Calibration of Interest Rates

Jakub Černý

Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic.

Abstract. In this contribution we study calibration methods of interest rate models. First, we assume that model parameters are constant and can be estimated by the maximum likelihood estimation or yield curve fitting methods. Next, we suppose that model parameters are random variables with their prior distributions. We present Markov Chain Monte Carlo algorithm to generate from posterior distribution using the Bayes theorem. Different methods of calibration based on real data are then applied on well-known Vašíček model with constant volatility.

Introduction

Every theoretical model, in our case financial model, has parameters that should be estimated. The parameter estimation based on a real data is called model calibration or calibrating model on a real data. The main question is which calibration method should we use. We may choose from wide range of statistical methods. From simpler and most used ones like the least squares estimation or maximum likelihood estimation. Or more sofisticated ones like the general method of moments or Markov Chain Monte Carlo (MCMC) algorithms. The MCMC methods in finance were studied in *Ratchev et al.* [2008], *Witzany* [2011] or *Eraker* [2001]. We were inspired by *Johannes and Polson* [2009] where an application on interest rate models is outlined.

The MCMC algorithms are based on Bayesian approach to parameter estimation. In Bayesian approach we assume that the parameter θ is a random variable, in contrast to classic parametric approach where $\theta \in \Theta \subset \mathcal{B}(\mathbb{R}^d)$. Cornerstone of Bayesian approach is the Bayesian theorem

Theorem 1 *Conditional density function $\pi(\theta|\mathbf{x})$ of vector θ for given vector $\mathbf{x}$ is equal to*

$$\pi(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int_\Theta f(\mathbf{x}|\theta)\pi(\theta)d\lambda(\theta)} \cdot \mathbb{1}_{\{\int_\Theta f(\mathbf{x}|\theta)\pi(\theta)d\lambda(\theta)\neq 0\}}$$

regarding Lebesgue measure λ.

The density function $\pi(\theta)$ is called the prior density and $\pi(\theta|\mathbf{x})$ is called the posterior density of θ. The function $f(\mathbf{x}|\theta)$ is called the likelihood or the likelihood function.

The MCMC algorithms are useful when we want to generate from complicated distributions. Idea of MCMC algorithms is a construction of Markov chain with stationary distribution same as the distribution which we search for. After a lot of steps we get a sample from that distribution.

One of the MCMC algorithms which we will use is called the Metropolis-Hastings algorithm. In the Metropolis-Hastings algorithm we assume that the limit distribution π has density f regarding σ-finite measure μ on a measurable space $(X, \mathcal{S})$ and we denote $X^+ = \{x \in X, f(x) > 0\}$. Let Q be the markov kernel on $(X, \mathcal{S})$, $Q(x, dy) = q(x, y)\mu(dy)$ and $Q(x, X^+) = 1$. The density q is called the proposal density. Then we can define the probability of proposal acceptance as

$$\alpha(x, y) = 1 \cdot \mathbb{1}_{\{f(x)q(x,y)\leq 0\}} + \min\left\{\frac{f(y)q(y,x)}{f(x)q(x,y)}, 1\right\} \cdot \mathbb{1}_{\{f(x)q(x,y)>0\}}$$

Algorithm.

1. Assign initial values $x^{(0)} \in X^+$ arbitralily and set $j = 0$.

2. Generate y from distribution $Q(x^{(j)}, \cdot)$. The proposal y is accepted with the probability $\alpha(x^{(j)}, y)$ then $x^{(j+1)} = y$. With probability $1 - \alpha(x^{(j)}, y)$ is the proposal y rejected and $x^{(j+1)} = x^{(j)}$.

3. If $j + 1 < J$ then set $j = j + 1$ and continue on step 2 else end algorithm.

The existence of a limit (resp. stationary) distribution of the Metropolis-Hastings algorithm is in detail proven in *Pawlas* [2007].

