of minimum B is, for any design,

$$(3.36) \qquad \min_T B = \underline{\beta}_2' \, (W_3 - W_2'W_1^{-1}W_2)\underline{\beta}_2 \; .$$

3.3 Variance Criterion

Equation (3.33) represents a set of sufficient conditions on the elements of T which generate a min B estimator. Subject to these conditions the secondary criterion, a variance criterion, is introduced to specify T, still independently of the choice of design. In this manner min B, and subject to min B, minimization of V is obtained for any experimental design.

From (3.6) and (3.11)

$$(3.37) \qquad \hat{y}(x) = \underline{x}_1'\underline{b}$$

$$= \underline{x}_1'T'\underline{y} \ ,$$

and therefore, the variance of the estimate $\hat{y}$, at the point x is, using the conditions on $\underline{\varepsilon}$,

$$(3.38) \qquad \text{var}\ \hat{y}(x) = \underline{x}_1'T'T\underline{x}_1\sigma^2 \ .$$

It is now required that subject to (3.35), var $\hat{y}(x)$ "averaged" over the region of interest be a minimum for choice of T. Thus, minimum B and subject to this, minimum V (min V | min B) is realized by minimizing

$$(3.39) \qquad V = \sigma^2 \int_R \underline{x}_1'T'T\underline{x}_1\,dx$$

with respect to T, and subject to (3.33).

It is now suggestive to write the variance of $\hat{y}$ as

$$(3.40) \qquad \text{var}\ \hat{y}(x;T) = \underline{x}_1'T'T\underline{x}_1\sigma^2$$

$$(3.41) \qquad = E[\hat{y}(x;T) - E\hat{y}(x;T)]^2$$

$$(3.42) \qquad = E[\underline{x}_1'T'\underline{y} - \underline{x}_1'T'X\underline{\beta}]^2 \ ,$$

where

$$(3.43) \qquad \underline{\beta}' = (\underline{\beta}_1' \ \vdots \ \underline{\beta}_2) \ .$$

Since $E(\underline{b}) = A\underline{\beta}$, var $\hat{y}(x;T)$ will be minimized with respect to T if $\underline{b} = T'\underline{y}$ is the minimum variance unbiased estimate of $A\underline{\beta}$, in a model which includes all the parameters. This, of course, follows from the Gauss-Markov theorem and yields therefore

(3.46) $\underline{b} = T'\underline{y} = A(X'X)^{-1}X'\underline{y}$.

Hence the T matrix which minimizes (3.40), for any x, is given by

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

If this minimizing T is denoted by T_0, then clearly

(3.48) $\text{var } \hat{y}(x;T_0) \leq \text{var } \hat{y}(x;T)$, for all x and T.

Therefore,

(3.49) $\int_R \text{var } \hat{y}(x;T_0)dx \leq \int_R \text{var } \hat{y}(x;T)dx$ for any T ,

and hence (3.47) gives the T matrix which produces the (min V | min B) estimator. Thus, the (min V | min B) estimator $\hat{y}(x)$ is given by

(3.50) $\hat{y}(x) = \underline{x}_1'A(X'X)^{-1}X'\underline{y}$.

Note that since (3.50) requires that $(X'X)$ be non-singular, a requirement is

(3.51) $N > d + k - 1 =$ degree of true model.

3.4 Some Properties of the Estimator

By using the estimator given by (3.50) minimum B and subject to minimum B, minimum V is realized for any experimental design. It is important to note that the $\hat{y}(x)$ so obtained depends on the degree of the assumed true model as reflected in the A and X matrices. Thus applying (3.50) enables an experimenter to fit an estimator satisfying certain criteria of optimality, while at the same time "guarding" against possible higher order effects. Section 4 will give an additional kind of protection in terms of experimental design conditions and for the present certain properties of the estimator are stated.

From (3.11), (3.8), (3.31), (3.32), (3.47) it is clear that

$$(3.52) \qquad E(\underline{b}) = E(T'\underline{y})$$

$$= T'E(\underline{y})$$

$$= T'(X_1\underline{\beta} + X_2\underline{\beta}_2)$$

$$= T'X_1\underline{\beta}_1 + T'X_2\underline{\beta}_2$$

$$(3.53) \qquad = \underline{\beta}_1 + W_1^{-1}W_2\underline{\beta}_2 \ .$$

Similarly,

$$(3.54) \qquad \text{Var }(\underline{b}) = \text{Var }(T'\underline{y})$$

$$= T'T\,\sigma^2$$

$$= A(X'X)^{-1}X'X(X'X)^{-1}A'\sigma^2$$

$$(3.55) \qquad = A(X'X)^{-1}A'\sigma^2 \ .$$

In terms of the estimator, (3.52) and (3.54) become

$$(3.56) \qquad E[\hat{y}(x)] = E(\underline{x}_1'\underline{b})$$

$$(3.57) \qquad = \underline{x}_1'(\underline{\beta}_1 + W_1^{-1}W_2\underline{\beta}_2) \ ,$$

$$(3.58) \qquad \text{Var } \hat{y}(x) = \text{Var }(\underline{x}_1'\underline{b})$$

$$(3.59) \qquad = \underline{x}_1'A(X'X)^{-1}A'\underline{x}_1\sigma^2 \ .$$

It is of interest, both from computational and certain design considerations, to develop a partitioned representation of the T matrix, and hence of the estimator.

From (3.47), (3.10), (3.9), and (3.34), T' can be written as

$$(3.60) \qquad T' = A(X'X)^{-1}X'$$

$$(3.61) \qquad T' = (I_d \mid W_1^{-1}W_2) \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}^{-1} \left(\frac{X_1'}{X_2'}\right) .$$

Let

$$S = X'X$$

$$(3.62) \qquad S_1 = X_1'X_1$$

$$S_2 = X_1'X_2$$

$$S_3 = X_2'X_2 .$$

Then (3.61) becomes

$$(3.63) \qquad T' = (I_d \mid W_1^{-1}W_2) \begin{pmatrix} S_1 & S_2 \\ S_2' & S_3 \end{pmatrix}^{-1} \left(\frac{X_1'}{X_2'}\right) .$$

Assuming S^{-1} exists, it can be written as

$$(3.64) \qquad S^{-1} = \begin{pmatrix} S_1^{-1}-S_1^{-1}S_2(S_2'S_1^{-1}S_2-S_3)^{-1}S_2'S_1^{-1} & S_1^{-1}S_2(S_2'S_1^{-1}S_2-S_3)^{-1} \\ (S_2'S_1^{-1}S_2-S_3)^{-1}S_2'S_1^{-1} & -(S_2'S_1^{-1}S_2-S_3)^{-1} \end{pmatrix}.$$

Then T' becomes

$$(3.65) \qquad T' = [S_1^{-1} + (W_1^{-1}W_2-S_1^{-1}S_2)(S_2'S_1^{-1}S_2-S_3)^{-1}S_2'S_1^{-1}]\, X_1'$$

$$- (W_1^{-1}W_2-S_1^{-1}S_2)(S_2'S_1^{-1}S_2-S_3)^{-1}X_2' .$$

From (3.65), the partitioned representation of the estimator is

$$(3.66) \qquad \hat{y}(x) = \underline{x}_1' \{[S_1^{-1} + (W_1^{-1}W_2-S_1^{-1}S_2)(S_2'S_1^{-1}S_2-S_3)^{-1}S_2'S_1^{-1}]\, X_1'$$

$$- (W_1^{-1}W_2-S_1^{-1}S_2)(S_2'S_1^{-1}S_2-S_3)^{-1}X_2'\} \underline{y} .$$

4. DERIVATION OF DESIGN MOMENT CONDITIONS FOR ONE DIMENSIONAL CASE

For a given degree of an estimating polynomial, represented by (d-1), and a given assumed degree of the true model, represented by (d+k-1), the estimator $\hat{y}(x)$ given by (3.50), or equivalently the matrix T' given by (3.47) guarantee the attainment of minimum B, and minimum V for this minimum B, for any experimental design. Additional criteria, i.e. criteria other than the bias and variance type criteria already satisfied by the estimation procedure, can now be introduced which may be satisfied by imposing certain conditions on the moments of the experimental design. For example, by requiring the condition that all of the odd moments vanish, symmetric designs are obtained. Similarly, from the partitioned representation of the T matrix as in (3.65), it is clear that if in addition to a (min V | min B) estimator a least squares estimator was a requirement, then it would be sufficient, if $(X'X)^{-1}$ exists, to have the design moments satisfy $W_1^{-1}W_2 = S_1^{-1}S_2$; because in this case T' would reduce to $S_1^{-1}X_1'$ and $\hat{y}(x)$ would then be the least squares estimator, $\underline{x}_1'(X_1'X_1)^{-1}X_1'\underline{y}$.

It should be mentioned that in all that follows in this chapter, and the next, the estimator used is the (min V | min B) estimator.

4.1 A Criterion for Choice of Design

The criterion introduced to select experimental designs is related to offering protection against different values of k. The derivation of the (min V | min B) estimator requires assumptions about the degree of the true model. For a given d, different estimators are obtained for different values of k. It is felt that a criterion consistent with and pertinent to the development of the (min V | min B) estimator is one that will generate the same estimator for various values of k protected against. Thus, protection against different values of k means that a particular estimator can be used, by choice of experimental design, which will be (min V | min B) for these different values of k. For example, if d=3 and k=1, i.e., if a quadratic estimator is used and the true model is assumed to be cubic, then applying (3.47) yields a certain T matrix, say T_1. If d=3 and k=2 is assumed, (3.47) yields another T matrix, say T_2. Protection against k=1 and k=2 simultaneously means choosing experimental designs which force $T_1=T_2$, and hence make the corresponding estimators $\hat{y}_1$ and $\hat{y}_2$ identical. Alternatively, in the response surface setting protection can be defined by supposing it is desired to approximate a particular response surface over a specified region of interest with a polynomial of degree d-1, denoted by $\hat{y}$. Then a $\hat{y}$ given by (3.50) can be determined, by choice of design, which will minimize B and subject to this will also minimize V for different "true η's" of higher degree than d-1.

It is important to note the relation of the above concept of protection to the requirement of least squares estimators, since the latter will be referred to and maintained as a familiar reference frame. As previously discussed, the design conditions leading to the least squares estimator are X'X non-singular and $S_1^{-1}S_2 = W_1^{-1}W_2$ where W_2 and S_2 depend on k, the number of terms assumed in the model but not included in the estimator $\hat{y}$. Consider the case k=r and partition the W_2 matrix, defined in (3.26), as

(4.1) $$W_2 = (\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_r) ,$$

where $\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_r$ are the (dxl) columns of W_2. Partition the (Nxr) X_2 matrix of (3.10) as

(4.2) $$X_2 = (\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_r) .$$

From (4.2) and (3.62)

(4.3) $$S_2 = X_1'(\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_r) .$$

Therefore, when k=r, the imposition of the least squares estimator as a requirement leads to the following conditions on the design moments,

(4.4) $$S_1^{-1} S_2 = W_1^{-1} W_2 .$$

From (4.1) and (4.3), these conditions are

(4.5) $$S_1^{-1} X_1'(\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_r) = W_1^{-1}(\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_r)$$

which means

(4.6) $$S_1^{-1} X_1' \underline{z}_k = W_1^{-1} \underline{w}_k \quad \text{for } k = 1,2,3,\ldots,r .$$

This means that a design leading to the least squares estimator, $\hat{y} = \underline{x}_1'(X_1'X_1)^{-1}X_1'\underline{y}$, as a (min V | min B) estimator for a true model with $X_2 = (\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_r)$ will automatically also be a least squares for any true model of lower degree. Since the estimator would be the same $\hat{y} = \underline{x}_1'(X_1'X_1)^{-1} X_1'\underline{y}$ we would automatically have $T_1 = T_2 = T_3 = \ldots = T_r$. That is, any design satisfying (4.5) will not only have a least squares estimator but will be a (min V | min B) design for which $T_1 = T_2 = T_3 = \ldots = T_r$.

However, recognizing and appreciating the result of the previous paragraph, it also shall be demonstrated that to realize (min V | min B), the "least squares" requirement may be more restrictive than necessary.

It is possible to satisfy $T_1 = T_2 = \ldots = T_r$ without necessarily using "least squares" designs. In some cases this additional flexibility may be useful to an experimenter.

The remainder of this chapter derives the protection conditions on the design moments for fitting linear, quadratic and cubic estimators, in the one dimension case, and protecting against higher order models up to and including quintic effects for each estimator.

One requirement initially imposed is that all odd moments arising in a particular "d and k" situation be zero. This requirement of symmetric designs is made mostly as a matter of making much of the algebra which follows more tractable. However, it should also be noted that symmetric designs are intuitively appealing to most experimenters.

It is convenient at this time to introduce some additional notation. It is standard to refer to the (Nxd) X_1 matrix as the design matrix. The ($N\times\overline{d+k}$) X matrix, defined in (3.34) as $X = (X_1|X_2)$, is frequently used for d+k = 3,4,5 and 6. Let

(4.7)

$$X_3 = \text{The } N\times 3 \text{ X matrix}$$

$$X_4 = \text{The } N\times 4 \text{ X matrix}$$

$$X_5 = \text{The } N\times 5 \text{ X matrix}$$

$$X_6 = \text{The } N\times 6 \text{ X matrix}.$$

That is, for example, X_4 is a matrix with the N columns $(1 \;\; x_i \;\; x_i^2 \;\; x_i^3)$, i = 1,2,...,N. Also let

$$(4.8)\qquad \begin{aligned} M_3 &= X_3'X_3 \\ M_4 &= X_4'X_4 \\ M_5 &= X_5'X_5 \\ M_6 &= X_6'X_6 \ . \end{aligned}$$

It is also standard to denote the j^{th} design moment by

$$(4.9)\qquad [1^j] = \frac{\sum_{u=1}^{N} x_u^{\,j}}{N} \ ,$$

Note that symmetric designs have $[1^j] = 0$ and j odd. Also let i index the values of k for the given d under study. Then T_i and A_i will denote the corresponding T and A matrices.

Of interest and used in the sequel are conditions for non-singularity of M_3, M_4, M_5 and M_6. These conditions are

(a) M_3 non-singular if

$$(4.10)\qquad [1^4] \neq [1^2]^2 \ ;$$

(b) M_4 non-singular if

$$(4.11)\qquad [1^4] \neq [1^2]^2, \quad [1^2][1^6] \neq [1^4]^2 \ ;$$

(c) M_5 non-singular is

$$(4.12)\qquad [1^2][1^6] \neq [1^4]^2,$$

$$(4.13)\qquad [1^8][1^4] + 2[1^4][1^6][1^2] - [1^2]^2[1^8] - [1^4]^3 - [1^6]^2 \neq 0;$$

(d) M_6 non-singular if

$$(4.14)\qquad [1^8][1^4] + 2[1^4][1^6][1^2] = [1^2]^2[1^8] - [1^4]^3 - [1^6]^2 \neq 0;$$

$$(4.15)\qquad [1^2][1^6][1^{10}] + 2[1^4][1^6][1^8] - [1^2][1^8]^2 - [1^4]^2[1^{10}] - [1^6]^3 \neq 0 \ .$$

4.2 Design Moment Conditions for d=2; k=1,2,3,4

When d=2, a linear estimator of the form $\hat{y} = b_0 + b_1 x$ is being fitted, where b_0 and b_1 are determined from (3.11) and (3.47) depending upon the assumed value of k.

Protection against quadratic (k=1) and cubic (k=2) effects means choosing designs which yield $T_1 = T_2$.

