Estimation of Integrated Volatility in the Presence of Noise and Jumps

Mathias Vetter (joint work with Mark Podolskij)
Ruhr-Universität Bochum

Amsterdam

June 8, 2007

Set up

- Classical framework: Observations $X_{\frac{i}{n}}$, $i = 0, \ldots, n$ of the process

$$(BSM) \qquad X_t = X_0 + \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}W_u \ , \qquad t \in [0,1] \tag{1}$$

  a is the (locally bounded) drift, σ is the càdlàg volatility and W is a standard Brownian motion.

- Generalisations:
  - Noisy diffusion model:

$$Y_{\frac{i}{n}} = X_{\frac{i}{n}} + U_i \qquad i = 0, \ldots, n \tag{2}$$

    with $(U_i)_{0 \leq i \leq n}$ i.i.d., independent of X; $EU = 0$ and $EU^2 = \omega^2$.
  - Noisy jump-diffusion model:

$$Z = Y + J \ , \tag{3}$$

    where J is a jump process.

Statement of the problem

- In this talk we propose a new methodology for the models (2) and (3) which solves the following estimation/test problems:
  - In model (2): Estimation of

    $$IV = \int_0^1 \sigma_u^2 du \ , \qquad IQ = \int_0^1 \sigma_u^4 du \ ,$$

    or even more generally $\int_0^1 |\sigma_u|^p du$, $p \geq 0$.
  - In model (3): Estimation of the joint quadratic variation

    $$\int_0^1 \sigma_u^2 du + \sum_{0 \leq u \leq 1} |\Delta J_u|^2$$

  - Test for the presence of jumps.
  - Estimation of σ_u^2 for $u \in [0,1]$.

Review: realised volatility and bipower variation

- Realised volatility and bipower variation at sampling frequency n are defined as:

$$RV_n = \sum_{i=1}^{n} |\Delta_i^n X|^2 \qquad \text{and} \qquad BV_n = \mu_1^{-2} \sum_{i=1}^{n-1} |\Delta_i^n X||\Delta_{i+1}^n X| \ , \tag{4}$$

  where $\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ and $\mu_r = \mathbb{E}\left(|\phi|^r\right)$, $\phi \sim N(0,1)$.

- RV_n, BV_n are both consistent for IV, i.e. as $n \to \infty$:

$$RV_n \xrightarrow{P} IV \ ,$$

$$BV_n \xrightarrow{P} IV. \tag{5}$$

Review: realised volatility and bipower variation

✘ Distribution theory for RV_n and BV_n:

$$n^{1/2}\left(RV_n - IV\right) \xrightarrow{d_{st}} MN\left(0, 2\, IQ\right)$$

$$n^{1/2}\left(BV_n - IV\right) \xrightarrow{d_{st}} MN\left(0, 2.6\, IQ\right) \tag{6}$$

where d_{st} denotes stable convergence. We refer to

- ❏ Jacod (1994)
- ❏ Barndorff-Nielsen & Shephard (2002)
- ❏ Barndorff-Nielsen, Graversen, Jacod, Podolskij & Shephard (2006)

for more details.

Review: realised volatility and bipower variation

- Consider a jump-diffusion model of the form:

$$\tilde{Z}_t = X_t + J_t \ , \qquad t \in [0,1] \tag{7}$$

- J is a jump process
- In the jump-diffusion framework we have

$$RV_n \xrightarrow{P} \int_0^1 \sigma_u^2 \mathrm{d}u + \sum_{u \leq 1} |\Delta J_u|^2 \ ,$$

while

$$BV_n \xrightarrow{P} \int_0^1 \sigma_u^2 \mathrm{d}u.$$

- Thus, $RV_n - BV_n$ estimates the jump component, if any, and may be transformed into a test of continuous sample paths (Barndorff-Nielsen & Shephard (2004, 2006)).

