# Varying Coefficient Models & Multivariate Parameters in Partial Differential Equation Models

Inference and Estimation

**Mohamed Ahkim**

Doctoral dissertation submitted to obtain the degrees of

– Doctor of Science: Mathematics at Universiteit Antwerpen

– Doctor of Science: Mathematics at Universiteit Hasselt

Supervisors:

Prof. Dr Anneleen Verhasselt

Prof. Dr Wim Vanroose

April 2017

# Acknowledgments

I would like to use this opportunity to thank the people who helped me the past four and a half years. It was a long journey.

My thesis would not have been possible without the help of my supervisors Anneleen and Wim. Thank you for guiding me along this journey with good advices and patience. I have learned a lot during our research discussions. Your advices on career perspectives is highly appreciated. Irène, thank you for the nice collaboration on the second paper. I have learned from you during my student years and as a coauthor. I would also like to thank the jury members. Your comments were valuable and lead to an improved readability of my thesis.

I had many colleagues at Universiteit Antwerpen and Universiteit Hasselt. I only have good memories of you. Mieke and Martine, thank you very much for providing administrative support and answering many questions. A huge thank you to my (previous) officemates: Tim, Yves, Aklilu, Nick, Michiel, Tapiwa, Kim and Mohammed. Nick, also thank you for the collaboration on the third project. It was really nice to have you all around, I wish you all the best. Yudhie and Stanislav, I enjoyed our countless discussions and activities when we met at conferences and meetings. You are wonderful people.

Uit mijn thuisomgeving zijn er veel mensen die ik dankbaar ben. Mijn vrienden met wie ik het kan hebben over zowel de serieuze als de minder serieuze dingen van het leven. Ik waardeer enorm jullie steun. Mijn broers en zussen die altijd voor mij klaar staan en vaak mijn toeverlaat zijn. Mijn neefjes en nichtjes die zomaar een lach op mijn gezicht kunnen toveren. En nog velen anderen.

Tenslotte wil ik de twee belangrijkste personen in mijn leven bedanken. Moeder Drifa El Faida and vader Ahmida Ahkim. Woorden kunnen niet uitdrukken hoeveel jullie voor mij betekenen. Het is onmogelijk om maar iets terug te doen van wat jullie voor mij gedaan hebben. Ik hoop dat jullie trots op me zijn.

Mohamed Ahkim, April 2017.

# Abstract

In this thesis we are interested in (unknown) functions which appear in statistical models, and testing procedures concerning these unknown functions. These functions are estimated flexibly (nonparametric) and not according to a prespecified (parametric) form. The nonparametric technique we consider is by spline approximations. Splines are used to estimate univariate as well as multivariate functions. Then, hypothesis testing about those unknown functions is translated to a testing procedure based on the spline estimation.

The first statistical model we consider is a varying coefficient model (VCM), which is an extension of the classical linear regression model in the sense that the regression coefficients are allowed to be functions, for example of time. Varying coefficient models (VCMs) are since many years popular in longitudinal data and panel data studies, and have been applied in fields such as finance, economics, ecology, epidemiology, health sciences, and so on. We estimate the coefficient by B-splines. An important question in a VCM is whether the coefficient has a particular parametric form, such as being constant or linear. This allows, on the one hand to draw conclusions on the effect of certain variables on the response variable. On the other hand, this could allow to propose a simpler model and strongly reduce the number of parameters in the model. We construct testing procedures to answer the former hypothesis, and give the supporting theoretical results for longitudinal data with correlated errors. Testing of such hypotheses in VCMs is studied in Chapter 2, with illustrations of the power through simulations and a data application.

In Chapter 3 we address our second hypothesis of VCMs. There, we are interested in whether a coefficient function is monotonic or convex, i.e. the shape. We develop testing procedures for monotonicity and convexity, with the necessary theoretical results. Moreover, we give procedures to test simultaneously the shapes of certain coefficient functions. The tests use

constrained and unconstrained regression splines. Application of our testing procedures on simulations reveal the effectiveness of our approach. Data applications are also given.

Chapter 4 studies parameters of partial differential equations (PDEs). Many complex dynamic systems are governed by PDEs, they appear in a vast number of scientific fields such as biology, physics and finance. PDEs are determined by their parameters. Often scientists face the challenge to determine unknown parameters of a PDE, and the need to estimate them from error prone measurements. In the statistical literature it is very often assumed that the parameters are constant, which restricts the application possibilities because in reality this assumption can be crude. In Chapter 4 we extend the parametric cascading method- which was shown to be very effective for PDE models with constant parameters- to the PDE setting where the coefficients vary with multiple variables. In the case of a linear PDE model, we show that our proposed estimator of the parameters is uniformly consistent.

In Chapter 1 we introduce further the concepts of this thesis with an overview of the relevant statistical literature. Finally, in Chapter 5 we conclude this thesis with a summary of the results and discuss future research perspectives.

# Abstract in Dutch

In deze thesis zijn we geïnteresseerd in (onbekende) functies die in statistische modellen voorkomen, en toetsingsprocedures omtrent deze onbekende functies. Deze functies worden flexibel geschat (niet-parametrisch) en niet volgens een voorgeschreven vorm (parametrisch). De niet-parametrische techniek die we beschouwen is via spline schattingen. Splines worden gebruikt om zowel univariate als multivariate functies te schatten. Vervolgens worden hypothesetoetsen over die onbekende functies vertaald naar een toetsingsprocedure gebaseerd op de spline schatting.

Het eerste statistisch model dat we beschouwen is een model met variërende coëfficiënten, wat een uitbreiding is van het klassieke lineaire regressie model in de zin dat de regressiecoëfficiënten functies mogen zijn, bijvoorbeeld van tijd. Modellen met variërende coëfficiënten (VCM) zijn sinds vele jaren populair in longitudinale data en paneldata studies, en zijn toegepast in domeinen als financiën, economie, ecologie, epidemiologie, gezondheidswetenschappen, etc. We schatten de coëfficiënten door middel van B-splines. Een belangrijke vraag in VCM is of de coëfficiënten een bepaalde parametrische vorm hebben, zoals constantheid of lineariteit. Dit laat toe om, enerzijds uitspraken te doen over de effecten van covariaten op de respons, anderzijds een simpeler model voor te stellen en het aantal parameters in het model sterk te reduceren. We construeren toetsingsprocedures voor zulke hypothesen, met theoretische onderbouwingen voor longitudinale data met gecorreleerde fouttermen. Zulke hypothesen toetsen in VCM worden bestudeerd in Hoofdstuk 2, met illustraties aan de hand van gesimuleerde data en toepassingen op reële data.

In Hoofdstuk 3 richten we ons tot andere soort hypothesen in VCM. Daar ligt onze interesse in het toetsen van monotoniciteit en convexiteit, d.i. de vorm. We ontwikkelen toetsingsprocedures voor monotoniciteit en convexiteit, met de nodige theoretische funderingen. Bovendien geven we ook procedures om simultaan de vorm van de coëfficiënten te toetsen, wat niet nodig

was in univariate regressiemodellen. Gesimuleerde data onthullen de effectiviteit van onze aanpak. We beschouwen ook een reële data toepassing.

Hoofdstuk 4 bestudeert parameters van modellen met partiële differentiaalvergelijkingen (PDEs). Verschillende complexe dynamische systemen zijn onderhevig aan PDEs, ze komen voor in wetenschappelijke domeinen zoals biologie, fysica, financiën, etc. PDEs worden bepaald door hun parameters. Vaak staan wetenschappers voor de uitdaging om onbekende parameters van PDEs te bepalen aan de hand van metingen die onderhevig zijn aan meetfouten. In de statistische literatuur wordt er heel vaak verondersteld dat de parameters van de PDEs constant zijn, wat de mogelijke toepassingen beperkt omdat in realiteit deze assumptie vaak te ruw is. In Hoofdstuk 4 breiden we een methode uit, waarvan reeds bewezen is dat deze effectief is in PDEs met constante parameters, naar PDEs met multivariate parameters. In het geval van lineaire PDEs tonen we aan dat onze schatter van de parameters uniform consistent is.

In Hoofdstuk 1 introduceren we verder de concepten van deze thesis met een overzicht van relevante statistische literatuur. Tenslotte, in Hoofdstuk 5 geven we een overzicht van de resultaten en lichten we enkele toekomstige onderzoeksperspectieven toe.

# Contents

# Chapter 1

# Introduction

The results in this thesis belong to the domain of nonparametric function estimation and hypothesis testing. Parametric function estimation assumes that the function of interest has a particular simple parametric expression, such as constant (i.e. polynomial of degree 0), linear (i.e. polynomial of degree 1) etc. On the other hand, nonparametric function estimation does not assume a prespecified expression of the function (Section 1.1).

The first goal of this thesis is to address the hypothesis whether coefficient functions in a varying coefficient model have a certain parametric form, such as taking a constant value (Section 1.2.1). In the same context we also construct effective testing procedures for the hypothesis that the coefficient functions are monotonic or convex (Section 1.2.2).

The second goal of this thesis is the estimation of varying parameters in a partial differential equation model (Section 1.3). The aim is to give good estimations of the varying parameters.

## 1.1 Spline approximation

We have mentioned that the coefficients of interest are estimated by nonparametric method. To introduce the concept of nonparametric function estimation we consider the univariate regression problem

$$Y = \mu(x) + \varepsilon, \tag{1.1}$$

where $\mu$ is an unknown mean function (depending on $x$) of the variable $Y$ with domain a finite interval $[a, b]$ and $\varepsilon$ is a mean zero stochastic variable which should be thought of as noise. It is not difficult to imagine that in general, assuming a parametric form (e.g. linear) for the mean function $\mu$ yields bad fits. Nonparametric estimation prevents this issue because the only assumption is on the smoothness of $\mu$. For example we assume that $\mu$ is continuous or that it has a bounded second derivative.

The function of interest ($\mu$ in (1.1)) is estimated by a spline function. Spline functions are characterized by a degree $q$ and knots $a = t_0 \leq t_1 \leq \ldots \leq t_K = b$.

**Definition 1.1** (Spline function)**.** *A function $S : [a, b] \to \mathbb{R}$ is a spline function of degree $q$ with knots $a = t_0 \leq t_1 \leq \ldots \leq t_K = b$ if*

$$S(x) = P_i(x) \quad \textit{on } [t_i, t_{i+1})$$

*for $i = 0, 1, \ldots, K-1$, and $S(t_K) = P_{K-1}(t_K)$, where $P_i$ is a polynomial of degree at most $q$ such that $S^{(q-1)}$ is continuous.*

The last condition –the function $S$ is $(q-1)$ continuously differentiable– means that the polynomials $P_i$ join smoothly at the knots. A few attractive features of spline functions (taken from Chapter 1 of Schumaker (2007)) are

- The spaces of spline functions are finite dimensional vector spaces with convenient bases (e.g. B-spline basis);
- Splines are relatively smooth functions;
- The derivatives and antiderivatives of splines are again splines;
- Every continuous function on the interval $[a, b]$ can be approximated arbitrarily well by splines with the degree $q$ fixed, provided a sufficient number of knots are allowed;
- Precise rates of convergence can be given for approximation of smooth functions by splines, not only are the functions themselves approximated, but their derivatives are as well.

Moreover, the space of spline functions has a conventient basis called (normalized) $B$-splines which are studied in great detail by De Boor (2001) and Schumaker (2007). B-splines have some useful properties of which we list a few:

- They are all nonzero;
- The support of any B-spline function is exactly given by $q+2$ consecutive knots;
- B-splines sum up to the constant function 1 on $[a,b]$;
- B-splines are defined recursively.

The recursive definition of B-splines allows to obtain the B-spline function of degree $q$ from B-splines of one degree lower. We define B-splines of degree 0 before we proceed to the recursive definition of B-splines. B-splines of degree 0 are given by

$$B_j(x;0) = \begin{cases} 1, & \text{if } t_{j-1} \leq x < t_j \\ 0, & \text{else,} \end{cases} \tag{1.2}$$

for $j = 1, \ldots, K$ and $B_K(t_K; 0) = 1$.

B-splines are defined recursively

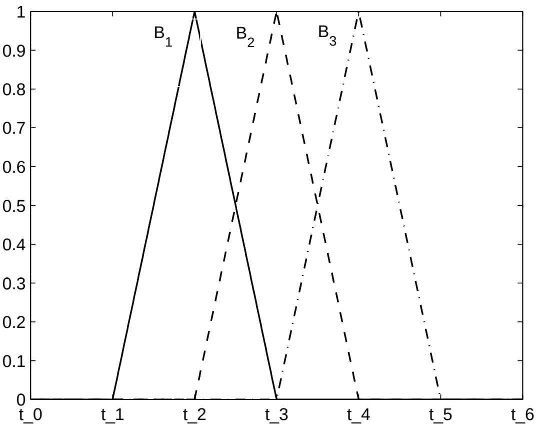
$$B_j(x;q) = \frac{x - t_{j-1}}{t_{j+q-2} - t_{j-1}} B_j(x;q-1) + (1 - \frac{x - t_j}{t_{j+q} - t_{j+1}}) B_{j+1}(x;q-1), \tag{1.3}$$

and it is applied for $j = -q+1, \ldots, K$ by adding knots $t_{-q} \leq t_{-q+1} \leq \ldots \leq t_{-1}$ left from $t_0$ and restricting to the interval $[t_0, t_K]$, to obtain a basis for the space of degree $q$ spline functions with knots $t_0, t_1, \ldots, t_K$. For the sake of presentation we let the indices of the B-splines go from 1 to $K+q$. In Figure 1.1 three B-splines of degree one and of degree two are drawn for the domain $[0,1]$. The B-splines of degree one is based on the knotset $\{0, \frac{1}{6}, \frac{2}{6}, \ldots, 1\}$, for the degree two B-spline basis the knots are $\{0, \frac{1}{7}, \frac{2}{7}, \ldots, 1\}$. Most properties of B-splines which are summed up above can directly be verified.

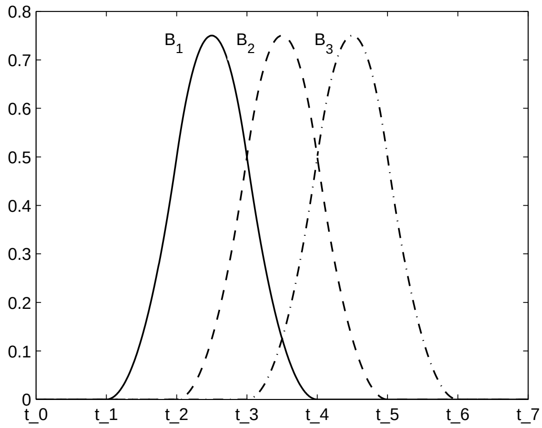
Let us return to model (1.1) and suppose we have data $(x_i, Y_i)$, for $i = 1, \ldots, n$, satisfying

$$Y_i = \mu(x_i) + \varepsilon_i \tag{1.4}$$

with $\mu$ a certain smooth function, and were $\varepsilon_i$ are independent and identically distributed (i.i.d.) random variables with finite variance $\sigma^2$. The function $\mu$ is modeled by a spline function $\mu(x) = \sum_{j=1}^{K+q} \alpha_j B_j(x;q)$. The estimator of $\boldsymbol{\alpha}$ is obtained by maximizing the likelihood function which is determined by the distribution of the error terms. When the error terms are normally




(a)




(b)

Figure 1.1: Three B-splines with equidistant knots on the unit interval (a) of degree 1 and knot distance $\frac{1}{6}$ and (b) of degree 2 with knot distance $\frac{1}{7}$.

distributed, the B-spline coefficients $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{K+q})^\top$ are estimated by finding the minimizer of the log-likelihood

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} \left( Y_i - \sum_{j=1}^{K+q} \alpha_j B_j(x_i; q) \right)^2 .$$

This approach is also used to estimate a multivariate $\mu$, which is achieved by using tensor product spline functions (see Chapter 4).

## 1.2 Varying coefficient models

We consider varying coefficient models (VCMs) to study longitudinal data. VCMs were developed by Hastie and Tibshirani (1993). Such models have been widely applied to many scientific areas: environmental science, ecology, econometrics, epidemiology, etc. VCMs are an extension of classic linear regression models where the coefficient corresponding to a covariate is assumed to be constant (independent of other variables). This assumption can lead to poor modeling when the data is for example time dependent. Therefore, the modeling strategy ought to be revised to increase flexibility and maintain interpretability (Fan and Wenyang (2008)). The extension consists of allowing the coefficients to depend on other variables. In Chapters 2 and

3 we consider the model

$$Y(t) = \mathbf{X}(t)^\top \boldsymbol{\beta}(t) + \varepsilon(t) = \sum_{p=0}^{d} X^{(p)}(t)\beta_p(t) + \varepsilon(t), \tag{1.5}$$

where $Y(t)$ is the response at time $t \in \mathcal{T} = [0,1]$; $\mathbf{X}(t) = (X^{(0)}(t), \ldots, X^{(d)}(t))^\top$ is the covariate vector at time $t$, with $X^{(0)}(t) \equiv 1$; $\boldsymbol{\beta}(t) = (\beta_0(t), \ldots, \beta_d(t))^\top$ is the vector of coefficient functions at time $t$. Note that $t$ can be any variable and the domain $\mathcal{T}$ can be any bounded interval. The stochastic error function $\varepsilon(t)$ has mean zero (conditioned on $\mathbf{X}(t)$). The $j$-th measurement of subject $i$ $(t_{ij}, Y_{ij}, \mathbf{X}_{ij})$ for $1 \leq i \leq n$ and $1 \leq j \leq N_i$, is a sample from $(t, Y(t), \mathbf{X}(t))$, where $t_{ij}$ is the observed time, $Y_{ij}$ is the observed response of the $i$th subject at time $t_{ij}$ and $\mathbf{X}_{ij} = (X_{ij}^{(0)}, \ldots, X_{ij}^{(d)})^\top$ is the corresponding observed covariate vector. The observed covariates and responses are used for nonparametric estimation of the coefficient functions $\beta_0, \beta_1, \ldots, \beta_d$. This can be achieved by several nonparametric techniques. Local polynomial techniques are discussed in Hoover et al. (1998) and Fan and Zhang (1999), among others. Huang et al. (2002) approximates the coefficient functions by spline functions using B-spline bases, and Antoniadis et al. (2012) use penalized splines approximation in a variable selection context. Note that the optimal choice of the smoothing level for coefficient estimation need not be the optimal choice for the hypothesis testing, for more on this note see Zhang and Mei (2012) (p. 1945-1946).

Consider the application of a VCM on the AIDS data which is a subset of the Multicenter AIDS Cohort Study, and which is also analysed in Chapter 2. This data set contains the repeated measurements of physical examinations, laboratory results and CD4 cell percentages of 283 homosexual men who became HIV-positive between 1984 and 1991. CD4 cells play an important role in the body's immune system. The HIV virus destroys CD4 cells. The fewer functioning CD4 cells, the weaker the immune system and therefore the more vulnerable a person is to infections and illnesses. The patients would have measurements taken every 6 months, but due to certain individual's missing their appointments and the random infection moment, the number of repeated measurements varied per individual. The aim of the statistical analysis is to describe the trend of the mean CD4 percentage depletion over time (in years) explained by the effects of cigarette smoking, age at HIV infection and pre-HIV infection CD4 percentage. For more details about the design, methods and medical applications see Kaslow et al. (1987). Consider the model

$$Y_{ij} = \beta_0(t_{ij}) + \beta_1(t_{ij})X_i^{(1)} + \beta_2(t_{ij})X_i^{(2)} + \beta_3(t_{ij})X_i^{(3)} + \varepsilon_{ij},$$

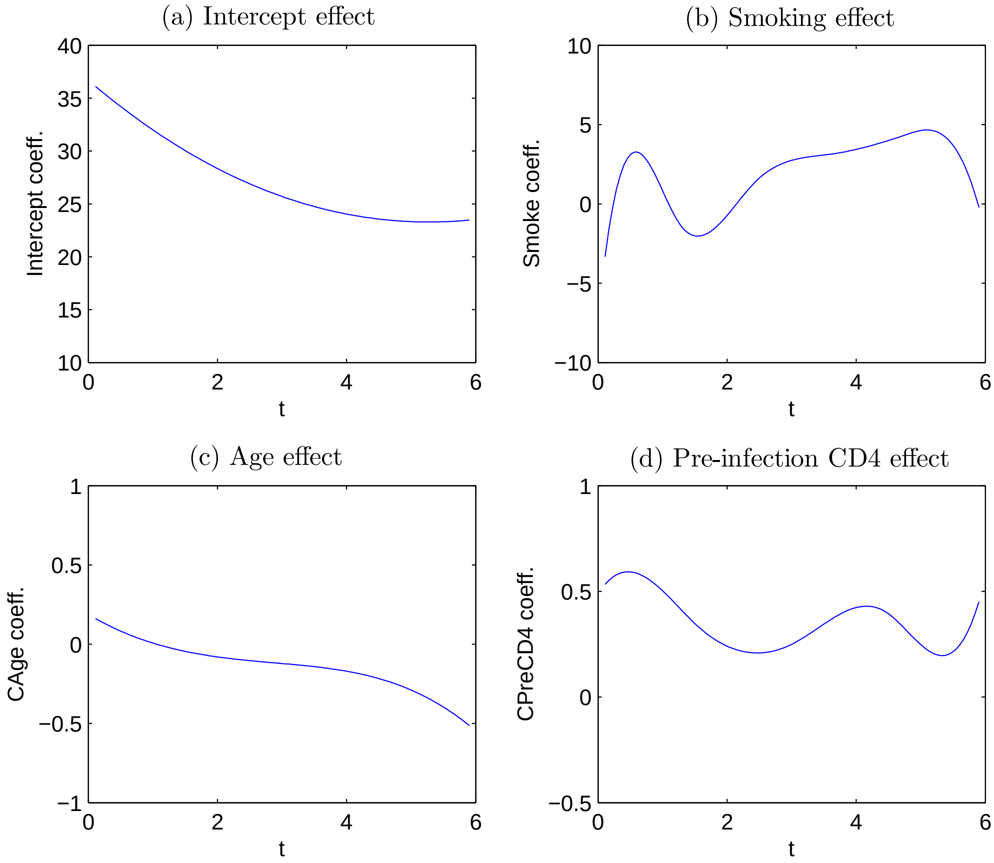



Figure 1.2: AIDS data. Spline estimations of the coefficient functions.

where $Y_{ij}$ is the $i$th individual CD4 percentage measured at time $t_{ij}$, $X_i^{(1)}$ is the smoking status of the $i$th individual; $X_i^{(1)}$ is 1 or 0 if the individual ever or never smoked, $X_i^{(2)}$ is the $i$th individual's centered age at HIV infection (obtained by subtracting the sample average age at HIV infection from the individual's age at HIV infection), and in a similar way we let $X_i^{(3)}$ be the $i$th individual's centered pre-HIV infection CD4 percentage. By centering, the intercept $\beta_0(t)$ represents the mean CD4 percentage $t$ years after HIV infection of a homosexual individual with an average age at HIV infection, an average pre-HIV CD4 rate and who has never smoked cigarettes. With our method which is described in Chapter 2, we test whether a linear regression model makes more sense. The p-value for this test is smaller than 1e$-$3, therefore we strongly reject the hypothesis that all coefficient functions are constants. Hence, it makes sense to state a varying coefficient model instead of the linear regression model. Figure 1.2 contains the coefficient function estimations which are obtained by finding the spline function approximation which conforms the data the most. The details of finding these spline function approximations are provided in Section 2.2.1.

The coefficients need not be time dependent. The second application is such an example, which is taken from the German Continental Deep Drill Program (KTB) which was designed to study the properties and processes of the deeper continental crust by means of a superdeep borehole. More background information and results from this program can be found in Emmermann and Lauterjung (1997). Of particular interest is the occurrence of cataclastic shear zones in the upper and middle earth crust. The amount of cataclastic rocks (CATR) which are revealed by the drill cuttings helps to address cataclastic shear zones. Antoniadis et al. (2012) explain the amount of CATR by other variables (such as $SiO_2$ content and $Na_2O$ content) through a VCM where the coefficients vary with depth up to 9.1 km, in a variable selection setting.

### 1.2.1 Testing for parametric forms of coefficients

In the context of model (1.5), it is important to know whether certain coefficient functions have a parametric form (i.e. a polynomial of prespecified degree), for example constant. This could mean that a covariate does not have a varying but a constant effect on the response and that effect should be modeled by a scalar. Similarly, when it is decided that a coefficient function is polynomial of a certain degree then a much simpler model can be proposed to both reduce computational costs and prevent overfitting. Chapter 2 establishes a method to test these kind of hypotheses for longitudinal data models. For a literature overview of hypothesis testing that a coefficient function is constant in a cross-sectional data model we refer to Li et al. (2011) and the references therein. For longitudinal data, Huang et al. (2002) constructed a test statistic based on the difference of the residual sum of squares under the null (coefficient is constant) and the alternative hypothesis (coefficient is varying), but do not acquire asymptotic results of their approach. They obtain critical values via a bootstrap strategy, which imposes the need of a relatively large sample size at a high computational cost. Zhang (2004) proposed generalized linear mixed models for inference in varying coefficient models that include models where $Y$ can be nonnormal such as binary or Poisson. However, their approach includes a strong parametric assumption through random effects, and typically these effects are assumed to be normal.

Our method extends the technique of Li et al. (2011) to longitudinal data with correlated error structures, where the coefficient functions are estimated based on a B-spline basis expansion. Their approach makes fully use of the nice properties of B-splines. A main advantage of this approach, besides its simplicity and high power, is that it can be extended to other interesting hypotheses. The test statistic follows (asymptotically if the coefficient functions are not spline functions) a Fisher distribution. The main difficulty is incorporating the weight matrix (when longitudinal data are used) and the correlation structure of the errors. The novelty of our approach furnishes this issue. We prove that our test statistic follows asymptotically a generalized Fisher (notation: F) distribution. This generalized F distribution is the exact null distribution if the coefficient functions are splines. We also discuss how we can test for parametric forms in other varying coefficient models.

### 1.2.2 Testing for shapes of coefficients

It can be of interest to derive some conclusions about the shapes of the coefficient functions. In Chapter 3, we develop tests for monotonicity and convexity. For example, a monotonically increasing coefficient of a time independent predictor indicates that the effect of this predicator on the response is increasing. This can be important in, among other fields, medical sciences. See for example our study of schizophrenia patients (Sect. 3.8), where the 'Severity of Illness' is modeled by a VCM with covariate the binary variable whether the patient received a drug, with coefficients depending on time (with week as unit of time). The general finding was that the drug improved the health of the patients considerably. Since we are employing a VCM we looked at the behaviour of the drug coefficient which revealed additional information on the evolution of the drug effect on the patients. To the best of our knowledge there is yet no effective testing procedure for monotonicity and convexity in varying coefficient models. Our approach for testing monotonicity is universal and can be applied to other varying coefficient models. In the context of univariate regression, methods of estimation under a monotonicity constraint and testing for monotonicity have been widely discussed, see Bowman et al. (1998), Ghosal et al. (2000),Wang and Meyer (2011) and references therein. In the context of varying coefficient models, not much has been written on this subject. However, Zhang et al. (2013) extended the SiZer map approach to varying coefficient models where the local polynomial estimation technique is used, which reveals the statistically significant features of the coefficient functions. The SiZer approach leads to a good explanatory tool, for example for choosing the level of smoothness of each coefficient function. We develop two testing procedures for monotonicity and convexity (concavity) using the nice properties of B-splines. We use a straightforward test statistic. For example, for testing whether a coefficient function is monotonically increasing the test statistic is the minimum of the derivative of the estimated coefficient function. Then, by using a bootstrap approach the hypothesis is rejected if the test statistic is significantly smaller than zero. Moreover, we develop testing procedures for testing simultaneously different coefficient functions. A side result of this work is that we have shown that the first few derivatives of the B-spline estimator are uniformly consistent.

## 1.3 Estimation of Multivariate parameters in Partial Differential Equation Models

Ordinary differential equations (ODE) are used to model dynamic processes which are common in real life. They have been applied in biology, physics, economy, engineering, etc. There is a considerable amount of literature on ODE models, dealing with parameter estimation and their statistical properties (see Xue et al. (2010) and references therein, Hong and Lian (2012) and Frasso et al. (2016), among others). In the overview below, we touch upon a few statistical methods available in the literature which have been used to estimate the parameters in an ODE.

Only in this section let $\mathbf{X}(t) = (X_1(t), X_2(t), \ldots, X_L(t))^\top$ denote an $L$-dimensional state variable vector with initial value $\mathbf{X}_0 = \mathbf{X}(t_0)$. We consider an initial value problem ODE model

$$\frac{d\mathbf{X}}{dt} = F(t, \mathbf{X}(t); \boldsymbol{\theta}(t)), \quad \forall t \in [t_0, T], \tag{1.6}$$

$$\mathbf{X}(t_0) = \mathbf{X}_0, \tag{1.7}$$

where $\boldsymbol{\theta}(t) = (\theta_1(t), \ldots, \theta_d(t))^\top$ is an unknown $d$-dimensional parameter vector with components (possibly constant) functions $\theta_p(t)$, $p = 1, \ldots, d$. Moreover, we assume that $F$ is a known smooth function with domain a subset of $I\!R^{L+1}$.

We assume having $n$ measurements of $\mathbf{X}(t)$ at random or fixed design points $t_1, \ldots, t_n$, and that we observe these measurements with an error:

$$\mathbf{Y}(t_i) = \mathbf{X}(t_i) + \boldsymbol{\varepsilon}(t_i), \quad i = 1, \ldots, n, \tag{1.8}$$

where the measurement errors $\boldsymbol{\varepsilon}(t_1), \boldsymbol{\varepsilon}(t_2), \ldots, \boldsymbol{\varepsilon}(t_n)$ are i.i.d. with mean zero and covariance $\boldsymbol{\Sigma}$, and $\mathbf{Y}(t_i) = (Y_1(t_i), \ldots, Y_L(t_i))^\top$.

Ramsay et al. (2007) study the model given by (1.6) and (1.7) where $\boldsymbol{\theta}(t)$ is a constant vector. They estimate the parameters by a penalized spline approach to assure both fidelity to the ODE model and the data (1.8).

Xue et al. (2010) give theoretical results of their parameter estimates where only one parameter is allowed to be time varying. To illustrate their estimation procedure we assume for simplicity that $\boldsymbol{\theta}(t)$ is a constant and write $\boldsymbol{\theta}$. When the exact solution of (1.6) and (1.7) is not known, a numerical solution is used. The numerical solution of $X_j$ is denoted by $\tilde{X}_j(t, \boldsymbol{\theta})$. Then the parameter $\boldsymbol{\theta}$ is obtained by solving the nonlinear least squares problem

$$\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^{n} \sum_{j=1}^{L} (Y_j(t_i) - \tilde{X}_j(t_i, \boldsymbol{\theta}))^2. \tag{1.9}$$

Hong and Lian (2012) solve linear ODE models ($F$ is linear in (1.6)) with time varying parameters $\boldsymbol{\theta}(t)$ by using the local polynomial smoothing technique. Their approach is a two-stage method. First, they use local polynomials to estimate $\mathbf{X}(t_i)$ from the data (1.8) to obtain the local polynomial estimator $\hat{\mathbf{X}}(t)$. Second, they solve, using the norm notation given in Appendix A,

$$\operatorname{argmin}_{\boldsymbol{\theta}(t) \in \mathcal{A}} \sum_{i=1}^{n} K((t_i - t_0)/h) \| \frac{d\hat{\mathbf{X}}}{dt}(t_i) - F(t_i, \hat{\mathbf{X}}(t_i); \boldsymbol{\theta}(t_i)) \|_2^2 \tag{1.10}$$

for a kernel function $K$ and bandwidth $h$, where $\mathcal{A}$ is the product space of polynomials of fixed degrees.

Some physical problems cannot be described by an ODE, but are described by partial differential equations (PDE). There are three main approaches to estimate the parameters of the PDE model. The first one is a two-stage method (similar to Hong and Lian (2012)). The second one is similar to the approach of Xue et al. (2010). Another approach is given by Xun et al. (2013) which is similar to Ramsay et al. (2007)'s approach. Xun et al. (2013) assume a multivariate process $g(\mathbf{t})$ setting ($\mathbf{t} = (t_1, \ldots, t_l)$), where we only observe

$$Y_i = g(\mathbf{t}_i) + \varepsilon_i, \quad i = 1, \ldots, n \tag{1.11}$$

and $\mathbf{t}_i$, $i = 1, \ldots, n$, where $\varepsilon_i$ are i.i.d. mean zero measurement errors. The latest research efforts are collected in Chapter 4 where we consider PDE models with unknown varying parameters which we want to estimate. Often scientists face the challenge to determine unknown parameters of a PDE, and the need to estimate them from error prone measurements. In the statistical literature it is generally assumed that the parameters are constant, which restricts the application possibilities because in reality this assumption can be crude. Below we give an example of heat diffusion which is described by a PDE and is discussed in Section 1.2 of Haberman (2004).

Consider the temperature function $g(t_1, t_2)$ of a one-dimensional rod of constant cross-sectional area with length $L$ which is made of one particular substance. The rod is oriented along the $t_1$-axis (from $t_1 = 0$ to $t_1 = L$). Suppose the lateral surface is insulated so that there is no transfer of heat energy in this direction. We observe the temperature $g(t_1, t_2)$ of the rod at positions $0 \leq t_1 \leq L$ and times $0 \leq t_2 \leq T$ with measurement errors. The question is now to find the unobservable source function $\theta(t_1, t_2)$ which is causing the rod to heat up and/or cool down. Since the rod is insulated, one can think of the source function to be the result of internal chemical reactions or electrical heating. This process is described by the PDE

$$\frac{\partial g}{\partial t_2}(t_1, t_2) + D\frac{\partial^2 g}{\partial t_1^2}(t_1, t_2) + \theta(t_1, t_2) = 0,$$

where $\theta$ is the source function which represents the internal heating source of the rod. Suppose we have Dirichlet boundary conditions:

$$\begin{cases} \frac{\partial g}{\partial t_2}(t_1, t_2) + D\frac{\partial^2 g}{\partial t_1^2}(t_1, t_2) + \theta(t_1, t_2) = 0 & \\ g(t_1, 0) = 0 & 0 \leq t_1 \leq L \\ g(0, t_2) = 0 & 0 \leq t_2 \leq T \\ g(L, t_2) = 0 & 0 \leq t_2 \leq T. \end{cases}$$

Figure 1.3 contains plots of $\theta$ and $g$ which satisfy the PDE model with $D = -\pi$. Our aim is to consistently estimate $\theta$ from error prone observations of $g$. This is the subject of Chapter 4.

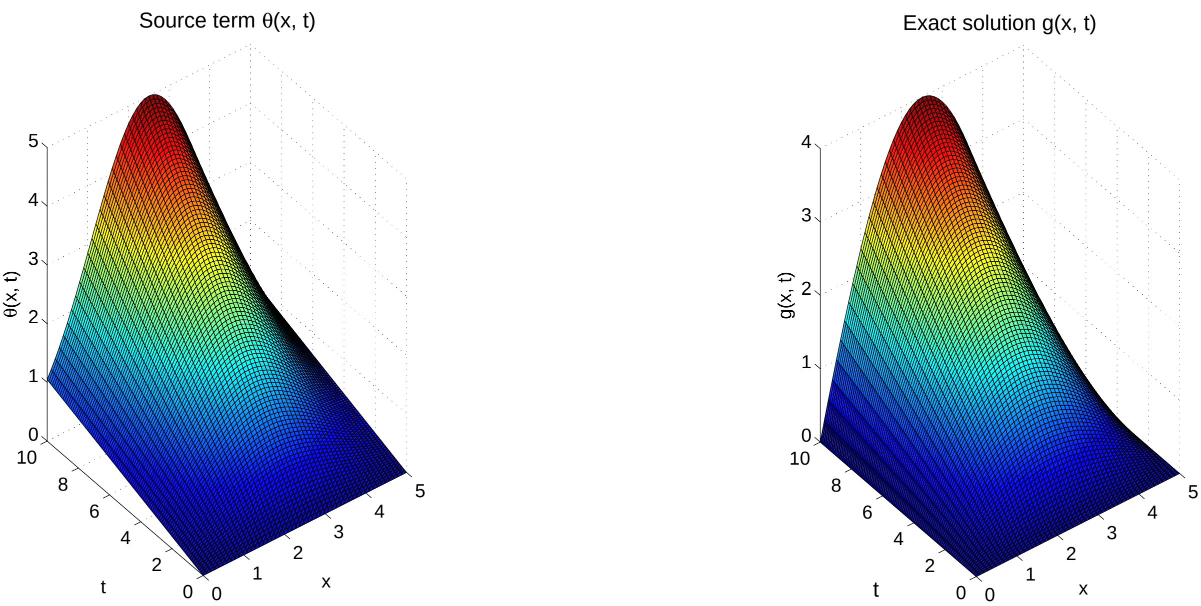



Figure 1.3: Heat diffusion. The heat solution $g$ of the PDE (right) corresponding to the source term $\theta$ (left).

# Chapter 2

# Testing for parametric forms in varying coefficient models

The content of this chapter is published in Ahkim and Verhasselt (2017).

## 2.1 Introduction

In this chapter we consider the varying coefficient model for longitudinal data

$$Y_{ij} = \sum_{p=0}^{d} X_{ij}^{(p)} \beta_p(t_{ij}) + \varepsilon(t_{ij}), \tag{2.1}$$

for $i = 1, \ldots, n$, and $j = 1, \ldots, N_i$. $Y_{ij}$ represents the response of individual $i$ at time $t_{ij}$, and $\mathbf{X}_{ij} = (X_{ij}^{(0)}, \ldots, X_{ij}^{(d)})^\top$ (with $X_{ij}^{(0)} \equiv 1$) the corresponding observed covariate vector. We develop procedures to test for parametric forms of the coefficients $\beta_p(\cdot),\ p = 0, 1, \ldots, d$. On the one hand this assures correct assessment of nonparametric covariate effects, if the hypothesis of parametric form is rejected. On the other hand, if the hypothesis is not rejected, a parametric model should be used. This would reduce the complexity of the model and prevent overfitting. For example, if we do not reject the hypothesis that $\beta_1(\cdot)$ is a polynomial of degree $q$, we model it by a degree $q$ spline function without internal knots (see the first alinea of Section 2.2.1), then the number of B-spline basis functions is $q + 1$ which is the dimension of the space of degree $q$ polynomials.

We study the test statistic which is a ratio of quadratic forms (RQF) and resembles the Fisher test static in simple linear regression. We assume homoscedastic normal errors with intrasubject correlation. If the correlation matrix is given we construct the RQF test statistic which follows a generalized F distribution under the null hypothesis. However, we show that imposing a misspecified correlation structure (i.e. assuming independence) the RQF method is still satisfying. It is a natural question whether the RQF approach also works when we do not assume the correlation structure to be given and instead use an estimated correlation matrix. In the conclusion of this chapter (Section 2.7) we explain why this approach (i.e. using a correlation estimate) is not successful. This led us to propose a bootstrap approach when there is no knowledge about the correlation structure and (or) the normality assumption does not hold.

When the response variable represents counts, a binary variable, etc. generalized varying coefficient models (GVCMs) are considered where it is assumed that the density function of the response variable comes from the exponential family. This is similar to the extension from the classic linear models to generalized linear models by McCullagh and Nelder (1989). We propose a bootstrap approach for hypothesis testing of parametric forms in GVCMs.

Further, this chapter is organized as follows. In Section 2.2 we describe the B-spline estimator. In Section 2.3 the testing procedure and asymptotic results are presented when error terms are normally distributed. The proofs are in Section 2.8. In Section 2.6 we discuss the extension to GVCMs. The performance of our method compared to Huang et al. (2002) are illustrated with numerical simulations in Section 2.4 and a data application is discussed in Section 2.5. Finally, we end with a conclusion in Section 2.7.

## 2.2 Spline estimation

### 2.2.1 B-spline estimator

In this section we briefly recall the B-spline estimator in varying coefficient models, see Huang et al. (2004). The assumption is that each component of $\boldsymbol{\beta}(t) = (\beta_0(t), \ldots, \beta_d(t))^\top$ can be approximated by a B-spline basis expansion, i.e., for each $p = 0, \ldots, d$, $\beta_p(t) \approx \sum_{l=1}^{m_p} \alpha_{pl} B_{pl}(t; q_p)$, where $\{B_{pl}(\cdot; q_p) : l = 1, \ldots, K_p + q_p = m_p\}$ is the normalized $\Big($i.e. $\sum_{j=1}^{m_p} B_j(\cdot; q_p) = 1\Big)$ $q_p$th degree B-spline basis with $K_p + 1$ equidistant knots $\xi_{p0}, \xi_{p1}, \ldots, \xi_{pK_p}$ in $\mathcal{T}$. Let $\mathbb{G}_p$ denote the space spanned by this basis.

The B-spline estimator $\hat{\boldsymbol{\alpha}} = (\hat{\boldsymbol{\alpha}}_0^\top, \ldots, \hat{\boldsymbol{\alpha}}_d^\top)^\top$ (with $\hat{\boldsymbol{\alpha}}_p = (\hat{\alpha}_{p1}, \ldots, \hat{\alpha}_{pm_p})^\top$) is obtained by minimizing the following expression with respect to $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_0^\top, \ldots, \boldsymbol{\alpha}_d^\top)^\top$, where $\boldsymbol{\alpha}_p = (\alpha_{p1}, \ldots, \alpha_{pm_p})^\top$ for $p = 0, \ldots, d$:

$$\sum_{i=1}^{n} w_i \sum_{j=1}^{N_i} \Big( Y_{ij} - \sum_{p=0}^{d} \sum_{l=1}^{m_p} X_{ij}^{(p)} B_{pl}(t_{ij}; q_p) \alpha_{pl} \Big)^2,$$

where $w_i$ denotes the weight for subject $i$, often $w_i = \frac{1}{N_i}$ is used. More compactly written, we solve

$$\min_{\boldsymbol{\alpha}} \sum_{i=1}^{n} (\mathbf{Y}_i - \mathbf{U}_i \boldsymbol{\alpha})^\top \mathbf{W}_i (\mathbf{Y}_i - \mathbf{U}_i \boldsymbol{\alpha}), \tag{2.2}$$

where

$$\begin{aligned}
\mathbf{Y}_i &= (Y_{i1}, \ldots, Y_{iN_i})^\top; \quad \mathbf{Y} = (\mathbf{Y}_1^\top, \ldots, \mathbf{Y}_n^\top)^\top; \\
\mathbf{B}(t) &= \begin{pmatrix} B_{01}(t; q_0) & \ldots & B_{0m_0}(t; q_0) & 0 \ldots 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & \ddots & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 \ldots 0 & B_{d1}(t; q_d) & \ldots & B_{dm_d}(t, q_d) \end{pmatrix} \in I\!R^{(d+1) \times dim}; \\
\mathbf{U}_{ij}^\top &= \mathbf{X}_{ij}^\top \mathbf{B}(t_{ij}) \in I\!R^{1 \times dim}; \\
\mathbf{U}_i &= (\mathbf{U}_{i1}, \ldots, \mathbf{U}_{iN_i})^\top \in I\!R^{N_i \times dim}, \quad \text{where } dim = \sum_{p=0}^{d} m_p; \\
\mathbf{U} &= (\mathbf{U}_1^\top, \ldots, \mathbf{U}_n^\top)^\top \in I\!R^{N \times dim}; \\
\mathbf{W}_i &= \operatorname{diag}\Big(w_i, \ldots, w_i\Big) \in I\!R^{N_i \times N_i} \quad \text{(a diagonal matrix with } N_i \text{ times} \\
&\quad\ w_i \text{ on the diagonal).} \\
\mathbf{W} &= \operatorname{diag}\Big(\mathbf{W}_1, \ldots, \mathbf{W}_n\Big) \in I\!R^{N \times N} \quad \text{(a block diagonal matrix} \\
&\quad\ \text{with the matrices } \mathbf{W}_i \text{ on the diagonal).}
\end{aligned}$$

If $(\mathbf{U}^\top \mathbf{W} \mathbf{U})$ is invertible, then (2.2) has a unique solution

$$\hat{\boldsymbol{\alpha}} = (\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \mathbf{U}^\top \mathbf{W} \mathbf{Y}. \tag{2.3}$$

Huang et al. (2004) proved that under Assumption 2.1.1−3 and Assumption 2.1.5 given in Section 2.8.1, the matrix $(\mathbf{U}^\top \mathbf{W} \mathbf{U})$ is invertible with probability tending to 1. Then, the

B-spline estimator of $\boldsymbol{\beta}(t)$ is

$$\hat{\boldsymbol{\beta}}(t) = \mathbf{B}(t)\hat{\boldsymbol{\alpha}} = (\hat{\beta}_0(t), \ldots, \hat{\beta}_d(t))^\top, \qquad \text{with } \hat{\beta}_p(t) = \sum_{l=1}^{m_p} \hat{\alpha}_{pl} B_{pl}(t; q_p).$$

### 2.2.2 Some properties of spline approximations

The motivation for our test statistics are based on the following nice properties of B-spline approximations.

