PEDRO JOSÉ DOS SANTOS PALHINHAS MOTA

Brownian Motion with Drift Threshold Model

Dissertation presented to obtain the Degree of Doctor in Mathematics, Speciality of Statistics, from the New University of Lisbon, Faculty of Sciences and Technology

LISBOA

2008

Acknowledgements

I would like to express my sincere gratitude to:

- Professor Doctor Mathieu Kessler and Professor Doctor Manuel Esquível, my supervisors for their help, guidance, encouragement and friendship that I have always received from them.
- To my wife Inês for all the patience, love and support and to my two daughters Margarida and Mafalda.
- This work was supported in part by the European Community's Human Potential Programme under contract HPRN-CT-2000-00100, DYNSTOCH and by PRODEP III (medida 5 - Acção 5.3), for which I am truly thankful.

Sumário

Nesta tese propomo-nos a implementar procedimentos de estimação com vista a obter estimadores dos limiares para modelos com limiares definidos a partir de equações diferenciais estocásticas. O primeiro procedimento a ser apresentado baseia-se na adequação do algoritmo EM (expectation-maximization ou esperança e maximização) à estimação de limiares no modelo com limiares construido a partir do processo Browniano com tendência. O segundo procedimento, repete uma das idéias fundamentais na estimação de limiares no contexto de séries temporais, estimação de minimos quadrados, ou seja o procedimento que iremos adoptar será o de estimar os limiares pelos valores que minimizam a soma do quadrado dos erros. Iremos implementar este procedimento não só para modelos com limiares baseados no processo Browniano com tendência mas também para modelos genéricos entre os quais se destacam os que são baseados nos processos de Ornstein-Uhlenbeck e Browniano geométrico. Ambos os procedimentos são sujeitos a uma implementação prática aplicada a dados simulados, sendo ainda o procedimento de estimação por minimos quadrados aplicado a dados reais respeitantes a cotações diárias de um conjunto de fundos financeiros internacionais. O primeiro fundo é o fundo PF-European Sustainable Equities-R da Pictet Funds e o segundo o Parvest Europe Dynamic Growth fund do BNP Paribas. Os dados para ambos os fundos são os preços diários do ano 2004. O último fundo a ser considerado é o fundo Converging Europe Bond da Schroder e os dados são os preços diários do ano 2005.

Summary

In this thesis we implement estimating procedures in order to estimate threshold parameters for the continuous time threshold models driven by stochastic differential equations. The first procedure is based on the EM (expectation-maximization) algorithm applied to the threshold model built from the Brownian motion with drift process. The second procedure mimics one of the fundamental ideas in the estimation of the thresholds in time series context, that is, conditional least squares estimation. We implement this procedure not only for the threshold model built from the Brownian motion with drift process but also for more generic models as the ones built from the geometric Brownian motion or the Ornstein-Uhlenbeck process. Both procedures are implemented for simulated data and the least squares estimation procedure is also implemented for real data of daily prices from a set of international funds. The first fund is the PF-European Sustainable Equities-R fund from the Pictet Funds company and the second is the Parvest Europe Dynamic Growth fund from the BNP Paribas company. The data for both funds are daily prices from the year 2004. The last fund to be considered is the Converging Europe Bond fund from the Schroder company and the data are daily prices from the year 2005.

Symbols and Notation

- $a \wedge b = \min\{a, b\}$
- $a \vee b = \max\{a, b\}$
- Δ, Δ_n-discretization interval
- $\mathbb{I}_A(x) = \begin{cases} 1 & \text{if } , x \in A \\ 0 & \text{otherwise} \end{cases}$
- $mod(a, b)$-remainder of the integer division of a by b
- $\tau_1, \ldots, \tau_k, ...$-threshold hitting times
- $[x] = \max\{n \in \mathbb{N}; n \leq x\}$
- $X_1, \ldots, X_n$ observation of the process.
- BMD-Brownian motion with drift
- CLS-conditional least squares
- EM-expectation maximization
- GBM-geometric Brownian motion
- $MCEM$-monte carlo expectation maximization
- OLS-ordinary least squares
- OU-Ornstein Uhlenbeck
- $SETAR$-self exciting threshold autoregressive model
- TAR-threshold autoregressive model

Contents

List of Figures

List of Tables

Preface

The past decades have witnessed major developments in the field of statistical inference for diffusion processes and time series analysis. In the time series the assumptions of linearity and stationarity have been abandoned and the study of nonlinear models is increasing. One class of nonlinear models, called threshold models can be found in [**Tong, 1990**], in this class the most popular is the threshold autoregressive model (TAR), or $TAR(m, p)$, where the process is divided into m regimes following in each regime an AR model. Our goal is to extend the notion of threshold processes to continuous time processes and obtain estimation methods for this kind of processes. A diffusion which experiences a regime change upon crossing upper (M) and lower (m) levels will be our generic model for the stochastic process. We want to study diffusions where changes in the drift parameter have a consequence on the trend of the process, that is, we want to consider diffusions with positive trend for some drift parameter μ_1 and with negative trend for drift parameter μ_2. For instance, the Brownian motion with drift is a diffusion of the suggested type and the thresholds are introduced, in the model, in the following way. Let us consider two thresholds, m and M, and suppose that we start with the process with positive trend (driven by μ_1) and the process continues in this first regime until it hits the upper threshold M, at that time a change occurs and the process follows with negative trend (driven by μ_2) and continues in this second regime until it hits the lower threshold m, starting all over again in the first regime. We call the resulting process a continuous time threshold model. Simple diffusion processes are often used for stochastic modelling in many areas as physics, biology or economics. In many applications a continuous time threshold model can be more useful than simple diffusion model, for instance in some cooling (or heating) system controlled by a thermostat we can observe this kind of behavior in the temperature evolution see [**Molina & et al., 2004**], in a biological system where the animal population increase until hits a threshold value that makes the population dynamics to change (for instance the lack of food) and the population decreases until it hits a lower threshold where the dynamics change (when the food is enough to support a small population) and the population increases once again or in a financial context we

can expect the price of some asset to have an increasing or decreasing dynamic between two thresholds. All these models involve unknown parameters which need to be estimated from observations of the processes. Different methods of estimation as maximum likelihood, least squares or martingale estimating functions are well studied for diffusion models. However, for thresholds diffusions models there are not (as to our knowledge) similar results regarding the estimation of the threshold parameters.

CHAPTER 1

Overview on threshold model

1.1. Introduction

In the time series context the assumptions of linearity and stationarity have been abandoned and the study of nonlinear models is increasing. One class of nonlinear models, called threshold models can be found in [**Tong, 1990**], in this class the most popular is the threshold autoregressive model (TAR), or $TAR(m,p)$, where the process is divided into m regimes following in each regime an AR model. This model can be represented by the following equation:

$$Y_t = \sum_{i=1}^{m} \left(a_{i,0} + a_{i,1}Y_{t-1} + \cdots + a_{i,p}Y_{t-p} + \varepsilon_{i,t}\right) \mathbb{I}_{\{r_{i-1} \leq Z_{t-1} < r_i\}}, \tag{1.1}$$

where for each regime i the nonlinear model is just ordinary $AR(p)$ process. The thresholds are real numbers $-\infty = r_0 < r_1 < \ldots < r_{m-1} < r_m = +\infty$ and $Z_{t-1} = Z(Y_1, ..., Y_{t-1})$ is the threshold variable that specifies the change in regime.

We want to study diffusions where changes in the drift parameter have a consequence on the trend of the process, that is, we want to consider diffusions with positive trend for some drift parameter μ_1 and with negative trend for drift parameter μ_2. Considering two thresholds, m and M, and suppose that we start with the process with positive trend (driven by μ_1) and the process continues in this first regime until it hits the upper threshold M, at that time a change occurs and the process follows with negative trend (driven by μ_2) and continues in this second regime until it hits the lower threshold m, starting all over again in the first regime. This model can be specified by the following equation:

$$dXt = a(\mu(t), X_t)dt + b(\sigma, X_t)dB_t, \tag{1.2}$$

with

$$\mu(t) = \sum_{k \geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k}, \tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1}, \tau_{2k+2}[}(t)\right],$$

where $\tau_0 < \tau_1 < \ldots < \tau_j < \ldots$ are the threshold hitting times, that is, $\tau_{2k+1} = \inf\{t > \tau_{2k}; X_t = M\}$ and $\tau_{2k+2} = \inf\{t > \tau_{2k+1}; X_t = m\}$ for $k \geq 0$ and $\tau_0 = 0$.

This thesis is mainly concerned with the estimation of the threshold parameters, m and M, in the threshold diffusion process. The original contributions of this thesis are contained in the next chapters. The thesis is organized as follows. The present chapter continues with a review on the estimating procedures for the threshold models in the time series context and in the diffusion context. Chapter 2 starts with an overview and some results on hitting times for some diffusion processes and follows with the introduction of a Monte Carlo Expectation Maximization type algorithm to estimate the parameters of the Brownian motion with drift threshold model, from discrete observations of the process, finishing with an implementation of the algorithm. In chapter 3, we start implementing a conditional least squares estimation procedure for the Brownian motion with drift threshold model, with decreasing step Δ_n and knowing the regime for each observation, after that we implement the procedure with unknown regimes. Finally, we generalize the estimation procedure to more general threshold models as the ones built, for example, from the Ornstein-Uhlenbeck and the geometric Brownian motion processes. For all the estimating procedures presented, simulation studies are carried out and the results are presented. In the end the procedure is also implemented with real data. Overall conclusions are presented in the end, along with some general comments on the estimation methods and suggestions for possible extensions and future research.

1.2. Overview on TAR models

The objective of this overview is to look at some estimating procedures and relevant questions in time-series threshold models. The threshold autoregressive model (1.1) can be written as,

$$Y_t = \begin{cases} a_{1,0} + a_{1,1}Y_{t-1} + \cdots + a_{1,p}Y_{t-p} + \varepsilon_{1,t}, & \text{if } Z_t < r_1 \\ a_{2,0} + a_{2,1}Y_{t-1} + \cdots + a_{2,p}Y_{t-p} + \varepsilon_{2,t}, & \text{if } r_1 \le Z_t < r_2 \\ \cdots & \\ a_{m,0} + a_{m,1}Y_{t-1} + \cdots + a_{m,p}Y_{t-p} + \varepsilon_{m,t}, & \text{if } r_{m-1} \le Z_t \end{cases}$$

with p the order of the autoregression and m the number of regimes. One particular well known case is the self-exciting threshold autoregressive model ($SETAR$), where the threshold variable Z_t is replaced by some lagged value of Y_t that is $Z_t = Y_{t-d}$ and d is called the delay. For the threshold model, with $A_i = (a_{i,0}, a_{i,1}, ..., a_{i,p})$, for $i = 1, ..., m$, $r = (r_1, ..., r_{m-1})$ and $\theta = (A_1, ..., A_m, r)$, the usual estimation procedure is conditional least squares (CLS), that is, the CLS estimator $\hat{\theta}_N = (\hat{A}_{1,N}, ..., \hat{A}_{m,N}, \hat{r}_N)$ is the one that

minimizes the conditional sum of squared errors,

$$L_N(\theta) = \sum_{t=p}^{N} (Y_t - \mathbb{E}_\theta[Y_t|\mathcal{F}_{t-1}])^2,$$

with $(\mathcal{F}_t)$ the natural filtration and N the number of observations.

The minimization is done in two steps.

(1) For fixed r, we can get by ordinary least squares (OLS) values for $\hat{A}_{i,N}(r)$, $i = 1, ..., m$ and write

$$S_N(r) = \sum_{i=1}^{m} S_{i,N}(\hat{A}_{i,N}, r)$$

where $S_{i,N}(\hat{A}_{i,N}, r) = \sum_{t=p}^{N} \left((Y_t - \hat{A}_{i,N}|Y)^2 \mathbb{I}_{\{r_{i-1} \leq Z_t < r_i\}} \right)$, $.|.$ stands for inner product and $Y = (1, Y_{t-1}, ..., Y_{t-p})$.

(2) Then we choose $\hat{r}_N$ that minimizes $S_N(r)$, that is

$$\hat{r}_N = \arg\min_r S_N(r) \tag{1.3}$$

and finally we put

$$\hat{A}_{i,N} = \hat{A}_{i,N}(\hat{r}_N), i = 1, ..., m.$$

Usually, it is assumed that $r \in R$ with R bounded set, and in the $SETAR$ case the delay parameter d is introduced and a double grid (over d and r) is introduced in (1.3). For the case of $SETAR$ model with $m = 2$,

$$Y_t = \begin{cases} a_{1,0} + a_{1,1}Y_{t-1} + \cdots + a_{1,p}Y_{t-p} + \varepsilon_{1,t}, & \text{if } Y_{t-d} \leq r \\ a_{2,0} + a_{2,1}Y_{t-1} + \cdots + a_{2,p}Y_{t-p} + \varepsilon_{2,t}, & \text{if } Y_{t-d} > r \end{cases} \tag{1.4}$$

where $\varepsilon_{i,t} = c_i\varepsilon_t$ with ε_t i.i.d. zero mean and unity variance, using the same estimation procedure (CLS), fixing r and d and performing a grid search Chan, in [**Chan, 1993**], showed that when the autoregressive function is discontinuous, that is, when $\exists W^* = (1, w_{p-1}, w_{p-2}, ..., w_0)$, with $w_{p-d} = r$, such that $(A_1 - A_2)|W^* \neq 0$, the CLS estimator $\hat{\theta}_N$ of θ is consistent. Moreover, in the same paper, the author shows that $\hat{r}_N$ is N consistent and $N(\hat{r}_N - r)$ converges in distribution to M_-, where $[M_-, M_+[$ is the unique random interval over which a compound Poisson process attains its global minimum. Furthermore, $N(\hat{r}_N - r)$ is asymptotically independent of $\sqrt{N}(\hat{A}_1 - A_1, \hat{A}_2 - A_2)$ and the latter is asymptotically normal with the same distribution as that for the case when r is known. Still in the $SETAR(2, p)$ case, when the autoregressive function is continuous, that is, if $(A_1 - A_2)|W^* = 0$ for all $W^* = (1, w_{p-1}, w_{p-2}, ..., w_0)$, where $w_{p-d} = r$, or in a

equivalent way, $a_{1,i} = a_{2,i}$ for $1 \leq i \neq d \leq p$ and $a_{1,0} + ra_{1,d} = a_{2,0} + ra_{2,d}$, that is, when (1.4) can be written as

$$Y_t = a_0 + \sum_{j=1,\neq d}^{p} a_j Y_{t-j} + \begin{cases} a_{1,d}(Y_{t-d} - r) + \varepsilon_{1,t} & \text{if } Y_{t-d} \leq r \\ a_{2,d}(Y_{t-d} - r) + \varepsilon_{2,t} & \text{if } Y_{t-d} > r \end{cases}$$

where $a_0 = a_{1,0} + ra_{1,d}$ and $a_j = a_{1,j} = a_{2,j}$ for $j \neq d$, Chan and Tsay, in the paper [**Chan & Tsay, 1998**], show that the CLS estimator, $\hat{\theta}_N$ of θ, is yet consistent. However, $\hat{r}_N$ is now $\sqrt{N}$ consistent and

$$\sqrt{N}(\hat{\theta}_N - \theta) \text{ is asymptotically } N(0, U^{-1}VU),$$

where $U = \mathbb{E}[H_t H_t^T]$ and $V = \mathbb{E}[e_t^2 H_t H_t^T]$. With $e_t(\theta) = Y_t - \mathbb{E}_\theta[Y_t|\mathcal{F}_{t-1}]$ and

$$\begin{aligned} H_t(\theta) =& (-1, -Y_{t-1}, ..., -Y_{t-d+1}, -(Y_{t-d} - r)_-, -(Y_{t-d} - r)_+, ..., -Y_{t-p}, \\ & a_{1,d}\mathbb{I}_{\{Y_{t-d} \leq r\}} + a_{2,d}\mathbb{I}_{\{Y_{t-d} > r\}})^T \end{aligned}$$

the partial derivatives of $e_t(\theta)$ with respect to θ, where $(y)_- = \min(y, 0)$ and $(y)_+ = \max(y, 0)$.

Hansen in a sequence of papers involving threshold models, studies several models. For instance, in [**Hansen, 2000**] the author develops asymptotic theory for the distribution of the regression estimates from the regression model, when $m = 2$,

$$Y_t = \begin{cases} A_1^T X_t + \varepsilon_t, & Z_t \leq r \\ A_2^T X_t + \varepsilon_t, & Z_t > r \end{cases},$$

with $X_t = (1, X_{1,t}, X_{2,t}, ..., X_{p,t})$, and when the regression errors form a martingale difference sequence. The estimation procedure is the same (CLS) and the fundamental difference between the assumptions in [**Hansen, 2000**] and the other authors is that he makes the assumption that

$$A_2 - A_1 = cN^{-\alpha},$$

that is, a decreasing threshold effect, with $c \neq 0$ and $0 < \alpha < 1/2$. Getting,

$$N^{1-2\alpha}(\hat{r}_N - r) \xrightarrow{d} \omega T$$

where

$$\omega = \frac{c^T V c}{(c^T D c)^2 f} \quad \text{and} \quad T = \arg\max_{s \in \mathbb{R}} \left[-\frac{1}{2}|s| + W(s) \right].$$

With, $D = D(r)$ when $D(s) = \mathbb{E}[X_t X_t^T | Z_t = s]$, $V = V(r)$ when $V(s) = \mathbb{E}[X_t X_t^T \varepsilon_t^2 | Z_t = s]$, $f = f(r)$ when $f(z)$ is the density of Z_t and

$$W(s) = \begin{cases} W^1_{-s}, & s < 0 \\ 0, & s = 0 \\ W^2_s, & s > 0 \end{cases},$$

where W^1_u and W^2_u are independent standard Brownian motion. The distribution for T is known and given by,

$$\mathbb{P}[T \leq x] = 1 + \sqrt{\frac{x}{2\pi}} \exp\left(-\frac{x}{8}\right) + \frac{3}{2}\exp(x)\phi\left(-\frac{3\sqrt{x}}{3}\right) - \left(\frac{x+5}{2}\right)\phi\left(-\frac{\sqrt{x}}{2}\right)$$

when $x \geq 0$ and ϕ cumulative standard normal distribution function. For $x < 0$, $\quad \mathbb{P}[T \leq x] = 1 - \mathbb{P}[T \leq -x]$.

In [**Hansen, 2000**], the author also builds likelihood ratio tests to test the hypothesis $H_0 : r = r_0$ under the assumption that ε_t are i.i.d. $N(0, \sigma^2)$. For the drift parameters in the regression, he shows that, with $\hat{\theta} = (\hat{A}_1^T, \hat{A}_2^T)^T$ and $\theta = (A_1^T, A_2^T)^T$,

$$\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V_\theta),$$

where V_θ is the standard asymptotic covariance matrix as if r were fixed and known.

Earlier, in [**Hansen, 1997**] the author proved similar results about the asymptotic distribution of the threshold and the likelihood ratio test for the $TAR(2, p)$ model, he also reviews the test for linearity suggested in [**Hansen, 1996**] to test the linear model against the threshold model and the same kind of test is studied in [**Hansen, 1999**].

In the paper [**Hansen & Seo, 2002**] the authors provide an estimation procedure for threshold co-integration in vector error-correction models with two regimes,

$$\Delta Y_t = \begin{cases} A_1^T X_{t-1}(\beta) + \varepsilon_t, & w_{t-1}(\beta) \leq r \\ A_2^T X_{t-1}(\beta) + \varepsilon_t, & w_{t-1}(\beta) > r \end{cases}.$$

Where Y_t is p-dimensional time series, $I(1)$ (that is, each of the series is non-stationary with unit root), which is co-integrated with cointegration vector β (so $\beta^T Y_t$ is stationary or $I(0)$). Where $w_t(\beta) = \beta^T Y_t$, A is $k \times p$ matrix and

$$X_{t-1}(\beta) = (1, w_{t-1}(\beta), \Delta Y_{t-1}, ..., \Delta Y_{t-l}).$$

The errors ε_t are assumed to form a martingale difference sequence with finite covariance matrix $\Sigma = \mathbb{E}[\varepsilon_t \varepsilon_t^T]$. Under the assumption that the errors ε_t are i.i.d. Gaussian, the

authors implement maximum likelihood estimation (MLE) of the threshold model. The Gaussian likelihood is

$$\mathfrak{L}_N(A_1, A_2, \Sigma, \beta, r) = -\frac{N}{2}\log|\Sigma| - \frac{1}{2}\sum_{t=1}^{N}\varepsilon_t(A_1, A_2, \Sigma, \beta, r)^T\Sigma^{-1}\varepsilon_t(A_1, A_2, \Sigma, \beta, r),$$

where

$$\varepsilon_t(A_1, A_2, \Sigma, \beta, r) = \Delta Y_t - A_1^T X_{t-1}(\beta)\mathbb{I}_{\{w_{t-1}(\beta)\leq r\}} - A_2^T X_{t-1}(\beta)\mathbb{I}_{\{w_{t-1}(\beta)>r\}}.$$

The MLE $(\hat{A}_1, \hat{A}_2, \hat{\Sigma}, \hat{\beta}, \hat{r})$ are the values that maximize $\mathfrak{L}_N(A_1, A_2, \Sigma, \beta, r)$, and the proposed estimation procedure is the following:

(1) Form a grid for β and r.
(2) For each value of β and r in the grid, calculate $\hat{A}_1(\beta, r)$, $\hat{A}_2(\beta, r)$ and $\hat{\Sigma}(\beta, r)$ as the constrained MLE for (A_1, A_2, Σ) (it coincides with OLS estimation).
(3) Find $(\hat{\beta}, \hat{r})$ as the values of (β, r) which yield the lowest value of

$$\log|\hat{\Sigma}(\beta, r)|$$

this is because the likelihood function is

$$\mathfrak{L}_N(\beta, r) = \mathfrak{L}_N(\hat{A}_1(\beta, r), \hat{A}_2(\beta, r), \hat{\Sigma}(\beta, r), \beta, r) = -\frac{N}{2}\log|\hat{\Sigma}(\beta, r)| - \frac{Np}{2}.$$

(4) Finally, set

$$\hat{\Sigma} = \hat{\Sigma}(\hat{\beta}, \hat{r}), \quad \hat{A}_1 = \hat{A}_1(\hat{\beta}, \hat{r}) \text{ and } \hat{A}_2 = \hat{A}_2(\hat{\beta}, \hat{r}).$$

Another process is studied by Tsay, in [**Tsay, 1998**], where the author considers a multivariate (open-loop) threshold regression model,

$$y_t = c_j + \sum_{i=1}^{p}\phi_i^j y_{t-i} + \sum_{i=1}^{q}\beta_i^j x_{t-i} + \varepsilon_t^j, \text{ if } r_{j-1} < z_{t-d} \leq r_j \tag{1.5}$$

with $j = 1, ..., s$, c_j constant vectors, y and x are k and v dimensional, respectively. Given observations $\{y_t, x_t, z_t\}$, $t = 1, ..., N$ the author provides a test to detect the threshold nonlinearity of y_t. Writing the linear model in a regression framework,

$$y_t^T = X_t^T\phi + \varepsilon_t^T, t = h+1, ..., N,$$

where $X_t = (1, y_{t-1}^T, ..., y_{t-p}^T, x_{t-1}^T, ..., x_{t-q}^T)^T$, ϕ is the parameter matrix and $h = \max(p, q, d)$. The threshold variable takes values in $S = \{z_{h+1-d}, ..., z_{n-d}\}$, denoting $z_{(i)}$ the i-th smallest element of S, and letting $t(i)$ be the time index of $z_{(i)}$. Then, the arranged regression

based in the increasing order of the threshold variable z_{t-d} is

$$y^T_{t(i)+d} = X^T_{t(i)+d}\phi + \varepsilon^T_{t(i)+d}, i = 1, ..., N - h. \tag{1.6}$$

To built the test the author uses predictive residuals. The idea is simply, if y_t is linear then the recursive LS estimator of (1.6) is consistent and the predictive residuals approach white noise and are uncorrelated with the regressor $X_{t(i)+d}$. On the other hand, if y_t follows a threshold model, the predictive residuals will no longer be white noise because the recursive LSE is biased and would be correlated with the regressor. The author also generalizes in a similar way the results of [**Chan, 1993**] and [**Hansen, 2000**] about the consistency of the CLS estimator of the parameters in model (1.5), and deduces the asymptotic normal distribution for the slope parameters.

The estimation of threshold parameters and confidence intervals when the model is $SETAR(m, p)$ with discontinuous autoregressive function, is considered in the paper [**Kapetanios, 2003**]. Using a different approach, the author proposes the use of generalized method of moments (GMM), specially in small samples. Using moment conditions of the form

$$\mathbb{E}[z_{j,t}(r)\varepsilon_t] = \mathbb{E}[z_t \mathbb{I}_{\{r_{j-1} \leq Y_{t-d} < r_j\}}\varepsilon_t] = 0$$

where z_t are variables that provides extraneous information about the threshold parameters. The author considers the loss function,

$$Loss(Y; r) = \bar{m}(y; r)^T \bar{m}(y; r) \tag{1.7}$$

where

$$\bar{m}(y, r)_j = \frac{1}{N} \sum_{i \in \Gamma_j(r)} z_{j,t}(r)\hat{\varepsilon}_t(r),$$

$r = (r_1, ..., r_{m-1})^T$ and $\Gamma_j(r)$ denotes the set of observations for which the j-th moment condition is specified to hold and $\hat{\varepsilon}_t$ are the regression residuals. The author proves that the estimator, $\hat{r}$, of the threshold parameter defined by the minimization of $Loss(Y; r)$ is consistent. The author also suggests the use of bootstrap (and sub-sampling) for construction of standard errors and confidence intervals for the threshold parameter.

The model $SETAR(2, p)$, is considered in [**Gonzalo & Wolf, 2005**] and the proposed solution for inference about the threshold parameter is sub-sampling when the estimation procedure for the model parameters is, once more, CLS. The basis of constructing confidence intervals for the threshold parameter r is the approximation of the

sampling distribution of $\hat{r}_N$, properly normalized. With

$$J_N(u) = \mathbb{P}[N^\beta(\hat{r}_N - r) \leq u]$$

for some positive β, the sub-sampling approximation to $J_N(u)$ is defined by

$$L_{N,b}(u) = \frac{1}{N-b+1} \sum_{a=1}^{N-b+1} \mathbb{I}_{\{b^\beta(\hat{r}_{b,a} - \hat{r}_N) \leq u\}},$$

where the integer $1 < b < N$ is the block size and $\hat{r}_{b,a} = \hat{r}(Y_a, ..., Y_{a+b-1})$ is the estimator of r computed in the sub-sample $Y_a, ..., Y_{a+b-1}$, that is, CLS estimation applied to the sub-sample.

In the paper [**Gonzalo & Pitarakis, 2002**] the model of interest is the threshold regression model, and the authors give conditions under which the threshold parameters converge to the true value when the CLS procedure is used in the sequential estimation of the parameters. Computer burden becomes substantial when $m > 2$ and to overcome this problem the authors propose a sequential estimation procedure for the threshold parameters. The authors consider at each step of the procedure a threshold regression model with two regimes, and follow the CLS estimation with a slight difference. Instead of estimating $\hat{r}_N$ as in (1.3), that is, $\hat{r}_N = \arg\min_r S_N(r)$ the authors define

$$J_N(r) = S_N - S_N(r) \tag{1.8}$$

where S_N is the sum of the square of the residuals when the model is supposed linear and estimate $\hat{r}_N$ has

$$\hat{r}_N = \arg\max_r J_N(r).$$

The authors deduce the (non-stochastic) limit $J_\infty(r)$ of $J_N(r)$ and give conditions for the convergence

$$\hat{r}_N \xrightarrow{p} r_{(1)}$$

where $r_{(1)} \in \{r_1, ..., r_{m-1}\}$ is the threshold parameter that dominates all the others in terms of their contribution for the maximization of $J_\infty(r)$. Writing $\hat{r}_N^{(1)}$ for this first estimate, the sequential procedure continues, a second threshold estimate, $\hat{r}_N^{(2)}$, is given by

$$\hat{r}_N^{(2)} = \arg\max_r J_N(r|\hat{r}_N^{(1)})$$

with

$$J_N(r|\hat{r}_N^{(1)}) = J_{1,N}(r|\hat{r}_N^{(1)})\mathbb{I}_{\left\{r<\hat{r}_N^{(1)}\right\}} + J_{2,N}(r|\hat{r}_N^{(1)})\mathbb{I}_{\left\{r>\hat{r}_N^{(1)}\right\}}.$$

The objective functions $J_{1,N}$ and $J_{2,N}$ are of the same kind of (1.8) but are built when is adjusted a threshold regression model with two regimes at each sub-sample, and where the sub-samples are selected using the condition, on the threshold variable Z, $Z_t < \hat{r}_N^{(1)}$ and $Z_t > \hat{r}_N^{(1)}$, respectively. Then,

$$\hat{r}_N^{(2)} \xrightarrow{p} r_{(2)}$$

and the procedure follows until all the thresholds are estimated. In this paper, the authors also discuss the small sample behavior of the estimators, and provide an estimating procedure for the number of regimes.

The Bayesian theory is also used for inference in TAR models as in the paper of [**Stramer & Lin, 2002**]. The authors consider a $SETAR(2, p_1, p_2)$, that is, they consider a two regime $SETAR$ model but where the autoregressive order is p_1 in the first regime and p_2 in the second regime, that is,

$$Y_t = \begin{cases} a_{1,0} + a_{1,1}Y_{t-1} + \cdots + a_{1,p}Y_{t-p_1} + c_1\varepsilon_t, & \text{if } Y_{t-d} \leq r \\ a_{2,0} + a_{2,1}Y_{t-1} + \cdots + a_{2,p}Y_{t-p_2} + c_2\varepsilon_t, & \text{if } Y_{t-d} > r \end{cases}$$

where ε_t is standard white noise. The authors consider the following prior distribution for the parameters,

(i) A_i^p are independent $N(0, v_i^{-1}I_i), i = 1, 2$ with I_i the identity matrix and v_i positive scalar.

(ii) c_i^2 are independent $IG(\alpha, \beta)$ where IG denotes the inverse gamma distribution.

(iii) r follows a $U[a, b]$ (uniform distribution) where a and b are the 25% and 75% empirical quantiles of the data.

(iv) d follows $U(1, ..., D)$, a discrete uniform distribution for some positive integer D.

(v) p_i are independent $U(0, .., K_i)$ for some positive integer K_i, $i = 1, 2$.

The authors then propose a reversible jump Markov-Chain-Monte-Carlo algorithm to jump between models with different p's and for dealing with models with different values of d.

Some authors use wavelets for inference in threshold models as the case in the papers [**Li & Xie, 1999**] and [**Ip, Wong, Li & Xie, 1999**]. In the first paper the authors assume the $SETAR(m, p)$ model,

$$Y_t = \sum_{i=1}^{m} (a_{i,0} + a_{i,1}Y_{t-1} + \cdots + a_{i,p}Y_{t-p} + \varepsilon_{i,t})\, \mathbb{I}_{\{r_{i-1} \leq Y_{t-d} < r_i\}}$$

with $-\infty < a < r_1 < \cdots < r_{m-1} < b < \infty$, and introduce the function,

$$T(y_1, ..., y_p) = \sum_{i=1}^{m} (a_{i,0} + a_{i,1}y_1 + \cdots + a_{i,p}y_p)\, \mathbb{I}_{\{r_{i-1} \leq y_d < r_i\}}.$$

They built p empirical wavelet coefficients for the function $T(.)$ and use its properties for the estimation of the delay d, the number of regimes m and the threshold coefficients. For instance, checking the ones that have large absolute values the d parameter is estimated. By further checking the empirical wavelets corresponding to the time delay across the fine scale levels, the thresholds and their levels are identified. Finally the authors show the consistency of the estimators.

In the second paper, [**Ip, Wong, Li & Xie, 1999**], the estimators are built in the same way, but the process considered is the open-loop threshold autoregressive model ($TARSO$)

$$Y_t = \sum_{i=1}^{m} (a_{i,0} + a_{i,1}Y_{t-1} + \cdots + a_{i,p}Y_{t-p} + b_{i,0}X_t + \cdots + b_{i,q}X_{t-q} + \varepsilon_{i,t})\, \mathbb{I}_{\{r_{i-1} \leq X_{t-d} < r_i\}}.$$

Then, they define the function T for this process as

$$T(z) = \sum_{i=1}^{m} (a_{i,0} + a_{i,1}y_1 + \cdots + a_{i,p}y_p + b_{i,0}x_t + b_{i,1}x_{t-1} + \cdots + b_{i,q}x_{t-q})\, \mathbb{I}_{\{r_{i-1} \leq x_d < r_i\}}$$

with $z = (y_1, ..., y_p, x_1, ..., x_q)$, and proceed as in the case of the $SETAR$ model.

