Threshold estimation for stochastic differential equations with jumps

Yasutaka Shimizu*
Osaka University, Osaka, JAPAN
E-mail: yasutaka@sigmath.es.osaka-u.ac.jp

Abstract

In recent years, the statistical inference for discretely observed jump processes is an important issue in finance and insurance. Due to the discreteness, it is unclear that an increment of neighboring data essentially comes from continuous or discontinuous shocks, which causes some difficulties for estimating unknowns in the underlying process. *Threshold Estimation* is one of the useful techniques to disentangle the continuous shocks and real jumps, where if an increment of data is smaller than a predetermined threshold, then we regard the increment does not include any jump. However, the choice of the threshold has optionality in practice, and the standard way has not been established yet. In this paper, we shall propose how to select some "optimal" thresholds from given data, and study the finite-sample performance by simulations.

Keywords: Jump-type processes, discretely observations, threshold estimation, threshold selection.

1 Introduction

Consider a (1-dim) stochastic process $X = (X_t)_{t\geq 0}$ satisfying a stochastic differential equation

$$dX_t = a(X_t, \mu)\, dt + b(X_t, \sigma)\, dW_t + c(X_{t-}, \delta) dZ_t, \tag{1}$$

where a, b and c are suitable functions with parameters (μ, σ, δ), W is a Wiener process, and Z is a compound Poisson process with intensity λ and a jump-distribution F. Suppose a situation that the process X is observed at discrete time points $t_i^n = i\Delta_n\, (i = 0, 1, \dots, n)$ for $\Delta_n > 0$. In the sequel, we suppose that $\Delta_n \to 0$ as $n \to \infty$.

Such a jump-type stochastic process is recently a standard tool, e.g., for modeling asset values in finance and insurance, and as a practical demand, we need to specify parameters $(\mu, \sigma, \delta, \lambda)$ and F from discrete samples of X. A useful technique to execute those inference is the *threshold estimation*, which is proposed by Mancini (2001) and also by Shimizu (2002), independently. The fundamental idea is doing the following judgement: for $\Delta_i^n X := X_{t_i^n} - X_{t_{i-1}^n}$ and a predetermined threshold $r_n > 0$,

$$\begin{aligned} |\Delta_i^n X| \leq r_n \quad &\Rightarrow \quad \text{No 'large' jump ocurred in } (t_{i-1}^n, t_i^n]; \\ |\Delta_i^n X| > r_n \quad &\Rightarrow \quad \text{A 'large' jump ocurred in } (t_{i-1}^n, t_i^n], \end{aligned}$$

and use $\{\Delta_i^n X\,;\, |\Delta_i^n X| \leq r_n\}$ to estimate a continuous part of X, and use $\{\Delta_i^n X\,;\, |\Delta_i^n X| > r_n\}$ to estimate jump part. see Shimizu and Yoshida (2006), or Ogihara and Yoshida (2011). The similar idea is also used by various authors in different contexts; see, e.g., Aït-Sahalia *et al.* (2009a,b,2011,2012), Gobbi and Mancini (2008), Cont and Mancini (2011), among others.

From a viewpoint of asymptotic theory, it is desirable that, e.g., $r_n = O(\Delta_n^\varpi)$ for $\varpi < 1/2$ as $\Delta_n \to 0$ to disentangle jumps and diffusion shocks, so most authors consider thresholds of the form $r_n = \alpha\Delta_n^\varpi$ for a constant $\alpha > 0$. However, how to choose constants (α, ϖ) is a big problem because most statistics based on this idea are very sensitive with choosing r_n. Some authors propose (α, ϖ), sometimes data-adaptively, via simulations, or *ad hoc* methods, but a standard way has not established yet. We shall propose an automatic algorithm to choose r_n from given data along the results by Shimizu (2010).

2 Thresholds via a bias correction

Consider the following statistic:

$$\widehat{IV}(r_n) := \sum_{i=1}^{n} (\Delta_i^n X)^2 \mathbf{1}_{\{|\Delta_i^n X| \le r_n\}},$$

which is a *threshold estimator* of the *integrated volatility* $\int_0^{T_n} b^2(X_t, \sigma)\,dt$; see, e.g., Mancini (2009). The idea to choose $r = r_n$ is to make the bias of $\widehat{IV}(r)$ as small as possible. Let

$$B_n(r) := \mathbb{E}\left[\widehat{IV}(r) - \int_0^{T_n} b^2(X_t)\,dt\right]$$

We say that $\widehat{IV}(r_n)$ is *asymptotically unbiased* if $B_n(r_n) \to 0$ as $n \to \infty$.