Fix $k \in \{0, \ldots, d\}$. Suppose that the function $\beta_k(t)$ is a constant $c_k$, then $\beta_k(t) = c_k = \sum_{l=1}^{m_k} \alpha_{kl} B_{kl}(t; q_k) \in \mathbb{G}_k$. The equation sign holds since constant functions on $\mathcal{T}$ are contained in $\mathbb{G}_k$. Moreover, we have that $\boldsymbol{\alpha}_k = (c_k, \ldots, c_k)^\top \in I\!R^{m_k \times 1}$, since normalized B-splines are used and the functions $B_{kl}(t; q_k)$ $(l = 1, \ldots, m_k)$ form a basis of $\mathbb{G}_k$. Therefore, the function $\beta_k(\cdot)$ is constant if and only if all the components of the vector of $\boldsymbol{\alpha}_k$ are equal. constant.

## 2.3 Testing constancy of coefficient functions

In this section we consider the problem of testing whether the $k$th coefficient $\beta_k(t)$ of a varying coefficient model is really varying. We develop a testing procedure to test for constancy, i.e. test

$$H_0 : \beta_k(\cdot) \text{ is a constant function} \qquad \text{versus} \qquad \neg H_0 : \beta_k(\cdot) \text{ is not a constant function.} \quad (2.4)$$

Li et al. (2011) consider hypothesis (2.4) in varying coefficient models with cross-sectional data. Their technique is based on the vector of first order differences $\mathbf{D}_1\hat{\boldsymbol{\alpha}}_k$ where

$$\mathbf{D}_1 = \begin{pmatrix} 1 & -1 & 0 & 0\ldots & 0 & 0 & 0 \\ 0 & 1 & -1 & 0\ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0\ldots & 0 & 1 & -1 \end{pmatrix} \in I\!R^{(m_k-1)\times m_k}.$$

They use linear splines. However, it should be noted that splines of any degree could be used (as noted in Section 2.2.2). We extend their approach to our longitudinal data model with correlated errors where the coefficient functions are estimated by B-splines of any degree.

We first give a test for the more restrictive hypothesis that all coefficient functions are constant:

$$H_1 : \beta_p(\cdot) \text{ is a constant function for } p = 0, \ldots, d \qquad \text{versus} \qquad \neg H_1. \tag{2.5}$$

Under hypothesis $H_1$ all coefficient functions are modeled by spline functions. Let us consider the model where all coefficient functions are spline functions, i.e. $\beta_p(t) = \sum_{l=1}^{m_k} \alpha_{kl} B_{kl}(t; q_k)$ and

$$\mathbf{Y} = \mathbf{U}\boldsymbol{\alpha} + \boldsymbol{\varepsilon},$$

with $\mathbf{Y} = (\mathbf{Y}_1^\top, \ldots, \mathbf{Y}_n^\top)^\top$, $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^\top, \ldots, \boldsymbol{\varepsilon}_n^\top)^\top$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iN_i})^\top$ for $i = 1, \ldots, n$. For this model, testing problem (2.5) is equivalent to

$$H_1^* : \mathbf{L}_1^\top \boldsymbol{\alpha} = \mathbf{0} \qquad \text{versus} \qquad \neg H_1^* : \mathbf{L}_1^\top \boldsymbol{\alpha} \neq \mathbf{0}, \tag{2.6}$$

where $\mathbf{L}_1^\top \in I\!R^{(dim-d-1)\times dim}$ takes the first order difference of consecutive components of $\boldsymbol{\alpha}_p$, $p = 0, \ldots, d$.

### 2.3.1 Construction of the test statistic

We assume a homoscedastic error structure, i.e. $\text{Cov}(\boldsymbol{\varepsilon}_i) = \sigma^2 \mathbf{R}_i$ where $\text{Corr}(\boldsymbol{\varepsilon}_i) = \mathbf{R}_i$ and $\text{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{R}$, where $\mathbf{R} = \text{diag}(\mathbf{R}_1, \ldots, \mathbf{R}_n)$. Recall the B-spline estimator of model (1.5)

$$\hat{\boldsymbol{\alpha}} = (\underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{U}})^{-1} \underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{Y}},$$

where $\underset{\sim}{\mathbf{U}} = \mathbf{W}^{\frac{1}{2}}\mathbf{U}$ and $\underset{\sim}{\mathbf{Y}} = \mathbf{W}^{\frac{1}{2}}\mathbf{Y}$. Let $P_{\underset{\sim}{\mathbf{U}}} = \underset{\sim}{\mathbf{U}}(\underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{U}})^{-1}\underset{\sim}{\mathbf{U}}^\top$ and $\tilde{\mathbf{Y}} = \text{E}(\mathbf{Y}|\mathcal{X})$, where $\mathcal{X} = \{(\mathbf{X}_{ij}, t_{ij}); i = 1, \ldots, n, j = 1, \ldots, N_i\}$. Throughout the remainder of this chapter we condition on $\mathcal{X}$. Let $\underset{\sim}{\mathbf{R}} := \text{Var}(\underset{\sim}{\mathbf{Y}}) = \sigma^2 \mathbf{W}^{1/2}\mathbf{R}\mathbf{W}^{1/2}$.

#### 2.3.1.1 Testing hypothesis (2.5)

*Remark* 2.1. If we would use Li et al. (2011)'s test statistic in our longitudinal case, then we would obtain the "test statistic"

$$\frac{dim - d - 1}{N - dim} \frac{\hat{\boldsymbol{\alpha}}^\top \mathbf{L}_1 (\mathbf{L}_1^\top (\underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{U}})^{-1} \mathbf{L}_1)^{-1} \mathbf{L}_1 \hat{\boldsymbol{\alpha}}}{\underset{\sim}{\mathbf{Y}}^\top (\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}}) \underset{\sim}{\mathbf{Y}}}, \tag{2.7}$$

where $\mathbf{I}_N$ is the identity matrix of dimension $N \times N$. Unlike the case in Li et al. (2011) where the statistic (2.7), forms a ratio of two *independent* $\chi^2$ variables, we here have a ratio of two *dependent* $\chi^2$ variables (see the proof of Theorem 2.1). The novelty of our approach allows to incorporate the weight matrix $\mathbf{W}$ and the correlation matrix $\mathbf{R}$ arising from longitudinal data models. This generalization is not straightforward.

Our test statistic is based on the fact that under hypothesis (2.5)

$$\mathrm{E}\left(\utilde{\mathbf{L}}_1^\top \hat{\utilde{\boldsymbol{\alpha}}} | \mathcal{X}\right) = \mathrm{E}\left(\mathbf{L}_1^\top \hat{\boldsymbol{\alpha}} | \mathcal{X}\right) = \mathbf{0},$$

where $\utilde{\mathbf{L}}_1^\top = \mathbf{L}_1^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{U}})^{-1} (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})$, $\hat{\utilde{\boldsymbol{\alpha}}} := (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{Y}}$. Note that $\hat{\utilde{\boldsymbol{\alpha}}}$ and $\utilde{\mathbf{L}}_1^\top$ are precisely introduced to obtain a ratio of independent quadratic forms of which the distribution is known, see below.

Since $\utilde{\mathbf{Y}} \sim N(\mathbf{W}^{\frac{1}{2}} \tilde{\mathbf{Y}}, \utilde{\mathbf{R}})$, we have that $\hat{\utilde{\boldsymbol{\alpha}}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \mathbf{W}^{1/2} \tilde{\mathbf{Y}} \quad \text{and} \quad \boldsymbol{\Sigma} = (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{U}} (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1}.$$

Next we define two quadratic forms in normal variables. The first is

$$Q_1 = \frac{1}{\sigma^2} \utilde{\mathbf{Y}}^\top (\mathbf{I}_N - P_{\utilde{\mathbf{U}}}) \utilde{\mathbf{Y}},$$

the second

$$Q_2 = \hat{\utilde{\boldsymbol{\alpha}}}^\top \utilde{\mathbf{L}}_1 (\utilde{\mathbf{L}}_1^\top \boldsymbol{\Sigma} \utilde{\mathbf{L}}_1)^{-1} \utilde{\mathbf{L}}_1^\top \hat{\utilde{\boldsymbol{\alpha}}}.$$

Our test statistic for hypothesis (2.5) is a ratio of these (stochastic) quadratic forms, namely

$$T_1 = \frac{dim - d - 1}{N - dim} \frac{Q_1}{Q_2}.$$

Such a test statistic will be termed by RQF test statistic. Note that $T_1$ does not depend on $\sigma^2$. Theorem 2.1 states the exact null distribution of $T_1$. When $t_1 \in I\!R$ is a realization of $T_1$, the p-value $p_1$ to test $H_1$ is defined to be

$$p_1 = F_{T_1}(t_1), \tag{2.8}$$

where $F_1$ is the distribution function of $T_1$ under the null hypothesis (2.6), since $Q_2$ is relatively small under the null hypothesis. By Theorem 2.1 we know that under the null hypothesis in (2.5), $T_1$ follows a generalized F distribution of the type

$$\frac{(\sum_{i=1}^{l} c_i X_i)/(\sum_{i=1}^{l} m_i)}{Y/n},$$

where the components of $(X_1, \ldots, X_l, Y)$ are independent with $X_i \sim \chi^2(m_i), Y \sim \chi^2(n)$ and all $c_i > 0$. Dunkl and Ramirez (2001) gave exact and numerically tractable expressions of the cumulative distribution function (cdf) for this kind of generalized F distributions. We have implemented the cdf of this distribution in Matlab to compute the p-values.

Note that in the case we have data without repeated measurements, i.e. the matrices $\mathbf{W}$ and $\mathbf{R}$ are the identity matrix, the test statistic

$$\frac{N - dim}{dim - d - 1} \frac{Q_2}{Q_1}$$

is exactly the test statistic used by Li et al. (2011) (see (2.7) in Remark 1) and follows the F distribution with degrees of freedom $(dim - d - 1, N - dim)$ under the null hypothesis in (2.5).

#### 2.3.1.2 Testing hypothesis (2.4)

Let us return to hypothesis (2.4):

$$H_0 : \beta_k(\cdot) \text{ is a constant function} \qquad \text{versus} \qquad \neg H_0 : \beta_k(\cdot) \text{ is not a constant function.}$$

This time we are only interested in the coefficients belonging to the coefficient function $\beta_k(.)$, therefore we apply the following transformation on $\hat{\boldsymbol{\alpha}}$

$$\mathbf{L}_2^\top \;=\; \begin{pmatrix} \mathbf{0} & \ldots & \mathbf{D}_1 & \ldots & \mathbf{0} \end{pmatrix} \in I\!R^{(m_k-1)\times dim}.$$

The hypothesis in terms of the B-spline coefficients becomes

$$H_0^* : \mathbf{L}_2^\top \boldsymbol{\alpha} = \mathbf{0} \qquad \text{versus} \qquad \neg H_0^* : \mathbf{L}_2^\top \boldsymbol{\alpha} \neq \mathbf{0}. \tag{2.9}$$

The test statistic for hypothesis (2.4) is

$$T_2 = \frac{m_k - 1}{N - dim} \frac{\frac{1}{\sigma^2} \utilde{\mathbf{Y}}^\top (\mathbf{I}_N - P_{\utilde{\mathbf{U}}}) \utilde{\mathbf{Y}}}{\utilde{\hat{\boldsymbol{\alpha}}}^\top \utilde{\mathbf{L}}_2 (\utilde{\mathbf{L}}_2^\top \boldsymbol{\Sigma} \utilde{\mathbf{L}}_2)^{-1} \utilde{\mathbf{L}}_2^\top \utilde{\hat{\boldsymbol{\alpha}}}},$$

where $\utilde{\mathbf{L}}_2^\top = \mathbf{L}_2^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{U}})^{-1} (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})$. Theorem 2.2 states that the null distribution function of $T_2$ denoted by $F_{T_2}$ is asymptotically (as $n \to \infty$) equal to the generalized F distribution

$$\frac{m_k - 1}{N - dim} \frac{\sum_{i=1}^{u} \lambda_i \chi^2(r_i)}{\chi^2(m_k - 1)},$$

with distribution function $F_2$, where $\lambda_i, r_i$ and $u$ are defined in Theorem 2.1. Suppose that $t_2$ is an observed value for $T_2$. As in (2.8), the p-value is

$$p_2 = F_2(t_2).$$

**Theorem 2.1.** *Assume that* $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{R})$ *in model (2.1). If hypothesis* $H_0$ *in (2.5) holds, then* $T_1$ *follows the distribution*

$$\frac{dim - d - 1}{N - dim} \frac{\sum_{i=1}^{u} \lambda_i \chi^2(r_i)}{\chi^2(dim - d - 1)},$$

*with distribution function* $F_1$, *where* $\lambda_1, \ldots, \lambda_u$ *denote the nonzero distinct eigenvalues of*

$$\mathbf{W}^{1/2}\mathbf{R}\mathbf{W}^{1/2}(\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}})$$

*with algebraic multiplicities* $r_1, \ldots, r_u$ *respectively, that satisfy* $\sum_{i=1}^{u} r_i = N - dim$, *and where* $\chi^2(r_1), \ldots, \chi^2(r_u), \chi^2(dim - d - 1)$ *are mutually independent.*

The proof of Theorem 2.1 is given in Section 2.8.3. The asymptotics of the following theorem holds when the number of subjects $n$ tends to infinity, the number of repeated measurements $N_i$ $(i = 1, \ldots, n)$ may or may not tend to infinity.

**Theorem 2.2.** *Assume that* $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{R})$ *in model (2.1). Define the random variable*

$$\frac{m_k - 1}{N - dim} \frac{\sum_{i=1}^{u} \lambda_i \chi^2(r_i)}{\chi^2(m_k - 1)},$$

*with distribution function* $F_2$, *where* $\lambda_1, \ldots, \lambda_u, r_1, \ldots, r_u$ *are defined in Theorem 2.1, and where* $\chi^2(r_1), \ldots, \chi^2(r_u), \chi^2(m_k - 1)$ *are mutually independent. Let* $\|\mathbf{A}\|$ *denote the Frobenius norm of a matrix* $\mathbf{A}$. *Under* $H_0$ *in (2.4)*

$$\|F_{T_2} - F_2\|_\infty =$$
$$O\left(\sqrt{M_{\xi_0}\left(N\rho_n^2\|\mathbf{R}^{-1/2}\|^2 + \sqrt{N}\rho_n\|\mathbf{R}^{-1/2}\|\right)} + \sqrt{M_{\eta_0}\left(Nw_{\max}\rho_n^2 + Nw_{\max}\rho_n\sqrt{\max_i w_i N_i^{1/2}}\right)}\right).$$

*as* $n \to \infty$, *where* $\|f\|_\infty$ *is the supremum norm of a function* $f$, $\rho_n$ *is the approximation error (see Appendix A),* $M_{\xi_0}$ *and* $M_{\eta_0}$ *are the maxima of the density function of* $\chi^2(m_k - 1)$ *and* $\sum_{i=1}^{u} \lambda_i \chi^2(r_i)$, *respectively.*

The proof of Theorem 2.2 is given in Section 2.8.4. A discussion on the bound stated in Theorem 2.2 is given in Section 2.9 which contains a bound in terms of the number of subjects, their repeated measurements and the number of knots.

#### 2.3.1.3 Testing a general hypothesis

From the proof of Theorem 2.2 we see that we can generalize the theorem for the test of any hypothesis of the following form

$$H_1^*: \mathbf{A}^\top \boldsymbol{\alpha} = \mathbf{a} \qquad \text{versus} \qquad \neg H_1^*: \mathbf{A}^\top \boldsymbol{\alpha} \neq \mathbf{a}, \tag{2.10}$$

where $\mathbf{A}$ is a known fixed nonzero matrix, and $\mathbf{a}$ is a known fixed vector, see (2.9). For instance, we can test whether $\beta_1(\cdot)$ is a polynomial of degree $q$ by using the derivative property of B-splines. The derivative of a spline function $g(t) = \sum_{j=1}^{m} \gamma_j B_j(t;q)$ having distance $\frac{1}{K}$ between the equidistant knots, is (De Boor (2001), page 116)

$$g'(t) = K \sum_{j=1}^{m-1} \Delta\gamma_{j+1} B_j(t; q-1), \tag{2.11}$$

where $\Delta\gamma_{j+1} = \gamma_{j+1} - \gamma_j$. Then, the matrix $\mathbf{A}^\top$ is defined as taking the $(q+1)$-th order differences of the B-spline coefficients corresponding to $\beta_1(\cdot)$ and $\mathbf{a} = \mathbf{0}$.

As before, to construct the RQF test statistic we define $\underset{\sim}{\mathbf{A}}^\top = \mathbf{A}^\top(\underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{R}}^{-1} \underset{\sim}{\mathbf{U}})^{-1}(\underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{U}})$. Let $r$ denote the number of rows of $\mathbf{A}^\top$. The test statistic for hypothesis (2.10) is

$$T_3 = \frac{r}{N - dim} \frac{\frac{1}{\sigma^2} \underset{\sim}{\mathbf{Y}}^\top (\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}}) \underset{\sim}{\mathbf{Y}}}{(\hat{\underset{\sim}{\boldsymbol{\alpha}}}^\top \underset{\sim}{\mathbf{A}} - \mathbf{a}^\top)(\underset{\sim}{\mathbf{A}}^\top \boldsymbol{\Sigma} \underset{\sim}{\mathbf{A}})^{-1}(\underset{\sim}{\mathbf{A}}^\top \hat{\underset{\sim}{\boldsymbol{\alpha}}} - \mathbf{a})}.$$

Denote its null distribution function by $F_{T_3}$. Define the random variable

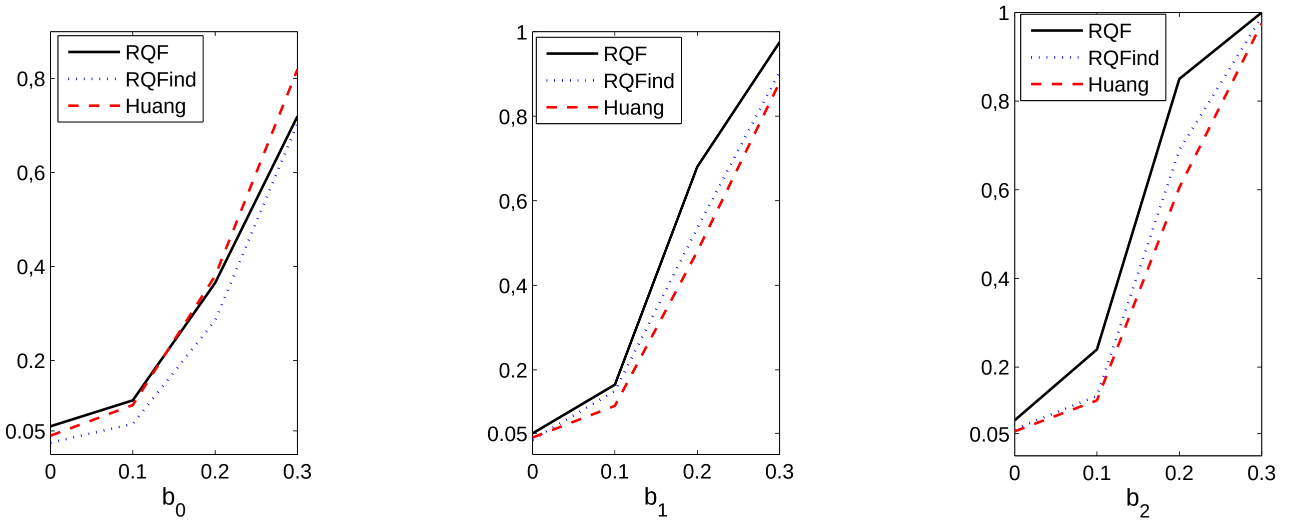
$$\frac{r}{N - dim} \frac{\sum_{i=1}^{u} \lambda_i \chi^2(r_i)}{\chi^2(r)}, \tag{2.12}$$

with distribution function $F_3$, where $\lambda_1, \ldots, \lambda_u, r_1, \ldots, r_u$ are defined in Theorem 2.2, and where $\chi^2(r_1), \ldots, \chi^2(r_u), \chi^2(r)$ are mutually independent. Suppose $t_3$ is a realization of $T_3$. As in (2.8), the p-value is

$$p_3 = F_3(t_3).$$

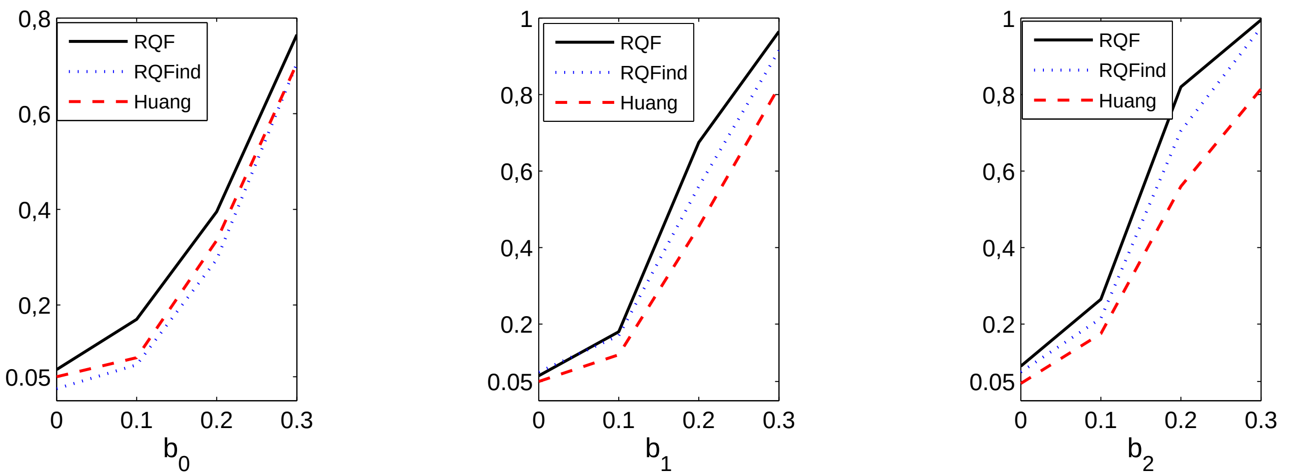
## 2.4 Simulation study

Here we discuss a simulation example where we illustrate the performance of our ratio of quadratic forms method (RQF) and compare it with Huang et al. (2002)'s bootstrap method




(a) (b) (c)

Figure 2.1: Simulation example ($n = 30$) with $\mathbf{R} = \mathbf{R}_{t1}$ and knots $(K_0, K_1, K_2) = (5, 5, 5)$. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (black line), RQFind (dotted line) and Huang (dashed line) respectively.




(a) (b) (c)

Figure 2.2: Simulation example ($n = 30$) with $\mathbf{R} = \mathbf{R}_{t1}$ and where knots are determined by CV. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (full line), RQFind (dotted line) and Huang (dashed line) respectively.

(Huang). We illustrate the importance of incorporating the correlation structure in the RQF method by also providing the RQF method where independence is assumed; referred to as RQFind.

We let the number of subjects be $n = 30$, the number of repeated measurements $N_i$ for individual $i$ is randomly generated from $\{9,\ldots,12\}$ for $i=1,\ldots,30$. For each individual $i$, the time points $t_{ij},\ j=1,\ldots,N_i$ are equidistant in $[0,1]$. We have a time dependent bivariate vector

$$\begin{pmatrix} X^{(1)}(t) \\ X^{(2)}(t) \end{pmatrix} \sim N(\mathbf{0},\boldsymbol{\Sigma}_X(t)), \quad \boldsymbol{\Sigma}_X(t) = \begin{pmatrix} \frac{3}{2} & 1/(2+t) \\ 1/(2+t) & 2 \end{pmatrix}.$$

We consider two types of intrasubject correlated errors, the first is $\mathbf{R}_{t1}$ defined by

$$\mathrm{Corr}(\varepsilon(t_{ij}),\varepsilon(t_{ik})) = \frac{1}{2}\exp(-|t_{ij}-t_{ik}|), \qquad 1 \le j \ne k \le N_i,$$

while the error terms of different subjects are mutually independent; the second is an exchangeable correlation structure $\mathbf{R}_{t2}$ defined by

$$\mathrm{Corr}(\varepsilon(t_{ij}),\varepsilon(t_{ik})) = 0.6, \qquad 1 \le j \ne k \le N_i,$$

while the error terms of different subjects are mutually independent. Furthermore, we use coefficient functions (with domain $[0,1]$):

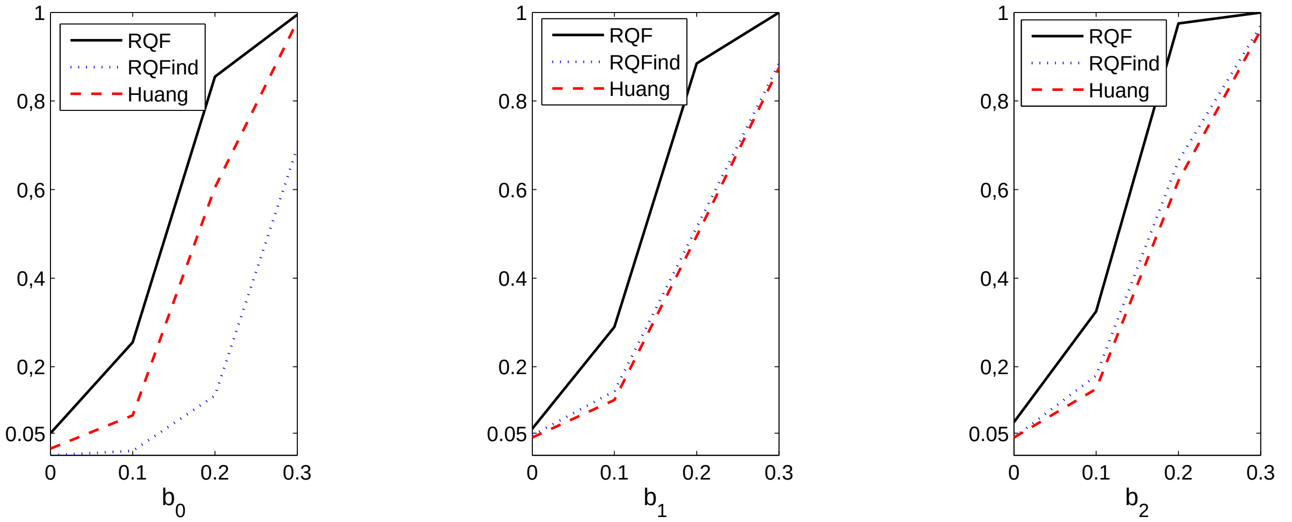
$$\beta_0(t) = 0.5(e-e^{-1}) + b_0(e^{2t-1} - 0.5(e-e^{-1})),$$

$$\beta_1(t) = 4/3 + b_1(8t(1-t)-4/3) \text{ and } \beta_2(t) = 1 + b_2(2\sin^2(2\pi t)-1),$$

where changing the parameters $b_0, b_1$ and $b_2$, changes the level of constancy. We measure the performance of our test by varying the deviation of each coefficient function $\beta_p$ ($p=0,1,2$) from a constant, that is $b_p$ varies from 0 to 1 while $b_j = 1$ for $j \in \{0,1,2\} \setminus \{p\}$.

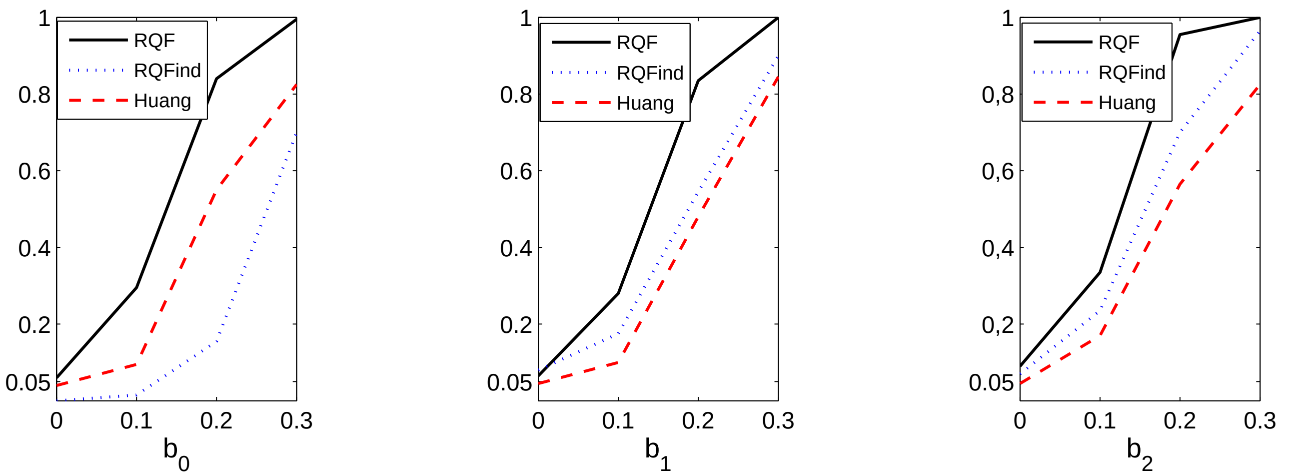
Note that we have introduced modeling bias since the coefficient functions are not spline functions (the modeling bias is of the order $O(K_n^{-4})$, where $K_n = \max_{i=0,\ldots,d} K_p$, as $n \to \infty$, see inequality (3.30) in Chapter 3). The results are based on 200 simulated data sets. In the simulations below, we follow two approaches concerning the choice of the knots. The first approach fixes the number of knots $(K_0,K_1,K_2) = (5,5,5)$. In the second approach we use a cross validation (CV) method to obtain the number of knots. Since $n=30$, it is feasible to employ the leave-one-subject-out cross-validation method (Huang et al. (2004) and references therein). The advantage of deleting the whole subject is preserving any intrasubject correlation. We delete subject $i$ from the original data to obtain the training data which we use to determine the B-spline estimator $\hat{\boldsymbol{\alpha}}^{-i}$. This is done for all the subjects $i=1,\ldots,n$, so that we can compute the cross validation score

$$CV(K_0,K_1,K_2) = \sum_{i=1}^{n} \|\mathbf{Y}_i - \mathbf{U}_i\hat{\boldsymbol{\alpha}}^{-i}\|_2^2. \tag{2.13}$$




(a) (b) (c)

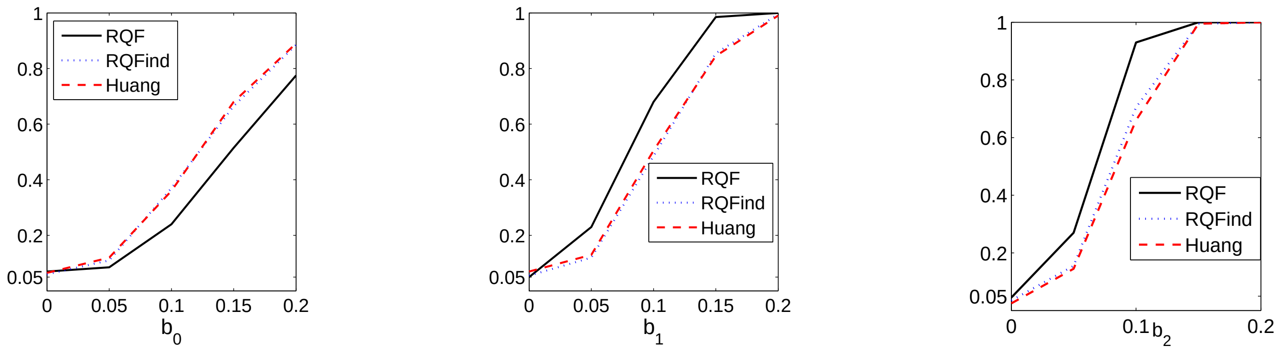
Figure 2.3: Simulation example ($n = 30$) with $\mathbf{R} = \mathbf{R}_{t2}$ and knots $(K_0, K_1, K_2) = (5, 5, 5)$. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (full line), RQFind (dotted line) and Huang (dashed line) respectively.




(a) (b) (c)

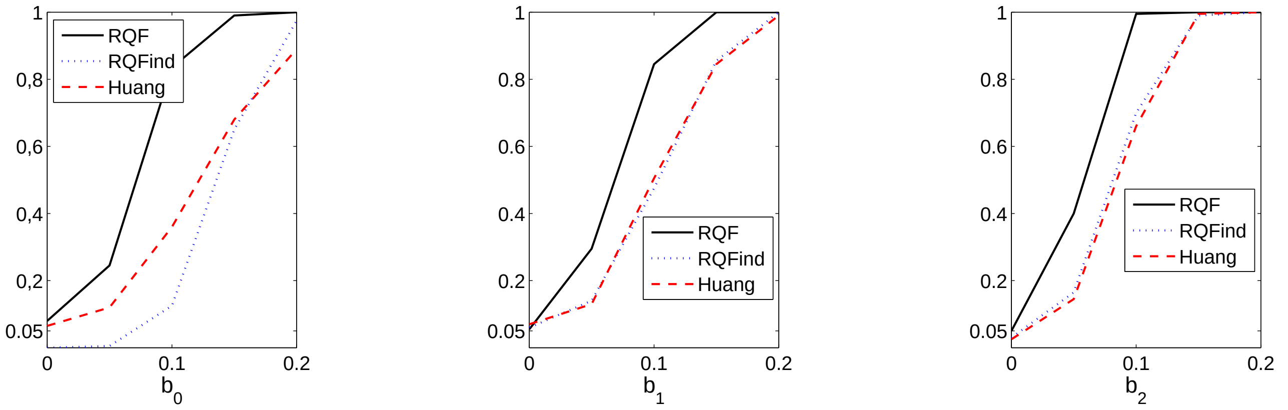
Figure 2.4: Simulation example ($n = 30$) with $\mathbf{R} = \mathbf{R}_{t2}$ and where knots are determined by CV. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (full line), RQFind (dotted line) and Huang (dashed line) respectively.

The desired $(K_0, K_1, K_2)$ is the minimizer of (2.13) where we let $(K_0, K_1, K_2)$ vary over $\{5, 6, 7, 8, 9\}^3$. The degree of the splines is fixed at 3.




(a) (b) (c)

Figure 2.5: Simulation example ($n = 60$) with $\mathbf{R} = \mathbf{R}_{t1}$ and knots $(K_0, K_1, K_2) = (5, 5, 5)$. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (full line), RQFind (dotted line) and Huang (dashed line) respectively.




(a) (b) (c)

Figure 2.6: Simulation example ($n = 60$) with $\mathbf{R} = \mathbf{R}_{t2}$ and knots $(K_0, K_1, K_2) = (5, 5, 5)$. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (full line), RQFind (dotted line) and Huang (dashed line) respectively.

The performance of the testing procedure is illustrated by the power, namely the probability $P(H_0 \text{ is rejected } |\neg H_0)$, that should be as close as possible to 1 under $\neg H_0$. In Figures 2.1-2.4 the power functions for each $b_p \in \{0, 0.1, 0.2, \dots, 1\}$, $p = 0, 1, 2$ are shown. The average computing time for a fixed knot vector of RQF is about 0.6 seconds, while Huangs bootstrap method took 22 seconds on average (bootstrap size B= 200).

When $b_p = 0$ the power functions attain approximately the theoretical level of 5%, and increase to 1 when $b_p$ increases. The RQF method performs better than Huang in all our examples, except in Figure 2.1(a) where Huang performs slightly better in the end. The RQFind approach

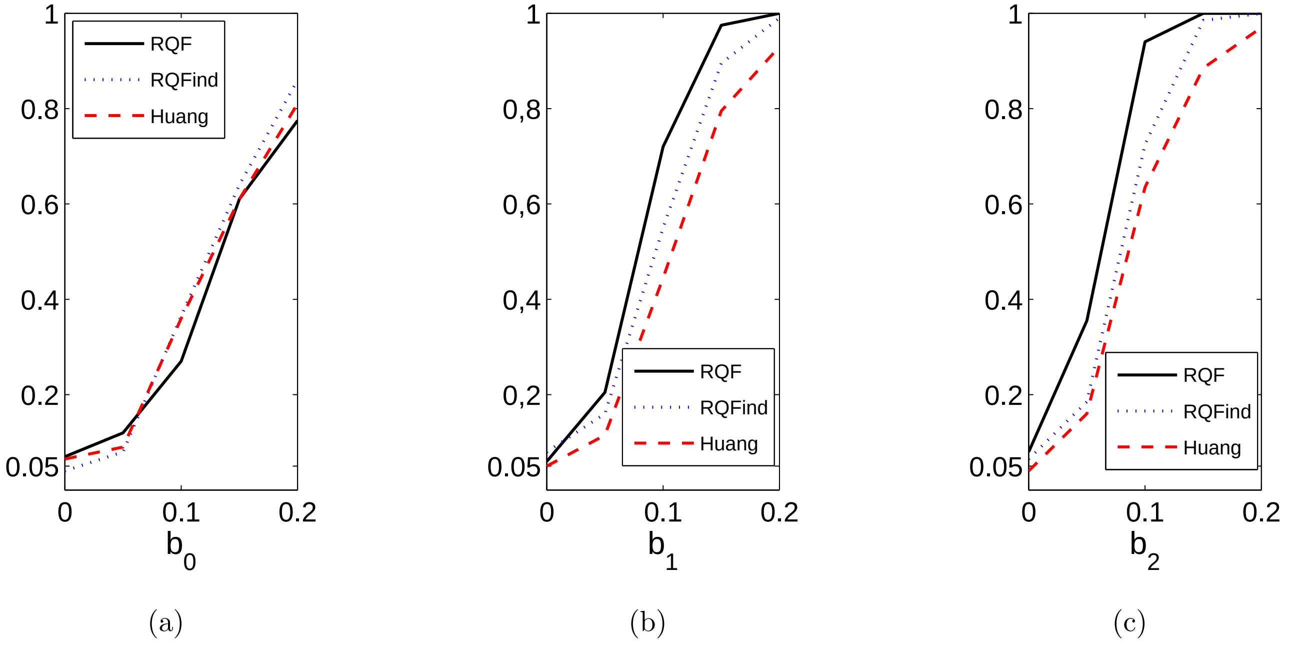



Figure 2.7: Simulation example ($n = 60$) with $\mathbf{R} = \mathbf{R}_{t1}$ and where knots are determined by CV. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (full line), RQFind (dotted line) and Huang (dashed line) respectively.

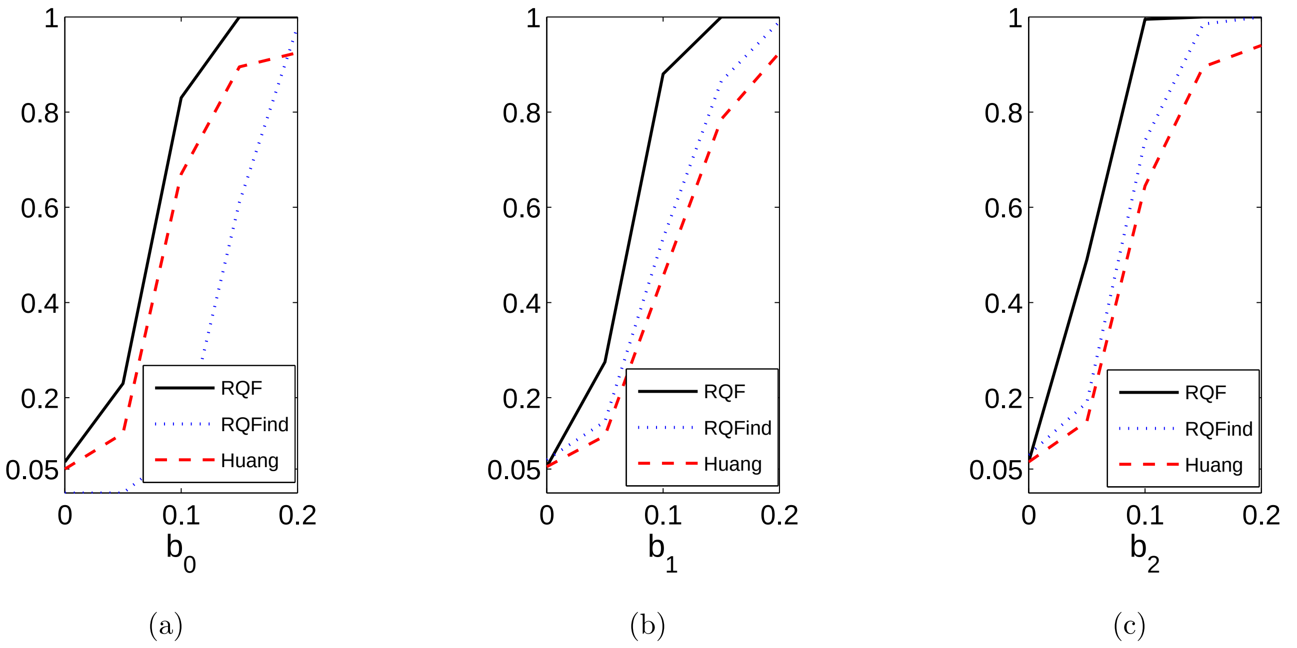



Figure 2.8: Simulation example ($n = 60$) with $\mathbf{R} = \mathbf{R}_{t2}$ and where knots are determined by CV. The power functions for the hypothesis that (a) $\beta_0(\cdot)$, (b) $\beta_1(\cdot)$ and (c) $\beta_2(\cdot)$ respectively, are constant for RQF (full line), RQFind (dotted line) and Huang (dashed line) respectively.

performs better than Huang when testing for $\beta_1(\cdot)$ and $\beta_2(\cdot)$, for $\beta_0(\cdot)$ it is the other way around. This discrepancy can be explained by the fact that $\beta_0(\cdot)$ is the intercept and thus has no covariate whereas $\beta_1(\cdot)$ and $\beta_2(\cdot)$ are multiplied with a covariate. Comparing RQF and RQFind, we see that for $\mathbf{R}_{t2}$ (Figures 2.3 and 2.4) the gaps between the power functions are bigger than for $\mathbf{R}_{t1}$ (Figures 2.1 and 2.2). The natural explanation is that the correlations in the case of $\mathbf{R}_{t1}$ (ranging from 0.18 to 0.46) are smaller than the correlations in $\mathbf{R}_{t2}$ (constant 0.6).

