Optimal Design in the Presence of Random or Fixed Block Effects

Dissertation
zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von **M.Sc. Jesús Alonso Cabrera**
geb. am 25.12.1966 in Usiacurí, Kolumbien

genehmigt durch die Fakultät für Mathematik
Der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Rainer Schwabe
Prof. Dr. Jesús López-Fidalgo

eingereicht am: 31.07.2013
Verteidigung am: 03.02.2014

Acknowledgment

First of all, I would like to express my most sincere gratitude to my supervisor, Prof.Dr. Rainer Schwabe, for his guidance, inspiration and supporting comments, throughout the process to do this thesis. I am further very grateful to the Professors and all my colleagues in our Institute of Mathematical Stochastic. I wish to thank Kerstin Altenkirch for her generous help.

I am deeply grateful to my wife Naira and our children Jesús David and Iván for their encouragement and love, I am also grateful to my parents, sisters, brother and his wife, nieces and nephews for their motivation and confidence.

Finally, I am much obliged to Universidad del Norte, for the financial support of my PhD program.

TO MY FATHER.

Zusammenfassung

Die vorliegende Dissertation entwickelt D-optimale Versuchspläne für lineare Modelle mit sowohl qualitativen als auch quantitativen Einflussfaktoren. Zu Beginn wird eine allgemeine Einführung in die Theorie der optimalen Versuchsplanung mit den am häufigsten verwendeten Optimalitätskriterien gegeben und das damit verbundene Allgemeines Äquivalenztheorem vorgestellt. Danach wird die Blockbildung als ein lineares Zweifaktormodell ohne Wechselwirkungen eingeführt, bei dem es keine Wechselwirkung zwischen den Blöcken und den Effekten des quantitativen Regressors gibt. In Abhängigkeit von den jeweiligen Versuchsbedingungen können zwei Arten von Blockeffekten, feste oder zufällige, unterschieden werden, die zu unterschiedlichen Modelle führen. In diesem Zusammenhang werden zwei Theoreme bewiesen, in denen die Optimalität und die wichtige Eigenschaft der Orthogonalität analysiert wird.
Im Hinblick auf die Zielsetzung dieser Arbeit wird eine Charakterisierung D-optimaler Pläne für die komplexe Structur ein gemischten zweifaktoriellen Modells mit gemeinsamem Basiswert gewonnen. Diese Charakterisierung erlaubt unter wenigen Annahmen die analytische Bestimmung der Gewichte des optimalen Versuchsplans vermittels konvexer Optimierung. Dabei ist zu beachten, dass diese optimalen Gewichte zwar vom Verhältnis der Varianzkomponenten zueinander abhängen, aber dass in praktischen Anwendungen die im Grenzübergang optimalen Pläne eine hohe Effizienz aufweisen, wenn der wert des Varianzverhältnisses gegen Null oder gegen unendlich strebt.

Summary

In the present thesis optimal experimental designs are developed for linear regression models with both qualitative and quantitative factors of influence. The exposition starts with a general introduction to optimal design theory in which the most popular optimality criteria and the corresponding General Equivalence Theorem are presented. After this brief introduction, we consider the blocked response surface experiments which can be regarded as two-factor linear models without interactions. Here as common in the literature we may assume that there is no interaction between the blocks and the effects of the quantitative regression factor. Depending on the nature of the experiment two types of blocking variables, fixed or random, have to be incorporated which then lead to essentially different models. Subsequently in this context two theorems are established in which the optimality and the key property of orthogonality are analyzed.
With respect to the aim of this work we generate a characterization of optimal designs for a more complex structure, a two-factorial mixed model with a common intercept. This characterization allows under few assumptions to find the weights of the optimal design analytically by means of convex optimization. It is worthwhile noting that the optimal weights depend on the ratio of variance components. However in this context, we show that in practical applications limiting optimal design shows a high efficiency, when the variance ratio approaches to zero or infinity.

Contents

Chapter 1

Introduction

In the framework of the statistical experiments the theory of optimal designs has been developed. In general the subject of this theory is that for an appropriate model, if we want to put emphasis on a special quality of the parameters estimate, the experimental settings should be chosen according to certain criteria with statistical sense, that by using a minimum amount of resources a maximum of information can be obtained. In the literature on optimal designs, a preeminent author was Kiefer(1959) who presented the main concepts, such as design measure and a variety of optimality criteria for this branch of experimental designs, Kiefer, in particular gave the name D-optimality to the criterion introduced by Wald(1943), which is the most commonly applied design criterion and is defined on the determinant of the covariance matrix. Kiefer and Wolfowitz(1960) made an essential contribution known as the first Equivalence Theorem, there they proved the equivalence between D-optimality and G-optimality and provided tools to verify the optimality of a given design. The monograph by Silvey (1980) and the book by Fedorov (1972) and more recently the books by Atkinson & Donev (1992) and Pukelshein (1993) where the authors made statistical and formal presentations of the optimal designs, also are recognized.
Very often, for a more realistic analysis of the data, a regression experiment has to be designed involving both qualitative and quantitative factors of

influence, for example, an intra-class regression model with the identical partial model in each class, see e.g. Searle(1971,p. 355) and Kurotschka (1984); a blocked response surface design , see e.g. Gilmour and Trinca (2000); Goos(2002) and Waite et al (2012) or more general, a two-factor linear model.
The problem of constructing an optimum experimental design for estimating the vector parameter of a two-factor linear model is more complex than for single factor models, however the question under what conditions we can find optimum designs for two-factor models in terms of optimum designs for their single factor models has been developed; for example, for multi-factor model with homoscedastic errors Schwabe(1996) presents optimal designs for a great variety of cases. Regarding these ideas, the aim of the present work is generating D-optimum designs for multi-factor models in the presence of random block effects. Of particular interest is the limiting behavior, when the variance of the random effects gets large and zero. Moreover we apply essentially analytical methods that involve convex optimization in continuous set up to models which involve discrete structure, a fact that has been neglected in the literature with few exceptions.
This thesis is organized as follows: in chapter 2, we present a general introduction to the optimal design of experiments, in particular the classical optimality criteria and the corresponding General Equivalence Theorem. In chapters 3 and 4 we deal with D-optimal designs for regression models in the presence of fixed and random block effects, respectively. We emphasize that the fixed and random block effects models are essentially different models, but in the two cases the block effects are considered nuisances parameters and we give special attention to optimality of orthogonal blocking. In chapter 5, we consider a two-factorial mixed model with a common intercept in which the factor effects are related by different interaction structure. Chapter 6 brings about optimal product design, suitable for analysis of the direct fixed effects as well as of the variances of the random effects. This thesis is closed with a discussion of the results and an outlook to possible future work.

Chapter 2

Optimal Designs in Linear Regression Models

The main aim of this chapter is to present a general introduction of several important topics on optimality theory of experimental designs in linear regression models, in particular the classical optimality criteria and an associated General Equivalence Theorem where the Kiefer-Wolfowitz equivalence theorem appear as a special case. This theory is based essentially on analytical methods that involve convex optimization in continuous setup. The foundation for this introductory chapter on optimal designs are the monograph by Silvey (1980) and Fedorov (1972) also by the books of Pukelsheim (1993) and Atkinson, Donev and Tobias (2007).

2.1 Classical Linear Regression Models

We consider an experiment where the response Y is a random variable with distribution of probability $\mathcal{P}$, such that Y is decomposed into the deterministic and known mean response function $\mu(\mathbf{x}, \boldsymbol{\beta})$ plus a random error ε, thus Y depend on r explanatory variables typically represented by the vector $\mathbf{x}^\top = (x_1, \ldots, x_r)$ (which it can be usually controlled) and $\boldsymbol{\beta}$ a vector of parameters which are constant but unknown to the experimenter; in this experimental situation $Y(\mathbf{x}) = \mu(\mathbf{x}, \boldsymbol{\beta}) + \varepsilon$ is a *linear* (*homoscedastic*) *model* whenever the expected value and variance of Y taken the following form

$$\mathrm{E}_{\mathcal{P}}(Y(\mathbf{x})) = \mu(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta}, \quad \mathrm{Var}_{\mathcal{P}}(Y(\mathbf{x})) = \mathrm{Var}(\varepsilon) = \sigma^2 \tag{2.1}$$

respectively, where $\mathbf{f} = (f_1, \ldots, f_p)^\top$ is a vector of p known, linearly independent and real valued regression function defined on the experimental region $\mathcal{X}$ which we assume is a compact set in $\mathbb{R}^r$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is a p-vector of unknown regression parameters for the effects of the explanatory variables and $\sigma^2 (> 0)$ is an unknown scalar parameter.

2.2 Experimental Designs and Information Matrices

In order to make statistical inference on the unknown parameters $\beta_1, \ldots, \beta_p$ or certain function of them, the experimenter is allowed to select N independent observations at the setting of control vectors $\mathbf{x}_1, \ldots, \mathbf{x}_N$ chosen from the experimental region $\mathcal{X}$.

Definition 2.1 *An experimental design of size N is a list of experimental settings $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in $\mathcal{X}$, not necessarily all distinct, denoted $\mathbf{d}_N := (\mathbf{x}_1, \ldots, \mathbf{x}_N)$.*

Typically the experimental design $\mathbf{d}_N$ is selected according to certain structure from the experimental region to answer the statistical question of interest. Let $\mathbf{Y}_N = (Y_1(\mathbf{x}_1), \ldots, Y_N(\mathbf{x}_N))^\top$ be the vector of observations at the experimental design $\mathbf{d}_N$, where it is assumed that the observations are independent and that the experimental runs are carried out under homogeneous conditions, then in matrix notation the uncorrelated linear model can be described by

$$\mathbf{Y}_N = \mathcal{F}(\mathbf{d}_N)\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.2}$$

where $\mathcal{F}(\mathbf{d}_N) = (\mathbf{f}(\mathbf{x}_1), \ldots, \mathbf{f}(\mathbf{x}_N))^\top$ is the $N \times p$ design matrix for the parameter $\boldsymbol{\beta}$, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)^\top$ is the N-vector of random error. Consequently, the expectation vector and covariance matrix of $\mathbf{Y}_N$ become

$$\mathrm{E}_{\mathcal{P}}(\mathbf{Y}_N) = \mathcal{F}(\mathbf{d}_N)\boldsymbol{\beta}, \quad \mathrm{Cov}_{\mathcal{P}}(\mathbf{Y}_N) = \mathrm{Cov}_{\mathcal{P}}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}_N, \tag{2.3}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix.
With these assumptions, we consider the general case, when the experimenter is interested in a linear aspect $\boldsymbol{\psi}$ of $\boldsymbol{\beta}$ identifiable under $\mathbf{d}_N$, that is a s-vector of linear combinations of the parameter $\boldsymbol{\beta}$, defined by

$$\boldsymbol{\psi}(\boldsymbol{\beta}) := \mathbf{L}_{\boldsymbol{\psi}}\boldsymbol{\beta}, \tag{2.4}$$

such that

$$\mathbf{L}_{\boldsymbol{\psi}} = \mathbf{K}(\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N)) \quad \text{for some known matrix } \mathbf{K} \in \mathbb{R}^{s\times p}.$$

On the other hand, in the space of symmetric $p \times p$ matrices, $\mathrm{Sym}(p)$, the subsets of nonnegative definite matrices, $\mathbf{NND}(p)$, and of positive definite matrices, $\mathbf{PD}(p)$, are key to the sequel. They are defined through quadratic forms

$$\mathbf{A} \in \mathbf{NND}(p) \iff \mathbf{A} \in \mathrm{Sym}(p) \text{ and } \mathbf{x}^\top\mathbf{A}\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^p$$

$$\mathbf{A} \in \mathbf{PD}(p) \iff \mathbf{A} \in \mathrm{Sym}(p) \text{ and } \mathbf{x}^\top\mathbf{A}\mathbf{x} > 0 \text{ for all } \mathbf{0} \neq \mathbf{x} \in \mathbb{R}^p.$$

Note that $\mathbf{PD}(p)$ is a subset of $\mathbf{NND}(p)$. In this work, we use the following characterization: For any $n \times p$ matrix $\mathbf{X}$, the matrix $\mathbf{X}^\top\mathbf{X} \in \mathbf{NND}(p)$.

Hence, $\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N)$ is an element of the set of nonnegative definite matrices $\mathbf{NND}(p)$, and by the Gauß-Markov theorem the best linear unbiased estimator for an identifiable linear aspect $\boldsymbol{\psi}$ is given by

$$\widehat{\boldsymbol{\psi}} = \mathbf{L}_{\boldsymbol{\psi}}(\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N))^-\mathcal{F}(\mathbf{d}_N)^\top\mathbf{Y}_N \tag{2.5}$$

with

$$\mathrm{Cov}(\widehat{\boldsymbol{\psi}}) = \sigma^2\mathbf{L}_{\boldsymbol{\psi}}(\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N))^-\mathbf{L}_{\boldsymbol{\psi}}^\top, \tag{2.6}$$

where for a matrix $\mathbf{A}$ we denote by $\mathbf{A}^-$ an arbitrary generalized inverse of $\mathbf{A}$, which satisfies $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$.
Thus $\widehat{\boldsymbol{\psi}}$ is invariant to the choice of the generalized inverse of $\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N)$. If $\mathrm{rank}(\mathcal{F}(\mathbf{d}_N)) = p \leq N$, then $\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N)$ is an element of the interior of $\mathbf{NND}(p)$ the set of positive definite matrices, $\mathbf{PD}(p)$, (see Pukelsheim, 2006,p.10) hence $\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N)$ has its standard inverse $(\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N))^{-1}$ for its unique generalized inverse and in particular, we can obtain the best linear unbiased estimator of the whole parameter vector $\boldsymbol{\beta}$ given by

$$\widehat{\boldsymbol{\beta}} = (\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N))^{-1}\mathcal{F}(\mathbf{d}_N)^\top\mathbf{Y}_N \tag{2.7}$$

with

$$\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) = \sigma^2(\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N))^{-1}, \tag{2.8}$$

hence the matrix $\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N)$, affects both the estimator $\widehat{\boldsymbol{\psi}}$ and its covariance matrix and the quality of a selected experimental design $\mathbf{d}_N$ is associated with the information provided for the linear combinations of the parameters because this information is contained in the $p \times p$-matrix $\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N)$.

Definition 2.2 *For a experimental design $\mathbf{d}_N$, the $p \times p$-matrix*

$$\frac{1}{N}\mathcal{F}(\mathbf{d}_N)^\top\mathcal{F}(\mathbf{d}_N) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{f}(\mathbf{x}_i)\mathbf{f}(\mathbf{x}_i)^\top := M_{\mathbf{f}}(\mathbf{d}_N)$$

is called the moment matrix of the experimental design $\mathbf{d}_N$.

As basic principles of regression experimental design is the execution of replicas as well as the randomization of the standard order of the experimental setting when the design is used, a experimental design $\mathbf{d}_N$ may contain a number, L with $L \leq N$, of distinct experimental control vectors $\mathbf{x}_1, \ldots, \mathbf{x}_L$ with frequencies $n_1, \ldots, n_L$ respectively such that $\sum_{l=1}^{L} n_l = N$ and we can consider the set of distinct design points $\{\mathbf{x}_1, \ldots, \mathbf{x}_L\}$ as the set of all support points of a discrete probability distribution δ_N on $\mathcal{X}$, which it defines

$$\delta_N(\mathbf{x}) = \begin{cases} \dfrac{n_l}{N} & \text{if} \quad \mathbf{x} = \mathbf{x}_l,\ l = 1, \ldots, L \\ \\ 0 & \qquad \text{otherwise,} \end{cases}$$

then every experimental design $\mathbf{d}_N$ has an associated discrete probability distribution δ_N, thus the moment matrix $M_{\mathbf{f}}(\mathbf{d}_N)$ now depends on δ_N and it can be written as an expected value

$$\begin{aligned} M_{\mathbf{f}}(\delta_N) &= \mathrm{E}_{\delta_N}\left(\mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top\right) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\,\mathbf{f}(\mathbf{x})^\top \delta_N(d\mathbf{x}) \\ &= \sum_{l=1}^{L} \delta_N(\mathbf{x}_l)\mathbf{f}(\mathbf{x}_l)\mathbf{f}(\mathbf{x}_l)^\top = M_{\mathbf{f}}(\mathbf{d}_N). \end{aligned} \tag{2.9}$$

2.3 Continuous and Exact Designs

The transition of an experimental design d_N for finite sample size N to a discrete probability distribution δ_N allows the generalization to designs for infinite sample size with the introduction of design measures, see for example, Kiefer(1959).

Definition 2.3 *δ is a design measure on the compact subset $\mathcal{X}$ of $\mathbb{R}^r$ if δ is a probability distribution on the Borel sets of $\mathcal{X}$.*
The set of all design measure on $\mathcal{X}$ will be denoted by $\mathcal{W}(\mathcal{B}_{\mathcal{X}})$.

$M_{\mathbf{f}}(\delta) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\, \mathbf{f}(\mathbf{x})^\top \delta(d\mathbf{x})$; *is the moment matrix of the design measure* δ.

Now the generalization of the experimental design to design measure, implies that the concept of identifiability has to be generalized too.

Definition 2.4 *A linear aspect* $\boldsymbol{\psi}$ *defined by* $\boldsymbol{\psi}(\boldsymbol{\beta}) := \mathbf{L}_{\boldsymbol{\psi}}\boldsymbol{\beta}$ *is identifiable under* δ *if there exists a matrix* $\mathbf{K} \in \mathbb{R}^{s\times p}$ *such that* $\mathbf{L}_{\boldsymbol{\psi}} = \mathbf{K}M_{\mathbf{f}}(\delta)$.

Definition 2.5 *If* ψ *is identifiable under* δ, *then the* $s \times s$*-matrix* $\mathbf{C}_\psi(\delta) := \mathbf{L}_{\boldsymbol{\psi}}(M_{\mathbf{f}}(\delta))^{-}\mathbf{L}_{\boldsymbol{\psi}}^\top$ *is called the covariance matrix of* ψ *corresponding to the design* δ.

We note that in the definition, the covariance matrix is independent of the special choice of the generalized inverse of the moment matrix $M_{\mathbf{f}}(\delta)$, also for a experimental design $\mathbf{d}_N$, the variance-covariance matrix of the best unbiased estimator of ψ, $\mathrm{Cov}(\psi)$ is directly proportional to $\mathbf{C}_\psi(\delta_N)$, with proportionality constant $\frac{1}{N}\,\sigma^2$ where δ_N is the corresponding associated design measure.

Definition 2.6 *If* ψ *is identifiable under* δ *with* $s \times p$ *coefficient matrix* $\mathbf{L}_{\boldsymbol{\psi}}$ *of full row rank* s, *then the* $s \times s$*-matrix*

$$\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\delta)) := \mathbf{C}_\psi(\delta)^{-1} = \left(\mathbf{L}_{\boldsymbol{\psi}}(M_{\mathbf{f}}(\delta))^{-}\mathbf{L}_{\boldsymbol{\psi}}^\top\right)^{-1}$$

is called the information matrix of ψ *corresponding to the design* δ.

In particular, if $M_{\mathbf{f}}(\delta)$ is positive definite, then the whole parameter vector $\boldsymbol{\beta}$ is identifiable under δ, and $\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{f}}(\delta)) = (\mathbf{I}_p(M_{\mathbf{f}}(\delta))^{-1}\mathbf{I}_p)^{-1} = M_{\mathbf{f}}(\delta)$, i.e. the information matrix of $\boldsymbol{\beta}$ coincide with the non singular moment matrix of δ.

Let $\mathfrak{M}$ be, the set of moment matrices, that is

$$\mathfrak{M} = \left\{ M_{\mathbf{f}}(\delta) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\, \mathbf{f}(\mathbf{x})^\top \delta(d\mathbf{x}); \quad \delta \in \mathcal{W}(\mathcal{B}_{\mathcal{X}}) \right\} \subset \mathbf{NND}(p)$$

We note that for δ_1 and δ_2 in $\mathcal{W}(\mathcal{B}_{\mathcal{X}})$ and all $t \in [0,1]$, we have that

$$t\delta_1 + (1-t)\delta_2 \in \mathcal{W}(\mathcal{B}_{\mathcal{X}}).$$

Hence the set of design measures is a convex set.
As a result, the set $\mathfrak{M}$ also is convex, indeed due to the linearity of the integral, it holds

$$tM_{\mathbf{f}}(\delta_1) + (1-t)M_{\mathbf{f}}(\delta_2) = M_{\mathbf{f}}(t\delta_1 + (1-t)\delta_2) \in \mathfrak{M}.$$

Moreover, if we assume that $\mathbf{f}$ is continuous on the compact set $\mathcal{X}$, then $\mathfrak{M}$ is a compact set (see Pukelsheim (1993),p.29), and according to Carathéodory's theorem (Silvey(1980),p.72) each element $M_{\mathbf{f}}(\delta)$ of $\mathfrak{M}$ can always be expressed as $M_{\mathbf{f}}(\xi)$, where ξ is a discrete design measure supported on at most $\frac{1}{2}p(p+1)+1$ points; that is, there always exists an *approximate* design ξ with a finite support which satisfies

$$M_{\mathbf{f}}(\delta) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\,\mathbf{f}(\mathbf{x})^{\top}\delta(d\mathbf{x}) = \sum_{\mathbf{x}\in supp\ \xi} \xi(\mathbf{x})\mathbf{f}(\mathbf{x})\,\mathbf{f}(\mathbf{x})^{\top} = M_{\mathbf{f}}(\xi).$$

This result suggests that it will be sufficient to consider only discrete designs, i.e. discrete probability distributions on $\mathcal{X}$ with a finite support. If the design ξ is supported at n distinct design points $\mathbf{x}_i \in \mathcal{X}$, it is denoted by

$$\xi = \left\{ \begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \\ w_1 & w_2 & \dots & w_n \end{array} \right\}, \tag{2.10}$$

where the first line gives the levels of the design points in the support set and the second line gives the weights associated with each design point. Since ξ is a discrete design measure,

$$\int_{\mathcal{X}} \xi(\mathrm{d}\mathbf{x}) = \sum_{i=1}^{n} \xi(\mathbf{x}_i) = \sum_{i=1}^{n} w_i = 1, \quad \text{and } \ 0 \le w_i \le 1, \text{ for all } \ i.$$

If the design weights $\xi(\mathbf{x}_i) = w_i$ are rational for all $i = 1, \dots, n$, then it will be possible to find an associated experimental design of size $N \ge p$, where each design point $\mathbf{x}_i$ is replicated an integer number $r_i = w_i N$ respectively, this designs are called exact designs and also are represented as ξ_N.

On the other hand, if the weights of a design ξ are not all rational, it will not be possible to find an exact design for any finite N, such design ξ is called continuous.

