The Korean Communications
in Statistics Vol. 3, No. 3, 1996
pp. 271-282

Convergence Properties of a Spectral Density Estimator

Gyeong Hye Shin[1], Hae Kyung Kim[2]

Abstract

This paper deals with the estimation of the power spectral density function of time series. A kernel estimator which is based on local average is defined and the rates of convergence of the pointwise, L_2-norm and L_∞-norm associated with the estimator are investigated by restricting as to kernels with suitable assumptions. Under appropriate regularity conditions, it is shown that the optimal rate of convergence for $0<r<1$ is N^{-r} both in the pointwise and L_2-norm, while $N^{r-1}(\log N)^{-r}$ is the optimal rate in the L_∞-norm. Some examples are given to illustrate the application of main results.

1. Introduction

The problem of estimating the second order characteristics, namely the covariance function $r_X(m)=E[X_t X_{t+m}]$, $m=0,\pm 1,\pm 2,\cdots$, or its Fourier transform, the spectral density $f_X(\omega)$, is one of the most important problems of time series analysis dealing with stationary random process $X=(X_t)$. For this reason, many different methods for estimating the density function $f_X(\omega)$ have developed and widely used in many fields of science. Especially, the periodogram approach based on the periodogram $I_N(\omega)=\frac{1}{N}|Y_N(\omega)|^2$, where

$$Y_N(\omega) = \sum_{k=1}^{N} X_k e^{-i\omega k}, \qquad \omega \in (-\pi,\pi),$$

which is constructed over the sample $X_1, X_2, \cdots, X_N$ plays a central role in the spectral analysis of time series since the introduction of the fast Fourier transform for evaluating

1) Lecturer, Department of mathematics, Yonsei University, Seoul, 120-749, Korea.
2) Professor, Department of mathematics, Yonsei University, Seoul, 120-749, Korea.

discrete Fourier transforms. This approach for estimating the spectral density is computational efficient and produces reasonable results for a large class of time series or random processes. In particular, if X is strictly stationary and summable second and fourth order cumulant functions, the periodogram $I_N(\omega)$ is asymptotically unbiased as an estimate of the spectral density. This means that the bias of the periodogram as an estimate of the spectral density $f_X(\omega)$ tends to zero rapidly as $N \to \infty$ under reasonable smoothness conditions on $f_X(\omega)$. However, in spite of the advantages, the periodogram is an unsatisfactory estimator of the spectral density since it is inconsistent in the sense that it does not converge to $f_X(\omega)$in mean square. Thus, this drawback of the periodogram is still troublesome when analyzing long data records. However, it is known that for any two fixed neighboring frequencies, ω_1, ω_2, the covariance $Cov\ [I_N(\omega_1), I_N(\omega_2)]$ tends to zero as $N \to \infty$. This suggests that an appropriate smoothing of the periodogram with a weight function, which does not restrict the possible form of function with the finite set of parameters, might lead to a reasonable estimate.

Now we consider a nonparametric estimator for the spectral density function: for $\omega \in [-\pi, \pi]$,

$$\hat{f}_X(\omega) = \frac{1}{2\pi} \sum_{h=-m_N}^{m_N} W_N(h) I_N(g(N,\omega)+\omega_h) \tag{1.1}$$

where $W_N(\cdot)$ is a real function which decreases at a suitable rate relative to N, $\omega_h = 2\pi h/N$is the Fourier frequency at $h = 0, 1, \cdots, [N/2]$, $g(N,\omega)$ is the nearest Fourier frequency to ω(the smaller one if there are two) and m_N is some positive integer which depends on N.

This estimator was introduced by Brockwell and Davis (1991) and further investgated by Shin (1994). As the usual methods of spectral analysis, the above estimator have associated with window functions W_N which are independent of the data or the properties of the time series which is analyzed. This window function relates the average estimated spectrum to the true spectrum. Instead of the fixed neighborhood, the estimator employs varying neighborhoods which depend on the fixed number of frequencies in the frequency domain. The estimation procedure first estimates the autocorrelation lags from the sample data, Fourier transforms the estimates to obtain the periodogram, and then windows the periodogram in an appropriate manner to obtain the power spectral density estimate.

