ELSEVIER

Signal Processing 81 (2001) 419–430

SIGNAL PROCESSING

www.elsevier.nl/locate/sigpro

A stabilized bandwidth selection method for kernel smoothing of the periodogram

Thomas C.M. Lee*

Department of Statistics, Colorado State University, Room 220, Statistics Building, Fort Collins, CO 80523-1877, USA

Received 23 March 2000; received in revised form 27 July 2000

Abstract

One popular method for nonparametric spectral density estimation is to perform kernel smoothing on the periodogram, and one important component of this method is the choice of the bandwidth (or span) for smoothing. This paper proposes a new bandwidth selection method that is based on a coupling of the so-called plug-in and the unbiased risk estimation ideas. This new method is easy to describe, simple to implement, and does not impose severe conditions on the unknown spectrum. Numerical results suggest that this new method often outperforms some other commonly used bandwidth selection methods. The new methodology is also applied to choose the bandwidth for log-periodogram smoothing. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Bandwidth selection; Log-periodogram and periodogram smoothing; Plug in; Spectral density estimation; Unbiased risk estimation

1. Introduction

This paper considers the problem of estimating the spectral density by nonparametrically smoothing the periodogram or the log-periodogram. Many approaches have been proposed. These include spline smoothing (e.g. [10,20]), kernel smoothing (or weighted local averaging) (e.g. [3,11,15,19]) and wavelet techniques (e.g. [4,13,21]). The approach that this paper is concerned with is kernel smoothing. Some appealing features of this approach are that it is simple to use, easy to understand and straightforward to interpret.

One important component of the kernel smoothing approach is the choice of the bandwidth (or span) for smoothing. The goal of this paper is to introduce a simple but effective technique for choosing the bandwidth. This technique is based on a coupling of the plug-in and the unbiased risk estimation ideas that are commonly found in the nonparametric probability density and regression estimation literature. This technique is termed PURE [12]. Our simulation results show that this PURE choice of bandwidth is superior to cross-validation and the method proposed in [11].

The rest of this paper is organized as follows. Section 2 provides some background material on the periodogram and the unbiased risk estimation technique. In Section 3 the new PURE choice of bandwidth for smoothing the periodogram is introduced, while Section 4 considers the smoothing of

* Tel.: + 1-970-491-2185; fax: + 1-970-491-7895.
E-mail address: tlee@stat.colostate.edu (T.C.M. Lee).

PII: S0165-1684(00)00218-8

the log-periodogram. Simulation results are reported in Section 5 and this paper ends with a concluding section.

2. Background

2.1. Periodogram smoothing

Suppose that $\{x_t\}$ is a real-valued, zero mean stationary process with unknown spectral density f, and that a finite-sized realization $x_0, \ldots, x_{2n-1}$ of $\{x_t\}$ is observed. The goal is to estimate f by using those observed x_t's. The periodogram is defined as

$$I(\omega) = \frac{1}{2\pi \times 2n} \left| \sum_{t=0}^{2n-1} x_t \exp(-\mathrm{i}\omega t) \right|^2,$$

$$\mathrm{i} = \sqrt{-1},\ \omega \in [0, 2\pi).$$

To simplify notation, write $\omega_j = 2\pi j/(2n)$. Since the spectral density f is symmetric about $\omega = \pi$, we shall focus our discussion on $f(\omega_j)$ for $j = 0, \ldots, n-1$. Also, as f is periodic with period 2π, we have $f(\omega_{-j}) = f(\omega_j)$ and $I(\omega_{-j}) = I(\omega_j)$ for $j = 1, \ldots, n-1$.

A frequently adopted model for $I(\omega_j)$ is (e.g. see [3,11,13,14])

$$I(\omega_j) = f(\omega_j)\varepsilon_j, \quad j = 0, \ldots, n-1, \tag{1}$$

where the ε_j's are independent standard exponential random variables. Thus $E\{I(\omega_j)\} = f(\omega_j)$ and $\mathrm{Var}\{I(\omega_j)\} = f(\omega_j)^2$. Due to its unacceptably large variance, $I(\omega_j)$ is seldom used as an estimate of $f(\omega_j)$.