Review: multiscale estimator and realised kernel

- The practical estimation of IV suffers from the presence of market microstructure noise such as bid-ask bounds or price discreteness (e.g., Ait-Sahalia, Mykland & Zhang (2005), Bandi & Russell (2005), Barndorff-Nielsen, Hansen, Lunde & Shephard (2006), and related papers.)
- A more realistic framework is the noisy diffusion model (2)

$$Y = X + U \tag{8}$$

- In this setting usual statistics are inconsistent, e.g.

$$\hat{\omega}^2 := \frac{1}{2n} \sum_{i=1}^{n} |\Delta_i^n Y|^2 \xrightarrow{P} \omega^2.$$

Review: multiscale estimator and realised kernel

- There are basically two different methods of estimation of IV in the noisy diffusion framework.
  - Zhang (2006) and Ait-Sahalia, Mykland & Zhang (2005) via a *multiscale estimator*.
  - Barndorff-Nielsen, Hansen, Lunde & Shephard (2006) via a *realised kernel estimator*.
- Both proved a CLT for a standardized version of the estimators with an optimal convergence rate $n^{-1/4}$. IQ appears in each theorem.
- However, both methods
  - are not robust to jumps
  - do not provide estimates for other powers of σ

Main results: Modulated bipower variation

- We choose a number $c > 0$ such that $K = c\,n^{\frac{1}{2}}$ and $M = \frac{n}{2K}$ are integers and define the *modulated bipower variation* by

$$MBV(Y, r, l)_n = n^{\frac{(r+l)}{4} - \frac{1}{2}} \sum_{m=1}^{M} |\bar{Y}_m|^r |\bar{Y}_{m+1}|^l \qquad r, l \geq 0$$

$$\bar{Y}_m = \hat{X}_{\frac{2K(m-1)}{n}} - \hat{X}_{\frac{2Km-K}{n}}$$

$$\hat{X}_a = \frac{1}{K} \sum_{l=1}^{K} Y_{a+\frac{l}{n}}.$$

- This approach can be generalised. It is possible to choose other weights for the random variables $Y_{a+\frac{l}{n}}$ as long as certain regularity conditions are fulfilled.

Main results: Consistency

- ✘ The number $K = c\, n^{\frac{1}{2}}$ controls the stochastic order of the term $\bar{Y}_m$, i.e.

$$\bar{U}_m = O_p(n^{-\frac{1}{4}}) \ , \qquad \bar{X}_m = O_p(n^{-\frac{1}{4}}) \ ,$$

and so the stochastic orders of $\bar{U}_m$ and $\bar{X}_m$ are balanced!

Theorem 1: If $E|U|^{2(r+l)+\epsilon} < \infty$ for some $\epsilon > 0$, as $n \to \infty$

$$MBV(Y,r,l)_n \xrightarrow{P} MBV(Y,r,l) = \frac{\mu_r \mu_l}{2c} \int_0^1 (\nu_1 \sigma_u^2 + \nu_2 \omega^2)^{\frac{r+l}{2}} \ du \qquad (9)$$

for some known constants ν_1 and ν_2 (which depend on c).

Main results: Consistency

- Theorem 1 shows that $MBV(Y,r,l)_n$ is inconsistent when estimating arbitrary (integrated) powers of volatility. Though, when $r+l$ is an even number (this condition is satisfied for the most interesting cases) a slight modification of $MBV(Y,r,l)_n$ turns out to be consistent.

- Consistent estimates of IV are given by

$$MRV(Y)_n := \frac{2cMBV(Y,2,0)_n - \nu_2 \hat{\omega}^2}{\nu_1} \xrightarrow{P} \int_0^1 \sigma_u^2 \, du$$

  and

$$MBV(Y)_n := \frac{\frac{2c}{\mu_1^2} MBV(Y,1,1)_n - \nu_2 \hat{\omega}^2}{\nu_1} \xrightarrow{P} \int_0^1 \sigma_u^2 \, du.$$

  Notice that the estimator $MBV(Y)_n$ is robust to jumps.

- Consistent estimates of IQ can be obtained similarly.

Main results: Central limit theorem

- To prove a CLT, we need stronger conditions:
  - σ is an Ito diffusion itself.
  - The noise process U has the representation

$$U_i = \sqrt{n}\,\omega\left(B_{\frac{i}{n}} - B_{\frac{i-1}{n}}\right), \qquad (10)$$

    where B is another Brownian motion independent of W.

- The above condition on U ensures that X and U are measurable with respect to the same type of filtration. This condition is important for the proof!
- The normal distribution of the noise induced by (10) is not crucial for the CLT! Other functions of rescaled increments of B can be considered. However, this leads to a slight modification of the CLT.

