Minimax Designs
for Comparing Treatment Means for Field Experiments

by

Beiyan Ou
B.Sc., University of Victoria, 2003

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

Minimax Designs
for Comparing Treatment Means for Field Experiments

by

Beiyan Ou
B.Sc., University of Victoria, 2003

Supervisory Committee

Supervisor

Dr. J. Zhou, (Department of Mathematics and Statistics)

Departmental Member

Dr. W. J. Reed, (Department of Mathematics and Statistics)

Departmental Member

Dr. M. Lesperance, (Department of Mathematics and Statistics)

Outside Member

Dr. M. Wilmut, (School of Earth and Ocean Sciences)

Supervisory Committee

Supervisor

Dr. J. Zhou, (Department of Mathematics and Statistics)

Departmental Member

Dr. W. J. Reed, (Department of Mathematics and Statistics)

Departmental Member

Dr. M. Lesperance, (Department of Mathematics and Statistics)

Outside Member

Dr. M. Wilmut, (School of Earth and Ocean Sciences)

ABSTRACT

This thesis studies the linear model, estimators of the treatment means, and optimality criteria for designs and analysis of spatially arranged experiments. Four types of commonly used spatial correlation structures are discussed, and a neighbourhood of covariance matrices is investigated. Various properties about the neighbourhood are explored. When the covariance matrix of the error process is unknown, but belongs to a neighbourhood of a covariance matrix, a modified generalized least squares estimator (GLSE) is proposed. This estimator seems more efficient than the ordinary least squares estimator in many practical applications. We also propose a criterion to find minimax designs that are efficient for a neighbourhood of correlations. When the number of plots is small, minimax designs can be computed exactly. When the number of plots is large, a simulated annealing algorithm is applied to find minimax or near minimax designs. Minimax designs for the least squares and generalized least squares estimators are compared in details. In general, we recommend using GLSE and the minimax design based on GLSE.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to acknowledge the contributions of some special persons, without their support, I would not have been successful in this endeavor.

My sincerest gratitude goes to my supervisor, Professor Julie Zhou of the Department of Mathematics and Statistics of the University of Victoria. Her broad knowledge, understanding, encouraging and personal guidance have provided a good basis for the present thesis.

My deepest appreciation goes to my parents; thank you mom and dad for bravely sending me out of China to Canada to see the other side of the world. Without that decision, I would not have the opportunity to pursue this master degree and to see the far-and-wide part of this land that God has kept so glorious-and-free.

I would like to gratefully acknowledge the faculty members and staff in the Department of Mathematics and Statistics who helped me over the years, and the financial support of the University of Victoria as well.

I would like to extend my warmest gratitude to my landlords, Richard and Ellen Hung, who have treated me as part of their family. Working full-time and preparing for my thesis were only possible with their care and hospitality.

It would be a long list to mention all my friends who have enriched my life in Victoria, but I am thankful for all of them nonetheless.

Last, but not least, I want to thank God for granting me the wisdom I need. With grateful heart, I hereby ascribe to Him all the glory and honor.

Beiyan Ou

September 2006

Chapter 1

Introduction

Observations from a field experiment containing several plots are usually correlated, since it is difficult to have uniform conditions for all the experiment plots. Observations from neighbouring plots tend to be positively correlated. If spatial correlation is taken into account at the design stage, the precision of the estimation procedure can be increased.

This chapter provides an introduction to spatial correlation in a field experiment. In Section 1.1, we show an example of the spatial dependence in an agricultural field experiment. In Section 1.2, we discuss linear models used for designs and analysis of spatially arranged experiments. In Section 1.3, we discuss two estimators of the treatment effects. In Section 1.4, we explain optimality criteria to construct optimal designs. We conclude this chapter with an overview of this thesis.

1.1 Field Experiments

The statistical theory of experimental design is concerned with obtaining as much information as possible from the resources available. In some applications, of which agriculture is the most obvious, the experimental material consists of sets of spatial material. Other spatial applications include forestry experiments, sheet metal production, and paper-making.

The agricultural experiment design has been the most explicitly developed part of the modern development of statistical experimental design since the 1920s. In agriculture, the objective of experiments usually is to test the relative productivity under different treatments, such as different varieties of crops, different fertilizers, and different soil fertilities, etc, and to determine if any one of the varieties yields more than others.

The fact that neighbouring plots tend to be more alike than those further apart, in other words that the observations are spatially dependent, was well known in the early days of agricultural experimentation. A simple way of partially overcoming this was to ensure that neighbouring plots should contain different treatments, and that plots having the same treatment should be well separated. This goes back at least to J. F. W. Johnson in 1849 (Cochran, 1960), who suggested that repetitions of a treatment should be a Knight's move apart, an idea that was popularized in the 1920s. The Knight's move square is a special Latin square such that all plots with the same

treatment can be traversed by a series of Knight's moves as on a chessboard. Figure 1.1 shows an example of Knight's move Latin squares from Martin (1996). Numbers 1, 2, ... , and 5 denote 5 treatments.

1	2	3	4	5
4	5	1	2	3
2	3	4	5	1
5	1	2	3	4
3	4	5	1	2

Figure 1.1: An example of Knight's move Latin squares from Martin (1996).

These layouts are systematic with each subsequent row being formed by cyclically shifting the treatments in the previous row by two plots. Note that *systematic* is used in general to mean that the design was selected purposely and was not the outcome of a valid randomization scheme. These systematic designs were very popular, however, there was no statistically correct way to compare treatments. Treatment effects were estimated by means, and variation was estimated using the (corrected) total sum of squares (Mead, 1988).

Fisher (1966) gave a detailed example of an agricultural experiment. In his example, the experimental area is divided into eight blocks, and each block is subdivided into five plots of equal dimensions, making forty plots in total (Figure 1.2). The

1	4	3	2	5
2	1	5	4	3
5	1	4	2	3
1	3	2	5	4
3	4	1	2	5
5	2	3	1	4
3	5	2	4	1
4	3	5	1	2

Figure 1.2: The 8x5 plots for Fisher's example

numbers in Figure 1.2 represent the different crop varieties. In this example there are five different crop varieties. Aside from the differences in crop variety used, it is assumed that uniform agricultural treatment is applied to the whole area. In each block, each of the five varieties under test is each assigned to one of the five plots, and this assignment is made at random. Figure 1.2 shows an example of the random assignments. His goal was to compare the yields of five crop varieties and evaluate any differences with a certain degree of precision.

However, sometimes the assumption of having uniform agricultural conditions across the field is not applicable. Fisher established the principles of randomization,

blocking, and replication to reduce error from unknown factors. By randomization we mean that both the allocation of the experimental material and the order in which the individual runs or trials of the experiment are to be performed are randomly determined. Some statistical methods require that the observations be independently distributed random variables. Randomization usually makes this assumption valid. Blocking helps provide more homogeneous experiment units to compare the treatment effect. By replication we mean a repetition of the basic experiment. Replication permits the experimenter to obtain a more precise estimate of the effect.

Yet, none of these methods helps remove the spatial dependence in this kind of agricultural field experiment. Randomization is not sufficient to neutralize the spatial correlation at spatial scales larger or smaller than the plot dimensions (Cressie, 1993). If the spatial correlation is taken into account at the design stage, the precision of the procedure may be increased since the structure is more suitably modeled. The design of comparative experiments is considered for the case that the units (or subsets of units) are spatially arranged, and the spatial or directional layout is itself important. Research in this area has greatly expanded recently, and is having an increasing impact on practical experimentation.

The layout of the experimental units considered in this thesis will be in two spatial dimensions. The spatial locations of the units and the spatial relationships between different units are important (Martin, 1996). Only rectangular units arranged in a

non-overlapping rectangular array are considered here.

1.2 Linear Model

In this thesis, there is only one response variable of interest. The following notation and terminology will be used. There will be mn plots and t treatments. Each treatment is applied to r plots, so that $tr = mn$. The layout can be considered as m rows each containing n plots. The model will consider only the effects of the treatments. The plots are equally spaced, so that the error correlation between any two plots depends on only the number of plots between them horizontally and vertically.

The vector model for fixed effects is usually in the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \tag{1.1}$$

where $\mathbf{y} = (y_{1,1}, \ldots, y_{1,n}, \ldots, y_{m,1}, \ldots, y_{m,n})'$ is the vector of the observed data (see Figure 1.3), and $\mathbf{X}$ is the Nxt design matrix with N = mn, $\boldsymbol{\theta}$ is the tx1 vector of treatment means, and $\boldsymbol{\epsilon} = (\epsilon_{1,1}, \ldots, \epsilon_{1,n}, \ldots, \epsilon_{m,1}, \ldots, \epsilon_{m,n})'$ is the vector of unknown errors. Entries in $\mathbf{X}$ are either 0 or 1. Each row of $\mathbf{X}$ consists of one 1 entry, and the remaining entries are 0's. The column number of the entry 1 corresponds to the treatment number.

The errors $\boldsymbol{\epsilon}$ in model (1.1) are usually assumed to be correlated, i.e. cov($\boldsymbol{\epsilon}$) =

$y_{1,1}$	$y_{1,2}$	$\cdots$	$y_{1,(n-1)}$	$y_{1,n}$
$y_{2,1}$	$y_{2,2}$	$\cdots$	$y_{2,(n-1)}$	$y_{2,n}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$y_{(m-1),1}$	$y_{(m-1),2}$	$\cdots$	$y_{(m-1),(n-1)}$	$y_{(m-1),n}$
$y_{m,1}$	$y_{m,2}$	$\cdots$	$y_{m,(n-1)}$	$y_{m,n}$

Figure 1.3: y_{ij} denotes the measured response on ith row and jth column, where $1 \leq i \leq m$, $1 \leq j \leq n$.

$\sigma^2\mathbf{V}$, where $\mathbf{V}$ is the correlation matrix. If the errors are independent and identically distributed (i.i.d.) with constant variance, then $\text{cov}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. In fact, the errors for the field experiments are likely to be correlated (Martin, 1982).

Often, we are interested in estimating $\boldsymbol{\theta}$, the treatment means. In field experiments, treatment means are usually the relative productivities under different treatments.

1.3 Estimation of Treatment Means θ

If the correlation matrix of the errors, $\mathbf{V}$, is known, the treatment means $\boldsymbol{\theta}$ can be estimated by generalized least squares estimator (GLSE), i.e.

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{G}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}, \tag{1.2}$$

The covariance matrix of $\hat{\boldsymbol{\theta}}_{\boldsymbol{G}}$ is, $\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{G}}) = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$. If the correlation matrix of the errors is unknown, the treatment means are usually estimated by the ordinary least-squares estimator (LSE) , i.e.

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \tag{1.3}$$

and the covariance matrix of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ is, $\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$.

1.4 Criteria for Optimal Designs

Designing to best estimate model parameters leads to optimal designs. The aim is to find arrangements of treatments into field plots that will provide the most accurate information about the true treatment means. Various authors have considered a wide variety of optimality criteria to construct optimal designs. If we are interested in the estimation of $\boldsymbol{\theta}$, we can minimize some scalar loss functions of the covariance matrix of an estimate, $\mathcal{L}(\text{cov}(\hat{\boldsymbol{\theta}}))$. When $\mathcal{L}$ = trace, the design is called A-optimal design; when $\mathcal{L}$ = determinant, the design is called D-optimal design (Becher, 1988).

Sometimes we are not only interested in estimating treatment means, but also comparing different treatment means and determining if any treatment means are significantly different from the others. For example, if we have t treatments, and want to find out if the control treatment mean, say θ_1, is significantly different from other treatment means, denoted $\theta_2, \ldots, \theta_t$, then instead of minimizing $\mathcal{L}(\text{cov}(\hat{\boldsymbol{\theta}}))$, we can minimize $\mathcal{L}(\text{cov}(\mathbf{C}\hat{\boldsymbol{\theta}}))$, in which $\mathbf{C}$ is a *(t-1)*xt contrast matrix, such that

$$\sum_{j=1}^{t} c_{i,j} = 0, \qquad i = 1, \cdots, t-1, \tag{1.4}$$

where $c_{i,j}$ is the element at the ith row and jth column of $\mathbf{C}$. Then $\sum_{j=1}^{t} c_{i,j}\ \theta_j$ is called a contrast, where $i = 1, \cdots, t-1$. In our example, the contrast matrix is

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & -1 & 0 \\ 1 & 0 & \cdots & 0 & -1 \end{pmatrix} \tag{1.5}$$

Notice that with $c_{1,1} = 1$, $c_{1,2} = -1$, and $c_{1,3} = \cdots = c_{1,t} = 0$, $\sum_{j=1}^{t} c_{1,j}\ \theta_j = \theta_1 - \theta_2$ is a contrast comparing θ_1 with θ_2.

1.5 Overview of the Thesis

In this chapter, we have discussed the importance of spatial dependence in an agricultural field experiment, introduced the linear model and estimators of the treatment means, and optimality criteria used for designs and analysis of spatially arranged experiments. The rest of this thesis is organized as follows: In Chapter 2, we will discuss four types of commonly used spatial correlation structures, and present some optimal designs with specified correlation structures. In Chapter 3, we will define a neighbourhood for a covariance matrix, and propose a criterion to find minimax designs that are efficient for a neighbourhood of correlations. In Chapter 4, we will introduce a numerical algorithm that helps us find the minimax designs and compare minimax designs for least squares and generalized least squares estimators.

Main contributions of this thesis are listed below.

1. A neighbourhood of a covariance matrix is discussed, and its properties are explored in details.

2. A robust criterion to find minimax designs is proposed. These designs are robust against misspecification of spatial correlation in field experiments.

3. When the covariance matrix of the error process is unknown, but belongs to a small neighbourhood of a specified covariance matrix, generalized least squares estimator performs better than the ordinary least squares estimator.

4. A simulated annealing algorithm is applied to compute minimax designs. Many minimax designs are derived and compared with optimal designs. We recommend using GLSE and the minimax design based on GLSE in practical applications.

Chapter 2

Spatial Correlation Structures and Optimal Designs

In the first section of this chapter, we discuss four types of spatial correlation structures. They are doubly-geometric correlation, discrete exponential correlation, moving average correlation, and the nearest neighbour correlation. Optimal designs are discussed in Section 2. Results are presented for designs with small number of plots and small number of treatments.

2.1 Spatial Correlation Structures

There are various types of spatial correlation structures proposed for the spatial nature of the experimental material. Here we review four commonly used correlation

structures.

2.1.1 Doubly-Geometric Process

The doubly-geometric process has its name because its correlation function is the product of two geometric terms. It is the extension of the first order autoregressive process in time to the plane. It is used to model the errors $\epsilon_{i,j}$ and has the form of

$$\epsilon_{i,j} = \lambda_1 \epsilon_{i-1,j} + \lambda_2 \epsilon_{i,j-1} - \lambda_1 \lambda_2 \epsilon_{i-1,j-1} + u_{i,j}, \tag{2.1}$$

where λ_1 and λ_2 are the parameters of the doubly-geometric process, and $u_{i,j}$ is i.i.d. with mean 0 and varriance σ^2. For square plots $\lambda_1 = \lambda_2 = \lambda$; for nonsquare plots, one choice is $\lambda_1 = \lambda^2$, $\lambda_2 = \sqrt{\lambda}$ which may be viewed as a rectangular plot of halved width and doubled length. Let g be the directed horizontal distance, and h be the directed vertical distance between two plots. The correlation between any two plots, denoted as $\rho_{g,h} = \text{corr}(\epsilon_{ij}, \epsilon_{kl}) = \lambda_1^{|g|} \lambda_2^{|h|}$, where $g = i - k$ and $h = j - l$, generalizes the autocorrelation structure of the one dimensional Markov process. The doubly-geometric process is used often because it has a few useful second-order properties (Martin, 1979). One of these properties is that the prediction estimate from the lower left corner is $\hat{\epsilon}_{i+h,j+k} = \lambda_1^h \lambda_2^k \epsilon_{i,j}$. This correlation structure will subsequently be denoted by DG.

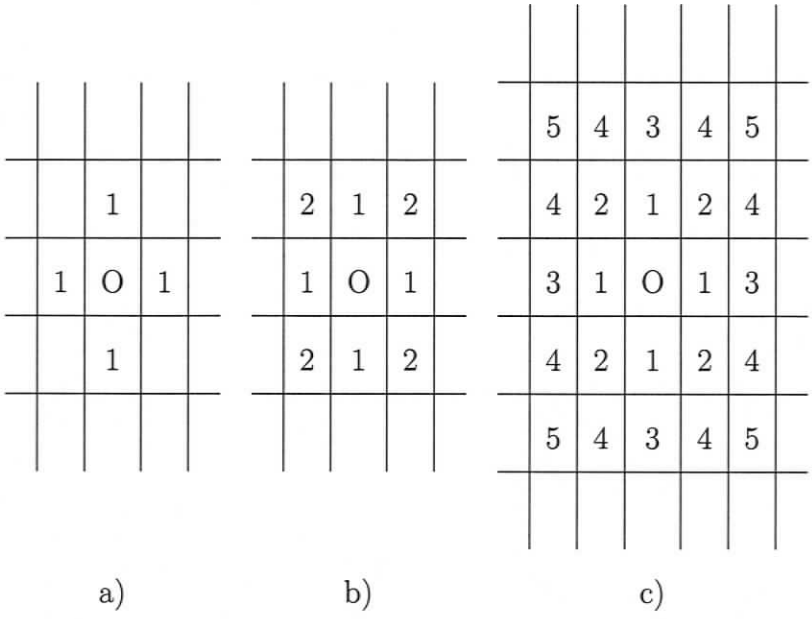


Figure 2.1: a) First-order neighbours. b) First and second-order neighbours. c) First to fifth-order neighbours. *The number represents the order of a neighbour is to a plot O.*

2.1.2 Discrete Exponential Process

The discrete exponential process is the discrete version of the continuous planar exponential process (Duby et al., 1997) where the correlation between two points on the plane is decreasing exponentially with the distance. The correlation between two plots can be represented by $\rho_{g,h} = \lambda^{\sqrt{\nu g^2 + (1/\nu)h^2}}$. It is assumed that $\nu = 1$ for square plots and $\nu \neq 1$ for nonsquare plots. For example, if the horizontal correlation is twice as strong as the vertical correlation, then ν is set to 2. This correlation structure will subsequently be denoted by DE.

2.1.3 Nearest Neighbour Correlation

Nearest neighour correlation structure (Kiefer, 1981) has been one of the most commonly used correlations in spatial layouts. Neighbours are the plots which are adjacent to a plot. It is reasonable to assume that there are no directional effects in the nearest neighbour structure. Thus, a left neighbour is equivalent to a right neighbour. Figure 2.1 presents the first-order, second-order, and fifth-order neighbours to a plot O. For the first-order nearest neighours (NN(1)) (see Figure 2.1a) , the correlation function $\rho_{1,0} = \rho_{0,1} = \rho_{-1,0} = \rho_{0,-1} = \rho$, $\rho_{g,h} = 0$ for $|g| + |h| > 1$, and in general, $|\rho| < \frac{1}{4}$.

2.1.4 Moving Average Correlation

The two-dimensional moving average generalization is found to be useful in the analysis of spatial correlation as well (Martin, 1986). A simple form of the moving average process (MA) is

$$\epsilon_{i,j} = \alpha u_{i-1,j} + \beta u_{i+1,j} + \gamma u_{i,j-1} + \delta u_{i,j+1} + u_{i,j}, \tag{2.2}$$

where α, β, γ and δ are constant coefficients. The autocorrelations are given in Table 2.1.

