Optimal experimental designs applied to Biostatistics.

Thesis for the Degree of Doctor of Mathematics

Santiago Campos Barreiro

Academic advisor:
Prof. Dr. Jesús López Fidalgo

Institute of Mathematics Applied to Science and Engineering
Department of Mathematics
University of Castilla-La Mancha

Ciudad Real, 2015
Assessment committee:

Dr. D. Fernando Plo Alastrué
University of Zaragoza.

Dra. Dña. María Jesús Rivas López
University of Salamanca.

Dr. D. Licesio J. Rodríguez Aragón
University of Castilla-La Mancha.
To Constanze
To my family
Acknowledgements

I would like to express my gratitude to various people without whom this work would not have been possible. Firstly, I would like to thank my supervisor, Prof. Dr. Jesús López-Fidalgo, for his advice and support during all these years. My deepest gratitude goes to my friend Víctor Ruiz for the help he provided with MATLAB and with physics concepts.

I would further like to express my gratitude to Hassan Mehdi, who made corrections to grammar and style, which led to a substantial improvement of this thesis. My grateful thanks are also extended to ENT consultant Dr. Eduardo Martín-Sanz for his invaluable advice with regard to clinical topics, to José San Millán for his help with LaTeX and to my very good friend Manuel Jiménez for answering questions relative to biological and veterinarian aspects.

Finally, thanks are also due to my colleagues of the group of Experimental Design at the University of Castilla-La Mancha, especially to Víctor Casero and Elvira Delgado, for their support throughout this project.
Contents

List of figures 11

List of tables 13

Summary 15

1 Optimal design of experiments theory 19

1.1 Introduction ........................................ 19

1.2 Linear model ........................................ 23

1.2.1 Estimates of the parameters .................. 24

1.3 Information matrix .................................. 27

1.3.1 Properties of the information matrix .......... 30

1.4 Optimality criteria ................................. 33

1.4.1 G-optimal criterion ............................. 35

1.4.2 D-optimal criterion ............................. 37

1.4.3 c-optimal criterion ............................. 38

1.5 Equivalence theorem ............................... 42

1.6 Non-linear model .................................. 49

2 Benign paroxysmal positional vertigo model 51

2.1 Introduction ........................................ 51

2.2 Derivation of a model for the maneuver ........ 53
2.3 A simple version of the model .................................. 58
  2.3.1 D-optimal design ........................................... 58
2.4 Optimal experimental designs ................................. 64
  2.4.1 D-optimal design for a discrete design space .......... 65
  2.4.2 D-optimal design for a continuous design space ...... 67
  2.4.3 c-optimal designs for a discrete design space ....... 68
  2.4.4 c-optimal designs for a continuous design space .... 70
2.5 Integral models ................................................. 71
  2.5.1 One-parameter model ...................................... 71
  2.5.2 Two-parameter model ...................................... 73

3 Growth model applied to a Friesian dairy farm .............. 81
  3.1 Introduction .................................................. 81
  3.2 Model for growth ............................................. 83
  3.3 Information matrix .......................................... 86
  3.4 Optimal designs .............................................. 91
  3.5 Robustness analysis ......................................... 95
    3.5.1 With respect to the choice of the nominal values .... 95
    3.5.2 Choice of the correlation structure ................... 95

4 Discrimination between growth models applied to a limousine
   beef farm ....................................................... 99
  4.1 Introduction .................................................. 99
  4.2 T-optimality and KL-optimality criterion .................. 101
  4.3 Discrimination between the two models .................... 104
4.3.1 Hypothesis test for discrimination .......................... 105
4.4 Algorithm ................................................................... 107
4.4.1 Calculation of the KL-optimal design ....................... 108
4.5 Robustness analysis .................................................... 110

Conclusions ..................................................................... 115

Bibliography ................................................................... 121

A How to apply the Elving’s method ................................. 131

B Navier-Stokes equations ............................................... 135

C Dix-Hallpike and Epley maneuvers ............................... 139
  C.1 Dix-Hallpike maneuver ........................................... 139
  C.2 Epley maneuver .................................................... 140

D Analytical primitive for equation (3.4) ......................... 143

E FIM for correlated observations .................................... 145

F Friesian and Limousine breeds ..................................... 149
  F.1 Friesian ............................................................... 149
    F.1.1 history .......................................................... 149
    F.1.2 Morphology .................................................... 149
    F.1.3 Productivity .................................................... 150
  F.2 Limousine ........................................................... 150
    F.2.1 Breed history .................................................. 150
G Data provided by the beef farm 153
List of Figures

1.1 Graphical method proposed by Elfving .......................... 41

2.1 Layout of the semicircular canals and ampulla .................... 52

2.2 Effects of active head movements .................................. 53

2.3 Composition of forces acting on the endolymph ................... 54

2.4 Plot of the function $d(t, \xi)$ ........................................ 60

2.5 Plot of the function $\text{eff}_r(\xi^*_0)$ ................................ 61

2.6 Elfving’s set. The functions $f(t)$ and $-f(t)$ are represented by
the continuous and dashed curves, respectively ...................... 63

2.7 Plot of the function $\eta(\Delta t, \alpha, \theta)$ for given nominal values .... 64

2.8 Plot of the generalized variance function for a continuous design
space ................................................................. 68

2.9 Elfving’s set for a discrete design space ............................. 69

2.10 Elfving’s set for a continuous design space ......................... 71

2.11 Plot of the generalized variance for example 1 .................... 76

2.12 Plot of the generalized variance for example 2 .................... 78

2.13 Plot of the Elfving’s set for example 2 ............................ 79

3.1 Relative efficiencies (in %) of the design $\xi^*_3$ for four different
values of the parameter $k$ and parameters $m_0$ and $M$ ............ 95

3.2 Plot of the correlation structures (3.14), (3.17), (3.18) and (3.19) 96
4.1 Relative efficiencies of the design $\xi^*$ for different values of the correlation and nugget parameters ........................................ 111

4.2 Relative efficiencies of the design $\xi^*$ for different values of the correlation and nugget parameters ........................................ 112

4.3 Relative efficiencies of the design $\xi^*$ for different values of the correlation and nugget parameters ........................................ 112

4.4 Relative efficiencies of the design $\xi^*$ for different values of the correlation and nugget parameters ........................................ 112

4.5 Plot of the correlation structures (4.3), (3.17), (3.18) and (3.19) 113

A.1 Convex hull of the set for the function $f(x) = (x(t), y(t))$. . . . 133

B.1 Strees rate variation over an element $dv$ of $V$. . . . . . . . . . 137

C.1 Dix-Hallpike maneuver. .................................................. 140

C.2 Epley maneuver. .......................................................... 141

F.1 Hostein-Friesian cattle. ................................................. 151

F.2 Limousine cattle. ......................................................... 152
List of Tables

2.1 Physical parameters ................................................. 65
2.2 Values of the generalized variance ................................. 66
2.3 Values of the efficiency ............................................. 67
2.4 Efficiency of the D-optimal design for example 1 ............... 75
2.5 Efficiency of the D-optimal design for example 2 ............... 78

3.1 Mean values of weights at their corresponding age from 250 heifers. These data have been provided by the growing facility .................. 89
3.2 Designs based on regular sequences. The four middle points are shown in the second column before summing up $t_1$ ......................... 93
3.3 Designs with different points at the support and their efficiencies with respect to $\xi_1^*$. The four middle points points are shown in the second column before summing up $t_1$ ......................... 93
3.4 Designs based on covariance structures (3.17), (3.18) and (3.19). The four middle points are shown in the second column before summing up $t_1$ ............. 97

4.1 Designs based on covariance structures (4.6), (4.7) and (4.8) and their corresponding efficiencies with respect o $\xi^*$. The six middle points are shown before summing up $t_1$ ................................................. 111
Summary

A well-designed experiment is an efficient method of learning about the behaviour of any stochastic process or particular system. Since even carefully controlled experiments cannot avoid random error, statistical methods are essential for their efficient design and analysis. The aim of this thesis is to design experiments based on mathematical foundations of optimal experimental design theory. These experiments belong to the field of Biostatistics.

In order to provide the reader with a clear idea about what an optimal design is, Chapter 1 is devoted to explain basic concepts, definitions, properties and theorems of this theory. The main optimal criteria as well as the Equivalence theorem are also presented as crucial tools to check the optimality of a design.

In Chapter 2 we deal with the pathology of the Benign Paroxysmal Positional Vertigo (BPPV), which is detected by a clinician through maneuvers consisting of a series of consecutive head turns that trigger the symptoms of vertigo in a patient. A biomechanical model is derived to provide a quantitative basis for the detection of BPPV and a simplification of the Navier-Stokes problem from the fluids theory has been used to construct the model. In addition, the same cubic splines that are commonly used in the kinematic control of robots were used to obtain an appropriate description of the different maneuvers.

This topic has never been treated in the literature of optimal designs and presents a novelty introduced through this thesis. In this chapter, D-optimal and c-optimal designs have been computed to obtain an optimal estimate of the model. These maneuvers consist of a series of specific head turns of duration $\Delta t$ and angle $\alpha$, that should be performed by the clinician on the patient. The designs indicate the duration and angle of the maneuver to be performed, as well as the corresponding proportion of replicates. The robustness study for the D-optimal design, with respect to the choice of the nominal values of the parameters, shows high efficiencies for small variations and provides a guide to the researcher. Furthermore, c-optimal designs give valuable assistance when checking how efficient the D-optimal design is for the estimation of each of
the parameters, allowing the physician to validate the model. The author has held consultations with an ENT consultant in order to align the outline more closely to practical scenarios.

In Chapter 3 we discuss design issues of West ontogenetic growth model applied to a Friesian dairy farm. The body mass growth of organisms is usually represented in terms of what is known as ontogenetic growth models, which represent the relation of dependence between the mass of the body and time. There are some theoretical works where optimal designs for growth models have been calculated. But in these cases, uncorrelated observations were assumed.

In this chapter we move beyond independent observations. A correlation structure has been included in the statistical model due to the fact that observations on a particular animal are not independent. The presence of correlation in the observations implies the information matrix is not additive anymore. Accordingly, we must restrict ourselves to a practical exact design, fixing a priori the number of time-points of the design. Moreover, in the covariance structure, a nugget term has been included since optimal design points may collapse under the presence of correlation when no nugget effect is present. As replicates of measurements at the same time, on the same heifer, do not make sense from a practical point of view, the designs calculated must contain $n$ distinct time-points.

A sensitivity analysis versus the choice of the correlation structure has been carried out. The choice of a robust correlation structure provides a methodology that can be used for any correlation structure. The experimental designs undertaken are computed to obtain an optimal fitting of the model, providing a tool to control the proper mass of heifers, which will help improve their productivity and, by extension, the competitiveness of the dairy farm.

Chapter 4 is concerned with a problem of finding an optimal experimental design for discriminating between two competing growth models applied to a calf growing facility. In order to optimize the efficiency of beef production systems, it is of great importance to know the behaviour of mass gain in cattle. The growth of beef-specialized breeds is characterized by models based on non-linear sigmoid curves. The shape and characteristics of these curves can vary
depending on factors such as the environment, production system, type of
breed and so on.

In this chapter, the problem of constructing optimal experimental designs for
discriminating between Brody and Richards models is considered. These two
models have already been compared for cattle, though in none of them this
comparison carried out using optimal designs. The T-optimal criterion is usu-
ally used for comparisons between models. But in this work, a criterion based
on the Kullback-Leibler distance is used the so called KL-optimal criterion.
We will also demonstrate that this criterion is an extension of T-optimality
when correlated observations are considered. Besides, a nugget term has been
included in the covariance structure, as we did in former chapter, and a sen-
sitivity analysis versus the choice of the correlation structure has been also
carried out. In order to compute optimal designs, the exchange-type algorithm
developed by Brinkulov et al. (1986) has been adapted to KL-optimality.

The final part of this thesis contains seven Appendices which can be read
to further supporting information and a greater level of detail. Following
that, the conclusions generated by this work are along with future research
topics. We finish showing general bibliographic references which have been
used throughout this thesis.
Chapter 1

Optimal design of experiments theory

1.1 Introduction

One of the main aims of Statistics consists of modeling the behaviour of any stochastic process or particular system. In order to carry this out, a mathematical model is always needed to explain the results obtained once the process has occurred, and predict accurately the future behaviour of that process or system. The mathematical models used in real situations depend on unknown parameters. In order to describe the way in which the results are expected to vary, we need to estimate these parameters with optimal accuracy. For that purpose, the Design of Experiments is used to help us design how to change the inputs of processes in order to observe and identify the reasons of the changes observed in the response $y$, which is usually expressed as follows,

$$y = \eta(x, \theta) + \varepsilon,$$

where $x$ represents the influence of a series of factors or experimental conditions, chosen from a given experimental domain $\chi$. In practice, $\chi$ will be a compact subset of an Euclidean Space, usually $\chi \subset \mathbb{R}$ or $\chi \subset \mathbb{R}^p$. The unknown parameters of the model are represented by $\theta$. The error $\varepsilon$ usually follows a Gaussian distribution with mean $0$ and constant variance $\sigma^2$, that is

$$E(y) = \eta(x, \theta) \quad \text{and} \quad \text{var}(y) = \sigma^2(x).$$

The problem of determining which set of observations to collect is what defines the design. It is common to say that the input variables are controlled by the researcher, while the unknown parameters are determined by nature.
In an experiment, we deliberately change one or more process variables (or factors) in order to observe the effect the changes have on one or more response variables. The statistical design of experiments is an efficient procedure for planning experiments so that the data obtained can be analyzed to yield valid and objective conclusions. An experimental design is the laying out of a detailed experimental plan in advance of doing the experiment. Well chosen experimental designs maximize the amount of information that can be obtained for a given amount of experimental effort.

There are many statistical issues to consider in the design of an empirical study. The questions about the control of unwanted variation and the validity of a study have been already discussed in the statistical literature by Cox (1958), Campbell and Stanley (1963) and Cook and Campbell (1979), among others. However, one important aspect that seems lacking in the discussion is the question whether the design is more or less efficient with regard to the objective of the study. Optimal design of experiments theory allows us to find the best design in the sense of obtaining an optimal estimate of the parameters of the model. Although this theory has been available for many years, scientists seem to have little exposure to its theory and potential applications in their fields. One of the aims of this thesis is to promote interest among researchers in the use of optimal design theory and enable them to design more efficient and less expensive studies.

The origin of this theory dates from the beginning of last century, when Smith (1918) proposed optimal designs for polynomial models. Since the famous work of Fisher (1935), statisticians have worked on optimal design problems. A small sample of references of these early years were Wald (1943), Hotelling (1944) and Elving (1952). There was a surge in research in this area in the early 1960s after the seminal papers by Kiefer and Wolfowitz (1959, 1960), showing that a design can be simultaneously optimal under two very different and useful objectives in the study.

The pioneering book, in English, on optimal experimental design was published by Fedorov (1972). Silvey (1980) provided a concise introduction to the central theory of the General Equivalence Theorem. The monographs by Kiefer (1985) was a voluminous collection of pioneering work in this area by the author and provided a good account of the chronological developments.
of optimal design theory. More theoretical treatments were given by Pázman (1986), Pukelsheim (1993), Schwabe (1996), Fedorov and Hackl (1997) and Walter and Pronzato (1997), which provided an introduction to the theory with engineering examples. However, despite the ubiquity of design issues in all studies, the number of statisticians working actively in this important area has always been relatively small, compared to the larger statistical community working on data analysis issues.

Among the most recently books dedicated to optimal design, stand out the works of Pistone, Riccomagno y Wynn (2000), implementing algebraic methods to the optimal design experiments, or Ucinski and Bogacka (2005), which is devoted to the design of experiments where measurements are taken over time, perhaps from measuring devices, or sensors, whose optimal spatial trajectories are to be determined.

Developments in the theory and practice of optimal experimental design can be followed in the proceedings volumes of the MODA conferences. The most recently, MODA10, was celebrated in 2013 in Lagów Lubuski (Poland). The main topics were about compartimental models, mixed-effect models and clinical essays.

The optimal design of experiments is based on a theory which, like any general theory, provides a unification of many separate results and a way of generating new results in novel situations. Part of this flexibility results from the algorithms derived from the General Equivalence Theorem, combined with computer-intensive search methods. The use of optimal or highly efficient designs in scientific or social research has advantages. Fewer observations and therefore smaller sample sizes are required to find real effects, thus reducing the costs of the study. Our examples in this thesis show that optimal designs can reduce the number of observations even in 40% in some cases when compared with the traditional used designs. This is especially beneficial in light of the ever-rising cost of conducting scientific or social studies. From an ethical viewpoint, a smaller sample is also highly desirable. For example, fewer patients may be required to undergo a controversial treatment or fewer animals need to be sacrificed in a toxicology study.

Initially, optimal experimental design theory was developed for linear regre-
ssion models although in practice, many real problems need to be modeled by using non-linear ones. Next, a brief description about the main aspects of this theory for linear and non-linear models is presented.
1.2 Linear model

If the function $\eta$ is a linear function on $\theta$, then the model is called *linear model*:

$$\eta(x, \theta) = f^t \theta = \sum_{j=1}^{m} f_j \theta_j.$$  

In this situation, the model (1.1) becomes

$$y = \sum_{j=1}^{m} f_j \theta_j + \varepsilon, \quad f_j : \chi \rightarrow \mathbb{R}, \quad (1.2)$$

where

$$E(y) = \eta(x, \theta) \quad \text{and} \quad \text{var}(y) = \sigma^2(x).$$

If the variance is constant, the model is called *homoscedastic model*. Otherwise, it will be called *heteroscedastic model*. In the experimental field it is plausible to suppose that variance of the response variable depends on the control variables through a known, well defined on $\chi$, bounded, continuous, and positive function $\lambda(x)$, that is:

$$\text{var}[y] = \sigma^2(z) = \frac{\sigma^2}{\lambda(x)}.$$ 

This function is called *efficiency function*.

**Remark.** As $\lambda(x)$ has the above mentioned properties, we can apply the following transformations to get a homoscedastic model,

$$y\sqrt{\lambda(x)} \rightarrow y^*; \quad f(x)\sqrt{\lambda(x)} \rightarrow f^*(x).$$

Due to simplicity of the transformations, we can suppose, without loss of generality, that the value of the efficiency function is $\lambda(x) = 1$. 
1.2.1 Estimates of the parameters

The experimental design consists of a planned collection of $N$ points $x_1, \ldots, x_N$ where the trials are going to be carried out. For each trial $x_i$, a response $y_i$ is obtained. Assuming the linear model, we can write

$$E[y_i] = \sum_{j=1}^{m} f_j(x_i) \theta_j, \quad i = 1, \ldots, N; \quad j = 1, \ldots, m.$$ 

Let us denote by

$$Y = (y_1, \ldots, y_N),$$

the observations vector and

$$X = (x_{ij}), \quad \text{where} \quad x_{ij} = f_j(x_i).$$

Therefore,

$$X = \begin{pmatrix}
  f_1(x_1) & f_2(x_2) & \cdots & f_m(x_1) \\
  \vdots & \vdots & \ddots & \vdots \\
  f_1(x_N) & f_2(x_N) & \cdots & f_m(x_N)
\end{pmatrix} \quad \text{matrix of the design.}$$

Using this matrix notation and assuming uncorrelated observations, the linear model becomes

$$E[Y] = X\theta, \quad \text{Cov}(Y|\theta, \sigma^2, X) = \sigma^2 I_N.$$ 

Parameters are estimated by using the least-squares method. Among all the estimators of $\theta$, we denote by $\hat{\theta}$ the one which minimizes

$$\min_{\theta} S(\theta) = \min_{\theta} \sum_{i=1}^{N} (y_i - f^t(x_i)\theta)^2.$$
1.2. LINEAR MODEL

Deriving $S(\theta)$ with respect to $\theta_j$, and setting the gradient to zero,

$$\frac{\partial S(\theta)}{\partial \theta_j} = \sum_{i=1}^{N} 2(y_i - f^t(x_i)\theta)(-f_j(x_i)) = 0.$$

The gradient equations become

$$\sum_{i=1}^{N} f_j(x_i)(f^t(x_i)\theta) = \sum_{i=1}^{N} f_j(x_i)y_i \quad j = 1, \ldots, m.$$

These are the equations of the equality

$$\begin{pmatrix} f_1(x_1) & \cdots & f_1(x_N) \\ \vdots & \ddots & \vdots \\ f_m(x_1) & \cdots & f_m(x_N) \end{pmatrix} \begin{pmatrix} f_1(x_1) \\ \vdots \quad \vdots \\ f_1(x_N) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \vdots \quad \vdots \\ \theta_m \end{pmatrix} = \begin{pmatrix} f_1(x_1) & \cdots & f_1(x_N) \\ \vdots & \ddots & \vdots \\ f_m(x_1) & \cdots & f_m(x_N) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \quad \vdots \\ y_m \end{pmatrix}.$$

In matrix form:

$$X^tX\hat{\theta} = X^tY$$

These equations are called normal equations system.

If $X^tX$ is non singular, then

$$\hat{\theta} = (X^tX)^{-1}X^tY,$$ (1.3)

and the covariance matrix is given by the following expression.

**Proposition.** It must be verified that:

$$\text{cov}(\hat{\theta} \mid \sigma^2) = \sigma^2 (X^tX)^{-1}.$$

**Proof.** Using that

$$\text{cov}(AY) = A\text{cov}(Y)A^t,$$

we have

$$\text{cov}(ABY) = AB\text{cov}(Y)(AB)^t.$$
If we denote $A$ and $B$ by

$$A = (X^t X)^{-1} \text{ and } B = X^t,$$

then

$$\text{cov}(ABY) = \text{cov}[(X^t X)^{-1} X^t Y] = (X^t X)^{-1} X^t \text{cov}(Y)[(X^t X)^{-1} X^t]^t.$$

As

$$\text{cov}(Y) = \sigma^2 I,$$

it results that

$$\text{cov}(\hat{\theta}) = \sigma^2 (X^t X)^{-1} X^t (X^t)^t [(X^t X)^{-1}]^t.$$

Moreover,

$$(X^t)^t = X^t (X^t)^{-1} = [(X^t)^t]^{-1} = (X^t X)^{-1},$$

so finally we have

$$\text{cov}(\hat{\theta}) = \sigma^2 (X^t X)^{-1} X^t X (X^t X)^{-1} = \sigma^2 (X^t X)^{-1}.$$

The value of the response $y$ at an unobserved point $x$ can be predicted as

$$\hat{y} = f^t(x) \hat{\theta},$$

and the variance of this estimate becomes

$$\text{var}(\hat{y}_x) = \sigma^2 f^t(x)(X^t X)^{-1} f(x).$$

The objective to be achieved is to find a design which gives somewhat the best estimation of the parameters (or linear functions of them) or a good prediction of $\hat{y}_x$. Through what it is defined as criterion $\Phi$, we measure the accuracy of the design and we can compare different designs of the same model.
1.3 Information matrix

**Definition.** An exact experimental design of size $N$ consists of a set of $N$ observations collected at points $V_N = \{x_1, \cdots, x_N\}$, in a given compact space $\chi$. Some of these $N$ points may be repeated, meaning that several observations are taken at the same value of $x$. This number of observations is usually predetermined by experimental cost constraints.

With these points the matrix of the design $X$ is formed, and from this, we calculate $X^tX$. This matrix is actually the Fisher information matrix for $\theta$, where observations are assumed normally distributed. If $X^tX$ is non-singular, its inverse is proportional to the covariance matrix of the estimator $\hat{\theta}$. The use of this matrix is very important when it comes to designing the experiment in an optimal way.

**Definition.** The information matrix of an exact design of $N$ points, where some of them may be repeated, may be written as

$$M(V_N) = \frac{1}{N} \sigma^{-2} X^t X = \frac{1}{N} \sum_{i=1}^{N} \sigma^{-2} f(x_i) f^t(x_i).$$

(1.4)

**Remark.** For the linear model, the matrix $X^tX$ results from the sum of $N$ matrices of rank one. We prove it for $N = 1$:

$$\begin{pmatrix} f_1(x_1) \\ f_2(x_1) \\ \vdots \\ f_m(x_1) \end{pmatrix} \begin{pmatrix} f_1(x_1) \\ \vdots \\ f_m(x_1) \end{pmatrix} = \begin{pmatrix} f_1(x_1)f_1(x_1) & \cdots & f_1(x_1)f_m(x_1) \\ \vdots & \ddots & \vdots \\ f_m(x_1)f_1(x_1) & \cdots & f_m(x_1)f_m(x_1) \end{pmatrix}.$$ 

If we write the determinant and remove $f_1(x_1)$ from the first row, $f_2(x_1)$ from
the second row and so on until removing \( f_m(x_1) \) from the k-th row,

\[
|X^tX| = f_1(x_1)f_2(x_1) \cdots f_m(x_1) = 0.
\]

**Remark.** As mentioned above, some of the N points may be repeated. But if we want the expression (1.4) only to contain not repeated points, each point can be associated with the number of times it appears repeated:

\[
X^tX = \sum_{i=1}^{N} f(x_i)f^t(x_i).
\]

If from the N points, k are different one another and each point repeats itself \( r_i \) times,

\[
X^tX = \sum_{i=1}^{N} f(x_i)f^t(x_i) = \sum_{i=1}^{k} r_if(x_i)f^t(x_i) = N\sum_{i=1}^{k} p_if(x_i)f^t(x_i),
\]

where \( p_i = \frac{r_i}{N} \).

Let us divide \( M(V_N) \) by N:

\[
\frac{1}{N}M(V_N) = \frac{1}{N} \sigma^{-2}X^tX = \sum_{i=1}^{k} \sigma^{-2} p_if(x_i)f^t(x_i).
\]

With this new matrix resulting from the division by N, it is easier to work than doing directly with \( X^tX \). This matrix is called *Average information matrix*, although in order to simplify the notation, it is called just *information matrix*.

Let us suppose now that the collection of weights \( p_i \) provide a probability measure on \( \chi \) supported on the non-repeated points \( x_1, \cdots, x_N \). This leads us to the definition of an *approximate design* as a discrete probability measure on \( \chi \) with a finite support. It is denoted by \( \xi_N \) (Kiefer, 1959). Now, the weights
$p_i$ can be irrational numbers. A design $\xi_N$ is represented as follows:

$$\xi_N = \left\{ \frac{x_1}{p_1}, \frac{x_2}{p_2}, \ldots, \frac{x_N}{p_N} \right\}, \quad \sum_{i=1}^{N} p_i = 1.$$

**Definition.** A *continuous design* is any probability measure $\xi$ on the Lebesgue $\sigma$-Algebra of $\chi$.

Let us then define the information matrix of a continuous design $\xi$ as follows:

$$M(\xi) = \int_{\chi} f(x) f^t(x) \xi(dx).$$

If the design is supported on a finite set, the information matrix is expressed as

$$M(\xi) = \int_{\chi} \sigma^{-2} f(x) f^t(x) \xi(dx) = \sum_{x \in \chi} \sigma^{-2} \xi(x) f(x) f^t(x),$$

where $\xi(x)$ is the weight on the point $x$ with respect to the measure $\xi$.

**Remark.** In practice, each design is an exact design since we always collect an integer number of observations. Therefore, approximate designs have no practical interpretation in terms of experiments although they are very convenient in order to search for optimal designs since they have very good properties. The discrete nature of the exact design causes serious difficulties to put into practice usual optimization techniques, particularly when $N$ is large (Atkinson, 2007). An analogous situation occurs when the minimum of a function defined on a set of integer numbers is wanted.

