MPRA
Munich Personal RePEc Archive

Parameter estimation from multinomial trees to jump diffusions with k means clustering

Lee, Kiseop and Xu, Mingxin

University of North Carolina at Charlotte

26 April 2007

Online at https://mpra.ub.uni-muenchen.de/3307/
MPRA Paper No. 3307, posted 24 May 2007 UTC

Parameter Estimation from Multinomial Trees to Jump Diffusions with K Means Clustering

Kiseop Lee*, Mingxin Xu†

April 26, 2007

Abstract

Ever since the pioneering work of Cox, Ross and Rubinstein [8], tree models have been popular among asset pricing methods. On the other hand, statistical estimation of parameters of tree models has not been studied as much. In this paper, we use K Means Clustering method to estimate the parameters of multinomial trees. By the weak convergence property of multinomial trees to continuous-time models, we show that this method can be in turn used to estimate parameters in continuous time models, illustrated by an example of jump-diffusion model.

1 Introduction

Since the seminal work by Black, Scholes and Merton on the geometrical Brownian motion model, various continuous time models were introduced as alternatives of the Black-Shcoles' model, such as Lévy pure-jump models, stochastic volatility models, and jump-diffusion models. These models were introduced to fix some unrealistic properties of the Black-Scholes' model, and have been successful in various degrees for the application to derivative pricing and hedging. On the other hand, an important practical problem about the estimation of parameters has not been addressed as extensively. A few exceptions are Ait-Sahalia [1], Ait-Sahalia and Kimmel [2], and

*Department of Mathematics, University of Louisville, Louisville, KY 40292, USA (kiseop.lee@louisville.edu)

†Department of Mathematics and Statistics, University of North Carolina at Charlotte, NC 28223, USA (mxu2@uncc.edu). The work of Mingxin Xu is supported by National Science Foundation under the grant SES-0518869 and John H. Biggs Faculty Fellowship.

Bhar, Chiarella, and To [4]. The existing methods mostly use maximum likelihood estimation and have turned out to be difficult to calculate and implement in jump models: in most cases clever numerical procedures are required in both estimating the likelihood function and finding the maximum.

Close relatives of these continuous time models in discrete time are multinomial trees. It is known that certain multinomial trees converge to continuous time models in distribution. Since the introduction of binomial trees as an approximation of the geometrical Brownian motion by Cox, Ross and Rubinstein [8], it has been popular in term structure modelling and other exotic derivative pricing. In fact, the estimation of the volatility parameter used in the geometrical Brownian motion model (σ) is fairly standard and is used for constructing the binomial tree ($u = e^{\sigma\sqrt{\Delta t}}, d = e^{-\sigma\sqrt{\Delta t}}$).

In this paper, we apply a simple but powerful statistical method called K Means Clustering to directly estimate the parameters in multinomial tree models. Then, using the weak convergence properties, we suggest that this method can be used to estimate parameters of continuous time models. The advantages of our approach in estimating the parameters in jump models are that it is a statistically well-established method and that it is easy to implement. We can avoid long numerical calculations, and instead use typical statistical softwares such as SAS and SPSS.

The paper is organized as follows. In section 2, we introduce the main problem. Section 3.1 explains the K Means Clustering Method. Section 3.2 explains how to use K Means Clustering to estimate parameters in the multinomial trees. In section 3.3, we find the parameter estimation in jump diffusion models that are the weak limits of the multinomial trees. We provide a couple of numerical examples in section 3.4. Section 4 concludes.

2 The Problem

Let us consider a multinomial tree with m time steps and k nodes at each time step. Formally, S_i denotes the price of the stock at time t_i, $i = 0, 1, ..., m$. The evolution of the stock prices process is

$$\frac{S_{i+1}}{S_i} = \xi, \quad i = 0, 1, ..., m-1, \tag{1}$$

where the multiplying factor ξ is a random variable that take different constant values with different probabilities as long as there is no arbitrage in the model, i.e., $\xi = \xi_j$ with probability p_j, $j = 1, 2, ..., k$.

