

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Optimum Designs for Identification and Discrimination within a Class of Competing Linear Regression Models

Permalink

<https://escholarship.org/uc/item/0005t018>

Author

Dutta, Santanu

Publication Date

2011

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Optimum Designs for Identification and Discrimination within a
Class of Competing Linear Regression Models

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Applied Statistics

by

Santanu Dutta

August 2011

Dissertation Committee:

Prof. Subir Ghosh, Chairperson

Prof. Barry Arnold

Prof. Bajis Dodin

Copyright by
Santanu Dutta
2011

The Dissertation of Santanu Dutta is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my advisor, Professor Subir Ghosh for being such a supportive advisor and such an excellent teacher. His dedication, commitment, and generosity have been instrumental behind the completion of this dissertation. Not only has he been a great mentor but he has always stretched his helpful hands in any difficulty that I faced in the course of the last five years.

I would also like to thank Professor Barry Arnold of the Statistics department and Professor Bajis Dodin of Operations Management & Management Science for their time and valuable input. I wish to thank all the faculty members of the Statistics department for enriching me with their vast knowledge in different fields of Statistics.

A special word of thanks goes out to all my friends in UCR who have always helped me in many ways during my stay in Riverside.

Finally, I wish to thank my wife Anushree who has always been there with me even if she is thousands of miles away from Riverside. Her support and encouragement have been my greatest source of strength and perseverance for the past few years. I would like to thank my parents, Mr. Saumendra Dutta and Late Mrs. Mamata Dutta who have been my real motivation for pursuing the Ph.D. program in the first place. Every bit of success that I have been able to achieve, I owe it to them. A special word of thanks to my elder brother Sudipta and my elder sister Shweta who have always taken pride in my successes.

To the loving memory of my mother, Mrs. Mamata Dutta.

ABSTRACT OF THE DISSERTATION

Optimum Designs for Identification and Discrimination within a
Class of Competing Linear Regression Models

by

Santanu Dutta

Doctor of Philosophy, Graduate Program in Applied Statistics
University of California, Riverside, August 2011
Prof. Subir Ghosh, Chairperson

We consider the problem of finding optimum designs for model identification and discrimination where the dependence of the response variable Y on an explanatory variable X can be described by at most a third order model. We therefore consider a class that includes all the models up to a maximum of third order with linear, quadratic, and cubic coefficients present. In addition all models have an intercept parameter. A general class of designs with 4 distinct points x_1, x_2, x_3 , and x_4 is considered with replications n_1, n_2, n_3 , and n_4 respectively, satisfying $n_1 + n_2 + n_3 + n_4 = n$ where n is known in advance. While discriminating between two models from the class of models considered, the true model may or may not be one of them. We define the predictive criterion function \mathbf{I} and the fitting criterion function \mathbf{J} . When the functions \mathbf{I} and \mathbf{J} are dependent on more than one model parameters, we define the additional criterion functions \mathbf{K}_I and \mathbf{K}_J . We use the proposed optimality criterion functions to obtain the optimal designs for the model identification and discrimination.

Contents

1	Introduction	1
1.1	The Class of Linear Models	2
1.2	The Discrimination Problem	5
1.3	Optimality Criterion Functions	6
1.3.1	Fitted Value Criterion: J Criterion	6
1.3.2	Predicted Value Criterion: I Criterion	7
1.3.3	T -optimality Criterion	8
1.3.4	K -Criterion	8
1.4	Literature Review	9
2	Linear vs. Special Cubic when the True Model Is Full Cubic	14
2.1	Introduction	14
2.2	Models and Associated Designs	14
2.3	Expression of J Criterion	17
2.4	Efficient Designs with respect to J Criterion	25
2.4.1	Finding the value of a for a given value of b	25
2.4.2	Finding the value of b	26
2.5	Expression of I Criterion	28

2.6	Efficient Designs with respect to \mathbf{I} Criterion	30
2.7	Some Comparisons between \mathbf{I} and \mathbf{J}	31
2.7.1	The $\mathbf{I}(\mathbf{p}_1)$ and $\mathbf{J}(\mathbf{p}_1)$ for Designs in \mathbf{D}_2	32
2.7.2	Exploring the Equality of $\mathbf{I}(p_1)$ and $\mathbf{J}(p_1)$ in \mathbf{D}_2	34
3	Linear vs. Special Cubic as the True Model	38
3.1	Introduction	38
3.2	Models and Associated Designs	38
3.3	Expression of \mathbf{J} Criterion	40
3.4	Expression of \mathbf{T} -Optimality Criterion	42
3.4.1	Two Interesting Observations	43
3.5	Efficient Designs with respect to \mathbf{J} Criterion	44
3.6	Expression of \mathbf{I} Criterion	51
3.7	Efficient Designs with respect to \mathbf{I} Criterion	52
4	Quadratic vs. Full Cubic as the True Model	53
4.1	Introduction	53
4.2	Models and Associated Designs	53
4.3	Expression of \mathbf{J} Criterion	55
4.4	Efficient Designs with respect to \mathbf{J} Criterion	59
4.5	Expression of \mathbf{I} Criterion	59
4.6	Efficient Designs with respect to \mathbf{I} Criterion	61
5	Linear vs. Quadratic when the True Model Is Full Cubic	62
5.1	Introduction	62
5.2	Models and Associated Designs	62
5.3	Expression of \mathbf{J} Criterion	65

5.4	Efficient Designs with respect to \mathbf{J} Criterion	71
5.5	Expression of \mathbf{I} Criterion	72
5.6	Efficient Designs with respect to \mathbf{I} Criterion	74
6	Linear vs. Full Cubic as the True Model	76
6.1	Introduction	76
6.2	Models and Associated Designs	76
6.3	Expression of \mathbf{J} Criterion	78
6.3.1	An Interesting Observation	83
6.4	Expression of \mathbf{K}_J Criterion	84
6.5	Expression of \mathbf{I} Criterion	85
6.6	Expression of \mathbf{K}_I Criterion	87
7	Quadratic vs. Special Cubic when the True Model Is Full Cubic	90
7.1	Introduction	90
7.2	Models and Associated Designs	90
7.3	Expression of \mathbf{J} Criterion	92
7.4	Expression of \mathbf{K}_J Criterion	98
7.5	Expression of \mathbf{I} Criterion	98
8	Conclusion	101
	Bibliography	103
	Appendix A	106
A.1	Maximization of \mathbf{J} for a given b	106
A.2	Maximization of \mathbf{J} for any b	108

Appendix B	109
B.1 Maximization of \mathbf{K}_J for a given b	109
B.2 Maximization of \mathbf{K}_J for any b	111

List of Figures

2.1	Plot of $\mathbf{J}(\mathbf{p}_1)$ against p_1 for different values of b	27
2.2	Plot of $\mathbf{I}(C)$ against C	30
2.3	Plot of $\mathbf{J}(\mathbf{p}_1)$ against p_1	33
2.4	Plot of $\mathbf{I}(p_1)$ against p_1	34
2.5	Plot of $\mathbf{I}(p_1)$ and $\mathbf{J}(p_1)$ against p_1	35
2.6	Plot of $F(p_1)$ against $p_1 \in (0, \frac{1}{2})$	36
5.1	Plot of $\mathbf{I}(B)$ against B	74
6.1	Plots of $\phi(A)$ against A and $\phi(D)$ against D	87

List of Tables

1.1	Presence of γ_i 's in models	4
1.2	Optimal design for estimating γ_{d+1}	9
1.3	Optimal designs for estimating one of the model parameters	10
1.4	Optimal designs for discrimination for different values of d	11
2.1	Some \mathbf{I} -optimal Designs	31
2.2	Some designs in \mathbf{D}_1 for which $\mathbf{I} = \mathbf{J}$	32
3.1	Study of the sign of F	47
3.2	Class of \mathbf{J} -optimal Designs	51
5.1	Some \mathbf{I} -optimal Designs	75
6.1	Finding Maximum of Criterion \mathbf{K}_I	88
6.2	Some \mathbf{K}_I -Optimal Designs in \mathbf{D}_1	89

Chapter 1

Introduction

We consider an experimental situation where the response variable Y is dependent on an explanatory variable X . We do not know the exact dependence of Y on X but we consider a class of possible models to explain this dependence. We assume that the true model is included within the class but we do not know which one is the true model. We consider the identification and discrimination between two competing models within the class when the true model is one of the competing models or a model different from the two competing models. We define the optimality criterion functions to maximize the difference between predicted or fitted values from two competing models. We obtain designs satisfying the criterion functions. Our designs have two aspects: the design points (the values of explanatory variable X) and the number of replications at each design point. We consider different classes of designs with respect to the unequal or different kinds of equal allocations of symmetric design points.

1.1 The Class of Linear Models

We consider an experimental situation where the dependence of the response variable Y on an explanatory variable X can be described by at most a third order model. We therefore consider a class that includes all the models up to a maximum of third order with linear, quadratic, and cubic coefficients present. In addition all models have an intercept parameter. To fit a maximum third order model, we need at least four distinct values of X . We consider that the data are collected at four distinct values of $X = x_1, x_2, x_3$, and x_4 in $[-1, 1]$ with replications n_1, n_2, n_3 , and n_4 respectively so that we can perform the least square fit of a third order model to the data. Let $y_j(x_i)$ represent the j^{th} observation at the i^{th} design point x_i , $j = 1, 2, \dots, n_i$, $i = 1, 2, 3, 4$. We assume that the total number of observations in the experiment is $n = n_1 + n_2 + n_3 + n_4$ which is known in advance and $-1 \leq x_1 < x_2 < x_3 < x_4 \leq 1$. We assume the model

$$\begin{aligned} E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \gamma_3 x_i^3, \\ \text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0, \end{aligned} \quad (1.1)$$

where $i, i' = 1, 2, 3, 4$; $j = 1, 2, \dots, n_i, j' = 1, 2, \dots, n_{i'}$, $(i, j) \neq (i', j')$, γ_u for $u = 0, 1, 2, 3$ are fixed unknown parameters, γ_0 is non-zero and at least one of γ_1, γ_2 , and γ_3 is non-zero, and σ^2 is unknown. We denote

$$\boldsymbol{\gamma} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} \quad \text{and} \quad \mathbf{X} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}'.$$

Thus the matrix representation of the model in (1.1) is given by

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\gamma}, \quad Var(\mathbf{y}) = \sigma^2\mathbf{I}. \quad (1.2)$$

The least square estimator (Rao (1973)) of $\boldsymbol{\gamma}$ is

$$\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_0 \quad \hat{\gamma}_1 \quad \hat{\gamma}_2 \quad \hat{\gamma}_3)' = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}.$$

The predicted value of y at a new x is given by $\hat{y}(x) = \hat{\gamma}_0 + \hat{\gamma}_1x + \hat{\gamma}_2x^2 + \hat{\gamma}_3x^3$. When we consider the predicted value $\hat{y}_j(x)$ at $x = x_i$, we get the fitted value from the model for the j^{th} observation at the i^{th} design point where $j = 1, 2, \dots, n_i$; $i = 1, 2, 3, 4$. In fact we get four distinct fitted values corresponding to the four distinct design points x_1, x_2, x_3 , and x_4 and they are repeated n_1, n_2, n_3 , and n_4 times respectively. Hence, when we consider the fitted values of all the observations, we get a vector of length n of the fitted values. Let the fitted values of y be given by $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\gamma}}$.

We consider the class of designs \mathbf{D} that consists of x_1, x_2, x_3 , and x_4 with their replications n_1, n_2, n_3 , and n_4 respectively with \mathbf{X} as a full rank matrix so that we can fit a third order model to the data. Mathematically, we denote the class of designs by

$$\mathbf{D} = \left\{ (x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : -1 \leq x_1 < x_2 < x_3 < x_4 \leq 1; \right. \\ \left. n_1 + n_2 + n_3 + n_4 = n, Rank(\mathbf{X}) = 4 \right\}. \quad (1.3)$$

We tabulate all possible models in (1.1) in Table-1.1 in which 0 and 1 represent

respectively absence and presence of a parameter in a model.

Table 1.1: Presence of γ_i 's in models

Model	γ_0	γ_1	γ_2	γ_3
1	1	1	0	0
2	1	0	1	0
3	1	0	0	1
4	1	1	1	0
5	1	1	0	1
6	1	0	1	1
7	1	1	1	1

We now pick any pair of models u_1 and u_2 where $u_1, u_2 = 1(1)7$ and $u_1 \neq u_2$. We denote them by MI and MII, MII being the higher order model and MI being the lower order model. For example if $u_1 = 1$ and $u_2 = 4$, then MI has $\gamma_2 = \gamma_3 = 0$ and MII has $\gamma_3 = 0$. Thus MI is a simple linear regression model and MII is a quadratic regression model. We assume that model 7 i.e. the full cubic model is always the unknown true model. Although we assume model 7 to be the unknown true model, we are interested in discriminating between any pair of models that we encounter in Table-1.1.

While discriminating between two models from the class of models considered in Table-1.1, the true model may or may not be one of them. We define the predicted value criterion function \mathbf{I} and the fitted value criterion function \mathbf{J} . When the functions \mathbf{I} and \mathbf{J} are dependent on more than one model parameters, we define the additional criterion functions \mathbf{K}_I and \mathbf{K}_J . We use the proposed optimality criterion functions to obtain the optimal designs for the model identification and discrimination. Some important pairs of models from Table-1.1 are chosen in different chapters and optimum designs are obtained for model identification and

discrimination.

Chapter-2 presents 1 vs. 5, assuming $\gamma_2 \neq 0$ in the true model 7. Chapter-3 represents 1 vs. 5 assuming $\gamma_2 = 0$ in the true model 7. Chapter-4 represents 4 vs. 7 (the true model). Chapter-5 represent 1 vs. 4 assuming $\gamma_3 \neq 0$ in the true model 7. Chapter-6 represents 1 vs. 7 (the true model). Chapter-7 represents 4 vs. 5 assuming 7 to be the true model.

1.2 The Discrimination Problem

Suppose the response variable Y is dependent on an explanatory variable X with two possible dependence as described by two simple linear regression models MI and MII. The exact nature of dependence is not known in advance but it is known though that one of MI and MII possibly describes adequately the dependence of Y on X . In the process of determination of the better model we have to discriminate one model from the other model in better describing the data. We consider the least square fit of a model to the data. At first we ensure that the design considered must identify all the models or in other words, the design considered must permit the least square fit for all models. The discrimination between two models MI and MII requires that the fitted values $\hat{y}^{(1)}$ and $\hat{y}^{(2)}$ under the two models should be different from each other. The model discrimination also requires that the predicted values $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ should be different from each other at the possible values of X . Here we consider the model identification and discrimination at the design stage of the experiment. Our goal is to determine the values of X from a given range judiciously. We also find the number of replications for the chosen values of X . These values of X together their replications have the ability of model

identification and discrimination. A design is called optimum within the class of all such available designs if its $\hat{\mathbf{y}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)}$ values have the maximum difference among the corresponding differences for other designs or its $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ values have the maximum difference for all possible values of x in some overall sense as described later.

1.3 Optimality Criterion Functions

We consider several optimality criterion functions and provide a brief discussion on them in this section.

1.3.1 Fitted Value Criterion: J Criterion

Let $\hat{\mathbf{y}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)}$ be the two vectors of fitted values of the two models MI and MII respectively. If the two vectors $\hat{\mathbf{y}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)}$ are very close to each other, it is difficult to discriminate between the two models. Therefore we maximize the difference between these two vectors to get the efficient design for discrimination purpose. As we discriminate between the two models at the design stage, we do not have any observations. So, we work with the expectations of the fitted values where we calculate $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$ under the assumed true model. But $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$ is a vector quantity with different elements within the vector. We cannot find a design that maximizes all the distinct elements of the vector simultaneously. We consider the criterion function $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})' E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$ which actually measures the square of the norm between the two vectors $E(\hat{\mathbf{y}}^{(1)})$ and $E(\hat{\mathbf{y}}^{(2)})$. We obtain an efficient design by maximizing this criterion function. We denote this criterion (Ghosh and Pal (2008)) by J where,

$$\mathbf{J} = E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})' E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)}).$$

1.3.2 Predicted Value Criterion: \mathbf{I} Criterion

Let $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ be the predicted values of the two models MI and MII respectively at $X = x$. We want these two values to be as far apart as possible. We calculate $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ under the assumed true model. Therefore, maximizing $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ would give us an efficient design for the discrimination between MI and MII. But this quantity depends on x and it is positive for some values of x , negative for some other values of x . Hence we consider the square of this quantity. We are interested in this quantity for some special values of x as well as for all possible values of x . When we consider the quantity with respect to all possible values of x we get an overall prediction measure for discrimination purpose. We find the overall measure by taking the weighted average of $[E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2$ on x . We denote this criterion by \mathbf{I} which is given by $\mathbf{I} = \int_{-1}^1 [E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 w(x) dx$,

where $\int_{-1}^1 w(x) dx = 1$. We assume equal weights throughout the region of x i.e.

$$\text{we consider } w(x) = \begin{cases} \frac{1}{2} & -1 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then we maximize \mathbf{I} with respect to the design points and their corresponding replications to get the most efficient design under this criterion. This criterion is a modified version of the Ghosh and Pal (2008) Predicted Value criterion in the sense that it is considering the squared difference, $[E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2$, instead of the absolute difference, $|E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))|$.

1.3.3 *T*-optimality Criterion

If model MII is the true model, the experiment should be designed to yield the maximum possible value of the lack of fit sum of squares (Atkinson and Fedorov (1975a)) and (Atkinson and Fedorov (1975b)) of model MI,

i.e. $SS_{LOF} = (\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})'(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$ should be maximum. But we do not have observations because we want to discriminate between these two models even before the data are collected. Therefore we consider the expected values of SS_{LOF} . As MI is linear in parameters here, $E(SS_{LOF})$ becomes proportional to the non-centrality parameter (***NCP***) of the χ^2 -distribution associated with the SS_{LOF} (Atkinson and Fedorov (1975a)). Now, considering the testing of hypotheses

H_0 : No lack of fit in MI vs. H_a : Lack of fit in MI,

we note that the power of the F-test for lack of fit of MI from MII is an increasing function of the ***NCP*** and thus the design should maximize the ***NCP*** to maximize the power of the test. A design which maximizes the ***NCP*** is called a *T*-optimal design.

1.3.4 *K*-Criterion

Sometimes the expression of ***J*** or ***I*** criterion is observed as a quadratic form in more than one unknown model parameters. Hence is it not possible to directly optimize ***J*** or ***I*** without the information on the model parameters. In that situation we propose a new criterion function ***K*** as the determinant of the matrix associated with the quadratic form. A design is said to be optimum with respect to criterion ***K*** if it maximizes the determinant of the matrix associated with the quadratic form for all designs in a class. Whenever we observe expression of ***J*** or ***I*** as a quadratic form in more than one unknown model parameters, we use criterion ***K*** to optimize

\mathbf{J} or \mathbf{I} . We denote them by \mathbf{K}_J and \mathbf{K}_I .

1.4 Literature Review

In the pioneering paper of Kiefer and Wolfowitz (1959), the optimum regression design was obtained for estimating one of the model parameters. Consider the regression model

$$E(y(x)) = \gamma_0 + \gamma_1x + \gamma_2x^2 + \dots + \gamma_{d+1}x^{d+1}.$$

They presented optimal designs to estimate one of the $(d+2)$ parameters optimally. The explanatory variable x is scaled such that $-1 \leq x \leq 1$. The proportion of observations taken at any point x is denoted by $\xi(x)$ where ξ is a probability distribution over $[-1, 1]$. The optimal design for estimating γ_{d+1} is given in Table 1.2. They used the *Chebyshev approximation* to find this optimal design. This

Table 1.2: Optimal design for estimating γ_{d+1}

x	$\xi(x)$
-1	$\frac{1}{2(d+1)}$
1	$\frac{1}{2(d+1)}$
$\cos\left(\frac{\pi}{d+1}\right)$	$\frac{1}{d+1}$
$\cos\left(\frac{2\pi}{d+1}\right)$	$\frac{1}{d+1}$
\vdots	\vdots
$\cos\left(\frac{d\pi}{d+1}\right)$	$\frac{1}{d+1}$

Chebyshev approximation approximates the term x^{d+1} with a polynomial in x of degree d and gives the $(d + 2)$ design points. In the next step, the weights of these $(d + 2)$ design points are obtained by solving $(d + 1)$ simultaneous linear equations. In the direct method, the variance of the best linear estimate of γ_{h+1} , which is a non-linear function of $(2d + 3)$ variables, has to be minimized. Therefore, the method given by them addresses to the computational challenges in finding the optimal design for this problem. We illustrate this further by giving designs for $d = 0, 1, 2$ in Table 1.3.

Table 1.3: Optimal designs for estimating one of the model parameters

d	Model	To Estimate	x	$\xi(x)$
0	$E(y) = \gamma_0 + \gamma_1x$	γ_1	-1 1	$\frac{1}{2}$ $\frac{1}{2}$
1	$E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2$	γ_2	-1 0 1	$\frac{1}{4}$ $\frac{1}{2}$ $\frac{1}{4}$
2	$E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2 + \gamma_3x^3$	γ_3	-1 $-\frac{1}{2}$ $\frac{1}{2}$ 1	$\frac{1}{6}$ $\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{6}$

Atkinson (1972) used the same optimal design for discriminating between the two models

$$E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2 + \dots + \gamma_dx^d,$$

and

$$E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2 + \dots + \gamma_dx^d + \gamma_{d+1}x^{d+1}.$$

The optimality criterion used for the discrimination purpose is to estimate γ_{d+1} with minimum variance. When $d = 1$, the problem reduces to discriminating between two simple linear regression models

$$E(y) = \gamma_0 + \gamma_1x \quad \text{and} \quad E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2$$

and the optimal design is $x_1 = -1, x_2 = 0, x_3 = 1, n_1 = n_3 = \frac{n}{4}$ and $n_2 = \frac{n}{2}$.

In Table-1.4 we give designs of this kind for $d = 0, 1, 2$.

Table 1.4: Optimal designs for discrimination for different values of d

d	Models	x_i 's	n_i 's
0	$E(y) = \gamma_0$	$x_1 = -1$	$n_1 = \frac{n}{2}$
	$E(y) = \gamma_0 + \gamma_1x$	$x_2 = 1$	$n_2 = \frac{n}{2}$
1	$E(y) = \gamma_0 + \gamma_1x$	$x_1 = -1$	$n_1 = \frac{n}{4}$
	$E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2$	$x_2 = 0$	$n_2 = \frac{n}{2}$
		$x_3 = 1$	$n_3 = \frac{n}{4}$
2	$E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2$	$x_1 = -1$	$n_1 = \frac{n}{6}$
		$x_2 = -\frac{1}{2}$	$n_2 = \frac{n}{3}$
	$E(y) = \gamma_0 + \gamma_1x + \gamma_2x^2 + \gamma_3x^3$	$x_3 = \frac{1}{2}$	$n_3 = \frac{n}{3}$
		$x_4 = 1$	$n_4 = \frac{n}{6}$

Atkinson (1972) further derived optimal designs for models involving more than one factor. Atkinson and Cox (1974) developed optimal designs for discriminating

between several (more than two) regression models extending the same idea.

Atkinson and Fedorov (1975a) proposed a locally optimum criterion called the T -optimality for discriminating between two rival models MI and MII. They assumed one of the models to be the true model and then maximized the noncentrality parameter of the χ^2 -distribution associated with the lack of fit sum of squares. They proposed two approaches: non-sequential designs and sequential designs, the non-sequential designs being described as the limits to which the sequential designs converge asymptotically.

Lauter (1974), Dette (1991, 1994, 1995), and Spruill (1990) also worked on this problem of discrimination using different optimality criteria. All these designs are classified as non-sequential in the sense that all observations are taken at one stage. The second approach is based on sequential methods. The observations are taken sequentially for the purpose of discrimination. This approach was followed by Atkinson and Cox (1974), Atkinson and Fedorov (1975a), Andrews (1971), Montepiedra and Yeh (1998), and many other authors. Biswas and Chaudhuri (2002) presented sequential designs for discriminating between two rival models MI and MII. They considered different mixtures of the D -optimum design for MI and the D -optimum design for MII at different stages.

