Robustness of the Separating Information Maximum Likelihood Estimation of Realized Volatility with Micro-Market Noise [a]

Naoto Kunitomo

University of Tokyo

August 2010
at Paris

---

[a] This talk is based on several unpublished papers which have been available at http://www.e.u-tokyo.ac.jp/cirje/research/dp : Kunitomo and Sato (2008a,b), (2010a) and a forthcoming paper : Kunitomo and Sato (2010b).

Outline of Presentation

1. Introduction
2. SIML (Separating Information Maximum Likelihood) estimation
3. Asymptotic Properties and Robustness
4. Simulations
5. Concluding remarks

1 Motivations of Study

1. Recently a considerable interest has been paid on the estimation problem of the realized volatility by using (ultra)high-frequency data. Several statistical methods have been developed by Anderson, T.G., Bollerslev, T. Diebold, F.K. and Labys, P. (2000 JASA), Gloter and Jacod (2001), Ait-Sahalia, Y., P. Mykland and L. Zhang (2005), Zhang, L., P. Mykland and Ait-Sahalia (2005), and Barndorff-Nielsen, O., P. Hansen, A. Lunde and N. Shepard (2006). Many of existing methods are rather complicated with micro-market noise.
2. Our aim is to develop a simple (non-parametric) estimation method for practical applications with micro-market noise. We have proposed a new estimation method called SIML : Separating Information Maximum Likelihood method in Kunitomo and Sato (2008a,b, 2010a,b)* ; Unpublished Discussion Papers : CIRJE-F-581,601,733 (http://www.e.u-tokyo.ac.jp/cirje/research/), and a forthcoming Paper : Kunitomo and Sato (2010b), Mathematics and Computers in Simulations, Elsevier.
3. We shall show that the SIML estimator has the consistency and the

asymptotic normality even when the noises are autocorrelated, and the efficient price process is endogenous, i.e. it can be correlated with the noise term. Besides, it has reasonable small sample properties.

2 SIML: Separating Information Maximum Likelihood) estimation

Let y_{ij} be the $i-$th observation of the $j-$th (log-) price at t_i^n for $i = 1, \cdots, n; j = 1, \cdots, p; 0 = t_0^n \leq t_1^n \leq \cdots \leq t_n^n = 1$. We set $\mathbf{y}_i = (y_{i1}, \cdots, y_{ip})^{'}$ be a $p \times 1$ vector and $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ be an $n \times p$ matrix of observations. The underlying continuous process $\mathbf{x}_i$ at t_i^n $(i = 1, \cdots, n)$ is not necessarily the same as the observed prices and let $\mathbf{v}_i^{'} = (v_{i1}, \cdots, v_{ip})$ be the vector of the additive micro-market noise at t_i^n, which is independent of $\mathbf{x}_i$. Then we have

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i$$

where $\mathbf{v}_i$ are a sequence of independent random variables with $\mathcal{E}(\mathbf{v}_i) = \mathbf{0}$ and $\mathcal{E}(\mathbf{v}_i \mathbf{v}_i^{'}) = \mathbf{\Sigma}_v$. In this paper we focus on the equi-distance case with $h_n = t_i^n - t_{i-1}^n = 1/n$ $(i = 1, \cdots, n)$.

We assume that

$$\mathbf{x}_t = \mathbf{x}_0 + \int_0^t \mathbf{C}_s^{(x)} d\mathbf{B}_s \quad (0 \le t \le 1),$$

where $\mathbf{B}_s$ is a $q \times 1$ $(q \ge 1)$ vector of the standard Brownian motions, and $\mathbf{C}_s^{(x)} = (c_{gh}^{(x)}(s))$ is a $p \times q$ matrix which is progressively measurable in $[0, s] \times \mathcal{F}_s$ and predictable. We write the instantaneous diffusion functions $\mathbf{\Sigma}_s^{(x)}$ $(= (\sigma_{gh}^{(x)(s)})) = \mathbf{C}_s^{(x)} \mathbf{C}_s^{(x)'}$ ($\mathcal{F}_s$ is the $\sigma-$field generated by $\{\mathbf{B}_r, r \le s\}$).

