A Rate-Optimal Estimator for the Integrated Covariance of Nonsynchronously Observed Diffusion Processes with Endogenous Noise

Yuta Koike

Graduate School of Mathematical Sciences

University of Tokyo

JAPAN

November 17, 2012

Outline

- Introduction
- Model
- Construction of the proposed estimator
- Main results
- Simulation study
- Conclusions

Introduction

- X^1, X^2: Latent log-price processes of two assets
- Assume that X^1 and X^2 are diffusion processes.
- Then we can define the integrated covariance process $[X^1, X^2]$ of two assets.
- As is well known, integrated covariance plays a central role for many issues in finance.

  $\Longrightarrow$ It is important to estimate this quantity from observation data of X^1 and X^2.

Introduction

- The aim of this talk Estimating the integrated covariance using high-frequency observation data

- I try to deal with the following problems:
  - Observation noise correlated with the latent log-returns (motivated by an empirical study of Hansen and Lunde (2006))
  - Nonsynchronicity of observation times
  - Dependence of observation times on observed processes (motivated by recent studies in the absence of noise; see e.g. Hayashi and Yoshida (2011))

Introduction

- The *Realized kernel* proposed in Barndorff-Nielsen et al. (2011) is known as a consistent estimator in such a situation.
- However, it does not attain the optimal convergence rate given in Gloter and Jacod (2001).
  $\Longrightarrow$ I propose an estimator which overcomes this issue.
- This is the first rate-optimal estimator in the presence of the (asymptotically non-degenerate) correlation between the latent log-returns and the observation noise (as far as I know).

Model: Latent log-prices

- X^1, X^2: Latent log-price processes of two assets

$$dX_t^k = a_t^k dt + \sigma_t^k dw_t^k, \qquad d[w^1, w^2]_t = \vartheta_t dt.$$

- $\mathcal{B}^{(0)} = (\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)}), P^{(0)})$: A stochastic basis
- w^k is a standard Wiener process on $\mathcal{B}^{(0)}$,
- a^k, σ^k and ϑ are càdlàg adapted processes.

Model: Observation times

- $\mathcal{I} = (S^i)_{i=0}^{\infty}$, $\mathcal{J} = (T^j)_{j=0}^{\infty}$: Sequences of $(\mathcal{F}_t^{(0)})$-stopping times satisfying $S^i \uparrow \infty$, $T^i \uparrow \infty$ as $i \to \infty$.

- $\mathcal{I}$ and $\mathcal{J}$ depend on a parameter $n \in \mathbb{N}$ which represents the frequency of the observation, and I assume that

$$n^{1-\varepsilon} \left[\sup_{i:S^i \leq t} (S^i - S^{i-1}) \vee \sup_{j:T^j \leq t} (T^j - T^{j-1}) \right] \to^p 0$$

as $n \to \infty$ for any $\varepsilon, t > 0$ ($S^{-1} = T^{-1} := 0$).

Model: Observation noise

- $\mathsf{X}^1_{S^i}, \mathsf{X}^2_{T^j}$: Noisy observation data of X^1 and X^2 observed at each times in $\mathcal{I}$ and $\mathcal{J}$ respectively:

$$\mathsf{X}^1_{S^i} = X^1_{S^i} + U^1_{S^i}, \qquad \mathsf{X}^2_{T^j} = X^2_{T^j} + U^2_{T^j}.$$

- The observation noise is of the form

$$U^1_{S^i} = \sqrt{n}(Z^1_{S^i} - Z^1_{S^{i-1}}) + \epsilon^1_{S^i},$$

$$U^2_{T^j} = \sqrt{n}(Z^2_{T^j} - Z^2_{T^{j-1}}) + \epsilon^2_{T^j}.$$

Model: Endogenous noise

- Z_t^1, Z_t^2: Diffusion processes govern by the following SDE:

$$\mathrm{d}Z_t^k = \underline{a}_t^k \mathrm{d}t + \underline{\sigma}_t^k \mathrm{d}\underline{w}_t^k, \qquad \mathrm{d}[\underline{w}^1, \underline{w}^2]_t = \underline{\vartheta}_t \mathrm{d}t,$$

$$\mathrm{d}[\underline{w}^k, w^l]_t = \underline{\vartheta}_t^{kl} \mathrm{d}t \ (l = 1, 2).$$

  - $\underline{w}^k$ is a standard Wiener process on $\mathcal{B}^{(0)}$,
  - $\underline{a}^k$, $\underline{\sigma}^k$, $\underline{\vartheta}$ and $\underline{\vartheta}^{kl}$ are càdlàg adapted processes.