As a special case $\alpha = \beta = \delta = \gamma$ and equation (2.2) becomes

$$\epsilon_{i,j} = \gamma u_{i-1,j} + \gamma u_{i+1,j} + \gamma u_{i,j-1} + \gamma u_{i,j+1} + u_{i,j}. \tag{2.3}$$

$\rho_{0,1} = \rho_{0,-1} = (\delta + \gamma)/K$	$\rho_{1,0} = \rho_{-1,0} = (\alpha + \beta)/K$
$\rho_{0,2} = \rho_{0,-2} = (\delta\gamma)/K$	$\rho_{2,0} = \rho_{-2,0} = (\alpha\beta)/K$
$\rho_{1,1} = \rho_{-1,-1} = (\alpha\delta + \beta\gamma)/K$	$\rho_{-1,1} = \rho_{1,-1} = (\alpha\gamma + \beta\delta)/K$
$\rho_{i,j} = 0$ for $\|i\| + \|j\| > 2$	$K = 1 + \alpha^2 + \beta^2 + \delta^2 + \gamma^2$

Table 2.1: Autocorrelation for the moving average model in Equation (2.2)

In this case, the autocorrelation function is given by $\rho_{1,0} = 2\gamma/(1 + 4\gamma^2) = \rho$, $\rho_{2,0} = \frac{1}{2}\gamma\rho$, and $\rho_{1,1} = \gamma\rho$. This particular correlation structure will subsequently be denoted by MA(1), and illustrated in Figure 2.2. Derivations are shown below:

		$\frac{1}{2}\gamma\rho$		
	$\gamma\rho$	ρ	$\gamma\rho$	
$\frac{1}{2}\gamma\rho$	ρ	O	ρ	$\frac{1}{2}\gamma\rho$
	$\gamma\rho$	ρ	$\gamma\rho$	
		$\frac{1}{2}\gamma\rho$		

Figure 2.2: Autocorrelations to a plot O for MA(1) model in Equation (2.3).

$$
\begin{aligned}
cov(\epsilon_{i,j}, \epsilon_{i,j}) &= var(\epsilon_{i,j}) \\
&= E((\gamma u_{i-1,j} + \gamma u_{i+1,j} + \gamma u_{i,j-1} + \gamma u_{i,j+1} + u_{i,j})^2) \\
&= (1 + 4\gamma^2)\sigma^2, \\
cov(\epsilon_{i,j}, \epsilon_{i+1,j}) &= cov(\gamma u_{i-1,j} + \gamma u_{i+1,j} + \gamma u_{i,j-1} + \gamma u_{i,j+1} + u_{i,j}, \\
&\quad \gamma u_{i,j} + \gamma u_{i+2,j} + \gamma u_{i+1,j-1} + \gamma u_{i+1,j+1} + u_{i+1,j}) \\
&= 2\gamma\sigma^2, \\
\rho_{1,0} &= \frac{cov(\epsilon_{i,j}, \epsilon_{i+1,j})}{var(\epsilon_{i,j})} = \frac{2\gamma\sigma^2}{(1 + 4\gamma^2)\sigma^2} = \rho \\
cov(\epsilon_{i,j}, \epsilon_{i+1,j+1}) &= cov(\gamma u_{i-1,j} + \gamma u_{i+1,j} + \gamma u_{i,j-1} + \gamma u_{i,j+1} + u_{i,j}, \\
&\quad \gamma u_{i,j+1} + \gamma u_{i+2,j+1} + \gamma u_{i+1,j} + \gamma u_{i+1,j+2} + u_{i+1,j+1}) \\
&= 2\gamma^2\sigma^2, \\
\rho_{1,1} &= \frac{cov(\epsilon_{i,j}, \epsilon_{i+1,j+1})}{var(\epsilon_{i,j})} = \frac{2\gamma^2\sigma^2}{(1 + 4\gamma^2)\sigma^2} = \gamma\rho \\
cov(\epsilon_{i,j}, \epsilon_{i+2,j}) &= cov(\gamma u_{i-1,j} + \gamma u_{i+1,j} + \gamma u_{i,j-1} + \gamma u_{i,j+1} + u_{i,j}, \\
&\quad \gamma u_{i+1,j} + \gamma u_{i+3,j} + \gamma u_{i+2,j-1} + \gamma u_{i+2,j+1} + u_{i+2,j}) \\
&= \gamma^2\sigma^2, \\
\rho_{2,0} &= \frac{cov(\epsilon_{i,j}, \epsilon_{i+2,j})}{var(\epsilon_{i,j})} = \frac{\gamma^2\sigma^2}{(1 + 4\gamma^2)\sigma^2} = \frac{1}{2}\gamma\rho.
\end{aligned}
$$

In addition, it can easily be shown that:

$$\begin{aligned}
\rho_{1,0} &= \rho_{0,-1} = \rho_{0,1} = \rho_{-1,0}, \\
\rho_{2,0} &= \rho_{0,-2} = \rho_{0,2} = \rho_{-2,0}, \\
\rho_{1,1} &= \rho_{-1,-1} = \rho_{-1,1} = \rho_{1,-1}, and \\
\rho_{i,j} &= 0 \; for \; |i| + |j| > 2.
\end{aligned}$$

2.2 Optimal Designs

Optimal designs usually depend on the criterion used, spatial correlation structures, the actual parameter values of spatial correlations, and estimation methods. Martin (1996) points out that optimal designs in most cases have been considered only for positive dependence, because negative dependence is unlikely in practice and leads to optimal designs which would not be used. Thus, we consider only positive dependence in this thesis as well.

Some of the optimal designs lead to Latin Square designs. A Latin Square is a txt array such that every one of t treatments appears exactly once in each row and in each column. Martin (1986) considered txt designs with $r = t$ for $t = 3, 4, 5$ using the D-optimal and A-optimal criteria. Usually the optimal design is a Latin Square, but for some processes with GLSE method the design shown in Figure 2.3 is optimal.

1	2	3
3	1	2
1	2	3

Figure 2.3: An example of optimal 3x3 design from Martin (1986)

1	2	3	4	1	2	3	4	1	2	1	2	1	2	3	4
3	4	1	2	3	4	1	2	3	4	3	4	2	1	4	3
2	1	4	3	1	2	3	4	4	3	4	3	3	4	1	2
4	3	2	1	3	4	1	2	2	1	2	1	4	3	2	1
1)				2)				3)				4)			

Figure 2.4: Some of optimal 4x4 designs

Becher (1988) gave numerical comparisons for square and nonsquare 4x4 designs. Some of the designs considered are given in Figure 2.4. For square plots, design 1 is A-optimal under the correlation structure DG for $\lambda \in (0,1)$. Design 2 is also A-optimal under the correlation structure DE for weak correlations, $\lambda < 0.221$; otherwise, design 1 is A-optimal. For nonsquare plots, design 2 is A-optimal for both correlation structures DG and DE under weak correlations. For stronger correlations ($\lambda > 0.250$) of DG design 1 is A-optimal, and for $\lambda > 0.135$ and DE design 3 is A-optimal.

The comparisons using the D-criterion give somewhat different results. For square plots, designs 1 and 4 are D-optimal under correlation structure DG. Under correlation structure DE design 2 is D-optimal for weak correlations ($\lambda < 0.275$), otherwise design 1 is D-optimal. For nonsquare plots design 3 is D-optimal for $\lambda \in (0,1)$ under DE and for $\lambda \in (0,0.563]$ under DG. For stronger correlations, designs 1 and 4 are again D-optimal under this correlation structure.

For small designs, complete enumeration and comparison might be possible. As the dimensions of the plots increase, the number of designs increases dramatically. In the following subsection, comparisons are made for all designs for 2x2 plots with 2 treatments, 2x3 plots with 2 treatments and 3x3 plots with 3 treatments, since there are 6 designs for 2x2 plots, 20 designs for 2x3 plots and 1680 designs for 3x3 plots. However, there are too many designs for $m > 3$ and $n > 3$ to consider all. There are 330,266 designs for 4x4 plots with 4 treatments, and it is feasible to consider only a

small subset of designs in detail.

Since we are interested in finding A-optimal and D-optimal designs for each of the four correlation structures mentioned previously, we will minimize loss functions $\mathcal{L} = \text{trace}(\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}))$ and $\mathcal{L} = \det(\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}))$ accordingly. In general, for MA(1), the range of γ is [0, 0.25] (Martin, 1986); for NN(1) the range of ρ is [0, 0.25] (Martin, 1986); for DG and DE, the range of λ is (0, 1) (Becher, 1988).

2.2.1 Designs for 2x2 plots

There are six designs in a 2x2 plots with $t = 2$ treatments. Figure 2.5 shows all the 2x2 designs.

1	1	2	2	1	2	2	1	1	2	2	1
2	2	1	1	1	2	2	1	2	1	1	2
1)		2)		3)		4)		5)		6)	

Figure 2.5: All 2x2 designs

If treatment labels are randomly assigned 1 or 2, then Design 1) is considered the same as Design 2), Design 3) is the same as Design 4), and Design 5) is the same as Design 6). It can be shown that Designs 1), 2), 3) and 4) yield the same covariance of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$: $\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$. So in the following comparisons, only

Designs 1) and 5) are considered. For Design 1):

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \tag{2.4}$$

and for Design 5,

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.5}$$

For correlation structure MA(1),

$$\mathbf{V} = \begin{pmatrix} 1 & \rho & \rho & \gamma\rho \\ \rho & 1 & \gamma\rho & \rho \\ \rho & \gamma\rho & 1 & \rho \\ \gamma\rho & \rho & \rho & 1 \end{pmatrix}. \tag{2.6}$$

For correlation structure DG, $\lambda_1 = \lambda_2 = \lambda$ for square plots, and

$$\mathbf{V} = \begin{pmatrix} 1 & \lambda & \lambda & \lambda^2 \\ \lambda & 1 & \lambda^2 & \lambda \\ \lambda & \lambda^2 & 1 & \lambda \\ \lambda^2 & \lambda & \lambda & 1 \end{pmatrix}. \tag{2.7}$$

For correlation structure DE, ν is set to 1 for square plots, and

$$\mathbf{V} = \begin{pmatrix} 1 & \lambda & \lambda & \lambda^{\sqrt{2}} \\ \lambda & 1 & \lambda^{\sqrt{2}} & \lambda \\ \lambda & \lambda^{\sqrt{2}} & 1 & \lambda \\ \lambda^{\sqrt{2}} & \lambda & \lambda & 1 \end{pmatrix}. \tag{2.8}$$

For correlation structures NN(1),

$$\mathbf{V} = \begin{pmatrix} 1 & \rho & \rho & 0 \\ \rho & 1 & 0 & \rho \\ \rho & 0 & 1 & \rho \\ 0 & \rho & \rho & 1 \end{pmatrix}. \tag{2.9}$$

Figures 2.6, 2.7, 2.8 and 2.9 show loss function versus correlation parameter for designs 1) and 5). It is obvious that Design 5) yields the minimum loss functions for all four correlation structures for all correlation parameter values. Therefore Design 5) is an optimal design and robust against correlation structures.

2.2.2 Designs for 2x3 plots

There are 20 designs in 2x3 plots with t = 2 treatments, and some are listed in Figure 2.10.

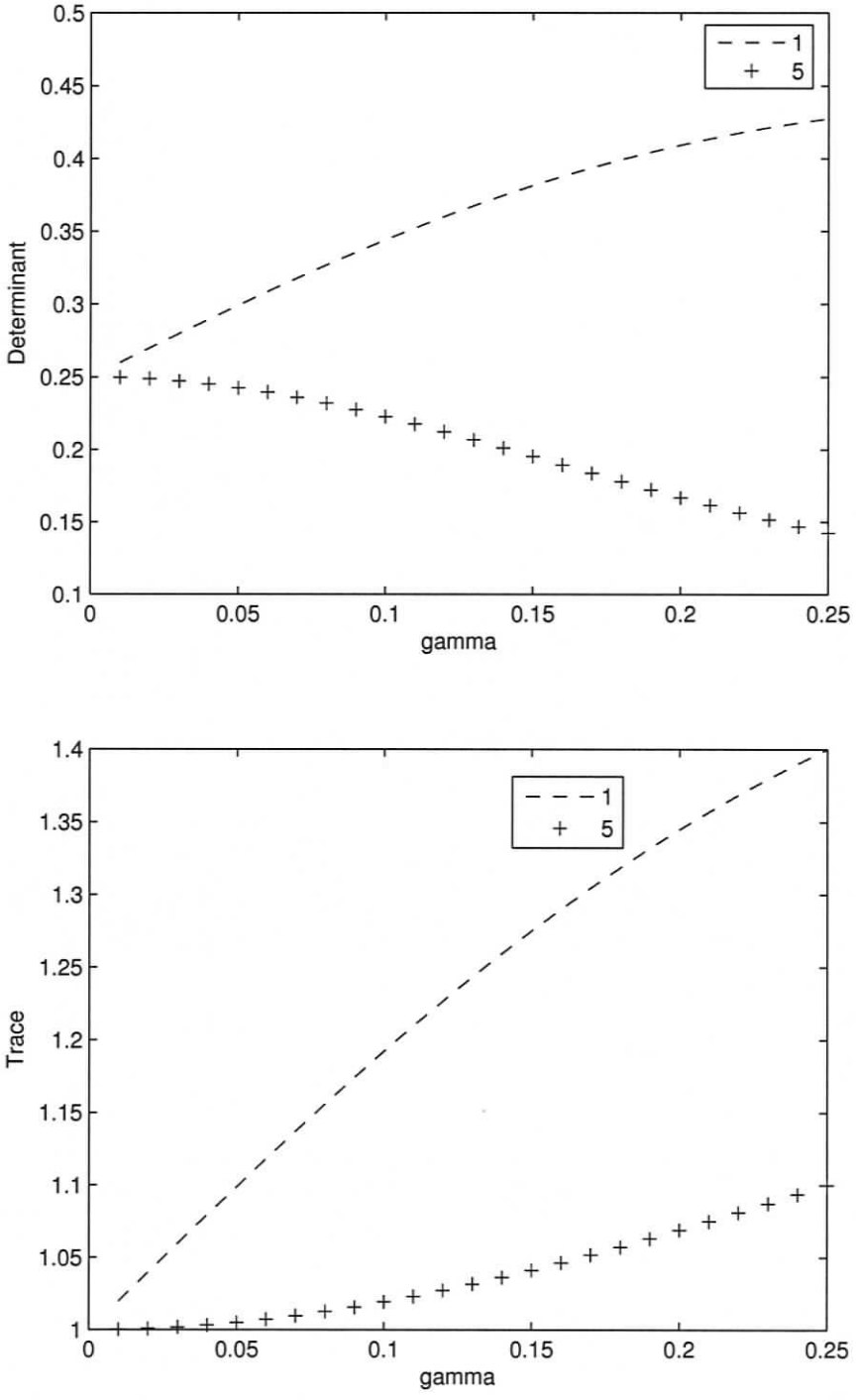


Figure 2.6: Loss functions of 2x2 designs for MA(1).

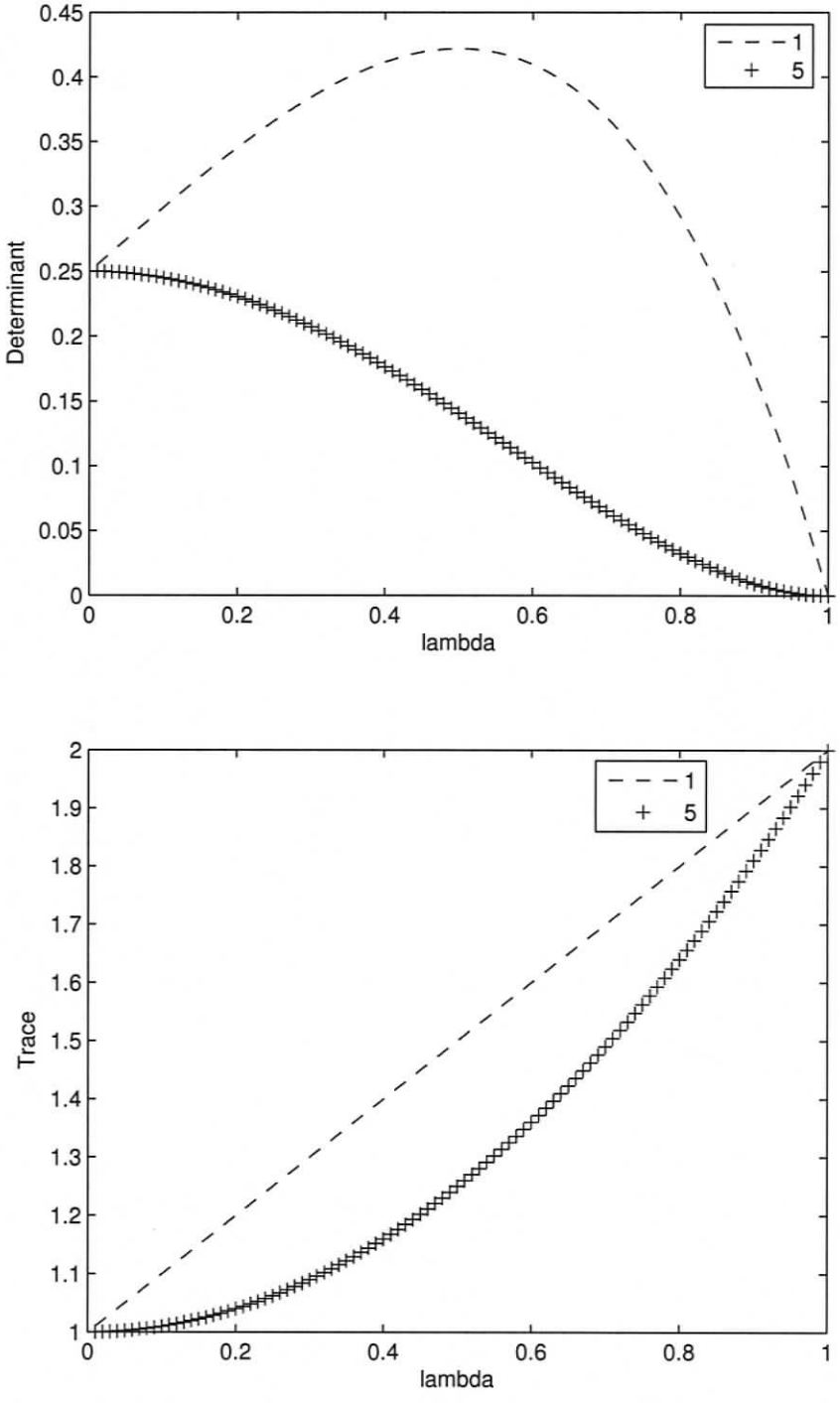


Figure 2.7: Loss functions of 2x2 designs for DG.