The problem is to estimate the quadratic variations and co-variations

$$\mathbf{\Sigma}_x = (\sigma_{gh}^{(x)}) = \int_0^1 \mathbf{\Sigma}_s^{(x)} ds$$

of the underlying continuous process $\{\mathbf{x}_t\}$ and also the variance-covariance $\mathbf{\Sigma}_v = (\sigma_{gh}^{(v)})$ of the noise process from the observed discrete time process $\mathbf{y}_i$ $(i = 1, \cdots, n)$. We use the notation that $\sigma_{gh}^{(x)}(s)$ and $\sigma_{gh}^{(v)}$ are the (g, h)-th element of $\mathbf{\Sigma}_s^{(x)}$ and $\mathbf{\Sigma}_v$, respectively.

In this paper three different situations on the instantaneous covariance function shall be considered. (i) When the coefficient matrix is constant, (i.e.

$\mathbf{C}_s^{(x)} = \mathbf{C}^{(x)}$), we call the standard case or the simple case. (ii) When the coefficient matrix is time-varying, but it is a deterministic function of time ($\mathbf{C}_s^{(x)}$), we call the deterministic time-varying case. (iii) When the coefficient matrix is time-varying and it is a stochastic function of time ($\mathbf{C}_s^{(x)}$), we call the stochastic case. We write the conditional covariance function of the (underlying) price returns without micro-market noise as

$$\mathcal{E}\left[(\mathbf{x}_i - \mathbf{x}_{i-1})(\mathbf{x}_i - \mathbf{x}_{i-1})^{'}|\mathcal{F}_{n,i-1}\right] = \int_{t_{i-1}}^{t_i} \mathbf{\Sigma}_s^{(x)} ds \ ,$$

which corresponds to $\frac{1}{n}\mathbf{\Sigma}_{t_{i-1}^n}^{(x)} = \frac{1}{n}\mathbf{C}_{t_{i-1}^n}^{(x)}\mathbf{C}_{t_{i-1}^n}^{(x)'}$, where $\mathbf{x}_i - \mathbf{x}_{i-1}$ is a sequence of martingale differences, $\mathbf{\Sigma}_s^{(x)}$ are the time-dependent (instantaneous) conditional variance and $\mathcal{F}_{n,i-1}$ is the $\sigma-$field generated by $\mathbf{x}_j \ (j \leq i-1)$ with (2.2) and $\mathbf{v}_j \ (j \leq i-1)$. More generally, as $n \to \infty$ we can consider the situation that the (true) realized covariance of the returns

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{\Sigma}_{t_{i-1}^n}^{(x)} \longrightarrow \mathbf{\Sigma}_x = \int_0^1 \mathbf{\Sigma}_s^{(x)} ds \ ,$$

which is a deterministic and constant matrix, $\mathbf{\Sigma}_0^{(x)}$ is the (fixed) initial condition and we assume $\sup_{0\le s\le 1}\|\mathbf{\Sigma}_s^{(x)}\| < \infty \ (a.s.)$ for the instantaneous covariance function.

The Standard Case

We first consider the situation when $\mathbf{x}_i$ and $\mathbf{v}_i$ $(i=1,\cdots,n)$ are independent with $\mathbf{\Sigma}_s^{(x)} = \mathbf{\Sigma}_x$ $(0 \le s \le 1)$, and $\mathbf{v}_i$ are independently, identically and normally distributed as $N_p(\mathbf{0},\mathbf{\Sigma}_v)$. Given the initial condition $\mathbf{y}_0$,

$$\mathbf{Y}_n \sim N_{n\times p}\left(\mathbf{1}_n \cdot \mathbf{y}_0^{'}, \mathbf{I}_n \otimes \mathbf{\Sigma}_v + \mathbf{C}_n\mathbf{C}_n^{'} \otimes h_n\mathbf{\Sigma}_x\right) ,$$