Model: Exogenous noise

- $(\epsilon^1_{S^i})$ and $(\epsilon^2_{T^j})$ are two i.i.d. random variables independent of $\mathcal{F}^{(0)}$. $\epsilon^1_{S^i}$ and $\epsilon^2_{T^j}$ are centered and mutually independent if $S^i \neq T^j$.
- Mathematical construction: Let Q be a probability measure on $\mathbb{R}^2$ with $\int zQ(\mathrm{d}z) = 0$. We endow the space $\Omega^{(1)} = (\mathbb{R}^2)^{[0,\infty)}$ with the product Borel σ-field $\mathcal{F}^{(1)}$. Set $P^{(1)}$ as $P^{(1)} = \otimes_{t\in[0,1]} P_t$ with $P_t = Q$ for all t. Then, $(\epsilon^1_t, \epsilon^2_t)$ is defined as the canonical process on $(\Omega^{(1)}, \mathcal{F}^{(1)}, (\mathcal{F}^{(1)}_t), P^{(1)})$ with $(\mathcal{F}^{(1)}_t)$ being the canonical filtration.

Model

- I work with the product stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ given by

$$\Omega = \Omega^{(0)} \times \Omega^{(1)}, \qquad \mathcal{F} = \mathcal{F}^{(0)} \otimes \mathcal{F}^{(1)},$$

$$\mathcal{F}_t = \cap_{s>t} \mathcal{F}_s^{(0)} \otimes \mathcal{F}_s^{(1)}, \qquad P = P^{(0)} \otimes P^{(1)}.$$

- The aim of this talk: Estimating

$$IC_T = \int_0^T \sigma_s^1 \sigma_s^2 \vartheta_s \mathrm{d}s$$

from the observation data $(\mathsf{X}_{S^i}, \mathsf{Y}_{T^j})_{i,j:S^i,T^j \leq T}$ in the time interval $[0, T]$ as $n \to \infty$.

Construction of the proposed estimator: Pre-averaging

- Choose a positive number θ.
- Choose a positive integer k_n satisfying

$$k_n = \theta\sqrt{n} + o(n^{1/4}).$$

  ($k_n = \lceil \theta\sqrt{n} \rceil$ for example)

- Choose a continuous real-valued function g on $[0, 1]$ which is piecewise C^1 with piecewise Lipschitz derivative g' (e.g. $g(x) = x \wedge (1 - x)$).

Construction of the proposed estimator: Pre-averaging

- The *pre-averaging in tick time*, which was first introduced in Christensen et al. (2010), is defined by

$$\overline{\mathsf{X}}^1(\mathcal{I})^i = \sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right)(\mathsf{X}^1_{S^{i+p}} - \mathsf{X}^1_{S^{i+p-1}}).$$

- For a technical reason (computing the asymptotic variance explicitly), I use the *pre-averaging in refresh time* instead of the pre-averaging in tick time.

- The refresh time was first introduced to this area in Barndorff-Nielsen et al. (2011).

Construction of the proposed estimator: Refresh time

- The *refresh times* are stopping times $R^0, R^1, \dots$ defined sequentially by $R^0 := S^0 \vee T^0$ and

$$R^k := \min\{S^i | S^i > R^{k-1}\} \vee \min\{T^j | T^j > R^{k-1}\}, \qquad (k = 1, 2, \dots)$$

- New sampling times $\widehat{S}^k$ and $\widehat{T}^k$ are defined by $\widehat{S}^0 := S^0$, $\widehat{T}^0 := T^0$ and

$$\widehat{S}^k := \min\{S^i | S^i > R^{k-1}\}, \qquad \widehat{T}^k := \min\{T^j | T^j > R^{k-1}\} \qquad (k = 1, 2, \dots)$$

Construction of the proposed estimator: Pre-averaging

- The pre-averaging in refresh time is defined by

$$\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i = \sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right)(\mathsf{X}^1_{\widehat{S}^{i+p}} - \mathsf{X}^1_{\widehat{S}^{i+p-1}}).$$

$\overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j$ is also defined in a similar manner.

Construction of the proposed estimator

- Our estimator $\widehat{PHY}^n_T$ is defined by

$$\widehat{PHY}^n_T = \frac{1}{(\psi_{HY} k_n)^2} \sum_{i,j:\widehat{S}^{i+k_n} \vee \widehat{T}^{j+k_n} \leq T} \overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i \overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j \bar{K}^{ij},$$

where

- $\psi_{HY} = \int_0^1 g(x)\mathrm{d}x$ (Normalizing factor),
- $\bar{K}^{ij} = 1_{\{[\widehat{S}^i,\widehat{S}^{i+k_n})\cap[\widehat{T}^j,\widehat{T}^{j+k_n})\neq\emptyset\}}$ (Hayashi-Yoshida type factor).

