Proceedings of the International MultiConference of Engineers and Computer Scientists 2011 Vol II,
IMECS 2011, March 16 - 18, 2011, Hong Kong

Application of MCMC Algorithm in Interest Rate Modeling

Xiaoxia Feng and Dejun Xie

***Abstract*—Interest rate modeling is a challenging but important problem in financial econometrics. This work is concerned with the parameter estimation of the short term interest models. In light of a recent development in Markov Chain Monte Carlo simulation techniques based on Gibbs sampling, numerical experimentations are carried out for finding an effective and convergent Beyesian estimation scheme. The optimal degree of data augmentation is probed on basis of sensitivity analysis in searching of maximum A-posteriori probability density. Our method is calibrated with both US Treasury bills and basic loan rates from Japanese market.**

***Index Terms*—Bayesian estimation, Gibbs sampler, MCMC method, MAP estimation, Data augumenation.**

I. Introduction

Interest rate is a key variable in economy and financial market. It determines or impacts the values of various financial contracts, options, and derivatives. Given an prescribed interest model, one may rely on both mathematical and simulation techniques to solve for the value of a financial contract. In reality, the value of a financial instrument, an index, or a contract, is often observable from market. What is drawing more attention is the inverse problem, i.e., to select a model, usually stochastic in nature, and to determine the parameters appearing in the model so that market participants may rely on the model in making future decisions. Because of this, there has been incessant literatures on the interest modeling and parameter estimation. For commonly used spot interest rate models, one can refer to [11], for instance.

Theoretical interest rate models, formulated in terms of stochastic differential equations, often assume continuous observations in time. The main merit of continuity is the flexibility for mathematical treatment. Among several types of mean-reverting processes, CIR model ([8]) is one of the most exploited short term rate models in literature. Accordingly, there have been considerable studies concerning the parameter estimation for the model.

Due to the complexity of the transitional probability density of CIR model, it is difficult to apply the maximum likelihood estimation (MLE) method to estimate the parameters. On the other hand, even if the MLE estimators are explicitly known, as is true for the Vasicek model, they are not necessarily unbiased [3]. Thus many approximation methods have been introduced. For example, in [4], [7], and [10], the continuous CIR process is approximated with finite discretizations. In [13],[14], and [5], it is suggested to approximate the drift term, or the diffusion term, or the transitional density function with nonparametric techniques. However, these approximation techniques, though plausible in theory, also have implementation difficulties, as argued in [6], for instance. These challenges may include bias from discrete sampling or slow convergence.

Xiaoxia Feng is with the Department of Mathematics, Xi'an Jiaotong University, Xi'an, 710048, China, e-mail: (x.x.f.09@stu.xjtu.edu.cn).
Dejun Xie is with the Department of Mathematical Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China, e-mail:(Dejun.Xie@xjtlu.edu.cn).

In a recent work ([12]), a Bayesian estimation approach is introduced with Markov Chain Monte Carlo (MCMC) algorithm for estimating the parameters of the CIR model. Gibbs sampler algorithm ([2]) based on Euler-Maruyama discritization is designed to simulate the posterior distribution of the latent data. In the mean time, a genetic algorithm is implemented to achieve the maximum A-posteriori (MAP hereinafter) estimation of the parameters.

However, there are several implementation issues to be answered about the method, including the its performance when applied to different markets. So one aim of our current work is to calibrate the relative convergence of the algorithm in terms of estimating model parameters. And also, we wish to find optimal degree of data augmentation for the MCMC scheme. In the following of this paper, a brief summary of our previous results (see[12]) is firstly provided. Through simulated data using non-central Chi-square approximations, the effects of the number of time intervals and the number of simulated paths on convergence of estimation are analyzed, based on which the optimal degree of data observation is obtained. Numerical results are provided using historical data of US Treasury bills and the basic loan rates from Japanese market.