Theorem 1. *Suppose that* a *and* b *are of linear growth, and that* $\lim_{n\to\infty} \sup_{t\le T_n} \mathbb{E}|X_t|^4 < \infty$. *Moreover suppose that*

$$\lim_{n\to\infty} \sqrt{T_n}\left(r_n + \frac{\Delta_n}{r_n}\right) = 0. \tag{2}$$

Then $\widehat{IV}(r_n)$ *is asymptotically unbiased.*

Remark 1. *The condition* (2) *implies that, in finite activity case, taking* $r_n = O(\Delta_n^{\varpi})$ *with* $\varpi < 1/2$ *is enough as* $T_n \equiv T$ *(fixed). As* $T_n \to \infty$*, we further need* $n\Delta_n^{1+\delta} \to 0$ *for some* $\delta \in (0, 1/2)$*. These facts are consistent with the results in earlier works; e.g., Mancini (2009), or Shimizu and Yoshida (2006).*

Theorem 2. *Suppose the same assumptions as in Theorem 1, and that* $\Delta_n \to 0$*. Then, for each* $r > 0$*,*

$$\left|B_n(r) - e^{-\lambda\Delta_n}\mathbb{E}[\widetilde{B}_n(r)]\right| = O(n\Delta_n^{3/2}), \quad n \to \infty,$$

where

$$\widetilde{B}_n(r) := \sum_{i=1}^{n}\left[I_i^n(r) - \int_{|z|>r} |z|^2 \Phi(dz; \Sigma_i^n)\right], \tag{3}$$

$\Phi(z; \Sigma)$ *is a Gaussian distribution function with mean zero and variance* Σ*, and*

$$\Sigma_i^n := \int_{t_{i-1}^n}^{t_i^n} b^2(X_s, \sigma)\,ds, \quad I_i^n(r) := \lambda\Delta_n \int_{|c(X_{t_{i-1}^n}, \delta)z| \le r} |c(X_{t_{i-1}^n}, \delta)z|^2 F(dz).$$

Due to the above result, $B_n \approx 0$ if $\widetilde{B}_n \approx 0$ when $n\Delta_n^{3/2} \to 0$. Now, our aim is to find $r = r_n$ that minimizes $\tilde{B}_n(r)$ in (3):

$$r_{opt} := \arg\min_{r\ge 0} |\widetilde{B}_n(r)|$$

However, $\widetilde{B}_n(r)$ still has some unknown parameters $(\mu, \sigma, \delta, \lambda, F)$, which are also to be estimated via a suitable threshold.

Let $\widehat{\Sigma}_i^n(r_n)$ and $\widehat{I}_i^n(r; r_n)$ be estimators of Σ_i^n and $I_i^n(r)$ with parameters replaced with some threshold estimators: e.g., $\widehat{\Sigma}_i^n(r_n) == \int_{t_{i-1}^n}^{t_i^n} b^2(X_s, \widehat{\sigma}_n(r_n))\,ds$ where $\widehat{\sigma}_n(r_n)$ is a threshold estimator of σ. How to construct threshold estimators, see, e.g., Shimizu and Yoshida (2006), Mancini (2009), Shimizu (2009), etc. Using such *plug-in estimators* $\widehat{\Sigma}_i^n(r_n)$ and $\widehat{I}_i^n(r; r_n)$, we propose the following algorithm:

Plug-in algorithm (Shimizu, 2010)

[Step 0] Choose a pilot threshold $r_n^{(0)} < \max_{1\le i\le n} |\Delta_i^n X|$, and calculate $\widehat{\Sigma}_i^n(r_n^{(0)})$ and $\widehat{I}_i^n(r; r_n^{(0)})$;

[Step *k*] For $k = 1, 2, \ldots$, iterate the following steps:

[Step k]-1: Calculate $\widehat{\Sigma}_i^n(r_n^{(k-1)})$ and $\widehat{I}_i^n(r; r_n^{(k-1)})$;

[Step k]-2: Find the root $r = r_n^{(k)}$ to the following equation:

$$\sum_{i=1}^n \left[\widehat{I}_i^n(r; r^{(k-1)}) - \int_{|z|>r} |z|^2 \Phi\left(dz; \widehat{\Sigma}_i^n(r^{(k-1)})\right) \right] = 0.$$

Iterate [Step k] until $\{r_n^{(k)}\}_{k\in\mathbb{N}}$ 'converges' to a constant r^*, which is used as a threshold.

