Linear regression of drift in continuous semimartingale models

Masaaki Fukasawa*
Department of Mathematics, Osaka University
1-1 Machikaneyama, Toyonaka, Osaka, 560-0043 JAPAN
fukasawa@math.sci.osaka-u.ac.jp

Abstract

We study statistical estimation of linear parameters in the drift coefficient of stochastic differential equation models. Explanatory variable is a continuous semimartingale. Observation is high-frequent and irregular. When sampling times are correlated to the explanatory variable, the standard least square estimator is shown to be biased up to first order. We prove a central limit theorem and propose a bias correction based on estimates of skewness in the explanatory variable.

1 Model

We observe discretely two processes X and Y and assume simple dynamics

$$\mathrm{d}Y_t = \vartheta X_t \mathrm{d}[Z]_t + \sigma \mathrm{d}Z_t$$

with unknown parameter $\vartheta \in \Theta \subset \mathbb{R}^p$, where Z is a continuous local martingale with strictly increasing quadratic variation $[Z]$ and $\sigma > 0$ is a constant. The regressand Y is one-dimensional and the regressor X is p-dimensional. We are interested in estimating ϑ from data

$$\{(X_{\tau_j}, Y_{\tau_j}, [Z]_{\tau_j}); j = 0, 1, 2, \dots, N-1\},$$

where $\tau = \{\tau_j\}$ is an increasing sequence of stopping times with $\tau_0 = 0$ and N is the number of data obtained in a fixed observation period, say $[0, T]$ with

*The author is appointed also by Center for the Study of Finance and Insurance, Osaka University and Institute of Statistical Mathematics. This work is supported by Japan Science and Technology Agency (CREST) and Japan Society for the Promotion of Science (KAKENHI) Grant Numbers 24684006 and 24300107.

a stopping time T. We do not suppose N is deterministic. Just for notational convenience, assume we observe Y_T and $[Z]_T$ additionally. The constant σ is either known or unknown. The canonical example for Z is a standard Brownian motion, where $[Z]_{\tau_j} = \tau_j$ are the sampling times naturally supposed to be observable. In financial applications, $[Z]$ is interpreted as the business time. In the literature $[Z]_{\tau_j}$ stands for, say, cumulative transaction volume at time τ_j. The absolute continuity of the drift part of Y with respect to $[Z]$ is a natural requirement in terms of the fundamental theorem of asset pricing.

Denote by $\mathcal{L}$ the set of the locally bounded and left-continuous adapted processes. We suppose the elements of X to be of the form

$$\mathrm{d}X_t^i = \hat{X}_t^i \mathrm{d}[M^i]_t + \mathrm{d}M_t^i,\ i = 1, 2, \dots, p$$

with $\hat{X}^i \in \mathcal{L}$ and continuous local martingales M^i which are nondegenerate in the sense that the Radon-Nikodym derivatives of $\mathrm{d}[M^i]$ with respect to $\mathrm{d}[Z]$ are positive and elements of $\mathcal{L}$. This framework includes, say, the Itô case with $p = 1$: Z is a standard Brownian motion and

$$\mathrm{d}X_t = \alpha_t \mathrm{d}t + \beta_t \mathrm{d}W_t,$$

where α, $\beta \in \mathcal{L}$ with $\beta > 0$, and W is a standard Brownian motion possibly correlated to Z. When $p > 1$, we additionally require that the elements of X are orthogonal, that is, $[X^i, X^j] = 0$ for all $i \neq j$ to avoid multicolinearity.

2 Likelihood

First we note that by time-change with respect to $[Z]$, the general model reduces to a simple case where Z is a standard Brownian motion with the sampling times $\hat{\tau}_j = [Z]_{\tau_j}$ and terminal observation time $\hat{T} = [Z]_T$, which are observable by the assumption. To fix ideas, in this section, we assume Z to be a standard Brownian motion and actually this can be done without loss of generality. Remark that even if the original sampling times are deterministic, say $\tau_j = jh$ with $h > 0$, they become stochastic after the time change.