Atkinson (2008) proposed DT -optimal designs for providing a balance between model discrimination and parameter estimation by taking a convex combination of D -criterion and T -criterion. Pukelsheim and Rosenberger (1993) presented designs considering several objectives simultaneously for discriminating between a second order and a third order polynomial models. They have used D -criterion and geometric mixtures of D -criteria. Dette and Kwiecien (2004) compared non-sequential designs with sequential designs to demonstrate that non-sequential designs provide

better model identification than the sequential designs.

Ghosh and Pal (2008) considered the issue of discriminating between a linear and a quadratic regression models. They proposed two criteria associated with the fitted and predicted observations. They showed that their design performs better than Kiefer-Wolfowitz design (Kiefer and Wolfowitz (1959)) under their predicted value criterion. They also evaluated the performance Biswas-Chaudhuri design (Biswas and Chaudhuri (2002)) with respect to the predicted value criterion.

Dette and Titoff (2009) derived various new properties of T -optimal designs, which in many circumstances allow an explicit determination of T -optimal designs. They also showed that in many cases T -optimal designs are not unique, and in that situation they gave a characterization of all T -optimal designs. Finally, they compared T -optimal designs with D -optimal designs using a simulation study.

Chapter 2

Linear vs. Special Cubic when the True Model Is Full Cubic

2.1 Introduction

We consider an experiment where the response variable Y is dependent on an explanatory variable X . We assume that the full cubic model MT is the unknown true model but at the same time we would like to have the ability to discriminate between two possible dependence as described by two models MI, a simple linear regression model and MII, a cubic regression model without the quadratic coefficient. So, our goal is to discriminate between these two models MI and MII at the design stage assuming MT to be the true model.

2.2 Models and Associated Designs

We consider the class of designs D in (1.3). We also consider the full cubic model from (1.1) as the true model and denote it by MT. Our aim is to discriminate

between the two models MI and MII assuming MT to be the true model.

The three models considered here are given by

$$\begin{aligned}
 \mathbf{MT:} \quad & E(y_j(x_i)) = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \gamma_3 x_i^3, \\
 & \text{Var}((y_j(x_i))) = \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
 \end{aligned} \tag{2.1}$$

$$\begin{aligned}
 \mathbf{MI:} \quad & E(y_j(x_i)) = \gamma_0 + \gamma_1 x_i, \\
 & \text{Var}((y_j(x_i))) = \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
 \end{aligned} \tag{2.2}$$

$$\begin{aligned}
 \mathbf{MII:} \quad & E(y_j(x_i)) = \gamma_0 + \gamma_1 x_i + \gamma_3 x_i^3, \\
 & \text{Var}((y_j(x_i))) = \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
 \end{aligned} \tag{2.3}$$

where $i, i' = 1, 2, 3, 4$; $j = 1, 2, \dots, n_i$, $j' = 1, 2, \dots, n_{i'}$, $(i, j) \neq (i', j')$.

The matrix representations of MT, MI, and MII are given by

$$E(\mathbf{y}) = \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{2.4}$$

$$E(\mathbf{y}) = \mathbf{X}^{(1)}\boldsymbol{\gamma}^{(1)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{2.5}$$

$$E(\mathbf{y}) = \mathbf{X}^{(2)}\boldsymbol{\gamma}^{(2)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{2.6}$$

where

$$\mathbf{X}^{(1)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(1)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix},$$

$$\mathbf{X}^{(2)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(2)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_3 \end{pmatrix},$$

and

$$\mathbf{X}^{(t)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(t)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_3 \\ \gamma_2 \end{pmatrix}.$$

We define

$$\mathbf{X}_2 = \begin{pmatrix} x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}',$$

$$\mathbf{X}_3 = \begin{pmatrix} x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}',$$

and

$$\mathbf{X}_{32} = \begin{pmatrix} x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}'.$$

Hence we get,

$$\mathbf{X}^{(2)} = \begin{pmatrix} \mathbf{X}^{(1)} & \vdots & \mathbf{X}_3 \end{pmatrix} \Rightarrow \mathbf{X}^{(2)'}\mathbf{X}^{(2)} = \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_3 \\ \mathbf{X}_3'\mathbf{X}^{(1)} & \mathbf{X}_3'\mathbf{X}_3 \end{pmatrix},$$

$$\begin{aligned}\mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(2)} & \vdots & \mathbf{X}_2 \end{pmatrix} \Rightarrow \mathbf{X}^{(t)'}\mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{X}^{(2)'}\mathbf{X}^{(2)} & \mathbf{X}^{(2)'}\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}^{(2)} & \mathbf{X}_2'\mathbf{X}_2 \end{pmatrix}, \\ \mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(1)} & \vdots & \mathbf{X}_{32} \end{pmatrix} \Rightarrow \mathbf{X}^{(t)'}\mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{32} \\ \mathbf{X}_{32}'\mathbf{X}^{(1)} & \mathbf{X}_{32}'\mathbf{X}_{32} \end{pmatrix}.\end{aligned}\quad (2.7)$$

2.3 Expression of J Criterion

We first consider the fitted value criterion J (ref. Chapter-1). Our goal is to obtain the efficient design within the class of designs \mathbf{D} in (1.3) for model selection and discrimination purposes. Let $\hat{\mathbf{y}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)}$ be the fitted values of the two models MI and MII respectively. First we have to find the expression of $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$ where the expectation is considered under the true model MT in (2.1). Then we have to find the expression of $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})'E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$.

The least square estimate of $\boldsymbol{\gamma}^{(1)}$ for MI is given by (Rao (1973))

$$\hat{\boldsymbol{\gamma}}^{(1)} = \begin{pmatrix} \hat{\gamma}_0^{(1)} \\ \hat{\gamma}_1^{(1)} \end{pmatrix} = (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{y}, \quad (2.8)$$

and the least square estimate of $\boldsymbol{\gamma}^{(2)}$ for MII is given by

$$\hat{\boldsymbol{\gamma}}^{(2)} = \begin{pmatrix} \hat{\gamma}_0^{(2)} \\ \hat{\gamma}_1^{(2)} \\ \hat{\gamma}_3^{(2)} \end{pmatrix} = (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{y}. \quad (2.9)$$

Now,

$$\begin{aligned}
E(\hat{\gamma}^{(1)}) &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}E(\mathbf{y}) \\
&= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\
&= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{32} \end{pmatrix}\boldsymbol{\gamma}^{(t)} \\
&= \begin{pmatrix} \mathbf{I}_2 & (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}_{32} \end{pmatrix}\boldsymbol{\gamma}^{(t)} \\
&= \begin{pmatrix} 1 & 0 & A & B \\ 0 & 1 & C & D \end{pmatrix}\boldsymbol{\gamma}^{(t)}, \tag{2.10}
\end{aligned}$$

and

$$\begin{aligned}
E(\hat{\gamma}^{(2)}) &= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}E(\mathbf{y}) \\
&= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\
&= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\begin{pmatrix} \mathbf{X}^{(2)'}\mathbf{X}^{(2)} & \mathbf{X}^{(2)'}\mathbf{X}_2 \end{pmatrix}\boldsymbol{\gamma}^{(t)} \\
&= \begin{pmatrix} \mathbf{I}_3 & (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{X}_2 \end{pmatrix}\boldsymbol{\gamma}^{(t)} \\
&= \begin{pmatrix} 1 & 0 & 0 & E \\ 0 & 1 & 0 & F \\ 0 & 0 & 1 & G \end{pmatrix}\boldsymbol{\gamma}^{(t)}, \tag{2.11}
\end{aligned}$$

where

$$A = \frac{1}{\text{Det}_1} \left[\sum_1^4 n_i x_i^2 \sum_1^4 n_i x_i^3 - \sum_1^4 n_i x_i \sum_1^4 n_i x_i^4 \right],$$

$$B = \frac{1}{\text{Det}_1} \left[\left(\sum_1^4 n_i x_i^2 \right)^2 - \sum_1^4 n_i x_i \sum_1^4 n_i x_i^3 \right],$$

$$C = \frac{1}{\text{Det}_1} \left[\sum_1^4 n_i \sum_1^4 n_i x_i^4 - \sum_1^4 n_i x_i \sum_1^4 n_i x_i^3 \right],$$

$$D = \frac{1}{\text{Det}_1} \left[\sum_1^4 n_i \sum_1^4 n_i x_i^3 - \sum_1^4 n_i x_i \sum_1^4 n_i x_i^2 \right],$$

$$\begin{aligned} \text{Det}_1 &= |\mathbf{X}^{(1)'} \mathbf{X}^{(1)}| \\ &= \left[n_1 n_2 (x_1 - x_2)^2 + n_1 n_3 (x_1 - x_3)^2 + n_1 n_4 (x_1 - x_4)^2 \right. \\ &\quad \left. + n_2 n_3 (x_2 - x_3)^2 + n_2 n_4 (x_2 - x_4)^2 + n_3 n_4 (x_3 - x_4)^2 \right], \end{aligned}$$

$$\begin{aligned} E &= \frac{-1}{\text{Det}_2} \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) x_1 x_2 x_3 \right. \\ &\quad + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) x_1 x_2 x_4 \\ &\quad + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) x_1 x_3 x_4 \\ &\quad \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) x_2 x_3 x_4 \right], \end{aligned}$$

$$\begin{aligned} F &= \\ &= \frac{1}{\text{Det}_2} \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) (x_1 x_2 + x_1 x_3 + x_2 x_3) \right. \\ &\quad + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) (x_1 x_2 + x_1 x_4 + x_2 x_4) \\ &\quad + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) (x_1 x_3 + x_1 x_4 + x_3 x_4) \\ &\quad \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) (x_2 x_3 + x_2 x_4 + x_3 x_4) \right], \end{aligned}$$

$$\begin{aligned}
G = \frac{1}{\text{Det}_2} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) \right],
\end{aligned}$$

$$\begin{aligned}
\text{Det}_2 & = |\mathbf{X}^{(2)'} \mathbf{X}^{(2)}| \\
& = \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3)^2 \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4)^2 \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4)^2 \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4)^2 \right].
\end{aligned}$$

We know that the fitted values of \mathbf{y} under MI and MII are expressed as $\hat{\mathbf{y}}^{(1)} = \mathbf{X}^{(1)} \hat{\boldsymbol{\gamma}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)} = \mathbf{X}^{(2)} \hat{\boldsymbol{\gamma}}^{(2)}$. Now assuming MT to be the true model the expected fitted values are given by

$$\begin{aligned}
E(\hat{\mathbf{y}}^{(1)}) & = \mathbf{X}^{(1)} E(\hat{\boldsymbol{\gamma}}^{(1)}) \\
& = \mathbf{X}^{(1)} (\mathbf{X}^{(1)'} \mathbf{X}^{(1)})^{-1} \mathbf{X}^{(1)'} \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\
& = \mathbf{H}_1 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \tag{2.12}
\end{aligned}$$

$$\begin{aligned}
E(\hat{\mathbf{y}}^{(2)}) & = \mathbf{X}^{(2)} E(\hat{\boldsymbol{\gamma}}^{(2)}) \\
& = \mathbf{X}^{(2)} (\mathbf{X}^{(2)'} \mathbf{X}^{(2)})^{-1} \mathbf{X}^{(2)'} \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\
& = \mathbf{H}_2 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \tag{2.13}
\end{aligned}$$

$$\Rightarrow E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)}) = (\mathbf{H}_1 - \mathbf{H}_2)\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \quad (2.14)$$

where $\mathbf{H}_1 = \mathbf{X}^{(1)}(\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}$ and $\mathbf{H}_2 = \mathbf{X}^{(2)}(\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}$. We note that \mathbf{H}_1 and \mathbf{H}_2 are symmetric and idempotent (Rao (1973)) i.e. $\mathbf{H}_1' = \mathbf{H}_1$, $\mathbf{H}_1^2 = \mathbf{H}_1$, $\mathbf{H}_2' = \mathbf{H}_2$ and $\mathbf{H}_2^2 = \mathbf{H}_2$.

Result 1. : $\mathbf{H}_2\mathbf{H}_1 = \mathbf{H}_1 = \mathbf{H}_1\mathbf{H}_2$

Proof:

$$\begin{aligned} \mathbf{H}_2\mathbf{H}_1 &= \mathbf{X}^{(2)}(\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{H}_1 \\ &= \mathbf{X}^{(2)} \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_3 \\ \mathbf{X}_3'\mathbf{X}^{(1)} & \mathbf{X}_3'\mathbf{X}_3 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} \\ \mathbf{X}_3'\mathbf{X}^{(1)} \end{pmatrix} (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'} \\ &= (\mathbf{X}^{(1)} \quad \mathbf{X}_3) \begin{pmatrix} \mathbf{I} \\ \mathbf{0} \end{pmatrix} (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'} \\ &= \mathbf{H}_1. \end{aligned}$$

Now

$$\begin{aligned} (\mathbf{H}_2\mathbf{H}_1)' &= \mathbf{H}_1' \\ \Leftrightarrow \mathbf{H}_1'\mathbf{H}_2' &= \mathbf{H}_1 \\ \Leftrightarrow \mathbf{H}_1\mathbf{H}_2 &= \mathbf{H}_1. \end{aligned}$$

Thus we have $\mathbf{H}_2\mathbf{H}_1 = \mathbf{H}_1 = \mathbf{H}_1\mathbf{H}_2$.

Clearly using (2.14) we get

$$\begin{aligned}
\mathbf{J} &= E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})' E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)}) \\
&= \boldsymbol{\gamma}^{(t)'} \mathbf{X}^{(t)'} (\mathbf{H}_1 - \mathbf{H}_2)^2 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\
&= \boldsymbol{\gamma}^{(t)'} \mathbf{X}^{(t)'} (\mathbf{H}_2 - \mathbf{H}_1) \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \quad \text{using Result-1.} \tag{2.15}
\end{aligned}$$

Now

$$\begin{aligned}
\mathbf{X}^{(t)'} \mathbf{H}_2 \mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(2)'} \\ \mathbf{X}_2' \end{pmatrix} \mathbf{H}_2 \begin{pmatrix} \mathbf{X}^{(2)} & \mathbf{X}_2 \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{X}^{(2)'} \\ \mathbf{X}_2' \mathbf{H}_2 \end{pmatrix} \begin{pmatrix} \mathbf{X}^{(2)} & \mathbf{X}_2 \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{X}^{(2)'} \mathbf{X}^{(2)} & \mathbf{X}^{(2)'} \mathbf{X}_2 \\ \mathbf{X}_2' \mathbf{X}^{(2)} & \mathbf{X}_2' \mathbf{H}_2 \mathbf{X}_2 \end{pmatrix} \\
&= \begin{pmatrix} \sum_1^4 n_i & \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^2 \\ \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^4 & \sum_1^4 n_i x_i^3 \\ \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 & \sum_1^4 n_i x_i^6 & \sum_1^4 n_i x_i^5 \\ \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^5 & S \end{pmatrix}, \tag{2.16}
\end{aligned}$$

$$\begin{aligned}
\mathbf{X}^{(t)'} \mathbf{H}_1 \mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(1)'} \\ \mathbf{X}_{32}' \end{pmatrix} \mathbf{H}_1 \begin{pmatrix} \mathbf{X}^{(1)} & \mathbf{X}_{32} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{X}^{(1)'} \\ \mathbf{X}_{32}' \mathbf{H}_1 \end{pmatrix} \begin{pmatrix} \mathbf{X}^{(1)} & \mathbf{X}_{32} \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{X}^{(t)'}\mathbf{H}_1\mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{32} \\ \mathbf{X}_{32}'\mathbf{X}^{(1)} & \mathbf{X}_{32}'\mathbf{H}_1\mathbf{X}_{32} \end{pmatrix} \\
&= \begin{pmatrix} \sum_1^4 n_i & \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^2 \\ \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^4 & \sum_1^4 n_i x_i^3 \\ \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 & \mathbf{X}_3'\mathbf{H}_1\mathbf{X}_3 & \mathbf{X}_3'\mathbf{H}_1\mathbf{X}_2 \\ \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 & \mathbf{X}_2'\mathbf{H}_1\mathbf{X}_3 & \mathbf{X}_2'\mathbf{H}_1\mathbf{X}_2 \end{pmatrix} \\
&= \begin{pmatrix} \sum_1^4 n_i & \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^2 \\ \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^4 & \sum_1^4 n_i x_i^3 \\ \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 & P & R \\ \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 & R & Q \end{pmatrix}, \tag{2.17}
\end{aligned}$$

where

$$P = \mathbf{X}_3'\mathbf{H}_1\mathbf{X}_3 = A \sum_1^4 n_i x_i^3 + C \sum_1^4 n_i x_i^4,$$

$$Q = \mathbf{X}_2'\mathbf{H}_1\mathbf{X}_2 = B \sum_1^4 n_i x_i^2 + D \sum_1^4 n_i x_i^3,$$

$$R = \mathbf{X}_3'\mathbf{H}_1\mathbf{X}_2 = A \sum_1^4 n_i x_i^2 + C \sum_1^4 n_i x_i^3,$$

$$S = \mathbf{X}_2'\mathbf{H}_2\mathbf{X}_2 = E \sum_1^4 n_i x_i^2 + F \sum_1^4 n_i x_i^3 + G \sum_1^4 n_i x_i^5,$$

and $A, B, C, D, E, F,$ and G are defined right after (2.11).

Now, using (2.15), (2.16), and (2.17) we get

$$\begin{aligned} \mathbf{J} &= \boldsymbol{\gamma}^{(t)'} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sum_1^4 n_i x_i^6 - P & \sum_1^4 n_i x_i^5 - R \\ 0 & 0 & \sum_1^4 n_i x_i^5 - R & S - Q \end{pmatrix} \boldsymbol{\gamma}^{(t)} \\ &= (S - Q)\gamma_2^2 + \left(\sum_1^4 n_i x_i^6 - P \right) \gamma_3^2 + 2\gamma_2\gamma_3 \left(\sum_1^4 n_i x_i^5 - R \right), \end{aligned} \quad (2.18)$$

where P, Q, R , and S are defined right after (2.17).

Now we consider the four distinct points as $-x_1 = x_4 = b$ and $-x_2 = x_3 = a$ where $0 < a < b \leq 1$. We also consider the allocation $n_1 = n_4$ and $n_2 = n_3$. These design points and their allocations give us the following subclass of designs in the general class of designs \mathbf{D} in (1.3)

$$\begin{aligned} \mathbf{D}_1 &= \left\{ (x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : \right. \\ &\quad -1 \leq x_1 = -b < x_2 = -a < x_3 = a < x_4 = b \leq 1, \\ &\quad \left. 0 < a < b \leq 1; n_1 = n_4, n_2 = n_3, n_1 + n_2 = \frac{n}{2}, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right\}. \end{aligned} \quad (2.19)$$

We note that under \mathbf{D}_1 we have

$$\begin{aligned} \text{Det}_1 &= 2n(b^2 n_1 + a^2 n_2), & \text{Det}_2 &= 4nn_1 n_2 a^2 b^2 (b^2 - a^2)^2, & A &= 0, \\ B &= \frac{2}{n}(b^2 n_1 + a^2 n_2), & C &= \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)}, & D &= 0, \\ E &= \frac{2}{n}(b^2 n_1 + a^2 n_2), & F &= 0, & G &= 0, \end{aligned}$$

$$P = \frac{2(b^4n_1 + a^4n_2)^2}{(b^2n_1 + a^2n_2)^2}, \quad Q = \frac{4}{n}(b^2n_1 + a^2n_2)^2 = S, \quad R = 0,$$

and hence the fitted value criterion takes the following form as defined by \mathbf{J}_b where

$$\begin{aligned} \mathbf{J}_b &= \frac{2\gamma_3^2 n_1 n_2 a^2 b^2 (b^2 - a^2)^2}{(b^2 n_1 + a^2 n_2)} \\ \Leftrightarrow \frac{\mathbf{J}_b}{n\gamma_3^2} &= \frac{2p_1(1 - 2p_1)a^2 b^2 (b^2 - a^2)^2}{2p_1 b^2 + (1 - 2p_1)a^2}. \end{aligned} \quad (2.20)$$

2.4 Efficient Designs with respect to J Criterion

Under the class of designs \mathbf{D}_1 in (2.19) the fitted value criterion takes the form \mathbf{J}_b in (2.20). We optimize the criterion function to obtain efficient designs with respect to the fitted value criterion.

2.4.1 Finding the value of a for a given value of b

Now, assuming b to be known if we maximize the fitted value criterion with respect to a and p_1 , we get $a = \frac{b}{2}$ and $p_1 = \frac{1}{6}$ (Appendix-A.1). This implies that we should consider x_2 and x_3 as the midpoints of the intervals $(-b, 0)$ and $(0, b)$ respectively and collect data with $\frac{1}{6}$ weight at each of the two end points, x_1 and x_4 , and with $\frac{1}{3}$ weight at each of the two middle points, x_2 and x_3 , to find the best design to maximize the fitted value criterion \mathbf{J}_b in the class of designs \mathbf{D}_1 in (2.19). We define the design below in \mathbf{D}_1 ,

$$d_1 = \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -b, x_2 = -\frac{b}{2}, x_3 = \frac{b}{2}, x_4 = b; \right. \\ \left. n_1 = n_4 = \frac{n}{6}, n_2 = n_3 = \frac{n}{3}, Rank(\mathbf{X}^{(t)}) = 4 \right], \quad (2.21)$$

and get the following theorem.

Theorem 1. *For a given value of b in the class of designs \mathbf{D}_1 , the design d_1 is optimum with respect to the criterion \mathbf{J} .*

2.4.2 Finding the value of b

Under the class of designs \mathbf{D}_1 in (2.19) when $-x_1 = x_4 = 1$ i.e. when $b = 1$, we obtain the fitted value criterion from (2.20) by replacing b with 1 and denote that by \mathbf{J}_1 where

$$\frac{\mathbf{J}_1}{n\gamma_3^2} = \frac{2p_1p_2a^2(1-a^2)^2}{(p_1+a^2p_2)}. \quad (2.22)$$

Clearly,

$$\mathbf{J}_1 - \mathbf{J}_b$$

$$= \frac{2n\gamma_3^2p_1p_2a^2(1-a^2)^2}{(p_1+a^2p_2)(b^2p_1+a^2p_2)} \left[p_1b^2 \left\{ 1 - \frac{(b^2-a^2)^2}{(1-a^2)^2} \right\} + p_2a^2 \left\{ 1 - b^2 \frac{(b^2-a^2)^2}{(1-a^2)^2} \right\} \right] \geq 0, \quad (2.23)$$

$$\text{because } 0 < b^2 \frac{(b^2-a^2)^2}{(1-a^2)^2} \leq \frac{(b^2-a^2)^2}{(1-a^2)^2} \leq 1.$$

We note that (2.23) holds for all $0 < a < b \leq 1$, $n_1 \neq 0$, and $n_2 \neq 0$. The ‘=’ holds only when $b = 1$. Thus the design with $-x_1 = x_4 = 1$, provides the maximum value of the fitted value criterion \mathbf{J} . We also see this in Appendix-A.2. Fig-2.1 also confirms the fact that \mathbf{J} attains the maximum possible value when $b = 1$. We considered $a = \frac{b}{2} = \frac{1}{2}$ while computing \mathbf{J} for the graph.

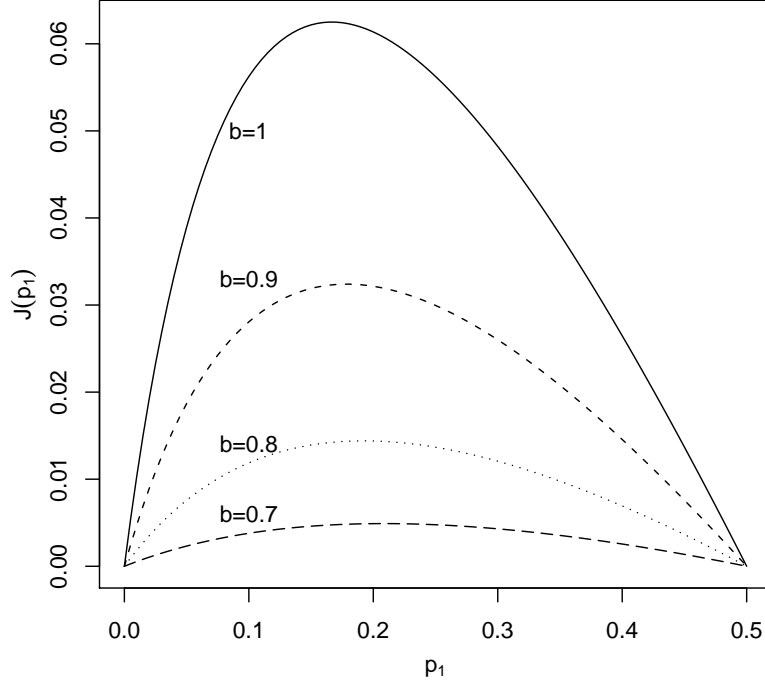


Figure 2.1: Plot of $\mathbf{J}(\mathbf{p}_1)$ against p_1 for different values of b

We define the design below in \mathbf{D}_1 ,

$$d_2 = \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}, x_4 = 1; \right. \\ \left. n_1 = n_4 = \frac{n}{6}, n_2 = n_3 = \frac{n}{3}, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right], \quad (2.24)$$

and obtain the following theorem.