**Proposition.** If $\Xi$ is the set of all designs defined on $\chi$, then $\Xi$ is convex.

**Proof.** Let us suppose we have two designs $\xi_1$ and $\xi_2$ supported on $N_1$ and $N_2$ points, respectively:

$$\xi_1 = \left\{ \frac{x_{1i}}{p_{1i}} \right\}_{i=1}^{N_1}, \quad \xi_2 = \left\{ \frac{x_{2i}}{p_{2i}} \right\}_{i=1}^{N_2}.$$
We define $\xi$ as a convex linear combination of $\xi_1$ and $\xi_2$:

$$\xi = (1 - \alpha)\xi_1 + \alpha\xi_2, \quad 0 \leq \alpha \leq 1.$$ 

The support of $\xi$ is the union of supports of $\xi_1$ and $\xi_2$ and then we put weight $(1 - \alpha)p_{1i}$ on $x_{1i}$ only if this point belongs to the support of $\xi_1$, $\alpha p_{2i}$ on $x_{2i}$ only if this point belongs to the support of $\xi_2$ and weight $(1 - \alpha)p_{1i} + \alpha p_{2i}$ for the points belonging to both $\xi_1$ and $\xi_2$. It is clear that $\xi$ is a design.

**Definition.** We define $\mathcal{M}$ as the set of all the information matrices,

$$\mathcal{M} = \{ M(\xi) : \xi \in \Xi \}.$$ 

### 1.3.1 Properties of the information matrix

The structure of $\mathcal{M}$ is less complex than the one of $\Xi$. We show now some of its more relevant properties.

**Remark.** From now on, it will be supposed that $\sigma^2 = 1$. Within the context of the optimal design that does not mean a loss of generality if $\sigma^2 f(x)$ is replaced by $f(x)$.

**Property.** The set $\mathcal{M}$ is convex.

**Proof.** If $\xi_1, \xi_2 \in \Xi$ and $0 < \lambda < 1$, then,

$$(1 - \lambda)M(\xi_1) + \lambda M(\xi_2) = \sum_{x \in \chi} (1 - \lambda)\xi_1(x)f(x)f^t(x) + \sum_{x \in \chi} \lambda\xi_1(x)f(x)f^t(x)$$

$$= \sum_{x \in \chi} [(1 - \lambda)\xi_1 + \lambda\xi_1(x)] f(x)f^t(x) = M((1 - \lambda)\xi_1 + \lambda\xi_2) \in \mathcal{M}.$$

**Property.** The information matrix is symmetric and non-negative definite.
Proof. The property of symmetry can be easily deduced from the definition of the information matrix. With respect to the second property, let a vector be $a \in \mathbb{R}^m$, $a \neq 0$, then
\[
a^t M(\xi)a = \int_\chi a^t f(x) f^t(x) a \xi(dx) = \int_\chi (a^t f(x))^2 \xi(dx) \geq 0.
\]

Lemma. The set $\mathcal{C} = \{f(x)f^t(x), \ x \in \chi\}$ has dimension $\frac{1}{2} m(m + 1)$, $m$ being the number of parameters of the model.

Proof. $\mathcal{C}$ is formed by symmetric matrices of order $m$, which has dimension $m^2$. the elements which can vary independently are those of the diagonal and those above it. This last set of elements is denoted by $\Delta$. Therefore,
\[
m^2 = 2 \dim(\Delta) + \dim(\text{diagonal}),
\]
and
\[
\dim(\Delta) = \frac{m^2 - m}{2}.
\]
Consequently,
\[
\dim(\mathcal{C}) = \dim(\Delta) + \dim(\text{diagonal}) = \frac{m^2 - m}{2} + m = \frac{1}{2} m(m + 1).
\]

Property. The set $\mathcal{M}$ is the convex hull of
\[
\mathcal{C} = \{f(x)f^t(x), \ x \in \chi\}.
\]

Proof. Direct from the definition of information matrix.

Property. $M(\xi)$ is singular if $\xi$ is supported on a number of points less than $m$. 

\[
\]
Proof. 

\[ M(\xi) = \sum_{i=1}^{r} \xi(x)f(x)f^{t}(x), \quad \text{being} \quad r < m. \]

Since the rank of \( f(x_i)f^{t}(x_i) = 1, \forall i, \) the rank of the sum is less than or equal to \( r \) and, therefore, \( M(\xi) \) is not regular. \( \square \)

Proposition. \textit{Given a design} \( \xi \), \textit{it exists a design} \( \zeta \) \textit{verifying that} \( M(\xi) = M(\zeta) \) \textit{and supported on, at most}, \( \frac{1}{2}m(m + 1) + 1 \) \textit{points}. \textit{For an element of} \( \mathcal{M} \) \textit{belonging to its boundary, this number of points is reduced in one unit.}

Proof. Since the set \( \mathcal{M} \) is the convex hull of \( \mathcal{C} \), each element of \( C_0(\mathcal{C}) \) can be associated univocally to an information matrix. Caratheodory’s theorem assures that exists

\[ \gamma_i \in [0, 1], \ x_i \in \chi, \ i = 1, \ldots, \frac{1}{2}m(m + 1) + 1 \]

such that

\[ M(\xi) = \sum_{i=1}^{j} \gamma_m f(x_i)f^{t}(x_i), \ j = \frac{1}{2}m(m + 1) + 1. \]

Therefore the design is

\[ \zeta = \begin{pmatrix} x_1 & x_2 & \cdots \\ \gamma_1 & \gamma_2 & \cdots \end{pmatrix}. \]

\( \square \)

Remark. Designs \( \xi \) y \( \zeta \) are equivalent, if and only if, \( M(\xi) = M(\zeta) \). (Pázman, 1986).
1.4. **OPTIMALITY CRITERIA**

### 1.4 Optimality criteria

The aim to be achieved is to find a design which gives the best estimation of the parameters (or linear functions of them) or an optimal estimation of the variable $y$ at an unobserved point. This is carried out by choosing $N$ points from the set $\chi$ which minimize $\text{cov}(\hat{\theta} | \sigma^2) \circ \text{var}(\hat{y}_x)$. Nevertheless, although one design can minimize the variance of certain linear functional, the variance of another one can be large. The choice of the optimal design depends on the interests of the researchers, on how easy become the calculations or on another subjective aspects. In this work, we focus on the first of this criterion, that is, the minimization of the unknown parameters of the model. Moreover, the matrix $\text{cov}(\hat{\theta})$ depends on the observations through the matrix of the design $M(\xi)$, therefore the optimality criteria must be defined over the set $M$. One important characteristic of the linear model is that the information matrix does not depend on the parameters of the model. For that reason the experiment is designed on the basis of this matrix.

**Definition.** A lower bounded function

$$\Phi : M \rightarrow \mathbb{R} \cup \{+\infty\}$$

is called a criterion function, if $M(\zeta) \leq M(\xi)$ implies that $\Phi[M(\xi)] \leq \Phi[M(\zeta)]$, the order $\leq$ being the Loewner order, that is,

$$A \leq B \text{ is and only if } A-B \text{ es positive definite.}$$

**Definition.** The design $\xi^* \in \Xi$ is called $\Phi$-optimal if it satisfies

$$\Phi[M(\xi^*)] = \min_{\xi \in \Xi} \Phi[M(\xi)].$$

The aim of the optimal design theory is to find the design $\xi^*$ which minimizes $\Phi[M(\xi)]$. The function $\Phi$ should verify these two important properties,

1. Convexity:

$$\Phi[(1-\alpha)M(\xi_1) + \alpha M(\xi_2)] \leq (1-\alpha)\Phi[M(\xi_1)] + \alpha \Phi[M(\xi_2)], \quad \alpha \in [0,1].$$
2. Positive homogeneity:

\[ \Phi[\delta M(\xi)] = \frac{1}{\delta} \Phi[M(\xi)], \quad \delta > 0. \]

Remark. Given an optimal design, each design with the same information matrix, is an optimal design, although there can be different information matrices which lead to optimal designs.

Notation. We denote

\[ \mathcal{M}_+ = \{M \in \mathcal{M} : \text{det}(M) > 0\}. \]
\[ \mathcal{M}_\Phi = \{M \in \mathcal{M} : \Phi(M) < +\infty\}. \]
\[ \Xi^*_\Phi = \{\xi \in \Xi : \xi \text{ is } \Phi\text{-optimal}\}. \]

Proposition. If \( \Phi \) is a convex function, then \( \Xi^*_\Phi \) is convex.

Proof. Let us suppose that

\[ \Phi[M(\xi_1)] = \Phi[M(\xi_2)] = \min_{\xi \in \Xi} \Phi[M(\xi)]. \]

Then, since \( \Phi \) is convex,

\[ \Phi\{(1 - \alpha)\xi_1 + \alpha\xi_2\} \leq (1 - \alpha)\Phi[M(\xi_1)] + \alpha\Phi[M(\xi_2)] = \min_{\xi \in \Xi} \Phi[M(\xi)]. \]

Definition. A function \( \Phi \) is a strictly decreasing criterion function if

\[ N \leq M \text{ y } N \neq M, \]

implies that

\[ \Phi(N) > \Phi(M). \]

Corollary. If \( \Phi \) is a strictly increasing criteria function, each \( \Phi \)-optimal design, \( \xi^* \), will have its information matrix in the boundary of the set \( \mathcal{M} \), that is, the number of points of the design \( \xi^* \) are between \( m \) and \( \frac{1}{2}m(m+1) \).
The goodness of any design $\xi$ with respect to the optimal design $\xi^*$ is measured by its efficiency. It is defined as

\[ \text{eff}_\Phi(\xi) = \frac{\Phi[M(\xi^*)]}{\Phi[M(\xi)]}, \]

where $\xi^*$ is the optimal design for the criteria $\Phi$. Obviously, the efficiency is between 0 and 1. Moreover, if for any design $\xi$ the number of observations is $N$ and its efficiency with respect to $\xi^*$ is $\rho$, then

\[ \rho = \frac{\Phi[M(\xi^*)]}{\Phi[M(\xi)]} = \frac{\Phi[N^{-1}(\sum_{\hat{\theta}})^{-1}]}{\Phi[N^{-1}(\sum_{\hat{\theta}})^{-1}]} = \frac{N\Phi[(\sum_{\hat{\theta}})^{-1}]}{N\Phi[(\sum_{\hat{\theta}})^{-1}]} = \frac{\Phi[(\sum_{\hat{\theta}})^{-1}]}{\Phi[(\sum_{\hat{\theta}})^{-1}]} , \]

$\sum_{\hat{\theta}}$ being the covariance matrix of $\hat{\theta}$.

**Lemma.** Given a design $\xi$ with efficiency $\rho$ with respect to the optimal design $\xi^*$, a number $N^* = N\rho$ of observations are enough for $\xi^*$ to have the same accuracy for the estimation as $\xi$ with $N$ observations.

**Proof.**

\[ 1 = \frac{\Phi[(\sum_{\hat{\theta}})^{-1}]}{\Phi[(\sum_{\hat{\theta}})^{-1}]} = \frac{(N^*)^{-1}\Phi[M(\xi^*)]}{N^{-1}\Phi[M(\xi)]} = \frac{(N^*)^{-1}}{N^{-1}} \rho. \]

Therefore, $N^* = N\rho$. \hfill \Box

**Remark.** If the efficiency of any design $\xi$ is of 60%, that means if we take the optimal design $\xi^*$, with the 60% of the observations which have been collected for $\xi$, the same accuracy for the estimation of the parameters is reached.

### 1.4.1 G-optimal criterion

This is a criterion based on the response $y$. The model can be chosen so that minimizes the variance of the response or maximizes its inverse. For the linear model and for one fixed $x$,

\[ \hat{y}_x = f^t(x)\hat{\theta}, \quad \text{var}(\hat{y}_x) \propto f^t(x)M^{-1}(\xi)f(x). \]

The generalized variance for a design $\xi$ is denoted by $d(x, \xi)$ and has the
following form:

\[
d(x, \xi) = f^t(x)M^{-1}(\xi)f(x).
\]

This function is proportional to the variance of the estimate of the response at point \(x\). The criterion consists in minimizing

\[
\max_{x \in \chi} d(x, \xi),
\]

that is, to minimize the largest possible value of the variance in \(\chi\):

\[
\Phi_G[M(\xi)] = \begin{cases} 
\max_{x \in \chi} f^t(x)M^{-1}(\xi)f(x) & \text{if } \det M(\xi) \neq 0 \\
\infty & \text{if } \det M(\xi) = 0.
\end{cases}
\]

The G-optimal criterion was suggested by Smith (1918).

**Proposition.**

1. \(\Phi_G\) is continuous in \(M\).
2. \(\Phi_G\) is convex in \(M\) and strictly convex in \(M_+\).

**Proof.** See Pázmán (1986).

**Remark.** Let \(\xi\) be a G-optimal design for the problem with regression function \(f(x)\) and the above defined efficiency function \(\lambda(x)\). Then \(\xi\) is also a G-optimal design for the problem designed with regression function \(g(x) = Lf(x)\) and efficiency function \(c\lambda(x)\), where \(L\) is a non-singular \(m \times m\) matrix and \(c > 0\) (Cook and Wonk, 1994).

**Definition.** Let \(M_m(\mathbb{R}) \cong \mathbb{R}^{m \times m}\) be the set of all real square matrices of order \(m\). If \(A \in M_m(\mathbb{R})\), then the trace of \(A\) is

\[
tr(A) = \sum_{j=1}^{m} [A]_{jj}.
\]

We denote by \(\langle A, B \rangle\) the inner product

\[
tr(A^T B),
\]
and by $\|A\|$ the norm in $\mathcal{M}_m(\mathbb{R})$ given by

$$\|A\| = \text{tr}(A^T A)^{1/2}.$$ 

**Definition.** Let $\Phi$ be a differentiable function defined in a neighborhood of a matrix $A$ in the space $\mathcal{M}_m(\mathbb{R})$. Then the gradient of the function $\Phi$ at a point $A$ is the $(m \times m)$ matrix

$$\{\nabla \Phi(A)\}_{ij} \equiv \left\{ \frac{\partial \Phi(A)}{\partial a_{ij}} \right\}_{i,j}$$

$i, j = 1, \ldots, m$.

**Example.** Let the function $\Psi$ be defined as

$$\Psi : \mathcal{M}_2(\mathbb{R}) \rightarrow \mathbb{R}$$

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \rightarrow x_{11}^2 - 2x_{11}x_{21} + x_{22}.$$ 

$\Psi$ is an scalar function and its gradient is

$$\nabla \Psi = \begin{bmatrix} \frac{\partial \Psi(A)}{\partial a_{ij}} \end{bmatrix}_{i,j} = \begin{bmatrix} \frac{\partial \Psi}{\partial x_{11}} & \frac{\partial \Psi}{\partial x_{12}} \\ \frac{\partial \Psi}{\partial x_{21}} & \frac{\partial \Psi}{\partial x_{22}} \end{bmatrix} = \begin{bmatrix} 2x_{11} - 2x_{21} & 0 \\ -2x_{11} & 1 \end{bmatrix}.$$ 

At some point $A$,

$$A = \begin{pmatrix} 3 & 2 \\ 4 & 0 \end{pmatrix} \in \mathbb{R}^4, \quad \nabla \Psi(A) = \begin{bmatrix} -2 & 0 \\ -6 & 1 \end{bmatrix}.$$ 

### 1.4.2 D-optimal criterion

This criterion was proposed by Wald (1943). It consists in minimizing the power $-1/m$ of the determinant:

$$\Phi[M(\xi)] = |M(\xi)|^{-1/m}.$$
Remark. This criteria is equivalent to maximize the determinant of the information matrix. The power is considered in order to become positive homogeneous. It is verified that

$$|\text{cov}(\hat{\theta})| \propto |M(\xi)|^{-1} = \prod_{j=1}^{m} \lambda_j^{-1}.$$  

where $\lambda_1, \cdots, \lambda_m$ are the eigenvalues of $M(\xi)$. Since the logarithm is an increasingly function, we can also define this criterion through the following criteria function:

$$\Phi_D[M(\xi)] = \begin{cases} 
\log \det \, M^{-1}(\xi) = -\log \det M(\xi) & \text{if } \det M(\xi) \neq 0 \\
\infty & \text{if } \det M(\xi) = 0.
\end{cases}$$

Proposition.

1. $\Phi_D$ is continuous on $\mathcal{M}$.

2. $\Phi_D$ is convex on $\mathcal{M}$ and strictly convex on $\mathcal{M}_+$.

3. If $\Phi_D$ is differentiable, its gradient will be:

$$\nabla[- \log \det M] = -M^{-1}.$$  

Proof. See Pázman (1986). □

Remark. The advantage of this criterion with respect to the others lies in its easy calculation.

1.4.3 c-optimal criterion

If we are interested in estimating a linear combination of the parameters, say $c^T \theta$, then we use the $c$-optimal criterion. It consists in minimizing the function

$$\Phi_c[M(\xi)] = c^t M^{-1}(\xi)c.$$  

This criteria is used since the variance of $c^T \theta$ is proportional to $c^T M^{-1}(\xi)c$. 
The disadvantage of this criteria is that they are frequently singular. An elegant way for finding an optimal design that estimates a linear combination of the parameters was given by Elfving (1952), and it is nicely illustrated and explained, e.g. by Chernoff (1972), Kitsos et al. (1988) or Pukelsheim (1993). Elfving proposed a graphical method which is applied, especially for two parameters (see Appendix A),

\[ y = f(x)\alpha + g(x)\beta + \epsilon(x), \quad \sigma^2(x) = \sigma^2, \quad x \in X. \]

Let us suppose we have a design \( \xi \) with \( l \) points \( x_1, \ldots, x_l \), \( p_i = \frac{N_i}{N} \), so we have \( N_i \) observations at the point \( x_i \), \( N \) being the total number of observations. Therefore, we have \( l \) points of the form \( X_1 = (f(x_1), g(x_1)), \ldots, X_l = (f(x_l), g(x_l)) \). The mean of the observations at \( x_i \) is expressed as follows,

\[ \bar{y}_i = f(x_i)\alpha + g(x_i)\beta + \bar{\epsilon}_i, \quad \text{var}(\bar{\epsilon}_i) = \frac{\sigma^2}{N_i}. \]

Let us denote \( \eta_i = \sqrt{p_i} \bar{\epsilon}_i \), so \( \text{var}(\eta_i) = \frac{\sigma^2}{N} \). In order to simplify the notational aspects, we can suppose, without loss of generality, that \( \frac{\sigma^2}{N} = 1 \), so \( \text{var}(\eta_i) = 1 \). We will focus on estimators of the form

\[ \hat{\phi} = \sum_{i=1}^{l} a_i \bar{y}_i \]

where \( \bar{y}_i \) is the most efficient linear estimator of \( f(x_i)\alpha + g(x_i)\beta \). In order this estimator to be unbiased,

\[ E[\hat{\phi}] = \sum_{i=1}^{l} a_i \left[ f(x_i)\alpha + g(x_i)\beta \right] = c_1\alpha + c_2\beta, \quad \forall \alpha, \beta \in \mathbb{R}, \]

has to be verified, that is, the equation

\[ \sum_{i=1}^{l} a_i X_i = c. \]
has to be satisfied. The variance of the estimator is

$$\text{var}(\hat{\phi}) = \sigma^2 \sum_i \frac{a_i^2}{N_i} = \frac{\sigma^2}{N} \sum_i \frac{a_i^2}{p_i} = \sum_i \frac{a_i^2}{p_i},$$

We have to minimize

$$\min \sum_i \frac{a_i^2}{p_i}, \quad \sum p_i = 1.$$

By using Lagrange multipliers, we should calculate $p_1, \ldots, p_l, \lambda$ which minimize

$$\sum_i \frac{a_i^2}{p_i} - \lambda(1 - \sum p_i),$$

From that, it results,

$$\text{var}(\hat{\phi}) = \sum_i \frac{a_i^2}{p_i} = \sum_i \frac{a_i^2}{|a_i|} = \sum_i \sum_j |a_j| \frac{a_i^2}{|a_i|} = \sum_j |a_j| \sum_i \sum_j |a_j| = \left( \sum_i |a_i| \right)^2.$$

Therefore, is to calculate the coefficients $a_i$ which minimizes

$$\min \left\{ \sum_i |a_i| : \sum_{i=1}^l a_i X_i = c \right\} = \min \left\{ \sum_i |a_i| : \sum_{i=1}^l w_i [\text{sgn}(a_i) X_i] = t c \right\},$$

where

$$w_i = \frac{|a_i|}{\sum_j |a_j|}, \quad t = \frac{1}{\sum_j |a_j|}.$$ 

We are searching for the greatest value of $t$, so that, $tc$ belongs to the convex hull of $x_1, \ldots, x_l$, therefore

$$c^* = tc = \sum_i a_i^* X_i = \sum_i t a_i X_i,$$

is a convex combination of vectors $\pm X_1, \ldots, \pm X_l$. 
The variance reaches its minimum value when the vertex of $c^*$ results from the intersection of the vector $c$ (or its prolongation) with the convex hull formed by the vectors $\pm X_1, \ldots, \pm X_l$. Then variance of $\hat{\varphi}$ and the optimal design are

$$\begin{pmatrix} x_1 & \cdots & x_l \\ \frac{|a_1|}{\sum_j |a_j|} & \cdots & \frac{|a_l|}{\sum_j |a_j|} \end{pmatrix}, \quad \text{var}(\hat{\varphi}) = \frac{1}{t^2} = \frac{\|c\|^2}{\|c^*\|^2}.$$

Figure 1.1 represents the procedure. The intersection point of the vector $c$ (or its prolongation) with the boundary of the convex hull is located at the segment which connects two of the points $\pm X_1, \ldots, \pm X_l$ (in this case, $X_2$ and $-X_1$). Therefore,

$$c^* = a_1^*(-X_1) + a_2^*X_2, \quad \text{with} \quad a_1^* + a_2^* = 1.$$  

As the point $c^*$ can be expressed as

$$X_2 + \lambda(-X_1 - X_2) = \lambda(-X_1) + (1 - \lambda)X_2, \quad 0 \leq \lambda \leq 1,$$

it results that $a_1^* (= \lambda)$ is proportional to the distance $d_1$ from $c^*$ to $X_2$ and $a_2^*$ is proportional to the distance $d_2$ from $c^*$ to $X_1$. This procedure can be applied for any points belonging to the design space. Therefore, the convex hull of $f(\chi) \cup -f(\chi)$ ($\chi$ being the design space) and the intersection of the straight line defined by the vector $c$ with the boundary of the convex hull.

Figure 1.1: Graphical method proposed by Elfving
1.5 Equivalence theorem

Kiefer y Wolfowitz (1960) proved the equivalence between D-optimal and G-optimal criteria.

**Theorem. (TGE).** If $\sigma^2$ is constant, $\forall x \in X$, a design $\xi^*$ is D-optimal if, and only if, is G-optimal. That is, the following statements are equivalent:

1. $\det M(\xi^*) = \max \{\det M(\xi) : \xi \in \Xi\}$.

2. $\max_x f^t(x)M^{-1}(\xi^*)f(x) = \min_{\xi} \max_x f^t(x)M^{-1}(\xi)f(x)$.

Moreover, the last statement is equal to $m$ (the number of parameters of the model).

**Proof.** Let $\xi^*, \xi_1 \in \Xi$ be a D-optimal and any other design, respectively. If we define the design $\xi_2$ as a linear convex combination of the others, that is,

$$\xi_2 = (1 - \alpha)\xi^* + \alpha\xi_1.$$ 

Then

$$M(\xi_2) = (1 - \alpha)M(\xi^*) + \alpha M(\xi_1).$$

Furthermore, the function

$$-\log \det \left[(1 - \alpha)M(\xi^*) + \alpha M(\xi_1)\right],$$

is convex and differentiable, and has a minimum at $\alpha = 0$. Therefore,

$$0 \leq \left(-\frac{\partial}{\partial \alpha} \log \det \left[(1 - \alpha)M(\xi^*) + \alpha M(\xi_1)\right]\right)_{\alpha=0}. \quad (1.5)$$

We have that

$$\frac{\partial}{\partial \alpha} \left(-\log \det \left[(1 - \alpha)M(\xi^*) + \alpha M(\xi_1)\right]\right) = \text{tr} \left[M^{-1}(\xi_2) \cdot \frac{\partial}{\partial \alpha} M(\xi_2)\right]$$

$$= -\text{tr} \left[M^{-1}(\xi_2)(M(\xi_1) - M(\xi^*))\right].$$
Evaluating at $\alpha = 0$, the expression (1.5) is equal to

$$-\text{tr} \left[ M^{-1}(\xi^*) (M(\xi_1) - M(\xi^*)) \right] = -\text{tr} \left[ M^{-1}(\xi^*) (M(\xi_1) - I_m) \right]$$

$$= - \left( \text{tr} M^{-1}(\xi^*) M(\xi_1) - \text{tr} I_m \right) = m - \text{tr} \left[ M^{-1}(\xi^*) M(\xi_1) \right] \geq 0,$$

and this implies that

$$\text{tr} M^{-1}(\xi^*) M(\xi_1) \leq m.$$ 

Let $\xi_x$ be the one design supported at only one point $x$. Then,

$$\text{tr} M^{-1}(\xi^*) M(\xi_1) - m = \text{tr} M^{-1}(\xi^*) M(\xi_x) - m = \text{tr} M^{-1}(\xi^*) f(x) f^t(x) - m$$

$$= f^t(x) M^{-1}(\xi^*) f(x) - m \leq 0,$$

so

$$d(x, \xi^*) \leq m, \quad \forall x \in \chi.$$ 

On the other hand, for any design $\xi$,

$$\int_{\chi} d(x, \xi) \xi(dx) = \int_{\chi} f^t(x) M^{-1}(\xi) f(x) \xi(dx) =$$

$$\int_{\chi} \text{tr} [f^t(x) M^{-1}(\xi) f(x)] \xi(dx) =$$

$$\int_{\chi} \text{tr} [M^{-1}(\xi) f(x) f^t(x)] \xi(dx) = \text{tr} [M^{-1}(\xi) \int_{\chi} f(x) f^t(x) \xi(dx)] =$$

$$\text{tr} M^{-1}(\xi) M(\xi) = \text{tr} I_m = m.$$

Therefore,

$$\int_{\chi} d(x, \xi) \xi(dx) = m \leq \max_x d(x, \xi) \cdot \int_{\chi} \xi(dx) = \max_x d(x, \xi).$$

This has been proved for any design $\xi$, so we have

$$\max d(x, \xi^*) \geq m, \quad \forall x.$$
On the other hand, we proved that
\[ d(x, \xi^*) \leq m, \forall x, \]
so the maximum of all of them is less than \( m \). It is obvious that the minimum among all the maxima is reached at \( \xi^* \) and its value is \( m \). Let us suppose now that \( \xi^* \) is a G-optimal but not a D-optimal design. Therefore, for some design \( \xi_1 \),
\[ \text{tr}M^{-1}(\xi^*)M(\xi_1) - m > 0. \]
Moreover, \( M(\xi_1) \) can be expressed as a linear convex combination of \( n \leq \frac{m(m+1)}{2} + 1 \) matrices of rank 1
\[ M(\xi_1) = \sum_{i=1}^{n} \xi(x_i)f(x_i)f^t(x_i). \]
On the other hand,
\[
\text{tr}M^{-1}(\xi)M(\xi) = \sum_{i=1}^{n} \xi(x_i)\text{tr}M^{-1}(\xi^*)f(x_i)f^t(x_i)
= \sum_{i=1}^{n} \xi(x_i)f^t(x_i)M^{-1}(\xi^*)f(x_i) = \sum_{i=1}^{n} \xi(x_i)d(x_i, \xi^*).
\]
Since \( \xi^* \) is G-optimal,
\[ \max d(x, \xi^*) = m, \forall x, \]
so we have
\[ d(x, \xi^*) \leq m, \forall x. \]
This implies that
\[
\sum_{i=1}^{n} \xi(x_i)d(x_i, \xi^*) - m \leq \sum_{i=1}^{n} \xi(x_i)m - m = m \sum_{i=1}^{n} \xi(x_i) - m = 0,
\]
and this falls into contradiction since we have proved that
\[ \text{tr}M^{-1}(\xi^*)M(\xi_1) - m > 0. \]

**Definition.** Given a real, convex function \( \Phi \), which is defined in a convex
subset belonging to an Euclidean space, and given two points from this subset, 
\(u\) and \(v\), we define the directional derivative of \(\Phi\) along a given vector \(v\) at a 
given point \(u\) as

\[
\partial \Phi(u, v) = \lim_{\beta \to 0} \frac{\Phi[u + \beta v] - \Phi(u)}{\beta}.
\]

The drawback here is that, if \(u\) and \(v\) are information matrices, it cannot 
be assured that \(u + \beta v\) is another information matrix. For that reason, it is 
desirable to take

\[(1 - \beta)u + \beta v = u + \beta(v - u),\]

because it is an information matrix. Although this directional derivative goes 
along the direction of \(v - u\), we write it as

\[
\partial \Phi(u, v) = \lim_{\beta \to 0} \frac{\Phi[(1 - \beta)u + \beta v] - \Phi(u)}{\beta}.
\]

This directional derivative always exists due to the convexity of \(\Phi\). We will 
prove that the function

\[
\varphi : (0, 1) \rightarrow \mathbb{R}
\]

\[
\beta \rightarrow \frac{\Phi[(1 - \beta)u + \beta v] - \Phi(u)}{\beta}
\]

is an increasingly function in \((0,1)\). In that case, the limit always exists since 
the function is monotonic in \((0,1)\). Let us see it is increasingly:

If we take \(0 < \beta_1 < \beta_2 < 1\), then,

\[
\Phi[(1 - \beta_1)u + \beta_1 v] - \Phi(u) = \Phi[\frac{\beta_1}{\beta_2}(1 - \beta_2)u + \beta_2 v)
\]

\[
+(1 - \frac{\beta_1}{\beta_2} u - \Phi(u) \leq \frac{\beta_1}{\beta_2} \Phi[(1 - \beta_2)u + \beta_2 v] + (1 - \frac{\beta_1}{\beta_2})\Phi[u] - \Phi(u)
\]

\[
= \frac{\beta_1}{\beta_2} \left( \Phi[(1 - \beta_2)u + \beta_2 v] - \Phi(u) \right).
\]

Therefore,
\[
\frac{\Phi[(1 - \beta_1)u + \beta_1 v]}{\beta_1} \leq \frac{\Phi[(1 - \beta_2)u + \beta_2 v]}{\beta_2}.
\]

Applying L’Hôpital rule, if \( \Phi[(1 - \beta)u + \beta v] \) is differentiable with respect to \( \beta \),

\[
\partial \Phi(u, v) = \lim_{\beta \to 0} \frac{\partial \Phi[(1 - \beta)u + \beta v]}{\partial \beta} = \left( \frac{\partial \Phi[(1 - \beta)u + \beta v]}{\partial \beta} \right)_{\beta=0}.
\]

**Definition.** \( \xi^* \) is a local optimal if for any design \( \xi \), the function

\[
\Phi : [0, 1) \to \mathbb{R}
\]

\[
\beta \to \Phi[(1 - \beta)M(\xi^*) + \beta M(\xi)]
\]

has a local minimum at \( \beta = 0 \).

