



Bayesian Estimation for Markov Modulated Asset Prices

Richard Emilion, Hamed Faires, Srikanth K. Iyer

► **To cite this version:**

Richard Emilion, Hamed Faires, Srikanth K. Iyer. Bayesian Estimation for Markov Modulated Asset Prices. 2009. hal-00464386

HAL Id: hal-00464386

<https://hal.archives-ouvertes.fr/hal-00464386>

Submitted on 26 Mar 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bayesian Estimation for Markov Modulated Asset Prices.

Richard Emilion*

MAPMO, UMR 6628 Université d'Orléans-CNRS,
BP. 6759 45067 Orléans Cedex 2, France,

and

Hafedh Faires †

Mathematics Department- Faculty of Sciences Al-Imam University
P.O.Box 90950 Riyadh 11623 KSA

and

Srikanth K. Iyer‡

Department of Mathematics Indian Institute of Science
Bangalore - 560012, India.

March 17, 2010

Abstract

A Stochastic Differential Equation (SDE) appearing in mathematical finance is considered in random environment by assuming that its two parameters are switched by an unobserved continuous-time Markov chain whose states represent the states of the market environment. A Dirichlet process is placed as a prior on the space of the sample paths

*corresponding author:richard.emilion@labomath.univ-orleans.fr

†hafedh.faires@univ-orleans.fr

‡Research supported in part by UGC SAP IV, skiyer@math.iisc.ernet.in.

of this chain, leading to a hierarchical Dirichlet model whose estimation is done both on simulated data and on a real data set from the Indian market.

Keywords: Dirichlet process, Markov regime switching, random environment, Stochastic Differential equation, Geometric Brownian motion, Asset price.

1 Introduction

Models in which parameters move between a fixed number of regimes with switching controlled by an unobserved stochastic process, are very popular in a great variety of domains (Finance, Biology, Meteorology, Networks, etc.). This is notably due to the fact that this additional flexibility allows the model to account for random regime changes in the environment. In this paper we consider the estimation problem for a model described by a stochastic differential equation (SDE) with Markov regime-switching (MRS), i.e., with parameters controlled by a finite state continuous-time Markov chain (CTMC). Such a model was used, for example, in Deshpande and Ghosh (2008) to price options in a regime switching market. In such a setting, the parameter estimation problem poses a real challenge, mainly due to the fact that the paths of the CTMC are unobserved. A standard approach consists in using the celebrated EM algorithm (Dempster, Laird, Rubin, 1977) as proposed for example in Hamilton (1990). Elliott, Malcolm and Tsoi (2003) study this problem using a filtering approach.

In the present paper, our estimation approach is Bayesian, the aim being to find a pair (parameters, CTMC path) with likelihood as large as possible. We refer the reader to Schnatter (2006) for a wider discussion on Markov switching models and the comparative advantages of the Bayesian approach. Standard priors are placed on the parameters space but, as the CTMC paths are unobserved, a large number of paths are drawn from a Dirichlet process placed as a prior on the path space of the CTMC. The complete model then appears

as a Hierarchical Dirichlet Model (HDM), as in Ishwaran, James and Sun (2000) and Ishwaran and James (2002). The estimation procedure for the model considered in this paper requires some rather nontrivial computations of posterior distributions due to the temporal level induced by the specific SDE and the CTMC. Using the well-known stick-breaking approximation, each set of iterations selects the pair with largest likelihood and then the Dirichlet process is updated in order to look for other paths which can further improve the likelihood.

The considered SDE is that of a geometric Brownian motion, a popular model for asset prices in mathematical finance which depends on two parameters, namely the drift and volatility. It is considered in the extended MRS setting so that the CTMC transitions correspond to regime changes in the market.

The rest of the paper is structured as follows. In Section 2 we present the stock price SDE with MRS and the complete HDM. Section 3 is devoted to the posterior computations. The estimation algorithm is described in Section 4. Numerical results are presented in Section 5 for one simulated data and on a data from the Indian market. We conclude with a summary in the last Section.

2 Markov regime switching with Dirichlet prior

Our model is specified in a mathematical finance setting but can be extended similarly to many similar models. We model the stock price using a geometric Brownian motion with drift and variance depending on the state of the market, the latter being modeled by a continuous time Markov chain. In what follows, the notation σ will be used to denote the variance rather than the standard deviation.

