INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

Some New Types of Designs with Applications in Conjoint Analysis

Yuan Ding

A Thesis
in
Special Individualized Programme

Presented in Partial Fulfilment of the requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

July 1997

Contents

Abstract

Some New Types of Designs with Applications in Conjoint Analysis

Yuan Ding, Ph.D.
Concordia University, 1997

Conjoint Analysis (CA) is a mainstream method in market research studying the consumer behaviour. This method entails choice of experimental designs often dealing with large numbers of attributes and attribute levels. This thesis introduces the use of Partially Augmented Design (PAD) in CA and develops the universal optimality of two specific PAD's for use with large numbers of attribute levels. An example is presented to show the efficiency of the universally optimal PAD.

Uniform design, based on the work of Fang and Wang (1981), is introduced as an alternative to PAD in CA. A theorem previously developed by the author concerning the restriction in use of uniform design is presented. Optimal uniform designs are proposed. These designs are not orthogonal and hence they are also useful in the context of CA, because substantial correlations may exist between attributes.

The next important step in CA is proper statistical analysis. Appropriateness of random and mixed effects models, which are different from traditional fixed effects models used for analyzing CA data, is highlighted in several examples.

ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Y. P. Chaubey, Dr. J. Tomberlin and Dr. D. Morin for all their guidance, encouragement and support in the preparation of this dissertation. I also thank the other two Ph.D's in my family, my husband, Weiming, and my daughter, Vivian Ding, for their encouragement and support, without which I would not have completed this research.

Chapter 1

Introduction and Overview

1.1 An Introduction to Conjoint Analysis and Review

Since the mid-1970's, probably no technique for modelling consumer preferences has received more attention than conjoint analysis (Steckel, DeSarbs and Mahajan, 1991; Gattin and Wittink, 1982; Green and Srinivasan, 1978). Areas of its applications cover products and services, as well as consumer, industrial and institutional markets. Currently, conjoint analysis and the related technique of experimental choice analysis represent the most widely applied methodologies for measuring and analyzing consumer preferences (Carroll and Green, 1995).

Since its introduction, conjoint analysis has been one of the fastest-growing marketing research techniques. Today, a variety of paradigms, methods, and analytical techniques are commonly used in academic and commercial applications of conjoint analysis. Cattin and Wittink (1982) did a study of commercial usage of conjoint analysis during its first decade and found approximately 600 commercial applications, of which more than 20 percent

occurred in the last year of the study. During the 1981-1985 period, they estimate that about 400 applications were performed annually (Wittink and Cattin, 1989).

Marketing research, similar to the business disciplines in general, has been a long time borrower of models, tools, and techniques from other sciences. Economists, statisticians, and operations researchers have made significant contributions to marketing, particularly in prescriptive model building. Over the past 30 years, psychometricians and mathematical psychologists have also provided a bounty of research riches in measurement and data analysis techniques (Carroll and Green, 1995). The seminal theoretical contribution to conjoint analysis was made by Luce, a mathematical psychologist, and Tukey, a statistician (Luce and Tukey, 1964). Early psychometric contributions to non-metric conjoint analysis were made by Kruskal (1965), Roskam (1968), Carroll (1969, 1973), and Young (1972). The evolution of conjoint analysis in marketing research and practice has been extensively documented in reviews by Green and Srinivasan (1978, 1990), Wittink and Cattin (1989), and Wittink, Vriens, and Burhenne (1994). In addition, Green and Krieger (1993) have surveyed conjoint methodology from the standpoint of new product design and optimization.

1.1.1 What is Conjoint Analysis (CA)?

Conjoint analysis is a multivariate technique used to understand how respondents develop preferences for products or services. It is based on the simple premise that consumers assess the value or utility of a product/service/idea (real or hypothetical) by combining the separate amounts of utility provided by each attribute. It is unique among multivariate methods in that the

researcher first constructs a set of hypothetical products or services by combining the possible attributes at various levels. These hypothetical products are then presented to respondents who provide only their overall evaluations of the hypothetical products/services. By constructing specific combinations (treatments or stimuli), the analyst is attempting to understand a respondent's preference structure. This information is used to determine which concepts of different segments would like best and whether a new concept is preferred over products that are currently on the market. It is also used to determine the importance of various attributes, including price. Conjoint analysis is closely related to traditional experimentation (Hair, Anderson, Tatham and Black, 1992). It is actually a "family" of techniques and methods, all theoretically based on the models of information integration and functional measurement (Louviere, 1988). Green and Srinivasan (1978) use the term *conjoint analysis* (CA) to refer to any decompositional method that estimates the structure of a consumer's preferences, given his/her overall evaluation of a set of alternatives that are prespecified in terms of levels of different attributes.

From a marketing management standpoint one is interested in the normative problem, namely, to find the point location that optimizes some objective function. The optimal product design problem requires procedures for not only estimating the value of the objective function for each point location of interest (the prediction problem) but also for searching the space systematically to find the specific location that results in the highest profits, revenue, share of choices, or whatever is being optimized. The first approach to *Adaptation of CA* (ACA) based methods to optimal product design was proposed by Zufryden (1977). His model assumed that the consumers' preference could be rated by an objective function or a utility function (Green et al, 1981)

such as following.

- total number of test brand first choices, each weighted by the frequency of product purchase
- total number of respondents not choosing a specific competitive product, conditioned on the test product
- total expected profits or total expected cash flow, in which both (subjective) purchase probability and conditional payoff are measured

That means, researchers could use a particular utility measure as a response variable in CA.

1.1.2 Steps and Assumptions Involved in Conjoint Analysis

The traditional method of CA is called the *full-profile method* (Green and Wind, 1975; Louviere, 1988; Wittink and Cattin, 1989). In this method, the consumer responds to complete descriptions (full profiles) of hypothetical objects.

For more detail see the following Table 1.1 from Green and Srinivasan (1990) which gives a general framework of the different steps in CA and the alternative methods of implementing each step.

Table 1.1 Steps Involved in Conjoint Analysis

	Step	Alternative Methods
1	Preference model	Vector model, ideal point model, part-worth function model, mixed model
2	Data collection method	Full profile, two-attributes-at-a-time (trade-off tables)
3	Stimulus set	Fractional factorial design, random sampling from constructing a multivariate distribution, pareto-optimal design
4	Stimulus presentation	verbal description (multiple-cue stimulus card), paragraph description, pictorial or three-dimensional model representations, physical products
5	Measurement scale for the dependent variable	Rating scale, rank order, paired comparisons, constant-sum paired comparisons, graded paired comparisons, category assignment
6	Estimation method	Metric methods (multiple regression); non-metric methods (LINMAP, MONANOVA, PREFMAP, Johnson's non-metric algorithm); choice-probability -based methods (logit, probit).

These related procedures are based on a number of assumptions (Clarke, 1987):

- It must make sense to view the product as a bundle of attribute levels– the product can be decomposed into separate features for which utilities

can be computed.

- The utility of the product is some simple function of the utilities of the product's attribute levels.
- A respondent will buy the product for which he/she has the highest utility. It isn't necessary to assume that this happens at every purchase occasion. It may just be a purchase probability, but there is an assumption that utilities are translated into real purchase behaviour.
- The attributes in the questionnaire are relevant as well as they adequately describe the product.
- There isn't any redundancy; that is, nothing gets counted twice.
- We must assume that, in real product-buying decisions, people act the way they behave in the interview.

These assumptions may never be completely valid in any situation, but they are similar to the assumptions made in a large number of marketing research studies, and they are reasonable approximations in many situations.

Montgomery (1985) presents some evidence of the ability of conjoint analysis to predict market shares. The results of Davidson's (1973) study of traffic between Montreal and downtown Ottawa and Robinson's (1980) study of the North Atlantic air travel market show that their market predictions are quite good (Table 1.2).

Table 1.2 Evidence of Predictive Validity of Conjoint Analysis

Forecast of Traffic	**Montreal to Downtown Ottawa**	
Transportation Mode	Actual Share, %	Conjoint Prediction, %
Air	3.2	3.2
Train	10.9	10.4
Bus	15.1	15.5
Car	70.7	70.7
North Atlantic	**Air Travel**	
Fare Type	Actual Share, %	Conjoint Prediction, %
Economy	25	23
Excursion	14	12
(14-21 day) Excursion	43	49
(22-45 day) GP Fares	18	16

1.2 Importance and Limitations of Existing Designs in CA

The application of formal experimental designs to marketing problems has increased markedly over the last three decades. Almost every marketing research text published during this period has devoted at least a chapter to the subject. Moreover, complete books on the topic have been written for the marketing research community.

In CA, the use of experimental design in the analysis of consumer decisions has two objectives (Hair et al, 1992):

(1) To determine the contribution of predictor variables and their respective values to the determination of consumer preferences.

(2) To establish a valid model of consumer judgements that is useful in predicting the consumer acceptance of any combination of attributes, even those not originally evaluated by consumers.

As is well known, full factorial designs provide maximum information per observation and allow the researcher to estimate interactions as well as main (or single factor) effects (Green, Carroll and Carmone, 1978). However, it is not always economical or possible to collect data on all of the experimental combinations that would be implied by a full factorial design. Therefore, statisticians have developed classes of designs called *fractional factorials* where only some of the experimental combinations are constructed.

In some product categories, consumers may know enough to make rational decisions, but not enough about individual brands to rate them on a number of attributes. For example, lots of consumers can make intelligent trade-offs between size, price, energy efficiency, and freezer space when choosing a refrigerator, but they would not be able to rate current models on these same attributes. The number of brands in a product class that a respondent may be familiar with usually is small. Furthermore, real brands and services are usually not distinctive enough to provide reliable estimates of parameters. For these reasons, CA is usually done with hypothetical stimulus descriptions. The most popular method of creating such descriptions is by using an orthogonal design, which chooses a good subset of all possible experiments or a complete set. As usual, a good experimental design should satisfy the following principles:

1. Uniformity: The experimental points should be scattered uniformly on the domain for experimentation such that these points are good representations.

2. Regularity: The experimental points satisfy some regularity condition so that it is convenient to do the analysis of variance for the experimental data.

The orthogonal design is produced according to these two principles. Due to its orthogonality, the orthogonal design yields a diagonal information matrix and uncorrelated estimates of main effects and interactions. Therefore the orthogonal design has been successfully applied in various fields, including CA. On the other hand, some limitations of orthogonal design in CA have been noted, such as the following.

(1) If there is a substantial amount of environmental correlation between some of the attributes (e.g., in the case of automobiles, acceleration, top speed, and size of engine are all positively correlated, and each is negatively related to gas mileage), an orthogonal design can produce some stimulus profiles that are not representive of the subject's environment (Steckel, DeSarbo and Mahajan, 1991). Under orthogonality, factor independence is to be sought wherever possible.

When some of the profiles turn out to be infeasible, various solutions can be implemented (Moore and Holbrook, 1990). First, one might permute the attribute levels to generate another orthogonal design. However, this solution works only when there is a very small number of combinations that are much less realistic than others. Second, one may merely remove the stimulus

that contain the infeasible attribute combinations. However, treating infeasible attribute combinations as missing observations is viewed as somewhat "artificial" and "hypocritical". Third, one can prepare a composite factor (super-attribute) for a number of highly correlated attributes, but it is no longer possible to separate the effects of the subfactors contained in the composite. Fourth, one might introduce designs that have lower environmental correlations, trading off some statistical efficiency for increased realism. This could be achieved by employing a random sampling procedure from a multivariate distribution.

(2)The early successes of conjoint analysis have led to industry demands for techniques that handle ever larger numbers of attributes and attribute levels. In addition, the number of levels may differ from factor to factor. Thus, too many experiments are required for an orthogonal design to satisfy the principles given above. This need, in turn, has prompted the development of data collection methods and models, such as, experimental choice analysis (Carson et al. 1994; Batsell and Louviere, 1991); adaptive conjoint analysis (ACA) (Zufryden, 1977; Johnson, 1987); Green's hybrid models (Green, 1984) and hierarchical Bayesian methods in CA (Allenby, Arora, and Ginter, 1995; Lenk and colleagues, 1994), etc. What appears to be lacking is constructing/introducing appropriate designs and convincing evidence of whether the newer conjoint methods for coping with larger numbers of attributes and levels are markedly superior to the older approaches.

1.3 Outline of the Thesis

This thesis deals with the problem of designing appropriate experiments for conjoint analysis. It has been highlighted in the previous section, choice of a design plays a very important role in conjoint analysis. There are indeed other concerns and issues in this important area of market research, the author is primarily concerned with the problem of design of experiment. Since, some of the issues mentioned here with respect to *conjoint analysis* may also prevail in other disciplines, the designs discussed here may also be applicable in other areas.

Because of the limitations of existing designs, two new classes of designs, *partially augmented designs* (PAD) and *uniform designs* (UD), are introduced for conjoint analysis in chapters 2 and 3, respectively. Within these two classes of designs, special designs have been proven by this author to be statistically optimal according to certain criteria. Universally optimal PADs and optimal UDs could be used to provide more flexible choices in the numbers of factors and levels with some optimal statistical properties in estimation. They may provide efficient achievements in the presence of environmental correlation. In analysis of marketing research data, they will give more accurate results in a wide range while requiring only a small number of stimuli. In addition, except for generating these new classes of stimuli, no new computer software is needed.

In section 2.1, the class of partially augmented designs is presented as a class of non-orthogonal designs. A new model related to PAD with some restrictions is then established. Section 2.2 presents a strong optimality criterion called *universal optimality*. Two useful types of PADs with universal

optimality under some conditions are proposed by the author in section 2.3. Detailed proofs are also developed. In Section 2.4 the author introduces these PADs into CA and illustrates their applications to CA. To assess the efficiency of the newly developed universally optimal PAD in CA a study of a real problem is conducted. The results of the analysis show that the new classes of designs have interesting potential power in CA.

Section 3.1 gives an alternative model, a regression model using uniform design. Section 3.2 introduces the construction of *uniform designs* based on "good lattice point sets". A theorem previously developed by the author concerning the restriction in the use of uniform designs is presented. Some designs with optimalities or uniformity are proposed in appendix A. The relationship between uniform design and orthogonal design is also presented in this section. To illustrate the application of uniform designs in CA, a uniform design for varying numbers of levels generated by a pseudo-level technique is used to creating stimulus descriptions for banking services.

In chapter 4, further work concerning the mixed effects model is discussed. Several examples show the potential research interests.

Chapter 2

Partially Augmented Designs with Universal Optimality

2.1 Introduction and Model

In this chapter, we introduce a class of designs called *partially augmented design* (PAD) proposed by Taguchi (1987). Using a different method of analysis, we establish a model with some restrictions. Under the model two special kinds of PADs with universal optimality are found in Theorem 2.4 and 2.5. Applying a universally optimal PAD to a real data set, the results of analysis show that this new class of designs has interesting potential applications to CA.

Now consider a factorial design with n factors. The j-th factor has s_j levels ($j = 1, ..., n$). We assume the expected value of an observation taken on the i_1-th level of the first factor, the i_2-th level of the second factor, ... , and the i_n-th level of the n-th factor is specified by

$$E(y_{i_1 i_2 ... i_n}) = \theta_{i_1}^{(1)} + ... + \theta_{i_{n-1}}^{(n-1)} + \theta_{i_n}^{(n)}, \tag{2.1}$$

where $\theta_{i_1}^{(1)}, ..., \theta_{i_{n-1}}^{(n-1)}$, and $\theta_{i_n}^{(n)}$ are unknown constants, $1 \leq i_j \leq s_j$ for $1 \leq j \leq n$, and all observations are uncorrelated with common variance. A design with N observations is a selection of N combinations of the levels of the n factors. We write $\mathcal{D}_N$ as the collection of all such designs with N observations. We restrict the number of levels in the first $n-1$ factors to the same s when discussing augmented designs, i.e., we assume that $s_1 = ... = s_{n-1} = s$ and $s_n = p$, where p may differ from s. We denote the design as $\tilde{\tilde{d}} \in \mathcal{D}_N$. We can write the model in matrix form as follows:

$$\begin{cases} E(\tilde{\tilde{Y}}) & = (X_1, X_2, \cdots, X_{n-1}, X_n)\tilde{\tilde{\theta}} \triangleq \tilde{\tilde{X}}\tilde{\tilde{\theta}} \\ Var(\tilde{\tilde{Y}}) & = \sigma^2 I_N, \end{cases} \tag{2.2}$$

where $\tilde{\tilde{Y}}$ is $N \times 1$ observation vector obtained under design $\tilde{\tilde{d}}$; X_j is an indicator matrix of the j-th $(j = 1, ..., n)$ factor specified by design $\tilde{\tilde{d}}$, i.e., $X_j = (x_{\alpha l}^{(j)})_{N \times s} (j = 1, ..., n-1)$ and $X_j = (x_{\alpha l}^{(j)})_{N \times p} (j = n)$ for

$$x_{\alpha l}^{(j)} = \begin{cases} 1, & l\text{-th level of the } j\text{-th factor appears in the } \alpha\text{-th experiment;} \\ 0, & l\text{-th level of the } j\text{-th factor does not appear in the } \alpha\text{-th experiment.} \end{cases}$$

Therefore, when $j = 1, ..., n-1, X_j \mathbf{1}_s = \mathbf{1}_N$, where $\mathbf{1}_s = (\underbrace{1, ..., 1}_{s})'$, and $X_n \mathbf{1}_p = \mathbf{1}_N$, $\tilde{\tilde{X}}$ is called the design matrix, $\tilde{\tilde{\theta}} = (\theta_1', \theta_2', ..., \tilde{\tilde{\theta}}_n')'$ (prime denotes transpose) is a $[(n-1)s + p]$ parameter vector with subvectors $\theta_j = (\theta_1^{(j)}, ..., \theta_s^{(j)})' (j = 1, ..., n-1)$ and $\tilde{\tilde{\theta}}_n = (\theta_1^{(n)}, ..., \theta_p^{(n)})'$, and I_N denotes the identity matrix of order N.