Main results: Consistency

- Let $N_t^n = \sum_{k=1}^{\infty} 1_{\{R^k \leq t\}}$ for each t.

Theorem 1

Suppose that $\int |z|^2 Q(\mathrm{d}z) < \infty$ and $N_T^n = O_p(n)$ as $n \to \infty$. Then

$$\widehat{PHY}_T^n \to^p IC_T$$

as $n \to \infty$.

- We need no restriction on the observation times except for the standard one (cf. Hayashi and Kusuoka (2008)).

Main results: Asymptotic mixed normality

- For each $k \geq 1$, set

$$\Gamma^k = [R^{k-1}, R^k), \qquad \check{I}^k = [\check{S}^k, \widehat{S}^k), \qquad \check{J}^k = [\check{T}^k, \widehat{T}^k),$$

where $\check{S}^k = \sup_{S^i < \widehat{S}^k} S^i$ and $\check{T}^k = \sup_{T^j < \widehat{T}^k} T^j$.

- Assume that for each n we have a filtration $(\mathcal{H}_t^n)$ of $\mathcal{F}$ to which $N^{n,1}$, $N^{n,2}$, σ^k, ϑ, $\underline{\sigma}^k$, $\underline{\vartheta}$ and $\underline{\vartheta}^{kl}$ $(k, l = 1, 2)$ are adapted, where

$$N_t^{n,1} = \sum_{i=1}^{\infty} 1_{\{S^i \leq t\}}, \qquad N_t^{n,2} = \sum_{j=1}^{\infty} 1_{\{T^j \leq t\}}.$$

Main results: Asymptotic mixed normality

- For each $\rho > 0$, define the processes $G(\rho)^n$, $F(\rho)^{n,1}$, $F(\rho)^{n,2}$, $F(1)^{n,1*2}$ and χ^n by

$$G(\rho)^n_s = E\left[\left(n|\Gamma^k|\right)^\rho \middle| \mathcal{H}^n_{R^{k-1}}\right], \quad F(\rho)^{n,1}_s = E\left[\left(n|\check{I}^{k+1}|\right)^\rho \middle| \mathcal{H}^n_{R^{k-1}}\right],$$

$$F(\rho)^{n,2}_s = E\left[\left(n|\check{J}^{k+1}|\right)^\rho \middle| \mathcal{H}^n_{R^{k-1}}\right], \quad \chi^n_s = P(\widehat{S}^k = \widehat{T}^k \middle| \mathcal{H}^n_{R^{k-1}})$$

and

$$F(1)^{n,1*2}_s = nE\left[|\check{I}^k \cap \check{J}^k| + |\check{I}^{k+1} \cap \check{J}^k| + |\check{I}^k \cap \check{J}^{k+1}| \middle| \mathcal{H}^n_{R^{k-1}}\right]$$

when $s \in \Gamma^k$ ($|\cdot|$ denotes the Lebesgue measure).

Main results: Asymptotic mixed normality

[A1] (i) There exists a càdlàg $\mathbf{F}^{(0)}$-adapted process G such that G and G_- do not vanish and $\sup_{0\le s\le T}|G(1)^n_s - G_s| \to^p 0$ as $n\to\infty$.

(ii) There exists a càdlàg $\mathbf{F}^{(0)}$-adapted process χ such that $\sup_{0\le s\le T}|\chi^n_s - \chi_s| \to^p 0$ as $n\to\infty$.

(iii) There exist càdlàg $\mathbf{F}^{(0)}$-adapted processes F^1, F^2 and F^{1*2} such that $\sup_{0\le s\le T}|F(1)^{n,l}_s - F^l_s| \to^p 0$ as $n\to\infty$ for every $l = 1, 2, 1*2$.

(iv) There exists a constant $\rho > 1$ such that the sequences

$$\left(\sup_{0\le s\le T} G(\rho)^n_s\right)_{n\in\mathbb{N}}, \quad \left(\sup_{0\le s\le T} F(\rho)^{n,1}_s\right)_{n\in\mathbb{N}}, \quad \left(\sup_{0\le s\le T} F(\rho)^{n,2}_s\right)_{n\in\mathbb{N}}$$

are tight.

