UNIVERSITY OF CALIFORNIA
RIVERSIDE

Characterization of Special Variance Structures for Designs in Model Identification and Discrimination

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

Doctor of Philosophy

in

Applied Statistics

by

Analisa Marielena Flores

August 2011

Dissertation Committee:
Dr. Subir Ghosh, Chairperson
Dr. Barry Arnold
Dr. Robert Hanneman

The Dissertation of Analisa Marielena Flores is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I owe my deepest gratitude and heartfelt thanks to many people who have supported me throughout this work.

First and foremost, I would like to thank my advisor, Dr. Subir Ghosh – a man of great strength, patience, dedication, knowledge, and kindness. I have learned a tremendous amount from him and am forever indebted to the many hours he has spent guiding my research.

I would also like to express my appreciation to Dr. Barry Arnold and Dr. Robert Hanneman for serving on my dissertation committee and giving their time, ideas, and comments.

There are also many people in the Statistics Department to thank. To the professors: I am thankful for the education, support, skills, and funding provided during my time as a graduate student. I am also very grateful to Dr. Linda Penas and Paula Lemire for their constant support and friendship. I would also like to thank Perla Fabelo for her help and direction while completing my graduate studies, as well as the many graduate students that I now call friends.

Most importantly, I am forever grateful to my family and friends who have been by my side throughout this incredible journey. To my parents, Jose and Karen Vega, whose unwavering love and support has allowed me to be where I am today. To my brother, Rocaciano Vega, whose friendship and support is deeply cherished and appreciated. To my best friends: Patty Lin-Koster, Vasilios Suarez, Tim Olsen, and Mark Jesunathadas, who have always made time to keep our friendship strong and for that, I am eternally grateful. I also thank the many other people in my life, old and new, that have shown their support.

Last but not least, I thank my husband, Mark Andrew Flores, for his love, friendship, and encouragement. His support has given me the strength and determination to complete this work. With him by my side, all things are possible.

ABSTRACT OF THE DISSERTATION

Characterization of Special Variance Structures
for Designs in Model Identification and Discrimination

by

Analisa Marielena Flores

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

Models containing the general mean, main effects, and all possible k two-factor interaction effects are considered for factorial experiments with m factors observed at two levels each. Specifically, fractional factorial designs consisting of n runs which permit the identification and discrimination of the models of interest are evaluated and classified. The classifications are dependent on a new property introduced in this dissertation, denoted P_g^V, $g \geq 1$, which relies on the structure of the variance-covariance matrix for the estimates of the model parameters. Designs with the property P_g^V, $g \geq 1$, permit to divide the models in the class considered into g groups so that all models in a group have equal variances for the least squares estimates of the k two-factor interaction effects. Ghosh-Tian optimum designs [Ghosh and Tian (2006)] for $m = 4$, $n = 6,...,11$, $k = 1$ and $m = 5$, $n = 7,...,16$, $k = 1$ are classified with respect to g values, $g \geq 1$, in the property P_g^V and presented through illustrative examples.

Several characterizations of designs with P_g^V, $g \geq 1$, are provided for the case when $k = 1$. Designs such as balanced designs, isomorphic designs and complementary designs with the property P_g^V are proven to have P_1^V for $k = 1$. Tables identifying all such balanced designs are provided. It is noted that although D9.2 and D14 are not balanced, these designs in fact take a special form resulting in P_1^V for k =1. These special forms are investigated in depth.

In addition, the construction of all designs giving P_1^V for $m = 3$, $n = 5,6,7,8$, $k = 1$ and $m = 4$, $n = 6,...,16$, $k = 1$ are described. Further, occurrences of P_1^V are presented for fractional factorial designs when $m = 5$, $n = 7,8,9$, $k = 1$.

Finally, additional characterizations of P_g^V, $g \geq 1$, when $k > 1$ are given and illustrated through various examples. Special designs are presented with the property P_1^V for $\max k$ which is the largest value of k in the models of interest that the design has the ability to identify and discriminate. It is shown that balanced designs will have P_1^V for $k = \binom{m}{2}$. Tables identifying all such balanced designs are provided.

Table of Contents

List of Tables

List of Figures

Chapter 1

General Linear Model

1.0 Introduction

Researchers often study the dependence of a response variable on multiple independent variables known as factors. To study this dependence, an experiment is performed. Then to obtain an adequate understanding of the results from this experiment, particularly the relationships of the factors with the response and among themselves, a model (or more likely, a set of models) is fitted to the collected data. The researcher does not know the true model for the data but will generally have an idea of a set of models which will possibly fit/explain the data well. From these possible models, the best model is then selected using a criterion function.

1.1 Model Information

The classical linear regression model is given as

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \varepsilon \qquad (1.1)$$

where y is the dependent variable (observed response), $x_1, \ldots, x_m$ are the independent variables (factors), β_0 is the intercept, $\beta_1, \ldots, \beta_m$ are the regression coefficients (effects of factors), and ε is the error due to effects of unspecified independent variables and/or a random element in the specified relationship. Because this model is the same for i, $i = 1, \ldots, n$, observations, it is common to write this model in matrix notation as

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon}, \tag{1.2}$$

where $\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & & x_{2m} \\ \vdots & & & & \\ 1 & x_{n1} & x_{n2} & & x_{nm} \end{bmatrix}$, $\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$, $\underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$,

with m representing the number of independent variables and n representing the number of observations obtained. Moreover, X is a known matrix with $rank(X) \leq m+1$ (for our discussion, we will have $rank(X) = m+1$) and it is assumed $E[\underline{\varepsilon}] = \underline{0}$ and $Var[\underline{\varepsilon}] = \sigma^2 I$ where σ^2 is unknown. With these assumptions, the model can be written as

$$E[\underline{y}] = X\underline{\beta},$$

$$Var[\underline{y}] = \sigma^2 I, \tag{1.3}$$

where $X\left(n\times(1+m)\right)$ is dependent on the design of the experiment and $\underline{\beta}\left((1+m)\times 1\right)$ is the vector of unknown parameters. These unknown parameters are estimated by fitting the data to the specified model.

1.2 Least Squares Estimation

The fitted model is written as

$$\underline{\hat{y}} = X\underline{\hat{\beta}}, \tag{1.4}$$

where the elements of $\underline{\hat{y}}$ are the fitted (predicted) values of the true values of the elements of $\underline{y}$ and $\underline{\hat{\beta}}$ is the estimate for the unknown vector $\underline{\beta}$. To estimate, $\underline{\beta}$, the method of Least Squares Estimation (LSE) is commonly employed. This method uses the data, $\underline{y}$, to estimate $\underline{\beta}$ such that the sums of squares of residuals/errors (SSE) is minimized where the residuals, r, are defined as $r_i = y_i - \hat{y}_i \quad i = 1,\ldots,n$; thus, $SSE = \sum_{i=1}^{n} r_i^2 = \underline{r}'\underline{r}$. This procedure gives

$$\hat{\underline{\beta}} = \left(X'X\right)^{-1} X'\underline{y}, \tag{1.5}$$

where $rank(X'X) = m+1$ (full rank), thus ensuring a solution exists for $\left(X'X\right)^{-1}$ [Draper and Smith (1998), Johnson and Wichern (2002)]. It can be shown that $\hat{\underline{\beta}}$ is an unbiased estimate of $\underline{\beta}$. Further, when the observations are normally distributed, $\hat{\underline{\beta}}$ is the best minimum variance unbiased estimate of $\underline{\beta}$. The variance-covariance matrix for $\hat{\underline{\beta}}$ is

$$Var\left[\hat{\underline{\beta}}\right] = \sigma^2 \left(X'X\right)^{-1}, \tag{1.6}$$

where the diagonal elements represent the variances of the estimates in $\hat{\underline{\beta}}$ and the off-diagonal elements are the covariances between these estimates. This variance-covariance matrix is a key component in comparing designs for an experiment.

A design for a given experiment is the set of treatments (combinations of factor levels) that will be examined in the experiment. We define a particular class of designs so that $rank\left(X\right) = m+1$ in (1.3), ensuring all parameters in the regression model can be estimated (i.e., $\hat{\underline{\beta}}$). Because there are several designs for a specified value of n which are included in this class, it is of interest to select the design which will optimize the amount of information contained in the experiment. In the following section, we define criterion functions for comparing the designs within this class.

1.3 Optimality Criterion Functions

The optimality criterion functions compare designs by considering the variance-covariance matrix of $\underline{\hat{\beta}}$ [Kiefer (1959)]. Because σ^2 is unknown, features of the $\left(X'X\right)^{-1}$ matrix are examined. In particular, the A-optimality criterion minimizes the trace of $\left(X'X\right)^{-1}$, the D-optimality criterion minimizes the determinant of $\left(X'X\right)^{-1}$, and the E-optimality criterion minimizes the maximum characteristic root (eigenvalue) of $\left(X'X\right)^{-1}$ [Kiefer (1959)].

Example 1.1

Consider an experiment conducted to determine the effect of machining factors on ceramic strength [Ives, Jahanmir, Gill, and Filliben (1998)]. The variables for the experiment are defined as

$y =$ mean ceramic strength

$x_1 =$ table speed (slow (0.025 m/s) or fast (0.125 m/s))

$x_2 =$ down feed rate (slow (0.05 mm) or fast (0.125mm))

$x_3 =$ wheel grit (140/170 or 80/100)

$x_4 =$ direction (longitudinal or transverse)

$x_5 =$ batch (1 or 2)

The researcher would like to compare two designs with eleven observations which permit the LSE of the parameters in equation (1.2) [Ghosh and Tian (2006)], Design A1 and Design A2. The designs are presented in Figure 1.1 where each row represents an observation, each column represents an independent variable, and a "–" indicates the variable is observed at the low level, while a "+" indicates the variable is observed at the high level. For example, consider the first observation in Design A1, (+,+,+,+,–). The ceramic strength would be observed while the table speed = 0.125 m/s, down feed rate = 0.125 mm, wheel grit = 80/100, direction = transverse, and batch = 1. To choose the better of the two designs, the A-, D-, and E-optimality criterion functions are considered and presented in Table 1.1 and Figure 1.2.

Design A1

+	+	+	+	–
+	+	+	–	+
+	+	–	+	+
+	–	+	+	+
–	+	+	+	+
+	–	–	–	–
–	+	–	–	–
–	–	+	–	–
–	–	–	+	–
–	–	–	–	+
–	–	+	+	+

Design A2

+	+	+	+	–
+	+	+	–	+
+	+	–	+	+
+	–	+	+	+
–	+	+	+	+
+	–	–	–	–
–	+	–	–	–
–	–	+	–	–
–	–	–	+	–
–	–	–	–	+
–	–	–	–	–

Figure 1.1: Design A1 and Design A2

Table 1.1: Comparison of Design A1 and Design A2 using A-,D-,E-optimality

Criterion	Design A1	Design A2
$\text{Trace}\left((X'X)^{-1}\right)$	0.605	0.637
$\text{Det}\left((X'X)^{-1}\right)$	7.926×10^{-7}	9.844×10^{-7}
$\text{MCR}\left((X'X)^{-1}\right)$	0.125	0.125

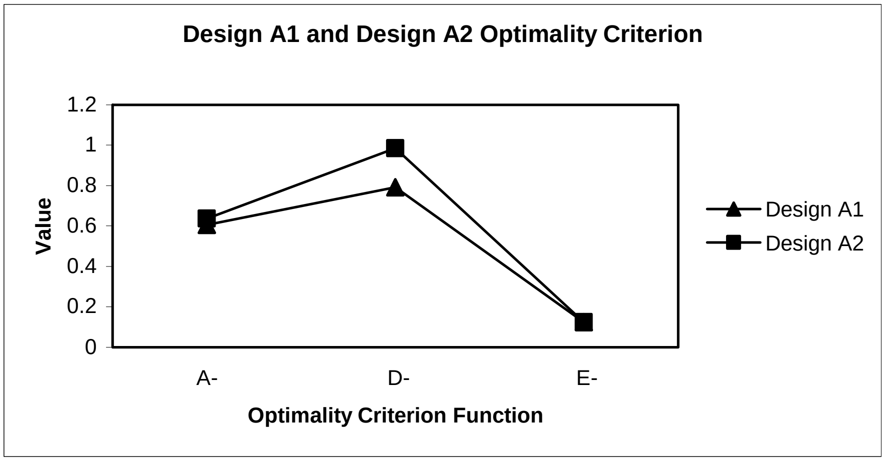

Figure 1.2: Comparison of Design A1 and Design A2 using A-,D-,E-optimality
Note: The values for the determinants have 10^{-6} dropped

The trace and determinant of the $\left(X'X\right)^{-1}$ matrix of Design A1 are smaller than those of Design A2, while the maximum characteristic root of $\left(X'X\right)^{-1}$ for the designs are equivalent. Therefore, Design A1 is better than Design A2 with respect to (w.r.t.) A- and D- optimality but the designs are equivalent with respect to E-optimality.

1.4 Research Objectives and Dissertation Outline

A crucial component to understanding and interpreting the results of research is fitting a model to the data of an experiment. In doing so, the effects of the independent variables on the dependent variable are realized. Although a researcher does not know prior to conducting the experiment which model will fit the data best, a set of possible models can be postulated to fit the data well. From this set of models, the best model can then be selected. However, it is necessary that the experiment is designed such that it will have the ability to identify and discriminate the various models which will later be used to describe the data. As such, it is a vital step in the research process to design an experiment with these capabilities.

A particular type of experimental design, called a factorial design, is employed when a researcher is not only interested in the effects of m independent variables on the dependent variable, but also possible interaction effects of the independent variables on the dependent variable. The researcher will generally have an idea of the order of interaction which exists between the independent variables or factors. Often, higher order interaction effects are assumed negligible. Nevertheless, the researcher does not know before the experiment which factors actually interact. Thus, it is important for the experiment to be designed not only with respect to overall optimality criterion functions, but such that all of the possible models, containing the various interaction effects, are fitted with equal precision.

In a 2^m fractional factorial experiment with n runs, there are $s = \binom{m}{2}$ two-factor interaction effects which can be considered. Then, the number of models, ν, to be considered for describing the data will depend on the number of interactions to be included in the model. If there are k two-factor interactions to be included in the model with the general mean and m main effects, there are $\nu = \binom{s}{k}$ possible models. Because the uncommon elements of any two models are the possible elements from k two-factor interactions, it is important to consider the measure of precision for which these elements are being estimated. Therefore, it is of interest to identify the fractional factorial designs which have the ability to unbiasedly estimate the $(1+m+k)$ parameters in each of the ν models while giving equal precision for the k two-factor interaction effect estimates. Such precision is measured using the estimated variance for each effect; the smaller the variance, the higher the precision.

In Chapter 2, fractional factorial designs and the associated linear models are introduced. In Chapter 3, the search linear model and optimality criterion functions are discussed, as well as a special class of designs known as balanced designs. In Chapter 4, a new property, denoted P_g^V, pertaining to the precision of the estimates is defined and illustrated. In Chapter 5, several characterizations of designs with P_g^V for g=1, denoted P_1^V, are given and explained through various proofs and examples. In Chapter 6, the fractional factorial designs for m=3,4 with P_1^V when $k = 1$ are identified through systematic decomposition of the full factorial designs. Additionally, the occurrences of P_1^V when $k = 1$ are presented in a table for m=5 when n= 7, 8, 9. Further, additional cases for P_g^V are discussed and illustrated by example; specifically for i) $k = 1, g > 1$, ii) $k > 1$, $g = 1$, and iii) $k > 1, g > 1$. In Chapter 7, tables of designs with P_1^V for m=3,4,5 for $k = 1$ are presented for the convenience and use of the reader. Chapter 8 presents the conclusions of the research contained here within, highlighting the most significant contributions.

Chapter 2

Experimental Designs

2.0 Introduction

Obtaining the best model is essential for describing data and being able to make appropriate inferences, allowing for the best understanding of the information contained in the results of an experiment. Thus, an experiment must be designed in such a way that these research objectives can be accomplished. A design for a given experiment is the set of treatments (combination of factor levels) that will be examined in the experiment.

2.1 Factorial Experiments

Factorial experiments are widely used in both the social and natural sciences. This type of experiment has multiple factors (independent variables) at multiple levels with each combination of those factor levels assigned to an experimental unit. A factorial experiment is designed to allow the researcher to examine the main effects of the factors and possible interactions between the factors.

Consider an experiment which contains m factors which can be observed at two levels each. This type of experiment is denoted as a 2^m factorial experiment and in order to estimate all possible effects (main and interaction), one would require 2^m runs. For example, suppose an experiment consists of three factors with two levels each. To run a full factorial experiment, $2^3 = 8$ treatments would be needed to estimate all possible effects. These effects include the main effects of the three factors, all possible two-factor interaction effects (there are three), and the three-factor interaction effect. In addition, replications of these treatments would be needed for complete analysis of the effects. Although the number of treatments needed here does not seem to be problematic, now consider a 2^4 factorial experiment. This experiment would require $2^4 = 16$ treatments (plus replications) to estimate and analyze all possible effects. Similarly, to estimate and analyze all possible effects in a 2^5 factorial experiment, 32 treatments (plus replications) are required. The number of runs needed in a full factorial experiment quickly becomes overwhelming and cumbersome as the number of factors (or the number of levels of a factor) increases.

2.2 Fractional Factorial Designs

Generally, higher order interaction effects in factorial experiments are assumed negligible and therefore do not need to be estimated. This allows for a reduction in the number of runs required, consequently reducing the expense of conducting the experiment. The experiment then becomes what is known as a fractional factorial experiment, where only a fraction (subset) of the combinations of factor levels is considered. However, the runs that are used must be chosen carefully so that the effects of interest can be estimated. Choosing these runs, or combination of factor levels (treatments), can be done at the design stage of the experiment and thus becomes a vital step in the experimental process. If the best treatments are evaluated in the experiment, the results will be more informative for model selection and analysis.

Example 2.1

Refer to Example 1.1 which considers an experiment investigating the effect of five machining factors on ceramic strength, where each factor can be observed at two levels. The five factors are labeled A, B, C, D, E where "–" and "+" indicate the factor is observed at it's low and high level, respectively, as illustrated in Table 2.1. This type of experiment is in fact a 2^5 factorial experiment. However, recall Design A1 and Design A2 consist of only eleven runs each (instead of $2^5 = 32$) and therefore are fractional factorial designs.

Table 2.1: Factors and levels for Example 2.1

	Factor A (table speed)	Factor B (down feed rate)	Factor C (wheel grit)	Factor D (direction)	Factor E (batch)
–	0.025 m/s	0.05mm	140/170	longitudinal	1
+	0.125 m/s	0.125mm	80/100	transverse	2

2.3 Main Effects and Interactions

Because factorial designs and fractional factorial designs have the ability to estimate the effects of the factors, as well as the effects of interactions between these factors, it is necessary to distinguish between the two types of effects. The main effect of a factor is defined to be the change in the response produced by a change in the level of the factor, averaged over the levels of the remaining factors. This idea is illustrated using a single Factor A in Figure 2.1.

Similarly, an interaction occurs when the difference in the response between the levels of one factor is not the same at all levels of the other factor, when averaged over the levels of the remaining factors. A two-factor interaction effect between Factor A and Factor B is demonstrated in Figure 2.2. This idea can be extended to higher order interaction effects.

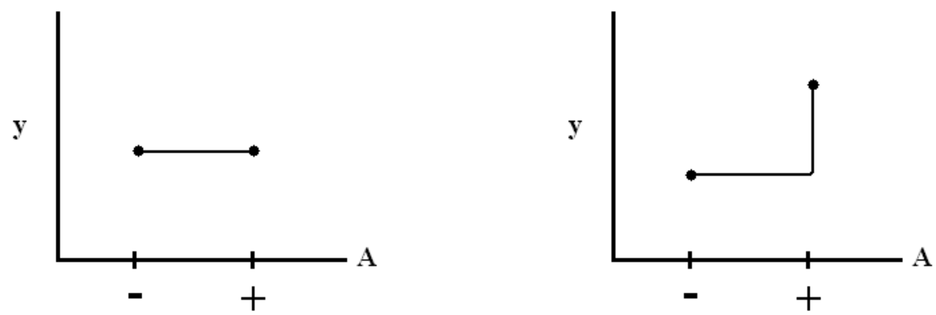

Figure 2.1: No main effect of Factor A (left). Main effect of Factor A (right).

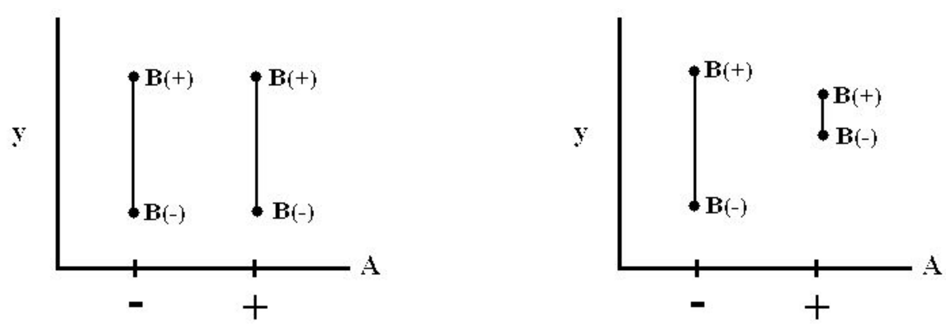

Figure 2.2: No interaction effect between Factor A and Factor B (left). Interaction effect between Factor A and Factor B (right).