Remark 2. *We will numerically check that* $r_n^{(k)} \to r^*$ *as* $k, n \to \infty$.

3 A modification of estimators of λ and F

For simplicity, suppose that $c(x, \delta) \equiv 1$, and that a distribution F is parametrized with $\zeta \in \mathbb{R}$:

$$F(dz) = F_\zeta(dz), \quad \zeta := \int_{\mathbb{R}} G(z) F_\zeta(dz) \quad \text{for a known function } G.$$

Then threshold estimators of λ and ζ is given by (e.g., Shimizu, 2009)

$$\widehat{\lambda}(r) = \frac{1}{T_n}\sum_{i=1}^n \mathbf{1}_{\{|\Delta_i^n X|>r\}}, \quad \widehat{\zeta}(r) = \frac{1}{\widehat{\lambda}(r) T_n}\sum_{i=1}^n G(\Delta_i^n X)\mathbf{1}_{\{|\Delta_i^n X|>r\}},$$

where $\widehat{\delta}(r)$ is also a threshold estimator of a parameter δ. Those are consistent estimators of λ and ζ, respectively, under some regularities as $n \to \infty$ and $r_n \to 0$ with a suitable rate. However, in practice, we construct estimators for given (fixed) $r = r_n > 0$, and 'small' jumps less than r_n are cut off. Hence the rate of expected information loss of $\widehat{\lambda}(r_n)$ seems approximately

$$\mathbb{E}\left[\widehat{\lambda}(r_n)/\lambda\right] \approx \int_{|z|>r_n} F_\zeta(dz) =: e_n.,$$

that is, $\widehat{\lambda}^* = e_n^{-1}\widehat{\lambda}_n(r_n)$ and $\widehat{\zeta}^* = e_n \widehat{\zeta}(r_n)$ would be compensate the information loss. The value of e_n is estimated as the root of the following estimated equation in $e \in (0, 1]$: for given threshold $r_n > 0$,

$$e = \int_{|z|>r_n} F_{e\widehat{\zeta}(r_n)}(dz). \tag{4}$$

4 Simulation

We try the algorithm in Section 2 as well as a modification in Section 3 by a simple O-U model as follows:

$$dX_t = -\mu X_t\, dt + \sigma\, dW_t + dZ_t^{(\lambda,\alpha,\beta)}, \quad X_0 = 0,$$

where Z is a compound Poisson with intensity λ and Gaussian jumps with mean α and variance β. For this model, we can construct *MLE-type threshold estimator* for all the parameters $(\mu, \sigma, \lambda, \alpha, \beta)$ according to Shimizu and Yoshida (2006). We compare our algorithm with several earlier methods: Lee and Mykland (2008) propose a *jump-detecting test* in each $(t_{i-1}^n, t_i^n]$ via the extreme value theory. Their method corresponds to use a data-adaptive threshold $r = r_i^n \in \mathcal{F}_{t_i^n}$ in $(t_{i-1}^n, t_i^n]$. Aït-Sahalia and Jacod (2009b) propose to use $r_n = \alpha\widehat{\sigma}\Delta_n^{\varpi}$ with, e.g., $3 \le \alpha \le 5$ and $\varpi = 0.47$ or 0.48, and $\widehat{\sigma}$ is an estimator of σ *without* using a threshold, for example, use a bipower variation etc.; see, e.g., Barndorff-Nielsen and Shephard (2004). Shimizu (2008) proposes an algorithm to choose r_n via a bias correction of $\widehat{\lambda}(r_n)$. This is also an automatic way to choose the threshold. In the end, we shall compare the following candidates. The symbol `*` means the result with a modification as in Section 3.

- `LM`: Test by Lee and Mykland (2008) with significance level of 5%[1].
- `S08`: By Shimizu (2008). The threshold is chosen so that the bias of $\widehat{\lambda}(r_n)$ is as small as possible.
- `AS-/AS*-`: $r_n = 3\widehat{\sigma}_B\Delta_n^{0.48}$ without/with the modification, where $\widehat{\sigma}_B^2$ is the realized bipower variation. This is the minimum threshold used by Aït-Sahalia and Jacod (2009b).
- `AS+/AS*+`: $r_n = 5\widehat{\sigma}_B\Delta_n^{0.47}$ without/with the modification, which is the maximum threshold used by Aït-Sahalia and Jacod (2009b).
- `S10/S10*`: Our method without/with the modification. The threshold is chosen so that the bias of $\widehat{IV}(r_n)$ is as small as possible.