Main results: Central limit theorem

Theorem 2: Assume that U is of the form (10), σ is an Ito-diffusion and $\sigma_s \neq 0$ for all s. Then we have

$$n^{\frac{1}{4}}\Big(MRV(Y)_n - \int_0^1 \sigma_u^2 \, du\Big) \xrightarrow{d_{st}} MN\Big(0, \frac{4c}{\nu_1^2}\widetilde{IQ}\Big) \,, \tag{11}$$

$$n^{\frac{1}{4}}\Big(MBV(Y)_n - \int_0^1 \sigma_u^2 \, du\Big) \xrightarrow{d_{st}} MN\Big(0, \frac{2c(\mu_2^2 + 2\mu_1^2\mu_2 - 3\mu_1^4)}{\mu_1^4\nu_1^2}\widetilde{IQ}\Big) \,, \tag{12}$$

where

$$\widetilde{IQ} = \int_0^1 (\nu_1\sigma_u^2 + \nu_2\omega^2)^2 \, du.$$

Main results: Central limit theorem

- ✘ The rate $n^{-1/4}$ is known to be optimal (see Gloter & Jacod (2001)).
- ✘ When σ is constant the conditional variance of $MRV(Y)_n$ is minimized at

  $$c = \frac{3\omega}{\sigma}$$

  and is approximately equal to

  $$21\sigma^3\omega.$$

  The corresponding expression for $MBV(Y)_n$ is given by $26\sigma^3\omega$. A natural lower bound is given by

  $$8\sigma^3\omega\ ,$$

  which is the variance of the maximum likelihood estimator (see Gloter & Jacod (2001)).

Modulated bipower variation: Modification

- For estimating the quadratic variation of $X + J$ we unfortunately have that

$$MRV(Z)_n \not\to \int_0^1 \sigma_u^2 du + \sum_{0 \le u \le 1} |\Delta J_u|^2$$

- However, a slight modification of the modulated bipower approach solves this problem. For $0 \le j \le 2K - 1$ we define the statistic

$$MRV(Y)_n^{(j)} := \frac{2cMBV(Y,2,0)_n^{(j)} - \nu_2 \hat{\omega}^2}{\nu_1} ,$$

where $MBV(Y,2,0)_n^{(j)}$ is the same quantity as $MBV(Y,2,0)_n$ but with starting point j/n. Finally, we consider the statistic

$$MRV(Y)_n^{ave} := \frac{1}{2K} \sum_{j=0}^{2K-1} MRV(Y)_n^{(j)} ,$$

and the quantity $MBV(Y)_n^{ave}$ is defined similarly.

Modulated bipower variation: Modification

- By construction we obtain the convergence in probability

$$MRV(Y)_n^{ave} \xrightarrow{P} \int_0^1 \sigma_u^2 du \,, \qquad MRV(Z)_n^{ave} \xrightarrow{P} \int_0^1 \sigma_u^2 du + \sum_{0\le u\le 1} |\Delta J_u|^2.$$

  Moreover, we can show a stable CLT for $MRV(Y)_n^{ave}$ (with convergence rate $n^{-1/4}$). When σ is constant the conditional variance of $MRV(Y)_n^{ave}$ (for the optimal choice of c) is approximately equal to

$$8.5\sigma^3\omega.$$

- For the bipower estimator of IV we have

$$MBV(Y)_n^{ave} \xrightarrow{P} \int_0^1 \sigma_u^2 du \,, \qquad MBV(Z)_n^{ave} \xrightarrow{P} \int_0^1 \sigma_u^2 du.$$

Applications: Confidence bands for quadratic variation

- For the noisy jump-diffusion model we can prove the stable convergence

$$n^{1/4}\Big(MRV(Z)_n^{ave} - \Big(\int_0^1 \sigma_u^2 du + \sum_{0\le u\le 1} |\Delta J_u|^2\Big)\Big) \stackrel{d_{st}}{\to} MN(0, V^2).$$

- By an estimation of the conditional (asymptotic) variance V^2 we can obtain a standard CLT, and so the confidence bands for the quadratic variation of $Y + J$.

Applications: Tests for jumps

- ✘ Since we have derived stable CLT's for $MRV(Y)_n^{ave}$ and $MBV(Y)_n^{ave}$ (i.e. when there are no jumps), we are able to test for jumps. More precisely, we reject the null hypothesis of no jumps for large values of

$$n^{1/4}\Big(MRV(Z)_n^{ave} - MBV(Z)_n^{ave}\Big)$$

or

$$n^{1/4}\Big(\frac{MRV(Z)_n^{ave}}{MBV(Z)_n^{ave}} - 1\Big).$$

- ✘ Another possibility is to apply the idea of Ait-Sahalia & Jacod (2006), which, however, envolves hard calculations in our case.