2.4 Model Information

As indicated in Section 2.1, factorial designs and fractional factorial designs are not only capable of estimating the m main effects of the independent variables in an experiment, but also possible effects of interactions between the factors. Since the model will have to account for the interactions, the model as stated in (1.1) must be expanded.

The linear model for a 2^m full factorial design with m factors includes the general mean (intercept), m main effects, and all possible interaction effects,

$$\begin{aligned} y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \beta_{12} x_1 x_2 + \cdots \\ + \beta_{(m-1)m} x_{m-1} x_m + \cdots + \beta_{1,2,\cdots,m} x_1 x_2 \cdots x_m + \varepsilon \end{aligned} \tag{2.1}$$

where y is the observed response, $x_1,...,x_m$ are the independent variables (factors) which are observed at –1 (low level) or +1 (high level) , β_0 is the general mean, $\beta_1,\ldots,\beta_m$ correspond to the main effects of factors, $\beta_{12},...\beta_{(m-1)m}$ correspond to the two-factor interaction effects, and so on, with $\beta_{1,2,...,m}$ corresponding to the $m-$factor interaction effect. Moreover, $rank(X) = 2^m$ and the same assumptions from Section 1.1 hold giving model (2.1) as model (1.3) where $X\left(n \times 2^m\right)$ is dependent on the design of the experiment and $\underline{\beta}(2^m \times 1)$ is the vector of unknown parameters.

Since unreplicated fractional factorial designs do not have the ability to estimate the general mean, main effects, and all interactions at the same time, higher order interactions are often assumed to be negligible. Thus, the model for a 2^m fractional factorial design will include the general mean, main effects, and a subset of possible interaction effects.

Assuming two-factor and higher order interactions are negligible, model (2.1) reduces to model (1.1). A fractional factorial design for this model, with corresponding matrix X having full rank ($rank(X) = 1 + m$), is known as a main effects design.

Assuming three-factor and higher order interactions are negligible, model (2.1) reduces to

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \beta_{12} x_1 x_2 + \cdots + \beta_{(m-1)m} x_{m-1} x_m + \varepsilon \ . \qquad (2.2)$$

The matrix notation for this model is model (1.3) where $X\left(n \times (1+m+s)\right)$, $s = \binom{m}{2}$, is the design matrix and $\underline{\beta}\left((1+m+s) \times 1\right)$ is the vector of unknown parameters corresponding to the general mean, m main effects, and all two-factor interaction effects. When $rank(X) = 1+m+s$, all parameters in model (2.2) can be estimated.

Example 2.2

Model (2.2) for $m = 4$ factors is written as

$$\begin{aligned} y = \beta_o &+ \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \\ &+ \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \varepsilon. \end{aligned} \qquad (2.3)$$

Consider Design A3 for a 2^4 fractional factorial experiment with $n = 11$ runs, presented in Figure 2.3. The design matrix, X, for this design has $rank(X) = 1+m+s = 1+4+6 = 11$, indicating it is capable of estimating all parameters in model (2.3).

In addition, model (2.2) for $m = 5$ factors is written as

$$\begin{aligned} y = \beta_o &+ \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \\ &+ \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{15} x_1 x_5 + \beta_{23} x_2 x_3 \\ &\quad + \beta_{24} x_2 x_4 + \beta_{25} x_2 x_5 + \beta_{34} x_3 x_4 + \beta_{35} x_3 x_5 + \beta_{45} x_4 x_5 + \varepsilon. \end{aligned} \qquad (2.4)$$

Considering Design A4 for a 2^5 fractional factorial experiment with $n = 16$ runs, the design matrix, X, has $rank(X) = 1+m+s = 1+5+10 = 16$. Therefore, this design is capable of estimating all parameters in model (2.4).

Moreover, it can be shown that Design A3 and Design A4 are optimal designs for $m = 4$ and $m = 5$, respectively, with respect to the optimality criterion described in Section 1.3 [Ghosh and Tian (2006)]. Note that Design A3 and Design A4 are optimum designs D7 and D18 from Ghosh-Tian. The values for these criterion are given in Table 2.2.

Design A3

−	−	−	−
+	+	−	−
+	−	+	−
+	−	−	+
−	+	+	−
−	+	−	+
−	−	+	+
+	+	+	−
+	+	−	+
+	−	+	+
−	+	+	+

Design A4

−	−	−	−	−
+	+	−	−	−
+	−	+	−	−
+	−	−	+	−
+	−	−	−	+
−	+	+	−	−
−	+	−	+	−
−	+	−	−	+
−	−	+	+	−
−	−	+	−	+
−	−	−	+	+
+	+	+	+	−
+	+	+	−	+
+	+	−	+	+
+	−	+	+	+
−	+	+	+	+

Figure 2.3: Design A3 and Design A4

Table 2.2: Design A3 and Design A4 A-, D- ,E-optimality criterion values

Criterion	Design A3	Design A4
$\text{Trace}\left((X'X)^{-1}\right)$	1.486	1.000
$\text{Det}\left((X'X)^{-1}\right)$	2.587×10^{-11}	5.421×10^{-20}
$\text{MCR}\left((X'X)^{-1}\right)$	0.25	0.0625

Chapter 3

Search Linear Models

3.0 Introduction

Often in fractional factorial experiments, it may be of interest to assume three-factor and higher order interactions are negligible while keeping the general mean, main effects, and a subset of two-factor interaction effects (as opposed to all two-factor interaction effects) in the model. This results in more than one possible model. Because the design must be chosen prior to conducting the experiment and the true model for describing the collected data-to-be is unknown, the design must be capable of estimating all possible models.

3.1 Search Linear Model

Consider the following linear model

$$E\left[\underline{y}\right] = X_1\underline{\beta_1} + X_2\underline{\beta_2},$$

$$Var\left[\underline{y}\right] = \sigma^2 I,$$

$$Rank\left(X_1\right) = p_1, \tag{3.1}$$

where $\underline{y}\left(n\times 1\right)$ is the vector of observed responses, $\underline{\beta_1}\left(p_1\times 1\right)$ is a vector of unknown parameters, $\underline{\beta_2}\left(p_2\times 1\right)$ is a vector unknown parameters with partial information on them, $X_1\left(n\times p_1\right)$ is the design matrix, $X_2\left(n\times p_2\right)$ is a matrix dependent on X_1, and σ^2 is an unknown constant. It is known that k of the p_2 parameters in $\underline{\beta_2}$ are non-zero and the remaining $(p_2 - k)$ parameters are zero. It is unknown which k parameters are non-zero. This model is called a search linear model [Srivastava (1975)] because it has the ability to search for the k non-zero parameters in $\underline{\beta_2}$. While a search linear model has the ability to identify the k non-zero parameters in $\underline{\beta_2}$, a search design has the ability to estimate the $(p_1 + p_2)$ unknown parameters. A search design satisfies

$$Rank\left[X_1 : X_{2j}\right] = p_1 + 2k \tag{3.2}$$

where X_{2j}, $j = 1, \ldots, \binom{p_2}{2k}$, are the $(n \times 2k)$ submatrices of X_2.

3.2 Model Information

Consider a special search linear model (3.1) as it applies to a fractional factorial design. Then $\underline{\beta_1}$ represents the vector of parameters corresponding to the general mean and m main effects with $p_1 = 1 + m$ and $\underline{\beta_2}$ represents the vector of parameters corresponding to the interaction effects. Particularly, assume three-factor and higher order interactions are assumed negligible so that $\underline{\beta_2}$ will consist of two-factor interactions only with $p_2 = \binom{m}{2} = s$.

There are $v = \binom{s}{k}$ possible models with the general mean, m main effects and k two-factor interaction effects which can each be written as model (1.3) where

$X\left(n \times (1+m+k)\right)$ is the design matrix and $\underline{\beta}((1+m+k) \times 1)$ is the vector of unknown parameters. For any two of the possible ν models, the general mean and main effects are common parameters, while some or all k two-factor interaction effects are different. A design which is capable of estimating the β parameters of these ν models is called a main effect plus k two-factor interactions design.

Moreover, the ν models can be written as

$$E\left(\underline{y}\right) = X_1 \underline{\beta_1} + X_2^{(u)} \underline{\beta_2^{(u)}},$$

$$Var\left[\underline{y}\right] = \sigma^2 I,$$

$$Rank\left(X_1, X_2^{(u)}\right) = 1+m+k, \; u = 1, \ldots, \nu, \tag{3.3}$$

where $\underline{\beta_1}\left((1+m) \times 1\right)$ is the vector with the unknown parameters corresponding to the general mean and m main effects, $\underline{\beta}_2^{(u)}\left(k \times 1\right)$ is the vector of $k \le s$ unknown parameters corresponding to the two-factor interaction effects, $X_1\left(n \times (1+m)\right)$ is the design matrix, $X_2^{(u)}\left(n \times k\right)$ is dependent on X_1 and the selected k two-factor interactions, and σ^2 is unknown. Letting $X^{(u)} = \left[X_1 \quad X_2^{(u)}\right]$ and $\underline{\beta}^{(u)} = \left[\beta_1 \quad \beta_2^{(u)}\right]'$, the LSE of $\underline{\beta}^{(u)}$ under (3.3) is denoted $\underline{\hat{\beta}}^{(u)}$ and $V^{(u)} = \sigma^{-2} Var\left[\hat{\beta}^{(u)}\right] = \left(X^{(u)\prime} X^{(u)}\right)^{-1}$. It should be noted that condition (3.3) allows for identification of the models, while condition (3.2) is necessary for discrimination between the models. A fractional factorial design meeting these conditions is a search design.

Example 3.1

Design A1 and Design A2 presented in Example 1.1 have the ability to identify the models with the general mean, $m=5$ main effects and $k=1$ two-factor interaction effect for a 2^5 fractional factorial experiment with 11 runs. These designs satisfy condition (3.3) giving $Rank\left(X_1, X_2^{(u)}\right)=1+5+1=7,\ u=1,\ldots,10$ and therefore are capable of estimating the $1+m+k=1+5+1=7$ parameters in each of $v=\begin{pmatrix} s \\ k \end{pmatrix}=\begin{pmatrix} 10 \\ 1 \end{pmatrix}=10$ models. Moreover, each design satisfies condition (3.2) with $Rank(X_1 : X_{2j})=1+m+2k=8,\ \forall j=1,\ldots,\begin{pmatrix} 10 \\ 2 \end{pmatrix}$ classifying Design A1 and Design A2 as search designs. Therefore, these designs have the ability to identify and discriminate between the ten models with the distinguishing factor between the models being the two-factor interaction effect. The two-factor interaction effect for the specified values of u are presented in Table 4.2 for $m=4,5$. Note that the remaining chapters will use these same specifications when referring to values of u .

Table 3.1: Specified two-factor interaction for values of *u*

Number of Factors	u									
	1	2	3	4	5	6	7	8	9	10
$m=4$	AB	AC	AD	BC	BD	CD				
$m=5$	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE

The ten models are given as

$$\text{Model 1:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{12}x_1x_2$$

$$\text{Model 2:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{13}x_1x_3$$

$$\text{Model 3:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{14}x_1x_4$$

$$\text{Model 4:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{15}x_1x_5$$

$$\text{Model 5:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{23}x_2x_3$$

$$\text{Model 6:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{24}x_2x_4$$

$$\text{Model 7:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{25}x_2x_5$$

$$\text{Model 8:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{34}x_3x_4$$

$$\text{Model 9:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{35}x_3x_5$$

$$\text{Model 10:}\quad E[y]=\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_{45}x_4x_5$$

After the experiment is performed, data are collected, and the parameters are estimated for all ten models, the model with the smallest SSE would be selected as the best model.

3.3 Optimality Criterion Functions

To select the best design from a class of designs that are capable of estimating the ν models, optimality criterion functions are considered. Similar to the optimality criterion functions discussed in Section 1.3, the variance-covariance matrix, $V^{(u)}$ as defined in Section 3.2 are considered, particularly the matrices $\left(X^{(u)\prime}X^{(u)}\right)^{-1}$ for $u=1,\ldots,\nu$. There are six optimality criterion functions [Srivastava (1977)] defined for a set of models; namely, AT, GT, AD, GD, AMCR, and GMCR. These criterion are defined as

AT = arithmetic mean of the trace of $V^{(u)},\ u=1,\ldots,\nu$

GT = geometric mean of the trace of $V^{(u)},\ u=1,\ldots,\nu$

AD = arithmetic mean of the determinants of $V^{(u)},\ u=1,\ldots,\nu$

GD = geometric mean of the determinants of $V^{(u)},\ u=1,\ldots,\nu$

AMCR = arithmetic mean of the maximum characteristic root of $V^{(u)},\ u=1,\ldots,\nu$

GMCR = geometric mean of the maximum characteristic root of $V^{(u)},\ u=1,\ldots,\nu$

The design with the minimum value for one of the criterion is said to be optimal with respect to that criterion. A design may be optimal with respect to one, some, or all of the criterion. Note that when $k=0$, there is $\nu=1$ model, and thus the criterion functions AT and GT become the A-optimality criterion function, AD and GD become the D-optimality criterion function, and AMCR and GMCR become the E-optimality criterion function.

Example 3.2

Consider Design A1 and Design A2 from Example 1.1, as well as the $\nu=10$ models presented in Example 3.1. To compare these designs with respect to the six criterion functions, the values of AT, GT, AD, GD, AMCR, and GMCR are calculated and presented in Table 3.2 and displayed in Figure 3.1.

Table 3.2: Comparison of Design A1 and Design A2 using the six criterion functions

Criterion	Design A1	Design A2
AT	0.7092888	0.7435897
GT	0.7092973	0.7435897
AD	7.733777×10^{-8}	9.781275×10^{-8}
GD	7.731824×10^{-8}	9.781275×10^{-8}
AMCR	0.125	0.125
GMCR	0.125	0.125

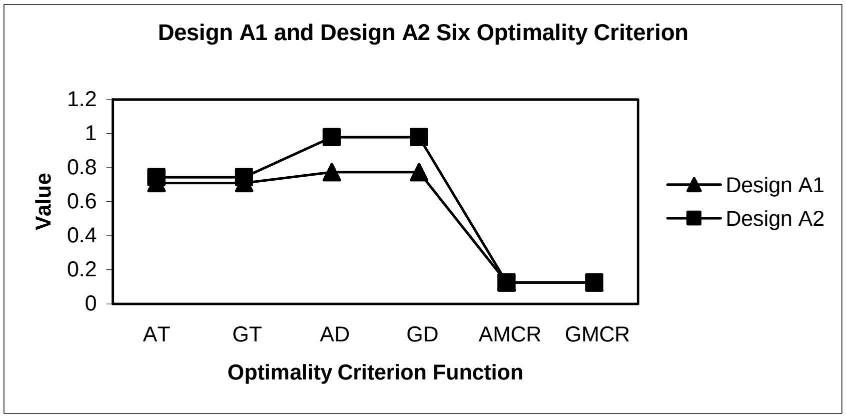

Figure 3.1: Comparison of Design A1 and Design A2 using the six criterion functions
Note: The values for AD and GD have 10^{-7} dropped

The values for AT, GT, AD, and GD for Design A1 are smaller than those for Design A2, while AMCR and GMCR for the designs are equivalent. Therefore, Design A1 is better than Design A2 with respect to AT, GT, AD, and GD but the designs are equivalent with respect to AMCR and GMCR.

3.4 Optimum Designs

Optimum designs for 2^m fractional factorial experiments with respect to some or all of the six criterion functions, obtained from the class of all designs with full estimation capacity, for various fixed values of m and n and the maximum value of k in model (3.3) are presented [Ghosh and Tian (2006)]. Recall, m represents the number of factors in a 2^m factorial experiment, n corresponds to the number of runs in the fractional factorial experiment, k is the maximum number of two-factor interaction effects included in the model, and v is the number of models. Moreover, each run is denoted by the position where the factor is observed at the high level. For example, consider the run (+, +, –, –) in a 2^4 factorial experiment. This run is denoted as 12. The run (+, –, +, +, –) in a 2^5 factorial experiment is denoted as 134. Note that 0 denotes the run where all m factors in a 2^m factorial experiment are observed at their low level.

Table 3.3: Optimum designs for *m* = 4,5; *Balanced design

m	Design	*n*	Max *k*	*v*	Runs	Optimum w.r.t.
4	D1*	5	0	1	1, 2, 3, 4, 1234	A–, D–, E–
4	D2	6	1	6	1, 2, 3, 4, 234, 1234	All six
4	D3*	7	1	6	0, 12, 13, 14, 23, 24, 34	All six
4	D4.1	8	2	15	1, 2, 3, 4, 124, 134, 234, 1234	AT, GT, AD, GD
4	D4.2	8	2	15	0, 1, 3, 12, 34, 123, 124, 1234	AMCR, GMCR
4	D5*	9	3	20	1, 2, 3, 4, 123, 124, 134, 234, 1234	All six
4	D6	10	5	6	0, 12, 13, 14, 23, 24, 34, 124, 134, 234	All six
4	D7*	11	6	1	0, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234	A–, D–, E–
5	D8	6	0	1	1, 2, 345, 1234, 1235, 1245	A–, D–, E–
5	D9.1	7	1	10	2, 5, 13, 345, 1234, 1235, 1245	AD
5	D9.2	7	1	10	0, 12, 1234, 1235, 1245, 1345, 2345	AT, AMCR
5	D9.3	7	1	10	1, 2, 3, 345, 1234, 1235, 1245	GT, GD, GMCR
5	D10	8	1	10	1, 2, 3, 4, 345, 1234, 1235, 1245	All six
5	D11	9	2	45	1, 2, 3, 4, 345, 1234, 1235, 1245, 1345	All six
5	D12*	10	3	120	1, 2, 3, 4, 5, 1234, 1235, 1245, 1345, 2345	All six
5	D13.1	11	3	120	1, 2, 3, 4, 5, 345, 1234, 1235, 1245, 1345, 2345	AT, GT, AD, GD
5	D13.2*	11	3	120	0, 1, 2, 3, 4, 5, 1234, 1235, 1245, 1345, 2345	AMCR, GMCR
5	D14	12	5	252	5, 12, 13, 14, 23, 24, 34, 1234, 1235, 1245, 1345, 2345	All six
5	D15	13	5	252	5, 12, 13, 14, 15, 23, 24, 34, 1234, 1235, 1245, 1345, 2345	All six
5	D16	14	7	120	0, 2, 4, 5, 12, 13, 14, 34, 35, 1234, 1235, 1245, 1345, 2345	All six
5	D17*	15	9	10	12, 13, 14, 15, 23, 24, 25, 34, 35, 45 1234, 1235, 1245, 1345, 2345	All six
5	D18*	16	10	1	0, 12, 13, 14, 15, 23, 24, 25, 34, 35, 45,1234, 1235, 1245, 1345, 2345	A–, D–, E–

3.5 Balanced Designs

Some designs contain runs which can be grouped according to the number of factors observed at the same level with the groups being denoted as S_i. A set of runs, S_i, is defined such that the set contains all runs with i factors observed at the low level (represented by $-$, or equivalently, -1) and $m-i$ factors observed at the high level (represented by +, or equivalently, 1). The number of runs in S_i is equal to $\binom{m}{i}$ and $S_i = -S_{m-i}$. These groups of treatments are illustrated for S_0, S_1, S_2 below.

$$S_0 = [1\;1\;1\;\ldots 1], \quad S_1 = \begin{bmatrix} -1 & 1 & 1 & \ldots & 1 \\ 1 & -1 & 1 & \ldots & 1 \\ 1 & 1 & -1 & \ldots & 1 \\ & \vdots & & & \\ 1 & 1 & 1 & \ldots & -1 \end{bmatrix}, \quad S_2 = \begin{bmatrix} -1 & -1 & 1 & \ldots & 1 & 1 \\ -1 & 1 & -1 & \ldots & 1 & 1 \\ -1 & 1 & 1 & \ldots & 1 & -1 \\ & \vdots & & & & \\ 1 & 1 & 1 & \ldots & -1 & -1 \end{bmatrix}.$$

A design which is composed of treatments from complete sets of S_i's is referred to as a balanced design. A balanced design has the property that the design will remain the same if the factors in the experiment are renamed [Srivastava and Chopra (1971)]. Note that some, but not all, of the optimum designs in Table 3.3 are balanced designs.

Example 3.3

Consider Design A1 and Design A2 from Figure 1.1 in Chapter 1. These designs are displayed again in Figure 3.2. Note that these two designs are equivalent to D13.1 and D13.2, respectively, from Table 3.3. It can be seen that D13.1 contains two complete sets of treatments, S_1 and S_4, with the final run in the design coming from S_2 (the complete set of S_2 is not present). However, D13.2 contains the complete sets of treatments for S_1, S_4, S_5. Thus, D13.2 is a balanced design.

D13.1

S_1	+	+	+	+	–
	+	+	+	–	+
	+	+	–	+	+
	+	–	+	+	+
	–	+	+	+	+
S_4	+	–	–	–	–
	–	+	–	–	–
	–	–	+	–	–
	–	–	–	+	–
	–	–	–	–	+
1 run from S_2	–	–	+	+	+

D13.2

S_1	+	+	+	+	–
	+	+	+	–	+
	+	+	–	+	+
	+	–	+	+	+
	–	+	+	+	+
S_4	+	–	–	–	–
	–	+	–	–	–
	–	–	+	–	–
	–	–	–	+	–
	–	–	–	–	+
S_5	–	–	–	–	–

Figure 3.2: Design A1 and Design A2 w.r.t. treatment sets, S_i

Chapter 4

New Property

4.0 Introduction

Because the k two-factor interaction effects are the possible distinguishing elements of the v possible models, a new property is defined which focuses on these k effects. Similar to the aforementioned optimality criterion functions, the new property considers $V^{(u)} = \left(X^{(u)\prime} X^{(u)}\right)^{-1}$, $u = 1,\ldots,v$. In particular, the last k diagonal elements of $V^{(u)}$, $u = 1,\ldots,v$, are studied. These elements correspond to the variances of the estimates for the k two-factor interaction effects. The variances represent the precision of the estimates; the smaller the variance, the higher the precision.