Assuming additional $k (1 \leq k < p)$ levels of the n-th factor, which will also be investigated, but there is no suitable orthogonal array for this type of design. We may keep the N combinations, and augment N_1 new combinations which include the k additional levels of the n-th factor. Then, the new design $\tilde{d}$ is a selection of $N + N_1$ combinations of the adjusted levels of

the n factors. Let $\mathcal{D}_{N+N_1}$ be the collection of all such designs with $N+N_1$ observations. We call such a $\tilde{d} \in \mathcal{D}_{N+N_1}$ an *augmented design.* The new design matrix is

$$\tilde{X} = \begin{pmatrix} X_1 & X_2 & \cdots & X_{n-1} & X_n & 0 \\ Z_1 & Z_2 & \cdots & Z_{n-1} & Z_n & Z_n^{(k)} \end{pmatrix},$$

where Z_j is a $N_1 \times s$ indicator matrix of the j-th ($j = 1, ..., n-1$) factor and $(Z_n | Z_n^{(k)})$ is a $N_1 \times (p+k)$ indicator matrix of the n-th factor in the N_1 augmented combinations. The corresponding new model is given by

$$\begin{cases} E(\tilde{Y}) & = \tilde{X}\tilde{\theta}, \\ Var(\tilde{Y}) & = \sigma^2 I_{N+N_1}, \end{cases} \tag{2.3}$$

where $\tilde{Y} = (\tilde{\tilde{Y}}' | Y_a')'$ is a $N+N_1$ observation vector including a N_1 augmented observation subvector Y_a and $\tilde{\theta} = (\tilde{\tilde{\theta}}', \theta_{p+1}^{(n)}, ..., \theta_{p+k}^{(n)})' \triangleq (\theta_1', ..., \theta_{n-1}', \theta_n')'$ where $\theta_n = (\theta_1^{(n)}, ..., \theta_p^{(n)}, \theta_{p+1}^{(n)}, ..., \theta_{p+k}^{(n)})'$. If we write $W_i = \begin{pmatrix} X_i \\ Z_i \end{pmatrix}$ for $i = 1, ..., (n-1)$ and $W_n = \begin{pmatrix} X_n & 0 \\ Z_n & Z_n^{(k)} \end{pmatrix}$, then, for the augmented design $\tilde{d} \in \mathcal{D}_{N+N_1}$, the design matrix is $\tilde{X} = (W_1, ..., W_n)$, and the coefficient matrix of the normal equations, i.e., "information matrix" of $\tilde{d}$ for estimating $\tilde{\theta}$ is

$$\begin{aligned} \tilde{X}'\tilde{X} &= \begin{pmatrix} W_1'W_1 & W_1'W_2 & \cdots & W_1'W_n \\ \vdots & \vdots & & \vdots \\ W_n'W_1 & W_n'W_2 & \cdots & W_n'W_n \end{pmatrix} \\ &= \begin{pmatrix} (W_1...W_{n-1})'(W_1...W_{n-1}) & (W_1...W_{n-1})'W_n \\ W_n'(W_1...W_{n-1}) & W_n'W_n \end{pmatrix}. \end{aligned} \tag{2.4}$$

To estimate some linear functions of a subvector of θ in model (2.3), particularly the contrast of level effects of any factor, $\Omega(\theta) = \bigcup_{j=1}^n \Omega(\theta_j) = \bigcup_{j=1}^n \{a_j'\theta_j : a_j'\mathbf{1} = 0\}$ is the set of the parameters which need to be estimated.

If we are interested in estimating $a_n'\theta_n$, by Gauss-Markov Theorem the b.l.u.e. of $a_n'\theta_n$ is $a_n'\hat{\theta}_n$ with

$$\begin{aligned} Var(a_n'\hat{\theta}_n) &= \sigma^2(0\vdots a_n')(\tilde{X}'\tilde{X})^- \begin{pmatrix} 0 \\ \cdots \\ a_n \end{pmatrix} \\ &= \sigma^2 a_n'(W_n' P_{(W_1 \ldots W_{n-1})^\perp} W_n)^- a_n \\ &\triangleq \sigma^2 a_n' C_n^-(\tilde{X}) a_n, \end{aligned}$$

where the $(0\vdots a_n')$ means $(\underbrace{0, \ldots, 0}_{(n-1)s}, a_n')$ and the orthogonal projection matrix

$$\begin{aligned} P_{(W_1 \ldots W_{n-1})^\perp} &= I - P_{(W_1 \ldots W_{n-1})} \\ &= I - (W_1 \ldots W_{n-1})[(W_1 \ldots W_{n-1})'(W_1 \ldots W_{n-1})]^-(W_1 \ldots W_{n-1})'. \end{aligned}$$

The coefficient matrix of the reduced normal equations, i.e., the "information matrix" of $\tilde{d}$ for θ_n

$$C_n(\tilde{X}) = W_n' P_{(W_1 \ldots W_{n-1})^\perp} W_n \tag{2.5}$$

is called the $C_{\tilde{d}}$-matrix of the design $\tilde{d}$ for estimating θ_n. $C_n^-(\tilde{X})$ is any generalized inverse of $C_n(\tilde{X})$.

It can be shown that

any $a_n'\theta_n$ in the $\Omega(\theta_n)$ is estimable $\Longleftrightarrow rkC_n(\tilde{X}) = (p+k) - 1$ for the $(p+k) \times (p+k)$ C-matrix $C_n(\tilde{X})$.

For this reason, we will discuss the optimalities only for the designs whose C-matrix $C_n(\tilde{X})$ have rank of $(p+k) - 1$.

2.2 Optimalities

In this section we use theorems, definitions and related interpretations in a paper of Hedayat (1981) to summarize the concepts and results regarding some optimalities.

We perform experiments mainly to efficiently estimate or test hypotheses about some specified unknown parameters of a given model. Different considerations lead to different criteria for the choice of "best" design. Here we limit our search to discrete designs which are supported on sets consisting of finite number of points.

In general, let d be a design and let Y be the vector of observations obtained under d, and assume

$$E(Y) = X\underline{\theta} = (X_1 \vdots X_2) \begin{pmatrix} \underline{\theta_1} \\ \cdots \\ \underline{\theta_2} \end{pmatrix}, Var(Y) = \sigma^2 I,$$

where $\underline{\theta_1}$ is a $v \times 1$ vector, its information matrix under d and above model is $C_d = X_1'X_1 - X_1'X_2(X_2'X_2)^{-}X_2X_1'$.

Suppose that

$\mathcal{B}_v$ = the class of all $v \times v$ nonnegative definite matrices,
$\mathcal{B}_{v,0}$ = $\{C : C \in \mathcal{B}_v;\ C\mathbf{1} = 0 \text{ and } C'\mathbf{1} = 0\}$,
$\mathcal{D}$ = the class of designs d under consideration,
$\mathcal{C}$ = $\{C_d, d \in \mathcal{D}\}$,

Also, let

$$\mu_{d1} \geq \mu_{d2} \geq \cdots \geq \mu_{dv}$$

be the eigenvalues of C_d. Note if $\mathcal{C} \subseteq \mathcal{B}_{v,0}$, $\mu_{dv} = 0$ for all $d \in \mathcal{D}$. If necessary, we let ξ denote an approximate design (a probability measure on the

experimental space) and C_ξ be the associated information matrix.

1. Some Well-Known Optimality Criteria

Assume $\mathcal{C} \subseteq \mathcal{B}_v$.

Definition 2.1 A design $\xi^* \in \mathcal{D}$ is *G-optimal* if and only if

$$min_{\xi \in \mathcal{D}} max_{X \in \mathcal{X}} Var_\xi \hat{EY}_X = max_{X \in \mathcal{X}} Var_{\xi^*} \hat{EY}_X,$$

where $\hat{EY}_X$ is the b.l.u.e. of EY_X and $\mathcal{X}$ is the experimental space. Kiefer called it *G-optimal* (for global or minmax), since we are minimizing the maximum variance of any predicted value over the experimental space.

Definition 2.2 A design $d^* \in \mathcal{D}$ is D-*optimal* if and only if C_{d^*} is non-singular and

$$min_{d \in \mathcal{D}} det(C_d^{-1}) = det(C_{d^*}^{-1}).$$

Here, "D-" stands for determinant.

This criterion has many appealing properties, such as:

(i) Under normality, if d^* is D-optimal, d^* minimizes:

a) The volume (or expected volume, if σ^2 is unknown, and the rank of C_d is invariant under d) of the smallest invariant confidence region on $\theta_1, ..., \theta_v$ for any given confidence coefficient.

b) The generalized variance of the estimators of parameters (Suppose $X = (X_1, X_2, ..., X_n)'$ is distributed as multivariate $N(\mu, V)$. The de-

terminant of V is called the generalized variance of X as defined by Wilks (1932)).

(ii) In the class of approximate designs, D-optimality $\Longleftrightarrow$ G-optimality whenever $\underline{\theta_1} = \underline{\theta}$.

(iii) The design remains D-optimal if one changes the scale of the parameters: Let $\theta_1', ..., \theta_v'$ be related to $\theta_1, ..., \theta_v$ by a non-singular linear transformation. If d^* is D-optimal for $\theta_1, ..., \theta_v$, then d^* is also D-optimal for $\theta_1', ..., \theta_v'$. The analogue for other criteria is false in even the simplest settings.

Definition 2.3 A design $d^* \in \mathcal{D}$ is A-*optimal* if and only if C_{d^*} is non-singular and

$$min_{d \in \mathcal{D}} Tr(C_d^{-1}) = Tr(C_{d^*}^{-1}).$$

"A-" stands for average.

If d^* is A-optimal, it minimizes the average variance of $\hat{\theta}_1, ..., \hat{\theta}_v$.

It seems natural to specify some optimality function ϕ on C, then, the problem is to find d to minimize $\phi(C_d)$. We call ϕ as an optimality criterion. Then the above criteria are as follows:

$$\begin{aligned} \text{A-optimality} : \phi_A(C_d) &= \textstyle\sum_{i=1}^{t} \mu_{di}^{-1}, \\ \text{D-optimality} : \phi_D(C_d) &= \textstyle\prod_{i=1}^{t} \mu_{di}^{-1}, \\ \text{E-optimality} : \phi_E(C_d) &= \mu_{dt}^{-1}, \end{aligned}$$

where t is the number of non-zero eigenvalues of C_d.

2. Universal Optimality

In Kiefer (1975), a strong optimality criterion *universal optimality* was considered, where assuming $\mathcal{C} \subseteq \mathcal{B}_{v,0}$.

Definition 2.4 We say $d^* \in \mathcal{D}$ is a universally optimal design, if d^* minimizes $\phi(C_d), d \in \mathcal{D}$ for every function $\phi : \mathcal{B}_{v,0} \to (-\infty, +\infty]$ satisfying:

(1) ϕ is convex,

(2) $\phi(bC)$ is nonincreasing in the scalar $b \geq 0$ for each $C \in \mathcal{B}_{v,0}$,

(3) ϕ is invariant under each permutation of rows and (the same on) columns.

Definition 2.5 A matrix M is called a *completely symmetric* (c.s.) matrix if $M = aI_v + bJ_v$, where a, b are scalars, I_v is the identity matrix and J_v is the $v \times v$ matrix consisting of all 1's.

The following theorem given by Kiefer (1975) is a simple tool in determining a universally optimal design.

Theorem 2.1 *Suppose a class* $\mathcal{C} = \{C_d : d \in \mathcal{D}\}$ *of matrices in* $\mathcal{B}_{v,0}$ *contains a* C_{d^*} *for which*

$$\begin{aligned} &(a)\ C_{d^*} \quad is \quad c.s. \\ &(b)\ trC_{d^*} \quad = \quad max_{d \in \mathcal{D}} trC_d \end{aligned} \tag{2.6}$$

Then d^* *is universally optimal in* $\mathcal{D}$. *(Since* $-tr(C)$ *satisfies the conditions*

in definition 2.4, it follows that (2.6)(b) is necessary for universal optimality).

Cheng refined Kiefer's criterion and defined a larger class of optimality criteria called *optimality criteria of type* $i(i = 1, 2)$ and *generalized optimality criteria of type* $i(i = 1, 2)$ that include A-, D- and E-criteria, and more. Referring to Hedayat (1981), if there is a universally optimal design over $\mathcal{D}$, then it is optimal with respect to a very general class of criteria including Cheng's criteria of type $i(i = 1, 2)$ and generalized criteria of type $i(i = 1, 2)$, etc. Obviously, if there is a universally optimal design over $\mathcal{D}$, then it is also a A-, E- and D-optimal design.

2.3 Partially Augmented Designs (PAD) with Universal Optimality

2.3.1 Orthogonal Array and its Universal Optimality

We define a *rectangular array* denoted by $(N, n, s_1 \times ... \times s_n)$ as an $N \times n$ matrix with entries in the i-th column from a set S_i of s_i elements, $1 \leq i \leq n$. A $(N, n, s_1 \times ... \times s_n)$ array is said to be an *orthogonal array (with variable numbers of symbols) of strength* d if for any selection of d columns, say the α_1-th,...,α_d-th, the number of times that the row vector $(i_1, ..., i_d)'(i_j \in S_{\alpha_j}, j = 1, ..., d)$ occurs in the $N \times d$ submatrix specified by the d selected columns is constant for all combinations $i_1 \in S_{\alpha_1}, ..., i_d \in S_{\alpha_d}$. The constant may, however, depend on the set of selected columns. We denote such an orthogonal array by $\mathrm{OA}(N; s_1, s_2..., s_n; d)$. When $s_1 = s_2 = ... = s_n = s$, we also write it as $\mathrm{OA}(N, n, s, d)$ and call it a symmetric orthogonal array. We will restrict ourselves to orthogonal array of strength 2 (i.e., orthogonal main effect plans) in this chapter, and denote it as $\mathrm{OA}(N; s_1, ..., s_n; 2)$, or

$L_N(s_1^{k_1}...s_r^{k_r})$ which has k_i columns with s_i levels, $n = k_1 + ... + k_r$ and s_i being unequal, if $r > 1$, the array is said to be *mixed* or to have mixed levels. Cheng (1980) gives following results for OA$(N; s_1, s_2, ..., s_n; 2)$:

Theorem 2.2 *If there is an* $OA(N; s_1, s_2, ..., s_n; 2)$ *under the model specified by (2.1), then it defines a factorial design* d^* *with* N *observations which minimizes* $\sum_{i=1}^{n} \phi_i(C_{di})$ *over* $\mathcal{D}_N$ *for all functions* $\phi_i : \mathcal{B}_{s_i,0} \to (-\infty, +\infty]$ *satisfying*

(1) ϕ_i *is convex,*

(2) for any fixed $C \in \mathcal{B}_{s_i,0}$, $\phi_i(bC)$ *is nonincreasing in the scalar* $b \geq 0$,

(3) ϕ_i *is invariant under each simultaneous permutation of rows and columns, where* $\mathcal{B}_{s_i,0}$ *is the set of all* $s_i \times s_i$ *nonnegative definite matrices with zero row and column sums;*

where C_{di} *is the coefficient matrix of the reduced normal equation for estimating* $(\theta(i)_1, ..., \theta(i)_{si})', i = 1, ..., n$.

By Kiefer's Definition 2.4, Cheng had proved that an orthogonal array is *universally optimal.*

2.3.2 Construction of Partially Augmented Designs

For $s_1 = ... = s_{n-1} = s, s_n = p + k, k \geq 1$ and p may differ from s, assuming there is no orthogonal array with $M (M < 2N)$ combinations for the n factors with the levels $s_1 = ... = s_{n-1} = s, s_n = p + k$ respectively, but there is a $L_N(s^{n-1}p)$ which is universally optimal according to Cheng's theorem, then, it is possible to construct an $(N + N_1) \times n$ array in order to add $k (1 \leq k < p)$

levels on the n-th factor.

Firstly, we arrange the first $n-1$ factors with s levels each and the n-th factor with its first p levels into 1st, 2nd,..., n-th columns of an orthogonal array $L_N(s^{n-1}p)$ respectively. Secondly, we use another $L_N(s^{n-1}p)$, where we keep the same settings for the first $n-1$ factors and some $p-k$ levels of the n-th factor as the former array $L_N(s^{n-1}p)$, then replace the other k original levels of this factor by its last k new levels in the n-th column. It is easy to see that there are only $kT \triangleq N_1$, where $T = N/p$, new rows in the second $L_N(s^{n-1}p)$ comparing to the first $L_N(s^{n-1}p)$. Finally, the N rows of the first $L_N(s^{n-1}p)$ and the $N_1, (N_1 < N)$ new rows from the second orthogonal array $L_N(s^{n-1}p)$ form an augmented $(N+N_1) \times n$ array, which defines an augmented design. We call such a design based on the orthogonal array a *partially augmented design* (PAD) and denote this $(N+N_1) \times n$ array as $PAD_{N+N_1}(s^{n-1}(p+k))$.

This type of designs can be used in CA when one factor has more levels than other factors or some extra levels of one factor need to be investigated after an orthogonal design.

Example 2.1 If a $3 \times 3 \times 3 \times 4$ factorial design is considered, but fewer than 15 combinations can be investigated. $L_9(3^4)$ is used for four factors with three levels each, where $N = 9, s = p = 3$ and $n = 4$, while we want to augment another (i.e. $k = 1$) level to the fourth factor but there is no suitable orthogonal array to use. By choosing $T = N/p = 3$ augmented runs from the $L_9(3^4)$, total of three possible $PAD_{12}(3^3 \times 4)$ with 12 combinations each are generated in Table 2.1.