Vašíček model

Vašíček model is an Ornstein-Uhlenbeck process with constant parameters. The stochastic differential equation (SDE) for short rate dynamics under the risk-neutral measure Q is following

$$dr(t) = a[b - r(t)]dt + \sigma dW^Q(t), \tag{1}$$

where a, b, σ are constants. The Vašíček model is the mean reversion model with the mean reverting value equal to b and the speed of reversion represented by a. The equation for short rate $r(t)$ calculated from (1) using Itô lemma is

$$r(t) = r(s)e^{-a(t-s)} + b\left(1 - e^{-a(t-s)}\right) + \sigma \int_s^t e^{-a(t-u)} dW^Q(u),$$

assuming $s < t$ and $r(s)$ is known.

It does not matter whether the measure is risk-neutral or real world if the market price of risk λ is constant. The SDE remains in the same form with modified parameter b, $b^* = b - \lambda\sigma$, and regarding the real world measure P instead of Q (for more details see *Brigo and Mercurio* [2006]).

The Vašíček model has an affine time structure, i.e. the value of zero coupon bond $P(t, T)$ at time t with maturity at time T under the risk neutral measure Q is in form

$$P(t, T) = \mathbb{E}^Q \left[\exp\left\{ -\int_t^T r(u)du \right\} \right] = \exp\{A(t, T) - B(t, T)r(t)\},$$

where $A(t, T)$ and $B(t, T)$ are deterministic functions

$$\begin{aligned} B(t, T) &= \frac{1 - e^{-a(T-t)}}{a}, \\ A(t, T) &= \left(b - \frac{\sigma^2}{2a^2}\right)[B(t, T) - (T - t)] - \frac{\sigma^2 B^2(t, T)}{4a}. \end{aligned}$$

The yield of zero coupon bond is

$$y(t, T) = -\frac{\log P(t, T)}{T - t} = \frac{1}{T - t}(B(t, T)r(t) - A(t, T)).$$

For evaluation of the zero coupon bond price or the zero coupon bond yield under the real world measure should be used modified parameter b^* when the market price of risk is constant (see *Málek* [2005]).

Calibration

The short rate r and the yield of the zero coupon bond y are random processes. After discretization we will get a hierarchical model in form

$$\begin{aligned} y(t,T_i) &= \beta_0+\beta_1 r(t)+\sigma^y\varepsilon_t^y, \quad i=1,\ldots,n \\ r(t) &= \alpha_0+\alpha_1 r(t-1)+\sigma^r\varepsilon_t^r, \end{aligned}$$

where

$$\begin{aligned} \beta_0 &= \frac{A(t,T_i)}{t-T_i}, & \alpha_0 &= b(1-e^{-a}), \\ \beta_1 &= \frac{B(t,T_i)}{T_i-t}, & \alpha_1 &= e^{-a}. \end{aligned}$$

Volatility σ^r is constant, a, b are random variables, and $\varepsilon_t^r|r_0,\ldots,r_{t-1} \sim N(0,1)$. Vector $\varepsilon_t^y|y(0,T_i),\ldots,y(t-1,T_i) \sim N(\mathbf{0},\mathbb{I})$ and σ^y is a variance matrix.

First we need to set a prior distribution of parameters α_0 and α_1, i.e. $\boldsymbol{\alpha}=(\alpha_0,\,\alpha_1)^T$. We will use maximum likelihood estimation $\widehat{\boldsymbol{\alpha}}=(\widehat{\alpha_0},\,\widehat{\alpha_1})^T=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{r}$, where

$$\mathbf{X}=\begin{pmatrix} 1, & 1, & \ldots, & 1 \\ r_0, & r_1, & \ldots, & r_{n-1}\end{pmatrix}^T, \qquad \boldsymbol{r}=(r_1,r_2,\ldots,r_n)^T$$

$$\widehat{\alpha_0}=\frac{1}{n}\sum_{i=1}^n r_i-\widehat{\alpha_1}\frac{1}{n}\sum_{i=1}^n r_{i-1}, \qquad \widehat{\alpha_1}=\frac{n\sum\limits_{i=1}^n r_i r_{i-1}-\sum\limits_{i=1}^n r_i\sum\limits_{i=1}^n r_{i-1}}{n\sum\limits_{i=1}^n r_{i-1}^2-\left(\sum\limits_{i=1}^n r_{i-1}\right)^2}$$