The following notations will be adopted:

1. n will denote the number of observed data and also the length of an observed path.
2. M will denote the number of states of the Markov chain.
3. The state space of the chain will be denoted by $S = \{i : 1 \leq i \leq M\}$.
4. N will denote the number of simulated paths.
5. Given a path of the CTMC, m will denote the number of distinct states in that path.

We now describe the model:

- Let (X_t) be a CTMC taking values in the set $S = \{i : 1 \leq i \leq M\}$ which represents M possible states of the market. The transition probabilities of this chain are denoted by p_{ij} , $i, j \in S$ and the transition rate matrix is $Q = (q_{ij})_{i,j \in S}$ with

$$\lambda_i > 0, \quad q_{ij} = \lambda_i p_{ij} \quad \text{if } i \neq j, \quad \text{and} \quad q_{ii} = -\sum_{j \neq i} q_{ij}, \quad i, j \in S.$$

- The stock price follows the following SDE:

$$\frac{dS_t}{S_t} = \beta(X_t)dt + \sqrt{\sigma(X_t)}dB_t, \quad t \geq 0,$$

where B_t is a standard Brownian motion. By Ito's formula, the process $Z_t = \log(S_t)$ satisfies the SDE,

$$dZ_t = \mu(X_t)dt + \sqrt{\sigma(X_t)}dB_t, \quad t \geq 0,$$

where $\mu(X_t) = \beta(X_t) - \frac{1}{2}\sigma(X_t)$. Let the observed data be denoted by Z_0, Z_1, \dots, Z_n , and define the log-returns as, $Y_t = Z_t - Z_{t-1} = \log(S_t/S_{t-1})$, $t = 1, 2, \dots, n$. Given the path $X = \{X_s, 0 \leq s \leq n\}$, let $T_j(t)$ be the time spent by the path X in state j in the time interval $[t-1, t]$. Define

$$\mu(t) := \sum_{j=1}^M \mu_j T_j(t); \quad \sigma(t) := \sum_{j=1}^M \sigma_j T_j(t), \quad (1)$$

were (μ_j, σ_j) , are the drift and volatility in regime j , $j = 1, \dots, M$. Then, conditional on the path X , the random variables $Y_t \sim \mathcal{N}(\mu(t), \sigma(t))$, $t = 1, 2, \dots, n$, and are independent. Here $\mathcal{N}(\mu, \sigma)$ denotes a Normal distribution with mean μ and variance σ .

- For each $i = 1, 2, \dots, M$, let the priors on μ_i and σ_i are given by

$$\mu_i \stackrel{ind}{\sim} \mathcal{N}(\theta, \tau^\mu), \quad \text{with} \quad \theta \sim \mathcal{N}(0, A), \quad A > 0, \quad (2)$$

$$\sigma_i \stackrel{ind}{\sim} \Gamma(\nu_1, \nu_2), \quad (3)$$

where $\Gamma(\nu_1, \nu_2)$ denotes a Gamma distribution with shape parameter ν_1 and scale parameter ν_2 .

- We now place a Dirichlet prior $\mathcal{D}(\alpha H)$, (see Ferguson, 1973) on the path space of the CTMC (X_t) , with precision parameter $\alpha > 0$ and mean H which is a probability measure governing a CTMC on the path space $D([0, \infty), S)$, the set of cad-lag functions. The initial distribution according to H is the uniform distribution $\pi_0 = (1/M, \dots, 1/M)$, and the transition matrix is P_0 with $p_{ij} = 1/(M-1)$, $i \neq j$ and $\lambda_i = \lambda > 0$, $i, j = 1, \dots, M$. Thus the Markov chain under H will spend an exponentially distributed time with mean $1/\lambda$ in any state i and then jump to state $j \neq i$ with probability $1/(M-1)$.

A realization of the Markov chain from the above prior is generated as follows: Generate a large number of paths $X_i = \{x_s^i : 0 \leq s \leq n\}$, $i = 1, 2, \dots, N$, from H . Generate the vector of probabilities $(p_i, i = 1, \dots, N)$ from a stick-breaking scheme with parameter α (see Sethuraman, 1994). Then draw a realization of the Markov chain from the distribution

$$p = \sum_{i=1}^N p_i \delta_{X_i}, \quad (4)$$

The parameter λ is chosen to be small so that the variance is large and hence we obtain

a large variety of paths to sample from at a later stage. The prior for α is given by,

$$\alpha \sim \Gamma(\eta_1, \eta_2). \quad (5)$$

3 Estimation

Roughly speaking, estimation will be done by simulating a large number of paths of the Markov chain, selecting one path according to a probability vector generated by stick-breaking, and then using the blocked Gibbs sampling technique. This technique requires the posterior distribution of the each parameter conditioned on the current values of the other parameters.