One possible way for obtaining better estimates for $f(\omega_j)$ is to smooth the $I(\omega_j)$'s. Here we consider the following kernel estimator for $f(\omega_j)$:

$$\hat{f}_h(\omega_j) = \sum_{k=-n}^{2n-1} K_h(\omega_k - \omega_j) I(\omega_k) \Bigg/ \sum_{l=-n}^{2n-1} K_h(\omega_l - \omega_j),$$

$$j = 0, \ldots, n-1. \tag{2}$$

In the above $K_h(\cdot) = (1/h)K(\cdot/h)$, where the kernel function K is (usually taken as) a symmetric probability density function and the bandwidth h is a nonnegative smoothing parameter that controls the amount of smoothing. It is well known that the choice of h is much more crucial than the choice of K (e.g. see [16] or [23]). Also, in most other kernel smoothing problems the limits of the two summations in (2) are 0 and $n-1$. However, since in the present setting boundary effects can be handled by periodic smoothing, the limits are changed from 0 and $n-1$ to $-n$ and $2n-1$, respectively.

The estimator $\hat{f}_h(\omega_j)$ can also be interpreted as a weighted average of the $I(\omega_j)$'s. It is because one could write

$$\hat{f}_h(\omega_j) = \sum_{k=-n}^{2n-1} W_{h,k}(j) I(\omega_k)$$

$$\text{with } W_{h,k}(j) = \frac{K_h(\omega_k - \omega_j)}{\sum_{l=-n}^{2n-1} K_h(\omega_l - \omega_j)}. \tag{3}$$

Notice that the weights $W_{h,k}(j)$'s sum to unity.

2.2. Unbiased risk estimation

As mentioned before, the choice of h is crucial. One such reasonable choice is to choose it as the minimizer of the risk function $R(h)$:

$$R(h) = E\left[\sum_{j=0}^{n-1} \{f(\omega_j) - \hat{f}_h(\omega_j)\}^2 \right].$$

Of course, in practice, this idealized choice of h cannot be obtained, as we do not know $f(\omega_j)$ and hence $R(h)$. In [11] an unbiased estimator $\hat{R}_{\text{SURE}}(h)$ of $R(h)$ is constructed using the so-called Stein's unbiased risk estimation (SURE) technique (e.g. see [17,18]). For the current formulation

$$\hat{R}_{\text{SURE}}(h) = \sum_{j=0}^{n-1} \{I(\omega_j) - \hat{f}_h(\omega_j)\}^2 - \frac{1 - 2W_{h,0}(0)}{2} \sum_{j=0}^{n-1} I(\omega_j)^2.$$

It is also suggested in [11] that h could be chosen as the minimizer of $\hat{R}_{\text{SURE}}(h)$. Good practical performances of this unbiased risk estimation approach for choosing h have been reported in [11,19].

However, in the nonparametric probability density (e.g. [2]) and regression estimation contexts (e.g. [1,7]), it has been shown that the variances of the risk estimators obtained by the SURE approach can sometimes be high. As a consequence, poor-quality bandwidths can sometimes be chosen. This phenomenon is expected to carry over to the context of spectral density estimation, and thus there is the need for seeking better (or more stable) estimates for $R(h)$; see the next section.

For comparative purposes, we also write down the following criterion derived using the method cross-validation:

$$\hat{R}_{\mathrm{CV}}(h) = \sum_{j=0}^{n-1} \{I(\omega_j) - \hat{f}_h(\omega_{-j})\}^2$$
$$= \sum_{j=0}^{n-1} \{I(\omega_j) - \hat{f}_h(\omega_j)\}^2 / \{1 - W_{h,0}(0)\}^2$$

(e.g. see [9,11] and references given therein). In the above $\hat{f}_h(\omega_{-j})$ is the usual "leave-one-out" estimate of $\hat{f}_h(\omega_j)$. It is straightforward to show that, as an estimator of $R(h)$, $\hat{R}_{\mathrm{CV}}(h)$ is biased (but the bias goes away when $n \to \infty$).

3. Stabilized risk estimation

This section presents the main contribution of this paper, namely, a new estimator for $R(h)$. This new estimator is relatively more stable than $\hat{R}_{\mathrm{SURE}}(h)$.