Then the prior distribution of $\boldsymbol{\alpha}$ is set as $N(\widehat{\boldsymbol{\alpha}},\underbrace{(\sigma^r)^2(\mathbf{X}^T\mathbf{X})^{-1}}_{\boldsymbol{\Sigma}})$.

Our goal is to generate from the posterior distribution $\pi(\alpha_0,\alpha_1|\boldsymbol{r},\boldsymbol{y})$. We will use the proposal distribution $\pi(\alpha_0,\alpha_1|\boldsymbol{r}):=\pi(\alpha_0,\alpha_1)$ computed from the likelihood function $f(\boldsymbol{r}|\alpha_0,\alpha_1)$ and the density of the prior distribution $p(\alpha_0,\alpha_1)$

$$\begin{aligned} f(\boldsymbol{r}|\alpha_0,\alpha_1) &= \prod_{i=0}^{n-1}\frac{1}{\sqrt{2\pi}\sigma^r}\exp\left\{-\frac{(r_{i+1}-\alpha_0-\alpha_1 r_i)^2}{2\,(\sigma^r)^2}\right\}, \\ p(\alpha_0,\alpha_1) &= \frac{1}{2\pi}|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{\alpha}-\widehat{\boldsymbol{\alpha}})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\alpha}-\widehat{\boldsymbol{\alpha}})\right\}, \\ \pi(\alpha_0,\alpha_1|\boldsymbol{r}) &\propto f(\boldsymbol{r}|\alpha_0,\alpha_1)p(\alpha_0,\alpha_1) \end{aligned}$$

From *Ratchev et al.* [2008] we know that the density $\pi(\alpha_0,\alpha_1)$ corresponds to a normal distribution and we set up proposal density parameters as $N((\boldsymbol{\Sigma}^{-1}+\mathbf{X}^T\mathbf{X})^{-1}(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\alpha}}+\mathbf{X}^T\boldsymbol{r}),(\boldsymbol{\Sigma}^{-1}+(\sigma^r)^{-2}\mathbf{X}^T\mathbf{X})^{-1})$.

We will use the Metropolis-Hastings algorithm to calibrate the Vašíček model

1. Set an initial vector of values $(\alpha_0^{(0)},\alpha_1^{(0)})$, $j=0$.

2. Generate $(\alpha_0^{(j+1)},\alpha_1^{(j+1)})$ from $\pi(\alpha_0^{(j)},\alpha_1^{(j)})$.

3. Accept $(\alpha_0^{(j+1)}, \alpha_1^{(j+1)})$ with the proposal acceptance probability α where

$$\alpha(\alpha_0^{(j+1)}, \alpha_1^{(j+1)}) = \min\left\{ \frac{f(\boldsymbol{y}|\alpha_0^{(j+1)}, \alpha_1^{(j+1)}, \boldsymbol{r}) q((\alpha_0^{(j+1)}, \alpha_1^{(j+1)}), (\alpha_0^{(j)}, \alpha_1^{(j)}))}{f(\boldsymbol{y}|\alpha_0^{(j)}, \alpha_1^{(j)}, \boldsymbol{r}) q((\alpha_0^{(j)}, \alpha_1^{(j)}), (\alpha_0^{(j+1)}, \alpha_1^{(j+1)}))}, 1 \right\}$$