A natural question arising is how to estimate ξ_j coupled with $p_j,\ j = 1, 2, ...k$. Since the multinomial tree is just an approximate model of the reality, we understand that statistically speaking, we observe the prices with some errors. In other words, what we really observe can be formulated as

$$\frac{S_{i+1}}{S_i} = \xi + \epsilon_i, \quad i = 0, 1, ..., m-1. \tag{2}$$

3 Main Result

3.1 K Means Clustering

When we have both input variables and output variables, we can build a model which explains the effects of inputs on outputs. Such a case is called supervised learning. On the other hand, if we have only outputs without inputs, then it becomes unsupervised learning. K Means Clustering is a popular unsupervised learning algorithm for finding clusters and cluster centers in a set of unlabeled data.[1] Suppose we already know that there are k different clusters, we use the following steps:

- Step 1 : Define k centers, one for each cluster.
- Step 2 : Each point is assigned to the cluster with the smallest distance.
- Step 3 : Once all points are assigned, recalculate the cluster centers.

We repeat Steps 1 to 3 until no more changes are done. In other words centers do not move any more. Detailed explanations of K Means Clustering Method and its implementation are given in standard textbooks such as Hastie et al. [11].

3.2 Application to the Multinomial Tree

Recall that we observe

$$\frac{S_{i+1}}{S_i} = \xi + \epsilon_i, , \quad i = 0, 1, ..., m-1, \tag{3}$$

where $\xi = \xi_j$ with probability p_j, and $j = 1, 2, ..., k$.

[1]There are other possible clustering methods such as Learning Vector Quantization and Gaussian Mixtures, and all of them, including K Means Clustering, have advantages and shortfalls. Interested readers may consult Hastie et al. [11].

We consider the data set composed of $\left\{\frac{S_{i+1}}{S_i}, i = 0, 1, 2, ...m-1\right\}$. Then estimating $\xi_j, j = 1, 2, ..., k$ is equivalent to finding k centers in K Means Clustering. Of course, a reasonable choice of initial centers is important. After finding the k centers, and assigning x_j number of points to the center ξ_j, we can estimate p_j with sample proportions $\hat{p}_j = \frac{x_j}{m}$.

3.3 Weak Convergence and Parameter Estimation in a Continuous Time Jump-Diffusion Model

For the simplest case, the convergence of Binomial approximation to the Black-Scholes model is studied by a classical work of Cox, Ross and Rubinstein [8]. There are other possible multinomial approximations derived from the PDE approach as shown in Heston and Zhou [12]. However, parameter estimation is fairly standard in this case and we will apply the K Means Clustering Method to a more interesting case of a simple jump-diffusion process.

A jump-diffusion model can be approximated by multinomial trees in a few different ways. One natural approach is through the PDE method explained in Chapters 2 and 3 of Prigent [15] and Chapter 3 of Clewlow and Strickland [6]. There exist extensive studies on numerical methods for the implementation of PDEs for option pricing such as the Finite Difference Method. In this section, we will adopt a trinomial tree from a direct approximation approach suggested by Nieuwenhuis and Vellekoop [14]. Let W_t be a standard Brownian motion, and N_t a Poisson process with constant intensity λ. The stock price process follows the stochastic differential equation

$$dS_t = S_t(\mu dt + \sigma dW_t + \alpha dN_t). \tag{4}$$

Suppose $\mu \in \mathbb{R}, \sigma > 0, \alpha \in \mathbb{R}, \lambda > 0$, and these are the parameters we are interested in estimating. On a finite-time interval $[0, T]$, define a stochastic processes $X_t^n = \begin{pmatrix} \tau_t^n \\ W_t^n \\ N_t^n \end{pmatrix}$ such that

$$X_t^n = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \text{ on } t \in \left[0, \tfrac{1}{n}\right); \quad X_t^n = \sum_{k=1}^{nT} \eta_k^n 1_{[T_k^n, T)}(t) \text{ on } t \in \left[\tfrac{1}{n}, T\right], \tag{5}$$

where $T_k^n = \frac{k}{n},\ k = 1, 2, ..., nT$, and

$$\eta_k^n = \begin{cases} \begin{pmatrix} \frac{1}{n} \\ \frac{1}{\sqrt{n}} \\ 0 \end{pmatrix} & \text{with probability } \frac{1}{2}(1 - \frac{\lambda}{n}); \\ \begin{pmatrix} \frac{1}{n} \\ -\frac{1}{\sqrt{n}} \\ 0 \end{pmatrix} & \text{with probability } \frac{1}{2}(1 - \frac{\lambda}{n}); \\ \begin{pmatrix} \frac{1}{n} \\ 0 \\ 1 \end{pmatrix} & \text{with probability } \frac{\lambda}{n}. \end{cases} \tag{6}$$