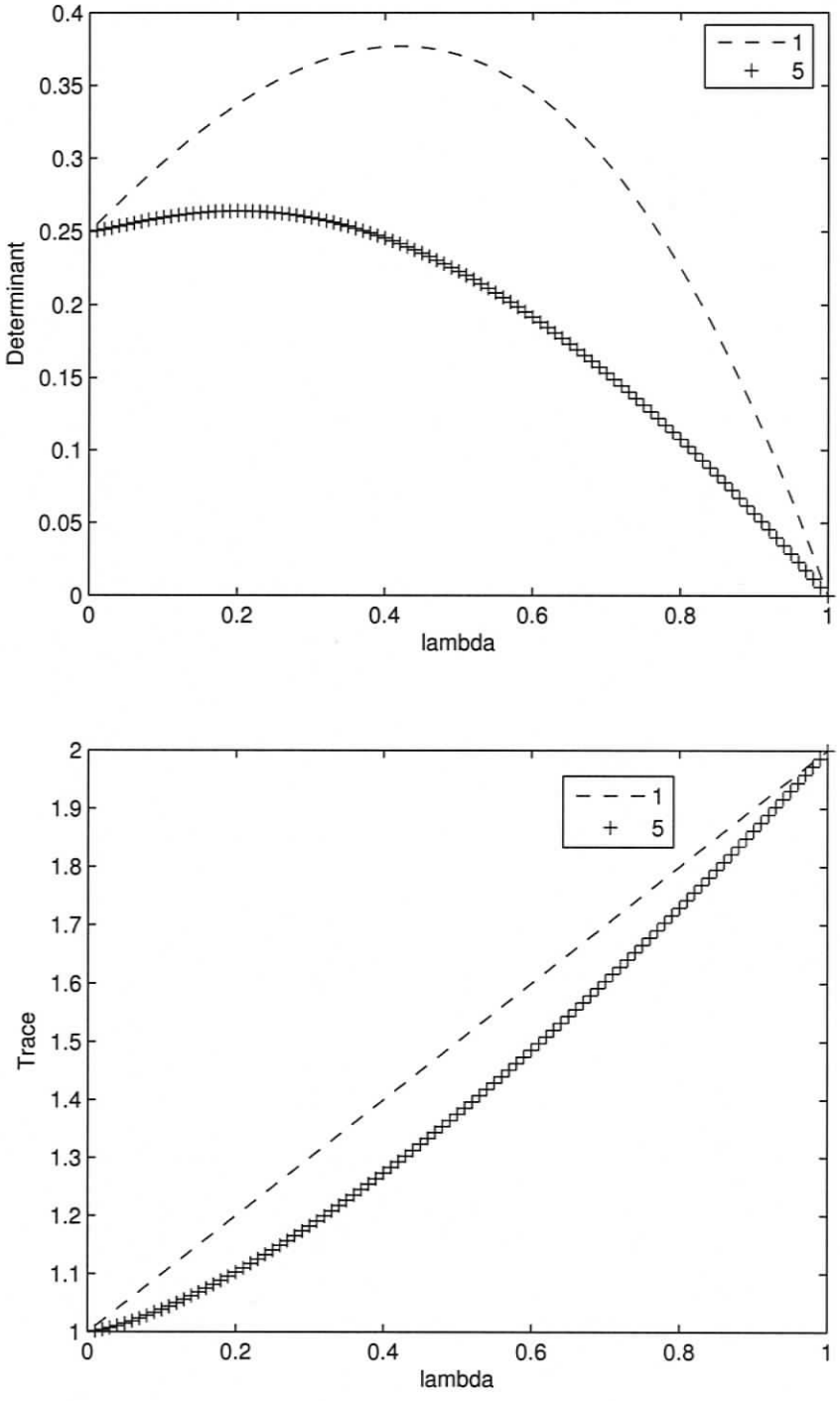


Figure 2.8: Loss functions of 2x2 designs for DE.

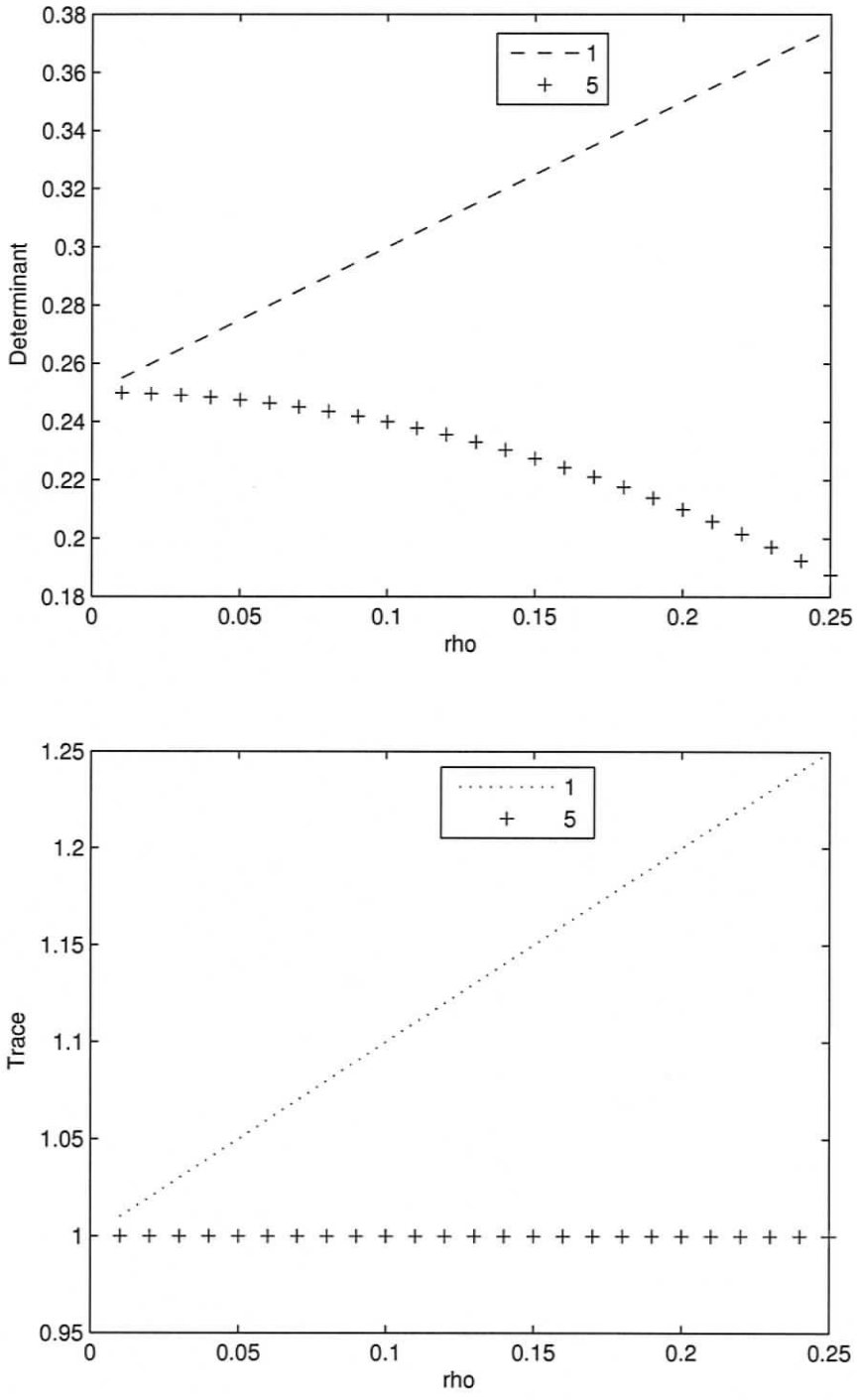


Figure 2.9: Loss functions of 2X2 designs for NN(1).

1	1	1	1	1	2	1	1	2	1	1	2	1	2	1
2	2	2	1	2	2	2	1	2	2	2	1	1	2	2
	1)			2)			3)			4)			5)	
1	2	1	1	2	1	1	2	2	1	2	2	1	2	2
2	1	2	2	2	1	1	1	2	1	2	1	2	1	1
	6)			7)			8)			9)			10)	

Figure 2.10: Some 2x3 designs

For Design 1),

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \tag{2.10}$$

and for Design 6),

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.11}$$

For correlation structures MA(1),

$$\mathbf{V} = \begin{pmatrix} 1 & \rho & \frac{1}{2}\gamma\rho & \rho & \gamma\rho & 0 \\ \rho & 1 & \rho & \gamma\rho & \rho & \gamma\rho \\ \frac{1}{2}\gamma\rho & \rho & 1 & 0 & \gamma\rho & \rho \\ \rho & \gamma\rho & 0 & 1 & \rho & \frac{1}{2}\gamma\rho \\ \gamma\rho & \rho & \gamma\rho & \rho & 1 & \rho \\ 0 & \gamma\rho & \rho & \frac{1}{2}\gamma\rho & \rho & 1 \end{pmatrix}. \tag{2.12}$$

For correlation structure DG, $\lambda_1 = \lambda_2 = \lambda$ for square plots, and

$$\mathbf{V} = \begin{pmatrix} 1 & \lambda & \lambda^2 & \lambda & \lambda^2 & \lambda^3 \\ \lambda & 1 & \lambda & \lambda^2 & \lambda & \lambda^2 \\ \lambda^2 & \lambda & 1 & \lambda^3 & \lambda^2 & \lambda \\ \lambda & \lambda^2 & \lambda^3 & 1 & \lambda & \lambda^2 \\ \lambda^2 & \lambda & \lambda^2 & \lambda & 1 & \lambda \\ \lambda^3 & \lambda^2 & \lambda & \lambda^2 & \lambda & 1 \end{pmatrix}. \tag{2.13}$$

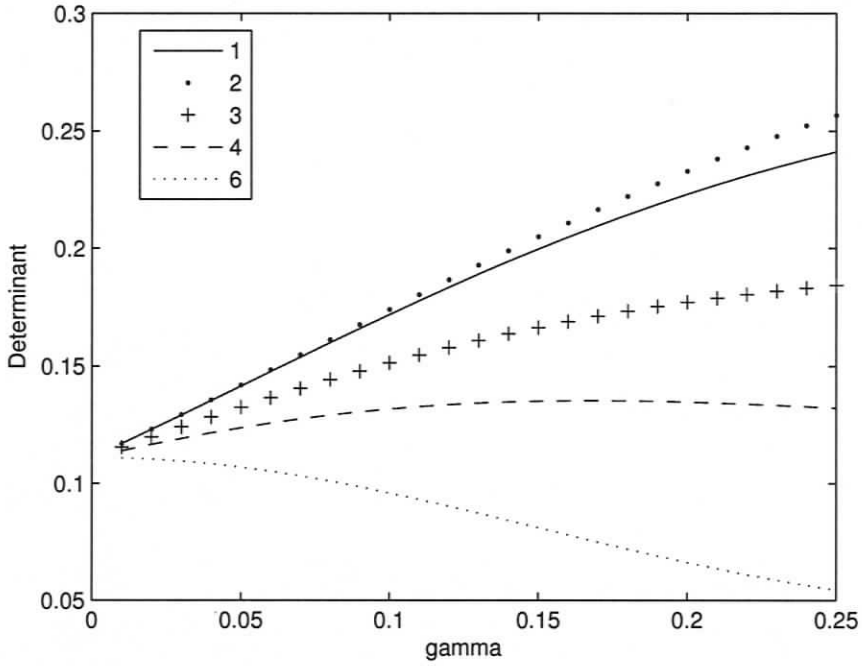
For correlation structure DE, ν is set to 1 for square plots, and

$$\mathbf{V} = \begin{pmatrix} 1 & \lambda & \lambda^2 & \lambda & \lambda^{\sqrt{2}} & \lambda^{\sqrt{5}} \\ \lambda & 1 & \lambda & \lambda^{\sqrt{2}} & \lambda & \lambda^{\sqrt{2}} \\ \lambda^2 & \lambda & 1 & \lambda^{\sqrt{5}} & \lambda^{\sqrt{2}} & \lambda \\ \lambda & \lambda^{\sqrt{2}} & \lambda^{\sqrt{5}} & 1 & \lambda & \lambda^2 \\ \lambda^{\sqrt{2}} & \lambda & \lambda^{\sqrt{2}} & \lambda & 1 & \lambda \\ \lambda^{\sqrt{5}} & \lambda^{\sqrt{2}} & \lambda & \lambda^2 & \lambda & 1 \end{pmatrix}. \tag{2.14}$$

For correlation structure NN(1),

$$\mathbf{V} = \begin{pmatrix} 1 & \rho & 0 & \rho & 0 & 0 \\ \rho & 1 & \rho & 0 & \rho & 0 \\ 0 & \rho & 1 & 0 & 0 & \rho \\ \rho & 0 & 0 & 1 & \rho & 0 \\ 0 & \rho & 0 & \rho & 1 & \rho \\ 0 & 0 & \rho & 0 & \rho & 1 \end{pmatrix}. \tag{2.15}$$

Again, because the treatment labels are randomly assigned 1 or 2, there are essentially 10 designs (Figure 2.10) to consider. It can be also shown that Designs 2) and 8) yield the same covariance of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$; Designs 3), 5), 7), and 9) yield the same covariance of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$; Designs 4) and 10) yield the same covariance of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$. As it is shown in Figures 2.11 - 2.14, Design 6) yields the minimum loss functions under the two criteria for all four correlation structures.



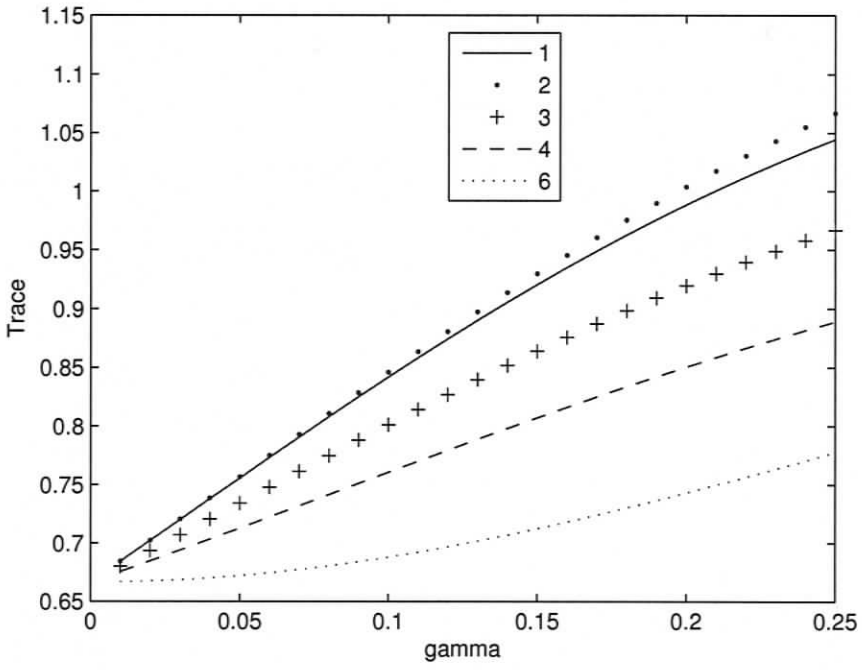


Figure 2.11: Loss functions of 2x3 designs for MA(1)