where $\mathbf{1}_n^{'} = (1,\cdots,1)$, $h_n = 1/n$ $(= t_i^n - t_{i-1}^n)$ and

$$\mathbf{C}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & 0 \\ 1 & \cdots & 1 & 1 & 1 \end{pmatrix} .$$

$$\mathbf{C}_n^{-1}\mathbf{C}_n^{'-1} = \mathbf{P}_n\mathbf{D}_n\mathbf{P}_n^{'} = 2\mathbf{I}_n - 2\mathbf{A}_n \ ,$$

where $\mathbf{D}_n$ is a diagonal matrix with $d_k = 2\left[1 - \cos(\pi(\frac{2k-1}{2n+1}))\right]$, and $\mathbf{P}_n = (p_{jk})$, $p_{jk} = \sqrt{\frac{2}{n+\frac{1}{2}}}\cos\left[\pi(\frac{2k-1}{2n+1})(j-\frac{1}{2})\right]$. We transform $\mathbf{Y}_n$ to $\mathbf{Z}_n$ $(= (\mathbf{z}_k^{'}))$ by

$$\mathbf{Z}_n = h_n^{-1/2}\mathbf{P}_n^{'}\mathbf{C}_n^{-1}\left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right)$$

where $\bar{\mathbf{Y}}_0 = \mathbf{1}_n \cdot \mathbf{y}_0^{'}$. The likelihood function under the Gaussian noise is given by

$$L_n^*(\boldsymbol{\theta}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{np}\prod_{k=1}^{n}|a_{kn}\boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_x|^{-1/2}e^{\left\{-\frac{1}{2}\mathbf{z}_k^{'}\left(a_{kn}\boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_x\right)^{-1}\mathbf{z}_k\right\}}$$

where $a_{kn} = 4n\sin^2\left[\frac{\pi}{2}\left(\frac{2k-1}{2n+1}\right)\right]$. Hence the maximum likelihood (ML) estimator can be defined as the solution of maximizing

$$L_n(\boldsymbol{\theta}) = \sum_{k=1}^{n}\log|a_{kn}\boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_x|^{-1/2} - \frac{1}{2}\sum_{k=1}^{n}\mathbf{z}_k^{'}[a_{kn}\boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_x]^{-1}\mathbf{z}_k \ .$$

From this representation we find that the ML estimator of unknown parameters is a rather complicated function of all observations in general because each a_{kn} terms depend on k as well as n. Let denote $a_{k_n,n}$ and then we can evaluate that $a_{k_n,n} \to 0$ as $n \to \infty$ when $k_n = O(n^{\alpha})$ $(0 < \alpha < \frac{1}{2})$ since $\sin x \sim x$ as $x \to 0$. Also $a_{n+1-l_n,n} = O(n)$ when $l_n = O(n^{\beta})$ $(0 < \beta < 1)$. When k_n is small, we expect that $a_{k_n,n}$ is small and we approximate $2 \times L_n(\boldsymbol{\theta})$ by

$$L_n^{(1)}(\boldsymbol{\theta}) = -m \log |\boldsymbol{\Sigma}_x| - \sum_{k=1}^{m} \mathbf{z}_k^{'} \boldsymbol{\Sigma}_x^{-1} \mathbf{z}_k \ .$$

Then the SIML estimator of $\hat{\boldsymbol{\Sigma}}_x$ is defined by

$$\hat{\boldsymbol{\Sigma}}_x = \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k^{'} \ .$$

On the other hand, when l_n is small and $k_n = n + 1 - l_n$, we expect that

$a_{n+1-l_n,n}$ is large and then we approximate $2 \times L_n(\boldsymbol{\theta})$ by

$$L_n^{(2)}(\boldsymbol{\theta}) = -\sum_{k=n+1-l}^{n} \log |a_{kn}\boldsymbol{\Sigma}_v| - \sum_{k=n+1-l}^{n} \mathbf{z}_k^{'}[a_{kn}\boldsymbol{\Sigma}_v]^{-1}\mathbf{z}_k \ .$$