Main results: Asymptotic mixed normality

- Example (S^i) (resp. (T^j)) are Poisson arrival times with intensity np^1, $p^1 > 0$ (resp. np^2, $p^2 > 0$).
- (S^i) and (T^j) are mutually independent as well as independent of X^k, Z^k and ϵ^k $(k = 1, 2)$.
- Then, [A1] holds with $(\mathcal{H}_t^n)$ being the filtration generated by N^n, σ^k, ϑ, $\underline{\sigma}^k$, $\underline{\vartheta}$ and $\underline{\vartheta}^{kl}$ $(k, l = 1, 2)$ and

$$G_s = \frac{1}{p^1} + \frac{1}{p^2} - \frac{1}{p^1 + p^2}, \qquad F_s^1 = \frac{1}{p^1}, \qquad F_s^2 = \frac{1}{p^2},$$

$$F_s^{1*2} = \frac{2}{p^1 + p^2}, \qquad \chi_s = 0.$$

Main results: Asymptotic mixed normality

- The following condition is an analogue of the strong predictability condition introduced in Hayashi and Yoshida (2011):

[A2] There exists a constant $\eta \in (0, \frac{1}{2})$ such that S^i and T^i are $(\mathcal{G}_t^{(n)})$-stopping times for every $n, i \in \mathbb{N}$, where

$$\mathcal{G}_t^{(n)} = \mathcal{F}^{(0)}_{(t-n^{-\eta})_+}$$

for each $t \in \mathbb{R}_+$.

Main results: Asymptotic mixed normality

- For a real-valued function x on $[0, T]$, put

$$w(x; \delta, T) = \sup\left\{|x(t) - x(s)|\Big| s, t \in [0, T], |s - t| \leq \delta\right\}$$

  for each $\delta > 0$.

- The following condition is satisfied when f is a Brownian semimartingale, for example.

[A3] For every $f = a^k$, σ^k, ϑ^k, $\underline{a}^k$, $\underline{\sigma}^k$, $\underline{\vartheta}^k$ and $\underline{\vartheta}^{kl}$ $(k, l = 1, 2)$, $w(f; h, T) = O_p(h^{\frac{1}{2} - \lambda})$ as $h \to 0$ for any $\lambda > 0$.

Main results: Asymptotic mixed normality

- The following condition is an analogue of the condition [A6] in Hayashi and Yoshida (2011):

[A4] $n\sum_{k:R^k\leq T}|\Gamma^k|^2 = O_p(1)$ as $n\to\infty$.

- The following condition is a regularity condition for the exogenous noise:

[N] $\int |z|^8 Q(\mathrm{d}z) < \infty$.

- Set $\Psi = \int zz^* Q(\mathrm{d}z)$.

Main results: Asymptotic mixed normality

- For two real-valued continuous functions α, β on $\mathbb{R}$, set

$$\psi_{\alpha,\beta}(x) = \int_0^1 \int_{x+u-1}^{x+u+1} \alpha(u)\beta(v)\mathrm{d}v\mathrm{d}u.$$

- We extend the functions g and g' to the whole real line by setting $g(x) = g'(x) = 0$ for $x \notin [0,1]$. Then we put

$$\kappa := \int_{-2}^{2} \psi_{g,g}(x)^2 \mathrm{d}x, \qquad \widetilde{\kappa} := \int_{-2}^{2} \psi_{g',g'}(x)^2 \mathrm{d}x,$$

$$\overline{\kappa} := \int_{-2}^{2} \psi_{g,g'}(x)^2 \mathrm{d}x.$$

Theorem 2

Suppose [A1]-[A4] and [N] hold. Then we have

$$n^{1/4}(\widehat{PHY}_T^n - IC_T) \to^{d_s} \sqrt{V_T}\zeta$$

as $n \to \infty$, where ζ is a standard normal random variable independent of $\mathcal{F}$ and $V_T = \int_0^T w_s^2 \mathrm{d}s$ with

$$\begin{aligned} w_s^2 =& \psi_{HY}^{-4}\Bigg[\theta\kappa(\sigma_s^1\sigma_s^2)^2(1+\vartheta_s^2)G_s + \theta^{-3}\widetilde{\kappa}\left\{\overline{\Psi}_s^{11}\overline{\Psi}_s^{22} + \left(\overline{\Psi}_s^{12}\right)^2\right\}\frac{1}{G_s} \\ &+ \theta^{-1}\overline{\kappa}\Big\{(\sigma_s^1)^2\overline{\Psi}_s^{22} + (\sigma_s^2)^2\overline{\Psi}_s^{11} + 2\sigma_s^1\sigma_s^2\vartheta_s\overline{\Psi}_s^{12} \\ &- \left(\underline{\sigma}_s^1\sigma_s^2\underline{\vartheta}_s^{12}F_s^1 - \underline{\sigma}_s^2\sigma_s^1\underline{\vartheta}_s^{21}F_s^2\right)^2\Big\}\Bigg], \\ \overline{\Psi}_s^{ll} =& \Psi^{ll} + \left(\underline{\sigma}_s^l\right)^2 F_s^l \ (l=1,2), \quad \overline{\Psi}_s^{12} = \Psi^{12}\chi_s + \underline{\sigma}_s^1\underline{\sigma}_s^2\underline{\vartheta}_s F_s^{1*2}. \end{aligned}$$

