



**Linnæus University**

Sweden

Degree Project

# CIR Modeling of Interest Rates



*Author:* Zan Miao

*Supervisor:* Roger Pettersson

*Examiner:* Karl-Olof Lindahl

*Date:* 2018-09-24

*Level:* Bachelor

Department Of Mathematics

## **Abstract**

Short-term interest rate models within one-year financing maturity are considered. In this thesis, we mainly study two short-term interest rate models, the Cox-Ingersoll-Ross model (CIR model) and the Vašíček model. The CIR model is evaluated by numerical simulations based on applying the Euler approximation method and an exact algorithm. By using an ordinary least squares method we can find an initial start value for implementation of a numerical estimate of parameters that maximize the likelihood. Similarly applying those methods to the Vašíček model, we compare the two models with empirical data based on three-month money market rates.

# Contents

<b>1 Introduction</b>	<b>3</b>
1.1 Preliminaries	4
1.2 Review of Interest Rate Models	4
1.2.1 The Vašíček Model	4
1.2.2 The Cox-Ingersoll-Ross Model	5
1.2.3 The CKLS Model	7
<b>2 The CIR Model</b>	<b>8</b>
2.1 Numerical Simulation for the CIR Model	8
2.1.1 The Euler Approximation Method	8
2.1.2 The Exact Algorithm	8
2.2 Parameter Estimation for the CIR Model	9
2.2.1 The Ordinary Least Squares	9
2.2.2 Maximum Likelihood Estimation	11
2.3 Main Results	12
2.3.1 Estimation Results	12
<b>3 The Vašíček Model</b>	<b>17</b>
3.1 Numerical Simulation for the Vašíček Model	17
3.2 Parameter Estimation for the Vašíček Model	17
<b>4 Comparison between the CIR Model and the Vašíček Model with Empirical Data</b>	<b>18</b>
4.1 Simulation Results	18
<b>5 Conclusion</b>	<b>20</b>
<b>References</b>	<b>20</b>
<b>A Matlab Codes</b>	<b>21</b>

# 1 Introduction

In financial markets, interest rates play a significant role. Interbank offered rates are of particular importance of which one of the most well-known example is the London Interbank Offered Rate (LIBOR) which is an averaged interest-rate between leading banks in London. Also, other nations have similar interbank offered rates. In the financial credit crunch 2008, many of those interbank loans were frozen which created a global financial uncertainty. This can motivate a study of models and data for those types of interbank offered rates. The traditional theory of interest rates studies qualitative perspectives, and the modern theory of interest rates quantitatively describes the dynamic characteristics of the term structure of interest rates which can be divided into static and dynamic models, [2]. In this thesis, we intend to study two general equilibrium models, the Cox-Ingersoll-Ross model and the Vašíček model, derived from the dynamic form of the term structure of interest rates. They both present the evolution of the interest rates and describe the interest rates movements over time. It is an interesting and important topic about how to estimate the movements of interest rates and predict the changes of future returns. In this thesis, the fitness of these two models with empirical interbank offered rate data is compared by the use of parameter estimates and numerical methods.

We present the Vašíček model, Cox-Ingersoll-Ross model (CIR model) and a more general model called CKLS model but are mainly focused on the CIR model. In chapter 1, we will give some preliminaries and introduce the three short-term interest rate models so the readers could understand the thesis better. In chapter 2, we will show how to numerically simulate the CIR model with the Euler approximation method and an exact algorithm, and estimate the parameters with ordinary least squares method and maximum likelihood estimation. In chapter 3, we will show how to simulate and estimate parameters for the Vašíček model in similar way. In chapter 4, we will compare the simulation results and summarize which model is fitting better with the empirical data we applied. Here we choose to use the Prague Interbank Offered Rate (PRIBOR) 3M data. PRIBOR is the average interest rates at which banks can lend to each other on the Czech interbank financial market and it is often used as a reference data sample. By the end of the thesis, there are references and MATLAB codes for the previous simulations and estimations.

## 1.1 Preliminaries

**Definition 1.** A standard one-dimensional Wiener process is a stochastic process  $\{W(t)\}_{t \geq 0}$  with  $W(0) = 0$  which satisfies the following properties:

- Each increment  $W(t) - W(s)$ ,  $s \leq t$ , is Gaussian with mean zero and variance  $t - s$ .
- The increments  $W(t_2) - W(t_1)$  and  $W(t_4) - W(t_3)$  are independent for any  $0 \leq t_1 < t_2 \leq t_3 < t_4$ .

For a stochastic process  $\{X(t)\}_{t \geq 0}$  such that  $X(t)$  belongs to the  $\sigma$ -algebra generated by the Wiener process up to time  $t$  and  $\int_0^T E[X(s)^2] ds < \infty$ , the stochastic integral  $\int_0^T X(s) dW(s)$  can be defined as

$$\int_0^T X(s) dW(s) = \lim_{n \rightarrow \infty} \sum_{i=0}^n X\left(\frac{i}{n}T\right) \Delta W_i,$$

where  $\Delta W_i = W\left(\frac{i+1}{n}T\right) - W\left(\frac{i}{n}T\right)$ , and the limit is in mean square sense. It means

$$E\left[\left(\sum_{i=0}^n X\left(\frac{i}{n}T\right) \Delta W_i - \int_0^T X(s) dW(s)\right)^2\right] \rightarrow 0,$$

as  $n \rightarrow \infty$ . That  $X(t)$  belongs to the  $\sigma$ -algebra generated by the Wiener process up to time  $t$  means heuristically that all the information of  $X(t)$  is given by  $\{W(s)\}_{s \leq t}$ . The stochastic integral  $\int_0^T X(s) dW(s)$  can be defined under weaker assumptions of  $\{X(t)\}_{t \geq 0}$ .

## 1.2 Review of Interest Rate Models

### 1.2.1 The Vašíček Model

The Vašíček model is a type of short rate model introduced by O. Vašíček in 1977, [3]. It uses a mean-reverting process to describe the evolution of the instantaneous interest rate that follows the stochastic differential equation (SDE)

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t, \quad r(0) = r_0,$$

where  $\{W_t = W(t)\}_{t \geq 0}$  is a standard Wiener process, and  $r_t = r(t)$  is a short-term interest rate. Here the diffusion term  $\sigma$  is an instantaneous volatility

and the drift term  $\alpha(\mu - r_t)$  represents a force pulling the interest towards its long term mean  $\mu$  with a speed parameter  $\alpha$ . The Vašíček SDE is linear with solution

$$r_t = r_s e^{-\alpha(t-s)} + \mu(1 - e^{-\alpha(t-s)}) + \sigma \int_s^t e^{-\alpha(t-u)} dW_u.$$

From this we see that it can be negative which typically is not a desirable interest rate property. The Vašíček interest rates are normal distributed with conditional expectation

$$E[r_t|r_s] = r_s e^{-\alpha(t-s)} + \mu(1 - e^{-\alpha(t-s)}), \quad s < t,$$

and conditional variance

$$\text{Var}[r_t|r_s] = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha(t-s)}), \quad s < t,$$

[3]. By introducing the symbol  $\sim$  meaning 'has the same distribution as', this can be expressed as

$$r_t|r_s \sim r_s e^{-\alpha(t-s)} + \mu(1 - e^{-\alpha(t-s)}) + \frac{\sigma}{\sqrt{2\alpha}} \sqrt{1 - e^{-2\alpha(t-s)}} Z, \quad (1)$$

where  $Z \sim \mathcal{N}(0, 1)$ .