Then the SIML estimator of $\hat{\boldsymbol{\Sigma}}^{(v)}$ is defined by

$$\hat{\boldsymbol{\Sigma}}_v = \frac{1}{l_n} \sum_{k=n+1-l_n}^{n} a_{kn}^{-1}\mathbf{z}_k\mathbf{z}_k^{'} \ .$$

For both $\hat{\boldsymbol{\Sigma}}_v$ and $\hat{\boldsymbol{\Sigma}}_x$, the number of terms m_n and l_n should be dependent on n. We need the order requirements that $m_n = O(n^{\alpha})$ $(0 < \alpha < \frac{1}{2})$ and $l_n = O(n^{\beta})$ $(0 < \beta < 1)$ for $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_v$, respectively.

Since we use a linear transformation, we alternatively write

$$\hat{\sigma}_{fg}^{(x)} = \sum_{i,j=1}^{n} c_{ij}(y_{fi} - y_{f,i-1})(y_{gj} - y_{g,j-1}) \ ,$$

where $s_{jk} = \cos\left[\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})\right]$ and

$$
\begin{aligned}
c_{ij} &= \frac{2}{m_n}\sum_{k=1}^{m_n} s_{ik}s_{jk} \;, \\
&= \frac{1}{m_n}\sum_{k=1}^{m_n}\left\{\cos\left[\frac{2\pi}{2n+1}(i+j-1)(k-\frac{1}{2})\right] + \cos\left[\frac{2\pi}{2n+1}(i-j)(k-\frac{1}{2})\right]\right\} \;.
\end{aligned}
$$

3 Summary of Asymptotic Properties of the SIML

We summarize the asymptotic properties of the SIML estimator when the sample size n is large. Kunitomo and Sato [2008a,b] have investigated the problem and have shown that the SIML estimator is consistent and it has the asymptotic normality under a set regularity conditions. For simplicity, we consider the scalar case and write $y_{fi} = x_{fi} + v_{fi}$ $(i = 1, \cdots, n)$ in the discrete time setting (and y_{ft}, x_{ft}, v_{ft} in the continuous setting) and write

$$x_{ft} = x_{f0} + \int_0^t c_s^{(ff)} dB_s^{(f)} .$$

Then the SIML estimator is

$$\hat{\sigma}_{ff}^{(x)} = \frac{1}{m_n} \sum_{k=1}^{m_n} z_{fk}^2 \ ,$$

where z_{fk} $(k = 1, \cdots, m_n)$, correspond to the transformed data for the returns $y_{fi} - y_{f,i-1}$ $(i = 1, \cdots, n)$. Then as $n \longrightarrow \infty$

$$\hat{\sigma}_{ff}^{(x)} - \sigma_{ff}^{(x)} \xrightarrow{p} 0$$

with $m_n = n^\alpha$ $(0 < \alpha < 1/2)$ and

$$\sqrt{m_n}\left[\hat{\sigma}_{ff}^{(x)} - \sigma_{ff}^{(x)}\right] \xrightarrow{d} N\left(0, 2[\sigma_{ff}^{(x)}]^2\right)$$

with $m_n^5/n^2 \to 0$.