Simulation study

- X^1, X^2: SV1F model (e.g. Barndorff-Nielsen et al. (2011))

$$\mathrm{d}X_t^k = \mu^k \mathrm{d}t + \rho^k \sigma_t^k B_t^k + \sqrt{1-(\rho^k)^2}\sigma_t^k W_t,$$

$$\sigma_t^k = \exp(\beta_0^k + \beta_1^k \varrho_t^k), \qquad \mathrm{d}\varrho_t^k = \alpha^k \varrho_t^k \mathrm{d}t + \mathrm{d}B_t^k,$$

where

- $(\mu^k, \beta_0^k, \beta_1^k, \alpha^k, \rho^k) = (0.03, -5/16, 1/8, -1/40, -0.3)$,
- (B^1, B^2, W) is a 3-dimensional standard Wiener processes,
- $\varrho_0^k \sim N(0, (-2\alpha^k)^{-1})$.

Simulation study

- (S^i) (resp. (T^j)): Poisson arrival times with intensity np^1 (resp. np^2)
- $n = 23400$ and $p^1 = 1/6$, $p^2 = 1/60$, $1/6$
- $Z^k = -\sqrt{\psi p^k} X^k$ and $\epsilon^k = 0$ $(k = 1, 2)$
- $\psi = 0, 0.001, 0.01$
- Number of repetition: 1000

Simulation study

- Benchmark estimators:
  - Pre-averaged HY estimator in tick time (PHY; proposed in Christensen et al. (2010))
  - Realized kernel (RK)
  - Modulated realized covariance based on refresh sampling (MRC; proposed in Christensen et al. (2010). This estimator is rate-optimal but inconsistent when $Z^k \neq 0$.)
- Tuning parameters: $\theta = 0.15$ and $g(x) = x \wedge (1-x)$ for $\widehat{PHY}$, PHY and MRC (following Christensen et al. (2011)). For RK, I followed Barndorff-Nielsen et al. (2009, 2011).

	$\psi = 0$		$\psi = 0.001$		$\psi = 0.01$	
$p^2 = \frac{1}{60}$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
$\widehat{PHY}$	-0.013	0.181	-0.013	0.184	0.007	0.249
PHY	-0.026	0.244	-0.034	0.240	-0.037	0.241
RK	-0.022	0.211	-0.035	0.210	-0.018	0.243
MRC	-0.044	0.112	-0.312	0.504	-0.064	0.445
$p^2 = \frac{1}{6}$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
$\widehat{PHY}$	-0.003	0.107	-0.002	0.110	0.013	0.197
PHY	-0.004	0.115	-0.004	0.116	0.005	0.148
RK	-0.006	0.127	-0.002	0.135	0.043	0.217
MRC	-0.002	0.052	0.080	0.129	1.469	2.084

Conclusions

- Combining the pre-averaged HY estimator with the refresh time, I proposed a new estimator for the integrated covariance.

  $\Longrightarrow$ It can deal with both of the nonsynchronicity of sampling schemes and endogenous observation noise.

- I have shown the consistency and the asymptotic mixed normality of the proposed estimator.
  - It attains the optimal convergence rate.

- The simulation study shows that the proposed estimator performs well in finite samples.

References

O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, and N. Shephard. Realized kernels in practice: trades and quotes. *Econom. J.*, 12:C1–C32, 2009.

O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, and N. Shephard. Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *J. Econometrics*, 162:149–169, 2011.

K. Christensen, S. Kinnebrock, and M. Podolskij. Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *J. Econometrics*, 159:116–133, 2010.

K. Christensen, M. Podolskij, and M. Vetter. On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. CREATES Research Paper 2011-53, Aarhus University, 2011.

A. Gloter and J. Jacod. Diffusions with measurement errors. II. Optimal estimators. *ESAIM Probab. Stat.*, 5:243–260, 2001.

P. R. Hansen and A. Lunde. Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, 24(2):127–161, 2006.

T. Hayashi and S. Kusuoka. Consistent estimation of covariation under nonsynchronicity. *Stat. Inference Stoch. Process.*, 11:93–106, 2008.

T. Hayashi and N. Yoshida. Nonsynchronous covariation process and limit theorems. *Stochastic Process. Appl.*, 121:2416–2454, 2011.