Table 2.1 Three $PAD_{12}(3^3 \times 4)$ based on $L_9(3^4)$

Stimulus	Factor 1	Factor 2	Factor 3	Factor 4
Original Design				
1	1	1	1	1
2	2	2	2	1
3	3	3	3	1
4	1	2	3	2
5	2	3	1	2
6	3	1	2	2
7	1	3	2	3
8	2	1	3	3
9	3	2	1	3
Possible augmented new runs (I)				
10	1	1	1	4
11	2	2	2	4
12	3	3	3	4
Possible augmented new rows (II)				
10'	1	2	3	4
11'	2	3	1	4
12'	3	1	2	4
Possible augmented new runs (III)				
10"	1	3	2	4
11"	2	1	3	4
12"	3	2	1	4

From Table 2.1, it is easy to see that any one of the three possible PADs with 12 runs is no longer orthogonal but has "partially orthogonality'. That is, a total of 9 pairs (1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3) in any two columns from the first three columns all appear but their frequencies

of appearances are different in the 12 rows. For example, the frequencies of each pair between factor 1 and 2 in the PAD (III), which includes the original orthogonal combinations No.1 to No.9 and the augmented combinations No.10" to No.12", are as follows:

Pair	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
Frequency	1	1	2	2	1	1	1	2	1

It is not hard to see, sometimes, the augmenting is equivalent to deleting. For example, from the following two tables we can see that their first 6 stimuli are the same when we augment one level of the third factor by adding two stimuli 5 and 6 to the orthogonal array $L_4(2^3)$, or delete one level of this factor by deleting two stimuli 7 and 8 from some columns of the orthogonal array $L_8(2^4 \times 4)$:

Stimulus	Factor 1	Factor 2	Factor 3
1	1	1	1
2	2	2	1
3	1	2	2
4	2	1	2
5	1	2	3
6	2	1	3

Stimulus	Factor 1	Factor 2	Factor 3
1	1	1	1
2	2	2	1
3	1	2	2
4	2	1	2
5	1	2	3
6	2	1	3
7	1	1	4
8	2	2	4

Therefore, the PA method might also be used in situation of deleting some unsatisfactory stimuli. For example, we can delete the fourth level of factor 3 from above $L_8(4 \times 2^4)$, i.e. the stimuli 7 and 8.

2.3.3 Partially Augmented Designs with Universal Optimality for the n-th Factor

By PA method we have the following design matrix in the linear model (2.3):

$$X = \begin{pmatrix} X_1 & X_2 & \cdots & X_{n-1} & X_n & 0 & \cdots & 0 \\ Z_1^{(1)} & Z_2^{(1)} & \cdots & Z_{n-1}^{(1)} & 0 & \mathbf{1}_T & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_1^{(k)} & Z_2^{(k)} & \cdots & Z_{n-1}^{(k)} & 0 & 0 & \cdots & \mathbf{1}_T \end{pmatrix}, \qquad (2.7)$$

where $Z_i^{(t)}(i = 1, ..., n-1; t = 1, ..., k)$ is the indicator matrix of i-th factor corresponding to the t-th augmented level of n-th factor in the second $L_N(s^{n-1}p)$.

For the special class of the augmented designs–partially augmented designs, we use design matrix X of (2.7) instead of $\tilde{X}$ in the model (2.3). The coefficient matrix $X'X$ of normal equation is then

$$\begin{pmatrix}
X_1'X_1+\sum_{t=1}^{k} Z_1^{(t)'}Z_1^{(t)} & \cdots & X_1'X_{n-1}+\sum_{t=1}^{k} Z_1^{(t)'}Z_{n-1}^{(t)} & X_1'X_n & Z_1^{(1)'}\mathbf{1}_T\cdots & Z_1^{(k)'}\mathbf{1}_T \\
X_2'X_1+\sum_{t=1}^{k} Z_2^{(t)'}Z_1^{(t)} & \cdots & X_2'X_{n-1}+\sum_{t=1}^{k} Z_2^{(t)'}Z_{n-1}^{(t)} & X_2'X_n & Z_2^{(1)'}\mathbf{1}_T\cdots & Z_2^{(k)'}\mathbf{1}_T \\
\vdots & & \vdots & \vdots & \vdots & \vdots \\
X_{n-1}'X_1+\sum_{t=1}^{k} Z_{n-1}^{(t)'}Z_1^{(t)} & \cdots & X_{n-1}'X_{n-1}+\sum_{t=1}^{k} Z_{n-1}^{(t)'}Z_{n-1}^{(t)} & X_{n-1}'X_n & Z_{n-1}^{(1)'}\mathbf{1}_T\cdots & Z_{n-1}^{(k)}\mathbf{1}_T \\
X_n'X_1 & \cdots & X_n'X_{n-1} & X_n'X_n & 0\cdots & 0 \\
\mathbf{1}_T'Z_1^{(1)} & \cdots & \mathbf{1}_T'Z_{n-1}^{(1)} & 0 & T\cdots & 0 \\
\vdots & & \vdots & \vdots & \vdots\ddots & \vdots \\
\mathbf{1}_T'Z_1^{(k)} & \cdots & \mathbf{1}_T'Z_{n-1}^{(k)} & 0 & 0\cdots & T
\end{pmatrix}$$

where for $i = 1, ..., n-1$,

$$\begin{aligned}
X_i'X_i &= \tfrac{N}{s} I_s, \\
X_n'X_n &= \tfrac{N}{p} I_p = T I_p, \\
X_i'X_j &= \tfrac{N}{s^2} J_s, i \neq j, j = 1, ..., n-1, \\
X_i'X_n &= \tfrac{N}{ps} J_{s,p}.
\end{aligned} \tag{2.8}$$

In fact, the matrix $Z_i^{(t)}$ is a $T \times s$ submatrix of X_i and corresponds to t-th augmented level of the n-th factor in the second $L_N(s^{n-1}p)$, so every level of the i-th factor should appear $\frac{T}{s}$ times in the T rows so that $\frac{T}{s} = \frac{N}{ps}$ must be a positive integer. We denote $m \triangleq \frac{T}{s}$ and have

$$\begin{aligned}
Z_i^{(t)'}Z_i^{(t)} &= mI_s, \\
Z_i^{(t)}\mathbf{1}_s &= \mathbf{1}_{ms}, \\
Z_i^{(t)'}\mathbf{1}_T &= m\mathbf{1}_s, \\
N_1 &= kT = kms,
\end{aligned} \tag{2.9}$$

for $i = 1, ..., n-1$. Using (2.8) and (2.9) the information matrix $X'X$ becomes:

$$\begin{pmatrix} \frac{N+kms}{s}I_s & \frac{N}{s^2}J_s+\sum_{t=1}^{k}Z_1^{(t)'}Z_2^{(t)} & \cdots & \frac{N}{s^2}J_s+\sum_{t=1}^{k}Z_1^{(t)'}Z_{n-1}^{(t)} & mJ_{s,(p+k)} \\ \frac{N}{s^2}J_s+\sum_{t=1}^{k}Z_2^{(t)'}Z_1^{(t)} & \frac{N+kms}{s}I_s & \cdots & \frac{N}{s^2}J_s+\sum_{t=1}^{k}Z_2^{(t)'}Z_{n-1}^{(t)} & mJ_{s,(p+k)} \\ \vdots & \vdots & & \vdots & \vdots \\ \frac{N}{s^2}J_s+\sum_{t=1}^{k}Z_{n-1}^{(t)'}Z_1^{(t)} & \frac{N}{s^2}J_s+\sum_{t=1}^{k}Z_{n-1}^{(t)'}Z_2^{(t)} & \cdots & \frac{N+kms}{s}I_s & mJ_{s,(p+k)} \\ mJ_{(p+k),s} & mJ_{(p+k),s} & \cdots & mJ_{(p+k),s} & msI_{p+k} \end{pmatrix} \tag{2.10}$$

Now our focus is to find some optimal design d^*s. i.e., to find some special $Z_i^{(t)}s$ which may lead to some optimal designs. Two new classes of PADs are presented below. It is shown that they universally optimal under certain conditions.

Lemma 2.3 *For a design $d \in \mathcal{D}_{N+N_1}$ under model (2.3) with design matrix $\tilde{X}$, the coefficient matrix $C_n(\tilde{X})$ of the reduced normal equation for estimating $a_n'\theta_n$ is $(p+k)\times(p+k)$ nonnegative definite, and $C_n(\tilde{X})\mathbf{1}=0$.*

Proof:
Since $C_n(\tilde{X}) = W_n' P_{(W_1 \ldots W_{n-1})^\perp} W_n$ by (2.5), it is nonnegative definite.

For convenience, we denote $W \triangleq (W_1, ..., W_{n-1})$, then $P_{(W_1 \ldots W_{n-1})^\perp} = P_{W^\perp}$ for $\tilde{X} = (W|W_n)$. Thus

$$\begin{aligned} C_n(\tilde{X})\mathbf{1}_{p+k} &= W_n' P_{W^\perp} W_n \mathbf{1}_{p+k} \\ &= W_n' P_{W^\perp} \mathbf{1}_{p+k}, \end{aligned}$$

this is because of $W_n \mathbf{1}_{p+k} = \mathbf{1}_{N+kms}$. Further, since $W\mathbf{1}_{(n-1)s} = \sum_{i=1}^{n-1} W_i \mathbf{1}_s = (n-1)\mathbf{1}_{N+kms}$,

$$\begin{aligned} C_n(\tilde{X})\mathbf{1}_{p+k} &= W_n' P_{W^\perp} \tfrac{1}{n-1} W \mathbf{1}_{(n-1)s} \\ &= 0. \end{aligned}$$

Theorem 2.4 *Under the model specified by (2.3) with design matrix (2.7), if there are positive integers $T = N/p, m = T/s$, and a partially augmented design d_1^* which is based on an orthogonal array $L_N(s^{n-1}p)$ and satisfies $\sum_{t=1}^{k} Z_i^{(t)'} Z_j^{(t)} = kmI_s$ for the $k(1 \leq k < p)$ augmented levels of the n-th factor and any $i, j = 1, ..., (n-1)(i \neq j)$, then d_1^* is a universally optimal design for $\Omega(\theta_n)$ in $\mathcal{D}_{N+N_1}$ with $N_1 = kT$.*

Proof:

First, we denote the design matrix of d_1^* by D_1. Since the d_1^* is a partially augmented design under model (2.3), the coefficient matrix $C_n(D_1)$ of the reduced normal equation for estimating $a_n'\theta_n$ is $(p+k) \times (p+k)$ nonnegative definite and $C_n(D_1)\mathbf{1} = 0$ by lemma (2.3). Therefore, $C_n(D_1) \in \mathcal{B}_{(p+k),0}$.

Second, $C_n(D_1)$ is c.s.:

If there are positive integers $T = N/p, m = T/s$, and a partially augmented design d_1^* which is based on an orthogonal array $L_N(s^{n-1}p)$ and satisfies $\sum_{t=1}^{k} Z_i^{(t)'} Z_j^{(t)} = kmI_s$ for the $k(1 \leq k < p)$ augmented levels of the n-th factor and any $i, j = 1, ..., (n-1)(i \neq j)$, that is, if a design d_1^* satisfies the conditions of Theorem (2.4), then the information matrix of (2.10) is

$$D_1'D_1 = \left(\begin{array}{cccc|c} \frac{N+kms}{s} I_s & kmI_s + \frac{N}{s^2} J_s & \cdots & kmI_s + \frac{N}{s^2} J_s & mJ_{s,(p+k)} \\ kmI_s + \frac{N}{s^2} J_s & \frac{N+kms}{s} I_s & \cdots & kmI_s + \frac{N}{s^2} J_s & mJ_{s,(p+k)} \\ \vdots & \vdots & & \vdots & \vdots \\ kmI_s + \frac{N}{s^2} J_s & kmI_s + \frac{N}{s^2} J_s & \cdots & \frac{N+kms}{s} I_s & mJ_{s,(p+k)} \\ \hline mJ_{(p+k),s} & mJ_{(p+k),s} & \cdots & mJ_{(p+k),s} & msI_{p+k} \end{array}\right) \quad (2.11)$$

$$\triangleq \begin{pmatrix} A_{n-1}, & D_{12} \\ D_{12}' & msI_{p+k} \end{pmatrix}$$

where $D_{12} = \mathbf{1}_{n-1} \otimes (mJ_{s,(p+k)})$, the $\otimes$ is Kronecker product, and A_{n-1} includes $(n-1) \times (n-1)$ submatrices.

Using equation (2.5), (2.4) and (2.11), we have

$$C_n(D_1) = msI_{p+k} - D_{12}' A_{n-1}^- D_{12},$$

where

$$\begin{aligned} D_{12}' A_{n-1}^- D_{12} = & \ (\mathbf{1}_{n-1}' \otimes m\mathbf{1}_{p+k}\mathbf{1}_s') A_{n-1}^- (\mathbf{1}_{n-1} \otimes m\mathbf{1}_s\mathbf{1}_{p+k}') \\ = & \ m^2 \mathbf{1}_{p+k}\mathbf{1}_{(n-1)s}' A_{n-1}^- \mathbf{1}_{(n-1)s}\mathbf{1}_{p+k}'. \end{aligned}$$

Since

$$A_{n-1} = \begin{pmatrix} \frac{N+kms}{s} I_s & kmI_s + \frac{N}{s^2} J_s & \cdots & kmI_s + \frac{N}{s^2} J_s \\ kmI_s + \frac{N}{s^2} J_s & \frac{N+kms}{s} I_s & \cdots & kmI_s + \frac{N}{s^2} J_s \\ \vdots & \vdots & & \vdots \\ kmI_s + \frac{N}{s^2} J_s & kmI_s + \frac{N}{s^2} J_s & \cdots & \frac{N+kms}{s} I_s \end{pmatrix}$$

$$= \begin{pmatrix} \frac{N}{s} I_s & \frac{N}{s^2} J_s & \cdots & \frac{N}{s^2} J_s \\ \frac{N}{s^2} J_s & \frac{N}{s} I_s & \cdots & \frac{N}{s^2} J_s \\ \vdots & \vdots & \ddots & \vdots \\ \frac{N}{s^2} J_s & \frac{N}{s^2} J_s & \cdots & \frac{N}{s} I_s \end{pmatrix} - (\mathbf{1}_{n-1} \otimes I_s)(-kmI_s)(\mathbf{1}_{n-1}' \otimes I_s)$$

$$\triangleq A_{11} - A_{12} A_{22}^- A_{21},$$

where $A_{12} = A_{21}' = \mathbf{1}_{n-1} \otimes I_s, A_{22}^- = -kmI_s$, by the formula of generalized inverse of partitioned matrix (Searle, 1982)

$$A_{n-1}^- = A_{11}^- + A_{11}^- A_{12} (A_{22} - A_{21} A_{11}^- A_{12})^- A_{21} A_{11}^-,$$

and by lemma (2.2) of Cheng (1978), we have $A_{11}^- = N^{-1} diag(sI_s, sI_s -$

$J_s, ..., sI_s - J_s)$, then

$$A_{22} - A_{21}A_{11}^{-}A_{12} = N^{-1}[(n-1)sI_s - (n-2)J_s],$$

and

$$(A_{22} - A_{21}A_{11}^{-}A_{12})^{-} = -\frac{Nkm}{N+(n-1)kms}[I_s + \frac{(n-2)km}{N+(n-1)kms}J_s].$$

Therefore,

$$\begin{aligned} & A_{12}(A_{22} - A_{21}A_{11}^{-}A_{12})^{-}A_{21}A_{11}^{-} \\ = & \ \frac{-km}{N+(n-1)kms}(J_{n-1} \otimes [I_s + \frac{(n-2)km}{N+kms}J_s])diag(sI_s, ..., sI_s - J_s) \\ = & \ \frac{-kms}{N+(n-1)kms}(\mathbf{1}_{n-1} \otimes c, \mathbf{1}_{n-1} \otimes (I_s - s^{-1}J_s), ..., \mathbf{1}_{n-1} \otimes (I_s - s^{-1}J_s)), \end{aligned}$$

where $c = I_s + \frac{(n-2)km}{N+kms}J_s$, and a generalized inverse of A_{n-1} is

$$A_{n-1}^{-} = a_0 \begin{pmatrix} a_1(I_s - a_2J_s) & -kms(I_s - s^{-1}J_s) & \cdots & -kms(I_s - s^{-1}J_s) \\ -kms(I_s - s^{-1}J_s) & a_1(I_s - s^{-1}J_s) & \cdots & -kms(I_s - s^{-1}J_s) \\ \vdots & \vdots & & \vdots \\ -kms(I_s - s^{-1}J_s) & -kms(I_s - s^{-1}J_s) & \cdots & -kms(I_s - s^{-1}J_s) \end{pmatrix},$$

where

$$a_0 = \frac{s}{N[N+(n-1)kms]}, \quad a_1 = N + (n-2)kms, \quad \text{and}$$

$$a_2 = \frac{(n-2)k^2m^2s}{(N+kms)[N+(n-2)kms]}.$$

Because $\mathbf{1}'_{(n-1)s}A_{n-1}^{-}\mathbf{1}_{(n-1)s}$ is the summation of all elements in A_{n-1}^{-}, and the summation of elements in $I_s - a_2J_s$ equals to $s - s^2a_2$ and the summation of elements in $I_s - s^{-1}J_s$ equals to 0, we have

$$\begin{aligned} D'_{12}A_{n-1}^{-}D_{12} = & \ m^2\mathbf{1}_{p+k}a_0[a_1(s - s^2a_2]\mathbf{1}'_{p+k} \\ = & \ \frac{m^2s^2}{N+kms}J_{p+k}. \end{aligned}$$

Thus,

$$C_n(D_1) = msI_{p+k} - \frac{m^2s^2}{N+kms}J_{p+k} \tag{2.12}$$

is c.s.

Finally, to prove that

$$trC_n(D_1) = \max_{X \in \mathcal{D}} trC_n(X),$$

$trC_n(D_1) = [(p+k)-1]ms$ follows directly from equation (2.12). On the other hand, by $\mathbf{1}'_{p+k}a_n = 0$, and $dim\mathcal{R}(\mathbf{1}_{p+k}) = 1$, there exist $p+k-1$ standardized orthogonal vectors in $\mathcal{R}^{\perp}(\mathbf{1}_{p+k})$. Let $\Gamma = (H|\frac{1}{\sqrt{p+k}}\mathbf{1}_{p+k})$, where the orthogonal matrix H consists of $p+k-1$ standardized orthogonal vectors of $\mathcal{R}^{\perp}(\mathbf{1}_{p+k})$, then we have (Ni and Ding, 1984)

$$trC_n(\tilde{X}) \leq trW_nHH'W_n' = (N+kms)(1-\frac{1}{p+k}) = (p+k-1)ms = trC_n(D_1).$$

Therefore, we proved that the d_1^* is universally optimal design for $\Omega(\theta_n)$ in the $\mathcal{D}_{N+kT}, T = ms$.