Theorem 2. *For the class of designs \mathbf{D}_1 , the design d_2 is optimum with respect to the criterion \mathbf{J} .*

2.5 Expression of I Criterion

We now consider the predicted value criterion (ref. Chapter-1) for model selection and discrimination purposes. Let $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ be the predicted values of the two models MI and MII at $X = x$. We first find the expression for $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ assuming MT to be the true model. We note that the quantity, $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$, may be positive for some values of x and negative for some other values of x . Therefore, we consider the squared value of $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ and find the criterion. Under the class of designs \mathbf{D} in (1.3) we have

$$\begin{aligned}
 E(\hat{y}^{(1)}(x)) &= \begin{pmatrix} 1 & x \end{pmatrix} E(\hat{\gamma}^{(1)}) \\
 &= \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} 1 & 0 & A & B \\ 0 & 1 & C & D \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (2.10)} \\
 &= \gamma_0 + \gamma_1 x + \gamma_2(B + Dx) + \gamma_3(A + Cx), \tag{2.25}
 \end{aligned}$$

and

$$\begin{aligned}
 E(\hat{y}^{(2)}(x)) &= \begin{pmatrix} 1 & x & x^3 \end{pmatrix} E(\hat{\gamma}^{(2)}) \\
 &= \begin{pmatrix} 1 & x & x^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & E \\ 0 & 1 & 0 & F \\ 0 & 0 & 1 & G \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (2.11)} \\
 &= \gamma_0 + \gamma_1 x + \gamma_2(E + Fx + Gx^3) + \gamma_3 x^3, \tag{2.26}
 \end{aligned}$$

where A, B, C, D, E, F , and G are defined right after (2.11). Thus from (2.25) and (2.26) we get

$$E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x)) = \gamma_2 [(B - E) + (D - F)x - Gx^3] + \gamma_3 [A + Cx - x^3] \quad (2.27)$$

The expression in (2.27) is intractable with respect to the general class of designs \mathbf{D} . Hence we consider the subclass of designs \mathbf{D}_1 in (2.19). Now under \mathbf{D}_1 , we have $B = E = \frac{2}{n}(b^2n_1 + a^2n_2)$, $A = D = F = G = 0$, and $C = \frac{(b^4n_1 + a^4n_2)}{(b^2n_1 + a^2n_2)}$. Hence under the class of designs \mathbf{D}_1 we get from (2.27),

$$[E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 = \gamma_3^2 x^2 (C - x^2)^2. \quad (2.28)$$

The discrimination between the models MI and MII will be the best when $\left[\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_3} \right]^2$ is maximum. The discrimination between the models MI and MII will not be possible with respect to predicted values when $\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_3} = 0$. We observe that at $x = 0, x = -\sqrt{C}$, and $x = \sqrt{C}$, $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x)) = 0$. Clearly, at $x = 0, x = -\sqrt{C}$, and $x = \sqrt{C}$, the discrimination between MI and MII will not be possible using the prediction criterion at the design stage. But we are interested in finding an overall prediction measure for discrimination purpose rather than evaluating at each x . So, we consider an overall measure by taking the weighted average of $\left[\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_3} \right]^2$ on x . Therefore the predicted value criterion

is given by,

$$\begin{aligned}
\mathbf{I} &= \frac{1}{2} \int_{-1}^1 [E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 dx \\
&= \frac{\gamma_3^2}{2} \int_{-1}^1 (Cx - x^3)^2 dx \\
&= \gamma_3^2 \int_0^1 (C^2x^2 - 2Cx^4 + x^6) dx \\
&= \gamma_3^2 \left[\frac{C^2}{3} - \frac{2C}{5} + \frac{1}{7} \right] \quad \text{where } 0 < C = \frac{(b^4n_1 + a^4n_2)}{(b^2n_1 + a^2n_2)} < 1.
\end{aligned} \tag{2.29}$$

2.6 Efficient Designs with respect to \mathbf{I} Criterion

We note that the criterion \mathbf{I} is a concave function with respect to C (*Fig-2.2*).

Designs with a very small or a very large value of C will perform well with respect

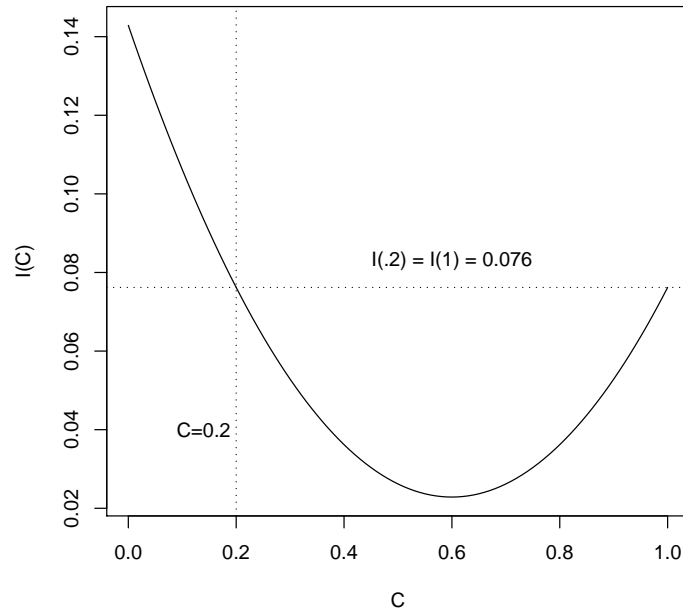


Figure 2.2: Plot of $\mathbf{I}(C)$ against C

to \mathbf{I} . We also note that $\mathbf{I}(0.2) = \mathbf{I}(1)$ but it is clear that $C \neq 1$ and $C \neq 0$. We obtain some choices of a, b , and p_1 numerically to present some designs (Table -2.1) in \mathbf{D}_1 in (2.19) which perform well with respect to \mathbf{I} .

Table 2.1: Some \mathbf{I} -optimal Designs

C	a	b	p_1	p_2	\mathbf{J}	\mathbf{I}
0.10	0.151	0.48	0.028	0.472	0.000	0.1062
0.10	0.209	0.439	0.060	0.440	0.000	0.1062
0.20	0.284	0.847	0.013	0.487	0.006	0.0762
0.20	0.307	0.749	0.024	0.476	0.004	0.0762
0.80	0.503	1.000	0.204	0.296	0.061	0.0362
0.80	0.537	1.000	0.212	0.288	0.060	0.0362
0.85	0.595	0.998	0.274	0.226	0.051	0.0437
0.90	0.404	0.995	0.288	0.212	0.042	0.0529
0.95	0.537	0.996	0.410	0.090	0.024	0.0637
0.99	0.424	1.000	0.471	0.029	0.007	0.0736

We note that the designs those perform well with respect to \mathbf{I} , do not perform well with respect to \mathbf{J} and vice versa. As a trade off we might be interested in designs which perform moderately well with respect to both \mathbf{I} and \mathbf{J} .

2.7 Some Comparisons between \mathbf{I} and \mathbf{J}

We now consider $b = 1$ and obtain some designs numerically in \mathbf{D}_1 which yield same numerical value for both the criteria and demonstrate them in Table-2.2. Then we

Table 2.2: Some designs in \mathbf{D}_1 for which $\mathbf{I} = \mathbf{J}$

C	a	b	p_1	p_2	\mathbf{J}	\mathbf{I}
0.8857	0.46998	1.00000	0.28119	0.21881	0.05007	0.05007
0.8857	0.47352	1.00000	0.28244	0.21757	0.05007	0.05007
0.8857	0.47438	1.00000	0.28272	0.21728	0.05007	0.05007
0.8857	0.47184	1.00000	0.28185	0.21815	0.05007	0.05007
0.8857	0.47450	1.00000	0.28276	0.21724	0.05007	0.05007
0.8857	0.47379	1.00000	0.28253	0.21748	0.05007	0.05007
0.8857	0.47479	1.00000	0.28286	0.21714	0.05007	0.05007
0.8857	0.47450	1.00000	0.28277	0.21723	0.05007	0.05007
0.8857	0.46966	1.00000	0.28108	0.21892	0.05007	0.05007
0.8857	0.47352	1.00000	0.28243	0.21757	0.05007	0.05007

consider the following subclass of designs in \mathbf{D}_1 ,

$$\mathbf{D}_2 = \left\{ (x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}, x_4 = 1; \right. \\ \left. n_1 = n_4, n_2 = n_3, n_1 + n_2 = \frac{n}{2}, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right\}. \quad (2.30)$$

We explore the performance of the designs in \mathbf{D}_2 with respect to criteria \mathbf{I} and \mathbf{J} .

2.7.1 The $\mathbf{I}(p_1)$ and $\mathbf{J}(p_1)$ for Designs in \mathbf{D}_2

Here we will discuss the properties of \mathbf{J} and \mathbf{I} criterion functions derived under the class of designs \mathbf{D}_2 . We substitute $b = 1$ and $a = \frac{1}{2}$ in (2.20) and obtain the fitted value criterion as

$$\mathbf{J}(p_1) = \frac{9p_1(1-2p_1)}{8(1+6p_1)}, \quad 0 < p_1 < \frac{1}{2}.$$

The $\mathbf{J}(p_1)$ can also be written as, $\mathbf{J}(p_1) = \left[\frac{1}{16} - \frac{3(p_1 - \frac{1}{6})^2}{8(p_1 + \frac{1}{6})} \right]$.

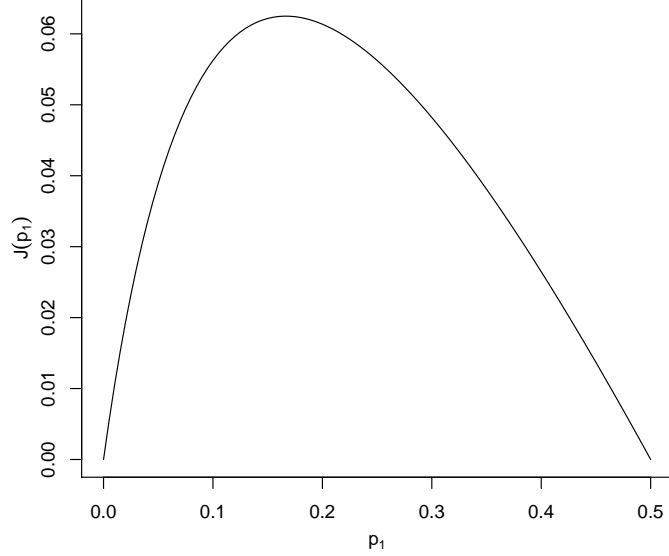


Figure 2.3: Plot of $\mathbf{J}(p_1)$ against p_1

- $\mathbf{J}(p_1)$ monotonically increases first and reaches its maximum at $p_1 = \frac{1}{6}$ and the maximum value is $\frac{1}{16}$, then it monotonically decreases to zero (*Fig-2.3*).
- Larger value of $\mathbf{J}(p_1)$ implies better discrimination.

We also substitute $b = 1$ and $a = \frac{1}{2}$ in (2.29) and obtain the predicted value criterion

as
$$\mathbf{I}(p_1) = \frac{(9900p_1^2 - 1068p_1 + 107)}{1680(1+6p_1)^2}, \quad 0 < p_1 < \frac{1}{2}.$$

Note that $\mathbf{I}(p_1)$ can also be written as,

$$\mathbf{I}(p_1) = \frac{55}{336} \left(\frac{A^2 + \frac{192}{1375}B}{A^2 + B} \right) \text{ where } A = \left(p_1 - \frac{7}{78} \right), B = \frac{20}{39} \left(p_1 + \frac{1}{26} \right).$$

Now, $A^2 \geq \frac{192}{1375}A^2$ where “=” holds iff $A = 0$ i.e iff $p_1 = \frac{7}{78}$

$$\Leftrightarrow \left(A^2 + \frac{192}{1375}B \right) \geq \frac{192}{1375} (A^2 + B)$$

$$\Leftrightarrow \mathbf{I}(p_1) \geq \frac{55}{336} \times \frac{192}{1375} = \frac{4}{175}.$$

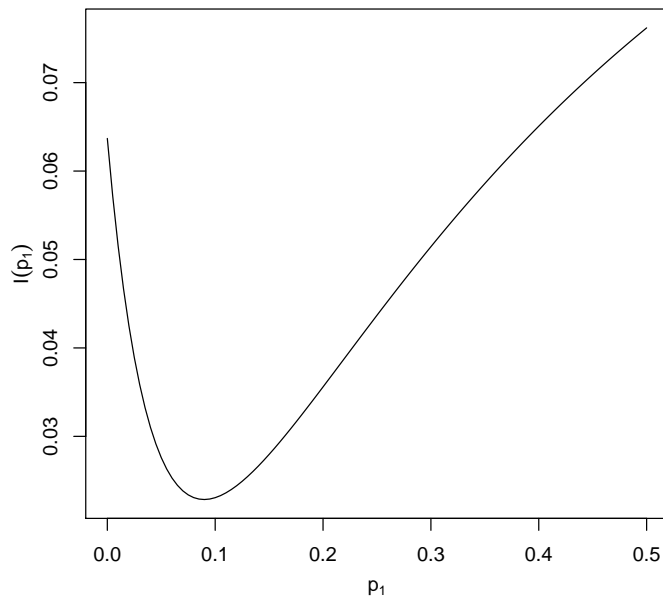


Figure 2.4: Plot of $\mathbf{I}(p_1)$ against p_1

- $\mathbf{I}(p_1)$ monotonically decreases to its minimum at $p_1 = \frac{7}{78}$ and the minimum value is $\frac{4}{175}$, then it monotonically increases (*Fig-2.4*).
- Larger value of $\mathbf{I}(p_1)$ implies better discrimination.

2.7.2 Exploring the Equality of $\mathbf{I}(p_1)$ and $\mathbf{J}(p_1)$ in D_2

We will study the following three situations first:

$$\mathbf{I}(p_1) = \mathbf{J}(p_1), \quad \mathbf{I}(p_1) > \mathbf{J}(p_1), \quad \text{and} \quad \mathbf{I}(p_1) < \mathbf{J}(p_1).$$

From (*Fig-2.5*) we note that

- $\mathbf{I}(p_1)$ values are in $(0.022857, 0.076190)$, $\mathbf{J}(p_1)$ values are in $(0, 0.0625)$.
- $\mathbf{I}(p_1)$ and $\mathbf{J}(p_1)$ intersect twice at $p_1 = 0.037709$ and $p_1 = 0.290214$.

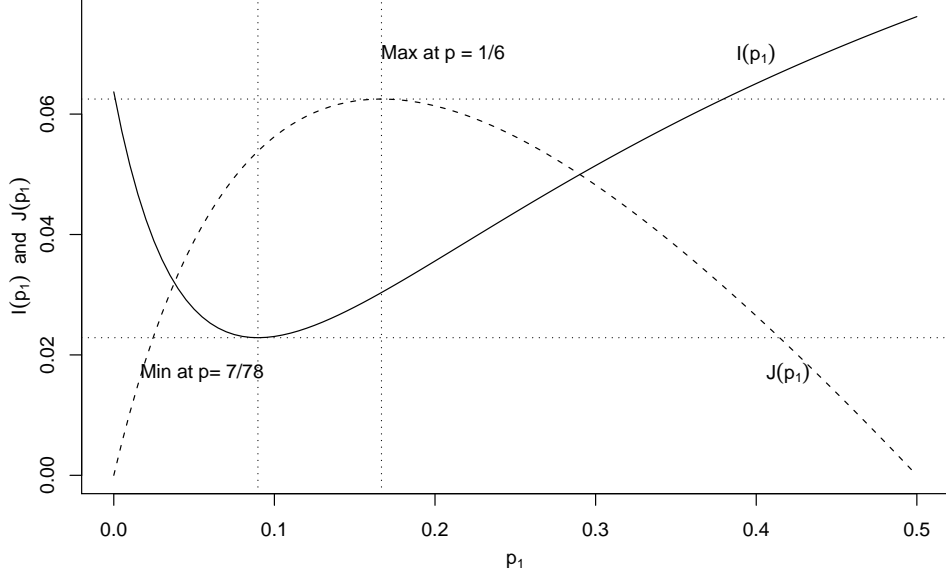


Figure 2.5: Plot of $\mathbf{I}(p_1)$ and $\mathbf{J}(p_1)$ against p_1

- $\mathbf{I}(p_1) > \mathbf{J}(p_1)$ when $0 < p_1 < 0.037709$ and $0.290214 < p < 0.5$.
- $\mathbf{I}(p_1) < \mathbf{J}(p_1)$ when $0.037709 < p_1 < 0.290214$.

Here we find the value(s) of p_1 for which $\mathbf{I}(p_1)$ and $\mathbf{J}(p_1)$ are equal. We already obtained these values graphically. We consider the function $F(p_1)$ where

$$\begin{aligned}
 F(p) &= \mathbf{I}(p_1) - \mathbf{J}(p_1) \\
 &= \frac{(9900p_1^2 - 1068p_1 + 107)}{1680(1 + 6p_1)^2} - \frac{9p_1(1 - 2p_1)}{8(1 + 6p_1)} \\
 &= \frac{(22680p_1^3 + 2340p_1^2 - 2958p_1 + 107)}{1680(1 + 6p_1)^2}. \tag{2.31}
 \end{aligned}$$

Clearly, $F(p_1) = 0$ will provide the value(s) of p_1 for which $\mathbf{I}(p_1) = \mathbf{J}(p_1)$. We note that $F(p_1) = 0$ in two points when $p_1 \in (0, \frac{1}{2})$: one in $0 < p_1 < 0.1$ and the other in

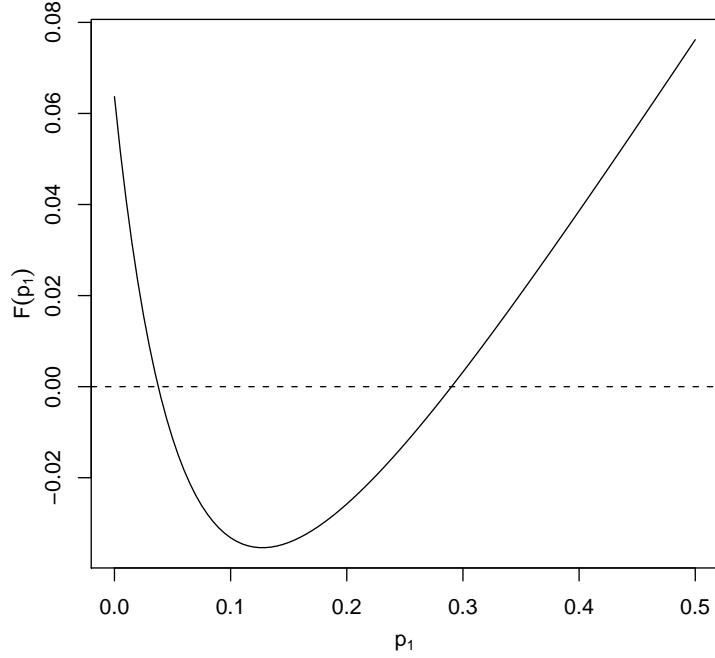


Figure 2.6: Plot of $F(p_1)$ against $p_1 \in (0, \frac{1}{2})$

$0.25 < p_1 < 0.30$ from (*Fig-2.6*). Now, using the Method of Bisection we can easily find those two points precisely as $p_1 = 0.037709$ and $p_1 = 0.290214$. We also note that only between these two points $F(p_1) < 0$ i.e. $\mathbf{I}(p_1) < \mathbf{J}(p_1)$.

We define two designs d_3 and d_4 in \mathbf{D}_2 in (2.30) as

$$d_3 = \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}, x_4 = 1; \right. \\ \left. n_1 = n_4 = 0.290214n, n_2 = n_3 = \frac{n}{2} - n_1, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right], \quad (2.32)$$

and

$$d_4 = \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}, x_4 = 1; \right. \\ \left. n_1 = n_4 = 0.037709n, n_2 = n_3 = \frac{n}{2} - n_1, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right]. \quad (2.33)$$

Thus we have the following theorems.

Theorem 3. *For the class of designs \mathbf{D}_2 , the design d_3 yield equal numerical value for both the criteria \mathbf{I} and \mathbf{J} .*

Theorem 4. *For the class of designs \mathbf{D}_2 , the design d_4 yield equal numerical value for both the criteria \mathbf{I} and \mathbf{J} .*

Theorem 5. *For the class of designs \mathbf{D}_2 , the design d_4 performs better than d_3 with respect to both the criteria \mathbf{I} and \mathbf{J} .*

We already have the design d_2 in (2.24) as the optimum design in \mathbf{D}_1 in (2.19) with respect to \mathbf{J} . Obviously it is optimum in \mathbf{D}_2 too because $\mathbf{D}_2 \subset \mathbf{D}_1$. Hence we have

Theorem 6. *For the class of designs \mathbf{D}_2 , the design d_2 is the optimum design with respect to \mathbf{J} .*

Now, we define the following subclass of designs in \mathbf{D}_2 (2.30),

$$\mathbf{D}_3 = \left\{ (x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}, x_4 = 1; \right. \\ \left. \frac{n}{6} < n_1 < \frac{n}{2}, n_2 = \frac{n}{2} - n_1, n_1 = n_4, n_2 = n_3, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right\} \quad (2.34)$$

and obtain the following theorem.

Theorem 7.

- (a) *All the designs in \mathbf{D}_3 perform better than the design d_2 with respect to the criterion \mathbf{I} .*
- (b) *The design d_2 performs better than all the designs in \mathbf{D}_3 with respect to the criterion \mathbf{J} .*

Chapter 3

Linear vs. Special Cubic as the True Model

3.1 Introduction

We again consider the experiment where the response variable Y is dependent on an explanatory variable X . So far we assumed the full cubic model MT to be the true model but here we assume $\gamma_2 = 0$. Therefore the two models MII and MT in Chapter-2 become identical here. We consider the other possible dependence as a simple linear regression model MI. Our goal is to discriminate between these two models MI and MII at the design stage assuming MII to be the true model.

3.2 Models and Associated Designs

We consider the class of designs D in (1.3). We also consider the cubic model without the quadratic term from (1.1) as the true model and denote it by MII. Our aim is to discriminate between the two models MI and MII.

The two models considered here are given by

$$\begin{aligned} \text{MI:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i, \\ \text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0, \end{aligned} \quad (3.1)$$

$$\begin{aligned} \text{MII:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_3 x_i^3, \\ \text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0, \end{aligned} \quad (3.2)$$

where $i, i' = 1, 2, 3, 4$; $j = 1, 2, \dots, n_i, j' = 1, 2, \dots, n_{i'}, (i, j) \neq (i', j')$.

The matrix representations of MII and MI are given by

$$E(\mathbf{y}) = \mathbf{X}^{(2)}\boldsymbol{\gamma}^{(2)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \quad (3.3)$$

$$E(\mathbf{y}) = \mathbf{X}^{(1)}\boldsymbol{\gamma}^{(1)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \quad (3.4)$$

where

$$\begin{aligned} \mathbf{X}^{(1)} &= \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(1)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix}, \\ \mathbf{X}^{(2)} &= \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(2)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_3 \end{pmatrix}. \end{aligned}$$

We define

$$\mathbf{X}_3 = \begin{pmatrix} x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}'.$$

Hence we get,

$$\mathbf{X}^{(2)} = \begin{pmatrix} \mathbf{X}^{(1)} & \vdots & \mathbf{X}_3 \end{pmatrix} \Rightarrow \mathbf{X}^{(2)'}\mathbf{X}^{(2)} = \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_3 \\ \mathbf{X}_3'\mathbf{X}^{(1)} & \mathbf{X}_3'\mathbf{X}_3 \end{pmatrix}. \quad (3.5)$$

3.3 Expression of J Criterion

We first consider the fitted value criterion J (ref. Chapter-1) and obtain the efficient design within the class of designs \mathbf{D} in (1.3) for model selection and discrimination purposes. Let $\hat{\mathbf{y}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)}$ be the fitted values of the two models MI and MII respectively. First we have to find the expression of $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$ where the expectation is considered under the true model MII in (3.2). Then we have to find the expression of $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})'E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$.

