Machine Learning Vasicek Model Calibration with Gaussian Processes

J. Beleza Sousa [a b] , M. L. Esquível [b] & R. M. Gaspar [c]

[a] M2A/DEETC, Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa , Lisboa , Portugal

[b] CMA/DM, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa , Caparica , Portugal

[c] Advance Research Center, Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa , Lisboa , Portugal
Published online: 01 Feb 2012.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis
Taylor & Francis Group

Machine Learning Vasicek Model Calibration with Gaussian Processes

J. BELEZA SOUSA[1,2], M. L. ESQUÍVEL[2], AND R. M. GASPAR[3]

[1]M2A/DEETC, Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, Lisboa, Portugal
[2]CMA/DM, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
[3]Advance Research Center, Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa, Lisboa, Portugal

In this article, we calibrate the Vasicek interest rate model under the risk neutral measure by learning the model parameters using Gaussian processes for machine learning regression. The calibration is done by maximizing the likelihood of zero coupon bond log prices, using mean and covariance functions computed analytically, as well as likelihood derivatives with respect to the parameters. The maximization method used is the conjugate gradients. The only prices needed for calibration are zero coupon bond prices and the parameters are directly obtained in the arbitrage free risk neutral measure.

1. Introduction

Calibration of interest rate models under the risk neutral measure typically entails the availability of some derivatives such as swaps, caps or swaptions.

In this article we present an alternative method for calibrating Gaussian models, namely, the Vasicek interest rate model (Vasicek, 1977), which requires zero coupon bond prices only.

The presented method has the following features.

• The only prices needed for calibration are zero coupon bond prices.
• All the model parameters are directly obtained in the risk neutral measure.

- The calibration method does not require a discrete model approximation nor the establishment of an objective measure dynamics.

The method is based on Gaussian processes for Machine Learning, and its main drawback is his applicability to Gaussian models only.

One key issue in using Gaussian processes for machine learning is to have enough prior information on the data, in order to specify mean and covariance functions. Under the Vasicek interest rate model, the risk neutral zero coupon bond prices follow a log normal distribution, which can easily be transformed into a Gaussian process by taking the logarithm of the zero coupon prices. The mean and covariance functions of this Gaussian process can be computed analytically making it suitable for Gaussian processes for machine learning regression.

2. Vasicek Interest Rate Model

In the Vasicek model, the interest rate follows an Ornstein-Uhlenbeck mean-reverting process, under the risk neutral measure, defined by the stochastic differential equation

$$dr(t) = k(\theta - r(t))dt + \sigma \, dW(t), \tag{1}$$

where k is the mean reversion velocity, θ is the mean interest rate level, σ is the volatility, and $W(t)$ the Wiener process, and k and σ are positive.

Let $s \leq t$. The solution of Eq. (1) is (Brigo and Mercurio, 2006)

$$r(t) = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}) + \sigma e^{-kt} \int_s^t e^{ku} dW(u). \tag{2}$$

The interest rate $r(t)$, conditioned on $\mathscr{F}_s$, is normally distributed with mean

$$E\{r(t) \,|\, \mathscr{F}_s\} = r(s)e^{-k(t-s)} + \theta\left(1 - e^{-k(t-s)}\right) \tag{3}$$

and variance

$$Var\{r(t) \,|\, \mathscr{F}_s\} = \frac{\sigma^2}{2k}\left(1 - e^{-2k(t-s)}\right).$$

Zero coupon bonds are interest rate derivatives, therefore, their market prices are observed in the risk neutral measure. The Vasicek model has affine term structure, which means that the T maturity zero coupon bond prices $p(t, T)$, observed in the risk neutral measure, are given by (Björk, 2004):

$$p(t, T) = e^{A(t,T) - B(t,T)r(t)}, \tag{4}$$

where

$$A(t, T) = \left(\theta - \frac{\sigma^2}{2k^2}\right)(B(t, T) - T + t) - \frac{\sigma^2}{4k}B(t, T)$$

and

$$B(t, T) = \frac{1}{k}(1 - e^{-k(T-t)}).$$

Equation (4) shows that the zero coupon bond prices $p(t, T)$ are log normal and consequently $\log(p(t, T))$ are normal.