One fundamental question considered in several papers, [**Caner & Hansen, 2001**], [**Bec, Guay & Guerre, 2008**], [**Gonzalo & Montesinos, 2004**], [**Hansen, 1996**], [**Hansen, 1999**], [**Hansen & Seo, 2002**], [**Tsay, 1998**], [**Wong & Li, 2000**], is the question of testing for linearity, that is, the question of testing the linear model against some kind of threshold model.

1.3. Overview on diffusion threshold models

As to our knowledge the only paper where a threshold diffusion process is considered, is in [**Freidlin & Pfeiffer, 1998**]. In this paper the authors consider Brownian motion with a drift that changes from positive (b_1) to negative (b_0) and use it to build a threshold model where the upper threshold (x_1) is not known and the lower threshold is zero. The regimes in the threshold model are driven by the positive and negative drifts, the process is supposed to be continuously observed and the estimation procedure relies on maximum likelihood estimation. The authors consider $X_t, 0 \leq t < \infty$, a one-dimensional process

defined as the continuous solution of the equation

$$dX_t = b_{i(t)}dt + \sigma dB_t, X_0 = x < 0,$$

where

$$i(t) = \begin{cases} 1, & \tau_{2k} \leq t < \tau_{2k+1} \\ 0, & \tau_{2k+1} \leq t < \tau_{2k+2} \end{cases}$$

when $\tau_{2k+1} = \min\{t > \tau_{2k} : X_t = x_1\}$ and $\tau_{2k+2} = \min\{t > \tau_{2k+1} : X_t = 0\}$. The authors define a cycle for the level $0 < h < x_1$ as the time interval $[\tau^*_{2k}, \tau^*_{2k+2}[$ with $\tau^*_{2k+1} = \min\{t > \tau^*_{2k} : X_t = h\}$, $\tau^*_{2k+2} = \min\{t > \tau^*_{2k+1} : X_t = 0\}$, that is, the time interval between two successive intersections of the trajectory X_t with the level $x = 0$ separated by crossing of the level h and define $N_{T,h}$ as the total number of cycles for the trajectory of X_t in $[0, T]$. Next, for each cycle k they observe $X_k = \max_{\tau^*_{2k}<t<\tau^*_{2k+2}} X_t$ and define, with $X_{(i)}$ the ith-order statistic of the $N_{T,h}$ cycle maxima, the estimator of x_1 as the value $X_{(l(N_{T,h}))}$ that satisfies,

$$X_{(l(N_{T,h}))}(N_{T,h} - l(N_{T,h})) = \max_{1\leq i\leq N_{T,h}} X_{(i)}(N_{T,h} - i).$$

The authors prove the consistency of this estimator.

CHAPTER 2

MCEM-algorithm

In this chapter we start with a first section dedicated to hitting times for some diffusion processes and the chapter follows with a second section where we use the information on hitting times to build an Expectation-Maximization type algorithm to estimate the model parameters in the Brownian motion with drift threshold model.

2.1. Hitting times distribution

In this section we start with hitting times for the Brownian motion and the Brownian motion with drift processes. The results being known will be stated and the proofs will be written in order to allow the generalization for some other processes, namely the geometric Brownian motion process. Given a stochastic process X with right-continuous paths, which is adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$ and considering a subset $\Gamma \in \mathcal{B}(\mathbb{R})$ of the state space of the process, the hitting time (in Γ) is defined as,

$$T_\Gamma(\omega) = \inf\{t \geq 0; X_t(\omega) \in \Gamma\}.$$

In the following the hitting times of interest will be in some predefined level or threshold.

2.1.1. Standard Brownian motion. Let $B = (B_t)_{t\geq 0}$ be standard Brownian motion and $T_b = \inf\{t \geq 0 : B_t = b\}, b \in \mathbb{R}$ the first hitting time of the level b by the process B. We want to study the distribution of T_b.

Definition 2.1. Define the running maximum, or the maximum to the date, M_t, of the Brownian motion by: $\forall t > 0, M_t = \sup_{0\leq s\leq t} B_s$

First we suppose $b > 0$, from the following proposition is easy to deduce the density function for T_b.

Proposition 2.2 ([**Hida, 1980**] or [**Revuz & York, 1991**]). The following equalities hold for $b \geq 0$ and $t > 0$:

$$\mathbb{P}[T_b \leq t] = \mathbb{P}[M_t \geq b] = 2\mathbb{P}[B_t \geq b]$$

We obtain the distribution function F_{T_b},

$$F_{T_b}(t) = \mathbb{P}[T_b \le t] = \mathbb{P}[M_t \ge b] = \frac{2}{\sqrt{2\pi t}} \int_b^{+\infty} e^{-\frac{y^2}{2t}} dy =$$
$$= \frac{2}{\sqrt{2\pi t}} \int_{\frac{b}{\sqrt{t}}}^{+\infty} e^{-\frac{x^2}{2}} \sqrt{t} dx.$$

doing the substitution $x = \frac{y}{\sqrt{t}}$. By differentiation the density of T_b is obtained,

$$f_{T_b}(t) = \frac{b}{\sqrt{2\pi t^3}} e^{-\frac{b^2}{2t}}, \quad t \ge 0, \quad b > 0.$$

In the case $b < 0$, we need the following definition.

Definition 2.3. Define the running minimum, or the minimum to the date, m_t, of the Brownian motion by: $\forall t > 0, m_t = \inf_{0 \le s \le t} B_s$.

Corollary 2.4. The following equalities hold for $b \le 0$ and $t > 0$:

$$\mathbb{P}[T_b \le t] = \mathbb{P}[m_t \le b] = 2\mathbb{P}[B_t \ge -b].$$

Proof. If $b < 0$ we have $\{T_b \le t\} = \{m_t \le b\}$, but

$$\{m_t \le b\} = \left\{ \inf_{0 \le s \le t} B_s \le b \right\} = \left\{ - \sup_{0 \le s \le t} (-B_s) \le b \right\}$$
$$= \left\{ \sup_{0 \le s \le t} B_s \ge -b \right\},$$

because $(-B_t)_{t \ge 0}$ is also a Brownian motion starting at the origin, and by proposition 2.2 we have

$$\mathbb{P}[m_t \le b] = \mathbb{P}[M_t \ge -b] = 2\mathbb{P}[B_t \ge -b].$$

□

For $b < 0$ we have $\{T_b \le t\} = \{m_t \le b\} = 2\mathbb{P}[B_t \ge -b]$ then in this case the density of T_b is given by

$$f_{T_b}(t) = \frac{-b}{\sqrt{2\pi t^3}} e^{-\frac{b^2}{2t}}, \quad t \ge 0, \quad b < 0,$$

and so we can finally write the density for $T_b, \quad b \in \mathbb{R}$,

$$f_{T_b}(t) = \frac{|b|}{\sqrt{2\pi t^3}} e^{-\frac{b^2}{2t}}, \quad t \ge 0.$$

The moment generating function for T_b is known and can be written as

$$\varphi_{T_b}(\alpha) = \int_0^{+\infty} \frac{|b| e^{-\alpha t}}{2\pi t^3} e^{-\frac{b^2}{2t}} dt = \int_0^{+\infty} \frac{|b|}{2\pi t^3} e^{-\frac{b^2 + 2\alpha t^2}{2t}} dt = e^{-|b|\sqrt{2\alpha}}, \quad \alpha > 0.$$

2.1.2. Brownian motion with drift μ. In this section we want to study the hitting times for the Brownian motion with drift process and we follow the construction in [**Karatzas & Shreve, 1991**]. Let us consider the Ito process given by,

$$d\tilde{B}_t = dB_t - \mu dt, \quad t \leq T, \quad \tilde{B}_0 = 0$$

for $T \in [0, +\infty[$ and B_t standard Brownian motion. We continue with, $T_b = \inf\{t \geq 0 : B_t = b\}$. Let,

$$M_t = \exp\left(\int_0^t \mu dB_s - \frac{1}{2}\int_0^t \mu^2 ds\right) = e^{\mu B_t - \frac{1}{2}\mu^2 t},$$

then by the Girsanov theorem ([**Øksendal, 1998**] or [**Karatzas & Shreve, 1991**]) we define the measure $\mathbb{P}^\mu$ in $(\Omega, \mathcal{F}_T)$ by,

$$d\mathbb{P}^\mu = M_T d\mathbb{P} = \exp\left(\mu B_T - \frac{1}{2}\mu^2 T\right) d\mathbb{P}$$

and w.r.t. $\mathbb{P}^\mu$, the process $\tilde{B}_t, t \leq T$ is a Brownian motion. Then we say that, under $\mathbb{P}^\mu$, $B_t = \mu t + \tilde{B}_t$ is a Brownian motion with drift μ.

Remark 2.5. We have that,

(1) $(M_t)_{0 \leq t \leq T}$ is a martingale w.r.t. $(\mathcal{F}_t)_{t \leq T}$ and $\mathbb{P}$, ([**Øksendal, 1998**](Ex. 4.4)), and so $M_T d\mathbb{P} = M_t d\mathbb{P}$ in $\mathcal{F}_t$.
(2) $\mathbb{P}^\mu[A] = \mathbb{E}[\mathbb{I}_A M_t], \forall A \in \mathcal{F}_t$ because $\mathbb{P}^\mu[A] = \int_\Omega \mathbb{I}_A d\mathbb{P}^\mu = \int_\Omega \mathbb{I}_A M_t d\mathbb{P} = \mathbb{E}[\mathbb{I}_A M_t]$.

With t fixed, on the set $\{T_b \leq t\} \in \mathcal{F}_t \cap \mathcal{F}_{T_b} = \mathcal{F}_{t\wedge T_b}$ we have $M_{t\wedge T_b} = M_{T_b}$ and the Optional Sampling theorem ([**Karatzas & Shreve, 1991**] or [**Revuz & York, 1991**]) with the remark imply

$$\begin{aligned}
\mathbb{P}^\mu[T_b \leq t] &= \mathbb{E}\left[\mathbb{I}_{\{T_b\leq t\}} M_t\right] = \mathbb{E}\left[\mathbb{E}[\mathbb{I}_{\{T_b\leq t\}} M_t | \mathcal{F}_{t\wedge T_b}]\right] = \mathbb{E}\left[\mathbb{I}_{\{T_b\leq t\}} \mathbb{E}[M_t | \mathcal{F}_{t\wedge T_b}]\right] \\
&= \mathbb{E}\left[\mathbb{I}_{\{T_b\leq t\}} M_{t\wedge T_b}\right] = \mathbb{E}\left[\mathbb{I}_{\{T_b\leq t\}} M_{T_b}\right] = \mathbb{E}\left[\mathbb{I}_{\{T_b\leq t\}} \exp\left(\mu B_{T_b} - \frac{1}{2}\mu^2 T_b\right)\right] \\
&= \mathbb{E}\left[\mathbb{I}_{\{T_b\leq t\}} \exp\left(\mu b - \frac{1}{2}\mu^2 T_b\right)\right] = \int_0^t e^{\mu b - \frac{1}{2}\mu^2 s} \mathbb{P}[T_b \in ds] \\
&= \int_0^t e^{\mu b - \frac{1}{2}\mu^2 s} \frac{|b|}{\sqrt{2\pi s^3}} e^{-\frac{b^2}{2s}} ds = \int_0^t \frac{|b|}{\sqrt{2\pi s^3}} e^{-\frac{(b-\mu s)^2}{2s}} ds.
\end{aligned} \tag{2.1}$$

So T_b has the density, under $\mathbb{P}^\mu$,

$$\mathbb{P}^\mu[T_b \in dt] = \frac{|b|}{\sqrt{2\pi t^3}} e^{-\frac{(b-\mu t)^2}{2t}} dt, \quad t > 0.$$

Because, under $\mathbb{P}^\mu$, $B_t = \mu t + \tilde{B}_t$ is a Brownian motion with drift μ, we can say that the density of T_b^μ is

$$f_{T_b^\mu}(t) = \frac{|b|}{\sqrt{2\pi t^3}} e^{-\frac{(b-\mu t)^2}{2t}}, \quad t > 0,$$

with $T_b^\mu = \inf\{t : \mu t + W_t = b\}$ were W_t is the standard Brownian motion. We will write B_t^μ for $\mu t + B_t$ and we get the moment generating function for T_b^μ,

$$\varphi_{T_b^\mu}(\alpha) = \mathbb{E}[e^{-\alpha T_b^\mu}] = \int_0^{+\infty} e^{-\alpha t} f_{T_b^\mu}(t)dt = e^{\mu b - |b|\sqrt{\mu^2+2\alpha}}, \quad \alpha > 0.$$

If we consider a Brownian motion with drift μ but starting at a instead of 0, represented as $B_{t,a}^\mu$, and because $T_{b,a}^\mu = \inf\{t : B_{t,a}^\mu = b\} = \inf\{t : B_t^\mu = b - a\} = T_{b-a}^\mu$ we get the density

$$f_{T_{b,a}^\mu}(t) = \frac{|b-a|}{\sqrt{2\pi t^3}} e^{-\frac{(b-a-\mu t)^2}{2t}}, \quad t > 0,$$

and the moment generating function,

$$\varphi_{T_{b,a}^\mu}(\alpha) = e^{\mu(b-a) - |b-a|\sqrt{\mu^2+2\alpha}}, \quad \alpha > 0.$$

2.1.3. Brownian motion with drift μ and diffusion coefficient σ. Let us consider the Brownian motion with drift μ and diffusion coefficient σ,

$$B_t^{\mu,\sigma} = \mu t + \sigma B_t$$

with B_t the standard Brownian motion. Starting with

$$B_t^{\mu/\sigma} = \frac{\mu}{\sigma} t + B_t,$$

then we have

$$B_t^{\mu,\sigma} = \sigma B_t^{\mu/\sigma}.$$

But

$$T_b^{\mu,\sigma} = \inf\left\{t : B_t^{\mu,\sigma} = b\right\} = \inf\left\{t : \sigma B_t^{\mu/\sigma} = b\right\} = \inf\left\{t : B_t^{\mu/\sigma} = \frac{b}{\sigma}\right\} = T_{\frac{b}{\sigma}}^{\mu/\sigma},$$

and because of what we have done in the last section, we obtain the density when $B_0^{\mu,\sigma} = a \Leftrightarrow B_0^{\mu/\sigma} = a/\sigma$

$$f_{T_{b,a}^{\mu,\sigma}}(t) = \frac{|b-a|}{\sqrt{2\pi\sigma^2 t^3}} e^{-\frac{(b-a-\mu t)^2}{2\sigma^2 t}}, \quad t > 0,$$

and, the moment generating function,

$$\varphi_{T_{b,a}^{\mu,\sigma}}(\alpha) = e^{\frac{\mu(b-a) - |b-a|\sqrt{\mu^2+2\alpha\sigma^2}}{\sigma^2}}, \quad \alpha > 0.$$

2.1.4. Transformation of the Brownian motion with drift. Let us consider the Brownian motion with drift, $B_t^{\mu,\sigma}$, we know the density for the first hitting time, $T_{b,a}^{\mu,\sigma}$. Define a new process by

$$Y_t = g\left(B_t^{\mu,\sigma}\right), \quad Y_0 = a,$$

for an invertible monotonous function g. Because

$$\begin{aligned} T_{b,a}^{Y} &= \inf\left\{t \geq 0 : Y_t = b\right\} = \inf\left\{t \geq 0 : g\left(B_t^{\mu,\sigma}\right) = b\right\} \\ &= \inf\left\{t \geq 0 : B_t^{\mu,\sigma} = g^{-1}(b)\right\} = T^{\mu,\sigma}_{g^{-1}(b),g^{-1}(a)}, \end{aligned}$$

we have the density for $T_{b,a}^{Y}$ given by:

$$f_{T_{b,a}^{Y}}(t) = \frac{\left|g^{-1}(b) - g^{-1}(a)\right|}{\sqrt{2\pi\sigma^2 t^3}}\, e^{-\frac{\left(g^{-1}(b)-g^{-1}(a)-\mu t\right)^2}{2\sigma^2 t}}, t > 0, \tag{2.2}$$

and the moment generating function

$$\varphi_{T_{b,a}^{Y}}(\alpha) = e^{\frac{\mu\left(g^{-1}(b)-g^{-1}(a)\right)-\left|g^{-1}(b)-g^{-1}(a)\right|\sqrt{\mu^2+2\alpha\sigma^2}}{\sigma^2}}, \quad \alpha > 0.$$

Remark 2.6. When we consider the process $Y_t = g\left(B_t^{\mu,\sigma}\right)$ by the Ito's formula, we get

$$dY_t = \left[g'\left(g^{-1}(Y_t)\right)\mu + \frac{\sigma^2}{2} g''\left(g^{-1}(Y_t)\right)\right] dt + \sigma g'\left(g^{-1}(Y_t)\right) dB_t, \tag{2.3}$$

and the results can be applied to the processes solution of this kind of s.d.e.

2.1.5. Geometric Brownian motion. If we consider the geometric Brownian motion, a positive process satisfying,

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad X_0 = a.$$

We know that the solution of the last s.d.e. is given by,

$$X_t = a\, e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t},$$

we can then write

$$X_t = g\left(B_t^{\mu - \frac{\sigma^2}{2}, \sigma}\right)$$

with

$$g(z) = ae^z, \quad g^{-1}(x) = \ln\left(\frac{x}{a}\right).$$

Then we can write, using (2.2), the density for the first hitting time of the geometric Brownian motion,

$$f_{T_{b,a}^{X}}(t) = \frac{\left|\ln(b) - \ln(a)\right|}{\sqrt{2\pi\sigma^2 t^3}}\, e^{-\frac{\left(\ln(b)-\ln(a)-\left(\mu-\frac{\sigma^2}{2}\right)t\right)^2}{2\sigma^2 t}}, t > 0.$$

and, the moment generating function,

$$\varphi_{T^X_{b,a}}(\alpha) = e^{\frac{1}{2}\left[\ln(a)-\ln(b)+\frac{2\mu(\ln(b)-\ln(a))-|\ln(b)-\ln(a)|\sqrt{8\alpha\sigma^2+(\sigma^2-2\mu)^2}}{\sigma^2}\right]}, \quad \alpha > 0.$$

2.1.6. Simple generalization. More generally, instead of starting with a Brownian motion, we can think in the process solution of the SDE

$$dX_t = \alpha(t)dB_t,\ X_0 = 0,$$

that is,

$$X_t = \int_0^t \alpha(s)dB_s,$$

where $\alpha(t)$ is a deterministic regular function. Our goal is to obtain the density for the hitting times of such a process. We know that X_t is a time changed Brownian motion

$$X_t = B_{\langle X\rangle_t} \text{ with } \langle X\rangle_t = \int_0^t \alpha(s)^2 ds,$$

and we need a reflection principle for this kind of processes.

Proposition 2.7. Let $X_t = \int_0^t \alpha(s)dB_s, X_0 = 0$, $\alpha(s)$ continuous function and $\forall t > 0, \quad \langle X\rangle_t \geq t$. With $M_t^X = \sup_{0\leq s\leq t} X_s$ and $T_b^X = \inf\{t \geq 0; X_t = b\}$ we have for $b \geq 0$

$$\mathbb{P}[T_b^X \leq t] = \mathbb{P}[M_t^X \geq b] = 2\mathbb{P}[X_t \geq b].$$

Proof. First observe that $\{T_b^X \leq t\} = \{M_t^X \geq b\}$ because B_t has continuous trajectories and α is continuous. On the other hand we have

$$\mathbb{P}[M_t^X \geq b] = \mathbb{P}[M_t^X \geq b, X_t \geq b] + \mathbb{P}[M_t^X \geq b, X_t < b]$$

because $\{X_t \geq b\} \subseteq \{M_t^X \geq b\}$ we get

$$\mathbb{P}[M_t^X \geq b, X_t \geq b] = \mathbb{P}[X_t \geq b].$$

Because $X_{T_b^X} = b$ we can write

$$\mathbb{P}[M_t^X \geq b, X_t < b] = \mathbb{P}[T_b^X \leq t, X_t < b] = \mathbb{P}[T_b^X \leq t, X_{T_b^X+(t-T_b^X)} - X_{T_b^X} < 0]$$

and with $s = t - T_b^X$

$$X_{T_b^X+s} - X_{T_b^X} = B_{\langle X\rangle_{(T_b^X+s)}} - B_{\langle X\rangle_{T_b^X}}.$$

Taking into account that $\langle X\rangle_t$ is increasing and in $\{T_b^X \leq t\}$ we have $s > 0$, so

$$\langle X\rangle_{T_b^X+s} \geq \langle X\rangle_{T_b^X},$$

and with B Brownian motion and the hypotheses made we get

$$B_{\langle X\rangle_{(T_b^X+s)}} - B_{\langle X\rangle_{T_b^X}} \perp \mathfrak{F}_{\langle X\rangle_{T_b^X}} \supseteq \mathfrak{F}_{T_b^X}.$$

Then

$$X_{T_b^X+s} - X_{T_b^X} \perp \mathfrak{F}_{T_b^X}$$

and

$$\mathbb{P}[M_t^X \geq b, X_t < b] = \mathbb{P}[T_b^X \leq t]\mathbb{P}[X_{T_b^X+s} - X_{T_b^X} < 0] = \frac{1}{2}\mathbb{P}[T_b^X \leq t],$$

because $X_{T_b^X+s} - X_{T_b^X}$ has Normal distribution with zero mean. Finally,

$$\mathbb{P}[M_t^X \geq b] = 2\mathbb{P}[X_t \geq b].$$

□

In the same way as earlier we can get the distribution for $T_b^X = \inf\{t \geq 0 : X_t = b\}, b \geq 0$, once $B_{\langle X\rangle_t} \sim N(0, \langle X\rangle_t)$,

$$F_{T_b^X}(t) = \mathbb{P}[M_t^X \geq b] = 2\mathbb{P}[B_{\langle X\rangle_t} \geq b] = \frac{2}{\sqrt{2\pi\langle X\rangle_t}}\int_b^{+\infty} e^{-\frac{y^2}{2\langle X\rangle_t}}dy =$$

$$= \frac{2}{\sqrt{2\pi\langle X\rangle_t}}\int_{\frac{b}{\sqrt{\langle X\rangle_t}}}^{+\infty} e^{-\frac{x^2}{2}}\sqrt{\langle X\rangle_t}dx,$$

doing the substitution

$$x = \frac{y}{\sqrt{\langle X\rangle_t}}.$$

By differentiation we obtain the density of T_b^X,

$$f_{T_b^X}(t) = \frac{b\langle X\rangle_t'}{\sqrt{2\pi\langle X\rangle_t^3}}e^{-\frac{b^2}{2\langle X\rangle_t}}, \quad t \geq 0, \quad b > 0.$$

For $b < 0$ we have the following analogous result.

Corollary 2.8. With the same assumptions of the last proposition, and $m_t^X = \inf_{0\leq s\leq t} X_s$ and $T_b^X = \inf\{t \geq 0; X_t = b\}$ we have for $b < 0$

$$\mathbb{P}[T_b^X \leq t] = \mathbb{P}[m_t^X \leq b] = 2\mathbb{P}[X_t \geq -b].$$

Proof. It is enough to see,

$$\{T_b^X \leq t\} = \{m_t^X \leq b\} = \{\inf_{0\leq s\leq t} X_s \leq b\} = \{\sup_{0\leq s\leq t}(-X_s) \geq -b\}$$

$$= \{\sup_{0\leq s\leq t} B_{\langle -X\rangle_s} \geq -b\} = \{\sup_{0\leq s\leq t} B_{\langle X\rangle_s} \geq -b\} = \{M_t^X \geq -b\}.$$

By proposition 2.7,

$$\mathbb{P}[T_b^X \leq t] = \mathbb{P}[M_t^X \geq -b] = 2\mathbb{P}[X_t \geq -b].$$

□

Becoming

$$f_{T_b^X}(t) = \frac{|b|\langle X\rangle_t'}{\sqrt{2\pi\langle X\rangle_t^3}} e^{-\frac{b^2}{2\langle X\rangle_t}}, \quad t \geq 0, \quad b \in \mathbb{R}. \tag{2.4}$$

REMARK 2.9. If we consider $\alpha(t) \equiv 1$ then $X_t = B_t$, $\langle X\rangle_t = t$ and the last density is just what we have computed in a previous section.

In order to generalize this result to a drifted process, let Y_t be the solution of

$$dY_t = -\mu\alpha(t)dt + dB_t,\ t \leq T,\ Y_0 = 0$$

with $\alpha(t)$ as in proposition 2.7, B_t a Brownian motion. Then by Girsanov's with

$$M_t = \exp\left(\int_0^t \mu\alpha(s)dB_s - \frac{1}{2}\int_0^t \mu^2\alpha^2(s)ds\right),\ t \leq T$$

we define the measure $\mathbb{Q}$ in $(\Omega, \mathfrak{F}_T)$ by,

$$d\mathbb{Q} = M_T d\mathbb{P},$$

with respect to which

$$Y_t = -\mu\int_0^t \alpha(s)ds + B_t$$

is a Brownian motion, for $t \leq T$. Let

$$X_t = \int_0^t \alpha(s)dB_s,\ X_0 = 0, \quad T_b^X = \inf\{t \geq 0; X_t = b\},$$

similarly to the case of the Brownian motion with drift, remark 2.5 and equation (2.1), because $X_t \in \mathfrak{F}_t$, $\{T_b^X \leq t\} \in \mathfrak{F}_t$ and $T_b^X \in \mathfrak{F}_{T_b^X}$.

$$\begin{aligned}
\mathbb{Q}[T_b^X \leq t] &= \mathbb{E}^{\mathbb{P}}\left[\mathbb{I}_{\{T_b^X \leq t\}} \exp\left(\mu\int_0^{T_b^X} \alpha(s)dB_s - \frac{1}{2}\mu^2\int_0^{T_b^X} \alpha^2(s)ds\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\mathbb{I}_{\{T_b^X \leq t\}} \exp\left(\mu X_{T_b^X} - \frac{\mu^2}{2}\langle X\rangle_{T_b^X}\right)\right] = \int_0^t e^{\mu b - \frac{\mu^2}{2}\langle X\rangle_s}\mathbb{P}[T_b^X \in ds] \\
&= \int_0^t e^{\mu b - \frac{\mu^2}{2}\langle X\rangle_s} \frac{|b|\langle X\rangle_s'}{\sqrt{2\pi\langle X\rangle_s^3}} e^{-\frac{b^2}{2\langle X\rangle_s}} ds = \int_0^t \frac{|b|\langle X\rangle_s'}{\sqrt{2\pi\langle X\rangle_s^3}} e^{-\frac{(b-\mu\langle X\rangle_s)^2}{2\langle X\rangle_s}} ds.
\end{aligned}$$

So the density w.r.t. $\mathbb{Q}$ of T_b^X is given by,

$$f_{T_b^X}(t) = \frac{|b|\langle X\rangle_t'}{\sqrt{2\pi\langle X\rangle_t^3}} e^{-\frac{(b-\mu\langle X\rangle_t)^2}{2\langle X\rangle_t}}.$$

We can write,

$$X_t = \int_0^t \alpha(s)dB_s = \int_0^t \alpha(s)dY_s + \mu\int_0^t \alpha^2(s)ds$$

also w.r.t. $\mathbb{Q}$, Y_t is Brownian motion, and we can conclude that for the process solution of the SDE

$$dX_t = \mu\alpha^2(t)dt + \alpha(t)dB_t,$$

with B_t Brownian motion, we have

$$f_{T_b^X}(t) = \frac{|b|\gamma'(t)}{\sqrt{2\pi\gamma^3(t)}} e^{-\frac{(b-\mu\gamma(t))^2}{2\gamma(t)}}, \quad \gamma(t) = \int_0^t \alpha^2(s)ds.$$

REMARK 2.10. Once more, writing $\alpha(t) \equiv 1$ then X_t is just a Brownian motion with drift μ, $\gamma(t) = t$ and

$$f_{T_b^X}(t) = \frac{|b|}{\sqrt{2\pi t^3}} e^{-\frac{(b-\mu t)^2}{2t}},$$

just as earlier for the Brownian motion with drift.

2.2. MCEM-algorithm in the threshold model

In this section we consider the Brownian motion with drift in order to built a threshold model. Given two thresholds, m and M, and supposing that the change occurs only in the drift coefficient we built a threshold model. Starting the process with a positive trend (driven by $\mu_1 > 0$), the process continues in this first regime until it hits the upper threshold M, at that time a change occurs and the process follows with negative trend (driven by $\mu_2 < 0$) and continues in this second regime until it hits the lower threshold m, starting all over again in the first regime. To estimate the thresholds m and M and the drift parameters we use an Expectation-Maximization (EM) type algorithm for incomplete data because we will work with observed data from the process but this data do not include observations of the threshold hitting times.

Let us consider the following model:

$$dXt = \mu(t)dt + \sigma dB_t, \quad X_0 = x_0 \in [m, M[,$$

where

$$\mu(t) = \sum_{k\geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k}, \tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1}, \tau_{2k+2}[}(t)\right], \quad \mu_1 > 0,\ \mu_2 < 0$$

where $0 = \tau_0 < \tau_1 < \ldots < \tau_k < \ldots$ are the threshold hitting times. That is, giving two thresholds m and M, τ_1 is the time needed for the Brownian motion with drift μ_1 to go from x_0 to M, $\tau_2 - \tau_1$ is the time needed for the process with drift μ_2 go from M to m and, more generally, $\tau_i - \tau_{i-1}$ is the time needed for the process to go from one threshold to the other. We suppose σ is known and we want to estimate $\theta = (\mu_1, \mu_2, m, M)$. Suppose that we observe $X_0, X_1, ..., X_n$ at equidistant times $0 = t_0, t_1, ..., t_n$ in the time interval $[0, T]$,

with $\Delta = t_j - t_{j-1}$ and $t_j = j\Delta$ for $j = 1, ..., n$. In both regimes the Brownian motion with drift is used to build the threshold model as explained before, however the resulting process is not a Brownian motion with drift in each of the regimes. Notice that, in the Brownian motion with drift process we have $X_{j+1} - X_j \sim N(\mu\Delta, \sigma^2\Delta)$ but in the corresponding threshold process this is not true, even if we know that is no change in regime between the observation times. The transition densities are not known for the threshold process and for that reason we use what we think that is a natural approximation, the transition densities of the Brownian motion with drift process. In the next chapter we will return again to the problem of the threshold process not being Brownian motion with drift in each regime. If we consider Brownian motion with drift μ_1 in the first regime of the process and Brownian motion with drift μ_2 in the second regime, then p_1 and p_2, the transition densities in regime 1 and 2, are

$$\begin{aligned} p_1(\Delta, x_i, x_{i+1}; \mu_1) &= \frac{1}{\sqrt{2\pi\sigma^2\Delta}} \exp\left(-\frac{(x_{i+1} - x_i - \mu_1\Delta)^2}{2\sigma^2\Delta}\right) \\ p_2(\Delta, x_i, x_{i+1}; \mu_2) &= \frac{1}{\sqrt{2\pi\sigma^2\Delta}} \exp\left(-\frac{(x_{i+1} - x_i - \mu_2\Delta)^2}{2\sigma^2\Delta}\right). \end{aligned}$$

The hitting times $\tau_1, ..., \tau_k$ are not observed and their number in the time interval $[0, T]$ is given by some random variable K. The densities for the difference between two consecutive hitting times $\tau_1, ..., \tau_k$ are known (subsection 2.1.3), and given by:

$$f_{T_1}(\tau_1; \mu_1, M) = \frac{M - x_0}{\sqrt{2\pi\sigma^2\tau_1^3}} \exp\left(-\frac{(M - x_0 - \mu_1\tau_1)^2}{2\sigma^2\tau_1}\right),$$

for $i = 2, 4, 6, ...,$

$$f_{T_i - T_{i-1}}(\tau_i - \tau_{i-1}; \theta) = \frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}} \exp\left(-\frac{(m - M - \mu_2(\tau_i - \tau_{i-1}))^2}{2\sigma^2(\tau_i - \tau_{i-1})}\right),$$

and for $i = 3, 5, 7, ...,$

$$f_{T_i - T_{i-1}}(\tau_i - \tau_{i-1}; \theta) = \frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}} \exp\left(-\frac{(M - m - \mu_1(\tau_i - \tau_{i-1}))^2}{2\sigma^2(\tau_i - \tau_{i-1})}\right).$$

Because the hitting times process is Markov, we can write the joint density for k hitting times as,

$$\begin{aligned} f_{T_1,...,T_k}(\tau_1, ..., \tau_k; \theta) &= \frac{M - x_0}{\sqrt{2\pi\sigma^2\tau_1^3}} \exp\left(-\frac{(M - x_0 - \mu_1\tau_1)^2}{2\sigma^2\tau_1}\right) \prod_{i=2}^{k} \frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}} \\ \times \prod_{i=1}^{[(k-1)/2]} &\exp\left(-\frac{(M - m - \mu_1(\tau_{2i+1} - \tau_{2i}))^2}{2\sigma^2(\tau_{2i+1} - \tau_{2i})}\right) \prod_{i=1}^{[k/2]} \exp\left(-\frac{(m - M - \mu_2(\tau_{2i} - \tau_{2i-1}))^2}{2\sigma^2(\tau_{2i} - \tau_{2i-1})}\right), \end{aligned}$$

where $[v] = \max_{l\in\mathbb{N}}\{l \leq v\}$. To get estimators for the drift and the threshold parameters we will use a Monte Carlo EM algorithm, because the hitting times are not observed and we will use Monte Carlo simulation in the expectation step.