Then, X_t^n converges weakly to $\begin{pmatrix} t \\ W_t \\ N_t \end{pmatrix}$ on $[0, T]$ as n goes to ∞. Let us define

$$S_t^n = S_0 \exp\left\{(\mu - \frac{\sigma^2}{2})\tau_t^n + \sigma W_t^n + \ln(1+\alpha)N_t^n\right\} = S_0 \exp\left\{ \begin{pmatrix} \mu - \frac{\sigma^2}{2} \\ \sigma \\ \ln(1+\alpha) \end{pmatrix}' X_t^n \right\}.$$

Then S_t^n also converges weakly to S_t as n goes to ∞. Therefore, the three branches of the corresponding trinomial tree should be

- $\exp\left\{(\mu - \frac{\sigma^2}{2})\frac{1}{n} + \frac{\sigma}{\sqrt{n}}\right\}$ with probability $\frac{1}{2}(1 - \frac{\lambda}{n})$,
- $\exp\left\{(\mu - \frac{\sigma^2}{2})\frac{1}{n} - \frac{\sigma}{\sqrt{n}}\right\}$ with probability $\frac{1}{2}(1 - \frac{\lambda}{n})$,
- $\exp\left\{(\mu - \frac{\sigma^2}{2})\frac{1}{n} + \ln(1+\alpha)\right\}$ with probability $\frac{\lambda}{n}$

By the K Means Clustering Method introduced in Sections 3.1 and 3.2, we can estimate u, m, d as the centers ξ_1, ξ_2, ξ_3, along with probabilities p_1, p_2, p_3 for p_u, p_m, p_d from the data set $\left\{\frac{S_{i+1}}{S_i}, i = 0, 1, 2, ...nT - 1\right\}$. After ordering ξ_1, ξ_2, ξ_3 from high to low, we assign them to u, m, d. From these six statistics $\xi_1, \xi_2, \xi_3, p_u, p_m, p_d$, we need to back out the parameters for the continuous time model $\mu, \sigma, \alpha, \lambda$.

For this purpose, we need to decide which two branches correspond to the Brownian motion and which branch corresponds to the jump among u, m, d. It is important to notice that the probabilities associated to the Brownian movements is symmetric with $\frac{1}{2}(1 - \frac{\lambda}{n})$. Therefore, in the numerical example we show below, we will choose the two closest probabilities for the Brownian part. In this way, we let the market data tell us whether it is a positive jump or negative one and what the associated intensity is. For illustration purpose, suppose the estimators p_m, p_d are closer in value, then we assign

$$u = \exp\left\{(\mu - \tfrac{\sigma^2}{2})\tfrac{1}{n} + \ln(1+\alpha)\right\}, \qquad p_u = \tfrac{\lambda}{n}; \tag{7}$$

$$m = \exp\left\{(\mu - \tfrac{\sigma^2}{2})\tfrac{1}{n} + \tfrac{\sigma}{\sqrt{n}}\right\}, \qquad p_m = \tfrac{1}{2}(1 - \tfrac{\lambda}{n}); \tag{8}$$

$$d = \exp\left\{(\mu - \tfrac{\sigma^2}{2})\tfrac{1}{n} - \tfrac{\sigma}{\sqrt{n}}\right\}, \qquad p_d = \tfrac{1}{2}(1 - \tfrac{\lambda}{n}). \tag{9}$$

From the three equations involving u, m, d in (7)-(9), we can solve for μ, σ, α:

$$\mu = \frac{n}{2}\ln(md) + \frac{n}{8}\left(\ln\frac{m}{d}\right)^2, \tag{10}$$

$$\sigma = \frac{\sqrt{n}}{2}\ln\frac{m}{d}, \tag{11}$$

$$\alpha = \frac{u}{\sqrt{md}} - 1. \tag{12}$$

The estimation of λ is not so straightforward and we will apply the Maximum Likelihood Method. Let X_u, X_m, X_d be numbers of ups, middles, and downs respectively. Then (X_u, X_m, X_d) follows the trinomial distribution with density function

$$P(X_u = x_u, X_m = x_m, X_d = x_d) = \binom{nT}{x_u\, x_m\, x_d}\left(\tfrac{\lambda}{n}\right)^{x_u}\left(\tfrac{1}{2}(1-\tfrac{\lambda}{n})\right)^{x_m}\left(\tfrac{1}{2}(1-\tfrac{\lambda}{n})\right)^{x_d},$$

where $x_u + x_m + x_d = nT$. We obtain the maximum likelihood estimator of λ by finding maximum of the likelihood function $L(\lambda) = P(X_u = x_u, X_m = x_m, X_d = x_d)$:

$$\hat{\lambda} = \frac{x_u}{T}.$$

Note that $p_u = \frac{x_u}{nT}$ and this is exactly

$$\hat{\lambda} = np_u = n(1 - p_m - p_d),$$

by checking the three equations involving λ from (7)-(9).