The following example demonstrates the existence of the universally optimal design.

Example 2.2 In the example 2.1, we have three $PAD_{12}(3^3 \times 4)$, where $N = 9, s = p = 3, k = 1$ and $T = 3$ which leads $m = 1$. For the augmented design (I) denoted by d^*, its total run includes orthogonal experimental combinations No.1 to 9 and augmented experimental combinations No. 10 to 12. Since its augmented indicator matrices are $Z_1^{(1)} = Z_2^{(1)} = Z_3^{(1)} = I_3$, hence $Z_i^{(1)'}Z_j^{(1)} = I_3$ for any $i, j = 1, 2, 3, i \neq j$; which satisfies the conditions in the

Theorem 2.4. Thus, this augmented design d^* is universally optimal for the $\Omega(\theta_4)$ in $\mathcal{D}_{12}$.

Theorem 2.5 *Under the model specified by (2.3) with design matrix (2.7), if there are positive integers $T = N/p, m = T/s$, and a partially augmented design d_2^* which is based on an orthogonal array $L_N(s^{n-1}p)$ and satisfies $\sum_{t=1}^{k} Z_i^{(t)'} Z_j^{(t)} = \frac{km}{s} J_s$ for the $k(1 \leq k < p)$ augmented levels of the n-th factor and any $i, j = 1, ..., (n-1)(i \neq j)$, then d_2^* is a universally optimal design for $\Omega(\theta_n)$ in $\mathcal{D}_{N+N_1}$ with $N_1 = kT$.*

Proof:

If there are positive integers $T = N/p, m = T/s$, and a partially augmented design d_2^* which is based on an orthogonal array $L_N(s^{n-1}p)$ and satisfies $\sum_{t=1}^{k} Z_i^{(t)'} Z_j^{(t)} = \frac{km}{s} J_s$ for the $k(1 \leq k < p)$ augmented levels of the n-th factor and any $i, j = 1, ..., (n-1)(i \neq j)$, then we have the information matrix of (2.10) for design matrix D_2 of the d_2^* as follows:

$$D_2' D_2 = \left(\begin{array}{cccc|c} \frac{N+kms}{s} I_s & \frac{N+kms}{s^2} J_s & \cdots & \frac{N+kms}{s^2} J_s & m J_{s,(p+k)} \\ \frac{N+kms}{s^2} J_s & \frac{N+kms}{s} I_s & \cdots & \frac{N+kms}{s^2} J_s & m J_{s,(p+k)} \\ \vdots & \vdots & & \vdots & \vdots \\ \frac{N+kms}{s^2} J_s & \frac{N+kms}{s^2} J_s & \cdots & \frac{N+kms}{s} I_s & m J_{s,(p+k)} \\ \hline m J_{(p+k),s} & m J_{(p+k),s} & \cdots & m J_{(p+k),s} & ms I_{p+k} \end{array}\right)$$

$$\triangleq \left(\begin{array}{cc} B_{n-1} & D_{12} \\ D_{12}' & ms I_{p+k} \end{array}\right)$$

where D_{12} is the same as in (2.11) and

$$B_{n-1} \triangleq \begin{pmatrix} B_{11} & B_{12} \\ B_{12}' & aI_s \end{pmatrix} = \left(\begin{array}{ccc|c} aI_s & \cdots & & as^{-1}J_s \\ \vdots & \ddots & & \vdots \\ \hline as^{-1}J_s & \cdots & & aI_s \end{array}\right).$$

Note that B_{n-1} denotes the matrix B with $(n-1) \times (n-1)$ submatrices, such as aI_s and $as^{-1}J_s$, $a = \frac{N+mks}{s}$ and $B_{11} = B_{n-2}$. Thus, $C_n(D_2)$, the coefficient matrix of the reduced normal equation for estimating $a_n'\theta_n$ is

$$C_n(D_2) = msI_{p+k} - D_{12}'B_{n-1}^{-}D_{12}.$$

By the same reason as in the proof of theorem (2.4), $C_n(D_2) \in \mathcal{B}_{(k+p),0}$.

To prove that $C_n(D_2)$ is c.s., we need to find the generalized inverse of B_{n-1} by the partitioned inverse rule for B_{n-1}. Since

$$\begin{aligned} B_{12}(aI_s)^{-1}B_{12}' &= (\mathbf{1}_{n-2} \otimes s^{-1}J_s)(aI_s)^{-1}(\mathbf{1}_{n-2}' \otimes as^{-1}J_s) \\ &= as^{-1}(J_{n-2} \otimes J_s) \\ &= as^{-1}J(n-2)s; \end{aligned}$$

$$\begin{aligned} E &\triangleq B_{11} - B_{12}(aI_s)^{-1}B_{12}' \\ &= B(n-2) - as^{-1}J(n-2)s \\ &= \begin{pmatrix} aI_s & \cdots & as^{-1}J_s \\ \vdots & \ddots & \vdots \\ as^{-1}J_s & \cdots & aI_s \end{pmatrix} - \begin{pmatrix} as^{-1}J_s & \cdots & as^{-1}J_s \\ \vdots & \ddots & \vdots \\ as^{-1}J_s & \cdots & as^{-1}J_s \end{pmatrix} \\ &= \begin{pmatrix} a(I_s - s^{-1}J_s) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a(I_s - s^{-1}J_s) \end{pmatrix} \\ &= I_{(n-2)} \otimes [a(I_s - s^{-1}J_s)]. \end{aligned}$$

Thus,

$$E^{-} = I_{(n-2)} \otimes a^{-1}(I_s - s^{-1}J_s).$$

$$\begin{aligned} B_{12}'E^- &= (\mathbf{1}_{n-2}' \otimes as^{-1}J_s)[I_{n-2} \otimes a^{-1}(I_s - s^{-1}J_s)] \\ &= s^{-1}[\mathbf{1}_{n-2}' \otimes (J_s - J_s)] \\ &= 0 \\ &= E^- B_{12}, \end{aligned}$$

and

$$B_{n-1}^- = \begin{pmatrix} E^- & 0 \\ 0 & (aI_s)^{-1} \end{pmatrix} = \begin{pmatrix} I_{n-2} \otimes a^{-1}(I_s - s^{-1}J_s) & 0 \\ 0 & a^{-1}I_s \end{pmatrix}.$$

Therefore,

$$\begin{aligned} & C_n(D_2) = msI_{p+k} - D_{12}'B^-(n-1)D_{12} \\ = \; & msI_{p+k} - (\mathbf{1}_{n-1}' \otimes mJ_{(p+k),s}) \begin{pmatrix} I_{n-2} \otimes a^{-1}(I_s - J_s/s) & 0 \\ 0 & a^{-1}I_s \end{pmatrix} (\mathbf{1}_{n-1} \otimes mJ_{s,(p+k)}) \\ = \; & msI_{p+k} - (\mathbf{1}_{n-2}' \otimes ma^{-1}J_{(p+k),s}(I_s - s^{-1}J_s) | ma^{-1}J_{(p+k),s}) \begin{pmatrix} \mathbf{1}_{n-2} \otimes mJ_{s,(p+k)} \\ mJ_{s,(p+k)} \end{pmatrix} \\ = \; & msI_{p+k} - m^2a^{-1}J_{(p+k),s}J_{s,(p+k)} \\ = \; & msI_{p+k} - \tfrac{m^2s^2}{N+kms}J_{p+k}. \end{aligned} \tag{2.13}$$

Which is a c.s. matrix.

The equations (2.12) and (2.13) show that the expressions for $C_n(D_1)$ and $C_n(D_2)$ are identical. Therefore, the d_2^* is also a universally optimal design for $\Omega(\theta_n)$ in the $\mathcal{D}_{N+N_1}$ with $N_1 = kT$.

Such kinds of universally optimal design also exist as demonstrated below.

Example 2.3 For a $2^3 \times 7$ factorial design, it seems there is no suitable orthogonal array with fewer than 32 runs to use. Based on an $L_{16}(2^{12} \times 4)$, since $T = N/p = 16/4 = 4, k = 7 - 4 = 3, N_1 = kT = 12$ we use its three columns with two levels each and one column with four levels to form a $PAD_{28}(2^3 \times 7)$ (Table 2.2) for this case:

Table 2.2 A universally optimal $PAD_{28}(2^3 \times 7)$

Stimulus	Factor 1	Factor 2	Factor 3	Factor 4
Original Design				
1	1	1	1	1
2	1	2	2	1
3	2	1	2	1
4	2	2	1	1
5	1	1	1	2
6	1	2	2	2
7	2	1	2	2
8	2	2	1	2
9	1	1	1	3
10	1	2	2	3
11	2	1	2	3
12	2	2	1	3
13	1	1	1	4
14	1	2	2	4
15	2	1	2	4
16	2	2	1	4
Augment combinations				
17	1	1	1	5
18	1	2	2	5
19	2	1	2	5
20	2	2	1	5
21	1	1	1	6
22	1	2	2	6
23	2	1	2	6
24	2	2	1	6
25	1	1	1	7
26	1	2	2	7
27	2	1	2	7
28	2	2	1	7

From the table 2.2 we have

$$Z_1^{(t)'} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad Z_2^{(t)'} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Z_3^{(t)'} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

for $t = 1, 2, 3$. Then

$$\sum_{t=1}^{3} Z_1^{(t)'} Z_2^{(t)} = \sum_{t=1}^{3} Z_1^{(t)'} Z_3^{(t)} = \sum_{t=1}^{3} Z_2^{(i)'} Z_3^{(t)} = 3J_2 = \frac{km}{s} J_s,$$

where $m = T/s = 2$. Therefore, this PAD d^* satisfies the conditions of theorem 2.5, so it is universally optimal for the $\Omega(\theta_4)$ in the $\mathcal{D}_{16+(3)(4)} = \mathcal{D}_{28}$.

2.3.4 The Efficiencies of the PADs for the First $n-1$ Factors

In the previous subsection, the Theorem 2.4 and 2.5 provide the universal optimality of PADs for the $\Omega(\theta_n)$ under the model (2.3) with design matrix (2.7). Although the optimality for the other $\Omega(\theta_j), j = 1, ..., n-1$, has not been investigated yet, because of the complexity in find the related generalized inverse, but we have studied the A-, D-, and E- efficiencies for the $n-1$ factors of some PADs by SAS program and the results are listed in the following table. Here, with no loss of generality we can partition the $D_1'D_1$ of (2.11) as $\begin{pmatrix} u_2 s^{-1} I_s & E_{12} \\ E_{12}' & Q_{n-1} \end{pmatrix}$ and the $D_2'D_2$ as $\begin{pmatrix} u_2 s^{-1} I_s & F_{12} \\ F_{12}' & R_{n-1} \end{pmatrix}$ for $u_2 = N + kms$, the total number of experiments in the PADs, respectively, so that the C-matrices of the design d^* which satisfied the conditions of Theorem 2.4 or 2.5 for estimating θ_i are $C_i(D_1) = C_1(D_1) = u_2 s^{-1} I_s - E_{12} Q_{n-1}^- E_{12}'$ or $C_i(D_2) = C_1(D_2) = u_2 S^{-1} I_s - F_{12} R_{n-1}^- F_{12}', i = 1, ..., n-1$, respectively; as a comparison, the full $s^{n-1} \times (p+k) (\triangleq u_1)$ factorial design (FFD), which is an orthogonal design with universal optimality, has its C-matrix

$C_f = s^{-1}NI_s - s^{-2}NJ_s$ (Cheng, 1978). Since the non-zero eigenvalues of the C-matrix C_d are $\mu_{d1} \geq ... \geq \mu_{d(s-1)}$, for $l = 1, 2$, we define the A-efficiency as $\frac{\phi_A(C_f)u_1}{\phi_A(C_1(D_l))u_2}$, the D-efficiency as $\frac{\phi_D(C_f)u_1^{s-1}}{\phi_D(C_1(D_l))u_2^{s-1}}$ and the E-efficiency as $\frac{\phi_E(C_f)u_1}{\phi_E(C_1(D_l))u_2}$.

The efficiencies of some PADs for the first $n-1$ factors

		Designs		
The types of Designs	FFD	$2^2 \times 3$	$3^5 \times 7$	$2^5 \times 7$
	PAD	$PAD_6(2^2 \times 3)$	$PAD_{21}(3^5 \times 7)$	$PAD_{56}(2^5 \times 7)$
No. of runs	u_1	12	1701	224
	u_2	6	21	56
C-matrix	FFD	$6I_2 - 3J_2$	$567I_3 - 189J_3$	$112I_2 - 56J_2$
	PAD	$\frac{8}{3}I_2 - \frac{4}{3}J_2$	$6.6I_3 - 2.2J_3$	$28I_2 - 14J_2$
Eigenvalue of the C-matrix	FFD	6, 0	567, 567, 0	112, 0
	PAD	$\frac{8}{3}$, 0	6.6, 6.6, 0	28, 0
Efficiency	A-	.89	.94	1
	D-	.89	.89	1
	E-	.89	.94	1

From the table, we can find that the A-, D- and E-efficiency are all greater than 0.89, from which we may conclude the PADs' approximate A-, D- and E- optimality for estimating $\theta_i (i = 1, ..., n-1)$ under the conditions of the Theorem 2.4 or 2.5.

2.4 Application of Universally Optimal PAD in CA

The PAD can be applied in a variety of situations of CA:

- There is one factor which includes more levels than the other factors but there is no a suitable orthogonal array to use.
- The number of experimental combinations must be smaller than the existing OA for some reasons.
- The survey needs to be separated by two stage designs, or after completing a set of experiments, some extra levels in one of the factors need to be compared statistically with former levels of the factor.

2.4.1 $\mathbf{PAD}_{21}(3^5 \times 7)$

Example 2.4 To examine preferences for various attributes of personal computers, the following table reports the 6 attributes and 22 attribute levels.

Description of Attributes Personal Computers

Attribute Name	Attribute Levels
1. Chip	1. Intel Microprocessor
	2. AMD Microprocessor
	3. Cyrix Microprocessor
2. Monitor	1. Average (VGA 640X480)
	2. Above average (SVGA 1024X788)
	3. Far above average (XVGA 1280X1024)
3. Storage	1. Average (80 meg Hard Drive)
	2. Above average (120 meg Hard Drive)
	3. Far above average (240 meg Hard Drive)
4. Upgradeability	1. Processor can not be upgraded
	2. Can upgrade to a faster processor in the further
	3. Can upgrade to a faster processor now
5. Price	1. \$1500/ \$2700
	2. \$2000/ \$3500
	3. \$2500/ \$4300
6. Brand	1. Brand A (Most preferred)
	2. Brand B
	3. Brand C
	4. Brand D
	5. Brand E
	6. Brand F
	7. Brand G (Least preferred)

(This example is adopted from Allenby, Arora and Ginter, 1995)

For this $3^5 \times 7$ factorial design, there is no a suitable orthogonal array which defines a fractional factorial design with a small number of runs. By the PA method we can augment one level to the sixth factor based on the 6 columns of $L_{18}(3^6 \times 6)$. Since $N = 18, s = 3, p = 6, k = 1$, and $T = N/p = 3$, three experimental combinations can be augmented to the original 18 exper-

imental combinations. According to the theorem 2.4 we can choose the first three combinations of the orthogonal array as the augmented part by change the original level 1 to level 7, that is, the PAD matrix includes total of 21 combinations, where one factor is with 7 levels and five factors are with 3 levels each as follows:

Table 2.3 A universally optimal $PAD_{21}(3^5 \times 7)$

	Chip	Monitor	Storage	Ability	Price	Brand
Original Design						
1	1	1	1	1	1	1
2	2	2	2	2	2	1
3	3	3	3	3	3	1
4	1	1	2	2	3	2
5	2	2	3	3	1	2
6	3	3	1	1	2	2
7	1	2	1	3	2	3
8	2	3	2	1	3	3
9	3	1	3	2	1	3
10	1	3	3	2	2	4
11	2	1	1	3	3	4
12	3	2	2	1	1	4
13	1	2	3	1	3	5
14	2	3	1	2	1	5
15	3	1	2	3	2	5
16	1	3	2	3	1	6
17	2	1	3	1	2	6
18	3	2	1	2	3	6
Augment Design						
19	1	1	1	1	1	7
20	2	2	2	2	2	7
21	3	3	3	3	3	7

2.4.2 A Real Example of Universally Optimal PAD

New developments of models and methods in CA are arriving so fast that even specialists find it difficult to keep up. In order to assess these new models and new methods, sometimes Monte Carlo simulation is used (Carroll and Green, 1995). Several authors (DeSarbo, Oliver and Rangaswamy, 1989; Wedel and Kistemaker,1989; Wedel and Steenkamp, 1989) have used small Monte Carlo studies to evaluate the performance of their methods and investigate the effects of factors hypothesized to influence performance. Vriens, Wedel and Wilms (1994) use Monte Carlo simulation to compare nine different models related to metric conjoint segmentation. Their ANOVA were based on 576 (64 data sets $\times$9 methods) combinations.

In this section, the author provides an example of universally optimal PAD for a real problem in a Monte-Carlo study in CA.

The six factors are as follows:

A (Corrtype)	two types of correlation (Pearson and Polychoric)
B (Item)	seven types of items (four of same category and three of mixture type)
C (J)	two test lengths (20 and 40)
D (Loading)	two types of item loadings
E (N)	two sample sizes (1000 and 1500)
F (ρ)	two levels of correlation between abilities ($r = 0.3\&0.7$)

Nandakumar, Yu, Li and Stout (1995) use a full factorial design ($2 \times 7 \times 2 \times 2 \times 2 \times 2 = 224$) to investigate the performance of a new software POLYDIMTEST which is used to assess the unidimensionality of polytomous test data. This is a simulation study in the test industry (for details see Li and

Stout, 1995). For each of the combinations, 100 replications were carried out and altogether 22400 replications were done. The response is the power of the POLY-DIMTEST.

In Table 2.4, the first two columns summarize the results of their simulation study, that is, the estimators of level effect (cell mean) and P-values for each factor by ANOVA. Their findings are included in the far right column of the same table. The detailed ANOVA is given in Table 2.7.