The least square estimate of $\boldsymbol{\gamma}^{(1)}$ for MI is given by (Rao (1973))

$$\hat{\boldsymbol{\gamma}}^{(1)} = \begin{pmatrix} \hat{\gamma}_0^{(1)} \\ \hat{\gamma}_1^{(1)} \end{pmatrix} = (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{y}, \quad (3.6)$$

and the least square estimate of $\boldsymbol{\gamma}^{(2)}$ for MII is given by

$$\hat{\boldsymbol{\gamma}}^{(2)} = \begin{pmatrix} \hat{\gamma}_0^{(2)} \\ \hat{\gamma}_1^{(2)} \\ \hat{\gamma}_3^{(2)} \end{pmatrix} = (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{y}. \quad (3.7)$$

We note that $\hat{\gamma}_0^{(2)}$, $\hat{\gamma}_1^{(2)}$, and $\hat{\gamma}_3^{(2)}$ are the least square estimates of γ_0 , γ_1 , and γ_3 respectively from MII. So, under MII these estimates are the best linear unbiased estimates (BLUE) (Rao (1973)) of the respective parameters i.e. $E(\hat{\boldsymbol{\gamma}}^{(2)}) = \boldsymbol{\gamma}^{(2)}$.

The expression of \mathbf{J} is readily available from (2.18). We just replace γ_2 by 0 and get the expression of the fitted value criterion as

$$\mathbf{J} = \left(\sum_1^4 n_i x_i^6 - P \right) \gamma_3^2, \quad (3.8)$$

where P is defined right after (2.17).

It can be checked that the detailed expression of \mathbf{J} in (3.8) is given by

$$\begin{aligned} \mathbf{J} = & \gamma_3^2 \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3)^2 \right. \\ & + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4)^2 \\ & + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4)^2 \\ & \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4)^2 \right] \\ & \left[n_1 n_2 (x_1 - x_2)^2 + n_1 n_3 (x_1 - x_3)^2 + n_1 n_4 (x_1 - x_4)^2 \right. \\ & \left. + n_2 n_3 (x_2 - x_3)^2 + n_2 n_4 (x_2 - x_4)^2 + n_3 n_4 (x_3 - x_4)^2 \right]. \end{aligned} \quad (3.9)$$

Considering $-x_1 = x_4 = b$ and $-x_2 = x_3 = a$ where $0 < a < b \leq 1$, \mathbf{J} reduces to,

$$\begin{aligned} \mathbf{J} &= \frac{4\gamma_3^2 a^2 b^2 (b^2 - a^2)^2 (n_1 n_2 n_3 + n_1 n_2 n_4 + n_1 n_3 n_4 + n_2 n_3 n_4)}{\left[(b+a)^2 (n_1 n_3 + n_2 n_4) + (b-a)^2 (n_1 n_2 + n_3 n_4) + 4a^2 n_2 n_3 + 4b^2 n_1 n_4 \right]} \\ &= \frac{4\gamma_3^2 a^2 b^2 (b^2 - a^2)^2 (n_1 n_2 n_3 + n_1 n_2 n_4 + n_1 n_3 n_4 + n_2 n_3 n_4)}{\left[\frac{n^2}{2} (b^2 + a^2) - \frac{1}{2} [a(n_1 - n_4) + b(n_2 - n_3)]^2 - \frac{1}{2} [b(n_1 - n_4) + a(n_2 - n_3)]^2 \right.} \\ & \quad \left. + 2(b^2 - a^2)(n_1 n_4 - n_2 n_3) \right]}. \end{aligned} \quad (3.10)$$

Now, the denominator of the expression of \mathbf{J} in (3.10) suggests a meaningful allocation $n_1 = n_4$ and $n_2 = n_3$ which reduces the \mathbf{J} considerably as follows:

$$\begin{aligned}
\mathbf{J} &= \frac{4\gamma_3^2 nn_1 n_2 a^2 b^2 (b^2 - a^2)^2}{\left[\frac{n^2}{2}(b^2 + a^2) + 2(b^2 - a^2)(n_1^2 - n_2^2) \right]} \\
&= \frac{4\gamma_3^2 nn_1 n_2 a^2 b^2 (b^2 - a^2)^2}{2n(n_1 b^2 + n_2 a^2)} \quad \text{since} \quad (n_1 + n_2) = \frac{n}{2} \\
\Leftrightarrow \frac{\mathbf{J}}{n\gamma_3^2} &= \frac{2p_1(1 - 2p_1)a^2 b^2 (b^2 - a^2)^2}{2p_1 b^2 + (1 - 2p_1)a^2} \quad \text{where} \quad (p_1 + p_2) = \frac{1}{2}. \tag{3.11}
\end{aligned}$$

Clearly the general design class \mathbf{D} in (1.3) reduces to a sub-class of designs \mathbf{D}_1 in (2.19) when we consider these design points with the special allocation scheme.

3.4 Expression of T -Optimality Criterion

We consider the fitted values of \mathbf{y} under MI and MII. Here $\hat{\mathbf{y}}^{(1)} = \mathbf{X}^{(1)}\hat{\boldsymbol{\gamma}}^{(1)} = \mathbf{H}_1\mathbf{y}$ and $\hat{\mathbf{y}}^{(2)} = \mathbf{X}^{(2)}\hat{\boldsymbol{\gamma}}^{(2)} = \mathbf{H}_2\mathbf{y}$ where $\hat{\boldsymbol{\gamma}}^{(1)}$ and $\hat{\boldsymbol{\gamma}}^{(2)}$ are defined in (3.6) and (3.7) respectively, and $\mathbf{H}_1 = \mathbf{X}^{(1)}(\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}$, $\mathbf{H}_2 = \mathbf{X}^{(2)}(\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}$.

We note that \mathbf{H}_1 and \mathbf{H}_2 are symmetric and idempotent (Rao (1973)) i.e. $\mathbf{H}_1' = \mathbf{H}_1$, $\mathbf{H}_1^2 = \mathbf{H}_1$ and $\mathbf{H}_2' = \mathbf{H}_2$, $\mathbf{H}_2^2 = \mathbf{H}_2$. Now, the lack-of-fit sum of squares is given by

$$\begin{aligned}
SS_{LOF} &= (\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)})'(\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)}) \\
&= \mathbf{y}'(\mathbf{H}_2 - \mathbf{H}_1)^2\mathbf{y} \\
&= \mathbf{y}'(\mathbf{H}_2 - \mathbf{H}_1)\mathbf{y}, \quad \text{by Result-1.} \tag{3.12}
\end{aligned}$$

Hence

$$\begin{aligned}
E(SS_{LOF}) &= E[\mathbf{y}'(\mathbf{H}_2 - \mathbf{H}_1)\mathbf{y}] = tr[(\mathbf{H}_2 - \mathbf{H}_1)E(\mathbf{y}\mathbf{y}')] \\
&= tr[(\mathbf{H}_2 - \mathbf{H}_1)\{\sigma^2\mathbf{I} + E(\mathbf{y})\mathbf{E}(\mathbf{y})'\}] \\
&= \sigma^2 tr(\mathbf{H}_2 - \mathbf{H}_1) + E(\mathbf{y})'(\mathbf{H}_2 - \mathbf{H}_1)\mathbf{E}(\mathbf{y}) \\
&= \sigma^2 + \boldsymbol{\gamma}'\mathbf{X}^{(2)'}(\mathbf{I} - \mathbf{H}_1)\mathbf{X}^{(2)}\boldsymbol{\gamma}.
\end{aligned} \tag{3.13}$$

Thus from (3.8) and (3.13) we get

$$\begin{aligned}
\mathbf{NCP} &= \frac{1}{\sigma^2}\boldsymbol{\gamma}'\mathbf{X}^{(2)'}(\mathbf{I} - \mathbf{H}_1)\mathbf{X}^{(2)}\boldsymbol{\gamma} \\
&= \frac{1}{\sigma^2}\gamma_3^2(\mathbf{X}_3'\mathbf{X}_3 - \mathbf{X}_3'\mathbf{H}_1\mathbf{X}_3) \\
&= \frac{1}{\sigma^2}\gamma_3^2\left(\sum_1^4 n_i x_i^6 - P\right) = \frac{\mathbf{J}}{\sigma^2}.
\end{aligned} \tag{3.14}$$

Hence the \mathbf{T} -optimality criterion is equivalent to the fitted value criterion- \mathbf{J} here.

3.4.1 Two Interesting Observations

Firstly we note the \mathbf{T} -optimality criterion is equivalent to the fitted value criterion \mathbf{J} here. We note that in Chapter-2 the \mathbf{T} -optimality criterion is not defined if we are interested to discriminate between the models MI and MII because the true model is a different model MT. But in this chapter it is defined because MII and MT are identical i.e. here we have to discriminate between two models MI and MII where MII is the true model.

Secondly we observe that under the subclass of designs \mathbf{D}_1 in (2.19), the fitted value criterion \mathbf{J} obtained in two different cases, one in Chapter-2 (in 2.20) and the other in this chapter (in 3.11), are identical. In Chapter-2 we have two fitted

models MI and MII whereas the true model is the full cubic model MT. In this chapter we have two models as MI and MII where the true model is MII. The class of designs \mathbf{D}_1 in (2.19) produces the identical value of the fitted value criterion \mathbf{J} in both the cases.

3.5 Efficient Designs with respect to \mathbf{J} Criterion

Since the expressions of \mathbf{J} for $\gamma_2 = 0$ (Chapter-3 in 3.11) and $\gamma_2 \neq 0$ (Chapter-2 in 2.20) are identical under the class of designs \mathbf{D}_1 , the theorems from Chapter-2 apply here.

Theorem 8. *For a given value of b in the class of designs \mathbf{D}_1 in (2.19), the design d_1 in (2.21) is optimum with respect to the criterion \mathbf{J} (or equivalently \mathbf{T}).*

We note that d_2 in (2.24) becomes the special Dette-Titoff design from the class of \mathbf{T} -optimal designs (Dette and Titoff (2009)). Hence we get following theorem.

Theorem 9. *For the class of designs \mathbf{D}_1 , the special Dette-Titoff design d_2 is optimum with respect to the criterion \mathbf{J} (or equivalently \mathbf{T}).*

We now consider a general allocation instead of the special allocation $n_1 = n_4$ and $n_2 = n_3$ and explore the new class of designs \mathbf{D}_4 to find the best design with respect to \mathbf{J} (or equivalently \mathbf{T} -optimality) criterion where

$$\mathbf{D}_4 = \left\{ (x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}, x_4 = 1 \right. \\ \left. n_1 + n_2 + n_3 + n_4 = n, Rank(\mathbf{X}^{(2)}) = 4 \right\}. \quad (3.15)$$

Under the class of designs \mathbf{D}_4 , the matrix representations of MI and MII are given

by

$$E(\mathbf{y}) = \mathbf{X}^{(1)}\boldsymbol{\gamma}^{(1)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I},$$

$$E(\mathbf{y}) = \mathbf{X}^{(2)}\boldsymbol{\gamma}^{(2)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I},$$

where

$$\mathbf{X}^{(1)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ -1 & \dots & -1 & -\frac{1}{2} & \dots & -\frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} & 1 & \dots & 1 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(1)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix},$$

$$\mathbf{X}_3 = \begin{pmatrix} -1 & \dots & -1 & -\frac{1}{8} & \dots & -\frac{1}{8} & \frac{1}{8} & \dots & \frac{1}{8} & 1 & \dots & 1 \end{pmatrix}',$$

and

$$\begin{aligned} \mathbf{X}^{(2)} &= \begin{pmatrix} \mathbf{X}^{(1)} & \mathbf{X}_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ -1 & \dots & -1 & -\frac{1}{2} & \dots & -\frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} & 1 & \dots & 1 \\ -1 & \dots & -1 & -\frac{1}{8} & \dots & -\frac{1}{8} & \frac{1}{8} & \dots & \frac{1}{8} & 1 & \dots & 1 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(2)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_3 \end{pmatrix}. \end{aligned}$$

We note that under \mathbf{D}_4 in (3.15), the fitted value criterion \mathbf{J} in (3.9) reduces to

$$\begin{aligned} \mathbf{J} &= \frac{\frac{9}{16}\gamma_3^2 [n_1n_2n_3 + n_1n_2n_4 + n_1n_3n_4 + n_2n_3n_4]}{\left[n \left[n_1 + n_4 + \frac{1}{4}(n_2 + n_3) \right] - \left[n_4 - n_1 + \frac{1}{2}(n_3 - n_2) \right]^2 \right]} \\ \Leftrightarrow \frac{16\mathbf{J}}{n\gamma_3^2} &= \frac{(n_4 + n_1)[(n_3 + n_2)^2 - (n_3 - n_2)^2] + (n_3 + n_2)[(n_4 + n_1)^2 - (n_4 - n_1)^2]}{\frac{4n}{9} \left[n \left[n_1 + n_4 + \frac{1}{4}(n_2 + n_3) \right] - \left[n_4 - n_1 + \frac{1}{2}(n_3 - n_2) \right]^2 \right]} \\ \Leftrightarrow \mathbf{J}_0 &= \frac{(t_1t_3 - t_1t_4^2 - t_2^2t_3)}{\frac{4}{9} \left[\left(t_1 + \frac{t_3}{4} \right) - \left(t_2 + \frac{t_4}{2} \right)^2 \right]}, \quad \text{assuming} \quad \mathbf{J}_0 = \frac{16\mathbf{J}}{n\gamma_3^2}, \end{aligned} \quad (3.16)$$

where

$$\begin{aligned} t_1 &= p_4 + p_1, & t_2 &= p_4 - p_1, \\ t_3 &= p_3 + p_2, & t_4 &= p_3 - p_2, \end{aligned} \tag{3.17}$$

$$t_1 + t_3 = 1, \quad 0 < t_1, t_3 < 1, \quad \text{and} \quad -1 < t_2, t_4 < 1. \tag{3.18}$$

Our goal is to maximize \mathbf{J} or equivalently \mathbf{J}_0 with respect to t_i 's i.e. with respect to p_i 's. Now we define

$$\begin{aligned} F_0 &= \text{Denominator of } \mathbf{J}_0 - \text{Numerator of } \mathbf{J}_0 \\ &= \frac{4}{9} \left[t_1 + \frac{t_3}{4} - \left(t_2 + \frac{t_4}{2} \right)^2 \right] - [t_1 t_3 - t_1 t_4^2 - t_2^2 t_3] \\ &= \frac{(3t_1 - 1)^2}{9} + \frac{2(t_4 - t_2)^2}{9} + \frac{(3t_1 - 1)(t_4^2 - t_2^2)}{3} \\ &= \frac{1}{9} [(3t_1 - 1)^2 + 2(t_4 - t_2)^2 + 3(3t_1 - 1)(t_4^2 - t_2^2)] \\ \Leftrightarrow F &= [(3t_1 - 1)^2 + 2(t_4 - t_2)^2 + 3(3t_1 - 1)(t_4^2 - t_2^2)], \end{aligned} \tag{3.19}$$

where $F = 9F_0$. We note that when $t_1 = \frac{1}{3}$ and $t_2 = t_4$, we have $F = 0$ implying Denominator = Numerator i.e. $\mathbf{J}_0 = 1$. We explore the function F to find if it is positive valued for all the values of t_1, t_2 , and t_4 . If $F \geq 0$ for all values of t_i 's, then \mathbf{J}_0 takes maximum value 1 when $F = 0$.

We define $u = (3t_1 - 1)$, $v = t_4 - t_2$, and $w = t_4 + t_2$. Thus from (3.19) we get

$$F = u^2 + 2v^2 + 3uvw \tag{3.20}$$

$$\Leftrightarrow F = \left(u + \frac{3}{2}vw \right)^2 - \frac{9}{4}v^2 \left(w^2 - \frac{8}{9} \right), \tag{3.21}$$

where $-1 < u < 2$, $-1 < v < 1$, and $-1 < w < 1$.

When $v \neq 0$ and $w^2 > \frac{8}{9}$, F in (3.21) can be written as,

$$F = \left(u + \frac{3}{2}vw + \frac{3}{2}\sqrt{v^2\left(w^2 - \frac{8}{9}\right)} \right) \left(u + \frac{3}{2}vw - \frac{3}{2}\sqrt{v^2\left(w^2 - \frac{8}{9}\right)} \right) \\ = (u - u_1)(u - u_2), \quad (3.22)$$

where $u_1 = -\frac{3}{2}vw - \frac{3}{2}\sqrt{v^2\left(w^2 - \frac{8}{9}\right)}$, $u_2 = -\frac{3}{2}vw + \frac{3}{2}\sqrt{v^2\left(w^2 - \frac{8}{9}\right)}$, and $u_2 > u_1$.

We now study the sign of F and tabulate our findings in Table-3.1.

Table 3.1: Study of the sign of F

u, v, w	Sign of F
$v = 0$	$F \geq 0$ by Theorem 10
$v \neq 0, w^2 \leq \frac{8}{9}$	$F \geq 0$ by Theorem 10
$v \neq 0, w^2 > \frac{8}{9}$	$F \geq 0$ or $F \leq 0$ by Theorem 10
$v \neq 0, w^2 > \frac{8}{9}$, and $DEN > 0$	$F > 0$ by Theorem 11
$v = 0, u = 0, w^2 < \frac{8}{9}$ and $DEN > 0$	$F = 0$ by Theorem 12

Theorem 10.

- (a) If $v = 0$, then $F \geq 0$.
- (b) If $v \neq 0$ and $w^2 \leq \frac{8}{9}$, then $F \geq 0$.
- (c) If $v \neq 0$ and $w^2 > \frac{8}{9}$, then $F \geq 0$ or $F \leq 0$.

Proof:

The (a) can be seen from the expression of F in (3.20) or in (3.21).

The (b) can be seen from the expression of F in (3.21).

From (3.22) we note that when $u_1 \leq u \leq u_2$, $F \leq 0$. But when $u \leq u_1$ or $u \geq u_2$ we have $F \geq 0$. This completes the proof of (c).

Theorem 11. *If $v \neq 0$, $w^2 > \frac{8}{9}$, and $DEN > 0$, then $F > 0$ where $DEN =$ Denominator of \mathbf{J}_0 in (3.16).*

Proof: From (3.16) we note that

$$\begin{aligned}
DEN &= \frac{4}{9} \left[t_1 + \frac{t_3}{4} - \left(t_2 + \frac{t_4}{2} \right)^2 \right] \\
&= \frac{4}{9} \left[\frac{3t_1 + 1}{4} - \frac{(2t_2 + t_4)^2}{4} \right] \\
&= \frac{1}{9} \left[u + 2 - \frac{(3w - v)^2}{4} \right] \\
&= \frac{1}{9} \left[\left(u + \frac{3}{2}vw \right) - \left(\frac{v^2}{4} + \frac{9}{4} \left(w^2 - \frac{8}{9} \right) \right) \right]. \tag{3.23}
\end{aligned}$$

$$DEN > 0 \Leftrightarrow \left(u + \frac{3}{2}vw \right) > \left[\frac{v^2}{4} + \frac{9}{4} \left(w^2 - \frac{8}{9} \right) \right]. \tag{3.24}$$

When $v \neq 0$, $w^2 > \frac{8}{9}$, and $DEN > 0$ we have,

$$\begin{aligned}
&\left(u + \frac{3}{2}vw \right) > \left[\frac{v^2}{4} + \frac{9}{4} \left(w^2 - \frac{8}{9} \right) \right] > 0 \\
&\Leftrightarrow \left(u + \frac{3}{2}vw \right)^2 > \left[\frac{v^2}{4} + \frac{9}{4} \left(w^2 - \frac{8}{9} \right) \right]^2 \\
&\Leftrightarrow \left(u + \frac{3}{2}vw \right)^2 > \left[\frac{v^2}{4} - \frac{9}{4} \left(w^2 - \frac{8}{9} \right) \right]^2 + \frac{9}{4}v^2 \left(w^2 - \frac{8}{9} \right) \\
&\Leftrightarrow F > \left[\frac{v^2}{4} - \frac{9}{4} \left(w^2 - \frac{8}{9} \right) \right]^2 \\
&\Rightarrow F > 0. \tag{3.25}
\end{aligned}$$

This completes the proof of Theorem 11.

Thus we have $F \geq 0$ for the admissible ranges of u, v and w .

Now we find when $F = 0$ holds. When $v \neq 0$ and $w^2 > \frac{8}{9}$, from (3.22) we have $F = 0$ at $u = u_1$ and $u = u_2$ and $F < 0$ when $u_1 < u < u_2$. But we already have seen that when $v \neq 0$ and $w^2 > \frac{8}{9}$,

$$DEN > 0$$

$$\begin{aligned} \Leftrightarrow \left(u + \frac{3}{2}vw\right) &> \left[\frac{v^2}{4} + \frac{9}{4}\left(w^2 - \frac{8}{9}\right)\right] \geq \frac{3}{2}\sqrt{v^2\left(w^2 - \frac{8}{9}\right)} \quad \because a^2 + b^2 \geq 2ab \\ \Rightarrow \left(u + \frac{3}{2}vw\right) &> \frac{3}{2}\sqrt{v^2\left(w^2 - \frac{8}{9}\right)} \\ \Leftrightarrow u &> u_2, \end{aligned}$$

and clearly when $u > u_2$, we have from (3.22) $F > 0$.

Hence, to satisfy the condition $DEN > 0$, u_1 and u_2 will not be realized as two distinct real roots of $(u^2 + 2v^2 + 3uvw) = 0$. Therefore we get,

$$\begin{aligned} u_1 = u_2 &= -\frac{3}{2}vw \\ \Leftrightarrow \frac{3}{2}\sqrt{v^2\left(w^2 - \frac{8}{9}\right)} &= 0 \\ \Leftrightarrow (I) : u &= -\frac{3}{2}vw, v = 0, 0 \leq w^2 < 1 \\ (II) : u &= -\frac{3}{2}vw, v \neq 0, w^2 = \frac{8}{9}. \end{aligned} \tag{3.26}$$

Theorem 12. *If $u = 0$, $v = 0$, $w^2 < \frac{8}{9}$, and $DEN > 0$, then $F = 0$ where $DEN =$ Denominator of \mathbf{J}_0 in (3.16).*

Proof: From (3.26) we note that $F = 0$ when

$$(I) : u = -\frac{3}{2}vw, v = 0, 0 \leq w^2 < 1 \quad \text{or}$$

(II) : $u = -\frac{3}{2}vw$, $v \neq 0$, $w^2 = \frac{8}{9}$.

Now, from (3.23) we note that under (II), $DEN = -\frac{v^2}{36} < 0$. Therefore, (II) is not possible. We also note that under (I), $DEN = -\frac{1}{4}(w^2 - \frac{8}{9})$ and thus if $w^2 \geq \frac{8}{9}$, then $DEN \leq 0$ but if $w^2 < \frac{8}{9}$, then $DEN > 0$. Therefore, $F = 0$ attains only when $u = v = 0$, and $w^2 < \frac{8}{9}$. This completes the proof of Theorem 12

Thus we conclude that $F_0 = (\text{Denominator of } \mathbf{J}_0 - \text{Numerator of } \mathbf{J}_0) \geq 0$ for all admissible values of (u, v, w) or equivalently of (t_1, t_2, t_4) . Clearly, \mathbf{J}_0 attains the maximum value 1, when $F_0 = 0$ i.e. when $F = 0$. We observed that $F = 0$ only when $u = 0$ and $v = 0$ equivalently when $t_1 = \frac{1}{3}$ and $t_2 = t_4$.

Now,

$$\begin{aligned} t_1 = \frac{1}{3} &\Leftrightarrow t_3 = \frac{2}{3} \\ \Leftrightarrow p_4 + p_1 &= \frac{1}{3} \end{aligned} \tag{3.27}$$

$$\text{and } p_3 + p_2 = \frac{2}{3} \tag{3.28}$$

$$\text{Also } t_4 = t_2 \Leftrightarrow p_3 - p_2 = p_4 - p_1. \tag{3.29}$$

Let us assume, $p_2 = p$.

$$\text{Then (3.28) provides, } p_3 = \frac{2}{3} - p,$$

$$\text{and (3.29) provides, } p_4 - p_1 = \frac{2}{3} - 2p. \tag{3.30}$$

$$\text{Now, (3.27) and (3.30) provide, } p_1 = p - \frac{1}{6} \quad \text{and} \quad p_4 = \frac{1}{2} - p.$$

As $p_i > 0$ for $i = 1, 2, 3, 4$; we have $p \in (\frac{1}{6}, \frac{1}{2})$. Thus there exists a subclass of

\mathbf{J} -optimal (or equivalently \mathbf{T} -optimal) designs within the class of designs \mathbf{D}_4 in (3.15) given in Table-3.2.