Although the SIML estimation was introduced under the Gaussian process and the standard model, it has reasonable finite sample properties as well as asymptotic properties under some volatility models and the non-Gaussian processes with

$$\mathcal{E}\left[(x_{fi} - x_{f,i-1})^2 | \mathcal{F}_{n,i-1}\right] = \int_{t_{i-1}}^{t_i} \sigma_{ff}^{(x)}(s)ds$$

or

$$\frac{1}{n}\sum_{i=1}^{n} \sigma_{ff}^{(x)}(t_{i-1}^n) \longrightarrow \sigma_{ff}^{(x)} = \int_0^1 \sigma_{ff}^{(x)}(s)ds \ .$$

As $n \longrightarrow \infty$, under a set of regularity conditions, the asymptotic distribution of the SIML estimator can be summarized as

$$\sqrt{m_n}\left[\hat{\sigma}_{ff}^{(x)} - \sigma_{ff}^{(x)}\right] \xrightarrow{d} N\left[0, V\right] \ ,$$

provided that we have the convergence of the asymptotic variance

$$\begin{aligned} V &= 2\left[\int_0^1 \sigma_{ff}^{(x)}(s)ds\right]^2 \\ &\quad +2\ \mathrm{plim}_{n\to\infty} \sum_{i,j=1}^{n} (m_n c_{ij}^2 - 1)\left[\int_{t_{i-1}}^{t_i} \sigma_{ff}^{(x)}(s)ds\right]\left[\int_{t_{j-1}}^{t_j} \sigma_{ff}^{(x)}(s)ds\right] \end{aligned}$$

and it is a positive constant when $m_n^5/n^2 \to 0$ (as $n \to \infty$). When V is a random variable, the convergence is in the sense of *stable convergence*.

4 Asymptotic Robustness of SIML

There is a natural question on the finite sample properties of the SIML estimation when the underlying assumptions are not valid, in particular when the micro-market noises are autocorrelated and endogenous.

Let $z_{in}^{(1)}$ and $z_{in}^{(2)}$ $(i = 1, \cdots, n)$ be the i-th elements of

$$\mathbf{z}_n^{(1)} = h_n^{-1/2} \mathbf{P}_n^{'} \mathbf{C}_n^{-1} (\mathbf{x}_n - \bar{\mathbf{y}}_0) \ , \ \mathbf{z}_n^{(2)} = h_n^{-1/2} \mathbf{P}_n^{'} \mathbf{C}_n^{-1} \mathbf{v}_n ,$$

respectively, where $\mathbf{x}_n^{(f)} = (x_{fi})$, $\mathbf{v}_n^{(f)} = (v_{fi})$ and $\mathbf{z}_n = (z_{in})$ are $n \times 1$

vectors with $z_{in} = z_{in}^{(1)} + z_{in}^{(2)}$. Then we use

$$\begin{aligned}
\sqrt{m_n}\left[\hat{\sigma}_{ff}^{(x)} - \sigma_{ff}^{(x)}\right] &= \sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n} z_{kn}^2 - \sigma_{ff}^{(x)}\right] \\
&= \sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n} z_{kn}^{(1)2} - \sigma_{ff}^{(x)}\right] + \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\mathcal{E}[z_{kn}^{(2)2}] \\
&\quad + \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\left[z_{kn}^{(2)2} - \mathcal{E}[z_{kn}^{(2)2}]\right] + 2\frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\left[z_{kn}^{(1)} z_{kn}^{(2)}\right] .
\end{aligned}$$

Then we shall investigate the conditions that three terms except the first one are $o_p(1)$. It is because we could estimate the realized volatility consistently as if there were no noise term in this situation.

Let $\mathbf{b}_k = \mathbf{e}_k^{'}\mathbf{P}_n^{'}\mathbf{C}_n^{-1} = (b_{kj})$ and $\mathbf{e}_k^{'} = (0,\cdots,1,0,\cdots)$ be an $n \times 1$ vector. We write $z_{kn}^{(2)} = \sum_{j=1}^{n} b_{kj} v_{fj}$ and notice that $\sum_{j=1}^{n} b_{kj} b_{k^{'}j} = \delta(k,k^{'}) a_{kn}$. Also we shall use the notation that K_i $(i \geq 1)$ are some positive constants. First we impose the condition

$$(\mathbf{I}) \quad \mathcal{E}[v_{fi} v_{fj}] = c_1 \rho^{|i-j|} \ (0 \leq \rho < 1) ,$$

where c_1 is a constant.