### 1.2.2 The Cox-Ingersoll-Ross Model

The Cox-Ingersoll-Ross model (CIR model) was presented in 1985 by J. C. Cox, J. E. Ingersoll and S. A. Ross as an alternative to the Vašíček model, [2]. The CIR model is described by the SDE

$$dr_t = \alpha(\mu - r_t)dt + \sigma\sqrt{r_t}dW_t, \quad r(0) = r_0, \quad (2)$$

where  $\{W_t = W(t)\}_{t \geq 0}$  is again a standard one-dimensional Wiener process,  $\alpha$  is a mean reversion speed,  $\mu$  is a long-run mean, and  $\sigma$  is a volatility rate. The unique solution to [2] is also known as the CIR process. By [2],

$$r_t = r_s + \int_s^t \alpha(\mu - r_u)du + \sigma \int_s^t \sqrt{r_u}dW_u, \quad s < t.$$

Hence

$$E[r_t|r_s] = r_s + \int_s^t \alpha(\mu - E[r_u|r_s])du, \quad s < t,$$

it means, with  $m_t = E[r_t|r_s]$ ,

$$\frac{d}{dt}m_t = \alpha(\mu - m_t), \quad s < t,$$

with solution

$$E[r_t|r_s] = m_t = r_s e^{-\alpha(t-s)} + \mu(1 - e^{-\alpha(t-s)}), \quad s < t.$$

Similarly, it can be shown that

$$\text{Var}[r_t|r_s] = \frac{r_s \sigma^2}{\alpha} (e^{-\alpha(t-s)} - e^{-2\alpha(t-s)}) + \frac{\mu \sigma^2}{2\alpha} (1 - e^{-\alpha(t-s)})^2,$$

see [4]. Now the conditional distribution of  $r_t$  given  $r_s$  will be discussed. We assume throughout that  $\alpha, \mu, \sigma$  and  $r_0$  are positive and  $2\alpha\mu > \sigma^2$  so that the zero boundary will not be reached, [6]. Then there exists a unique solution of [2] and the solution is positive. Furthermore, the transition density function of  $r_t$  at time  $t$  given  $r_s$  at time  $s$  is

$$p(r_s, s, r_t, t) = c e^{-(u+v)} \left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}), \quad (3)$$

where

$$\begin{aligned} c &= \frac{2\alpha}{\sigma^2(1 - e^{-\alpha\Delta t})}, \\ u &= c r_s e^{-\alpha\Delta t}, \\ v &= c r_t, \\ q &= \frac{2\alpha\mu}{\sigma^2} - 1, \\ \Delta t &= t - s \end{aligned} \quad (4)$$

and  $I_q(\cdot)$  is the modified Bessel function of the first kind of order  $q$  which can be defined as

$$I_q(x) = \sum_{j=0}^{\infty} \left(\frac{x}{2}\right)^{2j+q} \frac{1}{k! \Gamma(j+q+1)}, \quad (5)$$

[1], [2], [6], where  $\Gamma$  is the gamma function.

Let  $z_t = 2cr_t$ . Then the conditional distribution of  $z_t$  at time  $t$  given  $z_s$  at time  $s$  is non-central  $\chi_d^2(2u)$  distributed with  $d = 4\alpha\mu/\sigma^2$  degrees of freedom and non-central parameter  $\lambda = 2u$ , with  $u$  given by [4]. This means that  $z_t$  conditioned on  $z_s$  has the same distribution as

$$\chi_d^2(\lambda) = (Z + \sqrt{\lambda})^2 + \chi_{d-1}^2, \quad (6)$$

where  $\chi_{d-1}^2$  is a  $\chi^2$  random variable with  $d - 1$  degrees of freedom, and  $Z$  is an independent standard  $\mathcal{N}(0, 1)$  random variable. We can write this as

$$z_t|z_s \sim \chi_d^2(\lambda),$$

where

$$\lambda = 2u = 2cr_s e^{-\alpha\Delta t} = 2 \cdot \frac{2\alpha}{\mu^2(1 - e^{-\alpha\Delta t})} e^{-\alpha\Delta t} r_s,$$

where we recall that  $\Delta t = t - s$  from (4). Since  $z_t = 2cr_t$ ,  $r_t$  conditioned on  $r_s$  has the same distribution as  $z_t/2c$  conditioned on  $z_s/2c$ . In formulae,

$$r_t|r_s \sim \frac{z_t}{2c} | \frac{z_s}{2c} \sim \frac{1}{2c} \chi_d^2(\lambda). \quad (7)$$

In words, it means that  $r_t$  conditioned on  $r_s$  has the same distribution as  $\chi_d^2(\lambda)$  divided by  $2c$ .

### 1.2.3 The CKLS Model

K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders presented a more general short-term interest rate model called the CKLS model which includes both the Vaříček and the CIR model, [5]. The CKLS process is structured by the SDE

$$dr_t = \alpha(\mu - r_t)dt + \sigma r_t^\gamma dW_t, \quad (8)$$

where  $\gamma$  is an additional non-negative parameter. Some parameter examples for the CKLS model are given in Table 1:

Table 1: Some CKLS interest models.

Model	SDE
Vaříček	$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t$
Brennan-Schwartz	$dr_t = \alpha(\mu - r_t)dt + \sigma r_t dW_t$
CIR Square Root (SR)	$dr_t = \alpha(\mu - r_t)dt + \sigma \sqrt{r_t} dW_t$
CIR Variable-Rate (VR)	$dr_t = \sigma r_t^{\frac{3}{2}} dW_t$

In this thesis, the focus is mainly on the CIR model but with some comparisons to the Vaříček model. We first introduce simulations for the CIR model in order to later on compare empirical interest rates with CIR modeled interest rates.



## 2 The CIR Model

### 2.1 Numerical Simulation for the CIR Model

In this section, a simple Euler scheme and an exact algorithm are presented as simulation methods for the CIR model.

#### 2.1.1 The Euler Approximation Method

One of the simplest numerical approximations for the CIR model is to apply the Euler scheme, also called the Euler-Maruyama approximation as

$$\hat{r}(t_{i+1}) = \hat{r}(t_i) + \alpha(\mu - \hat{r}(t_i))\Delta t + \sigma\sqrt{|\hat{r}(t_i)|}\sqrt{\Delta t}Z_i, \quad (9)$$

where  $\hat{r}(0) = r(0)$  is given and  $Z_1, Z_2, \dots, Z_{n-1}$  are standard independent  $\mathcal{N}(0, 1)$  Gaussian variables, [4]. Here  $\hat{r}$  denotes a time-discretized approximation defined on a time partition  $t_1 < t_2 < \dots < t_i$ , and  $i = 1, 2, \dots, n-1$ . The  $\sqrt{\Delta t}Z_i$ -factor can be explained by that it has the same distribution as  $\Delta W_i = W(t_{i+1}) - W(t_i)$ . Observe that the Euler approximation of the CIR process [9] is a discrete version of the CIR SDE [2]. Even though the  $r$  given by the CIR model [2] is positive, an Euler trajectory may at some points be negative due to the time discretization error. Therefore the absolute value of  $\hat{r}$  is included in the dispersion term. Other modifications of the Euler method could also be applied such as

$$\hat{r}(t_{i+1}) = \max(\hat{r}(t_i) + \alpha(\mu - \hat{r}(t_i))\Delta t + \sigma\sqrt{\hat{r}(t_i)}\sqrt{\Delta t}Z_i, 0) \quad (10)$$

and

$$\hat{r}(t_{i+1}) = \hat{r}(t_i) + \alpha(\mu - \hat{r}(t_i))\Delta t + \sigma\sqrt{\max(\hat{r}(t_i), 0)}\sqrt{\Delta t}Z_i. \quad (11)$$

In this thesis we for simplicity only apply the Euler version [9]. It would be interesting to compare [9]-[11] but that is beyond the scope of this thesis.