Table 3.2: Class of \mathbf{J} -optimal Designs

Design Points	$x_1 = -1$	$x_2 = -\frac{1}{2}$	$x_3 = \frac{1}{2}$	$x_4 = 1$
Proportions	$p_1 = p - \frac{1}{6}$	$p_2 = p$	$p_3 = \frac{2}{3} - p$	$p_4 = \frac{1}{2} - p$

where $p \in (\frac{1}{6}, \frac{1}{2})$.

Clearly, when $p_2 = p = \frac{1}{3}$ we get the special Dette-Titoff design d_2 in (2.24) which is in fact the most efficient design in the class of designs \mathbf{D}_1 in (2.19). We denote the new subclass of \mathbf{J} -optimal (or equivalently \mathbf{T} -optimal) designs in \mathbf{D}_4 by

$$\mathbf{D}_5 = \left\{ (x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}, x_4 = 1; \right. \\ \left. p_1 = p - \frac{1}{6}, p_2 = p, p_3 = \frac{2}{3} - p, p_4 = \frac{1}{2} - p, \frac{1}{6} < p < \frac{1}{2}; \text{Rank}(\mathbf{X}^{(2)}) = 4 \right\}. \quad (3.31)$$

We note that \mathbf{J} -optimal design may not be unique. Any design in \mathbf{D}_5 gives the same optimal value of the fitted value criterion making each design to be \mathbf{J} -optimal here. Hence we have the following theorem.

Theorem 13. *For the class of designs in \mathbf{D}_4 in (3.15), all the design in \mathbf{D}_5 are optimum with respect to the criterion \mathbf{J} (or equivalently \mathbf{T}).*

3.6 Expression of \mathbf{I} Criterion

Here we derive the \mathbf{I} Criterion under the class of designs \mathbf{D} in (1.3). We denote the predicted values of y when $X = x$ for MI and MII as $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ respectively.

Then we find the expression for $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$. We note that the quantity, $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$, may be positive for some values of x and negative for some other values of x . Therefore, we consider the squared value of $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ and find the criterion. Now, the expression of $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ in this case is readily available from (2.27). We just need to replace γ_2 by 0 and hence we get the expression. Also under the class of designs \mathbf{D}_1 in (2.19) the expression $[E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2$ takes the form $\gamma_3^2 x^2 (C - x^2)^2$ which is exactly the same as what we obtained in Chapter-2. Consequently the predicted value criterion will be identical to the case of Chapter-2 which is given by

$$\mathbf{I} = \gamma_3^2 \left[\frac{C_b^2}{3} - \frac{2C_b}{5} + \frac{1}{7} \right] \quad \text{where} \quad C_b = \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)}. \quad (3.32)$$

3.7 Efficient Designs with respect to \mathbf{I} Criterion

Since the expressions of \mathbf{I} for $\gamma_2 = 0$ (Chapter-3 in 3.32) and $\gamma_2 \neq 0$ (Chapter-2 in 2.29) are identical under the class of designs \mathbf{D}_1 , the theorems from Chapter-2 apply here. From Chapter-2 we note that d_2 in (2.24) (which is the special Dette-Titoff \mathbf{T} -optimal or equivalently \mathbf{J} -optimal design here) doesn't perform well with respect to criterion \mathbf{I} . We consider the class of designs \mathbf{D}_3 in (2.34) and obtain the following theorem.

Theorem 14.

- (a) *All the designs in \mathbf{D}_3 perform better than the special Dette-Titoff design d_2 with respect to the criterion \mathbf{I} .*
- (b) *The special Dette-Titoff design d_2 performs better than all the designs in \mathbf{D}_3 with respect to the criterion \mathbf{J} (or equivalently criterion \mathbf{T}).*

Chapter 4

Quadratic vs. Full Cubic as the True Model

4.1 Introduction

Here the response variable Y is dependent on the explanatory variable X with two possible dependence as described by two models MI, a quadratic regression model and MT, a full cubic regression model. We assume that MT is the true model but it is unknown to us. We do not know whether MI or MT describes the dependence better. Our goal is to discriminate between these two models MI and MT by the design choice.

4.2 Models and Associated Designs

We consider the class of designs D in (1.3). We also consider the full cubic model from (1.1) as the true model and denote it by MT. Our aim is to discriminate between the two models MI and MT assuming MT to be the unknown true model.

The two models considered here are given by

$$\begin{aligned} \text{MI:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2, \\ \text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0, \end{aligned} \quad (4.1)$$

$$\begin{aligned} \text{MT:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \gamma_3 x_i^3, \\ \text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0, \end{aligned} \quad (4.2)$$

where $i, i' = 1, 2, 3, 4$; $j = 1, 2, \dots, n_i, j' = 1, 2, \dots, n_{i'}, (i, j) \neq (i', j')$. The matrix representations of MT and MI are given by

$$E(\mathbf{y}) = \mathbf{X}^{(1)}\boldsymbol{\gamma}^{(1)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \quad (4.3)$$

$$E(\mathbf{y}) = \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \quad (4.4)$$

where

$$\mathbf{X}^{(1)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(1)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix},$$

and

$$\mathbf{X}^{(t)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(t)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}.$$

We define

$$\mathbf{X}_3 = \left(x_1^3 \quad \dots \quad x_1^3 \quad x_2^3 \quad \dots \quad x_2^3 \quad x_3^3 \quad \dots \quad x_3^3 \quad x_4^3 \quad \dots \quad x_4^3 \right)'$$

We note that

$$\mathbf{X}^{(t)} = \left(\mathbf{X}^{(1)} \quad \vdots \quad \mathbf{X}_3 \right) \Rightarrow \mathbf{X}^{(t)'}\mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_3 \\ \mathbf{X}_3'\mathbf{X}^{(1)} & \mathbf{X}_3'\mathbf{X}_3 \end{pmatrix}. \quad (4.5)$$

4.3 Expression of J Criterion

We first consider the fitted value criterion J (ref. Chapter-1). We find the expression of J considering the class of designs D in (1.3) for model selection and discrimination purposes.

It can be checked that

$$E(\hat{\gamma}^{(t)}) = (\mathbf{X}^{(t)'}\mathbf{X}^{(t)})^{-1}\mathbf{X}^{(t)'}E(\mathbf{y}) = \gamma^{(t)}, \quad (4.6)$$

and

$$\begin{aligned} E(\hat{\gamma}^{(1)}) &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}E(\mathbf{y}) \\ &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}^{(t)}\gamma^{(t)} \\ &= \begin{pmatrix} 1 & 0 & 0 & A \\ 0 & 1 & 0 & B \\ 0 & 0 & 1 & C \end{pmatrix} \gamma^{(t)}, \end{aligned} \quad (4.7)$$

where

$$\begin{aligned}
A = \frac{1}{\text{Det}_1} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 x_2 x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 x_2 x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 x_3 x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 x_3 x_4) \right],
\end{aligned}$$

$$\begin{aligned}
B = \frac{-1}{\text{Det}_1} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 x_2 + x_1 x_3 + x_2 x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 x_2 + x_1 x_4 + x_2 x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 x_3 + x_1 x_4 + x_3 x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 x_3 + x_2 x_4 + x_3 x_4) \right],
\end{aligned}$$

$$\begin{aligned}
C = \frac{1}{\text{Det}_1} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) \right],
\end{aligned}$$

and

$$\begin{aligned}
\text{Det}_1 = |\mathbf{X}^{(1)'} \mathbf{X}^{(1)}| & = \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 \right]. \tag{4.8}
\end{aligned}$$

The fitted values of \mathbf{y} under MI and MT are expressed as $\hat{\mathbf{y}}^{(1)} = \mathbf{X}^{(1)}\hat{\boldsymbol{\gamma}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)} = \mathbf{X}^{(t)}\hat{\boldsymbol{\gamma}}^{(t)}$. Hence we have

$$\begin{aligned} E(\hat{\mathbf{y}}^{(1)}) &= \mathbf{X}^{(1)}E(\hat{\boldsymbol{\gamma}}^{(1)}) \\ &= \mathbf{X}^{(1)}(\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\ &= \mathbf{H}_1\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \end{aligned} \quad (4.9)$$

$$\begin{aligned} E(\hat{\mathbf{y}}^{(2)}) &= \mathbf{X}^{(t)}E(\hat{\boldsymbol{\gamma}}^{(t)}) \\ &= \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \end{aligned} \quad (4.10)$$

$$\Rightarrow E(\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)}) = (\mathbf{I} - \mathbf{H}_1)\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \quad (4.11)$$

where $\mathbf{H}_1 = \mathbf{X}^{(1)}(\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}$ and $\mathbf{H}_1' = \mathbf{H}_1$, $\mathbf{H}_1^2 = \mathbf{H}_1$. Now, using (4.11) we get

$$\begin{aligned} \mathbf{J} &= E(\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)})'E(\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)}) \\ &= \boldsymbol{\gamma}^{(t)'}\mathbf{X}^{(t)'}(\mathbf{I} - \mathbf{H}_1)\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}. \end{aligned} \quad (4.12)$$

It can be checked that

$$\mathbf{X}^{(t)'}\mathbf{H}_1\mathbf{X}^{(t)} = \begin{pmatrix} \sum_{i=1}^4 n_i & \sum_{i=1}^4 n_i x_i & \sum_{i=1}^4 n_i x_i^2 & \sum_{i=1}^4 n_i x_i^3 \\ \sum_{i=1}^4 n_i x_i & \sum_{i=1}^4 n_i x_i^2 & \sum_{i=1}^4 n_i x_i^3 & \sum_{i=1}^4 n_i x_i^4 \\ \sum_{i=1}^4 n_i x_i^2 & \sum_{i=1}^4 n_i x_i^3 & \sum_{i=1}^4 n_i x_i^4 & \sum_{i=1}^4 n_i x_i^5 \\ \sum_{i=1}^4 n_i x_i^3 & \sum_{i=1}^4 n_i x_i^4 & \sum_{i=1}^4 n_i x_i^5 & \mathbf{X}_3'\mathbf{H}_1\mathbf{X}_3 \end{pmatrix}. \quad (4.13)$$

Therefore

$$\mathbf{X}^{(t)'}(\mathbf{I} - \mathbf{H}_1)\mathbf{X}^{(t)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & P \end{pmatrix}, \quad (4.14)$$

and thus using (4.12) and (4.14) we get the fitted value criterion as

$$\mathbf{J} = P\gamma_3^2, \quad (4.15)$$

where

$$\begin{aligned} P &= \mathbf{X}_3'(\mathbf{I} - \mathbf{H}_1)\mathbf{X}_3 \\ &= \frac{n_1 n_2 n_3 n_4 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2}{\left[n_1 n_2 (x_1 - x_2)^2 \left\{ n_3 (x_1 - x_3)^2 (x_2 - x_3)^2 + n_4 (x_1 - x_4)^2 (x_2 - x_4)^2 \right\} \right. \\ &\quad \left. + n_3 n_4 (x_3 - x_4)^2 \left\{ n_1 (x_1 - x_3)^2 (x_1 - x_4)^2 + n_2 (x_2 - x_3)^2 (x_2 - x_4)^2 \right\} \right]}. \end{aligned} \quad (4.16)$$

We now calculate the \mathbf{J} criterion under the class of designs \mathbf{D}_1 in (2.19). Thus \mathbf{J} in (4.15) reduces to

$$\begin{aligned} \mathbf{J} &= \frac{2\gamma_3^2 n_1 n_2 a^2 b^2 (b^2 - a^2)^2}{(n_1 b^2 + n_2 a^2)} \\ \Leftrightarrow \frac{\mathbf{J}}{n\gamma_3^2} &= \frac{2p_1(1 - 2p_1)a^2 b^2 (b^2 - a^2)^2}{2p_1 b^2 + (1 - 2p_1)a^2} \quad \text{where} \quad (p_1 + p_2) = \frac{1}{2}. \end{aligned} \quad (4.17)$$

4.4 Efficient Designs with respect to J Criterion

It is interesting to note that under the class of designs \mathbf{D}_1 in (2.19) the expression of \mathbf{J} in this chapter is identical to the expression of \mathbf{J} in Chapter-2 (in 2.20) and also in Chapter-3 (in 3.11). Hence from Chapter-2 and 3 we obtain the following theorems.

Theorem 15. *For a given value of b in the class of designs \mathbf{D}_1 in (2.19), the design d_1 in (2.21) is optimum with respect to the criterion \mathbf{J} .*

We note that d_2 in (2.24) becomes the Kiefer and Wolfowitz (1959) optimal design in the setup of Chapter-4. Hence we get following theorem.

Theorem 16. *For the class of designs \mathbf{D}_1 , the Kiefer-Wolfowitz design d_2 is optimum with respect to the criterion \mathbf{J} .*

4.5 Expression of I Criterion

We now consider the predicted value criterion here. Let $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ be the predicted values of the two models MI and MT at $X = x$. We first find the expression for $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ assuming MT to be the unknown true model. Then we obtain the squared value of $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ and find the criterion.

Under the general class of designs \mathbf{D} in (1.3) we have

$$\begin{aligned}
E(\hat{y}^{(1)}(x)) &= \begin{pmatrix} 1 & x & x^2 \end{pmatrix} E(\hat{\gamma}^{(1)}) \\
&= \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & A \\ 0 & 1 & 0 & B \\ 0 & 0 & 1 & C \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (4.7)} \\
&= \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \gamma_3(A + Bx + Cx^2), \tag{4.18}
\end{aligned}$$

and

$$\begin{aligned}
E(\hat{y}^{(2)}(x)) &= \begin{pmatrix} 1 & x & x^2 & x^3 \end{pmatrix} E(\hat{\gamma}^{(t)}) \\
&= \begin{pmatrix} 1 & x & x^2 & x^3 \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (4.6)} \\
&= \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \gamma_3 x^3, \tag{4.19}
\end{aligned}$$

where A, B , and C are defined right after (4.7).

Thus from (4.18) and (4.19) we get

$$E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x)) = \gamma_3 [A + Bx + Cx^2 - x^3] \tag{4.20}$$

Now under \mathbf{D}_1 in (2.19) we have $A = C = 0$ and $B = \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)}$. Hence under this class of designs we get from (4.20),

$$[E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 = \gamma_3^2 x^2 (B - x^2)^2. \tag{4.21}$$

The discrimination between the models MI and MT will be the best when

$\left[\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_3}\right]^2$ is maximum. The discrimination between the models MI and MT will not be possible with respect to predicted values when $\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_3} = 0$. We observe that at $x = 0$, $x = -\sqrt{B}$, and $x = \sqrt{B}$, $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x)) = 0$. Clearly, at $x = 0$, $x = -\sqrt{B}$, and $x = \sqrt{B}$, the discrimination between MI and MT will not be possible using the prediction criterion at the design stage. We now consider an overall prediction measure by taking the weighted average of $\left[\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_3}\right]^2$ on x . Therefore the predicted value criterion is given by,

$$\begin{aligned}
\mathbf{I} &= \frac{1}{2} \int_{-1}^1 [E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 dx \\
&= \frac{\gamma_3^2}{2} \int_{-1}^1 (Bx - x^3)^2 dx \\
&= \gamma_3^2 \left[\frac{B^2}{3} - \frac{2B}{5} + \frac{1}{7} \right] \quad \text{where} \quad 0 < B = \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)} < 1.
\end{aligned} \tag{4.22}$$

4.6 Efficient Designs with respect to \mathbf{I} Criterion

Since the expressions of \mathbf{I} in (4.22) of Chapter-4 and in (2.29) of Chapter-2 are identical under the class of designs \mathbf{D}_1 , the theorems from Chapter-2 apply here. From Chapter-2 we note that d_2 in (2.24) (the special Dette-Titoff \mathbf{T} -optimal or equivalently \mathbf{J} -optimal design) doesn't perform well with respect to criterion \mathbf{I} . We consider the class of designs \mathbf{D}_3 in (2.34) and obtain the following theorem.

Theorem 17.

- (a) *All the designs in \mathbf{D}_3 perform better than the Kiefer-Wolfowitz design d_2 with respect to the criterion \mathbf{I} .*
- (b) *The Kiefer-Wolfowitz design d_2 performs better than all the designs in \mathbf{D}_3 with respect to the criterion \mathbf{J} (or equivalently criterion \mathbf{T}).*

Chapter 5

Linear vs. Quadratic when the True Model Is Full Cubic

5.1 Introduction

We consider an experiment where the response variable Y is dependent on an explanatory variable X by two models MI, a simple linear regression model and MII, a quadratic regression model. We assume that the full cubic model MT is the unknown true model. Our goal is to discriminate between the two models MI and MII at the design stage assuming MT to be the unknown true model.

5.2 Models and Associated Designs

We consider the class of designs D in (1.3). We also consider the full cubic model from (1.1) as the unknown true model and denote it by MT. The three models considered here are given by

$$\begin{aligned}
\text{MT:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \gamma_3 x_i^3, \\
\text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
\end{aligned} \tag{5.1}$$

$$\begin{aligned}
\text{MI:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i, \\
\text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
\end{aligned} \tag{5.2}$$

$$\begin{aligned}
\text{MII:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2, \\
\text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
\end{aligned} \tag{5.3}$$

where $i, i' = 1, 2, 3, 4$; $j = 1, 2, \dots, n_i, j' = 1, 2, \dots, n_{i'}, (i, j) \neq (i', j')$.

The matrix representations of MT, MI, and MII are given by

$$E(\mathbf{y}) = \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{5.4}$$

$$E(\mathbf{y}) = \mathbf{X}^{(1)}\boldsymbol{\gamma}^{(1)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{5.5}$$

$$E(\mathbf{y}) = \mathbf{X}^{(2)}\boldsymbol{\gamma}^{(2)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{5.6}$$

where

$$\mathbf{X}^{(1)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(1)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix},$$

$$\mathbf{X}^{(2)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(2)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix},$$

$$\mathbf{X}^{(t)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(t)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}.$$

We define

$$\mathbf{X}_2 = \begin{pmatrix} x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}',$$

$$\mathbf{X}_3 = \begin{pmatrix} x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}'$$

and

$$\mathbf{X}_{23} = \begin{pmatrix} x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}'.$$

Hence we get,

$$\begin{aligned} \mathbf{X}^{(2)} &= \begin{pmatrix} \mathbf{X}^{(1)} & : & \mathbf{X}_2 \end{pmatrix} \Rightarrow \mathbf{X}^{(2)'}\mathbf{X}^{(2)} = \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}^{(1)} & \mathbf{X}_2'\mathbf{X}_2 \end{pmatrix}, \\ \mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(2)} & : & \mathbf{X}_3 \end{pmatrix} \Rightarrow \mathbf{X}^{(t)'}\mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{X}^{(2)'}\mathbf{X}^{(2)} & \mathbf{X}^{(2)'}\mathbf{X}_3 \\ \mathbf{X}_3'\mathbf{X}^{(2)} & \mathbf{X}_3'\mathbf{X}_3 \end{pmatrix}, \\ \mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(1)} & : & \mathbf{X}_{23} \end{pmatrix} \Rightarrow \mathbf{X}^{(t)'}\mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{23} \\ \mathbf{X}_{23}'\mathbf{X}^{(1)} & \mathbf{X}_{23}'\mathbf{X}_{23} \end{pmatrix}. \end{aligned} \quad (5.7)$$

5.3 Expression of J Criterion

We first consider the fitted value criterion J (ref. Chapter-1) under the general class of designs D in (1.3) for model selection and discrimination purposes. Let $\hat{\mathbf{y}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)}$ be the fitted values of the two models MI and MII. First we have to find the expression of $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$ where the expectation is considered under the true model MT in (5.1).

The least square estimate of $\boldsymbol{\gamma}^{(1)}$ for MI is given by (Rao (1973))

$$\hat{\boldsymbol{\gamma}}^{(1)} = \begin{pmatrix} \hat{\gamma}_0^{(1)} \\ \hat{\gamma}_1^{(1)} \end{pmatrix} = (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{y}, \quad (5.8)$$

and the least square estimate of $\boldsymbol{\gamma}^{(2)}$ for MII is given by

$$\hat{\boldsymbol{\gamma}}^{(2)} = \begin{pmatrix} \hat{\gamma}_0^{(2)} \\ \hat{\gamma}_1^{(2)} \\ \hat{\gamma}_2^{(2)} \end{pmatrix} = (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{y}. \quad (5.9)$$

Now,

$$\begin{aligned} E(\hat{\boldsymbol{\gamma}}^{(1)}) &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}E(\mathbf{y}) \\ &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\ &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1} \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{23} \end{pmatrix} \boldsymbol{\gamma}^{(t)} \\ &= \begin{pmatrix} \mathbf{I}_2 & (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}_{23} \end{pmatrix} \boldsymbol{\gamma}^{(t)} \\ &= \begin{pmatrix} 1 & 0 & B & A \\ 0 & 1 & D & C \end{pmatrix} \boldsymbol{\gamma}^{(t)}, \end{aligned} \quad (5.10)$$

and

$$\begin{aligned}
E(\hat{\gamma}^{(2)}) &= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}E(\mathbf{y}) \\
&= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\
&= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\begin{pmatrix} \mathbf{X}^{(2)'}\mathbf{X}^{(2)} & \mathbf{X}^{(2)'}\mathbf{X}_3 \end{pmatrix}\boldsymbol{\gamma}^{(t)} \\
&= \begin{pmatrix} \mathbf{I}_3 & (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{X}_3 \end{pmatrix}\boldsymbol{\gamma}^{(t)} \\
&= \begin{pmatrix} 1 & 0 & 0 & E \\ 0 & 1 & 0 & F \\ 0 & 0 & 1 & G \end{pmatrix}\boldsymbol{\gamma}^{(t)}, \tag{5.11}
\end{aligned}$$

where A, B, C , and D are defined right after (2.11) and

$$\begin{aligned}
E &= \frac{1}{\text{Det}_2} \left[n_1n_2n_3(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2x_1x_2x_3 \right. \\
&\quad + n_1n_2n_4(x_1 - x_2)^2(x_1 - x_4)^2(x_2 - x_4)^2x_1x_2x_4 \\
&\quad + n_1n_3n_4(x_1 - x_3)^2(x_1 - x_4)^2(x_3 - x_4)^2x_1x_3x_4 \\
&\quad \left. + n_2n_3n_4(x_2 - x_3)^2(x_2 - x_4)^2(x_3 - x_4)^2x_2x_3x_4 \right],
\end{aligned}$$

$F =$

$$\begin{aligned}
&\frac{-1}{\text{Det}_2} \left[n_1n_2n_3(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2(x_1x_2 + x_1x_3 + x_2x_3) \right. \\
&\quad + n_1n_2n_4(x_1 - x_2)^2(x_1 - x_4)^2(x_2 - x_4)^2(x_1x_2 + x_1x_4 + x_2x_4) \\
&\quad + n_1n_3n_4(x_1 - x_3)^2(x_1 - x_4)^2(x_3 - x_4)^2(x_1x_3 + x_1x_4 + x_3x_4) \\
&\quad \left. + n_2n_3n_4(x_2 - x_3)^2(x_2 - x_4)^2(x_3 - x_4)^2(x_2x_3 + x_2x_4 + x_3x_4) \right],
\end{aligned}$$

$$\begin{aligned}
G = \frac{1}{\text{Det}_2} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) \right],
\end{aligned}$$

$$\begin{aligned}
\text{Det}_2 = |\mathbf{X}^{(2)'} \mathbf{X}^{(2)}| & = \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 \right].
\end{aligned}$$