There are many different forms of W_N which we can use, all of which lead to consistent estimates of $f_X(\omega)$. In choosing a weight function it is necessary to compromise between bias and variance of the spectral estimator. In order for the estimate $\hat{f}_X(\omega)$ to be optimal including consistency we impose basically the following assumptions on m_N and W_N :

Assumption A :

$A1$. $m_N \to \infty$ and $m_N/N \to 0$, *as* $N \to \infty$.

$A2$. $W_N(h) = W_N(-h)$ for *all* h.

$A3$. $\sum_{|h| \le m_N} W_N(h) \to 1$, *as* $N \to \infty$.

$A4$. $\sum_{|h| \le m_N} |W_N(h)|$ *is bounded* for *all* N.

$A5$. $\sum_{|h| \le m_N} W_N^2(h) \to 0$, *as* $N \to \infty$.

Furthermore, throughout this paper we restrict ourselves to a general linear process $X_t = \sum_{k=-\infty}^{\infty} a_k Z_{t-k}$ with $\sum_{k=-\infty}^{\infty} |a_k| |k|^{1/2} < \infty$ and $EZ_1^4 < \infty$, where Z_t are independent and identically distributed random variables with zero mean and variance σ^2. The condition $\sum_{k=-\infty}^{\infty} |a_k| |k|^{1/2} < \infty$ implies $\sum_{k=-\infty}^{\infty} a_k^2 < \infty$ which ensures that (X_t) is stationary with finite variance.

Under Assumption A, the estimator $\hat{f}_X(\omega)$ is weakly consistent for $f_X(\omega)$. In Brockwell and Davis (1991), it is shown that $\hat{f}_X(\omega)$ converges in mean square to $f_X(\omega)$ uniformly on $(-\pi, \pi)$. In fact, the assumptions $A1$ and $A3$ ensure that $E\hat{f}_X(\omega) \to f_X(\omega)$, and $A5$ ensures that $Var\, \hat{f}_X(\omega) \to 0$, uniformly in ω. The asymptotic distribution of $\hat{f}_X(\omega)$ was proved in Rosenblatt (1984) indirectly through the normality of $\hat{f}_X(\omega)$ under assumptions on weight function and cumulants which are similar to the assumptions $A1-A4$. Moreover, to gain more information on these asymptotic properties, Shin (1994) considered the rates of convergence of the estimators. Under some appropriate conditions, they proved that this estimator posses the optimal rate of convergence N^{-r} both pointwise and in the L_2-norm, and the optimal rate of convergence $N^{r-1}(\log N)^{-r}$ in the L_∞-norm.

In this paper we confine our attention to this optimal rate of convergence of the estimator and decide the bandwidth and the form of the spectral window achieving the optimal rate of convergence, mainly through some specific cases which are commonly used in time series analysis.

The rest of this paper is organized as follows. The asymptotic properties of the estimator

$\hat{f}_X(\omega)$ are investigated in Section 2, along with the results on rates of convergence. The bandwidth and the form of the spectral window achieving the optimal rate of convergence are given in Section 3. The proofs for the main theorems of Section 2 are given in Section 4.

2. Optimal Convergence

In general, the asymptotic optimal properties of the estimator $\hat{f}_X(\omega)$ depend more critically on the choice of bandwidth, m_N, than on the functional form of the spectral window W_N. In the particular frequency ω, therefore, the optimal rate of convergence of the estimators for each criterion mainly depends on the choice of window only through r. However, in the asymptotic theory of the spectrum the smoothness conditions of the spectral density function of the underlying process is basically required. We start this section by introducing some smoothness conditions on spectral density function which ensure the optimal rate of convergence properties of the estimator $\hat{f}_X(\omega)$. The spectral density function has to fulfill certain smoothness requirements to be specified below:

Assumption B :

B1. $f_X(\omega)$ *is continuous on* $[-\pi, \pi]$.

B2. $\sum_{|h|<\infty} |r_X(h)||h| = B < \infty$.

As measures of the quality or optimal criteria of the estimator $\hat{f}_X(\omega)$ we will use the pointwise, L_2-norm and L_∞-norm.