To implement the algorithm, we will augment the data $X_1, ..., X_n$ by considering the hitting times $T_1, ..., T_K$ in order to write the complete likelihood function,

$$L_c^K(\omega) = \sum_{k=1}^{+\infty} L_c(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta)\mathbb{I}_{\{K=k\}}(\omega)$$

with K the random variable that represents the number of hitting times in the interval $[0, T]$ (we don't consider the term were $K = 0$ because in that case there is no change in regime in $[0, T]$). Conditionally on the hitting times the process X_t is Markov and we have

$$\begin{aligned}
&L_c(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta) = f_{X_1,...,X_n,T_1,...,T_k}(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta) \\
&= f_{X_1,...,,X_n|T_1,...,T_k}(x_1, ..., x_n|\tau_1, ..., \tau_k)f_{T_1,...,T_k}(\tau_1, ..., \tau_k; \theta) \\
&= \prod_{i=0}^{j_1-1} p_1(\Delta, x_i, x_{i+1})p_1(\tau_1 - t_{j_1}, x_{j_1}, M)p_2(t_{j_1+1} - \tau_1, M, x_{j_1+1}) \prod_{i=j_1+1}^{j_2-1} p_2(\Delta, x_i, x_{i+1}) \\
&\times p_2(\tau_2 - t_{j_2}, x_{j_2}, m)p_1(t_{j_2+1} - \tau_2, m, x_{j_2+1}) \cdots p_{[1+mod(k+1,2)]}(\tau_k - t_{j_k}, x_{j_k}, \star) \\
&\times p_{[1+mod(k,2)]}(t_{j_k+1} - \tau_k, \star, x_{j_k+1}) \prod_{i=j_k+1}^{n-1} p_{[1+mod(k,2)]}(\Delta i, x_i, x_{i+1}) \\
&\times \frac{M - x_0}{\sqrt{2\pi\sigma^2\tau_1^3}} \exp\left(-\frac{(M - x_0 - \mu\tau_1)^2}{2\sigma^2\tau_1}\right) \prod_{i=2}^{k} \frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}} \\
&\times \prod_{i=1}^{[(k-1)/2]} \exp\left(-\frac{(M - m - \mu_1(\tau_{2i+1} - \tau_{2i}))^2}{2\sigma^2(\tau_{2i+1} - \tau_{2i})}\right) \prod_{i=1}^{[k/2]} \exp\left(-\frac{(m - M - \mu_2(\tau_{2i} - \tau_{2i-1}))^2}{2\sigma^2(\tau_{2i} - \tau_{2i-1})}\right),
\end{aligned}$$

where for all $i = 1, ..., k$, $j_i \in \{1, ..., n-1\}$ is, such that $\tau_i \in]t_{j_i}, t_{j_i+1}[$ and $j_1 < j_2 < \cdots < j_k$. With $\star = m \times mod(k+1, 2) + M \times mod(k, 2)$ and $mod(a, b)$ is the remainder of the integer division of a by b. Then we define the complete log-likelihood in a similar way,

$$\log(L_c^K)(\omega) = \sum_{k=1}^{+\infty} \log(L_c(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta))\mathbb{I}_{\{K=k\}}(\omega)$$

with

$$
\begin{aligned}
&\log(L_c(x_1,...,x_n,\tau_1,...,\tau_k;\theta)) = \\
&= -\left(\sum_{i=0}^{j_1-1} \frac{(x_{i+1}-x_i-\mu_1\Delta)^2}{2\sigma^2\Delta} + \sum_{i=j_1+1}^{j_2-1} \frac{(x_{i+1}-x_i-\mu_2\Delta)^2}{2\sigma^2\Delta} + \cdots\right. \\
&+ \sum_{i=j_k-1}^{n-1} \frac{(x_{i+1}-x_i-\mu_{[1+mod(k,2)]}\Delta)^2}{2\sigma^2\Delta} + \sum_{i=1}^{[(k+1)/2]} \frac{(M-x_{j_{2i-1}}-\mu_1(\tau_{2i-1}-t_{j_{2i-1}}))^2}{2\sigma^2(\tau_{2i-1}-t_{j_{2i-1}})} + \\
&+ \sum_{i=1}^{[(k+1)/2]} \frac{(x_{j_{2i-1}+1}-M-\mu_2(t_{j_{2i-1}+1}-\tau_{2i-1}))^2}{2\sigma^2(t_{j_{2i-1}+1}-\tau_{2i-1})} + \sum_{i=1}^{[k/2]} \frac{(m-x_{j_{2i}}-\mu_2(\tau_{2i}-t_{j_{2i}}))^2}{2\sigma^2(\tau_{2i}-t_{j_{2i}})} \\
&\left.+ \sum_{i=1}^{[k/2]} \frac{(x_{j_{2i}+1}-m-\mu_1(t_{j_{2i}+1}-\tau_{2i}))^2}{2\sigma^2(t_{j_{2i}+1}-\tau_{2i})}\right) + \log\left(\frac{M-x_0}{\sqrt{2\pi\sigma^2\tau_1^3}}\right) \\
&- \frac{(M-x_0-\mu_1\tau_1)^2}{2\sigma^2\tau_1} + \sum_{i=2}^{k} \log\left(\frac{M-m}{\sqrt{2\pi\sigma^2(\tau_i-\tau_{i-1})^3}}\right) \\
&- \sum_{i=1}^{[(k-1)/2]} \frac{(M-m-\mu_1(\tau_{2i+1}-\tau_{2i}))^2}{2\sigma^2(\tau_{2i+1}-\tau_{2i})} - \sum_{i=1}^{[k/2]} \frac{(m-M-\mu_2(\tau_{2i}-\tau_{2i-1}))^2}{2\sigma^2(\tau_{2i}-\tau_{2i-1})},
\end{aligned}
$$

apart from a term not involving θ.

2.2.1. E-step. Using the EM algorithm we need in the p-iteration of the E-step to compute the expectation of the log-likelihood of the augmented data given the observed data and current value $\theta_p = (\mu_{1,p}, \mu_{2,p}, m_p, M_p)$, that is

$$Q(\theta,\theta_p) = \mathbb{E}_{\theta_p}\left[\log\left(L_c^K(X_1,...,X_n,T_1,...,T_K)\right)|X_1,...,X_n\right].$$

We will approximate this expectation conditionally on the observed data through Monte Carlo.

Concretely, we simulate replicates $\tau_1^l,...,\tau_{k_l}^l, l = 1,...,L$, of $T_1,...,T_K$ and weight the replicates using importance sampling. Notice that, for each $l \in \{1,...,L\}$ we draw the sequence $\tau_1^l,...,\tau_{k_l}^l$ from

$$f_{T_1,...,T_{k_l}}(\tau_1,...,\tau_{k_l};\theta_p) = f_{T_1}(\tau_1;\mu_{1,p},M_p)f_{T_2-T_1}(\tau_2-\tau_1;\theta_p)f_{T_{k_l}-T_{k_l-1}}(\tau_{k_l}-\tau_{k_l-1};\theta_p)$$

in a sequential way. That is, we draw s_1^l from f_{T_1} taking $\tau_1^l = s_1^l$ then draw s_2^l from $f_{T_2-T_1}$ taking $\tau_2^l = \tau_1^l + s_2^l$ and so on (while we are in the $[0,T]$ interval). In this way we are drawing the $\tau's$ from the hitting time process and the number of hitting times from K.

Finally, we can approximate the conditional expectation of the log-likelihood of the augmented data writing,

$$\mathbb{E}_{\theta_p}\left[\log\left(L_c^K\right)|X_1,...,X_n\right] \approx \frac{1}{L}\sum_{l=1}^{L}\log\left(L_c^{k_l}(X_1,...,X_n,\tau_1^l,...,\tau_{k_l}^l)\right)\times w^l, \tag{2.5}$$

where the weight w^l is such that,

$$w^l \propto f_{X_1,...,X_n|T_1,...,T_{k_l}}(x_1,...,x_n|\tau_1^l,...,\tau_{k_l}^l)f_{T_1,...,T_{k_l}}(\tau_1^l,...,\tau_{k_l}^l).$$

2.2.2. M-step. In the M-step of the algorithm we want to maximize $Q(\theta,\theta_p)$ with respect to θ. Because of what we have done so far, this is just get the value of θ that maximizes,

$$Q_L(\theta,\theta_p) = \frac{1}{L}\sum_{l=1}^{L}\log\left(L_c^{k_l}(X_1,...,X_n,\tau_1^l,...,\tau_{k_l}^l)\right)\times w^l.$$

We start by differentiating $Q_L(\theta,\theta_p)$ with respect to θ and solving,

$$\frac{dQ_L(\theta,\theta_p)}{d\theta} = 0 \Leftrightarrow \begin{bmatrix} \frac{\partial Q_L(\theta,\theta_p)}{\partial\mu_1} \\ \frac{\partial Q_L(\theta,\theta_p)}{\partial\mu_2} \\ \frac{\partial Q_L(\theta,\theta_p)}{\partial m} \\ \frac{\partial Q_L}{\partial M} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

For μ_1,

$$\frac{\partial Q_L(\theta,\theta_p)}{\partial\mu_1} = \frac{1}{L}\sum_{l=1}^{L}\left[\frac{\partial\log\left(L_c^{k_l}(x_1,...,x_n,\tau_1^l,...,\tau_{k_l}^l;\theta)\right)}{\partial\mu_1}w^l\right]$$

Equating to zero we get,

$$\mu_1\left(\sum_{l=1}^{L}\left[a_{1,l}w^l\right]\right) - \sum_{l=1}^{L}\left[b_{1,l}w^l\right] = 0, \tag{2.6}$$

with

$$a_{1,l} = \Delta\left(j_1^l + \sum_{i=1}^{[(k_l-1)/2]}(j_{2i+1}^l - j_{2i}^l - 1) + (n - j_{k_l}^l - 1)mod(k_l+1,2)\right) + 2\tau_1^l$$
$$+ 2\sum_{i=1}^{[(k_l-1)/2]}(\tau_{2i+1}^l - \tau_{2i}^l) + \tau_{k_l}^l mod(k_l+1,2) + \sum_{i=1}^{[k_l/2]}(t_{j_{2i}^l+1} - t_{j_{2i-1}^l}) - t_{j_{k_l}^l}mod(k_l,2),$$

$$b_{1,l} = x_n mod(k_l+1,2) - 2x_0 + M(2[(k_l+1)/2]) - m(k_l-1),$$

and using the fact that $t_j = j\Delta$. For μ_2 we get,

$$\mu_2\left(\sum_{l=1}^{L}\left[a_{2,l}w^l\right]\right) - \sum_{l=1}^{L}\left[b_{2,l}w^l\right] = 0, \tag{2.7}$$

with

$$a_{2,l} = \Delta \left(\sum_{i=1}^{[k_l/2]} (j_{2i}^l - j_{2i-1}^l - 1) + (n - j_{k_l}^l - 1) mod(k_l, 2) \right) + 2 \sum_{i=1}^{[k_l/2]} (\tau_{2i}^l - \tau_{2i-1}^l)$$

$$- \tau_{k_l}^l mod(k_l, 2) + t_{j_1^l+1} + \sum_{i=1}^{[(k_l-1)/2]} (t_{j_{2i+1}^l+1} - t_{j_{2i}^l}) - t_{j_{k_l}^l} mod(k_l + 1, 2),$$

$$b_{2,l} = x_n mod(k_l, 2) + m(2[k_l/2]) - M k_l.$$

For m we get the equation,

$$(2.8) \quad \begin{aligned} & m^2 \left(\sum_{l=1}^{L} \left[(U_{m_p}^l + W^l) w^l \right] \right) \\ & + m \left(\sum_{l=1}^{L} \left[(-M(U_{m_p}^l + 2W^l) - V_{m_p}^l + \mu_1(k_l - 1) - \mu_2(2[k_l/2])) w^l \right] \right) \\ & + \left(\sum_{l=1}^{L} \left[(M(V_{m_p}^l - \mu_1(k_l - 1) + \mu_2(2[k_l/2])) - \sigma^2(k_l - 1) + M^2 W^l) w^l \right] \right) = 0. \end{aligned}$$

With

$$U_{m_p}^l = \sum_{i=1}^{[k_l/2]} \frac{\Delta}{(\tau_{2i}^l - t_{j_{2i}^l})(t_{j_{2i}^l+1} - \tau_{2i}^l)}$$

$$V_{m_p}^l = \sum_{i=1}^{[k_l/2]} \frac{x_{j_{2i}^l}(t_{j_{2i}^l+1} - \tau_{2i}^l) + x_{j_{2i}^l+1}(\tau_{2i}^l - t_{j_{2i}^l})}{(\tau_{2i}^l - t_{j_{2i}^l})(t_{j_{2i}^l+1} - \tau_{2i}^l)}$$

$$W^l = \sum_{i=2}^{k_l} \frac{1}{\tau_i^l - \tau_{i-1}^l}.$$

Doing the same with respect to M we get,

(2.9)

$$\begin{aligned} & M^3 \left(\sum_{l=1}^{L} \left[\left(-U_{M_p}^l - W^l - \frac{1}{\tau_1^l} \right) w^l \right] \right) + M^2 \left(\sum_{l=1}^{L} \left[\left((x_0 + m) U_{M_p}^l + V_{M_p}^l \right. \right. \right. \\ & \left. \left. \left. + \mu_1(2[(k_l + 1)/2]) - \mu_2 k_l + \frac{2x_0 + m}{\tau_1^l} + (x_0 + 2m) W^l \right) w^l \right] \right) \\ & + M \left(\sum_{l=1}^{L} \left[\left(-x_0 m U_{M_p}^l - (x_0 + m)(V_{M_p}^l + \mu_1(2[(k_l + 1)/2]) - \mu_2 k_l) + \sigma^2 k_l - \frac{x_0^2 + 2m x_0}{\tau_1^l} \right. \right. \right. \\ & \left. \left. \left. - (m^2 + 2x_0 m) W^l \right) w^l \right] \right) + \left(\sum_{l=1}^{L} \left[\left(x_0 m (V_{M_p}^l + \mu_1(2[(k_l + 1)/2]) - \mu_2 k_l \right. \right. \right. \\ & \left. \left. \left. + \frac{x_0}{\tau_1^l} + m W^l) - \sigma^2(m + (k_l - 1) x_0) \right) w^l \right] \right) = 0, \end{aligned}$$

with,

$$U_{M_p}^l = \sum_{i=1}^{[(k_l+1)/2]} \frac{\Delta}{(\tau_{2i-1}^l - t_{j_{2i-1}^l})(t_{j_{2i-1}^l+1} - \tau_{2i-1}^l)};$$

$$V_{M_p}^l = \sum_{i=1}^{[(k_l+1)/2]} \frac{x_{j_{2i-1}^l}(t_{j_{2i-1}^l+1} - \tau_{2i-1}^l) + x_{j_{2i-1}^l+1}(\tau_{2i-1}^l - t_{j_{2i-1}^l})}{(\tau_{2i-1}^l - t_{j_{2i-1}^l})(t_{j_{2i-1}^l+1} - \tau_{2i-1}^l)}.$$

Finally, we only need to solve the equations to get θ_{p+1} and start the $(p+1)$-iteration of the algorithm. If we consider the special case $\mu_1 = -\mu_2 = \mu$, then the equations are as following. For μ

$$\begin{aligned}\mu\left(\sum_{l=1}^{L}\left[\left(n\Delta+\tau_{k_l}^l\right)w^l\right]\right) - \sum_{l=1}^{L}\big[((2k_l-2)(M-m)+2(M-x_0) \\ +(M-x_n)mod(k_l,2)+(x_n-m)mod(k_l+1,2))w^l\big] = 0,\end{aligned} \tag{2.10}$$

for m,

$$\begin{aligned} & m^2\left(\sum_{l=1}^{L}\left[(U_{m_p}^l+W^l)w^l\right]\right) \\ & + m\left(\sum_{l=1}^{L}\left[(-M(U_{m_p}^l+2W^l)-V_{m_p}^l+\mu[2[k_l/2]+k_l-1])w^l\right]\right) \\ & + \left(\sum_{l=1}^{L}\left[(M(V_{m_p}^l-\mu[2[k_l/2]+k_l-1])-\sigma^2(k_l-1)+M^2W^l)w^l\right]\right) = 0,\end{aligned} \tag{2.11}$$

and for M we get,

(2.12)

$$\begin{aligned} & M^3\left(\sum_{l=1}^{L}\left[\left(-U_{M_p}^l - W^l - \frac{1}{\tau_1^l}\right)w^l\right]\right) + M^2\left(\sum_{l=1}^{L}\left[\left((x_0+m)U_{M_p}^l + V_{M_p}^l\right.\right.\right. \\ & \left.\left.\left. +\mu(2[(k_l+1)/2]+k_l) + \frac{2x_0+m}{\tau_1^l} + (x_0+2m)W^l\right)w^l\right]\right) \\ & + M\left(\sum_{l=1}^{L}\left[\left(-x_0 m U_{M_p}^l - (x_0+m)(V_{M_p}^l+\mu(2[(k_l+1)/2]+k_l)) + \sigma^2 k_l - \frac{x_0^2+2mx_0}{\tau_1^l}\right.\right.\right. \\ & \left.\left.\left. -(m^2+2x_0m)W^l\right)w^l\right]\right) + \left(\sum_{l=1}^{L}\left[\left(x_0m(V_{M_p}^l+\mu(2[(k_l+1)/2]+k_l)\right.\right.\right. \\ & \left.\left.\left. +\frac{x_0}{\tau_1^l} + mW^l) - \sigma^2(m+(k_l-1)x_0)\right)w^l\right]\right) = 0.\end{aligned}$$

Remark 2.11. To start the algorithm we need initial values for μ_1, μ_2, m and M. We can choose $\widehat{m}$ and $\widehat{M}$ as the α and $1-\alpha$ empirical quantile ($\alpha = 20\%$ for instance). After we fixed the thresholds one way to get values for $\tau_1, ..., \tau_{k_0}$ is choosing q_1 as the first

element in $\{1, ..., n\}$ such that $X_{q_1} \leq \widehat{M}$, $X_{q_1+1} > \widehat{M}$ and $\tau_1 = \lambda t_{q_1} + (1-\lambda)t_{q_1+1}$ for some $\lambda \in [0,1]$; q_2 the first element in $\{q_1+1, ..., n\}$ such that $X_{q_2} \geq \widehat{m}$, $X_{q_2+1} < \widehat{m}$ and $\tau_2 = \lambda t_{q_2} + (1-\lambda)t_{q_2+1}$. More generally, q_{2i+1} the first element in $\{q_{2i}+1, ..., n\}$ such that $X_{q_{2i+1}} \leq \widehat{M}$, $X_{q_{2i+1}+1} > \widehat{M}$ and $\tau_{2i+1} = \lambda t_{q_{2i+1}} + (1-\lambda)t_{q_{2i+1}+1}$, and q_{2i} the first element in $\{q_{2i-1}+1, ..., n\}$ such that $X_{q_{2i}} \geq \widehat{m}$, $X_{q_{2i}+1} < \widehat{m}$ and $\tau_{2i} = \lambda t_{q_{2i}} + (1-\lambda)t_{q_{2i}+1}$. Finally, we get $\widehat{\mu}_{1,0}$ and $\widehat{\mu}_{2,0}$ using the equations (2.6) and (2.7) in an appropriate way.

To implement the MCEM-algorithm in the way we do, it is needed to know the density of the hitting times and the density of the process observations conditioned to the hitting times. This is the justification for the generalizations of the hitting times density in the previous section being possible to get similar equations for the geometric Brownian motion or other processes remaining the problem of getting the true transition densities.

2.3. Simulation

In this section we implement a simulation study in order to test the estimation method. We start simulating 170 observations of the threshold model with $\Delta = .1$. The process is divided in two regimes, the first one being Brownian motion with the drift $\mu_1 = 1$ and the second Brownian motion with drift $\mu_2 = -1$, that is, the change in regime results from the change in the sign of the drift parameter, both regimes have a value of $\sigma = .4$. The change in regime happens when the process hits the upper threshold $M = 3$ if the process is in the first regime, or when the process hits the lower threshold $m = -3$ if the process is in the second regime. The process starts in the first regime with initial value $x_0 = 0$ and the trajectory is built summing (pseudo)random numbers generated from a normal distribution with parameters $\mu_1\Delta$ and $\sigma^2\Delta$ for the first regime and random numbers generated from a normal distribution with parameters $\mu_2\Delta$ and $\sigma^2\Delta$ for the second regime. The initial values for the parameters are computed, from the generated data, following remark 2.11 and are given by $\widehat{\mu}_{1,1} = 0.637295 = -\widehat{\mu}_{2,1}$, $\widehat{m}_1 = -1.14301$, and $\widehat{M}_1 = 1.99292$. At the p-th step of the algorithm, we compute $L = 500$ repetitions of hitting times sequences in order to approximate the Monte Carlo expectation (2.5). To draw the sequence of hitting times we use numerical methods to find the root τ of the equation,

$$\int_0^\tau \frac{\widehat{M}_p - \widehat{m}_p}{\sqrt{2\pi\sigma^2 t^3}} \, e^{-\frac{(\widehat{M}_p - \widehat{m}_p - \widehat{\mu}_{1,p} t)^2}{2\sigma^2 t}} dt = u$$

with u random number draw from computer $U[0,1]$ generator. Solutions of the last equation can be used to draw the hitting times not only from regime 1 to regime 2 but

also the opposite hitting times, because the difference between regimes is only in the sign of the drift parameter. From the last equation we draw values for the difference between consecutive hitting times and because of that, we need to sum these values to get the sequence of hitting times. Alternatively, it is possible to use the accept-reject method choosing for density candidate the gamma density with an appropriate choice of parameters. After the 500 repetitions are completed we can compute the $p+1$-th estimators $\widehat{\mu}_{1,p+1}, \widehat{m}_{p+1}$ and $\widehat{M}_{p+1}$ from equation (2.10), (2.11) and (2.12). The final values for the estimators, after 100 iterations of the algorithm, are $\widehat{\mu}_{1,101} = 1.05 = -\widehat{\mu}_{2,101}, \widehat{m}_{101} = -2.85$, and $\widehat{M}_{101} = 2.9$. Being the thresholds underestimated (in absolute value) and the drift overestimated. Notice that, if we estimate the drift parameter for the simulated data using the extra information about the regime for each observation, we get $\widehat{\mu} > 1$ and this can explain why, from the algorithm, we get an overestimated value $\widehat{\mu}_{1,101}$. The simulation was written in the Mathematica 4.1 program and the list of the instructions appears in appendix 2. One hundred iterations of the algorithm have been considered and the estimates are presented in the figures 2.1 , 2.2 and 2.3.

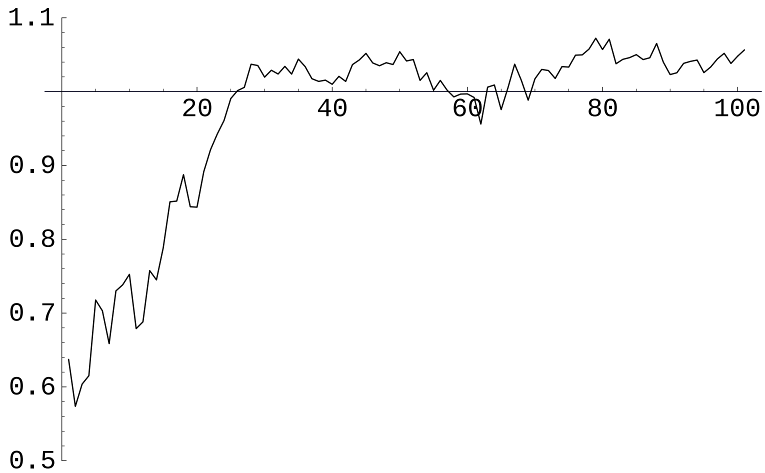

FIGURE 2.1. Estimated values for μ_1 from 100 iterations of MCEM

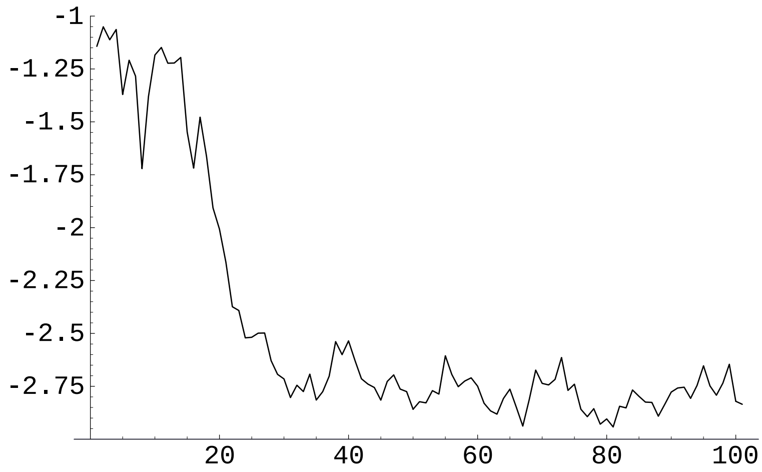

FIGURE 2.2. Estimated values for m from 100 iterations of MCEM

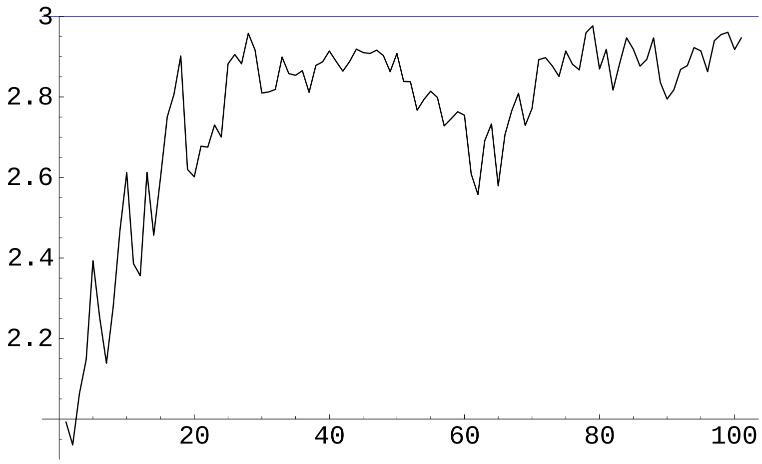

FIGURE 2.3. Estimated values for M from 100 iterations of MCEM

Repeating the procedure for 100 trajectories with the same conditions and considering the mean and standard deviation for the final estimators we get the following results:

TABLE 2.1. Mean and standard deviation for the drift and the threshold estimates from 100 replicates of MCEM

	$\widehat{\mu}$	$\widehat{m}$	$\widehat{M}$
mean	1.020	-2.940	2.950
s.d.	0.118	0.186	0.174

This method needs many computations and for that reason is computer time consuming, being impossible to increase the number of repetitions or the length of each trajectory in a personal computer. Once again, it should be possible to use the accept-reject method to simulate the hitting times, in order to reduce the computer time. We also apply the method to models with bigger volatility parameter but for the same order of repetitions we

have worst results and the only way to get them better should be increasing the number of repetitions with the handicap of time spending.

CHAPTER 3

LSE for Brownian motion with drift threshold model

In the previous chapter we implemented a MCEM algorithm in order to estimate the threshold parameters in the Brownian motion with drift threshold model based on a natural approximation for the transition density. Supposing that the transition density is the real one, we deduced the implementation of the algorithm but have not studied the properties of the estimators. In this chapter we will introduce and study another estimating procedure and get partial asymptotic results about the threshold estimators. In order to do this, we will implement a conditional least squares estimator procedure for the Brownian motion with drift threshold model and we will try to get more information about the law of the process in each regime in order to deduce the conditional expectation in the estimation procedure. We want to show that the conditional least square threshold estimator for the model,

$$dXt = \mu(t)dt + \sigma dB_t, \quad X_0 = x_0, \tag{3.1}$$

where

$$\mu(t) = \sum_{k\geq 0} \left[\mu_0 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) - \mu_0 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_0 > 0$$

and $0 = \tau_0 < \tau_1 < \ldots < \tau_j < \ldots$ are the threshold hitting times, is consistent. Our goal is to prove the consistency when the process is discretely observed and the regime to each observation belongs is unknown. We give some steps in order to prove the consistency. We will start with a first section where we sketch the proof of the consistency of the thresholds estimators in the conditions of decreasing discretization step size Δ and knowing the regime to which each observation of the process belongs. In the second section we will abandon the restriction of knowing the regimes for which each observation belongs to and we will introduce in the estimation procedure a practical way of computing the conditional regime classification, no asymptotical results on consistency will be proved for this case and a simulation study will be carried out. The third section is devoted to an implementation of a simulation study where we will simulate data from the more general models as the ones built from Geometric Brownian motion or the Ornstein-Uhlenbeck process and then we will implement the estimation procedure in order to estimate the

upper and lower thresholds. We also apply the estimating procedure to real financial data considering the three models Brownian motion with drift, Geometric Brownian motion and Ornstein-Uhlenbeck process.

3.1. Discretely observed process with known regimes

Consider the observations $X_1, \ldots, X_n$ of the Brownian motion threshold process in the time interval $[0, t_n]$ and introduce the r.v $R_1, \ldots, R_n$ where $R_i = 1$ if X_i is in regime 1, while $R_i = 2$ if X_i is in regime 2. Let $\hat{K}_i$ be the number of changes of value of the sequence $R_1, \ldots R_i$, we will assume in this section that we observe the values of R_i.

Then, we can define the estimator $\widehat{\mu}_i(m, M)$ of μ_0 in a natural way, as:

$$\widehat{\mu}_i(m, M) = \frac{\widehat{K}_i}{t_i}(M - m). \tag{3.2}$$

We have the following results.