(i) For k=1, from (3.25), (3.26) and (3.27)

$$W_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix} , \tag{4.16}$$

$$W_2 = \begin{bmatrix} 1/3 \\ 0 \end{bmatrix} , \tag{4.17}$$

$$W_3 = 1/5 , \tag{4.18}$$

and

$$T_1' = A_1 M_3^{-1} X_3' , \tag{4.19}$$

where

$$A_1 = \begin{bmatrix} 1 & 0 & 1/3 \\ 0 & 1 & 0 \end{bmatrix} \tag{4.20}$$

$$= \begin{bmatrix} \underline{a}_{01}' \\ \underline{a}_{11}' \end{bmatrix} , \tag{4.21}$$

$$X_3' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_N \\ x_1^2 & x_2^2 & \cdots & x_N^2 \end{bmatrix} \tag{4.22}$$

$$(4.23) \qquad M_3 = N\begin{bmatrix} 1 & 0 & [1^2] \\ & [1^2] & 0 \\ \text{(symmetric)} & & [1^4] \end{bmatrix},$$

and

$$(4.24) \qquad M_3^{-1} = \frac{1}{N[1^2]([1^4]-[1^2]^2)}\begin{bmatrix} [1^2][1^4] & 0 & -[1^2]^2 \\ & [1^4]-[1^2]^2 & 0 \\ \text{(symmetric)} & & [1^2] \end{bmatrix}.$$

It is convenient to write (4.19) as

$$(4.25) \qquad T_1' = \begin{bmatrix} \underline{t}_{01}' \\ \underline{t}_{11}' \end{bmatrix},$$

where from (4.18)

$$(4.26) \qquad \underline{t}_{01}' = \underline{a}_{01}' \; M_3^{-1}X_3' = \text{first row of } T_1'$$

$$(4.27) \qquad \underline{t}_{11}' = \underline{a}_{02}' \; M_3^{-1}X_3' = \text{second row of } T_1' \quad .$$

(ii) For k=2,

$$(4.28) \qquad T_2' = A_2M_4^{-1}X_4' ,$$

where

$$(4.29) \qquad A_2 = \begin{bmatrix} 1 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 3/5 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} A_1 & \begin{matrix} 0 \\ 3/5 \end{matrix} \end{bmatrix} \tag{4.30}$$

$$= \begin{bmatrix} \underline{a}'_{01} & 0 \\ \underline{a}'_{11} & 3/5 \end{bmatrix} \tag{4.31}$$

$$= \begin{bmatrix} \underline{a}'_{02} \\ \underline{a}'_{12} \end{bmatrix} , \tag{4.32}$$

$$X'_4 = \begin{bmatrix} 1 & 1 & \cdots, & 1 \\ x_1 & x_2 & \cdots & x_N \\ x_1^2 & x_2^2 & \cdots & x_N^2 \\ x_1^3 & x_2^3 & \cdots & x_N^3 \end{bmatrix} , \tag{4.33}$$

$$= \left[\frac{X'_3}{\underline{z}'_3}\right] , \tag{4.34}$$

with

$$\underline{z}'_3 = (x_1^3, x_2^3, \ldots, x_N^3) , \tag{4.35}$$

and

$$(4.36)\qquad M_4 = N\begin{bmatrix} 1 & 0 & [1^2] & 0 \\ & [1^2] & 0 & [1^4] \\ & & [1^4] & 0 \\ (\text{symmetric}) & & & [1^6] \end{bmatrix}$$

$$(4.37)\qquad = \begin{bmatrix} M_3 & \underline{u}_1 \\ \underline{u}_1' & N[1^6] \end{bmatrix},$$

where

$$(4.38)\qquad \underline{u}_1' = N(0 \quad [1^4] \quad 0)\,.$$

Let

$$(4.39)\qquad c_1 = (\underline{u}_1' M_3^{-1} \underline{u}_1 - N\;[1^6]\;)^{-1} = \text{a scalar},$$

and note that c_1 is well defined if $[1^4]^2 \neq [1^2][1^6]$, since $c_1 = -\,|M_3|\;/\;|M_4|$.

Then, from (4.34) and (4.35)

$$(4.40)\qquad M_4^{-1} = \begin{bmatrix} M_3^{-1}(I_3 - \underline{u}_1 c_1 \underline{u}_1' M_3^{-1}) & M_3^{-1}\underline{u}_1 c_1 \\ c_1 \underline{u}_1' M_3^{-1} & -c_1 \end{bmatrix},$$

where (4.11) gives the conditions for the existence of M_4^{-1}. Therefore, combining (4.28), (4.31), (4.34) and (4.40),

$$(4.41) \qquad T_2' = \begin{bmatrix} \underline{a}_{01}' & 0 \\ \\ \underline{a}_{11}' & 3/5 \end{bmatrix} \begin{bmatrix} M_3^{-1}(I_3-\underline{u}_1c_1\underline{u}_1'M_3^{-1}) & M_3^{-1}\underline{u}_1c_1 \\ \\ c_1\underline{u}_1'M_3^{-1} & -c_1 \end{bmatrix} \begin{bmatrix} X_3' \\ \\ \underline{z}_3' \end{bmatrix} .$$

From (4.41)

$$(4.42) \qquad \underline{t}_{02}' = \underline{a}_{01}'M_3^{-1}X_3' - \underline{a}_{01}'M_3^{-1}\underline{u}_1(c_1\underline{u}_1'M_3^{-1}X_3'-c_1\underline{z}_3')$$

$$= \text{first row of } T_2' ,$$

$$(4.43) \qquad \underline{t}_{12}' = \underline{a}_{11}'M_3^{-1}X_3'-(\underline{a}_{11}'M_3^{-1}\underline{u}_1-3/5)(c_1\underline{u}_1'M_3^{-1}X_3')+(\underline{a}_{11}'M_3^{-1}\underline{u}_1-3/5)c_1\underline{z}_3'$$

$$= \text{second row of } T_2' .$$

This leads to

Result I: A sufficient condition for $T_1' = T_2'$ is

$$(4.44) \qquad [1^4] = (3/5)\,[1^2] .$$

Proof: The two matrices T_1 and T_2 are equal if and only if $\underline{t}_{01}' = \underline{t}_{02}'$ and $\underline{t}_{11}' = \underline{t}_{12}'$. First it is shown that $\underline{t}_{01}' = \underline{t}_{02}'$. From (4.26) and (4.39)

$$\underline{t}_{01}' = \underline{a}_{01}'M_3^{-1}X_3'$$

$$\underline{t}_{02}' = \underline{a}_{01}'M_3^{-1}X_3' - \underline{a}_{01}'M_3^{-1}\underline{u}_1(c_1\underline{u}_1'M_3^{-1}X_3'-c_1\underline{z}_3') .$$

From (4.24) and (4.38), $(M_3^{-1}\underline{u}_1)'$ is a vector of the form (0, element, 0), and from (4.17) and (4.18), $\underline{a}_{01}' = (1 \;\; 0 \;\; 1/3)$. Hence, $\underline{a}_{01}'M_3^{-1}\underline{u}_1 = 0$ and therefore $\underline{t}_{01}' = \underline{t}_{02}'$. The remainder of the proof is to show that $\underline{t}_{11}' = \underline{t}_{12}'$ if $[1^4] = (3/5)\,[1^2]$. From (4.27) and (4.43)

$$\underline{t}_{11}' = \underline{a}_{11}' M_3^{-1} X_3' \qquad \text{and}$$

$$\underline{t}_{12}' = \underline{a}_{11}' M_3^{-1} X_3' - (\underline{a}_{11}' M_3^{-1} \underline{u}_1 - 3/5) c_1 \underline{u}_1' M_3^{-1} X_3' + (\underline{a}_{11}' M_3^{-1} \underline{u}_1 - 3/5) c_1 \underline{z}_3' \ .$$

From (4.20), (4.21), (4.24) and (4.38),

$$\underline{a}_{11}' M_3^{-1} \underline{u}_1 = (0 \quad 1 \quad 0)\, M_3^{-1} \underline{u}_1$$

$$= [1^4]/[1^2] \ .$$

Therefore, if $[1^4] = (3/5)[1^2]$, then $\underline{t}_{11}' = \underline{t}_{12}'$, which establishes result 1.

Combining (4.44) with the conditions for the non-singular matrices, protection against k=1,2, when d=2 is obtained when

$$(4.45) \qquad [1^4] = (3/5)[1^2]$$

$$[1^2] \neq 3/5, \ [1^6] \neq (9/25)[1^2] \ ,$$

It is of interest to note that the least squares design conditions for fitting linear estimators to assumed cubic models are, from (4.4) with d=2, r=2 and a symmetric design, $[1^2] = 1/3$ and $[1^4] = 1/5$; a special case of result 1.

(iii) For k=3, in order now to equate T_1, T_2 and T_3, along with (4.45) additional conditions are needed.

As before

$$(4.46) \qquad T_3' = A_3 M_5^{-1} X_5' \ ,$$

where

(4.47) $$A_3 = \begin{bmatrix} 1 & 0 & 1/3 & 0 & 1/5 \\ 0 & 1 & 0 & 3/5 & 0 \end{bmatrix}$$

(4.48) $$= \begin{bmatrix} \underline{a}'_{02} & 1/5 \\ \underline{a}'_{12} & 0 \end{bmatrix}$$

(4.49) $$= \begin{bmatrix} \underline{a}'_{03} \\ \underline{a}'_{13} \end{bmatrix} ,$$

(4.50) $$X'_5 = \begin{bmatrix} X'_4 \\ \hline \underline{z}'_4 \end{bmatrix} ,$$

with

(4.51) $$\underline{z}'_4 = (x_1^4, x_2^4, \ldots, x_N^4) ,$$

and

(4.52) $$M_5 = N \begin{bmatrix} 1 & 0 & [1^2] & 0 & [1^4] \\ & [1^2] & 0 & [1^4] & 0 \\ & & [1^4] & 0 & [1^6] \\ & & & [1^6] & 0 \\ \text{(symmetric)} & & & & [1^8] \end{bmatrix}$$

(4.53) $$= \begin{bmatrix} M_4 & \underline{u}_2 \\ \underline{u}'_2 & N[1^8] \end{bmatrix} ,$$

where

(4.54) $\quad \underline{u}_2' = N([1^4] \quad 0 \quad [1^6] \quad 0)$.

In a manner similar to the proof of result 1 it is seen that the additional condition for equality of T_1, T_2 and T_3 is

$$(4.55) \qquad [1^6] = (9/25)\,[1^2]\ .$$

However, (4.55) violates (4.45). Hence, the least squares solution, $[1^2] = 1/3$, $[1^4] = 1/5$ may be used with the additional restraint that $[1^6] \neq 3/25$, so that M_4 is non-singular, and also $(4/45)[1^8] + [1^6]((2/15) - [1^6]) \neq (1/5)^3$, so that M_5 is non-singular. These additional restraints are necessary because the (min V | min B) estimator requires that $(X'X)^{-1}$ exist and to show that designs for which $S_1^{-1}S_2 = W_1^{-1}W_2$ lead to the least squares estimator it is also necessary to have a non-singular $(X'X)$ matrix, for the full model.

Thus, in summary, protection is obtained for quadratic, cubic and quartic effects if

$$(4.56) \qquad \begin{aligned} &[1^2] = 1/3 \\ &[1^4] = 1/5 \\ &[1^6] \neq 3/25 \\ &(4/45)[1^8] + ((2/15) - [1^6])[1^6] \neq 1/125\ . \end{aligned}$$

(iv) The k=4 case is included for completeness, because if $T_1=T_2=T_3$ can only be satisfied by the least squares conditions, then obviously it is not possible to satisfy $T_1=T_2=T_3=T_4$ for other designs. The conditions are

$$[1^2] = 1/3$$

$$[1^4] = 1/5$$

(4.57) $$[1^6] = 1/7$$

$$[1^8] \neq \frac{129}{1225} \text{ (so that } M_5^{-1} \text{ exists)}$$

$$\frac{4}{525}[1^{10}] - [1^8](1/3\,[1^6] - 2/35) \neq 0 \text{ (so that } M_6^{-1} \text{ exists).}$$

4.3 Design Moment Conditions for d=3, k=1,2,3

When d=3, a quadratic estimator of the form $\hat{y} = b_0 + b_1x + b_2x^2$ is being used. Again, b_0, b_1 and b_2 are determined from (3.11) and (3.47) depending upon the assumed values of k.

(i) For k=1,

(4.58) $$T_1' = A_1M_4^{-1}X_4'$$

where

(4.59) $$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3/5 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(4.60) $$= \begin{bmatrix} \underline{a}_{01}' \\ \underline{a}_{11}' \\ \underline{a}_{21}' \end{bmatrix},$$

and X_4' is given by (4.30), M_4 is given by (4.33) and

$$(4.61) \qquad M_4^{-1} = \frac{1}{N}\begin{bmatrix} \frac{[1^4]}{[1^4]-[1^2]^2} & 0 & \frac{-[1^2]}{[1^4]-[1^2]^2} & 0 \\ & \frac{1^6}{[1^2][1^6]-[1^4]^2} & 0 & \frac{-[1^4]}{[1^2][1^6]-[1^4]^2} \\ (\text{symmetric}) & & \frac{1}{[1^4]-[1^2]^2} & 0 \\ & & & \frac{[1^2]}{[1^2][1^6]-[1^4]^2} \end{bmatrix} .$$

Writing (4.58) as

$$(4.62) \qquad T_1' = \begin{bmatrix} \underline{t}_{01}' \\ \underline{t}_{11}' \\ \underline{t}_{21}' \end{bmatrix}$$

means

$$(4.63) \qquad \underline{t}_{01}' = \underline{a}_{01}' M_4^{-1} X_4' = \text{first row of } T_1'$$

$$(4.64) \qquad \underline{t}_{11}' = \underline{a}_{11}' M_4^{-1} X_4' = \text{second row of } T_1'$$

$$(4.65) \qquad \underline{t}_{21}' = \underline{a}_{21}' M_4^{-1} X_4' = \text{third row of } T_1' \; .$$

(ii) For k=2

$$(4.66) \qquad T_2' = A_2 M_5^{-1} X_5' \; ,$$

where

$$A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & -3/35 \\ 0 & 1 & 0 & 3/5 & 0 \\ 0 & 0 & 1 & 0 & 6/7 \end{bmatrix} \tag{4.67}$$

$$= A_1 \begin{bmatrix} -3/35 \\ 0 \\ 6/7 \end{bmatrix} \tag{4.68}$$

$$= \begin{bmatrix} \underline{a}'_{01} & -3/35 \\ \underline{a}'_{11} & 0 \\ \underline{a}'_{21} & 6/7 \end{bmatrix} \tag{4.69}$$

$$= \begin{bmatrix} \underline{a}'_{02} \\ \underline{a}'_{12} \\ \underline{a}'_{22} \end{bmatrix} , \tag{4.70}$$

$$X'_5 = \left[\frac{X'_4}{\underline{z}'_4} \right] , \tag{4.71}$$

where

$$\underline{z}'_4 = (x_1^4, x_2^4, \ldots, x_N^4) , \tag{4.72}$$

$$(4.73) \qquad M_5 = N \begin{bmatrix} 1 & 0 & [1^2] & 0 & [1^4] \\ & [1^2] & 0 & [1^4] & 0 \\ & & [1^4] & 0 & [1^6] \\ & & & [1^6] & 0 \\ \text{(symmetric)} & & & & [1^8] \end{bmatrix}$$

$$(4.74) \qquad = \begin{bmatrix} M_4 & \underline{v}_1 \\ \underline{v}_1' & N[1^8] \end{bmatrix} ,$$

where

$$(4.75) \qquad \underline{v}_1' = N(\; [1^4] \quad 0 \quad [1^6] \quad 0 \;) .$$

Let

$$(4.76) \qquad e_1 = (\underline{v}_1' M_4^{-1} \underline{v}_1 - N[1^8])^{-1} = \text{a scalar,}$$

noting that $e_1 = - |M_4|/|M_5|$. Then, from (4.61), (4.62), and assuming (4.12) and (4.13)

$$(4.77) \qquad M_5^{-1} = \begin{bmatrix} M_4^{-1}(I_4 - \underline{v}_1 e_1 \underline{v}_1' M_4^{-1} & M_4^{-1}\underline{v}_1 e_1 \\ e_1 \underline{v}_1' M_4^{-1} & -e_1 \end{bmatrix} .$$

Combining (4.66), (4.69), (4.71) and (4.77) yields

$$(4.78) \qquad T_2' = \begin{bmatrix} \underline{a}_{01}' & -3/35 \\ \underline{a}_{11}' & 0 \\ \underline{a}_{21}' & 6/7 \end{bmatrix} \begin{bmatrix} M_4^{-1}(I_4 - \underline{v}_1 e_1 \underline{v}_1' M_4^{-1}) & M_4^{-1}\underline{v}_1 e_1 \\ e_1 \underline{v}_1' M_4^{-1} & -e_1 \end{bmatrix} \begin{bmatrix} X_4' \\ \\ \underline{z}_4' \end{bmatrix} .$$

<u>Result 2</u>: Sufficient conditions for $T_1' = T_2'$ are

$$[1^4] = (3/35)(10\ [1^2] - 1)$$

(4.82) $$[1^6] = (3/245)(53\ [1^2]) - 6)$$

$$[6/53) < [1^2] < 1\ .$$

Proof: $T_1' = T_2'$ if and only if $\underline{t}_{01}' = \underline{t}_{02}'$, $\underline{t}_{11}' = \underline{t}_{12}'$ and $\underline{t}_{21}' = \underline{t}_{22}'$. First, conditions ensuring that $\underline{t}_{01}' = \underline{t}_{02}'$ are derived. Then it is shown that $\underline{t}_{11}' = \underline{t}_{12}'$ and then conditions giving $\underline{t}_{21}' = \underline{t}_{22}'$ are found. From (4.63) and (4.79)

$$\underline{t}_{01}' = \underline{a}_{01}' M_4^{-1} X_4'$$

$$\underline{t}_{02}' = \underline{a}_{01}' M_4^{-1} X_4' - (\underline{a}_{01}' M_4^{-1} \underline{v}_1 + 3/35)(e_1 \underline{v}_1' M_4^{-1} X_4') + (\underline{a}_{01}' M_4^{-1} \underline{v}_1 + 3/35) e_1 \underline{z}_4'\ .$$

From (4.58), (4.59), (4.61) and (4.75)

$$\underline{a}_{01}'(M_4^{-1} \underline{v}_1) = (1 \quad 0 \quad 0 \quad 0)\ M_4^{-1} \underline{v}_1$$

$$= \frac{[1^4]^2 - [1^2][1^6]}{[1^4] - [1^2]^2}\ .$$

If $\underline{a}_{01}' M_4^{-1} \underline{v}_1 = -3/35$, then $\underline{t}_{01}' = \underline{t}_{02}'$. Hence, the condition

(4.83) $$\frac{[1^4]^2 - [1^2][1^6]}{[1^4] - [1^2]^2} = -3/35$$

yields $\underline{t}_{01}' = \underline{t}_{02}'$. From (4.64) and (4.80)

$$\underline{t}_{11}' = \underline{a}_{11}' M_4^{-1} X_4' \qquad \text{and}$$

$$\underline{t}_{22}' = \underline{a}_{21}' M_4^{-1} X_4' - (\underline{a}_{21}' M_4^{-1} \underline{v}_1 - 6/7)(e_1 \underline{v}_1' M_4^{-1} X_4') + (\underline{a}_{21}' M_4^{-1} \underline{v}_1 - 6/7) e_1 \underline{z}_4'\ .$$

From (4.58), (4.59), (4.61) and (4.75)

$$\underline{a}_{21}'M_4^{-1}\underline{v}_1 = \frac{[1^6] - [1^2][1^4]}{[1^4] - [1^2]^2} .$$

If $\underline{a}_{21}'M_4^{-1}\underline{v}_1 = 6/7$, then $\underline{t}_{21}' = \underline{t}_{22}'$. Hence the condition

$$(4.84) \qquad \frac{[1^6] - [1^2][1^4]}{[1^4] - [1^2]^2} = 6/7$$

yields $\underline{t}_{21}' = \underline{t}_{22}'$. Equations (4.83) and (4.84), after some algebra (see Appendix 9.1) reduce to the first two equations of (4.82), while the inequality ensures $0 < [1^4], [1^6] < 1$. Hence result 2 is established.

