# OPTIMAL DESIGNS FOR LINEAR MIXED MODELS


by

**LEGESSE KASSA DEBUSHO**

(MSc. Statistics)


Submitted in fulfilment of the requirements for the degree of

**DOCTOR OF PHILOSOPHY IN SCIENCE**


in the

School of Mathematics, Statistics and Information Technology

Faculty of Science and Agriculture

University of Kwazulu-Natal

Pietermaritzburg


August 2004

To the memory of

My dad Kassa Debusho

and

my mother-in-law Turuwerke Sileshi.

# Declaration

The work described in this thesis was carried out in the School of Mathematics, Statistics and Information Technology, University of Kwazulu-Natal, Pietermaritzburg, under the supervision of Professor Linda M. Haines and the co-supervision of Dr. Peter M. Njuho.

The thesis presents original work by the author and has not otherwise been submitted in any form for any degree or diploma to any University. Where use has been made of the work of others it is duly acknowledged in the text.

August, 2004.

**Student** :

Legesse Kassa Debusho

**Supervisor** :

Professor Linda M. Haines

**Co-supervisor** :

Dr Peter M. Njuho

## Acknowledgements

I would like to thank Professor Linda Haines, my supervisor, for her direction, advice and encouragement throughout the thesis. Her insight and patience made this thesis possible and coherent. It has been a privilege to work with her. I would also like to thank Dr. Peter Njuho, my co-supervisor, for all his input, direction and suggestions.

I should acknowledge Alemaya University, Ethiopia, for funding my studies at the University of Kwazulu-Natal and also thank the National Research Foundation of South Africa for financial support.

I am very grateful to Dr. Luc Duchateau for the use of the cattle data and for his valuable comments while he was at University of Natal on research leave. Grateful thanks are extended to Dr Principal Ndlovu of the School of Mathematics Statistics and Information Technology at the University of Kwazulu-Natal for the interest he showed in reading and commenting on the thesis at the earlier stages of the write up.

I do not forget to give special thanks to the Reverend Kebede Feyissa and his family and Demeke Mahteme for their kind words and consistent help through my study years.

My deepest thanks and appreciation go to my wife Rahel Genemew, to my mother Etenesh Berata, to my brothers Tadesse and Ashenafi and to my sisters Terefech and Tesfanesh for their love, support and encouragement. Their constant support enabled me to make this dream a reality.

Finally, I give honour to God in whom all things are possible.

# Abstract

The research of this thesis deals with the derivation of optimum designs for linear mixed models. The problem of constructing optimal designs for linear mixed models is very broad. Thus the thesis is mainly focused on the design theory for random coefficient regression models which are a special case of the linear mixed model. Specifically, the major objective of the thesis is to construct optimal designs for the simple linear and the quadratic regression models with a random intercept algebraically. A second objective is to investigate the nature of optimal designs for the simple linear random coefficient regression model numerically. In all models time is considered as an explanatory variable and its values are assumed to belong the set $\{0, 1, \ldots, k\}$. Two sets of individual designs, designs with non-repeated time points comprising up to $k+1$ distinct time points and designs with repeated time points comprising up to $k+1$ time points not necessarily distinct, are used in the thesis. In the first case there are $2^{k+1}-1$ individual designs while in the second case there are $2\begin{pmatrix} 2k+1 \\ k \end{pmatrix} - 1$ such designs. The problems of constructing population designs, which allocate weights to the individual designs in such a way that the information associated with the model parameters is in some sense maximized and the variances associated with the mean responses at a given vector of time points are in some sense minimized, are addressed. In particular $D$- and $V$-optimal designs are discussed. A geometric approach is introduced to confirm the global optimality of $D$- and $V$-optimal designs for the simple linear regression model with a random intercept. It is shown that for the simple linear regression model with a random intercept these optimal designs are robust to the choice of the variance ratio. A comparison of these optimal designs over the sets of individual designs with repeated and non-repeated points for that model is also made and indicates that the $D$- and $V$-optimal

population designs based on the individual designs with repeated points are more efficient than the corresponding optimal population designs with non-repeated points. Except for the one-point case, $D$- and $V$-optimal population designs change with the values of the variance ratio for the quadratic regression model with a random intercept. Further numerical results show that the $D$-optimal designs for the random coefficient models are dependent on the choice of variance components.

# Contents

# List of Tables

# List of Figures

# Notation

$\mathbf{X}$ : bold uppercase letters denote matrices

$\mathbf{J}$ : square matrix of ones

$\mathbf{I}$ : identity matrix

$\mathbf{y}$ : bold lower case letters denote vectors

$\boldsymbol{\beta}$ : vector of fixed effects

$\mathbf{b}$ : vector of random effects

$\mathbf{0}$ : vector of zeros

$\mathbf{1}$ : vector of ones

$\mathbf{M}_{\boldsymbol{\alpha}}$ : information matrix for vector of parameters $\boldsymbol{\alpha}$

$A, a$ : upper or lower case letters denote constants

$\mathbf{B} = \{b_{ij}\}_{i,j=1}^{n}$ : denotes an $n \times n$ matrix with elements $b_{ij}$, $i, j = 1, \ldots, n$

$\lambda_i$, $i = 1, \ldots, n$ : eigenvalues of an $n \times n$ matrix

$\mathbf{A}'$ : denotes the transpose of a matrix $\mathbf{A}$

$tr(\mathbf{A})$ : denotes the trace of a square matrix $\mathbf{A}$

$|\mathbf{A}|$ : denotes the determinant of a square matrix $\mathbf{A}$

$\mathbf{A} = diag\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m\}$ : denotes the block-diagonal matrix with blocks $\mathbf{A}_i$ on the main diagonal and zeros elsewhere

$\mathcal{N}_n$ : denotes the $n$-dimensional normal distribution

$\otimes$ : denotes the Kronecker product

# Chapter 1

# Introduction

Linear mixed models, which are models incorporating both fixed and random effects, are proving to be valuable and exciting tools for modelling a wealth of biological data. In particular, the concept of a linear mixed model draws together seemingly disparate structures such as split-plot, incomplete block and traditional agricultural models, models for repeated measures, longitudinal data and spatial statistics. This thesis deals with the use of linear mixed models for longitudinal data, which refers to data on individuals measured repeatedly at different times. Linear mixed models have attracted enormous interest in the statistical literature and present many fascinating challenges to researchers in applied statistics. One specific area within the context of mixed models which has considerable potential for good and meaningful research is that of constructing optimal experimental designs. The aim of this thesis is therefore to construct optimal designs for these models.

The thesis is organized as follows. A formal introduction of the linear mixed effects model, together with discussions on estimation of the model parameters and information matrices are given in Chapter 2. This chapter also consists of a discussion on optimal designs for linear mixed model and a review of related literature. The aims and objectives of the study, models of interest, designs and their spaces, design criteria and the data set that is used in this study are described in Chapter 3. The construction of $D$-optimal designs for estimating the parameters in the simple linear regression model with a random intercept as precisely as possible is discussed in Chapter 4. In Chapter 5, $V$-optimal population designs for the estimation of the mean responses at a vector of time points in the simple linear regression model with a random intercept are considered. Chapter 6 presents $D$-optimal population designs for estimation of the fixed effects in the quadratic regression model with a random intercept. $V$-optimal population designs for estimation of the mean responses in the quadratic regression model with a random intercept are discussed in Chapter 7. Optimal designs for the precise estimation of parameters in the random coefficient regression models are discussed numerically in Chapter 8. Finally, a summary of the results of this thesis with some open design problems in the theory of linear mixed effects models is given in Chapter 9.

Mathematica (Wolfram, 1999), a product of Wolfram Research Inc. and GAUSS were used extensively throughout chapters 4, 5, 6, 7 and 8.

# Chapter 2

# General Background

## 2.1 Linear mixed model

The linear mixed model methodology was first developed within the context of animal genetics and breeding research by Henderson, Kempthorne, Searle and Krosigk (1959). In recent years, however, the mixed model has also been introduced to analyze experiments with complex data structures in a variety of other disciplines, for example medicine (Brown and Prescott, 1999) and education (Goldstein, 1987). Many important statistical models can be expressed as mixed effects models or in other words models which incorporate both fixed effects and random effects. Examples of the underlying data sets include repeated measures data (Lindsey, 1993; Vonesh and Chinchilli, 1997), longitudinal data (Diggle, Liang and Zeger, 1994; Verbeke and Molenberghs, 1999), multilevel data (Goldstein, 1987), block designs (Goos, 2002) and pharmacokinetic data (Davidian and Giltinan, 1995, pages 262-272). Different names are also used in the statistical literature to describe the mixed model, reflecting the diversity of its use in many fields. These include hierarchical linear model, random effects model or variance components model and random coefficient regression model. The literature on linear mixed models is extensive and the basic results can be found in Searle (1971, 1987), Searle, Casella and McCulloch (1992) and McCulloch and Searle (2001). This thesis deals with the use of the linear mixed model for longitudinal data, which refers to data on individuals measured repeatedly at different times. A more general term is repeated measures data, which refers to data on individuals measured repeatedly either under different conditions or at different times.

In the present chapter the linear mixed model and the estimation of its parameters are briefly discussed and the related design problems and optimality criteria for the model are introduced. Specifically, the linear mixed model is described in Section 2.2 and the point estimation of the parameters of that model in Section 2.3. In Section 2.4, the information matrices for fixed effects parameters and variance-components are considered. Section 2.5 deals with random coefficient regression models, as a special case of the linear mixed model. Optimal design for the linear mixed model is discussed in Section 2.6. Finally, a brief review of the related literature on optimal design for linear mixed models is presented in Section 2.7.

## 2.2 Model

When repeated measures of responses are taken for individuals from a population two levels of variability arise, namely the between-individuals variability and the within-individual variability. Let $\mathbf{y}_i = (y_{i1}, \ldots, y_{id_i})'$ be the $d_i \times 1$ vector of responses from the $i$th individual for $i = 1, 2, \ldots, K$. In matrix notation the linear mixed model for the $i$th individual is given by

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i, \quad i = 1, \ldots, K \tag{2.1}$$

where

$\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown fixed effects which is common to the $K$ individuals,

$\mathbf{X}_i$ is a $d_i \times p$ design matrix associating $\boldsymbol{\beta}$ to $\mathbf{y}_i$,

$\mathbf{b}_i$ is a $q \times 1$ vector of random effects, i.e. the between-individual random effects,

$\mathbf{Z}_i$ is a $d_i \times q$ design matrix which relates $\mathbf{b}_i$ to the response $\mathbf{y}_i$, and

$\mathbf{e}_i$ is a $d_i \times 1$ vector of random errors, i.e. a vector of within-individual errors.

The random effects vectors $\mathbf{b}_i$ are assumed to be independently and normally distributed with mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{G}$, i.e. $\mathbf{b}_i \sim \mathcal{N}_q(\mathbf{0}, \mathbf{G})$, and the error vectors $\mathbf{e}_i$ are assumed to be independently and normally distributed with mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{R}_i$, i.e. $\mathbf{e}_i \sim \mathcal{N}_{d_i}(\mathbf{0}, \mathbf{R}_i)$, for $i = 1, 2, \ldots, K$. Note that $\mathbf{G}$ and $\mathbf{R}_i$ are $q \times q$ and $d_i \times d_i$ matrices respectively. In addition, both $\mathbf{b}_i$ and $\mathbf{e}_i$ are assumed to be independent within and between individuals.

Under the independence and normality assumptions for $\mathbf{b}_i$ and $\mathbf{e}_i$, the marginal distribution of the response $\mathbf{y}_i$ is normal with mean $\mathbf{X}_i \boldsymbol{\beta}$ and variance-covariance matrix $\mathbf{V}_i = \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i' + \mathbf{R}_i$, i.e. $\mathbf{y}_i \sim \mathcal{N}(\mathbf{X}_i \boldsymbol{\beta}, \mathbf{V}_i)$. Suppose that the matrices $\mathbf{V}_i$, $i = 1, \ldots, K$, depend on a vector of parameters $\boldsymbol{\theta}$. Specifically the parameter vector $\boldsymbol{\theta}$ consists of the $\dfrac{q\,(q+1)}{2} + \sum_{i=1}^{K} \dfrac{d_i\,(d_i+1)}{2}$ distinct variance-covariance elements of the matrices $\mathbf{G}$ and $\mathbf{R}_i$,

$i = 1, \ldots, K$. The elements of $\boldsymbol{\theta}$ are called variance components. The marginal likelihood function for the response of an individual can thus be expressed as

$$L(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i) = (2\pi)^{-d_i/2} |\mathbf{V}_i|^{-1/2} \exp\left\{ -\tfrac{1}{2}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})'\mathbf{V}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}) \right\} \tag{2.2}$$

for $i = 1, \ldots, K$.

The population model for all $K$ individuals has the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \mathbf{e},$$

where

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_K \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_K \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_K \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_K \end{pmatrix},$$

$\mathbf{Z} = \mathrm{diag}\{\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_K\}$, i.e. $\mathbf{Z}$ is the block-diagonal matrix with blocks $\mathbf{Z}_i$ on the main diagonal and zeros elsewhere, $Var(\mathbf{b}) = \mathrm{diag}\{\mathbf{G}, \mathbf{G}, \ldots, \mathbf{G}\}$ and $\mathbf{R} = Var(\mathbf{e}) = \mathrm{diag}\{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K\}$. The marginal distribution of $\mathbf{y}$ is $\mathcal{N}(\mathbf{X}\boldsymbol{\beta},\ \mathbf{V})$, where
$\mathbf{V} = \mathbf{Z}\,\mathrm{diag}\{\mathbf{G}, \mathbf{G}, \ldots, \mathbf{G}\}\,\mathbf{Z}' + \mathbf{R} = \mathrm{diag}\{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_K\}$. Therefore the marginal likelihood function associated with the full vector of responses $\mathbf{y}$ follows from the individual likelihood (2.2) as

$$L(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}) = (2\pi)^{-n/2} |\mathbf{V}|^{-\frac{1}{2}} \exp\left\{ -\tfrac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right\} \tag{2.3}$$

where $n = \sum_{i=1}^{K} d_i$.

The linear mixed model for repeated measures in the case where $Var(\mathbf{e}_i) = \sigma_e^2\mathbf{I}_{d_i}$, $i = 1, \ldots, K$, was introduced by Laird and Ware (1982). Lindstrom and Bates (1988) considered the more general case with $Var(\mathbf{e}_i) = \mathbf{R}_i$, $i = 1, \ldots, K$, where $\mathbf{R}_i$ is $d_i \times d_i$ matrix which does not depend on $i$ except for its dimension, i.e. it has the same structure for each individual. Unless stated otherwise it is assumed in this thesis that $\mathbf{R}_i = \sigma_e^2 \mathbf{I}_{d_i}$, $i = 1, \ldots, K$, and hence that $\mathbf{R} = \text{diag}\{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K\} = \sigma_e^2 \mathbf{I}_n$.

## 2.3 Estimation of the fixed effects and the variance components

In studies on linear mixed models it is usual to consider the estimation of the fixed effects $\boldsymbol{\beta}$ and the variance components $\boldsymbol{\theta}$, and also the prediction of the random effects $\mathbf{b}$. For a given data vector $\mathbf{y}$, the vector of random effects $\mathbf{b}$ is a realization of random variables which are usually unobservable and these effects must therefore necessarily be predicted from the data (Henderson, 1953). The prediction of the random effects as best linear unbiased predictors (BLUP's) is briefly discussed at the end of this section. More detailed discussions and examples are given in Robinson (1991) and Searle, Casella and McCulloch (1992, Chapter 7). The emphasis in this thesis is however on the precise estimation of the fixed effects and the variance components and attention is therefore confined to this problem.

### 2.3.1 Estimation of the fixed effects $\boldsymbol{\beta}$

Assume firstly that the variance components $\boldsymbol{\theta}$ are known. Then the fixed effects parameter $\boldsymbol{\beta}$ can be estimated by the method of maximum likelihood (ML). Specifically it follows from expressions (2.2) and (2.3) that the marginal log-likelihood function for the response vector

$\mathbf{y}$ is given by

$$\ell = \ln L(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}) = -\sum_{i=1}^{K} \left\{ \frac{d_i}{2} \ln(2\pi) + \frac{1}{2} \ln |\mathbf{V}_i| + \frac{1}{2} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta})' \mathbf{V}_i^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}) \right\}. \quad (2.4)$$

The maximum likelihood estimator of $\boldsymbol{\beta}$ is obtained by maximizing this function with respect to $\boldsymbol{\beta}$. In particular, differentiating $\ell$ with respect to $\boldsymbol{\beta}$ yields

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{K} (\mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{y}_i - \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \boldsymbol{\beta}). \quad (2.5)$$

Equating the derivatives in (2.5) to zero gives the maximum likelihood estimator of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}} = \left( \sum_{i=1}^{K} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{K} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{y}_i$$

with variance-covariance matrix

$$Var(\hat{\boldsymbol{\beta}}) = \left( \sum_{i=1}^{K} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \right)^{-1}.$$

The estimate of the fixed effects involves the variance components vector $\boldsymbol{\theta}$. When $\boldsymbol{\theta}$ is known, the estimate $\hat{\boldsymbol{\beta}}$ is a function of $\boldsymbol{\theta}$ through the variance-covariance matrices $\mathbf{V}_i$. In practice, however, $\boldsymbol{\theta}$ is not known and must be estimated from the data. Estimation of $\boldsymbol{\theta}$ is discussed in the following subsection.

### 2.3.2 Estimation of the variance components

The two most widely used methods for estimating the variance components are maximum likelihood, as introduced by Hartley and Rao (1967), and the restricted maximum likelihood (REML) approach of Patterson and Thompson (1971).

The ML estimator of $\boldsymbol{\theta}$ is obtained by maximizing the marginal log-likelihood function of $\mathbf{y}$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. Maximization of $\ell$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ requires solving the equations

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} = 0 \quad \text{and} \quad \frac{\partial \ell}{\partial \boldsymbol{\theta}} = 0$$

for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ simultaneously. The first equation yields

$$\hat{\boldsymbol{\beta}} = \left( \sum_{i=1}^{K} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{K} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{y}_i$$

as shown in Subsection 2.3.1. Thus $\hat{\boldsymbol{\beta}}$ is a function of $\boldsymbol{\theta}$ and can be written as $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$. Substituting $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$ into $\ell$ yields this log-likelihood as a function of $\boldsymbol{\theta}$, i.e. $\ell(\boldsymbol{\theta})$. The ML estimator of $\boldsymbol{\theta}$, say $\hat{\boldsymbol{\theta}}$, can be found by solving the equation $\dfrac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0$ (see Searle, Casella and McCulloch, 1992, pages 234-235; Verbeke and Molenberghs, 1999, page 42). In practice the equation $\dfrac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0$ is solved numerically by iteration, e.g. using the Newton Raphson method. Finally, substituting $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$ in $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$ gives the ML estimator of $\boldsymbol{\beta}$, i.e. $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})$.

One criticism of the ML approach to estimation of the variance components $\boldsymbol{\theta}$ is that it takes no account of the loss of degrees of freedom that results from estimating the fixed effects $\boldsymbol{\beta}$. As a result the estimates of the variance components are biased downward (Verbeke and Molenberghs, 1999, page 43). To overcome this limitation Patterson and Thompson (1971) proposed the restricted maximum likelihood approach. This approach applies ML estimation techniques to the likelihood function of a set of error contrasts defined as any linear combination $\mathbf{A}\,\mathbf{y}$ of the response $\mathbf{y}$ with zero expectation. The matrix $\mathbf{A}$ is taken to be an $(n-p) \times n$ matrix of full row rank and is orthogonal to the columns of $\mathbf{X}$, i.e. $\mathbf{A}\,\mathbf{X} = \mathbf{0}$. Thus the distribution of $\mathbf{A}\,\mathbf{y}$ is normal with mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{A}\,\mathbf{V}\,\mathbf{A}'$, i.e. $\mathbf{A}\,\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{A}\,\mathbf{V}\,\mathbf{A}')$ and does not depend on the fixed effects $\boldsymbol{\beta}$ (Harville, 1974). The likelihood function associated with the vector of error contrasts can be expressed as

$$L_{REML}(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}) = (2\pi)^{-(n-p)/2} \left|\mathbf{X}'\mathbf{X}\right|^{-1/2} \left|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right|^{-1/2} \left|\mathbf{V}\right|^{-1/2}$$

$$\times \exp\left\{ -\tfrac{1}{2}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right\}$$

where $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$. Then the REML estimates for $\boldsymbol{\theta}$, say $\hat{\boldsymbol{\theta}}_{REML}$, can be obtained by maximizing $L_{REML}(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y})$ using an iterative procedure (see Searle, Casella and McCulloch, 1992, Chapter 8). Observe that the likelihood function $L_{REML}(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y})$ and hence the REML estimator of $\boldsymbol{\theta}$ does not depend on the choice of the matrix $\mathbf{A}$. Finally substituting $\hat{\boldsymbol{\theta}}_{REML}$ into the expression $\hat{\boldsymbol{\beta}}$ yields the associated estimate for $\boldsymbol{\beta}$.

### 2.3.3 Prediction of the random effects b

The parameters $\boldsymbol{\beta}$ can be estimated and the random effects $\mathbf{b}$ predicted together by maximizing

$$\begin{aligned} f(\mathbf{y}, \mathbf{b}; \boldsymbol{\beta}, \boldsymbol{\theta}) &= \prod_{i=1}^{K} f(\mathbf{y}_i|\mathbf{b}_i)\, f(\mathbf{b}_i) \\ &= \prod_{i=1}^{K} \left\{ \frac{1}{(2\,\pi)^{d_i/2}\, |\mathbf{R_i}|^{1/2}} \right\} \exp\left\{ -\tfrac{1}{2}(\mathbf{y}_i - \mathbf{X}_i\,\boldsymbol{\beta} - \mathbf{Z}_i\,\mathbf{b}_i)'\mathbf{R_i}^{-1}(\mathbf{y}_i - \mathbf{X}_i\,\boldsymbol{\beta} - \mathbf{Z}_i\,\mathbf{b}_i) \right\} \\ &\quad \times \left\{ \frac{1}{(2\,\pi)^{q/2}\, |\mathbf{G}|^{1/2}} \right\} \exp\left\{ -\tfrac{1}{2}\,\mathbf{b}_i'\,\mathbf{G}^{-1}\,\mathbf{b}_i \right\} \end{aligned}$$

with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. Specifically, the hierarchical log-likelihood function for $f(\mathbf{y}, \mathbf{b}; \boldsymbol{\beta}, \boldsymbol{\theta})$ can be written as

$$\begin{aligned} \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}, \mathbf{b}) &= \sum_{i=1}^{K} \{\ln f(\mathbf{y}_i|\mathbf{b}_i) + \ln f(\mathbf{b}_i)\} \\ &= -\sum_{i=1}^{K} \left\{ \frac{d_i}{2} \ln(2\,\pi) + \frac{1}{2} \ln|\mathbf{R_i}| + \frac{1}{2}(\mathbf{y}_i - \mathbf{X}_i\,\boldsymbol{\beta} - \mathbf{Z}_i\,\mathbf{b}_i)'\,\mathbf{R_i}^{-1}(\mathbf{y}_i - \mathbf{X}_i\,\boldsymbol{\beta} - \mathbf{Z}_i\,\mathbf{b}_i) \right. \\ &\qquad \left. + \frac{q}{2} \ln(2\,\pi) + \frac{1}{2} \ln|\mathbf{G}| + \frac{1}{2}\,\mathbf{b}_i'\,\mathbf{G}^{-1}\,\mathbf{b}_i \right\}. \end{aligned} \tag{2.6}$$

Note that this function is not strictly a log-likelihood since the random effects $\mathbf{b}_i$ are not observed, and is therefore termed a hierarchical log-likelihood following Lee and Nelder (1996).

Differentiating $\ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}, \mathbf{b})$ with respect to $\boldsymbol{\beta}$ and $\mathbf{b}$ and equating the derivatives to zero gives the system of equations known as the mixed model equations

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}'$$

(Henderson *et al.*, 1959), where $\hat{\mathbf{b}}$ is the predictor of $\mathbf{b}$. Solving for $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{b}}_i$, yields the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}} = \left\{ \sum_{i=1}^{K} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \right\}^{-1} \sum_{i=1}^{K} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{y}_i$$

and the best linear unbiased predictors (BLUPs) of the $\mathbf{b}_i$ as

$$\hat{\mathbf{b}}_i = \mathbf{G}\,\mathbf{Z}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i - \mathbf{X}\hat{\boldsymbol{\beta}}), \quad i = 1, \ldots, K.$$

The BLUE of $\boldsymbol{\beta}$ requires knowing $\mathbf{V}_i$ whereas the BLUPs of $\mathbf{b}_i$ require knowing both $\mathbf{V}_i$ and $\mathbf{G}$. Note that the BLUP's are termed predictors in order to distinguish them from the estimators for the fixed effects.

## 2.4 Fisher information matrix

The information matrix for the parameters is equal to minus the expected value of the matrix of second-order derivatives of the log-likelihood function, where the derivatives are with respect to the parameters (see Azzalini, 1996, page 73). Specifically, suppose that $\boldsymbol{\alpha}$ is a vector of parameters in the model. Then the information matrix for $\boldsymbol{\alpha}$ is given by

$$\mathbf{I}_\alpha = -E\left(\frac{\partial^2 \ell(\boldsymbol{\alpha};\mathbf{y})}{\partial\boldsymbol{\alpha}\,\partial\boldsymbol{\alpha}'}\right)$$

where $\ell(\boldsymbol{\alpha};\mathbf{y})$ denotes the log-likelihood function for $\mathbf{y}$. The information matrix plays a crucial role in inference. For example, according to the Cramér-Rao theorem for a single parameter the inverse of the information matrix is proportional to the lower bound on the variance of any unbiased estimator of the parameter. Further, it is well known that the variance of the maximum likelihood estimator of the parameters can be approximated asymptotically by the inverse of the information matrix (Azzalini, 1996, page 83). The variance of the estimator of a parameter is a measure of its precision with small variance associated with high precision. Thus, more generally, the information matrix for the parameters, or some function of that matrix, can be used as a measure of precision. For this reason the information matrix forms the basis for optimality criteria in the design of experiments.

For the linear mixed model introduced in the previous section, the generic parameter $\boldsymbol{\alpha}$ corresponds to the parameters $\boldsymbol{\beta}$ or $\boldsymbol{\theta}$ or $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ together. Note that the variance-covariance matrices of the ML and the REML estimators are both approximated by the inverse of the information matrix (Verbeke and Molenberghs, 1999, page 64). The appropriate information matrices for the parameters in the linear mixed model are presented in the next subsections. The derivations of the information matrices are based on those presented in Searle, *et al.* (1992) and Verbeke and Molenberghs (1999).

### 2.4.1 Information matrix for $\boldsymbol{\beta}$

Recall that the marginal log-likelihood for $\mathbf{y}_i$, the response of the $i$th individual, is given by

$$\ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i) = -\frac{d_i}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{V}_i| - \frac{1}{2}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})'\mathbf{V}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}) \tag{2.7}$$

for $i = 1, \ldots, K$. Clearly, since the responses are independent, the marginal log-likelihood associated with the full vector of responses $\mathbf{y}$ is

$$\ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}) = \sum_{i=1}^{K} \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i). \tag{2.8}$$

Now the first- and second-order derivatives of the $i$th individual log-likelihood (2.7) with respect to $\boldsymbol{\beta}$ are given by

$$\frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial \boldsymbol{\beta}} = \mathbf{X}_i'\mathbf{V}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}) \tag{2.9}$$

and

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\beta}'} = -\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i$$

respectively. Therefore, for a given $\boldsymbol{\theta}$, the information matrix for $\boldsymbol{\beta}$ for the $i$th individual is

$$\begin{aligned} \mathbf{I}_{\beta}(\mathbf{X}_i, \mathbf{Z}_i) &= -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}\right) \\ &= \mathbf{X}_i'\,\mathbf{V}_i^{-1}\,\mathbf{X}_i = \mathbf{X}_i'\,(\mathbf{Z}_i\,\mathbf{G}\,\mathbf{Z}_i' + \mathbf{R}_i)^{-1}\,\mathbf{X}_i, \end{aligned} \tag{2.10}$$

where $\mathbf{R}_i = \sigma^2\,\mathbf{I}_{d_i}$, $i = 1, \ldots, K$. It thus follows from (2.8) and (2.10) that the information matrix for $\boldsymbol{\beta}$ over all individuals can be expressed succinctly as

$$\mathbf{I}_{\beta}(\mathbf{X}, \mathbf{Z}) = \sum_{i=1}^{K} \mathbf{X}_i'\mathbf{V}_i^{-1}\,\mathbf{X}_i. \tag{2.11}$$

### 2.4.2 Information matrix for $\boldsymbol{\theta}$

Let $\theta_j$ denote the $j$th element of the variance components vector $\boldsymbol{\theta}$, $j = 1, \ldots, \frac{q(q+1)}{2} + 1$, and let $\mathbf{w}_i = \mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}$, $i = 1, \ldots, K$. Note that $\mathbf{w}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_i)$. Then differentiating the log-likelihood given in expression (2.7) with respect to $\theta_j$ and using Results A.4.1 and A.4.2 from Appendix A yields

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial \theta_j} &= -\frac{1}{2}\left[\frac{\partial \ln|\mathbf{V}_i|}{\partial \theta_j} + \mathbf{w}_i' \frac{\partial \mathbf{V}_i^{-1}}{\partial \theta_j}\mathbf{w}_i\right] \\
&= -\frac{1}{2}\left[tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\right) - \mathbf{w}_i'\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\mathbf{V}_i^{-1}\mathbf{w}_i\right].
\end{aligned} \tag{2.12}$$

The second-order partial derivative of $\ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)$ with respect to the parameters $\theta_j$ and $\theta_k$ is thus given by

$$\begin{aligned}
&\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial \theta_j \partial \theta_k} = \\
&\frac{1}{2}\left[tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\right) - tr\left(\mathbf{V}_i^{-1}\frac{\partial^2 \mathbf{V}_i}{\partial \theta_j \partial \theta_k}\right) - 2\mathbf{w}_i'\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\mathbf{V}_i^{-1}\mathbf{w}_i\right. \\
&\qquad\qquad \left. + \mathbf{w}_i'\mathbf{V}_i^{-1}\frac{\partial^2 \mathbf{V}_i}{\partial \theta_j \partial \theta_k}\mathbf{V}_i^{-1}\mathbf{w}_i\right].
\end{aligned} \tag{2.13}$$

Since $E(\mathbf{w}_i\mathbf{w}_i') = \mathbf{V}_i$ it follows that

$$E\left(\mathbf{w}_i'\mathbf{V}_i^{-1}\frac{\partial^2 \mathbf{V}_i}{\partial \theta_j \partial \theta_k}\mathbf{V}_i^{-1}\mathbf{w}_i\right) = E\left[tr\left(\mathbf{V}_i^{-1}\frac{\partial^2 \mathbf{V}_i}{\partial \theta_j \partial \theta_k}\mathbf{V}_i^{-1}\mathbf{w}_i\mathbf{w}_i'\right)\right] = tr\left(\mathbf{V}_i^{-1}\frac{\partial^2 \mathbf{V}_i}{\partial \theta_j \partial \theta_k}\right)$$

and

$$\begin{aligned}
E\left(\mathbf{w}_i'\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\mathbf{V}_i^{-1}\mathbf{w}_i\right) &= E\left[tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\mathbf{V}_i^{-1}\mathbf{w}_i\mathbf{w}_i'\right)\right] \\
&= tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\right).
\end{aligned}$$

Substitution of these identities into the expectation of expression (2.13) yields

$$E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial \theta_j \partial \theta_k}\right) = -\frac{1}{2}tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\right).$$

Thus the $(j, k)$th element of the information matrix from the $i$th individual corresponding to parameters $\theta_j$ and $\theta_k$ is given by

$$\mathbf{I}_{\theta_j\,\theta_k}(\mathbf{X}_i, \mathbf{Z}_i) = -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial\theta_j\,\partial\theta_k}\right) = \frac{1}{2} tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial\theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial\theta_k}\right) \tag{2.14}$$

for $i = 1, \ldots, K$. Clearly, for the full data set $\mathbf{y}$, i.e. for all individuals, this becomes

$$\mathbf{I}_{\theta_j\,\theta_k}(\mathbf{X}, \mathbf{Z}) = \frac{1}{2}\sum_{i=1}^{K} tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial\theta_j}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial\theta_k}\right).$$

Thus, the information matrix for $\boldsymbol{\theta}$, $\mathbf{I}_\theta(\mathbf{X}, \mathbf{Z})$, over all individuals is a matrix with $(j, k)$th element equal to $\mathbf{I}_{\theta_j\,\theta_k}(\mathbf{X}, \mathbf{Z})$.

### 2.4.3 Information matrix for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$

Taking the partial derivative of expression (2.9) with respect to $\theta_j$ gives

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y}_i)}{\partial\boldsymbol{\beta}\,\partial\theta_j} = \mathbf{X}_i'\frac{\partial \mathbf{V}_i^{-1}}{\partial\theta_j}(\mathbf{y}_i - \mathbf{X}_i\,\boldsymbol{\beta})$$

and thus the information matrix for $\boldsymbol{\beta}$ and $\theta_j$ is

$$-E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \theta; \mathbf{y}_i)}{\partial\boldsymbol{\beta}\,\partial\theta_j}\right) = -\mathbf{X}_i'\frac{\partial \mathbf{V}_i^{-1}}{\partial\theta_j}E(\mathbf{y}_i - \mathbf{X}_i\,\boldsymbol{\beta}) = 0. \tag{2.15}$$

It thus follows that the information matrix for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ jointly can be expressed in block-diagonal form as

$$\mathbf{I}_{\beta,\theta}(\mathbf{X}, \mathbf{Z}) = \begin{pmatrix} \mathbf{I}_\beta(\mathbf{X}, \mathbf{Z}) & 0 \\ 0 & \mathbf{I}_\theta(\mathbf{X}, \mathbf{Z}) \end{pmatrix}, \tag{2.16}$$

where the matrices $\mathbf{I}_\beta(\mathbf{X}, \mathbf{Z})$ and $\mathbf{I}_\theta(\mathbf{X}, \mathbf{Z})$ are the information matrices over all individuals for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, respectively.

### 2.4.4 Information matrix for $\sigma_e^2$ and G

It is very common to consider $\mathbf{V}_i = \mathbf{Z}_i\,\mathbf{G}\,\mathbf{Z}_i' + \sigma_e^2\,\mathbf{I}_{d_i}$ with variance components corresponding to $\sigma_e^2$ and the distinct element of $\mathbf{G}$. The results pertaining to these components are a special case of those in Section 2.4.2 and they are presented in this subsection. Two cases are considered, one with $\mathbf{V}_i = \mathbf{Z}_i\,\mathbf{G}\,\mathbf{Z}_i' + \sigma_e^2\,\mathbf{I}_{d_i}$ and the other with $\mathbf{V}_i = \sigma_e^2\,(\mathbf{Z}_i\,\mathbf{G}^*\,\mathbf{Z}_i' + \mathbf{I}_{d_i})$, where $\mathbf{G}^* = \dfrac{1}{\sigma_e^2}\mathbf{G}$.

**Case 1:** $\mathbf{V}_i = \mathbf{Z}_i\,\mathbf{G}\,\mathbf{Z}_i' + \sigma_e^2\,\mathbf{I}_{d_i}$

Suppose the vector of variance components $\boldsymbol{\theta}$ composes the variance $\sigma_e^2$ and the $\dfrac{q\,(q+1)}{2}$ distinct elements of the variance-covariance matrix of $\mathbf{b}_i$, $\mathbf{G}$. Let $g_{rs}$ be the $(r,s)$th element of $\mathbf{G}$. Then

$$\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} = \mathbf{I}_{d_i}.$$

Thus, from expression (2.14) it follows that

$$\mathbf{I}_{(\sigma_e^2,\sigma_e^2)}(\mathbf{X}_i,\mathbf{Z}_i) = \frac{1}{2} tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2}\right) = \frac{1}{2} tr\left(\mathbf{V}_i^{-2}\right). \tag{2.17}$$

Furthermore,

$$\frac{\partial \mathbf{V}_i}{\partial g_{rs}} = \mathbf{Z}_i\,\frac{\partial \mathbf{G}}{\partial g_{rs}}\mathbf{Z}_i'.$$

Since $\mathbf{G}$ is a symmetric matrix

$$\frac{\partial \mathbf{G}}{\partial g_{rs}} = \begin{cases} \mathbf{u}_r\mathbf{u}_r' & \text{if } \; r = s \\ \mathbf{u}_r\mathbf{u}_s' + \mathbf{u}_s\mathbf{u}_r' & \text{if } \; r \neq s \end{cases}$$

where $\mathbf{u}_r$ represent the $r$th column of the identity matrix $\mathbf{I}_{d_i}$ (see Harville, 1997, pages

299-300). Thus

$$\frac{\partial \mathbf{V}_i}{\partial g_{rs}} = \begin{cases} \mathbf{Z}_i \mathbf{u}_r \mathbf{u}_r' \mathbf{Z}_i', & \text{if } r = s \\ \mathbf{Z}_i \mathbf{u}_r \mathbf{u}_s' \mathbf{Z}_i' + \mathbf{Z}_i \mathbf{u}_s \mathbf{u}_r' \mathbf{Z}_i', & \text{if } r \neq s \end{cases}$$

$$= \begin{cases} \mathbf{z}_{i,(r)} \mathbf{z}_{i,(r)}', & \text{if } r = s \\ \mathbf{z}_{i,(r)} \mathbf{z}_{i,(s)}' + \mathbf{z}_{i,(s)} \mathbf{z}_{i,(r)}', & \text{if } r \neq s \end{cases}$$

or, equivalently,

$$\frac{\partial \mathbf{V}_i}{\partial g_{rs}} = \frac{1}{2^{\delta_{rs}}} \{\mathbf{z}_{i,(r)} \mathbf{z}_{i,(s)}' + \mathbf{z}_{i,(s)} \mathbf{z}_{i,(r)}'\} \tag{2.18}$$

where $\mathbf{z}_{i,(r)}$ is the $r$th column of $\mathbf{Z}_i$ and $\delta_{rs}$ is the Kronecker delta defined as

$$\delta_{rs} = \begin{cases} 1, & \text{if } r = s \\ 0, & \text{if } r \neq s \end{cases} .$$

Thus, from expressions (2.14) and (2.18) it follows that

$$\begin{aligned}
\mathbf{I}_{g_{rs}, g_{tu}}(\mathbf{X}_i, \mathbf{Z}_i) &= \frac{1}{2} tr \left( \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial g_{rs}} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial g_{tu}} \right) \\
&= \frac{1}{2} \times \frac{1}{2^{\delta_{rs}} \times 2^{\delta_{tu}}} tr \left\{ \mathbf{V}_i^{-1} \left[ \mathbf{z}_{i,(r)} \mathbf{z}_{i,(s)}' + \mathbf{z}_{i,(s)} \mathbf{z}_{i,(r)}' \right] \mathbf{V}_i^{-1} \left[ \mathbf{z}_{i,(t)} \mathbf{z}_{i,(u)}' + \mathbf{z}_{i,(u)} \mathbf{z}_{i,(t)}' \right] \right\} \\
&= \frac{1}{2} \times \frac{1}{2^{\delta_{rs}} \times 2^{\delta_{tu}}} \{ (\mathbf{z}_{i,(u)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(r)})(\mathbf{z}_{i,(s)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(t)}) + (\mathbf{z}_{i,(t)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(r)})(\mathbf{z}_{i,(s)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(u)}) \\
&\quad + (\mathbf{z}_{i,(u)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(s)})(\mathbf{z}_{i,(r)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(t)}) + (\mathbf{z}_{i,(t)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(s)})(\mathbf{z}_{i,(r)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(u)}) \} \\
&= \frac{1}{2^{\delta_{rs}} \times 2^{\delta_{tu}}} \{ (\mathbf{z}_{i,(u)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(r)})(\mathbf{z}_{i,(s)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(t)}) + (\mathbf{z}_{i,(t)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(r)})(\mathbf{z}_{i,(s)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(u)}) \} .
\end{aligned} \tag{2.19}$$

Let $\mathbf{C}_i = \mathbf{Z}_i' \mathbf{V}_i^{-1} \mathbf{Z}_i$ and let $C_{i,ru} = \mathbf{z}_{i,(r)}' \mathbf{V}_i^{-1} \mathbf{z}_{i,(u)}$ denote the $(r, u)$th element of $\mathbf{C}_i$. Then (2.19) can be more written succinctly as

$$\mathbf{I}_{g_{rs}, g_{tu}}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2^{\delta_{rs}} \times 2^{\delta_{tu}}} \{ C_{i,ru} C_{i,st} + C_{i,su} C_{i,rt} \} . \tag{2.20}$$

Furthermore, expressions (2.14) and (2.18) yield

$$\mathbf{I}_{g_{rs},\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2} tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial g_{rs}}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2}\right) = \frac{1}{2} tr\left(\mathbf{V}_i^{-2}\frac{\partial \mathbf{V}_i}{\partial g_{rs}}\right)$$

$$= \frac{1}{2} \times \frac{1}{2^{\delta_{rs}}} tr\left\{\mathbf{z}'_{i,(s)}\mathbf{V}_i^{-2}\mathbf{z}_{i,(r)} + \mathbf{z}'_{i,(r)}\mathbf{V}_i^{-2}\mathbf{z}_{i,(s)}\right\} = \frac{1}{2^{\delta_{rs}}}\mathbf{z}'_{i,(r)}\mathbf{V}_i^{-2}\mathbf{z}_{i,(s)}.$$

Let $\mathbf{D}_i = \mathbf{Z}'_i\, \mathbf{V}_i^{-2}\, \mathbf{Z}_i$ and thus $D_{i,rs} = \mathbf{z}'_{i,(r)}\mathbf{V}_i^{-2}\mathbf{z}_{i,(s)}$ is the $(r, s)$th element of $\mathbf{D}_i$. Thus

$$\mathbf{I}_{g_{rs},\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2^{\delta_{rs}}} D_{i,rs}. \tag{2.21}$$

A convenient ordering for the elements of $\mathbf{G}$ was introduced by Mentré, Mallet and Baccar (1997). Specifically, let $r.s = s + \dfrac{r\,(r-1)}{2}$. Then the term $\mathbf{I}_{g_{rs},g_{tu}}(\mathbf{X}_i, \mathbf{Z}_i)$ can be taken to be the $(r.s, t.u)$th element of the information matrix for the distinct elements of $\mathbf{G}$, $\mathbf{I}_{\mathbf{G}}(\mathbf{X}_i, \mathbf{Z}_i)$. Similarly, the term $\mathbf{I}_{g_{rs},\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i)$ can be taken to be the $(r.s, 1)$th element of the joint information matrix for $\sigma_e^2$ and $\mathbf{G}$, $\mathbf{I}_{\mathbf{G},\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i)$.

Now, using expressions (2.17), (2.20) and (2.21), the information matrix for $\boldsymbol{\theta} = (\sigma_e^2, g_{11}, g_{21}, \ldots, g_{qq})$, i.e., for $\sigma_e^2$ and the $\dfrac{q\,(q+1)}{2}$ distinct elements of $\mathbf{G}$ for the $i$th individual is given by

$$\mathbf{I}_{\theta}(\mathbf{X}_i, \mathbf{Z}_i) = \begin{pmatrix} \mathbf{I}_{\sigma_e^2,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) & \mathbf{I}_{\mathbf{G},\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) \\ \mathbf{I}_{\mathbf{G},\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) & \mathbf{I}_{\mathbf{G}}(\mathbf{X}_i, \mathbf{Z}_i) \end{pmatrix}, \quad i = 1, \ldots, K. \tag{2.22}$$

Then the information matrix for $\sigma_e^2$ and $\mathbf{G}$ over all individuals is equal to

$$\mathbf{I}_{\theta}(\mathbf{X}, \mathbf{Z}) = \sum_{i=1}^{K} \mathbf{I}_{\theta}(\mathbf{X}_i, \mathbf{Z}_i).$$

Case 2: $\mathbf{V}_i = \sigma_e^2\,(\mathbf{Z}_i\, \mathbf{G}^*\, \mathbf{Z}'_i + \mathbf{I}_{d_i})$ where $\mathbf{G}^* = \dfrac{1}{\sigma_e^2}\,\mathbf{G}$

Let $g^*_{rs}$ be the $(r, s)$th element of $\mathbf{G}^*$. Then

$$\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} = \mathbf{Z}_i\,\mathbf{G}^*\,\mathbf{Z}_i' + \mathbf{I}_{d_i}$$

and

$$\frac{\partial \mathbf{V}_i}{\partial g^*_{rs}} = \sigma_e^2\,\mathbf{Z}_i\,\frac{\partial \mathbf{G}^*}{\partial g^*_{rs}}\,\mathbf{Z}_i.$$

So, from the result in expression (2.18) it follows that

$$\frac{\partial \mathbf{V}_i}{\partial g^*_{rs}} = \frac{\sigma_e^2}{2^{\delta_{rs}}}\left\{\mathbf{z}_{i,(r)}\mathbf{z}'_{i,(s)} + \mathbf{z}_{i,(s)}\mathbf{z}'_{i,(r)}\right\}. \tag{2.23}$$

Now, using expression (2.14)

$$\mathbf{I}_{\sigma_e^2,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2}tr\left\{\mathbf{V}_i^{-2}\,(\mathbf{Z}_i\,\mathbf{G}^*\,\mathbf{Z}_i' + \mathbf{I}_{d_i})^2\right\} = \frac{1}{2\,\sigma_e^4}. \tag{2.24}$$

Furthermore, from expression (2.14), (2.20) and (2.23) it follows that

$$\mathbf{I}_{g^*_{rs},g^*_{tu}}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{\sigma_e^4}{2^{\delta_{rs}} \times 2^{\delta_{tu}}}\left\{C_{i,ru}\,C_{i,st} + C_{i,su}\,C_{i,rt}\right\} = \sigma_e^4\,\mathbf{I}_{g_{rs},g_{tu}}(\mathbf{X}_i, \mathbf{Z}_i) \tag{2.25}$$

and

$$\begin{aligned}\mathbf{I}_{g^*_{rs},\sigma_e^2}(\mathbf{Z}_i) &= \frac{\sigma_e^2}{2 \times 2^{\delta_{rs}}}\left\{\mathbf{V}_i^{-1}\,(\mathbf{Z}_i\,\mathbf{G}^*\,\mathbf{Z}_i' + \mathbf{I}_{d_i})\left(\mathbf{z}_{i,(r)}\,\mathbf{V}_i^{-1}\,\mathbf{z}'_{i,(s)} + \mathbf{z}_{i,(s)}\,\mathbf{V}^{-1}\,\mathbf{z}'_{i,(r)}\right)\right\}, \\ &= \frac{\sigma_e^2}{2^{\delta_{rs}}}\,\mathbf{V}_i^{-1}\,(\mathbf{Z}_i\,\mathbf{G}^*\,\mathbf{Z}_i' + \mathbf{I}_{d_i})\,C_{i,rs} = \frac{1}{2^{\delta_{rs}}}\,C_{i,rs}\end{aligned}$$

Now using the Mentré *et al.* (1997) notation, the term $\mathbf{I}_{g^*_{rs},g^*_{tu}}(\mathbf{X}_i, \mathbf{Z}_i)$ can be taken to be the $(r.s, t.u)$th element of the information matrix for the distinct elements of $\mathbf{G}^*$, $\mathbf{I}_{\mathbf{G}^*}(\mathbf{X}_i, \mathbf{Z}_i)$. Similarly, the term $\mathbf{I}_{g^*_{rs},\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i)$ can be taken to be the $(r.s, 1)$th element of the joint information matrix for $\sigma_e^2$ and $\mathbf{G}^*$, $\mathbf{I}_{\mathbf{G}^*,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i)$. Thus, the information matrix for

$\boldsymbol{\theta}^* = (\sigma_e^2, g_{11}^*, g_{21}^*, \ldots, g_{qq}^*)$ for the $i$th individual, is given by

$$\mathbf{I}_{\theta^*}(\mathbf{X}_i, \mathbf{Z}_i) = \begin{pmatrix} \mathbf{I}_{\sigma_e^2,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) & \mathbf{I}_{\mathbf{G}^*,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) \\ \mathbf{I}_{\mathbf{G}^*,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) & \mathbf{I}_{\mathbf{G}^*}(\mathbf{X}_i, \mathbf{Z}_i) \end{pmatrix}, \; i = 1, \ldots, K \tag{2.26}$$

and the over all information matrix for $\sigma_e^2$ and $\mathbf{G}^*$ is equal to

$$\mathbf{I}_{\theta^*}(\mathbf{X}, \mathbf{Z}) = \sum_{i=1}^{K} \mathbf{I}_{\theta^*}(\mathbf{X}_i, \mathbf{Z}_i).$$

## 2.5 Random coefficient regression models

As mentioned in the introduction section of this chapter, this thesis focuses on the use of the linear mixed model for longitudinal data. Furthermore, it has been noted in Section 2.2 that such data have two sources of variation, within- and between-individuals. This variation can be introduced into a random coefficient regression model, which is a regression model with some or all of the parameters considered to be random effects. The random coefficient regression model can be written in terms of the linear mixed model as

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i,$$

where $\mathbf{Z}_i = \mathbf{X}_i$ or $\mathbf{Z}_i$ comprises a subset of the columns of $\mathbf{X}_i$ with this subset the same for all $i$, $i = 1, \ldots, K$. In the following subsections two cases are considered, first the special case of the random intercept model and then the more general case of the random coefficient model. Both of these models are considered in this thesis but particular attention is given to the random intercept model.

### 2.5.1 Random intercept model

The random intercept model is commonly used to describe longitudinal data. It is a special case of the random coefficient regression model with the first column of the design matrix $\mathbf{X}_i$ equal to $\mathbf{1}_{d_i}$ and $\mathbf{Z}_i = \mathbf{1}_{d_i}$, where $\mathbf{1}_{d_i}$ is the $d_i \times 1$ vector of 1's. Then under the linear mixed model formulation, the general random intercept model for the $i$th individual can be written as

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{1}_{d_i} b_i + \mathbf{e}_i \tag{2.27}$$

where the term $b_i$ is the random component of the intercept, and the other terms in the model are the same as those introduced in Section 2.2. The random terms $b_i$, $i = 1, \ldots, K$, are assumed to be independently normally distributed with mean zero and variance $\sigma_b^2$. Then the variance of $\mathbf{y}_i$ in (2.27) is given by

$$\mathbf{V}_i = \sigma_e^2 \mathbf{I}_{d_i} + \sigma_b^2 \mathbf{J}_{d_i} = \sigma_e^2(\mathbf{I}_{d_i} + \gamma \mathbf{J}_{d_i}) \tag{2.28}$$

where $\gamma = \dfrac{\sigma_b^2}{\sigma_e^2}$ is called the variance ratio (Longford, 1993, page 27) or the degree of correlation (Goos, 2002, page 79). The random intercept model in (2.27) is also called the compound symmetry model (Longford, 1993, page 28).

The observations within an individual are not independent because $\mathrm{Cov}(y_{ij}, y_{ij'}) = \sigma_b^2$ if $j \neq j'$. As $\mathrm{Var}(y_{ij}) = \sigma_b^2 + \sigma_e^2$, the observations within the same individual are correlated with intra-class correlation coefficient

$$\rho = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}.$$

The intra-class correlation measures the similarity between observations within an individual and it is also known as repeatability (Henderson *et al.*, 1959).

The information matrices for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ in the random intercept model can be obtained from the results for the linear mixed model. Recall from expression (2.10) that the information matrix $\mathbf{I}_\beta(\mathbf{X}_i, \mathbf{Z}_i)$ for the $i$th individual is given by

$$\mathbf{I}_\beta(\mathbf{X}_i, \mathbf{Z}_i) = \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i$$

for $i = 1, \ldots, K$. Let $\mathbf{X}_i = [\mathbf{1}_{d_i} \ \tilde{\mathbf{X}}_i]$, where $\tilde{\mathbf{X}}_i$ comprises the columns of $\mathbf{X}_i$ not corresponding to the intercept. Then the information matrix for the $i$th individual is given by

$$\mathbf{I}_\beta(\mathbf{X}_i, \mathbf{Z}_i) = \begin{pmatrix} \mathbf{1}_{d_i}' \\ \tilde{\mathbf{X}}_i' \end{pmatrix} \mathbf{V}_i^{-1} (\mathbf{1}_{d_i} \ \tilde{\mathbf{X}}_i) = \begin{pmatrix} \mathbf{1}_{d_i}'\mathbf{V}_i^{-1}\mathbf{1}_{d_i} & \mathbf{1}_{d_i}'\mathbf{V}_i^{-1}\tilde{\mathbf{X}}_i \\ \tilde{\mathbf{X}}_i'\mathbf{V}_i^{-1}\mathbf{1}_{d_i} & \tilde{\mathbf{X}}_i'\mathbf{V}_i^{-1}\tilde{\mathbf{X}}_i \end{pmatrix}.$$

Using Result A.2.1 from Appendix A with $\mathbf{V}_i$ specified in (2.28), it follows that

$$\mathbf{V}_i^{-1} = \frac{1}{\sigma_e^2}(\mathbf{I}_{d_i} - \frac{\gamma}{1 + d_i\gamma}\mathbf{J}_{d_i}).$$

Further

$$\mathbf{1}_{d_i}'\mathbf{V}_i^{-1}\mathbf{1}_{d_i} = \frac{1}{\sigma_e^2}\mathbf{1}_{d_i}'(\mathbf{I}_{d_i} - \frac{\gamma}{1 + d_i\gamma}\mathbf{J}_{d_i})\mathbf{1}_{d_i} = \frac{d_i}{\sigma_e^2(1 + d_i\gamma)},$$

$$\tilde{\mathbf{X}}_i'\mathbf{V}_i^{-1}\tilde{\mathbf{X}}_i = \frac{1}{\sigma_e^2}\tilde{\mathbf{X}}_i'(\mathbf{I}_{d_i} - \frac{\gamma}{1 + d_i\gamma}\mathbf{J}_{d_i})\tilde{\mathbf{X}}_i = \frac{1}{\sigma_e^2}\left[\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i - \frac{\gamma}{1 + d_i\gamma}\tilde{\mathbf{X}}_i'\mathbf{1}_{d_i}\mathbf{1}_{d_i}'\tilde{\mathbf{X}}_i\right]$$

and

$$\tilde{\mathbf{X}}_i'\mathbf{V}_i^{-1}\mathbf{1}_{d_i} = \frac{1}{\sigma_e^2}\tilde{\mathbf{X}}_i'(\mathbf{I}_{d_i} - \frac{\gamma}{1 + d_i\gamma}\mathbf{J}_{d_i})\mathbf{1}_{d_i} = \frac{\tilde{\mathbf{X}}_i'\mathbf{1}_{d_i}}{\sigma_e^2(1 + d_i\gamma)}.$$

Thus

$$\mathbf{I}_\beta(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{\sigma_e^2(1 + d_i\gamma)}\begin{pmatrix} d_i & \mathbf{1}_{d_i}'\tilde{\mathbf{X}}_i \\ \tilde{\mathbf{X}}_i'\mathbf{1}_{d_i} & (1 + d_i\gamma)\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i - \gamma\tilde{\mathbf{X}}_i'\mathbf{1}_{d_i}\mathbf{1}_{d_i}'\tilde{\mathbf{X}}_i \end{pmatrix} \tag{2.29}$$

for $i = 1, \ldots, K$. Clearly the overall information matrix for $\boldsymbol{\beta}$ is then given by

$$\mathbf{I}_\beta(\mathbf{X}, \mathbf{Z}) = \sum_{i=1}^{K}\mathbf{I}_\beta(\mathbf{X}_i, \mathbf{Z}_i).$$

Consider now the information matrix for $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$. Since $\mathbf{V}_i = \sigma_e^2 \mathbf{I}_i + \sigma_b^2 \mathbf{J}_i$ it follows that

$$\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} = \mathbf{I}_{d_i} \quad \text{and} \quad \frac{\partial \mathbf{V}_i}{\partial \sigma_b^2} = \mathbf{J}_{d_i}$$

and thus that from expression (2.14)

$$\mathbf{I}_{\sigma_e^2,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2} tr \left( \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} \right)^2 = \frac{1}{2\sigma_e^4} \left[ (d_i - 1) + \left( \frac{\sigma_e^2}{\sigma_e^2 + d_i \sigma_b^2} \right)^2 \right],$$

$$\mathbf{I}_{\sigma_b^2,\sigma_b^2}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2} tr \left( \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_b^2} \right)^2 = \frac{d_i^2}{2\,(\sigma_e^2 + d_i \sigma_b^2)^2}$$

and

$$\mathbf{I}_{\sigma_e^2,\sigma_b^2}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2} tr \left( \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_b^2} \right) = \frac{d_i}{2\,(\sigma_e^2 + d_i \sigma_b^2)^2}$$

Therefore the information matrix for $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$ from the $i$th individual is given by

$$\mathbf{I}_\theta(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2} \begin{pmatrix} \frac{1}{\sigma_e^4} \left[ (d_i - 1) + \left( \frac{\sigma_e^2}{\sigma_e^2 + d_i \sigma_b^2} \right)^2 \right] & \frac{d_i}{(\sigma_e^2 + d_i \sigma_b^2)^2} \\ \frac{d_i}{(\sigma_e^2 + d_i \sigma_b^2)^2} & \frac{d_i^2}{(\sigma_e^2 + d_i \sigma_b^2)^2} \end{pmatrix} \tag{2.30}$$

and this can be written in terms of $\sigma_e^2$ and the variance ratio $\gamma = \dfrac{\sigma_b^2}{\sigma_e^2}$ as

$$\mathbf{I}_\theta(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2\,\sigma_e^4} \begin{pmatrix} (d_i - 1) + \frac{1}{(1 + d_i \gamma)^2} & \frac{d_i}{(1 + d_i \gamma)^2} \\ \frac{d_i}{(1 + d_i \gamma)^2} & \frac{d_i^2}{(1 + d_i \gamma)^2} \end{pmatrix} \tag{2.31}$$

for $i = 1, \ldots, K$. Clearly the overall information matrix is given by

$$\mathbf{I}_\theta(\mathbf{X}, \mathbf{Z}) = \sum_{i=1}^{K} \mathbf{I}_\theta(\mathbf{X}_i, \mathbf{Z}_i).$$

Now consider the alternative parametrization of the variance components with $\boldsymbol{\theta} = (\sigma_e^2, \gamma)$ where $\gamma = \dfrac{\sigma_b^2}{\sigma_e^2}$. Then the variance-covariance matrix of the response for the $i$th individual $\mathbf{y}_i$ is given by expression (2.28), that is by

$$\mathbf{V}_i = \sigma_e^2(\mathbf{I}_{d_i} + \gamma \mathbf{J}_{d_i}).$$

Hence

$$\mathbf{V}_i^{-1} = \frac{1}{\sigma_e^2}(\mathbf{I}_{d_i} - \frac{\gamma}{1 + d_i\,\gamma}\mathbf{J}_{d_i}).$$

Moreover

$$\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} = \frac{1}{\sigma_e^2}\mathbf{V}_i = \mathbf{I}_{d_i} + \gamma \mathbf{J}_{d_i} \text{ and } \frac{\partial \mathbf{V}_i}{\partial \gamma} = \sigma_e^2 \mathbf{J}_{d_i}.$$

Thus

$$\mathbf{I}_{\sigma_e^2,\sigma_e^2}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2}\, tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2}\right)^2 = \frac{d_i}{2\sigma_e^4},$$
$$\mathbf{I}_{\gamma,\gamma}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2} tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \gamma}\right)^2 = \frac{d_i^2}{2\,(1 + d_i\gamma)^2}$$

and

$$\mathbf{I}_{\sigma_e^2,\gamma}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2} tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \gamma}\right) = \frac{d_i}{2\,\sigma_e^2\,(1 + d_i\gamma)}.$$

Thus the information matrix for $\boldsymbol{\theta} = (\sigma_e^2, \gamma)$ from the $i$th individual is given by

$$\mathbf{I}_\theta(\mathbf{X}_i, \mathbf{Z}_i) = \frac{d_i}{2}\begin{pmatrix} \dfrac{1}{\sigma_e^4} & \dfrac{1}{\sigma_e^2\,(1 + d_i\gamma)} \\ \\ \dfrac{1}{\sigma_e^2\,(1 + d_i\gamma)} & \dfrac{d_i}{(1 + d_i\gamma)^2} \end{pmatrix} \tag{2.32}$$

and the overall information matrix is equal to

$$\mathbf{I}_\theta(\mathbf{X}, \mathbf{Z}) = \sum_{i=1}^{K} \mathbf{I}_\theta(\mathbf{X}_i, \mathbf{Z}_i).$$

### 2.5.2 Random coefficient model

Consider the case when all parameters in the regression model are random, i.e. $\mathbf{Z}_i = \mathbf{X}_i$, for $i = 1, \ldots, K$. Then the linear mixed model can be rewritten as

$$\mathbf{y}_i = \mathbf{X}_i\,(\boldsymbol{\beta} + \mathbf{b}_i) + \mathbf{e}_i = \mathbf{X}_i\,\boldsymbol{\beta}_i + \mathbf{e}_i,$$

where $\boldsymbol{\beta}_i \sim \mathcal{N}(\boldsymbol{\beta}, \mathrm{G})$. Under the normality and independence assumptions relating to $\mathbf{b}_i$ and $\mathbf{e}_i$, the marginal distribution of $\mathbf{y}_i$ is given by $\mathbf{y}_i \sim \mathcal{N}(\mathbf{X}_i\,\boldsymbol{\beta}, \mathbf{V}_i)$ where $\mathbf{V}_i = \mathbf{X}_i\,\mathrm{G}\,\mathbf{X}_i' + \sigma_e^2\,\mathbf{I}_{d_i}$. Therefore by using Result A.2.4 from Appendix A the information matrix for $\boldsymbol{\beta}$ can be expressed as

$$\begin{aligned}\mathbf{I}_{\beta}(\mathbf{X}_i, \mathbf{Z}_i) &= \mathbf{X}_i'\left(\mathbf{X}_i\,\mathrm{G}\,\mathbf{X}_i' + \sigma_e^2\,\mathbf{I}_{d_i}\right)^{-1}\mathbf{X}_i = \left\{\mathbf{X}_i'\,(\sigma_e^2\,\mathbf{I}_{d_i})^{-1}\,\mathbf{X}_i + \mathrm{G}\right\}^{-1} \\ &= \frac{1}{\sigma_e^2}\left\{(\mathbf{X}_i'\,\mathbf{X}_i)^{-1} + \mathrm{G}_{\gamma}\right\}^{-1} \qquad (2.33)\end{aligned}$$

where $\mathrm{G}_{\gamma} = \dfrac{1}{\sigma_e^2}\mathrm{G}$. Hence

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \sigma_e^2\left\{\sum_{i=1}^{K}\left[(\mathbf{X}_i'\,\mathbf{X}_i)^{-1} + \mathrm{G}_{\gamma}\right]^{-1}\right\}^{-1}.$$

Furthermore, when $\mathbf{X}_i = \mathbf{Z}_i$ the $(r.s, t.u)$th element of the information matrix for G from the $i$th individual, $\mathbf{I}_{\mathrm{G}}(\mathbf{X}_i, \mathbf{Z}_i)$ in expression (2.20), becomes

$$\begin{aligned}(\mathbf{I}_{\mathrm{G}})_{r.s,t.u}(\mathbf{X}_i, \mathbf{Z}_i) = \frac{1}{2^{\delta_{rs}} \times 2^{\delta_{tu}}}&\left\{(\mathbf{x}_{i,(r)}'\,\mathbf{V}_i^{-1}\,\mathbf{x}_{i,(u)})\,(\mathbf{x}_{i,(s)}'\,\mathbf{V}_i^{-1}\,\mathbf{x}_{i,(t)})\right. \\ &\left. + (\mathbf{x}_{i,(s)}'\,\mathbf{V}_i^{-1}\,\mathbf{x}_{i,(u)})\,(\mathbf{x}_{i,(r)}'\,\mathbf{V}_i^{-1}\,\mathbf{x}_{i,(t)})\right\}, \qquad (2.34)\end{aligned}$$

where $\mathbf{x}_{i,(u)}'\mathbf{V}_i^{-1}\mathbf{x}_{i,(r)}$ is the $(r, u)$th element of $\mathbf{I}_{\beta}(\mathbf{X}_i, \mathbf{Z}_i)$ (Mentré *et al.*, 1997).

Consider now the case where $\mathbf{Z}_i$ comprises a subset of the columns of $\mathbf{X}_i$, i.e. when a subset of the coefficients $\boldsymbol{\beta}$ are random. Let $\boldsymbol{\beta} = [\boldsymbol{\beta}_1\ \boldsymbol{\beta}_2]'$, where $\boldsymbol{\beta}_1$ is a vector of fixed effects and $\boldsymbol{\beta}_2$ is a vector of random effects and let the matrix $\mathbf{X}_i$ be conformably partitioned as $\mathbf{X}_i = [\mathbf{X}_{1i}\ \mathbf{X}_{2i}]$ where $\mathbf{X}_{2i} = \mathbf{Z}_i$, $i = 1, 2, \ldots, K$. Then

$$\mathbf{y}_i = \mathbf{X}_{1i}\boldsymbol{\beta}_1 + \mathbf{X}_{2i}\left(\boldsymbol{\beta}_2 + \mathbf{b}_i\right) + \mathbf{e}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{e}_i,$$

where

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \begin{pmatrix} 0 \\ \mathbf{b}_i \end{pmatrix}.$$

Suppose that the variance-covariance matrix of $\mathbf{b}_i$ is $\mathbf{G}_2$. Then, under the assumption that

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \begin{pmatrix} 0 \\ \mathbf{b}_i \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\beta}, \mathbf{G})$$

where

$$\mathbf{G} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{G}_2 \end{pmatrix},$$

the distribution of the response $\mathbf{y}_i$ for the $i$th individual is normal with mean $\mathbf{X}_i\boldsymbol{\beta}$ and variance-covariance matrix $\mathbf{V}_i = \mathbf{X}_i\,\mathbf{G}\,\mathbf{X}_i' + \sigma_e^2\,\mathbf{I}_{d_i} = \mathbf{X}_{2i}\,\mathbf{G}_2\,\mathbf{X}_{2i}' + \sigma_e^2\,\mathbf{I}_{d_i}$. Therefore the information matrix for $\boldsymbol{\beta}$ from the $i$th individual can be written as

$$\begin{aligned}
\mathbf{I}_\beta(\mathbf{X}_i, \mathbf{Z}_i) &= \mathbf{X}_i'\left(\mathbf{X}_i\,\mathbf{G}\,\mathbf{X}_i' + \sigma_e^2\,\mathbf{I}_{d_i}\right)^{-1}\mathbf{X}_i = \left[\sigma_e^2\,(\mathbf{X}_i'\mathbf{X}_i)^{-1} + \mathbf{G}\right]^{-1} \\
&= \frac{1}{\sigma_e^2}\left[(\mathbf{X}_i'\mathbf{X}_i)^{-1} + \mathbf{G}_\gamma\right]^{-1} = \frac{1}{\sigma_e^2}\left\{(\mathbf{X}_i'\mathbf{X}_i)^{-1} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{G}_{2,\gamma} \end{pmatrix}\right\}^{-1}
\end{aligned}$$

where

$$\mathbf{G}_\gamma = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{G}_{2,\gamma} \end{pmatrix} \quad \text{and} \quad \mathbf{G}_{2,\gamma} = \frac{1}{\sigma_e^2}\mathbf{G}_2.$$

The information matrix for the variance components $\boldsymbol{\theta}$ can be derived from the results of Subsection 2.4.4, specifically from expression (2.26).

### 2.5.3 Invariance to linear transformation of $\mathbf{X}_i$

Consider the random coefficient model with $\mathbf{Z}_i = \mathbf{X}_i$, which can be written as

$$\mathbf{y}_i = \mathbf{X}_i\,\boldsymbol{\beta} + \mathbf{X}_i\,\mathbf{b}_i + \mathbf{e}_i,$$

where $\mathbf{b}_i \sim \mathcal{N}(0, \mathbf{G})$, $\mathbf{e}_i \sim \mathcal{N}(0, \sigma_e^2\,\mathbf{I})$ and $\mathbf{b}_i$ and $\mathbf{e}_i$ are independent, $i = 1, 2, \ldots, K$. Suppose that the columns of the design matrix $\mathbf{X}_i$ are linearly transformed as

$$\mathbf{X}_i^* = \mathbf{X}_i\,\mathbf{A}, \quad i = 1, \ldots, K$$

where $\mathbf{A}$ is a $p \times p$ nonsingular matrix. Then the transformed model can be written as

$$\mathbf{y}_i = \mathbf{X}_i^*\,\boldsymbol{\beta}^* + \mathbf{X}_i^*\,\mathbf{b}_i^* + \mathbf{e}_i,$$

where $\boldsymbol{\beta}^* = \mathbf{A}^{-1}\,\boldsymbol{\beta}$ and $\mathbf{b}_i^* = \mathbf{A}^{-1}\,\mathbf{b}_i$ with $\mathbf{b}_i^* \sim \mathcal{N}(0, \mathbf{G}^*)$ and $\mathbf{G}^* = \mathbf{A}^{-1}\,\mathbf{G}\,(\mathbf{A}^{-1})'$. Thus a linear transformation of the columns of $\mathbf{X}_i$ induces the transformation $\boldsymbol{\beta}^* = \mathbf{A}^{-1}\,\boldsymbol{\beta}$ in the fixed effects and $\mathbf{b}_i^* = \mathbf{A}^{-1}\,\mathbf{b}_i$ in the random effects. More particularly the structure of $\mathbf{G}$ after transformation depends on $\mathbf{A}$ and is very often not preserved (Longford 1993, pages 98-93).

For example, consider the simple linear regression model with $\mathbf{X} = [\mathbf{1}\ \ \mathbf{x}]$ where $\mathbf{x}$ is a column vector of observations $x$ and suppose that $x^* = a + b\,x$, where $a$ and $b$ are nonzero constants. Then

$$\mathbf{X}^* = \mathbf{X}\begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}, \quad \text{i.e. } \mathbf{A} = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}.$$

Let

$$\mathbf{G} = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix}.$$

Since

$$\mathbf{A}^{-1} = \begin{pmatrix} 1 & -a/b \\ 0 & 1/b \end{pmatrix}$$

it follows that

$$\mathbf{G}^* = \mathbf{A}^{-1}\,\mathbf{G}\,(\mathbf{A}^{-1})' = \begin{pmatrix} g_{11} - \frac{2\,a}{b}\,g_{12} + \frac{a^2}{b^2}\,g_{22} & \frac{g_{12}}{b} - \frac{a}{b^2}\,g_{22} \\ \frac{g_{12}}{b} - \frac{a}{b^2}\,g_{22} & \frac{g_{22}}{b^2} \end{pmatrix}.$$

Consider now the following cases.

(i) For the random intercept model with $g_{12} = g_{22} = 0$,

$$\mathbf{G} = \begin{pmatrix} g_{11} & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{G}^*$$

and thus the structure of $\mathbf{G}$ is preserved.

(ii) If $g_{11} = g_{12} = 0$, i.e. when the model has a random slope, then

$$\mathbf{G} = \begin{pmatrix} 0 & 0 \\ 0 & g_{22} \end{pmatrix} \quad \text{but} \quad \mathbf{G}^* = \frac{g_{22}}{b^2} \begin{pmatrix} a^2 & -a \\ -a & b \end{pmatrix}.$$

In this case the structure of $\mathbf{G}$ is totally changed. That is, for the random slope model the linear transformation may result in a different structure for $\mathbf{G}$ unless $a = 0$.

Note more generally that for the random coefficient model, the variance-covariance matrix of the response for the $i$th individual $\mathbf{y}_i$ does not change due to a linear transformation of the columns of $\mathbf{X}_i$. That is, $\mathbf{V}_i^* = \mathbf{X}_i^*\,\mathbf{G}^*\,(\mathbf{X}^*)' + \sigma_e^2\,\mathbf{I} = \mathbf{V}_i$. Therefore the information matrices

$$\mathbf{I}_{\beta^*}(\mathbf{X}_i^*) = (\mathbf{X}_i^*)'\,(\mathbf{V}_i^*)^{-1}\,\mathbf{X}_i^*$$

and

$$\mathbf{I}_\beta(\mathbf{X}_i) = \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i$$

are closely related through $\mathbf{I}_{\beta^*}(\mathbf{X}_i^*) = \mathbf{A}' \mathbf{I}_\beta(\mathbf{X}_i)\, \mathbf{A}$. However, due to the change that may occur to the variance-covariance matrix of the random effects on a linear transformation of the explanatory variables, the nature of the underlying model may change.

## 2.6 Optimum design for the linear mixed model

Methods of estimation for the fixed effects $\boldsymbol{\beta}$ and the variance components $\boldsymbol{\theta}$ for the linear mixed model have been discussed in Section 2.3. These methods are well-known and have been widely treated in the literature (see for example Verbeke and Molenberghs, 1999). However, few studies relating to the construction of optimal designs for the precise estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ have been reported. The goal of this thesis is to investigate such designs and the necessary background to the theory of optimal design for the linear mixed model is now presented in this section. Much of the material is based on the seminal papers of Mallet and Mentré (1988), Mentré *et al* (1995), Cheng (1995), Lohr (1995), Mentré *et al* (1997), Abt, Liski, Mandal and Sinha (1997), Abt, Gaffke, Liski and Sinha (1998), and Atkins and Cheng (1999).

### 2.6.1 Design problem

In optimal experimental design the general objective is to select values of the control variables so that the quantities of interest are estimated as precisely as possible. For the linear mixed model described in Section 2.2 attention focusses on the fixed effects $\boldsymbol{\beta}$ and the variance components $\boldsymbol{\theta}$. Thus designs which in some sense maximize the information on these parameters and which are constructed by an appropriate choice of the design matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$ for $i = 1, \ldots, K$ are sought. Note that for the random coefficient model the design matrix $\mathbf{Z}_i$ comprises a subset of the columns of $\mathbf{X}_i$ and thus the design problem reduces to a consideration of the design matrices $\mathbf{X}_i$, $i = 1, \ldots, K$. In this thesis only random intercept and random coefficient models are considered.

### 2.6.2 Exact and approximate population designs

Suppose that repeated measurements are taken on a group of $n$ individuals either at different times or on different conditions or different factor levels. Suppose further that the group can be divided into $r$ cohorts and that the $i$th cohort comprises $n_i$ individuals each with design matrix $\mathbf{X}_i$ corresponding to $d_i$ observations. Note that the total number of individuals is given by $n = \sum_{i=1}^{r} n_i$. The matrices $\mathbf{X}_i$ are distinct with dimension $d_i \times p$ where $p$ is the number of parameters in the parameters vector of interest, $\boldsymbol{\alpha}$. This specification defines an exact population design, $\xi_n$, which can be conveniently summarized as

$$\xi_n = \left\{ \begin{array}{ccc} \mathbf{X}_1, & \ldots, & \mathbf{X}_r \\ n_1, & \ldots, & n_r \end{array} \right\}.$$

The design matrices $\mathbf{X}_i$, $i = 1, \ldots, r$, which comprise the support of the population design, are taken from a set $S$ of all possible such matrices. The set $S$ can be finite, specifically $S = \{\mathbf{X}_1, \ldots, \mathbf{X}_K\}$, or infinite. Note also that $\sum_{i=1}^{r} n_i \, d_i$ observations are taken in total. Now using $\mathbf{I}_\alpha(\mathbf{X}_i)$ as the Fisher information matrix for cohort $i$ and dropping the dependence on $\mathbf{Z}_i$, the information matrix for the population design $\xi_n$ can be expressed as

$$\mathbf{I}_\alpha(\xi_n) = \sum_{i=1}^{r} n_i \mathbf{I}_\alpha(\mathbf{X}_i).$$

In the linear mixed model the parameter vector $\boldsymbol{\alpha}$ corresponds to $\boldsymbol{\beta}$ or $\boldsymbol{\theta}$ or both and $\mathbf{I}_\alpha(\mathbf{X}_i)$ to the associated information matrix.

An approximate population design is a probability measure $\xi$ on the space $S$ which assigns a weight $v_i$ to the $i$th individual design matrix $\mathbf{X}_i$, where $0 < v_i < 1$ and $\sum_{i=1}^r v_i = 1$, that is

$$\xi = \left\{ \begin{array}{ccc} \mathbf{X}_1, & \ldots, & \mathbf{X}_r \\ v_1, & \ldots, & v_r \end{array} \right\}, \quad 0 < v_i < 1 \text{ with } \sum_{i=1}^{r} v_i = 1.$$

It is often convenient to express the weights associated with an approximate design $\xi$ on a per observation rather than a per individual (or per cohort) basis. Specifically let $w_i$ be the weight per observation for the $i$th individual design matrix $\mathbf{X}_i$, $i = 1, \ldots, r$. Then, since there are $d_i$ observations associated with that individual, the weights $v_i$ are given by

$$v_i = \frac{w_i/d_i}{\sum_{i=1}^r w_i/d_i}$$

and conversely

$$w_i = \frac{v_i\, d_i}{\sum_{i=1}^r v_i\, d_i}$$

for $i = 1, \ldots, r$. The approximate design can thus be expressed as

$$\xi = \left\{ \begin{array}{ccc} \mathbf{X}_1, & \ldots, & \mathbf{X}_r \\ w_1, & \ldots, & w_r \end{array} \right\}, \quad 0 < w_i < 1 \text{ with } \sum_{i=1}^{r} w_i = 1.$$

The information matrix corresponding to $\mathbf{X}_i$ can be expressed on a per observation basis as

$$\mathbf{M}_\alpha(\mathbf{X}_i) = \frac{\mathbf{I}_\alpha(\mathbf{X}_i)}{d_i}$$

and is often referred to in this form as the standardized information matrix. Then the information matrix for the design $\xi$ on a per observation basis is given by

$$\mathbf{M}_\alpha(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\alpha(\mathbf{X}_i).$$

This formulation involving the weights per observation, $w_i, i = 1, \ldots, r$, is used throughout this thesis.

In the above discussions an approximate design was taken to be a probability measure on the set of design matrices $S$ which has finite support. This can be justified more generally as follows. Let $\mathcal{M}$ be the set of all information matrices of the form $\mathbf{M}_\alpha(\mathbf{X}_i)$ for $\mathbf{X}_i \in S$. This set is compact and convex (see Theorem 2.1.2, Fedorov, 1972, page 66). Then by Carathéodory's Theorem any information matrix can be represented as a weighted sum of $m$ information matrices where $m \leq \dfrac{p\,(p+1)}{2} + 1$ with $p$ the number of parameters in the vector $\boldsymbol{\alpha}$ (see Silvey, 1980, page 72). Thus, for any approximate population design $\xi$, there exists another approximate population design $\tilde{\xi}$ with finite support such that $\mathbf{M}_\alpha(\xi) = \mathbf{M}_\alpha(\tilde{\xi})$ (Mentré *et al*, 1997).

Once an approximate optimum design has been constructed there can be problems with its implementation in practice. Specifically the weights associated with the designs do not necessarily translate to integer numbers of individuals in a cohort and thus to integer numbers of observations. In other words, if there are $n$ individuals in total in an experiment, then the allocation of $n\,v_i$ individuals to cohort $i$ for $i = 1, \ldots, K$ may not be integer. In many studies the near-optimal exact design is obtained by using the closest integer assignment to the approximate allocation but this may not be optimal in the exact design sense. However, Cook and Nachtsheim (1980) and more recently Pukelsheim and Rieder

(1992) and Pukelsheim (1993, pages 325-327) have developed algorithms for constructing exact designs from given approximate optimum designs which are highly efficient.

### 2.6.3 Optimality criteria

Consider the linear mixed model with parameters $\boldsymbol{\beta}$ or $\boldsymbol{\theta}$ or both $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ corresponding to the generic parameter $\boldsymbol{\alpha}$. The aim of optimal experimental design is to in some sense maximize information on the parameters. Since it is not possible to maximize the information matrix itself, at least for $p > 1$, it is usual to consider a convex function of the information matrix $\mathbf{M}_\alpha(\xi)$, $\Psi\{\mathbf{M}_\alpha(\xi)\}$. Such a function $\Psi$ is called an optimality criterion. If it does not lead to ambiguity $\Psi(\xi)$ will be used to denote $\Psi\{\mathbf{M}_\alpha(\xi)\}$. There are a number of such criteria and those of $D$- and of $L$-optimality are discussed, within the context of the linear mixed model, in the next two subsections.

It is relevant at this point to introduce the notion of the directional derivative of a convex function $\Psi(\xi)$. Specifically let the design $\xi_X$ represent the design which puts unit mass at a single design matrix $\mathbf{X} \in S$ and let the design $\xi'$ be given by $\xi' = (1-\epsilon)\xi + \epsilon\,\xi_X$ for any $\epsilon$ with $0 \leq \epsilon \leq 1$. Then the derivative of the criterion $\Psi(\xi)$ at $\xi$ in the direction of $\xi_X$, is given by

$$\phi(\xi_X, \xi) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}\left[\Psi\{(1-\epsilon)\mathbf{M}_\alpha(\xi) + \epsilon\,\mathbf{M}_\alpha(\xi_X)\} - \{\Psi[\mathbf{M}_\alpha(\xi)]\}\right]. \qquad (2.35)$$

This directional derivative provides a powerful tool in the construction of optimal designs.

### 2.6.4 $D$-optimality

One of the best known and most widely used optimality criteria is that of $D$-optimality. In the context of the linear mixed model this criterion is specified as the determinant of the population information matrix, $|\mathbf{M}_\alpha(\xi)|$, where $\mathbf{M}_\alpha(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\alpha(\mathbf{X}_i)$. Maximization of $|\mathbf{M}_\alpha(\xi)|$ is equivalent to maximization of $\ln|\mathbf{M}_\alpha(\xi)|$ and the design $\xi^*$ is said to be $D$-optimal if it maximizes the function

$$\Psi(\xi) = \ln|\mathbf{M}_\alpha(\xi)|\,. \tag{2.36}$$

The General Equivalence Theorem for $D$-optimality presented in the seminal paper of Kiefer and Wolfowitz (1959) relates to the standard optimal designs for regression models but can be modified to accommodate population designs. Specifically suppose that all the information matrices $\mathbf{M}_\alpha(\mathbf{X}_i)$, where $\mathbf{X}_i \in S$, are nonsingular. Then the following lemma is a necessary precursor to the formulation of the Equivalence Theorem for population designs and is based on Lemmas 2.2.2 and 2.2.3 in Fedorov (1972, page 71).

**Lemma 2.6.1** *The function* $\ln|\mathbf{M}_\alpha(\xi)|$ *is concave on the set of information matrices* $\mathcal{M}$*. Furthermore, for a population design* $\xi$*, the directional derivative of* $\Psi(\xi) = \ln|\mathbf{M}_\alpha(\xi)|$ *at* $\xi$ *in the direction of* $\xi_{\mathbf{x}}$ *is given by*

$$\phi(\xi_{\mathbf{x}}, \xi) = tr\{\mathbf{M}_\alpha(\xi)^{-1}\mathbf{M}_\alpha(\xi_{\mathbf{x}})\} - p,$$

*where* $p$ *is the number of parameters specified in the vector* $\boldsymbol{\alpha}$*.*

**Proof**

It follows from the convexity of the set of information matrices that a design $\xi = (1-\epsilon)\,\xi_1 + \epsilon\,\xi_{\mathbf{x}}, 0 < \epsilon < 1$, has information matrix

$$\mathbf{M}(\xi) = (1-\epsilon)\,\mathbf{M}_\alpha(\xi_1) + \epsilon\,\mathbf{M}_\alpha(\xi_{\mathbf{x}}). \tag{2.37}$$

Thus from Result A.2.2 of Appendix A it follows that

$$|\mathbf{M}_\alpha(\xi)| \geq |\mathbf{M}_\alpha(\xi_1)|^{1-\epsilon}\,|\mathbf{M}_\alpha(\xi_{\mathbf{x}})|^{\epsilon}$$

and hence that

$$\ln|\mathbf{M}_\alpha(\xi)| \geq (1-\epsilon)\,\ln|\mathbf{M}_\alpha(\xi_1)| + \epsilon\,\ln|\mathbf{M}_\alpha(\xi_{\mathbf{x}})|\,.$$

Thus $\ln|\mathbf{M}_\alpha(\xi)|$ is concave on the set $\mathcal{M}$.

