One Factor Machine Learning Gaussian Short Rate

J. Beleza Sousa
M2A/ADEETC, Instituto Superior de Engenharia de Lisboa, Instituto Politecnico de Lisboa CMA/DM, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
Email: jsousa@deetc.isel.ipl.pt
Address: Rua Conselheiro Emidio Navarro, 1, 1959-007 Lisboa, Portugal, Quinta da Torre, 2829-516 Caparica, Portugal
Telephone: +351 218317210

M. L. Esquvel
CMA/DM, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
Email: mle@fct.unl.pt
Address: Quinta da Torre, 2829-516 Caparica, Portugal
Telephone: +351 212948388

R. M. Gaspar
CEMAPRE, Instituto Superior de Economia e Gestão, Universidade Tecnica de Lisboa
Email: rmgaspar@iseg.utl.pt
Address: Rua Miguel Lupi 20, 1249-078 Lisboa, Portugal
Telephone: +351 213925923
In this paper we model the short rate, under the risk neutral measure, as a Gaussian process, conditioned on market observed zero coupon bonds log prices. The model is based on Gaussian processes for machine learning, using a single Vasicek factor as prior.

All model parameters are learned directly under the risk neutral measure, using zero coupon bonds log prices only. The model supports observations of zero coupon bonds with distinct maturities limited to one observation per time instant. All the supported observations are automatically fitted. We derive the model's SDE and model the Euribor using real data.

**KEYWORDS:**
Gaussian short rate, Gaussian processes for machine learning, risk neutral measure.
1 Introduction

In our previous work [8] we calibrated the Vasicek [9] short rate model, for a single $T$ maturity zero coupon bond, directly under the risk neutral measure, using zero coupon bond prices only. The method is based on conditioning the Vasicek zero coupon bond log prices Gaussian process, to market observed zero coupon bond log prices, using the Gaussian processes for Machine Learning framework [6]. In this paper we recognize the conditioned on market observed zero coupon log prices underlying Gaussian process as an alternative one factor Gaussian short rate model by itself. We extend the calibration method to several $T$ maturity zero coupon bonds, limited to one observation per time instant, and derive the model’s stochastic differential equation.
Consider the Gaussian process $g(x)$, with mean function $m(x)$ and covariance function $\text{cov}(x_i, x_j)$,

$$g(x) \sim \mathcal{GP}(m(x), \text{cov}(x_i, x_j)).$$

(1)

Consider also, data $\mathcal{D} = (X, y)$, where the matrix $X$ collects a set of vectors $\{x_1^c, x_2^c, \ldots, x_n^c\}$ where the value $y^c = g(x^c)$ was observed, and vector $y$ collects the corresponding set of observed values $\{y_1^c, y_2^c, \ldots, y_n^c\}$.

The process, $g_D(x)$, defined by all trajectories of $g(x)$ that pass through data $\mathcal{D}$ is also Gaussian [6], with mean function $m_D(x)$ and covariance function $\text{cov}_D(x_i, x_j)$,

$$g_D(x) \sim \mathcal{GP}(m_D(x), \text{cov}_D(x_i, x_j)).$$

(2)

The process $g(x)$ is called the prior, $g_D(x)$ is called the conditioned on data process, $\mathcal{D}$ is called the training set and $m_D(x)$, $\text{cov}_D(x_i, x_j)$ are given by [5]

$$m_D(x) = m(x) + K_\top X^{-1}(y - m)$$

(3)

and

$$\text{cov}_D(x_i, x_j) = \text{cov}(x_i, x_j) - K_\top X^{-1}K_{X,x}$$

(4)

where, $m$ is the prior training set mean vector, $K$ is the prior training set covariance matrix and $K_{X,x}$ is a prior covariance vector between every training vector and $x$.

The "One Factor Machine Learning Gaussian Short Rate" is a single factor arbitrage free Vasicek short rate prior, conditioned on market observed zero coupon bonds log prices.

2 Short rate prior

The model’s prior is that, under the arbitrage free, risk neutral measure, the short rate, $r(t)$, follows a Vasicek Ornstein-Uhlenbeck mean-reverting process, defined by the stochastic differential equation (SDE) [9]:

$$dr(t) = k(\theta - r(t))dt + \sigma dW(t).$$

(5)

Parameter $k$ is the mean reversion velocity, parameter $\theta$ is the mean interest rate level, parameter $\sigma$ is the volatility and $W(t)$ is the Wiener process. Parameters $k$ and $\sigma$ are strictly positive.
Let \( s \) be the initial time, with \( 0 < s < t \). The solution of Equation 5 is given by

\[
r(t) = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}) + \sigma e^{-kt} \int_s^t e^{ku} dW(u). \tag{6}
\]

The initial short rate value, \( r(s) \), is considered as an extra model parameter since its value must be obtained under the risk neutral measure.

