MODELLING OF STOCK PRICES BY THE MARKOV CHAIN MONTE CARLO METHOD

Mantas LANDAUSKAS
Kaunas University of Technology, Faculty of Fundamental Sciences
Studentų 50, LT - 51368 Kaunas, Lithuania
mantas.landauskas@ktu.lt

Eimutis VALAKEVIČIUS
Kaunas University of Technology, Faculty of Fundamental Sciences
Studentų 50, LT - 51368 Kaunas, Lithuania
eimval@ktu.lt

Abstract. This paper presents a universal approach to modelling stock prices. The technique involves Markov Chain Monte Carlo (MCMC) sampling from piecewise-uniform distribution.

Today’s financial models are based on assumptions which make them inadequate in many cases. One of the most important issues is determining the distribution of a stock price, its return or other financial mean. The approach proposed in this paper removes almost all presumptions from a distribution of a stock price. The probability density must be evaluated using some nonparametric estimates. The kernel density estimate (KDE) suits well for that purpose. It gives a smooth and presentable estimate.

MCMC was chosen due to its versatility and is applied to KDE using piecewise-linear distribution as proposal density. The proposal density is constructed according to the KDE. Such link between the piecewise-linear distribution’s simplicity and relative massiveness of KDE balances together.

Involving the kernel density estimate and the methodology to sample from it makes the technique universal for modelling any real stochastic system while having empirical data only and barely any assumptions about the distribution of it.

JEL classification: C10, C15, C46, C65.
Keywords: Stock prices, Markov chain, Monte Carlo method, MCMC, kernel density, piecewise-linear distribution.
Reikšminiai žodžiai: akcijų kainos, Markovo grandinė, Monte Karlo metodas, branduolinių tankis, dalimis tolygusis skirstinys.
Introduction

Classical model of stock prices has some assumptions about financial data. It cannot be applied to model the stock price having returns which are not log normally distributed.

There are several approaches to model difficult quantities, but they specialize in different areas. The purpose of this paper is to present a universal technique for modelling stock prices. This technique consists of special numerical methods and is suitable for any empirical data.

The Markov chain Monte Carlo method is used to sample from empirical probability density of a stock price. The technique is flexible and requires just the ability to calculate probability at any given point. Furthermore, MCMC was successfully applied to one-factor models for the interest rate (B. Eraker, 2001). This also acts as the reasoning for choosing it for this approach of modelling stock prices.

It is also needed to approximately evaluate empirical probability density. This is performed using kernel density estimation. The link between these two methods is considered and this leads to apply it on every financial data.

1. Monte Carlo Modelling of Stock Prices

The process of a stock price is treated as a Brownian motion. Thus its value satisfies the equation:

\[
\frac{dS}{S} = \mu dt + \sigma dz.
\]  

Consider a financial mean with log normally distributed returns. The random walk of price of such a financial mean is modeled according this formula (P. Wilmott, 2007):

\[
S(t+\Delta t) = S(t) e^{(\delta - \frac{1}{2} \sigma^2) \Delta t + \sigma \sqrt{\Delta t} Z}.
\]

Here random value \(Z \sim N(0,1)\) follows standard normal distribution, \(\delta\) is annual risk free return and \(\sigma\) is annual standard deviation of the logarithm of a stock price.

2. Markov chain Monte Carlo (MCMC)

Suppose it is needed to generate \(x_i \sim \pi(x)\). When \(x_i \sim \pi(x)\) is difficult to sample from, MCMC sampling technique could be performed. In fact MCMC is a set of techniques used for this purpose. The main idea of it is to construct a Markov chain \(\{X_i\}_{i=0}^{\infty}\), such that

\[
\lim_{i \to \infty} P(X_i = x) = \pi(x).
\]
A Markov chain is predefined by an initial state $P(X_0 = x_0) = g(x_0)$ and the transition kernel $P(y|x) = P(X_{i+1} = y|X_i = x)$. Stationary distribution $\pi(x) = \lim_{i \to \infty} f(x_i)$ is unique if the chain is ergodic. Then:

$$\pi(y) = \sum_{x \in \Omega} x \pi(x) P(y|x), \forall y \in \Omega.$$  

(2.2)

