Efficient estimation of Markov-switching model with application in stock price classification

Farshid Mehrdoust[1], **Idin Noorani**[2], **Mahdi Khavari**[3]

[1] Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran
fmehrdoust@guilan.ac.ir

[2] Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran
idinnoorani@yahoo.com

[3] Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran
mahdikhavari29@gmail.com

Abstract:
In this paper, we discuss the calibration of geometric Brownian motion model equipped with Markov-switching factor. Since the motivation for this research comes from a recent stream of literature in stock economics, we propose an efficient estimation method to sample series of stock prices based on the expectation-maximization algorithm. We also implement an empirical application to evaluate the performance of the suggested model. For this purpose, based on the proposed Markov-switching model, we classify market data under various economic regimes by estimating the smoothed probabilities of hidden Markov chain states. Numerical results through the classification of the data set show that the proposed Markov-switching model fits the actual stock prices and reflects the main stylized facts of market dynamics.

Keywords: Regime-switching model, Parameter estimation, Expectation-maximization algorithm.
MSC2010 Classifications: 62M10, 65M70, 91G80

[1]Corresponding author
Received: 2021-11-23 Approved: 2021-12-28

Introduction

There are several researches in the literature, which their purpose to present a more realistic reflection of financial markets. The stochastic models can be broadly divided into: local volatility, stochastic volatility, volatility as an unknown process and Markov-switching models. The stochastic volatility model as an unknown process was first presented by [2]. However, as pointed out in [3], these models complicate the solution of financial market derivatives, because they involve additional nonlinear factors in ordinary differential equations.

From [8] and [6], local volatility models have been proposed by deterministic function of asset price and time horizon. Fast calibration of these volatility models types was extensively studied by [10]. As we know, the model requires a smooth and continuous implied volatility, while numerical methods considered for the local volatility models can yield unstable results. According to studies conducted by [7] and [16], the performance of the local volatility model is not satisfactory because these models are highly restrictive. The stochastic models were generally investigated by [1]. It should be noted that the stochastic models are too popular for pricing financial derivatives and these models are also able to produce implied volatility smile well (see [12], [26] and [11]). However, stochastic models are not able to take into account different economic states without the regime-switching factor. Under such circumstances, these models may not reflect well the significant events that occur in the dynamics of financial time series.

Many financial time series sporadically show significant interruptions in their behavior that are associated with events such as war, climate change, recession, inflation, and so on (see [29] and [4]). In these situations, economists tend to use variables that change the behavior of time series dynamics. The model that can analyze these changes is called the Markov-switching model. This model can consider the intermittent and repetitive changes of economic regimes endogenously, while in the other models, these changes are usually as specific and exogenous. In most researches, there may be little information about the times that the parameters change. Therefore, we need to make results for milestones so that the change of parameters is significant. In this regard, some researchers first considered models that only one regime change occurs in the data, then models with more than one regime were designed. The probability of switching depends on the path of the asset. This dependence introduced as Markov-switching models (see [19]). Markov-switching (regime-switching) models are among the models that take the sudden changes into account and have been widely used to evaluate asset prices. Such models have recently been considered not only in econometrics but also in other areas such as population dynamics, river flow analysis, and speech recognition (see [21], [14] and [17]). Further, a comparison of various types of Markov-switching models for exchange rate can be found in [28]. Newly, [15] and [23] have experimentally shown that adding the Markov-switching factor to the volatility dynamics leads to more non-Gaussianity stock returns.

In order to compare the prices obtained by the model and the actual market data, calibration by model parameters estimation is evermore implicated (see [25], [13] and [24]). This challenging process under the Markov-switching model is the focus of this paper. Given the latent switching mechanism, in which the model parameters change based on sudden changes over time, it is necessary to estimate these parameters by deriving model parameters and the state values of the hidden Markov chain process simultaneously. In this paper, we calibrate the geometric Brownian motion (GBM) model equipped with the hidden Markov chain. To do this, we use the method proposed by [20], in which the parameters of the hidden Markov model were estimated using the expectation-maximization (EM) algorithm. The proposed model actually demonstrates the Markov-switching GBM (MSGBM) dynamics, such that the stock price can be classified based on its calibration.

The MSGBM model framework

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $[0, T]$ is a time interval with a maturity time of $T > 0$. Also, let $Y := \{Y(t)\}_{t\in[0,T]}$ be a hidden Markov chain with N state on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that the set of chain states is $\mathcal{E} = \{e_1, \ldots, e_N\}$. Without loss of the generality, suppose that the hidden Markov chain states are considered as singular vectors; That is, for every $j = 1, \ldots, N$, jth component of e_j is one and the rest is zero. Let $\Pi = (\pi_{jk})_{j,k=1,\ldots,N}$ be a hidden Markov chain rate matrix, where π_{jk} is the intensity of the chain transition from state e_j (regime j)

into state e_k (regime k). Such that for every $j, k = 1, \ldots, N$ we have

$$\pi_{jk} = \begin{cases} \pi_{jk} \geq 0 & j \neq k, \\ -\sum_{j=1, j\neq k}^{N} \pi_{jk} & j = k. \end{cases}$$

Consider the hidden Markov chain transition matrix Y as $P = (p_{jk})$, which includes the probabilities $p_{jk} = \mathbb{P}(Y(t) = e_k | Y(t-1) = e_j)$, such that p_{jk} means switching regime j (state e_j) at time $t-1$ to regime k (state e_k) at time t. For instance, if the hidden Markov chain has two states ($N = 2$), then

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ p_{21} & 1 - p_{21} \end{pmatrix}.$$

Here, the considered Markov chain in the Markov-switching model has two-state, i.e., $N = 2$. Generalizing the model to more states is similar to the two-state model.