Now, we shall state the main theorems of this paper, which give more local consistency results of the estimator and the proofs will be given later on. For the variables a_N and b_N we shall use the notation $a_N \sim b_N$ if $a_N / b_N \to 1$ as $N \to \infty$. Given random variables X_N, $N \geq 1$, let $X_N = O_p(a_N)$ mean that the random variables $a_N^{-1} X_N$, $N \geq 1$ are bounded in probability or, equivalently, that

$$\lim_{c \to \infty} \limsup_N P(|X_N| > c a_N) = 0.$$

Theorem 2.1. Suppose that Assumption B is fulfilled and that the following conditions hold: for $\delta_N \sim N^{-r}$ with $0 < r < 1$,

(i) $m_N \sim N \delta_N$.

(*ii*) $|\sum_{|h|\leq m_N} W_N(h) - 1| = O(N^{-r})$,

(*iii*) $\sum_{|h|\leq m_N} W_N^2(h) = O(N^{-2r})$.

Then $|\hat{f}_X(\omega) - f_X(\omega)| = O_p(N^{-r}), \quad \omega \in [-\pi, \pi]$.

Theorem 2.2. Suppose $\delta_N \sim N^{-r}$ and that the conditions in Theorem 2.1 are fulfilled. Then $\| \hat{f}_X(\omega) - f_X(\omega) \|_2 = O_p(N^{-r})$.

Theorem 2.3. Suppose that Assumption B is fulfilled and that the following conditions hold: for $\delta_N \sim N^{r-1}(\log N)^{-r}$ with $0 < r < 1$,

(*i*) $m_N \sim N\delta_N$.

(*ii*) $|\sum_{|h|\leq m_N} W_N(h) - 1| = O(\delta_N)$,

(*iii*) $\sum_{|h|\leq m_N} W_N^2(h) = O(N^{-s})$ for $s > x+2$ where x is the proper positive integer. Then there exists a positive constant C such that for every $\omega \in [-\pi, \pi]$,

$$\lim_N P(\| \hat{f}_X(\omega) - f_X(\omega) \|_\infty \geq CN^{r-1}(\log N)^{-r}) = 0.$$

Remark. The results in the above theorems are optimal in the minimax sense according to Stone (1982). That is, N^{-r} is said to be achievable rate of convergence and the estimate $\hat{f}_X(\omega)$ is said to be asymptotically optimal.

3. Some Windows with Optimality

In this section, we consider the bandwidth and the form of the spectral window achieving the optimal rate of convergence, mainly through some specific cases which are commonly used in time series analysis.

As observed before, the asymptotic optimal properties of the estimator $\hat{f}_X(\omega)$ depend more critically on the choice of bandwidth, m_N, than on the functional form of the spectral window W_N. Therefore, in the particular frequency ω the optimal rate of convergence of the estimators for each criterion mainly depends on the choice of window only through r.

First we consider the approximating form of the estimator $\hat{f}_X(\omega)$. From Assumption A, the function W_N in (1.1) is symmetric with it weights inversely proportional to the magnitude of the lag h, so that we can express

$$\hat{f}_X(\omega) \cong \frac{1}{2\pi} \sum_{|h|\le[\frac{N}{2}]} W_N(h) I_N(g(N,\omega)+\omega_h)$$

which is approximately equal to

$$\frac{1}{2\pi} \sum_{|h|\le[\frac{N}{2}]} W(\omega_h) I_N(\omega+\omega_h) \frac{2\pi}{N}$$

$$\cong \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\lambda) I_N(\omega+\lambda)\, d\lambda = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\omega-\lambda) I_N(\lambda)\, d\lambda \tag{3.1}$$

where $W(\omega_h)=\frac{N}{2\pi} W_N(h)$, $|h|\le[N/2]$. Let T_N be a function of N such that $T_N\to\infty$ and $T_N/N\to 0$ as $N\to\infty$. Define $w(\frac{h}{T_N}) = \int_{-\pi}^{\pi} W(\omega)e^{i\omega h}\, d\omega$; $|h|\le T_N$, so that $W(\omega) = \frac{1}{2\pi}\sum_{|h|\le T_N} w\left(\frac{h}{T_N}\right)e^{-i\omega h}$, then the right hand side of (3.1) is

$$\left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \sum_{|h|\le T_N} w\left(\frac{h}{T_N}\right) e^{-i(\omega-\lambda)h} I_N(\lambda)\, d\lambda$$

$$= \frac{1}{2\pi}\sum_{|h|\le T_N} w\left(\frac{h}{T_N}\right) \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\lambda h} I_N(\lambda)\, d\lambda\, e^{-i\omega h}.$$