We denote by μ and σ , the current values of the vectors $(\mu_1, \mu_2, \dots, \mu_n)$ and $(\sigma_1, \sigma_2, \dots, \sigma_n)$, respectively. Let Y be the vector of observed data (Y_1, \dots, Y_n) . Given the current path $X = (x_s, 0 \leq s \leq n)$ of the Markov chain, let $X^* = (x_1^*, \dots, x_m^*)$ be the distinct values in X .

3.1 Modifying the observed data set

In order to obtain the conditional distribution of the parameters, we first need to extract the change in the log-returns between the jump times of the Markov chain. Let $0 = t_0 < t_1 < t_2 < \dots < t_J$ be the times at which the path X changes state. Define the log-returns between the jump times, $W_k = \log(S_{t_k}/S_{t_{k-1}})$, $k = 1, 2, \dots, J$. To obtain realizations of the W_k from the observed Y process, we need to simulate Gaussian random variables conditioned on their sums.

Consider any $t \in \{0, 1, \dots, n\}$ for which the chain changes state atleast once in the time interval $[t-1, t]$. So for some p, k we have $t_{k-1} < t-1 \leq t_k < \dots < t_{k+p} \leq t < t_{k+p+1}$. Let $V_t^1 = \log(S_{t_k}/S_{t_{k-1}})$ and $V_t^2 = \log(S_t/S_{t_{k+p}})$. Then,

$$Y_t = V_t^1 + \sum_{i=1}^p W_{k+i} + V_t^2. \quad (6)$$

Suppose that the chain X is in state j_i in the time interval $[t_{k+i-1}, t_{k+i})$, $i = 0, 1, \dots, p+1$. Set $s_0 = t_k - (t-1)$, $s_i = t_{k+i} - t_{k+i-1}$, $i = 1, 2, \dots, p$, and $s_{p+1} = t - t_{k+p}$. Let $m_j = \mu_{j_i} s_i$ and $v_j = \sigma_{j_i} s_i$, $i = 0, 1, \dots, p+1$. Recall that $Y_t \sim \mathcal{N}(\mu(t), \sigma(t))$, where $\mu(t), \sigma(t)$ are as defined in (1). It is easy to see that the joint conditional density of $(V_t^1, W_{k+1}, \dots, W_{k+p})$ given $Y_t = y$ will be

$$f(u_0, u_1, \dots, u_p) = C \prod_{i=0}^p \exp \left(-\frac{1}{2} \frac{v_i + v_{p+1}}{v_i v_{p+1}} \left(u_i - \frac{v_{p+1} m_i + v_i (y - m_{p+1})}{v_i + v_{p+1}} \right)^2 \right), \quad (7)$$

where C is a constant that depends on y and the parameters. Thus, one can simulate the variables $V_t^1, W_k, W_{k+1}, \dots, W_{k+p}$ from independent Gaussian distributions and then obtain V_t^2 using (6).

Using the above procedure, we can obtain a realization for all W_k for which $[t_{k-1}, t_k] \subseteq [t-1, t]$, for some $t \in \{0, 1, \dots, n\}$. Now for any k for which there is a $q \geq 0$, such that $t-1 \leq t_{k-1} < t < t+1 < \dots < t+q \leq t_k < t+q+1$, we can obtain W_k using the relation

$$W_k = V_t^2 + \sum_{i=1}^q Y_{t+i} + V_{t+q+1}^1. \quad (8)$$

Note that the W values depend on the path X and the parameter values μ, σ and hence are to be computed in each iteration of the Gibbs sampling procedure which we describe next.

3.2 Gibbs sampling procedure

We are now ready to estimate the posterior distributions of the parameters using Gibbs sampling. The procedure consists of two nested iterations. The outer loop is initiated using the parameters chosen for the priors. Each iteration of the outer loop consists of a large number of iterations of the inner loop in which a sample is drawn recursively for each parameter conditioned on the current values of the other parameters and the data. At the end of an iteration of the inner loop we obtain one realization of the parameters from their approximate posterior distribution.

We now derive the conditional distribution for each parameter conditioned on the other parameters and the data. Recall the X^* is the set of distinct values observed in the path X of the CTMC.