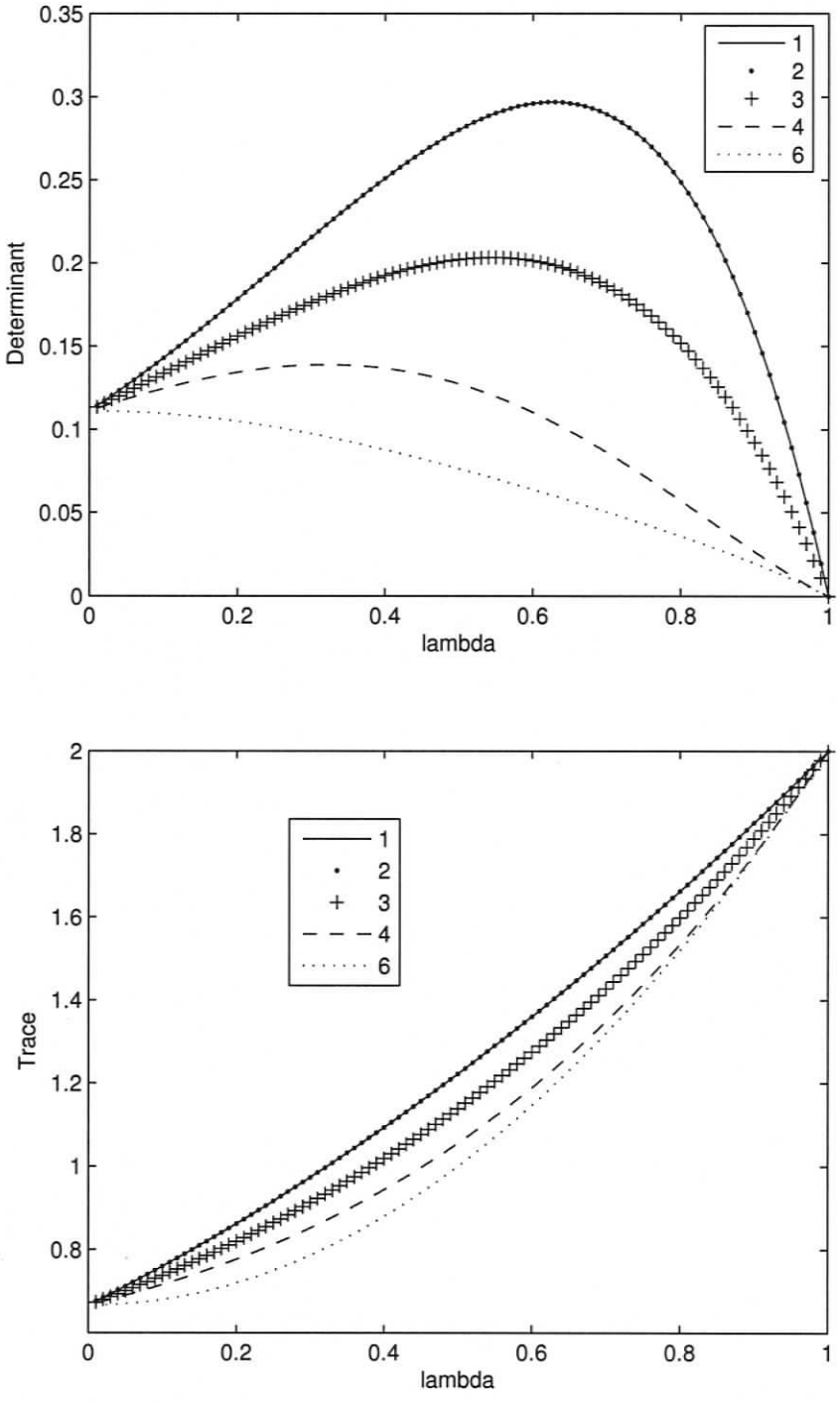


Figure 2.12: Loss functions of 2x3 designs for DG

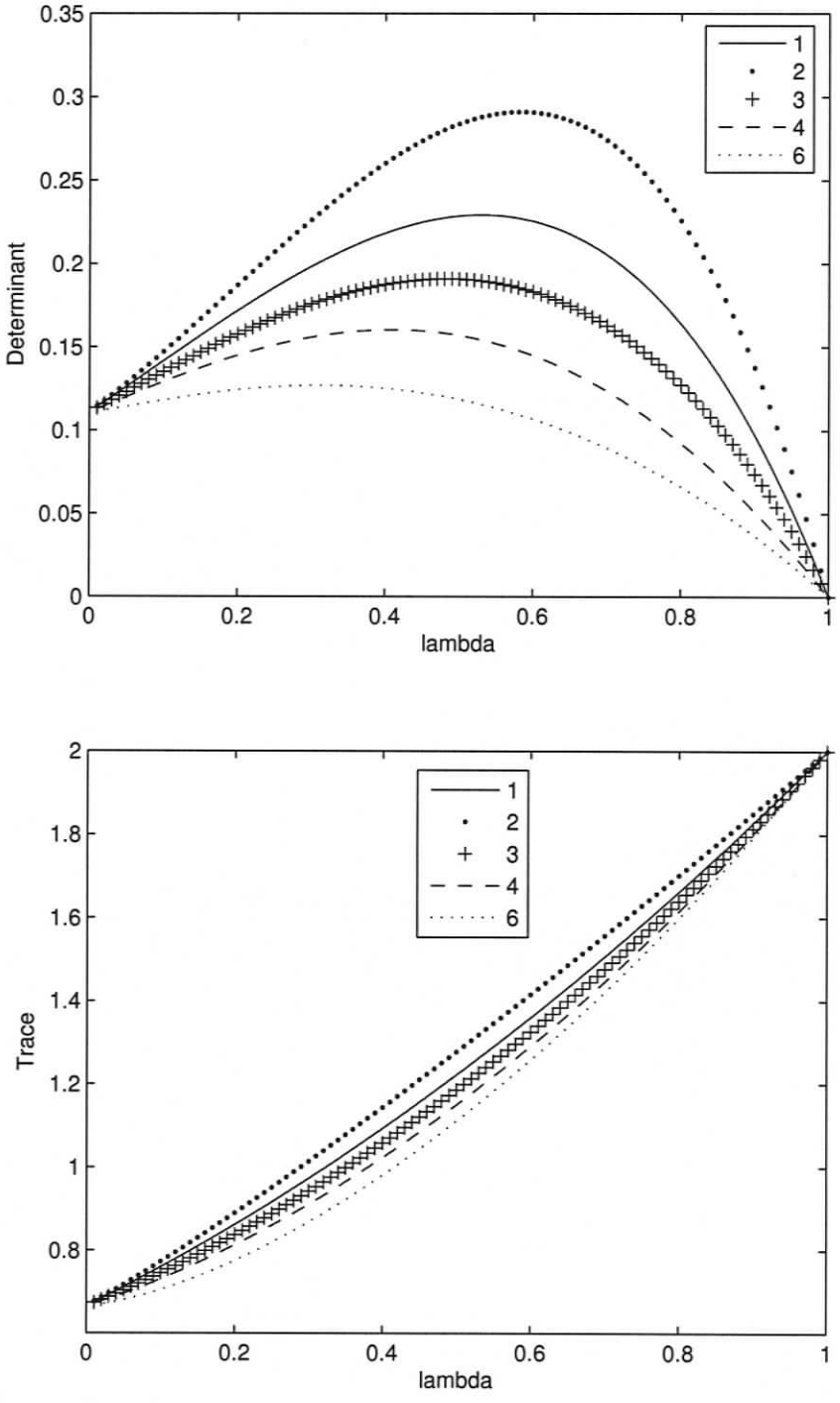


Figure 2.13: Loss functions of 2x3 designs for DE

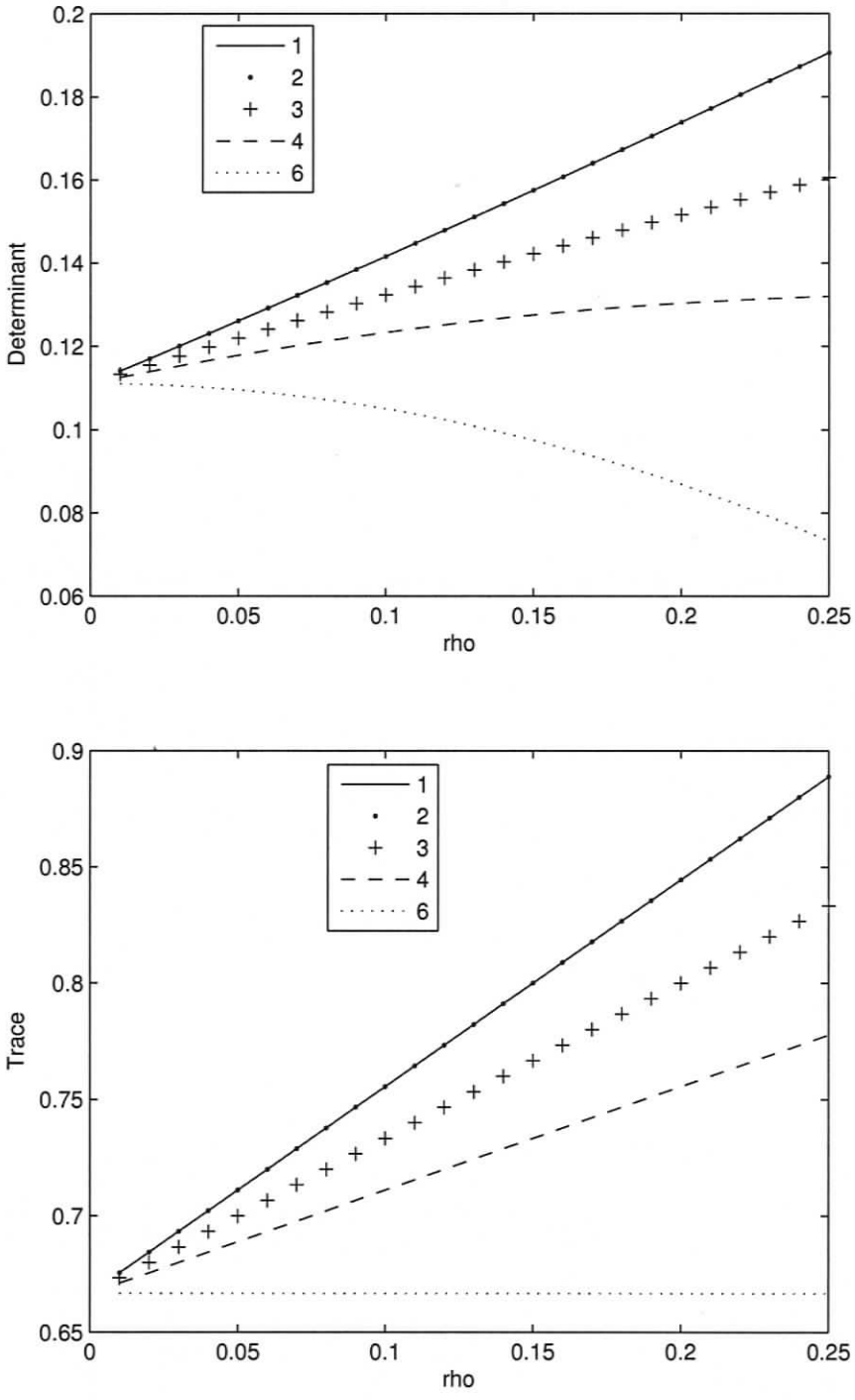


Figure 2.14: Loss functions of 2x3 designs for NN(1)

2.2.3 Designs for 2xn plots

If we extend n to any positive integer, with $t = 2$ treatments, we have the following general results for NN(1).

Theorem 2.1. *For 2xn plots with $t = 2$ treatments, under NN(1) the A-optimal designs minimizing trace(cov($\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$)) are the designs in which every plot has a treatment different from its nearest neighbours.*

Proof. Denote the 2xn plots as following:

1	2	$\cdots$	n-1	n
n+1	n+2	$\cdots$	2n-1	2n

The design matrix

$$\mathbf{X}' = \begin{pmatrix} x_{1,1} & x_{2,1} & \cdots & x_{2n-1,1} & x_{2n,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{2n-1,2} & x_{2n,2} \end{pmatrix}, \tag{2.16}$$

where $x_{i,j}$ is either 0 or 1, and $x_{i,1} + x_{i,2} = 1$. When $x_{i,1} = 1$, treatment 1 is applied to plot i; when $x_{i,2} = 1$, treatment 2 is applied to plot i. The correlation matrix under NN(1), $\mathbf{V}$, can be partitioned into four nxn submatrices, i.e.,

$$\mathbf{V} = \begin{pmatrix} \mathbf{V_1} & \mathbf{V_2} \\ \mathbf{V_2} & \mathbf{V_1} \end{pmatrix}, \tag{2.17}$$

where

$$\mathbf{V_1} = \begin{pmatrix} 1 & \rho & 0 & \cdots & 0 \\ \rho & 1 & \rho & \cdots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \cdots & \rho & 1 & \rho \\ 0 & \cdots & 0 & \rho & 1 \end{pmatrix} \quad and \quad \mathbf{V_2} = \rho \mathbf{I_n}.$$

The covariance matrix of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ is, $\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) = \sigma^2(\mathbf{X'X})^{-1}\mathbf{X'VX}(\mathbf{X'X})^{-1}$. Here,

$$\mathbf{X'X} = \begin{pmatrix} n & 0 \\ 0 & n \end{pmatrix}. \tag{2.18}$$

Let $\mathbf{A} = \mathbf{X'VX} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$, where

$$\begin{aligned} a_{1,1} &= \sum_{i=1}^{2n} x_{i,1}^2 + 2\rho[\sum_{i=1}^{n-1} x_{i,1}x_{i+1,1} + \sum_{i=n+1}^{2n-1} x_{i,1}x_{i+1,1} \\ &\quad + \sum_{i=1}^{n} x_{i,1}x_{i+n,1}] \\ &:= \sum_{i=1}^{2n} x_{i,1}^2 + 2\rho s_1 \\ &= n + 2\rho s_1, \end{aligned}$$

$$
\begin{aligned}
a_{2,2} &= \sum_{i=1}^{2n} x_{i,2}^2 + 2\rho[\sum_{i=1}^{n-1} x_{i,2}x_{i+1,2} + \sum_{i=n+1}^{2n-1} x_{i,2}x_{i+1,2} \\
&\quad + \sum_{i=1}^{n} x_{i,2}x_{i+n,2}] \\
&= \sum_{i=1}^{2n} x_{i,2}^2 + 2\rho s_2 \\
&:= n + 2\rho s_2, \\
a_{1,2} &= \sum_{i=1}^{2n} x_{i,1}x_{i,2} + \rho[\sum_{i=1}^{n-1}(x_{i,1}x_{i+1,2} + x_{i,2}x_{i+1,1}) \\
&\quad + \sum_{i=n+1}^{2n-1}(x_{i,1}x_{i+1,2} + x_{i,2}x_{i+1,1}) + \sum_{i=1}^{n}(x_{i,1}x_{i+n,2} + x_{i,2}x_{i+n,1})] \\
&:= \sum_{i=1}^{2n} x_{i,1}x_{i,2} + \rho s_3 \\
&= \rho s_3, \\
a_{2,1} &= a_{1,2}, \\
\min_{\mathbf{x}} trace(cov(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) &= \frac{\sigma^2}{n^2} \min_{\mathbf{x}} trace(\mathbf{X}'\mathbf{V}\mathbf{X}) \\
&= \frac{\sigma^2}{n^2} \min_{\mathbf{x}}(a_{1,1} + a_{2,2}) \\
&= \frac{\sigma^2}{n^2} \min_{\mathbf{x}}(2n + 2\rho(s_1 + s_2))
\end{aligned}
$$

For $\rho > 0$,

$$
trace(cov(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) \geq \frac{2\sigma^2}{n}, \tag{2.19}
$$

where the equality holds when

$$\begin{aligned} s_1 + s_2 &= \sum_{j=1}^{2}\sum_{i=1}^{n-1} x_{i,j}x_{i+1,j} + \sum_{j=1}^{2}\sum_{i=n+1}^{2n-1} x_{i,j}x_{i+1,j} \\ &\quad + \sum_{j=1}^{2}\sum_{i=1}^{n} x_{i,j}x_{i+n,j} \\ &= 0. \end{aligned}$$

If every plot has a different treatment from its nearest neighbours, then $trace(cov(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) = \frac{2\sigma^2}{n}$. Therefore, the A-optimal 2xn designs under NN(1) are the designs in which every plot has a treatment different from its neighbours. □

Theorem 2.2. *For* 2xn *plots with* $t = 2$ *treatments, under NN(1) the D-optimal designs minimizing* $det(cov(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}))$ *are the designs in where every plot has a treatment different from its nearest neighbours.*

Proof. Matrices $\mathbf{X}$, $\mathbf{V}$ and $\mathbf{A}$ have the same definitions as in the proof of Theorem 2.1.

$$\begin{aligned} \min_{\mathbf{x}} det(cov(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) &= \frac{\sigma^4}{n^4} \min_{\mathbf{x}} det(\mathbf{X'VX}) \\ &= \frac{\sigma^4}{n^4} \min_{\mathbf{x}} (a_{1,1}a_{2,2} - a_{1,2}a_{2,1}) \\ &= \frac{\sigma^4}{n^4} \min_{\mathbf{x}} [(n + 2\rho s_1)(n + 2\rho s_2) - (\rho s_3)^2] \end{aligned}$$

For $\rho > 0$,

$$det(cov(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) \geq \frac{\sigma^4}{n^2}(1 - \rho^2(3 - \frac{2}{n})^2), \tag{2.20}$$

where the equality holds when $s_1 = s_2 = 0$ and $s_3 = 2n - 2 + n = 3n - 2$. If every plot has a treatment different from its nearest neighbours, then $det(cov(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) = \frac{\sigma^4}{n^2}(1-\rho^2(3-\frac{2}{n})^2)$. Therefore, the D-optimal 2xn designs under NN(1) are the designs in which every plot has a treatment different from its neighbours. □

We can extend this result to any mxn designs with $t = 2$ treatments. The A-optimal and D-optimal designs under NN(1) are the designs in which every plot has a treatment different from its neighbours. Figure 2.15 shows an example of the A-optimal and D-optimal design for mxn designs with $t = 2$ treatments.

$$\begin{array}{ccccc} 1 & 2 & \cdots & 1 & 2 \\ 2 & 1 & \cdots & 2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 2 & \cdots & 1 & 2 \\ 2 & 1 & \cdots & 2 & 1 \end{array}.$$

Figure 2.15: An example of the A-optimal mxn designs with $t = 2$ treatments

2.2.4 Designs for 3x3 plots

There are 1,680 designs for 3 treatments with 3 replicates arranged in a 3x3 square. Figure 2.16 shows an example of this.

1 2 3

1 2 3

2 3 1

Figure 2.16: An example of 3x3 designs with 3 treatments

To generate all the possible designs, we number the plots from 1 to 9, and translate the 3x3 matrix structure into a 1x9 array (Figure 2.17). There are 3 treatments and each has 3 replicates. First, 3 elements are chosen and assigned to the first treatment; another 3 elements are chosen and assigned to the second treatment; the 3 elements left in the array are assigned to the third treatment. Thus there are $\binom{9}{3}\binom{6}{3}\binom{3}{3} = 1{,}680$ possible designs for 3 treatments. Algorithms from Reingold et al. (1977) are used to generate all of these designs.

1	2	3
4	5	6
7	8	9

a)

1	2	3	4	5	6	7	8	9

b)

Figure 2.17: a) The 3x3 matrix. b) The 1x9 array.

The n! permutation of a set $\{a_1, a_2, \ldots, a_n\}$ of distinct elements is among the most commonly generated combinatorial objects. Without loss of generality, we assume the

elements of the set to be the integers 1 to 9 in our case; that is, we consider only permutations of the set $1, 2, \ldots, 9$. The permutations are generated in a sequence of lexicographic order. The sequence of permutations of $\{1, 2, \ldots, n\}$ in lexicographic order is the sequence of permutations in increasing numerical order. For example, the lexicographic sequence of permutations on three elements is 123, 132, 213, 231, 312, 321.

In general, not all the subsets of $\{a_1, a_2, \ldots, a_n\}$ are required. Only those that satisfy some restrictions are required; the most common restriction is the size of the subsets, which applies to our case. Of particular interest are the subsets of size k, the $\binom{n}{k}$ combinations of n things taken k at a time.

To apply this algorithm in our case with 9 elements, we first generate all combinations of size 3 of the integers $\{1, 2, \ldots, 9\}$. For every combination we assign them to treatment number 1. Then we generate combinations of size 3 again with the remaining 6 integers; these combinations of size 3 are assigned to treatment number 2. For example, a subset of size 3 of the integers $\{1, 2, \ldots, 9\}$ can be $\{1, 4, 8\}$; a subset of size 3 from the remaining integers $2, 3, 5, 6, 7, 9$ can be $\{5, 6, 9\}$. Then the design looks like this:

1	3	3
1	2	2
3	1	2

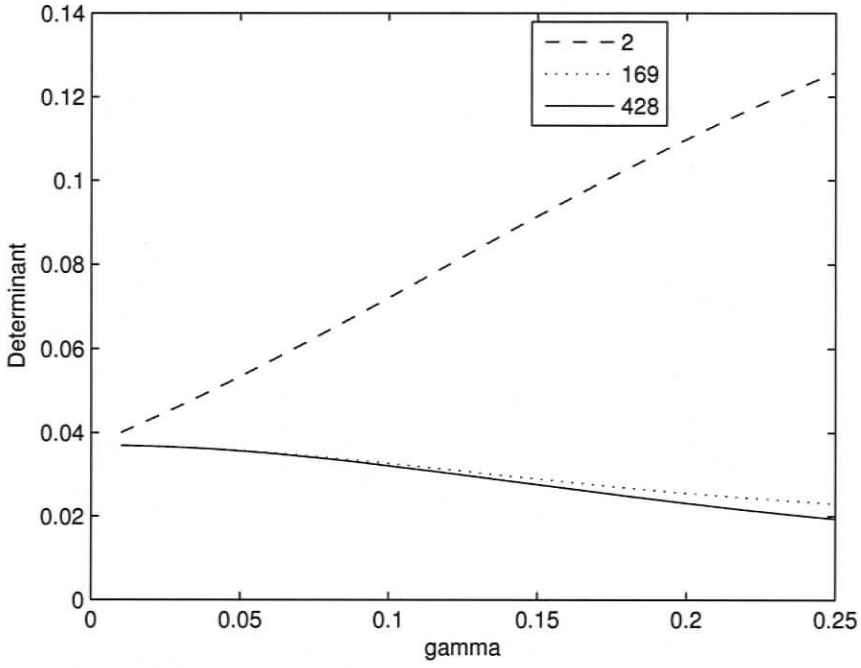
Once all the designs are generated, we calculate the covariance of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ for each design, and find the design with the minimum scalar loss function.

There are 2 designs which are optimal under different optimality criteria. In lexicographic order, they are the 169th and 428th designs. They are not unique as the treatment number is assigned randomly with 1, 2 or 3. There are also other designs similar to these two designs having the same optimality properties. To avoid redundancy, we consider only these two designs (Figure 2.18). Design 2 is considered because it is one of the pessimal designs under all the optimality criteria.

1	1	1	1	2	1	1	2	3
2	2	3	3	1	3	3	1	2
2	3	3	2	3	2	2	3	1
	a)			b)			c)	

Figure 2.18: a) Design 2. b)Design 169. c) Design 428.

Figures 2.19, 2.20, 2.21, and 2.22 show loss functions of these three 3x3 designs for different correlation structures. Design 428 is both A-optimal and D-optimal under MA(1), DG and DE. For NN(1), both Design 169 and 428 are A-optimal, and Design 169 is D-optimal. However Design 428 is very close to D-optimal.