Thus we have an approximation

$$\hat{f}_X(\omega) \cong \frac{1}{2\pi}\sum_{|h|\le m_N} w\left(\frac{h}{m_N}\right)\hat{r}(h)e^{-i\omega h}.$$

Next, we decide the bandwidth m_N through r for optimal convergence rate of the estimator $\hat{f}_X(\omega)$ for three criteria. From the definitions we have

$$\sum_{|h|\le m_N} W_N^2(h) = \left(\frac{2\pi}{N}\right)^2 \sum_{|h|\le m_N} W^2(\omega_h)$$

$$\le \frac{2\pi}{N}\sum_{|h|\le[\frac{N}{2}]} W^2(\omega_h)\frac{2\pi}{N} \tag{3.2}$$

$$\cong \frac{2\pi}{N}\int_{-\pi}^{\pi} W^2(\omega)\, d\omega.$$

Furthermore, if we extend w as a bounded continuous even function on R^1 such that $|w(x)|\le 1$, $w(0)=1$, $w(x)=w(-x)$, $w(x)=0$; $|x|>1$ then the last expression with the integral term of (3.2) is equal to

$$\frac{2\pi}{N}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\sum_{|h|\le T_N} w\left(\frac{h}{T_N}\right)e^{-i\omega h}\right)^2 d\omega$$

$$= \frac{1}{N}\sum_{|h|\le T_N} w^2\left(\frac{h}{T_N}\right) = \frac{T_N}{N}\int_{-1}^{1} w^2(x)\, dx.$$

Now, let $m_N = N/T_N$. Then we obtain

$$N^{2r}\sum_{|h|\le m_N} W_N^2(h) \le \frac{N^{2r}}{m_N}\int_{-1}^{1} w^2(x)\, dx \le N^{3r-1}\cdot R. \tag{3.3}$$

which is bounded only when $r=1/3$ since the order of m_N is $O(N^{1-r})$, where R is some constant.

Thus, if the spectral window W_N (or the corresponding lag window) has the lag window generator $w(x)$, and moreover if $m_N=O(N^{2/3})$ in the estimator $\hat{f}_X(\omega)$, then the main theorems guarantee that the estimator possesses the optimal rate of convergence $N^{-1/3}$ both pointwise and in the L_2-norm; and the optimal rate of convergence $(N^2\log N)^{-1/3}$ in the L_∞-norm. Fortunately, almost all the windows including Truncated, Bartlett, Daniell, Parzen, Blackman-Tukey, Bartlett-Priestley window etc, which are commonly used in time series analysis are of the scale parameter form of $w(x)$, the some exceptions being the Whittle, Daniell and Bartlett window.

Finally, for the practical application, we consider a simple example (the moving average smoothing);

$$W_N(h) = \frac{1}{2m_N+1}\cdot J_{\{|h|\le m_N\}}(h)$$

where J_A is the indicator function on A. Assume that the time series X_t is the linear process satisfying the underlying conditions. Evidently, the weight function satisfies Assumption A. Moreover, since $\sum_{|h|\le m_N} W_N^2(h)=\frac{1}{2m_N+1}$, if we impose $m_N=O(N^{1-r})$ with $r=1/3$ then $N^{2r}\sum_{|h|\le m_N} W_N^2(h) = \frac{N^{2r}}{2N^{1-r}+1}$ is bounded. The main results in Section 3 therefore guarantee that the optimal convergence rate of the estimator $\hat{f}_X(\omega)$ in both pointwise and L_2 sense is $N^{-1/3}$ for any frequency ω on $[-\pi,\pi]$.

4. Proofs of Main Theorems

This section contains proofs for the main theorems of Section 2, which are the result of the rates of convergence for the pointwise, L_2-norm and L_∞-norm.

For notational simplicity we shall write m for m_N, the dependence on N being understood.