2.1. *Zero Coupon Bond Log Prices Mean Function*

Since

$$\log(p(t, T)) = A(t, T) - B(t, T)r(t) \tag{5}$$

the mean function $\mu(t, T)$ of $\log(p(t, T))$ is given by

$$
\begin{aligned}
\mu(t, T) &= E\left\{\log(p(t, T)) \,|\, \mathscr{F}_s\right\} \\
&= E\left\{A(t, T) - B(t, T)r(t) \,|\, \mathscr{F}_s\right\} \\
&= A(t, T) - B(t, T)E\left\{r(t) \,|\, \mathscr{F}_s\right\}.
\end{aligned}
$$

Considering the initial instant $s = 0$, and using Eq. (3) for $E\{r(t) \,|\, \mathscr{F}_s\}$ we get

$$
\begin{aligned}
\mu(t, T) &= A(t, T) - B(t, T)\left(r_0 e^{-kt} + \theta\left(1 - e^{-kt}\right)\right) \\
&= \frac{1}{4k^3}e^{-kT}\left(-4\left(e^{k(-t+T)} - 1\right)k^2(r_0 - \theta)\right. \\
&\quad - 4e^{kT}k^3(T - t)\theta + 2(e^{kt} - e^{kT})\sigma^2 \\
&\quad \left. + k\left(e^{kt} + e^{kT}(2T - 2t - 1)\right)\sigma^2\right),
\end{aligned} \tag{6}
$$

where r_0 stands for the initial interest rate value, the value of the interest rate $r(t)$, at $t = 0$.

2.2. *Zero Coupon Bond Log Prices Covariance Function*

The covariance function $cov(t_1, t_2, T)$ of $\log(p(t, T))$ is given by

$$
\begin{aligned}
cov(t_1, t_2, T) &= E\left\{(\log(p(t_1, T)) - \mu(t_1, T))\right. \\
&\quad \left. \times (\log(p(t_2, T)) - \mu(t_2, T)) \,|\, \mathscr{F}_s\right\} \\
&= E\left\{\log(p(t_1, T))\log(p(t_2, T)) \,|\, \mathscr{F}_s\right\} - \mu(t_1, T)\mu(t_2, T).
\end{aligned} \tag{7}
$$

Using Eq. (5), the term $E\left\{\log(p(t_1, T))\log(p(t_2, T)) \,|\, \mathscr{F}_s\right\}$, is given by

$$
\begin{aligned}
&E\left\{\log(p(t_1, T))\log(p(t_2, T)) \,|\, \mathscr{F}_s\right\} \\
&= E\left\{(A(t_1, T) - B(t_1, T)r(t_1))(A(t_2, T) - B(t_2, T)r(t_2)) \,|\, \mathscr{F}_s\right\} \\
&= A(t_1, T)A(t_2, T) - A(t_1, T)B(t_2, T)E\left\{r(t_2) \,|\, \mathscr{F}_s\right\} \\
&\quad - B(t_1, T)A(t_2, T)E\left\{r(t_1) \,|\, \mathscr{F}_s\right\} + B(t_1, T)B(t_2, T)E\left\{r(t_1)r(t_2) \,|\, \mathscr{F}_s\right\} \\
&= A(t_1, T)A(t_2, T) - A(t_1, T)B(t_2, T)\mu(t_2, T) \\
&\quad - B(t_1, T)A(t_2, T)\mu(t_1, T) + B(t_1, T)B(t_2, T)E\left\{r(t_1)r(t_2) \,|\, \mathscr{F}_s\right\}.
\end{aligned} \tag{8}
$$

Using the Vasicek SDE solution Eq. (2), with $s = 0$, the term $E\{r(t_1)r(t_2)\,|\,\mathscr{F}_s\}$ is given by

$$
\begin{aligned}
&E\left\{r(t_1)r(t_2)\,|\,\mathscr{F}_s\right\} \\
&= E\left\{\left(r_0 e^{-kt_1} + \theta\left(1 - e^{-kt_1}\right) + \sigma e^{-kt_1}\int_0^{t_1} e^{ku}dW(u)\right)\right. \\
&\qquad \left. \times \left(r_0 e^{-kt_2} + \theta\left(1 - e^{-kt_2}\right) + \sigma e^{-kt_2}\int_0^{t_2} e^{ku}dW(u)\right)\right\} \\
&= r_0^2 e^{-k(t1+t2)} + r_0 e^{-kt_1}\theta\left(1 - e^{-kt_2}\right) \\
&\quad + \theta\left(1 - e^{-kt_1}\right)r_0 e^{-kt_2} + \theta^2\left(1 - e^{-kt_1}\right)\left(1 - e^{-kt_2}\right) \\
&\quad + \sigma^2 e^{-k(t1+t2)}E\left\{\int_0^{t_1} e^{ku}dW(u)\int_0^{t_2} e^{ku}dW(u)\right\}.
\end{aligned}
\tag{9}
$$