II. Review of Main Methodology

The CIR short term interest rate process is defined by the stochastic differential equation

$$dy(t) = \{\alpha - \beta y(t)\}dt + \sigma\sqrt{y(t)}dW(t) \qquad (1)$$

where $\{W(t), t \geq 0\}$ is a standard Brownian motion and α, β, $\sigma > 0$ are the constant model parameters. The majors steps of the MCMC based iterative conditional sampling algorithm for parameter estimation are summarized as follows.

A. The Fully Conditional Posterior Distribution

Suppose there are T observations, and M augmented data points between each pair of adjacent observations. Let Y=$(y_1, \cdots, y_T)$ denote the set of all observed historical data and $Y^* = (y_1^*, \cdots, y_{T-1}^*)$ the set of all augmented data, where $y_t^* = \{y_{t,1}^*, \cdots, y_{t,M}^*\}$. According to the general theory of Beyesian estimation (see [1], for instance), we may assume that the priori density function of the parameters is proportional to the inverse of σ. The following are the results about the fully conditional posterior distribution derived in ([12]).

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

For the augmented data Y^*, we have

$$f(y_t^*|Y,\theta) = \sum_{j=0}^{M} f(y_{t,j+1}^* \mid y_{t,j}^*, \theta) \tag{2}$$

where $\theta = (\alpha, \beta, \sigma)$, $y_{t,0}^* = y_t$, $y_{t,M+1}^* = y_{t+1}$ and

$$y_{t,j+1}^*|y_{t,j}^*, \theta \sim N(y_{t,j}^* + (\alpha - \beta y_{t,j}^*)\triangle, \sigma^2 \triangle y_{t,j}^*). \tag{3}$$

For the drift parameters $\psi = (\alpha, \beta)$, we have

$$\psi|Y, Y^*, \sigma^2 \sim N(\mu, \Lambda^{-1}) \tag{4}$$

where $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, $\Lambda = \begin{pmatrix} \frac{\triangle}{\sigma^2}A & -\frac{\triangle}{\sigma^2}n \\ -\frac{\triangle}{\sigma^2}n & \frac{\triangle}{\sigma^2}B \end{pmatrix}$ and

$$\begin{aligned}
\mu_1 &= \frac{\sigma^2}{\triangle}\frac{BC + nD}{AB - n^2}, \\
\mu_2 &= \frac{\sigma^2}{\triangle}\frac{nC + AD}{AB - n^2}, \\
n &= (T-1)(M+1), \\
A &= \sum_{t=1}^{T-1}\sum_{j=0}^{M}\frac{1}{y_{t,j}^*}, \\
B &= \sum_{t=1}^{T-1}\sum_{j=0}^{M} y_{t,j}^* \\
C &= -\sum_{t=1}^{T-1}\sum_{j=0}^{M}\frac{y_{t,j}^* - y_{t,j+1}^*}{y_{t,j}^*}, \\
D &= \sum_{t=1}^{T-1}\sum_{j=0}^{M}(y_{t,j}^* - y_{t,j+1}^*).
\end{aligned}$$

B. The Sampling Algorithm

Step 1: Initialize y_0, θ and use Gibbs sampler to generate the initial values of $y_1^*, \cdots, y_{T-1}^*$.
Step 2: Use Gibbs sampler to
(a) Update α, β from $f(\psi|Y^*, Y, \sigma^2)$ where $y_1^*, \cdots, y_{T-1}^*$ are obtained from the previous iteration.
(b) Update σ^2 from $f(\sigma^2|Y^*, Y, \alpha, \beta)$ where $y_1^*, \cdots, y_{T-1}^*$ are obtained from the previous iteration and α, β are given by (a)
Step 3: Update $y_1^*, \cdots, y_{T-1}^*$ from $f(y_t^*|y_t, \alpha, \beta, \sigma^2)$
Step 4: Repeat Step 2 until the prescribed sampling size N is reached.

