Analyzing high-frequency stock price data and learning volatility using neural networks

Teppei Ogihara

Graduate School of Information Science and Technology, University of Tokyo

Abstract

Forecasting variances of stocks and covariances of stock pairs is an important task to control the loss from stock assets for many financial institutions which hold a huge amount of stocks. The study of high-frequency data becomes more important because huge information of high-frequency data enable us to forecast stock variances and covariances more accurately. However, there are two problems on statistical analysis of high-frequency data; market microstructure noise and nonsynchronous observations.

We consider a two-dimensional stochastic process $X = \{X_t\}_{0\le t\le T}$ satisfying the following stochastic differential equation:

$$dX_t = \mu(t, X_t, \sigma_*)dt + b(t, X_t, \sigma_*)dW_t, \quad t \in [0, T],$$

where $\{W_t\}_{0\le t\le T}$ is a two-dimensional standard Wiener process, μ is a $\mathbb{R}^2$-valued Borel function, $b = \{b^{ij}\}_{i,j=1}^2$ is a $\mathbb{R}^2 \otimes \mathbb{R}^2$-valued Borel function, and $\sigma_* \in \mathbb{R}^d$ is a parameter. For this model, we consider noisy, nonsynchronous observations of X_t^1 and X_t^2. [1] construct a maximum likelihood estimator from the quasi-likelihood function and introduce properties such as asymptotic mixed normality, local asymptotic normality.

We develop asymptotic theory in cases where the parametric model does not include the true parameters (misspecified model) and incorporate neural networks into this quasi-likelihood analysis to learn the intraday stock price model from high-frequency data. We will study asymptotic theory of a maximum-likelihood-type estimator for misspecified model. In this setting, the maximum-likelihood-type estimator cannot attain the optimal convergence rate $n^{-1/4}$ due to the asymptotic bias. We construct a new estimator which attains the optimal rate by using a bias correction, and show the asymptotic mixed normality.

References

[1] Ogihara, T. (2018): Parametric Inference for Nonsynchronously Observed Diffusion Processes in the Presence of Market Microstructure Noise, Bernoulli, 24(4B) 3318–3383.

[2] Ogihara, T. (2021): Misspecified diffusion models with high-frequency observations and an application to neural networks, Stochastic Process. Appl. 142, 245–292.