3.1. Derivation of the new risk estimator

First, we begin with $E[\{f(\omega_j) - \hat{f}_h(\omega_j)\}^2]$. Using the identity $E(X^2) = \{E(X)\}^2 + \mathrm{Var}(X)$ for any random variable X, we have

$$E[\{f(\omega_j) - \hat{f}_h(\omega_j)\}^2]$$
$$= [E\{f(\omega_j) - \hat{f}_h(\omega_j)\}]^2 + \mathrm{Var}\{f(\omega_j) - \hat{f}_h(\omega_j)\}$$
$$= [f(\omega_j) - E\{\hat{f}_h(\omega_j)\}]^2 + \mathrm{Var}\{\hat{f}_h(\omega_j)\}$$
$$= [f(\omega_j) - E\{\hat{f}_h(\omega_j)\}]^2 + \mathrm{Var}\left\{\sum_k W_{h,k}(j) I(\omega_j)\right\}$$
$$= [f(\omega_j) - E\{\hat{f}_h(\omega_j)\}]^2 + \sum_k W_{h,k}^2(j)\,\mathrm{Var}\{I(\omega_j)\}$$
$$= [f(\omega_j) - E\{\hat{f}_h(\omega_j)\}]^2 + \sum_k W_{h,k}^2(j) f^2(\omega_j)$$

and hence

$$R(h) = E\left[\sum_{j=0}^{n-1} \{f(\omega_j) - \hat{f}_h(\omega_j)\}^2\right]$$
$$= \sum_{j=0}^{n-1} [f(\omega_j) - E\{\hat{f}_h(\omega_j)\}]^2$$
$$+ \sum_{j=0}^{n-1} \sum_k W_{h,k}^2(j) f^2(\omega_j).$$

Therefore, if f is known, $R(h)$ for a given h can be estimated by

$$\sum_{j=0}^{n-1} \{f(\omega_j) - \hat{f}_h(\omega_j)\}^2 + \sum_{j=0}^{n-1} \sum_k W_{h,k}^2(j) f^2(\omega_j). \tag{4}$$

Certainly the above expression is of no practical use, as we do not know f. One way to construct an estimator of $R(h)$ using this expression is first to obtain a *pilot* estimate $\hat{f}_{h_{\mathrm{P}}}$ (with a *pilot* bandwidth h_{P}; see below) and then plug-in this pilot estimate into (4):

$$\hat{R}_{\mathrm{PURE}}(h) = \sum_{j=0}^{n-1} \{\hat{f}_{h_{\mathrm{P}}}(\omega_j) - \hat{f}_h(\omega_j)\}^2$$
$$+ \sum_{j=0}^{n-1} \sum_k W_{h,k}^2(j) \hat{f}_{h_{\mathrm{P}}}^2(\omega_j). \tag{5}$$

Here we propose to choose h as the minimizer of $\hat{R}_{\mathrm{PURE}}(h)$. Notice that $\hat{R}_{\mathrm{PURE}}(h)$ is constructed by using both the plug-in and the unbiased risk estimation ideas, and hence the name PURE. Also notice that no "higher-order" quantities like f'' are required in using $\hat{R}_{\mathrm{PURE}}(h)$.

3.2. Some remarks

By comparing the expressions for $\hat{R}_{\mathrm{SURE}}(h)$ and $\hat{R}_{\mathrm{PURE}}(h)$, perhaps one can gain some insights about why $\hat{R}_{\mathrm{PURE}}(h)$ is a more stable estimator. In the expression for $\hat{R}_{\mathrm{PURE}}(h)$, those high-variance $I(\omega_j)$'s are, in a way, replaced by the pilot estimates $\hat{f}_{h_{\mathrm{P}}}(\omega_j)$. These pilot estimates are themselves smoothed versions of $I(\omega_j)$'s, and hence should have lower variances. In other words, the first step of "pre-smoothing" the $I(\omega_j)$'s introduces a stabilizing effect to the estimation, but perhaps at a small expense of increasing the bias. This is because $\hat{f}_{h_P}(\omega_j)$ is generally a slightly biased estimator of $f(\omega_j)$ and this would make $\hat{R}_{\text{PURE}}(h)$ biased for $R(h)$; see (4) and (5).

There are two additional attractive features of the proposed PURE bandwidth selection procedure: it is easy to describe and simple to implement. It only requires the selections of two bandwidths, and these selections can be easily performed via solving two simple minimization problems. Also, this PURE idea to risk estimation has been, sometimes under different names, applied with great success to tackle different nonparametric probability density and regression estimation problems [12]. Those different names that have been used include stabilized selectors [1,2], smoothed cross-validation [6], double smoothing [8] and exact risk approach [22].

