Journal of the Korean Data & Information Science Society
2011, **22**(3), 549–553

한국데이터정보과학회지

Deciding a sampling length for estimating the parameters in Geometric Brownian Motion

Junmo Song[1]

[1]KYOBO Securities Co., Ltd.

Received 7 February 2011, revised 25 March 2011, accepted 30 March 2011

Abstract

In this paper, we deal with the problem of deciding the length of data for estimating the parameters in geometric Brownian motion. As an approach to this problem, we consider the change point test and introduce simple test statistic based on the cumulative sum of squares test (cusum test). A real data analysis is performed for illustration.

Keywords: Cusum test, geometric Brownian motion, sampling length.

1. Introduction

Since Black and Scholes (1973) used geometric Brownian motion (GBM) to derive the famous option-pricing formula, this process has been widely used in describing the dynamics of underlying assets in finance. Its easy implementation and convenience in estimating the parameters have most practitioners prefer to work with GBM for modeling their underlying assets, though a large number of financial models have been developed; see Merton (1976), Heston (1993), Heston and Nandi (2000), Schoutens (2003) and references therein.

In this paper, we deal with the problem of deciding the length of data for estimating the parameters in GBM. Since too long-period data tends to overlook recent trend, it is significant to use data of proper length, especially in financial data which often suffer from structural changes in underlying models owing to changes of monetary policy and critical social events.

As an approach to this problem, we consider the change point problem since the data from the last change point to the last observed point can be regarded as a good alternative for sampling length. For a review of the change point problem, see e.g. Hinkley (1971), Inclan and Tiao (1994) and Csörgö and Horváth (1997). In particular for time series models, we refer to Lee *et al.* (2003). Recently, Gregorio and Iacus (2008), Song and Lee (2009) and Lee *et al.* (2010) considered the parameter change test for diffusion processes based on discretely observed sample. Particularly, their papers focused on the parameter change in diffusion coefficient. Since the volatility of the assets plays a crucial role in finance, their studies are noteworthy, especially in application to interest rate models, such as the Vasicek

[1] Manager, OTC Products Dealing Team, KYOBO Securities Co., Ltd., Youido-Dong, Yongdungpo-Gu, Seoul 150-737, Korea. E-mail: junmo.song@gmail.com

model and the Cox-Ingersoll-Ross model. However, there are a few limits in applying their methods to GBM. For example, Song and Lee (2009)'s method can not be used directly, because GBM, or its log-transformed process, is not an ergodic process.

In this paper, we introduce simple procedure to detect the change point in GBM by employing the cusum test statistics. The proposed method can test whether both of the parameters in drift and diffusion coefficients changed or not. In Gregorio and Iacus (2008), Song and Lee (2009) and Lee *et al.* (2010), sampling interval needs to converge to zero as number of data increases in order for obtaining the asymptotic distribution of each test statistic. On the other hand, our test does not impose this condition. In some cases, the condition could be a restriction in practical applications.

The paper is organized as follows. Section 2 introduces cusum statistics for GBM and a real data analysis on KOSPI200 is presented in Section 3.

2. Cusum test for deciding sampling length

Let us consider the following geometric Brownian motion S_t:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{2.1}$$

where μ and σ are unknown parameters and $\{W_t : t >= 0\}$ is a standard Wiener process.

We intend to find the last change point by using the cusum statistics based on discretely observed sample, say, S_{t_i}, $t_i = i\Delta t$, $i = 1, \cdots, n$. Here we emphasize that Δt is not necessary to converge to zero as $n \to \infty$. Δt is assumed to be fixed. Note that the solution of stochastic differential equation (2.1) can be represented as

$$\log(S_{t_i}/S_{t_{i-1}}) = (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma(W_{t_i} - W_{t_{i-1}}). \tag{2.2}$$

Due to the properties of Wiener process, we can see that $\log(S_{t_i}/S_{t_{i-1}})$'s are independent and identical random variables with Normal distribution.