LEMMA 3.1. With K_i the true number of changes in regime in $[0, t_i]$, we have:

$$\lim_{i \to +\infty} \frac{K_i}{t_i} = \frac{\mu_0}{M_0 - m_0}, \quad a.s.$$

PROOF. Let us write $\tau_1, ..., \tau_{K_i}$ for the thresholds hitting times. The difference between regimes is just the sign of the drift parameter and it is easy to conclude from the initial sections in chapter 2 that $\tau_{j+1} - \tau_j, j = 1, ...$ is a sequence of independent identically distributed random variables with mean $\frac{M_0 - m_0}{\mu_0}$. Following the same idea as in the proof of the strong law of large numbers for renewal processes, looking at K_i as $K_i = \max\{n : \tau_n \le t_i\}$, we have:

$$\tau_1 + \sum_{j=1}^{K_i - 1} (\tau_{j+1} - \tau_j) = \tau_{K_i} \le t_i < \tau_{K_i + 1} = \tau_1 + \sum_{j=1}^{K_N} (\tau_{j+1} - \tau_j), \forall i$$

and so,

$$\frac{\tau_1 + \sum_{j=1}^{K_i - 1} (\tau_{j+1} - \tau_j)}{K_i} \le \frac{t_i}{K_i} < \frac{\tau_1 + \sum_{j=1}^{K_i} (\tau_{j+1} - \tau_j)}{K_i}.$$

Now since $0 < \mathbb{E}[\tau_{j+1} - \tau_j] = \frac{M_0 - m_0}{\mu_0} < \infty$ we have:

$$K_i \to \infty$$

and from the strong law of large numbers,

$$\lim_{i \to \infty} \frac{\tau_1 + \sum_{j=1}^{K_i} (\tau_{j+1} - \tau_j)}{K_i} = \lim_{i \to \infty} \frac{1}{K_i} \sum_{j=1}^{K_i} (\tau_{j+1} - \tau_j) = \frac{M_0 - m_0}{\mu_0}.$$

In the same way we get,

$$\lim_{i\to\infty} \frac{\tau_1 + \sum_{j=1}^{K_i-1}(\tau_{j+1} - \tau_j)}{K_i} = \frac{M_0 - m_0}{\mu_0},$$

and that completes the proof. □

We also have a similar result on convergence for the ratio $\frac{\hat{K}_i}{t_i}$.

Notice that a change in regime is not detected, when we compute $\hat{K}_i$, in one of the following cases:

(1) we have in the sequence $..., X_j \in R_1, X_{j+1} \in R_1, ...$ but we have $2p$, for some $p \geq 1$, changes in regime in the interval $]t_j, t_{j+1}[$,

(2) we have in the sequence $..., X_j \in R_1, X_{j+1} \in R_2, ...$ but we have $2p + 1$, for some $p \geq 1$, changes in regime in the interval $]t_j, t_{j+1}[$.

In both cases more than one change in regime in $]t_j, t_{j+1}[$ is needed.

Lemma 3.2. Being $\hat{K}_i$ the number of changes of value of the sequence $R_1, \ldots R_i$. We have that $\lim_{i\to\infty} \frac{K_i - \hat{K}_i}{t_i} = 0$ almost surely.

Proof. Let us consider

$$E_n = \{\omega : R_n = R_{n+1}, \text{ but } \exists t \in [t_n, t_{n+1}[\text{ with } R_t \neq R_n\},$$

the event corresponding to the first situation of change in regime in $[t_n, t_{n+1}]$ not detected. With $\Delta_n = t_{n+1} - t_n$, we begin by checking that:

$$\mathbb{P}[E_n] < \frac{2}{\sqrt{2\pi}} \frac{e^{-\frac{(M_0 - m_0 - \mu_0\Delta_n)^2}{2\Delta_n}}}{(M_0 - m_0 - \mu_0\Delta_n)} \sqrt{\Delta_n}. \tag{3.3}$$

This is easy to prove if we first notice that,

(3.4)

$$E_n \subseteq \{\text{more than one change of regime in } [t_n, t_{n+1}[\wedge R_n = 1\} \subseteq \{\min_{0\leq s\leq\Delta_n} B_s^{-\mu_0} < m_0 - M_0\}$$

$$\subseteq \{\min_{0\leq s\leq\Delta_n} B_s^0 < m_0 - M_0 + \mu_0\Delta_n\} \stackrel{\mathcal{L}}{=} \{\max_{0\leq s\leq\Delta_n} B_s^0 > M_0 - m_0 - \mu_0\Delta_n\}$$

or

(3.5)

$$E_n \subseteq \{\text{more than one change of regime in } [t_n, t_{n+1}[\wedge R_n = 2\} \subseteq \{\max_{0\leq s\leq\Delta_n} B_s^{\mu_0} > M_0 - m_0\}$$

$$\subseteq \{\max_{0\leq s\leq\Delta_n} B_s^0 > M_0 - m_0 - \mu_0\Delta_n\}$$

where $(B_s^{\mu_0})_{s\geq 0}$ denotes the Brownian motion with drift (drift coefficient μ_0, note in particular that $(B_s^0)_{s\geq 0}$ is the standard Brownian motion) starting from 0. Then we can write

$$\mathbb{P}[E_n] \leq \mathbb{P}[\max_{0\leq s\leq \Delta_n} B_s^0 > M_0 - m_0 - \mu_0\Delta_n],$$

and we use the law of the maximum of the Brownian motion (see chapter 2, proposition 2.2) together with the well known inequality,

$$\forall x \text{ let } Z \sim \mathcal{N}(0,1), \quad \mathbb{P}(Z \geq x) < \frac{1}{\sqrt{2\pi}} \frac{e^{-x^2/2}}{x}, \tag{3.6}$$

to get

$$\mathbb{P}[E_n] \leq \frac{2}{\sqrt{2\pi}} \frac{e^{-\frac{(M_0-m_0-\mu_0\Delta_n)^2}{2\Delta_n}}}{(M_0 - m_0 - \mu_0\Delta_n)} \sqrt{\Delta_n}.$$

Then, from the previous bound, we get, that exists constants $c_1, c_2 > 0$ such that,

$$\sum_{n=1}^{\infty} \mathbb{P}[E_n] \leq c_1 \sum_{n=1}^{\infty} e^{-\frac{c_2}{\Delta_n}}. \tag{3.7}$$

For the sum to be finite it is enough to consider a decreasing Δ_n as for instance $\Delta_n = 1/n$. Finally, from Borel-Cantelli lemma ([**Adams, 1996**]) we get that,

$$P[\limsup E_n] = \mathbb{P}[E_n; \text{infinitely often}] = 0.$$

We get the same result for the second kind of changes in regime in $[t_n, t_{n+1}]$ that are not detected, and we can conclude that

$$\lim_{i\to\infty} \frac{\hat{K}_i}{t_i} = \lim_{i\to\infty} \frac{K_i}{t_i}, \; a.s.$$

□

We also have.

LEMMA 3.3. Defining $\hat{\mu}_i = \hat{\mu}_i(m_0, M_0)$ as in equation (3.2), when m_0 and M_0 are the true values of the thresholds, we have:

$$\lim_{i\to+\infty} \hat{\mu}_i = \mu_0, \quad a.s.$$

PROOF. From (3.2) we can write,

$$\hat{\mu}_i = \frac{\hat{K}_i(M_0 - m_0)}{t_i}, \tag{3.8}$$

then from the previous lemmas we get,

$$\lim_{i\to\infty} \frac{\hat{K}_i(M_0 - m_0)}{t_i} = \mu_0,$$

and that completes the proof of this lemma. $\square$

Ideally, we define the conditional least square errors function as,

$$LSE_n(m, M) = \sum_{i=1}^{n-1} \left(X_{i+1} - \mathbb{E}_{\mu,m,M}\left[X_{i+1} | X_i, ..., X_1 \right] \right)^2,$$

but we do not know the conditional expectation for this process and we will therefor consider an approximation of that expectation, and introduce the auxiliary conditional least squares contrast:

(3.9)

$$LS_n(m, M) = \sum_{i=1}^{n-1} (X_{i+1} - X_i - \hat{\mu}_i(m, M)\Delta)^2 1_{R_i = R_{i+1} = 1} + \sum_{i=1}^{n-1} (X_{i+1} - X_i + \hat{\mu}_i(m, M)\Delta)^2 1_{R_i = R_{i+1} = 2},$$

where, as before,

$$\hat{\mu}_i(m, M) = \frac{\hat{K}_i}{t_i}(M - m). \tag{3.10}$$

In order to justify this auxiliary conditional least squares contrast, we prove the following lemma.

LEMMA 3.4. We have:

$$\mathbb{E}_{\mu_0,m_0,M_0}[X_{i+1} - X_i | X_i, R_i] 1_{R_i = R_{i+1} = 1} = \mu_0 \Delta 1_{R_i = R_{i+1} = 1} + A(m_0, M_0, \Delta),$$

where

$$\mathbb{E}\left[A(m_0, M_0, \Delta) \right] < 2\mu_0 \Delta \times \left[\frac{2\sqrt{\Delta}}{\sqrt{2\pi}} \frac{e^{-\frac{(M_0 - m_0 - \mu_0 \Delta)^2}{2\Delta}}}{(M_0 - m_0 - \mu_0 \Delta)} \right]^{1/2}.$$

PROOF. On the event $R_i = R_{i+1} = 1$, we have,

$$\begin{aligned} X_{i+1} - X_i &= (X_{i+1} - X_i) 1_{\text{no change of regime}} + (X_{i+1} - X_i) 1_{\text{more than one change of regime}} \\ &= (B_\Delta^0 + \mu_0 \Delta) 1_{\text{no change of regime}} + (X_{i+1} - X_i) 1_{\text{more than one change of regime}} \\ &= B_\Delta^0 + \mu_0 \Delta + (X_{i+1} - X_i - B_\Delta^0 - \mu_0 \Delta) 1_{\text{more than one change of regime}}. \end{aligned} \tag{3.11}$$

We need to compute $\mathbb{E}[(X_{i+1} - X_i - B_\Delta^0 - \mu_0 \Delta) 1_{\text{more than one change of regime}}]$. We have from Cauchy-Schwarz,

(3.12)

$$\left| \mathbb{E}[(X_{i+1} - X_i - B_\Delta^0 - \mu_0 \Delta) 1_{\text{more than one change of regime}}] \right| \leq \mathbb{E}\left[\left((X_{i+1} - X_i - B_\Delta^0 - \mu_0 \Delta) \right)^2 \right]^{\frac{1}{2}} \times$$
$$\times (\mathbb{P}[\text{more than one change of regime}])^{\frac{1}{2}}.$$

Now

$$\mathbb{E}\left[\left((X_{i+1} - X_i - B_\Delta^0 - \mu_0 \Delta) \right)^2 \right]^{1/2} \leq 2\mu_0 \Delta, \tag{3.13}$$

because we always have that,

$$X_i - \mu_0\Delta + B^0_\Delta \leq X_{i+1} \leq X_i + \mu_0\Delta + B^0_\Delta \Leftrightarrow -2\mu_0\Delta \leq X_{i+1} - X_i - \mu_0\Delta - B^0_\Delta \leq 0. \tag{3.14}$$

We also have,

$$\mathbb{P}[\text{more than one change of regime}] \leq 2\frac{1}{\sqrt{2\pi}}\frac{e^{-(M_0-m_0-\mu_0\Delta)^2/2\Delta}}{(M_0-m_0-\mu_0\Delta)}\sqrt{\Delta}, \tag{3.15}$$

as shown in the proof of lemma 3.2. We deduce the required inequality by combining equations (3.11), (3.12), (3.13) and (3.15). □

From the auxiliary conditional least squares contrast equation (3.9) we can estimate the difference $M - m$.

LEMMA 3.5. The estimator of the difference $M - m$ is given by:

$$\widehat{M-m} = \frac{\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i)1_{R_i=R_{i+1}=1} - \sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i)1_{R_i=R_{i+1}=2}}{\Delta\sum_{i=1}^{n-1}\frac{\hat{K}_i^2}{t_i^2}1_{R_i=R_{i+1}}} \tag{3.16}$$

PROOF. It is enough to compute,

$$\begin{aligned}
&\frac{\partial LS_n(m,M)}{\partial(M-m)} = 0 \\
\Leftrightarrow &-\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}\left(X_{i+1}-X_i-\frac{\hat{K}_i}{t_i}(M-m)\Delta\right)1_{R_i=R_{i+1}=1} \\
&+\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}\left(X_{i+1}-X_i+\frac{\hat{K}_i}{t_i}(M-m)\Delta\right)1_{R_i=R_{i+1}=2} = 0 \\
\Leftrightarrow &-\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i)1_{R_i=R_{i+1}=1} + \sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i)1_{R_i=R_{i+1}=2} \\
&+(M-m)\Delta\sum_{i=1}^{n-1}\frac{\hat{K}_i^2}{t_i^2}\left(1_{R_i=R_{i+1}=1}+1_{R_i=R_{i+1}=2}\right) = 0 \\
\Leftrightarrow &M-m = \frac{\left(\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i)1_{R_i=R_{i+1}=1} - \sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i)1_{R_i=R_{i+1}=2}\right)}{\Delta\sum_{i=1}^{n-1}\frac{\hat{K}_i^2}{t_i^2}1_{R_i=R_{i+1}}}.
\end{aligned} \tag{3.17}$$

□

Now we can write the proof of the consistency of the estimators, solution of the auxiliary conditional least squares contrast.

THEOREM 3.6. When $\Delta_n \to 0$ and $n\Delta_n \to +\infty$, the estimator $\widehat{M-m}$ that minimizes the $LS_n(m, M)$ function is consistent.

PROOF. It is easy to obtain the decomposition:

(3.18)

$$LS_n(m, M) = LS_n(m_0, M_0) - (M - m - (M_0 - m_0)) \times A_1(n) + (M - m - (M_0 - m_0))^2 A_2(n),$$

where

$$A_1(n) = 2\Delta \sum_{i=1}^{n-1} \frac{\hat{K}_i}{t_i} \{(X_{i+1} - X_i - \hat{\mu}_i(m_0, M_0)\Delta) 1_{R_i = R_{i+1} = 1} - (X_{i+1} - X_i + \hat{\mu}_i(m_0, M_0)\Delta) 1_{R_i = R_{i+1} = 2}\}$$

$$A_2(n) = \Delta^2 \sum_{i=1}^{n-1} \frac{\hat{K}_i^2}{t_i^2} 1_{R_i = R_{i+1}}.$$

We therefore deduce that the estimator of the difference $M - m$ that minimizes $LS_n(m, M)$ can also be written as,

$$\widehat{M-m} = M_0 - m_0 + 2\frac{A_1(n)}{A_2(n)}.$$

From lemma 3.2 we have:

$$\frac{\hat{K}_i}{t_i} \to \frac{\mu_0}{M_0 - m_0} \quad a.s. \tag{3.19}$$

Taking Césaro means ([**Williams, 1991**]), we deduce that

$$\frac{1}{n\Delta^2} A_2(n) \to \frac{\mu_0^2}{(M_0 - m_0)^2}, \quad a.s. \tag{3.20}$$

because we can write,

$$\frac{1}{n\Delta^2} A_2(n) = \frac{1}{n} \sum_{i=1}^{n-1} \frac{\hat{K}_i^2}{t_i^2} 1_{R_i = R_{i+1}} = \frac{b_n}{n} \times \frac{1}{b_n} \sum_{i=1}^{n-1} (b_{i+1} - b_i) \frac{\hat{K}_i^2}{t_i^2}$$

with $b_n = \sum_{i=1}^{n-1} 1_{R_i = R_{i+1}}$ and because

$$\frac{b_n}{n} = \frac{n - \sum_{i=1}^{n-1} 1_{R_i \neq R_{i+1}}}{n} = 1 - \frac{\hat{K}_n}{n} = 1 - \frac{\hat{K}_n}{t_n} \frac{t_n}{n} = 1 - \frac{\hat{K}_n}{t_n} \Delta$$

and once more we just need to use the fact that $\frac{\hat{K}_n}{t_n} \to \frac{\mu_0}{M_0 - m_0}$ and then when $\Delta \to 0$ we get the desired result.

As a consequence, to conclude the consistency of $\widehat{M-m}$, we are left to prove

$$\frac{1}{n\Delta^2} A_1(n) \to 0. \tag{3.21}$$

We will prove

$$\frac{1}{n\Delta} \sum_{i=1}^{n-1} \frac{\hat{K}_i}{t_i} (X_{i+1} - X_i - \hat{\mu}_i(m_0, M_0)\Delta) 1_{R_i = R_{i+1} = 1} \xrightarrow[n\to\infty]{\mathbb{P}_{\theta_0}} 0, \tag{3.22}$$

the proof being similar for the second term in $A_1(n)$.

Let us write

$$\frac{1}{n\Delta}\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i-\hat{\mu}_i(m_0,M_0)\Delta)1_{R_i=R_{i+1}=1}=A_{11}(n)+A_{12}(n),$$

where

$$\begin{aligned} A_{11}(n) &= \frac{1}{n\Delta}\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i-\mu_0\Delta)1_{R_i=R_{i+1}=1} \\ A_{12}(n) &= \frac{1}{n\Delta}\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}\Delta\left(\mu_0-\hat{\mu}_i(m_0,M_0)\right)1_{R_i=R_{i+1}=1}. \end{aligned}$$

$A_{12}(n)$ converges to zero *a.s.* as deduced from lemmas 3.2 and 3.3, after taking the Césaro mean.

As for $A_{11}(n)$, we introduce

$$\zeta_i=\frac{1}{n\Delta}\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i-\mu_0\Delta)1_{R_i=R_{i+1}=1}.$$

To deduce the convergence of $A_{11}(n)$ to zero, it is enough to prove, see Lemma 9 of [**Genon-Catalot & Jacod, 1993**],

$$\sum_{i=1}^{n-1}\mathbb{E}[\zeta_i|\mathcal{F}_i]\xrightarrow[n\to\infty]{\mathbb{P}_{\theta_0}}0, \tag{3.23}$$

$$\sum_{i=1}^{n-1}\mathbb{E}[\zeta_i^2|\mathcal{F}_i]\xrightarrow[n\to\infty]{\mathbb{P}_{\theta_0}}0. \tag{3.24}$$

To prove (3.23), we use lemma 3.4 to obtain

$$\sum_{i=1}^{n-1}\mathbb{E}[\zeta_i|\mathcal{F}_i]\le\frac{2\mu_0}{n}\sum_{i=1}^{n-1}\frac{\hat{K}_i}{t_i}\left[\frac{2}{\sqrt{2\pi}}\frac{e^{-\frac{(M_0-m_0-\mu_0\Delta)^2}{2\Delta}}}{(M_0-m_0-\mu_0\Delta)}\sqrt{\Delta}\right]^{1/2}$$

because,

$$\begin{aligned} \mathbb{E}[\zeta_i|\mathcal{F}_i] &= \frac{1}{n\Delta}\mathbb{E}\left[\frac{\hat{K}_i}{t_i}(X_{i+1}-X_i-\mu_0\Delta)1_{R_i=R_{i+1}=1}\middle|\mathcal{F}_i\right]= \\ &= \frac{1}{n\Delta}\frac{\hat{K}_i}{t_i}\mathbb{E}\left[(X_{i+1}-X_i-\mu_0\Delta)1_{R_i=R_{i+1}=1}|\mathcal{F}_i\right]=\frac{1}{n\Delta}\frac{\hat{K}_i}{t_i}A(m_0,M_0,\Delta) \end{aligned}$$

and then the sum in (3.23) goes to zero when Δ goes to zero because of the exponential bound. To prove the convergence in (3.24), we can get, from the proof of lemma 3.4,

equation (3.11), the decomposition,

$$\sum_{i=1}^{n-1}\mathbb{E}[\zeta_i^2|\mathcal{F}_i] = \frac{1}{n^2\Delta^2}\sum_{i=1}^{n-1}\mathbb{E}\left[\left(\frac{\hat{K}_i}{t_i}\right)^2 (X_{i+1}-X_i-\mu_0\Delta)^2 1_{R_i=R_{i+1}=1}\middle|\mathcal{F}_i\right] =$$

$$=\frac{1}{n^2\Delta^2}\sum_{i=1}^{n-1}\left(\frac{\hat{K}_i}{t_i}\right)^2\mathbb{E}\Big[\big(B_\Delta^0+(X_{i+1}-X_i-B_\Delta^0-\mu_0\Delta)1_{\text{more than one change of regime}}\big)^2|\mathcal{F}_i\Big]$$

and then use the Cauchy-Schwarz inequality to get,

$$\leq \frac{2}{n^2\Delta}\sum_{i=1}^{n-1}\left(\frac{\hat{K}_i}{t_i}\right)^2 + \frac{2}{n^2\Delta^2}\sum_{i=1}^{n-1}\left(\frac{\hat{K}_i}{t_i}\right)^2 (2\mu_0\Delta)^2\left[\frac{2\sqrt{\Delta}}{\sqrt{2\pi}}\frac{e^{-\frac{(M_0-m_0-\mu_0\Delta)^2}{2\Delta}}}{(M_0-m_0-\mu_0\Delta)}\right]^{1/2}$$

$$= \frac{2}{n\Delta}\frac{1}{n}\sum_{i=1}^{n-1}\left(\frac{\hat{K}_i}{t_i}\right)^2 + \frac{2}{n^2}\sum_{i=1}^{n-1}\left(\frac{\hat{K}_i}{t_i}\right)^2 (2\mu_0)^2\left[\frac{2\sqrt{\Delta}}{\sqrt{2\pi}}\frac{e^{-\frac{(M_0-m_0-\mu_0\Delta)^2}{2\Delta}}}{(M_0-m_0-\mu_0\Delta)}\right]^{1/2}$$

and this concludes the proof by taking Césaro's mean and making $n\to\infty$. □

Notice that the contrast (3.9) only depends on M and m through the difference $M-m$. One way to extend the dependence, in order to estimate M and m, could be to expand further the conditional expectation in such a way that incorporate new terms depending on each threshold, separately. This is done in the following lemma:

LEMMA 3.7. Define for a Brownian motion with drift μ, starting from 0, τ_b^μ to be the first hitting time of b. We have:

$$\mathbb{E}_{\mu_0,m_0,M_0}[(X_{i+1}-X_i)|X_i,R_i]1_{R_i=1} = \mu_0\Delta 1_{R_i=1} + 2\mu_0\mathbb{E}[(\tau_b^{\mu_0}-\Delta)1_{0<\tau<\Delta}]_{b=M_0-X_i}1_{R_i=1} + A(m_0,M_0,\Delta),$$

where, for all $x<M_0$ and Δ,

$$\mathbb{E}[A(m_0,M_0,\Delta)] < 2\mu_0\Delta\times\left[\frac{2}{\sqrt{2\pi}}\frac{e^{-\frac{(M_0-m_0-\mu_0\Delta)^2}{2\Delta}}}{(M_0-m_0-\mu_0\Delta)}\sqrt{\Delta}\right]^{1/2}.$$

Remarks.

(1) It is not easy to obtain an analytical expression for the second term in the expansion: $\mathbb{E}[(\tau_b^{\mu_0}-\Delta)1_{0<\tau<\Delta}]_{b=M_0-X_i}$.

(2) It is important to note that when $X_i=M_0$, the second term of the expansion

$$2\mu_0\mathbb{E}[(\tau_b^{\mu_0}-\Delta)1_{0<\tau<\Delta}]_{b=0} = -2\mu_0\Delta.$$

Proof. We have

$$\begin{aligned}X_{i+1}-X_i =&(X_{i+1}-X_i)1_{\text{no change of regime}}+(X_{i+1}-X_i)1_{\text{exactly one change of regime}}\\&+(X_{i+1}-X_i)1_{\text{more than one change of regime}}.\end{aligned}$$

On the event $R_i=1$, we then decompose:

$$\begin{aligned}X_{i+1}-X_i =&B^0_\Delta+\mu_0\Delta 1_{\text{no change of regime}}+[\mu_0\tau_b^{\mu_0}-\mu_0(\Delta-\tau_b^{\mu_0})]1_{\text{one change of regime}}\\&+(X_{i+1}-X_i-B^0_\Delta)1_{\text{more than one change of regime}}\end{aligned}$$

which can be rewritten

$$\begin{aligned}X_{i+1}-X_i =&B^0_\Delta+\mu_0\Delta+2[\mu_0(\tau_b^{\mu_0}-\Delta)]1_{\text{one change of regime}}\\&+(X_{i+1}-X_i-\mu_0\Delta-B^0_\Delta)1_{\text{more than one change of regime}}.\end{aligned}$$

We are finished by proving that

$$\mathbb{E}[(X_{i+1}-X_i-\mu_0\Delta-B^0_\Delta)1_{\text{more than one change of regime}}]<2\mu_0\Delta\left[\frac{2}{\sqrt{2\pi}}\frac{e^{-\frac{(M_0-m_0-\mu_0\Delta)^2}{2\Delta}}}{(M_0-m_0-\mu_0\Delta)}\sqrt{\Delta}\right]^{\frac{1}{2}},$$

which is not difficult using Cauchy-Schwarz and the fact that,

$$\mathbb{P}[\text{more than one change of regime}]\le\frac{2}{\sqrt{2\pi}}\frac{e^{-\frac{(M_0-m_0-\mu_0\Delta)^2}{2\Delta}}}{(M_0-m_0-\mu_0\Delta)}\sqrt{\Delta}.$$

as we proved before. □

With this new lemma we can define a new auxiliary conditional least squares contrast with terms that depend on each of the thresholds opening the possibility of defining consistent estimators for each of the thresholds.

(3.25)

$$\begin{aligned}LS_n(m,M)=&\sum_{i=1}^{n-1}(X_{i+1}-X_i-\hat\mu_i(m,M)\Delta-2\hat\mu_i\mathbb{E}[(\tau_b^{\hat\mu_i}-\Delta)1_{0<\tau<\Delta}]_{b=M-X_i})^2 1_{R_i=1}\\&+\sum_{i=1}^{n-1}(X_{i+1}-X_i+\hat\mu_i(m,M)\Delta+2\hat\mu_i\mathbb{E}[(\tau_b^{-\hat\mu_i}-\Delta)1_{0<\tau<\Delta}]_{b=m-X_i})^2 1_{R_i=2}.\end{aligned}$$

In the terms $\mathbb{E}[(\tau_b^{\hat\mu_i}-\Delta)1_{0<\tau<\Delta}]_{b=M-X_i}$ and $\mathbb{E}[(\tau_b^{-\hat\mu_i}-\Delta)1_{0<\tau<\Delta}]_{b=m-X_i}$, m and M are not explicit and so the previous approach to get the estimator for the thresholds difference and the ideas for the proof of the consistency are not applicable. For now we have no complete way of defining threshold estimators. However, in order to get some practical

way of estimating the thresholds we will in the next section introduce an alternatively estimating procedure and then we will check this ideas through simulation.

3.2. Discretely observed process with unknown regimes

In this section our purpose is to implement a practical way of getting estimator for each threshold when we do not know the regime of each observation. The first step of the procedure consists of the classification of the observations in regimes in order to built an auxiliary conditional least squares contrast function as we did in the last section. For each n the procedure is implemented as follows.

(1) For fixed thresholds m and M, we split the observations in two regimes, $\widehat{R}_1(m, M)$ and $\widehat{R}_2(m, M)$ corresponding to the regime with positive trend and negative trend respectively. The regime classification is done in the following way: we start with the first observation in regime 1 and we consider all observations in regime 1 until the first observation that is bigger or equal than M, being that observation considered still in regime 1 and simultaneously also in regime 2, then we consider all the observations from this time on in regime 2 until we get the first one that is smaller or equal than m being that observation considered in regime 2 and also in regime 1, the classification process continues in this way until all the observations be considered in one of the regimes. We also get a value for $\widehat{K}_n$ the number of changes in regime.

(2) Next, we can compute the conditional estimator $\widehat{\mu}_n$ of μ_0 in a similar way as before but in this case using all the information, that is the number of estimated changes in regime in $[0, t_n]$ and not only on $[0, t_i]$. We define the auxiliary conditional least squares contrast function in the same way, but with the indicator functions depending only on the regime of the previous observation instead of the regimes for consecutive observations (it was needed before for the proof of the consistency).

(3.26)

$$LS_n(m, M) = \sum_{i=1}^{n-1} \left(X_{i+1} - X_i - \widehat{\mu}_n(m, M)\Delta\right)^2 1_{\widehat{R}_i=1} + \sum_{i=1}^{n-1} \left(X_{i+1} - X_i + \widehat{\mu}_n(m, M)\Delta\right)^2 1_{\widehat{R}_i=2}.$$

(3) Finally we choose for the threshold estimates the value of $(\widehat{m}, \widehat{M})$ that minimizes $LS_n(m, M)$, that is,

$$(\widehat{m}_n, \widehat{M}_n) = \text{argmin}_{(m,M)} LS_n(m, M),$$

and we put

$$\widehat{\mu}_n = \widehat{\mu}_n(\widehat{m}_n, \widehat{M}_n).$$

REMARK 3.8. Notice that the least squares contrast function (3.26) depends on the difference $M - m$ through $\widehat{\mu}_n$ as before, but it also depends on M and m through the sequence $\widehat{R}_1, ..., \widehat{R}_n$. With this dependence we can estimate the thresholds m and M separately because to different pairs (m, M) corresponds different sequences $\widehat{R}_1, ..., \widehat{R}_n$ and in consequence different values for $LS_n(m, M)$.

REMARK 3.9. For discrete observations of a continuous time process we have an extra difficulty comparing with the time series theory, because we do not know what happens between observations. In time series, the change in regime happens only at some observed value of the process but that is not the case in the continuous processes because the change in regime can happen between observed values. However, from the way we split the observations in the two regimes, it is obvious that we can restrict our computations of the conditional least squares error to the values of m and M that belong to the observed values of the process. This happens because if $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ are the order statistics of the observations then for $m \in]X_{p,n}, X_{p+1,n}]$ and for $M \in]X_{q,n}, X_{q+1,n}]$, $(p < q)$ with constant value of $M - m$, the sum $LS_n(m, M)$ is constant because the sequence $\widehat{R}_1, ..., \widehat{R}_n$ is always the same. Moreover, in order to reduce the computational burden, and because $m < M$, we can restrict our choice of m to the values $X_{1,n} \leq \cdots \leq X_{pn,n}$ and M to the values $X_{n-pn,n} \leq \cdots \leq X_{n,n}$ for some $p \in]0, 1[$.

We think that this procedure will give consistent estimators when Δ decreases and simultaneously the interval where we observe the process, $[0, t_n]$ increases. To get asymptotical results we should need a decreasing Δ because that is the only way to ensure that the regime classification of the observations converges to the true one for some sequence of thresholds $(m_n, M_n)_{n \in \mathbb{N}}$. At the same time we need to increase the observation interval $[0, t_n]$ to ensure that the number of changes in regime increases because those changes are fundamental in the estimation of the thresholds.

We are now in condition to illustrate this procedure on simple examples through simulations. First we start implementing the procedure for the process with fixed step size Δ considered in the LS_n function. In this implementation we will work with observations in a fixed length $[0, T]$ interval and we will simulate 500 repetitions of independent trajectories. In the second example we implement the procedure for the same process with

decreasing step size between observations. In both cases we use a smaller discretization step in the simulation of the trajectory than the one used to get a set of observations of the process.

EXAMPLE 3.10. We consider the Brownian motion with drift threshold model with $\mu_0 = 1$, $m_0 = -2$ and $M_0 = 2$, and for $\sigma = .4$ and $\sigma = .9$. The trajectory of the process is generated using a discretization step of .01 and we start the procedure with the conditional regime classification. The following figure illustrates a section of the trajectories under consideration for the two values of σ. The simulations were done using

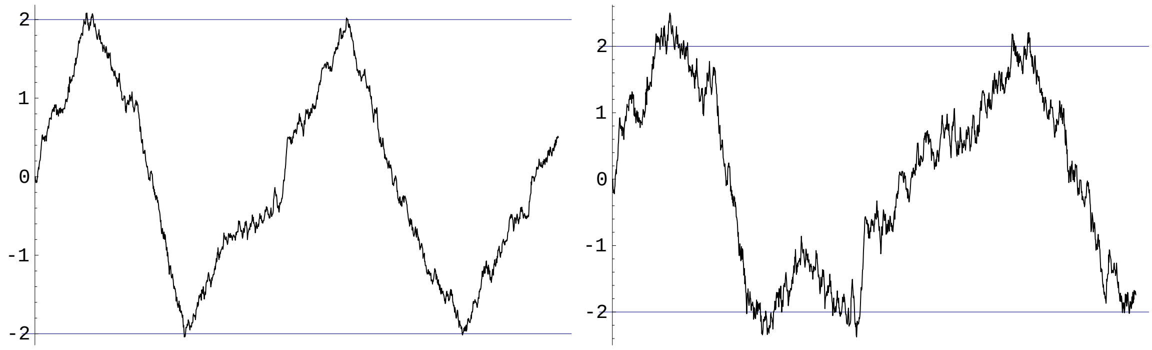

FIGURE 3.1. Threshold BMD trajectories with $\mu_0 = 1$, $m_0 = -2$, $M_0 = 2$ and $\sigma = .4$ vs $\sigma = .9$

the Mathematica program and first we start with 500 repetitions for the interval $[0, 100]$ (with 1000 observations). For the estimating procedure we fix $\Delta = .1$ and we use the LS_n (3.26) function. We perform a grid search for m in $[-2.3, -1]$ and for M in $[1, 2.3]$ and the step of the grid is .01. The results for the both cases (different values of σ) under consideration are presented in the next table.