To achieve higher precision, we can allow different jump sizes where the stock price process is driven by a compound Poisson process and can be represented as

$$dS_t = S_t(\mu dt + \sigma dW_t + \sum_{i=1}^{k}\alpha_i dN_{i,t}), \tag{13}$$

where each $N_{i,t}$ is a standard Poisson process with constant intensity λ_i for $i = 1, ..., k$. The total intensity of the compound Poisson process is $\lambda = \sum_{i=1}^{k}\lambda_i$. The corresponding multinomial tree will have $k+2$ branches in each time step, and we can still define $X_t^n = \begin{pmatrix} \tau_t^n \\ W_t^n \\ N_{1,t}^n \\ \cdots \\ N_{k,t}^n \end{pmatrix}$ with (5), and the expanded

$$\eta_j^n = \begin{cases} \begin{pmatrix} \frac{1}{n} \\ \frac{1}{\sqrt{n}} \\ 0 \\ \cdots \\ 0 \end{pmatrix} & \text{with probability } \frac{1}{2}(1-\frac{\lambda}{n}), \\ \begin{pmatrix} \frac{1}{n} \\ -\frac{1}{\sqrt{n}} \\ 0 \\ \cdots \\ 0 \end{pmatrix} & \text{with probability } \frac{1}{2}(1-\frac{\lambda}{n}), \\ \begin{pmatrix} \frac{1}{n} \\ 0 \\ \cdots \\ 1 \\ \cdots \\ 0 \end{pmatrix} & \text{with probability } \frac{q_i\lambda}{n},\ i=1,2,...,k, \end{cases} \tag{14}$$

where 1 is placed at $(i+2)$th component, and $q_i = \frac{\lambda_i}{\lambda}$. Then the resulting discrete process

$$S_t^n = S_0 \exp\left\{ (\mu - \frac{\sigma^2}{2})\tau_t^n + \sigma W_t^n + \sum_{i=1}^{k} \log(1+\alpha_i) N_{i,t}^n \right\}$$

converges weakly to the continuous version (13), and the corresponding branches in the $k+2$-multinomial tree are

- $\exp\left\{ (\mu - \frac{\sigma^2}{2})\frac{1}{n} + \frac{\sigma}{\sqrt{n}} \right\}$ with probability $\frac{1}{2}(1-\frac{\lambda}{n})$,
- $\exp\left\{ (\mu - \frac{\sigma^2}{2})\frac{1}{n} - \frac{\sigma}{\sqrt{n}} \right\}$ with probability $\frac{1}{2}(1-\frac{\lambda}{n})$,
- $\exp\left\{ (\mu - \frac{\sigma^2}{2})\frac{1}{n} + \ln(1+\alpha_i) \right\}$ with probability $\frac{q_i\lambda}{n}$, $i=1,2,...,m$.

There are $k+2$ equations that relate the $k+2$ estimated centers ξ_i in the K Means Clustering Method to solve for the $k+2$ parameters $\mu, \sigma, \alpha_1, ..., \alpha_k$. There are additional $k+2$ equations that relate estimated probabilities $p_1, .., p_{k+2}$ associated to the centers ξ_i by $p_i = \frac{x_i}{nT}$ to estimate the k intensities λ_i through q_i. Since the probabilities p_i sum up to 1, we have one more degree of freedom than the number of variables we would like to solve. As in the trinomial model, the maximum likelihood estimators of the intensities are

$$\hat{\lambda}_i = \frac{x_i}{T},$$

which turns out to be a feasible solution that satisfies all $k+2$ constraints.

node	frequency (x_i)	estimated probability (p_i)	estimated center (ξ_i)
up	35	0.2846	1.0116
middle	79	0.6423	0.9986
down	9	0.0731	0.9825

Table 1: Trinomial approximation of IBM daily closing prices.

node	frequency (x_i)	estimated probability (p_i)	estimated center (ξ_i)
up	42	0.3717	1.0152
middle	55	0.4472	0.9979
down	16	0.1301	0.9779

Table 2: Trinomial approximation of S&P 500 index weekly closing prices.