#### 2.1.2 The Exact Algorithm

By [6] and [7], we can simulate  $r_t$  given  $r_s$  exactly. It means that for this scheme, there is no time discretization error. Given  $r_s$ , we don't know the value of  $r_t$  but we know the conditional distribution of  $r_t$  from which we can simulate outcomes. For the Euler method we only know an approximate conditional distribution of a future CIR process value.

## 2.2 Parameter Estimation for the CIR Model

### 2.2.1 The Ordinary Least Squares

There are many methods to estimate the parameters of the CIR model, for example, the ordinary least squares estimation method (OLSE), maximum likelihood estimation method (MLE), method of moments, etc. In this thesis we apply the MLE method which is solved numerically with initial parameter values given by the OLSE method. A version of the OLSE method will here be presented. The idea is to get approximate estimates of the parameters  $\alpha, \mu$  and  $\sigma$ . For that we consider a discretized version of (2), which is similar to the Euler discretization (9),

$$r(t_{i+1}) - r(t_i) = \alpha(\mu - r(t_i))\Delta t + \sigma\sqrt{|r(t_i)|}\Delta W_i,$$

where  $i = 0, 1, \dots, n-1$ ,  $\Delta t = t - s$  and  $\Delta W_i = W(t_{i+1}) - W(t_i)$ . Observe that here  $r(t_1), \dots, r(t_n)$  are empirical data while in (9),  $\hat{r}(t_1), \dots, \hat{r}(t_n)$  are simulated data. We obtain

$$\begin{aligned} \frac{r(t_{i+1}) - r(t_i)}{\sqrt{|r(t_i)|}} &= \frac{\alpha}{\sqrt{|r(t_i)|}}(\mu - r(t_i))\Delta t + \sigma\Delta W_i \\ &= \alpha\mu\frac{\Delta t}{\sqrt{|r(t_i)|}} - \alpha\sqrt{|r(t_i)|}\Delta t + \sigma\Delta W_i. \end{aligned} \tag{12}$$

By letting

$$\begin{aligned} y_i &= \frac{r(t_{i+1}) - r(t_i)}{\sqrt{|r(t_i)|}}, \\ \beta_1 &= \alpha\mu, \\ \beta_2 &= -\alpha, \\ z_{1i} &= \frac{\Delta t}{\sqrt{|r(t_i)|}}, \\ z_{2i} &= \sqrt{|r(t_i)|}\Delta t, \\ \epsilon_i &= \sigma\Delta W_i, \end{aligned} \tag{13}$$

(12) can be written as

$$y_i = \beta_1 z_{1i} + \beta_2 z_{2i} + \epsilon_i,$$

which is equivalent to

$$Y = Z\beta + \epsilon,$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix}, Z = \begin{bmatrix} \frac{\Delta t}{\sqrt{|r_1|}} & \sqrt{|r_1|}\Delta t \\ \frac{\Delta t}{\sqrt{|r_2|}} & \sqrt{|r_2|}\Delta t \\ \vdots & \vdots \\ \frac{\Delta t}{\sqrt{|r_{n-1}|}} & \sqrt{|r_{n-1}|}\Delta t \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \epsilon = \sigma\sqrt{\Delta t} \begin{bmatrix} \mathcal{N}_1(0, 1) \\ \mathcal{N}_2(0, 1) \\ \vdots \\ \mathcal{N}_{n-1}(0, 1) \end{bmatrix}, \quad (14)$$

where  $\mathcal{N}_1(0, 1), \dots, \mathcal{N}_{n-1}(0, 1)$  are independent  $\mathcal{N}(0, 1)$ .

The OLSE of  $\beta$  is then

$$\hat{\beta} = \arg \min_{\beta} \|Y - Z\beta\|^2,$$

where  $\|\cdot\|$  denotes Euclidean distance. It is well known that

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (Z^T Z)^{-1} Z^T Y,$$

[7, p 364]. From (13),  $\alpha$  and  $\mu$  can be estimated as

$$\begin{aligned} \hat{\alpha} &= -\hat{\beta}_2, \\ \hat{\mu} &= \frac{\hat{\beta}_1}{\hat{\alpha}}. \end{aligned}$$

In [1], there are explicit formulae for  $\hat{\alpha}$  and  $\hat{\mu}$  in terms of intricate sums. Those sums are however not needed for this thesis.

By [7, p 371], and (14),

$$\widehat{\sigma^2 \Delta t} = \frac{1}{n} \|Y - Z\hat{\beta}\|^2,$$

i.e.

$$\hat{\sigma} = \frac{1}{\sqrt{\Delta t n}} \|Y - Z\hat{\beta}\|,$$

which is a standardized distance between  $Y$  and the fitted  $Z\hat{\beta}$ . The above estimates of  $\alpha, \mu$  and  $\sigma$  serve as initial estimates for a numerical optimization of the likelihood which will be discussed below.

### 2.2.2 Maximum Likelihood Estimation

When  $r(t_i)$  is given at time  $t_i$ , the transition density function (3) for  $r(t_{i+1})$ , given  $r(t_i)$  is known. The likelihood function describing the joint density for the data  $r(t_1), r(t_2), \dots, r(t_i)$  is

$$L(\theta) = p(r(t_1)) \prod_{i=1}^{n-1} p(r(t_{i+1})|r(t_i), \theta),$$

where  $\theta = (\alpha, \mu, \sigma)$  are the parameters,  $p(r(t_1))$  is the initial density of  $r(t_1)$  for simplicity assumed to not depend on  $\theta$ , and  $p(r(t_{i+1})|r(t_i), \theta) = p(r(t_i), t_i, r(t_{i+1}), t_{i+1})$  given by (3) is the conditional density of  $r(t_{i+1})$  given  $r(t_i)$  which depends on  $\theta$ .<sup>1</sup> The log-likelihood function is

$$\ln L(\theta) = \ln p(r(t_1)) + \sum_{i=1}^{n-1} \ln p(r(t_{i+1})|r(t_i), \theta).$$

Generally, the log-likelihood function is easier to handle than the likelihood function transforming products to sums. For the CIR model, we can by (3) easily get the log-likelihood function of the CIR process as

$$\begin{aligned} \ln L(\theta) = & \ln p(r(t_1)) + (n-1) \ln c + \\ & \sum_{i=1}^{n-1} \left\{ -u(t_i) - v(t_{i+1}) + \frac{q}{2} \frac{v(t_{i+1})}{u(t_i)} + \ln I_q(2\sqrt{u(t_i)v(t_{i+1})}) \right\}, \end{aligned} \quad (15)$$

where  $u(t_i) = cr(t_i)e^{-\alpha\Delta t}$  and  $v(t_{i+1}) = cr(t_{i+1})$ . The MLE method is to maximize the log-likelihood function (15) with respect to  $\theta$ , i.e.