We know that the fitted values of \mathbf{y} under MI and MII are expressed as $\hat{\mathbf{y}}^{(1)} = \mathbf{X}^{(1)} \hat{\boldsymbol{\gamma}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)} = \mathbf{X}^{(2)} \hat{\boldsymbol{\gamma}}^{(2)}$. Now assuming that MT to be the true model the expected values of the fitted values are given by

$$\begin{aligned}
E(\hat{\mathbf{y}}^{(1)}) & = \mathbf{X}^{(1)} E(\hat{\boldsymbol{\gamma}}^{(1)}) \\
& = \mathbf{X}^{(1)} (\mathbf{X}^{(1)'} \mathbf{X}^{(1)})^{-1} \mathbf{X}^{(1)'} \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\
& = \mathbf{H}_1 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \tag{5.12}
\end{aligned}$$

$$\begin{aligned}
E(\hat{\mathbf{y}}^{(2)}) & = \mathbf{X}^{(2)} E(\hat{\boldsymbol{\gamma}}^{(2)}) \\
& = \mathbf{X}^{(2)} (\mathbf{X}^{(2)'} \mathbf{X}^{(2)})^{-1} \mathbf{X}^{(2)'} \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\
& = \mathbf{H}_2 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \tag{5.13}
\end{aligned}$$

$$\Rightarrow E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)}) = (\mathbf{H}_1 - \mathbf{H}_2) \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \tag{5.14}$$

where $\mathbf{H}_1 = \mathbf{X}^{(1)}(\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}$ and $\mathbf{H}_2 = \mathbf{X}^{(2)}(\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}$. We note that \mathbf{H}_1 and \mathbf{H}_2 are symmetric and idempotent (Rao (1973)) i.e. $\mathbf{H}_1' = \mathbf{H}_1$, $\mathbf{H}_1^2 = \mathbf{H}_1$, $\mathbf{H}_2' = \mathbf{H}_2$ and $\mathbf{H}_2^2 = \mathbf{H}_2$. Also by Result 1 we have $\mathbf{H}_2\mathbf{H}_1 = \mathbf{H}_1 = \mathbf{H}_1\mathbf{H}_2$. Clearly using (5.14) we get

$$\begin{aligned}
\mathbf{J} &= E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})' E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)}) \\
&= \boldsymbol{\gamma}^{(t)'} \mathbf{X}^{(t)'} (\mathbf{H}_1 - \mathbf{H}_2)^2 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\
&= \boldsymbol{\gamma}^{(t)'} \mathbf{X}^{(t)'} (\mathbf{H}_2 - \mathbf{H}_1) \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \quad \text{using Result-1.} \tag{5.15}
\end{aligned}$$

Now,

$$\begin{aligned}
\mathbf{X}^{(t)'} \mathbf{H}_2 \mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(2)'} \\ \mathbf{X}_3' \end{pmatrix} \mathbf{H}_2 \begin{pmatrix} \mathbf{X}^{(2)} & \mathbf{X}_3 \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{X}^{(2)'} \\ \mathbf{X}_3' \mathbf{H}_2 \end{pmatrix} \begin{pmatrix} \mathbf{X}^{(2)} & \mathbf{X}_3 \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{X}^{(2)'} \mathbf{X}^{(2)} & \mathbf{X}^{(2)'} \mathbf{X}_3 \\ \mathbf{X}_3' \mathbf{X}^{(2)} & \mathbf{X}_3' \mathbf{H}_2 \mathbf{X}_3 \end{pmatrix} \\
&= \begin{pmatrix} \sum_1^4 n_i & \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 \\ \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 \\ \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 & \sum_1^4 n_i x_i^5 \\ \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 & \sum_1^4 n_i x_i^5 & S \end{pmatrix}, \tag{5.16}
\end{aligned}$$

$$\begin{aligned}
\mathbf{X}^{(t)'}\mathbf{H}_1\mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(1)'} \\ \mathbf{X}_{23}' \end{pmatrix} \mathbf{H}_1 \begin{pmatrix} \mathbf{X}^{(1)} & \mathbf{X}_{23} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{X}^{(1)'} \\ \mathbf{X}_{23}'\mathbf{H}_1 \end{pmatrix} \begin{pmatrix} \mathbf{X}^{(1)} & \mathbf{X}_{23} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{23} \\ \mathbf{X}_{23}'\mathbf{X}^{(1)} & \mathbf{X}_{23}'\mathbf{H}_1\mathbf{X}_{23} \end{pmatrix} \\
&= \begin{pmatrix} \sum_1^4 n_i & \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 \\ \sum_1^4 n_i x_i & \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 \\ \sum_1^4 n_i x_i^2 & \sum_1^4 n_i x_i^3 & P & R \\ \sum_1^4 n_i x_i^3 & \sum_1^4 n_i x_i^4 & R & Q \end{pmatrix}, \tag{5.17}
\end{aligned}$$

where

$$P = \mathbf{X}_2'\mathbf{H}_1\mathbf{X}_2 = B \sum_1^4 n_i x_i^2 + D \sum_1^4 n_i x_i^3,$$

$$Q = \mathbf{X}_3'\mathbf{H}_1\mathbf{X}_3 = A \sum_1^4 n_i x_i^3 + C \sum_1^4 n_i x_i^4,$$

$$R = \mathbf{X}_2'\mathbf{H}_1\mathbf{X}_3 = A \sum_1^4 n_i x_i^2 + C \sum_1^4 n_i x_i^3,$$

$$S = \mathbf{X}_3'\mathbf{H}_2\mathbf{X}_3 = E \sum_1^4 n_i x_i^3 + F \sum_1^4 n_i x_i^4 + G \sum_1^4 n_i x_i^5,$$

and $A, B, C,$ and D are defined right after (2.11) and $E, F,$ and G are defined right after (5.11).

Now, using (5.15), (5.16), and (5.17) we get

$$\begin{aligned} \mathbf{J} &= \boldsymbol{\gamma}^{(t)'} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sum_1^4 n_i x_i^4 - P & \sum_1^4 n_i x_i^5 - R \\ 0 & 0 & \sum_1^4 n_i x_i^5 - R & S - Q \end{pmatrix} \boldsymbol{\gamma}^{(t)} \\ &= \left(\sum_1^4 n_i x_i^4 - P \right) \gamma_2^2 + (S - Q) \gamma_3^2 + 2\gamma_2 \gamma_3 \left(\sum_1^4 n_i x_i^5 - R \right), \end{aligned} \quad (5.18)$$

where $P, Q, R,$ and S are defined right after (5.17).

Now we consider class of designs \mathbf{D}_1 in (2.19). We note that under \mathbf{D}_1 we have

$$\begin{aligned} Det_1 &= 2n(b^2 n_1 + a^2 n_2), & Det_2 &= 8(b^2 - a^2)^2 n_1 n_2 (b^2 n_1 + a^2 n_2), & A &= 0, \\ B &= \frac{2}{n}(b^2 n_1 + a^2 n_2), & C &= \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)}, & D &= 0, \\ E &= 0, & F &= \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)}, & G &= 0, \\ P &= \frac{4}{n}(b^2 n_1 + a^2 n_2)^2, & Q &= \frac{2(b^4 n_1 + a^4 n_2)^2}{(b^2 n_1 + a^2 n_2)} = S, & R &= 0, \end{aligned}$$

and hence the fitted value criterion in (5.18) is given by

$$\begin{aligned} \mathbf{J} &= \frac{4}{n} \gamma_2^2 n_1 n_2 (b^2 - a^2)^2 \\ \Leftrightarrow \frac{\mathbf{J}}{n \gamma_2^2} &= 2p_1(1 - 2p_1)(b^2 - a^2)^2. \end{aligned} \quad (5.19)$$

5.4 Efficient Designs with respect to J Criterion

For any given a and b , \mathbf{J} in (5.19) is maximized at $p_1 = \frac{1}{4}$. We define the design d_5 in \mathbf{D}_1 in (2.19) as the following:

$$d_5 = \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -b, x_2 = -a, x_3 = a, x_4 = b; \right. \\ \left. n_1 = n_2 = n_3 = n_4 = \frac{n}{4}, Rank(\mathbf{X}^{(t)}) = 4 \right]. \quad (5.20)$$

Thus we have the following theorem.

Theorem 18. *For any given a and b in the class of designs \mathbf{D}_1 , the design d_5 is optimum with respect to the criterion \mathbf{J} .*

Also for a given a , \mathbf{J} in (5.19) is maximized when $p_1 = \frac{1}{4}$ and $b = 1$. We define design d_6 in \mathbf{D}_1 as the following:

$$d_6 = \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -a, x_3 = a, x_4 = 1; \right. \\ \left. n_1 = n_2 = n_3 = n_4 = \frac{n}{4}, Rank(\mathbf{X}^{(t)}) = 4 \right]. \quad (5.21)$$

Hence obtain the following theorem.

Theorem 19. *For a given value of a in the class of designs \mathbf{D}_1 , the design d_6 is optimum with respect to the criterion \mathbf{J} .*

Special Note: When $\gamma_3 = 0$ in MT, then MT and MII are identical models representing the full quadratic model. In that case the problem reduces to discrimination between a linear and a quadratic model and we need at least 3 distinct design points for discrimination purpose. We note that this setup is identical to the

setup explained in Ghosh and Pal (2008). From the expression of \mathbf{J} in (5.19) we note that $2p_1(1 - 2p_1)$ is maximum when $p_1 = \frac{1}{4}$ and $(b^2 - a^2)^2$ is maximum when $b = 1$ and $a = 0$ and thus the \mathbf{J} is maximum for the design with $x_1 = -1, x_2 = 0$, and $x_3 = 1$ with replications $n_1 = \frac{n}{4}, n_2 = \frac{n}{2}$, and $n_3 = \frac{n}{4}$. This is the Kiefer and Wolfowitz (1959) optimal design for estimating γ_2 .

5.5 Expression of I Criterion

Here we consider the predicted value criterion under the general class of designs \mathbf{D} in (1.3). Let $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ be the predicted values of the two models MI and MII at $X = x$. We first find the expression for $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ assuming MT to be the true model. We know that

$$\begin{aligned} E(\hat{y}^{(1)}(x)) &= \begin{pmatrix} 1 & x \end{pmatrix} E(\hat{\gamma}^{(1)}) \\ &= \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} 1 & 0 & B & A \\ 0 & 1 & D & C \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (5.10)} \\ &= \gamma_0 + \gamma_1 x + \gamma_2(B + Dx) + \gamma_3(A + Cx), \end{aligned} \tag{5.22}$$

and

$$\begin{aligned} E(\hat{y}^{(2)}(x)) &= \begin{pmatrix} 1 & x & x^2 \end{pmatrix} E(\hat{\gamma}^{(2)}) \\ &= \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & E \\ 0 & 1 & 0 & F \\ 0 & 0 & 1 & G \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (5.11)} \\ &= \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \gamma_3(E + Fx + Gx^2), \end{aligned} \tag{5.23}$$

where A, B, C , and D are defined right after (2.11) and E, F , and G are defined right after (5.11). Thus from (5.22) and (5.23) we get

$$E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x)) = \gamma_2 [B + Dx - x^2] + \gamma_3 [(A - E) + (C - F)x - Gx^2] \quad (5.24)$$

Now considering the class of designs \mathbf{D}_1 in (2.19), we have $B = \frac{2}{n}(b^2n_1 + a^2n_2)$, $A = D = E = G = 0$, and $C = F = \frac{(b^4n_1 + a^4n_2)}{(b^2n_1 + a^2n_2)}$. Hence under \mathbf{D}_1 we have

$$[E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 = \gamma_2^2 (B - x^2)^2. \quad (5.25)$$

The discrimination between the models MI and MII will be the best when $\left[\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_2}\right]^2$ is maximum. The discrimination between the models MI and MII will not be possible with respect to predicted values when $\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_2} = 0$. We observe that at $x = -\sqrt{B}$ and $x = \sqrt{B}$, $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x)) = 0$. Clearly, at $x = -\sqrt{B}$ and $x = \sqrt{B}$, the discrimination between MI and MII will not be possible using the prediction criterion at the design stage. But we are interested in finding an overall prediction measure for discrimination purpose rather than evaluating at each x . So, we consider an overall measure by taking the weighted average of $\left[\frac{E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))}{\gamma_2}\right]^2$ on x . Therefore the predicted value criterion is given by,

$$\begin{aligned} \mathbf{I} &= \frac{\gamma_2^2}{2} \int_{-1}^1 (B - x^2)^2 dx \\ &= \gamma_2^2 \left(B^2 - \frac{2B}{3} + \frac{1}{5} \right) \end{aligned} \quad (5.26)$$

where $0 < B = \frac{2}{n}(b^2n_1 + a^2n_2) < 1$.

5.6 Efficient Designs with respect to I Criterion

We note that I in (5.26) is a concave function with respect to B (Fig-5.1). Designs with a very high or a very small value of B will perform well with respect to I . But designs with very high values of B are better than those with very small values

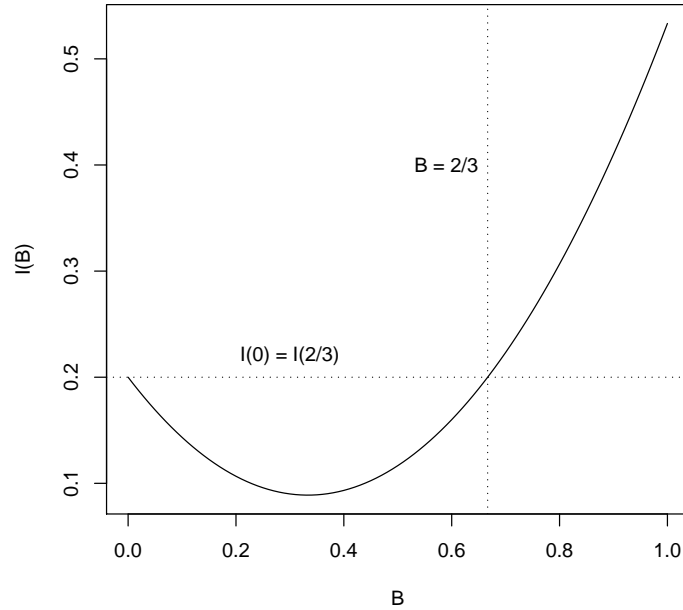


Figure 5.1: Plot of $I(B)$ against B

of B . We also note that $I(0) = I(\frac{2}{3})$ but it is clear that $C \neq 1$ and $C \neq 0$. We fix $b = 1$ and obtain some choices of a and p_1 numerically to present some designs (Table -5.1) in D_1 in (2.19) which perform well with respect to I .

We note that the designs those perform well with respect to I , do not perform well with respect to J and vice versa. As expected the designs with high values of B perform better than those with smaller values of B with respect to I . As the value of p_1 gets closer to 0.25 the performance of the design with respect to J is

Table 5.1: Some \mathbf{I} -optimal Designs

B	b	a	p_1	p_2	\mathbf{J}	\mathbf{I}
0.970	1	0.707	0.470	0.030	0.01410	0.49423
0.970	1	0.754	0.465	0.035	0.01206	0.49423
0.970	1	0.800	0.458	0.042	0.00990	0.49423
0.900	1	0.535	0.430	0.070	0.06143	0.41000
0.900	1	0.561	0.427	0.073	0.05856	0.41000
0.900	1	0.600	0.422	0.078	0.05400	0.41000
0.800	1	0.200	0.396	0.104	0.15200	0.30667
0.800	1	0.300	0.390	0.110	0.14200	0.30667
0.800	1	0.400	0.381	0.119	0.12800	0.30667
0.800	1	0.500	0.367	0.133	0.11000	0.30667
0.800	1	0.600	0.344	0.156	0.08800	0.30667
0.625	1	0.500	0.250	0.250	0.14063	0.17396

the best. We define the following subclass in \mathbf{D}_1

$$\mathbf{D}_6 = \left\{ (x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -a, x_3 = a, x_4 = 1; \right. \\ \left. \frac{1}{4} < p_1 = p_4 < \frac{1}{2}, p_2 = p_3, p_1 + p_2 = \frac{1}{2}, \text{Rank}(\mathbf{X}^{(2)}) = 4 \right\}. \quad (5.27)$$

We obtain the following theorem.

Theorem 20.

- (a) For a given value of a , all the designs in \mathbf{D}_6 perform better than the special d_6 in (5.21) with respect to the criterion \mathbf{I} .
- (b) For a given value of a , the design d_2 performs better than all the designs in \mathbf{D}_3 with respect to the criterion \mathbf{J} .

Chapter 6

Linear vs. Full Cubic as the True Model

6.1 Introduction

Here the response variable Y is dependent on the explanatory variable X with two possible dependence as described by two models MI, a simple linear regression model and MT, a full cubic regression model. We assume that MT is the true model but it is unknown to us. We do not know whether MI or MT describes the dependence better. Our goal is to discriminate between these two models MI and MT by the design choice.

6.2 Models and Associated Designs

We consider the class of designs D in (1.3). We also consider the full cubic model from (1.1) as the true model and denote it by MT. Our aim is to discriminate between the two models MI and MT assuming MT to be the unknown true model.

The two models considered here are given by

$$\begin{aligned} \mathbf{MI}: \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i, \\ \text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0, \end{aligned} \quad (6.1)$$

$$\begin{aligned} \mathbf{MT}: \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \gamma_3 x_i^3, \\ \text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0, \end{aligned} \quad (6.2)$$

where $i, i' = 1, 2, 3, 4$; $j = 1, 2, \dots, n_i$, $j' = 1, 2, \dots, n_{i'}$, $(i, j) \neq (i', j')$. The matrix representations of MT and MI are given by

$$E(\mathbf{y}) = \mathbf{X}^{(1)}\boldsymbol{\gamma}^{(1)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \quad (6.3)$$

$$E(\mathbf{y}) = \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \quad (6.4)$$

where

$$\mathbf{X}^{(1)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(1)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix},$$

$$\mathbf{X}^{(t)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(t)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}.$$

We define

$$\begin{aligned}\mathbf{X}_2 &= \begin{pmatrix} x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}', \\ \mathbf{X}_3 &= \begin{pmatrix} x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \\ \mathbf{X}_{23} &= \begin{pmatrix} x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}'.\end{aligned}$$

We note that

$$\mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{X}^{(1)} & \vdots & \mathbf{X}_{23} \end{pmatrix} \Rightarrow \mathbf{X}^{(t)'}\mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{23} \\ \mathbf{X}_{23}'\mathbf{X}^{(1)} & \mathbf{X}_{23}'\mathbf{X}_{23} \end{pmatrix}. \quad (6.5)$$

6.3 Expression of J Criterion

We first consider the fitted value criterion J (ref. Chapter-1). We find the expression of J considering the class of designs D in (1.3) for model selection and discrimination purposes. It can be checked that

$$E(\hat{\gamma}^{(t)}) = (\mathbf{X}^{(t)'}\mathbf{X}^{(t)})^{-1}\mathbf{X}^{(t)'}E(\mathbf{y}) = \gamma^{(t)}, \quad (6.6)$$

and

$$\begin{aligned}E(\hat{\gamma}^{(1)}) &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}E(\mathbf{y}) \\ &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}^{(t)}\gamma^{(t)} \\ &= \begin{pmatrix} 1 & 0 & A & B \\ 0 & 1 & C & D \end{pmatrix} \gamma^{(t)},\end{aligned} \quad (6.7)$$

where

$$A = \frac{1}{Det_1} \left[\left(\sum_{i=1}^4 n_i x_i^2 \right)^2 - \sum_{i=1}^4 n_i x_i \sum_{i=1}^4 n_i x_i^3 \right],$$

$$B = \frac{1}{Det_1} \left[\sum_{i=1}^4 n_i x_i^2 \sum_{i=1}^4 n_i x_i^3 - \sum_{i=1}^4 n_i x_i \sum_{i=1}^4 n_i x_i^4 \right],$$

$$C = \frac{1}{Det_1} \left[\sum_{i=1}^4 n_i \sum_{i=1}^4 n_i x_i^3 - \sum_{i=1}^4 n_i x_i \sum_{i=1}^4 n_i x_i^2 \right],$$

$$D = \frac{1}{Det_1} \left[\sum_{i=1}^4 n_i \sum_{i=1}^4 n_i x_i^4 - \sum_{i=1}^4 n_i x_i \sum_{i=1}^4 n_i x_i^3 \right],$$

and

$$Det_1 = |\mathbf{X}^{(1)'} \mathbf{X}^{(1)}| = \left[n_1 n_2 (x_1 - x_2)^2 + n_1 n_3 (x_1 - x_3)^2 + n_1 n_4 (x_1 - x_4)^2 + \right. \\ \left. n_2 n_3 (x_2 - x_3)^2 + n_2 n_4 (x_2 - x_4)^2 + n_3 n_4 (x_3 - x_4)^2 \right].$$

Now the fitted values of \mathbf{y} under MI and MT are expressed as $\hat{\mathbf{y}}^{(1)} = \mathbf{X}^{(1)} \hat{\boldsymbol{\gamma}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)} = \mathbf{X}^{(t)} \hat{\boldsymbol{\gamma}}^{(t)}$. Hence we have

$$\begin{aligned} E(\hat{\mathbf{y}}^{(1)}) &= \mathbf{X}^{(1)} E(\hat{\boldsymbol{\gamma}}^{(1)}) \\ &= \mathbf{X}^{(1)} (\mathbf{X}^{(1)'} \mathbf{X}^{(1)})^{-1} \mathbf{X}^{(1)'} \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\ &= \mathbf{H}_1 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \end{aligned} \tag{6.8}$$

$$E(\hat{\mathbf{y}}^{(2)}) = \mathbf{X}^{(t)} E(\hat{\boldsymbol{\gamma}}^{(t)}) = \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \quad (6.9)$$

$$\Rightarrow E(\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)}) = (\mathbf{I} - \mathbf{H}_1) \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}, \quad (6.10)$$

where $\mathbf{H}_1 = \mathbf{X}^{(1)} (\mathbf{X}^{(1)'} \mathbf{X}^{(1)})^{-1} \mathbf{X}^{(1)'}$ and $\mathbf{H}_1' = \mathbf{H}_1$, $\mathbf{H}_1^2 = \mathbf{H}_1$. Now, using (6.10) we get

$$\begin{aligned} \mathbf{J} &= E(\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)})' E(\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)}) \\ &= \boldsymbol{\gamma}^{(t)'} \mathbf{X}^{(t)'} (\mathbf{I} - \mathbf{H}_1)^2 \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)} \\ &= \boldsymbol{\gamma}^{(t)'} \mathbf{X}^{(t)'} (\mathbf{I} - \mathbf{H}_1) \mathbf{X}^{(t)} \boldsymbol{\gamma}^{(t)}. \end{aligned} \quad (6.11)$$

Now,

$$\begin{aligned} \mathbf{X}^{(t)'} \mathbf{H}_1 \mathbf{X}^{(t)} &= \begin{pmatrix} \mathbf{X}^{(1)'} \\ \mathbf{X}_{23}' \end{pmatrix} \mathbf{H}_1 \begin{pmatrix} \mathbf{X}^{(1)} & \mathbf{X}_{23} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{X}^{(1)'} \mathbf{X}^{(1)} & \mathbf{X}^{(1)'} \mathbf{X}_{23} \\ \mathbf{X}_{23}' \mathbf{X}^{(1)} & \mathbf{X}_{23}' \mathbf{H}_1 \mathbf{X}_{23} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^4 n_i & \sum_{i=1}^4 n_i x_i & \sum_{i=1}^4 n_i x_i^2 & \sum_{i=1}^4 n_i x_i^3 \\ \sum_{i=1}^4 n_i x_i & \sum_{i=1}^4 n_i x_i^2 & \sum_{i=1}^4 n_i x_i^3 & \sum_{i=1}^4 n_i x_i^4 \\ \sum_{i=1}^4 n_i x_i^2 & \sum_{i=1}^4 n_i x_i^3 & P & R \\ \sum_{i=1}^4 n_i x_i^3 & \sum_{i=1}^4 n_i x_i^4 & R & Q \end{pmatrix}, \end{aligned} \quad (6.12)$$

where

$$P = \mathbf{X}_2' \mathbf{H}_1 \mathbf{X}_2 = A \sum_{i=1}^4 n_i x_i^2 + C \sum_{i=1}^4 n_i x_i^3,$$

$$Q = \mathbf{X}_3' \mathbf{H}_1 \mathbf{X}_3 = B \sum_{i=1}^4 n_i x_i^3 + D \sum_{i=1}^4 n_i x_i^4,$$

$$R = \mathbf{X}_2' \mathbf{H}_1 \mathbf{X}_3 = A \sum_{i=1}^4 n_i x_i^3 + C \sum_{i=1}^4 n_i x_i^4,$$

and $A, B, C,$ and D are defined right after (6.7). Therefore

$$\mathbf{X}^{(t)'} (\mathbf{I} - \mathbf{H}_1) \mathbf{X}^{(t)} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{23}' (\mathbf{I} - \mathbf{H}_1) \mathbf{X}_{23} \end{pmatrix},$$

and

$$\begin{aligned} \mathbf{X}_{23}' (\mathbf{I} - \mathbf{H}_1) \mathbf{X}_{23} &= \begin{pmatrix} \sum_{i=1}^4 n_i x_i^4 - P & \sum_{i=1}^4 n_i x_i^5 - R \\ \sum_{i=1}^4 n_i x_i^5 - R & \sum_{i=1}^4 n_i x_i^6 - Q \end{pmatrix} \\ &= \frac{1}{\text{Det}_1} \begin{pmatrix} P^* & R^* \\ R^* & Q^* \end{pmatrix} = \mathbf{S}_J \quad (\text{say}). \end{aligned} \tag{6.13}$$

where $P, Q,$ and R are defined right after (6.12) and

$$\begin{aligned} P^* &= n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 \\ &\quad + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 \\ &\quad + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 \\ &\quad + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2, \end{aligned}$$

$$\begin{aligned} Q^* &= n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3)^2 \\ &\quad + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4)^2 \\ &\quad + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4)^2 \\ &\quad + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4)^2, \end{aligned}$$

$$\begin{aligned}
R^* &= n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) \\
&\quad + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) \\
&\quad + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) \\
&\quad + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4),
\end{aligned}$$

and

$$\begin{aligned}
Det_1 &= |\mathbf{X}^{(1)'} \mathbf{X}^{(1)}| = n_1 n_2 (x_1 - x_2)^2 + n_1 n_3 (x_1 - x_3)^2 + n_1 n_4 (x_1 - x_4)^2 \\
&\quad + n_2 n_3 (x_2 - x_3)^2 + n_2 n_4 (x_2 - x_4)^2 + n_3 n_4 (x_3 - x_4)^2.
\end{aligned}$$

Using (6.11), (6.12), and (6.13) we get

$$\begin{aligned}
\mathbf{J} &= \gamma^{(t)'} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_J \end{pmatrix} \gamma^{(t)} \\
&= \frac{1}{Det_1} [P^* \gamma_2^2 + Q^* \gamma_3^2 + 2\gamma_2 \gamma_3 R^*], \tag{6.14}
\end{aligned}$$

We now consider the fitted value criterion \mathbf{J} in (6.14) under the class of designs \mathbf{D}_1 in (2.19). We note that under \mathbf{D}_1 ,

$$R^* = 0, \quad \frac{P^*}{Det_1} = \frac{2(b^2 - a^2)^2 n_1 n_2}{(n_1 + n_2)}, \quad \text{and} \quad \frac{Q^*}{Det_1} = \frac{2a^2 b^2 (b^2 - a^2)^2 n_1 n_2}{(n_1 b^2 + n_2 a^2)}.$$

Thus under \mathbf{D}_1 , \mathbf{J} takes the form

$$\mathbf{J} = \gamma_2^2 \left[\frac{2(b^2 - a^2)^2 n_1 n_2}{(n_1 + n_2)} \right] + \gamma_3^2 \left[\frac{2a^2 b^2 (b^2 - a^2)^2 n_1 n_2}{(n_1 b^2 + n_2 a^2)} \right]. \tag{6.15}$$

The fitted value criterion involves two unknown model parameters γ_2 and γ_3 . As a consequence it is not possible to maximize it directly with respect to the design parameters only. But we can maximize \mathbf{J} using \mathbf{K} criterion as discussed in the next section.