4.1 Property P_g^V

Definition: A design is said to have the property P_g^V where g is a positive integer, if there exist g groups of models so that the variances of k two-factor interactions for all models within a group are identical with each other.

From the above definition, we observe that P_g^V, $g \geq 1$, represents the property of a design giving common entries of $\underline{V}_2^{(u)}$, $u = 1,2,\ldots,v$ for g groups of models where $\underline{V}_2^{(u)}$ is a $(k \times 1)$ vector of the diagonal elements of $\sigma^{-2}\left(Var\left[\hat{\beta}_2^{(u)}\right]\right)$. The value of the identical entries of $\underline{V}_2^{(u)}$ is denoted as simply, $V_2^{(u)}$. This property is written as $P_g^{\nu}\left(\nu_1,...,\nu_g\right)$, $\nu_1+...+\nu_g=\nu$ where ν_i is the number of models in Group i, $i=1,...,g$ and ν=total number of models. Moreover, the common variance for the ν_1 models in Group 1 is smaller than the common variance for the ν_2 models in Group 2 and so on with Group g consisting of the ν_g models with the largest common variance.

Example 4.1

Consider Design A1 and Design A2 from Example 1.1, as well as the $\nu=10$ models presented in Example 3.1 for $k=1$. Recall that the two designs, A1 and A2, are equivalent to D13.1 and D13.2, respectively, from Table 3.3. To illustrate P_g^V, $\underline{V}_2^{(u)}$, $u=1,...,10$, is calculated for each design and presented in Table 4.1 and Figure 4.1. Note that when $k=1$, $\underline{V}_2^{(u)}$ is a scalar and denoted as $V_2^{(u)}$.

Design D13.1 has common values of $V_2^{(u)}$ for two groups with six models ($u=2-7$) giving a smaller value of $V_2^{(u)}$=0.095 than the remaining four models ($u=1,8,9,10$) with $V_2^{(u)}$=0.100. Thus, the property is expressed as $P_2^V\left(6,4\right)$. Similarly, D13.2 has a common value of $V_2^{(u)}$=0.099 for all ten models and the property is denoted as $P_1^V\left(10\right)$. Recall, the two-factor interaction corresponding to the specified values of u can be found in Table 3.1

Table 4.1: Values of $V_2^{(u)}$ for D13.1 and D13.2

	u									
Design	1	2	3	4	5	6	7	8	9	10
D13.1	0.100	0.095	0.095	0.095	0.095	0.095	0.095	0.100	0.100	0.100
D13.2	0.099	0.099	0.099	0.099	0.099	0.099	0.099	0.099	0.099	0.099

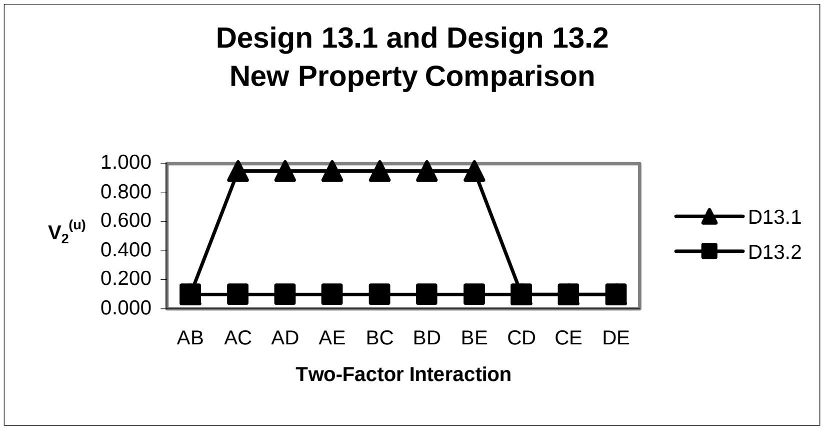

Figure 4.1: Values of $V_2^{(u)}$ for D13.1 and D13.2

It is observed from Table 4.1 that the values of $V_2^{(u)}$ for $u = 2-7$ in D13.1 are smaller than those of D13.2. This indicates the precisions of the two-factor interaction estimates for these models are higher, making D13.1 optimal with respect to precision for these particular models. For the remaining models, $u = 1,8,9,10$, D13.2 is optimal. However, at the design stage of the experiment it is unknown which model will be chosen as the best fit for the data collected.

4.2 P_g^V Classification for Optimum Designs

Although P_g^V can be used to evaluate and choose a design when the precision of the distinguishing elements of the models is of importance, it is likely that the experimenter will also be interested in the optimum criterion functions described in Section 3.3. For the optimum designs given in Table 3.3, we now consider the model in (3.3) with $k = 1$. Table 4.2 and Table 4.3 present $P_g^v(v_1,\ldots,v_g)$ for these designs for $m = 4$ and $m = 5$, respectively. Recall, the values of u are defined in Table 3.1.

Table 4.2: P_g^V Characterization for optimum designs, *m* =4, *k* = 1;*Balanced design

Design	n	v	$P_g^V(v_1, v_2, \ldots v_g)$	$V_2^{(u)}$	u
D1*	5	0	N/A	N/A	N/A
D2	6	6	$P_2^V(3,3)$	0.313 1.250	1 – 3 4 – 6
D3*	7	6	$P_1^V(6)$	0.188	All
D4.1	8	6	$P_2^V(3,3)$	0.136 0.188	1, 2, 4 3, 5, 6
D4.2	8	6	$P_2^V(2,4)$	0.125 0.167	3, 4 1, 2, 5, 6
D5*	9	6	$P_1^V(6)$	0.116	All
D6	10	6	$P_2^V(3,3)$	0.113 0.132	3, 5, 6 1, 2, 4
D7*	11	6	$P_1^V(6)$	0.109	All

Table 4.3: P_g^V Characterization for optimum designs, *m* =5, *k* = 1;*Balanced design

Design	n	v	$P_g^V\left(v_1, v_2, \ldots v_g\right)$	$V_2^{(u)}$	u
D8	6	0	N/A	N/A	N/A
D9.1	7	10	$P_2^V(2,8)$	0.188 0.750	5, 10 1 – 4, 6 – 9
D9.2	7	10	$P_1^V(10)$	0.625	All
D9.3	7	10	$P_4^V(2,4,3,1)$	0.234 0.417 0.938 3.750	8, 9 3, 4, 6, 7 1, 2, 5 10
D10	8	10	$P_4^V(1,2,6,1)$	0.137 0.199 0.243 0.547	8 9, 10 2 – 7 1
D11	9	10	$P_6^V(1,2,1,1,4,1)$	0.123 0.145 0.160 0.166 0.182 0.188	8 2, 3 4 7 5, 6, 9, 10 1
D12*	10	10	$P_1^V(10)$	0.104	All
D13.1	11	10	$P_2^V(6,4)$	0.096 0.100	2 – 7 1, 8 – 10
D13.2*	11	10	$P_1^V(10)$	0.099	All
D14	12	10	$P_1^V(10)$	0.097	All
D15	13	10	$P_2^V(4,6)$	0.085 0.090	1 – 4 5 – 10
D16	14	10	$P_4^V(2,3,3,2)$	0.077 0.082 0.083 0.090	4, 5 6, 7, 10 1, 2, 9 3, 8
D17*	15	10	$P_1^V(10)$	0.069	All
D18*	16	10	$P_1^V(10)$	0.063	All

From Table 4.2, it is seen that Design D7 has the property $P_1^V(6)$ with $V_2^{(u)}$= 0.109. This indicates that all $u = 6$ models containing the general mean, $m = 4$ main effects, and $k = 1$ two-factor interaction effect belong to a single group, i.e., $g = 1$, with the variances of the two-factor interactions being identical for all models; namely 0.109.

It is observed in Table 4.3 that Design D11 has the property $P_6^V\left(1,2,1,1,4,1\right)$. Here, the $v = 10$ models are divided into six groups, i.e., $g = 6$, with the variances of the two-factor interaction being the same in each of the groups. The group with the smallest variance, $V_2^{(u)}$=0.123, consists of a single model ($u = 8$) containing the general mean, $m = 5$ main effects, and $k = 1$ two-factor interaction effect corresponding to factors C and D. The next group, with $V_2^{(u)}$=0.145, consists of two models ($u = 2,3$) and the following group, with $V_2^{(u)}$=0.160, consists of a single model ($u = 4$). The last three groups consist of a single model ($u = 7$), four models ($u = 5,6,9,10$), and a single model ($u = 1$), respectively. Note that the last group listed has the largest variance, $V_2^{(u)}$=0.188, and consists of the single model containing the general mean, $m = 5$ main effects, and $k = 1$ two-factor interaction effect corresponding to factors A and B.

Chapter 5

P_g^V Characterization: k = 1, g = 1

5.0 Introduction

As observed in Table 4.1, it is possible for a design to be optimal with respect to the variance criterion. Design D13.1 is better than Design D13.2 for models 2 – 7. However, D13.2 is better than D13.1 for models 1 and 8-10. Because it is unknown at the design stage which model will be deemed appropriate at the end of an experiment, it is important for the distinguishing parameter between models (in this case, the k interaction effects) to be estimated with the same precision. Thus, the most important case of the property, P_g^V where $g = 1$, is examined. More specifically, the case where $k = 1$ is considered in detail. It is of prime interest to be able to identify a design with P_1^V using characterizations of the design. In this chapter, a variety of characterizations are presented to identify a design with the property P_1^V.

5.1 Optimum Designs with P_1^V for k = 1

To characterize designs with P_1^V, the optimum designs presented in Chapter 3 are regularly referenced as examples, as well as for illustration of the characterization described. Table 5.1 presents the optimum designs with P_1^V considered in this chapter where $k = 1$, along with the value of $V_2^{(u)}$ for all $u = 1,...,\nu$.

Table 5.1: Optimum designs with P_1^V for k = 1; *Balanced designs

m	Design	n	ν	Optimum w.r.t. the criterion functions	$P_g^V(\nu_1,\nu_2,...\nu_g)$	$V_2^{(u)}$
4	D3*	7	6	All six	$P_1^V(6)$	0.1875
4	D5*	9	6	All six	$P_1^V(6)$	0.1161
4	D7*	11	6	A–, D–, E–	$P_1^V(6)$	0.1094
5	D9.2	7	10	AT, AMCR	$P_1^V(10)$	0.6250
5	D12*	10	10	All six	$P_1^V(10)$	0.1042
5	D13.2*	11	10	AMCR, GMCR	$P_1^V(10)$	0.0994
5	D14	12	10	All six	$P_1^V(10)$	0.0969
5	D17*	15	10	All six	$P_1^V(10)$	0.0694
5	D18*	16	10	A–, D–, E–	$P_1^V(10)$	0.0625

5.2 Inverse Matrix Characterization

Consider $V^{(u)} = \sigma^{-2} Var\left[\hat{\beta}^{(u)}\right] = \left(X^{(u)\prime} X^{(u)}\right)^{-1}$, $u = 1,...,\nu$, which is a $(m+2)\times(m+2)$ matrix with the diagonal elements representing the variances of the parameter estimates and the off-diagonal elements representing the co-variances between the parameter estimates. Then $V_2^{(u)}$ is the $(m+2)^{th}$ diagonal element of $V^{(u)}$ and designs with P_1^V will have this value constant for all ν models.

It can be seen [Ghosh, Deng, and Luan (2007), Rao (2002)] that

$$\left(X^{(u)\prime} X^{(u)}\right)^{-1} = \begin{bmatrix} W^{-1} + W^{-1}\underline{w}_u V_2^{(u)} \underline{w}'_u W^{-1} & -W^{-1}\underline{w}_u V_2^{(u)} \\ -V_2^{(u)} \underline{w}'_u W^{-1} & V_2^{(u)} \end{bmatrix}, \quad u = 1,...,\nu \tag{5.1}$$

where

$$W = X_1{}'X_1,$$
$$\underline{w}'_u = \underline{X}_2^{(u)\prime} X_1,$$
$$w_{uu} = \underline{X}_2^{(u)\prime}\, \underline{X}_2^{(u)},$$
$$V_2^{(u)} = (w_{uu} - \underline{w}'_u W^{-1} \underline{w}_u)^{-1} \qquad u = 1, \ldots, v. \tag{5.2}$$

It follows that

$$\left(V_2^{(u)}\right)^{-1} = w_{uu} - \underline{w}'_u W^{-1} \underline{w}_u, \quad w_{uu} = n, \quad u = 1, \ldots, v. \tag{5.3}$$

Thus, the following results are true.

Theorem 1: A design has P_1^V for $k = 1$ iff (if and only if) $\underline{w}'_u W^{-1} \underline{w}_u$ is constant for all u.

Corollary 1.1: A design has P_1^V for $k = 1$ if $\underline{w}_u$ are identical for all u.

Corollary 1.2: A design has P_1^V for $k = 1$ if for any u_1 and u_2 there is a $(m+1)\times(m+1)$ matrix P_{u_1,u_2} satisfying

i) $\underline{w}_{u_1} = P_{u_1,u_2} \underline{w}_{u_2}$,

ii) $P'_{u_1,u_2} W^{-1} P_{u_1,u_2} = W^{-1}$.

Note: P_{u_1,u_2} may or may not be a permutation matrix.

Example 5.1 for *Corollary 1.1*

Consider D12 from Table 5.1 for $m = 5$ which has $P_1^V(10)$ for $k = 1$ two-factor interaction effects included in the model with the general mean and main effects. This is a balanced design composed of S_1 and S_4 for which $\underline{w}'_u W^{-1} \underline{w}_u = 0.4$ for $u = 1,...,10$ where

$$W^{-1} = \left[\begin{array}{c|ccccc} a & b & b & b & b & b \\ \hline b & d & c & c & c & c \\ b & c & d & c & c & c \\ b & c & c & d & c & c \\ b & c & c & c & d & c \\ b & c & c & c & c & d \end{array}\right] \text{ where } \begin{array}{l} a = 0.100 \\ b = 0 \\ c = -0.013 \\ d = 0.111 \end{array} \text{ and } \underline{w}_u = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \forall u.$$

For this particular design, it is seen that $\underline{w}_u$ is identical for all u resulting in constant $\underline{w}'_u W^{-1} \underline{w}_u$ for all models. Other examples of designs with this same feature are D3, D5, D13.2, D17, D18 from Table 5.1; all of which are balanced designs. However, not all balanced designs have $\underline{w}_u$ the same for all u. Moreover, it is not necessary for $\underline{w}_u$ to be identical for all u to have $\underline{w}'_u W^{-1} \underline{w}_u$ constant (as seen in the next example).

Example 5.2 for *Corollary 1.2*

Consider D7 from Table 5.1 for $m = 4$ which has $P_1^V(6)$ for $k = 1$ two-factor interaction effect included in the model with the general mean and main effects. Recall that this is a balanced design; in particular, it is composed of the complete sets S_1, S_2, and S_4. This design has $\underline{w}'_u W^{-1} \underline{w}_u = 1.8571$ for $u = 1,\ldots,6$ where

$$W^{-1} = \left[\begin{array}{c|cccc} a & -b & -b & -b & -b \\ \hline -b & a & b & b & b \\ -b & b & a & b & b \\ -b & b & b & a & b \\ -b & b & b & b & a \end{array}\right] \text{ where } \begin{array}{l} a = 0.095 \\ b = 0.011 \end{array}$$

$$\underline{w}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -3 \\ -3 \end{bmatrix}, \quad \underline{w}_2 = \begin{bmatrix} -1 \\ 1 \\ -3 \\ 1 \\ -3 \end{bmatrix}, \quad \underline{w}_3 = \begin{bmatrix} -1 \\ 1 \\ -3 \\ -3 \\ 1 \end{bmatrix}, \quad \underline{w}_4 = \begin{bmatrix} -1 \\ -3 \\ 1 \\ 1 \\ -3 \end{bmatrix}, \quad \underline{w}_5 = \begin{bmatrix} -1 \\ -3 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \quad \underline{w}_6 = \begin{bmatrix} -1 \\ -3 \\ -3 \\ 1 \\ 1 \end{bmatrix}$$

It is observed that each $\underline{w}_u$, $u = 1,\ldots,6$, are permutations of the remaining $\underline{w}_u$ vectors. Therefore $\underline{w}_{u_1} = P_{u_1,u_2} \underline{w}_{u_2}$ where P_{u_1,u_2} is an $(m+1)\times(m+1)$ permutation matrix. If

$W^{-1} = P'_{u_1,u_2} W^{-1} P_{u_1,u_2}$, then $\underline{w}'_{u_1} W^{-1} \underline{w}_{u_1} = \underline{w}'_{u_2} P'_{u_1,u_2} W^{-1} P_{u_1,u_2} \underline{w}_{u_2}$ is constant for u_1 and u_2. As an illustration, consider $u_1 = 1$ and $u_2 = 2$ where

$$P_{1,2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

It can be shown that $W^{-1} = P'_{1,2} W^{-1} P_{1,2}$ giving $\underline{w}'_1 W^{-1} \underline{w}_1 = \underline{w}'_2 P'_{1,2} W^{-1} P_{1,2} \underline{w}_2$ constant for $u = 1,2$. Similarly, a permutation matrix can be found for the remaining pairs of u_1 and u_2 resulting in $\underline{w}'_u W^{-1} \underline{w}_u$ constant for all u.

Example 5.3 for *Theorem 1*

Consider D9.2 from Table 5.1 for $m = 5$ which has $P_1^V(10)$ for $k = 1$ two-factor interaction effect included in the model with the general mean and main effects. This design is not a balanced design as it is composed of both complete and incomplete sets of treatments; in particular, S_1, S_5 and one treatment from S_2. This design has $\underline{w}'_u W^{-1} \underline{w}_u =$ 5.4 for $u = 1,...,10$ where W^{-1} and $\underline{w}_u$, $u = 1,...,10$ are calculated and presented on the following page. It is observed that $\underline{w}_u$, $u = 2,...,10$, are permutations of each other while $\underline{w}_1$ is different than the remaining $\underline{w}_u$ vectors. However, note that in all $\underline{w}_u$, the following properties are true: $(w_1 + w_2 + w_3) = -(w_4 + w_5 + w_6)$ and $(w_1^2 + w_2^2 + w_3^2) = (w_4^2 + w_5^2 + w_6^2)$ where $\underline{w}_u{}' = [w_1 \quad w_2 \quad w_3 \quad w_4 \quad w_5 \quad w_6]$.

$$W^{-1} = \left[\begin{array}{ccc|ccc} a & b & b & c & c & c \\ b & a & b & c & c & c \\ b & b & a & c & c & c \\ \hline c & c & c & a & b & b \\ c & c & c & b & a & b \\ c & c & c & b & b & a \end{array}\right] \text{ where } \begin{array}{l} a = 0.193 \\ b = -0.056 \\ c = -0.006 \end{array}$$

$$\underline{w}_1 = \begin{bmatrix} 3 \\ 3 \\ 3 \\ -3 \\ -3 \\ -3 \end{bmatrix},\ \underline{w}_2 = \begin{bmatrix} 1 \\ 1 \\ -3 \\ 3 \\ -1 \\ -1 \end{bmatrix},\ \underline{w}_3 = \begin{bmatrix} 1 \\ 1 \\ -3 \\ -1 \\ 3 \\ -1 \end{bmatrix},\ \underline{w}_4 = \begin{bmatrix} 1 \\ 1 \\ -3 \\ -1 \\ -1 \\ 3 \end{bmatrix},\ \underline{w}_5 = \begin{bmatrix} 1 \\ -3 \\ 1 \\ 3 \\ -1 \\ -1 \end{bmatrix},$$

$$\underline{w}_6 = \begin{bmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 3 \\ -1 \end{bmatrix},\ \underline{w}_7 = \begin{bmatrix} 1 \\ -3 \\ 1 \\ -1 \\ -1 \\ 3 \end{bmatrix},\ \underline{w}_8 = \begin{bmatrix} 3 \\ -1 \\ -1 \\ 1 \\ 1 \\ -3 \end{bmatrix},\ \underline{w}_9 = \begin{bmatrix} 3 \\ -1 \\ -1 \\ 1 \\ -3 \\ 1 \end{bmatrix},\ \underline{w}_{10} = \begin{bmatrix} 3 \\ -1 \\ -1 \\ -3 \\ 1 \\ 1 \end{bmatrix}.$$

As a result of the common form of $\underline{w}_u$ for all u, it is seen that $\underline{w}_u' W^{-1} \underline{w}_u$ is calculated as

$$\begin{aligned} \underline{w}_u' W^{-1} \underline{w}_u &= (a-b)\left(\left(w_1^2 + w_2^2 + w_3^2\right) + \left(w_4^2 + w_5^2 + w_6^2\right)\right) \\ &\quad + b\left(\left(w_1 + w_2 + w_3\right)^2 + \left(w_4 + w_5 + w_6\right)^2\right) \\ &\quad + 2c\left(w_1 + w_2 + w_3\right)\left(w_4 + w_5 + w_6\right) \\ &= 2(a-b)\left(w_1^2 + w_2^2 + w_3^2\right) + 2(b-c)\left(w_1 + w_2 + w_3\right)^2, \quad u = 1,\ldots,10 \end{aligned}$$

where a,b,c are the specified values from W^{-1}. For $\underline{w}_u$, $u = 2,\ldots,10$, $\left(w_1^2 + w_2^2 + w_3^2\right)$ and $\left(w_1 + w_2 + w_3\right)^2$ are the same and therefore $\underline{w}_u' W^{-1} \underline{w}_u$ is the same. Now consider $\underline{w}_1$ as it compares to $\underline{w}_u$, $u = 2,\ldots,10$. It can be shown that $a + 4b - 5c = 0$ must be satisfied for $\underline{w}_u' W^{-1} \underline{w}_u$ to be constant for all u. For design D9.2, this is indeed the case. Thus, it is seen that $\underline{w}_u$ need not be constant for all u, nor must the $\underline{w}_u$'s be permutations of one another for $\underline{w}_u{}' W^{-1} \underline{w}_u$ to be constant for all u.