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \ln L(\theta).$$

For optimization of the CIR likelihood function (15) in Matlab, we can use the code `fminsearch`, but there may be some numerical problems since  $I_q(2\sqrt{uv})$  tends quickly to plus infinity as  $uv \rightarrow +\infty$  which by (3) and (4) happens if  $\sigma^2 \rightarrow 0$ . Due to this reason, the code `besseli(q, 2 * sqrt(u.*v))` which calculates  $I_q(2\sqrt{uv})$  may make the MLE method fail in this situation. Instead, let  $I_q^1(2\sqrt{uv}) = I_q(2\sqrt{uv})e^{-2\sqrt{uv}}$ . The exponentially rescaled

---

<sup>1</sup>In statistics, there is often an ambiguity with notations. By  $p(r(t_1))$  means here more precisely the density function of a random variable  $R(t_1)$  evaluated at  $r(t_1)$ . Similarly by  $p(r(t_{i+1})|r(t_i), \theta)$  means the conditional density evaluated at  $r(t_{i+1})$  for a random variable  $R(t_{i+1})$  given that  $R(t_i) = r(t_i)$ , where  $r(t_i)$  is the empirical data at time  $t_i$ . Here  $R(t)$  is the value of the CIR process at time  $t$ .

Besseli function  $I_q^1(2\sqrt{uv})$  can be implemented in Matlab with the code `besseli(q, 2 * sqrt(u.*v), 1)` which reduces divergence problem since though  $I_q(2\sqrt{uv})$  explodes for large  $\sqrt{uv}$ , the factor  $e^{-\sqrt{uv}}$  reduces that phenomena. We therefore rearrange the log-likelihood function (15) into

$$\ln L(\theta) = \ln p(r(t_1)) + (n-1) \ln c + \sum_{i=1}^{n-1} \left\{ -u(t_i) - v(t_{i+1}) + \frac{q}{2} \frac{v(t_{i+1})}{u(t_i)} + \ln I_q^1\left(2\sqrt{u(t_i)v(t_{i+1})}\right) + 2\sqrt{u(t_i)v(t_{i+1})} \right\}.$$

Note that it is not elementary to algebraically deduce whether  $\ln L$  is a strictly concave function of  $\theta = (\alpha, \mu, \sigma)$ . However, for the data which is studied in this thesis it seems that  $\ln L$  is indeed a strictly concave function of  $\theta$ .

## 2.3 Main Results

### 2.3.1 Estimation Results

In this thesis, we will estimate CIR modeled parameters with the empirical data sample PRIBOR (Prague interbank offered rate) for the short-term 3 months (3M) by applying the MLE method. A standard convention is that one year is 250 working days. We set the time difference  $\Delta t = \frac{1}{250}$  for the time series of 3589 working-day observations of PRIBOR 3M shown in Figure 1.

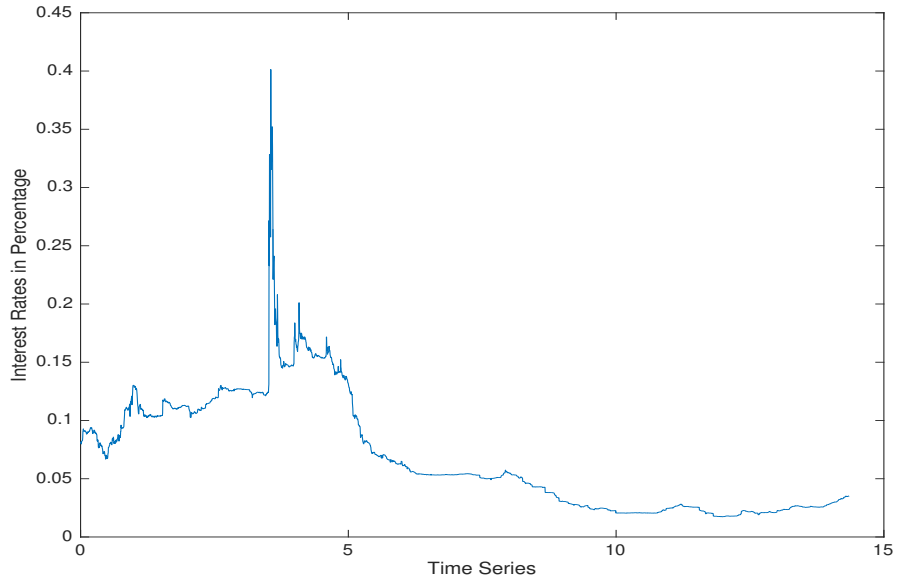


Figure 1: PRIBOR 3M.

Before applying the MLE method, we need to find the estimated initial parameter values by using the OLSE method. The initial values given by the OLSE method which is shown in the following table are quite close to the optimal values given by MLE method.

Table 2: Estimation Results.

Methods	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\sigma}$
OLSE (initial parametric values)	0.1209	0.0423	0.1642
MLE (optimal parameters)	0.4363	0.0613	0.1491

Figure 2 illustrates  $\ln L$  as a function of  $\alpha$  given optimal parameters  $\mu$  and  $\sigma$ . From the graph, it is indicated that  $\ln L$  is a strictly concave function of  $\alpha$  giving a unique optimal  $\alpha$ -value.

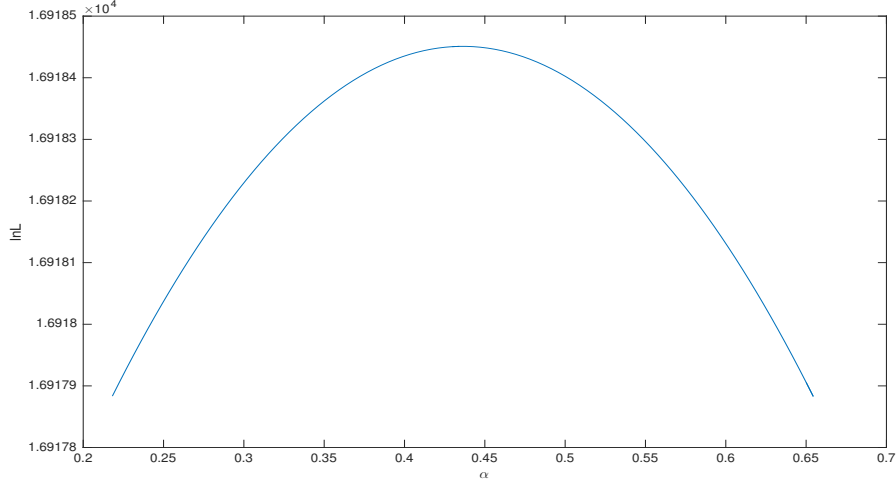


Figure 2: The log-likelihood  $\ln L$  as a function of  $\alpha$  given optimal  $\mu$  and  $\sigma$ .

Similarly, with Figure 3, given optimal  $\alpha$  and  $\sigma$ ,  $\ln L$  is plotted as a log-likelihood function of  $\mu$ . Visually,  $\ln L$  is also indicated to be a strictly concave function of  $\mu$  giving a unique optimal  $\mu$ -value.

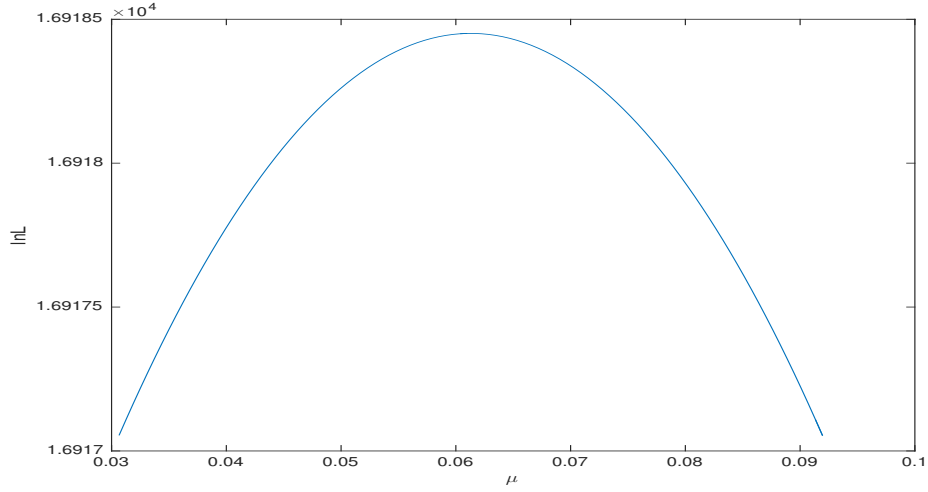


Figure 3: The log-likelihood  $\ln L$  as a function of  $\mu$  given optimal  $\alpha$  and  $\sigma$ .