**Theorem.** Let \( \xi^* \) and \( \Phi \) be any design and a convex criteria function, respectively, so that

\[
\partial \Phi[M(\xi^*), M(\xi)] > -\infty, \quad \xi \in \Xi.
\]

The following statements are equivalent:

1. \( \xi^* \) is \( \Phi \)-optimal.
2. \( \xi^* \) is locally optimal.
3. \( \partial \Phi[M(\xi^*), M(\xi)] \geq 0, \quad \xi \in \Xi. \)

**Proof.** If \( \xi^* \) is a \( \Phi \)-optimal design, it will be locally optimal. Now, if \( \xi^* \) is locally \( \Phi \)-optimal, then for each \( \xi \in \Xi \), it exists \( \delta > 0 \) so that \( \forall \beta \in (0, \delta), \)

\[
\Phi[M(\xi^*)] \leq \Phi[(1 - \beta)M(\xi^*) + \beta M(\xi)].
\]

Dividing by \( \beta \):

\[
\frac{\Phi[(1 - \beta)M(\xi^*) + \beta M(\xi)] - \Phi[M(\xi^*)]}{\beta} \geq 0 \implies \partial \Phi[M(\xi^*), M(\xi)] \geq 0.
\]
In order to prove the last equivalence, let us suppose that $\xi^*$ is not $\Phi$-optimal. In that case, there will be a design $\xi$ so that $\Phi[M(\xi^*)] < \Phi[M(\xi)]$. Then $\forall \beta \in (0, 1)$,

$$\frac{\Phi[(1 - \beta)M(\xi^*) + \beta M(\xi)] - \Phi[M(\xi^*)]}{\beta} \leq \frac{(1 - \beta)\Phi[M(\xi^*)] + \beta \Phi[M(\xi)] - \Phi[M(\xi^*)]}{\beta} = \Phi[M(\xi)] - \Phi[M(\xi^*)].$$

And this falls into contradiction with the equivalence 3.

**Remark.** This theorem gives a general criterion to check if a design is or not $\Phi$-optimal, no matter if the criterion function is differentiable or not. But when it is differentiable, we can use the following theorem.

**Theorem.** If $\Phi$ is differentiable in a neighbourhood of $M(\xi)$, the following statements are equivalent:

1. $\xi^*$ is $\Phi$-optimal.

2. $f^t(x)\nabla \Phi[M(\xi^*)]f(x) \geq trM(\xi^*)\nabla \Phi[M(\xi^*)]$, $x \in \chi$.

3. $\min_x f^t(x)\nabla \Phi[M(\xi^*)]f(x) = \sum_{x \in \chi} f^t(x)\nabla \Phi[M(\xi^*)]f(x)\xi^*(x)$.

**Proof.** See Pázman (1986).

**Corollary 4.** Let the function $\Psi$ be

$$\Psi(x, \xi) = f^t(x)\nabla \Phi[M(\xi)]f(x) - trM(\xi)\nabla \Phi[M(\xi)].$$

It is satisfied that

1. $\Psi(x, \xi^*) \geq 0$, $\forall \xi \in \Xi$.

2. $\Psi(x, \xi^*) = 0$, $\forall x \in \chi$.

3. $\frac{\partial \Psi(x, \xi^*)}{\partial x} = 0$, $\forall x \in \chi \cap \Xi$. 


Remark. If $\xi^*$ is a D-optimal design, $\Phi_D(M(\xi^*)) = M^{-1}(\xi^*)$. Therefore, for any information matrix $N$,

$$\partial \Phi_D(M(\xi^*), N) = \text{tr}[\nabla \Phi_D(M(\xi^*)) (M(\xi^*) - N)] = \text{tr}[-M^{-1}(\xi^*)(N - M(\xi^*))]$$

$$= \text{tr}M^{-1}(\xi^*)M(\xi^*) - \text{tr}M^{-1}(\xi^*)N = m - \text{tr}M^{-1}(\xi^*)N \geq 0,$$

so for any information matrix $N$,

$$\text{tr}M(\xi^*)N \leq m.$$

By using a one-point design, $\xi_x$, we have

$$M_x = f(x)f^t(x) \quad \text{and} \quad N = \sum_x f(x)f^t(x)\xi(x).$$

It is obvious that

$$\text{tr}[M^{-1}(\xi^*) \sum_x f(x)f^t(x)\xi(x)] \leq m, \quad \xi \in \Xi,$$

satisfying the equivalence between,

$$\sum_x \text{tr}[M^{-1}(\xi^*)f(x)f^t(x)]\xi(x) \leq m, \quad \forall \xi \in \Xi,$$

and

$$\text{tr}[M^{-1}(\xi^*)f(x)f^t(x)] \leq m, \quad \forall x.$$

Therefore,

$$\text{tr}[M^{-1}(\xi^*)f(x)f^t(x)] = f^t(x)M^{-1}(\xi^*)f(x) \leq m, \quad \forall x.$$
The theory described above cannot be applied to any regression model. When the expected value of the response, \( \eta(x, \theta) \), is non-linear on \( \theta \), the model is called non-linear model. A feature common to all non-linear models is that the optimal design depends upon the value of the parameter \( \theta \). Since the purpose of the design is to estimate \( \theta \), the dependence of the design on the value of the parameter is unfortunate, but unavoidable for optimal designs with non-linear models. For that reason, it is necessary to use a prior estimator \( \theta^{(0)} \), called nominal value, which usually represents the best guess for the parameter \( \theta \) at the beginning of the experiment, and then to consider designs which minimizes the criterion function. The resulting design is called locally optimal design (Chernoff, 1953).

A sensitivity analysis is then convenient to evaluate the impact in guessing wrongly the nominal values of the parameters. Ratkowsky (1990) provided an introduction to non-linear regression models, with emphasis on the properties of models with one or more factors. Box and Lucas (1959) described the D-optimal designs for non-linear models but they did not mention that the D-optimality criterion or the Equivalence Theorem proves the optimality of the designs found. Atkinson and Fedorov (1975) described with details the compartmental model. Seber and Wild (1989) provided a complete overview of non-linear regression. More recently, Pázman (2002) gives a survey of optimal design of non-linear models with parameter constrains.

Let us suppose we have a non-linear model which belongs to an exponential parametric family of distributions of the form \( g(y|x, \theta) \). The Fisher information matrix

\[
I(x, \theta) = \left( E \left[ -\frac{\partial^2 \log g(y|x, \theta)}{\partial \theta_i \partial \theta_j} \right] \right)_{ij},
\]

is symmetric and positive definite. Besides, when we have a large enough number of observations, \( N \), the covariance matrix of the estimates is asymptotically proportional to the inverse of the Fisher information matrix.
**Definition.** The information matrix of a design $\xi$ for a non-linear model is defined as follows,

$$M(\xi, \theta) = \sum_{x \in \chi} \xi(x)I(x, \theta).$$

**Remark.** As mentioned above, this matrix depends on the parameters $\theta$. For this reason, nominal values have to be taken.

Most of the current approaches to design a study for a non-linear model simply linearize the non-linear mean function via a first-order Taylor’s approximation and work with the simplified model. For a nominal value of $\theta$, $\theta^{(0)}$,

$$E(y_x) = \eta(x, \theta) = \eta(x, \theta^{(0)}) + \sum_{i=1}^{m} \frac{\partial \eta(x, \theta)}{\partial \theta_i} \Big|_{\theta=\theta^{(0)}} (\theta_i - \theta_i^{(0)}).$$

(1.6)

Then, if

$$f_i(x) = \frac{\partial \eta(x, \theta)}{\partial \theta_i} \Big|_{\theta=\theta^{(0)}},$$

we have,

$$E(y_x - \eta(x, \theta^{(0)})) = \sum_{i=1}^{m} \frac{\partial \eta(x, \theta)}{\partial \theta_i} \Big|_{\theta=\theta^{(0)}} (\theta_i - \theta_i^{(0)}) = \sum_{i=1}^{m} f_i(x)(\theta_i - \theta_i^{(0)}),$$

and the linearization will have been carried out.

**Remark.** In the linearized model the differences $(\Delta \theta)_i = \theta_i - \theta_i^{(0)}$ are estimated. Due to this linearization, the estimators of $(\Delta \theta)_i$ are biased.
Chapter 2

Benign paroxysmal positional vertigo model

2.1 Introduction

First described by Bárány (1906), Benign Paroxysmal Positional Vertigo is the most common vestibular disorder leading to vertigo. These vestibular symptoms are precipitated when the orientation of the head or body is changed relative to gravity, provoking brief periods (2-3 minutes) of vertigo, imbalance, and nausea. These changes can occur during daily activities such as lying down in bed or reaching up to retrieve an object from a high shelf. Because people with Benign Paroxysmal Positional Vertigo (BPPV) often feel dizzy and unsteady when they tip their heads back to look up, is commonly called top-shelf vertigo (Squires et al., 2004).

Contrary to what is widely believed, this type of vertigo is caused by a disorder in an organ of the inner ear, called the semicircular canal, which regulates balance. Figure 2.1 illustrates the three canals and how they are arranged in a similar way to the three cartesian axes. Each canal is filled with a fluid called endolymph and contains motion sensors within the fluids. At the base of each canal, the bony region of the canal has a dilated sac at one end called the ampulla. Within the ampulla is a mound of hair cells and supporting cells called crista ampullaris. These hair cells are composed of many cilia and are embedded in a gelatinous structure called cupula. As the head rotates, the duct moves, but the endolymph lags behind. This deflects the cupula and bends the cilia within. The bending of these cilia alters an electric signal that is transmitted to the brain, which sends this information to the eyes, provoking the corresponding vestibular movement which helps us keep our balance.
When does BPPV occur? Semicircular canals are identified as the origin of BPPV. There, calcium carbonate particles ($CaCO_3$) called *otoliths*, which are normally affixed to the canal walls, are detached by the aforementioned head or body changes. This extra mass floating in the endolymph causes an abnormal movement of the cupula, since these particles displace more volume of endolymph than usual. The brain misinterprets this displacement and sends erroneous information to the eyes, provoking a characteristic ocular nystagmus and the subsequent vertigo.

Dix and Hallpike (1952) were pioneers in developing maneuvers which led to the detection of BPPV. These maneuvers consist of a series of consecutive head turns that trigger ocular responses in the patient, on the basis of which clinicians can determine whether a patient suffers from BPPV or not (see Appendix C). Rabbitt and Rajguru (2004) developed a model which calculates the volume of endolymph displaced when the Dix and Hallpike maneuver is put into practice. In this model, a mathematical approximation consisting of a curve crossing the different angular positions was used. This is a theoretical model never validated with real data as far as the authors know. In this work, the model is particularized to a specific real situation and an experimental plan is produced. In our case, a maneuver composed of two consecutive turns has been developed: a positive pitch (turning the head back) from the upright position, followed by a positive roll (turning the head right), as these are the most common two head movements that trigger the above-mentioned nystagmus. In order to reflect a more realistic situation, a cubic interpolation between points has been carried out. Another advantage of carrying out the cubic interpolation with respect to Rabbit model, is to obtain an analytical expression of the curve. This permits us to design statistical experiments aimed
2.2. DERIVATION OF A MODEL FOR THE MANEUVER

at deriving an optimal estimation of the unknown parameters of the model.

The model to be derived will be used to predict the endolymph volume displacement in response to a maneuver composed by two consecutive turns: a positive pitch (turning the head back) from the upright position followed by a positive roll (turning the head right). The standard definitions are used for head rotations: pitch rotates about an axis out the ear (y-axis), roll rotates about an axis out the eye (x-axis) and yaw rotates about an axis pointing out the top of the head (z-axis). We are only concerned with the horizontal canal since usually the particles are located inside that area (Martín-Sanz, 2006).

Therefore, considering $Q(\Delta t, \alpha, \theta)$ to be the volume of endolymph displaced after the first turn, where $\Delta t$ and $\alpha$ are the variables which represent duration and angle of each turn, respectively, our model is formulated as follows,

$$y = Q(\Delta t, \alpha, \theta) + \varepsilon, \quad \Delta t \in [0.5, 1.5], \quad \alpha \in [\pi/6, \pi/2]. \quad (2.1)$$

The error $\varepsilon$ follows a Gaussian distribution with mean 0 and constant variance $\sigma^2$, that is

$$E(y) = Q(\Delta t, \alpha, \theta) \quad \text{and} \quad \text{var}(y) = \sigma^2.$$
particle and endolymph and the gravity force $\vec{g}$ acting on the particle (see Figure 2.3). After considering the forces mentioned, as well as the simplifications corresponding to the fluid, the Navier-Stokes equations must be applied with the corresponding boundary and initial conditions (Hain et al., 2005). Thus, the equations which relate the volume flow rate of endolymph to the pressure and inertial forces at time $t$ are

$$\left\{\begin{array}{l}
\theta_1 \frac{\partial}{\partial t} Q(t, \alpha, \theta) + \theta_2 Q(t, \alpha, \theta) = F_i + F_n, \\
Q(0, \alpha, \theta) = 0.
\end{array}\right.$$ (2.2)

![Figure 2.3: Composition of forces acting on the endolymph](image)

The unknown parameters, $\theta^T = (\theta_1, \theta_2)$, are related to the damping and stiffness of the canal, respectively. The right side of the first equation is represented by two forces. The inertial forcing due to the angular acceleration of the head-fixed system relative to the ground-fixed inertial frame is

$$F_i = \int_0^{l_n} \rho (\ddot{\Omega}(t) \wedge \vec{R}(s)) \, ds,$$

where $\ddot{\Omega}(t) \wedge \vec{R}(s)$ stands for the tangential acceleration, $\ddot{\Omega}(t)$ being the angular acceleration of the head relative to the ground-fixed inertial frame resolved into the head-fixed frame and $\vec{R}(s)$ the vector running from the head-fixed coordinate system’s origin to the centerline of the canal. The parameterization of $\vec{R}(s)$ is made with respect to the arc length $s$ (also called *natural parameter*).
The head-fixed coordinate system was defined when the subject was in the upright position prior to movement of the head. The constant $\rho$ stands for the density of the canal and $l_n$ is the length covered by the otolith.

The second term, $F_n$, is the result of the interaction drag forces due to the particle moving relative to the fluid,

$$F_n = \frac{A_s}{A_e} N \left[ \frac{4}{3} a (\rho_s - \rho_e) (\vec{g} - \vec{\Omega}(t) \wedge \vec{R}(s)) \cdot \vec{n} + \frac{6}{a} \mu \left( \frac{\partial}{\partial t} Q(t, \alpha, \theta) \right) \right].$$

In this equation, $\vec{g}$ is the gravitational acceleration, $\vec{n}$ is the unit normal tangent vector to the canal centerline and $\dot{\xi}$ is the velocity of the particle. The constants $A_s$, $N$, $A_e$, $a$, $\rho_s$, $\rho_e$ and $\mu_e$ stand for frontal area of the particle, number of particles which are floating inside the canal, cross-sectional area of the canal, radius of the particle, density of the particle, density of the endolymph and endolymph viscosity, respectively.

The manner in which the maneuver has been carried out was established by the clinician. The angle of turn in each time $t$ is represented by $\Omega(t)$. The movements of the head determine the magnitude and direction of the vectors $\vec{\Omega}(t)$ and $\vec{g}$. It is important to note that the model equations refer to the non-inertial system, therefore, the linear and angular acceleration must be resolved into this system. This is done by using

$$\vec{\Omega}(t) = M(t) \vec{\Omega}_I(t),$$

where $\vec{\Omega}_I(t)$ is the angular acceleration referred to the inertial system (e.g. the clinician who makes the maneuvers) and $M(t)$ a rotation matrix. Since these maneuvers consist of a pitch (y-axis) followed by a roll movement (x-axis), the vector $\vec{\Omega}_I$ for the pitch is $(0, \Omega(t), 0)^T$ and for the roll it is $(\Omega(t), 0, 0)^T$.

The first rotation is a positive pitch of duration $\Delta t$ (starting at time $t = 0$).
The rotation matrix for the first turn is then,

\[ M_1(t) = \begin{pmatrix} \cos \Omega(t) & 0 & \sin \Omega(t) \\ 0 & 1 & 0 \\ -\sin \Omega(t) & 0 & \cos \Omega(t) \end{pmatrix} \quad t \leq \Delta t. \]

The second rotation is a positive roll of duration \( \Delta t \) (starting at time \( t = \Delta t \)). The rotation matrix for the second turn is then,

\[ M_2(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \Omega(t) & -\sin \Omega(t) \\ 0 & \sin \Omega(t) & \cos \Omega(t) \end{pmatrix} \begin{pmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{pmatrix}. \]

where \( \Delta t < t \leq 2\Delta t \). The angle of turn, \( \Omega(t) \), is calculated through cubic interpolation:

\[ \Omega(t) = \begin{cases} \Omega_1(t) = a + bt + ct^2 + dt^3 & t \leq \Delta t, \\ \Omega_2(t) = a' + b't + c't^2 + d't^3 & \Delta t < t \leq 2\Delta t. \end{cases} \]

For \( \Omega_1 \) and \( \Omega_2 \), the boundary conditions are, respectively:

\[ \begin{align*} 
\Omega_1(0) &= 0 \\
\Omega_1(\Delta t) &= \alpha \\
\dot{\Omega}_1(0) &= 0 \\
\dot{\Omega}_1(\Delta t) &= 0 \\
\end{align*} \quad \text{and} \quad \begin{align*} 
\Omega_2(\Delta t) &= 0 \\
\Omega_2(2\Delta t) &= \alpha \\
\dot{\Omega}_2(\Delta t) &= 0 \\
\dot{\Omega}_2(2\Delta t) &= 0 \\
\end{align*} \]

Once the former boundary conditions are imposed, \( \Omega(t) \) is written as follows,

\[ \Omega(t) = \begin{cases} \Omega_1(t) = \frac{3\alpha}{\Delta t^2} t^2 - \frac{2\alpha}{\Delta t^3} t^3, & t \leq \Delta t, \\ \Omega_2(t) = 5\alpha - \frac{12\alpha}{\Delta t} t + \frac{9\alpha}{\Delta t^2} t^2 - \frac{2\alpha}{\Delta t^3} t^3, & \Delta t < t \leq 2\Delta t. \end{cases} \]

In order to obtain the expressions for the forcing terms \( F_i \) and \( F_n \), we assume that the geometry of the posterior canal is described by a circle of radius \( r \).
Therefore,

\[ F_i = \int_0^{l_n} \rho (\ddot{\Omega}(t) \wedge \ddot{R}(s)) \, ds = \rho \, r \, \ddot{\Omega}_2(t) \sin \Omega_2(t) \, l_n \]

\[ = \rho \, r \, l_n \sin(5\alpha - \frac{12\alpha}{\Delta t} t + \frac{9\alpha}{\Delta t^2} t^2 - \frac{2\alpha}{\Delta t^3} t^3) \left( \frac{18\alpha}{\Delta t^2} - \frac{12\alpha}{\Delta t^3} t \right), \]

and

\[ F_n = \frac{a^2 N}{b^2} \left[ \frac{4}{3} a (\rho_s - \rho_e) \left( -r \sin \alpha \sin \Omega_2(t) \dot{\Omega}_2(t) - g \cos(\xi/r) \right) + \frac{6 \mu \xi}{a} \right] \]

\[ + \frac{a N 6 \mu_e}{\pi b^2 (b^2 - a^2)} \dot{Q}, \]

where \( A_s = \pi a^2, A_e = \pi b^2, \) \( b \) stands for the radius of the cross-sectional area of the canal and \( \dot{\xi} = 0.02 \, cm/s \) (Squires et al., 2004).

After imposing the initial condition, the solution of equations (2.2) is written as

\[ Q(t, \alpha, \theta) = \frac{\exp\{-\theta_2 t/c(\theta_1, t)\}}{c(\theta_1, t)} \int_0^t \exp\{\theta_2 s/c(\theta_1, s)\} F(s, \alpha) \, ds, \]

where

\[ F(t, \alpha) = F_i + F_n, \]

and

\[ c(\theta_1, t) = \frac{a N 6 \mu_e}{\pi b^2 (b^2 - a^2)} \dot{Q} + \theta_1. \]
2.3 A simple version of the model

An approximation to equation (2.2) is considered here replacing $F(s, \alpha)$ with $x$, assuming this variable is now under the control of the experiment. This should only be considered as a kind of toy example and will permit us to have an explicit solution of the differential equation and therefore, we will not have to resort to numerical algorithms to calculate the D-optimal design. The solution is expressed as,

$$Q(t, x, \theta) = \frac{x\left(1 + \exp\left\{-\frac{\theta_2 t}{\theta_1}\right\}\left(\theta_2 - 1\right)\right)}{\theta_2}.$$  \hspace{1cm} (2.3)

2.3.1 D-optimal design

Caratheodory’s theorem assures that for a model with two parameters, the D-optimal design is supported, at most, at three points. Nevertheless, frequently for models with two parameters, the design has only two support points.

Theorem. Let a model be

$$E(y) = x \mu(t, \theta), \quad t \in [0, A], \quad x \in [0, B].$$

The D-optimal design is $\xi_B^* \otimes \xi_t^*$, where

$$\xi_B^* = \left\{ x^* = B \right\} \quad \text{and} \quad \xi_t^* = \left\{ t_1^*, \ldots, t_m^* \atop p_1, \ldots, p_m \right\},$$

where $\otimes$ stands for the Kronecker product and $\xi_t^*$ is the D-optimal design for the model $\mu(t, \theta)$. Therefore,

$$\xi_t^* \otimes \xi_B^* = \left\{ (t_1^*, B), \ldots, (t_m^*, B) \atop p_1, \ldots, p_m \right\}.$$
2.3. A SIMPLE VERSION OF THE MODEL

Proof. Differentiating $Q(t, x, \theta)$ with respect to $\theta$, we get

$$
\frac{\partial}{\partial \theta} Q(t, x, \theta) = x \frac{\partial}{\partial \theta} \mu(t, \theta) \equiv xf(t, \theta).
$$

Moreover,

$$
det M(\xi) = \det \left[ \sum_{t,x} x^2 \xi(t, x) f(t, \theta) f^T(t, \theta) \right]
\leq \det \left[ B^2 \sum_t \sum_x \xi(t, x) f(t, \theta) f^T(t, \theta) \right] = B^{2m} \det \left[ \sum_t \xi_t(t) f(t, \theta) f^T(t, \theta) \right].
$$

where $\xi_t(t)$ is the marginal design of $\xi(\cdot, t)$ and $\sum_t \xi_t(t) f(t, \theta) f^T(t, \theta)$ is its information matrix whose determinant has to be maximized in order to give $\xi_t^*$. 

Therefore, we only need to calculate the D-optimal design for the model,

$$
\mu(t, \theta) = \frac{1 + \exp\{-\theta_2 t/\theta_1\} (\theta_2 - 1)}{\theta_2} \quad t \in [0, 1].
$$

So we have

$$
\xi_D^* = \begin{cases} 
  t^* & 1/2 \\
  1/2 & 1/2
\end{cases}, \quad t^* = \frac{(e^r - 1) - r}{r (e^r - 1)}, \quad r = \frac{\theta_2}{\theta_1}, \quad (2.4)
$$

where $t^*$ gives the maximum value of the determinant of the information matrix. Moreover, it may be proved numerically that the function

$$
\psi(t, \xi) = 2 - d(t, \xi),
$$

being,

$$
d(t, \xi) = f^t(t, \theta) M^{-1}(t, \xi) f(t, \theta),
$$
is greater than or equal to zero for all \( t \in [0, 1] \), so it satisfies the Equivalence Theorem and it is D-optimal.