Example 5.4 for *Theorem 1*

Consider D14 from Table 5.1 for $m = 5$ which has $P_1^V(10)$ for $k = 1$ two-factor interaction effect included in the model with the general mean and main effects. This design is an unbalanced design, as it is composed of both complete and incomplete sets of treatments; in particular, S_1, six of ten treatments from S_2 and one of five treatments from S_4. This design has $\underline{w}'_u W^{-1} \underline{w}_u =$ 1.6774 for $u = 1,...,10$ where W^{-1} and $\underline{w}_u$, $u = 1,...,10$ are calculated and presented below. It is observed that $\underline{w}_u$, $u = 1,2,3,5,6,8$, are permutations of each other while $\underline{w}_u$, $u = 4,7,9,10$, are permutations of each other but are different than $\underline{w}_u$, $u = 1,2,3,5,6,8$.

$$W^{-1} = \left[\begin{array}{c|ccccc} a & b & b & b & b & -b \\ \hline b & d & c & c & c & -c \\ b & c & d & c & c & -c \\ b & c & c & d & c & -c \\ b & c & c & c & d & -c \\ -b & -c & -c & -c & -c & d \end{array}\right] \quad \text{where} \quad \begin{array}{l} a = 0.096 \\ b = -0.016 \\ c = 0.002 \\ d = 0.086 \end{array}$$

$$\underline{w}_1 = \begin{bmatrix} 0 \\ 2 \\ 2 \\ -2 \\ -2 \\ 2 \end{bmatrix}, \ \underline{w}_2 = \begin{bmatrix} 0 \\ 2 \\ -2 \\ 2 \\ -2 \\ 2 \end{bmatrix}, \ \underline{w}_3 = \begin{bmatrix} 0 \\ 2 \\ -2 \\ -2 \\ 2 \\ 2 \end{bmatrix}, \ \underline{w}_4 = \begin{bmatrix} 0 \\ -2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}, \ \underline{w}_5 = \begin{bmatrix} 0 \\ -2 \\ 2 \\ 2 \\ -2 \\ 2 \end{bmatrix},$$

$$\underline{w}_6 = \begin{bmatrix} 0 \\ -2 \\ 2 \\ -2 \\ 2 \\ 2 \end{bmatrix}, \ \underline{w}_7 = \begin{bmatrix} 0 \\ 2 \\ -2 \\ 2 \\ 2 \\ 2 \end{bmatrix}, \ \underline{w}_8 = \begin{bmatrix} 0 \\ -2 \\ -2 \\ 2 \\ 2 \\ 2 \end{bmatrix}, \ \underline{w}_9 = \begin{bmatrix} 0 \\ 2 \\ 2 \\ -2 \\ 2 \\ 2 \end{bmatrix}, \ \underline{w}_{10} = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 2 \\ -2 \\ 2 \end{bmatrix}$$

It can be seen that $\underline{w}_u{}' W^{-1} \underline{w}_u$ for all u is calculated as

$$\underline{w}_u{}' W^{-1} \underline{w}_u = (d-c)\left(w_2^2 + w_3^2 + w_4^2 + w_5^2 + w_6^2\right) + c\left(w_2 + w_3 + w_4 + w_5 - w_6\right)^2$$

where $\underline{w}_u{}' = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 & w_5 & w_6 \end{bmatrix}$ and c, d are the specified values from W^{-1}. Note that $\left(w_2^2 + w_3^2 + w_4^2 + w_5^2 + w_6^2\right)$ is constant for all $\underline{w}_u$ since $w_i = \pm 2,\ i = 2,\ldots,6$. The second summation in $\underline{w}_u{}' W^{-1} \underline{w}_u$ can be broken down into two cases as presented in Table 5.2 where it is seen $\left(w_2 + w_3 + w_4 + w_5 - w_6\right) = \pm 2$ for all u. Therefore, $\underline{w}_u{}' W^{-1} \underline{w}_u$ is constant for all u for D14.

Table 5.2: Calculated values of w_u vector entries for D14

u	$w_2 + w_3 + w_4 + w_5$	$-w_6$	$w_2 + w_3 + w_4 + w_5 - w_6$	$\left(w_2 + w_3 + w_4 + w_5 - w_6\right)^2$
1,2,3,5,6,8	0	–2	–2	4
4,7,9,10	4	–2	2	4

5.3 Characterization by Determinants

Consider $V^{(u)} = \sigma^{-2} Var\left[\hat{\beta}^{(u)}\right] = \left(X^{(u)\prime} X^{(u)}\right)^{-1}$, $u = 1,...,v$. It can be seen that

$$\left(X^{(u)\prime} X^{(u)}\right)^{-1} = \begin{bmatrix} X_1{}'X_1 & X_1{}'X_2^{(u)} \\ X_2^{(u)\prime}X_1 & X_2^{(u)\prime}X_2^{(u)} \end{bmatrix}^{-1} = \left[\begin{array}{c|c} - - - & - - - \\ & W^{(u)} \end{array}\right] \tag{5.4}$$

where $W^{(u)} = \sigma^{-2}\left(Var\left[\hat{\beta}_2^{(u)}\right]\right)$. Thus, $\underline{V}_2^{(u)}$, the variances of the estimates for the k two-factor interactions, are represented by the diagonal elements of $W^{(u)}$. Specifically, the diagonal elements of $W^{(u)}$ are denoted as $w^{(u,j)}$ and can be calculated as

$$w^{(u,j)} = \frac{\left|X^{(u)\prime} X_{(-j)}^{(u)}\right|}{\left|X^{(u)\prime} X^{(u)}\right|} \tag{5.5}$$

where $X^{(u)\prime} X_{(-j)}^{(u)}$ denotes the $X^{(u)\prime} X^{(u)}$ matrix with the j^{th} row and column deleted, $j = (m+2),(m+3),...,(m+k+1)$. Then for designs with P_g^V, $w^{(u,j)}$ is the same for g groups of models for all j.

Consider the specific case of this chapter; namely, P_1^V for $k = 1$. Then, from equation (5.5), the calculation for the variance of the two-factor interaction effect, $V_2^{(u)}$, becomes

$$W^{(u)} = V_2^{(u)} = \frac{\left|X^{(u)\prime} X_{(-m+2)}^{(u)}\right|}{\left|X^{(u)\prime} X^{(u)}\right|} = \frac{\left|X_1{}'X_1\right|}{\left|X^{(u)\prime} X^{(u)}\right|} \tag{5.6}$$

where $V_2^{(u)}$ is constant for all $u = 1,...,v$ models. Thus, the following result holds for designs with P_1^V,

Theorem 2: A design will have P_1^V for $k = 1$ iff $\left|X^{(u)\prime}X^{(u)}\right|$ is constant for all u.

Moreover, it is possible to recognize designs with P_1^V from the optimality criterion, AD and GD, from Section 3.3 which are calculated using the determinants of the inverse of the $X^{(u)\prime}X^{(u)}$ matrix for all u. AD is the defined as the arithmetic mean of the determinants of $V^{(u)}$, $u = 1,...,v$, or in other words,

$$\text{AD} = \frac{1}{v}\sum_{u=1}^{v}\left|\left(X^{(u)\prime}X^{(u)}\right)^{-1}\right| \qquad (5.7)$$

and GD is defined as the geometric mean of the determinants of $V^{(u)}$, $u = 1,...,v$, or in other words,

$$\text{GD} = \left(\prod_{u=1}^{v}\left|\left(X^{(u)\prime}X^{(u)}\right)^{-1}\right|\right)^{1/v}. \qquad (5.8)$$

From ***Theorem 2***, designs with P_1^V have $\left|X^{(u)\prime}X^{(u)}\right|$ constant for all u, implying $\left|\left(X^{(u)\prime}X^{(u)}\right)^{-1}\right|$ is also constant for all u since $\left|\left(X^{(u)\prime}X^{(u)}\right)^{-1}\right| = \left|X^{(u)\prime}X^{(u)}\right|^{-1}$. When $\left|\left(X^{(u)\prime}X^{(u)}\right)^{-1}\right|$ is constant for all v models, AD = GD by the Inequality of arithmetic and geometric means (AM-GM Inequality). AM-GM Inequality states that the arithmetic mean of a list of non-negative real numbers is greater than or equal to the geometric mean of the same list; and further, that the two means are equal if and only if every number in the list is the same. Thus, the following result holds for designs with P_1^V.

Corollary 2.1: A design will have P_1^V for $k=1$ iff AD = GD.

Example 5.5

Consider the designs from Table 5.1 for $m=4,5$ which have P_1^V for $k=1$ two-factor interaction effect included in the model with the general mean and main effects. The determinants for calculating $V_2^{(u)}$ for these designs are presented in Table 5.3, along with the calculated values of AD and GD. It is seen that $\left|X^{(u)\prime}X^{(u)}\right|$ is constant for all $u=1,\ldots,v$ models for each of the designs. Additionally, the values for the optimality criterion functions AD and GD are equivalent.

Table 5.3: Determinants for optimum designs with P_1^V for k = 1; *Balanced designs

m	Design	n	v	$V_2^{(u)}$	$\left\|X_1'X_1\right\|$	$\left\|X^{(u)\prime}X^{(u)}\right\|$	AD = GD
4	D3*	7	6	0.1875	12,288	65,536	1.5258×10^{-5}
4	D5*	9	6	0.1161	53,248	458,752	2.1798×10^{-6}
4	D7*	11	6	0.1094	145,152	1,327,104	7.5352×10^{-7}
5	D9.2	7	10	0.6250	40,960	65,536	1.5258×10^{-5}
5	D12*	10	10	0.1042	737,280	7,077,888	1.4128×10^{-7}
5	D13.2*	11	10	0.0994	1,015,808	10,223,616	9.7812×10^{-8}
5	D14	12	10	0.0969	2,571,264	26,542,080	3.7676×10^{-8}
5	D17*	15	10	0.0694	10,485,760	150,994,944	6.6227×10^{-9}
5	D18*	16	10	0.0625	16,777,216	268,435,456	3.7252×10^{-9}

5.4 Eigenvalue Characterization

Consider $\left|X^{(u)\prime}X^{(u)}\right|$ for $u=1,\ldots,v$ models consisting of the general mean, main effects, and $k=1$ two-factor interaction effect. From ***Theorem 2***, if a design has $\left|X^{(u)\prime}X^{(u)}\right|$ constant for all u, then the design has P_1^V. It is known that

$$\left|X^{(u)\prime}X^{(u)}\right|=\prod_{i=1}^{m+k+1}\lambda_i^{(u)},\quad u=1,\ldots,v \tag{5.9}$$

where $\lambda_i^{(u)}$, $i=1,\ldots,(m+k+1)$, are the eigenvalues of $X^{(u)\prime}X^{(u)}$, $u=1,\ldots,v$. Thus, designs with P_1^V can be characterized in the following manner,

Corollary 2.2 A design will have P_1^V for $k=1$ if the set of eigenvalues of $X^{(u)\prime}X^{(u)}$ are the same for all u.

This result is illustrated in Example 5.6. Note that it is not necessary for a design with P_1^V to have the same set of eigenvalues for all $X^{(u)\prime}X^{(u)}$ as illustrated in Example 5.7.

Example 5.6

Consider D14 from Table 5.1 for $m=5$ which has $P_1^V(10)$ for $k=1$ two-factor interaction effect included in the model with the general mean and main effects. The eigenvalues are calculated, along with their product which is equal to $\left|X^{(u)\prime}X^{(u)}\right|$, $u=1,\ldots,v$, by equation (5.9) and presented in Table 5.4.

Table 5.4: Eigenvalues for D14, *k* = 1

u	$X^{(u)\prime}X^{(u)}$ Eigenvalues						$\left\|X^{(u)\prime}X^{(u)}\right\|$
1-10	16.8989	16	12	12	12	8	7.1010	26,542,080

It is observed that the set of eigenvalues of $X^{(u)\prime}X^{(u)}$ are the same for all ten models. Thus, by ***Corollary 2.2***, D14 has P_1^V.

Example 5.7

Consider D9.2 from Table 5.1 for $m=5$ which has $P_1^V(10)$ for $k=1$ two-factor interaction effect included in the model with the general mean and main effects. The eigenvalues are calculated, along with their product which is equal to $\left|X^{(u)\prime}X^{(u)}\right|$, $u=1,...,v$, and presented in Table 5.5.

Table 5.5: Eigenvalues for D9.2, *k* = 1

u	$X^{(u)\prime}X^{(u)}$ Eigenvalues							$\left\|X^{(u)\prime}X^{(u)}\right\|$
1	16	16	4	4	4	4	1	65,536
2-10	16	10.8706	9.5102	4	4	4	0.6190	65,536

Although the set of eigenvalues of $X^{(u)\prime}X^{(u)}$ differ for $u=1$ compared to the set of eigenvalues for $u=2,...,10$, the value of $\left|X^{(u)\prime}X^{(u)}\right|$ is the same for all u. Thus, by ***Theorem 2***, D9.2 has P_1^V.

5.5 Complementary Designs

Consider a design D which consists of n treatments created by a combination of m factors observed at their high level or low level, represented by 1 and –1, respectively.

The complement of design D, denoted D^C, is defined as $\mathrm{D}^C = -\mathrm{D}$. Therefore, each of the factors observed at the low level in design D will be observed at the high level in design D^C and each of the factors observed at the high level in design D will be observed at the low level in design D^C. For designs which have the ability to identify and discriminate the models consisting of the general mean, main effects, and k two-factor interaction effects, the following result holds.

Theorem 3: For a general k, the complement of a design D, denoted D^C, has P_1^V iff design D has P_1^V.

Proof: This result can be derived from ***Theorem 2*** which states that a design will have P_1^V iff $\left|X^{(u)\prime}X^{(u)}\right|$ is constant for all u when $k=1$. Note that $X^{(u)} = \left[\underline{j} \quad \mathrm{D} \quad X_2^{(u)}\right]$ where $\underline{j}$ is a vector of length n with all entries equal to 1, D is the $n\times(m+k)$ design matrix with entries equal to 1 or –1, and $X_2^{(u)}$ is a $n\times k$ matrix which is dependent on D. Consider such a design D and WLOG let $X^{(u)} = \left[\underline{j} \quad X_2^{(u)} \quad \mathrm{D}\right]$. Then $X^{(u)\prime}X^{(u)}$ can be calculated as

$$X^{(u)\prime}X^{(u)} = \left[\begin{array}{cc|c} \underline{j}'\underline{j} & \underline{j}'X_2^{(u)} & \underline{j}'\mathrm{D} \\ X_2^{(u)\prime}\underline{j} & X_2^{(u)\prime}X_2^{(u)} & X_2^{(u)\prime}\mathrm{D} \\ \hline \mathrm{D}'\underline{j} & \mathrm{D}'X_2^{(u)} & \mathrm{D}'\mathrm{D} \end{array}\right] = \begin{bmatrix} A & B \\ B' & C \end{bmatrix} \tag{5.10}$$

where A is a $(k+1)\times(k+1)$ matrix, B is a $(k+1)\times m$ matrix, and C is a $m\times m$ matrix.

Similarly, for the complement of design D, D^C, $\hat{X}^{(u)} = \begin{bmatrix} \underline{j} & X_2^{(u)} & -D \end{bmatrix}$ where $X_2^{(u)}$ is the same $n \times k$ matrix in $X^{(u)}$ for design D (since we are considering two-factor interactions only, this matrix is not affected by taking the complement of the design) and its corresponding $\hat{X}^{(u)\prime}\hat{X}^{(u)}$ can be calculated as

$$\hat{X}^{(u)\prime}\hat{X}^{(u)} = \left[\begin{array}{cc|c} \underline{j}'\underline{j} & \underline{j}'X_2^{(u)} & -\underline{j}'D \\ X_2^{(u)\prime}\underline{j} & X_2^{(u)\prime}X_2^{(u)} & -X_2^{(u)\prime}D \\ \hline -D'\underline{j} & -D'X_2^{(u)} & D'D \end{array}\right] = \begin{bmatrix} A & -B \\ -B' & C \end{bmatrix} \tag{5.11}$$

where $A, B,$ and C are the same matrices as those in equation (5.10). Then from

$$\left|\begin{bmatrix} A & B \\ B' & C \end{bmatrix}\right| = |A|\left|C - B'A^{-1}B\right| = \left|\begin{bmatrix} A & -B \\ -B' & C \end{bmatrix}\right|, \tag{5.12}$$

design D and its complement D^C have $\left|X^{(u)\prime}X^{(u)}\right|$ the same. Since D was assumed to have P_1^V giving $\left|X^{(u)\prime}X^{(u)}\right|$ constant for all u, its complement D^C also has $\left|X^{(u)\prime}X^{(u)}\right|$ constant for all u, and thus has P_1^V. This completes the proof.

Example 5.8

It is known that D13.2 for $m = 5$ has P_1^V when estimating the ten models consisting of the general mean, main effects, and $k = 1$ two-factor interaction effect. Consider this design, along with its complement $D^C13.2$ presented in Figure 5.1. Table 5.6 displays the values of $\left|X^{(u)\prime}X^{(u)}\right|$ for each of the designs. It is observed that $\left|X^{(u)\prime}X^{(u)}\right|$ is not only constant for all u in each design, but it is also the same for D13.2

and $\mathrm{D}^{C}13.2$. Moreover, the value of the variance of the two-factor interaction effect, $V_2^{(u)}$, is the same for both designs. It is concluded that $\mathrm{D}^{C}13.2$ has P_1^{V}, just as D13.2 does.

D13.2

+	+	+	+	–
+	+	+	–	+
+	+	–	+	+
+	–	+	+	+
–	+	+	+	+
+	–	–	–	–
–	+	–	–	–
–	–	+	–	–
–	–	–	+	–
–	–	–	–	+
–	–	–	–	–

$\mathrm{D}^{C}13.2$

–	–	–	–	+
–	–	–	+	–
–	–	+	–	–
–	+	–	–	–
+	–	–	–	–
–	+	+	+	+
+	–	+	+	+
+	+	–	+	+
+	+	+	–	+
+	+	+	+	–
+	+	+	+	+

Figure 5.1: D13.2 and D^{C}13.2

Table 5.6: Determinant and $V_2^{(u)}$ for D13.2 and D^{C}13.2

Design	$\left\lvert X^{(u)\prime} X^{(u)} \right\rvert$	$V_2^{(u)}$
D13.2	10,223,616	0.0994
$\mathrm{D}^{C}13.2$	10,223,616	0.0994

5.6 Isomorphic Designs

Definition: Two fractional factorial designs are called isomorphic to each other if one can be obtained from the other by renaming of factors (column permutations) or by changing the treatment orders (row permutations).

Consider two isomorphic fractional factorial designs, D_1 and D_2, for m factors which consist of n treatments each. Letting $X_{\mathrm{D}_1}^{(u)} = \left[\underline{j} \quad \mathrm{D}_1 \quad X_2^{(u)}\right]$ and $X_{\mathrm{D}_2}^{(u)} = \left[\underline{j} \quad \mathrm{D}_2 \quad X_2^{(u)}\right]$

as seen in Section 5.5, it is possible to write $X^{(u)}_{\mathrm{D}_1} = P_r X^{(u)}_{\mathrm{D}_2} P_c$ where P_r is an $n \times n$ permutation matrix which will permute the rows of $X^{(u)}_{\mathrm{D}_2}$ and P_c is a $(m+k+1)\times(m+k+1)$ permutation matrix which permutes the columns of $X^{(u)}_{\mathrm{D}_2}$. Then

$$\begin{aligned} X^{(u)\prime}_{\mathrm{D}_1} X^{(u)}_{\mathrm{D}_1} &= \left(P_r X^{(u)}_{\mathrm{D}_2} P_c\right)' \left(P_r X^{(u)}_{\mathrm{D}_2} P_c\right) \\ &= P_c' X^{(u)\prime}_{\mathrm{D}_2} P_r' P_r X^{(u)}_{\mathrm{D}_2} P_c \\ &= P_c' X^{(u)\prime}_{\mathrm{D}_2} X^{(u)}_{\mathrm{D}_2} P_c . \end{aligned} \quad (5.13)$$

Suppose D$_1$ is known to have the property P_1^V. By ***Theorem 2***, $\left|X^{(u)\prime}_{\mathrm{D}_1} X^{(u)}_{\mathrm{D}_1}\right|$ is constant for all u and from equation (5.13) it follows that

$$\left|X^{(u)\prime}_{\mathrm{D}_1} X^{(u)}_{\mathrm{D}_1}\right| = \left|P_c' X^{(u)\prime}_{\mathrm{D}_2} X^{(u)}_{\mathrm{D}_2} P_c\right| = \left|P_c'\right| \left|X^{(u)\prime}_{\mathrm{D}_2} X^{(u)}_{\mathrm{D}_2}\right| \left|P_c\right| = \left|X^{(u)\prime}_{\mathrm{D}_2} X^{(u)}_{\mathrm{D}_2}\right| \quad (5.14)$$

where $|P_c| = \pm 1$ and $|P_c| = |P_c'|$. Thus, the following result is true.