In an idealized situation that X and Y are continuously observed, the log-likelihood is explicitly given as

$$\log \frac{\mathrm{d}P^\vartheta}{\mathrm{d}P^0} = \frac{1}{\sigma^2} \int_0^T \vartheta X_t \mathrm{d}Y_t - \frac{1}{2\sigma^2} \int_0^T |\vartheta X_t|^2 \mathrm{d}t$$

and so, the likelihood and Bayesian analyses run straightforward. The maximum likelihood estimator is $\hat{\vartheta} = \Gamma^{-1}\Delta$, where

$$\Delta^i = \int_0^T X_t^i \mathrm{d}Y_t,\ \ \Gamma^{ij} = \int_0^T X_t^i X_t^j \mathrm{d}t.$$

Bayesian estimators are also written in terms of Δ and Γ. Here we have only discrete samples and naturally consider their discretized versions

$$\Delta^i[\tau] = \int_0^T X^i[\tau]_t \mathrm{d}Y_t, \;\; \Gamma^{ij}[\tau] = \int_0^T X_t^i[\tau] X_t^j[\tau] \mathrm{d}t,$$

where we define $X[\tau]$ as $X[\tau]_t = X_{\tau_j}$ for $t \in [\tau_j, \tau_{j+1})$ and $\tau_N = T$. Notice that the discretized maximum likelihood estimator $\hat{\vartheta}[\tau] = \Gamma[\tau]^{-1}\Delta[\tau]$ coincides with the exact maximum likelihood estimator for a discrete time model

$$Y_{\tau_{j+1}} = Y_{\tau_j} + \vartheta\delta_j X_{\tau_j} + \sigma\sqrt{\delta_j}\varepsilon_{j+1}, \;\; \delta_j = \tau_{j+1} - \tau_j, \; j = 0, 1, \ldots, N-1,$$

with $\{\varepsilon_j\}_{j=1}^{\infty}$ being an iid sequence of the standard normal distribution. Note also that $\hat{\vartheta}[\tau]$ is the least square estimator in that it minimizes

$$\sum_{j=0}^{N-1} \delta_j^{-1} |Y_{\tau_{j+1}} - Y_{\tau_j} - \vartheta\delta_j X_{\tau_j}|^2.$$

If the sampling scheme τ is sufficiently high-frequent, we expect estimators based on these discretized integrals to perform reasonably well. In order to evaluate the discretization error, in the next section we consider a sequence τ^n of sampling schemes with

$$\int_0^1 \|X_t - X[\tau^n]_t\|^2 \mathrm{d}t \to 0$$

in probability as $n \to \infty$ and study the limit distribution of, say, $\Delta - \Delta[\tau^n]$.

3 Central limit theorem

In this section we state mathematical results in a way that they are true in the general situation of Z being not necessarily a Brownian motion. For their proofs, as already discussed, we may assume Z is so without loss of generality. In the original time-scale, the statistics are written as

$$\Delta^i[\tau] = \int_0^T X^i[\tau]_t \mathrm{d}Y_t, \;\; \Gamma^{ij}[\tau] = \int_0^T X_t^i[\tau] X_t^j[\tau] \mathrm{d}[Z]_t.$$

Theorem 3.1 *If there exist $P = (P^1, \ldots, P^p)$, $P^i, Q^i \in \mathcal{L}$, $i = 1, \ldots, p$ and a sequence ε_n with $\varepsilon_n \to 0$ such that*

$$\begin{aligned} \varepsilon_n^{-1} \int_0^{s\wedge T} (X_t^i - X^i[\tau^n]_t) \mathrm{d}[Z]_t &\to \int_0^s P_t^i \mathrm{d}[Z]_t, \\ \varepsilon_n^{-2} \int_0^{s\wedge T} |X_t^i - X^i[\tau^n]_t|^2 \mathrm{d}[Z]_t &\to \int_0^s Q_t^i \mathrm{d}[Z]_t \end{aligned} \tag{1}$$

in probability for all $s \geq 0$*, then* $\varepsilon_n^{-1}(\Delta - \Delta[\tau^n])$ *converges stably to*

$$\int_0^{\cdot} P_t \mathrm{d}Y_t + \sigma \hat{Z},$$

where $\hat{Z}$ *is a* p*-dimensional conditionally Gaussian martingale with*

$$[\hat{Z}^i, \hat{Z}^j] = \delta^{ij} \int_0^{\cdot} Q_t^i \mathrm{d}[Z]_t - \int_0^{\cdot} P_t^i P_t^j \mathrm{d}[Z]_t, \quad \delta^{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Proof: By the Girsanov-Maruyama transformation, we may suppose Y is a local martingale. Then apply Jacod and Shiryaev (2002), Theorem IX.7.3. ////

The main message here is that the natural discretization $\Delta[\tau]$ suffers from a bias if $P \neq 0$. The next result shows when it is the case.

