Computational Statistics and Data Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Optimal design of Fourier estimator in the presence of microstructure noise

Fangfang Wang *

Department of Information and Decision Sciences, University of Illinois at Chicago, 601 South Morgan Street, Chicago, IL 60607-7124, United States

ARTICLE INFO

Article history:
Received 22 October 2012
Received in revised form 4 August 2013
Accepted 4 August 2013
Available online xxxx

Keywords:
Discrete Fourier transform
Quadratic variation
Microstructure noise
Periodogram
Spectral density
Whittle likelihood

ABSTRACT

The Fourier estimator of Malliavin and Mancino depends on both sample size and a so-called cutting frequency. The latter controls the number of Fourier coefficients to be included, and it also determines how the Fourier estimator responds to market microstructure noise. By examining the finite sample properties of the Fourier estimator, an easy-to-implement procedure is developed for the optimal cutting frequency which minimizes the mean squared error in the presence of the microstructure noise, along with a modified Whittle likelihood approach for the estimation of the signal-to-noise ratio.

1. Introduction

The study on the integrated volatility of a log-price process begins with the analysis of realized volatility which is a sum of finely sampled squared intra-daily returns (see for instance Andersen and Bollerslev (1998), Andersen et al. (2001), Meddahi (2002), Barndorff-Nielsen and Shephard (2002), among others). In theory, the realized volatility provides a consistent estimate of integrated volatility under certain regularity conditions. The justification relies on the observability of the true price process sampled at any frequency. The presence of market microstructure noise, however, prevents us from sampling the price process too frequently—due to price discreteness, bid–ask bounce, etc. The realized volatility calculated from ultra-high frequency returns turns out to be an estimate of the variation of the noise, rather than the latent true price process (see for instance Aït-Sahalia et al. (2005)).

A noise-adjusted (or noise-robust) estimator of integrated volatility in the presence of microstructure noise has been studied extensively in the literature. See the work of Zhou (1996), Andersen et al. (1999), Malliavin and Mancino (2002), Barndorff-Nielsen and Shephard (2004), Zhang et al. (2005), Zhang (2006), Hansen and Lunde (2006), Kalnina and Linton (2008), Barndorff-Nielsen et al. (2008), Bandi and Russell (2008), Mancino and Sanfelici (2008), Malliavin and Mancino (2009), Olhede et al. (2009), Jacod et al. (2009), Mykland and Zhang (2009), Bandi and Russell (2011), Høg and Lunde (2003), Capobianco (2004), Fan and Wang (2007), Subbotin (2008), Barunik and Vacha (2012), among many others.

In this paper, we will revisit the Fourier estimator of integrated volatility proposed by Malliavin and Mancino (2002). Assuming that the true log-price process is a continuous semi-martingale, Malliavin and Mancino (2002) introduces the continuous time Fourier transform of the price process and volatility process over a finite time interval, and links them via the Bohr convolution. In particular, the integrated volatility is the Fourier transform of the volatility process at zero

* Tel.: +1 312 996 2679; fax: +1 312 413 0385.
E-mail address: ffwang@uic.edu.

0167-9473/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2013.08.003

frequency. The Fourier estimator of integrated volatility is constructed by discretizing the Fourier transform of the price process at the discretely observed time points. Properties of the Fourier estimator are further explored by Barucci and Reno (2002a,b), Mancino and Sanfelici (2008), Nielsen and Frederiksen (2008), Malliavin and Mancino (2009), among others.

The Fourier estimator is determined by two factors: the sample size n, and a so-called cutting frequency N. The latter controls the number of Fourier coefficients to be included in the Fourier transform. As pointed out by Mancino and Sanfelici (2008), 'the choice of a suitable N allows to render the Fourier estimator invariant to short-run noise introduced by market microstructure effects, with consequent efficiency gains'. Mancino and Sanfelici (2008) provides a numerical method to find the optimal cutting frequency which minimizes the mean squared error (MSE). The minimization is performed by comparing the computed MSE values over distinct integer-valued N. Through a simulation study, Mancino and Sanfelici (2008) suggest that the optimal cutting frequency should be derived using quote-to-quote returns. In this paper, we will present a mathematical framework for the derivation of the optimal cutting frequency. By fixing the sample size, we give an explicit asymptotic expression for the optimal cutting frequency which is of order $n^{1/3}$. We also study theoretically and numerically the behavior of the mean squared error at the optimal cutting frequency.