For Figure 4, optimal values of  $\alpha$  and  $\mu$  are given and  $\ln L$  is plotted as a function of  $\sigma$ . From (3) and (4), it is seen that numerical problems may

appear for small  $\sigma$ -values. Therefore Figure 4 is not symmetric around the optimal  $\sigma$ .

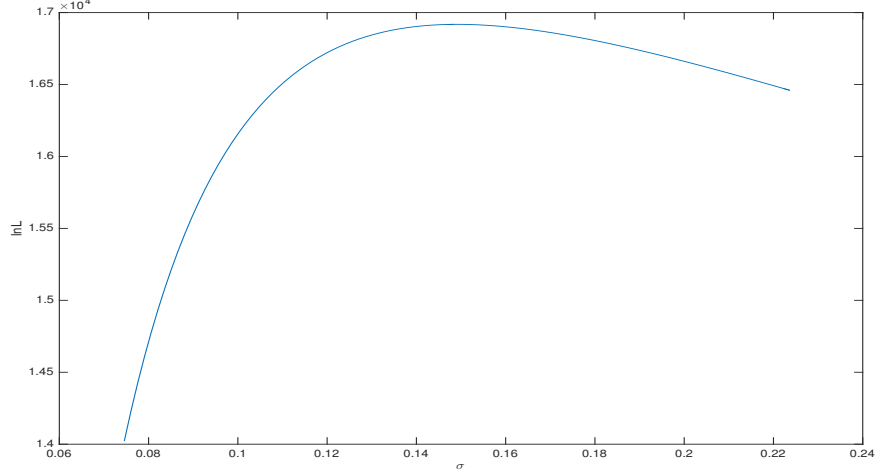


Figure 4: The log-likelihood  $\ln L$  as a function of  $\sigma$  given optimal  $\alpha$  and  $\mu$ .

In Figure 5, given optimal  $\sigma$ ,  $\ln L$  is plotted as a function of  $\alpha$  and  $\mu$ . In Figure 6, given optimal  $\alpha$ ,  $\ln L$  is plotted as a function of  $\sigma$  and  $\mu$ . In Figure 7, given optimal  $\mu$ ,  $\ln L$  is plotted as a function of  $\alpha$  and  $\sigma$ . In [1], the author observed that the log-likelihood functions are quite flat given two drift parameters and it is hard to determine the optimal parametric values. But in this thesis, given two parameters, the log-likelihood appears to be clearly strict concave in Figures 5-7 by changing the ranges of coordinate axes. Also by using the Matlab code `surf` when plotting Figures 5-7, we can easily observe the optimal values according to the level curves.



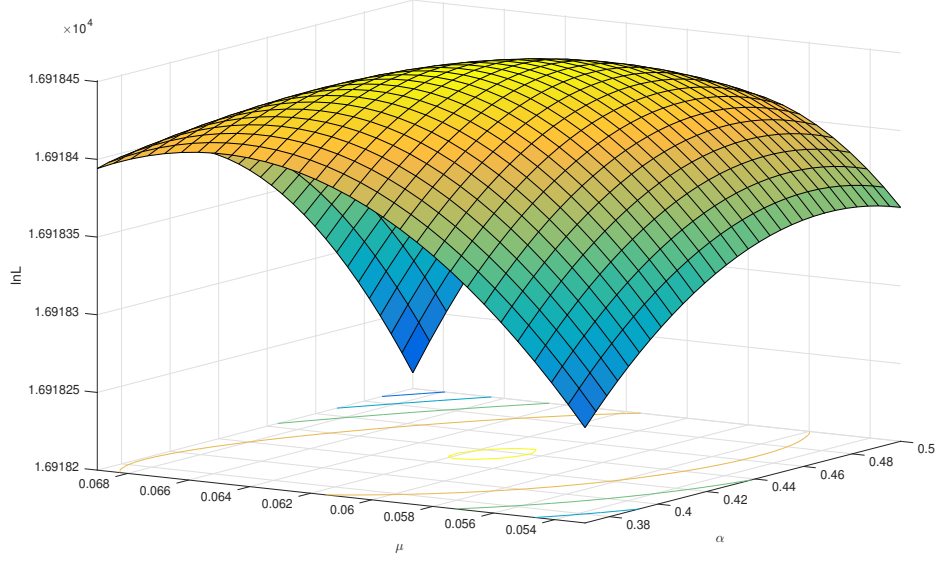


Figure 5: The log-likelihood  $\ln L$  as a function of  $\alpha$  and  $\mu$  given optimal  $\sigma$ .

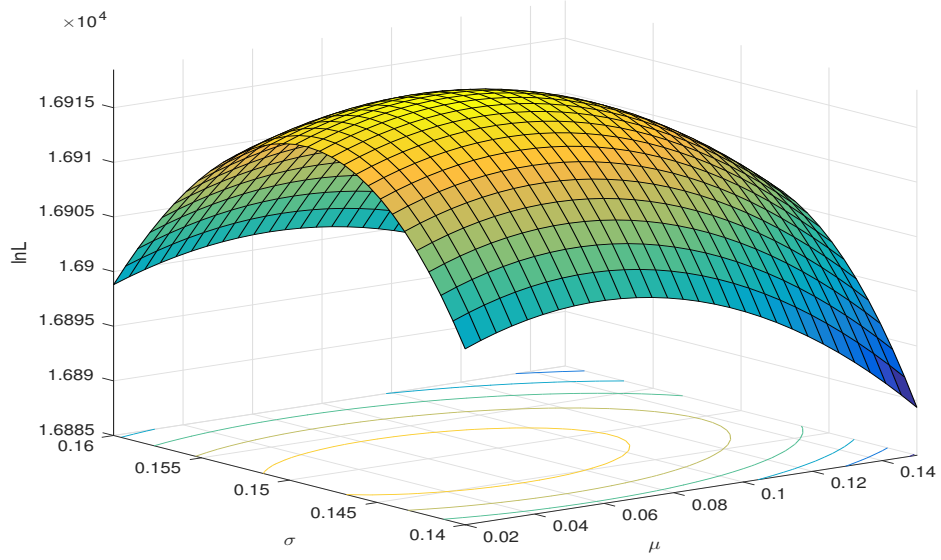


Figure 6: The log-likelihood  $\ln L$  as a function of  $\mu$  and  $\sigma$  given optimal  $\alpha$ .