We now consider the same example, but with a bigger sample size. We let $n = 60$ and the number of repeated measurements $N_i$ are chosen randomly from $[18, 24]$. The RQF method takes on average 12.3 seconds for fixed knots, while Huang needs 260 seconds on average (bootstrap size B= 200). As for the knot selection we use leave 10 subjects-out cross-validation (also denoted by CV) where we divide the data in 6 fixed parts . The results can be found in Figures 2.7-2.8 with similar conclusions as for the smaller data example. Note that the power functions increase faster in this case as could be expected.

## 2.5 AIDS data

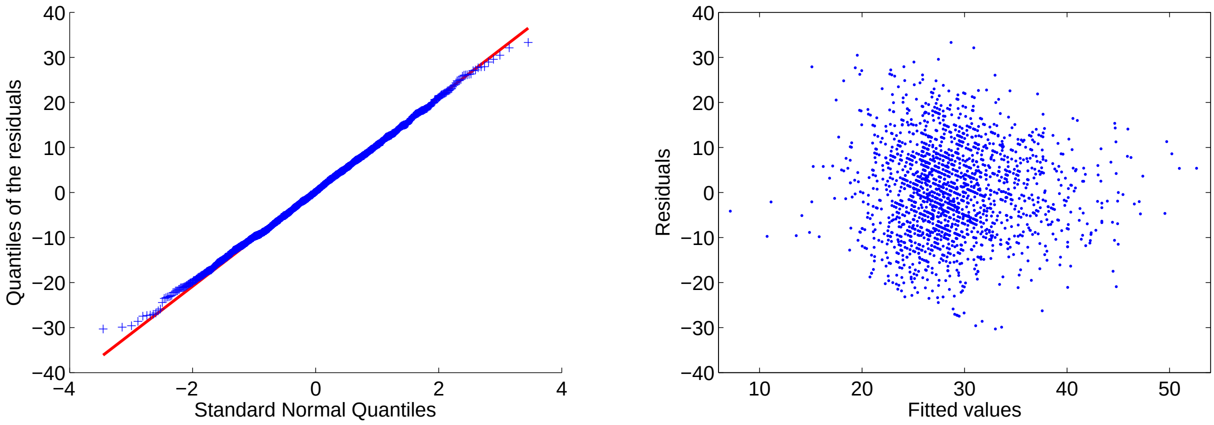
We apply our testing methodology to the AIDS data which is a subset of the Multicenter AIDS Cohort Study (MACS). We introduced the data in Section 1.2. The model we consider is

$$Y_{ij} = \beta_0(t_{ij}) + \beta_1(t_{ij})X_i^{(1)} + \beta_2(t_{ij})X_i^{(2)} + \beta_3(t_{ij})X_i^{(3)} + \varepsilon_{ij}, \tag{2.14}$$

where $Y_{ij}$ is the $i$th individual CD4 percentage measured at time $t_{ij}$, $X_i^{(1)}$ is the smoking status of the $i$th individual; $X_i^{(1)}$ is 1 or 0 if the individual ever or never smoked, $X_i^{(2)}$ is the $i$th individual's centered age at HIV infection (obtained by subtracting the sample average age at HIV infection from the individual's age at HIV infection), and in a similar way we let $X_i^{(3)}$ be the $i$th individual's centered pre-HIV infection CD4 percentage. By centering, the intercept $\beta_0(t)$ represents the mean CD4 percentage $t$ years after HIV infection of a homosexual individual with an average age at HIV infection, an average pre-HIV CD4 rate and who has never smoked cigarettes.

In our analysis we use the same B-spline bases as Huang et al. (2002) for the estimation of the coefficient functions, i.e. $(K_0, K_1, K_2, K_3) = (1, 6, 2, 4)$ and $(q_0, q_1, q_2, q_3) = (3, 3, 3, 3)$. Figure

1.2 in Chapter 1 contains the coefficient estimations.




(a) (b)

Figure 2.9: AIDS data: residual plots. (a) Q-Q plot; (b) plot of residuals against fitted values from model (2.14).

The RQF method requires the normality assumption to hold. The Q-Q plot (Figure 2.9(a)) indicates that a normal structure is quite accurate. From Figure 2.9(b) we see that up to few outlying fitted values it is reasonable to assume homoscedasticity. However, the residual plots reveal that there is intrasubject correlation. Also, the Durbin-Watson statistic is approximately 0.7 which is a sign there is considerable correlation. Therefore we also give a bootstrap version of the RQF method, which is presented in Section 2.6.2. In the bootstrap approach we work with the knots $(K_0, K_1, K_2, K_3) = (5, 5, 5, 5)$. Before setting the pseudo responses we define residuals

$$\hat{\epsilon}_{ij} = Y_{ij} - \sum_{p=0}^{3} X_{ij}^{(p)} \hat{\beta}_p(t_{ij}),$$

where $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$ is the B-spline estimator with $(K_0, K_1, K_2, K_3) = (5, 5, 5, 5)$. Let

$$Y_{ij}^{ps} = \sum_{p=0}^{3} X_{ij}^{(p)} \hat{\beta}_p^{H_0}(t_{ij}) + \hat{\epsilon}_{ij} \quad \text{for } i = 1, \ldots, n \quad \text{and } j = 1, \ldots, N_i,$$

be a set of pseudo responses under the null hypothesis. For example, for the hypothesis that $\beta_0(\cdot)$ is a constant the estimators $(\hat{\beta}_0^{H_0}, \hat{\beta}_1^{H_0}, \hat{\beta}_2^{H_0}, \hat{\beta}_3^{H_0})$ are obtained by using degrees $(q_0, q_1, q_2, q_3) = (0, 3, 3, 3)$ and knots $(K_0, K_1, K_2, K_3) = (1, 5, 5, 5)$. The bootstrap size is B=500.

Let us test whether a linear regression model makes more sense, hence we want to test hypothesis (2.5). The p-value (2.8) for this test is smaller than 1e−3 for RQFind as well as the

| | RQFind | Bootstrap | Huang et al. (2002) |
|---|---|---|---|
| Null hypothesis | **p-value** | **p-value** | **p-value** |
| $\beta_0(\cdot)$ is constant | $<$ 1e−3 | $<$ 1e−3 | $<$ 1e−3 |
| $\beta_1(\cdot)$ is constant | 0.495 | 0.146 | 0.176 |
| $\beta_2(\cdot)$ is constant | 0.153 | 0.280 | 0.301 |
| $\beta_3(\cdot)$ is constant | 0.575 | 0.066 | 0.059 |

Table 2.1: AIDS data: p-values from several methods.

bootstrap approach. Therefore we strongly reject the hypothesis that all coefficient functions are constants. In Table 2.1 the p-values are presented for testing the constancy of each coefficient function by RQFind and the bootstrap approach (denoted by Bootstrap), it also includes the corresponding results of Huang et al. (2002).

Table 2.1 shows that the results for testing the constancy of $\beta_0(\cdot), \beta_1(\cdot), \beta_2(\cdot)$ and $\beta_3(\cdot)$ are the same for both methods with significance level 0.05. Note that the p-values from our bootstrap approach and Huang's are very close. The bootstrap approaches are the same, however, we use a multidimensional instead of a one-dimensional test statistic. For $\beta_3(\cdot)$, the p-values from our bootstrap approach and Huang are on the border of being significant, while RQFind "strongly" does not reject the hypothesis that $\beta_3(\cdot)$ is a constant. This may be caused by the misspecification of the correlation structure in RQFind. This analysis suggests that the change in mean CD4 cell percentages is accounted for only by the intercept $\beta_0(\cdot)$, since the covariates are not time dependent and only $\beta_0(\cdot)$ is not constant.

## 2.6 Extension to generalized varying coefficient models

There are situations where the response $Y(t)$ is not a continuous random variable, such is the case when $Y(t)$ represents counts or categories. In such cases a generalized varying coefficient model(GVCM) is proposed, see for example Cai et al. (2000) and Verhasselt (2014). GVCMs were proposed to estimate the conditional mean of a response $Y(t)$ which is not neccessarely normal distributed. That is, the density function of the random variable $Y(t)$ at time $t$ is assumed to belong to the exponential family

$$f(Y;\theta,\phi) = \exp\left(\frac{Y\theta - b(\theta)}{a(\phi)} + c(Y,\phi)\right), \tag{2.15}$$

where $a(\cdot), b(\cdot)$ and $c(\cdot)$ are known functions, $\phi$ is a scale parameter and $\theta$ is the canonical parameter. In our setting $\theta$ depends on $\mathbf{X}(t)$ and $a(\cdot)$ is bounded. Let $\mu(\mathbf{X}(t)) = \mathrm{E}(Y(t)|\mathbf{X}(t))$ and $\sigma(\mathbf{X}(t))^2 = \mathrm{Var}(Y(t)|\mathbf{X}(t))$, then the following properties hold

$$\mu(\mathbf{X}(t)) = \frac{db}{d\theta}(\theta(\mathbf{X}(t))) \quad \text{and} \quad \sigma(\mathbf{X}(t))^2 = \frac{d^2b}{d\theta^2}(\theta(\mathbf{X}(t)))a(\phi).$$

In a GVCM we link the function

$$\kappa(\mathbf{X}(t)) = \sum_{p=0}^{d} X^{(p)}(t)\beta_p(t)$$

to the mean $\mu(\mathbf{X}(t))$ through a link function $g$: $\kappa(\mathbf{X}(t)) = g(\mu(\mathbf{X}(t)))$. The link function $g$ is called canonical if $g(\mu(\mathbf{X}(t))) = \theta(\mathbf{X}(t))$. As before, we have longitudinal data $(t_{ij}, \mathbf{X}_{ij}, Y_{ij})$ with subjects $i = 1,\ldots,n$, and number of repeated measurement $j = 1,\ldots,N_i$. We use the canonical link function, thus the model is

$$\theta(\mathbf{X}(t)) = \sum_{p=0}^{d} X^{(p)}(t) \sum_{l=1}^{m_p} B_{pl}(t;q_p)\alpha_{pl}. \tag{2.16}$$

### 2.6.1 B-spline estimator of $\boldsymbol{\alpha}$

Below we give the B-spline estimator of $\boldsymbol{\alpha}$ which is obtained by following the P-spline approach in Verhasselt (2014), by discarding the penalisation terms. The B-spline estimator is obtained by maximizing the log likelihood given by (2.15). Thus we maximize with respect to $\boldsymbol{\alpha}$

$$S(\boldsymbol{\alpha}) = -2\sum_{i=1}^{n}\frac{1}{N_i}\sum_{j=1}^{N_i}\left(\frac{Y_{ij}\theta_{ij} - b(\theta_{ij})}{a(\phi)} + c(Y_{ij},\phi)\right), \quad \text{where } \theta_{ij} = \mathbf{U}_{ij}\boldsymbol{\alpha},$$

which is equivalent to maximizing the following with respect to $\boldsymbol{\alpha}$

$$\begin{aligned} S_2(\boldsymbol{\alpha}) &= -2\sum_{i=1}^{n}\frac{1}{N_i}\sum_{j=1}^{N_i}(Y_{ij}\mathbf{U}_{ij}\boldsymbol{\alpha} - b(\mathbf{U}_{ij}\boldsymbol{\alpha})) \\ &= -2\sum_{i=1}^{n}\frac{1}{N_i}\left(\mathbf{Y}_i^\top\mathbf{U}_i\boldsymbol{\alpha} - \mathbf{1}_{N_i}^\top b(\mathbf{U}_i\boldsymbol{\alpha})\right) \\ &= -2\left(\mathbf{Y}^\top\mathbf{W}\mathbf{U}\boldsymbol{\alpha} - \mathbf{1}_N^\top\mathbf{W}b(\mathbf{U}\boldsymbol{\alpha})\right), \end{aligned}$$

where the definitions of $\mathbf{Y}, \mathbf{Y}_i, \mathbf{U}, \mathbf{U}_i$ and $\mathbf{W}$ are as before, and $\mathbf{1}_{N_i}^\top = (1,\ldots,1) \in I\!R^{N_i\times 1}$. One can use an iterative method to maximize the previous expression, for example Newton-Raphson.

### 2.6.2 Testing for parametric forms in GVCM

We briefly describe a bootstrap approach to test hypothesis (2.10) in GVCMs which is inspired by the bootstrap approach of Huang et al. (2002). The essential step is to create pseudo data $\{(Y_{ij}^{ps}, X_{ij}, t_{ij}) : i = 1, \ldots, n, j = 1, \ldots, N_i\}$ which satisfies the null hypothesis in (2.10). For the most general hypothesis ($\mathbf{A}^\top \boldsymbol{\alpha} = \mathbf{a}$), $S_2(\boldsymbol{\alpha})$ should be maximized under the constraint $\mathbf{A}^\top \boldsymbol{\alpha} = \mathbf{a}$ to obtain the estimator under the null hypothesis. When the hypothesis is that a particular coefficient function is a polynomial of degree $q$, the constraint is easily imposed by modelling that coefficient function as described in Section 2.1, so that it is estimated by a polynomial of degree $q$.

Denote by $\hat{\boldsymbol{\alpha}}^{cs}$ the estimator we obtain under the constraints imposed by the null hypothesis. Pseudo data $\{(Y_{ij}^{ps}, X_{ij}, t_{ij}) : i = 1, \ldots, n, j = 1, \ldots, N_i\}$ are simulated by using $\hat{\boldsymbol{\alpha}}^{cs}$, the model (2.16) and the density function (2.15). The test statistic is $\mathbf{A}^\top \hat{\boldsymbol{\alpha}} - \mathbf{a}$, where $\hat{\boldsymbol{\alpha}}$ is the estimator obtained without constraints. The null hypothesis is either rejected or not based on the bootstrap procedure given below.

- Step 1: Resample $n$ subjects with replacement from
$$\{(Y_{ij}^{ps}, X_{ij}, t_{ij}) : i = 1, \ldots, n, j = 1, \ldots, N_i\}$$
to obtain the bootstrap sample $\{(Y_{ij}^{ps*}, X_{ij}^*, t_{ij}^*) : i = 1, \ldots, n, j = 1, \ldots, N_i^*\}$.

- Step 2: Repeat the above resampling procedure $B$ times.

- Step 3: Obtain the test statistic vector from each bootstrap sample and derive the center of mass $\boldsymbol{\mu}_M$ and the sample covariance $\boldsymbol{\Sigma}_M$ of all test statistic vectors obtained from all the bootstrap samples. Then determine the sample distribution of all Mahalanobis distances.

- Step 4: Take the $(1-\alpha)$ percentile $M_{1-\alpha}$ of the Mahalanobis distances obtained in Step 3 and reject the null hypothesis (2.10) if $(\mathbf{A}^\top \hat{\boldsymbol{\alpha}} - \mathbf{a} - \boldsymbol{\mu}_M)^\top \boldsymbol{\Sigma}_M (\mathbf{A}^\top \hat{\boldsymbol{\alpha}} - \mathbf{a} - \boldsymbol{\mu}_M) > M_{1-\alpha}$, else do not reject the null hypothesis.

In the next chapter we apply a similar bootstrap approach (i.e. where the test statistic is a vector) to simultaneously hypothesis testing of coefficient functions.

## 2.7 Conclusion

The RQF method was introduced in the VCM setting by Li et al. (2011) as well as its theoretical motivation. They illustrated based on simulations that it is competitive with other methods in the literature. In this chapter, we extended it to VCM models for longitudinal data with intrasubject correlation. This method stands on its own due to its simplicity.

The simulations showed that the RQF method is more powerful with considerably less computing time than the bootstrap method of Huang et al. (2002). We applied the RQF method to the AIDS data. The analysis indicated that the change in mean CD4 cell percentages is only accounted for by the intercept function.

Moreover, the RQF method allows to test a series of hypotheses, by adjusting the transformation matrix on the coefficients, see (2.10). For example one could test simultaneously whether certain coefficient functions are constant, constant with a prespecified constant, polynomial, etc.

A drawback of the RQF method is that the normality assumption should not be violated and that the correlation structure should be known. When no prior information is available on the variance, it is natural to consider the RQF method where we plug in an estimate of the true variance $\mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{V}$. We have pursued this approach by estimating the variance as described in Huang et al. (2004), yielding bad results. There are several reasons for such an unsuccessful attempt. Estimating $\mathbf{V}$ is not sufficient, what is needed is a good estimation of $(\underset{\sim}{\mathbf{U}}'\mathbf{W}^{1/2}\mathbf{V}\mathbf{W}^{1/2}\underset{\sim}{\mathbf{U}})^{-1}$. Also, Theorems 2.1 and 2.2 only hold if $\mathbf{V}$ is positive definite, so more research is needed to enforce this. Instead, we proposed a bootstrap approach when there is little information on $\mathbf{V}$, which can also be applied to hypothesis testing in generalized varying coefficient models.

## 2.8 Proofs

### 2.8.1 Assumptions

**Assumption 2.1.**

1. The observation times $t_{ij},\ j = 1, \ldots, N_i,\ i = 1, \ldots, n$, are chosen independently according to a distribution function $F_T(t)$ on $\mathcal{T}$. Moreover, they are independent of the response and the covariate process $\{(Y_i(t), X_i(t))\},\ i = 1, \ldots, n$. The distribution function $F_T(t)$ has a Lebesgue density $f_T(t)$ that is bounded away from zero and infinity, uniformly over all $t \in \mathcal{T}$, that is, there exist positive constants $M_1$ and $M_2$ such that $M_1 \leqslant f_T(t) \leqslant M_2$ for $t \in \mathcal{T}$.

2. The eigenvalues $\omega_0(t), \ldots, \omega_d(t)$ of $\mathbf{\Sigma}(t) = \mathrm{E}(\mathbf{X}(t)\mathbf{X}(t)^\top)$ are bounded away from zero and infinity, uniformly over all $t \in \mathcal{T}$, that is, there exist positive constants $M_3$ and $M_4$ such that $M_3 \leqslant \omega_0(t) \leqslant \ldots \leqslant \omega_d(t) \leqslant M_4$ for $t \in \mathcal{T}$.

3. There exists a positive constant $M_5$ such that $|X_p(t)| \leqslant M_5$ for $t \in \mathcal{T}$ and $p = 0, \ldots, d$.

4. There exists a positive constant $M_6$ such that $\mathrm{E}(\varepsilon(t)^2) \leqslant M_6 < \infty$ for $t \in \mathcal{T}$.

5. $\limsup_n \frac{\max_p K_p}{\min_p K_p} < \infty$.

These conditions are commonly used (e.g. Huang et al. (2004)) and are satisfied in many practical examples. Let $K_n = \max_{i=0,\ldots,d} K_p$. As for Assumption 2.1, when dealing with deterministic time points we can replace this assumption by

$$\sup_{t \in \mathcal{T}} |F_n(t) - F_T(t) = o(1/K_n)$$

for some distribution function $F_T$ having a lebesgue density function $f_T$ which is bounded away from zero and infinity, uniformly over $t \in \mathcal{T}$, where $F_n(t) = \frac{1}{n}\sum_{i=1}^n \frac{1}{N_i}\sum_{j=1}^{N_i} 1_{t_{ij}<t}$ and $1_{t_{ij}<t}$ is the indicator function (Huang et al. (2004)). Note that we do not assume zero modeling bias, since we allow the knots to increase to infinity.

### 2.8.2 Theorem of Tan (1977)

In the proof of Theorem 3 and 4 we need the following lemma, based on Theorem 3.1 of Tan (1977).

**Lemma 2.1.** *Let $\mathbf{Z} \sim N(\boldsymbol{\mu}, \mathbf{V})$ with $\mathbf{V}$ invertible and $Q = \mathbf{Z}^\top \mathbf{A}\mathbf{Z}$, where $\mathbf{A}$ is a real symmetric matrix. Then $Q = \sum_{i=1}^z \lambda_i \chi^2(r_i, \theta_i^2)$ where $\chi^2(r_i, \theta_i^2)$ are independent noncentral chi-square* *variables, $\lambda_1, \ldots, \lambda_z$ are the nonzero distinct eigenvalues of* $\mathbf{VA}$ *with algebraic multiplicities* $r_1, \ldots, r_z$ *respectively, and*

$$\theta_j^2 = \boldsymbol{\mu}^\top \mathbf{V}^{-1} \mathbf{E}_j \boldsymbol{\mu},$$

*where* $\mathbf{VA}$ *has the spectral decomposition* $\mathbf{VA} = \sum_{j=1}^z \lambda_j \mathbf{E}_j$. *Moreover, we have that*

$$\boldsymbol{\mu}^\top \mathbf{A} \boldsymbol{\mu} = \sum_{j=1}^{z} \lambda_j \theta_j^2.$$

### 2.8.3 Proof of Theorem 2.1

*Proof.* Under hypothesis $H_1$ we have that $\beta_p(t) = \sum_l \alpha_{pl} B_{pl}(t; q_p)$ and $\alpha_{pl} = c_p$ for $l = 1, \ldots, m_p;\ p = 0, \ldots, d$. Therefore $\mathrm{E}(\underset{\sim}{\mathbf{Y}}|\mathcal{X}) = \underset{\sim}{\mathbf{U}}\boldsymbol{\alpha}$ and

$$\underset{\sim}{\mathbf{L}}_1^\top \hat{\underset{\sim}{\boldsymbol{\alpha}}} \sim N_{dim-d-1}(\mathbf{0}, \underset{\sim}{\mathbf{L}}_1^\top (\underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{U}})^{-1} \underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{R}}^{-1} \underset{\sim}{\mathbf{U}} (\underset{\sim}{\mathbf{U}}^\top \underset{\sim}{\mathbf{U}})^{-1} \underset{\sim}{\mathbf{L}}_1) =: N_{dim-d-1}(\mathbf{0}, \boldsymbol{\Sigma}_1),$$

hence we find that

$$Q_2 = \hat{\underset{\sim}{\boldsymbol{\alpha}}}^\top \underset{\sim}{\mathbf{L}}_1 \boldsymbol{\Sigma}_1^{-1} \underset{\sim}{\mathbf{L}}_1^\top \hat{\underset{\sim}{\boldsymbol{\alpha}}} \sim \chi^2(dim - d - 1).$$

The specified distribution of $Q_1 \sim \sum_{i=1}^u \lambda_i \chi^2(r_i, \theta_i^2)$ follows from Lemma 2.1 with $0 = \sum_{i=1}^u \lambda_i \theta_i^2$. We now show that $\sum_{i=1}^u r_i = N - dim$ and that all $\theta_i = 0$. Note that the idempotent matrix $(\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}})$ has eigenvalues 0 and 1. Therefore we have the decomposition $I\!R^N = \mathcal{E}_0 + \mathcal{E}_1$, where $\mathcal{E}_b$ is the eigenspace corresponding to the eigenvalue $\lambda = b$ of the matrix $(\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}})$. Moreover, $\mathcal{E}_1$ has dimension $\mathrm{trace}(\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}}) = N - dim$. Denote by $\mathcal{E}_0'$ the eigenspace of the eigenvalue $\lambda = 0$ of the matrix $\underset{\sim}{\mathbf{R}}\frac{(\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}})}{\sigma^2}$. One can verify that $\mathcal{E}_0 = \mathcal{E}_0'$. Hence, in order to find the eigenvectors corresponding to a nonzero eigenvalue we can restrict to the space $\mathcal{E}_1 \subset I\!R^N$. This also means that the $\lambda_i$ are eigenvalues of $\underset{\sim}{\mathbf{R}}$. Since $\underset{\sim}{\mathbf{R}}$ is positive definite and the fact $0 = \sum_{i=1}^u \lambda_i \theta_i^2$, we obtain that all $\theta_i = 0$. The eigenspace of $\underset{\sim}{\mathbf{R}}$ has dimension $N$, therefore

$$\sum_{i=1}^{u} r_i = N - dim.$$

It remains to show that $Q_1$ and $Q_2$ are independent. By Theorem 3.2 of Tan (1977) $Q_1$ and $Q_2$ are independent if and only if

$$\underset{\sim}{\mathbf{R}}(\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}})\underset{\sim}{\mathbf{R}}\underset{\sim}{\mathbf{R}}^{-1}\underset{\sim}{\mathbf{U}}(\underset{\sim}{\mathbf{U}}^\top\underset{\sim}{\mathbf{U}})^{-1}\underset{\sim}{\mathbf{L}}_1(\underset{\sim}{\mathbf{L}}_1^\top(\underset{\sim}{\mathbf{U}}^\top\underset{\sim}{\mathbf{U}})^{-1}\underset{\sim}{\mathbf{U}}^\top\underset{\sim}{\mathbf{R}}^{-1}\underset{\sim}{\mathbf{U}}(\underset{\sim}{\mathbf{U}}^\top\underset{\sim}{\mathbf{U}})^{-1}\underset{\sim}{\mathbf{L}}_1)^{-1}$$
$$\underset{\sim}{\mathbf{L}}_1^\top(\underset{\sim}{\mathbf{U}}^\top\underset{\sim}{\mathbf{U}})^{-1}\underset{\sim}{\mathbf{U}}^\top\underset{\sim}{\mathbf{R}}^{-1}\underset{\sim}{\mathbf{R}} = \mathbf{0}. \quad (2.17)$$

It takes a small effort to verify the equation above by noting that $P_{\underset{\sim}{\mathbf{U}}}\underset{\sim}{\mathbf{U}} = \underset{\sim}{\mathbf{U}}$. ∎

### 2.8.4 Proof of Theorem 2.2

*Proof.* The proof of this theorem is along the same lines as the proof of Theorem 3 in Li et al. (2011), some of the details are however different due to our longitudinal setting. Recall the definition of $\boldsymbol{\alpha}^*$ (see Appendix A). Set $\boldsymbol{\delta} = \mathrm{E}[\mathbf{Y}|\mathcal{X}] - \mathbf{U}\boldsymbol{\alpha}^*$, then $||\boldsymbol{\delta}||_\infty = O(\rho_n)$. We can also write $\utilde{\mathbf{Y}} = \utilde{\mathbf{U}}\boldsymbol{\alpha}^* + \mathbf{W}^{1/2}\boldsymbol{\delta} + \mathbf{W}^{1/2}\boldsymbol{\varepsilon}$, so that under hypothesis $H_0$ we obtain

$$\utilde{\mathbf{Y}} \sim N(\utilde{\mathbf{U}}\boldsymbol{\alpha}^* + \mathbf{W}^{1/2}\boldsymbol{\delta}, \utilde{\mathbf{R}}).$$

Note that $\mathbf{L}_2^\top \boldsymbol{\alpha}^* = 0$ under $H_0$, hence

$$\utilde{\mathbf{L}}_2^\top \hat{\utilde{\boldsymbol{\alpha}}} = \utilde{\mathbf{L}}_2^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \mathbf{W}^{1/2}\boldsymbol{\varepsilon} + \utilde{\mathbf{L}}_2^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \mathbf{W}^{1/2}\boldsymbol{\delta},$$

so

$$\utilde{\mathbf{L}}_2^\top \hat{\utilde{\boldsymbol{\alpha}}} \;\sim\; N_{m_k-1}(\utilde{\mathbf{L}}_2^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \mathbf{W}^{1/2}\boldsymbol{\delta}, \utilde{\mathbf{L}}_2^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{U}} (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{L}}_2)$$

Denote $\boldsymbol{\Sigma}_2 := \mathrm{Cov}(\utilde{\mathbf{L}}_2^\top \hat{\utilde{\boldsymbol{\alpha}}})$. We define

$$\xi_0 := (\mathbf{W}^{1/2}\boldsymbol{\varepsilon})^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{U}} (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{L}}_2 \boldsymbol{\Sigma}_2^{-1} \utilde{\mathbf{L}}_2^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \mathbf{W}^{1/2}\boldsymbol{\varepsilon} \sim \chi^2(m_k - 1).$$

Using Lemma 2.1, we obtain that

$$\begin{aligned}
\eta_0 &:= (\mathbf{W}^{1/2}\boldsymbol{\varepsilon})^\top \frac{(\mathbf{I}_N - P_{\utilde{\mathbf{U}}})}{\sigma^2} (\mathbf{W}^{1/2}\boldsymbol{\varepsilon}) \sim \sum_{i=1}^{u} \lambda_i \chi^2(r_i), \\
\xi_1 &:= \utilde{\mathbf{Y}}^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{U}} (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{L}}_2 \boldsymbol{\Sigma}_2^{-1} \utilde{\mathbf{L}}_2^\top (\utilde{\mathbf{U}}^\top \utilde{\mathbf{U}})^{-1} \utilde{\mathbf{U}}^\top \utilde{\mathbf{R}}^{-1} \utilde{\mathbf{Y}} \sim \chi^2(m_k - 1, \gamma^2), \\
\eta_1 &:= \utilde{\mathbf{Y}}^\top \frac{(\mathbf{I}_N - P_{\utilde{\mathbf{U}}})}{\sigma^2} \utilde{\mathbf{Y}} \sim \sum_{i=1}^{u} \lambda_i \chi^2(r_i, \theta_i^2),
\end{aligned}$$

where $\gamma^2$ and $\theta_i^2$ are specified in Lemma 2.1. Denote $\tau_0 = \frac{\eta_0}{\xi_0}$ and $\tau_1 = \frac{\eta_1}{\xi_1}$. To prove Theorem 2.2, we need to show that

$$\lim_{n\to\infty} (F_{\tau_1}(t) - F_{\tau_0}(t)) = 0 \quad \text{uniformly in} \;\; t > 0. \tag{2.18}$$

Some mathematical preparation is needed to prove (2.18). The Takagi factorization of $(\mathbf{I}_N - P_{\utilde{\mathbf{U}}})$ leads to a matrix $\mathbf{G} \in I\!R^{(N-dim)\times N}$ such that

$$\mathbf{G}^\top \mathbf{G} = \frac{(\mathbf{I}_N - P_{\utilde{\mathbf{U}}})}{\sigma^2}, \;\; \mathbf{G}\mathbf{G}^\top = \mathbf{I}_{N-dim}.$$

Throughout $\|\mathbf{A}\|$ ($\|\mathbf{c}\|$) denotes the Frobenius (Euclidean) norm of a matrix $\mathbf{A}$ (vector $\mathbf{c}$), and $\langle \mathbf{a}, \mathbf{b}\rangle$ denotes the standard inproduct of vectors $\mathbf{a}, \mathbf{b}$. Let $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \ldots, \zeta_{N-dim})^\top = \mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}$, then $\eta_0 = ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}||_2^2 = \sum_{i=1}^{N-dim} \zeta_i^2$ where

$$\boldsymbol{\zeta} \sim N(\mathbf{0}, \mathbf{G}\underset{\sim}{\mathbf{R}}\mathbf{G}^\top).$$

Let $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_{N-dim})^\top = \mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}$. Note that if $\boldsymbol{\delta} = \mathbf{0}$, then there is nothing to prove since in that case $\xi_0 = \xi_1$ and $\eta_0 = \eta_1$, so we proceed with the case $\boldsymbol{\delta} \neq \mathbf{0}$. We also have that

$$N - dim = Rank(\mathbf{G}^\top\mathbf{G}) \leq \min(Rank(\mathbf{G}^\top), Rank(\mathbf{G})) = Rank(\mathbf{G}) \leq N - dim,$$

from which it follows that $\boldsymbol{\nu} \neq \mathbf{0}$. Define an orthogonal transformation $\mathbf{T} \in I\!R^{(N-dim)\times(N-dim)}$ with first row equal to $\boldsymbol{\nu}^\top/||\boldsymbol{\nu}||$ and let $\boldsymbol{\zeta}^* = (\zeta_1^*, \zeta_2^*, \ldots, \zeta_{N-dim}^*)^\top = \mathbf{T}\boldsymbol{\zeta}$. We obtain the expressions

$$\begin{aligned}
\eta_0 &= ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}||_2^2 = \sum_{i=1}^{N-dim} (\zeta_i^*)^2 \\
\eta_1 &= ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon} + \mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||_2^2 = ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}||^2 + ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 + 2\langle \mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}, \mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}\rangle \\
&= ||\boldsymbol{\zeta}||^2 + ||\boldsymbol{\nu}||^2 + 2\langle \boldsymbol{\zeta}, \boldsymbol{\nu}\rangle \\
&= ||\boldsymbol{\zeta}||^2 + ||\boldsymbol{\nu}||^2 + 2\zeta_1^*||\boldsymbol{\nu}|| \\
&= (||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}|| + \zeta_1^*)^2 + \sum_{i=2}^{N-dim} (\zeta_i^*)^2 .
\end{aligned}$$

Therefore

$$\begin{aligned}
|\eta_1 - \eta_0| &\leq ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 + 2||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}|||\zeta_1^*| \\
\mathrm{E}(|\eta_1 - \eta_0|) &\leq ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 + 2||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||\,\mathrm{E}(|\zeta_1^*|) \\
&= ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 + 2||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||\sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}(\zeta_1^*)}, \qquad (2.19)
\end{aligned}$$

since for a mean zero normal variable $Z$ we have the property $\mathrm{E}(|Z|) = \sqrt{\frac{2}{\pi}\mathrm{Var}(Z)}$. Now $\mathrm{Var}(\boldsymbol{\zeta}^*) = \mathrm{Var}(\mathbf{T}\boldsymbol{\zeta}) = \mathbf{T}\mathbf{G}\underset{\sim}{\mathbf{R}}\mathbf{G}^\top\mathbf{T}^\top$ and $\mathbf{T}\mathbf{G}\mathbf{G}^\top\mathbf{T}^\top = \mathbf{I}_{N-dim}$. We want to bound $\mathrm{Var}(\zeta_1^*)$. Let $\mathbf{b} = (b_1, b_2, \ldots, b_N)$ denote the first row of the orthogonal matrix $\mathbf{T}\mathbf{G}$, then we know $\|\mathbf{b}\| = 1$, also denote by $\mathbf{c}_1, \ldots, \mathbf{c}_N$ the columns of $\underset{\sim}{\mathbf{R}}$. Using the fact $\langle \mathbf{b}, \mathbf{c}_i\rangle \leq \sigma^2 \max_{i=1}^n w_i\sqrt{N_i}$ which is obtained by the Cauchy-Schwarz inequality, we have that

$$\mathrm{Var}(\zeta_1^*) = \sum_{i=1}^N b_i\langle \mathbf{b}, \mathbf{c}_i\rangle \leq \sum_{i=1}^N |b_i\langle \mathbf{b}, \mathbf{c}_i\rangle| \leq \sum_{i=1}^N |\langle \mathbf{b}, \mathbf{c}_i\rangle| \leq \sigma^2 N \max_{i=1}^n w_i\sqrt{N_i}$$

Using the previous inequality, we can continue from equation (2.19) to obtain

$$\mathrm{E}(|\eta_1 - \eta_0|) \leq ||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 + 2||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||\sqrt{\frac{2}{\pi}}\sigma\sqrt{N\max_{i=1}^{n} w_i\sqrt{N_i}}. \tag{2.20}$$

Let $\mathbf{H} = \frac{1}{\sigma}\boldsymbol{\Sigma}_2^{-1/2}\underset{\sim}{\mathbf{L}}_2^\top(\underset{\sim}{\mathbf{U}}^\top\underset{\sim}{\mathbf{U}})^{-1}\underset{\sim}{\mathbf{U}}^\top\mathbf{W}^{-1/2}\mathbf{R}^{-1/2}$, then $\mathbf{H}\mathbf{H}^\top = \mathbf{I}_{m_k-1}$, $\xi_0 = ||\mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\varepsilon}/\sigma||^2$ and $\xi_1 = ||\mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\varepsilon}/\sigma + \mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\delta}/\sigma||^2$. Analogously as in (2.20) we obtain

$$\mathrm{E}(|\xi_1 - \xi_0|) \leq ||\mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\delta}/\sigma||^2 + 2||\mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\delta}/\sigma||\sqrt{\frac{2}{\pi}}, \tag{2.21}$$

since for any orthogonal transformation $\mathbf{T}_2 \in I\!R^{(m_k-1)\times(m_k-1)}$, the variance of the first component of $\boldsymbol{\kappa}^* := \mathbf{T}_2\boldsymbol{\kappa}$, where $\boldsymbol{\kappa} = \mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\varepsilon}/\sigma$ is obtained by the entry with index $(1,1)$ of the matrix

$$\mathrm{Cov}(\boldsymbol{\kappa}^*) = \frac{1}{\sigma^2}\mathbf{T}_2\mathbf{H}\mathbf{R}^{-1/2}\sigma^2\mathbf{R}\mathbf{R}^{-1/2}\mathbf{H}^\top\mathbf{T}_2^\top = \mathbf{I}_{m_k-1}.$$

Note that $\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}$ and $\mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\varepsilon}/\sigma$ are independent multivariate normal random vectors, because on the one hand

$$\mathrm{Cov}(\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}, \mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\varepsilon}/\sigma) = \mathbf{G}\mathbf{W}^{1/2}\mathbf{R}^{1/2}\mathbf{H}^\top,$$

on the other hand, by the same argument as in (2.17)

$$\mathbf{G}^\top\mathbf{G}\mathbf{W}^{1/2}\mathbf{R}^{1/2}\mathbf{H}^\top\mathbf{H} = \mathbf{0},$$

from which we find that

$$\mathbf{G}\mathbf{G}^\top\mathbf{G}\mathbf{W}^{1/2}\mathbf{R}^{1/2}\mathbf{H}^\top\mathbf{H}\mathbf{H}^\top = \mathbf{G}\mathbf{W}^{1/2}\mathbf{R}^{1/2}\mathbf{H}^\top = \mathbf{0}.$$

Hence

$$\mathrm{Cov}(\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\varepsilon}, \mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\varepsilon}/\sigma) = \mathbf{0}.$$

Fix a $t > 0$, then

$$\begin{aligned} F_{\tau_1}(t) - F_{\tau_0}(t) &= P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_0}{\xi_0} < t\right) \\ &= P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_1}{\xi_0} < t\right) + P\left(\frac{\eta_1}{\xi_0} < t\right) - P\left(\frac{\eta_0}{\xi_0} < t\right) \\ &\leq P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_1}{\xi_0} < t\right). \end{aligned} \tag{2.22}$$

For the last inequality, since $\eta_1$ and $\xi_1$ are independent, and $\eta_1$ and $\xi_0$ are independent, we have that

$$P\left(\frac{\eta_1}{\xi_0} < t\right) = \mathrm{E}_{\xi_0}\{P(\eta_1 \leq t\xi_0)|\xi_0\}$$

$$\begin{aligned}
&= \mathrm{E}_{\xi_0}\{\int_{||\mathbf{x}+\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 \le t\xi_0} f(\mathbf{x})d\mathbf{x} \quad |\xi_0\} \\
&= \mathrm{E}_{\xi_0}\{\int_{||\mathbf{x}||^2 \le t\xi_0} f(\mathbf{x}-\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta})d\mathbf{x} \quad |\xi_0\} \\
&\le \mathrm{E}_{\xi_0}\{\int_{||\mathbf{x}||^2 \le t\xi_0} f(\mathbf{x})d\mathbf{x} \quad |\xi_0\} \\
&= P\left(\frac{\eta_0}{\xi_0} < t\right),
\end{aligned}$$

where $f$ is the density function of the multivariate normal distribution $N_{N-dim}(0, \mathbf{G}\underset{\sim}{\mathbf{R}}\mathbf{G}^\top)$.

Continuing from equation (2.22) with $c$ a positive real number

$$\begin{aligned}
P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_1}{\xi_0} < t\right) &= P(\xi_0 \le \eta_1/t) - P(\xi_1 \le \eta_1/t) \\
&= P(\xi_1 \le \eta_1/t, \xi_0 \le \eta_1/t) + P(\xi_1 > \eta_1/t, \xi_0 \le \eta_1/t) - P(\xi_1 \le \eta_1/t) \\
&\le P(\xi_1 > \eta_1/t, \xi_0 \le \eta_1/t) \\
&= P(\xi_1 > \eta_1/t, \eta_1/t - c \le \xi_0 \le \eta_1/t) + P(\xi_1 > \eta_1/t, \xi_0 < \eta_1/t - c) \\
&\le P(\eta_1/t - c \le \xi_0 \le \eta_1/t) + P(\xi_0 - \xi_1 < -c) \\
&\le M_{\xi_0} c + \frac{1}{c}\mathrm{E}(|\xi_0 - \xi_1|), \qquad (2.23)
\end{aligned}$$

where $M_{\xi_0}$ is the maximum of the density function of $\xi_0$ (the Markov inequality is applied in (2.23)). Substitute

$$c = \sqrt{\frac{\mathrm{E}(|\xi_0 - \xi_1|)}{M_{\xi_0}}}$$

in (2.23) to find that

$$P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_1}{\xi_0} < t\right) \le 2\sqrt{M_{\xi_0}\,\mathrm{E}(|\xi_0 - \xi_1|)},$$

and by (2.22) we obtain that for all $t \ge 0$

$$F_{\tau_1}(t) - F_{\tau_0}(t) \le 2\sqrt{M_{\xi_0}\,\mathrm{E}(|\xi_0 - \xi_1|)}.$$

On the other hand, we obtain in a similar fashion where $c$ denotes a positive real number

$$\begin{aligned}
F_{\tau_1}(t) - F_{\tau_0}(t) &= P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_0}{\xi_0} < t\right) \\
&= P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_0}{\xi_1} < t\right) + P\left(\frac{\eta_0}{\xi_1} < t\right) - P\left(\frac{\eta_0}{\xi_0} < t\right) \\
&\ge P\left(\frac{\eta_1}{\xi_1} < t\right) - P\left(\frac{\eta_0}{\xi_1} < t\right)
\end{aligned}$$

$$
\begin{aligned}
&= -\left(P(\eta_0 \le t\xi_1) - P(\eta_1 \le t\xi_1)\right) \\
&= -\left(P(\eta_1 \le t\xi_1, \eta_0 \le t\xi_1) + P(\eta_1 > t\xi_1, \eta_0 \le t\xi_1) - P(\eta_1 \le t\xi_1)\right) \\
&\ge -\left(P(\eta_1 > t\xi_1, \eta_0 \le t\xi_1)\right) \\
&= -\left(P(\eta_1 > t\xi_1, t\xi_1 - c \le \eta_0 \le t\xi_1) + P(\eta_1 > t\xi_1, \eta_0 \le t\xi_1 - c)\right) \\
&\ge -\left(P(t\xi_1 - c \le \eta_0 \le t\xi_1) + P(\eta_0 - \eta_1 < -c)\right) \\
&\ge -\left(M_{\eta_0} c + \frac{1}{c}\,\mathrm{E}(|\eta_0 - \eta_1|)\right), \qquad (2.24)
\end{aligned}
$$

where $M_{\eta_0}$ is the maximum of the density function of the random variable $\eta_0$. Substitute in (2.24)

$$c = \sqrt{\frac{\mathrm{E}(|\eta_0 - \eta_1|)}{M_{\eta_0}}}$$

to finally establish

$$\forall t > 0 : -2\sqrt{M_{\eta_0}\,\mathrm{E}(|\eta_0 - \eta_1|)} \le F_{\tau_1}(t) - F_{\tau_0}(t) \le 2\sqrt{M_{\xi_0}\,\mathrm{E}(|\xi_0 - \xi_1|)}. \qquad (2.25)$$

Note that

$$
\begin{aligned}
||\mathbf{H}\mathbf{R}^{-1/2}\boldsymbol{\delta}/\sigma||^2 &\le ||\mathbf{R}^{-1/2}\boldsymbol{\delta}/\sigma||^2 = O(N\rho_n^2\|\mathbf{R}^{-1/2}\|^2) \\
||\mathbf{G}\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 &\le ||\mathbf{W}^{1/2}\boldsymbol{\delta}||^2 = O\left(N\rho_n^2 w_{\max}\right)
\end{aligned}
$$

since $\mathbf{H}^\top\mathbf{H}$ and $\mathbf{G}^\top\mathbf{G}$ are idempotent matrices, thus 0 and 1 are the only eigenvalues. Then by (2.19),(2.21) and (2.25), it follows that

$$
\begin{aligned}
&|F_{\tau_1}(t) - F_{\tau_0}(t)| = \\
&O\left(\sqrt{M_{\xi_0}\left(N\rho_n^2\|\mathbf{R}^{-1/2}\|^2 + \sqrt{N}\rho_n\|\mathbf{R}^{-1/2}\|\right)} + \sqrt{M_{\eta_0}\left(Nw_{\max}\rho_n^2 + Nw_{\max}\rho_n\sqrt{\max_i w_i N_i^{1/2}}\right)}\right).
\end{aligned}
$$

The proof is complete since $F_2 = F_{\tau_0}$ and $F_{T_2} = F_{\tau_1}$. ■

## 2.9 Rate of convergence

In Theorem 2.2 we assume (4.17). We shed more light on this rate by assuming that $\frac{N_{\max}}{N_{\min}}$ is bounded ($N_{\max} = \max_{i=1,\ldots,n} N_i$ and $N_{\min} = \min_{i=1,\ldots,n} N_i$), $\frac{N_{\max}^{3/2}}{n} = o(1)$ and $\frac{dim}{n} = o(1)$. Suppose that subjects with equal number of repeated measurements have the same time points, we do not need this assumption if the correlation structure does not depend on time, such is the case with any time independent correlation structure.