Let $\mu = (\mu_1, \mu_2) \in \mathbb{R}^2$ and $\sigma = (\sigma_1, \sigma_2) \in \mathbb{R}^2$ be dependent on $Y(t)$. We define

$$\begin{cases} \mu(t) := \mu(Y(t)) = \langle \mu, Y(t) \rangle = \sum_{j=1}^{2} \mu_j \langle Y(t), e_j \rangle, \\ \sigma(t) := \sigma(Y(t)) = \langle \sigma, Y(t) \rangle = \sum_{j=1}^{2} \sigma_j \langle Y(t), e_j \rangle, \quad \sigma_j > 0, \quad j = 1, 2, \end{cases} \tag{1}$$

where $\langle \cdot, \cdot \rangle$ represents the inner product in $\mathbb{R}^2$. On the other hand, as expressed in [9], $Y(t)$ has the following expression

$$dY(t) = \Pi Y(t)dt + dV(t), \tag{2}$$

where $\mathbb{R}^2$-value process $\{V(t)\}_{t\in[0,T]}$ is an $\mathcal{F}^Y$-martingale process, such that $\mathcal{F}^Y := \mathcal{F}^Y\{(t)\}_{t\in[0,T]}$ indicates the filtration generated by the hidden Markov chain Y. We define $f(t, y) = e^{-\Pi t} y$. Then from Eq. (2) and by applying Ito's lemma to $f(t, y)$, we obtain

$$\begin{aligned} df(s, y) &= f_s ds + f_y dY(s) + \frac{1}{2} f_{yy} dY(s) dY(s) \\ &= -\Pi e^{-\Pi s} Y(s) ds + e^{-\Pi s} (\Pi Y(s) ds + dV(s)) \\ &= e^{-\Pi s} dV(s). \end{aligned}$$

Integrating with respect to s from 0 to t and multiplying the sides by $e^{\Pi t}$ we have

$$Y(t) = e^{\Pi t}Y(0) + \int_0^t e^{\Pi(t-s)}dV(s).$$

Therefore

$$\begin{aligned}\mathbb{E}[Y(t)] =& e^{\Pi t}\mathbb{E}[Y(0)]\\ =& e^{\Pi t}\left(e_1 \times \mathbb{P}(Y(0) = e_1) + e_2 \times \mathbb{P}(Y(0) = e_2)\right).\end{aligned}$$

We put $m := (e_1 \times \mathbb{P}(Y(0) = e_1) + e_2 \times \mathbb{P}(Y(0) = e_2))$. In this case, using Eq. (1), we have

$$\begin{aligned}\mathbb{E}[r(Y(t))] &= \langle \mu, \mathbb{E}[Y(t)]\rangle = \langle \mu, me^{\Pi t}\rangle,\\ \mathbb{E}[\sigma(Y(t))] &= \langle \sigma, \mathbb{E}[Y(t)]\rangle = \langle \sigma, me^{\Pi t}\rangle.\end{aligned} \tag{3}$$

Suppose that $S(t)$ is the stock price at the moment $t \in [0, T]$. In this case, the RSGBM model under the accurate probability measure $\mathbb{P}$ can be expressed as follows

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dB(t), \quad S(0) = s,$$

wherein $(\mu(t))_{t\in[0,T]}$ and $(\sigma(t))_{t\in[0,T]}$ are the average rates of return and market volatility, respectively, which depend on the hidden Markov chain $Y(t)$. Also, $B(t)$ is a Brownian motion under the probability measure $\mathbb{P}$.

Recently [27], showed that the RSGBM model under the neutral risk probability measure $\mathbb{Q}$ could be expressed as follows

$$dS(t) = r(t)S(t)dt + \sigma(t)S(t)dW(t), \quad S(0) = s, \tag{4}$$

where $(r(t))_{t\in[0,T]}$ is the interest rate that depends on the hidden Markov chain $Y(t)$. $W(t)$ is also the standard Brownian motion under the probability measure $\mathbb{Q}$. In general, $S(t)$ can be expressed as follows

$$S(t) = \begin{cases} S_1(t) & \text{if } Y(t) \text{ in state } e_1 \text{ (regime 1)},\\ \\ S_2(t) & \text{if } Y(t) \text{ in state } e_2 \text{ (regime 2)},\end{cases} \tag{5}$$

where,

$$dS_i(t) = r_iS_i(t)dt + \sigma_iS_i(t)dW(t), \quad i = 1 \vee 2.$$

Here and everywhere, it is always assumed that the Markov chain process Y and the Brownian motion W are independent of each other. Therefore, Y is considered as an external factor of market information. Assume that $\mathcal{F}_t^W$ is a filtration generated by Brownian motion W. We define $\mathcal{F}_t := \mathcal{F}_t^W \vee \mathcal{F}_t^Y$ as a global filtration.

Implementation of the EM algorithm

The expectation-maximization algorithm or EM algorithm is an efficient iterative method for calculating the maximum likelihood estimation in the presence of missing or hidden information. In contrast, in the maximum likelihood estimation method, the parameters are estimated in the presence of visible data. [5] first introduced the EM algorithm. Each step of the EM algorithm consists of two steps: expectation (E) and maximization (M). In step E, the confidential data are estimated based on a conditional expectation based on the observed data and the available model parameters. In step M, the likelihood function is maximized under the assumption that the lost data is known.

Assume that Z is a random vector derived from a parametric family to derive the EM algorithm. We want to find Λ so that $\mathbb{P}(Z|\Lambda)$ is maximized. The logarithm likelihood function is defined as follows

$$L(\Lambda) = \log(\mathbb{P}(Z|\Lambda)), \tag{6}$$

where the likelihood function is considered as a function of the Λ parameter for X data. Since the logarithm function is a strictly ascending function, the value Λ, which maximizes $\mathbb{P}(Z|\Lambda)$, will also maximize $L(\Lambda)$.

As mentioned, the EM algorithm is an iterative method for maximizing $L(\Lambda)$, in which $\Lambda^{(m)}$ is an estimate of Λ after mth iteration. Since the goal is to maximize $L(\Lambda)$, we want to calculate the updated estimate of Λ, such that,

$$L(\Lambda) > L(\Lambda^{(m)}).$$

It means that we can maximize the difference between the terms $L(\Lambda)$ and $L(\Lambda^{(m)})$ and we can write

$$L(\Lambda) - L(\Lambda^{(m)}) = \log(\mathbb{P}(Z|\Lambda)) - \log(\mathbb{P}(Z|\Lambda^{(m)})). \tag{7}$$

So far we have not considered any missing or hidden variables, the EM algorithm provides a framework for including such information in problems where there is such information. Suppose that Y is a hidden random vector composed of random variables y. Now we write the total probability $\mathbb{P}(Z|\Lambda)$ based on the hidden variable y. In this case, we have

$$\mathbb{P}(Z|\Lambda) = \sum_y \mathbb{P}(Z|y, \Lambda)\mathbb{P}(y|\Lambda). \tag{8}$$

Therefore, Eq. (7) can be rewritten as follows

$$L(\Lambda) - L(\Lambda^{(m)}) = \log\left(\sum_y \mathbb{P}(Z|y,\Lambda)\mathbb{P}(y|\Lambda)\right) - \log(\mathbb{P}(Z|\Lambda^{(m)})). \tag{9}$$