### 2.1 Short rate mean and covariance

Given Equation 6 the short rate prior mean and covariance functions, \( m_r(t) \) and \( \text{cov}_r(t_1, t_2) \), are given by

\[
m_r(t) = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}) \tag{7}
\]

and

\[
\text{cov}_r(t_1, t_2) = \sigma^2 e^{-k(t_1 + t_2)} \frac{1}{2k} \left( e^{2k \min(t_1, t_2)} - e^{2ks} \right). \tag{8}
\]

### 2.2 Zero coupon bond log prices mean and covariance

Under the risk neutral measure, the price \( p \), at time \( t \), of a zero coupon bond that pays 1 at maturity \( T \), is given by [2]

\[
p(t, T) = E \left[ e^{-\int_t^T r(u)du} \right]. \tag{9}
\]

Under the Vasicek model \( p(t, T) \) is given by

\[
p(t, T) = e^{A(t,T)-B(t,T)r(t)} \tag{10}
\]

where

\[
B(t,T) = \frac{1}{k} \left(1 - e^{-k(T-t)}\right) \tag{11}
\]

and

\[
A(t,T) = \left( \theta - \frac{\sigma^2}{2k^2} \right) (B(t,T) - T + t) - \frac{\sigma^2}{4k} B^2(t, T). \tag{12}
\]

It is clear from Equation 10 that model has an affine term structure and that the logarithm of the zero coupon bonds prices is Gaussian.
\[
\log p(t, T) \sim \mathcal{GP}(m_p(t, T), \text{cov}_p(t_i, T_i, t_j, T_j)).
\]

Since the logarithm of zero coupon bonds prices is given by
\[
\log p(t, T) = A(t, T) - B(t, T)r(t)
\]
the mean and covariance \( m_p(t, T) \) and \( \text{cov}_p(t_1, T_1, t_2, T_2) \) are given by
\[
m_p(t, T) = A(t, T) - B(t, T)m_r(t)
\]
\[
= \left( \theta - \frac{\sigma^2}{2k^2} \right) \left( \frac{1 - e^{-k(T-t)}}{k} - T + t \right)
- \frac{\sigma^2 (1 - e^{-k(T-t)})^2}{4k^3}
\]
\[
- \frac{(1 - e^{-k(T-t)}) \left( r(s)e^{-k(t-s)} + \theta \left( 1 - e^{-k(t-s)} \right) \right)}{k}
\]

and
\[
\text{cov}_p(t_1, T_1, t_2, T_2) = B(t_1, T_1)B(t_2, T_2)\text{cov}_r(t_1, t_2)
\]
\[
= \frac{\sigma^2 e^{-k(t_1+t_2)} (1 - e^{-k(T_1-t_1)}) (1 - e^{-k(T_2-t_2)}) \left( e^{2k \min(t_1, t_2)} - e^{2k s} \right)}{2k^3}
\]

3 One Factor Machine Learning Gaussian Short Rate

Let:
- \( \mathbf{x} = [t \ T]^\top; \)
- \( y = \log p(\mathbf{x}) = \log p(t, T); \)
- \( m_p(\mathbf{x}) = m_p(t, T); \)
- \( \text{cov}_p(\mathbf{x}_i, \mathbf{x}_j) = \text{cov}_p(t_i, T_i, t_j, T_j). \)

Following Section 1, let matrix \( \mathbf{X} \) collect a set of vectors \( \mathbf{x} \) where the values of zero coupon log prices were observed, and let vector \( \mathbf{y} \) collect the
corresponding values \( y^\circ = \log p(\mathbf{x}^\circ) \). Recall from Section 1 that \( \mathcal{D} = (\mathbf{X}, \mathbf{y}) \) is the training set.

The “One Factor Machine Learning Gaussian Short Rate” is the Gaussian short rate process, \( r_{\mathcal{D}}(t) \), underlying the zero coupon bond prices

\[
p_{\mathcal{D}}(t, T) = E \left[ e^{-\int_t^T r_{\mathcal{D}}(u) du} \right]
\]

where \( \log p_{\mathcal{D}}(t, T) = \log p_{\mathcal{D}}(\mathbf{x}) \) is the conditioned on zero coupon bonds log prices Gaussian process

\[
\log p_{\mathcal{D}}(\mathbf{x}) \sim \mathcal{GP} \left( m_{p_{\mathcal{D}}}(\mathbf{x}), \text{cov}_{p_{\mathcal{D}}}(\mathbf{x}_i, \mathbf{x}_j) \right)
\]

using \( \log p(\mathbf{x}) \) as prior.

