

Estimation of the long-run variance of nonlinear time series with an application to change point analysis

Vaidotas Characiejus^a

Joint work with Piotr Kokoszka^b and Xiangdong Meng^b

EcoSta 2025 / Tokyo, Japan, August 22, 2025

^aDepartment of Mathematics and Computer Science, University of Southern Denmark, Denmark

^bDepartment of Statistics, Colorado State University, USA

Outline

Smoothed periodogram estimator of the long-run variance

Weakly dependent nonlinear processes

Limiting distribution of the estimator

Application to change point analysis

Summary

Smoothed periodogram estimator of the long-run variance

Stationary sequences and their autocovariances

Definition

$\{X_t\}_{t \in \mathbb{Z}}$ is a stationary sequence of random variables if

- (a) $E |X_t|^2 < \infty$ for all $t \in \mathbb{Z}$;
- (b) $E X_t = \mu$ for all $t \in \mathbb{Z}$;
- (c) $\text{Cov}(X_r, X_s) = \text{Cov}(X_{r+t}, X_{s+t})$ for all $r, s, t \in \mathbb{Z}$.

Definition

For a stationary sequence $\{X_t\}_{t \in \mathbb{Z}}$,

$$\gamma(h) := \text{Cov}(X_h, X_0) \quad \text{with} \quad h \in \mathbb{Z}$$

is the sequence of autocovariances.

Long-run variance

Definition

For a stationary sequence $\{X_t\}_{t \in \mathbb{Z}}$ with $\sum_{h=1}^{\infty} |\gamma(h)| < \infty$,

$$\sigma_{\infty}^2 := \sum_{h \in \mathbb{Z}} \gamma(h) = \gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h)$$

is the long-run variance (LRV) of $\{X_t\}_{t \in \mathbb{Z}}$.

- For an uncorrelated sequence $\{X_t\}_{t \in \mathbb{Z}}$,

$$\text{Var} \bar{X}_n = \frac{\gamma(0)}{n} \quad \text{for } n \geq 1.$$

- For a stationary sequence $\{X_t\}_{t \in \mathbb{Z}}$ with $\sum_{h=1}^{\infty} |\gamma(h)| < \infty$,

$$\text{Var} \bar{X}_n \sim \frac{\sigma_{\infty}^2}{n} \quad \text{as } n \rightarrow \infty.$$

Spectral density

The LRV is typically estimated nonparametrically in the frequency domain.

Definition

For a stationary sequence $\{X_t\}_{t \in \mathbb{Z}}$ with $\sum_{h=1}^{\infty} |\gamma(h)| < \infty$,

$$f(\lambda) := \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-ih\lambda} \quad \text{for } \lambda \in [-\pi, \pi]$$

is the spectral density of $\{X_t\}_{t \in \mathbb{Z}}$, where $i = \sqrt{-1}$.

Observe that

$$\sigma_{\infty}^2 = 2\pi f(0)$$

and hence we can estimate the LRV by estimating $f(0)$.

Discrete Fourier transform and the periodogram

Definition

The discrete Fourier transform (DFT) and the periodogram of $X_1, \dots, X_n$ with $n \geq 1$ are defined as

$$\mathcal{X}_n(\omega_j) := \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n X_t e^{-it\omega_j} \quad \text{and} \quad I_n(\omega_j) := |\mathcal{X}_n(\omega_j)|^2,$$

respectively, where $\omega_j = 2\pi j/n$ with $j \in \{-(n-1)/2, \dots, \lfloor n/2 \rfloor\}$ are the Fourier frequencies and $i = \sqrt{-1}$.

The periodogram can be viewed as the sample spectral density since

$$I_n(\omega_j) = \frac{1}{2\pi} \sum_{h=-(n-1)}^{n-1} \hat{\gamma}(h) e^{-ih\omega_j},$$

where $\hat{\gamma}(h) := n^{-1} \sum_{t=1}^{n-h} X_{t+h} X_t$ for $h \geq 0$ and $\hat{\gamma}(h) = \hat{\gamma}(-h)$ for $h < 0$.

Smoothed periodogram estimator

- The periodogram is not a consistent estimator of the spectral density function.
- A simple consistent estimator of the spectral density function is obtained by smoothing of the periodogram in the following way

$$\hat{f}(\lambda) = \frac{1}{2\pi} \frac{1}{2m+1} \sum_{|j| \leq m} I_n(g(n, \lambda) + \omega_j),$$

where $m = m_n \rightarrow \infty$ but $m = o(n)$ as $n \rightarrow \infty$ and $g(n, \lambda)$ is the multiple of $2\pi/n$ closest to $\lambda \in [-\pi, \pi]$.