The rest of the paper is organized as follows. Section 2 gives a review on the Fourier estimator of Malliavin and Mancino. Its connection with periodograms is discussed in Section 3. Section 4 talks about how to find the optimal cutting frequency. We give a periodogram-based estimation approach in Section 5. Section 6 considers a real data application. Section 7 concludes the paper. Technical details and the proofs are deferred to an Appendix.

2. Fourier analysis of integrated volatility

In this section, we give a review on the Fourier estimator of integrated volatility introduced by Malliavin and Mancino (2002, 2009).

Suppose that the log-price process $\{X(t), 0 \le t \le T\}$ is a Brownian semi-martingale satisfying the following stochastic differential equation

$$dX(t) = \mu(t)dt + \sigma(t)dW(t), \tag{1}$$

where W is a standard Brownian motion, and μ, σ are adapted stationary processes such that $E[\int_0^T \mu^2(t)dt] < \infty$ and $E[\int_0^T \sigma^4(t)dt] < \infty$. Define the Fourier coefficients of the instantaneous volatility $\sigma(t)$ and the log-price process $X(t)$ at $k \in \mathbb{Z}$ as follows: $\mathcal{F}(\sigma^2)(k) = \frac{1}{T}\int_0^T e^{-2\pi ikt/T}\sigma^2(t)dt$ and $\mathcal{F}(dX)(k) = \frac{1}{T}\int_0^T e^{-2\pi ikt/T}dX(t)$, where $i = \sqrt{-1}$. Malliavin and Mancino (2002, 2009) point out that,

$$\mathcal{F}(\sigma^2)(k) = \lim_{N\to\infty} \frac{T}{c(N_0, N)} \sum_{N_0 \le |s| \le N} \mathcal{F}(dX)(s)\mathcal{F}(dX)(k-s) \quad \text{in probability},$$

for any integer k, where $N_0 \ge 0$ and $c(N_0, N) = 2N + 1 - \max(2N_0 - 1, 0)$. Malliavin and Mancino (2002) considers the case where $N_0 > 0$, while $N_0 = 0$ is discussed by Malliavin and Mancino (2009). When $k = 0$,

$$\int_0^T \sigma^2(t)dt = \lim_{N\to\infty} \frac{T^2}{c(N_0, N)} \sum_{N_0 \le |s| \le N} \mathcal{F}(dX)(s)\mathcal{F}(dX)(-s) \quad \text{in probability}. \tag{2}$$

(2) is pragmatically applicable only if the whole sample path of $\{X(t), 0 \le t \le T\}$ is observable.

Suppose that the log-price process $X(t)$ is observed at discrete time points $\{0 = t_0 < t_1 < \cdots < t_n = T\}$ where $t_j = Tj/n$ for $j = 0, 1, \ldots, n$. Let $\Delta X_j = X(t_j) - X(t_{j-1})$ be the return over $[t_{j-1}, t_j]$ for $j = 1, 2, \ldots, n$. The Fourier estimator is defined as follows:

$$\hat{\sigma}^2_{N,n}(X) = \frac{T^2}{c(N_0, N)} \sum_{N_0 \le |s| \le N} c_s(X)c_{-s}(X), \tag{3}$$

where $c_k(X) = T^{-1}\sum_{j=1}^n e^{-2\pi ikt_{j-1}/T}\Delta X_j$. Malliavin and Mancino (2002, 2009) show that $\lim_{n,N\to\infty} \hat{\sigma}^2_{N,n}(X) = \int_0^T \sigma^2(t)dt$ in probability. The consistency of $\hat{\sigma}^2_{N,n}(X)$ suggests using returns sampled over fine intervals to construct an 'accurate' estimator of the integrated volatility. However, the presence of microstructure noise prevents us from sampling prices too frequently. When the price is sampled over fine intervals such as every 30 s, the observed price deviates from the efficient price due to the imperfection of the trading process.

We will consider the situation where microstructure noise matters. Let $\{Y_j\}_{j=0}^n$ be the noisy observations of $X(t)$ at time $\{t_j\}_{j=0}^n$. Namely,

$$Y_j = X(t_j) + \epsilon_j, \quad j = 0, 1, 2, \ldots, n.$$

The observation errors $\{\epsilon_j\}$ satisfy the following assumption.

Assumption 2.1. $\{\epsilon_j\}$ is a white noise process with variance σ^2_ϵ, $E(\epsilon_j^4) = \eta\sigma_\epsilon^4 < \infty$. The noise process $\{\epsilon_j\}$ is uncorrelated with X at all lags and leads.

Please cite this article in press as: Wang, F., Optimal design of Fourier estimator in the presence of microstructure noise. Computational Statistics and Data Analysis (2013), http://dx.doi.org/10.1016/j.csda.2013.08.003