Given equations 3 and 4, \( m_{p_{\mathcal{D}}}(\mathbf{x}) \) and \( \text{cov}_{p_{\mathcal{D}}}(\mathbf{x}_i, \mathbf{x}_j) \) are given by

\[
m_{p_{\mathcal{D}}}(\mathbf{x}) = m_p(\mathbf{x}) + \mathbf{K}_x^\top \mathbf{X} \mathbf{K}_x^{-1} (\mathbf{y} - \mathbf{m})
\]

and

\[
\text{cov}_{p_{\mathcal{D}}}(\mathbf{x}_i, \mathbf{x}_j) = \text{cov}_p(\mathbf{x}_i, \mathbf{x}_j) - \mathbf{K}_x^\top \mathbf{X} \mathbf{K}_x^{-1} \mathbf{K}_x
\]

where

- \( \mathbf{m} \) is the prior mean on the training set. It results from applying \( m_p(\mathbf{x}) \) function (Equation 16) on all \( \mathbf{X} \) collected vectors;
- \( \mathbf{K} \) is the prior covariance matrix on the training set. It results from applying \( \text{cov}_p(\mathbf{x}_i, \mathbf{x}_j) \) function (Equation 18) on all pairs of \( \mathbf{X} \) collected vectors;
- \( \mathbf{K}_x \) is the prior covariance between every vector in the training set and \( \mathbf{x} \). It results from applying \( \text{cov}_p(\mathbf{x}_i, \mathbf{x}_j) \) function (Equation 18) on all pairs composed by each \( \mathbf{X} \) collected vector, and the \( \mathbf{x} \) vector.

### 3.1 Properties

1. The model supports a single observation \( y \) on each vector \( \mathbf{x} = [t\ T]^\top \).

Given Equation 18, the observation of more than one log price \( y \), of a \( T \) maturity zero coupon bond, at time \( t \), would result in two equal lines in matrix \( \mathbf{K} \), which would not be invertible, as required by Equations 21 and 22.
2. The model only supports the observation of a single \( T \) maturity zero coupon bond log price on each time \( t \).

As in the previous property, given Equation 17, the observation of more than one \( T \) maturity zero coupon bond log price, on each time \( t \), would result in two equal lines in matrix \( K \).

3. Despite the parameters values, all supported observations are automatically fitted.

Given Equation 21, the model mean, \( m_{pD}(x) \), on each training observation, \( y^\diamond = \log p(x^\diamond) \), equals \( y^\diamond \):

\[
m_{pD}(x^\diamond) = y^\diamond.
\]  

(23)

Furthermore, given Equation 22, the model variance on each training observation equals zero:

\[
cov_{pD}(x^\diamond, x^\diamond) = 0.
\]  

(24)

3.2 SDE

In order to obtain the "One Factor Machine Learning Gaussian Short Rate" SDE we assume, at a first moment, that the conditioned on data short rate follows, under the arbitrage free, risk neutral measure, the deterministic time dependent parameters SDE of the generalized Hull and White\(^1\) model.

Then, writing the generalized Hull and White model zero coupon bond log prices mean and covariance functions, making them equal to the corresponding "One Factor Machine Learning Gaussian Short Rate" functions, and solving in order to the deterministic time dependent parameters, we get the SDE parameters and confirm our initial assumption.

Let the "One Factor Machine Learning Gaussian Short Rate" short rate, \( r_D(t) \), follow, under the arbitrage free, risk neutral measure, the generalized Hull and White model SDE [4]

\[
dr_D(t) = (\theta(t) - \alpha(t)r_D(t))dt + \sigma(t)dW(t).
\]  

(25)

The solution of Equation 25 is

\(^1\)The generalized Hull and White model is also known as the extended Vasicek model, as it is a Vasicek model extended with time dependent parameters.
\( r_D(t) = r_D(s) e^{H(s) - H(t)} + e^{-H(t)} \int_s^t e^{H(u)} \theta(u) du + e^{-H(t)} \int_s^t e^{H(u)} \sigma(u) dW(u) \)  

(26)

where

\[ H(t) = \int_0^t \alpha(u) du. \]  