Smoothed periodogram estimator of LRV

- We consider the smoothed periodogram estimator of the LRV given by

$$Q_n = \frac{1}{m} \sum_{j=1}^m I_n(\omega_j),$$

where $m = m_n \rightarrow \infty$ but $m = o(n)$ as $n \rightarrow \infty$.

- We extend the asymptotic normality of Q_n to general nonlinear moving averages, a result previously available only for linear processes.
- Because Q_n relies on a narrow band of local Fourier frequencies, it is particularly relevant for local Whittle estimation of the Hurst exponent. We illustrate this with an application to change point analysis.

Weakly dependent nonlinear processes

Linear process

Definition

$\{X_t\}_{t \in \mathbb{Z}}$ is a linear process if

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \quad \text{for each } t \in \mathbb{Z},$$

where

- $\{a_j\}_{j \geq 0} \subset \mathbb{R}$ is such that $\sum_{j=0}^{\infty} a_j^2 < \infty$;
- $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is a sequence of independent and identically distributed random variables with $\mathbf{E} \varepsilon_0 = 0$ and $\mathbf{E} \varepsilon_0^2 < \infty$.

The asymptotic normality of Q_n under the assumption of linearity follows from the central limit theorem for weighted sums of periodogram ordinates established by Giraitis and Koul (2013).

Nonlinear moving average

Many commonly considered stationary processes are nonlinear (for example, bilinear, threshold, GARCH, stochastic volatility models) and not covered by the currently existing asymptotic results.

Definition

$\{X_t\}_{t \in \mathbb{Z}}$ is a nonlinear moving average if

$$X_t = g(\varepsilon_t, \varepsilon_{t-1}, \dots) \quad \text{for each } t \in \mathbb{Z}, \quad (1)$$

where

- $g : S^\infty \rightarrow \mathbb{R}$ is a measurable function;
- $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is a sequence of independent and identically distributed random elements with values in a measurable space S .

L^p - m -approximable nonlinear moving average

Consider a modified version of the nonlinear moving average in (1) defined by

$$\chi_t^{(m)} = g(\varepsilon_t, \dots, \varepsilon_{t-m+1}, \varepsilon'_{t-m}, \varepsilon'_{t-m-1}, \dots),$$

where $\{\varepsilon'_t\}_{t \in \mathbb{Z}}$ is an independent copy of $\{\varepsilon_t\}_{t \in \mathbb{Z}}$.

Definition

A nonlinear moving average is L^p - m -approximable if

$$\sum_{t=1}^{\infty} \|X_t - \chi_t^{(t)}\|_p < \infty$$

with $p \geq 1$, where $\|\cdot\|_p = (\mathbf{E} |\cdot|^p)^{1/p}$.

Hörmann and Kokoszka (2010)

Limiting distribution of the estimator

Assumption 1

$\{X_t\}_{t \in \mathbb{Z}}$ is a nonlinear moving average and

$$\|X_t - X_t^{(m)}\|_4 = O(m^{-\eta}) \quad \text{as } m \rightarrow \infty \quad (2)$$

with some $\eta > 3$.

Assumption 1 implies that

- $\{X_t\}_{t \in \mathbb{Z}}$ is L^4 - m -approximable;
- $\sum_{h=1}^{\infty} h^r |\gamma(h)| < \infty$ for $0 \leq r < \eta - 1$.

Theorem 1

Let $\{X_t\}_{t \in \mathbb{Z}}$ be a nonlinear moving average such that $EX_0 = 0$ and $E|X_0|^4 < \infty$. Suppose that Assumption 1 is satisfied and $m \rightarrow \infty$ but $m = o(n^{4/5})$ as $n \rightarrow \infty$. Then

$$\frac{\sqrt{m}}{f(0)}[Q_n - f(0)] \xrightarrow{D} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

The key component of the proof is a general central limit theorem for quadratic forms established by Liu and Wu (2010) (see Theorem 6 therein).

Application to change point analysis

Long-range dependent and change point models

- Long-range dependent and change point models can produce similar patterns in time plots, autocovariances, and spectra despite being fundamentally different.
- Several approaches have been proposed to distinguish between long-range dependence and change-point models (for example, Baek and Pipiras (2012) and Norwood and Killick (2018)).
- We show that the approach of Baek and Pipiras (2012) is applicable also to nonlinear moving averages satisfying the assumptions of Theorem 1.

Change point model

The observations $X_1, \dots, X_n$ satisfy

$$X_t = \begin{cases} \mu + R_t, & 1 \leq t \leq n^*, \\ \mu + \Delta + R_t, & n^* + 1 \leq t \leq n, \end{cases} \quad (3)$$

where

- μ is the base level mean;
- Δ is the magnitude of change at an unknown change point n^* ;
- $\{R_t\}_{t \in \mathbb{Z}}$ is a weakly dependent stationary sequence.