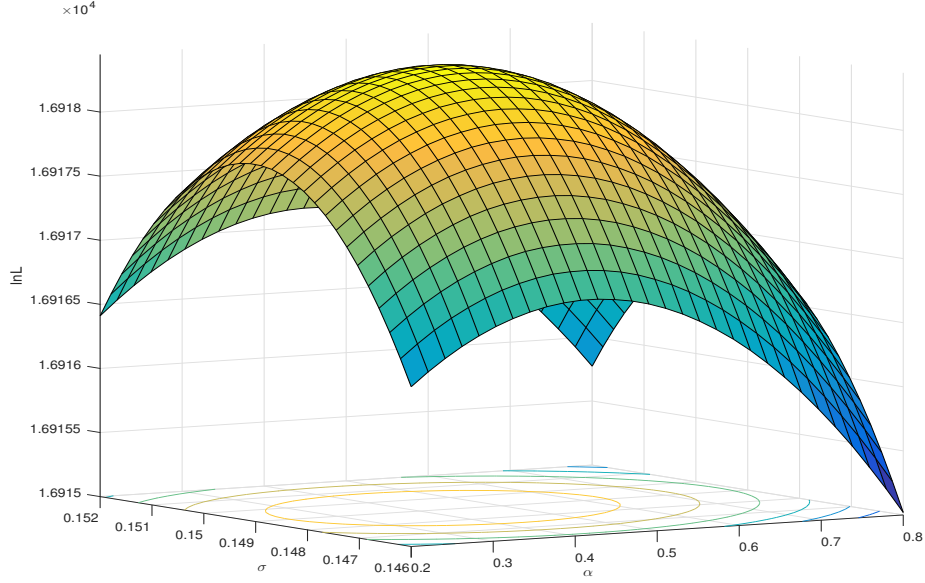


Figure 7: The log-likelihood  $\ln L$  as a function of  $\alpha$  and  $\sigma$  given optimal  $\alpha$ .

### 3 The Vašíček Model

#### 3.1 Numerical Simulation for the Vašíček Model

We choose to apply the exact algorithm to the Vašíček Model by using

$$r(t_{i+1}) = r(t_i)e^{(-\alpha\Delta t)} + \mu(1 - e^{(-\alpha\Delta t)}) + \sigma\sqrt{\frac{1 - e^{-2\alpha\Delta t}}{2\alpha}}Z_i, \quad (16)$$

where  $\Delta t = r(t_{i+1}) - r(t_i)$  and  $Z_i$  are standard independent  $\mathcal{N}(0, 1)$  Gaussian variables with  $i = 1, 2, \dots, n - 1$ . The algorithm (16) is explained by (1).

#### 3.2 Parameter Estimation for the Vašíček Model

We applied the same parameter estimation methods as we did to the CIR model for the same PRIBOR 3M data. The estimates for the Vašíček model are shown in the following table:

Table 3: Estimation Results.

Methods	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\sigma}$
OLSE (initial parametric values)	1.1086	0.0657	0.0738
MLE (optimal parameters)	1.1110	0.0657	0.0739

Note that for the Vašíček model, the OLSE results and the MLE results are almost identical but not exactly equal. It is beyond the scope of this thesis to explain this phenomena.

## 4 Comparison between the CIR Model and the Vašíček Model with Empirical Data

### 4.1 Simulation Results

At Figure 8, the PRIBOR 3M data are plotted together with simulated CIR trajectories with parameters given by the MLE.

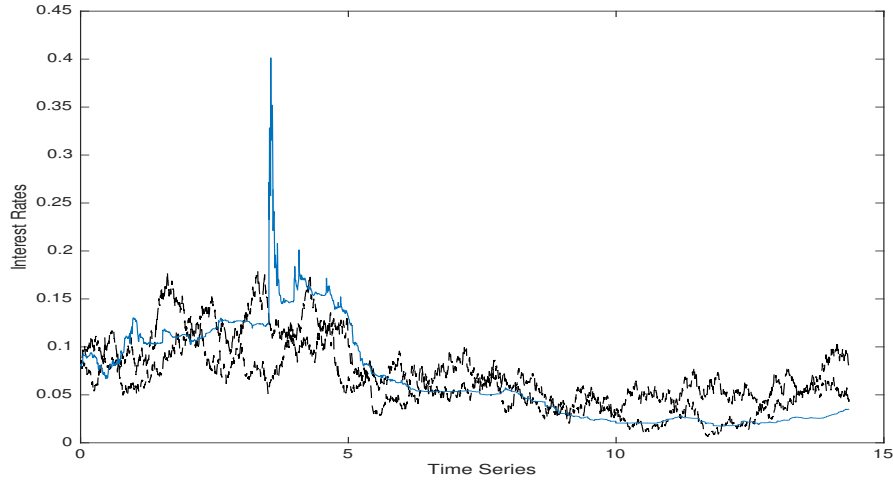


Figure 8: Empirical PRIBOR 3M data and simulated CIR data by using the exact algorithm.

Note that the CIR simulated data share the properties of the PRIBOR 3M data by non-negative interest rates and that the fluctuation of data is larger for larger interest rates. Perhaps an even better model than the CIR model would be the CKLS model (8) with  $\gamma > 1/2$ , since then the fluctuations would be even larger for large interest rates and smaller for small interest rates. However, if  $\gamma > 1/2$ , there is, at least to the knowledge of the author, no general transition density formula. In situations for which there is no explicit transition density, there exists difficulties to estimate the parameters as well simulate the process. That can to some extent be resolved by estimating parameters by a generalised method of moments, [8].

At Figure 9, the PRIBOR 3M data are plotted together with simulated Vašíček trajectories with parameters given by the MLE.

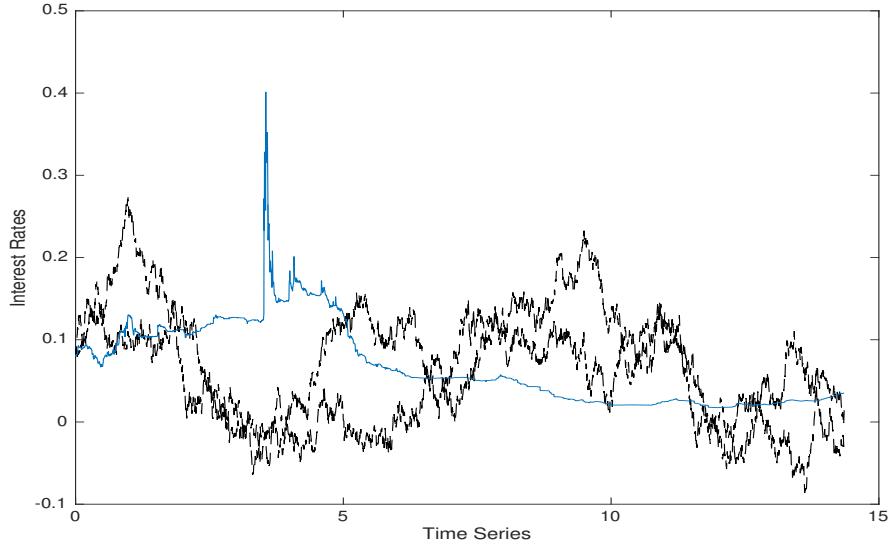


Figure 9: Empirical PRIBOR 3M data and the simulated Vašíček data by using the exact algorithm.

For the Vašíček model, the interest rates can be negative and the fluctuations of interest rate does not depend on the magnitude of the interest rates. From Figures 8 and 9 and the discussion above we see that the CIR model is more suitable than the Vašíček model with this PRIBOR 3M data sample.

## 5 Conclusion

We basically studied and evaluated the CIR model and the Vašíček model for empirical interest rate data. We wanted to find out which of the models was the best fitting with the empirical data. The empirical interest rate data were non-negative and fluctuated more for larger interest rates than for interest rates close to zero. That behaviour was captured by the CIR model but not by the Vašíček model. Hence the CIR model seemed to be a more appropriate interest rate model than the Vašíček model at least what concerned interest rate data of the type studied in this thesis. It was also observed that a CKLS model (8) with  $\gamma > 1/2$  might fit empirical interest rates even more, but then it might be harder to estimate the parameters due to the lack of explicit transition densities. Then other methods for parameter estimations are needed, see for instance 8.