In order to compute $E\{\int_0^{t_1} e^{ku}dW(u)\int_0^{t_2} e^{ku}dW(u)\}$, we first consider $t_1 < t_2$. In this case, we have

$$
\begin{aligned}
&E\left\{\int_0^{t_1} e^{ku}dW(u)\int_0^{t_2} e^{ku}dW(u)\right\} \\
&= E\left\{\left(\int_0^{t_1} e^{ku}dW(u)\right)\left(\int_0^{t_1} e^{ku}dW(u) + \int_{t_1}^{t_2} e^{ku}dW(u)\right)\right\} \\
&= E\left\{\left(\int_0^{t_1} e^{ku}dW(u)\right)^2\right\}.
\end{aligned}
\tag{10}
$$

In case $t_2 < t_1$, we have

$$
\begin{aligned}
&E\left\{\int_0^{t_1} e^{ku}dW(u)\int_0^{t_2} e^{ku}dW(u)\right\} \\
&= E\left\{\left(\int_0^{t_2} e^{ku}dW(u) + \int_{t_2}^{t_1} e^{ku}dW(u)\right)\left(\int_0^{t_2} e^{ku}dW(u)\right)\right\} \\
&= E\left\{\left(\int_0^{t_2} e^{ku}dW(u)\right)^2\right\}.
\end{aligned}
\tag{11}
$$

Given Eqs. (10) and (11), we get

$$
E\left\{\int_0^{t_1} e^{ku}dW(u)\int_0^{t_2} e^{ku}dW(u)\right\} = E\left\{\left(\int_0^{min(t_1,t_2)} e^{ku}dW(u)\right)^2\right\}.
$$

Finally, using Itô isometry

$$
\begin{aligned}
E\left\{\int_0^{t_1} e^{ku}dW(u)\int_0^{t_2} e^{ku}dW(u)\right\} &= E\left\{\left(\int_0^{min(t_1,t_2)} e^{ku}dW(u)\right)^2\right\} \\
&= \int_0^{min(t_1,t_2)} E\left\{\left(e^{ku}\right)^2\right\} du
\end{aligned}
$$

$$= \int_0^{min(t_1, t_2)} e^{2ku} du$$

$$= \frac{1}{2k} \left(e^{2k min(t_1, t_2)} - 1 \right). \tag{12}$$

Using Eqs. (6)–(9) and (12), the covariance function $cov(t_1, t_2, T)$ of $\log(p(t, T))$ is given by

$$cov(t_1, t_2, T) = \frac{1}{2k^3} e^{-k(2T + t1 + t2)} \left(e^{2k min(t1, t2)} - 1 \right)$$

$$\left(e^{kT} - e^{kt1} \right) \left(e^{kT} - e^{kt2} \right) \sigma^2. \tag{13}$$

3. Gaussian Processes for Machine Learning

The goal of Gaussian processes for machine learning is to find the nonlinear unknown mapping $y = f(\mathbf{x})$, from data $(\mathbf{X}, \mathbf{y})$, using Gaussian distributions over functions[1] (Rasmussen and Williams, 2005)

$$\mathscr{GP} \sim \mathcal{N}(\mu(\mathbf{x}), cov(\mathbf{x_1}, \mathbf{x_2})).$$

The pair $(\mathbf{X}, \mathbf{y})$ is the training set. The matrix $\mathbf{X}$ collects a set of n vectors $\mathbf{x}$ where the value $y = f(\mathbf{x})$ was observed. The corresponding y values are collected in vector $\mathbf{y}$.

The set of vectors $\mathbf{x}^\star$ where the values $y^\star = f(\mathbf{x}^\star)$ were not observed, is collected in matrix $\mathbf{X}^\star$. The matrix $\mathbf{X}^\star$ is the test set.