The class of all *approximate* design, that is all discrete design measure(exact and continuous) on $\mathcal{X}$ will be denoted by $\mathfrak{A}$. The subclass of $\mathfrak{A}$ of all design ξ for which the linear aspect $\psi(\boldsymbol{\beta}) = \mathbf{L}_\psi \boldsymbol{\beta}$ is identifiable is denoted by $\mathfrak{A}_\psi$, thus we can write

$$\mathfrak{M}_\psi = \{M_{\mathbf{f}}(\xi);\ \ \xi \in \mathfrak{A}_\psi\} \subset \mathfrak{M}.$$

The following step is to find designs ξ^* such that the performance of the estimator $\hat{\psi}$ is optimum. As a remark, the performance of the designs is typically valued in terms of the moment matrix of ξ^* through the information matrix of ψ, however the moment matrices are not necessarily comparable, therefore a uniform optimization is not possible in general, but for instance, in the linear context under discussion, we have that if the responses are normally distributed, the related $100(1-\alpha)\%$ Scheffé confidence ellipsoid for the identifiable linear aspect $\psi(\boldsymbol{\beta}) = \mathbf{L}_\psi \boldsymbol{\beta}$ with $\text{rank}(\mathbf{L}_\psi) = s$ will be

$$\widehat{S}(\xi_N) = \left\{ \psi \in \mathbb{R}^s : \left(\psi - \widehat{\psi}\right)^\top \boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi_N)) \left(\psi - \widehat{\psi}\right) \leq s\, f_{s,v,1-\alpha}\, \widehat{\sigma^2} \right\}$$

for all $\xi_N \in \mathfrak{A}_\psi$, where $\widehat{\sigma^2}$ is the usual estimator of σ^2, i.e. the residual sum of squares divided by $v = N - \text{rank} M_{\mathbf{f}}(\xi_N)$ and $f_{s,v,1-\alpha}$ is the $(1-\alpha)$ quantile of the F distribution with numerator degrees of freedom s and denominator degrees of freedom v .

The quality of this confidence ellipsoid depends on the probability distribution ξ_N through the information matrix $\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi_N))$, because it is as precise as its volume small is, and the volume of the ellipsoid is inversely proportional to the square root of the determinant of the information matrix.

Therefore, if we want to put emphasis on the quality of the parameter estimates a natural way or *criteria* is to find a design ξ which maximizes the determinant of the information matrix.

2.4 Classical Optimality criteria

Let ϕ be a real-valued function on the whole set of $\mathbf{NND}(s)$. Given $\psi(\boldsymbol{\beta}) = \mathbf{L}_\psi\boldsymbol{\beta}$ a linear aspect with the coefficient matrix $\mathbf{L}_\psi$ of full row rank s and a design $\xi \in \mathfrak{A}_\psi$,. When the function ϕ is evaluated, with statistical meaning, in the information matrix $\boldsymbol{\mathcal{I}}_\psi(M_\mathbf{f}(\xi)) = \left(\mathbf{L}_\psi M_\mathbf{f}(\xi)^- \mathbf{L}_\psi^\top\right)^{-1}$, then we have a design criteria and a design $\xi^* \in \mathfrak{A}_\psi$ is called ϕ-optimal for ψ if

$$\phi(\boldsymbol{\mathcal{I}}_\psi(M_\mathbf{f}(\xi^*))) = \max_{\xi \in \mathfrak{A}_\psi} \phi(\boldsymbol{\mathcal{I}}_\psi(M_\mathbf{f}(\xi))). \tag{2.11}$$

That is, the optimality properties of designs ξ are determined by their moment matrix $M_\mathbf{f}(\xi)$.

Many different criteria can be found in the optimal design literature and each of these criteria capture particular statistical aspects; in the following we will present only some of the most important and popular design criteria.

2.5 D_ψ-optimality

We start with the determinant criterion for a linear aspect ψ, denoted D_ψ-criterion, where $\psi(\boldsymbol{\beta}) = \mathbf{L}_\psi\boldsymbol{\beta}$ is an identifiable linear aspect with $\text{rank}\mathbf{L}_\psi = s \leq p$, this criterion determines the design that maximizes the determinant of the information matrix $\boldsymbol{\mathcal{I}}_\psi(M_\mathbf{f}(\xi))$. This corresponds geometrically to minimize the volume of the confidence ellipsoid for the linear aspect ψ of the unknown parameter vector $\boldsymbol{\beta}$ in the linear model (2.1) under the assumption of normality of errors.

Definition 2.7 *A design ξ^* is called D_ψ-optimal or ϕ_0-optimal for ψ if it maximizes*

$$\phi_0\left(\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))\right) = \det\left(\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))\right) \quad \textit{for all} \quad \xi \in \mathfrak{A}_\psi,$$

or equivalently

$$\det\left(\mathbf{L}_\psi M_{\mathbf{f}}(\xi^*)^- \mathbf{L}_\psi^\top\right) = \min_{\xi \in \mathfrak{A}_\psi} \det\left(\mathbf{L}_\psi M_{\mathbf{f}}(\xi)^- \mathbf{L}_\psi^\top\right).$$

One of the most distinctive properties of the D_ψ-criterion is that the optimal design remains invariant to regular linear transformations of the linear aspect. Indeed, suppose that an aspect is re-parametrized according to $\widetilde{\psi}(\boldsymbol{\beta}) = \mathbf{H}\mathbf{L}_\psi\boldsymbol{\beta}$, with $\mathbf{H}$ a nonsingular $s \times s$ matrix, then provided the identity

$$\begin{aligned}
\boldsymbol{\mathcal{I}}_{\mathbf{H}\psi}(M_{\mathbf{f}}(\xi)) &= \left(\mathbf{H}\,\mathbf{L}_\psi M_{\mathbf{f}}(\xi)^- \mathbf{L}_\psi^\top\, \mathbf{H}^\top\right)^{-1} \\
&= \left(\mathbf{H}\left(\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))\right)^{-1}\mathbf{H}^\top\right)^{-1} \\
&= \left(\mathbf{H}^\top\right)^{-1}\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))\,\mathbf{H}^{-1},
\end{aligned}$$

the maximization of

$$\begin{aligned}
\det \boldsymbol{\mathcal{I}}_{\mathbf{H}\psi}(M_{\mathbf{f}}(\xi)) &= \det\left(\left(\mathbf{H}^\top\right)^{-1}\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))\,\mathbf{H}^{-1}\right) \\
&= \frac{\det \boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))}{\det \mathbf{H}^2}
\end{aligned} \tag{2.12}$$

implies that a design D_ψ-optimal for the linear aspect $\psi(\boldsymbol{\beta}) = \mathbf{L}_\psi\boldsymbol{\beta}$, is also $D_{\widetilde{\psi}}$-optimal with $\widetilde{\psi} = \mathbf{H}\psi$, because $(\det \mathbf{H})^2$ is independent of ξ.

As an illustration of D_ψ-optimality, we consider in the following section a specially important case.

2.6 D_s-optimality

In many situations we are interested in estimating merely a subset of s of the p parameters of the whole vector $\boldsymbol{\beta}$. Hence without loss of generality, we can assume that the components of $\mathbf{f}(\mathbf{x})$ are arranged in such a way that

$$\mathrm{E}(Y) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} = \mathbf{f_1}(\mathbf{x})^\top \boldsymbol{\beta_1} + \mathbf{f_2}(\mathbf{x})^\top \boldsymbol{\beta_2} \tag{2.13}$$

where the components of $\boldsymbol{\beta}_1$ are the s parameters of interest. The $p-s$ elements of $\boldsymbol{\beta}_2$ are usually treated as nuisance parameter.
If ξ is a design with moment matrix $M_{\mathbf{f}}(\xi)$, then the volume of the confidence ellipsoid for the parameter vector $\boldsymbol{\beta}_1$ is inversely proportional to the square root of the determinant of the information matrix of $\psi(\boldsymbol{\beta}) = \boldsymbol{\beta}_1 \in \mathbb{R}^s$ given by

$$\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}_1}(M_{\mathbf{f}}(\xi)) \quad = \quad \left(\mathbf{L}_s\, M_{\mathbf{f}}(\xi)^{-}\, \mathbf{L}_s^\top\right)^{-1}, \tag{2.14}$$

taking $\psi(\boldsymbol{\beta}) = \boldsymbol{\beta}_1 = \mathbf{L}_s\boldsymbol{\beta}$, where $\mathbf{L}_s = [\mathbf{I}_s \;\; \mathbf{0}] \in \mathbb{R}^{s\times p}$ and $\mathbf{I}_s$ is the $s \times s$ identity matrix. Therefore the natural criterion is

$$\phi_0(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}_1}(M_{\mathbf{f}}(\xi))) = \det\left(\mathbf{L}_s\{M_{\mathbf{f}}(\xi)\}^{-}\mathbf{L}_s^\top\right)^{-1} \tag{2.15}$$

which is called D_s-criterion. To obtain an alternative formula for the definition of ϕ_0, in this case the information matrix can be partitioned according to $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ as

$$M_{\mathbf{f}}(\xi) = \begin{pmatrix} M_{11}(\xi) & M_{12}(\xi) \\ \\ M_{12}(\xi)^\top & M_{22}(\xi) \end{pmatrix}$$

where

$$M_{ij}(\xi) = \int_{\mathcal{X}} \mathbf{f}_i(\mathbf{x})\, \mathbf{f}_j(\mathbf{x})^\top \xi(\mathrm{d}\mathbf{x}), \quad i,j \in \{1,2\}.$$

The inverse of the information matrix $\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}_1}(M_{\mathbf{f}}(\xi))$ is the covariance matrix for the least squares estimate of $\boldsymbol{\beta}_1$ that is $(\mathbf{L}_s\{M_{\mathbf{f}}(\xi)\}^{-}\mathbf{L}_s^\top)$, the $s \times s$ upper left submatrix of $M_{\mathbf{f}}(\xi)^{-}$, and by rules for inverting partitioned nonnegative definite symmetric matrices (see e.g.,Fedorov 1972, p.19), we have that

$$\left(\mathbf{L}_s\{M_{\mathbf{f}}(\xi)\}^{-}\mathbf{L}_s^{\top}\right) = \{M_{11}(\xi) - M_{12}(\xi)M_{22}(\xi)^{-}M_{12}(\xi)^{\top}\}^{-1}. \qquad (2.16)$$

Hence, a design ξ^* is called D_s-optimal or ϕ_0-optimal for $\boldsymbol{\beta}_1$ if $M_{\mathbf{f}}(\xi^*)$ maximizes

$$\begin{aligned}\phi_0(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}_1}(M_{\mathbf{f}}(\xi))) &= \det\left(\mathbf{L}_s\{M_{\mathbf{f}}(\xi)\}^{-}\mathbf{L}_s^{\top}\right)^{-1} \\ &= \det\left(M_{11}(\xi) - M_{12}(\xi)M_{22}(\xi)^{-}M_{12}(\xi)^{\top}\right)\end{aligned}$$

On the other hand if ξ is a design for which the moment matrix $M_{\mathbf{f}}(\xi)$ is nonsingular, then by the formula for the determinant of partitioned symmetric matrices we obtain

$$\det(M_{\mathbf{f}}(\xi)) = \det(M_{22}(\xi))\det\left(M_{11}(\xi) - M_{12}(\xi)M_{22}(\xi)^{-1}M_{12}(\xi)^{\top}\right). \qquad (2.17)$$

Hence, in this case, a design ξ^* is D_s-optimal or ϕ_0-optimal for $\boldsymbol{\beta}_1$ if $M_{\mathbf{f}}(\xi^*)$ maximizes

$$\begin{aligned}\phi_0(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}_1}(M_{\mathbf{f}}(\xi))) &= \det\left(\mathbf{L}_s M_{\mathbf{f}}(\xi)^{-1}\mathbf{L}_s^{\top}\right)^{-1} \\ &= \det\left(M_{11}(\xi) - M_{12}(\xi)M_{22}(\xi)^{-1}M_{12}(\xi)^{\top}\right) \\ &= \frac{\det(M_{\mathbf{f}}(\xi))}{\det(M_{22}(\xi))}\end{aligned}$$

on $\mathfrak{M}_{\boldsymbol{\beta}_1}$ or equivalently

$$\det((\mathbf{L}_s\{M_{\mathbf{f}}(\xi^*)\}^{-1}\mathbf{L}_s^{\top})^{-1}) = \max_{\xi\in\mathfrak{A}_{\boldsymbol{\beta}_1}}\{\det(M_{\mathbf{f}}(\xi))/\det(M_{22}(\xi))\}. \qquad (2.18)$$

Now we assume that in the linear model (2.1) there is an explicit constant term or intercept β_0, then we can write (2.13) as follows

$$\mathrm{E}(Y) = \mathbf{f}(\mathbf{x})^{\top}\boldsymbol{\beta} = \mathbf{f_1}(\mathbf{x})^{\top}\boldsymbol{\beta_1} + \beta_{\mathbf{0}} \qquad (2.19)$$

As a result, the D_s-optimal design for estimating $\boldsymbol{\beta}_1$ coincides with the ϕ_0-optimal design for estimating $\boldsymbol{\beta}$, indeed as $\mathbf{f}_2(\mathbf{x}) \equiv \mathbf{1}$, then $\mathbf{M}_{22}(\xi) = 1$ for all $\xi \in \mathfrak{M}$ and therefore

$$\begin{aligned} \det\left(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{f}}(\xi^*))\right) &= \det M_{\mathbf{f}}(\xi^*) \\ &= \max_{\xi \in \mathfrak{A}} \det M_{\mathbf{f}}(\xi) \\ &= \max_{\xi \in \mathfrak{A}} \{\det(M_{\mathbf{f}}(\xi))/\det(M_{22}(\xi))\} \\ &= \det\left(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}_1}(M_{\mathbf{f}}(\xi^*))\right) \end{aligned}$$

where we use the notation $\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}_1}(M_{\mathbf{f}}(\xi)) = (\mathbf{L}_{(p-1)} M_{\mathbf{f}}(\xi)^{-1} \mathbf{L}_{(p-1)}^{\top})^{-1}$ with $\mathbf{L}_{(p-1)} = [\mathbf{I}_{(p-1)}\ \mathbf{0}]$ for the information matrix of $\boldsymbol{\beta}_1$.

2.7 A_ψ-optimality

Given an estimable linear aspect $\psi(\boldsymbol{\beta}) = \mathbf{L}_\psi \boldsymbol{\beta}$ with coefficient matrix $\mathbf{L}_\psi$ of full row rank s. When the components of ψ have a definite physical meaning, then a reasonable option is to definite an optimality criteria directly attached to the standardized variances of them.

Definition 2.8 *A design ξ^* is called A_ψ-optimal or ϕ_1-optimal for ψ if it maximizes*

$$\phi_1\left(\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))\right) = \left(\frac{1}{p}\,\mathrm{tr}\left(\boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi))\right)^{-1}\right)^{-1} \quad \textit{for all} \quad \xi \in \mathfrak{A}_\psi,$$

or equivalently

$$\mathrm{tr}\left(\mathbf{L}_\psi \{M_{\mathbf{f}}(\xi^*)\}^{-} \mathbf{L}_\psi^{\top}\right) = \min_{\xi \in \mathfrak{A}_\psi} \mathrm{tr}\left(\mathbf{L}_\psi M_{\mathbf{f}}(\xi)^{-} \mathbf{L}_\psi^{\top}\right).$$

If we prefer to think in terms of variance-covariance matrix rather than information matrix, then the A_ψ-optimality criterion minimizes the trace of the covariance matrix $\mathbf{L}_\psi M_{\mathbf{f}}(\xi)^{-} \mathbf{L}_\psi^{\top}$, which implies to choose designs ξ that minimize the expected mean squared deviation of the estimates components of the linear aspect $\psi(\boldsymbol{\beta})$.

2.8 G-optimality

A natural interest for the experimenter, that arises from practical situations, lies in predicting point-wise the mean value for the response over the experimental region. Thus for a design $\xi \in \mathfrak{A}_{\boldsymbol{\beta}}$ and a particular explanatory vector $\mathbf{x} \in \mathcal{X}$, the variance of the point-wise prediction of the response function, associated with ξ is proportional to the standardized variance function

$$\begin{aligned} d(\mathbf{x}, \xi) : &= \mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-1} \mathbf{f}(\mathbf{x}) \\ &= \mathbf{f}(\mathbf{x})^\top (\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{f}}(\xi)))^{-1} \mathbf{f}(\mathbf{x}) \end{aligned}$$

As $d(\mathbf{x}, \xi)$ has a maximum over the experimental region $\mathcal{X}$, because it is a compact set and on the other hand, holds $\max\{d\} = \min\{-d\}$, then the next criterion choose a design to maximize this minimum.

Definition 2.9 *A design ξ^* with moment matrix $M_{\mathbf{f}}(\xi^*)$, positive definite, is called* **G-optimal** *if and only if maximizes*

$$\phi_G\left(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{f}}(\xi))\right) = \min_{\mathbf{x} \in \mathcal{X}} \{-d(\mathbf{x}, \xi)\} = \min_{\mathbf{x} \in \mathcal{X}} \{-\mathbf{f}(\mathbf{x})^\top \left(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{f}}(\xi))\right)^{-1} \mathbf{f}(\mathbf{x})\},$$

or equivalently

$$\max_{\mathbf{x} \in \mathcal{X}} (\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi^*)^{-1} \mathbf{f}(\mathbf{x})) = \min_{\xi \in \mathfrak{A}_{\boldsymbol{\beta}}} \max_{\mathbf{x} \in \mathcal{X}} (\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-1} \mathbf{f}(\mathbf{x})).$$

2.9 Convex Optimization for Linear Regression Design

The design criterion described in the above section, are only examples of a general class of functions $\phi : \mathbf{NNP}(s) \longrightarrow \mathbb{R}$, which satisfies the following three properties,

1. Monotonicity. If $\mathbf{0} \leq \mathbf{M}_1 \leq \mathbf{M}_2$, then $\phi(\mathbf{M}_1) \leq \phi(\mathbf{M}_2)$, here $\mathbf{M}_1 \leq \mathbf{M}_2$ $[\mathbf{M}_1 < \mathbf{M}_2]$ denotes that $\mathbf{M}_2 - \mathbf{M}_1$ is a nonnegative [positive] definite matrix. For $i = 1, 2$ let $\mathbf{M}_i = \boldsymbol{\mathcal{I}}_\psi(M_{\mathbf{f}}(\xi_i))$ be, in this case relative to the criterion ϕ the design ξ_2 for ψ is at least as good as the design ξ_1.

2. Concavity. For all $\mathbf{M}_1,\ \mathbf{M}_2 \in \mathbf{NNP}(s)$ and $t \in [0, 1]$, holds

$$\phi(t\mathbf{M}_1 + (1-t)\mathbf{M}_2) \geq t\phi(\mathbf{M}_1) + (1-t)\phi(\mathbf{M}_2)$$

When ϕ is strictly concave and finite, the ϕ-optimal moment matrix is unique in $\mathfrak{M}$. This, however, is the most that we can guarantee. Because of the possibility that two designs ξ and ξ' can have the same information matrix.

3. Differentiability. ϕ is differentiable, that is the Fréchet directional derivative of $\phi(\cdot)$ at all $\mathbf{M}_1 > \mathbf{0}$ in the direction of $\mathbf{M}_2$ defined as

$$F_\phi(\mathbf{M}_1, \mathbf{M}_2) = \lim_{t \to 0^+} \frac{1}{t}[\phi\{(1-t)\mathbf{M}_1 + t\mathbf{M}_2\} - \phi(\mathbf{M}_1)]$$

is linear in its second argument, in other words, it satisfies

$$F_\phi(\mathbf{M}_1, \sum a_i \mathbf{M}_i) = \sum a_i F_\phi(\mathbf{M}_1, \mathbf{M}_i),$$

where the a_i are real numbers such that $\sum_i a_i = 1$; (see, e.g., Silvey(1980, Appendix 3) and Rockafeller (1970, p.241)). With these properties we are in conditions to establish central theoretical results in the theory of the optimum design of experiment, for more details the following theorems can be seen in Silvey(1980),pp.19.

Theorem 2.10 (cf. Silvey(1980), Theorem 3.7) *Let $\delta_{\mathbf{x}}$ be the Dirac measure supported at $\mathbf{x}$ and ϕ a design criterion differentiable at $M_{\mathbf{f}}(\xi^*)$. Then the design ξ^* is ϕ-optimal for $\boldsymbol{\beta}$ if and only if*

$$F_\phi(M_{\mathbf{f}}(\xi^*), M_{\mathbf{f}}(\delta_{\mathbf{x}})) = F_\phi(M_{\mathbf{f}}(\xi^*), \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top) \leq 0 \quad \textit{for all } \mathbf{x} \in \mathcal{X}.$$

2.10 The Equivalence Theorem

The formulation of the design problem proposed above, in most real applications is computationally complex; in the practice we can use better the next theorems which provide tools for the construction and checking the optimality of a candidate approximate design.
First of all, if the attention is restricted to the full parameter $\boldsymbol{\beta}$, we consider the known in the literature as the **Equivalence Theorem**.

Theorem 2.11 (cf. Silvey(1980), Theorem 3.9) *If there exists a design with moment matrix positive definite and it is ϕ-optimal for $\boldsymbol{\beta}$, then the following two statements are equivalent*

1. *The approximate design ξ^* is ϕ-optimal for $\boldsymbol{\beta}$,*
2. $\max\limits_{\mathbf{x}\in\mathcal{X}} F_\phi(M_{\mathbf{f}}(\xi^*), \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top) = \min\limits_{\xi\in\mathfrak{A}} \max\limits_{\mathbf{x}\in\mathcal{X}} F_\phi(M_{\mathbf{f}}(\xi), \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top)$.

As an illustration consider the case where ϕ is the $D_{\boldsymbol{\beta}}$-criterion, then the Fréchet derivative in the direction $\mathbf{f}(\mathbf{x})\,\mathbf{f}(\mathbf{x})^\top$ of this function is given by

$$\begin{aligned} F_{\phi_0}(M_{\mathbf{f}}(\xi), \mathbf{f}(\mathbf{x})\,\mathbf{f}(\mathbf{x})^\top) &= [\mathrm{tr}\{\mathbf{f}(\mathbf{x})\,\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-1})\} - p]\det(M_{\mathbf{f}}(\xi)) \\ &= [\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-1})\mathbf{f}(\mathbf{x}) - p]\det(M_{\mathbf{f}}(\xi)). \end{aligned}$$

We have by definition 2.9 that an approximate design is G-optimal if

$$\max_{\mathbf{x}\in\mathcal{X}}(\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi^*)^{-1}\mathbf{f}(\mathbf{x})) = \min_{\xi\in\mathfrak{A}}\max_{\mathbf{x}\in\mathcal{X}}(\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-1}\mathbf{f}(\mathbf{x})).$$

Therefore a special case of the Theorems 2.10 and 2.11 is the essential content of the following well-known and praised first Equivalence Theorem of Kiefer and Wolfowitz (1960)

Theorem 2.12 (cf. Kiefer and Wolfowitz (1960)) *If the regression range* $\mathbf{f}(\mathcal{X}) \subseteq \mathbb{R}^p$ *and spans* $\mathbb{R}^p$*, then for a design* ξ *with moment matrix* $M(\xi)$*, positive definite, the following four statements are equivalent*

1. *the design* ξ *is* $D_{\boldsymbol{\beta}}$*-optimal*
2. *the design* ξ *is* G*-optimal*
3. $\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-1}\mathbf{f}(\mathbf{x}) \leq p \quad$ *for all* $\mathbf{x} \in \mathcal{X}$
4. $\max_{\mathbf{x}\in\mathcal{X}} \mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-1}\mathbf{f}(\mathbf{x}) = p.$

In case of optimality,

$$\mathbf{f}(\mathbf{x}_i)^\top M_{\mathbf{f}}(\xi)^{-1}\mathbf{f}(\mathbf{x}_i) = p, \quad \xi(\mathbf{x}_i) \leq \frac{1}{p}, \quad \text{for all } \mathbf{x}_i \in Supp(\xi).$$

Parallel versions of the Equivalence Theorem can be obtained for other criteria taking its particular form, for example the equivalence theorem for the A-criterion,

Theorem 2.13 (cf. Pukelsheim (2006),Theorem 9.7) *If the regression range* $\mathbf{f}(\mathcal{X}) \subseteq \mathbb{R}^p$ *and spans* $\mathbb{R}^p$*, then a design* ξ *with moment matrix* $M_{\mathbf{f}}(\xi)$*, positive definite, is* A*- or* ϕ_1*-optimal for* $\boldsymbol{\beta}$ *if and only if*

$$\mathbf{f}(\mathbf{x})^\top M_{\mathbf{f}}(\xi)^{-2}\mathbf{f}(\mathbf{x}) \leq \mathrm{tr}(M_{\mathbf{f}}(\xi)^{-1}) \quad \text{for all } \mathbf{x} \in \mathcal{X}$$

In case of optimality, any support point $\mathbf{x}_i$ *of the design* $\xi \in \mathfrak{A}_{\boldsymbol{\beta}}$ *satisfies*

$$\begin{aligned} \mathbf{f}(\mathbf{x}_i)^\top M_{\mathbf{f}}(\xi)^{-2}\,\mathbf{f}(\mathbf{x}_i) &= \mathrm{tr}(M_{\mathbf{f}}(\xi)^{-1}), \\ \xi(\mathbf{x}_i) &\leq \frac{\lambda_{\max}((M_{\mathbf{f}}(\xi)^{-1}))}{\mathrm{tr}(M_{\mathbf{f}}(\xi)^{-1})}. \end{aligned}$$

Where $\lambda_{\max}(\mathbf{B})$ *is the largest eigenvalue of the matrix* $\mathbf{B}$.