(27)

Under this assumption:

1. The short rate mean, \( m_{r_D}(t) \), and covariance, \( \text{cov}_{r_D}(t_1, t_2) \), functions are given by

\[ m_{r_D}(t) = r_D(s) e^{H(s) - H(t)} + e^{-H(t)} \int_s^t e^{H(u)} \theta(u) du \]  

(28)

and

\[ \text{cov}_{r_D}(t_1, t_2) = e^{-H(t_1) - H(t_2)} \int_s^{\min(t_1, t_2)} e^{2H(u)} \sigma^2(u) du; \]  

(29)

2. The zero coupon bond prices are given by [1]

\[ p_D(t, T) = e^{A_D(t, T) - B_D(t, T) r_D(t)} \]  

(30)

where

\[ B_D(t, T) = e^{H(t)} \int_t^T e^{-H(u)} du \]  

(31)

and

\[
A_D(t, T) = \int_t^T \int_t^s e^{-H(u) - H(s)} \int_t^u e^{2H(v)} \sigma^2(v) dv du ds \\
- \int_t^T e^{-H(u)} \int_t^u e^{H(v)} \theta(v) dv du; \]

(32)

3. The model is affine and the zero coupon bond log prices, \( \log p_D(t, T) \), are Gaussian

\[ \log p_D(x) \sim \mathcal{GP}(m_{p_D}(x), \text{cov}_{p_D}(x_i, x_j)) \]  

(33)
4. The zero coupon bond log prices mean and covariance are given by

\[
m_{\text{pD}}(t, T) = AD(t, T) - B\text{D}(t, T) m_{rD}(t) \tag{34}
\]

and

\[
cov_{\text{pD}}(t_1, T_1, t_2, T_2) = BD(t_1, T_1)BD(t_2, T_2)cov_{rD}(t_1, t_2). \tag{35}
\]

### 3.2.1 Parameter \( \alpha(t) \)

In order to get parameter \( \alpha(t) \) we first note that, under the risk neutral measure, the initial value of the short rate conditioned on data, \( r_D(s) \), equals the prior initial value, \( r(s) \), because the observations that distinguish the two processes occur, by definition of the initial time, at times greater than \( s \).

Then, expanding the zero coupon bond log prices mean conditioned on data, \( m_{\text{pD}}(t, T) \), in Equations 34 and 21, and making the term with \( r_D(s) \) equal to the term with \( r(s) \) we get

\[
BD(t, T)r_D(s)e^{H(s)-H(t)} = B(t, T)r(s)e^{-k(t-s)} \tag{36}
\]

\[
BD(t, T)r(s)e^{H(s)-H(t)} = B(t, T)r(s)e^{-k(t-s)}. \tag{37}
\]

Making the exponential functions of \( s \) equal

\[
\Leftrightarrow e^{H(s)-H(t)} = e^{-k(t-s)}
\]

\[
\Leftrightarrow \int_s^t \alpha(u)du = k(t - s)
\]

\[
\Leftrightarrow \alpha(t) = k \tag{38}
\]

Therefore,

\[
H(t) = \int_0^t \alpha(u)du = \int_0^t kdu = kt. \tag{39}
\]

Also, substituting Equation 39 in Equation 31 shows that \( BD(t, T) \) becomes equal to \( B(t, T) \)
\[ B_D(t, T) = e^{H(t)} \int_t^T e^{-H(u)} du = e^{kt} \int_t^T e^{-ku} du = \frac{1}{k} (1 - e^{-k(T-t)}) = B(t, T). \] (40)

### 3.2.2 Parameter $\sigma(t)$

Given Equations 35 and 40, and setting $\text{cov}_{pD}(t_1, T_1, t_2, T_2)$, in Equations 35 and 22, equal

\[
B(t_1, T_1)B(t_j, T_j)\text{cov}_{rD}(t_i, t_j) = B(t_1, T_1)B(t_j, T_j)\text{cov}_{r}(t_i, t_j) - \mathbf{K}_{X, [t, T]}\mathbf{K}^{-1}\mathbf{K}_{X, [t, T]}^\top.
\] (41)

Inserting Equation 17,

\[
B(t_1, T_1)B(t_j, T_j)\text{cov}_{rD}(t_i, t_j) = B(t_1, T_1)B(t_j, T_j)\text{cov}_{r}(t_i, t_j) - \mathbf{K}_{X, [t, T]}\mathbf{K}^{-1}\mathbf{K}_{X, [t, T]}^\top
\Rightarrow \text{cov}_{rD}(t_i, t_j) = \text{cov}_{r}(t_i, t_j) - \frac{\mathbf{K}_{X, [t, T]}\mathbf{K}^{-1}\mathbf{K}_{X, [t, T]}^\top}{B(t_1, T_1)B(t_j, T_j)}
\] (42)