Given the Jensen inequality for each convex function such as f and for nonnegative Λ_j, that is applied in the condition $\sum_{j=1}^{n} \Lambda_j = 1$, we have

$$f\left(\sum_{j=1}^{n} \Lambda_j x_j\right) \leq \sum_{j=1}^{n} \Lambda_j f(x_j).$$

In this case, using Jensen inequality for Eq. (9), we obtain

$$\begin{aligned}
L(\Lambda) - L(\Lambda^{(m)}) &= \log\left(\sum_y \mathbb{P}(Z|y,\Lambda)\mathbb{P}(y|\Lambda)\right) - \log(\mathbb{P}(Z|\Lambda^{(m)}))\\
&= \log\left(\sum_y \mathbb{P}(Z|y,\Lambda)\mathcal{P}(y|\Lambda)\frac{\mathbb{P}(y|Z,\Lambda^{(m)})}{\mathbb{P}(y|Z,\Lambda^{(m)})}\right) - \log(\mathbb{P}(Z|\Lambda^{(m)}))\\
&= \log\left(\sum_y \mathbb{P}(y|Z,\Lambda^{(m)})\left(\frac{\mathbb{P}(Z|y,\Lambda)\mathbb{P}(y|\Lambda)}{\mathbb{P}(y|Z,\Lambda^{(m)})}\right)\right) - \log(\mathbb{P}(Z|\Lambda^{(m)}))\\
&\geq \sum_y \left(\mathbb{P}(y|Z,\Lambda^{(m)})\log\left(\frac{\mathbb{P}(Z|y,\Lambda)\mathbb{P}(y|\Lambda)}{\mathbb{P}(y|Z,\Lambda^{(m)})}\right)\right) - \log(\mathbb{P}(Z|\Lambda^{(m)}))\\
&= \sum_y \left(\mathbb{P}(y|Z,\Lambda^{(m)})\log\left(\frac{\mathbb{P}(Z|y,\Lambda)\mathbb{P}(y|\Lambda)}{\mathbb{P}(y|Z,\Lambda^{(m)})\mathbb{P}(Z|\Lambda^{(m)})}\right)\right)\\
&\triangleq \Delta(\Lambda|\Lambda^{(m)}).
\end{aligned} \tag{10}$$

As a result

$$L(\Lambda) \geq L(\Lambda^{(m)}) + \Delta(\Lambda|\Lambda^{(m)}). \tag{11}$$

From $l\left(\Lambda|\Lambda^{(m)}\right) \triangleq L(\Lambda^{(m)}) + \Delta(\Lambda|\Lambda^{(m)})$, the inequality of expression (11) is written as follows

$$L(\Lambda) \geq l(\Lambda|\Lambda^{(m)}). \tag{12}$$

Therefore $l(\Lambda|\Lambda^{(m)})$ is an upper bound for the maximum likelihood function $L(\Lambda)$.

Now we show that the function $l(\Lambda|\Lambda^{(m)})$ and $L(\Lambda)$ are equals for each

$\Lambda = \Lambda^{(m)}$.

$$
\begin{aligned}
l(\Lambda^{(m)}|\Lambda^{(m)}) &= L(\Lambda^{(m)}) + \Delta(\Lambda|\Lambda^{(m)}) \\
&= L(\Lambda^{(m)}) + \sum_y \mathbb{P}(y|Z,\Lambda^{(m)}) \log\left(\frac{\mathbb{P}(Z|y,\Lambda^{(m)})\mathbb{P}(y|\Lambda^{(m)})}{\mathbb{P}(y|Z,\Lambda^{(m)})\mathbb{P}(Z|\Lambda^{(m)})}\right) \\
&= L(\Lambda^{(m)}) + \sum_y \mathbb{P}(y|Z,\Lambda^{(m)}) \log\left(\frac{\mathbb{P}(Z,y|\Lambda^{(m)})}{\mathbb{P}(Z,y|\Lambda^{(m)})}\right) \\
&= L(\Lambda^{(m)}). \qquad (13)
\end{aligned}
$$

As can be seen, the function $l(\Lambda|\Lambda^{(m)})$ is an upper bound for the likelihood function $L(\Lambda)$, which is equal to $\Lambda = \Lambda^{(m)}$, and the value of $l(\Lambda|\Lambda^{(m+1)})$ is maximized for $\Lambda = \Lambda^{(m+1)}$. Since $l(\Lambda) \geq l(\Lambda|\Lambda^{(m)})$, an increase of $l(\Lambda|\Lambda^{(m)})$ is guaranteed and the value of the likelihood function $L(\Lambda)$ increases with each step.

Now we just need to calculate the value $\Lambda^{(m+1)}$. Therefore we have

$$
\begin{aligned}
\Lambda^{(m+1)} &= \arg\max_{\Lambda}\{l(\Lambda|\Lambda^{(m)}\} \\
&= \arg\max_{\Lambda}\left\{L(\Lambda^{(m)}) + \sum_y \mathbb{P}(y|Z,\Lambda^{(m)}) \log\left(\frac{\mathbb{P}(Z|y,\Lambda)\mathbb{P}(y|\Lambda)}{\mathbb{P}(y|Z,\Lambda^{(m)})\mathbb{P}(Z|\Lambda^{(m)})}\right)\right\}. \qquad (14)
\end{aligned}
$$

Excluding sentences that are fixed relative to Λ, we have

$$
\begin{aligned}
\Lambda^{(m+1)} &= \arg\max_{\Lambda}\left\{\sum_y \mathbb{P}(y|Z,\Lambda^{(m)}) \log\left(\mathbb{P}(Z|y,\Lambda)\mathbb{P}(y|\Lambda)\right)\right\} \\
&= \arg\max_{\Lambda}\left\{\sum_y \mathbb{P}(y|Z,\Lambda^{(m)}) \log\left(\frac{\mathbb{P}(Z,y,\Lambda)\mathbb{P}(y,\Lambda)}{\mathbb{P}(y,\Lambda)\mathbb{P}(\Lambda)}\right)\right\} \\
&= \arg\max_{\Lambda}\left\{\sum_y \mathbb{P}(y|Z,\Lambda^{(m)}) \log\left(\mathbb{P}(Z,y|\Lambda)\right)\right\} \\
&= \arg\max_{\Lambda}\left\{\mathbb{E}_{Y|Z,\Lambda^{(m)}}\left[\log\left(\mathbb{P}(Z,y|\Lambda)\right)\right]\right\}. \qquad (15)
\end{aligned}
$$

As we can see, in Eq. (15), the phases of expectation and maximization have appeared. Thus the EM algorithm includes the following iterations

(i) Step E: Determine the conditional expectation $\mathbb{E}_{Y|Z,\Lambda^{(m)}}[\log \mathbb{P}(Y,y|\Lambda)]$.