2.11 D-optimal Designs for Polynomial Models

A notable remark is that the Equivalence Theorem of Kiefer & Wolfowitz can be used for constructing $D_{\boldsymbol{\beta}}$-optimal designs, $\boldsymbol{\beta}$ the full parameter vector, for the $p = (q+1)$th order polynomial regression in a single control variable. The model is

$$\mathrm{E}(Y(\mathbf{x})) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} = \beta_0 + \beta_1 x + \cdots + \beta_q x^q, \tag{2.20}$$

$$\mathrm{Var}(Y(\mathbf{x})) = \mathrm{Var}(\varepsilon) = \sigma^2 \tag{2.21}$$

where the vector valued function $\mathbf{f}(x) = (1, x, \ldots, x^q)^\top \in \mathbb{R}^{q+1}$, the experimental region is the closed interval $[-1, 1] = \mathcal{X}$ in $\mathbb{R}$, $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_q)^\top$ is the full vector of unknown parameters and $\sigma^2 (> 0)$ is an unknown scalar parameter.

In this case, (see, e.g. Fedorov (1972, p. 89) Silvey (1980, p. 42) and Pukelsheim (2006, p.213)) the $D_{\boldsymbol{\beta}}$-optimal design is the uniform design on the support set $\{x_{(1)}, \ldots, x_{(q+1)}\}$ of $q+1$ solutions of the equation

$$(1 - x^2) P_q'(x) = 0, \tag{2.22}$$

where $P_q'(x)$ is the derivative of the qth order Legendre Polynomial, which has explicit representation given by the formula

$$P_q(x) = \frac{1}{2^q} \sum_{k=0}^{q} \binom{q}{k}^2 (x-1)^{q-k} (x+1)^k.$$

Because of the first factor of the equation (2.22), two solutions are the boundary points $x_{(1)} = -1$ and $x_{(q+1)} = 1$. When a total of N observations are taken and $N = m * (q+1)$, we have that a $D_{\boldsymbol{\beta}}$-optimal experimental design can be constructed where each design point $x_{(i)}, \quad i = 1, \ldots q+1$; is replicated the same number m of times.

For example, when $q = 2$, the quadratic regression, the design points are $x_{(1)} = -1,\ x_{(3)} = 1$ and the value for which the derivative of $P_2(x) = \dfrac{3x^2-1}{2}$ is zero, that is, $x_{(2)} = 0$. Thus the corresponding $D_{\boldsymbol{\beta}}$-optimal design for estimating $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^\top$ is

$$\xi^* = \left\{ \begin{array}{ccc} -1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{array} \right\}, \tag{2.23}$$

with moment matrix given by

$$\begin{aligned} M_{\mathbf{f}}(\xi^*) &= \int_{[-1,1]} \mathbf{f}(x)\ \mathbf{f}(x)^\top \xi^*(\mathrm{d}x) \\ &= \sum_{x \in \{-1,0,1\}} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} \begin{pmatrix} 1 & x & x^2 \end{pmatrix} 1/3 = 1/3 \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} > \mathbf{0} \end{aligned}$$

Now we obtain the standardized variance function of ξ^* at $x \in [-1, 1]$,

$$\begin{aligned} d(x, \xi^*) &= \mathbf{f}(x)^\top M_{\mathbf{f}}(\xi^*)^{-1} \mathbf{f}(x) \\ &= 3 \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1/2 & 0 \\ -1 & 0 & 3/2 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} \\ &= \frac{3}{4}[4 - 6x^2(1-x^2)] \leq \mathbf{3} \end{aligned}$$

and we can verify that it has maximum occurring at -1, 0 and 1, the points of support of the approximate design ξ^*. In Figure 2.1, we can see the curve of the standardized variance function $d(x, \xi^*)$ for the $D_{\boldsymbol{\beta}}$-optimal approximate design in the quadratic regression model.

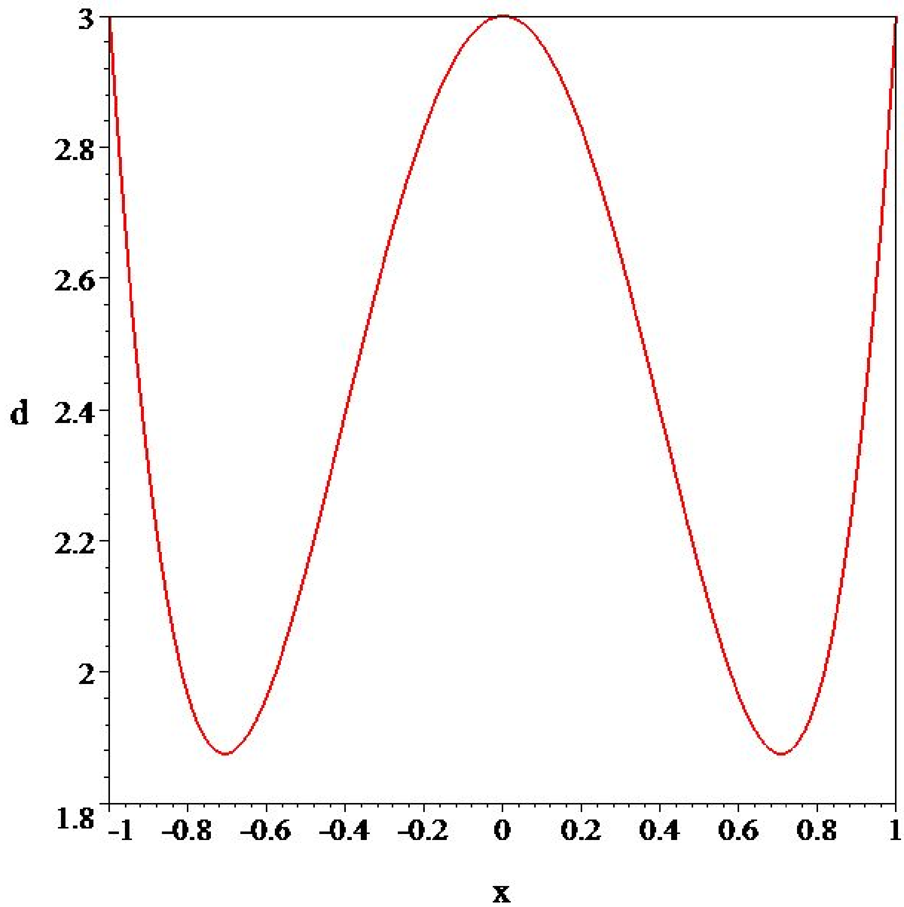

Figure 2.1: Standardized variance function $d(x, \xi^*)$ for the D-optimal design; quadratic regression

Table 2.1: Optimal designs for some polynomial regression models

Model E(Y)	Optimality criterion	Experimental region	Optimal design
$\beta_0 + \beta_1 x$	D, G, A	$[-1, 1]$	$\xi^* = \begin{Bmatrix} -1 & 1 \\ 1/2 & 1/2 \end{Bmatrix}$
$\beta_0 + \beta_1 x + \beta_2 x^2$	D, G	$[-1, 1]$	$\xi^* = \begin{Bmatrix} -1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{Bmatrix}$
$\beta_0 + \beta_1 x + \beta_2 x^2$	A, D_{β_2}	$[-1, 1]$	$\xi^* = \begin{Bmatrix} -1 & 0 & 1 \\ 1/4 & 1/2 & 1/4 \end{Bmatrix}$
$\beta_0 + \beta_1 x + \beta_2 x^2$	$D_{(\beta_2 - \beta_1)}$	$[-1, 1]$	$\xi^* = \begin{Bmatrix} -1 & 0 \\ 1/2 & 1/2 \end{Bmatrix}$
$\beta_1 x + \beta_2 x^2$	D, G	$[0, 1]$	$\xi^* = \begin{Bmatrix} \sqrt{2} - 1 & 1 \\ \sqrt{2}/2 & 1 - \sqrt{2}/2 \end{Bmatrix}$

Chapter 3

Blocking Response Surface Experiments

3.1 Introduction

Very often in experimental practice, the usual linear model (2.1) presents a variant, because, for example, the experimental runs cannot be performed under the assumption of homogeneous conditions, hence for a more realistic analysis of the data in these cases, the experiment has to be *blocked*, that is, it identifies groups or blocks of experimental units within which the homogeneity of conditions can be assumed. Blocks can be, for example, days, batches, or test subjects. Thus now the experiment involve a blocking variable, qualitative variable, which unlike the quantitative explanatory variable, is not under the direct control of the experimenter, but it can be adjusted to a finite number of levels (blocks) and the variation produced by the different blocks in the experiment is accounted for by including new parameters called, block effects, in the statistical model (see e.g. Khuri(1992) and Goos (2002), besides others).

We consider henceforward, the following assumptions:

1. The effect of the qualitative variable is purely additive, that is, we have a two-factor linear model without interaction, in our case there is no interaction between blocks and the experimental quantitative factors effects.

2. There is an explicit term constant or intercept included in the model, which depends on the considered problem.

3. The block effects into the model will be treated as nuisance parameters.

4. Depending on the nature of the experiment two types of blocking variables, fixed or random, can be assumed.

3.2 The block effects model

The expression for a statistical linear model with intercept for a blocked experiment that consist in b blocks, $i = 1, \dots, b$ with m_i observations each one, can be written as

$$Y_{ij} \quad = \quad \beta_0 + \mathbf{f}(\mathbf{x}_{ij})^\top \boldsymbol{\beta} + \gamma_i + \epsilon_{ij}, \tag{3.1}$$

where for $j = 1, \dots, m_i$; Y_{ij} is the response of the jth observation, among the m_i at the block i, of the experimental setting $\mathbf{x}_{ij}$, $\mathbf{f} = (f_1, \dots, f_q)^\top$ is a vector of q known regression function defined on some compact subset $\mathcal{X}$ of $\mathbb{R}^r$, β_0 is the intercept, $\boldsymbol{\beta}$ is a q- vector of parameters, that contains all quantitative factor effects, the term γ_i denotes the additive i-th block effect, which ensures the block-to-block variation in the responses and it assumes that ϵ_{ij}, the experimental random errors of the run j on block i, are independent and identically normal distributed with expected value zero and variance σ^2,

$$\epsilon_{ij} \sim N(0, \sigma^2).$$

We assume that the block sizes are balanced and fixed, that is the number of observations per block is constant, i.e. $m_i \equiv m$, also for each block i of observations, there is associated an experimental design on $\mathcal{X}$, denoted by $\mathcal{B}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{im})$.

As a remark, the experimental settings $\mathbf{x}_{ij}$ for each block i need not be all distinct and the performance does not depend on the order of the observations within the blocks, hence we can rewrite $\mathcal{B}_i$ in terms of its *distinct* settings, $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iS_i}$ together with their respective number $m_{i1}, \ldots, m_{iS_i}$ of replications such that $\sum_{s=1}^{S_i} m_{is} = m$. Thus we have now the exact design for block i

$$\mathcal{B}_i^{(m)} = \left\{ \begin{array}{cccc} \mathbf{x}_{i1} & \mathbf{x}_{i2} & \ldots & \mathbf{x}_{iS_i} \\ \\ m_{i1} & m_{i2} & \ldots & m_{iS_i} \end{array} \right\}, \tag{3.2}$$

and we will denote the set of all exact design for block with m observations by $\mathfrak{B}^{(m)}$.
Now, we specify the experimental settings for the whole sample of b blocks. Similar to the experimental design defined for the ordinary linear model in Chapter 2, a *block design* $\mathcal{B}$ of size b is a set of designs for block $\mathcal{B}_i^{(m)} \in \mathfrak{B}^{(m)}$ $i = 1, \ldots, b$ not necessarily all distinct.

On the other side, we can regard, for example, that the points $\mathbf{x}_{is}$, for all $i = 1, \ldots, b$; $s = 1, \ldots, S_i$; are elements of a set of distinct points, or "treatments", say $\mathbf{x}_1, \ldots, \mathbf{x}_T$; with $T > S_i$, therefore the blocking response surface experiment can be considered as an incomplete block design. Hence when a same "treatment" $\mathbf{x}_t$ for a determinate t, $1 \leq t \leq T$; is assigned to more than one block, and we can have, for instance, $\mathbf{x}_t = \mathbf{x}_{is} = \mathbf{x}_{i's'}$ for $i \neq i'$ and in this case m_{is} is the number of appearances of treatment $\mathbf{x}_t = \mathbf{x}_{is}$ in block i.

The m observations in each block at $\mathcal{B}_i^{(m)}$ can be regarded as a m-variate response. Thus the vector of observations for block becomes

$$\begin{aligned} \mathbf{Y}_i &= (\mathbf{1}_m, \mathbf{F}_i)(\beta_0, \boldsymbol{\beta}^\top)^\top + \mathbf{1}_m \gamma_i + \boldsymbol{\epsilon}_i \\ &= \mathbf{G}_i \boldsymbol{\theta} + \mathbf{1}_m \gamma_i + \boldsymbol{\epsilon}_i \end{aligned} \tag{3.3}$$

here $\mathbf{1}_m$ is a vector of length m with all entries equal to one, $\mathbf{G}_i = (\mathbf{1}_m, \mathbf{F}_i)$ is the design matrix for block i, which is partitioned into the first column of ones corresponding to constant intercept and the design matrix

$$\mathbf{F}_i = \mathcal{F}(\mathcal{B}_i^{(m)}) = (\underbrace{\mathbf{f}(\mathbf{x}_{i1}), \ldots, \mathbf{f}(\mathbf{x}_{i1})}_{m_{i1}\ times}, \ldots, \underbrace{\mathbf{f}(\mathbf{x}_{iS_i}), \ldots, \mathbf{f}(\mathbf{x}_{iS_i})}_{m_{iS_i}\ times})^\top \tag{3.4}$$

for the vector parameter $\boldsymbol{\beta}$, $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}^\top)^\top$ and $\boldsymbol{\epsilon}_i$ is the vector of corresponding observational random errors.

Depending on the nature of the experiment two types of blocking variables, fixed or random, can be considered, which imply different statistical models.

1. When the block effects are regarded as fixed, because there are no available inter-block information, the distributional assumption is associated only with the vector of random errors,
   $\forall\ i,\ i' \in \{1, \ldots, b\}$

   $$\mathrm{E}(\boldsymbol{\epsilon}_i) = \mathbf{0}_m, \quad \mathrm{Cov}(\boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_{i'}) = \delta_{ii'}\sigma^2\mathbf{I}_m \tag{3.5}$$

   and the linear identifiability condition $\gamma_b = -\sum_{i=1}^{b-1}\gamma_i$ of the block effects of the model is imposed, with the advantage that it preserves the interpretation of the intercept β_0 as the overall average response across the blocks.
   Here, $\delta_{ii'} = \begin{cases} 1 & \text{if } i = i' \\ 0 & \text{if } i \neq i' \end{cases}$, which is known as the Kronecker delta.

2. The blocking variable brings random effects to the model (3.3), when the blocks can be identifiable as random choice from a population of blocks. Thus now in vector notation the model contains the random block effects γ_i together with the vector of random errors $\boldsymbol{\epsilon}_i$, where the random block effects are independently of each other and of the vector of random errors and they have expected value zero and same variance σ_γ^2, thus the distributional assumptions of the model are given by

$\forall\, i,\; i' \in \{1, \ldots, b\}$

$$\mathrm{E}(\gamma_i) = 0, \quad \mathrm{Cov}(\gamma_i, \gamma_{i'}) = \delta_{ii'}\sigma_\gamma^2, \tag{3.6}$$

$$\mathrm{E}(\boldsymbol{\epsilon}_i) = \mathbf{0}_m, \quad \mathrm{Cov}(\boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_{i'}) = \delta_{ii'}\sigma^2\mathbf{I}_m \tag{3.7}$$

$$\text{and} \qquad \mathrm{Cov}(\gamma_i, \boldsymbol{\epsilon}_{i'}) = \mathbf{0}_m. \tag{3.8}$$

3.3 Designs for Fixed Block Effects Model

The observations of an experiment with a blocking variable, where we assume that the blocks under study are chosen directly by the experimenter, because they are the only blocks of interest, can be analyzed using fixed block effects. Hence the estimation of the regression parameter in the fixed block effects model is a special case of standard analytical technique of ordinary least squares (OLS) regression.

Let $\boldsymbol{\gamma}_f = (\gamma_{f1}, \ldots, \gamma_{f(b-1)})^\top$ be the $(b-1)$-vector of fixed block effects and the last block effect $\gamma_{fb} = -\mathbf{1}_{(b-1)}^\top \boldsymbol{\gamma}_f$, then the model(3.1) involves an intercept, a q-vector of quantitative factor effects and b fixed block effects, thus the vector of observations for block $i = 1, \ldots, b$ becomes

$$\begin{aligned} \mathbf{Y}_i &= (\mathbf{1}_m, \mathbf{F}_i)(\beta_0, \boldsymbol{\beta}^\top)^\top + \mathbf{1}_m\gamma_{fi} + \boldsymbol{\epsilon}_i \\ &= \mathbf{G}_i\boldsymbol{\theta} + \mathbf{H}(i)\boldsymbol{\gamma}_f + \boldsymbol{\epsilon}_i \\ &= \big(\mathbf{G}_i, \mathbf{H}(i)\big)\big(\boldsymbol{\theta}^\top,\, \boldsymbol{\gamma}_f^\top\big)^\top + \boldsymbol{\epsilon} \end{aligned} \tag{3.9}$$

where

$$\mathbf{H}(i) = \big(H_1(i), \ldots, H_{(b-1)}(i)\big),$$

$$H_k(i) = \begin{cases} \mathbf{1}_m & \text{if } \; i = k \\ -\mathbf{1}_m & \text{if } \; i = b \\ \mathbf{0} & \text{otherwise} \end{cases}$$

If the observations of the whole sample of blocks are summarized as

$$\mathbf{Y} \;=\; \left(\mathbf{Y}_1^\top, \mathbf{Y}_2^\top, \ldots, \mathbf{Y}_b^\top\right)^\top$$

in matrix notation the $f$ixed block effects model becomes

$$\begin{aligned}\mathbf{Y} &= \mathbf{G}\boldsymbol{\theta} + \mathbf{H}_f\boldsymbol{\gamma}_f + \boldsymbol{\epsilon}\\ &= \left(\mathbf{G}, \mathbf{H}_f\right)\left(\boldsymbol{\theta}^\top, \, \boldsymbol{\gamma}_f^\top\right)^\top + \boldsymbol{\epsilon}\end{aligned} \tag{3.10}$$

where

$$\mathbf{G} = \left(\mathbf{G}_1^\top, \ldots, \mathbf{G}_b^\top\right)^\top = \begin{pmatrix} \mathbf{1}_m & \mathbf{F}_1 \\ \mathbf{1}_m & \mathbf{F}_2 \\ \vdots & \vdots \\ \mathbf{1}_m & \mathbf{F}_b \end{pmatrix} = \left(\mathbf{1}_{bm}, \, \mathbf{F}\right),$$

$$\mathbf{F} = \left(\mathbf{F}_1^\top, \ldots, \mathbf{F}_b^\top\right)^\top,$$

$$\begin{aligned}\mathbf{H}_f &= \begin{pmatrix} \mathbf{H}(1) \\ \mathbf{H}(2) \\ \vdots \\ \mathbf{H}(b-1) \\ \mathbf{H}(b) \end{pmatrix} = \begin{pmatrix} \mathbf{1}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m \\ \mathbf{0}_m & \mathbf{1}_m & \cdots & \mathbf{0}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{1}_m \\ -\mathbf{1}_m & -\mathbf{1}_m & \cdots & -\mathbf{1}_m \end{pmatrix}_{bm\times(b-1)}\\ &= \left(\mathbf{I}_{b-1}, -\mathbf{1}_{b-1}\right)^\top \otimes \mathbf{1}_m\end{aligned}$$

and

$$\boldsymbol{\epsilon} \;=\; \left(\boldsymbol{\epsilon}_1^\top, \boldsymbol{\epsilon}_2^\top, \ldots, \boldsymbol{\epsilon}_b^\top\right)^\top.$$

Here we use the subscript f for the effects-type coding of the block effects, $\mathbf{H}_f$ is the design matrix corresponding to the (b-1)-vector $\boldsymbol{\gamma}_f$ of the fixed block effects and $\otimes$ denotes the Kronecker product of two matrices or vectors.