Result 2 along with (4.13) yields the constraint

(4.85)

$$[1^8] \neq \frac{(828/1715)[1^2]^3-(28179/60025)[1^2]^2+(5346/60025)[1^2]-(1431/300125)}{[1^2]^2 - (6/7)\,[1^2] + 3/35} .$$

It is again of interest to note that the least squares protection conditions are from (4.4) with d=3 and r=2, $[1^2] = 1/3$, $[1^4] = 1/5$, $[1^6] = 1/7$; a special case of result 2.

(iii) For k=3, additional conditions are needed to equate T_1, T_2 and T_3 and thereby protect against quintic effects as well as cubic and quartic. Now

$$(4.86) \qquad T_3' = A_3M_6^{-1}X_6' ,$$

where

$$A_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & -3/35 & 0 \\ 0 & 1 & 0 & 3/5 & 0 & 3/7 \\ 0 & 0 & 1 & 0 & 6/7 & 0 \end{bmatrix} \tag{4.87}$$

$$= \begin{bmatrix} & 0 \\ A_2 & 3/7 \\ & 0 \end{bmatrix} \tag{4.88}$$

$$= \begin{bmatrix} \underline{a}'_{02} & 0 \\ \underline{a}'_{12} & 3/7 \\ \underline{a}'_{22} & 0 \end{bmatrix} \tag{4.89}$$

$$= \begin{bmatrix} \underline{a}'_{03} \\ \underline{a}'_{13} \\ \underline{a}'_{23} \end{bmatrix} , \tag{4.90}$$

$$X'_6 = \left[\frac{X'_5}{\underline{z}'_5} \right] , \tag{4.91}$$

where X'_5 is given by (4.71) and

$$\underline{z}'_5 = (x_1^5, x_2^5, \ldots, x_N^5) , \tag{4.92}$$

$$M_6 = \begin{bmatrix} M_5 & \underline{v}_2 \\ \underline{v}'_2 & N[1^{10}] \end{bmatrix} , \tag{4.93}$$

where

$$(4.94)\qquad \underline{v}_2' = N(0 \quad [1^6] \quad 0 \quad [1^8] \quad 0\ ) \ .$$

Letting

$$(4.95)\qquad e_2 = (\underline{v}_2' M_5^{-1} \underline{v}_2 - N[1^{10}])^{-1} = \text{a scalar},$$

from (4.93), (4.94), (4.95) and assuming (4.14) and (4.15),

$$(4.96)\qquad M_6^{-1} = \begin{bmatrix} M_5^{-1}(I_5 - \underline{v}_2 e_2 \underline{v}_2' M_5^{-1} & M_5^{-1}\underline{v}_2 e_2 \\ e_2 \underline{v}_2' M_5^{-1} & -e_2 \end{bmatrix},$$

and hence

$$(4.97)\qquad T_3' = \begin{bmatrix} \underline{a}_{02}' & 0 \\ \underline{a}_{12}' & 3/7 \\ \underline{a}_{22}' & 0 \end{bmatrix} \begin{bmatrix} M_5^{-1}(I_5 - \underline{v}_2 e_2 \underline{v}_2' M_5^{-1} & M_5^{-1}\underline{v}_2 e_2 \\ e_2 \underline{v}_2' M_5^{-1} & -e_2 \end{bmatrix} \begin{bmatrix} X_5' \\ \\ \underline{z}_5' \end{bmatrix} .$$

Also

$$(4.98)\qquad M_5^{-1} = \frac{1}{N} \begin{bmatrix} \frac{[1^8][1^4]-[1^6]^2}{h_1} & 0 & \frac{[1^4][1^6]-[1^8][1^2]}{h_1} & 0 & \frac{[1^2][1^6]-[1^4]^2}{h_1} \\ & \frac{[1^6]}{h_2} & 0 & \frac{-[1^4]}{h_2} & 0 \\ & & \frac{[1^8]-[1^4]^2}{h_1} & 0 & \frac{[1^2][1^4]-[1^6]}{h_1} \\ & \text{(symmetric)} & & \frac{[1^2]}{h_2} & 0 \\ & & & & \frac{[1^4]-[1^2]^2}{h_1} \end{bmatrix},$$

where

$$h_1 = [1^4][1^8] + 2[1^4][1^6][1^2] - [1^2]^2[1^8] - [1^4]^3 - [1^6]^2 \tag{4.99}$$

$$h_2 = [1^6][1^2] - [1^4]^2 \quad . \tag{4.100}$$

Thus, from (4.97)

$$\underline{t}'_{03} = \underline{a}'_{02}M_5^{-1}X'_5 - \underline{a}'_{02}M_5^{-1}\underline{v}_2(e_2\underline{v}'_2M_5^{-1}X'_2 - e_2\underline{z}'_5) = \text{first row of } T'_3 \ . \tag{4.101}$$

$$\underline{t}'_{13} = \underline{a}'_{12}M_5^{-1}X'_5 - (\underline{a}'_{12}M_5^{-1}\underline{v}_2 - 3/7)(e_2\underline{v}'_2M_5^{-1}X'_5) + (\underline{a}'_{12}M_5^{-1}\underline{v}_2 - 3/7)e_2\underline{z}'_5$$
$$= \text{second row of } T'_3 \ . \tag{4.102}$$

$$\underline{t}'_{23} = \underline{a}'_{22}M_5^{-1}X'_5 - \underline{a}'_{22}M_5^{-1}\underline{v}_2(e_2\underline{v}'_2M_5^{-1}X'_5 - e_2\underline{z}'_5) = \text{third row of } T'_3 \ . \tag{4.103}$$

Result 3: Sufficient conditions for $T'_1 = T'_2 = T'_3$ are

$$[1^4] = 3/35\ (10\ [1^2] - 1)$$

$$[1^6] = 3/245\ (53\ [1^2] - 6)$$

$$[1^8] = 9/8575\ (460\ [1^2] - 53) \tag{4.104}$$

$$53/460 < [1^2] < 1$$

$$[1^2] \neq 1/3 \quad .$$

Proof: It is sufficient to find conditions yielding $T'_2 = T'_3$ and then including result 2. First $\underline{t}'_{03} = \underline{t}'_{02}$ because (4.98) and (4.79) imply $\underline{t}'_{03} = \underline{t}'_{02} - \underline{a}'_{02}M_5^{-1}\underline{v}_2(e_2\underline{v}'_2M_5^{-1}X'_5 - e_2\underline{z}'_5)$ and from (4.67), (4.70), (4.98) and (4.94) $\underline{a}'_{02} = (1 \quad 0 \quad 0 \quad 0 \quad -3/35)$ and $(M_5^{-1}\underline{v}_2)'$ is of the form (0, element, 0, element, 0). Hence $\underline{a}'_{02}M_5^{-1}\underline{v}_2 = 0$. Furthermore, from (4.102) and (4.66) $\underline{t}'_{12} = \underline{a}'_{12}M_5^{-1}X'_5$ and $\underline{t}'_{13} = \underline{a}'_{12}M_5^{-1}X'_5 - (\underline{a}'_{12}M_5^{-1}\underline{v}_2 - 3/7)(e_2\underline{v}'_2M_5^{-1}X'_5) + (\underline{a}'_{12}M_5^{-1}\underline{v}_2 - 3/7)e_2\underline{z}'_5$. From (4.70), (4.98) and (4.94), $\underline{t}_{12}' = \underline{t}_{13}'$ if

$$(4.105) \qquad \frac{[1^6]^2 - [1^4][1^8] + (3/5)([1^2][1^8] - [1^4][1^6]}{[1^6]\,[1^2] - [1^4]^2} = 3/7 ,$$

since the left hand side of (4.105) is $\underline{a}_{12}'M_5^{-1}\underline{v}_2$. Similarly, (4.53) and (4.90) yield $\underline{t}_{22}' = \underline{a}_{22}'M_5^{-1}X_5'$ and $\underline{t}_{23}' = \underline{a}_{22}'M_5^{-1}X_5'-a_{22}'M_5^{-1}\underline{v}_2(e_2\underline{v}_2'M_5^{-1}X_5'-e_2\underline{z}_5')$, and (4.59), (4.60) and (4.61) imply $\underline{a}_{22}'M_5^{-1}\underline{v}_2 = 0$. Hence $\underline{t}_{22}' = \underline{t}_{23}'$. It remains to make (4.105) compatible with result 2. Altering (4.105) yields

$$(4.106) \qquad [1^8]((3/5)[1^2]-[1^4]) = (3/7)([1^2][1^6]-[1^4]^2)-[1^6]([1^6]-(3/5)[1^4]).$$

Notice that if $[1^2] = 1/3$, then result 2 implies $[1^4] = 1/5$, which makes the left hand side of (4.106) zero, and hence $[1^8]$ and $[1^{10}]$ may take on any value staisfying (4.14) and (4.15). In this case, then

$$(4.107) \qquad (3/7)([1^2][1^6] - [1^4]^2) = [1^6]([1^6] - (3/5)[1^4]) ;$$

but if $[1^2] = 1/3$ and $[1^4] = 1/5$ then (4.107) yields $[1^6] = 1/7$ or $[1^6] = 3/25$. Since $[1^6] = 3/25$ is not compatible with result 2 when $[1^2] = 1/3$, this value of $[1^2]$ is excluded when using result 2. Simplifying (4.107) and using result 2, yields after some algebra (see Appendix 9.2), the third equation of (4.104), where the inequality on $[1^2]$ ensures the moments lying between 0 and 1. This establishes result 3.

4.4 Design Moment Conditions for d=4; k=1,2

When d=4, a cubic estimator of the form $\hat{y} = b_0 + b_1x + b_2x^2 + b_3x^3$ is being used where the b's are determined from (3.11) and (3.72) depending upon the value of k.

(i) For k=1, the set-up is

$$(4.108)\qquad T_1' = A_1 M_5^{-1} X_5' \quad ,$$

where

$$(4.109)\qquad A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & -3/35 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 6/7 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$(4.110)\qquad = \begin{bmatrix} \underline{a}_{01}' \\ \underline{a}_{11}' \\ \underline{a}_{21}' \\ \underline{a}_{31}' \end{bmatrix}$$

and X_5' is given by (4.71) and M_5^{-1} is given by (4.98).

Writing (4.108) as

$$(4.111)\qquad T_1' = \begin{bmatrix} \underline{t}_{01}' \\ \underline{t}_{11}' \\ \underline{t}_{21}' \\ \underline{t}_{31}' \end{bmatrix}$$

yields

(4.112) $\underline{t}_{01}' = \underline{a}_{01}' M_5^{-1} X_5' =$ first row of T_1'

(4.113) $\underline{t}_{11}' = \underline{a}_{11}' M_5^{-1} X_5' =$ second row of T_1'

(4.114) $\underline{t}_{21}' = \underline{a}_{21}' M_5^{-1} X_5' =$ third row of T_1'

(4.115) $\underline{t}_{31}' = \underline{a}_{31}' M_5^{-1} X_5' =$ fourth row of T_1' .

(ii) For k=2,

(4.116) $T_2' = A_2 M_6^{-1} X_6'$,

where

$$A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & -3/35 & 0 \\ 0 & 1 & 0 & 0 & 0 & -5/21 \\ 0 & 0 & 1 & 0 & 6/7 & 0 \\ 0 & 0 & 0 & 1 & 0 & 10/9 \end{bmatrix} \tag{4.117}$$

$$= \begin{bmatrix} & 0 \\ & -5/21 \\ A_1 & 0 \\ & 10/9 \end{bmatrix} \tag{4.118}$$

$$= \begin{bmatrix} \underline{a}_{01}' & 0 \\ \underline{a}_{11}' & -5/21 \\ \underline{a}_{21}' & 0 \\ \underline{a}_{31}' & 10/9 \end{bmatrix} \tag{4.119}$$

$$(4.120)\qquad A_2 = \begin{bmatrix} \underline{a}'_{02} \\ \underline{a}'_{12} \\ \underline{a}'_{22} \\ \underline{a}'_{32} \end{bmatrix} ,$$

$$(4.121)\qquad X'_6 = \begin{bmatrix} X'_5 \\ \underline{z}'_5 \end{bmatrix} ,$$

as in (4.91) and (4.92) and M_6^{-1} is given by (4.96). Thus, from (4.116), (4.119), (4.96) and (4.121)

$$(4.122)\qquad T'_2 = \begin{bmatrix} \underline{a}'_{01} & 0 \\ \underline{a}'_{11} & -5/21 \\ \underline{a}'_{21} & 0 \\ \underline{a}'_{31} & 10/9 \end{bmatrix} \begin{bmatrix} M_5^{-1}(I_5 - \underline{v}_2 e_2 \underline{v}'_2 M_5^{-1}) & M_5^{-1}\underline{v}_2 e_2 \\ e_2 \underline{v}'_2 M_5^{-1} & -e_2 \end{bmatrix} \begin{bmatrix} X'_5 \\ \underline{z}'_5 \end{bmatrix} .$$

and therefore

$$(4.123)\qquad \underline{t}'_{02} = \underline{a}'_{01} M_5^{-1} X'_5 - \underline{a}'_{01} M_5^{-1} \underline{v}_2 (e_2 \underline{v}'_2 M_5^{-1} X'_5 - e_2 \underline{z}'_5)$$

$= \text{first row of } T'_2$,

$$(4.124)\qquad \underline{t}'_{12} = \underline{a}'_{11} M_5^{-1} X'_5 - (\underline{a}'_{11} M_5^{-1} \underline{v}_2 + 5/21)(e_2 \underline{v}'_2 M_5^{-1} X'_5) + (\underline{a}'_{11} M_5^{-1} \underline{v}_2 + 5/21) e_2 \underline{z}'_5$$

$= \text{second row of } T'_2$,

$$(4.125)\qquad \underline{t}'_{22} = \underline{a}'_{21} M_5^{-1} X'_5 - \underline{a}'_{21} M_5^{-1} \underline{v}_2 (e_2 \underline{v}'_2 M_5^{-1} X'_5 - e_2 \underline{z}'_5)$$

$= \text{third row of } T'_2$,

$$(4.126)\qquad \underline{t}'_{32} = \underline{a}'_{31} M_5^{-1} X'_5 - (\underline{a}'_{31} M_5^{-1} \underline{v}_2 - 10/9)(e_2 \underline{v}'_2 M_5^{-1} X'_5) + (\underline{a}'_{31} M_5^{-1} \underline{v}_2 - 10/9) e_2 \underline{z}'_5$$

$= \text{fourth row of } T'_2$.