Under the Vasicek interest rate model the zero coupon bonds log prices $\log(p(t, T))$ are normal

$$\mathscr{GP} \sim \mathcal{N}(\mu(t, T), cov(t_1, t_2, T)),$$

where $\mu(t, T)$ is given by Eq. (6) and $cov(t_1, t_2, T)$ is given by Eq. (13).

Since T, the bond maturity, is a bond feature, in this case the mapping we are interested in is the scalar mapping

$$y = f(t),$$

where y stands for the zero coupon bonds log prices. This reduces the training set to the pair of vectors $(\mathbf{t}, \mathbf{y})$, and the test set to vector $\mathbf{t}^\star$.

Since the process is Gaussian (Rasmussen and Williams, 2005),

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}^\star \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu}^\star \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{K}_\star \\ \mathbf{K}_\star^T & \mathbf{K}_{\star\star} \end{bmatrix} \right)$$

and

$$p(\mathbf{y}^\star \,|\, \mathbf{t}^\star, \mathbf{t}, \mathbf{y}) \sim \mathcal{N} \left(\boldsymbol{\mu}^\star + \mathbf{K}_\star^T \mathbf{K}^{-1}(\mathbf{y} - \boldsymbol{\mu}), \mathbf{K}_{\star\star} - \mathbf{K}_\star^T \mathbf{K}^{-1} \mathbf{K}_\star \right),$$

[1]See Rasmussen (2004) for a short introduction to the Gaussian distributions over functions framework.

where μ and $\mu^{\star}$ are mean vectors of train and test sets, $\mathbf{K}$ is the train set covariance matrix, $\mathbf{K}_{\star}$ the train-test covariance matrix, and $\mathbf{K}_{\star\star}$ the test set covariance matrix.

The conditional distribution

$$p(\mathbf{y}^{\star} \,|\, \mathbf{t}^{\star}, \mathbf{t}, \mathbf{y})$$

corresponds to the posterior process on the data

$$\mathcal{GP}_{\mathcal{D}} \sim \mathcal{N}(m_{\mathcal{D}}(t), cov_{\mathcal{D}}(t_1, t_2)),$$

where

$$m_{\mathcal{D}}(t) = m(t) + \mathbf{K}_{\mathbf{t},t}^T \mathbf{K}^{-1}(\mathbf{y} - \mu) \tag{14}$$

and

$$cov_{\mathcal{D}}(t_1, t_2) = cov(t_1, t_2) - \mathbf{K}_{\mathbf{t},t_1}^T \mathbf{K}^{-1} \mathbf{K}_{\mathbf{t},t_2}, \tag{15}$$

where $\mathbf{K}_{\mathbf{t},t}$ is a correlation vector between every training instant and t.

Equation (14) is the regression function while Eq. (15) is the regression confidence. Equations (14) and (15) are the central equations of Gaussian processes for machine learning regression.

In order to learn the model parameters $\Theta = \{r_0, k, \theta, \sigma\}$ from data, the likelihood of the training data given the parameters (closed form) (Rasmussen, 2004)

$$\begin{aligned} L &= \log p(\mathbf{y} \,|\, \mathbf{t}, \Theta) \\ &= -\frac{1}{2} \log |\mathbf{K}| - \frac{1}{2}(\mathbf{y} - \mu)^T \mathbf{K}^{-1}(\mathbf{y} - \mu) - \frac{n}{2}\log(2\pi) \end{aligned}$$

is maximized, based on the derivatives of L with respect to each of the parameters (closed forms).

Note that, since we want to learn the parameters in the arbitrage free risk neutral measure, the initial interest rate value r_0, is considered a parameter, like k, θ, and σ, to be learned from the zero coupon bond log prices.

Denoting each of the parameters in set Θ by Θ_i, and since

$$\frac{\partial}{\partial \Theta_i} \log |\mathbf{K}| = tr\left(\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \Theta_i}\right)$$

and

$$\frac{\partial}{\partial \Theta_i}\mathbf{K}^{-1} = -\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \Theta_i}\mathbf{K}^{-1}$$

the derivatives $\frac{\partial L}{\partial \Theta_i}$ are given by

$$\begin{aligned} \frac{\partial L}{\partial \Theta_i} = &-\frac{1}{2}tr\left(\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \Theta_i}\right) \\ &+ \frac{1}{2}(\mathbf{y} - \mu)^T \mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \Theta_i}\mathbf{K}^{-1}(\mathbf{y} - \mu) + (\mathbf{y} - \mu)^T \mathbf{K}^{-1}\frac{\partial \mu}{\partial \Theta_i}. \end{aligned}$$