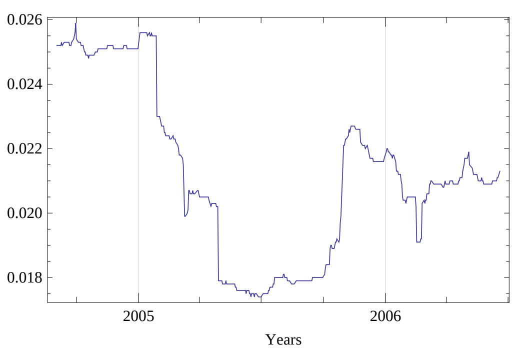
and $\boldsymbol{y}$ is a vector of yields at different maturities (yield curve). Function $f(\boldsymbol{y}|\alpha_0, \alpha_1, \boldsymbol{r})$ is the likelihood function of $\boldsymbol{y}$ with given parameters α_0, α_1 and the short rate $\boldsymbol{r}$

$$f(\boldsymbol{y}|\alpha_0, \alpha_1, \boldsymbol{r}) = \prod_{i=1}^{n} \left(\sqrt{2\pi}\right)^m |\sigma^y|^{-1} \exp\left\{ -\frac{1}{2} (\boldsymbol{y} - \beta_0 - \beta_1 \boldsymbol{r})^T (\sigma^y)^{-1} (\boldsymbol{y} - \beta_0 - \beta_1 \boldsymbol{r}) \right\},$$

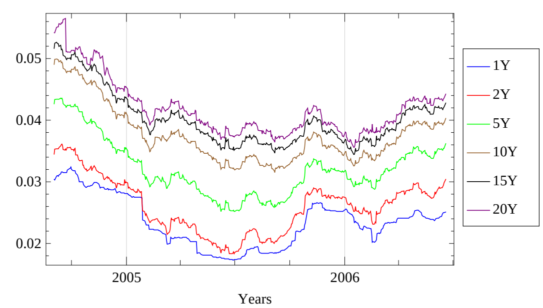
where $\boldsymbol{y} = \{y\}_{i,j}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, m is the number of rates forming yield curve and n is the number of historical yield curve observations. Values of parameters (β_0, β_1) are obtained from the accepted values (α_0, α_1).

Application

To calibrate the Vašíček model we will use swap rates with maturities 1Y-10Y, 12Y, 15Y, 20Y and as a short rate we will use 3M PRIBOR. Both observed from 2.9.2004 to 19.6.2006. Zero coupon bond yields are easily obtained from swap rates by bootstrapping (see *Hull* [2008]). Figures below show historical development of the short rate, i.e. 3M PRIBOR, and zero coupon bond yields.



(a) Development of 3M PRIBOR



(b) Development of the zero coupon bond yields with maturities 1Y,2Y,5Y,10Y,15Y and 20Y

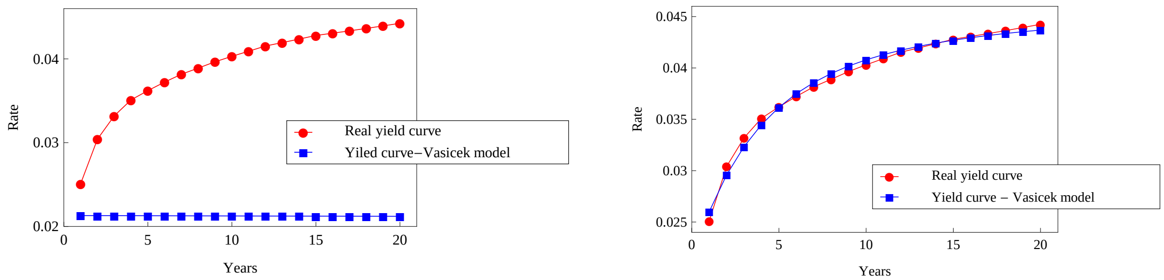
Figure 1.: Development of 3M PRIBOR and zero-coupon bond yields from 2.9.2004 to 19.6.2006