<u>Result 4</u>: Sufficient conditions for $T_1' = T_2'$ are

$$[1^6] = (5/21)((14/3)\ [1^4] - [1^2])$$

(4.127) $$[1^8] = (50/189)((113/30)\ [1^4] - [1^2]\ ,$$

$$(30/113)[1^2] < 1^4 < (9/10) + (3/14)[1^2]\ .$$

Proof: Since $T_1' = T_2'$ if and only if $\underline{t}_{01}' = \underline{t}_{02}'$, $\underline{t}_{11}' = \underline{t}_{12}'$, $\underline{t}_{21}' = \underline{t}_{22}'$ and $\underline{t}_{31}' = \underline{t}_{32}'$, the proof shall follow the same procedure as in results 1, 2 and 3. First $\underline{t}_{01}' = \underline{t}_{02}'$, because, from (4.123) and (4.112) it is sufficient to show that $\underline{a}_{01}' M_5^{-1} \underline{v}_2 = 0$. From (4.98) and (4.94), as seen previously $(M_5^{-1}\underline{v}_2)'$ is of the form (0, element, 0, element, 0) and from (4.109) and (4.110), $\underline{a}_{01}' = (1 \quad 0 \quad 0 \quad 0 \quad -3/35)$. Hence $\underline{t}_{01}' = \underline{t}_{02}'$. Similarly, from (4.109) and (4.110), $\underline{a}_{21}' = (0 \quad 0 \quad 1 \quad 0 \quad 6/7)$ and therefore $\underline{a}_{21}' M_5^{-1} \underline{v}_2 = 0$, which implies $\underline{t}_{21}' = \underline{t}_{22}'$ from (4.125) and (4.114). It remains to find the conditions which will equate the second and fourth rows of T_1' and T_2'. From (4.113) and (4.124)

$$\underline{t}_{11}' = \underline{a}_{11}' M_5^{-1} X_5'$$

$$\underline{t}_{12}' = \underline{a}_{11}' M_5^{-1} X_5' - (\underline{a}_{11}' M_5^{-1} \underline{v}_2 + 5/21)(e_2 \underline{v}_2' M_5^{-1} X_5') + (\underline{a}_{11}' M_5^{-1} \underline{v}_2 + 5/21) e_2 \underline{z}_5'.$$

From (4.109), (4.110), (4.98) and (4.94)

$$\underline{a}_{11}' M_5^{-1} \underline{v}_2 = \frac{[1^6]^2 - [1^4][1^8]}{[1^6][1^2] - [1^4]^2}\ .$$

Thus, $\underline{t}_{11}' = \underline{t}_{12}'$ if

(4.128) $$\frac{[1^6]^2 - [1^4][1^8]}{[1^6][1^2] - [1^4]^2} = \frac{-5}{21}\ .$$

Similarly, from (4.115) and (4.126)

$$\underline{t}_{31}' = \underline{a}_{31}' M_5^{-1} X_5'$$

$$\underline{t}_{32}' = \underline{a}_{31}' M_5^{-1} X_5' - (\underline{a}_{31}' M_5^{-1} \underline{v}_2 - 10/9)(e_2 \underline{v}_2' M_5^{-1} X_5') + (\underline{a}_{31}' M_5^{-1} \underline{v}_2 - 10/9) e_2 \underline{z}_5' .$$

Again, from (4.109), (4.110), (4.98) and (4.94)

$$\underline{a}_{31}' M_5^{-1} \underline{v}_2 = \frac{[1^2][1^8] - [1^4][1^6]}{[1^6][1^2] - [1^4]^2} ,$$

so that $\underline{t}_{31}' = \underline{t}_{32}'$ if

$$\frac{[1^2][1^8] - [1^4][1^6]}{[1^6][1^2] - [1^4]^2} = \frac{10}{9} . \qquad (4.129)$$

After some algebra (see Sppendix 9.3), (4.128) and (4.129) reduce to the two equations of (4.114). From these equations, the inequality ensures that $[1^6]$ and $[1^8]$ are between 0 and 1. Hence result 4 is established.

It is convenient and useful for the following section to summarize the important results of this section as in Table 4.1.

Table 4.1. Summary of design moment requirements for linear, quadratic and cubic estimators and protection against certain higher order effects

d	Values of k Protected against	Moment Conditions[a]
2	1,2	$[1^4] = (3/5)[1^2]$
2	1,2,3,4	$[1^2] = 1/3,\ [1^4] = 1/5,\ [1^6] = 1/7$
3	1,2	$[1^4] = (3/35)(10[1^2]-1)$
		$[1^6] = (3/245)(53[1^2]-6)$
		$(6/53) < [1^2] < 1$
3	1,2,3	$[1^4] = (3/35)(10[1^2]-1)$
		$[1^6] = (3/245)(53[1^2]-6)$
		$[1^8] = (9/8575)(460[1^2]-53)$
		$(53/460) < [1^2] < 1$
4	1,2	$[1^6] = (5/21)((14/3)[1^4]-[1^2])$
		$[1^8] = (50/189)((113/30)[1^4]-[1^2])$
		$(30/113)[1^2] < [1^4] < (9/10) + (3/14)[1^2]$

[a] The condition, to be found in the text, for non-singular (X'X) matrices are not included here.

5. SOME DESIGNS

For the design moment conditions generated in the previous section, actual designs are derived for specified experimental situations. Designs are developed for certain numbers of experimental levels presumed of interest to experimenters. Aside from the particular designs much of this material is to exhibit techniques, approaches, interpretations and meanings of resolving the design moment conditions. It is clear that the admissible values of solutions to equations involving the experimental levels will lead to additional requirements on the design moments. Some of the design possibilities will offer the experimenter more design moment flexibility than the earlier approaches.

An important solution technique is the application of the elementary symmetric functions, see Householder (1953), to finding experimental levels satisfying the design moment requirements. This technique was exploited for similar problems and discussed in detail by Manson (1966). Since this methodology is very useful, the particular type of problem involved and the corresponding method of solution is described below.

Consider the problem of finding $u_1, u_2, \ldots, u_t$ such that the equations

$$(5.1) \qquad s_p = \sum_{i=1}^{t} u_i^p = g_p(\gamma) \qquad \text{for} \quad p = 1, 2, \ldots, t ,$$

are satisfied for given γ and known g_p. Newton's identities relate the s_p to the elementary symmetric functions σ_j of the roots of the t^{th} degree polynomial $P(u)$. Thus, given γ, the σ_j are determined and represent the coefficients of u^{t-j} in $P(u)$, i.e.

$$(5.2) \qquad P(u) = u^t - \sigma_1 u^{t-1} + \sigma_2 u^{t-2} + \ldots + (-1)^t \sigma_t \ .$$

Hence, the u_i satisfying (5.1) are the roots of $P(u) = 0$.

For deriving the designs of this chapter σ_1, σ_2, σ_3 and σ_4 are needed. They are given by, see Manson (1966),

$$(5.3) \qquad \sigma_1 = s_1$$

$$(5.4) \qquad \sigma_2 = \frac{s_1^2 - s_2}{2!}$$

$$(5.5) \qquad \sigma_3 = \frac{s_1^3 - 3s_1s_2 + 2s_3}{3!}$$

$$(5.6) \qquad \sigma_4 = \frac{s_1^4 - 6s_1^2s_2 + 8s_1s_3 + 3s_2^2 - 6s_4}{4!} \ .$$

It is important also to note that a necessary condition for all the roots of $P(u)$ to be real and non-negative is that all the σ_j be non-negative.

For convenience of notation, denote $[1^2]$ by a, $[1^4]$ by b, $[1^6]$ by c, $[1^8]$ by e.

Most of the numerical results presented in this chapter and the next were obtained using an IBM 360 computer system.

5.1 Designs for d=2

In this section a class of designs is derived for d=2 with protection against k=1 and 2. The conditions which must be satisfied are

$b = 3a/5$, $a \neq 3/5$, $c \neq 9a/25$, $[1] = [1^3] = [1^5] = 0$ and $N \geq 4$. The class of designs is defined by consideration of experiments involving five levels $\ell_1, \ell_2, \ell_3, \ell_4, 0$.

Letting $\ell_3 = -\ell_1$ and $\ell_4 = -\ell_2$ satisfies the symmetry requirements and it is also convenient to write $N = 4m+n$, where $m = 1,2,3,\ldots;$ $n = 0,1,2,3$ and to define

$$(5.7) \qquad \alpha = (4m+n)/m = N/m = 4+n/m \;.$$

The design will then be completely specified by placing m observations at each of the four experimental values $x = \pm\ell_1$, $x = \pm\ell_2$ and n at $x=0$, where ℓ_1 and ℓ_2 are to be determined.

By the definition of a and from (5.7)

$$(5.8) \qquad a = \sum_{u=1}^{N} x_u^2/N$$

$$(5.9) \qquad = \frac{m\ell_1^2 + m\ell_2^2 + m\ell_3^2 + m\ell_4^2 + n(0)}{N}$$

$$(5.10) \qquad = \frac{2}{\alpha}(\ell_1^2 + \ell_2^2) \;,$$

or

$$(5.11) \qquad \ell_1^2 + \ell_2^2 = \frac{\alpha}{2}\, a \;.$$

Similarly

$$(5.12) \qquad \ell_1^4 + \ell_2^4 = \frac{\alpha}{2}\, b \;.$$

The condition $b = 3a/5$ and (5.12) yield

(5.13) $$b = \frac{2}{\alpha}(\ell_1^4 + \ell_2^4) = \frac{3}{5}a ,$$

which means

(5.14) $$\ell_1^4 + \ell_2^4 = 3\alpha\, a/10 .$$

Thus, (5.11) and (5.14) are the two equations

$$\ell_1^2 + \ell_2^2 = \frac{\alpha}{2}a$$

$$\ell_1^4 + \ell_2^4 = \frac{3\alpha}{10}a ,$$

and the elementary symmetry functions can now be introduced.

Setting

(5.15) $$s_1 = \ell_1^2 + \ell_2^2$$

(5.16) $$s_2 = \ell_1^4 + \ell_2^4 ,$$

it is clear that (5.11), (5.14), (5.15), (5.16) can be represented by the equations (5.1) where $t = 2$, $u_1 = \ell_1^2$, $u_2 = \ell_2^2$, $\gamma = a$, $g_1(\gamma) = \alpha\gamma/2$, $g_2(\gamma) = 3\alpha\gamma/10$. Hence using (5.3) and (5.4)

(5.17) $$\sigma_1 = \frac{\alpha}{2}a$$

(5.18) $$\sigma_2 = \frac{\alpha}{4}a\left(\frac{\alpha}{2}a - \frac{3}{5}\right) ,$$

and therefore ℓ_1^2 and ℓ_2^2 are the two roots of the equation

(5.19) $$u^2 - \sigma_1 u + \sigma_2 = 0 .$$

The necessity for σ_1 and σ_2 to be non-negative yields the additional constraint

(5.20) $\qquad a \geq 6/5\alpha$.

Substituting (5.17) and (5.18) in (5.19) and solving the quadratic equation yields

(5.21) $$u = \frac{\alpha}{4} a \pm \frac{1}{2} [\alpha a \, (\frac{3}{5} - \frac{\alpha}{4} a)]^{\frac{1}{2}} .$$

Since the two solutions are ℓ_1^2 and ℓ_2^2, it is necessary that $3/5 - \alpha a/4 \geq 0$ and this implies

(5.22) $\qquad a \leq 12/5\alpha$.

Hence, it is sufficient to set ℓ_1^2 and ℓ_2^2 at the values given in (5.21). To ensure that $c \neq 9a/25$, note

(5.23) $$c = \frac{2}{\alpha} (\ell_1^6 + \ell_2^6)$$

(5.24) $$= a^2 \, \alpha(\frac{9}{20} - \frac{\alpha}{8} a) ,$$

which means that $c = 9a/25$ if $a = 6/5\alpha$ or $12/5\alpha$, and therefore the equalities in (5.20) and (5.22) are not allowed.

Hence, summarizing the above development, the desired class of designs is given by:

$N = 4m+n$, $= N/m$, $m = 1,2,3,\ldots,$; $n = 0,1,2,3$.

m observations at $+\ell_1$

m observations at $-\ell_1$

m observations at $+\ell_2$

m observations at $-\ell_2$

n observations at 0 ,

where

$$(5.25) \qquad \ell_1^2 = \frac{\alpha}{4} a + \frac{1}{2} [\alpha a (\frac{3}{5} - \frac{\alpha}{4} a)]^{\frac{1}{2}}$$

$$(5.26) \qquad \ell_2^2 = \frac{\alpha}{4} a - \frac{1}{2} [\alpha a (\frac{3}{5} - \frac{\alpha}{4} a)]^{\frac{1}{2}}$$

$$(5.27) \qquad \frac{6}{5\alpha} < a < \frac{12}{5\alpha} .$$

An alternative way of looking at the design is by the explicit relation between ℓ_1 and ℓ_2. For simplicity take $\alpha = 4$, i.e. suppose N is divisible by 4. Since $\ell_1^4 + \ell_2^4 = (3/5)(\ell_1^2 + \ell_2^2)$, solving for ℓ_2^2 yields

$$(5.28) \qquad \ell_2^2 = \frac{3}{10} \pm \frac{1}{2} (\frac{9}{25} - 4\ell_1^4 + \frac{12}{5} \ell_1^2)^{\frac{1}{2}} .$$

Furthermore $(9/25) - 4\ell_1^4 + 12\ell_1^2/5 \geq 0$ implies

$$(5.29) \qquad \ell_1^2 \leq \frac{3}{10} (1 + \sqrt{2}) .$$

(This can also be found from $\max_a \ell_1^2$ in (5.25)).

Also, from (5.27), $3/5 < \ell_1^2 + \ell_2^2 < 6/5$ and hence all admissible values of ℓ_1 and ℓ_2 must lie between the two circles of radius $(.6)^{\frac{1}{2}}$ and $(1.2)^{\frac{1}{2}}$. Figure 5.1 shows the admissible values of ℓ_1 and ℓ_2.

5.2 Designs for d=3, k=1,2

From result 2, the conditions which must be satisfied are

$$b = \frac{3}{35} (10 a - 1)$$

$$c = \frac{3}{245} (53 a - 6)$$

$$\frac{6}{53} < a < 1 .$$

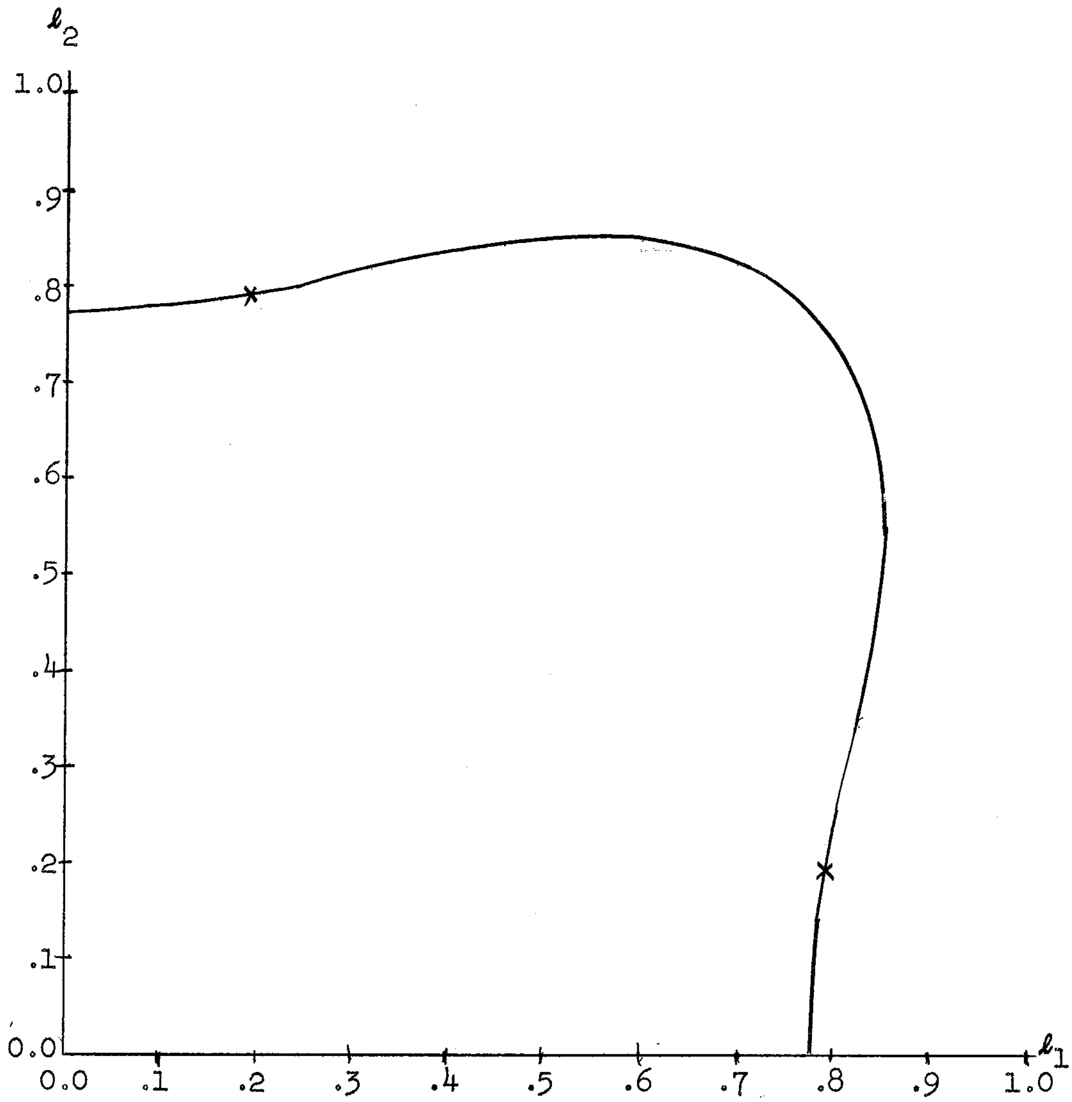

Figure 5.1. Values of ℓ_1 and ℓ_2 with N/4 observations at $\pm\ \ell_1$ and $\pm\ \ell_2$, for d=2, k=1,2 (X: the least squares design)

The class of designs is defined for experiments where N is divisible by six or six levels $\ell_1, \ell_2, \ldots, \ell_6$ are involved, with N/6 experiments at each level.