$\mathbf{K}_{X, [t, T]}^\top$ is a vector of prior covariances, $\text{cov}_{p}(t_i, T_i, t_j, T_j)$, between every $T^o$ maturity zero coupon bond log price at time $t^o$ in the training set, and a $T$ maturity zero coupon bond log price at time $t$. Given equation 17,

\[
\frac{\text{cov}_{p}(t^o, T^o, t, T)}{B(t, T)} = \frac{B(t^o, T^o)B(t, T)\text{cov}_{r}(t^o, t)}{B(t, T)}
\Rightarrow \frac{\text{cov}_{p}(t^o, T^o, t, T)}{B(t, T)} = B(t^o, T^o)\text{cov}_{r}(t^o, t).
\] (43)

Therefore

\[
\text{cov}_{rD}(t_i, t_j) = \text{cov}_{r}(t_i, t_j) - \mathbf{V}_{X, t_i}^\top \mathbf{K}^{-1} \mathbf{V}_{X, t_j}
\] (44)
where each element $v_{\{t^o,T\}^\top,t}$ of vector $V_{X,t}$ is given by
\[
v_{\{t^o,T\}^\top,t} = B(t^o,T^o) \text{cov}_t(t^o,t).
\] (45)

Given Equations 29 and 44
\[
e^{-k(t_i+t_j)} \int_{s}^{\min(t_i,t_j)} e^{2ku^2} du = \frac{\sigma^2 e^{-k(t_i+t_j)}}{2k} \left( e^{2k\min(t_i,t_j)} - e^{2ks} \right) - V_{X,t_i}^\top K^{-1} V_{X,t_j}.
\] (46)

Equation 46 applies to every $t_i$ and $t_j$, in particular if $t_i = t_j = t$
\[
e^{-2kt} \int_{s}^{t} e^{2ku^2} du = \frac{\sigma^2 e^{-2kt}}{2k} \left( e^{2kt} - e^{2ks} \right) - V_{X,t}^\top K^{-1} V_{X,t}.
\] (47)

Therefore
\[
\int_{s}^{t} e^{2ku^2} du = \frac{\sigma^2}{2k} \left( e^{2kt} - e^{2ks} \right) - \frac{V_{X,t}^\top K^{-1} V_{X,t}}{e^{-kt}}
\]
\[
= \frac{\sigma^2}{2k} \left( e^{2kt} - e^{2ks} \right) - U_{X,t}^\top K^{-1} U_{X,t}
\] (48)

where each element $u_{\{t^o,T^o\}^\top,t}$ of vector $U_{X,t}$ is given by
\[
u_{\{t^o,T^o\}^\top,t} = B(t^o,T^o) e^{-kt} \text{cov}_t(t^o,t)
\]
\[
= B(t^o,T^o) \sigma^2 e^{-k(t^o+t)} \frac{1}{2k} \left( e^{2k\min(t^o,t)} - e^{2ks} \right)
\]
\[
= B(t^o,T^o) \sigma^2 e^{-k(t^o)} \frac{1}{2k} \left( e^{2k\min(t^o,t)} - e^{2ks} \right)
\] (49)

Differentiating both sides of Equation 48 w.r.t. $t$
\[
e^{2kt} \sigma^2(t) = \frac{d}{dt} \left( \frac{\sigma^2}{2k} \left( e^{2kt} - e^{2ks} \right) - U_{X,t}^\top K^{-1} U_{X,t} \right)
\]
\[
= \frac{\sigma^2}{2k} \left( e^{2kt} - 2U_{X,t}^\top K^{-1} \frac{d}{dt} U_{X,t} \right)
\]
\[
= \sigma^2 e^{2kt} - 2U_{X,t}^\top K^{-1} Q_{X,t}
\] (50)

where each element $q_{\{t^o,T^o\}^\top,t}$ of vector $Q_{X,t}$ is given by
\begin{align}
q_{[t^o \ T^o] \ T^o, t} &= 1_{\mathbb{R}^+} (t^o - t) B(t^o, T^o) \sigma^2 e^{-kt^o} e^{2kt}. \tag{51}
\end{align}