Because the size of the experiment is huge when a large number of factors is to be studied, simulation studies involving full factorial design are usually extremely time-consuming. For this reason, the author proposes a $\text{PAD}_{56}(2^5 \times 7)$ with universal optimality for the same study. Another reason for using the PAD is that there is no an orthogonal array which is appropriate to reduce the size of experiment for this study. We choose specific six columns from a $L_{32}(2^{28} \times 4)$ including the unique column with 4 levels to arrange the 6 factors, where only the first 4 levels of factor B (Item) are used, $N = 32, s = 2, p = 4$. We use this 32×6 orthogonal array as our first 32 experiments labeled No.1 to No.32. Since $p + k = 7$, so $k = 3$ and $N_1 = kT = k(N/p) = (3)(32/4) = 24$. Using the same 6 columns from another $L_{32}(2^{28} \times 4)$ we keep almost all settings as the same as in the first 32 experiments but use level 5, 6, 7 and 4 of the factor B instead of the level 1, 2, 3 and 4 in the 4 levels column respectively. Comparing with the first 32 experiments we find that there are 24 new combinations, which is augmented part with label No.33 to No. 56. Merging these 24 new experiments into the 32×6 array, a $\text{PAD}_{56}(2^5 \times 7)$ is formed in Table 2.5. It can be shown that this $\text{PAD}_{56}(2^5 \times 7)$ meets the conditions of Theorem 2.5, hence it is a universally optimal PAD denoted as $\text{UOPAD}_{56}(2^5 \times 7)$.

As the comparison the estimators of level effect for all factors in the $\text{UOPAD}_{56}(2^5 \times 7)$ and related F-test are also listed in Table 2.4. The conclusions based on the $\text{UOPAD}_{56}(2^5 \times 7)$ are almost the same as those based on full factorial design $\text{L}_{224}(2^5 \times 7)$, except for the p-value of factor A. The important fact is that 3/4 of the 22400 replications are saved by use of the PAD.

In addition, if we observe the detailed ANOVA given in Table 2.8 and compare with Table 2.7, it can be seen that the residual under the full factorial design $\text{L}_{224}(2^5 \times 7)$ is 79.494, and for the $\text{UOPAD}_{56}(2^5 \times 7)$ is 64.087. The efficiency $\frac{MSE_{r56}/56}{MSE_{r224}/224} = 3.225$. As a comparison, we also give another design $\text{PAD}_{56}(2^5 \times 7)$ (see Table 2.6), which is also a partially augmented design but does not satisfy either conditions of Theorem 2.4 and 2.5. The related results of ANOVA are given in Table 2.9.

This example shows that the use of the PAD with universal optimality provides an efficient alternative of the full factorial design with advantages in small number of stimuli, some statistical optimalities, and general tool of analysis, etc.

Table 2.4 The summary based on data of Nandakumar et al

Design	Full		UOPAD		Nandakumar's Description
replications	22,400		5,600		
factor	Cell Mean	P	Cell Mean	P	
A (Corr)		$< .01$		$> .05$	Pearson correlations led to
Pearson	70.67		70.96		equal or better performance
Polychoric	65.74		66.25		than Polychoric correlations.
B (Item)		$< .01$		$< .05$	The power decreased slightly
3	83.72		83.13		for tests containing mixed
4	79.06		75.50		category items.
2	71.56		71.00		
1	62.22		66.23		
6	62.13		62.75		
5	60.34		61.00		
7	58.41		60.25		
C (J)		$< .01$		$< .01$	The power increased as the
40	77.03		79.07		number of item increased.
20	59.38		58.14		
D (Load)		$< .01$		$< .01$	When tests contain mixed-
1 or 2	75.36		78.14		ability items, the power
1 and 2	61.05		59.07		decreases.
E (N)		$< .01$		$< .01$	The power increased as the
1500	73.06		73.11		sample size increased.
1000	63.35		63.11		
F (ρ)		$< .10$		$< .01$	The power increased as the
0.3	83.3		85.17		correlation between the
0.7	53.11		51.50		abilities decreased.

Table 2.5 $UOPAD_{56}(2^5 \times 7)$ based on data of Nandakumar et al

No.	Load	J	ρ	N	Item	Corr	No.	Load	J	ρ	N	Item	Corr
1	1	1	1	1	1	1	29	1	2	1	1	4	2
2	1	2	2	2	1	1	30	1	1	2	2	4	2
3	2	2	1	2	1	1	31	2	1	1	2	4	2
4	2	1	2	1	1	1	32	2	2	2	1	4	2
5	1	2	1	1	1	2	33	1	1	1	1	5	1
6	1	1	2	2	1	2	34	1	2	2	2	5	1
7	2	1	1	2	1	2	35	2	2	1	2	5	1
8	2	2	2	1	1	2	36	2	1	2	1	5	1
9	1	1	1	1	2	1	37	1	2	1	1	5	2
10	1	2	2	2	2	1	38	1	1	2	2	5	2
11	2	2	1	2	2	1	39	2	1	1	2	5	2
12	2	1	2	1	2	1	40	2	2	2	1	5	2
13	1	2	1	1	2	2	41	1	1	1	1	6	1
14	1	1	2	2	2	2	42	1	2	2	2	6	1
15	2	1	1	2	2	2	43	2	2	1	2	6	1
16	2	2	2	1	2	2	44	2	1	2	1	6	1
17	1	1	1	1	3	1	45	1	2	1	1	6	2
18	1	2	2	2	3	1	46	1	1	2	2	6	2
19	2	2	1	2	3	1	47	2	1	1	2	6	2
20	2	1	2	1	3	1	48	2	2	2	1	6	2
21	1	2	1	1	3	2	49	1	1	1	1	7	1
22	1	1	2	2	3	2	50	1	2	2	2	7	1
23	2	1	1	2	3	2	51	2	2	1	2	7	1
24	2	2	2	1	3	2	52	2	1	2	1	7	1
25	1	1	1	1	4	1	53	1	2	1	1	7	2
26	1	2	2	2	4	1	54	1	1	2	2	7	2
27	2	2	1	2	4	1	55	2	1	1	2	7	2
28	2	1	2	1	4	1	56	2	2	2	1	7	2

Table 2.6 $PAD_{56}(2^5 \times 7)$ based on data of Nandakumar et al

No.	Load	J	ρ	N	Item	Corr	No.	Load	J	ρ	N	Item	Corr
1	1	1	1	1	1	1	29	2	2	1	1	4	2
2	2	2	1	2	1	1	30	1	1	1	2	4	2
3	2	1	2	1	1	1	31	1	2	2	1	4	2
4	1	2	2	2	1	1	32	2	1	2	2	4	2
5	2	1	1	1	1	2	33	1	1	1	1	5	1
6	1	2	1	2	1	2	34	2	2	1	2	5	1
7	1	1	2	1	1	2	35	2	1	2	1	5	1
8	2	2	2	2	1	2	36	1	2	2	2	5	1
9	2	1	1	1	2	1	37	2	1	1	1	5	2
10	1	2	1	2	2	1	38	1	2	1	2	5	2
11	1	1	2	1	2	1	39	1	1	2	1	5	2
12	2	2	2	2	2	1	40	2	2	2	2	5	2
13	1	1	1	1	2	2	41	2	1	1	1	6	1
14	2	2	1	2	2	2	42	1	2	1	2	6	1
15	2	1	2	1	2	2	43	1	1	2	1	6	1
16	1	2	2	2	2	2	44	2	2	2	2	6	1
17	2	2	1	1	3	1	45	1	1	1	1	6	2
18	1	1	1	2	3	1	46	2	2	1	2	6	2
19	1	2	2	1	3	1	47	2	1	2	1	6	2
20	2	1	2	2	3	1	48	1	2	2	2	6	2
21	1	2	1	1	3	2	49	2	2	1	1	7	1
22	2	1	1	2	3	2	50	1	1	1	2	7	1
23	2	2	2	1	3	2	51	1	2	2	1	7	1
24	1	1	2	2	3	2	52	2	1	2	2	7	1
25	1	2	1	1	4	1	53	1	2	1	1	7	2
26	2	1	1	2	4	1	54	2	1	1	2	7	2
27	2	2	2	1	4	1	55	2	2	2	1	7	2
28	2	1	2	2	4	1	56	1	1	2	2	7	2

Table 2.7 ANOVA Table for Full Factorial Design (N=224)

* * * A N A L Y S I S O F V A R I A N C E * * *

POWER
by CORRTYPE
ITEM
J
LOADING
N
RHO

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	105809.750	11	9619.068	121.003	.000
CORRTYPE	1360.286	1	1360.286	17.112	.000
ITEM	19214.429	6	3202.405	40.285	.000
J	5284.571	1	5284.571	66.477	.000
LOADING	11457.161	1	11457.161	144.125	.000
N	17431.143	1	17431.143	219.275	.000
RHO	51062.161	1	51062.161	642.337	.000
Explained	105809.750	11	9619.068	121.003	.000
Residual	16852.804	212	79.494		
Total	122662.554	223	550.056		

Table 2.8 ANOVA Table for $UOPAD_{56}(2^5 \times 7)$

* * * A N A L Y S I S O F V A R I A N C E * * *

POWER
by CORRTYPE
ITEM
J
LOADING
N
RHO

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	32497.536	11	2954.321	46.099	.000
CORRTYPE	311.143	1	311.143	4.855	.033
ITEM	3439.607	6	573.268	8.945	.000
J	1134.000	1	1134.000	17.695	.000
LOADING	5092.071	1	5092.071	79.456	.000
N	6132.071	1	6132.071	95.684	.000
RHO	16388.643	1	16388.643	255.726	.000
Explained	32497.536	11	2954.321	46.099	.000
Residual	2819.821	44	64.087		
Total	35317.357	55	642.134		

Table 2.9. ANOVA Table for $PAD_{56}(2^5 \times 7)$

* * * A N A L Y S I S O F V A R I A N C E * * *

POWER
by CORRTYPE
ITEM
J
LOADING
N
RHO

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	27141.394	11	2467.399	21.890	.000
CORRTYPE	279.018	1	279.018	2.475	.123
ITEM	6142.357	6	1023.726	9.082	.000
J	2040.537	1	2040.537	18.103	.000
LOADING	2405.161	1	2405.161	21.338	.000
N	4376.823	1	4376.823	38.830	.000
RHO	10892.161	1	10892.161	96.632	.000
Explained	27141.394	11	2467.399	21.890	.000
Residual	4959.588	44	112.718		
Total	32100.982	55	583.654		

Chapter 3

Uniform Designs

3.1 Introduction

Generally, conjoint analysis gives the analyst three alternatives to choose ranging from the most restrictive (a linear relationship) to the least restrictive (separate part-worths), with the ideal-point or quadratic model falling in between. The linear model is the simplest yet of most restricted form, because we estimate only a single part-worth (similar to regression coefficient) that is multiplied by the value of level. In the quadratic form, also known as the ideal model, the assumption of strict linearity is relaxed, such that we have a simple curvilinear relationship. Finally, the part-worth form is most general, allowing for each level to have its own part-worth estimate. When using the separate part-worths, the number of estimated values increases quickly as we add factors and levels because each new level has a separate part-worth estimate. The type of relationship can be specified for each factor separately, thus allowing for mixture if needed. This choice does not affect how the treatments or stimuli are created, but it does impact on how and what types of part-worths are estimated by CA.

Green (1974) pointed out a *problem* in the design field of conjoint studies. He assumed that twelve different airline carriers are involved in a conjoint study. After 20 years, Carroll and Green (1995) wrote:

> Two trends have been noted in the application of conjoint analysis to business problems. First, the early successes of conjoint analysis have led to industry demands for techniques that handle ever larger numbers of attributes and attribute levels. ...

They considered that coping with larger numbers of attributes and levels still is a problem in CA. In fact, with growing emphasis on marketing studies that deal with the measurement of consumer preferences (e.g., conjoint analysis) it is not unusual for some researchers to deal with twelve or more factors, each involving two to six or more levels.

Several experimental designs, including completely randomized, randomized block, incomplete block, orthogonal designs, are based on an analysis of variance (ANOVA) model. This model has the advantage of allowing a quite general dependence of the response on the various factors and levels.

Suppose there are s factors, and for simplicity we assume an equal number of levels q of each factor. Carrying out such a experiment would be expensive when using orthogonal array even if q is of moderate size. The constraint of orthogonality requires $N \geq q^2$ for the number N of experiments. Moreover, an even larger number of experiments is required to estimate all the main and interaction effects. Hence, most authors recommend the use of $L_N(2^s)$ and $L_N(3^s)$. But sometimes the restriction of two or three levels is not adequate.

An alternative approach is taken to consider some design d, such as optimal designs, by assuming some regression model

$$E(Y) = X\beta, \qquad Var(Y) = \sigma^2 I_N. \tag{3.1}$$

where Y is an $N \times 1$ response vector, X is an $N \times s$ design matrix, $\beta' = (\beta_1, ..., \beta_s)$ is a parameter vector. The least square estimate of β is $\hat{\beta} = (X'X)^{-1}X'Y$ with covariance matrix $Var(\hat{\beta}) = (X'X)^{-1}$, where $C_d = X'X$ is information matrix. The ANOVA model can be considered as a special case of (3.1), where the elements of X consist only of zero and ones. Since one disadvantage of the ANOVA model is that the number of parameters to be estimated is quite large, therefore, we may consider the regression model (3.1).

According to these concerns a new class of experimental designs called *uniform design(UD)*, generated by *number-theoretic methods* and based on the regression model (Fang and Wang, 1994), will be introduced for conjoint analysis in this chapter. The UD is proposed by Fang (1980), Wang and Fang (1981) and has been widely applied in China to problems in system engineering, pharmaceutics, chemical engineering, and natural sciences. We may use the UD in situations of larger number of attributes and levels, or of environmental correlation between attributes under regression model for CA.

3.2 Construction of Uniform Designs (UD)

If we drop the principle (b) of a good experimental design and only keep the principle (a) described in chapter 1, then we may use a uniformly scattered set of points on the experimental domain. Wang and Fang (1981) have given

the following definitions:

Definition 3.1 A uniform-type design (U-type design for simplicity), denoted by $U_N(N^t)$, is a $N \times t$ matrix with each column being a permutation of $1, 2, ..., N$.

A U-type design $U_N(N^t)$ provides an experimental design with N experimental points for t factors at N levels each. Where the notation $U_N(N^t)$ has a similar meaning to the orthogonal array $L_N(s^t)$: "U" denotes the U-type design; "N" the total number of experiments, which equals the number of levels for each factor; "t" the maximum number of possible columns. There is a basic difference between orthogonal design and U-type design: in the orthogonal design, any $q < t$ columns are equivalent, i.e., any q columns are orthogonal, but in the U-type design, the efficiencies may be different for distinct choices of q columns, i.e., the uniformity may differ for different q columns. Therefore, a recommendation of best q columns for any given number $q < t$ is needed. Given a $U_N(N^t) = (u_{ij})$, let $x_{ij} = (u_{ij} - 0.5)/N$ for $i = 1, ..., N$ and $j = 1, ..., t$, then $X = (x_{ij})$ is an $N \times t$ induced matrix denoted by $X_N(N^t)$, where $x_{ij} \in (0, 1)$. There is an one-to-one correspondence between the $U_N(N^t)$ and $X_N(N^t)$. Each row in the matrix X also can be treated as an experimental point. Let $\mathcal{U}_{N,t}$ be a set of all U-type designs for given N, $\mathcal{X}_{N,t}$ be a set of induced matrices of all U-type designs and D be a measure of uniformity defined on the $\mathcal{X}_{N,t}$. D maps $\mathcal{X}_{N,t}$ into $R^+ = [0, \infty)$. Suppose that the lower D corresponds to the better uniformity for the N experimental points of X (or say, $U_N(N^t)$).

Definition 3.2 A U-type design $U_N(N^t) = (u_{ij})$ is called *uniform design(UD)* for given N if its induced matrix $X_N(N^t)$ has smallest D-value

among $\mathcal{X}_{N,t}$.

It is clear that there are different uniform designs corresponding to different measures of uniformity, and any criterion of uniformity should be invariant for exchanging rows and columns of a U-type design. Therefore, the first column of $U_N(N^t)$ can always be taken $(1, 2, ..., N)'$. There are $N! - 1$ possible permutations of $\{1, 2, ..., N\}$ for the second column, $N! - 2$ choices for the third column, and so on. Finding a uniform design $U_N(N^t)$ is a difficult task for even moderate N and t, e.g., $N \geq 10, t \geq 5$, the load of computation for choosing the best t columns from all possibilities is very heavy. Thus, many authors have restricted consideration to some *subsets* of $\mathcal{U}_{N,t}$.

3.2.1 Uniform Designs Based on Good Lattice Point Sets

Wang and Fang(1981) suggested to use good lattice point (*glp*) sets as the subset of $\mathcal{U}_{N,t}$. For given N and $s \leq t$ they defined a subset $\mathcal{UG}_{N,s}$ of the $\mathcal{U}_{N,s}$ as follows:

Definition 3.3 Let $\{h_1, ..., h_t\}(h_1 < h_2 < ... < h_t < N)$ be a set of all positive integers, which satisfies the conditions that the great common divisor (g.c.d.) $(N, h_i) = 1$ for $i = 1, ..., t$. The maximum number t can be determined by the Euler function $\phi(N)$ (Hua, 1956). $\phi(N) = N(1 - \frac{1}{p_1})(1 - \frac{1}{p_2})...(1 - \frac{1}{p_r})$, where $N = p_1^{t_1}...p_r^{t_r}$ is the prime decomposition of N. Let $u_{ij} \equiv ih_j(modN)$ and $\mathbf{u}_j = (u_{1j}, ..., u_{Nj})', i = 1, ..., N, j = 1, ..., t$, where the multiplicative operation modulo N is modified such that $1 \leq u_{ij} \leq N$, e.g., $u_{ij} = N$ if $ih_j = mN$ for some integer m. Let G_N be a set of $\mathbf{u}_j, j = 1, ..., t$.