To prove the second part of the lemma, take the logarithm of the determinant of the information matrix specified in (2.37) as

$$\ln|\mathbf{M}_\alpha(\xi)| = \ln|(1-\epsilon)\,\mathbf{M}_\alpha(\xi_1) + \epsilon\,\mathbf{M}_\alpha(\xi_{\mathbf{x}})|\,.$$

Differentiating with respect to $\epsilon$ yields

$$\begin{aligned}\frac{\partial}{\partial\epsilon}\ln|\mathbf{M}_\alpha(\xi)| &= tr\{[(1-\epsilon)\,\mathbf{M}_\alpha(\xi_1) + \epsilon\mathbf{M}_\alpha(\xi_{\mathbf{x}})]^{-1}\frac{\partial}{\partial\epsilon}[(1-\epsilon)\,\mathbf{M}_\alpha(\xi_1) + \epsilon\mathbf{M}_\alpha(\xi_{\mathbf{x}})]\} \\ &= tr\{\mathbf{M}_\alpha(\xi)^{-1}[\mathbf{M}_\alpha(\xi_{\mathbf{x}}) - \mathbf{M}_\alpha(\xi_1)]\}.\end{aligned}$$

Then from Result A.4.1 of Appendix A and from the fact that $\mathbf{M}_\alpha(\xi) = \mathbf{M}_\alpha(\xi_1)$ when $\epsilon = 0$ it follows that

$$\phi(\xi_{\mathbf{x}}, \xi) = \lim_{\epsilon\to 0}\frac{\partial}{\partial\epsilon}\ln|\mathbf{M}_\alpha(\xi)| = tr\{\mathbf{M}_\alpha(\xi)^{-1}\mathbf{M}_\alpha(\xi_{\mathbf{x}})\} - p.$$

□

The General Equivalence Theorem for $D$-optimal population designs is now presented.

**Theorem 2.6.1** *The following three conditions on the D-optimal population design $\xi^*$ are equivalent:*

1. *The design $\xi^*$ maximizes* $\ln|\mathbf{M}_\alpha(\xi)|$.
2. *The design $\xi^*$ minimizes* $\max_{\mathbf{X}\in S} tr\{\mathbf{M}_\alpha^{-1}(\xi)\,\mathbf{M}_\alpha(\mathbf{X})\}$.
3. *The support of $\xi^*$ is contained in the set of design matrices, $\mathbf{X}$, such that*

$$tr\{\mathbf{M}_\alpha^{-1}(\xi^*)\mathbf{M}_\alpha(\mathbf{X})\} = p$$

*where $p$ is the number of parameters in $\boldsymbol{\alpha}$.*

The proof of this theorem follows from Lemma 2.6.1 and is given in Fedorov (1972, pages 71-73), Cheng (1995) and Mentré *et al.* (1997).

The Equivalence Theorem provides methods for the construction of and for verifying the $D$-optimality of a design. However, it says nothing about the uniqueness of the optimum designs. In this thesis, the uniqueness of optimal designs is not considered in any detail. Thus designs constructed numerically and found to be optimal are described and would appear to be unique. However it is recognized that other designs which are optimal and which are based on more or fewer support designs may well exist.

$D$-optimal population designs can be compared on the basis of their efficiencies. Specifically, the $D$-efficiency of an arbitrary design $\xi$ with respect to a $D$-optimal population design $\xi^*$ is defined as

$$D_{eff} = \left(\frac{|\mathbf{M}_\alpha(\xi)|}{|\mathbf{M}_\alpha(\xi^*)|}\right)^{1/p} \tag{2.38}$$

where $p$ is the number of parameters in the vector $\boldsymbol{\alpha}$ (Atkinson and Donev, 1992, page 116). The ratio is raised to the power $\dfrac{1}{p}$ in order for the efficiency to be normalized with respect to the number of parameters of interest.

Recall the discussions on invariance to linear transformation from Subsection 2.5.3. Specifically, for a linear transformation of the form $\mathbf{X}_i^* = \mathbf{X}_i\,\mathbf{A}$ the information matrices for $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$, $\mathbf{I}_{\beta}(\mathbf{X}_i) = \mathbf{X}_i'\,\mathbf{V}_i^{-1}\,\mathbf{X}_i$ and $\mathbf{I}_{\beta^*}(\mathbf{X}_i^*) = (\mathbf{X}_i^*)'\,(\mathbf{V}_i^*)^{-1}\,\mathbf{X}_i^*$, are related through $\mathbf{I}_{\beta^*}(\mathbf{X}_i^*) = \mathbf{A}'\,\mathbf{I}_{\beta}(\mathbf{X}_i)\,\mathbf{A}$. Therefore the standardized information matrices for $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$ at the population design $\xi$

$$\mathbf{M}_{\beta^*}(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_{\beta^*}(\mathbf{X}_i^*) = \sum_{i=1}^{r} w_i \frac{\mathbf{I}_{\beta^*}(\mathbf{X}_i^*)}{d_i}$$

and

$$\mathbf{M}_{\beta}(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_{\beta}(\mathbf{X}_i) = \sum_{i=1}^{r} w_i \frac{\mathbf{I}_{\beta}(\mathbf{X}_i)}{d_i}$$

are also related through $\mathbf{M}_{\beta^*}(\xi) = \mathbf{A}'\,\mathbf{M}_{\beta}(\xi)\,\mathbf{A}$. Thus

$$|\mathbf{M}_{\beta^*}(\xi)| = |\mathbf{A}|^2 |\mathbf{M}_{\beta}(\xi)|.$$

Since $\mathbf{A}$ is a constant, maximizing $\ln |\sum_{i=1}^{r} w_i\,\mathbf{M}_{\beta^*}(\mathbf{X}_i^*)|$ is the same as maximizing $\ln |\sum_{i=1}^{r} w_i\,\mathbf{M}_{\beta}(\mathbf{X}_i)|$ and thus the $D$-optimal population design for $\boldsymbol{\beta}$ is invariant to a linear transformation.

Recall from expression (2.14) that the $(j,k)$th element of the information matrix from the $i$th individual corresponding to parameters $\theta_j$ and $\theta_k$ is given by

$$\mathbf{I}_{\theta_j\,\theta_k}(\mathbf{X}_i) = \frac{1}{2} tr \left( \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_j} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \right)$$

for $i = 1, \ldots, K$. Note also that the overall variance of the $i$th response $\mathbf{V}_i$ does not change with a linear transformation $\mathbf{X}_i^* = \mathbf{X}_i\,\mathbf{A}$. This implies that

$$\mathbf{I}_{\theta_j^*\,\theta_k^*}(\mathbf{X}_i^*) = \frac{1}{2} tr \left( (\mathbf{V}_i^*)^{-1} \frac{\partial \mathbf{V}_i^*}{\partial \theta_j^*} (\mathbf{V}_i^*)^{-1} \frac{\partial \mathbf{V}_i^*}{\partial \theta_k^*} \right) = \mathbf{I}_{\theta_j\,\theta_k}(\mathbf{X}_i)$$

because $\mathbf{V}_i = \mathbf{V}_i^*$. Therefore the standardized information matrices for $\boldsymbol{\theta}$

$$\mathbf{M}_{\theta}(\xi) = \sum_{i=1}^{r} w_i \frac{\mathbf{I}_{\theta}(\mathbf{X}_i)}{d_i}$$

and

$$\mathbf{M}_{\theta^*}(\xi) = \sum_{i=1}^{r} w_i \frac{\mathbf{I}_{\theta^*}(\mathbf{X}_i^*)}{d_i}$$

are the same. Thus the $D$-optimal population designs for $\boldsymbol{\theta}$ are also invariant to linear transformation.

However, it is important again to emphasize that for the random coefficient regression model, the variance-covariance matrix for the random effects in a linearly transformed model $\mathbf{G}^*$ can be structurally different from $\mathbf{G}$. In other words, the nature of the random terms in the underlying model can change with a linear transformation and hence a particular $D$-optimal population design for $\boldsymbol{\beta}$ or $\boldsymbol{\theta}$ can be optimal for two quite different model settings.

### 2.6.5 $L$-optimality

Let $\mathbf{A}$ and $\mathbf{B}$ be $p \times p$ matrices. Then a function $L$ which satisfies the following three conditions

i. $L(\mathbf{A} + \mathbf{B}) = L(\mathbf{A}) + L(\mathbf{B})$,

ii. $L(c\,\mathbf{A}) = c\,L(\mathbf{A})$, where $c$ is a constant, and

iii. for a positive semi-definite matrix $\mathbf{A}$, $L(\mathbf{A}) \geq 0$

is called a linear function. Optimality criteria which are based on such linear functions of $\mathbf{M}_\alpha^{-1}(\xi)$ have been widely used in many design problems. According to Fedorov (1972, page 122) a design $\xi^*$ is said to be linear-optimal or $L$-optimal if it minimizes a criterion of the form

$$\Psi_L(\xi) = L\left\{\mathbf{M}_\alpha^{-1}(\xi)\right\}.$$

The following Lemma is a necessary precursor to the formulation of the Equivalence Theorem for $L$-optimal population designs and is based on Lemma 2.9.1 in Fedorov (1972, pages 123-124).

**Lemma 2.6.2** *The function $L\left\{\mathbf{M}_\alpha^{-1}(\xi)\right\}$ is a convex function on the set of information matrices $\mathcal{M}$. Furthermore, for a population design $\xi$, the directional derivative of $\Psi_L(\xi) = L\left\{\mathbf{M}_\alpha^{-1}(\xi)\right\}$ at $\xi$ in the direction of $\xi_\mathbf{x}$ is given by*

$$\phi_L(\xi_\mathbf{x}, \xi) = L\left\{\mathbf{M}_\alpha^{-1}(\xi)\right\} - L\left\{\mathbf{M}_\alpha^{-1}(\xi)\,\mathbf{M}_\alpha(\xi_\mathbf{x})\,\mathbf{M}_\alpha^{-1}(\xi)\right\}.$$

**Proof**

As for Lemma 2.6.1, the proof follows from the convexity of the set of information matrices, that is from expression (2.37). Thus from Result A.2.2 of Appendix A it follows that

$$\mathbf{M}_\alpha^{-1}(\xi) = \{(1-\epsilon)\,\mathbf{M}_\alpha(\xi_1) + \epsilon\,\mathbf{M}_\alpha(\xi_\mathbf{x})\}^{-1} \leq (1-\epsilon)\,\mathbf{M}_\alpha^{-1}(\xi_1) + \epsilon\,\mathbf{M}_\alpha^{-1}(\xi_\mathbf{x}).$$

Then from the conditions defining $L$-optimality,

$$L\{[(1-\epsilon)\,\mathbf{M}_\alpha(\xi_1) + \epsilon\,\mathbf{M}_\alpha(\xi_\mathbf{x})]^{-1}\} \leq (1-\epsilon)\,L\left\{\mathbf{M}_\alpha^{-1}(\xi_1)\right\} + \epsilon\,L\left\{\mathbf{M}_\alpha^{-1}(\xi_\mathbf{x})\right\}$$

and thus $L\left\{\mathbf{M}_\alpha^{-1}(\xi)\right\}$ is convex on the set $\mathcal{M}$.

To prove the second part of the lemma, note that by Result A.4.2 from Appendix A

$$\frac{\partial \mathbf{M}_\alpha^{-1}(\xi)}{\partial \epsilon} = -\mathbf{M}_\alpha^{-1}(\xi)\,\left[\mathbf{M}_\alpha(\xi_\mathbf{x}) - \mathbf{M}_\alpha(\xi_1)\right]\,\mathbf{M}_\alpha^{-1}(\xi).$$

Then, in view of the linearity of the function $L$,

$$\phi_L(\xi_\mathbf{x}, \xi) = \lim_{\epsilon \to 0} L\left(\frac{\partial \mathbf{M}_\alpha^{-1}(\xi)}{\partial \epsilon}\right) = L\left\{\mathbf{M}_\alpha^{-1}(\xi)\right\} - L\left\{\mathbf{M}_\alpha^{-1}(\xi)\,\mathbf{M}_\alpha(\xi_\mathbf{x})\,\mathbf{M}_\alpha^{-1}(\xi)\right\}$$

where $\mathbf{M}_\alpha(\xi) = \mathbf{M}_\alpha(\xi_1)$ when $\epsilon = 0$. □

The General Equivalence Theorem for $L$-optimal population design $\xi^*$ is now presented.

**Theorem 2.6.2** *The following three conditions on the $L$-optimal population design $\xi^*$ are equivalent:*

1. *The design $\xi^*$ minimizes $L\{\mathbf{M}_\alpha^{-1}(\xi)\}$.*
2. *The design $\xi^*$ minimizes $\max_{\mathbf{X}\in S} L\{\mathbf{M}_\alpha^{-1}(\xi)\, \mathbf{M}_\alpha(\xi_\mathbf{x})\, \mathbf{M}_\alpha^{-1}(\xi)\}$.*
3. *The support of $\xi^*$ is contained in the set of design matrices $\mathbf{X}$ such that*

$$L\{\mathbf{M}_\alpha^{-1}(\xi^*)\, \mathbf{M}_\alpha(\xi_\mathbf{x})\, \mathbf{M}_\alpha^{-1}(\xi^*)\} = L\{\mathbf{M}_\alpha^{-1}(\xi^*)\}.$$

The proof of this theorem follows from Lemma 2.6.2 and is given by Fedorov (1972, pages 125-127).

Note that in Theorem 2.6.2 if $L\{\mathbf{M}_\alpha^{-1}(\xi)\} > 0$ the $L$-optimal population design $\xi^*$ will be unique (see Fedorov, 1972, page 128).

More recently Atkinson and Donev (1992, page 113) defined the $L$-optimal criterion as a special case of that introduced by Fedorov (1972) and specifically as

$$\Psi_L(\xi) = tr\{\mathbf{M}_\alpha^{-1}(\xi)\mathbf{Q}\},$$

where $\mathbf{Q}$ is a $p \times p$ matrix of coefficients. Note that if $\mathbf{Q}$ is nonnegative definite of rank $s \leq p$ it can be expressed in the form $\mathbf{Q} = \mathbf{B}\mathbf{B}'$ where $\mathbf{B}$ is a $p \times s$ matrix of full column rank. It then follows that

$$\Psi_L(\xi) = tr\{\mathbf{M}_\alpha^{-1}(\xi)\mathbf{Q}\} = tr\{\mathbf{M}_\alpha^{-1}(\xi)\mathbf{B}\mathbf{B}'\} = tr\{\mathbf{B}'\mathbf{M}_\alpha^{-1}(\xi)\mathbf{B}\}.$$

If $\mathbf{Q} = \mathbf{c}\mathbf{c}'$, then $tr\{\mathbf{M}_{\alpha}^{-1}(\xi)\,\mathbf{c}\,\mathbf{c}'\}$ equals $\mathbf{c}'\,\mathbf{M}_{\alpha}^{-1}(\xi)\,\mathbf{c}$, the variance of the estimator of the linear function $\mathbf{c}'\,\boldsymbol{\alpha}$, and thus the $L$-criterion reduces to that of $c$-optimality. Also if $\mathbf{Q} = \mathbf{I}$, then the criterion $\Psi_L(\xi) = tr\{\mathbf{M}_{\alpha}^{-1}(\xi)\}$ reduces to $A$-optimality.

Suppose that interest centers on the precise estimation of the mean responses at a given set of design points. Suppose further that these design points are assembled in a matrix $\mathbf{X}_g$ of full row rank. The maximum likelihood estimator of the mean responses at $\mathbf{X}_g$ in model (2.1) is equal to $\mathbf{X}_g\,\hat{\boldsymbol{\beta}}$ and its asymptotic variance based on a design $\xi$ is given by $\mathbf{X}_g\,\mathbf{M}_{\beta}^{-1}(\xi)\mathbf{X}_g'$. Thus a suitable optimality criterion is that of average variance, termed $V$-optimality, and expressed proportionally as

$$\Psi_V(\xi) = tr\left\{\mathbf{X}_g\,\mathbf{M}_{\beta}^{-1}(\xi)\,\mathbf{X}_g'\right\} = tr\left\{\mathbf{M}_{\beta}^{-1}(\xi)\,\mathbf{X}_g'\mathbf{X}_g\right\}.$$

This is clearly a special case of the $L$-optimality criterion introduced by Atkinson and Donev (1992) and thus, in turn, of that introduced by Fedorov (1972, page 122). The subscript $V$ is used to indicate that the criterion relates to the variances of mean responses. Thus a $V$-optimal design is that design for which $\Psi_V(\xi)$ is minimized over the set of all possible population designs. $V$-optimality is an important and widely used criterion and was introduced within the context of the linear mixed model by Abt *et al.* (1997). It is used extensively in this thesis.

Note immediately that the $V$-optimality criterion is invariant to linear transformation (see Subsection 2.5.3). This follows directly from the fact that the mean response $\mathbf{X}\,\boldsymbol{\beta}$ is itself invariant to such transformation. Recall however that in case of the random coefficient regression model the nature of the random terms in the underlying model can change with a linear transformation and hence that a particular $V$-optimal design can be optimal for two quite different model settings.

Since $V$-optimality is a special case of $L$-optimality, it follows immediately from Lemma 2.6.2 that the directional derivative of $\Psi_V(\xi) = tr\{\mathbf{M}_\beta^{-1}(\xi)\,\mathbf{X}_g'\mathbf{X}_g\}$ at $\xi$ in the direction of $\xi_{\mathbf{x}}$ is given by

$$\phi_V(\xi_{\mathbf{x}}, \xi) = tr\left\{\mathbf{M}_\beta^{-1}(\xi)\,\mathbf{X}_g'\mathbf{X}_g\right\} - tr\left\{\mathbf{M}_\beta^{-1}(\xi)\,\mathbf{M}_\beta(\xi_{\mathbf{x}})\,\mathbf{M}_\beta^{-1}(\xi)\,\mathbf{X}_g'\mathbf{X}_g\right\}.$$

Further the General Equivalence Theorem for $V$-optimal population designs is given in the following theorem as a special case of Theorem 2.6.2.

**Theorem 2.6.3** *The following three conditions on the $V$-optimal population design $\xi^*$ are equivalent:*

1. *The design $\xi^*$ minimizes* $tr\left\{\mathbf{M}_\beta^{-1}(\xi)\,\mathbf{X}_g'\mathbf{X}_g\right\}$.
2. *The design $\xi^*$ minimizes* $\max_{\mathbf{x}\in S} tr\left\{\mathbf{M}_\beta^{-1}(\xi)\,\mathbf{M}_\beta(\xi_{\mathbf{x}})\,\mathbf{M}_\beta^{-1}(\xi)\,\mathbf{X}_g'\mathbf{X}_g\right\}$.
3. *The support of $\xi^*$ is contained in the set of design matrices $\mathbf{X}$ such that*

$$tr\left\{\mathbf{M}_\beta^{-1}(\xi^*)\,\mathbf{M}_\beta(\xi_{\mathbf{x}})\,\mathbf{M}_\beta^{-1}(\xi^*)\,\mathbf{X}_g'\mathbf{X}_g\right\} = tr\left\{\mathbf{M}_\beta^{-1}(\xi^*)\,\mathbf{X}_g'\mathbf{X}_g\right\}.$$

$V$-optimal population designs can be compared on the basis of their efficiencies. Specifically, the $V$-efficiency of an arbitrary design $\xi$ with respect to a $V$-optimal population design $\xi^*$ is defined as

$$V_{eff} = \frac{tr\{\mathbf{M}_\beta^{-1}(\xi^*)\mathbf{X}_g'\,\mathbf{X}_g\}}{tr\{\mathbf{M}_\beta^{-1}(\xi)\mathbf{X}_g'\,\mathbf{X}_g\}}.$$

## 2.7 A brief review of the related literature

### 2.7.1 Optimal designs for the fixed effects

There are few studies reported in the literature on the construction of optimal designs for mixed models and many of these have already been referred to in Section 2.6. These literature once again are drawn together here for the sake of completeness. At the end of this subsection some general studies, which are not discussed in Section 2.6, are briefly reviewed.

The concept of $D$-optimal population designs was first introduced by Mallet and Mentré (1988) for estimation of parameters in nonlinear mixed model. They also introduced an Equivalence Theorem for $D$-optimal population designs. Later Mentré *et al.* (1995) and Mentré *et al.* (1997) discussed this concept in more details for the same model. Mallet and Mentré (1988) and Mentré *et al.* (1997) propose an algorithm to construct $D$-optimal population designs. Further they compute $D$-optimal population designs based on a finite set of individual designs for a linearized nonlinear mixed model, which is in effect a random coefficient model, numerically. The individual designs in this set have sizes which vary from 1 to $m$ where $m$ is the number of time points in the experiment. In addition, they introduce cost functions which are based on the number and nature of observations in an individual, and on the duration of experiment for each individual.

Cheng (1995) and Atkins and Cheng (1999) consider optimal population designs for the quadratic regression model with a random intercept on the interval [-1, 1], where the random intercept in this model is a random block effect. Cheng as well as Atkins and

Cheng discuss the Equivalence Theorem for $D$-optimal population designs. Cheng gives a candidate two-point $D$-optimal population design (that is, design with blocks of size two) for fitting quadratic regression model and shows its global optimality by invoking the Equivalence Theorem.

Atkins and Cheng (1999) extend the results of Cheng's two-point $D$-optimal designs to general block size for fitting quadratic model. Further they discuss $A$-optimal population designs and show that in some cases knowing the $D$- and $A$-optimal designs when all of the random errors are uncorrelated can help to find $D$- and $A$-optimal population designs for fitting random intercept models.

Abt, Liski, Mandal and Sinha (1997) present optimal population designs for the simple linear regression model with errors having a compound symmetry. This model is equivalent to a random intercept model. Abt *et al.* (1997) construct approximate optimal population designs for precise estimation of the slope parameter $\beta_1$ in the simple linear regression model based on specified sets of individual designs defined on the set $\{-k, -k+1, \ldots, 0, \ldots, k-1, k\}$ algebraically. They use the design criterion which minimizes the variance of the generalized least squares estimator of $\beta_1$. Further, they derive $V$-optimal population designs for growth prediction algebraically. However, they do not introduce Equivalence Theorem to check the global optimality of the $V$-optimal designs and optimal designs for $\beta_1$.

In a second paper, Abt, Gaffke, Liski and Sinha (1998) consider optimal population designs for quadratic regression model with errors having a compound symmetry. They calculate $A$-, $D$- and $E$-optimal approximate population designs for precise estimation of the slope and quadratic parameters based again on specified sets of individual designs defined on the set $\{-k, -k+1, \ldots, 0, \ldots, k-1, k\}$ numerically. They also construct $V$-

optimal population designs for growth prediction based on these sets numerically and in addition they examine depends of optimal designs on the variance ratio. However, they do not introduce Equivalence Theorem for any of the criterion used.

Tan and Berger (1999) compute D-optimal designs for polynomial regression models with a random intercept numerically and compare them to designs with equally spaced time points. Verbeke and Lesaffre (1999) discuss the design aspect of longitudinal experiments when drop-out is to be expected.

The construction of optimal block designs has been considered by many authors, for example, Pukelsheim (1983), Bagchi (1987), Yah (1988), Atkinson and Donev (1989), Shah and Sinha (1989), Mejza and Kageyama (1995), Morgan (1997). More recently Goos and Vandebroek (2001) and Goos (2002) have derived results on exact D-optimal response surface designs in the presence of random block effects.

Several authors have discussed optimal designs for generalized linear mixed models, for example, Snijders and Bosker (1993; 1999, pages 141-154), Moerbeek, Van Breukelen and Berger (2001) and Moerbeek and Ausems (2003). Ouwens, Tan and Berger (2002) compute locally $D$-optimal designs for first- and second-degree polynomial random coefficient models with first-order autoregressive serial correlation numerically.

Jones and Wang (1999) and Jones, Wang, Jarvis and Byrom (1999) have computed D-optimal discrete designs for nonlinear mixed model numerically.

### 2.7.2 Optimal designs for the variance components

Khuri (2000) provides a comprehensive review of the literature on designs for estimating variance components. In this subsection, studies which are related to this thesis are discussed.

The information matrix for variance components and hence the design criteria are functions of the unknown variance components, for example, see expressions in (2.30) and (2.32). Thus, the choice of an efficient design cannot be made without some knowledge of these variance components. In this case there are two approaches to compute the optimal designs, local and Bayesian optimality criteria. The local optimality criteria require specifying some prior values of the variance components (Chernoff, 1953). Mukerjee and Huda (1988) and Giovagnoli and Sebastiani (1989) discuss locally optimal designs for estimation of variance components in multifactor and one-way random effects models, respectively. The Bayesian optimality criteria allow one to put a prior on the unknown parameter and thus avoid the overdependence on a single value of the parameter (Atkinson and Donev, 1992, page 197). Lohr (1995) gives Bayesian optimal designs for estimation of functions of variance components and the variance components themselves in one-way random effects model.

# Chapter 3

# Aims and Objectives of the Study

## 3.1 Introduction

The theory of optimal design for the linear model was mainly developed for responses that are independently and identically distributed. This thesis is concerned with the derivation of optimum designs for linear mixed models, which are the models that are appropriate for describing certain correlated responses. The problem of constructing optimal designs for linear mixed models is, however, very broad. Thus, this thesis is mainly focussed on optimal design theory for random coefficient regression models which are special cases of the linear mixed model. The broad aim of the study is to obtain explicit expressions for optimal designs for these models.

This Chapter is organized as follows. In Section 3.2, the design problem and the data set that is used in this thesis are described. The models, designs, design criteria of interest, and the aims and objective of the thesis are discussed in Section 3.3.

## 3.2 Problem

In longitudinal studies, it is common for a researcher to take measurements on individuals at equally spaced time points. This design is very often not optimal for longitudinal experiments. For instance, if the researcher knows that the relation between the response and the time points is a simple linear regression with uncorrelated errors then it would be more effective to take measurements at the extreme time points in order to estimate the regression parameters precisely. However, for longitudinal studies, measurements on an individual at different time points are very often correlated. Therefore, the selection of optimal time points should consider the correlation structure of the measurements. An example of a longitudinal experiment is an animal experiment in which a researcher may be interested in collecting data at a fixed number of time points, say two, three or more, due to some constraints such as one on the total number of measurements. Then the design questions posed by the researcher are at which time points to take measurements, how many repeated measurements to be made on each animal, how many animals to use in the study and how to optimally allocate the time points to the animal so as to estimate the parameters of the model as precisely as possible.

This study was motivated by a data set from an experiment undertaken at the International Livestock Research Institute (ILRI), Kenya (Duchateau, Janssen and Rowlands, 1998, page 13). The objective of the research was to compare breed differences in susceptibility to the disease trypanosomosis. The animals used in the experiment are from two different cattle breeds, N'Dama and Boran. Six animals were taken from each breed. The percentage packed cell volume (PCV), which is the percentage of the volume of the blood serum taken up by the red blood cells, was measured for each animal at a series of fourteen different time points following experimental infection with trypanosomosis. The data are shown in Table 3.1. Duchateau, Janssen and Rowlands (1998) fitted different models to this data, including a simple linear regression model with a random intercept where the animals are included as a random effect.

If the researchers were to redesign the experiment, some of the questions to be considered from the design point of view would be: how many animals should be allocated to different designs $\mathbf{t}$, where $\mathbf{t}$ is a vector of time points, and how many measurements should be taken at each time point within a design $\mathbf{t}$ in order to estimate the fixed effects, the variance components and the mean responses of an appropriate linear mixed model as precisely as possible. These questions inspired the present study and capture the aim and essence of this thesis, which is to provide answers to these questions.

## 3.3 Aims and Objectives of the study

The optimal design problem introduced in Section 3.2 is now formulated within a more specific framework.

### 3.3.1 Models

Consider a longitudinal experiment with $K$ individuals. Suppose that each of the $K$ individuals provides measurements $y_{ij}$ at $d_i$ time points $t_{ij}$ taken from the set $\{0, 1, \ldots, k\}$, where $j = 1, \ldots, d_i$, $i = 1, \ldots, K$ and $k$ is an integer. Furthermore, suppose that the time

Table 3.1: PCV (%) at a series of 14 different time points following experimental trypanosomal infection in cattle from the N'Dama and Boran breeds.

| | | Days following infection | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Breed | Animal | 0 | 2 | 4 | 7 | 9 | 14 | 17 | 18 | 21 | 23 | 25 | 29 | 31 | 35 |
| Boran | 1 | 36.2 | 35.9 | 35.3 | 35.4 | 35.4 | 31.5 | 25.5 | 34.4 | 34.1 | 25.8 | 28.7 | 21.6 | 21.3 | 17.8 |
| | 2 | 35.9 | 38.5 | 35.9 | 36.0 | 36.3 | 36.3 | 25.2 | 31.5 | 30.6 | 28.7 | 29.0 | 23.9 | 21.3 | 18.1 |
| | 3 | 29.5 | 33.3 | 29.2 | 29.9 | 29.0 | 29.9 | 21.3 | 27.4 | 25.5 | 25.5 | 24.8 | 23.6 | 22.6 | 20.4 |
| | 4 | 28.5 | 27.6 | 27.9 | 27.7 | 29.3 | 26.7 | 21.3 | 26.7 | 25.2 | 23.6 | 23.6 | 20.0 | 19.4 | 17.2 |
| | 5 | 30.4 | 29.5 | 28.8 | 28.7 | 28.7 | 27.1 | 20.7 | 25.2 | 22.9 | 23.2 | 22.9 | 20.7 | 19.1 | 18.5 |
| | 6 | 33.7 | 36.2 | 33.3 | 32.2 | 30.9 | 29.6 | 22.6 | 30.3 | 28.3 | 25.8 | 24.5 | 21.6 | 17.5 | 15.9 |
| N'Dama | 1 | 30.4 | 33.0 | 33.3 | 31.9 | 30.6 | 31.2 | 27.7 | 28.0 | 28.3 | 27.7 | 25.8 | 26.1 | 24.5 | 22.6 |
| | 2 | 37.5 | 37.8 | 36.5 | 35.7 | 35.7 | 33.8 | 33.4 | 31.5 | 32.5 | 34.7 | 30.6 | 31.5 | 25.2 | 28.7 |
| | 3 | 32.4 | 30.4 | 31.7 | 31.2 | 31.5 | 27.7 | 27.1 | 27.4 | 29.0 | 28.0 | 27.4 | 28.3 | 26.1 | 22.9 |
| | 4 | 34.3 | 33.0 | 27.5 | 36.3 | 34.1 | 30.6 | 27.7 | 29.9 | 28.0 | 27.1 | 26.7 | 28.7 | 23.9 | 22.6 |
| | 5 | 30.4 | 32.1 | 32.1 | 30.9 | 30.6 | 29.6 | 23.6 | 29.0 | 29.6 | 28.7 | 27.1 | 25.8 | 26.1 | 24.2 |
| | 6 | 40.4 | 37.5 | 38.8 | 37.0 | 38.9 | 31.9 | 27.1 | 31.5 | 31.2 | 33.1 | 35.4 | 30.6 | 28.7 | 26.1 |

$t_{ij}$ is the only explanatory variable for the response $y_{ij}$. The following three models for this experimental setting are considered in the thesis.

1. **Simple linear regression with a random intercept**

   The simple linear regression model with a random intercept is given by

$$y_{ij} = \beta_o + \beta_1 t_{ij} + b_i + e_{ij}, \quad j = 1, \ldots, d_i, \ i = 1, \ldots, K \tag{3.1}$$

   where it is assumed that the $i$th individual effect $b_i$ is $\mathcal{N}(0, \sigma_b^2)$, that the random error associated with the $j$th observation on the $i$th individual $e_{ij}$ is $\mathcal{N}(0, \sigma_e^2)$, and that $b_i$ and $e_{ij}$ are independent for $j = 1, \ldots, d_i$ and $i = 1, \ldots, K$. The intercept $\beta_o$ and slope $\beta_1$ are the fixed effects and $\sigma_b^2$ and $\sigma_e^2$ comprise the variance components.

   The matrix form of model (3.1) is

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{1}_{d_i} b_i + \mathbf{e}_i, \quad i = 1, \ \ldots, \ K,$$

   where $\mathbf{y}_i = (y_{i1} \ \ y_{i2} \ldots \ y_{id_i})'$, $\mathbf{X}_i = [\mathbf{1}_{d_i} \ \ \mathbf{t}_i]$ is the design matrix with $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{id_i})'$ a vector of time points, $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ and $\mathbf{e}_i = (e_{i1}, e_{i2}, \ldots, e_{id_i})'$. Furthermore $b_i \sim \mathcal{N}(0, \sigma_b^2)$, $\mathbf{e}_i \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I}_{d_i})$ and $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$ is the vector of variance components.

2. **Quadratic regression with a random intercept**

   The quadratic regression model with a random intercept is given by

$$y_{ij} = \beta_0 + b_i + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + e_{ij}, \quad j = 1, \ \ldots \ d_i, \quad i = 1, \ \ldots, \ K. \tag{3.2}$$

   It is again assumed that $b_i \sim \mathcal{N}(0, \sigma_b^2)$, that $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$ and that $b_i$ and $e_{ij}$ are independent for $j = 1, \ldots, d_i$ and $i = 1, \ldots, K$. The matrix form of this model is similar to that for the simple linear regression case except that here $\boldsymbol{\beta} = (\beta_o, \beta_1, \beta_2)'$ and $\mathbf{X}_i = [\mathbf{1}_i\ \mathbf{t}_i\ (\mathbf{t}_i^{(2)})]$, where $\mathbf{t}_i^{(2)} = (t_{i1}^2, t_{i2}^2, \ldots, t_{id_i}^2)'$.

3. **Simple linear random coefficient regression model**

The simple linear random coefficient regression model is given by

$$y_{ij} = (\beta_o + b_{0i}) + (\beta_1 + b_{1i})\, t_{ij} + e_{ij}, \quad j = 1, 2, \ldots, d_i, \quad i = 1, 2, \ldots, K \tag{3.3}$$

where $b_{0i}$ and $b_{1i}$ represent the occurrence of random effects. It is assumed that $b_{0i} \sim \mathcal{N}(0, \sigma_{b_0}^2)$, that $b_{1i} \sim \mathcal{N}(0, \sigma_{b_1}^2)$, that $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$ and that $b_{0i}$ and $b_{1i}$ are correlated with $Cov(b_{0i}, b_{1i}) = \sigma_{b_0 b_1}$. Further it is assumed that the error terms $e_{ij}$ are independent of $b_{0i}$ and $b_{1i}$ within and between individuals. Therefore, the vector of variance components is given as $\boldsymbol{\theta} = (\sigma_e^2, \sigma_{b_0}^2, \sigma_{b_1}^2, \sigma_{b_0 b_1})$. The matrix form of model (3.3) is

$$\mathbf{y}_i = \mathbf{X}_i\, \boldsymbol{\beta} + \mathbf{Z}_i\, \mathbf{b}_i + \mathbf{e}_i$$

where $\mathbf{Z}_i = \mathbf{X}_i = [\mathbf{1}_{d_i}\ \ \mathbf{t}_i]$ with $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{id_i})'$, $\boldsymbol{\beta} = (\beta_0, \beta_1)'$, $\mathbf{b}_i = (b_{0i}, b_{1i})'$ and $\mathbf{e}_i = (e_{i1}, e_{i2}, \ldots, e_{id_i})'$. Let $\mathbf{b}_i = (b_{0i}, b_{1i})'$ be the vector of the random effects. Then for the above model

$$Var(\mathbf{b}_i) = \mathbf{G} = \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma_{b_1}^2 \end{pmatrix},$$

$\mathbf{b}_i \sim \mathcal{N}(0, \mathbf{G})$ and $\mathbf{e}_i \sim \mathcal{N}(0, \sigma_e^2\, \mathbf{I}_{d_i})$.

### 3.3.2 Designs

Consider the longitudinal experiment described in the previous subsection. The $d$-point design $\mathbf{t} = (t_1, \ldots, t_d)$, where $t_j \in \{0, 1, \ldots, k\}$, which puts equal weight on each point is termed a $d$-point individual design. The space of all such designs for $d = 1, \ldots, k+1$ is relevant to the design problems in this thesis and specifically, the optimal population designs are constructed from this space and its subspaces. Two sets of individual designs are considered in this thesis; the set of designs with non-repeated points, i.e. with $0 \leq t_1 < \ldots < t_d \leq k$ and the set of designs with repeated points, i.e. with $0 \leq t_1 \leq \ldots \leq t_d \leq k$. The spaces of designs for both cases are described below.

### Space of designs with non-repeated time points

Consider the set consisting of all possible $d$-point individual designs $\mathbf{t} = (t_1, t_2, \ldots, t_d)$ which put equal weights on the distinct time points $t_1, t_2, \ldots, t_d$ with

$$t_j \in \{0, 1, \ldots, k\} \text{ and } 0 \leq t_1 < t_2 < \ldots < t_d \leq k \tag{3.4}$$

where $k$ is an integer with $k \geq 1$ and $d \leq k+1$. The space of designs for non-repeated points can thus be defined as the set

$$S_{d,k} = \{\mathbf{t} : \mathbf{t} = (t_1, t_2, \ldots, t_d),\ t_j \in \{0, 1, \ldots, k\},\ j = 1, \ldots, d,\ 0 \leq t_1 < t_2 < \ldots < t_d \leq k\}.$$

There are clearly $\begin{pmatrix} k+1 \\ d \end{pmatrix}$ designs in the space $S_{d,k}$ and these are summarized for all values of $d$ from 1 to $k+1$ in Table 3.2. Note that there is a total of

$$N = \begin{pmatrix} k+1 \\ 1 \end{pmatrix} + \begin{pmatrix} k+1 \\ 2 \end{pmatrix} + \ldots + \begin{pmatrix} k+1 \\ k+1 \end{pmatrix} = 2^{k+1} - 1$$

individual designs.

Table 3.2: List of individual designs with non-repeated time points from the set $\{0, 1, \ldots, k\}$.

| $d$ | Individual designs ($\mathbf{t}$) | Number of possible designs |
|---|---|---|
| 1 | $(0), (1), \ldots, (k)$ | $\binom{k+1}{1} = k+1$ |
| 2 | $(0, 1), \ldots, (k-1, k)$ | $\binom{k+1}{2} = \frac{k\,(k+1)}{2}$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $d$ | $(0, 1, \ldots, d), \ldots,$ $(k-(d-1), \ldots, k)$ | $\binom{k+1}{d}$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $k+1$ | $(0, 1, \ldots, k)$ | $\binom{k+1}{k+1} = 1$ |
| Total number of designs | | $2^{k+1} - 1$ |

### Space of designs with repeated time points

Consider the set consisting of all possible $d$-point individual designs $\mathbf{t} = (t_1, t_2, \ldots, t_d)$ which put equal weights on the time points $t_1, t_2, \ldots, t_d$ which are not necessarily distinct, that is

$$t_j \in \{0, 1, \ldots, k\} \quad \text{and} \quad 0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k \tag{3.5}$$

where $k$ is an integer with $k \geq 1$ and $d \leq k + 1$. The space of designs for repeated points can thus be defined as the set

$$T_{d,k} = \{\mathbf{t} : \mathbf{t} = (t_1, t_2, \ldots, t_d),\, t_j \in \{0, 1, \ldots, k\},\ j = 1, \ldots, d,\ 0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k\}.$$

Observe that the set of designs with non-repeated points $S_{d,k}$ is a subset of $T_{d,k}$.

The number of all possible $d$-point designs $\mathbf{t} \in T_{d,k}$ can be found using occupancy arguments (see Feller, 1968, pages 38-39). Thus the process of listing all possible $d$-point designs is equivalent to the placement of $d$ balls into $k + 1$ cells. Then by occupancy arguments there are

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

possible $d$-point designs. An alternative derivation, also based on occupancy arguments, proceeds as follows. Suppose that a $d$-point design comprises $r$ distinct numbers where $1 \leq r \leq d$. Then there are $\binom{k+1}{r}$ ways of choosing these numbers. Furthermore, for any $r$ distinct numbers there are $\binom{d-1}{r-1}$ ways of selecting the numbers of repeats of each of the points. Overall, therefore, it follows from expression (11.9) in Feller (1968, page

62) that the total number of $d$-point designs is given by

$$\sum_{r=1}^{d} \begin{pmatrix} d-1 \\ r-1 \end{pmatrix} \begin{pmatrix} k+1 \\ r \end{pmatrix} = \begin{pmatrix} k+d \\ d \end{pmatrix}.$$

Finally, the total number of all possible $d$-point designs in the set $T_{d,k}$ is equal to

$$N = \sum_{d=1}^{k+1} \begin{pmatrix} k+d \\ d \end{pmatrix} = \frac{\Gamma(k+\frac{3}{2})\,4^{k+1}}{\Gamma(\frac{1}{2})\,\Gamma(k+2)} - 1 = 2 \begin{pmatrix} 2k+1 \\ k \end{pmatrix} - 1,$$

a result based on equation (12.8) of Feller (1968, page 64). As an example suppose that $k = 2$ and thus that $t_j \in \{0, 1, 2\}$. Then there are a total of $2 \begin{pmatrix} 5 \\ 2 \end{pmatrix} - 1 = 19$ designs in the set $T_{d,2}$, $1 \leq d \leq 3$. These designs are listed in Table 3.3.

Now, using the notation of Subsection 2.6.2, an approximate population design based on $r$ individual designs can be summarized succinctly as

$$\xi = \left\{ \begin{array}{ccc} \mathbf{t}_1, & \ldots, & \mathbf{t}_r \\ w_1, & \ldots, & w_r \end{array} \right\},$$

where

$$0 < w_i < 1 \quad \text{and} \quad \sum_{i=1}^{r} w_i = 1.$$

Then the information matrix for a parameter vector $\boldsymbol{\alpha}$ in terms of the population design $\xi$ is given by

$$\mathbf{M}_{\alpha}(\xi) = \sum_{i=1}^{r} w_i\, \mathbf{M}_{\alpha}(\mathbf{t}_i)$$

where $\mathbf{M}_{\alpha}(\mathbf{t}_i)$ is the standardized information matrix for $\boldsymbol{\alpha}$ at the individual design $\mathbf{t}_i$, $i = 1, \ldots, r$. In this thesis the approximate designs on a per point basis are sought because they are easier to construct with the help of the General Equivalence Theorem than the corresponding exact designs.

Table 3.3: List of individual designs with repeated points from the set $\{0,1,2\}$.

| $d$ | $r$ | Number of $d$-point designs | Individual designs |
|---|---|---|---|
| 1 | 1 | $\binom{1-1}{1-1}\binom{3}{1}=3$ | (0),(1),(2) |
| 2 | 1 | $\binom{2-1}{1-1}\binom{3}{1}=3$ | (0,0),(1,1),(2,2) |
| | 2 | $\binom{2-1}{2-1}\binom{3}{2}=3$ | (0,1),(0,2),(1,2) |
| 3 | 1 | $\binom{3-1}{1-1}\binom{3}{1}=3$ | (0,0,0),(1,1,1),(2,2,2) |
| | 2 | $\binom{3-1}{2-1}\binom{3}{2}=6$ | (0,0,1),(0,0,2),(0,1,1),<br>(0,2,2),(1,1,2),(1,2,2) |
| | 3 | $\binom{3-1}{3-1}\binom{3}{2}=1$ | (0,1,2) |

### 3.3.3 Design criteria

This study focusses on $D$- and $V$-optimality criteria for the linear mixed models described in Subsection 3.3.1. Specifically consider the information matrix for the parameter vector $\boldsymbol{\alpha}$ at the design $\xi$, namely $\mathbf{M}_\alpha(\xi) = \sum_{i=1}^{r} w_i\, \mathbf{M}_\alpha(\mathbf{t}_i)$, where $\boldsymbol{\alpha}$ denotes the fixed effects $\boldsymbol{\beta}$ or the variance components $\boldsymbol{\theta}$. Then the $D$-optimality criterion is given by

$$\Psi_D(\xi) = \ln \left| \sum_{i=1}^{r} w_i\, \mathbf{M}_\alpha(\mathbf{t}_i) \right|$$

and the $D$-optimal design $\xi_D^*$ is that design which maximizes $\Psi_D(\xi)$ over the set of all approximate population designs.

For $V$-optimality, consider a vector of time points $\mathbf{t}_g$, where the elements of $\mathbf{t}_g$ are taken from the set $\{0, 1, \ldots, k\}$ and are assembled in the design matrix $\mathbf{X}_g$ in accord with the linear mixed model of interest. Then the design problem is to estimate the mean responses at the selected vector $\mathbf{t}_g$ as precisely as possible. The maximum likelihood estimator of the mean response is $\mathbf{X}_g\, \hat{\boldsymbol{\beta}}$ and its asymptotic variance based on the population design $\xi$ is equal to $\mathbf{X}_g\, \mathbf{M}_\beta^{-1}(\xi)\, \mathbf{X}_g'$. Therefore, the $V$-optimality criterion is given by

$$\Psi_V(\xi) = tr\{\mathbf{M}_\beta^{-1}(\xi)\, \mathbf{X}_g'\, \mathbf{X}_g\}.$$

### 3.3.4 Aims and objectives

The major objective of this thesis is to construct $D$- and $V$-optimal designs for the random intercept models (3.1) and (3.2) algebraically and to confirm their global optimality by invoking the Equivalence Theorem.

The secondary objective is to investigate the nature of optimal designs for the simple linear random coefficient regression model (3.3) numerically, taking cognisance of the fact that the structure of the variance-covariance matrix of the random effects **G** in such a model may change due to a linear transformation of the time points.

The construction of optimal population designs for longitudinal models has been extensively studied in the design literature in various contexts. In particular Cheng (1995) and Atkins and Cheng (1999) gave results on optimal population designs for the quadratic regression model with a random intercept over the design space [-1, 1]. Their results, however, cannot be directly translated to design spaces comprising a finite number of time points, especially when the number of such points is small. Abt *et al.* (1997) derived optimal population designs for estimation of the slope parameter and for growth prediction in the simple linear regression model with a random intercept algebraically. They assumed the time points available in the design space to be equally spaced. In a second paper, Abt *et al.* (1998) studied optimal designs for estimation of the linear and quadratic coefficients and for growth prediction in the quadratic regression model with a random intercept numerically. However, in both papers, the authors only considered a limited number of individual designs on which to base the population designs and in addition did not apply the Equivalence Theorem to check the global optimality of the designs. Moreover, as will be discussed in Chapters 5 and 7, their designs are not always optimal.

The approach used in this thesis is different from the above studies and provides solutions to the problems discussed in this chapter. The results obtained in this thesis will be compared with the above studies in a detailed manner.

# Chapter 4

# $D$-optimal Designs for the Simple Linear Regression Model with a Random Intercept

## 4.1 Introduction

In this Chapter optimal designs for the precise estimation of the intercept and slope parameters and of the variance components for the simple linear regression model with a random intercept are examined. Recall from Subsection 3.3.1 that the simple linear regression model with a random intercept is given by

$$y_{ij} = \beta_o + \beta_1 t_{ij} + b_i + e_{ij},\ j = 1, \ldots, d_i \ \text{ and } \ i = 1, \ldots, K \tag{4.1}$$

where $y_{ij}$ is the $j$th observation on individual $i$, $t_{ij} \in \{0, 1, \ldots, k\}$ with $k \geq 1$ and $K$ is the number of individuals. The intercept $\beta_o$ and slope $\beta_1$ are fixed effects. It is assumed that

$b_i \sim \mathcal{N}(0, \sigma_b^2)$, that $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$, and that $b_i$ and $e_{ij}$ are independent for $j = 1, \ldots, d_i$ and $i = 1, \ldots, K$. Furthermore, the matrix form of model (4.1) is given by

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{1}_{d_i} b_i + \mathbf{e}_i$$

where $\mathbf{y}_i = (y_{i1} \;\; y_{i2} \ldots \; y_{id_i})'$, $\mathbf{X}_i = [\mathbf{1}_{d_i} \;\; \mathbf{t}_i]$ is the design matrix with $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{id_i})'$ a vector of time points and $\mathbf{e}_i = (e_{i1}, e_{i2}, \ldots, e_{id_i})'$, $i = 1, \ldots, K$.

The design problems of this chapter are that of optimally choosing the time points $\mathbf{t}_i$ for individuals $i = 1, \ldots, K$ so as to estimate the fixed effects $\boldsymbol{\beta}$, the slope parameter $\beta_1$ and the variance components $\boldsymbol{\theta}$ in model (4.1) as precisely as possible. The designs specified by $\mathbf{t}_i$, $i = 1, \ldots, K$ are taken to be elements of the space of designs $S_{d,k}$, i.e. designs with non-repeated time points, or elements of the space of designs $T_{d,k}$ which comprises designs with repeated time points.

In Section 4.2, some preliminary results on the geometry of the space of designs $S_{d,k}$ and $T_{d,k}$ are discussed. These results play an important role in the computation of the optimum designs for the simple linear regression model with a random intercept. In Section 4.3, the construction of $D$-optimal designs for estimation of the fixed effects on the space of designs with non-repeated time points is discussed. The construction of $D$-optimal designs for the estimation of fixed effects in the case where replication of time points in an individual design is possible is discussed in Section 4.4. The efficiencies of $D$-optimal population designs with repeated points relative to designs with non-repeated points are considered in Section 4.5. Optimal design for estimation of the slope parameter in model (4.1) is presented in Section 4.6 and the $D$-optimal designs for variance components are considered in Section 4.7. Finally, in Section 4.8 the results of the Chapter are illustrated by means of an example.

## 4.2 Geometry of the spaces of $d$-point designs

### 4.2.1 Non-repeated time points

Consider the space of $d$-point designs with non-repeated time points introduced in Section 3.3.2 and specified by

$$S_{d,k} = \{\mathbf{t} : \mathbf{t} = (t_1, \ldots, t_d),\ t_j \in \{0, 1, \ldots, k\},\ j = 1, \ldots, d,\ 0 \le t_1 < t_2 < \ldots < t_d \le k\}.$$

Then it follows that the elements of the $d$-point designs $\mathbf{t}$ in the set $S_{d,k}$ satisfy the $d + 1$ inequalities

$$\begin{cases} t_1 \ge 0 \\ t_2 \ge t_1 + 1 \\ \vdots \\ t_{d-1} \ge t_{d-2} + 1 \\ t_d \ge t_{d-1} + 1 \\ t_d \le k \end{cases} \tag{4.2}$$

for $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$. Thus the designs correspond to a lattice of points in $\mathbb{R}^d$ enclosed by the polytope $P_{d,k}$ described by the system of equations

$$\begin{pmatrix} -1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & -1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -1 & \ldots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \ldots & 1 & -1 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{d-1} \\ x_d \end{pmatrix} \le \begin{pmatrix} 0 \\ -1 \\ -1 \\ \vdots \\ -1 \\ k \end{pmatrix} \tag{4.3}$$

for $x_j \in [0, k]$, $j = 1, \ldots, d$. Note that this polytope is a subset of the hypercube defined by the Cartesian product $[0, k] \times [0, k] \times \ldots \times [0, k] = [0, k]^d$ in $\mathbb{R}^d$ and denoted by $C_{d,k}$.

The $d + 1$ vertices of the polytope $P_{d,k}$ are defined as solutions to the $d + 1$ subsets of $d$ equations taken from the system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is the $(d + 1) \times d$ coefficient matrix on the left hand side of (4.3), $\mathbf{x} = (x_1, x_2, \ldots, x_{d-1}, x_d)'$ and $\mathbf{b} = (0, -1, \ldots, -1, k)'$ (see e.g. Cook, Cunningham, Pulleyblank and Schrijver, 1998, Proposition 6.7, page 205). These vertices are thus given by

$$\begin{aligned} \mathbf{v}_1 &= (0, 1, \ldots, d-2, d-1) \\ \mathbf{v}_2 &= (0, 1, \ldots, d-2, k) \\ &\vdots \\ \mathbf{v}_{d+1} &= (k-d+1, k-d, \ldots, k) \end{aligned} \tag{4.4}$$

and clearly correspond to designs in the set $S_{d,k}$. Note that the vertices $\mathbf{v}_{j+1}$, $j = 0, 1, \ldots, d$, can be generated systematically as the union of the first $d - j$ elements of $\mathbf{v}_1$ and the last $j$ elements of $\mathbf{v}_{d+1}$. Thus in particular

$$\begin{aligned} \mathbf{v}_1 &= \underbrace{(0, 1, \ldots, d-2, d-1)}_{d \text{ elements}} \\ \mathbf{v}_{j+1} &= \underbrace{(0, 1, \ldots, d-j-1,}_{d-j \text{ elements}} \underbrace{k-j+1, \ldots, k-1, k)}_{j \text{ elements}}, \; j = 1, \ldots, d-1 \end{aligned}$$

and

$$\mathbf{v}_{d+1} = (k-d+1, k-d, \ldots, k).$$

The vertices can be paired according to their distance from the center of the hypercube $C_{d,k} = [0, k]^d$, i.e. from the point $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$. For example, the vertices

$\mathbf{v}_1 = (0, 1, \ldots, d-2, d-1)$ and $\mathbf{v}_{d+1} = (k-d+1, \ldots, k-1, k)$ have equal squared distances from $\mathbf{x}_c$ which are given by

$$S_1 = \left(\frac{k}{2}\right)^2 + \left(\frac{k}{2} - 1\right)^2 + \ldots + \left(\frac{k}{2} - d + 1\right)^2.$$

The vertices $\mathbf{v}_2 = (0, 1, \ldots, d-2, k)$ and $\mathbf{v}_d = (0, k-d+2, \ldots, k-1, k)$ also have equal squared distances from $\mathbf{x}_c$ of

$$S_2 = 2\left(\frac{k}{2}\right)^2 + \left(\frac{k}{2} - 1\right)^2 + \ldots + \left(\frac{k}{2} - d + 2\right)^2.$$

Clearly $S_2$ is greater than $S_1$ and thus the two vertices $\mathbf{v}_2$ and $\mathbf{v}_d$ are further away from the center of the hypercube $C_{d,k}$ than $\mathbf{v}_1$ and $\mathbf{v}_{d+1}$.

Consider now, more generally, the pair of vertices $\mathbf{v}_{j+1} = (0, 1, \ldots, d-j-1, k-j+1, \ldots, k-1, k)$ and $\mathbf{v}_{d-j+1} = (0, \ldots, j-1, k-d+j+1, \ldots, k-1, k)$ for any $d$. There are $\frac{d+1}{2}$ pairs of vertices specified by $\mathbf{v}_{j+1}$ and $\mathbf{v}_{d-j+1}$ for $j = 1, \ldots, \frac{d-1}{2}$ when $d$ is odd. For $d$ even there are $\frac{d}{2}$ pairs of vertices specified by $\mathbf{v}_{j+1}$ and $\mathbf{v}_{d-j+1}$ for $j = 1, \ldots, \frac{d}{2} - 1$ and a single vertex $\mathbf{v}_{\frac{d}{2}+1}$. These pairs of vertices are equal distances from the center of the hypercube $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$ with the squared distance given by

$$\begin{aligned} S_{j+1} &= \sum_{i=1}^{d-j} \left[\frac{k}{2} - (i-1)\right]^2 + \sum_{i=d-j+1}^{d} \left[\frac{k}{2} - (d-i)\right]^2 \\ &= \frac{1}{12} d\,(2 + 4\,d^2 + 6\,k + 3\,k^2 - 6\,d\,(1+k)) - (k-d+1)(j^2 - d\,j). \end{aligned} \tag{4.5}$$

Note that

$$\sum_{i=d+1}^{d} \left(\frac{k}{2} - (d-i)\right)^2 = \sum_{i=1}^{0} \left(\frac{k}{2} - (i-1)\right)^2 = 0.$$

For given values of $k$ and $d$, the expression in (4.5) is a quadratic in the index $j$. Solving the equation

$$\frac{\partial S_{j+1}}{\partial j} = (d - 2\,j)\,(k - d + 1) = 0$$

for $j$ gives the solution $\frac{d}{2}$ and, since

$$\frac{\partial^2 S_{j+1}}{\partial j^2} = -2\,(k-d+1) < 0,$$

the squared distance $S_{j+1}$ is therefore maximized at $j = \frac{d}{2}$. Thus when the number of points in the design, $d$, is an even integer the vertex furthest from $\mathbf{x}_c$ is indexed by $j = \frac{d}{2}$ and is thus given by

$$\mathbf{v}_{\frac{d}{2}+1} = (0, 1, \ldots, \frac{d}{2} - 1, k - \frac{d}{2} + 1, \ldots, k-1, k).$$

This vertex is referred to, somewhat informally, as the "extreme vertex". When $d$ is an odd integer, $\frac{d}{2}$ is not an integer. However since $S_{j+1}$ is quadratic in $j$, its maximum over the integers $j = 1, \ldots, d-1$ occurs at an integer closest to $\frac{d}{2}$, namely $\frac{d-1}{2}$ and $\frac{d+1}{2}$. Thus the two vertices

$$\mathbf{v}_{\frac{d+1}{2}} = (0, 1, \ldots, \frac{d-3}{2}, k - \frac{d-1}{2}, \ldots, k-1, k)$$

and

$$\mathbf{v}_{\frac{d+3}{2}} = (0, 1, \ldots, \frac{d-1}{2}, k - \frac{d-3}{2}, \ldots, k-1, k)$$

with indices $j = \frac{d-1}{2}$ and $\frac{d+1}{2}$ respectively are furthest from the point $\mathbf{x}_c$. These are again termed as the "extreme" vertices. The following example helps to fix the above ideas.

**Example 4.2.1** Consider exact two-point individual designs, i.e. $d = 2$ with $k \geq 1$. Then the polytope $P_{2,k}$ is a triangle with vertices $(0,1), (0,k)$ and $(k-1,k)$. This triangle encloses all possible two-point designs $(t_1, t_2)$ such that $0 \leq t_1 < t_2 \leq k$ with $t_j \in \{0, 1, \ldots, k\}$, $j = 1, 2$. Note that the extreme vertex is $(0,k)$ and its squared distance from $\mathbf{x}_c = (\frac{k}{2}, \frac{k}{2})$ is $\frac{k^2}{2}$. The squared distance from $\mathbf{x}_c$ for the other two vertices is $\frac{k^2}{2} - k + 1$ which is less than $\frac{k^2}{2}$ for $k \geq 1$. The polytope $P_{2,5}$ and the design points in $S_{2,5}$ are shown in Figure 4.1.

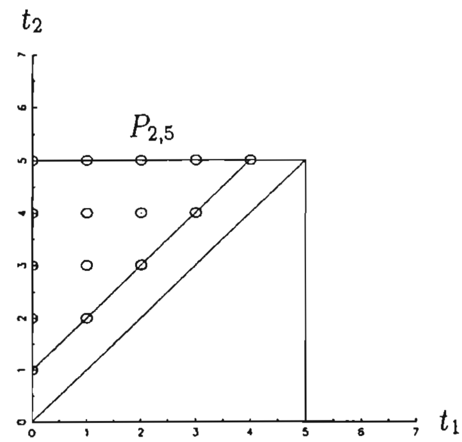


Figure 4.1: Two-point individual designs for $k = 5$. The symbol $\circ$ represents a two-point design in $S_{2,5}$ and the polytope $P_{2,5}$ is the triangle with vertices (0, 1), (0,5) and (4,5).

### 4.2.2 Repeated time points

Consider now the set of $d$-point designs with repeated time points discussed in Section 3.3.2 and specified by

$$T_{d,k} = \{\mathbf{t} : \mathbf{t} = (t_1, \ldots, t_d),\ t_j \in \{0, 1, \ldots, k\},\ j = 1, \ldots, d,\ 0 \le t_1 \le t_2 \le \ldots \le t_d \le k\}.$$

The elements of the designs $\mathbf{t}$ in the set $T_{d,k}$ satisfy the $d + 1$ inequalities

$$\left\{ \begin{aligned} -t_1 &\le 0 \\ t_1 - t_2 &\le 0 \\ &\vdots \\ t_{d-1} - t_d &\le 0 \\ t_d &\le k \end{aligned} \right. \tag{4.6}$$

for $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$. Thus the designs form a lattice of points in $\mathbb{R}^d$ enclosed by the polytope $Q_{d,k}$ described by the system

$$\begin{pmatrix} -1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & -1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -1 & \ldots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \ldots & 1 & -1 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{d-1} \\ x_d \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ k \end{pmatrix} \tag{4.7}$$

for $x_i \in [0, k]$, i=1,..., d. Note that this polytope is a subset of the hypercube $C_{d,k}$.

The $d + 1$ vertices of the polytope $Q_{d,k}$ are defined as solutions to the $d + 1$ subsets of $d$ equations from the system (4.7) and are specified as

$$\begin{aligned} \mathbf{v}_1^* &= (k, k, k, \ldots, k, k) \\ \mathbf{v}_2^* &= (0, k, k, \ldots, k, k) \\ &\vdots \\ \mathbf{v}_d^* &= (0, 0, 0, \ldots, 0, k) \\ \mathbf{v}_{d+1}^* &= (0, 0, 0, \ldots, 0, 0) \end{aligned}$$

Thus the $d + 1$ vertices have a general expression

$$\mathbf{v}_{j+1}^* = (\underbrace{0, 0, \ldots, 0, 0}_{j \text{ elements}}, \underbrace{k, k, \ldots, k, k}_{d-j \text{ elements}})$$

where $j = 0, 1, \ldots, d$ and these vertices correspond to designs in the set $T_{d,k}$. These vertices are equal distances from the center of the hypercube $C_{d,k}$, i.e. $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$, with the squared distance given by

$$S_{j+1} = j\,\frac{k^2}{4} + (d-j)\,\frac{k^2}{4} = \frac{d\,k^2}{4}.$$

**Example 4.2.2** Consider exact two-point individual designs, i.e. $d = 2$ with $k \geq 1$. Then the polytope $Q_{2,k}$ is a triangle with vertices $(0,0), (0,k)$ and $(k,k)$. This triangle encloses all possible two-point individual designs $(t_1, t_2)$ such that $0 \leq t_1 \leq t_2 \leq k$ with $t_j \in \{0, 1, \ldots, k\}$, $j = 1, 2$. The three vertices have equal squared distance $\frac{k^2}{2}$ from $x_c = (\frac{k}{2}, \frac{k}{2})$. For $k = 5$, the polytope $Q_{2,5}$ and the design points in $T_{2,5}$ are shown in Figure 4.2.

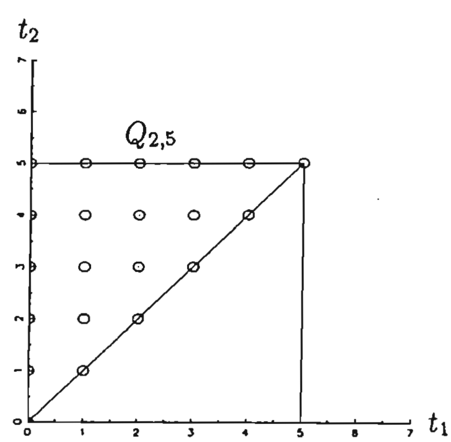


Figure 4.2: Two-point individual designs for $k = 5$. The symbol $\circ$ represents a two-point design in $T_{2,5}$ and the polytope $Q_{2,5}$ is the triangle with vertices (0, 0), (0,5) and (5,5).

## 4.3 $D$-optimal designs for the fixed effects based on designs with non-repeated time points

In this section, the $D$-optimal designs for the fixed effects, $\boldsymbol{\beta} = (\beta_o, \beta_1)'$, i.e. designs which maximize the determinant of the information matrix for $\boldsymbol{\beta}$ are discussed. The optimal designs are based on individual designs which put equal weights on non-repeated time points, i.e. designs from the set

$$S_{d,k} = \{\mathbf{t} : \mathbf{t} = (t_1, \ldots, t_d),\ t_j \in \{0, 1, \ldots, k\},\ j = 1, \ldots, d,\ 0 \leq t_1 < t_2 < \ldots < t_d \leq k\}.$$

### 4.3.1 $d$-point $D$-optimal individual designs

Consider a $d$-point individual designs $\mathbf{t} = (t_1, t_2, \ldots, t_d)$ having equal weights on the distinct points $t_1, t_2, \ldots, t_d$ and $0 \leq t_1 < t_2 < \ldots < t_d \leq k$. Then the standardized information matrix per observation for $\boldsymbol{\beta}$ at the individual design $\mathbf{t}$ can be derived from the expression (2.29) and has a generic form

$$\mathbf{M}_\beta(\mathbf{t}) = \frac{1}{\sigma_e^2} \begin{pmatrix} \dfrac{1}{1 + d\gamma} & \dfrac{\mathbf{1}'\mathbf{t}}{d(1 + d\gamma)} \\ & \\ \dfrac{\mathbf{t}'\mathbf{1}}{d(1 + d\gamma)} & \dfrac{1}{d}\mathbf{t}'(\mathbf{I} - \dfrac{\gamma}{1 + d\gamma}\mathbf{J})\mathbf{t} \end{pmatrix}. \tag{4.8}$$

Now

$$\frac{1}{d}\mathbf{t}'(\mathbf{I} - \frac{\gamma}{1 + d\gamma}\mathbf{J})\mathbf{t} = \frac{1}{d}\left(\frac{\mathbf{t}'\mathbf{t} + d\gamma\,\mathbf{t}'\mathbf{t} - \gamma\,(\mathbf{t}'\mathbf{1})^2}{1 + d\gamma}\right) = \frac{\sum_{j=1}^{d} t_j^2 + \gamma\, d\, SS(\mathbf{t})}{d(1 + d\gamma)}$$

where $SS(\mathbf{t}) = \mathbf{t}'\mathbf{t} - \frac{1}{d}(\mathbf{t}'\mathbf{1})^2 = \sum_{j=1}^{d} t_j^2 - \frac{1}{d}(\sum_{j=1}^{d} t_j)^2$, is the sum of squares of the elements of $\mathbf{t}$. Thus

$$\mathbf{M}_{\beta}(\mathbf{t}) = \frac{1}{\sigma_e^2\, d\,(1 + d\,\gamma)} \begin{pmatrix} d & \sum_{j=1}^{d} t_j \\ \sum_{j=1}^{d} t_j & \sum_{j=1}^{d} t_j^2 + \gamma\, d\, SS(\mathbf{t}) \end{pmatrix}. \tag{4.9}$$

Note that the error variance $\sigma_e^2$ factors out in (4.9) and can be taken to be 1 without loss of generality.

Then the $d$-point design $\mathbf{t}^*$ is an exact $D$-optimal individual design if it maximizes the determinant

$$|\mathbf{M}_{\beta}(\mathbf{t})| = \frac{d\,\sum_{j=1}^{d} t_j^2 + d^2\,\gamma\, SS(\mathbf{t}) - (\sum_{j=1}^{d} t_j)^2}{d^2\,(1 + d\,\gamma)^2} = \frac{SS(\mathbf{t})}{d\,(1 + d\,\gamma)}. \tag{4.10}$$

Since $d\,(1 + d\,\gamma)$ factors out in (4.10) the exact $D$-optimal individual designs for the parameters $\boldsymbol{\beta}$ are independent of $\gamma$ and simply maximize $SS(\mathbf{t})$. In other words, these optimal designs are robust to the choice of variance ratio $\gamma$ and are thus the same as for the uncorrelated case with $\gamma = 0$.

It is well known from the statistical literature that when the errors in a simple linear regression model are uncorrelated the approximate $D$-optimal design puts equal weight on the two extreme points, 0 and $k$ (Atkinson and Donev, 1992, page 60). It follows immediately therefore that the two-point $D$-optimal individual design for $\boldsymbol{\beta}$ when $\gamma > 0$ coincides with this approximate $D$-optimal design and is given by $(0, k)$.

The general results for exact $D$-optimal individual designs based on $d$ points are presented in the following theorem.

**Theorem 4.3.1** *Consider the set of all d-point individual designs* $\mathbf{t} \in S_{d,k}$ *which put equal weights on the distinct time points* $t_1, t_2, \ldots, t_d$ *where* $t_j \in \{0, 1, \ldots, k\}, j = 1, \ldots, d$ *and* $d$ *is a positive integer in the interval* $[2, k+1]$. *Then the d-point D-optimal individual designs for the fixed effects* $\boldsymbol{\beta}$ *in the simple linear regression model with a random intercept are given by*

$$\mathbf{t}_e^* = (0, 1, \ldots, \frac{d}{2} - 1, k - \frac{d}{2} + 1, \ldots, k-1, k)$$

*for d even and either*

$$\mathbf{t}_{o1}^* = (0, 1, \ldots, \frac{d-3}{2}, k - \frac{d-1}{2}, \ldots, k-1, k)$$

*or*

$$\mathbf{t}_{o2}^* = (0, 1, \ldots, \frac{d-1}{2}, k - \frac{d-3}{2}, \ldots, k-1, k),$$

*for d odd.*

**Proof**

The sum of squares

$$SS(\mathbf{t}) = \mathbf{t}'\,(\mathbf{I} - \frac{1}{d}\,\mathbf{J})\,\mathbf{t}$$

is a quadratic form in $\mathbf{t}$. The embedded matrix $(\mathbf{I} - \frac{1}{d}\,\mathbf{J})$ is symmetric and has one eigenvalue of zero and $d-1$ eigenvalues each equal to 1 (Harville, 1997, pages 516-519). Thus $(\mathbf{I} - \frac{1}{d}\,\mathbf{J})$ is positive semi-definite and the quadratic form $SS(\mathbf{t})$ is convex, but not strictly convex, on the polytope $P_{d,k}$. It now follows immediately from a standard result in the theory of nonlinear optimization (see e.g. Bazaraa and Shetty, 1979, page 92) that the maximum of $SS(\mathbf{t})$ occurs at one or more of the vertices of the polytope $P_{d,k}$.

Recall from Section 4.2 that the vertices of the polytope $P_{d,k}$ can be expressed as $\mathbf{v}_{j+1} = (0, 1, \ldots, d-j-1, k-j+1, \ldots, k-1, k)$ where $j = 0, 1, \ldots, d$. Then the general expression for $SS(\mathbf{t})$ for a vertex $\mathbf{v}_{j+1}$ is given by

$$SS(\mathbf{t}) = \frac{(k-d+1)\,(k+1)\,j\,(d-j)}{d} + \frac{d\,(d^2-1)}{12}$$

for $j = 0, 1, \ldots, d$. Solving the equation

$$\frac{\partial SS(\mathbf{t})}{\partial j} = \frac{1}{d}(d-2\,j)\,(k-d+1)\,(1+k) = 0$$

for the index $j$ gives the solution $\frac{d}{2}$. Also

$$\frac{\partial^2 SS(\mathbf{t})}{\partial j^2} = -\frac{1}{d}\,2\,(k-d+1)\,(k+1) < 0$$

for $d \le k+1$. Therefore the sum of squares $SS(\mathbf{t})$ attains a maximum at $j = \frac{d}{2}$. Thus when $d$ is an even integer the sum of squares $SS(\mathbf{t})$ is maximized at the extreme vertex $\mathbf{v}_{\frac{d}{2}+1} = \mathbf{t}_e^*$ of the polytope $P_{d,k}$. When $d$ is an odd integer it follows from the fact that $SS(\mathbf{t})$ is a quadratic in $j$ that the maximum occurs at the integers $\frac{d-1}{2}$ and $\frac{d+1}{2}$ and thus at the extreme vertices $\mathbf{v}_{\frac{d+1}{2}} = \mathbf{t}_{o1}^*$ and $\mathbf{v}_{\frac{d+3}{2}} = \mathbf{t}_{o2}^*$. □

Note that when $d$ is an even integer the information matrix for $\boldsymbol{\beta}$ at the optimal design $\mathbf{t}_e^*$ has the determinant

$$|\mathbf{M}_\beta(\mathbf{t}_e^*)| = \frac{1}{12\,(1+d\,\gamma)}\,\{2 + d^2 + 6\,k + 3\,k^2 - 3\,d\,(1+k)\}$$

while for $d$ an odd integer the information matrices at the designs $\mathbf{t}_{o1}^*$ and $\mathbf{t}_{o2}^*$ have determinants

$$|\mathbf{M}_\beta(\mathbf{t}_{o1}^*)| = |\mathbf{M}_\beta(\mathbf{t}_{o2}^*)| = \frac{1}{12\,d^2\,(1+d\,\gamma)}(d^2-1)\,\{d^2 - 3\,d\,(1+k) + 3\,(1+k)^2\}.$$

For example, consider the case of $d = 3$. Then the three-point $D$-optimal individual designs are either $\mathbf{t}^*_{o1} = (0, k-1, k)$ or $\mathbf{t}^*_{o2} = (0, 1, k)$ and the determinant of the information matrix for $\boldsymbol{\beta}$ at these designs is given by

$$|\mathbf{M}_\beta(\mathbf{t}^*_{o1})| = |\mathbf{M}_\beta(\mathbf{t}^*_{o2})| = \frac{2\,(k^2 - k + 1)}{9\,(1 + 3\,\gamma)}.$$

Note that the determinant of the information matrix depends on the variance ratio $\gamma$ but that the actual optimal designs do not.

The geometry associated with the results in the proof of Theorem 4.3.1 when $d$ is an even integer is interesting. To understand the geometry of $SS(t)$ it is helpful to start with the two-point case.

Consider the line $L$ defined by the vector $\mathbf{1}_2$ and equation $t_1 = t_2$. Let a point $\mathbf{t}_p$, which is on the line $L$ has a position vector $\mathbf{t} = (t_1, t_2)$ (see Figure 4.3). Now projection of $\mathbf{t}$ onto the line $L$ is the vector $\mathbf{t}_p = \frac{1}{2}(\mathbf{t}'\mathbf{1}_2)\mathbf{1}_2$. Thus the squared perpendicular distance of $\mathbf{t}$ from the line $L$ is equal to

$$\mathbf{t}'\,\mathbf{t} - \frac{1}{2}(\mathbf{t}'\,\mathbf{1}_2)^2$$

which is equal to $SS(t)$. Thus $SS(t)$ is the square of the length of the projection of $\mathbf{t}$ onto the plane orthogonal to $\mathbf{1}_2$ and passing through $\mathbf{t}$.

Consider now a $d$-point design $\mathbf{t} \in S_{d,k}$. Therefore the sum of squares for $\mathbf{t}$ is

$$SS(t) = \mathbf{t}'\,(\mathbf{I} - \frac{1}{d}\,\mathbf{J})\,\mathbf{t}.$$

The matrix $(\mathbf{I} - \frac{1}{d}\,\mathbf{J})$ has one eigenvalue of 0 corresponding to the eigenvector $\mathbf{1}_d$ and $d-1$ eigenvalues of 1 corresponding to eigenvectors defining the hyperplane orthogonal to $\mathbf{1}_d$.

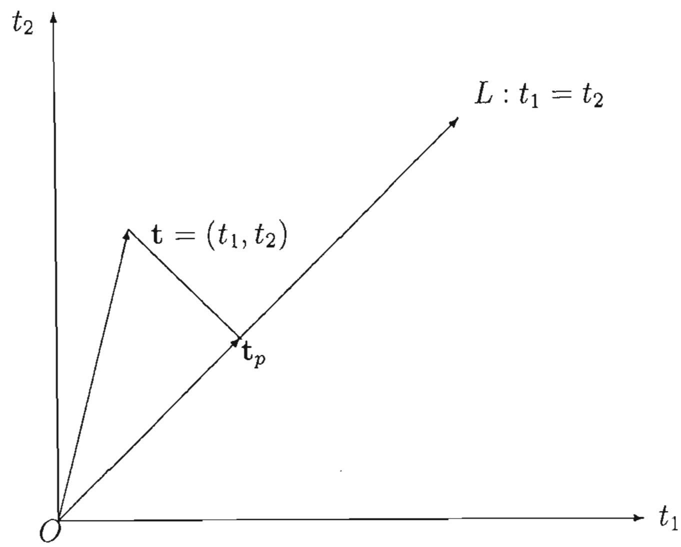


Figure 4.3: Geometry of $SS(t)$ for a two-point design.

Thus points $\mathbf{t}$ with constant sum of squares $SS(t)$ fall on hypercylinders with a common axis $\mathbf{1}_d$ (see Figure 4.3 for $d = 2$).

When $d = 2$ a point $(0, k)$ has the largest $SS(t)$. Thus the ball $B$ through $\mathbf{v}_2 = (0, k)$ with center $\mathbf{x}_c = (\frac{k}{2}, \frac{k}{2})$, i.e. through the extreme vertex, clearly encloses the polytope $P_{2,k}$ and hence all design points in the set $S_{2,k}$. Note that the sum of the elements of $\mathbf{v}_2$ and $\mathbf{x}_c$ are equal. Therefore the vector $(\mathbf{v}_2 - \mathbf{x}_c)$ is orthogonal to $\mathbf{1}_2$ and the ball is in turn enclosed by the hypercylinder passing through $\mathbf{v}_2$ with axis $\mathbf{1}_2$. Thus all design points have squared distances of projection onto the plane orthogonal to $\mathbf{1}_2$ which are less than that of $\mathbf{v}_2$. Thus $\mathbf{v}_2$ has the largest $SS(t)$. This geometry is illustrated for $k = 5$ and $d = 2$ in Figure 4.4.

Now consider $d$ even with $d > 2$. Then the design $\mathbf{v}_{\frac{d}{2}+1}$ is the vertex of the polytope $P_{d,k}$ furthest from the point $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$. Thus the ball $B$ centered at $\mathbf{x}_c$ and having boundary point $\mathbf{v}_{\frac{d}{2}+1}$ encloses all designs in the set $S_{d,k}$. Further, since the sum of the elements of $\mathbf{v}_{\frac{d}{2}+1}$ and $\mathbf{x}_c$ are the same and equal to $\dfrac{kd}{2}$, the vector $(\mathbf{v}_{\frac{d}{2}+1} - \mathbf{x}_c)$ is orthogonal

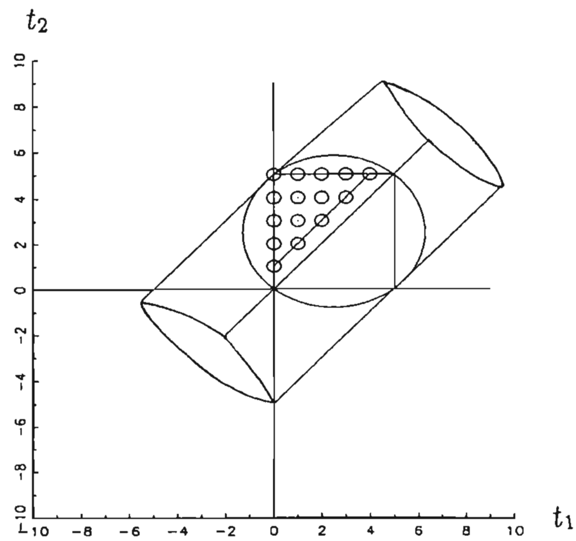


Figure 4.4: Geometry of $SS(t)$ for $k = 5$ and $d = 2$. The symbol $\circ$ represents a two-point design in $S_{2,5}$ and the polytope $P_{2,5}$ is the triangle with vertices (0, 1), (0,5) and (4,5).

to the vector $\mathbf{1}_d$. It thus follows that the hypercylinder through $\mathbf{v}_{\frac{d}{2}+1}$ with axis $\mathbf{1}_d$ contains the ball $B$ and hence that the maximum of $SS(t)$ occurs at $\mathbf{v}_{\frac{d}{2}+1}$.

The above argument does not hold when $d$ is an odd integer because the vertices $\mathbf{v}_{\frac{d+1}{2}}$ and $\mathbf{v}_{\frac{d+3}{2}}$ do not fall on the hyperplane through $\mathbf{x}_c$ and orthogonal to $\mathbf{1}_d$. Note that the ball through these vertices centered at $\mathbf{x}_c$ encloses all design points in $S_{d,k}$. However the hypercylinder with axis $\mathbf{1}_d$ containing the vertices $\mathbf{v}_{\frac{d+1}{2}}$ and $\mathbf{v}_{\frac{d+3}{2}}$ does not in turn enclose this ball.

### 4.3.2 $D$-optimal population designs based on d-point individual designs

Consider all $d$-point designs $\mathbf{t} = (t_1, \ldots, t_d)$ with $t_j \in \{0, 1, \ldots, k\}, j = 1, \ldots, d$ and $0 \leq t_1 < \ldots, < t_d \leq k$, i.e. consider the set of designs $S_{d,k}$. Recall that there are $\begin{pmatrix} k+1 \\ d \end{pmatrix}$ $d$-point individual designs in the set $S_{d,k}$. Then a population design over the set $S_{d,k}$ puts weights $w_i$ on the designs $\mathbf{t}_i$ for $i = 1, \ldots, r$ respectively and can be summarized as

$$\xi = \begin{Bmatrix} \mathbf{t}_1, & \ldots, & \mathbf{t}_r \\ w_1, & \ldots, & w_r \end{Bmatrix} \quad \text{with } 0 < w_i < 1 \text{ and } \sum_{i=1}^{r} w_i = 1. \tag{4.11}$$

Note that $r$ can range from 1 to $\begin{pmatrix} k+1 \\ d \end{pmatrix}$. The information matrix for $\boldsymbol{\beta}$ at a population design $\xi$ is therefore given by

$$\mathbf{M}_\beta(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\beta(\mathbf{t}_i)$$

where $\mathbf{M}_\beta(\mathbf{t}_i)$ is the standardized information matrix for $\boldsymbol{\beta}$ at the $d$-point individual design $\mathbf{t}_i$ and has the generic form

$$\mathbf{M}_\beta(\mathbf{t}) = \frac{1}{\sigma_e^2} \begin{pmatrix} \dfrac{1}{1+d\gamma} & \dfrac{\mathbf{1}'\mathbf{t}}{d(1+d\gamma)} \\ \dfrac{\mathbf{t}'\mathbf{1}}{d(1+d\gamma)} & \dfrac{1}{d}\mathbf{t}'(\mathbf{I} - \dfrac{\gamma}{1+d\gamma}\mathbf{J})\mathbf{t} \end{pmatrix}.$$

Note that the variance parameter $\sigma_e^2$ factors out of this expression and it can be taken to be 1 without loss of generality.

The $d$-point $D$-optimal population design is the design $\xi_D^*$ which maximizes the criterion

$$\Psi_D(\xi) = \ln|\mathbf{M}_\beta(\xi)| = \ln\left|\sum_{i=1}^{r} w_i \mathbf{M}_\beta(\mathbf{t}_i)\right|$$

over the set of all population designs specified by (4.11). Furthermore it follows immediately from the Equivalence Theorem introduced in Subsection 2.6.4 that the design $\xi_D^*$ is $D$-optimal if and only if

$$\phi(\mathbf{t}, \xi_D^*) = tr\{\mathbf{M}_\beta^{-1}(\xi_D^*)\,\mathbf{M}_\beta(\mathbf{t})\} - 2 \leq 0$$

for all individual designs $\mathbf{t}$ in the space of designs $S_{d,k}$ with equality holding at the support designs of $\xi_D^*$ where $\phi(\mathbf{t}, \xi_D^*)$ is the directional derivative of $\Psi_D(\xi) = \ln |\mathbf{M}_\beta(\xi)|$ at $\xi_D^*$ in the direction of $\mathbf{t}$.

The $D$-optimal population designs based on $d$-point individual designs follow immediately from the exact $D$-optimal individual designs given in Subsection 4.3.1 and are presented in the following two theorems.

**Theorem 4.3.2** *Consider the set of all d-point individual designs which put equal weights on the distinct time points $t_1, t_2, \ldots, t_d$ with $t_j \in \{0, 1, \ldots, k\}, j = 1, \ldots, d$, and $0 \leq t_1 < t_2 < \ldots < t_d \leq k$ for d an even integer greater than or equal to 2. Then*

$$\xi_{D_e}^* = \left\{ \begin{array}{c} (0, 1, \ldots, \frac{d}{2} - 1, k - \frac{d}{2} + 1, \ldots, k - 1, k) \\ 1 \end{array} \right\}$$

*is the D-optimal population design for the fixed effects $\boldsymbol{\beta}$ in the model (4.1) over this set for all $\gamma \geq 0$.*

**Proof**

Note that for $d = k + 1$ with $k$ odd there is only one $d$-point individual design so this necessarily comprises the required $D$-optimal population design. In the remainder of the proof $d$ is therefore taken to be strictly less than $k + 1$.