It follows that

$$\begin{aligned}
\mathrm{E}(\mathbf{Y}) &= (\mathbf{G},\, \mathbf{H}_f)(\boldsymbol{\theta}^\top, \boldsymbol{\gamma}_f^\top)^\top = (\mathbf{1}_{bm},\, \mathbf{F},\, \mathbf{H}_f)(\beta_0,\ \boldsymbol{\beta}^\top,\ \boldsymbol{\gamma}_f^\top)^\top && (3.11)\\
\mathrm{Cov}(\mathbf{Y}) &= \sigma^2 \mathbf{I}_{bm} && (3.12)
\end{aligned}$$

As a result, when we use fixed block effects, then it requires the estimation of as many as block effects as we have blocks in our experiment.
On the other hand, if the columns of $\mathbf{F}$ are linearly independent of $\mathbf{1}_{bm}$ and the columns of $\mathbf{H}_f$, then the partitioned matrix $(\mathbf{1}_{bm},\, \mathbf{F},\, \mathbf{H}_f)$ is of full rank $1 + q + b - 1 = q + b$, the number of parameters. As a result the ordinary least squares estimators of the intercept β_0 and the quantitative factor effects $\boldsymbol{\beta}$, components of the vector $\boldsymbol{\theta}$ and the $(b-1)$ fixed block effects $\boldsymbol{\gamma}_f$ are given by

$$\begin{aligned}
\begin{pmatrix} \widehat{\boldsymbol{\theta}} \\ \widehat{\gamma}_f \end{pmatrix} &= ((\mathbf{G},\, \mathbf{H}_f)^\top(\mathbf{G},\, \mathbf{H}_f))^{-1}(\mathbf{G},\, \mathbf{H}_f)^\top \mathbf{Y} \\
&= \begin{pmatrix} \mathbf{G}^\top\mathbf{G} & \mathbf{G}^\top\mathbf{H}_f \\ \mathbf{H}_f^\top\mathbf{G} & \mathbf{H}_f^\top\mathbf{H}_f \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{G}^\top \\ \mathbf{H}_f^\top \end{pmatrix} \mathbf{Y}. && (3.13)
\end{aligned}$$

Thus the information matrix of $\left(\boldsymbol{\theta}^\top, \boldsymbol{\gamma}_f^\top\right)^\top$ coincide with the moment matrix of the experimental block design $\mathcal{B}$ with b exact design for blocks $\mathcal{B}_i^{(m)}$ $i = 1, \ldots, b$.

$$bmM(\mathcal{B}) = \begin{pmatrix} \mathbf{G}^\top\mathbf{G} & \mathbf{G}^\top\mathbf{H}_f \\ \mathbf{H}_f^\top\mathbf{G} & \mathbf{H}_f^\top\mathbf{H}_f \end{pmatrix} \tag{3.14}$$

where in particular

$$\begin{aligned}
\mathbf{G}^\top\mathbf{G} &= \sum_{i=1}^{b} \mathbf{G}_i^\top \mathbf{G}_i \\
&= \sum_{i=1}^{b} (\mathbf{1}_m,\, \mathbf{F}_i)^\top (\mathbf{1}_m,\, \mathbf{F}_i) \\
&= \sum_{i=1}^{b} \begin{pmatrix} m & \mathbf{1}_m^\top \mathbf{F}_i \\ \mathbf{F}_i^\top \mathbf{1}_m & \mathbf{F}_i^\top \mathbf{F}_i \end{pmatrix}
\end{aligned}$$

when we use the following notation

$$\mathbf{F}_i^\top \mathbf{F}_i = \sum_{s=1}^{S_i} m_{is} \mathbf{f}(\mathbf{x}_{is}) \mathbf{f}(\mathbf{x}_{is})^\top = m\widehat{\mathbf{F}}_i^\top \mathbf{W}_i \widehat{\mathbf{F}}_i \tag{3.15}$$

and

$$\mathbf{F}_i^\top \mathbf{1}_m = \sum_{s=1}^{S_i} \mathbf{f}(\mathbf{x}_{is}) m_{is} = m\widehat{\mathbf{F}}_i^\top \mathbf{W}_i \mathbf{1}_{S_i} \tag{3.16}$$

where $\mathbf{W}_i$ is the diagonal matrix with the proportions $m_{is}/m = w_{is}, \;\; s = 1, \ldots, S_i$ as diagonal entries and the design matrix $\widehat{\mathbf{F}}_i$ is evaluated at the support experimental settings $\mathbf{x}_{i1}, \mathbf{x}_{i2} \ldots, \mathbf{x}_{S_i}$. Then this representation can be also used for approximate design for block, that is where now in $\mathcal{B}_i^{(m)}$ the $m_{is} \in \mathbb{R}, \;\; \sum_{s=1}^{S_i} m_{is} = m;\, m_{is} > 0$.

For the construction of D-optimal design for this model (3.12) and since the fixed block effects $\boldsymbol{\gamma}_f$ are not of primary interest, we can use the particular criterion D_s-optimality, in our case $D_{\boldsymbol{\theta}}$-optimality, because the interests is in the effects of the parameter $\boldsymbol{\theta}$ only.
Using $D_{\boldsymbol{\theta}}$-optimality requires maximization of

$$\begin{aligned}
\det \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}(M(\mathcal{B})) &= \frac{\det(M(\mathcal{B}))}{\det\left(\mathbf{H}_f^\top \mathbf{H}_f\right)} \\
&= \det\left(\mathbf{G}^\top\mathbf{G} - \mathbf{G}^\top \mathbf{H}_f (\mathbf{H}_f^\top \mathbf{H}_f)^{-1} \mathbf{H}_f^\top \mathbf{G}\right)
\end{aligned}$$

and due to that the block size are fixed, then $\det\left(\mathbf{H}_f^\top \mathbf{H}_f\right)$ is independent of the experimental setting. Therefore maximizing $\det \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}(M(\mathcal{B}))$ is equivalent to maximizing $\det(M(\mathcal{B}))$. Hence, D- and $D_{\boldsymbol{\theta}}$-optimality designs coincide in fixed block effects model.

3.4 The Fixed Block Experiments Viewed as a Two-Factor Model

The linear regression model in the presence of fixed block effects can be viewed as a two kind factor model without interaction between the factor effects, indeed the response of the fixed effects model (3.1) at the ith block can be rewritten as

$$\begin{aligned} Y_{ij}(i, \mathbf{x}_{ij}^\top) &= \beta_0 + \gamma_{fi} + \mathbf{f}(\mathbf{x}_{ij})^\top \boldsymbol{\beta} + \epsilon_{ij} \\ &= (1_{\{1\}}(i), \ldots, 1_{\{b\}}(i))(\beta_0 + \gamma_{f1}, \ldots, \beta_0 + \gamma_{fb})^\top + \mathbf{f}(\mathbf{x}_{ij})^\top \boldsymbol{\beta} + \epsilon_{ij} \\ &= (\mathbf{a}(i)^\top, \mathbf{f}(\mathbf{x}_{ij})^\top)(\beta_0 + \gamma_{f1}, \ldots, \beta_0 + \gamma_{fb}, \, \boldsymbol{\beta}^\top)^\top + \epsilon_{ij} \\ &= \mathbf{g}(i, \mathbf{x}_{ij}^\top)^\top \left(\boldsymbol{\mu}_f^\top, \, \boldsymbol{\beta}^\top\right)^\top + \epsilon_{ij}, \end{aligned} \tag{3.17}$$

where the $1_{\{k\}}(i)$ are the indicators function, thus $\mathbf{a}(i)$ is a vector of length b with its ith entry equal to one and all other entries equal to zero, $\mathbf{g}(i, \mathbf{x}_{ij}^\top)^\top = (\mathbf{a}(i)^\top, \mathbf{f}(\mathbf{x}_{ij})^\top)$ and $\boldsymbol{\mu}_f = (\beta_0 + \gamma_{f1}, \ldots, \beta_0 + \gamma_{fb})^\top$. Let $\mathcal{X}_b = \{1, \ldots, b\}$ be the index set of blocks and the Cartesian product set $\mathcal{X}_b \times \mathcal{X}$ the induced new experimental region.

We consider an approximate design $\mathcal{B}$ on $\mathcal{X}_b \times \mathcal{X}$ with bm observations, which can be written as

$$\mathcal{B}^{(bm)}(i, \mathbf{x}) = (bm)^{-1} \, \mathcal{B}_i^{(m)}(\mathbf{x}),$$

Hence the moment matrix of the design $\mathcal{B}^{(bm)}$ on $\mathcal{X}_b \times \mathcal{X}$ for the model (3.17) is presented in the following form

$$
\begin{aligned}
M_{\mathbf{g}}(\mathcal{B}^{(bm)}) &= \int_{\mathcal{X}_b \times \mathcal{X}} \mathbf{g}(i, \mathbf{x}^\top)\mathbf{g}(i, \mathbf{x}^\top)^\top \mathcal{B}^{(bm)}(\mathrm{d}(i, \mathbf{x})) \\
&= (bm)^{-1} \sum_{i=1}^{b} \left(\int_{\mathcal{X}} \mathbf{g}(i, \mathbf{x}^\top)\mathbf{g}(i, \mathbf{x}^\top)^\top \, \mathcal{B}_i^{(m)}(\mathrm{d}\mathbf{x}) \right) \\
&= (bm)^{-1} \sum_{i=1}^{b} \, M_{\mathbf{g}}(\mathcal{B}_i^{(m)})
\end{aligned}
$$

where, for $i = 1, \ldots, b$

$$
M_{\mathbf{g}}(\mathcal{B}_i^{(m)}) = \int_{\mathcal{X}} \mathbf{g}(i, \mathbf{x}^\top)\mathbf{g}(i, \mathbf{x}^\top)^\top \, \mathcal{B}_i^{(m)}(\mathrm{d}\mathbf{x})
$$

$$
\begin{aligned}
M_{\mathbf{g}}(\mathcal{B}_i^{(m)}) &= \int_{\mathcal{X}} (\mathbf{a}(i)^\top, \mathbf{f}(\mathbf{x})^\top)^\top \, (\mathbf{a}(i)^\top, \mathbf{f}(\mathbf{x})^\top) \, \mathcal{B}_i^{(m)}(\mathrm{d}\mathbf{x}) \\
&= \begin{pmatrix} \mathbf{a}(i)\mathbf{a}(i)^\top & \int_{\mathcal{X}} \mathbf{a}(i) \, \mathbf{f}(\mathbf{x})^\top \, \mathcal{B}_i^{(m)}(\mathrm{d}\mathbf{x}) \\ \int_{\mathcal{X}} \mathbf{f}(\mathbf{x}) \, \mathbf{a}(i)^\top \, \mathcal{B}_i^{(m)}(\mathrm{d}\mathbf{x}) & \int_{\mathcal{X}} \mathbf{f}(\mathbf{x}) \, \mathbf{f}(\mathbf{x})^\top \, \mathcal{B}_i^{(m)}(\mathrm{d}\mathbf{x}) \end{pmatrix} \\
&= m \begin{pmatrix} \mathbf{a}(i)\mathbf{a}(i)^\top & \mathbf{a}(i) \, \sum_{s=1}^{S_i} \mathbf{f}(\mathbf{x}_{is})^\top w_{is} \\ \mathbf{a}(i)^\top \, \sum_{s=1}^{S_i} \mathbf{f}(\mathbf{x}_{is}) w_{is} & \sum_{s=1}^{S_i} w_{is} \mathbf{f}(\mathbf{x}_{is}) \mathbf{f}(\mathbf{x}_{is})^\top \end{pmatrix} \\
&= m \begin{pmatrix} \mathbf{a}(i)\mathbf{a}(i)^\top & \mathbf{a}(i) \, \mathbf{1}_{S_i}^\top \mathbf{W}_i \widehat{\mathbf{F}}_i \\ \mathbf{a}(i)^\top \, \widehat{\mathbf{F}}_i^{\top} \mathbf{W}_i \mathbf{1}_{S_i} & \widehat{\mathbf{F}}_i^{\top} \mathbf{W}_i \widehat{\mathbf{F}}_i \end{pmatrix}.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
M_{\mathbf{g}}(\mathcal{B}^{(bm)}) &= (bm)^{-1}\sum_{i=1}^{b} M_{\mathbf{g}}(\mathcal{B}_i^{(m)}) \\
&= b^{-1}\begin{pmatrix} \sum_{i=1}^{b}\mathbf{a}(i)\mathbf{a}(i)^\top & \sum_{i=1}^{b}\mathbf{a}(i)\,\mathbf{1}_{S_i}^\top\mathbf{W}_i\widehat{\mathbf{F}}_i \\ \sum_{i=1}^{b}\mathbf{a}(i)^\top\,\widehat{\mathbf{F}}_i^\top\mathbf{W}_i\mathbf{1}_{S_i} & \sum_{i=1}^{b}\widehat{\mathbf{F}}_i^\top\mathbf{W}_i\widehat{\mathbf{F}}_i \end{pmatrix} \\
&= b^{-1}\left(\begin{array}{ccc|c} & & & \mathbf{1}_{S_1}^\top\mathbf{W}_1\widehat{\mathbf{F}_1} \\ & \mathbf{I}_b & & \vdots \\ & & & \mathbf{1}_{S_b}^\top\mathbf{W}_b\widehat{\mathbf{F}_b} \\ \hline \widehat{\mathbf{F}_1}^\top\mathbf{W}_1\mathbf{1}_{S_1} & \cdots & \widehat{\mathbf{F}_b}^\top\mathbf{W}_b\mathbf{1}_{S_b} & \sum_{i=1}^{b}\widehat{\mathbf{F}}_i^\top\mathbf{W}_i\widehat{\mathbf{F}}_i \end{array}\right).
\end{aligned}
$$

The following equation of determinants holds, because $M_{\mathbf{g}}(\mathcal{B}^{(bm)})$ is a partitioned positive definite symmetric matrix

$$
\det\left(M_{\mathbf{g}}(\mathcal{B}^{(bm)})\right) \;\propto\; \det\left(\mathbf{J}_{22} - \mathbf{J}_{12}^\top\,\mathbf{J}_{12}\right)
$$

where

$$
\begin{aligned}
\mathbf{J}_{12}^\top = \mathbf{J}_{21} &= \sum_{i=1}^{b}\mathbf{a}(i)^\top\,\widehat{\mathbf{F}}_i^\top\mathbf{W}_i\mathbf{1}_{S_i} \\
&= \left(\widehat{\mathbf{F}_1}^\top\mathbf{W}_1\mathbf{1}_{S_1},\cdots,\widehat{\mathbf{F}_b}^\top\mathbf{W}_b\mathbf{1}_{S_b}\right)
\end{aligned}
$$

$$
\mathbf{J}_{22} = \sum_{i=1}^{b}\widehat{\mathbf{F}}_i^\top\mathbf{W}_i\widehat{\mathbf{F}}_i
$$

and by (2.17) we have

$$
\det\left(M_{\mathbf{g}}(\mathcal{B}^{(bm)})\right) \;\propto\; \det\left(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{g}}(\mathcal{B}^{(bm)}))\right)
$$

Hence, D- and $D_{\boldsymbol{\beta}}$-optimality designs coincide in fixed block effects model, a known fact in the model without blocks. Thus we have the following

Lemma 3.1 *When the effects of a blocking variable and the block size are assumed fixed, a given uniform block structure* $\mathcal{B}$, *which maximizes* $\det(\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{g}}(\mathcal{B})))$, *then also maximizes* $\det(M_{\mathbf{g}}(\mathcal{B}))$.

But yet the regression parameter $\boldsymbol{\theta}$ [$\boldsymbol{\beta}$] cannot be estimated independently of the fixed block effects $\boldsymbol{\gamma}_f$ [$\boldsymbol{\mu}_f$] in the model. However sometimes we can avoid this fact when is possible to choose an orthogonal blocking design.

3.5 Orthogonal Blocking

By definition, a design is orthogonally blocked if the columns of the design matrix $\mathbf{G}$ for the regression parameter $\boldsymbol{\theta}$ are orthogonal to those of the design matrix $\mathbf{H}_f$ for $\boldsymbol{\gamma}$, the $(b-1)$-vector of fixed block effects, that is if

$$\mathbf{G}^\top \mathbf{H}_f := \mathbf{0}_{(q+1)\times(b-1)}$$

where $\mathbf{0}_{(q+1)\times(b-1)}$ is a $(q+1)\times(b-1)$ matrix of zeros. Now,

$$\begin{aligned}
\mathbf{G}^\top \mathbf{H}_f &= \begin{pmatrix} \mathbf{1}_b^\top \otimes \mathbf{1}_m^\top \\ \\ (\mathbf{F}_1^\top, \ldots, \mathbf{F}_b^\top) \end{pmatrix} \left(\begin{pmatrix} \mathbf{I}_{b-1} & -\mathbf{1}_{b-1} \end{pmatrix}^\top \otimes \mathbf{1}_m \right) \\
&= \begin{pmatrix} \mathbf{1}_{b-1}^\top \otimes \mathbf{1}_m^\top & \mathbf{1}_m^\top \\ \\ (\mathbf{F}_1^\top, \ldots, \mathbf{F}_{b-1}^\top) & \mathbf{F}_b^\top \end{pmatrix} \begin{pmatrix} \mathbf{I}_{b-1} \otimes \mathbf{1}_m \\ \\ -\mathbf{1}_{b-1}^\top \otimes \mathbf{1}_m \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{1}_{b-1}^\top m - \mathbf{1}_{b-1}^\top m \\ \\ (\mathbf{F}_1^\top \mathbf{1}_m, \ldots, \mathbf{F}_{b-1}^\top \mathbf{1}_m) - \mathbf{1}_{b-1}^\top \otimes \mathbf{F}_b^\top \mathbf{1}_m \end{pmatrix} = \mathbf{0}_{(q+1)\times(b-1)}
\end{aligned}$$

this implies that,

$$\mathbf{F}_i^\top \mathbf{1}_m = \mathbf{F}_b^\top \mathbf{1}_m \quad \text{for all } i = 1, \ldots, b-1$$

As a result, the condition for orthogonality holds if

$$(\mathbf{F}_1^\top, \ldots, \mathbf{F}_b^\top)(\mathbf{1}_b \otimes \mathbf{1}_m) = b\mathbf{F}_i^\top \mathbf{1}_m \quad \text{for all } i = 1, \ldots, b$$

that is

$$\frac{1}{m}\mathbf{F}_i^\top \mathbf{1}_m = \frac{1}{bm}\mathbf{F}^\top \mathbf{1}_{bm} \quad \text{for all } i = 1, \ldots, b$$

or equivalently, in case of a approximate design for block

$$\begin{aligned} \frac{1}{m}\mathbf{F}_i^\top \mathbf{1}_m &= \frac{1}{bm}\mathbf{F}^\top \mathbf{1}_{bm} \\ \widehat{\mathbf{F}}_i^\top \mathbf{W}_i \mathbf{1}_{S_i} &= \frac{1}{b}\sum_{i=1}^{b} \widehat{\mathbf{F}}_i^\top \mathbf{W}_i \mathbf{1}_{S_i} \\ &= \widehat{\mathbf{F}}^\top \mathbf{W} \mathbf{1}_{S^*} \end{aligned}$$

where

$$\widehat{\mathbf{F}} = (\widehat{\mathbf{F}}_1^\top, \ldots, \widehat{\mathbf{F}}_b^\top)^\top, \quad \mathbf{W} = \frac{1}{b}\mathrm{diag}(\mathbf{W}_1, \ldots, \mathbf{W}_b) \quad \text{and} \quad S^* = \sum_{i=1}^{b} S_i.$$

Thus, in an orthogonally blocking design for experiments involving quantitative variables, the average columns of all design matrix for block is the same for all blocks and it is equal to the average columns of the total design matrix.

Also we have the following remark, if we consider the orthogonal block design

$$\begin{aligned} \mathcal{B} = (\mathcal{B}_1^{(m)}, \ldots, \mathcal{B}_b^{(m)}) &:= \begin{pmatrix} \mathbf{x}_{11} & \cdots & \mathbf{x}_{ij} & \cdots & \mathbf{x}_{bS_b} \\ m_{11} & \cdots & m_{ij} & \cdots & m_{bS_b} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_l & \cdots & \mathbf{x}_{S^*} \\ m_1 & \cdots & m_l & \cdots & m_{S^*} \end{pmatrix} = \mathcal{E} \end{aligned}$$

$$\text{with} \quad l = \sum_{i=1}^{b} S_{i-1} + j; \quad j = 1, \ldots, S_i; \quad S_0 = 0;$$

then, the information matrix for $\boldsymbol{\beta}$ corresponding to $\mathcal{B}$ given by

$$\begin{aligned}
\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{g}}(\mathcal{B})) &= \widehat{\mathbf{F}}^\top \mathbf{W}\widehat{\mathbf{F}} - \frac{1}{b}\sum_{i=1}^{b} \widehat{\mathbf{F}}_i^\top \mathbf{W}_i \mathbf{1}_{S_i}\mathbf{1}_{S_i}^\top \mathbf{W}_i \widehat{\mathbf{F}}_i \\
&= \widehat{\mathbf{F}}^\top \mathbf{W}\widehat{\mathbf{F}} - \frac{1}{b}\sum_{i=1}^{b} \widehat{\mathbf{F}}^\top \mathbf{W}\mathbf{1}_{S^*}\mathbf{1}_{S^*}^\top \mathbf{W}\widehat{\mathbf{F}} \\
&= \widehat{\mathbf{F}}^\top \mathbf{W}\widehat{\mathbf{F}} - \widehat{\mathbf{F}}^\top \mathbf{W}\mathbf{1}_{S^*}\mathbf{1}_{S^*}^\top \mathbf{W}\widehat{\mathbf{F}} \\
&= \mathcal{F}(\mathcal{E})^\top \mathbf{W}\mathcal{F}(\mathcal{E}) - \mathcal{F}^\top(\mathcal{E})\mathbf{W}\mathbf{1}_{S^*}\mathbf{1}_{S^*}^\top \mathbf{W}\mathcal{F}(\mathcal{E}) \\
&= \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{(1,\mathbf{f}^\top)^\top}(\mathcal{E})),
\end{aligned}$$

is equal to the information matrix for $\boldsymbol{\beta}$ corresponding to the population experimental design $\mathcal{E}$ for the linear model without block effects, where $\mathcal{F}(\mathcal{E}) = (\mathbf{f}(\mathbf{x}_1), \ldots, \mathbf{f}(\mathbf{x}_{S^*}))^\top$. In other words, a orthogonally blocked design, conserve the information on the regression parameter $\boldsymbol{\beta}$ of the linear model without the presence of fixed block effects. However we cannot ignore the fixed block effects in the model, because we are inflating in this case the variance of the experimental error $\boldsymbol{\epsilon}$, indeed we have that in general

$$\begin{aligned}
\mathrm{Var}(\boldsymbol{\epsilon}|\mathcal{B}) &= \mathrm{Var}\big(\mathbf{Y} - \mathbf{G}\hat{\boldsymbol{\theta}} - \mathbf{H}_f\hat{\boldsymbol{\gamma}}_f\big) \\
&= \mathrm{Var}\big(\mathbf{Y} - \mathbf{G}\hat{\boldsymbol{\theta}}\big) + \mathrm{Var}\big(\mathbf{H}_f\hat{\boldsymbol{\gamma}}_f\big) - 2\,\mathrm{Cov}\big(\mathbf{Y} - \mathbf{G}\hat{\boldsymbol{\theta}},\, \mathbf{H}_f\hat{\boldsymbol{\gamma}}_f\big) \\
&= \mathrm{Var}(\boldsymbol{\epsilon}|\mathcal{E}) + \mathrm{Var}\big(\mathbf{H}_f\hat{\boldsymbol{\gamma}}_f\big) - 2\,\mathrm{Cov}\big(\mathbf{H}_f\hat{\boldsymbol{\gamma}}_f + (\boldsymbol{\epsilon}|\mathcal{B}), \mathbf{H}_f\hat{\boldsymbol{\gamma}}_f\big) \\
&= \mathrm{Var}(\boldsymbol{\epsilon}|\mathcal{E}) + \mathrm{Var}\big(\mathbf{H}_f\hat{\boldsymbol{\gamma}}_f\big) - 2\,\mathrm{Var}\big(\mathbf{H}_f\hat{\boldsymbol{\gamma}}_f\big) \\
&= \mathrm{Var}(\boldsymbol{\epsilon}|\mathcal{E}) - \mathbf{H}_f\mathrm{Var}\big(\hat{\boldsymbol{\gamma}}_f\big)\mathbf{H}_f^\top
\end{aligned}$$

Hence, unlike of an only block design, in an orthogonally blocked design, no information on $\hat{\boldsymbol{\beta}}$, in the linear model without blocks is lost, and additionally they will help us in a better interpretation of the results.
Also it can announce (Goos and Vandebroek (2001)) the following

Lemma 3.2 *When the effects of a blocking variable are assumed fixed, an exact design $\mathcal{E}$ which is $D_{\boldsymbol{\beta}}$-optimal and is orthogonally blocked also is $D_{(\boldsymbol{\mu}^\top\ \boldsymbol{\beta}^\top)^\top}$-optimal for a given block structure $\mathcal{B}$.*

Chapter 4

Optimal Designs in the Presence of Random Block Effects

In this chapter, we focus on the construction of D-optimal designs of blocked experiments when their block effects are regarded as random, that is the blocks can be identifiable as a random sample from a larger population of blocks, hence is possible to make predictions about future observations, it assumes also the existence of correlation between the responses measured within any given block in order to get precise factor effects estimate (see e.g. Kunert(1994), Cheng (1995), Atkins and Cheng(1999) and Schmelter and Schwabe (2008)).
Since this model contains fixed regression effects and random block effects we have to use the more computationally analytical technique of linear mixed models as generalized least squares (GLS) estimation of the factor effects, however in our case the result already known for the D-criterion and the A-criterion in chapter 1, can be applied (analogously) to the mixed models information matrix for known covariance-variance matrix. Detailed descriptions and proves of these generalizations can be found in the paper, on the optimality of single-group Designs in linear mixed models, by Thomas

Schmelter(2007).