**Remark.** When \( \theta_2 \) is much larger than \( \theta_1 \), the value of \( t^* \) is approximately zero. On the other hand, when the value of \( \theta_1 \) is much larger than \( \theta_2 \), \( t^* \) approaches 1/2. Besides, the solution \( t^* \) is convex and decreasing for \( r > 0 \).

**Example.** A D-optimal design for model (2.3) is calculated, with nominal values of the parameters \( \theta_1^{(0)} = 0.085 \) and \( \theta_2^{(0)} = 0.2 \). By using (2.4),

\[
\xi^*_D = \begin{pmatrix} 0.3199 & 1 \\ 1/2 & 1/2 \end{pmatrix}.
\]

Figure 2.4 shows the plot of the function \( d(t, \xi) \),

\[
d(t, \xi) = e^{-4.71t} \left( 8.57 + 8.5e^{4.71} + e^{2.35t} \left( -17.14 - 89.4t \right) + t(89.49 + 296.5t) \right),
\]

Directly from its definition and considering the D-optimal design, we can define the efficiency as a function depending on \( r \):

\[
\text{eff}_r(\xi^*_r) = \frac{(\det M(\xi^*_r, r))^{1/2}}{(\det M(\xi^*_r, r))^{1/2}}.
\]
2.3. A SIMPLE VERSION OF THE MODEL

The expression obtained is the following:

\[ eff_r(\xi^*_r) = \frac{e^{-(1+t^*_0)r}\left(-1 + e^{t^*_0r} + t^*_0 - t^*_0e^r\right)^{1/2}}{e^{-(1+t^*)r}\left(-1 + e^{t^*r} + t^* - t^*e^r\right)^{1/2}}, \]

and it is only valid for designs supported at two points. The value \( t^* \) stands for the D-optimal for a certain value \( r = \theta_2/\theta_1 \),

\[ t^* = \frac{(e^r - 1) - r}{r(e^r - 1)}. \]

Hence, for nominal values \( r_0 = 0.2/0.085 \) and \( t^*_0 = 0.3199 \).

Figure 2.5 shows how robust the design is with respect to the real value of the parameters. The closer the nominal values of the parameters are to their real values, the more efficient the design is. The robustness of the design is checked for different possible values of \( r \) between 1 and 3.5, where the efficiency remains over 90%.

Figure 2.5: Plot of the function \( eff_r(\xi^*_r) \)
2.3.1 c-optimal designs

With nominal values $\theta_1^{(0)} = 0.2$ and $\theta_2^{(0)} = 0.085$ and using as function $f(t)$ the partial derivatives of the function given by (2.3), the Elfving’s set is obtained by plotting the convex hull of $\{f(t) \cup -f(t)\}$, $t \in [0, 1]$ (see Figure 2.6).

Figure 2.6 also illustrates how the convex hull has one tangent line (taking its symmetry into consideration), so when the straight line resulting from the prolongation of the vector $c$ cuts that tangent, the c-optimal design is always composed by the same points, $t_0 = 0.257$ and 1. In any case, the weight for these points depends on the linear combination of the parameters $c^T \theta$, that is, it depends on the vector $c$.

In order to estimate separately the parameters $\theta_1$ and $\theta_2$, it must be set $c^T = (1, 0)$ and $c^T = (0, 1)$, calling the resulting optimal designs, $c_{\theta_1}$ and $c_{\theta_2}$-optimal designs, respectively. The tangential points $t_0$ and 1 (extreme points of the tangential segment) define the support points for the $c_{\theta_1}$ and $c_{\theta_2}$-optimal designs. The cut points of this straight line with both axes, $c^*_{\theta_1} = (x^*, 0)$ and $c^*_{\theta_2} = (0, y^*)$, define the weights for the $c_{\theta_1}$ and $c_{\theta_2}$-optimal designs, respectively. Therefore, to estimate $\theta_1$ the same procedure as before must be followed.

The straight line defined by the vector $c^T = (1, 0)$ cuts the boundary of the set $G$ at $(x^*, 0)$ which can be obtained as a linear convex combination of the points $f(t_0)$ and $f(1)$, so the support points of the c-optimal design are $t_0 = 0.257$ and 1. The weights of the design are the weights in the convex combination, that is:

$$
\xi^*_{c_{\theta_1}} = \begin{pmatrix} 0.257 & 1 \\ 0.2 & 0.8 \end{pmatrix}.
$$

It follows that the c-optimal design to estimate the parameter $\theta_2$ is of the form

$$
\xi^*_{c_{\theta_2}} = \begin{pmatrix} 0.257 & 1 \\ 0.4 & 0.6 \end{pmatrix}.
$$
2.3. A SIMPLE VERSION OF THE MODEL

Since the shape of the convex hull does not change for different nominal values of $r = \theta_2/\theta_1$, the two-point c-optimal design is always supported at the points $t_0 = 0$ and certain $t$ obtained by solving numerically the equation

$$e^{-rt} + e^{-r}(1 + r - rt) + 2(rt - 1) = 0.$$ 

One important aspect to bear in mind is to check the efficiency of the D-optimal design with respect to the c-optimal design.

The formula of the efficiency

$$\text{eff}_c(\xi^*_D) = \frac{c^T M^{-1}(\xi_c)c}{c^T M^{-1}(\xi^*_D)c},$$  \hspace{1cm} (2.5)$$

provides a way to see how good the D-optimal design is at estimating each of the parameters. In the case of $\theta_1$, $c^T = (1, 0)$, the efficiency is around 75%, while for $\theta_2$, $c^T = (1, 0)$, we have a efficiency around 90%.

---

**Figure 2.6: Elfving’s set.** The functions $f(t)$ and $-f(t)$ are represented by the continuous and dashed curves, respectively.
2.4 Optimal experimental designs

Optimal designs for the model given in (2.1) are computed, with nominal values of the parameters $\theta_1^{(0)} = 0.85$ and $\theta_2^{(0)} = 0.2$ (Rabbitt, 1999). It is assumed that six particles are detached from the canal wall (Rajguru and Rabbitt, 2007). On the other hand, it is known that $A_s/A_e \approx 10^{-4}$ and that, approximately, the values of $r$ and $l_n$ are 0.1 cm and 0.5 cm, respectively (Rajguru et al., 2005). The values used for the constants are listed in Table 2.1. Figure 2.7 shows the plot of the function given by the above mentioned $Q(\Delta t, \alpha, \theta)$. This is plotted as a function of $\Delta t$ and $\alpha$ for the given nominal values. The information matrix is expressed as

$$M(\xi, \theta) = \sum_{\Delta t, \alpha \in \chi} \xi(z) f(\Delta t, \alpha, \theta) f^T(\Delta t, \alpha, \theta),$$

where

$$f^T(\Delta t, \alpha, \theta) = (f_1(\Delta t, \alpha, \theta), f_2(\Delta t, \alpha, \theta)) = \left( \frac{\partial Q(\Delta t, \alpha, \theta)}{\partial \theta_1}, \frac{\partial Q(\Delta t, \alpha, \theta)}{\partial \theta_2} \right).$$

Figure 2.7: Plot of the function $\eta(\Delta t, \alpha, \theta)$ for given nominal values

The two components are written as
$f_1(\Delta t, \alpha, \theta) =$

$$
\left( \frac{\partial}{\partial t} \exp\left\{ -\theta_2 \frac{\Delta t}{c(\theta_1, \Delta t)} \right\} \right) \cdot \int_0^{\Delta t} \exp\{\theta_2 s/c(\theta_1, s)\} F(s, \alpha) ds + 
\exp\left\{ -\theta_2 \frac{\Delta t}{c(\theta_1, \Delta t)} \right\} \int_0^{\Delta t} \frac{\theta_2 s}{c^2(\theta_1, s)} \exp\{\theta_2 s/c(\theta_1, s)\} F(s, \alpha) ds.
$$

and

$f_2(\Delta t, \alpha, \theta) =$

$$
- \frac{\Delta t}{c^2(\theta_1, \Delta t)} \exp\left\{ -\theta_2 \frac{\Delta t}{c(\theta_1, \Delta t)} \right\} \int_0^{\Delta t} \exp\{\theta_2 s/c(\theta_1, s)\} F(s, \alpha) ds + 
\frac{1}{c(\theta_1, \Delta t)} \exp\left\{ -\theta_2 \frac{\Delta t}{c(\theta_1, \Delta t)} \right\} \int_0^{\Delta t} \frac{s}{c(\theta_1, s)} \exp\{\theta_2 s/c(\theta_1, s)\} F(s, \alpha) ds.
$$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_s$ : Frontal area of the particle</td>
<td>$3.14 \times 10^{-4} \text{ cm}^2$</td>
</tr>
<tr>
<td>$\rho$ : density of the canal</td>
<td>$1.0 \text{ g cm}^{-3}$</td>
</tr>
<tr>
<td>$\rho_s$ : density of the particle</td>
<td>$2.7 \text{ g cm}^{-3}$</td>
</tr>
<tr>
<td>$\rho_e$ : density of the endolymph</td>
<td>$1.0 \text{ g cm}^{-3}$</td>
</tr>
<tr>
<td>$\mu_e$ : viscosity of the endolymph</td>
<td>$8.5 \times 10^{-3} \text{ dyn cm s}^{-1}$</td>
</tr>
<tr>
<td>$g$ : gravitational acceleration</td>
<td>$981 \text{ cm s}^{-2}$</td>
</tr>
</tbody>
</table>

Table 2.1: Physical parameters

2.4.1 D-optimal design for a discrete design space

D-optimal design for model (2.1) are obtained in order to estimate simultaneously the parameters, $\theta_1$ and $\theta_2$. Points ($\alpha, \Delta t$) are taken from the design space $[0.5, 1.5] \times [\pi/6, \pi/2]$. Since it is more realistic to ask the clinician to carry out the maneuvers for typical values of angles and times, a finite design space is being considered,

$$
\Delta t \in \{0.5, 1, 1.5\} \quad \text{and} \quad \alpha \in \{\pi/6, \pi/4, \pi/3, 5\pi/12, \pi/2\}.
$$
Taking this finite set into consideration and maximizing the determinant of the information matrix, the optimal design is

\[
\xi_1^* = \left\{(1.5, \pi/6), (0.5, \pi/4), (0.5, \pi/2)\right\}.
\]

Table 2.2 shows that all the values of the generalized variance are smaller than the number of parameters, and therefore the design obtained satisfies numerically the Equivalence Theorem. The design indicates the duration and angle of the maneuvers to be carried out, but not how many observations the sample will have. That will be determined by the researcher through other means (for instance, the budget). The weights give us the proportion of different maneuvers to be performed. For instance, if the sample contains \( N = 100 \) observations, the clinician will have to repeat the maneuver consisting of a positive 90° pitch from the upright position followed by a positive 90° roll, both with a duration of half a second, 37 times.

\[
\begin{array}{cccccc}
\pi/6 & \pi/4 & \pi/3 & 5\pi/12 & \pi/2 \\
0.5 & 1.96 & 2 & 1.24 & 1.53 & 2 \\
1 & 1.93 & 1.34 & 0.19 & 0.45 & 0.4 \\
1.5 & 2 & 1.38 & 0.06 & 0.29 & 0.24 \\
\end{array}
\]

Table 2.2: Values of the generalized variance

As mentioned above, the goodness of any design \( \xi \) is measured by its efficiency. We will see how efficient the D-optimal design \( \xi_1^* \) is by checking the value of

\[
\text{eff}_\theta(\xi_1^*) = \frac{\Phi_D[M(\xi_{\theta^*}, \theta)]}{\Phi_D[M(\xi_1^*, \theta)]} = \frac{[\det M(\xi_{\theta^*}, \theta)]^{-1/m}}{[\det M(\xi_1^*, \theta^{(o)})]^{-1/m}},
\]

where \( \xi_{\theta^*} \) is the D-optimal design calculated for a possible real value of the parameter \( \theta \). This efficiency shows how robust the design is with respect to the true (unknown) value of the parameters. To check the robustness of a design, we want to check how the quality of the estimation would be affected by a wrong choice of the nominal value. The efficiency can sometimes be multiplied by 100 and be reported in percentage terms. If, for instance, we
take as nominal value $\theta^{(0)} = 0.5$, the true value being $\theta = 0.6$ and the efficiency being 50%, then the design $\xi^*_\theta(0)$ would need to double the total number of observations to perform as well as the optimal design calculated with the true value $\theta = 0.6$. Thus, our design would not be very robust. Table 2.3 illustrates the sensitivity of the D-optimal design with respect to the choice of the parameters $\theta_1$ (horizontal values) and $\theta_2$ (vertical values), i.e., how robust the design is with respect to the true values of those parameters. As can be observed, small variations of the nominal values do not affect the quality of the estimation much.

\[
\begin{array}{ccccc}
0.1 & 0.7 & 0.85 & 0.9 & 2 \\
0.015 & 37 \% & 98\% & 98\% & 78\% & 28\% \\
0.2 & 35\% & 97\% & 100\% & 80\% & 31\% \\
0.9 & 29\% & 84\% & 90\% & 80\% & 30\%
\end{array}
\]

Table 2.3: Values of the efficiency

### 2.4.2 D-optimal design for a continuous design space

If instead of the finite set

$$\chi = \{\pi/6, \pi/4, \pi/3, 5\pi/12, \pi/2\} \times \{0.5, 1, 1.5\},$$

we would have chosen a continuous design space $\chi = [0.5, 1.5] \times [\pi/6, \pi/2]$, the D-optimal design obtained would have been

$$\xi^*_2 = \begin{Bmatrix}
(1.5, 0.55)
(0.5, 0.72)
(0.5, \pi/2)
0.47
0.16
0.37
\end{Bmatrix}.$$  

As can be observed, the results would have been very similar to $\xi_1$, having an efficiency of around 95%. But in practice, the clinician cannot carry out any turns that are smaller than 15° or shorter than half a second. Figure 2.8 shows how the generalized variance function for a continuous design space satisfies the Equivalence Theorem, that is, for all the values of $\Delta t$ and $\alpha$ within the space design, the values of the generalized variance must be lower or equal to
the number of parameters to estimate. At the points of the design, the value of the function must be equal to the number of parameters.

![Figure 2.8: Plot of the generalized variance function for a continuous design space](image)

### 2.4.3 c-optimal designs for a discrete design space

Considering the same finite set as before, the interest is again in estimating one parameter, but now both of them are unknown. Elfving’s set is displayed in Figure 2.9. In order to estimate the parameter $\theta_1$, let $c^T = (1, 0)$. The straight line defined by this vector cuts the boundary of the Elfving’s set at point $(x^*, 0)$ and it defines the $c_{\theta_1}$-optimal design to estimate the parameter $\theta_1$. This point is expressed as a convex combination of the points $-f(\Delta t_1, \alpha_1) = -f(0.5, 5\pi/12)$ and $-f(\Delta t_2, \alpha_2) = -f(0.5, \pi/2)$ with weights $p_1 = 0.86$ and $p_2 = 0.14$, respectively. Therefore, the optimal design is

$$\xi^*_3 = \begin{pmatrix} 0.86 \\ 0.14 \end{pmatrix}.$$  

On the other hand, to estimate the parameter $\theta_2$, let $c^T = (0, 1)$. The straight line defined by this vector cuts the boundary of the Elfving’s set at point $(0, y^*)$, and it defines the $c_{\theta_2}$-optimal design to estimate the parameter $\theta_2$. This point is expressed as a convex combination of points $f(\Delta t_1, \alpha_1)$ and $-f(\Delta t_4, \alpha_4) = -f(1.5, \pi/6)$ with weights $p_1 = 0.3$ and $p_2 = 0.7$, respectively. Therefore, the optimal design is

$$\xi^*_4 = \begin{pmatrix} 0.3 \\ 0.7 \end{pmatrix}.$$
Finally, to obtain the c-optimal design for the estimation of $\theta_1 + \theta_2$ we use $c^T = (1, 1)$ to define the point $c^*$. This point is expressed as a convex combination of the points $-f(\Delta t_4, \alpha_4) = -f(1.5, 5\pi/12)$ and $-f(\Delta t_3, \alpha_3) = -f(1, \pi/6)$ with weights $p_1 = 0.8$ and $p_2 = 0.2$, respectively. So the c-optimal design is now

$$
\xi^*_c = \begin{cases} 
(1.5, 5\pi/12) & (1, \pi/6) \\
0.8 & 0.2 
\end{cases}.
$$

One important aspect is to check the efficiency of the D-optimal design ($\xi^*_D$) with respect to the c-optimal design ($\xi^*_c$). The formula of the efficiency,

$$
\text{eff}_c(\xi^*_D) = \frac{c^T M^{-1}(\xi^*_c, \theta^{(0)}) c}{c^T M^{-1}(\xi^*_D, \theta^{(0)}) c},
$$

provides a way to see how good the D-optimal design is at estimating each of the parameters. In the case of $\theta_1$, the D-optimal design $\xi^*_1$ and c-optimal design $\xi^*_3$ are compared and the efficiency is around 75%, while for $\theta_2$, $\xi^*_1$ and $\xi^*_4$ are compared, having an efficiency of around 85%. These results show the D-optimal design is more efficient for estimating $\theta_2$ than for estimating $\theta_1$, that is, with this design, the test power for testing \( H_0 : \theta_2 = 0 \) will be greater than the test power for \( H_0 : \theta_1 = 0 \).

![Figure 2.9: Elfving’s set for a discrete design space](image)
2.4.4 c-optimal designs for a continuous design space

For a continuous design space, the Elfving’s set obtained is shown in Figure 2.10 by plotting the convex hull of the surface \( \{ f(\Delta t, \alpha, \theta(0)) \cup -f(\Delta t, \alpha, \theta(0)) \} \). In this case \( t \in [0.5, 1.5] \times [\pi/6, \pi/2] \). With the help of the method proposed by López-Fidalgo and Rodríguez-Díaz (2004), we will calculate the c-optimal design to estimate the parameters \( \theta_1, \theta_2 \) and an example of a linear combination of them, \( \theta_1 + \theta_2 \).

The straight line defined by the vector \( c^T = (1, 0) \) cuts the Elfving’s set at point \((x^*, 0) = (0.144, 0)\) and it defines the \( c_{\theta_1} \)-optimal design to estimate the parameter \( \theta_1 \). This point is expressed as a convex combination of \( -f(\Delta t_2, \alpha_2) = -f(0.49, 1.29) \) and \( -f(\Delta t_3, \alpha_3) = -f(0.5, 1.57) \) with weights \( p_1 = 0.8 \) and \( p_2 = 0.2 \), respectively. Therefore, the optimal design is

\[
\xi_6^* = \begin{cases} 
(0.49, 73.9^\circ) & (0.5, 89^\circ) \\
0.8 & 0.2 
\end{cases}
\]

The straight line defined by vector \( c^T = (0, 1) \) cuts the boundary of the Elfving’s set at point \((0, y^*) = (0, 0.057)\) and it defines the \( c_{\theta_2} \)-optimal design to estimate the parameter \( \theta_2 \). This point is expressed as a convex combination of \( -f(\Delta t_1, \alpha_1) = -f(1.5, 0.49) \) and \( -f(\Delta t_3, \alpha_3) = -f(0.49, 1.29) \) with weights \( p_1 = 0.75 \) and \( p_2 = 0.25 \), respectively. Therefore, the optimal design is

\[
\xi_7^* = \begin{cases} 
(0.49, 73.9^\circ) & (1.5, 28^\circ) \\
0.75 & 0.25 
\end{cases}
\]

The efficiency of the D-optimal design for estimating the parameter \( \theta_1 \) is around 75%, while for \( \theta_2 \), it is around 90%. Finally, to obtain the c-optimal design for the estimation of \( \theta_1 + \theta_2 \), vector \( c^T = (1, 1) \) is considered. In this case, the straight line defines the point \( c^* = (0.084, 0.084) = -f(1.3, 0.6) \), so the c-optimal design is supported at one point:

\[
\xi_8^* = \begin{cases} 
(1.3, 34^\circ) \\
1 
\end{cases}
\]
As we can observe, the results obtained are quite similar to the case concerning the discrete design space, except for the estimation of $\theta_1 + \theta_2$, where for the continuous case the $c$-optimal design is only supported at one point, although in both cases the angle of turn is similar. If the number of support points in a $c$-optimal design is less than the number of parameters, this design allows the computation of the maximum likelihood estimate of this linear combination. But, in this case, not all the parameters are identifiable individually, that is, some of them cannot be estimated.

![Elfving's set for a continuous design space](image)

**Figure 2.10:** Elfving’s set for a continuous design space

### 2.5 Integral models

These models are frequently used in Biology to calculate morphologic or anatomic parameters in animals. In this cases it is not possible to find an explicit primitive function of them, therefore the differential equation cannot be solved analytically. In this section, exponential models with one and two parameters are considered.

#### 2.5.1 One-parameter model

Let us consider the model:

$$E[y(t, \theta)] = \eta(t, \theta), \quad var(y) = \sigma^2, \quad t \in \chi = [0, 1], \quad \theta \in \mathbb{R}.$$
where the function $\eta(t, \theta)$ can be expressed in any of these two forms:

$$
\eta(t, \theta) = \int_0^t e^{\theta s^2} ds \quad \text{or} \quad \frac{\partial \eta(t, \theta)}{\partial t} = e^{\theta t^2}, \quad \text{with} \quad \eta(0, \theta) = 0.
$$

The optimal design is a one-point design

$$
\xi^* = \left\{ \begin{array}{c} t^* \\ 1 \end{array} \right\},
$$

and the information matrix

$$
M(\xi) = f^2(t) \in \mathbb{R},
$$

where

$$
f(t) = \frac{\partial}{\partial \theta} \left( \int_0^t e^{\theta s^2} ds \right) = \int_0^t s^2 e^{\theta s^2} ds \in \mathbb{R}.
$$

In order to find the optimal design we will have to calculate $t^*$ such that

$$
t^* = \arg \max_{t \in \chi} f^2(t).
$$

Since the function $f(t)$ always takes positive values,

$$
t^* = \arg \max_{t \in \chi} f(t).
$$

Besides, $f(t)$ is increasing in its domain. $t^*$ is then the maximum over the set $\chi$. In this case, $t^* = 1$, so the optimal design is

$$
\xi^* = \left\{ \begin{array}{c} 1 \\ 1 \end{array} \right\}.
$$
2.5.2 Two-parameter model

Let us consider the model:

$$E[y(t, \theta)] = \eta(t, \theta), \quad \text{var}(y) = \sigma^2, \quad t \in \chi = [0, 1], \quad \theta^T \in \mathbb{R}^+ \times \mathbb{R}^+,$$

where the function $\eta(t, \theta)$ can be expressed in any of these two forms:

$$\eta(t, \theta) = \int_0^t \theta_1 e^{-\theta_2 s^2} \, ds \quad \text{or} \quad \frac{\partial \eta(t, \theta)}{\partial t} = \theta_1 e^{-\theta_2 t^2}, \quad \text{with} \quad \eta(0, \theta) = 0.$$

The information matrix is expressed as

$$M(\xi) = \sum_{t \in \chi} \xi(t) f(t) f^T(t),$$

being

$$f^T(t) = (f_1(t), f_2(t)) = \left( \frac{\partial \eta(t, \theta)}{\partial \theta_1}, \frac{\partial \eta(t, \theta)}{\partial \theta_2} \right). \quad (2.8)$$

The first component is,

$$f_1(t) = \frac{\partial}{\partial \theta_1} \left( \int_0^t \theta_1 e^{-\theta_2 s^2} \, ds \right) = \int_0^t e^{-\theta_2 s^2} \, ds.$$

Since we cannot obtain a primitive of $e^{-\theta_2 s^2}$, we must write $f_1(t)$ as a normal cumulative distribution function. As we have

$$\int_0^t e^{-\theta_2 s^2} \, ds = \sqrt{\frac{\pi}{\theta_2}} \int_0^t \frac{1}{\sqrt{2\pi \sqrt{1/2\theta_2}}} \exp \left\{ -\frac{1}{2} \left( \frac{s}{\sqrt{1/2\theta_2}} \right)^2 \right\} \, ds,$$

the integrand is the density probability function of $\mathcal{N}(0, \sqrt{1/2\theta_2})$. Therefore, $f_1(t)$ can be expressed as

$$f_1(t) = \sqrt{\frac{\pi}{\theta_2}} \left( \Phi(t\sqrt{2\pi\theta_2}) - \frac{1}{2} \right),$$
Φ(t) being the standard normal cumulative distribution function. The second component, \( f_2(t) \), is expressed as a gamma cumulative distribution function:

\[
f_2(t) = \int_0^t -\theta_1 s^2 e^{-\theta_2 s^2} \, ds = -\theta_1 \int_0^t s^2 e^{-\theta_2 s^2} \, ds.
\]

By taking \( s^2 = y \), we have

\[
\int_0^t s^2 e^{-\theta_2 s^2} \, ds = \int_0^{t^2} \frac{1}{2} y^{1/2} e^{-\theta_2 y} \, dy,
\]

which is written as

\[
\int_0^{t^2} \frac{1}{2} y^{1/2} e^{-\theta_2 y} \, dy = -\frac{\theta_1}{2} \frac{\Gamma(3/2)}{\theta_2^{3/2}} \int_0^{t^2} \frac{1}{\Gamma(3/2)} \theta_2^{3/2} y^{3/2-1} e^{-\theta_2 y} \, dy.
\]

Then,

\[
f_2(t) = -\frac{\theta_1}{2} \frac{\Gamma(3/2)}{\theta_2^{3/2}} G_a(t^2) = -\frac{\sqrt{\pi}}{4} \frac{\theta_1}{\theta_2^{3/2}} G_a(t^2),
\]

\( G_a(t) \) being the cumulative distribution function of the gamma distribution \( \Gamma(\theta_2, 3/2) \). Since the model is linear in \( \theta_1 \), the D-optimal design does not depend on this parameter. The information matrix is then

\[
I(t) = f(t) f^t(t) = \begin{pmatrix}
\frac{\pi}{\theta_2^2} g_1^2(t) & -\frac{\pi \theta_1}{\theta_2^2} g_1(t) g_2(t) \\
-\frac{\pi \theta_1}{\theta_2^2} g_1(t) g_2(t) & \frac{\pi \theta_1^2}{\theta_2^2} g_2^2(t)
\end{pmatrix},
\]

where

\[
g_1(t) = \Phi(t\sqrt{2\pi \theta_2}) - 1/2 \quad \text{and} \quad g_2(t) = \frac{1}{4} G(t^2).
\]

**Example 1.** Most of the times the D-optimal design generated to estimate
the parameters \( \theta_1 \) and \( \theta_2 \) is a two-point design of the form

\[
\xi^*_D = \left\{ \begin{array}{l}
t^*_1 \\
1/2 \\
t^*_2 \\
1/2
\end{array} \right\}.
\]

It is only necessary to take a nominal value for \( \theta_2 \); say \( \theta_2 = 0.085 \). The D-optimal design is then supported at the points \( t^*_1 = 0.026 \) and \( t^*_2 = 0.63 \),

\[
\xi^*_D = \left\{ 0.026 \\
1/2 \\
0.63 \\
1/2 \right\}.
\]

and, as we can see in Figure 2.11, the generalized variance function

\[
\psi(t) = 11, 2 \left( Erf(3, 43t) \right)^2 - 19, 16 Erf(3, 4t) \Gamma(0, 085, 0, 2t^2/3) +
\]

\[
+ 9, 95 \left( \Gamma(0, 085, 0, 2t^2/3) \right)^2,
\]

satisfies the Equivalence theorem, where

\[
Erf(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-s^2} ds
\]

is the Gauss error function and

\[
\Gamma(a, z_0, z_1) = \int_{z_0}^{z_1} s^{a-1} e^{-s} ds
\]

being the Incomplete Gamma function. Table 2.4 illustrates the sensibility of the D-optimal design with respect to the nominal value of the parameter \( \theta_2 \).