6.3.1 An Interesting Observation

We define $\boldsymbol{\gamma}^{(*)} = \begin{pmatrix} \gamma_2 & \gamma_3 \end{pmatrix}'$ and note that (Rao (1973))

$$\begin{aligned}
\text{Var}(\hat{\boldsymbol{\gamma}}^{(t)}) &= \sigma^2 [\mathbf{X}^{(t)'}\mathbf{X}^{(t)}]^{-1} \\
\Leftrightarrow \begin{bmatrix} \text{Var}(\hat{\boldsymbol{\gamma}}^{(1)}) & \text{Cov}(\hat{\boldsymbol{\gamma}}^{(1)}, \hat{\boldsymbol{\gamma}}^{(*)}) \\ \text{Cov}(\hat{\boldsymbol{\gamma}}^{(*)}, \hat{\boldsymbol{\gamma}}^{(1)}) & \text{Var}(\hat{\boldsymbol{\gamma}}^{(*)}) \end{bmatrix} &= \sigma^2 \begin{bmatrix} \mathbf{X}^{(1)'}\mathbf{X}^{(1)} & \mathbf{X}^{(1)'}\mathbf{X}_{23} \\ \mathbf{X}'_{23}\mathbf{X}^{(1)} & \mathbf{X}'_{23}\mathbf{X}_{23} \end{bmatrix}^{-1} \\
\Rightarrow \text{Var}(\hat{\boldsymbol{\gamma}}^{(*)}) &= \sigma^2 [\mathbf{X}_{23}'(\mathbf{I} - \mathbf{H}_1)\mathbf{X}_{23}]^{-1} \\
\Leftrightarrow \text{Var}(\hat{\boldsymbol{\gamma}}^{(*)}) &= \sigma^2 \mathbf{S}_J^{-1}, \tag{6.16}
\end{aligned}$$

where \mathbf{S}_J is defined in (6.13). We also note that

$$|\mathbf{S}_J| = |\mathbf{X}_{23}'(\mathbf{I} - \mathbf{H}_1)\mathbf{X}_{23}| = \frac{|\mathbf{X}^{(t)'}\mathbf{X}^{(t)}|}{|\mathbf{X}^{(1)'}\mathbf{X}^{(1)}|}$$

where

$$|\mathbf{X}^{(t)'}\mathbf{X}^{(t)}| = n_1 n_2 n_3 n_4 [(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4)]^2.$$

It is to be noted that maximizing $|\mathbf{S}_J|$ implies minimizing the volume of the confidence ellipsoid of the parameter estimates $(\hat{\boldsymbol{\gamma}}^{(*)})$.

6.4 Expression of K_J Criterion

Under the class of designs \mathbf{D} in (1.3) we have,

$$\begin{aligned}
 K_J &= |\mathbf{S}_J| \\
 &= \frac{n_1 n_2 n_3 n_4 [(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4)]^2}{\left[n_1 n_2 (x_1 - x_2)^2 + n_1 n_3 (x_1 - x_3)^2 + n_1 n_4 (x_1 - x_4)^2 \right. \\
 &\quad \left. + n_2 n_3 (x_2 - x_3)^2 + n_2 n_4 (x_2 - x_4)^2 + n_3 n_4 (x_3 - x_4)^2 \right]}
 \end{aligned} \tag{6.17}$$

Considering the subclass of designs \mathbf{D}_1 in (2.19) the criterion reduces to,

$$\begin{aligned}
 K_J &= \frac{8a^2 b^2 (b^2 - a^2)^4 n_1^2 n_2^2}{n (n_1 b^2 + n_2 a^2)} \quad \text{since} \quad (n_1 + n_2) = \frac{n}{2} \\
 \Leftrightarrow \frac{K_J}{n^2} &= \frac{8a^2 b^2 (b^2 - a^2)^4 p_1^2 p_2^2}{(p_1 b^2 + p_2 a^2)} \quad \text{since} \quad (p_1 + p_2) = \frac{1}{2}.
 \end{aligned} \tag{6.18}$$

We note that for a given value of b , K_J attains its maximum when $p_1 = \frac{1}{5}$ and $a = \frac{b}{\sqrt{6}}$ (Appendix-B.1). We define the design below in \mathbf{D}_1 ,

$$\begin{aligned}
 d_7 &= \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -b, x_2 = -\frac{b}{\sqrt{6}}, x_3 = \frac{b}{\sqrt{6}}, x_4 = b; \right. \\
 &\quad \left. n_1 = n_4 = \frac{n}{5}, n_2 = n_3 = \frac{3n}{10}, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right],
 \end{aligned} \tag{6.19}$$

and get the following theorem.

Theorem 21. *For a given value of b in the class of designs \mathbf{D}_1 , the design d_7 is optimum with respect to the criterion K_J .*

Now from Appendix B.2 we note that criterion \mathbf{K}_J attains its maximum possible value when $b = 1$. So, we define the following design in \mathbf{D}_1

$$d_8 = \left[(x_1, x_2, x_3, x_4; n_1, n_2, n_3, n_4) : x_1 = -1, x_2 = -\frac{1}{\sqrt{6}}, x_3 = \frac{1}{\sqrt{6}}, x_4 = 1; \right. \\ \left. n_1 = n_4 = \frac{n}{5}, n_2 = n_3 = \frac{3n}{10}, \text{Rank}(\mathbf{X}^{(t)}) = 4 \right], \quad (6.20)$$

and get the following theorem.

Theorem 22. *For the class of designs \mathbf{D}_1 , the design d_8 is optimum with respect to the criterion \mathbf{K}_J .*

6.5 Expression of I Criterion

Under the class of designs \mathbf{D} in (1.3) we have

$$\begin{aligned} E(\hat{y}^{(1)}(x)) &= \begin{pmatrix} 1 & x \end{pmatrix} E(\hat{\gamma}^{(1)}) \\ &= \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} 1 & 0 & A & B \\ 0 & 1 & C & D \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (6.7)} \\ &= \gamma_0 + \gamma_1 x + \gamma_2(A + Cx) + \gamma_3(B + Dx), \end{aligned} \quad (6.21)$$

and

$$\begin{aligned} E(\hat{y}^{(2)}(x)) &= \begin{pmatrix} 1 & x & x^2 & x^3 \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (6.6)} \\ &= \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \gamma_3 x^3, \end{aligned} \quad (6.22)$$

where A, B, C , and D are defined right after (6.6). From (6.21) and (6.22) we get

$$E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x)) = \gamma_2 [A + Cx - x^2] + \gamma_3 [B + Dx - x^3] \quad (6.23)$$

Now, under the subclass of designs \mathbf{D}_1 in (2.19) we have

$$\begin{aligned} Det_1 &= 2n(b^2n_1 + a^2n_2), & B &= C = 0, \\ A &= \frac{2}{n}(b^2n_1 + a^2n_2), & D &= \frac{(b^4n_1 + a^4n_2)}{(b^2n_1 + a^2n_2)}. \end{aligned}$$

Thus from (6.23) we have,

$$[E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 = [\gamma_2 (A - x^2) + \gamma_3 (Dx - x^3)]^2. \quad (6.24)$$

Hence the predicted value criterion under \mathbf{D}_1 is given by

$$\begin{aligned} \mathbf{I} &= \frac{1}{2} \int_{-1}^1 [\gamma_2 (A - x^2) + \gamma_3 (Dx - x^3)]^2 dx \\ &= \gamma_2^2 \left[A^2 - \frac{2A}{3} + \frac{1}{5} \right] + \gamma_3^2 \left[\frac{D^2}{3} - \frac{2D}{5} + \frac{1}{7} \right] \end{aligned} \quad (6.25)$$

where $0 < A = \frac{(b^2n_1 + a^2n_2)}{(n_1 + n_2)} < 1$ and $0 < D = \frac{(b^4n_1 + a^4n_2)}{(b^2n_1 + a^2n_2)} < 1$. The expression of criterion \mathbf{I} is still dependent on two model parameters γ_2 and γ_3 . We optimize \mathbf{I} using \mathbf{K} criterion which is discussed in the next section.

6.6 Expression of K_I Criterion

From (6.25) we obtain the K_I criterion for the predicted values as

$$K_I = \phi(A)\phi(D), \quad (6.26)$$

where $\phi(A) = \left(A^2 - \frac{2A}{3} + \frac{1}{5}\right)$, $\phi(D) = \left(\frac{D^2}{3} - \frac{2D}{5} + \frac{1}{7}\right)$, $0 < A = \frac{(b^2n_1+a^2n_2)}{(n_1+n_2)} < 1$, and $0 < D = \frac{(b^4n_1+a^4n_2)}{(b^2n_1+a^2n_2)} < 1$. Now, to maximize K_I , we have to simultaneously maximize $\phi(A)$ and $\phi(D)$. From *Fig-6.1*) we note that $\phi(A)$ is maximum when

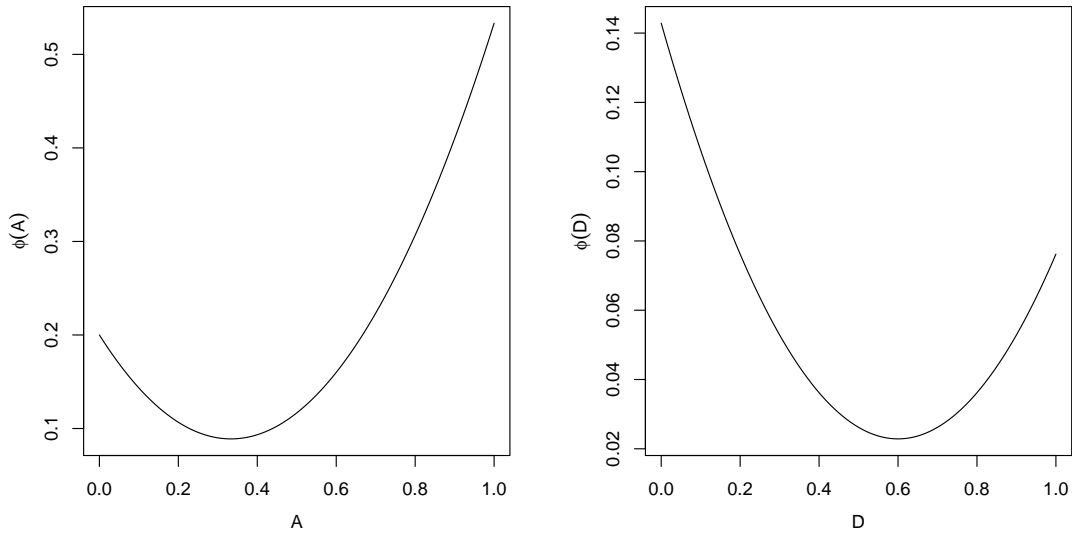


Figure 6.1: Plots of $\phi(A)$ against A and $\phi(D)$ against D

$A \rightarrow 1$ and $\phi(D)$ is maximum when $D \rightarrow 0$. Therefore when $A \rightarrow 1$ and $D \rightarrow 0$ we should have the maximum value of K_I . The other possibilities of finding a maximum value of K_I is given in Table-6.1. But some choices are not realized here because A and D change in the same direction: if A increases so does D and if A decreases so does D . Therefore the maximum possible value of K_I is attained when

Table 6.1: Finding Maximum of Criterion \mathbf{K}_I

Value of A, D	\mathbf{K}_I	Realizable
$A \rightarrow 1, D \rightarrow 1$	0.04063492	Yes
$A \rightarrow 0, D \rightarrow 0$	0.02857143	Yes
$A \rightarrow 1, D \rightarrow 0$	0.07619048	No
$A \rightarrow 0, D \rightarrow 1$	0.01523810	No

$A \rightarrow 1$ and $D \rightarrow 1$. We assume $b = 1$ and obtain some numerical choices for a and p_1 to maximize A and D to optimize \mathbf{K}_I . Some optimal designs in \mathbf{D}_1 with respect to \mathbf{K}_I criterion are proposed in Table-6.2. We note that designs those perform well with respect to \mathbf{K}_I do not perform well with respect to \mathbf{K}_J and vice versa. We consider the design d_8 of (6.20) here and we note that it performs the best in the lot with respect to \mathbf{K}_J but not well with respect to \mathbf{K}_I . We also consider the design with points $x_1 = -1, x_2 = -\frac{1}{2}, x_3 = \frac{1}{2}$, and $x_4 = 1$ with equal allocation and note that it performs better with respect to \mathbf{K}_J than \mathbf{K}_I .

Table 6.2: Some K_I -Optimal Designs in D_1

b	a	p_1	p_2	A	D	K_J	K_I
1	0.20	0.48	0.02	0.961600	0.9984027	0.000052	0.03664073
1	0.24	0.48	0.02	0.962304	0.9977437	0.000070	0.03662306
1	0.26	0.48	0.02	0.962704	0.9973811	0.000078	0.03661450
1	0.30	0.48	0.02	0.963600	0.9966002	0.000094	0.03659924
1	0.80	0.47	0.03	0.978400	0.9858708	0.000035	0.03660706
1	0.70	0.48	0.02	0.979600	0.9897958	0.000050	0.03723345
1	0.72	0.48	0.02	0.980736	0.9898174	0.000042	0.03734432
1	0.74	0.48	0.02	0.981904	0.9899080	0.000034	0.03746759
1	0.75	0.48	0.02	0.982500	0.9899809	0.000031	0.03753414
1	0.76	0.47	0.03	0.974656	0.9849807	0.000060	0.03614353
1	0.78	0.47	0.03	0.976504	0.9853611	0.000047	0.03636415
1	0.80	0.48	0.02	0.985600	0.9906494	0.000016	0.03792034
1	$\frac{1}{\sqrt{6}}$	$\frac{1}{5}$	$\frac{3}{10}$	0.500000	0.8333333	0.009259	0.00478395
1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0.625000	0.8500000	0.007910	0.00760032

Chapter 7

Quadratic vs. Special Cubic when the True Model Is Full Cubic

7.1 Introduction

Here we consider that the response variable Y is dependent on the explanatory variable X with two possible dependence as described by two models MI, a quadratic linear regression model and MII, a cubic regression model without the quadratic coefficient. We assume the full cubic regression model MT to be the true model. So, our goal is to discriminate between these two models MI and MII at the design stage assuming MT to be the unknown true model.

7.2 Models and Associated Designs

We consider the general class of designs \mathbf{D} in (1.3). We also consider the full cubic model from (1.1) as the true model and denote it by MT. Our aim is to discriminate between the two models MI and MII assuming MT to be the unknown true model.

The three models considered here are given by

$$\begin{aligned}
\mathbf{MT:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \gamma_3 x_i^3, \\
\text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
\end{aligned} \tag{7.1}$$

$$\begin{aligned}
\mathbf{MI:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2, \\
\text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
\end{aligned} \tag{7.2}$$

$$\begin{aligned}
\mathbf{MII:} \quad E(y_j(x_i)) &= \gamma_0 + \gamma_1 x_i + \gamma_3 x_i^3, \\
\text{Var}((y_j(x_i))) &= \sigma^2, \quad \text{Cov}((y_j(x_i)), (y_{j'}(x_{i'}))) = 0,
\end{aligned} \tag{7.3}$$

where $i, i' = 1, 2, 3, 4$; $j = 1, 2, \dots, n_i$, $j' = 1, 2, \dots, n_{i'}$, $(i, j) \neq (i', j')$. The matrix representations of MT, MI, and MII are given by

$$E(\mathbf{y}) = \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{7.4}$$

$$E(\mathbf{y}) = \mathbf{X}^{(1)}\boldsymbol{\gamma}^{(1)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{7.5}$$

$$E(\mathbf{y}) = \mathbf{X}^{(2)}\boldsymbol{\gamma}^{(2)}, \quad \text{Var}(\mathbf{y}) = \sigma^2\mathbf{I}, \tag{7.6}$$

where

$$\mathbf{X}^{(1)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}', \quad \boldsymbol{\gamma}^{(1)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix},$$

$$\mathbf{X}^{(2)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \boldsymbol{\gamma}^{(2)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_3 \end{pmatrix},$$

$$\mathbf{X}^{(t)} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_3 & x_4 & \dots & x_4 \\ x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \\ x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}', \boldsymbol{\gamma}^{(t)} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}.$$

We define

$$\mathbf{X}_2 = \begin{pmatrix} x_1^2 & \dots & x_1^2 & x_2^2 & \dots & x_2^2 & x_3^2 & \dots & x_3^2 & x_4^2 & \dots & x_4^2 \end{pmatrix}',$$

$$\mathbf{X}_3 = \begin{pmatrix} x_1^3 & \dots & x_1^3 & x_2^3 & \dots & x_2^3 & x_3^3 & \dots & x_3^3 & x_4^3 & \dots & x_4^3 \end{pmatrix}'.$$

7.3 Expression of J Criterion

We first consider the fitted value criterion J (ref. Chapter-1) and obtain the efficient design within the class of designs \mathbf{D} in (1.3) for model selection and discrimination purposes. Let $\hat{\mathbf{y}}^{(1)}$ and $\hat{\mathbf{y}}^{(2)}$ be the fitted values of the two models MI and MII. First we find the expression of $E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})$.

The least square estimate of $\boldsymbol{\gamma}^{(1)}$ for MI is given by (Rao (1973))

$$\hat{\boldsymbol{\gamma}}^{(1)} = \begin{pmatrix} \hat{\gamma}_0^{(1)} \\ \hat{\gamma}_1^{(1)} \\ \hat{\gamma}_2^{(1)} \end{pmatrix} = (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{y}, \quad (7.7)$$

and the least square estimate of $\boldsymbol{\gamma}^{(2)}$ for MII is given by

$$\hat{\boldsymbol{\gamma}}^{(2)} = \begin{pmatrix} \hat{\gamma}_0^{(2)} \\ \hat{\gamma}_1^{(2)} \\ \hat{\gamma}_3^{(2)} \end{pmatrix} = (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{y}. \quad (7.8)$$

It can be checked that

$$\begin{aligned} E(\hat{\boldsymbol{\gamma}}^{(1)}) &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}E(\mathbf{y}) \\ &= (\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\ &= \begin{pmatrix} 1 & 0 & 0 & A \\ 0 & 1 & 0 & B \\ 0 & 0 & 1 & C \end{pmatrix} \boldsymbol{\gamma}^{(t)}, \end{aligned} \quad (7.9)$$

and

$$\begin{aligned} E(\hat{\boldsymbol{\gamma}}^{(2)}) &= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}E(\mathbf{y}) \\ &= (\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\ &= \begin{pmatrix} 1 & 0 & E & 0 \\ 0 & 1 & F & 0 \\ 0 & 0 & G & 1 \end{pmatrix} \boldsymbol{\gamma}^{(t)}, \end{aligned} \quad (7.10)$$

where

$$\begin{aligned}
A = \frac{1}{\text{Det}_1} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 x_1 x_2 x_3 \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 x_1 x_2 x_4 \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 x_1 x_3 x_4 \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 x_2 x_3 x_4 \right],
\end{aligned}$$

$$\begin{aligned}
B = \frac{-1}{\text{Det}_1} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 x_2 + x_1 x_3 + x_2 x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 x_2 + x_1 x_4 + x_2 x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 x_3 + x_1 x_4 + x_3 x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 x_3 + x_2 x_4 + x_3 x_4) \right],
\end{aligned}$$

$$\begin{aligned}
C = \frac{1}{\text{Det}_1} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) \right],
\end{aligned}$$

$$\begin{aligned}
\text{Det}_1 & = |\mathbf{X}^{(1)'} \mathbf{X}^{(1)}| = \\
& \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 \right. \\
& \left. + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 \right],
\end{aligned}$$

$$\begin{aligned}
E = \frac{-1}{\text{Det}_2} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) x_1 x_2 x_3 \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) x_1 x_2 x_4 \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) x_1 x_3 x_4 \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) x_2 x_3 x_4 \right],
\end{aligned}$$

$$F =$$

$$\begin{aligned}
\frac{1}{\text{Det}_2} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) (x_1 x_2 + x_1 x_3 + x_2 x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) (x_1 x_2 + x_1 x_4 + x_2 x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) (x_1 x_3 + x_1 x_4 + x_3 x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) (x_2 x_3 + x_2 x_4 + x_3 x_4) \right],
\end{aligned}$$

$$\begin{aligned}
G = \frac{1}{\text{Det}_2} & \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3) \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4) \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4) \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4) \right],
\end{aligned}$$

$$\begin{aligned}
\text{Det}_2 = |\mathbf{X}^{(2)'} \mathbf{X}^{(2)}| & = \left[n_1 n_2 n_3 (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 + x_2 + x_3)^2 \right. \\
& + n_1 n_2 n_4 (x_1 - x_2)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_1 + x_2 + x_4)^2 \\
& + n_1 n_3 n_4 (x_1 - x_3)^2 (x_1 - x_4)^2 (x_3 - x_4)^2 (x_1 + x_3 + x_4)^2 \\
& \left. + n_2 n_3 n_4 (x_2 - x_3)^2 (x_2 - x_4)^2 (x_3 - x_4)^2 (x_2 + x_3 + x_4)^2 \right].
\end{aligned}$$