(ii) Step M: Maximize the expression $\mathbb{E}_{Y|Z,\Lambda^{(m)}}[\log \mathbb{P}(Y,y|\Lambda)]$ relative to Λ.

Estimation of the MSGBM model parameters

In this section, we use the EM algorithm based on the actual market data to estimate the Markov-switching times of the MSGBM model, the transition probabilities of the Markov chain, and the parameters that depend on the hidden Markov chain, i.e., $(r(t), \sigma(t))$. Using the same strategy, the Markov-switching model parameter with more states can be estimated.

Before estimating the Markov-switching model parameters in Eq. (4), we consider its discrete form by using the Euler method as follows

$$S(t_j) = S(t_{j-1}) + r(t)S(t_{j-1}) + \sigma(t_j)S(t_{j-1})\sqrt{\delta}\epsilon_j, \tag{16}$$

where $\{0 = t_0 < t_1 < \ldots < t_n = T\}$ is a partition of interval $[0, T]$ with step length $\delta := t_j - t_{j-1}$. Also $\epsilon_1, \ldots, \epsilon_n$ is a random instance of the standard normal distribution. Since the ϵ_j has a standard normal distribution, then $S(t_j)$ has a normal distribution with mean $S(t_{j-1}) + r(t_j)S(t_{j-1})$ and variance $\sigma^2(t_j)S^2(t_{j-1})\delta$, conditional upon the $S(t_{j-1})$ and $Y(t_j)$. Therefore, assuming that the number of Markov chain states is two, the density function of $S(t_j)$ conditional upon the $S(t_{j-1})$ and $Y(t_j)$ is expressed as follows

$$\begin{aligned} f\left(S(t_j)|S(t_{j-1}), Y(t_j) = e_k, \Lambda\right) = &\left(\frac{1}{\sqrt{2\pi\sigma_k^2 S^2(t_{j-1})\delta}}\right) \\ &\times \exp\left\{-\frac{\left(S(t_j) - S(t_{j-1}) - r_k S(t_{j-1})\right)^2}{2\sigma_k^2 S^2(t_{j-1})\delta}\right\}, \quad k = 1, 2, \end{aligned} \tag{17}$$

where $\Lambda := \{r_1, r_2, \sigma_1, \sigma_2, P\}$ is a set of parameters that we want to estimate. Consider the likelihood function based on the parameter Λ as follows

$$L(\Lambda|\mathcal{S}, \mathcal{Y}) = \prod_{j=1}^{n} f\left(S(t_j)|S(t_{j-1}), Y(t_j) = e_k, \Lambda\right), \tag{18}$$

where $\mathcal{S} = (S(t_1), \ldots, S(t_n))$ and $\mathcal{Y} = (Y(t_1), \ldots, Y(t_n))$. In this case, the logarithm likelihood function is obtained as follows

$$
\begin{aligned}
\mathcal{L}(\Lambda|\mathcal{S},\mathcal{Y}) &= \ln\left(\prod_{j=1}^{n} f\left(S(t_j)|S(t_{j-1}),Y(t_j)=e_k,\Lambda\right)\right)\\
&= \sum_{j=1}^{n} \ln f\left(S(t_j)|S(t_{j-1}),Y(t_j)=e_k,\Lambda\right)\\
&= -\frac{1}{2}\sum_{j=1}^{n}\ln(2\pi\sigma_k^2 S^2(t_{j-1})\delta) - \sum_{j=1}^{n}\left[\frac{(S(t_j)-S(t_{j-1})-r_kS(t_{j-1}))^2}{2\sigma_k^2S^2(t_{j-1})\delta}\right]\\
&= -\frac{1}{2}\sum_{j=1}^{n}\left[\ln(2\pi\sigma_k^2 S^2(t_{j-1})\delta) + \frac{(S(t_j)-S(t_{j-1})-r_kS(t_{j-1}))^2}{\sigma_k^2S^2(t_{j-1})\delta}\right]. \qquad (19)
\end{aligned}
$$

As mentioned before, the EM algorithm maximizes the probability function (4) for models with missing observations or unobserved variables. Its purpose is to maximize the probability function in the presence of unobserved data. This algorithm is an iterative method consisting of the following two steps in $(m+1)$th iteration.

(i) Step E: Determine the following conditional expectation

$$
\mathbb{E}_{\mathcal{Y}|\mathcal{S},\Lambda^{(m)}}\left[\ln\mathbb{P}\left(\mathcal{S},\mathcal{Y}\mid\Lambda^{(m+1)}\right)\right] := \sum_{k=1}^{2}\sum_{j=1}^{n}\ln\mathbb{P}\left(\mathcal{S},\mathcal{Y}\mid\Lambda^{(m+1)}\right)\mathbb{P}\left(Y(t_j)=e_k\mid\mathcal{S},\Lambda^{(m)}\right). \qquad (20)
$$

(ii) Step M: Maximize the problem (20) relative to Λ. In this case, the logarithm likelihood function of Eq. (19) is maximized relative to the model parameters, and $\Lambda^{(m+1)}$ is obtained in $(m+1)$th iterations.

Next, using Theorems 0.3 and 0.4, we specify steps E and M of the EM algorithm, respectively, to estimate the parameters of the MSGBM model.

Theorem 0.3. *Conditional expectation $\mathbb{E}_{\mathbb{Y}|\mathbb{S},\Lambda^{(m)}}[\ln\mathbb{P}(\mathbb{S},\mathbb{Y}|\Lambda^{(m+1)})]$, in step E of the EM algorithm is as follows*

$$
\begin{aligned}
\mathbb{E}_{\mathcal{Y}|\mathcal{S},\Lambda^{(m)}}\left[\ln\mathbb{P}\left(\mathcal{S},\mathcal{Y}\mid\Lambda^{(m+1)}\right)\right] &= \sum_{k=1}^{2}\sum_{j=1}^{n}\mathbb{P}\left(Y(t_j)=e_k|\mathcal{S},\Lambda^{(m)}\right) \qquad (21)\\
&\quad\times \ln f\left(S(t_j)|S(t_{j-1}),Y(t_j)=e_k,\Lambda^{(m+1)}\right).
\end{aligned}
$$

Proof. See [22]. □

The value $\mathbb{E}\left[Y(t_j)=e_k|S(t_j),\Lambda\right]$ is known as the "filtering inference" and is a linear combination of the observation vector $S(t_j)$ and the probability $\mathbb{P}\left(Y(t_j)=e_k|S(t_j),\Lambda\right)$ is calculated. Since $Y(t)$ is hidden and not directly visible, the expected values of the Markov chain process can be observed by the observation vector $\mathbb{E}\left[Y(t_j)=e_k|S(t_j),\Lambda\right]$ to be calculated. These expected results are known as "smoothed inference" and calculate the conditional probability $\mathbb{P}\left(Y(t_j)=e_k|\mathcal{S},\Lambda\right)$. In order to calculate the smoothed probability, the filtering step must be completed.