- **Conditional for μ .** For each $j \in X^*$ draw

$$(\mu_j | \theta, \tau^\mu, \sigma, X, W) \stackrel{ind}{\sim} \mathcal{N}(\mu_j^*, \sigma_j^*), \quad (9)$$

where

$$\mu_j^* = \sigma_j^* \left(\sum_{k: X_{t_{k-1}}=j} \frac{W_k}{\sigma_j(t_k - t_{k-1})} + \frac{\theta}{\tau^\mu} \right),$$

$$\sigma_j^* = \left(\frac{n_j}{\sigma_j} + \frac{1}{\tau^\mu} \right)^{-1},$$

and n_j being the number of times j occurs in X . For each $j \in X \setminus X^*$, independently simulate $\mu_j \sim \mathcal{N}(\theta, \tau^\mu)$.

- **Conditional for σ .** For each $j \in X^*$ draw

$$(\sigma_j | \mu, \nu, X, W) \stackrel{ind}{\sim} \Gamma(\nu_1 + \frac{n_j}{2}, \nu_{2,j}^*), \quad (10)$$

where

$$\nu_{2,j}^* = \nu_{2,j} + \sum_{k: X_{t_{k-1}}=j} \frac{(W_k - \mu_j(t_k - t_{k-1}))^2}{2(t_k - t_{k-1})}.$$

Also for each $j \in X \setminus X^*$, independently simulate $\sigma_j \sim \Gamma(\nu_1, \nu_2)$.

- **Conditional for parameters of H**

The reestimation of the initial distribution and the transition matrix from a given path X^i can be done by any standard MLE procedure for CTMC.

- **Conditional for X .**

$$(X | p) \sim \sum_{i=1}^N p_i^* \delta_{X_i}, \quad (11)$$

where

$$p_i^* \propto \prod_{j=1}^m \left(\prod_{\{k: x_{t_{k-1}}^{i,*} = j\}} \frac{1}{(2\pi\sigma_j(t_k - t_{k-1}))^{1/2}} e^{-\frac{1}{2\sigma_j^2}(W_k^i - \mu_j(t_k - t_{k-1}))^2} \right) p_i, \quad (12)$$

where $(x_1^{i,*}, \dots, x_m^{i,*})$ denote the current $m = m(i)$ unique values of the states and t_k^i , W_k^i are as defined in subsection 3.1 for the path X_i , $i = 1, \dots, N$.

- **Conditional for p .**

$$p_1 = V_1^*, \text{ and } p_k = (1 - V_1^*) \cdots (1 - V_{k-1}^*) V_k^*, \quad k = 2, 3, \dots, N - 1, \quad (13)$$

where

$$V_k^* \stackrel{ind}{\sim} \beta(1 + r_k, \alpha),$$

r_k equal 1 if $i = k$ and 0 otherwise.

- **Conditional for α .**

$$(\alpha|p) \sim \Gamma \left(N + \eta_1 - 1, \eta_2 - \sum_{i=1}^{N-1} \log(1 - V_i^*) \right),$$

where the V^* values are those obtained in the simulation of p in the above step.

- **Conditional for θ .**

$$(\theta|\mu) \sim \mathcal{N}(\theta^*, \tau^*), \quad (14)$$

where

$$\theta^* = \frac{\tau^*}{\tau^\mu} \sum_{j=1}^M \mu_j,$$

and

$$\tau^* = \left(\frac{M}{\tau^\mu} + \frac{1}{A} \right)^{-1}.$$

Proof.

(a) The computations of the posterior distributions for μ , σ and θ follow in the same manner as in Ishwaran and James (2002) and Ishwaran and Zarepour (2000). Here, $X_t = s$ means that the class variable is equal to s .

(b) Conditional for X :

$$P\{X = X_i \mid p, \mu, \sigma, W\} = P\{W \mid p, \sigma, X = X_i, \mu\}P\{X = X_i \mid \sigma, \mu, p\}P\{\mu, \sigma\}$$

$$\propto \prod_{j=1}^m \left(\prod_{\{k: x_{t_{k-1}}^{i,*} = j\}} \frac{1}{(2\pi\sigma_j(t_k - t_{k-1}))^{1/2}} e^{-\frac{1}{2\sigma_j^2}(W_k^i - \mu_j(t_k - t_{k-1}))^2} \right) p_i$$

where $X_i = (x_1^i, \dots, x_n^i)$ and $(x_1^{i,*}, \dots, x_m^{i,*})$ denote the current m unique values in the path X_i .