Recall from Subsection 2.6.4 that $D$-optimal designs for random intercept models are invariant to linear transformations. Thus without loss of generality, let an individual design $\mathbf{t}$ be linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$ where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$ is the center of the hypercube $C_{d,k}$. This is equivalent to moving each point $t_j$ in $\mathbf{t}$ to $\tilde{t}_j = t_j - \frac{k}{2}$, $j = 1 \ldots, d$. Thus the space of designs in the transformed coordinates is given by

$$\tilde{S}_{d,k} = \{\tilde{\mathbf{t}} : \tilde{\mathbf{t}} = (\tilde{t}_1, \ldots, \tilde{t}_d),\, \tilde{t}_j \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}-1, \frac{k}{2}\},\ j = 1, \ldots, d,$$
$$-\frac{k}{2} \leq \tilde{t}_1 < \tilde{t}_2 < \ldots < \tilde{t}_d \leq \frac{k}{2}\}.$$

Then the proposed optimum design $\xi^*_{D_e}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{D_e} = \left\{ \begin{array}{c} (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d}{2}-1, \frac{k}{2}-\frac{d}{2}+1, \ldots, \frac{k}{2}-1, \frac{k}{2}) \\ 1 \end{array} \right\}.$$

The standardized information matrix for $\boldsymbol{\beta}$ at the design $\tilde{\xi}^*_{D_e}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}^*_{D_e}) = \begin{pmatrix} \dfrac{1}{1+d\gamma} & 0 \\ 0 & \dfrac{1}{12}H \end{pmatrix}$$

where $H = d^2 - 3\,d\,(k+1) + 3\,k^2 + 6\,k + 2$. Note that the term $H$ is positive because the information matrix $\mathbf{M}_\beta(\tilde{\xi}^*_D)$ is necessarily positive definite. It then follows that

$$\mathbf{M}_\beta(\tilde{\xi}^*_{D_e})^{-1} = \begin{pmatrix} 1+d\gamma & 0 \\ 0 & \dfrac{12}{H} \end{pmatrix}$$

and that

$$\left|\mathbf{M}_\beta(\tilde{\xi}^*_{D_e})\right| = \frac{H}{12\,(1+d\gamma)}. \tag{4.12}$$

Consider now the directional derivative of $\Psi_D(\tilde{\xi}) = \ln\left|\mathbf{M}_\beta(\tilde{\xi})\right|$ at the population design $\tilde{\xi}^*_{D_e}$ in the direction of the single design $\tilde{\mathbf{t}}$, that is

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = tr\left[\mathbf{M}_\beta(\tilde{\xi}^*_{D_e})\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\right] - 2$$

where

$$\mathbf{M}_\beta(\tilde{\mathbf{t}}) = \begin{pmatrix} \dfrac{1}{1+d\gamma} & \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d(1+d\gamma)} \\ \\ \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d(1+d\gamma)} & \dfrac{1}{d}\tilde{\mathbf{t}}'(\mathbf{I} - \dfrac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}} \end{pmatrix}$$

is the standardized information matrix for $\boldsymbol{\beta}$ at a $d$-point design $\tilde{\mathbf{t}}$. Then this derivative is given explicitly by

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = \frac{1}{H}\, 12\, \{\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}}\} - 1. \tag{4.13}$$

By the Equivalence Theorem for $D$-optimal population designs, the design $\tilde{\xi}^*_{D_e}$ is $D$-optimal if and only if $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) \leq 0$ for all $d$-point designs $\tilde{\mathbf{t}}$ in $\tilde{S}_{d,k}$ with equality holding at the support designs of $\tilde{\xi}^*_{D_e}$. Consider therefore the nature of $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e})$. The matrix $(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})$ is symmetric with $(d-1)$ eigenvalues equal to 1 and one eigenvalue equal to $\dfrac{1}{1+d\gamma}$. Thus the quadratic form $\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\,\tilde{\mathbf{t}}$ is positive definite and the directional derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e})$ is a convex function over the polytope $\tilde{P}_{d,k}$ enclosing all the $d$-point designs in $\tilde{S}_{d,k}$. It now follows immediately from a well known result in nonlinear optimization theory that the maxima of $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e})$ occur at one or more of the vertices of the polytope $\tilde{P}_{d,k}$ (Bazaraa and Shetty, 1979, page 92).

The optimality of $\tilde{\xi}^*_{D_e}$ can therefore be confirmed by examining the values of the directional derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e})$ at the vertices of $\tilde{P}_{d,k}$. Recall from Section 4.2 that these vertices are the designs $\mathbf{v}_1, \ldots, \mathbf{v}_{d+1}$ defined in expression (4.4) and are given in transformed coordinates as

$$\tilde{\mathbf{v}}_1 = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+d-2, -\frac{k}{2}+d-1)$$

$$\vdots$$

$$\tilde{\mathbf{v}}_{j+1} = \left(-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+d-j-1, \frac{k}{2}-j+1, \ldots, \frac{k}{2}-1, \frac{k}{2}\right)$$

where $j = 1, \ldots, d-1$ and

$$\tilde{\mathbf{v}}_{d+1} = \left(\frac{k}{2}-d+1, \frac{k}{2}-d, \ldots, \frac{k}{2}-1, \frac{k}{2}\right).$$

It is straightforward to show that the directional derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e})$ at $\tilde{\mathbf{v}}_{j+1}$ is given explicitly by

$$\phi(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_D) = -\frac{1}{H} 3\,(d-2j)^2\,(k+1-d)\,\{1+(k+1)\,\gamma\}. \tag{4.14}$$

Consider the terms in this expression. Clearly $1+(k+1)\,\gamma > 0$ and also $k+1-d > 0$ since $d < k+1$. In addition $(d-2j)^2 \geq 0$ for $j = 0, 1, \ldots, d$ and as has been explained previously the denominator $H$ is always positive. Thus it follows that $\phi(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{D_e}) \leq 0$ at the vertices $\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_{d+1}$ and thus at all possible $d$-point designs $\tilde{\mathbf{t}}$ in $\tilde{S}_{d,k}$. Equality holds at $j = \frac{d}{2}$ and thus at the support design of $\tilde{\xi}^*_{D_e}$. Thus by the Equivalence Theorem of Subsection 2.6.4 $\xi^*_{D_e}$ is the $D$-optimal population design over the set $S_{d,k}$ for $\boldsymbol{\beta}$ when $d$ is even for all $\gamma \geq 0$.

□

**Theorem 4.3.3** *Consider the set of all d-point individual designs which put equal weights on the distinct time points $t_1, t_2, \ldots, t_d$ with $t_j \in \{0, 1, \ldots, k\}, j = 1, \ldots, d$ and $0 \leq t_1 < t_2 < \ldots < t_d \leq k$ for $d$ an odd integer greater than or equal to 3. Then*

$$\xi^*_{D_o} = \left\{ \begin{array}{cc} (0, 1, \ldots, \frac{d-1}{2}, k-\frac{d-3}{2}, \ldots, k) & (0, 1, \ldots, \frac{d-3}{2}, k-\frac{d-1}{2}, \ldots, k) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

*is the $D$-optimal population design for the fixed effects $\boldsymbol{\beta}$ in the model (4.1) over this set for all $\gamma \geq 0$ .*

**Proof**

The proof of this theorem is similar to that of Theorem 4.3.2 and it outlined briefly below.

Note that for $d = k + 1$ with $k$ even there is only one $d$-point individual design so this is necessarily optimal. In the remainder of the proof $d$ is taken to be strictly less than $k + 1$.

Consider the individual designs $\mathbf{t}$ linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$, where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$ is the center of the hypercube $C_{d,k}$. Then the proposed optimum design $\xi^*_{D_o}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{D_o} = \left\{ \begin{array}{cc} (-\frac{k}{2}, 1 - \frac{k}{2}, \ldots, -\frac{k}{2} + \frac{d-1}{2}, \frac{k}{2} - \frac{d-3}{2}, \ldots, \frac{k}{2}) & \\ \frac{1}{2} & \\ & (-\frac{k}{2}, 1 - \frac{k}{2}, \ldots, -\frac{k}{2} + \frac{d-3}{2}, \frac{k}{2} - \frac{d-1}{2}, \ldots, \frac{k}{2}) \\ & \frac{1}{2} \end{array} \right\}.$$

The standardized information matrix for $\boldsymbol{\beta}$ at the population design $\tilde{\xi}^*_{D_o}$ can be expressed as

$$\mathbf{M}_\beta(\tilde{\xi}^*_{D_o}) = \begin{pmatrix} \dfrac{1}{1 + d\gamma} & 0 \\ 0 & \dfrac{H}{12\, d\, (1 + d\gamma)} \end{pmatrix}$$

where

$$H = d^3 - 3\,(k+1)\,(d^2+1) + (3\,k^2 + 6\,k + 5)\,d + (d^2 - 1)\,\{d^2 + 3\,(k+1)\,(k - d + 1)\}\,\gamma.$$

Note immediately that the term $H$ is greater than zero because the information matrix $\mathbf{M}_\beta(\tilde{\xi}^*_{D_o})$ is necessarily positive definite. Further note that

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_o}) = \begin{pmatrix} 1 + d\gamma & 0 \\ 0 & \dfrac{12\, d\, (1 + d\gamma)}{H} \end{pmatrix}$$

and that

$$\left|\mathbf{M}_\beta(\tilde{\xi}^*_{D_o})\right| = \frac{H}{12\,d\,(1+d\,\gamma)^2}. \tag{4.15}$$

The directional derivative of $\Psi_D(\tilde{\xi}) = \ln\left|\mathbf{M}_\beta(\tilde{\xi})\right|$ at $\tilde{\xi}^*_{D_o}$ in the direction of a $d$-point individual design $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$ is given by

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_o}) = tr\left[\mathbf{M}_\beta(\tilde{\xi}^*_{D_o})^{-1}\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\right] - 2 = \frac{1}{H}\,12(1+d\,\gamma)\{\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}}\} - 1 \tag{4.16}$$

where $\mathbf{M}_\beta(\tilde{\mathbf{t}})$ is the standardized information matrix for $\boldsymbol{\beta}$ at a $d$-point design $\tilde{\mathbf{t}}$.

The expression for $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_o})$ in (4.16) is a convex function on the polytope $\tilde{P}_{d,k}$. Thus the maxima of $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_o})$ occur at the vertices of this polytope. Now the general expression for a vertex in the transformed coordinates is given by $\tilde{\mathbf{v}}_{j+1} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+d-j-1, \frac{k}{2}-j+1, \ldots, \frac{k}{2})$, $j = 0, 1, \ldots, d$ and the directional derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_o})$ at $\tilde{\mathbf{v}}_{j+1}$ can be expressed explicitly by

$$\phi(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{D_o}) = -\frac{1}{H}\,3\,(d-2j-1)\,(d-2j+1)\,(k+1-d)\,\{1+(k+1)\,\gamma\}.$$

Now consider the terms in $\phi(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{D_o})$. Clearly $1+(k+1)\,\gamma > 0$ and also $k+1-d > 0$ since $d < k+1$. In addition $(d-2j-1)\,(d-2j+1) = (d-2j)^2 - 1 \geq 0$ for $j = 0, 1, \ldots, d$ since $d \geq 3$ and as mentioned earlier the denominator $H$ is always positive. Thus it follows that $\phi(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{D_o}) \leq 0$ at the vertices $\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_{d+1}$ and thus at all possible $d$-point designs $\tilde{\mathbf{t}}$ in $\tilde{S}_{d,k}$. Equality holds only at $j = \dfrac{d-1}{2}$ and $\dfrac{d+1}{2}$ and thus at the support designs of $\tilde{\xi}^*_{D_o}$.
$\square$

The population designs in Theorems 4.3.2 and 4.3.3 are optimal for all $\gamma \geq 0$, so that no prior knowledge of the variance components is needed. In other words, the $D$-optimal population designs based on the $d$-point individual designs are robust to the choice of variance components.

**Example 4.3.1** Consider the simple linear regression model with a random intercept. For $k = 4$ and $d = 3$ the design

$$\xi^*_{D_o} = \left\{ \begin{array}{cc} (0,1,4) & (0,3,4) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

is the $D$-optimal population design for $\gamma \geq 0$. In Figure 4.6 a graph of directional derivative $\phi(\mathbf{t}, \xi^*_D)$ against the individual designs $(0,1,2)$, $(0,1,3)$, $(0,1,4)$, $(0,2,3)$, $(0,2,4)$,$(0,3,4)$, $(1,2,3)$, $(1,2,4)$, $(1,3,4)$ and $(2,3,4)$, labelled for convenience 1 through 10 respectively, is presented. It is clear from this figure that the condition $\phi(\mathbf{t}, \xi^*_{D_o}) \leq 0$ is satisfied for all designs $\mathbf{t} \in S_{3,4}$ and that equality holds at the support designs of $\xi^*_{D_o}$. Thus $\xi^*_{D_o}$ is the $D$-optimal population design.

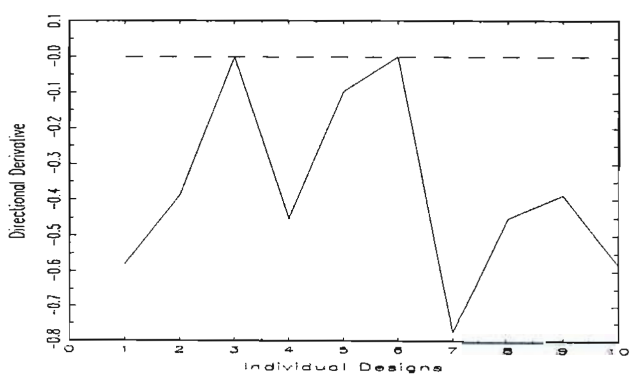


Figure 4.5: Plot of the directional derivative $\phi(\mathbf{t}, \xi^*_{D_o})$ against the individual designs $\mathbf{t}$.

The geometry pertaining to the proofs of Theorem 4.3.2 is interesting. Consider first the case of $d = 2$ for which the $D$-optimal population design $\tilde{\xi}^*_{D_e}$ puts weight 1 at the design $(-\frac{k}{2}, \frac{k}{2})$. The directional derivative at $\tilde{\xi}^*_{D_e}$ in the direction of $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2)$, where $\tilde{t}_j = t_j - \frac{k}{2}, j = 1, 2$, has the form

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = \frac{2}{k^2}\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+2\gamma}\mathbf{J})\tilde{\mathbf{t}} - 1.$$

For $\gamma = 0$ this simplifies to

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = \frac{2}{k^2}(\tilde{t}_1^2 + \tilde{t}_2^2) - 1.$$

Then the contours

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = \frac{2}{k^2}(\tilde{t}_1^2 + \tilde{t}_2^2) - 1 = c$$

for some constant $c$ are concentric circles of radius $k\sqrt{\frac{c+1}{2}}$ for $c \geq -1$. Thus the circle through $\tilde{\mathbf{t}} = (-\frac{k}{2}, \frac{k}{2})$ has $c = 0$ and encloses the square $\tilde{C}_{2,k}$ and hence the polytope $\tilde{P}_{2,k}$ and thus all designs in $\tilde{S}_{2,k}$ have $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) \leq 0$. Otherwise, for $\gamma > 0$, $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = c$ is an ellipse. The major axis of this ellipse is defined by the eigenvector $\mathbf{1}_2$ and has length $\frac{k}{2}\sqrt{1+2\gamma}\sqrt{c+1}$. The minor axis is defined by the vector $(1, \quad -1)'$ and has length $\frac{k}{2}\sqrt{c+1}$. Then the contours $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = c$ for $c \geq -1$ are concentric ellipsoids. Thus the ellipse through $(-\frac{k}{2}, \frac{k}{2})$ has $c = 0$ and encloses the square $\tilde{C}_{2,k}$ and hence the polytope $\tilde{P}_{2,k}$ and thus all designs in $\tilde{S}_{2,k}$ have $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) \leq 0$. Figure 4.6 illustrates this for $k = 10$ and $\gamma = 1$.

In general, for $d$ even the vertex $\tilde{\mathbf{v}}_{\frac{d}{2}+1} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d}{2}-1, \frac{k}{2}-\frac{d}{2}+1, \frac{k}{2})$ of the polytope $\tilde{P}_{d,k}$ is further from the origin than all other vertices. Consider the ellipse defined by $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) = c$ for some constant $c$, which passes through $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$. The vertex $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ lies on the plane perpendicular to $\mathbf{1}_d$ and is the only such vertex of $\tilde{P}_{d,k}$. Thus $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ is the closet point on that ellipse to the origin. It thus follows that all vertices of $\tilde{P}_{d,k}$ fall inside that ellipsoid and hence $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_e}) \leq 0$ for all design points $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$.

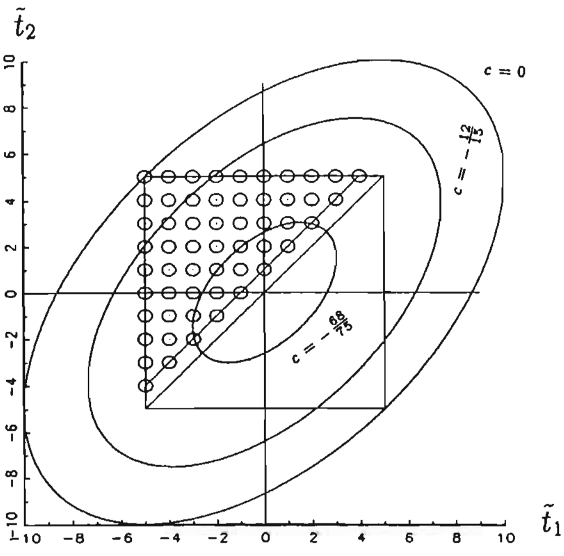


Figure 4.6: Contours of constant $\phi(\tilde{\mathbf{t}}, \tilde{\xi}_D^*)$ for $k = 10$ and $\gamma = 1$. The square $\tilde{C}_{2,10}$ is defined by the vertices (-5, -5), (-5, 5), (5, 5) and (5, -5). The symbol $\circ$ represents a two-point design in $\tilde{S}_{2,10}$ and the polytope $\tilde{P}_{2,10}$ is the triangle with vertices (-5,-4), (-5, 5) and (4, 5).

When $d$ is odd similar arguments to the $d$ even case can be used. However, the two vertices $\tilde{\mathbf{v}}_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}_{\frac{d+3}{2}}$ of $\tilde{\xi}_{D_o}^*$ do not lie on the plane perpendicular to $\mathbf{1}_d$ but are closest to that plane. These vertices are farthest from the origin.

### 4.3.3 The best $D$-optimal population design for the fixed effects

In this section, the best $D$-optimal population design on a per observation basis is discussed. For example, consider model (4.1) for $k = 6$. From Theorem 4.3.2 and Theorem 4.3.3 it follows that the two-, three-, four-, five-, six- and seven-point $D$-optimal population designs are

$$\xi^*_{D_2} = \left\{ \begin{array}{c} (0,6) \\ 1 \end{array} \right\}, \qquad \xi^*_{D_3} = \left\{ \begin{array}{cc} (0,1,6) & (0,5,6) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\},$$

$$\xi^*_{D_4} = \left\{ \begin{array}{c} (0,1,5,6) \\ 1 \end{array} \right\} \qquad \xi^*_{D_5} = \left\{ \begin{array}{cc} (0,1,2,5,6) & (0,1,4,5,6) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

$$\xi^*_{D_6} = \left\{ \begin{array}{c} (0,1,2,4,5,6) \\ 1 \end{array} \right\} \quad \text{and} \quad \xi^*_{D_7} = \left\{ \begin{array}{c} (0,1,2,3,4,5,6) \\ 1 \end{array} \right\}.$$

Note that there is only one seven-point design $\xi^*_{D_7}$ for $k = 6$ and it is necessarily optimal. The associated determinants of the above designs are

$$\left|\mathbf{M}_\beta(\xi^*_{D_2})\right| = \frac{9}{1+2\gamma}, \qquad \left|\mathbf{M}_\beta(\xi^*_{D_3})\right| = \frac{2\,(11+31\,\gamma)}{3\,(1+3\,\gamma)^2},$$

$$\left|\mathbf{M}_\beta(\xi^*_{D_4})\right| = \frac{13}{2\,(1+4\,\gamma)}, \qquad \left|\mathbf{M}_\beta(\xi^*_{D_5})\right| = \frac{27+134\,\gamma}{5\,(1+5\,\gamma)^2},$$

$$\left|\mathbf{M}_\beta(\xi^*_{D_6})\right| = \frac{14}{3\,(1+6\,\gamma)} \quad \text{and} \quad \left|\mathbf{M}_\beta(\xi^*_{D_7})\right| = \frac{4}{1+7\,\gamma}.$$

Since

$$\frac{\left|\mathbf{M}_\beta(\xi^*_{D_2})\right|}{\left|\mathbf{M}_\beta(\xi^*_{D_3})\right|} = \frac{27\,(1+3\,\gamma)^2}{2\,(1+2\,\gamma)\,(11+31\,\gamma)} > 1, \qquad \frac{\left|\mathbf{M}_\beta(\xi^*_{D_2})\right|}{\left|\mathbf{M}_\beta(\xi^*_{D_4})\right|} = \frac{18\,(1+4\,\gamma)}{13\,(1+2\,\gamma)} > 1$$

$$\frac{\left|\mathbf{M}_\beta(\xi^*_{D_2})\right|}{\left|\mathbf{M}_\beta(\xi^*_{D_5})\right|} = \frac{45\,(1+5\,\gamma)^2}{(1+2\,\gamma)\,(27+134\,\gamma)} > 1, \qquad \frac{\left|\mathbf{M}_\beta(\xi^*_{D_2})\right|}{\left|\mathbf{M}_\beta(\xi^*_{D_6})\right|} = \frac{27\,(1+6\,\gamma)}{14\,(1+2\,\gamma)} > 1$$

and

$$\frac{\left|\mathbf{M}_\beta(\xi^*_{D_2})\right|}{\left|\mathbf{M}_\beta(\xi^*_{D_7})\right|} = \frac{9\,(1+7\,\gamma)}{4\,(1+2\,\gamma)} > 1$$

the $D$-optimal population design $\xi^*_{D_2}$ is best for estimating $\boldsymbol{\beta}$ when it compared with $\xi^*_{D_3}$, $\xi^*_{D_4}$, $\xi^*_{D_5}$, $\xi^*_{D_6}$ and $\xi^*_{D_7}$. From the above example and other numeric studies for a range of $k$

it can be speculated that the population design which puts weight 1 on $(0, k)$ is $D$-optimal for the fixed effects $\boldsymbol{\beta}$ in model (4.1) over the set of all possible population designs. This is proved in the following theorem.

**Theorem 4.3.4** *Consider the set of population designs based on all possible individual designs* $\mathbf{t}$ *which put equal weights on the distinct time points* $t_1, t_2, \ldots, t_d$ *with* $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$ *and* $0 \leq t_1 < t_2 < \ldots < t_d \leq k$ *for* $d$ *any positive integer less than or equal to* $k+1$. *Then the design*

$$\xi^*_{D_2} = \left\{ \begin{array}{c} (0,k) \\ \\ 1 \end{array} \right\}$$

*is the* $D$*-optimal population design for the fixed effects* $\boldsymbol{\beta}$ *in model (4.1) over this set for all* $\gamma \geq 0$.

**Proof**

Consider the individual designs $\mathbf{t}$ linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$, where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$ is the center of the hypercube $C_{d,k}$, i.e. $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$. Then the proposed optimum design $\xi^*_{D_2}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{D_2} = \left\{ \begin{array}{c} (-\frac{k}{2}, \frac{k}{2}) \\ \\ 1 \end{array} \right\}.$$

Note immediately that the information matrix for $\boldsymbol{\beta}$ at $\tilde{\xi}^*_{D_2}$ is equal to

$$\mathbf{M}_{\beta}(\tilde{\xi}^*_{D_2}) = \frac{1}{\sigma_e^2} \begin{pmatrix} \dfrac{1}{1+2\gamma} & 0 \\ 0 & \dfrac{k^2}{4} \end{pmatrix}.$$

Since $\dfrac{1}{\sigma_e^2}$ factors out $\sigma_e^2$ can be taken to be 1 without the loss of generality and hence

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_2}) = \begin{pmatrix} 1+2\gamma & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}.$$

Now it follows from the Equivalence Theorem of Subsection 2.6.4 that to prove the present theorem, it is only necessary to show that the directional derivative of the criterion $\Psi_D(\tilde{\xi}) = \ln|\mathbf{M}_\beta(\tilde{\xi})|$ at $\tilde{\xi}^*_{D_2}$ in the direction of an individual design $\tilde{\mathbf{t}}$, that is

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2}) = tr\left[\mathbf{M}_\beta(\tilde{\xi}^*_{D_2})^{-1}\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\right] - 2,$$

is less than or equal to zero with equality holding at the support design of $\tilde{\xi}^*_{D_2}$. Further, recall that for a $d$-point individual design $\tilde{\mathbf{t}}$

$$\mathbf{M}_\beta(\tilde{\mathbf{t}}) = \begin{pmatrix} \dfrac{1}{1+d\gamma} & \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d(1+d\gamma)} \\ \\ \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d(1+d\gamma)} & \dfrac{1}{d}\tilde{\mathbf{t}}'(\mathbf{I} - \dfrac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}} \end{pmatrix}$$

so that

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2}) = \frac{4\,\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\,\tilde{\mathbf{t}}}{d\,k^2} - \frac{2\,(d-1)\,\gamma + 1}{1+d\gamma}. \tag{4.17}$$

Consider first an individual design comprising a single point $\tilde{t}$ where $-\frac{k}{2} \le \tilde{t} \le \frac{k}{2}$. Then the directional derivative of the criterion $\Psi_D(\tilde{\xi})$ at $\tilde{\xi}^*_{D_2}$ in the direction of $\tilde{t}$ is

$$\phi(\tilde{t}, \tilde{\xi}^*_{D_2}) = \frac{(2\,\tilde{t} + k)\,(2\,\tilde{t} - k)}{(1+\gamma)\,k^2}$$

and this is less than or equal to zero since $|\tilde{t}| \le \frac{k}{2}$.

For the case $d = 2$, it has already been shown in Theorem 4.3.2 that the design $\tilde{\xi}^*_{D_2}$ is the $D$-optimal population design on the set of two-point individual designs, i.e. on the space of designs $\tilde{S}_{2,k}$. It thus follows that $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2}) \le 0$ for any two-point design $\tilde{\mathbf{t}} \in \tilde{S}_{2,k}$.

Finally, consider $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2})$ for all $d$-point designs $\tilde{\mathbf{t}}$ with $d \geq 3$. In particular consider the geometry of $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2})$ in expression (4.17). Then the values of $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2})$ associated with contours

$$\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\,\tilde{\mathbf{t}} = c$$

where $c$ is a constant are of exactly the same form as those described in the proofs of Theorem 4.3.2 and Theorem 4.3.3. So it follows immediately that the largest value of $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2})$ occurs at the vertex $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ when $d$ is even and at the vertices $\tilde{\mathbf{v}}_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}_{\frac{d+3}{2}}$ when $d$ is odd. Thus it is only necessary to check the condition $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2}) \leq 0$ at these vertices.

Consider first the case of $d$ even and $d > 2$. Recall that the vertex

$$\tilde{\mathbf{v}}_{\frac{d}{2}+1} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d}{2}-1, \frac{k}{2}-\frac{d}{2}+1, \frac{k}{2}-\frac{d}{2}+2, \ldots, \frac{k}{2}-1, \frac{k}{2})$$

is the support design for the $D$-optimal population design in Theorem 4.3.2 expressed in the transformed coordinates. At $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ the directional derivative is given by

$$\phi(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*_{D_2}) = -\frac{d-2}{3\,k^2\,(1+d\,\gamma)}\,\{3\,k^2\,\gamma + (3\,k - d + 1)\,(d\,\gamma + 1)\}$$

and this is strictly less than 0 since $d - 2 > 0$, $3\,k^2\,\gamma + (3\,k - d + 1)\,(d\,\gamma + 1) > 0$ and $3\,k^2\,(1 + d\,\gamma) > 0$ for $d \leq k+1$, $k \geq 2$ and $\gamma \geq 0$.

Consider now the case of $d$ odd with $d \geq 3$. Recall also that the vertices

$$\tilde{\mathbf{v}}_{\frac{d+1}{2}} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d-1}{2}, \frac{k}{2}-\frac{d-3}{2}, \ldots, \frac{k}{2}-1, \frac{k}{2})$$

and

$$\tilde{\mathbf{v}}_{\frac{d+3}{2}} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d-3}{2}, \frac{k}{2}-\frac{d-1}{2}, \ldots, \frac{k}{2}-1, \frac{k}{2})$$

are the support designs for the $D$-optimal population design in Theorem 4.3.3 in the transformed coordinates. At both $\tilde{\mathbf{v}}_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}_{\frac{d+3}{2}}$, the directional derivative is

$$\phi(\tilde{\mathbf{v}}_{\frac{d+1}{2}}, \tilde{\xi}^*_{D_2}) = \phi(\tilde{\mathbf{v}}_{\frac{d+3}{2}}, \tilde{\xi}^*_{D_2}) = \frac{d-1}{3\,d\,k^2\,(1+d\,\gamma)}\,\{C_1(d)\,\gamma + C_0(d)\}$$

where

$$C_0(d) = d^2 - (3\,k+2)\,d + 3\,(k+1)$$

and

$$C_1(d) = d^3 - d^2\,(3\,k+2) - 3\,d\,(k-1)\,k + 3\,(k+1)^2.$$

The term $C_1(d)$ is a cubic in $d$ with stationary points given by the solutions to the equation

$$\frac{\partial C_1(d)}{\partial d} = 3\,d^2 - 3\,(k-1)\,k - 2\,d\,(3\,k+2) = 0$$

and thus by $d_1 = \frac{1}{3}(3\,k+2-\sqrt{4+3\,k+18\,k^2})$ and $d_2 = \frac{1}{3}(3\,k+2+\sqrt{4+3\,k+18\,k^2})$. Since

$$\left.\frac{\partial^2 C_1(d)}{\partial\,d^2}\right|_{d=d_1} = -2\,\sqrt{4+3\,k+18\,k^2} < 0$$

and

$$\left.\frac{\partial^2 C_1(d)}{\partial\,d^2}\right|_{d=d_2} = 2\,\sqrt{4+3\,k+18\,k^2} > 0$$

$C_1(d)$ attains its maximum and minimum at $d_1$ and $d_2$ respectively. Since $3\,k+2>0$ and $\sqrt{4+3\,k+18\,k^2}>0$, the inequality $(3\,k+2)^2-(4+3\,k+18\,k^2) = -9\,k\,(k-1) < 0$ implies that $3\,k+2 < \sqrt{4+3\,k+18\,k^2}$. Thus the root $d_1$ is always negative. Clearly, the root $d_2$ is always positive. Furthermore, since

$$d_2 - (k+1) = \frac{1}{3}(\sqrt{4+3\,k+18\,k^2} - 1) > 0 \;\; \text{for} \;\; k \geq 2$$

$d_2$ is greater than $k+1$. So both $d_1$ and $d_2$ are not in the range of interest for $d$, that is, the interval $[3, k+1]$. Therefore the maximum and minimum of $C_1(d)$ fall outside $[3, k+1]$. At

$d$ values of 3 and $k+1$ and for $k \geq 2$, $C_1(d)$ has the form $C_1(3) = -6\,(k^2+2\,k-2) \leq 0$ and $C_1(k+1) = -(5\,k+2)\,(k+1)\,(k-1) \leq 0$ respectively. Thus $C_1(d) \leq 0$ for all $d \in [3, k+1]$. The graph of $C_1(d)$ against $d$ for $k = 5$ is shown in Figure 4.7. The minimum of $C_1(d)$ in the graph occurs at -1.5521 and the maximum at 12.8855 but these values fall outside the prescribed range for $d$ of [3, 6].

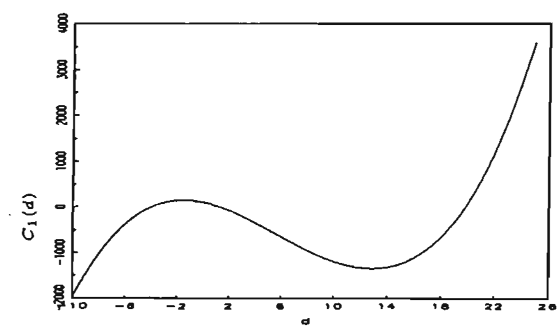

Figure 4.7: Graph of $C_1(d)$ for $k = 5$.

The term $C_0(d)$ is a quadratic in $d$ and, since the coefficient of $d^2$ is positive, represents a parabola which opens upwards. Furthermore, $C_0(3) = -6\,(k-1) < 0$ and $C_0(k+1) = -2\,(k+1)\,(k-1) < 0$ for $k \geq 2$. Thus $C_0(d)$ is necessarily less than 0 for $d \in [3, k+1]$.

Overall, since both $C_1(d)$ and $C_0(d)$ are negative for $d \in [3, k+1]$, the directional derivatives, $\phi(\tilde{\mathbf{v}}_{\frac{d+1}{2}}, \tilde{\xi}^*_{D_2})$ and $\phi(\tilde{\mathbf{v}}_{\frac{d+3}{2}}, \tilde{\xi}^*_{D_2})$, are less than or equal to zero. $\square$

### 4.3.4 $D$-efficiencies

The population designs $\xi$ based on $d$-point individual designs can be compared with the best $D$-optimal design $\xi^*_{D_2}$ according to efficiency. For $D$-optimality, the efficiency of the population design $\xi$ is defined as

$$D_{eff} = \left\{ \frac{|\mathbf{M}_\beta(\xi)|}{|\mathbf{M}_\beta(\xi^*_{D_2})|} \right\}^{\frac{1}{2}}.$$

Recall from Subsection 4.3.2 that the determinant of the information matrix for the fixed effects $\boldsymbol{\beta}$, $\mathbf{M}_\beta(\xi)$, at the optimal designs $\xi^*_{D_2}$, $\xi^*_{D_e}$ and $\xi^*_{D_o}$ respectively are

$$|\mathbf{M}_\beta(\xi^*_{D_2})| = \frac{k^2}{4\,(1+2\,\gamma)},$$

$$|\mathbf{M}_\beta(\xi^*_{D_e})| = \frac{1}{12\,(1+d\,\gamma)}\{d^2 - 3\,d\,(k+1) + 3\,k^2 + 6\,k + 2\}$$

and

$$|\mathbf{M}_\beta(\xi^*_{D_o})| = \frac{H}{12\,d\,(1+d\,\gamma)^2}$$

where

$$H = d^3 - 3\,(k+1)\,(d^2+1) + (3\,k^2 + 6\,k + 5)\,d + (d^2-1)\,\{d^2 + 3\,(k+1)\,(k-d+1)\}\,\gamma.$$

Thus, the efficiencies of the $D$-optimal population designs $\xi^*_{D_e}$ and $\xi^*_{D_o}$ with respect to the best $D$-optimal design $\xi^*_{D_2}$ are given by

$$D_{eff(1)} = \left\{ \frac{|\mathbf{M}_\beta(\xi^*_{D_e})|}{|\mathbf{M}_\beta(\xi^*_{D_2})|} \right\}^{\frac{1}{2}} = \left\{ \frac{\{d^2 - 3\,d\,(k+1) + 3\,k^2 + 6\,k + 2\}(1+2\,\gamma)}{3\,k^2\,(1+d\,\gamma)} \right\}^{\frac{1}{2}}$$

and

$$D_{eff(2)} = \left\{ \frac{|\mathbf{M}_\beta(\xi^*_{D_o})|}{|\mathbf{M}_\beta(\xi^*_{D_2})|} \right\}^{\frac{1}{2}} = \left\{ \frac{H\,(1+2\,\gamma)}{3\,k^2\,d\,(1+d\,\gamma)^2} \right\}^{\frac{1}{2}},$$

respectively.

Plots of the $D$-efficiencies $D_{eff(1)}$ and $D_{eff(2)}$ against $\gamma$ for $k = 10$ and appropriate values of $d$ in the interval [3, 11] are presented in Figures 4.8 and 4.9 respectively. It is clear from these plots that for fixed $d$ the $D$-efficiencies decrease as $\gamma$ increases and that for fixed $\gamma$ the $D$-efficiencies increase as the number of time points $d$ decreases. However, the gain in efficiency achieved by taking designs with small $d$ decreases as $\gamma$ increases. This point is explored in more detail in the next subsection. Note that for all values of $k$, the $d$-point $D$-optimal population designs themselves do not depend on $\gamma$ but their attendant $D$-efficiencies do.

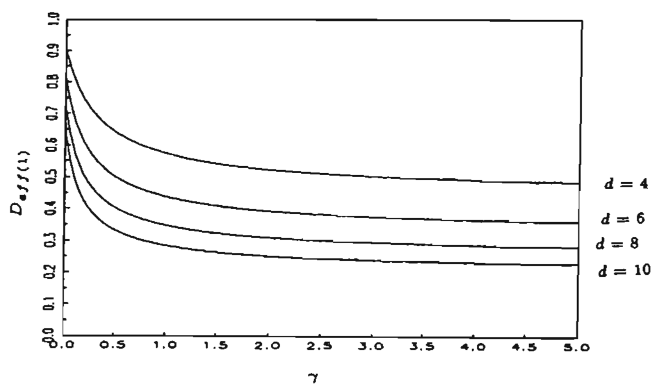


Figure 4.8: $D$-efficiencies of the $d$-point $D$-optimal population designs against $\gamma$ for $k = 10$ and $d = (4, 6, 8, 10)$.

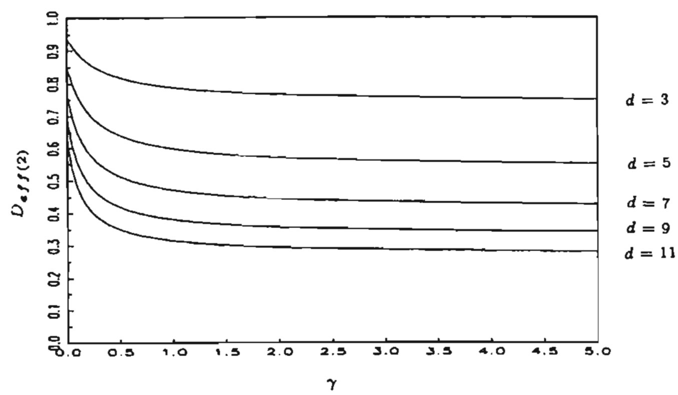


Figure 4.9: $D$-efficiencies of the $d$-point $D$-optimal population designs plotted against $\gamma$ for $k = 10$ and $d = (3, 5, 7, 9, 11)$.

### 4.3.5 Comparison of $D$-optimal population designs

In Subsection 4.3.3 it was proved that the population design comprising the design $(0, k)$ is the best $D$-optimal population design over all $d$-point population designs for $d$ any positive integer less than or equal to $k + 1$. Sometimes however the researcher may be interested in estimating model parameters as precisely as possible based on measurements taken at more than two time points. For such cases the above result does not answer the researcher's needs. For example, consider a comparison between the three- and four-point $D$-optimal population designs $\xi^*_{D_3}$ and $\xi^*_{D_4}$. For $\xi^*_{D_3}$ to be more efficient than $\xi^*_{D_4}$ on a per observation basis the determinant

$$|\mathrm{M}_\beta(\xi^*_{D_3})| = \frac{4 + 8\gamma - 4k(1 + 2\gamma) + k^2(3 + 8\gamma)}{12(1 + 3\gamma)^2}$$

must be greater than the determinant

$$|\mathbf{M}_\beta(\xi^*_{D_4})| = \frac{2 - 2\,k + k^2}{4\,(1 + 4\,\gamma)}.$$

Thus the condition

$$\frac{|\mathbf{M}_\beta(\xi^*_{D_3})|}{|\mathbf{M}_\beta(\xi^*_{D_4})|} = \frac{(1 + 4\,\gamma)\,\{4 + 8\,\gamma - 4\,k\,(1 + 2\,\gamma) + k^2\,(3 + 8\,\gamma)\}}{3\,(2 - 2\,k + k^2)\,(1 + 3\,\gamma)^2} > 1$$

and equivalently

$$(1 + 4\,\gamma)\,\{4 + 8\,\gamma - 4\,k\,(1 + 2\,\gamma) + k^2\,(3 + 8\,\gamma)\} - 3\,(2 - 2\,k + k^2)\,(1 + 3\,\gamma)^2 > 0$$

and thus

$$k^2\,\gamma\,(2 + 5\,\gamma) + 2\,(k - 1)\,(1 + 6\,\gamma + 11\,\gamma^2) > 0$$

must hold. For $k \geq 2$ with $\gamma \geq 0$ this latter inequality is clearly always true. Thus the parameters $\boldsymbol{\beta}$ in model (4.1) are estimated more precisely by $\xi^*_{D_3}$ than $\xi^*_{D_4}$. The general result for $d \geq 3$ is presented in the following theorem.

**Theorem 4.3.5** *Let the constants $d_e$ and $d_o$ be even and odd positive integers both greater than or equal to 3. Then the $D$-optimal population designs $\xi^*_{D_e}$ and $\xi^*_{D_o}$ are the most efficient designs on a per observation basis over the set of $d$-point individual designs defined on the space of designs $S_{d,k}$ with $d > d_e$ and $d > d_o$, respectively. Furthermore, the $D$-efficiency decreases on a per observation basis as $d$ increases.*

**Proof**

Let $m$ be a positive integer greater than or equal to one such that $d = 2\,m \leq k + 1$ when $d$ is an even positive integer and $d = 2\,m + 1 \leq k + 1$ when $d$ is an odd positive integer. To prove the theorem it is necessary to show that the inequalities

$$\left|\mathbf{M}_\beta(\xi^*_{D_{e(2m)}})\right| > \left|\mathbf{M}_\beta(\xi^*_{D_{o(2m+1)}})\right| > \left|\mathbf{M}_\beta(\xi^*_{D_{e(2m+2)}})\right| \tag{4.18}$$

hold for all $m$ where $\xi^*_{D_{e(2m)}}$, $\xi^*_{D_{e(2m+1)}}$ and $\xi^*_{D_{e(2m+2)}}$ are the $D$-optimal population designs based on individual designs with $2m$, $2m+1$ and $2m+2$ points respectively.

Consider first the inequality

$$\left|\mathbf{M}_\beta(\xi^*_{D_{e(2m)}})\right| > \left|\mathbf{M}_\beta(\xi^*_{D_{o(2m+1)}})\right|.$$

This inequality is true if and only if the ratio

$$\frac{\left|\mathbf{M}_\beta(\xi^*_{D_{e(2m)}})\right|}{\left|\mathbf{M}_\beta(\xi^*_{D_{o(2m+1)}})\right|} = \frac{D_1}{D_2} > 1 \tag{4.19}$$

where

$$D_1 = \{3\,k\,(k-2\,m) + 6\,(k-m) + 4\,m^2 + 2\}(2\,m+1)\,\{1 + (1+2\,m)\,\gamma\}^2$$

and

$$D_2 = \{4\,m\,(m+1)\,\{3\,k\,(k-2\,m) + 2\,m\,(2\,m-1) + 3\,k + 1\}\,\gamma + 6\,k\,m\,(k-2\,m)$$
$$+3\,k^2 + 4\,m\,(2\,m^2+1)\}\,(1 + 2\,m\,\gamma).$$

Observe that for $k \geq 2\,m$ and $m > 1$, it follows that $k - 2\,m \geq 0$ and $2\,m - 1 > 0$. Therefore for any $\gamma \geq 0$ and $m > 1$ the denominator $D_2$ is greater than zero. Thus the inequality in (4.19) is true if and only if

$$D_1 - D_2 = C_2\,\gamma^2 + C_1\,\gamma + C_0 > 0$$

where

$$C_0 = 2\,(1 + 3\,k - 4\,m)\,(m+1),$$

$$C_1 = 2\,(m+1)\,\{3\,k^2 + 6\,m\,(k-2\,m) + 2\,(3\,k-m) + 2\}$$

and

$$C_2 = 2\,(3\,k^2 - 4\,m^2)\,m\,(2\,m+3) + 4\,m^2\,(3\,k-4) + 3\,k^2 + 6\,k + 6\,m\,(5\,k+1) + 2.$$

Now, if $C_0 > 0$, $C_1 > 0$ and $C_2 > 0$ then $C_2\,\gamma^2 + C_1\,\gamma + C_0 > 0$ because $\gamma \geq 0$. Consider therefore the coefficients $C_0$, $C_1$ and $C_2$. Since $k \geq 2\,m$ and thus $k^2 \geq 4\,m^2$, it follows that $3\,k > 4\,m$ and hence that $C_0 > 0$, that $3\,k > m$ and hence that $C_1 > 0$ and that $3\,k^2 > 4\,m^2$ and hence that $C_2 > 0$.

Consider now the inequality

$$\left|\mathbf{M}_\beta(\xi^*_{D_{o(2m+1)}})\right| > \left|\mathbf{M}_\beta(\xi^*_{D_{e(2m+2)}})\right|.$$

This inequality is true if and only if the inequality

$$\frac{\left|\mathbf{M}_\beta(\xi^*_{D_{o(2m+1)}})\right|}{\left|\mathbf{M}_\beta(\xi^*_{D_{e(2m+2)}})\right|} = \frac{D_3}{D_4} > 1 \qquad (4.20)$$

holds, where

$$D_3 = \{4\,m\,(m+1)\,\{3\,k\,(k-2\,m) + 2\,m\,(2\,m-1) + 3\,k + 1\}\,\gamma + 6\,k\,m\,(k-2\,m)$$
$$+3\,k^2 + 4\,m\,(2\,m^2+1)\}\,\{2\,\gamma\,(m+1)+1\}$$

and

$$D_4 = (2\,m+1)\,\{3\,k\,(k-2\,m) + 2\,m\,(2\,m+1)\}\,\{1 + (2\,m+1)\,\gamma\}^2.$$

Observe here that for $k \geq 2\,m+1$, it follows that $k - 2\,m > 0$. Therefore for any $\gamma \geq 0$ and $m > 1$ the denominator $D_4$ is greater than zero. Thus the inequality (4.20) is true if and only if

$$D_3 - D_4 = C_2\,\gamma^2 + C_1\,\gamma + C_0 > 0$$

where

$$C_0 = 2\,(1 + 3\,k - 4\,m)\,m,$$

$$C_1 = 2\,m\{6\,m(k - 2\,m) + 2\,(6\,k - 5\,m) + 3\,k^2 + 4\}$$

and

$$C_2 = 4\,m^2\,\{(3\,k^2 - 4\,m^2) + (9\,k - 10\,m)\} + 3\,k^2\,(2\,m - 1) + 6\,m + 2\,m(15\,k - 8\,m).$$

Since $k \geq 2\,m$ and thus $k^2 \geq 4\,m^2$, it follows that $3\,k > 4\,m$ and hence that $C_0 > 0$, that $6\,k > 5\,m$ and hence that $C_1 > 0$ and that $3\,k^2 > 4\,m^2$, $9\,k > 10\,m$ and $15\,k > 8\,m$ and hence that $C_2 > 0$. □

## 4.4 $D$-optimal designs for the fixed effects $\beta$ based on designs with repeated time points

In the previous section the construction of $D$-optimal individual and population designs based on designs with non-repeated time points were considered. In this section, the problem of constructing the corresponding optimal designs based on individual designs for which replications of the same time point are possible, i.e. based on the space of designs given by the set

$$T_{d,k} = \{\mathbf{t} : \mathbf{t} = (t_1, \ldots, t_d),\ t_j \in \{0, 1, \ldots, k\},\ j = 1, \ldots, d,\ 0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k\}$$

is now examined.

### 4.4.1 $d$-point $D$-optimal individual designs

Consider a $d$-point individual design $\mathbf{t} = (t_1, t_2, \ldots, t_d)$ with $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$ not necessarily distinct, i.e. $0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k$. Recall from Subsection 4.2.2 that there are $\begin{pmatrix} k+d \\ d \end{pmatrix}$ such $d$-point individual designs in the set $T_{d,k}$. Recall also from Subsection 4.3.1 that the $d$-point design $\mathbf{t}^*$ is an exact $D$-optimal individual design if it maximizes

$$|\mathbf{M}_\beta(\mathbf{t})| = \frac{SS(t)}{d\,(1 + d\,\gamma)}.$$

Since $d\,(1 + d\,\gamma)$ factors out, maximizing the criterion involves maximizing

$$SS(t) = \mathbf{t}'\,(\mathbf{I} - \frac{1}{d}\,\mathbf{J})\,\mathbf{t}$$

independent of $\gamma$.

The general results for exact $D$-optimal individual designs based on $d$-points are presented in the following theorem.

**Theorem 4.4.1** *Consider the set of all d-point individual designs* $\mathbf{t} = (t_1, t_2, \ldots, t_d)$ *with* $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k$ *and d an integer in the interval* $[2, k+1]$. *Then the d-point D-optimal individual designs for the fixed effects* $\boldsymbol{\beta}$ *in the model (4.1) over this set are given by*

$$\mathbf{t}_e^* = (\underbrace{0, \ldots, 0}_{\frac{d}{2}\ times},\ \underbrace{k, \ldots, k}_{\frac{d}{2}\ times})$$

*for d even and either*

$$\mathbf{t}_{o1}^* = (\ \underbrace{0, \ldots, 0}_{\frac{d-1}{2}\ times}\ ,\ \ \underbrace{k, \ldots, k}_{\frac{d+1}{2}\ times}\ )$$

*or*

$$\mathbf{t}^*_{o2} = (\underbrace{0,\ldots,0}_{\frac{d+1}{2}\ times}, \underbrace{k,\ldots,k}_{\frac{d-1}{2}\ times})$$

*for d odd.*

**Proof**

Consider first the case of $d$ an even integer greater than or equal to 2. Since the $D$-optimal approximate design for the fixed effects in a simple linear regression model with uncorrelated errors puts weight $\frac{1}{2}$ on 0 and k (Atkinson and Donev, 1992, page 60) it thus follows from Atkins and Cheng (1999) that the best individual $D$-optimal design for any $\gamma$ when $d$ is an even integer is given by

$$\mathbf{t}^*_e = (\underbrace{0,\ldots,0}_{\frac{d}{2}\text{ times}}, \underbrace{k,\ldots,k}_{\frac{d}{2}\text{ times}}).$$

Consider now $d$ an odd integer greater than or equal to 3. Note from Subsection 4.3.1 that

$$SS(t) = \mathbf{t}'\,(\mathbf{I} - \frac{1}{d}\,\mathbf{J})\,\mathbf{t}$$

is a convex function on a polytope $Q_{d,k}$. Thus it is only necessary to check the maximum of $SS(t)$ at the vertices of $Q_{d,k}$. Recall from Section 4.2 that the vertices of the polytope $Q_{d,k}$ can be expressed as

$$\mathbf{v}^*_{j+1} = (\underbrace{0,\ldots,0}_{j\text{ times}}, \underbrace{k,\ldots,k}_{(d-j)\text{ times}})$$

where $j = 0, 1, 2, \ldots, d$. Then the general expression for $SS(t)$ for a vertex $\mathbf{v}^*_{j+1}$ is given by

$$SS(t) = \frac{1}{d}\,k^2\,(d-j)\,j \quad \text{for } j = 0, 1, \ldots, d.$$

Solving the equation

$$\frac{\partial SS(t)}{\partial j} = \frac{1}{d} k^2 (d - 2j) = 0$$

for the index $j$ yields $j = \frac{d}{2}$. Since the second derivative

$$\frac{\partial^2 SS(t)}{\partial j^2} = -\frac{2}{d} k^2 < 0$$

$SS(t)$ attains a maximum at $j = \frac{d}{2}$. When $d$ is an odd integer it follows from the fact that $SS(t)$ is a quadratic in $j$ that the maximum of $SS(t)$ occurs at the integers $\frac{d-1}{2}$ and $\frac{d+1}{2}$ and thus at the extreme vertices

$$\mathbf{v}^*_{\frac{d+1}{2}} = (\underbrace{0, \ldots, 0}_{\frac{d-1}{2} \text{ times}}, \underbrace{k, \ldots, k}_{\frac{d+1}{2} \text{ times}})$$

and

$$\mathbf{v}^*_{\frac{d+3}{2}} = (\underbrace{0, \ldots, 0}_{\frac{d+1}{2} \text{ times}}, \underbrace{k, \ldots, k}_{\frac{d-1}{2} \text{ times}}).$$

Observe that $\mathbf{v}^*_{\frac{d+1}{2}} = \mathbf{t}^*_{o1}$ and $\mathbf{v}^*_{\frac{d+3}{2}} = \mathbf{t}^*_{o2}$. This proof also holds for $d$ even. □

### 4.4.2 $D$-optimal population designs based on $d$-point individual designs

From the results for $D$-optimal individual designs presented in Theorem 4.4.1 it would seem intuitively reasonable to assume that a $D$-optimal population design based on $d$-point individual designs comprises the single design $\mathbf{t}^*_e$ for $d$ even and puts equal weights on the designs $\mathbf{t}^*_{o1}$ and $\mathbf{t}^*_{o2}$ for $d$ odd. This is proved in the following theorem.

**Theorem 4.4.2** *Consider the set of all d-point individual designs* $\mathbf{t}$ *with* $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k$ *and* $d$ *an integer in the interval* $[2, k+1]$. *Then the D-optimal population designs for the fixed effects* $\boldsymbol{\beta}$ *in model (4.1) over this set for all* $\gamma \geq 0$ *are given by*

$$\xi^*_{D_{re}} = \left\{ \begin{array}{c} (\underbrace{0, \ldots, 0}_{\frac{d}{2}\ times}, \underbrace{k, \ldots, k}_{\frac{d}{2}\ times}) \\ 1 \end{array} \right\}$$

*for d even and*

$$\xi^*_{Dro} = \left\{ \begin{array}{cc} (\underbrace{0, \ldots, 0}_{\frac{d+1}{2}\ times}, \underbrace{k, \ldots, k}_{\frac{d-1}{2}\ times}), & (\underbrace{0, \ldots, 0}_{\frac{d-1}{2}\ times}, \underbrace{k, \ldots, k}_{\frac{d+1}{2}\ times}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

*for d odd.*

**Proof**

Consider first the case of $d$ an even integer. Atkins and Cheng (1999) show that if a $D$-optimal approximate design for a polynomial model with uncorrelated errors corresponds to an exact design involving $m$ points, say $\xi^*_{(m)}$, then the $D$-optimal population design based the set of $mr$-point individual designs where $r = 1, 2, 3, \ldots$ for the corresponding model with a random intercept comprises a single design in which the exact design $\xi^*_{(m)}$ is replicated $r$ times. The approximate $D$-optimal design for a simple linear regression model with uncorrelated error associates weights 0.5 with the points 0 and $k$ and therefore corresponds to the exact designs comprising the two points 0 and $k$. It thus follows immediately from the result of Atkins and Cheng (1999) that a population design based on $d$-point individual designs where $d$ is an even integer comprises the single design in which 0 and $k$ are each repeated $\frac{d}{2}$ times. Thus

$$\xi^*_{D_{re}} = \left\{ \begin{array}{c} (\underbrace{0,\ldots,0}_{\frac{d}{2}\text{ times}}, \underbrace{k,\ldots,k}_{\frac{d}{2}\text{ times}}) \\ 1 \end{array} \right\}$$

is the $D$-optimal population design for $d$ even.

Consider now the theorem when $d$ is an odd integer. Let an individual design $\mathbf{t}$ be linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$ where $\mathbf{x}_c = (\frac{k}{2},\ldots,\frac{k}{2})$ is the center of the hypercube $C_{d,k}$. Thus the space of designs in the transformed coordinates is given by

$$\tilde{T}_{d,k} = \{\tilde{\mathbf{t}} : \tilde{\mathbf{t}} = (\tilde{t}_1,\ldots,\tilde{t}_d),\ \tilde{t}_j \in \{-\frac{k}{2}, -\frac{k}{2}+1,\ldots,\frac{k}{2}\},\ j = 1,\ldots,d,$$
$$-\frac{k}{2} \leq \tilde{t}_1 \leq \tilde{t}_2 \leq \ldots \leq \tilde{t}_d \leq \frac{k}{2}\}.$$

Then the proposed optimal design $\xi^*_{D_{ro}}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{D_{ro}} = \left\{ \begin{array}{cc} (\underbrace{-\frac{k}{2},\ldots,-\frac{k}{2}}_{\frac{d+1}{2}\text{ times}}, \underbrace{\frac{k}{2},\ldots,\frac{k}{2}}_{\frac{d-1}{2}\text{ times}}), & (\underbrace{-\frac{k}{2},\ldots,-\frac{k}{2}}_{\frac{d-1}{2}\text{ times}}, \underbrace{\frac{k}{2},\ldots,\frac{k}{2}}_{\frac{d+1}{2}\text{ times}}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}.$$

Note immediately that the standardized information matrix for $\beta$ at the design $\tilde{\xi}^*_{D_{ro}}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}^*_{D_{ro}}) = \begin{pmatrix} \frac{1}{1+d\gamma} & 0 \\ 0 & \frac{k^2\,(d-\gamma+d^2\,\gamma)}{4\,d\,(1+d\,\gamma)} \end{pmatrix}$$

and hence that

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_{ro}}) = \begin{pmatrix} 1+d\gamma & 0 \\ 0 & \frac{4\,d\,(1+d\,\gamma)}{k^2\,(d-\gamma+d^2\,\gamma)} \end{pmatrix}.$$

Note also that

$$|\mathbf{M}_\beta(\tilde{\xi}^*_{D_{ro}})| = \frac{k^2\,(d-\gamma+d^2\,\gamma)}{4\,d\,(1+d\,\gamma)^2}.$$

Consider the directional derivative of the criterion $\Psi_D(\tilde{\xi}) = \ln|\mathbf{M}_\beta(\tilde{\xi})|$ at $\tilde{\xi}^*_{D_{ro}}$ in the direction of an individual design $\tilde{\mathbf{t}}$, i.e.

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_{ro}}) = tr\left[\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_{ro}})\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\right] - 2$$

where

$$\mathbf{M}_\beta(\tilde{\mathbf{t}}) = \begin{pmatrix} \dfrac{1}{1+d\,\gamma} & \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d\,(1+d\,\gamma)} \\ \\ \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d\,(1+d\,\gamma)} & \dfrac{1}{d}\tilde{\mathbf{t}}'\,(\mathbf{I}-\dfrac{\gamma}{1+d\,\gamma}\,\mathbf{J})\,\tilde{\mathbf{t}} \end{pmatrix}$$

is the standardized information matrix for a $d$-point design $\tilde{\mathbf{t}} \in \tilde{T}_{d,k}$. By the Equivalence Theorem for $D$-optimal population designs, the design $\tilde{\xi}^*_{D_{ro}}$ is $D$-optimal if and only if $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_{ro}}) \le 0$ for all $d$-point designs $\tilde{\mathbf{t}} \in \tilde{T}_{d,k}$ with equality holding at the support design of $\tilde{\xi}^*_{D_{ro}}$. At the design $\tilde{\xi}^*_{D_{ro}}$ the derivative is given by

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_{ro}}) = \frac{4\,(1+d\,\gamma)}{k^2\,\{(d^2-1)\,\gamma+d\}}\,\tilde{\mathbf{t}}'\,(\mathbf{I}-\frac{\gamma}{1+d\,\gamma}\,\mathbf{J})\,\tilde{\mathbf{t}} - 1.$$

The derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_{ro}})$ is a convex function over the polytope $\tilde{Q}_{d,k}$. Thus it is only necessary to examine $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_{ro}})$ at the vertices of $\tilde{Q}_{d,k}$. Recall from Subsection 4.2.2 that these vertices are the designs $\mathbf{v}_1^*, \mathbf{v}_2^*, \ldots, \mathbf{v}_{d+1}^*$ and are given in the transformed coordinates as

$$\tilde{\mathbf{v}}^*_{j+1} = (\underbrace{-\frac{k}{2},\ldots,-\frac{k}{2}}_{j\ \text{times}},\ \underbrace{\frac{k}{2},\ldots,\frac{k}{2}}_{(d-j)\ \text{times}})\ \text{ for }\ j=0,1,\ldots,d.$$

At the general vertex $\tilde{\mathbf{v}}^*_{j+1}$, the directional derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_{ro}})$ is given by

$$\phi(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{D_{ro}}) = -\frac{(d-2j-1)\,(d-2j+1)\,\gamma}{\{(d^2-1)\,\gamma+d\}}$$

for $j = 0, 1, \ldots, d$. The denominator $\{(d^2-1)\,\gamma+d\}$ is greater than zero for all $\gamma \geq 0$ and $d \geq 1$. The term $(d-2j-1)\,(d-2j+1) = (d-2j)^2-1$ in the numerator is clearly greater than or equal to zero for $j = 0, 1, \ldots, d$ and $d = 1, \ldots, k+1$. Moreover $\phi(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{D_{ro}}) = 0$ at $j = \dfrac{d-1}{2}$ and $j = \dfrac{d+1}{2}$, and thus at the support designs of $\tilde{\xi}^*_{D_{ro}}$. Thus, it follows that $\phi(\tilde{\mathbf{v}}^*_{j+1}, \tilde{\xi}^*_{D_{ro}}) \leq 0$ for $j = 0, 1, \ldots, d$ and for all $\gamma \geq 0$. $\square$

### 4.4.3 The best $D$-optimal population design for the fixed effects

Observe from Theorem 4.4.2 that the design

$$\xi^*_{D_2} = \left\{ \begin{array}{c} (0, k) \\ \\ 1 \end{array} \right\}$$

is the $D$-optimal population design over the set of two-point individual designs with repeated points. It has also been proved in Theorem 4.3.4 that $\xi^*_{D_2}$ is $D$-optimal over the set of population designs defined on $S_{d,k}$, i.e. on the space of designs with non-repeated time points. It is therefore not unreasonable to assume that the design $\xi^*_{D_2}$ is optimal over all population designs based on individual designs with repeated points, i.e. $0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k$ and this is demonstrated in the proof of the following theorem.

**Theorem 4.4.3** *Consider the set of population designs based on all possible individual designs* $\mathbf{t}$ *which put equal weights on the time points* $t_1, t_2, \ldots, t_d$ *with* $t_j \in \{0, 1, \ldots, k\}, j = 1, \ldots, d$ *and* $0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k$ *for* $d$ *any positive integer less than or equal to* $k+1$. *Then the design*

$$\xi^*_{D_2} = \left\{ \begin{array}{c} (0, k) \\ \\ 1 \end{array} \right\}$$

*is the D-optimal population design for the fixed effects $\beta$ in model (4.1) over this set for all $\gamma \geq 0$.*

**Proof**

Recall from Subsection 4.3.3 that $\xi^*_{D_2}$ can be written in linearly transformed coordinates as

$$\tilde{\xi}^*_{D_2} = \left\{ \begin{array}{c} (-\frac{k}{2}, \frac{k}{2}) \\ 1 \end{array} \right\}$$

and that

$$\mathbf{M}_\beta^{-1}(\xi^*_{D_2}) = \begin{pmatrix} 1+2\gamma & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}.$$

Thus the directional derivative of the criterion $\Psi_D(\tilde{\xi}) = \ln |\mathbf{M}_\beta(\tilde{\xi})|$ at the population design $\tilde{\xi}^*_{D_2}$ in the direction of the $d$-point individual design $\tilde{\mathbf{t}} \in \tilde{T}_{d,k}$ is given explicitly by

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2}) = \frac{4\,\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\,\tilde{\mathbf{t}}}{d\,k^2} - \frac{2\,(d-1)\,\gamma + 1}{1+d\,\gamma}$$

and is a convex function on the polytope $\tilde{Q}_{d,k}$. Thus the largest value of $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2})$ occurs at the extreme vertex $\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}$ of the polytope $\tilde{Q}_{d,k}$ when $d$ is even and at the extreme vertices $\tilde{\mathbf{v}}^*_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}^*_{\frac{d+3}{2}}$ of $\tilde{Q}_{d,k}$ when $d$ is odd. It is therefore only necessary to check that the condition $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2}) \leq 0$ holds at these vertices.

Consider first the case of $d$ even. Then at $\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}$ the derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2})$ is given by

$$\phi(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*_{D_2}) = -\frac{(d-2)\,\gamma}{1+d\,\gamma}$$

and this expression is less than or equal to zero for $d \geq 2$ and $\gamma \geq 0$.

Consider now the case of $d$ odd. The derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_2})$ at the vertices $\tilde{\mathbf{v}}^*_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}^*_{\frac{d+3}{2}}$ has the form

$$\phi(\tilde{\mathbf{v}}_{\frac{d+1}{2}}, \tilde{\xi}^*_{D_2}) = \phi(\tilde{\mathbf{v}}_{\frac{d+3}{2}}, \tilde{\xi}^*_{D_2}) = -\frac{(d-1)^2\,\gamma}{d\,(1+d\,\gamma)}$$

and is less than or equal to zero for $d \geq 1$ and $\gamma \geq 0$. Thus the design $\xi^*_{D_2}$ is the $D$-optimal population design for the fixed effects $\beta$ in model (4.1) over the set of all population designs defined on the set $T_{d,k}$ for all $\gamma \geq 0$. $\square$

### 4.4.4 $D$-efficiencies

Recall from Subsection 4.4.2 that when $d$ is even, the information matrix for the fixed effects $\beta$ at the $D$-optimal population design $\tilde{\xi}^*_{D_{re}}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}^*_{D_{re}}) = \begin{pmatrix} \dfrac{1}{1+d\gamma} & 0 \\ 0 & \dfrac{k^2}{4} \end{pmatrix}$$

and hence that

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_{re}}) = \begin{pmatrix} 1+d\gamma & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}.$$

Note also that

$$|\mathbf{M}_\beta(\tilde{\xi}^*_{D_{re}})| = \frac{k^2}{4\,(1+d\,\gamma)}$$

and recall that the determinant of $\mathbf{M}_\beta(\xi)$ for the optimal design $\xi^*_{D_2}$ is

$$|\mathbf{M}_\beta(\xi^*_{D_2})| = \frac{k^2}{4\,(1+2\,\gamma)}.$$

Thus the efficiency of the $d$-point $D$-optimal population design $\xi^*_{D_{re}}$ with respect the best $D$-optimal population design $\xi^*_{D_2}$ is given by

$$D_{eff(1)} = \left\{ \frac{|\mathbf{M}_\beta(\xi^*_{D_{re}})|}{|\mathbf{M}_\beta(\xi^*_{D_2})|} \right\}^{\frac{1}{2}} = \sqrt{\frac{1+2\,\gamma}{1+d\,\gamma}}.$$

This efficiency has the limiting value

$$\lim_{\gamma \to \infty} D_{eff(1)} = \sqrt{\frac{2}{d}}.$$

Similarly, for $d$ odd, the determinant of $\mathbf{M}_\beta(\xi)$ for the optimal design $\xi^*_{D_{ro}}$ is

$$|\mathbf{M}_\beta(\xi^*\xi^*_{D_{ro}})| = \frac{k^2\,(d-\gamma+d^2\,\gamma)}{4\,d\,(1+d\,\gamma)^2}$$

and thus the efficiency of the $d$-point $D$-optimal population design $\xi^*_{D_{ro}}$ with respect to the best $D$-optimal population design $\xi^*_{D_2}$ is given by

$$D_{eff(2)} = \left\{\frac{|\mathbf{M}_\beta(\xi^*_{D_{ro}})|}{|\mathbf{M}_\beta(\xi^*_{D_2})|}\right\}^{\frac{1}{2}} = \sqrt{\frac{(d-\gamma+d^2\,\gamma)\,(1+2\,\gamma)}{d\,(1+d\,\gamma)^2}}.$$

This has the limiting value

$$\lim_{\gamma\to\infty} D_{eff(2)} = \sqrt{\frac{2\,(d^2-1)}{d^3}}.$$

When $\gamma = 0$ however both efficiencies $D_{eff(1)}$ and $D_{eff(2)}$ are equal to 1. Thus for $\gamma = 0$ the $d$-point $D$-optimal population designs are fully efficient.

Plots of the $D$-efficiencies $D_{eff(1)}$ and $D_{eff(2)}$ of the $d$-point ($d \geq 3$) $D$-optimal population designs $\xi^*_{D_{re}}$ and $\xi^*_{D_{ro}}$ with respect to the optimal design $\xi^*_{D_2}$ against the variance ratio $\gamma$ for $k = 10$ are presented in Figures 4.10 and 4.11 respectively. The plots exhibit a similar trend to those for the non-repeated points case in that for a specified value of $d$ the efficiencies decreases with increasing $\gamma$. It is also clear that for a given $\gamma$ the efficiencies decrease as $d$ increases.

### 4.4.5 Comparison of $D$-optimal population designs

$D$-optimal population designs based on $d$-point individual designs with repeated time points for $d \geq 3$ are now compared in terms of their $D$-efficiencies. The general result is presented in the following theorem.

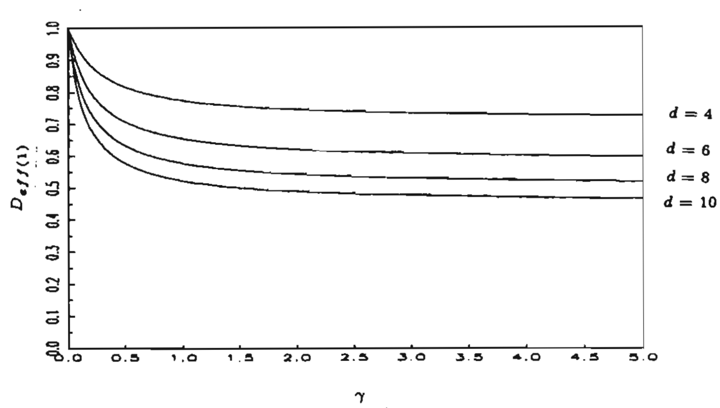


Figure 4.10: $D$-efficiency of the $d$-point $D$-optimal population design $\xi^*_{D_{re}}$ with respect to $\xi^*_{D_2}$ plotted against $\gamma$ for $k = 10$ and $d = (4, 6, 8, 10)$.

**Theorem 4.4.4** *Let the constants $d_e$ and $d_o$ be even and odd positive integers both greater than or equal to 3. Then the D-optimal population designs $\xi^*_{D_{re}}$ and $\xi^*_{D_{ro}}$ are the most efficient designs on a per observation basis over the set of d-point individual designs defined on the space of designs $T_{d,k}$ with $d > d_e$ and $d > d_o$, respectively. Furthermore, the D-efficiency decreases on a per observation basis as d increases.*

### Proof

Consider $d_e = 2m$ even and $d_o = 2m + 1$ odd for $m$ is a positive integer greater than or equal to 1. To prove the theorem it is necessary to show that the inequalities

$$\left|\mathrm{M}_\beta(\xi^*_{D_{re(2m)}})\right| > \left|\mathrm{M}_\beta(\xi^*_{D_{ro(2m+1)}})\right| > \left|\mathrm{M}_\beta(\xi^*_{D_{re(2m+2)}})\right|$$

or

$$\left|\mathrm{M}_\beta(\xi^*_{D_{re(2m)}})\right| - \left|\mathrm{M}_\beta(\xi^*_{D_{ro(2m+1)}})\right| > 0 \text{ and } \left|\mathrm{M}_\beta(\xi^*_{D_{ro(2m+1)}})\right| - \left|\mathrm{M}_\beta(\xi^*_{D_{re(2m+2)}})\right| > 0$$

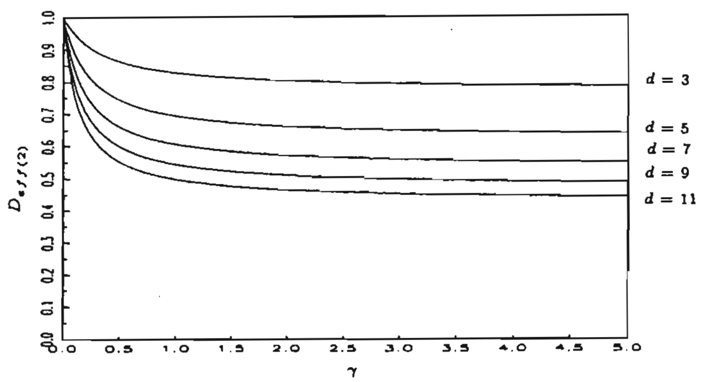


Figure 4.11: $D$-efficiencies of the $d$-point $D$-optimal population design $\xi^*_{D_{ro}}$ with respect to $\xi^*_{D_2}$ against $\gamma$ for $k = 10$ and $d = (3, 5, 7, 9, 11)$.

hold for all $m$ where $\xi^*_{D_{re(2m)}}$, $\xi^*_{D_{re(2m+1)}}$ and $\xi^*_{D_{re(2m+2)}}$ are the $D$-optimal population designs based on individual designs with $2m$, $2m+1$ and $2m+2$ points respectively.

Substituting $d$ by $2m$ and $2m+1$ in the determinants of $\mathbf{M}_\beta(\xi^*_{D_{re}})$ and $\mathbf{M}_\beta(\xi^*_{D_{ro}})$ respectively yield

$$|\mathbf{M}_\beta(\xi^*_{D_{re(2m)}})| = \frac{k^2}{4\,(1 + 2\,m\,\gamma)}$$

and

$$|\mathbf{M}_\beta(\xi^*_{D_{ro(2m+1)}})| = \frac{k^2\,\{1 + 2\,m + 4\,m\,\gamma\,(m+1)\}}{4\,(1 + 2\,m)\,\{1 + \gamma\,(1 + 2\,m)\}^2}.$$

The difference

$$|\mathbf{M}_\beta(\xi^*_{D_{re(2m)}})| - |\mathbf{M}_\beta(\xi^*_{D_{ro(2m+1)}})| = \frac{k^2\,\gamma\,\{2 + 2\,m + \gamma(1 + 6\,m + 4\,m^2)\}}{4\,(1 + 2\,m)\,(1 + 2\,m\,\gamma)\,\{1 + \gamma(1 + 2\,m)\}^2} > 0$$

for $m \geq 1$ and $\gamma > 0$, implying that $|\mathbf{M}_\beta(\xi^*_{D_{re(2m)}})| > |\mathbf{M}_\beta(\xi^*_{D_{ro(2m+1)}})|$ for $m \geq 1$ and $\gamma > 0$.

Now let $d = 2m + 2 \leq k + 1$. Then the determinant of $\mathbf{M}_\beta(\xi^*_{D_{re(2m+2)}})$ is given by

$$|\mathbf{M}_\beta(\xi^*_{D_{re(2m+2)}})| = \frac{k^2}{4\{1 + 2\gamma(m+1)\}}.$$

Then

$$|\mathbf{M}_\beta(\xi^*_{D_{re(2m+2)}})| - |\mathbf{M}_\beta(\xi^*_{D_{ro(2m+1)}})| = \frac{k^2\gamma\{2m + \gamma(2m-1) + 4m^2\gamma\}}{4(1+2m)\{1+\gamma(1+2m)\}^2\{1+2\gamma(1+m)\}}$$

which is greater than zero since both the numerator and denominator are positive for $m \geq 1$ and $\gamma > 0$. □

## 4.5 Efficiencies of population designs based on individual designs with repeated points relative to those with non-repeated points

In this section $D$-optimal population designs based on $d$-point individual designs with repeated and with non-repeated points are compared. Note immediately that for $d = 1$ and $d = 2$ such designs are identical. For designs with $d \geq 3$ the comparison is based on $D$-efficiencies and thus on the determinants of the appropriate information matrices and is presented formally in the following theorem.

**Theorem 4.5.1** *For $d \geq 3$, $D$-optimal population designs based on $d$-point individual designs with repeated time points are more efficient than the corresponding $D$-optimal population designs with non-repeated points for all $\gamma \geq 0$.*

**Proof**

Consider first the case of $d$ even. Recall from Subsections 4.3.2 and 4.4.2 that the determinants of the information matrices for $\boldsymbol{\beta}$ at $D$-optimal population designs based on $d$-point individual designs with non-repeated and with repeated points are, respectively,

$$|\mathbf{M}_\beta(\xi^*_{D_e})| = \frac{1}{12(1+d\gamma)}\{d^2 - 3d(k+1) + 3k^2 + 6k + 2\}$$

and

$$|\mathbf{M}_\beta(\xi^*_{D_{re}})| = \frac{k^2}{4\,(1+d\,\gamma)}.$$

Therefore the difference between these determinants is given by

$$|\mathbf{M}_\beta(\xi^*_{D_{re}})| - |\mathbf{M}_\beta(\xi^*_{D_e})| = \frac{(d-2)\,(3\,k+1-d)}{12\,(1+d\,\gamma)}$$

and since $d-2>0$ for $d \geq 3$ and $3\,k+1-d>0$ for $d \leq k+1$ this difference is greater than zero for all $\gamma \geq 0$. Thus $|\mathbf{M}_\beta(\xi^*_{D_{re}})| > |\mathbf{M}_\beta(\xi^*_{D_e})|$.

Suppose now that $d$ is an odd integer. Then the determinants of the information matrices $\mathbf{M}_\beta(\xi^*_{D_o})$ and $\mathbf{M}_\beta(\xi^*_{D_{ro}})$ are, respectively,

$$|\mathbf{M}_\beta(\xi^*_{D_o})| = \frac{H}{12\,d\,(1+d\,\gamma)^2}$$

where

$$H = d^3 - 3\,(k+1)\,(d^2+1) + (3\,k^2+6\,k+5)\,d + (d^2-1)\,\{d^2+3\,(k+1)\,(k-d+1)\}\,\gamma$$

and

$$|\mathbf{M}_\beta(\xi^*_{D_{ro}})| = \frac{k^2\,\{d+\gamma\,(d^2-1)\}}{4\,d\,(1+d\,\gamma)^2}.$$

Therefore the difference between these determinants is equal to

$$|\mathbf{M}_\beta(\xi^*_{D_{ro}})| - |\mathbf{M}_\beta(\xi^*_{D_o})| = \frac{(d-1)\{(d+1)\,C_1\,\gamma + C_0\}}{12\,d\,(1+d\,\gamma)}$$

where

$$C_0 = 3\,d\,k - 3\,k - d^2 + 2\,d - 3 \text{ and } C_1 = 3\,d\,(k+1) - 6\,k - d^2 - 3.$$

Observe that $C_0$ and $C_1$ are quadratic in $d$. For $d \geq 3$ and $d \leq k+1$

$$\frac{\partial C_0}{\partial d} = 3\,k - 2\,d + 2 > 0.$$

Further at $d = 3$ and $d = k+1$, $C_0$ is given by the positive expressions $6\,(k-1)$ and $2\,(k^2-1)$ respectively. Thus $C_0$ is necessarily positive on the interval $[3, k+1]$. Similarly for $d \geq 3$ and $d \leq k+1$

$$\frac{\partial C_1}{\partial d} = 3\,k - 2\,d + 3 > 0$$

and at $d = 3$ and $d = k+1$ $C_1$ is given by $3\,(k-1)$ and $2\,k(k-1) - 1$ respectively, both of which are positive for $k > 1$. Thus $C_1 > 0$ on the interval $[3, k+1]$. Overall therefore $|\mathbf{M}_\beta(\xi^*_{D_{ro}})| - |\mathbf{M}_\beta(\xi^*_{D_o})| > 0$ for all $\gamma \geq 0$. □

Thus, it follows immediately from the above result that $D$-optimal population designs based on $d$-point individual designs with repeated time points are more efficient than the corresponding $D$-optimal population designs with non-repeated points for all $\gamma \geq 0$. Note that the result in Theorem 4.5.1 also follows from the fact that the set of designs with non-repeated points $S_{d,k}$ is a subset of the set of designs with repeated points $T_{d,k}$. The efficiency that will be lost by using $D$-optimal population designs with non-repeated points can easily be calculated using (2.38). For example, for $k = 10$ and $d = 4$

$$D_{eff)} = \left\{ \frac{|\mathbf{M}_\beta(\xi^*_{D_e})|}{|\mathbf{M}_\beta(\xi^*_{D_{re}})|} \right\}^{\frac{1}{2}} = \sqrt{0.82} = 0.9055.$$

Thus, 9.45% of efficiency will be lost if design (0, 1, 9, 10) used to estimate $\boldsymbol{\beta}$.

## 4.6 Optimal design for the slope parameter, $\beta_1$

If the main interest of inference is to estimate the slope parameter $\beta_1$ in model (4.1) as precisely as possible, it is natural to find designs that minimize the length of the confidence interval of the estimator $\hat{\beta}_1$, or equivalently, that minimizes the variance of the estimator $\hat{\beta}_1$, $Var(\hat{\beta}_1)$ .

Let the information matrix for the fixed effects $\boldsymbol{\beta}$ of the simple linear regression model with a random intercept specified by (4.1) at the population design

$$\xi = \begin{Bmatrix} \mathbf{t}_1, & \ldots, & \mathbf{t}_r \\ w_1, & \ldots, & w_r \end{Bmatrix} \quad \text{with } 0 < w_i < 1 \text{ and } \sum_{i=1}^{r} w_i = 1$$

be expressed as

$$\mathbf{M}_\beta(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\beta(\mathbf{t}_i) = \frac{1}{1 + d\gamma} \begin{pmatrix} 1 & m_{12}(\xi) \\ m_{12}(\xi) & m_{22}(\xi) \end{pmatrix}$$

where for a $d$-point design $\mathbf{t}_i$

$$\mathbf{M}_\beta(\mathbf{t}_i) = \frac{1}{d\,(1 + d\gamma)} \begin{pmatrix} d & \sum_{j=1}^{d} t_j \\ \sum_{j=1}^{d} t_j & \sum_{j=1}^{d} t_j^2 + \gamma\, d\, SS(t) \end{pmatrix}, \qquad i = 1, \ldots, r$$

assuming that $\sigma_e^2 = 1$. Then the variance of the maximum likelihood estimate $\hat{\beta}_1$ of the slope parameter $\beta_1$ is given by

$$Var(\hat{\beta}_1) = \frac{1 + d\gamma}{m_{22}(\xi) - m_{12}^2(\xi)}.$$

However, the determinant of the information matrix for $\boldsymbol{\beta}$ at the population design $\xi$ is

$$|\mathbf{M}_\beta(\xi)| = \frac{m_{22}(\xi) - m_{12}^2(\xi)}{(1 + d\gamma)^2}.$$

Thus for model (4.1) minimizing $Var(\hat{\beta}_1)$ is equivalent maximizing the determinant $|\mathbf{M}_\beta(\xi)|$. In other word, for the simple linear regression model with a random intercept, the $D$-optimal population designs for $\boldsymbol{\beta}$ derived in previous sections are also optimal for the precise estimation of the slope parameter $\beta_1$. This is a special case of a result given by Goos (2002, page 15).

## 4.7 $D$-optimal designs for variance components

In the linear random intercept model (4.1) there are only two sources of variation, random intercept and the random error. Therefore the variance components parameter $\boldsymbol{\theta}$ is a vector $(\sigma_e^2, \sigma_b^2)$ where $\sigma_b^2$ and $\sigma_e^2$ are the variances of random intercept and random error, respectively.