For the first part we use the fact that $M_{\xi_0} = O\left(\frac{1}{\sqrt{m_k - 1}}\right)$ (Li et al. (2011)), thus the first part is bounded by

$$\sqrt{\left(\frac{N\rho_n^2\|\mathbf{R}^{-1/2}\|^2}{\sqrt{m_k}} + \frac{\sqrt{N}\rho_n\|\mathbf{R}^{-1/2}\|}{\sqrt{m_k}}\right)}.$$

**Bounding $M_{\eta_0}$**

For the second part, we note that there is no closed form expression of the density function of a linear combination of chi-square variables (see Bausch (2013) among others). However, we obtain a reasonable bound on $M_{\eta_0}$ which is the maximum of the density of $\sum_{i=1}^{u} \lambda_i \chi^2(r_i)$.

First, it does not hold that $r_i = 1$ for all $i$. To prove this, suppose otherwise, i.e. $r_i = 1$ for all $i$. Then, by Theorem 2.1, we have $u = \sum_{i=1}^{u} r_i = N - dim$. Next, we obtain a bound on $u$. We argue, as in the proof of Theorem 2.1, that to find a bound on $u$ we restrict to the eigen space $\mathcal{E}_1 \subset I\!R^N$ of eigenvalue 1 of $(\mathbf{I}_N - P_{\underset{\sim}{\mathbf{U}}})$. Thus, restricting to $\mathcal{E}_1$, we only look at at the number of positive eigenvalues of $\mathbf{W}^{1/2}\mathbf{R}\mathbf{W}^{1/2}$ which is a block diagonal matrix. By the restriction on the time points (see above), $\mathbf{W}^{1/2}\mathbf{R}\mathbf{W}^{1/2}$ contains at most $N_{\max} - N_{\min} + 1$ different block matrices with dimensions not exceeding $N_{\max}$. Hence, the number of different positive eigenvalues does not exceed $N_{\max}(N_{\max} - N_{\min} + 1)$, i.e. $u \leq N_{\max}(N_{\max} - N_{\min} + 1)$. By assumption all $r_i = 1$, thus it should hold

$$N - dim = \sum_{i=1}^{u} r_i = u \leq N_{\max}(N_{\max} - N_{\min} + 1). \tag{2.26}$$

Divide (2.26) by $N$, since $N_{\max}/N_{\min}$ is bounded by $C > 0$ and $N_{\max}/n \to 0$, we obtain from the previous inequality using also the fact $N \geq nN_{\min}$, that the left hand side is $1 + o(1)$ while the right hand side is $o(1)$. This is a contradiction. Hence, there is a $1 \leq j \leq k$ such that $r_j > 1$.

Also, we can write $\sum_{i=1}^{u} \lambda_i \chi^2(r_i)$ as a sum of a scaled chi-square distribution $\lambda_{\max}\chi^2(r_{\lambda_{\max}})$ and the remaining part, where $\lambda_{\max} := \max_i \lambda_i$ is assumed to be an eigenvalue of a vector in $\mathcal{E}_1$. Moreover, we assume that $r_{\lambda_{\max}} > 1$. The density of this sum is a convolution which is bounded by $O(\frac{1}{\lambda_{\max}})$ (after a small calculation). Moreover, by Theorem 2.1 of Wolkowicz and

Styan (1980) we know that

$$\lambda_{\max} \geq \frac{Tr(\mathbf{W}^{1/2}\mathbf{R}\mathbf{W}^{1/2})}{N} = \sum_{i=1}^{n} N_i w_i / N \geq w_{\min}$$

since $\mathbf{R}$ contains only ones on its diagonal. Hence we derived $M_{\eta_0} = O(1/w_{\min})$.

### Bound on (4.17)

By the discussion above, we have the following bound on (4.17)

$$\sqrt{\left(\frac{N\rho_n^2\|\mathbf{R}^{-1/2}\|^2}{\sqrt{m_k}} + \frac{\sqrt{N}\rho_n\|\mathbf{R}^{-1/2}\|}{\sqrt{m_k}}\right)} + \sqrt{\left(\frac{w_{\max}}{w_{\min}}N\rho_n^2 + N\frac{w_{\max}}{w_{\min}}\rho_n\sqrt{\max_i w_i N_i^{1/2}}\right)}. \tag{2.27}$$

We consider a particular case for this bound to obtain a bound which solely depends on the number of subjects, their number of repeated measurements and the number of knots. Let us evaluate $\|\mathbf{R}^{-1/2}\|^2$ for the case that the intrasubject correlation is a constant $0 \leq c \leq 1$ and all number of repeated measurements are equal, i.e. $N_1 = N_2 = \ldots = N_n$. Since $\mathbf{R}^{-1/2}$ is a real symmetric matrix

$$\|\mathbf{R}^{-1/2}\| = \sqrt{Tr(\mathbf{R}^{-1})}. \tag{2.28}$$

Note that $\mathbf{R} = (\mathbf{R}_1, \ldots, \mathbf{R}_n)$ is a block diagonal matrix with all blocks being equal, thus $\mathbf{R}^{-1}$ is the block diagonal matrix composed of inverses of the block matrices in $\mathbf{R}$. We want to compute the trace of $\mathbf{R}^{-1}$, therefore we focus on $\mathbf{R}_1^{-1}$. By the Sherman-Morrison formula (equation (16) in Bartlett (1951))

$$\mathbf{R}_1^{-1} = \frac{1}{1-c}\mathbf{I}_{N_1} - \frac{c}{1-c}\frac{\mathbf{1}_{N_1}\mathbf{1}_{N_1}^\top}{1+c(N_1-1)},$$

where $\mathbf{I}_{N_1} \in I\!R^{N_1 \times N_1}$ is the identity matrix and $\mathbf{1}_{N_1} \in I\!R^{N_1 \times 1}$ is a column vector with unit components. Then $Tr(\mathbf{R}_1^{-1}) = O(N_1)$, and by (2.28) $\|\mathbf{R}^{-1/2}\|^2 = Tr(\mathbf{R}^{-1}) = O(nN_1)$.

Suppose the coefficients $\beta_p(\cdot)$, $p = 0, \ldots, d$, have bounded fourth derivatives, then by the arguments after (3.37) and Assumption 2.1.5 we know that $\rho_n = O(m_k^{-4})$. To summarise, when the weights are $w_i = \frac{1}{N_i}$, bound (2.27) becomes

$$\sqrt{\frac{nN_1}{m_k^{9/2}}\left(\frac{nN_1}{m_k^4} + 1\right)} + \sqrt{\frac{nN_1^{3/4}}{m_k^4}\left(\frac{N_1^{1/4}}{m_k^4} + 1\right)}. \tag{2.29}$$

A sufficient condition for (2.29) to go to zero as $n \to \infty$ is $\frac{nN_1}{m_k^4} = o(1)$.

# Chapter 3

# Monotonicity testing in varying coefficient models

The content of this chapter is published in Ahkim et al. (2016).

## 3.1 Introduction

As in the previous chapter, we consider the varying coefficient model (2.1). We develop procedures to test for monotonicity and convexity of coefficients $\beta_p(\cdot),\ p = 0, \ldots, d$. This can be of interest when making inference on certain covariates. One concrete example is given by our study of schizophrenic patients (Section 3.8) whose 'Severity of Illness' are modeled by a VCM with covariate the binary variable whether the patient received a drug, with coefficients depending on time (with week as unit of time). In the previous chapter we obtained the null distribution of the test statistic by assuming normality and homoscedasticity of the errors, with the correlation structure known. In this chapter the test statistic involves taking the minimum of a stochastic vector. Under normality and homoscedasticity with given correlation structure the p-value can be computed by evaluating a multivariate integral (see Section 3.4.1.2). However, this approach becomes computationally complicated for simultaneous testing. Therefore we opt for a bootstrap approach (similar as the bootstrap approach proposed in Chapter 2) without the normality assumption and the need of invoking the correlation.

The remaining of the chapter is as follows. In Section 3.2 we discuss the flexible B-spline estimator (Huang et al. (2004)), which is followed by a section with some general spline properties as how to impose monotonicity on a spline function. The testing procedures for monotonicity are presented in Section 3.4, followed by the section on testing for convexity (concavity). When there are several covariates in the model, it can be of interest to test simultaneously the shape of different coefficient functions. This is discussed in Section 3.6. We illustrate the performances of the testing procedures on simulated data in Section 3.7, and apply it to data in Section 3.8. Section 3.9 contains a short conclusion of this chapter.

## 3.2 Spline estimation

As in the previous chapter, we work with equidistant knots and assume that $\lim_{n\to\infty}\rho_n = 0$, i.e., the unknown function $\boldsymbol{\beta}$ can be uniformly approximated by spline functions of certain fixed degrees as the number of subjects $n$ and the number of knots increase.

We call the estimator $\hat{\boldsymbol{\beta}}$ uniform consistent if $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_\infty = o_P(1)$. Under certain conditions we have the uniform consistency of $\hat{\boldsymbol{\beta}}$ and its derivatives. Denote by $B^l([0,1])$ the set of real functions with domain $[0,1]$ who have a bounded $l$-th derivative. Let $K_n = \max_{p=0,\ldots,d} K_p$.

**Theorem 3.1.** *Suppose $\beta_p \in B^{q_p+1}([0,1])$. Set*

$$r_n^2 = \frac{K_n^2}{n^2}\sum_{i=1}^n \left(\frac{1}{N_i}\left(1-\frac{1}{K_n}\right)+\frac{1}{K_n}\right).$$

*Then, under Assumption 2.1,*

$$\|\hat{\beta}_p^{(v)} - \beta_p^{(v)}\|_\infty = O_P(K_n^v\rho_n + K_n^{v-q_p-1} + K_n^v r_n),$$

*for $v = 0,\ldots,q_p$, where $\beta_p^{(v)}$ is the $v$-th derivative of $\beta_p$.*

The consistency of the testing procedures is based on the approximation power given in Theorem 3.1, which is proved in Section 3.10.2. It follows immediately from Theorem 3.1 that $\hat{\boldsymbol{\beta}}$ and its derivatives are also uniform consistent.

**Corollary 3.1.** *Suppose $\beta_p \in B^{q_p+1}([0,1])$ for $p = 0,\ldots,d$. Then, under Assumption 2.1,*

$$\|\hat{\boldsymbol{\beta}}^{(v)} - \boldsymbol{\beta}^{(v)}\|_\infty = O_P(K_n^v\rho_n + K_n^{v-\min_{p=0,\ldots,d} q_p-1} + K_n^v r_n),$$

*for $v = 0,\ldots,\min_{p=0,\ldots,d} q_p$.*

## 3.3 Preliminaries

Recall that the derivative of a spline function $g(t) = \sum_{j=1}^{m} \gamma_j B_j(t;q)$ having distance $\frac{1}{K}$ between the equidistant knots, is given by (2.11). In the lemma below it is established that when $q = 2$, monotonicity of $g(t)$ in the knots $\xi_0, \cdots, \xi_K$ is equivalent to monotonicity on the whole domain $[\xi_0, \xi_K]$.

**Lemma 3.1.** *If $q = 2$, then $g'(t) \geq 0$ for all $t \in [0,1]$ if and only if $g'(\xi_i) \geq 0$ for $i = 0, 1, \ldots, K$.*

*Proof.* To see this, suppose $q = 2$, hence $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_m)^\top \in I\!R^{K+2}$ and

$$g'(t) = K \sum_{j=1}^{m-1} \Delta\gamma_{j+1} B_j(t;1).$$

Degree one B-splines have a support of 3 knots, are zero at the end knots of the support. Moreover, the sum of all B-splines evaluated in a point in $[\xi_0, \xi_K]$ is 1. Hence, $g'(\xi_0) = K\Delta\gamma_2, g'(\xi_1) = K\Delta\gamma_3, \cdots, g'(\xi_K) = K\Delta\gamma_{K+2}$. This proves the equivalences "derivative is positive on all knots" $\Leftrightarrow$ "the differences $\Delta\gamma_j, j = 2, \ldots, K+2$, are positive" $\Leftrightarrow$ "the derivative is positive on the interval $[\xi_0, \xi_K]$." ■

### 3.3.1 Constrained splines

In Section 3.4 we describe our testing procedures for the hypothesis that a certain coefficient function is increasing (decreasing). The estimation of the null distribution is based on a truthful estimation of the relevant function. Below we discuss the constraints on the B-spline coefficients which need to be added to obtain an increasing (decreasing) estimator.

#### 3.3.1.1 Quadratic splines

The derivative of a quadratic spline function $g(t) = \sum_{j=1}^{m} \gamma_j B_j(t;2)$ with B-spline basis $B_1(\cdot;2), \ldots, B_m(\cdot;2)$ which are based on the equidistant knots $\xi_0, \ldots, \xi_K$ is a linear spline function $g'(t) = K\sum_{j=1}^{m-1} \Delta\gamma_{j+1} B_j(t;1)$ with B-spline basis $B_1(\cdot;1), \ldots, B_{m-1}(\cdot;1)$. Define the matrix $\mathbf{S} \in I\!R^{(K+1)\times(K+2)}$ which consists of B-spline derivatives at the knots; $\mathbf{S}_{ij} = B_j'(\xi_{i-1};2)$. By Lemma 3.1, the function $g$ is increasing if and only if

$$\mathbf{S}\boldsymbol{\gamma} \geq \mathbf{0} \in I\!R^{K+1}, \tag{3.1}$$

where $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)^\top$. When the objective is to estimate an increasing function, we solve (2.2) under the constraint (3.1) to obtain the constrained estimator $\hat{\boldsymbol{\alpha}}^{cs}$.

#### 3.3.1.2 Cubic splines

Unlike for quadratic spline estimation where linear constraints at the knots can impose monotonicity, for cubic spline estimation we are required to impose quadratic constraints at the knots. We use a result of Karlin and Studden (1966) formulated by Daouia et al. (2016) as follows.

**Proposition 3.1.** *Let $p(x) = p_0 + p_1 x + p_2 x^2$ be a quadratic polynomial. Then $p(x) \geq 0$ for all $x \in [0, 1]$ if and only if there exists $y_0 \geq 0$ such that $(p_0 + p_2 + y_0, p_0 - p_2 - y_0, p_1 - y_0)^\top \in \mathcal{Q}_3$, where $\mathcal{Q}_{k+1} = \{(z_0, \dots, z_k) : \ z_0 \geq \|(z_1, \dots, z_k)^\top\|_2\}$ is the $(k+1)$-dimensional second order cone.*

The idea is to apply Proposition 3.1 to the derivative of a cubic spline function, which is by definition a second degree polynomial between consecutive knots, to obtain the constraints on the B-spline coefficients which ensure that the cubic spline is an increasing function. Note that Proposition 3.1 needs to be adjusted so that it can be applied to a second degree polynomial with domain a subinterval of $[0, 1]$. Daouia et al. (2016) used this approach to obtain the constraints when one works with the truncated power function basis. In Proposition 3.2 we give the constraints on $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_m)^\top$ to ensure $g(t) = \sum_{j=1}^{m} \gamma_j B_j(t; 3)$ is monotonically increasing.

**Proposition 3.2.** *The function $g'$ is nonnegative on $[\xi_0, \xi_K]$ if and only if there exists a vector $\mathbf{h} = (h_0, h_1, \dots, h_{K-1})^\top \in I\!R^{K \times 1}$ with positive components such that*

$$\mathbf{A}_j \mathbf{v} \geq \|(\mathbf{B}_j \mathbf{v}, \mathbf{C}_j \mathbf{v})^\top\|_2, \quad \textit{for} \quad j = 0, \dots, K-1, \tag{3.2}$$

*where $\mathbf{v} = ((\mathbf{D}_1 \boldsymbol{\gamma})^\top, \mathbf{h}^\top)^\top$ and $\mathbf{D}_1$ is the matrix which takes the first order differences of $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_m)^\top$, and where*

$$\mathbf{A}_j = (\underbrace{0, \dots, 0}_{j}, \frac{2}{K^2}, -\frac{1}{K^2}, \frac{1}{K^2}, \underbrace{0, \dots, 0}_{m-4}, 1, \underbrace{0, \dots, 0}_{K-j-1}) \in I\!R^{1 \times (m-1+K)}$$

$$\mathbf{B}_j = (\underbrace{0, \dots, 0}_{j}, 0, \frac{3}{K^2}, -\frac{1}{K^2}, \underbrace{0, \dots, 0}_{m-4}, -1, \underbrace{0, \dots, 0}_{K-j-1}) \in I\!R^{1 \times (m-1+K)}$$

$$\mathbf{C}_j = (\underbrace{0,\ldots,0}_{j}, \frac{-2}{K^2}, \frac{2}{K^2}, 0, \underbrace{0,\ldots,0}_{m-4}, -1, \underbrace{0,\ldots,0}_{K-j-1}) \in I\!R^{1\times(m-1+K)}$$

*for* $j = 0, \ldots, K-1$.

The proof is relegated to Section 3.10.1. We obtain the constrained cubic spline estimator $\hat{\boldsymbol{\alpha}}^{cs}$ by solving (2.2) under the corresponding constraints (3.2). We use the same notation for the constrained quadratic spline estimator when there is no ambiguity.

### 3.3.2 Selection number of knots

Here, the B-spline estimator is attained by fixing the degree vector and allowing only the knot vector to vary. We resolve to a cross validation method (2.13) to obtain a desired knot vector $(K_0, \ldots, K_d)$. The desired $K = (K_0, \ldots, K_d)$ is the minimizer of (2.13). One can also resolve to the $v$-fold cross validation method. Here we partition the data in equal parts (with respect to the subjects) $P_1, P_2, \ldots, P_v$ where all the information of one subject is contained in only one part. Then a training data is formed by deleting one part $P_i$ with which we determine the B-spline estimator $\hat{\boldsymbol{\alpha}}^{-P_i}$ and compute the cross validation score for the deleted part. The total cross validation score which we seek to minimize is

$$CV_v(K_0, \ldots, K_d) = \sum_{i=1}^{v} \|\mathbf{Y}_{P_i} - \mathbf{U}_{P_i}\hat{\boldsymbol{\alpha}}^{-P_i}\|_2^2. \tag{3.3}$$

### 3.3.3 Conditional variance estimation

Let $\mathcal{X} = \{(t_{ij}, \mathbf{X}_{ij}) : i = 1, \ldots, n, j = 1, \ldots, N_i\}$. Conditioning on $\mathcal{X}$, we obtain by (2.3)

$$\mathrm{Cov}(\hat{\boldsymbol{\alpha}}\ |\mathcal{X}) = (\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}\mathbf{U}^\top\mathbf{W}\mathbf{V}\mathbf{W}\mathbf{U}(\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}, \tag{3.4}$$

where the only unknown is $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \mathbf{V} = \mathrm{diag}(\mathbf{V}_1, \ldots, \mathbf{V}_n)$, and $\mathrm{Cov}(\boldsymbol{\varepsilon}_i) = \mathbf{V}_i$ with $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iN_i})^\top$. More explicitly

$$(\mathbf{V}_i)_{jj'} = \mathrm{Cov}(\varepsilon_i(t_{ij}), \varepsilon_i(t_{ij'})), \quad 1 \le j, j' \le N_i. \tag{3.5}$$

Huang et al. (2004) estimate $\mathrm{Cov}(\varepsilon_i(t_{ij}), \varepsilon_i(t_{ij'}))$ by a tensor product spline on $[0,1] \times [0,1]$ (Chapter 12, Schumaker (2007)), that is

$$\mathrm{Cov}(\varepsilon_i(t), \varepsilon_i(s)) \approx \sum_{k,l=1}^{m_\varepsilon} u_{kl} B_k(t, q_\varepsilon) B_l(s, q_\varepsilon), \quad t, s \in [0,1],\ t \neq s, \tag{3.6}$$

where we use a fixed set of B-splines $\{B_1(\cdot; q_\varepsilon), B_2(\cdot; q_\varepsilon), \ldots, B_{m_\varepsilon}(\cdot; q_\varepsilon)\}$ with degrees $q_\varepsilon$ and equidistant knots in $[0,1]$, and let $u_{kl} = u_{lk}$. We impose the approximation in (3.6) only when $t \neq s$, since the covariance function $\mathrm{Cov}(\varepsilon_i(t), \varepsilon_i(s))$ is not necessarily continuous at $t = s$, that is, $\lim_{s \to t} \mathrm{Cov}(\varepsilon_i(t), \varepsilon_i(s)) \neq \mathrm{Cov}(\varepsilon_i(t), \varepsilon_i(t))$, see Diggle and Verbyla (1998) and Diggle (1988) for example. Moreover, $E(\varepsilon_i(t_{ij})\varepsilon_i(t_{ij'})) = \mathrm{Cov}(\varepsilon(t_{ij}), \varepsilon(t_{ij'}))$, therefore we could estimate the coefficients $u_{kl}$ by finding the minimizer of

$$\sum_{i=1}^{n} \sum_{j=1}^{N_i} \sum_{j'=j+1}^{N_i} \left( \varepsilon_i(t_{ij})\varepsilon_i(t_{ij'}) - \sum_{k,l=1}^{m_\varepsilon} u_{kl} B_k(t_{ij}; q_\varepsilon) B_l(t_{ij'}; q_\varepsilon) \right)^2 \tag{3.7}$$

if the error terms $\varepsilon_i(t_{ij})$ were observed. Since they are not observed, we replace them by the residuals $\hat{\varepsilon}_i(t_{ij}) = Y_{ij} - \mathbf{X}_{ij}^\top \hat{\boldsymbol{\beta}}(t_{ij})$ to obtain the minimizer $\hat{u}_{kl}$ $(k, l = 1, \ldots, m_\varepsilon)$.

For the estimation of $\sigma^2(t) = \mathrm{Cov}(\varepsilon(t), \varepsilon(t))$ we use the approximation $\sigma^2(t) \approx \sum_k v_k B_k(t; q_\varepsilon)$. As before, we minimize

$$\sum_{i=1}^{n} \sum_{j=1}^{N_i} \left( \varepsilon_i^2(t_{ij}) - \sum_{k=1}^{m_\varepsilon} v_k B_k(t_{ij}; q_\varepsilon) \right)^2 \tag{3.8}$$

conditioned by $v_k \geq 0$ to obtain $\hat{v}_k,\ k = 1, \ldots, m_\varepsilon$, and define the variance estimate $\hat{\sigma}^2(t) = \sum_{k=1}^{m_\varepsilon} \hat{v}_k B_k(t; q_\varepsilon)$. Under mild conditions, this yields a consistent estimator for the covariance function (Huang et al. (2004)).

## 3.4 Monotonicity tests in VCM

We test whether $\beta_k(\cdot)$ (for a fixed $k \in \{0, \ldots, d\}$) is increasing. Hence, the hypothesis

$$H_0 : \beta_k'(t) \geq 0 \text{ for all } t \text{ in } [0,1], \qquad \text{versus} \qquad \neg H_0. \tag{3.9}$$

When $q_k = 2$ we use the idea of Wang and Meyer (2011) who worked with quadratic splines to test monotonicity in the univariate case.

To test whether $\beta_k(\cdot)$ is decreasing we use the varying coefficient model where we replace $X^{(k)}$ by $-X^{(k)}$. Then, we test whether the corresponding coefficient, which is equal to $-\beta_k(\cdot)$, is increasing.

### 3.4.1 Quadratic splines approximation

For quadratic spline functions, the monotonicity constraint is translated into a linear constraint on the B-spline coefficients (see Section 3.3.1.1). Define $\mathbf{C} = (\mathbf{0}_1 \mathbf{S} \mathbf{0}_3)$ where $\mathbf{0}_1 \in I\!R^{(K_k+1)\times\sum_{j=0}^{k-1} m_j}$,$\mathbf{0}_3 \in I\!R^{(K_k+1)\times\sum_{j=k+1}^{d} m_j}$ are matrices with entries 0 and $\mathbf{S} \in I\!R^{(K_k+1)\times(K_k+2)}$ is the matrix of derivatives at the knots of B-splines corresponding to the coefficient $\beta_k$: $\mathbf{S}_{ij} = B'_{kj}(\xi_{k(i-1)}; 2)$, see (3.1). Hence, the estimate $\hat{\beta}_k$ is increasing if and only if

$$\mathbf{C}\hat{\boldsymbol{\alpha}} \geq 0.$$

Then it is natural to take as test statistic $\min(\mathbf{C}\hat{\boldsymbol{\alpha}})$. The pseudo algorithm to test hypothesis $H_0$ is as follows.

1. Determine the unconstrained estimator $\hat{\boldsymbol{\alpha}}$, and calculate $s_{\min}$, the minimum of the slopes at the knots; $s_{\min} = \min(\mathbf{C}\hat{\boldsymbol{\alpha}})$. This is the test statistic.
2. If $s_{\min}$ is non-negative, we do not reject $H_0$.
3. If $s_{\min} < 0$, determine the distribution of $s_{\min}$ under the null hypothesis and calculate the $\alpha$ percentile (see below) $Q_\alpha$.
4. If $s_{\min}$ is smaller than the $\alpha$ percentile, then we reject $H_0$.

Below we discuss two approaches to determine the null distribution of $s_{\min}$.

#### 3.4.1.1 Bootstrap method

We use a bootstrap method to determine the null distribution of $s_{\min}$. We start with setting residuals

$$\hat{\epsilon}_{ij} = Y_{ij} - \sum_{p=0}^{d} X_{ij}^{(p)} \hat{\beta}_p(t_{ij})$$

where $\hat{\boldsymbol{\beta}}$ is the unconstrained B-spline estimator and let

$$Y_{ij}^{ps} = \sum_{p=0}^{d} X_{ij}^{(p)} \hat{\beta}_p^{cs}(t_{ij}) + \hat{\epsilon}_{ij} \quad \text{for } i = 1, \ldots, n \quad \text{and } j = 1, \ldots, N_i$$

be a set of pseudo responses under the null hypothesis with $\hat{\boldsymbol{\beta}}^{cs} = (\hat{\beta}_0^{cs}, \ldots, \hat{\beta}_d^{cs})^\top$ the constrained estimate putting the constraint on $\beta_k$. The bootstrap procedure to determine the null distribution of $s_{\min}$ goes as follows.

- Step 1: Resample $n$ subjects (with all its repeated measurements) with replacement from

$$\{(Y_{ij}^{ps}, X_{ij}, t_{ij}) : i = 1, \ldots, n, j = 1, \ldots, N_i\}$$

  to obtain the bootstrap sample $\{(Y_{ij}^{ps*}, X_{ij}^*, t_{ij}^*) : i = 1, \ldots, n, j = 1, \ldots, N_i^*\}$

- Step 2: Repeat the above sampling procedure $B$ times.

- Step 3: Obtain the test statistic $s_{\min}^*$ from each bootstrap sample and derive the empirical distribution based on all $s_{\min}^*$.

- Step 4: Take the $\alpha$ percentile $\hat{Q}_\alpha$ of the empirical distribution in Step 3 and reject the null hypothesis if $s_{\min} < \hat{Q}_\alpha$, else do not reject the null hypothesis.

#### 3.4.1.2 Multivariate normal method

This approach is useful when we have normal errors and is also considered in Wang and Meyer (2011). Assume normal errors $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n)^\top$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V}). \tag{3.10}$$

We need the function $Pr(s_{\min} \leq r),\ r \in I\!R$. Since $\mathrm{E}(\mathbf{Y}|\mathcal{X}) \approx \mathbf{U}\boldsymbol{\alpha}$ and $s_{\min} = \min(\mathbf{C}\hat{\boldsymbol{\alpha}})$, we have that $\mathbf{C}\hat{\boldsymbol{\alpha}}$ is, conditioned on $\mathcal{X}$, approximately normal with mean $\mathbf{C}\boldsymbol{\alpha}$ and covariance

$$\boldsymbol{\Sigma} = \mathbf{C}(\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}\mathbf{U}^\top\mathbf{W}\mathbf{V}\mathbf{W}\mathbf{U}(\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}\mathbf{C}^\top. \tag{3.11}$$

We obtain the expression

$$Pr(s_{\min} \leq r) = 1 - \int \cdots \int_{\{\mathbf{z}|\mathbf{z}-r\mathbf{1}\geq 0\}} \phi(\mathbf{z}; \mathbf{C}\boldsymbol{\alpha}, \boldsymbol{\Sigma})d\mathbf{z}, \text{ where } \mathbf{z}, \mathbf{1} = (1, 1, \ldots, 1)^\top \in I\!R^{(K_k+1)\times 1} \tag{3.12}$$

where $\phi(\cdot\ ; \mathbf{C}\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ denotes the multivariate normal density function with mean $\mathbf{C}\boldsymbol{\alpha}$ and covariance $\boldsymbol{\Sigma}$. We can compute (3.12) only if $\boldsymbol{\alpha}$ and $\mathbf{V}$ are known. It is clear that we can only approximate $\boldsymbol{\alpha}$ and $\mathbf{V}$. The suggestion is to use $\hat{\boldsymbol{\alpha}}^{cs}$, the constrained estimator, for approximating $\boldsymbol{\alpha}$, and the unconstrained $\hat{\boldsymbol{\alpha}}$ for estimating $\mathbf{V}$ as in Section 3.3.3; if we would use $\hat{\boldsymbol{\alpha}}^{cs}$ instead, it would lead to a biased estimate for $\mathbf{V}$ under the alternative hypothesis. Therefore, we use

$$P_{\mathbf{C}\hat{\boldsymbol{\alpha}}^{cs}, \hat{\boldsymbol{\Sigma}}}(r) = 1 - \int \cdots \int_{\{\mathbf{z}|\mathbf{z}-r\mathbf{1}\geq 0\}} \phi(\mathbf{z}; \mathbf{C}\hat{\boldsymbol{\alpha}}^{cs}, \hat{\boldsymbol{\Sigma}})d\mathbf{z}, \tag{3.13}$$

where

$$\hat{\mathbf{\Sigma}} = \mathbf{C}(\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}\mathbf{U}^\top\mathbf{W}\hat{\mathbf{V}}\mathbf{W}\mathbf{U}(\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}\mathbf{C}^\top. \tag{3.14}$$

The estimated $\alpha$ percentile is determined as follows

$$\hat{Q}_\alpha = \inf\{r \mid P_{\mathbf{C}\hat{\boldsymbol{\alpha}}^{cs},\hat{\mathbf{\Sigma}}}(r) \geq \alpha\}. \tag{3.15}$$

#### 3.4.1.3 Consistency of the test for quadratic splines

The following theorem states that the type II error tends to 0 when the coefficient function function is strictly increasing.

**Theorem 3.2.** *Assume that* $K_n\rho_n + K_n^{q_k} + K_n r_n = o(1)$*. Under Assumption 2.1, if* $\inf_{t\in[0,1]}\beta_k'(t) = \delta > 0$*, then* $P(s_{min} < \min(0, \hat{Q}_\alpha)) = o(1)$*.*

The proof of this theorem can be found in Section 3.10.3.

### 3.4.2 Cubic splines approximation

For cubic splines we look at the minimum of the derivative of

$$\hat{\beta}_k'(t) = K_k \sum_{j=1}^{m_k-1} \Delta\hat{\alpha}_{k(j+1)} B_j(t; q_k - 1)$$

for $t \in [0,1]$, see equation (2.11), where $q_k = 3$. The degrees for other coefficient functions can be arbitrary. Hypothesis $H_0$ holds if and only if $\beta_k'$ is non-negative on its domain. In practice we work with a grid $G$ of $[0,1]$, say $G = \{0, 0.001, 0.002, \cdots, 1\}$. Then we determine the minimum of $\hat{\beta}_k$ over the grid $G$ which will be the test statistic, i.e. the test statistic is $\hat{\beta}_k(c)$ for gridpoint $c$. When the test statistic is nonnegative we do not reject $H_0$. In the other case we want to measure how plausible the negative test statistic is. Therefore we look at the $\alpha$ percentile of the null distribution of $\hat{\beta}_k(c)$. The pseudo algorithm for this approach is as follows.

1. Compute $s = \min_{g\in G}\hat{\beta}_k'(g)$

2. If $s \geq 0$, do not reject $H_0$.

3. If $s < 0$, choose a $c \in G$ such that $s = \hat{\beta}_k'(c) < 0$. Determine the null distribution of $s$ (see below) and check whether $s$ is smaller than the $\alpha$ percentile $\tilde{Q}_\alpha$ .

4. If $s < \tilde{Q}_\alpha$, reject $H_0$, else do not reject $H_0$.

We are left with determining $\tilde{Q}_\alpha$ in Step 3, hence it suffices to find the null distribution of $s$.

Note that we could have used the same approach when working with quadratic splines. However, in that case we know that the minimum over the grid $G$ of the linear spline function $\hat{\beta}_k'$ is attained at a knot. This fact makes that this approach is not appropriate when we use quadratic spline estimation.

#### 3.4.2.1 Bootstrap method

The bootstrap method to determine the null distribution of $s$ is similar as before (Section 3.4.1.1).

#### 3.4.2.2 Asymptotic normality

Another approach to estimate the null distribution is motivated by the asymptotic normality of $\hat{\beta}_k'(t)$. Define

$$\mathbf{b}(t; q_k - 1) = (B_1(t; q_k - 1), B_2(t; q_k - 1), B_3(t; q_k - 1), \cdots B_{m_k-1}(t; q_k - 1))^\top \in I\!R^{(m_k-1)\times 1}, \tag{3.16}$$

and, let $\mathbf{D} \in I\!R^{(m_k-1)\times dim}$ denote the matrix such that

$$\mathbf{D}\hat{\boldsymbol{\alpha}} = (\Delta\hat{\alpha}_{k2}, \Delta\hat{\alpha}_{k3}, \cdots, \Delta\hat{\alpha}_{km_k})^\top.$$

Hence, we need to find the null distribution of $\hat{\beta}_k'(c) = K_k \mathbf{b}(c; q_k - 1)^\top \mathbf{D}\hat{\boldsymbol{\alpha}}$. This leads to the following result, for which the proof is given in Section 3.10.4.

**Theorem 3.3.** *Suppose the process $\varepsilon(t)$ can be decomposed as the sum of two independent stochastic processes, $\varepsilon^{(1)}(t)$ and $\varepsilon^{(2)}(t)$, where $\varepsilon^{(1)}(t)$ is an arbitrary mean zero process and $\varepsilon^{(2)}(t)$ is a process of measurement errors that are independent at different time points and have mean zero and constant variance $\sigma^2$. Under Assumption 2.1 in Section 2.8.1, where $\beta_p$*

*has a bounded fourth derivative with $q_p \geq 3$ for all $p$. Suppose $q_k = 3$ and $\lim_n \frac{K_n^9}{n \max_i N_i} = \infty$, then*

$$\frac{\hat{\beta}_k'(c) - \beta_k'(c)}{\sqrt{\text{Var}\left(\hat{\beta}_k(c) - \beta_k'(c)\right)}} \xrightarrow{d} N(0,1) \qquad \textit{as } n \to \infty. \tag{3.17}$$

We can use this asymptotic normality result if we can estimate the variance and $\beta_k'(c)$. For the variance, we have

$$\begin{aligned}\text{Var}\left(\hat{\beta}_k(c) - \beta_k'(c)\right) &= \text{Var}(K_k \mathbf{b}(c; q_k - 1)^\top \mathbf{D}(\hat{\boldsymbol{\alpha}} - E(\hat{\boldsymbol{\alpha}})) = K_k^2 \mathbf{b}(c; q_k - 1)^\top \mathbf{D}(\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \\ &\qquad \mathbf{U}^\top \mathbf{W} \mathbf{V} \mathbf{W} \mathbf{U}(\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \mathbf{D}^\top \mathbf{b}(c; q_k - 1) \in I\!R.\end{aligned} \tag{3.18}$$

We can estimate $\mathbf{V}$ by $\hat{\mathbf{V}}$ as it is described in Section 3.3.3 to obtain the following estimate of the variance (3.18)

$$\hat{v}_k = K_k^2 \mathbf{b}(c; q_k - 1) \mathbf{D}(\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \mathbf{U}^\top \mathbf{W} \hat{\mathbf{V}} \mathbf{W} \mathbf{U}(\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \mathbf{D}^\top \mathbf{b}(c; q_k - 1)^\top. \tag{3.19}$$

We now have an estimate for the variance, but we still need to estimate $\beta_k'(c)$ to obtain the (estimated) null distribution of $\hat{\beta}'(c)$. Therefore we take $K_k \mathbf{b}(c; q_k - 1) \mathbf{D} \hat{\boldsymbol{\alpha}}^{cs}$ as a good approximation for $\beta_k'(c)$, where $\hat{\boldsymbol{\alpha}}^{cs}$ is the constrained cubic spline estimator.

The estimated $\alpha$ percentile $\tilde{Q}_\alpha$ is the $\alpha$ percentile of a $N(K_k \mathbf{b}(c; q_k - 1) \mathbf{D} \hat{\boldsymbol{\alpha}}^{cs}, \hat{v}_k)$.

#### 3.4.2.3 Consistency

Suppose we estimate $\mathbf{V}$ by $\hat{\sigma}^2 \mathbf{I}_N$ where

$$\hat{\sigma}^2 = \frac{1}{N - dim} (\mathbf{Y} - \mathbf{U}\hat{\boldsymbol{\alpha}})^\top (\mathbf{Y} - \mathbf{U}\hat{\boldsymbol{\alpha}}).$$

When there is correlation, even with this misspecified estimator of $\mathbf{V}$, Theorem 3.4 says that as $n$ increases to infinity we correctly reject the null hypothesis with probability tending to 1 when the coefficient $\beta_k(\cdot)$ is strictly decreasing in a point in the domain. Also, when $\beta_k(\cdot)$ is strictly increasing we do not reject the null hypothesis with probability tending to 1.

**Theorem 3.4.** *We construct a grid $G_n$ which depends on $n$, such that $\min G_n = 0$, $\max G_n = 1$ and the supremum distance between two consecutive grid points tends to zero as $n \to \infty$. Under Assumption 2.1 and the condition $K_n \rho_n + K_n^{-q_k} + K_n r_n = o(1)$ such that $\lim_n \frac{K_n^3}{n} = 0$ we have the following:*

1. *Suppose* $\inf_{t\in[0,1]} \beta_k'(t) = \delta < 0$, *then* $\lim_{n\to\infty} P(s \geq \min(0, \tilde{Q}_\alpha)) = 0$.

2. *Suppose* $\inf_{t\in[0,1]} \beta_k'(t) = \delta > 0$, *then* $\lim_{n\to\infty} P(s < \min(0, \tilde{Q}_\alpha)) = 0$.

The proof of this theorem is deferred to Section 3.10.5. Theorem 3.4 states the effectiveness of this method under certain conditions when we consider for example the space $\mathcal{S} = \{\beta_k \in B^4([0,1])|\ \inf_{t\in[0,1]} \beta_k'(t) \neq 0\}$. This function space can be seen as the space of smooth functions without increasing functions with a flat spot.

## 3.5 Testing convexity

It is also of interest to test for convexity or impose convexity when estimating a coefficient function. We want to test the hypothesis that $\beta_k$ is convex, hence

$$H_1 : \beta_k''(t) \geq 0 \text{ for all } t \text{ in } [0,1], \qquad \text{versus} \qquad \neg H_1. \tag{3.20}$$

The testing procedures are analogous to these in Section 3.4, except for the obvious adjustments since we now work with the second derivative. Moreover, the consistency results in Section 3.4 are carried over under the appropriate adjustments of the conditions.

### 3.5.1 Cubic spline

When $q_k = 3$, the second derivative of the B-spline estimate $\hat{\beta}_k$ is a linear spline function. The second derivative of $\hat{\beta}_k$ is nonnegative if and only if (as in (3.1))

$$\mathbf{T}\boldsymbol{\alpha}_k \geq 0, \tag{3.21}$$

where $\mathbf{T} \in I\!R^{(K_k+1)\times(K_k+2)}$ is the matrix of second derivatives at the knots of B-splines corresponding to the coefficient $\beta_k$: $\mathbf{T}_{ij} = B_{kj}''(\xi_{k(i-1)}; 3)$. Define $\mathbf{H} = (\mathbf{0}_1\mathbf{T}\mathbf{0}_3)$ where $\mathbf{0}_1 \in I\!R^{(K_k+1)\times\sum_{j=0}^{k-1} m_j}$,$\mathbf{0}_3 \in I\!R^{(K_k+1)\times\sum_{j=k+1}^{d} m_j}$ are matrices with entries 0. Then we proceed as in Section 3.4.1.1 using the test statistic $\min(\mathbf{H}\hat{\boldsymbol{\alpha}})$. Moreover, the estimator $\hat{\boldsymbol{\alpha}}^{cs}$ is obtained under constraint (3.21).

### 3.5.2 Quartic spline

When $q_k = 4$ we use the bootstrap method as in Section 3.4.2.1. For the asymptotic normality approach (Section 3.4.2.2) we need the $\alpha$ percentile of $N(K_k^2 \mathbf{b}(c_2; q_k - 2)\mathbf{D}_2 \hat{\boldsymbol{\alpha}}^{cs}, \hat{v}_{2k})$ denoted by $\hat{Q}_{2,\alpha}$, where $c_2 = \text{argmin}_{t \in G} \hat{\beta}_k''(t)$ and $\mathbf{D}_2$ is the matrix which takes the second order differences of $\hat{\boldsymbol{\alpha}}_k^{cs}$ and

$$\hat{v}_{2k} = K_k^4 \mathbf{b}(c_2; q_k - 2)\mathbf{D}_2(\mathbf{U}^\top \mathbf{W}\mathbf{U})^{-1}\mathbf{U}^\top \mathbf{W}\hat{\mathbf{V}}\mathbf{W}\mathbf{U}(\mathbf{U}^\top \mathbf{W}\mathbf{U})^{-1}\mathbf{D}_2^\top \mathbf{b}(c_2; q_k - 2)^\top. \tag{3.22}$$

The constrained estimator $\hat{\boldsymbol{\alpha}}_k^{cs}$ is obtained as in Section 3.3.1.2 since we need to constrain a quadratic spline function. The only difference is that we work with second order differences because we work with the second order derivative of a quartic spline function.