TABLE 3.1. Estimates for 500 replicates for the threshold BMD process in $[0, 100]$ with $n = 1000$, $\mu_0 = 1$, $m_0 = -2$ and $M_0 = 2$, and for $\sigma = .4$ and $\sigma = .9$

	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$	$mean(\widehat{\mu}_n)$	$sd(\widehat{\mu}_n)$
$\sigma = .4$	-1.989	0.020	1.990	0.020	0.980	0.040
$\sigma = .9$	-1.996	0.045	2.003	0.050	0.971	0.088

As we can see the results suggest that the procedure works well, getting good approximations for the original values and as expected, the standard deviation for the estimators are larger when $\sigma = 0.9$.

In order to increase the number of changes in regime we also apply the procedure to the observation interval $[0, 500]$ ($n = 5000$) but this time just for 100 replicates and for both values of σ. The results are the following:

TABLE 3.2. Estimates for 100 replicates for the threshold BMD process in $[0, 500]$ with $n = 5000$, $\mu_0 = 1$, $m_0 = -2$ and $M_0 = 2$, and for $\sigma = .4$ and $\sigma = .9$

	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$	$mean(\widehat{\mu}_n)$	$sd(\widehat{\mu}_n)$
$\sigma = .4$	-2.000	0.0002	2.000	0.0002	0.985	0.016
$\sigma = .9$	-1.997	0.0140	1.994	0.0170	0.967	0.044

In the next example, we apply the estimating procedure to the process with decreasing discretization step and increasing time interval $[0, t_n]$.

EXAMPLE 3.11. Once more we consider the one dimensional Brownian motion with drift process, we will consider $\mu_0 = 1$, $m_0 = -2$ and $M_0 = 2$, and $\sigma = .9$. First, in order to study the asymptotic behavior of the estimators when Δ decreases we apply the procedure for 100 independent trajectories with decreasing Δ for a fixed observation interval $[0, 200]$. We compute the sum LS_n for the values of $(m, M) \in [-2.5, 1.0] \times [1.0, 2.5]$ in a grid with grid step .1. It will be interesting to simulate more intensively the process but we need a large amount of computer time to do that. We put the results in the following table.

TABLE 3.3. Estimates for 100 replicates from BMD process with decreasing Δ in the fixed observation interval $[0, 200]$. We have $\Delta = 200/n$ and $\mu_0 = 1$, $m_0 = -2$, $M_0 = 2$, $\sigma = .9$

	$mean(\widehat{\mu}_n)$	$sd(\widehat{\mu}_n)$	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$
$n = 100, \Delta = 2$	0.701	0.037	-0.801	0.132	0.762	0.157
$n = 200, \Delta = 1$	0.823	0.032	-1.298	0.118	1.293	0.123
$n = 400, \Delta = 1/2$	0.884	0.027	-1.556	0.082	1.585	0.075
$n = 800, \Delta = 1/4$	0.922	0.026	-1.732	0.056	1.733	0.062
$n = 1600, \Delta = 1/8$	0.940	0.026	-1.820	0.049	1.822	0.054
$n = 3200, \Delta = 1/16$	0.976	0.025	-1.970	0.046	1.959	0.049

REMARK 3.12. We know that, on one hand, it is not possible to estimate consistently the drift parameters if the observation interval is kept fixed, and similarly we do not expect to be able to estimate consistently the threshold parameters in this asymptotic

framework. Indeed, for a fixed observation interval the true number of changes in regime is constant and the changes in regime are fundamental for the consistent estimation of the parameters. We therefore do not expect the standard deviations of the estimators in table 3.3 to decrease to zero.

We also repeat the procedure for decreasing discretization step and increasing time interval in order to see the asymptotic behavior of the estimators. The results of this simulation are based in the following. The trajectory is initial simulated with step $\delta = 2^{-7}$ in the interval $[0, 128]$ and in the estimating procedure the discretization step is given by $\Delta_k = \frac{1}{2^k}$ with $k = 4, ..., 7$ corresponding to the observation intervals $[0, 2^k]$ and the number of observations 4^k. In the figure 3.2 we show one trajectory for $k = 5$ and in table 3.4 we present the results and as we can see the standard deviation of the estimators decreases to zero sugesting the consistency of the estimators.

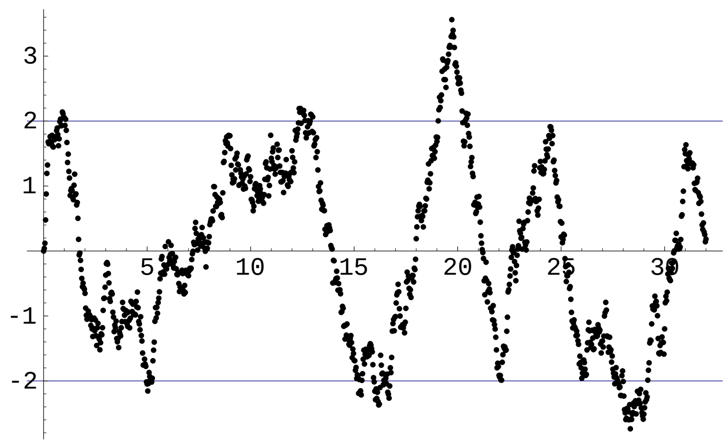

FIGURE 3.2. Threshold BMD process trajectory for $\Delta = 1/32$, observation interval $[0, 32]$ and $\mu_0 = 1$, $m_0 = -2$, $M_0 = 2$, $\sigma = .9$

TABLE 3.4. Estimates for 100 replicates from BMD process with decreasing $\Delta_k = 2^{-k}$, increasing observation interval $[0, 2^k]$, number of observations $n_k = 4^k$ and $\mu_0 = 1$, $m_0 = -2$, $M_0 = 2$, $\sigma = .9$

	$mean(\widehat{\mu}_n)$	$sd(\widehat{\mu}_n)$	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$
$k = 4$	0.981	0.216	-1.954	0.103	1.944	0.115
$k = 5$	1.011	0.161	-1.956	0.075	1.955	0.080
$k = 6$	1.010	0.107	-1.983	0.052	1.976	0.055
$k = 7$	0.990	0.077	-1.986	0.015	1.984	0.016

A final remark, it will be interesting to compare intensively the least square error estimates with the ones we get from the MCEM algorithm in chapter 2. In a first attempt we applied the least square errors procedure to the simulated data of chapter 2 and get similar results with values $\widehat{\mu}_n = 1.04387$, $\widehat{m}_n = -2.995$ and $\widehat{M}_n = 2.845$.

3.3. LSE for general diffusions discretely observed

We now investigate the generalization of the estimation procedure of the threshold estimators to other threshold diffusion processes than the ones built from the Brownian motion with drift. Processes that can change from an increasing to a decreasing regime after a change in the drift parameter, defined by the hitting of predefined thresholds, are the ones that we want to study. The estimating procedure will be the same as the one for the threshold process built from the Brownian motion with drift in the last section, that is, we will use the least squares estimation approach in order to get the threshold estimators. First we will define the LSE function for general threshold processes and then introduce auxiliary least contrast functions for the threshold processes built from the Ornstein-Uhlenbeck and geometric Brownian motion processes. Second, we will implement the procedure to simulated data from threshold models built using the Ornstein-Uhlenbeck and geometric Brownian motion processes and we will also apply the estimation procedure to real data from a set of several international financial funds.

We will study the general threshold diffusion model:

$$dXt = a(\mu(t), X_t)dt + b(\sigma, X_t)dB_t, \tag{3.27}$$

with

$$\mu(t) = \sum_{k\geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_1 \in \Theta_1, \mu_2 \in \Theta_2$$

where $0 = \tau_0 < \tau_1 < \ldots < \tau_j < \ldots$ are the threshold hitting times, that is, the instants where the process in the increasing regime hits the upper threshold M_0 or the process in the decreasing regime hits the lower threshold m_0. The sets Θ_1 and Θ_2 are such that, for $\mu \in \Theta_1$, $X_t(\mu)$ is in regime 1 (increasing regime) and for $\mu \in \Theta_2$, $X_t(\mu)$ is in regime 2 (decreasing regime). We define the conditional least square errors function as,

$$LSE_n(m, M) = \sum_{i=1}^{n-1} \left(X_{i+1} - \mathbb{E}_{\mu,m,M}\left[X_{i+1}|X_i, R_i\right]\right)^2,$$

for the general process, but as before for each of the processes that we will study an auxiliary function will be needed.

3.3.1. Geometric Brownian motion threshold process. Using the geometric Brownian motion process we built a threshold model.

$$dXt = \mu(t)X_t dt + \sigma X_t dB_t, \quad X_0 = x_0, \tag{3.28}$$

where

$$\mu(t) = \sum_{k\geq 0} \left[\mu_0 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) - \mu_0 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_0 > \sigma^2/2.$$

For the particular case of this threshold process we will consider the auxiliary conditional least squares contrast:

$$LS_n(m,M) = \sum_{i=1}^{n-1}(X_{i+1} - X_i e^{\hat{\mu}_n(m,M)\Delta})^2 1_{\widehat{R}_i=1} + \sum_{i=1}^{n-1}(X_{i+1} - X_i e^{-\hat{\mu}_n(m,M)\Delta})^2 1_{\widehat{R}_i=2}, \tag{3.29}$$

where we use the approximation

$$\mathbb{E}_{\mu,m,M}\left[X_{i+1}|X_i, R_i\right] = X_i e^{\hat{\mu}_n(m,M)\Delta}$$

or

$$\mathbb{E}_{\mu,m,M}\left[X_{i+1}|X_i, R_i\right] = X_i e^{-\hat{\mu}_n(m,M)\Delta}$$

depending on the regime, because that is the conditional expectation for the Geometric Brownian motion without thresholds and as we suspect for the case of the Brownian motion with drift threshold model this should be a good approximation when Δ is small. We also use the same ideas to define the drift estimator as,

$$\hat{\mu}_n(m,M) = \frac{\sigma^2}{2} + \frac{\hat{K}_n}{t_n}\ln\left(\frac{M}{m}\right). \tag{3.30}$$

where σ^2 should be replaced by $\hat{\sigma}^2$ if is unknown (being estimated in the same way as for the simple geometric Brownian motion) and where $\hat{K}_n$ is the estimated number of changes in regime in $[0, t_n]$. Notice that, we write $1_{\widehat{R}_i}$ instead of $1_{\widehat{R}_i=\widehat{R}_{i+1}}$ once again. For consistency proof purposes we should replace $\hat{\mu}_n(m,M)$ with $\hat{\mu}_i(m,M)$ in order to work with a function that is measurable with respect to $\sigma(X_1,...,X_i)$. If we want to define the change in regime as a change from $\mu_1 > \sigma^2/2$ to $\mu_2 < \sigma^2/2$ we should compute the estimators $\hat{\mu}_{1,n}$, $\hat{\mu}_{2,n}$ of μ_1 and μ_2 replacing t_n in equation(3.30) by $2\Delta\sum_{i=1}^{n-1} 1_{\widehat{R}_i=1}$ and $2\Delta\sum_{i=1}^{n-1} 1_{\widehat{R}_i=2}$, respectively.

Once again we start by repeating the procedure for 100 independent trajectories with decreasing Δ and for a fixed observation interval $[0, 200]$. We consider, $\mu_1 = 1, \mu_2 = -1$

and $\sigma = .6$, for the thresholds $m_0 = 5$ and $M_0 = 15$. We compute LS_n for $(m, M) \in [4, 7] \times [11, 16]$ in a grid with step .1, we get the results presented in the table 3.5.

TABLE 3.5. Estimates for 100 replicates from GBM process with decreasing Δ in the fixed observation interval $[0, 200]$. We have $\Delta = 200/n$ and $\mu_1 = 1 = -\mu_2$, $m_0 = 5$, $M_0 = 15$, $\sigma = .6$

	$mean(\widehat{\mu}_{1,n})$	$sd(\widehat{\mu}_{1,n})$	$mean(\widehat{\mu}_{2,n})$	$sd(\widehat{\mu}_{2,n})$	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$
$n = 100$	0.195	0.096	-0.073	0.093	6.451	0.203	11.052	0.408
$n = 200$	0.400	0.088	-0.155	0.085	6.210	0.182	11.960	0.371
$n = 400$	0.616	0.064	-0.341	0.069	5.965	0.131	12.495	0.320
$n = 800$	0.773	0.061	-0.576	0.062	5.670	0.098	13.310	0.212
$n = 1600$	0.822	0.050	-0.772	0.051	5.495	0.069	14.130	0.160
$n = 3200$	0.903	0.044	-0.914	0.047	5.150	0.058	14.840	0.117

In order to implement the simulation study for decreasing discretization step and increasing observation interval, we built the initial trajectory with discretization step $\delta = 2^{-7}$ and apply the procedure for each $k = 4, ..., 7$ with $\Delta_k = \frac{1}{2^k}$, the observation intervals $[0, 2^k]$ and the number of observations 4^k.

We show one trajectory in figure 3.3 and get the results presented in table 3.6.

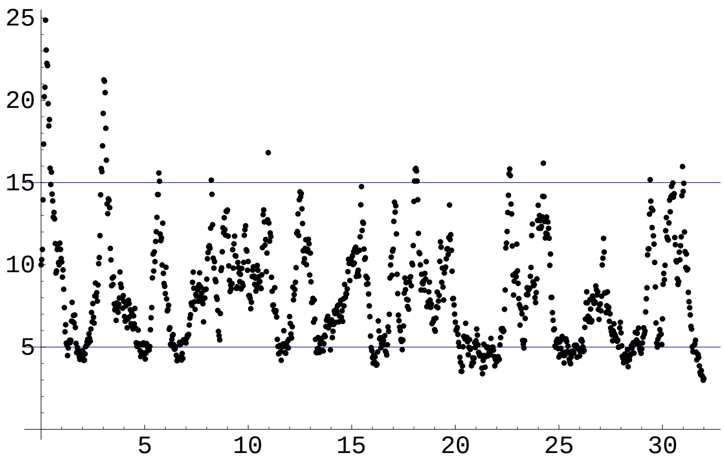

FIGURE 3.3. Trajectory for the Geometric Brownian motion threshold process with $\Delta = 1/32$, observation interval $[0, 32]$ and $\mu_1 = 1 = -\mu_2$, $m_0 = 5$, $M_0 = 15$, $\sigma = .6$

As we can see the procedure seems to work well to get the threshold estimators, regarding the drift parameters, the convergence seems to be slower.

TABLE 3.6. Estimates for 100 replicates from GBM process with decreasing $\Delta_k = 2^{-k}$, increasing observation interval $[0, 2^k]$, number of observations $n_k = 4^k$ and $\mu_1 = 1 = -\mu_2$, $m_0 = 5$, $M_0 = 15$, $\sigma = .6$

	$mean(\widehat{\mu}_{1,n})$	$sd(\widehat{\mu}_{1,n})$	$mean(\widehat{\mu}_{2,n})$	$sd(\widehat{\mu}_{2,n})$	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$
$k=4$	0.830	0.161	-0.949	0.295	5.412	0.296	14.364	0.499
$k=5$	0.920	0.136	-0.959	0.184	5.374	0.160	14.450	0.358
$k=6$	0.961	0.099	-0.974	0.126	5.236	0.137	14.522	0.276
$k=7$	0.984	0.059	-1.007	0.089	5.022	0.076	14.958	0.067

3.3.2. Ornstein-Uhlenbeck. Finally, we built from the Ornstein-Uhlenbeck process a threshold model.

$$dXt = \mu(t)X_t dt + \sigma dB_t, \quad X_0 = x_0, \tag{3.31}$$

where

$$\mu(t) = \sum_{k\geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_1 > 0, \mu_2 < 0$$

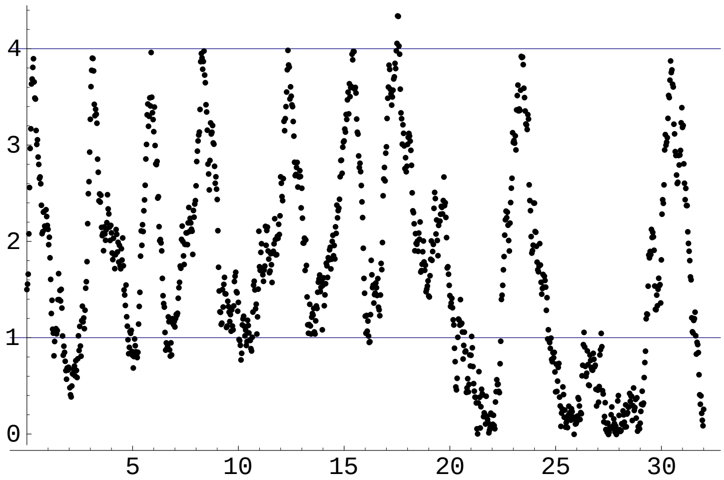

FIGURE 3.4. One trajectory for the Ornstein-Uhlenbeck threshold process with $\Delta = 1/32$, observation interval $[0, 32]$ and $\mu_1 = 1 = -\mu_2$, $m_0 = 1$, $M_0 = 4$, $\sigma = 1$

For the particular case of this threshold process we will consider the auxiliary conditional least squares contrast similar to the one used for the Geometric Brownian motion threshold process:

$$LS_n(m, M) = \sum_{i=1}^{n-1}(X_{i+1} - X_i e^{\hat{\mu}_{1,n}(m,M)\Delta})^2 1_{\widehat{R}_i=1} + \sum_{i=1}^{n-1}(X_{i+1} - X_i e^{\hat{\mu}_{2,n}(m,M)\Delta})^2 1_{\widehat{R}_i=2}, \tag{3.32}$$

because the conditional expectation for the original Ornstein-Uhlenbeck process without thresholds is the same as the one for the Geometric Brownian motion and once more we expect that for small Δ this should be a good approximation.

This time we define the estimator of μ_1 and μ_2 using the usual estimators for the Ornstein-Uhlenbeck process but considering the observations in each regime to estimate each one of the parameters.

$$\hat{\mu}_1(m,M) = \frac{1}{\Delta} \ln \left(\frac{\sum_{i=1}^{n-1} (X_i X_{i+1}) 1_{\hat{R}_i(m,M)=1}}{\sum_{i=1}^{n-1} X_i^2 1_{\hat{R}_i(m,M)=1}} \right), \tag{3.33}$$

and

$$\hat{\mu}_2(m,M) = \frac{1}{\Delta} \ln \left(\frac{\sum_{i=1}^{n-1} (X_i X_{i+1}) 1_{\hat{R}_i(m,M)=2}}{\sum_{i=1}^{n-1} X_i^2 1_{\hat{R}_i(m,M)=2}} \right). \tag{3.34}$$

where the dependency from m and M is given from the conditional classification $\hat{R}_i(m,M) = 1$ or $\hat{R}_i(m,M) = 2$.

We apply the procedure for 100 independent trajectories with decreasing Δ and for the fixed observation interval $[0,200]$. Considering $\mu_1 = 1, \mu_2 = -1$ and $\sigma = 1$, and for the thresholds $m_0 = 1$ and $M_0 = 4$. We compute LS_n for the values of $(m,M) \in [0.5, 2.4] \times [2.6, 4.5]$ in a grid with step .1. We get the results presented in table 3.7.

TABLE 3.7. Estimates for 100 replicates from O.U. process with decreasing Δ in the fixed observation interval $[0,200]$. We have $\Delta = 200/n$ and $\mu_1 = 1 = -\mu_2$, $m_0 = 1$, $M_0 = 4$, $\sigma = 1$

	$mean(\widehat{\mu}_{1,n})$	$sd(\widehat{\mu}_{1,n})$	$mean(\widehat{\mu}_{2,n})$	$sd(\widehat{\mu}_{2,n})$	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$
$n = 100$	0.176	0.089	-0.245	0.077	1.978	0.191	2.125	0.214
$n = 200$	0.492	0.068	-0.548	0.067	1.806	0.129	2.347	0.178
$n = 400$	0.646	0.065	-0.688	0.062	1.613	0.124	2.900	0.127
$n = 800$	0.762	0.064	-0.770	0.058	1.404	0.085	3.338	0.092
$n = 1600$	0.811	0.058	-0.832	0.049	1.266	0.075	3.621	0.067
$n = 3200$	0.876	0.039	-0.944	0.043	1.038	0.049	3.920	0.013

REMARK 3.13. Notice that the reported values for the standard deviations of the estimators are influenced by the fact that, in order to reduce the computer time, we restricted our search in the minimization procedure to derive the value of the thresholds estimators by using a grid with step 0.1. As mentioned in Remark 3.12, these standard deviations should not tend to zero, a higher number of observation is required to check the behavior of these quantities in tables 3.5 and 3.7.

To implement the simulation study for the decreasing discretization step and increasing observation interval, we built the initial trajectory as before. We apply the procedure for each $k = 4, ..., 7$ with $\Delta_k = \frac{1}{2^k}$, the observation intervals $[0, 2^k]$ and the number of observations 4^k. We get the results presented in table 3.8. As already happens for the

TABLE 3.8. Estimates for 100 replicates from OU process with decreasing $\Delta_k = 2^{-k}$, increasing observation interval $[0, 2^k]$, number of observations $n_k = 4^k$ and $\mu_1 = 1 = -\mu_2$, $m_0 = 1$, $M_0 = 4$, $\sigma = 1$

	$mean(\widehat{\mu}_{1,n})$	$sd(\widehat{\mu}_{1,n})$	$mean(\widehat{\mu}_{2,n})$	$sd(\widehat{\mu}_{2,n})$	$mean(\widehat{m}_n)$	$sd(\widehat{m}_n)$	$mean(\widehat{M}_n)$	$sd(\widehat{M}_n)$
$k = 4$	0.941	0.147	-0.915	0.191	1.118	0.078	3.742	0.091
$k = 5$	0.954	0.118	-0.934	0.101	1.098	0.056	3.830	0.058
$k = 6$	0.963	0.094	-0.942	0.079	1.056	0.039	3.890	0.043
$k = 7$	0.984	0.062	-0.971	0.044	1.001	0.009	3.990	0.010

Geometric Brownian motion, the procedure seems to work well to get the threshold estimators, but the convergence for the drift parameters seems to be slower.

3.3.3. Real data. Now we will apply the procedure to real data. We do not perform any kind of adjustment test and so we compute the threshold estimators of the considered data supposing the model is of any of the three kinds considered before. The data is the daily prices of three different funds, two of them mainly consists of stocks while the other mainly of bonds.

EXAMPLE 3.14. The first fund to be considered is the PF-European Sustainable Equities-R fund from the Pictet Funds company. The data we consider is from the year 2004. We only consider the data from the first ten months because as we can see in the figure 3.5, the last two months do not follow the same dynamics. Supposing that the threshold model is from, Brownian motion with drift (BMD), geometric Brownian motion (GBM) or Ornstein-Uhlenbeck (OU), respectively, the estimating procedure is implemented. We compute LS_n for the values of $(m, M) \in [114, 118] \times [118, 124]$ in a grid with step .1.

The results can be observed in table 3.9 where the value of LS_n could be of interest just to have a rough idea of what model should have a better adjustment. The threshold estimates are the same, and this make us suspect that it is enough to use a reasonable adjusted model to estimate the thresholds.

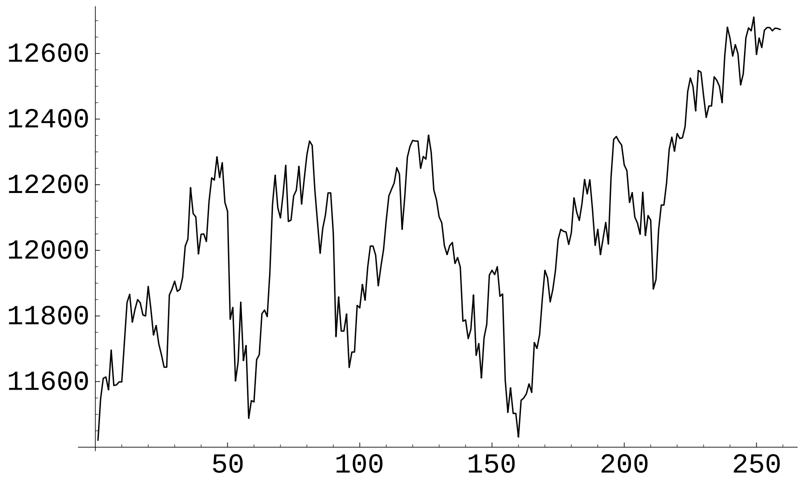

FIGURE 3.5. 2004 daily prices from European Sustainable Equities-R fund from Pictet Funds

TABLE 3.9. Threshold estimates from PF-European Sustainable Equities-R

	$\widehat{\mu}_{1,n}$	$\widehat{\mu}_{2,n}$	$\widehat{m}_n$	$\widehat{M}_n$	LS_n
BMD	0.203	-0.365	116.500	122.800	148.337
GBM	0.002	-0.003	116.500	122.800	148.453
OU	0.002	-0.003	116.500	122.800	148.450

EXAMPLE 3.15. The second fund to be considered is the Parvest Europe Dynamic Growth fund from the BNP Paribas company. The data is from the year 2004. Once

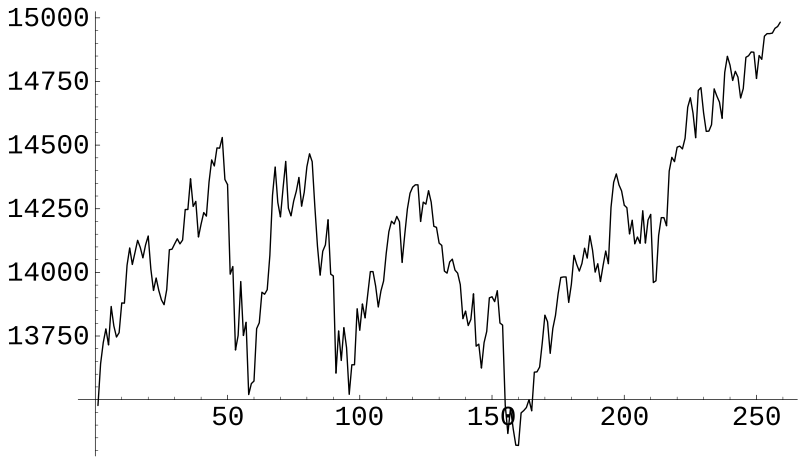

FIGURE 3.6. 2004 daily prices from Parvest Europe Dynamic Growth fund from the BNP Paribas

again we only consider the data from the first ten months and we will consider the three models under study. We compute LS_n for the values of $(m, M) \in [134, 139] \times [140, 145]$ in a grid with step .1, for this example, we get the values that appear in table 3.10.

TABLE 3.10. Threshold estimates from Parvest Europe Dynamic Growth

	$\widehat{\mu}_{1,n}$	$\widehat{\mu}_{2,n}$	$\widehat{m}_n$	$\widehat{M}_n$	LS_n
BMD	0.314	−0.296	135.300	143.400	197.854
GBM	0.002	−0.002	135.300	143.400	197.881
OU	0.002	−0.002	135.300	143.400	197.878

Once more the threshold estimates are the same for all models and the LSE values are very similar.

EXAMPLE 3.16. The last fund to be considered is the Converging Europe Bond fund from the Schroder company. The data we have considered is from the year 2005. This time

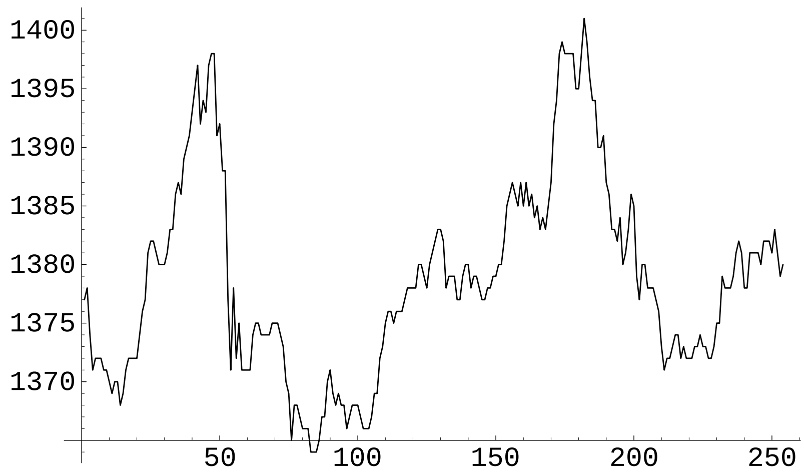

FIGURE 3.7. 2005 daily prices from Converging Europe Bond fund from the Schroder company

we consider the data from the entire year as it is presented and we will consider the three models under study. We compute LS_n for the values of $(m, M) \in [1364, 1380] \times [1386, 1401]$ in a grid with step .5.

The threshold estimates are the same for all models and once again the LSE values are similar as we can see in table 3.11.

TABLE 3.11. Threshold estimates from Converging Europe Bond

	$\widehat{\mu}_{1,n}$	$\widehat{\mu}_{2,n}$	$\widehat{m}_n$	$\widehat{M}_n$	LS_n
BMD	0.3153	−1.250	13.710	13.980	916.823
GBM	0.0002	−0.001	13.710	13.980	916.781
OU	0.0002	−0.001	13.710	13.980	916.781

Conclusion and future research

In conclusion, in this thesis we have developed and implemented two estimating procedures for the problem of diffusion processes threshold estimation. The first procedure, MCEM, is one adaptation of the Expectation-Maximization algorithm to our problem. The second procedure, least squares estimation, has been applied to several threshold diffusions but results about the consistency have only been proved for the threshold difference $M - m$. During this thesis we also tried other ideas like building tests to detect the change in regime or using other approaches to prove the consistency of the estimators. For future research we point out some topics. Regarding the least squares estimators, is desirable to prove asymptotic consistency and to study the asymptotic distribution of the estimators. Another possibility is to consider, not only a change in the drift parameter from one regime to the other, but a complete change of process, that is, in one regime the process is driven by one diffusion and in the other regime we consider another diffusion. Like in the time series threshold models (see chapter 1), other estimating procedures can be built using maximum likelihood theory, the Bayesian theory, generalized method of moments or Wavelets and it can also be used sub-sampling or sequential estimation. Another question mentioned in the time series threshold models is the problem of testing for linearity, that is, the problem of testing the threshold model against the linear model. For our model this question is in open, and for future research it could be of interest to build a test for testing a threshold diffusion model against simple diffusion model.

Resumo em Português

O modelo auto-regressivo com limiares (threshold-TAR) está bem documentado e estudado no caso de séries temporais. Não sendo este o caso quando se consideram processos contínuos e em particular difusões, propomo-nos estudar algumas difusões, em que estas sofrem uma mudança de um regime com tendência positiva para um regime com tendência negativa quando o processo atinge um limiar superior "M" e sofrem uma mudança de um regime com tendência negativa para um regime com tendência positiva quando o processo atinge um limiar inferior "m". Nesta tese propomo-nos a implementar procedimentos de estimação com vista a obter estimadores dos limiares para modelos com limiares definidos a partir de equações diferenciais estocásticas. O primeiro procedimento a ser apresentado baseia-se na adequação do algoritmo EM (expectation-maximization ou esperança e maximização) à estimação de limiares no modelo com limiares construido a partir do processo Browniano com tendência. O segundo procedimento, repete uma das idéias fundamentais na estimação de limiares no contexto de séries temporais, estimação de minimos quadrados, ou seja o procedimento que iremos adoptar será o de estimar os limiares pelos valores que minimizam a soma do quadrado dos erros. Iremos implementar este procedimento não só para modelos com limiares baseados no processo Browniano com tendência mas também para modelos genéricos entre os quais se destacam os que são baseados nos processos de Ornstein-Uhlenbeck e Browniano geométrico. Ambos os procedimentos são sujeitos a uma implementação prática aplicada a dados simulados, sendo ainda o procedimento de estimação por minimos quadrados aplicado a dados reais respeitantes a cotações diárias de um conjunto de fundos financeiros internacionais. O primeiro fundo é o fundo PF-European Sustainable Equities-R da Pictet Funds e o segundo o Parvest Europe Dynamic Growth fund do BNP Paribas. Os dados para ambos os fundos são os preços diários do ano 2004. O último fundo a ser considerado é o fundo Converging Europe Bond da Schroder e os dados são os preços diários do ano 2005.