Therefore
\begin{align}
\sigma^2(t) &= \sigma^2 - 2U^\top_{X,t} K^{-1} R_{X,t} \tag{52}
\end{align}

where each element \( r_{[t^o \ T^o] \ T^o, t} \) of vector \( R_{X,t} \) is given by
\begin{align}
r_{[t^o \ T^o] \ T^o, t} &= 1_{\mathbb{R}^+} (t^o - t) B(t^o, T^o) \sigma^2 e^{-kt^o} \tag{53}
\end{align}

and
\begin{align}
\sigma(t) &= (\sigma^2 - 2U^\top_{X,t} K^{-1} R_{X,t})^{\frac{1}{2}}. \tag{54}
\end{align}

### 3.2.3 Parameter \( \Theta(t) \)

Expanding \( m_{\text{mD}}(t, T) \) in Equations 34 and 21
\begin{align}
A_D(t, T) - B_D(t, T) m_{\text{mD}}(t) &= m_p(t, T) + K^\top_{X,[t \ T]} K^{-1} (y - m) \tag{55}
\end{align}

using Equations 34 \((A_D(t, T))\), 40 \((B_D(t, T)\) equal to \(B(t, T)\)), 28 \((m_{\text{mD}}(t))\), 38 \((\alpha(t) = k)\) and 16 \((m_p(t, T))\), canceling the terms \(B(t, T)r(s)e^{-k(t-s)}\) on both sides (Equation 36), keeping the terms with \( \Theta(t) \) in the left hand side and moving all other terms to the right hand side, equality of Equation 55 becomes
\begin{align}
&- \int_t^T e^{-ku} \int_t^u e^{kv} \theta(v) dv du - \frac{1 - e^{-k(T-t)}}{k} e^{-kt} \int_t^t e^{ku} \theta(u) du \\
&= - \int_t^T \int_t^u e^{-k(u+s)} \int_t^u e^{2kv} \sigma^2(v) dv du ds \\
&+ \left( \theta - \frac{\sigma^2}{2k^2} \right) \left( 1 - \frac{e^{-k(T-t)}}{k} - T + t \right) \\
&- \frac{\sigma^2 (1 - e^{-k(T-t)})^2}{4k^3} \\
&- \frac{1 - e^{-k(T-t)}}{k} \theta (1 - e^{-k(t-s)}) \\
&+ K^\top_{X,[t \ T]} K^{-1} (y - m). \tag{56-61}
\end{align}
Differentiating the left hand side in Expression 56 w.r.t. $T$

\[
\frac{d}{dT} \left( -\int_{t}^{T} e^{-ku} \int_{t}^{u} e^{kv} \theta(v) dv du - \frac{1 - e^{-k(T-t)}}{k} e^{-kt} \int_{s}^{t} e^{ku} \theta(u) du \right)
\]

\[
= -e^{-kT} \int_{t}^{T} e^{ku} \theta(v) dv - e^{-kT} \int_{s}^{t} e^{ku} \theta(u) du
\]

(62)

Equation 62 applies to every $t \leq T$, in particular, if $t = T$ becomes

\[
\left. \left( \frac{d}{dT} \left( -\int_{t}^{T} e^{-ku} \int_{t}^{u} e^{kv} \theta(v) dv du - \frac{1 - e^{-k(T-t)}}{k} e^{-kt} \int_{s}^{t} e^{ku} \theta(u) du \right) \right) \right|_{t=T}
\]

\[
= -e^{-kT} \int_{s}^{T} e^{ku} \theta(u) du.
\]

(63)

Regarding the right hand side part in Expression 57, differentiating w.r.t. $T$, and making $t = T$, cancel this part

\[
\left( -\frac{d}{dT} \int_{t}^{T} e^{-k(u+s)} \int_{t}^{u} e^{2kv} \sigma^2(v) dv dus \right) \bigg|_{t=T}
\]

\[
= \left( -\int_{t}^{T} e^{-k(u+T)} \int_{t}^{u} e^{2kv} \sigma^2(v) dv du \right) \bigg|_{t=T}
\]

\[
= 0.
\]

(64)

Proceeding with the right hand side part in Expressions 58, 59 and 60, differentiating w.r.t. $T$, and making $t = T$,

\[
\left( \frac{d}{dT} \left( \theta - \frac{\sigma^2}{4K^2} \left( \frac{1 - e^{-k(T-t)}}{k} - T + t \right) \right. \right.
\]

\[
\left. \left. - \frac{\sigma^2}{k} \left( 1 - e^{-k(T-t)} \right) \right)^2 \right|_{t=T}
\]

\[
= \theta \left( e^{-k(T-s)} - 1 \right).
\]