Theorem 4: If two designs, D_1 and D_2, are isomorphic, then D_2 will have P_1^V iff D_1 has P_1^V.

Example 5.9

Consider D9.2 consisting of $n = 7$ treatments for $m = 5$ factors, as well as design DA5 as presented in Figure 5.2.

D9.2

–	–	–	–	–
+	+	–	–	–
+	+	+	+	–
+	+	+	–	+
+	+	–	+	+
+	–	+	+	+
–	+	+	+	+

Design A5

–	–	–	–	–
–	+	+	+	+
+	–	+	+	+
+	+	–	+	+
+	+	+	–	+
+	+	+	+	–
–	–	+	–	+

Figure 5.2: D9.2 and Design A5

These designs are capable of identifying and discriminating the $v = 10$ models which contain the general mean, main effects and $k = 1$ two-factor interaction effect. It can be shown that D9.2 and Design A5 are isomorphic designs with

$$P_r = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } P_c = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

such that $X^{(u)}_{\text{D9.2}} = P_r X^{(u)}_{\text{DA5}} P_c$ for all u. It is known that D9.2 has $P_1^V(10)$ and thus by ***Theorem 4***, Design A5 also has $P_1^V(10)$. Table 5.7 displays the values of $\left| X^{(u)\prime} X^{(u)} \right|$ for each of the designs. Not only is $\left| X^{(u)\prime} X^{(u)} \right|$ constant for all u in each design, but it is also the same for D9.2 and DA5. Also, the value of the variance of the two-factor interaction effect, $V_2^{(u)}$, is the same for both designs.

Table 5.7: Determinant and $V_2^{(u)}$ for D9.2 and Design A5

Design	$\left\| X^{(u)\prime} X^{(u)} \right\|$	$V_2^{(u)}$
D9.2	65,536	0.6250
DA5	65,536	0.6250

5.7 Balanced Designs

Recall from Section 3.5, a design which consists of complete sets of treatments, S_i, $i = 0,1,...,m$, is referred to as a balanced design. Denote such a design as D^{B}, where each set S_i, $i = 0,1,...,m$, is replicated r_i, $i = 0,1,...,m$, times with $r_i \geq 0$ and an integer such that $n \geq 1 + m + k$. It can be shown that $X^{(u)\prime} X^{(u)}$ consists of five distinct values, $n, \varphi_1, \varphi_2, \varphi_3, \varphi_4$ which have the following positions

$$
\begin{array}{ll}
n = \underline{1}'\underline{1} = \underline{x}_i'\underline{x}_i = \underline{x}_{ij}'\underline{x}_{ij} & \\
\varphi_1 = \underline{1}'\underline{x}_i = \underline{x}_i'\underline{x}_{ij} = \underline{x}_j'\underline{x}_{ij} & \\
\varphi_2 = \underline{1}'\underline{x}_{ij} = \underline{x}_i'\underline{x}_{i'} = \underline{x}'_{ij}\underline{x}_{ij'} = \underline{x}'_{ij}\underline{x}_{ji'} & i \neq i',\ j \neq j' \\
\varphi_3 = \underline{x}_i'\underline{x}_{i'j} & i \neq i', j \\
\varphi_4 = \underline{x}'_{ij}\underline{x}_{i'j'} & i \neq i' \neq j \neq j'
\end{array}
$$

where $X^{(u)} = \begin{bmatrix} \underline{1} & X_1 & X_2^{(u)} \end{bmatrix}$, $\underline{1}$ is a vector of length n with all entries equal to 1, $X_1 = \begin{bmatrix} \underline{x}_1 & \underline{x}_2 & \cdots & \underline{x}_m \end{bmatrix}$ is the $n \times m$ design matrix, and $X_2^{(u)}$ is an $n \times k$ matrix which depends on X_1.

For example, consider a fractional factorial design with m factors which has the ability to identify and discriminate ν models which contain the general mean, main effects, and $k = 1$ two-factor interaction effect. Specifically, consider the two-factor interaction, AB. Then $X^{(1)\prime} X^{(1)}$ becomes

$$
X^{(1)\prime} X^{(1)} =
\begin{array}{c}
 \\ \mu \\ A \\ B \\ C \\ D \\ \vdots \\ \\ AB
\end{array}
\begin{array}{c}
\begin{array}{cccccccc} \mu & A & B & C & D & \cdots & & AB \end{array} \\
\begin{bmatrix}
n & \varphi_1 & \varphi_1 & \varphi_1 & \varphi_1 & \cdots & \varphi_1 & \varphi_2 \\
\varphi_1 & n & \varphi_2 & \varphi_2 & \varphi_2 & \cdots & \varphi_2 & \varphi_1 \\
\varphi_1 & \varphi_2 & n & \varphi_2 & \varphi_2 & \cdots & \varphi_2 & \varphi_1 \\
\varphi_1 & \varphi_2 & \varphi_2 & n & \varphi_2 & \cdots & \varphi_2 & \varphi_3 \\
\varphi_1 & \varphi_2 & \varphi_2 & \varphi_2 & n & & \varphi_2 & \varphi_3 \\
\vdots & \vdots & \vdots & \vdots & & \ddots & \varphi_2 & \vdots \\
\varphi_1 & \varphi_2 & \varphi_2 & \varphi_2 & \varphi_2 & \varphi_2 & n & \varphi_3 \\
\varphi_2 & \varphi_1 & \varphi_1 & \varphi_3 & \varphi_3 & \cdots & \varphi_3 & n
\end{bmatrix}
\end{array},
$$

where

$$n = \sum_{i=0}^{m} r_i \binom{m}{i}$$

$$\varphi_1 = n - 2\sum_{i=1}^{m} r_i \binom{m-1}{i-1}$$

$$\varphi_2 = n - 4\sum_{i=1}^{m-1} r_i \binom{m-2}{i-1}$$

$$\varphi_3 = n - 2\left[\sum_{i=3}^{m} r_i \binom{m-3}{i-3} + 3\sum_{i=1}^{m-2} r_i \binom{m-3}{i-1}\right]$$

$$\varphi_4 = n - 8\left[\sum_{i=3}^{m-1} r_i \binom{m-4}{i-3} + \sum_{i=1}^{m-3} r_i \binom{m-4}{i-1}\right].$$

Note that each $X^{(u)}$, $u = 1,...,\nu$, differ only with respect to the two-factor interaction effect and since all sets of treatments, S_i, are complete, it is possible to permute the rows and columns of $X^{(u)}$ such that $X^{(u)}$ will be the same for all models. This property of balanced designs indicates that the variance-covariance matrix of the estimates, $V^{(u)} = \sigma^{-2} Var\left[\hat{\beta}^{(u)}\right] = \left(X^{(u)\prime} X^{(u)}\right)^{-1}$, is invariant under permutation of the factor names [Srivastava and Chopra (1971)]. As a consequence of this, the following holds

Theorem 5: A balanced design, denoted D^B, will have P_1^V for $k = 1$.

To illustrate, Table 5.8, Table 5.9, and Table 5.10 present the balanced designs for $m = 3, 4, 5$, respectively, which have the ability to identify and discriminate the models of interest. Along with identifying the sets of treatments used in each design, the tables present the number of treatments, n, as well as the value of the constant $V_2^{(u)}$ (to four decimal places). Here, each set of treatments, S_i, $i = 1,...,m$, is replicated at most one time, i.e., $r_i \leq 1$, $i = 1,...,m$. Note that the complements for each of these designs will also have P_1^V by ***Theorem 3*** and thus are not included in the tables.

Table 5.8: Balanced designs with P_1^V for m=3 when k=1

n	S_0	S_1	S_2	S_3	$V_2^{(u)}$
5	•	•		•	0.5000
6		•	•		0.1875
7	•	•	•		0.1667
8	•	•	•	•	0.1250

Table 5.9: Balanced designs with P_1^V for m=4 when k=1

n	S_0	S_1	S_2	S_3	S_4	$V_2^{(u)}$
6	•	•			•	0.8750
7	•		•			0.1875
8		•		•		0.1250
8	•		•		•	0.1250
9	•	•		•		0.1160
10		•	•			0.1250
10	•	•		•	•	0.1041
11	•	•	•			0.1219
11		•	•		•	0.1094
12	•	•	•		•	0.0938
14		•	•	•		0.0729
15	•	•	•	•		0.0687
16	•	•	•	•	•	0.0625

Table 5.10: Balanced Designs with P_1^V for m=5 when k=1

n	S_0	S_1	S_2	S_3	S_4	S_5	$V_2^{(u)}$
7	•	•				•	1.3750
7	•		•			•	1.3750
10		•			•		0.1041
11	•		•				0.1250
11	•			•			0.1250
11	•	•			•		0.0993
12	•	•			•	•	0.0937
15		•	•				0.1041
15		•		•			0.0694
16	•	•	•				0.1033
16	•	•		•			0.0683
16	•		•		•		0.0625
16	•			•	•		0.0917
17	•	•	•			•	0.0824
17	•	•		•		•	0.0598
20			•	•			0.0520
20		•	•		•		0.0526
21	•		•	•			0.0504
21	•	•	•		•		0.0501
21	•	•		•	•		0.0511
22	•		•	•		•	0.0458
22	•	•		•	•	•	0.0481
25		•	•	•			0.0425
26	•	•	•	•			0.0418
26	•		•	•	•		0.0398
27	•		•	•	•	•	0.0382
30		•	•	•	•		0.0335
31	•	•	•	•	•		0.0325
32	•	•	•	•	•	•	0.0313

Example 5.10

Consider the balanced design for general m consisting of S_0, S_1, S_{m-1}, S_m which has $n = 1 + \binom{m}{1} + \binom{m}{m-1} + 1 = 2(m+1)$ runs. It is of interest to identify and discriminate the $v = \binom{m}{2}$ models which contain the general mean, main effects, and $k = 1$ two-factor interaction effect. To show that this design has estimation capabilities for general $m \geq 3$, it must be proven that $X^{(u)\prime} X^{(u)}$ is invertible; or similarly that $\det\left(X^{(u)\prime} X^{(u)}\right) = \left|X^{(u)\prime} X^{(u)}\right|$ is non-zero. Letting $X^{(u)} = \left[X_1 \quad \underline{1} \quad \underline{X}_2^{(u)}\right] = \left[X_1 \quad Y^{(u)}\right]$, it can be shown for all u,

$$X^{(u)\prime} X^{(u)} = \begin{bmatrix} X_1{}' X_1 & X_1{}' Y^{(u)} \\ Y^{(u)\prime} X_1 & Y^{(u)\prime} Y^{(u)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} n & & & \\ & n & (2m\text{-}6) & \\ & & \ddots & \\ (2m\text{-}6) & & n & \\ & & & n \end{bmatrix} & O \\ O' & \begin{bmatrix} n & (2m\text{-}6) \\ (2m\text{-}6) & n \end{bmatrix} \end{bmatrix} \tag{5.15}$$

where $X_1{}' X_1$ is an $m \times m$ matrix, O is an $m \times 2$ zero matrix and $Y^{(u)\prime} Y^{(u)}$ is a 2×2 matrix. Then from equation (5.15), the determinant of $X^{(u)\prime} X^{(u)}$ is calculated as

$$\left|X^{(u)\prime} X^{(u)}\right| = \left[n^2 - (2m-6)^2\right]\left[n + m(2m-6) - (2m-6)\right]\left[n - (2m-6)\right]^{m-1} > 0. \tag{5.16}$$

Therefore, $X^{(u)\prime}X^{(u)}$ is invertible and it is concluded for general $m \geq 3$ that the design consisting of treatments from S_0, S_1, S_{m-1}, S_m is capable of estimating the $v = \binom{m}{2}$ models of interest.

Additionally, the estimates of the general mean, main effects, and $k = 1$ two-factor interaction for the u^{th} model, found using the method of least squares as presented in Section 1.2, are given as

$$\hat{\beta}_0 = \frac{1}{n^2 - (2m-6)^2}\left[n\sum_{i=1}^{n} y - (2m-6)\underline{X}_2^{(u)\prime}\underline{y} \right],$$

$$\hat{\beta}_2^{(u)} = \frac{1}{n^2 - (2m-6)^2}\left[n\underline{X}_2^{(u)\prime}\underline{y} - (2m-6)\sum_{i=1}^{n} y \right],$$

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_m \end{bmatrix} = (X_1{}'X_1)^{-1}X_1{}'\underline{y}. \tag{5.17}$$

Moreover, as seen in equation (5.16), $\left| X^{(u)\prime}X^{(u)} \right|$ is the same for all u, implying $V_2^{(u)}$ is constant for all u, and thus the balanced design consisting of S_0, S_1, S_{m-1}, S_m has P_1^V.

5.8 Augmented Designs

Suppose it is of interest to augment a design D, which has P_1^V by adding a treatment/run, denoted a_D, to be observed in the experiment, where a_D represents a treatment that is not already included in the design (it is not of interest to replicate the treatments). The new design, denoted D_{+1}, may or may not have P_1^V.

Example 5.11

Consider design D9.2 from Example 5.3. Recall that this is an unbalanced design for $m = 5$ and $n = 7$ with the property P_1^V with $V_2^{(u)}$=0.625 for all u when estimating the parameters in the models containing the general mean, main effects, and $k = 1$ two-factor interaction effect. The treatments for this design can be seen in Table 3.3. Suppose an additional treatment, $a_D = [1, -1, 1, -1, -1]$, is chosen to be included in this design. The augmented design for $m = 5$ with $n = 8$ is denoted $\mathrm{D9.2}_{+1}^1$. Table 5.11 presents the values for the two-factor interaction effect for each of the ten models. It can be seen that the augmented design does not have the property P_1^V. Instead, $\mathrm{D9.2}_{+1}^1$ has the property $P_2^V(4,6)$.

Table 5.11: Values of $V_2^{(u)}$ for $\mathrm{D9.2}_{+1}^1$

u									
1	2	3	4	5	6	7	8	9	10
0.375	0.375	0.375	0.375	0.375	0.250	0.250	0.250	0.250	0.375

Example 5.12

As in Example 5.11, we again consider design D9.2 with the property P_1^V. We now add the treatment $a_D = [1,1,1,1,1]$ to the design. Denote this augmented design as $\mathrm{D9.2}_{+1}^2$. Table 5.12 presents the values for the two-factor interaction effects for each of the ten models where each model includes the general mean, main effects, and $k = 1$ two-factor interaction effect. This augmented design $D9.2^2_{+1}$ for $m=5$ with $n=8$ has the property P_1^V, just as the original un-augmented design D9.2 for $m=5$ with $n=7$ has the property P_1^V.

Table 5.12: Values of $V_2^{(u)}$ for $D9.2^2_{+1}$

				u					
1	2	3	4	5	6	7	8	9	10
0.430	0.430	0.430	0.430	0.430	0.430	0.430	0.430	0.430	0.430

In fact, it is possible to systematically choose a_D such that the augmented design will have P_1^V. There are two types of designs which are of particular interest: 1) balanced designs and 2) unbalanced designs.

Consider a balanced design D^B which has P_1^V by ***Theorem 5*** for experiments which are estimating the effects of the general mean, m main effects, and $k=1$ two-factor interaction effect. Then D^B_{+1} will have P_1^V for $a_D=[1,1,...,1]$ or $a_D=[-1,-1,...,-1]$. Obviously, D^B_{+1} will remain a balanced design since $a_D=[1,1,...,1]=S_0$ and $a_D=[-1,-1,...,-1]=S_m$.

Now consider an unbalanced design, denoted D^U. Using the notation from equation (5.2), suppose D^U has the associated particular form

$$W^{-1}=\left(X_1'X_1\right)^{-1}=\left[\begin{array}{c|c}(a-b)I_s+bJ_s & cJ_s\\ \hline cJ_s & (a-b)I_s+bJ_s\end{array}\right] \qquad (5.18)$$

where $X_1 = \left[\underline{j} \quad \mathrm{D}^{\mathrm{U}}\right]$ represents the $n \times (1+m)$ design matrix with $\underline{j}$ being a vector of all ones, a,b,c are real-valued constants, I_s is an identity matrix of size s, J_s is a square matrix of size s with all entries equal to one, and $1+m=2s$. Then for estimating the effects of the general mean, m main effects, and $k=1$ two-factor interactions, $\underline{w}_u$ will be a vector of length $2s$ for all $u=1,\ldots,v$ models. Consider a vector $\underline{w}_u$ of the following type

$$\underline{w}_u{}' = \left[\underline{w}_s{}' \;\vdots\; \underline{w}_s{}' P'\right] \text{ for all } u \qquad (5.19)$$

where P is an $s \times s$ permutation matrix and a,b,c are such that $\underline{w}_u{}' W^{-1} \underline{w}_u$ is constant for all u, then D^{U} will have P_1^V by ***Theorem 1***. It is of interest to augment such a design, D^{U}, with property P_1^V by adding an additional treatment such that the resulting design, $\mathrm{D}_{+1}^{\mathrm{U}}$, will also have the property P_1^V.

Theorem 6: For $\underline{w}_u$ satisfying (5.19), the following results are true where $\underline{w}_u$ is denoted as $\underline{w}$ WLOG since all u models give the same form:

(a) $\underline{j}'\left(X_1'X_1\right)^{-1} = constant \times \underline{j}'$

(b) $\underline{j}'\underline{w} = \underline{w}'\underline{j} = 0$

(c) $\underline{j}'\left(X_1'X_1\right)^{-1}\underline{w} = \underline{w}'\left(X_1'X_1\right)^{-1}\underline{j} = 0$

(d) $\underline{w}'\left(X_1'X_1\right)^{-1}\underline{w} = 2(a-b)\left(w_1^2 + \cdots + w_s^2\right) + 2(b-c)\left(w_1 + \cdots w_s\right)^2$

(e) $\underline{w}'\left(X_1'X_1 + J_{2s}\right)^{-1}\underline{w} = \underline{w}'\left(X_1'X_1\right)^{-1}\underline{w}$

(f) $\underline{j}'\left(X_1'X_1 + J_{2s}\right)^{-1}\underline{w} = \underline{w}'\left(X_1'X_1 + J_{2s}\right)^{-1}\underline{j} = 0$

(g) $\underline{j}'\left(X_1'X_1+J_{2s}\right)^{-1}\underline{j}=\dfrac{\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}{1+\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}$

(h) $\left(\delta\underline{j}+\underline{w}\right)'\left(X_1'X_1+J_{2s}\right)^{-1}\left(\underline{w}+\delta\underline{j}\right)=$

$$\underline{w}'\left(X_1'X_1\right)^{-1}\underline{w}+\delta^2\frac{\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}{1+\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}$$

Proof: (a) and (b) are easily proven.

(c) Follows from (a) and (b).

(d)
$$\begin{aligned}\underline{w}'\left(X_1X_1\right)^{-1}\underline{w}&=\left[\underline{w}_s' \;\vdots\; -\underline{w}_s'P'\right]\left[\begin{array}{c:c}(a-b)I_s+bJ_s & cJ_s\\ \hdashline cJ_s & (a-b)I_s+bJ_s\end{array}\right]\left[\begin{array}{c}\underline{w}_s\\ \hdashline -P\underline{w}_s\end{array}\right]\\
&=\underline{w}_s'\left((a-b)I_s+bJ_s\right)\underline{w}_s+\underline{w}_s'P'\left((a-b)I_s+bJ_s\right)P\underline{w}_s-2c\underline{w}_s'J_sP\underline{w}_s\\
&=2\underline{w}_s'\left((a-b)I_s+bJ_s\right)\underline{w}_s-2c\underline{w}_s'J_s\underline{w}_s\\
&=2\left[(a-b)\underline{w}_s'\underline{w}_s+b\underline{w}_s'\underline{j}\,\underline{j}'\underline{w}_s\right]-2c\underline{w}_s'\underline{j}\,\underline{j}'\underline{w}_s\\
&=2\left[(a-b)\left(w_1^2+\cdots+w_s^2\right)+b\left(w_1+\cdots+w_s\right)^2\right]-2c\left(w_1+\cdots+w_s\right)^2\\
&=2(a-b)\left(w_1^2+\cdots+w_s^2\right)+2(b-c)\left(w_1+\cdots+w_s\right)^2\end{aligned}$$

(e)
$$\begin{aligned}\left(X_1'X_1+J_{2s}\right)^{-1}&=\left(X_1'X_1+\underline{j}\,\underline{j}'\right)^{-1}\\
&=\left(X_1'X_1\right)^{-1}-\frac{\left(X_1'X_1\right)^{-1}\underline{j}\,\underline{j}'\left(X_1'X_1\right)^{-1}}{1+\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}\end{aligned}$$

From (c), $\underline{w}'\left(X_1'X_1\right)^{-1}\underline{j}=\underline{j}'\left(X_1'X_1\right)^{-1}\underline{w}=0.$

Hence, (e) is true.

(f) By (c).

(g) $$\underline{j}'\left(X_1'X_1+J_{2s}\right)^{-1}\underline{j}=\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}-\frac{\left(\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}\right)^2}{1+\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}$$
$$=\frac{\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}{1+\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}$$

(h) By (c), (e), (f), and (g). This completes the proof.