## References

- [1] Kamil Kladvko. (2007). Maximum Likelihood Estimation of the Cox-Ingersoll-Ross Process: The Matlab Implementation. Working paper, Technical Computing Prague.
- [2] John C. Cox, Jonathan E. Ingersoll, Jr., Stephen A. Ross. (March, 1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, Vol. 53, No. 2, pp. 385-407.
- [3] Oldrich Vasicek. (November, 1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, Vol. 5, Issue 2, pp. 177-188.
- [4] Leif B.G. Andersen, Peter Jackel, Christian Kähl. (May, 2010). Simulation of Square-Root Processes. <https://doi.org/10.1002/9780470061602.eqf13009>.
- [5] K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders. (July, 1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *The Journal of Finance*, VOL. XLVII, NO. 3.
- [6] Stefano M. Iacus. (2008). *Simulation and Inference for Stochastic Differential Equations*. Springer-Verlag New York.
- [7] Richard A. Johnson, Dean W. Wichern. (2008). *Applied Multivariate Statistical Analysis*, 6th Edition. Pearson.

- [8] Fredrik Niveman. (April, 2010). Parameter Estimation in Discretely Sampled Diffusions Experiencing Volatility Induced Stationarity, Master Thesis, Chalmers University of Technology, University of Gothenburg.

## A Matlab Codes

```

1 % The transition density function for CIR model
2 function p = CIRtdf(rS,rT,T,alpha,mu,sigma)
3 c = (2*alpha)/((sigma^2)*(1-exp(-alpha*T))); % T=T-S
4 u = c*exp(-alpha*T)*rS;
5 v = c*rT;
6 q = ((2*alpha*mu)/(sigma^2))-1;
7 z = 2*sqrt(u.*v);
8 bf= besseli(q,z,1); % Besseli function
9 p= c*exp(-u-v).*(v./u).^ (q/2).*bf.*exp(z);
10 end

```

```

1 % The log-likelihood function for CIR model
2 function logL = CIRlog(data,dt,alpha,mu,sigma)
3 Nsteps = length(data);
4 rs = data(1:Nsteps-1); %empirical data
5 rt = data(2:Nsteps);
6
7 c = (2*alpha)/((sigma^2)*(1-exp(-alpha*dt))); %dt=t-s
8 u = c*exp(-alpha*dt)*rs;
9 v = c*rt;
10 q = ((2*alpha*mu)/(sigma^2))-1;
11 z = 2*sqrt(u.*v);
12 bf = besseli(q,z,1);
13 logL = (Nsteps-1)*log(c) - ...
        sum(u+v-0.5*q*log(v./u)-log(bf)-z);
14 % the transition density function p= ...
        c*exp(-u-v).*(v./u).^ (q/2).*bf.*exp(z)
15 end

```

```

1 % The OLSE method for CIR model
2 function initialvalues = CIRols(data,dt)
3
4 Nsteps = length(data);
5 rs = data(1:Nsteps-1);
6 rt = data(2:Nsteps);
7
8 dr = diff(data)./sqrt(rs);
9 reg = [dt./sqrt(rs) dt*sqrt(rs)];
10 theta = reg\dr;

```

```

11 alpha0 = -theta(2);
12 mu0 = theta(1)/alpha0;
13 sigma0 = std(dr-reg*theta)/sqrt(dt);
14 initialvalues = [alpha0, mu0, sigma0];
15 % OLSE of the initial parameter values
16 end

```

```

1 % Paramter Estimation for CIR model
2 dataPribor3M = ...
    load('/Users/MUSEMZ/Documents/MATLAB/Mycode/
3 Pribor3M.mat','-regexp');
4 data = dataPribor3M.Pribor3M; % importing empirical data
5
6 dt=1/250; % a standard convention is 250 working days ...
    per year
7 initialvalues = CIRols(data,dt);
8 alpha0 = initialvalues(1);
9 mu0 = initialvalues(2);
10 sigma0 = initialvalues(3);
11
12 params_optim = fminsearch(@(params)-CIRlog(data, dt, ...
    params(1), params(2), params(3)), initialvalues);
13 % MLE method
14 alpha_optim = params_optim(1);
15 mu_optim = params_optim(2);
16 sigma_optim = params_optim(3);
17 [alpha0,mu0,sigma0; alpha_optim,mu_optim,sigma_optim]
18 % estimated initial paramter values and optimal parameters
19 [CIRlog(data,dt,alpha0,mu0,sigma0); ...
    CIRlog(data,dt,alpha_optim,mu_optim,sigma_optim)];
20
21 % plot Figures 2-4 in the thesis
22 figure(1)
23 ngrid1= 100;
24
25 % plot ln L as a function of alpha given optimal mu ...
    and sigma in Figure 2
26 subplot(3,1,1)
27 Alpha=alpha_optim/2:alpha_optim/(ngrid1-1):3*alpha_optim/2;
28 LnlAlpha=zeros(1,ngrid1);
29 for i=1:ngrid1
30     LnlAlpha(i)=CIRlog(data, dt, Alpha(i), mu_optim, ...
        sigma_optim);
31 end
32 plot(Alpha,LnlAlpha)
33 xlabel('\alpha');ylabel('lnL');
34
35 % plot ln L as a function of mu given optimal alpha ...

```

```

    and sigma in Figure 3
36 subplot(3,1,2)
37 Mu=mu_optim/2:mu_optim/(ngrid1-1):3*mu_optim/2;
38 LnlMu=zeros(1,ngrid1);
39 for i=1:ngrid1
40     LnlMu(i)=CIRlog(data, dt, alpha_optim, Mu(i), ...
        sigma_optim);
41 end
42 plot(Mu,LnlMu)
43 xlabel('\mu');ylabel('lnL');
44
45 % plot ln L as a function of sigma given optimal mu ...
    and alpha in Figure 4
46 subplot(3,1,3)
47 Sigma=sigma_optim/2:sigma_optim/(ngrid1-1):3*sigma_optim/2;
48 LnlSigma=zeros(1,ngrid1);
49 for i=1:ngrid1
50     LnlSigma(i)=CIRlog(data,dt,alpha_optim,mu_optim,Sigma(i));
51 end
52 plot(Sigma,LnlSigma)
53 xlabel('\sigma');ylabel('lnL');
54
55 % plot Figures 5-7 in the thesis
56 figure(2)
57 ngrid=30;
58
59 % plot ln L as a function of alpha and mu given ...
    optimal sigma in Figure 5
60 subplot(3,1,1)
61 Lnl=zeros(ngrid,ngrid);
62     [Alpha, ...
        Mu]=meshgrid(0.37:(0.50-0.37)/(ngrid-1):0.50, ...
63     0.053:(0.069-0.053)/(ngrid-1):0.069);
64 for i=1:ngrid
65     for j=1:ngrid
66         Lnl(i,j)=CIRlog(data, dt, Alpha(i,j), ...
            Mu(i,j), sigma_optim);
67     end
68 end
69 surfc(Alpha,Mu,Lnl)
70 xlabel('\alpha');ylabel('\mu');zlabel('lnL');
71
72 % plot ln L as a function of sigma and mu given ...
    optimal alpha in Figure 6
73 subplot(3,1,2)
74 Lnl=zeros(ngrid,ngrid);
75 [Mu, Sigma]=meshgrid(0.02:(0.15-0.02)/(ngrid-1):0.15, ...
76     0.14:(0.16-0.14)/(ngrid-1):0.16);
77 for i=1:ngrid