Now for the choice of the pilot bandwidth h_P. In other similar problems it has been shown that the choice for the pilot bandwidth is not a crucial issue (e.g., see the references cited in the previous paragraph). Therefore, for simplicity, we suggest choosing h_P as the minimizer of $\hat{R}_{\text{SURE}}(h)$. This simple choice for h_P performed very well in our simulations.

As a summary, our PURE-based spectral density estimate $\hat{f}_{h_F}$ can be obtained by the following steps:

1. choose the *pilot* bandwidth h_P as the minimizer of $\hat{R}_{\text{SURE}}(h)$,
2. use (3) to compute $\hat{f}_{h_P}(\omega_j)$, $j = 1, \ldots, n-1$, with h_P as the bandwidth,
3. substitute the computed $\hat{f}_{h_P}(\omega_j)$'s into expression (5) for $\hat{R}_{\text{PURE}}(h)$,
4. choose the *final* bandwidth h_F as the minimizer of $\hat{R}_{\text{PURE}}(h)$,
5. use (3) to compute the final spectral density estimate $\hat{f}_{h_F}(\omega_j)$, $j = 1, \ldots, n-1$, with h_F as the bandwidth.

4. Log-periodogram smoothing

The PURE methodology can also be applied to choose the bandwidth for log-periodogram smoothing. In this case the natural risk function that one would like to minimize is

$$R'(h) = E\left[\sum_{j=0}^{n-1} \{\log f(\omega_j) - \log \hat{f}_h(\omega_j)\}^2\right].$$

The first step is to transform the multiplicative model (1) into an additive model by taking a logarithmic transform:

$$y(\omega_j) = \log I(\omega_j) + \gamma = \log f(\omega_j) + \xi_j,$$
$$j = 0, \ldots, n-1,$$

where the ξ_j's are independent zero mean random variables with variance $\pi^2/6$ and $\gamma = 0.57721$ is the Euler's constant.

Let $\hat{g}_h(\omega_j) = \sum_{k=-n}^{2n-1} W_{h,k}(j) y(\omega_k)$ be the kernel estimate of $\log f(\omega_j)$. Using the same technique as before, one can show that

$$\hat{R}'_{\text{PURE}}(h) = \sum_{j=0}^{n-1} \{\hat{g}_{h_P}(\omega_j) - \hat{g}_h(\omega_j)\}^2 + \frac{n\pi^2}{6} \sum_k W_{h,k}^2(0)$$

is a PURE-based estimator of $R'(h)$. In the above h_P is a pilot bandwidth, which we recommend choosing as the minimizer of the SURE-based estimator of $R'(h)$ proposed by [11]:

$$\hat{R}'_{\text{SURE}}(h) = \sum_{j=0}^{n-1} \{y(\omega_j) - \hat{g}_h(\omega_j)\}^2 - \frac{n\pi^2}{6}\{1 - 2W_{h,0}(0)\}.$$

For completeness we also list the following cross-validation criterion for log-periodogram smoothing:

$$\hat{R}'_{\text{CV}}(h) = \sum_{j=0}^{n-1} \{y(\omega_j) - \hat{g}_h(\omega_j)\}^2 / \{1 - W_{h,0}(0)\}^2.$$

5. Simulation

This section reports the results of two simulation experiments that we have conducted.

5.1. Periodogram smoothing

The first experiment was concerned with the smoothing of the periodogram. Four test examples and three different sample sizes were used. The three sample sizes were $n = 256$, 512 and 1024, and the four test examples were from the ARMA(α,β) model

$$x_t + a_1 x_{t-1} + \cdots + a_\alpha x_{t-\alpha} = \tau_t + b_1 \tau_{t-1} + \cdots + b_\beta \tau_{t-\beta}, \quad \tau_t \sim \mathrm{N}(0,1)$$

with parameters given by

Example 1. AR(3) with $a_1 = -1.5$, $a_2 = 0.7$ and $a_3 = -0.1$.

Example 2. AR(3) with $a_1 = 0.9$, $a_2 = 0.8$ and $a_3 = 0.6$.

Example 3. MA(3) with $b_1 = 0.9$, $b_2 = 0.8$ and $b_3 = 0.6$.

Example 4. MA(4) with $b_1 = -0.3$, $b_2 = -0.6$, $b_3 = -0.3$ and $b_4 = 0.6$.