Theorem 7: An augmented design, D_{+1}^{U}, has P_1^V iff $\left(\delta\underline{j}+\underline{w}_u\right)'\left(X_1'X_1+J_{2s}\right)^{-1}\left(\underline{w}_u+\delta\underline{j}\right)$ is constant for all u models.

Corollary 7.1: An augmented design, D_{+1}^{U}, has P_1^V for $a_D=[1,1,...,1]$ or $a_D=[-1,-1,...,-1]$.

Proof: From ***Theorem 6 (h)***, letting $\delta=\pm1$ indicates $a_D=\delta\underline{j}$ can be chosen as $[1,1,...,1]$ or $[-1,-1,...,-1]$ and D_{+1}^{U} will have P_1^V by ***Theorem 1***. This can be seen easily, as $\left(\delta\underline{j}+\underline{w}\right)'\left(X_1'X_1+J_{2s}\right)^{-1}\left(\underline{w}+\delta\underline{j}\right)$ will remain constant for all u models since $\underline{w}'\left(X_1'X_1\right)^{-1}\underline{w}$ was constant for D^U and $\dfrac{\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}{1+\underline{j}'\left(X_1'X_1\right)^{-1}\underline{j}}$ is independent of u.

Example 5.13

Design D9.2 for $m=5$ and $n=7$ is an unbalanced design with the specified form presented in this discussion. D9.2 has P_1^V for $k=1$ two-factor interaction effect included in the model with the general mean and main effects. Thus, it is possible to add a run, either $[1,1,1,1,1]$ or $[-1,-1,-1,-1,-1]$, to D9.2 and P_1^V will still hold. As an illustration, let $\delta = 1$, giving $a_D = \delta \underline{j} = [1,1,1,1,1]$ since the original design already includes $[-1,-1,-1,-1,-1]$. From ***Theorem 6 (h)***, it can be calculated

$$\left(\delta \underline{j} + \underline{w}_u\right)' \left(X_1{}'X_1 + J_{2s}\right)^{-1} \left(\underline{w}_u + \delta \underline{j}\right) = \underline{w}_u{}' \left(X_1{}'X_1\right)^{-1} \underline{w}_u + \delta^2 \frac{\underline{j}' \left(X_1{}'X_1\right)^{-1} \underline{j}}{1 + \underline{j}' \left(X_1{}'X_1\right)^{-1} \underline{j}}$$

$$\left(\underline{j} + \underline{w}_u\right)' \left(X_1{}'X_1 + J_{2s}\right)^{-1} \left(\underline{w}_u + \underline{j}\right) = \underline{w}_u{}' \left(X_1{}'X_1\right)^{-1} \underline{w}_u + \frac{\underline{j}' \left(X_1{}'X_1\right)^{-1} \underline{j}}{1 + \underline{j}' \left(X_1{}'X_1\right)^{-1} \underline{j}}$$

$$5.6727 = 5.4 + \frac{0.375}{1.375}$$

$$5.6727 = 5.4 + 0.2727 \qquad \text{all } u$$

where $W^{-1} = \left(X_1{}'X_1\right)^{-1}$ and $\underline{w}_u$, $u = 1,\ldots,10$, are given in Example 5.3. Therefore, D9.2_{+1} for $m = 5$ and $n = 8$ has $P_1^V(10)$ with $V_2^{(u)} = \left(n - \left(\underline{j} + \underline{w}_u\right)' \left(X_1{}'X_1 + J_{2s}\right)^{-1} \left(\underline{w}_u + \underline{j}\right)\right)^{-1}$ $= \left(8 - 5.6727\right)^{-1} = 0.4296$ by equation (5.2).

Chapter 6

Finding Designs with P_g^V

6.0 Introduction

All designs can be classified according to the property P_g^V. It is of particular interest to characterize and construct designs having the property P_g^V where $g = 1$. This case guarantees that the uncommon parameters of the possible models are estimated with the same precision across models. In this chapter, a theorem is given which provides an approach for finding designs having the property P_1^V for general m. Additionally, methods for systematically constructing designs with P_1^V are detailed for $m = 3$ and $m = 4$ for all values of $n = m+2,...,\binom{m}{2}$ and $k = 1$. Other occurrences of the property P_g^V for $k > 1,\ g > 1$ are also detailed.

6.1 Design Decomposition where P_1^V Holds

Consider a design D for m factors having n treatments and the capability of estimating p parameters in the model

$$E\left(\underline{y}\right) = X^{(u)}\underline{\beta}^{(u)} \text{ for } u = 1,...,v \qquad (6.1)$$

where $\underline{y}\left(n\times 1\right)$ is the vector of observed responses, $X^{(u)}(n\times p)$ is a matrix dependent on D and model of interest, and $\underline{\beta}^{(u)}\left(p\times 1\right)$ is the vector of unknown parameters. We define a new design, denoted D$_{-t}$, such that t observations are removed from design D. Therefore, D$_{-t}$ is composed of $n-t$ treatments which are a subset of the n treatments in D. We partition the components of equation (6.1) as

$$\underline{y} = \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix}, \quad X^{(u)} = \begin{pmatrix} X_1^{(u)} \\ X_2^{(u)} \end{pmatrix} \text{ all } u, \tag{6.2}$$

where $\underline{y}_1 \left((n-t)\times 1\right)$ and $\underline{y}_2 \left(t\times 1\right)$ are vectors consisting of the observed responses and $X_1^{(u)}\left((n-t)\times p\right)$ and $X_1^{(u)}\left(t\times p\right)$ are sub-matrices of $X^{(u)}(n\times p)$. Then it is possible to calculate the variance-covariance matrix [Rao (2002)] for D$_{-t}$ as

$$\left(X_1^{(u)\prime}X_1^{(u)}\right)^{-1} = \left(X^{(u)\prime}X^{(u)}\right)^{-1} + \left(X^{(u)\prime}X^{(u)}\right)^{-1} X_2^{(u)\prime}\left(I_t - M\right)^{-1} X_2^{(u)}\left(X^{(u)\prime}X^{(u)}\right)^{-1} \tag{6.3}$$

where $I_t\left(t\times t\right)$ is an identity matrix and $M\left(t\times t\right)$ is a matrix calculated as $X_2^{(u)}\left(X^{(u)\prime}X^{(u)}\right)^{-1} X_2^{(u)\prime}$.

There exist two special cases for designs which have the ability to identify and discriminate the models of interest which allow for (6.3) to be simplified.

Case 1: A design D giving $X^{(u)\prime}X^{(u)} = nI_p$. For designs of this form, (6.3) is calculated as $\left(X_1^{(u)\prime}X_1^{(u)}\right)^{-1} = \frac{1}{n}I_p + \frac{1}{n^2}X_2^{(u)\prime}\left(I_t - M\right)^{-1}X_2^{(u)}$ where $M = \frac{1}{n}X_2^{(u)}X_2^{(u)\prime}$.

Case 2: A design D$_{-t}$ for $t=1$ giving $X^{(u)\prime}X^{(u)} = nI_p$. Then M in (6.3) becomes $M = \frac{p}{n}$ and (6.3) is calculated as $\left(X_1^{(u)\prime}X_1^{(u)}\right)^{-1} = \frac{1}{n}I_p + \frac{1}{n(n-p)}X_2^{(u)\prime}X_2^{(u)}$.

Theorem 8: When $X^{(u)\prime}X^{(u)} = nI_p$ and $k = 1$, the design D has the property P_1^V.

Theorem 9: When $t = 1$, $X^{(u)\prime}X^{(u)} = nI_p$ and $k = 1$, the design D_{-1} has the property P_1^V.

Proof: We have $Var\left(\hat{\beta}_2^{(u)}\right) = \sigma^2\left(\frac{1}{n} + \frac{1}{n(n-p)}\right) =$ a constant independent of u. This completes the proof.

Using the above results, as well as other methods for systematically removing treatments from the full factorial designs (in other words, methodically selecting particular treatments from the full design) to create fractional factorial designs, it is possible to find designs having the property P_1^V.

6.2 Designs with P_g^V for *m*=3, *k*=1, *g*=1

Consider an experiment which has $m = 3$ factors (A, B, C) with two levels each. By dividing the full factorial design into two groups, it is possible to identify all designs with P_1^V. Although $m = 3$ is a simple case of a 2^m factorial experiment, its investigation and explanation of composing fractional factorial designs giving P_1^V allows insight for working with experiments which have larger values of m.

6.2.1 Full Design Decomposition

A 2^3 full factorial design has $n = 8$ treatments. These eight treatments can be separated into two groups with four treatments in each group as seen in Figure 6.1.

A	B	C	
+	+	+	Group 1 = S_0, S_2
+	–	–	
–	+	–	
–	–	+	
–	–	–	Group 2 = S_1, S_3
–	+	+	
+	–	+	
+	+	–	

Figure 6.1: 2^3 Full factorial design

Note that the treatments in Group 1 are negative of the treatments in Group 2. The treatments are then divided into four pairs with each pair containing the two treatments that are negative of each other; one treatment is from Group 1 and one treatment is from Group 2. Moreover, each treatment in a pair is assigned as *a* or *b* as seen in Figure 6.2.

		A	B	C
P_1	*a*	+	+	+
	b	–	–	–
P_2	*a*	+	–	–
	b	–	+	+
P_3	*a*	–	+	–
	b	+	–	+
P_4	*a*	–	–	+
	b	+	+	–

Figure 6.2: Pairs 1-4 for a 2^3 factorial design

Each treatment can then be referenced according to its pair number and letter. For example, treatment 1*a* corresponds to (+, +, +) and treatment 4*b* corresponds to (+, +, –).

Note that the full factorial design can be assembled by combining Group 1 and Group 2 or combining P_1–P_4. These divisions (or a combination of) can be used to form fractional factorial designs for various values of n which are capable of identifying and discriminating between models containing the general mean, main effects, and $k = 1$ two-factor interaction effect with P_1^V.

6.2.2 Number of 2^3 Fractional Factorial Designs with P_1^V

Table 6.1 summarizes fractional factorial designs with specified value of n for a 2^3 factorial experiment. The *Number of Possible Designs* is found by taking $\binom{2^3}{n} = \binom{8}{n}$. This gives the total number of ways that n treatments can be selected from the full factorial design. The *Number of Capable Designs* represents the number of designs with the ability to identify and discriminate the models containing the general mean, main effects, and $k = 1$ two-factor interaction effect. The *Number of Designs with* P_1^V is then given, followed by the value of $V_2^{(u)}$, the common variance of the two-factor interaction effect. Because the value of $V_2^{(u)}$ may differ for the designs, the *Number of Designs with the Specified* $V_2^{(u)}$ is stated. Note that the sum of the values in the last column of the table should equal the value in the *Number of Designs with* P_1^V column.

Table 6.1: Summary of 2^3 Fractional Factorial Designs for *n*=5-8

n	# of Possible Designs	# of Capable Designs	# of Designs with P_1^V	Value of $V_2^{(u)}$	# of Designs with Specified $V_2^{(u)}$
5	56	8	8	0.5	All
6	28	16	16	0.1875 0.25	4 12
7	8	8	8	0.1667	All
8	1	1	1	0.125	All

It is seen in Table 6.1 that all designs which are capable of identifying and discriminating the models consisting of the general mean, main effects and one two-factor interaction effect also give the same $V_2^{(u)}$ value for all three models.

6.2.3 *n=5*

Consider fractional factorial designs with $n=5$ treatments for a 2^3 factorial experiment having the property P_1^V. There are eight such designs giving $V_2^{(u)}$=0.5 for all $u=1,2,3$. Four of these designs can be constructed by combining Group 1 and one treatment from Group 2. Similarly, the remaining four designs can be constructed by combining Group 2 and one treatment from Group 1. Therefore, the eight designs having P_1^V with $n=5$ can be found systematically using the design decomposition given in Section 6.2.1.

6.2.4 *n=6*

There are sixteen fractional factorial designs with P_1^V for $n=6$. Of these 16 designs, 4 designs have $V_2^{(u)}$=0.1875 and 12 designs have $V_2^{(u)}$=0.25. The construction for the designs giving the smallest $V_2^{(u)}$ value are considered first.

The fractional factorial designs giving $V_2^{(u)}$=0.1875 can be constructed by selecting three unique pairs from P_1, P_2, P_3, and P_4. There are $\binom{4}{3}=4$ possible ways to make these selections, thus resulting in the four fractional factorial designs with P_1^V and

$V_2^{(u)}$=0.1875. Note that these designs are in a special class of designs known as foldover designs.

The twelve fractional factorial designs giving $V_2^{(u)}$=0.25 are formed by choosing a group of treatments (Group 1 or Group 2) and selecting two treatments from the other group. For example, it is possible to start with Group 1 and add treatments 1*b* and 2*b* from Group 2. When starting with Group 1, there are $\binom{4}{2}=6$ ways to select two treatments from Group 2. Similarly, when starting with Group 2, there are six ways to choose two treatments from Group 1. Using these two methods, there are twelve resulting fractional factorial designs with $n=6$, P_1^V, and $V_2^{(u)}$=0.25.

6.2.5 *n=7*

Table 6.1 reveals that there are $\binom{8}{7}=8$ possible designs with $n=7$ and that all of these designs have P_1^V with a $V_2^{(u)}$ value of 0.1667 (rounded). Note that the 8 designs can be found by removing $t=1$ treatment from the full factorial design for all eight treatments. Then by ***Theorem 9***, these eight designs have P_1^V.

6.2.6 *n=8*

The design for $n=8$ is of the form given in equation (6.3). Therefore, by ***Theorem 8***, this design has P_1^V and $V_2^{(u)}$ is calculated as 0.125.

6.2.7 Summary

Table 6.2 is presented to summarize the *Number of Designs Identified using the Decomposition* of the 2^3 full factorial design and the methods described in this chapter. Note that all but the last column of this table are identical to Table 6.1.

Table 6.2: Summary of 2^3 fractional factorial designs identified using the decomposition

n	# of Possible Designs	# of Capable Designs	# of Designs with P_1^V	Value of $V_2^{(u)}$	# of Designs with Specified $V_2^{(u)}$	# of Designs Identified using the Decomposition
5	56	8	8	0.5	All	All
6	28	16	16	0.1875 0.25	4 12	All
7	8	8	8	0.1667	All	All
8	1	1	1	0.125	All	All

6.3 Designs with P_g^V for m=4, k=1, g=1

Consider an experiment which has $m = 4$ factors (A, B, C, D) with two levels each. It is found that by knowing only two fractional factorial designs from a 2^4 factorial experiment, it is possible to identify all possible designs with the property P_1^V.

6.3.1 Full Design Decomposition

Consider a 2^4 factorial design with $n = 16$ treatments. The full factorial design can be separated into two distinct fractional factorial designs consisting of eight treatments each, as seen in Figure 6.3. Note that each of the fractional factorial designs formed will consist of complete sets of treatments, denoted S_i, as defined in Section 3.5.

A	B	C	D	
+	–	–	–	Design I = S_1, S_3
–	+	–	–	
–	–	+	–	
–	–	–	+	
–	+	+	+	
+	–	+	+	
+	+	–	+	
+	+	+	–	
+	+	+	+	Design II = S_0, S_2, S_4
+	+	–	–	
+	–	+	–	
+	–	–	+	
–	–	–	–	
–	–	+	+	
–	+	–	+	
–	+	+	–	

Figure 6.3: 2^4 Full Factorial Design

The first of these designs is then separated into two groups, Group 1 and Group 2, as shown in Figure 6.4.

A	B	C	D	
+	–	–	–	Group 1 = S_3
–	+	–	–	
–	–	+	–	
–	–	–	+	
–	+	+	+	Group 2 = S_1
+	–	+	+	
+	+	–	+	
+	+	+	–	

Figure 6.4: Design I

Note that the four treatments in Group 1 are negative of the four treatments in Group 2. Using this fact, the treatments in Design I are coupled to get four pairs, P_5, P_6, P_7, and P_8. The treatments in each pair are further labeled as *a* & *b* and are referenced according to the pair number and the letter label as seen in Figure 6.5. For example, treatment 5*a* is (+,–,–,–) and treatment 7*b* is (+,+,–,+).

		A	B	C	D
P_5	*a*	+	–	–	–
	b	–	+	+	+
P_6	*a*	–	+	–	–
	b	+	–	+	+
P_7	*a*	–	–	+	–
	b	+	+	–	+
P_8	*a*	–	–	–	+
	b	+	+	+	–

Figure 6.5: Pairs 5-8 from Design I

Similarly, consider the remaining eight treatments from the full factorial design (Figure 6.3) which were not included in Design I (Figure 6.4). This experimental design is labeled Design II and as done with the first design, its eight treatments are divided into two groups, Group 3 and Group 4, as shown in Figure 6.6.

A	B	C	D	
+	+	+	+	Group 3
+	+	–	–	
+	–	+	–	
+	–	–	+	
–	–	–	–	Group 4
–	–	+	+	
–	+	–	+	
–	+	+	–	

Figure 6.6: Design II

Again, the treatments in Group 3 are negative of the treatments in Group 4 and this is used to pair the treatments into four pairs, P_1, P_2, P_3, and P_4, with each pair having a treatment *a* and *b* as illustrated in Figure 6.7.

		A	B	C	D
P_1	*a*	+	+	+	+
	b	–	–	–	–
P_2	*a*	+	+	–	–
	b	–	–	+	+
P_3	*a*	+	–	+	–
	b	–	+	–	+
P_4	*a*	+	–	–	+
	b	–	+	+	–

Figure 6.7: Pairs 1-4 from Design II

Note that the full factorial design can be assembled by combining (1) Design I and Design II, (2) Group 1-Group 4, or (3) P_1-P_8. These divisions (or a combination of) can be used to form fractional factorial designs for various values of n which are capable of identifying and discriminating between models containing the general mean, main effects, and $k = 1$ two-factor interaction effect with P_1^V.

6.3.2 Number of 2^4 Fractional Factorial Designs with P_1^V

Table 6.3 summarizes fractional factorial designs having P_1^V with specified value of n for a 2^4 factorial experiment. The *Number of Possible Designs* is found by taking $\binom{2^4}{n} = \binom{16}{n}$. This gives the total number of ways that n treatments can be selected from the full factorial design. The remaining columns summarize the designs having the property P_1^V as described in Section 6.2.2.

Table 6.3: Summary of 2^4 fractional factorial designs for *n*=6-16

n	# of Possible Designs	# of Capable Designs	# of Designs With P_1^V	Value of $V_2^{(u)}$	# of Designs with Specified $V_2^{(u)}$
6	8008	272	16	0.875	All
7	11,440	1920	80	0.1875	All
8	12,870	4954	2	0.125	All
9	11,440	7104	80	0.116071 0.140625	16 64
10	8008	6520	24	0.104167 0.125	8 16
11	4368	4080	96	0.100446 0.109375 0.121875	64 16 16
12	1820	1796	32	0.09375	All
13	560	560	64	0.084375	All
14	120	120	8	0.07291	All
15	16	16	16	0.06875	All
16	1	1	1	0.0625	All

Since there are such a large number of designs which have the ability to identify and discriminate the models consisting of the general mean, main effects and one two-factor interaction effect while giving the same $V_2^{(u)}$ value for all six models, it is of interest to develop a method for systematically constructing these fractional factorial designs.

6.3.3 *n*=6

Consider fractional factorial designs composed of $n = 6$ treatments from the full factorial design. There are 16 designs with P_1^V with all designs giving $V_2^{(u)}$=0.875 for all u models.

All sixteen of these designs can be constructed using the following method. Start by selecting a pair of treatments from Design I (*P*5, *P*6, *P*7, *P*8). Let x_1, x_2, x_3, x_4 represent the level of Factors A, B, C, D, respectively, and $x_i = \alpha$, $i = 1,2,3,4$ where $\alpha = 1,-1$. Then by defining a relationship between x_1, x_2, x_3, x_4 for the selected pair, the remaining four treatments are chosen from Design II (*P*1, *P*2, *P*3, *P*4) such that the relationship is satisfied. Similarly, one can start by selecting a pair of treatments from Design II and choosing the remaining four treatments from Design I (satisfying the relationship in the selected pair from Design II). For each initial pair of treatments selected, there will be two designs ($\alpha = 1,-1$) resulting from this procedure. Thus, $2 \times \binom{4}{1} \times 2 = 16$ designs are obtained for *n*=6 which give $V_2^{(u)}$=0.875.

For example, consider constructing a design for *n*=6 by selecting the pair P_5 from Design I. The relationship observed in P_5 can be written as $\{x_1\} \neq \{x_2 = x_3 = x_4\}$. Four treatments are then chosen from Design II (P_1, P_2, P_3, P_4) such that $\{x_1\} \neq \{x_2 = x_3 = x_4\}$. For example, consider $x_1 = 1$. The treatments from Design II which meet this criterion are 1*a*, 2*a*, 3*a*, and 4*a*. There are no treatments satisfying $\{x_2 = x_3 = x_4\} = -1$. Therefore, the resulting fractional factorial design for $n = 6$ is P_5, 1*a*, 2*a*, 3*a*, 4*a*. Similarly, a fractional factorial design for $n = 6$ can be constructed by selecting this same pair, P_5, with the same defining relationship as specified above, but letting $x_1 = -1$ and $\{x_2 = x_3 = x_4\} = 1$. The resulting design will be P_5, 1*b*, 2*b*, 3*b*, 4*b*. Thus, two designs are composed from the pair P_5.

6.3.4 *n*=7

Consider fractional factorial designs composed of $n = 7$ treatments from the full factorial design. There are 80 designs with P_1^V with all designs giving $V_2^{(u)}$=0.1875 for all u models.