RP-1. Resumo do Capítulo 1

Nas últimas décadas temos assistido a grandes desenvolvimentos na área da inferência estatística quer no campo das séries temporais quer no campo das difusões. No contexto das séries temporais as condições de linearidade e de estacionaridade têm sido abandonadas e o estudo de modelos não lineares tem aumentado. Uma classe de mo-delos não lineares, chamados de modelos autoregressivos com limiares (TAR-Threshold autoregressive) está amplamente estudado em [**Tong, 1990**], nestes modelos um processo é dividido em m regimes sendo cada regime um processo autoregressivo. Este tipo de modelo pode ser representado pela seguinte equação:

$$Y_t = \sum_{i=1}^{m} (a_{i,0} + a_{i,1}Y_{t-1} + \cdots + a_{i,p}Y_{t-p} + \varepsilon_{i,t})\, \mathbb{I}_{\{r_{i-1} \leq Z_{t-1} < r_i\}}. \tag{RP-1}$$

Os limiares são números reais $-\infty = r_0 < r_1 < \ldots < r_{m-1} < r_m = +\infty$ e $Z_{t-1} = Z(Y_1, ..., Y_{t-1})$ é a variável limiar que especifica a mudança de regime. Neste capítulo é feito um apanhado dos diferentes métodos de estimação e de outras questões relevantes no estudo de modelos com limiares no contexto das séries temporais. Como já referimos o nosso objectivo passa por generalizar a noção de processos com limiares a processos contínuos. Uma difusão que experimenta uma mudança de regime quando ultrapassa um limiar superior (M) ou um limiar inferior (m) pode ser considerada como um modelo genérico do processo estocástico. Considerando-se dois limiares m e M e supondo-se que se inicia com a difusão de tendência positiva (definida por um parâmetro de tendência μ_1) o processo continuará neste regime até ao instante em que atinge o limiar superior M ocorrendo então uma mudança de regime (que corresponde à alteração do parâmetro de tendência para μ_2) mantendo-se o processo neste segundo regime até ao instante em que atinge o limiar inferior m, voltando o processo ao regime inicial. Este modelo pode ser representado pela seguinte equação:

$$dXt = a(\mu(t), X_t)dt + b(\sigma, X_t)dB_t, \tag{RP-2}$$

com

$$\mu(t) = \sum_{k\geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k}, \tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1}, \tau_{2k+2}[}(t)\right],$$

onde $0 = \tau_0 < \tau_1 < \ldots < \tau_j < \ldots$ são os instantes de contacto dos limiares, isto é, $\tau_{2k+1} = \inf\{t > \tau_{2k}; X_t = M\}$ e $\tau_{2k+2} = \inf\{t > \tau_{2k+1}; X_t = m\}$ para $k \geq 0$ e $\tau_0 = 0$.

RP-2. Resumo do Capítulo 2

Este capítulo começa por uma primeira secção dedicada ao estudo dos instantes de contacto (hitting times) para algumas difusões, nomeadamente para o processo Browniano com tendência e para o processo Browniano geométrico. Prossegue com uma segunda secção dedicada à implementação do algoritmo MCEM (Monte-Carlo Expectation-Maximization) ao modelo de limiares construido com base no processo Browniano com tendência e termina com uma secção dedicada à simulação do modelo e implementação do algoritmo com vista à estimação dos parâmetros do modelo.

RP-2.1. Distribuição dos instantes de contacto. Nesta secção apresentaremos resultados sobre a distribuição dos instantes de contacto para alguns processos.

Dado um processo estocástico X com trajectorias contínuas e adaptado a uma filtração $(\mathcal{F}_t)_{t\geq 0}$, considerando um subconjunto $\Gamma \in \mathcal{B}(\mathbb{R})$ do espaço de estados do processo, o instante de contacto (em Γ) é definido por,

$$T_\Gamma(\omega) = \inf\{t \geq 0; X_t(\omega) \in \Gamma\}.$$

No que se segue os instantes de contacto de interesse serão os instante de contacto com limiares predefinidos.

RP-2.1.1. *Processo Browniano com tendência μ e coeficiente de difusão σ.* Consideremos o processo Browniano com tendência μ e coeficiente de difusão σ,

$$B_t^{\mu,\sigma} = \mu t + \sigma B_t$$

com B_t processo Browniano standard. Começando com

$$T_{b,a}^{\mu,\sigma} = \inf\left\{t : B_t^{\mu,\sigma} = b\right\}, \quad B_0^{\mu,\sigma} = a,$$

obtemos a densidade

$$f_{T_{b,a}^{\mu,\sigma}}(t) = \frac{|b-a|}{\sqrt{2\pi\sigma^2 t^3}}\, e^{-\frac{(b-a-\mu t)^2}{2\sigma^2 t}}, \quad t > 0.$$

RP-2.1.2. *Processo Browniano geométrico.* Se considerarmos o processo Browniano geométrico, um processo positivo satisfazendo a equação diferencial estocástica,

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad X_0 = a.$$

Sabe-se que a solução desta equação é dada por,

$$X_t = a\, e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t},$$

pelo que podemos escrever

$$X_t = g\left(B_t^{\mu - \frac{\sigma^2}{2}, \sigma}\right)$$

com

$$g(z) = ae^z, \quad g^{-1}(x) = \ln\left(\frac{x}{a}\right).$$

Pelo que estamos perante uma transformação do processo Browniano com tendência, obtendo-se a densidade para o primeiro instante de contacto,

$$f_{T_{b,a}^X}(t) = \frac{|\ln(b) - \ln(a)|}{\sqrt{2\pi\sigma^2 t^3}} \, e^{-\frac{\left(\ln(b)-\ln(a)-\left(\mu-\frac{\sigma^2}{2}\right)t\right)^2}{2\sigma^2 t}}, t > 0.$$

RP-2.2. Algoritmo MCEM. Nesta secção iremos implementar um procedimento de estimação dos parâmetros do modelo com limiares baseado no algoritmo MCEM (Monte-Carlo Expectation-Maximization) que nos permite fazer estimação de parâmetros em modelos com dados imncompletos como será o nosso caso pois os instantes de contacto não são conhecidos. Consideraremos o processo Browniano com tendência e dados dois limiares m e M iremos supor que a mudança ocorre apenas no coeficiente de tendência.

Considere-se o seguinte modelo:

$$dXt = \mu(t)dt + \sigma dB_t, \quad X_0 = x_0 \in [m, M[,$$

com

$$\mu(t) = \sum_{k \geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k}, \tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1}, \tau_{2k+2}[}(t)\right], \quad \mu_1 > 0, \ \mu_2 < 0$$

onde $0 = \tau_0 < \tau_1 < \ldots < \tau_k < \ldots$ são os instantes de contacto. Isto é, dados dois limiares m e M, τ_1 é o tempo necessário para o processo com tendência positiva μ_1 ir de x_0 até M, $\tau_2 - \tau_1$ é o tempo necessário para o processo com tendência negativa μ_2 ir de M até m, genericamente, $\tau_i - \tau_{i-1}$ é o tempo necessário para o processo ir de um limiar até ao outro. Supomos que σ é conhecido e queremos estimar $\theta = (\mu_1, \mu_2, m, M)$ observando-se $X_0, X_1, ..., X_n$ em intervalos equidistantes $0 = t_0, t_1, ..., t_n$ no intervalo $[0, T]$, com $\Delta = t_j - t_{j-1}$ e $t_j = j\Delta$ para $j = 1, ..., n$. Usamos o processo Browniano com tendência μ_1 para gerar o primeiro regime e o processo Browniano com tendência μ_2 para gerar o segundo regime, no entanto o processo resultante não é um processo Browniano com tendência. Apesar das densidades de transição não serem conhecidas para este processo com limiares iremos implementar o algoritmo EM utilizando o que nos parece uma aproximação natural, ou seja, as densidades de transição do processo Browniano com tendência. Então p_1 e p_2,

as densidades de transição nos regimes 1 e 2 consideradas, serão respectivamente,

$$p_1(\Delta, x_i, x_{i+1}; \mu_1) = \frac{1}{\sqrt{2\Pi\sigma^2\Delta}} \exp\left(-\frac{(x_{i+1} - x_i - \mu_1\Delta)^2}{2\sigma^2\Delta}\right)$$
$$p_2(\Delta, x_i, x_{i+1}; \mu_2) = \frac{1}{\sqrt{2\Pi\sigma^2\Delta}} \exp\left(-\frac{(x_{i+1} - x_i - \mu_2\Delta)^2}{2\sigma^2\Delta}\right).$$

Os instantes de contacto $\tau_1, ..., \tau_k$ não são observados e o seu número no intervalo $[0, T]$ é dado por uma variável aleatória K. As densidades para a diferença entre instantes de contacto consecutivos são conhecidas e dadas por:

$$f_{T_1}(\tau_1; \mu_1, M) = \frac{M - x_0}{\sqrt{2\pi\sigma^2\tau_1^3}} \exp\left(-\frac{(M - x_0 - \mu_1\tau_1)^2}{2\sigma^2\tau_1}\right),$$

para $i = 2, 4, 6, ...,$

$$f_{T_i - T_{i-1}}(\tau_i - \tau_{i-1}; \theta) = \frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}} \exp\left(-\frac{(m - M - \mu_2(\tau_i - \tau_{i-1}))^2}{2\sigma^2(\tau_i - \tau_{i-1})}\right),$$

e para $i = 3, 5, 7, ...,$

$$f_{T_i - T_{i-1}}(\tau_i - \tau_{i-1}; \theta) = \frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}} \exp\left(-\frac{(M - m - \mu_1(\tau_i - \tau_{i-1}))^2}{2\sigma^2(\tau_i - \tau_{i-1})}\right).$$

Para este modelo teremos de considerar a função de verosimilhança completa ou total, ou seja, a função de verosimilhança para a totalidade dos dados (os conhecidos e os desconhecidos),

$$L_c^K(\omega) = \sum_{k=1}^{+\infty} L_c(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta)\mathbb{I}_{\{K=k\}}(\omega).$$

Condicionando aos instantes de contacto o processo X_t é markoviano e teremos

$$L_c(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta) = f_{X_1,...,X_n,T_1,...,T_k}(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta)$$
$$= f_{X_1,...,X_n|T_1,...,T_k}(x_1, ..., x_n|\tau_1, ..., \tau_k) f_{T_1,...,T_k}(\tau_1, ..., \tau_k; \theta)$$
$$= \prod_{i=0}^{j_1-1} p_1(\Delta, x_i, x_{i+1}) p_1(\tau_1 - t_{j_1}, x_{j_1}, M) p_2(t_{j_1+1} - \tau_1, M, x_{j_1+1}) \prod_{i=j_1+1}^{j_2-1} p_2(\Delta, x_i, x_{i+1})$$
$$\times\, p_2(\tau_2 - t_{j_2}, x_{j_2}, m) p_1(t_{j_2+1} - \tau_2, m, x_{j_2+1}) \cdots p_{[1+mod(k+1,2)]}(\tau_k - t_{j_k}, x_{j_k}, \star)$$
$$\times\, p_{[1+mod(k,2)]}(t_{j_k+1} - \tau_k, \star, x_{j_k+1}) \prod_{i=j_k+1}^{n-1} p_{[1+mod(k,2)]}(\Delta i, x_i, x_{i+1})$$
$$\times\, \frac{M - x_0}{\sqrt{2\pi\sigma^2\tau_1^3}} \exp\left(-\frac{(M - x_0 - \mu\tau_1)^2}{2\sigma^2\tau_1}\right) \prod_{i=2}^{k} \frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}}$$
$$\times \prod_{i=1}^{[(k-1)/2]} \exp\left(-\frac{(M - m - \mu_1(\tau_{2i+1} - \tau_{2i}))^2}{2\sigma^2(\tau_{2i+1} - \tau_{2i})}\right) \prod_{i=1}^{[k/2]} \exp\left(-\frac{(m - M - \mu_2(\tau_{2i} - \tau_{2i-1}))^2}{2\sigma^2(\tau_{2i} - \tau_{2i-1})}\right).$$

Onde para todo o $i = 1, ..., k$, $j_i \in \{1, ..., n-1\}$ é, tal que, $\tau_i \in]t_{j_i}, t_{j_i+1}[$ e $j_1 < j_2 < \cdots < j_k$. Com $\star = m \times mod(k+1, 2) + M \times mod(k, 2)$ e $mod(a, b)$ é o resto da divisão inteira de a por b. Definimos também a função logaritmo de verosimilhança completa de forma similar por,

$$\log(L_c^K)(\omega) = \sum_{k=1}^{+\infty} \log(L_c(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta))\mathbb{I}_{\{K=k\}}(\omega)$$

com

$$\begin{aligned}
&\log(L_c(x_1, ..., x_n, \tau_1, ..., \tau_k; \theta)) = \\
&= -\sum_{j_1,..,j_k=1}^{n-1} \left\{ \sum_{i=0}^{j_1-1} \frac{(x_{i+1} - x_i - \mu_1\Delta)^2}{2\sigma^2\Delta} + \sum_{i=j_1+1}^{j_2-1} \frac{(x_{i+1} - x_i - \mu_2\Delta)^2}{2\sigma^2\Delta} + \cdots \right. \\
&+ \sum_{i=j_k-1}^{n-1} \frac{(x_{i+1} - x_i - \mu_{[1+mod(k,2)]}\Delta)^2}{2\sigma^2\Delta} + \sum_{i=1}^{[(k+1)/2]} \frac{(M - x_{j_{2i-1}} - \mu_1(\tau_{2i-1} - t_{j_{2i-1}}))^2}{2\sigma^2(\tau_{2i-1} - t_{j_{2i-1}})} + \\
&+ \sum_{i=1}^{[(k+1)/2]} \frac{(x_{j_{2i-1}+1} - M - \mu_2(t_{j_{2i-1}+1} - \tau_{2i-1}))^2}{2\sigma^2(t_{j_{2i-1}+1} - \tau_{2i-1})} + \sum_{i=1}^{[k/2]} \frac{(m - x_{j_{2i}} - \mu_2(\tau_{2i} - t_{j_{2i}}))^2}{2\sigma^2(\tau_{2i} - t_{j_{2i}})} \\
&\left. + \sum_{i=1}^{[k/2]} \frac{(x_{j_{2i}+1} - m - \mu_1(t_{j_{2i}+1} - \tau_{2i}))^2}{2\sigma^2(t_{j_{2i}+1} - \tau_{2i})} \right\} \times \prod_{i=1}^{k} \mathbb{I}_{[t_{j_i}, t_{j_i+1}[}(\tau_i) + \log\left(\frac{M - x_0}{\sqrt{2\pi\sigma^2\tau_1^3}}\right) \\
&- \frac{(M - x_0 - \mu_1\tau_1)^2}{2\sigma^2\tau_1} + \sum_{i=2}^{k} \log\left(\frac{M - m}{\sqrt{2\pi\sigma^2(\tau_i - \tau_{i-1})^3}}\right) \\
&- \sum_{i=1}^{[(k-1)/2]} \frac{(M - m - \mu_1(\tau_{2i+1} - \tau_{2i}))^2}{2\sigma^2(\tau_{2i+1} - \tau_{2i})} - \sum_{i=1}^{[k/2]} \frac{(m - M - \mu_2(\tau_{2i} - \tau_{2i-1}))^2}{2\sigma^2(\tau_{2i} - \tau_{2i-1})},
\end{aligned}$$

excepto por um termo que não envolve θ.

RP-2.2.1. *Fase E.* Neste algoritmo, precisamos na p-ésima iteraçao do passo E, de calcular a esperança matemática da função logaritmo de verosimilhança completa condicionada aos dados e aos valores correntes de $\theta_p = (\mu_{1,p}, \mu_{2,p}, m_p, M_p)$, isto é,

$$Q(\theta, \theta_p) = \mathbb{E}_{\theta_p}\left[\log\left(L_c^K\right)|X_1 = x_1, ..., X_n = x_n\right]$$

De seguida aproxima-se este valor esperado usando Monte Carlo. Concretamente, simulamos réplicas $\tau_1^l, ..., \tau_{k_l}^l$ de $T_1, ..., T_K$ e ponderamos essas réplicas. Repare-se que para cada $l \in \{1, ..., L\}$ extraimos a sequência $\tau_1^l, ..., \tau_{k_l}^l$ de

$$f_{T_1,...,T_{k_l}}(\tau_1, ..., \tau_{k_l}; \theta_p) = f_{T_1}(\tau_1; \mu_{1,p}, M_p) f_{T_2-T_1}(\tau_2 - \tau_1; \theta_p) f_{T_{k_l}-T_{k_l-1}}(\tau_{k_l} - \tau_{k_l-1}; \theta_p)$$

de forma sequencial. Isto é, extraimos s_1^l de f_{T_1} fazendo $\tau_1^l = s_1^l$ depois extraimos s_2^l de $f_{T_2-T_1}$ fazendo $\tau_2^l = \tau_1^l + s_2^l$ e continua-se enquanto estivermos no intervalo $[0, T]$. Desta

forma não só construimos uma sequência de instantes de contacto mas também obtemos uma observação do número de instantes de contacto K. Finalmente poderemos aproximar a esperança condicional da função logaritmo de verosimilhança completa por,

$$\mathbb{E}_{\theta_p}\left[\log\left(L_c^K\right)|X_1,...,X_n\right] \approx \frac{1}{L}\sum_{l=1}^{L}\log\left(L_c^{k_l}(X_1,...,X_n,\tau_1^l,...,\tau_{k_l}^l)\right)\times w^l, \tag{RP-3}$$

onde o peso w^l é tal que,

$$w^l \propto f_{X_1,...,X_n|T_1,...,T_{k_l}}(x_1,...,x_n|\tau_1^l,...,\tau_{k_l}^l)f_{T_1,...,T_{k_l}}(\tau_1^l,...,\tau_{k_l}^l).$$

RP-2.2.2. *Fase M.* Nesta fase queremos maximizar $Q(\theta,\theta_p)$ com respeito a θ, ou seja, queremos obter o valor de θ que maximiza,

$$Q_L(\theta,\theta_p) = \frac{1}{L}\sum_{l=1}^{L}\log\left(L_c^{k_l}(X_1,...,X_n,\tau_1^l,...,\tau_{k_l}^l)\right)\times w^l.$$

Derivando $Q_L(\theta,\theta_p)$ em ordem a θ e resolvendo a equação,

$$\frac{dQ_L(\theta,\theta_p)}{d\theta} = 0 \Leftrightarrow \begin{bmatrix} \frac{\partial Q_L(\theta,\theta_p)}{\partial\mu_1} \\ \frac{\partial Q_L(\theta,\theta_p)}{\partial\mu_2} \\ \frac{\partial Q_L(\theta,\theta_p)}{\partial m} \\ \frac{\partial Q_L}{\partial M} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Para μ_1,

$$\frac{\partial Q_L(\theta,\theta_p)}{\partial\mu_1} = \frac{1}{L}\sum_{l=1}^{L}\left[\frac{\partial\log\left(L_c^{k_l}(x_1,...,x_n,\tau_1^l,...,\tau_{k_l}^l;\theta)\right)}{\partial\mu_1}w^l\right]$$

Igualando a zero e resolvendo obtem-se,

$$\mu_1\left(\sum_{l=1}^{L}\left[a_{1,l}w^l\right]\right) - \sum_{l=1}^{L}\left[b_{1,l}w^l\right] = 0, \tag{RP-4}$$

com

$$\begin{aligned} a_{1,l} = \Delta&\left(j_1^l + \sum_{i=1}^{[(k_l-1)/2]}(j_{2i+1}^l - j_{2i}^l - 1) + (n - j_{k_l}^l - 1)mod(k_l+1,2)\right) + 2\tau_1^l \\ &+ 2\sum_{i=1}^{[(k_l-1)/2]}(\tau_{2i+1}^l - \tau_{2i}^l) + \tau_{k_l}^l mod(k_l+1,2) + \sum_{i=1}^{[k_l/2]}(t_{j_{2i}^l+1} - t_{j_{2i-1}^l}) - t_{j_{k_l}^l}mod(k_l,2), \\ b_{1,l} = {}& x_n mod(k_l+1,2) - 2x_0 + M(2[(k_l+1)/2]) - m(k_l-1) \end{aligned}$$

Para μ_2 obtemos,

$$\mu_2\left(\sum_{l=1}^{L}\left[a_{2,l}w^l\right]\right) - \sum_{l=1}^{L}\left[b_{2,l}w^l\right] = 0, \tag{RP-5}$$

com

$$a_{2,l} = \Delta\left(\sum_{i=1}^{[k_l/2]}(j^l_{2i} - j^l_{2i-1} - 1) + (n - j^l_{k_l} - 1)mod(k_l, 2)\right) + 2\sum_{i=1}^{[k_l/2]}(\tau^l_{2i} - \tau^l_{2i-1})$$

$$- \tau^l_{k_l}mod(k_l, 2) + t_{j^l_1+1} + \sum_{i=1}^{[(k_l-1)/2]}(t_{j^l_{2i+1}+1} - t_{j^l_{2i}}) - t_{j^l_{k_l}}mod(k_l + 1, 2),$$

$$b_{2,l} = x_n mod(k_l, 2) + m(2[k_l/2]) - Mk_l.$$

Para m temos a equação,

$$\begin{aligned} &m^2\left(\sum_{l=1}^{L}\left[(U^l_{m_p} + W^l)w^l\right]\right) \\ &+ m\left(\sum_{l=1}^{L}\left[(-M(U^l_{m_p} + 2W^l) - V^l_{m_p} + \mu_1(k_l - 1) - \mu_2(2[k_l/2]))w^l\right]\right) \\ &+ \left(\sum_{l=1}^{L}\left[(M(V^l_{m_p} - \mu_1(k_l - 1) + \mu_2(2[k_l/2])) - \sigma^2(k_l - 1) + M^2W^l)w^l\right]\right) = 0. \end{aligned} \tag{RP-6}$$

Com

$$U^l_{m_p} = \sum_{i=1}^{[k_l/2]}\frac{\Delta}{(\tau^l_{2i} - t_{j^l_{2i}})(t_{j^l_{2i}+1} - \tau^l_{2i})}$$

$$V^l_{m_p} = \sum_{i=1}^{[k_l/2]}\frac{x_{j^l_{2i}}(t_{j^l_{2i}+1} - \tau^l_{2i}) + x_{j^l_{2i}+1}(\tau^l_{2i} - t_{j^l_{2i}})}{(\tau^l_{2i} - t_{j^l_{2i}})(t_{j^l_{2i}+1} - \tau^l_{2i})}$$

$$W^l = \sum_{i=2}^{k_l}\frac{1}{\tau^l_i - \tau^l_{i-1}}.$$

Fazendo o mesmo com respeito a M, obtemos,

(RP-7)

$$\begin{aligned} &M^3\left(\sum_{l=1}^{L}\left[\left(-U^l_{M_p} - W^l - \frac{1}{\tau^l_1}\right)w^l\right]\right) + M^2\left(\sum_{l=1}^{L}\left[\left((x_0 + m)U^l_{M_p} + V^l_{M_p}\right.\right.\right. \\ &\left.\left.\left.+\mu_1(2[(k_l + 1)/2]) - \mu_2k_l + \frac{2x_0 + m}{\tau^l_1} + (x_0 + 2m)W^l\right)w^l\right]\right) \\ &+ M\left(\sum_{l=1}^{L}\left[\left(-x_0mU^l_{M_p} - (x_0 + m)(V^l_{M_p} + \mu_1(2[(k_l + 1)/2]) - \mu_2k_l) + \sigma^2k_l - \frac{x_0^2 + 2mx_0}{\tau^l_1}\right.\right.\right. \\ &\left.\left.\left.-(m^2 + 2x_0m)W^l\right)w^l\right]\right) + \left(\sum_{l=1}^{L}\left[\left(x_0m(V^l_{M_p} + \mu_1(2[(k_l + 1)/2]) - \mu_2k_l\right.\right.\right. \\ &\left.\left.\left.+\frac{x_0}{\tau^l_1} + mW^l) - \sigma^2(m + (k_l - 1)x_0)\right)w^l\right]\right) = 0, \end{aligned}$$

com,

$$U^l_{M_p} = \sum_{i=1}^{[(k_l+1)/2]} \frac{\Delta}{(\tau^l_{2i-1} - t_{j^l_{2i-1}})(t_{j^l_{2i-1}+1} - \tau^l_{2i-1})};$$
$$V^l_{M_p} = \sum_{i=1}^{[(k_l+1)/2]} \frac{x_{j^l_{2i-1}}(t_{j^l_{2i-1}+1} - \tau^l_{2i-1}) + x_{j^l_{2i-1}+1}(\tau^l_{2i-1} - t_{j^l_{2i-1}})}{(\tau^l_{2i-1} - t_{j^l_{2i-1}})(t_{j^l_{2i-1}+1} - \tau^l_{2i-1})}.$$

Bastando, finalmente, resolver as equações para se obter θ_{p+1} e recomeçar a $(p+1)$-ésima iteração do algoritmo.

RP-2.3. Simulação. Iremos agora proceder à implementação do algoritmo usando dados si-mulados. Começamos por simular 170 observações do modelo de limiares com $\Delta = .1$. O processo é dividido em dois regimes, o primeiro será Browniano com tendência $\mu_1 = 1$ e o segundo Browniano com tendência $\mu_2 = -1$, para ambos os regimes teremos $\sigma = .4$. A mudança de regime ocorre quando o processo atinge o limiar superior $M = 3$ se o processo está no regime 1, ou quando atinge o limiar inferior $m = -3$ se o processo está no regime 2. Na p-ésima iteração do algoritmo, calculamos $L = 500$ repetições de sequências de instantes de contacto de forma a aproximar o valor esperado (RP-3) por Monte Carlo. Para construir a sequência de instantes de contacto usamos métodos numéricos para calcular a raiz τ da equação,

$$\int_0^\tau \frac{\widehat{M}_p - \widehat{m}_p}{\sqrt{2\pi\sigma^2 t^3}}\, e^{-\frac{(\widehat{M}_p - \widehat{m}_p - \widehat{\mu}_{1,p} t)^2}{2\sigma^2 t}} dt = u$$

com u número pseudo aleatório gerado a partir do gerador de $U[0,1]$ do programa. Em alternativa também é possível utilizar o método de aceitação-rejeição para construir a sequência de instantes de contacto. Após as 500 repetições estarem terminadas podemos calcular os $(p+1)$-ésimos estimadores $\widehat{\mu}_{1,p+1}, \widehat{m}_{p+1}$ e $\widehat{M}_{p+1}$ a partir das equações (RP-4), (RP-5), (RP-6) e (RP-7). 100 iterações do algoritmo foram consideradas e as sequências de estimadores obtidas são representadas nas próximas figuras.

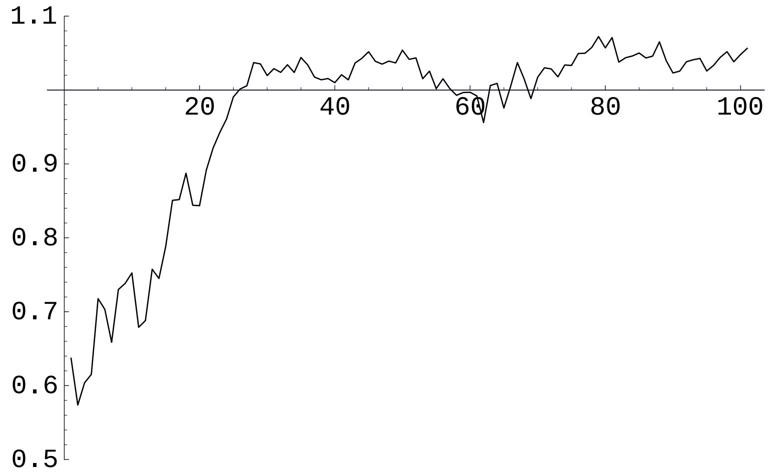

Figura 1. Valores estimados de μ_1 a partir de 100 iterações de MCEM

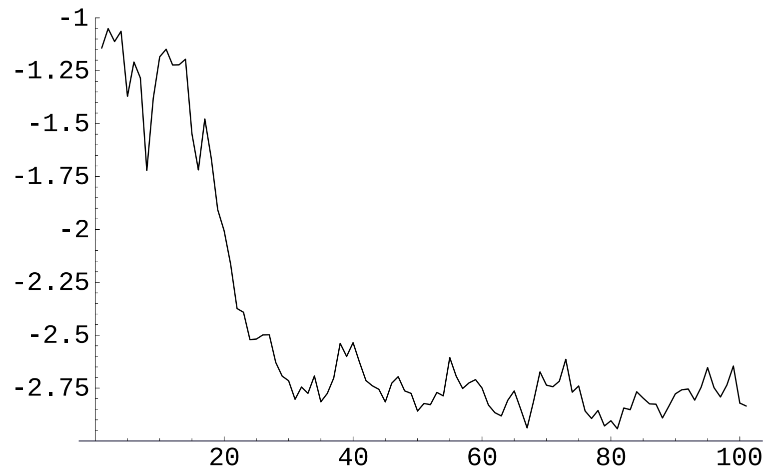

Figura 2. Valores estimados de m a partir de 100 iterações de MCEM

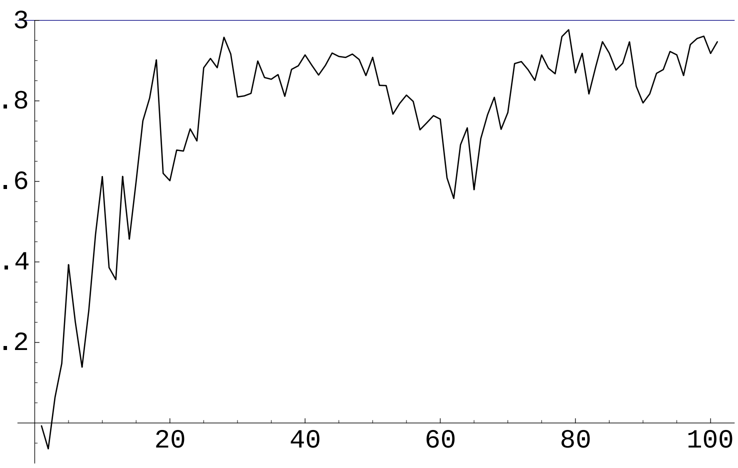

Figura 3. Valores estimados de M a partir de 100 iterações de MCEM

O valores finais dos estimadores, após 100 iterações do algoritmo, são $\widehat{\mu}_{1,101} = 1.05 = -\widehat{\mu}_{2,101}$, $\widehat{m}_{101} = -2.85$, and $\widehat{M}_{101} = 2.9$. Tendo-se obtido sub-estimação dos limiares e sobre-estimação do parâmetro de tendência. A explicação pode passar pelo facto de nos dados considerados o estimador real (podemos estimar directamente o parâmetro, uma vez que sabemos exactamente onde houve a mudança de regime) do parâmetro de tendência ser de $\widehat{\mu} = 1.2 > 1 = \mu_1$. Mais à frente iremos comparar estes resultados com os obtidos por outro procedimento de estimação. A simulação da trajectória e a implementação do algoritmo foram produzidas no programa Mathematica 4.1 e a lista de instrucções aparece no anexo 2. Repetindo o procedimento para 100 trajectórias nas mesmas condições obtiveram-se os seguintes resultados:

Tabela 1. Média e desvio padrão para os estimadores da tendência e dos limiares para 100 repetições do algoritmo MCEM

	$\widehat{\mu}$	$\widehat{m}$	$\widehat{M}$
média	1.020	-2.940	2.950
d.p.	0.118	0.186	0.174

Este método necessita de muita computação tornando-se impraticável um grande número de repetições.

RP-3. Resumo do Capítulo 3

Neste capítulo iremos introduzir um procedimento de estimação dos limiares baseado no método de estimação de mínimos quadrados e demonstrar resultados parciais de consistência dos estimadores daí resultantes. O modelo em análise pode ser representado por,

$$dXt = \mu(t)dt + \sigma dB_t, \quad X_0 = x_0, \tag{RP-8}$$

onde

$$\mu(t) = \sum_{k\geq 0} \left[\mu_0 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) - \mu_0 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_0 > 0$$

e $0 = \tau_0 < \tau_1 < \ldots < \tau_j < \ldots$ são os instantes de contacto do processo com os limiares. O nosso objectivo é provar a consistência dos estimadores quando temos um conjunto discreto de observações do processo e o regime ao qual cada uma pertence é desconhecido. Começaremos com uma primeira secção onde teremos uma prova parcial da consistência sob as hipóteses de diminuição do intervalo de discretização Δ e conhecimento do regime ao qual cada observação pertence. Na segunda secção abandonaremos a restrição de

se conhecer o regime ao qual cada observação pertence e introduziremos no procedimento de estimação uma forma prática de classificar as observações em regimes. Estudos através de simulações serão realizadas. Na terceira secção iremos alargar o procedimento de estimação a outros modelos como os construidos a partir dos processos Browniano geométrico e Ornstein-Uhlenbeck. Finalmente aplicaremos o procedimento a dados reais provenientes de fundos internacionais.