Consider symmetric designs with $\ell_4 = -\ell_1$, $\ell_5 = -\ell_2$, $\ell_6 = -\ell_3$. As in the preceding section, the second, fourth and sixth moments can be written

$$a = (\ell_1^2 + \ell_2^2 + \ell_3^2)/3 \tag{5.30}$$

$$b = (\ell_1^4 + \ell_2^4 + \ell_3^4)/3 \tag{5.31}$$

$$c = (\ell_1^6 + \ell_2^6 + \ell_3^6)/3 \quad . \tag{5.32}$$

Substitution of the requirements on b and c in terms of a yields the three equations

$$\ell_1^2 + \ell_2^2 + \ell_3^2 = 3a \tag{5.33}$$

$$\ell_1^4 + \ell_2^4 + \ell_3^4 = \frac{9}{35}(10\,a - 1) \tag{5.34}$$

$$\ell_1^6 + \ell_2^6 + \ell_3^6 = \frac{9}{245}(53\,a - 6) \quad . \tag{5.35}$$

In order to solve these equations for ℓ_1, ℓ_2, ℓ_3 in terms of the second moment, the elementary symmetric functions are used, where now the application of (5.1) is for $t = 3$, $u_i = \ell_i^2$, $\gamma = a$ and g_1, g_2, g_3 are the right hand sides of (5.33), (5.34), (5.35). Hence, from (5.3), (5.4) and (5.5)

(5.36) $\sigma_1 = s_1 = 3a$

(5.37) $$\sigma_2 = \frac{s_1^2 - s_2}{21} = \frac{9}{2}\left(a^2 - \frac{2}{7}a + \frac{1}{35}\right)$$

(5.38) $$\sigma_3 = \frac{s_1^3 - 3s_1s_2 + 2s_3}{31} = \frac{9}{2}\left(a^3 - \frac{6}{7}a^2 + \frac{169}{735}a - \frac{4}{245}\right) .$$

For $a > 6/53$, σ_1, σ_2, σ_3 are greater than zero. The requires values of ℓ_1^2, ℓ_2^2, ℓ_3^2 are therefore the roots of the cubic equation

(5.39) $$u^3 - \sigma_1 u^2 + \sigma_2 u - \sigma_3 = 0 ,$$

where σ_1, σ_2, σ_3 are given by (5.36), (5.37) and (5.38) in terms of a.

Varying a over the range (6/53,1) in increments of .01, and solving (5.39) leads to the results in Table 5.1. Note that only the range of a = .33 to a = .53 is admissible, since all other values generated 2 complex roots. It is also of interest to observe that if minimum bias subject to least squares estimation is desired, then for this experimental set-up, only the levels $\ell_1 = .86624682$, $\ell_2 = .26663540$, $\ell_3 = .42251865$, corresponding to a = 1/3, could be used.

5.3 Designs for d=4, k=1,2

When fitting a cubic and protecting against quartic and quintic true models, result 4 gives the conditions

$$c = \frac{5}{63}(14b - 3a)$$

$$e = \frac{5}{567}(113b - 30a) ,$$

Table 5.1. Designs for d=3, k=1,2 with N/6 experiments at each of $\pm \ell_1$, $\pm \ell_2$, $\pm \ell_3$

a	ℓ_1	ℓ_2	ℓ_3
.33	.86441137	.28063902	.40501201
.34	.86978621	.24632224	.45033021
.35	.87477345	.22563634	.48359037
.36	.87938347	.21185459	.51166624
.37	.88362364	.20295407	.53667393
.38	.88749837	.19796810	.55960278
.39	.89100901	.19628338	.58101272
.40	.89415364	.19742269	.60125996
.41	.89692670	.20097117	.62059091
.42	.89931834	.20655578	.63918795
.43	.90131348	.21384387	.65719465
.44	.90289038	.22254580	.67473130
.45	.90401836	.23241593	.69190580
.46	.90465420	.24325064	.70882290
.47	.90473611	.25488431	.72559393
.48	.90417304	.26718429	.74235009
.49	.90282402	.28004567	.75926492
.50	.90045340	.29338651	.77660030
.51	.89661579	.30714374	.79482252
.52	.89025754	.32136974	.81500139
.53	.87683336	.33572962	.84169405

with $0 < a,b,c,e < 1$. Consider designs using eight non-zero levels $\ell_1,\ell_2,\ldots,\ell_8$, placing N/8 observation at each level and symmetrized by setting $\ell_5 = -\ell_1$, $\ell_6 = -\ell_2$, $\ell_7 = -\ell_3$, $\ell_8 = -\ell_4$.

The second, fourth, sixth and eighth moments are

$$a = (\ell_1^2 + \ell_2^2 + \ell_3^2 + \ell_4^2)/4 \tag{5.40}$$

$$b = (\ell_1^4 + \ell_2^4 + \ell_3^4 + \ell_4^4)/4 \tag{5.41}$$

$$c = (\ell_1^6 + \ell_2^6 + \ell_3^6 + \ell_4^6)/4 \tag{5.42}$$

$$e = (\ell_1^8 + \ell_2^8 + \ell_3^8 + \ell_4^8)/4 \; . \tag{5.43}$$

In terms of the protection conditions, these equations are

$$\ell_1^2 + \ell_2^2 + \ell_3^2 + \ell_4^2 = 4a \tag{5.44}$$

$$\ell_1^4 + \ell_2^4 + \ell_3^4 + \ell_4^4 = 4b \tag{5.45}$$

$$\ell_1^6 + \ell_2^6 + \ell_3^6 + \ell_4^6 = 4c = \frac{20}{63}(14b - 3a) \tag{5.46}$$

$$\ell_1^8 + \ell_2^8 + \ell_3^8 + \ell_4^8 = 4e = \frac{20}{567}(113b - 30a) \; . \tag{5.47}$$

For specified values of a and b, equations (5.44) - (5.47) represent a system of four equations in the four unknowns ℓ_1, ℓ_2, ℓ_3, ℓ_4. Application of (5.1) is now for $t = 4$, $u_i = \ell_i^2$, $\gamma = (a,b)$ and g_1, g_2, g_3, g_4 the respective right hand sides of (5.44), (5.45), (5.46) and (5.47). Thus the four elementary symmetric functions are, from (5.3), (5.4), (5.5), and (5.6), (5.3) $\sigma_1 = s_1$

(5.48) $= 4a$

(5.4) $\sigma_2 = \frac{s_1^2 - s_2}{2!}$

(5.49) $= 2(4a^2-b)$

(5.5) $\sigma_3 = \frac{s_1^3 - 3s_1s_2 + 2s_3}{3!}$

(5.50) $= \frac{32}{3} a^3 - 8ab + \frac{40}{27} b - \frac{20}{63} a$

(5.6) $\sigma_4 = \frac{s_1^4 - 6s_1^2s_2 + 8s_1s_3 + 3s_2^2 - 6s_4}{4!}$

(5.51) $= \frac{32}{3} a^4 - 16a^2b - \frac{80}{63} a^2 + 2b^2 + \frac{1120}{189} ab + \frac{50}{189} a - \frac{565}{567} b$.

Solving the quartic equation

(5.52) $u^4 - \sigma_1 u^3 + \sigma_2 u^2 - \sigma_3 u + \sigma_4 = 0$

yields the required values of $\ell_1^2, \ell_2^2, \ell_3^2, \ell_4^2$. Since solutions are desired for various values of a and b in order to generate a family of designs, the region of admissible (a,b) values is at first confined to that region where σ_1, σ_2, σ_3 and σ_4 are non-negative; necessary for all the roots of (5.52) to be real. This, of course, is not sufficient for all the roots to be real and does not ensure that the roots lie between zero and one, but the conditions reduce the region of admissible designs in the (a,b) plane. Figure 5.2 shows this initially admissible region. (It is of interest to note that the minimum bias conditions on

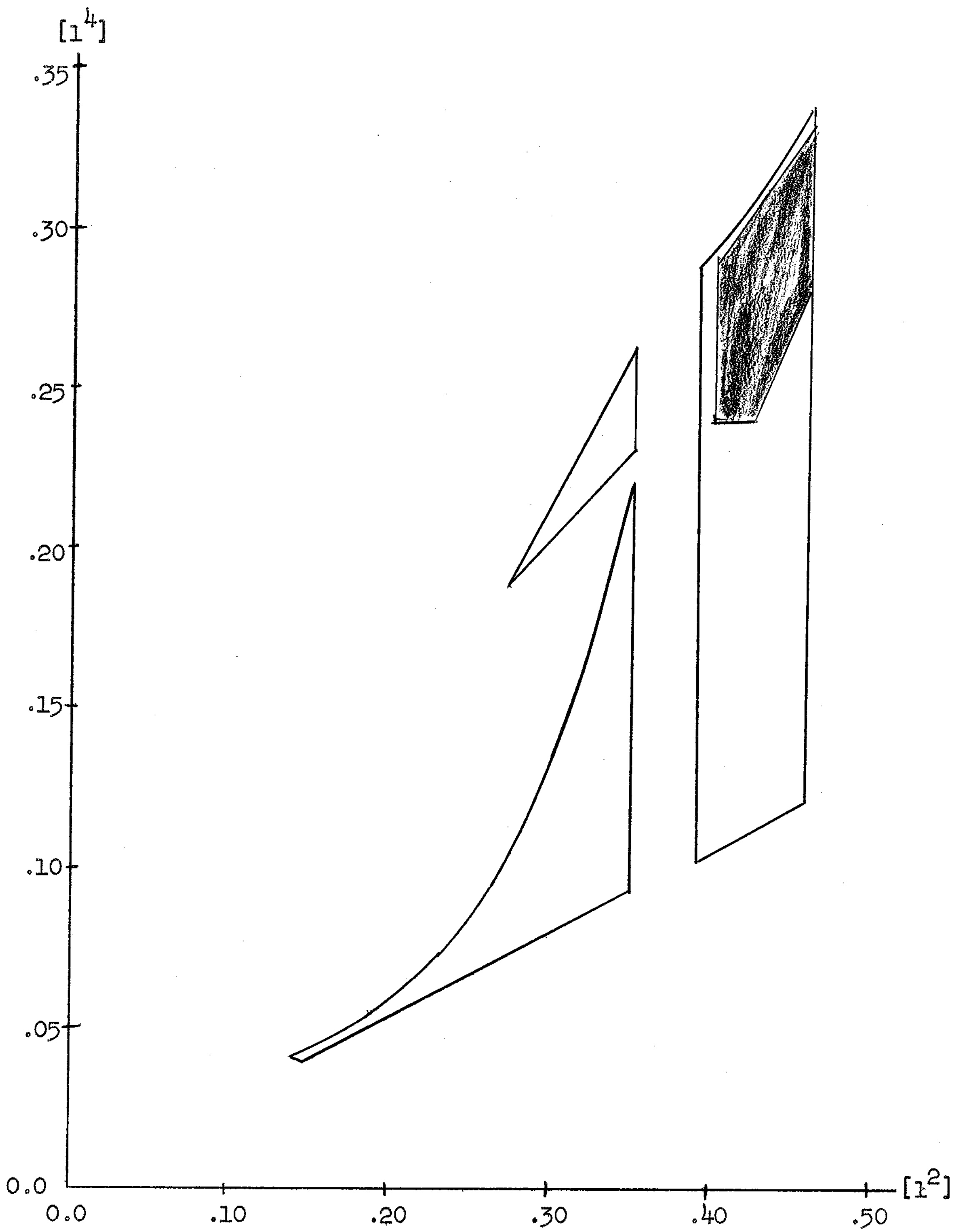

Figure 5.2. Initial admissible region on $[1^2]$ and $[1^4]$ for d=4, k=1,2, N/8 observations at $\pm \ell_1$, $\pm \ell_2$, $\pm \ell_3$, $\pm \ell_4$ (shaded area shows final admissible region)

the design to have a least squares estimator lead to an inadmissible result for this class of designs. This is because the least squares requirements $a = 1/3$, $b = 1/5$, $c = 1/7$, $e = 1/9$, mean that $\sigma_4 = -.00101$ and hence minimum bias designs cannot be attained, for this class, using the least squares estimator). Table 5.2 presents designs, obtained by "sweeping out" the admissible region of Figure 5.2, which satisfy the requirements, along with the corresponding values of the moments a and b.

5.4 d=3, k=1,2,3

Designs for this situation must satisfy, from result 3,

$$b = \frac{3}{35}(10\,a - 1)$$

$$c = \frac{3}{245}(53\,a - 6)$$

$$e = \frac{9}{8575}(460\,a - 53)\ ,$$

with $(53/460) < a < 1$. Thus, for example, with N experiments, 2m distinct non-zero levels $\ell_1, \ell_2, \ldots, \ell_m, \ell_{m+1} = -\ell_1, \ell_{m+2} = -\ell_2, \ldots, \ell_{2m} = -\ell_m$, and N divisible by 2m so that there are N/2m experiments at each ℓ_i, the problem is to solve the system of four equations in the m variables $\ell_1, \ell_2, \ldots, \ell_m$, given by

$$(5.53)\qquad \begin{aligned} \ell_1^2 + \ell_2^2 + \ldots + \ell_m^2 &= ma \\ \ell_1^4 + \ell_2^4 + \ldots + \ell_m^4 &= \frac{3m}{35}(10a-1) \\ \ell_1^6 + \ell_2^6 + \ldots + \ell_m^6 &= \frac{3m}{245}(53a-6) \\ \ell_1^8 + \ell_2^8 + \ldots + \ell_m^8 &= \frac{9m}{8575}(460a-53)\ . \end{aligned}$$

Table 5.2. Designs for d=4, k=1,2 with N/8 experiments at each of $\pm \ell_1$, $\pm \ell_2$, $\pm \ell_3$, $\pm \ell_4$

a	b	ℓ_1	ℓ_2	ℓ_3	ℓ_4
.400	.24	.63962956	.53607600	.91665500	.25147596
.405	.24	.62677741	.55653767	.91584868	.28042306
.410	.24	.60713511	.58092670	.91497807	.31100843
.425	.24	.61145621	.55324569	.91169107	.43458014
.400	.25	.68408360	.48759560	.92196057	.21040169
.405	.25	.67802394	.50505011	.92128764	.23756475
.410	.25	.67154472	.52055505	.92058420	.26565931
.415	.25	.66501213	.53344259	.91982855	.29514963
.420	.25	.69523584	.54223378	.91899549	.32686072
.425	.25	.65571176	.54391114	.91805546	.36245954
.430	.25	.65627060	.53207281	.91697346	.40665353
.400	.26	.71535163	.44873534	.92647406	.16898061
.405	.26	.71102403	.46816266	.92589535	.19490088
.410	.26	.70631837	.48561049	.92531553	.22155806
.415	.26	.70135707	.50097799	.92471612	.24903697
.420	.26	.69636526	.51387679	.92407580	.27765085
.425	.26	.69173659	.52347907	.92337002	.30799019
.430	.26	.68810545	.52817935	.92257036	.34117648
.435	.26	.68634597	.52462531	.92164349	.37982993
.440	.26	.68735276	.50024504	.92054976	.43576281
.400	.27	.74157156	.41115062	.93021857	.12538022
.405	.27	.73823905	.43369316	.92970125	.15022959

Table continued on next page

Table 5.2 (continued)

a	b	ℓ_1	ℓ_2	ℓ_3	ℓ_4
.410	.27	.73452774	.45384940	.92921024	.17623299
.415	.27	.73049120	.47188822	.92872925	.20289359
.420	.27	.72622119	.48784667	.92823957	.23017323
.425	.27	.72186731	.50153321	.92771981	.25827888
.430	.27	.71767022	.51246153	.92714564	.28763451
.435	.27	.71399825	.51967945	.92648916	.31899469
.440	.27	.71137240	.52132255	.92571814	.35386159
.445	.27	.71043226	.51278242	.92479472	.39635202
.400	.28	.76502898	.37172152	.93317941	.07569649
.405	.28	.76241015	.39871545	.93269370	.09919282
.410	.28	.75940922	.42231280	.93226172	.12584764
.415	.28	.75605404	.44330222	.93186875	.15325163
.420	.28	.75238953	.46208039	.93149773	.18084158
.425	.28	.74848615	.47879925	.93112923	.20990154
.430	.28	.74445267	.49340897	.93074129	.23676671
.435	.28	.74045435	.50564013	.93030922	.26566927
.440	.28	.73673625	.51492129	.92980516	.29586844
.445	.28	.73364586	.52017211	.92919734	.32894129
.450	.28	.73163883	.51918886	.92844900	.36486985
.455	.28	.73123857	.50535362	.92751670	.41185028
.400	.29	.78683340	.32727363	.93529479	.03148292
.410	.29	.78235124	.38949268	.93441003	.05567666
.415	.29	.77954263	.41402384	.93407917	.09161686
.420	.29	.77638695	.43578513	.93380203	.12380788

Table continued on next page

Table 5.2 (continued)

a	b	ℓ_1	ℓ_2	ℓ_3	ℓ_4
.425	.29	.77292301	.45527657	.93356001	.15420395
.430	.29	.76920819	.47273869	.93333243	.18365040
.435	.29	.76532664	.48822369	.93309589	.21270829
.440	.29	.76140109	.50160689	.93282479	.24186118
.445	.29	.75760848	.51255226	.93249063	.27162620
.450	.29	.75419794	.52041477	.93206154	.30268006
.455	.29	.75150472	.52399643	.93150125	.33611578
.420	.30	.79914275	.40839097	.93502042	.01801430
.425	.30	.79606953	.43078313	.93488949	.08173641
.430	.30	.79268143	.45088422	.93480883	.12038282
.435	.30	.78902817	.46899659	.93475501	.15397987
.440	.30	.78517905	.48524339	.93470258	.18537477
.445	.30	.78123100	.49959724	.93462397	.21577478
.450	.30	.77721976	.51187191	.93448934	.24594881
.455	.30	.77363367	.52167642	.93426601	.27657128
.435	.31	.81216194	.44846262	.93505383	.07034634
.440	.31	.80845103	.46710577	.93523142	.11645309
.445	.31	.80452150	.48398292	.93541922	.15328593
.450	.31	.80045470	.49913147	.93558625	.18659561
.455	.31	.79636032	.51246231	.93570110	.21830274
.445	.32	.82801485	.46671702	.93445284	.05800463
.450	.32	.82385947	.48409700	.93498633	.11272187
.455	.32	.81954550	.49988351	.93549674	.15266795
.455	.33	.84390157	.48532938	.93286994	.04515757

The solution must be such that ℓ_1^2, ℓ_2^2, ..., ℓ_m^2 are all real numbers lying between zero and one, for some values of a lying between (53/460) and 1. Attempts at solving (5.53), using the previously discussed solution techniques, have been made for m=4 and 5, and for values of m from 6 to 10 allowing experiments at zero with four non-zero levels, and no admissible solutions have yet been found.