## 3.6 Simultaneous testing

We address how to conduct a simultaneous shape test, i.e. test simultaneously whether certain coefficient functions are monotone and/or convex. Suppose we want to test for $b$ shapes (monotonicity and/or convexity). For example, we test whether $\beta_{i_1}(\cdot)$ is monotonic, $\beta_{i_2}(\cdot)$ is convex,..., $\beta_{i_b}(\cdot)$ is convex where $i_1, \ldots, i_b$ is contained in $\{0, 1, \ldots, d\}$. For simplicity we require that $i_1, i_2, \ldots, i_b$ is a mutually different sequence. Let $\mathbf{s}$ denote the column vector with length $b$ of the corresponding test statistics. Thus for the example above

$$\mathbf{s} = (\min_{t \in G} \hat{\beta}_{i_1}'(t), \min_{t \in G} \hat{\beta}_{i_2}''(t), \ldots, \min_{t \in G} \hat{\beta}_{i_b}''(t))^\top,$$

for a suitable grid $G \subset [0, 1]$ .

The first way is to test for the relevant coefficient functions based on the Bonferroni correction. Thus, we do the test for each relevant coefficient function with level $\alpha/b$. We do not reject the null hypothesis if all individual tests give a positive answer, i.e. all individual null hypotheses are not rejected. When we use cubic splines for testing monotonicity (Section 3.4) and quartic splines for testing convexity (Section 3.5), the Bonferroni correction method yields a consistent multiple test when we restrict to the appropriate space.

Otherwise, it is quite straightforward to use the bootstrap approach as before. If all the components of $\mathbf{s}$ are nonnegative, then we do not reject the null hypothesis. In the other case, we use a bootstrap method to determine how plausible such a test statistic is observed under the null hypothesis. As before we start with setting residuals

$$\hat{\epsilon}_{ij} = Y_{ij} - \sum_{p=0}^{d} X_{ij}^{(p)} \hat{\beta}_p(t_{ij})$$

where $\hat{\boldsymbol{\beta}}$ is the unconstrained B-spline estimator and let

$$Y_{ij}^{ps} = \sum_{p=0}^{d} X_{ij}^{(p)} \hat{\beta}_p^{cs}(t_{ij}) + \hat{\epsilon}_{ij} \quad \text{for } i = 1, \dots, n \quad \text{and } j = 1, \dots, N_i$$

be a set of pseudo responses under the null hypothesis with $\hat{\boldsymbol{\beta}}^{cs} = (\hat{\beta}_0^{cs}, \dots, \hat{\beta}_d^{cs})^\top$ the constrained estimate which is obtained by adding the appropriate constraints. Thus, for the example above we put the monotonicity constraint on $\beta_{i_1}(\cdot)$, the convexity constraint on $\beta_{i_2}(\cdot)$, ..., the convexity constraint on $\beta_{i_b}(\cdot)$. Then the bootstrap procedure to determine whether we do not reject the null hypothesis is as follows.

- Step 1: Resample $n$ subjects with replacement from

$$\{(Y_{ij}^{ps}, X_{ij}, t_{ij}) : i = 1, \dots, n, j = 1, \dots, N_i\}$$

  to obtain the bootstrap sample $\{(Y_{ij}^{ps*}, X_{ij}^*, t_{ij}^*) : i = 1, \dots, n, j = 1, \dots, N_i^*\}$

- Step 2: Repeat the above resampling procedure $B$ times.

- Step 3: Obtain the test statistic vector from each bootstrap sample and derive the center of mass $\boldsymbol{\mu}_M$ and the sample covariance $\boldsymbol{\Sigma}_M$ of all test statistic vectors obtained from all the bootstrap samples. Then determine the sample distribution of all Mahalanobis distances.

- Step 4: Take the $(1-\alpha)$ percentile $M_{1-\alpha}$ of the Mahalanobis distances obtained in Step 3 and reject the null hypothesis if $(\mathbf{s} - \boldsymbol{\mu}_M)^\top \boldsymbol{\Sigma}_M (\mathbf{s} - \boldsymbol{\mu}_M) > M_{1-\alpha}$, else do not reject the null hypothesis.

## 3.7 Simulation examples

In this section we evaluate the testing procedures on simulated data with significance level $\alpha = 0.05$. We assume a longitudinal model with mean zero errors. Let the number of subjects be $n$, the number of repeated measurements for subject $i$ is $N_i$ which is randomly generated from $\{\lfloor 3n/10\rfloor, \dots, \lfloor 4n/10\rfloor\}$ for $i = 1, \dots, n$. For each subject $i$, the time points $t_{ij},\ j = 1, \dots, N_i$ are equidistant in $[0, 1]$ and blurred by adding a random variable with distribution $N(0, 5\cdot 10^{-5})$. For the error structure we consider two settings. In the first setting $\varepsilon(t_{ij}) \sim N(0, 0.6^2)$ and in the second $\varepsilon(t_{ij}) \sim Un[-1.73, 1.73]$. In both settings, the error terms from different subjects are independent and the intrasubject correlation is

$$\mathrm{Corr}(\varepsilon(t_{ij}), \varepsilon(t_{ik})) = 0.2, \qquad 1 \le i \le n,\ 1 \le j, k \le N_i,\ j \ne k. \tag{3.23}$$

Throughout the simulations, the signal-to-noise ratio (SNR) is around 7. The SNR is defined by

$$\frac{\mathrm{Var}\left(\beta_0(t) + \sum_{p=0}^{d} X^{(p)}(t)\beta_p(t)\right)}{\mathrm{Var}\left(\varepsilon(t)\right)}.$$

In practice the SNR is estimated by its sample version. Moreover, the simulation results are based on 200 samples. First we study the performances for the test whether a certain coefficient function is increasing, followed by a study for testing the convexity of one coefficient function and for simultaneous testing.

### 3.7.1 Monotononicity tests

We consider $n = 50$ and $n = 100$. We use coefficient functions $\beta_0(t) = 0.25 + 2t,\ \beta_1(t) = f(t),\ \beta_2(t) = -0.5 + 10(t - 0.5)^2$. The hypothesis (3.9) is tested for $\beta_1(\cdot)$ using several functions $f$. We consider

- $f_{1,a}(t) = -2 + 2(1 + t - a\exp(-50(t - 0.5)^2))$;
- $f_2(t) = 1.1$;
- $f_3(t) = 5(t - 0.25)^2$;
- $f_4(t) = 5(t - 0.25)_+^2$.

The function $f_{1,a}(\cdot)$ is taken from Bowman et al. (1998). This function is strictly monotone for $a = 0.15$, whereas a dip appears when $a = 0.30$ and more profoundly when $a = 0.45$. The function $f_3(\cdot)$ is a parabola which is strictly decreasing on $[0, 0.25]$ and strictly increasing on

$[0.25, 1]$. Wang and Meyer (2011) used $f_4(\cdot)$, which is the same as the function $f_3(\cdot)$ but is zero for $t \leq 0.25$, to show that certain testing procedures reject too often the null hypothesis in (3.9) when the increasing function has flat parts.

We employ the time dependent bivariate vector

$$\begin{pmatrix} X^{(1)}(t) \\ X^{(2)}(t) \end{pmatrix} \sim N(\mathbf{0}, \mathbf{\Sigma}(t)), \quad \mathbf{\Sigma}(t) = \begin{pmatrix} 1 & 1/(4+t) \\ 1/(4+t) & 1 \end{pmatrix}. \tag{3.24}$$

We set for the cubic spline testing procedure $K_0 = K_1 = K_2 = 4$ with $q_0 = q_1 = q_2 = 3$, whereas for the quadratic spline testing procedure $K_0 = K_2 = 4, K_1 = 5$ with $q_0 = q_2 = 3, q_1 = 2$. The number of knots are chosen in such a way that the number of B-splines for each coefficient is the same irrespective of the degree we use.

The rejection rates are collected in Table 3.1. The quadratic spline testing procedure is based on the bootstrap method (Section 3.4.1.1) with bootstrap size equal to 200. For the cubic spline approach we look at both the bootstrap method and the method based on asymptotic normality (Sections 3.4.2.1 and 3.4.2.2) where we restrict the grid $G$ to $[0.05, 0.95]$ using 100 equidistant grid points. In the asymptotic normality approach (AN) there is the obstacle of estimating the covariance (see (3.18)). Note that Theorem 3.4 holds when we replace $\hat{\mathbf{V}}$ in (3.18) by $\hat{\sigma}^2\mathbf{I}_N$. The rejection rates in the asymptotic normality approach where we use the true covariance matrix are almost equal (differences of less than 1%) to the rejection rates using the (misspecified) covariance $\hat{\sigma}^2\mathbf{I}_N$, therefore we only report the results based on $\hat{\sigma}^2\mathbf{I}_N$.

From the results in Table 3.1 we see that the bootstrap outcomes (B) are comparable to each other and consistent as $n$ increases, i.e. the rejection rates when the function is not increasing ($f_{1,0.3}(\cdot), f_{1,0.45}(\cdot)$ and $f_3(\cdot)$) tend to 1 while the rejection rate is about 0.05 (the testing level, as the sample size increases) for increasing functions ($f_{1,0.15}(\cdot), f_2(\cdot)$ and $f_4(\cdot)$). The asymptotic normality approach (AN) rejects too often for increasing functions which are not contained in $\mathcal{S}$ (see Section 3.4.2.3 for the definition of $\mathcal{S}$), i.e. the constant function and the increasing function with a flat start (bold numbers in Table 3.1). This seems to be inherent to the approach, since the simulations for a much bigger $n = 500$ reveal extremely high rejection rates (0.995 for the constant function and 1 for the increasing function with a flat start). However, for functions in $\mathcal{S}$ ($f_{1,0.15}(\cdot), f_{1,0.30}(\cdot), f_{1,0.45}(\cdot)$ and $f_3(\cdot)$) the asymptotic normality approach performs better than the bootstrap methods. We could expect this bad performance for functions with flat parts

| $f, n$ | Characteristics of $f$ | Normal errors | | | Uniform errors | | |
|---|---|---|---|---|---|---|---|
| | | 2 | 3 (B) | 3 (AN) | 2 | 3(B) | 3(AN) |
| $f_{1,0.15}(\cdot), 50$ | increasing | 0.045 | 0.010 | 0.045 | 0.025 | 0.015 | 0.025 |
| $f_{1,0.15}(\cdot), 100$ | ” | 0.030 | 0.005 | 0.020 | 0.015 | 0.005 | 0.010 |
| $f_{1,0.3}(\cdot), 50$ | small dip | 0.445 | 0.340 | 0.900 | 0.450 | 0.340 | 0.955 |
| $f_{1,0.3}(\cdot), 100$ | ” | 0.945 | 0.990 | 1.000 | 0.965 | 0.990 | 1.000 |
| $f_{1,0.45}(\cdot), 50$ | large dip | 0.910 | 0.920 | 1.000 | 0.910 | 0.940 | 1.000 |
| $f_{1,0.45}(\cdot), 100$ | ” | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| $f_2(\cdot), 50$ | constant | 0.075 | 0.045 | **0.335** | 0.055 | 0.040 | **0.385** |
| $f_2(\cdot), 100$ | ” | 0.050 | 0.045 | **0.705** | 0.055 | 0.065 | **0.625** |
| $f_3(\cdot), 50$ | parabolic | 0.510 | 0.545 | 0.600 | 0.480 | 0.545 | 0.595 |
| $f_3(\cdot), 100$ | ” | 0.875 | 0.935 | 1.000 | 0.835 | 0.950 | 1.000 |
| $f_4(\cdot), 50$ | increasing with flat part | 0.050 | 0.045 | **0.300** | 0.045 | 0.050 | **0.265** |
| $f_4(\cdot), 100$ | ” | 0.020 | 0.045 | **0.530** | 0.035 | 0.040 | **0.515** |

Table 3.1: The rejection rates for the hypothesis that $\beta_1(\cdot)$ is increasing are stated based on 200 simulations. The functions in the most left column are consecutively substituted in $\beta_1(\cdot)$. In the case of cubic splines we differentiate between the bootstrap method (denoted by B) and the method based on asymptotic normality (AN).

since the AN approach is based on the pointwise asymptotic normality result given in Theorem 3.3, and for increasing functions with flat parts the test statistic $s$ is attained at different grid points while the AN approach assumes the grid point where the minimum is attained to be unique. Furthermore, the results for normal errors and uniform errors are comparable. The first two rows of Table 3.1 suggest that the test level is not reached as $n$ increases. For the AN approach this seems to be in conflict with the result (3.17) which uses the true coefficient function and the true variance with increasing number of knots. However, we estimate both the coefficient function and the variance with fixed number of knots (which do not increase as $n$ increase). By our remark at the end of the previous paragraph, the ‘problem’ seems to lie with the estimation of the coefficient function (under the null hypothesis). The results of the bootstrap approach which uses the same coefficient function estimation agrees with this

| Hypothesis | $n = 50$ | $n = 100$ | $n = 150$ |
|---|---|---|---|
| $\beta_0(\cdot)$ is convex (case one) | 0.04 | 0.025 | 0 |
| $\beta_1(\cdot)$ is convex (case one) | 0 | 0.47 | 1 |
| $\beta_0(\cdot)$ and $\beta_2(\cdot)$ are increasing (case two) | 1 | 1 | 1 |
| $\beta_0(\cdot)$ and $\beta_2(\cdot)$ are increasing (case three) | 0.265 | 0.915 | 0.955 |

Table 3.2: The rejection rates stated are based on 200 simulations with normal errors. These results are based on the bootstrap method using cubic splines for monotonicity testing and quartic splines for convexity testing. Three cases are considered. In each case $\beta_0(t) = 0.25 + 2t$, and further in case one $\beta_1(t) = f_{1,0.45}(t)$, $\beta_2(t) = -0.5 + 10(t - 0.5)^2$; for case two we take $\beta_1(t) = f_{1,0.30}(t)$, $\beta_2(t) = -0.5 + 10(t - 0.5)^2$; for case three we use $\beta_1(t) = -0.5 + 10(t - 0.5)^2$, $\beta_2(t) = f_{1,0.30}(t)$.

conclusion. Therefore, in practice the recommendation is to use the bootstrap approach.

### 3.7.2 Convexity and simultaneous tests

In addition, we test the convexity of $\beta_0(t) = 0.25 + 2t$ in first instance and $\beta_1(\cdot) = f_{1,0.45}(\cdot)$ in second instance, using quartic splines and the bootstrap method. To illustrate the effectiveness of our simultaneous approach, we also test whether $\beta_0(\cdot)$ and $\beta_2(\cdot)$ are both increasing using cubic splines and the bootstrap method. Table 3.2 contains the results where we have fixed the knots $K_0 = K_1 = K_2 = 4$. As before, we see that the powers tend to one and we do not reject too often when the null hypothesis holds.

## 3.8 Data applications

### 3.8.1 PBC

We consider a database of 424 patients having primary biliary cirrhosis (PBC) established and collected by the Mayo Clinic between January 1974 and May 1984. PBC is a fatal chronic liver disease of unknown cause. The database is available in R (Package ‘survival’). These 424 patients met standard eligibility criteria for a randomized, double-blinded, placebo-controlled,

| Hypothesis | PBC data | | |
|---|---|---|---|
| | 2 | 3 (B) | 3 (AN) |
| $\beta_0(\cdot)$ is increasing | 0.17 | 0.37 | 0.48 |
| $\beta_0(\cdot)$ is decreasing | 0.07 | 0.14 | 0.44 |
| $\beta_1(\cdot)$ is increasing | 0.05 | 0.05 | 0.37 |
| $\beta_1(\cdot)$ is decreasing | 0.42 | 0.38 | 0.18 |
| $\beta_2(\cdot)$ is increasing | 0.49 | 0.37 | 0.50 |
| $\beta_2(\cdot)$ is decreasing | 0.24 | 0.28 | 0.44 |

Table 3.3: The p-values concerning the PBC data. In the case of cubic splines we differentiate between the bootstrap method (denoted by B) and the method based on asymptotic normality (AN)

.

clinical trial of the drug D-penicillamine (DPCA). There was randomization in 312 of 424 cases and complete follow up to July, 1986 was attempted. Our study will be based on these 312 patients. Further information about this dataset can be found in Fleming and Harrington (1991), who gave a detailed description of the database with a thorough study. We suppose a VCM with age dependent coefficient functions. In our study we omit few youngest and oldest patients so that the data has a more dense number of observations at the boundaries and we omit the patients who were censored or received kidney transplantation, leading to 122 patients that are at least 33 and at most 71 years old. The response $Y$ denotes the logarithm of days between registration and death. The covariates are the logarithm of bilirubin in mg/dl (serum bilirubin is well established as an independent predictor of prognosis in PBC, see for example Lammers et al. (2014)) denoted by $\log(B)$ and presence of edema (0: no edema and no diuretic therapy for edema, 0.5: edema present for which no diuretic therapy was given or edema resolved with diuretic therapy, 1: edema despite diuretic therapy) denoted by $E$, edema represents the accumulation of fluids in the tissue. The logarithm of bilirubin is used in order to have covariates of a similar scale. For interpretability reasons, we translate the $\log(B)$ values such that the minimal $\log(B)$ value is 0. We are interested in the influence of these covariates and their effects on the survival time of the patient as the patient's age varies. This is studied


(a) (b)

Figure 3.1: In (a) the estimations of mean survival time of the VCM (3.25) as a function of age for patients with Edema=0 and three different bilirubin levels are depicted. In (b) the analogous case is depicted with Edema=1.

by by considering the VCM

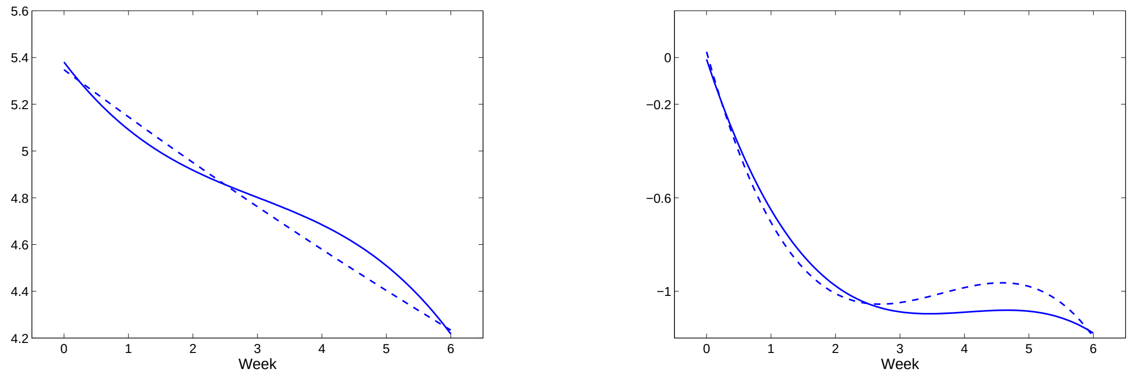
$$Y = \beta_0(age) + \beta_1(age)\log(B) + \beta_2(age)E + \varepsilon(age). \tag{3.25}$$

The question is whether the coefficients $\beta_0(\cdot), \beta_1(\cdot)$ and $\beta_2(\cdot)$ are increasing or decreasing. The number of knots are determined by a 3-fold cross validation where the number of knots vary from 3 to 8. The results are given in Table 3.4. For completeness, we also include the asymptotic normality results. However, as we have concluded in the end of Section 3.7 we base our judgments on the bootstrap approach.

From Table 3.4 we see that for the intercept ($\beta_0(\cdot)$) and the coefficient of edema ($\beta_2(\cdot)$) we do not reject the hypothesis that they are increasing, neither do we reject that they are decreasing. Therefore, it is likely that $\beta_0(\cdot)$ and $\beta_2(\cdot)$ are constant functions, i.e. age independent. For $\beta_1(\cdot)$ (coefficient of log bilirubin) we conclude, on the boundary, that it is decreasing. This reveals us that bilirubin has a decreasing impact on the survival time of the patient as the age increases. This result is depicted in Figure 3.1(b), where we only see the combined intercept and bilirubin effect on the survival time $Y$. There we see that for higher bilirubin levels the decreasing trend on the survival time is more eminent. The same conclusion holds for Figure 3.1(a) with $E = 1$ revealing the combined effects of the intercept and the covariates.

### 3.8.2 Schizophrenia data

The second data example is from the National Institute of Mental Health Schizophrenia Collaborative Study. Specifically, we study Item 79, ‘Severity of Illness’, of the Inpatient Multidi-




(a) (b)

Figure 3.2: Schizophrenia data. The full lines are the cubic spline estimations (i.e. the degree vector is $(3,3)$) of the coefficients $\beta_0(\cdot)$ and $\beta_1(\cdot)$, the dashed lines are the estimations when the degree vector is $(3,2)$. (a) contains the estimations of $\beta_0(\cdot)$, and (b) the estimations of $\beta_1(\cdot)$.

mensional Psychiatric Scale (IMPS; Lorr and Klett (1966)). Item 79 was originally measured on a numerical scale ranging from 1 (normal, not at all ill) to 7 (among the most extremely ill). In this study, most patients were measured at weeks 0, 1, 3 and 6; however, a few patients were additionally measured at weeks 2, 4 and 5. The $n = 437$ patients were randomly assigned to either receive a drug or a placebo. The data are available in R (Package 'vcrpart'). Previously, these data were studied by for example Hedeker and Gibbons (1997) who used a random-effects pattern-mixture model for the changes of the 'Severity of Illness' measurements. Here, we study the changing of Item 79 ($= Y$) with the VCM

$$Y(week) = \beta_0(week) + \beta_1(week)Drug + \varepsilon(week), \tag{3.26}$$

$Drug$ is a binary variable where $Drug = 1$ denotes a patient who received the drug, and $Drug = 0$ means that the patient received a placebo. The number of knots are determined by a 4-fold cross validation where the number of knots vary from 1 to 8. This yields the knot vector $(1,1)$ for both the degree vectors $(3,3)$ and $(3,2)$. As for the degree vector $(2,3)$, the knot vector obtained is $(2,1)$. In Figure 3.2 we depict the cubic spline estimators for $\beta_0(\cdot)$ and $\beta_1(\cdot)$ (the solid curves in respectively Figures 3.2(a) and (b)). We also present, as dashed curves, the spline estimators when using the degree vector $(2,3)$. The question of interest is how the drug affects the illness of the patients, that is the coefficient $\beta_1(\cdot)$. A negative $\beta_1(\cdot)$ which is decreasing indicates that the drug is effective. The full line in Figure 3.2(b) suggests that such is the case. Moreover, from Figure 3.2(b) the drug effect drops quickly to reach a

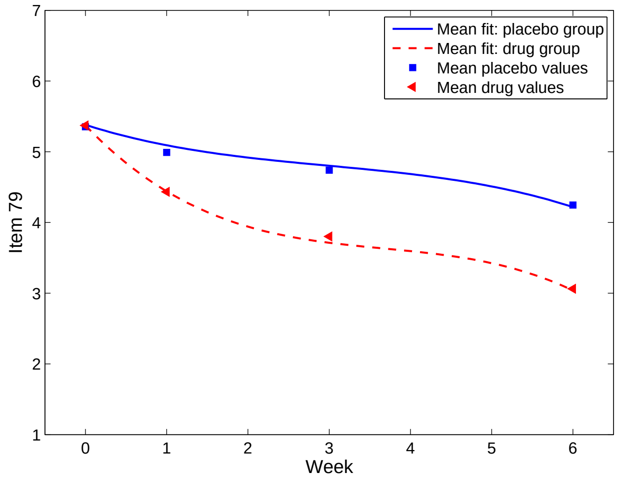



Figure 3.3: Schizophrenia data. The mean fits from model (3.26) are shown of the placebo group and the drug group. The squares and triangles are the mean Item 79 measurements at weeks 0, 1, 3 and 6, of the placebo group and drug group, respectively.

steady effect of $-1$ from week 3 onwards. Figure 3.2(a) shows an overall mildly decreasing trend for the intercept function $\beta_0(\cdot)$, revealing a little improvement of the illness over time (in weeks).

Figure 3.3 contains the mean fits, i.e. $\hat{\beta}_0(week) + \hat{\beta}_1(week)Drug$, for the placebo group and the drug group. We see that the varying of mean Item 79 measurements are well described by model (3.26) for both groups.

Table 3.4 contains the results of the monotonicity tests. The asymptotic normality results are included for completeness. However, as we have concluded in the end of Section 3.7, we base our judgments on the bootstrap approach. For test level 0.10 we reject the hypothesis that $\beta_0(\cdot)$ is increasing, moreover, the high p-values for the decreasing hypothesis indicates that $\beta_0(\cdot)$ is decreasing. For $\beta_1(\cdot)$ we totally reject the hypothesis that it is increasing. For cubic splines we do not reject the hypothesis that $\beta_1(\cdot)$ is decreasing. On contrary is the result for quadratic splines, because the cubic fit is monotonically decreasing while the quadratic fit is not from week 4 onwards (see Figure 3.2(b)). This discrepancy can be explained by noting that very few measurements were taken on weeks 4 and 5.

| Hypothesis | Schizophrenia data | | |
|---|---|---|---|
| | 2 | 3 (B) | 3 (AN) |
| $\beta_0(\cdot)$ is increasing | 0.08 | 0.05 | 0.21 |
| $\beta_0(\cdot)$ is decreasing | 0.69 | 0.65 | 0.50 |
| $\beta_1(\cdot)$ is increasing | 0.00 | 0.00 | 0.02 |
| $\beta_1(\cdot)$ is decreasing | 0.03 | 0.70 | 0.35 |

Table 3.4: The p-values concerning the Schizophrenia data with bootstrap size B= 500. In the case of cubic splines we differentiate between the bootstrap method (denoted by B) and the method based on asymptotic normality (AN).

## 3.9 Conclusion

In this chapter we presented two approaches for (simultaneous) shape testing in varying coefficient models. We showed in simulation examples that in general the bootstrap approach is consistent and yields the best results. One application modeled the 'Severity of Ilness' measurements from the Schizophrenia data by a varying coefficient model with covariate a binary variable which indicates whether the patient received a drug, and where the coefficients depend on the week number. Our interest was in the drug coefficient and whether it is decreasing, which would reveal how affective the drug is. It turned out that the drug coefficient is negative and mainly monotonically decreasing as the age increases.

Varying coefficient models are also applied in other contexts, in particular in the generalized context (see Section 2.6) and in survival analysis (see Fan and Wenyang (2008) for an overview). It should be noted that the bootstrap approach is quite universal because you mainly need an estimation of the B-spline coefficients which satisfies the null hypothesis, i.e. you need to add the relevant constraints on the B-spline coefficients to the optimization problem.

# 3.10 Proofs

## 3.10.1 Proof of Proposition 3.2

*Proof.* The function $g(t)$ is a cubic polynomial between consecutive knots $[\xi_j, \xi_{j+1}]$. Below we show that the constraint to be monotonically increasing on $[\xi_j, \xi_{j+1}]$ coincides with (3.2) for this particular $j$.

We proceed with the derivative of a cubic spline function $g(t) = \sum_{j=1}^{m} \gamma_j B_j(t;3)$ with B-spline basis $B_1(\cdot;3), \dots, B_m(\cdot;3)$. Note that $g'(t) = K \sum_{j=1}^{m-1} \Delta\gamma_{j+1} B_j(t;2)$ with B-spline basis $B_1(\cdot;2), \dots, B_{m-1}(\cdot;2)$ and knots $\xi_0, \xi_1, \dots, \xi_K$. Next, we use the recursive definition of B-splines to determine $B_1(\cdot;2), \dots, B_{m-1}(\cdot;2)$. We let $B_j(\cdot;3)$ start at knot $\xi_{j-4}$, when $j < 4$ this is achieved by adding equidistant knots at the left of $\xi_0$. Hence, $B_j(\cdot;2)$ starts at knot $\xi_{j-3}$ and ends in $\xi_j$, for $j = 1 \dots, m-1$. Then, the relevant B-splines for this interval are $B_{j+1}(\cdot;2), B_{j+2}(\cdot;2)$ and $B_{j+3}(\cdot;2)$. By equation (14) on page 90 in De Boor (2001), the recursive B-spline relations with equidistant knots are

$$
\begin{aligned}
B_j(t;2) &= \frac{(t-\xi_{j-3})}{2/K} B_j(t;1) + \frac{(\xi_j - t)}{2/K} B_{j+1}(t;1) \\
B_j(t;1) &= \frac{(t-\xi_{j-3})}{1/K} B_j(t;0) + \frac{(\xi_{j-1} - t)}{1/K} B_{j+1}(t;0) \\
B_j(t;0) &= \begin{cases} 1, & \text{if } \xi_{j-3} \le t < \xi_{j-2} \\ 0, & \text{else,} \end{cases}
\end{aligned} \tag{3.27}
$$

where it should be noted that in De Boor (2001) the indices are different because we use a different indexation and only use equidistant single knots (i.e. all knots have multiplicity one). Using (3.27), we find that restricted to the interval $[\xi_j, \xi_{j+1}]$

$$
\begin{aligned}
B_{j+1}(t;2) &= \frac{K^2}{2}(t^2 - 2\xi_{j+1} t + \xi_{j+1}^2) \\
B_{j+2}(t;2) &= \frac{K^2}{2}(-2t^2 + t(\xi_{j-1} + \xi_j + \xi_{j+1} + \xi_{j+2}) - \xi_{j-1}\xi_{j+1} - \xi_j \xi_{j+2}) \\
B_{j+3}(t;2) &= \frac{K^2}{2}(t^2 - 2\xi_j t + \xi_j^2),
\end{aligned}
$$

where the coefficients of $B_{j+1}(\cdot;2), B_{j+2}(\cdot;2)$ and $B_{j+3}(\cdot;2)$, are $(\gamma_{j+2} - \gamma_{j+1}), (\gamma_{j+3} - \gamma_{j+2})$ and $(\gamma_{j+4} - \gamma_{j+3})$, respectively. Thus, the polynomial expression of $\frac{2g'}{K^3}$ on $[\xi_j, \xi_{j+1}]$ is

$$
\frac{2}{K^2}\big((\gamma_{j+2} - \gamma_{j+1}) B_{j+1}(t;2) + (\gamma_{j+3} - \gamma_{j+2}) B_{j+2}(t;2) + (\gamma_{j+4} - \gamma_{j+3}) B_{j+3}(t;2)\big) =
$$

$$t^2\big((\gamma_{j+2}-\gamma_{j+1})-2(\gamma_{j+3}-\gamma_{j+2})+(\gamma_{j+4}-\gamma_{j+3})\big)+$$
$$t\big(-2(\gamma_{j+2}-\gamma_{j+1})\xi_{j+1}+(\gamma_{j+3}-\gamma_{j+2})(\xi_{j-1}+\xi_j+\xi_{j+1}+\xi_{j+2})-2(\gamma_{j+4}-\gamma_{j+3})\xi_j\big)+$$
$$(\gamma_{j+2}-\gamma_{j+1})\xi_{j+1}^2+(\gamma_{j+3}-\gamma_{j+2})(-\xi_{j-1}\xi_{j+1}-\xi_j\xi_{j+2})+(\gamma_{j+4}-\gamma_{j+3})\xi_j^2$$
$$=: q_{j2}t^2+q_{j1}t+q_{j0}.$$

To apply Proposition 3.1 we need to rewrite the previous equation as a composition with the function $[0,1]\to[\xi_j,\xi_{j+1}] : z\mapsto \xi_j+z/K$. Therefore we consider the function

$$z^2\frac{q_{j2}}{K^2}+z(\frac{2\xi_j q_{j2}+q_{j1}}{K})+\xi_j^2 q_{j2}+\xi_j q_{j1}+q_{j0} =: p_{j2}z^2+p_{j1}z+p_{j0}. \tag{3.28}$$

Moreover, using that $\xi_j-\xi_{j-1}=\frac{1}{K}$ for all $j$, we obtain

$$\begin{aligned}
p_{j2} &= \frac{(\gamma_{j+2}-\gamma_{j+1})-2(\gamma_{j+3}-\gamma_{j+2})+(\gamma_{j+4}-\gamma_{j+3})}{K^2}\\
p_{j1} &= \frac{-2(\gamma_{j+2}-\gamma_{j+1})+2(\gamma_{j+3}-\gamma_{j+2})}{K^2}\\
p_{j0} &= \frac{(\gamma_{j+2}-\gamma_{j+1})+(\gamma_{j+3}-\gamma_{j+2})}{K^2}.
\end{aligned}$$

By Proposition 3.1, the function $g'$ is positive on $[\xi_j,\xi_{j+1}]$ if and only if there exists a positive $h_j$ such that

$$p_{j0}+p_{j2}+h_j \geq \|(p_{j0}-p_{j2}-h_j, p_{j1}-h_j)^\top\|_2. \tag{3.29}$$

Inequality (3.29) is the constraint we need when we restrict to the interval $[\xi_j,\xi_{j+1}]$. We need $K$ such constraints which are given by (3.2). This completes the proof. ■

### 3.10.2 Proof of Theorem 3.1

*Proof.* By the triangle inequality

$$\|\hat\beta_p^{(v)}-\beta_p^{(v)}\|_\infty \leq \|\hat\beta_p^{(v)}-\tilde\beta_p^{(v)}\|_\infty+\|\tilde\beta_p^{(v)}-\beta_p^{*(v)}\|_\infty+\|\beta_p^{*(v)}-\beta_p^{(v)}\|_\infty,$$

where $\tilde\beta_p(t)=\mathrm{E}(\hat\beta_p(t)|\mathcal{X})$. We give the proof by bounding each part of the right side. We subsequently deal with the third, second and first term of this upper bound.

- By Corollary 6.21 and (2.120) of Theorem 2.59 in Schumaker (2007), there exist a spline function $\beta_p^*$ of degree $q_p$ with equidistant knots $\xi_{p0}=0,\xi_{p1},\dots,\xi_{pK_p}=1$, such that

$$\|\beta_p^{(v)}-\beta_p^{*(v)}\|_\infty \leq C_1 K_p^{v-q_p-1}\|\beta_p^{(q_p+1)}\|_\infty \tag{3.30}$$

for $v = 0, \ldots, q_p$, where $C_1$ only depends on $q_p$. Hence $\|\beta_p^{(v)} - \beta_p^{*(v)}\|_\infty = O(K_p^{v-q_p-1})$ for $v = 0, \ldots, q_p$.

- Let $\tilde{\boldsymbol{\alpha}}_p = \mathrm{E}(\hat{\boldsymbol{\alpha}}_p|\mathcal{X})$. The derivative formula for B-splines gives

$$\hat{\beta}_p^{(v)}(t) - \tilde{\beta}_p^{(v)}(t) = K_p^v \mathbf{b}(t; q_p - v)^\top \mathbf{D}_v(\hat{\boldsymbol{\alpha}}_p - \tilde{\boldsymbol{\alpha}}_p),$$

where $\mathbf{D}_v$ denotes the matrix which takes the $v$-th order differences of a vector. Now,

$$\hat{\boldsymbol{\alpha}}_p - \tilde{\boldsymbol{\alpha}}_p = (\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \mathbf{U}^\top \mathbf{W}(\mathbf{Y} - \tilde{\mathbf{Y}}) = (\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \mathbf{U}^\top \mathbf{W} \boldsymbol{\varepsilon},$$

it is shown in Lemma A.4 of Huang et al. (2004), that

$$\|(\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1} \mathbf{U}^\top \mathbf{W} \boldsymbol{\varepsilon}\|_2^2 = O_P(r_n^2),$$

which yields

$$\|\hat{\beta}_p^{(v)} - \tilde{\beta}_p^{(v)}\|_\infty = O_P(K_p^v r_n), \tag{3.31}$$

since $\|\mathbf{b}(t; q_p - v)^\top \mathbf{D}_v\|_\infty$ is bounded by the properties of B-splines and the fact that $\mathbf{D}_v$ is a band matrix with bandwidth $v$.

- Let $\boldsymbol{\alpha}_p^*$ denote the B-spline coefficients of $\beta_p^*$, then using the derivative formula for B-splines we get

$$\begin{aligned} |\tilde{\beta}_p^{(v)}(t) - \beta_p^{*(v)}(t)| &= \|K_p^v \mathbf{b}(t; q_p - v)^\top \mathbf{D}_v(\tilde{\boldsymbol{\alpha}}_p - \boldsymbol{\alpha}_p^*)\|_\infty \\ &= O(K_p^v \|\tilde{\boldsymbol{\alpha}}_p - \boldsymbol{\alpha}_p^*\|_\infty). \end{aligned}$$

By Lemma A.11 of Huang et al. (2004) $\|\tilde{\boldsymbol{\alpha}}_p - \boldsymbol{\alpha}_p^*\|_\infty = O_P(\rho_n)$, therefore

$$\|\tilde{\beta}_p^{(v)} - \beta_p^{*(v)}\|_\infty = O_P(K_n^v \rho_n). \tag{3.32}$$

Equations (3.30) to (3.32) and Assumption 2.1.5 establish

$$\|\hat{\beta}_p^{(v)} - \beta_p^{(v)}\|_\infty = O_P(K_n^v \rho_n + K_n^{v-q_p-1} + K_n^v r_n).$$

The proof is complete. ■

### 3.10.3 Proof of Theorem 3.2

*Proof.* The outline of the proof is similar to the proof given for the univariate case in Wang and Meyer (2011). It is essentially based on the uniform consistency of the B-spline estimator and the fact

$$|\min_{t\in\mathcal{U}} f(t) - \min_{t\in\mathcal{U}} g(t)| \leq \max_{t\in\mathcal{U}} |f(t) - g(t)|, \tag{3.33}$$

for functions $f$ and $g$ of which $\mathcal{U}$ is a subset of the domain. Let $\Xi_k = \{\xi_{k0}, \xi_{k1}, \ldots, \xi_{kK_k}\}$. For $n$ sufficiently large $\min_{t\in\Xi_k} \beta_k'(t) > 0$ and

$$\begin{aligned}
P(s_{\min} < \min(0, \hat{Q}_\alpha)) &\leq P(\min_{t\in\Xi_k} \hat{\beta}_k'(t) < 0) \\
&= P(\min_{t\in\Xi_k} \hat{\beta}_k'(t) - \min_{t\in\Xi_k} \beta_k'(t) < -\min_{t\in\Xi_k} \beta_k'(t)) \\
&\leq P(|\min_{t\in\Xi_k} \hat{\beta}_k'(t) - \min_{t\in\Xi_k} \beta_k'(t)| > \min_{t\in\Xi_k} \beta_k'(t))) \\
&\leq P(\max_{t\in\Xi_k} |\hat{\beta}_k'(t) - \beta_k'(t)| > \min_{t\in\Xi_k} \beta_k'(t)) \\
&\leq P(\|\hat{\beta}_k' - \beta_k'\|_\infty > \min_{t\in\Xi_k} \beta_k'(t)).
\end{aligned}$$

The uniform consistency of $\hat{\beta}_k$ yields

$$\lim_{n\to\infty} P(\|\hat{\beta}_k' - \beta_k'\|_\infty > \min_{t\in\Xi_k} \beta_k'(t))) = 0,$$

since $\lim_{n\to\infty} \min_{t\in\Xi_k} \beta_k'(t) = \delta > 0$. This completes the proof. ∎

### 3.10.4 Proof of Theorem 3.3

*Proof.* By Lemma A.8 in Huang et al. (2004) we have

$$\frac{K_k \mathbf{b}(c; q_k - 1)^\top \mathbf{D}\hat{\boldsymbol{\alpha}} - \tilde{\beta}_k'(c)}{\sqrt{var}} \xrightarrow{d} N(0,1),$$

where $var = \mathrm{Var}(\mathbf{b}(c; q_k - 1)^\top (K_k \mathbf{D}\hat{\boldsymbol{\alpha}} - \tilde{\beta}_k'(c)))$. We argue that this asymptotic normality still holds when $\tilde{\beta}_k'(c)$ is replaced by $\beta_k'(c)$ by showing that

$$var^{-1/2} |\tilde{\beta}_k'(c) - \beta_k'(c)| = o_P(1).$$

Therefore we determine a lower bound on $var$. Recall the expression for $var$ (3.18).

Inequality (A.5) in Huang et al. (2004) establishes

$$\boldsymbol{\lambda}^\top \mathbf{U}^\top \mathbf{W}\mathbf{V}\mathbf{W}\mathbf{U}\boldsymbol{\lambda} \gtrsim \frac{\|\boldsymbol{\lambda}\|_2^2 n}{K_n \max_i N_i} \tag{3.34}$$

for any vector $\boldsymbol{\lambda} \in I\!R^{dim\times 1}$, where for sequences $a_n$ and $b_n$ we write $a_n \gtrsim b_n$ or $b_n \lesssim a_n$ when the sequence $b_n/a_n$ is bounded. When both $a_n \gtrsim b_n$ and $b_n \gtrsim a_n$ we write $a_n \asymp b_n$. Now we set $\boldsymbol{\lambda} = K_k(\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}\mathbf{D}^\top\mathbf{b}(c; q_k-1)$. By Lemma A.3 in Huang et al. (2004) $\|(\mathbf{U}^\top\mathbf{W}\mathbf{U})^{-1}\|_2 \asymp \frac{K_n}{n}$, therefore

$$\|\boldsymbol{\lambda}\|_2^2 \gtrsim K_k^2 \frac{K_k^2}{n^2}\|\mathbf{D}^\top\mathbf{b}(c; q_k)\|_2^2.$$

Note that

$$\|\mathbf{D}^\top\mathbf{b}(c; q_k-1)\|_2^2 = \sum_{j=0}^{m_k-1}(B_j(c; q_k-1) - B_{j+1}(c; q_k-1))^2, \tag{3.35}$$

with the convention $B_0(c; q_k-1) = B_{m_k}(c; q_k-1) = 0$. To determine a lower bound of (3.35) we can assume without loss of generality that $c \in [\xi_0, \xi_1] = [0, \frac{1}{K_k}]$ due to the properties of B-splines with equidistant knots. Then

$$\|\mathbf{D}^\top\mathbf{b}(c; q_k-1)\|_2^2 \geq (B_1(c; q_k-1) - B_2(c; q_k-1))^2 + (B_3(c; q_k-1) - B_2(c; q_k-1))^2 \tag{3.36}$$

Using the explicit formulas of B-splines given in Section 3.10.1, we obtain that the function

$$(B_1(c; q_k-1) - B_2(c; q_k-1))^2 + (B_3(c; q_k-1) - B_2(c; q_k-1))^2$$

on $[0, \frac{1}{K_k}]$ has a positive minimum at $c = \frac{1}{2K_k}$. This minimum does not depend on the number of knots. Therefore

$$\boldsymbol{\lambda}^\top \mathbf{U}^\top \mathbf{W}\mathbf{V}\mathbf{W}\mathbf{U}\boldsymbol{\lambda} \gtrsim \frac{\|\boldsymbol{\lambda}\|_2^2 n}{K_n \max_i N_i} \gtrsim \frac{\frac{K_n^4}{n^2} n}{K_n \max_i N_i} = \frac{K_n^3}{n \max_i N_i}. \tag{3.37}$$

By (3.30), Assumption 2.1.5 and the fact $q_k = 3$ it follows that $\rho_n = O(K_n^{-4})$ and $\|\beta_k^{*'} - \beta_k'\|_\infty = O(K_p^{-3})$. Then, by (3.32) and the triangle inequality $\|\tilde{\beta}_k' - \beta_k'\|_\infty \leq \|\tilde{\beta}_k' - \beta_k^{*'}\|_\infty + \|\beta_k^{*'} - \beta_k'\|_\infty = O(K_p^{-3})$, therefore

$$var^{-1/2}|\tilde{\beta}'(c) - \beta'(c)| \lesssim \frac{\sqrt{n \max_i N_i}}{\sqrt{K_n^9}}.$$

■

### 3.10.5 Proof of Theorem 3.4

*Proof.* We use similar steps as in the proof of Proposition 2 of Wang and Meyer (2011). Denote $\delta_n = \min_{t\in G_n} \beta_k'(t)$, then $\lim_{n\to\infty} \delta_n = \delta$. Let $c_n \in G_n$ such that $\min_{t\in G_n} \hat{\beta}_k'(t) = \hat{\beta}_k'(c_n)$.