The expected fitted values from the two models are given by

$$\begin{aligned}
E(\hat{\mathbf{y}}^{(1)}) &= \mathbf{X}^{(1)}E(\hat{\boldsymbol{\gamma}}^{(1)}) \\
&= \mathbf{X}^{(1)}(\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\
&= \mathbf{H}_1\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)},
\end{aligned} \tag{7.11}$$

$$\begin{aligned}
E(\hat{\mathbf{y}}^{(2)}) &= \mathbf{X}^{(2)}E(\hat{\boldsymbol{\gamma}}^{(2)}) \\
&= \mathbf{X}^{(2)}(\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\
&= \mathbf{H}_2\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)},
\end{aligned} \tag{7.12}$$

$$\Rightarrow E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)}) = (\mathbf{H}_1 - \mathbf{H}_2)\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}, \tag{7.13}$$

where $\mathbf{H}_1 = \mathbf{X}^{(1)}(\mathbf{X}^{(1)'}\mathbf{X}^{(1)})^{-1}\mathbf{X}^{(1)'}$ and $\mathbf{H}_2 = \mathbf{X}^{(2)}(\mathbf{X}^{(2)'}\mathbf{X}^{(2)})^{-1}\mathbf{X}^{(2)'}$. We note that \mathbf{H}_1 and \mathbf{H}_2 are symmetric and idempotent (Rao (1973)) i.e. $\mathbf{H}_1' = \mathbf{H}_1$, $\mathbf{H}_1^2 = \mathbf{H}_1$, $\mathbf{H}_2' = \mathbf{H}_2$ and $\mathbf{H}_2^2 = \mathbf{H}_2$. It is to be noted that here $\mathbf{H}_2\mathbf{H}_1 \neq \mathbf{H}_1$. Clearly using (7.13) we get

$$\begin{aligned}
\mathbf{J} &= E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)})'E(\hat{\mathbf{y}}^{(1)} - \hat{\mathbf{y}}^{(2)}) \\
&= \boldsymbol{\gamma}^{(t)'}\mathbf{X}^{(t)'}(\mathbf{H}_1 - \mathbf{H}_2)^2\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)} \\
&= \boldsymbol{\gamma}^{(t)'}\mathbf{X}^{(t)'}(\mathbf{H}_2 + \mathbf{H}_1 - \mathbf{H}_1\mathbf{H}_2 - \mathbf{H}_2\mathbf{H}_1)\mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}.
\end{aligned} \tag{7.14}$$

Now we consider the subclass of designs \mathbf{D}_1 in (2.19) which is a subclass of \mathbf{D}

in (1.3). We note that under \mathbf{D}_1 we have

$$\begin{aligned} Det_1 &= 8n_1n_2(b^2 - a^2)^2(b^2n_1 + a^2n_2), & B &= \frac{(b^4n_1 + a^4n_2)}{(b^2n_1 + a^2n_2)} & A &= 0, & C &= 0 \\ Det_2 &= 4nn_1n_2a^2b^2(b^2 - a^2)^2, & E &= \frac{2}{n}(b^2n_1 + a^2n_2), & F &= 0, & G &= 0, \end{aligned}$$

and the fitted value criterion is given by

$$\begin{aligned} \mathbf{J} &= \gamma^{(t)'} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2n_1n_2(b^2-a^2)^2}{(n_1+n_2)} & 0 \\ 0 & 0 & 0 & \frac{2n_1n_2a^2b^2(b^2-a^2)^2}{(b^2n_1+a^2n_2)} \end{pmatrix} \gamma^{(t)} \\ &= \gamma_2^2 \left[\frac{2n_1n_2(b^2 - a^2)^2}{(n_1 + n_2)} \right] + \gamma_3^2 \left[\frac{2n_1n_2a^2b^2(b^2 - a^2)^2}{(b^2n_1 + a^2n_2)} \right]. \end{aligned} \quad (7.15)$$

It is interesting to observe that under the class of designs \mathbf{D}_1 , the expressions of \mathbf{J} in (6.15) of Chapter-6 and in (7.15) of Chapter-7 are identical. The fitted value criterion \mathbf{J} involves two unknown model parameters γ_2 and γ_3 and as a consequence it is not possible to maximize it directly with respect to the design parameters only.

We define

$$\mathbf{S}_J = \begin{pmatrix} \frac{2n_1n_2(b^2-a^2)^2}{(n_1+n_2)} & 0 \\ 0 & \frac{2n_1n_2a^2b^2(b^2-a^2)^2}{(b^2n_1+a^2n_2)} \end{pmatrix},$$

and obtain the \mathbf{K}_J criterion as $|\mathbf{S}_J|$ in the next section.

7.4 Expression of K_J Criterion

The criterion K_J is obtained under the class of designs D_1 in (2.19) as

$$\begin{aligned}
 K_J &= |S_J| \\
 &= \frac{8a^2b^2(b^2 - a^2)^4 n_1^2 n_2^2}{n(n_1 b^2 + n_2 a^2)} \quad \text{since } (n_1 + n_2) = \frac{n}{2} \\
 \Leftrightarrow \frac{K_J}{n^2} &= \frac{8a^2b^2(b^2 - a^2)^4 p_1^2 p_2^2}{(p_1 b^2 + p_2 a^2)} \quad \text{since } (p_1 + p_2) = \frac{1}{2}. \quad (7.16)
 \end{aligned}$$

Clearly, the criterion K_J observed in (7.16) is identical to the criterion K_J in (6.18) of Chapter-6. Hence from Chapter-6 the theorems follow.

Theorem 23. *For a given value of b in the class of designs D_1 , the design d_5 in (6.19) is optimum with respect to the criterion K_J .*

Theorem 24. *For the class of designs D_1 , the design d_6 in (6.20) is optimum with respect to the criterion K_J .*

7.5 Expression of I Criterion

Here we consider the predicted value criterion under the class of designs D in (1.3).

Let $\hat{y}^{(1)}(x)$ and $\hat{y}^{(2)}(x)$ be the predicted values of the two models MI and MII at $X = x$. We first find the expression for $E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))$ assuming MT to be the true model.

We know that

$$\begin{aligned}
E(\hat{y}^{(1)}(x)) &= \begin{pmatrix} 1 & x & x^2 \end{pmatrix} E(\hat{\gamma}^{(1)}) \\
&= \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & A \\ 0 & 1 & 0 & B \\ 0 & 0 & 1 & C \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (7.9)} \\
&= \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \gamma_3(A + Bx + Cx^2), \tag{7.17}
\end{aligned}$$

and

$$\begin{aligned}
E(\hat{y}^{(2)}(x)) &= \begin{pmatrix} 1 & x & x^3 \end{pmatrix} E(\hat{\gamma}^{(2)}) \\
&= \begin{pmatrix} 1 & x & x^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & E & 0 \\ 0 & 1 & F & 0 \\ 0 & 0 & G & 1 \end{pmatrix} \boldsymbol{\gamma}^{(t)} \quad \text{from (7.10)} \\
&= \gamma_0 + \gamma_1 x + \gamma_2(E + Fx + Gx^3) + \gamma_3 x^3, \tag{7.18}
\end{aligned}$$

where A, B, C, D, E, F , and G are defined right after (7.10).

Thus from (7.17) and (7.18) we get

$$E(\hat{y}^{(1)}(x)) - \hat{y}^{(2)}(x) = \gamma_2 [x^2 - E - Fx - Gx^3] + \gamma_3 [A + Bx + Cx^2 - x^3] \tag{7.19}$$

Now considering the class of designs \mathbf{D}_1 in (2.19), we have $A = C = F = G = 0$, $B = \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)}$, and $E = \frac{2}{n}(b^2 n_1 + a^2 n_2)$. Therefore under this class of designs we

get from (7.19),

$$\begin{aligned} & [E(\hat{y}^{(1)}(x) - \hat{y}^{(2)}(x))]^2 \\ &= [\gamma_2 (x^2 - E) + \gamma_3 (Bx - x^3)]^2. \end{aligned} \quad (7.20)$$

Hence the predicted value criterion under \mathbf{D}_1 is given by

$$\begin{aligned} \mathbf{I} &= \frac{1}{2} \int_{-1}^1 [\gamma_2 (x^2 - E) + \gamma_3 (Bx - x^3)]^2 dx \\ &= \gamma_2^2 \left[E^2 - \frac{2E}{3} + \frac{1}{5} \right] + \gamma_3^2 \left[\frac{B^2}{3} - \frac{2B}{5} + \frac{1}{7} \right] \end{aligned} \quad (7.21)$$

where $0 < E = \frac{(b^2 n_1 + a^2 n_2)}{(n_1 + n_2)} < 1$ and $0 < B = \frac{(b^4 n_1 + a^4 n_2)}{(b^2 n_1 + a^2 n_2)} < 1$. The expression of criterion \mathbf{I} depends on two unknown model parameters γ_2 and γ_3 . Hence it is not possible to directly optimize \mathbf{I} with respect to design parameters only. We use criterion \mathbf{K} to optimize \mathbf{I} in this case. But it is interesting to note that the expression of \mathbf{I} is identical to the expression of \mathbf{I} in (6.25) of Chapter-6. So all the \mathbf{K}_I optimal designs obtained in Chapter-6 apply here.

Chapter 8

Conclusion

In this dissertation we consider the problem of model identification and discrimination for the class of models describing the dependence of the response variable Y on an explanatory variable X by at most a third order polynomial regression model. Hence the class consists of models up to a maximum of third order with linear, quadratic, and cubic terms present. We include an intercept parameter for all the models. A general class of designs with replicated four distinct points is considered. While discriminating between the two models within the class, the unknown true model may or may not be one of them. Ghosh and Pal (2008) proposed two optimality criterion functions \mathbf{J} and \mathbf{I} for the model identification and discrimination. The fitted value criterion \mathbf{J} and a modified predicted value criterion \mathbf{I} by replacing the absolute difference with the squared difference, are chosen as the optimality criterion functions. When the criterion functions \mathbf{J} and \mathbf{I} are dependent on more than one model parameter, we define a new criterion \mathbf{K} to optimize \mathbf{J} and \mathbf{I} and denote them by \mathbf{K}_J and \mathbf{K}_I .

We observe that the \mathbf{T} -optimality criterion (Atkinson and Fedorov (1975a))

is not meaningful when the unknown true model is not one of the two models that we consider for model identification and discrimination (Chapter-2, 5, and 7). When one of the two models considered for identification and discrimination is the unknown true model, the \mathbf{J} -optimality criterion is identical to the \mathbf{T} -optimality criterion as shown in Chapter-3.

For discrimination between a linear and a special cubic regression models with the latter being the unknown true model (Chapter-3), we obtain a class of designs that are better than the Dette-Titoff \mathbf{T} -optimal designs (Dette and Titoff (2009)) under the criterion \mathbf{I} . However the Dette-Titoff design is naturally better than our class of designs under the criterion \mathbf{T} (or equivalently \mathbf{J}). We also obtain a class of \mathbf{J} -optimal designs indicating the non-uniqueness of the \mathbf{J} -optimal design.

For discrimination between a quadratic and a full cubic regression models with the latter being the unknown true model, we obtain a class of designs that are better than the Kiefer-Wolfowitz optimum design (Kiefer and Wolfowitz (1959)) under the criterion \mathbf{I} . However the Kiefer-Wolfowitz optimum design performs better than our class under the criterion \mathbf{J} .

For discrimination between a quadratic and a cubic regression models with no quadratic term ($\gamma_2 = 0$) with the unknown true model being the full cubic model, we introduce the \mathbf{K}_J and \mathbf{K}_I criterion functions to obtain optimal designs for the model identification and discrimination (Chapter-7). For discrimination between a linear and a full cubic regression models with the latter being the unknown true model, the optimal designs are also obtained with the proposed \mathbf{K}_J and \mathbf{K}_I criterion functions (Chapter-6).

Bibliography

Andrews, D. F. (1971). Sequentially designed experiments for screening out bad models with f tests. *Biometrika*, **58**:427–432.

Atkinson, A. C. (1972). Planning experiments to detect inadequate regression models. *Biometrika*, **59**:275–293.

Atkinson, A. C. (2008). DT-optimum designs for model discrimination and parameter estimation. *Journal of Statistical Planning and Inference*, **138**:56–64.

Atkinson, A. C. and Cox, D. (1974). Planning experiments for discriminating between models (with discussion). *Journal of the Royal Stat. Soc.*, **B**:321–348.

Atkinson, A. C. and Fedorov, V. V. (1975a). The design of experiments for discriminating between two rival models. *Biometrika*, **62**:57–70.

Atkinson, A. C. and Fedorov, V. V. (1975b). Optimal design: Experiments for discriminating between several models. *Biometrika*, **62**:289–303.

Biswas, A. and Chaudhuri, P. (2002). An efficient design for model discrimination and parameter estimation in linear models. *Biometrika*, **89**:709–718.

Dette, H. (1991). A note on robust designs for polynomial regression. *Journal of Statistical Planning and Inference*, **28**:223–232.

- Dette, H. (1994). Discrimination designs for polynomial regression on a compact interval. *Annals of Statistics*, **22**:890–904.
- Dette, H. (1995). Optimal designs for identifying the degree of a polynomial regression. *Annals of Statistics*, **23**:1248–1267.
- Dette, H. and Kwiecien, R. (2004). A comparison of sequential and non-sequential designs for discrimination between nested regression models. *Biometrika*, **91**:165–176.
- Dette, H. and Titoff, S. (2009). Optimal discrimination designs. *The Annals of Statistics*, 37(4):2056–2082.
- Ghosh, S. and Pal, R. (2008). A note on model identification and discrimination for simple linear regression. *Sankhya*, **70-B**:0–11.
- Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. *Ann. Math. Statistics*, **30**:271–294.
- Lauter, E. (1974). Experimental design in a class of models. *Math. Oper. Statist.*, **5**:379–398.
- Montepiedra, G. and Yeh, A. B. (1998). A two-stage strategy for the construction of d-optimal experimental designs. *Communication in Statistics: Simulation and Computation*, **27**:377–401.
- Pukelsheim, F. and Rosenberger, J. L. (1993). Experimental designs for model discrimination. *Journal of the American Statistical Association*, **88**:642–649.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Second edition. John Wiley & Sons Inc., New York.

Spruill, M. (1990). A note on model identification and discrimination for simple linear regression. *Sankhya*, **B-52**:67–74.

Appendix A

A.1 Maximization of J for a given b

$$\frac{\mathbf{J}_b}{n\gamma_3^2} = \frac{2p_1p_2a^2b^2(b^2 - a^2)^2}{(b^2p_1 + a^2p_2)} \quad \text{where} \quad (p_1 + p_2) = \frac{1}{2}. \quad (\text{A.1})$$

Taking the common logarithm of both sides and assuming $F = \log\left(\frac{\mathbf{J}_b}{n\gamma_3^2}\right)$ we get

$$F = \log 2 + \log p_1 + \log\left(\frac{1}{2} - p_1\right) + 2\log a + 2\log b + 2\log(b^2 - a^2) - \log\left[(b^2 - a^2)p_1 + \frac{a^2}{2}\right]. \quad (\text{A.2})$$

Now, differentiating (A.2) with respect to a and p_1 we get F_a and F_{p_1} respectively where

$$F_a = \frac{2}{a} - \frac{4a}{(b^2 - a^2)} - \frac{2a(1 - 2p_1)}{[2(b^2 - a^2)p_1 + a^2]}, \quad (\text{A.3})$$

$$F_{p_1} = \frac{1}{p_1} - \frac{2}{1 - 2p_1} - \frac{2(b^2 - a^2)}{[2(b^2 - a^2)p_1 + a^2]}. \quad (\text{A.4})$$

Equating each of (A.3) and (A.4) to zero, we get

$$F_a = 0 \Rightarrow \frac{2(b^2 - 3a^2)}{a^2} = \frac{2(b^2 - a^2)(1 - 2p_1)}{[2(b^2 - a^2)p_1 + a^2]} \quad (\text{A.5})$$

$$\begin{aligned} F_{p_1} = 0 &\Rightarrow \frac{(1 - 4p_1)}{p_1} = \frac{2(b^2 - a^2)(1 - 2p_1)}{[2(b^2 - a^2)p_1 + a^2]} & (\text{A.6}) \\ &\Rightarrow \frac{2(b^2 - 3a^2)}{a^2} = \frac{(1 - 4p_1)}{p_1} \\ &\Rightarrow p_1 = \frac{a^2}{2(b^2 - a^2)}. \end{aligned}$$

Hence from (A.5) we get

$$\frac{2(b^2 - 3a^2)}{a^2} = \frac{2(b^2 - a^2)(b^2 - 2a^2)}{2a^2(b^2 - a^2)} \Rightarrow a = \frac{b}{2} \quad \text{and hence} \quad p_1 = \frac{1}{6}. \quad (\text{A.7})$$

Therefore the function F has an optimum at $p_1 = \frac{1}{6}$ and $a = \frac{b}{2}$. We have to check if this is a maximum or a minimum. We first differentiate (A.3) with respect to a and p_1 and obtain F_{aa} and F_{ap_1} respectively, where

$$F_{aa} = -\frac{2}{a^2} - \frac{4(b^2 + a^2)}{(b^2 - a^2)^2} - \frac{2(1 - 2p_1)[2(b^2 + a^2)p_1 - a^2]}{[2(b^2 - a^2)p_1 + a^2]^2}, \quad (\text{A.8})$$

$$F_{ap_1} = \frac{4ab^2}{[2(b^2 - a^2)p_1 + a^2]^2}. \quad (\text{A.9})$$

Then differentiating (A.4) with respect to a and p_1 we obtain F_{p_1a} and $F_{p_1p_1}$ respectively, where

$$F_{p_1a} = \frac{4ab^2}{[2(b^2 - a^2)p_1 + a^2]^2}, \quad (\text{A.10})$$

$$F_{p_1p_1} = -\frac{1}{p_1^2} - \frac{4}{(1 - 2p_1)^2} + \frac{4(b^2 - a^2)^2}{[2(b^2 - a^2)p_1 + a^2]^2}. \quad (\text{A.11})$$

The Hessian matrix $H(F)$ obtained in this case is given by

$$H(F) = \begin{bmatrix} F_{aa} & F_{ap_1} \\ F_{p_1a} & F_{p_1p_1} \end{bmatrix}. \quad (\text{A.12})$$

At $p_1 = \frac{1}{6}$ and $a = \frac{b}{2}$, $H(F)$ takes the value:

$$H(F) = \begin{bmatrix} -\frac{160}{9b^2} & \frac{8}{b} \\ \frac{8}{b} & -36 \end{bmatrix},$$

which is a negative definite matrix. Hence, F is maximum at $p_1 = \frac{1}{6}$ and $a = \frac{b}{2}$.

A.2 Maximization of J for any b

The fitted value criterion \mathbf{J}_b in (A.1) is maximum when $p_2 = 2p_1 = \frac{1}{3}$ and $a = \frac{b}{2}$ and the maximum value is $\mathbf{J}_{b[\max]} = \frac{nb^6}{16}$. Clearly $\mathbf{J}_{b[\max]}$ is monotonically increasing in b where $0 < b \leq 1$. Hence the maximum value of \mathbf{J}_b i.e. $\mathbf{J}_{b[\max]}$ attains its maximum when $b = 1$.

Appendix B

B.1 Maximization of K_J for a given b

The criterion K_J in Chapter-6 is given by

$$\frac{K_J}{n^2} = \frac{8a^2b^2(b^2 - a^2)^4 p_1^2 p_2^2}{(p_1 b^2 + p_2 a^2)} \quad \text{where} \quad (p_1 + p_2) = \frac{1}{2}. \quad (\text{B.1})$$

Taking the common logarithm of both sides and assuming $F = \log\left(\frac{K_J}{n^2}\right)$ we get

$$F = \log 8 + 2 \log a + 2 \log b + 4 \log(b^2 - a^2) + 2 \log p_1 + 2 \log\left(\frac{1}{2} - p_1\right) - \log\left[(b^2 - a^2)p_1 + \frac{a^2}{2}\right]. \quad (\text{B.2})$$

Now, differentiating (B.2) with respect to a and p_1 we get F_a and F_{p_1} respectively where

$$F_a = \frac{2}{a} - \frac{8a}{(b^2 - a^2)} - \frac{2a(1 - 2p_1)}{[2(b^2 - a^2)p_1 + a^2]}, \quad (\text{B.3})$$

$$F_{p_1} = \frac{2}{p_1} - \frac{4}{1 - 2p_1} - \frac{2(b^2 - a^2)}{[2(b^2 - a^2)p_1 + a^2]}. \quad (\text{B.4})$$

Equating each of (B.3) and (B.4) to zero, we get

$$F_a = 0 \Rightarrow \frac{(b^2 - 5a^2)}{a^2} = \frac{(b^2 - a^2)(1 - 2p_1)}{[2(b^2 - a^2)p_1 + a^2]} \quad (\text{B.5})$$

$$\begin{aligned} F_{p_1} = 0 &\Rightarrow \frac{(1 - 4p_1)}{p_1} = \frac{(b^2 - a^2)(1 - 2p_1)}{[2(b^2 - a^2)p_1 + a^2]} \quad (\text{B.6}) \\ &\Rightarrow \frac{(b^2 - 5a^2)}{a^2} = \frac{(1 - 4p_1)}{p_1} \\ &\Rightarrow p_1 = \frac{a^2}{(b^2 - a^2)}. \end{aligned}$$

Now, from (B.5) we get

$$\begin{aligned} \frac{(b^2 - 5a^2)}{a^2} &= \frac{(b^2 - 3a^2)}{3a^2} \\ \Rightarrow a &= \frac{b}{\sqrt{6}} \quad \text{and hence} \quad p_1 = \frac{1}{5}. \end{aligned} \quad (\text{B.7})$$

Therefore for a given b , F is optimum when $a = \frac{b}{\sqrt{6}}$ and $p_1 = \frac{1}{5}$. Therefore the function F has an optimum at $p_1 = \frac{1}{5}$ and $a = \frac{b}{\sqrt{6}}$. We have to check if this is a maximum or a minimum. We first differentiate (B.3) with respect to a and p_1 and obtain F_{aa} and F_{ap_1} respectively, where

$$F_{aa} = -\frac{2}{a^2} - \frac{8(b^2 + a^2)}{(b^2 - a^2)^2} - \frac{2(1 - 2p_1)[2(b^2 + a^2)p_1 - a^2]}{[2(b^2 - a^2)p_1 + a^2]^2}, \quad (\text{B.8})$$

$$F_{ap_1} = \frac{4ab^2}{[2(b^2 - a^2)p_1 + a^2]^2}. \quad (\text{B.9})$$

Then differentiating (B.4) with respect to a and p_1 we obtain $F_{p_1 a}$ and $F_{p_1 p_1}$ respectively, where

$$F_{p_1 a} = \frac{4ab^2}{[2(b^2 - a^2)p_1 + a^2]^2}, \quad (\text{B.10})$$

$$F_{p_1 p_1} = -\frac{2}{p_1^2} - \frac{8}{(1 - 2p_1)^2} + \frac{4(b^2 - a^2)^2}{[2(b^2 - a^2)p_1 + a^2]^2}. \quad (\text{B.11})$$

The Hessian matrix $H(F)$ obtained in this case is given by

$$H(F) = \begin{bmatrix} F_{aa} & F_{ap_1} \\ F_{p_1 a} & F_{p_1 p_1} \end{bmatrix}. \quad (\text{B.12})$$

At $p_1 = \frac{1}{5}$ and $a = \frac{b}{\sqrt{6}}$, $H(F)$ takes the value:

$$H(F) = \begin{bmatrix} -\frac{672}{25b^2} & \frac{16}{b\sqrt{6}} \\ \frac{16}{b\sqrt{6}} & -\frac{550}{9} \end{bmatrix},$$

which is a negative definite matrix. Hence, F is maximum at $p_1 = \frac{1}{5}$ and $a = \frac{b}{\sqrt{6}}$.

B.2 Maximization of K_J for any b

The criterion K_J in (B.1) is maximum when $p_1 = \frac{1}{5}$, $p_2 = \frac{3}{10}$ and $a = \frac{b}{\sqrt{6}}$ and the maximum value is $K_{J[max]} = \frac{b^8}{108}$. Clearly $K_{J[max]}$ is monotonically increasing in b where $0 < b \leq 1$. Hence the maximum value of K_J i.e. $K_{J[max]}$ attains its maximum when $b = 1$.