For instance, we use

$$\frac{1}{m_n}\sum_{k=1}^{m_n} a_{kn} = \frac{1}{m_n} 2n \sum_{k=1}^{m_n}\left[1-\cos(\pi\frac{2k-1}{2n+1})\right]$$

$$= \frac{n}{m_n}\left[2m_n - \frac{\sin\pi\frac{2m_n}{2n+1}}{\sin\pi\frac{1}{2n+1}}\right] = O(\frac{m_n^2}{n})$$

Then the second term becomes

$$\frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\mathcal{E}[z_{kn}^{(2)}]^2 \le K_1 \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n} a_{kn} = O(\frac{m_n^{5/2}}{n}),$$

which is negligible if $0 < \alpha < 0.4$. For the third term, we need to consider the variance of

$$z_{kn}^{(2)2} - \mathcal{E}[z_{kn}^{(2)2}] = \sum_{j,j'=1}^{n} b_{kj}b_{k,j'}\left[v_{fj}v_{f,j'} - \mathcal{E}(v_{fj}v_{f,j'})\right]$$

and we impose the additional condition

$$(\mathbf{II}) \qquad \mathcal{E}\left[(v_{fi}v_{f,i'} - \mathcal{E}(v_{fi}v_{f,i'}))(v_{f,i''}v_{f,i'''} - \mathcal{E}(v_{f,i''}v_{f,i'''}))\right]$$
$$= c_2 \rho^{\frac{1}{2}(|i-i^{'}|+|i^{''}-i^{'''})|} \ (0 \leq \rho < 1) \ ,$$

where c_2 is a constant. The condition (II) is satisfied for the (weakly dependent) linear processes on $\{v_{fj}\}$ with bounded 4th order moments.

Theorem 1 : Assume Conditions (I) and (II), the moment condition that $\mathcal{E}(v_{fi}^4)$ are bounded, and $0 < \alpha < 1/3$ for $m_n = O(n^\alpha)$. Then the asymptotic distribution is asymptotically $(m_n, n \to \infty)$ equivalent to that of $\sqrt{m_n}\left[(1/m_n)\sum_{k=1}^{m_n} z_{kn}^{(1)2} - \sigma_{ff}^{(x)}\right]$.

In our derivation the only term involved in the correlation of noise and signal is the fourth term. and then it is interesting to find the condition that they can be ignored for estimating the realized volatility and covariance. Then the sufficient condition we need is

$$(\mathbf{I}^{'}) \quad \mathcal{E}[v_{fi}v_{fj}|r_k, k=1,\cdots,n] = c_3\rho_1^{|i-j|}\ (0 \leq \rho_1 < 1)\ a.s.,$$

where c_3 is bounded.

Theorem 2 : Instead of Condition (I) in Theorem 1, assume Conditions $(\mathrm{I})^{'}$ and (II), $0 < \alpha < 1/3$ for $m_n = O(n^{\alpha})$, and relax the independence assumption between the signal term $\{x_{fi}\}$ and noise term $\{v_{fi}\}$. Then the results of Theorem 1 hold.

5 Simulations

For the illustrative purpose, we give two Tables when the micro-market noises have MA(1) structure. Although MA(1) (the 1st order moving average process) is a simple stochastic process, the results of Tables 6 and 7 indicate the general properties of the SIML estimation. We note that if we knew the true process MA(1) in advance, it is certainly better to use the maximum likelihood (ML) estimation which should be efficient, but we had experiments for the situations when we do not have such information on the noise process.

Second, we give one table when the micro-market noise is correlated with the true price process. There can be many possible ways to formulate the dependence structure between the micro-market noise and the efficient market prices. As a representative example, we give some results when the current market noise depends on the innovation of impact of the efficient market price process at the same period and the market noise term has the MA(1) structure. The SIML estimation gives reasonable and stable results even in this situation. (Kunitomo and Sato (2010a)).