Proof of Theorem 2.1. It sufficies to show that for sufficiently large C,

$$P\left(|\hat{f}_X(\omega) - f_X(\omega)| \geq CN^{-r}\right) \to 0, \quad as \quad N \to \infty.$$

Note that for all ω, $|\hat{f}_X(\omega) - f_X(\omega)|$ is less than or equal to

$$\left|\sum_{|h|\leq m} W_N(h)\left\{\frac{1}{2\pi} I_N(g(N,\omega) + \omega_h) - f_X(g(N,\omega) + \omega_h)\right\}\right|$$

$$+ \left|\sum_{|h|\leq m} W_N(h)\right|\left| f_X(g(N,\omega) + \omega_h) - f_X(\omega)\right| \tag{4.1}$$

$$+ | f_X(\omega)| \left|\sum_{|h|\leq m} W_N(h) - 1\right|.$$

Since $f_X(\omega)$ is continuous, the third term of (4.1) is $O(N^{-r})$ in virtue of (ii). By the definition of spectral density function $f_X(\omega)$, $| f_X(g(N,\omega) + \omega_h) - f_X(\omega)|$ becomes:

$$\left|\frac{1}{2\pi}\sum_{h=-\infty}^{\infty} r_X(h)\exp(-ih\omega)\ [\exp\{-ih(g(N,\omega) + \omega_h - \omega)\} - 1]\right|$$

which is less than or equal to

$$\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}|r_X(h)||\exp\{-ih(g(N,\omega) + \omega_h - \omega)\} - 1|$$

$$\leq \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}|r_X(h)||h||g(N,\omega) + \omega_h - \omega|$$

$$\leq \frac{B}{2\pi}\left(\frac{\pi}{N} + \frac{2\pi m}{N}\right).$$

Now we may deduce from (i) that the second term of (4.1) is equivalent to $O(N^{-r})$. Therefore, it suffices to show that as $N \to \infty$

$$P\left(\left|\sum_{|h|\leq m} W_N(h)\left\{\frac{1}{2\pi} I_N(g(N,\omega) + \omega_h) - f_X(g(N,\omega) + \omega_h)\right\}\right| \geq CN^{-r}\right) \to 0 .$$

Put $\frac{1}{2\pi} I_N(g(N,\omega) + \omega_h) - f_X(g(N,\omega) + \omega_h) = U_h f_X(g(N,\omega) + \omega_h)$ where the sequence $\{U_h\}$is approximately $WN(0,1)$ for large N by Theorem 10.3.2 in Brockwell and Davis (1991).

For sufficiently large N,

$$P\Big(\big| \sum_{|h| \le m} W_N(h) U_h f_X(g(N,\omega) + \omega_h) \big| \ge CN^{-r} \Big)$$

$$\le \frac{\sum_{|h| \le m} W_N^2(h) f_X^2(g(N,\omega) + \omega_h)}{(CN^{-r})^2} .$$

The last expression is less than or equal to

$$\max_{|h| \le m} f_X^2(g(N,\omega) + \omega_h) \sum_{|h| \le m} W_N^2(h)$$

which converges to 0 by (iii) and the continuity of $f_X(\omega)$. Therefore the proof is complete.

Proof of Theorem 2.2. We show that

$$P\Big(\Big\{ \int_0^\pi | \hat{f}_X(\omega) - f_X(\omega)|^2 \, d\omega \Big\}^{1/2} \ge CN^{-r} \Big) \to 0 \quad as \quad N \to \infty.$$

By the asymptotic unbiasedness of $\hat{f}_X$, it is enough to show that

$$P\Big(\int_0^\pi | \hat{f}_X(\omega) - E\, \hat{f}_X(\omega)|^2 \, d\omega \ge C^2 N^{-2r} \Big) \to 0 \quad as \quad N \to \infty.$$

Note that

$$P\Big(\int_0^\pi | \hat{f}_X(\omega) - E\, \hat{f}_X(\omega)|^2 \, d\omega \ge C^2 N^{-2r} \Big) \le \frac{E \int_0^\pi | \hat{f}_X(\omega) - E\, \hat{f}_X(\omega)|^2 \, d\omega}{C^2 N^{-2r}} . \tag{4.2}$$

The right hand side of (4.2) is also less than or equal to

$$\frac{\pi \Big\{ \sum_{|h| \le m} W_N^2(h) f_X^2(\omega) + o\Big(\sum_{|h| \le m} W_N^2(h) \Big) + \frac{C_2}{(2\pi)^2 N} \sum_{|h| \le m} W_N^2(h)(2m+1) \Big\}}{C^2 N^{-2r}} .$$

The inequality (4.2) is obtained by Markov's inequality and the last expression is obtained by Theorem 10.4.1 in Brockwell and Davis (1991). Now the result follows by the conditions (i), (iii) in Theorem 2.1.