C. the posterior density and the MAP estimation

$$p(\sigma^2|Y) = \frac{1}{N - N_0}\sum_{j=N_0+1}^{N} p(\sigma^2|Y, Y_{(j)}^*, \alpha_{(j)}, \beta_{(j)})$$

$$p(\alpha, \beta|Y) = \frac{1}{N - N_0}\sum_{j=N_0+1}^{N} p(\sigma^2|Y, Y_{(j)}^*, \sigma_{(j)}^2)$$

where the subscript j refers to the jth iteration and the first N_0 iterations are cast away in order to mitigate the effects of initial conditions. The MAP estimates of the parameters are the points where $p(\alpha, \beta|Y)$ and $p(\sigma^2|Y)$, respectively, have the maximum values. The genetic algorithm in MATLAB can help to realize the optimization process.

TABLE I
SUMMARY STATISTICS FOR THE PARAMETERS WHEN N IS FIXED

parameters	MAP	correlation coefficient		
M = 3				
α	0.0692	1.0000	0.8010	-0.0368
β	1.3721	0.8010	1.0000	-0.0305
σ^2	0.0906	-0.0368	-0.0305	1.0000
M = 6				
α	0.0583	1.0000	0.8199	-0.0795
β	1.346	0.8199	1.0000	-0.0766
σ^2	0.0668	-0.0795	-0.0766	1.0000
M = 7				
α	0.0506	1.0000	0.8089	-0.0596
β	1.1051	0.8089	1.0000	-0.0507
σ^2	0.0624	-0.0596	-0.0507	1.0000
M = 8				
α	0.0379	1.0000	0.8049	-0.0692
β	0.8748	0.8049	1.0000	-0.0671
σ^2	0.0614	-0.0692	-0.0671	1.0000
M = 10				
α	0.0269	1.0000	0.8201	-0.0856
β	0.6184	0.8201	1.0000	-0.0760
σ^2	0.0527	-0.0856	-0.0760	1.0000
M = 20				
α	0.0265	1.0000	0.8313	-0.1292
β	0.6118	0.8313	1.0000	-0.1274
σ^2	0.0391	-0.1292	-0.1274	1.0000
M = 25				
α	0.0162	1.0000	0.8322	-0.1293
β	0.3764	0.8322	1.0000	-0.1391
σ^2	0.0351	-0.1293	-0.1391	1.0000

III. NUMERICAL EXPERIMENTATION

In this section, we use simulated data to carry out numerical experimentations. To test the robustness of the method, sensitivity analysis is provided for different time intervals and various number of iterations. We take $y_0 = 1$ and obtain 100 observations using a noncentral chi-squared distribution with $\alpha = 0.05, \beta = 1, \sigma = 0.25$. The following Figure 1 gives the time series of the simulated observations. The value of $\triangle^\dagger$ indicating the time interval between two observations is taken to be 1/52 and N_0 is set at 100 for all the computations. The time interval changes with M. We first set N=5100 and run our program for different values of M, which yields different estimates of the parameters. The results of the these estimations are presented in TABLE I.

From the results reported in table 1, we see that the MAP procedure achieves most accurate estimates of the true

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

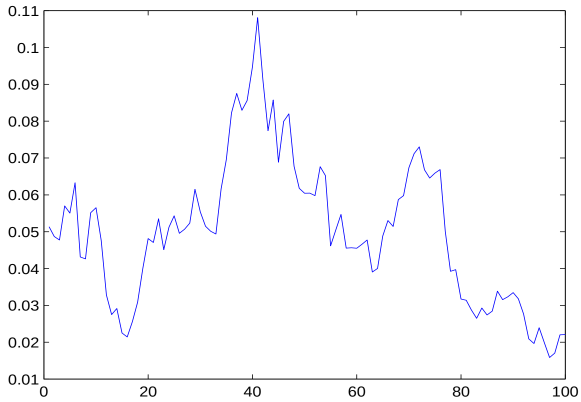

Figure 1:Time series of 100 simulated observations.