Suppose that there are $K$ individuals, that the $i$th individual has a design based on $d_i$ points, $i = 1, \ldots, K$, and that the total number of observations $N$ is fixed. Thus the design is specified by the allocation $d_1, \ldots, d_K$ with $N = \sum_{i=1}^{K} d_i$ fixed and from Subsection 2.5.1 the overall information matrix for $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$ for this allocation is then given by

$$\mathbf{M}_\theta(\mathbf{d}) = \frac{1}{2}\begin{pmatrix} \frac{1}{\sigma_e^4}\sum_{i=1}^{K}\left[(d_i - 1) + \left(\frac{\sigma_e^2}{\sigma_e^2 + d_i\sigma_b^2}\right)^2\right] & \sum_{i=1}^{K}\frac{d_i}{(\sigma_e^2 + d_i\sigma_b^2)^2} \\ \sum_{i=1}^{K}\frac{d_i}{(\sigma_e^2 + d_i\sigma_b^2)^2} & \sum_{i=1}^{K}\frac{d_i^2}{(\sigma_e^2 + d_i\sigma_b^2)^2} \end{pmatrix}$$

where $\mathbf{d} = (d_1, d_2, \ldots, d_K)'$. Thus

$$|\mathbf{M}_\theta(\mathbf{d})| = \frac{1}{4}\left\{\left(\frac{1}{\sigma_e^4}\sum_{i=1}^{K}\left[(d_i - 1) + \left(\frac{\sigma_e^2}{\sigma_e^2 + d_i\sigma_b^2}\right)^2\right]\right)\left(\sum_{i=1}^{K}\frac{d_i^2}{(\sigma_e^2 + d_i\sigma_b^2)^2}\right)\right.$$

$$- \left( \sum_{i=1}^{K} \frac{d_i}{(\sigma_e^2 + d_i \sigma_b^2)^2} \right)^2 \Bigg\} . \tag{4.21}$$

The information matrix $\mathbf{M}_\theta(\mathbf{d})$ and hence the determinant $|\mathbf{M}_\theta(\mathbf{d})|$ depend only on the number of points $d_i$ in the design and not on the actual values of the points. The matrix $\mathbf{M}_\theta(\mathbf{d})$ and $|\mathbf{M}_\theta(\mathbf{d})|$ also depend on the unknown parameter $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$. Therefore to calculate the optimum design for $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$ it is assumed here following Chernoff (1953) that a best guess is taken for $\sigma_e^2$ and $\sigma_b^2$. The above design allocation is equivalent to the design problem of a one-way model as described by Giovagnoli and Sebastiani (1989). It was previously reached by Anderson (1975, 1981), Mukerjee and Huda (1988) and Giovagnoli and Sebastiani (1989) that the balanced allocation, in which the same number of measurements are allocated to each individual, is optimal for this model. Therefore the interest of this section is to examine such an allocation.

Under the balanced allocation, i.e. $d_1 = d_2 = \ldots = d_K$, the determinant of the information matrix $\mathbf{M}_\theta(\mathbf{d})$ in (4.21) is

$$|\mathbf{M}_\theta(d)| = \frac{N^2\,(d-1)}{4\,(\sigma_e^2)^2\,(\sigma_e^2 + d\,\sigma_b^2)^2}$$

where $N$ is fixed. Then solving the equation

$$\frac{\partial |\mathbf{M}_\theta(d)|}{\partial d} = \frac{N^2\,\{\sigma_e^2 - (d-2)\,\sigma_b^2\}}{4\,(\sigma_e^2)^2\,(\sigma_e^2 + d\,\sigma_b^2)^3} = 0$$

yields

$$d^* = 2 + \frac{\sigma_e^2}{\sigma_b^2}$$

as a solution for $d$. Since

$$\left.\frac{\partial^2 |\mathbf{M}_\theta(d)|}{\partial d^2}\right|_{d=d^*} = -\frac{N^2\,\sigma_b^2}{32\,(\sigma_e^2)2\,(\sigma_e^2 + \sigma_b^2)^3} < 0$$

the determinant $|\mathbf{M}_\theta(d)|$ attains its maximum at $d^*$. For a given number of observations, $N$, when there is large within-individual variation relative to between-individual variation, the result suggests taking a large number of observations on a few individuals, i.e. the optimal $d$ is large and the associated $K$ is small.

Suppose now that the variance components parameter $\boldsymbol{\theta}$ comprises the random error variance $\sigma_e^2$ and the variance ratio $\gamma = \frac{\sigma_b^2}{\sigma_e^2}$, i.e. $\boldsymbol{\theta} = (\sigma_e^2, \gamma)$. Then from Subsection 2.5.1 the overall information matrix of $K$ individuals each with $d_i$ points, $i = 1, \ldots, K$, is given by

$$\mathbf{M}_\theta(\mathbf{d}) = \frac{1}{2}\begin{pmatrix} \dfrac{N}{\sigma_e^4} & \dfrac{1}{\sigma_e^2}\displaystyle\sum_{i=1}^{K}\frac{d_i}{1+d_i\gamma} \\ \dfrac{1}{\sigma_e^2}\displaystyle\sum_{i=1}^{K}\frac{d_i}{1+d_i\gamma} & \displaystyle\sum_{i=1}^{K}\frac{d_i^2}{(1+d_i\gamma)^2} \end{pmatrix}$$

and

$$|\mathbf{M}_\theta(\mathbf{d})| = \frac{1}{4\sigma_e^4}\left\{N\sum_{i=1}^{K}\frac{d_i^2}{(1+d_i\gamma)^2} - \left(\sum_{i=1}^{K}\frac{d_i}{1+d_i\gamma}\right)^2\right\} \tag{4.22}$$

where the total number of observations $N = \sum_{i=1}^{K} d_i$ is fixed. Since $|\mathbf{M}_\theta(\mathbf{d})|$ depends on the unknown parameter $\boldsymbol{\theta} = (\sigma_e^2, \gamma)$ it is again assumed that a best guess of this parameter is available to calculate the optimum $d$. For a balanced allocation the determinant in (4.22) simplifies to

$$|\mathbf{M}_\theta(d)| = \frac{N^2\,(d-1)}{4\,(\sigma_e^2)^2\,(\sigma_e^2 + d\,\gamma)^2}.$$

Then solving the equation

$$\frac{\partial|\mathbf{M}_{\theta_2}|}{\partial d} = \frac{N^2\,\{1-(d-2)\,\gamma\}}{2\,(\sigma_e^2)^2\,(1+d\,\gamma)^4} = 0$$

yields

$$d^* = 2 + \frac{1}{\gamma}$$

as a solution for $d$ and it is equal to the optimal value of $d$ in the $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$ case. This result also implies that for a fixed $N$, when $\gamma$ is small, a large number of observations per individual is preferable. As $\gamma \to \infty$, $d^* = 2$. Thus the optimum number of observations $d^*$ must be restricted to lie in the interval $[2, N]$ (Giovagnoli and Sebastiani, 1989).

## 4.8 Trypanosmosis example

The results of this chapter are applied to the data from the experiment in susceptibility to trypanosmosis reported by Duchateau, Janssen and Rowlands (1998, page 13) and introduced in Chapter 3. Here only the data corresponding to the N'Dama breed are used. In the experiment there are six animals each with observations taken at 14 time points 0, 2, 4, 7, 9, 14, 17, 18, 21, 23, 25, 29, 31 and 35 giving 84 observations in all. Figure 4.13 contains a plot of changes in PCV against time for the six animals.

The plots show that the observations for each animal decrease approximately linearly with time and further that the observations for each animal appear to be correlated. This indicates that the simple linear regression model with a random intercept is a suitable model for the data. Specifically the PCV for the $j$th measurement on the $i$th animal at time $t_j$ can be modelled as

$$y_{ij} = \beta_o + \beta_1 t_j + b_i + e_{ij}, \quad j = 1, 2, \ldots, 14, \quad i = 1, \ldots, 6$$

where $y_{ij}$ is the PCV for animal $i$ at time $t_j$, $\beta_o$ and $\beta_1$ are the fixed effects, $b_i$ is a random effect relating to the $i$th animal and $e_{ij}$ is the random error associated with measurement $j$ on animal $i$. It is assumed that $b_i \sim N(0, \sigma_b^2)$, that $e_{ij} \sim N(0, \sigma_e^2)$ and that the $b_i$ and the $e_{ij}$ are independent. Note that the covariance structure of the fourteen observations

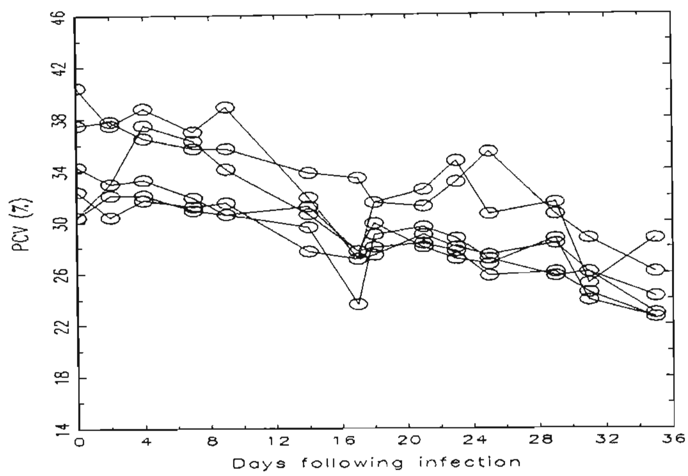


Figure 4.12: Changes in PCV following infection of N'Dama cattle.

$\mathbf{y}_i$ on the same animal is given by $\mathbf{V}_i = \sigma_e^2 \mathbf{I}_{14} + \sigma_b^2 \mathbf{J}_{14}$. Maximum likelihood estimates of the fixed effect parameters and the variance components were obtained from the data using PROC MIXED in SAS (Littell, Milliken, Stroup and Wolfinger, 1996) and are given by $\hat{\beta}_o = 35.077$, $\hat{\beta}_1 = -0.276$, $\hat{\sigma}_b^2 = 4.181$ and $\hat{\sigma}_e^2 = 3.595$.

Assume that the values of the variance components parameters are the maximum likelihood estimates obtained from the data and also that only 84 observations on the cattle are affordable in the experiment. The objective is then to estimate the model parameters as precisely as possible. Suppose that all 36 days labelled 0, 1, ..., 35 are available for taking measurements and that the researchers are interested in taking measurements for a maximum of six animals. Then for this pool of cattle the researcher can take 42 observations on 2 animals based on design (0, 35); or 28 observations on 3 animals, i.e. 14 observations based on design (0, 1, 35) and the other 14 based on design (0, 34, 35); or 21 observations on 4 animals based on design (0, 1, 34, 35), and so on following the results of this chapter. These results, i.e. $\xi_D^*$, the determinant of $\mathbf{M}_\beta(\xi_D^*)$ and the $D$-efficiency relative to the best population design $\xi_{D_2}^*$ are presented in Table 4.1. It is easy to see from this table that $D$-efficiencies relative to $\xi_{D_2}^*$ decrease as the number of time points $d$ increases and that the $D$-optimum population design $\xi_{D_d}^*$ with small $d$ is more efficient than the one with large $d$. Note that in Table 4.1 the $D$-optimal population design, for $d = 5$ is not included since 42 measurements cannot be allocated equally among five animals.

Consider now the original experiment, that is suppose that there are only the 14 days listed in the experiment available for taking measurements. Since the days are not equally spaced the results of this Chapter to calculate the $D$-optimum population designs to estimate the fixed effects $\boldsymbol{\beta}$ as precisely as possible do not apply. Therefore a GAUSS program has been written to compute a $D$-optimal population design based on the set of $d$-point individual designs for $1 \leq d \leq 14$. The program is given in the file labelled "doptinte" on the CD provided with this thesis. The program calculates a $d$-point $D$-optimal population design for a given value of $\gamma$ and it can be used also to compute the best $D$-optimal population design. The $d$-point $D$-optimal population designs for the experiment obtained from this program with $\gamma = 1.163$ are presented in Table 4.2.

Observe that, in contrast to the results in Table 4.1, the design weight for the $D$-optimal designs when $d$ is an odd integer is not 0.5. For example, the design weight in $\xi_{D_3}^*$ is 0.89 for design (0,2,35) and 0.19 for design (0,31,35). However, the $D$-efficiencies relative to $\xi_{D_2}^*$ exhibit a similar trend to those for Table 4.1 in that the efficiencies decrease as the number

Table 4.1: $d$-point $D$-optimal population designs for $\beta$ in the simple linear regression model with random intercept for the trypanosmosis data based on the set of points $\{0, 1, \ldots, 35\}$

| $d$ | $\xi^*_{D_d}$ | $w$ | $\lvert\mathbf{M}_\beta(\xi^*_{D_d})\rvert$ | $D$-efficiency |
|---|---|---|---|---|
| 2 | (0,35) | 1 | 92.0776 | 1.0000 |
| 3 | (0,1,35) | 0.5 | | |
| | (0,34,35) | 0.5 | 60.4601 | 0.8103 |
| 4 | (0,1,34,35) | 1 | 51.1766 | 0.7455 |
| 6 | (0,1,2,33,34,35) | 1 | 34.2087 | 0.6095 |
| 7 | (0,1,2,3,33,34,35) | 0.5 | | |
| | (0,1,2,32,33,34,35) | 0.5 | 28.4589 | 0.5559 |
| 14 | (0,1,2,3,4,5,6,29,30,31,32,33,34,35) | 1 | 12.3973 | 0.3669 |

Table 4.2: $d$-point $D$-optimal population designs for $\beta$ in the simple linear regression model with random intercept for the trypanosmosis data in the actual study

| $d$ | $\xi^*_{D_d}$ | $w$ | $\lvert\mathbf{M}_\beta(\xi^*_{D_d})\rvert$ | $D$-efficiency |
|---|---|---|---|---|
| 2 | (0,35) | 1 | 92.0776 | 1.0000 |
| 3 | (0,2,35) | 0.81 | | |
| | (0,31,35) | 0.19 | 57.5429 | 0.7905 |
| 4 | (0,2,34,35) | 1 | 45.7360 | 0.7048 |
| 6 | (0,2,4,29,31,35) | 1 | 28.1364 | 0.5528 |
| 7 | (0,2,4,7,29,31,35) | 1 | 22.3438 | 0.4926 |
| 14 | (0,2,4,7,9,14,17,18,21,23,25,29,31,35) | 1 | 6.7633 | 0.2710 |

of time points $d$ increases. The $D$-optimum population design $\xi^*_{D_3}$, for instance, indicates that since there are a total of 84 measurements the researcher has to take $84 \times 0.81 = 68$ based on design (0,2,35) and $84 \times 0.19 = 16$ based on design (0, 31,35).

Recall from Section 4.7 that the optimum value of the number of time points $d$ per individual under a balanced allocation for estimating the variance components $\boldsymbol{\theta} = (\sigma_e^2, \sigma_b^2)$ as precisely as possible is given by $d^* = 2 + \frac{1}{\gamma}$. Assuming that the variance components parameters values are the maximum likelihood estimates obtained from the data the optimum number of time points required per individual is $d^* = 2 + \frac{1}{1.163} = 2.8598 \approx 3$. Thus 3-point designs are recommended for precise estimation of $\sigma_e^2$ and $\sigma_b^2$.

# Chapter 5

# $V$-optimal Population Designs for the Simple Linear Regression Model with a Random Intercept

## 5.1 Introduction

One of the design problems addressed in this thesis is the estimation of mean responses specified by a linear mixed model as precisely as possible. The objective of this Chapter is therefore to describe the construction of $V$-optimal designs for the simple linear regression model with a random intercept.

Recall from Subsection 3.3.1 that the matrix form of the simple linear regression model with a random intercept is given by

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{1}_{d_i} b_i + \mathbf{e}_i \tag{5.1}$$

where $\mathbf{y}_i$ is a $d_i \times 1$ vector of observations for the $i$th individual at time points $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{id_i})'$, $\mathbf{X}_i = [1 \;\; \mathbf{t}_i]$, $\boldsymbol{\beta} = (\beta_o, \beta_1)'$, $b_i$ is a random intercept for the $i$th individual and $\mathbf{e}_i$ is a random error vector, $i = 1, \ldots, K$. Further it is assumed that $b_i \sim \mathcal{N}(0, \sigma_b^2)$, that $\mathbf{e}_i \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I}_{d_i})$, and that $b_i$ and the elements of $\mathbf{e}_i$ are independent within and between individuals. Under these assumptions the population mean response for the individual $i$ is equal to $E(\mathbf{y}_i) = \mathbf{X}_i \boldsymbol{\beta}$ where $i = 1, \ldots, K$.

Suppose that interest centres on the estimation of the mean response at a given vector of time points $\mathbf{t}_g$, where the elements of $\mathbf{t}_g$ are taken from the set $\{0, 1, \ldots, k\}$. Suppose also that the elements of $\mathbf{t}_g$ are assembled in the design matrix $\mathbf{X}_g$ in accord with the linear mixed model of interest. Then the population mean response at $\mathbf{t}_g$ is given by

$$\boldsymbol{\mu}_g = \mathbf{X}_g \boldsymbol{\beta}$$

where $\mathbf{X}_g = (1 \;\; \mathbf{t}_g)$. The maximum likelihood estimator of the mean response $\boldsymbol{\mu}_g$ is $\mathbf{X}_g \hat{\boldsymbol{\beta}}$ and its asymptotic variance based on the population design

$$\xi = \left\{ \begin{array}{ccc} \mathbf{t}_1, & \ldots, & \mathbf{t}_r \\ w_1, & \ldots, & w_r \end{array} \right\} \quad \text{with } 0 < w_i < 1 \text{ and } \sum_{i=1}^{r} w_i = 1 \tag{5.2}$$

is equal to $\mathbf{X}_g \mathbf{M}_\beta^{-1}(\xi) \mathbf{X}_g'$. Note that the information matrix for $\boldsymbol{\beta}$ at $\xi$ is given by

$$\mathbf{M}_\beta(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\beta(\mathbf{t}_i)$$

where $\mathbf{M}_\beta(\mathbf{t}_i)$ is the standardized information matrix for $\boldsymbol{\beta}$ at the individual design $\mathbf{t}_i$ and is specified in expression (4.9). Therefore the $V$-optimality criterion, which is the average of the variances of the estimators of the mean responses, can be formulated as

$$\Psi_V(\xi) = tr\{\mathbf{M}_\beta^{-1}(\xi) \mathbf{X}_g' \mathbf{X}_g\}$$

and the population design $\xi_V^*$ is $V$-optimal if it minimizes this criterion. Since the within individual variance $\sigma_e^2$ factors out of the expression for $\mathbf{M}_\beta(\xi)$ it will also factor out of the criterion $\Psi_V(\xi)$ and $\sigma_e^2$ can thus be taken to be 1 without loss of generality. Furthermore, it follows immediately from the Equivalence Theorem given in Theorem 2.6.3 that the design $\xi_V^*$ is $V$-optimal if and only if

$$\phi_V(\mathbf{t},\xi_V^*) = tr\{\mathbf{M}_\beta^{-1}(\xi_V^*)\mathbf{X}_g'\mathbf{X}_g\,\mathbf{M}_\beta^{-1}(\xi_V^*)\mathbf{M}_\beta(\mathbf{t})\} - tr\{\mathbf{M}_\beta^{-1}(\xi_V^*)\mathbf{X}_g'\mathbf{X}_g\} \leq 0 \tag{5.3}$$

for all individual designs $\mathbf{t}$ in the space of designs of interest, with equality holding at the support designs of $\xi_V^*$. In the present study $\mathbf{t}_g$ is assumed to be the vector $(0, 1, \ldots, k)'$, unless stated otherwise, and thus

$$\mathbf{X}_g = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & k \end{pmatrix}.$$

The organization of this chapter as follows. The construction of $V$-optimal population designs based on $d$-point individual designs with non-repeated time points is discussed in Section 5.2. These designs are compared in Section 5.3 and the construction of the best $V$-optimal population design over all such designs is discussed in Section 5.4. $V$-optimal population designs based on $d$-point individual designs with repeated points are considered in Section 5.5 and comparisons of these designs with the corresponding designs based on non-repeated time points are given in Section 5.6. Finally, in Section 5.7 the results of the chapter are illustrated by using the trypanosmosis data.

## 5.2 $V$-optimal population designs based on designs with non-repeated time points

In this section $V$-optimal population designs based on the set $S_{d,k}$ are discussed. First the case of $d = 1$ is considered and then that for general $d$ with $d$ even and $d$ odd follows.

### 5.2.1 Designs based on one-point individual designs

**Theorem 5.2.1** *Consider model (5.1) and the set of all possible one-point designs* $t \in \{0, 1, \ldots, k\}$. *Then*

$$\xi^*_{V_1} = \left\{ \begin{array}{cc} (0) & (k) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

*is the $V$-optimal population design for the mean responses* $\boldsymbol{\mu}_g$ *at* $\mathbf{t}_g = (0, 1, \ldots, k)'$ *over this set for all* $\gamma \geq 0$.

**Proof**

Recall from Subsection 2.6.5 that $V$-optimal designs for random intercept models are invariant to linear transformation. Thus without loss of generality, let an individual design $t$ be linearly transformed according to $\tilde{t} = t - \frac{k}{2}$. Then the proposed optimum design $\xi^*_{V_1}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{V_1} = \left\{ \begin{array}{cc} (-\frac{k}{2}) & (\frac{k}{2}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}.$$

Note immediately that the standardized information matrix for $\boldsymbol{\beta}$ at the design $\tilde{\xi}^*_{V_1}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}^*_{V_1}) = \frac{1}{1+\gamma}\begin{pmatrix} 1 & 0 \\ 0 & \dfrac{k^2}{4} \end{pmatrix}$$

and hence that

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_1}) = (1+\gamma)\begin{pmatrix} 1 & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}.$$

Note also that the matrix $\mathbf{X}_g'\mathbf{X}_g$ in the transformed coordinates is given by

$$\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g = \begin{pmatrix} \mathbf{1}' \\ \tilde{\mathbf{t}}_g' \end{pmatrix}\begin{pmatrix} \mathbf{1} & \tilde{\mathbf{t}}_g \end{pmatrix} = \begin{pmatrix} k+1 & 0 \\ 0 & \dfrac{k(k+2)(k+1)}{12} \end{pmatrix}. \tag{5.4}$$

It then follows that the $V$-optimality criterion evaluated at the design $\tilde{\xi}^*_{V_1}$ can be expressed as

$$\Psi_V(\tilde{\xi}^*_{V_1}) = tr\left\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_1})\,\tilde{\mathbf{X}}_g'\,\tilde{\mathbf{X}}_g\right\} = \frac{2}{3\,k}(1+\gamma)\,(2\,k+1)\,(k+1).$$

Now in order to show that $\tilde{\xi}^*_{V_1}$ and hence $\xi^*_{V_1}$ is the $V$-optimal design it is necessary to invoke the appropriate Equivalence Theorem. In particular, it necessary to show that $\phi_V(\tilde{t}, \tilde{\xi}^*_{V_1}) \leq 0$ for all single-point designs $\tilde{t} \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}\}$. From expression (4.9) the standardized information matrix for $\boldsymbol{\beta}$ at a one-point design $\tilde{t}$ is given by

$$\mathbf{M}_\beta(\tilde{t}) = \frac{1}{1+\gamma}\begin{pmatrix} 1 & \tilde{t} \\ \tilde{t} & \tilde{t}^2 \end{pmatrix}.$$

Substituting the appropriate expressions for $\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_1})$, $\mathbf{M}_\beta(\tilde{t})$ and $\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g$ into that for the directional derivative, namely (5.3), yields

$$\phi_V(\tilde{t}, \tilde{\xi}^*_{V_1}) = -\frac{(1+k)\,(2+k)\,(k-2\,\tilde{t})\,(k+2\,\tilde{t})\,(1+\gamma)}{3\,k^3}.$$

Now $(k - 2\tilde{t})(k + 2\tilde{t}) \geq 0$ for $|\tilde{t}| \leq \dfrac{k}{2}$ and $\dfrac{(1+k)(2+k)(1+\gamma)}{3k^3} > 0$ for all $k > 0$ and $\gamma \geq 0$. Thus $\phi_V(\tilde{t}, \tilde{\xi}^*_{V_1}) \leq 0$ for all $\tilde{t} \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}\}$. Furthermore equality holds at the support designs of $\tilde{\xi}^*_{V_1}$. Thus by the Equivalence Theorem of Theorem 2.6.3, $\xi^*_{V_1}$ is the $V$-optimal design for the mean responses $\boldsymbol{\mu}_g$ over the set of all possible one-point designs for all $\gamma \geq 0$. □

Thus the $V$-optimal population design based on one-point individual designs allocates the measurements equally to designs based on the extreme points 0 and $k$. Further it is robust to the choice of variance ratio $\gamma$.

### 5.2.2 Designs based on $d$-point individual designs

The $V$-optimal population designs based on $d$-point individual designs with non-repeated points, where $d$ is an integer in the interval $[2, k+1]$, are presented in the following two theorems.

**Theorem 5.2.2** *Consider model (5.1) and the set of all d-point individual designs which put equal weights on the distinct time points $t_1, t_2, \ldots, t_d$ with $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$, and $0 \leq t_1 < t_2 < \ldots < t_d \leq k$ for d an even integer greater than or equal to 2. Then*

$$\xi^*_{V_e} = \left\{ \begin{array}{c} (0, 1, \ldots, \frac{d}{2} - 1, k - \frac{d}{2} + 1, \ldots, k-1, k) \\ 1 \end{array} \right\}$$

*is the V-optimal population design for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ over this set for all $\gamma \geq 0$.*

### Proof

Note that for $d = k+1$ with $k$ odd there is only one $d$-point individual design so this necessarily comprises the required $V$-optimal population design. In the remainder of the proof $d$ is therefore taken to be strictly less than $k+1$.

Consider the individual designs $\mathbf{t}$ linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$ where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$ is the center of the hypercube $C_{d,k}$. This is equivalent to moving each element $t_j$ in $\mathbf{t}$ to $\tilde{t}_j = t_j - \frac{k}{2}$, $j = 1\ldots, d$. Then the proposed optimum design $\xi^*_{V_e}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{V_e} = \left\{ \begin{array}{c} (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d}{2}-1, \frac{k}{2}-\frac{d}{2}+1, \ldots, \frac{k}{2}-1, \frac{k}{2}) \\ 1 \end{array} \right\}.$$

Recall from the proof of Theorem 4.3.2 that the inverse of the information matrix for $\boldsymbol{\beta}$ at the design $\tilde{\xi}^*_{V_e}$ is given by

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}^*_{V_e}) = \begin{pmatrix} 1+d\gamma & 0 \\ 0 & \dfrac{12}{H} \end{pmatrix} \tag{5.5}$$

where $H = d^2 - 3\,d\,(k+1) + 3\,k^2 + 6\,k + 2$. Note that the term $H$ is positive because the information matrix $\mathbf{M}_{\beta}(\tilde{\xi}^*_{V_e})$ is necessarily positive definite. Then the criterion $\Psi_V(\tilde{\xi}_V)$ at $\tilde{\xi}^*_{V_e}$ is given by

$$\Psi_V(\tilde{\xi}^*_{V_e}) = (k+1)\,(1+d\,\gamma) + \frac{1}{H}\,k\,(k+1)\,(k+2).$$

Recall from expression (4.9) that the standardized information matrix for $\boldsymbol{\beta}$ at the $d$-point design $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$ is given by

$$\mathbf{M}_{\beta}(\tilde{\mathbf{t}}) = \begin{pmatrix} \dfrac{1}{1+d\gamma} & \dfrac{1}{d\,(1+d\gamma)}\mathbf{1}'\tilde{\mathbf{t}} \\ \dfrac{1}{d\,(1+d\gamma)}\mathbf{1}'\tilde{\mathbf{t}} & \dfrac{1}{d}\tilde{\mathbf{t}}(\mathbf{I} - \dfrac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}} \end{pmatrix}. \tag{5.6}$$

Substituting the appropriate expressions for $\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_e})$, $\mathbf{M}_\beta(\tilde{\mathbf{t}})$ and $\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g$ into that for the directional derivative, namely (5.3), yields

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_e}) = \frac{12}{d\,H^2}\,k\,(1+k)\,(2+k)\,\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\,\tilde{\mathbf{t}} - \frac{1}{H}\,k\,(k+1)\,(k+2). \tag{5.7}$$

The derivative $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_e})$ in (5.7) is a convex function on the polytope $\tilde{P}_{d,k}$. Thus it is only necessary to check the condition $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_e}) \le 0$ at the vertices of the polytope $\tilde{P}_{d,k}$. Now the general expression for a vertex in the transformed coordinates is given by $\tilde{\mathbf{v}}_{j+1} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+d-j-1, \frac{k}{2}-j+1, \ldots, \frac{k}{2})$, $j = 0, 1, \ldots, d$ and the directional derivative (5.7) at that vertex can be expressed explicitly as

$$\phi_V(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{V_e}) = -\frac{3}{d\,H^2\,(1+d\,\gamma)}\,(d-2\,j)^2\,(k-d+1)\,k\,(k+1)\,(k+2)\,(1+\gamma+k\,\gamma).$$

Since the expressions $(d-2\,j)^2\,(k-d+1)\,k\,(k+1)\,(k+2)\,(1+\gamma+k\,\gamma)$ and $d\,H^2\,(1+d\,\gamma)$ are greater than zero for $2 \le d < k+1$, $j = 0, 1, \ldots, d$ and $\gamma \ge 0$, it then follows that $\phi_V(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{V_e}) \le 0$ . Furthermore, equality holds at $j = \dfrac{d}{2}$ and the vertex $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ is the support design of $\tilde{\xi}^*_{V_e}$. □

**Theorem 5.2.3** *Consider model (5.1) and the set of all d-point individual designs which put equal weights on the distinct time points $t_1, t_2, \ldots, t_d$ with $t_j \in \{0, 1, \ldots, k\}$, $j = 0, 1, \ldots, d$ and $0 \le t_1 < t_2 < \ldots < t_d \le k$ for $d$ an odd integer such that $1 \le d \le k+1$. Then*

$$\xi^*_{V_o} = \left\{ \begin{array}{cc} (0, 1, \ldots, \frac{d-1}{2}, k - \frac{d-3}{2}, \ldots, k) & (0, 1, \ldots, \frac{d-3}{2}, k - \frac{d-1}{2}, \ldots, k) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

*is the $V$-optimal population design for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ over this set for all $\gamma \ge 0$.*

**Proof**

Note that for $d = k+1$ with $k$ even there is only one $d$-point individual design so this is necessarily $V$-optimal. In the remainder of the proof $d$ is therefore taken to be strictly less than $k+1$.

Consider a linear transformation of individual designs $\mathbf{t}$, i.e. $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$ where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$. Then the proposed optimum design $\xi^*_{V_o}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{V_o} = \left\{ \begin{array}{cc} (-\frac{k}{2}, 1-\frac{k}{2}, \ldots, -\frac{k}{2}+\frac{d-1}{2}, \frac{k}{2}-\frac{d-3}{2}, \ldots, \frac{k}{2}) & (-\frac{k}{2}, 1-\frac{k}{2}, \ldots, -\frac{k}{2}+\frac{d-3}{2}, \frac{k}{2}-\frac{d-1}{2}, \ldots, \frac{k}{2}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}.$$

Recall from the proof of Theorem 4.3.3 that the inverse of the information matrix for $\boldsymbol{\beta}$ at $\tilde{\xi}^*_{V_o}$ is given by

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}^*_{V_o}) = \begin{pmatrix} 1+d\gamma & 0 \\ 0 & \dfrac{12\,d\,(1+d\gamma)}{H} \end{pmatrix} \tag{5.8}$$

where

$$H = d^3 - 3\,(k+1)\,(d^2+1) + (3\,k^2+6\,k+5)\,d + (d^2-1)\,\{d^2+3\,(k+1)\,(k-d+1)\}\,\gamma.$$

Note that the term $H$ is greater than zero because the information matrix $\mathbf{M}_{\beta}(\tilde{\xi}^*_{V_o})$ is necessarily positive definite. It now follows immediately from this expression for $\mathbf{M}_{\beta}(\tilde{\xi}^*_{V_o})$, and from the expressions for $\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g$ in (5.4) and $\mathbf{M}_{\beta}(\tilde{\mathbf{t}})$ in (5.6), that the $V$-optimality criterion at $\tilde{\xi}^*_{V_o}$ is given by

$$\Psi_V(\tilde{\xi}^*_{V_o}) = (k+1)\,(d\gamma+1)\left\{1+\frac{1}{H}\,d\,k\,(k+2)\right\}$$

and that the directional derivative of $\Psi_V(\tilde{\xi})$ at $\tilde{\xi}^*_{V_o}$ in the direction of a $d$-point individual design $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$ can be written as

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_o}) = \frac{1}{H} d\, k\,(k+1)\,(k+2)(d\gamma+1)\left\{\frac{12\,d\,(d\gamma+1)}{H}\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}} - 1\right\}. \quad (5.9)$$

The expression for $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_o})$ in (5.9) is a convex function on the polytope $\tilde{P}_{d,k}$ and thus it is only necessary to consider the condition $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_o}) \leq 0$ at the vertices $\tilde{\mathbf{v}}_{j+1}$, $j = 0, 1, \ldots, d$ of that polytope. Now $\tilde{\mathbf{v}}_{j+1} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+d-j-1, \frac{k}{2}-j+1, \ldots, \frac{k}{2})$ and the directional derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_o})$ at that vertex has the form

$$\phi_V(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{V_o}) = -\frac{3}{H^2}\{d\,[(d-2\,j)^2-1]\,(k-d+1)\,k\,(k+1)\,(k+2)\,(1+d\gamma)\,(1+\gamma+k\gamma)\}.$$

The sign of $\phi_V(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{V_o})$ depends on the sign of $[(d-2\,j)^2-1]\,(k-d+1)$ since

$$\frac{3}{H^2}\,k\,(k+1)\,(k+2)\,(1+d\gamma)\,(1+\gamma+k\gamma) > 0$$

for $\gamma \geq 0$. Now since $d < k+1$, $k-d+1 > 0$. Also $(d-2\,j)^2 - 1 \geq 0$ for $j = 0, 1, \ldots, d$. Therefore $[(d-2\,j)^2-1]\,(k-d+1) \geq 0$ for $j = 0, 1, \ldots, d$ and $1 \leq d < k+1$. Thus $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_o}) \leq 0$ for any $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$ and for all $\gamma \geq 0$. Furthermore, the equality $\phi_V(\tilde{\mathbf{v}}_{j+1}, \tilde{\xi}^*_{V_o}) = 0$ holds only at $j = \dfrac{d-1}{2}$ and $\dfrac{d+1}{2}$ and hence at the support designs of $\tilde{\xi}^*_{V_o}$. $\square$

The $V$-optimal population designs derived in Theorems 5.2.2 and 5.2.3 are the same as the corresponding $D$-optimal population designs of Theorems 4.3.2 and 4.3.3 respectively. This coincidence is based on the fact that the directional derivatives for both $V$- and $D$-optimality are proportional to the quadratic form $\tilde{\mathbf{t}}'(\mathbf{I} - \dfrac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}}$.

## 5.3 Comparison of $V$-optimal population designs

The design $\xi_1$ is more efficient than the design $\xi_2$ in estimating mean responses $\boldsymbol{\mu}_g$ if

$$tr\{\mathbf{M}_\beta^{-1}(\xi_1)\mathbf{X}_g'\mathbf{X}_g\} < tr\{\mathbf{M}_\beta^{-1}(\xi_2)\mathbf{X}_g'\mathbf{X}_g\}.$$

Thus $\Psi_V(\xi) = tr\{\mathbf{M}_\beta^{-1}(\xi)\mathbf{X}_g'\mathbf{X}_g\}$ is a measure of the $V$-efficiency of a particular design $\xi$ for the mean responses $\boldsymbol{\mu}_g$ corresponding to $\mathbf{X}_g$. The $V$-optimal population designs derived in Theorems 5.2.2 and 5.2.3 are compared on a per point basis in this way and the results are presented in the following theorem.

**Theorem 5.3.1** *Let the constants $d_e$ and $d_o$ be even and odd integers with $2 \leq d_e \leq k+1$ and $3 \leq d_o \leq k+1$. Then the $V$-optimal population designs $\xi^*_{V_e}$ and $\xi^*_{V_o}$ are the most efficient designs on a per observation basis for mean responses $\boldsymbol{\mu}_g$ corresponding to $\mathbf{t}_g = (0, 1, \ldots, k)'$ over the set of $d$-point individual designs defined on the space of designs $S_{d,k}$ with $d \geq d_o$ and $d \geq d_e$, respectively. Furthermore, the $V$-efficiency decreases on a per observation basis as $d$ increases.*

**Proof**

Let $d_e$ be a positive even integer. To prove the theorem it is only necessary to show that the inequalities

$$\Psi(\tilde{\xi}^*_{V_{d_e}}) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{d_e}})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\} < tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{(d_e+1)}})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\} = \Psi(\tilde{\xi}^*_{V_{(d_e+1)}})$$

for $2 \leq d_e \leq k$ and

$$\Psi(\tilde{\xi}^*_{V_{(d_e+1)}}) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{(d_e+1)}})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\} < tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{(d_e+2)}})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\} = \Psi(\tilde{\xi}^*_{V_{(d_e+2)}})$$

for $0 < d_e \leq k-1$ hold, where $\tilde{\xi}^*_{V_{d_e}}$, $\tilde{\xi}^*_{V_{(d_e+1)}}$ and $\tilde{\xi}^*_{V_{(d_e+2)}}$ are the $V$-optimal designs in the transformed coordinates based on individual designs with $d_e$, $d_{e+1}$ and $d_{e+2}$ respectively. Equivalently it is only necessary to show that the differences

$$D_1 = \Psi(\tilde{\xi}^*_{V_{d_e}}) - \Psi(\tilde{\xi}^*_{V_{(d_e+1)}})$$

and

$$D_2 = \Psi(\tilde{\xi}^*_{V_{(d_e+1)}}) - \Psi(\tilde{\xi}^*_{V_{(d_e+2)}})$$

are less than zero for $2 \leq d_e \leq k$.

Consider first the difference

$$D_1 = -\frac{(k+1)\{A_2\,\gamma^2 + A_1\,\gamma + A_0\}}{W_1\,\{d_e\,(d_e+2)\,W_3 + W_2\}}$$

where

$$W_1 = (d_e-2)\,(d_e-1) + 3\,k\,(k-d_e) + 6\,k,$$

$$W_2 = 3\,k\,d_e\,(k-d_e) + 3\,k^2 + d_e^3 + 2\,d_e,$$

$$W_3 = 1 - d_e + d_e^2 + 3\,k + 3\,k\,(k-d_e)$$

since $d_e \leq k$, $W_i > 0$, $i = 1, 2, 3$. Furthermore,

$$A_0 = k\,(k+2)\,(d_e+2)\,(3\,k-2\,d_e+1) > 0, \quad \text{since } d_e \leq k,$$

$$\begin{aligned} A_1 = {} & 4\,(k+d_e) + 20\,k^2 + 4\,k\,d_e + 6\,d_e\,(k-d_e) + 14\,k^3 + 8\,k^2\,d_e + 16\,k\,(k^2-d_e^2) + 4\,d_e^3 \\ & + 3\,(k^4-d_e^4) + 8\,k^2\,(k^2-d_e^2) + 12\,k^2\,d_e\,(k-d_e) + 11\,k\,d_e^3 + d_e^5 + 6\,k\,d_e^3\,(k-d_e) \\ & + k^2\,d_e\,(9\,k^2 - 15\,k\,d_e + 7\,d_e^2) > 0 \end{aligned}$$

since $2 \le d_e \le k$,

$$9\,k^2 - 15\,k\,d_e + 7\,d_e^2 = (3\,k - \frac{5}{2}\,d_e)^2 + \frac{3}{4}\,d_e^2 > 0$$

by completing the square and

$$A_2 = W_3\,d_e\,(d_e+2)\,\{2+3\,k+d_e^2+3\,(k+1)\,(k-d_e)\} > 0$$

since $2 \le d_e \le k$. Thus $D_1 < 0$ for $2 \le d_e \le k$.

Consider now the difference

$$D_2 = -\frac{2\,(k+1)\{B_2\,\gamma^2 + B_1\,\gamma + B_0\}}{W_4\,\{d_e\,(d_e+2)\,W_3 + W_2\}}$$

where

$$W_4 = d_e\,(d_e+1)+3\,k\,(k-d_e) > 0, \text{ since } d_e \le k-1$$

$$B_0 = k\,(k+2)\,(3\,k - 2\,d_e + 1)\,d_e > 0, \text{ since } d_e \le k-1$$

$$\begin{aligned} B_1 &= 6\,k^3\,(k-1)+k^2\,d_e\,(3\,k-\frac{5}{2}d_e)^2+k\,d_e\,(k-d_e)\,(14+7\,d_e+6\,d_e^2)+k^2\,d_e\,(14+d_e+\frac{3}{4}d_e^2) \\ &\quad +2\,k\,d_e + 2\,d_e^2 + 2\,d_e^2 + d_e^4 + d_e^5 > 0 \end{aligned}$$

since $d_e \le k-1$ and

$$B_2 = d_e\,(d_e+2)\,W_3\,W_4 > 0, \text{ since } W_3 > 0 \text{ and } W_4 > 0.$$

Thus the difference $D_2 < 0$ for $0 \le d_e \le k-1$. □

The result in Theorem 5.3.1 demonstrates that the mean responses for a given vector of time points can be estimated more precisely by using an optimum design with small number of time points in its support.

Theorem 5.3.1 compares $V$-optimal designs for $d \geq 2$ and the results are always true for all $\gamma \geq 0$. However, for $V$-optimal population designs based on one- and two-point individual designs the theorem does not hold for $\gamma = 0$. Consider the set of all one-and two-point individual designs. In Subsection 5.2.1 it has been proved that

$$\xi^*_{V_1} = \left\{ \begin{array}{cc} (0) & (k) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

is the $V$-optimal population design over the set of one-point individual designs for all $\gamma \geq 0$. Further, from Theorem 5.2.2, the design

$$\xi^*_{V_2} = \left\{ \begin{array}{c} (0, k) \\ 1 \end{array} \right\}$$

is the $V$-optimal population design over the set of two-point individual designs for all $\gamma \geq 0$. The linearly transformed versions of these designs yield the criteria

$$\Psi_V(\tilde{\xi}^*_{V_1}) = \frac{2}{3\,k}\,(1+\gamma)\,(2\,k+1)\,(k+1)$$

and

$$\Psi_V(\tilde{\xi}^*_{V_2}) = \frac{2}{3\,k}\,(k+1)\,(1+2\,k+3\,k\,\gamma).$$

The ratio

$$\frac{\Psi_V(\tilde{\xi}^*_{V_1})}{\Psi_V(\tilde{\xi}^*_{V_2})} = \frac{(1+\gamma)\,(2\,k+1)}{3\,k\,\gamma+2\,k+1} \leq 1$$

for any $k \geq 1$ and $\gamma \geq 0$ and equality holds only for $\gamma = 0$. Thus $\xi^*_{V_1}$ is more efficient than $\xi^*_{V_2}$ only for $\gamma > 0$.

**Example 5.3.1** Consider model (5.1) and let $k = 6$. Then it follows from Theorems 5.2.1, 5.2.2 and 5.2.3 that the $V$-optimal population designs based on $d$-point individual designs where $1 \leq d \leq 6$ are given by

$$\xi^*_{V_1} = \left\{ \begin{array}{cc} (0) & (6) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}, \xi^*_{V_2} = \left\{ \begin{array}{c} (0,6) \\ 1 \end{array} \right\}, \quad \xi^*_{V_3} = \left\{ \begin{array}{cc} (0,1,6) & (0,5,6) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\},$$

$$\xi^*_{V_4} = \left\{ \begin{array}{c} (0,1,5,6) \\ 1 \end{array} \right\}, \qquad \xi^*_{V_5} = \left\{ \begin{array}{cc} (0,1,2,5,6) & (0,1,4,5,6) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

$$\xi^*_{V_6} = \left\{ \begin{array}{c} (0,1,2,4,5,6) \\ 1 \end{array} \right\} \qquad \text{and} \qquad \xi^*_{V_7} = \left\{ \begin{array}{c} (0,1,2,3,4,5,6) \\ 1 \end{array} \right\}.$$

Note that for $k = 6$ there is only one seven-point design $\xi^*_{V_7}$ and it is necessarily $V$-optimal.

The variance of an estimated response at time $t$ for a design $\xi^*_{V_d}$ is given by

$$\Psi(\xi^*_{V_d}) = \mathbf{x}'_g \, \mathbf{M}^{-1}_\beta(\xi^*_{V_d}) \, \mathbf{x}_g$$

where $\mathbf{x}_g = [1 \quad t]'$ and $\mathbf{M}^{-1}_\beta(\xi^*_{V_d})$ is given in transformed coordinates by expression (5.5) for $d$ even and by expression (5.8) for $d$ odd in the Theorems of Section 5.2. Therefore for the above designs the variances are

$$\begin{aligned}
Var(\xi^*_{V_1}, t) &= \frac{1}{9}(18 - 6\,t + t^2)\,(1+\gamma), \\
Var(\xi^*_{V_2}, t) &= \frac{1}{9}(18 - 6\,t + t^2 + 18\,\gamma), \\
Var(\xi^*_{V_3}, t) &= \frac{(1+3\,\gamma)\,(49 - 18\,t + 3\,t^2 + 62\,\gamma)}{2\,(11+31\,\gamma)}, \\
Var(\xi^*_{V_4}, t) &= \frac{1}{13}(31 - 12\,t + 2\,t^2 + 52\,\gamma), \\
Var(\xi^*_{V_5}, t) &= \frac{(1+5\,\gamma)\,(72 - 30\,t + 5\,t^2 + 134\,\gamma)}{27 + 134\,\gamma}, \\
Var(\xi^*_{V_6}, t) &= \frac{1}{14}(41 - 18\,t + 3\,t^2 + 84\,\gamma)
\end{aligned}$$

and

$$Var(\xi^*_{V_7}, t) = \frac{1}{4}(13 - 6\,t + t^2 + 28\,\gamma).$$

Note that the variances depend on the variance ratio $\gamma$ but that the actual optimal designs do not. Plots of these variances against $t$ for $\gamma = 0.25$ and $t = \{0, 1, \ldots, 6\}$ are presented in Figure 5.1. The plots show that for fixed $\gamma$ the variances decrease as the number of time points $d$ decreases.

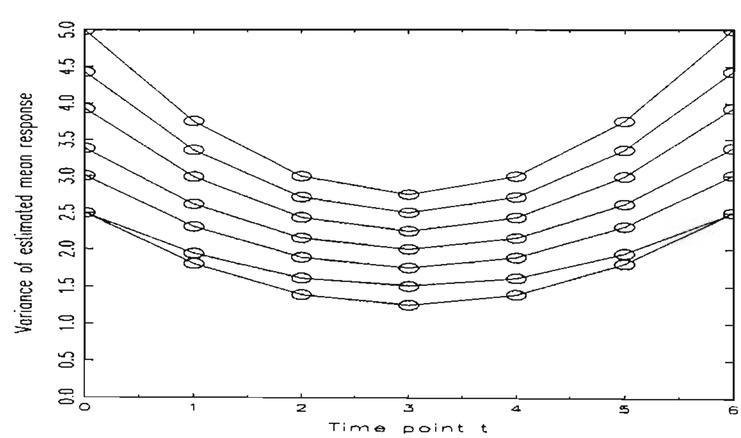


Figure 5.1: Variances of the estimated mean responses for the $V$-optimal population designs based on $d$-point individual designs with $\gamma = 0.25$ and $k = 6$ as a function of time $t$. The uppermost curve corresponds to $\xi^*_{V_7}$ and the lowest to $\xi^*_{V_1}$.

## 5.4 Best $V$-optimal population designs based on designs with non-repeated time points

In Chapter 4 it was shown that

$$\xi_D^* = \left\{ \begin{array}{c} (0, k) \\ \\ 1 \end{array} \right\}$$

is the $D$-optimal population design for the fixed effects $\boldsymbol{\beta}$ in the simple linear regression model with a random intercept over all population designs. Further, from Theorem 5.2.2, this design is also the $V$-optimal population design based on the set of all two-point individual designs. However, it is not the $V$-optimal population design over all population designs as is demonstrated in the following discussion.

Recall from Subsection 5.2.2 that the inverse of the information matrix for $\boldsymbol{\beta}$ at the design

$$\tilde{\xi}_{V_2}^* = \left\{ \begin{array}{c} (-\frac{k}{2}, \frac{k}{2}) \\ \\ 1 \end{array} \right\}$$

is

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}_{V_2}^*) = \begin{pmatrix} 1+2\gamma & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}$$

and recall also from expression (5.4) that

$$\tilde{\mathbf{X}}_g' \tilde{\mathbf{X}}_g = \begin{pmatrix} k+1 & 0 \\ 0 & \dfrac{k(k+2)(k+1)}{12} \end{pmatrix}.$$

Then the directional derivative of $\Psi_V(\tilde{\xi}) = tr\left\{\mathbf{M}_{\beta}^{-1}(\tilde{\xi})\, \tilde{\mathbf{X}}_g' \tilde{\mathbf{X}}_g\right\}$ at $\tilde{\xi}_{V_2}^*$ in the direction of a one-point design $\tilde{t}$ is equal to

$$\phi_V(\tilde{t}, \tilde{\xi}_{V_2}^*) = \frac{(k+1)\left\{4(k+2)\,\tilde{t}^2 + k^2\left[6\,k\,\gamma^2 + 2\,\gamma\,(k-1) - (k+2)\right]\right\}}{3\,k^3\,(1+\gamma)}.$$

At the one-point support designs of the $V$-optimal population design

$$\tilde{\xi}^*_{V_1} = \left\{ \begin{array}{cc} (-\frac{k}{2}) & (\frac{k}{2}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

this directional derivative is given by

$$\phi_V(-\frac{k}{2}, \tilde{\xi}^*_{V_2}) = \phi_V(\frac{k}{2}, \tilde{\xi}^*_{V_2}) = \frac{2\,(k+1)\,\gamma\,(3\,k\,\gamma + k - 1)}{3\,k\,(1+\gamma)} > 0$$

for any $k \geq 1$ and $\gamma > 0$. Therefore $\tilde{\xi}^*_{V_2}$ and hence $\xi^*_{V_2}$ is not optimal over the set of one-point individual designs for all $\gamma > 0$.

Consider also the population design $\tilde{\xi}^*_{V_1}$. The inverse of the information matrix for $\boldsymbol{\beta}$ at $\tilde{\xi}^*_{V_1}$ is given by

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}^*_{V_1}) = (1+\gamma) \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}.$$

and the directional derivative of $\Psi_V(\tilde{\xi})$ at $\tilde{\xi}^*_{V_1}$ in the direction of a two-point design $\tilde{\mathbf{t}} = (-\frac{k}{2}, \frac{k}{2})$ is equal to

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) = \frac{2\,\gamma\,(1+\gamma)\,(1+k)\,\{(1-k) + \gamma\,(k+2)\}}{3\,k\,(1+2\,\gamma)}.$$

The sign of this expression is determined by the sign of $(1-k) + \gamma\,(k+2)$ since $\dfrac{2\,\gamma\,(1+\gamma)\,(1+k)}{3\,k\,(1+2\,\gamma)} > 0$ for $k \geq 1$ and $\gamma > 0$. For a given $k$, $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) \leq 0$ if and only if $(1-k) + \gamma\,(k+2) \leq 0$ and thus if and only if $\gamma \leq \dfrac{k-1}{k+2}$. Therefore the design $\xi^*_{V_1}$ is only $V$-optimal over the set of two-point designs when $\gamma \leq \dfrac{k-1}{k+2}$.

From the above discussion it can be argued, at least intuitively, that in order to estimate the mean responses precisely an optimum population design which combines both $\tilde{\xi}^*_{V_1}$ and $\tilde{\xi}^*_{V_2}$ is needed. This is now demonstrated in the following theorem.

**Theorem 5.4.1** *Consider the set of population designs based on all possible $d$-point individual designs $\mathbf{t}$ which put equal weights on the distinct time points $t_1, t_2, \ldots, t_d$ with $t_j \in \{0, 1, \ldots, k\}$, $j = 1, 2, \ldots, d$, and $0 \leq t_1 < t_2 \ldots t_d \leq k$ for $d$ a positive integer in $[1, k+1]$. Then the $V$-optimal population design for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ in model (5.1) over this set is given by*

$$\xi^*_{V_1} = \left\{ \begin{array}{cc} (0) & (k) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\},$$

*for $\gamma \leq \gamma(k) = \dfrac{k-1}{k+2}$ and by*

$$\xi^*_{V_c} = \left\{ \begin{array}{ccc} (0) & (k) & (0, k) \\ w & w & 1 - 2w \end{array} \right\},$$

*where*

$$w = \frac{(1+\gamma)\left\{k\,(2+3\,\gamma) + 1 - \sqrt{3k(2+k)(1+2\,\gamma)}\right\}}{2\,\gamma\,(3\,k\,\gamma + k - 1)}$$

*for $\gamma > \gamma(k)$.*

**Proof**

Consider the linear transformation of a single point design $t_j$ as $\tilde{t}_j = t_j - \frac{k}{2}$, $j = 1, \ldots, d$, so that a $d$-point individual design is transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$ where $\mathbf{x}_c$ is the $d \times 1$ vector $(\frac{k}{2}, \ldots, \frac{k}{2})$. Then the proposed optimal designs $\xi^*_{V_1}$ and $\xi^*_{V_c}$ in the transformed coordinates are given by

$$\tilde{\xi}^*_{V_1} = \left\{ \begin{array}{cc} (-\frac{k}{2}) & (\frac{k}{2}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

and

$$\tilde{\xi}^*_{V_c} = \left\{ \begin{array}{ccc} (-\frac{k}{2}) & (\frac{k}{2}) & (-\frac{k}{2}, \frac{k}{2}) \\ w & w & 1 - 2w \end{array} \right\}$$

respectively. Note also that in the transformed coordinates the mean responses are estimated at $\tilde{\mathbf{t}}_g = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2})'$ and thus that

$$\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g = \begin{pmatrix} k+1 & 0 \\ 0 & \dfrac{k(k+2)(k+1)}{12} \end{pmatrix}.$$

The proof of the theorem is accomplished in two steps. The first step shows that $\tilde{\xi}^*_{V_1}$ is $V$-optimal over all possible $d$-point population designs for $\gamma \leq \gamma(k) = \dfrac{k-1}{k+2}$ and then the second step shows that $\tilde{\xi}^*_{V_c}$ is $V$-optimal over all possible $d$-point population designs for $\gamma > \dfrac{k-1}{k+2}$.

**Step I**

Recall from Subsection 5.2.1 that the inverse of the standardized information matrix for $\boldsymbol{\beta}$ and the criterion function at the population design $\tilde{\xi}^*_{V_1}$ are, respectively

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}^*_{V_1}) = (1+\gamma)\begin{pmatrix} 1 & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}$$

and

$$\Psi_V(\tilde{\xi}^*_{V_1}) = \frac{2}{3\,k}\,(1+\gamma)\,(2\,k+1)\,(k+1).$$

Note that the within individual variance $\sigma_e^2$ is equal to 1 by assumption. Therefore the directional derivative of the criterion $\Psi_V(\tilde{\xi}) = tr\{\mathbf{M}_{\beta}^{-1}(\tilde{\xi})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\}$ at $\tilde{\xi}^*_{V_1}$ in the direction of a $d$-point individual design $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$ has the form

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) = \frac{(1+\gamma)\,(k+1)}{3\,k}\left\{\frac{4\,(k+2)\,(1+\gamma)}{k^2\,d}\,\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\,\gamma}\mathbf{J})\,\tilde{\mathbf{t}} - \frac{2\,(d\,\gamma+1)\,(2\,k+1) - 3\,k\,(1+\gamma)}{(d\,\gamma+1)}\right\}. \qquad (5.10)$$

Now to prove that $\tilde{\xi}^*_{V_1}$ is the $V$-optimal population design it is only necessary to show that the condition $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) \leq 0$ holds for any $d$-point individual design $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$. Since $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ is a convex function on the polytope $\tilde{P}_{d,k}$ it is only necessary to check that $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) \leq 0$ at the extreme vertices of $\tilde{P}_{d,k}$.

(1) For the case $d = 1$, it has already been proved in Theorem 5.2.1 that the population design $\xi^*_{V_1}$ is $V$-optimal over the set of all one-point individual designs for all $\gamma \geq 0$ and thus for $\gamma \leq \dfrac{k-1}{k+2}$.

(2) When $d = 2$, the extreme vertex of $\tilde{P}_{2,k}$ is $\tilde{\mathbf{t}}_2 = (-\frac{k}{2}, \frac{k}{2})$. Then

$$\phi_V(\tilde{\mathbf{t}}_2, \tilde{\xi}^*_{V_1}) = \frac{2\gamma(1+\gamma)(1+k)\{1-k+\gamma(k+2)\}}{3k(1+2\gamma)}.$$

and, as has been shown earlier, for a given $k$ $\phi_V(\tilde{\mathbf{t}}_2, \tilde{\xi}^*_{V_1}) \leq 0$ if and only if $\gamma \leq \dfrac{k-1}{k+2}$.

(3) Next, consider $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ in (5.10) for all $d$-point designs $\tilde{\mathbf{t}}$ with $d \geq 3$. The values of $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ associated with the contours

$$\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}} = c$$

where $c$ is a constant, are of exactly the same form as those described in the proof of Theorem 5.2.3. Thus the largest value of $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ over the space of designs $\tilde{S}_{d,k}$ is attained at the vertex $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ when $d$ is even and at the vertices $\tilde{\mathbf{v}}_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}_{\frac{d+3}{2}}$ when $d$ is odd and it is only necessary to check the condition $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) \leq 0$ at these vertices.

Consider first the case of $d$ even and $d > 2$. Recall that the vertex

$$\tilde{\mathbf{v}}_{\frac{d}{2}+1} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d}{2}-1, \frac{k}{2}-\frac{d}{2}+1, \frac{k}{2}-\frac{d}{2}+2, \ldots, \frac{k}{2}-1, \frac{k}{2})$$

is the support design for the $V$-optimal population design expressed in the transformed coordinates and derived in Theorem 5.2.2. At $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ the directional derivative is given by

$$\phi_V(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_1}) = \frac{(1+k)\,(1+\gamma)}{9\,k^3\,(1+d\,\gamma)}\,F \tag{5.11}$$

where

$$\begin{aligned} F = {} & \gamma\,(1+\gamma)\,(k+2)\,d^3 - (1+\gamma)\,(k+2)\,\{3\,(k+1)\,\gamma - 1\}\,d^2 + \{(k+2)\,(2+6\,k+3\,k^2)\,\gamma^2 \\ & - (3\,k-2)\,(3\,k^2+k-1)\,\gamma - 3\,(k+1)\,(k+2)\}\,d + 2\,(2+7\,k+6\,k^2+6\,k^3)\,\gamma \\ & + 2\,(k+2)\,(3\,k+1). \end{aligned}$$

The sign of the directional derivative in (5.11) depends on the sign of $F$ because $\dfrac{(1+k)\,(1+\gamma)}{9\,k^3\,(1+d\,\gamma)}$ $> 0$ for $k \geq 1$ and $\gamma \geq 0$.

Consider now the function $F$ which is cubic in $d$. The equation

$$\frac{\partial F}{\partial d} = 0$$

yields the stationary points of $F$ located at

$$d_1 = \frac{A - B}{3\,\gamma\,(2+k)\,(1+\gamma)}$$

and

$$d_2 = \frac{A + B}{3\,\gamma\,(2+k)\,(1+\gamma)},$$

where

$$A = (2+k)\,(1+\gamma)\,\{3\,(1+k)\,\gamma - 1\}$$

and

$$B = \sqrt{\{2+k+(8+10\,k+3\,k^2)\,\gamma + 3\,(4+4\,k+7\,k^2+12\,k^3)\,\gamma^2 + 3\,(2+k)\,\gamma^3\}}$$

$$\times \sqrt{(1+\gamma)\,(k+2)}.$$

Note that $d_1 < d_2$ since $B > 0$. The second derivative of $F$ with respect to $d$ has the values

$$\left.\frac{\partial^2 F}{\partial d^2}\right|_{d=d_1} = -2\,B < 0$$

and

$$\left.\frac{\partial^2 F}{\partial d^2}\right|_{d=d_2} = 2\,B > 0$$

at $d_1$ and $d_2$, respectively. Therefore the function $F$ attains its maximum at $d_1$ and its minimum at $d_2$.

Consider first the root $d_1$. Then it can be shown that $d_1 \leq 0$ for $0 \leq \gamma \leq \dfrac{k-1}{k+2}$. Specifically, since $B > 0$, then for $A < 0$, and thus $\gamma < \dfrac{1}{3\,(k+1)}$, it follows immediately that the root $d_1$ is negative. Consider now the case of $A \geq 0$, and thus $\gamma \geq \dfrac{1}{3\,(k+1)}$. Observe that $A^2 - B^2 < 0$ implies that $A - B < 0$. Now

$$A^2 - B^2 = 3\,(2+k)\,\gamma\,(1+\gamma)\,f(\gamma) < 0$$

where

$$f(\gamma) = (3\,k^2 + 6\,k + 2)\,(k+2)\,\gamma^2 - (3\,k^2 + k - 1)\,(3\,k - 2)\,\gamma - 3\,(k+2)\,(k+1)$$

and clearly the sign of $A^2 - B^2$ is determined by that of $f(\gamma)$. The function $f(\gamma)$ is quadratic in $\gamma$ and, since $(3\,k^2 + 6\,k + 2)\,(k+2) > 0$, represents a parabola which opens upwards. Furthermore

$$f\left(\frac{1}{3\,(k+1)}\right) = -\frac{56 + 166\,k + 207\,k^2 + 150\,k^3 + 54\,k^4}{9\,(k+1)^3} < 0$$

and

$$f\left(\frac{k-1}{k+2}\right) = -\frac{(2\,k+1)(8 + 13\,k - 6\,k^2 + 3\,k^3)}{k+2} < 0.$$

Note that in the latter case the term $8 + 13\,k - 6\,k^2 + 3\,k^3$ has two imaginary and one negative root, its first derivative with respect to $k$, $3\,k\,(k-2) + 13$, is greater than zero for $k \geq 1$ and approaches $\infty$ as $k$ approaches $\infty$ and is therefore positive for $k \geq 1$. Thus $f(\gamma)$ is necessarily less than 0 in the range $\dfrac{1}{3\,(k+1)} \leq \gamma \leq \dfrac{k-1}{k+2}$ and this in turn implies that $A^2 - B^2 < 0$ and hence $A - B < 0$. Thus $d_1$ itself are negative over that range.

Now compare the stationary point $d_2$ with the maximum possible value of $d$, i.e. with $k+1$. Then

$$d_2 - (k+1) = \frac{B - (k+2)\,(1+\gamma)}{3\,\gamma\,(2+k)\,(1+\gamma)}.$$

Since

$$B^2 - (k+2)^2\,(1+\gamma)^2 = 3\,(2+k)\,\gamma\,(1+\gamma)\,\{2 + 3\,k + k^2 + (4 + 4\,k + 7\,k^2 + 12\,k^3)\,\gamma + (2+k)\,\gamma^2\}$$

is strictly greater than zero for $k \geq 1$ and $\gamma > 0$, it follows that $B > (k+2)\,(1+\gamma)$ which implies that $d_2 > k+1$. Therefore the solutions $d_1$ and $d_2$ do not fall in the interval $[1, k+1]$.

To complete the proof for $d$ even, consider the property of $F$. Recall from the earlier discussion that the function $F$ has a maximum at $d_1$ and a minimum at $d_2$. Further $d_1 < 0$ for $0 \leq \gamma \leq \dfrac{k-1}{k+2}$ and $d_2 > k+1$. Thus $F$ is decreasing on the interval $[d_1, d_2]$ and hence on the interval $[2, k+1]$. The maximum of $F$ in the second interval occurs at $d = 2$ and is given by

$$F_2 = 6\,k^2\,\gamma\,\{1 - k + \gamma\,(k+2)\}$$

and clearly $F_2 \leq 0$ for $\gamma \leq \dfrac{k-1}{k+2}$. It thus follows that the function $F$ is less than or equal to zero in the interval $[2, k+1]$ for $0 \leq \gamma \leq \dfrac{k-1}{k+2}$ and hence $\phi_V(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*{}_{V_1}) \leq 0$.

Consider now the case of $d$ odd with $d \geq 3$. Recall also that the vertices

$$\tilde{\mathbf{v}}_{\frac{d+1}{2}} = (-\frac{k}{2}, -\frac{k}{2} + 1, \ldots, -\frac{k}{2} + \frac{d-1}{2}, \frac{k}{2} - \frac{d-3}{2}, \ldots, \frac{k}{2} - 1, \frac{k}{2})$$

and

$$\tilde{\mathbf{v}}_{\frac{d+3}{2}} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+\frac{d-3}{2}, \frac{k}{2}-\frac{d-1}{2}, \ldots, \frac{k}{2}-1, \frac{k}{2})$$

are the support designs for the $V$-optimal population design in the transformed coordinates and derived in Theorem 5.2.3. At both $\tilde{\mathbf{v}}_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}_{\frac{d+3}{2}}$, the directional derivative is given by

$$\phi_V(\tilde{\mathbf{v}}_{\frac{d+1}{2}}, \tilde{\xi}^*_{V_1}) = \phi_V(\tilde{\mathbf{v}}_{\frac{d+3}{2}}, \tilde{\xi}^*_{V_1}) = \frac{(d-1)\,(k+1)\,(1+\gamma)}{9\,d\,k^3\,(1+d\,\gamma)}\,H \tag{5.12}$$

where

$$H = \gamma\,(1+\gamma)\,(k+2)\,d^3 - (1+\gamma)\,(k+2)\{(3\,k+2)\,\gamma - 1\}\,d^2 + \{3\,k\,(k+1)\,(k+2)\,\gamma^2$$
$$-\,(3\,k+2)\,(2-2\,k+3\,k^2)\,\gamma - (k+2)\,(3\,k+2)\}\,d + 3\,(1+\gamma)\,(k+1)\,(k+2)\{(k+1)\,\gamma + 1\}$$

which is a cubic function in $d$. Since the terms $(d-1)\,(k+1)\,(1+\gamma)$ and $9\,d\,k^3\,(1+d\,\gamma)$ are positive for $3 \leq d \leq k+1$ and $\gamma > 0$, the sign of the expression in (5.12) takes the sign of $H$. Therefore it is only necessary to examine $H$. The equation

$$\frac{\partial H}{\partial d} = 0$$

yields the stationary points of $H$ as

$$d_1 = \frac{A - C}{3\,\gamma\,(2+k)\,(1+\gamma)}$$

and

$$d_2 = \frac{A + C}{3\,\gamma\,(2+k)\,(1+\gamma)},$$

where

$$A = (2+k)\,(1+\gamma)\,\{3\,(2+k)\,\gamma - 1\}$$

and

$$C = \sqrt{\{2+k+3\,(k+1)\,(k+2)\,\gamma + 6\,(3\,k+2)\,(2\,k^2+1)\,\gamma^2 + (8+10\,k+3\,k^2)\,\gamma^3\}}$$

$$\times \sqrt{(1+\gamma)\,(k+2)}.$$

Note that since $C > 0$, $d_1 < d_2$. The second derivative of $H$ with respect to $d$ has the values

$$\left.\frac{\partial^2 H}{\partial d^2}\right|_{d=d_1} = -2\,C < 0$$

and

$$\left.\frac{\partial^2 H}{\partial d^2}\right|_{d=d_2} = 2\,C > 0$$

at $d_1$ and $d_2$, respectively. Therefore $H$ attains its maximum at $d_1$ and its minimum at $d_2$.

The properties of the roots $d_1$ and $d_2$ depend on the sign of expression $A$. If $A < 0$ and thus $\gamma < \dfrac{1}{3\,(k+2)}$ then $d_1 < 0$. Consider now the case of $A \geq 0$ and thus $\gamma \geq \dfrac{1}{3\,(k+2)}$. Now

$$A^2 - C^2 = 3\,(2+k)\,\gamma\,(1+\gamma)\,g(\gamma)$$

where

$$g(\gamma) = 3\,k\,(k+1)\,(k+2)\,\gamma^2 - (3\,k+2)\,(2-2\,k+3\,k^2)\,\gamma - (k+2)\,(3\,k+2)$$

and clearly the sign of $A^2 - C^2$ is determined by that of $g(\gamma)$. The function $g(\gamma)$ is quadratic in $\gamma$ and, since $3\,k\,(k+1)\,(k+2) > 0$, represents a parabola which opens upwards. Furthermore

$$g\left(\frac{1}{3\,(k+2)}\right) = -\frac{20+61\,k+41\,k^2+18\,k^3}{3\,(k+2)} < 0$$

and

$$g\left(\frac{k-1}{k+2}\right) = -\frac{20+61\,k+45\,k^2+k^3+8\,k^4}{3\,(k+2)} < 0.$$

Thus $g(\gamma)$ is necessarily less than zero in the range $\dfrac{1}{3\,(k+2)} \leq \gamma \leq \dfrac{k-1}{k+2}$ and this in turn implies that $A^2 - C^2 < 0$ and hence $A - C < 0$ and $d_1$ itself are negative over that range.

Now compare the stationary point $d_2$ with the maximum possible value of $d$, i.e. with $k+1$. Then

$$d_2 - (k+1) = \frac{D - (k+2)\,(1+\gamma)^2}{3\,\gamma\,(1+\gamma)\,(2+k)}.$$

Since $D > 0$ and $(k+2)\,(1+\gamma)^2 > 0$

$$D^2 - (k+2)^2\,(1+\gamma)^4 = 3\,\gamma\,(1+\gamma)\,(k+2)\{k\,(k+2) + (1+2\,k)\,(6\,k^2+k+2)\,\gamma$$
$$+(k+1)\,(k+2)\,\gamma^2\} > 0$$

and this implies that $D > (k+2)\,(1+\gamma)^2 > 0$ which in turn implies that $d_2 > k+1$. Therefore the solutions $d_1$ and $d_2$ do not fall in the interval $[3, k+1]$. Since $H$ has a maximum at $d_1$ and a minimum at $d_2$, $H$ is decreasing on the interval $[d_1, d_2]$ and hence on the interval $[3, k+1]$. Also at $d = 3$, $H$ is given by

$$H_3(\gamma) = 6\{2\,(k+2)\,(k^2-k+1)\,\gamma^2 - (4\,k^3+2\,k^2+3\,k-6)\,\gamma - \,(k-1)\,(k+2)\}$$

which is quadratic in $\gamma$. Since $12\,(k+2)\,(k^2-k+1) > 0$ for $k \geq 1$, $H_3(\gamma)$ represents a parabola which opens upwards. Further $H_3(0) = -6\,(k-1)\,(k+2) < 0$ and

$$H_3\left(\frac{k-1}{k+2}\right) = -\frac{6\,k\,(k-1)\,(k+3)\,(2\,k+1)}{k+2} < 0.$$

This implies that $H_3(\gamma) < 0$ for $0 \leq \gamma \leq \dfrac{k-1}{k+2}$.

From the discussions presented in (1), (2) and (3), it follows that the design $\tilde{\xi}^*_{V_1}$ is $V$-optimal over the set of population designs defined on $\tilde{S}_{d,k}$ for all $d$ whenever $0 \leq \gamma \leq \dfrac{k-1}{k+2}$.

**Step II**

Consider now the design

$$\tilde{\xi}^*_{V_c} = \left\{ \begin{array}{ccc} (-\frac{k}{2}) & (\frac{k}{2}) & (-\frac{k}{2}, \frac{k}{2}) \\ w & w & 1-2w \end{array} \right\}$$

for $\gamma > \dfrac{k-1}{k+2}$. It follows from $\mathbf{M}_\beta(\xi) = \sum_{i=1}^r w_i \mathbf{M}_\beta(\mathbf{t}_i)$ that for the design $\tilde{\xi}^*_{V_c}$ with weight $w$ on each of the design points $-\dfrac{k}{2}$ and $\dfrac{k}{2}$ and weight $1-2w$ on the design point $(-\dfrac{k}{2}, \dfrac{k}{2})$, the information matrix for $\boldsymbol{\beta}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}^*_{V_c}) = \frac{1}{4(2\gamma+1)(\gamma+1)} \begin{pmatrix} 4\{1+\gamma(1+2\,w)\} & 0 \\ 0 & k^2\{1+\gamma(1-2\,w)\}(2\gamma+1) \end{pmatrix}$$

and hence that

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c}) = (1+\gamma) \begin{pmatrix} \dfrac{1+2\,\gamma}{1+\gamma(1+2\,w)} & 0 \\ 0 & \dfrac{4}{k^2\{1+\gamma(1-2\,w)\}} \end{pmatrix}.$$

The weight $w$ must be chosen to minimize the criterion function

$$\begin{aligned} \Psi_V(\tilde{\xi}^*_{V_c}) &= tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\} \\ &= \frac{2\,(1+k)\,(1+\gamma)\,\{1+2\,k+(1+5\,k+2\,w-2\,k\,w)\,\gamma-3\,k\,(2\,w-1)\,\gamma^2\}}{3\,k\,\{1+(1-2\,w)\,\gamma\}\,(1+\gamma+2\,w\,\gamma)}, \end{aligned} \tag{5.13}$$

where

$$\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g = \begin{pmatrix} k+1 & 0 \\ 0 & \dfrac{k(k+2)(k+1)}{12} \end{pmatrix}.$$

Taking the first derivative of the function (5.13) with respect to $w$ and equating it to zero gives the following two solutions for $w$

$$w_1 = \frac{(1+\gamma)\{k(2+3\,\gamma)+1-\sqrt{3k(2+k)(1+2\,\gamma)}\}}{2\gamma(3\,k\,\gamma+k-1)}$$

and

$$w_2 = \frac{(1+\gamma)\,\{k(2+3\,\gamma)+1+\sqrt{3k(2+k)(1+2\,\gamma)}\}}{2\,\gamma(3\,k\,\gamma+k-1)}.$$

Note that

$$w_2 - \frac{1}{2} = \frac{(1+2\,\gamma)\,(1+2\,k)+(1+\gamma)\,\sqrt{3k(2+k)(1+2\,\gamma)}}{2\gamma(3\,k\,\gamma+k-1)} > 0$$

for $k \geq 1$. Thus $w_2$ is not an acceptable weight for $\tilde{\xi}^*_{V_c}$.

First it is necessary to show that $w_1$ is a meaningful weight and in particular that $0 < w_1 < \frac{1}{2}$ for $\gamma > \dfrac{k-1}{k+2}$. For any $\gamma > 0$ and $k \geq 1$, $w_1 > 0$ if and only if

$$k(2+3\,\gamma)+1-\sqrt{3k(2+k)(1+2\,\gamma)} > 0.$$

This inequality is always true since

$$\{k(2+3\,\gamma)+1\}^2 - 3k(2+k)(1+2\,\gamma) = (3\,k\,\gamma+k-1)^2 \geq 0$$

for any $\gamma > 0$ and $k \geq 1$, and hence for $\gamma > \dfrac{k-1}{k+2}$. Further $w_1 < \frac{1}{2}$ if and only if

$$w_1 - \frac{1}{2} = \frac{(1+2\,\gamma)\,(2\,k+1)-(1+\gamma)\,\sqrt{3k(2+k)(1+2\,\gamma)}}{2\gamma(3\,k\,\gamma+k-1)} < 0. \qquad (5.14)$$

Clearly since $3\,k\,\gamma+k-1 > 0$ for $k \geq 1$ and $\gamma > 0$, the inequality in expression (5.14) is less than zero if and only if

$$f(k,\gamma) = (1+2\,\gamma)\,(2\,k+1)-(1+\gamma)\,\sqrt{3k(2+k)(1+2\,\gamma)} < 0.$$

Now since

$$(1+2\,\gamma)^2\,(2\,k+1)^2 - 3\,k\,(1+\gamma)^2\,(2+k)(1+2\,\gamma) =$$

$$-(1+2\,\gamma)\,\{\gamma\,(k+2)-(k-1)\}\,(3\,k\,\gamma+k-1) < 0$$

for $\gamma > \dfrac{k-1}{k+2}$, it follows that $f(k,\gamma) < 0$ for $\gamma > \dfrac{k-1}{k+2}$. Thus $w_1 < \frac{1}{2}$ for $k \geq 1$ and $\gamma > \dfrac{k-1}{k+2}$.

The second derivative of $\Psi_V(\tilde{\xi}^*_{V_c})$ with respect to $w$ has the form

$$\left.\frac{\partial^2 \Psi_V(\tilde{\xi}^*_{V_c})}{\partial w^2}\right|_{w_1} = \frac{\left\{\begin{array}{c} 32(k+1)\,(k+2)\,\gamma^2(1+2\,\gamma)(3\,k\,\gamma+k-1)^4\{k(2+3\,\gamma)+1 \\ -\sqrt{3k(k+2)(1+2\gamma)}\} \end{array}\right\}}{\left\{\begin{array}{c} k^{\frac{3}{2}}(1+\gamma)^2\{\sqrt{3k(k+2)(1+2\gamma)}-k-2\}^3\,\{3\,\sqrt{k}(1+2\,\gamma) \\ -\sqrt{3(2+k)(1+2\gamma)}\}^3 \end{array}\right\}}$$

at $w_1$. Observe that

i. $k(2+3\,\gamma)+1-\sqrt{3k(k+2)(1+2\gamma)} > 0$ since

$$\{k(2+3\,\gamma)+1\}^2 - 3k(k+2)(1+2\gamma) = (k-1+3\,k\gamma)^2 > 0,$$

ii. $3\,\sqrt{k}(1+2\,\gamma) - \sqrt{3(2+k)(1+2\gamma)} > 0$ since

$$9\,k\,(1+2\,\gamma)^2 - 3\,(k+2)(1+2\gamma) = 6\,(1+2\,\gamma)\,(k-1+3\,k\gamma) > 0$$

and

iii. $\sqrt{3\,k\,(2+k)(1+2\gamma)} - (k+2) > 0$ since

$$3\,k\,(2+k)(1+2\gamma) - (k+2)^2 = 2\,(k+2)\,(k-1+3\,k\gamma) > 0.$$

Therefore it follows from (i), (ii) and (iii) that $\left.\dfrac{\partial^2 \Psi_v(\tilde{\xi}^*_{V_c})}{\partial w^2}\right|_{w_1} > 0$ and hence that $w = w_1$ minimizes $\Psi_V(\tilde{\xi}^*_{V_c})$.

Substituting the weight $w_1$ for $w$ in $\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c})$ yields

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c}) = \begin{pmatrix} \dfrac{(1+2\gamma)(3k\gamma+k-1)}{3k(1+2\gamma)-A} & 0 \\ 0 & \dfrac{4(3k\gamma+k-1)}{k^2(A-k-2)} \end{pmatrix}$$

where $A = \sqrt{3k(k+2)(1+2\gamma)}$. Then the directional derivative of $\Psi_V(\tilde{\xi}^*_{V_c})$ at the design $\tilde{\mathbf{t}}$ is given by

$$\begin{aligned}\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c}) &= tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c})\,\tilde{\mathbf{X}}_g'\,\tilde{\mathbf{X}}_g\,\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c})\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\} - \Psi_V(\tilde{\xi}^*_{V_c}) = \\ &- \frac{(1+2\gamma)(1+k)\{1+2\gamma(d-1)\}(3k\gamma+k-1)^2}{(1+d\gamma)\{A-3k(1+2\gamma)\}^2} \\ &+ \frac{4(k+1)(k+2)(3k\gamma+k-1)^2}{3k^3 d(k+2-A)^2}\tilde{\mathbf{t}}(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}},\end{aligned}$$

where

$$\mathbf{M}_\beta(\tilde{\mathbf{t}}) = \begin{pmatrix} \dfrac{1}{1+d\gamma} & \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d(1+d\gamma)} \\ \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d(1+d\gamma)} & \dfrac{1}{d}\tilde{\mathbf{t}}'(\mathbf{I} - \dfrac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}} \end{pmatrix}$$

and

$$\Psi_V(\tilde{\xi}^*_{V_c}) = -\frac{2(1+k)\sqrt{k(k+2)(1+2\gamma)}(3k\gamma+k-1)^2}{\sqrt{3}k(-2-k+A)\{A-3k(1+2k\gamma)\}}.$$

Observe that the directional derivative $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c})$ is a convex function over the polytope $\tilde{P}_{d,k}$. Thus it is only necessary to examine $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c})$ at the extreme vertices of $\tilde{P}_{d,k}$ for all possible values of $d$.

(1) For the case $d=1$, $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c}) = 0$ at $\frac{k}{2}$ and $-\frac{k}{2}$ and similarly for $d=2$, $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c}) = 0$ at $(-\frac{k}{2}, \frac{k}{2})$. In fact these results are to be expected since these designs are the support designs of the population design $\tilde{\xi}^*_{V_c}$.

(2) Consider now all $d$-point designs $\tilde{\mathbf{t}}$ with $d \geq 3$. The values of $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c})$ associated with contours

$$\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\,\gamma}\mathbf{J})\tilde{\mathbf{t}} = c$$

where $c$ is a constant, are of exactly the same form as those described in Theorem 5.2.2 and Theorem 5.2.3. So it follows immediately that the largest value of $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c})$ occurs at the vertex $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ when $d$ is even and at the vertices $\tilde{\mathbf{v}}_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}_{\frac{d+3}{2}}$ when $d$ is odd. Thus it is only necessary to check the condition $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c}) \leq 0$ at these vertices.

Consider first the case for which $d$ is an even integer. Then at the vertex $\tilde{\mathbf{v}}_{\frac{d}{2}+1}$ the directional derivative is given by

$$\phi_V(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_c}) = \frac{2\,(d-2)\,(k+1)\,(k+2)\,(1+2\,\gamma)\,(3\,k\,\gamma+k-1)^2\,(1+2\,k+3\,k\,\gamma-A)}{3\,(1+d\,\gamma)\,k^2\,\{A-(k+2)\}^2\,\{A-3\,k\,(1+2\,\gamma)\}^2}\,f(d)$$

where $A = \sqrt{3\,k\,(k+2)\,(1+2\,\gamma)}$ and $f(d) = d^2\,\gamma + d\,(1-\gamma-3\,k\,\gamma) - 3\,k^2\,\gamma - 3\,k - 1$, a quadratic in $d$. It has been shown previously in Step I that $(1+2\,k+3\,k\,\gamma-A) > 0$. Note that the expression $(d-2)\,(k+1)\,(k+2)\,(1+2\,\gamma)\,(3\,k\,\gamma+k-1)^2$ in the numerator of $\phi_V(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_c})$ is greater than or equal to zero for $d \geq 2$. Note also that the denominator is positive. Therefore the sign of $\phi_V(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_c})$ depends on the sign of the function $f(d)$. So it is necessary to examine the properties of $f(d)$. The function $f(d)$ is a quadratic in $d$, since $\gamma > 0$, and represents a parabola opening upwards. Furthermore $f(1) = -3\,k\,(1+\gamma+k\,\gamma) < 0$ and $f(k+1) = -k\,(2+(2+5\,k)\,\gamma) < 0$. This implies that $f(d) < 0$ for $1 \leq d \leq k+1$. Thus $\phi_V(\tilde{\mathbf{v}}_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_c}) < 0$ for all even $d$ in the interval $[1, k+1]$.

When $d$ is an odd integer and $d \geq 3$ the directional derivatives at the vertices $\tilde{\mathbf{v}}_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}_{\frac{d+3}{2}}$ are given by

$$\phi_V(\tilde{\mathbf{v}}_{\frac{d+1}{2}}, \tilde{\xi}^*_{V_c}) = \phi_V(\tilde{\mathbf{v}}_{\frac{d+3}{2}}, \tilde{\xi}^*_{V_c})$$

$$\frac{2\,(d-1)\,(k+1)\,(k+2)\,(1+2\,\gamma)\,(3\,k\,\gamma+k-1)^2\,(1+2\,k+3\,k\,\gamma-A)}{3\,d\,(1+d\,\gamma)\,k^2\,\{A-(k+2)\}^2\,\{A-3\,(1+2\,\gamma)\}^2}\,h(d)$$

where $A=\sqrt{3\,k\,(k+2)\,(1+2\,\gamma)}$ and

$$h(d)=d^3\,\gamma+d^2\,(1-2\,\gamma-3\,\gamma\,k)+d\,\{-2-3\,k(1-\gamma+k\,\gamma)\}+3\,(k+1)\,(1+\gamma+k\,\gamma).$$

It has been proved earlier that $(1+2\,k+3\,k\,\gamma-A)>0$. Furthermore the expressions $(d-1)\,(k+1)\,(k+2)\,(1+2\,\gamma)\,(3\,k\,\gamma+k-1)^2$ and $3\,d\,(1+d\,\gamma)\,k^2\,\{A-(k+2)\}^2\,\{A-3\,(1+2\,\gamma)\}^2$ are positive for $d\geq 3$ and $\gamma>0$. Therefore the sign of $\phi_V(\tilde{\mathbf{v}}_{\frac{d+1}{2}},\tilde{\xi}^*_{V_c})$ takes the sign of $h(d)$ and it is only necessary to examine $h(d)$.

