An extension to the applicability of adaptive bandwidth choice

Dehao Dai∗
June 10, 2022

Abstract
In this paper, we extend the applicability of the bandwidth choice method of Politis (2003) by relaxing the conditions of his Theorem 2.3.

1 Introduction
Assume \(X_1, \cdots, X_N\) are data from the strictly stationary time series \(\{X_n, n \in \mathbb{Z}\}\) with mean \(\mu = \mathbb{E}X_t\), and autocovariance \(\gamma(k) = \mathbb{E}(X_t - \mu)(X_{t+|k|} - \mu)\); here both \(\mu\) and \(\gamma(\cdot)\) are unknown. Now we consider the problem of estimating the spectral density function \(f(\omega) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} e^{ik\omega}\gamma(k)\), for \(\omega \in [-\pi, \pi]\) with kernel \(\Lambda_h\). The corresponding kernel estimator may be written as:

\[
\hat{f}(\omega) = \Lambda_h \ast I_N(\omega)
\]

where \(I_N(\omega) = (2\pi)^{-1} \sum_{k=-N+1}^{N-1} e^{ik\omega}\gamma(k)\) is the periodogram and \(\ast\) denotes convolution; here, the real number \(h\) is the bandwidth. In Politis (2003), the kernel \(\Lambda_h(\omega)\) is defined as

\[
\Lambda_h(\omega) = \frac{1}{2\pi} \sum_{k=-M}^{M} \lambda^T(k/M)e^{ik\omega}
\]

where \(h = 1/M\) is regarded as the bandwidth and the function \(\lambda^T(\cdot)\) has a trapezoidal shape symmetric around zero, i.e.

\[
\lambda^T(t) = \begin{cases} 
1 & \text{if } |t| \in [0, c] \\
\frac{1}{1-c}(1 - |t|) & \text{if } t \in [c, 1] \\
0 & \text{else.}
\end{cases}
\]

∗Department of Mathematics, Univ. of California, San Diego; email: ddai@ucsd.edu
Note that $\lambda^T$ depends on the breaking point $c$.

An alternative form to represent $\hat{f}(w)$ in Politis (2003) is the Fourier transform of the kernel, namely $\lambda_h(k) = \lambda^T(hk)$

$$
\hat{f}(w) = (2\pi)^{-\frac{1}{2}} \sum_{k=-\infty}^{\infty} e^{iwk} \lambda^T(kh) \hat{\gamma}(k)
$$

where $\hat{\gamma}(k) = N^{-1} \sum_{i=1}^{N-|k|} (X_i - \bar{X}_N)(X_{i+|k|} - \bar{X}_N)$ is the lag-$k$ sample autocovariance for $|k| < N$ and $\hat{\gamma}(k)$ is defined to be zero for $|k| \geq N$.

Politis (2003) proposed the following empirical rule of picking $M$ with $c = 1/2$.

**EMPIRICAL RULE OF PICKING M:** Let $\rho(k) = \gamma(k)/\gamma(0)$, and $\hat{\rho}(k) = \hat{\gamma}(k)/\hat{\gamma}(0)$. Let $\hat{m}$ be the smallest positive integer such that $|\hat{\rho}(\hat{m}+k)| < c\sqrt{\log N/N}$, for $k = 1, \ldots, K_N$, where $c > 0$ is a fixed constant, and $K_N$ is a positive, nondecreasing integer-valued function of $N$ such that $K_N = o(\log N)$. Then, let $M = \hat{m}/c = 2\hat{m}$.

Theorem 2.2 of Politis (2003) showed that the above empirical rule performs well in three cases for $\rho(k)$, namely (i) polynomial decay, (ii) exponential decay, and (iii) hard cut-off. In the next section, we complement his Theorem 2.2 by providing two weaker formulations of cases (i) and (ii).

## 2 Main result

**Corollary:** Assume conditions strong enough to ensure that for all finite $n$,

$$
\max_{k=1, \ldots, n} |\hat{\rho}(s+k) - \rho(s+k)| = O_p(1/\sqrt{N})
$$

uniformly in $s$, and

$$
\max_{k=0,1, \ldots, n-1} |\hat{\rho}(k) - \rho(k)| = O_p\left(\sqrt{\frac{\log N}{N}}\right).
$$

Also assume that the sequence $\gamma(k)$ does not have more than $K_N - 1$ zeros in its first $k_0$ lags (i.e., for $k = 1, \ldots, k_0$).

(i) Assume that $\gamma(k) = \sum_{i=1}^{l} C_i k^{-d_i} \cos(a_i k + \theta_i)$ for $k > k_0$, and for some $C_i > 0$ and positive integers $d_i$ and $l$, and some distinct constants $a_i \geq \frac{\pi}{K_N}, \theta_i \in [0, 2\pi]$ for all $i = 1, \ldots, l$. Then,

$$
\hat{M} \overset{P}{\sim} A_1 \left(\frac{N}{\log N}\right)^{1/2d}
$$

for some positive constant $A_1$, and $d = \min_{i \in [l]} \{d_i\}$.
(ii) Assume $\gamma(k) = \sum_{i=1}^{l} C_i \xi_k \cos(a_i k + \theta_i)$ for $k > k_0$, where $C > 0$, $|\xi| < 1$, $a_i \geq \frac{\pi}{K_N}$, $\theta_i \in [0, 2\pi]$ and positive integer $l$ are some constants. Then,

$$\hat{M} \overset{p}{\sim} A_2 \log N$$

where $A_2 = -\frac{1}{\log |\xi|}$ with $|\xi| = \max_{i \in [l]} \{|\xi_i|\}$.