Table 6 : Estimation of Realized Volatility (MA(1) noise, $a = 0.5$)

n=300	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.03E-04	3.53E-06	1.87E-03	2.03E-04	5.18E-07	3.68E-04	1.99E-04	1.86E-07
SD	9.62E-05	5.12E-07	1.99E-04	9.69E-05	7.45E-08	3.21E-05	9.35E-05	2.69E-08
MSE	9.27E-09	2.61E-12		9.40E-09	1.07E-13		8.74E-09	3.44E-14
AVAR	8.17E-09	8.34E-14		8.17E-09	8.34E-16		8.17E-09	8.34E-20
n=5000	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.04E-04	3.52E-06	2.82E-02	2.00E-04	3.62E-07	3.00E-03	2.00E-04	1.38E-08
SD	5.27E-05	1.63E-07	7.61E-04	5.12E-05	1.69E-08	7.94E-05	5.09E-05	6.45E-10
MSE	2.79E-09	2.35E-12		2.62E-09	2.65E-14		2.59E-09	1.39E-16
AVAR	2.65E-09	8.79E-15		2.65E-09	8.79E-17		2.65E-09	8.79E-21
n=20000	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.05E-04	3.56E-06	1.12E-01	1.98E-04	3.57E-07	1.14E-02	2.00E-04	6.09E-09
SD	3.95E-05	9.59E-08	1.53E-03	4.00E-05	9.58E-09	1.54E-04	3.92E-05	1.62E-10
MSE	1.58E-09	2.43E-12		1.61E-09	2.48E-14		1.54E-09	1.67E-17
AVAR	1.52E-09	2.90E-15		1.52E-09	2.90E-17		1.52E-09	2.90E-21

Table 7 : Estimation of Realized Volatility (MA(1) noise, $a = -0.5$)

n=300	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.07E-04	8.40E-07	9.17E-04	2.02E-04	2.47E-07	2.71E-04	2.02E-04	1.82E-07
SD	9.59E-05	1.25E-07	8.22E-05	9.65E-05	3.65E-08	2.26E-05	9.40E-05	2.66E-08
MSE	9.24E-09	1.36E-12		9.32E-09	3.53E-15		8.83E-09	3.33E-14
AVAR	8.17E-09	8.34E-14		8.17E-09	8.34E-16		8.17E-09	8.34E-20
n=5000	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.05E-04	4.96E-07	1.22E-02	2.02E-04	5.89E-08	1.40E-03	2.00E-04	1.08E-08
SD	5.21E-05	2.34E-08	2.75E-04	5.30E-05	2.81E-09	3.10E-05	5.29E-05	4.88E-10
MSE	2.75E-09	2.26E-12		2.81E-09	1.99E-14		2.79E-09	7.71E-17
AVAR	2.65E-09	8.79E-15		2.65E-09	8.79E-17		2.65E-09	8.79E-21
n=20000	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$	H-vol	$\sigma_{ff}^{(x)}$	$\sigma_{ff}^{(v)}$
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.06E-04	4.52E-07	4.82E-02	2.01E-04	4.75E-08	5.00E-03	2.01E-04	2.99E-09
SD	4.01E-05	1.23E-08	5.47E-04	3.84E-05	1.26E-09	5.58E-05	4.01E-05	8.07E-11
MSE	1.64E-09	2.40E-12		1.48E-09	2.33E-14		1.61E-09	9.88E-19
AVAR	1.52E-09	2.90E-15		1.52E-09	2.90E-17		1.52E-09	2.90E-21