4.1 The random block effects model

If in a block experiment the block effects are assumed as realizations of a random qualitative factor , then we have the random block effects, this situation in the vector notation (3.3) has the form:

$$\begin{aligned} \mathbf{Y}_i(i, \mathcal{B}_i^{(m)}) &= \big(Y_{i1}(\mathbf{x}_{i1}), \ldots, Y_{im}(\mathbf{x}_{im})\big)^\top \\ \mathbf{Y}_i &= \big(\mathbf{1}_m, \mathbf{F}_i\big)\big(\beta_0, \boldsymbol{\beta}^\top\big)^\top + \mathbf{1}_m\gamma_i + \boldsymbol{\epsilon}_i \\ &= \mathbf{G}_i\boldsymbol{\theta} + \mathbf{1}_m\gamma_i + \boldsymbol{\epsilon}_i \end{aligned}$$

we assume that the random parameter of different block, the random effects $\gamma_i, \quad i = 1, \ldots, b$ are normally distributed, with zero means and variances σ_γ^2, independently of each other $\big(\text{Cov}(\gamma_i, \gamma_{i'}) = 0$ for $i \neq i'\big)$ and of the vector of observational errors $\big(\text{Cov}(\gamma_i, \boldsymbol{\epsilon}_{i'}) = \mathbf{0}_m\big)$. The observational errors are assumed to be homoscedastic and independent $\big(\text{Cov}(\boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_{i'}) = \delta_{ii'}\sigma^2\mathbf{I}_m$ here, $\delta_{ii'}$ is the Kronecker delta$\big)$ and they also have each one the normal distribution, such that $\boldsymbol{\epsilon}_i \sim N\big(\mathbf{0}, \sigma^2\mathbf{I}_m\big)$.
It follows that

$$\text{E}\big(\mathbf{Y}_i\big) = \big(\mathbf{1}_m, \mathbf{F}_i\big)\big(\beta_0, \boldsymbol{\beta}^\top\big)^\top = \mathbf{G}_i\boldsymbol{\theta} \tag{4.1}$$

$$\begin{aligned} \text{Cov}\big(\mathbf{Y}_i\big) &= \text{Cov}\big(\mathbf{G}_i\boldsymbol{\theta} + \mathbf{1}_m\gamma_i + \boldsymbol{\epsilon}_i\big) = \text{Cov}\big(\mathbf{1}_m\gamma_i + \boldsymbol{\epsilon}_i\big) \\ &= \text{Cov}\big(\mathbf{1}_m\gamma_i\big) + \text{Cov}\big(\boldsymbol{\epsilon}_i\big) = \mathbf{1}_m\text{Cov}\big(\gamma_i\big)\mathbf{1}_m^\top + \text{Cov}\big(\boldsymbol{\epsilon}_i\big) \\ &= \sigma_\gamma^2\mathbf{1}_m\mathbf{1}_m^\top + \sigma^2\,\mathbf{I}_m = \sigma^2\big(\mathbf{I}_m + d\mathbf{1}_m\mathbf{1}_m^\top\big) \\ &= \sigma^2\,\mathbf{V}, \end{aligned} \tag{4.2}$$

where $d = \sigma_\gamma^2/\sigma^2$ is the variance ratio and $\mathbf{V} = \mathbf{I}_m + d\mathbf{1}_m\mathbf{1}_m^\top$ is a symmetric nonsingular matrix which is independent on the experimental setting.

4.2 Methods of Estimation

We first consider the case that the fixed effects design matrix for block $\mathbf{G}_i$ is of full column rank, the variance radio d is known, which implies the knowledge of $\mathbf{V}$, and $\boldsymbol{\theta}$ is estimated on the exact design for block $\mathcal{B}_i^{(m)}$, with support treatments $\mathbf{x}_{is}, \quad s = 1, \ldots, S_i$ by minimizing the generalized squared distance of the observed values from the predicted value of the correlated linear model:

$$\mathbf{L}_{GLS}\big(\boldsymbol{\theta}, \mathbf{Y}_i\big) = \big(\mathbf{Y}_i - \mathbf{G}_i\boldsymbol{\theta}\big)^\top \mathbf{V}^{-1}\big(\mathbf{Y}_i - \mathbf{G}_i\boldsymbol{\theta}\big) \longrightarrow \min_{\boldsymbol{\theta}\in\mathbb{R}^{q+1}} \tag{4.3}$$

The GLS estimator of $\boldsymbol{\theta}$ for block i is then

$$\hat{\boldsymbol{\theta}}_{GLS,i} = \big(\mathbf{G}_i^\top\mathbf{V}^{-1}\mathbf{G}_i\big)^{-1}\mathbf{G}_i^\top\mathbf{V}^{-1}\mathbf{Y}_i, \tag{4.4}$$

On the other hand,

$$\begin{aligned}
\mathbf{G}_i^\top\mathbf{V} &= \mathbf{G}_i^\top\big(\mathbf{I}_m + d\mathbf{1}_m\mathbf{1}_m^\top\big) \\
&= \big(\mathbf{1}_m,\, \mathbf{F}_i\big)^\top \left(\mathbf{I}_m + d\big(\mathbf{1}_m,\, \mathbf{F}_i\big)\begin{pmatrix} 1 & \mathbf{0}_q^\top \\ \mathbf{0}_q & \mathbf{0}_{q\times q}\end{pmatrix}\big(\mathbf{1}_m,\, \mathbf{F}_i\big)^\top\right) \\
&= \big(\mathbf{1}_m,\, \mathbf{F}_i\big)^\top + d\big(\mathbf{1}_m,\, \mathbf{F}_i\big)^\top\big(\mathbf{1}_m,\, \mathbf{F}_i\big)\begin{pmatrix} 1 & \mathbf{0}_q^\top \\ \mathbf{0}_q & \mathbf{0}_{q\times q}\end{pmatrix}\big(\mathbf{1}_m,\, \mathbf{F}_i\big)^\top \\
&= \left(\mathbf{I}_{q+1} + d\big(\mathbf{1}_m,\, \mathbf{F}_i\big)^\top\big(\mathbf{1}_m,\, \mathbf{F}_i\big)\begin{pmatrix} 1 & \mathbf{0}_q^\top \\ \mathbf{0}_q & \mathbf{0}_{q\times q}\end{pmatrix}\right)\big(\mathbf{1}_m,\, \mathbf{F}_i\big)^\top \\
&= \mathbf{U}_i\mathbf{G}_i^\top
\end{aligned}$$

In the given $(q+1)\times(q+1)$ regular matrix $\mathbf{U}_i$, the subscript i indicates the dependence on the experimental setting in the ith block.
With the above, it can be shown that the generalized least squares estimator for block $\hat{\boldsymbol{\theta}}_{GLS,i}$ coincides with the ordinary least squares estimator $\hat{\boldsymbol{\theta}}_{OLS,i}$; indeed

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{GLS,i} &= \left(\mathbf{G}_i^\top \mathbf{V}^{-1}\mathbf{G}_i\right)^{-1}\mathbf{G}_i^\top \mathbf{V}^{-1}\mathbf{Y}_i \\
&= \left(\mathbf{U}_i^{-1}\mathbf{U}_i\mathbf{G}_i^\top \mathbf{V}^{-1}\mathbf{G}_i\right)^{-1}\mathbf{U}_i^{-1}\mathbf{U}_i\mathbf{G}_i^\top \mathbf{V}^{-1}\mathbf{Y}_i \\
&= \left(\mathbf{U}_i^{-1}\mathbf{G}_i^\top \mathbf{V}\mathbf{V}^{-1}\mathbf{G}_i\right)^{-1}\mathbf{U}_i^{-1}\mathbf{G}_i^\top \mathbf{V}\mathbf{V}^{-1}\mathbf{Y}_i \\
&= \left(\mathbf{U}_i^{-1}\mathbf{G}_i^\top \mathbf{G}_i\right)^{-1}\mathbf{U}_i^{-1}\mathbf{G}_i^\top \mathbf{Y}_i \\
&= \left(\mathbf{G}_i^\top \mathbf{G}_i\right)^{-1}\mathbf{U}_i\mathbf{U}_i^{-1}\mathbf{G}_i^\top \mathbf{Y}_i \\
&= \left(\mathbf{G}_i^\top \mathbf{G}_i\right)^{-1}\mathbf{G}_i^\top \mathbf{Y}_i = \hat{\boldsymbol{\theta}}_{OLS,i}.
\end{aligned}$$

Hence the estimators for block $\hat{\boldsymbol{\theta}}_i$ do not require the knowledge of the variance ratio d, however by this fact and because they ignore the information that can be obtained from the other blocks in the block design, individually, the estimators for block $\hat{\boldsymbol{\theta}}_i$ are not the best linear unbiased estimators for $\boldsymbol{\theta}$.

If the number b of block that should be observed are summarized to

$$\mathbf{Y} = \left(\mathbf{Y}_1^\top, \dots, \mathbf{Y}_b^\top\right)^\top,$$

in matrix notation the model becomes

$$\mathbf{Y} = \mathbf{G}\boldsymbol{\theta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \tag{4.5}$$

where

$$\mathbf{G} = \left(\mathbf{G}_1^\top, \dots, \mathbf{G}_b^\top\right)^\top = \left(\mathbf{1}_b \otimes \mathbf{1}_m,\ \left(\mathbf{F}_1^\top, \dots, \mathbf{F}_b^\top\right)^\top\right) = \left(\mathbf{1}_{bm},\ \mathbf{F}\right),$$

$$
\begin{aligned}
\mathbf{Z} &= (\mathbf{I}_b \otimes \mathbf{1}_m) = \begin{pmatrix} \mathbf{1}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m \\ \mathbf{0}_m & \mathbf{1}_m & \cdots & \mathbf{0}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{1}_m \end{pmatrix}_{bm\times b} \\
\boldsymbol{\gamma} &= (\gamma_1, \ldots, \gamma_b)^\top \sim \mathcal{N}\big(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}_b\big),
\end{aligned}
$$

and

$$
\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1^\top, \ldots, \boldsymbol{\epsilon}_b^\top)^\top \sim \mathcal{N}\big(\mathbf{0}, \sigma^2 \mathbf{I}_{bm}\big).
$$

Here $\mathbf{G}$ denotes the design matrix of the explanatory variables for the fixed regression (or analysis of variance, or analysis of covariance) parameter $\boldsymbol{\theta}$ which is partitioned into the fixed effect design matrix component $\mathbf{G}_i$ corresponding to the ith level of the blocking variable. The design matrix $\mathbf{Z}$ contain the indicator variables for the random block effects. By independence of observations in different blocks of the experiments and properties of the Kronecker product of matrices, it follows then that the expected value and covariance matrix of $\mathbf{Y}$ are respectively

$$
\mathrm{E}\big(\mathbf{Y}\big) = \big(\mathbf{1}_{mb}, \mathbf{F}\big)\big(\beta_0, \boldsymbol{\beta}^\top\big)^\top, \tag{4.6}
$$

$$
\begin{aligned}
\mathrm{Cov}\big(\mathbf{Y}\big) &= \mathrm{Cov}\big(\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}\big) \\
&= \mathbf{Z}\,\mathrm{Cov}\big(\boldsymbol{\gamma}\big)\,\mathbf{Z}^\top + \mathrm{Cov}\big(\boldsymbol{\epsilon}\big) \\
&= \sigma_\gamma^2\,\mathbf{Z}\,\mathbf{Z}^\top + \sigma^2\,\mathbf{I}_{bm} \\
&= \sigma_\gamma^2\,\big(\mathbf{I}_b \otimes \mathbf{1}_m\big)\big(\mathbf{I}_b \otimes \mathbf{1}_m\big)^\top + \sigma^2\,\big(\mathbf{I}_b \otimes \mathbf{I}_m\big) \\
&= \sigma^2(\mathbf{I}_b \otimes \mathbf{V}).
\end{aligned} \tag{4.7}
$$

A known variance ratio d implies the knowledge of $\mathbf{I}_b \otimes \mathbf{V}$ and provided that the design matrix $\mathbf{G}$ is of full rank, then the population parameter, the vector of fixed effects $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}^\top)^\top$, in the random block effects model (4.5) can be estimated on the population basis by the generalized least squares estimator

$$\hat{\boldsymbol{\theta}} = \big(\mathbf{G}^\top(\mathbf{I}_b \otimes \mathbf{V})^{-1}\mathbf{G}\big)^{-1}\mathbf{G}^\top(\mathbf{I}_b \otimes \mathbf{V})^{-1}\mathbf{Y} \tag{4.8}$$

obtained by minimizing the generalized squared distance of the observed values from the predicted value of the correlated linear model:

$$\mathbf{L}_{GLS}\big(\boldsymbol{\theta}, \mathbf{Y}\big) = \big(\mathbf{Y} - \mathbf{G}\boldsymbol{\theta}\big)^\top(\mathbf{I}_b \otimes \mathbf{V})^{-1}\big(\mathbf{Y} - \mathbf{G}\boldsymbol{\theta}\big) \longrightarrow \min_{\boldsymbol{\theta}\in\mathbb{R}^{q+1}}$$

The GLS estimator $\hat{\boldsymbol{\theta}}$ is the best linear unbiased estimator for $\boldsymbol{\theta}$ and as remark the GLS-estimation is a distribution free methods.

Note that in the case where all design matrices component for blocks $\mathbf{G}_i$ have full rank $q + 1$, then individual models can be adjusted uniquely for block and the generalized least squared estimator $\hat{\boldsymbol{\theta}}_{GLS}$ is a matrix weighted mean of the individually estimated parameter $\hat{\boldsymbol{\theta}}_{GLS,i}$

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{GLS} &= \big(\mathbf{G}^\top(\mathbf{I}_b \otimes \mathbf{V})^{-1}\mathbf{G}\big)^{-1}\,\mathbf{G}^\top(\mathbf{I}_b \otimes \mathbf{V})^{-1}\mathbf{Y} \\
&= \left(\sum_{i=1}^{b}\mathbf{G}_i^\top\mathbf{V}^{-1}\mathbf{G}_i\right)^{-1}\sum_{i=1}^{b}\mathbf{G}_i^\top\mathbf{V}^{-1}\mathbf{Y}_i \\
&= \left(\sum_{i=1}^{b}\mathbf{G}_i^\top\mathbf{V}^{-1}\mathbf{G}_i\right)^{-1}\sum_{i=1}^{b}\mathbf{G}_i^\top\mathbf{V}^{-1}\mathbf{G}_i\,\hat{\boldsymbol{\theta}}_{GLS,i}
\end{aligned}$$

under the assumptions given above of d known, $\mathbf{G}$ of full rank together with the normality distributed random effects and the independence of the observations of different blocks, the Maximum Likelihood Estimation can be applied for fitting the linear model to the data, that is given the observation vector $\mathbf{Y}$, the likelihood function

$$\mathbf{L}_{\mathbf{Y}} : \mathbb{R}^{q+1} \longrightarrow [0, \infty),$$

such that

$$\mathbf{L}_{\mathbf{Y}}(\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{bm/2}(\det \mathbf{V})^{b/2}} \exp\left[\frac{-1}{2\sigma^2}\big(\mathbf{Y} - \mathbf{G}\boldsymbol{\theta}\big)^\top (\mathbf{I}_b \otimes \mathbf{V})^{-1}\big(\mathbf{Y} - \mathbf{G}\boldsymbol{\theta}\big)\right]$$

is maximized with respect to the parameter $\boldsymbol{\theta}$, thus a maximum likelihood estimator for $\boldsymbol{\theta}$ is a vector $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{q+1}$ with

$$\mathbf{L}_{\mathbf{Y}}(\hat{\boldsymbol{\theta}}) = \max_{\boldsymbol{\theta} \in \mathbb{R}^{q+1}} \mathbf{L}_{\mathbf{Y}}(\boldsymbol{\theta})$$

the result is that the maximum likelihood estimator and the generalized least estimator coincide, therefore

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{ML} &= \big(\mathbf{G}^\top (\mathbf{I}_b \otimes \mathbf{V})^{-1}\mathbf{G}\big)^{-1}\mathbf{G}^\top (\mathbf{I}_b \otimes \mathbf{V})^{-1}\mathbf{Y} \\
&= \left(\sum_{i=1}^{b} \mathbf{G}_i^\top \mathbf{V}^{-1}\mathbf{G}_i\right)^{-1} \sum_{i=1}^{b} \mathbf{G}_i^\top \mathbf{V}^{-1}\mathbf{Y}_i \\
&= \left(\sum_{i=1}^{b} (\mathbf{1}_m,\, \mathbf{F}_i)^\top \mathbf{V}^{-1}(\mathbf{1}_m,\, \mathbf{F}_i)\right)^{-1} \sum_{i=1}^{b} (\mathbf{1}_m,\, \mathbf{F}_i)^\top \mathbf{V}^{-1}\mathbf{Y}_i, \quad (4.9)
\end{aligned}$$

with variance-covariance matrix given by

$$\begin{aligned}
\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{ML}) &= \sigma^2 \left(\mathbf{G}^\top (\mathbf{I}_b \otimes \mathbf{V})^{-1}\mathbf{G}\right)^{-1} \\
&= \sigma^2 \left(\sum_{i=1}^{b} (\mathbf{1}_m,\, \mathbf{F}_i)^\top \mathbf{V}^{-1}(\mathbf{1}_m,\, \mathbf{F}_i)\right)^{-1} \quad (4.10)
\end{aligned}$$

4.3 Information Matrices and Optimal Designs

The information matrix of the unknown fixed model parameter $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}^\top)^\top$ or moment matrix corresponding to the block design

$\mathcal{B} = (\mathcal{B}_1^{(m)}, \ldots, \mathcal{B}_b^{(m)})$ will depend on the variance ratio d through the matrix $\mathbf{V}$, hence we write

$$\begin{aligned}\mathcal{M}(\mathcal{B}, d) &= \mathbf{G}^\top (\mathbf{I}_b \otimes \mathbf{V})^{-1} \mathbf{G} \\ &= \sum_{i=1}^{b} \mathbf{G}_i^\top \mathbf{V}^{-1} \mathbf{G}_i \\ &= \sum_{i=1}^{b} (\mathbf{1}_m,\, \mathbf{F}_i)^\top \mathbf{V}^{-1} (\mathbf{1}_m,\ \mathbf{F}_i) \end{aligned} \tag{4.11}$$

We use the fact that

$$\mathbf{V}^{-1} = \mathbf{I}_m - \frac{d}{1+md} \mathbf{1}_m \mathbf{1}_m^\top, \tag{4.12}$$

then we have

$$\mathcal{M}(\mathcal{B}, d) = \sum_{i=1}^{b} \mathbf{G}_i^\top \mathbf{G}_i - \frac{d}{1+md} \sum_{i=1}^{b} \mathbf{G}_i^\top \mathbf{1}_m \mathbf{1}_m^\top \mathbf{G}_i, \tag{4.13}$$

or partitioning the information matrix according to β_0 and $\boldsymbol{\beta}$ yields,

$$\begin{aligned}\mathcal{M}(\mathcal{B}, d) &= \sum_{i=1}^{b} (\mathbf{1}_m,\, \mathbf{F}_i)^\top \left(\mathbf{I}_m - \frac{d}{1+md} \mathbf{1}_m \mathbf{1}_m^\top \right) (\mathbf{1}_m,\, \mathbf{F}_i) \\ &= \frac{1}{1+md} \sum_{i=1}^{b} \left(\begin{array}{c|c} m & \mathbf{1}_m^\top \mathbf{F}_i \\ \hline \mathbf{F}_i^\top \mathbf{1}_m & (1+md)\mathbf{F}_i^\top \mathbf{F}_i - d\, \mathbf{F}_i^\top \mathbf{1}_m \mathbf{1}_m^\top \mathbf{F}_i \end{array} \right)\end{aligned}$$

But the problem of determining D-optimum designs, selecting b exact design for blocks of size m, $(\mathcal{B}_1^{(m)}, \ldots, \mathcal{B}_b^{(m)})$ not necessarily all distinct, for estimating the fixed, but unknown $(q+1)$vector parameter $\boldsymbol{\theta}$ is difficult, because the evaluation of the information matrix is expensive still using numerical methods.

Due to for all $i = 1, \dots, b$

$$\mathrm{E}\big(\mathbf{Y}_i\big) = \big(\mathbf{1}_m, \mathbf{F}_i\big)\big(\beta_0, \boldsymbol{\beta}^\top\big)^\top \tag{4.14}$$

$$\mathrm{Cov}\big(\mathbf{Y}_i\big) = \sigma^2 \mathbf{V} \tag{4.15}$$

we can do analysis of estimations of the parameter with only the distinct designs for blocks among the b ones, for example Atkins and Cheng(1999) sketched an approach based on the approximate block designs on the m-dimensional experimental design $\mathcal{X}^m$. In general an approximate block design, similar to the approximate designs considerate in the chapter 2, can be written as

$$\boldsymbol{\xi} = \left\{ \begin{array}{ccc} \xi_1^{(m)} & \dots & \xi_L^{(m)} \\ \\ g_1 & \dots & g_L \end{array} \right\}, \tag{4.16}$$

the exact design for blocks , $\xi_l^{(m)}$ $\quad l = 1, \dots, L$; that appear in the block design $\boldsymbol{\xi}$ are called the support of the design, thus $\mathrm{Supp}(\boldsymbol{\xi}) = \{\xi_1^{(m)}, \dots, \xi_L^{(m)}\}$, is a set of L different exact design for blocks among the b ones, additionally they will be observed under the blocks design $\boldsymbol{\xi}$ with weights or frequencies $g_1 \dots g_L$; respectively, so that $\mid \{i : \mathcal{B}_i^{(m)} = \xi_l^{(m)}\} \mid \approx b g_l$. Since $\boldsymbol{\xi}$ is a measure, the weights must satisfy the constraints, for all l, $\quad 0 \leq g_l \leq 1$, with $\displaystyle\sum_{l=1}^{L} g_l = 1$.

The information matrix of $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}^\top)^\top$ or moment matrix corresponding to the approximate block design $\boldsymbol{\xi}$ on $\mathcal{X}^m$ is then,

$$\mathcal{M}(\boldsymbol{\xi}, d) = \sum_{l=1}^{L} g_l (\mathbf{1}_m, \, \mathbf{F}_l)^\top \mathbf{V}^{-1} (\mathbf{1}_m, \, \mathbf{F}_l). \tag{4.17}$$

As an alternative way (see Schmelter (2007)), the information contributed by the observations in the block l can be represented by the moment matrix for the corresponding exact design for block

$$
\begin{aligned}
\mathcal{M}(\xi_l^{(m)}, d) &= (\mathbf{1}_m,\, \mathbf{F}_l)^\top \mathbf{V}^{-1} (\mathbf{1}_m,\, \mathbf{F}_l) \\
&= \frac{1}{1+md} \left(\begin{array}{c|c} m & \mathbf{1}_m^\top \mathbf{F}_l \\ \hline \mathbf{F}_l^\top \mathbf{1}_m & (1+md)\mathbf{F}_l^\top \mathbf{F}_l - d\, \mathbf{F}_l^\top \mathbf{1}_m \mathbf{1}_m^\top \mathbf{F}_l \end{array} \right)
\end{aligned}
$$

Now similar to the approximate designs for the ordinary linear model, in a generalized setup, we allow proportions m_{ls}/m non rational by means of the following definition.