Theorem 3.2 *Suppose* $E[[M^i]_T^6] < \infty$ *for* $i = 1, \dots, p$*. Denote the conditional moments of increments by*

$$G_{ij}^k[\tau] = E[(M_{\tau_{j+1}}^i - M_{\tau_j}^i)^k | \mathcal{F}_{\tau_j}].$$

If there exist $P^i, Q^i \in \mathcal{L}$*,* $i = 1, \dots, p$ *and a sequence* ε_n *with* $\varepsilon_n \to 0$ *such that*

$$G_{ij}^3[\tau^n]/G_{ij}^2[\tau^n] = 3\varepsilon_n P_{\tau_j^n}^i + o_p(\varepsilon_n), \quad G_{ij}^4[\tau^n]/G_{ij}^2[\tau^n] = 6\varepsilon_n^2 Q_{\tau_j^n}^i + o_p(\varepsilon_n^2)$$

and $G_{ij}^6[\tau^n]/G_{ij}^2[\tau^n] = o_p(\varepsilon_n^4)$*,* $G_{ij}^{12}[\tau^n]/G_{ij}^2[\tau^n] = o_p(\varepsilon_n^8)$ *for* $i = 1, \dots p$*, then*

$$\varepsilon_n^{-1} \sum_{j=0}^{N-1} (X_{\tau_{j+1}^n}^i - X_{\tau_j^n}^i)^3 \to 3 \int_0^T P_t^i \mathrm{d}[X^i]_t,$$
$$\varepsilon_n^{-2} \sum_{j=0}^{N-1} (X_{\tau_{j+1}^n}^i - X_{\tau_j^n}^i)^4 \to 6 \int_0^T Q_t^i \mathrm{d}[X^i]_t$$

in probability and (1) holds.

Proof: This is a multi-dimensional version of Fukasawa (2011), Theorem 2.6. ////

From the above results, the bias term appears in cases where the increments of the explanatory variable X exhibit nonzero skewness. Many examples of sampling schemes τ^n which yield nonzero P are given in Fukasawa (2010).

Lemma 3.1 *In addition to the assumptions of Theorem 3.2, suppose that there exist* $R^{ab} \in \mathcal{L}$*,* $a, b = 1, \dots, p$ *such that*

$$F_{ab,j}^3[\tau^n]/F_{ab,j}^2[\tau^n] = 3\varepsilon_n R_{\tau_j^n}^{ab} + o_p(\varepsilon_n), \quad F_{ab,j}^6[\tau^n]/F_{ab,j}^2[\tau^n] = o_p(\varepsilon_n^2)$$

for $a, b = 1, \dots, p$, where

$$F^k_{ab,j}[\tau] = E[(M^a_{\tau_{j+1}} M^b_{\tau_{j+1}} - M^a_{\tau_j} M^b_{\tau_j})^k | \mathcal{F}_{\tau_j}].$$

Then

$$\varepsilon_n^{-1} \sum_{j=0}^{N-1} (X^a_{\tau^n_{j+1}} X^b_{\tau^n_{j+1}} - X^a_{\tau^n_j} X^b_{\tau^n_j})^3 \to 3 \int_0^T R^{ab}_t \mathrm{d}[X^a X^b]_t,$$
$$\varepsilon_n^{-1} \int_0^T (X^a_t X^b_t - X^a[\tau^n]_t X^b[\tau^n]_t) \mathrm{d}[Z]_t \to \int_0^T R^{ab}_t \mathrm{d}[Z]_t$$

in probability.

Theorem 3.3 *Under the assumptions of Theorem 3.2 and Lemma 3.1, the scaled discretization error $\varepsilon_n^{-1}(\vartheta - \vartheta[\tau^n])$ of the maximum likelihood estimator converges stably to a mixed normal distribution with mean*

$$\Gamma^{-1} \int_0^T P_t \mathrm{d}Y_t + \Gamma^{-1} \int_0^T R_t \mathrm{d}[Z]_t \Gamma^{-1} \Delta$$

and covariance matrix $V = [V^{ij}]$ defined as

$$V^{ij} = \delta^{ij} \int_0^T Q^i_t \mathrm{d}[Z]_t - \int_0^T P^i_t P^j_t \mathrm{d}[Z]_t.$$

For bias correction, estimates of P and R are required. The both P and R are defined in terms of conditional moments and so, the standard kernel technique can be applied to obtain their estimates.

References

[1] M. Fukasawa: Realized volatility with stochastic sampling, Stochastic Process. Appl. 120 (2010) 829-852.

[2] M. Fukasawa: Discretization error of stochastic integrals, Ann. Appl. Probab. 21 (2011) 1436-1465.

[3] J. Jacod and A. Shiryaev: *Limit theorems for stochastic processes*, 2nd ed., Springer-Verlag (2002).