TABLE II
SUMMARY STATISTICS FOR THE PARAMETERS WHEN M IS FIXED

N = 4100

parameters	MAP	correlation coefficient		
α	0.0432	1.0000	0.8101	-0.0801
β	1.014	0.8101	1.0000	-0.0716
σ^2	0.0651	-0.0801	-0.0716	1.0000

N = 5100

parameters	MAP	correlation coefficient		
α	0.0506	1.0000	0.8089	-0.0596
β	1.1051	0.8089	1.0000	-0.0507
σ^2	0.0624	-0.0596	-0.0507	1.0000

N = 5600

parameters	MAP	correlation coefficient		
α	0.0408	1.0000	0.8109	-0.0456
β	0.9233	0.8109	1.0000	-0.0443
σ^2	0.0643	-0.0456	-0.0443	1.0000

parameter values when M=7. Other values of M away from 7 tend to generate large bias. Small value of M means large discretization interval, leading to large bias. On the other hand, misrepresentation of true data structure may occur due to the large number of inserted latent data points when M is large. In the above simulated example, M=7 is optimal.

In addition, we find the estimated long-term mean of the CIR model $\frac{\alpha}{\beta}$ is always close to the mean of the true data for different values of M. Another interesting finding is that α and β are always positive correlated while both of them have negative correlations with σ. When we fix M=7, and run the program with different values of N, we find the estimated results are more or less very close to each other, as long as N is large enough. Such results, as shown in TABLE II, tend to suggest that the method outlined in Section II will provide convergent estimations when the number of iteration paths are sufficiently large, provided that an an optimal M is chosen.

IV. Application to Historical Financial Data

We wish to implement the method outlined in Section II using the historical data of US Treasury bills using the method outlined in Section II. Here, we first select 8 years of weekly data of the yields of the US Treasury bills from November 2002 to October 2010. The time span is long enough to accommodate the usual market fluctuations. And it has the same time interval between two observations with the above simulated example. Figure 2 gives the time series of historical yields of the US 6-month Treasury bills used in our example. The iteration results for parameter estimations are plotted in Figure 3. The MAP estimates for the three parameters appearing in the CIR model are, respectively, $\alpha = 7.0660e-4, \beta = 3.5546e-4$, and $\sigma = 0.002$.

To compare, we also apply our method to the basic loan rate from Japanese market. Here we use the 132 monthly observations of Japan's basic loan rates from Jan. 2000 to Dec. 2010. To choose a relatively reasonable value for M, we first compute the time interval between each pair of consecutive observations, size of the data in terms of number of observations, sample expectation, and standard deviation of the data under study. We use the observed mean of data to approximate the initial simulation of $\frac{\alpha}{\beta}$ and use the observed variance to approximate σ, although they are not exactly the same in theory. We simulate the same number of sample points with the real data, and then we apply our method outlined in Section III to get a rational M as our selection for analyzing the historical data. The time series of the historical monthly data of 10 years of Japanese basic loan rate is plotted in Figure 4. According to our approximation, a rational choice of M is 8.

Implementation of our method yields the iteration results for parameter estimations as plotted in Figure 5. The MAP estimates for the parameters are respective $\alpha = 3.0202e-4, \beta = 0.0012$, and $\sigma^2 = 0.01742$.