Definition 4.1 *An approximate design for block l of size m, denoted $\xi_l^{(m)}$, is a set of distinct experimental setting* $\mathbf{x}_{l1}, \ldots, \mathbf{x}_{lS_l} \in \mathcal{X}$ *together with the corresponding proportions $m_{ls}/m := w_{lt} \in [0, 1]$, satisfying* $\sum_{s=1}^{S_l} m_{ls} = m$*. We write*

$$
\xi_l^{(m)} = \left\{ \begin{array}{cccc} \mathbf{x}_{l1} & \mathbf{x}_{l2} & \ldots & \mathbf{x}_{lS_l} \\ w_{l1} & w_{l2} & \ldots & w_{lS_l} \end{array} \right\}. \tag{4.18}
$$

Applying the above definition we obtain that

$$
\mathbf{F}_l^\top \mathbf{F}_l = m \sum_{s=1}^{S_l} w_{ls} \mathbf{f}(\mathbf{x}_{ls}) \mathbf{f}(\mathbf{x}_{ls})^\top = m \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \widehat{\mathbf{F}}_l \tag{4.19}
$$

and

$$
\mathbf{1}_m^\top \mathbf{F}_l = m \sum_{s=1}^{S_l} \mathbf{f}(\mathbf{x}_{ls}) w_{(ls)} = m \mathbf{1}_{S_l}^\top \mathbf{W}_l \widehat{\mathbf{F}}_l \tag{4.20}
$$

where $\mathbf{W}_l$ is the diagonal matrix with the proportions $m_{lt}/m = w_{lt}$, $t = 1, \ldots, ls_l$ as diagonal entries and the design matrix $\widehat{\mathbf{F}}_l$ is evaluated at the support experimental settings $\mathbf{x}_{l1}, \mathbf{x}_{l2} \ldots, \mathbf{x}_{ls_l}$.
Thus the moment matrix of an approximate design for a single block, $\xi_l^{(m)}$, can be written as

$\mathcal{M}(\xi_l^{(m)}, d)$

$$= \frac{1}{1+md}\left(\begin{array}{c|c} 1 & \mathbf{1}_{s_l}^\top \mathbf{W}_l \widehat{\mathbf{F}}_l \\ \hline \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \mathbf{1}_{s_l} & (1+md)\widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \widehat{\mathbf{F}}_l - md\, \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \mathbf{1}_{s_l} \mathbf{1}_{s_l}^\top \mathbf{W}_l \widehat{\mathbf{F}}_l \end{array}\right).$$

Therefore, the block design $\boldsymbol{\xi}$ can be identified as a probability measure on the set of approximate designs for block of size m such that $\boldsymbol{\xi}$ has support equal to the set $\{\xi_1^{(m)}, \ldots, \xi_L^{(m)}\}$,

and the information matrix of $\boldsymbol{\theta}$ corresponding to the design $\boldsymbol{\xi}$ can be decomposed as a weighted sum of information matrices for single blocks, Thus we have
$\mathcal{M}(\boldsymbol{\xi}, d) = \sum\limits_{l=1}^{L} g_l \mathcal{M}(\xi_l^{(m)}, d)$

$$= \frac{1}{1+md}\sum_{l=1}^{L} g_l \left(\begin{array}{c|c} 1 & \mathbf{1}_{s_l}^\top \mathbf{W}_l \widehat{\mathbf{F}}_l \\ \hline \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \mathbf{1}_{s_l} & (1+md)\widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \widehat{\mathbf{F}}_l - md\, \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \mathbf{1}_{s_l} \mathbf{1}_{s_l}^\top \mathbf{W}_l \widehat{\mathbf{F}}_l \end{array}\right)$$

$g_l \in [0,1], \quad \text{with } \sum_{l=1}^{L} g_l = 1.$

The information matrix results in a considerable compact notation upon introducing the vector

$$\begin{aligned} \boldsymbol{\varpi}_l &:= \sqrt{\mathbf{W}_l}\, \mathbf{1}_{S_l} \\ &= \left(\sqrt{w_{l1}}, \sqrt{w_{l2}}, \ldots, \sqrt{w_{lS_l}}\right)^\top, \end{aligned} \tag{4.21}$$

and the matrix

$$\begin{aligned} \widetilde{\mathbf{F}}_l &:= \sqrt{\mathbf{W}_l}\, \widehat{\mathbf{F}}_l \\ &= \left(\sqrt{w_{l1}}\mathbf{f}(\mathbf{x}_{l1}), \ldots, \sqrt{w_{(lS_l)}}\mathbf{f}(\mathbf{x}_{lS_l})\right)^\top \end{aligned} \tag{4.22}$$

The information matrix, thus can be rewritten as

$\mathcal{M}(\boldsymbol{\xi}, d)$

$$= \frac{1}{1+md}\sum_{l=1}^{L} g_l \left(\begin{array}{c|c} 1 & \boldsymbol{\varpi}_l^\top \widetilde{\mathbf{F}}_l \\ \hline \widetilde{\mathbf{F}}_l^\top \boldsymbol{\varpi}_l & (1+md)\widetilde{\mathbf{F}}_l^\top \widetilde{\mathbf{F}}_l - md\, \widetilde{\mathbf{F}}_l^\top \boldsymbol{\varpi}_l \boldsymbol{\varpi}_l^\top \widetilde{\mathbf{F}}_l \end{array} \right). \quad (4.23)$$

If there is special interest for the parameter of the regression fixed effects $\boldsymbol{\beta}$, then by the properties of partitioned matrices the corresponding partial information matrix is

$$\begin{aligned} \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, d)) \;&=\; \sum_{l=1}^{L} g_l\, \widetilde{\mathbf{F}}_l^\top \widetilde{\mathbf{F}}_l - \frac{md}{1+md} \sum_{l=1}^{L} g_l\, \widetilde{\mathbf{F}}_l^\top \boldsymbol{\varpi}_l \boldsymbol{\varpi}_l^\top \widetilde{\mathbf{F}}_l \\ &\quad - \frac{1}{1+md} \sum_{l=1}^{L} g_l\, \widetilde{\mathbf{F}}_l^\top \boldsymbol{\varpi}_l \sum_{l'=1}^{L} g_{l'}\, \boldsymbol{\varpi}_{l'}^\top \widetilde{\mathbf{F}_{l'}} \end{aligned} \quad (4.24)$$

and by the formula for evaluate the determinant of partitioned positive definite symmetric matrices, we obtain

$$\det \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, d)) = \frac{1}{1+md}\, \det(\mathcal{M}(\boldsymbol{\xi}, d)). \quad (4.25)$$

Hence, D- and $D_{\boldsymbol{\beta}}$-optimal designs go together also in random block effects as in fixed block effects. So we have shown

Lemma 4.2 *When the effects of a blocking variable are assumed random, an approximate block design* $\boldsymbol{\xi}$ *on the set of approximate designs for block is* $D_{(\beta_0, \boldsymbol{\beta}^\top)^\top}$*-optimal if only if it is* $D_{\boldsymbol{\beta}}$*-optimal.*

4.4 Limiting Models

We consider now in the random block effects model (4.5) the limiting of the partial information matrix of $\boldsymbol{\beta}$ corresponding to the block design $\boldsymbol{\xi}$,

$\mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, d))$, for $d \longrightarrow 0$ and $d \longrightarrow \infty$, respectively.
For $d \longrightarrow 0$ we obtain that

$$\mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, d)) \longrightarrow \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \widetilde{\mathbf{F}}_l - \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \boldsymbol{\varpi}_l \sum_{l'=1}^{L} g_{l'} \boldsymbol{\varpi}_{l'}^{\top} \widetilde{\mathbf{F}_{l'}} =: \mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, 0)).$$

For $d \longrightarrow \infty$, it is obtained

$$\mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, d)) \longrightarrow \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \widetilde{\mathbf{F}}_l - \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \boldsymbol{\varpi}_l \boldsymbol{\varpi}_l^{\top} \widetilde{\mathbf{F}}_l =: \mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, \infty)).$$

With the above and (4.24), we obtain the following lemma, which compares $\mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, d))$ with a convex combination of the two partial limiting information matrices related to it

Lemma 4.3 *When the effects of a blocking variable are assumed random and $\boldsymbol{\xi}$ is a block design, then*

$$\mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, d)) = \frac{1}{1+md} \mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, 0)) + \frac{md}{1+md} \mathcal{I}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, \infty)).$$

In particular, given a block design $\boldsymbol{\xi}$ for the random block effects model (4.5) where for all $l = 1, \ldots, L$ we have the following two facts, first, the designs for block

$$\xi_l^{(m)} = \left\{ \begin{array}{cccc} \mathbf{x}_{l1} & \mathbf{x}_{l2} & \ldots & \mathbf{x}_{lS_l} \\ \\ w_{l1} & w_{l2} & \ldots & w_{lS_l} \end{array} \right\}. \tag{4.26}$$

are exact, that is the proportions $w_{ls} = m_{ls}/m$ are rational numbers for all $s = 1, \ldots, S_l$; and second, the bg_l are integer numbers, then the block design $\boldsymbol{\xi}$ for the random block effects model will be taken as an exact block design. We suppose that $\boldsymbol{\xi}$ is an exact block design for the model (4.5) where

$bg_l = b_l$, then

$$\begin{aligned}\sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \widetilde{\mathbf{F}}_l &= \sum_{l=1}^{L} \frac{b_l}{b} \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \widehat{\mathbf{F}}_l \\ &= \sum_{l=1}^{L} \frac{b_l}{bm} \mathbf{F}_l^{\top} \mathbf{F}_l = \frac{1}{bm} \sum_{i=1}^{b} \mathbf{F}_i^{\top} \mathbf{F}_i\end{aligned}$$

and

$$\begin{aligned}\sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \boldsymbol{\varpi}_l &= \sum_{l=1}^{L} \frac{b_l}{b} \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \mathbf{1}_{s_l} \\ &= \sum_{l=1}^{L} \frac{b_l}{bm} \mathbf{F}_l^{\top} \mathbf{1}_m = \frac{1}{bm} \sum_{i=1}^{b} \mathbf{F}_i^{\top} \mathbf{1}_m\end{aligned}$$

therefore

$$\begin{aligned}\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi}, 0)) &= \frac{1}{bm} \left(\sum_{i=1}^{b} \mathbf{F}_i^{\top} \mathbf{F}_i - \frac{1}{bm} \sum_{i=1}^{b} \mathbf{F}_i^{\top} \mathbf{1}_m \sum_{i'=1}^{b} \mathbf{1}_m^{\top} \mathbf{F}_{i'} \right) \\ &= \frac{1}{bm} \left(\mathbf{F}^{\top} \mathbf{F} - \frac{1}{bm} \mathbf{F}^{\top} \mathbf{1}_{bm} \mathbf{1}_{bm}^{\top} \mathbf{F} \right)\end{aligned}$$

which is identical to the partial information matrix $\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M(\mathcal{E}))$ for the uncorrelated fixed effects model without block effects

$$\mathbf{Y} = (\mathbf{1}_b \otimes \mathbf{1}_m, \mathbf{F})(\beta_0, \boldsymbol{\beta}^{\top})^{\top} + \epsilon. \tag{4.27}$$

Also we have that

$$\begin{aligned}\sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \boldsymbol{\varpi}_l \boldsymbol{\varpi}_l^{\top} \widetilde{\mathbf{F}}_l &= \sum_{l=1}^{L} \frac{b_l}{b} \widehat{\mathbf{F}}_l^{\top} \mathbf{W}_l \mathbf{1}_{s_l} \mathbf{1}_{s_l}^{\top} \mathbf{W}_l \widehat{\mathbf{F}}_l \\ &= \sum_{l=1}^{L} \frac{b_l}{bm^2} \mathbf{F}_l^{\top} \mathbf{1}_m \mathbf{1}_m^{\top} \mathbf{F}_l \\ &= \frac{1}{bm^2} \sum_{i=1}^{b} \mathbf{F}_i^{\top} \mathbf{1}_m \mathbf{1}_m^{\top} \mathbf{F}_i\end{aligned}$$

therefore

$$\begin{aligned}
\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi},\infty)) &= \frac{1}{bm}\left(\sum_{i=1}^{b}\mathbf{F}_i^\top\mathbf{F}_i - \frac{1}{m}\sum_{i=1}^{b}\mathbf{F}_i{}^\top\mathbf{1}_m\mathbf{1}_m^\top\mathbf{F}_i\right)\\
&= \frac{1}{bm}\left(\mathbf{F}^\top\mathbf{F} - \frac{1}{m}\mathbf{F}^\top(\mathbf{I}_b\otimes\mathbf{1}_m)(\mathbf{I}_b\otimes\mathbf{1}_m)^\top\mathbf{F}\right)
\end{aligned}$$

this expression is the same of the partial information matrix $\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(M_{\mathbf{g}}(\mathcal{B}))$ for the fixed block effects model (3.17), which in matrix notation can be written as

$$\mathbf{Y} = \mathbf{A}\boldsymbol{\mu} + \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{A} = \mathbf{I}_b \otimes \mathbf{1}_m$ and $\boldsymbol{\mu}_f = \mathbf{1}_b\beta_0 + \left(\boldsymbol{\gamma}_f^\top,\ -\mathbf{1}_{(b-1)}^\top\boldsymbol{\gamma}_f\right)^\top$.

Hence this fact show that, when $d \longrightarrow \infty$, the $D_{\boldsymbol{\beta}}$-optimal design for the random block effects model is identical with the $D_{\boldsymbol{\beta}}$-optimal design for the fixed block effects model, and by Lemmas (3.1) and (4.2) this is true also for D-optimal designs.

4.5 Optimal and Orthogonal Block Design

In this section, we will see how in a random blocked experiment the estimation of the fixed parameter for the effects of the experimental setting and the interpretation of the result are simplified when the design are orthogonally blocked (Khuri(1992)).
We have an advantage if the random block effects model (4.5)

$$\mathbf{Y} = (\mathbf{1}_b \otimes \mathbf{1}_m, \mathbf{F})(\beta_0, \boldsymbol{\beta}^\top)^\top + (\mathbf{I}_b \otimes \mathbf{1}_m)\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

can be rewritten as

$$\begin{aligned}
\mathbf{Y} &= \mathbf{1}_{bm}\beta_0 + \mathbf{F}\boldsymbol{\beta} + (\mathbf{I}_b \otimes \mathbf{1}_m)\boldsymbol{\gamma} + \boldsymbol{\epsilon} \\
&= \mathbf{1}_{bm}\beta_0 + \frac{1}{bm}\mathbf{1}_{bm}\mathbf{1}_{bm}^\top\big(\mathbf{I}_b \otimes \mathbf{1}_m\big)\boldsymbol{\gamma} + \mathbf{F}\boldsymbol{\beta} \\
&\qquad + (\mathbf{I}_b \otimes \mathbf{1}_m)\boldsymbol{\gamma} - \frac{1}{bm}\mathbf{1}_{bm}\mathbf{1}_{bm}^\top\big(\mathbf{I}_b \otimes \mathbf{1}_m\big)\boldsymbol{\gamma} + \boldsymbol{\epsilon} \\
&= \mathbf{1}_{bm}\Big(\beta_0 + \frac{1}{bm}\mathbf{1}_{bm}^\top\big(\mathbf{I}_b \otimes \mathbf{1}_m\big)\boldsymbol{\gamma}\Big) + \mathbf{F}\boldsymbol{\beta} \\
&\qquad + \Big(\mathbf{I}_{bm} - \frac{1}{bm}\mathbf{1}_{bm}\mathbf{1}_{bm}^\top\Big)\big(\mathbf{I}_b \otimes \mathbf{1}_m\big)\boldsymbol{\gamma} + \boldsymbol{\epsilon} \\
&= \boldsymbol{a}\mathbf{1}_{bm} + \mathbf{F}\boldsymbol{\beta} + \widetilde{\mathbf{Z}}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \qquad (4.28)
\end{aligned}$$

where

$$\boldsymbol{a} = \beta_0 + \frac{1}{bm}\mathbf{1}_{bm}^\top\big(\mathbf{I}_b \otimes \mathbf{1}_m\big)\boldsymbol{\gamma}. \qquad (4.29)$$

and

$$\widetilde{\mathbf{Z}} = \Big(\mathbf{I}_{bm} - \frac{1}{bm}\mathbf{1}_{bm}\mathbf{1}_{bm}^\top\Big)\big(\mathbf{I}_b \otimes \mathbf{1}_m\big) \qquad (4.30)$$

It is important to observe that the elements of each column of $\widetilde{\mathbf{Z}}$ sum 0. Hence $\mathbf{1}_{bm}^\top\widetilde{\mathbf{Z}} = \mathbf{0}_{b\times 1}^\top$ and as by definition, a design is orthogonally blocked if the column of $(\mathbf{1}_{bm}\ \mathbf{F}) = \mathbf{G}$ are orthogonal to those of $\widetilde{\mathbf{Z}}$, that is if

$$\big(\mathbf{1}_{bm}\ \ \mathbf{F}\big)^\top\widetilde{\mathbf{Z}} = \mathbf{0}_{(q+1)\times b}$$

or equivalently

$$\begin{aligned}
\mathbf{F}^\top\widetilde{\mathbf{Z}} &= \mathbf{F}^\top\Big(\mathbf{I}_{bm} - \frac{1}{bm}\mathbf{1}_{bm}\mathbf{1}_{bm}^\top\Big)\big(\mathbf{I}_b \otimes \mathbf{1}_m\big) \qquad (4.31) \\
&= \mathbf{0}_{q\times b} \qquad (4.32)
\end{aligned}$$

where $\mathbf{0}_{q\times b}$ is a $q\times b$ matrix of zeros. From this condition we have

$$\begin{aligned}
\mathbf{F}^\top\big(\mathbf{I}_b\otimes\mathbf{1}_m\big) &= \frac{1}{bm}\mathbf{F}^\top\mathbf{1}_{bm}\mathbf{1}_{bm}^\top\big(\mathbf{I}_b\otimes\mathbf{1}_m\big)\\
\big(\mathbf{F}_1^\top,\ldots,\mathbf{F}_b^\top\big)\big(\mathbf{I}_b\otimes\mathbf{1}_m\big) &= \frac{1}{bm}\mathbf{F}^\top m\mathbf{1}_{bm}\mathbf{1}_b^\top\\
\left(\frac{1}{m}\mathbf{F}_1^\top\mathbf{1}_m,\ldots,\frac{1}{m}\mathbf{F}_b^\top\mathbf{1}_m\right) &= \big(\frac{1}{bm}\mathbf{F}^\top\mathbf{1}_{bm},\ldots,\frac{1}{bm}\mathbf{F}^\top\mathbf{1}_{bm}\big)
\end{aligned}$$

or equivalently this condition becomes

$$\frac{1}{m}\mathbf{F}_i^\top\mathbf{1}_m=\frac{1}{bm}\mathbf{F}^\top\mathbf{1}_{bm};\qquad \text{for all}\quad i=1,\ldots,b; \tag{4.33}$$

where $\mathbf{F}_i$ is the block fixed effect design matrix component of $\mathbf{F}$ corresponding to the ith level of the blocking variable.

In the case of an exact block design $\boldsymbol{\xi}$ for the model (4.28), then we can show that $\boldsymbol{\xi}$ becomes an orthogonally blocked design, when it satisfies the following condition

$$g_l\widetilde{\mathbf{F}_l}^\top\boldsymbol{\varpi}_l=\sum_{l'=1}^{L}g_{l'}\widetilde{\mathbf{F}_{l'}}^\top\boldsymbol{\varpi}_{l'};\quad \text{for all}\quad l=1,\ldots,L. \tag{4.34}$$

Indeed under this condition we have that the

$$\begin{aligned}
\frac{1}{bm}\mathbf{F}^\top\mathbf{1}_{bm} &= \frac{1}{bm}\sum_{i=1}^{b}\mathbf{F}_i{}^\top\mathbf{1}_m\\
&= \sum_{l=1}^{L}\frac{b_l}{bm}\mathbf{F}_l{}^\top\mathbf{1}_m=\sum_{l=1}^{L}\frac{b_l}{b}\widehat{\mathbf{F}_l}^\top\mathbf{W}_l\mathbf{1}_{s_l}\\
&= \sum_{l=1}^{L}g_l\widetilde{\mathbf{F}_l}^\top\boldsymbol{\varpi}_l
\end{aligned}$$

On the other hand, if $l'\in\{1,\ldots,L\}$, then

$$\begin{aligned}
\widetilde{\mathbf{F}_{l'}}^\top\boldsymbol{\varpi}_{l'} &= \widehat{\mathbf{F}_{l'}}^\top\mathbf{W}_{l'}\mathbf{1}_{s_l'}\\
&= \frac{1}{m}\mathbf{F}_{l'}{}^\top\mathbf{1}_m=\frac{1}{m}\mathbf{F}_i^\top\mathbf{1}_m\quad i\in\{1,\ldots,b\}
\end{aligned}$$

This conditions for orthogonal blocking defined by the equations (4.34) can be replaced in (4.24), thus we obtain

$$
\begin{aligned}
\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi},0)) &= \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \widetilde{\mathbf{F}}_l - \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \boldsymbol{\varpi}_l \sum_{l'=1}^{L} g_{l'} \boldsymbol{\varpi}_{l'}^{\top} \widetilde{\mathbf{F}}_{l'} \\
&= \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \widetilde{\mathbf{F}}_l - \sum_{l=1}^{L} g_l \widetilde{\mathbf{F}}_l^{\top} \boldsymbol{\varpi}_l \boldsymbol{\varpi}_l^{\top} \widetilde{\mathbf{F}}_l = \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi},\infty))
\end{aligned}
$$

therefore by lemma(4.3) $\boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi},d)) = \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}}(\mathcal{M}(\boldsymbol{\xi},0))$, thus the information matrix for an orthogonally blocked design is independent on the variance ratio d and with help of the Lemma (4.2) we have shown the following

Theorem 4.4 *An exact design $\boldsymbol{\xi}$ which is D-optimal for the uncorrelated linear model (4.27) and is orthogonally blocked, is a D-optimal block design for the random block effects model (4.28).*

4.6 Example

We consider the uncorrelated quadratic regression model in two explanatory variables without interactions

$$Y_l = \beta_0 + \beta_1 x_{1l} + \beta_2 x_{2l} + \beta_{11} x_{1l}^2 + \beta_{22} x_{2l}^2 + \epsilon_l; \qquad (x_{1l}, x_{2l}) \in [-1,1] \times [-1,1].$$

The design $\boldsymbol{\xi}$ which assigns equal weights $\frac{1}{9}$ to the four corner points $(\pm 1, \pm 1)$, to the four center points of the sides $(0, \pm 1)$; $(\pm 1, 0)$ to the center point $(0,0)$ of the square experimental region is D-optimum for this model. If the design $\boldsymbol{\xi}$ is orthogonally blocked as follow

$$\xi_1 = \left\{ \begin{array}{lll} (-1,0) & (0,1) & (1,-1) \\ 1/3 & 1/3 & 1/3 \end{array} \right\}, \quad \xi_2 = \left\{ \begin{array}{lll} (-1,-1) & (0,0) & (1,1) \\ 1/3 & 1/3 & 1/3 \end{array} \right\},$$

$$\xi_3 = \left\{ \begin{array}{lll} (-1,1) & (1,0) & (0,-1) \\ 1/3 & 1/3 & 1/3 \end{array} \right\}.$$

Then by theorem(4.4) $\boldsymbol{\xi}$ is a D-optimal block design for the adequate random block effects model with responses
$Y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_{11} x_{1ij}^2 + \beta_{22} x_{2ij}^2 + \gamma_i + \epsilon_{ij}$,
at the jth run on block i, ($i = 1, 2, 3;\ \ j = 1, 2, 3$) with wights given by $\boldsymbol{\xi}(i, (x_{1ij}, x_{2ij})) = \frac{1}{3}\xi_i(x_{1ij}, x_{2ij})$. Further this orthogonally blocked D-optimum design do not depend on the variance ratio d.