1. For $n$ sufficiently large we have $\delta_n < 0$. Moreover, by (3.33)

$$\begin{aligned}
P(s < \min(0, \hat{Q}_\alpha)) &= P(\min_{t\in G_n} \hat{\beta}'_k(t) < \min(0, \hat{Q}'_\alpha)) \\
&= P(\min_{t\in G_n} \hat{\beta}'_k(t) - \min_{t\in G_n} \beta'_k(t) < -\min_{t\in G_n} \beta'_k(t) + \min(0, \hat{Q}'_\alpha)) \\
&\geq P(\max_{t\in G_n} |\hat{\beta}'_k(t) - \beta'_k(t)| < |\delta_n| + \min(0, \hat{Q}'_\alpha)) \\
&= P(\max_{t\in G_n} |\hat{\beta}'_k(t) - \beta'_k(t)| < |\delta_n| + \min(0, \inf\{r \; |P_{\hat{\beta}^{cs\prime}_k(c_n), \hat{v}_k}(r) \geq \alpha\}) \\
&\geq P(\max_{t\in G_n} |\hat{\beta}'_k(t) - \beta'_k(t)| < |\delta_n| - |\inf\{r \; |P_{0,\hat{v}_k}(r) \geq \alpha\}|) \\
&= P(\max_{t\in G_n} |\hat{\beta}'_k(t) - \beta'_k(t)| < |\delta_n| - q_\alpha \sqrt{\hat{v}_k}) \\
&\geq P(\|\hat{\beta}'_k - \beta'_k\|_\infty < |\delta_n| - q_\alpha \sqrt{\hat{v}_k}),
\end{aligned}$$

where $q_\alpha$ is the $\alpha$ quantile of the standard normal variable.

Before we proceed with the last inequality we need to bound $\hat{v}_k$. Recall

$$\begin{aligned}
\|\hat{v}_k\|_2 &= \|K_k^2 \hat{\sigma}^2 \mathbf{b}(c)^\top \mathbf{D}(\mathbf{U}^\top \mathbf{W}\mathbf{U})^{-1} \mathbf{U}^\top \mathbf{W}^2 \mathbf{U}(\mathbf{U}^\top \mathbf{W}\mathbf{U})^{-1} \mathbf{D}^\top \mathbf{b}(c)\|_2 \\
&\leq K_k^2 \hat{\sigma}^2 \|\mathbf{b}(c)^\top \mathbf{D}\|_2^2 \|(\mathbf{U}^\top \mathbf{W}\mathbf{U})^{-1}\|_2^2 \|\mathbf{U}^\top \mathbf{W}^2 \mathbf{U}\|_2.
\end{aligned}$$

Using Markov's inequality, we see that the estimate $\hat{\sigma}^2$ is bounded in probability if we prove that $\mathrm{E}(\hat{\sigma}^2 | \; \mathcal{X})$ is bounded. Hence, we start with

$$\begin{aligned}
\mathrm{E}(\hat{\sigma}^2 | \; \mathcal{X}) &= \mathrm{E}\left(\frac{1}{N - dim} (\mathbf{Y} - \mathbf{U}\hat{\boldsymbol{\alpha}})^\top (\mathbf{Y} - \mathbf{U}\hat{\boldsymbol{\alpha}}) | \; \mathcal{X}\right) \\
&= \frac{1}{N - dim} \mathrm{E}\left(\sum_{i=1}^{n} \sum_{j=1}^{N_i} \left(Y_{ij} - \sum_{p=0}^{d} X_{ij}^{(p)} \hat{\beta}_p(t_{ij})\right)^2 | \; \mathcal{X}\right) \\
&= \frac{1}{N - dim} \mathrm{E}\left(\sum_{i=1}^{n} \sum_{j=1}^{N_i} \left(\varepsilon_{ij}^2 + 2\varepsilon_{ij} \sum_{p=0}^{d} X_{ij}^{(p)} (\beta_p(t_{ij}) - \hat{\beta}_p(t_{ij}))\right.\right. \\
&\left.\left. + \sum_{p=0}^{d} X_{ij}^{(p)2} (\beta_p(t_{ij}) - \hat{\beta}_p(t_{ij}))^2\right) | \; \mathcal{X}\right).
\end{aligned} \tag{3.38}$$

Due to the following facts

- Assumption 2.1.4, which also means that $\mathrm{E}(|\varepsilon_{ij}| \; |\mathcal{X})$ is bounded by a uniform constant for all $i, j$;
- In the proof of Theorem 3.1 we showed $\|\beta_p - \hat{\beta}_p\|_\infty = O_P(\rho_n + r_n)$. Therefore, by the condition in the statement of this theorem $\|\beta_p - \hat{\beta}_p\|_\infty = o_P(1)$, moreover $\mathrm{E}(\|\beta_p - \hat{\beta}_p\|_\infty \; |\mathcal{X}) = o(1)$;

- Assumption 2.1.3;

expression (3.38) is bounded by $O(\frac{N}{N-dim})$, where $dim = \sum_{p=0}^{d} m_p$. Clearly $dim/N = o(1)$, which establishes $\mathrm{E}(\hat{\sigma}^2 | \mathcal{X}) = O(1)$.

Then, due to the properties of B-splines $\|\mathbf{b}(c)^\top \mathbf{D}\|_2^2$ is bounded, and by Lemma A.3 in Huang et al. (2004) we have that $\|(\mathbf{U}^\top \mathbf{W} \mathbf{U})^{-1}\|_2 \asymp \frac{K_n}{n}$ and $\|(\mathbf{U}^\top \mathbf{W} \mathbf{U})\|_2 \lesssim \frac{n}{K_n}$ from which we obtain, by first noting that $\|(\mathbf{U}^\top \mathbf{W}^2 \mathbf{U})\|_2 \leq \|(\mathbf{U}^\top \mathbf{W} \mathbf{U})\|_2$,

$$\|\hat{v}_k\|_2 = O_P\left(K_n^3/n\right).$$

By assumption $K_n^3/n \to 0$ and the fact that $\delta_n \to \delta$, we obtain by Theorem 3.1 that

$$\lim_{n\to\infty} P(\|\hat{\beta}'_k - \beta'_k\|_\infty < |\delta_n| - q_\alpha \sqrt{\hat{v}_k}) = 1$$

which completes the proof of the first part.

2. For the second part, let $n$ be sufficiently large so that $\delta_n > 0$, then

$$\begin{aligned}
P(s < \min(0, \hat{Q}_\alpha)) &= P(\min_{t\in G_n} \hat{\beta}'_k(t) < 0) \\
&= P(\min_{t\in G_n} \hat{\beta}'_k(t) - \min_{t\in G_n} \beta'_k(t) < -\min_{t\in G_n} \beta'_k(t)) \\
&= P(\min_{t\in G_n} \hat{\beta}'_k(t) - \min_{t\in G_n} \beta'_k(t) < -\delta_n) \\
&\leq P(|\min_{t\in G_n} \hat{\beta}'_k(t) - \min_{t\in G_n} \beta'_k(t)| > \delta_n) \\
&\leq P(\max_{t\in G_n} |\hat{\beta}'_k(t) - \beta'_k(t)| > \delta_n) \\
&\leq P(\|\hat{\beta}'_k - \beta'_k\|_\infty > \delta_n).
\end{aligned}$$

By the uniform consistency of $\hat{\beta}'_k$ and since $\lim_n \delta_n = \delta$ we have

$$\lim_{n\to\infty} P(\|\hat{\beta}'_k - \beta'_k\|_\infty > \delta_n) = 0.$$

This completes the second part of the proof.

■

# Chapter 4

# Estimating multivariate parameters in PDE models

This chapter is based on Ahkim et al. (2017) (manuscript).

## 4.1 Introduction

Many scientists study dynamic systems abiding by PDEs which are governed by certain parameters. Often these parameters are unknown and the interest is to shed light on them based on error prone measurements of the state variables. The PDE parameters are viewed as a source of information about the dynamic process beyond what is revealed by the state variables, hence the interest in their value. There is a considerable amount of literature on estimating constant PDE parameters. In contrast is the situation for PDEs with multivariate parameters. Applications of PDE models to real life problems crucially depend on methods to acquire precise measurements. In the last three decades many successful advances have been made. Hence the importance to provide a general framework for effectively estimating multivariate parameters in PDE models and giving theoretical foundations. That is the purpose of this chapter.

Our model assumes a multivariate state variable $g(\mathbf{t})$ setting $(\mathbf{t} = (t_1, \ldots, t_l))$, where we only observe states $Y_i$

$$Y_i = g(\mathbf{t}_i) + \varepsilon_i, \quad i = 1, \ldots, n \tag{4.1}$$

for $\mathbf{t}_i = (t_{i1}, \ldots, t_{il})$ $(i = 1, \ldots, n)$ and with $\varepsilon_i|\mathbf{t}_i$ i.i.d. mean zero measurement errors. The multivariate state variable $g$ is modeled by a known PDE

$$F\left(\mathbf{t}, g, \frac{\partial g}{\partial t_1}, \ldots, \frac{\partial g}{\partial t_l}, \frac{\partial^2 g}{\partial t_1 \partial t_1}, \ldots, \frac{\partial^2 g}{\partial t_1 \partial t_l}, \ldots, \frac{\partial^2 g}{\partial t_l \partial t_l}; \boldsymbol{\theta}\right) = 0, \tag{4.2}$$

where the multivariate parameter $\boldsymbol{\theta}$. Equation (4.2) might suggest that the PDE must be of second order, however a PDE of any order can be considered (by allowing higher order derivatives). Also, for simplicity we write $F(g(\mathbf{t}); \boldsymbol{\theta})$ to refer to (4.2).

There are three main approaches to estimate the constant parameters of a general PDE. The first approach is a two-stage method where unknown PDEs are modeled by multivariate polynomials. Based on this approximation the PDE parameters are obtained by using a least squares approximation (Müller and Timmer (2004) and Bär et al. (1999)). In the second approach, the PDE is first solved with a numerical method and then the parameters are estimated by solving another least squares problem (Müller and Timmer (2002)). The third approach is a penalized smoothing method introduced by Ramsay et al. (2007) in the ODE context. This approach can be seen as an extension of the penalized spline method for estimating mean functions of Eilers and Marx (1996). In the context of PDEs, this approach is followed by Xun et al. (2013) and Frasso et al. (2015) who estimate constant parameters of linear PDE models. In this chapter we extend this approach to PDE models with multivariate parameters and give consisteny results in case of linear PDEs. Note that in general a linear PDE model does not imply closed form expressions of the estimators. However, numerical optimization techniques are to be used in general. For a further introduction and considerations of PDE modeling with splines we refer to Frasso et al. (2015).

To the best of our knowledge there is no theoretical literature on estimating multivariate parameters of a PDE. Our goal is to estimate (consistently) multivariate parameters based on the flexible third approach. Xun et al. (2013) showed that this approach performs better than the two-stage method (see Figure 1 in the same article) and it has the advantage of not needing to find numerical solutions of a PDE. Moreover, they showed the asymptotic normality of the parameter estimator.

The relevance of this work is highlighted by many recent PDE applications in the scientific literature. For illustration purposes we consider two biological applications. Hartung et al. (2014) model tumor growth and metastatic spreading which is described by a transport equation with varying parameters. Metastatic spreading is the process whereby one initial tumor grows and starts to spread cancer cells causing new growing tumors. The PDE which describes this process is

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial}{\partial x}\left(g_m(x)\rho(x,t)\right) = 0,$$

where $\rho(x,t)$ is the number of size $x$ tumors per unit length at time $t$, $g_m(x)$ is the growth rate. Hartung et al. (2014) model the $g_m(x)$ by parametric functions which are formulated based on experience. Nonparametric estimation of the parameters relieves the scientist from formulating parametric models. The second example is the modeling of brain glioma growth using a modified reaction-diffusion equation by Jianjun et al. (2013). In our simulation example we discuss the heating and cooling of a rod, which is described by the reaction-diffusion equation.

We approximate $g(\mathbf{t})$ and $\boldsymbol{\theta}(\mathbf{t})$ by tensor product spline functions and give asymptotic results when the number of observations $n$ tends to infinity. Unlike Xun et al. (2013), we do not assume the tensor product spline approximation to be exact, however, $g(\mathbf{t})$ as well as $\boldsymbol{\theta}(\mathbf{t})$ are assumed to be smooth functions. It should also be noted that Xun et al. (2013) do not consider incorporating boundary conditions. From our experience we know that omitting boundary conditions leads to a very poor estimation of $g(\mathbf{t})$ on the boundaries, hence the unacceptable bias of the parameter estimation at the boundaries. We show that due to the nice B-spline properties, the boundary conditions are easily translated into linear constraints on the tensor product B-spline coefficients (as it is mentioned in Frasso et al. (2015)). We also establish the uniform consistency of our parameter estimator $\hat{\boldsymbol{\theta}}(\mathbf{t})$ under certain assumptions.

The overview of this chapter is as follows. In Section 4.2 we will describe the B-spline modeling of the multivariate functions. In Section 4.3 two methods for estimating the parameters are presented. In the case of a linear PDE model, we show explicit expressions for the estimators in Section 4.4, where we also give asymptotic results. The proofs can be found in Section 4.7. In Section 4.5 we apply both estimation methods to the 1D heat equation, where we wish to describe the source term which causes the heating and/or cooling of the bar. The results of the simulation study are discussed in Section 4.5.3. We summarize our findings in Section 4.6.

## 4.2 Modelling multivariate PDE parameters

Using the same notation as before we consider the PDE model

$$F\left(g(\mathbf{t});\boldsymbol{\theta}(\mathbf{t})\right)=0 \tag{4.3}$$

with given boundary conditions on $g$. The methods in this chapter still work if we assume no given boundary conditions, and in such a case if it is reasonable to assume constant boundary values for $g$, we include the constraints which ensure that the derivatives along the boundaries are zero. We observe $(\mathbf{t}_i, Y_i)$ for $i=1,\ldots,n$, and assume that the domain of $g$ and $\boldsymbol{\theta}$ is a compact rectangle $\mathcal{H}\subset I\!R^l$.

The multivariate state variable $g$ is estimated by a tensor product spline function $s(\mathbf{t};\boldsymbol{\alpha}) = \mathbf{B}_g^\top(\mathbf{t})\boldsymbol{\alpha}$ (see Chapter 11 of Schumaker (2007)), where $\mathbf{B}_g(\mathbf{t})$ is a vector of basis functions of dimension $m_g$ evaluated in $\mathbf{t}$. Thus

$$\mathbf{B}_g(\mathbf{t})=\left(\mathbf{B}_{g1}(t_1)^\top\otimes\mathbf{B}_{g2}(t_2)^\top\otimes\ldots\otimes\mathbf{B}_{gl}(t_l)^\top\right)^\top,$$

where $\mathbf{B}_{gj}(t_j)\in I\!R^{m_{gj}\times 1}$ denotes the column vector of B-splines for the $t_j$ direction evaluated in $t_j$, and $m_g=\prod_{j=1}^l m_{gj}$. For the sake of presentation we assume without loss of generality that $\mathcal{H}$ is a unit rectangle, i.e. $\mathcal{H}=[0,1]^l$. Also, we omit the index $g$ in $\mathbf{B}_g(\mathbf{t})$ if it is clear from the context. The design points $\mathbf{t}_i$ are deterministic and we assume there exists a cumulative distribution function $G(\mathbf{t})$, with a positive and continuous density function on $[0,1]^l$ such that

$$\sup_{\mathbf{t}\in[0,1]^l}|G_n(\mathbf{t})-G(\mathbf{t})|=o\left(m_g^{-1}\right), \tag{4.4}$$

where

$$G_n(\mathbf{t})=\frac{1}{n}\sum_{i=1}^n 1_{\prod_{j=1}^l[0,t_{ij}]}(\mathbf{t})$$

is the empirical distribution function of $\mathbf{t}_1,\ldots,\mathbf{t}_n$ with $1_A$ the indicator function of the set $A$. We also assume that the knots $0=\xi_{j0}<\xi_{j1}<\ldots<\xi_{jK_{gj}}=1$ corresponding to $\mathbf{B}_{gj}, j=1,\ldots,l$, are quasi-uniform. This means that at each $j=1,\ldots,l$, with consecutive knot distances $\delta_{jk}=\xi_{jk}-\xi_{j(k-1)}$ we have a positive constant $C$ such that

$$\frac{\max_{1\le k\le K_{gj}}\delta_{jk}}{\min_{1\le k\le K_{gj}}\delta_{jk}}<C. \tag{4.5}$$

Assumptions (4.4) and (4.5) are also used in Yoo and Ghosal (2016) among others.

The PDE parameter $\boldsymbol{\theta}(\mathbf{t}) = (\theta_1(\mathbf{t}), \ldots, \theta_d(\mathbf{t}))$ is also modeled by tensor product spline functions. Let $\boldsymbol{\gamma}_p \in I\!R^{m_{\theta_p} \times 1}$ denote the vector of tensor product spline coefficients corresponding to the modeling of $\theta_p(\mathbf{t}) \approx \mathbf{B}_{\theta_p}^\top(\mathbf{t})\boldsymbol{\gamma}_p$, for $p = 1, \ldots, d$, where

$$\mathbf{B}_{\theta_p}(\mathbf{t}) = \left(\mathbf{B}_{\theta_p 1}(t_1)^\top \otimes \mathbf{B}_{\theta_p 2}(t_2)^\top \otimes \ldots \otimes \mathbf{B}_{\theta_p l}(t_l)^\top\right)^\top,$$

with $\mathbf{B}_{\theta_p j}(t_j) \in I\!R^{m_{\theta_p j} \times 1}$ denoting the column vector of B-splines for the $t_j$ direction evaluated in $t_j$, and let $m_{\theta_p} = \prod_{j=1}^{l} m_{\theta_p j}$. Let $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \ldots, \boldsymbol{\gamma}_d^\top)^\top \in I\!R^{m_{\boldsymbol{\theta}} \times 1}$ where $m_{\boldsymbol{\theta}} = \sum_{p=1}^{d} m_{\theta_p}$ and let the notation $F(\boldsymbol{\alpha}; \boldsymbol{\gamma}) = 0$ refer to the spline modeling of (4.3). In theory, we let $g$ and $\theta_p$ for $p = 1, \ldots, d$, be approximated by tensor product splines which have a fixed degree and an increasing number of quasi-uniform knots such that the approximation errors tend to zero as the sample size $n \to +\infty$.

## 4.3 Estimating multivariate PDE parameters

In this section, we present two different methodologies for estimating the multivariate parameters $\boldsymbol{\theta}(\mathbf{t})$. The first one extends the method used by Xun et al. (2013) and consists of two sequential minimization problems, which is why we refer to it as the "two-step method". The second method considers one big minimization problem and we'll refer to it as the "one-step method".

### 4.3.1 The two-step method

The two-step method obtains an estimator for $\boldsymbol{\alpha}$ in function of $\boldsymbol{\gamma}$. An estimator for $\boldsymbol{\gamma}$ is then obtained by minimizing a penalized least squares measure of fit for the given data $\mathbf{Y} = (Y_1, \ldots, Y_n)^\top$. Finally, using $\hat{\boldsymbol{\gamma}}$ we obtain the estimate $\hat{\boldsymbol{\alpha}}$.

#### 4.3.1.1 Estimation of $\alpha$ for a fixed $\gamma$

To estimate $\boldsymbol{\alpha}$, we first treat $\boldsymbol{\gamma}$ as fixed and minimize a penalized least squares criterion:

$$J(\boldsymbol{\alpha} \mid \boldsymbol{\gamma}) = \sum_{i=1}^{n} (Y_i - \mathbf{B}_g(\mathbf{t}_i)\boldsymbol{\alpha})^2 + \lambda \int F(\boldsymbol{\alpha}; \boldsymbol{\gamma})^2 \, d\mathbf{t} + \mu \left( \sum_{j=1}^{\mathcal{C}} \left\| \mathbf{V}_j \boldsymbol{\alpha} \right\|_2^2 \right). \tag{4.6}$$

The first term models the fidelity to the measured data and the second term to the PDE model (4.3). The final term enforces the boundary (and initial) conditions. In Section 4.5 we explain how such conditions can always be translated into simple linear constraints on the tensor product spline coefficients $\boldsymbol{\alpha}$ which we write as

$$\left\|\mathbf{V}_1\boldsymbol{\alpha}\right\|_2^2 + \ldots + \left\|\mathbf{V}_{\mathcal{C}}\boldsymbol{\alpha}\right\|_2^2 = 0.$$

Here, $\mathcal{C}$ is the number of constraints, $\mathbf{V}_i$ $(i = 1, \ldots, \mathcal{C})$ are known matrices and $\|\cdot\|_2$ is the Frobenius norm (see Appendix A).

Hence, in first instance we minimize $J(\boldsymbol{\alpha}|\boldsymbol{\gamma})$ to obtain $\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})$. The integral in (4.6) can be approximated by numerical integration methods.

#### 4.3.1.2 Estimation of $\boldsymbol{\gamma}$

Once an estimator $\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})$ is obtained, we obtain an estimator $\hat{\boldsymbol{\gamma}}$ by minimizing the penalized least squares measure of fit

$$H(\boldsymbol{\gamma}) = \sum_{i=1}^{n} \left(Y_i - \mathbf{B}_g(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})\right)^2 + \sum_{j=1}^{l}\sum_{p=1}^{d} \lambda_{pj}\left\|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\right\|_2^2. \tag{4.7}$$

The penalty matrices $\mathbf{P}_{pj}$ encourage smooth estimates of $\hat{\boldsymbol{\theta}}(\mathbf{t})$ by penalizing the roughness of the coefficients. They will be discussed in more detail in Section 4.3.3. Using $\hat{\boldsymbol{\gamma}}$, we can now finally calculate the tensor product spline estimator $\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\gamma}})$.

Until now we have treated the regularization parameters $(\lambda, \mu, \{\lambda_{pj}|\ p = 1, \ldots, d,\ j = 1, \ldots, l\})$ as fixed. In Section 4.3.4 we propose a criterion to determine these regularization parameters.

### 4.3.2 The one-step method

Instead of solving two minimization problems, the one-step method combines (4.6) and (4.7) into one big minimization problem

$$\begin{aligned} K(\boldsymbol{\alpha}, \boldsymbol{\gamma}) = &\sum_{i=1}^{n} \left(Y_i - \mathbf{B}_g(\mathbf{t}_i)\boldsymbol{\alpha}\right)^2 + \lambda \int F(\boldsymbol{\alpha}; \boldsymbol{\gamma})^2 \, d\mathbf{t} + \mu \left(\sum_{j=1}^{\mathcal{C}} \left\|\mathbf{V}_j\boldsymbol{\alpha}\right\|_2^2\right) \\ &+ \sum_{j=1}^{l}\sum_{p=1}^{d} \lambda_{pj}\left\|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\right\|_2^2. \end{aligned} \tag{4.8}$$

Note that because

$$\frac{\partial K}{\partial \boldsymbol{\alpha}} = \frac{\partial J}{\partial \boldsymbol{\alpha}}$$

the estimator $\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})$ will be the same for both methods. However, because the same does not hold for the partial derivatives with respect to $\boldsymbol{\gamma}$, the methods will find different estimators $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$. This is illustrated in Section 4.8 and 4.9 where these calculations are done for the heat equation problem discussed in Section 4.5.

In Section 4.3.4 we propose a criterion to determine the regularization parameters $(\lambda, \mu, \{\lambda_{pj} | \, p = 1, \ldots, d, \; j = 1, \ldots, l\})$.

### 4.3.3 Ensuring smoothness of tensor product spline functions

In this section we define the penalization matrices used in (4.7) for $\boldsymbol{\gamma}_1$. For the sake of clarity we assume $l = 2$ and denote spline coefficients

$$\boldsymbol{\gamma}_1 = (\boldsymbol{\Gamma}_{11}, \ldots, \boldsymbol{\Gamma}_{1m_{\theta_1 2}}, \boldsymbol{\Gamma}_{21}, \ldots, \boldsymbol{\Gamma}_{2m_{\theta_1 2}}, \ldots, \boldsymbol{\Gamma}_{m_{\theta_1 1}1}, \ldots, \boldsymbol{\Gamma}_{m_{\theta_1 1}m_{\theta_1 2}})^\top.$$

Consider

$$\mathbf{B}_{\theta_1}(\mathbf{t})^\top \boldsymbol{\gamma}_1 = \sum_{i=1}^{m_{\theta_1 1}} \sum_{j=1}^{m_{\theta_1 2}} \boldsymbol{\Gamma}_{ij} B_{\theta_1 1,i}(t_1) B_{\theta_1 2,j}(t_2). \tag{4.9}$$

Then, $\boldsymbol{\gamma}_1$ is the subsequent concatenation of the rows of $\boldsymbol{\Gamma}$, where $\boldsymbol{\Gamma}$ is the matrix with $(i,j)$-th element $\boldsymbol{\Gamma}_{ij}$. Our way to achieve smoothness of (4.9), which is also described in Marx and Eilers (2005), is by penalizing both the rows and the columns of $\boldsymbol{\Gamma}$. Penalization of the columns (rows) ensures smoothing in the $t_1$ ($t_2$) direction. Suppose we penalize the first order differences of the rows and the second order differences of the columns of $\boldsymbol{\Gamma}$, this is equivalent with penalizing $\boldsymbol{\gamma}_1$ by the matrices

$$\mathbf{P}_{11} = \mathbf{D}_1 \otimes \mathbf{I}_{m_{\theta_1 2}}$$

and

$$\mathbf{P}_{12} = \mathbf{I}_{m_{\theta_1 1}} \otimes \mathbf{D}_2,$$

respectively, where the matrix $\mathbf{D}_1$ ($\mathbf{D}_2$) takes the first (second) order differences of the columns (transposed rows) of $\boldsymbol{\Gamma}$ and $\mathbf{I}_m$ denotes the identity matrix of size $m$. For example when $m_{\theta_1 1} = m_{\theta_1 2} = 3$

$$\mathbf{D}_1 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad \mathbf{D}_2 = \begin{pmatrix} -1 & 2 & 1 \end{pmatrix},$$

and hence the penalization matrices for the $t_1$- and $t_2$-direction are respectively given by

$$\mathbf{P}_{11} = \mathbf{D}_1 \otimes \mathbf{I}_3 = \begin{pmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix},$$

and

$$\mathbf{P}_{12} = \mathbf{I}_3 \otimes \mathbf{D}_2 = \begin{pmatrix} -1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & 1 \end{pmatrix}.$$

### 4.3.4 Determining the regularization parameters

The estimators $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\theta}}$ depend on the regularization parameters $\lambda, \mu, \lambda_{pj}, p = 1, \ldots, d, j = 1, \ldots, l$. The optimal regularization parameters need to fulfill three main goals: goodness-of-fit to the data and fidelity to both the PDE equation and its boundary conditions. Fidelity to the PDE equation is assessed by $F(\mathbf{t}_i; \hat{\boldsymbol{\alpha}}; \hat{\boldsymbol{\gamma}}), i = 1, \ldots, n$, which should have a small mean. Fidelity to the boundary conditions is measured by

$$\sum_{j=1}^{\mathcal{C}} \left\|\mathbf{V}_j \hat{\boldsymbol{\alpha}}\right\|_2^2 = \left\|(\sum_{j=1}^{\mathcal{C}} \mathbf{V}_j^\top \mathbf{V}_j)^{1/2} \hat{\boldsymbol{\alpha}}\right\|_2^2$$

and it should be as small as possible. Xun et al. (2013) proposed a criterion where a trade-off between goodness-of-fit and fidelity to the PDE model is made. However, one important consideration is overlooked, i.e. the fidelity to the PDE equation which will imply a relatively small variance for $F(\mathbf{t}_i; \hat{\boldsymbol{\alpha}}; \hat{\boldsymbol{\gamma}})$ $(i = 1, \ldots, n)$, especially when we have multivariate parameters. When this variance is included in the criterion, we found that the estimator $\hat{\boldsymbol{\theta}}$ approximates $\boldsymbol{\theta}$ smoothly and hence more precise. A similar consideration holds for the boundary conditions.

We therefore propose to choose the regularization parameters such that the following criterion is minimized:

$$\frac{\sum_{i=1}^n (Y_i - \mathbf{B}(\mathbf{t}_i)\hat{\boldsymbol{\alpha}})^2}{n\hat{\sigma}_\varepsilon^2} + \frac{\sum_{i=1}^n F(\mathbf{t}_i; \hat{\boldsymbol{\alpha}}; \hat{\boldsymbol{\gamma}})^2}{n\hat{\sigma}_F^2} + \frac{\|\sqrt{\sum_{j=1}^{\mathcal{C}} \mathbf{V}_j^\top \mathbf{V}_j} \hat{\boldsymbol{\alpha}}\|_2^2}{m_g \hat{\sigma}_b^2} + \hat{\sigma}_F^2 + \hat{\sigma}_b^2. \tag{4.10}$$

Here $\hat{\sigma}_\varepsilon^2$ is the sample variance of $(Y_i - \mathbf{B}(\mathbf{t}_i)\hat{\boldsymbol{\alpha}})$, $\hat{\sigma}_F^2$ is the sample variance of $F(\mathbf{t}_i; \hat{\boldsymbol{\alpha}}; \hat{\boldsymbol{\gamma}})$ $(i = 1, \ldots, n)$ and $\hat{\sigma}_b^2$ is the sample variance of $\sqrt{\sum_{j=1}^{\mathcal{C}} \mathbf{V}_j^\top \mathbf{V}_j \hat{\boldsymbol{\alpha}}}$. The averaging and rescaling in the first three terms of (4.10) serve to weight equally the three goals mentioned above. Note that we do not add $\hat{\sigma}_\varepsilon^2$ in (4.10) because $\hat{\sigma}_\varepsilon^2$ should be approximately $\sigma^2 = \mathrm{Var}(\varepsilon_i)$, the noise level.

In practice, a grid is set for each regularization parameter, and the regularization parameters that minimize criterion (4.10) over these grids are chosen.

## 4.4 Linear PDE models

For linear PDEs we derive more explicit formulas for the one- and the two-step method. The most general linear PDE can be written as

$$F(g; \boldsymbol{\theta}) = \sum_{i=1}^{r} h_i(\boldsymbol{\theta}) O_i(g) = 0, \tag{4.11}$$

where $O_i(\cdot)$ denotes an operator which either takes a partial derivative of $g$, takes $g$ to itself or to the constant function 1. Let $\mathcal{O}$ denote the order of the PDE (4.11). Note that partial derivatives of $s(\mathbf{t}; \boldsymbol{\alpha}) = \mathbf{B}_g^\top(\mathbf{t})\boldsymbol{\alpha}$ are still linear in $\boldsymbol{\alpha}$.

We assume that (without loss of generality) the coefficient of the operator which takes $g$ to the function 1 is $\theta_d$. Then we can write (4.11) in terms of the tensor product spline coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ as follows:

$$F(\boldsymbol{\alpha}; \boldsymbol{\gamma}) = \mathbf{f}^\top(\mathbf{t}; \boldsymbol{\gamma})\boldsymbol{\alpha} + \mathbf{Q}_d^\top(\mathbf{t})\boldsymbol{\gamma}.$$

Here $\mathbf{f}(\mathbf{t}; \boldsymbol{\gamma}) \in I\!R^{m_g \times 1}$ is a (column) vector and $\mathbf{Q}_d(\mathbf{t}) \in I\!R^{m_{\boldsymbol{\theta}} \times 1}$ is the (column) vector such that $\mathbf{Q}_d^\top(\mathbf{t})\boldsymbol{\gamma}$ is the tensor product spline modeling of $\theta_d(\mathbf{t})$.

### 4.4.1 Estimator for the two-step method

As it is described in Section 4.3.1, to estimate $(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ we first minimize

$$\begin{aligned} J(\boldsymbol{\alpha} \mid \boldsymbol{\gamma}) = {} & \left\|\mathbf{Y} - \mathcal{B}\boldsymbol{\alpha}\right\|_2^2 + \lambda\left(\boldsymbol{\alpha}^\top \mathbf{R}(\boldsymbol{\gamma})\boldsymbol{\alpha} + 2\boldsymbol{\gamma}^\top \mathbf{T}(\boldsymbol{\gamma})\boldsymbol{\alpha} + \boldsymbol{\gamma}^\top \mathbf{Z}(\boldsymbol{\gamma}_d)\boldsymbol{\gamma}\right) \\ & + \mu\left(\sum_{j=1}^{\mathcal{C}} \left\|\mathbf{V}_j\boldsymbol{\alpha}\right\|_2^2\right) \end{aligned} \tag{4.12}$$

with respect to $\boldsymbol{\alpha}$. Here $\mathbf{R}(\boldsymbol{\gamma}) = \int \mathbf{f}(\mathbf{t};\boldsymbol{\gamma})\mathbf{f}^\top(\mathbf{t};\boldsymbol{\gamma})\ d\mathbf{t}$, $\mathbf{T}(\boldsymbol{\gamma}) = \int \mathbf{Q}_d(\mathbf{t})\mathbf{f}^\top(\mathbf{t};\boldsymbol{\gamma})\ d\mathbf{t}$, $\mathbf{Z}(\boldsymbol{\gamma}_d) = \int \mathbf{Q}_d(\mathbf{t})\mathbf{Q}_d^\top(\mathbf{t})\ d\mathbf{t}$ and

$$\mathcal{B} = \begin{pmatrix} \mathbf{B}^\top(\mathbf{t}_1) \\ \mathbf{B}^\top(\mathbf{t}_2) \\ \vdots \\ \mathbf{B}^\top(\mathbf{t}_n) \end{pmatrix} \in I\!R^{n\times m_g}. \tag{4.13}$$

When the integrals are difficult to compute explicitly, we can approximate them numerically. See for example Burden and Faires (2005), who suggest to use a composite Simpson's rule as an adequate approximation of these integrals.

Continuing with the minimization problem, we find that

$$\begin{aligned}\frac{\partial J(\boldsymbol{\alpha} \mid \boldsymbol{\gamma})}{\partial \boldsymbol{\alpha}} =& -2\mathbf{Y}^\top\mathcal{B} + 2\boldsymbol{\alpha}^\top\mathcal{B}^\top\mathcal{B} + \lambda\left(2\boldsymbol{\alpha}^\top\mathbf{R}(\boldsymbol{\gamma}) + 2\boldsymbol{\gamma}^\top\mathbf{T}(\boldsymbol{\gamma})\right) \\ &+ 2\mu\boldsymbol{\alpha}^\top\left(\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right).\end{aligned} \tag{4.14}$$

The solution to $\frac{\partial J(\boldsymbol{\alpha}|\boldsymbol{\gamma})}{\partial\boldsymbol{\alpha}} = 0$ is therefore given by

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}) = \left(\mathcal{B}^\top\mathcal{B} + \lambda\mathbf{R}(\boldsymbol{\gamma}) + \mu\left(\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)\right)^{-1}(\mathcal{B}^\top\mathbf{Y} - \lambda\mathbf{T}(\boldsymbol{\gamma})^\top\boldsymbol{\gamma}). \tag{4.15}$$

Second, to obtain the estimator of $\boldsymbol{\gamma}$ we minimize

$$H(\boldsymbol{\gamma}) = \left\|\mathbf{Y} - \mathcal{B}\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})\right\|_2^2 + \sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\left\|\mathbf{P}_{pj}\gamma_p\right\|_2^2, \tag{4.16}$$

with respect to $\boldsymbol{\gamma}$. In Section 4.8 we compute the minimizer of $H(\boldsymbol{\gamma})$ for the Heat Example (Section 4.5).

### 4.4.2 Estimator for the one-step method

As it is described in Section 4.3.2, the one-step approach minimizes

$$\begin{aligned}K(\boldsymbol{\alpha},\boldsymbol{\gamma}) =& \left\|\mathbf{Y} - \mathcal{B}\boldsymbol{\alpha}\right\|_2^2 + \lambda\left(\boldsymbol{\alpha}^\top\mathbf{R}(\boldsymbol{\gamma})\boldsymbol{\alpha} + 2\boldsymbol{\gamma}^\top\mathbf{T}(\boldsymbol{\gamma})\boldsymbol{\alpha} + \boldsymbol{\gamma}^\top\mathbf{Z}(\boldsymbol{\gamma}_d)\boldsymbol{\gamma}\right) \\ &+ \mu\left(\sum_{j=1}^{\mathcal{C}}\left\|\mathbf{V}_j\boldsymbol{\alpha}\right\|_2^2\right) + \sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\left\|\mathbf{P}_{pj}\gamma_p\right\|_2^2.\end{aligned}$$

Thus the following system of equations needs to be solved

$$\begin{cases} \dfrac{\partial K}{\partial \boldsymbol{\alpha}} = \mathbf{0} \\ \dfrac{\partial K}{\partial \boldsymbol{\gamma}} = \mathbf{0}. \end{cases}$$

As mentioned before, the partial derivative of $K$ w.r.t $\boldsymbol{\alpha}$ is the same as the partial derivative of $J$ w.r.t. $\boldsymbol{\alpha}$ which was given by (4.14). See Section 4.9 where we solve this system of equations for the Heat Example (Section 4.5).

### 4.4.3 Practical considerations

The estimator $\hat{\boldsymbol{\alpha}}$ for both methods depends on the inverse of $(\mathcal{B}^\top \mathcal{B})$ with $\mathcal{B} \in I\!R^{n \times m_g}$. Hence for $(\mathcal{B}^\top \mathcal{B})$ to be invertible we must have $n \geq m_g$. This means that when the number of observations is relatively small, the number of knots in each direction must be chosen accordingly. We also need the quasi-uniformity of the knots and condition (4.4) which says that the measurements must be fairly distributed over the whole domain, see (4.23) onwards for the proof. It is also required that $n + d \cdot l \geq m_{\boldsymbol{\theta}}$, this becomes clear by looking at the dimension of the Jacobian matrix when one applies the Gauss-Newton method to minimize (4.7).

When the number of multivariate parameters is high, the computational cost to determine the $(d \cdot l + 2)$ regularization parameters becomes a burden. Assuming $\lambda_{p1}, \lambda_{p2}, \ldots, \lambda_{pl}$ to be equal for $p = 1, \ldots, d$, would reduce the number of regularization parameters to $d + 2$. However, our experience is that such an assumption delivers up a large amount of accuracy in the final multivariate parameter estimates, because in general multivariate parameters vary in their degree of smoothness in various directions. This issue is part of still ongoing research.

### 4.4.4 Asymptotic results (two-step method)

In this section we present the theoretical results for the two-step method with the full basis. Similar techniques can be used to obtain asymptotic results for the one-step method. Proposition 4.1 gives the rate of convergence of our estimator of $g$. This rate involves the approximation error term $\rho_n = \inf_{f \in \mathbb{G}} ||g - f||_\infty$ where $\mathbb{G} = \{\mathbf{B}_g(\mathbf{t})\boldsymbol{\alpha} \; | \boldsymbol{\alpha} \in I\!R^{m_g}\}$ is the space of tensor product spline functions which we employ.

Under certain smoothness conditions on $g$ we can say more about $\rho_n$. To this end, we first introduce the following notion. Set $0 < \kappa \le 1$, a function $h$ on $\mathcal{H}$ is said to satisfy a Hölder condition with exponent $\kappa$ if there exists a positive number $\kappa^*$ such that $|h(\mathbf{t}) - h(\mathbf{t}_0)| \le \kappa^* \|\mathbf{t} - \mathbf{t}_0\|_2^\kappa$ for $\mathbf{t}, \mathbf{t}_0 \in \mathcal{H}$. Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_l)$ denote an $l$-tuple of nonnegative integers, we define $[\boldsymbol{\nu}] = \nu_1 + \ldots + \nu_l$. For such $l$-tuples we let $D^\nu$ denote the differential operator defined by

$$D^{\boldsymbol{\nu}} = \frac{\partial^{[\boldsymbol{\nu}]}}{\partial t_1^{\nu_1} \ldots \partial t_l^{\nu_l}}.$$

We refer to $[\boldsymbol{\nu}]$ as the order of $D^{\boldsymbol{\nu}}$. Now, let $k$ be a nonnegative integer and set $p = k + \kappa$. A function $h$ on $\mathcal{H}$ is called $p$–smooth if it is $k$-times continuously differentiable on $\mathcal{H}$ and $D^{\boldsymbol{\nu}} h$ satisfies a Hölder condition with exponent $\kappa$ for all $\boldsymbol{\nu}$ with $[\boldsymbol{\nu}] = k$.

In the nonparametric estimation literature the $p$–smooth condition is often used (see Huang (1998), page 251). Suppose the B-spline degrees in all directions are equal to $q$ and $m_g^{1/l} = m_{gj}$ for all directions $j$, then, under the $p$–smoothness condition $\rho_n \lesssim m_g^{-p/l}$ if $q \ge p-1$ (see (13.69) and Theorem 12.8 of Schumaker (2007)).

Recall that $\mathcal{O}$ is the highest order of the partial derivatives in the linear PDE (4.11).

**Proposition 4.1.** *Assume that (4.4),(4.5) and* ***A.1*** *to* ***A.8*** *(defined in Section 4.7) hold. If* $\frac{m_g^2}{n}\left(\lambda m_g^{2\mathcal{O}/l} + \mu\right) = o(1)$ *as* $n \to \infty$, *then*

$$\frac{1}{n}\sum_{i=1}^{n}(g(\mathbf{t}_i) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}))^2 = O_P\left(\rho_n^2 + \frac{m_g}{n}(\lambda^2 m_g^{1+2\mathcal{O}/l}/n + 1)\right) \quad \text{as } n \to \infty. \tag{4.17}$$

The proof is given in Section 4.7. Next is the main theorem which states that the two-step estimator of the multivariate parameters is uniform consistent with probability tending to one as $n \to \infty$. When we write $\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ for some function $\boldsymbol{\theta}$, we mean the estimator we find by fixing the coefficient $\boldsymbol{\theta}$ in (4.6) where $F(\boldsymbol{\alpha}; \boldsymbol{\gamma})$ is replaced by $F(\boldsymbol{\alpha}; \boldsymbol{\theta})$. Thus $\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ depends on the regularization parameters $\lambda, \mu$. Let $\boldsymbol{\theta}_0$ denote the true parameter function of the PDE model (4.11).