(65)

Regarding the part in Expression 61, recall that each element $k_{l^\circ T^\circ}^\top [l T]^\top$ of vector $K_{X[l^\circ T^\circ]}^\top$ is given by the covariance $\text{cov}_p(x_i, x_j)$, in Equation 17, between each vector $[l^\circ T^\circ]^\top$ in the training set, and $[l T]^\top$
\[ k_{[t^o, T^o]^\top, [t, T]^\top} = B(t^o, T^o)B(t, T)\text{cov}_r(t^o, t). \]  
\[ (66) \]

Given that
\[ \left. \frac{d}{dT} B(t, T) \right|_{t=T} = 1, \]
\[ (67) \]
differentiating w.r.t. \( T \), and making \( t = T \), the part in Expression 61
\[ \left. \frac{d}{dT} \left( \mathbf{K}_{X,[t, T]^\top} \mathbf{K}^{-1}(y - m) \right) \right|_{t=T} = \mathbf{N}_{X,T}^\top \mathbf{K}^{-1}(y - m) \]
\[ (68) \]
where each element, \( n_{[t^o, T^o]^\top, T} \), of vector \( \mathbf{N}_{X,T} \) is given by
\[ n_{[t^o, T^o]^\top, T} = B(t^o, T^o)\text{cov}_r(t^o, T) \]
\[ = \left( 1 - e^{-k(T^o - t^o)} \right) k \sigma^2 e^{-kT} \frac{1}{2} e^{2k\min(t^o, T) - e^{2ks}}. \]
\[ (69) \]

Grouping together the results in Equations 63, 65 and 68
\[ -e^{kT} \int_s^T e^{ku}(u) du = \theta \left( e^{-k(T-s)} - 1 \right) + \mathbf{N}_{X,T}^\top \mathbf{K}^{-1}(y - m) \]
\[ \Leftrightarrow \int_s^T e^{ku}(u) du = -\theta e^{ks} + \theta e^{kT} - \frac{\mathbf{N}_{X,T}^\top \mathbf{K}^{-1}(y - m)}{e^{-kT}}. \]
\[ (70) \]
Differentiating w.r.t. \( T \), Equation 70
\[ e^{kT}\theta(T) = k\theta e^{kT} - \frac{d}{dT} \left( \frac{\mathbf{N}_{X,T}^\top \mathbf{K}^{-1}(y - m)}{e^{-kT}} \right) \]
\[ (71) \]
and given that
\[ \frac{d}{dT} \left( \frac{n_{[t^o, T^o]^\top, T}}{e^{-kT}} \right) \]
\[ = \frac{d}{dT} \left( \frac{\left( 1 - e^{-k(T^o - t^o)} \right) k \sigma^2 e^{-kT} \frac{1}{2} e^{2k\min(t^o, T) - e^{2ks}}}{e^{-kT}} \right) \]
\[ = \frac{d}{dT} \left( \frac{\left( 1 - e^{-k(T^o - t^o)} \right)}{k} \sigma^2 e^{-kT} e^{2kT} \right), \]
\[ (72) \]
\( \theta(T) \) is given by

\[
\theta(T) = k\theta - Z_{X,T}^\top K^{-1}(y - m) \quad (73)
\]

where each element \( z_{[t^o \ T^o]}^\top, T \) of vector \( Z_{X,T} \) is given by

\[
z_{[t^o \ T^o]}^\top, T = \mathbb{1}_{\mathbb{R}^+}(t^o - T) \left( 1 - e^{-k(T^o - t^o)} \right) \sigma^2 e^{-kt} e^{kT}.
\] (74)

Finally, since Equation 73 applies to all \( T > s \), it can be rewritten as a function of \( t \)

\[
\theta(t) = k\theta - Z_{X,t}^\top K^{-1}(y - m). \quad (75)
\]

### 3.3 Learning the parameters

The prior parameters, \( r(s) \), \( \theta \), \( k \) and \( \sigma \), are obtained, directly under the risk neutral measure, by maximizing the prior log likelihood, \( L \), of these parameters, given the market observed zero coupon bonds log prices in the training set \( D \)

\[
L = -\frac{1}{2} \log |K| - \frac{1}{2} (y - m)^T K^{-1} (y - m) - \frac{n}{2} \log(2\pi) \quad (76)
\]

using the closed forms of the log likelihood derivative w.r.t. to each of the parameters [8].