Hurst exponent

Definition 2

Suppose $\{X_t\}_{t \in \mathbb{Z}}$ is a stationary sequence of random variables. If its spectral density satisfies

$$f(\lambda) = |\lambda|^{1-2H}g(\lambda)$$

with some $0 < H < 1$, where g is a continuous function on $[-\pi, \pi]$ with $g(\lambda) > 0$ in the neighbourhood of $\lambda = 0$, then H is called the Hurst exponent of $\{X_t\}_{t \in \mathbb{Z}}$.

- If $H = 1/2$, $\{X_t\}_{t \in \mathbb{Z}}$ is called short-range or weakly dependent.
- If $H > 1/2$, $\{X_t\}_{t \in \mathbb{Z}}$ is called long-range dependent.
- Observe that if $H > 1/2$, then $f(\lambda) \rightarrow \infty$ as $\lambda \rightarrow 0$.

Change point versus long-range dependence

The objective is to test

H_0 : model (3) holds with $\{R_t\}_{t \in \mathbb{Z}}$ satisfying Definition 2 with $H = 1/2$

versus

H_1 : $\{X_t\}_{t \in \mathbb{Z}}$ satisfies Definition 2 with $H > 1/2$.

Local Whittle estimator of the Hurst exponent

Definition

The local Whittle estimator (LWE) of the Hurst exponent H is defined as

$$\hat{H} = \arg \min_{H \in \Theta} L(H)$$

where $\Theta = [\Delta_1, \Delta_2]$ with $0 < \Delta_1 < \Delta_2 < 1$,

$$L(H) = \log \left\{ \frac{1}{m} \sum_{l=1}^m \omega_l^{2H-1} I_n(\omega_l) \right\} - (2H-1) \frac{1}{m} \sum_{l=1}^m \log \omega_l,$$

and m is the number of low frequencies used in the estimation.

Δ_1 and Δ_2 can be taken arbitrarily near 0 and 1, or set to reflect prior beliefs (e.g., $\Delta_1 = 1/2$).

Change point

A potential change point can severely bias the LWE so its effect is removed by considering the residuals

$$\hat{R}_t = \begin{cases} X_t - \hat{n}^{-1} \sum_{i=1}^{\hat{n}} X_i, & 1 \leq t \leq \hat{n}, \\ X_t - (n - \hat{n})^{-1} \sum_{i=\hat{n}+1}^n X_i, & \hat{n} + 1 \leq t \leq n, \end{cases}$$

where $\hat{n}$ is the standard CUSUM estimator

$$\hat{n} = \min \left\{ k : \left| \sum_{j=1}^k X_j - \frac{k}{n} \sum_{j=1}^n X_j \right| = \max_{1 \leq s \leq n} \left| \sum_{j=1}^s X_j - \frac{s}{n} \sum_{j=1}^n X_j \right| \right\}.$$

The test statistic is

$$T^{(\hat{R})} = 2\sqrt{m}\left(\hat{H}^{(\hat{R})} - \frac{1}{2}\right),$$

where $\hat{H}^{(\hat{R})}$ is the LWE based on the residuals $\{\hat{R}_t\}_{1 \leq t \leq n}$.

Theorem 3

Suppose that model (3) holds with $\{R_t\}_{1 \leq t \leq n}$ satisfying $E R_0 = 0$ and $E |R_0|^4 < \infty$ and Assumption 1. In addition, assume that

- (a) $n^* = \lfloor n\theta \rfloor$ with some $\theta \in (0, 1)$;
- (b) the change in mean level $\Delta = \Delta(n)$ and the change point estimator $\hat{n}$ satisfy

$$n\Delta^2 \rightarrow \infty, \quad \Delta^2 |\hat{n} - n^*| = O_p(1) \quad \text{as } n \rightarrow \infty;$$

(c)

$$\frac{1}{m} + \frac{m^5(\log m)^2}{n^4} + \frac{m(\log m)^2}{n\Delta^2} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Then $T^{(\hat{R})} \xrightarrow{D} \mathcal{N}(0, 1)$ as $n \rightarrow \infty$.

Idea of the proof of Theorem 3

The proof of Theorem 3 can be reduced to known arguments, Theorem 1 and the following convergence that we establish for nonlinear moving averages and under certain additional assumptions

$$\frac{1}{\sqrt{m}} \sum_{j=1}^m \nu_{l,m} \left(\frac{l_n(\omega_j)}{f(0)} - 1 \right) \xrightarrow{D} \mathcal{N}(0, 1) \quad \text{as } n \rightarrow \infty,$$

where

$$\nu_{l,m} = \log l - \frac{1}{m} \sum_{l'=1}^m \log l'$$

are the specific weights that follow from the LWE.