The expression $\mathbb{P}\left(Y(t_j)=e_k|S(t_j),\Lambda\right)$ in the Theorem 0.3, consists of two stages of filtering and smoothing. Let the parameter vector be obtained by step M in step $(m-1)$, in this case, as examined by [20], the two-step filtering and smoothing algorithm to obtain $\mathbb{P}\left(Y(t_j)=e_k|S(t_j),\Lambda^{(m)}\right)$ is expressed as follows

(i) Filtering: For $j=1,\ldots,n$, as long as $\mathbb{P}\left(Y(t_n)=e_k|S(t_n),\Lambda^{(m)}\right)$ is obtained, we calculate the following probability

$$\begin{aligned}&\mathbb{P}\left(Y(t_i)=e_k|S(t_j),\Lambda^{(m)}\right)\\&=\frac{\mathbb{P}\left(Y(t_j)=e_k|S(t_{j-1}),\Lambda^{(m)}\right)f\left(S(t_j)|Y(t_j)=e_k,S(t_{j-1}),\Lambda^{(m)}\right)}{\sum_{i=1}^{2}\left[\mathbb{P}\left(Z(t_j)=e_i|S(t_{j-1}),\Lambda^{(m)}\right)f\left(S(t_j)|Y(t_j)=e_i,S(t_{j-1}),\Lambda^{(m)}\right)\right]},\end{aligned}$$

where

$$\mathbb{P}\left(Y(t_{j+1})=e_k|S(t_j),\Lambda^{(m)}\right)=\sum_{i=1}^{2}P_{ki}^{(m)}\mathbb{P}\left(Y(t_j)=e_i|S(t_j),\Lambda^{(m)}\right).$$

(ii) Smoothing: For $j=n-1,\ldots,1$, we calculate the following probability

$$\begin{aligned}&\mathbb{P}\left(Y(t_j)=e_k|\mathcal{S},\Lambda^{(m)}\right)\\&=\sum_{i=1}^{2}\left[\frac{\mathbb{P}\left(Y(t_{j+1})=e_i|\mathcal{S},\Lambda^{(m)}\right)\mathbb{P}\left(Y(t_j)=e_k|S(t_j),\Lambda^{(m)}\right)P_{ik}^{(m)}}{\mathbb{P}\left(Y(t_{j+1})=e_i|S(t_j),\Lambda^{(m)}\right)}\right],\end{aligned}$$

where, the probability of transition is as follows

$$P_{ik}^{(m+1)}=\frac{\sum_{j=1}^{n}\left[\mathbb{P}\left(Y(t_{j+1})=e_i\mid S(t_{j+1}),\Lambda^{(m)}\right)\left(\frac{P_{ik}^{(m)}\mathbb{P}\left(Y(t_j)=e_k|\mathcal{S},\Lambda^{(m)}\right)}{\mathbb{P}\left(Y(t_{j+1})=e_i|S(t_j),\Lambda^{(m)}\right)}\right)\right]}{\sum_{j=1}^{n}\mathbb{P}\left(Y(t_j)=e_k\mid\mathcal{S},\Lambda^{(m)}\right)}.$$

In addition, as a starting point of the algorithm we have

$$h_j^{(m)}=\mathbb{P}\left(Y(t_1)=e_j|S(t_0),\Lambda^{(m)}\right),$$

and we will have the following iterations

$$h_j(m+1) = \mathbb{P}\left(Y(t_1) = e_j|\mathcal{S}, \Lambda^{(m)}\right),$$

where $h_j^{(0)}$ for each $j = 1, 2$ must be given as input to the algorithm (see [18]).

Theorem 0.4. *The maximum value of the relationship*

$$\mathbb{E}_{\mathcal{Y}|\mathcal{S},\Lambda^{(m)}}\left[\ln \mathbb{P}\left(\mathcal{S}, \mathcal{Y}|\Lambda^{(m+1)}\right)\right]$$

In step M of the EM algorithm, under the MSGBM model, the following points are obtained

$$\hat{r}_k = \frac{\sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)\left(\frac{S(t_j)-S(t_{j-1})}{S(t_{j-1})}\right)}{\sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)}, \tag{22}$$

$$\hat{\sigma}_k = \sqrt{\frac{\sum_{j=1}^{n}\left[\left(\mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)\right)\left(\frac{\left(S(t_j)-S(t_{j-1})-\hat{r}_k S(t_{j-1})\right)^2}{S^2(t_{j-1})\delta}\right)\right]}{\sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)}}. \tag{23}$$

Proof. According to Theorem 0.3, we have

$$\mathbb{E}_{\mathcal{Y}|\mathcal{S},\Lambda^{(m)}}\left[\ln \mathbb{P}\left(\mathcal{S}, \mathcal{Y} \mid \Lambda^{(m+1)}\right)\right] = \sum_{k=1}^{2}\sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)$$
$$\times \ln f\left(S(t_j)|S(t_{j-1}), Y(t_j) = e_k, \Lambda^{(m+1)}\right).$$

Substituting Eq. (19) in the expression above gives

$$\begin{aligned}\mathbb{E}_{\mathcal{Y}|\mathcal{S},\Lambda^{(m)}}\left[\ln \mathbb{P}\left(\mathcal{S}, \mathcal{Y} \mid \Lambda^{(m+1)}\right)\right] = \sum_{k=1}^{2}\sum_{j=1}^{n} \Bigg[\mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right) \\ \times \left(-\frac{1}{2}\left[\ln\left(2\pi\sigma_k^2 S^2(t_{j-1})\delta\right)\right.\right. \\ \left.\left. + \frac{\left(S(t_j) - S(t_{j-1}) - r_k S(t_{j-1})\right)^2}{\sigma_k^2 S^2(t_{j-1})\delta}\right]\right)\Bigg].\end{aligned} \tag{24}$$