### 4 Simulation

In order to evaluate the ability of learning parameters from data, we executed the following simulation procedure:

1. Fix the initial time to zero, and the time period to 1 year. Fix the time increment to 1/260 (assuming quotes on 5 working days per week, 52 weeks per year);

2. Fix the set of possible maturities to 7, 14 and 21 days, and 1 to 12 months (assuming 30 days months);

3. Fix the set of prior parameters to \( r(0) = 0.035 \), \( k = 0.26 \), \( \theta = 0.08 \) and \( \sigma = 0.04 \) (the approximate values obtained in the real data case described in Section 5);
Table 1: Prior parameters $r(0)$, $k$, $\theta$ and $\sigma$, mean, standard deviation and 95% confidence interval, learned from 1000 simulated data calibrations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fixed Value</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r(0)$</td>
<td>0.035</td>
<td>0.035</td>
<td>0.0016</td>
<td>0.032 to 0.038</td>
</tr>
<tr>
<td>$k$</td>
<td>0.26</td>
<td>0.26</td>
<td>0.027</td>
<td>0.20 to 0.031</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.08</td>
<td>0.081</td>
<td>0.006</td>
<td>0.070 to 0.093</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.04</td>
<td>0.039</td>
<td>0.0014</td>
<td>0.037 to 0.042</td>
</tr>
</tbody>
</table>

4. Use Equations 7 and 8 simulate one prior short rate trajectory;

5. Use Equation 14 to compute one, randomly selected maturity, zero coupon bond log price for each day;

6. Learn the prior parameters using the simulated data, the method described in Section 3.3 and the Conjugate Gradient method, available in Wolfram Mathematica 9 [7];

7. Repeat the previous steps 4 to 6, for 1000 trajectories.

Figure 1 shows the learned parameters histograms and Table 1 the corresponding mean, standard deviation, and 95% confidence intervals.

As it can be observed in Table 1, all the confidence intervals include the corresponding fixed value in step 3.
Figure 2: Short rate SDE deterministic time dependent parameters $\alpha(t)$, $\theta(t)$ and $\sigma(t)$, for one of the simulated trajectories.

For purposes of illustration, Figure 2 shows the short rate SDE deterministic time dependent parameters, $\alpha(t)$, $\theta(t)$ and $\sigma(t)$, of Equations 38, 75, and 54, respectively, of one of the simulated trajectories, computed from the learned prior parameters and the simulated data. These are the deterministic time dependent parameters of the short rate SDE in Equation 25 that exactly fit the simulated zero coupon bond log prices observations in the training set (the additional version of $\sigma(t)$, in the vicinity of 1.0, shows the detailed evolution of $\sigma(t)$, which can not be observed with the original time scale).

5 Real data

In this section we model the Euribor rates, quoted by the Portuguese bank Caixa Geral de Depósitos (CGD), which belongs to the Euribor panel banks. All the data used is publicly available at the Euribor Internet site\(^2\).

\(^2\)http://www.euribor-ebf.eu/
Prior Parameter | \( r(0) \) | \( k \) | \( \theta \) | \( \sigma \)  
---|---|---|---|---
Learned Value | 0.0349 | 0.2661 | 0.0826 | 0.0381  

Table 2: Euribor model prior parameters, learned from one randomly selected maturity quote per day, from CGD bank, during 2007 and 2008.

We choose the crisis years of 2007 and 2008 as the period to model.

Given the model limitation in Property 2, Section 3.1, we randomly selected one of the 15 Euribor maturities to get one observation for each day in the chosen period. The selected Euribor rates were converted to the equivalent zero coupon bond log prices and used as the training set.

Table 2 shows the prior parameters learned from the real data used as the training set.

Figure 3 illustrates the short rate SDE deterministic time dependent parameters, \( \alpha(t) \), \( \theta(t) \) and \( \sigma(t) \), in Equation 25, computed from the learned prior parameters and data, using Equations 38, 75, and 54, respectively.

6 Conclusions

In this paper we propose to model the short rate, under the arbitrage free risk neutral measure, as a conditioned on zero coupon bonds log prices Gaussian process.

All model parameters are learned directly under the risk neutral measure, using zero coupon bonds log prices only.

The model supports observations of zero coupon bonds with distinct maturities limited to one observation per time instant. All the supported observations are automatically fitted.

References


Figure 3: Short rate SDE deterministic time dependent parameters $\alpha(t)$, $\theta(t)$ and $\sigma(t)$, for the Euribor, quoted by CGD during 2007 and 2008, using a single randomly selected maturity per day.