The derivative of $h(d)$ with respect to $d$ is less than zero, that is

$$\frac{\partial h(d)}{\partial d}=-\gamma\,\{3\,d\,(2\,k-d)+4\,d+3\,k\,(k-1)\}+2\,(d-1)-3\,k<0$$

for $3\leq d\leq k+1$ and hence $h(d)$ is decreasing on the interval $[3,k+1]$. Furthermore $h(3)=-6\,\{\gamma\,(k^2+2\,k-2)+k-1\}<0$ and $h(k+1)=-(k-1)\,(k+1)\{\gamma\,(5\,k+2)+2\}<0$. Thus $h(d)<0$ for $d\in[3,k+1]$ and this in turn implies that $\phi_V(\tilde{\mathbf{v}}_{\frac{d+1}{2}},\tilde{\xi}^*_{V_c})\leq 0$ for $3\leq d\leq k+1$.

Thus from the discussions (1) and (2), it follows that the design $\tilde{\xi}^*_{V_c}$ is $V$-optimal over the set of population designs defined on $\tilde{S}_{d,k}$ for all $d$ values whenever $\gamma>\dfrac{k-1}{k+2}$.
□

The design weights in $\xi^*_{V_c}$ refer to the proportion of *observations* that should be taken at the individual designs. Thus, for a total of $N$ observations, the individual allocation is $N\,w$ observations to each of the time points 0 and $k$, with the remaining $\frac{N}{2}\,(1-2\,w)$ observations allocated to $(0,k)$. On normalizing the allocation of observations on a per subject basis it follows that a proportion of individuals $\frac{2\,w}{1+2\,w}$ is allocated to each of the single-point designs

0 and $k$ and the residual proportion $\frac{1-2w}{1+2w}$ to the design $(0, k)$. For example for $w = \frac{1}{4}$ and a total of $N = 200$ observations, there are 50 individuals allocated to each of the designs 0 and $k$ and another 50 to the design $(0, k)$.

The dependence of the optimal weight $w$ on $\gamma$ is also of interest. Consider the case of $k = 10$. Then from Theorem 5.4.1, the optimum designs are

$$\xi^*_{V_1} = \begin{Bmatrix} (0) & (10) \\ 0.5 & 0.5 \end{Bmatrix},$$

for $\gamma \leq \frac{3}{4}$, and

$$\xi^*_{V_c} = \begin{Bmatrix} (0) & (10) & (0, 10) \\ w & w & 1 - 2w \end{Bmatrix}$$

where

$$w = \frac{(1+\gamma)\{7 + 10\gamma - 2\sqrt{10(1 + +2\gamma)}\}}{2\gamma(3 + 10\gamma)}$$

for $\gamma > \frac{3}{4}$. The weight $w$ is plotted against $\gamma$ in Figure 5.2. Note that for $0 \leq \gamma \leq \frac{3}{4}$, $w = \frac{1}{2}$. For $\gamma > \frac{3}{4}$ the weight decreases to minimum of 0.4073 with increasing $\gamma$ and then approaches $\frac{1}{2}$ asymptotically from below as $\gamma$ tends to infinity.

Abt *et al.* (1997) considered optimal designs for prediction under the linear growth model with a random intercept. Their results indicate that the design

$$\tilde{\xi}^*_{V_1} = \begin{Bmatrix} (-\frac{k}{2}) & (\frac{k}{2}) \\ \frac{1}{2} & \frac{1}{2} \end{Bmatrix}$$

is optimal for growth prediction unless the intraclass correlation $\rho = \dfrac{\gamma}{1+\gamma}$ is very high. For instance, when $k = 10$ the authors claim that the design $\tilde{\xi}^*_{V_1}$ is optimal for $\rho \leq 0.88$.

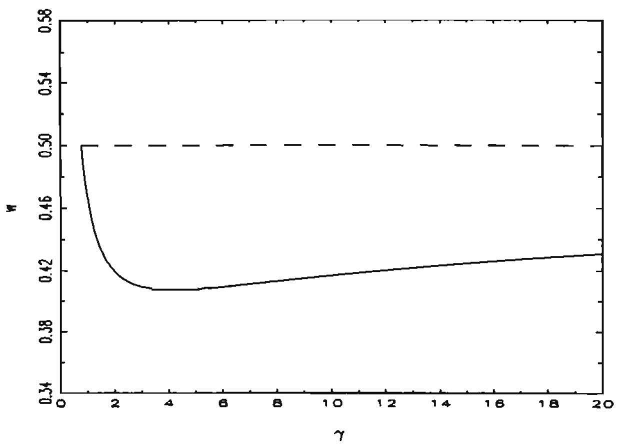


Figure 5.2: Plot of the optimal design weight $w$ against $\gamma$ for $k = 10$.

However, according to Theorem 5.4.1, whenever $\gamma > \gamma(10) = \frac{3}{4}$, and equivalently when $\rho > \frac{3}{7}$, the design

$$\tilde{\xi}^*_{V_c} = \left\{ \begin{array}{ccc} (-5) & (5) & (-5,5) \\ w & w & 1-2w \end{array} \right\},$$

with

$$w = \frac{(1+\gamma)\{7+10\gamma+1-2\sqrt{10(1+2\gamma)}\}}{2\gamma(3+10\gamma)}$$

is optimal over all possible population designs defined on the set $\{-5,\dots,-1,0,1,\dots,5\}$. Therefore, the use of the design $\tilde{\xi}^*_{V_1}$ for growth prediction when $\frac{3}{7} < \rho < 0.88$ will result in loss of efficiency. Similar situations occur for other values of $k$.

## 5.5 $V$-optimal population designs for the mean responses based on designs with repeated time points

### 5.5.1 Designs based on $d$-point individual designs

In the previous sections the construction of $V$-optimal population designs based on designs with non-repeated time points was considered. In this section, the problem of constructing these optimal designs based on individual designs for which replications of the time points are permitted, is considered. It has been shown in Section 5.2 that, for the simple linear regression model with a random intercept, the $D$-optimal design relating to the fixed effects $\boldsymbol{\beta}$ and based on $d$-point individual designs with non-repeated time points is also the $V$-optimal design for the mean responses on the set $\{0, 1, \ldots, k\}$. An equivalent result is also found for designs with repeated points and this is presented in the following theorem.

**Theorem 5.5.1** *Consider the set of all $d$-point individual designs $\mathbf{t}$ with $t_j \in \{0, 1, \ldots, k\}$, $j = 1, \ldots, d$, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_d \leq k$ and $d$ an integer in the interval $[2, k+1]$. Then the $V$-optimal population designs for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ in model (5.1) over this set for all $\gamma \geq 0$ are given by*

$$\xi^*_{V_{re}} = \left\{ \begin{array}{c} (\underbrace{0, \ldots, 0}_{\frac{d}{2}\ times}, \underbrace{k, \ldots, k}_{\frac{d}{2}\ times}) \\ 1 \end{array} \right\}$$

*for $d$ even and*

$$\xi^*_{V_{ro}} = \left\{ \begin{array}{cc} (\underbrace{0, \ldots, 0}_{\frac{d+1}{2}\ times}, \underbrace{k, \ldots, k}_{\frac{d-1}{2}\ times}), & (\underbrace{0, \ldots, 0}_{\frac{d-1}{2}\ times}, \underbrace{k, \ldots, k}_{\frac{d+1}{2}\ times}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

*for d odd.*

**Proof**

Recall from Subsection 2.6.5 that $V$-optimal population designs for random intercept models are invariant to linear transformations in the explanatory variables. Thus consider transforming the time points as $\tilde{t}_j = t_j - \frac{k}{2}, j = 1, 2, \ldots, d$, so that a $d$-point individual design is transformed as $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$ where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$ is the center of the hypercube $C_{d,k}$. Thus the space of designs in the transformed coordinates is given by

$$\tilde{T}_{d,k} = \{\tilde{\mathbf{t}} : \tilde{\mathbf{t}} = (\tilde{t}_1, \ldots, \tilde{t}_d),\ \tilde{t}_j \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}\},\ j = 1, \ldots, d,$$
$$-\frac{k}{2} \leq \tilde{t}_1 \leq \tilde{t}_2 \leq \ldots \leq \tilde{t}_d \leq \frac{k}{2}\}.$$

Then the proposed optimal designs $\xi^*_{V_{re}}$ and $\xi^*_{V_{ro}}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{V_{re}} = \left\{ \begin{array}{c} (\underbrace{-\frac{k}{2}, \ldots, -\frac{k}{2}}_{\frac{d}{2}\text{ times}}, \underbrace{\frac{k}{2}, \ldots, \frac{k}{2}}_{\frac{d}{2}\text{ times}}) \\ 1 \end{array} \right\}$$

and

$$\tilde{\xi}^*_{V_{ro}} = \left\{ \begin{array}{cc} (\underbrace{-\frac{k}{2}, \ldots, -\frac{k}{2}}_{\frac{d+1}{2}\text{ times}}, \underbrace{\frac{k}{2}, \ldots, \frac{k}{2}}_{\frac{d-1}{2}\text{ times}}), & (\underbrace{-\frac{k}{2}, \ldots, -\frac{k}{2}}_{\frac{d-1}{2}\text{ times}}, \underbrace{\frac{k}{2}, \ldots, \frac{k}{2}}_{\frac{d+1}{2}\text{ times}}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

respectively. Recall from the proof of Theorem 4.4.2 that the inverses of the standardized information matrices for $\boldsymbol{\beta}$ at the designs $\tilde{\xi}^*_{V_{re}}$ and $\tilde{\xi}^*_{V_{ro}}$ are, respectively,

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}^*_{re}) = \begin{pmatrix} 1 + d\gamma & 0 \\ 0 & \frac{4}{k^2} \end{pmatrix}$$

and

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{ro}) = \begin{pmatrix} 1+d\gamma & 0 \\ 0 & \dfrac{4\,d\,(1+d\,\gamma)}{k^2\,(d-\gamma+d^2\,\gamma)} \end{pmatrix}.$$

Consider first the case of $d$ even. Then the directional derivative of the criterion $\Psi_V(\tilde{\xi}) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\}$ at $\tilde{\xi}^*_{V_{re}}$ in the direction of a $d$-point design $\tilde{\mathbf{t}} \in \tilde{T}_{d,k}$ is given by

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_{re}}) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{re}})\,\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\,\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{re}})\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\} - \Psi_V(\tilde{\xi}^*_{V_{re}})$$

where

$$\mathbf{M}_\beta(\tilde{\mathbf{t}}) = \begin{pmatrix} \dfrac{1}{1+d\gamma} & \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d\,(1+d\gamma)} \\ \dfrac{\mathbf{1}'\tilde{\mathbf{t}}}{d\,(1+d\gamma)} & \dfrac{1}{d}\tilde{\mathbf{t}}'\,(\mathbf{I} - \dfrac{\gamma}{1+d\,\gamma}\,\mathbf{J})\,\tilde{\mathbf{t}} \end{pmatrix},$$

$$\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g = \begin{pmatrix} k+1 & 0 \\ 0 & \dfrac{k\,(k+1)(k+2)}{12} \end{pmatrix},$$

and

$$\Psi_V(\tilde{\xi}^*_{re}) = \frac{(1+k)\,(2+4\,k+3\,d\,k\,\gamma)}{3\,k}.$$

This derivative can be expressed explicitly as

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_{re}}) = \frac{4\,(k+1)\,(k+2)}{3\,k^3\,d}\,\tilde{\mathbf{t}}'\,(\mathbf{I} - \frac{\gamma}{1+d\,\gamma}\,\mathbf{J})\,\tilde{\mathbf{t}} - \frac{(k+1)\,(k+2)}{3\,k}$$

and is clearly a convex function on the polytope $\tilde{Q}_{d,k}$. Thus it is only necessary to check the condition $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_{re}}) \leq 0$ at the vertices of the polytope $\tilde{Q}_{d,k}$. Recall from Subsection 4.2.2 that these vertices are the designs $\mathbf{v}_1^*, \mathbf{v}_2^*, \ldots, \mathbf{v}_{d+1}^*$, written in the transformed coordinates as

$$\tilde{\mathbf{v}}^*_{j+1} = (\underbrace{-\frac{k}{2}, \ldots, -\frac{k}{2}}_{j\ \text{times}}, \underbrace{\frac{k}{2}, \ldots, \frac{k}{2}}_{(d-j)\ \text{times}})$$

for $j = 0, 1, \ldots, d$, and that the directional derivative is given by

$$\phi_V(\tilde{\mathbf{v}}^*_{j+1}, \tilde{\xi}^*_{V_{re}}) = -\frac{(d-2j)^2\,(k+1)\,(k+2)\,\gamma}{3\,d\,k\,(1+d\,\gamma)}$$

for $j = 0, 1, \ldots, d$. Clearly $\phi_V(\tilde{\mathbf{v}}^*_{j+1}, \tilde{\xi}^*_{V_{re}}) \leq 0$ for $j = 0, 1, \ldots, d$, $d \geq 2$ and $\gamma \geq 0$. Furthermore, $\phi_V(\tilde{\mathbf{v}}^*_{j+1}, \tilde{\xi}^*_{V_{re}}) = 0$ at $j = \dfrac{d}{2}$ and thus at the support design of $\tilde{\xi}^*_{V_{re}}$. Thus, from the appropriate Equivalence Theorem, it now follows immediately that the design $\tilde{\xi}^*_{V_{re}}$ and hence $\xi^*_{V_{re}}$ is $V$-optimal for $\gamma \geq 0$.

Consider now the case of $d$ odd. The directional derivative of $\Psi_V(\tilde{\xi}_{V_{ro}}) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\}$ at $\tilde{\xi}^*_{V_{ro}}$ in the direction of a $d$-point design $\tilde{\mathbf{t}}$ is given by

$$\begin{aligned}\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_{ro}}) &= tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{ro}})\,\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\,\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{ro}})\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\} - \Psi_V(\tilde{\xi}^*_{V_{ro}})\\ &= \frac{4\,d\,(d\,\gamma+1)^2\,(k+1)\,(k+2)}{3\,k^3\,\{\gamma\,(d^2-1)+d\}^2}\,\tilde{\mathbf{t}}'\,(\mathbf{I} - \frac{\gamma}{1+d\,\gamma}\,\mathbf{J})\,\tilde{\mathbf{t}} - \frac{d\,(k+1)\,(k+2)\,(d\,\gamma+1)}{3\,k\,\{\gamma\,(d^2-1)+d\}}\end{aligned}$$

where

$$\Psi_V(\tilde{\xi}^*_{V_{ro}}) = \frac{(1+k)\,(1+d\,\gamma)\,(2\,d+4\,d\,k-3\,k\,\gamma+3\,d^2\,k\,\gamma)}{3\,k\,\{\gamma\,(d^2-1)+d\}}$$

and is a convex function on the polytope $\tilde{Q}_{d,k}$. So here it is again only necessary to examine $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_{ro}})$ at the vertices of $\tilde{Q}_{d,k}$. At the general vertex $\tilde{\mathbf{v}}^*_{j+1} = (-\frac{k}{2}, -\frac{k}{2}+1, \ldots, -\frac{k}{2}+d-j-1, \frac{k}{2}-j+1, \ldots, \frac{k}{2})$ the directional derivative has the form

$$\phi_V(\tilde{\mathbf{v}}^*_{j+1}, \tilde{\xi}^*_{V_{ro}}) = -\frac{d\,\{(d-2\,j)^2-1\}\,(k+1)\,(k+2)\,(1+d\,\gamma)\,\gamma}{3\,k\,\{\gamma\,(d^2-1)+d\}^2}$$

which is less than or equal to zero for $j = 0, 1, \ldots, d$. For $d \geq 1$ and $\gamma \geq 0$ it is clear that $3\,k\,\{\gamma\,(d^2-1)+d\}^2 > 0$, that $(2+3\,k+k^2)\,(1+d\,\gamma)\,\gamma \geq 0$ and also that for $j = 0, 1, \ldots, d$ $(d-2\,j)^2-1 \geq 0$. Moreover $\phi_V(\tilde{\mathbf{v}}^*_{j+1}, \tilde{\xi}^*_{V_{ro}}) = 0$ at $j = \dfrac{d-1}{2}$ and $j = \dfrac{d+1}{2}$ and thus at the support designs of $\tilde{\xi}^*_{V_{ro}}$. Thus it follows that $\phi_V(\tilde{\mathbf{v}}^*_{j+1}, \tilde{\xi}^*_{V_{ro}}) \leq 0$ for $j = 0, 1, \ldots, d$ and for $\gamma \geq 0$ and hence that the design $\tilde{\xi}^*_{V_{ro}}$ is $V$-optimal. $\square$

Note that the $D$-optimal population designs relating to the fixed effects $\boldsymbol{\beta}$ in model (5.1) and based on $d$-point individual designs with repeated time points are also $V$-optimal for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$.

### 5.5.2 Comparison of $V$-optimal population designs

$V$-optimal population designs based on $d$-point individual designs with repeated time points for $d \geq 3$ are now compared in terms of their $V$-efficiencies and thus in effect in terms of their criteria values. The general result is presented in the following theorem.

**Theorem 5.5.2** *The efficiency, on a per observation basis, of $V$-optimal population design based on $d$-point individual designs with repeated time points decreases as $d$ increases.*

**Proof**

Let $d_e$ be a positive even integer. To prove the theorem it is only necessary to show that the inequalities

$$\Psi_V(\tilde{\xi}^*_{V_{r(d_e)}}) < \Psi_V(\tilde{\xi}^*_{V_{r(d_e+1)}}) < \Psi_V(\tilde{\xi}^*_{V_{r(d_e+2)}}),$$

hold for all $d_e$ where $\tilde{\xi}^*_{V_{r(d_e)}}$, $\tilde{\xi}^*_{V_{r(d_e+1)}}$ and $\tilde{\xi}^*_{V_{r(d_e+2)}}$ are the $V$-optimal designs in the transformed coordinates based on individual designs with $d_e$, $d_e + 1$ and $d_e + 2$ points respectively. Equivalently it is only necessary to show that the differences

$$D_1 = \Psi_V(\tilde{\xi}^*_{V_{r(d_e+1)}}) - \Psi_V(\tilde{\xi}^*_{V_{r(d_e)}})$$

and

$$D_2 = \Psi_V(\tilde{\xi}^*_{V_{r(d_e+2)}}) - \Psi_V(\tilde{\xi}^*_{V_{r(d_e+1)}})$$

are greater than or equal to zero for $1 \leq d_e \leq k$.

Substituting $d_e$ and $d_e + 1$ for $d$ in the expressions for $\Psi_v(\tilde{\xi}^*_{V_{re}})$ and $\Psi_v(\tilde{\xi}^*_{V_{ro}})$ respectively yields

$$\Psi_V(\tilde{\xi}^*_{V_{r(d_e)}}) = \frac{(1+k)\,(2+4\,k+3\,k\,d_e\,\gamma)}{3\,k}$$

and

$$\Psi_V(\tilde{\xi}^*_{V_{ro(d_e+1)}}) = \frac{(1+k)\,\{1+\gamma(d_e+1)\}\,\{2+2\,d_e+4\,k\,(1+d_e)+3\,k\,d_e\,\gamma\,(2+d_e)\}}{3\,k\,\{1+d_e\,(1+2\,\gamma+d_e\,\gamma)\}}$$

respectively. Then the difference

$$D_1 = \frac{(1+k)\,\gamma\,\{2+4\,k+3\,k\,d_e\,(1+2\,\gamma+d_e\,\gamma)\}}{3\,k\,\{1+d_e+d_e\,\gamma\,(2+d_e)\}}$$

is clearly greater than zero for $1 \leq d_e \leq k$, $k \geq 2$ and $\gamma > 0$.

Similarly substituting $d_e + 2$ for $d$ in the expression for $\Psi_V(\tilde{\xi}^*_{re})$ gives

$$\Psi_V(\tilde{\xi}^*_{V_{r(d_e+2)}}) = \frac{(1+k)\,\{2+4\,k+3\,k\,\gamma\,(2+d_e)\}}{3\,k}$$

and thus the difference

$$D_2 = \frac{(1+k)\,\gamma\,\{-2+2\,k+3\,k\,d_e+3\,k\,d_e\,\gamma\,(d_e+2)\}}{3\,k\,\{1+d_e+d_e\,\gamma\,(2+d_e)\}}$$

is greater than zero for $1 \leq d_e \leq k$, $k \geq 2$ and $\gamma > 0$. □

The result in Theorem 5.5.2 mirrors that found for the $V$-optimal population designs based on individual designs with non-repeated time points.

### 5.5.3 Best $V$-optimal population designs

It has been shown in Subsection 4.5.3 that the best $D$-optimal design over the set of population designs with non-repeated points is also best over the set of population designs with repeated points. Note that the one- and two-point $V$-optimal population designs based on designs with repeated points are identical to those of designs with non-repeated points. It is therefore not unreasonable to assume that the $V$-optimal designs in Theorem 5.4.1 are $V$-optimal over all population designs based on individual designs with repeated points and this demonstrated in following theorem.

**Theorem 5.5.3** *Consider the set of population designs based on all possible individual designs* $\mathbf{t}$ *which put equal weights on the time points* $t_1, t_2, \ldots, t_d$ *with* $t_j \in \{0, 1, \ldots, k\}$ *and* $0 \le t_1 \le t_2 \le \ldots, \le t_d$ *for* $d$ *any positive integer less than or equal to* $k+1$. *Then the* $V$*-optimal population designs for the mean responses* $\boldsymbol{\mu}_g$ *at* $\mathbf{t}_g = (0, 1, \ldots, k)'$ *in model (5.1) over this set are given by*

$$\xi^*_{V_1} = \left\{ \begin{array}{cc} (0) & (k) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

*for* $0 \le \gamma \le \dfrac{k-1}{k+2}$ *and by*

$$\xi^*_{V_c} = \left\{ \begin{array}{ccc} (0) & (k) & (0,k) \\ w & w & 1-2w \end{array} \right\},$$

*where*

$$w = \frac{(1+\gamma)\,(k\,(2+3\,\gamma)+1-\sqrt{3k(2+k)(1+2\,\gamma)})}{2\,\gamma\,(3\,k\,\gamma+k-1)}$$

*for* $\gamma > \dfrac{k-1}{k+2}$.

**Proof**

Recall from the proof of Theorem 5.4.1 that the designs $\xi^*_{V_1}$ and $\xi^*_{V_c}$ can be written in the linearly transformed coordinates as

$$\tilde{\xi}^*_{V_1} = \left\{ \begin{array}{cc} (-\frac{k}{2}) & (\frac{k}{2}) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$$

and

$$\tilde{\xi}^*_{V_c} = \left\{ \begin{array}{ccc} -\frac{k}{2} & \frac{k}{2} & (-\frac{k}{2}, \frac{k}{2}) \\ w & w & 1-2w \end{array} \right\}$$

respectively and that the associated inverses of the standardized information matrices for $\boldsymbol{\beta}$ are given by

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}^*_{V_1}) = (1+\gamma) \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{4}{k^2} \end{pmatrix}$$

and

$$\mathbf{M}_{\beta}^{-1}(\tilde{\xi}^*_{V_c}) = \begin{pmatrix} \dfrac{(1+2\gamma)(3k\gamma+k-1)}{3k(1+2\gamma)-A} & 0 \\ 0 & \dfrac{4(3k\gamma+k-1)}{k^2(A-k-2)} \end{pmatrix}$$

where $A = \sqrt{3k(k+2)(1+2\gamma)}$.

Consider first the optimality of the design $\tilde{\xi}^*_{V_1}$. The directional derivative of the criterion $\Psi_V(\tilde{\xi}) = tr\{\mathbf{M}_{\beta}^{-1}(\tilde{\xi})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\}$ at $\tilde{\xi}^*_{V_1}$ in the direction of a $d$-point design $\tilde{\mathbf{t}} \in \tilde{T}_{d,k}$ is given by

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) = \frac{(1+\gamma)(k+1)}{3k} \left\{ \frac{4(k+2)(1+\gamma)}{k^2 d} \tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\gamma}\mathbf{J})\tilde{\mathbf{t}} - \frac{2(d\gamma+1)(2k+1) - 3k(1+\gamma)}{d\gamma+1)} \right\}$$

and is a convex function on the polytope $\tilde{Q}_{d,k}$. Thus to prove that $\tilde{\xi}^*_{V_1}$ is the $V$-optimal population design it is only necessary to check that the condition $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) \leq 0$ holds at the extreme vertices of the polytope $\tilde{Q}_{d,k}$.

(1) For the cases $d = 1$ and $d = 2$, the proof of optimality for $\tilde{\xi}^*_{V_1}$ is identical to that given in Theorem 5.4.1.

(2) Now consider $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ for all $d$-point designs $\tilde{\mathbf{t}} \in \tilde{T}_{d,k}$ with $d \geq 3$. Then

$$\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) = \frac{4\,(k+1)(k+2)(1+\gamma)^2}{3\,k^3}\,\tilde{\mathbf{t}}'(\mathbf{I} - \frac{\gamma}{1+d\,\gamma}\mathbf{J})\tilde{\mathbf{t}} + \frac{(k+1)\,(1+\gamma)\,(2+k+2\,d\,\gamma-3\,k\,\gamma+4\,d\,k\,\gamma)}{3\,(k+d\,k\,\gamma)}$$

and is of exactly the same form as that described in the proof of Theorem 5.4.1. Thus the largest value of $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ over the space of individual designs with points in $\tilde{T}_{d,k}$ is attained at the vertex $\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}$ when $d$ is even and at the vertices $\tilde{\mathbf{v}}^*_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}^*_{\frac{d+3}{2}}$ when $d$ is odd. Thus, in order to show that $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1}) \leq 0$ it is only necessary to prove this inequality for the extreme vertices.

When $d$ is even and $d > 2$, the derivative at $\tilde{\xi}^*_{V_1}$ in the direction of the extreme vertex $\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}$ is given by

$$\phi_V(\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_1}) = \frac{(1+k)\,\gamma\,(1+\gamma)\,(2+4\,k-3\,d\,k+2\,d\,\gamma+d\,k\,\gamma)}{3\,k\,(1+d\,\gamma)}. \tag{5.15}$$

Clearly, since $\gamma \geq 0$, the expression (5.15) is less than or equal to zero if and only if

$$2+4\,k-3\,d\,k+2\,d\,\gamma+d\,k\,\gamma \leq 0$$

and thus if and only if

$$\gamma \leq \frac{3\,k\,d-2\,(2\,k+1)}{d\,(k+2)} = f(d).$$

Since $\dfrac{\partial f(d)}{\partial d} = \dfrac{2\,(2\,k+1)}{d^2\,(k+2)} > 0$ for all $d$, $f(d)$ is increasing on the interval $(2, k+1]$. Thus the function $f(d)$ has a minimum at $d = 2$ which is equal to $\dfrac{k-1}{k+2}$ and this minimum provides the upper bound for $\gamma$. Thus $\phi_V(\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_1}) \leq 0$ whenever $0 \leq \gamma \leq \dfrac{k-1}{k+2}$.

Consider now $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ when $d$ is odd and $d \geq 3$. At the extreme vertices $\tilde{\mathbf{v}}^*_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}^*_{\frac{d+3}{2}}$, the directional derivative $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_1})$ is given by

$$\begin{aligned} \phi_V(\tilde{\mathbf{v}}^*_{\frac{d+1}{2}}, \tilde{\xi}^*_{V_1}) &= \phi_V(\tilde{\mathbf{v}}^*_{\frac{d+3}{2}}, \tilde{\xi}^*_{V_1}) \\ &= \frac{(d-1)\,(1+k)\,\gamma\,(1+\gamma)\,\{k\,[1+d\,(\gamma-3)+\gamma]+2\,(1+\gamma+d\,\gamma)\}}{3\,d\,k\,(1+d\,\gamma)}. \end{aligned} \tag{5.16}$$

Clearly for $d \geq 3$ and $\gamma \geq 0$, expression (5.16) less than or equal to zero if and only if

$$k\,[1+d\,(\gamma-3)+\gamma]+2\,(1+\gamma+d\,\gamma) \leq 0$$

and thus if and only if

$$\gamma \leq \frac{3\,d\,k-k-2}{(d+1)\,(k+2)} = g(\gamma).$$

Since $\dfrac{\partial g(d)}{\partial d} = \dfrac{2\,(2\,k+1)}{(d+1)^2\,(k+2)} > 0$ for all $d$, $g(d)$ is increasing on the interval $[3, k+1]$ and it has a minimum at $d=3$ which is equal to $\dfrac{4\,k-1}{2\,(k+2)}$. This minimum provides the upper bound for $\gamma$ and it is greater than $\dfrac{k-1}{k+2}$ for all $k \geq 2$. Thus the expression (5.16) is less than or equal to zero for $0 \leq \gamma \leq \dfrac{k-1}{k+2}$.

Now consider the directional derivative of $\Psi_V(\tilde{\xi}) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi})\tilde{\mathbf{X}}_g'\,\tilde{\mathbf{X}}_g\}$ at the composite $V$-optimal design $\tilde{\xi}^*_{V_c}$ in the direction of a $d$-point design $\tilde{\mathbf{t}} \in \tilde{T}_{d,k}$, that is

$$\begin{aligned} \phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c}) &= tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c})\,\tilde{\mathbf{X}}_g'\,\tilde{\mathbf{X}}_g\,\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_c})\,\mathbf{M}_\beta(\tilde{\mathbf{t}})\} - \Psi_V(\tilde{\xi}^*_{V_c}) \\ &= -\frac{(1+2\,\gamma)\,(1+k)\,\{1+2\,\gamma\,(d-1)\}\,(3\,k\,\gamma+k-1)^2}{(1+d\,\gamma)\,\{A-3\,k(1+2\,\gamma)\}^2} \\ &\quad + \frac{4(k+1)(k+2)(3\,k\,\gamma+k-1)^2}{3\,k^3\,d\,(k+2-A)^2}\,\tilde{\mathbf{t}}(\mathbf{I}-\frac{\gamma}{1+d\,\gamma}\mathbf{J})\tilde{\mathbf{t}}. \end{aligned}$$

Since $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c})$ is a convex function over the polytope $\tilde{Q}_{d,k}$ the inequality $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c}) \leq 0$ need only be checked at the extreme vertices of $\tilde{Q}_{d,k}$.

For $d$ even and $d > 2$, the directional derivative $\phi_V(\tilde{\mathbf{t}}, \tilde{\xi}^*_{V_c})$ at the extreme vertex $\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}$ is given by

$$\phi_V(\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_c}) = -\frac{(d-2)\,(k+1)\,\gamma\,(3\,k\,\gamma + k - 1)^2}{6\,k\,(1+d\,\gamma)\,(3\,k\,\gamma + 2\,k + 1 - A)}$$

where $A = \sqrt{3\,k\,(2+k)\,(1+2\,\gamma)}$. It has been proved in Section 5.4 that $3\,k\,\gamma + 2\,k + 1 > A$. Thus $\phi_V(\tilde{\mathbf{v}}^*_{\frac{d}{2}+1}, \tilde{\xi}^*_{V_c}) \leq 0$ for $d \geq 2$ and $k \geq 2$.

When $d$ is odd and $d \geq 3$ the directional derivative at both vertices $\tilde{\mathbf{v}}^*_{\frac{d+1}{2}}$ and $\tilde{\mathbf{v}}^*_{\frac{d+3}{2}}$ is given by

$$\phi_V(\tilde{\mathbf{v}}^*_{\frac{d+1}{2}}, \tilde{\xi}^*_{V_c}) = \phi_V(\tilde{\mathbf{v}}^*_{\frac{d+3}{2}}, \tilde{\xi}^*_{V_c}) = -\frac{(d-1)^2\,(k+1)\,\gamma\,(3\,k\,\gamma + k - 1)^2}{6\,d\,k\,(1+d\,\gamma)\,(3\,k\,\gamma + 2\,k + 1 - A)}.$$

Clearly this is less than or equal to zero for $d \geq 3$ and $k \geq 2$. □

## 5.6 Efficiencies of $V$-optimal population designs based on individual designs with repeated points relative to those with non-repeated time points

The relative efficiency of the $V$-optimal population design based on $d$-point individual designs with repeated time points to the corresponding optimal design with non-repeated time points is defined by

$$\frac{\Psi_V(\tilde{\xi}^*_{V_{r(d)}})}{\Psi_V(\tilde{\xi}^*_{V_d})} = \frac{tr[\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_{r(d)}})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g]}{tr[\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_d})\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g]}$$

where $\tilde{\xi}^*_{V_{r(d)}}$ and $\tilde{\xi}^*_{V_d}$ are the $V$-optimal designs in the transformed coordinates based on $d$-point individual designs with repeated and non-repeated time points respectively. Clearly if this ratio is less than 1 or, equivalently, if $\Psi_V(\tilde{\xi}^*_{V_d}) - \Psi_V(\tilde{\xi}^*_{V_{r(d)}}) > 0$, then the optimal design with repeated points is more efficient than that for non-repeated points. The general result for $d \geq 3$ is presented in the following theorem.

**Theorem 5.6.1** *For $d \geq 3$, $V$-optimal population designs based on d-point individual designs with repeated time points are more efficient than the corresponding $V$-optimal population designs with non-repeated points for all $\gamma \geq 0$.*

**Proof**

Consider first the case of $d$ even. Recall from Subsection 5.2.2 and from Section 5.5 that the criterion $\Psi_V(\tilde{\xi})$ at the $V$-optimal population designs $\xi^*_{V_e}$ and $\xi^*_{V_{re(d)}}$ is given by

$$\Psi_V(\tilde{\xi}^*_{V_e}) = (k+1)\,(1+d\,\gamma) + \frac{1}{H}\,k\,(k+1)\,(k+2)$$

where $H = d^2 - 3\,d\,(k+1) + 3\,k^2 + 6\,k + 2$, and by

$$\Psi_V(\tilde{\xi}^*_{V_{re}}) = \frac{(k+1)\,(2+4\,k+3\,d\,k\,\gamma)}{3\,k}$$

respectively. Recall from Subsection 5.2.2 that $H > 0$. Therefore the difference between these criteria is given by

$$\Psi_V(\tilde{\xi}^*_{V_e}) - \Psi_V(\tilde{\xi}^*_{V_{re}}) = \frac{(d-2)\,(3\,k-d+1)\,(1+k)\,(2+k)}{3\,k\,H}.$$

Clearly this difference is greater than zero for $d \geq 3$ and $\gamma \geq 0$. Thus $\Psi_V(\tilde{\xi}^*_{V_e}) > \Psi_V(\tilde{\xi}^*_{V_{re}})$.

Consider now the case of $d$ odd. Recall from Subsection 5.2.2 and Section 5.5 that the criterion $\Psi_V(\tilde{\xi})$ at the $V$-optimal population designs $\tilde{\xi}^*_{V_o}$ and $\tilde{\xi}^*_{V_{ro}}$ is given by

$$\Psi(\tilde{\xi}^*_{V_o}) = (k+1)\,(d\,\gamma+1)\{1 + \frac{1}{H}\,d\,k\,(k+2)\}$$

where

$$H = d^3 - 3\,(k+1)\,(d^2+1) + (3\,k^2+6\,k+5)\,d + (d^2-1)\,\{d^2+3\,(k+1)\,(k-d+1)\}\,\gamma$$

and by

$$\Psi_V(\tilde{\xi}^*_{V_{ro}}) = \frac{(k+1)\,(1+d\,\gamma)\,(2\,d+4\,d\,k-3\,k\,\gamma+3\,d^2\,k\,\gamma)}{3\,k\,(d-\gamma+d^2\,\gamma)}$$

respectively. Recall from Subsection 5.2.2 that $H > 0$. Therefore the difference between these criteria is equal to

$$\Psi_V(\tilde{\xi}^*_{V_o}) - \Psi_V(\tilde{\xi}^*_{V_{ro}}) = \frac{d\,(d-1)\,(k+1)\,(k+2)\,(d\,\gamma+1)\,\{-(d+1)\,C_1(d)\,\gamma + C_0(d)\}}{3\,k\,\{(d^2-1)\,\gamma+d\}\,H}$$

where

$$C_0(d) = -d^2 + d\,(3\,k+2) - 3\,(k+1)$$

and

$$C_1(d) = d^2 - 3\,d\,(k+1) + 3\,(2\,k+1).$$

Since $\frac{\partial C_0(d)}{\partial d} = 3\,k - 2\,d + 2 > 0$ for $d \geq 3$ and $k \geq 2$, $C_0(3) = 6\,(k-1) > 0$ and $C_0(k+1) = 2\,(k+1)\,(k-1) > 0$, then $C_0(d) > 0$ on the interval $[3, k+1]$. Similarly, since $\frac{\partial C_1(d)}{\partial d} = 2\,d - 3\,(k+1) < 0$, $C_1(3) = -3\,(k-1) < 0$ and $C_1(k+1) = -2\,k\,(k-1) + 1 < 0$ for $d \geq 3$ and $k \geq 2$ the function $C_1(d) < 0$ on the interval $[3, k+1]$. Thus, over all, $\Psi_V(\tilde{\xi}^*_{V_o}) - \Psi_V(\tilde{\xi}^*_{V_{ro}}) > 0$ for all $\gamma \geq 0$. □

## 5.7 Trypanosomosis example

The results of this chapter are applied to the data from the experiment in susceptibility to trypanosmosis. This example was introduced in Chapter 3 and used in Section 4.8

to calculate $D$-optimal population designs. Here also only the data corresponding to the N'Dama breed are used. Assume that the variance components are the maximum likelihood estimates obtained from the data, i.e. $\sigma_b^2 = 4.181$, $\sigma_e^2 = 3.595$ and hence $\gamma = 1.163$. Assume also that only 84 observations on the cattle are affordable in the experiment and that all 36 days labelled 0, 1, ..., 35 are available for taking measurements. The objective is to estimate the mean responses at a vector of time points $\mathbf{t}_g = (0, 1, \ldots, 35)$ as precisely as possible. Using Theorems 5.2.2 and 5.2.3 two-, three-, four-, six-point, seven- and fourteen-point $V$-optimal population designs $\xi^*_{V_d}$ are presented in Table 5.1. The table also provides the values of $\Psi_V(\xi^*_{V_d}) = tr\{\mathbf{M}_\beta^{-1}(\xi^*_{V_d})\mathbf{X}'_g\mathbf{X}_g\}$ and the $V$-efficiency relative to the $V$-optimal design $\xi^*_{V_2}$. Observe that the $V$-efficiencies relative to $\xi^*_{V_2}$ decrease as the number of time points $d$ increases and that the $V$-optimum population design $\xi^*_{V_d}$ with small $d$ is more efficient than the one with large $d$.

Consider now the original experiment, that is suppose now that there are only the 14 days listed in the experiment are available for taking measurements. Since the days are not equally spaced the results of this chapter to calculate the $V$-optimum population designs for estimating the mean responses at a vector of time points as precisely as possible do not apply. Therefore a GAUSS program has been written to compute a $V$-optimal population design based on the set of $d$-point individual designs for $1 \leq d \leq 14$. The program is given in the file labelled "voptinte" on the CD provided with this thesis. The program calculates a $d$-point $V$-optimal population design for a given value of $\gamma$. The $d$-point $V$-optimal population designs for the experiment with $\gamma = 1.163$ and $d = 2, 3, 4, 6, 7, 14$ are presented in Table 5.2.

Observe that in contrast to the results in Table 5.1, the design weight for $V$-optimal

Table 5.1: $d$-point $V$-optimal population designs for the mean responses in the simple linear regression model with random intercept for the trypanosmosis data based on the set of points $\{0, 1, \ldots, 35\}$

| $d$ | $\xi^*_{V_d}$ | $w$ | $\Psi_V(\xi^*_{V_d})$ | $V$-efficiency |
|---|---|---|---|---|
| 2 | (0,35) | 1 | 168.422 | 1.0000 |
| 3 | (0,1,35) | 0.5 | | |
| | (0,34,35) | 0.5 | 216.54 | 0.7778 |
| 4 | (0,1,34,35) | 1 | 255.019 | 0.6604 |
| 6 | (0,1,2,33,34,35) | 1 | 341.84 | 0.4927 |
| 7 | (0,1,2,3,33,34,35) | 0.5 | | |
| | (0,1,2,32,33,34,35) | 0.5 | 386.391 | 0.4359 |
| 14 | (0,1,2,3,4,5,6,29,30,31,32,33,34,35) | 1 | 691.744 | 0.2435 |

Table 5.2: $d$-point $V$-optimal population designs for the mean responses in the simple linear regression model with random intercept for the trypanosmosis data in the actual study

| $d$ | $\xi^*_{V_a}$ | $w$ | $\Psi_V(\xi^*_{V_a})$ | $V$-efficiency |
|---|---|---|---|---|
| 2 | (0,35) | 1 | 51.9305 | 1.0000 |
| 3 | (0,2,35) | 0.5683 | | |
| | (0,31,35) | 0.4317 | 69.215 | 0.7503 |
| 4 | (0,2,34,35) | 1 | 85.4607 | 0.6077 |
| 6 | (0,2,4,39,31,35) | 1 | 118.982 | 0.4365 |
| 7 | (0,2,4,7,29,31,35) | 0.6245 | | |
| | (0,2,4,25,29,31,35) | 0.3755 | 136.044 | 0.3817 |
| 14 | (0,2,4,7,9,14,17,18,21,23,25,29,31,35) | 1 | 255.948 | 0.2029 |

designs when $d$ is an odd integer is not 0.5. For example, the design weight in $\xi^*_{V_3}$ is 0.5683 for design (0,2,35) and 0.4317 for design (0,31,35). However, $V$-efficiencies relative to $\xi^*_{V_2}$ exhibit a similar trend to those for Table 5.2 in that the efficiencies decrease as the number of time points $d$ increases.

# Chapter 6

# $D$-optimal Population Designs for the Quadratic Regression Model with a Random Intercept

## 6.1 Introduction

The quadratic regression model with additive errors is widely used in Statistics. It is well known (see for example Atkinson and Donev (1992, page 99)) that when the errors in this model are uncorrelated, the $D$-optimal approximate design for estimating the regression parameters over the design space $[-1, 1]$ puts equal weights of $\frac{1}{3}$, on the values -1, 0 and 1. When the number of observations, say $n$, is a multiple of 3 this design provides an exact $D$-optimal design. In this Chapter, $D$- and $D_s$-optimal population designs for estimation of the fixed effects in the quadratic regression model with a random intercept are now considered.

The definition of the quadratic regression model with a random intercept and the general form of the information matrix for the fixed effects are presented in Section 6.2. In Sections 6.3 and 6.4 the construction of $D$-optimal population designs for estimation of the fixed effects based on one- and two-point individual designs respectively are discussed algebraically. These optimal designs are identical for designs with non-repeated and repeated time points and therefore only designs in the set $S_{d,k}$ are considered in these sections. $D_s$-optimal population designs for estimation of the linear and quadratic regression coefficients $\beta_1$ and $\beta_2$ are discussed in Section 6.5. Finally, in Section 6.6 the $D$-optimal population designs based on $d$-point individual derived with $d \geq 3$ are derived numerically and discussed.

## 6.2 Preliminaries

Recall from Subsection 3.3.1 that the quadratic regression model with a random intercept which is of interest in this thesis is defined by

$$y_{ij} = \beta_0 + b_i + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + e_{ij}, \quad j = 1, \ldots d_i \text{ and } i = 1, \ldots, K \tag{6.1}$$

where $y_{ij}$ is the $j$th observation on individual $i$, $\beta_0$, $\beta_1$ and $\beta_2$ are the fixed effects, $t_{ij} \in \{0, 1, \ldots, k\}$ with $k \geq 2$ and $K$ is the number of individuals. It is assumed that $b_i \sim \mathcal{N}(0, \sigma_b^2)$, that $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$ and that $b_i$ and $e_{ij}$ are independent. The matrix form of this model is similar to that for the simple linear case except that here $\boldsymbol{\beta} = (\beta_o, \beta_1, \beta_2)'$, $\mathbf{X}_i = [\mathbf{1}_i \ \mathbf{t}_i \ \mathbf{t}_i^{(2)}]$ and $\tilde{\mathbf{X}}_i = [\mathbf{t}_i \ \mathbf{t}_i^{(2)}]$ where $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{id_i})'$ and $\mathbf{t}_i^{(2)}$ represents a column vector with elements equal to the squares of the time points.

Consider the population design

$$\xi = \left\{ \begin{array}{ccc} \mathbf{t}_1, & \ldots, & \mathbf{t}_r \\ w_1, & \ldots, & w_r \end{array} \right\} \quad \text{with } 0 < w_i < 1 \text{ and } \sum_{i=1}^{r} w_i = 1.$$

Then the information matrix for the parameters $\boldsymbol{\beta}$ in model (6.1) at a population design $\xi$ is given by

$$\mathbf{M}_\beta(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\beta(\mathbf{t}_i)$$

where

$$\mathbf{M}_\beta(\mathbf{t}_i) = \frac{1}{\sigma_e^2\, d_i\,(1 + d_i\,\gamma)} \begin{pmatrix} d_i & \sum_{j=1}^{d_i} t_{ij} & \sum_{j=1}^{d_i} t_{ij}^2 \\ \sum_{j=1}^{d_i} t_{ij} & A_1 & A_2 \\ \sum_{j=1}^{d_i} t_{ij}^2 & A_2 & A_3 \end{pmatrix} \tag{6.2}$$

with

$$A_1 = (1 + d_i\,\gamma) \sum_{j=1}^{d_i} t_{ij}^2 - \gamma \left( \sum_{j=1}^{d_i} t_{ij} \right)^2,$$

$$A_2 = (1 + d_i\,\gamma) \sum_{j=1}^{d_i} t_{ij}^3 - \gamma \left( \sum_{j=1}^{d_i} t_{ij} \right) \left( \sum_{j=1}^{d_i} t_{ij}^2 \right)$$

and

$$A_3 = (1 + d_i\,\gamma) \sum_{j=1}^{d_i} t_{ij}^4 - \gamma \left( \sum_{j=1}^{d_i} t_{ij}^2 \right)^2,$$

is the standardized information matrix for $\boldsymbol{\beta}$ at the design $\mathbf{t}_i$ derived from expression (2.29). Note that the error variance $\sigma_e^2$ factors out of expression (6.2) and it can be taken to be 1 without loss of generality.

The $D$-optimal population design based on $d$-point individual designs for estimation of $\boldsymbol{\beta}$ in the quadratic regression model with a random intercept, i.e. model (6.1), is that design which maximizes

$$\Psi(\xi) = \ln |\mathbf{M}_\beta(\xi)| = \ln |\sum_{i=1}^{r} w_i \mathbf{M}_\beta(\mathbf{t}_i)|$$

over the set of all population designs specified by $\xi$. Furthermore, it follows immediately from the Equivalence Theorem introduced in Subsection 2.6.4 that the design $\xi_D^*$ is $D$-optimal if and only if

$$\phi(\mathbf{t}, \xi_D^*) = tr\{\mathbf{M}_\beta^{-1}(\xi_D^*)\,\mathbf{M}_\beta(\mathbf{t})\} - 3 \leq 0 \tag{6.3}$$

for all individual designs $\mathbf{t}$ in the space of designs of interest, with equality holding at the support designs of $\xi_D^*$. Note that $\phi(\mathbf{t}, \xi_D^*)$ is the directional derivative of $\Psi(\xi) = \ln |\mathbf{M}_\beta(\xi)|$ at $\xi_D^*$ in the direction of $\mathbf{t}$.

## 6.3 $D$-optimal population designs based on one-point individual designs

Consider now the equally spaced time points, $0, 1, 2, \ldots, k$, where $k$ is a positive integer greater than or equal to 2. Then the space of one-point individual designs consists of these $k + 1$ time points. Let $t \in \{0, 1, \ldots, k\}$ be a single time point. Then it follows immediately from expression (6.2) that the standardized information matrix for $\boldsymbol{\beta}$ at $t$ on a per observation basis is given by

$$\mathbf{M}_\beta(t) = \frac{1}{1+\gamma}\begin{pmatrix} 1 & t & t^2 \\ t & t^2 & t^3 \\ t^2 & t^3 & t^4 \end{pmatrix}.$$

The term $(1 + \gamma)$ factors out of this expression and hence out of $\mathbf{M}_\beta(\xi)$ where $\xi$ is the one-point population design. Therefore the one-point $D$-optimal population designs for the quadratic random intercept model do not depend on the variance ratio $\gamma$. Thus the design problem is reduced to that of constructing $D$-optimal designs for a quadratic regression model with uncorrelated errors. For $k$ even the $D$-optimal design is

$$\xi^*_{D_e} = \left\{ \begin{array}{ccc} (0) & (\frac{k}{2}) & (k) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}$$

(Atkinson and Donev, 1992, page 99). Note that for the quadratic model at least three distinct time points are needed in order to estimate the three parameters in $\boldsymbol{\beta}$ and thus $k$ must be greater than or equal to 2. For $k$ odd the results relating to the one-point $D$-optimal population design are presented in the following Theorem.

**Theorem 6.3.1** *Consider the set of all one-point designs $t \in \{0, 1, \ldots, k\}$ where $k$ is an odd integer greater than or equal to 3. Then*

$$\xi^*_{D_o} = \left\{ \begin{array}{cccc} (0) & (\frac{k-1}{2}) & (\frac{k+1}{2}) & (k) \\ w & \frac{1}{2} - w & \frac{1}{2} - w & w \end{array} \right\},$$

*where $0 < w < \frac{1}{2}$ and*

$$w = \frac{k^2 - 2 + \sqrt{1 - k^2 + k^4}}{6\,(k^2 - 1)}$$

*is the $D$-optimal population design for the fixed effects $\boldsymbol{\beta}$ in model (6.1) over this set for all $\gamma \geq 0$.*

**Proof**

Recall from Subsection 2.6.4 that $D$-optimal designs for random intercept models are invariant to linear transformations. Thus without loss of generality, let a one-point individual design $t$ be linearly transformed according to $\tilde{t} = t - \frac{k}{2}$. Thus the space of designs of interest in the transformed coordinates is given by

$$\tilde{S}_{1,k} = \{\tilde{t} : \tilde{t} \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}-1, \frac{k}{2}\}\}$$

and the proposed optimum design $\xi^*_{D_o}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{D_o} = \left\{ \begin{array}{cccc} (-\frac{k}{2}) & (-\frac{1}{2}) & (\frac{1}{2}) & (\frac{k}{2}) \\ w & \frac{1}{2}-w & \frac{1}{2}-w & w \end{array} \right\}.$$

The proof is accomplished in two steps; the first step deals with the calculation of the weight $w$ and the second step shows that the proposed design is optimal.

### Step 1. Calculation of the weight $w$

The standardized information matrix for $\boldsymbol{\beta}$ at the design $\tilde{\xi}^*_{D_o}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}^*_{D_o}) = \frac{1}{1+\gamma} \begin{pmatrix} 1 & 0 & B_1 \\ 0 & B_1 & 0 \\ B_1 & 0 & B_2 \end{pmatrix} \tag{6.4}$$

where $B_1 = \frac{1}{4}(1 - 2\,w + 2\,k^2\,w)$ and $B_2 = \frac{1}{16}(1 - 2\,w + 2\,k^4\,w)$. Then the weight $w$ must be chosen to maximize the determinant of $\mathbf{M}_\beta(\tilde{\xi}^*_{D_o})$, that is $w$ is chosen to maximize

$$\left|\mathbf{M}_\beta(\tilde{\xi}^*_{D_o})\right| = \frac{(k-1)^2\,(1+k)^2\,w\,(1-2\,w)\,(1-2\,w+2\,k^2\,w)}{32\,(1+\gamma)^3}. \tag{6.5}$$

Taking the first derivative of expression (6.5) with respect to $w$ and equating it to zero gives the two solutions for $w$ as

$$w_1 = \frac{k^2 - 2 + \sqrt{1 - k^2 + k^4}}{6\;(k^2 - 1)}$$

and

$$w_2 = \frac{k^2 - 2 - \sqrt{1 - k^2 + k^4}}{6\,(k^2 - 1)}.$$

Consider the solution $w_2$. For $k \geq 3$, it follows that $k^2 - 2 > 0$, $k^2 - 1 > 0$ and $1 - k^2 + k^4 > 0$. Therefore $(k^2 - 2)^2 - (1 - k^2 + k^4) = -3\,(k^2 - 1) < 0$ for $k \geq 3$ and this in turn implies that the numerator of $w_2$, $k^2 - 2 - \sqrt{1 - k^2 + k^4}$ is less than zero. The denominator of $w_2$, $6\,(k^2 - 1)$ is clearly positive for $k \geq 3$. Thus $w_2 < 0$ and is not an acceptable weight.

Consider now the solution $w_1$. The terms $k^2 - 2$ and $k^2 - 1$ in $w_1$ are positive for $k \geq 3$. Therefore $w_1$ is strictly positive for $k \geq 3$. It is clear from the nature of the proposed design $\xi^*_{D_o}$ that a weight $w$ is only acceptable provided $w \leq \frac{1}{2}$. Consider now the difference

$$\frac{1}{2} - w_1 = \frac{2\,k^2 - 1 - \sqrt{1 - k^2 + k^4}}{6\,(k^2 - 1)}.$$

The term $2\,k^2 - 1$ is positive for $k \geq 3$ and since

$$(2\,k^2 - 1)^2 - (k^4 - k^2 + 1) = 3\,k^2\,(k^2 - 1) > 0$$

for $k \geq 3$, it follows that $2\,k^2 - 1 - \sqrt{1 - k^2 + k^4} > 0$. Thus $w_1 \leq \frac{1}{2}$. Further observe that

$$\left.\frac{\partial^2 \left|\mathbf{M}_\beta(\tilde{\xi}^*_{D_o})\right|}{\partial w^2}\right|_{w=w_1} = -\frac{(k-1)^2\,(k+1)^2\,\sqrt{1 - k^2 + k^4}}{8\,(1+\gamma)^3} < 0$$

so that $\left|\mathbf{M}_\beta(\tilde{\xi}^*_{D_o})\right|$ is a maximum at $w = w_1$.

**Step 2. Optimality of the proposed design**

Substituting $w_1$ for $w$ in (6.4) and inverting the resultant matrix gives the directional derivative of $\Psi(\tilde{\xi}) = \ln|\mathbf{M}_\beta(\tilde{\xi})|$ at $\tilde{\xi}^*_{D_o}$ in the direction of a one-point design $\tilde{t} \in \tilde{S}_{1,k}$ as

$$\phi(\tilde{t}, \tilde{\xi}^*_{D_o}) = tr[\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_o})\mathbf{M}_\beta(\tilde{t})] - 3$$

$$= -\frac{9\,(k-2\tilde{t})\,(k+2\tilde{t})\,(2\tilde{t}-1)\,(2\tilde{t}+1)}{(2\,k^2-1-\sqrt{1-k^2+k^4})\,(k^2-2+\sqrt{1-k^2+k^4})}. \tag{6.6}$$

It has been shown earlier that the terms $2\,k^2-1-\sqrt{1-k^2+k^4}$ and $k^2-2$ are positive for $k \geq 3$, so the denominator in (6.6) is greater than zero. Furthermore, $(k-2\tilde{t})\,(k+2\tilde{t}) \geq 0$ since $-\frac{k}{2} \leq \tilde{t} \leq \frac{k}{2}$. Thus the sign of directional derivative $\phi(\tilde{t}, \tilde{\xi}^*_{D_o})$ is determined by the sign of $-(2\tilde{t}-1)\,(2\tilde{t}+1)$. So $\phi(\tilde{t}, \tilde{\xi}^*_{D_o}) \leq 0$ if and only if $-(2\tilde{t}-1)\,(2\tilde{t}+1) \leq 0$ and thus if and only if $\tilde{t} \leq -\frac{1}{2}$ or $\tilde{t} \geq \frac{1}{2}$. However, for $k$ odd the points $\tilde{t}$ such that $-\frac{1}{2} < \tilde{t} < \frac{1}{2}$ are not in the set $\tilde{S}_{1,k}$. Thus $\phi(\tilde{t}, \tilde{\xi}^*_{D_o}) \leq 0$ for all single point designs in $\tilde{S}_{1,k}$. Furthermore, $\phi(\tilde{t}, \tilde{\xi}^*_{D_o}) = 0$ at the support points of $\tilde{\xi}^*_{D_o}$. Thus by the Equivalence Theorem of Subsection 2.6.4, $\xi^*_{D_o}$ is the $D$-optimal population design for the fixed effects $\boldsymbol{\beta}$ over the set $S_{1,k}$ for all $\gamma \geq 0$. □

Recall that the optimum design weight in the above theorem is

$$w = \frac{k^2-2+\sqrt{1-k^2+k^4}}{6\;(k^2-1)}$$

where $k$ is an odd integer greater than or equal to 3. From the expression for the weight $w$ it is clear that $w$ depends on $k$. Differentiating $w$ with respect to $k$ yields

$$\frac{\partial w}{\partial k} = \frac{k\,\{2\,\sqrt{1-k^2+k^4}-(k^2+1)\}}{6\,(k^2-1)^2\,\sqrt{1-k^2+k^4}}.$$

For $k \geq 3$, the terms $2\,\sqrt{1-k^2+k^4}$ and $k^2+1$ are strictly positive. Therefore $4\,(1-k^2+k^4)-(k^2+1)^2 = 3\,(k^2-1)^2 > 0$ implies that $2\,\sqrt{1-k^2+k^4}-(k^2+1) > 0$. Further the denominator $6\,(k^2-1)^2\,\sqrt{1-k^2+k^4} > 0$. So $\dfrac{\partial w}{\partial k} > 0$ for $k \geq 3$ and thus $w$ is increasing monotonically with $k$ for $k$ an odd integer greater than or equal to 3.

For $k = 3$, $w = \frac{1}{48}(7+\sqrt{73}) = 0.323833$. Furthermore, observe that

$$\lim_{k\to\infty} w = \lim_{k\to\infty} \left\{ \frac{(1-2/k^2+\sqrt{1-1/k^2+1/k^4})}{6\,(1-1/k^2)} \right\} = \frac{1}{3}.$$

Thus overall the optimum design weight $w$ increases from 0.323833 asymptotically to $\frac{1}{3}$ as $k \to \infty$. The graph of the weight $w$ against $k$ is presented in Figure 6.1.

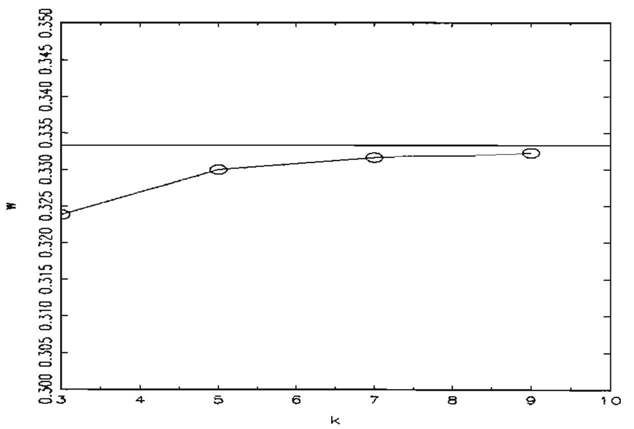


Figure 6.1: Plot of the design weight $w$ defined in Theorem 6.3.1 against $k$.

Suppose that the discrete time points in $\tilde{S}_{1,k}$ are scaled to lie between -1 and 1, that is each point in $\tilde{S}_{1,k}$ is multiplied by a factor $\frac{2}{k}$. Then the $D$-optimal population design $\tilde{\xi}^*_{D_o}$ becomes

$$\tilde{\xi}^*_{D_o} = \left\{ \begin{array}{cccc} (-1) & (-\frac{1}{k}) & (\frac{1}{k}) & (1) \\ w & \frac{1}{2} - w & \frac{1}{2} - w & w \end{array} \right\}.$$

Now as $k \to \infty$, so $w \to \frac{1}{3}$ and $\frac{1}{k} \to 0$ and this design simplifies to

$$\tilde{\xi}^*_{D} = \left\{ \begin{array}{ccc} (-1) & (0) & (1) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}$$

which coincides with the $D$-optimal approximate design for the quadratic regression model with uncorrelated errors based on the design space [-1, 1].

**Example 6.3.1** Consider the quadratic regression model with a random intercept as specified by (6.1) and let $k = 9$. By Theorem 6.3.1 the design

$$\xi^*_{D_o} = \left\{ \begin{array}{cccc} (0) & (4) & (5) & (9) \\ 0.3323 & 0.1677 & 0.1677 & 0.3323 \end{array} \right\}$$

with $w = \frac{1}{480}(79 + \sqrt{6481}) = 0.1677$ is the $D$-optimal population design for all $\gamma \geq 0$. It is straightforward to show that $\phi(\tilde{t}, \tilde{\xi}^*_{D_o}) = tr[\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_o})\mathbf{M}_\beta(\tilde{t})] - 3$ can be expressed as

$$\phi(\tilde{t}, \tilde{\xi}^*_{D_o}) = \frac{9\,(2\,\tilde{t} - 9)\,(9 + 2\,\tilde{t})\,(2\,\tilde{t} - 1)\,(1 + 2\,\tilde{t})}{(161 - \sqrt{6481})\,(79 + \sqrt{6481})},$$

where $\tilde{\xi}^*_{D_o}$ is the linearly transformed version of $\xi^*_{D_o}$. Observe that $\phi(\tilde{t}, \tilde{\xi}^*_{D_o}) \leq 0$ for $\tilde{t} \in \{-\frac{9}{2}, \ldots, -\frac{1}{2}, \frac{1}{2}, \ldots, \frac{9}{2}\}$ and that equality holds at the support designs of $\tilde{\xi}^*_{D_o}$ for all $\gamma \geq 0$. In Figure 6.2 a graph of the directional derivative $\phi(t, \xi^*_{D_o})$ against the individual designs $t = 0, 1, \ldots, 9$ is presented. The figure shows that $\phi(t, \xi^*_{D_o}) \leq 0$ for all one-point designs in the space of designs $S_{1,9}$, that equality holds at the support designs of $\xi^*_{D_o}$ and thus that $\xi^*_{D_o}$ is the $D$-optimal population design.

## 6.4 $D$-optimal population designs based on two-point individual designs

In this Section, the construction of $D$-optimal designs for the precise estimation of the parameters $\boldsymbol{\beta}$ in model (6.1) when the individual designs consist of two time points is discussed. Before the general results for two-point designs are given a lemma on directional derivatives of symmetric designs will be presented. A two-point population design $\xi$ is called

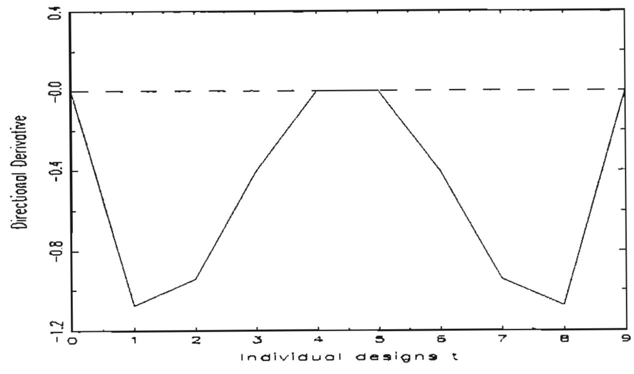


Figure 6.2: Plot of the directional derivative $\phi(t, \xi^*_{D_o})$ against the individual designs $t \in S_{1,9}$ for $k = 9$.

symmetric if it contains individual designs $(x_1, x_2)$ and $(-x_2, -x_1)$ as support designs with equal weight. The lemma will show that for symmetric population designs the directional derivative associated with $D$-optimality in the direction of a two-point design attains its maximum along a portion of the boundary of the design region for two-point designs or at the origin. The lemma is presented below and the proof is based on that of Lemma 3.1 in Cheng (1995) and, more specifically, of Theorem 3.1 in Atkins and Cheng (1999).

**Lemma 6.4.1** *Consider a symmetric population design $\xi_s$ comprising two-point designs $\mathbf{x} = (x_1, x_2)$ with $x_i \in [-\frac{k}{2}, \frac{k}{2}]$, $i = 1, 2$. Then the directional derivative associated with $D$-optimality at $\xi_s$ in the direction of $\mathbf{x}$, $\phi(\mathbf{x}, \xi_s)$, attains its maximum on that portion of the boundary of the design region for $\mathbf{x}$ defined by $x_1 = -\dfrac{k}{2}$ and $0 \leq x_2 \leq \dfrac{k}{2}$ or at the origin (0, 0).*

**Proof**

For a symmetric design $\xi_s$ the inverse of the information matrix $\mathbf{M}_\beta(\xi_s)$ has the form

$$\mathbf{M}_\beta^{-1}(\xi_s) = \begin{pmatrix} E_1 & 0 & E_2 \\ 0 & E_3 & 0 \\ E_2 & 0 & E_4 \end{pmatrix}.$$

Note that since $\mathbf{M}_\beta^{-1}(\xi_s)$ is a variance matrix $E_1, E_3$ and $E_4$ are greater than zero. Then the directional derivative of $\Psi(\xi) = \ln|\mathbf{M}_\beta(\xi)|$ at $\xi_s$ in the direction of the design $\mathbf{x} = (x_1, x_2)$ has the form

$$\begin{aligned}\phi(\mathbf{x}, \xi_s) &= tr[\mathbf{M}_\beta^{-1}(\xi_s)\mathbf{M}_\beta(\mathbf{x})] - 3 \\ &= \frac{1}{2\,(1+2\,\gamma)} \{E_4\,(1+\gamma)\,(x_1^4 + x_2^4) + \{2\,E_2 + E_3\,(1+\gamma)\}\,(x_1^2 + x_2^2) - 2\,E_4\,\gamma\,x_1^2\,x_2^2 \\ &\qquad - 2\,E_3\,\gamma\,x_1\,x_2 + 2\,(E_1 - 6\,\gamma - 3)\} \qquad (6.7)\end{aligned}$$

where $\mathbf{M}_\beta(\mathbf{x})$ is the information matrix for $\boldsymbol{\beta}$ at $\mathbf{x} = (x_1, x_2)$ on a per observation basis. It is immediately obvious from this expression that

$$\phi((x_1, x_2), \xi_s) = \phi((x_2, x_1), \xi_s)$$

and that

$$\phi((x_1, x_2), \xi_s) = \phi((-x_2, -x_1), \xi_s)$$

and hence that $\phi(\mathbf{x}, \xi_s)$ is symmetric about $x_1 = x_2$ and $x_1 = -x_2$. Therefore the search for a maximum of $\phi(\mathbf{x}, \xi_s)$ can be restricted to the region defined by $-\frac{k}{2} \leq x_1 \leq 0$ and $x_1 \leq x_2 \leq -x_1$. Furthermore it is clear that since $E_3 > 0$ and $\gamma \geq 0$, the coefficient of $x_1\,x_2$ in $\phi(\mathbf{x}, \xi_s)$, $-2\,E_3\,\gamma$, is less than or equal to zero. Thus $\phi((x_1, x_2), \xi_s) > \phi((x_1, -x_2), \xi_s)$

when $x_1 x_2 < 0$ and the search for the maximum of the directional derivative can be further restricted to the region defined by $-\frac{k}{2} \leq x_1 \leq 0$ and $0 \leq x_2 \leq -x_1$.

Thus to complete the proof now it is only necessary to look for a maximum value of $\phi((x_1, -c\,x_1), \xi_s)$ where $x_1 \in [-\frac{k}{2}, 0]$ and $c \in [0, 1]$. Consider therefore rays starting at the origin and defined by $(x_1, -c\,x_1)$ for $-\frac{k}{2} \leq x_1 \leq 0$ and $c \in [0, 1]$ which cover the region of interest. Then it is only necessary to investigate the maximum value of the directional derivative along such a ray, in other words $\phi((x_1, -c\,x_1), \xi_s)$. Now for $x_2 = -c\,x_1$,

$$\begin{aligned}
\phi((x_1, -c\,x_1), \xi_s) &= f(x_1) \\
&= \frac{1}{2\,(1+2\,\gamma)} \left\{E_4\,[(1+c^4) + (c^2-1)^2\,\gamma]\,x_1^4 + [(1+c^2)\,(2\,E_2 + E_3)\right. \\
&\qquad \left. + E_3\,(c+1)^2\,\gamma]\,x_1^2 + 2\,(E_1 - 6\,\gamma - 3)\right\}.
\end{aligned}$$

The function $f(x_1)$ is quartic in $x_1$ and has the form $A\,x_1^4 + B\,x_1^2 + C$. Clearly $f(x_1)$ is symmetric in $x_1$ about zero. The coefficient of $x_1^4$ in $f(x_1)$ is $A = E_4\,[(1+c^4) + (c^2-1)^2\,\gamma]$ and is greater than zero since $E_4 > 0$ and $\gamma \geq 0$. Thus $f(x_1)$ approaches $\infty$ as $x_1$ approaches $\infty$ or $-\infty$.

Differentiating $f(x_1)$ with respect to $x_1$ yields

$$\frac{\partial f(x_1)}{\partial x_1} = \frac{2\,x_1\,(A\,x_1^2 + B)}{1+2\,\gamma}.$$

This expression is equal to zero if and only if $x_1 = 0$ or $2\,A\,x_1^2 + B = 0$, where $A = E_4\,[(1+c^4) + (c^2-1)^2\,\gamma]$ and $B = [(1+c^2)\,(2\,E_2+E_3) + E_3\,(c+1)^2\,\gamma]$, and thus if and only if $x_1 = 0$ or $x_1 = \pm\sqrt{-\frac{B}{2\,A}}$. Thus if $\frac{B}{2\,A} \geq 0$, $A\,x_1^4 + B\,x_1^2 + C$ has one turning point at zero and if $\frac{B}{2\,A} < 0$ then $A\,x_1^4 + B\,x_1^2 + C$ has three turning points at $-\sqrt{-\frac{B}{2\,A}}$, 0 and $\sqrt{-\frac{B}{2\,A}}$. Consider now $x_1$ restricted to the interval $[-\frac{k}{2}, 0]$. It then follows that the maximum value of $\phi((x_1, -cx_1), \xi_s)$ is either at $(0,0)$ or at $(-\frac{k}{2}, x_2)$ where $0 \le x_2 \le \frac{k}{2}$. Note that whether the maximum of the directional derivative $\phi((x_1, x_2), \xi_s)$ is at $(0,0)$ or $(-\frac{k}{2}, x_2)$ where $0 \le x_2 \le \frac{k}{2}$ depends on the nature of the design. $\square$

**Corollary 6.4.1** *Consider lattice points* $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2)$ *with* $\tilde{t}_j \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}-1, \frac{k}{2}\}$, $j = 1, 2$ *in the region bounded by* $\tilde{t}_2 = -\tilde{t}_1$, $\tilde{t}_1 = -\frac{k}{2}$ *and* $\tilde{t}_2 = 0$. *Then by Lemma 6.4.1 the maxima for the directional derivative* $\phi(\tilde{\mathbf{t}}, \xi_s)$ *associated with* $D$*-optimality at a symmetric population design* $\xi_s$ *in the direction of* $\tilde{\mathbf{t}}$ *occur either at the end points of the segments of the rays containing lattice points, or at the closest lattice points to the boundary* $\tilde{t}_1 = -\frac{k}{2}$ *or at the origin (0, 0).*

For example, consider the design region for $k = 10$. Suppose that the design region is bounded by a ray $\tilde{t}_2 = -\frac{2}{5}\tilde{t}_1$, $\tilde{t}_1 = -5$ and $\tilde{t}_2 = 0$, and that the directional derivative $\phi((-5, \tilde{t}_2), \xi_s)$ is positive for $\tilde{t}_2 \in [1, 2]$ (see Figure 6.3). There are two lattice points (-3,1) and (-4,1) in the region. Thus by the above corollary $\phi((\tilde{t}_1, \tilde{t}_2), \xi_s)$ has to be examined at these points for its maxima. However, if $\phi((-5, \tilde{t}_2), \xi_s) > 0$ for $\tilde{t}_2 \in [0, 1]$ on the design region bounded by the ray $\tilde{t}_2 = -\frac{1}{5}\tilde{t}_1$, $x_1 = -5$ and $\tilde{t}_2 = 0$ there is no lattice point to check for a maximum of $\phi((-5, \tilde{t}_2), \xi_s)$.

The $D$-optimal population designs based on two-point individual designs with $k$ even are presented in the following three theorems. The above lemma and corollary are used extensively in the proofs of these results.

**Theorem 6.4.1** *Consider the set of all two-point individual designs* $\mathbf{t} = (t_1, t_2)$ *which put equal weights on the distinct time points* $t_1$ *and* $t_2$ *with* $t_j \in \{0, 1, \ldots, k\}, j = 1, 2,$ *and*

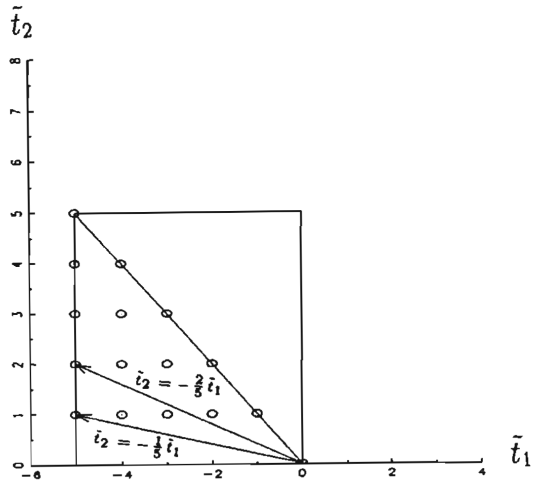


Figure 6.3: The design region for two-point designs with $k = 10$. The symbol o represents a lattice point.

*$0 \leq t_1 < t_2 \leq k$ for $k$ an even integer greater than or equal to 2. Then*

$$\xi_D^* = \left\{ \begin{array}{ccc} (0, \frac{k}{2}) & (0, k) & (\frac{k}{2}, k) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\},$$

*is the $D$-optimal population design for the fixed effects $\boldsymbol{\beta}$ in the model (6.1) over this set when (i) $k \leq 4$ and for all $\gamma \geq 0$, and (ii) $k > 4$ provided that $\gamma \leq \gamma(k) = \dfrac{3\,(k+2)}{k^2 - 3\,k - 6}$.*

**Proof**

Consider the individual designs $\mathbf{t} = (t_1, t_2)$ be linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$, where $\mathbf{x}_c = (\frac{k}{2}, \frac{k}{2})$. Thus the space of designs of interest in the transformed coordinates is given by

$$\tilde{S}_{2,k} = \{\tilde{\mathbf{t}} : \tilde{\mathbf{t}} \in (\tilde{t}_1, \tilde{t}_2),\ \tilde{t}_i \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}-1, \frac{k}{2}\},\ i=1,2\ \ -\frac{k}{2} \leq \tilde{t}_1 < \tilde{t}_2 \leq \frac{k}{2}\}$$

and the proposed optimum design $\xi_D^*$ can be written in the transformed coordinates as

$$\tilde{\xi}_D^* = \left\{ \begin{array}{ccc} (-\frac{k}{2}, 0) & (-\frac{k}{2}, \frac{k}{2}) & (0, \frac{k}{2}) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}.$$

The standardized information matrix for $\boldsymbol{\beta}$ at the design $\tilde{\xi}_D^*$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}_D^*) = \frac{1}{6\,(1+2\,\gamma)} \begin{pmatrix} 6 & 0 & k^2 \\ 0 & \frac{1}{2}\,k^2\,(2+3\,\gamma) & 0 \\ k^2 & 0 & \frac{1}{8}\,k^4\,(2+\gamma) \end{pmatrix}.$$

It then follows that

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}_D^*) = \frac{3\,(1+2\,\gamma)}{2+3\,\gamma} \begin{pmatrix} 2+\gamma & 0 & -\frac{8}{k^2} \\ 0 & \frac{4}{k^2} & 0 \\ -\frac{8}{k^2} & 0 & \frac{48}{k^4} \end{pmatrix} \tag{6.8}$$

and that

$$|\mathbf{M}_\beta(\tilde{\xi}^*)| = \frac{k^6\,(2+3\,\gamma)^2}{1728\,(1+2\,\gamma)^3}.$$

Then the directional derivative of $\Psi(\tilde{\xi}) = \ln|\mathbf{M}_\beta(\tilde{\xi})|$ at the population design $\tilde{\xi}_D^*$ in the direction of a two-point design $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2) \in \tilde{S}_{2,k}$ is given by

$$\begin{aligned}\phi(\tilde{\mathbf{t}}, \tilde{\xi}_D^*) &= tr[\mathbf{M}_\beta^{-1}(\tilde{\xi}^*)\mathbf{M}_\beta(\tilde{\mathbf{t}})] - 3 \\ &= \frac{1}{k^4\,(2+3\,\gamma)}\,\{72\,(1+\gamma)\,(\tilde{t}_1^4+\tilde{t}_2^4) - 144\,\gamma\,\tilde{t}_1^2\,\tilde{t}_2^2 + 6\,k^2\,(\gamma-3)\,(\tilde{t}_1^2+\tilde{t}_2^2) \\ &\qquad -12\,k^2\,\gamma\,\tilde{t}_1\,\tilde{t}_2 - 6\,k^4\,\gamma\}\end{aligned}$$

where $\mathbf{M}_{\beta}(\tilde{\mathbf{t}})$ is the standardized information matrix for $\boldsymbol{\beta}$ in expression (6.2) at the individual design $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2)$.

In the first part of the theorem there are only two cases, $k = 2$ and $k = 4$. When $k = 2$, the proof is trivial since there is only one population design

$$\tilde{\xi}_D^* = \left\{ \begin{array}{ccc} (-1,0) & (-1,1) & (0,1) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}$$

and this is necessarily optimal. In the case of $k = 4$, the proposed optimum population design in the transformed coordinates is

$$\tilde{\xi}_D^* = \left\{ \begin{array}{ccc} (-2,0) & (-2,2) & (0,2) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}$$

and it is symmetric. Thus by Corollary 6.4.1 it is only necessary to check $\phi(\tilde{\mathbf{t}}, \tilde{\xi}_D^*)$ at the design $\tilde{\mathbf{t}} = (-2, 1)$. At this design

$$\phi((-2,1), \tilde{\xi}_D^*) = -\frac{3\,(9+\gamma)}{32\,(2+3\,\gamma)} < 0$$

for $\gamma \geq 0$. Further, the derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}_D^*)$ is equal to zero at each of the support designs (-2, 0), (-2, 2) and (0, 2).

Consider now $\phi(\tilde{\mathbf{t}}, \tilde{\xi}_D^*)$ for the case of $k > 4$. Assume that $\tilde{\mathbf{t}} = \mathbf{x} = (x_1, x_2)$ where $x_i \in [-\frac{k}{2}, \frac{k}{2}], i = 1, 2$. Observe that $\tilde{\xi}_D^*$ is symmetric. Therefore by Lemma 6.4.1 the maximum of the directional derivative $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ occurs on that portion of the boundary of the design region defined by $x_1 = -\frac{k}{2}$ and $0 \leq x_2 \leq \frac{k}{2}$ or at the origin.

Consider first the directional derivative $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ at the origin. At the origin (0, 0) the derivative

$$\phi((0,0), \tilde{\xi}_D^*) = -\frac{6\,\gamma}{2+3\,\gamma} \leq 0$$

for $\gamma \geq 0$.

Consider now the derivative $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ on the design region defined by $x_1 = -\frac{k}{2}$ and $0 \leq x_2 \leq \frac{k}{2}$. For $x_1 = -\frac{k}{2}$

$$\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) = \frac{6}{k^4\,(2+3\,\gamma)}\,(k-2\,x_2)\,x_2\,\{k^2\,\gamma - 3\,k\,(1+\gamma)\,x_2 - 6\,(1+\gamma)\,x_2^2\}.$$

This is quartic in $x_2$. It can be easily seen that $\dfrac{k}{2}$ and 0 are the two zero points of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$. Since $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ is a quartic in $x_2$, it has four zero points not necessarily distinct. Solving the equation

$$f(x_2) = k^2\,\gamma - 3\,k\,(1+\gamma)\,x_2 - 6\,(1+\gamma)\,x_2^2 = 0$$

gives the other two zero points of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ as

$$r_1 = \frac{k\,\{\sqrt{3\,(3+14\,\gamma+11\,\gamma^2)} - 3\,(1+\gamma)\}}{12\,(1+\gamma)}$$

and

$$r_2 = -\frac{k\,\{\sqrt{3\,(3+14\,\gamma+11\,\gamma^2)} + 3\,(1+\gamma)\}}{12\,(1+\gamma)}.$$

Clearly the solution $r_2 < 0$ and so $r_2$ do not fall in the interval $[0, \frac{k}{2}]$. Thus consider only the solution $r_1$. Since $3\,(1+\gamma) > 0$ and $\sqrt{3\,(3+14\,\gamma+11\,\gamma^2)} > 0$ it follows that

$$3\,(3+14\,\gamma+11\,\gamma^2) - 9\,(1+\gamma)^2 = 24\,\gamma\,(1+\gamma) \geq 0$$

which implies that $\sqrt{3\,(3+14\,\gamma+11\,\gamma^2)} - 3\,(1+\gamma) \geq 0$. Thus $r_1 \geq 0$ for all $\gamma \geq 0$ and equality holds at $\gamma = 0$. Since

$$\frac{\partial r_1}{\partial \gamma} = \frac{k}{(1+\gamma)\,\sqrt{3\,(3+14\,\gamma+11\,\gamma^2)}} > 0$$

for $k > 4$ and $\gamma \geq 0$, $r_1$ is monotonically increasing with $\gamma$. Consider also the difference

$$\frac{k}{2} - r_1 = \frac{k\,\{9\,(1+\gamma) - \sqrt{3\,(3+14\,\gamma+11\,\gamma^2)}\}}{12\,(1+\gamma)}.$$

Since $9\,(1+\gamma) > 0$, $81\,(1+\gamma)^2 - 3\,(3+14\,\gamma+11\,\gamma^2) = 24\,(3+5\,\gamma+2\,\gamma) > 0$ implies that $9\,(1+\gamma) - \sqrt{3\,(3+14\,\gamma+11\,\gamma^2)} > 0$. Therefore $r_1 \in [0, \frac{k}{2})$.

Note that since $\gamma \geq 0$ the coefficient of $x_2^4$ in $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$, $\dfrac{72\,(1+\gamma)}{k^4\,(2+3\,\gamma)}$, is greater than zero. Thus $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ approaches $\infty$ as $x_2$ approaches to $\infty$ or $-\infty$. Furthermore, since $0 \leq x_2 \leq \frac{k}{2}$ the term $k - 2x_2$ and the function $f(x_2)$ are greater than or equal to zero for $r_2 \leq x_2 \leq r_1$. The form of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ is shown in Figure 6.4. When $0 \leq x_2 < r_1$, $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) > 0$ and when $r_1 < x_2 \leq \frac{k}{2}$, $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) \leq 0$.

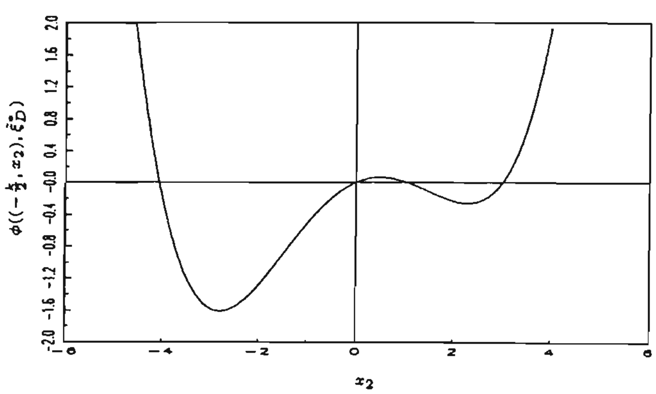


Figure 6.4: Plot of the directional derivative $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ against $x_2$ for $k = 6$ and $\gamma = 2.5$.

For $\tilde{\xi}_D^*$ to be optimal $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) \leq 0$ for $x_2 \in [0, \frac{k}{2}]$. Looking at Figure 6.4, for the inequality $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) \leq 0$ to hold it is necessary that $r_1 \leq 1$. For $r_1 < x_2 \leq 1$, $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) \leq 0$ and thus $\phi((x_1, x_2), \tilde{\xi}_D^*) < 0$ for all points $(x_1, x_2)$ in the region bounded by a line $x_2 = -x_1$, a ray $x_2 = -\frac{k}{2} x_1$ and $x_2 = -\frac{k}{2}$. For $0 < x_2 < r_1$, $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) > 0$

and hence by Corollary 6.4.1 the sign of $\phi((x_1, x_2), \tilde{\xi}_D^*)$ should be examined at the lattice points $(x_1, x_2)$ in the region bounded by a ray $x_2 = -\frac{k}{2}\, x_1$, $x_2 = -\frac{k}{2}$ and $x_2 = 0$. However there is no lattice point in that region. Thus the design $\tilde{\xi}_D^*$ is optimal if and only if $r_1 \leq 1$.