These testing examples have previously been used by many authors; e.g., see [3,9,11,14,20]. The kernel function used was $K(x) = \frac{3}{4}(1 - x^2)$, $x \in [0,1]$. It is the optimal kernel of order (0,2) derived in [5]. It is also known as the Epanechnikov kernel, and it is optimal in the sense that it minimizes the risk $R(h)$ for a given h as $n \to \infty$.

For each combination of test example and sample size (totally there are $4 \times 3 = 12$ such combinations), 500 independent series were simulated, and the corresponding periodograms were also computed. For each of these generated periodograms the following four bandwidths were obtained:

h_{OPT}: the optimal bandwidth that minimizes $R(h)$ (in practice h_{OPT} is unobtainable);
h_{PURE}: the bandwidth chosen by the PURE procedure (i.e., minimizer of $\hat{R}_{\mathrm{PURE}}(h)$);
h_{SURE}: the bandwidth chosen by the SURE procedure (i.e., minimizer of $\hat{R}_{\mathrm{SURE}}(h)$);
h_{CV}: the bandwidth chosen by cross-validation (i.e., minimizer of $\hat{R}_{\mathrm{CV}}(h)$);

In addition, three risk ratios were computed: $r_{\mathrm{PURE}} = R(h_{\mathrm{PURE}})/R(h_{\mathrm{OPT}})$, $r_{\mathrm{SURE}} = R(h_{\mathrm{SURE}})/R(h_{\mathrm{OPT}})$ and $r_{\mathrm{CV}} = R(h_{\mathrm{CV}})/R(h_{\mathrm{OPT}})$. These risk ratios are used for comparing the relative merits of the bandwidth selection methods: the smaller the risk ratio, the better the method.

The average values of the risk-ratio differences $r_{\mathrm{SURE}} - r_{\mathrm{PURE}}$, $r_{\mathrm{CV}} - r_{\mathrm{PURE}}$ and $r_{\mathrm{CV}} - r_{\mathrm{SURE}}$ are listed in Table 1. Paired t-tests were conducted to test if the differences are statistically significant at 5%

Table 1
Average values of various risk-ratio differences for the periodogram smoothing experiment. Numbers in parenthesis are estimated standard errors

Example	n	$r_{\mathrm{SURE}} - r_{\mathrm{PURE}}$	$r_{\mathrm{CV}} - r_{\mathrm{PURE}}$	$r_{\mathrm{CV}} - r_{\mathrm{SURE}}$
1	256	2.923 (0.631)[a]	3.08 (0.776)[a]	0.157 (0.946)
2	256	0.028 (0.025)	0.306 (0.08)[a]	0.278 (0.075)[a]
3	256	2.445 (0.789)[a]	2.180 (1.087)[a]	−0.264 (1.696)
4	256	0.165 (0.045)[a]	0.526 (0.129)[a]	0.361 (0.121)[a]
1	512	2.385 (0.875)[a]	3.031 (1.345)[a]	0.646 (1.971)
2	512	0.086 (0.03)[a]	0.368 (0.072)[a]	0.282 (0.059)[a]
3	512	1.001 (0.207)[a]	2.159 (0.569)[a]	1.158 (0.633)
4	512	0.205 (0.044)[a]	0.774 (0.216)[a]	0.568 (0.204)[a]
1	1024	2.060 (0.459)[a]	4.148 (1.328)[a]	2.088 (1.308)
2	1024	0.229 (0.052)[a]	0.273 (0.099)[a]	0.044 (0.109)
3	1024	1.779 (0.601)[a]	1.476 (0.597)[a]	−0.303 (0.982)
4	1024	0.178 (0.034)[a]	0.476 (0.137)[a]	0.298 (0.139)[a]

[a] Indicate that the difference is significant at 5% level.