Now, let $X_1, \cdots, X_n$ be a sequence of i.i.d. random variables with finite fourth moment. Then it is well known that

$$\frac{1}{\sqrt{n \cdot var(X_1^2)}} \max_{1\le k\le n} \left| \sum_{i=1}^{k} X_i^2 - \frac{k}{n}\sum_{i=1}^{n} X_i^2 \right| \to_d \sup_{0\le t\le 1}|B_t^0| \text{ as } n \to \infty, \tag{2.3}$$

where $\{B_t^0\}$ is the Brownian bridge; see Lee *et al.* (2010). Since the previous sequence $\{\log(S_{t_i}/S_{t_{i-1}})\}$ satisfies the conditions for establishing (2.3), we have the following theorem.

Theorem 2.1 Let $\{S_t : t >= 0\}$ be the price process from the GBM in (2.1) and S_{t_i} be the observation at time t_i, where $t_i = i\Delta t$ and $i = 1, \cdots, n$. If sampling interval Δt is fixed and the parameters μ and σ do not change, then we have

$$T_n := \frac{1}{\sqrt{n}\hat{\tau}_n} \max_{1\le k\le n} \left| \sum_{i=1}^{k} r_i^2 - \frac{k}{n}\sum_{i=1}^{n} r_i^2 \right| \to_d \sup_{0\le t\le 1}|B_t^0| \text{ as } n \to \infty,$$

where $r_i = \log(S_{t_i}/S_{t_{i-1}})$, $\widehat{\tau_n^2} = \sum_{i=1}^n r_i^4/n - (\sum_{i=1}^n r_i^2/n)^2$ and $\{B_t^0\}$ is the Brownian bridge.

If T_n is large, we consider that there is a change in the parameters μ and σ during the observation, and then we estimate the change point to be the point that maximizes $|\sum_{i=1}^k r_i^2 - k/n \sum_{i=1}^n r_i^2|$ over k. As mentioned earlier, we decide a sampling length to be from the last change point to the last observed point.

Remark 2.1. The above theorem holds for the process driven by $dS_t = \mu S_t dt + \sigma S_t dL_t$, where $\{L_t : t >= 0\}$ is a Levy process of which distribution has finite fourth moment.

Remark 2.2. In contrast to Gregorio and Iacus (2008), Song and Lee (2009) and Lee *et al.* (2010), any estimator for μ and σ is not necessary for building the test statistics.

3. Real data analysis

In this section, we provide a real data analysis. The data employed is KOSPI200 covering the period ranging from February 28, 2005 to February 25, 2011, for a total of 1494 observations, and is assumed to be generated from the GBM in (2.1) with $\Delta t = 1/252$, where 252 is the average trading days per year. We perform cusum test at a nominal level 0.1, based on log-return series, $r_i = \log(S_{t_i}/S_{t_{i-1}})$, to find the last change point. The empirical quantile value for this nominal level is 1.488; the empirical quantile values for $\sup_{0\le t\le 1}|B_t^0|$ are provided in Table 1 of Lee *et al.* (2003).

The results are presented in Table 3.1. We find seven change points during the period and the last change point is found to be on Feb. 25, 2010. Thus, as one of proper sampling lengths for estimating μ and σ, we can recommend the observations from Feb. 26, 2010 to Feb. 25, 2011. In this period, μ and σ are estimated to be 0.236 and 0.15 respectively.

Table 3.1 Estimated change points, T_n, σ, and μ

No.	Period			T_n	σ	μ
1	2005-03-02	∼	2006-01-16	1.644	17.1%	39.2%
2	2006-01-17	∼	2006-07-20	2.026	22.8%	-17.3%
3	2006-07-21	∼	2007-07-12	3.278	14.2%	41.0%
4	2007-07-13	∼	2008-02-14	1.785	31.1%	-15.8%
5	2008-02-15	∼	2008-08-29	3.892	21.2%	-22.9%
6	2008-09-01	∼	2009-03-04	2.108	53.5%	-48.2%
7	2009-03-05	∼	2010-02-25	1.846	21.6%	43.5%
8	2010-02-26	∼	2011-02-25		15.0%	23.6%
Total	2005-03-02	∼	2011-02-25		24.7%	14.6%