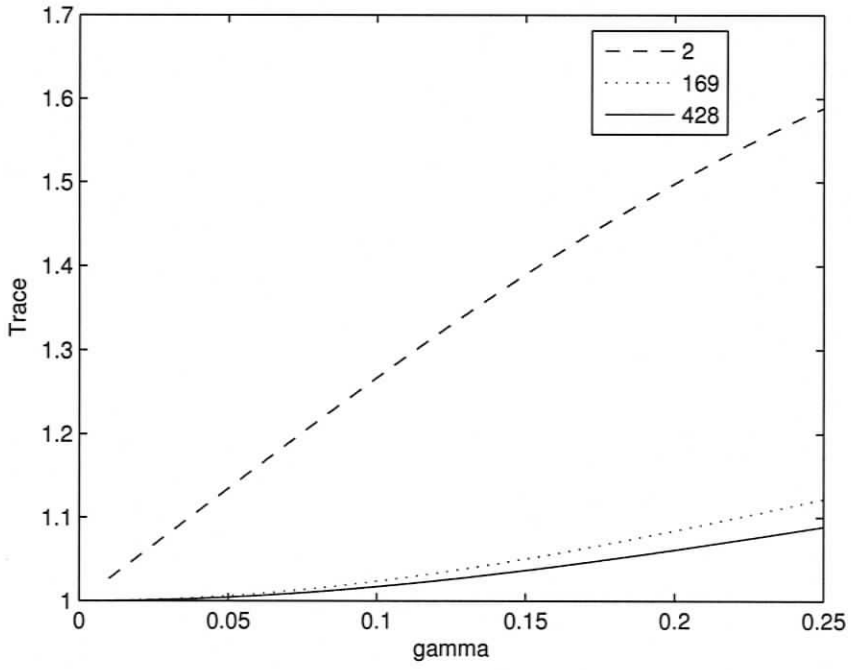


Figure 2.19: Loss functions of some 3x3 designs for MA(1)

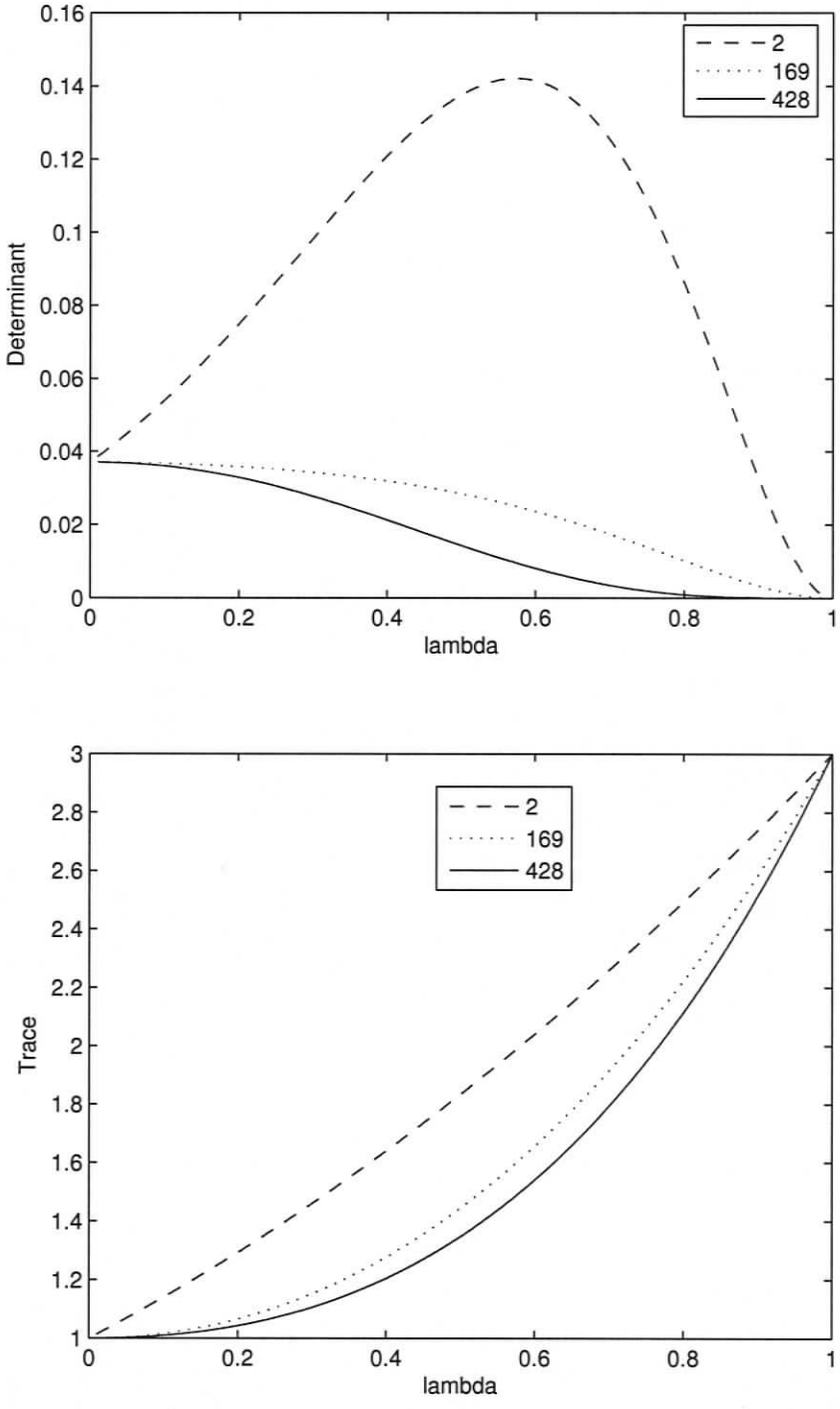


Figure 2.20: Loss functions of some 3x3 designs for DG

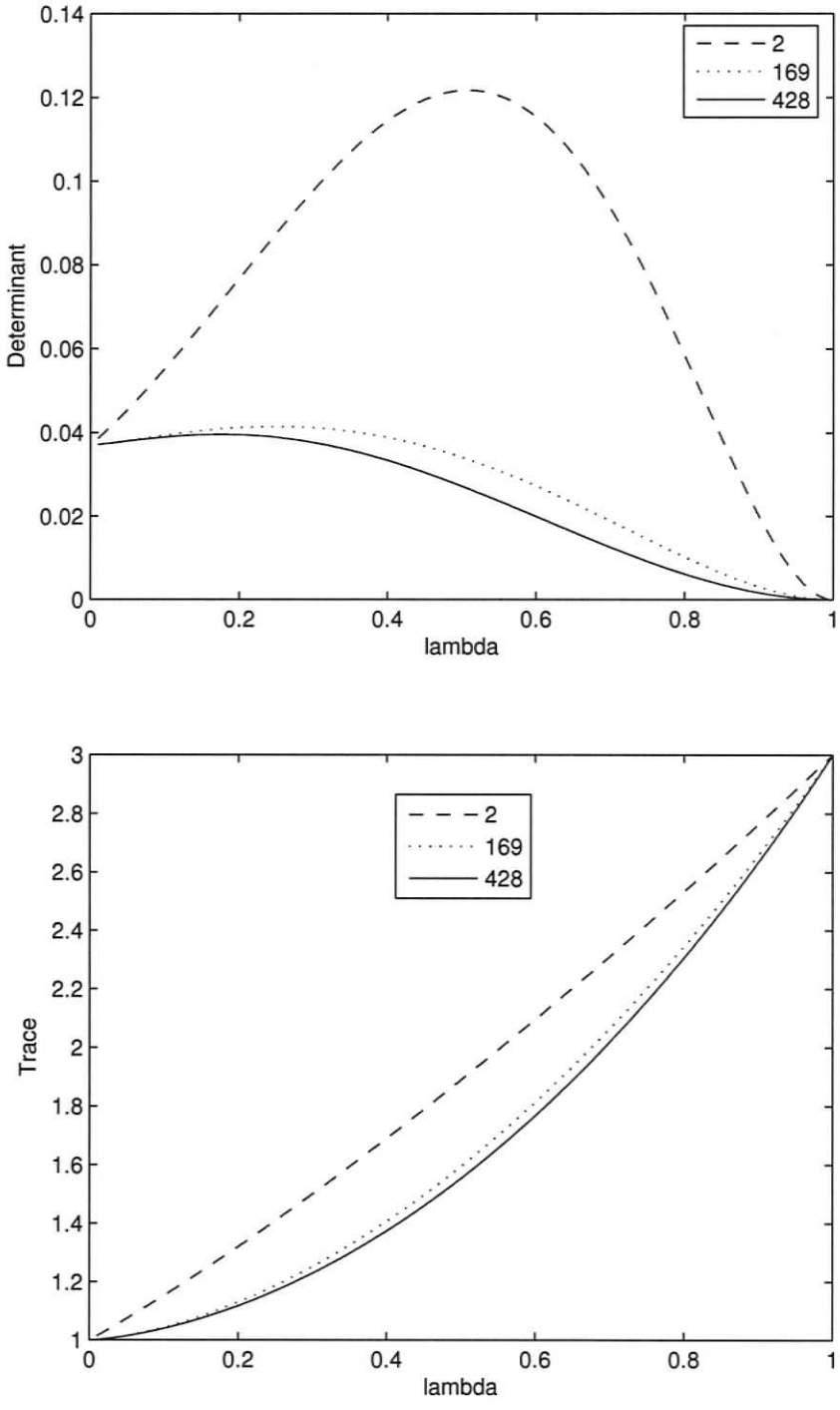


Figure 2.21: Loss functions of some 3x3 designs for DE

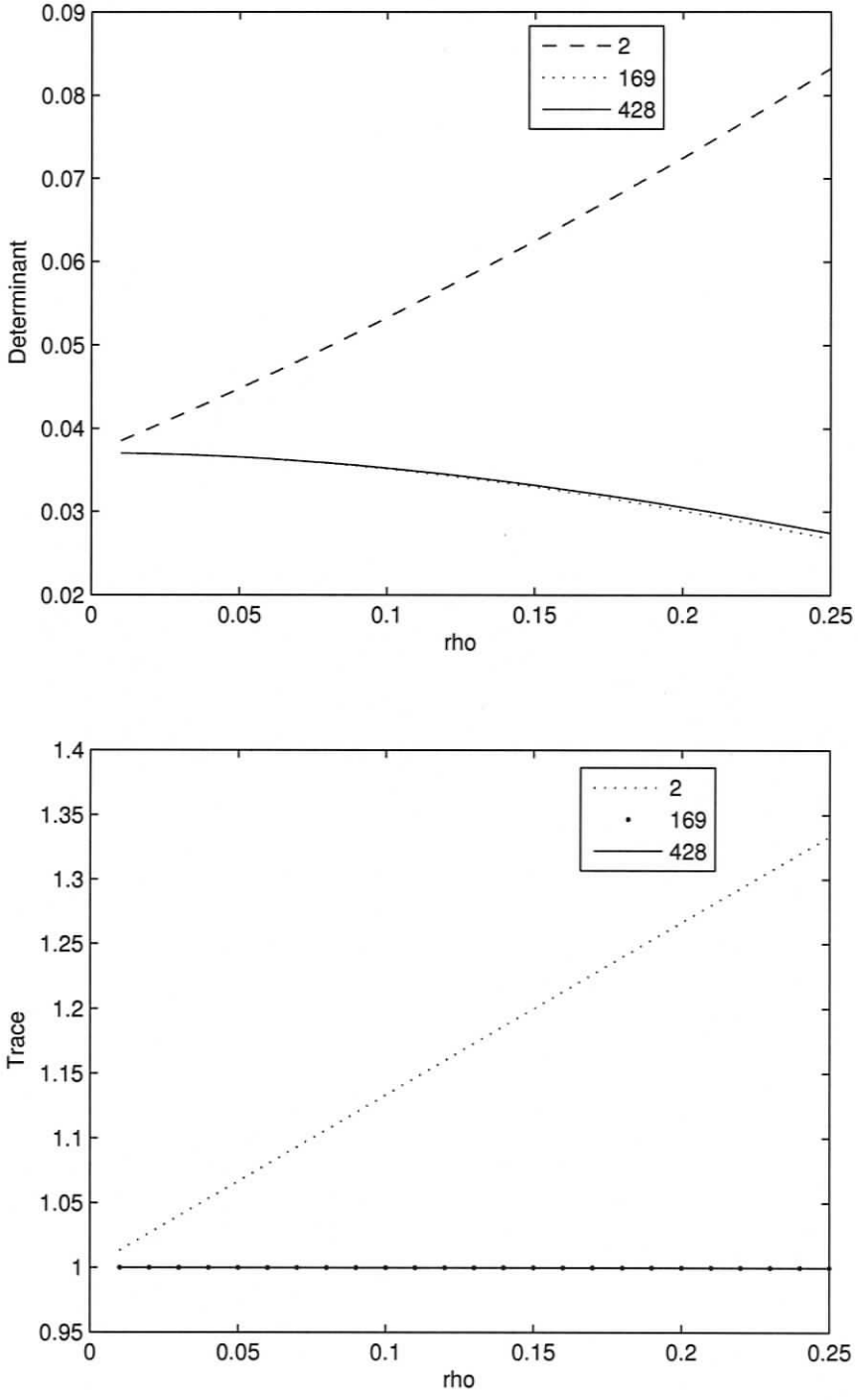


Figure 2.22: Loss functions of some 3x3 designs for NN(1)

2.2.5 Designs for 4x4 plots

Becher (1988) careful examined on 14 designs for 4x4 plots (Figure 2.23). Designs 1) to 12) are Latin Square designs. Design 13) was chosen because the minimum average distance between plots with the same treatment is maximized. Design 14) gives maximal average distance for two treatments. Only square plots are considered.

Figures 2.24, 2.25, 2.26, and 2.27 show the optimal and pessimal loss functions of these 4x4 designs for different correlation structures. For MA(1), Design 14) is D-optimal and Design 8 is A-optimal. For NN(1), Design 14) is D-optimal and Designs 1) to 14) are A-optimal. For DG, Design 8) is both D-optimal and A-optimal, and Design 1) is also D-optimal. For DE, Design 13) is D-optimal when $\lambda \leq 0.27$, otherwise, Design 8) is D-optimal; Design 13) is A-optimal when $\lambda \leq 0.22$, otherwise, Design 8) is A-optimal. These examples show that Latin Squares are not always optimal.

1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
2	1	4	3	3	1	4	2	2	1	4	3	4	1	2	3	2	1	4	3
3	4	1	2	2	4	1	3	4	3	1	2	3	4	1	2	4	3	2	1
4	3	2	1	4	3	2	1	3	4	2	1	2	3	4	1	3	4	1	2
	1)				2)				3)				4)				5)		
1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
3	1	4	2	2	4	1	3	3	4	1	2	3	4	1	2	4	3	2	1
4	3	2	1	3	1	4	2	2	1	4	3	4	1	2	3	2	1	4	3
2	4	1	3	4	3	2	1	4	3	2	1	2	3	4	1	3	4	1	2
	6)				7)				8)				9)				10)		
1	2	3	4	1	2	3	4	1	2	3	4	1	2	1	2				
4	3	1	2	4	3	2	1	3	4	1	2	3	4	3	4				
3	4	2	1	3	4	1	2	1	2	3	4	4	3	4	3				
2	1	4	3	2	1	4	3	3	4	1	2	2	1	2	1				
	11)				12)				13)				14)						

Figure 2.23: Some 4x4 designs

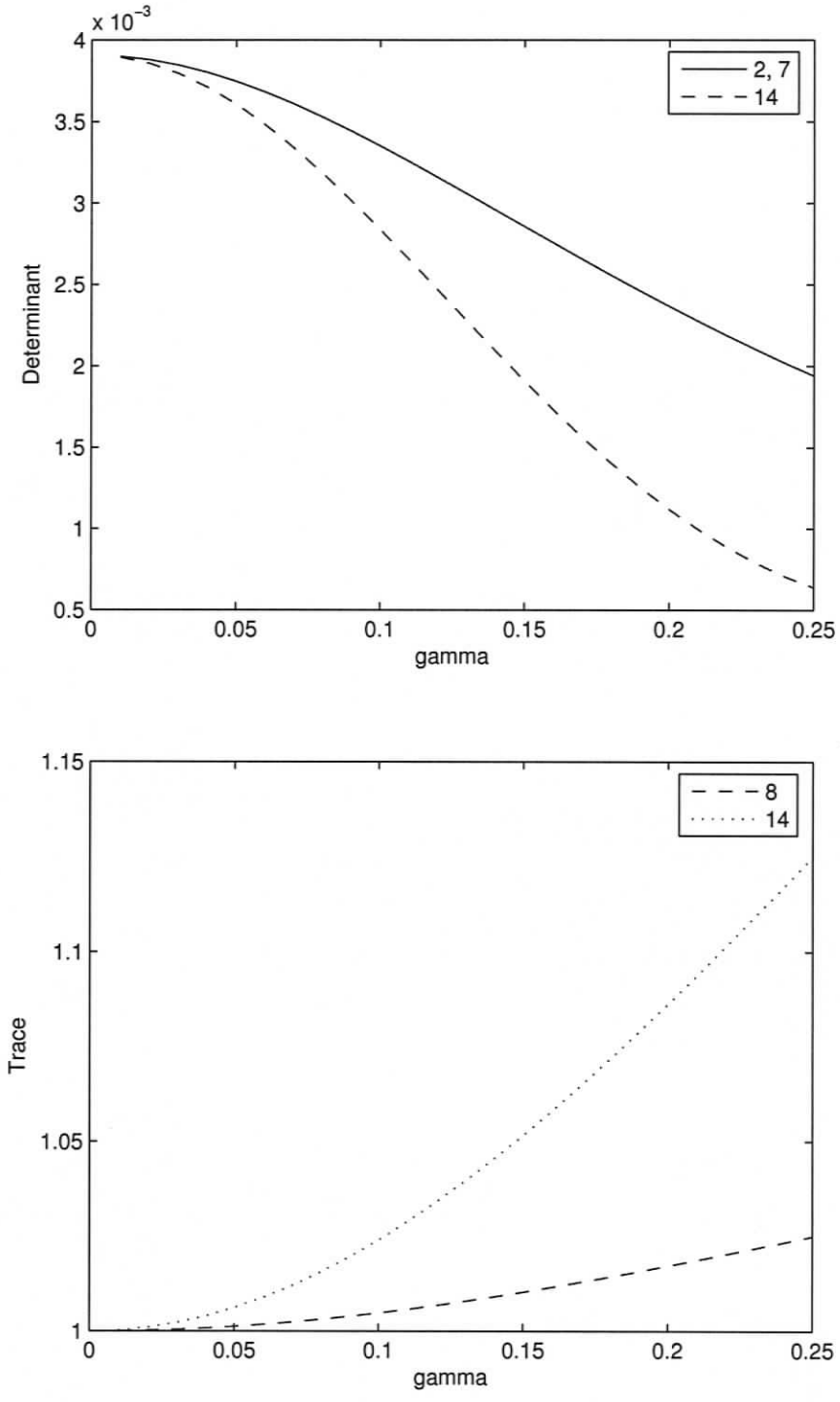


Figure 2.24: Loss functions of some 4x4 designs for MA(1)