V. Conclusion

This paper focuses on the numerical simulation and implementation of our previous work on Bayesian estimation of the parameters. Using simulated examples, we have shown that the approach can generate good estimates, although the performance of the algorithm relies on the selection of the number of the inserted points. Moreover, we have proposed

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

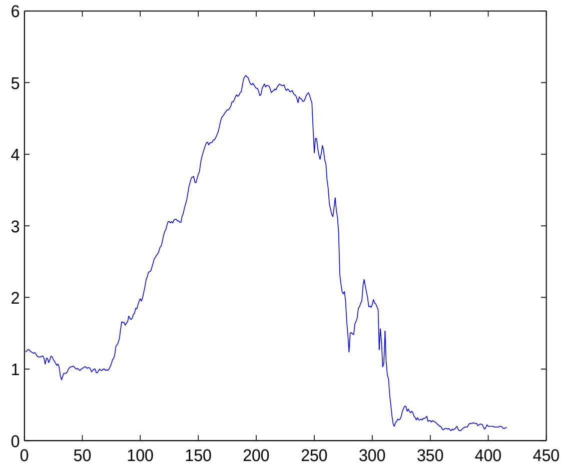

Figure 2: Time series of US 6-month Treasury bills over a duration of 8 years.

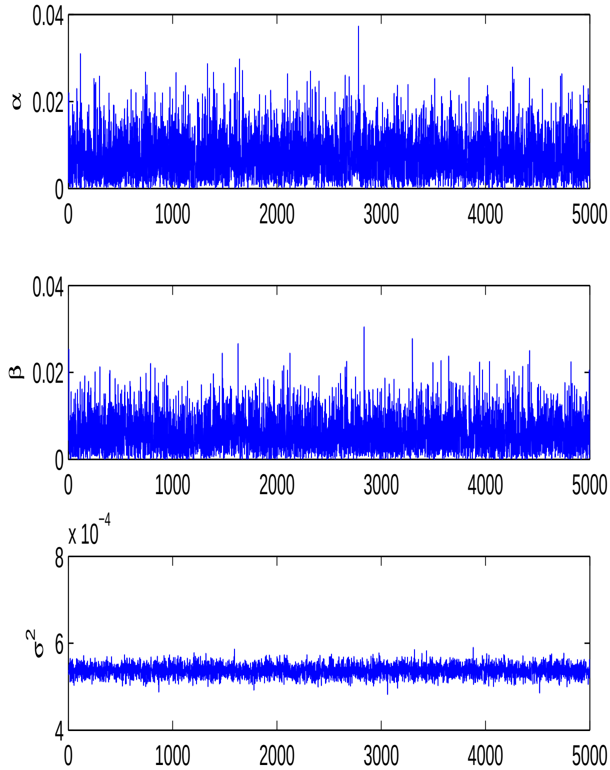

Figure 3: The estimated values of the parameters α, β, σ^2 from the fully conditional posterior distribution in each iteration with 8 years of observations. Here $M = 7, N = 5100$.

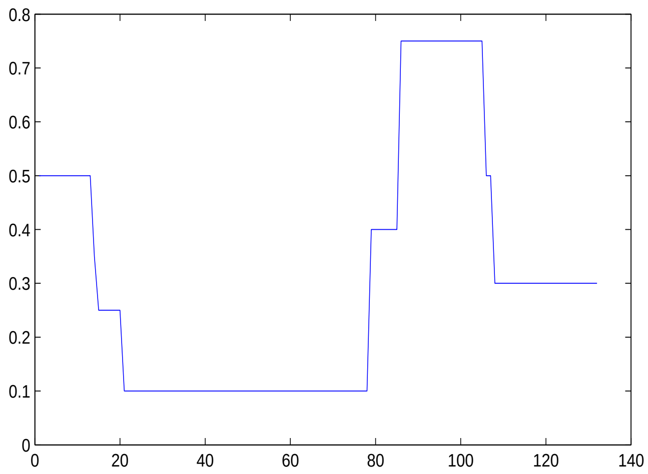

Figure 4:Time series of Japan basic loan rate over a duration of 10 years.