6. GENERALIZATION TO HIGHER DIMENSIONS

This chapter includes extension of the previous development to fitting polynomial responses involving p controllable factors. Given the new set-up, the optimum estimator is the "natural extension" of the estimator derived in section 3, where the derivation did not explicitly depend on the dimension of the factor space. Examples of the design problem are also presented in this section.

6.1 The p-Dimensional Case

As in section 3.1, an entire region of operability is assumed, over which the true model η at the point $\underline{x}' = (x_1, x_2, \ldots, x_p)$ is a polynomial of degree d+k-1 in the p factors $x_1, x_2, \ldots, x_p$. The true model is again written

$$(6.1) \qquad \eta(\underline{x}) = \underline{x}_1'\underline{\beta}_1 + \underline{x}_2'\beta_2 \quad ,$$

where now

$$(6.2) \qquad \underline{x}_1' = (1; x_1, x_2, \ldots, x_p;\ x_1^2, x_2^2, \ldots, x_{p-1}x_p;\ x_1^3, x_2^3, \ldots, x_{p-2}x_{p-1}x_p;\ \ldots; \ldots; x_1^{d-1}, x_2^{d-1}, \ldots, x_{p-d+2}\cdots x_{p-1}x_p)$$

$$(6.3) \qquad \underline{x}_2' = (x_1^d, x_2^d, \ldots, x_{p-d+1}\cdots x_{p-1}x_p;\ \ldots;\ x_1^{d+k-1}, x_2^{d+k-1}, \ldots, x_{p-d-k+2}\cdots x_{p-1}x_p)$$

$$(6.4) \qquad \underline{\beta}_1' = (\beta_0;\ \beta_1, \beta_2, \ldots, \ldots, \beta_p;\ \beta_{11}, \beta_{22}, \ldots, \beta_{p-1,p};\ \beta_{111}, \beta_{222}, \ldots, \beta_{p-2,p-1,p};\ \ldots;\ \beta_{\underbrace{11\ldots1}_{d-1}}, \beta_{\underbrace{22\ldots2}_{d-1}}, \beta_{p-d+2,\ldots,p-1,p})$$

(6.5) $$\underline{\beta}_2' = (\beta_{\underbrace{11\ldots1}_{d}}, \beta_{\underbrace{22\ldots2}_{d}}, \ldots, \beta_{p-d+1,\ldots,p-1,p}, \ldots;$$

$$\ldots, \beta_{\underbrace{11\ldots1}_{d+k-1}}, \beta_{\underbrace{22\ldots2}_{d+k-1}}, \ldots, \beta_{p-d-k+2,\ldots,p}) .$$

Note that equations (6.2), (6.3), (6.4) and (6.5) correspond to equations (3.2), (3.3), (3.4) and (3.5).

Since a polynomial of degree q in r variables contains $\binom{q+r}{q}$ terms, $\eta(\underline{x})$ contains

(6.6) $$\theta_1 = \binom{p+d+k-1}{d+k-1}$$

terms, $\underline{x}_1$ and $\underline{\beta}_1$ are $(\theta_2 x 1)$ vectors, where

(6.7) $$\theta_2 = \binom{p+d-1}{d-1} ,$$

and $\underline{x}_2$ and $\underline{\beta}_2$ are $(\theta_3 x 1)$ vectors, where

(6.8) $$\theta_3 = \theta_1 - \theta_2 .$$

Over the region of interest R, η is to be fitted by the estimator

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

where the θ_2 elements of $\underline{b}$ are to be determined. The observational structure is as in (3.7) and (3.8) and the "X" matrices are now

$$(6.10) \qquad X_1 = \begin{bmatrix} \underline{x}'_{11} \\ \underline{x}'_{21} \\ \cdot \\ \cdot \\ \cdot \\ \underline{x}'_{N1} \end{bmatrix}$$

and

$$(6.11) \qquad X_2 = \begin{bmatrix} \underline{x}'_{12} \\ \underline{x}'_{22} \\ \cdot \\ \cdot \\ \cdot \\ \underline{x}'_{N2} \end{bmatrix}$$

so that X_1 is $(N \times \theta_2)$ and X_2 is $(N \times \theta_3)$. In (6.10) and (6.11) $\underline{x}'_{i1}$ and $\underline{x}'_{i2}$ represent the i^{th} observation on vectors of the type in (6.2) and (6.3).

As in (3.11) attention is restricted to those $\underline{b}$'s of the form $\underline{b} = T'\underline{y}$, where now T is the $(N \times \theta_2)$ matrix to be determined. Remembering the generalized set-up, the derivation of the (min V | min B) estimator is exactly as in section 3, noting that integration is now in p-space over the region R, and the W matrices in (3.25), (3.26) and (3.27) are extended to integrals over R with respect to $x_1, x_2, \ldots, x_p$. The problem is now to choose T to minimize

$$(6.12) \qquad B = \Omega \int_R \left\{ E[\hat{y}(\underline{x})] - \eta(\underline{x}) \right\}^2 d\underline{x} ,$$

where $\Omega^{-1} = \int_R d\underline{x}$, and subject to this, to minimize

(6.13) $$V = \Omega \int_R \text{Var}\ \hat{y}\ (\underline{x})d\underline{x}\ .$$

This leads to the T matrix given by (3.47), in the generalized form, and the (min V | min B) estimator, as in (3.50),

(6.14) $$\hat{y}(\underline{x}) = \underline{x}_1' A(X'X)^{-1}X'\underline{y}\ .$$

6.2 Two Examples of the Design Problem

Example 1. Suppose there are two controllable factors x_1 and x_2 and suppose that $\hat{y}(\underline{x})$ is a linear function and that protection is required against the true model being quadratic or cubic. Thus, p=2, d=2 and it is necessary to choose a design which equates the (min V | min B) estimators for k=1 and k=2. It is also assumed that R is the region defined by $-1 \leq x_1,\ x_2 \leq 1$.

When k=1; $\theta_1 = 6$, $\theta_2 = 3$, $\theta_3 = 3$ and

(6.15) $$\underline{x}_1' = (1;\ x_1, x_2)$$

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

(6.17) $$\underline{\beta}_1' = (\beta_0, \beta_1, \beta_2)$$

(6.18) $$\underline{\beta}_2' = (\beta_{11}, \beta_{22}, \beta_{12})\ ,$$

i.e.

(6.19) $$\eta(\underline{x}) = \eta(x_1, x_2)$$

(6.20) $$= \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2 + \beta_{12}x_1x_2\ .$$

The estimator is

$$\hat{y}(\underline{x}) = \underline{x}_1'\underline{b} \tag{6.21}$$

$$= b_0 + b_1x_1 + b_2x_2 \quad , \tag{6.22}$$

where

$$\underline{b} = T_1'\underline{y} \ , \tag{6.23}$$

and

$$T_1' = A_1(X_a'X_a)^{-1}X_a' \quad , \tag{6.24}$$

where A_1 and X_a' and $(X_a'X_a)^{-1}$ are defined below.

In this case

$$W_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}, \tag{6.25}$$

$$W_2 = \begin{bmatrix} 1/3 & 1/3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{6.26}$$

and therefore

$$W_1^{-1}W_2 = \begin{bmatrix} 1/3 & 1/3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad , \tag{6.27}$$

which means

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{6.28}$$

$$= \begin{bmatrix} \underline{a}'_{01} \\ \underline{a}'_{11} \\ \underline{a}'_{21} \end{bmatrix} . \tag{6.29}$$

Furthermore, here

$$X'_a = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ x_{11}^2 & x_{12}^2 & \cdots & x_{1N}^2 \\ x_{21}^2 & x_{22}^2 & \cdots & x_{2N}^2 \\ x_{11}x_{21} & x_{12}x_{22} & \cdots & x_{1N}x_{2N} \end{bmatrix} , \tag{6.30}$$

so that, assuming all odd moments are zero,

$$X'_aX_a = N \begin{bmatrix} 1 & 0 & 0 & [1^2] & [2^2] & 0 \\ 0 & [1^2] & 0 & 0 & 0 & 0 \\ 0 & 0 & [2^2] & 0 & 0 & 0 \\ [1^2] & 0 & 0 & [1^4] & [1^22^2] & 0 \\ [2^2] & 0 & 0 & [1^22^2] & [2^4] & 0 \\ 0 & 0 & 0 & 0 & 0 & [1^22^2] \end{bmatrix} , \tag{6.31}$$

and

(6.32) $\quad (X_a'X_a)^{-1} =$

$$\frac{1}{N}\begin{bmatrix} \frac{[1^4][2^4]-[1^22^2]^2}{h} & 0 & 0 & \frac{[1^22^2][2^2]-[1^2][2^4]}{h} & \frac{[1^22^2][1^2]-[1^4][2^2]}{h} & 0 \\ & \frac{1}{[1^2]} & 0 & 0 & 0 & 0 \\ & & \frac{1}{[2^2]} & 0 & 0 & 0 \\ \text{(symmetric)} & & & \frac{[2^4]-[2^2]^2}{h} & \frac{[1^2][2^2]-[1^22^2]}{h} & 0 \\ & & & & & \frac{1}{[1^22^2]} \end{bmatrix}$$

where

(6.33) $\quad h = [1^4][2^4] + 2[1^2][2^2][1^22^2]-[1^22^2]^2 - [1^4][2^2]^2 - [1^2]^2[2^4].$

Thus, using (6.29), (6.24) can be written

(6.34) $$T_1' = \begin{bmatrix} \underline{t}_{01}' \\ \underline{t}_{11}' \\ \underline{t}_{21}' \end{bmatrix}$$

(6.35) $$= \begin{bmatrix} \underline{a}_{01}' (X_a'X_a)^{-1}X_a' \\ \underline{a}_{11}' (X_a'X_a)^{-1}X_a' \\ \underline{a}_{21}' (X_a'X_a)^{-1}X_a' \end{bmatrix} .$$

When k=2; $\theta_1 = 10$, $\theta_2 = 3$, $\theta_3 = 7$ and now

$$(6.36) \qquad \underline{x}_2' = (x_1^2,\ x_2^2,\ x_1x_2,\ x_1^3,\ x_2^3,\ x_1x_2^2,\ x_1^2x_2) .$$

$$(6.37) \qquad \underline{\beta}_2' = (\beta_{11},\ \beta_{22},\ \beta_{12},\ \beta_{111},\ \beta_{222},\ \beta_{122},\ \beta_{112}) ,$$

and $\underline{x}_1'$ and $\underline{\beta}_1'$ are as in (6.15) and (6.17). The estimator is $\hat{y}(\underline{x}) = \underline{x}_1' T_2' \underline{y}$ where

$$(6.38) \qquad T_2' = A_2 (X_b' X_b)^{-1} X_b' \quad ,$$

where A_2 and X_b' are defined below.

In this case W_1 is as in (6.25) and

$$(6.39) \qquad W_2 = \begin{bmatrix} 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/5 & 0 & 1/9 & 0 \\ 0 & 0 & 0 & 0 & 1/5 & 0 & 1/9 \end{bmatrix}$$

so that

$$(6.40) \qquad W_1^{-1} W_2 = \begin{bmatrix} 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3/5 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 3/5 & 0 & 1/3 \end{bmatrix}$$

and hence

$$(6.41) \qquad A_2 = \begin{bmatrix} 1 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 3/5 & 0 & 1/3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 3/5 & 0 & 1/3 \end{bmatrix}$$

$$(6.42) \qquad A_2 = \begin{bmatrix} & 0 & 0 & 0 & 0 \\ A_1 & 3/5 & 0 & 1/3 & 0 \\ & 0 & 1/3 & 0 & 3/5 \end{bmatrix}$$

$$(6.43) \qquad = \begin{bmatrix} \underline{a}'_{01} & 0 & 0 & 0 & 0 \\ \underline{a}'_{11} & 3/5 & 0 & 1/3 & 0 \\ \underline{a}'_{21} & 0 & 1/3 & 0 & 3/5 \end{bmatrix}$$

$$(6.44) \qquad = \begin{bmatrix} \underline{a}'_{02} \\ \underline{a}'_{12} \\ \underline{a}'_{22} \end{bmatrix} .$$

Also

$$(6.45) \qquad X'_b = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ x_{11}^2 & x_{12}^2 & \cdots & x_{1N}^2 \\ x_{21}^2 & x_{22}^2 & \cdots & x_{2N}^2 \\ x_{11}x_{21} & x_{12}x_{22} & \cdots & x_{1N}x_{2N} \\ x_{11}^3 & x_{12}^3 & \cdots & x_{1N}^3 \\ x_{21}^3 & x_{22}^3 & \cdots & x_{2N}^3 \\ x_{11}x_{21}^2 & x_{12}x_{22}^2 & \cdots & x_{1N}x_{2N}^2 \\ x_{11}^2x_{21} & x_{12}^2x_{22} & \cdots & x_{1N}^2x_{2N} \end{bmatrix}$$

$$(6.46) \qquad X_b' = \left[\frac{X_a'}{U'}\right] ,$$

where U' is the $(4xN)$ matrix of the last four rows of X_b'. Let

$$(6.47) \qquad E = X_a'X_a$$

$$(6.48) \qquad F = X_a'U$$

$$(6.49) \qquad G = U'U \quad .$$

Then

$$(6.50) \qquad X_b'X_b = \begin{bmatrix} E & F \\ F' & G \end{bmatrix} ,$$

and

$$(6.51) \qquad (X_b'X_b)^{-1} = \begin{bmatrix} E^{-1}-E^{-1}FBF'E^{-1} & E^{-1}FB \\ B'F'E^{-1} & -B \end{bmatrix} ,$$

where

$$(6.52) \qquad B = (F'E^{-1}F-G)^{-1} ,$$

and E^{-1} is assumed to exist. From (6.38), (6.44), (6.47), and (6.51),

$$(6.53) \qquad \begin{bmatrix} \underline{t}_{02}' \\ \underline{t}_{12}' \\ \underline{t}_{22}' \end{bmatrix} = \begin{bmatrix} \underline{a}_{01}' & 0 & 0 & 0 & 0 \\ \underline{a}_{11}' & 3/5 & 0 & 1/3 & 0 \\ \underline{a}_{21}' & 0 & 1/3 & 0 & 3/5 \end{bmatrix} \begin{bmatrix} E^{-1}-E^{-1}FBF'E^{-1} & E^{-1}FB \\ B'F'E^{-1} & -B \end{bmatrix} \begin{bmatrix} X_a' \\ U' \end{bmatrix},$$

where

$$(6.54) \qquad T_2' = \begin{bmatrix} \underline{t}_{02}' \\ \underline{t}_{12}' \\ \underline{t}_{22}' \end{bmatrix} .$$

The problem is now to find sufficient conditions for equating the three rows of (6.53) and the three rows of (6.34).

Result 5: Sufficient conditions for $T_1' = T_2'$ are

$$(6.55) \qquad \begin{aligned} [1^4] &= 3/5\ [1^2] \\ [1^2 2^2] &= 1/3\ [1^2] \\ [1^2] &= [2^2] \\ [1^4] &= [2^4] , \end{aligned}$$

and the $X'X$ matrices in (6.32) and (6.51) are non-singular.