To obtain the maximum points of Eq. (24), we differentiate from r_k and σ_k $(k = 1, 2)$ and then set it to zero. Differentiating Eq. (24) with respect to r_k and set it equal to zero, we obtain

$$\frac{1}{2\sigma_k^2\delta}\sum_{j=1}^{n}\left[\mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)\left(\frac{2S(t_{j-1})\left(S(t_j) - S(t_{j-1}) - r_k S(t_{j-1})\right)}{S^2(t_{j-1})}\right)\right] = 0$$
$$\Rightarrow \sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)\left(\frac{S(t_j) - S(t_{j-1}) - r_k S(t_{j-1})}{S(t_{j-1})}\right) = 0$$
$$\Rightarrow \sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right)\left(\frac{S(t_j) - S(t_{j-1})}{S(t_{j-1})}\right) = r_k \sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k|\mathcal{S}, \Lambda^{(m)}\right).$$

Therefore, estimation of the parameter r_k for $k = 1, 2$ is obtained as follows

$$\hat{r}_k = \frac{\sum_{j=1}^{n} \left[\left(\mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right) \right) \left(\frac{S(t_j) - S(t_{j-1})}{S(t_{j-1})} \right) \right]}{\sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right)},$$

which indicates the interest rate parameter is in the kth regime.

Now we differentiate from Eq. (24) with respect to σ_k and then set it to zero, in which case we have

$$\sum_{j=1}^{n} \left[\mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right) \left(-\frac{1}{2} \left[\frac{2}{\sigma_k} - \frac{2\left(S(t_j) - S(t_{j-1}) - \hat{r}_k S(t_{j-1}) \right)^2}{\sigma_k^3 S^2(t_{j-1})\delta} \right] \right) \right] = 0$$

$$\Rightarrow \frac{1}{\sigma_k^3} \sum_{j=1}^{n} \left[\mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right) \left(\sigma_k^2 - \frac{\left(S(t_j) - S(t_{j-1}) - \hat{r}_j S(t_{j-1}) \right)^2}{S^2(t_{j-1})\delta} \right) \right] = 0$$

$$\Rightarrow \sum_{j=1}^{n} \left[\left(\mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right) \right) \left(\frac{\left(S(t_j) - S(t_{j-1}) - \hat{r}_k S(t_{j-1}) \right)^2}{S^2(t_{j-1})\delta} \right) \right]$$

$$= \sigma_k^2 \sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right).$$

Therefore, estimation of the parameter σ_k for $k = 1, 2$ is obtained as follows

$$\hat{\sigma}_k = \sqrt{\frac{\sum_{j=1}^{n} \left[\left(\mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right) \right) \left(\frac{\left(S(t_j) - S(t_{j-1}) - \hat{r}_k S(t_{j-1}) \right)^2}{S^2(t_{j-1})\delta} \right) \right]}{\sum_{j=1}^{n} \mathbb{P}\left(Y(t_j) = e_k | \mathcal{S}, \Lambda^{(m)} \right)}},$$

which indicates the market volatility is in the kth regime. $\square$

After estimating the Markov-switching model parameters, the data should be classified according to the smoothed probabilities in regimes 1 and 2. More precisely, if the probability $\mathbb{P}\left(Y(t_j) = e_2 | \mathcal{S} \right)$ is greater than $\mathbb{P}\left(Y(t_j) = e_1 | \mathcal{S} \right)$, then the model dynamic is currently in regime 2, otherwise it is in regime 1. In other words, since $\mathbb{P}\left(Y(t_j) = e_2 | \mathcal{S} \right) + \mathbb{P}\left(Y(t_j) = e_1 | \mathcal{S} \right) = 1$, if the smoothed probability for state e_2 at time t_j is greater than 0.5, that is, if $\mathbb{P}\left(Y(t_j) = e_2 | \mathcal{S} \right) > 0.5$, then the model dynamic at time t_j is in regime 2, otherwise at time t_j it is currently in regime 1.

To estimate the parameters of the RSGBM model, we use the daily stock price data of Microsoft and Intel companies from 01/01/2017 to 01/01/2020 (Figures 1 and 2). We set the maximum number of iterations of the EM algorithm to 20. The likelihood estimation results are reported

Table 1: Estimation of the MSGBM model parameters by using the EM algorithm.

Market	Regime 1		Regime 2		Transition probability			
	$\hat{r}$	$\hat{\sigma}$	$\hat{r}$	$\hat{\sigma}$	P_{11}	P_{12}	P_{21}	P_{22}
Microsoft	0.00201	0.05292	0.00082	0.17487	0.94606	0.05394	0.05654	0.94346
Intel	0.00140	0.06050	-0.00025	0.21297	0.89718	0.10282	0.10431	0.89569

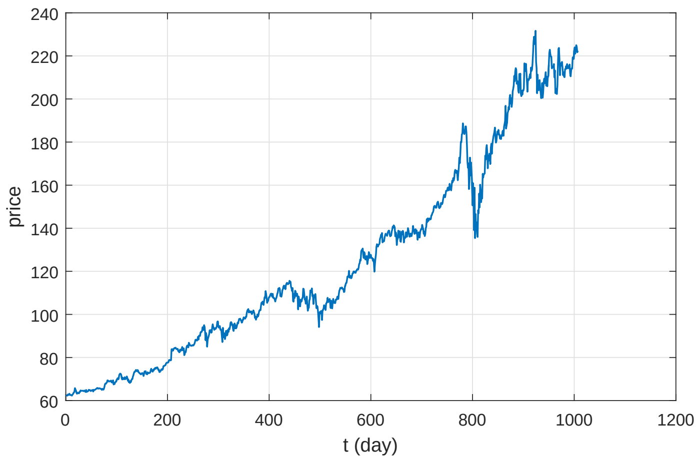

Figure 1: Microsoft stock price from 01/01/2017 to 01/01/2020.