Consider now $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ at $x_2 = 1$, that is consider

$$\phi((-\frac{k}{2}, 1), \tilde{\xi}_D^*) = \frac{1}{k^4\,(2 + 3\,\gamma)}\{6\,(k-2)\,[k^2\,\gamma - 3\,k\,(1+\gamma) - 6\,(1+\gamma)]\}.$$

This is less or equal to zero if and only if $\gamma \leq \dfrac{3\,(k+2)}{k^2 - 3\,k - 6}$. Note that $r_1$ is monotonically increasing with $\gamma$. Furthermore, $\phi((-\frac{k}{2}, \tilde{t}_2), \tilde{\xi}_D^*) = 0$ at $x_2 = 0$ and $x_2 = \frac{k}{2}$, i.e. at the support designs of $\tilde{\xi}_D^*$. □

The condition $k > 4$ in the theorem is a result of $\gamma(k)$. Since $\gamma$ is a variance ratio its value must be greater than or equal to zero. This holds if $k$ is an integer greater than 4 since $k^2 - 3\,k - 6 > 0$ for $k > 4.37228$.

Cheng (1995) and Atkins and Cheng (1999) show for model (6.1) with $t_j \in [-1, 1], j = 1, 2$ that the $D$-optimal population designs comprising two support points put equal weights at (-1, 1), (-1, 0) and (0, 1) as $\gamma \to 0$. Suppose that the points in the individual designs $\tilde{\mathrm{t}} \in \tilde{S}_{2,k}$ are multiplied by $\dfrac{2}{k}$ so that the points lie between -1 and 1. Then the optimum design in Theorem 6.4.1 is expressed as

$$\tilde{\xi}_D^* = \left\{ \begin{array}{ccc} (-1,0) & (-1,1) & (0,1) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}.$$

Observe that $\lim_{k\to\infty} \gamma(k) = 0$. Thus for large $k$ when the time points are scaled to lie between -1 and 1, $\tilde{\xi}_D^*$ of Theorem 6.4.1 coincides with the asymptotic results of Cheng (1995) and Atkins and Cheng (1999).

**Theorem 6.4.2** *Consider the set of all two-point individual designs* $\mathbf{t} = (t_1, t_2)$ *which put equal weights on the distinct time points* $t_1$ *and* $t_2$ *with* $t_j \in \{0, 1, \ldots, k\}, j = 1, 2$ *and* $0 \leq t_1 < t_2 \leq k$ *for* $k$ *an even integer greater than or equal to 6. Then*

$$\xi_D^* = \left\{ \begin{array}{ccc} (0, \frac{k}{2}+1) & (\frac{k}{2}-1, k) & (0, k) \\ w & w & 1-2w \end{array} \right\},$$

*where*

$$w = \frac{B - \sqrt{A}}{3\,(k-2)\,\{2 + k + \gamma\,(3\,k+2)\}}$$

$$A = 16 + k^2\,(k^2-4) + 2\,(2+k)\,(16 - 8\,k^2 + 3\,k^3)\,\gamma + (96 + 96\,k - 60\,k^2 - 28\,k^3 + 15\,k^4)\,\gamma^2$$
$$+ 2\,(k-2)\,(2+3\,k)\,(-8 - 4\,k + 3\,k^2)\,\gamma^3 + (k-2)^2\,(2+3\,k)^2\,\gamma^4,$$

*and*

$$B = (k-2)\,(2+3\,k)\,\gamma^2 + 2\,(3\,k^2 - 2\,k - 4)\,\gamma + 2\,(k^2-2)$$

*is the D-optimal population design for the fixed effects* $\boldsymbol{\beta}$ *in the model (6.1) over this set whenever* $\gamma_c \leq \gamma < \gamma_d$ *where* $\gamma_c$ *is the only positive root of the cubic*

$$(k^2 - 3\,k - 6)\,(k^2 + 9\,k + 2)\,\gamma^3 + (k^4 - 61\,k^2 - 116\,k - 52)\,\gamma^2$$
$$-\ (k+2)\,(2\,k^2 + 21\,k + 26)\,\gamma - 3\,(k+2)^2 = 0 \tag{6.9}$$

*and* $\gamma_d$ *is the only positive root of the cubic*

$$(-24 - 46\,k - 9\,k^2 + k^3)\,(72 + 122\,k + 27\,k^2 + k^3)\,\gamma^3 + (-5184 - 15072\,k - 13316\,k^2$$
$$-\ 4188\,k^3 - 437\,k^4 + k^6)\,\gamma^2 - 3\,(2+k)\,(4+k)\,(216 + 314\,k + 63\,k^2 + 2\,k^3)\,\gamma$$
$$- 27\,(2+k)^2\,(4+k)^2 = 0. \tag{6.10}$$

**Proof**

Consider the individual designs $\mathbf{t} = (t_1, t_2)$ be linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$, where $\mathbf{x}_c = (\frac{k}{2}, \frac{k}{2})$. Then the proposed optimal design $\xi_D^*$ can be written in the transformed coordinates as

$$\tilde{\xi}_D^* = \left\{ \begin{array}{ccc} (-\frac{k}{2}, 1) & (-1, \frac{k}{2}) & (-\frac{k}{2}, \frac{k}{2}) \\ w & w & 1 - 2\,w \end{array} \right\}.$$

The proof of the theorem is accomplished in two steps. The first step deals with the calculation of the weight $w$ and then the second step shows that the proposed design is optimal.

**Step 1. Calculation of the weight $w$**

It follows from $\mathbf{M}_\beta(\xi) = \sum_{i=1}^{r} w_i\, \mathbf{M}_\beta(\mathbf{t}_i)$ that at the population design $\tilde{\xi}_D^*$ with weight $w$ on each of the individual designs $(-\frac{k}{2}, 1)$ and $(-1, \frac{k}{2})$ and weight $1 - 2\,w$ on the individual design $(-\frac{k}{2}, \frac{k}{2})$, the standardized information matrix for $\boldsymbol{\beta}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}_D^*) = \frac{1}{4\,(1+2\gamma)} \begin{pmatrix} 4 & 0 & k^2 + w\,(4-k^2) \\ 0 & D_1 & 0 \\ k^2 + w\,(4-k^2) & 0 & D_2 \end{pmatrix} \tag{6.11}$$

where

$$D_1 = k^2\,(1 + 2\,\gamma) - (k-2)\,(2 + k + 2\,\gamma + 3\,k\,\gamma)\,w$$

and

$$D_2 = \frac{1}{4}\,\{k^4 + (k^2 - 4)\,[k^2\,(-1+\gamma) - 4\,(1+\gamma)]\,w\}.$$

The weight $w$ must be chosen to maximize the determinant of $\mathbf{M}_\beta(\tilde{\xi}_D^*)$, that is $w$ is chosen to maximize

$$\left|\mathbf{M}_\beta(\tilde{\xi}_D^*)\right| = \frac{1}{64\,(1+2\,\gamma)^3}\left\{(k^2-4)^2\,w\,(-1+w-\gamma)\,[-4\,k\,w\,\gamma\right.$$
$$\left.-4\,w\,(1+\gamma)+k^2\,(-1+w-2\,\gamma+3\,w\,\gamma)]\right\}. \qquad (6.12)$$

Taking the first derivative of the function (6.12) with respect to $w$ and equating it to zero gives the two solutions for $w$ as

$$w_1 = \frac{B-\sqrt{A}}{3\,(k-2)\,\{2+k+\gamma\,(3\,k+2)\}}$$

and

$$w_2 = \frac{B+\sqrt{A}}{3\,(k-2)\,\{2+k+\gamma\,(3\,k+2)\}}$$

where

$$A = 16+k^2\,(k^2-4)+2\,(2+k)\,(16-8\,k^2+3\,k^3)\,\gamma+(96+96\,k-60\,k^2-28\,k^3+15\,k^4)\,\gamma^2$$
$$+2\,(k-2)\,(2+3\,k)\,(-8-4\,k+3\,k^2)\,\gamma^3+(k-2)^2\,(2+3\,k)^2\,\gamma^4$$

and

$$B = (k-2)\,(2+3\,k)\,\gamma^2+2\,(3\,k^2-2\,k-4)\,\gamma+2\,(k^2-2).$$

The coefficients of $\gamma^m$, $m = 0, 1, 2, 3, 4$, in $A$ and $B$ are positive and are given in Appendix B.1. Thus $A$ and $B$ are positive.

Since

$$\left.\frac{\partial^2\left|\mathbf{M}_\beta(\tilde{\xi}_D^*)\right|}{\partial w^2}\right|_{w_1} = -\frac{(k^2-4)\,\sqrt{A}}{32\,(1+2\,\gamma)^3} < 0$$

and

$$\frac{\partial^2 \left|\mathbf{M}_\beta(\tilde{\xi}_D^*)\right|}{\partial w^2}\Bigg|_{w_2} = \frac{(k^2-4)\sqrt{A}}{32\,(1+2\,\gamma)^3} > 0$$

therefore $w_1$ maximizes $\left|\mathbf{M}_\beta(\tilde{\xi}_D^*)\right|$. Since

$$B^2 - A = 3\,(k-2)\,k^2\,(1+\gamma)\,(1+2\,\gamma)\,(2+k+2\,\gamma+3\,k\,\gamma) > 0$$

for $\gamma \geq 0$ and $k \geq 6$ it follows that $B - \sqrt{A} > 0$ and hence that the weight $w_1$ is strictly positive. Now

$$\frac{1}{2} - w_1 = \frac{2\sqrt{A} - B_1}{6\,(k-2)\,\{2+k+\gamma\,(2+3\,k)\}},$$

where

$$B_1 = 4 + k^2 + (2+k)\,(3\,k-2)\,\gamma + 2\,(k-2)\,(2+3\,k)\,\gamma^2.$$

Clearly $B_1 > 0$ for $\gamma \geq 0$ and $k \geq 6$. Consider now

$$4\,A - B_1^2 = 3\,(k-2)\,\{2+k+\gamma\,(2+3\,k)\}\,\{(k^2-4)+(k+2)\,(3\,k-10)\,\gamma+4\,(k^2-4\,k-4)\,\gamma^2\}.$$

The maximum root for $k^2-4\,k-4$ is 4.82843 and the function $k^2-4\,k-4$ is greater than zero for $k > 4.82843$ and thus for $k \geq 6$. Moreover, $3\,k-10 > 0$ for $k \geq 6$. It follows that $4\,A - B_1^2 > 0$ for $\gamma \geq 0$ and $k \geq 6$ and hence that $2\sqrt{A} - B_1 > 0$. Thus $0 < w_1 \leq \frac{1}{2}$.

Consider the derivative of $w_1$ with respect to $\gamma$, that is

$$\frac{\partial w_1}{\partial \gamma} = \frac{C_1\sqrt{A} - C_0}{3\,\{2+k+\gamma\,(2+3\,k)\}^2\sqrt{A}},$$

where

$$C_0 = 2\,(k^3-2\,k^2-8\,k-8) + 2\,(2+k)\,(3\,k^3+k^2-20\,k-16)\,\gamma + 3\,(2+k)\,(2+3\,k)$$
$$\times\,(3\,k^2-4\,k-8)\,\gamma^2 + (2+3\,k)^2\,(5\,k^2-4\,k-16)\,\gamma^3 + (k-2)\,(2+3\,k)^3\,\gamma^4$$

and

$$C_1 = 4\,(1+k) + 2\,(2+k)\,(2+3\,k)\,\gamma + (2+3\,k)^2\,\gamma^2.$$

Since the coefficients of $\gamma^m$, $m = 0, 1, 2, 3, 4$, in $C_0$ and $C_1$ are positive, it follows that $C_0$ and $C_1$ are positive as shown in Appendix B.2. Now since

$$C_1^2\,A - C_0^2 = 12\,k^2\,(1+\gamma)\,\{2+k+\gamma\,(2+3\,k)\}^3\,\{(k-2) + 2\,(2\,k^2-2)\,\gamma$$

$$+ (k-1)\,(2+3\,k)\,\gamma^2\} > 0$$

for $\gamma \geq 0$ and $k \geq 0$, it follows that $C_1\,\sqrt{A} - C_0 > 0$ for $k \geq 6$ and $\gamma \geq 0$ and thus that $w_1$ increasing monotonically with $\gamma$.

**Step 2. Optimality of the proposed design**

Assume that $(\tilde{t}_1, \tilde{t}_2) = \mathbf{x} = (x_1, x_2)$ with $x_j \in [0, \frac{k}{2}]$, $j = 1, 2$. Observe that the proposed optimal design $\tilde{\xi}_D^*$ is symmetric. Therefore by Lemma 6.4.1 the maximum of the directional derivative of $\Psi(\tilde{\xi}) = \ln|\mathbf{M}_\beta(\tilde{\xi})|$ at $\tilde{\xi}_D^*$ in the direction of $\mathbf{x}$, $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ occurs on that portion of the boundary of the design region defined by $x_1 = -\frac{k}{2}$ and $0 \leq x_2 \leq \frac{k}{2}$ or at the origin.

Consider first the derivative $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ at the origin. Substituting the weight $w_1$ for $w$ in (6.11) and inverting the resultant matrix gives the directional derivative $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ at the origin (0, 0) as

$$\phi((0,0), \tilde{\xi}_D^*) = \frac{H_1\,\sqrt{A} + H_0}{T},$$

where

$$H_0 = -6\,k^2\,\{6\,(2+k)^2\,(4+k^2) - (2+k)\,(-224 - 272\,k - 64\,k^2 - 28\,k^3 + 4\,k^4 + k^5)\,\gamma$$

$$+\ (832 + 2048\,k + 1392\,k^2 + 168\,k^3 - 76\,k^4 - 38\,k^5 - 3\,k^6)\,\gamma^2 + (2+k)\,(384 + 1056\,k$$

$$+ 576\,k^2 - 260\,k^3 - 92\,k^4 + 9\,k^5)\,\gamma^3 + (k-2)\,(2+3\,k)\,(-88 - 272\,k - 194\,k^2 + 2\,k^3$$

$$+ 15\,k^4)\,\gamma^4 + (k-2)^2\,(2+k)\,(2+3\,k)^2\,(2+5\,k)\,\gamma^5\}$$

and

$$H_1 = 6\,k^2\,\{-6\,(k+2)^2 + (2+k)\,(-32 - 44\,k + 4\,k^2 + k^3)\,\gamma + (-56 - 160\,k - 82\,k^2$$

$$+ 22\,k^3 + 9\,k^4)\,\gamma^2 + (k-2)\,(2+k)\,(2+3\,k)\,(2+5\,k)\,\gamma^3\}.$$

At $\gamma = 0$

$$\phi((0,0), \tilde{\xi}_D^*) = \frac{36\,k^2}{(2\,k^2 - 4 - \sqrt{16 - 4\,k^2 + k^4})\,(k^2 - 8 + \sqrt{16 - 4\,k^2 + k^4})} > 0$$

since $(2\,k^2 - 4)^2 - (16 - 4\,k^2 + k^4) = 3\,k^2\,(k^2 - 4) > 0$ for $k \geq 6$, implies that $2\,k^2 - 4 - \sqrt{16 - 4\,k^2 + k^4} > 0$. However, $\gamma$ in this theorem is a positive number and therefore this particular case does not affect the result. For any even integer $k \geq 6$

$$\lim_{\gamma \to \infty} \phi((0,0), \tilde{\xi}_D^*) = -2.$$

Further the derivative of $\phi((0,0), \tilde{\xi}_D^*)$ with respect to $\gamma$ has the form

$$\frac{\partial \phi((0,0), \tilde{\xi}_D^*)}{\partial \gamma} = \frac{\sum_{i=0}^{12} f_i(k)\,\gamma^i + \sum_{i=0}^{10} g_i(k)\,\gamma^i\,\sqrt{A}}{(2+k)^2\,\sqrt{A}\,T^2}$$

where $f_i(k)$ and $g_i(k)$ are polynomials in $k$. It is shown in Appendix B.3 that $f_i(k) < 0$, $i = 0, 1, \ldots, 12$, $g_i(k) < 0$, $i = 0, 1, \ldots, 5$ and $g_i(k) > 0$, $i = 6, 7, \ldots, 10$ for $k \geq 6$. Now

$$\left\{\sum_{i=6}^{10} g_i(k)\,\gamma^i\right\}^2 A - \left\{\sum_{i=6}^{12} f_i(k)\,\gamma^i\right\}^2 = \sum_{i=12}^{20} h_i(k)\,\gamma^i.$$

Since all $h_i(k) < 0$, $i = 12, \ldots, 20$ for $k \geq 6$ it follows that $\sum_{i=12}^{20} h_i(k)\,\gamma^i < 0$ for $k \geq 6$ and $\gamma > 0$ as shown in Appendix B.3. This implies that $\dfrac{\partial \phi((0,0), \tilde{\xi}_D^*)}{\partial \gamma} < 0$ for $\gamma > 0$. Thus $\phi((0,0), \tilde{\xi}_D^*) < 0$ for $\gamma > 0$.

Consider now the derivative $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ on the design region defined by $x_1 = -\frac{k}{2}$ and $0 \leq x_2 \leq \frac{k}{2}$. Substituting $w_1$ for $w$ in (6.11) and inverting the resultant matrix gives the directional derivative $\phi(\mathbf{x}, \tilde{\xi}_D^*)$ at a two-point design $(-\frac{k}{2}, x_2)$ as

$$\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) = \frac{1}{T} 6\,(2\,x_2 - k)\,(x_2 - 1)\,Q(x_2), \quad 0 \leq x_2 \leq \frac{k}{2} \tag{6.13}$$

where

$$T = (k+2)^2\,\{(4+k^2)+2\,(2+k)\,\gamma-(k-2)\,(2+3\,k)\,\gamma^2+\sqrt{A}\}\,\{B-\sqrt{A}\}\,\{\sqrt{A}+(k^2-8)$$
$$+\,2\,(-8-4\,k+3\,k^2)\,\gamma+2\,(k-2)\,(2+3\,k)\,\gamma^2\},$$

$$Q(x_2) = C_2\,x_2^2+C_1\,x_2+C_0,$$

with

$$C_0 = B_1\,\sqrt{A}+B_0,$$

$$C_1 = \frac{k+2}{2}\,C_2,$$

$$C_2 = 6\,(1+\gamma)\,(2+k+\gamma\,(2+3\,k)^2\,\{(k^2+4)+4\,(k+2)\,\gamma-(k-2)\,(2+3\,k)\,\gamma^2+\sqrt{A}\},$$

$$B_0 = 3\,k\,(2+k)^2\,(4+k^2)-k\,(2+k)\,(-104-132\,k-46\,k^2-21\,k^3+k^4)\,\gamma$$
$$-\,4\,k\,(-88-232\,k-194\,k^2-64\,k^3-7\,k^4+3\,k^5)\,\gamma^2-4\,k\,(-72-264\,k-294\,k^2$$
$$-\,92\,k^3+11\,k^4+9\,k^5)\,\gamma^3-k\,(2+3\,k)\,(-56-196\,k-106\,k^2+37\,k^3+12\,k^4)\,\gamma^4$$
$$-\,(k-2)\,k\,(2+3\,k)^2\,(2+9\,k+k^2)\,\gamma^5\}$$

and

$$B_1 = 3\,k\,(2+k)^2-k\,(2+k)\,(-14-21\,k+k^2)\,\gamma+k\,(20+68\,k+41\,k^2)\,\gamma^2$$
$$+\,k\,(2+3\,k)\,(2+9\,k+k^2)\,\gamma^3.$$

The denominator $T$ in expression (6.13) is positive (see Appendix B.2). Therefore the sign of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ takes the sign of $(2\,x_2 - k)\,(x_2 - 1)\,Q(x_2)$. Observe that $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ in (6.13) is quartic in $x_2$ and it can be easily seen that 1 and $\frac{k}{2}$ are the two zero points of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$. Since $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ is a quartic in $\tilde{t}_2$ it has four zero points not necessarily distinct. Solving the equation $Q(x_2) = 0$ the other two zero points of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ are

$$r_1 = -\frac{k+2}{4} + \frac{\sqrt{3}}{12}\sqrt{\frac{E_1\,\sqrt{A} + E_0}{(1+\gamma)\,\{2 + k + \gamma\,(2 + 3\,k)\}^2}}$$

and

$$r_2 = -\frac{k+2}{4} - \frac{\sqrt{3}}{12}\sqrt{\frac{E_1\,\sqrt{A} + E_0}{(1+\gamma)\,\{2 + k + \gamma\,(2 + 3\,k)\}^2}}$$

where

$$\begin{aligned} E_0 = {} & 3\,(k-2)^2\,(2+k)^2 + (2+k)\,(72 - 20\,k - 66\,k^2 + 13\,k^3)\,\gamma + (144 + 64\,k - 248\,k^2 \\ & - 128\,k^3 + 45\,k^4)\,\gamma^2 + (k-2)\,(2+3\,k)\,(17\,k^2 + 4\,k - 12)\,\gamma^3 \end{aligned}$$

and

$$E_1 = 16\,k\,(2+k)\,\gamma > 0.$$

It has been shown in Appendix B.4 that the coefficients of $\gamma^m$, $m = 0, 1, 2, 3$, in $E_0$ are positive so that $E_0 > 0$. Thus the expression in the square root of $r_1$ is positive. Clearly $r_2 < 0$ and thus $r_2$ does not fall in the interval $[0, \frac{k}{2}]$.

Consider now the root $r_1$. The root $r_1$ has the following properties

(i) When $\gamma = 0$, $r_1 = -1$.

(ii) $\lim_{\gamma \to \infty} r_1 = \dfrac{1}{4}\left\{-(k+2) + \sqrt{\dfrac{(k-2)\,(-4 + 12\,k + 11\,k^2)}{2 + 3\,k}}\right\}$.

(iii) $\dfrac{\partial r_1}{\partial \gamma} = \dfrac{\sqrt{3}\,k\,(k+2)\,(F_1\,\sqrt{A} + F_0)}{3\,(1+\gamma)^2\,\{2 + k + \gamma\,(1 + 3\,k)\}^3\,\sqrt{F}}$, where

$$F_0 = 2\,\{(2+k)\,(k^2\,(k^2 - 4) + 16)\} + 4\,(80 + 72\,k - 20\,k^2 - 24\,k^3 + 5\,k^4 + 3\,k^5)\,\gamma$$

$$+\ 4\,(160+208\,k-32\,k^2-88\,k^3+2\,k^4+9\,k^5)\,\gamma^2+2\,(160+192\,k-112\,k^2$$

$$-\,94\,k^3+33\,k^4)\,\gamma^3+8\,(2+3\,k)\,(20+12\,k-17\,k^2-5\,k^3+3\,k^4)\,\gamma^4$$

$$+4\,(k-2)\,(2+3\,k)^2\,(-2+k^2)\,\gamma^5\},$$

$$F_1=-(2+k)\,(4+k^2)+(-24-20\,k-6\,k^2+3\,k^3)\,\gamma+(-24-28\,k+6\,k^2+15\,k^3)\,\gamma^2$$

$$+\,(2+3\,k)\,(-4+5\,k^2)\,\gamma^3$$

and

$$F=\frac{E_1\,\sqrt{A}+E_0}{(1+\gamma)\,\{2+k+\gamma\,(2+3\,k)\}^2}.$$

All the coefficients of $\gamma^m$, $m=0,1,\ldots,5$, in $F_0$ are positive as shown in Appendix B.4. This implies that $F_0>0$. Observe that except for the constant term $-(2+k)^2\,(4+k^2)$ all the coefficients of $\gamma^m$, $m=1,2,3$ in $F_1$ are nonnegative. Note also that $A$ is a polynomial of degree 4 in $\gamma$.

Let

$$f_0(k)=2\,\sqrt{3}\,k\,(2+k)^2\,\{k^2\,(k^2-4)+16\},$$

$$f_1(k)=4\,\sqrt{3}\,k\,(2+k)\,(80+72\,k-20\,k^2-24\,k^3+5\,k^4+3\,k^5)$$

and

$$f_2(k)=4\,\sqrt{3}\,k\,(2+k)\,(160+208\,k-32\,k^2-88\,k^3+2\,k^4+9\,k^5).$$

Now

$$Z=\{f_0(k)+f_1(k)\,\gamma+2\,f_2(k)\,\gamma^2\}^2-\{\sqrt{3}\,k\,(2+k)^2\,(4+k^2)\,\sqrt{A}\}^2$$

$$=9\,(k-2)^2\,k^2\,(2+k)^6\,\{k^2\,(k^2-4)+16\}+6\,k^2\,(2+k)^3\,(9216+8192\,k-5376\,k^2$$

$$-\ 5568\,k^3+1920\,k^4+\ 1776\,k^5-336\,k^6-292\,k^7+36\,k^8+21\,k^9)\,\gamma$$

$$+ 3\,k^2\,(2+k)^2\,(178176 + 319488\,k + 41216\,k^2 - 199680\,k^3 - 62208\,k^4$$

$$+ 62976\,k^5 + 23328\,k^6 - 10368\,k^7 - 3832\,k^8 + 768\,k^9 + 273\,k^{10})\,\gamma^2$$

$$+ 6\,k^2\,(2+k)^2\,(202752 + 445440\,k + 144128\,k^2 - 278016\,k^3 - 155648\,k^4$$

$$+ 79104\,k^5 + 51296\,k^6 - 12192\,k^7 - 7512\,k^8 + 804\,k^9 + 423\,k^{10})\,\gamma^3$$

$$+ 3\,k^2\,(2+k)^2\,(672 + 880\,k - 128\,k^2 - 352\,k^3 + 6\,k^4 + 33\,k^5)\,(608 + 784\,k$$

$$- 128\,k^2 - 352\,k^3 + 10\,k^4 + 39\,k^5)\,\gamma^4.$$

Clearly the constant term $9\,(k-2)^2\,k^2\,(2+k)^6\,\{k^2\,(k^2-4)+16\}$ in $Z$ is greater than zero for $k \geq 6$. Further it has been shown in Appendix B.5 that the coefficients of $\gamma^m$, $m = 1, 2, 3, 4$ in $Z$ are positive for $k \geq 6$. Thus $Z > 0$ for $k \geq 6$ and $\gamma > 0$. This implies that

$$\{f_0(k) + f_1(k)\,\gamma + 2\,f_2(k)\,\gamma^2\} - \{\sqrt{3}\,k\,(2+k)^2\,(4+k^2)\,\sqrt{A}\} > 0$$

and hence that $\sqrt{3}\,k\,(2+k)\,(F_1\,\sqrt{A} + F_0) > 0$. Furthermore, the denominator in $\dfrac{\partial r_1}{\partial \gamma}$ is positive. Thus $\dfrac{\partial r_1}{\partial \gamma} > 0$. Overall, from (i), (ii) and (iii) it follows that the root $r_1$ is monotonically increasing with $\gamma$.

(iv) $$\frac{k}{2} - \lim_{\gamma\to\infty} r_1 = \frac{k}{2} - \frac{1}{4}\left\{-(k+2) + \sqrt{\frac{(k-2)\,(-4+12\,k+11\,k^2)}{2+3\,k}}\right\}$$

$$= \frac{1}{4}\left\{2 + 3\,k - \sqrt{\frac{(k-2)\,(-4+12\,k+11\,k^2)}{2+3\,k}}\right\}.$$

Since $(2+3\,k)^2 - \dfrac{(k-2)\,(-4+12\,k+11\,k^2)}{2+3\,k} = \dfrac{16\,k\,(2+k)^2}{2+3\,k} > 0$ for $k \geq 6$, $\dfrac{1}{4}\left\{2 + 3\,k - \sqrt{\dfrac{(k-2)\,(-4+12\,k+11\,k^2)}{2+3\,k}}\right\} > 0$ for $k \geq 6$. Thus $r_1 < \dfrac{k}{2}$.

Consider now the properties of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ for different values of $x_2 = r_1$. Recall that

$$\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) = \frac{6}{T}\,(2\,x_2 - k)\,(x_2 - 1)\,Q(x_2),$$

where $x_2 \in [0, \frac{k}{2}]$. Note that $\phi((-\frac{k}{2}, 1), \tilde{\xi}_D^*) = \phi((-\frac{k}{2}, \frac{k}{2}), \tilde{\xi}_D^*) = 0$. In fact these results are to be expected since $(-\frac{k}{2}, 1)$ and $(-\frac{k}{2}, \frac{k}{2})$ are the support designs of $\tilde{\xi}_D^*$. So the roots of $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ are $r_2$, 1, $r_1$ and $\frac{k}{2}$. The derivative $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ has one of the forms shown in Figures 6.5, 6.6 and 6.7.

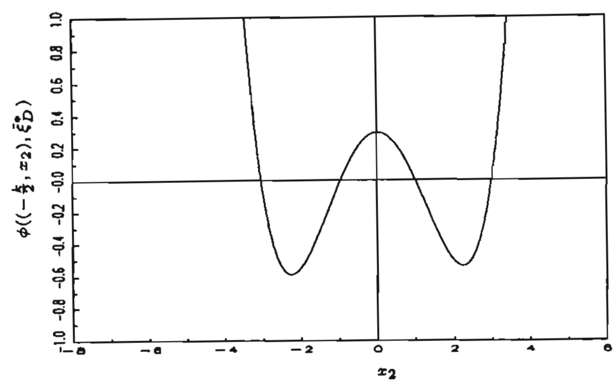


Figure 6.5: Plot of the directional derivative $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ against $x_2$ for $k = 6$ and $\gamma = 0.025$.

When $r_1 < 0$, the derivative $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) > 0$ as seen for example in Figure 6.5 and hence by Corollary 6.4.1, $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ needs to be examined at lattice points in the design region of interest. However, there is no lattice point in that design region except a point $(-\frac{k}{2}, 0)$ which lies on the boundary. Therefore for the inequality $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) \leq 0$ to hold it is necessary that $r_1 \geq 0$. In other words, the proposed design $\tilde{\xi}_D^*$ is optimal if and only if $r_1 \geq 0$. Since $r_1$ is monotonically increasing with $\gamma$ the proposed design $\tilde{\xi}_D^*$ is optimal if and only if $\gamma \geq \gamma_c$ where $\gamma_c$ is the only positive root of $r_1 = 0$, that is $\gamma_c$ is the positive root

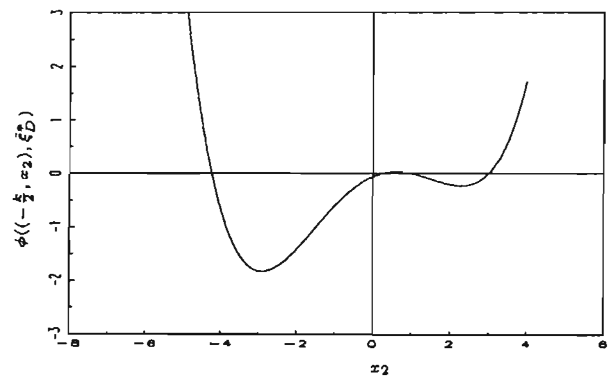


Figure 6.6: Plot of the directional derivative $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ against $x_2$ for $k = 6$ and $\gamma = 7.5$.

of

$$\frac{\sqrt{3}}{12}\sqrt{\frac{E_1\sqrt{A}+E_0}{(1+\gamma)\{2+k+\gamma(2+3k)\}^2}} = \frac{k+2}{4}.$$

The solution set for this equation is in that of

$$E_1\sqrt{A}+E_0 = (k+2)^2(1+\gamma)\{2+k+\gamma(2+3k)\}^2$$

which satisfies

$$E_1\sqrt{A} = (k+2)^2(1+\gamma)\{2+k+\gamma(2+3k)\}^2 - E_0.$$

This solution is a subset of the solution set of

$$E_1^2 A = \{(k+2)^2(1+\gamma)\{2+k+\gamma(2+3k)\}^2 - E_0\}^2.$$

Equivalently, the solution set of

$$3k^2(1+\gamma)\{2+k+\gamma(2+3k)\}^2\{(k^2-3k-6)(k^2+9k+2)\gamma^3+(k^4-61k^2$$

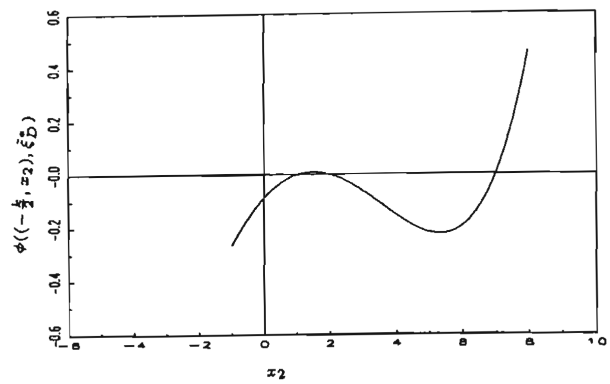


Figure 6.7: Plot of the directional derivative $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ against $x_2$ for $k = 14$ and $\gamma = 8.9$.

$$- \, 116\,k - 52)\,\gamma^2 - \,(k+2)\,(2\,k^2 + 21\,k + 26)\,\gamma - 3\,(k+2)^2\} = 0$$

is the solution set of the cubic

$$(k^2 - 3\,k - 6)\,(k^2 + 9\,k + 2)\,\gamma^3 + (k^4 - 61\,k^2 - 116\,k - 52)\,\gamma^2$$
$$- \,(k+2)\,(2\,k^2 + 21\,k + 26)\,\gamma - 3\,(k+2)^2 = 0.$$

In Figure 6.6, $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) > 0$ for $r_1 < x_2 \leq 1$. Therefore by Corollary 6.4.1 $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ needs to be examined at lattice points on the design region of interest. However, there is no lattice point in the region bounded by a ray $x_2 = -\frac{2}{k}\,x_1$, $x_1 = -\frac{k}{2}$ and $x_2 = 0$. Further $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) = 0$ at $x_2 = 1$ for all $\gamma \geq 0$. Note that $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) < 0$ for all design points $\mathbf{x} = (x_1, x_2)$ in the region bounded by a ray $x_2 = -\frac{2}{k}\,x_1$, $x_2 = x_1$ and $x_1 = -\frac{k}{2}$.

In Figure 6.7, $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*) > 0$ for $1 < r_1 < 2$. There are lattice points in the region bounded by the rays $x_2 = -\frac{2}{k}x_1$ and $x_2 = -\frac{4}{k}x_1$, and $x_1 = -\frac{k}{2}$. These lattice points have the form $(x_1, 1)$ such that $-\frac{k}{2} + 1 \leq x_1 \leq -2$ and thus lie on the line $x_2 = 1$. By Corollary 6.4.1 $\phi((-\frac{k}{2}, x_2), \tilde{\xi}_D^*)$ should be examined for maxima at the lattice points which are close to either to the boundary $x_2 = -\frac{k}{2}$ or to the origin. A lattice point $(-\frac{k}{2} + 1, 1)$ is the closest to the boundary $x_2 = -\frac{k}{2}$ and the lattice point $(-2, 1)$ is the closest to the origin. At $(-\frac{k}{2} + 1, 1)$ the directional derivative is given by

$$\phi((-\frac{k}{2} + 1, 1), \tilde{\xi}_D^*) = \frac{6\,(T_1\sqrt{A} + T_0)}{T}$$

where

$$\begin{aligned} T_0 = {} & -3\,(k-4)\,(k-1)\,k\,(2+k)^2\,(4+k^2) - (2+k)\,(256+448\,k-264\,k^2-484\,k^3 \\ & +90\,k^4 - 65\,k^5 + 25\,k^6)\,\gamma - 8\,(256 + 512\,k - 20\,k^2 - 572\,k^3 - 221\,k^4 + 56\,k^5 \\ & +10\,k^6 + 6\,k^7)\,\gamma^2 + 8\,(-384 - 912\,k - 84\,k^2 + 1036\,k^3 + 507\,k^4 - 201\,k^5 - 70\,k^6 \\ & +9\,k^7)\,\gamma^3 + (2+3\,k)\,(-1024 - 1440\,k + 1288\,k^2 + 1684\,k^3 - 414\,k^4 - 415\,k^5 + 105\,k^6)\,\gamma^4 \\ & +(k-2)\,(2+3\,k)^2\,(64 + 64\,k - 90\,k^2 - 49\,k^3 + 29\,k^4)\,\gamma^5 \end{aligned}$$

and

$$\begin{aligned} T_1 = {} & -3\,(k-4)\,(k-1)\,k\,(2+k)^2 - (2+k)\,(64+64\,k-46\,k^2-77\,k^3+25\,k^4)\,\gamma \\ & +(-256 - 464\,k + 84\,k^2 + 412\,k^3 + 89\,k^4 - 81\,k^5)\,\gamma^2 \\ & -(2+3\,k)\,(64 + 64\,k - 90\,k^2 - 49\,k^3 + 29\,k^4)\,\gamma^3. \end{aligned}$$

Since $T > 0$ the sign of $\phi((-\frac{k}{2} + 1, 1), \tilde{\xi}_D^*)$ depends on the sign of $T_1\sqrt{A} + T_0$. It has been shown in Appendix B.6 that the coefficients $T_0$ and $T_1$ are less than zero for $k \geq 6$ and

$\gamma > 0$ and hence that $\phi((-\frac{k}{2}+1,1),\tilde{\xi}_D^*) < 0$ for $k \geq 6$ and $\gamma > 0$. Similarly, at (−2,·1) the directional derivative is given by

$$\phi((-2,1),\tilde{\xi}_D^*) = \frac{6\,(k-4)\,(U_1\sqrt{A}+U_0)}{T}$$

where

$$\begin{aligned}U_0 = &-9\,(2+k)^2\,(4+k)\,(4+k^2) - (2+k)\,(1280+1784\,k+692\,k^2+358\,k^3+85\,k^4\\ &+8\,k^5+k^6)\,\gamma - (4480+11168\,k+8864\,k^2+3464\,k^3+1296\,k^4+348\,k^5\\ &+50\,k^6+3\,k^7)\,\gamma^2 + (-3840-11616\,k-11040\,k^2-3992\,k^3-1272\,k^4-452\,k^5\\ &-38\,k^6+9\,k^7)\,\gamma^3 + (2+3\,k)\,(-800-1560\,k-660\,k^2-98\,k^3-97\,k^4+32\,k^5+15\,k^6)\,\gamma^4\\ &+(k-2)\,(2+3\,k)^2\,(32+38\,k+17\,k^2+22\,k^3+5\,k^4)\,\gamma^5\end{aligned}$$

and

$$\begin{aligned}U_1 = &-9\,(2+k)^2\,(4+k) - (2+k)\,(176+266\,k+69\,k^2+8\,k^3+k^4)\,\gamma\\ &-(272+700\,k+584\,k^2+195\,k^3+58\,k^4+9\,k^5)\,\gamma^2\\ &-(2+3\,k)\,(32+38\,k+17\,k^2+22\,k^3+5\,k^4)\,\gamma^3.\end{aligned}$$

Since $T > 0$ the sign of $\phi((-2,1),\tilde{\xi}_D^*)$ takes the sign of $U_1\sqrt{A}+U_0$. It has been shown in Appendix B.7 that the coefficients $U_0$ and $U_1$ are less than zero for $k \geq 6$ and $\gamma > 0$ and hence that $\phi((-2,1),\tilde{\xi}_D^*) < 0$.

Consider now the case of $r_1 \geq 2$. Since $r_1$ is monotonically increasing with $\gamma$ the proposed design $\tilde{\xi}_D^*$ is optimal if and only if $\gamma \geq \gamma_d$, where $\gamma_d$ is the only positive root of $r_1 = 2$, that is $\gamma_d$ is the positive root of

$$\frac{\sqrt{3}}{12}\sqrt{\frac{E_1\sqrt{A}+E_0}{(1+\gamma)\,\{2+k+\gamma\,(2+3\,k)\}^2}} = \frac{k+10}{4}.$$

The solution set of this equation is in that of

$$E_1\sqrt{A}+E_0=(k+10)^2\,(1+\gamma)\,\{2+k+\gamma\,(2+3\,k)\}^2$$

which satisfies

$$E_1\sqrt{A}=(k+10)^2\,(1+\gamma)\,\{2+k+\gamma\,(2+3\,k)\}^2-E_0.$$

This solution is a subset of the solution set of

$$E_1^2\,A=\{(k+10)^2\,(1+\gamma)\,\{2+k+\gamma\,(2+3\,k)\}^2-E_0\}^2.$$

Equivalently, the solution set of

$$\begin{aligned}3\,(1+\gamma)\,\{2+k+\gamma\,(2+3\,k)\}^2\,\{(-24-46\,k-9\,k^2+k^3)\,(72+122\,k+27\,k^2+k^3)\,\gamma^3\\+(-5184-15072\,k-13316\,k^2-4188\,k^3-437\,k^4+k^6)\,\gamma^2\\-3\,(2+k)\,(4+k)\,(216+314\,k+63\,k^2+2\,k^3)\,\gamma-27\,(2+k)^2\,(4+k)^2\}=0\end{aligned}$$

is the solution set of the cubic equation

$$\begin{aligned}(-24-46\,k-9\,k^2+k^3)\,(72+122\,k+27\,k^2+k^3)\,\gamma^3\\+(-5184-15072\,k-13316\,k^2-4188\,k^3-437\,k^4+k^6)\,\gamma^2\\-3\,(2+k)\,(4+k)\,(216+314\,k+63\,k^2+2\,k^3)\,\gamma-27\,(2+k)^2\,(4+k)^2=0.\end{aligned}$$

□

**Theorem 6.4.3** *Consider the set of all two-point individual designs* $\mathbf{t}=(t_1,t_2)$ *which put equal weights on the distinct time points* $t_1$ *and* $t_2$ *with* $t_j\in\{0,1,\ldots,k\}, j=1,2$ *and* $0\le t_1<t_2\le k$ *for* $k$ *an even integer greater than or equal to 6. Then*

$$\xi_D^*=\left\{\begin{array}{ccccc}(0,\frac{k}{2}) & (0,\frac{k}{2}+1) & (0,k) & (\frac{k}{2}-1,k) & (\frac{k}{2},k)\\ w_1 & w_2 & 1-2\,w_1-2\,w_2 & w_2 & w_1\end{array}\right\},$$

*where*

$$w_1 = \frac{1}{32\,k^2\,\gamma^2}\{3\,(2+k)^2 + (2+k)\,(26+21\,k+2\,k^2)\,\gamma + (52+116\,k+61\,k^2-k^4)\,\gamma^2$$
$$-(-6-3\,k+k^2)\,(2+9\,k+k^2)\,\gamma^3\}$$

*and*

$$w_2 = \frac{1}{32\,(k-2)\,(2+k)\,\gamma^2}\{-3\,(2+k)^2 - (2+k)\,(42+21\,k+2\,k^2)\,\gamma +$$
$$(-180-180\,k-45\,k^2+k^4)\,\gamma^2 + (3+k)\,(6+k)\,(-6-3\,k+k^2)\,\gamma^3\}$$

*is the D-optimal population design for the fixed effects $\boldsymbol{\beta}$ in the model (6.1) over this set provided that $\gamma(k) < \gamma < \gamma_c$, where $\gamma(k)$ and $\gamma_c$ are as defined in Theorem 6.4.1 and Theorem 6.4.2, respectively.*

**Proof**

Consider the individual designs $\mathbf{t} = (t_1, t_2)$ be linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$, where $\mathbf{x}_c = (\frac{k}{2}, \frac{k}{2})$. Then the proposed optimal design $\xi_D^*$ can be written in the transformed coordinates as

$$\tilde{\xi}_D^* = \left\{ \begin{array}{ccccc} (-\frac{k}{2}, 0) & (-\frac{k}{2}, 1) & (-\frac{k}{2}, \frac{k}{2}) & (-1, \frac{k}{2}) & (0, \frac{k}{2}) \\ w_1 & w_2 & 1 - 2\,w_1 - 2\,w_2 & w_2 & w_1 \end{array} \right\}.$$

The information matrix for $\boldsymbol{\beta}$ at the population design $\tilde{\xi}_D^*$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}_D^*) = \frac{1}{4\,(1+2\gamma)} \begin{pmatrix} 4 & 0 & k^2\,(1-w_1-w_2)+4\,w_2 \\ 0 & D_1 & 0 \\ k^2\,(1-w_1-w_2)+4\,w_2 & 0 & D_2 \end{pmatrix} \qquad (6.14)$$

where

$$D_1 = k^2\,\{1 - w_1 - w_2 - \gamma\,(2 - 3\,w_1 - 3\,w_2)\} - 4\,w_2\,(k\,\gamma + \gamma + 1)$$

and

$$D_2 = k^4\,\{1 - w_1 - w_2 + \gamma\,(w_1 + w_2)\} - 8\,k^2\,w_2\,\gamma + 16\,w_2\,(1 + \gamma).$$

The wights $w_1$ and $w_2$ must be chosen to maximize $|\mathbf{M}_\beta(\tilde{\xi}_D^*)|$, that is to maximize

$$|\mathbf{M}_\beta(\tilde{\xi}_D^*)| = \frac{1}{64\,(1+2\,\gamma)^3}\{(k^2 - k^2\,w_1 + 4\,w_2 - k^2\,w_2 + 2\,k^2\,\gamma - 3\,k^2\,w_1\,\gamma + 4\,w_2\,\gamma$$

$$+4\,k\,w_2\,\gamma - 3\,k^2\,w_2\,\gamma)\,(k^4\,w_1 - k^4\,w_1^2 + 16\,w_2 - 8\,k^2\,w_2 + k^4\,w_2 + 8\,k^2\,w_1\,w_2$$

$$-2\,k^4\,w_1\,w_2 - 16\,w_2^2 + 8\,k^2\,w_2^2 - k^4\,w_2^2 + k^4\,w_1\,\gamma + 16\,w_2\,\gamma - 8\,k^2\,w_2\,\gamma + k^4\,w_2\,\gamma)\}.$$

Solving the equations

$$\frac{\partial|\mathbf{M}_\beta(\tilde{\xi}_D^*)|}{\partial w_1} = 0$$

and

$$\frac{\partial|\mathbf{M}_\beta(\tilde{\xi}_D^*)|}{\partial w_2} = 0$$

simultaneously yields the pairs of solutions for $w_1$ and $w_2$

(1) $$w_{11} = \frac{1}{32\,k^2\,\gamma^2}\{3\,(2+k)^2 + (2+k)\,(26 + 21\,k + 2\,k^2)\,\gamma + (52 + 116\,k + 61\,k^2 - k^4)\,\gamma^2$$

$$-(-6 - 3\,k + k^2)\,(2 + 9\,k + k^2)\,\gamma^3\}$$

and

$$w_{21} = \frac{1}{32\,(k-2)\,(2+k)\,\gamma^2}\{-3\,(2+k)^2 - (2+k)\,(42 + 21\,k + 2\,k^2)\,\gamma +$$

$$(-180 - 180\,k - 45\,k^2 + k^4)\,\gamma^2 + (3+k)\,(6+k)\,(-6 - 3\,k + k^2)\,\gamma^3\};$$

(2) $$w_{12} = \frac{\{(k+2) + \gamma\,(2 + 3\,k)\}\,\sqrt{A} + B_1}{8\,(k-2)\,k^2\,\gamma^2\,(1 + 3\,\gamma)}$$

and

$$w_{22} = \frac{1}{(k-2)^2\,\gamma^2}\,\{\sqrt{A} - (k-2)\,[(2+k) + (14+7\,k+k^2)\,\gamma + 15\,(2+k)\,\gamma^2$$
$$-\,3\,(-6-3\,k+k^2)\,\gamma^3]\},$$

and

$$(3)\quad w_{13} = \frac{\{(k+2)+\gamma\,(2+3\,k)\}\,\sqrt{A} + B_2}{8\,(k-2)\,k^2\,\gamma^2\,(1+3\,\gamma)}$$

and

$$w_{23} = \frac{1}{8\,(k-2)^2\,\gamma^2}\,\{-\sqrt{A} + (k-2)\,[-(2+k) - (14+7\,k+k^2)\,\gamma - 15\,(2+k)\,\gamma^2$$
$$+\,3\,(-6-3\,k+k^2)\,\gamma^3]\},$$

where

$$A = (k-2)^2\,(1+\gamma)\,(1+3\,\gamma)^2\,\{(k+2)^2 + (2+k)\,(14+7\,k+2\,k^2)\,\gamma + (60+60\,k$$
$$+\,15\,k^2+k^4)\,\gamma^2 + (-6-3\,k+k^2)^2\,\gamma^3\},$$

$$B_1 = (k-2)\,(1+3\,\gamma)\,\{-(k+2)^2 - (2+k)^2\,(5+k)\,\gamma - (28+44\,k+19\,k^2)\,\gamma^2$$
$$+\,(2+3\,k)\,(-6-3\,k+k^2)\,\gamma^3\}$$

and

$$B_2 = (k-2)^2\,(1+\gamma)\,(1+3\,\gamma)\,\{-(k+2)^2 - (2+k)^2\,(5+k)\,\gamma - (28+44\,k+19\,k^2)\,\gamma^2$$
$$+\,(2+3\,k)\,(-6-3\,k+k^2)\,\gamma^3\}.$$

The weights in (2) and (3) are not acceptable weights for $\gamma \in (\gamma(k), \gamma_c)$. For example for $k = 6$ and $\gamma = 2.05$ the weights in (2) and (3) give $\{w_{12} = 10.952, w_{22} = -13.446\}$ and $\{w_{13} = -5.099, w_{23} = 7.634\}$ respectively. Note that $w_{21} = 0$ for $\gamma = \dfrac{3\,(k+2)}{k^2-3\,k-6)}$ but $w_{21} < 0$ for $\gamma < \dfrac{3\,(k+2)}{k^2-3\,k-6)}$. Note also that $w_{11} = 0$ for $\gamma = \gamma_c$ and $w_{11} < 0$ for $\gamma > \gamma_c$.

Substituting $w_{11}$ and $w_{21}$ for $w_1$ and $w_2$, respectively in (6.14) and inverting the resultant matrix the directional derivative of $\Psi(\tilde{\xi}) = \ln|\mathbf{M}_\beta(\tilde{\xi})|$ at $\tilde{\xi}^*_D$ in the direction of the two-point design $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2) \in \tilde{S}_{2,k}$ is given by

$$\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_D) = \frac{8\,\gamma^2\,\{f_1(\tilde{t}_1, \tilde{t}_2) + f_2(\tilde{t}_1, \tilde{t}_2)\}}{(1+\gamma)\,F}$$

where

$$f_1(\tilde{t}_1, \tilde{t}_2) = 8\,(1+\gamma)\,(\tilde{t}_1^4 + \tilde{t}_2^4) + 2\,\{-4 - 2\,k - k^2 - \gamma\,(4 + 2\,k - k^2)\}\,(\tilde{t}_1^2 + \tilde{t}_2^2) - 16\,\gamma\,\tilde{t}_1^2\,\tilde{t}_2^2,$$

$$f_2(\tilde{t}_1, \tilde{t}_2) = -4\,(k+2)\,(1+\gamma)\,\tilde{t}_1\,\tilde{t}_2$$

and

$$F = (2+k)^2 + (2+k)\,(14 + 7\,k + 2\,k^2)\,\gamma + (60 + 60\,k + 15\,k^2 + k^4)\,\gamma^2 + (-6 - 3\,k + k^2)^2\,\gamma^3.$$

Since $\tilde{t}_i^2 \leq \dfrac{k^2}{4}$, $\tilde{t}_i^4 \leq \dfrac{k^4}{16}$, $i = 1, 2$ and $\tilde{t}_1\,\tilde{t}_2 \leq \dfrac{k^2}{4}$

$$f_1(\tilde{t}_1, \tilde{t}_2) \leq -k^2\,(k+2)\,(1+\gamma)$$

and

$$f_2(\tilde{t}_1, \tilde{t}_2) \leq k^2\,(k+2)\,(1+\gamma).$$

Thus

$$8\,\gamma^2\,\{f_1(\tilde{t}_1, \tilde{t}_2) + f_2(\tilde{t}_1, \tilde{t}_2)\} \leq 0.$$

Furthermore, $(1+\gamma)\,F > 0$ for $\gamma(k) \leq \gamma < \gamma_c$ and $k \geq 6$. Thus $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_D) \leq 0$ for $(\tilde{t}_1, \tilde{t}_2) \in \tilde{S}_{2,k}$ and equality holds at the support designs of $\tilde{\xi}^*_D$. □

**Remark**: Consider the optimum design in Theorem 6.4.2 for $k = 14$. The cubic in expression (6.10) has only one positive solution $\gamma_d = 8.45718$. The directional derivative at that optimal design, $\phi((-\frac{k}{2}, \tilde{t}_2), \tilde{\xi}^*_D)$, is positive for some $\tilde{t}_2 \in [-\frac{k}{2}, \frac{k}{2}]$ and for some $\gamma \geq \gamma_d$ when

$k = 14$. For example, for $k = 14$ and $\gamma = 8.9$ a plot of directional derivative is presented in Figure 6.7. The figure shows that $\phi((-\frac{k}{2}, \tilde{t}_2), \tilde{\xi}_D^*) > 0$ for $1 < \tilde{t}_2 < 2$. However, when $\tilde{t}_2 = 2$, $\phi((-\frac{k}{2}, \tilde{t}_2), \tilde{\xi}_D^*) = 0$. Therefore, when $k = 14$, for some $\gamma$ values such that $\gamma \geq \gamma_d$, the design $(-\frac{k}{2}, 2)$ is one of the support design of $D$-optimal population design. In general, it is believed that more optimal designs can be discovered for $k \geq 14$ and $\gamma \geq \gamma_d$ and these can proved following the steps in the proof of Theorem 6.4.2. However, the proofs will be algebraically longer and more complicated than that of Theorem 6.4.2. In such cases, the GAUSS program of Section 4.8 can be used to compute a two-point $D$-optimal population design for a given $\gamma$ numerically. For example, for $k = 14$ and $\gamma = 8.9$ the program yields the following $D$-optimal population design based on two-point individual designs

$$\tilde{\xi}_D^* = \left\{ \begin{array}{ccc} (-2, 7) & (-7, 7) & (-7, 2) \\ 0.4213 & 0.1574 & 0.4213 \end{array} \right\}.$$

In the following two theorems, the $D$-optimal population designs based on two-point individual designs when $k$ is an odd integer are presented for different values of $\gamma$. The proofs of the theorems are essentially the same as the proof of Theorem 6.4.2 and are therefore not given here.

Let $k_o = 2\,m + 1$, where $m$ is a positive integer greater than or equal to 1, that is, $k_o$ is an odd integer greater than or equal to 3. The $D$-optimum population design for the fixed effects $\boldsymbol{\beta}$ when $k$ an odd integer can be computed from Theorem 6.4.2 by substituting $k$ by $2\,k_o$ for some $\gamma$ values and are presented in the following theorems.

**Theorem 6.4.4** *Consider the set of all two-point individual designs* $\mathbf{t} = (t_1, t_2)$ *which put equal weights on the distinct time points* $t_1$ *and* $t_2$ *with* $t_j \in \{0, 1, \ldots, k_o\}, j = 1, 2$ *and*

$0 \leq t_1 < t_2 \leq k_o$ *for* $k_o$ *an odd integer greater than or equal to 3. Then*

$$\xi_D^* = \left\{ \begin{array}{ccc} (0, \frac{k_o+1}{2}) & (\frac{k_o-1}{2}, k_o) & (0, k_o) \\ w & w & 1-2w \end{array} \right\},$$

*where*

$$w = \frac{B - \sqrt{A}}{3\,(k-1)\,\{1 + k_o + \gamma\,(1 + 3\,k_o)\}},$$

$$A = 1 - k_o^2 + k_o^4 + 2\,(1 + k_o)\,(2 + k_o^2\,(3\,k_o - 4)\,\gamma + (6 + 12\,k_o - 15\,k_o^2 - 14\,k_o^3 + 15\,k_o^4)\,\gamma^2$$
$$+2\,(k_o - 1)\,(1 + 3\,k_o)\,(-2 - 2\,k_o + 3\,k_o^2)\,\gamma^3 + (k_o - 1)^2\,(1 + 3\,k_o)^2\,\gamma^4$$

*and*

$$B = -1 + 2\,k_o^2 + 2\,(3\,k_o^2 - k_o - 1)\,\gamma + (k_o - 1)\,(3\,k_o + 1)\,\gamma^2$$

*is the D-optimal population design for the fixed effects* $\boldsymbol{\beta}$ *in the model (6.1) over this set when (i)* $k_o \leq 7$ *and for all* $\gamma \geq 0$ *and (ii)* $k_o \geq 9$ *provided that* $0 \leq \gamma < \gamma_r$ *where* $\gamma_r$ *is the only positive root of the cubic*

$$(-6 - 21\,k_o - 6\,k_o^2 + k_o^3)\,(18 + 59\,k_o + 18\,k_o^2 + k_o^3)\,\gamma^3 + (-324 - 1752\,k_o - 2719\,k_o^2$$
$$-1288\,k_o^3 - 190\,k_o^4 + k_o^6)\,\gamma^2 - 4\,(1 + k_o)\,(3 + k_o)\,(27 + 73\,k_o + 21\,k_o^2 + k_o^3)\,\gamma$$
$$-12\,(1 + k_o)^2\,(3 + k_o)^2 = 0.$$

**Theorem 6.4.5** *Consider the set of all two-point individual designs* $\mathbf{t} = (t_1, t_2)$ *which put equal weights on the distinct time points* $t_1$ *and* $t_2$ *with* $t_j \in \{0, 1, \ldots, k_o\}, j = 1, 2$ *and* $0 \leq t_1 < t_2 \leq k_o$ *for* $k_o$ *an odd integer greater than or equal to 9. Then*

$$\xi_D^* = \left\{ \begin{array}{ccc} (0, \frac{k_o+3}{2}) & (\frac{k_o-3}{2}, k_o) & (0, k_o) \\ w & w & 1-2w \end{array} \right\},$$

*where*

$$w = \frac{B - \sqrt{A}}{3\,(k_o - 3)^3\,(3 + k_o)^2\,\{3 + k_o + 3\,\gamma\,(1 + k_o)\}},$$

$$A = \frac{1}{(k_o - 1)^4\,(1 + k_o)^4}\{(k_o - 3)^4\,(3 + k_o)^4\,[81 - 9\,k_o^2 + k_o^4 + 6\,(3 + k_o)\,(18 - 4\,k_o^2 + k_o^3)\,\gamma$$

$$+\ 3\,(162 + 108\,k_o - 45\,k_o^2 - 14\,k_o^3 + 5\,k^4)\,\gamma^2 + 18\,(k_o - 3)\,(1 + k_o)\,(-6 - 2\,k_o + k_o^2)\,\gamma^3$$

$$+9\,(k_o - 3)^2\,(1 + k_o)^2]\}$$

*and*

$$B = (k_o - 3)^2\,(3 + k_o)^2\,\{81 - 9\,k_o^2 + k_o^4 + 6\,(3 + k_o)\,(18 - 4\,k_o^2 + k_o^3)\,\gamma + 3\,(162 + 108\,k_o$$

$$-45\,k_o^2 - 14\,k_o^3 + 5\,k_o^4)\,\gamma^2\}$$

*is the D-optimal population design for the fixed effects $\boldsymbol{\beta}$ in the model (6.1) over this set provided that $\gamma \geq \gamma_s$ where $\gamma_s$ is the only positive root of the cubic*

$$(-6 - 21\,k_o - 6\,k_o^2 + k_o^3)\,(18 + 27\,k_o + 18\,k_o^2 + k_o^3)\,\gamma^3 + (-324 - 1368\,k_o - 1791\,k_o^2$$

$$-\ 1032\,k_o^3 - 222\,k_o^4 + k_o^6)\,\gamma^2 - 4\,(1 + k_o)\,(3 + k_o)\,(27 + 57\,k_o + 21\,k_o^2 + k_o^3)\,\gamma$$

$$-\ 12\,(1 + k_o)^2\,(3 + k_o)^2 = 0.$$

Furthermore, numerical computations show that in the set of all two-point individual designs $\mathbf{t} = (t_1, t_2)$ which put equal weights on the distinct time points $t_1$ and $t_2$ with $t_j \in \{0, 1, \ldots, k_o\}, j = 1, 2$ and $0 \leq t_1 < t_2 \leq k_o$ for $k_o$ an odd integer greater than or equal to 9, the design

$$\xi_D^* = \left\{ \begin{array}{ccccc} (0, \frac{k_o+1}{2}) & (0, \frac{k_o+3}{2}) & (0, k_o) & (\frac{k_o-3}{2}, k_o) & (\frac{k_o-1}{2}, k_o) \\ w_1 & w_2 & 1 - 2\,w_1 - 2\,w_2 & w_2 & w_1 \end{array} \right\},$$

is the $D$-optimal population design for the fixed effects $\boldsymbol{\beta}$ over this set provided that $\gamma \in (\gamma_r, \gamma_s)$. The design weights $w_1$ and $w_2$ such that $w_1 + w_2 < \frac{1}{2}$ are obtained numerically for a given $k_o$ and $\gamma \in (\gamma_r, \gamma_s)$.

## 6.5 Optimal design for estimation of linear and quadratic coefficients

In this section, $D_s$-optimal population design for precise estimation of the linear and quadratic coefficients $\beta_1$ and $\beta_2$ in model (6.1) is discussed. A $D_s$-optimal population design minimizes the determinant of the submatrix $\mathbf{V}_{22}$ of $\mathbf{M}_\beta^{-1}(\xi)$ that corresponds to the linear and quadratic coefficients. Suppose the information matrix for $\boldsymbol{\beta}$, $\mathbf{M}_\beta(\xi)$, at the population design

$$\xi = \begin{Bmatrix} \mathbf{t}_1, & \ldots, & \mathbf{t}_r \\ w_1, & \ldots, & w_r \end{Bmatrix} \text{ with } 0 < w_i < 1 \text{ and } \sum_{i=1}^{r} w_i = 1$$

be defined in block as

$$\mathbf{M}_\beta(\xi) = \begin{pmatrix} \mathbf{M}_{11}(\xi) & \mathbf{M}_{12}(\xi) \\ \mathbf{M}_{12}'(\xi) & \mathbf{M}_{22}(\xi) \end{pmatrix}$$

where $\mathbf{M}_{11}(\xi)$ is a submatrix corresponding to the intercept $\beta_o$. The submatrix $\mathbf{V}_{22}$ is proportional to the inverse of $\mathbf{M}_{22}(\xi) - \mathbf{M}_{12}'(\xi)\mathbf{M}_{11}^{-1}(\xi)\mathbf{M}_{12}(\xi)$. Thus, the design $\xi_{D_s}^*$ is called a $D_s$-optimal design if

$$\left|\mathbf{M}_\beta(\xi_{D_s}^*)\right| = \max_\xi \left|\mathbf{M}_{22}(\xi) - \mathbf{M}_{12}'(\xi)\mathbf{M}_{11}^{-1}(\xi)\mathbf{M}_{12}\right| = \max_\xi \frac{|\mathbf{M}_\beta(\xi)|}{|\mathbf{M}_{11}(\xi)|},$$

among all designs $\xi$ defined on the space of designs of interest, where $|\mathbf{M}_\beta(\xi)| \neq 0$ and $|\mathbf{M}_{11}(\xi)| \neq 0$.

The General Equivalence Theorem for $D_s$-optimal designs of Atkinson and Donev (1992, page 110) can be modified to accommodate population designs, following the approach in Subsection 2.6.4. Thus the design $\xi^*_{D_s}$ is $D_s$-optimal population design if and only if

$$\phi_s(\mathbf{t}, \xi^*_{D_s}) = tr\{\mathbf{M}_\beta^{-1}(\xi^*_{D_s})\mathbf{M}_\beta(\mathbf{t})\} - tr\{\mathbf{M}_{11}^{-1}(\xi^*_{D_s})\mathbf{M}_{11}(\mathbf{t})\} \leq 2. \tag{6.15}$$

for all $\mathbf{t}$ in the space of designs of interest, with equality holding at the support designs of $\xi^*_{D_s}$ where $\mathbf{M}_\beta(\mathbf{t})$ is the standardized information matrix for $\boldsymbol{\beta}$ at $\mathbf{t}$ and $\mathbf{M}_{11}(\mathbf{t}) = \dfrac{d}{d(1+d\gamma)}$, i.e. the (1, 1)th element of $\mathbf{M}_\beta(\mathbf{t})$.

For a quadratic model with a random intercept

$$\mathbf{M}_{11}(\xi^*_{D_s}) = \mathbf{M}_{11}(\mathbf{t})$$

therefore (6.15) yields

$$\phi_s(\mathbf{t}, \xi^*_{D_s}) = tr[\mathbf{M}_\beta^{-1}(\xi^*_{D_s})\mathbf{M}_\beta(\mathbf{t})] \leq 2 + tr[\mathbf{M}_{11}^{-1}(\xi^*_{D_s})\mathbf{M}_{11}(\mathbf{t})] = 3.$$

Thus the $D$-optimal population designs derived in the previous sections are also optimal for the precise estimation of $\beta_1$ and $\beta_2$ in model (6.1).

## 6.6 $D$-optimal population design based on $d$-point individual designs

The algebraic derivation of the optimum designs described Sections 6.3 and 6.4 were only based on one-and two-point individual designs. It is our future research plan to extend this approach to designs based on $d$-point individuals designs with $d \geq 3$.

The GAUSS program "doptinte" of Section 6.4 can be used to compute a $D$-optimal population design based on the set of $d$-point individual designs for $d \in [3, k+1]$ for given $k$ and $\gamma$ numerically. The program can also be used to calculate the best $D$-optimal population design. The best $D$-optimal population design for the fixed effects $\boldsymbol{\beta}$ in model (6.1) for $k$ even, however, can be obtained from Atkins and Cheng (1999). Since the $D$-optimal population design based on the set of one-point individual designs puts equal weights of $\frac{1}{3}$ on the design points 0, $\frac{k}{2}$ and $k$ and therefore corresponds to the exact design comprising three points 0, $\frac{k}{2}$ and $k$. It thus follows from Atkins and Cheng (1999) that the best $D$-optimal population design for $\boldsymbol{\beta}$ is the design which put a weight 1 on $(0, \frac{k}{2}, k)$. This is now demonstrated in the following theorem.

**Theorem 6.6.1** *Consider the set of population designs based on all possible individual designs* $\mathbf{t}$ *which put equal weights on the distinct time points* $t_1, t_2, \ldots, t_d$, $j = 1, \ldots, d$ *and* $0 \leq t_1 < t_2 < \ldots < t_d \leq k$ *for* $d$ *a positive integer less than* $k+1$. *Then*

$$\xi^*_{D_b} = \left\{ \begin{array}{c} (0, \frac{k}{2}, k) \\ 1 \end{array} \right\}$$

*is the* $D$*-optimal population design for the fixed effects* $\boldsymbol{\beta}$ *in model (6.1) over this set for any* $\gamma \geq 0$ *when* $k$ *is an even integer.*

**Proof**

Consider the individual designs $\mathbf{t}$ to be linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$, where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$. Then the proposed optimum design $\xi^*_{D_b}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{D_b} = \left\{ \begin{array}{c} (-\frac{k}{2}, 0, \frac{k}{2}) \\ 1 \end{array} \right\}.$$

Note immediately that the information matrix for $\boldsymbol{\beta}$ at $\tilde{\xi}^*_{D_b}$ is equal to

$$\mathbf{M}_\beta(\tilde{\xi}^*_{D_b}) = \frac{1}{6\,(1+3\,\gamma)} \begin{pmatrix} 6 & 0 & k^2 \\ 0 & (1+3\,\gamma)\,k^2 & 0 \\ k^2 & 0 & \frac{1}{4}\,k^4\,(1+\gamma) \end{pmatrix}$$

and hence that

$$\mathbf{M}^{-1}_\beta(\tilde{\xi}^*_{D_b}) = \begin{pmatrix} 3\,(1+\gamma) & 0 & -\dfrac{12}{k^2} \\ 0 & \dfrac{6}{k^2} & 0 \\ -\dfrac{12}{k^2} & 0 & \dfrac{72}{k^4} \end{pmatrix}.$$

Recall from expression (6.2) that the information matrix for $\boldsymbol{\beta}$ at a $d$-point design $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_d)$ is given by

$$\mathbf{M}_\beta(\tilde{\mathbf{t}}) = \frac{1}{d\,(1+d\,\gamma)} \begin{pmatrix} d & \sum_{j=1}^d \tilde{t}_j & \sum_{j=1}^d \tilde{t}_j^2 \\ \sum_{j=1}^d \tilde{t}_j & A_1 & A_2 \\ \sum_{j=1}^d \tilde{t}_j^2 & A_2 & A_3 \end{pmatrix}$$

where

$$A_1 = (1+d\,\gamma) \sum_{j=1}^d \tilde{t}_j^2 - \gamma \left( \sum_{j=1}^d \tilde{t}_j \right)^2,$$

$$A_2 = (1+d\,\gamma) \sum_{j=1}^d \tilde{t}_j^3 - \gamma \left( \sum_{j=1}^d \tilde{t}_j \right) \left( \sum_{j=1}^d \tilde{t}_j^2 \right)$$

and

$$A_3 = (1+d\,\gamma) \sum_{j=1}^d \tilde{t}_j^4 - \gamma \left( \sum_{j=1}^d \tilde{t}_j^2 \right)^2.$$

Then using the identity

$$\left(\sum_{j=1}^{d} \tilde{t}_j\right)^2 = d \sum_{j=1}^{d} \tilde{t}_j^2 - \sum_{j<j'}^{d} (\tilde{t}_j - \tilde{t}_{j'})^2$$

the directional derivative of $\Psi(\tilde{\xi}) = \ln |\mathbf{M}_\beta(\tilde{\xi})|$ at $\tilde{\xi}^*_{D_b}$ in the direction of $\tilde{\mathbf{t}} \in \tilde{S}_{d,k}$ is given by

$$\begin{aligned}\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_b}) &= tr[\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{D_b})\mathbf{M}_\beta(\tilde{\mathbf{t}})] - 3 \\ &= \frac{1}{k^4\, d\,(1 + d\,\gamma)} \left\{ -3\, d\,(d-1)\, k^4\, \gamma - 18\, k^2 \sum_{j=1}^{d} \tilde{t}_j^2 + 72 \sum_{j=1}^{d} \tilde{t}_j^4 + 6\, k^2\, \gamma \sum_{j<j'}^{d} (\tilde{t}_j - \tilde{t}_{j'})^2 \right. \\ &\qquad \left. + 72\, \gamma \sum_{j<j'}^{d} (\tilde{t}_j^2 - \tilde{t}_{j'}^2)^2 \right\}.\end{aligned}$$

Since for $-\frac{k}{2} \le \tilde{t}_j < \tilde{t}_{j'} \le \frac{k}{2}$

$$\sum_{j<j'}^{d} (\tilde{t}_j^2 - \tilde{t}_{j'}^2)^2 \le \frac{d\,(d-1)\, k^4}{32},$$

$$\sum_{j<j'}^{d} (\tilde{t}_j - \tilde{t}_{j'})^2 \le \frac{d\,(d-1)\, k^4}{8}$$

and

$$\sum_{j=1}^{d} \tilde{t}_j^4 \le \frac{1}{4} k^2 \sum_{j=1}^{d} \tilde{t}_j^2$$

it then follows that

$$72\, \gamma \sum_{j<j'}^{d} (\tilde{t}_j^2 - \tilde{t}_{j'}^2)^2 \le \frac{9}{4}\, d\,(d-1)\, k^4\, \gamma,$$

$$6\, k^2\, \gamma \sum_{j<j'}^{d} (\tilde{t}_j - \tilde{t}_{j'})^2 \le \frac{3}{4}\, d\,(d-1)\, k^4\, \gamma$$

for $\gamma \ge 0$ and

$$72 \sum_{j=1}^{d} \tilde{t}_j^4 \le 18\, k^2 \sum_{j=1}^{d} \tilde{t}_j^2.$$

Thus $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_b}) \le 0$. Moreover, equality holds at the support design of $\tilde{\xi}^*_{D_b}$. □

**Example 6.6.1** Consider the quadratic regression model with a random intercept as specified by model (6.1) and let $k = 5$. The best $D$-optimum population designs for $\boldsymbol{\beta}$, calculated numerically using the GAUSS program for $\gamma = 0.05$ and $\gamma = 2.5$ are given by

$$\xi^*_{D_{b1}} = \left\{ \begin{array}{ccc} (0,3) & (0,5) & (2,5) \\ 0.3433 & 0.3134 & 0.3433 \end{array} \right\}$$

and by

$$\xi^*_{D_{b2}} = \left\{ \begin{array}{ccccc} (0,3) & (0,5) & (2,5) & (0,2,5) & (0,3,5) \\ 0.1825 & 0.0660 & 0.1825 & 0.2845 & 0.2845 \end{array} \right\}$$

respectively. Observe that the best $D$-optimal population designs change with $\gamma$ for odd $k$. The plots of the directional derivatives $\phi_D(\mathbf{t}, \xi^*_{D_{b1}})$ and $\phi_D(\mathbf{t}, \xi^*_{D_{b2}})$ against the individual designs $\mathbf{t} \in S_{d,5}$ where $1 \leq d \leq 6$, i.e. against the designs (0), (1), ..., (5), (0, 1), ..., (1, 2, 3, 4, 5, 6) labelled 1 through 63, are presented in Figures 6.8 and 6.9. The figures show that there are three maxima for $\phi_D(\mathbf{t}, \xi^*_{D_{b1}})$ and five maxima for $\phi(\mathbf{t}, \xi^*_{D_{b2}})$ at the design points which are equal to zero. These maxima occur at the support designs of $\xi^*_{D_{bi}}, i = 1, 2$ and thus the designs are optimum.

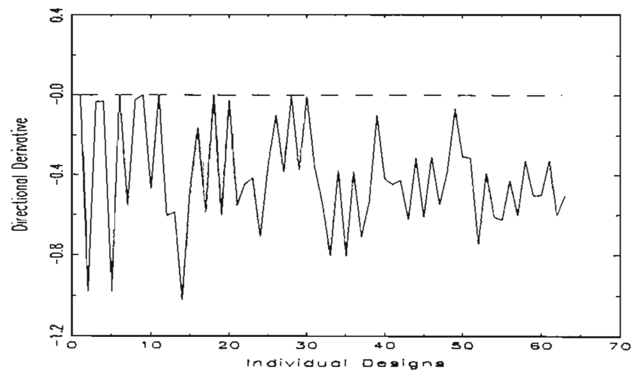


Figure 6.8: Plot of the directional derivative $\phi_D(\mathbf{t}, \xi^*_{D_{b1}})$ against the individual design $\mathbf{t}$.

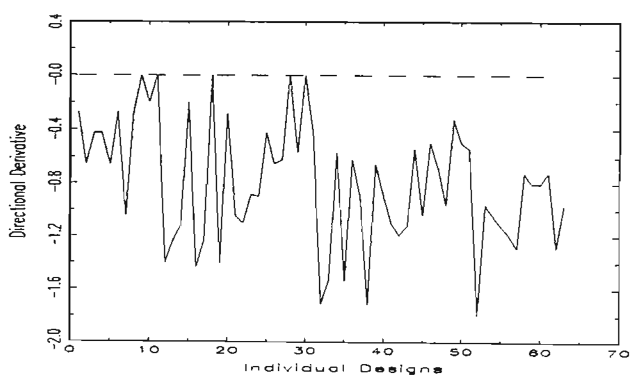


Figure 6.9: Plot of the directional derivative $\phi_D(\mathbf{t}, \xi^*_{D_{b2}})$ against the individual design $\mathbf{t}$.

# Chapter 7

# $V$-optimal Population Designs for the Quadratic Regression Model with a Random Intercept

## 7.1 Introduction

In this Chapter the problem of constructing $V$-optimal population designs for the quadratic regression model with a random intercept is considered.

Recall from Subsection 3.3.1 that the matrix form of the quadratic regression model with a random intercept is given by

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{1}\, b_i + \mathbf{e}_i \tag{7.1}$$

where $\mathbf{y}_i$ is a $d_i \times 1$ vector of observations for the $i$th individual at time points $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{id_i})'$, $\mathbf{X}_i = (\mathbf{1}\ \mathbf{t}_i\ \mathbf{t}_i^{(2)})$, where $\mathbf{t}_i^{(2)}$ is a column vector with elements equal to the squares of the time points, $\boldsymbol{\beta} = (\beta_o, \beta_1, \beta_2)'$, $b_i$ is a random intercept for the $i$th individual and $\mathbf{e}_i$ is a random error vector, $i = 1, \ldots, K$. Further it is assumed that $b_i \sim \mathcal{N}(0, \sigma_b^2)$, that $\mathbf{e}_i \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbf{I}_{d_i})$, and that $b_i$ and the elements of $\mathbf{e}_i$ are independent within and between individuals. Under these assumptions the population mean response for the individual $i$ is equal to $E(\mathbf{y}_i) = \mathbf{X}_i \boldsymbol{\beta}$ where $i = 1, \ldots .K$.

Then, as in Chapter 5, the design problem is to estimate the population mean responses $\boldsymbol{\mu}_g$ at a given vector of time points $\mathbf{t}_g$, where the elements of $\mathbf{t}_g$ are taken from the set $\{0, 1, \ldots, k\}$, as precisely as possible. The population mean response at $\mathbf{t}_g$, $\boldsymbol{\mu}_g$, is equal to $\mathbf{X}_g \boldsymbol{\beta}$ where $\mathbf{X}_g = (1 \;\; \mathbf{t}_g \;\; \mathbf{t}_g^{(2)})$ and $\mathbf{t}_g^{(2)}$ is a column vector with elements equal to the square of the time points in $\mathbf{t}_g$. The maximum likelihood estimator of $\boldsymbol{\mu}_g$ is $\mathbf{X}_g \hat{\boldsymbol{\beta}}$ with asymptotic variance based on the population design $\xi$ equal to $\mathbf{X}_g \mathbf{M}_\beta^{-1}(\xi) \mathbf{X}_g'$. Therefore the $V$-optimal population design for estimation of $\boldsymbol{\mu}_g$ minimizes the criterion function

$$\Psi_V(\xi) = tr \left\{ \mathbf{M}_\beta^{-1}(\xi) \mathbf{X}_g' \mathbf{X}_g \right\}$$

over the set of population designs defined on the space of designs of interest. Furthermore, by the Equivalence Theorem of Theorem 2.6.3 the population design $\xi_V^*$ is $V$-optimal if and only if

$$\phi_V(\mathbf{t}, \xi_V^*) = tr\{\mathbf{M}_\beta^{-1}(\xi_V^*) \mathbf{X}_g' \mathbf{X}_g \mathbf{M}_\beta^{-1}(\xi_V^*) \mathbf{M}_\beta(\mathbf{t})\} - tr\{\mathbf{M}_\beta^{-1}(\xi_V^*) \mathbf{X}_g' \mathbf{X}_g\} \leq 0 \qquad (7.2)$$

for all individual designs $\mathbf{t}$ in the space of designs of interest where $\mathbf{M}_\beta(\mathbf{t})$ is the standardized information matrix for $\boldsymbol{\beta}$ given in expression (6.2). Note that equality in (7.2) holds at the support designs of $\xi_V^*$. In this chapter it is assumed that $\mathbf{t}_g = (0, 1, \ldots, k)'$ unless stated otherwise and thus that

$$\mathbf{X}_g = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & k & k^2 \end{pmatrix}.$$

Note that $V$-optimal population designs for the linear random intercept model are invariant to a linear transformation of the design points. It is thus convenient to consider the individual designs $\mathbf{t}$ linearly transformed according to $\tilde{\mathbf{t}} = \mathbf{t} - \mathbf{x}_c$, where $\mathbf{x}_c = (\frac{k}{2}, \ldots, \frac{k}{2})$. Then the matrix $\mathbf{X}_g = (\mathbf{1}\ \mathbf{t}_g\ \mathbf{t}_g^{(2)})$ can be written in the transformed coordinates as

$$\tilde{\mathbf{X}}_g = (\mathbf{1}\ \tilde{\mathbf{t}}_g\ \tilde{\mathbf{t}}_g^{(2)}) = \begin{pmatrix} 1 & -\frac{k}{2} & \frac{k^2}{4} \\ 1 & -\frac{k}{2}+1 & \frac{k^2}{4}-k+1 \\ \vdots & \vdots & \vdots \\ 1 & \frac{k}{2} & \frac{k^2}{4} \end{pmatrix}$$

and thus

$$\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g = \begin{pmatrix} k+1 & 0 & B_1 \\ 0 & B_1 & 0 \\ B_1 & 0 & B_2 \end{pmatrix} \tag{7.3}$$

where $B_1 = \frac{1}{12}k(k+2)(k+1)$ and $B_2 = \frac{1}{240}k(k+2)(k+1)(3k^2+6k-4)$.

This Chapter is organized as follows. The $V$-optimal population designs based on one- and two-point individual designs are discussed in Sections 7.2 and 7.3 respectively. The one- and two-point $V$-optimal population designs are identical for designs with non-repeated and repeated time points and therefore designs in the set $S_{d,k}$ need only be considered in these sections. Finally, in Section 7.4 the $V$-optimal population designs based on $d$-point individual designs for $d \geq 3$ are discussed and selected examples constructed numerically.

## 7.2 $V$-optimal population designs based on one-point individual designs

Consider now the equally spaced time points, $0, 1, 2, \ldots, k$, where $k$ is a positive integer greater than or equal to 2. Then the space of one-point individual designs consists of these $k+1$ time points. Let $t \in \{0, 1, \ldots, k\}$ be a single time point. Then it follows from expression (6.2) that the standardized information matrix for $\boldsymbol{\beta}$ at $t$ on a per observation basis is given by

$$\mathbf{M}_{\beta}(t) = \frac{1}{1+\gamma}\begin{pmatrix} 1 & t & t^2 \\ t & t^2 & t^3 \\ t^2 & t^3 & t^4 \end{pmatrix}.$$

The term $(1+\gamma)$ factors out of this expression and hence out of $\mathbf{M}_{\beta}(\xi)$ where $\xi$ is a one-point population design. Clearly it follows that the term $(1+\gamma)$ factors out of the $V$-optimal criterion

$$\Psi_V(\xi) = tr\{\mathbf{M}_{\beta}^{-1}(\xi)\,\mathbf{X}_g'\mathbf{X}_g\}.$$

Therefore the $V$-optimal population designs based on one-point individual designs for the quadratic random intercept model do not depend on the variance ratio $\gamma$.

For $k$ even the $V$-optimal population design based on one-point individual designs is presented in the following theorem.

**Theorem 7.2.1** *Consider the set of all one-point designs $t \in \{0, 1, \ldots, k\}$ where $k$ is an even integer greater than or equal to 2. Then*

$$\xi_{V_1}^{*} = \left\{\begin{array}{ccc} (0) & (\frac{k}{2}) & (k) \\ w & 1-2w & w \end{array}\right\}$$

*where*

$$w = \frac{(k+2)\,(4\,k^2+3\,k-2) - 2\,\sqrt{B}}{30\,k^2}$$

*with*

$$B = (k-1)\,(2+k)\,(1+k^2)\,(4\,k^2+3\,k-2)$$

*is the $V$-optimal population design for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ in model (7.1) over this set for all $\gamma \geq 0$.*

**Proof**

Note that $V$-optimal population designs for the random intercept models are invariant to linear transformation of design points. Thus let a one-point individual design $t$ be linearly transformed according to $\tilde{t} = t - \frac{k}{2}$. Then the space of designs of interest in the transformed coordinates is given by

$$\tilde{S}_{1,k} = \{\tilde{t} : \tilde{t} \in \{-\frac{k}{2}, -\frac{k}{2}+1, \ldots, \frac{k}{2}-1, \frac{k}{2}\}$$

and the proposed optimum design $\xi^*_{V_1}$ can be written in the transformed coordinates as

$$\tilde{\xi}^*_{V_1} = \left\{ \begin{array}{ccc} (-\frac{k}{2}) & (0) & (\frac{k}{2}) \\ w & 1-2\,w & w \end{array} \right\}.$$

The proof of the theorem is accomplished in two steps. The first step deals with the calculation of the weight $w$ and the second step shows that the proposed design is optimal.