Proof of Theorem 2.3. In order to prove the theorem it is convenient to restrict $\hat{f}_X(\omega)$ to a grid $G_N \subset [0, \pi]$ and define a new estimator $\bar{f}_X(\omega)$ on all of $[0, \pi]$. Set $D_N = [\delta_N^{-1}]$, where [ ] denotes the greatest integer function. Let G_N be the collection of $(D_N + 1)$ points in $[0, \pi]$ each of whose coordinates is of the form $j\pi / D_N$ for some integer j such that $0 \le j \le D_N$. Correspondingly, $[0, \pi]$ can be written as the union of D_N subintervals, each having length $\lambda_N = D_N^{-1}$ and all of its vertices in G_N. Hence for each $\omega \in [0, \pi]$, there is a subinterval Q_t with center t such that $\omega \in Q_t$.

Let C_N denote the collection of centers of these subintervals. Then

$$P(\sup_{\omega\in[0,\pi]} | \hat{f}_X(\omega) - f_X(\omega)| \geq CN^{r-1}(\log N)^{-r})$$
$$= P(\max_{t\in C_N}\sup_{\omega\in Q_t} | \hat{f}_X(\omega) - f_X(\omega)| \geq CN^{r-1}(\log N)^{-r}).$$

Since $|f_X(t) - f_X(\omega)| \leq C_1|t-\omega| \leq C_1\lambda_N$ for some constant C_1, $|f_X(t) - f_X(\omega)| = O(\delta_N)$. Therefore it is enough to show that

$$P(\max_{t\in C_N}\sup_{\omega\in Q_t} | \hat{f}_X(\omega) - f_X(t)| \geq CN^{r-1}(\log N)^{-r}) \to 0 \quad as \quad N\to\infty.$$

Set $K = [m + N\lambda_N]$. Define for each $t \in C_N$,

$$\bar{f}_X(t) = \frac{1}{2\pi}\sum_{|h|\leq K} W_N(h) I_N(g(N,t) + \omega_h).$$

It is sufficient to show that there is a positive constant C such that

$$P\Big(\max_{t\in C_N}\sup_{\omega\in Q_t} | \hat{f}_X(\omega) - \bar{f}_X(t)| \geq CN^{r-1}(\log N)^{-r}\Big) \to 0 \quad as \quad N\to\infty \tag{4.3}$$

and

$$P\Big(\max_{t\in C_N} | \bar{f}_X(t) - f_X(t)| \geq CN^{r-1}(\log N)^{-r}\Big) \to 0 \quad as \quad N\to\infty. \tag{4.4}$$

Note that

$$| \hat{f}_X(\omega) - \bar{f}_X(t)| \leq \left|\frac{1}{2\pi}\sum_{|h|\leq m} W_N(h)\, I_N(g(N,\omega) + \omega_h) - I_N(g(N,t) + \omega_h)\right|$$
$$+ \left|\frac{1}{2\pi}\sum_{h=m+1}^{K} W_N(h) I_N(g(N,t) + \omega_h)\right| + \left|\frac{1}{2\pi}\sum_{h=-K}^{-m-1} W_N(h) I_N(g(N,t) + \omega_h)\right|$$
$$= (I) + (II) + (III).$$

By the mean value theorem,

$$|I_N(g(N,\omega) + \omega_h) - I_N(g(N,t) + \omega_h)| = |I_N'(\xi_h)||g(N,\omega) - g(N,t)|$$

for a proper ξ_h which lies between $g(N,\omega)+\omega_h$ and $g(N,t)+\omega_h$. Hence

$$(I) = \frac{1}{2\pi}\sum_{|h|\leq m} | W_N(h)||I_N'(\xi_h)||g(N,\omega) - g(N,t)|$$
$$\leq \frac{1}{2\pi}\max_{|h|\leq m} |I_N'(\xi_h)||g(N,\omega) - g(N,t)| \sum_{|h|\leq m} | W_N(h)|$$
$$\leq M_o\, |g(N,\omega) - g(N,t)|$$

where $M_o = \frac{1}{2\pi}\max_{|h|\leq m}|I_N'(\xi_h)| \ \sup_N(\sum_{|h|\leq m}|W_N(h)|)$.