Then the subset $\mathcal{UG}_{N,s}$ of $\mathcal{U}_{N,s}$ is defined as

$$\mathcal{UG}_{N,s} = \{U : U \in \mathcal{U}_{N,s}, \quad \text{each column of U belongs to } G_N\}.$$

A design in $\mathcal{UG}_{N,s}$ is called a UG-type design. A UG-type design with smallest D-value for its induced matrix over $\mathcal{UG}_{N,s}$ is called a G-uniform design and denoted by $UG_{N,s}$.

For examples, since $12 = 3 \times 2^2$, i.e., $p_1 = 2$, $p_2 = 3$, so $\phi(12) = 12(1 - \frac{1}{2})(1 - \frac{1}{3}) = 4$. In fact, the 4 possible associated h_i with $(h_i, N) = 1$ are 1,5,7 and 11; and $\phi(7) = 7(1 - \frac{1}{7}) = 6$ with 1,2,3,4,5,6 as h_is. Therefore for given N, we have $t = \phi(N)$.

Fang (1980), Wang and Fang (1981) have given a detailed discussion on properties and generation of UG-type designs. Since computation required to find the G-uniform designs is burdensome even when the $\phi(N)$ is moderate, there are two ways to reduce the computing load. The first is to choose $h_1 = 1$, which reduces the possible designs from $\binom{\phi(N)}{s}$ to $\binom{\phi(N)-1}{s-1}$. The second is to consider $(h_1, ..., h_t)$ to be of the form $(1, a, a^2, ..., a^{t-1})(mod N)$, where a is an integer satisfying $1 \leq a < N$ and $a^i \not\equiv a^j (mod N), 1 \leq i < j \leq t - 1$. (Niederreiter, 1977, and Fang and Wang, 1994). Since the a has at most $N - 1$ possible choices and we only need find the best a after the comparing, the computing load reduced dramatically. Although this restriction only yields nearly G-uniform designs, most design given by Fang (1980) and others were obtained in this way. They obtained G-uniform designs based on the discrepancy for $4 \leq N \leq 31$ and $s \leq t = \phi(N)$.

Fang and Wang (1994) suggested to use only the first N rows of $U_{N+1}((N+1)^s)$ notated as $U_N^*(N^s)$ for even N to improve the uniformity of G-uniform

designs and $\mathcal{UG}^{*}_{N,s}$ was defined similarly. Such modification can lead to designs with lower discrepancy than existing nearly G-uniform designs.

The author point out that the G-uniform designs are available for any number of levels N and any number of attributes s only when $s \leq \frac{t}{2}+1$ for $U_N(N^t)$ and $U^{*}_N(N^t)$.

Why should the number s of attributes be less than or equal $\frac{t}{2}+1$ when total t columns is included in the $U_N(N^t)$? As an answer, the author proposes the following theorem concerning the ranks of $U_N(N^t)$ and $U^{*}_N(N^t)$:

Theorem 3.1 *For any UG-type design, the ranks of matrices $U_N(N^t)$ and $U^{*}_N(N^t)$ $(N \geq 4)$ are both less than or equal to $\frac{t}{2}+1$.*

Proof: According to the definition (3.3), we consider the following three cases:

(1) N is a prime number. So, the Euler function $\phi(N) = N-1$. Thus, there is an integer $k, k < t = \phi(N)$ such that $t = rk$. For any given $h_i = i (i \leq k)$ of the first row $(h_1, h_2, ..., h_t) = (1, 2, ..., N-1)$ in the $U_N(N^t)$, there must be another $h_{N-i} = N - i$ in the same row corresponding to it. The m-th row of the $U_N(N^t)$ is $(mh_1, mh_2, ..., mh_t)(modN), (1 \leq m \leq N)$. For $h_i = i$, we can write that $mh_i = Nl + p$, where l and p are positive integers with $0 < p \leq N$, then $mh_{N-i} = m(N-i) = mN - mi$, that means, $mh_{N-i} \equiv -mi(modN)$, where

$$-mi = -(Nq+p) = N(-q-1) + (N-p) \equiv \begin{cases} N-p, & \text{if } p \neq N \\ N, & \text{if } p = N \end{cases}$$

Therefore, for any positive integer $m, 1 \leq m \leq N$,

$$mh_i(modN) + mh_{N-i}(modN) = \begin{cases} p + (N-p) = N, & p \neq N \\ p + N = N + N = 2N, & p = N \end{cases}$$

Performing elementary transformation of columns of the matrix of $U_N(N^t)$ by adding the i-th ($i = 1, 2, ..., k$) column of matrix $U_N(N^t)$ to its $N - i$-th column, respectively. We get k new columns which are all same as $(\underbrace{N, N, ..., N}_{N-1}, 2N)'$, so, the rank of the new matrix is less than or equal to $\frac{t}{2} + 1$. Therefore, the rank of matrix $U_N(N^t)$ also is less than or equal to $\frac{t}{2} + 1$.

(2) N is a non-prime number. In this situation, we have $(h_1, h_2, ..., h_t)$, the set of all positive integers which satisfies the conditions of g.c.d. $(N, h_i) = 1$ and $h_i < N$ for $i = 1, ..., t$ and $t = \phi(N)$, as first row of the $U_N(N^t)$. Obviously, the $t < N - 1$.

Since $(N, h_i) = 1$ for any number h_i in the $(h_1, h_2, ..., h_t)$, and the necessary and sufficient condition for $(N, h_i) = 1$ is that there are two integers a and b which satisfy $aN + bh_i = 1$. It is follows that $(a+b)N + (-b)(N - h_i) = aN + bh_i = 1$, that is, $(N, N - h_i) = 1$. Thus, $N - h_i$ also is in the set $(h_1, h_2, ..., h_t)$. Therefore, as in case (1), we can prove that the rank of matrix $U_N(N^t)$ is less than or equal to $\frac{t}{2} + 1$.

(3) When N is even. The matrix of the UG-design could be either as a special case of (2) or as a remainder of $U_{N+1}[((N+1)^t]$. In the second situation, the N rows of the matrix are the first N rows of $U_{N+1}[(N+1)^t]$ and denoted as $U_N^*(N^t)$. Using same procedures as in cases (1) and (2) to number $N + 1$, we found that after same elementary transformation of columns the k new

columns in the first N rows of the $U_{N+1}[(N+1)^t]$ are all $(\underbrace{N, N, ..., N}_{N})'$, so the rank of matrix $U_N(N^t)$ is still less than or equal to $\frac{t}{2}+1$, where $t = \phi(N+1)$. The theorem is proved.

Corollary 3.2 *For regression analysis, at most* $\frac{t}{2}+1$ *factors can be arranged by UG-type design* $U_N(N^t)$ *or* $U_N^*(N^t)$.

From now on, for simplicity, we still use the notations of matrices $U_N(N^t)$ and $U_N^*(N^t)$ provided by Fang (1980), Wang and Fang (1981), but no more than $\frac{t}{2}+1$ columns of the t columns in the both matrices are effective for use under regression model.

Furthermore, according to the author's results of computation for $4 \leq N \leq 19$ (see Appendix A), the ranks of matrices $U_N(N^t)$ and $U_N^*(N^t)$ both equal to $\frac{t}{2}+1$.

3.2.2 UG-type Designs with Uniformity and Optimalities

Let $\mathcal{P} = \{\mathbf{x}_1, ..., \mathbf{x}_N\}$ be some set of N points on unit cube C^s. Generating a uniform design requires a measure of uniformity. Several useful measures for generating an uniform design have been proposed (Fang and Hickernell, 1995). To find out the G-uniform design we need to find out t columns with some lowest measure D for the number of factors $s = 2, ..., \frac{t}{2}+1$. We introduce following *discrepancy* as the measure of uniformity:

Definition 3.4 Let $F_N(\mathbf{x})$ be the empirical distribution of $\mathcal{P}$, that is,

$$F_N(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N} I\{\mathbf{x}_i \leq \mathbf{x}\},$$

where $I\{.\}$ is the indicator function and all inequalities are understood to be with respect to the componentwise order of R^s. Then

$$D_p(\mathcal{P}) = [\int_{C^s} |F_N(\mathbf{x}) - F(\mathbf{x})|^p \, d\mathbf{x}]^{1/p}$$

is called L_p-discrepancy of $\mathcal{P}$, where $F(\mathbf{x})$ is the uniform distribution function on C^s. Setting $p = 1, 2$ and ∞ give the L_1-, L_2-, and L_∞-discrepancies, respectively.

The L_∞-discrepancy is usually simply called the *discrepancy*, the more popular measure in UD. It can be written as

$$D(\mathcal{P}) = \sup_{\mathbf{x} \in C^s} |F_N(\mathbf{x}) - F(\mathbf{x})|.$$

In fact, the discrepancy is the Kolmogorov-Smirnov statistic for goodness-of fit. One disadvantage of the discrepancy is that it is expensive to compute. Bundschuh and Zhu (1993) gave an algorithm suitable for a small number of factors. A better approach is used by Winker and Fang (1995).

What are the statistical advantages of the G-uniform design based on above discrepancy except the uniformity of its experimental points on the domain (Ding, 1986)? In other words, can we find a class of designs with some statistical optimalities while keeping the uniformity?

Recall the definition of the G-optimality and compare it with above concept of the *discrepancy*, their statistical meanings seem to be related. Because of the equivalence of G-optimal and D-optimal designs, according to

the Kiefer and Wolfowitz (1960) theorem, a special result of the general equivalence theorem for optimum designs (see Kiefer, 1974), it may not be surprising if optimality and uniformity are satisfied at the same time in many cases.

Now, assuming

$$\mathcal{C} = \{C_{\mathcal{P}}, \mathcal{P} \in \mathcal{D}\} \subseteq \mathcal{B}_s, \quad \mathcal{D} \triangleq \mathcal{UG}_{N,s},$$

we denote a design in $\mathcal{D}$ with uniformity as $\mathcal{P}_U$, a design with D-optimality as $\mathcal{P}_D$, a design with A-optimality as $\mathcal{P}_A$, and a design with both A-optimality and uniformity as $\mathcal{P}_{AU}$, etc., respectively.

Suppose that we use an UG-type design $\mathcal{P}$ to arrange an experiment with $s \leq \frac{t}{2}+1$ factors having N levels each under following linear regression model (Ding, 1986):

$$\begin{aligned} y_\alpha &= \beta_0 + \beta_1(x_{\alpha 1} - \bar{x}_1) + ... + \beta_s(x_{\alpha s} - \bar{x}_s) + \varepsilon_\alpha \\ &= (\beta_0 - \beta_1\bar{x}_1 - ... - \beta_s\bar{x}_s) + \beta_1 x_{\alpha 1} + ... + \beta_s x_{\alpha s} + \varepsilon_\alpha, \end{aligned} \tag{3.1}$$

where $\bar{x}_i = \frac{1}{N}\sum_{\alpha=1}^{N} x_{\alpha i}, i = 1, ..., s, \alpha = 1, ..., N$, the design matrix

$$X_0 = \left(\begin{array}{c|ccc} 1 & x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1s} - \bar{x}_s \\ 1 & x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2s} - \bar{x}_s \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{N1} - \bar{x}_1 & x_{N2} - \bar{x}_2 & \cdots & x_{Ns} - \bar{x}_s \end{array}\right) \triangleq (\mathbf{1}|X_{N\times s}),$$

the information matrix

$$C_{\mathcal{P}} = X_0'X_0 \triangleq \begin{pmatrix} N & 0 \\ 0 & L_{s\times s} \end{pmatrix}.$$

It is easy to see the advantage of the model (3.3) which decreases computing load of the related traces and determinants by reducing the order of the

$C_{\mathcal{P}}$-matrix, i.e., by $C_{\mathcal{P}}^{-1} = \begin{pmatrix} N^{-1} & 0 \\ 0 & L_{s\times s}^{-1} \end{pmatrix}$. Therefore, we can comparing $Tr(L^{-1})$ (or $det(L^{-1})$) instead of compare $Tr(C_{\mathcal{P}}^{-1})$ (or $det(C_{\mathcal{P}}^{-1})$). Due to this reason we may also simply call the X and L as design matrix and information matrix, respectively.

In appendix A, we have given $\mathcal{P}_{ADU}$, $\mathcal{P}_{DU}$, $\mathcal{P}_{AD}$ with Fang's $\mathcal{P}_U$ for the UG-type design for $N < 20$. From this appendix one may find indeed that most of the designs provide A- and D-optimality while keeping the uniformity.

3.2.3 Uniform Design Related to an Orthogonal Array

It is possible to relate the UD to the orthogonal array by using the pseudo-level technique. Suppose that we want to use a design with N levels in a practical problem with only q interesting levels when $N = rq$ for some integer r. We may reduce the original N levels to q levels by *pseudo-level technique* as follows: merge the first r levels of the N levels into level 1, the $(r+1)$-th to $(r+r)$-th levels of the N levels into level 2, and so forth. For example, applying this pseudo-level technique to a U-type design $U_9(9^4)$ in Table 3.1, we get an orthogonal array $L_9(3^4)$ in Table 3.2.

In general, Fang et al (1995) gives the following theorem in this connection.

Theorem 3.3 *Any orthogonal array $L_N(q^t)$, if it exists, can be generated from a U-type design $U_N(N^t)$ by the use of the pseudo-level technique, assuming N is a multiple of q.*

Figure 3.1 shows the plots of first two columns of the Table 3.1 and 3.2. The 9 circle points from $L_9(3^4)$ lie on a 3×3 grid. By shifting each of these points onto a finer 9×9 grid, we can obtain 9 star points that form a $U_9(9^4)$ design which may have lower discrepancy than $L_9(3^4)$. One may think of this as the reverse of the pseudo-level technique. Therefore, Fang (1995) proposed the following U-type design $U_N(N^t)$ based on the OA.

Definition 3.5 A $O_{N,t,q}$ design $U_N(N^t)$ is a U-type design that can be reduced to an orthogonal design $L_N(q^t)$ by the pseudo-level technique where N is some multiple of q. The design with lowest D-value is called an $O_{N,t,q}$-uniform design and denoted by $UO_{N,t,q}$.

For a given $L_N(q^t)$ there are many possible $O_{N,t,q}$ designs $U_N(N^t)$. In each column of $L_N(q^t)$ we can replace level $k(1 \leq k \leq q)$ by any permutation of $(k-1)r+1, (k-1)r+2, ..., kr$. A search over possible permutations yields a $O_{N,t,q}$ uniform design. Fang (1995) found that $O_{N,t,q}$ uniform designs often have lower discrepancy than G-uniform design and OA. Table 3.3 compares the discrepancy for these three types of designs for $N = 9$ and $t = 2, 3$ and 4. But Fang only obtained the $O_{N,t,q}$ uniform designs for a few cases of small N and t. Furthermore, the relationships of the UD and OD has not been fully studied.

Table 3.1 $L_9(3^4)$

No.	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Table 3.2 $U_9(9^4)$

No.	1	2	3	4
1	1	2	1	3
2	2	5	4	5
3	3	9	8	7
4	4	3	6	9
5	5	4	7	1
6	6	7	2	6
7	7	1	9	4
8	8	6	3	8
9	9	8	5	2

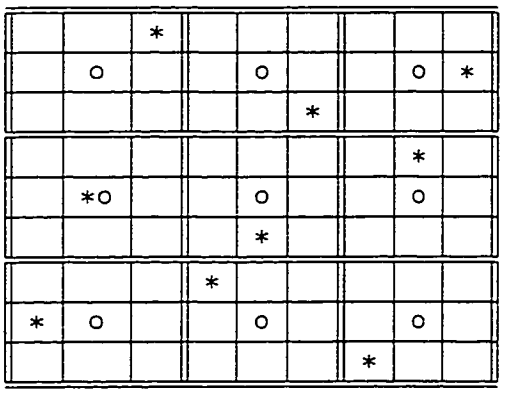

Figure 3.1. Comparison of the experimental points $L_9(3^4)$ (o) and $U_9(9^4)$ (*).

Table 3.3 Discrepancies for $L_9(3^s)$, $UG_{9,s}$ and $UO_{9,s,3}$ designs

t	$L_9(3^s)$	$UG_{9,s}$	$UO_{9,s,3}$
2	0.3056	0.1574	0.1451
3	0.4213	0.2885	0.2114
4	0.5177	0.4066	0.2739

3.3 Uniform Designs with Different Number of Levels

Although the uniform design has the advantage of requiring fewer experiments with large number of levels, it is usual that some of the factors must be restricted in a certain number of levels. For example, the only possible numbers of seats in a car are 4, 5, or 6. Therefore, we need the uniform design with different numbers of level for different factors. In fact, Fang and Li (1994) generate a number of designs with various numbers of levels derived from *glp* sets using the pseudo-level technique.

Example 3.1 For redesigning its checking account services, a financial institution may consider four attributes with different numbers of level listed in Table 3.4.

We may choose $U_6(6^3 \times 3)$ (Table 3.5), the G-uniform design, to arrange the four factors, where the factors A(cost of checking account), C(accessibility to banking service), D(hours) have six levels each while factor B(quality of service) has three levels. By linear regression model, 5-4(1)=1 degree is for residual, and we can estimate single part-worth, or say, regression coefficient.

Table 3.4 Four attributes in a design of financial service

	Attribute name		Attribute levels
A.	Cost of checking account	1.	$0.60 a check, monthly statement along with the checks processed
		2.	First six checks each month are free, subsequent checks are $0.85 each
		3.	First two checks each month are free, subsequent checks are $1.25 each, earn interest on every dollar.
		4.	$1,000 minimum balance in a checking or savings account all the time
		5.	Free checks by keeping $1,000 minimum monthly balance, otherwise, $0.60 a check
		6.	Absolutely free checking service
B.	Quality of service	1.	Service is less friendly than average. Bank personnel are less likely to help you than those in most banks.
		2.	Service is average in friendliness. Bank personnel are average in their willingness to help you.
		3.	Service is above average in friendliness. Bank personnel are more likely to help you than those in most bank.
C.	Accessibility to banking service	1.	20-minute drive from home
		2.	15-minute drive from home
		3.	10-minute drive from home
		4.	5-minute drive from home
		5.	10-minute walk from home
		6.	5-minute walk from home
D.	Hours	1.	Weekdays 9am-4pm
		2.	Weekdays 9am-4pm and evenings 4pm-8pm twice a week
		3.	Weekdays 9am-4pm and Saturdays 9am-12:00noon
		4.	Weekdays 9am-5pm
		5.	Weekdays 9am-5pm and Saturdays 9am-12:00noon
		6.	Monday through Saturday 9am-4pm.