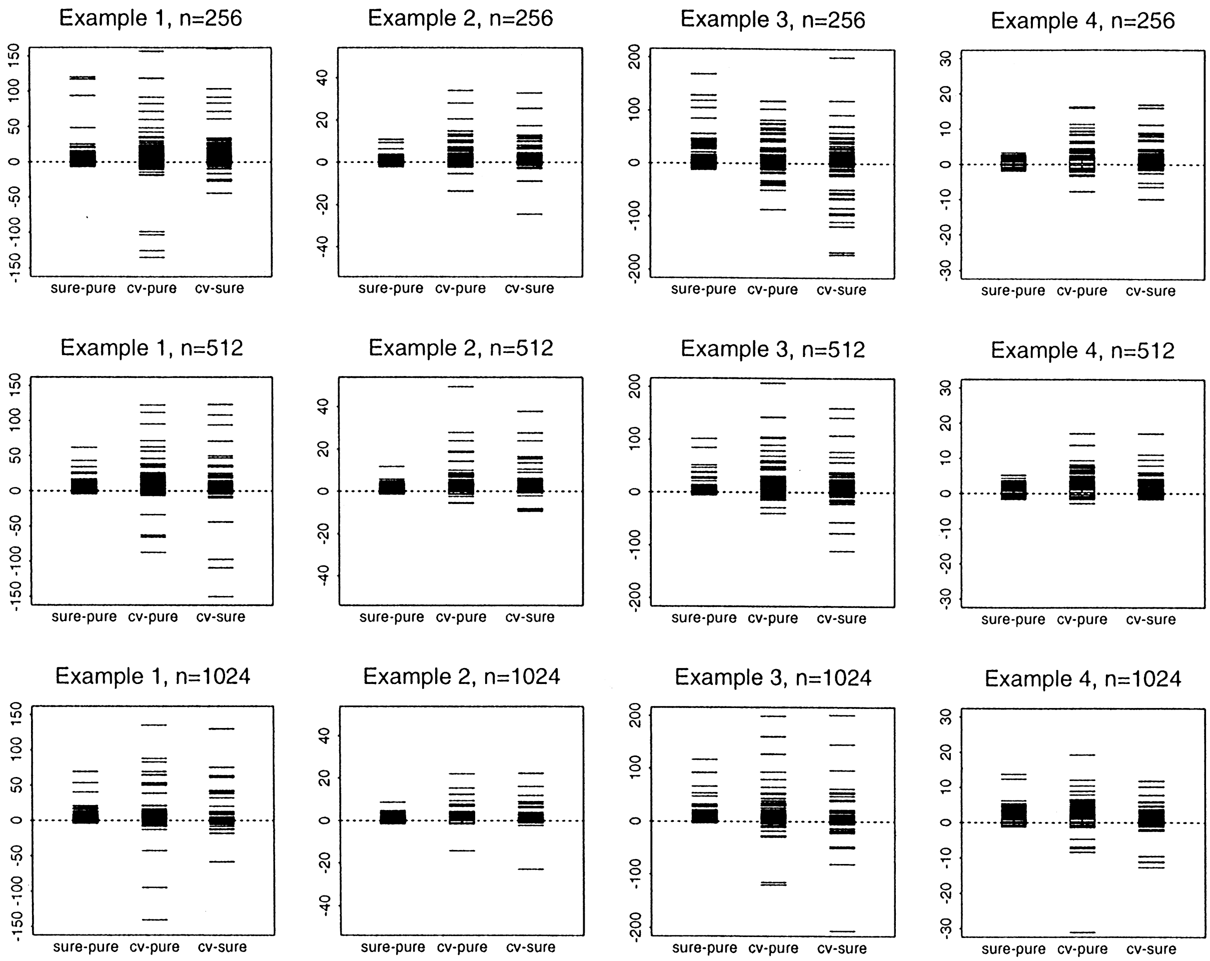

Fig. 1. Boxplots of the three risk-ratio differences $r_{\mathrm{SURE}} - r_{\mathrm{PURE}}$, $r_{\mathrm{CV}} - r_{\mathrm{PURE}}$ and $r_{\mathrm{CV}} - r_{\mathrm{SURE}}$ for the periodogram smoothing experiment.

significance level. In addition, boxplots of these three risk-ratio differences are given in Fig. 1, while plots of different spectrum estimates are provided in Figs. 2–5 for the purpose of visually evaluating the different bandwidth selection methods.

From Table 1 and Figs. 1–5 one can see that the PURE method outperformed the other two methods in all cases except for Example 2 with $n = 256$, where PURE and SURE gave similar performance. Also, in most cases the PURE spectrum estimates contain less wiggles than others. Another observation is that SURE is slightly superior to CV.

5.2. Log-periodogram smoothing

In order to compare the relative merits of the PURE, SURE and cross-validation bandwidth selection methods described in Section 4 for log-periodogram smoothing, a similar numerical experiment was performed. The set-up was essentially the same as in the previous section, but the target risk function was $R'(h)$, and the risk ratios are defined in terms of $R'(h)$. Results are reported, in the same format as before, in Table 2 and Figs. 6–10. From these results, one can see that PURE also, but to a lesser extent, outperforms SURE in the context of log-periodogram smoothing.