RP-3.1. Conjunto discreto de observações do processo BMD com limiares e com regimes conhecidos. Considerem-se as observações $X_1, \ldots, X_n$ do processo Browniano com limiares no intervalo $[0, t_n]$ e introduzam-se as variáveis aleatórias $R_1, \ldots, R_n$ onde $R_i = 1$ se X_i pertence ao regime 1, enquanto $R_i = 2$ se X_i pertence ao regime 2. Seja $\hat{K}_i$ o número de mudanças de valor na sequência $R_1, \ldots R_i$, iremos assumir nesta secção que são observados os valores de R_i.

Podemos definir o estimador $\widehat{\mu}_i(m, M)$ de μ_0 de uma forma natural, como:

$$\widehat{\mu}_i(m, M) = \frac{\widehat{K}_i}{t_i}(M - m). \tag{RP-9}$$

Temos os seguintes resultados.

Lema RP-3.1. Com K_i o verdadeiro número de mudanças de regime no intervalo $[0, t_i]$, temos:

$$\lim_{i \to +\infty} \frac{K_i}{t_i} = \frac{\mu_0}{M_0 - m_0}, \quad q.c.$$

Proof. Ver capítulo 3, lema 3.1. □

Temos um resultado similar para $\frac{\hat{K}_i}{t_i}$.

Note-se que uma mudança de regime não é detectada, quando se calcula $\hat{K}_i$, se ocorrer uma das situações seguintes:

(1) temos na sequência $..., X_j \in R_1, X_{j+1} \in R_1, ...$ mas temos $2p$, para algum $p \geq 1$, mudanças de regime no intervalo $]t_j, t_{j+1}[$,

(2) temos na sequência $..., X_j \in R_1, X_{j+1} \in R_2, ...$ mas temos $2p + 1$, para algum $p \geq 1$, mudanças de regime no intervalo $]t_j, t_{j+1}[$.

Em ambos os casos tem de ocorrer mais do que uma mudança de regime no intervalo $]t_j, t_{j+1}[$.

Lema RP-3.2. Sendo $\hat{K}_i$ o número de mudanças de valor na sequência $R_1, \ldots R_i$. Temos que $\lim_{i \to \infty} \frac{K_i - \hat{K}_i}{t_i} = 0$ q.c.

Proof. Ver capítulo 3, lema 3.2. □

Temos também.

Lema RP-3.3. Definindo $\hat{\mu}_i = \hat{\mu}_i(m_0, M_0)$ como na equação (RP-9), quando m_0 e M_0 são os verdadeiros valores dos limiares, teremos:

$$\lim_{i \to +\infty} \hat{\mu}_i = \mu_0, \quad q.c.$$

Proof. Ver capítulo 3, lema 3.3. □

Num cenário ídeal definiriamos a soma dos quadrados do erros (condicional) como,

$$LSE_n(m, M) = \sum_{i=1}^{n-1} \left(X_{i+1} - \mathbb{E}_{\mu,m,M} \left[X_{i+1} | X_i, ..., X_1 \right] \right)^2,$$

mas a esperança matemática não é conhecida pelo que teremos de considerar uma aproximação e introduzir uma função soma do quadrados dos erros auxiliar:

$$LS_n(m, M) = \sum_{i=1}^{n-1} (X_{i+1} - X_i - \hat{\mu}_i(m, M)\Delta)^2 1_{R_i=R_{i+1}=1} + \sum_{i=1}^{n-1} (X_{i+1} - X_i + \hat{\mu}_i(m, M)\Delta)^2 1_{R_i=R_{i+1}=2}, \tag{RP-10}$$

onde, como anteriormente,

$$\hat{\mu}_i(m, M) = \frac{\hat{K}_i}{t_i}(M - m). \tag{RP-11}$$

Para justificar esta aproximação prova-se o seguinte lema.

Lema RP-3.4. Temos:

$$\mathbb{E}_{\mu_0,m_0,M_0}[X_{i+1} - X_i | X_i, R_i] 1_{R_i=R_{i+1}=1} = \mu_0 \Delta 1_{R_i=R_{i+1}=1} + A(m_0, M_0, \Delta),$$

onde

$$\mathbb{E}\left[A(m_0, M_0, \Delta)\right] < 2\mu_0\Delta \times \left[\frac{2\sqrt{\Delta}}{\sqrt{2\pi}} \frac{e^{-\frac{(M_0 - m_0 - \mu_0\Delta)^2}{2\Delta}}}{(M_0 - m_0 - \mu_0\Delta)} \right]^{1/2}.$$

Proof. Ver capítulo 3, lema 3.4. □

Da função estimadora (RP-10) podemos estimar a diferença $M - m$.

Lema RP-3.5. O valor da diferença $M - m$ que minimiza a função estimadora (RP-10) é dado por:

$$\widehat{M - m} = \frac{\sum_{i=1}^{n-1} \frac{\hat{K}_i}{t_i}(X_{i+1} - X_i) 1_{R_i=R_{i+1}=1} - \sum_{i=1}^{n-1} \frac{\hat{K}_i}{t_i}(X_{i+1} - X_i) 1_{R_i=R_{i+1}=2}}{\Delta \sum_{i=1}^{n-1} \frac{\hat{K}_i^2}{t_i^2} 1_{R_i=R_{i+1}}} \tag{RP-12}$$

PROOF. Ver capítulo 3, lema 3.5. □

Podemos agora escrever a prova da consistência para o estimador da diferença $M_0 - m_0$.

TEOREMA RP-3.6. O estimador $\widehat{M-m}$ que minimiza a função $LS_n(m, M)$ é consistente.

PROOF. Ver capítulo 3, teorema 3.6. □

RP-3.2. Conjunto discreto de observações do processo BMD com limiares e com regimes desconhecidos. Nesta secção o nosso objectivo é implementar uma forma prática de estimação dos estimadores para cada um dos limiares quando não se conhece o regime ao qual cada observação pertence. O primeiro passo será por isso o de classificar as observações em regimes de modo a se construir a função da soma do quadrados dos erros como na secção anterior. O procedimento de estimação será implementado da seguinte forma.

(1) Para limiares fixos m e M, classificaremos as observações em regimes, $\widehat{R}_1(m, M)$ e $\widehat{R}_2(m, M)$ correspondentes aos regimes com tendência positiva e com tendência ne-gativa, respectivamente. A classificação em regimes é feita da seguinte forma: fixando (m, M), as observações vão sendo classificadas no regime 1 até encontrarmos a primeira observação que é maior ou igual a M, sendo essa observação classificada no regime 1 e simultâneamente no regime 2, todas as observações daí em diante serão classificadas no regime 2 até encontrarmos uma observação que seja menor ou igual a m sendo essa observação classificada, ainda, no regime 2 mas simultâneamente no regime 1, o processo de classificação continua até todas as observações estarem classificadas. Também se obtém um valor $\widehat{K}_n$, o número estimado de mudanças de regime.
(2) De seguida, e com base nesta classificação calcula-se o valor $\widehat{\mu}_n$ o estimador condicional de μ_0 de uma forma similar ao que já foi feito na secção anterior mas considerando agora toda a informação, isto é, o número estimado de mudanças de regime em $[0, t_n]$ e não apenas no intervalo $[0, t_i]$. Define-se a função da soma do quadrados dos erros de uma forma análoga mas com as funções indicatrizes a dependerem apenas do regime da última observação em vez de dependerem dos

regimes de duas observações consecutivas.

(RP-13)

$$LS_n(m,M)=\sum_{i=1}^{n-1}(X_{i+1}-X_i-\widehat{\mu}_n(m,M)\Delta)^2\,1_{\widehat{R}_i=1}+\sum_{i=1}^{n-1}(X_{i+1}-X_i+\widehat{\mu}_n(m,M)\Delta)^2\,1_{\widehat{R}_i=2}.$$

(3) Finalmente escolhem-se para estimadores dos limiares os valores $(\widehat{m},\widehat{M})$ que minimizam $LS_n(m,M)$, ou seja,

$$(\widehat{m}_n,\widehat{M}_n)=\text{argmin}_{(m,M)}LS_n(m,M),$$

e estima-se o coeficiente de tendência por

$$\widehat{\mu}_n=\widehat{\mu}_n(\widehat{m}_n,\widehat{M}_n).$$

Observação RP-3.7. Repare-se que função soma do quadrados dos erros (RP-13) depende da diferença $M-m$ através de $\widehat{\mu}_n$ como anteriormente, mas também depende de M e m por intermédio da sequência $\widehat{R}_1,...,\widehat{R}_n$. Devido a esta dependência podemos estimar os limiares m e M separadamente pois a diferentes pares (m,M) correspondem diferentes sequências $\widehat{R}_1,...,\widehat{R}_n$ e consequentemente diferentes valores de $LS_n(m,M)$.

Pensamos que este procedimento deverá permitir obter estimadores consistentes quando Δ decresce para zero e simultaneamente o intervalo onde se observa o processo, $[0,t_n]$ aumenta. Estamos agora em condições de implementar este procedimento através de simulações. A aplicação do procedimento foi realizado sob diferentes condições desde valor fixo para Δ e para o intervalo de observações $[0,t]$, passando por Δ decrescente e intervalo de observações fixo e terminando com $\Delta=\Delta_n$ decrescente e intervalo de observações $[0,t_n]$ crescente. Dos diferentes cenários apenas se apresenta aqui o último estudado podendo consultar-se o capítulo 3 para mais informação.

Exemplo RP-3.8. Consideramos o processo Browniano com tendência e com limiares, e os valores $\mu_0=1$, $m_0=-2$ e $M_0=2$, e $\sigma=.9$. Os resultados da simulaçã são baseados no seguinte. Cada trajectória é inicialmente gerada com discretização $\delta=2^{-7}$ no intervalo $[0,128]$ para o procedimento de estimação consideraremos as observações com discretização $\Delta_k=\frac{1}{2^k}$ para $k=4,...,7$ correspondendo ao intervalo de observações $[0,2^k]$ e a um número de observações 4^k. Calcula-se LS_n para valores de $(m,M)\in[-2.5,1]\times[1,2.5]$ usando uma grelha com passo .1. Na figura 4 apresentamos uma trajectória para $k=5$ e na tabela 2 são apresentados os resultados obtidos.

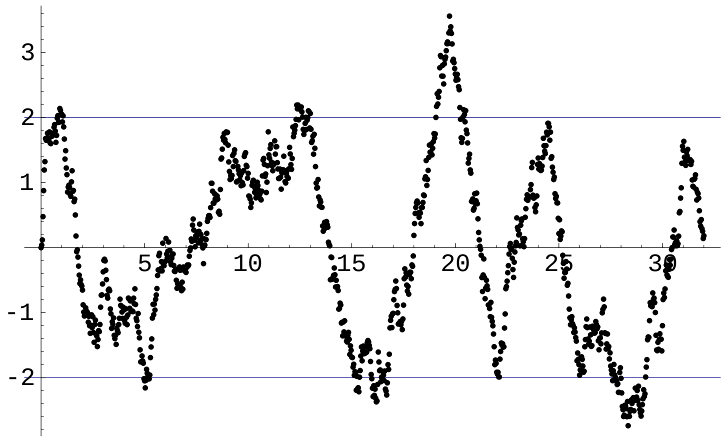

Figura 4. Trajectória do processo BMD com limiares para $\Delta = 1/32$, intervalo de observações $[0, 32]$ e $\mu_0 = 1$, $m_0 = -2$, $M_0 = 2$, $\sigma = .9$

Tabela 2. Estimativas para 100 repetições do processo BMD com limiares, $\Delta_k = 2^{-k}$ decrescente, intervalo de observações $[0, 2^k]$ crescente, número de observações $n_k = 4^k$ e $\mu_0 = 1$, $m_0 = -2$, $M_0 = 2$, $\sigma = .9$

	$média(\widehat{\mu}_n)$	$dp(\widehat{\mu}_n)$	$média(\widehat{m}_n)$	$dp(\widehat{m}_n)$	$média(\widehat{M}_n)$	$dp(\widehat{M}_n)$
$k = 4$	0.981	0.216	-1.954	0.103	1.944	0.115
$k = 5$	1.011	0.161	-1.956	0.075	1.955	0.080
$k = 6$	1.010	0.107	-1.983	0.052	1.976	0.055
$k = 7$	0.990	0.077	-1.986	0.015	1.984	0.016

Uma observação final, seria interessante comparar de forma intensiva os estimadores de minímos quadrados com os estimadores obtidos do algoritmo MCEM estudado no capítulo 2. Numa primeira tentativa aplicou-se o procedimento de minimização da soma dos quadrados dos erros estudado neste capítulo aos dados simulados no capítulo 2 e obtiveram-se resultados similares com valores $\widehat{\mu}_n = 1.04387$, $\widehat{m}_n = -2.995$ e $\widehat{M}_n = 2.845$.

RP-3.3. Estimadores de mínimos quadrados para difusões. Iremos agora generalizar o procedimento de estimação dos limiares a outros processos que não apenas o construido do processo Browniano com tendência. Iremos apresentar as funções estimadoras para processos com limiares construidos a partir dos processos Ornstein-Uhlenbeck e Browniano geométrico. Iremos aplicar o procedimento de estimação a estes processos e também a dados reais provenientes de um conjunto de fundos internacionais.

Pretende-se estudar o modelo geral com limiares:

$$dXt = a(\mu(t), X_t)dt + b(\sigma, X_t)dB_t, \tag{RP-14}$$

com

$$\mu(t) = \sum_{k\geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_1 \in \Theta_1, \mu_2 \in \Theta_2$$

onde $0 = \tau_0 < \tau_1 < \ldots < \tau_j < \ldots$ são os instantes de contacto nos limiares . Os conjuntos Θ_1 e Θ_2 são tais que, para $\mu \in \Theta_1$, $X_t(\mu)$ está no regime 1 (regime crescente) e para $\mu \in \Theta_2$, $X_t(\mu)$ está no regime 2 (regime decrescente). Define-se a função soma do quadrado dos erros (condicional),

$$LSE_n(m, M) = \sum_{i=1}^{n-1} \left(X_{i+1} - \mathbb{E}_{\mu,m,M}\left[X_{i+1}|X_i, R_i\right]\right)^2,$$

para o processo geral, mas como anteriormente teremos de usar funões auxiliares.

RP-3.3.1. *Processo Browniano geométrico (GBM) com limiares.* Partindo do processo Browniano geométrico constrói-se o processo com limiares correspondente.

$$dXt = \mu(t)X_t dt + \sigma X_t dB_t, \quad X_0 = x_0, \tag{RP-15}$$

onde

$$\mu(t) = \sum_{k\geq 0} \left[\mu_0 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) - \mu_0 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_0 > \sigma^2/2.$$

Para este modelo em particular iremos utilizar a função estimadora auxiliar:

$$LS_n(m, M) = \sum_{i=1}^{n-1} (X_{i+1} - X_i e^{\hat{\mu}_n(m,M)\Delta})^2 1_{\widehat{R}_i=1} + \sum_{i=1}^{n-1} (X_{i+1} - X_i e^{-\hat{\mu}_n(m,M)\Delta})^2 1_{\widehat{R}_i=2}, \tag{RP-16}$$

onde se usa a aproximação

$$\mathbb{E}_{\mu,m,M}\left[X_{i+1}|X_i, R_i\right] = X_i e^{\hat{\mu}_n(m,M)\Delta}$$

ou

$$\mathbb{E}_{\mu,m,M}\left[X_{i+1}|X_i, R_i\right] = X_i e^{-\hat{\mu}_n(m,M)\Delta}$$

dependendo do regime, porque esta é a esperança condicional do processo Browniano geométrico sem limiares. Usando as mesmas idéias da secção anterior definiremos o estimador,

$$\hat{\mu}_n(m, M) = \frac{\sigma^2}{2} + \frac{\hat{K}_n}{t_n} \ln\left(\frac{M}{m}\right). \tag{RP-17}$$

onde σ^2 deve ser substituido por $\hat{\sigma}^2$ se for desconhecido e onde $\hat{K}_n$ continua a ser o valor estimado de mudanças de regime no intervalo $[0, t_n]$.

Uma vez mais a aplicação do procedimento foi realizado sob diferentes condições mas apenas se apresenta aqui o caso de $\Delta = \Delta_n$ decrescente e intervalo de observações $[0, t_n]$ crescente. Podendo consultar-se o capítulo 3 para mais informação.

Consideramos o processo Browniano geométrico com limiares, e os valores $\mu_1 = 1, \mu_2 = -1$, $m_0 = 5$ e $M_0 = 15$, e $\sigma = .6$. Os resultados da simulação são baseados no seguinte. Cada trajectória é inicialmente gerada com discretização $\delta = 2^{-7}$ no intervalo $[0, 128]$ para o procedimento de estimação consideraremos as observações com discretização $\Delta_k = \frac{1}{2^k}$ para $k = 4, ..., 7$ correspondendo ao intervalo de observações $[0, 2^k]$ e a um número de observações 4^k. Calcula-se LS_n para valores de $(m, M) \in [4, 7] \times [11, 16]$ usando uma grelha com passo .1. Na figura 5 apresentamos uma trajectória para $k = 5$ e na tabela 3 são apresentados os resultados obtidos.

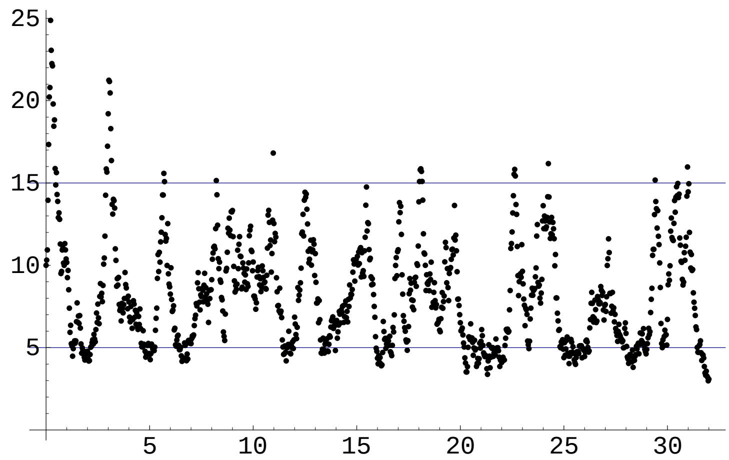

Figura 5. Trajectória do processo GBM com limiares, $\Delta = 1/32$, intervalo de observações $[0, 32]$ e $\mu_1 = 1 = -\mu_2$, $m_0 = 5$, $M_0 = 15$, $\sigma = .6$

Como se verifica o procedimento parece funcionar bem.

Tabela 3. Estimativas para 100 repetições do processo GBM com limiares $\Delta_k = 2^{-k}$ decrescente, intervalo de observações $[0, 2^k]$, número de observações $n_k = 4^k$ e $\mu_1 = 1 = -\mu_2$, $m_0 = 5$, $M_0 = 15$, $\sigma = .6$

	$média(\widehat{\mu}_{1,n})$	$dp(\widehat{\mu}_{1,n})$	$média(\widehat{\mu}_{2,n})$	$dp(\widehat{\mu}_{2,n})$	$média(\widehat{m}_n)$	$dp(\widehat{m}_n)$	$média(\widehat{M}_n)$	$dp(\widehat{M}_n)$
$k = 4$	0.830	0.161	−0.949	0.295	5.412	0.296	14.364	0.499
$k = 5$	0.920	0.136	−0.959	0.184	5.374	0.160	14.450	0.358
$k = 6$	0.961	0.099	−0.974	0.126	5.236	0.137	14.522	0.276
$k = 7$	0.984	0.059	−1.007	0.089	5.022	0.076	14.958	0.067

RP-3.3.2. *Processo Ornstein-Uhlenbeck (OU) com limiares.* Partindo do processo Ornstein-Uhlenbeck constrói-se o processo com limiares correspondente.

$$dXt = \mu(t)X_t dt + \sigma dB_t, \quad X_0 = x_0, \tag{RP-18}$$

onde

$$\mu(t) = \sum_{k\geq 0} \left[\mu_1 \mathbb{I}_{[\tau_{2k},\tau_{2k+1}[}(t) + \mu_2 \mathbb{I}_{[\tau_{2k+1},\tau_{2k+2}[}(t)\right], \quad \mu_1 > 0, \mu_2 < 0$$

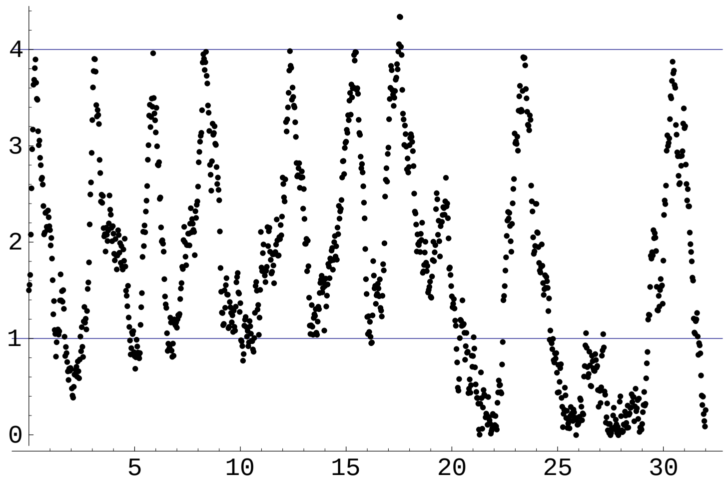

Figura 6. Trajectória do processo Ornstein-Uhlenbeck com limiares com $\Delta = 1/32$, intervalo de observações $[0, 32]$ e $\mu_1 = 1 = -\mu_2$, $m_0 = 1$, $M_0 = 4$, $\sigma = 1$

Para este modelo em particular iremos utilizar a função estimadora auxiliar:
(RP-19)

$$LS_n(m, M) = \sum_{i=1}^{n-1}(X_{i+1} - X_i e^{\hat{\mu}_{1,n}(m,M)\Delta})^2 1_{\widehat{R}_i=1} + \sum_{i=1}^{n-1}(X_{i+1} - X_i e^{\hat{\mu}_{2,n}(m,M)\Delta})^2 1_{\widehat{R}_i=2},$$

porque a esperança condicional para o processo Ornstein-Uhlenbeck sem limiares é a mesma que a do processo Browniano geométrico e uma vez mais pensamos que para valores pequenos de Δ esta deve ser uma boa aproximação. Para este modelo definem-se os estimadores de μ_1 e μ_2 usando os estimadores usuais para o processo Ornstein-Uhlenbeck mas considerando as observações em cada regime para estimar os parâmetros correspondentes.

(RP-20)
$$\hat{\mu}_1(m, M) = \frac{1}{\Delta} \ln\left(\frac{\sum_{i=1}^{n-1}(X_i X_{i+1}) 1_{\hat{R}_i(m,M)=1}}{\sum_{i=1}^{n-1} X_i^2 1_{\hat{R}_i(m,M)=1}}\right),$$

e

(RP-21)
$$\hat{\mu}_2(m, M) = \frac{1}{\Delta} \ln\left(\frac{\sum_{i=1}^{n-1}(X_i X_{i+1}) 1_{\hat{R}_i(m,M)=2}}{\sum_{i=1}^{n-1} X_i^2 1_{\hat{R}_i(m,M)=2}}\right).$$

onde a dependência de m e M advém da classificação $\hat{R}_i(m, M) = 1$ ou $\hat{R}_i(m, M) = 2$.

Consideramos o processo Ornstein-Uhlenbeck com limiares, e os valores $\mu_1 = 1, \mu_2 = -1$, $m_0 = 1$ e $M_0 = 4$, e $\sigma = 1$. Calcula-se LS_n para valores de $(m, M) \in [0.5, 2.4] \times$

$[2.6, 4.5]$ usando uma grelha com passo .1. Na tabela 4 são apresentados os resultados obtidos.

Tabela 4. Estimativas para 100 repetições do processo OU com limiares $\Delta_k = 2^{-k}$ decrescente, intervalo de observações $[0, 2^k]$, número de observações $n_k = 4^k$ e $\mu_1 = 1 = -\mu_2$, $m_0 = 1$, $M_0 = 4$, $\sigma = 1$

	$média(\widehat{\mu}_{1,n})$	$dp(\widehat{\mu}_{1,n})$	$média(\widehat{\mu}_{2,n})$	$dp(\widehat{\mu}_{2,n})$	$média(\widehat{m}_n)$	$dp(\widehat{m}_n)$	$média(\widehat{M}_n)$	$dp(\widehat{M}_n)$
$k = 4$	0.941	0.147	−0.915	0.191	1.118	0.078	3.742	0.091
$k = 5$	0.954	0.118	−0.934	0.101	1.098	0.056	3.830	0.058
$k = 6$	0.963	0.094	−0.942	0.079	1.056	0.039	3.890	0.043
$k = 7$	0.984	0.062	−0.971	0.044	1.001	0.009	3.990	0.010

Também neste caso o procedimento parece funcionar bem.

RP-3.3.3. *Dados reais.* Também se aplicou o procedimento de estimação a um conjunto de dados reais. Não procedemos a nenhum tipo de teste de ajustamento dos modelos e optámos por aplicar o procedimento aos três modelos considerados anteriormente. Os dados são respeitantes a cotações diárias de três fundos diferentes, dois dos quais fundos predominantemente compostos por acções e um terceiro composto predominantemente por obrigações.

Exemplo RP-3.9. O primeiro fundo a ser considerado é o fundo PF-European Sustainable Equities-R gerido pela empresa Pictet Funds. Os dados considerados são respeitantes ao ano de 2004.

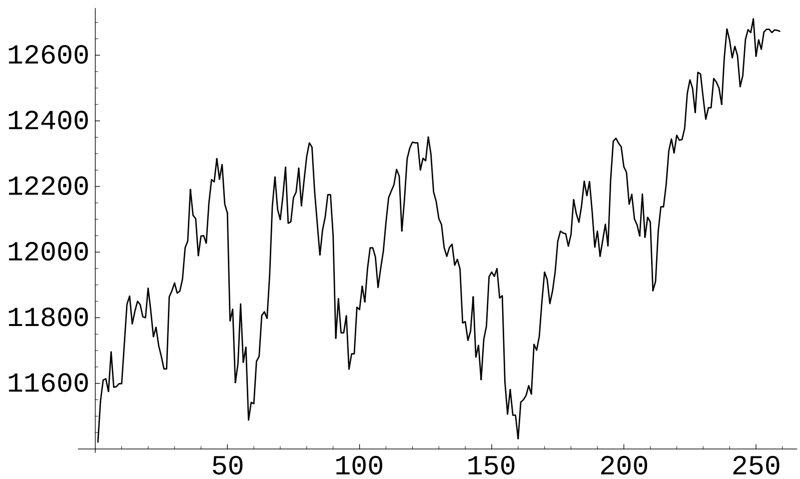

Figura 7. Dados de 2004 do fundo European Sustainable Equities-R da Pictet Funds

Apenas consideramos os dados dos primeiros 10 meses, porque como se pode ver na figura 7, nos dois últimos meses de 2004 houve uma alteração na dinâmica do fundo. Supondo

que o modelo com limiares subjacente à dinâmica do fundo é Browniano com tendência (BMD), Browniano geometrico (GBM) ou Ornstein-Uhlenbeck (OU), respectivamente. Calculando-se LS_n para valores de $(m, M) \in [114, 118] \times [118, 124]$ numa grelha com malha .1. Obtiveram-se os seguintes resultados:

Tabela 5. Estimativas dos parâmetros para o fundo European Sustainable Equities-R

	$\widehat{\mu}_{1,n}$	$\widehat{\mu}_{2,n}$	$\widehat{m}_n$	$\widehat{M}_n$	LS_n
BMD	0.203	-0.365	116.500	122.800	148.337
GBM	0.002	-0.003	116.500	122.800	148.453
OU	0.002	-0.003	116.500	122.800	148.450

Os resultados obtidos não diferem de um modelo para outro, o que nos leva a suspeitar que mesmo que não se considere o modelo mais correcto pode-se obter bons estimadores para os limiares desde que se considere um modelo que possa ter um ajustamento aceitável. Os valores de LSE_n serão de interesse apenas para nos dar uma indicação de qual modelo melhor se ajusta aos dados.

Exemplo RP-3.10. O segundo fundo a ser considerado é o fundo Parvest Europe Dynamic Growth da empresa BNP Paribas. Os dados considerados são do ano de 2004.

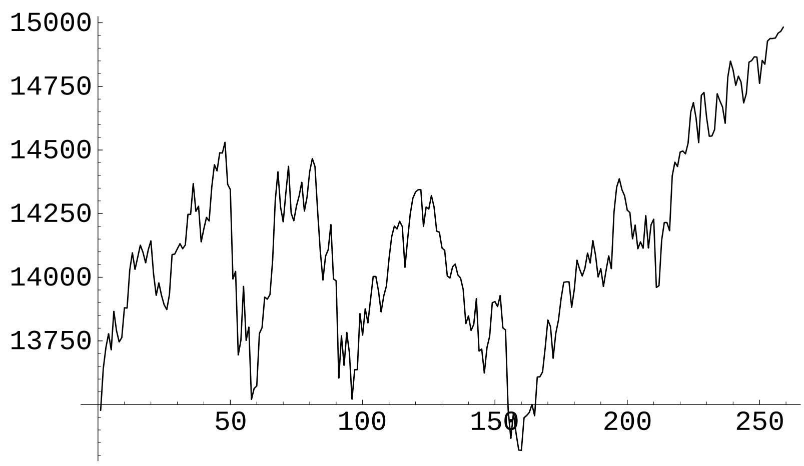

Figura 8. Dados de 2004 do fundo Parvest Europe Dynamic Growth de BNP Paribas

Uma vez mais apenas consideramos os dados dos primeiros 10 meses de 2004, e calculou-se LS_n para os valores de $(m, M) \in [134, 139] \times [140, 145]$ numa grelha de malha .1.

Obtiveram-se os resultados apresentados na próxima tabela.

Tabela 6. Estimativas dos parâmetros para o fundo Parvest Europe Dynamic Growth

	$\widehat{\mu}_{1,n}$	$\widehat{\mu}_{2,n}$	$\widehat{m}_n$	$\widehat{M}_n$	LS_n
BMD	0.314	-0.296	135.300	143.400	197.854
GBM	0.002	-0.002	135.300	143.400	197.881
OU	0.002	-0.002	135.300	143.400	197.878

Uma vez mais os estimadores são idênticos para todos os modelos.

Exemplo RP-3.11. O último fundo a ser considerado é o fundo Converging Europe Bond da empresa Schroder. Os dados considerados são do ano de 2005. Desta vez considerou-se os dados da totalidade do ano e calculou-se LS_n para valores de $(m, M) \in [1364, 1380] \times [1386, 1401]$ numa grelha com malha .5.

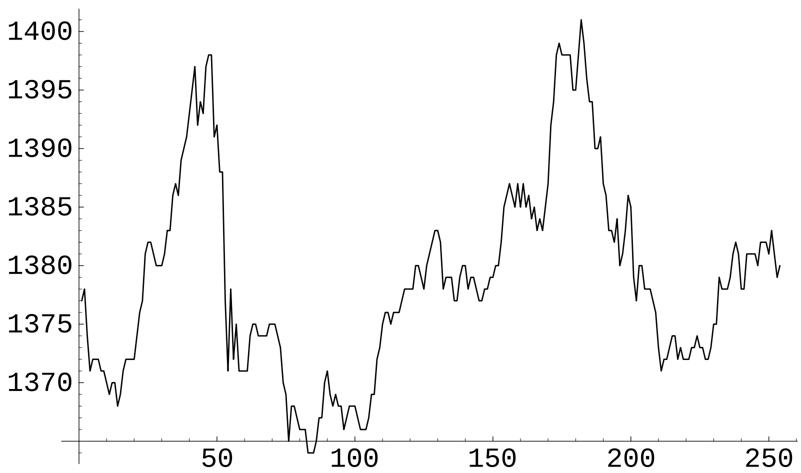

Figura 9. Dados de 2005 referentes ao fundo Converging Europe Bond da Schroder

Obteve-se:

Tabela 7. Estimativas dos parâmetros para o fundo Converging Europe Bond

	$\widehat{\mu}_{1,n}$	$\widehat{\mu}_{2,n}$	$\widehat{m}_n$	$\widehat{M}_n$	LS_n
BMD	0.3153	-1.250	13.710	13.980	916.823
GBM	0.0002	-0.001	13.710	13.980	916.781
OU	0.0002	-0.001	13.710	13.980	916.781

Uma vez mais os estimadores dos limiares são idênticos para todos os modelos.