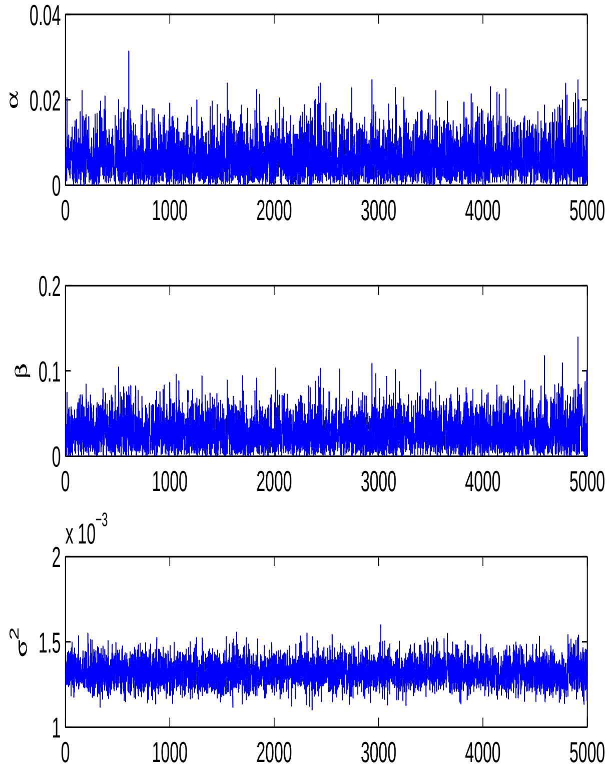

Figure 5: The estimated values of the parameters α, β, σ^2 from the fully conditional posterior distribution in each iteration with 10 years of observations. Here $M = 8, N = 5100$.

a method for finding the acceptable M, based on empirical experiments, and applied it to the real historical data of Japanese basic loan rates. These results provide insightful hints for further studies on how to improve the accuracy and convergence of the method.

References

[1] C. S. Jones, "Bayesian Estimation of Continuous-Time Finance Models," Working paper,1999.
[2] D. Sorensen and D. Gianola, "Likelihood,Bayesian,and MCMC Methods in Quantitative Genetics," New York,Springer,2002.
[3] D. Xie, "Parametric Estimation for Treasury Bills" International Research J. of Finance and Economics, 17, 2008, pp. 27–32.
[4] G. Deelstra and G. Parker, "A Covariance Equivalent Discretization of The CIR Model," 5th AFIR International Colloquium,1995.
[5] G. J. Jiang and J. L. Knight, "A Nonparametric Approach to the Estimation of Diffusion Processes,with an application to a short–term interest rate model," Economic theory, vol.13,1997,pp.615–645.
[6] G. O. Robert and O. Stramer, "On Inference for Parcially Observed Nonlinear Diffusion Models using the Metropolis-Hastings Algorithm," Biometrika,vol.88,2001,pp.603–621.
[7] I. Shoji and T. Ozaki, "Estimation for Nonlinear Stochastic Differencial Equations by a Local Linearization Method," Stochastic Analysis and applications, vol.16,1998,pp.733–752.
[8] J. C. Cox,J. E. Ingersoll and S. A. Ross, " An Intertemporalgenaral Equilibrium Model of Asset Prices," Econometrica, vol.53,1985,pp.363–384.
[9] J. S. Liu, "Monte Carlo Strategies in Scientific Computing," Springer Series in Statistics, Springer New York, 2001.
[10] J. Yu and P. C. B. Phillips, "A Gaussian Approach for Estimating Continuouse Time Models of Short Term Interest Rate," The Econometrics Journal,vol.4,2001,pp.210–224.
[11] P. Willmott, "Derivatives, the theory and practice of financial engineering," John Wiley & Sons, New York, 1999.
[12] X. Feng and D. Xie," Beyesian Estimation of CIR Model," XJTLU Working Paper.
[13] Y. Aït-Sahalia, "Transition Densities for Interest Rate and Other Nonlinear Diffusions," Journal of Finance,vol.54,1999 ,pp.1361–1395.
[14] Y. Aït-Sahalia, "Maximum likelihood estimation of discretely sample diffusion: A close form approximation approach," Econometrica, vol.70,2002,pp.223–262.

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)