```



```

78     for j=1:ngrid
79         Lnl(i,j)=CIRlog(data, dt, alpha_optim, ...
                        Mu(i,j), Sigma(i,j));
80     end
81 end
82 surfc(Mu, Sigma, Lnl)
83 xlabel('\mu'); ylabel('\sigma'); zlabel('lnL');
84
85 % plot ln L as a function of alpha and sigma given ...
    optimal mu in Figure 7
86 subplot(3,1,3)
87 Lnl=zeros(ngrid,ngrid);
88 [Alpha, Sigma]=meshgrid(0.2:(0.8-0.2)/(ngrid-1):0.8, ...
                        0.146:(0.152-0.146)/(ngrid-1):0.152);
89 for i=1:ngrid
90     for j=1:ngrid
91         Lnl(i,j)=CIRlog(data, dt, Alpha(i,j), ...
                        mu_optim, Sigma(i,j));
92     end
93 end
94 end
95 surfc(Alpha, Sigma, Lnl)
96 xlabel('\alpha'); ylabel('\sigma'); zlabel('lnL');

```

```

1 % Empirical Pribor 3M data
2 figure(1)
3
4 alpha = alpha_optim;
5 mu = mu_optim;
6 sigma = sigma_optim;
7 T = (length(data)-1)/250;
8 x0 = data(1);
9
10 treal=linspace(0,T,length(data));
11 plot(treal,data)
12 xlabel('Time Series'); ylabel('Interest Rates in ...
    Percentage');

```

```

1 % The Euler approximation method for CIR model
2 function [t, yEuler] = ...
    Euler(alpha,mu,sigma,T,x0,Nsteps,nsim)
3 dt = T/(Nsteps-1); % arbitrary increment Δ t
4 y = zeros(Nsteps,nsim); % create a Nsteps-by-nsim ...
    array of zeros
5 y(1, :) = x0*ones(1,nsim); % nsim is the number of ...
    simulations
6 for i = 1:(Nsteps-1);
7     y(i+1,:)=y(i,:)+alpha*(mu-y(i,:))*dt+sigma*

```

```

8      sqrt(abs(y(i,:)).*randn(1,nsim)*sqrt(dt));
9  end
10 t = linspace(0,T,Nsteps); % generating Nsteps points ...
    and the space between the points is  $\Delta t$ 
11 yEuler = y;
12 end

```

```

1 % The exact algorithm for CIR model
2 function [t, xExact] = ...
    Exact(alpha,mu,sigma,T,x0,Nsteps,nsim)
3 dt = T/(Nsteps-1);
4 n = ...
    (4*alpha*exp(-alpha*dt))/((sigma^2)*(1-exp(-alpha*dt)));
5 % x(T) or n(t+dt)
6 df = (4*alpha*mu)/(sigma^2); % degrees of freedom
7 x = zeros(Nsteps,nsim);
8 x(1, :) = x0*ones(1,nsim);
9 for t = 1:Nsteps-1;
10     ncp = x(t,:)*n; % non-centrality parameter ...
        x(t)x(T) or x(t)n(t+dt)
11     % ncsrd = ncx2rnd(d, ncp, 1, 1); % non-central ...
        chi-square random var
12     Po = poissrnd(ncp/2);
13     v = df+2*Po;
14     % chi2prime = (randn+sqrt(ncp)^2)+ncsrd when d ...
        greater than 1
15     % x(t+1) = chi2prime*
16     chi2v = chi2rnd(v);
17     x(t+1,:) = chi2v*(exp(-alpha*dt)/n);
18 end
19 t = linspace(0,T,Nsteps);
20 xExact = x;
21 end

```

```

1 % Simulation of Exact and RealData
2 alpha = alpha_optim;
3 mu = mu_optim;
4 sigma = sigma_optim;
5 % optimal parameters for CIR model
6
7 T = (length(data)-1)/250;
8 x0 = data(1);
9 Nsteps = length(data)-1;
10 nsim = 2; %numbers of simulation
11
12 [tExact, xExact] = Exact(alpha,mu,sigma,T,x0,Nsteps,nsim);
13 treal=linspace(0,T,length(data));

```

```

14 plot(tExact,xExact,'k--',treal,data)
15 xlabel('Time Series');ylabel('Interest Rates');

```

```

1 % The log-likelihood function for Vasicek model
2 function Lnl = Vasiceklog(data,dt,a,b,sigma)
3
4 Nsteps = length(data);
5 rs = data(1:Nsteps-1);
6 rt = data(2:Nsteps);
7
8 alpha = exp(-a*dt);
9 beta = b*(1-alpha);
10 Vsqr = (sigma^2)/(2*a)*(1-exp(-2*a*dt));
11
12 Lnl =
13 (-1/2)*sum(log(2*pi)+log(Vsqr)+(rt-alpha*rs-beta).^2/Vsqr);
14 end

```

```

1 % The OLSE method for Vasicek model
2 function initialvalues = Vasicekols(data,dt)
3
4 Nsteps = length(data);
5 rs = data(1:Nsteps-1);
6 rt = data(2:Nsteps);
7
8 dr = diff(data);
9 reg = [dt*ones(Nsteps-1,1) dt*rs];
10 theta = reg\dr;
11 a0 = -theta(2);
12 b0 = theta(1)/a0;
13 sigma0 = std(dr-reg*theta)/sqrt(dt);
14 initialvalues = [a0, b0, sigma0];
15
16 end

```

```

1 % Paramter Estimation for CIR model
2 dataPribor3M = ...
    load('/Users/MUSEMZ/Documents/MATLAB/Vasicek ...
        Model/Pribor3M.mat','-regexp');
3 data = dataPribor3M.Pribor3M;
4
5 dt=1/250;
6 initialvalues = Vasicekols(data,dt);
7 a0 = initialvalues(1);
8 b0 = initialvalues(2);
9 sigma0 = initialvalues(3);
10

```

```

11 params_optim = fminsearch(@(params)-Vasiceklog(data, ...
    dt, params(1), params(2), params(3)), initialvalues);
12 a_optim = params_optim(1);
13 b_optim = params_optim(2);
14 s_optim = params_optim(3);
15 [a0,b0,sigma0; a_optim,b_optim,s_optim]
16 % estimated initial parameter values and optimal parameters
17 [Vasiceklog(data,dt,a0,b0,sigma0); ...
    Vasiceklog(data,dt,a_optim,b_optim,s_optim)];

```

```

1 % The exact algorithm for Vasicek model
2 function [t, r] = VasicekExact(a,b,sigma,T,x0,Nsteps,nsim)
3 dt = T/(Nsteps-1);
4 r = zeros(Nsteps,nsim);
5 r(1, :) = x0*ones(1,nsim);
6 for i = 1:(Nsteps-1);
7     r(i+1,:) = r(i,:) * exp(-a*dt) + b*(1-exp(-a*dt)) + sigma *
8         sqrt((1-exp(-2*a*dt))/(2*a)) * randn(1,nsim);
9 end
10 t = linspace(0,T,Nsteps);
11 end

```

```

1 %Vasicek Exact and Realdata Simulation
2 a = a_optim;
3 b = b_optim;
4 s = s_optim;
5 % optimal parameters for Vasicek model
6 T = (length(data)-1)/250;
7 x0 = data(1);
8 Nsteps = length(data)-1;
9 nsim = 2; %numbers of simulation
10
11 [tExact, xExact] = VasicekExact(a,b,s,T,x0,Nsteps,nsim);
12 treal=linspace(0,T,length(data));
13 plot(tExact,xExact,'k--',treal,data)
14 xlabel('Time Series');ylabel('Interest Rates');

```