Proof: As in the previous results conditions are derived row by row. From (6.35) and (6.47)

$$\underline{t}_{01}' = \underline{a}_{01}'\ E^{-1}X_a'$$

and from (6.53)

$$\underline{t}_{02}' = \underline{a}_{01}'E^{-1}X_a' - \underline{a}_{01}'E^{-1}FBF'E^{-1}X_a' + \underline{a}_{01}'E^{-1}FBU' ,$$

but from (6.28), (6.29), (6.32), (6.45), (6.47), (6.48), $\underline{a}_{01}'E^{-1}F = \underline{0}'$, where $\underline{0}$ is a (4x1) null vector. Therefore $\underline{t}_{02}' = \underline{t}_{01}'$. Similarly, from (6.35) and (6.47)

$$\underline{t}_{11}' = \underline{a}_{11}'E^{-1}X_a' ,$$

and from (6.53), and noting that $B = B'$ in (6.52),

$$\underline{t}_{12}' = \underline{a}_{11}'E^{-1}X_a' - (\underline{a}_{11}'E^{-1}F-\underline{x}')BF'E^{-1}X_a' + (\underline{a}_{11}'E^{-1}F-\underline{s}')BU ,$$

where $\underline{s}' = (3/5 \quad 0 \quad 1/3 \quad 0)$. Hence if $\underline{a}_{11}'E^{-1}F = \underline{s}'$, then $\underline{t}_{12}' = \underline{t}_{11}'$. From (6.28) and (6.29) and the definition of $\underline{s}'$, this means the second row of $E^{-1}F$ must equal $\underline{s}'$. From (6.32), (6.45), (6.47) and (6.48), the second row of $E^{-1}F$ is

$$([1^4]/[1^2] \quad 0 \quad [1^2 2^2]/[1^2] \quad 0) .$$

Thus, if $([1^4]/[1^2]) = 3/5$ and $([1^2 2^2]/[1^2]) = 1/3$, then $\underline{t}_{12}' = \underline{t}_{11}'$. Finally, as above, $\underline{t}_{22}' = \underline{t}_{21}'$ if the third row of $E^{-1}F$ is the same as $(0 \quad 1/3 \quad 0 \quad 3/5)$. This will occur if $([2^4]/[2^2]) = 3/5$ and $([1^2 2^2]/[2^2]) = 1/3$. Since this means $\dfrac{[1^2 2^2]}{[1^2]} = \dfrac{[1^2 2^2]}{[2^2]}$, then $[1^2] = [2^2]$, and further, $[2^4] = 3/5\ [2^2] = 3/5\ [1^2] = [1^4]$. Hence result 5 is established.

It is again of interest to note that the least squares minimum bias conditions force $[1^2] = [2^2] = 1/3$, $[1^4] = [2^4] = 1/5$ and $[1^2 2^2] = 1/9$; a special case of (6.55).

For the design moment conditions given in result 5, actual designs are generated for different configurations of levels of the two factors. All the examples will be for symmetric designs and comparisons with the least squares minimum bias designs will be made, if the latter exist. Combinations of factor levels shall be denoted by (ℓ_i, ℓ_j) say, meaning x_1 is at level ℓ_i and x_2 is at level ℓ_j.

Design 1: This is a twelve point configuration made up of four experiments at $(\ell_1, 0)$, $(0, \ell_1)$ $(-\ell_1, 0)$, $(0, -\ell_1)$ and a 3^2 factorial with $-\ell_2$, 0 and ℓ_2 as the levels, with the center point deleted (center points may, however, be added with no change in the resulting design). For this design

(6.56) $$\sum_{u=1}^{12} x_{1u}^2 = \sum_{u=1}^{12} x_{2u}^2 = 2\ell_1^2 + 6\ell_2^2$$

(6.57) $$\sum_{u=1}^{12} x_{1u}^4 = \sum_{u=1}^{12} x_{2u}^4 = 2\ell_1^4 + 6\ell_2^4$$

(6.58) $$\sum_{u=1}^{12} x_{1u}^2 x_{2u}^2 = 4\ell_2^4 ,$$

and conditions (6.55) yield

(6.59) $$4\ell_2^4 = (1/3)(2\ell_1^2 + 6\ell_2^2)$$

(6.60) $$2\ell_1^4 + 6\ell_2^4 = (3/5)(2\ell_1^2 + 6\ell_2^2) .$$

The solution of (6.59) and (6.60) is

(6.61) $$\ell_1 = (3/5)^{1/4} \frac{1}{\sqrt{2}} (1 + \frac{1}{\sqrt{15}})^{1/2}$$

(6.62) $$\ell_2 = \frac{1}{\sqrt{2}} (1 + \frac{1}{\sqrt{15}})^{1/2} = (\frac{5}{3})^{1/4} \ell_1 ,$$

or

$$\ell_1 = .69806757$$

$$\ell_2 = .79315789 .$$

Note that there is no least squares minimum bias design for this configuration. Figure 6.1 shows the design.

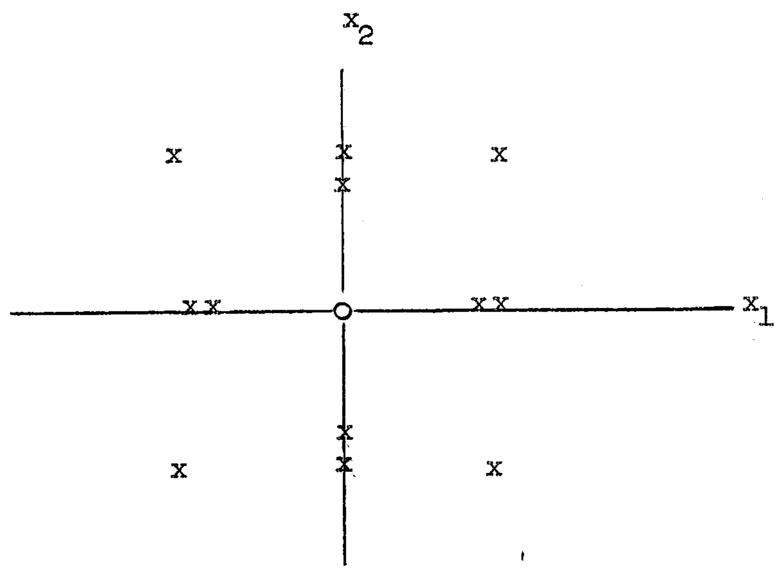

Figure 6.1. Design 1, experimental configuration, $\ell_1 = .698$, $\ell_2 = .793$

Design 2: This is also a twelve point configuration (center points again may or may not be added with no change in the other levels) with eight points at (ℓ_1,ℓ_2), $(-\ell_1,\ell_2)$, $(-\ell_1,-\ell_2)$, $(\ell_1,-\ell_2)$, (ℓ_2,ℓ_1), $(-\ell_2,\ell_1)$, $(-\ell_2,-\ell_1)$ and $(\ell_2,-\ell_1)$, and four at (ℓ_3,ℓ_3), $(-\ell_3,\ell_3)$, $(-\ell_3,-\ell_3)$, $(\ell_3,-\ell_3)$. Figure 6.2 shows a typical configuration.

Here

(6.63) $$\sum_{u=1}^{12} x_{1u}^2 = \sum_{u=1}^{12} x_{2u}^2 = 4(\ell_1^2 + \ell_2^2 + \ell_3^2)$$

(6.64) $$\sum_{u=1}^{12} x_{1u}^4 = \sum_{u=1}^{12} x_{2u}^4 = 4(\ell_1^4 + \ell_2^4 + \ell_3^4)$$

(6.65) $$\sum_{u=1}^{12} x_{1u}^2 x_{2u}^2 = 8\ell_1^2\ell_2^2 + 4\ell_3^2 ,$$

and applying conditions (6.55) yields

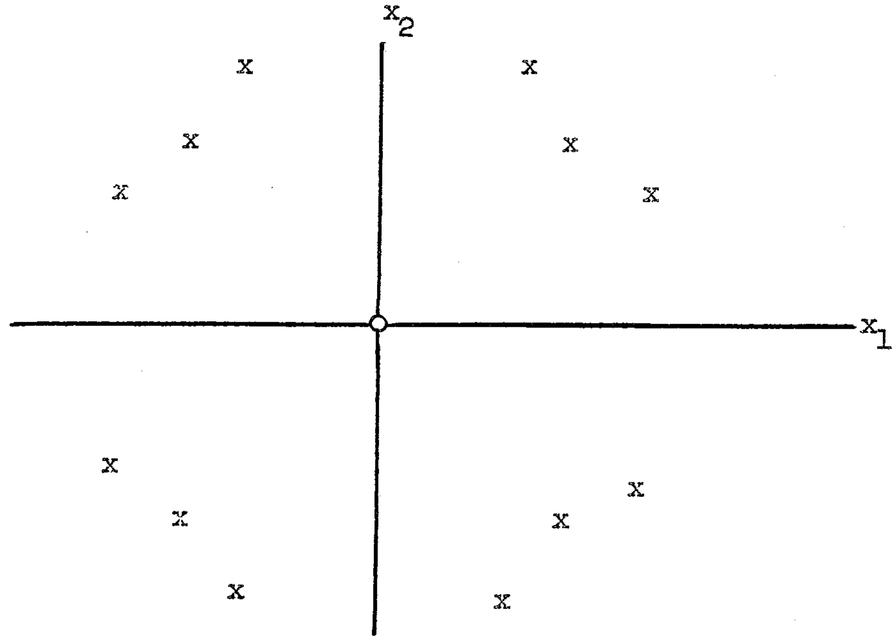

Figure 6.2. Typical design 2, experimental configuration

(6.66) $$4(\ell_1^4 + \ell_2^4 + \ell_3^4) = \frac{12}{5} (\ell_1^2 + \ell_2^2 + \ell_3^2)$$

(6.67) $$8\ell_1^2\ell_2^2 + 4\ell_3^2 = \frac{4}{3} (\ell_1^2 + \ell_2^2 + \ell_3^2) .$$

A class of admissible solutions is defined as those solutions such that ℓ_1, ℓ_2, ℓ_3 are all real and such that $0 \leq \ell_1^2, \ell_2^2, \ell_3^2 \leq 1$. Varying the quantity $(\ell_1^2 - \ell_2^2)$ over the range .50 to .61, and thus ensuring admissible solutions, in steps of .005 yields the designs in Table 6.1. The least squares conditions require $\ell_1 = .47350358$, $\ell_2 = .86058319$, $\ell_3 = .18759247$.

Design 3: Consider a full 4^2 factorial with levels $-\ell_1$, $-\ell_2$, ℓ_2, ℓ_1. Then

(6.68) $$\sum_{u=1}^{16} x_{1u}^2 = \sum_{u=1}^{16} x_{2u}^2 = 8(\ell_1^2 + \ell_2^2)$$

Table 6.1. Designs for the configuration of design 2, and moments

ℓ_1	ℓ_2	ℓ_3	$\sum_{u=1}^{12} x_u^2$	$\sum_{u=1}^{12} x_u^4$	$\sum_{u=1}^{12} x_{1u}^2 x_{2u}^2$
.46658851	.84717462	.04572007	3.750000	2.250000	1.250000
.46884063	.85135864	.10826207	3.825375	2.295225	1.275125
.47097069	.85546092	.14747279	3.901500	2.340900	1.300500
.47296960	.85947673	.17942483	3.978375	2.387025	1.326125
.47482705	.86340068	.20755371	4.056000	2.433600	1.352000
.47653127	.86722664	.23330164	4.134375	2.480625	1.378125
.47806868	.87094756	.25744063	4.213500	2.528500	1.431125
.48057743	.87804024	.30263951	4.374000	2.624400	1.458000
.48150850	.88139119	.32425742	4.455375	2.673225	1.485125
.48219063	.88459471	.34548430	4.537500	2.722500	1.512500
.48259206	.88763455	.36647451	4.620375	2.772225	1.540125
.48267362	.89049078	.38736591	4.704000	2.822400	1.568000
.48238606	.89313846	.40829048	4.788375	2.873025	1.596125
.48166601	.89554573	.42938410	4.873500	2.924100	1.624500
.48042966	.89767068	.45079756	4.959375	2.975625	1.653125
.47856220	.89945638	.47271179	5.046000	3.027600	1.682000
.47589923	.90082189	.49536209	5.133375	3.080025	1.711125
.47219054	.90164511	.51908303	5.221500	3.132900	1.740500
.46702127	.90172549	.54440428	5.310375	3.186225	1.770125
.45960553	.90068710	.57229845	5.400000	3.240000	1.800000
.44804725	.89763374	.60506287	5.490375	3.294225	1.830125
.42315481	.88829049	.65364747	5.581500	3.348900	1.860500

$$(6.69) \qquad \sum_{u=1}^{16} x_{1u}^4 = \sum_{u=1}^{16} x_{2u}^4 = 8(\ell_1^2 + \ell_2^2)$$

$$(6.70) \qquad \sum_{u=1}^{16} x_{1u}^2 x_{2u}^2 = 4(\ell_1^4 + \ell_2^4) + 8\ell_1^2\ell_2^2 \ .$$

Result 5 yields the system

$$(6.71) \qquad \ell_1^4 + \ell_2^4 = \frac{3}{5}(\ell_1^2 + \ell_2^2)$$

$$(6.72) \qquad \ell_1^4 + \ell_2^4 + 2\ell_1^2\ell_2^2 = \frac{2}{3}(\ell_1^2 + \ell_2^2) \ .$$

However (6.72) can be written

$$(6.73) \qquad (\ell_1^2 + \ell_2^2)^2 = \frac{2}{3}(\ell_1^2 + \ell_2^2)$$

or

$$(6.74) \qquad \ell_1^2 + \ell_2^2 = \frac{2}{3} \ ,$$

which means, from (6.71)

$$(6.75) \qquad \ell_1^4 + \ell_2^4 = 2/5 \ .$$

Equations (6.74) and (6.75) are exactly the same conditions as the least squares requirement yields, since least squares requires

$$(6.76) \qquad 8(\ell_1^2 + \ell_2^2) = 16/3$$

$$(6.77) \qquad 8(\ell_1^4 + \ell_2^4) = 16/5$$

$$(6.78) \qquad 4(\ell_1^4 + \ell_2^4) + 8\ell_1^2\ell_2^2 = 16/9 \ ,$$

where (6.78) is simply (6.76) squared. Hence for this 4^2 design the least squares solution is also the non-least squares solution. This solution is

$$\ell_1 = \sqrt{(1/3)\,(1 + \frac{2}{\sqrt{5}})}$$

$$\ell_2 = \sqrt{(1/3)\,(1 - \frac{2}{\sqrt{5}})}$$

or

$$\ell_1 = .79465447$$

$$\ell_2 = .18759248 \ .$$

However, suppose one (or more than one) center point at (0,0) is added to the design. Then the above result will, of course, still satisfy (6.55) but is no longer a least squares solution; i.e. adding a center point (s) makes a least squares design impossible since the least squares equations (for one center point)

(6.79) $$8(\ell_1^2 + \ell_2^2) = 17/3$$

(6.80) $$8(\ell_1^4 + \ell_2^4) = 17/5$$

(6.81) $$4(\ell_1^4 + \ell_2^4) + 8\ell_1^2\ell_2^2 = 17/9$$

are inconsistent.

Design 4: Consider adding a zero level to the 4^2 yielding a 5^2 with $-\ell_1$, $-\ell_2$, 0, ℓ_2, ℓ_1. Then

(6.82) $$\sum_{u=1}^{25} x_{1u}^2 = \sum_{u=1}^{25} x_{2u}^2 = 10(\ell_1^2 + \ell_2^2)$$

$$(6.83) \qquad \sum_{u=1}^{25} x_{1u}^4 = \sum_{u=1}^{25} x_{2u}^4 = 10(\ell_1^4 + \ell_2^4)$$

$$(6.84) \qquad \sum_{u=1}^{25} x_{1u}^2 x_{2u}^2 = 4(\ell_1^4 + \ell_2^4) + 8\ell_1^2\ell_2^2 ,$$

and the conditions to be satisfied are

$$(6.85) \qquad \ell_1^4 + \ell_2^4 = (3/5)(\ell_1^2 + \ell_2^2)$$

$$(6.86) \qquad 2(\ell_1^4 + \ell_2^4) + 4\ell_1^2\ell_2^2 = (5/3)(\ell_1^2 + \ell_2^2) .$$

As for the 4^2 factorial, this design, given by

$$\ell_1 = \sqrt{(1/12)(5 + \sqrt{11})}$$

$$\ell_2 = \sqrt{(1/12)(5 - \sqrt{11})}$$

or

$$\ell_1 = .69305207$$

$$\ell_2 = .14028127 ,$$

coincides with the least squares minimum bias design. However, if an additional center point is included the design satisfies result 5 but there is no least squares design in this case either.

Design 5: Suppose four experiments, at $(\ell_3,0)$, $(0,\ell_3)$, $(-\ell_3,0)$ and $(0,-\ell_3)$ are added to the 5^2 factorial. Then the equations which must be satisfied are

$$(6.87) \qquad 5(\ell_1^4 + \ell_2^4) + \ell_3^4 = (3/5)[5(\ell_1^2 + \ell_2^2) + \ell_3^2]$$

$$(6.88) \qquad 2(\ell_1^4 + \ell_2^4) + 4\ell_1^2\ell_2^2 = (1/3)[5(\ell_1^2 + \ell_2^2) + \ell_3^2] .$$

In order for a solution to represent an admissible design, as previously defined, the solution must satisfy

(6.89) $5/6 \leq \ell_1^2 + \ell_2^2 \leq 1$,

since

$$\ell_3^2 = 3(\ell_1^2 + \ell_2^2)[2(\ell_1^2 + \ell_2^2) - 5/3] .$$

Table 6.2 gives admissible designs, and moments, obtained by varying $\ell_1^2 + \ell_2^2$ from .84 to 1.0 in steps of .01. The least squares design requires $\ell_1^2 + \ell_2^2 = \sqrt{29}/6$ and here $\ell_1 = .85184513$, $\ell_2 = .41459298$, $\ell_3 = .58795918$ is the design.