In order to compute the vector of derivatives, $\frac{\partial \boldsymbol{\mu}}{\partial \Theta_i}$, and the matrix of derivatives $\frac{\partial \mathbf{K}}{\partial \Theta_i}$, the derivatives of the mean function $\mu(t, T)$ (Eq. 6), and the derivatives of the covariance function $cov(t_1, t_2, T)$ (Eq. 13) with respect to the parameters are used, namely:

$$\frac{\partial \mu(t, T)}{\partial r_0} = \frac{e^{-kT} - e^{-kt}}{k};$$

$$\frac{\partial \mu(t, T)}{\partial k} = \frac{1}{4k^4} e^{-kT} \left(4k^3 (e^{k(T-t)} t - T)(r_0 - \theta) - 6(e^{kt} - e^{kT})\sigma^2 \right.$$
$$+ 2k(e^{kT}(2t - 2T + 1) + e^{kt}(t - T - 1))\sigma^2$$
$$\left. + k^2(-4r_0 + 4e^{k(T-t)}(r_0 - \theta) + 4\theta + e^{kt} t \sigma^2 - e^{kT} T \sigma^2) \right);$$

$$\frac{\partial \mu(t, T)}{\partial \theta} = \frac{e^{-kt} - e^{-kT} + kt - kT}{k};$$

$$\frac{\partial \mu(t, T)}{\partial \sigma} = \frac{\left(e^{k(t-T)}(k + 2) + 2kT - 2kt - k - 2 \right) \sigma}{2k^3};$$

$$\frac{\partial cov(t_1, t_2, T)}{\partial r_0} = 0;$$

$$\frac{\partial cov(t_1, t_2, T)}{\partial k} = \frac{1}{2k^4} e^{-k(t1+t2+2T)} \left(e^{k(t1+t2)}(3 + 2kT) + e^{2kT}(3 + k(t1 + t2)) \right.$$
$$- e^{k(t1+T)}(3 + k(t2 + T)) - e^{k(t2+T)}(3 + k(t1 + T))$$
$$+ e^{k(t1+2min(t1,t2)+T)}(3 + k(t2 - 2min(t1, t2) + T))$$
$$+ e^{k(t2+2min(t1,t2)+T)}(3 + k(t1 - 2min(t1, t2) + T))$$
$$+ e^{k(t1+t2+2min(t1,t2))}(-3 + 2k(min(t1, t2) - T))$$
$$\left. + e^{2k(min(t1,t2)+T)}(-3 - k(t1 + t2 - 2min(t1, t2))) \right) \sigma^2;$$

$$\frac{\partial cov(t_1, t_2, T)}{\partial \theta} = 0;$$

$$\frac{\partial cov(t_1, t_2, T)}{\partial \sigma} = \frac{1}{k^3} e^{-k(t1+t2+2T)} \left(e^{kT} - e^{kt1} \right) \left(e^{kT} - e^{kt2} \right)$$
$$\times \left(e^{2kmin(t1,t2)} - 1 \right) \sigma.$$

4. Simulation Results

In order to test the proposed calibration method, we used Eqs. (2) and (5), with fixed parameters values, to simulate 1,000 sequences of zero coupon bond log prices.

The parameters values used were: initial interest rate $r_0 = 0.5$; mean interest rate level $\theta = 0.1$; mean reversion velocity $k = 2$; and volatility $\sigma = 0.2$.

We considered the zero coupon bond maturity of one year, $T = 1$, and simulated one year daily prices sequences by considering 260 prices per sequence (5 working days prices per week, 52 weeks per year).