**Proof** of (i): First note that from the condition that the sequence $\gamma(k)$ does not have more that $K_N - 1$ zeros in its first $k_0$ lags with (1), it follows that $\hat{m} > k_0$ with high probability, so we can focus on the part of the correlogram to the right of $k_0$, i.e., look at $\gamma(k)$ for $k > k_0$ only.

Now condition $a_i \geq \frac{\pi}{K_N}$ implies that $K_N \geq \max\{\pi/a_i\}$, i.e. $K_N$ is bigger than all the half-periods of the $\cos(a_i k + \theta)$ for $i = 1, \cdots, l$. Hence, (2) implies that

$$\max_{k=1,\cdots,K_N} |\hat{\rho}(\hat{m} + k)| \geq C(\hat{m} + K_N)^{-d} + O_P(\sqrt{\log N/N}),$$

(3) where $d = \min_{i \in [l]} \{d_i\}$, since

$$\sum_{i=1}^{l} C_i k^{-d_i} = C k^{-d} \left( \frac{C_{L}}{C} + \sum_{i=1, i \neq L}^{l} \frac{C_i k^{-d_i}}{C} \right) \geq C k^{-d}$$

where $L = \arg\min_{i \in [l]} \{d_i\}$ and $C = \max_{i \in [l]} \{C_i\}$.

From part (i) assumption with (2), it follows that

$$|\hat{\rho}(k)| = |C k^{-d}| + O_P(\sqrt{\log N/N}) \quad \text{uniformly in } k,$$

where $d = \min_{i \in [l]} \{d_i\}$, i.e.,

$$|\hat{\rho}(\hat{m})| = |C \hat{m}^{-d}| + O_P(\sqrt{\log N/N}).$$

(4)

The empirical rule for picking $M$ implies that $|\hat{\rho}(\hat{m})| \geq c \sqrt{\log N/N}$, whereas

$$\max_{k=1,\cdots,K_N} |\hat{\rho}(\hat{m} + k)| < c \sqrt{\log N/N}.$$

Thus, the above two statements with (3) and (4)

$$A_1 \left( \frac{N}{\log N} \right)^{1/2d} - K_N \leq \hat{m} \leq A_1 \left( \frac{N}{\log N} \right)^{1/2d}$$

with high probability. Since $K_N = o(\log N)$,

$$\hat{M} \overset{p}{\sim} A_1 \left( \frac{N}{\log N} \right)^{1/2d},$$

3
where $A_1$ is some positive constant and $d = \min_{i \in [l]} \{d_i\}$.

**Proof** of (ii): It is similar to proof of part (i). First, Hence, (2) implies that

$$
\max_{k=1,\ldots,K_N} |\hat{\rho}(\hat{m} + k)| \geq C|\xi|^{|\hat{m}| + K_N} + O_p(\sqrt{\log N/N}) \tag{5}
$$

where $|\xi| = \max_{i \in [l]} \{|\xi_i|\}$, since

$$
\sum_{i=1}^l C_i|\xi_i|^k = C|\xi|^k \left( C_L/C + \sum_{i=1,i \neq L} C_i \frac{|\xi_i|}{|\xi|}^k \right) \geq C|\xi|^k
$$

where $L = \arg \max_i \{|\xi_i|\}$ and $C = \max_{i \in [l]} \{C_i\}$.

From part (ii) assumption with (2), it follows that

$$
|\hat{\rho}(k)| = |C\xi^k| + O_p(\sqrt{\log N/N}) \text{ uniformly in } k,
$$

where $|\xi| = \max_{i \in [l]} \{|\xi_i|\}$, i.e.,

$$
|\hat{\rho}(\hat{m})| = |C\xi^{\hat{m}}| + O_p(\sqrt{\log N/N}) \tag{6}
$$

Thus, the two statements of the empirical rule for picking $M$ mentioned in part (i) with (5) and (6) implies

$$A_2 \log N - K_N \leq \hat{m} \leq A_2 \log N$$

with high probability. Thus,

$$\hat{M} \overset{P}{\sim} A_2 \log N$$

where $A_2 = -1/\log |\xi|$ with $|\xi| = \max_{i \in [l]} \{|\xi_i|\}$.

Note that a general sufficient condition guaranteeing eq. (1) and (2) has been given by Xiao & Wu (2012).

**References**

Politis, D. N. (2003), ‘Adaptive Bandwidth Choice’, *Journal of Nonparametric Statistics* **15**(4-5), 517–533.

Politis, D. N. (2011), ‘Higher-order accurate, positive semidefinite estimation of large-sample covariance and spectral density matrices’, *Econometric Theory* **27**(4), 703–744.

Xiao, H. & Wu, W. B. (2012), ‘Covariance matrix estimation for stationary time series’, *The Annals of Statistics* **40**(1), 466–493.