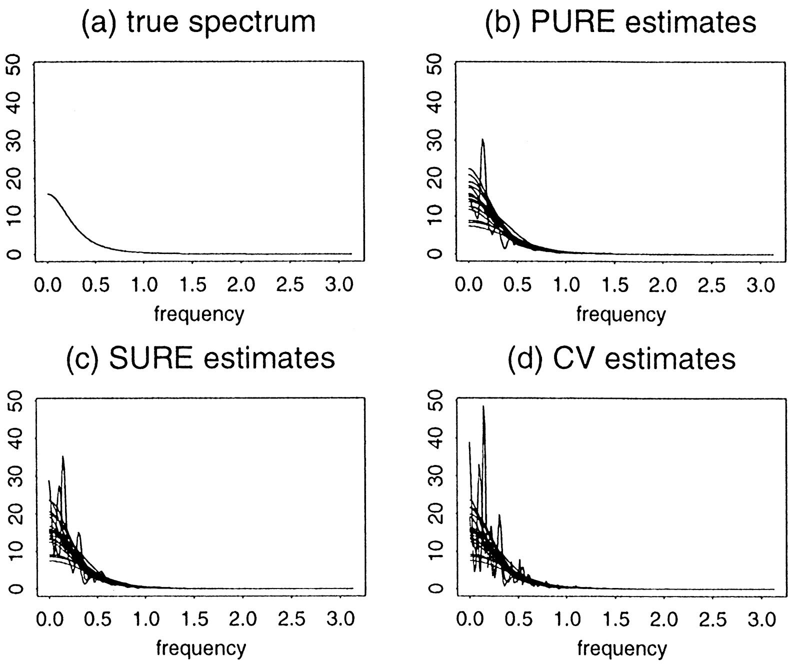

Fig. 2. Smoothing of periodogram, Example 1 with $n = 512$: plots of the (a) spectrum and twenty estimates obtained by the (b) PURE, (c) SURE, and the (d) CV methods.

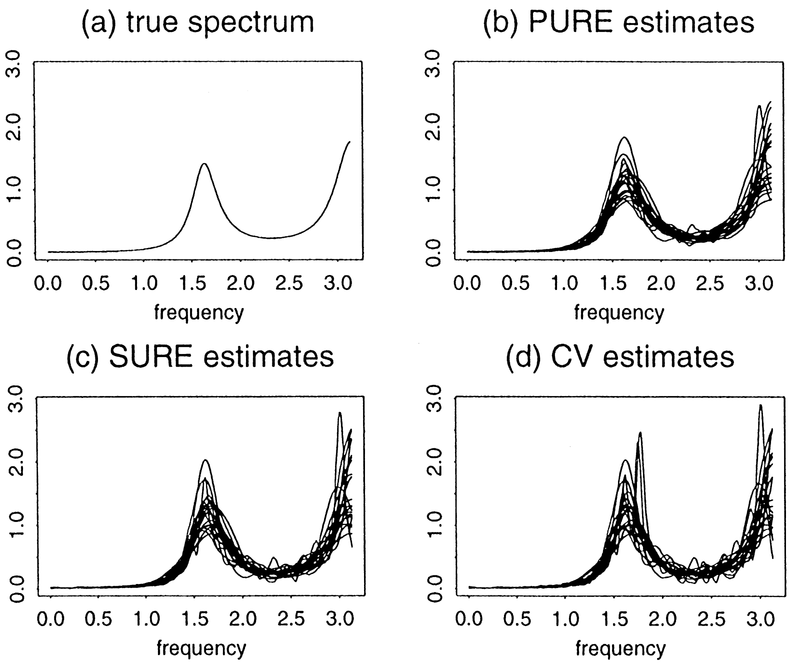

Fig. 3. Similar to Fig. 2 but for Example 2.