Figure 1 illustrates a simulated sequence of zero coupon bond log prices, as well as the mean and variance functions.
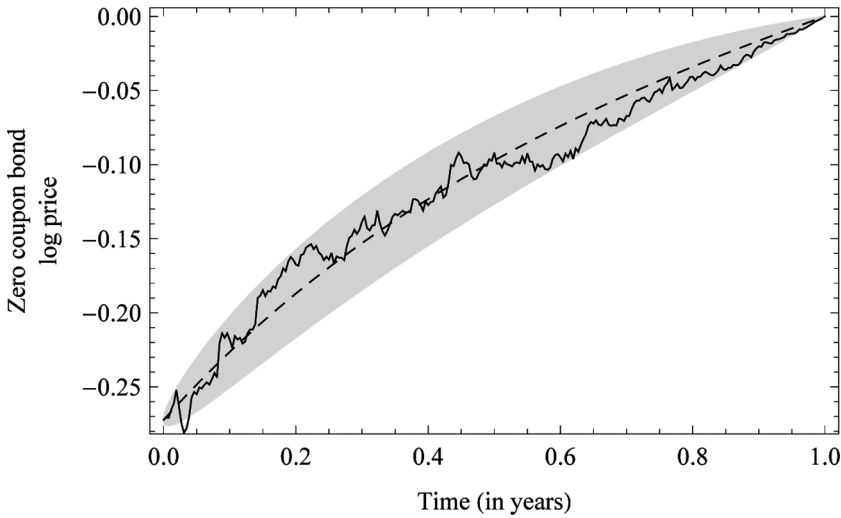
Figure 1. Zero coupon bond log prices simulated sequence (solid black), mean (dashed black), and two standard deviations interval (light gray).

We applied the calibration procedure by maximizing the likelihood of each one of the zero coupon bond log prices sequences, using Wolfram Mathematica 7 (Wolfram Research, 2009) conjugate gradients implementation with default configuration parameters.

Figure 2 illustrates the 50 bins parameters histograms obtained from the 1,000 calibrations performed, and Table 1 shows the corresponding mean, standard deviation and 95% confidence intervals.
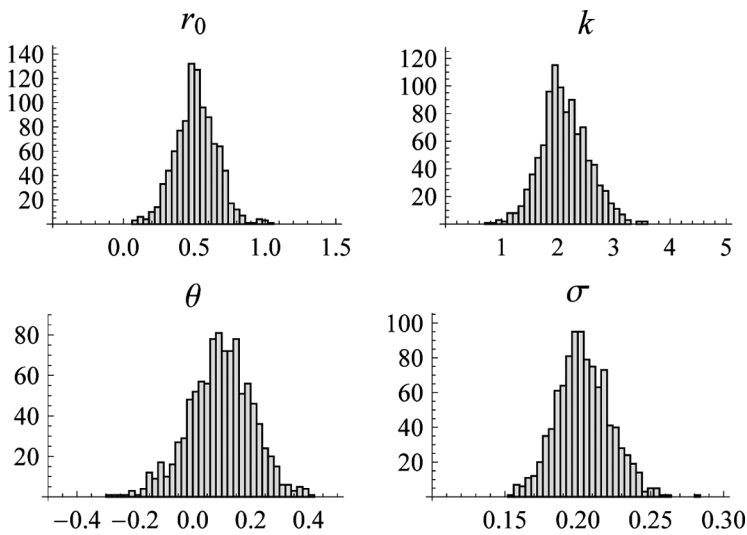

Figure 2. Learned parameters 50 bins histograms.

Table 1
Parameters r_0, k, θ, and σ, 1000 calibrations mean, standard deviation and 95% confidence interval

Parameter	Value	Mean	Std. Dev.	95% CI		
r_0	0.5	0.527	0.482	−0.418	to	1.472
k	2.0	2.098	0.855	0.419	to	3.776
θ	0.1	0.083	0.443	−0.786	to	0.952
σ	0.2	0.203	0.039	0.126	to	0.280

As it can be observed, all parameters 95% confidence intervals contain the fixed parameter value used. Despite this fact, results in Table 1 also show that while the 1,000 calibrations sample mean estimations, given by the parameters means showed in third column of Table 1, are quite accurate, the confidence intervals for each individual calibration are quite large. Therefore calibration with the proposed method should include as much zero coupon bond prices sequences as possible. Care should be taken when calibrating the model with just one price sequence, because the estimation errors can be quite large.

5. Calibration to Real Data

In real life there are only a few zero coupon bonds and their great majority is traded over the counter (OTC). Those bonds are traded between two market players instead of openly being traded in an exchange and data is scarce. Zero coupon bond prices can also be implicitly extracted from the prices of other fixed income products, but in this case the prices would be theoretical and dependent on the assumptions underlying the term structure fitting.

We had, thus, two choices: either use actual market data on one of the few zero coupon bonds actually traded on the market and rely on OTC quotes, or use the price on other fixed income products and extract from those theoretical zero coupon bond prices. We chose the first approach. At the time we looked for data there was a two-year maturity zero coupon in the OTC market live for approximately one year. We took this bond and used a quote service that delivers market prices aggregated from different dealers responsible for trading (market makers) this particular bond.