Table 3.5 G-uniform design $U_6(6^3 \times 3)$

	A	C	D	B
	1	2	3	4
1	1	2	3	2
2	2	4	6	1
3	3	6	2	3
4	4	1	5	1
5	5	3	1	3
6	6	5	4	2
Discrepancy	0.3581			

Chapter 4

Conclusions and Directions for Further Research

4.1 Concluding Comments

It is no exaggeration to say that a necessary condition for a new method or model to receive widespread application is for it to be convenient and easy-to-use. By summarising the advantages of uniform designs, we can see that the uniform designs provide many more flexibilities and economies for CA:

1. Uniform design is a new alternative tool for coping with large numbers of attributes and levels within attribute. For example, one may arrange large numbers of attributes, say more than 6 or 10, and large numbers of levels up to 31.

2. Uniform designs use the regression model which estimates the single coefficient for every attribute so that the number of combinations (stimulus) is reduced dramatically. e.g., it falls from OA's $N \geq q^2$ to UD's $N = q$. It can be a single part-worth (i.e., regression coefficient of linear term) that is multiplied by the level's value, and the quadratic terms,

the interaction terms (the coefficient of cross term), etc. It allows us to predict the total worth.

3. It is easier to use various existing software of the regression for the analysis of UD, where no more tools are required.

4. When the true environmental correlations are high, uniform design in conjoint analysis may avoid some infeasible attribute combinations (e.g., low-priced luxury cars; very powerful, low-wattage air condition).

5. UD also can provide some designs with mixed levels, that is, UD allows different attributes to have different number of levels, especially, while few attributes may have larger number of levels and others may have less.

On other hand, the PAD also can be applied in a variety of situations of CA:

1. There is one factor which includes more levels than other factors but there does not exist a suitable orthogonal array to use.

2. The number of experimental combinations must be smaller than the existing OA for some reasons.

3. The survey needs to be separated by two stage designs, or after completing a set of experiments, some extra levels in one of the factors need to be compared statistically with former levels of the factor.

Universally optimal PADs and UG-type optimal designs are proposed as new designs for conjoint analysis in this thesis. Like any new technique in its infancy, there is a lot of room for improvement. First, coping with

larger numbers of attributes and levels with different levels still is a big challenge. For example, when two attributes have twelve levels while the remaining two or more than two attributes have three levels, we have not yet provided related design, such as UG-type optimal designs or PAD. Second, Designs as described in this thesis are used for fixed effects linear model and $Var(Y) = \sigma^2 I$, the questions are whether the same optimality holds when the $Var(Y) = \sigma^2\Sigma$ for $\Sigma > 0$ or under mixed model?

In section 4.2 and 4.3, we will briefly discuss optimality for experimental designs under mixed model, and highlight two new analytic methods for CA under mixed model.

4.2 Optimal Designs in Mixed Model

In CA, it is not unusual that the levels of some attributes come from a large population and our interest is in this population, not in specific levels. Then, we need mixed model:

$$Y = X\beta + \varepsilon, \quad \varepsilon = U_1\xi_1 + ... + U_r\xi_r = U\xi,$$

where Y is an n-vector of observable random variables, X, the design matrix, and $U_i, i = 1, ..., r$, are known matrices, β is an m-vector of fixed parameters, and $\xi_i, i = 1, ..., r$ are unobservable structural random variables such that

$$\xi_i \sim (0, \sigma^2 I_{ni}), \quad cov(\xi_i, \xi_j) = 0, i \neq j,$$

and

$$\varepsilon \sim (0, \sigma_1^2 V_1 + ... + \sigma_r^2 V_r), \quad V_i = U_i U_i'.$$

When some attributes are considered random variables, we have one or more additional sources of information for estimating treatment effects. If the

analysis is based on an appropriate mixed model, in most cases, we shall attempt to show that a design which is optimal under the fixed effects model continues to remain so under the mixed model. These results are non-trivial because the comparative position of two given designs with respect to a given criterion such as A- or D-optimal may not remain the same under the two models (Shah and Sinha, 1989).

It is important to recognize that the analysis of a mixed model is based on the estimation of the variance components and that the estimates of the treatment effects are affected by the errors in the estimation of these variance components. Shah and Sinha (1989) ignore these errors, i.e., they regard the variance component as known up to a scalar multiplier for the purpose of optimality. Then the primary interest is in optimality results which hold for all values of the unknown variance components. Thus, the optimality with respect to the fixed effects model is indeed an apriori necessary condition for optimality with respect to the mixed model. For this reason, even though the designs which are disconnected in the fixed effects model are connected in the mixed model, these could not be optimal for all values of the variance components and hence need not be considered.

Optimality of block designs under the mixed model was first considered by Sinha (1980). Then, several authors established optimality of certain designs. For instance, the *Balanced Block Designs* (BBDs) is universally optimal in the extended sense with respect to the mixed model. Work in this area is still in the early stages and many important problems yet remain unresolved (Shah and Sinha, 1989). Little is currently known about the optimality based on factorial and fractional factorial designs in the mixed model.

4.3 Disaggregate Level Analysis

We propose a mixed model as an alternative of part-worth function model, which involves limiting assumption of fixed effects. Two examples are used to illustrate how the results based on ANOVA II (mixed) model are congruent with that of traditional procedure.

Example 4.1 One may try to decide what attributes affect choice in the canned dog food market (Hair, Jr. et al, 1992). The three attributes (factors) selected with two values (levels) each as affecting the purchase decision are as follows:

Factor	**Level**
Brand name	Arf versus Mr.Dog
Ingredients	All meat versus meat and fiber supplements
Can size	6 ounce versus 12 ounce

Following table lists the range of the part-worth estimate based on respondent 1 (Hair, Jr. et al, 1992) and the order of importance.

	size	ingredients	brand
range	1.512	3.022	0.756
%	29	57	14
order	2	1	3
MINQUE	2.00	8.00	0.5

Suppose we consider these three factor effects as random effects and run a variance component analysis. Estimates of these three factors are displayed in the last lines of the table. The order of importance for three factors is exactly same as the order in part-worth analysis, i.e. ingredients, size and brand.

Example 4.2 In Chapter 2, we compared a PAD with a complete design based on the simulation study data set. Here we assume all the six factors are random. Table 4.1 shows the estimates of variances.

Table 4.1 The results of variance component analysis
for the study of Nandakumar et al

Source	E(MS)	Estimate	Order
A (Corrtype)	Var(e)+112 Var(Corrtype)	11.44	6
B (Item)	Var(e)+ 32 Var(Item)	97.59	4
C (J)	Var(e)+112 Var(J)	46.47	5
D (Loading)	Var(e)+112 Var(loading)	101.59	3
E (N)	Var(e)+112 Var(N)	154.93	2
F (ρ)	Var(e)+112 Var(ρ)	455.20	1
error	Var(e)+112 Var(error)	29.49	

Comparing the six estimates of variances with error variance, only the variance of factor A, i.e. correlation type, is less than error variance. This means that we may ignore the difference between these two types of correlation in terms of their power. This conclusion is almost the same as Nandakumar et al (1995) (see Table 2.4). According to the order of variances estimates, the importance of the factors, are: ρ, N, loading, item and J.

4.4 Aggregate Level Analysis

The general conjoint segmentation model is formulated by Hagerty (1985) as:

$$YP(P'P)^{-1}P' = XB + E$$

where

i = the consumers, $1, 2, ...N$;

j = the profiles, $1, 2, ...n$;

k = the conjoint design variables, $1, 2, ...K$;

s = the segments, $1, 2, ...S$;

X = a $(n \times K)$ matrix containing the K conjoint design dummy variables for the n profiles;

Y = a $(N \times S)$ matrix containing the responses of the N consumers to the n profiles;

P = a $(N \times S)$ matrix representing a general partitioning scheme for assigning consumers to segments;

B = a $(K \times N)$ matrix of regression coefficients; and

E = a $(n \times N)$ matrix of random error (Vriens, *et al* 1996).

This equation accommodates individual-level analysis ($S = N$), segment-level analysis ($1 < S < N$), and aggregate-level analysis ($S = 1$). Since economists have generally been most interested in the aggregate implications of multi-attribute utility structures and less considered with estimation of individual utility function per se (Green et al, 1978), we shall use a general method for analyzing data of covariance structures at the aggregate level in CA.

Joreskog (1974) describes a general method for analyzing data according to a general model of covariance structures. The model assumes that the population variance-covariance matrix $\mathbf{\Sigma}(p \times p) = (\sigma_{ij})$ of a set of variables has the form

$$\mathbf{\Sigma} = \mathbf{B}(\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{'} + \mathbf{\Psi^2})\mathbf{B}' + \mathbf{\Theta^2}, \tag{4.1}$$

where $\mathbf{B}(p \times p) = (\beta_{ik}), \mathbf{\Lambda}(q \times r) = (\lambda_{km})$, the symmetric matrix $\mathbf{\Phi}(r \times r) =$

(φ_{mn}), and the diagonal matrices $\mathbf{\Psi}(q \times q) = (\delta_{kl}\psi_k)$ and $\mathbf{\Theta}(p \times p) = (\delta_{ij}\theta_i)$ are parameter matrices (δ_{ij} denotes Kronecker's delta). It is assumed that the mean vector of the variables is unconstrained so that the information about the covariance structure is provided by usual sample variance-covariance matrix $\mathbf{S}(p \times p) = (s_{ij})$, which may be taken to be a correlation matrix if the model is scale free and if the units of measurements in the variables are arbitrary or irrelevant.

The covariance structure of equation (4.1) arises when the observed variables $\mathbf{x}(p \times 1)$ are of the form

$$\mathbf{x} = \mu + \mathbf{B\Lambda}\xi + \mathbf{B}\zeta + \mathbf{e}, \tag{4.2}$$

where $\xi(r \times 1)$, $\zeta(q \times 1)$, and $\mathbf{e}(p \times 1)$ are uncorrelated random latent vectors, in general unobserved, with zero mean vector and dispersion matrices $\mathbf{\Phi}$, $\mathbf{\Psi}^2$, and $\mathbf{\Theta}^2$, respectively, and where μ is the mean vector of $\mathbf{x}$.

To determine the estimates of the unknown parameters, two different methods of fitting the model to the observed data may be used. One is the *generalized least squares method* (GLS) that minimizes

$$\mathbf{G} = tr(\mathbf{I} - \mathbf{S}^{-1}\mathbf{\Sigma})^2; \tag{4.3}$$

the other is the *maximum-likelihood method* (ML) that minimizes

$$\mathbf{M} = log|\mathbf{\Sigma}| + tr(\mathbf{S\Sigma}^{-1}) - log|\mathbf{S}| - p. \tag{4.4}$$

In a large sample of size N, $(N-1)$ times the minimum value of $\mathbf{G}$ or $\mathbf{M}$ may be used as a χ^2 to test the goodness of fit of the model and for both methods, approximate standard errors may be obtained for each estimated

parameter by computing the inverse of the information matrix.

Wiley, Schmidt and Bramble (1973) suggested the study of a general class of components of covariance models. This class of models is a special case of Equation (4.1), namely, when $\mathbf{B}$ is diagonal, $\mathbf{\Lambda}$ is known a priori, $\mathbf{\Phi}$ is symmetric and positive definite, and $\mathbf{\Psi}$ or $\mathbf{\Theta}$ are either zero or diagonal. The covariance matrix $\mathbf{\Sigma}$ will then be of the form

$$\mathbf{\Sigma} = \mathbf{\Delta A \Phi A' \Delta} + \mathbf{\Theta^2} \tag{4.5}$$

or

$$\mathbf{\Sigma} = \mathbf{\Delta}(\mathbf{A \Phi A'} + \mathbf{\Psi^2})\mathbf{\Delta}. \tag{4.6}$$

The matrix $\mathbf{A}(p \times k)$ is assumed to be known and gives the coefficient of the linear functions connecting the manifest and latent variables, $\mathbf{\Delta}$ is a $p \times p$ diagonal matrix of unknown scale factors, $\mathbf{\Phi}$ is the $k \times k$ symmetric and positive definite covariance matrix of the latent variables and $\mathbf{\Psi^2}$ and $\mathbf{\Theta^2}$ are $p \times p$ diagonal matrix of error variances.

Within this class of models eight different special cases are of interest. These are generated by the combination of the following set of conditions:

$$\begin{aligned}
\text{on}\mathbf{\Delta}: &\quad \left\{ \begin{array}{ll} \mathbf{\Delta} & = \mathbf{I} \\ \mathbf{\Delta} & \neq \mathbf{I}; \end{array} \right\} \\
\text{on}\mathbf{\Phi}: &\quad \left\{ \begin{array}{ll} \mathbf{\Phi} & \text{is diagonal} \\ \mathbf{\Phi} & \text{is not diagonal;} \end{array} \right\} \\
\text{on}\mathbf{\Psi^2}\text{or}\mathbf{\Theta^2}: &\quad \left\{ \begin{array}{ll} \mathbf{\Psi^2}\text{or}\mathbf{\Theta^2} & = \sigma^2\mathbf{I} \\ \mathbf{\Psi^2}\text{or}\mathbf{\Theta^2} & \text{general diagonal.} \end{array} \right\}
\end{aligned}$$

Selection of these eight combinations depends on the scale of measurement and the homogeneity of variance.

Joreskog (1974) introduced an example taken from Wiley, Schmidt and Bramble (1973). This example was reanalyzed through LISREL. The data consist of 51 education students on a test designed to assess teachers' judgements. The items used were designed according to three factors: Grade Level (G), Teacher Approach (T) and Teaching Method (M), which were hypothesized to influence learning and teaching. We use this example to CA for purpose of explaining basic idea associated with aggregate analysis. There are two levels for each factor. In this case, the observed variables $\mathbf{x}(p \times p)$ are of the form:

$$\mathbf{x}_\nu = \mu + \mathbf{A}\mathbf{u}_\nu + \mathbf{e}_\nu, \tag{4.7}$$

where μ is the mean vector and $\mathbf{e}_\nu$ is a random error vector both of the same form as $\mathbf{x}_\nu$. The variance-covariance matrix of $\mathbf{x}_\nu$ is

$$\mathbf{\Sigma} = \mathbf{A}\mathbf{\Phi}\mathbf{A}' + \mathbf{\Theta}^2, \tag{4.8}$$

where $\mathbf{A}$ is a design matrix based on factorial designs which has one random way of classification $\nu = 1, 2, ..., N$, for aggregate level analysis in CA.

A 2^3 factorial design

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{pmatrix}.$$

can be used. The $N = 51$ respondents would then be asked either to rank or rate eight stimuli.

In equation (4.8), $\mathbf{\Phi}$ is general diagonal, or $\mathbf{\Phi} = (\varphi_{ij})$ is a square matrix, and $\mathbf{\Theta}^2$ is general diagonal or $\mathbf{\Theta}^2 = \sigma^2\mathbf{I}$. By the sample covariance matrix $\mathbf{S}$ for these data, $\Delta = I$ was considered appropriate. Then the data were analyzed under each of the four remaining cases (Joreskog, 1974). Table 4.2 shows the χ^2 values and their degrees of freedom. In the margin of this table, χ^2 values are given for testing the hypotheses of uncorrelated components and of homogeneous error variances. It appears that these hypotheses were both rejected by the tests. The only model that seems reasonable is the one that assumes both correlated components and heterogeneous error variances. The maximum-likelihood estimates of variances and covariances of the components and their error variance, together with their standard errors, are also given in the Table. The value 9.16 in $\hat{\Phi}$ shows that the variance of subjects is the largest one and factor M (teaching method) is another major source (5.21).

Table 4.2 χ^2 for testing the fit of four models

	$\Theta^2 = \sigma^2 I$	$\Theta^2 \neq \sigma^2 I$	
Φ diagonal	$\chi^2_{31} = 68.25$	$\chi^2_{24} = 46.16$	$\chi^2_7 = 22.09$
Φ not diagonal	$\chi^2_{25} = 51.00$	$\chi^2_{18} = 25.98$	$\chi^2_7 = 25.02$
	$\chi^2_6 = 17.25$	$\chi^2_6 = 20.18$	

Maximum-Likelihood Estimates of ϕ and θ^2

(Standard Errors in Parentheses)

$$\hat{\mathbf{\Phi}} = \begin{bmatrix} 9.16(1.95) & & & \\ 0.75(0.48) & 0.70(0.34) & & \\ 0.63(0.43) & -0.05(0.33) & 0.43(0.91) & \\ -0.62(1.10) & -0.51(0.81) & 1.13(0.51) & 5.21(1.58) \end{bmatrix}$$

$$\theta^2 = diag(1.52(0.83), 4.95(1.41), 8.25(1.88), 5.58(1.60), 1.95(0.96), 5.76(1.21), 2.52(0.92)).$$

Although the data is from an educational study, one of the purposes of the study is to identify the most important factor(s) which is/are the same as in conjoint analysis. The author suggests applying structural equation modelling (SEQ) or structural analysis of covariance matrices to aggregate level analysis in conjoint analysis.