<table>
<thead>
<tr>
<th>( \theta_2^{(0)} )</th>
<th>( eff_{D}(\xi) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>99.95</td>
</tr>
<tr>
<td>0.1</td>
<td>99.64</td>
</tr>
<tr>
<td>0.09</td>
<td>99.95</td>
</tr>
<tr>
<td>0.08</td>
<td>99.95</td>
</tr>
<tr>
<td>0.04</td>
<td>99.95</td>
</tr>
</tbody>
</table>

Table 2.4: Efficiency of the D-optimal design for example 1
Example 2. Taking as a nominal value $\theta_2^{(0)} = 10$, the two-point D-optimal design obtained is supported at $t_1^* = 0.8467$ and $t_2^* = 1$:

$$\xi_D^* = \begin{cases} 
0.8467 & 1 \\
1/2 & 1/2 
\end{cases}.$$

We observe in the Figure 2.12 how the generalized variance, 

$$\psi(t) = 25, 5 \left( Erf \left( \frac{t}{\sqrt{10}} \right) \right)^2 - 6, 97.10^9 Erf \left( \frac{t}{\sqrt{10}} \right) \Gamma(10, 0, 2t^2/3)$$

$$+ 7, 7.10^{17} \left( \Gamma(10, 0, 2t^2/3) \right)^2,$$

satisfies the Equivalence Theorem. Table 2.5 shows the sensibility of the D-optimal design with respect to the nominal values of the parameter $\theta_2$.

If we take the same nominal values as in example 2, that is, $\theta_1^{(0)} = 1$ and $\theta_2^{(0)} = 0.085$ and use as function $f(t)$ the one given in (2.8), we obtain the set represented on Figure 2.13, where the functions $f(t)$ and $-f(t)$ are represented by the continuous and dashed curve, respectively.

If we are only interested in estimating the parameters $\theta_1$ and $\theta_2$, we must set $c^\tau = (1, 0)$ and $c^\rho = (0, 1)$, calling the resulting optimal design, $c_{\theta_1}$ and $c_{\theta_2}$.
optimal design, respectively. The tangential points define the support points for the $c_{\theta_1}$ and $c_{\theta_2}$-optimal designs. The cut points of this straight line with both axes, $c_{\theta_1}^* = (x^*, 0)$ and $c_{\theta_2}^* = (0, y^*)$, define the weights for the $c_{\theta_1}$ and $c_{\theta_2}$-optimal designs, respectively. Therefore, to estimate $\theta_1$ we will proceed as follows:

Point $(x^*, 0)$ can be obtained as a linear convex combination of the points $f(t_1)$ and $f(t_0)$. Therefore, the support points of the $c$-optimal design will be $t_1 = 0.017$ and $t_0 = 0.5$. The weights of the design are the weights in the convex combination

$$
\xi_{c_{\theta_1}}^* = \begin{cases} 
0.017 & 0.5 \\
0.64 & 0.36
\end{cases}.
$$

It follows that the $c$-optimal design to estimate the parameter $\theta_2$ will be

$$
\xi_{c_{\theta_2}}^* = \begin{cases} 
0.017 & 0.5 \\
0.94 & 0.06
\end{cases}.
$$

In order to estimate the rest of the linear combinations, it is enough to calculate the design by setting $c^T = (a, 1)$. Due to symmetry of the convex hull, it is enough to consider the vector on the first two quadrants. As we can observe in the Figure 2.13, for different values of $a$ we have a different two-point design.

If $a \in (f_1(t_1)/f_2(t_1), \infty)$, the $c$-optimal design will be supported at $\{t_1, t_0\}$. In case $a \in (f_1(1)/f_2(1), f_1(t_2)/f_2(t_2))$, being $t_2 = 0.067$, the support points will be $\{1, t_2\}$. And finally, if $a \in (-\infty, f_1(t_0)/f_2(t_0)]$, the $c$-optimal design will be supported again at $\{t_1, t_0\}$. In any case, the weight for the first point of the design will be given by $p = 0.0064 + 0.012/(0.04 + 0.313a)$. It becomes obvious that $1-p$ will be the weight for the second point.

In order to see how good the D-optimal design is for estimating each of the parameters $\theta_1$ and $\theta_2$, we must resort to the formula of the efficiency

$$
eff f_c(\xi_D^*) = \frac{c^T M^{-1}(\xi_c) c}{c^T M^{-1}(\xi_D^*) c}.$$
If we take $c^T = (1, 0)$, we will check how good the D-optimal design is for estimating the parameter $\theta_1$:

$$
\text{eff}_{c}(\xi_D^*) = \frac{c^T M^{-1}(\xi_{c\theta_1}) c}{c^T M^{-1}(\xi_D^*) c} \cdot 100 = 89.59\%
$$

On the other hand, if we are interested in seeing how precise the D-optimal design is for estimating $\theta_2$, we should take the vector $c^T = (0, 1)$, whereupon the efficiency is:

$$
\text{eff}_{c}(\xi_D^*) = \frac{c^T M^{-1}(\xi_{c\theta_2}) c}{c^T M^{-1}(\xi_D^*) c} \cdot 100 = 57.56\%
$$

![Figure 2.12: Plot of the generalized variance for example 2](image)

**Table 2.5: Efficiency of the D-optimal design for example 2**

<table>
<thead>
<tr>
<th>$\theta_2^{(0)}$</th>
<th>$\text{eff}_D(\xi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>99.82</td>
</tr>
<tr>
<td>8</td>
<td>99.19</td>
</tr>
<tr>
<td>3.5</td>
<td>81.17</td>
</tr>
<tr>
<td>15</td>
<td>98.13</td>
</tr>
<tr>
<td>25</td>
<td>93.89</td>
</tr>
<tr>
<td>89.6</td>
<td>87.76</td>
</tr>
</tbody>
</table>
Figure 2.13: Plot of the Elfving's set for example 2
CHAPTER 3

Growth model applied to a Friesian dairy farm

3.1 Introduction

Animal husbandry has been practiced for thousands of years since the first domestication of animals. Since then, the technology involved in breeding livestock has undergone a significant development, resulting in high productivity rates in farms. The study of the growth of animals in these industries is an area of intense interest since the evolution of the animals’ weight plays an essential role in their productivity. Therefore, having a good understanding of models which provide an estimated value of the mass of an animal as a function of its age, is of great importance.

This study concerns a specific variety of cow called Friesian. Originating in what is now the Netherlands and more specifically in the two northern provinces of North Holland and Friesland. The animals were the regional cattle of the Frisians and the Saxons. The Dutch breeders bred and oversaw the development of the breed with the goal of obtaining animals that could best use grass, the area’s most abundant resource. Over the centuries, the result was a high-producing, black-and-white dairy cow. It is black and white due to artificial selection by the breeders. Nowadays this breed of cow is the world’s leading milk producers with an average of 6,000 liters per year (Prendiville et al., 2011). The work was carried out in a dairy farm called A Devesa, located in the northwest of the region of Galicia, in Spain. The farm is one of the largest and most technologically advanced dairy farms in Spain.

The farm abides by both the European and the Spanish law related to good
practices when treating animals (Council Regulation, EC, No. 1/2005 of 22nd of December 2004 on the protection of animals during transport and related operations and amending Directives 64/432/EEC and 93/119/EC and Regulation, EC, No 1255/97 (OJ L 3 of 5.1.2005) and Spanish Royal Decree No. 692/2010 of 20th of May 2010). Furthermore, the farm is aware of the fact that animal welfare is not only affected by veterinarian cares but also by implementing an ethical code by which animals are going to feel in a comfortable environment.

From the farm, heifers are sent to a growing facility between ten and twenty days after birth, where they remain for two years. After this period, they are sent back to the dairy farm where they give birth for the first time and are ready then to be milked. During the two years at the facility, the animals’ weight must be kept within set control limits. This permits to make the best choice regarding the type and amount of fodder to give to the animal, based on its developmental stage and, in turn, will influence the quality and quantity of the milk produced (Zanton and Heinrichs, 2008). For that reason, the way in which the animals are fed is of great importance.

The weight control limits used are set by using a growth model. As with any other mathematical model, it depends on an unknown number of parameters. This work aims to find an optimal estimation of these parameters. For this purpose, a statistical experiment has been designed.

Some examples of where optimal design theory has been applied to growth models can be found in the literature (Dette and Pepelyshev, 2008; López-Fidalgo et. al, 2011) but in these cases, optimal designs have been calculated for uncorrelated observations. In contrast, this study measures the weight of the same animal at different points in time. These are repeated measurements and a correlation between them must be considered.

There is also an extensive literature on optimal design of experiments for correlated observations. Sacks and Ylvisaker (1970), accomplished the study for regression models from a theoretical point of view, while Müller (2007) worked on the framework of spatial statistics. An example of a numerical method for the construction of optimal designs for time-dependent models in the presence of correlation is shown in Uciński and Atkinson (2004). Zhigljavsky et. al
(2010) introduced a new design methodology for constructing asymptotic optimal designs for correlated data. Recently, Dette et al. (2013) made some progress providing explicit results on optimal designs for linear regression models with correlated observations that are not restricted to the location scale model. However, to the authors’ knowledge, the literature does not address optimal design of experiments for a growth model with correlated observations.

3.2 Model for growth

The model to be used at the heifer growing facility was derived by West et al. (2001). It is a general model for ontogenetic growth in organisms based on principles for the allocation of metabolic energy between the maintenance of existing tissue and the production of new ones (Nicholls and Ferguson, 2004). Ontogenetic development is fuelled by metabolism and occurs primarily by cell division. Incoming energy and materials from the environment are transported through hierarchical branching network systems to supply all cells. These resources are transformed into metabolic energy, which is used for life-sustaining activities. During growth, some fraction of this energy is allocated to the production of new tissue. Thus, the rate of energy transformation is the sum of two terms, one of which represents the maintenance of existing tissue, and the other, the creation of new tissue. This is expressed by the conservation of energy equation:

\[ B = \sum_c \left[ N_c B_c + E_c \frac{dN_c}{dt} \right]. \quad (3.1) \]

The incoming rate of energy flow, \( B \), is the average resting metabolic rate of the whole organism at time \( t \), \( B_c \) is the metabolic rate of a single cell, \( E_c \) is the metabolic energy required to create a cell and \( N_c \) is the total number of cells, the sum is over all types of tissue.

Possible differences between tissues are ignored and some average typical cell is taken as the fundamental unit. The first term, \( N_c B_c \), is the power needed to sustain the organism in all of its activities, whereas the second is the power allocated to production of new cells and therefore to growth. \( E_c \), \( B_c \) and the
mass of a cell, \( m_c \), are assumed to be independent of \( m \) remaining constant throughout growth and development.

At any time \( t \) the total body mass is \( m = m_c N_c \), so equation (3.1) can be written as

\[
\frac{dm}{dt} = \left( \frac{m_c}{E_c} \right) B - \left( \frac{B_c}{E_c} \right) m. \tag{3.2}
\]

Now, if \( B = B_0 m^{3/4} \), where \( B_0 \) is constant for a given taxon, then

\[
\frac{dm}{dt} = k m^{3/4} - bm, \tag{3.3}
\]

with \( a = B_0 m_c / E_c \) and \( b = B_c / E_c \). A model was developed for understanding the \( 3/4 \) exponent (West et. al, 1997). It is based on the premise that the tendency of natural selection to optimize energy transport has led to the evolution of fractal-like distribution networks. The \( 3/4 \) exponent was shown to be related to the scaling of the total number (\( N_t \)) of terminal units (capillaries) in the network: \( B \propto N_t \propto m^{3/4} \). In contrast, the total number of cells, \( N_c \propto m \). Thus, the reason for the different exponents of \( m \) in the two terms on the right-hand side of equation (3.3) is that the network constrains the total number of supply units (capillaries) to scale differently from the total number of cells supplied (Peters, 1983). This imbalance between supply and demand ultimately limits growth. If the exponents were the same, then \( dm/dt \neq 0 \) and organisms would continue to grow indefinitely, so an asymptotic maximum body size (\( M \)) is reached. This occurs when \( dm/dt = 0 \), giving \( M = (k/b)^4 = (B_0 m_c / B_c)^4 \). Thus, the variation in \( M \) among species within a taxon, where \( B_0 \) and \( m_c \) do not change, is determined by the systematic variation of the in vivo cellular metabolic rate, \( B_c \). Within a taxon, \( B_0 \), \( m_c \) and \( E_c \) are approximately constant, so \( k \) should be approximately independent of \( M \), whereas \( b \) should scale as \( M^{-1/4} \). Between groups, however, \( k \) should vary, principally reflecting variations in \( B_0 \). Equation (3.3) can therefore be re-expressed as

\[
\frac{dm}{dt} = k m^{3/4} \left[ 1 - \left( \frac{m}{M} \right)^{1/4} \right]. \tag{3.4}
\]
Therefore, the mass $m$ of the organism is obtained by integrating at any time $t$ (see Appendix D),

$$
m = M \left[ 1 - \left( 1 - \left( \frac{m_0}{M} \right)^{1/4} \right) \exp \left\{ - \frac{kt}{M^{1/4}} \right\} \right]^4, \tag{3.5}
$$

where $m_0$, $M$ and $k$ are the mass at birth ($t = 0$), the asymptotic maximum body mass and a parameter governing the growth, respectively. Thus, according to this equation the body mass will increase continuously ($dm/dt > 0$ at any time $t$) with an asymptote at mass $m = M$, which in practice means the body mass tends to be steady after some time.

Heifers are sent from the dairy farm to the growing facility when they are between ten and twenty days old. All of them are weighed upon their arrival. Accordingly, one cannot determine a priori with total exactitude the age at which the animals will be weighed for the first time. As the distribution of birth can be considered uniform over time, this design specifies that the first measure will be taken at time $t_1 \sim \mathcal{U}(10, 20)$, so the variability in this interval (standard deviation) is only 2.8 days. Animals are also weighed when they are approximately 60, 120, 160, 240, 450 days and 2 years old, just before being sent back to the dairy farm.

The goal to be achieved is to find an optimal estimation of the vector parameter $\alpha = (m_0, M, k)'$ through the D-optimality criterion. The design $\xi$ will consist of measuring at times

$$\{t_1, t_2, t_3, t_4, t_5, t_6, t_7\},$$

where

$$t_1 \sim \mathcal{U}(10, 20), \quad t_2 = t_1 + 45, \quad t_7 = t_1 + 715, \quad (3.6)$$

and for the rest of the times,

$$t_i = t_1 + 45 + \sum_{j=3}^{i} h_j, \quad i = 3, \ldots, 6, \quad \sum_{j=3}^{6} h_j \leq 715 - 45 = 670. \quad (3.7)$$

The values of $h_3, h_4, h_5, h_6 > 0$ have to be optimized. When these heifers are
two months old, their weaning takes place. This is the process of gradually introducing the animal to what will be its adult diet and withdrawing the supply of its mother’s milk. After 2 years, the heifers are expected to give birth and start their adult life back at the dairy farm. For these two reasons, times $t_2$ and $t_7$ are fixed in practice.

### 3.3 Information matrix

The presence of correlation has been considered because the observations on a single heifer may not be independent. When a correlation structure is considered, the statistical model is defined as follows,

$$ y = \eta(t, \theta) + \varepsilon \quad t \in \chi, $$

where $\eta(t, \theta)$ is the expected value of $y$, $\theta$ represents the $r$-dimensional vector of unknown parameters and $t$ represents the time-point at which the response is observed. These times vary in a compact design space $\chi$. The error $\varepsilon$ follows a Gaussian process with zero mean and a covariance structure of $y$ depending on the period of time between measurements (isotropic),

$$ \text{Cov} \left( y(t_i), y(t_j) \right) = c(|t_i - t_j|, \theta), \quad (3.8) $$

where $c(., \theta)$ is called the covariance function. Vector $\theta$ includes all the parameters of the model for both the mean and the covariance in either case of overlapping or distinct parameters, say $\alpha$ for the mean, $\eta$, and $\gamma$ for the covariance.

As mentioned above, optimal design of experiments theory allows us to find the best design in the sense of obtaining an optimal estimator of the parameters of the model by minimizing a function of the variance-covariance matrix of $\hat{\theta}$ (Pázman, 1986).

The maximum-likelihood estimator (MLE) of $\theta$ from the $n$ experimental ob-
servations is given by

$$\hat{\theta} = \arg \max_\theta \log f(y|\xi, \theta),$$

where

$$\log f(y|\xi, \theta) = -\frac{1}{2} \left[(y - \eta(\xi, \theta))' \Sigma^{-1} (y - \eta(\xi, \theta)) + \log \det \Sigma + n \log(2\pi)\right],$$

$y = (y(t_1), \ldots, y(t_n))'$, $\xi = (t_1, \ldots, t_n)'$, $\eta(\xi, \theta) = (\eta(t_1, \theta), \ldots, \eta(t_n, \theta))'$ and $\Sigma$ being the variance-covariance matrix whose generic $(i,j)$ entry is defined as

$$\Sigma_{ij} = \text{Cov} (y(t_i), y(t_j)) = c(|t_i - t_j|, \theta). \quad (3.9)$$

After some calculations (see Appendix E) one obtains the information matrix of the design $\xi$,

$$M(\xi, \theta) = \frac{\partial \eta(\xi, \theta)}{\partial \theta} \Sigma^{-1} \frac{\partial \eta(\xi, \theta)}{\partial \theta}' + \frac{1}{2} \text{tr} \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta}' \right), \quad (3.10)$$

with $\frac{\partial \Sigma}{\partial \theta}$ being a $n \times n$ matrix whose elements are $r \times r$ matrices. Thus the resulting product of the second term can be seen as a $n \times n$ diagonal block matrix where each block has dimension $r \times r$. The symbol “tr” is applied here as the sum of these blocks, resulting in a final $r \times r$ matrix.

In our case, $\eta(t, \theta)$ plays the role of $m$ in model (3.5) and the observations of the heifer’s masses will be referred as $y$. We define the covariance function as we made in (3.8) but in this case, the strength of the correlation between two observations depends strongly on the temporal lag between them, rather than on the magnitude of the response itself. Therefore, the covariance structure and the mean will not share common parameters. This statement is supported by the data provided by the growing facility (see table 3.1). In that table, the last column shows the mean of the weights of the heifers when they have reached their adult life (4 years). This last measurement has been made at the dairy farm. In this situation, it is natural to define a covariance structure by using a function which exponentially decays with increasing time-distance
between the measurements,

\[ \text{Cov}(y_{t_i}, y_{t_j}) = \sigma^2 \exp\{-\beta |t_i - t_j|\}. \]  

(3.11)

The presence of correlation in the observations implies the information matrix is not additive anymore. Accordingly, we must restrict ourselves to a practical exact design, fixing a priori the number of time-points of the design,

\[ \xi^* = \{t_1, t_1 + 45, t_1 + 45 + h_3, t_1 + 45 + h_3 + h_4, t_1 + 45 + \sum_{j=3}^{5} h_j, t_1 + 45 + \sum_{j=3}^{6} h_j, t_1 + 715\}. \]

Since the function \( m \) given in (3.5) and the covariance structure (3.11) do not share common parameters (\( m \) only depends on parameters \( m_0, M \) and \( k \), and \( \Sigma \) is only governed by the parameters \( \beta \)) the information matrix for the design \( \xi \) given in (3.10) adopts the form of the following 4 \( \times \) 4 diagonal block matrix, where the first block has dimension 3 \( \times \) 3 and the second one has dimension 1 \( \times \) 1 (Pázman, 2004; Pázman, 2007),

\[
M(\xi, \theta) = E_{t_1} \begin{pmatrix}
\frac{\partial m'}{\partial \alpha} \Sigma^{-1} \frac{\partial m}{\partial \alpha'} & 0_{3 \times 2} \\
0_{2 \times 3} & \frac{1}{2} \text{tr} \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \beta} \Sigma^{-1} \frac{\partial \Sigma}{\partial \beta'} \right)
\end{pmatrix},
\]  

(3.12)

where \( \theta = (\alpha, \beta) \) and \( \alpha = (m_0, M, k) \).

The original information matrix depends of \( t_1, h_3, h_4, h_5 \) and \( h_6 \), but \( t_1 \) will come at random. Thus, the information matrix considered is then the expectation of it with respect to the prior distribution assumed for \( t_1 \). Therefore, to obtain the D-optimal design, we must calculate the values of \( h_3, h_4, h_5 \) and \( h_6 \) which give the minimum of \[ \det M(\xi, \theta) \]^{-1/4}.

Matrix \( M(\xi, \theta) \) depends on the parameters \( m_0, M, k \) and \( \beta \) because the model is non-linear, so we have to use nominal values of the parameters to carry out the optimization. The dairy farm’s owner provided the weights of 250 heifers which were sent to the growing facility (Table 3.1). The table also
contains the weights of these animals when they were four years old. This last measurement took place in the dairy farm and will help calculate a nominal value of the asymptotic maximum body mass, $M$. Nominal values have been obtained by using the maximum-likelihood estimation,

$$\theta^{(0)} = (\alpha^{(0)}, \beta^{(0)}) = \arg \max_{\alpha, \beta} \log f(y \mid t, \theta),$$

where

$$\log f(y \mid t, \theta) = -\frac{1}{2} \left[ (y - m)' \Sigma^{-1} (y - m) + \log \det \Sigma + 8 \log(2\pi) \right] \quad i = 1, \ldots, 8.$$ 

<table>
<thead>
<tr>
<th>days</th>
<th>10-20</th>
<th>55-65</th>
<th>110-120</th>
<th>160-170</th>
<th>230-240</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight (kg)</td>
<td>42</td>
<td>69</td>
<td>139</td>
<td>182</td>
<td>265</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>days</th>
<th>420-430</th>
<th>720-730</th>
<th>1450-1460</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight (kg)</td>
<td>401</td>
<td>645</td>
<td>736</td>
</tr>
</tbody>
</table>

Table 3.1: Mean values of weights at their corresponding age from 250 heifers. These data have been provided by the growing facility.

The values of $y = (y_1, \ldots, y_n)$ are taken from the second row of the Table 3.1 and $m = (m_1, \ldots, m_n)$ is calculated by using (3.5) for times $t_1, \ldots, t_n$, which are taken from the first row of the same Table, being calculated as the midpoint of the interval. Once the Maximum-Likelihood method has been carried out,

$$\theta^{(0)} = (m_0^{(0)}, M^{(0)}, k^{(0)}, \beta^{(0)}) = (40.3, 759.1, 0.1, 10.6). \quad (3.13)$$

These values are used as nominal values for computing a locally optimal design.

The maximum of $\det M(\xi, \theta)$ is calculated numerically. In order to reduce the high computational cost, it is necessary to accomplish a reparameterization of the model given in (3.5). In this context, the D-optimal design is invariant since the maximization of the determinant of the information matrix given in (3.12)

$$\det M(\xi, \theta) = \frac{1}{2} \det \left[ tr \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \beta} \Sigma^{-1} \frac{\partial \Sigma}{\partial \beta'} \right) \right] \det \left[ \frac{\partial m'}{\partial \alpha} \Sigma^{-1} \frac{\partial m}{\partial \alpha'} \right],$$
is not affected by any reparameterization of the type \( \tilde{m}(\tau) = m(\tau(\alpha)) \), \( \tau(\alpha) \) being a one-to-one differential mapping. This is because

\[
\frac{\partial \tilde{m}}{\partial \tau} = \frac{\partial \alpha}{\partial \tau} \frac{\partial m}{\partial \alpha},
\]

and therefore,

\[
\frac{\partial \tilde{m}'}{\partial \tau} \Sigma^{-1} \frac{\partial \tilde{m}}{\partial \tau'} = \frac{\partial \alpha}{\partial \tau} \frac{\partial m'}{\partial \alpha} \Sigma^{-1} \frac{\partial m}{\partial \alpha} \frac{\partial \alpha}{\partial \tau'}.
\]

By taking \( \tau = (a, b, c)' \),

\[
c = M^{1/4}, \quad b = 1 - \left( \frac{m_0}{M} \right)^{1/4} \quad \text{and} \quad a = \frac{k}{M^{1/4}},
\]

the model \( m \) given in (3.5) becomes

\[
\tilde{m} = \left[ c \left( 1 - b \exp\{-a t\} \right) \right]^4.
\]

This reparameterized model will enormously reduce the computational cost.
3.4. **OPTIMAL DESIGNS**

**Optimal designs**

At the growing facility, heifers are weighed at seven different ages, so our design is also composed of seven measurements. We will demonstrate how these measurements are more efficient for estimating the parameters (in the sense of D-optimality) than the ones carried out at the facility. Note, that since we are in the correlated errors setting, the instruments from classical approximate design theory are not available to us, unless we adopt a very different definition of design measures (Müller and Pázman, 2003). Thus, in this work we confine ourselves to finding exact designs.

The calculation of the optimal design consists of finding the values of $h_3$, $h_4$, $h_5$ and $h_6$ (4.5) which minimizes the value of $\left[ \det M_\xi(t, \theta) \right]^{-1/4}$. Once this minimization has been carried out, we have the locally optimal design,

$$\xi^*_1 = \{t_1, t_1 + 45, t_1 + 45, t_1 + 300, t_1 + 400, t_1 + 715, t_1 + 715\}.$$ 

Since $t_1$ is a random time, we cannot control a priori the exact age at which the heifers will be weighed. Thus, we will optimize the periods of time between measurements. Then once the first has been carried out, the rest of the weighing times are clear.

We defined the efficiency of any design $\xi$ compared with another $\zeta$ as

$$\text{eff}_{\xi, \zeta} = \frac{\Phi[M(\xi, \theta)]}{\Phi[M(\zeta, \theta)]} = \frac{[\det M(\xi, \theta)]^{-1/r}}{[\det M(\zeta, \theta)]^{-1/r}},$$

$r$ being the number of parameters of the model. The efficiency can sometimes be multiplied by 100 and be reported in percentage terms.

In order to compare $\xi^*_1$ with the measurements taken at growing facility (from now on expressed as $\xi_f$), we check the relative efficiency of this design with respect to the other. Through the efficiency we measure how better $\xi^*_1$ is compared to $\xi_f$,

$$\text{eff}_{\xi^*_1, \xi_f} = \frac{[\det M_{\xi^*_1}(t, \theta^{(0)})]^{-1/4}}{[\det M_{\xi_f}(t, \theta^{(0)})]^{-1/4}} = \frac{0.00245}{0.00324} = 75\%.$$
The main drawback to $\xi^*_1$ is that carrying out a measurement two times at $t_1 + 45$ and at $t_1 + 715$ on the same animal has no utility from a practical point of view. An alternative would be to consider optimal regular sequences of four time-points between $t_1 + 45$ and $t_1 + 715$ which, along with the three fixed time-points (3.6), form the design. We consider arithmetic, geometric, inverse and harmonic progressions. Table 3.2 shows the efficiency of these designs compared to $\xi^*_1$. The second column represents the values of the other four points of the corresponding design that should be added to the value of $t_1$. Only the geometric design is more efficient than $\xi_f$ but there is not much difference between them. A second alternative consists in considering the interval

$$[t_1 + 45 + \epsilon, t_1 + 715 - \delta],$$

and then calculating for different values of $\epsilon$ and $\delta$, the four different time-points belonging to that interval which minimizes the value of the determinant $[\det M_\xi(t, \theta^{(0)})]^{-1/4}$. These points, plus the ones described in (3.6), form the design.