RP-4. Conclusão

Em conclusão, nesta tese desenvolvemos e implementámos dois procedimentos de estimação distintos para o problema dos modelos com limiares do tipo difusão. O primeiro procedimento MCEM é apenas uma adaptação do algoritmo Expectation Maximization e o segundo baseia-se no procedimento de estimação baseado na minimização do quadrado dos erros. Para investigação futura podermos referir alguns temas como sejam: a demonstração da consistência dos estimadores dos limiares individualmente, a generalização do modelo a processos onde a mudança de regime seja conseguida através de uma alteração completa do processo, isto é, por exemplo num regime Browniano com tendência e noutro Browniano geométrico. Outros procedimentos de estimação podem ser considerados, como teoria Bayesiana ou outros métodos estudados no contexto das séries temporais. Há ainda a questão da construção de testes para testar o modelo de limiares contra o modelo simples, e finalmente, tendo em conta os resultados obtidos pela análise de dados reais pode-se considerar o problema de usar o Browniano com tendência para estimar os limiares de outros processos com limiares, mesmo que os diferentes regimes não sejam processos Brownianos com tendência.

Appendix 1: Math. 4.1 instructions to generate a threshold trajectory

$<<Statistics`NormalDistribution`;$

$<<Statistics`ConfidenceIntervals`;$

$<<Statistics`DescriptiveStatistics`;$

$RandomNormal[\mu\_, \sigma\_] := Random[NormalDistribution[\mu, \sigma]];$

$\alpha 1 = 1; \beta 1 = 1; \alpha 2 = -1; \beta 2 = 1; m0 = 1; M0 = 4; nsimu = 9;$

$a[\theta\_, s\_, x\_] := \theta * x; b[\theta\_, s\_, x\_] := \theta;$

$SeedRandom[1]; x0 = 1.5; Xt = x0; t = 0; \delta = \frac{1}{2^{nsimu}}; T = 2^{nsimu};$

$lista1inicial = \{\}; lista2 = \{\}; interruptor = 0;$

$While[t <= T, If[interruptor == 0, \{While[And[Xt < M0, t <= T],$

$\{lista1inicial = Insert[lista1inicial, \{N[t], Xt\}, -1];$

$Xt = Xt + a[\alpha 1, t, Xt] * \delta + \beta 1 * RandomNormal[0, \sqrt{\delta}]; t = t + \delta\}];$

$interruptor = 1; lista2 = Insert[lista2, N[t], -1]; Xt = M0\},$

$\{While[And[Xt > m0, t <= T], \{lista1inicial = Insert[lista1inicial, \{N[t], Xt\}, -1];$

$Xt = Xt + a[\alpha 2, t, Xt] * \delta + \beta 2 * RandomNormal[0, \sqrt{\delta}]\}; t = t + \delta];$

$interruptor = 0; lista2 = Insert[lista2, N[t], -1]; Xt = m0\}]];$

Appendix 2: Math. 4.1 instructions to compute the BMD MCEM estimators

$<< Statistics`NormalDistribution`; << Statistics`DescriptiveStatistics`;$

$RandomNormal[\mu\_, \sigma\_] := Random[NormalDistribution[\mu, \sigma]];$

$\delta = .1; T = 17; \alpha 1 = 1; \beta 1 = .4; \alpha 2 = -1; \beta 2 = .4; m0 = -3; M0 = 3; L = 500;$

$MaxStep = 100; x0 = 0; Xt = x0; t = 0; a[\theta\_, s\_, x\_] := \theta; b[\theta\_, s\_, x\_] := \theta;$

$SeedRandom[0]; lista1 = \{\}; lista2 = \{\}; lista3 = \{\}; interruptor = 0;$

$While[t <= T, If[interruptor == 0, \{While[And[Xt <= M0, t <= T],$

$\{lista1 = Insert[lista1, Xt, -1]; t = t + \delta;$

$Xt = RandomNormal[Xt + a[\alpha 1, t, Xt] * \delta, b[\beta 1, t, Xt] * \sqrt{\delta}]\}]; interruptor = 1;$

$lista2 = Insert[lista2, N[t] + .05, -1]; Xt = M0\},$

$\{While[And[Xt >= m0, t <= T], \{lista1 = Insert[lista1, Xt, -1]; t = t + \delta;$

$Xt = RandomNormal[Xt + a[\alpha 2, t, Xt] * \delta, b[\beta 2, t, Xt] * \sqrt{\delta}]\}]; interruptor = 0;$

$lista2 = Insert[lista2, N[t] + .05, -1]; Xt = m0\}]]; ListPlot[lista1];$

$lista2 = Drop[lista2, -1]; \sigma = \beta 1; c = Length[lista1]; d = Length[lista2];$

$q1 = Quantile[lista1, .2]; q2 = Quantile[lista1, .8]; m1 = q1; M1 = q2;$

$xx1 = Last[lista1]; xx2 = Extract[lista1, c - 1]; i = 0;$

$If[Mod[d, 2] == 0, While[Not[xx2 >= q1 > xx1], i = i + 1;$

$xx1 = Extract[lista1, c - i]; xx2 = Extract[lista1, c - i - 1]],$

$While[Not[xx2 < q2 <= xx1], i = i + 1;$

$xx1 = Extract[lista1, c - i]; xx2 = Extract[lista1, c - i - 1]]];$

$\mu 1 = 1/c * \delta + (c - i - 1/10 + .05)((2 * d - 2) * (M1 - m1) + 2 * (M1 - x0) +$

$(M1 - Last[lista1]) * Mod[d, 2] + (Last[lista1] - m1) * Mod[d + 1, 2]);$

$Print[''\theta 0 = \{'', \mu 1, '', '', m1, '', '', M1, ''\}'']; PutAppend[\{0, \mu 1, m1, M1\}, ''estimadores''];$

$For[p = 1, p <= MaxStep, SeedRandom[p - 1]; For[l = 1, l <= L, u = Random[];$

$z = w/.FindRoot[\int_0^w \frac{M1 - x0}{\sqrt{2\pi\sigma^2 s^3}} \exp\left[-\frac{(M1 - x0 - \mu 1 s)^2}{2 * \sigma^2 * s}\right] \partial s == u, \{w, M1 - x0\}];$

$hit = z; w[l] = \frac{M1 - x0}{\sqrt{2\pi\sigma^2 z^3}} \exp\left[-\frac{(M1 - x0 - \mu 1 z)^2}{2 * \sigma^2 * s}\right]; lista5 = \{hit\}; While[hit < c/10,$

$u = Random[]; z = w/.FindRoot[\int_0^w \frac{M1 - m1}{\sqrt{2\pi\sigma^2 s^3}} * \exp[-\frac{(M1 - m1 - \mu 1 s)^2}{2\sigma^2 s} \partial s == u,$

$\{w, M1 - m1\}]; hit = hit + z; w[l] = w[l] * \frac{M1 - m1}{\sqrt{2\pi\sigma^2 z^3}} * \exp[-\frac{(M1 - m1 - \mu 1 z)^2}{2\sigma^2 z};$

$lista5 = Insert[lista5, hit, -1]]; If[Last[lista5] * 10 > c,$

$lista5 = Drop[lista5, -1]]; d = Length[lista5]; k_m[l] = IntegerPart[d/2];$

$lista2 = Table[IntegerPart[10 * Extract[lista5, i]]/10 + .05, \{i, d\}];$

$Print[l, ''\ '', lista2]; k_M[l] = IntegerPart[(d + 1)/2]; k[l] = k_m[l] + k_M[l]; f_{X|T}[l] = 1;$

$For[j = 1, j <= IntegerPart[10 * Extract[lista2, 1]] - 1,$

$f_{X|T}[l] = f_{X|T}[l] * 1/\sqrt{2 * \pi * \sigma^2 * \delta} *$

$Exp[-1/2 * \sigma^2 * \delta((Extract[lista1, j + 1] - Extract[lista1, j] - \mu 1 * \delta)^2)]; j + +];$

$For[i = 1, i <= k_M[l] - 1, For[j = IntegerPart[10 * Extract[lista2, 2 * i]] + 1,$

$j <= IntegerPart[10 * Extract[lista2, 2 * i + 1]] - 1,$

$f_{X|T}[l] = f_{X|T}[l] * 1/\sqrt{2 * \pi * \sigma^2 * \delta} *$

$Exp[-1/2 * \sigma^2 * \delta((Extract[lista1, j + 1] - Extract[lista1, j] - \mu 1 * \delta)^2)]; j + +]; i + +];$

$For[i = 1, i <= k_m[l], For[j = IntegerPart[10 * Extract[lista2, 2 * i - 1]] + 1,$

$j <= IntegerPart[10 * Extract[lista2, 2 * i]] - 1,$

$f_{X|T}[l] = f_{X|T}[l] * 1/\sqrt{2 * \pi * \sigma^2 * \delta} *$

$Exp[-1/2 * \sigma^2 * \delta((Extract[lista1, j + 1] - Extract[lista1, j] + \mu 1 * \delta)^2)]; j + +]; i + +];$

$If[Mod[k[l], 2] == 0, For[j = IntegerPart[10 * Last[lista2]] + 1, j <= c - 1,$

$f_{X|T}[l] = f_{X|T}[l] * 1/\sqrt{2 * \pi * \sigma^2 * \delta} *$

$Exp[-1/2 * \sigma^2 * \delta((Extract[lista1, j + 1] - Extract[lista1, j] - \mu 1 * \delta)^2)]; j + +],$

$For[j = IntegerPart[10 * Last[lista2]] + 1, j <= c - 1,$

$f_{X|T}[l] = f_{X|T}[l] * 1/\sqrt{2 * \pi * \sigma^2 * \delta} *$

$Exp[-1/2 * \sigma^2 * \delta((Extract[lista1, j + 1] - Extract[lista1, j] + \mu 1 * \delta)^2)]; j + +]];$

$For[i = 1, i <= k_M[l], jotaimpar = Extract[lista2, 2 * i - 1];$

$f_{X|T}[l] = f_{X|T}[l] * 1/\sqrt{2 * \pi * \sigma^2 * (jotaimpar - Floor[10 * jotaimpar]/10)} *$

$Exp[-1/2 * \sigma^2 * (jotaimpar - Floor[10 * jotaimpar]10)$

$((M1 - Extract[lista1, IntegerPart[10 * jotaimpar]]$

$- \mu 1 * (jotaimpar - Floor[10 * jotaimpar]/10))^2)] *$

$* 1/\sqrt{2 * \pi * \sigma^2 * (Ceiling[10 * jotaimpar]/10 - jotaimpar)} *$

$Exp[-1/2 * \sigma^2 * (Ceiling[10 * jotaimpar]/10 - jotaimpar)$

$((Extract[lista1, IntegerPart[10 * jotaimpar] + 1] - M1$

$+ \mu 1 * (Ceiling[10 * jotaimpar]/10 - \ jotaimpar))^2)]; i + +];$

$For[i = 1, i <= k_m[l], jotapar = Extract[lista2, 2 * i];$

$f_{X|T}[l] = f_{X|T}[l] * 1/\sqrt{2 * \pi * \sigma^2 * (jotapar - Floor[10 * jotapar]/10)} *$

$Exp[-1/2 * \sigma^2 * (jotapar - Floor[10 * jotapar]10)$

$((m1 - Extract[lista1, IntegerPart[10 * jotapar]]$

$+ \mu 1 * (jotapar - Floor[10 * jotapar]/10))^2)]$

$* 1/\sqrt{2 * \pi * \sigma^2 * (Ceiling[10 * jotapar]/10 - jotapar)} *$

$Exp[-1/2 * \sigma^2 * (Ceiling[10 * jotapar]/10 - jotapar)$

$((Extract[lista1, IntegerPart[10 * jotapar] + 1] - m1$

$- \mu 1 * (Ceiling[10 * jotapar]/10 - \ jotapar))^2)]; i + +];$

$Peso[l] = f_{X|T}[l] * w[l]; a1[l] = c * \delta + Last[lista2];$

$$U_m[l] = \sum_{i=1}^{k_m[l]} \delta/((Extract[lista2, 2 * i] - Floor[10 * Extract[lista2, 2 * i]]/10) *$$

$(Ceiling[10 * Extract[lista2, 2 * i]]/10 - Extract[lista2, 2 * i]));$

$$V_m[l] = \sum_{i=1}^{k_m[l]} (Extract[lista1, IntegerPart[10 * Extract[lista2, 2 * i]]] *$$

$(Extract[lista2, 2 * i] - Floor[10 * Extract[lista2, 2 * i]]/10)$

$+ Extract[lista1, IntegerPart[10 * Extract[lista2, 2 * i]] - 1]*$

$(Ceiling[10 * Extract[lista2, 2 * i]]/10 - Extract[lista2, 2 * i]))/((Extract[lista2, 2 * i]$

$- Floor[10 * Extract[lista2, 2 * i]]/10)*$

$(Ceiling[10 * Extract[lista2, 2 * i]]/10 - Extract[lista2, 2 * i]));$

$$U_M[l] = \sum_{i=1}^{k_M[l]} \delta/((Extract[lista2, 2 * i - 1] - Floor[10 * Extract[lista2, 2 * i - 1]]/10)*$$

$(Ceiling[10 * Extract[lista2, 2 * i - 1]]/10 - Extract[lista2, 2 * i - 1]));$

$$V_M[l] = \sum_{i=1}^{k_M[l]} (Extract[lista1, IntegerPart[10 * Extract[lista2, 2 * i - 1]]]*$$

$(Extract[lista2, 2 * i - 1] - Floor[10 * Extract[lista2, 2 * i - 1]]/10)$

$+ Extract[lista1, IntegerPart[10 * Extract[lista2, 2 * i - 1]] - 1]*$

$(Ceiling[10 * Extract[lista2, 2 * i - 1]]/10$

$- Extract[lista2, 2 * i - 1]))/((Extract[lista2, 2 * i - 1]$

$- Floor[10 * Extract[lista2, 2 * i - 1]]/10)*$

$(Ceiling[10 * Extract[lista2, 2 * i - 1]]/10 - Extract[lista2, 2 * i - 1]));$

$$W[l] = \sum_{i=2}^{k[l]} -11/(Extract[lista2, i + 1] - Extract[lista2, i]) + 1/(Extract[lista2, 1]);$$

$T1[l] = (Extract[lista2, 1]); l + +]; \theta p = ReplaceAll[\{\mu, m, M\},$

$$FindRoot[\{(\sum_{j=1}^{L} (a1[j] * Peso[j])) * \mu - \sum_{j=1}^{L} (((2 * k[j] - 2) * (M - m)$$

$+ 2 * (M - x0) + (M - Last[lista1]) * Mod[k[j], 2]+$

$(Last[lista1] - m) * Mod[k[j] + 1, 2]) * Peso[j]) == 0,$

$$(\sum_{j=1}^{L} ((U_m[j] + W[j]) * Peso[j])) * m^2 + (\sum_{j=1}^{L} (((-U_m[j] - 2 * W[j]) * M$$

$+ \mu * (2 * k_m[j] + k[j] - 1) - V_m[j]) * Peso[j])) * m$

$$+ \sum_{j=1}^{L} (((V_m[j] - \mu * (2 * k_m[j] + k[j] - 1)) * M$$

$- \sigma^2 * (k[j] - 1) + W[j] * M^2) * Peso[j]) == 0,$

$$(\sum_{j=1}^{L}((-U_M[j] - W[j] - 1/T1[j]) * Peso[j])) * M^3$$

$$+ (\sum_{j=1}^{L}(((x0 + m) * U_M[j] + W[j] * (x0 + 2 * m) + \mu * (2 * k_M[j] + k[j])$$

$$+ V_M[j] + 2 * x0 + m/T1[j]) * Peso[j])) * M^2$$

$$+ (\sum_{j=1}^{L}((-x0 * m * U_M[j] - (x0 + m) * (V_M[j] + \mu * (2 * k_M[j] + k[j]))$$

$$+ \sigma^2 * k[j] - x0^2 + 2 * m * x0/T1[j] - W[j] * (m^2 + 2 * x0 * m)) * Peso[j])) * M$$

$$+ \sum_{j=1}^{L}((x0 * m * (V_M[j] + \mu * (2 * k_M[j] + k[j]) + x0/T1[j] + W[j] * m)$$

$$- \sigma^2 * (m - (k[j] - 1) * x0)) * Peso[j]) == 0\},$$

$$\{\mu, \mu 1\}, \{m, m1\}, \{M, M1\}, MaxIterations- > 30]];$$

$$Print["\theta", p, " = ", \theta p]; \mu 1 = Extract[\theta p, 1]; m1 = Extract[\theta p, 2]; M1 = Extract[\theta p, 3];$$

$$PutAppend[\{p, \mu 1, m1, M1\}, "estimadores"]; p + +];$$

Appendix 3: Math. 4.1 instructions to compute the BMD least squares estimators

$lista = ReadList["C : \backslash DocumentsandSettings \backslash PictetEuropeEquitities.txt"];$

$lista1 = Table[\{i, Extract[lista, i]/100\}, \{i, 1, 210\}];$

$x0 = Extract[lista1, 1]; Xt = x0;$

$ListPlot[lista1, PlotJoined- > True]; d = Length[lista1]; lista4 = \{\};$

$For[l = 114, l <= 118, For[p = 118, p <= 124, m1 = l; M1 = p;$

$lista3 = \{\}; S11 = 0; S12 = 0; S21 = 0; S22 = 0; R = 1;$

$For[i = 1, i < d, If[And[R == 1,$

$Extract[lista1, \{i, 2\}] < M1 <= Extract[lista1, \{i + 1, 2\}]],$

$R = 2; lista3 = Insert[lista3, Extract[lista1, \{i + 1, 1\}], -1],$

$If[And[R == 2, Extract[lista1, \{i, 2\}] > m1 >= Extract[lista1, \{i + 1, 2\}]], R = 1;$

$lista3 = Insert[lista3, Extract[lista1, \{i + 1, 1\}], -1]]]; i + +];$

$k_{m,M} = Length[lista3]; X_T = Extract[Last[lista1], 2]; lista3 = Insert[lista3, 1, 1];$

$lista3 = Insert[lista3, d - 2, -1]; For[i = 1, i <= k_{m,M},$

$For[j = Extract[lista3, i], j < Extract[lista3, i + 1],$

$S11 = S11 + Extract[lista1, \{j + 1, 2\}] - Extract[lista1, \{j, 2\}];$

$S12 = S12 + 1; j = j + 1]; i = i + 2]; If[S12 > 0, \hat{\mu}_T[1] = S11/S12, \hat{\mu}_T[1] = 0];$

$For[i = 1, i <= k_{m,M}, For[j = Extract[lista3, i + 1], j < Extract[lista3, i + 2],$

$S21 = S21 + Extract[lista1, \{j + 1, 2\}] - Extract[lista1, \{j, 2\}];$

$S22 = S22 + 1; j = j + 1]; i = i + 2]; If[S22 > 0, \hat{\mu}_T[2] = S21/S22, \hat{\mu}_T[2] = 0];$

$G_n = 0; For[i = 0, i <= k_{m,M}, For[j = Extract[lista3, i + 1], j < Extract[lista3, i + 2],$

$G_n = G_n + (Extract[lista1, \{j + 1, 2\}] - Extract[lista1, \{j,\ 2\}] - \hat{\mu}_T[Mod[i, 2] + 1])^2;$

$j = j + 1]; i + +];$

$lista4 = Insert[lista4, \{G_n, m1, M1, \hat{\mu}_T[1], \hat{\mu}_T[2], k_{m,M}\}, -1];$

$p = p + .1]; l = l + .1]; lista4 = Sort[lista4];$

$Print["LSE = ", Extract[lista4, \{1, 1\}]]; Print["m0 = ", Extract[lista4, \{1, 2\}]];$

$Print["M0 = ", Extract[lista4, \{1, 3\}]]; Print["\mu_1 = ", Extract[lista4, \{1, 4\}]];$

$Print["\mu_2 = ", Extract[lista4, \{1, 5\}]];$

Appendix 4: Math. 4.1 instructions to compute the GBM least squares estimators

$lista = ReadList["C:\backslash Documents and Settings \backslash PictetEuropeEquitities.txt"];$

$lista1 = Table[\{i, Extract[lista, i]/100\}, \{i, 1, 210\}];$

$x0 = Extract[lista1, 1]; Xt = x0; ListPlot[lista1, PlotJoined- > True];$

$d = Length[lista1]; lista4 = \{\};$

$For[l = 114, l <= 118, For[p = 118, p <= 124, m1 = l; M1 = p; lista3 = \{\}; lista5 = \{\};$

$S11 = 0; S12 = 0; S13 = 0; S21 = 0; S22 = 0; S23 = 0; R = 1;$

$For[i = 1, i < d, If[And[R == 1,$

$Extract[lista1, \{i, 2\}] < M1 <= Extract[lista1, \{i + 1, 2\}]],$

$R = 2; lista3 = Insert[lista3, Extract[lista1, \{i + 1, 1\}], -1],$

$If[And[R == 2, Extract[lista1, \{i, 2\}] > m1 >= Extract[lista1, \{i + 1, 2\}]], R = 1;$

$lista3 = Insert[lista3, Extract[lista1, \{i + 1, 1\}], -1]]]; i + +];$

$k_{m,M} = Length[lista3]; lista3 = Insert[lista3, 1, 1];$

$lista3 = Insert[lista3, d - 2, -1]; For[i = 1, i <= k_{m,M},$

$For[j = Extract[lista3, i], j < Extract[lista3, i + 1],$

$S11 = S11 + Log[Extract[lista1, \{j + 1, 2\}]/Extract[lista1, \{j, 2\}]];$

$S12 = S12 + Log[Extract[lista1, \{j + 1, 2\}]/Extract[lista1, \{j, 2\}]]^2;$

$S13 = S13 + 1; j = j + 1]; i = i + 2];$

$\beta_N[1] = S12/S13 - 1 - S11^2/S13 * (S13 - 1); \alpha_N[1] = S11/S13 + \beta_N[1]/2;$

$For[i = 1, i <= k_{m,M}, For[j = Extract[lista3, i + 1], j < Extract[lista3, i + 2],$

$S21 = S21 + Log[Extract[lista1, \{j + 1, 2\}]/Extract[lista1, \{j, 2\}]];$

$S22 = S22 + Log[Extract[lista1, \{j + 1, 2\}]/Extract[lista1, \{j, 2\}]]^2;$

$S23 = S23 + 1; j = j + 1]; i = i + 2];$

$\beta_N[2] = S22/S23 - 1 - S21^2/S23 * (S23 - 1); \alpha_N[2] = S21/S23 + \beta_N[2]/2;$

$G_n = 0; For[i = 0, i <= k_{m,M},$

$For[j = Extract[lista3, i + 1], j < Extract[lista3, i + 2],$

$G_n = G_n + (Extract[lista1, \{j + 1, 2\}] - Extract[lista1, \{j,\ 2\}] * Exp[\alpha_N[Mod[i, 2] + 1]])^2;$

$j = j + 1]; i + +];$

$lista4 = Insert[lista4, \{G_n, m1, M1, \alpha_N[1], \alpha_N[2], k_{m,M}\}, -1]; p = p + .1];$

$l = l + .1]; lista4 = Sort[lista4];$

$Print["LSE = ", Extract[lista4, \{1, 1\}]]; Print["m0 = ", Extract[lista4, \{1, 2\}]];$

$Print["M0 = ", Extract[lista4, \{1, 3\}]]; Print["\alpha_T[1] = ", Extract[lista4, \{1, 4\}]];$

$Print["\alpha_T[2] = ", Extract[lista4, \{1, 5\}]];$

Appendix 5: Math. 4.1 instructions to compute the OU least squares estimators

$lista = ReadList["C : \backslash DocumentsandSettings\backslash PictetEuropeEquitities.txt"];$

$lista1 = Table[\{i, Extract[lista, i]/100\}, \{i, 1, 210\}]; x0 = Extract[lista1, 1]; Xt = x0;$

$ListPlot[lista1, PlotJoined- > True]; d = Length[lista1]; lista4 = \{\};$

$For[l = 114, l <= 118, For[p = 118, p <= 124, m1 = l; M1 = p; lista3 = \{\}; lista5 = \{\};$

$S11 = 0; S12 = 0; S21 = 0; S22 = 0; R = 1;$

$For[i = 1, i < d, If[And[R == 1,$

$Extract[lista1, \{i, 2\}] < M1 <= Extract[lista1, \{i + 1, 2\}]],$

$R = 2; lista3 = Insert[lista3, Extract[lista1, \{i + 1, 1\}], -1],$

$If[And[R == 2, Extract[lista1, \{i, 2\}] > m1 >= Extract[lista1, \{i + 1, 2\}]], R = 1;$

$lista3 = Insert[lista3, Extract[lista1, \{i + 1, 1\}], -1]]]; i + +];$

$k_{m,M} = Length[lista3]; X_T = Extract[Last[lista1], 2];$

$lista3 = Insert[lista3, 1, 1]; lista3 = Insert[lista3, d - 2, -1];$

$For[i = 1, i <= k_{m,M}, For[j = Extract[lista3, i], j < Extract[lista3, i + 1],$

$S11 = S11 + Extract[lista1, \{j, 2\}] * Extract[lista1, \{j + 1, 2\}];$

$S12 = S12 + Extract[lista1, \{j, 2\}]^2; j = j + 1]; i = i + 2];$

$If[S12 > 0, \alpha_T[1] = Log[S11/S12], \alpha_T[1] = 0]; For[i = 1, i <= k_{m,M},$

$For[j = Extract[lista3, i + 1], j < Extract[lista3, i + 2],$

$S21 = S21 + Extract[lista1, \{j, 2\}] * Extract[lista1, \{j + 1, 2\}];$

$S22 = S22 + Extract[lista1, \{j, 2\}]^2; j = j + 1]; i = i + 2];$

$If[S22 > 0, \alpha_T[2] = Log[S21/S22], \alpha_T[2] = 0]; G_n = 0;$

$For[i = 0, i <= k_{m,M}, For[j = Extract[lista3, i + 1], j < Extract[lista3, i + 2],$

$G_n = G_n+(Extract[lista1, \{j + 1, 2\}]-Extract[lista1, \{j,\ 2\}]*Exp[\alpha_T[Mod[i, 2]+1]])^2;$

$j = j + 1]; i + +]; lista4 = Insert[lista4, \{G_n, m1, M1, \alpha_T[1], \alpha_T[2], k_{m,M}\}, -1];$

$p = p + .1]; l = l + .1]; lista4 = Sort[lista4];$

$Print["LSE = ", Extract[lista4, \{1, 1\}]]; Print["m0 = ", Extract[lista4, \{1, 2\}]];$

$Print["M0 = ", Extract[lista4, \{1, 3\}]]; Print["\alpha_T[1] = ", Extract[lista4, \{1, 4\}]];$

$Print["\alpha_T[2] = ", Extract[lista4, \{1, 5\}]];$

Bibliography

[Adams, 1996] Malcolm Adams & Victor Guillemin. Measure theory and probability. Birkhäuser.

[Bec, Guay & Guerre, 2008] Frédérique Bec, Alain Guay & Emmanuel Guerre. Adaptative consistent unit root tests based on autoregressive threshold model. Journal of Econometrics Volume 142, Issue 1, Pages 94-133.

[Caner & Hansen, 2001] Mehmet Caner & Bruce E. Hansen. Threshold autoregression with a unit root. Econometrica 69(6), 1555-1596.

[Chan, 1993] K.S. Chan. Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. The Annals of Statistics 21(1), 520-533.

[Chan & Tsay, 1998] K.S. Chan & Ruey S. Tsay. Limiting properties of the least squares estimator of a continuous threshold autoregressive model. Biometrika 85(2), 413-426.

[Freidlin & Pfeiffer, 1998] M. Freidlin & Ruth Pfeiffer. A threshold estimation problem for processes with hysteresis. Statistics & Probability Letters 36, 337-347.

[Genon-Catalot & Jacod, 1993] V. Genon-Catalot & J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. Ann. Inst. Henry Poincaré (Proba. Statist.) 29, 119-151.

[Gonzalo & Montesinos, 2004] Jesús Gonzalo & Raquel Montesinos. Threshold stochastic unit root models. Unpublished manuscript, Department of Statistics and Econometrics, Universidad Carlos III de Madrid.

[Gonzalo & Pitarakis, 2002] Jesús Gonzalo & Jean-Yves Pitarakis. Estimation and model selection based inference in single and multiple threshold models. Journal of Econometrics 110(2), 319-352.

[Gonzalo & Wolf, 2005] Jesús Gonzalo & Michael Wolf. Subsampling inference in threshold autoregressive models. Journal of Econometrics 127(2), 201-224.

[Hansen, 1996] Bruce E. Hansen. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64(2), 413-430.

[Hansen, 1997] Bruce E. Hansen. Inference in TAR models. Studies in Nonlinear Dynamics and Econometrics 2(1), 1-14.

[Hansen, 1999] Bruce E. Hansen. Testing for linearity. Journal of Economic Surveys 13(5), 551-576.

[Hansen, 2000] Bruce E. Hansen. Sample splitting and threshold estimation. Econometrica 68(3), 575-603.

[Hansen & Seo, 2002] Bruce E. Hansen & Byeongseon Seo. Testing for two-regime threshold cointegration in vector error-correction models. Journal of Econometrics 110, 293-318.

[Hida, 1980] T. Hida, Brownian Motion. Springer-Verlag.

[Ip, Wong, Li & Xie, 1999] Way-Cheung Ip, Heung Wong, Yuan Li & Zhongjie Xie. Threshold variable selection by wavelets in open-loop threshold autoregressive models. Statistics & Probability Letters 42, 375-392.

[Kapetanios, 2001] George Kapetanios. Model selection in threshold models. Journal of Time Series Analisys 22(6), 733-754.

[Kapetanios, 2003] George Kapetanios. Using extraneous information and GMM to estimate threshold parameters in TAR models. Work in Progress.

[Karatzas & Shreve, 1991] I. Karatzas & S. Shreve, Brownian Motion and Stochastic Calculus-second edition. Springer.

[Li & Xie, 1999] Yuan Li & Zhongjie Xie. The wavelet identification of thresholds and time delay of threshold autoregressive models. Statistica Sinica 9, 153-166.

[Molina & et al., 2004] A. Molina, A. Gabaldón, M. Kessler, J.A. Fuentes, E. Gómez. Application of smoothing techniques to solve the cooling and heating residential load aggregation problem. The International Journal for Computation and Mathematics in Electrical and Electronic Engineering; Volume: 23 Issue: 1.

[Petruccelli, 1996] Joseph D. Petruccelli. On the consistency of least squares estimators for a threshold AR(1) model. Journal of Time Series Analysis vol. 7, n-4, 269-278.

[Øksendal, 1998] B. Øksendal, Stochastic Differential Equations-fifth edition. Springer.

[Revuz & York, 1991] D. Revuz & M. York, Continuous Martingales and Brownian Motion-third edition. Springer.

[Stramer & Lin, 2002] Osnat Straner & Yu-Jau Lin. On inference for threshold autoregressive models. Sociedad de Estadística e Investigación Operativa, Test 11(1), 55-71.

[Tong, 1990] H. Tong, Non-linear Time Series - A dynamical system approach. Oxford University Press.

[Tsay, 1998] Ruey S. Tsay. Testing and modeling multivariate threshold models. Journal of the American Statistical Association 93(443), 1188-1202.

[Williams, 1991] David Williams. Probability with martingales. Cambridge University Press.

[Wong & Li, 2000] C. S. Wong & W. K. Li. Testing for double threshold autoregressive conditional heteroscedastic model. Statistica Sinica 10, 173-189.