Latter equality could be rewritten as a set of $(n - 1)$ linear equations:

$$\begin{align*}
\pi(x_2) &= \pi(x_1) P(x_2|x_1) + \pi(x_2) P(x_2|x_2) + \ldots + \pi(x_n) P(x_2|x_n) \\
\pi(x_n) &= \pi(x_1) P(x_n|x_1) + \pi(x_2) P(x_n|x_2) + \ldots + \pi(x_n) P(x_n|x_n)
\end{align*}$$  

(3.3)

here $n := |\Omega|$. There are a total number of $(n - 1)$ equations and $n(n - 1)$ transition probabilities $P(y|x_k)$, $k = 1, n$, $j = 1, n - 1$. Thus there exist an infinite number of transition kernels $P(y|x)$, such that the stationary distribution of the Markov chain is $\pi(x)$.

One of the techniques used for constructing such a transition kernel is Metropolis-Hastings algorithm (J.S. Dagpunar, 2007). The idea of it is to choose any other transition kernel $Q(y|x)$. Then there exists a probability that $Q(y|x)$ is equal to $P(y|x)$.

$$P(y|x) = Q(y|x) \alpha(y|x), \ y \neq x, \ \alpha(y|x) \in [0; 1].$$  

(2.4)

Considering the detailed balance condition of a time-homogeneous Markov chain yields:

$$\pi(x) Q(y|x) \alpha(y|x) = \pi(y) Q(x|y) \alpha(x|y), \ \forall x \neq y.$$  

(2.5)

The general solution for (2.5) is $\alpha(y|x) = r(x, y)\pi(y)Q(x|y)$. It is necessary to have a higher acceptance ratio when sampling random numbers, therefore by adjusting $r(x, y)$ and considering higher acceptance ratio while sampling random numbers (V. Prokaj, 2009) it is shown that:

$$\alpha(y|x) = \min\left(1, \frac{\pi(y)Q(x|y)}{\pi(x)Q(y|x)}\right).$$  

(2.6)

3. **Nonparametric probability density estimation**

Consider a sample consisting of random independent and identically distributed values $X_i$. Kernel density estimate is chosen for evaluate the probability density of $X_i$.

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i), \ K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right).$$  

(3.1)

here $K(\cdot)$ is the kernel function, $h$ is its width.
Below are some kernel functions that are frequently used. The triangular kernel function is useful if the data has sharp edged distribution. Gaussian kernel makes the estimate’s PDF plot very smooth.

\[
K(x) = \begin{cases} 
1 - |x|, & |x| \leq 1, \\
0, & |x| > 1.
\end{cases} \quad \text{(triangular), (3.3)}
\]

\[
K(x) = \frac{3}{4} \left( 1 - x^2 \right), \quad |x| \leq 1, \quad \text{(Yapanichnikov), (3.4)}
\]

\[
K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{(Gauss). (3.5)}
\]

Basically, such probability density estimation is about assigning kernel density to each \( X_i \) and including weighted sum of all other assignations. The contribution of any other \( X_j \) to the probability value at \( X_i \) is smaller if \( X_i - X_j \) is bigger.

Figure 1 shows the probability density estimation from 5 given points while applying Gaussian kernel. The estimate is absolutely smooth. The only drawback of such
estimation is the necessity of using all the points from the sample while evaluating the probability at a particular point.

4. A New Approach to Modelling Stock Prices

4.1. Evaluation of Distribution Function

This chapter presents the approach to model stock prices or any other statistical data (the method is universal enough) without knowing analytical probability density function.

First of all kernel density estimation must be performed and construct an estimate to the return of a stock price. At this point there could be a discussion if this estimate is accurate, but it is assumed to be exact. And there is no need to look for analytical functions which best fit in a particular case. It is not necessary to think about the shape at all, it forms itself according the data. The only question is the width of the kernel function.