4.7 Single-Block Design in Random Blocks Effects Model

In this section we consider a particular orthogonally block designs $\boldsymbol{\xi}$, for the linear model in the presence of random block effects (4.5). If $\boldsymbol{\xi}$ is uniform across the blocks , then all b blocks are observed under the same conditions, that is, the experimental settings are the same for each block, $\xi_i^{(m)} = \xi_1^{(m)}$ for all $i = 1, \ldots, b$ thus the design can be written as

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1^{(m)} \\ 1 \end{pmatrix}. \tag{4.35}$$

These class of block designs are called *single − block designs* and we can note that the orthogonality condition (4.34), with $L = 1$, is satisfied.

Lemma 4.5 *If* $\boldsymbol{\xi}$ *is a single − block design, then* $\boldsymbol{\xi}$ *is an orthogonal block design.*

As an illustration we have the following theorem, which without considering the orthogonality, was proved by Atkins and Cheng (1999),

Theorem 4.6 (Theorem 2.1.,Atkins and Cheng (1999)) *Suppose a D-optimal approximate design ξ^* under the uncorrelated model (4.27) is supported on s points $\mathbf{x}_1, \ldots, \mathbf{x}_s \in \mathcal{X}$. If $m\xi^*(\mathbf{x}_t)$ is an integer for all $t = 1\ldots, s$, then the single block design $\boldsymbol{\xi} = \begin{pmatrix} \xi^* \\ 1 \end{pmatrix}$ is D-optimal under the random block effects model (4.5).*

The above theorem identifies situations where D-optimal population designs in the presence of random block effects do not depend on the variance ratio d and they can be obtained from optimal designs under the uncorrelated model, but that is because the design so defined are orthogonally blocked.

Chapter 5

Linear Regression Model in the Presence of a Partially Interacting Qualitative Factor and Random Block Effects

5.1 Introduction

Many regression experiments are designed involving both discrete and continuous factors of influence, for instance, a experimenter can apply a treatment to randomly selected experimental units, which belong to a finite number p of group or classes and then compares the group means for some quantitative response Y. Due to the p groups or classes can be regarded as levels k of a qualitative factor, such that the rearrangements of the ordering of the levels, do not affect the performance of the experiment, then a one-way model with group effects $\alpha_1, \dots, \alpha_p$ can be considered. Now we assume a adaptation to the model, where additionally to the quantitative variable Y there exist a quantitative control variable $\mathbf{x}$, which can be chosen independently of the levels of the qualitative factor, and Y is linearly related

to the variable $\mathbf{x}$. The above describes an intra-class regression model with the same model in each class, see e.g. Searle(1971,p. 355) and Kurotschka (1984). However we can see this experiment also as a particular example of a two-factor linear model.
The problem of constructing optimum experimental design for estimating the vector parameter of a two-factor linear model is more complicated than for single factor models, however the question under what conditions it can find optimum designs for two-factor models in terms of optimum designs for their single factor models has been developed, for example for multi-factor model with homoscedastic errors, Schwabe(1996) presents optimal designs for a great variety of cases, but the optimum design for multi-factor models involving one blocking random variable, that is in presence of random block effects has been less studied, because of this fact in this work we focus on the construction of product designs for a two-factor model given by a one-way layout partly interacting with covariates and additionally in presence of random blocks effects.
In connection with the above, in the present chapter we consider a linear model with three explicit kinds of factors and different interaction structures.
We start with the introduction of the marginal single models described by their corresponding marginal response functions.
The first and second marginal models are one-way layout models, where the qualitative factors are adjusted to a p- and b different levels, but the effects of each single level $k = 1, \ldots, p$ associated to the first marginal model are assumed fixed while the effects of each single level $i = 1, \ldots, b$ associated to the second marginal model are assumed random, thus

$$Y_{kj} = \alpha_k + \epsilon_{kj}, \quad j = 1, \ldots, n \tag{5.1}$$

and

$$Y_{ij} = \tau_i + \epsilon_{ij}, \quad j = 1, \ldots, m \tag{5.2}$$

where $\tau_i \overset{iid}{\sim} \mathcal{N}(0, \sigma_\tau^2)$ and $\text{Cov}(\tau_l, \epsilon_{ij}) = 0$.
The third marginal model is the usual regression model with intercept,

quantitative factors and mean response

$$\mu_3(\mathbf{x}) \quad = \quad \beta_0 + \sum_{l=1}^{q} f_l(\mathbf{x})\beta_l = \beta_0 + \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} \tag{5.3}$$

$\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^r$.

5.2 Regression Models with Qualitative Factor: Common Intercept

When an experiment also includes qualitative factors, the effects between the quantitative and qualitative factors should be taken into consideration. This section is devoted to the following particular situation. With the above first and third marginal models, we consider now a two-factor model, with lth observation at treatment k given by

$$Y_{kl} \quad = \quad \beta_0 + \mathbf{f}(\mathbf{x}_{kl})^\top \boldsymbol{\beta}^{(k)} + \epsilon_{kl}, \tag{5.4}$$

where $k = 1, \dots, p; \quad l = 1, \dots, N; \quad \mathbf{x}_{kl} \in \mathcal{X}; \quad \beta_0 \in \mathbb{R},$
$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_q(\mathbf{x}))^\top; \quad \boldsymbol{\beta}^{(k)} = (\beta_{1k}, \dots, \beta_{qk})^\top \in \mathbb{R}^q, \quad \epsilon_{kl} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2),$

hence the parameters $\boldsymbol{\beta}^{(k)}$ depend on k in some way to be specified, but $\mathbf{f}(\mathbf{x})$ is the same in each group level k and the intercept β_0 is common, that is the quantitative variable $\mathbf{x}$ can be choosen independently of k and the intercept β_0 is a invariant factor effects.
In this work, we focus on the construction of designs when the response depend on a random blocking variable, hence now suppose that the observations in each group k are blocked into b balanced random blocks of size m, thus the blocks are nested into each group level and the total of observations for group is $n = b * m$, then the jth observation at block i in group k at setting $\mathbf{x}_{kij}$ for the explanatory variable takes the form

$$Y_{kij} \quad = \quad \beta_0 + \gamma_{ki} + \mathbf{f}(\mathbf{x}_{kij})^\top \boldsymbol{\beta}^{(k)} + \epsilon_{kij} \tag{5.5}$$

$k = 1, \ldots, p; \quad i = 1, \ldots, b; \quad j = 1, \ldots, m, \quad \gamma_{ki}$ is the $i-th$ random block effect in treatment k and ϵ_{kij} is the random error, also it is assumed that $\forall\, i,\ i' \in \{1, \ldots, b\}, \quad E(\boldsymbol{\epsilon}_{ki}) = \mathbf{0}_m, \quad \mathrm{Cov}(\boldsymbol{\epsilon}_{ki}, \boldsymbol{\epsilon}_{ki'}) = \delta_{ii'}\sigma^2\mathbf{I}_m;$

$E(\gamma_{ki}) = 0, \quad \mathrm{Cov}(\gamma_{ki}, \gamma_{ki'}) = \delta_{ii'}\sigma_\gamma^2$ and $\mathrm{Cov}(\gamma_{ki}, \boldsymbol{\epsilon}_{ki'}) = \mathbf{0}_m.$

Here, $\delta_{ii'}$ is the Kronecker delta.

We denote each block of observations i in group k, similar to the random block effects model in chapter 4 as

$$\mathbf{Y}_{ki} \;=\; \mathbf{1}_m\beta_0 + \mathbf{F}_{ki}\boldsymbol{\beta}^{(k)} + \mathbf{1}_m\gamma_{ki} + \boldsymbol{\epsilon}_{ki}, \tag{5.6}$$

where $\mathcal{B}^{(m)}_{(ki)}$ is an exact design of size m associated to the block i in the group k, that is a set of distinct experimental setting $\mathbf{x}_{ki1}, \ldots, \mathbf{x}_{kiS_{ki}} \in \mathcal{X}$ together with the corresponding numbers of replications m_{kis} satisfying $\sum_{s=1}^{S_{ki}} m_{kis} = m$. We write

$$\mathcal{B}^{(m)}_{(ki)} = \left\{ \begin{array}{cccc} \mathbf{x}_{ki1} & \mathbf{x}_{ki2} & \ldots & \mathbf{x}_{kiS_{ki}} \\ \\ m_{ki1} & m_{ki2} & \ldots & m_{kiS_{ki}} \end{array} \right\}. \tag{5.7}$$

Also

$$\mathbf{F}_{ki} = diag(\mathbf{1}_{m_{ki1}}, \ldots, \mathbf{1}_{m_{kiS_{ki}}})(\mathbf{f}(\mathbf{x}_{ki1}), \ldots, \mathbf{f}(\mathbf{x}_{kiS_{ki}}))^\top$$

it follows that the expected value is

$$\begin{aligned} \mathrm{E}\left(\mathbf{Y}_{ki}\right) &= \mathbf{1}_m\beta_0 + \mathbf{F}_{ki}\boldsymbol{\beta}^{(k)} \\ &= \left(\mathbf{1}_m, \mathbf{a}_1(k)^\top \otimes \mathbf{F}_{ki}\right)\left(\beta_0, \boldsymbol{\beta}^{(1)\top}, \ldots, \boldsymbol{\beta}^{(p)\top}\right)^\top \end{aligned} \tag{5.8}$$

and the variance-covariance matrix is

$$\begin{aligned} \mathrm{Cov}\left(\mathbf{Y}_{ki}\right) &= \sigma^2\mathbf{I}_m + \sigma_\gamma^2\mathbf{1}_m\mathbf{1}_m^\top \\ &= \sigma^2\left(\mathbf{I}_m + d\mathbf{1}_m\mathbf{1}_m^\top\right) \\ &= \sigma^2\mathbf{V} \end{aligned} \tag{5.9}$$

where $\boldsymbol{\beta}^{(k)} \in \mathrm{R}^q, \quad d = \sigma_\gamma^2/\sigma^2$ is the variance ratio and $\mathbf{V} = \mathbf{I}_m + d\mathbf{1}_m\mathbf{1}_m^\top$ is a positive definite symmetric matrix.

5.3 Moment Matrix and Optimal Design

We can do analysis of estimations of the parameter $\boldsymbol{\beta}^{(k)}$, if we assume that the block designs are treated equally across the groups and with only the distinct designs for blocks among the b ones. In general an approximate block design for group k, can be written as

$$\boldsymbol{\xi}_{=} \left\{ \begin{array}{ccc} \xi_1^{(m)} & \dots & \xi_L^{(m)} \\ \\ g_1 & \dots & g_L \end{array} \right\} \in \Xi^{(b)}, \tag{5.10}$$

The exact design for blocks , $\xi_l^{(m)}$ $\quad l = 1, \dots, L$; that appear in the block design for group k, $\boldsymbol{\xi}$, are called the support of the design, additionally they will be observed with weights or frequencies $g_1 \dots g_L$; respectively, so that $| \{i : \mathcal{B}_i^{(m)} = \xi_l^{(m)}\} | \approx b g_l$. Since $\boldsymbol{\xi}$ is a measure, the weights must satisfy the constraints, for all l, $\quad 0 \leq g_l \leq 1$, with $\sum_{l=1}^{L} g_l = 1$.

Let $\mathcal{W}(\mathcal{X}_p)$ be the class of designs on the subsets of $\mathcal{X}_p$. We regard now a product design denoted by $\eta \times \boldsymbol{\xi} \in \mathcal{W}(\mathcal{X}_p) \times \Xi^{(b)}$,

then the corresponding moment matrix or information matrix for the full parameter in the random block effects model (5.6) is given by

$$\mathcal{M}(\eta \times \boldsymbol{\xi}) \quad = \quad \int_{\boldsymbol{\Xi}} (\mathbf{1}_m, \mathbf{a}_1^\top(k) \otimes \mathbf{F})^\top \mathbf{V}^{-1} (\mathbf{1}_m, \mathbf{a}_1^\top(k) \otimes \mathbf{F}) \, \mathrm{d}(\eta \times \boldsymbol{\xi})$$

Symmetry reasons (Schwabe,(1996, p.23-24)) implies that such product designs can be shown to be optimal if the marginal design η is considered as the uniform design $\eta^*(k) = \frac{1}{p}$ for all $k \in \mathcal{X}_p$, thus we have

$$\mathcal{M}(\eta^* \times \boldsymbol{\xi}) = \begin{pmatrix} \mathbf{M}_{11}(\boldsymbol{\xi}) & \frac{1}{p}\mathbf{1}_p \otimes \mathbf{M}_{12}(\boldsymbol{\xi}) \\ \\ \frac{1}{p}\mathbf{1}_p^\top \otimes \mathbf{M}_{12}(\boldsymbol{\xi})^\top & \frac{1}{p}\mathbf{I}_p \otimes \mathbf{M}_{22}(\boldsymbol{\xi}) \end{pmatrix}$$

where

$$\mathbf{M}_{11}(\boldsymbol{\xi}) = \int \mathbf{1}_m^\top \mathbf{V}^{-1} \mathbf{1}_m \,\mathrm{d}\boldsymbol{\xi}, \quad \mathbf{M}_{12}(\boldsymbol{\xi}) = \int \mathbf{1}_m^\top \mathbf{V}^{-1} \mathbf{F} \,\mathrm{d}\boldsymbol{\xi}$$

$$\text{and} \quad \mathbf{M}_{22}(\boldsymbol{\xi}) = \int \mathbf{F}^\top \mathbf{V}^{-1} \mathbf{F} \,\mathrm{d}\boldsymbol{\xi}$$

As a remark, we can see that the moment matrix of the design $\boldsymbol{\xi}$

$$\mathcal{M}(\boldsymbol{\xi}) \quad = \quad \begin{pmatrix} \mathbf{M}_{11}(\boldsymbol{\xi}) & \mathbf{M}_{12}(\boldsymbol{\xi}) \\ & \\ \mathbf{M}_{12}(\boldsymbol{\xi})^\top & \mathbf{M}_{22}(\boldsymbol{\xi}) \end{pmatrix}. \tag{5.11}$$

is associated with the mixed model for $p = 1$

$$\mathbf{Y}_i \quad = \quad \mathbf{1}_m \beta_0 + \mathbf{F}_i \boldsymbol{\beta} + \mathbf{1}_m \gamma_i + \boldsymbol{\epsilon}_i \tag{5.12}$$

then the determinant of $\mathcal{M}(\eta^* \times \boldsymbol{\xi})$ can be obtained by the rules for partitioned positive definite symmetric matrices and properties of the Kronecker product as follows,

$$\begin{aligned} &\det\left(\mathcal{M}(\eta^* \times \boldsymbol{\xi})\right) \\ &\quad \propto \Big[\det \mathbf{M}_{22}(\boldsymbol{\xi})\Big]^p \cdot \det\Big(\mathbf{M}_{11}(\boldsymbol{\xi}) - \mathbf{M}_{12}(\boldsymbol{\xi})^\top \mathbf{M}_{22}(\boldsymbol{\xi})^{-1} \mathbf{M}_{12}(\boldsymbol{\xi})\Big) \\ &\quad \propto \Big[\det \mathbf{M}_{22}(\boldsymbol{\xi})\Big]^{p-1} \det\Big(\mathcal{M}(\boldsymbol{\xi})\Big) \end{aligned}$$

Thus we have proved the following

Theorem 5.1 *Consider the two-factor mixed model (5.8)-(5.9). The pro duct design $\eta^* \times \boldsymbol{\xi}^*$ is D-optimal in the class $\mathcal{W}(\mathcal{X}_p) \times \boldsymbol{\Xi}^{(b)}$ if η^* is a uniform design on $\mathcal{X}_p$, and $\boldsymbol{\xi}^*$ maximizes*

$$\det(\mathcal{M}(\boldsymbol{\xi})) \, \det\left(\int \mathbf{F}^\top \mathbf{V}^{-1} \mathbf{F} \,\mathrm{d}\boldsymbol{\xi}\right)^{p-1}.$$

Chapter 6

Optimal Design for a Linear Model with Interacting Treatment Factor and Random Block Effects

The model of the previous chapter can be used, for example, in situations where the treatment effects $\boldsymbol{\beta}^{(k)}$ in p mutually exclusive groups or classes are to be compared, the value $\mathbf{x}$ of the quantitative factor corresponds to the doses of the treatment in the group chosen; but additionally the observations are taken in random blocks within treatment, thus the term $\beta_0 + \gamma_{ki}$ denotes the baseline of block i in group k, which are assumed to come from the same population for all treatment groups.

6.1 Optimal design

We can do analysis of D-optimal estimations of the parameters $\boldsymbol{\beta}^{(k)}$, according to Theorem 5.1, if we assume that the design for blocks $\mathcal{B}_i^{(m)}$ are uniform across the groups and further all b blocks within each group are observed under the same conditions, that is, the experimental settings are the same for each design for block, $\mathcal{B}_i^{(m)} = \mathcal{B}_1^{(m)}$ for all $k = 1, \dots, p$; $i = 1, \dots, b$; then the block designs for group can be written as

$$\boldsymbol{\xi}(i, \mathbf{x}) = \frac{1}{b}\mathcal{B}_1^{(m)}(\mathbf{x}) \tag{6.1}$$

or for short,

$$\boldsymbol{\xi} = \begin{pmatrix} \mathcal{B}_1^{(m)} \\ \\ 1 \end{pmatrix}. \tag{6.2}$$

These class of block designs for group are the *single − block designs*.

Now in this context, let δ_m be the standardized design for block $\mathcal{B}_1^{(m)}/m$, thus

$$\delta_m = \begin{pmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_S \\ w_0 & w_1 & \dots & w_S \end{pmatrix} \tag{6.3}$$

where m is the size of the block, $\mathbf{x}_s$ are the distinct settings and $w_s = m_{1\{s+1\}}/m$ the respectively weights, with $\sum_{s=0}^{S} w_s = 1$.
Now, it follows that $\boldsymbol{\xi} = \begin{pmatrix} \delta_m \\ 1 \end{pmatrix}$ and by (4.21),(4.22) and (4.23) we have

$$\boldsymbol{\varpi} = \sqrt{\mathbf{W}}\,\mathbf{1}_S; \quad \widetilde{\mathbf{F}} = \sqrt{\mathbf{W}}\,\widehat{\mathbf{F}}; \quad \widehat{\mathbf{F}} = (\mathbf{f}(\mathbf{x}_0), \dots, \mathbf{f}(\mathbf{x}_S))^\top$$

$$\mathbf{W} = \mathrm{diag}(w_0, w_1, \dots, w_S)$$

and
$\det(\mathcal{M}(\boldsymbol{\xi}))$

$$= \left(\frac{1}{1+md}\right)^{q+1} \det\left(\begin{array}{c|c} 1 & \boldsymbol{\varpi}^\top\widetilde{\mathbf{F}} \\ \hline \widetilde{\mathbf{F}}^\top\boldsymbol{\varpi} & (1+md)\widetilde{\mathbf{F}}^\top\widetilde{\mathbf{F}} - md\,\widetilde{\mathbf{F}}^\top\boldsymbol{\varpi}\boldsymbol{\varpi}^\top\widetilde{\mathbf{F}} \end{array}\right)$$

$$= \left(\frac{1}{1+md}\right)^{q+1} \det\left((1+md)\widetilde{\mathbf{F}}^\top\widetilde{\mathbf{F}} - md\,\widetilde{\mathbf{F}}^\top\boldsymbol{\varpi}\boldsymbol{\varpi}^\top\widetilde{\mathbf{F}} - \widetilde{\mathbf{F}}^\top\boldsymbol{\varpi}\boldsymbol{\varpi}^\top\widetilde{\mathbf{F}}\right)$$

$$= \frac{1}{1+md}\det\left(\widetilde{\mathbf{F}}^\top\widetilde{\mathbf{F}} - \widetilde{\mathbf{F}}^\top\boldsymbol{\varpi}\boldsymbol{\varpi}^\top\widetilde{\mathbf{F}}\right)$$

$$= \frac{1}{1+md}\det\left((\boldsymbol{\varpi},\,\widetilde{\mathbf{F}}_{\delta_m})^\top(\boldsymbol{\varpi},\,\widetilde{\mathbf{F}})\right)$$

$$= \frac{1}{1+md}\det\left((\mathbf{1}_{(S+1)},\widehat{\mathbf{F}})^\top\mathbf{W}(\mathbf{1}_{(S+1)},\widehat{\mathbf{F}})\right) \tag{6.4}$$

Assumption 1. We assume throughout the remainder that $\mathbf{x}_0 = \mathbf{0} \in \mathcal{X}$, $S = q$ and $\mathbf{f}(\mathbf{0}) = \mathbf{0_q}$,. The class of all single-block designs with these conditions is denote by $\boldsymbol{\Theta}_1$, also we have the following partial result for $\boldsymbol{\xi} \in \boldsymbol{\Theta}_1$:

$$\det(\mathcal{M}(\boldsymbol{\xi})) = \frac{1}{1+md}\left(\det(\mathbf{1}_{(q+1)},\widehat{\mathbf{F}})\right)^2 \det(\mathbf{W})$$

$$= \frac{1}{1+md}\left(\det\widehat{\mathbf{F}}_{\mathscr{G}}\right)^2\left(\prod_{s=0}^{q} w_s\right) \tag{6.5}$$

since we can write $\det(\mathbf{1}_{(q+1)},\widehat{\mathbf{F}}) = \det\widehat{\mathbf{F}}_{\mathscr{G}}$, where $\widehat{\mathbf{F}}_{\mathscr{G}} = (\mathbf{f}(\mathbf{x}_1),\ldots,\mathbf{f}(\mathbf{x}_q))^\top$.