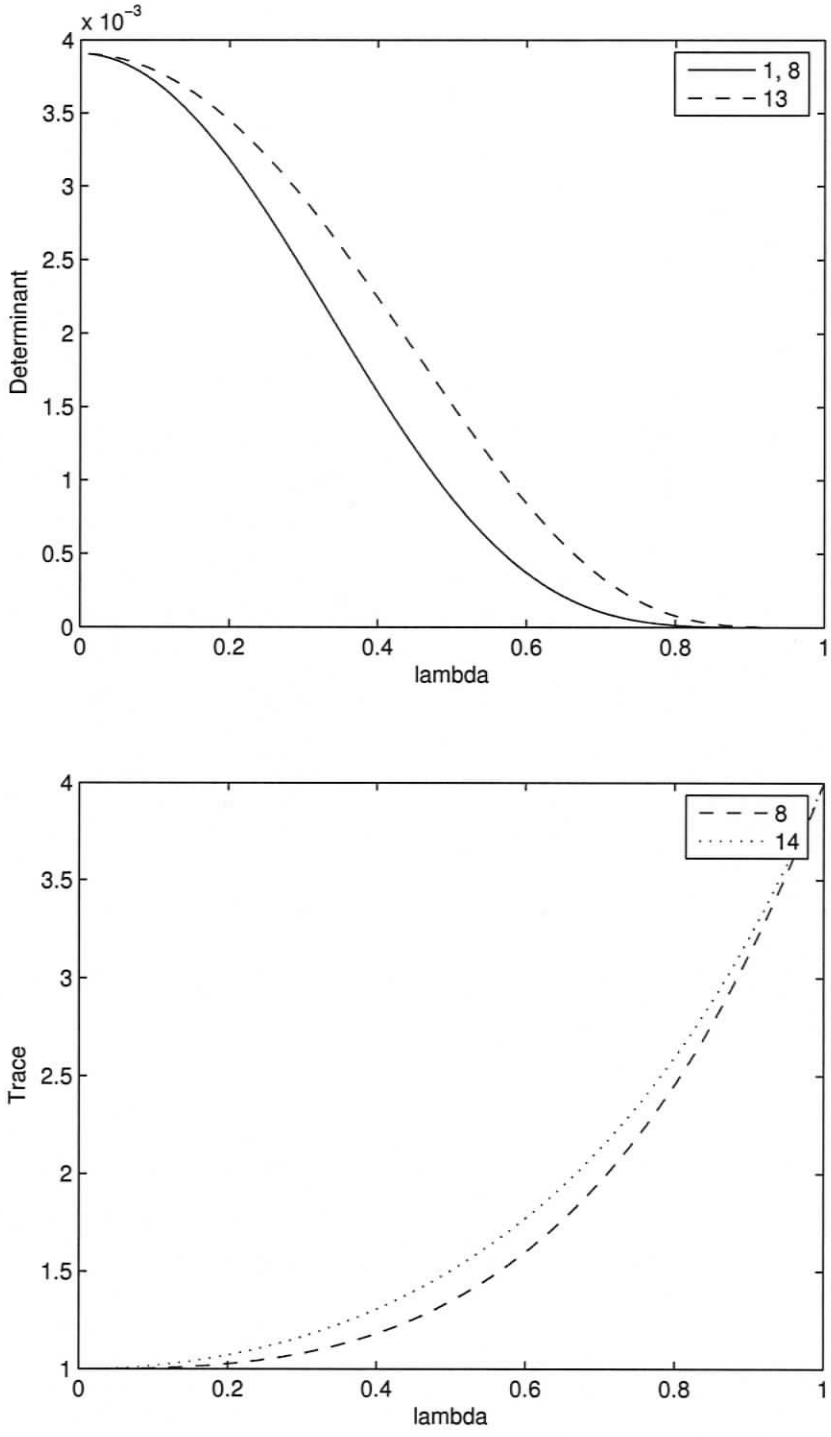


Figure 2.25: Loss functions of some 4x4 designs for DG

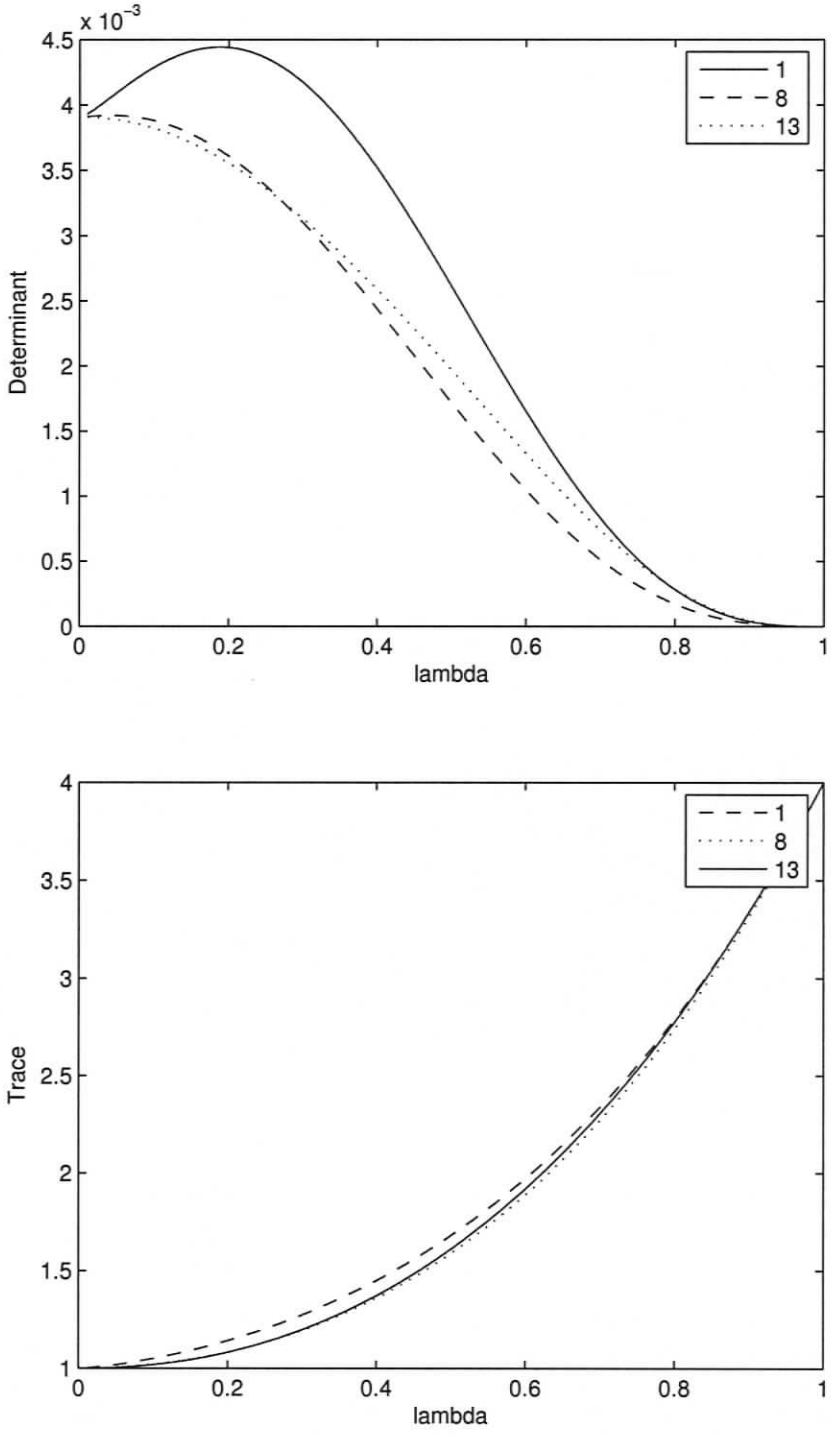


Figure 2.26: Loss functions of some 4x4 designs for DE

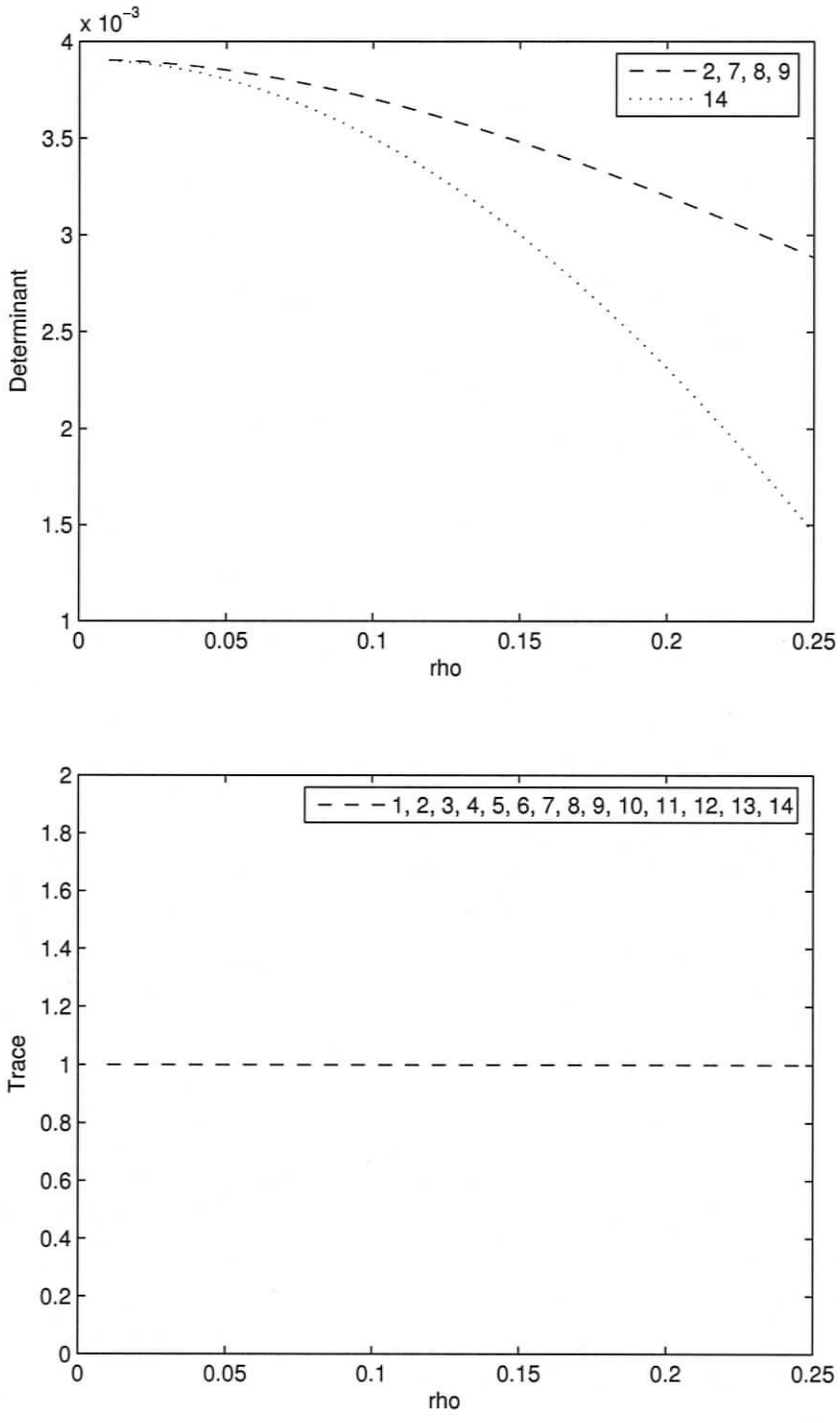


Figure 2.27: Loss functions of some 4x4 designs for NN(1)

Chapter 3

Minimax Designs

All optimal designs shown in the previous chapter are for specified correlation structures. However, often in practice, the correlation structure is either unknown or estimated with uncertainty. Thus, we need to find optimal designs which are robust in a sense that they are efficient for a neighbourhood of correlations. Let $\mathbf{R_0}$ be a given covariance matrix of the errors in model (1.1), which equals the correlation matrix multiplied by σ^2. In Section 3.1 we define a neighbourhood of $\mathbf{R_0}$, and give specific examples of the neighbourhoods of the correlation structures mentioned in Section 2.1. In Section 3.2, we propose a criterion for robust designs, and in Section 3.3, we show some results for robust designs.

3.1 Neighbourhood of Covariance Matrix

A neighbourhood of covariance matrix $\mathbf{R_0}$ is defined by

$$\boldsymbol{R}_{\beta,\mathbf{K}} = \{\mathbf{R}|0 \leq \mathbf{R} \leq \mathbf{R_0} + \beta\mathbf{K}\}, \beta > 0, \mathbf{K} \geq \mathbf{0}, \tag{3.1}$$

where the matrix order is according to semi-positive definiteness, i.e., $\mathbf{R_1} \leq \mathbf{R_2}$ means $\mathbf{R_2} - \mathbf{R_1}$ is semi-positive definite. Matrix $\mathbf{R_0}$ is the ideal covariance matrix under a specified spatial process, $\beta > 0$ is a constant and $\mathbf{K} \geq \mathbf{0}$ can be either $\mathbf{I}_N$ or $\mathbf{R_0}$ in our investigation. Some useful properties of the neighbourhood are :

- It is obvious that $\mathbf{R_0} \in \boldsymbol{R}_{\beta,\mathbf{K}}$.
- Constant β controls the size of the neighbourhood. Smaller β leads to a smaller neighbourhood.

Neighbourhoods for the correlation structures mentioned in Section 2.1 are investigated, and numeric results are shown in examples 3.1 - 3.5, for plot size $m = 5$ and $n = 6$.

Example 3.1: Suppose $\mathbf{R_0}$ is the covariance matrix for NN(1) with correlation parameter $\rho_0 = 0.1$ and $\sigma^2 = 1$. A neighbourhood of $\mathbf{R_0}$, $\boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{R_0}}$, with $\beta = 0.07$ contains all covariance matrices $\mathbf{R}$ from NN(1) with $\rho \in [0.09, 0.12]$. When $\beta = 1.00$, all covariance matrices $\mathbf{R}$ from NN(1) with $\rho \in [0.01, 0.24]$ are in the neighbourhood $\boldsymbol{R}_{\beta=1.00,\mathbf{K}=\mathbf{R_0}}$.

Example 3.2: Suppose $\mathbf{R_0}$ is the covariance matrix for MA(1) with correlation parameter $\rho_0 = 0.192$ and $\sigma^2 = 1$. A neighbourhood of $\mathbf{R_0}$, $\boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{R_0}}$, with $\beta = 0.50$ contains all covariance matrices $\mathbf{R}$ from MA(1) with $\rho \in [0.137, 0.357]$. When $\beta = 1.00$, all covariance matrices $\mathbf{R}$ with $\rho \in [0.06, 0.481]$ are in the neighbourhood $\boldsymbol{R}_{\beta=1.00,\mathbf{K}=\mathbf{R_0}}$.

Example 3.3: Suppose $\mathbf{R_0}$ is the covariance matrix for DG with correlation parameter $\lambda_0 = 0.10$ and $\sigma^2 = 1$. A neighbourhood of $\mathbf{R_0}$, $\boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{I_N}}$, with $\beta = 0.50$ contains all covariance matrices $\mathbf{R}$ from DG with $\lambda \in [0.01, 0.18]$. When $\beta = 1.00$, all covariance matrices $\mathbf{R}$ from DG with $\lambda \in [0.01, 0.25]$ are in the neighbourhood $\boldsymbol{R}_{\beta=1.00,\mathbf{K}=\mathbf{I_N}}$.

Example 3.4: Suppose $\mathbf{R_0}$ is the covariance matrix for DE with correlation parameter $\lambda_0 = 0.10$ and $\sigma^2 = 1$. A neighbourhood of $\mathbf{R_0}$, $\boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{I_N}}$, with $\beta = 0.50$ contains all covariance matrices $\mathbf{R}$ from DE with $\lambda \in [0.01, 0.16]$. When $\beta = 1.00$, all covariance matrices $\mathbf{R}$ from DE with $\lambda \in [0.01, 0.22]$ are in the neighbourhood $\boldsymbol{R}_{\beta=1.00,\mathbf{K}=\mathbf{I_N}}$.

Example 3.5: Suppose $\mathbf{R_0}$ is the covariance matrix for NN(1) with correlation parameter $\rho_0 = 0.10$ and $\sigma^2 = 1$. A neighbourhood of $\mathbf{R_0}$, $\boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{R_0}}$, with $\beta = 0.70$ contains all covariance matrices $\mathbf{R}$ from MA(1) with $\rho \in [0.02, 0.29]$. When $\beta = 1.00$, all covariance matrices $\mathbf{R}$ from MA(1) with $\rho \in [0.02, 0.36]$ are in the neighbourhood $\boldsymbol{R}_{\beta=1.00,\mathbf{K}=\mathbf{R_0}}$.

In addition, analytic results for the neighbourhood of $\mathbf{R_0}$ from NN(1) are found.

Theorem 3.1. *The neighbourhood of* $\mathbf{R_0}$ *from NN(1) with correlation parameter* ρ_0 *and variance* σ^2, $\boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{R_0}}$, *contains all covariance matrices* $\mathbf{R}$ *from NN(1) with variance* σ^2 *and correlation parameter* ρ *satisfying*

$$\max\left\{0, (1+\beta)\rho_0 - \frac{1}{4}\beta\right\} \leq \ \rho \ \leq \min\left\{(1+\beta)\rho_0 + \frac{1}{4}\beta, \frac{1}{4}\right\}.$$

Proof. For NN(1), the correlation parameter $\rho_{10} = \rho_{01} = \rho_{-1,0} = \rho_{0,-1} = \rho$, $\rho_{gh} = 0$ for $|g| + |h| > 1$, and $|\rho| < \frac{1}{4}$. For any $\mathbf{R}$ from NN(1) with correlation parameter ρ, $\mathbf{R} \in \boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{R_0}}$ means

$$\mathbf{R} \ \leq \ (1+\beta)\mathbf{R_0}.$$

Thus we want to show $(1+\beta)\mathbf{R_0} - \mathbf{R} \geq \mathbf{0}$ for $(1+\beta)\rho_0 - \frac{1}{4}\beta \leq \rho \leq (1+\beta)\rho_0 + \frac{1}{4}\beta$. Define $\tilde{\mathbf{R}} = ((\mathbf{1}+\beta)\mathbf{R_0} - \mathbf{R})/\beta$, then

$$\tilde{\mathbf{R}} = \sigma^{\mathbf{2}} \begin{pmatrix} 1 & \tilde{\rho} & 0 & \cdots & \tilde{\rho} & 0 & 0 & \cdots \\ \tilde{\rho} & 1 & \tilde{\rho} & \cdots & 0 & \tilde{\rho} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \cdots & 0 & \tilde{\rho} & 0 & \cdots & \tilde{\rho} & 1 & \tilde{\rho} \\ \cdots & 0 & 0 & \tilde{\rho} & \cdots & 0 & \tilde{\rho} & 1 \end{pmatrix}, \tag{3.2}$$

where $\tilde{\rho} = ((1+\beta)\rho_0 - \rho)/\beta$. If $(1+\beta)\rho_0 - \frac{1}{4}\beta \leq \rho \leq (1+\beta)\rho_0 + \frac{1}{4}\beta$, then $|\tilde{\rho}| < \frac{1}{4}$. $\tilde{\mathbf{R}}$ is a valid covariance matrix for the first order NN if $|\tilde{\rho}| < \frac{1}{4}$. Therefore, $\tilde{\mathbf{R}} \geq \mathbf{0}$, i.e., $(1+\beta)\mathbf{R_0} - \mathbf{R} \geq \mathbf{0}$. This completes the proof. □

Theorem 3.2. *The neighbourhood of* $\mathbf{R_0}$ *from NN(1) with correlation parameter* ρ_0 *and variance* σ^2, $\boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{I_N}}$, *contains all covariance matrices* $\mathbf{R}$ *from NN(1) with variance* σ^2 *and correlation parameter* ρ *satisfying*

$$\max\left\{0, \rho_0 - \frac{\beta}{4\sigma^2}\right\} \leq \ \rho \ \leq \min\left\{\rho_0 + \frac{\beta}{4\sigma^2}, \frac{1}{4}\right\}.$$

Proof. For any $\mathbf{R}$ from NN(1) with parameter ρ, $\mathbf{R} \in \boldsymbol{R}_{\beta,\mathbf{K}=\mathbf{I_N}}$ means

$$\mathbf{R} \ \leq \ \mathbf{R_0} + \beta\mathbf{I_N}.$$

Thus we want to show $\mathbf{R_0} + \beta\mathbf{I_N} - \mathbf{R} \geq \mathbf{0}$ for $\rho_0 - \frac{\beta}{4\sigma^2} \leq \rho \leq \rho_0 + \frac{\beta}{4\sigma^2}$. Define $\tilde{\mathbf{R}} = (\mathbf{R_0} + \beta\mathbf{I_N} - \mathbf{R})/\beta$, then $\tilde{\mathbf{R}}$ has the same structure as it is in equation (3.2) with $\tilde{\rho} = \sigma^2(\rho_0 - \rho)/\beta$. If $\rho_0 - \frac{\beta}{4\sigma^2} \leq \rho \leq \rho_0 + \frac{\beta}{4\sigma^2}$, then $|\tilde{\rho}| < \frac{1}{4}$, and $\tilde{\mathbf{R}}$ is a valid covariance matrix for the first order NN. Therefore, $\tilde{\mathbf{R}} \geq \mathbf{0}$, i.e., $\mathbf{R_0} + \beta\mathbf{I_N} - \mathbf{R} \geq \mathbf{0}$. This completes the proof. □

3.2 Minimax Design Criterion

Because the correlation structure is usually unknown or estimated with uncertainty, either the assumed structure is correct but the parameter value is not, or the assumed structure is incorrect and the parameter value is correct or not. Thus, we need to find optimal designs which are robust in a sense that they are efficient for a neighbourhood of $\mathbf{R_0}$. In Section 1.4, we minimize some scalar functions of the covariance matrix,

$\mathcal{L}(\text{cov}(C\hat{\theta}))$, to find optimal designs. Here, we propose minimax designs which are solutions to the following minimax problem

$$\min_{\mathbf{X}} \max_{\mathbf{R}\in R_{\beta,\mathbf{K}}} \mathcal{L}(cov(\mathbf{C}\hat{\boldsymbol{\theta}})), \tag{3.3}$$

where $\mathbf{R} = \text{cov}(\epsilon) = \sigma^2\mathbf{V}$, $\hat{\boldsymbol{\theta}}$ is an estimator of $\boldsymbol{\theta}$ and matrix $\mathbf{C}$ is constant, such as a contrast matrix in equation (1.5). Minimax designs are robust against misspecification of spatial correlation structure.