(c) Conditional for p : The Sethuraman stick-breaking scheme can be extended to the two-parameter Beta distributions, see Walker and Muliere (1997, 1998):

Let $V_k \stackrel{ind}{\sim} \beta(a_k, b_k)$, for each $k = 1, \dots, N$. Let

$$p_1 = V_1, \text{ and } p_k = (1 - V_1) \cdots (1 - V_{k-1})V_k, \quad k = 2, 3, \dots, N - 1.$$

We will write the above random vector, in short as

$$p \sim SB(a_1, b_2, \dots, a_{N-1}, b_{N-1}).$$

By Connor and Mosimann (1969), the density of p is

$$\left(\prod_{k=1}^{N-1} \frac{\Gamma(a_k - b_k)}{\Gamma(a_k)\Gamma(b_k)} \right) p_1^{a_1-1} \cdots p_{N-1}^{a_{N-1}-1} p_N^{b_{N-1}-1} \times$$

$$\times (1 - P_1)^{b_1 - (a_2 - b_2)} \cdots (1 - P_{N-2})^{b_{N-2} - (a_{N-1} - b_{N-1})},$$

where $P_k = p_1 + \dots + p_k$.

From this, it easily follows that the distribution is conjugate for multinomial sampling, and consequently the posterior distribution of p given X , when $a_k = 1$ and $b_k = \alpha$ for each k , is

$$SB(a_1^*, b_2^*, \dots, a_{N-1}^*, b_{N-1}^*),$$

where

$$\begin{aligned} b_k^* &= \alpha \\ a_k^* &= 1 + r_k, \end{aligned}$$

and r_k equal 1 if $i = k$ and 0 otherwise, $k=1, \dots, N-1$. □

4 Implementation

The algorithm presented in the previous section was implemented in C language. The implementation includes:

- functions that simulate standard probability distributions: Uniform, Normal, Gamma, Beta, Exponential.
- a function that returns an index $\in \{1, \dots, n\}$ according to a vector of probability p_1, \dots, p_n .
- a function that simulates a probability vector according to stick-breaking scheme.
- a function that simulates n paths of a Markov chain.
- a function that records the number of times a state appears in a path.
- a function that chooses one of the paths according to a vector of probability.
- a function that modifies the parameters of prior distributions according to the formulas of the above posteriori distributions.

We first simulated a large number of paths given π_0 and Q_0 . Then at each iteration a path is selected according to the weights drawn by a stick breaking scheme and the parameters are updated using posteriori formulas. The pair (parameters, path) is saved if it improves

the likelihood. After a large number of iteration we use the pair with maximum likelihood to re-estimate π and Q_0 from the corresponding path. The above procedure is repeated a large number of times expecting an improvement of the likelihood.

4.1 Simulated data

We fit the model, using the algorithm developed above, to a simulated series of length $n = 480$, $M = 4$ regimes, the mean and variance in each state being chosen as follows:

$$\begin{aligned}(\mu_1, \sigma_1) &= (-1.15, 0.450) \\(\mu_2, \sigma_2) &= (-0.93, 0.450) \\(\mu_3, \sigma_3) &= (-0.60, 0.440) \\(\mu_4, \sigma_4) &= (1.40, 0.500).\end{aligned}$$

We carry out the estimation procedure for the series simulated using the above parameters with number of states $M = 10$, number of paths $N = 100$ and run it for 25,000 iterations. It was observed that the algorithm was able to put most of the mass (in terms of the stationary distribution of the MC) on 4 regimes, which are close to the ones chosen above.

At the end of the iterations we compute a confidence interval for the mean and for the variance in each regime. The confidence intervals for the mean and the variance are given below.

Regime 1:

$$I_m = [-1.208, -1.12423] \quad \text{and} \quad I_v = [0.431, 0.4738].$$

Regime 2:

$$I_m = [-0.9351, -0.9296] \quad \text{and} \quad I_v = [0.442, 0.4538].$$

Regime 3:

$$I_m = [-0.63446, -0.5140] \quad \text{and} \quad I_v = [0.4319, 0.4491].$$

Regime 4:

$$I_m = [1.30114, 1.43446] \quad \text{and} \quad I_v = [0.4949, 0.5081].$$

Based on the confidence intervals given above, we can conclude that the algorithm is able to identify the parameters of the simulated data set.

4.2 Bombay Stock Exchange Data

We applied our algorithm to the daily closing prices of the mid-cap index of the Bombay stock exchange from 21/12/2006 to 15/11/2007 (www.bseindia.com). For this data set we have, $n = 250$, and we run our algorithm with $N = 100$ paths and choose a Gamma(2, 4) distribution as the prior for α .