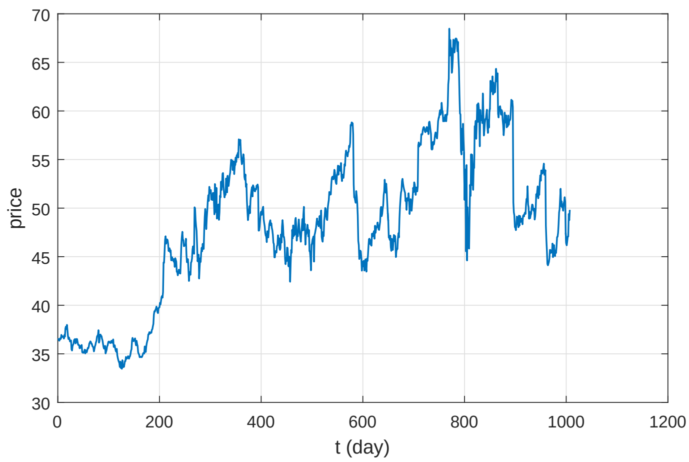

Figure 2: Intel stock price from 01/01/2017 to 01/01/2020.

in the last iteration of the EM algorithm in Table 1 for the Microsoft and Intel markets. We can also see the trend of this algorithm in estimating the parameters of the RSGBM model and transition probability of the Markov chain states (transition probability of hidden Markov chain) with various iterations for Microsoft and Intel markets in Figures 3 and 4, respectively. As it is known, the values of the parameters converge to a certain level after several iterations, which indicate the estimated value of the parameter.

Figures 5 and 6 show Microsoft and Intel stock price data classification based on smoothed probability in state e_1 (regime 1) and state e_2 (regime 2), respectively. Significant results are obtained from each regime and give each one a real economic concept. As shown in Table 1, regime 1 corresponds to a low-volatility regime, and regime 2 represents a high-stress state constantly changing between periods. The reflection of this feature for the fundamental market data of Microsoft and Intel from 01/01/2018 to 01/01/2019 are shown in Figures 5 and 6.

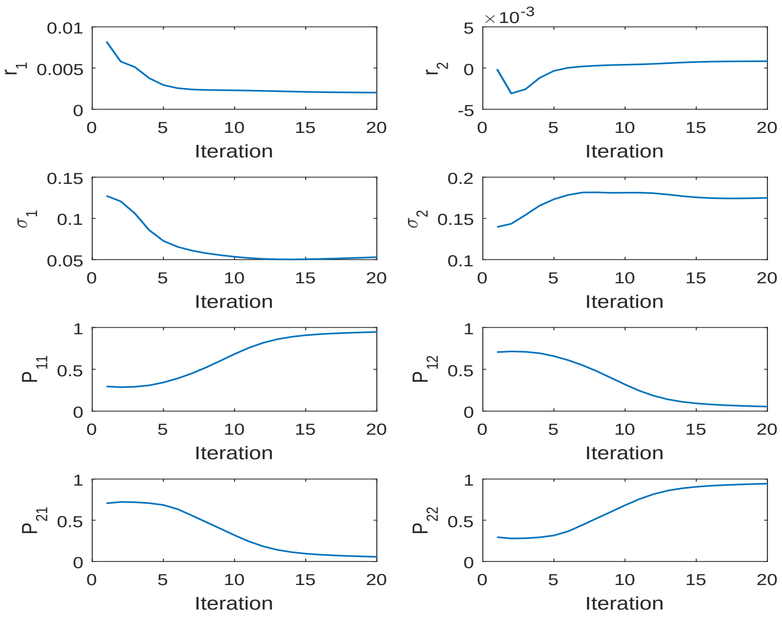

Figure 3: Estimation of the MSGBM model parameters for Microsoft data by the EM algorithm with various iterations.

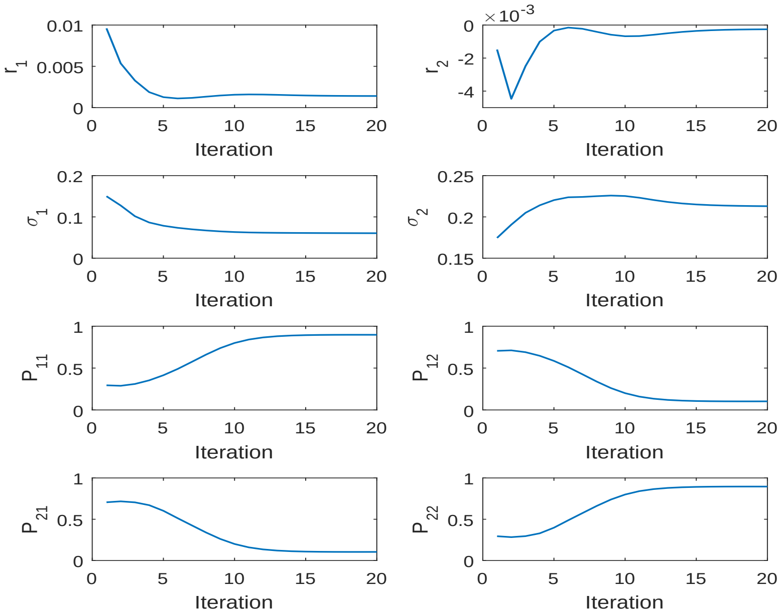

Figure 4: Estimation of the MSGBM model parameters for Intel data by the EM algorithm with various iterations.

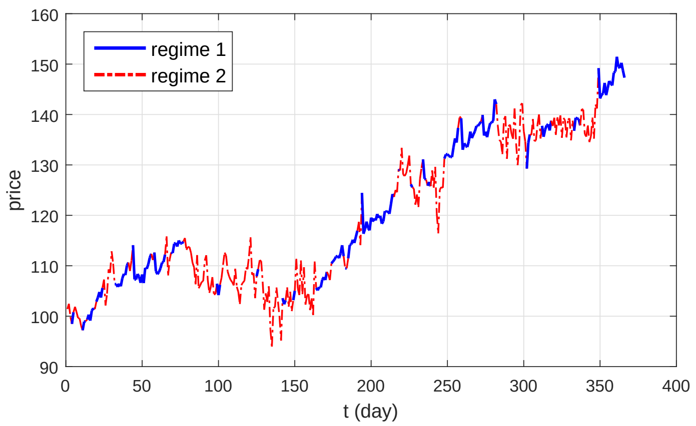

Figure 5: Microsoft stock price classification from 01/01/2018 to 01/01/2019 by regimes 1 and 2 under the MSGBM model.