If the covariance matrix of the errors is known, the treatment means $\boldsymbol{\theta}$ can be estimated by $\hat{\boldsymbol{\theta}}_{\boldsymbol{G}}$ in equation (1.2), and $\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{G}}) = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$. If the covariance matrix $\mathbf{R}$ of the errors is unknown, the treatment means are usually estimated by $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ in equation (1.3), and $\text{cov}(\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$. If the covariance matrix $\mathbf{R}$ is unknown, but belongs to a neighbourhood of $\mathbf{R_0}$, we propose a generalized least squares estimator $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ as $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}} = (\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{y}$, and $\text{cov}(\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = (\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{R}\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}$. Robust designs based on $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ will be studied and compared in this thesis.

3.3 Some Results Based on Minimax Design Criterion

In this section, we show some analytic results based on the minimax design criterion. In Theorem 3.3 we show that for LSE the maximum loss over a neighbourhood of

$\mathbf{R_0}$, $\boldsymbol{R}_{\beta,\mathbf{K}}$, occurs when $\mathbf{R} = \mathbf{R_0} + \beta\mathbf{K}$.

Theorem 3.3. *For $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$, the maximum loss over a neighbourhood of $\mathbf{R_0}$ is*

$$\mathcal{L}(\mathbf{X}, \mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = \max_{\mathbf{R}\in\boldsymbol{R}_{\beta,\mathbf{K}}} \text{trace}(\text{cov}(\mathbf{C}\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) \;=\; trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})|\mathbf{R} = \mathbf{R_0} + \beta\boldsymbol{K}),$$

where $\beta > 0$, and $\mathbf{K} \geq 0$.

Proof. For any $\mathbf{R} \in \boldsymbol{R}_{\beta,\mathbf{K}}$,

$$\begin{aligned}
trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\} \\
&= trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{R_0} + \beta\mathbf{K} + \mathbf{R} - \mathbf{R_0} - \beta\mathbf{K})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\} \\
&= trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{R_0} + \beta\mathbf{K})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\} \\
&\quad + trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{R} - \mathbf{R_0} - \beta\mathbf{K})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\} \\
&\leq trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{R_0} + \beta\mathbf{K})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\}, \\
&\quad since\ \mathbf{R} - \mathbf{R_0} - \beta\mathbf{K} \leq \mathbf{0}.
\end{aligned}$$

Thus $\max_{\mathbf{R}\in\boldsymbol{R}_{\beta,\mathbf{K}}} trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}))$=$trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})|\mathbf{R} = \mathbf{R_0} + \beta\mathbf{K})$. □

Theorem 3.4 shows that estimator $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ has a similar property to that of $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

Theorem 3.4. *For $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$, the maximum loss over a neighbourhood of $\mathbf{R_0}$ is*

$$\mathcal{L}(\mathbf{X}, \mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = \max_{\mathbf{R}\in\boldsymbol{R}_{\beta,\mathbf{K}}} \text{trace}(\text{cov}(\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})) \;=\; trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})|\mathbf{R} = \mathbf{R_0} + \beta\boldsymbol{K}),$$

where $\beta > 0$, and $\mathbf{K} \geq 0$.

Proof. For any $\mathbf{R} \in \boldsymbol{R}_{\beta,\mathbf{K}}$,

$$\begin{aligned}
trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})) &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{R}\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\} \\
&= trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{R}_0^{-1}\mathbf{X}'(\mathbf{R_0} + \beta\mathbf{K} \\
&\quad +\mathbf{R} - \mathbf{R_0} - \beta\mathbf{K})\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\} \\
&= trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}(\mathbf{R_0} + \beta\mathbf{K})\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\} \\
&\quad +trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}(\mathbf{R} - \mathbf{R_0} - \beta\mathbf{K})\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\} \\
&\leq trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}(\mathbf{R_0} + \beta\mathbf{K})\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\}, \\
&\quad since\ \mathbf{R} - \mathbf{R_0} - \beta\mathbf{K} \leq \mathbf{0}.
\end{aligned}$$

Thus $\max_{\mathbf{R}\in\boldsymbol{R}_{\beta,\mathbf{K}}}$ trace (cov($\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$)) = trace (cov($\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$) $|\mathbf{R} = \mathbf{R_0} + \beta\mathbf{K}$). □

Furthermore, we can simplify $\mathcal{L}$ ($\mathbf{X}$,$\mathbf{C}\hat{\boldsymbol{\theta}}$) for $\mathbf{K} = \mathbf{I}_N$, and $\mathbf{R_0}$. When $\mathbf{K} = \mathbf{R_0}$,

$$\begin{aligned}
\mathcal{L}(\mathbf{X}, \mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{R_0} + \beta\mathbf{R_0})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\} \\
&= (1+\beta)trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\},
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{L}(\mathbf{X}, \mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}(\mathbf{R_0} + \beta\mathbf{R_0})\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\} \\
&= (1+\beta)trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{R_0}\mathbf{R}_0^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\} \\
&= (1+\beta)trace\{\mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'\}.
\end{aligned}$$

In other words, when $\mathbf{K} = \mathbf{R_0}$, optimal designs with specified covariance matrix $\mathbf{R_0}$

are robust designs under our design criterion, since

$$\max_{\mathbf{R}\in R_{\beta,\mathbf{K}=\mathbf{R_0}}} trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}})) = (1+\beta)trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}})|\mathbf{R_0}). \tag{3.4}$$

When $\mathbf{K}=\mathbf{I_N}$,

$$\begin{aligned}\mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{R_0}+\beta\mathbf{I_N})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\}\\ &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\}\\ &\quad +\beta trace\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\},\end{aligned}$$

and

$$\begin{aligned}\mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0^{-1}}(\mathbf{R_0}+\beta\mathbf{I_N})\mathbf{R_0^{-1}}\mathbf{X}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{C}'\}\\ &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{R_0}\mathbf{R_0^{-1}}\mathbf{X}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{C}'\}\\ &\quad +\beta trace\{\mathbf{C}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{R_0^{-1}}\mathbf{X}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{C}'\}\\ &= trace\{\mathbf{C}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{C}'\}\\ &\quad +\beta trace\{\mathbf{C}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0^{-2}}\mathbf{X}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{C}'\}.\end{aligned}$$

Therefore, when $\mathbf{K}=\mathbf{I_N}$,

$$\begin{aligned}\max_{\mathbf{R}\in R_{\beta,\mathbf{K}=\mathbf{I_N}}} trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})) &= trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})|\mathbf{R_0}) + \beta\mathbf{trace}\{\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\},\\ \max_{\mathbf{R}\in R_{\beta,\mathbf{K}=\mathbf{I_N}}} trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})) &= trace(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})|\mathbf{R_0})\\ &\quad +\beta trace\{\mathbf{C}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0^{-2}}\mathbf{X}(\mathbf{X}'\mathbf{R_0^{-1}}\mathbf{X})^{-1}\mathbf{C}'\}.\end{aligned}$$

For mxn designs with t treatments, $\mathbf{X'X} = \mathrm{r}\mathbf{I_t}$ is a constant matrix. Thus, when $\mathbf{K} = \mathbf{I_N}$, optimal designs with specified covariance matrix $\mathbf{R_0}$ are also robust designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ as well.

In the next chapter, we present a numerical algorithm for finding classical optimal designs and minimax designs, and show some examples of minimax designs for estimating $\boldsymbol{\theta}$ by using $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$.

Chapter 4

Numerical Algorithm for Finding Minimax Designs

In Chapter 3, we have found the maximum trace of $\text{cov}(\mathbf{C}\hat{\boldsymbol{\theta}})$. In this chapter, we investigate designs that minimize the maximum trace of $\text{cov}(\mathbf{C}\hat{\boldsymbol{\theta}})$. For small number of design plots, complete enumeration and comparison may be possible. As the dimensions of the plots increase, the number of possible designs increases dramatically. In Section 4.1, we introduce a simulated annealing algorithm that helps us find the minimax designs. In Section 4.2, we present examples of minimax designs of the least squares and generalized least squares estimators. These designs are then compared.

4.1 Simulated Annealing Algorithm

Algorithms based on simulated annealing types of routines have been used to find optimal designs, e.g., robust run-order designs (Zhou, 2001) and designs of factorial experiments (Elliott, Eccleston, and Martin, 1999). In this section, we describe a simulated annealing algorithm for finding minimax designs. Our algorithm consists of two phases: construction of an initial design and the set up of search parameters; and a simulated annealing routine. This algorithm is based on an interchange procedure, and does not guarantee finding the global optimum, but only a local optimum. So to mitigate obtaining a local optimum, we run a number of trials with different starting designs.

In the algorithm we have developed, a design is specified by the design matrix, $\mathbf{X}$. The algorithm begins with an initial design which is a random allocation of the different treatments. The objective function, f, is defined as the maximum value of $(\text{cov}(\mathbf{C}\hat{\boldsymbol{\theta}}))$, i.e., $f=\mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}})$. A new feasible design is obtained by an interchange between two allocations of two different treatments. The change in the objective function, say $\triangle f$, is the change in $\mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}})$. We accept the new design with the following probability p:

$$p = \begin{cases} 1, & \triangle f \leq 0 \\ exp(-\triangle f/T), & \triangle f > 0, \end{cases}$$

where T corresponds to a temperature. Designs with greater optimality value can

be accepted with probability $exp(-\triangle f/T)$. This allows the algorithm to move away from local optimum.

Initially, the temperature is set high so there is a high probability that designs with greater objective values are accepted. It is then slowly decreased or cooled down, after a specified number of interchanges, say 1,000, have been performed at each temperature. This process is repeated until the probability of accepting a 'worse' design is small, the temperature, T, is below the stopping temperature, say 10^{-6}.

Our algorithm uses geometric cooling (Elliott, Eccleston, and Martin, 1999), which simply reduces the temperature at a given time, T_i, by ξ, so that $T_{i+1} = \xi T_i$. The cooling factor, ξ, is between 0 and 1.

At the end of the annealing routine the algorithm carries out a steepest descent routine (Elliott, Eccleston, and Martin, 1999). This routine performs all possible pair-wise interchanges of the existing best design, calculates f after each interchange. At the end of this routine, if the best design found previously still has the smallest f, then we say this is a locally optimum design. Otherwise, we carry out the steepest descent routine again on the new best design, until the best designs before and after the routine are consistent. For smaller designs this step is not necessary since the annealing routine generally converges to the optimal design, but for larger designs it is a way to ensure that at least a locally optimum design is returned.

The algorithm can be briefly summarized in the following steps:

1. Select an initial design, $\mathbf{X}$. Obtain an initial temperature, T_0. Let $\mathbf{X}^* = \mathbf{X}$, where $\mathbf{X}^*$ denotes the best current design.

2. Initialize the iteration number $j = 1$.

3. Generate a new design by interchanging allocations of two different treatments, denote it by $\mathbf{X}^{(1)}$, and calculate $\triangle f$.

4. Let $\mathbf{X} = \mathbf{X}^{(1)}$ with probability as given by the probability rule stated above.

5. If $\mathbf{X}$ is better than $\mathbf{X}^*$, i.e., $\mathbf{X}$ has a smaller $\mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}})$, then let $\mathbf{X}^* = \mathbf{X}$.

6. Increment j by 1 and return to step 3 until the specified number of iterations has elapsed.

7. Cool the temperature: $T_{i+1} = \xi T_i$. If the temperature has reached the stopping temperature, then stop; otherwise, repeat steps 2-6.

8. Execute the steepest descent routine.

4.2 Minimax Designs

Using the algorithm described in Section 4.1, we hope to find minimax designs that satisfy our robust design criterion. As mentioned in Section 3.3, when $\mathbf{K} = \mathbf{R_0}$, optimal designs with specified covariance matrix $\mathbf{R_0}$ are also robust minimax designs

2	5	6	5	1	3		6	2	3	1	6	1
5	2	4	2	6	2		2	6	2	4	1	3
4	1	2	6	2	6		1	3	6	3	2	1
1	3	1	3	6	3		4	6	4	5	3	5
4	1	3	4	3	5		3	4	5	2	5	1
1	5	4	5	6	4		4	2	4	5	6	5

(1) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = 1.4$ (2) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 1.3117$

Figure 4.1: Minimax designs with size 6x6 and $t = 6$ under NN(1), (1) for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

for both $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ and $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$. When $\mathbf{K} = \mathbf{I_N}$, optimal designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ are robust as well. However, when $\mathbf{K} = \mathbf{I_N}$, minimax designs for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ may be different from the classical optimal designs. In this section, we show examples of classical optimal designs and minimax designs for various covariance matrices $\mathbf{R_0}$. The designs reported in the following examples were the best obtained from 20 runs of the algorithm. In these examples, we set $\sigma^2 = 1$. Square plots are assumed in the following examples. In Examples 4.1 - 4.6, $\mathbf{K} = \mathbf{I_N}$ and $\mathbf{C} = \mathbf{I_t}$; in example 4.7, $\mathbf{K} = \mathbf{R_0}$ and $\mathbf{C} = \mathbf{I_t}$; in example 4.8 - 4.10, $\mathbf{K} = \mathbf{I_N}$ but $\mathbf{C} \neq \mathbf{I_t}$. Minimax designs are not unique in these examples.

Example 4.1: For designs with size 6x6 and $t = 6$ treatments, $\mathbf{R_0}$ is the covari-

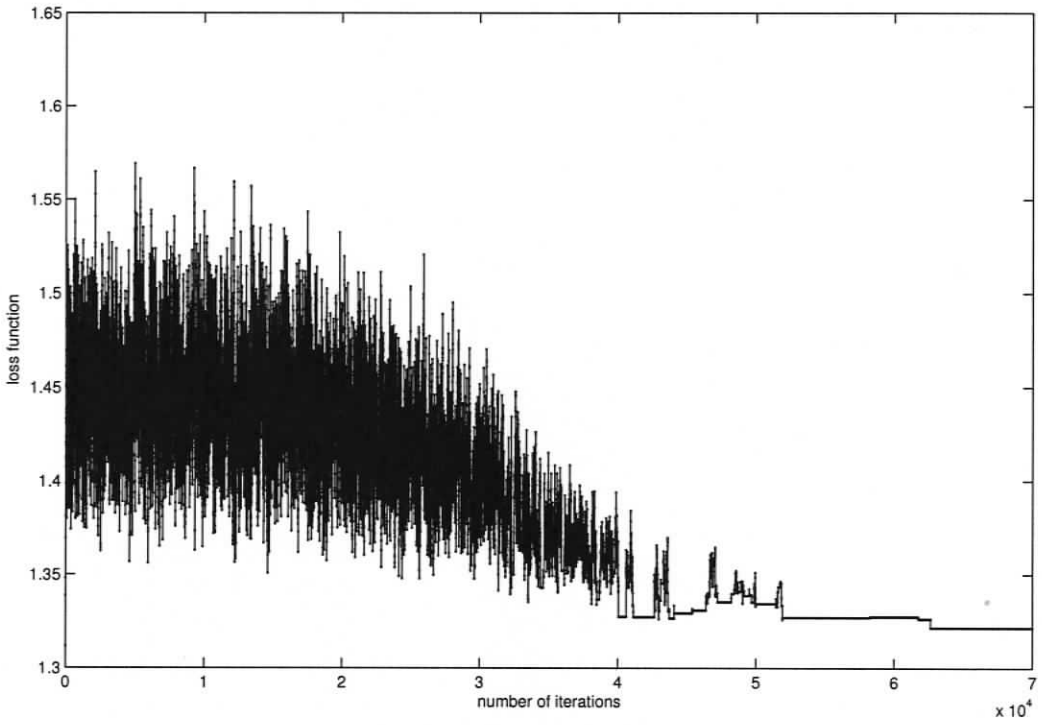


Figure 4.2: Loss function $\mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ Vs. number of iterations in Example 4.1

ance matrix for NN(1) with correlation parameter $\rho = 0.2$. When $\beta = 0.4$, minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are shown in Figure 4.1.

In the computation, initial value for T is 0.3 and it is reduced progressively by a factor of 0.9 after each 1000 iterations. Figure 4.2 plots the loss function $f = \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$ versus the number of iterations. We can see from Figure 4.2 that the loss function f converges after $6.3\text{x}10^4$ runs.

As noted before, minimax design for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ is the optimal design for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ with

specified covariance structure $\mathbf{R_0}$. The minimax designs for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ in Figure 4.1 (2) is optimal for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$, and yields the same $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ value as that of Figure 4.1 (1). Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ for the same minimax design, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

When covariance matrix $\mathbf{R_0}$ is from NN(1), minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ are the designs in which every plot receives a treatment different from its nearest neighbours.