**Step 1. Calculation of the weight $w$**

Note immediately that the standardized information matrix for $\boldsymbol{\beta}$ at the design $\tilde{\xi}^*_{V_1}$ is given by

$$\mathbf{M}_\beta(\tilde{\xi}^*_{V_1}) = \frac{1}{(1+\gamma)} \begin{pmatrix} 1 & 0 & \dfrac{k^2 w}{2} \\ 0 & \dfrac{k^2 w}{2} & 0 \\ \dfrac{k^2 w}{2} & 0 & \dfrac{k^4 w}{8} \end{pmatrix}$$

and hence that

$$\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_1}) = (1+\gamma) \begin{pmatrix} \dfrac{1}{1-2w} & 0 & -\dfrac{4}{k^2(1-2w)} \\ 0 & \dfrac{2}{k^2 w} & 0 \\ -\dfrac{4}{k^2(1-2w)} & 0 & \dfrac{8}{k^2 w(1-2w)} \end{pmatrix}.$$

It then follows from expression (7.3) and $\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_1})$ that the $V$-optimality criterion $\Psi_V(\tilde{\xi})$ evaluated at $\tilde{\xi}^*_{V_1}$ is given by

$$\Psi_V(\tilde{\xi}^*_{V_1}) = \frac{(1+k)\{(k+2)(4k^2+3k-2) - 30k^2 w\}(1+\gamma)}{15k^3 w(1-2w)}.$$

The weight $w$ is necessarily in the interval $[0, \frac{1}{2}]$ and must be chosen to minimize $\Psi_V(\tilde{\xi}^*_{V_1})$. The equation

$$\frac{\partial \Psi_V(\tilde{\xi}^*_{V_1})}{\partial w} = \frac{(1+k)\{4 - 16w + 4k(k^2+1)(4w-1) - k^2(11 - 44w + 60w^2)\}(1+\gamma)}{15k^3(1-2w)^2 w^2} = 0$$

yields the solutions

$$w_1 = \frac{(k+2)(4k^2+3k-2) - 2\sqrt{B}}{30k^2}$$

and

$$w_2 = \frac{(k+2)(4k^2+3k-2) + 2\sqrt{B}}{30k^2}$$

where $B = (k-1)(2+k)(1+k^2)(-2+3k+4k^2)$.

Consider the solution $w_2$. Since

$$w_2 - \frac{1}{2} = \frac{2\,(k-1)\,(k^2+1) + \sqrt{B}}{30\,k^2} > 0$$

for $k \geq 2$ it follows that $w_2 > \frac{1}{2}$ and thus that $w_2$ is not an acceptable weight.

Consider now the solution $w_1$. Observe that $(k+2)\,(4\,k^2+3\,k-2)$ and $B$ are greater than zero for $k \geq 2$. Thus

$$(k+2)^2\,(4\,k^2+3\,k-2)^2 - 4\,B = 15\,k^2\,(k+2)\,(4\,k^2+3\,k-2) > 0$$

implies that $(k+2)\,(4\,k^2+3\,k-2) - 2\,\sqrt{B} > 0$ and that $w_1 > 0$. Consider also the difference

$$w_1 - \frac{1}{2} = \frac{2\,(k-1)\,(k^2+1) - \sqrt{B}}{15\,k^2}.$$

For $k \geq 2$ it follows that $(k-1)\,(k^2+1) > 0$. Therefore

$$(k-1)^2\,(k^2+1)^2 - B = -15\,(k-1)\,k^2\,(1+k^2) < 0$$

for $k \geq 2$ and this implies that $(k-1)\,(k^2+1) - \sqrt{B}$ is less than zero. Thus $w_1 < \frac{1}{2}$.

Further, consider the second derivative

$$\left.\frac{\partial^2 \Psi_V(\tilde{\xi}^*_{V_1})}{\partial w^2}\right|_{w=w_1} =$$

$$\frac{8\,(k+2)\,(k^2-1)\,(k^2+1)\,(4\,k^2+3\,k-2)\,(1+\gamma)\,(8\,k^3+7\,k^2+8\,k-8-4\,\sqrt{B})}{3375\,w_1^3\,(1-2\,w_1)^3}.$$

Since $0 \leq w_1 < \frac{1}{2}$ for $k \geq 2$ the sign of $\left.\frac{\partial^2 \Psi_V(\tilde{\xi}^*_{V_1})}{\partial w^2}\right|_{w=w_1}$ depends on the sign of $8\,k^3+7\,k^2+8\,k-8-4\,\sqrt{B}$. However, since $8\,k^3+7\,k^2+8\,k-8 > 0$ for $k \geq 2$

$$(8\,k^3+7\,k^2+8\,k-8)^2 - 16\,B = 225\,k^4 > 0$$

implies that $8\,k^3 + 7\,k^2 + 8\,k - 8 - 4\sqrt{B} > 0$ and hence that $\left.\dfrac{\partial^2 \Psi_V(\tilde{\xi}^*_{V_1})}{\partial w^2}\right|_{w=w_1} > 0$. Thus $w = w_1$ minimizes $\Psi_V(\tilde{\xi}^*_{V_1})$.

**Step 2. Optimality of the proposed design**

By substituting $w_1$ for $w$ in $\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_1})$, it follows that the directional derivative of the criterion $\Psi_V(\tilde{\xi}_V) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}_V)\,\tilde{\mathbf{X}}_g'\,\tilde{\mathbf{X}}_g\}$ at $\tilde{\xi}^*_{V_1}$ in the direction of a one-point design $\tilde{t}$ is given by

$$\phi_V(\tilde{t}, \tilde{\xi}^*_{V_1}) = -\frac{E_1}{E_2}\,\tilde{t}^2\,(k - 2\,\tilde{t})\,(k + 2\,\tilde{t})(1 + \gamma)$$

where

$$E_1 = 900\,(k^2 - 1)\,(2 + k)\,\{(k^2 + 1)\,(8\,k^2 + 11\,k - 4) - 2\,(1 + 2\,k)\,\sqrt{B}\}$$

and

$$E_2 = \{2\,(1 - k + k^2 - k^3) + \sqrt{B}\}^2\,\{4 - 4\,k - 11\,k^2 - 4\,k^3 + 2\,\sqrt{B}\}^2.$$

Consider the terms in $\phi_V(\tilde{t}, \tilde{\xi}^*_{V_1})$. It is clear that $E_2 > 0$ and that $\tilde{t}^2\,(k - 2\,\tilde{t})\,(k + 2\,\tilde{t}) \geq 0$ for $|\tilde{t}| \leq \frac{k}{2}$ and $k \geq 2$. Further, since for $k \geq 2$ both $(k^2 + 1)\,(8\,k^2 + 11\,k - 4)$ and $2\,(1 + 2\,k)\,\sqrt{B}$ are positive it follows that

$$(k^2 + 1)^2\,(8\,k^2 + 11\,k - 4)^2 - 4\,(1 + 2\,k)^2\,B = 15\,k\,(1 + k^2)\,(-8 + 11\,k + 20\,k^2 + 7\,k^3) > 0.$$

This implies that $(k^2 + 1)\,(8\,k^2 + 11\,k - 4) - 2\,(1 + 2\,k)\,\sqrt{B} > 0$ and hence that $E_1 > 0$. Thus $\phi_V(\tilde{t}, \tilde{\xi}^*_{V_1}) \leq 0$ for $|\tilde{t}| \leq \frac{k}{2}$ and $\gamma \geq 0$. Observe also that equality holds at $\tilde{t} = -\frac{k}{2}, 0$ and $\frac{k}{2}$. Thus by Theorem 2.6.3 the design $\tilde{\xi}^*_{V_1}$ and hence $\xi^*_{V_1}$ is the $V$-optimal population design for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ in model (7.1) over the set of one-point individual designs $t \in \{0, 1, \ldots, k\}$ for all $\gamma \geq 0$. □

The plot of the optimal weight $w$ of Theorem 7.2.1 against $k$ is presented in Figure 7.1. The graph shows that as $k \to \infty$ the weight $w$ approaches $\dfrac{1}{4}$. Thus, as $k \to \infty$, the $V$-optimal population design $\xi^*_{V_1}$ converges to a design which puts 25% of the weight at the designs with extreme time points 0 and $k$, and the remaining 50% at the design with time point $\dfrac{k}{2}$. Note that for $k = 2$ the weight $w = \frac{1}{3}$.

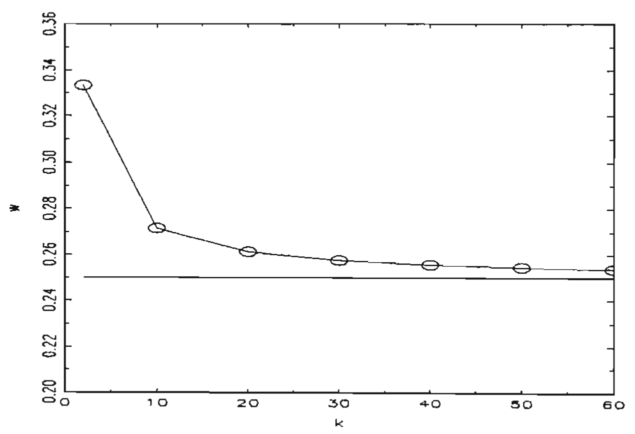


Figure 7.1: Plot of the design weight $w$ in Theorem 7.2.1 against $k$.

When $k$ is an odd integer greater than or equal to 3 the $V$-optimal population designs based on one-point individual designs were computed numerically for various values of $k$ and $\gamma$. The Equivalence Theorem in conjunction with numerical calculations showed that the design of the form

$$\xi^*_{V_o} = \left\{ \begin{array}{cccc} (0) & (\frac{k-1}{2}) & (\frac{k+1}{2}) & (k) \\ w & \frac{1}{2} - w & \frac{1}{2} - w & w \end{array} \right\}$$

where $0 < w < \frac{1}{2}$, is the $V$-optimal population design for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ over the set of one-point individual designs $t \in \{0, 1, \ldots, k\}$ for all $\gamma \geq 0$. Further the weight $w$ is chosen to minimize

$$\Psi_V(\tilde{\xi}^*_{V_o}) = 4\,(1+\gamma)\,\frac{G_1}{G_2}$$

where

$$\tilde{\xi}^*_{V_o} = \left\{ \begin{array}{cccc} (-\frac{k}{2}) & (\frac{1}{2}) & (\frac{1}{2}) & (\frac{k}{2}) \\ w & \frac{1}{2}-w & \frac{1}{2}-w & w \end{array} \right\}$$

is the design $\xi^*_{V_o}$ in the linearly transformed coordinates,

$$G_1 = 15\,(k^2-1)\,(2\,k-1)\,(k+1)^2\,w^2 + k\,(k^2-1)\,(4\,k^2-15\,k-1)\,w - 2\,k^3\,(k^2+1)$$

and

$$G_2 = 15\,(k^2-1)\,w\,(1-2\,w)\,\{k^2\,(1-2\,w)+2\,w\}.$$

Solving the equation

$$\frac{\partial \Psi_V(\tilde{\xi}^*_{V_o})}{\partial w} = 4\,(1+\gamma)\,\frac{G_3}{G_4} = 0$$

where

$$G_3 = 60\,(k-1)^2\,(k+1)^3\,(2\,k-1)\,w^4 + 8\,(k-1)^2\,k\,(k+1)^2\,(4\,k^2-15\,k-1)\,w^3$$

$$-(k-1)\,k\,(k+1)\,(40\,k^4-30\,k^3+27\,k^2+15\,k+2)\,w^2 + 8\,k^3\,(k^2+1)\,(2\,k^2-1)\,w - 8\,k^5\,(k^2+1)$$

and

$$G_4 = 15\,(k^2-1)\,w^2\,(1-2\,w)^2\,\{k^2\,(1-2\,w)+2\,w\}^2,$$

which is a quartic in $w$, the design weight $w$ that minimizes $\Psi_V(\tilde{\xi}^*_{V_o})$ is given by the following expression

$$w = A_0 + \frac{1}{2}\,\{-\sqrt{B_1} + \sqrt{B_2 - B_3}\}$$

where

$$A_0 = -\frac{k\,(-1-15\,k+4\,k^2)}{30\,(1+k)\,(2\,k-1)},$$

$$B_1 = A_1+\frac{A_2}{A_3\,(A_4+\sqrt{A_5})^{1/3}}+\frac{(A_4+\sqrt{A_5})^{1/3}}{180\,\times\,2^{\frac{1}{3}}\,(k-1)^2\,(1+k)^3\,(2\,k-1)},$$

$$B_2 = 2\,A_1-\frac{A_2}{A_3\,(A_4+\sqrt{A_5})^{1/3}}-\frac{(A_4+\sqrt{A_5})^{1/3}}{180\,\times\,2^{\frac{1}{3}}\,(k-1)^2\,(1+k)^3\,(2\,k-1)},$$

and

$$B_3 = \frac{A_7}{4\,\sqrt{B_1}}$$

with

$$A_1 = \frac{k\,(-10-57\,k-43\,k^2+46\,k^3+174\,k^4+128\,k^5+32\,k^6)}{450\,(k-1)\,(1+k)^2\,(2\,k-1)^2},$$

$$A_2 = (k-1)^4\,k^2\,(1+k)^5\,(2+k)^2\,(2\,k-1)^3\,(1+9\,k+4\,k^2)^2,$$

$$A_3 = 90\,\times\,2^{\frac{2}{3}}\,(k-1)^4\,(1+k)^6\,(2\,k-1)^2,$$

$$\begin{aligned} A_4 = {} & 2\,(k-1)^2\,k^3\,(1+k)^3\,(2+k)\,(4+84\,k+417\,k^2-33540\,k^3+20520\,k^4+15150\,k^5 \\ & +\;48855\,k^6+9930\,k^7-14340\,k^8+5880\,k^9+11424\,k^{10}+4224\,k^{11}+512\,k^{12}), \end{aligned}$$

$$\begin{aligned} A_5 = {} & 17280\,(k-1)^4\,k^9\,(1+k)^6\,(2+k)^2\,(3+k)\,(1+k^2)\,(-5+6\,k+3\,k^2)\,(4+84\,k \\ & +\;417\,k^2-17340\,k^3+6480\,k^4+15150\,k^5+31575\,k^6-6270\,k^7-17580\,k^8+5880\,k^9 \\ & +\;11424\,k^{10}+4224\,k^{11}+512\,k^{12}), \end{aligned}$$

$$A_6 = -30-2389\,k+5333\,k^2-2791\,k^3+489\,k^4-684\,k^5-216\,k^6+2224\,k^7+1408\,k^8+256\,k^9$$

and

$$A_7 = -\frac{2\,k^2\,A_6}{3375\,(k-1)^2\,(1+k)^3\,(2\,k-1)^3}.$$

The plot of the weight $w$ in $\xi^*_{V_o}$ against $k$ is presented in Figure 7.2. As for the $k$ even case, as $k \to \infty$ the design weight $w$ approaches $\frac{1}{4}$. Thus, as $k \to \infty$, $\xi^*_{V_o}$ converges to a design which puts equal weights at the designs corresponding to the two extreme points 0 and $k$, and the two middle points $(\frac{k-1}{2})$ and $(\frac{k+1}{2})$.

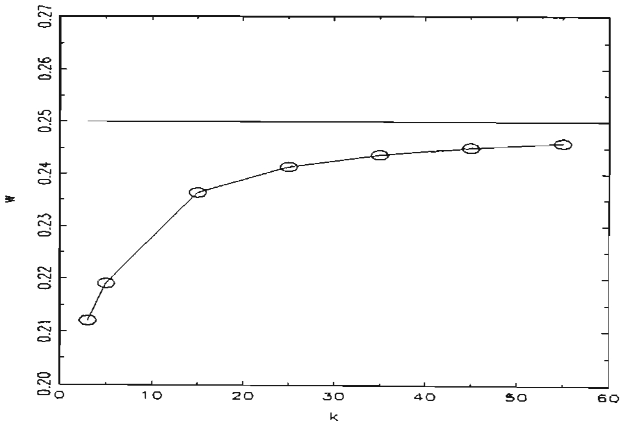


Figure 7.2: Plot of design weight $w$ of the $V$-optimal population design $\xi^*_{V_o}$ against $k$.

**Example 7.2.1** Consider the quadratic regression model with a random intercept as specified by model (7.1). Then for $k = 5$ and $k = 9$ the designs

$$\tilde{\xi}^*_{V_o} = \left\{ \begin{array}{cccc} (0) & (2) & (3) & (5) \\ 0.2809 & 0.2191 & 0.2191 & 0.2809 \end{array} \right\}$$

and

$$\tilde{\xi}^*_{V_o} = \left\{ \begin{array}{cccc} (0) & (4) & (5) & (9) \\ 0.2707 & 0.2293 & 0.2293 & 0.2707 \end{array} \right\}$$

respectively are proposed as the $V$-optimal population designs over the set of one-point individual designs for all $\gamma \geq 0$. The plots of the directional derivatives $\phi_V(\tilde{t}, \tilde{\xi}^*_{V_o})$ against the individual designs $t = 0, 1, 2, 3, 4, 5$ for $k = 5$, and against the individual designs $t = 0, 1, \ldots, 9$ for $k = 9$ are presented in Figures 7.3 and 7.4 respectively. The figures show that the condition $\phi_V(\tilde{t}, \tilde{\xi}^*_{V_o}) \leq 0$ holds for all single point design $\tilde{t}$ in the design space of interest and further equality holds at the support points of $\tilde{\xi}^*_{V_o}$. Thus the designs appear to be $V$-optimal.

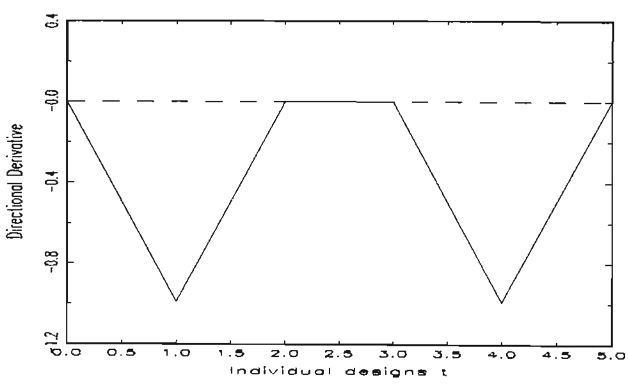


Figure 7.3: Plot of the directional derivative $\phi(\mathbf{t}, \xi^*_{V_o})$ against the individual design ($t$) for $k = 5$.

In the second part of their study, Abt, Gaffke, Liski and Sinha (1998) constructed optimal designs for prediction under the quadratic growth model with a random intercept. Their numerical results indicate that for $0.6 < \rho < 1$ the design

$$\tilde{\xi}^*_1 = \left\{ \begin{array}{ccc} (-\frac{k}{2}) & (0) & (\frac{k}{2}) \\ w_1 & w_2 & w_3 \end{array} \right\} \quad \text{with } 0 \leq w_i < 1 \text{ and } \sum_{i=1}^{3} w_i = 1$$

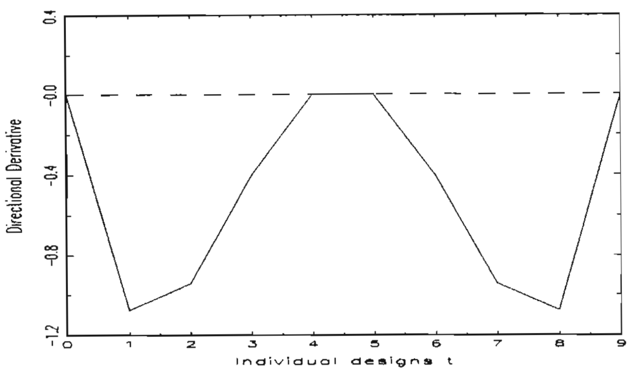


Figure 7.4: Plot of the directional derivative $\phi(\mathbf{t}, \xi^*_{V_o})$ against the individual design $(t)$ for $k = 9$.

is optimal for growth prediction, whereas for $0 \leq \rho \leq 0.6$ the design

$$\tilde{\xi}^*_2 = \left\{ \begin{array}{cc} (0) & (\frac{k}{2}) \\ w_4 & w_5 \end{array} \right\} \quad \text{with} \quad 0 < w_i < 1 \;\text{ and }\; w_4 + w_5 = 1$$

is optimal for growth prediction, where $\rho$ is the intra-class correlation $\frac{\gamma}{1+\gamma}$. In Table 7.1 some findings from Abt *et al.* (1998) are compared with the optimum design of Theorem 7.2.1. The optimum weights $w_i$, $i = 1, 2, 3$ in $\tilde{\xi}^*_1$ are from Table 2 of Abt *et al.* (1998), whereas $w$ is from Theorem 7.2.1. The criteria values of both designs $\Psi_V(\tilde{\xi}^*_1)$ and $\Psi_V(\tilde{\xi}^*_{V_1})$ are also given in Table 7.1. The criterion value $\Psi_V(\tilde{\xi}^*_{V_1})$ is less than $\Psi_V(\tilde{\xi}^*_1)$ in all cases. Thus the optimum design of Theorem 7.2.1 is more efficient than that of Abt *et al.* (1998).

Table 7.1: Optimal weights and criteria values for some combinations of $k$ and $\rho$ from Abt *et al.* (1998) and Theorem 7.2.1

| $k$ | $\rho$ | $w_1$ | $w_2$ | $w_3$ | $w$ | $\Psi_V(\tilde{\xi}_1^*)$ | $\Psi_V(\tilde{\xi}_{V_1}^*)$ |
|---|---|---|---|---|---|---|---|
| 10 | 0.75 | 0.1881 | 0.4045 | 0.4074 | 0.2715 | 111.27 | 102.18 |
| | 0.9 | 0.2305 | 0.4173 | 0.3522 | 0.2715 | 262.96 | 255.42 |
| 20 | 0.75 | 0.1356 | 0.4432 | 0.4212 | 0.2612 | 220.73 | 187.09 |
| | 0.9 | 0.2107 | 0.4485 | 0.3408 | 0.2612 | 482.96 | 467.72 |
| 40 | 0.75 | 0.1051 | 0.4636 | 0.4314 | 0.2557 | 460.76 | 357.55 |
| | 0.9 | 0.2001 | 0.4647 | 0.3351 | 0.2557 | 925.67 | 893.86 |
| 60 | 0.75 | 0.0940 | 0.4705 | 0.4355 | 0.2539 | 712.26 | 528.14 |
| | 0.9 | 0.1965 | 0.4702 | 0.3333 | 0.2539 | 1368.83 | 1320.35 |

## 7.3 $V$-optimal population designs based on two-point individual designs

In the previous section it was observed that the design weight for the $V$-optimal population design based on the set of one-point individual designs for $k$ odd could only be found numerically and that an algebraic proof of the optimality of the design was not possible. The weights of $V$-optimal population designs based on the set of two-point individual designs are of a similar nature. Therefore in this section the $V$-optimal designs based on two-point individual designs are investigated numerically. First, the numerical results for the $k$ even case are discussed and then those for $k$ odd follow.

Consider the set of all two-point individual designs $\mathbf{t} = (t_1, t_2)$ which put equal weights on the distinct time points $t_1$ and $t_2$ with $t_j \in \{0, 1, \ldots, k\}, j = 1, 2$ and $0 \leq t_1 < t_2 \leq k$ for

$k$ an even integer greater than or equal to 2. The $V$-optimal population designs over this set were computed numerically for various values of $k$ and $\gamma$. The Equivalence Theorem in conjunction with numerical calculations showed that the design of the form

$$\xi^*_{V_e} = \left\{ \begin{array}{ccc} (0, \frac{k}{2}) & (0, k) & (\frac{k}{2}, k) \\ w & 1-2w & w \end{array} \right\} \tag{7.4}$$

is the $V$-optimal population design for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ over this set for all $\gamma \geq 0$. Further the weight $w$ is chosen to minimize

$$\Psi_V(\tilde{\xi}^*_{V_e}) = tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_e})\,\tilde{\mathbf{X}}'_g\tilde{\mathbf{X}}_g\} = -(1+k)\,(1+2\,\gamma)\frac{H_1}{H_2}$$

where

$$\tilde{\xi}^*_{V_e} = \left\{ \begin{array}{ccc} (-\frac{k}{2}, 0) & (-\frac{k}{2}, \frac{k}{2}) & (0, \frac{k}{2}) \\ w & 1-2w & w \end{array} \right\}$$

is the design $\xi^*_{V_e}$ in the linearly transformed coordinates,

$$H_1 = 15\,k^2\,(1{+}4\,\gamma{+}3\,k\,\gamma^2)\,w^2 - 2\{15\,k^3\,\gamma^2 + (12-12\,k+37\,k^2-7\,k^3)\,\gamma$$

$$+\,(4-4\,k+19\,k^2-4\,k^3)\}\,w - 8\,(k-1)\,(1+k^2)\,(1+2\,\gamma)$$

and

$$H_2 = 15\,k^3\,w\,\{-(2{+}6\,\gamma{+}3\,\gamma^2)\,w + (1{+}\gamma)\,(1{+}2\,\gamma)\}.$$

For fixed $k$ and $\gamma$, the optimal weight is a positive number less than $\frac{1}{2}$ and is a solution of the equation

$$\frac{\partial tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_e})\,\tilde{\mathbf{X}}'_g\tilde{\mathbf{X}}_g\}}{\partial w} = (1+k)\,(1+2\,\gamma)\frac{H_3}{H_4} = 0 \tag{7.5}$$

where

$$H_3 = 15\,k^2\,(1{+}3\,\gamma)(2{+}4\,\gamma{+}3\,k\,\gamma^2)\,w^4 - 4\,(1{+}3\,\gamma)\{15\,k^3\,\gamma^2 + (12-12\,k+37\,k^2$$

$$- 7\,k^3)\,\gamma + 4 - 4\,k + 19\,k^2 - 4\,k^3\}\,w^3 + \{3\,(24 - 24\,k + 34\,k^2 + k^3)\,\gamma^3 - 3\,(-104 + 104\,k$$

$$- 154\,k^2 + 79\,k^3)\,\gamma^2 - 2\,(-108 + 108\,k - 173\,k^2 + 98\,k^3)\,\gamma - 10\,(-4 + 4\,k - 7\,k^2$$

$$+ 4\,k^3)\}\,w^2 + 16\,(k-1)\,(1+k^2)\,(1+2\,\gamma)\,(2 + 6\,\gamma + 3\,\gamma^2)\,w$$

$$- 8\,(k-1)\,(1+k^2)\,(1+\gamma)\,(1+2\,\gamma)^2$$

and

$$H_4 = 15\,k^3\,w^2\,(w-1-\gamma)^2\,\{-1+w+\gamma\,(3\,w-2)\}^2.$$

The equation (7.5) is quartic in $w$. It is not possible to solve for $w$ analytically, at least in general, and therefore for a given $k$ and $\gamma$ the optimal value of $w$ is computed numerically.

Similarly, when $k$ is an odd integer greater than or equal to 3, the $V$-optimal population designs based on the set of two-point individual designs were computed numerically for various values of $k$ and $\gamma$. The population designs of the form

$$\xi^*_{V_o} = \left\{ \begin{array}{ccc} (0, \frac{k+1}{2}) & (0, k) & (\frac{k-1}{2}, k) \\ v & 1 - 2\,v & v \end{array} \right\} \tag{7.6}$$

were always $V$-optimal for the mean responses $\boldsymbol{\mu}_g$ at $\mathbf{t}_g = (0, 1, \ldots, k)'$ over this set for all $\gamma \geq 0$. The weight $v$ is chosen to minimize

$$\Psi_V(\tilde{\xi}^*_{V_o}) = tr\{\mathbf{M}^{-1}_{\beta}(\tilde{\xi}^*_o)\,\tilde{\mathbf{X}}'_g\tilde{\mathbf{X}}_g\} = -(1+2\,\gamma)\frac{H_5}{H_6}$$

where

$$\tilde{\xi}^*_{V_o} = \left\{ \begin{array}{ccc} (-\frac{k}{2}, \frac{1}{2}) & (-\frac{k}{2}, \frac{k}{2}) & (-\frac{1}{2}, \frac{k}{2}) \\ v & 1 - 2\,v & v \end{array} \right\}$$

is the design $\xi^*_{V_o}$ in the linearly transformed coordinates,

$$H_5 = 5\,(k^2-1)\,\{3\,(k^2-1)\,(1+3\,k)\,\gamma^2 + 2\,(k-1)\,(3+7\,k)\,\gamma + 3\,(1+k)\,(2\,k-1)\}\,v^2$$

$$- 2\,k\,(k-1)\,\{15\,k\,(k+1)^2\,\gamma^2 + (1+23\,k+31\,k^2-7\,k^3)\,\gamma - (k+1)\,(-1-15\,k+4\,k^2)\}v$$
$$- 8\,k^3\,(1+k^2)\,(1+2\,\gamma)$$

and

$$H_6 = 15\,(k^2-1)\,v\{\{-(k-1)\,(1+3\,k)\,\gamma^2 - 2\,(-1-k+3\,k^2)\,\gamma + 1 - 2\,k^2\}\,v + k^2(1+\gamma)\,(1+2\,\gamma)\}.$$

For fixed $k$ and $\gamma$, the optimal weight $v$ is a positive number less than $\frac{1}{2}$ and is a solution to the equation

$$\frac{\partial tr\{\mathbf{M}_\beta^{-1}(\tilde{\xi}^*_{V_o})\,\tilde{\mathbf{X}}_g'\tilde{\mathbf{X}}_g\}}{\partial v} = (1+2\,\gamma)\frac{H_7}{H_8} = 0 \tag{7.7}$$

where

$$H_7 = 5\,(k-1)^2\,(k+1)\,\{1+k+\gamma\,(3\,k+1)\}\,\{3\,(k+1)\,(2\,k-1)+2\,(k-1)\,(3+7\,k)\,\gamma$$
$$+\ 3\,(k-1)\,(k+1)\,(3\,k+1)\,\gamma^2\}\,v^4 - 4\,(k-1)^2\,k\,\{1+k+\gamma\,(3\,k+1)\}\,\{-(k+1)$$
$$\times\ (4\,k^2-15\,k-1) + (1+23\,k+31\,k^2-7\,k^3)\,\gamma + 15\,k\,(k+1)\,\gamma^2\}\,v^3$$
$$+\ \{(k-1)\,k\,\{-(k+1)\,(2+15\,k+27\,k^2-30\,k^3+40\,k^4) + (-6-39\,k-156\,k^2-15\,k^3$$
$$+ 40\,k^4 - 196\,k^5)\,\gamma + (-6-27\,k-122\,k^2-24\,k^3+224\,k^4-237\,k^5)\,\gamma^2 + (3\,k-1)\,(2+11\,k+41\,k^2$$
$$+\ 65\,k^3+k^4)\,\gamma^3\}\,v^2 + 16\,k^3\,(k^2+1)\,(1+2\,\gamma)\,\{-1+2\,k^2+2\,(3\,k^2-k-1)\,\gamma$$
$$+\ (k-1)\,(3\,k+1)\,\gamma^2\}\,v - 8\,k^5\,(k^2+1)\,(1+\gamma)\,(1+2\,\gamma)^2$$

and

$$H_8 = 15\,(k^2-1)\,v^2\,(1-v+\gamma)^2\,\{2\,k\,v\,\gamma + v\,(1+\gamma) - k^2\,(-1+v-2\,\gamma+3\,v\,\gamma)\}^2.$$

This equation is quartic in $v$ and it is not possible to solve for $v$ analytically at least in general. Therefore for a given $k$ and $\gamma$ the optimal value of $v$ is computed numerically.

The weights $w$ and $v$ for the optimal designs $\xi^*_{V_e}$ and $\xi^*_{V_o}$ respectively are given for some selected values of $k$ and $\gamma$ in Tables 7.2 and 7.3. For comparison purpose the criterion values $tr\{\mathbf{M}_\beta^{-1}(\xi_e)\,\mathbf{X}_g'\,\mathbf{X}_g\}$ of the design $\xi_e$ putting equal weights on $(0, \frac{k}{2})$, $(0, k)$ and $(\frac{k}{2}, k)$ are presented with the values of criterion $tr\{\mathbf{M}_\beta^{-1}(\xi^*_{V_e})\,\mathbf{X}_g'\,\mathbf{X}_g\}$ in Table 7.2. Similarly, the criterion values $tr\{\mathbf{M}_\beta^{-1}(\xi_o)\,\mathbf{X}_g'\,\mathbf{X}_g\}$ of the design $\xi_o$ putting equal weights on $(0, \frac{k+1}{2})$, $(0, k)$ and $(\frac{k-1}{2}, k)$ are presented with the values of criterion $tr\{\mathbf{M}_\beta^{-1}(\xi^*_{V_o})\,\mathbf{X}_g'\,\mathbf{X}_g\}$ in Table 7.3. Furthermore, the efficiencies of $\xi^*_{V_e}$ with respect to $\xi_e$ and $\xi^*_{V_o}$ with respect to the $\xi_o$ are presented in the last column of the respective tables. For $V$-optimality, for instance for $k$ even, the efficiency of the design $\xi_e$ relative to the optimal design $\xi^*_{V_e}$ is defined by $\dfrac{tr\{\mathbf{M}_\beta^{-1}(\xi^*_{V_e})\,\mathbf{X}_g'\,\mathbf{X}_g\}}{tr\{\mathbf{M}_\beta^{-1}(\xi_e)\,\mathbf{X}_g'\,\mathbf{X}_g\}}$.

There are a few immediate implications from the tables. The results suggest that the designs $\xi_e$ and $\xi_o$ which put equal weights on the support designs of the $V$-optimal population designs are about as efficient as the $V$-optimal designs $\xi^*_{V_e}$ and $\xi^*_{V_o}$ respectively for large variance ratios. It can also be observed that the design $\xi_e$ and $\xi_o$ are almost as efficient as the respective $V$-optimal designs when $k$ is small (as for instance, $k = 4$ and $k = 5$), for all $\gamma$ values. However, as $k$ gets larger the efficiency of $V$-optimal designs decreases.

The variances of the estimated mean response $\hat{\mu} = \hat{\beta}_0 + \hat{\beta}_1\,t + \hat{\beta}_2\,t^2$ at time $t$ for the designs $\xi^*_{V_e}$ and $\xi^*_{V_o}$ are given by

$$Var(\xi^*_{V_e}) = \mathbf{x}_g'\,\mathbf{M}_\beta^{-1}(\xi^*_{V_e})\,\mathbf{x}_g$$

and

$$Var(\xi^*_{V_o}) = \mathbf{x}_g'\,\mathbf{M}_\beta^{-1}(\xi^*_{V_o})\,\mathbf{x}_g$$

respectively, where $\mathbf{x}_g = [1\ t\ t^2]'$ and $\mathbf{M}_\beta^{-1}(\xi^*_{V_e})$ and $\mathbf{M}_\beta^{-1}(\xi^*_{V_o})$ are the inverses of the standardized information matrices for $\boldsymbol{\beta}$ at $\xi^*_{V_e}$ and $\xi^*_{V_o}$ respectively. For example, the variances

Table 7.2: Two-point $V$-optimal designs for estimation of the mean responses at $\mathbf{t}_g = (0, 1, \ldots, k)'$ in model (7.1) computed using equation (7.5)

| $k$ | $\gamma$ | $w$ | $tr\{\mathbf{M}_\beta^{-1}(\xi_{V_e}^*)\mathbf{X}_g'\mathbf{X}_g\}$ | $tr\{\mathbf{M}_\beta^{-1}(\xi_e)\mathbf{X}_g'\mathbf{X}_g\}$ | Efficiency |
|---|---|---|---|---|---|
| 4 | 0.1 | 0.401 | 14.376 | 14.674 | 0.980 |
| | 0.5 | 0.392 | 19.199 | 19.500 | 0.985 |
| | 1 | 0.387 | 24.677 | 24.975 | 0.988 |
| | 5 | 0.380 | 65.467 | 65.757 | 0.996 |
| 10 | 0.1 | 0.450 | 28.433 | 30.206 | 0.941 |
| | 0.5 | 0.432 | 38.862 | 40.626 | 0.957 |
| | 1 | 0.423 | 50.815 | 52.557 | 0.967 |
| | 5 | 0.410 | 140.396 | 142.091 | 0.988 |
| 20 | 0.1 | 0.468 | 52.249 | 56.661 | 0.922 |
| | 0.5 | 0.447 | 72.080 | 76.457 | 0.943 |
| | 1 | 0.436 | 94.859 | 99.180 | 0.956 |
| | 5 | 0.421 | 265.804 | 270.017 | 0.984 |
| 30 | 0.1 | 0.475 | 76.138 | 83.217 | 0.915 |
| | 0.5 | 0.452 | 105.382 | 112.400 | 0.938 |
| | 1 | 0.441 | 138.989 | 145.920 | 0.953 |
| | 5 | 0.425 | 391.308 | 398.070 | 0.983 |

Table 7.3: Two-point $V$-optimal designs for estimation of the mean responses at $\mathbf{t}_g = (0, 1, \ldots, k)'$ in model (7.1) calculated by solving equation (7.7)

| $k$ | $\gamma$ | $v$ | $tr\{\mathbf{M}_\beta^{-1}(\xi^*_{V_o})\mathbf{X}'_g\mathbf{X}_g\}$ | $tr\{\mathbf{M}_\beta^{-1}(\xi_o)\mathbf{X}'_g\mathbf{X}_g\}$ | Efficiency |
|---|---|---|---|---|---|
| 5 | 0.1 | 0.440 | 16.869 | 17.689 | 0.954 |
| | 0.5 | 0.441 | 18.978 | 23.325 | 0.957 |
| | 1 | 0.439 | 28.671 | 29.799 | 0.962 |
| | 5 | 0.437 | 77.250 | 78.567 | 0.983 |
| 15 | 0.1 | 0.459 | 40.285 | 43.531 | 0.925 |
| | 0.5 | 0.451 | 55.097 | 58.516 | 0.942 |
| | 1 | 0.443 | 72.273 | 75.765 | 0.954 |
| | 5 | 0.431 | 202.208 | 205.807 | 0.983 |
| 21 | 0.1 | 0.472 | 54.565 | 59.389 | 0.919 |
| | 0.5 | 0.454 | 75.030 | 80.004 | 0.938 |
| | 1 | 0.443 | 72.273 | 75.765 | 0.952 |
| | 5 | 0.431 | 202.208 | 205.807 | 0.982 |
| 29 | 0.1 | 0.476 | 73.661 | 80.602 | 0.914 |
| | 0.5 | 0.456 | 101.655 | 108.727 | 0.935 |
| | 1 | 0.446 | 133.991 | 141.090 | 0.950 |
| | 5 | 0.431 | 377.845 | 384.968 | 0.982 |

of $\hat{\mu}$ for $k = 6$, and $\gamma = 0.1, \gamma = 0.5$ $\gamma = 1$ and $\gamma = 5$ as a function of the time point $t$ are given by

$$Var(\xi^*_{V_e}, t) = 0.052\,(29.625 - 10.521\,t + t^2)\,(2.501 - 1.479\,t + t^2) \quad \text{for} \quad \gamma = 0.1,$$

$$Var(\xi^*_{V_e}, t) = 0.055\,(31.213 - 10.732\,t + t^2)\,(2.818 - 1.268\,t + t^2) \quad \text{for} \quad \gamma = 0.5,$$

$$Var(\xi^*_{V_e}, t) = 0.057\,(32.831 - 10.944\,t + t^2)\,(3.169 - 1.056\,t + t^2) \quad \text{for} \quad \gamma = 1$$

and

$$Var(\xi^*_{V_e}, t) = 0.061\,(41.349 - 11.960\,t + t^2)\,(5.588 - 0.040\,t + t^2) \quad \text{for} \quad \gamma = 5$$

respectively. Similarly, the variances of $\hat{\mu}$ for $k = 7$, and $\gamma = 0.1, \gamma = 0.5$, $\gamma = 1$ and $\gamma = 5$ as a function of the time point $t$ are given by

$$Var(\xi^*_{V_o}, t) = 0.029\,(39.964 - 12.220\,t + t^2)\,(3.426 - 1.780\,t + t^2) \quad \text{for} \quad \gamma = 0.1,$$

$$Var(\xi^*_{V_o}, t) = 0.030\,(42.065 - 12.448\,t + t^2)\,(3.929 - 1.552\,t + t^2) \quad \text{for} \quad \gamma = 0.5,$$

$$Var(\xi^*_{V_o}, t) = 0.030\,(44.326 - 12.695\,t + t^2)\,(4.458 - 1.305\,t + t^2) \quad \text{for} \quad \gamma = 1$$

and

$$Var(\xi^*_{V_o}, t) = 0.032\,(56.412 - 13.927\,t + t^2)\,(7.920 - 0.073\,t + t^2) \quad \text{for} \quad \gamma = 5$$

respectively. These variances are displayed in Figure 7.5 for $k = 6$ and in Figure 7.6 for $k = 7$. A similar pattern is observed for both k even and k odd in that for all $\gamma$ the variance is large when estimation of $\mu = \beta_0 + \beta_1\,t + \beta_2\,t^2$ has been made at early or late times. In general the estimation of $\mu$ is more precise for small $\gamma$.

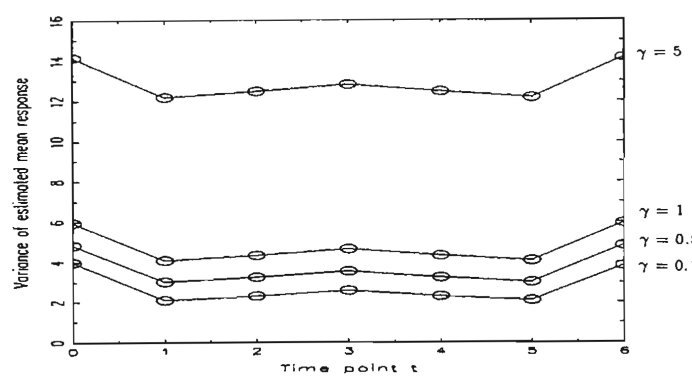


Figure 7.5: Variances of estimated mean responses for two-point $V$-optimal population designs $\xi^*_{V_e}$ as a function of $t$ when $k = 6$ and $\gamma = 0.1, 0.5, 1$ and 5.

## 7.4 $V$-optimal population design based on $d$-point individual designs

The GAUSS program described in Section 5.7 can be used to extend the numerical computation of $V$-optimal designs to those based on $d$-points where $3 \leq d \leq k+1$ for $k \geq 3$. Specifically $V$-optimal population designs based on $d$-point individual designs for a given value of $\gamma$ either over the space of designs $S_{d,k}$ or over $T_{d,k}$ can be readily found. The program can also be used to compute the best $V$-optimal population design over the set of all population designs.

As an example, for $k = 6$ and $\gamma = 0.05$ the following optimal designs were obtained:

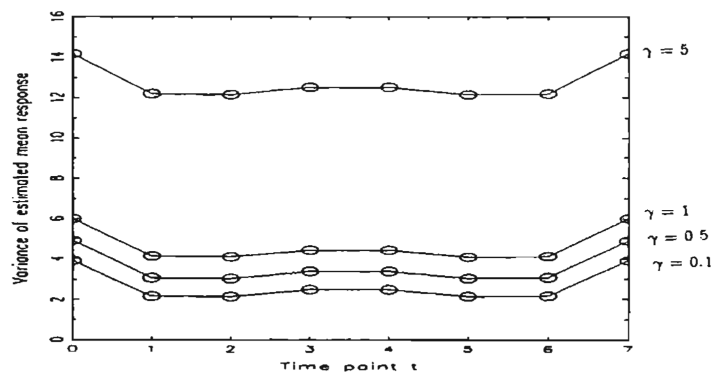


Figure 7.6: Variances of estimated mean responses for two-point $V$-optimal population designs $\xi^*_{V_o}$ as a function of $t$ when $k = 7$ and $\gamma = 0.1, 0.5, 1$ and $5$.

(i) One-point design

$$\xi^*_{V_1} = \left\{ \begin{array}{ccc} (0) & (3) & (6) \\ 0.2840 & 0.4320 & 0.2840 \end{array} \right\}$$

with criterion value $\Psi(\xi^*_{V_1}) = 17.9961$.

(ii) Two-point design

$$\xi^*_{V_2} = \left\{ \begin{array}{ccc} (0,3) & (0,6) & (3,6) \\ 0.4293 & 0.1414 & 0.4293 \end{array} \right\}$$

with criterion value $\Psi(\xi^*_{V_2}) = 18.0860$.

(iii) Three-point design

$$\xi^*_{V_3} = \left\{ \begin{array}{ccc} (0,3,4) & (0,3,6) & (2,3,6) \\ 0.1387 & 0.7226 & 0.1387 \end{array} \right\}$$

with criterion value $\Psi(\xi^*_{V_3}) = 18.4632$.

(iv) Four-point design

$$\xi^*_{V_4} = \left\{ \begin{array}{cc} (0,2,3,6) & (0,3,4,6) \\ 0.5000 & 0.5000 \end{array} \right\}$$

with criterion value $\Psi(\xi^*_{V_4}) = 19.1427$.

(v) Five-point design

$$\xi^*_{V_5} = \left\{ \begin{array}{ccc} (0,1,3,4,6) & (0,2,3,4,6) & (0,2,3,5,6) \\ 0.2528 & 0.4944 & 0.2528 \end{array} \right\}$$

with criterion value $\Psi(\xi^*_{V_5}) = 20.6267$.

(vi) Six-point design

$$\xi^*_{V_6} = \left\{ \begin{array}{cccc} (0,1,2,3,4,6) & (0,1,2,3,5,6) & (0,1,3,4,5,6) & (0,2,3,4,5,6 \\ 0.3478 & 0.1522 & 0.1522 & 0.3478 \end{array} \right\}$$

with criterion value $\Psi(\xi^*_{V_6}) = 22.1062$.

(vii) Best $V$-optimal design

$$\xi^*_{V_b} = \left\{ \begin{array}{cc} (3) & (0,6) \\ 0.4327 & 0.5673 \end{array} \right\}$$

with criterion value $\Psi(\xi^*_{V}) = 17.9352$.

Observe from the criterion values that the optimal designs with small $d$ are more efficient than those with large $d$ or, in other words, for a given variance ratio $\gamma$ the efficiency decreases as $d$ increases.

The $V$-optimal population designs change with the variance ratio $\gamma$. For example, when $k = 5$, the best $V$-optimal population designs for $\gamma = 0.005$, $\gamma = 0.05$ and $\gamma = 2$ are

$$\xi^*_{V_{b1}} = \left\{ \begin{array}{ccc} (2) & (3) & (0,\,5) \\ 0.219 & 0.219 & 0.562 \end{array} \right\},$$

$$\xi^*_{V_{b2}} = \left\{ \begin{array}{ccccc} (2) & (3) & (0,\ 3) & (0,\ 5) & (2,5) \\ 0.208 & 0.208 & 0.023 & 0.538 & 0.023 \end{array} \right\}$$

and

$$\xi^*_{V_{b3}} = \left\{ \begin{array}{ccccccc} (0) & (2) & (3) & (5) & (0,\ 3) & (0,\ 5) & (2,5) \\ 0.078 & 0.140 & 0.118 & 0.083 & 0.212 & 0.163 & 0.206 \end{array} \right\}$$

respectively. The $V$-optimality of these designs was checked by using the Equivalence Theorem for $V$-optimal population designs. The plots of the directional derivatives $\phi_V(\mathbf{t}, \xi^*_{V_{bi}}), i = 1,2,3$ against the individual designs $\mathbf{t} \in S_{d,5}$ where $1 \leq d \leq 6$, i.e. (0), (1), ..., (5), (0, 1), ..., (1, 2, 3, 4, 5, 6), labelled for convenience 1 through 63, are presented in Figures 7.7, 7.8 and 7.9. There are three maxima for $\phi_V(\mathbf{t}, \xi^*_{V_{b1}})$, five maxima for $\phi(\mathbf{t}, \xi^*_{V_{b2}})$ and seven maxima for $\phi_V(\mathbf{t}, \xi^*_{V_{b3}})$ over the design spaces $S_{d,k}, d = 1,2,3,4,5$ and the maxima are equal to zero. These maxima occurred at the support designs of $\xi^*_{V_{bi}}, i = 1,2,3$, and thus the designs would seem, at least numerically, to be optimum.

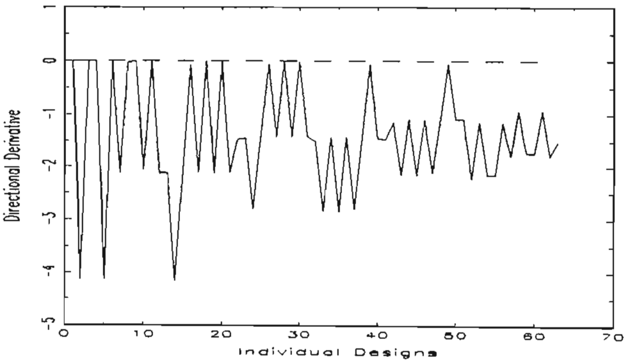


Figure 7.7: Plot of the directional derivative $\phi_V(\mathbf{t}, \xi^*_{V_{b1}})$ against the individual designs $\mathbf{t}$.

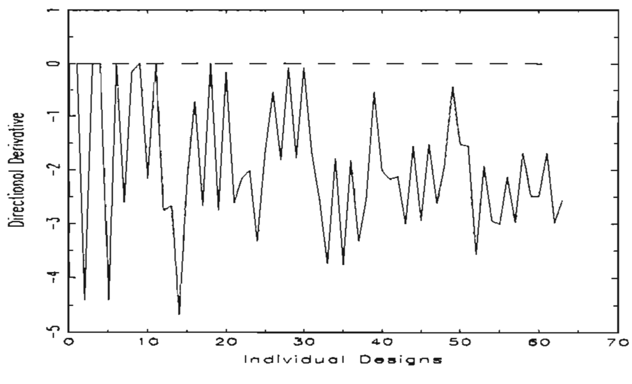


Figure 7.8: Plot of the directional derivative $\phi_V(\mathbf{t}, \xi^*_{V_{b2}})$ against the individual designs $\mathbf{t}$.

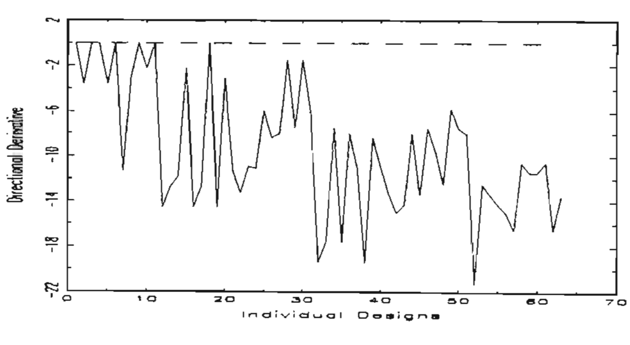


Figure 7.9: Plot of the directional derivative $\phi_V(\mathbf{t}, \xi^*_{V_{b3}})$ against the individual designs $\mathbf{t}$.

# Chapter 8

# Optimal Designs for Random Coefficient Regression Models

## 8.1 Introduction

In the previous chapters the discussions were on the construction of optimal designs for the precise estimation of the parameters and mean responses under the linear model with a random intercept and in particular, with the simple linear and the quadratic regression models with a random intercept. In this Chapter the problems of constructing optimal designs for the simple linear regression model with a random slope and for the simple linear random coefficient regression model are considered. In the first model a random effect is attached to the slope parameter, whereas in the second model random effects are attached to both the intercept and the slope parameters.

Before presenting some properties and numerical results for the optimal designs for these models, the impact of a linear transformation of the columns of design matrices $\mathbf{X}_i$ on the variance structure of the random effects, i.e. on the matrix $\mathbf{G}$, is discussed. Specifically, recall from Subsection 2.5.3 that for the random coefficient model with $\mathbf{Z}_i = \mathbf{X}_i$, $i = 1, \ldots, K$ the linear transformation $\mathbf{X}_i^* = \mathbf{X}_i\,\mathbf{A}$ of the columns of $\mathbf{X}_i$, where $\mathbf{A}$ is a non-singular matrix, yields $\mathbf{G}^* = \mathbf{A}^{-1}\,\mathbf{G}\,(\mathbf{A}^{-1})'$ as the variance of random effects. In fact, this linear transformation induces the transformation $\mathbf{b}_i^* = \mathbf{A}^{-1}\,\mathbf{b}_i$ in the random effects.

Consider now the simple linear random coefficient regression model specified by expression (3.3), that is by

$$y_{ij} = (\beta_o + b_{0i}) + (\beta_1 + b_{1i})\,t_{ij} + e_{ij}, \qquad j = 1, 2, \ldots, d_i \ \text{ and } \ i = 1, 2, \ldots, K \tag{8.1}$$

where

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma_{b_1}^2 \end{pmatrix}\right].$$

The complete description of this model will be presented later in Section 8.3. Further, let

$$\mathbf{G}_\gamma = \frac{1}{\sigma_e^2}\begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma_{b_1}^2 \end{pmatrix} = \begin{pmatrix} \gamma_{b_0} & \gamma_{b_0 b_1} \\ \gamma_{b_0 b_1} & \gamma_{b_1} \end{pmatrix} \tag{8.2}$$

where $\sigma_e^2 = Var(e_{ij})$. Suppose that a time point $t$ is linearly transformed as

$$t_j^* = u + v\,t_j, \quad j = 1, \ldots, d$$

where $u$ and $v$ are constants and the subscripts relating to individuals are ignored for convenience. This transformation can be represented in matrix form as $\mathbf{X}^* = \mathbf{X}\,\mathbf{A}$ where

$\mathbf{X} = (\mathbf{1}_d \ \mathbf{t})$ with $\mathbf{t} = (t_1, \ldots, t_d)'$ and $\mathbf{A} = \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix}$. Since

$$\mathbf{A}^{-1} = \begin{pmatrix} 1 & -\frac{u}{v} \\ 0 & \frac{1}{v} \end{pmatrix}$$

it follows that the variance-covariance matrix for the random effects in the transformed model corresponds to $\mathbf{G}^* = \sigma_e^2 \, \mathbf{G}_\gamma^*$ where

$$\mathbf{G}_\gamma^* = \mathbf{A}^{-1} \, \mathbf{G}_\gamma \, (\mathbf{A}^{-1})' = \begin{pmatrix} \gamma_{b_0} - 2\frac{u}{v}\gamma_{b_0b_1} + \frac{u^2}{v} & \frac{\gamma_{b_0b_1}}{v} - \frac{u}{v^2}\gamma_{b_1} \\ \frac{\gamma_{b_0b_1}}{v} - \frac{u}{v^2}\gamma_{b_1} & \frac{\gamma_{b_1}}{v^2} \end{pmatrix}. \tag{8.3}$$

As explained in Subsection 2.5.3, the structure of $\mathbf{G}_\gamma^*$ may be different from that of $\mathbf{G}_\gamma$ and thus the underlying model for the random effects may change. The optimum designs in the following sections are discussed taking into account this possible change.

The organization of the rest of this Chapter as follows. The numerical construction of $D$-optimal designs for the simple linear regression model with a random slope is discussed in Section 8.2 and $D$-optimal designs for the simple linear random coefficient regression model are considered in Section 8.3. Some concluding remarks are given in Section 8.4. In this chapter only the set of individual designs with non-repeated time points, i.e. individual designs $\mathbf{t} = (t_1, \ldots, t_d)'$ with $t_1 < t_2 < \ldots < t_d$, are considered. However the results obtained can easily be extended to designs with repeated time points.

# 8.2 $D$-optimal designs for the simple linear regression model with a random slope

In this section $D$-optimal designs for precise estimation of the fixed effects and the variance components in the simple linear regression model with a random slope are discussed. In the first subsection, the simple linear regression model with a random slope and appropriate information matrices are presented. Then $D$-optimal individual designs for the fixed effects are considered in the second subsection and $D$-optimal population designs for estimation of the fixed effects and the variance components are discussed in the third subsection.

## 8.2.1 Model

Recall the simple linear random coefficient regression model specified by expression (8.1). Then the simple linear regression model with a random slope is a special case of (8.1) and is defined as

$$y_{ij} = \beta_o + (\beta_1 + b_{1i})\,t_{ij} + e_{ij}, \quad j = 1, 2, \ldots, d_i \ \text{ and } i = 1, 2, \ldots, K. \tag{8.4}$$

The parameters $\beta_o$ and $\beta_1$ have their usual significance and $b_{1i}$ represents the occurrence of a random effect, namely a random slope. It is assumed that $b_{1i} \sim N(0,\ \sigma_{b_1}^2)$, that $e_{ij} \sim (0,\ \sigma_e^2)$, and that all of the $b_{1i}$'s and $e_{ij}$'s are uncorrelated within and between individuals for $i = 1, \ldots, K$. Note that under the linear mixed model formulation of Chapter 2, model (8.4) comprises $\boldsymbol{\beta} = (\beta_o,\ \beta_1)'$, $\mathbf{b}_i = b_{1i}$, $\mathbf{X}_i = [\mathbf{1}_{d_i} \ \ \mathbf{t}_i]$, $\mathbf{Z}_i = \mathbf{t}_i$ with $\mathbf{t}_i = (t_{i1},\ t_{i2},\ \ldots,\ t_{id_i})'$ and $\mathbf{e}_i = (e_{i1},\ e_{i2},\ \ldots,\ e_{id_i})'$. Therefore under the assumption of normality for $b_{1i}$ and $e_{ij}$, the marginal distribution of $\mathbf{y}_i$ is $\mathcal{N}(\mathbf{X}_i\boldsymbol{\beta},\ \mathbf{V}_i)$ with $\mathbf{V}_i = \sigma_e^2\mathbf{I}_{d_i} + \sigma_{b_1}^2\mathbf{t}_i\mathbf{t}_i' = \sigma_e^2(\mathbf{I}_{d_i} + \delta\mathbf{t}_i\mathbf{t}_i')$ where $\delta = \dfrac{\sigma_{b_1}^2}{\sigma_e^2}$ is the variance ratio.

Observe that, if $\gamma_{b_0} = 0$ and $\gamma_{b_0 b_1} = 0$ and thus a random slope model is adopted, then the matrix $\mathbf{G}_\gamma$ in (8.2) corresponds to

$$\begin{pmatrix} 0 & 0 \\ 0 & \delta \end{pmatrix}.$$

Therefore for the linear transformation $t^* = u + v\,t$ the matrix in (8.3) simplifies to

$$\mathbf{G}_\gamma^* = \frac{\gamma_{b_1}}{v^2}\begin{pmatrix} u^2 & -u \\ -u & 1 \end{pmatrix}.$$

Thus it can easily be seen from $\mathbf{G}_\gamma^*$ that the random slope structure will be preserved in model (8.4) if and only if $u = 0$, that is, if and only if there is no change in location of the time points. Otherwise, the model takes the structure of the simple linear random coefficient regression model. This means that if a design which is optimal for a random slope model is linearly transformed it is not necessarily optimal for the associated transformed random slope model.

The information matrix for the parameters $\boldsymbol{\beta}$ in the random slope model at an individual design $\mathbf{t}_i$ is derived as follows. Recall from expression (2.10) that the $i$th individual information matrix for $\boldsymbol{\beta}$ in the linear mixed model is

$$\mathbf{I}_\beta(\mathbf{X}_i) = \mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i, \quad i = 1, \ldots, K.$$

Now in the present case, using Result A.2.1 from Appendix A, the inverse of the variance-covariance matrix

$$\mathbf{V}_i = \sigma_e^2\,(\mathbf{I}_{d_i} + \delta\,\mathbf{t}_i\,\mathbf{t}_i')$$

is given by

$$\mathbf{V}_i^{-1} = \frac{1}{\sigma_e^2}\,(\mathbf{I}_{d_i} + \delta\,\mathbf{t}_i\,\mathbf{t}_i')^{-1} = \frac{1}{\sigma_e^2}\left(\mathbf{I}_{d_i} - \frac{\delta\,\mathbf{t}_i\,\mathbf{t}_i'}{1+\delta\,\mathbf{t}_i'\,\mathbf{t}_i}\right).$$

Then using this inverse matrix and $\mathbf{X}_i = [\mathbf{1}_{d_i}\ \ \mathbf{t}_i]$ in $\mathbf{X}_i'\,\mathbf{V}_i^{-1}\,\mathbf{X}_i$ yield

$$\mathbf{I}_\beta(\mathbf{X}_i) = \begin{pmatrix} \mathbf{1}_{d_i}'\,\mathbf{V}_i^{-1}\,\mathbf{1}_{d_i} & \mathbf{1}_{d_i}'\,\mathbf{V}_i^{-1}\,\mathbf{t}_i \\ \mathbf{t}_i'\,\mathbf{V}_i^{-1}\,\mathbf{1}_{d_i} & \mathbf{t}_i'\,\mathbf{V}_i^{-1}\,\mathbf{t}_i \end{pmatrix}.$$

Further,

$$\mathbf{1}_{d_i}'\,\mathbf{V}_i^{-1}\,\mathbf{1}_{d_i} = \frac{1}{\sigma_e^2}\,\mathbf{1}_{d_i}'\left(\mathbf{I}_{d_i} - \frac{\delta\,\mathbf{t}_i\,\mathbf{t}_i'}{1+\delta\,\mathbf{t}_i'\,\mathbf{t}_i}\right)\mathbf{1}_{d_i} = \frac{d_i[1+\delta\,SS(t)]}{\sigma_e^2\,(1+\delta\sum_{j=1}^{d_i} t_{ij}^2)}$$

where $SS(t) = \mathbf{t}_i'\,\mathbf{t}_i - \dfrac{1}{d_i}\,(\mathbf{1}_{d_i}'\,\mathbf{t}_i)^2$,

$$\mathbf{1}_{d_i}'\,\mathbf{V}_i^{-1}\,\mathbf{t}_i = \frac{1}{\sigma_e^2}\,\mathbf{1}_{d_i}'\left(\mathbf{I}_{d_i} - \frac{\delta\,\mathbf{t}_i\,\mathbf{t}_i'}{1+\delta\,\mathbf{t}_i'\,\mathbf{t}_i}\right)\mathbf{t}_i = \frac{\sum_{j=1}^{d_i} t_{ij}}{\sigma_e^2(1+\delta\sum_{j=1}^{d_i} t_{ij}^2)}$$

and

$$\mathbf{t}_i'\,\mathbf{V}_i^{-1}\,\mathbf{t}_i = \frac{1}{\sigma_e^2}\,\mathbf{t}_i'\left(\mathbf{I}_{d_i} - \frac{\delta\,\mathbf{t}_i\,\mathbf{t}_i'}{1+\delta\,\mathbf{t}_i'\,\mathbf{t}_i}\right)\mathbf{t}_i = \frac{\sum_{j=1}^{d_i} t_{ij}^2}{\sigma_e^2\,(1+\delta\sum_{j=1}^{d_i} t_{ij}^2)}.$$

Therefore the standardized information matrix for the fixed effects $\boldsymbol{\beta}$ for the $i$th individual with $\mathbf{t}_i = (t_{i1},\ldots,t_{id_i})'$ is given by

$$\mathbf{M}_\beta(\mathbf{t}_i) = \frac{\mathbf{I}_\beta(\mathbf{t}_i)}{d_i} = \frac{1}{\sigma_e^2}\,\frac{1}{d_i\,(1+\delta\sum_{j=1}^{d_i} t_{ij}^2)}\begin{pmatrix} d_i[1+\delta\,SS(t)] & \sum_{j=1}^{d_i} t_{ij} \\ \sum_{j=1}^{d_i} t_{ij} & \sum_{j=1}^{d_i} t_{ij}^2 \end{pmatrix}. \tag{8.5}$$

Since $\dfrac{1}{\sigma_e^2}$ factors out in (8.5), $\sigma_e^2$ can be taken to be 1 without loss of generality. The same expression for $\mathbf{I}_\beta(\mathbf{X}_i)$ can be obtained using expression (2.33), i.e. $[(\mathbf{X}_i'\mathbf{X}_i)^{-1}+\mathbf{G}_\gamma]^{-1}$ where

$$\mathbf{X}_i'\mathbf{X}_i = \begin{pmatrix} d_i & \sum_{j=1}^{d_i} t_{ij} \\ \sum_{j=1}^{d_i} t_{ij} & \sum_{j=1}^{d_i} t_{ij}^2 \end{pmatrix} \quad \text{and} \quad \mathbf{G}_\gamma = \begin{pmatrix} 0 & 0 \\ 0 & \delta \end{pmatrix}.$$

Consider now the information matrix for the variance component vector $\boldsymbol{\theta} = (\sigma_e^2, \delta)$ in the random slope model at an individual design $\mathbf{t}_i$. Since $\mathbf{V}_i = \sigma_e^2(\mathbf{I}_{d_i} + \delta\mathbf{t}_i\mathbf{t}_i')$, it follows that

$$\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} = \mathbf{I}_{d_i} + \delta \mathbf{t}_i \mathbf{t}_i' \quad \text{and} \quad \frac{\partial \mathbf{V}_i}{\partial \delta} = \sigma_e^2 \mathbf{t}_i \mathbf{t}_i'.$$

Then the elements of the information matrix for the variance component vector $\boldsymbol{\theta} = (\sigma_e^2, \delta)$ are

$$\mathbf{I}_{\sigma_e^2,\sigma_e^2}(\mathbf{t}_i) = \frac{1}{2} tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2}\right)^2 = \frac{1}{2}\, tr\left\{\frac{1}{\sigma_e^2}\left(\mathbf{I}_{d_i} + \delta\, \mathbf{t}_i\, \mathbf{t}_i'\right)^{-1}\left(\mathbf{I}_{d_i} + \delta \mathbf{t}_i \mathbf{t}_i'\right)\right\}^2 = \frac{1}{2\,\sigma_e^4},$$

$$\mathbf{I}_{\sigma_e^2,\delta}(\mathbf{t}_i) = \frac{1}{2}\, tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \delta}\right) = \frac{1}{2\,\sigma_e^2}\left(\frac{\sum_{j=1}^{d_i} t_{ij}^2}{1+\delta \sum_{j=1}^{d_i} t_{ij}^2}\right)$$

and

$$\mathbf{I}_{\delta,\delta}(\mathbf{t}_i) = \frac{1}{2}\, tr\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \delta}\right)^2 = \frac{1}{2}\left(\frac{\sum_{j=1}^{d_i} t_{ij}^2}{1+\delta \sum_{j=1}^{d_i} t_{ij}^2}\right)^2$$

since

$$\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \sigma_e^2} = \frac{1}{\sigma_e^2}$$

and

$$\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \delta} = \mathbf{t}_i\mathbf{t}_i' - \frac{\delta\,(\mathbf{t}_i\mathbf{t}_i'\mathbf{t}_i\mathbf{t}_i')}{1+\delta\, \mathbf{t}_i'\mathbf{t}_i}.$$

On assembling the above results, the information matrix for the variance components $\boldsymbol{\theta} = (\sigma_e^2,\ \delta)$ for the $i$th individual with $\mathbf{t}_i = (t_{i1}, \ldots, t_{id_i})'$ is therefore given by

$$\mathbf{I}_{\theta}(\mathbf{t}_i) = \frac{1}{2}\begin{pmatrix} \dfrac{1}{\sigma_e^4} & \dfrac{1}{\sigma_e^2} \times \dfrac{\sum_{j=1}^{d_i} t_{ij}^2}{1+\delta \sum_{j=1}^{d_i} t_{ij}^2} \\ \\ \dfrac{1}{\sigma_e^2} \times \dfrac{\sum_{j=1}^{d_i} t_{ij}^2}{1+\delta \sum_{j=1}^{d_i} t_{ij}^2} & \left(\dfrac{\sum_{j=1}^{d_i} t_{ij}^2}{1+\delta \sum_{j=1}^{d_i} t_{ij}^2}\right)^2 \end{pmatrix}$$

and hence the standardized information matrix for $\boldsymbol{\theta}$ for the $i$th individual is equal to

$$\mathbf{M}_{\theta}(\mathbf{t}_i) = \frac{\mathbf{I}_{\theta}(\mathbf{t}_i)}{d_i}$$

for $i = 1, \ldots, K$.

### 8.2.2 $d$-point $D$-optimal individual designs for $\boldsymbol{\beta}$

Recall from expression (8.5) that the standardized information matrix for $\boldsymbol{\beta}$ at a $d$-point individual design $\mathbf{t} = (t_1, t_2, \ldots, t_d)'$ with $t_1 < t_2 < \ldots < t_d$ is

$$\mathbf{I}_\beta(\mathbf{t}) = \frac{1}{\sigma_e^2} \frac{1}{d\,(1 + \delta \sum_{j=1}^d t_j^2)} \begin{pmatrix} d[1 + \delta\, SS(t)] & \sum_{j=1}^d t_j \\ \sum_{j=1}^d t_j & \sum_{j=1}^d t_j^2 \end{pmatrix}$$

where $\delta = \dfrac{\sigma_{b_1}^2}{\sigma_e^2}$ and $SS(t) = \mathbf{t}'\mathbf{t} - \dfrac{1}{d}(\mathbf{1}'\mathbf{t})^2$. Then the $d$-point exact $D$-optimal individual design for $\boldsymbol{\beta}$ maximizes the determinant of the information matrix for $\boldsymbol{\beta}$, that is it maximizes

$$|\mathbf{I}_\beta(\mathbf{t})| = \frac{SS(t)}{d\,(1 + \delta \sum_{j=1}^d t_j^2)}. \tag{8.6}$$

Note that $\sigma_e^2 = 1$ by assumption. Observe that both the numerator and denominator in (8.6) depend on the time points, and that unlike the case of the simple linear regression model with a random intercept $|\mathbf{I}_\beta(\mathbf{t})|$ depends on the variance ratio $\delta$. Therefore only locally $D$-optimal designs, i.e. designs for a best guess of the parameter $\delta$, are considered in this section. A numerical example to illustrate the property of the $D$-optimal individual designs is given below.

**Example 8.2.1** Consider the simple linear regression model with a random slope as specified by model (8.4) and let $k = 4$. Therefore there are $\begin{pmatrix} 5 \\ 2 \end{pmatrix} = 10$ two-point, $\begin{pmatrix} 5 \\ 3 \end{pmatrix} = 10$ three-point and $\begin{pmatrix} 5 \\ 4 \end{pmatrix} = 5$ four-point individual designs. Consider first the two-point individual designs. The determinants of the information matrices for $\boldsymbol{\beta}$ for the two-point individual designs are given by

$$|\mathbf{I}_\beta((0,1))| = \frac{1}{4\,(1+\delta)}, \qquad |\mathbf{I}_\beta((0,2))| = \frac{1}{1 + 4\,\delta},$$

$$|\mathbf{I}_\beta((0,3))| = \frac{9}{4\,(1+9\,\delta)}, \qquad |\mathbf{I}_\beta((0,4))| = \frac{4}{1+16\,\delta}$$
$$|\mathbf{I}_\beta((1,2))| = \frac{1}{4\,(1+5\,\delta)}, \qquad |\mathbf{I}_\beta((1,3))| = \frac{1}{1+10\,\delta},$$
$$|\mathbf{I}_\beta((1,4))| = \frac{9}{4\,(1+17\,\delta)}, \qquad |\mathbf{I}_\beta((2,3))| = \frac{1}{4\,(1+13\,\delta)},$$
$$|\mathbf{I}_\beta((2,4))| = \frac{1}{1+20\,\delta} \text{ and } |\mathbf{I}_\beta((3,4))| = \frac{1}{4\,(1+25\,\delta)}.$$

Clearly the two-point $D$-optimal individual design is (0, 4) for all values of the variance ratio $\delta \geq 0$.

The determinants of the information matrices for $\boldsymbol{\beta}$ for the three-point individual designs are

$$|\mathbf{I}_\beta((0,1,2))| = \frac{2}{3\,(1+5\,\delta)}, \qquad |\mathbf{I}_\beta((0,1,3))| = \frac{14}{9\,(1+10\,\delta)},$$
$$|\mathbf{I}_\beta((0,1,4))| = \frac{26}{9\,(1+17\,\delta)}, \qquad |\mathbf{I}_\beta((0,2,3))| = \frac{14}{9\,(1+13\,\delta)},$$
$$|\mathbf{I}_\beta((0,2,4))| = \frac{8}{3\,(1+20\,\delta)}, \qquad |\mathbf{I}_\beta((0,3,4))| = \frac{26}{9\,(1+25\,\delta)},$$
$$|\mathbf{I}_\beta((1,2,3))| = \frac{2}{3\,(1+14\,\delta)}, \qquad |\mathbf{I}_\beta((1,2,4))| = \frac{14}{9\,(1+21\,\delta)},$$
$$|\mathbf{I}_\beta((1,3,4))| = \frac{14}{9\,(1+26\,\delta)}, \text{ and } |\mathbf{I}_\beta((2,3,4))| = \frac{2}{3\,(1+29\,\delta)}.$$

Thus (0,1,4) is the three-point $D$-optimal individual design for all $\delta > 0$ and when $\delta = 0$ both (0, 1, 4) and (0, 3, 4) are $D$-optimal.

Finally, the determinants of the information matrices for $\boldsymbol{\beta}$ for the four-point individual designs are

$$|\mathbf{I}_\beta((0,1,2,3))| = \frac{5}{4\,(1+4\,\delta)}, \qquad |\mathbf{I}_\beta((0,1,2,4))| = \frac{35}{16\,(1+21\,\delta)},$$
$$|\mathbf{I}_\beta((0,1,3,4))| = \frac{5}{2\,(1+26\,\delta)}, \qquad |\mathbf{I}_\beta((0,2,3,4))| = \frac{35}{16\,(1+29\,\delta)}$$

and

$$|\mathbf{I}_\beta((1,2,3,4))| = \frac{5}{4\,(1+30\,\delta)}.$$

Consider the difference

$$|\mathbf{I}_\beta((0,1,2,4))| - |\mathbf{I}_\beta((0,1,3,4))| = \frac{5\,(14\,\delta-1)}{16\,(1+21\,\delta)\,(1+26\,\delta)}$$

which is greater than zero if and only if $\delta > \frac{1}{14}$. Thus, if $\delta < \frac{1}{14}$ then the $D$-optimal individual design is (0,1,3,4), when $\delta > \frac{1}{14}$ (0,1,2,4) is the $D$-optimal individual design and when $\delta = \frac{1}{14}$ both (0,1,3,4) and (0,1,2,4) are $D$-optimal.

Suppose now that the time points in the above example are transformed as $\tilde{t}_j = t_j - 2$, $j = 1\ldots,d$. Then the determinants of the information matrices for $\boldsymbol{\beta}$ for the two-point individual designs in the transformed coordinates are

$$|\mathbf{I}_\beta((-2,-1))| = |\mathbf{I}_\beta((1,2))| = \frac{1}{4+20\,\delta},$$

$$|\mathbf{I}_\beta((-2,0))| = |\mathbf{I}_\beta((0,2))| = \frac{1}{1+4\,\delta},$$

$$|\mathbf{I}_\beta((-2,1))| = |\mathbf{I}_\beta((-1,2))| = \frac{9}{4\,(1+5\,\delta)},$$

$$|\mathbf{I}_\beta((-1,0))| = |\mathbf{I}_\beta((0,1))| = \frac{1}{4\,(1+\delta)},$$

$$|\mathbf{I}_\beta((-1,1))| = \frac{1}{1+2\,\delta} \quad \text{and} \quad |\mathbf{I}_\beta((-2,2))| = \frac{4}{1+8\,\delta}.$$

So, the two-point $D$-optimal individual design in the transformed coordinates is (-2, 2) for all $\delta \geq 0$.

The determinants of the information matrices for $\boldsymbol{\beta}$ for the three-point individual designs in the transformed coordinates are

$$|\mathbf{I}_\beta((-2,-1,0))| = |\mathbf{I}_\beta((0,1,2))| = \frac{2}{3\,(1+5\,\delta)},$$

$$|\mathbf{I}_\beta((-2,-1,1))| = |\mathbf{I}_\beta((-2,-1,1))| = \frac{14}{9\,(1+6\,\delta)},$$

$$|\mathbf{I}_\beta((-2,-1,2))| = |\mathbf{I}_\beta((-2,1,2))| = \frac{26}{9\,(1+9\,\delta)},$$

$$|\mathbf{I}_\beta((-1,0,2))| = |\mathbf{I}_\beta((-2,0,1))| = \frac{14}{9\,(1+5\,\delta)},$$

$$|\mathbf{I}_\beta((-1,0,1))| = \frac{2}{3\,(1+2\,\delta)} \quad \text{and} \quad |\mathbf{I}_\beta((-2,0,2))| = \frac{8}{3\,(1+8\,\delta)}.$$

Consider the difference

$$|\mathbf{I}_\beta((-2,0,2))| - |\mathbf{I}_\beta((-2,-1,2))| = \frac{2\,(4\,\delta-1)}{9\,(1+8\,\delta)(1+9\,\delta)}.$$

which is greater than zero if and only if $\delta > \frac{1}{4}$. Thus, if $\delta > \frac{1}{4}$ then the three-point $D$-optimal individual design is (-2,0,2), when $\delta < \frac{1}{4}$ (-2, -1, 2) or (-2, 1, 2) is the three-point $D$-optimal individual design and when $\delta = \frac{1}{4}$ (-2,0,2), (-2,-1,2) and (-2, 1, 2) are $D$-optimal.

Finally, the determinants of the information matrices for $\boldsymbol{\beta}$ for the four-point individual designs in the transformed coordinates are

$$|\mathbf{I}_\beta((-2,-1,0,1))| = |\mathbf{I}_\beta((-1,0,1,2))| = \frac{5}{4\,(1+6\,\delta)},$$

$$|\mathbf{I}_\beta((-2,-1,0,2))| = |\mathbf{I}_\beta((-2,0,1,2))| = \frac{35}{16\,(1+9\,\delta)}$$

and

$$|\mathbf{I}_\beta((-2,-1,1,2))| = \frac{5}{2\,(1+10\,\delta)}.$$

Thus (-2,-1,1,2) is the four-point $D$-optimal individual design for all $\delta \geq 0$.

The results of this example illustrate the fact that exact $D$-optimal individual designs for estimation of the fixed effects in the simple linear regression model with a random slope are not necessarily invariant to linear transformation. It is clear from the example that the optimal designs with a change of locations $\tilde{t}_j = t_j - \frac{k}{2}$ can be different and do not always map onto each other.

Note that since $|\mathbf{I}_\theta(\mathbf{t}_i)| = 0$, the $D$-optimal individual designs can not be computed for $\boldsymbol{\theta}$.

### 8.2.3 $D$-optimal population designs for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$

At a population design

$$\xi = \begin{Bmatrix} \mathbf{t}_1, & \ldots, & \mathbf{t}_r \\ w_1, & \ldots, & w_r \end{Bmatrix} \quad \text{with } 0 < w_i < 1 \text{ and } \sum_{i=1}^{r} w_i = 1$$

where $\mathbf{t}_i = (t_{i1}, \ldots, t_{id})'$ with $t_{i1} < t_{i2} < \ldots < t_{id}$, $i = 1, 2, \ldots, r$, the information matrix for the parameters vector $\boldsymbol{\alpha}$ is given by

$$\mathbf{M}_\alpha(\xi) = \sum_{i=1}^{r} w_i \mathbf{M}_\alpha(\mathbf{t}_i)$$

where $\mathbf{M}_\alpha(\mathbf{t}_i)$ is the standardized information matrix for $\boldsymbol{\alpha}$ at $\mathbf{t}_i$. In model (8.4) the parameters vector $\boldsymbol{\alpha}$ denotes either the fixed effects $\boldsymbol{\beta}$ or the variance components $\boldsymbol{\theta}$. Then the $D$-optimal population design for estimation of $\boldsymbol{\alpha}$ maximizes

$$\Psi_D(\xi) = \ln \left| \sum_{i=1}^{r} w_i \mathbf{M}_\alpha(\mathbf{t}_i) \right| .$$

Furthermore, from the Equivalence Theorem introduced in Subsection 2.6.4 a design $\xi_D^*$ is $D$-optimal if and only if the directional derivative of $\Psi_D(\xi)$ at $\xi_D^*$ in the direction of $\mathbf{t}$ is less than or equal to zero, that is

$$\phi_D(\mathbf{t}, \xi^*_{D(\alpha)}) = tr\{\mathbf{M}_\alpha^{-1}(\xi^*_{D(\alpha)})\, \mathbf{M}_\alpha(\mathbf{t})\} - p \leq 0$$

for all individual designs $\mathbf{t}$ in the space of designs of interest, with equality holding at the support designs of $\xi^*_{D(\alpha)}$ where $p$ is the number of parameters in $\boldsymbol{\alpha}$. In model (8.4) $p = 2$ for both $\boldsymbol{\alpha} = \boldsymbol{\beta}$ and $\boldsymbol{\alpha} = \boldsymbol{\theta}$. Observe that $\frac{1}{\sigma_e^4}$ factors out in $\Psi_D(\xi)$ for both $\boldsymbol{\alpha} = \boldsymbol{\beta}$ and $\boldsymbol{\alpha} = \boldsymbol{\theta}$. Therefore $\sigma_e^2$ can be taken to be 1 without loss of generality. Recall from the previous subsection that the determinant $|\mathbf{I}_\beta(\mathbf{t})|$ depends on the variance ratio $\delta$ and since $\mathbf{M}_\beta(\mathbf{t}) = \frac{\mathbf{I}_\beta(\mathbf{t})}{d}$ the criterion $\Psi_D(\xi) = \ln|\sum_{i=1}^r w_i \mathbf{M}_\beta(\mathbf{t}_i)|$ also depends on $\delta$. Note that this is also true for $\Psi_D(\xi) = \ln|\sum_{i=1}^r w_i \mathbf{M}_\theta(\mathbf{t}_i)|$. Therefore the optimum designs cannot be found algebraically, at least in general, but can be obtained numerically for a best guess of the parameter $\delta$. GAUSS programs were written to compute the $D$-optimal population designs for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ for a given value of $\delta$ and the programs are given in the files labelled "dslope" and "dvarslop" respectively on the CD provided with this thesis.

**Example 8.2.2** Consider the simple linear regression model with a random slope as specified by (8.4) and let $k = 4$ and $\delta = 0.1$. The $D$-optimal population designs for $\boldsymbol{\beta}$ over the sets of one-, two-, three- and four-point individual designs are

$$\xi^*_{D_1(\beta)} = \left\{ \begin{array}{cc} (0) & (4) \\ 0.5 & 0.5 \end{array} \right\}, \qquad \xi^*_{D_2(\beta)} = \left\{ \begin{array}{c} (0,4) \\ 1 \end{array} \right\},$$

$$\xi^*_{D_3(\beta)} = \left\{ \begin{array}{c} (0,1,4) \\ 1 \end{array} \right\} \quad \text{and} \quad \xi^*_{D_4(\beta)} = \left\{ \begin{array}{c} (0,1,2,4) \\ 1 \end{array} \right\}$$

respectively. Suppose now that the time points in the above example are transformed as $\tilde{t}_j = t_j - 2, j = 1, \ldots, d$. The $D$-optimal population designs for $\boldsymbol{\beta}$ over the sets of one-, two-, three- and four-point individual designs in the transformed coordinates are given by

$$\tilde{\xi}^*_{D_1(\beta)} = \left\{ \begin{array}{cc} (-2) & (2) \\ 0.5 & 0.5 \end{array} \right\}, \qquad \tilde{\xi}^*_{D_2(\beta)} = \left\{ \begin{array}{c} (-2,2) \\ 1 \end{array} \right\},$$

$$\tilde{\xi}^*_{D3(\beta)} = \left\{ \begin{array}{cc} (-2,-1,2) & (-2,1,2) \\ 0.5 & 0.5 \end{array} \right\} \quad \text{and} \quad \tilde{\xi}^*_{D4(\beta)} = \left\{ \begin{array}{c} (-2,-1,1,2) \\ 1 \end{array} \right\}$$

respectively. As for the exact $D$-optimal individual designs of the previous subsection, the numerical results show that the $D$-optimal population designs for $\boldsymbol{\beta}$ in the random slope models with a change of location $\tilde{t} = t - \frac{k}{2}$ do not always map onto each other.