Of these 80 designs, 40 can be assembled by taking three pairs from Design II (P_1, P_2, P_3, P_4) and adding one treatment from the remaining ten treatments in the full factorial design. For example, it is possible to choose the treatments in P_1, P_2, P_3 and treatment 4*a* or P_2, P_3, P_4 and treatment 6*b* to get designs for $n = 7$ with P_1^V .Similarly, it is possible to take three pairs from Design I (P_5, P_6, P_7, P_8) and add one treatment from the remaining ten treatments in the full factorial design. This results in the remaining 40 designs for $n = 7$ which have the ability to identify and discriminate the models and give $V_2^{(u)}$ constant.

Thus, the 80 designs for $n = 7$ giving P_1^V with $V_2^{(u)}$=0.1875 can be found using the pairs P_5, P_6, P_7, P_8 from Design I and P_1, P_2, P_3, P_4 from Design II.

6.3.5 *n=8*

Consider fractional factorial designs with $n=8$ treatments from the full factorial design. Design I, as illustrated in Figure 6.4, is such a design. Further, this design has P_1^V with $V_2^{(u)}$=0.125 for all u models.

Similarly, consider Design II from Figure 6.6. This design also has P_1^V with $V_2^{(u)}$=0.125 for all u models.

Furthermore, as seen in Table 6.3, it is found that of all possible combinations of $n=8$ treatments from the full factorial design (a total of 12,870 designs), Design I and Design II are the only such designs for $n=8$ giving P_1^V. Additionally, these designs can be characterized as foldover designs.

6.3.6 *n=9*

For fractional factorial designs consisting of $n=9$ treatments, there are 80 designs with P_1^V. Of these eighty designs, 16 give $V_2^{(u)}$=0.116071 (rounded) and 64 give $V_2^{(u)}$=0.140625. The 16 fractional factorial designs giving the smaller $V_2^{(u)}$ of 0.116071 are considered first.

Consider Design I with $n=8$. If one treatment from Group 3 or Group 4 (the treatments forming Design II) are added to this design, it is possible to construct 8 designs with $n=9$. These 8 designs have P_1^V with $V_2^{(u)}$=0.116071.

Similarly, consider Design II with $n = 8$ and add one treatment from Group 1 or Group 2 (the treatments forming Design I) to obtain an additional 8 designs with $n = 9$ and P_1^V with $V_2^{(u)}$=0.116071.

Thus, the 16 designs for $n = 9$ satisfying P_1^V with the smaller $V_2^{(u)}$ of 0.116071 can be found using Design I and Design II.

The remaining 64 designs for $n = 9$ which give $V_2^{(u)}$=0.140625 can also be found from Design I and Design II by utilizing a comparable approach to that seen for the 16 designs for *n*=7. Start by choosing three pairs from Design I (P_5, P_6, P_7, P_8). A relationship is then defined between the treatments in the pair that was not chosen from this design. Using this relationship, three treatments are selected from Design II to get a fractional factorial design for $n = 9$.

Similarly, three pairs from Design II (P_1, P_2, P_3, P_4) can be chosen and the remaining three treatments are selected from Design I such that the relationship defined by the un-chosen pair in Design II is satisfied. For each set of initial three pairs chosen, there will be eight designs resulting from this procedure. Thus resulting in $2 \times \binom{4}{3} \times 8 = 64$ designs for $n = 9$ which give $V_2^{(u)}$=0.140625.

As an illustration of this method, suppose P_1, P_2, and P_3 are chosen. The relationship observed in P_4 (the pair not selected) can be written as $\{x_1 = x_4\} \neq \{x_2 = x_3\}$. Three treatments are then chosen from Design I (P_5, P_6, P_7, P_8) such that $\{x_1 = x_4\} \neq \{x_2 = x_3\}$. For example, consider $x_1 = 1$. The treatments from Design I which meet this criterion are 5*a*, 6*b*, 7*b*, and 8*b*. However, because 8*b* does not satisfy the relationship $\{x_1 = x_4\} \neq \{x_2 = x_3\}$, it is dropped and the resulting fractional factorial design for $n = 9$ is P_1, P_2, P_3, 5*a*, 6*b*, 7*b*.

6.3.7 *n=10*

Consider fractional factorial designs with $n = 10$ treatments. There are 24 designs with P_1^V where 8 designs give $V_2^{(u)}$=0.104167 (rounded) and 16 designs give $V_2^{(u)}$=0.125. The 8 designs giving the smaller $V_2^{(u)}$ value are considered first.

Take Design I with $n = 8$ and add one pair from Design II for all four pairs to get 4 designs with $n = 10$, P_1^V, and $V_2^{(u)}$=0.104167.

Similarly, take Design II with $n = 8$ and add a pair from Design I for all four pairs to obtain the remaining 4 designs with $n = 10$, P_1^V, and $V_2^{(u)}$=0.104167.

Thus, it is seen that the 8 designs for $n = 10$ P_1^V, and $V_2^{(u)}$=0.104167 are found using Design I, Design II, and their respective pairs.

The remaining 16 designs which satisfy the property of interest giving $V_2^{(u)}$=0.125 can be obtained from Design I and Design II using a comparable approach to that seen for the 64 designs for $n = 9$. Begin by choosing three pairs from Design I (P_5, P_6, P_7, P_8) and using the pair that was not chosen to define a relationship between its treatments which the remaining four treatments chosen from Design II must satisfy. Similarly, three pairs from Design II (P_1, P_2, P_3, P_4) are chosen with the remaining four treatments from Design I selected such that the relationship defined by the un-chosen pair in Design II is satisfied. For each set of initial three pairs chosen, there will be two designs resulting from this procedure. Thus $2\times\binom{4}{3}\times 2=16$ designs are obtained for $n=10$ which give $V_2^{(u)}$=0.125.

For example, suppose P_5, P_6, and P_7 are selected. The defining relationship of P_8 (the pair not selected) can be written as $\{x_1=x_2=x_3\}\neq\{x_4\}$. Four treatments from Design II are then chosen such that $\{x_1=x_2=x_3\}\neq\{x_4\}$. When considering $x_1=x_2=x_3=1$, treatment 1*a* meets the criteria and when considering $x_4=-1$, the satisfying treatments are 2*a*, 3*a*, and 4*b*. Thus the resulting design is P_5, P_6, P_7, 1*a*, 2*a*, 3*a*, 4*b*. The second design corresponding with the pairs P_5, P_6, and P_7 is then found by considering $x_1=x_2=x_3=-1$ and $x_4=1$.

6.3.8 *n=11*

There are 96 designs which give a constant $V_2^{(u)}$ value for fractional factorial designs consisting of $n=11$ treatments. Of these 96 designs, there are 64 designs which give $V_2^{(u)}$=0.100446, 16 which give $V_2^{(u)}$=0.109375, with the remaining 16 giving $V_2^{(u)}$=0.121875. It is of interest to first construct the 64 designs which give the smallest constant $V_2^{(u)}$ value.

By using Design I as a base, only 3 more treatments are needed to get a design of $n=11$ treatments. These three treatments are selected from Design II in such a way that no two treatments come from the same pair. For example, treatments 1*a*, 2*a*, 3*b* can be added to Design I to get a design which satisfies P_1^V. However, treatments 1*a*, 1*b*, 2*a* cannot be selected since 1*a* and 1*b* are both in P_1; this will not give a design which gives a constant $V_2^{(u)}$. Using this criteria, we are able to construct 32 designs giving P_1^V.

Similarly, it is possible to start with Design II as a base and add 3 treatments from Design I. Again, there cannot be a complete pair of treatments included in the three added treatments. This technique will give another 32 designs having the property of interest.

Therefore, the 64 designs with P_1^V and giving the smallest $V_2^{(u)}$ for $n = 11$ treatments can be constructed using Design I, Design II, and their respective pairs.

The remaining 32 designs for $n = 11$ with P_1^V, 16 with $V_2^{(u)}$=0.109375 and 16 with $V_2^{(u)}$=0.121875, are slightly more complicated to assemble. To begin, consider the 16 designs which give $V_2^{(u)}$=0.109375. First, select three pairs from Design I and one treatment from the pair not selected, resulting in seven treatments. Then choose four treatments from Design II, conditioned on the relationship defined by the treatment not selected in Design I. For example, suppose P_5, P_6, P_7, 8*a* are chosen. Then select four treatments from Design II such that the relationship defined by 8*b* is satisfied. Similarly, it is possible to start by choosing the initial seven treatments from Design II and the remaining four treatments from Design I. This method produces $2 \times \binom{4}{3} \times 2 = 16$ fractional factorial designs for $n = 11$ giving a constant $V_2^{(u)}$ value of 0.109375.

The final 16 designs for $n = 11$ which give $V_2^{(u)}$=0.121875 can be obtained similarly to the method above. From Design I, three pairs and one treatment from the pair not chosen are selected to get seven treatments. Instead of selecting the remaining four treatments from Design II conditioned on the relationship of the treatment not chosen in Design I, the selected treatments are conditioned on the treatment chosen in Design I. For example, suppose P_5, P_6, P_7, 8*a* are chosen. The remaining four treatments are selected from Design II such that the relationship defined by 8*a* is satisfied (instead of 8*b* as seen previousely). Similarly, this method can be used starting with Design II and selecting the remaining four treatments from Design I. This will produce $2 \times \binom{4}{3} \times 2 = 16$ designs for $n = 11$ satisfying P_1^V with $V_2^{(u)}$=0.121875.

6.3.9 *n=12*

There are 32 designs with P_1^V where $V_2^{(u)}$=0.09375 for fractional factorial designs with $n = 12$ treatments.

Starting with Design I, four additional treatments are required to get a design with $n = 12$ treatments. To satisfy the property of interest, four additional treatments from Design II are selected such that no complete pair is included. There are 16 such designs.

Similarly, starting with Design II and adding four treatments from Design I such that no complete pair is included, another set of 16 designs are formed. Each of these 16 designs have P_1^V.

Thus, the 32 designs with P_1^V for $n = 12$ can be found using Design I, Design II, and their respective pairs.

6.3.10 *n=13*

There are 64 fractional factorial designs for $n = 13$ that have P_1^V. These 64 designs give $V_2^{(u)}$=0.084375.

Half of the 64 designs mentioned above can be found by starting with the 8 treatments from Design I and adding five treatments in the following way: choose one pair from Design II and one treatment from each of the remaining pairs of Design II. This forms a design of $n = 13$ treatments, all of which have P_1^V. To get the other half of the 64 designs, start with Design II and add one pair from Design I and one treatment from each of remaining pairs of Design I.

Hence, the 64 designs with P_1^V for $n = 13$ can be constructed by using Design I, Design II, and their respective pairs.

6.3.11 *n=14*

Consider fractional factorial designs for $n = 14$ treatments which have P_1^V. There are 8 such designs, all of which give $V_2^{(u)}$=0.07291 (rounded).

Using Design I as a base, 6 additional treatments are needed to get a design with $n = 14$. These 6 treatments can be selected by choosing any three pairs from Design II. This can be done in four ways: (1) P_1, P_2, P_3 (2) P_1, P_2, P_4 (3) P_1, P_3, P_4 (4) P_2, P_3, P_4. By adding these pairs to Design I, 4 of the 8 designs mentioned above are produced.

Similarly, starting with Design II and choosing any three pairs from Design I, results in a design with P_1^V. This can be done in four ways: (1) P_5, P_6, P_7 (2) P_5, P_6, P_8 (3) P_5, P_7, P_8 (4) P_6, P_7, P_8.

Therefore, the 8 designs fulfilling the property of interest for $n=14$ can be found using Design I, Design II, and their respective pairs.

6.3.12 *n=15*

It is possible to construct 16 designs by removing $t=1$ treatment from the full factorial design. These are the 16 designs presented in Table 6.3. By ***Theorem 9***, these designs have P_1^V with a calculated value of $V_2^{(u)}$= 0.06875.

6.3.13 *n=16*

The design for $n=16$ is of the form given in equation (6.3). Therefore, by ***Theorem 8***, this design has the property P_1^V with $V_2^{(u)}$ calculated as 0.0625.

6.3.14 Summary

The final column of Table 6.4 summarizes the *Number of Designs Identified using Design I and II* using the methods described in this chapter.

Table 6.4: Summary of 2^4 fractional factorial designs identified using Design I and Design II

n	# of Possible Designs	# of Capable Designs	# of Designs with P_1^V	Value of $V_2^{(u)}$	# of Designs with Specified $V_2^{(u)}$	# of Designs Identified using Design I and II
6	8008	272	16	0.875	All	All
7	11,440	1920	80	0.1875	All	All
8	12,870	4954	2	0.125	All	All
9	11,440	7104	80	0.116071 0.140625	16 64	All All
10	8008	6520	24	0.104167 0.125	8 16	All All
11	4368	4080	96	0.100446 0.109375 0.121875	64 16 16	All All All
12	1820	1796	32	0.09375	All	All
13	560	560	64	0.084375	All	All
14	120	120	8	0.07291	All	All
15	16	16	16	0.06875	All	All
16	1	1	1	0.0625	All	All

6.4 Number of Designs with P_g^V for m=5, k=1, g=1

Consider an experiment which has $m=5$ factors (A, B, C, D, E) with two levels each. Table 6.5 summarizes the number of fractional factorial designs for specified values of n with $k=1$; particularly, identifying the number of designs having the property P_1^V. This table is similar to Table 6.1 and Table 6.3 which summarize the fractional factorial designs for $m=3$ and $m=4$, respectively. Recall from previous sections, the *Number of Possible Designs* is found by taking $\binom{2^m}{n}=\binom{32}{n}$ for $m=5$ factors. This gives the total number of ways that n treatments can be selected from the full factorial design. The remaining columns summarize the capable designs with the property P_1^V as described in Section 6.2.2.

Table 6.5: Summary of 2^5 fractional factorial designs for *n*=7-9

n	# of Possible Designs	# of Capable Designs	# of Designs with P_1^V	Value of $V_2^{(u)}$	# of Designs with Specified $V_2^{(u)}$
7	3,365,856	54,336	352	0.625 1.375	320 32
8	10,518,300	803,040	4960	0.375 0.4296875 0.5	4320 320 320
9	28,048,800	5,321,760	0	N/A	N/A

The complexity of analyzing fractional factorial designs for $m=5$ quickly becomes cumbersome. It can be seen from Table 6.5 that even for designs consisting of the smallest number of treatments ($n=7$) which may or may not be capable of estimating the general mean, main effects, and 1 two-factor interaction effect, the number of possible designs which must be analyzed is 3,365,856. Moreover, by adding only two additional treatments (considering designs with $n=9$), there are more than 28 million designs which must be analyzed. However, it is worth noting that the property P_1^V still holds as seen in previous chapters.

6.5 P_g^V $k \geq 1, g \geq 1$

Although it is most practical to require a constant precision for the distinguishing elements of the various models being fit to the data, i.e., the design has P_1^V, this cannot always be satisfied. For the case with $k=1$, which has been described in detail thus far, designs with P_1^V can be easily chosen and identified. However, when $k>1$, designs with P_1^V occur less often and in fact are most likely to be seen when $k=\binom{m}{2}$ as is discussed in Section 6.5.1. Moreover, alternative cases for P_g^V are considered; specifically, Section 6.5.2 considers P_g^V for $g>1$ when $k=1$ and Section 6.5.3 considers P_g^V for $g>1$ when $k>1$.

6.5.1 $k > 1, g = 1$

Consider a balanced design which has the ability to unbiasedly estimate the parameters in the model containing the general mean, m main effects, and $k=\binom{m}{2}$ two-factor interaction effects. For this design to have P_1^V, the last k diagonal elements of $V^{(u)}=\left(X^{(u)\prime}X^{(u)}\right)^{-1}$, $u=1,\dots,\nu$ where $\nu=1$ since there is only one distinct model containing all two-factor interaction effects, must be equal. Since the design is balanced, it can be shown that $V^{(u)}$ consists of at most ten distinct values with the positions corresponding to the following positions of $X^{(u)\prime}X^{(u)}$ [Srivastava and Chopra (1971)]

$v_1=\underline{1}'\underline{1}$	$v_6=\underline{x}_j'\underline{x}_{ij}=\underline{x}_i'\underline{x}_{ij}$	
$v_2=\underline{1}'\underline{x}_i$	$v_7=\underline{x}_i'\underline{x}_{i'j}$	$i\neq i', j$
$v_3=\underline{1}'\underline{x}_{ij}$	$v_8=\underline{x}_{ij}'\underline{x}_{ij}$	
$v_4=\underline{x}_i'\underline{x}_i$	$v_9=\underline{x}_{ij}'\underline{x}_{ij'}=\underline{x}_{ij}'\underline{x}_{i'j}$	$i\neq i', j\neq j'$
$v_5=\underline{x}_i'\underline{x}_{i'}$ $\quad i\neq i'$	$v_{10}=\underline{x}_{ij}'\underline{x}_{i'j'}$	$i\neq i'\neq j\neq j'$

As an illustration, consider a balanced fractional factorial design for $m = 4$ factors. Then the variance-covariance matrix corresponding to the model containing the general mean, main effects, and $k = \binom{4}{2} = 6$ two-factor interaction effects will be of the following form

$$V^{(u)} = \left(X^{(u)\prime} X^{(u)}\right)^{-1} = \left[\begin{array}{c|cccc|cccccc} v_1 & v_2 & v_2 & v_2 & v_2 & v_3 & v_3 & v_3 & v_3 & v_3 & v_3 \\ \hline v_2 & v_4 & v_5 & v_5 & v_5 & v_6 & v_6 & v_6 & v_7 & v_7 & v_7 \\ v_2 & v_5 & v_4 & v_5 & v_5 & v_6 & v_7 & v_7 & v_6 & v_6 & v_7 \\ v_2 & v_5 & v_5 & v_4 & v_5 & v_7 & v_6 & v_7 & v_6 & v_7 & v_6 \\ v_2 & v_5 & v_5 & v_5 & v_4 & v_7 & v_7 & v_6 & v_7 & v_6 & v_6 \\ \hline v_3 & v_6 & v_6 & v_7 & v_7 & v_8 & v_9 & v_9 & v_9 & v_9 & v_{10} \\ v_3 & v_6 & v_7 & v_6 & v_7 & v_9 & v_8 & v_9 & v_9 & v_{10} & v_9 \\ v_3 & v_6 & v_7 & v_7 & v_6 & v_9 & v_9 & v_8 & v_{10} & v_9 & v_9 \\ v_3 & v_7 & v_6 & v_6 & v_7 & v_9 & v_9 & v_{10} & v_8 & v_9 & v_9 \\ v_3 & v_7 & v_6 & v_7 & v_6 & v_9 & v_{10} & v_9 & v_9 & v_8 & v_9 \\ v_3 & v_7 & v_7 & v_6 & v_6 & v_{10} & v_9 & v_9 & v_9 & v_9 & v_8 \end{array}\right], \quad u = 1$$

where v_i, $i = 1, \ldots, 10$ represent the variances and covariances between the parameter estimates in the model. Then, it is apparent that the design has P_1^V with $V_2^{(u)} = v_8$ where $u = 1$ and the following result is established.

Theorem 10: A balanced design will have P_1^V for $k = \binom{m}{2}$.

Table 6.6, Table 6.7, and Table 6.8 present the balanced designs for $m = 3, 4, 5$, respectively, which have the ability to identify and discriminate the models containing the general mean, main effects, and all two-factor interaction effects. Along with identifying the sets of treatments used in each design, the tables present the number of treatments, n, as well as the value of the constant $V_2^{(u)}$ (to four decimal places) where $u=1$. Here, each set of treatments, S_i, $i=1,...,m$, is replicated at most one time. Note that the complements for each of these designs will also have P_1^V by ***Theorem 3*** and thus are not included in the tables.

Table 6.6: Balanced designs with P_1^V for m=3 when k=$\binom{m}{2}$

n	S_0	S_1	S_2	S_3	$V_2^{(u)}$
7	•	•	•		0.2500
8	•	•	•	•	0.1250

Table 6.7: Balanced designs with P_1^V for m=4 when k=$\binom{m}{2}$

n	S_0	S_1	S_2	S_3	S_4	$V_2^{(u)}$
11	•	•	•			0.2500
11		•	•		•	0.1389
12	•	•	•		•	0.1250
14		•	•	•		0.1250
15	•	•	•	•		0.0750
16	•	•	•	•	•	0.0625

Table 6.8: Balanced Designs with P_1^V for m=5 when $k=\binom{m}{2}$

n	S_0	S_1	S_2	S_3	S_4	S_5	$V_2^{(u)}$
16	•	•	•				0.2500
16	•	•		•			0.0972
16	•		•		•		0.0625
16	•			•	•		0.1389
17	•	•	•			•	0.1352
17	•	•		•		•	0.0605
20		•	•		•		0.0601
21	•		•	•			0.0816
21	•	•	•		•		0.0527
21	•	•		•	•		0.0544
22	•		•	•		•	0.0694
22	•	•		•	•	•	0.0508
25		•	•	•			0.0812
26	•	•	•	•			0.0535
26	•		•	•	•		0.0430
27	•		•	•	•	•	0.0410
30		•	•	•	•		0.0375
31	•	•	•	•	•		0.0332
32	•	•	•	•	•	•	0.0313

It is also possible to have P_1^V for a design where the model contains the general mean, m main effects, and k two-factor interaction effects where $1 < k \le \binom{m}{2}$. This was observed in two very special fractional factorial designs from Table 3.3, namely, D17 and D18.