Appendix A

UG-type Designs with Uniformity and Optimalities

Name of design	No. of factors	Effective columns under giving number of factors			
		$\mathcal{P}_{ADU}$	$\mathcal{P}_{DU}$	$\mathcal{P}_{AD}$	$\mathcal{P}_{U}$
$U_4^*(4^4)$	2	1,2			
$U_5(5^4)$	2			1,4	1,2
	3	1,2,4			
$U_6^*(6^6)$	2	1,3			
	3	1,2,3			
$U_7(7^6)$	2	1,3			
	3	1,2,3			
	4		1,2,3,6	1,2,4,6	
$U_8^*(8^6)$	2	1,3			
	3	1,3,5			
$U_9(9^6)$	2	1,3			
	3	1,3,5			
	4		1,2,3,5	1,2,4,5	
$U_{10}^*(10^{10})$	2	1,7			
	3		1,5,7	1,3,5	
	4	1,2,5,7			
	5	1,2,3,5,7			
$U_{11}(11^{10})$	2			1,5	1,7

Name of design	No. of factors	Effective columns under giving number of factors			
		$\mathcal{P}_{ADU}$	$\mathcal{P}_{DU}$	$\mathcal{P}_{AD}$	$\mathcal{P}_{U}$
$U_{11}(11^{10})$	3	1,5,7			
	4	1,2,5,7			
	5	1,2,3,5,7			
	6		1,2,3,5,7,10	1,2,4,5,8,10	
$U_{12}^{*}(12^{12})$	2	1,5			
	3	1,3,9			
	4	1,6,8,10			
	5	1,6,8,9,10			
	6	1,2,6,8,9,10			
$U_{13}(13^{12})$	2			1,6	1,5
	3			1,6,10	1,3,4
	4	1,6,8,10			
	5	1,6,8,9,10			
	6	1,2,6,8,9,10			
	7		1,2,6,8,9, 10,12	1,2,3,4, 6,8,12	
$U_{14}^{*}(14^{8})$	2	1,6			
	3	1,3,4			
	4	1,3,4,7			
$U_{15}(15^{8})$	2	1,6			
	3	1,3,4			
	4	1,3,4,7			
	5		1,2,3,4,7	1,2,3,5,7	
$U_{16}^{*}(16^{16})$	2			1,4	1,10
	3	1,10,15			
	4		1,10,14,15		
	5			1,2,4,5,11	1,4,10,14,15
	6			1,2,4,5,10,11	1,4,6,10,14,15
	7	1,4,6,9,10,14,15			
	8	1,4,5,6,9,10,14,15			
$U_{17}(17^{16})$	2			1,11	1,10
	3	1,10,15			
	4	1,10,14,15			

Name of design	No. of factors	Effective columns under giving number of factors			
		$\mathcal{P}_{ADU}$	$\mathcal{P}_{DU}$	$\mathcal{P}_{AD}$	$\mathcal{P}_{U}$
$U_{17}(17^{16})$	5	1,4,10,14,15			
	6	1,4,6,10,14,15			
	7	1,4,6,9,10,14 15			
	8	1,4,5,6,9 10,14,15			
	9		1,4,5,6,9, 10,14,15,16	1,3,5,9,10 11,13,15,16	
$U_{18}^{*}(18^{18})$	2	1,8			
	3	1,7,8			
	4	1,6,8,14			
	5		1,6,8,14,17	1,2,4,8,16	
	6			1,6,7,8, 10,14	1,6,8,10, 14,17
	7			1,2,4,7,8, 13,16	1,6,7.8, 10,14,17
	8	1,3,6,7,8,10 14,17			
	9	1,3,4,6,7,8 10,14,17			
$U_{19}(19^{18})$	2			1,6	1,8
	3			1,6,14	1,7,8
	4	1,6,8,14			
	5	1,6,8,14,17			
	6	1,6,8,10,14,17			
	7	1,6,7,8,10,14,17			
	8	1,3,6,7,8,10, 14,17			
	9	1,3,4,6,7,8, 10,14,17			
	10		1,3,4,6,7,8, 10,14,17,18	1,3,5,6,10,11, 12,15,17,18	

Bibliography

[1] Allenby, G. M., Arora, N. and Ginter, J. L. (1995), Incorporating prior knowledge into the analysis of conjoint studies, *Journal of Marketing Research*, **32** (May), 156-62.

[2] Batsell, R. R. and Louviere, J. J. (1991), Experimental analysis of choice, *Marketing Letters*, **2** (August), 199-214.

[3] Blais, J., and Laurier, M. (1995), Methodological considerations in using DIMTEST to assess unidimensionality. Paper presented at annual meeting of the American Educational Research Association, San Francisco, CA.

[4] Bundschuh, P. and Zhu, Y. C. (1993), A method for exact calculation of the discrepancy of low-dimensional finite point sets (I), Abhandlungen aus dem Math. Seminar der Univ. Hamburg, Bd. 63.

[5] Camilli, G. (1995), The effects of dimensionality on equating the law school admission test, *Journal of Educational Measurement*, **32**, 79-96.

[6] Carroll, J. D. (1969), *Categorical conjoint measurement*, paper presented at Meeting of Mathematical Psychology, Ann Arbor, MI, (August).

[7] Carroll, J. D. (1973), Models and algorithms for multidimensional scaling, conjoint measurement, and related techniques, in *Multiattribute Decisions in Marketing*, Green, P. E., Wind, Y. eds. Hinsdale, IL: Dryden Press, 335-37; 341-48.

[8] Carroll, J. D. and Green, P. E. (1995), Psychometric methods in Marketing Research: Part I, Conjoint Analysis, *Journal of Marketing Research*, **32**, 385-391.

[9] Carson, R. T. et al. (1994), Experimental analysis of choice, *Marketing Letters*, **5** (October), 351-68.

[10] Cattin, P. and Wittink, D. R. (1982), Commercial use of conjoint analysis: A survey, *Journal of Marketing*, **46**, 44-53.

[11] Cheng, C. S.(1978), Optimal designs for the elimination of multi-way heterogeneity, *Ann. Statist.* **6**, 1262-1272.

[12] Cheng, C. S. (1980), Orthogonal array with variable numbers of symbols, *Ann. Statist.* **8**, 447-453.

[13] Clarke, D. G. (1987), *Marketing Analysis and Decision making: Text and Cases with Lotus 123*, Redwood City, Calif.: Scientific Press.

[14] Davidson, J. D. (1973), Forecasting traffic on STOL, *Operational Research Quarterly*, 561-69.

[15] DeSarbo, W. S.,Oliver, R. L. and Rangaswamy, A. (1989), A simulated annealing methodology for clusterwise linear regression, *Psychometrika*, **54 (4)**, 707-36.

[16] Ding, Y., Lin, J. G. and Zhou, J. X. (1975), *Methods of Orthogonal Experimental Designs*, Shanghai: Shanghai People's Press.

[17] Ding, Y. (1986), An exploration of optimality of uniform design, *Journal of Applied Probability and Statistics*, **2**, 153-160.

[18] Ding, Y. and Chaubey, Y. P. (1997), Optimality of partially augmented designs, *Technique Report*, **2**, Dept. of Math. and Stat., Concordia University.

[19] Donoghur, Holland, and Thayer (1995), A Monte Carlo study of factors that affect the Mantel-Haenszel and standardization measures of differential item functioning. In P. W. Holland and H. Wainer (Ed.), Differential Item Functioning (pp137-166), New Jersey: Hillsdale.

[20] Fang, K. T. (1980), The uniform design: application of number-theoretic methods in experimental design, *Acta Math. Appl. Sinica*, **3**, 363-72.

[21] Fang, K. T. and Li, J. K. (1994), Some new results on uniform design, *Chinese Science Bulletin*, **39**, 1921-4.

[22] Fang, K. T. and Hickernell, F. J. (1995), *The uniform design and its applications*, paper presented at meeting of ISI, 26-42.

[23] Fang, K. T. and Wang, Y. (1994), *Number-theoretic Methods in Statistics*, Published by Chapman & Hall, London.

[24] Geramite, A. V. and Seberry, J. (1979), *Orthogonal Designs, Quadratic forms and Hadamard matrix*, Marcel Dekker Inc. NY.

[25] Gessaroli, M. and De Champlain, A. (1996), Using an approximate Chi-square statistic to test the number of dimensions underlying the responses to a set of items, *Journal of Educational Measurement*, **33**, 157-179.

[26] Goldberg. S. M., Green, P. E. and Wind, Y. (1984), Conjoint analysis of price premiums for hotel amenities, *Journal of Business*, **57** (January), S111-S132.

[27] Green, P. E. (1974), On the designs of choice experiments involving multifactor alternatives, *Journal of Consumer Research*, **1** (September), 61-8.

[28] Green, P. E. (1984), Hybrid models for conjoint analysis: an expository review, *Journal of Marketing Research*, **21** (May), 155-69.

[29] Green, P.E. and Carmone, F. J. (1970), *Multidimensional scaling and related techniques in marketing analysis*, Boston: Allyn and Bacon.

[30] Green, P.E., Carroll, J. D. and Carmone, F. J. (1978), Some new types of fractional factorial designs for marketing experiments, in *Research in Marketing*, 1, J. N. Sheth, Greenwich, ed. JAI Press, 99-122.

[31] Green, P.E., Carroll, J. D. and Goldberg, S. M. (1981), A general approach to product design optimization via conjoint analysis, Journal of Marketing, **45** (summer), 17-37.

[32] Green, P. E. and Krieger (1993), Conjoint analysis with product-positioning applications, in *Handbooks in OR & MS*, **5**, Eliashberg J. and Lilien G. L., eds. New York: Elsevier Science Publishers.

[33] Green, P. E. and Rao V. R. (1971), Conjoint Measurement for quantifying judgmental data, *Journal of Marketing Research*, **8** (August), 355-63.

[34] Green, P. E. and Srinivasan, V. (1978), Conjoint analysis in consumer research: Issues and outlook, *Journal of Consumer Research*, **5**, 103-23.

[35] Green, P. E. and Srinivasan, V. (1990), Conjoint analysis in marketing: New developments with implications for research and practice, *Journal of Marketing*, **54** (October), 3-19.

[36] Green, P.E. and Wind, Y. (1975), New way to measure consumers' judgement, *Haverd Business Review*, July-August, 107-117.

[37] Hagerty, M. R. (1985), Improving the predictive power of conjoint analysis: the use of factor analysis and cluster analysis, *Journal of Marketing Research*, **22** (May), 168-84.

[38] Hair, J. F. Jr., Anderson, P. E., Tatham, R. L. and Black, W. C. (1992), *Multivariate Data Analysis with Readings*, Macmillam Publishing Company, New York.

[39] Hlawka, E. (1962), Zur angenaherten Berechnung mehrfacher Integrale, *Monatsh, Math.*, **66**, 140-151.

[40] Hua, L. K. (1956), *Introduction to Number Theory*, Science Press, Beijing.

[41] Joreskog K. G. (1974), Analyzing psychological data by structural analysis of covariance matrices, in *Contemporary Developments in Mathematical Psychology*, Atkinson, R. C., Krantz, D. H., Luce, R. D. and Suppes, P. eds. San Francisco: Freeman, 1-56.

[42] Johnson, R.M. (1987) Adaptive Conjoint Analysis, in *Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis and Computer Interviewing*, M. Meterrano, ed. Ketchum, ID: Sawtooth Software, 253-65.

[43] Kamakure, W. (1988), A least squares procedure for benefit segmentation with conjoint experiments, *Journal of Marketing Research*, **25**, 157-67.

[44] Kiefer, J. (1959), Optimum experimental designs, *Journal of Royal Statistics(B)*, **21**, 272-319.

[45] Kiefer, J. and Wolfowitz, J. (1960), The equivalence of two extremum problems, *Canad. J. Math.*, **12**, 363-366.

[46] Kiefer, J. (1974), General equivalence theory for optimum designs (approximate theory), *The Annals of Statistics*, **2**, 849-879.

[47] Kiefer, J. (1975), Construction and optimality of generalized Youden designs, *A survey of Statistical Design and Linear Models* (ed. by J.N. Srivastava), 333-353, North-Holland Pub. Co, Amsterdam.

[48] Korobov, N. M. (1959a), The approximate computation of multiple integrals, *Dokl. Akad. Nauk. SSSR*, **124**, 1207-1210.

[49] Korobov, N. M. (1959b), Computation of multiple integrals by the method of optimal coefficients, *Vestnik Moskow Univ. Ser. Mat. Astr. Fiz. Him.*, **4**, 19-25.

[50] Kruskal, J. B. (1965), Analysis of factorial experiments by estimating monotone transformations of the data, *Journal of the Royal Statistical Society, Series B*, **27**, 251-63.

[51] Lenk, P. J., DeSarbo, W. S., Green, P. E. and Young, M. R. (1994), Hierarchical Bayes conjoint analysis: recovery of part-worth heterogeneity from incomplete designs in conjoint analysis, working paper, School of Business Administration. University of Michigan.

[52] L, H. H. and Stout, W. F. (1995), Assessment of unidimentionality for mixed polytomous and dichotomous item data: Refinements of Poly-DIMTEST. paper presented at the annual NCME meeting, San Francisco, CA.

[53] Louviere, J.J. (1988) *Analyzing Decision Making: Metric Conjoint Analysis*, Newbury Park, Calif.: SAGE Publications.

[54] Louviere, J.J. and Woodworth, G. G. (1988), On the design and analysis of correlated conjoint experiments using difference designs, in *Advances in Consumer Research*, **15**, ed. Michael J. Houston, Provo, UT: Association for Consumer Research, 510-517.

[55] Luce, D. R. and Tukey, J. W. (1964), Simultaneous conjoint measurement: A new type of fundamental measurement, *Journal of Mathematical Psychology*, **1**, 1-27.

[56] Montgomery, D. B. (1985), Conjoint calibration of the customer/competitor interface in industrial markets, *Marketing Science Institute Report*, 85-112.

[57] Moore, W. L. (1980), Levels of aggregation in conjoint analysis: an empirical comparison, *Journal of Marketing Research*, **17** (November), 516-23.

[58] Moore, W. L. and Pessemier, E. A. (1993), *Product Planning and Management: Designing and Delivering Value*, New York: McGraw-Hill Series in Marketing, 160-174.

[59] Moore, William L. and Holbrook, M. B. (1990), Conjoint analysis on objects with environmentally correlated attributes: the questionable importance of representative design, *Journal of Consumer Research*, **16** (March), 490-497.

[60] Nandakumar, R., Yu, F., Li, H. H. and Stout W. (1995), Assessing unidimensionality of polytomous data, working paper, University of Delaware, Newark.

[61] Neter, J., Kutner, M. H., Nachtsheim, C. J. and Wasserman, W. (1996), *Applied Linear statistical Models.* (4th Ed.). Chicago: Irwin.

[62] Niederreiter, H. (1977), Pseudo-random number and optimal coefficients, *Advances in Math.*, **26**, 99-181.

[63] Ni, G. X. and Ding, S. L. (1984), The Generalized Optimality of Orthogonal Array with Any Order Interaction, *KeXue TongBao*, **29**, 452-453.

[64] Robinson, P. J. (1980), *Applications of conjoint analysis to pricing problems*, in Market Measurement and Analysis: Proceedings of the 1979 ORSA/TIMS Conference on marketing, Montgomery D. B. and Wittink D. R. (eds.), Cambridge, Mass.: Marketing Science Institute, 183-205.

[65] Roskam, E. C. I. (1968), *Metric analysis of ordinal data in Psychology*, Netherlands: Vam Voorschoten.

[66] Searle, S. R. (1982), *Matrix Algebra Useful for Statistics*, John Wiley & Sons, Inc.

[67] Shah, K. R. and Sinha, B. K. (1989), Theory of Optimal Designs in *Lecture Notes in Statistics*, **54**, Berger, J., Fienberg, S., Gani, J., Krickberg, K. and Singer, B. (eds.), Springer-Verlag Berlin Heideberg.

[68] Sinha, B. K. (1980), Optimal Block Designs, unpublished seminar notes. Indian statistical Institute, Calcutta, India.

[69] Steckel, J. H., DeSarbo, W. S. and Mahajan, V. (1991), On the creation of feasible conjoint analysis experimental designs, *Decision Sciences*, **22**, 435-442.

[70] Taguchi, G. (1987), *Experimental designs*, Mechanical Industry Press, Beijing.

[71] Vriens, M., Wedel, M. and Wilms, T. J. (1994), Metric conjoint segmentation methods: A Monte Carlo comparison, working paper, Faculty of Economics, University of Groningen.

[72] Wang, J. C. and Wu, C. F. J. (1992), Nearly Orthogonal Arrays with Mixed Levels and Small Runs, *Technometrics*, **34**, No. 4, 409-419.

[73] Wang, Y. and Fang, K. T. (1981), A note on uniform distribution and experimental design, *Kexue Tongba*, **26**, 485-489.

[74] Wang, Y. and Fang, K. T. (1990a), Number theoretic methods in applied statistics, *Chinese Ann. Math. Ser. B*, **11**, 41-55.

[75] Wedel, M. and Kistemaker, C. K. (1989), Consumer benefit segmentation using clusterwise linear regression, *International Journal of Research in Marketing*, **6**, 45-9.

[76] Wedel, M. and Steenkamp, E. M. (1989), Fuzzy clusterwise regression approach to benefit segmentation, *International Journal of Research in Marketing*, **6**, 245-58..

[77] Wiley, D. E., Schmidt, W. H. and Bramble, W. J. (1973), Studies of a class of covariance structure models, *Journal of the American Statistical Association*, **68**, 317-23.

[78] Winker, P. and Fang, K. T. (1995), Application of threshold accepting to the evaluation of the discrepancy of a set of points, Technical Report Math-067, Hong Kong Baptist University.

[79] Wittink, D. R. and Cattin P. (1989), Commercial use of conjoint analysis: An update, *Journal of Marketing*, **53** (July), 91-6.

[80] Wittink, D. R. and Montgomery, D. B.(1979), Predictive validity of tradeoff analysis for alternative segmentation schemes, in *AMA Educators' Conference Proceedings*, 131-8.

[81] Wittink, D. R., Vriens, M. and Burhenne, W. (1994), Commercial use of conjoint in Europe: results and critical reflections, *International Journal of Research in Marketing*, **11**, 41-52.

[82] Young, F. W. (1972), A model of polynomial conjoint analysis algorithms, in *Multidimensional scaling: Theory and applications in the behavioural sciences*, **1**, Shepard, R. N., Romney, A. K. and Nerlove, S. eds. New York: Academic Press, 69-104.

[83] Zufryden, F. S. (1977), A conjoint measurement-based approach for optimal new product design and market segmentation, in *Analytic approaches to product and market planning*, A. D. Shocker, ed., Cambridge, MA: Marketing Science Institute, 100-114.