4.2. Special technique for constructing a proposal density

The target probability density is now constructed. In order to model it a special technique is required, because there are no inverse cumulative density function or one cannot represent the estimate using known analytical PDF’s. MCMC is a solution but it could not be applied directly to the PDF estimate mentioned before.

Probably the biggest advantage of MCMC is the ability to generate required density using the proposal density, which should be similar in shape to target density. No other requirements to proposal density. Thus the complexity of proposal density is as simple as it is needed. Consider a histogram, which is relatively fast and simple non-parametric estimate for target density. It is possible to use it as proposal density therefore. But the assumption about target density not being discrete must be taken in mind, there are no set of values to construct a histogram from. The idea of the technique presented in this paper is to construct a piecewise-uniform distribution according to the kernel density estimate. A piecewise-uniform distribution is defined in eq. (4.2.1).

![Proposal density as a piecewise-linear probability distribution.](image-url)
Modelling of Stock Prices by the Markov Chain Monte Carlo Method

\[ q(x) = \begin{cases} 
q_1, & x_0 < x \leq x_1, \\
q_2, & x_1 < x \leq x_2, \\
\vdots & \vdots \\
q_n, & x_{n-1} < x \leq x_n.
\end{cases} \]  

(4.2.1)

The area below the probability density function must be equal to 1, thus:

\[ \sum_{i=1}^{n} q_i = \frac{1}{x_n - x_0}. \]  

(4.2.2)

This distribution is treated as a proposal density. Generating random numbers from this distribution is fast and simple.

![Generating random numbers using inverse CDF](image)

Fig. 3. Generating random numbers using inverse CDF.

Sampling from \( q(x) \) requires application of a search procedure. Firstly a \( u \sim U(0;1) \) is drawn. Then it is required to find the interval \((x_i, x_{i-1}]\), \(i = 1, n\), to which \( u \) belongs to. Since the number of intervals is going to be small, this step does not require many calculation steps. Then \( u \) is mapped to \( x \) according to the CDF of \( q(x) \) like in figure 3. CDF of \( q(x) \) is obtained by calculating the area below target density in each of the intervals.

Using \( q(x) \) as the proposal density and kernel density estimate as a target distribution implies random values \( y_i \) having distribution equal to \( f(y) \). It must be noted that acceptance ratio for \( x_i \) is now

\[ \alpha(y|x) = \min \left( 1, \frac{f(y)q(x)}{f(x)q(y)} \right). \]  

(4.2.3)
The sampling technique is called the independence Metropolis-Hastings when \( q(x|y) = q(y|x) \). The independence sampler has one significant advantage compared to traditional Metropolis-Hastings: the sequence \( \{x_i\} \) has no memory effect. Each random value accepted in simulation process does not depend on previous value. Thus there is no importance in what was \( x_0 \) generated. A brief description of Metropolis-Hastings techniques could be found in (M. Johannes, 2006).

5. Calibration of the model

Every model should give adequate results and compare to other known models or techniques. Making the model hold this is called a calibration. In this case, the new technique for modelling stock prices must give similar results as traditional Monte Carlo if stock returns are log normally distributed. Again the hypotheses about the normality of the logarithms of the stock returns are going to be tested.

![Graph](image.png)

**Fig. 4.** Yahoo! Inc. historical share prices.

Yahoo! Inc. (YHOO) share prices from 2010 01 04 to 2010 09 27 were chosen for performing the calibration. Historical share prices are depicted in figure 4. By performing the Kolmogorov-Smirnov test on the logarithms of the prices’ returns \( p = 0.992 \) and \( D = 0.0742 \) were obtained. \( D < p \) shows that the logarithms are normally distributed and leads data to be suitable for classical stock price model.
200 trajectories (figure 5) were modeled for each technique. According to the classical Monte Carlo approach the mean value of a price after 50 days will be 18.08 $. The newly proposed technique gave it 18.10 $ per share. This is actually expected, because the trend was considered.