We will simplify our model and fix the historical volatilities as constants. The standard deviation of absolute short rate yields is $\sigma^r = 0.000232074$. The covariance matrix of zero coupon bond yields is large so we will show just a part of it

$$\sigma^y = \begin{pmatrix} 0.000016739 & 0.0000174346 & \ldots & 0.0000151689 \\ 0.0000174346 & 0.0000191944 & \ldots & 0.0000166125 \\ \vdots & & \ddots & \vdots \\ 0.0000151689 & 0.0000166125 & \ldots & 0.0000208994 \end{pmatrix}$$

Correlations between rates are very high (not less than 81%).

First, we will show the MLE based only on historical observations of 3M PRIBOR and the yield curve fitting method. Although the MLE is a correct method for parameters estimation, interest rate models should also fit the yield curve. From following figures we may see that the model calibrated by MLE based on the historical data is flat because MLE does not take into account yield curve data, only short rate, i.e. 3M PRIBOR. MLE yield curve lies under the real yield curve because of the historical development of 3M PRIBOR (see Figure 1). Yield curve fitting method fits the yield curve but not precisely as we can see from the figure below. Description of both methods can be found in *Brigo and Mercurio* [2006].



(a) Yield curve estimated by MLE ($a = 0.0056$, $b = 0.0196$)

(b) Yield curve estimated by yield curve fitting ($a = 0.4213$, $b = 0.0467$)

Figure 2.: Yield curves estimated by MLE and yield curve fitting

We have generated the posterior distribution (10 000 samples) from which we want to generate parameters regarding the data. The following pictures show the convergence of parameters a and b calculated from α_0 and α_1.

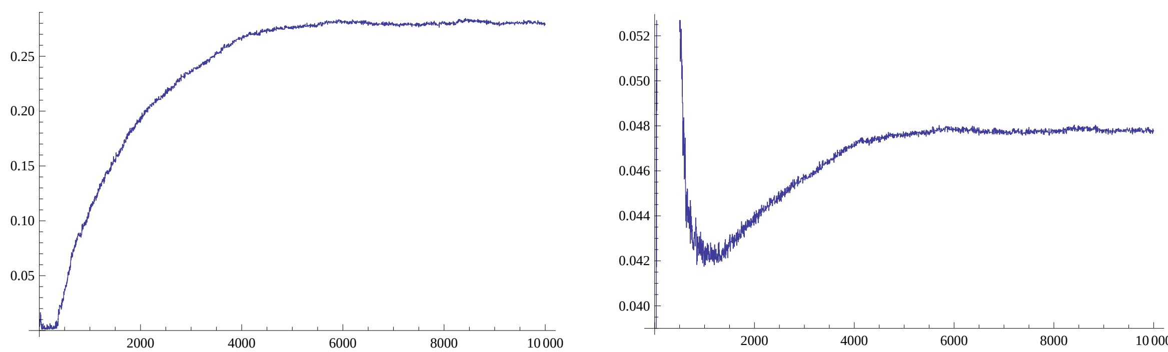


Figure 3.: Convergence of the parameter a (left figure) and b (right figure)

From the previous pictures can be seen that burn-in will be about first 4000 samples. Burn-in has been set up to 5000 samples and adjusted distributions of parameters are on the following pictures.