Table 8 : Estimation of Realized Volatility (MA(1) and Endogenous noise, $a = 0.5, \rho = 0.5, l = 0$)

n=300	σ_x^2	σ_v^2	H-vol	σ_x^2	σ_v^2	H-vol	σ_x^2	σ_v^2
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.01E-04	2.55E-06	1.35E-03	2.04E-04	5.68E-07	3.81E-04	1.99E-04	2.05E-07
SD	9.55E-05	3.79E-07	1.44E-04	9.55E-05	8.31E-08	3.50E-05	9.29E-05	3.02E-08
MSE	9.11E-09	4.48E-13		9.14E-09	1.42E-13		8.63E-09	4.23E-14
AVAR	8.17E-09	8.34E-14		8.17E-09	8.34E-16		8.17E-09	8.34E-20
n=5000	σ_x^2	σ_v^2	H-vol	σ_x^2	σ_v^2	H-vol	σ_x^2	σ_v^2
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.01E-04	1.95E-06	1.55E-02	2.01E-04	2.44E-07	2.00E-03	2.01E-04	1.78E-08
SD	5.21E-05	9.14E-08	4.20E-04	5.14E-05	1.13E-08	5.24E-05	5.18E-05	8.43E-10
MSE	2.72E-09	1.09E-14		2.64E-09	2.03E-15		2.68E-09	2.51E-16
AVAR	2.65E-09	8.79E-15		2.65E-09	8.79E-17		2.65E-09	8.79E-21
n=20000	σ_x^2	σ_v^2	H-vol	σ_x^2	σ_v^2	H-vol	σ_x^2	σ_v^2
true-val	2.00E-04	2.00E-06		2.00E-04	2.00E-07		2.00E-04	2.00E-09
mean	2.00E-04	1.87E-06	5.87E-02	2.00E-04	2.10E-07	6.60E-03	1.99E-04	7.25E-09
SD	4.03E-05	5.12E-08	8.14E-04	3.92E-05	5.60E-09	8.82E-05	3.87E-05	1.96E-10
MSE	1.63E-09	1.98E-14		1.54E-09	1.22E-16		1.50E-09	2.76E-17
AVAR	1.52E-09	2.90E-15		1.52E-09	2.90E-17		1.52E-09	2.90E-21

Data generating process:

$$y_t = x_t + \sqrt{\sigma_v^2/(1+a^2)}v_t$$
$$x_t = x_{t-1} + \sqrt{\sigma_x^2/n}u_t$$
$$v_t = \epsilon_t - a\epsilon_{t-1}$$
$$\epsilon_t = (1-\rho)w_t + \rho u_{t-l}$$
$$u_t \sim i.i.d.N(0,1), w_t \sim i.i.d.N(0,1)$$

6 Conclusions

1. The SIML estimator is simple and it has reasonable statistical properties.
2. We show the asymptotic robustness of the SIML estimator by simulations. We have compared SIML with the realized kernel method by Bandorff-Nielsen et al. (2008), which needs the information variance-ratio for determining the bandwidth parameter in advance.
3. The SIML estimator is also simple and useful for multivariate high frequency series.
4. We have applied the SIML method to the Nikkei-225 futures and spot indexes with the realizing hedging problem, which are major stock indexes in Japan.

References

[1] Kunitomo, N. and S. Sato (2008a), Separating Information Maximum Likelihood Estimation of Realized Volatility and Covariance with Micro-Market Noise, CIRJE Discussion Paper F-581, University of Tokyo, 2008a, (http://www.e.u-tokyo.ac.jp/cirje/research/).

[2] Kunitomo, N. and S. Sato (2008b), Realized Volatility, Covariance and Hedging Coefficient of the Nikkei-225 Futures with Micro-Market Noise, CIRJE Discussion Paper F-601, University of Tokyo.

[3] Kunitomo, N. and S. Sato (2010a), Robustness of the Separating Information Maximum Likelihood Estimation of Realized Volatility with Micro-Market Noise, CIRJE Discussion Paper F-733, University of Tokyo.

[4] Kunitomo, N. and S. Sato (2010b), "The SIML Estimation of Realized Volatility of Nikkei-225 Futures and Hedging Coefficient with Micro-Market Noise," forthcoming in *Mathematics and Computers in Simulation*, Elsevier.