**Theorem 4.1.** *Assume that (4.4),(4.5) and* ***A.1*** *to* ***A.8*** *(defined in Section 4.7) hold. Suppose the bound which is given in (4.17) tends to zero as* $n \to \infty$, *where* $\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}_0)$ *depends on* $\lambda, \mu$ *(see the comments before the statement of this theorem). Write* $\hat{\boldsymbol{\theta}}_n$ *since* $\hat{\boldsymbol{\theta}}$ *depends on* $n$. *Assume the first order partial derivatives of the sequence* $(\hat{\boldsymbol{\theta}}_n)$ *are bounded by a positive number with* *probability one. If* $\frac{m_{\theta_p}\lambda_{pj}}{n} = o(1)$ *for* $p = 1, \ldots, d, j = 1, \ldots, l$, *then*

$$\left\| \hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right\|_\infty = o_P(1).$$

The proof is in Section 4.7.

## 4.5 Example: The heat equation

Let us revisit the evolution of the temperature of a rod introduced in Section 1.3. Mathematically, this results in the well known 1D heat equation on a bar of length $L$ with Dirichlet boundary conditions:

$$\begin{cases} \dfrac{\partial g}{\partial t_2}(t_1, t_2) + D\dfrac{\partial^2 g}{\partial t_1^2}(t_1, t_2) + \theta(t_1, t_2) = 0 & \\ \qquad g(t_1, 0) = g_0(t_1) & 0 \leq t_1 \leq L \\ \qquad g(0, t_2) = 0 & 0 \leq t_2 \leq T \\ \qquad g(L, t_2) = 0 & 0 \leq t_2 \leq T. \end{cases} \tag{4.18}$$

Here, $g : [0, L] \times [0, T] \longrightarrow I\!R$ is the temperature of the bar at each point in space and time, $\theta : [0, L] \times [0, T] \longrightarrow I\!R$ is a source term heating or cooling the bar and $D$ is the diffusion coefficient which is assumed to be known. In this example $l = 2$ and the number of unknown parameters is $d = 1$.

*Remark* 4.1. By considering the substitutions

$$\begin{aligned} g(t_1, t_2) &\longleftarrow g(t_1, t_2) - g_0(t_1) \\ \theta(t_1, t_2) &\longleftarrow \theta(t_1, t_2) + D\frac{\partial^2 g_0}{\partial t_1^2}(t_1), \end{aligned}$$

we can restrict ourselves to the special case $g_0(t_1) = 0$.

It should be noted that we can implement general boundary conditions $g(0, t_2) = b_0(t_2)$ and $g(L, t_2) = b_L(t_2)$, see Remark 4.2.

### 4.5.1 Boundary constraints on the spline coefficients

The boundary conditions have to be satisfied by the spline estimator of $g$, which we enforce by imposing linear constraints on the tensor product B-spline coefficients. We use cubic splines.

Since cubic B-splines have a support which is spanned by 4 consecutive knots and vanish at endpoints, we have

$$\begin{aligned} s(t_1, 0) &= \sum_{i=1}^{m_{g1}} \sum_{j=1}^{m_{g2}} \alpha_{ij} B_{g1,i}(t_1) B_{g2,j}(0) \\ &= \sum_{i=1}^{m_{g1}} \big(\alpha_{i1} B_{g2,1}(0) + \alpha_{i2} B_{g2,2}(0) + \alpha_{i3} B_{g2,3}(0)\big) B_{g1,i}(t_1). \end{aligned}$$

Then $s(t_1, 0) = 0$ implies that

$$\alpha_{i1} B_{g2,1}(0) + \alpha_{i2} B_{g2,2}(0) + \alpha_{i3} B_{g2,3}(0) = 0, \quad \forall i \in \{1, \dots, m_{g1}\}, \tag{4.19}$$

because the B-splines constitute a basis. Using the tensor product notation, this can be written as

$$\mathbf{V}_1 \boldsymbol{\alpha} = \mathbf{0},$$

where $\mathbf{V}_1 \in \mathbf{R}^{m_{g1} \times (m_{g1} m_{g2})}$ which has on row $i$ and on columns $(i-1)m_{g2}+1, (i-1)m_{g2}+2, (i-1)m_{g2}+3$ the values $B_{g2,1}(0), B_{g2,2}(0), B_{g2,3}(0)$, respectively. For example, when $m_{g1} = m_{g2} = 4$,

$$\mathbf{V}_1 = \mathbf{I}_4 \otimes \begin{pmatrix} B_{g2,1}(0) & B_{g2,2}(0) & B_{g2,3}(0) & 0 \end{pmatrix} \in I\!R^{4 \times 16}.$$

Similarly, the boundary conditions $s(0, t_2) = 0, s(L, t_2) = 0$ translate to $\mathbf{V}_2 \boldsymbol{\alpha} = \mathbf{0}, \mathbf{V}_3 \boldsymbol{\alpha} = \mathbf{0}$, respectively.

*Remark* 4.2. We can implement general boundary conditions, for example when the boundary conditions are $g(0, t_2) = b_0(t_2)$ and $g(L, t_2) = b_L(t_2)$, by finding the spline approximations of $b_0$ and $b_L$. Then, by analogous reasoning, the right hand side of (4.19) is equal to the B-spline coefficients of $b_0$ and $b_L$, respectively.

### 4.5.2 Boundary constraints by using multiple knots B-spline

Throughout we use B-splines which are based on singular equidistant knots. However, when we use multiple knots at the boundary, we can enforce the boundary conditions (4.18) to hold in our modeling of $g$. Suppose we use cubic splines where the multiplicity of the initial knot (at $t_1 = 0$) is 4 and all the other knots are singular (for clarity of the presentation we allow the misuse of notation and denote the B-spline basis by $B_1, B_2, \dots, B_m$). Then $B_1(0) = 1$ and $B_i(0) = 0$ for all $2 \le i \le m$, see Figure 4.1(b). This means that for example the condition $g(0, t_2) = 0$ can be enforced by using for the $t_1$ direction the reduced basis $B_2, B_3, \dots, B_m$.

Thus to enforce all the boundary conditions we just employ multiple end knots and discard the first (and or last) B-spline.

This approach ensures the boundary conditions (4.18) hold exactly. Also, since the penalty term based on the boundary conditions can be removed from the definition of $J, H$ and $K$. This means that we need to determine one less regularization parameter, resulting in a lower computational cost.

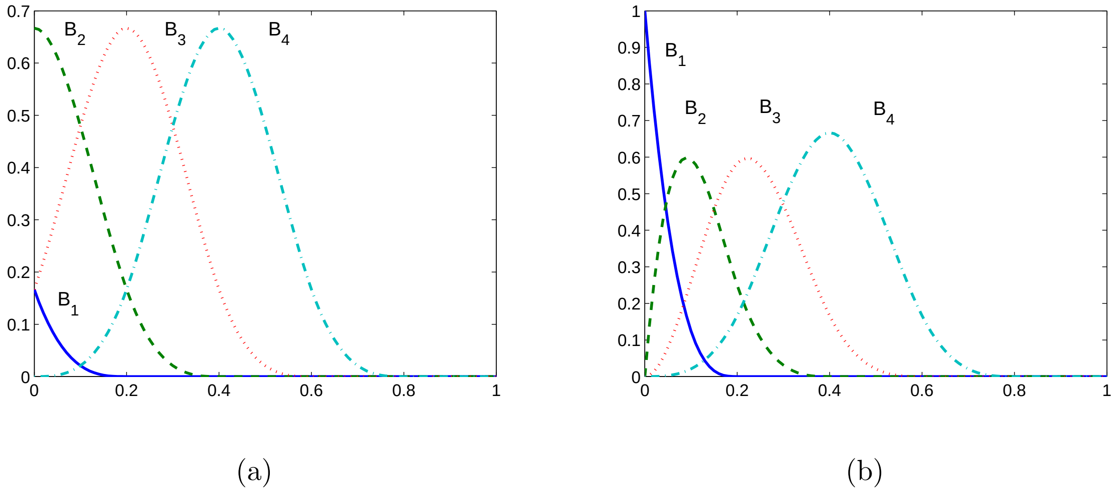



Figure 4.1: Single and multiple knots cubic B-splines on the interval $[0, 1]$. Left: first four B-splines with equidistant single knots. Right: first four B-splines with equidistant interior knots and multiple initial knot.

Other more advanced techniques to enforce boundary conditions by basis adjustments exist in the literature. These include Web-splines (Höllig (2003)) and I-splines (Sanches et al. (2011)). Implementing such techniques is beyond the scope of this chapter.

### 4.5.3 Simulation: heat equation

We conduct a simulation study of the heat equation (4.18). We assume that the diffusion coëfficiënt $D = -\pi$ is known, we set the length of the rod $L = 5$ and simulate measurements up to $T = 10$. We now wish to estimate the unknown source term $\theta$ by the one- and by the two-step method. We run 200 simulations which consist of $n = 5151$ observations on the rectangular grid

$$\{0, 0.1, 0.2, \ldots, 5\} \times \{0, 0.1, 0.2, \ldots, 10\}.$$

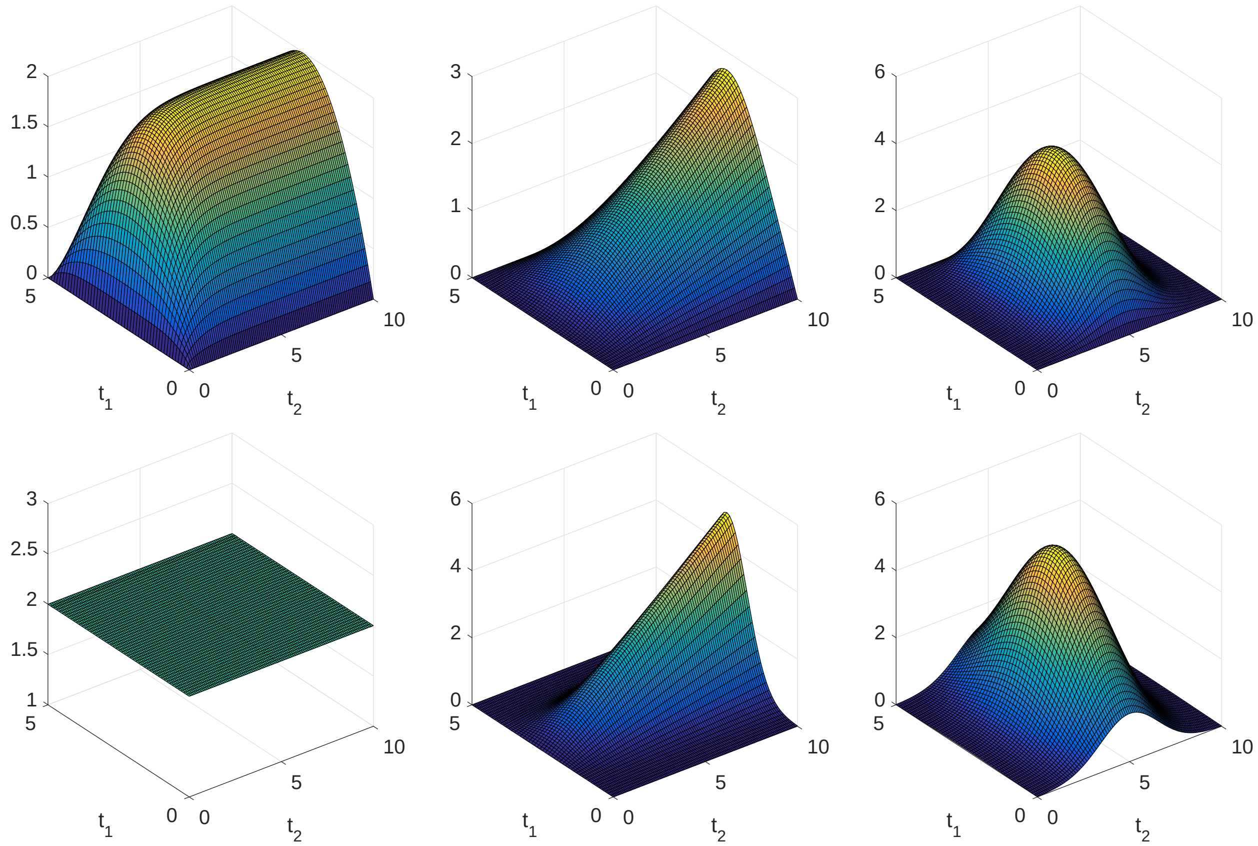



Figure 4.2: The solutions $g$ (top row) of the heat equation (4.18) for the various sources $\theta$ (bottom row). Cases 1, 2 and 3 are displayed column-wise from left to right.

The measurements themselves are generated based on the Fourier expansion of the solution with i.i.d. additive noise terms $\varepsilon_i \sim N(0, \sigma^2)$. We consider two different noise levels $\sigma = 0.05$ and $\sigma = 0.1$ and three different source terms $\theta$:

- $\theta_1(t_1, t_2) = 2$,
- $\theta_2(t_1, t_2) = \frac{1}{2} t_2 \exp\left(-(t_1 - \frac{5}{2})^2\right)$,
- $\theta_3(t_1, t_2) = 5 \exp\left(-\frac{1}{5}(\mathbf{t} - \mathbf{u})(\mathbf{t} - \mathbf{u})^\top\right)$, with $\mathbf{u} = (5/2, 10/2)$.

These functions and the corresponding solution $g$ can be seen in Figure 4.2.

We choose a generous number of knots to increase flexibility, keeping in mind that the regularization parameters ensure smoothness (Marx and Eilers (2005)). The temperature function $g$ is modeled by cubic tensor product B-splines with 20 equidistant knots in both directions. The source term $\theta$ is modeled in a similar fashion with 15 equidistant knots in both directions. We choose less knots for the multivariate parameters because we only observe these parameters indirectly and we do not need their partial derivatives. The regularization parameters $\lambda, \mu, \lambda_1, \lambda_2$ are chosen on the grid

$$\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 10^2, 10^3, 10^4, 10^5, 10^6\}$$

based on the criterion described in Section 4.3.4.

The effectiveness of the methods is shown by the relative errors of the source term $\theta$ and the temperature function $g$:

$$\frac{\sqrt{\sum_{i=1}^{n}(\hat{\theta}(\mathbf{t}_i) - \theta(\mathbf{t}_i))^2}}{\sqrt{\sum_{i=1}^{n}\theta(\mathbf{t}_i)^2}}, \qquad \frac{\sqrt{\sum_{i=1}^{n}(\hat{g}(\mathbf{t}_i) - g(\mathbf{t}_i))^2}}{\sqrt{\sum_{i=1}^{n} g(\mathbf{t}_i)^2}}. \tag{4.20}$$

The results are collected in Table 4.1 and 4.2. These results reveal that the full basis (of B-splines with equidistant single knots) yields more accurate results in all cases except for the estimation of $g$ when $\theta = \theta_2$ and $\sigma = 0.05$. This is explained by the fact that by construction the reduced basis approach is less flexible for the estimation of the boundary regions of $g$. Then, the partial derivatives in the boundary regions of $g$ are not accurate which leads to a very low accuracy in boundary estimations of $\theta$. This is confirmed by the figures in Section 4.10. For the full basis, the overall conclusion is that the two-step method yields better results than the one-step method. As it is expected, Table 4.1 and 4.2 show that a higher measurement precision of $g$ generally gives better estimates. Some examples of the state variable $g$ are shown in Section 4.10 and illustrate again that the full basis approach gives much better estimates than the reduced basis approach, especially at the boundaries of $\theta$. Also these estimates mimic very well the true shapes depicted in Figure 4.2.

## 4.6 Conclusion

In this chapter estimating multivariate parameters of linear PDEs from measurements of the state variable is considered. We proposed to model the PDE parameters and the state variable by tensor product splines, then the estimators are obtained by considering a trade-off between least squares measure of fit, a measure of fidelity to the PDE model and boundary conditions.

We investigated two avenues for enforcing boundary conditions on the modelling of the state variable. The first enforces the boundary conditions exactly trough a reduced B-spline basis:

| | Two-step method | | One-step method | |
|---|---|---|---|---|
| | Full basis | Reduced basis | Full basis | Reduced basis |
| $g$ | 6.45e−3(1.35e−3) | 1.50e−2(5.78e−3) | 6.99e−3(1.42e−3) | 9.83e−3(2.97e−3) |
| $\theta = \theta_1$ | 4.57e−2(2.30e−2) | 4.36e−1(2.30e−1) | 6.41e−2(1.96e−2) | 5.06e−1(8.87e−2) |
| $g$ | 1.19e−2(5.52e−3) | 1.28e−2(4.68e−3) | 1.48e−2(3.84e−3) | 1.55e−2(5.83e−3) |
| $\theta = \theta_2$ | 2.83e−1(7.91e−2) | 4.44e−1(4.43e−2) | 3.25e−1(5.53e−2) | 4.16e−1(4.44e−2) |
| $g$ | 4.55e−3(6.76e−4) | 9.96e−3(3.61e−3) | 1.55e−2(2.66e−2) | 1.11e−1(3.88e−2) |
| $\theta = \theta_3$ | 3.55e−1(4.25e−3) | 4.95e−1(1.03e−1) | 3.55e−1(6.15e−3) | 7.34e−1(1.14e−1) |

Table 4.1: Mean and standard deviation (in brackets) of the relative error for 200 runs with noise generated for $\sigma = 0.10$.

| | Two-step method | | One-step method | |
|---|---|---|---|---|
| | Full basis | Reduced basis | Full basis | Reduced basis |
| $g$ | 5.03e−3(7.47e−4) | 1.33e−2(5.53e−3) | 6.65e−3(1.64e−3) | 7.76e−3(1.00e−3) |
| $\theta = \theta_1$ | 3.63e−2(1.39e−2) | 4.52e−1(2.13e−1) | 7.04e−2(8.53e−3) | 5.55e−1(4.56e−2) |
| $g$ | 9.41e−3(6.47e−3) | 7.81e−3(3.61e−3) | 1.39e−2(4.12e−3) | 1.27e−2(4.32e−3) |
| $\theta = \theta_2$ | 2.49e−1(9.72e−2) | 4.71e−1(5.16e−2) | 3.20e−1(6.60e−2) | 4.25e−1(2.81e−2) |
| $g$ | 2.35e−3(3.59e−4) | 7.55e−3(3.05e−3) | 1.23e−2(2.58e−2) | 1.17e−1(2.64e−2) |
| $\theta = \theta_1$ | 3.53e−1(2.24e−3) | 5.49e−1(1.06e−1) | 3.54e−1(5.71e−3) | 7.51e−1(6.68e−2) |

Table 4.2: Mean and standard deviation (in brackets) of the relative error for 200 runs with noise generated for $\sigma = 0.05$.

we use a B-spline basis with multiple knots at the (relevant) boundaries where we omit the first and (or) last basis function. The second is by adding constraints on the spline coefficients. Also, asymptotic results were included.

We compared different estimation methods on simulated data which describe the heating and cooling of a rod due to an unknown source term. The source term was allowed to be space (one dimensional) and time dependent. We found that our best estimating method (i.e. the method with complete bases and added boundary constraints) succeeds in reconstructing the source term well, for several source shapes.

## 4.7 Proofs

Matrix and function norms are defined in Appendix A. We use the notation $a_n \asymp b_n$ to denote that $a_n/b_n$ and $b_n/a_n$ are bounded. Let $C(\mathcal{H})$ the space of continuous functions $\mathcal{H} \to I\!R^d$. Recall that $\boldsymbol{\alpha}$ ($m_g$) and $\boldsymbol{\gamma}$ ($m_{\boldsymbol{\theta}}$) depend on $n$, i.e. they are allowed to increase in dimension as $n \to \infty$. The consistency results are based on few assumptions which are needed in the proof of Theorem 4.1, which establishes the consistency of our multivariate parameter estimator.

**A.1** The function $g : \mathcal{H} \to I\!R$ is bounded measurable, i.e. $\|g\|_\infty < \infty$ and $g$ is integrable.

**A.2** The functions $h_i(\boldsymbol{\theta})$ are bounded for all bounded $\boldsymbol{\theta}$.

**A.3** The tensor product spline coefficients satisfy $\boldsymbol{\gamma}_p \in \mathcal{K}^{m_{\theta_p}}$ for some compact $\mathcal{K} \subset I\!R$.

**A.4** The spline dimensions $m_{gi}$ satisfy $m_{gi} \asymp m_{gj}$ $\forall i, j \in \{1, \ldots, l\}$, hence $m_{gi} \asymp m_g^{1/l}$ for all $i \in \{1, \ldots, l\}$.

**A.5** $\frac{1}{n}\sum_{i=1}^{n} (\mathbf{B}(\mathbf{t}_i)(\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}_1) - \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}_2)))^2$ is continuous in $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ ($C(\mathcal{H})$ is equipped with $\|\cdot\|_\infty$) and converges (with probability tending to one) uniformly over $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \in C(\mathcal{H}) \times C(\mathcal{H})$.

**A.6** $Q(\boldsymbol{\theta}) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} (\mathbf{B}(\mathbf{t}_i)(\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}_0)))^2$ has a unique minimum in $\boldsymbol{\theta} = \boldsymbol{\theta}_0$

**A.7** $\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{B}(\mathbf{t}_i)(\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}_0) - \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta})) \xrightarrow{P} 0$ for any $\boldsymbol{\theta} \in C(\mathcal{H})$.

**A.8** There exists a positive constant $M$ such that $P(\varepsilon_i < M) = 1$ for all $i \in \{1, \ldots, n\}$.

**A.1**-**A.4** are very general and are satisfied in practice. In the literature it is often assumed that unknown parameters are contained in some compact space, our equivalent is **A.3**. Assumptions **A.5** and **A.6** are inspired by the assumptions 1 and 2 in Yu and Ruppert (2002). **A.5** says that $\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}_1)$ is a smooth function of $\boldsymbol{\theta}_1$, **A.6** is needed for identifiability reasons. **A.8** states that the magnitude of the measurement errors are, sensibly so, bounded.

In Proposition 4.1 we show the consistency of the estimator of $g$. In the proof of this proposition we need the following proposition (on the inverse of the sum of two matrices) which is derived from Corollary 5.6.16 Horn and Johnson (1988).

**Proposition 4.2.** *Let $\mathbf{G}_1, \mathbf{G}_2$ be $m \times m$ square matrices such that $\mathbf{G}_1$ is invertible. If there exists a matrix norm $\|\cdot\|$ such that $\|\mathbf{G}_1^{-1}\mathbf{G}_2\| < 1$, then*

$$(\mathbf{G}_1 + \mathbf{G}_2)^{-1} = \sum_{j=0}^{\infty} (-1)^j \mathbf{G}_1^{-1} (\mathbf{G}_2 \mathbf{G}_1^{-1})^j.$$

*Proof.* By Corollary 5.6.16 of Horn and Johnson (1988), we have that for an invertible matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$, where $\|\mathbf{I}_m - \mathbf{A}\| < 1$ and where $\mathbf{I}_m$ is the identity matrix,

$$\mathbf{A}^{-1} = \sum_{j=0}^{\infty} (\mathbf{I}_m - \mathbf{A})^j.$$

Apply the previous equation to $\mathbf{A} = \mathbf{I}_m + \mathbf{G}_2 \mathbf{G}_1^{-1}$ to find

$$(\mathbf{I}_m + \mathbf{G}_2 \mathbf{G}_1^{-1})^{-1} = \sum_{j=0}^{\infty} (-1)^j (\mathbf{G}_2 \mathbf{G}_1^{-1})^j,$$

and left multiply both sides with $\mathbf{G}_1^{-1}$ to find

$$(\mathbf{G}_1 + \mathbf{G}_2)^{-1} = \sum_{j=0}^{\infty} (-1)^j \mathbf{G}_1^{-1} (\mathbf{G}_2 \mathbf{G}_1^{-1})^j,$$

which is the desired equation. ■

Before we proceed to the proof of Proposition 4.1 we obtain a bound on $\|\mathbf{f}^\top(\mathbf{t};\boldsymbol{\gamma})\|_\infty$. By **A.2** it suffices to bound B-splines and their derivatives since the components of $\mathbf{f}^\top(\mathbf{t};\boldsymbol{\gamma})$ are composed of B-splines and their derivatives up to the order $\mathcal{O}$ (this follows clearly from the tensor product notation). B-splines are bounded: they take values between 0 and 1. The derivative of a B-spline $B_j(t;q)$ having distance $\frac{1}{K}$ between the equidistant knots, is (De Boor (2001), page 116)

$$B_j'(t;q) = K(B_{j-1}(t;q-1) - B_j(t;q-1)).$$

Hence, the derivative of B-splines is bounded by the number of knots. We can reapply the previous property to find that the $\mathcal{O}$th derivative of B-splines is bounded by the $\mathcal{O}$th power of the number of knots. Thus $\|\mathbf{f}(\mathbf{t};\boldsymbol{\gamma})\|_\infty = O\left(m_g^{\mathcal{O}/l}\right)$ and $\|\mathbf{f}^\top(\mathbf{t};\boldsymbol{\gamma})\|_\infty = O\left(m_g^{\mathcal{O}/l+1}\right)$.

*Proof of Proposition 4.1.* Let $\hat{\boldsymbol{\alpha}}_{un}$ denote the unpenalized spline estimator, i.e. $\lambda = 0 = \mu$ in (4.6). We prove (4.17) by using the triangle inequality

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(g(\mathbf{t}_i) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}))^2} \leq \sqrt{\frac{1}{n}\sum_{i=1}^{n}(g(\mathbf{t}_i) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_{un})^2}$$

$$+\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathbf{B}^\top(\mathbf{t}_i)(\hat{\boldsymbol{\alpha}}_{un}-\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})))^2}. \quad (4.21)$$

We subsequently bound the first part and the second part of (4.21). A bound on the first part follows from Theorem 1 of Huang (1998)

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(g(\mathbf{t}_i)-\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_{un})^2}=O_P(\sqrt{\frac{m_g}{n}+\rho_n^2}). \quad (4.22)$$

Moreover, Lemma 8.9 of Yoo and Ghosal (2016) establishes that

$$\boldsymbol{\alpha}^\top(\mathcal{B}^\top\mathcal{B})\boldsymbol{\alpha}\asymp nm_g^{-1}\|\boldsymbol{\alpha}\|_2^2, \quad (4.23)$$

if (4.4) and (4.5) hold. Thus for $n$ sufficiently large, the eigenvalues of $(\mathcal{B}^\top\mathcal{B})$ are contained in $[C_1nm_g^{-1}, C_2nm_g^{-1}]$ for positive constants $C_1 < C_2$. Therefore $(\mathcal{B}^\top\mathcal{B})$ is invertible for $n$ sufficiently large. Thus, under such conditions $\hat{\boldsymbol{\alpha}}_{un}$ (is unique and) equals

$$\hat{\boldsymbol{\alpha}}_{un}=(\mathcal{B}^\top\mathcal{B})^{-1}\mathcal{B}^\top\mathbf{Y}.$$

For the second part of (4.21), we bound

$$\begin{aligned}
\mathbf{B}^\top(\mathbf{t})\hat{\boldsymbol{\alpha}}_{un}-\mathbf{B}^\top(\mathbf{t})\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}) &= \mathbf{B}^\top(\mathbf{t})(\hat{\boldsymbol{\alpha}}_{un}-\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}))\\
&= \mathbf{B}^\top(\mathbf{t})\left((\mathcal{B}^\top\mathcal{B})^{-1}-\left(\mathcal{B}^\top\mathcal{B}+\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)^{-1}\right)\mathcal{B}^\top\mathbf{Y}\\
&\quad+\mathbf{B}^\top(\mathbf{t})\left(\mathcal{B}^\top\mathcal{B}+\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)^{-1}\lambda\mathbf{T}^\top(\boldsymbol{\gamma})\boldsymbol{\gamma}\\
&= \mathbf{a}^\top\mathbf{Y}+\mathbf{B}^\top(\mathbf{t})\left(\mathcal{B}^\top\mathcal{B}+\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)^{-1}\lambda\mathbf{T}^\top(\boldsymbol{\gamma})\boldsymbol{\gamma}, \quad (4.24)
\end{aligned}$$

in three steps.

- **Step 1**: $\left((\mathcal{B}^\top\mathcal{B})^{-1}-\left(\mathcal{B}^\top\mathcal{B}+\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)^{-1}\right)=O\left(\frac{m_g^3}{n^2}(\lambda m_g^{2\mathcal{O}/l}+\mu)\right)$.

  By Proposition 4.2,

  $$\left(\mathcal{B}^\top\mathcal{B}+\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)^{-1}=\sum_{j=0}^{\infty}(-1)^j(\mathcal{B}^\top\mathcal{B})^{-1}\left(\left(\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)(\mathcal{B}^\top\mathcal{B})^{-1}\right)^j$$

  if the series converges. This series converges if

  $$\left\|\left((\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j)(\mathcal{B}^\top\mathcal{B})^{-1}\right)\right\|_\infty<1. \quad (4.25)$$

Condition (4.25) holds if we demonstrate that

$$\left\|\left((\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j)(\mathcal{B}^\top\mathcal{B})^{-1}\right)\right\|_\infty = O\left(\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l}+\mu)\right). \tag{4.26}$$

This follows from three facts. First, the fact that

$$\begin{aligned}\|\mathbf{R}(\boldsymbol{\gamma})\|_\infty &\le \int_{\mathcal{H}}\|\mathbf{f}(\mathbf{t};\boldsymbol{\gamma})\|_\infty\|\mathbf{f}^\top(\mathbf{t};\boldsymbol{\gamma})\|_\infty\, d\mathbf{t}\\ &\le \int_{\mathcal{H}} m_g^{\mathcal{O}/l}m_g m_g^{\mathcal{O}/l}\, d\mathbf{t} = O(m_g^{1+2\mathcal{O}/l})\end{aligned}$$

because of the comments before this proof and the compactness of $\mathcal{H}$ , second,

$$\left\|\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right\|_\infty = O(m_g),$$

since all $\mathbf{V}_j$ are bounded because their terms are B-splines which are bounded by 1, and third,

$$\|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty = O(\frac{m_g}{n}), \tag{4.27}$$

which follows from the observation after (4.23) and Lemma 8.4 of Yoo and Ghosal (2016).

Then

$$\left\|(\mathcal{B}^\top\mathcal{B})^{-1}-(\mathcal{B}^\top\mathcal{B}+\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j)^{-1}\right\|_\infty = O\left(\frac{m_g^3}{n^2}(\lambda m_g^{2\mathcal{O}/l}+\mu)\right) \tag{4.28}$$

follows from (4.26) which tends to zero by assumption in the proposition statement, via

$$\begin{aligned}&\left\|\sum_{j=1}^{\infty}(-1)^j(\mathcal{B}^\top\mathcal{B})^{-1}\left((\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j)(\mathcal{B}^\top\mathcal{B})^{-1}\right)^j\right\|_\infty\\ &\le \|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty\sum_{j=1}^{\infty}\left\|\left(\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\left(\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j\right)\right)\right\|_\infty^j\|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty^j\\ &= \|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty\left(\frac{1}{1-\|(\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j)\|_\infty\|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty}-1\right)\\ &= \|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty\frac{\|(\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j)\|_\infty\|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty}{1-\|(\lambda\mathbf{R}(\boldsymbol{\gamma})+\mu\sum_{j=1}^{\mathcal{C}}\mathbf{V}_j^\top\mathbf{V}_j)\|_\infty\|(\mathcal{B}^\top\mathcal{B})^{-1}\|_\infty}\\ &= O\left(\frac{m_g^3}{n^2}(\lambda m_g^{2\mathcal{O}/l}+\mu)\right)\end{aligned}$$

- **Step 2**: $\mathbf{a}^\top\mathbf{Y} = O_P\left(\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l}+\mu)\right)$. We have by the normality property of B-splines $\|\mathbf{B}^\top(\mathbf{t})\|_\infty = 1$. By Lemma 8.3 of Yoo and Ghosal (2016) $\|\mathcal{B}^\top\|_\infty = O(\frac{n}{m_g})$, recalling the definition of $\mathcal{B}$ (see (4.13)). Therefore, using (4.28), (4.24) and **A.1**, we obtain

$$\|\mathbf{a}^\top\|_\infty = O\left(\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l} + \mu)\right), \quad \|\mathbf{a}^\top \mathbf{g}\|_\infty = O\left(\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l} + \mu)\right), \tag{4.29}$$

where $\mathbf{g} = (g(\mathbf{t}_1), \ldots, g(\mathbf{t}_n))^\top$. Also, with $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^\top$,

$$\|\mathbf{a}^\top \boldsymbol{\varepsilon}\|_\infty \leq \|\mathbf{a}^\top\|_\infty \|\boldsymbol{\varepsilon}\|_\infty = O_P\left(\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l} + \mu)\right)$$

by (4.29) and **A.8**. Hence

$$\|\mathbf{a}^\top \mathbf{Y}\|_\infty = O_P\left(\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l} + \mu)\right), \tag{4.30}$$

since $\mathbf{Y} = \mathbf{g} + \boldsymbol{\varepsilon}$.

- **Step 3**:

$$\mathbf{B}^\top(\mathbf{t})\left(\mathcal{B}^\top\mathcal{B} + \lambda \mathbf{R}(\boldsymbol{\gamma}) + \mu \sum_{j=1}^{\mathcal{C}} \mathbf{V}_j^\top \mathbf{V}_j\right)^{-1} \lambda \mathbf{T}^\top(\boldsymbol{\gamma})\boldsymbol{\gamma} = O\left(\frac{\lambda m_g^{1+\mathcal{O}/l}}{n} + \frac{\lambda m_g^{3+\mathcal{O}/l}}{n^2}(\lambda m_g^{2\mathcal{O}/l} + \mu)\right).$$

Apply the triangle inequality to (4.27) and (4.28) to find

$$\left\|\mathbf{B}^\top(\mathbf{t})(\mathcal{B}^\top\mathcal{B} + \lambda \mathbf{R}(\boldsymbol{\gamma}) + \mu \sum_{j=1}^{\mathcal{C}} \mathbf{V}_j^\top \mathbf{V}_j)^{-1}\right\|_\infty = O_P\left(\frac{m_g}{n} + \frac{m_g^3}{n^2}(\lambda m_g^{2\mathcal{O}/l} + \mu)\right). \tag{4.31}$$

and

$$\begin{aligned} \|\lambda \mathbf{T}^\top(\boldsymbol{\gamma})\boldsymbol{\gamma}\|_\infty &\leq \lambda \|\mathbf{T}^\top(\boldsymbol{\gamma})\|_\infty \|\boldsymbol{\gamma}\|_\infty \\ &\leq \lambda \|\boldsymbol{\gamma}\|_\infty \int_{\mathcal{H}} \|\mathbf{f}(\mathbf{t};\boldsymbol{\gamma})^\top\|_\infty \|\mathbf{Q}_d^\top(\mathbf{t})\|_\infty d\mathbf{t} \\ &= O(\lambda m_g^{\mathcal{O}/l}), \end{aligned} \tag{4.32}$$

$\|\mathbf{Q}_d^\top(\mathbf{t})\|_\infty = 1$ since B-splines sum up to 1.

Finally, by **Step 2** and **Step 3** we have shown that uniformly in $\mathbf{t}$

$$|\mathbf{B}^\top(\mathbf{t})\hat{\boldsymbol{\alpha}}_{un} - \mathbf{B}^\top(\mathbf{t})\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})| = O_P\left(\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l} + \mu) + \lambda \frac{m_g^{1+\mathcal{O}/l}}{n} + \lambda \frac{m_g^{3+\mathcal{O}/l}}{n^2}(\lambda m_g^{2\mathcal{O}/l} + \mu)\right).$$

By the triangle inequality and (4.22) we have established the desired result

$$\frac{1}{n}\sum_{i=1}^{n}(g(\mathbf{t}_i)-\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}))^2 =$$
$$O_P\left(\frac{m_g}{n}+\rho_n^2+\left(\frac{m_g^4}{n^2}+\lambda^2\frac{m_g^{6+2\mathcal{O}/l}}{n^4}\right)(\lambda m_g^{2\mathcal{O}/l}+\mu)^2+\lambda^2\frac{m_g^{2+2\mathcal{O}/l}}{n^2}\right),$$

using the assumption $\frac{m_g^2}{n}(\lambda m_g^{2\mathcal{O}/l}+\mu)\to 0$ we deduce that

$$\frac{1}{n}\sum_{i=1}^{n}(g(\mathbf{t}_i)-\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}))^2 = O_P\left(\rho_n^2+\frac{m_g}{n}(\lambda^2 m_g^{1+2\mathcal{O}/l}/n+1)\right).$$

■

The next lemma shows that $\hat{\boldsymbol{\gamma}}$ is a stochastic variable that exists and which is essentially Lemma 2 of Jennrich (1969). Moreover, to stress the dependence of $\boldsymbol{\gamma}$ on $n$, we add the subscript $n$ to $\boldsymbol{\gamma}$ (and $\hat{\boldsymbol{\gamma}}$) in the upcoming proofs. Also, we use the notation $\boldsymbol{\theta}_n(\mathbf{t})$ ($\hat{\boldsymbol{\theta}}_n(\mathbf{t})$) to denote the spline function which corresponds to $\boldsymbol{\gamma}_n$ ($\hat{\boldsymbol{\gamma}}_n$).

**Lemma 4.1.** *The estimator $\hat{\boldsymbol{\gamma}}_n$ which minimizes*

$$\sum_{i=1}^{n}\left(Y_i-\mathbf{B}(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}_n)\right)^2+\sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\|_2^2$$

*exists.*

*Proof.* Define the real valued function

$$Q_n(\boldsymbol{\gamma}_n,Y_1,Y_2,\ldots,Y_n)=\sum_{i=1}^{n}\left(Y_i-\mathbf{B}(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}_n)\right)^2+\sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\|_2^2.$$

Note that $(Y_1,\ldots,Y_n)\in\mathcal{M}$ where $\mathcal{M}$ is the product of measurable space $I\!R^n$. Apply Lemma 2 of Jennrich (1969) with real valued function $Q_n$ and compact subset $\mathcal{K}^m$ to show the existence of the statistic $\hat{\boldsymbol{\gamma}}_n=(\hat{\boldsymbol{\gamma}}_1^\top,\ldots,\hat{\boldsymbol{\gamma}}_d^\top)^\top$.

■

Next, we show the uniform consistency of our parameter estimator. The proof is inspired to some extent by the proof of Theorem 1 Yu and Ruppert (2002). We do not assume $\boldsymbol{\theta}_0$ to be a collocation of tensor product spline functions, that causes us to use in the proof the Arzelà-Ascoli Theorem with the assumption that the components of $\hat{\boldsymbol{\theta}}$ have bounded partial derivatives. In the other case $\boldsymbol{\theta}_0$ is the collocation of tensor product spline functions, hence the number of knots are fixed, and therefore the conditions of the Arzelà-Ascoli Theorem are satisfied by our comments before the proof of Proposition 4.1. By $\hat{\boldsymbol{\gamma}}_p$ we mean the spline coefficients estimator corresponding to the parameter function $\hat{\boldsymbol{\theta}}_p$.

*Proof of Theorem 4.1.* Let $\mathcal{F} \subset C(\mathcal{H})$ denote a compact subspace. Define for $\boldsymbol{\theta} \in \mathcal{F}$

$$Q_n(\boldsymbol{\theta};\boldsymbol{\gamma}_n) = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{Y}_i - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}))^2 + \frac{1}{n}\left(\sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\|\mathbf{P}_{pi}\boldsymbol{\gamma}_p\|_2^2\right), \tag{4.33}$$

which can be written as

$$\begin{aligned}
Q_n(\boldsymbol{\theta};\boldsymbol{\gamma}_n) &= \frac{1}{n}\sum_{i=1}^{n}\big((\mathbf{Y}_i - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0)) + (\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}))\big)^2 \\
&+ \frac{1}{n}\left(\sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\|_2^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\big(\mathbf{Y}_i - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0)\big)^2 \\
&+ \frac{1}{n}\sum_{i=1}^{n}\big(\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})\big)^2 \\
&+ \frac{2}{n}\sum_{i=1}^{n}\big(\mathbf{Y}_i - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0)\big)\big(\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})\big) \\
&+ \frac{1}{n}\left(\sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\|_2^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i + f_i)^2 + \frac{1}{n}\sum_{i=1}^{n}\big(\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})\big)^2 \\
&+ \frac{2}{n}\sum_{i=1}^{n}(\varepsilon_i + f_i)\big(\mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})\big) \\
&+ \frac{1}{n}\left(\sum_{j=1}^{l}\sum_{p=1}^{d}\lambda_{pj}\|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\|_2^2\right) \\
&= A_1 + A_2 + A_3 + A_4,
\end{aligned}$$

where $f_i$ arises as the difference $f_i = g(\mathbf{t}_i) - \mathbf{B}^\top(\mathbf{t}_i)\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0)$.

We demonstrate that $Q_n(\boldsymbol{\theta};\boldsymbol{\gamma}_n) = A_1 + A_2 + A_3 + A_4 \xrightarrow{P} \sigma^2 + Q(\boldsymbol{\theta})$ as $n \to \infty$ uniformly in $\boldsymbol{\theta} \in \mathcal{F}$. For $A_1$ we have

$$A_1 = \frac{1}{n}\sum_{i}\varepsilon_i^2 + \frac{2}{n}\sum_{i}\varepsilon_i f_i + \frac{1}{n}\sum_{i} f_i^2 \xrightarrow{P} \sigma^2 \tag{4.34}$$

using the (strong) law of large numbers, the fact $f_i \to 0$ (by Proposition 4.1 and (**A.7**)) and the Cauchy-Schwarz inequality. By **A.5** , $A_2 \xrightarrow{P} Q(\boldsymbol{\theta})$. Then

$$A_3 = \frac{2}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{B}^\top(\mathbf{t}_i)\left(\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})\right) + \frac{2}{n}\sum_{i=1}^{n} f_i \mathbf{B}^\top(\mathbf{t}_i)\left(\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})\right), \tag{4.35}$$

where uniformly in $\boldsymbol{\theta} \in \mathcal{F}$, the first part converges in probability to 0 by Lemma 4.2 (given below) and the second part gives by the Cauchy-Schwarz inequality

$$\begin{aligned} \frac{2}{n}\sum_{i=1}^{n} \left| f_i \mathbf{B}^\top(\mathbf{t}_i)\left(\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})\right)\right| &\le 2\left(\frac{1}{n}\sum_{i}^{n} f_i^2\right)^{1/2} \left(\frac{1}{n}\sum_{i}^{n}\left(\mathbf{B}^\top(\mathbf{t}_i)(\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}))\right)^2\right)^{1/2} \\ &= o_P(1). \end{aligned}$$

Finally, we show that $A_4 \to 0$ using **A.3** and the fact that $\|\mathbf{P}_{pj}^\top\|_\infty, \|\mathbf{P}_{pj}\|_\infty$ are bounded (see Section 4.3.3),

$$\begin{aligned} \|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\| &= \|\mathbf{P}_{pj}\boldsymbol{\gamma}_p\|_\infty \\ &\le \|\boldsymbol{\gamma}_p^\top\|_\infty \|\mathbf{P}_{pj}^\top\|_\infty \|\mathbf{P}_{pj}\|_\infty \|\boldsymbol{\gamma}_p\|_\infty = O(m_{\gamma p}) \end{aligned}$$

Thus $A_4 = O(\frac{1}{n}\sum_{j=1}^{l}\sum_{p=1}^{d} m_{\gamma p}\lambda_{pj})$. Therefore

$$Q_n(\boldsymbol{\theta}; \boldsymbol{\gamma}_n) \xrightarrow{P} Q(\boldsymbol{\theta}) + \sigma^2, \tag{4.36}$$

uniformly in $\boldsymbol{\theta} \in \mathcal{F}$.