Theorem 4.1. *For mxn plots with t treatments and r replicates, when the covariance matrix* $\mathbf{R_0}$ *is from NN(1), minimax designs for* $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ *are the designs in which every plot has a treatment different from its nearest neighbours.*

Proof. Note that $\mathbf{X'X} = r\mathbf{I}_t$ is a constant matrix. Let

$$\mathbf{X'} = \begin{pmatrix} \mathbf{x'_1} \\ \vdots \\ \mathbf{x'_t} \end{pmatrix}, \tag{4.1}$$

then

$$\begin{aligned} trace(\mathbf{X'R_0X}) &= \sum_{i=1}^{t} \mathbf{x'_i R_0 x_i} \\ &= \sigma^2 \sum_{i=1}^{t} \mathbf{x'_i x_i} + \mathbf{2} \sum_{\mathbf{i=1}}^{\mathbf{t}} \mathbf{x'_i (R_0 - \sigma^2 I) x_i} \\ &= \sigma^2 tr + 2\sigma^2 \rho s, \end{aligned}$$

where s=total number of times in which a plot receives the same treatment as one of

its nearest neighbours. Thus,

$$trace\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\} \geq \frac{\sigma^2 tr}{r^2} = \frac{\sigma^2 t}{r}. \tag{4.2}$$

When every plot has a treatment different from its nearest neighbours and if $\mathbf{K} = \mathbf{I_N}$, then

$$\begin{aligned} \mathcal{L}(\mathbf{X}, \hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) &= trace\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\} \\ &\quad +\beta trace\{(\mathbf{X}'\mathbf{X})^{-1}\} \\ &= \sigma^2\frac{t}{r} + \frac{\beta t}{r} = (\sigma^2+\beta)\frac{t}{r}; \end{aligned}$$

if $\mathbf{K} = \mathbf{R_0}$, then

$$\begin{aligned} \mathcal{L}(\mathbf{X}, \hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) &= (1+\beta)trace\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R_0}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\} \\ &= (1+\beta)\sigma^2\frac{t}{r}. \end{aligned}$$

□

For $t = r = m = n$, a Latin Square design is obviously an optimal minimax design for NN(1), because in a Latin Square design every one of t treatments appears exactly once in each row and in each column, and the same treatment does not occur in neighbouring plots.

Example 4.2: For designs with size 6x6 and $t = 6$ treatments, $\mathbf{R_0}$ is the covariance matrix for MA(1) with correlation parameter $\gamma = 0.2$, when $\beta = 0.4$, minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are shown in Figure 4.3. Minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

2	4	6	3	4	2		6	4	1	5	4	6
5	1	2	5	6	1		5	6	2	1	6	3
3	6	4	1	3	5		4	2	5	2	3	6
1	5	3	2	4	6		2	4	2	3	5	1
6	4	1	6	5	3		1	2	4	5	1	3
3	2	5	4	1	2		4	3	1	3	6	5

(1) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = 1.4$ (2) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 1.3284$

Figure 4.3: Minimax designs with size 6x6 and $t = 6$ under MA(1), (1) for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

are different. Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

When covariance matrix $\mathbf{R_0}$ is from MA(1), minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ are the designs in which every plot has a treatment different from its twelve neighbours shown in Figure 2.2.

Theorem 4.2. *For mxn plots with t treatments and r replicates, when the covariance matrix* $\mathbf{R_0}$ *is from MA(1), minimax designs for* $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ *are the designs in which every plot has a treatment different from its twelve neighbours shown in Figure 2.2.*

Proof. Note that $\mathbf{X}'\mathbf{X} = r\mathbf{I}_t$ is a constant matrix. Let $\mathbf{X}'$ has the same representation

as it is in equation (4.1), then

$$\begin{aligned} trace(\mathbf{X'R_0X}) &= \sum_{i=1}^{t} \mathbf{x'_i R_0 x_i} \\ &= \sigma^2 \sum_{i=1}^{t} \mathbf{x'_i x_i} + \mathbf{2} \sum_{\mathbf{i=1}}^{\mathbf{t}} \mathbf{x'_i (R_0 - \sigma^2 I) x_i} \\ &= \sigma^2 tr + 2\sigma^2 \rho s_1 + 2\sigma^2 \gamma \rho s_2 + \sigma^2 \gamma \rho s_3, \end{aligned}$$

where s_i = total number of times in which a plot receives the same treatment as one of its ith neighbours, $i = 1, 2$ and 3, referring to Figure 2.1. Thus,

$$trace\{(\mathbf{X'X})^{-1}\mathbf{X'R_0X}(\mathbf{X'X})^{-1}\} \geq \frac{\sigma^2 tr}{r^2} = \frac{\sigma^2 t}{r}. \tag{4.3}$$

When every plot has a treatment different from its twelve neighbours shown in Figure 2.2 and if $\mathbf{K} = \mathbf{I_N}$, then

$$\begin{aligned} \mathcal{L}(\mathbf{X}, \hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) &= trace\{(\mathbf{X'X})^{-1}\mathbf{X'R_0X}(\mathbf{X'X})^{-1}\} \\ &\quad + \beta trace\{(\mathbf{X'X})^{-1}\} \\ &= \sigma^2 \frac{t}{r} + \frac{\beta t}{r} = (\beta + \sigma^2)\frac{t}{r}; \end{aligned}$$

if $\mathbf{K} = \mathbf{R_0}$, then

$$\begin{aligned} \mathcal{L}(\mathbf{X}, \hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) &= (1 + \beta) trace\{(\mathbf{X'X})^{-1}\mathbf{X'R_0X}(\mathbf{X'X})^{-1}\} \\ &= (1 + \beta)\sigma^2 \frac{t}{r}. \end{aligned}$$

□

1	5	3	4	2	6		4	2	1	3	5	2
4	6	2	1	3	5		5	3	6	2	3	5
3	2	6	5	4	1		6	5	2	6	1	4
5	1	4	3	6	2		2	6	5	4	2	1
6	3	5	2	1	4		3	4	6	1	4	3
2	4	1	6	5	3		6	3	1	5	1	4

(1) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}) = 1.6967$ (2) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 1.4974$

Figure 4.4: Minimax designs with size 6x6 and $t = 6$ under DG, (1) for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$.

We can conclude that when the covariance matrix $\mathbf{R_0}$ is from MA(1), the values of $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}})$ is the same as when $\mathbf{R_0}$ is from NN(1). For $t = r = m = n$, a Latin Square design with Knight's move is obviously an optimal minimax design for MA(1) if such design exists. Mentioned in Section 1.1, the Knight's move Latin square is a special Latin square such that all plots with the same treatment can be traversed by a series of Knight's moves as on a chessboard. Figure 1.1 shows an example of 5x5 Knight's move Latin squares. Because of this special feature of minimax designs for MA(1), they are also minimax designs for NN(1).

Example 4.3: For designs with size 6x6 and $t = 6$ treatments, $\mathbf{R_0}$ is the covariance matrix for DG with correlation parameter $\lambda = 0.6$, when $\beta = 0.1$, minimax

2	3	5	6	2	3		6	3	1	2	4	3
6	1	4	1	5	4		4	2	5	3	1	5
4	5	2	3	6	1		5	3	4	6	2	6
3	6	1	4	5	2		1	6	1	5	4	1
5	2	3	2	6	3		4	5	2	6	3	2
1	4	6	5	1	4		2	6	3	1	4	5

(1) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = 2.0481$ (2) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 1.7891$

Figure 4.5: Minimax designs with size 6x6 and $t = 6$ under DE with $\beta = 0.1$, (1) for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are shown in Figure 4.4. Minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are different. Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

Example 4.4: For designs with size 6x6 and $t = 6$ treatments, $\mathbf{R_0}$ is the covariance matrix for DE with correlation parameter $\lambda = 0.6$, when $\beta = 0.1$, minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are shown in Figure 4.5. Minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are different. Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

Example 4.5: For designs with size 6x6 and $t = 6$ treatments, $\mathbf{R_0}$ is the covariance matrix for DE with correlation parameter $\lambda = 0.6$, some representations of

4	1	2	5	3	6		3	2	1	3	5	6
5	6	3	1	4	2		5	4	6	2	4	1
3	2	4	6	5	1		1	5	3	1	5	2
1	6	5	2	4	3		6	4	2	6	4	3
4	3	1	3	1	6		3	5	1	3	5	6
5	2	6	4	5	2		2	4	6	2	4	1

(1) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = 3.4481$ (2) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 3.5242$

Figure 4.6: Minimax designs with size 6x6 and $t = 6$ under DE with $\beta = 1.5$, (1) for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

β	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$
0.1	1.6967	1.4974
0.7	2.6481	2.5441
1.5	3.4481	3.5242
2.0	3.9481	4.1257

Table 4.1: Loss functions for 6x6 designs and $t = 6$ under DE with different β values

β	Neighbourhood for $\lambda = 0.6$ in DE	Neighbourhood for $\lambda = 0.6$ in DG
0.05	[0.58,0.60]	[0.59,0.60]
0.10	[0.55,0.60]	[0.57,0.60]
0.30	[0.40,0.60]	[0.46,0.61]
0.60	[0.09,0.61]	[0.20,0.61]
1.00	[0.01,0.62]	[0.01,0.63]
1.50	[0.01,0.64]	[0.01,0.64]
2.00	[0.01,0.65]	[0.01,0.66]

Table 4.2: Some examples of the neighbourhoods for λ for DG and DE.

minimax loss functions are shown in Table 4.1. Minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ for $\beta = 1.5$ are shown in Figure 4.6. Minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are different. From this example, we see that when β is small, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$; $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ when β is large. For this particular example, 1.5 is the smallest β value such that $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$. Table 4.2 shows some examples of the neighbourhoods of λ for DG and DE. As we can see, as β gets larger, the size of the neighbourhood gets larger as well.

Example 4.6: For designs with size 3x15 and $t = 3$ treatments, $\mathbf{R_0}$ is the covariance matrix from NN(1) with correlation parameter $\rho_0 = 0.12$, when $\beta = 0.1$, minimax design for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is shown in Figure 4.7. The same minimax design

3 1 3 1 3 1 2 1 2 1 2 3 2 3 2

1 3 1 3 1 3 1 2 1 2 3 2 3 2 3

2 1 3 1 3 1 2 1 2 1 2 3 2 3 2

Figure 4.7: Minimax design with size 3x15 and $t = 3$ under NN(1) when $\mathbf{K} = \mathbf{I_N}$, for both $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

is optimal for both $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ and $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$. For this design, $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ is 0.2200 and $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 0.2092$. Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ for the same minimax design, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

Example 4.7: Consider designs with size 3x15 and $t = 3$ treatments, and $\mathbf{R_0}$ is the covariance matrix from NN(1) with correlation parameter $\rho_0 = 0.12$, when $\beta = 0.1$. The difference between Example 4.7 and Example 4.6 is that $\mathbf{K} = \mathbf{R_0}$ in Example 4.7, and $\mathbf{K} = \mathbf{I_N}$ in Example 4.6. Event though $\mathbf{K}$ is different, the minimax design shown in Figure 4.7 is still optimal for both $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$. Here $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ is 0.2200, coincidentally the same as that when $\mathbf{K} = \mathbf{I_N}$, and $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 0.2069$. Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ for the same minimax design, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

Example 4.8: Suppose there are $t = 3$ treatments, and we want to compare the control treatment mean, say θ_1, with other treatment means θ_2 and θ_3. Let the

	r_1	r_2	r_3	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$
1)	4	4	4	0.7914	0.6694
2)	6	3	3	0.7521	0.6711
3)	2	5	5	1.3634	1.1641
4)	8	2	2	1.1897	1.0023
5)	10	1	1	2.4614	2.1530

Table 4.3: Loss functions for some of the ways to distribute 3 treatments into 12 plots under MA(1). r_i represents the number of replicates that treatment i receives.

contrast matrix

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}. \tag{4.4}$$

In previous examples, all treatments have the same number of replicates. However, if we want to compare treatment means, it is not necessary that all treatments have the same number of replicates. In this example, we want to find out the best way to distribute treatment replicates. Suppose $m = 3$ and $n = 4$. Some of the ways to distribute treatment replicates are shown in Table 4.3.

Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ for these five combinations. In this case, Combination 1) yields the least $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$ and Combination 2) yields the least $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$, where $\mathbf{R_0}$ is the covariance matrix from MA(1) with correlation parameter $\rho_0 = 0.2$, when $\beta = 0.3$.

3	1	2	1		2	1	2	3
1	3	1	3		1	3	1	2
2	1	2	1		3	1	2	3

(1) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = 0.7521$ (2) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 0.6694$

Figure 4.8: Minimax designs with size 3x4 and $t = 3$ under MA(1), (1) for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

Minimax design for the smallest $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ and $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$ are shown in Figure 4.8.

Example 4.9: Suppose there are $t = 8$ treatments, and we want to compare the control treatment mean, say θ_1, with other treatment means θ_2, ... , θ_8. Let the contrast matrix

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}. \tag{4.5}$$

In this example, we want to find out the best way to distribute treatment replicates.

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$
1)	3	3	3	3	3	3	3	3	5.2000	4.8890
2)	10	2	2	2	2	2	2	2	4.5450	4.3510
3)	17	1	1	1	1	1	1	1	8.5651	8.1668

Table 4.4: Loss functions for some of the ways to distribute 8 treatments into 24 plots under NN(1), where r_i represents the number of replicates that treatment i receives.

Suppose $m = 3$ and $n = 8$. Some of the ways to distribute treatment replicates are shown in Table 4.4.

Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ for these three combinations. In this case, Combination 2) yields the least $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}})$. Here $\mathbf{R_0}$ is the covariance matrix from NN(1) with correlation parameter $\rho_0 = 0.15$, when $\beta = 0.2$. Minimax design yields the smallest $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ and $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$ is shown in Figure 4.9. For this design, $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = 4.5450$ and $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 4.3510$.

Example 4.10: All the conditions are the same as those in Example 4.9, except that $\mathbf{R_0}$ is the covariance matrix from DG with correlation parameter $\lambda = 0.15$, when $\beta = 0.2$. Values of $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$ and $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$ for some of the ways to distribute treatment replicates are shown in Table 4.5.

Since $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) < \min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$

2 3 1 2 1 5 1 5

7 1 3 1 4 1 6 1

8 7 1 8 1 4 1 6

Figure 4.9: Minimax design with size 3x8 and $t = 8$ under NN(1) for both $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$	$\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$
1)	3	3	3	3	3	3	3	3	5.1573	5.0208
2)	10	2	2	2	2	2	2	2	4.5796	4.4573
3)	17	1	1	1	1	1	1	1	8.5490	8.2074

Table 4.5: Loss functions for some of the ways to distribute 8 treatments into 24 plots under DG, where r_i represents the number of replicates that treatment i receives.

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

(1) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}}) = 4.5796$ (2) $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}) = 4.4573$

Figure 4.10: Minimax designs with size 3x8 and $t = 8$ under DG, (1) for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and (2) for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$

for these three combinations. In this case, Combination 2) yields the least $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}})$. Minimax design for the smallest $\min_{\mathbf{X}} \hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ and $\min_{\mathbf{X}} \hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ are shown in Figure 4.10.

From the examples above, we can see that β controls the size of the neighbourhood. When β is small, the neighbourhood is small and $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is a more efficient estimator than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$. When β is large, the neighbourhood is large and $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$ may be a more efficient estimator than $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$. All treatments can have the same number of replicates when we want to estimate treatment means. When we want to compare treatment means with a control, having all treatments with the same number of replicates does not necessary lead to the most efficient treatment allocation.

Chapter 5

Conclusion

The objective of field experiments usually is to estimate and compare treatment means, $\boldsymbol{\theta}$. Spatial correlation is an important factor to be considered at the design stage. We presented a linear model that takes spatial correlation into account. In this thesis, we used the least squares and generalized least squares estimation methods to estimate treatment means. We used some scalar loss functions of the covariance matrix, $\mathcal{L}(\text{cov}(\hat{\boldsymbol{\theta}}))$, as our optimality criteria to construct optimal designs.

In Chapter 2, we discussed four types of spatial correlation structures. They are doubly-geometric correlation, discrete exponential correlation, moving average correlation, and the nearest neighbour correlation. We also presented optimal designs found in the literature and optimal designs for small number of plots and small number of treatments through complete enumeration and comparison. All optimal designs

shown in this chapter have specified correlation structures.

In Chapter 3, we proposed an approach to find robust optimal designs when the correlation structure is either unknown or estimated with uncertainty. First, we defined and gave examples of a neighbourhood of a covariance matrix $\mathbf{R}$, $\boldsymbol{R}_{\beta,\mathbf{K}}$. Then we proposed minimax designs which are solutions to the minimax problem, $\min_{\mathbf{X}} \max_{\mathbf{R}\in\boldsymbol{R}_{\beta,\mathbf{K}}} \mathcal{L}\ (\ \mathrm{cov}(\mathbf{C}\hat{\boldsymbol{\theta}})\ )$. In addition, we proposed a generalized least squares estimator $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$, which can be applied when $\mathbf{R}$ is not known exactly.

In Chapter 4, we introduced a simulated annealing algorithm that helps us find minimax designs when complete enumeration and comparison may seem impossible. Using this numerical algorithm, we can compute minimax designs for the least squares and generalized least squares estimation methods. By comparisons, it seems that when β is small, $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R_0}})$ is smaller than $\min_{\mathbf{X}} \mathcal{L}(\mathbf{X},\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{LS}})$. In other words, when the neighbourhood size is small, $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ is a more efficient estimator than $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

For practical applications, it is recommended to use minimax designs for $\hat{\boldsymbol{\theta}}_{\mathbf{R_0}}$ to estimate treatment means when we have some information about the covariance matrix of the spatial error process. If the covariance matrix is estimated with great uncertainties, then it may be better to use minimax designs for $\hat{\boldsymbol{\theta}}_{\boldsymbol{LS}}$.

Bibliography

[1] Becher, H. (1988), On Optimal Experimental Design under Spatial Correlation Structures for Squares and Nonsquare Plot Designs, *Communications in Statistics, Part B – Simulation and Computation*, **17**, 771-780.

[2] Cressie, N. (1993), *Statistics for Spatial Data*, Wiley, New York.

[3] Cochran, W.G. and Cox, G.M. (1960), *Experimental Designs*, Wiley, New York.

[4] Duby, C., Guyon, X. and Prum, B. (1977), The Precision of Different Experimental Designs for a Random Field, *Biometrika*, **64**, 59-66.

[5] Elliott, L.J., Eccleston, J.A., and Martin, R.J. (1999), An Algorithm for the Design of Factorial Experiments When the Data are Correlated, *Statistics and Computing*, **9**, 195-201.

[6] Fisher, R.A. (1966), *The Designs of Experiments*, Oliver and Boyd, Edinburgh.

[7] Haining, R.P. (1977), The Moving Average Model for Spatial Interaction, *Transactions of the Institute of British Geographer*, **3**, 202-25.

[8] Kiefer, J. and Wynn, H.P. (1981), Optimum Balanced Block and Latin Square designs for Correlated Observations, *Ann. Statist* , **9**, 737-57.

[9] Martin, R.J. (1979), A Subclass of Lattice Processes Applied to a Problem in Planar Sampling, *Biometrika*, **66**, 209-217.

[10] Martin, R.J. (1982), Some Aspects of Experimental Design and Analysis when Erros are Correlated, *Biometrika*, **69**, 597-612.

[11] Martin, R.J. (1986), On the Design of Experiments Under Spatial Correlation, *Biometrika*, **73**, 247-277.

[12] Martin, R.J. (1996), Spatial Experimental Design, *Design and Analysis of Experiments, [Handbook of Statistics 13]*, 477-514.

[13] Mead, R. (1988), *The Designs of Experiments: Statistical Principles for Practical Applications*, Cambridge University Press, Cambridge.

[14] Reingold, E., Nievergelt, J., and Deo, N. (1977), *Combinatiorial Algorithms (Theory and Practice)*, Prentice Hall, N.J..

[15] Zhou, J. (2001), A Robust Criterion for Experimental Designs for Serially Correlated Observations, *Technometrics*, **43**, 462-467.