If we are seeking out a D-optimal design for the model (5.6), then according to Theorem 5.1 we need to calculate also

$$\mathcal{M}_{22}(\boldsymbol{\xi}) = \int \mathbf{F}^\top\mathbf{V}^{-1}\mathbf{F}\,\mathrm{d}\boldsymbol{\xi}$$

$$= \widetilde{\mathbf{F}}^\top\widetilde{\mathbf{F}} - \frac{md}{1+md}\,\widetilde{\mathbf{F}}^\top\boldsymbol{\varpi}\boldsymbol{\varpi}^\top\widetilde{\mathbf{F}}$$

$$= \widehat{\mathbf{F}}^\top\mathbf{W}\widehat{\mathbf{F}} - \frac{md}{1+md}(\widehat{\mathbf{F}}^\top\mathbf{W}\mathbf{1}_{q+1})(\mathbf{1}_{q+1}^\top\mathbf{W}\widehat{\mathbf{F}}) \tag{6.6}$$

now due to the assumption 1 we can write

$$\begin{aligned}\widehat{\mathbf{F}}^\top\mathbf{W}\widehat{\mathbf{F}} &= (\mathbf{0}_q, \widehat{\mathbf{F}}_{\not 0}^\top)\mathrm{diag}(w_0, \widetilde{\mathbf{W}})(\mathbf{0}_q, \widehat{\mathbf{F}}_{\not 0}^\top)^\top \\ &= (\mathbf{0}_q, \widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}})(\mathbf{0}_q, \widehat{\mathbf{F}}_{0}^\top)^\top = \widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}}\widehat{\mathbf{F}}_{\not 0}, \qquad (6.7)\end{aligned}$$

and

$$\begin{aligned}\widehat{\mathbf{F}}^\top\mathbf{W}\mathbf{1}_{q+1} &= (\mathbf{0}_q, \widehat{\mathbf{F}}_{\not 0}^\top)\mathrm{diag}(w_0, \widetilde{\mathbf{W}})\,(1, \mathbf{1}_q^\top)^\top \\ &= (\mathbf{0}_q, \widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}})(1, \mathbf{1}_q^\top)^\top = \widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}}\mathbf{1}_q \qquad (6.8)\end{aligned}$$

where $\widetilde{\mathbf{W}} = \mathrm{diag}(w_1, \ldots, w_q)$, then we have that replacing (6.7) and (6.8) in (6.6) results in

$$\mathcal{M}_{22}(\boldsymbol{\xi}) = \left(\widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}}\widehat{\mathbf{F}}_{\not 0} - \frac{md}{1+md}(\widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}}\mathbf{1}_q)(\mathbf{1}_q^\top\widetilde{\mathbf{W}}\widehat{\mathbf{F}}_{\not 0})\right)$$

Now we use the fact that for any positive definite matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ and vector $\mathbf{b} \in \mathbb{R}^n$, it holds the following identity

$$\det(\mathbf{A} - \mathbf{b}\mathbf{b}^\top) = \det(\mathbf{A})[1 - \mathbf{b}^\top\mathbf{A}^{-1}\mathbf{b}], \qquad (6.9)$$

then the base of the second factor in the theorem 5.1 can be written in the form

$$\begin{aligned}&\det\left(\mathcal{M}_{22}(\boldsymbol{\xi})\right) \\ &= \det(\widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}}\widehat{\mathbf{F}}_{\not 0})\left(1 - \frac{md}{1+md}(\mathbf{1}_q^\top\widetilde{\mathbf{W}}\widehat{\mathbf{F}}_{\not 0})(\widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}}\widehat{\mathbf{F}}_{\not 0})^{-1}(\widehat{\mathbf{F}}_{\not 0}^\top\widetilde{\mathbf{W}}\mathbf{1}_q)\right) \\ &\propto (\det\widehat{\mathbf{F}}_{\not 0})^2\left(\prod_{s=1}^q w_s\right)(1 + md - md\mathbf{1}_q^\top\widetilde{\mathbf{W}}\mathbf{1}_q)\end{aligned}$$

Now the third factor in the former expression can be simplified as

$$
\begin{aligned}
(1 + md - md\mathbf{1}_q^\top \widetilde{\mathbf{W}} \mathbf{1}_q) &= \left(1 + md - md \sum_{s=1}^{q} w_s\right) \\
&= \left(1 + md \left(1 - \sum_{s=1}^{q} w_s\right)\right) = (1 + mdw_0)\,.
\end{aligned}
$$

that is

$$
\det\left(\mathcal{M}_{22}(\boldsymbol{\xi})\right) \;\propto\; (\det \widehat{\mathbf{F}}_{\mathcal{Q}})^2 \left(\prod_{s=1}^{q} w_s\right)(1 + mdw_0)\,. \tag{6.10}
$$

Thus as we want to apply the Theorem 5.1, we find that

$$
\begin{aligned}
&\det(\mathcal{M}(\boldsymbol{\xi}))\ \det\left(\mathcal{M}_{22}(\boldsymbol{\xi})\right)^{p-1} \\
&\quad\propto\; \left(\det \widehat{\mathbf{F}}_{\mathcal{Q}}\right)^2 \left(\prod_{s=0}^{q} w_s\right)\left(\left(\det \widehat{\mathbf{F}}_{\mathcal{Q}}\right)^2 \left(\prod_{s=1}^{q} w_s\right)(1 + mdw_0)\right)^{p-1} \\
&\quad\propto\; \left(\det \widehat{\mathbf{F}}_{\mathcal{Q}}\right)^{2p} w_0 \left(\prod_{s=1}^{q} w_s\right)^{p} (1 + mdw_0)^{p-1}.
\end{aligned} \tag{6.11}
$$

As a result, the maximization of this product of determinants can be se parated into two functions respectively, one of the unknown support points $\mathbf{x}_1, \ldots, \mathbf{x}_q$, and independent of md, and the other of the design weights $w_0, \ldots, w_q$, the second function do dependent of md.
We suppose that $(w_0, w_1, \ldots, w_q) \in [0,1]^{q+1}$, then continuous analytical methods allow finding that the maximum occurs when the proportions are

$$
w_0^* = \frac{1}{2(q+1)}\left(1 - \frac{pq+1}{pmd} + \sqrt{\left(1 - \frac{pq+1}{pmd}\right)^2 + \frac{4(q+1)}{pmd}}\right),
$$

which is increasing in d, and $w_s^* = \dfrac{1 - w_0^*}{q}, \quad s = 1, \ldots, q.$

Thus we have shown the following

Theorem 6.1 *The product design* $\eta^* \times \boldsymbol{\xi}^*$ *with* $\boldsymbol{\xi}^* = \begin{pmatrix} \delta^*_{m,d} \\ 1 \end{pmatrix}$ *in the regression model (5.8)-(5.9) with* $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ *and* $\delta^*_{m,d}$ *supported on* $\mathbf{0}$ *and* q *points,* $\mathbf{x}^*_1, \ldots, \mathbf{x}^*_q$*, which maximizes* $\det \widehat{\mathbf{F}}_{\mathcal{Q}}$*, it is D-optimal in* $\mathcal{W}(\mathcal{X}_p) \times \mathbf{\Theta}_1$ *if an only if* η^* *is the uniform design on* $\mathcal{X}_p$ *and*

$$w^*_0(d) = \frac{1}{2(q+1)}\left(1 - \frac{pq+1}{pmd} + \sqrt{\left(1 - \frac{pq+1}{pmd}\right)^2 + \frac{4(q+1)}{pmd}}\right)$$

and $\quad w^*_s(d) = \dfrac{1 - w^*_0(d)}{q}, \quad s = 1, \ldots, q,$ *respectively.*

.

6.2 Limiting Models

We regard the limiting cases for $d \longrightarrow 0$, and $d \longrightarrow \infty$.

For $d \longrightarrow 0$, we approach a two-factor model with common intercept and without block effects:

$$Y_{kj} = \beta_0 + \mathbf{f}(\mathbf{x}_j)^\top \boldsymbol{\beta}^{(k)} + \epsilon_{kj}$$

$k \in \mathcal{X}_p, \ \mathbf{x}_j \in \mathcal{X}$ and $\mathbf{f}(\mathbf{0}) = \mathbf{0}_q$.
Let $\mathcal{W}(\mathcal{B}_{\mathcal{X}})$ according to definition 2.3, a product design $\eta^* \times \delta^*$ with δ^* supported on $\mathbf{0}$ and the q points, $\mathbf{x}^*_1, \ldots, \mathbf{x}^*_q$, which maximize $\det \widehat{\mathbf{F}}_{\mathcal{Q}}$, it is D-optimal in $\mathcal{W}(\mathcal{X}_p) \times \mathcal{W}(\mathcal{B}_{\mathcal{X}})$ if and only if η^* is the uniform design on $\mathcal{X}_p$, $w^*_0 = \dfrac{1}{pq+1}$, and $w^*_s = \dfrac{p}{pq+1} \quad s = 1, \ldots, q$ respectively.

For $d \longrightarrow \infty$ we obtain $w^*_{s,\infty} = \dfrac{1}{q+1}$, for all $s = 0, 1, \ldots, q$ thus the design $\eta^* \times \boldsymbol{\xi}^*_\infty$ is identical to the $D_{\boldsymbol{\beta}}$-optimum design in $\mathcal{W}(\mathcal{X}_p) \times \mathbf{\Theta}_1$ for the two-factor model in the presence of *fixed* block effects: common intercept:

$$\mathbf{Y}_{ki} \;=\; \mathbf{1}_m \beta_0 + \mathbf{F}_i \boldsymbol{\beta}^{(k)} + \mathbf{1}_m \gamma_{ki} + \boldsymbol{\epsilon}_{ki}, \tag{6.12}$$

$\epsilon_{kij} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$, the fixed block effects satisfy the identifiability condition $\sum_{k=1}^{p} \sum_{i=1}^{b} \gamma_{ki} = 0$

6.3 Example

As an example, we will construct a product design which is optimum for a one-way layout with $k = 1, \ldots, p$ levels combined with polynomial regression in one explanatory variable and measurements at baseline in the presence of random block effects: Common intercept.
Thus we have the following two-factor mixed model (5.6)

$$\mathbf{Y}_{ki} = \mathbf{1}_m \beta_0 + \mathbf{F}_i \boldsymbol{\beta}^{(k)} + \mathbf{1}_m \gamma_{ki} + \boldsymbol{\epsilon}_{ki}, \tag{6.13}$$

with single observations:

$$Y_{kij} = \beta_0 + \gamma_{ki} + \sum_{s=1}^{q} \beta_{sk} x_j^s + \epsilon_{kij} \tag{6.14}$$

$k \in \mathcal{X}_p, \quad x \in \mathcal{X} = [0, 1] \quad \text{and} \quad \mathbf{f}(x) = (x, x^2, \ldots, x^q)^\top.$

Let η^* be the uniform design on $\mathcal{X}_p$, $\boldsymbol{\xi} = \begin{pmatrix} \delta_{m,d} \\ 1 \end{pmatrix} \in \boldsymbol{\Theta}_1$ with

$$\delta_{m,d} = \begin{pmatrix} 0 & x_1 & \ldots & x_{q-1} & 1 \\ w_0 & w_1 & \ldots & w_{q-1} & w_q \end{pmatrix}$$

where the distinct settings x_s are ordered: $x_0 = 0 < x_1 < \cdots < x_q \leq 1$.

As we are seeking a D-optimal design and we are in the situation of the previous section, we can apply Theorem 6.1.

$$\det(\mathcal{M}(\eta^* \times \boldsymbol{\xi})) \propto \left(\det \widehat{\mathbf{F}}_{\mathcal{Y}}\right)^{2p} w_0 \left(\prod_{s=1}^{q} w_s\right)^p (1 + mdw_0)^{p-1}$$

where

$$\det \widehat{\mathbf{F}}_{\mathcal{G}} = \begin{vmatrix} x_1 & \dots & x_1^q \\ \vdots & \ddots & \vdots \\ x_{q-1} & \dots & x_{q-1}^q \\ 1 & \dots & 1 \end{vmatrix} = \left(\prod_{s=1}^{q-1} x_s(1-x_s)\right)\left(\prod_{s<s'}(x_s - x_{s'})\right).$$

Then continuous analytical methods show that the maximum is achieved when the design points $x_1^*, \dots, x_{q-1}^*$ are solutions of the equation $P_q'(2x-1) = 0$, where

$$P_q(2x-1) = (-1)^q \sum_{s=0}^{q} \binom{q}{s}\binom{q+s}{s}(-x)^s$$

is the explicit representation of the qth shifted and rescaled Legendre polynomial on the interval $[0,1]$, thus the support points of $\delta_{m,d}^*$ are independent of m, d and p, and on the other hand, the optimal proportions are

$$w_0^*(d) = \frac{1}{2(q+1)}\left(1 - \frac{pq+1}{pmd} + \sqrt{\left(1 - \frac{pq+1}{pmd}\right)^2 + \frac{4(q+1)}{pmd}}\right),$$

and $\quad w_s^*(d) = \dfrac{1 - w_0^*(d)}{q}, \quad s = 1, \dots, q$ respectively.

Remark. For the comparison of $p \geq 2$ treatments, but with only $q = 1$, that is $\mathbf{f}(x) = x$, optimal designs have been obtained in the model (6.14) by Schwabe(1996) and Schmelter(2008).

Consider an experimental situation which fits into the underlying model, where the partly interacting qualitative factor has $p = 2$ levels and $\mathbf{f}(x) = (x, x^2)^\top, \quad x \in [0,1]$, then we have that a D-optimal design for the full parameter

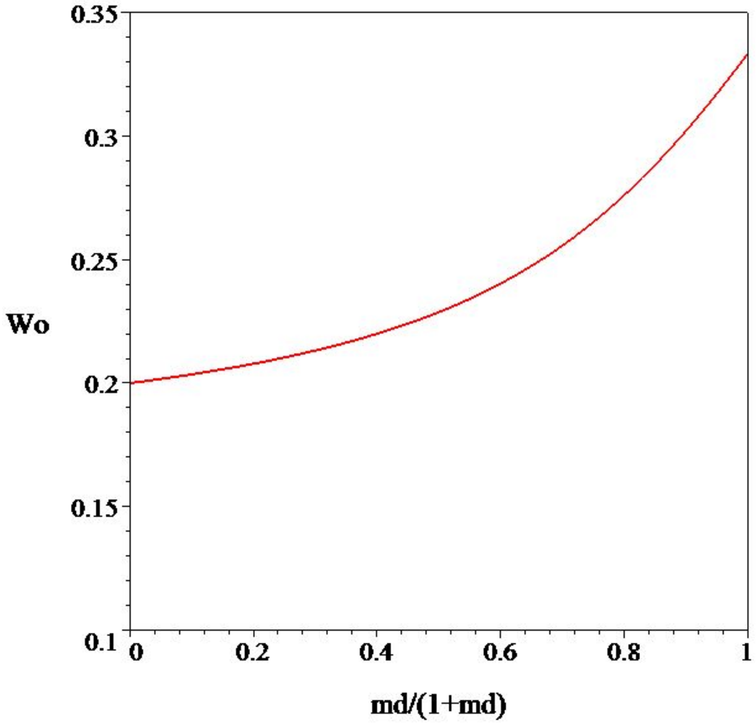

Figure 6.1; D-optimal proportion $w_0^*(d)$ for quadratic regression and p=2

is obtained when the single-block design $\boldsymbol{\xi}^*$ for the random block effects model is supported on an approximate design given by

$$\delta^*_{m,d} = \begin{pmatrix} 0 & 1/2 & 1 \\ w_0^*(d) & w_1^*(d) & w_2^*(d) \end{pmatrix}$$

where

$$\begin{aligned} w_0^*(d) &= \frac{1}{6}\left(1 - \frac{5}{2md} + \sqrt{\left(1 - \frac{5}{2md}\right)^2 + \frac{6}{md}}\right) \\ w_1^*(d) &= w_2^*(2) = \frac{1 - w_0^*(d)}{2}. \end{aligned}$$

The dependence of the D-optimal weight at $x = 0$ on the variance ratio d is presented in Figure(6.1), where to cover the whole range of d in a finite interval we rescale the horizontal axis with the transformation given by

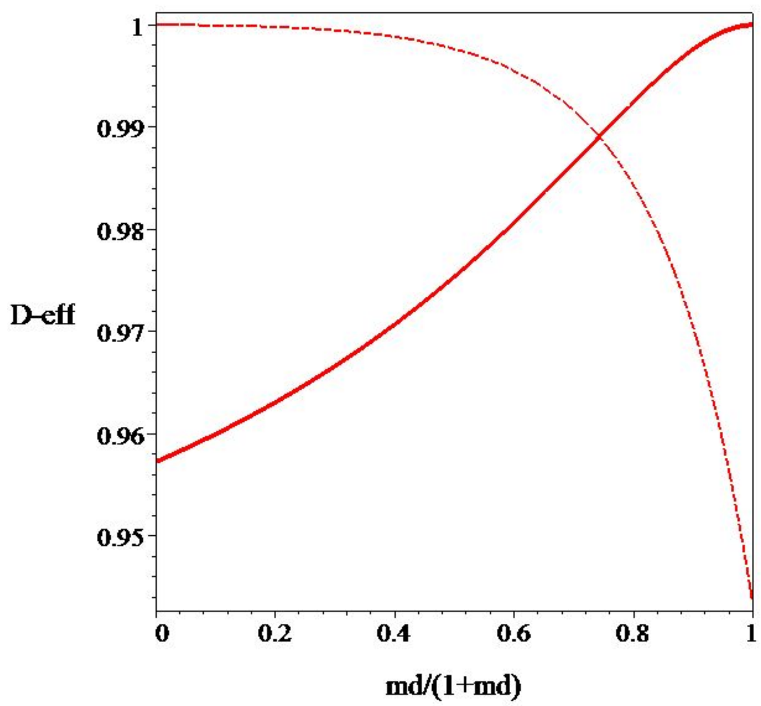

Figure 6.2; D-efficiencies: $\eta^* \times \boldsymbol{\xi}_0^*$ (dashed line) and $\eta^* \times \boldsymbol{\xi}_\infty^*$ (solid line) for quadratic regression and p=2 groups

$$d \longrightarrow \frac{md}{1+md}.$$

In the design literature, in order to calculate the efficiency of an arbitrary design $\eta^* \times \boldsymbol{\xi}$ with respected to the optimal design $\eta^* \times \boldsymbol{\xi}^*$ for the D-criteria, the D-efficiency of a design $\eta^* \times \boldsymbol{\xi}$ is defined as

$$D_{eff}(\eta^* \times \boldsymbol{\xi}) = \left(\frac{\det(\mathcal{M}(\eta^* \times \boldsymbol{\xi}))}{\det(\mathcal{M}(\eta^* \times \boldsymbol{\xi}^*))} \right)^{1/5}$$

where $5 = 1 + pq$ is the number of model parameters.

In the example under consideration, also we have plotted (in Figure 6.2), the D-efficiencies, respectively in dependence on $\frac{md}{1+md}$, of the limiting optimal designs, that is when the single block design for group are $\boldsymbol{\xi}_0^*$ and $\boldsymbol{\xi}_\infty^*$ support, respectively, on

$$\delta_{m,0}^* = \begin{pmatrix} 0 & 1/2 & 1 \\ 1/5 & 2/5 & 2/5 \end{pmatrix}$$

and

$$\delta_{m;\infty}^* = \begin{pmatrix} 0 & 1/2 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

and η^* is the discrete uniform design on the group levels set $\{1, \ldots, p\}$
As we can see, the figure shows that, the optimal block design based on the optimal design for the uncorrelated linear model without blocks effects is highly efficient for the two-factor mixed model over the whole range of d with a minimal D-efficiency of 0.9432 when the variance ratio d becomes large. On the other hand the D-optimum limiting fixed block effects design has a good performance for the two-factor model with random block for all d values with a minimal D-efficiency of 0.957248 at $d = 0$.

Chapter 7

Discussion and Outlook

In the present thesis we have developed mainly D- optimal designs for a two-factor linear model, where we have taken into consideration the two structures, without interactions and partial interactions between the effects of the qualitative and quantitative explanatory factors. An example of the first case that is of practical importance is the linear model with a constant term and additive block effects (we assume that the number of observations is the same for all blocks). Thus in chapters 3 and 4, using matrix algebra, we have presented that, if a $D_{\boldsymbol{\beta}}$-optimal design for estimating the regression parameter of the usual uncorrelated linear model can be [orthogonally] blocked, then it is D-optimal for the underling two-factor linear model. In the case of a fixed block effects model if the design is orthogonally blocked, then the regression parameter $\boldsymbol{\beta}$ is estimated independent of the block effects. In the case of a random block effects model, where we had to use the more computational technique of the linear mixed effects model as generalized least estimation, due to intra-block correlation, if the design is orthogonally blocked, then the regression parameter $\boldsymbol{\beta}$ is estimated independent of the ratio of variance components d. A particular class of orthogonally blocked designs is the class of single-block designs, where all blocks are observed under the same experimental setting.

With respect to the aim of this work we have presented a characterization of

D-optimal designs for a more complex structure, a two-factor mixed model with intercept given by a linear regression model where only the intercept is invariant of the qualitative fixed factor in the presence of random block effects. This characterization , under few assumptions, allowed to find analytically by means of convex optimization the weights of the optimal design. It is worthwhile noting that the optimal weights depend on the variance ratio. However, in this context if optimal single-block designs are computed where the linear regression is the polynomial regression in one explanatory variable, we show that the limiting optimal design are highly efficient, when the variance ratio value approaches to zero or infinity.
Future work could research the performance of partially interacting models involving more than one blocking factor using no single- block designs or generating designs in practical applications for multi-factor models in the presence of random block effects with different assumptions or experimental regions.

Bibliography

[1] Atkins J.E., Cheng, C.-S.(1999): Optimal regression design in the presence of random block effects. *J. Statist. Plann. Inference*, 77: 321-335.

[2] Atkinson, A. C., Donev, A. N. and Tobias, R. D. (2007). *Optimum Experimental Designs, with S.A.S.* Oxford University Press, Oxford.

[3] Cheng, C.-S.(1995): Optimal regression design under random block effects models. *Statistica Sinica*, 5: 485-497.

[4] Entholzner, M., Benda, N., Schmelter, T. and Schwabe, R. (2005): A note on designs for estimating population parameters. *Biometrical Letters* 42, 25-41.

[5] Fedorov, V.V.,(1972): *Theory of Optimal Experiments*, Academic Press, New York.

[6] Gilmour, S. G. and Trinka, L. A. (2000). Some practical advice on polynomial regression analysis from blocked response surface designs, *Communications in Statistics : Theory and Methods* 29, 2157-2180.

[7] Goos, P. (2002): *The Optimal Design of Blocked and Split – Plot Experiments*. Springer, New York.

[8] Goos, P., Vandebroek, M.(2001).D-optimal response surface designs in the presence of random block effects.*Comput. Statist. Data Analys.* 37, 433-453.

[9] Khuri, A.I. (1992). Response surface models with random block effects. *Technometrics*, 34, 26-37.

[10] Kiefer, J. (1959). Optimum experimental designs (with discussion). *Journal of the Royal Statistical Society*, Series B, 21, 272-319.

[11] Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* 2 849–879.

[12] Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, 12,363-366.

[13] Kunert, J.(1994): Optimality of block designs with variable block sizes and random block effects. *Metrika* 41:71-81.

[14] Kurotschka, V. G. (1984). A general approach to optimum design of experiments with qualitative and quantitative factors. In: J. K. Ghosh and J. Roy (Eds): *Statistics : Applications and New Directions : Proceedings of the Indian Statistical Institute Golden Jubilee International Conference Calcutta* 1981, 353-368.

[15] Pukelshein, F. (1993): *Optimal Design of Experiments*, John Wiley and Sons, New York.

[16] Rockafeller, R.T. (1970): *Convex analysis*, Princeton U.P.,Princeton N.J.

[17] Schmelter, T., (2007): The optimality of single-group designs for certain mixed models. *Metrika* 65 ,183-193.

[18] Schmelter, T., Schwabe, R., (2008): On optimal designs in random intercept models. *Tatra Mt. Math. Publ.* 39 : 145-153.

[19] Schwabe, R.(1996):*Optimum Design for Multi − Factor Models*. Springer, New York.

[20] Searle, S. R.(1971):*Linear Models*. Wiley, New York.

[21] Silvey, S. D.(1980):*Optimum Design*. Chapman and Hall, London.

[22] Waite, T.W., Woods, D.C. and Waterhouse, T.H. (2012): *Designs for generalized linear models with random block effects*. Southampton, GB, Southampton Statistical Sciences Research Institute, 23pp. (Southampton Statistical Sciences Research Institute Methodology Working Papers, M12/01).

[23] Wald, A. (1943). On the efficient design of statistical investigations, *Annals of Mathematical Statistics*, 14, 134-140.