We set the true value as $(\mu, \sigma, \lambda, \alpha, \beta) = (0.2, 0.7, 10.0, 0.0, 0.5)$. Table 1 shows means and standard deviations (s.d.) of each estimator through 5000 times experiments in the case where $n = 3000$, $\Delta_n = 1/250$, and Figure 1 shows relative errors of estimators of σ by `AJ*`$\pm$ and our proposal `S10*` for different values of $\sigma = 0.05 \sim 0.8$. We see that our proposal `S10*` always choose a suitable threshold such that the estimation of σ is so nice. From those results, we see a superiority of our automatic 'plug-in' algorithm.

5 Conclusions

- Threshold estimators are very sensitive for a choice of a threshold. They are some *ad hoc* proposals to select a threshold, which highly depend on the true model. Therefore we need to establish a way to choose it suitably from given data without optionality.
- Our method can automatically determine a suitable threshold that can give us accurate estimation of parameters, if needed, via a suitable *modification* in Section 3.
- If the diffusion component is $b(x, \sigma) \equiv 0$, the estimator $\widehat{IV}(r_n)$ with a threshold given by our method becomes close to zero, which would be available for testing the existence of a jump from discrete data to observe an estimator of a diffusion term.
- In this paper, we used a model with compound Poisson jumps, but our method can be applied to models with infinite activity jumps in principle; Shimizu (2010) reports the details with numerical experiments.

References

[1] Aït-Sahalia, Y. and Jacod, J. (2009a). Estimating the degree of activity of jumps in high frequency data, *Ann. Statist.*, **37**, 2202–2244.

[2] Aït-Sahalia, Y. and Jacod, J. (2009b). Testing for jumps in a discretely observed process, *Ann. Statist.*, **37**, 184–222.

[1]We used $\beta^* = -\log(-\log(0.95))$ and $K = \lfloor n^{1/2} \rfloor$ in their notation.

	LM	S08	AJ+/AJ*+	AJ-/AJ*-	S10/S10*	true
r_n	NA	0.114	0.332	0.199	0.141	0.139
(s.d.)	NA	(0.005)	(0.023)	(0.014)	(0.003)	(r_{opt})
μ	0.209	**0.198**	0.215	0.214	0.210	0.2
(s.d.)	(0.093)	(0.082)	(0.095)	(0.087)	(0.085)	
σ	0.776	0.674	0.782	0.717	**0.699**	0.7
(s.d.)	(0.033)	(0.014)	(0.027)	(0.012)	(0.010)	
λ	6.782	11.07	6.284/9.861	7.648/9.837	8.603/**10.11**	10
(s.d.)	(0.667)	(1.690)	(0.624/1.229)	(0.724/1.023)	(0.897/1.077)	
α	0.002	0.0016	0.0017/0.0015	0.0011/0.0016	**0.0009**/0.0014	0.0
(s.d.)	(0.093)	(0.055)	(0.0997/0.0627)	(0.0826/0.0639)	(0.0731/0.0598)	
β	0.713	0.459	0.767/0.489	0.644/0.508	0.577/**0.493**	0.5
(s.d.)	(0.092)	(0.080)	(0.100/0.081)	(0.085/0.074)	(0.076/0.071)	

Table 1: Joint estimation using "Threshold Estimation". The bold type means the *best* estimate. The mean (s.d.) of $\hat{\sigma}_B$ through the experiments was 0.8891 (s.d. $= 0.0620$), which has a large bias from the truth 0.7.