D17 is a balanced design for $m = 5$, consisting of the complete sets of treatments from S_1 and S_3, which has the ability to identify and discriminate the models containing the general mean, main effects, and k two-factor interaction effects with max $k = 9$. For each of the possible values of k, D17 has P_1^V with the values of $V_2^{(u)}$, $u = 1,...,\nu$, where $\nu = \begin{pmatrix} \begin{pmatrix} m \\ 2 \end{pmatrix} \\ k \end{pmatrix}$, as displayed in Table 6.9.

Table 6.9: D17 values of $V_2^{(u)}$ for k≤9

k	u	$V_2^{(u)}$
1	10	0.0694
2	45	0.0703
3	120	0.0714
4	210	0.0729
5	252	0.0750
6	210	0.0781
7	120	0.0833
8	45	0.0938
9	10	0.1250

Similarly, D18 is a balanced design for $m = 5$, consisting of the complete sets of treatments from S_1, S_3, and S_5, which has the ability to identify and discriminate the models containing the general mean, main effects, and k two-factor interaction effects with max $k = \begin{pmatrix} m \\ 2 \end{pmatrix} = 10$. It should be noted that D18 is a regular one-half orthogonal fraction of a 2^5 factorial experiment [Ghosh and Tian (2006)]. Moreover, as presented in Table 6.10, D18 has P_1^V with the values of $V_2^{(u)}$, $u = 1,...,\nu$, constant for all possible values of k.

Table 6.10: D18 values of $V_2^{(u)}$ for k≤10

k	u	$V_2^{(u)}$
1	10	0.0625
2	45	0.0625
3	120	0.0625
4	210	0.0625
5	252	0.0625
6	210	0.0625
7	120	0.0625
8	45	0.0625
9	10	0.0625
10	1	0.0625

6.5.2 $k = 1, g > 1$

The number of designs with the ability to unbiasedly estimate the general mean, m main effects, and $k = 1$ two-factor interaction effects which have P_g^V, $g > 1$, far outnumber such designs with P_1^V. With the exception of $m = 3$ as summarized and displayed in Table 6.1, this can be seen in the summary tables for $m = 4$ and $m = 5$ which are presented in Table 6.3 and Table 6.5, respectively. The number of designs with P_g^V, $g > 1$, is equal to the number of capable designs minus the number of designs with P_1^V. By definition of P_g^V, these designs will give g groups of values for the variances of the distinguishing two-factor interaction effect estimate. Thus, many of the characterizations from Chapter 5 which hold for designs with constant values, or $g = 1$, of the variances of the two-factor interaction estimates can be generalized for designs with $g > 1$.

For example, consider ***Theorem 1*** which states that a design will have P_1^V for $k=1$ iff $\underline{w}'_u W^{-1} \underline{w}_u$ is constant for all u. For designs with $g>1$, this theorem can be generalized to state that a design will have P_g^V for $k=1$ iff $\underline{w}'_u W^{-1} \underline{w}_u$ is constant for g groups of models where $u=1,...,\nu$. Similarly, ***Theorem 2*** states that a design will have P_1^V for $k=1$ iff $\left|X^{(u)\prime} X^{(u)}\right|$ is constant for all u. To characterize designs with P_g^V, $g>1$, this can be generalized to state that a design will have P_g^V for $k=1$ iff $\left|X^{(u)\prime} X^{(u)}\right|$ is constant for g groups of models where $u=1,...,\nu$.

Example 6.1

Consider D13.1 (also labeled Design A1 in Example 4.1) for $m=5$ which has the ability to identify and discriminate the models consisting of the general mean, main effects, and $k=1$ two-factor interaction effect. As can be seen from Table 4.1, this design has two groups of values for $V_2^{(u)}$, $u=1,...,10$. Specifically, $V_2^{(u)}=0.1003$ (rounded) for $u=1,8,9,10$ and $V_2^{(u)}=0.0958$ (rounded) for $u=2,...,7$. Therefore, D13.1 is said to have $P_2^V(6,4)$. To characterize this design as such, consider the calculations for $\underline{w}'_u W^{-1} \underline{w}_u$, $u=1,...,10$.

D13.1 gives $\underline{w}'_u W^{-1} \underline{w}_u$=1.0260 for $u=1,8,9,10$ and $\underline{w}'_u W^{-1} \underline{w}_u$= 0.5584 for $u=2,...,7$ where W^{-1} and $\underline{w}_u$, $u=1,...,10$ are calculated and presented on the following page.

$$W^{-1}=\left[\begin{array}{c|cc|ccc} a & b & b & c & c & c \\ \hline b & d & e & f & f & f \\ b & e & d & f & f & f \\ \hline c & f & f & g & h & h \\ c & f & f & h & g & h \\ c & f & f & h & h & g \end{array}\right] \text{ where } \begin{array}{ll} a=0.094 & e=-0.025 \\ b=0.008 & f=-0.005 \\ c=-0.007 & g=0.104 \\ d=0.100 & h=-0.021 \end{array}$$

$$\underline{w}_u=\begin{bmatrix} 3 \\ -1 \\ -1 \\ 1 \\ 1 \\ 1 \end{bmatrix},\ u=1,8,9,10 \qquad \underline{w}_u=\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \end{bmatrix},\ u=2,\dots,7$$

For this particular design, it is seen that $\underline{w}_u$ is identical for the first group of models $u=1,8,9,10$ and also identical for the remaining models, $u=2,\dots,7$, in group two, resulting in the two distinct values of $V_2^{(u)}$.

Alternatively, it is possible to characterize D13.1 as having $P_2^V(6,4)$ by considering $\left|X^{(u)\prime}X^{(u)}\right|$, $u=1,\dots,10$. The number of distinct values of $\left|X^{(u)\prime}X^{(u)}\right|$ will determine the value of g in P_g^V. The calculated values are presented in Table 6.11 and it is seen that D13.1 has $P_2^V(6,4)$.

Table 6.11: Determinants for D13.1 when $k = 1$

Model u	$V_2^{(u)}$	$\left\|X_1{}'X_1\right\|$	$\left\|X^{(u)\prime}X^{(u)}\right\|$
1,8,9,10	0.1003	1,261,568	12,582,912
2-7	0.0958	1,261,568	13,172,736

6.5.3 $k > 1, g > 1$

Consider a design which has the ability to identify and discriminate the models containing the general mean, m main effects, and $k > 1$ two-factor interaction effects. Except for designs with very special structures (as indicated in Section 6.5.1), the examples of this study indicate that designs will have P_g^V, $g > 1$, for $k > 1$ when the designs have P_1^V for $k-1$ two-factor interactions included in the model. Otherwise, P_g^V will not hold for any value of g, i.e., $\underline{V}_2^{(u)}$ will not consist of common entries for each of the models.

Example 6.2

Suppose D5 from Table 3.3 is being used to model the general mean, $m = 4$ main effects, and $k = 2$ two-factor interaction effects. There will be $v = \binom{6}{2} = 15$ models with the two-factor interaction effects in each model representing the distinguishing elements. The entries of $\underline{V}_2^{(u)}$ are calculated and presented in Table 6.12. It is seen that three of the models have $V_2^{(u)} = 0.4375$, $u = 5,8,10$, and the remaining twelve models have $V_2^{(u)} = 0.1167$. Thus, D5 has $P_2^V(12,3)$. Note from Table 4.2 that D5 has P_1^V when it is used to estimate the models consisting of the general mean, main effects, and $k = 1$ two-factor interaction effect. Therefore, by estimating an additional two-factor interaction effect in the model, D5 goes from having P_g^V with $g = 1$ to having P_g^V with $g = 2$. When modeling $k = 3$ two-factor interactions, D5 does not have P_g^V for any value of g.

Table 6.12: $\underline{V}_2^{(u)}$ entries for D5 when k=2

Model u	Two-factor Interactions	$\underline{V}_2^{(u)}$ Entries	
1	AB, AC	0.1167	0.1167
2	AB, AD	0.1167	0.1167
3	AB, BC	0.1167	0.1167
4	AB, BD	0.1167	0.1167
5	AB, CD	0.4375	0.4375
6	AC, AD	0.1167	0.1167
7	AC, BC	0.1167	0.1167
8	AC, BD	0.4375	0.4375
9	AC, CD	0.1167	0.1167
10	AD, BC	0.4375	0.4375
11	AD, BD	0.1167	0.1167
12	AD, CD	0.1167	0.1167
13	BC, BD	0.1167	0.1167
14	BC, CD	0.1167	0.1167
15	BD, CD	0.1167	0.1167

Chapter 7

Tables of Designs with P_1^V for m=3,4,5 and k=1

For the convenience of the reader and user, designs from Table 6.1, Table 6.3, and Table 6.5 having the property P_1^V for the various values of $V_2^{(u)}$ are presented in this chapter. For each combination of m, n, and $V_2^{(u)}$ a single design is presented. Note that the remaining designs can be constructed systematically, as described in Chapter 6. Recall, m represents the number of factors in a 2^m factorial experiment, n corresponds to the number of runs in the fractional factorial experiment, k is the number of two-factor interaction effects included in the model, ν is the number of models calculated as $\binom{m}{2}$, and $V_2^{(u)}$ is the value of the common variance for the design having property P_1^V.

Each run in the design is denoted by the position where the factor is observed at the high level as described in Section 3.4. For example, consider the run (–, +, –, –) in a 2^4 factorial experiment. This run is denoted as 2. The run (+, –, +, –, –) in a 2^5 factorial experiment is denoted as 13. Note that 0 denotes the run where all m factors in a 2^m factorial experiment are observed at their low level.

Table 7.1: Designs for m = 3 with k=1 having P_1^V

n	$V_2^{(u)}$	Runs
5	0.500	0, 12, 13, 23, 123
6	0.188	1, 2, 3, 12, 13, 23
6	0.250	0, 1, 2, 12, 13, 23
7	0.167	1, 2, 3, 12, 13, 23, 123
8	0.125	0, 1, 2, 3, 12, 13, 23, 123

Table 7.2: Designs for *m* = 4 with *k*=1 having P_1^V

n	$V_2^{(u)}$	Runs
6	0.875	0, 123, 124, 134, 234, 1234
7	0.188	0, 12, 13, 14, 23, 24, 34
8	0.125	1, 2, 3, 4, 123, 124, 134, 234
9	0.116	0, 1, 2, 3, 4, 123, 124, 134, 234
9	0.141	12, 13, 14, 23, 24, 34, 123, 124, 134
10	0.104	0, 1, 2, 3, 4, 123, 124, 134, 234, 1234
10	0.125	12, 13, 14, 23, 24, 34, 123, 124, 134, 234
11	0.100	1, 2, 3, 4, 12, 13, 14, 123, 124, 134, 234
11	0.109	0, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234
11	0.122	12, 13, 14, 23, 24, 34, 123, 124, 134, 234, 1234
12	0.094	0, 1, 2, 3, 4, 12, 13, 14, 123, 124, 134, 234
13	0.084	0, 1, 2, 3, 4, 12, 13, 14, 123, 124, 134, 234, 1234
14	0.073	1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234
15	0.069	0, 1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234
16	0.063	0, 1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234, 1234

Table 7.3: Designs for *m* = 5 with *k*=1 having P_1^V

n	$V_2^{(u)}$	Runs
7	0.625	0, 12, 1234, 1235, 1245, 1345, 2345
7	1.375	0, 1234, 1235, 1245, 1345, 2345, 12345
8	0.375	0, 1, 3, 25, 245, 1345, 1234, 12345
8	0.430	0, 12, 1234, 1235, 1245, 1345, 2345, 12345
8	0.500	1, 2, 12, 345, 1234, 1235, 1245, 12345
10	0.104	1, 2, 3, 4, 5, 1234, 1235, 1245, 1345, 2345
11	0.099	0, 1, 2, 3, 4, 5, 1234, 1235, 1245, 1345, 2345
12	0.097	5, 12, 13, 14, 23, 24, 34, 1234, 1235, 1245, 1345, 2345
15	0.069	12, 13, 14, 15, 23, 24, 25, 34, 35, 45, 1234, 1235, 1245, 1345, 2345
16	0.063	0, 12, 13, 14, 15, 23, 24, 25, 34, 35, 45, 1234, 1235, 1245, 1345, 2345

Chapter 8

Conclusions

Fractional factorial designs with n runs from factorial experiments with m factors at two levels each are studied. It is of particular interest to examine the class of designs with the ability to identify and discriminate the models consisting of the general mean, main effects, and k two-factor interaction effects, $1 \le k \le \binom{m}{2}$. In this class of models, it is generally assumed that only some of the two-factor interaction effects are important, i.e., $k < \binom{m}{2}$. However, at the design stage of the experiment it is unknown which two-factor interactions these will be. Because it is of interest to identify the best model for describing the data, all models in the class must be considered and therefore must be estimated. To estimate the parameters in these models, a design must be carefully chosen at the design stage of the experiment. It is possible to select the design based on some optimality criterion function such as AT, GT, AD, GD, AMCR, GMCR as discussed in Section 3.3. However, within these optimum designs it is observed that some of the designs have advantages over other designs when considering the precision of the estimates for the uncommon elements of the models (i.e., the k two-factor interaction effects) in the class of interest. In particular, some designs allow for equal precision in estimating all of the uncommon elements in the models for each of the models in the class, while other designs have a grouping effect for the precisions of the estimates of the uncommon parameters. Each of these scenarios should be considered at the design stage of the experiment. If there are no assumptions placed on the importance of the factors, then it is unknown which model is likely to describe the data best. In such a case, it is reasonable to require that all uncommon parameters in the models be estimated with equal precision. Alternatively, if the experimenter believes that some of the effects (and therefore some of the models) are more likely to accurately describe the data, it may be of interest to consider designs that will group the precisions of the estimates, where some of the precisions may be higher for some models than others. The experimenter can then choose a design with the grouping effect which gives higher precision to the most likely models. To consider this phenomenon, a new property is introduced which has the ability to classify designs according to the structure of the precisions for the estimates of the uncommon elements in the models of a particular class. The precision of the estimates are evaluated using the variances where the smaller the variance, the higher the precision.

The new property, denoted P_g^V and defined in Chapter 4, utilizes the variance-covariance matrix of the estimates for each model to classify the designs which are capable of estimating the k two-factor interaction effects with equal precision. Specifically, a design is said to have the property P_g^V where g is a positive integer, if there exist g groups of models so that the variances of k two-factor interactions for all models within a group are identical with each other. Following from this definition, we observe that a design has this property when it gives common entries of $\underline{V}_2^{(u)}$, $u = 1, 2, ..., v$, for g groups of models where $\underline{V}_2^{(u)}$ is a $(k \times 1)$ vector of the diagonal elements of $\sigma^{-2}\left(Var\left[\hat{\beta}_2^{(u)}\right]\right)$. The value of the identical entries of $\underline{V}_2^{(u)}$ is denoted as simply $V_2^{(u)}$. This property is written as $P_g^{\nu}\left(\nu_1,...,\nu_g\right)$, $\nu_1+...+\nu_g=\nu$ where ν_i is the number of models in Group i, $i=1,...,g$, and ν is the total number of models. There are $s=\binom{m}{2}$ possible two-factor interaction effects and $\nu=\binom{s}{k}$ models of interest.

During the design stage of the experiment, it is most likely that all models in the class will be given equal consideration. Therefore, it is important to consider designs having the property $P_1^{\nu}\left(\nu\right)$, simply denoted as P_1^V. Such designs will have equal precision for the uncommon elements in the models of interest and are studied in detail in Chapter 5 and Chapter 6. Moreover, it is possible to compare designs having P_1^V by considering the level of precision of the estimates for the uncommon parameters in the models; some designs may have better precision than other designs with the same number of treatments.

In Chapter 5, several characterizations of designs having P_1^V for $k=1$ are identified. Particularly, special types of designs such as balanced designs, isomorphic designs (under row and column permutations), and complementary designs are proven to have P_1^V for $k=1$. Most notably, there exist designs which are not of a special form which have P_1^V . These designs, D9.2 and D14 from Table 3.3, are studied in detail in Example 5.3 and Example 5.4, respectively, for $k=1$. These two examples demonstrate that it is possible to achieve P_1^V, even when the design is not balanced.

In addition, Chapter 6 presents the construction of all designs with the property P_1^V for $m=3$, $n=5,6,7,8$, $k=1$ and $m=4$, $n=6,...,16$, $k=1$. The number of designs with the property P_1^V are also presented for fractional factorial designs when $m=5$, $n=7,8,9$, $k=1$. For a 2^5 factorial experiment, fractional factorial designs consisting of $n=9$ treatments, requires the analysis of $\binom{2^5}{9}=\binom{32}{9}=28,048,800$ possible designs.

Chapter 6 also investigates special designs which have the property P_1^V for $\max k$ which is the largest value of k in the models of interest that the design has the ability to identify and discriminate. In the case when $k=\binom{m}{2}$ it is found that balanced designs will have P_1^V.

Finally, characterizations of P_1^V when $k=1$ are generalized for cases when $g>1$ and $k \geq 1$. Application of these characterizations are not as significant as in the case when $g=1$ since most designs which have the ability to unbiasedly estimate the models of interest will be of this kind. There are only a limited number of designs which will have P_1^V and thus this special case was examined in much greater detail.

References

Atkinson, A.C., Cox, D.R., (1974). Planning experiments for discriminating between models. *Journal of the Royal Statistical Society. Series B (Methodological)*. 36(3), 321-348.

Atkinson, A.C., Fedorov, V.V., 1975. Optimal design: experiments for discriminating between several models. *Biometrika*, 62(2), 289-303.

Biswas, A., Chaudhuri, P., 2002. An efficient design for model discrimination and parameter estimation in linear models. *Biometrika*, 89(3), 709-718.

Cotter, S.C., 1979. A screening design for factorial experiments with interactions. *Biometrika*, 66(2), 317-320.

Dette, H., Kwiecien, R., 2004. A comparison of sequential and non-sequential designs for discriminating between nested regression models. *Biometrika,* 91(1), 165-176.

Dey, A., Mukerjee, R., 1999. *Fractional Factorial Plans*. New York: John Wiley & Sons, Inc.

Draper, N. R., Smith, H. 1998. *Applied Regression Analysis* (3rd ed.). New York: John Wiley & Sons, Inc.

Federer, W.T., Raghavarao, D., 1975. On Augmented Designs. *Biometrics.* 31(1), 29-35.

Ghosh, S., Deng, H., Luan, L., 2007. Error sum of squares and power comparisons for model search, identification, and discrimination. *Journal of Statistical Planning and Inference*, 137, 3475-3482.

Ghosh, S., Teschmacher, L., 2002. Comparisons of search designs using search probabilities. *Journal of Statistical Planning and Inference*, 104, 439-458.

Ghosh, S., Tian, Y., 2006. Optimum two level fractional factorial plans for model identification and discrimination. *Journal of Multivariate Analysis*, 97, 1437-1450.

Hinkelmann, K., Kempthorne, O., 2005. *Design and Analysis of Experiments* (Vol. 2). New Jersey: John Wiley & Sons, Inc.

Hoaglin, D. C., Welsch, R. E., 1978. The hat matrix in regression and anova. *The American Statistician,* 32(1), 17-22.

Hocking, R.R., 2003. *Methods and Applications of Linear Models: Regression and the Analysis of Variance* (2nd ed.), New Jersey: John Wiley & Sons, Inc.

Ives, L.K., Jahanmir, S. , Gill, L., Filliben, J.J., 1998. High performance ceramics: effect of grinding on strength of sintered reaction bonded silicon nitride. *International Ceramic Review*, 47(3) 168-173.

Johnson, R. A., Wichern, D. W., 2002. *Applied Multivariate Statistical Analysis* (5th ed.). New Jersey: Prentice Hall.

Kiefer, J., 1959. Optimum experimental designs. *Journal of the Royal Statistical Society. Series B (Methodological)*, 21(2), 272-319.

Li, W., Nachtsheim, C.J., 2000. Model-robust factorial designs. *Technometrics,* 42(4), 345-352.

Montgomery, D.C., 2001. *Design and Analysis of Experiments* (5th ed.). New York: John Wiley & Sons, Inc.

Montgomery, D. C., Peck, E. A., Vining, G. G., 2001. *Introduction to Linear Regression Analysis* (3rd ed.). New York: John Wiley & Sons, Inc.

Poole, M. A., O'Farrell, P. N., 1971. The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*, 52, 145-158.

Pukelsheim, F., Rosenberger, J.L., 1993. Experimental designs for model discrimination. *Journal of the American Statistical Association,* 88(422), 642-649.

Rao, C. R., 2002. *Linear Statistical Inference and Its Applications* (2nd ed.). New York: John Wiley & Sons, Inc.

Scheffe, H., 1959. *The Analysis of Variance*. New York: John Wiley & Sons, Inc.

Shirakura, T., 1976. Optimal balanced fractional 2^m factorial designs of resolution VII, $6 \le m \le 8$. *The Annals of Statistics*, 4(3), 515-531.

Srivastava, J.N., 1970. Optimal balanced 2^m fractional factorial designs, In: Essays in Probability and Statistics (R.C. Bose et al., Eds), Univ. of North Carolina Press, Chapel Hill and Indian Statistical Institute, 689-706.

Srivastava, J.N., Chopra, D.V., 1971. Balanced optimal 2^m fractional factorial designs of resolution V, $m \le 6$. *Technometrics*, 13, 257-269.

Taubman, S.B., Addelman, S., 1978. Some two-level multiple-response factorial plans. *Journal of the American Statistical Association,* 73(363), 607-612.