Since $|g(N,\omega)-g(N,t)|=O(N^{r-1}(\log N)^{-r})+O(N^{-1})$, (I) is equivalent to $O(N^{r-1}(\log N)^{-r})$. By similar manner, (II) and (III) are equivalent to $O(N^{r-1}(\log N)^{-r})$. Hence (4.3) is valid. Observe that

$$\begin{aligned}\bar{f}_X(t) - f_X(t) &= \sum_{|h|\le K} W_N(h)\left\{\frac{1}{2\pi} I_N(g(N,t)+\omega_h) - f_X(g(N,t)+\omega_h)\right\} \\ &\quad + \sum_{|h|\le K} W_N(h)\, f_X(g(N,t)+\omega_h) - f_X(t) \\ &\quad + \{\sum_{|h|\le K} W_N(h) - 1\} f_X(t).\end{aligned}$$

By the condition (i), we have

$$\begin{aligned}\max_{t\in C_N}|\sum_{|h|\le K} W_N(h) - 1||f_X(t)| &= |\sum_{|h|\le K} W_N(h) - 1| \max_{t\in C_N} f_X(t) \\ &= O(N^{r-1}(\log N)^{-r}).\end{aligned}$$

By the condition (i) and continuity of $f_X(\omega)$, we obtain

$$\begin{aligned}\max_{t\in C_N}\sum_{|h|\le K}|W_N(h)||f_X(g(N,t)+\omega_h) - f_X(t)| &\le B_0 C_1 \max_{t\in C_N}|g(N,t)+\omega_h - t| \\ &= O(N^{r-1}(\log N)^{-r})\end{aligned}$$

where $B_0=\sup_N(\sum_{|h|\le K}|W_N(h)|)$.

Let $\#(A)$ denote the number of elements in a set A. We know that there is a positive constant x such that $\#(C_N)\le N^x$. Now

$$\begin{aligned}&P\left(\max_{t\in C_N}\left|\sum_{|h|\le K} W_N(h)\left\{\frac{1}{2\pi} I_N(g(N,t)+\omega_h) - f_X(g(N,t)+\omega_h)\right\}\right| \ge CN^{r-1}(\log N)^{-r}\right) \\ &\le P\left(\bigcup_{t\in C_N}|\sum_{|h|\le K} W_N(h)\, U_h f_X(g(N,t)+\omega_h)| \ge CN^{r-1}(\log N)^{-r}\right) \\ &\le \#(C_N)\, P\left(\max_{|h|\le K} f_X(g(N,t)+\omega_h)\sum_{|h|\le K}|W_N(h)\, U_h| \ge CN^{r-1}(\log N)^{-r}\right) \\ &\le N^x P\left(\sum_{|h|\le K}|W_N(h)\, U_h| \ge \frac{CN^{r-1}(\log N)^{-r}}{\max_{|h|\le K} f_X(g(N,t)+\omega_h)}\right) \\ &\le N^x E\,(\sum_{|h|\le K}|W_N(h)\, U_h|)^2 \Big/ \left(\frac{CN^{r-1}(\log N)^{-r}}{\max_{|h|\le K} f_X(g(N,t)+\omega_h)}\right)^2 \\ &\le \frac{N^{x+2-2r-s}O(1)}{C^2(\log N)^{-2r}} \\ &= o(1)\end{aligned}$$

for sufficiently large N, where the third inequality is obtained by using Markov's inequality and the fourth by applying the property of the sequence $\{U_h\}$ defined in the proof of Theorem 2.1.

References

[1] Brockwell, P. J. and Davis, R. A. (1991). *Time Series : Theory and Methods*, Springer-Verlag, New York.

[2] Rosenblatt, M. (1984). Asymptotic normality, strong mixing and spectral density estimates, *Annals of Probability*, Vol. 4 , 1167-1180.

[3] Shin, G. H. (1994). Asymptotic convergence properties in spectral estimation, Ph.D. Thesis, Yonsei Univ.

[4] Stone, C. J. (1982). Optimal global rates of convergence for nonparametric regression, *Annals of Statistics*, Vol. 10, 1040-1053.

[5] Truong, Y. K. (1991). Nonparametric curve estimation with time series errors, *Journal of Statistical Planning and Inference*, Vol. 28, 167-183.