The greater the values of $\epsilon$ and $\delta$ are, the less efficiency the design will have (see Table 3.3).

$$\xi^*_2 = \{t_1, t_1 + 45, t_1 + 60, t_1 + 300, t_1 + 660, t_1 + 700, t_1 + 715\}.$$

Design $\xi^*_1$ is more efficient, but $\xi^*_2$ has no repeated time-points. Comparing $\xi^*_2$ with $\xi_f$ we can see that $\xi^*_2$ is more efficient with an efficiency of around 86%.

Finally, the introduction of a nugget effect is considered. In the covariance structure a nugget term is included since optimal design points may collapse under the presence of correlation when no nugget effect is present. As replicates of measurements at the same time on the same heifer do not make sense from a practical point of view, the designs calculated must contain $n$ distinct time-points.

The conception of the nugget term was first introduced in Geostatistics by Matheron (1962). It is also widely used in Gaussian processes (Pepelyshev, 2010) and Spatial Statistics (Cressie, 1993; Ripley, 1981). For an isotropic correlation structure the variance-covariance matrix for two observations tends
3.4. **OPTIMAL DESIGNS**

<table>
<thead>
<tr>
<th>sequence</th>
<th>$\left[ \det M_\xi(t, \theta) \right]^{-1/4}$</th>
<th>$\text{eff}_\xi^*(\cdot)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic ${189, 313, 447, 581}$</td>
<td>0.00369</td>
<td>66%</td>
</tr>
<tr>
<td>Geometric ${84, 148, 254, 428}$</td>
<td><strong>0.00300</strong></td>
<td><strong>82%</strong></td>
</tr>
<tr>
<td>Inverse ${162, 267, 380, 516}$</td>
<td>0.00330</td>
<td>74%</td>
</tr>
<tr>
<td>Harmonic ${58, 80, 119, 211}$</td>
<td>0.00340</td>
<td>72%</td>
</tr>
</tbody>
</table>

Table 3.2: Designs based on regular sequences. The four middle points are shown in the second column before summing up $t_1$

<table>
<thead>
<tr>
<th>Interval</th>
<th>sequence</th>
<th>$\left[ \det M_\xi(t, \theta) \right]^{-1/4}$</th>
<th>$\text{eff}_\xi^*(\cdot)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1 + [60, 700]$</td>
<td>${60, 300, 660, 700}$</td>
<td><strong>0.00281</strong></td>
<td><strong>87%</strong></td>
</tr>
<tr>
<td>$t_1 + [75, 685]$</td>
<td>${75, 315, 660, 685}$</td>
<td>0.00298</td>
<td>82%</td>
</tr>
<tr>
<td>$t_1 + [90, 625]$</td>
<td>${90, 335, 625, 625}$</td>
<td>0.00315</td>
<td>77%</td>
</tr>
<tr>
<td>$t_1 + [120, 600]$</td>
<td>${120, 340, 600, 600}$</td>
<td>0.00339</td>
<td>72%</td>
</tr>
</tbody>
</table>

Table 3.3: Designs with different points at the support and their efficiencies with respect to $\xi^*_1$. The four middle points points are shown in the second column before summing up $t_1$

to a singular form when the distance tends to zero. This behavior is due to the lack of microvariation allowed for by the assumed covariance function. Then optimal designs tend to avoid collapsing points. If a so-called nugget effect is introduced in the covariance structure more meaningful and practically relevant designs arise. In particular, sometimes it may be proved that the distance between the points of a two-point D-optimal design is an increasing function of the nugget effect (Stehlík et al., 2008). These correlation functions are typically used in the literature (Cressie, 1993). Amo et al. (2013) provided a general result to obtain a large class of feasible models for a covariance structure that is a function of the mean. We will define the covariance structure as follows,

$$\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} 
\sigma^2 \rho \exp\{-\beta |t_i - t_j|\} & \text{for } t_i \neq t_j, \\
\sigma^2 (1 - \rho) & \text{for } t_i = t_j.
\end{cases} \quad (3.14)$$

where $\rho$ is the nugget term (Stehlík et al., 2008). Now, with the introduction of the nugget term, the correlation structure (3.14) is depends on the parameters $\beta$ and $\rho$, so the information matrix adopts the form of the following $5 \times 5$ diagonal block matrix, where the first block has dimension $3 \times 3$ and the second one has dimension $2 \times 2$. 
\[ M(\xi, \theta) = E_{t_1} \begin{pmatrix} \frac{\partial m'}{\partial \alpha} \Sigma^{-1} \frac{\partial m}{\partial \alpha'} & 0_{3 \times 2} \\ 0_{2 \times 3} & \frac{1}{2} \text{tr} \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \gamma} \Sigma^{-1} \frac{\partial \Sigma}{\partial \gamma'} \right) \end{pmatrix}, \quad (3.15) \]

where \( \theta = (\alpha, \gamma) \), \( \alpha = (m_0, M, k) \) and \( \gamma = (\beta, \rho) \). In order to obtain the D-optimal design, we must calculate the values of \( h_3, h_4, h_5 \) and \( h_6 \) which give the minimum of \( \left[ \det M(\xi, \theta) \right]^{-1/5} \). Again, we have calculated the nominals values by using the maximum-likelihood estimation,

\[
\theta^{(0)} = (\alpha^{(0)}, \gamma^{(0)}) = (m_0^{(0)}, M^{(0)}, k^{(0)}, \beta^{(0)}, \rho^{(0)}) = (43, 751, 0.08, 4, 0.95). \quad (3.16)
\]

Once this minimization has been carried out, we have the locally optimal design,

\[ \xi^*_3 = \{t_1, t_1 + 45, t_1 + 60, t_1 + 340, t_1 + 360, t_1 + 700, t_1 + 715\}. \]

Through the efficiency we measure how much better \( \xi^*_3 \) is compared to \( \xi_f \),

\[
\text{eff}_{\xi^*_3, \xi_f} = \frac{[\det M(\xi^*_3, \theta^{(0)})]^{-1/5}}{[\det M(\xi_f, \theta^{(0)})]^{-1/5}} = \frac{0.001887}{0.002360} = 80.5\%. \]
Robustness analysis

3.5.1 With respect to the choice of the nominal values

In this section it will be checked how the quality of the optimal design would be affected by a wrong choice of the nominal value. Assuming \( \theta^* \) are possible true values of the parameters and \( \theta^{(0)} \) are the nominal values used for the computation of the D-optimal design \( \xi_{\theta^{(0)}}^* \), the efficiency

\[
\text{eff}_{\xi_{\theta^{(0)}}^* \xi_{\theta^{(0)}}^*} = \frac{\left[ \det M(\xi_{\theta^{(0)}}^*, \theta^*) \right]^{-1/5}}{\left[ \det M(\xi_{\theta^{(0)}}^*, \theta^*) \right]^{-1/5}},
\]

measures the goodness of the design \( \xi_{\theta^{(0)}}^* \) obtained under the nominal values, where \( \xi_{\theta^*}^* \) is the actual optimal design. Figure 3.1 illustrates the robustness of the D-optimal design \( \xi_3^* \) with respect to the choice of the parameters \( m_0, M \) and \( k \). As can be observed, differences in the order of \( 10^{-2} \) for the parameter \( k^{(0)} \) keeps the value of the efficiency over 75%, even when the variations of the parameters \( M^{(0)} \) and \( m_0^{(0)} \) are high.

![Figure 3.1: Relative efficiencies (in %) of the design \( \xi_3^* \) for four different values of the parameter \( k \) and parameters \( m_0 \) and \( M \)](image)

3.5.2 Choice of the correlation structure

We have considered a well known and widely used model for the trend of the growing weight of animals. We claim a correlation structure has to be considered when there are repeated measurements. This is convenient from
both the practical and the statistical point of view, resulting in information gains and cost reductions. A novelty that this work introduces is the choice of such a correlation structure. The one used here is rather usual within this framework, but other may also be suitable. We compare the efficiency of the locally optimal designs obtained with respect to the choice of these three typical covariance structures:

(a) Dagum function,

\[
\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} 
\sigma^2 \left[ \rho \left( 1 - \frac{(t_i - t_j)^\beta}{1 + (t_i - t_j)^\beta} \right) \right] & \text{for } t_i \neq t_j, \\
\sigma^2 (1 - \rho) & \text{for } t_i = t_j.
\end{cases}
\]

(b) Cauchy function,

\[
\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} 
\sigma^2 \rho \left( 1 + (t_i - t_j)^\beta \right)^{-\gamma} & \text{for } t_i \neq t_j, \\
\sigma^2 (1 - \rho) & \text{for } t_i = t_j.
\end{cases}
\]

(c) Gaussian model,

\[
\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} 
\sigma^2 \rho \exp\{-\beta^2 (t_i - t_j)^2\} & \text{for } t_i \neq t_j, \\
\sigma^2 (1 - \rho) & \text{for } t_i = t_j.
\end{cases}
\]

Figure 3.2: Plot of the correlation structures (3.14), (3.17), (3.18) and (3.19)
sequence \[ \text{det} M(\xi, \theta) \] \[ -1/5 \] eff(\xi^*_1) \\
\xi_{dag} \{70, 300, 345, 705\} 0.002734 69 \%
\xi_{ca} \{70, 320, 350, 675\} 0.002195 86 \%
\xi_{ga} \{60, 70, 90, 350\} 0.002328 81 \%

Table 3.4: Designs based on covariance structures (3.17), (3.18) and (3.19). The four middle points are shown in the second column before summing up $t_1$.

Table 3.4 shows the efficiencies of the locally optimal design $\xi^*_3$ with respect to the designs $\xi^*_{dag}$, $\xi^*_{ca}$ and $\xi^*_{ga}$, calculated for model (3.5) with the covariance structures (3.17), (3.18) and (3.19). The similar behavior of correlation structures (3.18) and (3.19) (Figure 3.2) permits us to compare the designs obtained with them. The efficiency is barely affected by the choice of these two correlation structures (see Table 3.4). Even for the Dagum function, which is quite far from the nominal as shown in the Figure 3.2, the loss of efficiency is only about 30%.
Chapter 4

Discrimination between growth models applied to a limousine beef farm

4.1 Introduction

This chapter is concerned with problem of finding an optimal experimental design for discriminating between two competing growth models applied to a beef farm. T-optimal design is usually used for this discrimination and the models involved are usually homoscedastic models with normally distributed observations. But in this case, a criterion based on the Kullback-Leibler distance is proposed, since correlated observations must be considered due to the fact that observations on a particular animal are not independent.

In order to optimize the efficiency of beef production systems, it is of great importance to know the behaviour of weight gain in cattle throughout time. The growth of beef-specialized breeds is characterized by models based on non-linear sigmoid curves. The most popular are (Brody, 1945; Richards, 1959; Bertalanffy, 1934; Nelder, 1961 (generalized logistic) and Gompertz, 1925). The shape and characteristics of these curves can vary depending on factors such as the environment, production system, type of breed and so on.

This study has been carried out in a beef farm called Navalázar, located in the northwest of the region of Córdoba, Spain, and concerns a specific beef cattle breed called Limousine. This farm also abides by both the European and the Spanish law above mentioned and is aware of the fact that animal welfare is not only affected by veterinarian cares but also by implementing an ethical code by which animals are going to feel in a comfortable environment.

After weaning, which happens around six months after born, calves are sent
from the farm to the growing facility, where they remain for approximately 12 months before being sent to the abattoir. During this period, the animal’s weight must be kept under control. This permits to make the best choice regarding the type and amount of fodder to give, based on its developmental stage and, in turn, will influence the quality and quantity of the meat.

The weight control is adjusted by using growth models such as the one mentioned. This work considers the problem of constructing optimal experimental designs for discriminating between Brody and Richards models. These two models are nested (the extended model reduces to the simpler model for a specific choice of a subset of the parameters) and appear frequently in livestock research (Vergara et al., 2003; Goonewardene et al., 1981; Hirooka and Yamada, 1990; Cue et al., 2012).

Several studies have compared growth models for cattle (Brown et al., 1976; López de la Torre et al., 1992; Johnson et al., 1990; Fitzhugh Jr., 1976) whereas DeNise and Brinks (1985) and Beltrán et al.(1992) compared Brody and Richards curves. The Brody equation has been the most used in beef cattle studies because of its ease of computation and its goodness of fit (Stewart and Martin, 1983; Nelsen et al., 1982; Kaps et al., 1999), though in some studies Richards model was reported to fit data better than Brody’s (DeNise and Brinks, 1985; Beltrán et al., 1992; López de la Torre et al., 1992). Although this chapter is focused on discriminating between these two models, a decision-making problem with more than two models may be considered in further research.
4.2 T-optimality and KL-optimality criterion

In order to determine an optimal design for discriminating between two rival models \( \eta_1(t, \theta_1) \) and \( \eta_2(t, \theta_2) \), Atkinson and Fedorov (1975) proposed to fix one of them, say \( \eta(t, \theta) = \eta_1(t, \theta_1) \) (more precisely its corresponding parameters \( \theta_1 \)), considering it as the “true” model, and then to determine the design which maximizes

\[
T_{21}(\xi) = \min_{\theta_2 \in \Omega_2} \int \left[ \eta(t) - \eta_2(t, \theta_2) \right]^2 \xi(dx),
\]

where \( \eta(t) = \eta_1(t, \theta_1^{(0)}) \) is completely determined using some nominal values of \( \theta_i \in \Omega_1 \), i.e. \( \theta_i = \theta_i^{(0)} \). This design \( \xi \) is called T-optimal design. This criterion has been studied by numerous authors (Ponce de León and Atkinson, 1991; Ućinski and Bogacka, 2005; Dette and Titoff, 2009 or López-Fidalgo et al., 2007, among others).

The T-optimal design is essentially a maximin problem. The minimization is carried out since we first assume the worst-case scenario, that is, when \( \eta_2(t, \theta_2) \) is as close as possible to the “true” model. We then maximize \( T_{21}(\xi) \) to find the best among those worst possible situations. Except for very simple models, T-optimal discriminating designs are not easy to find and even their numerical determination is a very challenging task. As mentioned above, an important drawback of this approach consists of the fact that the criterion and, as a consequence, the corresponding optimal discriminating designs depend sensitively on the parameters of one of the competing models. In contrast to other optimality criteria, this dependence appears even for linear models. Therefore, T-optimal designs are locally optimal since they can only be implemented if some prior information regarding these parameters is available.

For the correlated case, the definition of \( T_{21}(\xi) \) can be given as follows (Amo-Salas et al., 2012),

\[
T_{21}(\xi) = \min_{\theta_2 \in \Omega_2} \left( \eta(t) - \eta_2(t, \theta_2) \right)^\prime \Sigma^{-1} \left( \eta(t) - \eta_2(t, \theta_2) \right), \tag{4.1}
\]
where $\Sigma$ is the covariance matrix whose generic $(i, j)$ entry is defined as in (3.9). It is a natural generalization of the T-optimality criterion function for the correlated case. Optimal exact designs are computed by maximizing this criterion. This criterion is again a particular case of KL-optimality and therefore it maximizes the test power for discrimination.

Let $f_1(y, t, \theta_1)$ and $f_2(y, t, \theta_2)$ be two rival density functions, where $f_1(y, t, \theta_1^{(0)})$ is assumed to be the true model. With this notation, the KL distance between the true model and $f_2(y, t, \theta_2)$ is defined as

$$I(f_1, f_2, t, \theta_2) = \int f_1(y, t, \theta_1^{(0)}) \log \left( \frac{f_1(y, t, \theta_1^{(0)})}{f_2(y, t, \theta_2)} \right) dy, \quad t \in \chi,$$

where the integral is computed over the sample space of the possible observations. Kullback and Leibler (1951) developed this quantity, motivated by considerations of information theory. They used the notation $I(f_1, f_2, \ldots)$ as a measure of the loss of information when $f_2$ is fitted to approximate $f_1$. Therefore, the KL-optimality criterion is defined as follows (López-Fidalgo et al., 2007),

$$I_{12}(\xi) = \min_{\theta_2 \in \Omega_2} \left\{ \int_{\chi} I(f_1, f_2, t, \theta_2) \xi(dt) \right\}. \quad (4.2)$$

A design which maximizes $I_{12}(\xi)$ is called KL-optimal design.

**Theorem.** Given two competing gaussian processes with means $\eta_1(t, \theta_1^{(0)})$ and $\eta_2(t, \theta_2)$, and covariance structures $\Sigma_1$ and $\Sigma_2$, respectively, the KL-optimality criterion leads to the expression,

$$I(f_1, f_2, t, \theta_2) = -\frac{1}{2} \log \left| \frac{\Sigma_1}{\Sigma_2} \right| - \frac{1}{2} n + tr \left( \Sigma_2^{-1} \Sigma_1 \right) +$$

$$\frac{1}{2} \left( \eta_1(t, \theta_1^{(0)}) - \eta_2(t, \theta_2) \right)' \Sigma_2^{-1} \left( \eta_1(t, \theta_1^{(0)}) - \eta_2(t, \theta_2) \right).$$

**Proof.**

$$I(f_1, f_2, t, \theta_2) = \int f_1(y, t, \theta_1^{(0)}) \log \left( \frac{f_1(y, t, \theta_1^{(0)})}{f_2(y, t, \theta_2)} \right) dy = E_1 \left[ \log \frac{f_1(y, t, \theta_1^{(0)})}{f_2(y, t, \theta_2)} \right].$$

As $f_1(y, t, \theta_1^{(0)})$ and $f_2(y, t, \theta_2)$ follow a Gaussian distribution,
4.2. T-OPTIMALITY AND KL-OPTIMALITY CRITERION

\[ E_1 \left[ \log \frac{f_1(y, t, \theta_1^{(0)})}{f_2(y, t, \theta_2)} \right] = E_1 \left[ \log f_1(y, t, \theta_1^{(0)}) - \log f_2(y, t, \theta_2) \right] \]
\[ = -\frac{1}{2} E_1 \left[ \log \frac{\Sigma_1}{\Sigma_2} \right] - \frac{1}{2} E_1 \left[ (y - \eta_1(t, \theta_1^{(0)}))^\prime \Sigma_1^{-1}(y - \eta_1(t, \theta_1^{(0)})) \right] \]
\[ + \frac{1}{2} E_1 \left[ (y - \eta_2(t, \theta_2))^\prime \Sigma_2^{-1}(y - \eta_2(t, \theta_2)) \right]. \]

For simplicity, let denote \( \eta_1 = \eta_1(t, \theta_1^{(0)}) \) and \( \eta_2 = \eta_2(t, \theta_2) \), respectively. The second term of the expectation \( E_1 \) is,

\[ E_1 \left[ (y - \eta_1)^\prime \Sigma_1^{-1}(y - \eta_1) \right] = -\frac{1}{2} \text{tr} \left( \Sigma_1^{-1} E_1 \left[ (y - \eta_1)^\prime (y - \eta_1) \right] \right) \]
\[ = \text{tr} \left( \Sigma_1^{-1} \Sigma_1 \right) = n. \]

And the third,

\[ E_1 \left[ (y - \eta_2)^\prime \Sigma_2^{-1}(y - \eta_2) \right] \]
\[ = E_1 \left[ [(y - \eta_1) + (\eta_1 - \eta_2)]^\prime \Sigma_2^{-1} [(y - \eta_1) + (\eta_1 - \eta_2)] \right] \]
\[ = \text{tr} \left( \Sigma_2^{-1} \Sigma_1 \right) + 2 E_1 \left[ (y - \eta_1)^\prime \Sigma_2^{-1}(\eta_1 - \eta_2) \right] + E_1 \left[ (\eta_1 - \eta_2)^\prime \Sigma_2^{-1}(\eta_1 - \eta_2) \right] \]
\[ = \text{tr} \left( \Sigma_2^{-1} \Sigma_1 \right) + 2 \Sigma_2^{-1}(\eta_1 - \eta_2) E_1[y - \eta_1] \]
\[ + (\eta_1 - \eta_2)^\prime \Sigma_2^{-1}(\eta_1 - \eta_2) = \text{tr} \left( \Sigma_2^{-1} \Sigma_1 \right) + 0 + (\eta_1 - \eta_2)^\prime \Sigma_2^{-1}(\eta_1 - \eta_2). \]

Therefore,

\[ I(f_1, f_2, t, \theta_2) = -\frac{1}{2} \log \left| \frac{\Sigma_1}{\Sigma_2} \right| - \frac{1}{2} n + \frac{1}{2} \text{tr} \left( \Sigma_2^{-1} \Sigma_1 \right) + \frac{1}{2} (\eta_1 - \eta_2)^\prime \Sigma_2^{-1}(\eta_1 - \eta_2). \]
Remark. For $\Sigma_1 = \Sigma_2 = \Sigma$,

$$I(f_1, f_2, t, \theta_2) = \frac{1}{2}(\eta_1 - \eta_2)'\Sigma^{-1}(\eta_1 - \eta_2),$$

which corresponds to the criterion (4.1). Therefore, the criterion defined in (4.2) is an extension of the extended T-optimality criterion for correlated observations when the covariance matrix is assumed equal for the rival models.

### 4.3 Discrimination between the two models

As mentioned above, these two models have already been compared for cattle, though in none of them this comparison were carried out by using optimal designs. They are general models for ontogenetic growth in organisms based on principles for the allocation of metabolic energy between the maintenance of existing tissue and the production of new ones (Nicholls and Ferguson, 2004). Richards model provides the mass of the organism at any time $t$:

$$\eta(t, \theta) = M \left( 1 - B \exp\{-kt\} \right)^A.$$

where $t$ is the age, $M$ represents the asymptotic maximum body mass (asymptotic mature weight), $B$ is a time scale parameter and $k$ and $A$ being the rate of approach to mature weight and a shape parameter that allows for a variable inflection point, respectively. Brody model is nested within Richards since it is a particular case of it when $A = 1$.

The presence of correlation has been considered because the observations on a single calf may not be independent. The fact of carrying out more than one measurement at the same time on the same animal has no utility from a practical point of view. Therefore, we will introduce a nugget effect in the covariance structure in order to avoid collapsing of design points. This effect produces a shift in these points which leads to an optimal design without replicated points. The conception of the nugget term was first introduced in Geostatistics by Matheron (1962). It is also widely used in Gaussian processes (Pepelyshev, 2010) and Spatial Statistics (Cressie, 1993 and Ripley, 1981). For an isotropic correlation structure, the variance-covariance matrix for two
4.3. DISCRIMINATION BETWEEN THE TWO MODELS

observations tends to a singular form when the distance tends to zero. This behavior is due to the lack of microvariation allowed for by the assumed covariance function. Optimal designs tend then to avoid collapsing points. If a nugget effect is introduced in the covariance structure more meaningful and practically relevant designs arise. In particular, sometimes it may be proved that the distance between the points of a two-point D-optimal design is an increasing function of the nugget effect (Stehlík et al., 2008). These correlation functions are typically used in the literature (Cressie, 1993). Amo-Salas et al. (2013) provided a general result to obtain a large class of feasible models for a covariance structure. We define the covariance structure by using a function which exponentially decays with increasing time-distance between the measurements,

\[
\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} 
\sigma^2 \rho \exp\{-\beta |t_i - t_j|\} & \text{for } t_i \neq t_j, \\
\sigma^2 (1 - \rho) & \text{for } t_i = t_j,
\end{cases}
\]

where \(\rho\) is the nugget term (Stehlík et al., 2008).

4.3.1 Hypothesis test for discrimination

Let us consider two competing Gaussian processes with means \(\eta_1(t, \theta_1)\) and \(\eta_2(t, \theta_2)\) given by Richards and Brody functions, respectively,

\[
\eta_1(t, \theta_1) = M_1 (1 - B_1 \exp\{-k_1 t\})^{A_1}, \\
\eta_2(t, \theta_2) = M_2 (1 - B_2 \exp\{-k_2 t\}),
\]

with correlation structures defined by (4.3). In this situation, the density functions associated to these two processes are

\[
f_k(y, t, \theta_k) = \frac{1}{(2\pi)^{n/2} |\Sigma_k|^{1/2}} \exp\{-(y - \eta_k(t, \theta_k))'\Sigma_k^{-1}(y - \eta_k(t, \theta_k))\} \quad k = 1, 2,
\]

where \(\Sigma_k\) is the variance-covariance matrix whose generic \((i, j)\) entry is defined as in (4.3).

In order to discriminate between Richards and Brody models, the following
hypothesis test may be considered:

\[ \begin{align*}
H_0 : f_2(y, t, \theta_2) \\
H_1 : f_1(y, t, \theta_1^{(o)})
\end{align*} \]

where \( \theta_1^{(o)} \) are nominal values of the parameter \( \theta_1 \). In this test the alternative hypothesis is assumed to be ”true” (this means Richards model is assumed to be ”true”) since we want to maximize the test power. The likelihood ratio for an observation \( y \) at time \( t \) will be

\[ L = \frac{f_2(y, t, \theta_2)}{f_1(y, t, \theta_1^{(o)})}, \]

and a common statistical test is that based on the statistic

\[ R = -2 \log(L) = 2 \log \left\{ \frac{f_1(y, t, \theta_1^{(o)})}{f_2(y, t, \theta_2)} \right\}, \]

in such a way that the hypothesis \( H_0 \) will be rejected for large values of \( R \). The expectation of this statistic for one design point, under \( H_1 \), is

\[ E_{H_1}(R) = 2 \int f_1(y, t, \theta_1^{(o)}) \log \left[ \frac{f_1(y, t, \theta_1^{(o)})}{f_2(y, t, \theta_2)} \right] dy = 2 \mathcal{I}(f_1, f_2, t, \theta_2). \quad (4.4) \]

The larger \( E_{H_1}(R) \) and \( \mathcal{I}(f_1, f_2, t, \theta_2) \) are, the larger the power function of \( R \) is. This is because hypothesis \( H_0 \) is rejected when this statistic is greater than a critical value. Using equation (4.4) for an exact design and the corresponding observations, we obtain,

\[ \mathcal{I}_{12}(\xi) = \min_{\theta_2 \in \Omega_2} \left\{ \int_X \int f_1(y, t, \theta_1^{(o)}) \log \left[ \frac{f_1(y, t, \theta_1^{(o)})}{f_2(y, t, \theta_2)} \right] dy \xi(dt) \right\}, \]

this expression being proportional to \( \min_{\theta_2 \in \Omega_2} \{E_{H_1}(R)\} \). Therefore, the KL-optimal design maximizes the power function in the worst case (López-Fidalgo et al., 2007).
4.4. ALGORITHM

Algorithm

In order to compute optimal designs, the numerical algorithm developed by Brinkkulov et al. (1986) is adapted to KL-optimality. It is an exchange-type algorithm that starts from an arbitrary initial \( n \)-points design. In case of exact designs, this number of points is fixed by the practitioner and none of them are repeated. At each iteration one support point is deleted from the current design and a new point is included in its place to maximize the value of the criterion function. Next, the algorithm is detailed:

**Step 1.** Select an initial design \( \xi_n^{(0)} = \{t_1^{(0)}, \ldots, t_n^{(0)}\} \) such that \( t_i^{(0)} \neq t_j^{(0)}, i, j \in \{1, 2, \ldots, n\}, i \neq j \).