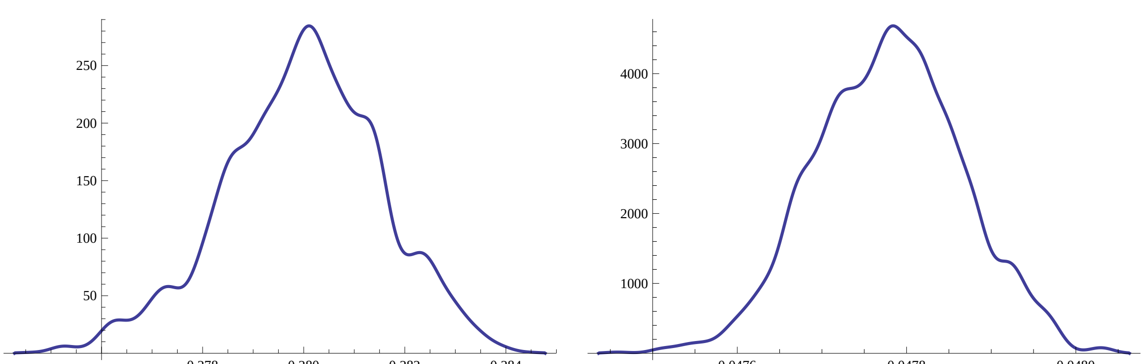


Figure 4.: Distribution of the parameter a (left figure) and b (right figure) after burn-in

Empirical characteristics of the parameters distribution are below. We can see that the variance after burn-in is very low and the mean is close to the yield curve fitting method results.

$$\mathrm{mean}(a,b) \doteq (0.280, 0.0478), \quad \mathrm{var}(a,b) \doteq \begin{pmatrix} 2.42\times10^{-6} & 5.18\times10^{-8} \\ 5.18\times10^{-8} & 7.43\times10^{-9} \end{pmatrix}$$

The advantage of using the Metropolis-Hastings algorithm is inclusion of all available short rate and yield curve data. This is most of the information we can use to calibrate the interest rate model. The yield curve using the mean of estimated parameters a and b by MCMC (dashed line) and present yield curve (full line) are shown in the following figure.

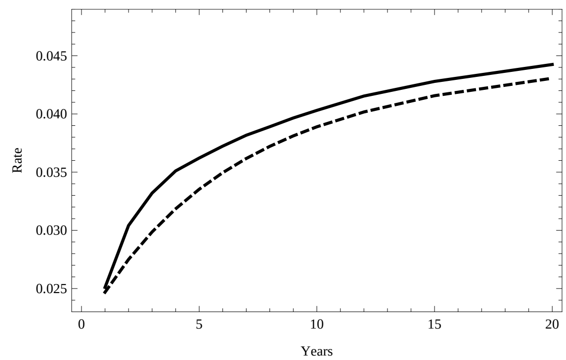


Figure 5.: Present yield curve observed from the real data (full line) and the yield curve based on mean of simulated parameters (dashed line)

Underestimation of the present yield curve is caused by historical development of zero coupon bond yields which were most of the time lower than the yield at 19.6.2006 (see Figure 1).

References

Brigo D., Mercurio F.: Interest Rate Model - Theory and Practice, Springer Finance, Berlin, 2006.

Eraker B.: MCMC Analysis of Diffusion Models with Application to Finance, *Journal of Business & Economic Statistics*, vol. 19, no. 2, 177-191, 2001.

Hull J. C.: Options, Futures, and Other Derivatives, 7th Edition, Pearson Prentice Hall, Englewood Cliffs, 2008.

Johannes M., Polson N.: MCMC Methods for Financial Econometrics, in Handbook of Financial Econometrics, Volume 2 (eds Ait-Sahalia and L.P. Hansen), North Holland, pp. 1-72, 2009.

Málek J.: Dynamika úrokových měr a úrokové deriváty, EKOPRESS, Praha, 2005.

Pawlas Z.: MCMC algoritmy, Lecture notes for MCMC algorithm lectures, Charles University, available at `http://www.karlin.mff.cuni.cz/~pawlas/2007/STP139/mcmc.pdf`, pp.11.

Ratchev S.T., Hsu J.S., Bagasheva B.S., Fabozzi F.J.: Bayesian Methods in Finance, The Frank J. Fabozzi Series, Wiley, 2008.

Witzany J.: Estimating Correlated Jumps and Stochastic Volatilities, IES Working Paper, Prague, No. 35., pp. 33, 2011.