These designs were shown numerically to be the $D$-optimal population designs by invoking the appropriate Equivalence Theorem. For instance, the plot of the directional derivative $\phi_D(\mathbf{t}, \xi^*_{D3(\beta)})$ against the individual designs (0, 1, 2), (0, 1, 3), (0, 1, 4), (0, 2, 3), (0, 2, 4), (0, 3, 4), (1, 2, 3), (1, 2, 4), (1, 3, 4) and (2, 3, 4), labelled 1 through 10, is presented in Figure 8.1. Similarly, the plot of the directional derivative $\phi_D(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D3(\beta)})$ against the individual designs (-2, -1, 0), (-2, -1, 1), (-2, -1, 2), (-2, 0, 1), (-2, 0, 2), (-1, 1, 2), (-1, 0, 1), (-1, 0, 2), (-1, 1, 2) and (0, 1, 2), labelled 1 through 10, is presented in Figure 8.2. Note that in both cases the maxima, which are equal to zero, occurred at the support design(s).

**Example 8.2.3** Consider again the simple linear regression model with a random slope as specified by (8.4) and let $k = 4$ and $\delta = 0.1$. The $D$-optimal population designs for $\boldsymbol{\theta} = (\sigma_e^2, \delta)$ over the sets of two-, three- and four-point individual designs were found to be

$$\xi^*_{D2(\theta)} = \left\{ \begin{array}{cc} (0,1) & (3,4) \\ 0.5 & 0.5 \end{array} \right\}, \quad \xi^*_{D3(\theta)} = \left\{ \begin{array}{cc} (0,1,2) & (2,3,4) \\ 0.5 & 0.5 \end{array} \right\}$$

and

$$\xi^*_{D4(\theta)} = \left\{ \begin{array}{cc} (0,1,2,3) & (1,2,3,4) \\ 0.5 & 0.5 \end{array} \right\}$$

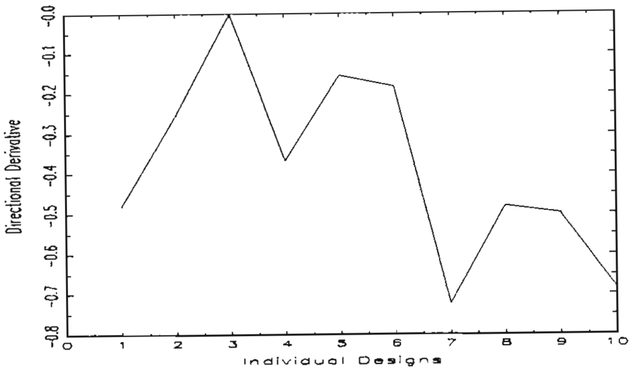


Figure 8.1: Plot of the directional derivative $\phi(\mathbf{t}, \xi^*_{D_3})$ against the three-point individual designs $\mathbf{t}$ for $\delta = 0.1$.

respectively. In contrast the $D$-optimal population designs for $\boldsymbol{\theta} = (\sigma_e^2, \delta)$ over the set of two-, three- and four-point individual designs in the transformed coordinates are

$$\tilde{\xi}^*_{D_2(\theta)} = \left\{ \begin{array}{cc} (-2,2) & (-1,0) \\ 0.5 & 0.5 \end{array} \right\}, \; \tilde{\xi}^*_{D_3(\theta)} = \left\{ \begin{array}{ccc} (-2,-1,2) & (-2,1,2) & (-1,0,1) \\ 0.298 & 0.202 & 0.5 \end{array} \right\}$$

and

$$\tilde{\xi}^*_{D_4(\theta)} = \left\{ \begin{array}{cc} (-2,-1,0,1) & (-2,-1,1,2) \\ 0.5 & 0.5 \end{array} \right\}$$

respectively. Observe that the optimum designs $\tilde{\xi}^*_{D_2(\theta)}$, $\tilde{\xi}^*_{D_3(\theta)}$ and $\tilde{\xi}^*_{D_4(\theta)}$ are not the linearly transformed versions of $\xi^*_{D_2(\theta)}$, $\xi^*_{D_3(\theta)}$ and $\xi^*_{D_4(\theta)}$. That is, a situation similar to that for the $D$-optimal population designs for $\boldsymbol{\beta}$ is observed here. As for the $D$-optimal population designs for $\boldsymbol{\beta}$, the optimality of the above designs was shown numerically by invoking the appropriate Equivalence Theorem. For instance, the plot of the directional derivative $\phi_D(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_3(\theta)})$ against the individual designs (-2, -1, 0), (-2, -1, 1), (-2, -1, 2), (-2, 0, 1), (-2, 0,

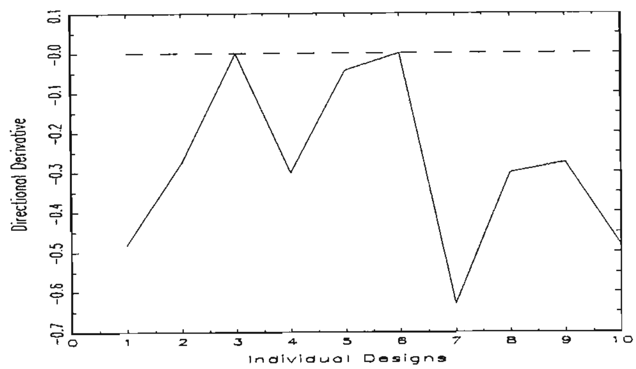


Figure 8.2: Plot of the directional derivative $\phi(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_3})$ against the three-point individual designs $\tilde{\mathbf{t}}$ for $\delta = 0.1$.

2), (-1, 1, 2), (-1, 0, 1), (-1, 0, 2), (-1, 1, 2) and (0, 1, 2), labelled 1 through 10 respectively, is presented in Figure 8.3. The figure shows that there are three maxima for $\phi_D(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D_3(\theta)})$ which are equal to zero and they occurred at the support designs of $\tilde{\xi}^*_{D_3(\theta)}$.

## 8.3 $D$-optimal population designs for the simple linear random coefficient regression model

In this section, $D$-optimum designs for the simple linear random coefficient regression model, which is the simple linear regression model with both intercept and slope random, are discussed. It is interesting to note that in this case, as for the model setting in the previous section, the $D$-optimality criteria are functions of the variance ratios. Thus, the optimal

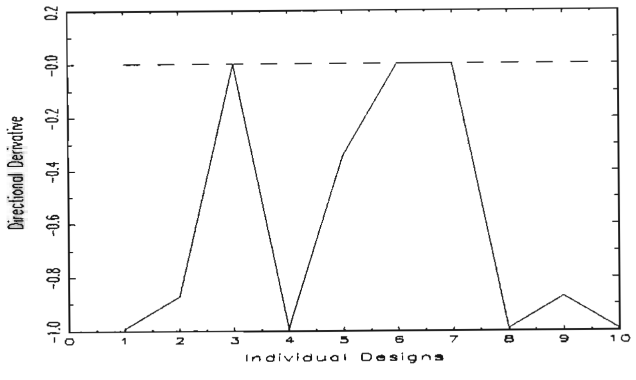


Figure 8.3: Plot of the directional derivative $\phi_D(\tilde{\mathbf{t}}, \tilde{\xi}^*_{D3(\theta)})$ against the three-point individual designs $\tilde{\mathbf{t}}$ for $\delta = 0.1$.

designs derived here are locally optimal. Recall that a linear transformation of the explanatory variables in the random coefficient model may change the structure of the variance matrix of the random effects and as a result the underlying model may also change.

### 8.3.1 Model

Recall from expression (8.1) that the simple linear random coefficient regression model for the $j$th observation on the $i$th individual at the time point $t_{ij}$ is given by

$$y_{ij} = (\beta_o + b_{0i}) + (\beta_1 + b_{1i})\, t_{ij} + e_{ij}, \quad j = 1, 2, \ldots, d_i \text{ and } i = 1, 2, \ldots, K. \qquad (8.7)$$

where $y_{ij}$ is the $j$th observation on individual $i$, $t_{ij} \in \{0, 1, \ldots, k\}$ with $k \geq 1$ and $K$ is the number of individuals. The intercept $\beta_o$ and slope $\beta_1$ are the fixed effects. It is assumed that $b_{0i} \sim \mathcal{N}(0, \sigma^2_{b_0})$, that $b_{1i} \sim \mathcal{N}(0, \sigma^2_{b_1})$ , that $e_{ij} \sim \mathcal{N}(0, \sigma^2_e)$ and that $Cov(e_{ij}, b_{0i}) =$

$Cov(e_{ij}, b_{1i}) = Cov(e_{ij}, e_{ij'}) = 0$ and $Cov(b_{0i}, b_{1i}) = \sigma_{b_0 b_1}$. Furthermore, under the linear mixed model formulation of Chapter 2,

$$\mathbf{X}_i = \mathbf{Z}_i = \begin{pmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \vdots \\ 1 & t_{id_i} \end{pmatrix},$$

$\boldsymbol{\beta} = (\beta_o, \beta_1)'$ and $\mathbf{b}_i = (b_{0i}, b_{1i})'$. Thus the expectation vector and variance-covariance matrix of $\mathbf{y}_i = (y_{i1} \;\; y_{i2} \ldots \; y_{id_i})'$ are $E(\mathbf{y}_i) = \mathbf{X}_i \boldsymbol{\beta}$ and $Var(\mathbf{y}_i) = \mathbf{V}_i = \mathbf{X}_i \, \mathbf{G} \, \mathbf{X}_i' + \sigma_e^2 \, \mathbf{I}_{d_i}$ respectively, where

$$\mathbf{G} = \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma_{b_1}^2 \end{pmatrix}.$$

The appropriate information matrices for the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ in the simple linear random coefficient regression model at an individual design $\mathbf{t}_i$ can be obtained from the results for the linear mixed model and they are given below. Consider first the information matrix for $\boldsymbol{\beta}$. The model (8.7) is of the form

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \, \mathbf{b}_i + \mathbf{e}_i$$

with $\mathbf{X}_i = \mathbf{Z}_i$. Therefore the information matrix for $\boldsymbol{\beta}$ for the $i$th individual is given by expression (2.10), that is by

$$\mathbf{I}_{\beta}(\mathbf{t}_i) = \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i = \begin{pmatrix} \mathbf{1}_{d_i}' \, \mathbf{V}_i^{-1} \, \mathbf{1}_{d_i} & \mathbf{1}_{d_i}' \, \mathbf{V}_i^{-1} \, \mathbf{t}_i \\ \mathbf{t}_i' \, \mathbf{V}_i^{-1} \, \mathbf{1}_{d_i} & \mathbf{t}_i' \, \mathbf{V}_i^{-1} \, \mathbf{t}_i \end{pmatrix},$$

with $\mathbf{X}_i = (\mathbf{1}_d, \; \mathbf{t}_i)$, $\mathbf{V}_i^{-1} = \dfrac{1}{\sigma_e^2} (\mathbf{X}_i \, \mathbf{G}_\gamma \, \mathbf{X}_i' + \mathbf{I}_{d_i})^{-1}$, $i = 1, \ldots, K$ and

$$\mathbf{G}_\gamma = \begin{pmatrix} \gamma_{b_0} & \gamma_{b_0 b_1} \\ \gamma_{b_0 b_1} & \gamma_{b_1} \end{pmatrix},$$

where $\gamma_{b_0} = \dfrac{\sigma_{b_0}^2}{\sigma_e^2}$, $\gamma_{b_1} = \dfrac{\sigma_{b_1}^2}{\sigma_e^2}$ and $\gamma_{b_0b_1} = \dfrac{\sigma_{b_0b_1}}{\sigma_e^2}$. Thus the standardized information matrix for $\boldsymbol{\beta}$ at the individual design $\mathbf{t}_i = (t_{i1}, \ldots, t_{id_i})'$ is given by

$$\mathbf{M}_\beta(\mathbf{t}_i) = \frac{\mathbf{I}_\beta(\mathbf{t}_i)}{d_i}.$$

Consider now the information matrix for $\boldsymbol{\theta}$. Assume that $\sigma_e^2$ is a nuisance parameter. Then the variance components in $\boldsymbol{\theta}$ are specified by $\gamma_{b_0}, \gamma_{b_1}$ and $\gamma_{b_0b_1}$. Recall from expression (2.14) that the information matrix of the variance components $\boldsymbol{\theta}$ from the $i$th individual corresponding to parameters $\theta_j$ and $\theta_k$ is

$$\mathbf{I}_\theta(\mathbf{X}_i) = \frac{1}{2}\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_l}\right), i = 1, \ldots, K.$$

For $\mathbf{V} = \sigma_e^2\,(\mathbf{X}\,\mathbf{G}_\gamma\,\mathbf{X}' + \mathbf{I})$ and ignoring the index $i$ for an individual for convenience,

$$\frac{\partial \mathbf{V}}{\partial \gamma_{b_0}} = \sigma_e^2\,\mathbf{X}\,\frac{\partial \mathbf{V}}{\partial \gamma_{b_0}}\mathbf{X}' = \sigma_e^2\,\mathbf{1}_d\,\mathbf{1}_d',$$

$$\frac{\partial \mathbf{V}}{\partial \gamma_{b_1}} = \sigma_e^2\,\mathbf{X}\,\frac{\partial \mathbf{V}}{\partial \gamma_{b_1}}\mathbf{X}' = \sigma_e^2\,\mathbf{t}\,\mathbf{t}'$$

and

$$\frac{\partial \mathbf{V}}{\partial \gamma_{b_0b_1}} = \sigma_e^2\,\mathbf{X}\,\frac{\partial \mathbf{V}}{\partial \gamma_{b_0b_1}}\mathbf{X}' = \sigma_e^2\,(\mathbf{t}\,\mathbf{1}_d' + \mathbf{1}_d\,\mathbf{t}')$$

since $\mathbf{X} = (\mathbf{1}_d,\ \mathbf{t})$ for model (8.7). Therefore the elements of the information matrix for $\boldsymbol{\theta}$ are given by

$$h_{(\gamma_{b_0})^2} = \frac{1}{2}tr\left(\mathbf{V}^{-1}\sigma_e^2\,\mathbf{1}_d\,\mathbf{1}_d'\right)^2 = \frac{1}{2}\left(\mathbf{I}_\beta(\mathbf{t})\right)_{11}^2,$$

$$h_{(\gamma_{b_1})^2} = \frac{1}{2}tr\left(\mathbf{V}^{-1}\sigma_e^2\,\mathbf{t}\,\mathbf{t}'\right)^2 = \frac{1}{2}\left(\mathbf{I}_\beta(\mathbf{t})\right)_{22}^2,$$

$$h_{(\gamma_{b_0b_1})^2} = \frac{1}{2}tr\left\{\mathbf{V}^{-1}\sigma_e^2\,(\mathbf{t}'\,\mathbf{1}_d + \mathbf{1}_d'\,\mathbf{t})\right\}^2 = (\mathbf{I}_\beta(\mathbf{t}))_{12}^2 + (\mathbf{I}_\beta(\mathbf{t}))_{11}\,(\mathbf{I}_\beta(\mathbf{t}))_{22},$$

$$h_{(\gamma_{b_0},\gamma_{b_1})} = \frac{1}{2}tr\left(\{\mathbf{V}^{-1}(\sigma_e^2\,\mathbf{1}_d\,\mathbf{1}_d')\,\mathbf{V}^{-1}(\sigma_e^2\,\mathbf{t}\,\mathbf{t}')\}\right) = \frac{1}{2}\left(\mathbf{I}_\beta(\mathbf{t})\right)_{12}^2,$$

$$h_{(\gamma_{b_0},\gamma_{b_0b_1})} = \frac{1}{2}tr\left\{\mathbf{V}^{-1}(\sigma_e^2\,\mathbf{1}_d\,\mathbf{1}_d')\,\mathbf{V}^{-1}(\mathbf{t}\,\mathbf{1}_d' + \mathbf{1}_d\,\mathbf{t}')\right\} = (\mathbf{I}_\beta(\mathbf{t}))_{11}\,(\mathbf{I}_\beta(\mathbf{t}))_{12},$$

and

$$h_{(\gamma_{b_1},\gamma_{b_0b_1})} = \frac{1}{2}tr\left\{\mathbf{V}^{-1}(\sigma_e^2\,\mathbf{t}\,\mathbf{t}')\,\mathbf{V}^{-1}(\mathbf{t}\,\mathbf{1}_d' + \mathbf{1}_d\,\mathbf{t}')\right\} = (\mathbf{I}_\beta(\mathbf{t}))_{22}\,(\mathbf{I}_\beta(\mathbf{t}))_{12}.$$

Thus on assembling the above results, the information matrix for $\boldsymbol{\theta} = (\gamma_{b_0}, \gamma_{b_1}, \gamma_{b_0b_1})$ at the $i$th individual design $\mathbf{t}_i$ can be written as

$$\mathbf{I}_\theta(\mathbf{t}_i) = \begin{pmatrix} h_{(\gamma_{b_0})^2} & h_{(\gamma_{b_0},\gamma_{b_1})} & h_{(\gamma_{b_0},\gamma_{b_0b_1})} \\ h_{(\gamma_{b_0},\gamma_{b_1})} & h_{(\gamma_{b_1})^2} & h_{(\gamma_{b_1},\gamma_{b_0b_1})} \\ h_{(\gamma_{b_0},\gamma_{b_0b_1})} & h_{(\gamma_{b_1},\gamma_{b_0b_1})} & h_{(\gamma_{b_0b_1})^2} \end{pmatrix}$$

and hence for the $i$th individual the standardized information matrix for $\boldsymbol{\theta}$ is given by

$$\mathbf{M}_\theta(\mathbf{t}_i) = \frac{\mathbf{I}_\theta(\mathbf{t}_i)}{d_i}$$

for $i = 1, \ldots, K$.

### 8.3.2 $D$-optimal population designs for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$

Consider the information matrix for the parameters vector $\boldsymbol{\alpha}$ at a population design $\xi$, i.e. $\mathbf{M}_\alpha(\xi) = \sum_{i=1}^{r} w_i\,\mathbf{M}_\alpha(\mathbf{t}_i)$, where $\boldsymbol{\alpha}$ denotes the fixed effects $\boldsymbol{\beta}$ or the variance components $\boldsymbol{\theta} = (\gamma_{b_0}, \gamma_{b_1}, \gamma_{b_0b_1})$. Then the $D$-optimal population design for $\boldsymbol{\alpha}$ maximizes

$$\Psi_D(\xi) = \ln\left|\sum_{i=1}^{r} w_i\,\mathbf{M}_\alpha(\mathbf{t}_i)\right|$$

over the set of population designs. Furthermore, it follows from the Equivalence Theorem given in Theorem 2.6.1 that the design $\xi^*_{D(\alpha)}$ is the $D$-optimal population design if and only if

$$\phi_D(\mathbf{t}, \xi^*_{D(\alpha)}) = tr\{\mathbf{M}_\alpha^{-1}(\xi^*_{D(\alpha)})\,\mathbf{M}_\alpha(\mathbf{t})\} - p \leq 0$$

with equality holding at the support designs of $\xi^*_{D(\alpha)}$, where $p$ is the number of parameters in $\boldsymbol{\alpha}$. In model (8.7) $p = 2$ when $\boldsymbol{\alpha}$ denotes $\boldsymbol{\beta}$ and $p = 3$ when $\boldsymbol{\alpha}$ denotes $\boldsymbol{\theta} = (\gamma_{b_0}, \gamma_{b_1}, \gamma_{b_0b_1})$. Note that $\sigma_e^2$ factors out in the information matrices $\mathbf{M}_\beta(\mathbf{t}_i)$ and $\mathbf{M}_\theta(\mathbf{t}_i)$ and thus is taken to be 1 without the loss of generality.

Recall from the previous section that the simple linear regression model with a random slope, i.e. model (8.4), is a special case of model (8.7) and that the $D$-optimality criterion for the parameters in that model is a function of the variance ratio $\delta$. As for model (8.4) the criterion $\Psi_D(\xi) = \ln|\sum_{i=1}^{r} w_i\,\mathbf{M}_\alpha(\mathbf{t}_i)|$ for the parameters in model (8.7) is a function of the variance ratios $\gamma_{b_0}, \gamma_{b_1}$ and $\gamma_{b_0b_1}$ and thus the optimum designs for the precise estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ cannot be found algebraically. Thus the optimal designs of this section were found as locally optimal designs for best guesses of $\gamma_{b_0}, \gamma_{b_1}$ and $\gamma_{b_0b_1}$. The $D$-optimal designs for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ for a given variance-covariance matrix of the random effects, $\mathbf{G}_\gamma$ were computed numerically using the GAUSS programs "dbetarc" and "dvarrcm", respectively. The programs are on the CD provided with this thesis. The variance-covariance matrix

$$\mathbf{G}_\gamma = \begin{pmatrix} 1 & -0.05 \\ -0.05 & 0.25 \end{pmatrix}$$

was used to illustrate the numerical results.

**Example 8.3.1** Consider the simple linear random coefficient regression model in (8.7) with $k = 4$ and $\mathbf{G}_\gamma = \begin{pmatrix} 1 & -0.05 \\ -0.05 & 0.25 \end{pmatrix}$. The $D$-optimal population designs for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$

over the set of three-point individual designs $\mathbf{t} = (t_1, t_2, t_3)$ with $t_1 < t_2 < t_3$ were found to be the same and are given by

$$\xi^*_{D3(\beta)} = \left\{ \begin{array}{c} (0,1,4) \\ 1 \end{array} \right\} = \xi^*_{D3(\theta)}.$$

Suppose that the time points in the above example are transformed as $\tilde{t}_j = t_j - 2, j = 1,2,3$. The $D$-optimal population designs for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ over the set of three-point individual designs $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2, \tilde{t}_3)$ with $\tilde{t}_1 < \tilde{t}_2 < \tilde{t}_3$ are

$$\tilde{\xi}^*_{D3(\beta)} = \left\{ \begin{array}{c} (-2,1,2) \\ 1 \end{array} \right\}$$

and

$$\tilde{\xi}^*_{D3(\theta)} = \left\{ \begin{array}{cc} (-2,-1,2) & (-2,1,2) \\ 0.013 & 0.987 \end{array} \right\}$$

respectively. Observe here that the optimum designs $\tilde{\xi}^*_{D3(\beta)}$ and $\tilde{\xi}^*_{D3(\theta)}$ are not the linearly transformed version of $\xi^*_{D3(\beta)}$ and $\xi^*_{D3(\theta)}$ respectively. The plots of the directional derivatives $\phi_D(\mathbf{t}, \xi^*_{D3(\beta)})$ and $\phi_D(\mathbf{t}, \xi^*_{D3(\theta)})$ against the individual designs (-2, -1, 0), (-2, -1, 1), (-2, -1, 2), (-2, 0, 1), (-2, 0, 2), (-1, 1, 2), (-1, 0, 1), (-1, 0, 2), (-1, 1, 2) and (0, 1, 2), labelled 1 through 10, are presented in Figures 8.4 and 8.5. It is clear from the figures that the conditions $\phi_D(\mathbf{t}, \xi^*_{D3(\beta)}) \leq 0$ and $\phi_D(\mathbf{t}, \xi^*_{D3(\theta)}) \leq 0$ are satisfied for all designs $\tilde{\mathbf{t}}$ and that in both cases equality holds at the support designs. Thus the designs are $D$-optimal.

The optimal designs were computed numerically for various values of $\mathbf{G}_\gamma$ and they change with $\mathbf{G}_\gamma$. In general, the optimal designs for the simple linear random coefficient regression model specified by (8.7) computed over the set of individual designs $\mathbf{t}$ with $t_1 < t_2 < \ldots < t_d$

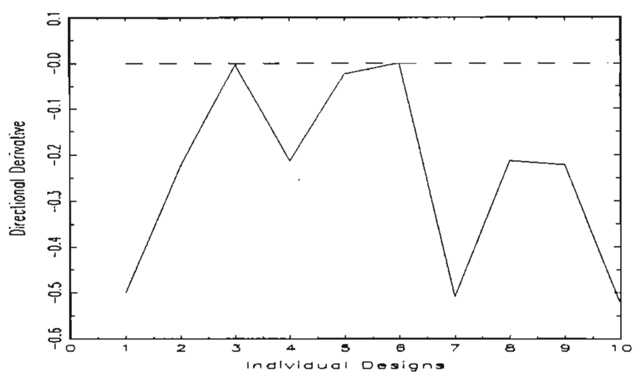


Figure 8.4: Plot of the directional derivative $\phi(\mathbf{t}, \tilde{\xi}^*_{D3(\beta)})$ against the three-point individual designs $\tilde{\mathbf{t}}$.

for a given $\mathbf{G}_\gamma$ are optimal for the simple linear random coefficient regression model with a particular $\mathbf{G}^*_\gamma = \mathbf{A}^{-1}\,\mathbf{G}_\gamma\,(\mathbf{A}^{-1})$ over the set of individual designs $\tilde{\mathbf{t}}$ with $\tilde{t}_1 < \tilde{t}_2 < \ldots < \tilde{t}_d$, and vice versa. For instance, in the above example the three-point design $\xi^*_{D3(\beta)}$ is optimal for model (8.7) with

$$\mathbf{G}_\gamma = \begin{pmatrix} 1 & -0.05 \\ -0.05 & 0.25 \end{pmatrix} \quad \text{and} \quad \mathbf{G}^*_\gamma = \begin{pmatrix} 1.8 & 0.45 \\ 0.45 & 0.25 \end{pmatrix}.$$

## 8.4 Conclusions

In this chapter optimal designs for simple linear regression models with a random slope and with both intercept and slope random are investigated numerically. Unlike the optimal designs of Chapters 4 and 5, the optimal designs depend on the values of the variance components and thus were calculated numerically for the given values of variance components,

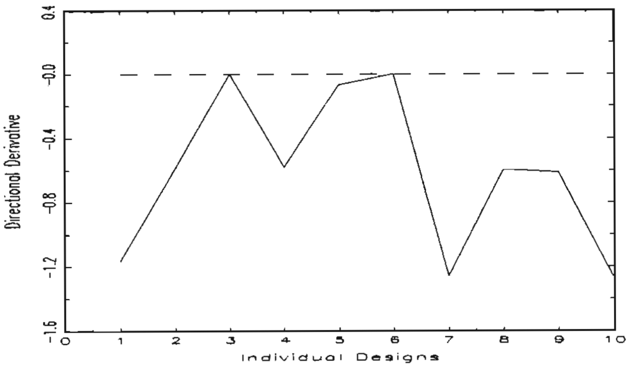


Figure 8.5: Plot of the directional derivative $\phi(\mathbf{t}, \tilde{\xi}^*_{D_3(\theta)})$ against the three-point individual designs $\tilde{\mathbf{t}}$.

i.e. they are locally optimal. Furthermore, it is observed that the $D$-optimal designs for these models with a change of location do not always map onto each other. Specifically, the optimal designs for the above models computed over the set of individual designs $\mathbf{t}$ with $t_1 < t_2 < \ldots < t_d$ for a given variance matrix of the random effects, $\mathbf{G}_\gamma$, are optimal for a particular $\mathbf{G}^*_\gamma = \mathbf{A}^{-1}\,\mathbf{G}_\gamma\,(\mathbf{A}^{-1})$ over the set of individual designs $\tilde{\mathbf{t}}$ with $\tilde{t}_1 < \tilde{t}_2 < \ldots < \tilde{t}_d$ and vice versa.

# Chapter 9

# Conclusions

In this thesis, $D$- and $V$-optimal designs for the linear mixed effects model have been investigated. This was done with special reference to longitudinal data, that is data measured repeatedly at time points. However, the results can be used for other applications which fit within the framework of this model.

In Chapters 4 and 5, $D$- and $V$-optimal designs were developed algebraically for the precise estimation of the fixed effects and of the mean responses at a given vector of time points in the simple linear regression model with a random intercept. The optimal designs are based on designs with non-repeated and repeated time points where the time points are assumed to be equally spaced. The results of these chapters show that (i) the $D$-optimal population designs based on $d$-point individual designs for estimation of the fixed effects are optimal for estimation of the slope parameter and also $V$-optimal for estimation of the mean responses at the time vector $\mathbf{t}_g = (0, 1, \ldots, k)'$, (ii) the $D$- and $V$-optimal population designs based on $d$-point individual designs with repeated time points are always more efficient than the corresponding $D$- and $V$-optimal population designs with non-repeated time points and (iii) the optimal designs are robust to the choice of variance ratio.

Chapters 6 and 7 contain the results of constructing $D$- and $V$-optimal population designs for the quadratic regression model with a random intercept. Only $D$-optimal population designs based on one- and two-point individual designs and the $V$-optimal population design based on one-point individual designs for even number of time points were computed algebraically. However GAUSS programs have been written to calculate numerically the optimal designs for any number of time points in individual designs $d$ $(1 \leq d \leq k+1)$ and for a given variance ratio $\gamma$. Except for the optimal designs based on one-point individual designs, both $D$- and $V$-optimal population designs for the quadratic regression model with a random intercept change with the variance ratio.

Linear transformations of the design matrices in the random coefficient model may result in structural changes in the variance matrices of random effects. With this in mind, $D$-optimal designs for parameters were calculated for the simple linear regression model with a random slope and for the simple linear random coefficient regression model numerically and are discussed in Chapter 8. The numerical results show that, in contrast to the random intercept model, the optimal designs for these models are not invariant to a linear transformation of time points. This means that if a design which is optimal for a random coefficient model is linearly transformed it is not necessarily optimal for the associated transformed random coefficient model.

The results of this thesis are based on the assumption that the variation of observations within individuals is the same, that is, it is assumed that the degree to which observations are correlated is equal for every pair of observations within an individual. However, it is common for observations on an individual measured at time points closer to each other are more correlated than observations measured at time points which are well separated. This knowledge could be employed for developing optimal designs within the framework of linear mixed models. Furthermore, the models used in this work were restricted to one explanatory variable, time. However, the linear mixed effects model can accommodate more variables. Therefore further research is also needed into the construction of optimal designs for such cases.

Apart from the above extensions, several other design problems for the linear mixed model are open for investigation. For example, optimal designs for simultaneous estimation of the fixed effects and the variance components, the $D$- and $V$-optimal population designs for higher degree polynomial regression models with a random intercept and the extension of the population designs to other criteria such as $A$ and $E$ criteria all merit attention.

# Bibliography


Abt, M., Gaffke, N., Liski, E.P., Sinha, B.K. (1998). Optimal designs in growth curve models: II Correlated model for quadratic growth: Optimal designs for slope parameter estimation and growth prediction. *Journal of Statistical Planning and Inference*, **67**, 287 – 296.

Abt, M., Liski, E.P., Mandal, N.K., Sinha, B.K. (1997). Optimal designs in growth curve models: I Correlated model for linear growth: Optimal designs for slope parameter estimation and growth prediction. *Journal of Statistical Planning and Inference*, **64**, 141 – 150.

Anderson, R.L. (1975). Designs and estimators for variance components. In *A Survey of Statistical Design and Linear Models*, 1 – 29. Amsterdam: North Holland.

Anderson, R.L. (1981). Recent developments in designs and estimators for variance components. In *Statistics and Related Topics*, 3 – 22. Amesterdam: North Holland.

Atkins, J.E. and Cheng, C.S. (1999). Optimal regression designs in the presence of random block effects. *Journal of Statistical Planning and Inference*, **77**, 321 – 335.

Atkinson A.C. and Donev A.N. (1989). The construction of exact $D$-optimum experimental designs with application to blocking response surface designs. *Biometrika*, **76**, $515 - 526$.

Atkinson A.C. and Donev A.N. (1992). *Optimum Experimental Designs.* Oxford University Press: Oxford.

Azzalini, A. (1996). *Statistical Inference: Based on the likelihood.* Chapman and Hall: London.

Bagchi, S. (1987). On the optimalities of the MBGDD's under mixed effects model. *Communication in Statistics A, Theory and Methods*,**16**, $3565 - 3576$.

Bazaraa, M.S. and Shety, C.M. (1979). *Nonlinear Programming: Theory and Algorithms.* Wiley: New York.

Bischoff, W. (1993). On $D$-optimal designs for linear models under correlated observations with an application to a linear model with multiple response. *Journal of Statistical Planning and Inference*, **37**, $69 - 805$.

Brown, H.K. and Prescott, R. (1999). *Applied Mixed Models in Medicine.* Wiley: New York.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review. *Statistical Science*, **10**, $273 - 304$.

Cheng, C.S. (1995). Optimal regression designs under random block-effects models.

*Statistica Sinica*, **5**, 485 – 497.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics*, **24**, 586 – 602.

Cook, R.D. and Nachtsheim. C.J. (1980). A comparison of algorithms for constructing exact $D$-optimal designs. *Technometrics*, **22**, 315 – 324.

Cook, W.J., Cunningham W.H., Pulleyblank W.R. and Schrijver A. (1998). *Combinatorial Optimization.* Wiley: New York.

Davidian, M. and Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measurement Data.* Chapman and Hall: London.

Diggle, P.J., Liang, K-Y and Zeger, S.L. (1994). *Analysis of Longitudinal Data* . Oxford University Press: Oxford.

Duchateau, L., Janssen, P. and Rowlands, G.J. (1998). *Linear mixed models: An introduction with applications in veterinary research.* International Livestock Research Institute, Nairobi, Kenya.

Ehrenfeld, S. (1955). On the efficiency of experimental designs. *Annals of Mathematical Statistics*, **26**, 247 – 255.

Eisenhart,C. (1947). The assumptions underlying the analysis of variance. *Biometrics*,**3**, 1 – 21.

Fedorov, V.V. (1972). *Theory of Optimal Experiments.* Academic Press: New York and London.

Fedorov, V.V. and Hackl, P. (1997). *Model-Oriented Design of Experiments.* Springer: New York.

Feller, W. (1968). *An Introduction to Probability Theory and its Applications.* Vol. I., Wiley: New York.

Giovagnoli, A. and Sebastiani, P. (1989). Experimental designs for mean and variance estimation in variance components models. *Computational Statistics and Data Analysis*,**8**, 21 – 28.

Goldstein, H. (1987). *Multilevel Models in Educational and Social Research.* Griffin, London.

Goos, P. and Vandebroek, M. (2001). *D*-optimal response surface designs in the presence of random block effects. *Computational Statistics and Data Analysis*, **37**, 433 – 453.

Goos, P. (2002). *The Optimal Design of Blocked and Split-Plot Experiments.* Springer-Verlag: New York.

Guest, P.G. (1958). The spacing of observations in polynomial regression. *Annals of Mathematical Statistics*, **29**, 294 – 299.

Guiard, G.H., Sumpf, D. and Nürnberg, G. (2000). About optimal designs for estimating variance components with ANOVA in one-way-classification under non-normality.

*Journal of Statistical Planning and Inference*, **89**, 269 – 285.

Hartley, H.O. and Rao, J.N.K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika*, **54**, 93 – 108.

Harville, D.A. (1974). Bayesian inference for variance components using only error contrasts. *Biometrika*, **61**, 383 – 385.

Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective.* Springer-Verlag: New York.

Healy, M.J.R. (1986). *Matrices for Statistics.* Clarendon Press: Oxford.

Henderson, C.R. (1953). Estimation of variance and covariance components. *Biometrics*, **9**, 226 – 252.

Henderson, C.R., Kempthorne, O., Searle, S.R., and Krosigk, C.M. (1959). The estimation of environmental and genetic trends from records subject to culling. *Biometrics*, **15**, 192 – 218.

Hoel, P.G. (1958). Efficiency problems in polynomial estimation. *Annals of Mathematical Statistics*, **29**, 1134 – 1145.

Jones, B. and Wang, J. (1999). Constructing optimal designs for fitting pharmacokinetic models. *Statistics and Computing*, **9**, 209 – 218.

Jones, B., Wang, J., Jarvis, P. and Byrom, W. (1999). Design of cross-over trials for pharmacokinetic studies. *Journal of Statistical Planning and Inference*, **78**, 307 – 316.

Khuri, A.I. (2000). Design for variance components estimation: Past and present. *International Statistical Review*, **68**, 311 – 322.

Kiefer, J. (1959). Optimum experimental designs. *Journal of Royal Statistical Society (B)*, **21**, 272 – 304.

Kiefer, J. (1961). Optimum designs in regression problems, II. *Annals of Mathematical Statistics*, **32**, 298 – 325.

Kiefer, J. (1962). An extremum result. *Canadian Journal of Mathematics*, **14**, 597 – 601.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Annals of Statistics*, **2**, 849 – 879.

Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems, I. *Annals of Mathematical Statistics*, **30**, 271 – 294.

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

Laird, N.M. and Ware, J.H. (1982). Random effects models for longitudinal data. *Biometrics.*, **38**, 963 – 974.

Lee, Y. and Nelder, J.A. (1996). Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society, Series (B)*, **58**, 619 – 678.

Lindsey, J.K. (1993). *Models for Repeated Measurements.* Oxford University Press:

Oxford.

Lindstrom, M.J. and Bates, D.M. (1988). Newton-Raphson and EM algorithms. *Journal of the American Statistical Association*, **83**, 1014 – 1022

Littell, R.C., Milliken, G.A., Stroup, W.W. and Wolfinger, R.D. (1996). *SAS System for Mixed Models*. SAS Institute, Cary, North Carolina.

Lohr, S.L. (1995). Optimal Bayesian design of experiments for the one-way random effects model. *Biometrika*, **82**, 175 – 186.

Longford, N.T. (1993). *Random Coefficient Models*. Oxford University Press: Oxford.

Mallet, A. and Mentré, F. (1988). An approach to the design of experiments for estimating the distribution of parameters in random models. In *Proceedings of the 12th IMACS World Congress. Villeneuve d'Asq, Edition Gerdfin-Sité Scientifique*, 5, 134 – 137.

McCulloch, C.E. and Searle, S.R. (2001). *Generalized, Linear and Mixed Models*. Wiley: New York.

Mejza, S. and Kageyama, S. (1995). On the optimality of certain nested block designs under a mixed effects model. In *Proceedings of MODA4*, Physica Verlag, Heidelberg, Ed. Kitsos, C.P. and Müller, W.G., pages 157-164.

Mentré, F., Burtin, P., Merlé, Y., van Bree, J. and Steimer, J.L. (1995). Sparse-sampling optimal designs in pharmacokinetics and toxicokinetics. *Drug Information*

*Journal*, **29**, 997 – 1019.

Mentré, F., Mallet, A. and Baccar, D. (1997). Optimal design in random effects regression models. *Biometrika*, **84**, 429 – 442.

Moerbeek, M., Van Breukelen, G.J.P. and Berger, M.P.F. (2001). Optimal experimental designs for multilevel logistic models. *The Statistician*, **50**, 17 – 30.

Moerbeek, M. and Ausems, M. (2003). Optimal sample sizes in experimental designs with individuals nested within clusters. *Understanding Statistics*, **2**, 151 – 175.

Morgan, J.P. (1997). Optimal design for interacting blocks with OAVS incidence. *Metrika*, **45**, 67 – 83.

Mukerjee, R. and Huda, S. (1988). Optimal design for the estimation of variance components. *Biometrika*, **75**, 75 – 80.

Ouwens, M.J.N.M., Tan, F.E.S. and Berger, M.P.F. (2002). Maxmin $D$-optimal designs for longitudinal mixed effects models. *Biometrics*, **58**, 735 – 741.

Patterson, H.D. and Thompson, R. (1971). Recovery of inter-block information when block size are unequal. *Biometrika*, **55**, 545 – 554.

Pukelsheim, F. (1983).On optimality properties of simple block designs in the approximate design theory. *Journal of Statistical Planning and Inference*, **8**, 193 – 208.

Pukelsheim, F. (1993). *Optimal Design of Experiments.* Wiley: New York.

Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, **79**, 763 – 770.

Robinson, G.K. (1991). That BLUP is a good thing-the estimation of random effects. *Statistical Science*. **6**, 15 – 51.

Searle, S.R. (1971). Topics in variance components estimation. *Biometrics*, **27**, 1 – 76.

Searle, S.R. (1987). *Linear Models for Unbalanced Data*. Wiley: New York.

Searle, S.R. (1988). Mixed models and unbalanced data: where from, whereat and whereto. *Communications in Statistics A: Theory and Methods* **17**, 935 – 968.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992). *Variance Components*. Wiley: New York.

Shah, K.R. and Sinha, J.N. (1989). *Theory of optimal designs*. Springer-Verlag: New York.

Silvey, S. D. (1980). *Optimal Design*. Chapman and Hall: London.

Smith, K. (1918). On the standard derivations of adjusted and interpolated values of an observed polynomial function and its constants and guidance they give towards a proper choice of the distribution of observations. *Biometrika*, **12**, 1 – 85.

Snijders, T.A.B. and Bosker, R.J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational Statistics*, **18**, 237 – 259.

Snijders, T.A.B. and Bosker, R.J. (1999). *Multilevel Analysis: An introduction to basic and advanced multilevel modeling.* SAGE: London.

Sprott, D.A. (1975). Marginal and conditional sufficiency. *Biometrika*, **62**, 599 – 605.

Studden, W.J., (1980). $D_s$-optimal designs for polynomial regression using continued fractions. *Annals of Statistics*, **8**, 1132 – 1141.

Tan, E.S. and Berger, M.P.F. (1999). Optimal allocation of time points for the random effects model. *Communications in Statistics - Simulations*,**28**, 517 – 540.

Thompson, R. (1977). The estimation of heritability with unbalanced data. I. Observations available on parents and offspring. *Biometrics*, **33**, 485 – 495.

Verbeke, G. and Lesaffre, E. (1999). The effect of drop-out on the efficiency of longitudinal experiments. *Applied Statistics*, **48**, 363 – 375.

Verbeke, G. and Molenberghs, G. (1999). *Linear Mixed Models for Longitudinal Data.* Springer-Verlag: New York.

Vonesh, E. F. and Chinchilli, V. M. (1996). *Linear and Nonlinear Models for the Analysis of Repeated Measurements.* Marcel Dekker: New York.

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

Wolfram, S. (1999). *The Mathematica Book*, 4th edition, Wolfram Media/Cambridge University Press: Cambridge.

Wynn, H.P. (1970). The sequential generation of $D$-optimum experimental designs. *Annals of Mathematical Statistics*, **41**, 1655 – 1664.

Yah, C.-M. (1988). A class of universally optimal binary block designs. *Journal of Statistical Planning and Inference*, **18**, 355 – 361.

# Appendix A

# Results in Matrix Algebra

## A.1 Kronecker product

Let $\mathbf{A}$ and $\mathbf{B}$ be matrices of dimension $r \times c$ and $s \times d$ respectively. Then the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$ is given as

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\,\mathbf{B} & a_{12}\,\mathbf{B} & \dots & a_{1c}\,\mathbf{B} \\ a_{21}\,\mathbf{B} & a_{22}\,\mathbf{B} & \dots & a_{2c}\,\mathbf{B} \\ \vdots & \vdots & \vdots & \\ a_{r1}\,\mathbf{B} & a_{r2}\,\mathbf{B} & \dots & a_{rc}\,\mathbf{B} \end{pmatrix}.$$

The order of $\mathbf{A} \otimes \mathbf{B}$ is $r\,s \times c\,d$.

## A.2 Results for inverses of matrices

### A.2.1 Result 1

Suppose $a$ and $b$ are nonzero constants. Then

$$(a\,\mathbf{I} + b\,\mathbf{J})^{-1} = \frac{1}{a}\{\mathbf{I} - \frac{b}{a + n\,b}\mathbf{J}\}$$

where $\mathbf{I}$ and $\mathbf{J}$ are the $n \times n$ identity matrix and the matrix having every element equal to unity respectively (Searle, Casella and McCulloch, 1992, page 443).

### A.2.2 Result 2

If $\mathbf{A}$ and $\mathbf{B}$ are positive-definite matrices and $0 < \alpha < 1$, then

1. $\alpha\,\mathbf{A}^{-1} + (1-\alpha)\,\mathbf{B}^{-1} \geq \{\alpha\,\mathbf{A} + (1-\alpha)\,\mathbf{B}\}^{-1}$ and

2. $|\alpha\,\mathbf{A} + (1-\alpha)\,\mathbf{B}| \geq |\mathbf{A}|^{\alpha}\,|\mathbf{B}|^{1-\alpha}$.

Moreover, the equality sign in 1 and 2 holds only if $\mathbf{A} = \mathbf{B}$ (Fedorov, 1972, page 19-20).

### A.2.3 Result 3

Suppose that a positive-definite matrix $A$ is partitioned as

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}.$$

Then

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}$$

where

$$\begin{aligned}
\mathbf{B}_{11} &= \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} \\
\mathbf{B}_{12} &= -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}^{-1} \\
\mathbf{B}_{21} &= -\mathbf{B}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} \\
\mathbf{B}_{22} &= \{\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\}^{-1}
\end{aligned} \tag{A.1}$$

(Graybill, 1983, page 184).

### A.2.4 Result 4

Let $\mathbf{R}$ represent an $n \times n$ matrix, $\mathbf{S}$ an $n \times m$ matrix, $\mathbf{T}$ an $m \times m$ matrix and $\mathbf{U}$ an $m \times n$ matrix. Suppose that $\mathbf{R}$ and $\mathbf{T}$ are nonsingular. Then

$$(\mathbf{R} + \mathbf{S}\,\mathbf{T}\,\mathbf{U})^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{S}\,\{\mathbf{T}^{-1} + \mathbf{U}\,\mathbf{R}^{-1}\mathbf{S}\}^{-1}\mathbf{U}\mathbf{R}^{-1}$$

(Harville, 1997, page 424).

## A.3 Matrix differentiation

### A.3.1 Differentiation of a matrix product

Let $\mathbf{F}$ and $\mathbf{G}$ represent $p \times q$ and $q \times r$ matrices of functions defined on a set $S$ of a vector $\mathbf{x} = (x_1, \ldots, x_m)'$ of $m$ variables. Then, at any interior point of $S$ at which $\mathbf{F}$ and $\mathbf{G}$ are continuously differentiable, $\mathbf{F}\,\mathbf{G}$ is continuously differentiable and

$$\frac{\partial \mathbf{F}\,\mathbf{G}}{\partial x_j} = \mathbf{F}\,\frac{\partial \mathbf{G}}{\partial x_j} + \frac{\partial \mathbf{F}}{\partial x_j}\mathbf{G}, \quad j = 1, \ldots, m$$

(Harville, 1997, page 297).

### A.3.2 Differentiation of a matrix with respect to its elements

Suppose that $\mathbf{X} = \{x_{st}\}_{s=1,t=1}^{m}{}^{m}$, is a symmetric matrix of dimensions $m \times m$, where $x_{st}$ is the element that is in the $i$th row and $j$th column of $\mathbf{X}$. Then

$$\frac{\partial \mathbf{x}_{st}}{\partial x_{ii}} = \begin{cases} 1, & \text{if } \; s = t = i, \\ 0, & \text{otherwise}, \qquad\qquad i = 1, 2, \ldots, m \end{cases}$$

and for $j < i$

$$\frac{\partial \mathbf{x}_{st}}{\partial x_{ii}} = \begin{cases} 1, & \text{if } \; s = i, \text{ and } t = j \text{ or } s = j \text{ and } t = i, \\ 0, & \text{otherwise.} \end{cases}$$

In matrix notation

$$\frac{\partial \mathbf{X}}{\partial x_{ii}} = \mathbf{u}_i \, \mathbf{u}_i'$$

and, for $j < i$ (or alternatively for $j > i$),

$$\frac{\partial \mathbf{X}}{\partial x_{ij}} = \mathbf{u}_i \, \mathbf{u}_j' + \mathbf{u}_j \, \mathbf{u}_i'$$

where $\mathbf{u}_i$ represent the $i$th column of an identity matrix (Harville, 1997, page 299).

### A.3.3 Differentiation of the trace of a matrix

Let $\mathbf{F}$ represent $p \times p$ matrix of functions defined on a set $S$ of a vector $\mathbf{x} = (x_1, \ldots, x_m)'$ of $m$ variables. Then, at any interior point of $S$ at which $\mathbf{F}$ is continuously differentiable, $tr(\mathbf{F})$ is continuously differentiable and

$$\frac{\partial tr(\mathbf{F})}{\partial x_j} = tr\left(\frac{\partial \mathbf{F}}{\partial x_j}\right)$$

(Harville, 1997, page 300).

## A.4 First-order partial derivatives of the determinant and the inverse of a matrix

### A.4.1 Determinants

**Result 1**

Suppose $\mathbf{A} = \{a_{ij}\}_{i=1,j=1}^{m \ \ m}$ is a square matrix of order $m$ having elements that are not functionally related. Then denoting the cofactor of $a_{ij}$ in $|\mathbf{A}|$ by $|\mathbf{A}_{ij}|$, the first-order partial derivatives of the determinant at $a_{ij}$ is

$$\frac{\partial|\mathbf{A}|}{\partial a_{ij}} = |\mathbf{A}_{ij}|, \quad i, j = 1, 2, \ldots, m.$$

When $\mathbf{A}$ is symmetric

$$\frac{\partial|\mathbf{A}|}{\partial a_{ij}} = \begin{cases} |\mathbf{A}_{ij}|, & \text{if } j = i, \\ 2\,|\mathbf{A}_{ij}|, & \text{if, } j < i \end{cases}$$

(Harville, 1997, page 304-306).

**Result 2**

Suppose the elements of the square matrix $\mathbf{A}$ are functions of the scalar $t$. Then

$$\frac{\partial \log|\mathbf{A}|}{\partial t} = tr\left(\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial t}\right)$$

(Harville, 1997, page 305).

### A.4.2 Inverses

Suppose $\mathbf{A}$ is a nonsingular matrix and its elements are functions of the scalar $t$. Then

$$\frac{\partial \mathbf{A}^{-1}}{\partial t} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial t}\mathbf{A}^{-1}$$

(Harville, 1997, page 307).

# Appendix B

# Proofs of inequalities in Theorem 6.4.2

## B.1 $A > 0$ and $B > 0$

Observe the following:

i. The maximum root for $f_1(k) = 3\,k^2 - 2\,k - 4$ is 1.53518. $f_1(k) > 0$ for $k > 1.53518$ and so for $k \geq 6$.

ii. $f_2(k) = 16 - 8\,k^2 + 3\,k^3$ has only one real root -1.17782. $f_2(k) > 0$ for $k > -1.17782$ and so for $k \geq 6$.

iii. $f_3(k) = 96 + 96\,k - 60\,k^2 - 28\,k^3 + 15\,k^4$ has only two real roots -1.58943 and -0.811993. $f_3(k) > 0$ for $k > -0.811993$ and so for $k \geq 6$.

iv. The maximum root for $f_4(k) = -8 - 4\,k + 3\,k^2$ is 2.4305. $f_4(k) > 0$ for $k > 2.4305$ and so for $k \geq 6$.

Furthermore, since $k \geq 6$, $k-2$, $k^2-2$ and $k^2-4$ are strictly greater than zero. Thus $A > 0$ and $B > 0$ for $k \geq 6$ and $\gamma \geq 0$.

## B.2 $C_1 > 0$, $C_0 > 0$ and $T > 0$

Observe the following:

i. There is only one real root 4.2958 for $f_1(k) = k^3 - 2k^2 - 8k - 8$. $f_1(k) > 0$ for $k > 4.2958$ and hence for $k \geq 6$.

ii. The maximum root for $f_2(k) = 3k^3 + k^2 - 20k - 16$ is 2.76936. $f_2(k) > 0$ for $k > 2.76936$ and hence for $k \geq 6$.

iii. The maximum root for $f_3(k) = 3k^2 - 4k - 8$ is 2.4305. $f_3(k) > 0$ for $k > 2.4305$ and hence for $k \geq 6$.

iv. The maximum root for $f_4(k) = 5k^2 - 4k - 16$ is 2.23303. $f_4(k) > 0$ for $k > 2.23303$ and hence for $k \geq 6$.

Thus from (i) to (iv) it follows that $C_0 > 0$ for $k \geq 6$ and $\gamma \geq 0$. Observe also that $C_1 > 0$ for $k \geq 6$ and $\gamma \geq 0$.

Consider the sign of $T$. Observe the following

i. $k^2 - 8 > 0$ for $k \geq 6$.

ii. The maximum root for $f_1(k) = -8 - 4k + 3k^2$ is 2.4305 and $f_1(k) > 0$ for $k > 2.4305$.

So

$$\sqrt{A} + (k^2 - 8) + 2(-8 - 4k + 3k^2)\gamma + 2(k-2)(2+3k)\gamma^2 > 0$$

for $k \geq 6$ and $\gamma \geq 0$. It has already been shown that $B - \sqrt{A} > 0$ in the weight $w_1$. Thus the sign of $T$ depends on the sign of

$$(4+k^2) + 2\,(2+k)\,\gamma - (k-2)\,(2+3\,k)\,\gamma^2 + \sqrt{A}.$$

Then consider the difference

$$A - \{(k-2)\,(2+3\,k)\,\gamma^2\}^2 = k^2\,(k^2-4) + 16 + 2\,(2+k)\,(16 - 8\,k^2 + 3\,k^3)\,\gamma$$

$$+\,(96 + 96\,k - 60\,k^2 - 28\,k^3 + 15\,k^4)\,\gamma^2 + 2\,(k-2)\,(2+3\,k)\,(3\,k^2 - 4\,k - 8)\,\gamma^3.$$

There are only two negative real roots, -1.58943 and -0.811993, for the function $f(k) = 96+96\,k-60\,k^2-28\,k^3+15\,k^4$ and as $f(k) > 0$ for $k > -0.811993$, $f(k) > 0$ for $k \geq 6$. It is also easy to see that other terms in $A - \{(k-2)\,(2+3\,k)\,\gamma^2\}^2$ are nonnegative for $k \geq 6$ and $\gamma > 0$. Therefore $A - \{(k-1)\,(3\,k+1)\,\gamma^2\}^2 > 0$. This implies that $\sqrt{A} - (k-1)\,(3\,k+1)\,\gamma^2 > 0$. Thus $T > 0$.

## B.3 Coefficients in $\phi((0,0),\tilde{\xi}_D^*)$

Consider

$$\frac{\partial\phi((0,0),\tilde{\xi}^*)}{\partial\gamma} = \frac{\sum_{i=0}^{12} f_i(k)\,\gamma^i + (\sum_{i=0}^{10} g_i(k)\,\gamma^i\,\sqrt{A})}{(2+k)^2\,\sqrt{A}\,T^2},$$

where

$$f_0(k) = -24\,(k-2)\,k^2\,(2+k)\,(-8192-6144\,k-384\,k^4-48\,k^5+8\,k^6+12\,k^7$$

$$-\,18\,k^8 + 12\,k^{10} + 6\,k^{11} + k^{12}),$$

$$f_1(k) = -24\,(k-2)\,k^2\,(-188416-348160\,k-159744\,k^2+10240\,k^3+4800\,k^4$$

$$- 7264\,k^5 - 512\,k^6 + 896\,k^7 + 60\,k^8 - 254\,k^9 + 210\,k^{10} + 323\,k^{11} + 128\,k^{12} + 18\,k^{13}),$$

$$f_2(k) = -6\,(k-2)\,k^2\,(2+k)\,(-1982464-3874816\,k-1347584\,k^2+718848\,k^3$$
$$+ 111488\,k^4 - 97088\,k^5 + 15584\,k^6 + 8480\,k^7 + 760\,k^8 - 2604\,k^9 + 2282\,k^{10}$$
$$+ 2338\,k^{11} + 561\,k^{12})$$

$$f_3(k) = -6\,(k-2)\,k^2\,(-12615680-38748160\,k-35561472\,k^2-3248128\,k^3$$
$$+ 8293120\,k^4 + 1387392\,k^5 - 796736\,k^6 + 26912\,k^7 + 85872\,k^8 - 6184\,k^9$$
$$-2732\,k^{10} + 18462\,k^{11} + 12714\,k^{12} + 2475\,k^{13})$$

$$f_4(k) = -6\,(k-2)\,k^2\,(-27033600-100024320\,k-117669888\,k^2-26075136\,k^3$$
$$+ 35136768\,k^4 + 14267904\,k^5 - 4068416\,k^6 - 1400992\,k^7 + 422384\,k^8 + 65728\,k^9$$
$$- 28172\,k^{10} + 22142\,k^{11} + 27072\,k^{12} + 6777\,k^{13})$$

$$f_5(k) = -12\,(k-2)\,k^2\,(-20545536-89210880\,k-129073152\,k^2-47185920\,k^3$$
$$+ 43526784\,k^4 + 30727872\,k^5 - 4696960\,k^6 - 5078464\,k^7 + 458776\,k^8 + 365852\,k^9$$
$$- 43972\,k^{10} - 6618\,k^{11} + 16731\,k^{12} + 5913\,k^{13}),$$

$$f_6(k) = -36\,(k-2)^2\,k^2\,(2+3\,k)\,(1892352+7569408\,k+10192896\,k^2+3956736\,k^3$$
$$- 1982624\,k^4 - 1424864\,k^5 + 181928\,k^6 + 180908\,k^7 - 5648\,k^8 - 4801\,k^9$$
$$+ 2382\,k^{10} + 756\,k^{11}),$$

$$f_7(k) = -12\,(k-2)^2\,k^2\,(2+3\,k)\,(4595712+21491712\,k+34578432\,k^2$$
$$+ 17215488\,k^3 - 8431392\,k^4 - 8993520\,k^5 + 305888\,k^6 + 1446584\,k^7 + 10766\,k^8$$
$$- 89211\,k^9 + 4491\,k^{10} + 2511\,k^{11}),$$

$$f_8(k) = -6\,(k-2)^2\,k^2\,(2+3\,k)^2\,(2703360+10475520\,k+11710464\,k^2-428544\,k^3$$

$$- 6788448\,k^4 - 1305984\,k^5 + 1466816\,k^6 + 215712\,k^7 - 139790\,k^8 - 3600\,k^9 + 3429\,k^{10}),$$

$$f_9(k) = -6\,(k-2)^3\,k^2\,(2+3\,k)^3\,(-281600-1013760\,k-999424\,k^2+75776\,k^3$$
$$+ 439968\,k^4 + 42160\,k^5 - 62700\,k^6 - 2184\,k^7 + 2331\,k^8),$$

$$f_{10}(k) = -6\,(k-2)^4\,k^2\,(2+3\,k)^4\,(19712+66176\,k+55296\,k^2-6400\,k^3-15980\,k^4$$
$$- 124\,k^5 + 969\,k^6),$$

$$f_{11}(k) = -48\,(k-2)^5\,k^2\,(2+3\,k)^5\,(-104-328\,k-214\,k^2+28\,k^3+27\,k^4),$$

$$f_{12}(k) = -24\,(k-2)^6\,k^2\,(2+k)\,(2+3\,k)^6\,(2+5\,k),$$

$$g_0(k) = -24\,(k-2)\,k^2\,(2+k)\,(4+k^2)\,(-512-384\,k+64\,k^2+48\,k^3-36\,k^4-12\,k^5$$
$$+ 10\,k^6 + 6\,k^7 + k^8),$$

$$g_1(k) = -24\,(k-2)\,k^2\,(-38912-72704\,k-38656\,k^2-7040\,k^3-2720\,k^4-816\,k^5$$
$$+ 400\,k^6 + 424\,k^7 + 422\,k^8 + 301\,k^9 + 106\,k^{10} + 15\,k^{11}),$$

$$g_2(k) = -18\,(k-2)\,k^2\,(2+k)\,(-110592-221184\,k-93696\,k^2+14080\,k^3-960\,k^4$$
$$+ 288\,k^5 + 2016\,k^6 + 912\,k^7 + 816\,k^8 + 502\,k^9 + 123\,k^{10}),$$

$$g_3(k) = -24\,(k-2)\,k^2\,(-417792-1314816\,k-1289472\,k^2-282496\,k^3+155648\,k^4$$
$$+ 48192\,k^5 + 6592\,k^6 + 9216\,k^7 + 4820\,k^8 + 3014\,k^9 + 1476\,k^{10} + 297\,k^{11}),$$

$$g_4(k) = -12\,(k-2)\,k^2\,(-1376256-5246976\,k-6526464\,k^2-2089472\,k^3+1268736\,k^4$$
$$+ 706048\,k^5 - 18272\,k^6 - 3872\,k^7 + 16072\,k^8 + 6556\,k^9 + 3810\,k^{10} + 1053\,k^{11}),$$

$$g_5(k) = -216\,(k-2)^2\,k^2\,(2+3\,k)\,(21504+75264\,k+88256\,k^2+33824\,k^3-3944\,k^4$$
$$- 2948\,k^5 + 670\,k^6 + 267\,k^7 + 50\,k^8 + 18\,k^9),$$

$$g_6(k) = 12\,(k-2)^2\,k^2\,(2+3\,k)\,(-301056-1268736\,k-1809024\,k^2-813568\,k^3$$

$$+\,237360\,k^4+218688\,k^5-5872\,k^6-14176\,k^7+411\,k^8),$$

$$g_7(k)=24\,(k-2)^2\,k^2\,(2+3\,k)^2\,(-39936-138240\,k-130272\,k^2+11168\,k^3+52864\,k^4$$
$$+\,4504\,k^5-6070\,k^6-84\,k^7+135\,k^8),$$

$$g_8(k)=18\,(k-2)^3\,k^2\,(2+3\,k)^3\,(4608+14976\,k+12192\,k^2-1280\,k^3-3180\,k^4$$
$$-\,28\,k^5+147\,k^6),$$

$$g_9(k)=24\,(k-2)^4\,k^2\,(2+3\,k)^4\,(-176-544\,k-352\,k^2+40\,k^3+39\,k^4)$$

and

$$g_{10}(k)=24\,(k-2)^5\,k^2\,(2+k)\,(2+3\,k)^5\,(2+5\,k).$$

By using the approach of the previous sections it can be shown that $f_i(k)<0$ where $i=0,1,\ldots,12$, $g_i(k)<0$ where $i=0,1,\ldots,5$ and $g_i(k)>0$ where $i=6,\ldots,10$ for $k\geq 6$.

Similarly, it can be shown that the following functions are negative for $k\geq 6$:

$$h_{12}(k)=-144\,(k-2)^4\,k^4\,(2+3\,k)^2\,(307788093849600+2456089798901776\,k$$
$$+\,8200370069176632\,k^2+14453234403901444\,k^3+13342262042296632\,k^4$$
$$+\,3914893193379844\,k^5-3807447419453444\,k^6-3494559549358088\,k^7$$
$$-\,1257517924659200\,k^8+86621651507200\,k^9+21767178315264\,k^{10}$$
$$-\,10086985614592\,k^{11}-3337782319040\,k^{12}+1007942676224\,k^{13}+274387115728\,k^{14}$$
$$-110647624912\,k^{15}-20183334888\,k^{16}+10327521328\,k^{17}+2231622317\,k^{18}$$
$$-\,271052988\,k^{19}-14435613\,k^{20}+32414256\,k^{21}+5143824\,k^{22})$$

$$h_{13}(k)=-288\,(k-2)^4\,k^4\,(2+3\,k)^2\,(22420333264896+193413953617920\,k$$

$$+\,700709137809408\,k^2+1347122884509696\,k^3+1369489873305600\,k^4$$

$$+\,462946005811200\,k^5-450152313421824\,k^6-492374479355904\,k^7$$

$$-\,56664888763392\,k^8+126373527188992\,k^9+45128900338432\,k^{10}-15205419092992\,k^{11}$$

$$-\,8199717869824\,k^{12}+1313693555712\,k^{13}+775846082496\,k^{14}-150329717232\,k^{15}$$

$$-\,57682822924\,k^{16}+14758140394\,k^{17}+4575399873\,k^{18}-663313905\,k^{19}-195442794\,k^{20}$$

$$+\,27796284\,k^{21}+5694948\,k^{22})$$

$$h_{14}(k)=-144\,(k-2)^4\,k^4\,(2+3\,k)^2\,(36017125982208+334339044802560\,k$$

$$+\,1307476846706688\,k^2+2725413849661440\,k^3+3029229493813248\,k^4$$

$$+\,1166364714074112\,k^5-1116461748977664\,k^6-1419607298801664\,k^7$$

$$-\,250084478407680\,k^8+390586007232512\,k^9+180528836921088\,k^{10}$$

$$-\,46569507290112\,k^{11}-37010885742976\,k^{12}+3233229520896\,k^{13}+4025455584096\,k^{14}$$

$$-\,374885913792\,k^{15}-312635538212\,k^{16}+47670743344\,k^{17}+23528240307\,k^{18}$$

$$-\,2707095384\,k^{19}-1314069291\,k^{20}+81439749\,k^{21}+28979937\,k^{22})$$

$$h_{15}(k)=-144\,(k-2)^4\,k^4\,(2+3\,k)^3\,(9223777812480+78092068454400\,k+269874200838144\,k^2$$

$$+\,467039505350656\,k^3+354682751549440\,k^4-72151828463616\,k^5-318645294268416\,k^6$$

$$-\,145374599331840\,k^7+76865183357952\,k^8+73688998502912\,k^9-4199231840768\,k^{10}$$

$$-\,16013303996672\,k^{11}-733523581696\,k^{12}+2019651579520\,k^{13}+62755685984\,k^{14}$$

$$-\,177672984848\,k^{15}+5303816964\,k^{16}+13548002934\,k^{17}-500664564\,k^{18}-850175298\,k^{19}$$

$$+\,16374879\,k^{20}+22086513\,k^{21})$$

$$h_{16}(k)=-144\,(k-2)^4\,k^4\,(2+3\,k)^4\,(1577343516672+12128492716032\,k+36377594953728\,k^2$$

$$+49076422836224\,k^3+15018381017088\,k^4-35184368975872\,k^5-34254512701440\,k^6$$

$$+3864974000128\,k^7+16180810702080\,k^8+2406249280000\,k^9-3846440590336\,k^{10}$$

$$-817547652352\,k^{11}+566761785952\,k^{12}+96403249536\,k^{13}-58417332944\,k^{14}$$

$$-4777094768\,k^{15}+4775634005\,k^{16}+216725118\,k^{17}-336001797\,k^{18}-8597772\,k^{19}$$

$$+11332305\,k^{20})$$

$$h_{17}(k)=-288\,(k-2)^5\,k^4\,(2+3\,k)^5\,(-45405437952-337612111872\,k-964072833024\,k^2$$

$$-1207448109056\,k^3-287027388416\,k^4+823845601280\,k^5+665022375936\,k^6$$

$$-123115071488\,k^7-271595445376\,k^8-16679188096\,k^9+52317308000\,k^{10}$$

$$+5067583936\,k^{11}-5868978600\,k^{12}-235254184\,k^{13}+487340482\,k^{14}-4533552\,k^{15}$$

$$-39885867\,k^{16}-56052\,k^{17}+1842183\,k^{18})$$

$$h_{18}(k)=-144\,(k-2)^6\,k^4\,(2+3\,k)^6\,(3148873728+22691708928\,k+61430562816\,k^2$$

$$+70476464128\,k^3+11299758080\,k^4-45161603072\,k^5-28927748096\,k^6+7511601152\,k^7$$

$$+9659469376\,k^8-153115392\,k^9-1339548448\,k^{10}+13639392\,k^{11}+110617716\,k^{12}$$

$$-9540040\,k^{13}-10814760\,k^{14}+540252\,k^{15}+738963\,k^{16})$$

$$h_{19}(k)=-5184\,(k-2)^7\,k^4\,(2+3\,k)^7\,(-1376256-9633792\,k-24571904\,k^2-25374720\,k^3$$

$$-1948160\,k^4+14820352\,k^5+7162368\,k^6-2278912\,k^7-1771328\,k^8+114848\,k^9$$

$$+134560\,k^{10}-32552\,k^{11}-18346\,k^{12}+3540\,k^{13}+2403\,k^{14})$$

and

$$h_{20}(k)=-139968\,(k-2)^8\,k^{14}\,(2+k)\,(2+3\,k)^8\,(2+5\,k).$$

## B.4 Coefficients in $E_0$, $F_0$ and $F_1$

Observe the following:

i. The maximum root of $f_1(k) = 72 - 20\,k - 66\,k^2 + 13\,k^3$ is 5.16723. $f_1(k) > 0$ for $k \geq 5.16723$ and hence for $k \geq 6$.

ii. $f_2(k) = 144 + 64\,k - 248\,k^2 - 128\,k^3 + 45\,k^4$ has only two positive roots 0.7778 and 4.0662. $f_2(k) > 0$ for $k > 4.0662$ and hence for $k \geq 6$.

iii. The maximum root of $f_3(k) = -12 + 4\,k + 17\,k^2$ is 0.73072. $f_3(k) > 0$ for $k \geq 0.73072$ and hence for $k \geq 6$.

Observe that

i. There is only one real root -3.03461 for $f_1(k) = 80 + 72\,k - 20\,k^2 - 24\,k^3 + 5\,k^4 + 3\,k^5$ and $f_1(k) > 0$ for $k > -3.03461$.

ii. There are only three real roots -2.57762, -1.2713 and -1.022423 for $f_2(k) = 160 + 208\,k - 32\,k^2 - 88\,k^3 + 2\,k^4 + 9\,k^5$ and $f_2(k) > 0$ for $k > -1.022423$.

iii. The maximum root for $f_3(k) = 160 + 192\,k - 112\,k^2 - 94\,k^3 + 33\,k^4$ is 3.19156 and $f_3(k) > 0$ for $k > 3.19156$.

iv. The maximum root for $f_4(k) = 20 + 12\,k - 17\,k^2 - 5\,k^3 + 3\,k^4$ is 2.8733 and $f_4(k) > 0$ for $k > 2.8733$.

v. There is only one real root 4.10118 for $f_5(k) = -24 - 20\,k - 6\,k^2 + 3\,k^3$ and $f_5(k) > 0$ for $k > 4.10118$.

vi. There is only one real root 1.52022 for $f_6(k) = -24 - 28\,k + 6\,k^2 + 15\,k^3$ and $f_6(k) > 0$ for $k > 1.52022$.

## B.5 Coefficients in $Z$

Observe the following:

i. There is only one real root -3.29339 for

$$f_1(k) = 9216 + 8192\,k - 5376\,k^2 - 5568\,k^3 + 1920\,k^4 + 1776\,k^5 - 336\,k^6 - 292\,k^7$$
$$+36\,k^8 + 21\,k^9.$$

Thus $f_1(k) > 0$ for $k > -3.29339$ and hence for $k \geq 6$.

ii. There is no real root for

$$f_2(k) = 178176 + 319488\,k + 41216\,k^2 - 199680\,k^3 - 62208\,k^4 + 62976\,k^5$$
$$+23328\,k^6 - 10368\,k^7 - 3832\,k^8 + 768\,k^9 + 273\,k^{10}.$$

The positive terms of $f_2(k)$ dominate the function and therefore $f_2(k) > 0$ for $k \geq 6$.

iii. The maximum root for

$$f_3(k) = 202752 + 445440\,k + 144128\,k^2 - 278016\,k^3 - 155648\,k^4 + 79104\,k^5$$
$$+51296\,k^6 - 12192\,k^7 - 7512\,k^8 + 804\,k^9 + 423\,k^{10}$$

is -1.03145. $f_3(k) > 0$ for $k > -1.03145$ and hence for $k \geq 6$.

iv. The maximum root for

$$f_4(k) = (672 + 880\,k - 128\,k^2 - 352\,k^3 + 6\,k^4 + 33\,k^5)\,(608 + 784\,k - 128\,k^2 \\ -352\,k^3 + 10\,k^4 + 39\,k^5)$$

is -0.95881. $f_4(k) > 0$ for $k > -0.95881$ and hence for $k \geq 6$.

Furthermore, the constant term $f_0(k) = 9\,(k-2)\,k^2\,(2+k)^6\,[k^2\,(k^2-4)+16] > 0$ for $k \geq 6$.

## B.6 $T_0 > 0$ and $T_1 > 0$

Observe the following:

i. The largest positive root for the function $f_0(k) = -3\,(k-4)\,(k-1)\,k\,(2+k)^2\,(4+k^2)$ is 4. $f_0(k) < 0$ for $k > 4$.

ii. The largest positive root for the function $f_1(k) = -(2+k)\,(256 + 448\,k - 264\,k^2 - 484\,k^3 + 90\,k^4 - 65\,k^5 + 25\,k^6)$ is 3.35845. $f_1(k) < 0$ for $k > 3.35845$.

iii. The largest positive root for the function $f_2(k) = -8\,(256 + 512\,k - 20\,k^2 - 572\,k^3 - 221\,k^4 + 56\,k^5 + 10\,k^6 + 6\,k^7)$ is 2.91027 and $f_2(k) < 0$ for $k > 2.91027$.

iv. The largest positive root for $f_3(k) = 8\,(-384 - 912\,k - 84\,k^2 + 1036\,k^3 + 507\,k^4 - 201\,k^5 - 70\,k^6 + 9\,k^7)$ is 9.38154 and $f_3(k) > 0$ for $k > 9.38154$.

v. The largest positive root for $f_4(k) = (2+3\,k)\,(-1024 - 1440\,k + 1288\,k^2 + 1684\,k^3 - 414\,k^4 - 415\,k^5 + 105\,k^6)$ is 3.68476 and $f_4(k) > 0$ for $k > 3.68476$.

vi. The largest positive root for $f_5(k) = f_5(k) = (k-2)\,(2+3\,k)^2\,(64+64\,k-90\,k^2-49\,k^3+29\,k^4)$ is 2.43899 and $f_5(k) > 0$ for $k > 2.43899$.

vii. The largest positive root for $g_0(k) = -3\,(k-4)\,(k-1)\,k\,(2+k)^2$ is 4 and $g_0(k) < 0$ for $k > 4$.

viii. The largest positive root for $g_1(k) = -(2+k)\,(64+64\,k-46\,k^2-77\,k^3+25\,k^4)$ is 3.33245 and $g_1(k) < 0$ for $k > 3.33245$.

ix. The largest positive root for $g_2(k) = (-256-464\,k+84\,k^2+412\,k^3+89\,k^4-81\,k^5)$ is 2.75345 and $g_2(k) < 0$ for $k > 2.75345$, and

x. The largest positive root for $g_3(k) = -(2+3\,k)\,(64+64\,k-90\,k^2-49\,k^3+29\,k^4)$ is 2.43899 and $g_3(k) < 0$ for $k > 2.43899$.

Consider now

$$\left\{\sum_{i=3}^{5} f_i(k)\,\gamma^i\right\} - A\left\{\sum_{i=0}^{3} g_i(k)\,\gamma^i\right\} = \sum_{i=0}^{9} h_i(k)\,\gamma^i$$

where

$$h_0(k) = -9\,(k-4)^2\,(k-1)^2\,k^2\,(2+k)^4\,(16-4\,k^2+k^4),$$

$$h_1(k) = -12\,(k-4)\,(k-1)\,k\,(2+k)^3\,(512+896\,k-592\,k^2-1224\,k^3+420\,k^4+336\,k^5$$
$$-115\,k^6-55\,k^7+17\,k^8),$$

$$h_2(k) = -2\,(2+k)^2\,(32768+212992\,k+106496\,k^2-501760\,k^3-337120\,k^4+481888\,k^5$$
$$+\;254296\,k^6-252632\,k^7-47872\,k^8+61694\,k^9+1001\,k^{10}-6974\,k^{11}+1073\,k^{12}),$$

$$h_3(k) = -2\,(2+k)\,(524288+2260992\,k+1732608\,k^2-4052992\,k^3-5356032\,k^4$$

$$+\ 2341248\,k^5 + 4742656\,k^6 - 653424\,k^7 - 1900272\,k^8 + 235208\,k^9 + 383892\,k^{10}$$

$$- 58917\,k^{11} - 32740\,k^{12} + 6825\,k^{13}),$$

$$h_4(k) = -7340032-32112640\,k-35217408\,k^2+36761600\,k^3+90827264\,k^4$$

$$+11232768\,k^5 - 72194432\,k^6 - 28289600\,k^7 + 27375888\,k^8 + 12411200\,k^9$$

$$- 6416368\,k^{10} - 2402508\,k^{11} + 893963\,k^{12} + 180986\,k^{13} - 57789\,k^{14},$$

$$h_5(k) = -4\,(3670016+16515072\,k+19337216\,k^2-16954368\,k^3-48481536\,k^4$$

$$- 10395264\,k^5 + 38174976\,k^6 + 19195920\,k^7 - 14373936\,k^8 - 8790608\,k^9 + 3431592\,k^{10}$$

$$+ 1820941\,k^{11} - 541590\,k^{12} - 148389\,k^{13} + 42372\,k^{14}),$$

$$h_6(k) = -8912896+-44171264\,k-60260352\,k^2+36765696\,k^3+140711424\,k^4$$

$$+\ 32030208\,k^5 - 133160960\,k^6 - 64826688\,k^7 + 72190608\,k^8 + 44915264\,k^9$$

$$-\ 20422840\,k^{10} - 12635964\,k^{11} + 3620059\,k^{12} + 1274934\,k^{13} - 342909\,k^{14},$$

$$h_7(k) = -2\,(2+3\,k)\,(524288+2097152\,k+1449984\,k^2-3831808\,k^3-3846912\,k^4$$

$$+\ 5796352\,k^5 + 6560000\,k^6 - 4791312\,k^7 - 6138368\,k^8 + 1558080\,k^9 + 2269276\,k^{10}$$

$$- 395195\,k^{11} - 318408\,k^{12} + 74835\,k^{13}),$$

$$h_8(k) = -9\,k^3\,(2+3\,k)^2\,(4096+53760\,k+98304\,k^2-40624\,k^3-160704\,k^4-6504\,k^5$$

$$+ 86408\,k^6 - 1335\,k^7 - 18990\,k^8 + 4057\,k^9),$$

and

$$h_9(k) = -18\,(k-2)\,k^4\,(2+3\,k)^3\,(7\,k^2-11\,k-14)\,(64+64\,k-90\,k^2-49\,k^3+29\,k^4).$$

It is easy to show that $h_i(k) < 0$ for all $k \geq 6$. Therefore it follows from the above discussions that $T_0 < 0$ for all $k \geq 6$ and $\gamma \geq 0$ .

Consider now the coefficients in $T_1$. Let

$$f_0(k) = -3\,(k-4)\,(k-1)\,k\,(2+k)^2,$$

$$f_1(k) = -(2+k)\,(64+64\,k-46\,k^2-77\,k^3+25\,k^4),$$

$$f_2(k) = -256-464\,k+84\,k^2+412\,k^3+89\,k^4-81\,k^5,$$

and

$$f_3(k) = -(2+3\,k)\,(64+64\,k-90\,k^2-49\,k^3+29\,k^4).$$

It is clear that the functions $f_0(k)$, $f_1(k)$ and $f_3(k)$ are less than zero for $k \geq 6$. Further the largest positive root for $f_2(k)$ is 2.75345 and $f_2(k) < 0$ for $k > 2.75345$. Thus it follows that $T_1 < 0$ for all $k \geq 6$ and $\gamma \geq 0$. Overall therefore $T_1\sqrt{A} + T_0 < 0$ for all $k \geq 6$ and $\gamma \geq 0$.

## B.7 $U_1\sqrt{A} + U_0 < 0$

Let

$$f_0(k) = -9\,(2+k)^2\,(4+k)\,(4+k^2),$$

$$f_1(k) = -(2+k)\,(1280+1784\,k+692\,k^2+358\,k^3+85\,k^4+8\,k^5+k^6),$$

$$f_2(k) = -(4480+11168\,k+8864\,k^2+3464\,k^3+1296\,k^4+348\,k^5+50\,k^6+3\,k^7),$$

$$f_3(k) = -3840-11616\,k-11040\,k^2-3992\,k^3-1272\,k^4-452\,k^5-38\,k^6+9\,k^7,$$

$$f_4(k) = (2+3\,k)\,(-800-1560\,k-660\,k^2-98\,k^3-97\,k^4+32\,k^5+15\,k^6),$$

$$f_5(k) = (k-2)\,(2+3\,k)^2\,(32+38\,k+17\,k^2+22\,k^3+5\,k^4),$$

$$g_0(k) = -9\,(2+k)^2\,(4+k),$$

$$g_1(k) = -(2+k)\,(176+266\,k+69\,k^2+8\,k^3+k^4),$$

$$g_2(k) = -272-700\,k-584\,k^2-195\,k^3-58\,k^4-9\,k^5$$

and

$$g_3(k) = -\,(2+3\,k)\,(32+38\,k+17\,k^2+22\,k^3+5\,k^4).$$

It is easy to see that $f_i(k) < 0$, $i = 0, 1, 2$, $f_5(k) > 0$ and $g_i(k) < 0$, $i = 0, 1, 2, 3$ for all $k \geq 6$. The largest positive root for $f_3(k)$ is 10.6534. Also $f_3(k) < 0$ for $k < 10.6534$ and $f_3(k) > 0$ for $k > 10.6534$. Further the largest positive root for $f_4(k)$ is 3.16283 and $f_4(k) < 0$ for all $k \geq 6$. Consider now

$$\left\{\sum_{i=3}^{5} f_i(k)\,\gamma^i\right\} - A\left\{\sum_{i=0}^{3} g_i(k)\,\gamma^i\right\} = \sum_{i=0}^{9} h_i(k)\,\gamma^i$$

where

$$h_0(k) = -81\,(2+k)^4\,(4+k)^2\,(16-4\,k^2+k^4),$$

$$h_1(k) = -18\,(2+k)^3\,(4+k)\,(5120+7136\,k+400\,k^2-1800\,k^3-120\,k^4+378\,k^5 + 92\,k^6 + 8\,k^7 + k^8),$$

$$h_2(k) = -(2+k)^2\,(2928640+8105984\,k+6732864\,k^2+575808\,k^3-1410192\,k^4 - 269840\,k^5 + 253040\,k^6 + 127196\,k^7 + 28746\,k^8 + 4704\,k^9 + 468\,k^{10} + 16\,k^{11} + k^{12}),$$

$$h_3(k) = -2\,(2+k)\,(6840320+26462208\,k+36616192\,k^2+19008640\,k^3-1369600\,k^4 - 4371440\,k^5 - 579248\,k^6 + 669768\,k^7 + 367396\,k^8 + 110179\,k^9 + 22526\,k^{10} + 2650\,k^{11} + 182\,k^{12} + 12\,k^{13}),$$

$$h_4(k) = -41574400-205434880\,k-394657280\,k^2-355039744\,k^3-120313728\,k^4$$
$$+31671744\,k^5+34897424\,k^6+5140768\,k^7-4017056\,k^8-2879932\,k^9$$
$$-1083669\,k^{10}-250944\,k^{11}-35214\,k^{12}-3472\,k^{13}-234\,k^{14},$$

$$h_5(k) = -4\,(10723328+58770432\,k+127155456\,k^2+133227392\,k^3+60708096\,k^4$$
$$-3150448\,k^5-14560656\,k^6-5449376\,k^7-167544\,k^8+875141\,k^9+549986\,k^{10}$$
$$+166540\,k^{11}+30096\,k^{12}+3972\,k^{13}+306\,k^{14}),$$

$$h_6(k) = -15646720-92512256\,k-214548992\,k^2-228981248\,k^3-68828160\,k^4$$
$$+88631488\,k^5+106712336\,k^6+55273440\,k^7+16779096\,k^8+595348\,k^9$$
$$-2289405\,k^{10}-1047040\,k^{11}-258836\,k^{12}-44484\,k^{13}-3762\,k^{14},$$

$$h_7(k) = -2\,(2+3\,k)\,(573440+2683904\,k+4333312\,k^2+849920\,k^3-5987072\,k^4$$
$$-8356880\,k^5-5655904\,k^6-2561712\,k^7-551988\,k^8+199137\,k^9+159696\,k^{10}$$
$$+55174\,k^{11}+12054\,k^{12}+1125\,k^{13}),$$

$$h_8(k) = -9\,k^2\,(2+3\,k)^2\,(-6144-62464\,k-127152\,k^2-104160\,k^3-56280\,k^4$$
$$-22440\,k^5+3313\,k^6+5488\,k^7+2764\,k^8+768\,k^9+79\,k^{10})$$

and

$$h_9(k) = -18\,(k-2)\,k^3\,(2+3\,k)^3\,(18+11\,k+6\,k^2+k^3)\,(32+38\,k+17\,k^2+22\,k^3+5\,k^4).$$

Observe that $h_i(k)<0$, $i=0,1,2,4,9$ for all $k\geq 6$. Observe also the following

i. The largest root for $h_3(k)$ is -2 and $h_3(k)<0$ for $k>-2$.

ii. The largest root for $h_5(k)$ is -0.7293 and $h_5(k)<0$ for $k>-0.7293$.

iii. The largest root for $h_7(k)$ is 2.54034 and $h_7(k)<0$ for $k>2.54034$.

iv. The largest root for $h_8(k)$ is 2.19337 and $h_8(k)<0$ for $k>-0.7293$.

Thus overall $U_1\sqrt{A}+U_0<0$ for all $k\geq 6$ and $\gamma\geq 0$.