Rejection rule

Provided that certain assumptions hold,

$$T^{(\hat{R})} \xrightarrow{D} \mathcal{N}(0, 1) \quad \text{as } n \rightarrow \infty$$

under the null hypothesis and hence we reject the null hypothesis if

$$T^{(\hat{R})} > z_{1-\alpha},$$

where $z_{1-\alpha}$ is the $(1 - \alpha)$ th quantile of the standard normal distribution.

Definition

$\{R_t\}_{t \in \mathbb{Z}}$ is said to be a GARCH(1, 1) process if it satisfies the equations

$$R_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

where

- $\alpha_0 > 0, \alpha_1, \beta_1 \geq 0, \alpha_1 + \beta_1 < 1$;
- $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ are independent and identically distributed random variables such that $E \varepsilon_0 = 0$, $E \varepsilon_0^2 = 1$, and $E \varepsilon_0^4 < \infty$.
- $\alpha_1 + \beta_1 < 1 \Rightarrow$ the GARCH(1, 1) process has a stationary solution.
- $E R_0^4 \iff E |\beta_1 + \alpha_1 \varepsilon_0^2|^2 < 1$
- $E |\beta_1 + \alpha_1 \varepsilon_0^2|^2 < 1 \implies$ Assumption 1.

Empirical size

Nominal size	1%			5%			10%		
n	1000	5000	10000	1000	5000	10000	1000	5000	10000
$\Delta = 0.30, \theta = 0.25$									
$m = n^{0.60}$	0.90	1.12	1.68	4.00	5.00	6.48	7.86	9.96	11.56
$m = n^{0.70}$	0.92	1.18	1.46	3.98	5.00	5.58	7.86	9.68	9.90
$\Delta = 0.30, \theta = 0.50$									
$m = n^{0.60}$	0.64	0.68	0.78	3.14	3.36	4.20	6.26	6.58	8.28
$m = n^{0.70}$	0.80	0.70	0.98	3.38	3.96	4.06	6.68	7.36	7.66
$\Delta = 0.30, \theta = 0.75$									
$m = n^{0.60}$	0.76	1.08	1.72	3.66	5.02	5.96	7.26	10.00	11.08
$m = n^{0.70}$	0.80	1.12	1.62	3.88	4.90	5.66	7.80	9.38	10.18
$\Delta = 0.60, \theta = 0.25$									
$m = n^{0.60}$	1.68	1.76	1.58	6.24	6.04	6.00	11.12	11.10	11.32
$m = n^{0.70}$	1.52	1.44	1.64	5.70	6.06	6.04	10.78	11.04	10.24
$\Delta = 0.60, \theta = 0.5$									
$m = n^{0.60}$	0.78	0.74	0.72	3.82	3.74	4.32	7.68	7.24	8.62
$m = n^{0.70}$	0.86	0.84	1.00	3.98	4.18	4.20	7.60	7.82	8.12
$\Delta = 0.60, \theta = 0.75$									
$m = n^{0.60}$	1.58	1.82	1.66	6.16	6.32	6.26	10.84	11.14	11.46
$m = n^{0.70}$	1.32	1.58	1.94	5.86	6.12	5.98	10.86	10.48	10.66

Summary

Summary

- We consider the smoothed periodogram estimator of the long-run variance.
- The limiting distribution of the estimator is derived for nonlinear moving averages, a result previously available only for linear processes.
- We demonstrate how our result can be used in local Whittle estimation of the Hurst exponent and in distinguishing long-range dependence from change point models.
- For more details, see Characiejus, Kokoszka, and Meng (2025).

<https://imada.sdu.dk/u/characiejus/>

References



Baek, Changryong and Vladas Pipiras (Jan. 2012). “Statistical tests for a single change in mean against long-range dependence”. In: *Journal of Time Series Analysis* 33.1, pp. 131–151.



Characiejus, Vaidotas, Piotr Kokoszka, and Xiangdong Meng (2025). “Estimation of the Long-Run Variance of Nonlinear Time Series With an Application to Change Point Analysis”. In: *Journal of Time Series Analysis*.



Giraitis, Liudas and Hira L. Koul (2013). “On asymptotic distributions of weighted sums of periodograms”. In: *Bernoulli* 19.5B, pp. 2389–2413.



Hörmann, Siegfried and Piotr Kokoszka (2010). “Weakly dependent functional data”. In: *The Annals of Statistics* 38.3, pp. 1845–1884.



Liu, Weidong and Wei Biao Wu (2010). “ASYMPTOTICS OF SPECTRAL DENSITY ESTIMATES”. In: *Econometric Theory* 26.4, pp. 1218–1245.



Norwood, Ben and Rebecca Killick (2018). “Long memory and changepoint models: a spectral classification procedure”. In: *Statistics and Computing* 28.2, pp. 291–302.



Robinson, P. M. (1995). “Gaussian Semiparametric Estimation of Long Range Dependence”. In: *The Annals of Statistics* 23.5, pp. 1630–1661.