**Step 2.** Compute

\[
\tilde{\theta}_2^{(0)} = \arg \min_{\theta_2 \in \Omega_2} \mathcal{I}(f_1, f_2, t, \theta_2) \quad \text{and} \quad \Delta(\xi_n^{(0)}) = \mathcal{I}(f_1, f_2, t, \tilde{\theta}_2^{(0)})
\]

**Step 3.** Determine

\[
(i^*, t^*) = \arg \max_{(i, t) \in I \times \chi} \Delta(\xi_n^{(0)}),
\]

where \( \Delta(\xi_n^{(0)}, t_i \leftrightarrow t) \) means that the support point \( t_i \) in the design \( \xi_n^{(0)} \) is exchanged by \( t \in \chi \). If

\[
\frac{\Delta(\xi_n^{(0)}, t_i \leftrightarrow t^*) - \Delta(\xi_n^{(0)})}{\Delta(\xi_n^{(0)})} < \delta,
\]

where \( \delta \) is the given tolerance, then STOP. Otherwise,

\[
\xi_n^{(1)} = \{t_1^{(0)}, \ldots, t_i^*, \ldots, t_n^{(0)}\},
\]

and we go to step 1, taking \( \xi_n^{(1)} \) as initial design.

Before calculating the value of \( \tilde{\theta}_2^{(0)} \) we must know the nominal value of \( \theta_1 \), \( \theta_i^{(0)} \). This nominal value has been obtained by using the Maximum-Likelihood
Estimation from historical data (see Appendix G),

\[
\theta_1^{(0)} = \arg \max_{\theta_1} \log \left( \frac{1}{(2\pi)^{n/2}\mid \Sigma_1^{1/2} \mid} \exp \left\{ \left( y - \eta_1(t, \theta_1) \right) \Sigma_1^{-1} (y - \eta_1(t, \theta_1)) \right\} \right) .
\]

The values of \( y = (y_1, \ldots, y_n) \) correspond to the weight of a single calf at eight different ages. They were provided by Navalázar farm. Once the maximum-likelihood method has been carried out,

\[
\theta_1^{(0)} = (M_1^{(0)}(1), B_1^{(0)}(1), k_1^{(0)}(1), A_1^{(0)}(1), \beta_1^{(0)}(1), \rho_1^{(0)}(1)) = (796, 0.66, 0.0044, 3.89, 0.04, 0.95).
\]

These values of \( \theta_1^{(0)} \) are used as nominal values for computing a locally optimal design.

### 4.4.1 Calculation of the KL-optimal design

As mentioned above, calves are sent to the growing facility just after their weaning, all of them being weighed upon their arrival. Accordingly, one cannot determine a priori exactly the age at which the animals will be weighed for the first time. As the distribution of birth can be considered uniform over time, this design specifies that the first measure after weaning will be taken at time \( t_1 \sim U(170, 190) \), since approximately every 20 days a group of animals are sent to the growing facility. Around eighteen months after birth (540 days), the yearlings are sent to the abattoir where they will be killed for consumption as food. Therefore, the first measurement will be made as soon as possible after weaning, that is, \( t_1 \sim U(170, 190) \) and the last one when they are about 540 days old, that is, \( t_8 = t_1 + 540 - 180 = t_1 + 360 \). The design \( \xi \) will consist then of measuring at times

\[
\{ t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8 \},
\]

where \( t_1 \sim U(170, 190) \), \( t_8 = t_1 + 360 \) and for the rest of the times,

\[
t_i = t_1 + \sum_{j=2}^{i} h_j \quad i = 2, \ldots, 7, \quad \sum_{j=2}^{7} h_j \leq 540 - 180 = 360.
\]
The values of $h_3, h_4, h_5, h_6, h_7 > 0$ have to be optimized by using the algorithm. Since $t_1$ is a random time, we cannot control a priori the exact age at which the calf will be weighed. Thus, we will optimize the periods of time between measurements. Once this minimization has been carried out, we have the locally KL-optimal design,

$$\xi^* = \{t_1, t_1 + 30, t_1 + 60, t_1 + 80, t_1 + 90, t_1 + 110, t_1 + 240, t_1 + 360\}.$$  

The relative efficiency of any design $\xi$ compared with another $\zeta$ is computed by dividing the values of the KL-optimality criterion. We compare then the values

$$\text{eff}_{\xi, \zeta} = \frac{I_{12}(\xi)}{I_{12}(\zeta)}.$$  

The efficiency can sometimes be multiplied by 100 and be reported in percentage terms. If this efficiency is higher than 1 then the power test for discrimination between the two models is higher with the design $\xi$ than with the design $\zeta$. We intend to compare the relative efficiency of $\xi$ with respect to the measurements taken at the growing facility, which from now on will be expressed as $\xi_f$:

$$\xi_f = \{t_1, t_1 + 50, t_1 + 100, t_1 + 150, t_1 + 205, t_1 + 255, t_1 + 310, t_1 + 360\},$$  

where $t_1 \sim \mathcal{U}(170, 190)$. This design consists of eight points representing the age at which the calves were weighed at the growing facility. Through the efficiency we measure how much better $\xi^*$ is compared to $\xi_f$,

$$\text{eff}_{\xi_f, \xi^*} = \frac{I_{12}(\xi_f)}{I_{12}(\xi^*)} = 66\%.$$  

In this section it will be checked how the quality of the optimal design would be affected by a wrong choice of the nominal value. Let us call $\theta^*$ as any possible true value of the parameters and $\theta^{(0)}$ being the nominal values used for the computation of the KL-optimal design $\xi_{\theta^{(0)}}^*$. The efficiency

$$\text{eff}_{\xi_{\theta^{(0)}}^*, \xi_{\theta^*}^*} = \frac{I_{12}(\xi_{\theta^{(0)}}^*)}{I_{12}(\xi_{\theta^*}^*)}$$

measures the goodness of the design $\xi_{\theta^{(0)}}^*$ obtained under the nominal values, where $\xi_{\theta^*}^*$ is the actual optimal design. Figures 4.1, 4.2 and 4.3 illustrate the robustness of the KL-optimal design $\xi^*$ with respect to the choice of the parameters $M_1$, $B_1$, $k_1$ and $A_1$. Shifting around 10% the parameters $k_1$, $A_1$ and $B_1$ keeps the efficiency over 70%, even when the variations of the parameter $M_1$ is large (from 756 to 835 kg). On the other hand, Figure (4.4) shows the robustness of $\xi^*$ with respect to the choice of the parameters $\rho$ and $\beta$. The higher the value of $\rho$ is, the greater the decrease in the value of efficiency will be.

We have considered a well-known and widely used model for the trend of the growing weight of animals. We claim a correlation structure has to be considered when there are repeated measurements. This is convenient from both the practical and the statistical points of view, resulting in information gain and cost reduction. A novelty that this work introduces is the choice of such a correlation structure. The one used here is rather usual within this framework, but others may also be suitable. We compare the efficiency of the locally optimal designs obtained with respect to the choice of these three typical covariance structures:

(a) Dagum function

$$\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} \sigma^2 \left[ \rho \left( 1 - \frac{(t_i - t_j)\beta}{1 + (t_i - t_j)\beta} \right)^\gamma \right] & \text{for } t_i \neq t_j \\ \sigma^2 (1 - \rho) & \text{for } t_i = t_j \end{cases}$$

(4.6)
4.5. **ROBUSTNESS ANALYSIS**

(b) Cauchy function,

\[
\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} 
\sigma^2 \rho \left(1 + (t_i - t_j)^\beta\right)^{-\gamma} & \text{for } t_i \neq t_j \\
\sigma^2 (1 - \rho) & \text{for } t_i = t_j.
\end{cases}
\]  

(4.7)

(c) Gaussian model,

\[
\text{Cov}(y_{t_i}, y_{t_j}) = \begin{cases} 
\sigma^2 \rho \exp\{-\beta^2 (t_i - t_j)^2\} & \text{for } t_i \neq t_j \\
\sigma^2 (1 - \rho) & \text{for } t_i = t_j.
\end{cases}
\]  

(4.8)

Table (4.1) shows the efficiencies of the locally optimal design \(\xi^*\) with respect to the designs \(\xi^*_{\text{dag}}, \xi^*_{\text{ca}}\) and \(\xi^*_{\text{ga}}\), which have been calculated assuming covariance structures (4.6), (4.7) and (4.8), respectively. The similar behavior of the correlation structures (Figure 4.5) allows us to compare the designs obtained with them. The efficiency is not substantially affected by the choice of these two correlation structures (always over 75%).

<table>
<thead>
<tr>
<th>KL-optimal design</th>
<th>eff((\xi^*))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\xi^*_{\text{dag}}) ({145, 155, 170, 180, 250, 300})</td>
<td>87%</td>
</tr>
<tr>
<td>(\xi^*_{\text{ca}}) ({50, 60, 70, 210, 260, 330})</td>
<td>75%</td>
</tr>
<tr>
<td>(\xi^*_{\text{ga}}) ({30, 60, 80, 90, 110, 240})</td>
<td>77%</td>
</tr>
</tbody>
</table>

Table 4.1: Designs based on covariance structures (4.6), (4.7) and (4.8) and their corresponding efficiencies with respect \(\xi^*\). The six middle points are shown before summing up \(t_1\)

![Figure 4.1](image)

Figure 4.1: Relative efficiencies of the design \(\xi^*\) for different values of the correlation and nugget parameters
Discrimination between growth models

Figure 4.2: Relative efficiencies of the design $\xi^*$ for different values of the correlation and nugget parameters

Figure 4.3: Relative efficiencies of the design $\xi^*$ for different values of the correlation and nugget parameters

Figure 4.4: Relative efficiencies of the design $\xi^*$ for different values of the correlation and nugget parameters
Figure 4.5: Plot of the correlation structures (4.3), (3.17), (3.18) and (3.19)
Discrimination between growth models
Conclusions

This work is mainly focused on practical aspects. Apart from the introduction of some theoretical methodology we also wanted to apply the optimal design theory to three practical scenarios.

The biomechanical model was derived to provide a quantitative basis for the detection of BPPV. This model is based on a maneuver consisting of two consecutive head turns which are the most common head movements leading to vertigo symptoms. We would like to remark that although the model can only be applied for this specific maneuver, it could be also extended to other types. The experiment, that is, the duration and angle of the head movements to be applied to the patients should be based on the design provided by the D-optimal design since it helps us estimate the parameters simultaneously, thereby minimizing the confidence ellipsoid. The c-optimal design is used either for estimating linear combinations of the parameters, or for estimating the parameters separately. In this case, c-optimal design also provides valuable assistance to check how efficient the D-optimal design is for the estimation of each of the parameters. This provides an interesting check of the sensitivity, since a D-optimal design could be quite efficient for estimating a particular parameter but quite inefficient for estimating another one.

The covariance matrix of the estimates is asymptotically proportional to the inverse of the Fisher information matrix. Theoretically, the application of the maneuvers which are specified in the design, along with their corresponding proportions, will assure that an objective function of the covariance matrix of the estimators

$$\Sigma_{\hat{\theta}} = \begin{pmatrix} \text{var } \hat{\theta}_1 & \text{cov } (\hat{\theta}_1, \hat{\theta}_2) \\ \text{cov } (\hat{\theta}_1, \hat{\theta}_2) & \text{var } \hat{\theta}_2 \end{pmatrix} \propto M^{-1}(\xi, \theta),$$

will be minimized. Symbol $\propto$ stands for "asymptotically proportional" in this case. We would also like to point out that, as far as we know, the models
found in the published works describing these sorts of maneuvers have not been validated with data yet. Clinicians hold that the extra volume of endolymph displaced by the otoliths are directly related to the eye movements provoked in the patient under vertigo symptoms. Therefore, to validate the model, response $y$ should be measured through a variable related to eye movement.

The study conducted in Chapter 3 provides the farm with a restricted design which helps estimate the unknown parameters in an optimal way (following D-optimality). This will be used to keep the heifer’s weight under control. On the other hand, it is important to point out that the results obtained cannot be extended to other areas of Spain or Europe; not even to other Friesian dairy farms, due to the wide variability of this breed. This means a local fitting has to be performed and used for each individual farm, but the procedure applied remain the same.

Although the West model provides a good fit, other growth models such as Brody, Gompertz or Logistic could have been used. For that purpose, optimal criteria will be considered in further research to discriminate between those models (Atkinson and Fedorov, 1975; López-Fidalgo et al., 2007).

The introduction of the nugget effect to the correlation structure has avoided the collapse of some optimal design points which has led to a design without repeated points. The exponential correlation used here is one of the most common within this framework, but others may also be suitable. We have performed a robustness analysis in order to show the importance of a right choice of the structure of correlation. This work provides the methodology, which can be used for any correlation structure, as well as the growing model. The efficiency of the design used in practice with respect to our design is around 80%. This means, design $\xi^*_1$ may save 20% of the observations to get the same results as $\xi_f$.

We have performed a simulation to implement the design $\xi^*_1$ at the dairy farm. The number of simulations was 10,000 and for each of the simulations, 100 observations were generated for each time-point. The mean of the estimators of $\theta$ obtained in the simulations was $(43.41, 750.2, 0.082, 4.1, 0.95)$, which is very near the nominal values (3.13). This means the estimator is almost unbiased. On the other hand, the value of the determinant of the variance-covariance
matrix, $\Sigma_\theta$, is quite similar to the inverse of the information matrix. Thus, its use to compute optimal designs is justified. We have considered a well-known and widely used model for the trend of the growing weights of animals.

The example considered in this chapter (in which constant variance is assumed) agrees with the usual treatment of the problem in the literature (West et al., 2001; Zanton and Heinrichs, 2008). Nevertheless, a non–constant variance may be considered, breaking the property of isotropy, or introducing something similar to the nugget effect. This has been considered recently by Boukouvalas et al. (2014), where the efficiency for repeated measurements will be higher under the presence of heteroscedastic variance.

Another way of dealing with the correlation of the observations would be through mixed models with random coefficients. There is an increasing interest in finding optimal designs for regression models with random effects, see e.g. Tommasi et al. (2014) for a recent example. In a different context, Goos and Jones (2011) considered models with random effects, but always tried to find the covariance structure behind them (e.g. in pages 155 or 270). Additionally the distinction between fixed and random effects is sometimes not clear (see e.g. McCulloch and Searle, 2001, examples in Chapter 1). Furthermore, Demidenko (2004) devotes Chapter 4 to a growth model analyzing different cases for the covariance matrix showing that our approach is not unusual in this context. As such mixed models are gaining in popularity and deserve further research in optimal experimental design.

Finally, in Chapter 4 we have computed a restricted optimal design for discriminating between two well-known and widely used models for the trend of the growing weights of animals. We would like to remark that although this chapter is focused on discriminating between these two models, a decision-making problem with more than two models may be considered in further research.

The criterion used in Section 2 of this chapter generalizes the T-optimality criterion for correlated observations. In this case, the results obtained cannot be extended either to other areas of Spain or Europe; not even to other Limousine farms due to the wide variability of this breed. Furthermore, the design depends on the prior values of the parameters of the model assumed to be the true one. This means a local fitting has also to be performed. Moreover, at the
Navalázar farm, the calves are sent to the abattoir when they are 18 months old. Not every beef farm operates this way, since, despite the quality of the meat being high at this age, it is also more expensive and more difficult to place this product on the market. Therefore, we should not extrapolate from these particular outcomes to other farms.

As in previous chapter, the introduction of the nugget effect to the correlation structure has avoided the collapsing of some optimal design points, which has led to an optimal design without replications. The efficiency of the design used in practice (measurements taken at the growing facility), with respect to the computed design $\xi^{\text{ast}}$ is around 66%, which is substantially lower. Thus, the restricted optimal design computed implies an important gain with respect to the traditional one. The choice of a robust correlation structure is an important contribution, since there is not much literature on optimal design for discrimination between models in the context of correlated observations. This work therefore provides a methodology that can be used for any correlation structure.

As mentioned above, the design computed here is mainly for discrimination between two rival models. Several issues arise in this respect:

(a) Optimality criteria to estimate the parameters of undertaken predictions may be considered using compound criteria for both purposes (May and Tommasi, 2014).

(b) If the choice has to be made between more than two rival models, different criteria derived from KL-optimality may be used. The authors are currently working on this topic.

(c) The designs computed or mentioned here are for statistical inference purposes. In practical terms, the farmer would like to know what the best times for an optimal control of the weight are. This is not exactly the same topic, although very much related.

Finally, we would like to remark that the realization of these designs would comply both with Spanish and the European regulations and would not affect to the animals welfare since this implementation only implies weighing the heifers at different points in time from those already carried out at the growing
facility under normal circumstances.
Bibliography


[80] Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika, 12*, 1.


How to apply the Elving’s method

Let the regression model be

\[ y = \beta e^{-x} + \gamma e^x, \quad x \in \chi = [0, 1]. \]

We are interested in estimating a linear combination of the parameters, say \( c^T \theta \), where \( \theta = (\beta, \gamma) \). In this case, c-optimal designs for three concrete combinations are going to be calculated, that is,

\[ c = (0, 1)^T, \ c = (1, 0)^T, \ c = (1, 1)^T \text{ and } c = (1, 2)^T. \]

We will calculate the convex hull of the set \( f(\chi) \cup -f(\chi) \), where \( f(x) = (e^{-x}, e^x) \). Let the curve in parametric form

\[
\begin{align*}
x(t) &= e^{-t} \\
y(t) &= e^t
\end{align*}
\]

\( t \in [0, 1] \).

Figure A.1 represents the Elfving’s set for the function \( f(x) = (x(t), y(t)) \). Depending on what linear combination we take, we will obtain different designs. If the prolongation of the vector \( c \) cuts the segment, we will have a two-points design. Otherwise the design will only have one point.

**Case a) \( c = (0, 1)^T \)**

The design is

\[
\xi = \begin{bmatrix} 0 & 1 \\ p & 1 - p \end{bmatrix},
\]
In order to calculate the value of \( p \), we have to solve the following equation

\[
(0, a^*) = (1 - p)(e^{-1}, e^1) + p(-e^0, -e^0),
\]

which has solution \( p = \frac{1}{1 + e} \).

Therefore, the design is

\[
\xi = \begin{cases} 
0 & 1 \\
\frac{1}{1 + e} & \frac{1}{1 + e}
\end{cases},
\]

**Case b)** \( c = (0, 1)^T \)

The design is

\[
\xi = \begin{cases} 
0 & 1 \\
1 - q & q
\end{cases},
\]

where \( q \) is the solution of the equation

\[
(0, b^*) = (1 - p)(e^0, e^0) + p(-e^{-1}, -e^1),
\]

Therefore, the design is

\[
\xi = \begin{cases} 
0 & 1 \\
\frac{e}{1 + e} & \frac{1}{1 + e}
\end{cases}.
\]

**Case c)** \( c = (1, 1)^T \)

In this case, the \( c \)-optimal is the one-point design

\[
\xi = \begin{cases} 
0 \\
1
\end{cases}.
\]
Case d) $c = (1, 2)^r$

In this case, the $c$-optimal is the one-point design

$$\xi = \begin{cases} d^* \\ 1 \end{cases},$$

where $t^*$ is the value which satisfies

$$(x(t^*), y(t^*)) = d^*.$$
APPENDIX A. HOW TO APPLY THE ELVING’S METHOD
Navier-Stokes equations

The Navier-Stokes equations are a set of second-order partial differential equations relating first and second derivatives of fluid velocity. These equations deal mainly with the relationship between the properties of pressure \( p \) and velocity \( u \). We are going to apply the law of conservation of mass to a fluid.

If we have an arbitrary volume \( V \). Let \( dS \) be a surface element on the surface \( \partial V \) of \( V \). On \( dS \), the density \( \rho \) of the fluid is constant, as is the velocity \( u \). If \( \eta_s \) is the outward unit normal to \( dS \), then in a time \( dt \), the volume of fluid that flows through \( dS \) is \( u \cdot \eta_s \, dS \, dt \), so the mass that leaves through \( dS \) is

\[
\frac{dm}{dt} = \rho \, u \cdot \eta_s \, dS.
\]

Integrating over the surface and using the divergence theorem, we have the entire mass flux across \( \partial V \) equal to the rate of change of the mass within \( V \):

\[
\frac{dm}{dt} = -\int_{\partial V} \rho \, u \cdot \eta_s \, dS = -\int \int \int_V \left[ \nabla \cdot (\rho \, u) \right] \, dv.
\]

But

\[
m = \int \int \int_V \rho \, dv \implies \frac{dm}{dt} = \int \int \int_V \frac{d\rho}{dt} \, dv,
\]

so this implies

\[
0 = \int \int \int_V \left[ \frac{d\rho}{dt} + \nabla \cdot (\rho \, u) \right] \, dv.
\]

This is true for every \( V \), so we have that in everywhere

\[
\frac{d\rho}{dt} + \nabla \cdot (\rho \, u) = 0. \quad \text{(B.1)}
\]
In this work we have assumed incompressibility, that is, given a container, it is impossible to ever pack more fluid into it or take fluid out without changing the volume. That means the density is constant. This is equivalent to saying that $\nabla \cdot u = 0$ (B.1).

For a body, the stress at a point acting on a planar surface can be decomposed twice: the stress vector, into 3 components, and the surface into three planes each perpendicular to an axis. The $i^{th}$ component ($i = 1, \ldots, 3$) of the stress $T$ on a surface with unit normal $n_j$ ($j = 1, \ldots, 3$), can therefore be described by $T_i = \sigma_{ij} n_j$, where $\sigma_{ij}$ is a second-degree tensor known as the stress tensor.

If we consider an incompressible, viscous fluid subject to an external body force $f$ and a volume element $dv$, the total body force acting in the $x_i$ direction on $dv$ is due to $f_i$ and to forces caused by the stress tensor $\sigma_{ij}$. Let consider these stress forces. At $x$, the component $\sigma_{ij}$ of the stress tensor represents the force per unit area in the $x_i$ direction acting at a point on a plane normal to the $x_j$ direction. There are three components of the stress tensor pointing in the $x_i$ direction: $\sigma_{ij}$. Each of these stresses varies over the element $dv$ in the $x_j$ direction by some small amount, $\partial \sigma_{ij}$. Since $dv$ has side lengths of $dx_i$, the rate of stress variation is thus $\partial \sigma_{ij} / \partial x_i$ (see Fig. B.1). This represents a body force on $dv$. Therefore, in the $x_i$ direction, the total force on $dv$ is

$$F_i = f_i \ dv + \sum_j \frac{\partial \sigma_{ij}}{\partial x_i} \ dv.$$ 

The momentum of the above differential volume element in the $x_i$ direction is

$$P_i = \rho v_i \ dv.$$ 

Using the Newton’s second law, $F = \frac{dP}{dt}$,

$$f_i \ dv + \sum_j \frac{\partial \sigma_{ij}}{\partial x_j} \ dv = \rho \left( \frac{\partial u_i}{\partial t} + \sum_j u_j \frac{\partial u_i}{\partial x_j} \right) \ dv.$$ 

Integrating over a volume $V$ and considering the integrand is continuous and
the volume is arbitrary, we have the Cauchy momentum equation,

$$f_i + \sum_j \frac{\partial\sigma_{ij}}{\partial x_j} = \rho \left( \frac{\partial u_i}{\partial t} + \sum_j u_j \frac{\partial u_i}{\partial x_j} \right). \tag{B.2}$$

Since the endolymph is a newtonian fluid, the stress on a volume element is linearly related to the velocity gradient, the proportionality constant being the viscosity ($\mu$). If we consider a fluid element $dv$, the stresses acting on it can be broken down into two components: a normal uniform stress, known as pressure (non-deviatoric stress), which is the average of all the normal stresses

$$p = -\frac{1}{3} \sum \sigma_{ii},$$

and a deviatoric stress $\tau_{ij} = \sigma_{ij} - p \delta_{ij}$ (matrix of the tensor with zeros on the diagonal). The deviatoric stress is responsible for the deformation of the volume element, and is therefore related to the velocity gradient.

Here, we apply the assumption that our fluid is Newtonian: the deviatoric stress tensor is proportional to the rate of strain tensor, or $\tau_{ij} = \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$. 

![Stress rate variation over an element $dv$ of $V$.](image-url)
This permits us to simplify the Cauchy momentum equation, and in particular the second term on the left side,

\[
\sum_j \frac{\partial \sigma_{ij}}{\partial x_j} = \sum_j \frac{\partial}{\partial x_j} \left( \tau_{ij} + p \delta_{ij} \right) = \sum_j \frac{\partial}{\partial x_j} \left( \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right) = \frac{\partial p}{\partial x_i} + \mu \sum_j \frac{\partial}{\partial x_j} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \frac{\partial p}{\partial x_i} + \mu \sum_j \left( \frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial^2 u_j}{\partial x_j \partial x_i} \right) = \frac{\partial p}{\partial x_i} + \mu \nabla^2 u_i + \mu \frac{\partial}{\partial x_i} \nabla \cdot u = \nabla p + \mu \nabla^2 u_i,
\]

having used, variously, the continuity of second derivatives and the fact that our fluid is incompressible.

Setting density to 1 for convenience and substituting back into the momentum equation, we finally have the Navier-Stokes equations for an incompressible Newtonian fluid,

\[
\frac{\partial u_i}{\partial t} + \sum_j u_j \frac{\partial u_i}{\partial x_j} = f_i + \mu \nabla^2 u_i - \frac{\partial p}{\partial x_i} + \sum_j \frac{\partial u_j}{\partial x_j} = 0.
\]

One can also write the above Navier-Stokes equations more succinctly and in more familiar form as follows,

\[
\begin{align*}
\nabla \cdot u &= 0 & \text{Mass Continuity equation} \\
\frac{\partial u}{\partial t} + u \cdot \nabla u &= f + \mu \nabla^2 u - \nabla p & \text{Cauchy Momentum equations}
\end{align*}
\]
Appendix C

Dix-Hallpike and Epley maneuvers

In this appendix we describe how the Dix-Hallpike and Epley maneuvers are performed. Dix-Hallpike test is used by the clinicians to diagnose the pathology whereas Epley test allows free floating particles from the affected semicircular canal to be relocated.

C.1 Dix-Hallpike maneuver

This is a diagnostic maneuver used to identify benign paroxysmal positional vertigo. It was developed by Dix and Hallpike (1952). As can be observed in Figure C.1, it consists in moving the patient’s head in a way to trigger vertigo’s symptoms. It is described as follows:

- The patient is brought from a sitting to a supine position, with the head turned 45 degrees to one side. The clinician then helps the patient quickly lie back down, so that the head hangs backward over the end of the table, at an incline of about 20 degrees.

- The clinician watches the eyes for nystagmus. The timing and appearance of the eye movements will identify the cause of vertigo as either the inner ear or the brain.

- After the patient sits upright for a few minutes to recover from the vertigo, the procedure is repeated with the head turned in the opposite direction.

A positive Dix-Hallpike tests consists of a burst of nystagmus (jumping of the eyes). It is likely that the vertigo is caused by an inner ear or brain problem, depending on the way the patient reacted to the test.
C.2 Epley maneuver

This is a repositioning maneuver. First time described by Epley (1980). As observed in Figure C.2, it consists of 4 turns. The aim of this maneuver is to relocate the floating particles. It consists in four movements:

- The clinician will move the patient approximately 45 degrees in the direction of the affected ear.