In figure 6 the histograms of classical Monte Carlo and MCMC are compared. They represent the distribution of stock prices at the end of the modelling process. The modelling process contained 1000 paths of a stock price and simulated 100 days. Thus it required 100000 random stock returns to be performed.
The biggest difference between the two histograms exists at about mean value. The tails match better. Classical Monte Carlo converges to stock price distribution when the number of paths is increasing; the method proposed in this paper should also. Checking if the new method matches Monte Carlo is equivalent to checking if it converges to the distribution of a stock price. While evaluating the difference between two probability densities often an integral of an absolute value of their difference is used. Now consider an estimate:

$$\Delta_{MC} = \sum_{j=1}^{m} |h_{MC}(X_j) - h_{MCMC}(X_j)|,$$

(5.1)

$h_{MC}()$ and $h_{MCMC}()$ is the histograms of a stock price at the end of the modelling, $m$ is number of bars and $X_j$ represents the center point of the $j$-th bar.

Table 1. Differences between the histograms of the stock prices modeled by Monte Carlo and MCMC

<table>
<thead>
<tr>
<th>No. of bars in $g(\lambda)$</th>
<th>No. of trajectories</th>
<th>No. of random values</th>
<th>$\Delta_{MC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>50</td>
<td>5000</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10000</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20000</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>50000</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>100000</td>
<td>0.064</td>
</tr>
</tbody>
</table>

Table 1 shows how $\Delta_{MC}$ changes if the number of a stock price paths $N$ increases. The bigger $N$ the more Monte Carlo and MCMC results are alike. MCMC proved to be suitable for modelling stock prices. The number of bars in $g(\lambda)$ is equal to a question of peaks and distance between them in target distribution. Since the returns of stock prices have a distribution similar
in shape to normal distribution, \( g(x) \) should have a small odd number of bars in order to best match the target distribution.

Table 2. Differences between the histograms of the stock prices modeled by Monte Carlo and MCMC

<table>
<thead>
<tr>
<th>No. of bars in ( g(x) )</th>
<th>No. of trajectories</th>
<th>No. of random values</th>
<th>( \Delta_{MC} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>50</td>
<td>5000</td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>10000</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20000</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>50000</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>100000</td>
<td>0.050</td>
</tr>
</tbody>
</table>

As the table 2 shows choosing 5 bars in proposal density results in more precise distribution of stock prices. Accuracy increases but the calculation time is higher also. This is due to more calculation steps required to find the interval of \( g(x) \) to which a particular random number belongs to.

6. Modelling stock prices

Here is an example when classical Monte Carlo method cannot be applied to model stock prices.

![Distribution of normalized logarithms of continuous day returns.](image)

The histogram of normalized logarithms of continuous day returns \( R_t \) of Tesco Corporation (TESO) is depicted in figure 8. Although the hypothesis of normality is accepted, there exist two peaks. If one is confident about the shape of histogram, the assumption of normality should be rejected and standard Monte Carlo cannot be applied.
Constructing kernel density estimate for $r_i = \frac{S_i - S_{i-1}}{S_{i-1}}$ using Gaussian kernel function also gives PDF with 2 peaks (figure 9). MCMC with piecewise-linear distribution as a proposal density was applied to this PDF.

Average share price after 50 days resulted in $14.75. All the prices generated are distributed according kernel density estimate. Sampling is based entirely on empirical data and has no assumptions about PDF.

7. Conclusions

1. While estimating the probability density of a custom stock return with kernel density, each return in the sample is considered.
2. Proposed technique for modelling stock prices leads for average path of the stock price having small dispersion. The same holds for the Monte Carlo method.
3. The higher number of intervals used for constructing piecewise-uniform probability density leads to better accuracy of distribution modeled, but requires more time to perform the method.
4. Combining MCMC with kernel density estimate leads the technique for being able to model any real system. Thus empirical probability density is constructed using particular statistical information. This could be value of a financial mean, product quality measures and so on. Thus the technique is universal.

References

Mantas Landauskas is a postgraduate student in applied mathematics at Faculty of Fundamental Sciences, Kaunas University of Technology. Master's degree research area: modeling of stochastic systems by MCMC method.