With the above choice, we obtain 5 regimes for which the estimates for the mean, variance and stationary probabilities are given in Table 1.

| | R 1 | R 2 | R 3 | R 4 | R 5 |
|----------|------------|-----------|-----------|-----------|-----------|
| μ | -0.001 | 0.001 | -0.003 | 0.001 | -0.003 |
| σ | 0.1982 e-4 | 1.0917e-4 | 2.9624e-4 | 0.6382e-4 | 1.0502e-4 |
| π | 0.21 | 0.12 | 0.1 | 0.43 | 0.12 |

Table 1. Estimated drift, volatility and stationary probabilities of the regimes.

The corresponding Markov chain path with best likelihood, its parameters λ_i s and the matrix of transition probability $(p_{i,j})_{i,j=1,\dots,5,i \neq j}$ are given in Tables 2,3 respectively.

| λ_1 | λ_2 | λ_3 | λ_4 | λ_5 |
|-------------|-------------|-------------|-------------|-------------|
| 0.03 | 0.03 | 0.05 | 0.02 | 0.06 |

Table 2. Estimated reciprocal mean sojourn times in the regimes.

$$\begin{pmatrix} & 0.5 & 0.2 & 0 & 0.3 \\ 0.43 & & 0.14 & 0.22 & 0.21 \\ 0.55 & 0.45 & & 0 & 0 \\ 0.18 & 0.27 & 0.17 & & 0.38 \\ 0 & 0.17 & 0.08 & 0.75 & \end{pmatrix}$$

Table 3. Estimated transition probability matrix.

We may classify the regimes on the basis of volatility as low (R_1), moderate (R_2, R_4, R_5) and high (R_3). Observe that in the high volatility state, the index has a negative drift as is usually observed in analysis of empirical data. The moderate volatility regime (R_4) is most persistent with an average sojourn time of 50 days whereas the high volatility regime persists on average for 20 days. The high volatility state R_3 is followed almost always by the states R_1 or R_2 .

A by-product of our algorithm is the distribution of the current state of the volatility, which is required to compute the price of an option (see [2] and references therein).

5 Conclusion

A Bayesian approach to estimation for a regime switching geometric Brownian motion is proposed. The algorithm while being computationally intensive is able to segregate the different regimes based on the drift and volatility, thus giving useful insights into the behavior of the market. It has been observed empirically that markets fluctuate between periods of high, moderate and low volatilities. The above estimation procedure provides a clear quantitative picture of the number of regimes and an estimate of the drifts and volatilities in these regimes. Estimation of current market state is also easier using the algorithm proposed compared to models using continuous stochastic volatility models. Given an estimate of the regime, the

algorithm also gives an idea of likely duration for which the regime is likely to persist and the distribution of the regimes that may follow.

Acknowledgment: We would like to thank Chiranjit Mukhopadhyay for sharing his data with us.

References

- [1] DEMPSTER, A.P., LAIRD, R. M. AND RUBIN, D. B. (1977) Maximum likelihood from incomplete data via EM algorithm. *J. Royal Stat. Soc. B.* **39**, 1-38.
- [2] DESHPANDE, A. AND GHOSH, M. K. (2007). Risk Minimizing Option Pricing in a Regime Switching Market. *Stochastic Analysis and Applications*, **28**, 2008.
- [3] ELLIOTT, R.J., MALCOLM, W.P., AND TSOI, A.H. (2003). Robust parameter estimation for asset price models with Markov modulated volatilities. *Journal of Economics Dynamics and Control*, **27 (8)**, 1391-1409.
- [4] FERGUSON, T.S (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209-230.
- [5] HAMILTON, J. A. (1990). Analysis of time series subject to changes of regime. *J. Econometrics*, **45**, 39-70.
- [6] ISHWARAN H., JAMES, L.F AND SUN, J. (2000). Bayesian Model Selection in Finite Mixtures by Marginal Density Decompositions, *conditionally accepted by Journal of the American Statistical Association*.
- [7] ISHWARAN, H. AND ZAREPOUR, M. (2000). Markov Chain Monte Carlo in Approximate Dirichlet and Beta Two- Parameter Process Hierarchical Models, *EJP* **1**, 1-28.

- [8] ISHWARAN, H. and JAMES, L.F. (2002). Approximate Dirichlet processes computing in finite normal mixtures: smoothing and prior information. *J. Comp. Graph. Stat.* **11** 209-230.
- [9] SCHNATTER, S. F. (2006). Finite mixture and Markov switching models. Springer.
- [10] SETHURAMAN, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica*, **4**, 639-650.