- The clinician will tilt the patient backward so that he is both maintaining the original 45-degree position and leaning backward horizontally. The clinician continues to hold the head until the vertigo stops.

- After the vertigo stops, the head is turned about 90 degrees in the direction of the unaffected ear.

- The clinician helps the patient back into a seated position and tilts the head down 30 degrees.

Figure C.1: Dix-Hallpike maneuver.
Figure C.2: Epley maneuver.
APPENDIX C. DIX-HALLPIKE AND EPLEY MANEUVERS
Analytical primitive for equation (3.4)

Let the equation 3.4 be

\[
\frac{dm}{dt} = a m^{3/4} \left[ 1 - \left( \frac{m}{M} \right)^{1/4} \right].
\]

Integrating in both sides,

\[
\int \frac{dm}{m^{3/4} \left[ 1 - \left( \frac{m}{M} \right)^{1/4} \right]} = \int a \, dt = at \quad (*).
\]

By doing

\[x = 1 - \left( \frac{m}{M} \right)^{1/4} \Rightarrow dx = -\frac{1}{4} M^{1/4} m^{-3/4} \, dm.\]

Therefore,

\[
\int \frac{dm}{m^{3/4} \left[ 1 - \left( \frac{m}{M} \right)^{1/4} \right]} = -\frac{1}{4} M^{1/4} \int \frac{1}{x} \, dx = -\frac{1}{4} M^{1/4} \ln(|x|) + x_0
\]

\[= -\frac{1}{4} M^{1/4} \ln(|1 - \left( \frac{m}{M} \right)^{1/4}|) + x_0.\]

From (*) we have

\[-\frac{1}{4} M^{1/4} \ln(|1 - \left( \frac{m}{M} \right)^{1/4}|) + x_0 = at,
\]
APPENDIX D. ANALYTICAL PRIMITIVE FOR EQUATION (3.4)

and consequently,

\[
\exp\{\ln(1 - \left(\frac{m}{M}\right)^{1/4}) + x_0\} = \exp\left\{\frac{-at}{4M^{1/4}}\right\}.
\]

Therefore,

\[
\left(\frac{m}{M}\right)^{1/4} = 1 - \frac{1}{x_0} \exp\left\{\frac{-at}{4M^{1/4}}\right\}.
\]

For \(t = 0\), \(m\) becomes \(m_0\) and \(\frac{1}{x_0} = 1 - \left(\frac{m_0}{M}\right)^{1/4}\). Resulting from this we get the model (3.5),

\[
\left(\frac{m}{M}\right)^{1/4} = 1 - \left[1 - \left(\frac{m_0}{M}\right)^{1/4}\right] \exp\left\{\frac{-at}{4M^{1/4}}\right\}.
\]
FIM for correlated observations

The derivation of $M(\xi, \theta)$ is made in detail to the more general case, that is, when the mean of the response and the covariance structure may share common parameters (Pázman, 2004).

The information matrix for any design $\xi$ is equal to

$$M(\xi, \theta) = E_y \left[ -\frac{\partial^2 \log f(y|t, \theta)}{\partial \theta^2} \right],$$

and, assuming Gaussianity, the negative of the first order derivative of the log-likelihood function being

$$-\frac{\partial \log f(y|t, \theta)}{\partial \theta_i} = \frac{1}{2} \left\{ \frac{\partial}{\partial \theta_i} \left[ (y - \eta(t, \theta))' \Sigma^{-1}(\theta)(y - \eta(t, \theta)) \right] + \frac{\partial}{\partial \theta_i} \left[ \log \det \Sigma(\theta) \right] \right\}.$$

The first term is

$$-2(y - \eta(t, \theta))' \Sigma^{-1}(\theta) \frac{\partial \eta(t, \theta)}{\partial \theta_i} - (y - \eta(t, \theta))' \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \Sigma^{-1}(\theta)(y - \eta(t, \theta)),$$

and the second,

$$\frac{\partial \det \Sigma(\theta)}{\partial \theta_i} = \frac{\text{tr} \left[ \text{adj}(\Sigma(\theta)) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \right]}{\det \Sigma(\theta)} = \text{tr} \left[ \frac{\text{adj}(\Sigma(\theta))}{\det \Sigma(\theta)} \frac{\partial \Sigma(\theta)}{\partial \theta_i} \right] = \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \right].$$
The first term of the second order derivative is

\[
\frac{\partial}{\partial \theta_i} \frac{\partial}{\partial \theta_j} \left[ \frac{1}{2} (y - \eta(t, \theta))^\prime \Sigma^{-1}(\theta) (y - \eta(t, \theta)) \right] = \left( \frac{\partial \eta(t, \theta)}{\partial \theta_j} \right)^\prime \Sigma^{-1}(\theta) \frac{\partial \eta(t, \theta)}{\partial \theta_i} +
\]

\[
(y - \eta(t, \theta))^\prime \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_j} \Sigma^{-1}(\theta) \frac{\partial \eta(t, \theta)}{\partial \theta_i} - (y - \eta(t, \theta))^\prime \Sigma^{-1}(\theta) \frac{\partial \eta(t, \theta)}{\partial \theta_i} \frac{\partial \eta(t, \theta)}{\partial \theta_j} +
\]

\[
(y - \eta(t, \theta))^\prime \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \Sigma^{-1}(\theta) \frac{\partial \eta(t, \theta)}{\partial \theta_j} -
\]

\[
\frac{1}{2} (y - \eta(t, \theta))^\prime \Sigma^{-1}(\theta) \frac{\partial \eta(t, \theta)}{\partial \theta_i} \frac{\partial \eta(t, \theta)}{\partial \theta_j} \Sigma^{-1}(\theta) (y - \eta(t, \theta)).
\]

and the second,

\[
\frac{\partial}{\partial \theta_i} \frac{\partial}{\partial \theta_j} \left[ \frac{1}{2} \log \det \Sigma(\theta) \right] = -\frac{1}{2} \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_j} \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \right]
\]

\[
+ \frac{1}{2} \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma}{\partial \theta_i} \frac{\partial \Sigma}{\partial \theta_j} \right].
\]

Considering that

\[
E[(y - \eta(t, \theta))] = 0, \ E[(y - \eta(t, \theta))^\prime (y - \eta(t, \theta))] = \Sigma(\theta),
\]

and for any vector \( x \) and any symmetric matrix \( A \),

\[
E[x'Ax] = \text{tr}[A \Sigma],
\]

the information matrix, obtained by calculating the expected value of the sum of the first and second term of the second order derivative, is expressed as
follows,
\[
(M(\xi, \theta))_{ij} = \frac{\partial \eta'(t, \theta)}{\partial \theta_j} \frac{\partial \eta(t, \theta)}{\partial \theta_i} + \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_j} \right] \\
- \frac{1}{2} \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i \partial \theta_j} \right] - \frac{1}{2} \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_j} \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \right] \\
+ \frac{1}{2} \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i \partial \theta_j} \right]
\]
\[
= \frac{\partial \eta'(t, \theta)}{\partial \theta_j} \Sigma^{-1}(\theta) \frac{\partial \eta(t, \theta)}{\partial \theta_i} + \frac{1}{2} \text{tr} \left[ \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_i} \Sigma^{-1}(\theta) \frac{\partial \Sigma(\theta)}{\partial \theta_j} \right].
\]
Friesian and Limousine breeds

F.1 Friesian

F.1.1 history

Originating in Europe, Friesians were bred in what is now the Netherlands and more specifically in the two northern provinces of North Holland and Friesland, and northern Germany. The original stock were the black animals and white animals of the Batavians and Friesians, migrant European tribes who settled in the Rhine Delta region about 2000 years ago. For many years, Friesians were bred and strictly culled to obtain animals which would make best use of grass, the area’s most abundant resource. The intermingling of these animals evolved into an efficient, high-producing black-and-white dairy cow. After World War II the Friesian was selected for a high-yielding, large-framed, single-purpose dairy cow with an exceptionally good udder. Not only did Europe convert to Friesian dairy cattle, but the conversion spread all over the world, wherever the climate and conditions are suited for European dairy cattle.

F.1.2 Morphology

Friesians are a dairy breed. The selection performed by the Dutch led to a very beautiful animal. The spanish Friesian is black and white colour with a quite long head and shows more bone and less fat and muscle because they have been specifically selected to produce milk. Dairy cattle are always thinner and more angular than beef cattle. The cow is about 150-160 cm high at the withers, weighing approximately 650-750 kg.
F.1.3 Productivity

This breed is the world’s leading milk producers. In Spain the population of cows in dairy farms is about 1200000. From that population, around 97% are Friesians. The spanish breed produces an average of 6700 liters per year with 3.36% of milk fat and 3.11% of protein.

F.2 Limousine

F.2.1 Breed history

These cattle are native to the western part of the Massif Central in France, from the central and south-western part of the country, in the region of Limousine. These areas are characterized by significant precipitation and poor soils. In these unfavourable conditions, the Limousine breed has been developing for thousands of years. As a result of their environment, evolved into a breed of unusual sturdiness, health and adaptability. During the early times of animal power, Limousine gained a well-earned reputation as work animals in addition to their beef qualities, being universally renown and esteemed both as beasts of burden and beef cattle. At the end of their work life these animals were then fattened for slaughter. Since those early days the breed has developed from a working meat animal into a highly specialized beef producing animal and are referred to as the ”butcher's animal”. Nowadays, the Limousine breed is common throughout the world, and in numerous countries the population of this beef breed occupies the leading position.

F.2.2 Morphology

Limousine cattle, as a result of their environment, evolved into a breed of unusual sturdiness, health and adaptability. It is a breed of highly muscled. This golden-red cattle is characterized by a large size, a uniform and bright coat, a small and short head with a broad forehead, being the neck short. Its chest is broad and rounded. The amount of fat and bones in the body is low in comparison to other breeds. Cows weigh on average 650 kg and are over
F.2. LIMOUSINE

135 cm high at the withers. Bulls weigh over 900 kg and are about 145 cm high at the withers.

F.2.3 Productivity

Limousine is a very economically profitable beef cattle. These animals can make good use of feeds provided at the growing facility and are also known for high fertility and splendid results in the scope of slaughter capacity. A significant advantage of this breed is longevity, the cows often bring up between and ten and twenty calves during their lifetime. It produces beef with a low proportion of bone and fat, a top killing-out percentage and a high yield of saleable red meat. The carcass yield is around 65%. Limousine cattle is considered the most efficient and fastest of all breeds at converting feed into saleable meat even though Limousine’s live weight growth was the slowest. Limousine’s ability to adapt to different environments contributed greatly to the breed’s current success. Today, the breed is present in about 70 countries around the world, and in all latitudes ranging from Finland in the north to South Africa in the south.

Figure F.1: Hostein-Friesian cattle.
Figure F.2: Limousine cattle.
# Appendix G

## Data provided by the beef farm

The following data were provided by Navalázar farm. Each row corresponds to the weight of a single calf at its corresponding age. The first column refers to the weight of the calf upon arrival at the growing facility, with the last column referring to its weight before being sent to the abattoir.

<table>
<thead>
<tr>
<th>calf</th>
<th>day</th>
<th>kg</th>
<th>day</th>
<th>kg</th>
<th>day</th>
<th>kg</th>
<th>day</th>
<th>kg</th>
<th>day</th>
<th>kg</th>
<th>day</th>
<th>kg</th>
<th>day</th>
<th>kg</th>
<th>day</th>
<th>kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>170</td>
<td>215</td>
<td>236</td>
<td>291</td>
<td>277</td>
<td>337</td>
<td>333</td>
<td>402</td>
<td>389</td>
<td>478</td>
<td>434</td>
<td>539</td>
<td>493</td>
<td>617</td>
<td>537</td>
<td>677</td>
</tr>
<tr>
<td>2</td>
<td>170</td>
<td>160</td>
<td>238</td>
<td>266</td>
<td>285</td>
<td>333</td>
<td>344</td>
<td>414</td>
<td>386</td>
<td>459</td>
<td>436</td>
<td>502</td>
<td>494</td>
<td>537</td>
<td>534</td>
<td>565</td>
</tr>
<tr>
<td>3</td>
<td>170</td>
<td>184</td>
<td>229</td>
<td>298</td>
<td>276</td>
<td>366</td>
<td>338</td>
<td>465</td>
<td>383</td>
<td>517</td>
<td>443</td>
<td>596</td>
<td>493</td>
<td>652</td>
<td>541</td>
<td>720</td>
</tr>
<tr>
<td>4</td>
<td>170</td>
<td>161</td>
<td>228</td>
<td>228</td>
<td>287</td>
<td>299</td>
<td>340</td>
<td>367</td>
<td>386</td>
<td>423</td>
<td>440</td>
<td>491</td>
<td>496</td>
<td>555</td>
<td>531</td>
<td>588</td>
</tr>
<tr>
<td>5</td>
<td>171</td>
<td>209</td>
<td>226</td>
<td>307</td>
<td>288</td>
<td>422</td>
<td>324</td>
<td>477</td>
<td>394</td>
<td>579</td>
<td>441</td>
<td>631</td>
<td>488</td>
<td>684</td>
<td>541</td>
<td>731</td>
</tr>
<tr>
<td>6</td>
<td>171</td>
<td>196</td>
<td>233</td>
<td>276</td>
<td>286</td>
<td>344</td>
<td>338</td>
<td>417</td>
<td>393</td>
<td>500</td>
<td>444</td>
<td>560</td>
<td>493</td>
<td>619</td>
<td>541</td>
<td>655</td>
</tr>
<tr>
<td>7</td>
<td>171</td>
<td>187</td>
<td>224</td>
<td>261</td>
<td>286</td>
<td>327</td>
<td>330</td>
<td>374</td>
<td>393</td>
<td>434</td>
<td>429</td>
<td>469</td>
<td>498</td>
<td>524</td>
<td>540</td>
<td>543</td>
</tr>
<tr>
<td>8</td>
<td>171</td>
<td>245</td>
<td>227</td>
<td>301</td>
<td>288</td>
<td>374</td>
<td>338</td>
<td>443</td>
<td>378</td>
<td>502</td>
<td>429</td>
<td>607</td>
<td>480</td>
<td>710</td>
<td>530</td>
<td>838</td>
</tr>
<tr>
<td>9</td>
<td>171</td>
<td>149</td>
<td>224</td>
<td>252</td>
<td>276</td>
<td>348</td>
<td>328</td>
<td>430</td>
<td>382</td>
<td>505</td>
<td>433</td>
<td>560</td>
<td>482</td>
<td>619</td>
<td>535</td>
<td>654</td>
</tr>
<tr>
<td>10</td>
<td>171</td>
<td>225</td>
<td>240</td>
<td>329</td>
<td>285</td>
<td>396</td>
<td>328</td>
<td>452</td>
<td>383</td>
<td>537</td>
<td>429</td>
<td>583</td>
<td>483</td>
<td>667</td>
<td>535</td>
<td>719</td>
</tr>
<tr>
<td>11</td>
<td>172</td>
<td>163</td>
<td>224</td>
<td>253</td>
<td>286</td>
<td>358</td>
<td>334</td>
<td>429</td>
<td>379</td>
<td>495</td>
<td>437</td>
<td>583</td>
<td>481</td>
<td>646</td>
<td>531</td>
<td>718</td>
</tr>
<tr>
<td>12</td>
<td>172</td>
<td>240</td>
<td>224</td>
<td>308</td>
<td>276</td>
<td>394</td>
<td>336</td>
<td>488</td>
<td>387</td>
<td>567</td>
<td>428</td>
<td>625</td>
<td>497</td>
<td>717</td>
<td>544</td>
<td>768</td>
</tr>
<tr>
<td>13</td>
<td>172</td>
<td>168</td>
<td>225</td>
<td>284</td>
<td>274</td>
<td>371</td>
<td>334</td>
<td>458</td>
<td>379</td>
<td>512</td>
<td>445</td>
<td>568</td>
<td>480</td>
<td>597</td>
<td>530</td>
<td>631</td>
</tr>
<tr>
<td>14</td>
<td>172</td>
<td>206</td>
<td>229</td>
<td>303</td>
<td>278</td>
<td>385</td>
<td>338</td>
<td>468</td>
<td>381</td>
<td>518</td>
<td>445</td>
<td>574</td>
<td>495</td>
<td>612</td>
<td>537</td>
<td>640</td>
</tr>
<tr>
<td>15</td>
<td>173</td>
<td>174</td>
<td>229</td>
<td>288</td>
<td>289</td>
<td>394</td>
<td>340</td>
<td>483</td>
<td>376</td>
<td>536</td>
<td>435</td>
<td>647</td>
<td>489</td>
<td>713</td>
<td>538</td>
<td>784</td>
</tr>
<tr>
<td>16</td>
<td>173</td>
<td>218</td>
<td>238</td>
<td>312</td>
<td>285</td>
<td>384</td>
<td>331</td>
<td>453</td>
<td>379</td>
<td>511</td>
<td>435</td>
<td>576</td>
<td>488</td>
<td>634</td>
<td>532</td>
<td>663</td>
</tr>
<tr>
<td>17</td>
<td>173</td>
<td>180</td>
<td>224</td>
<td>251</td>
<td>286</td>
<td>345</td>
<td>326</td>
<td>404</td>
<td>386</td>
<td>490</td>
<td>437</td>
<td>556</td>
<td>495</td>
<td>636</td>
<td>539</td>
<td>678</td>
</tr>
<tr>
<td>18</td>
<td>173</td>
<td>259</td>
<td>229</td>
<td>324</td>
<td>275</td>
<td>375</td>
<td>324</td>
<td>453</td>
<td>394</td>
<td>557</td>
<td>433</td>
<td>627</td>
<td>484</td>
<td>706</td>
<td>536</td>
<td>803</td>
</tr>
<tr>
<td>19</td>
<td>173</td>
<td>199</td>
<td>228</td>
<td>259</td>
<td>273</td>
<td>322</td>
<td>334</td>
<td>403</td>
<td>382</td>
<td>491</td>
<td>430</td>
<td>579</td>
<td>482</td>
<td>677</td>
<td>548</td>
<td>806</td>
</tr>
<tr>
<td>20</td>
<td>173</td>
<td>243</td>
<td>236</td>
<td>338</td>
<td>290</td>
<td>412</td>
<td>331</td>
<td>466</td>
<td>389</td>
<td>535</td>
<td>433</td>
<td>585</td>
<td>489</td>
<td>650</td>
<td>546</td>
<td>714</td>
</tr>
<tr>
<td>21</td>
<td>173</td>
<td>160</td>
<td>222</td>
<td>305</td>
<td>287</td>
<td>440</td>
<td>331</td>
<td>517</td>
<td>388</td>
<td>597</td>
<td>434</td>
<td>666</td>
<td>491</td>
<td>739</td>
<td>530</td>
<td>785</td>
</tr>
<tr>
<td>Calf</td>
<td>Weight</td>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>174.204</td>
<td>239.321</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>174.214</td>
<td>237.296</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>175.186</td>
<td>222.274</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>175.180</td>
<td>237.302</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>175.133</td>
<td>226.234</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>176.230</td>
<td>233.297</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>176.230</td>
<td>227.332</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>176.180</td>
<td>240.328</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>177.228</td>
<td>238.346</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>177.202</td>
<td>225.382</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>177.194</td>
<td>228.398</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>178.194</td>
<td>239.381</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>178.187</td>
<td>222.377</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>178.207</td>
<td>223.384</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>178.215</td>
<td>234.313</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>178.258</td>
<td>230.338</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>179.240</td>
<td>230.326</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>179.241</td>
<td>237.327</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>179.186</td>
<td>228.289</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>180.187</td>
<td>223.289</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>180.213</td>
<td>227.274</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>180.268</td>
<td>235.346</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>180.195</td>
<td>239.272</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>180.230</td>
<td>231.304</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>181.206</td>
<td>230.288</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>181.197</td>
<td>234.331</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>181.193</td>
<td>227.256</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>181.207</td>
<td>227.267</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>181.191</td>
<td>234.310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>182.221</td>
<td>236.297</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>182.204</td>
<td>233.286</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>182.219</td>
<td>240.288</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>182.225</td>
<td>229.292</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>182.250</td>
<td>235.328</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>182.228</td>
<td>232.263</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

APPENDIX G. DATA PROVIDED BY THE BEEF FARM
<table>
<thead>
<tr>
<th>calf</th>
<th>183</th>
<th>229</th>
<th>280</th>
<th>344</th>
<th>376</th>
<th>444</th>
<th>496</th>
<th>545</th>
</tr>
</thead>
<tbody>
<tr>
<td>57</td>
<td>230</td>
<td>357</td>
<td>428</td>
<td>504</td>
<td>549</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>183</td>
<td>232</td>
<td>286</td>
<td>337</td>
<td>379</td>
<td>429</td>
<td>498</td>
<td>533</td>
</tr>
<tr>
<td>58</td>
<td>399</td>
<td>473</td>
<td>523</td>
<td>586</td>
<td>659</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>183</td>
<td>245</td>
<td>283</td>
<td>333</td>
<td>378</td>
<td>436</td>
<td>495</td>
<td>547</td>
</tr>
<tr>
<td>59</td>
<td>459</td>
<td>504</td>
<td>586</td>
<td>667</td>
<td>732</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>183</td>
<td>226</td>
<td>287</td>
<td>326</td>
<td>378</td>
<td>439</td>
<td>485</td>
<td>543</td>
</tr>
<tr>
<td>60</td>
<td>387</td>
<td>445</td>
<td>517</td>
<td>574</td>
<td>655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>183</td>
<td>209</td>
<td>285</td>
<td>325</td>
<td>383</td>
<td>440</td>
<td>497</td>
<td>546</td>
</tr>
<tr>
<td>61</td>
<td>376</td>
<td>442</td>
<td>517</td>
<td>574</td>
<td>655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>184</td>
<td>222</td>
<td>278</td>
<td>337</td>
<td>388</td>
<td>430</td>
<td>480</td>
<td>539</td>
</tr>
<tr>
<td>62</td>
<td>302</td>
<td>377</td>
<td>454</td>
<td>521</td>
<td>592</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>184</td>
<td>213</td>
<td>285</td>
<td>325</td>
<td>383</td>
<td>440</td>
<td>497</td>
<td>546</td>
</tr>
<tr>
<td>63</td>
<td>376</td>
<td>442</td>
<td>517</td>
<td>574</td>
<td>655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td>196</td>
<td>277</td>
<td>332</td>
<td>387</td>
<td>440</td>
<td>497</td>
<td>549</td>
</tr>
<tr>
<td>64</td>
<td>258</td>
<td>305</td>
<td>377</td>
<td>444</td>
<td>503</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td>209</td>
<td>290</td>
<td>344</td>
<td>386</td>
<td>444</td>
<td>490</td>
<td>533</td>
</tr>
<tr>
<td>65</td>
<td>272</td>
<td>473</td>
<td>536</td>
<td>606</td>
<td>664</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td>213</td>
<td>285</td>
<td>339</td>
<td>395</td>
<td>433</td>
<td>495</td>
<td>543</td>
</tr>
<tr>
<td>66</td>
<td>287</td>
<td>381</td>
<td>454</td>
<td>505</td>
<td>566</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td>206</td>
<td>278</td>
<td>329</td>
<td>386</td>
<td>430</td>
<td>489</td>
<td>539</td>
</tr>
<tr>
<td>67</td>
<td>391</td>
<td>476</td>
<td>533</td>
<td>592</td>
<td>655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td>164</td>
<td>281</td>
<td>329</td>
<td>389</td>
<td>431</td>
<td>480</td>
<td>545</td>
</tr>
<tr>
<td>68</td>
<td>322</td>
<td>393</td>
<td>456</td>
<td>512</td>
<td>592</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td>206</td>
<td>282</td>
<td>339</td>
<td>395</td>
<td>433</td>
<td>495</td>
<td>544</td>
</tr>
<tr>
<td>69</td>
<td>329</td>
<td>399</td>
<td>476</td>
<td>533</td>
<td>638</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>208</td>
<td>275</td>
<td>342</td>
<td>388</td>
<td>429</td>
<td>484</td>
<td>540</td>
</tr>
<tr>
<td>70</td>
<td>339</td>
<td>451</td>
<td>518</td>
<td>564</td>
<td>633</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>321</td>
<td>276</td>
<td>329</td>
<td>393</td>
<td>427</td>
<td>488</td>
<td>533</td>
</tr>
<tr>
<td>71</td>
<td>393</td>
<td>574</td>
<td>638</td>
<td>694</td>
<td>781</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>240</td>
<td>287</td>
<td>331</td>
<td>387</td>
<td>433</td>
<td>495</td>
<td>543</td>
</tr>
<tr>
<td>72</td>
<td>381</td>
<td>431</td>
<td>506</td>
<td>555</td>
<td>631</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>243</td>
<td>277</td>
<td>336</td>
<td>387</td>
<td>437</td>
<td>495</td>
<td>535</td>
</tr>
<tr>
<td>73</td>
<td>311</td>
<td>394</td>
<td>476</td>
<td>533</td>
<td>638</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>202</td>
<td>290</td>
<td>332</td>
<td>391</td>
<td>435</td>
<td>494</td>
<td>545</td>
</tr>
<tr>
<td>74</td>
<td>393</td>
<td>574</td>
<td>638</td>
<td>694</td>
<td>781</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>246</td>
<td>275</td>
<td>334</td>
<td>382</td>
<td>438</td>
<td>486</td>
<td>538</td>
</tr>
<tr>
<td>75</td>
<td>377</td>
<td>466</td>
<td>515</td>
<td>570</td>
<td>634</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>187</td>
<td>228</td>
<td>283</td>
<td>327</td>
<td>378</td>
<td>432</td>
<td>495</td>
<td>535</td>
</tr>
<tr>
<td>76</td>
<td>369</td>
<td>422</td>
<td>482</td>
<td>539</td>
<td>626</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>188</td>
<td>234</td>
<td>281</td>
<td>326</td>
<td>380</td>
<td>435</td>
<td>490</td>
<td>535</td>
</tr>
<tr>
<td>77</td>
<td>319</td>
<td>437</td>
<td>515</td>
<td>590</td>
<td>726</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>188</td>
<td>200</td>
<td>273</td>
<td>331</td>
<td>389</td>
<td>441</td>
<td>483</td>
<td>535</td>
</tr>
<tr>
<td>78</td>
<td>367</td>
<td>428</td>
<td>482</td>
<td>539</td>
<td>669</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>188</td>
<td>187</td>
<td>278</td>
<td>335</td>
<td>395</td>
<td>461</td>
<td>495</td>
<td>538</td>
</tr>
<tr>
<td>79</td>
<td>351</td>
<td>424</td>
<td>480</td>
<td>522</td>
<td>586</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>188</td>
<td>221</td>
<td>287</td>
<td>335</td>
<td>376</td>
<td>436</td>
<td>491</td>
<td>540</td>
</tr>
<tr>
<td>80</td>
<td>353</td>
<td>420</td>
<td>472</td>
<td>563</td>
<td>730</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>189</td>
<td>258</td>
<td>287</td>
<td>325</td>
<td>381</td>
<td>428</td>
<td>488</td>
<td>545</td>
</tr>
<tr>
<td>81</td>
<td>370</td>
<td>429</td>
<td>487</td>
<td>524</td>
<td>626</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>189</td>
<td>200</td>
<td>279</td>
<td>337</td>
<td>380</td>
<td>433</td>
<td>492</td>
<td>540</td>
</tr>
<tr>
<td>82</td>
<td>352</td>
<td>430</td>
<td>482</td>
<td>541</td>
<td>631</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>