3.4 Examples

3.4.1 Trinomial Approximation of the IBM Daily Stock Price

In this subsection, we illustrate the K Means Clustering approximation method by using the daily closing prices of the IBM stock from Oct 2, 2006 to March 31, 2007. There were 124 trading days, so we observed 123 data points for $\frac{S_{t+1}}{S_t}$. We use trinomial tree as our approximation model. Table 1 shows the results for K Means Clustering with three centers.

The corresponding parameters for the trinomial trees is $u = 1.0116$ with probability $p_u = 0.2846$, $m = 0.9986$ with probability $p_m = 0.6423$ and $d = 0.9825$ with probability $p_d = 0.0731$. Percentage-wise, the IBM stock moves up by 1.16% daily, moves down by 0.14% or 1.75%.

3.4.2 Jump Diffusion Approximation of the S&P 500 Index

Next, we approximate the S&P500 index weekly closing data with jump-diffusion model (4) and estimate its parameters through a trinomial tree parameter estimation from K Means Clustering Method. We collected data form the first week of January 2005 to the last week of March 2007. There are 114 weeks in total, which gives 113 data points for $\frac{S_{t+1}}{S_t}$. The time unit is a year, so $n = 52$. Table 2 shows the result of K Means Clustering Method. Since the up and middle frequencies are closer than the down frequency, we will assign the 'down' branch to the jump. The parameters for the jump diffusion model is thus estimated as

$$
\begin{aligned}
\sigma &= \frac{\sqrt{n}}{2} \ln \frac{u}{m} = 0.0620, \\
\mu &= \frac{n}{2} \ln(um) + \frac{n}{8} \left(\ln \frac{u}{m} \right)^2 = 0.3395, \\
\alpha &= \frac{d}{\sqrt{um}} - 1 = -0.0284, \\
\lambda &= \frac{x_d}{T} = 7.3628.
\end{aligned} \tag{15}
$$

This implies a 2.84% downward jump of S&P with a frequency of about 7.36 times per year, while the volatility coming from the Brownian motion is 6.2%.

4 Conclusion

We have studied how to estimate parameters in multinomial tree models using the K Means Clustering Method. This is a simple, but powerful statistical method which can be easily done by standard software such as SAS and SPSS. This method was then applied to parameter estimation in continuous time jump-diffusion models as explained in Section 3.3 and an example in Section 3.4.2. However, this methodology can be applied more widely to other continuous time models that are weak limits of multinomial trees. For stochastic volatility models the reader can consult Ait-Sahalia and Kimmel [2] and Florescu and Viens [10] about their approximation by multinomial models.

References

[1] Ait-Sahalia, Y.: Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach, Econometrica, **70**, 223-262 (2002)

[2] Ait-Sahalia, Y., Kimmel, R.: Maximum likelihood estimation of stochastic volatility models, Journal of Financial Economics, **83**, 413-452 (2007)

[3] Albanese, C., Kuznetsov, A.: Discretization schemes for subordinated processes, working paper (2007)

[4] Bhar, R., Chiarella, C., To, T.: A maxium likelihood approach to estimation of Heath-Jarrow-Morton model, preprint

[5] Chang, L. B., Palmer, K.: Smooth convergence in the binomial model, Finance and Stochastics **11**, 91–105 (2007)

[6] Clewlow, L., Strickland, C.: Implementing Derivatives Models, Wiley Financial Engineering, (1999)

[7] Ramezani, C., Zeng, Y.: An Empirical Assessment of Double Exponential Jump-Diffusion Process, Forthcoming in Annals of Finance

[8] Cox, J. C., Ross, S. A., Rubinstein, M.: Option pricing: a simple approach, The Journal of Financial Economics **7**, 229–263 (1979)

[9] Durham, G., Gallant, R.: Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes, preprint

[10] Florescu, I., Viens, F.: Stochastic volatility: option pricing using a multinomial recombining tree, preprint (2006)

[11] Hastie, T., Tibshirani, R., Freidman, J.: The Elements of Statistical Learning, Springer Series in Statistics (2001)

[12] Heston, S., Zhou, G.: On the rate of convergence of discrete-time contingent claims, Mathematical Finance **10**, 53–75 (2000)

[13] Li, M., pearson, N., Posteshman, A.: Conditional estimation of diffusion processes, Journal of Financial Economics **74**, 31-66 (2004)

[14] Nieuwenhuis, J. W., Vellekoop, M. H.: Weak convergence of tree methods, to price options on defaultable assets, Decisions in Economics and Finance **27**, 87–107 (2004)

[15] Prigent, J-L.: Weak Convergence of Financial Markets, Springer Finance, (2003)