In order to compare the above volatilities with the result of other volatility model, we perform an additional analysis of KOSPI200 using GARCH(1,1) model. GARCH volatilities, say $\sqrt{\mathrm{V_i^G}}$, are specified as follows:

$$V_i^G \Delta t = w + \alpha r_{i-1}^2 + \beta V_{i-1}^G \Delta t, \tag{3.1}$$

where $r_i = \ln(S_{t_i}/S_{t_{i-1}}) = \sqrt{V_i^G \Delta t}\epsilon_i$ and $\{\epsilon_i\}$ is a sequence of *i.i.d* random variables with zero mean and unit variance. Δt is an sampling interval, and here we take 1/252. Table 3.2 presents the estimation results of parameters in (3.1), and Figure 3.1 shows KOSPI200 series, GARCH volatilities and the estimated σ, that is volatility of GBM, in each subdivided interval. As seen in Figure 3.1, each estimated volatility of GBM in subinterval seems to behave like a local mean-reverting level of GARCH volatilities. Thus, we can check indirectly that the whole period is subdivided adequately.

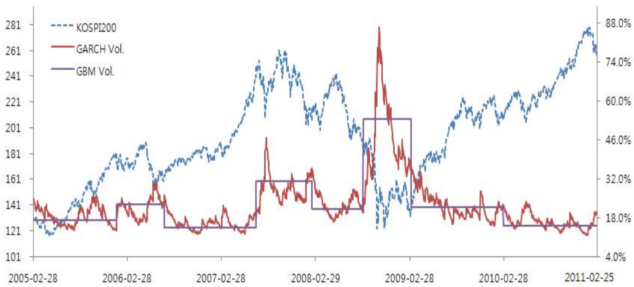

Figure 3.1 KOSPI200 series, GARCH volatilities, and GBM volatilities from Feb. 28, 2005 to Feb. 25, 2011

Table 3.2 Estimation results of GARCH parameters

	Estimate	Std. Error	t value	Pr(> | t |)
w	2.646e-06	8.884e-07	2.978	0.002898
α	7.538e-02	1.161e-02	6.493	8.4e-11
β	9.123e-01	1.278e-02	71.385	< 2e-16

4. Concluding remarks

In this paper, we propose the procedure for deciding a sampling length for estimating the parameters in GBM. This procedure includes the change point test for finding the last change point, and we employ cusum test among other change point tests. The proposed test is simple to implement because any estimator is not needed in constructing the test statistics. It is expected that this method would be helpful in finance, since many financial assets are still modelled by GBM. For instance, our method could be a solution for deciding a sampling length for estimating the volatility in GBM, which is an important issue in valuation of derivatives.

References

Black, F. and Scholes, M. S. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy,* **81**, 637-654.

Cs-rg-, M. and Horv-a- th, L. (1997). *Limit theorems in change-point analysis*, Wiley, New York.

Gregorio, A. and Iacus, S. (2008). Least squares volatility change point estimation for partially observed diffusion processes. *Communications in Statistics: Theory and Methods*, **37**, 2342-2357.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, **6**, 327-343.

Heston, S. L. and Nandi, S. (2000). A closed-form GARCH option valuation model. *The Review of Financial Studies*, **13,** 585-625.

Hinkley, D. V. (1971). Inference about the change-point from cumulative sum tests. *Biometrika*, **58**, 509-523.

Inclan, C. and Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of change of variance. *Journal of the American Statistical Association*, **89**, 913-923.

Lee, S., Ha, J., Na, O. and Na, S. (2003). The cusum test for parameter change in time series models. *Scandinavian Journal of Statistics*, **30**, 781-796.

Lee, S., Lee, T. and Na, O. (2010). Cusum of squares test for discretely observed sample from diffusion processes. *Journal of the Korean Data & Information Science Society*, **21**, 179-183.

Na, O., Ko, B. and Lee, S. (2010). Cusum of squares test for discretely observed sample from multidimensional diffusion processes. *Journal of the Korean Data & Information Science Society*, **21**, 547-554.

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, **3**, 125-144.

Park, S. and Lee, S. (2006). Simulation study on the scale change test for autoregressive models with heavy-tailed innovations. *Journal of the Korean Data & Information Science Society*, **17**, 1397-1403.

Park, S. and Lee, S. (2007). Modelling KOSPI200 data based on GARCH(1,1) parameter change test. *Journal of the Korean Data & Information Science Society*, **18**, 11-16.

Schoutens, W. (2003). *Levy processes in finance: Pricing financial derivatives*, John Wiley & Sons Ltd.

Song, J. and Lee, S. (2009). Test for parameter change in discretely observed diffusion processes. *Statistical Inference for Stochastic Processes*, **12**, 165-183.