Conclusion

In this paper, a stochastic model based on the two-state hidden Markov chain is studied. The proposed model represents a Markov-switching

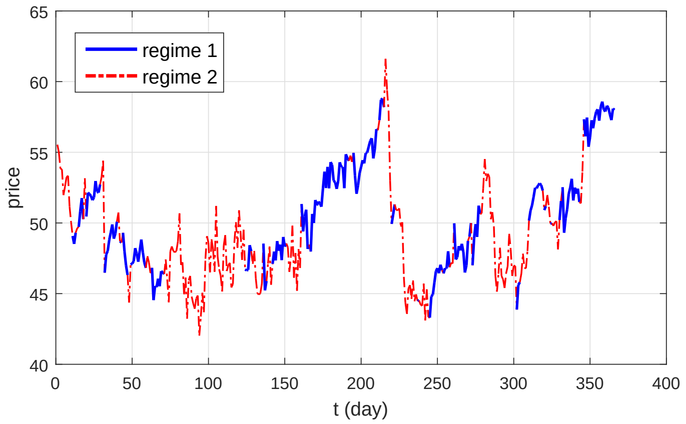

Figure 6: Intel stock price classification from 01/01/2018 to 01/01/2019 by regimes 1 and 2 under the MSGBM model.

model whose parameters depend on the hidden Markov chain and change over time. We estimated the model parameters based on the actual data using the EM algorithm, where the conditional expectation of E-step, is obtained by the two-step filtering-smoothing algorithm. According to the obtained numerical results, in the first and second regimes, the interest rates are high and low, respectively, and also in these two regimes, the market volatility is low and high, respectively. These cases, which represent a healthy stock economy in the first regime and a sick stock economy in the second regime, were evaluated by classifying real data. Numerical results show that the proposed MSGBM model can well reflect these market realities.

Bibliography

[1] C. Alexander, L. M. Nogueira, *Hedging with stochastic local volatility*, 2004.

[2] M. Avellaneda, A. Levy, and A. Parás, *Pricing and hedging derivative securities in markets with uncertain volatilities*, Applied Mathematical Finance, 2 (1995), pp. 73-88.

[3] M. Avellaneda, R. Buff, *Combinatorial implications of nonlinear uncertain volatility models: the case of barrier options*, Applied Mathematical Finance, 6 (1999), pp. 1-18.

[4] T. Davig, *Regime-switching debt and taxation*, Journal of Monetary Economics, 51 (2004), pp. 837-859.

[5] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society: Series B (Methodological), 39 (1977), pp. 1-22.

[6] E. Derman, I. Kani, *The volatility smile and its implied tree*, Goldman Sachs Quantitative Strategies Research Notes, 2 (1994), pp. 45-60.

[7] B. Dumas, J. Fleming, and R. E. Whaley, *Implied volatility functions: Empirical tests*, The Journal of Finance, 53 (1998), pp. 2059-2106.

[8] B. Dupire, *Pricing with a smile*, Risk, 7 (1994), pp. 18-20.

[9] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov models: estimation and control*, Springer Science & Business Media, 2008.

[10] B. Engelmann, M. R. Fengler, M. Nalholm, and P. Schwendner, *Static versus dynamic hedges: an empirical comparison for barrier options*, Review of Derivatives Research, 9 (2006), pp. 239-264.

[11] S. Fallah, F. Mehrdoust, *American option pricing under double Heston stochastic volatility model: simulation and strong convergence analysis*, Journal of Statistical Computation and Simulation, 89 (2019), pp. 1322-1339.

[12] S. Fallah, F. Mehrdoust, *On the existence and uniqueness of the solution to the double Heston model equation and valuing Lookback option*, Journal of Computational and Applied Mathematics, 350 (2019), pp. 412-422.

[13] S. Fallah, F. Mehrdoust, *CEV model equipped with the long-memory*, Journal of Computational and Applied Mathematics, 389 (2021), pp. 113359.

[14] G. A. Fink, *Markov models for pattern recognition: from theory to applications*, Springer Science & Business Media, 2014.

[15] S. Goutte, A. Ismail, and H. Pham, *Regime-switching stochastic volatility model: estimation and calibration to VIX options*, Applied Mathematical Finance, 24 (2017), pp. 38-75.

[16] P. S. Hagan, D. Kumar, A. S. Lesniewski, and D. E. Woodward, *Managing smile risk*, The Best of Wilmott, 1 (2002), pp. 249-296.

[17] M. Hahn, S. Frühwirth-Schnatter, and J. Sass, *Estimating models based on Markov jump processes given fragmented observation series*, AStA Advances in Statistical Analysis, 93 (2009), pp. 403-425.

[18] J. D. Hamilton, *Analysis of time series subject to changes in regime*, Journal of Econometrics, 45 (1990), PP. 39-70.

[19] J. D. Hamilton, *Regime switching models. In Macroeconometrics and time series analysis*, Palgrave Macmillan, London, 2010.

[20] C. J. Kim, *Dynamic linear models with Markov-switching*, Journal of Econometrics, 60 (1994), pp. 1-22.

[21] Q. Luo, X. Mao, *Stochastic population dynamics under regime switching* Journal of Mathematical Analysis and applications, 334 (2007), pp. 69-84.

[22] F. Mehrdoust, I. Noorani, *Forward price and fitting of electricity Nord Pool market under regime-switching two-factor model*, Mathematics and Financial Economics, (2021), pp. 1-43.

[23] F. Mehrdoust, I. Noorani, *Pricing S&P500 barrier put option of American type under Heston-CIR model with regime-switching*, International Journal of Financial Engineering, 6 (2019), pp. 1950014.

[24] F. Mehrdoust, S. Fallah, *On the calibration of fractional two-factor stochastic volatility model with non-Lipschitz diffusions*, Communications in Statistics-Simulation and Computation, (2020), pp. 1-20.

[25] F. Mehrdoust, I. Noorani, and A. Hamdi, *Calibration of the double Heston model and an analytical formula in pricing American put option*, Journal of Computational and Applied Mathematics, 392 (2021), pp. 113422.

[26] F. Mehrdoust, A. R. Najafi, S. Fallah, and O. Samimi, *Mixed fractional Heston model and the pricing of American options*, Journal of Computational and Applied Mathematics, 330 (2018), pp. 141-154.

[27] I. Noorani, F. Mehrdoust, and A. Nasroallah, *A generalized antithetic variates Monte-Carlo simulation method for pricing of Asian option in a Markov regime-switching model*, Mathematics and Computers in Simulation, 181 (2021), pp. 1-15.

[28] M. Pourrafiee, S. M. Pourmohammad Azizi, M. Mohammadi Larijani, and A. Pahlevannezhad, *Comparing the different types of Markov switching model for Euro to Iran Rial exchange rate*, Journal of Mathematics and Modeling in Finance, 1 (2021), pp. 49-58.

[29] C. A. Sims, T. Zha, *Were there regime switches in US monetary policy?*, American Economic Review, 96 (2006), pp. 54-81.