Example 2: This example is presented to show the effect of a change in the region of interest on the protection design conditions. Consider the same set-up as in the first example with R now being defined by the circle of radius r and center at the origin, so that the region of interest is $0 \leq x_1^2 + x_2^2 \leq r^2$.

The only changes in the estimators for k=1 and k=2 occur in the "A" matrices. In this case

(6.90) $$W_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2/4 & 0 \\ 0 & 0 & r^2/4 \end{bmatrix} ,$$

and when k=1

(6.91) $$W_2 = \begin{bmatrix} r^2/4 & r^2/4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ,$$

so that

Table 6.2. Designs and moments for design 5

ℓ_1	ℓ_2	ℓ_3	$29[1^2]$	$29[1^4]$	$29[1^2 2^2]$
.83569958	.37630600	.18330303	8.4672	5.08032	2.8224
.84001712	.37996216	.29154759	8.6700	5.20200	2.8900
.84373643	.38483912	.37094474	8.8752	5.32512	2.9584
.84683347	.39098906	.43749286	9.0828	5.44968	3.0276
.84927605	.39840958	.49638695	9.2928	5.57568	3.0976
.85102223	.40713778	.55009092	9.5052	5.70312	3.1684
.85201806	.41721125	.60000000	9.7200	5.83200	3.2400
.85219420	.42867825	.64699304	9.9372	5.96232	3.3124
.85146084	.44160440	.69166466	10.1568	6.09408	3.3856
.84969984	.45608134	.73443856	10.3788	6.22728	3.4596
.84675224	.47224002	.77562878	10.6032	6.36192	3.5344
.84239699	.49027269	.81547532	10.8300	6.49800	3.6100
.83631231	.51047205	.85416626	11.0592	6.63552	3.6864
.82799841	.53330914	.89185201	11.2908	6.77448	3.7636
.81659967	.55961146	.92865494	11.5248	6.91488	3.8416
.80039637	.59107161	.96467611	11.7612	7.05672	3.9204
.77459667	.63245553	1.0000000	12.0000	7.20000	4.0000

$$(6.92) \qquad A_1 = \begin{bmatrix} 1 & 0 & 0 & r^2/4 & r^2/4 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$(6.93) \qquad = \begin{bmatrix} \underline{a}'_{01} \\ \underline{a}'_{11} \\ \underline{a}'_{21} \end{bmatrix} .$$

When k=2

$$(6.94) \qquad W_2 = \begin{bmatrix} r^2/4 & r^2/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r^2/8 & 0 & r^4/24 & 0 \\ 0 & 0 & 0 & 0 & r^4/24 & 0 & r^4/8 \end{bmatrix}$$

and therefore,

$$(6.95) \qquad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ r^2/2 & 0 & r^2/6 & 0 \\ 0 & r^2/6 & 0 & r^2/2 \end{bmatrix}$$

$$= \begin{bmatrix} \underline{a}'_{02} \\ \underline{a}'_{12} \\ \underline{a}'_{22} \end{bmatrix} .$$

Then proceeding as in example 1, and using the new "A" matrices, leads to:

Result 6: Sufficient conditions for $T_1' = T_2'$ when R is $0 \le x_1^2 + x_2^2 \le r^2$ are

$$[1^2] = [2^2]$$

$$[1^4] = [2^4]$$

$$[1^4] = \frac{r^2}{2}[1^2]$$

$$[1^2 2^2] = \frac{r^2}{6}[1^2] \quad ,$$

and the X'X matrices in (6.32) and (6.51) are non-singular.

Note that the least squares conditions $[1^2] = [2^2]$, $[1^4] = [2^4]$, $[1^2] = r^2/4$, $[1^4] = r^4/8$, $[1^2 2^2] = r^4/24$ are a special case of (6.96)

7. SUMMARY

The problem of exploring polynomial response surfaces in the presence of assumed model inadequacies has been considered. Over the entire region of factor operability it was assumed that the true response is a polynomial of degree d+k-1. However, within the region of experimental interest it was assumed that a polynomial of degree d-1 would be used as an estimator of the true response. In this setting the important statistical problems which arise are those of designing experiments and choosing estimators for fitting the response surface. It was argued, in Section 2, that optimality criteria, to be useful in attacking these problems, must be meaningful in terms of the underlying problem and also reflect the association between the two problems.

The criteria chosen were 1) minimum squared bias integrated over the region of interest, 2) minimum variance integrated over the region of interest and 3) protection against higher order effects, meaning satisfaction of 1) and 2) for various higher order true models.

Because of the relationship between the design problem and the estimating procedure, it was decided to initially apply the criterion considered most important, the bias criterion, to the estimation problem. Subject to satisfying the bias criterion independently of the unknown parameter values in the true model, estimators are considered optimum if the variance criterion is also satisfied. This led, as developed in Section 3, to the (min V | min B) estimator. This estimator, given the degree of the assumed true model, satisfies the bias and variance requirements for any experimental design. This result enables minimum B, and subject to this, minimum V, to be attained without any restrictions on the design moments. Prior to this outcome minimum B could be reached, but only if certain design moment requirements were met. This is because the previous works confined attention to the least squares estimation procedure, a somewhat questionable criterion in the presence of recognized bias, as discussed in Section 2. Using least squares estimators, the bias criterion can be satisfied only by imposing restrictions on the design. It was shown that this arises as a special case of the more general development.

It is at this stage, i.e. after an estimation procedure has been derived, that additional criteria should be applied to distinguish between different experimental designs. Given that the (min V | min B) estimator is to be used, the protection criterion was introduced in Section 4. When applied to the experimental design problem, this protection criterion yields, by choice of design, the same estimator for different higher order effects in the true model. This procedure involved determining design moment conditions which resulted in the same estimator being used for different higher order models. In Section 4, for one independent variable, design moment conditions were derived for using linear, quadratic and cubic polynomial approximating functions and protecting against certain higher order effects up to quintic. Comparisons with the least squares minimum bias design conditions were made and it was found that greater flexibility for the values of the pertinent moments could be achieved by using the non-least squares estimator. Where, for example, the second moment must always be 1/3 to attain minimum B using least squares, no such restriction was required for the (min V | min B) procedure. Similar results also pertained to higher moments, reflecting greater freedom in the design moment conditions.

In Section 5, using the elementary symmetric functions of the roots of polynomials, actual experimental designs were generated for the design moment conditions derived in Section 4. Some of the difficulties inherent in resolving the resulting systems of non-linear equations into actual experimental designs were evident in this section. Complex valued factor levels and levels lying outside the region of interest are inadmissible as applied to experimental designs. However, for the experimentally meaningful situations, detailed results giving combinations of design levels satisfying the requirements were obtained and displayed in Figure 5.1 and Tables 5.1 and 5.2. This section was of importance in demonstrating how the derivations may be applied to experimental situations.

Generalizations of both the estimator and design conditions were made, in Section 6, for the multi-factor case, with examples of the design problem given for two factors with a linear estimator and protection against quadratic and cubic effects. When the design problem is considered for more than one independent variable the choice of the region of interest and the experimental design configuration both offer additional flexibility. In the example considered, the effect on the moment requirements, of changing the region of interest from a unit cube to a unit sphere was seen by comparing results 5 and 6, in Section 6. For the cube the fourth moment must equal 3/5 of the second moment and the fourth order product moment must be 1/3 the second moment. For the sphere these values were 1/2 and 1/6 respectively. The effect on the problem, of different configurations was considered, and specific designs were tabulated.

In summary then, the development presented in this work has employed an approach to a particular response surface problem involving model inadequacies which enables a certain amount of flexibility over the earlier approaches. The experimenter can guard, in the (min V | min B) sense, against a particular higher order effect by using the appropriate estimator with any experimental design, and can protect against a number of higher order effects by choice of design.

It is hoped that future work in this problem domain will closely relate both the design and estimation problems and relate both these problems to underlying problem objectives. More specifically future work can be suggested in the following areas:

1) Extensions of the derivation of protection conditions on the design moments to multi-factor set-ups involving more than two factors. It is conjectured that the results for two factors will extend "naturally".

2) Relaxation of certain symmetry requirements. With modern computing facilities available it may be desirable to explore the effect of not requiring certain odd moments to vanish. All of the theory and general approaches in this work will still be applicable in such non-symmetric designs.

3) Changing the form of the true model and/or estimator from polynomials to other kinds of functions. In this realm also it would be interesting to change the regions of interest and apply the procedures to different combinations of true models, estimators, and regions.

4) Introduce prior information on the unknown parameters in the true model. Equation (3.36) gives the value of minimum B actually attained, as a quadratic form in these parameters. If it were pertinent, a prior distribution could be attached to these parameters and additional optimality criteria introduced.

5) Applying other forms of bias, variance and protection criteria. Clearly, the procedures and approaches in this work could be used with other criteria for estimators and designs. The literature review, in Section 2, contains many criteria which may be useful and of interest in a particular problem setting.

6) Relating optimality criteria to cope with the wide choice of design configurations. This would be appropriate in the multi-dimensional set-up and for example could lead to certain classifications of designs, based on the criteria used.

7) Weighting functions emphasizing different subspaces of the factor space. This essentially incorporates into the optimality criteria the concept of prior information on the relative importance of different sub-regions in the region of interest.

8) Extend the values of d and k for which designs were actually obtained. This suggestion may be of more academic or numerical interest than of experimental value. The question of convergence of the derived design moment conditions to the least squares requirements for larger values of d and k is included here.

8. LIST OF REFERENCES

Anderson, R. L. 1953. Recent advances in finding best operating conditions. J. Am. Stat. Assoc. 48:789-798.

Box, G. E. P. 1954. The exploration and exploitation of response surfaces: some general considerations and examples. Biometrics 10:16-60.

Box, G. E. P., and N. R. Draper. 1959. A basis for the selection of a response surface design. J. Am. Stat. Assoc. 54:622-654.

Box, G. E. P., and N. R. Draper. 1963. The choice of second order rotatable designs. Biometrika 50:335-352.

Box, G. E. P., and J. S. Hunter. 1957. Multi-factor experimental designs for exploring response surfaces. Ann. Math. Stat. 28:195-241.

Box, G. E. P., and K. B. Wilson. 1951. On the experimental attainment of optimum conditions. J. Roy. Stat. Soc. B13:1-45.

Chernoff, H. 1953. Locally optimal designs for estimating parameters. Ann. Math. Stat. 24:586-602.

David, H. A., and B. E. Arens. 1959. Optimal spacing in regression analysis. Ann. Math. Stat. 30:1072-1081.

de la Garza, A. 1954. Spacing of information in polynomial regression. Ann. Math. Stat. 25:123-130.

Draper, N. R. 1963. Ridge analysis of response surfaces. Technometrics 5:469-479.

Elfving, G. 1952. Optimum allocation in linear regression theory. Ann. Math. Stat. 23:255-262.

Folks, J. L. 1958. Comparisons of designs for exploration of response relationships. Unpublished Ph.D. thesis, Department of Statistics, Iowa State College, Ames, Iowa.

Friedman, M., and L. J. Savage. 1947. Planning experiments seeking maxima, pp. 363-372. In C. Eisenhart, M. W. Hastay, and W. A. Wallis (ed.), Selected Techniques of Statistical Analysis. McGraw-Hill Book Co., Inc., New York.

Hoel, P. G. 1958. Efficiency problems in polynomial regression. Ann. Math. Stat. 29:1134-1145.

Hoel, P. G. 1961. Some properties of optimal spacing in polynomial estimation. Ann. Inst. Stat. Math. (Tokyo) 13:1-8.

Hoel, P. G., and A. Levine. 1964. Optimal spacing and weighting in polynomial prediction. Ann. Math. Stat. 33:1553-1560.

Hotelling, H. 1941. The experimental determination of the maximum of a function. Ann. Math. Stat. 12:20-45.

Householder, A. S. 1953. Principles of Numerical Analysis. McGraw-Hill Book Co., Inc., New York.

Kiefer, J. 1959. Optimum experimental designs. J. Roy. Stat. Soc. (London) B21:272-304.

Kiefer, J. 1961. Optimum experimental designs v, with applications to systematic and rotatable designs, pp. 381-405. In J. Neyman (ed.), Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, University of California Press, Berkeley and Los Angeles.

Kiefer, J., and J. Wolfowitz. 1959. Optimum designs in regression problems. Ann. Math. Stat. 30:271-294.

Lawrence, W. E. 1964. Experimental designs which minimize model inadequacies. Unpublished Ph.D. thesis, Department of Statistics, University of Wisconsin, Madison, Wisconsin.

Manson, A. R. 1966. Minimum bias designs for an exponential response. Oak Ridge National Laboratory, Oak Ridge, Tennessee, ORNL-3873.

9. APPENDICES

9.1 Some Algebra Leading to (4.82)

Equations (4.83) and (4.84) are

$$\frac{[1^4]^2 - [1^2][1^6]}{[1^4] - [1^2]^2} = -3/35$$

$$\frac{[1^6] - [1^2][1^4]}{[1^4] - [1^2]^2} = 6/7 \quad .$$

For each of handling, let $a = [1^2]$, $b = [1^4]$, $c = [1^6]$. Then (4.83) and (4.84) become

(9.1) $$(b^2-ac)/(b-a^2) = -3/35$$

(9.2) $$(c-ab)/(b-a^2) = 6/7 \quad ,$$

which yield

$$-(35/3)(b^2-ac) = (7/6)(c-ab)$$

or

$$ac - c/10 = b^2 - ab/10$$

or upon solving for c,

(9.3) $$c = b(10b-a)/(10a-1) \ ; \quad a \neq 10.$$

Putting (9.3) into (9.1) yields

$$b^2 - ab(10b-a)/(10a-1) = -(3/35)(b-a^2) \ ,$$

which reduces to

(9.4) $$b = (3/35)(10a-1) \quad .$$

Substituting (9.4) in (.93) yields

$$c = (3/35)(10a-1)(10b-a)/(10a-1)$$

$$= (3/35)(10b-a)$$

$$= (3/35)[(6/7)(10a-1)-a] \; ,$$

which becomes

$$(9.5) \qquad c = (3/245)(53a-b).$$

9.2 Some Algebra Leading to (4.104)

As in 9.1, set $d = [1^8]$. Then (4.105) is

$$(9.6) \qquad c^2 - bd + (3/5)(ad-bc) = (3/7)(ac-b^2) \; .$$

This becomes

$$(9.7) \qquad d[(3/5)a-b] = (3/7)(ac-b^2)-c(c-3b/5) \; .$$

Substituting (9.4) and (9.5) into the terms of (9.7) yields

$$(9.8) \qquad (3a/5)-b = (3a/5) = (3/35)(10a-1)$$

$$= -(3/35)(1-3a)$$

$$(9.9) \qquad ac-b^2 = (3a/245)(53a-6) - (3/35)^2(10a-1)^2$$

$$= -(3/35)(a^2 - 6a/7 + 3/35)$$

$$(9.10) \qquad c(c-3b/5) = (3/245)(53a-6)[(3/245)(53a-6) - (9/175)(10a-1)]$$

$$= [9/(245)(1225)] \; (2915a^2 - 807a + 54) \; ,$$

and combining (9.8), (9.9) and (9.10) yields, from (9.7)

$$d = \frac{-(3/7)(3/35)(a^2-6a/7+3/35)-[9/(245)(1225)](2915a^2-807a+54)}{(3/35)(1-3a)}$$

$$= \frac{[9/(7)(1225)]\ (1380a^2-619a+53)}{3a-1}$$

$$= (9/8575)(3a-1)(460a-53)/(3a-1)$$

$$= (9/8575)(460a-53)\ .$$

(9.11)

9.3 Some Algebra Leading to (4.127)

As in 9.1 and 9.2, with $a = [1^2]$, $b = [1^4]$, $c = [1^6]$, $d = [1^8]$, (4.128) and (4.129) are

(9.12) $$(c^2-bd)/(ac-b^2) = -5/21$$

(9.13) $$(ad-bc)/(ac-b^2) = 10/9\ ,$$

which yield

$$-(21/5)(c^2-bd) = (9/10)(ad-bc)$$

or

$$bd-3ad/14 = c^2-3bc/14\ ,$$

and solving for d yields

(9.14) $$d = c(14c-3b)/(14b-3a)\ .$$

Substituting (9.14) in (9.12) yields

$$c^2-bc(14c-3b)/(14b-3a) = -(5/21)(ac-b^2)\ ,$$

which reduces to

(9.15) $$c = (5/63)(14b-3a)\ .$$

Solving for d in terms of a and b, from (9.15) and (9.14) yields

$$d = (5/63)(14b-3a)[(10/9)(14b-3a)-3b]/(14b-3a) ,$$

which reduces to

$$d = (5/567)(113b-30a) . \qquad (9.16)$$