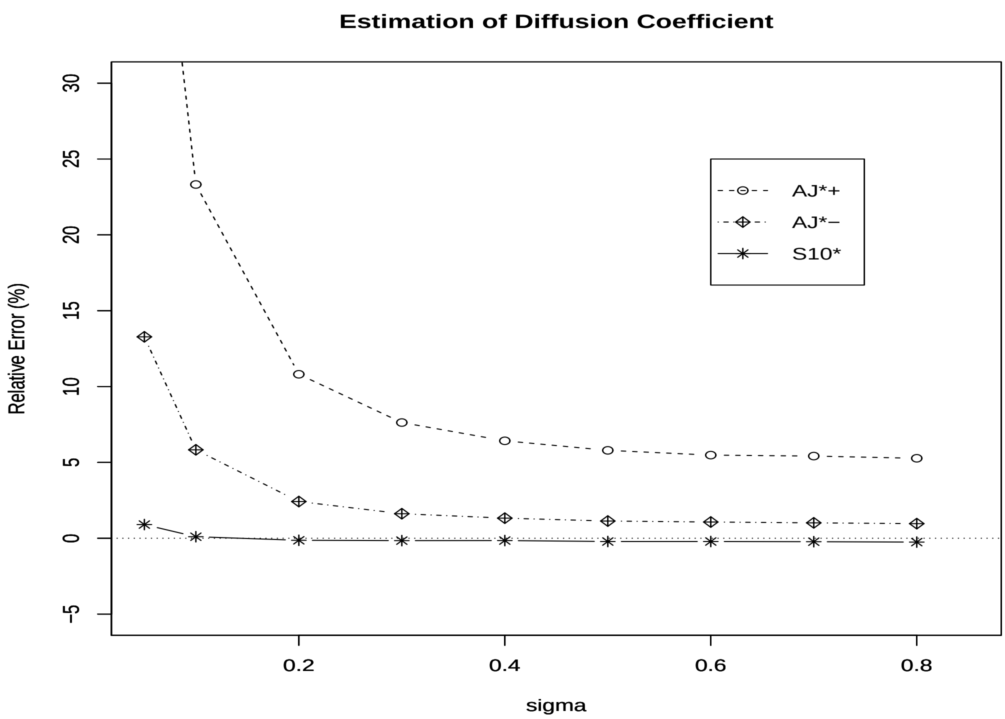

Figure 1: Relative errors of mean of $\widehat{\sigma}(r)$ through 5000 simulations: $\approx \frac{\mathbb{E}[\widehat{\sigma}(r)]-\sigma}{\sigma} \times 100$ (%), via `AJ*`$\pm$ and our proposal `S10*`. Although `AJ*` has a wide range of values, our proposal `S10*` is always accurate.

[3] Aït-Sahalia, Y. and Jacod, J. (2011). Testing whether jumps have finite and infinite activity, *Ann. Statist.*, **39**, no. 3, 1689–1719.

[4] Aït-Sahalia, Yacine; Jacod, Jean and Li, J. (2012). Testing for jumps in noisy high frequency data, *J. Econometrics* **168**, no. 2, 207–222.

[5] Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps (with discussion), *J. Financial Econ.*, **2**, 1–48.

[6] Cont, R. and Mancini, C. (2011). Nonparametric tests for pathwise properties of semimartingales, *Bernoulli*, **17**, no. 2, 781–813.

[7] Gobbi, F. and Mancini, C. (2008). Estimating the diffusion part of the covariation between two volatility models with jumps of Lévy type, *Applied and industrial mathematics in Italy II*, Ser. Adv. Math. Appl. Sci., **75**, 399–409.

[8] Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales, *Stochastic Process. Appl.*, **118**, no. 4, 517–559.

[9] Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Rev. Financial Studies*, **21**, 2535–2563.

[10] Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion, *Giornale dell'Istituto Italiano degli Attuari*, Volume LXIV, Roma, 19–47.

[11] Mancini, C. (2009). Nonparametric threshold estimation for models with stochastic diffusion coefficient and jumps, *Scand. J. Statist.*, 2009, **36**, 270–296.

[12] Ogihara, T. and Yoshida, N. (2011). Quasi-likelihood analysis for the stochastic differential equation with jumps. *Stat. Inference Stoch. Process.*, **14**, no. 3, 189–229.

[13] Shimizu, Y. (2002). Estimation of diffusion processes with jumps from discrete observations, Zenkin Tenkai, University of Tokyo, December 4, 2002; Master thesis 2003, Graduate School of Mathematical Sciences, The University of Tokyo.

[14] Shimizu, Y. (2008). A practical inference for discretely observed jump-diffusions from finite samples, *J. Japan Statist. Soc.*, **38**, no. 3, 391–413.

[15] Shimizu, Y. (2009). Functional estimation for Levy measures of semimartingales with Poissonian jumps, *J. Multivariate Anal.*, **100**, no. 6, 1073–1092

[16] Shimizu, Y. (2010). Threshold selection in jump-discriminant filter for discretely observed jump processes, *Statist. Methods and Appl.*, **19**, no. 3, 355–378.

[17] Shimizu, Y. and Yoshida, N. (2006). Estimation of parameters for diffusion processes with jumps from discrete observations, *Statist. Infer. Stochastic Proc.*, **9**, no. 3, 227–277.