Table 2 shows the parameters learned and standard deviations obtained from the square root of the diagonal of the inverse Fisher information matrix, as a measure of the calibration errors. Figure 3 illustrates the corresponding mean and variance functions, along with the log prices sequence itself. It should be noted that the learned parameters along with Eqs. (14) and (15) allow the plot in Fig. 3 of the learned mean and the learned two standard deviations interval, not only in the

Table 2
Learned parameters r_0, k, θ, and σ, for a real, two-year maturity, zero coupon bond, calibrated with approximately one year of available prices

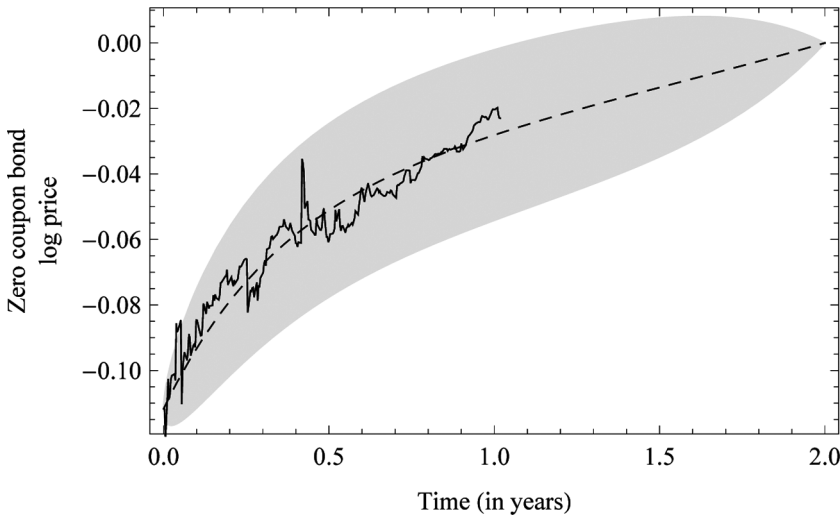|	r_0	k	θ	σ
Learned	0.212	2.925	0.025	0.195
Std. Dev.	0.110	1.897	0.022	0.119

Figure 3. Real, two-year maturity, zero coupon bond log prices sequence (solid black), learned mean (dashed black), and learned two standard deviations interval (light gray).

time interval where there are available prices, but for the all t between the initial time $t = 0$ and maturity $t = T$.

As it can be observed in Fig. 3, the mean and variance functions adjust quite well to the particular price sequence used. However, the large standard deviations obtained do confirm that the errors in the learned parameters can be quite large when calibrating with just one price sequence.

6. Conclusions

In this article we presented a calibration procedure of the Vasicek interest rate model under the risk neutral measure by learning the model parameters using Gaussian processes for machine learning regression with zero coupon bond log prices mean and covariance functions computed analytically.

Compared with other calibration procedures, in this one all the parameters are obtained in the arbitrage free risk neutral measure and the only prices needed for calibration are zero coupon bond prices. On the other hand, this calibration procedure makes no discrete model approximation and makes no simplifications that possibly allow arbitrage opportunities, as it happens when calibrating using interest rate trees (Brigo and Mercurio, 2006). It also does not require the establishment of an objective measure dynamics for the interest rate, as it happens when applying the classical maximum likelihood estimation directly to the interest rate (Brigo and Mercurio, 2006).

References

Björk, T. (2004). *Arbitrage Theory in Continuous Time*. 2nd ed. Oxford: Oxford University Press.

Brigo, D., Mercurio, F. (2006). *Interest Rate Models – Theory and Practice: With Smile, Inflation and Credit*. 2nd ed. Heidelberg: Springer.

Rasmussen, C. E. (2004). Gaussian processes in machine learning. In: Bosquet, O., von Luxburg, U., Rätsch, G., eds. *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Revised Lectures*. Springer-Verlag volume 3176 of *Lecture Notes in Computer Science*, pp. 63–71.

Rasmussen, C. E., Williams, C. K. I. (2005). *Gaussian Processes for Machine Learning*. Cambridge, MA: The MIT Press.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* 5:177–188.

Wolfram Research, I. (2009). *Mathematica Edition: Version* 7.01.0.