Now we show the uniform consistency of $\hat{\boldsymbol{\theta}}_n$. By Lemma 4.1 $\hat{\gamma}_n$ that minimizes $Q_n(\boldsymbol{\theta}_n; \boldsymbol{\gamma}_n)$ exists. The theory above is developed around compact spaces. The space of interest is

$$\mathcal{A} = \{\hat{\boldsymbol{\theta}}_n \mid n \ge n_0\},$$

for a fixed $n_0 \in \mathbb{N}$. By **A.3** and the properties of B-splines $\mathcal{A}$ is a set of uniformly bounded functions. By assumption the partial derivatives of $\hat{\boldsymbol{\theta}}_n$ are bounded by the same constant, and by the mean value inequality the set $\mathcal{A}$ is Lipschitz continuous with one Lipschitz constant. Hence $\mathcal{A}$ is a set of equicontinuous functions. Let $\bar{\mathcal{A}} \subset C(\mathcal{H})$ denote the closure of $\mathcal{A}$. Then $\bar{\mathcal{A}}$ is closed by definition, it is uniformly bounded and equicontinuous because $\mathcal{A}$ is uniformly bounded and equicontinuous. By the Arzelà-Ascoli Theorem $\bar{\mathcal{A}}$ is compact, thus every sequence in $\bar{\mathcal{A}}$ has a converging subsequence. Thus $(\hat{\boldsymbol{\theta}}_n)$ has a subsequence $(\hat{\boldsymbol{\theta}}_{n_k})$ which converges uniformly to a function $\boldsymbol{\theta}'$, i.e. $\|\hat{\boldsymbol{\theta}}_{n_k} - \boldsymbol{\theta}'\|_\infty = o_P(1)$. By **A.5**, $Q(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$, thus

$$Q(\hat{\boldsymbol{\theta}}_{n_k}) \xrightarrow{P} Q(\boldsymbol{\theta}'). \tag{4.37}$$

Write

$$Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k};\hat{\gamma}_{n_k}) - Q(\boldsymbol{\theta}') - \sigma^2 = [Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k};\hat{\gamma}_{n_k}) - Q(\hat{\boldsymbol{\theta}}_{n_k}) - \sigma^2] + [Q(\hat{\boldsymbol{\theta}}_{n_k}) - Q(\boldsymbol{\theta}')], \tag{4.38}$$

where the first term converges to zero in probability since

$$|Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k};\hat{\gamma}_{n_k}) - Q(\hat{\boldsymbol{\theta}}_{n_k}) - \sigma^2| \leq \sup_{\boldsymbol{\theta}\in\bar{\mathcal{A}}} |Q_{n_k}(\boldsymbol{\theta};\hat{\gamma}_{n_k}) - Q(\boldsymbol{\theta}) - \sigma^2|,$$

and the latter tends to zero in probability by (4.36). The second term also converges to zero in probability, because $Q$ is continuous in $\boldsymbol{\theta}$ by **A.4** and **A.5**, and $\|\hat{\boldsymbol{\theta}}_{n_k} - \boldsymbol{\theta}'\|_\infty = o_P(1)$. Thus

$$Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k};\hat{\gamma}_{n_k}) \xrightarrow{P} Q(\boldsymbol{\theta}') + \sigma^2. \tag{4.39}$$

On the other side there is a sequence of spline functions $(\tilde{\boldsymbol{\theta}}_{n_k})$ (with coefficients $(\tilde{\gamma}_{n_k})$) from the same spline spaces as the sequence $(\hat{\boldsymbol{\theta}}_{n_k})$ such that $\|\tilde{\boldsymbol{\theta}}_{n_k} - \boldsymbol{\theta}_0\|_\infty \to 0$ when $n_k \to \infty$ (Theorem 12.8 in Schumaker (2007)). Next, since $Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k};\hat{\gamma}_{n_k}) \leq Q_{n_k}(\tilde{\boldsymbol{\theta}}_{n_k};\tilde{\gamma}_{n_k})$ we obtain by taking the limit $k \to \infty$ of the previous inequality and applying (4.39) twice

$$Q(\boldsymbol{\theta}') + \sigma^2 \leq Q(\boldsymbol{\theta}_0) + \sigma^2,$$

which gives $Q(\boldsymbol{\theta}') \leq Q(\boldsymbol{\theta}_0)$. By **A.6** we must have $\boldsymbol{\theta}' = \boldsymbol{\theta}_0$.

Now we argue that $(\hat{\boldsymbol{\theta}}_n)$ converges to $\boldsymbol{\theta}_0$ with probability tending to one. It is clear by the discussion above that if it converges it must converge to $\boldsymbol{\theta}_0$. Suppose $(\hat{\boldsymbol{\theta}}_n)$ does not converge. There exists a neighbourhood $\mathcal{V}$ of $\boldsymbol{\theta}_0$ such that $\mathcal{V}$ does not contain infinitely many functions from the sequence $(\hat{\boldsymbol{\theta}}_n)$. These infinitely many functions cannot have a subsequence which converges to $\boldsymbol{\theta}_0$, which is a contradiction. We have shown the desired result

$$\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\|_\infty = o_P(1)$$

■

The following lemma is based on Theorem 4 of Jennrich (1969) and is adapted to fit our setting. Let $\mathcal{F} \subset C(\mathcal{H})$ denote a compact subspace.

**Lemma 4.2.** *Assume $\boldsymbol{\varepsilon}_i$ is i.i.d. and the conditions **A.5**,**A.7** hold. Define*

$$p_n(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) := \frac{1}{n}\sum_{i=1}^{n} \left(\mathbf{B}(\mathbf{t}_i)(\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_1) - \hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_2))\right)^2$$

*which is a sequence of continuous functions in* $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \in \mathcal{F} \times \mathcal{F}$ *(consequence of* ***A.5****). We have the convergence*

$$h_n(\boldsymbol{\theta}) := \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \mathbf{B}(\mathbf{t}_i)(\hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta}_0) - \hat{\boldsymbol{\alpha}}_n(\boldsymbol{\theta})) \xrightarrow{P} 0$$

*uniformly over* $\boldsymbol{\theta} \in \mathcal{F}$*.*

*Proof.* By **A.5** the sequence $|p_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)|$ converges uniformly in $\boldsymbol{\theta}_1 \in C(\mathcal{H})$ (with probability tending to one) for any $\boldsymbol{\theta}_2 \in C(\mathcal{H})$ to a continuous function in $\boldsymbol{\theta}_1$. Hence there exists a neighbourhood $\mathcal{V}$ of $\boldsymbol{\theta}_2$ such that (with probability tending to one) for all $\boldsymbol{\theta}_1 \in \mathcal{V}$

$$|p_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)| \leq \delta \tag{4.40}$$

for a given $\delta > 0$, when $n$ is sufficiently large. Then, consider the inequality

$$|h_n(\boldsymbol{\theta}_1)| \leq \sqrt{p_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}\|\boldsymbol{\varepsilon}\|_2/\sqrt{n} + |h_n(\boldsymbol{\theta}_2)| \tag{4.41}$$

which follows from the triangle and the Cauchy-Schwarz inequalities. By **A.7**,

$$h_n(\boldsymbol{\theta}_2) \xrightarrow{P} 0.$$

Thus by inequalities (4.40)-(4.41) and the law of large numbers,

$$\sup_{\boldsymbol{\theta}_1 \in \mathcal{V}} |h_n(\boldsymbol{\theta}_1)| = O_P(\sqrt{\delta}\sigma + \delta).$$

Therefore $C(\mathcal{H})$ is covered by neighbourhoods $\mathcal{V}$ such that

$$\sup_{\boldsymbol{\theta} \in \mathcal{V}} |h_n(\boldsymbol{\theta})| = O_P(\sqrt{\delta}\sigma + \delta)$$

. This leads to such a finite subcover of the compact subspace $\mathcal{F}$ so that

$$\sup_{\boldsymbol{\theta} \in \mathcal{F}} |h_n(\boldsymbol{\theta})| = O_P(\sqrt{\delta}\sigma + \delta).$$

Since $\delta > 0$ is taken arbitrary we have shown the desired result

$$|h_n(\boldsymbol{\theta})| \xrightarrow{P} 0$$

uniformly in $\boldsymbol{\theta} \in \mathcal{F}$. ∎

## 4.8 Heat Example: Two-step method computation

In this section we derive the expressions for the estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ for the two-step method applied to the heat diffusion problem described in Section 4.5.

The method consists of minimizing

$$\begin{aligned} J(\boldsymbol{\alpha} \mid \boldsymbol{\gamma}) = \|\mathbf{Y} - \mathcal{B}\boldsymbol{\alpha}\|^2 + \lambda \left(\boldsymbol{\alpha}^\top \mathbf{R}\boldsymbol{\alpha} + 2\boldsymbol{\gamma}^\top \mathbf{T}\boldsymbol{\alpha} + \boldsymbol{\gamma}^\top \mathbf{Z}\boldsymbol{\gamma}\right) \\ + \mu \left(\|\mathbf{V}_1\boldsymbol{\alpha}\|^2 + \|\mathbf{V}_2\boldsymbol{\alpha}\|^2 + \|\mathbf{V}_3\boldsymbol{\alpha}\|^2\right) \end{aligned} \tag{4.42}$$

with respect to $\boldsymbol{\alpha}$, which yields $\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma})$, and

$$H(\boldsymbol{\gamma}) = \|\mathbf{Y} - \mathcal{B}\hat{\boldsymbol{\alpha}}\left(\boldsymbol{\gamma}\right)\|^2 + \lambda_1 \|\mathbf{P}_1\boldsymbol{\gamma}\|^2 + \lambda_2 \|\mathbf{P}_2\boldsymbol{\gamma}\|^2 \tag{4.43}$$

with respect to $\boldsymbol{\gamma}$.

From (4.42) it now follows that

$$\frac{\partial J}{\partial \boldsymbol{\alpha}} = -2\left(\mathcal{B}^\top \mathbf{Y} - \mathcal{B}^\top \mathcal{B}\boldsymbol{\alpha}\right)^\top + 2\lambda\left(\boldsymbol{\alpha}^\top \mathbf{R} + \boldsymbol{\gamma}^\top \mathbf{T}\right) + 2\mu\boldsymbol{\alpha}^\top \left(\mathbf{V}_1^\top \mathbf{V}_1 + \mathbf{V}_2^\top \mathbf{V}_2 + \mathbf{V}_3^\top \mathbf{V}_3\right).$$

This is equal to zero if and only if

$$\hat{\boldsymbol{\alpha}}\left(\boldsymbol{\gamma}\right) = \mathcal{D}^{-1}\left(\mathcal{B}^\top \mathbf{Y} - \lambda \mathbf{T}^\top \boldsymbol{\gamma}\right),$$

where $\mathcal{D} = \mathcal{B}^\top \mathcal{B} + \lambda \mathbf{R} + \mu\left(\mathbf{V}_1^\top \mathbf{V}_1 + \mathbf{V}_2^\top \mathbf{V}_2 + \mathbf{V}_3^\top \mathbf{V}_3\right)$. Note that $\mathcal{D} = \mathcal{D}^\top$ and

$$\frac{\partial \hat{\boldsymbol{\alpha}}}{\partial \boldsymbol{\gamma}} = -\lambda \mathcal{D}^{-1}\mathbf{T}^\top.$$

From (4.43) it now follows that

$$\frac{\partial H}{\partial \boldsymbol{\gamma}} = -2\left(\mathbf{Y} - \mathcal{B}\hat{\boldsymbol{\alpha}}\left(\boldsymbol{\gamma}\right)\right)^\top \mathcal{B}\frac{\partial \hat{\boldsymbol{\alpha}}}{\partial \boldsymbol{\gamma}} + 2\boldsymbol{\gamma}^\top \left(\lambda_1 \mathbf{P}_1^\top \mathbf{P}_1 + \lambda_2 \mathbf{P}_2^\top \mathbf{P}_2\right).$$

Substituting $\hat{\boldsymbol{\alpha}}\left(\boldsymbol{\gamma}\right)$ and its derivative into this equation, we find that this is equal to zero if and only if

$$\left(\lambda \mathbf{T}\mathcal{D}^{-1}\mathcal{B}^\top \mathcal{B}\mathcal{D}^{-1}\mathbf{T}^\top + \frac{\lambda_1}{\lambda}\mathbf{P}_1^\top \mathbf{P}_1 + \frac{\lambda_2}{\lambda}\mathbf{P}_2^\top \mathbf{P}_2\right)\boldsymbol{\gamma} = \left(\mathbf{T}\mathcal{D}^{-1}\mathcal{B}^\top \mathcal{B} - \mathbf{T}\right)\mathcal{D}^{-1}\mathcal{B}^\top \mathbf{Y}.$$

## 4.9 Heat Example: One-step method computation

In this section we derive the expressions for the estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ for the one-step method applied to the heat diffusion problem discribed in Section 4.5.

The method consist of minimizing:

$$\begin{aligned} K(\boldsymbol{\alpha},\boldsymbol{\gamma}) &= \|\mathbf{Y}-\mathcal{B}\boldsymbol{\alpha}\|^2 + \lambda\left(\boldsymbol{\alpha}^\top\mathbf{R}\boldsymbol{\alpha} + 2\boldsymbol{\gamma}^\top\mathbf{T}\boldsymbol{\alpha} + \boldsymbol{\gamma}^\top\mathbf{Z}\boldsymbol{\gamma}\right) \\ &\quad + \mu\left(\|\mathbf{V}_1\boldsymbol{\alpha}\|^2 + \|\mathbf{V}_2\boldsymbol{\alpha}\|^2 + \|\mathbf{V}_3\boldsymbol{\alpha}\|^2\right) + \lambda_1\|\mathbf{P}_1\boldsymbol{\gamma}\|^2 + \lambda_2\|\mathbf{P}_2\boldsymbol{\gamma}\|^2 \end{aligned} \tag{4.44}$$

with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$.

From (4.44) it follows that

$$\begin{cases} \dfrac{\partial K}{\partial \boldsymbol{\alpha}} = -2\left(\mathcal{B}^\top\mathbf{Y} - \mathcal{B}^\top\mathcal{B}\boldsymbol{\alpha}\right)^\top + 2\lambda\left(\boldsymbol{\alpha}^\top\mathbf{R} + \boldsymbol{\gamma}^\top\mathbf{T}\right) + 2\mu\boldsymbol{\alpha}^\top\left(\mathbf{V}_1^\top\mathbf{V}_1 + \mathbf{V}_2^\top\mathbf{V}_2 + \mathbf{V}_3^\top\mathbf{V}_3\right) \\ \dfrac{\partial K}{\partial \gamma} = 2\lambda\left(\boldsymbol{\alpha}^\top\mathbf{T}^\top + \boldsymbol{\gamma}^\top\mathbf{Z}\right) + 2\boldsymbol{\gamma}^\top\left(\lambda_1\mathbf{P}_1^\top\mathbf{P}_1 + \lambda_2\mathbf{P}_2^\top\mathbf{P}_2\right). \end{cases}$$

This implies that $\nabla K$ is equal to zero if and only if

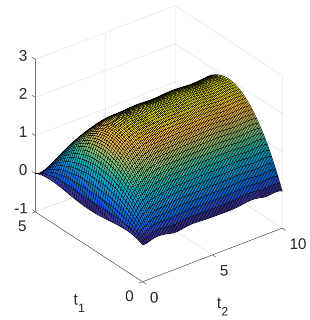
$$\begin{aligned} &\begin{cases} \mathcal{D}\boldsymbol{\alpha} + \lambda\mathbf{T}^\top\boldsymbol{\gamma} = \mathcal{B}^\top\mathbf{Y} \\ \lambda\mathbf{T}\boldsymbol{\alpha} + \left(\lambda\mathbf{Z} + \lambda_1\mathbf{P}_1^\top\mathbf{P}_1 + \lambda_2\mathbf{P}_2^\top\mathbf{P}_2\right)\boldsymbol{\gamma} = 0 \end{cases} \\ \Leftrightarrow &\begin{bmatrix} \mathcal{D} & \lambda\mathbf{T}^\top \\ \lambda\mathbf{T} & \left(\lambda\mathbf{Z} + \lambda_1\mathbf{P}_1^\top\mathbf{P}_1 + \lambda_2\mathbf{P}_2^\top\mathbf{P}_2\right) \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} \mathcal{B}^\top\mathbf{Y} \\ 0 \end{bmatrix}, \end{aligned}$$

where $\mathcal{D}$ is the same as in Section 4.8. Moreover, it is easy to see that we find the same expression for the estimator of $\boldsymbol{\alpha}$ as in the two-step method, but that the estimator of $\boldsymbol{\gamma}$ is different.

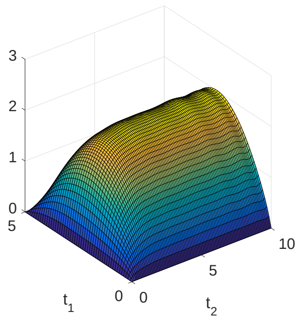
## 4.10 Figures

In all the images below, the first row of images contains the estimates of the state variable $g$ and the second row the corresponding estimates of $\theta$ for $\sigma = 0.1$. The columns are as follows:

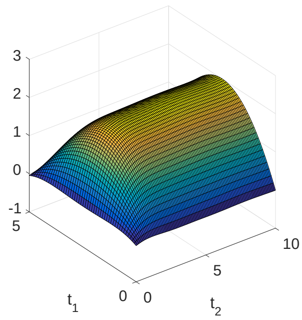
- Column 1: the two-step method using the full basis.
- Column 2: the two-step method using the reduced basis.
- Column 3: the one-step method using the full basis.
- Column 4: the one-step method using the reduced basis.



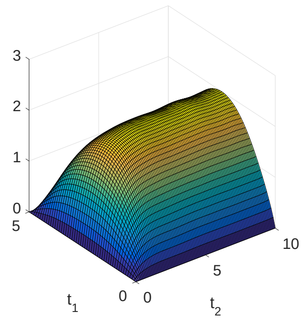

(a) 6.59e−3



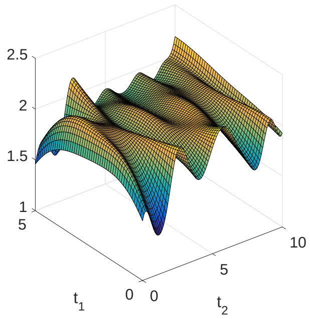

(b) 2.00e−2



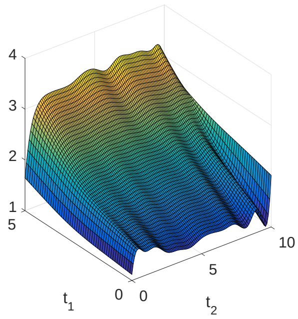

(c) 8.43e−3



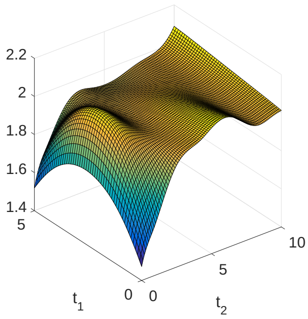

(d) 8.34e−3



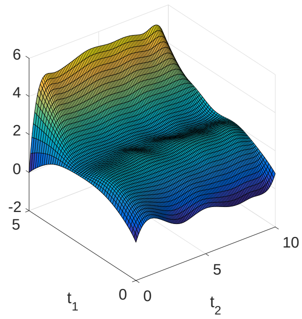

(e) 4.96e−2




(f) 2.72e−1

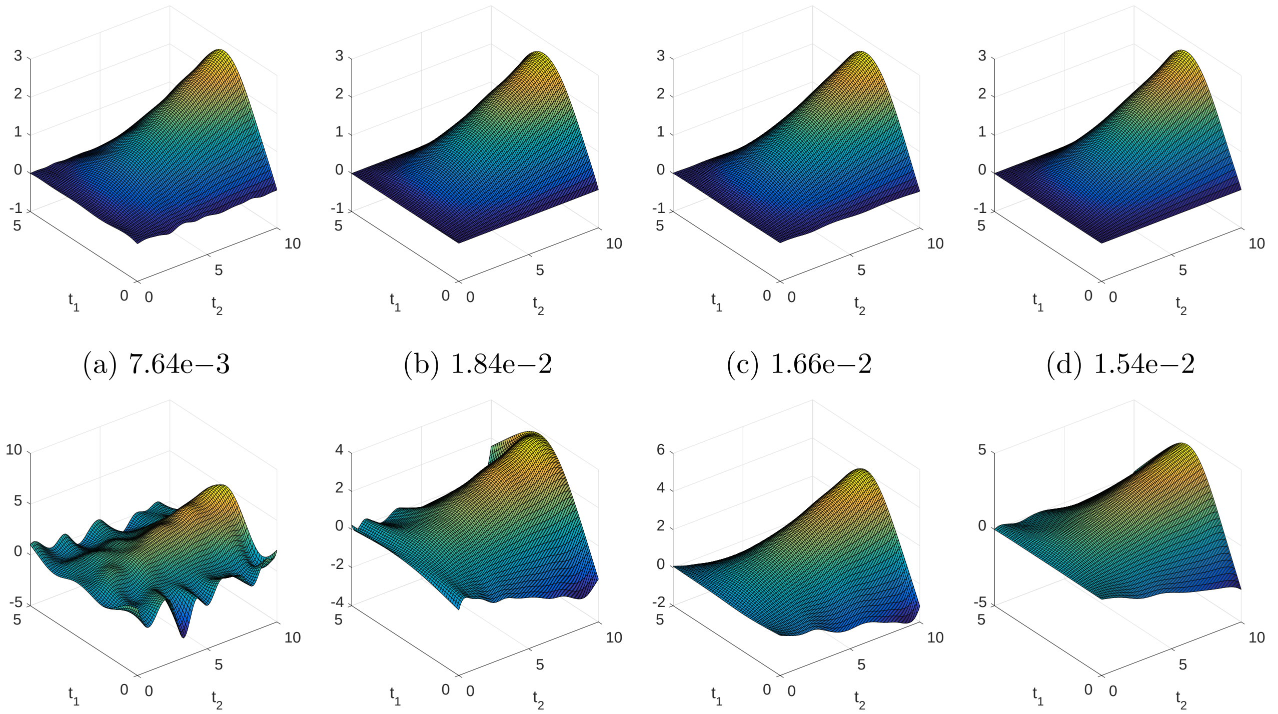



(g) 3.65e−2




(h) 5.31e−1

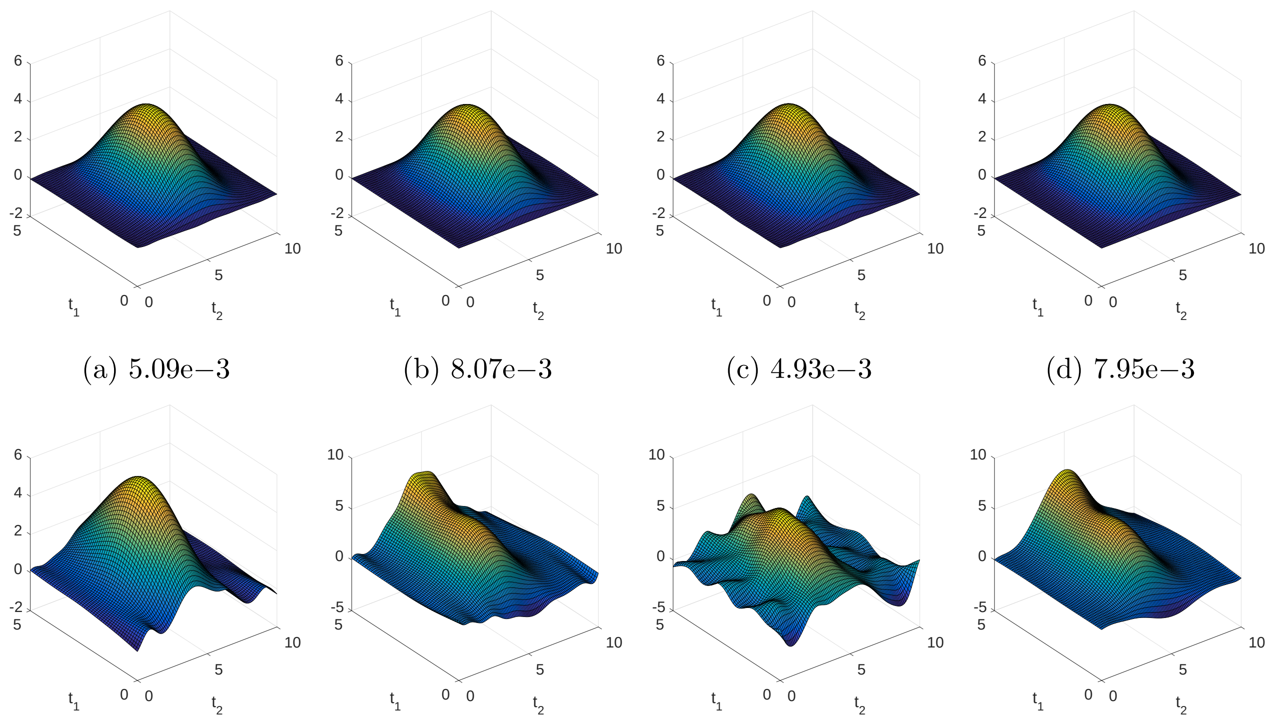
Figure 4.3: Example of the estimations found when using $\theta_1$ as the source function for the heat equation. The numbers mentioned below each figure are the relative error with respect to the exact solution.



(a) 7.64e−3 (b) 1.84e−2 (c) 1.66e−2 (d) 1.54e−2

(e) 2.00e−1 (f) 4.08e−1 (g) 3.38e−1 (h) 4.00e−1

Figure 4.4: Example of the estimations found when using $\theta_2$ as the source function for the heat equation. The numbers mentioned below each figure are the relative error with respect to the exact solution.




(a) 5.09e−3 (b) 8.07e−3 (c) 4.93e−3 (d) 7.95e−3

(e) 3.60e−1 (f) 4.93e−1 (g) 3.59e−1 (h) 4.92e−1

Figure 4.5: Example of the estimations found when using $\theta_3$ as the source function for the heat equation. The numbers mentioned below each figure are the relative error with respect to the exact solution.

# Chapter 5

# General conclusions and research perspectives

In this thesis attention is devoted to estimating unknown functions by splines and their inferences. We focused on varying coefficient models (VCMs) and PDE models.

In the classic linear regression model the coefficients are constants. For many data examples this parametric assumption is not likely. VCMs are an extension of the linear regression model, i.e. the coefficients are allowed to be functions, maintaining interpretability and increasing flexibility. In linear regression models F-tests (which are ratios of quadratic forms) are used for inferences on the coefficients. In Chapter 2 we generalized the F-test approach to longitudinal data VCMs by using the nice properties of B-splines. We showed that the test statistic follows (asymptotically) a generalized F distribution. As such we can test whether the coefficient functions have a parametric form, for example constant or linear. We conducted an extensive simulation study to show the good performances of our generalized F-tests, and gave a data application. In the context of generalized varying coefficient models (GVCMs) we provided a bootstrap approach to test similar hypotheses.

Chapter 3 dealt with monotonicity and convexity testing in VCMs. We used a testing approach which has proven to be very effective in univariate regression models and extended it to VCMs, with theoretical foundations. A broad simulated study is given where we also included testing simultaneously monotonicity and convexity of coefficient functions. Moreover, we gave data applications.

Our testing methods could be applied to other VCMs such as quantile estimation models (Andriyana et al. (2014)), recurrent event data models (Eshaghi et al. (2016)) and time series data models (Huang and Shen (2004)). A challenging hypothesis to study in the context of VCMs is whether a coefficient function is piecewise constant. The objective of piecewise testing is to discover abrupt changes, this is also termed change point detection. Such hypotheses are popular in time series data. There is statistical literature on how to estimate the jump locations of a piecewise constant function in several models (Lebarbier (2005) and Kolar et al. (2009) among others). It would be interesting to study such hypotheses in VCMs.

In Chapter 4 we studied linear PDE models. PDEs are used to describe a vast number of dynamic processes, they appear in scientific fields such as biology, physics, finance, etc. PDEs are determined by their parameters, which are often unknown and of interest for scientists. In the statistical literature it as assumed that these parameters are constant, this can be a crude assumption in practice. In Chapter 4 we showed how to model and estimate multivariate parameters of linear PDEs. We demonstrated that our estimators are uniformly consistent. As an illustration we simulated the temperature of a rod which is heating up and cooling down due to an unknown bivariate source term which we estimated from error prone temperature measurements of the rod. We found that our approach managed to estimate the unknown source term accurately.

Further research in this context includes variance estimation of a multivariate parameter estimate $\hat{\theta}$, i.e. estimate $\text{Var}(\hat{\theta}(\mathbf{t}))$ for a point $\mathbf{t}$ in the domain of $\theta$. This boils down to estimating the variance of the tensor product spline coefficients of $\hat{\theta}(\mathbf{t})$ and would allow to perform statistical inference of the unknown parameter $\boldsymbol{\theta}$. One approach for variance estimation was suggested by Rodriguez-Fernandez et al. (2006) and Xue et al. (2010), by using the pseudo-information matrix given by the inverse hessian matrix of (4.7). Frasso et al. (2015) found that this yields a reasonable approximation of the covariance matrix of constant PDE parameters. A second approach which is also suggested by Xue et al. (2010) is the weighted bootstrap method (Ma and Kosorok (2005)). This approach suggests to solve (4.7) repeatedly by adding iid weights with mean zero and unit variance.

We mainly focused on linear PDEs, our approach can be extended to include nonlinear PDE models by employing numerical optimization techniques. Note that the term which ensures a good fit of the measured data, i.e. the term sum of squared residuals in (4.6) and (4.7), should be replaced by the log-likelihood to include nonnormal error structures. We could use a Gauss-Newton method as in Ramsay et al. (2007), who estimate constant parameters in ODE models by a two-step method, to obtain the solutions of minimizing (4.6) and (4.7). Inference in PDE models did not receive considerable attention in the literature. Methods for hypothesis testing of the multivariate parameters could be proposed by using the experiences from the VCMs setting. One could test whether a parameter is time independent, constant, etc. which can help to fine tune PDE equations.

# Appendix A

# Notation

1. For a real matrix $\mathbf{A} \in I\!R^{n_1 \times n_2}$, $\|\mathbf{A}\|_2$ denotes the Frobenius norm: $\|\mathbf{A}\|_2 = \sqrt{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \mathbf{A}_{ij}^2}$. The norm $\|\cdot\|_\infty$ is defined by $\|\mathbf{A}\|_\infty = \max_{i=1,\ldots,n_1} \sum_{j=1}^{n_2} |\mathbf{A}_{ij}|$.

2. For a real valued function $h$ on $\mathcal{T}$, $||h||_\infty = \sup_{t\in\mathcal{T}} |h(t)|$ denotes its supremum norm, while for a real vector valued function $\mathbf{h} = (h_1, \ldots, h_m)^\top$, we let its supremum norm be $||\mathbf{h}||_\infty = \max_{1\leqslant i\leqslant m} ||h_i||_\infty$.

3. Let $\mathcal{G} = \mathbb{G}_0 \times \ldots \times \mathbb{G}_d$. Define the function $\mathbf{g}^*(t) = (g_0^*(t), \ldots, g_d^*(t))^\top$ such that $||\boldsymbol{\beta} - \mathbf{g}^*||_\infty = \rho_n = \inf_{\mathbf{g}\in\mathcal{G}} ||\boldsymbol{\beta} - \mathbf{g}||_\infty$. Let $\boldsymbol{\alpha}^*$ denote the corresponding coefficient vector, i.e. $\mathbf{g}^*(t) = \mathbf{B}(t)\boldsymbol{\alpha}^*$.

# Bibliography


Ahkim, M., Gijbels, I. and Verhasselt, A. (2016), 'Shape testing in varying coefficient models', *TEST* . DOI:10.1007/s11749-016-0518-y.

Ahkim, M., Schenkels, N., Vanroose, W. and Verhasselt, A. (2017), 'Estimating multivariate parameters in pde models', *Manuscript* .

Ahkim, M. and Verhasselt, A. (2017), 'Testing for constancy in varying coefficient models', *Communications in Statistics - Theory and Methods* . DOI:10.1080/03610926.2017.1300271.

Andriyana, Y., Gijbels, I. and Verhasselt, A. (2014), 'P-splines quantile regression estimation in varying coefficient models', *TEST* **23**(1), 153–194.

Antoniadis, A., Gijbels, I. and Verhasselt, A. (2012), 'Variable selection in varying coefficient models using p-splines', *Journal of Computational and Graphical Statistics* **21**(3), 638–661.

Bär, M., Hegger, R. and Kantz, H. (1999), 'Fitting partial differential equations to space-time dynamics', *Physical Review E* **59**(1), 336–341.

Bartlett, M. S. (1951), 'An inverse matrix adjustment arising in discriminant analysis', *The Annals of Mathematical Statistics.* **22**(1), 107–111.

Bausch, J. (2013), 'On the Efficient Calculation of a Linear Combination of Chi-Square Random Variables with an Application in Counting String Vacua', *Journal of Physics A: Mathematical and Theoretical* **46**(1), 505202.

Bowman, A. W., Jones, M. C. and Gijbels, I. (1998), 'Testing monotonicity of regression', *Journal of Computational and Graphical Statistics* **7**(4), 489–500.

Burden, R. L. and Faires, J. D. (2005), *Numerical Analysis*, Thompson.

Cai, Z., Fan, J. and Li, R. (2000), 'Efficient estimation and inferences for varying-coefficient models', *Journal of the American Statistical Association* **95**(451), 888–902.

Daouia, A., Noh, H. and Park, B. U. (2016), 'Data envelope fitting with constrained polynomial splines', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **78**(1), 3–30.

De Boor, C. (2001), *A Practical Guide to Splines*, number v. 27 *in* 'Applied Mathematical Sciences', Springer.

Diggle, P. J. (1988), 'An approach to the analysis of repeated measurements', *Biometrics* **44**(4), 959–971.

Diggle, P. and Verbyla, A. (1998), 'Non-parametric estimation of covariance structure in longitudinal data', *Biometrics* **54**, 401–415.

Dunkl, C. F. and Ramirez, D. E. (2001), 'Computation of the generalized f distribution', *Australian & New Zealand Journal of Statistics* **43**(1), 21–31.

Eilers, P. H. and Marx, B. D. (1996), 'Flexible smoothing with b-splines and penalties', *Statistical science* **11**(2), 89–102.

Emmermann, R. and Lauterjung, J. (1997), 'The german continental deep drilling program ktb: Overview and major results', *Journal of Geophysical Research: Solid Earth* **102**(B8), 18179–18201.

Eshaghi, E., Baghishani, H. and Shahsavani, D. (2016), 'Time-varying coefficients models for recurrent event data when different varying coefficients admit different degrees of smoothness: application to heart disease modeling', *Statistics in Medicine* **35**(23), 4166–4182.

Fan, J. and Wenyang, Z. (2008), 'Statistical methods with varying coefficient models', *Statistics and its interface* **1**(1), 179–195.

Fan, J. and Zhang, W. (1999), 'Statistical estimation in varying coefficient models', *The Annals of Statistics* **27**(5), 1491–1518.

Fleming, T. R. and Harrington, D. P. (1991), *Counting Processes and Survival Analysis*, John Wiley & Sons.

Frasso, G., Jaeger, J. and Lambert, P. (2015), 'Parameter estimation and inference in dynamic systems described by linear partial differential equations', *AStA Advances in Statistical Analysis* **100**(3), 1–29.

Frasso, G., Jaeger, J. and Lambert, P. (2016), 'Inference in dynamic systems using b-splines and quasilinearized ode penalties', *Biometrical Journal* **58**(3), 691–714.

Ghosal, S., Sen, A. and van der Vaart, A. W. (2000), 'Testing monotonicity of regression', *The Annals of Statistics* **28**(4), 1054–1082.

Haberman, R. (2004), *Applied Partial Differential Equations with Fourier Series and Boundary Value Problems*, Pearson Education.

Hartung, N., Mollard, S., Barbolosi, D., Benabdallah, A., Chapuisat, G., Henry, G., Giacometti, S., Iliadis, A., Ciccolini, J., Faivre, C. and Hubert, F. (2014), 'Mathematical modeling of tumor growth and metastatic spreading: Validation in tumor-bearing mice', *Cancer Research* **74**(22), 6397–6407.

Hastie, T. and Tibshirani, R. (1993), 'Varying-coefficient models', *Journal of the Royal Statistical Society. Series B (Methodological)* **55**(4), 757–796.

Hedeker, D. and Gibbons, R. D. (1997), 'Application of random-effects pattern-mixture models for missing data in longitudinal studies', *Psychological Methods* **2**(6), 64–78.

Höllig, K. (2003), *Finite element methods with B-splines*, SIAM Frontiers in Applied Mathematics.

Hong, Z. and Lian, H. (2012), 'Time-varying coefficient estimation in differential equation models with noisy time-varying covariates', *Journal of Multivariate Analysis* **103**(1), 58 – 67.

Hoover, D. R., Rice, J. A., Wu, C. O. and Yang, L.-P. (1998), 'Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data', *Biometrika* **85**(4), 809–822.

Horn, R. A. and Johnson, C. R. (1988), *Matrix Analysis*, Cambridge University Press.

Huang, J. Z. (1998), 'Projection estimation in multiple regression with application to functional anova models', *The Annals of Statistics* **26**(1), 242–272.

Huang, J. Z. and Shen, H. (2004), 'Functional coefficient regression models for non-linear time series: A polynomial spline approach', *Scandinavian Journal of Statistics* **31**(4), 515–534.

Huang, J. Z., Wu, C. O. and Zhou, L. (2002), 'Varying-coefficient models and basis function approximations for the analysis of repeated measurements', *Biometrika* **89**(1), 111–128.

Huang, J. Z., Wu, C. O. and Zhou, L. (2004), 'Polynomial spline estimation and inference for varying coefficient models with longitudinal data', *Statistica Sinica* **14**(3), 763–788.

Jennrich, R. I. (1969), 'Asymptotic properties of non-linear least squares estimators', *The Annals of Mathematical Statistics* **40**(2), 633–643.

Jianjun, Y., Liu, L. and Hu, Q. (2013), 'Mathematical modeling of brain glioma growth using modified reaction-diffusion equation on brain mr images', *Computers in Biology and Medicine* **43**(12), 2007–2013.

Karlin, S. and Studden, W. J. (1966), 'Optimal experimental designs', *The Annals of Mathematical Statistics* **37**(4), 783–815.

Kaslow, R., Ostrow, D., Detel, R., Phair, J., Polk, B. and Rinaldo, C. (1987), 'The Multicenter AIDS Cohort Study : rationale, organization, and selected characteristics of the participants', *American Journal of Epidemiology* **126**, 310–318.

Kolar, M., Song, L. and Xing, E. P. (2009), 'Sparsistent learning of varying-coefficient models with structural changes', pp. 1006–1014.

Lammers, W. J., van Buuren, H. R., Hirschfield, G. M., Janssen, H. L., Invernizzi, P., Mason, A. L., Ponsioen, C. Y., Floreani, A., Corpechot, C., Mayo, M. J., Battezzati, P. M., Pars, A., Nevens, F., Burroughs, A. K., Kowdley, K. V., Trivedi, P. J., Kumagi, T., Cheung, A., Lleo, A., Imam, M. H., Boonstra, K., Cazzagon, N., Franceschet, I., Poupon, R., Caballeria, L., Pieri, G., Kanwar, P. S., Lindor, K. D. and Hansen, B. E. (2014), 'Levels of alkaline phosphatase and bilirubin are surrogate end points of outcomes of patients with primary biliary cirrhosis: An international follow-up study', *Gastroenterology* **147**(6), 1338 – 1349.

Lebarbier, E. (2005), 'Detecting multiple change-points in the mean of gaussian process by model selection', *Signal Processing* **85**(4), 717 – 736.

Li, N., Xu, X. and Liu, X. (2011), 'Testing the constancy in varying-coefficient regression models', *Metrika* **74**(3), 409–438.

Lorr, M. and Klett, C. J. (1966), *Inpatient Multidimensional Psychiatric Scale: Manual.*, Palo Alto, CA: Consulting Psychologists Press.

Ma, S. and Kosorok, M. R. (2005), 'Robust semiparametric m-estimation and the weighted bootstrap', *Journal of Multivariate Analysis* **96**(1), 190 – 217.

Marx, B. D. and Eilers, P. H. (2005), 'Multidimensional penalized signal regression', *Technometrics* **47**(1), 13–22.

McCullagh, P. and Nelder, J. A. (1989), *Generalized linear models (Second edition)*, London: Chapman & Hall.

Müller, T. and Timmer, J. (2002), 'Fitting parameters in partial differential equations from partially observed noisy data', *Physica D: Nonlinear Phenomena* **171**(12), 1 – 7.

Müller, T. and Timmer, J. (2004), 'Parameter identification techniques for partial differential equations', *International Journal of Bifurcation and Chaos* **14**(6), 2053–2060.

Ramsay, J. O., Hooker, G., Campbell, D. and Cao, J. (2007), 'Parameter estimation for differential equations: a generalized smoothing approach', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**(5), 741–796.

Rodriguez-Fernandez, M., Egea, J. A. and Banga, J. R. (2006), 'Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems', *BMC Bioinformatics* **7**(1), 483.

Sanches, R., Bornemann, F. and Cirak, F. (2011), 'Immersed b-spline (i-spline) finite element methods for geometrically complex domains', *Computer Methods in Applied Mechanics and Engineering* **200**(13-16), 1432–1445.

Schumaker, L. (2007), *Spline Functions: Basic Theory*, Cambridge University Press.

Tan, W. Y. (1977), 'On the distribution of quadratic forms in normal random variables', *Canadian Journal of Statistics* **5**(2), 241–250.

Verhasselt, A. (2014), 'Generalized varying coefficient models: A smooth variable selection technique', *Statistica Sinica* **24**(18), 147–171.

Wang, J. C. and Meyer, M. C. (2011), 'Testing the monotonicity or convexity of a function using regression splines', *Canadian Journal of Statistics* **39**, 89–107.

Wolkowicz, H. and Styan, G. P. (1980), 'More bounds for elgenvalues using traces', *Linear Algebra and its Applications* **31**(1), 1 – 17.

Xue, H., Miao, H. and Wu, H. (2010), 'Sieve estimation of constant and time-varying coefficients in nonlinear ordinary differential equation models by considering both numerical error and measurement error', *Ann. Statist.* **38**(4), 2351–2387.
**URL:** *http://dx.doi.org/10.1214/09-AOS784*

Xun, X., Cao, J., Mallick, B., Maity, A. and Carroll, R. J. (2013), 'Parameter estimation of partial differential equation models', *Journal of the American Statistical Association* **108**(503), 1009–1020.

Yoo, W. W. and Ghosal, S. (2016), 'Supremum norm posterior contraction and credible sets for nonparametric multivariate regression', *The Annals of Statistics* **44**(3), 1069–1102.

Yu, Y. and Ruppert, D. (2002), 'Penalized spline estimation for partially linear single-index models', *Journal of the American Statistical Association* **97**(460), 1042–1054.

Zhang, D. (2004), 'Generalized linear mixed models with varying coefficients for longitudinal data', *Biometrics* **60**(1), 8–15.

Zhang, H.-G. and Mei, C.-L. (2012), 'Sizer inference for varying coefficient models', *Communications in Statistics - Simulation and Computation* **41**(10), 1944–1959.

Zhang, H.-G., Mei, C.-L. and Wang, H.-L. (2013), 'Robust sizer approach for varying coefficient models', *Mathematical Problems in Engineering* **2013**(Article ID 547874), 13 pages.