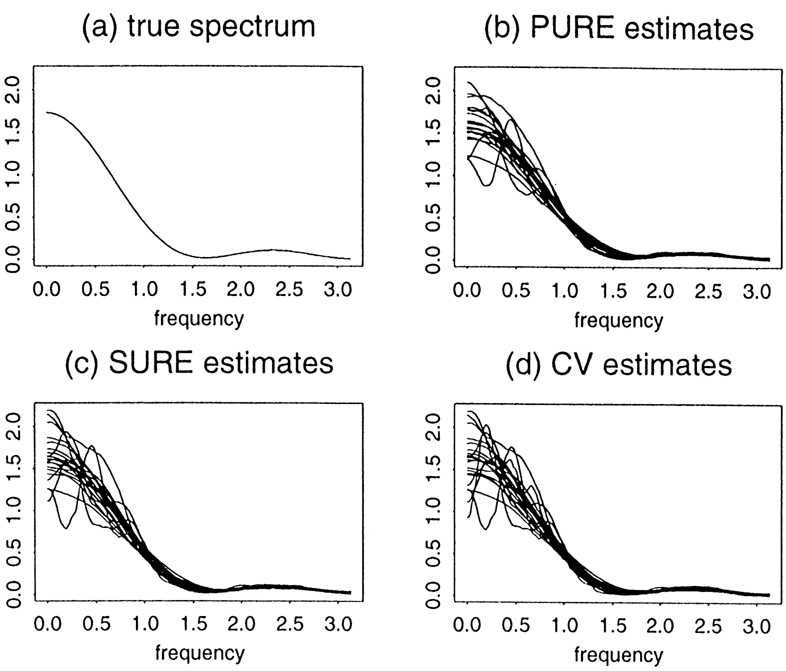

Fig. 4. Similar to Fig. 2 but for Example 3.

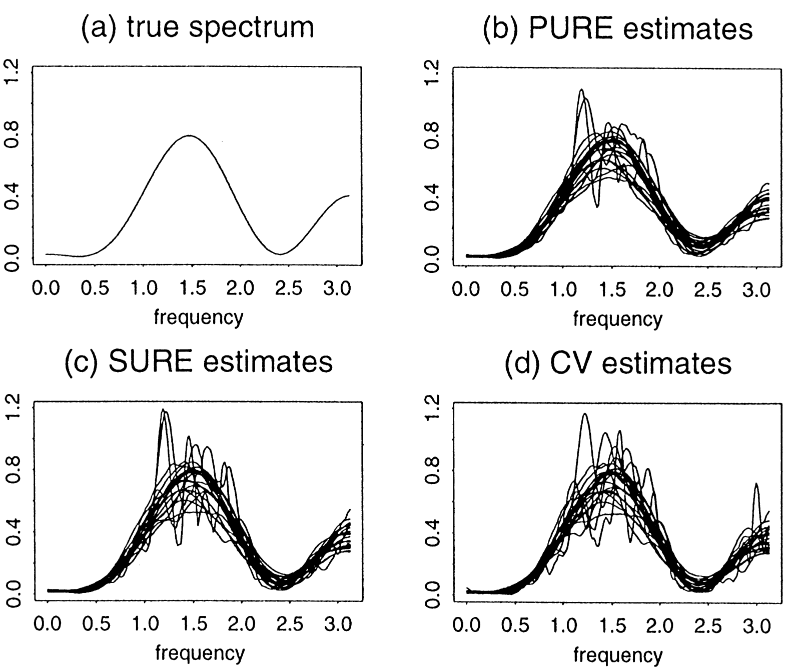

Fig. 5. Similar to Fig. 2 but for Example 4.

Table 2
Similar to Table 1 but for log-periodogram smoothing

Example	n	$r_{\text{SURE}} - r_{\text{PURE}}$	$r_{\text{CV}} - r_{\text{PURE}}$	$r_{\text{CV}} - r_{\text{SURE}}$
1	256	0.107 (0.034)[a]	– 0.042 (0.069)	– 0.149 (0.089)
2	256	0.018 (0.014)	– 0.011 (0.022)	– 0.028 (0.022)
3	256	0.048 (0.019)[a]	0.020 (0.044)	– 0.028 (0.047)
4	256	0.020 (0.019)	– 0.050 (0.029)	– 0.070 (0.038)
1	512	0.070 (0.021)[a]	0.048 (0.047)	– 0.022 (0.052)
2	512	0.008 (0.013)	0.024 (0.026)	0.015 (0.023)
3	512	0.004 (0.014)	0.002 (0.023)	– 0.002 (0.024)
4	512	0.045 (0.016)[a]	0.008 (0.033)	– 0.038 (0.04)
1	1024	0.053 (0.017)[a]	0.081 (0.025)[a]	0.028 (0.023)
2	1024	0.002 (0.011)	0.009 (0.015)	0.006 (0.011)
3	1024	0.007 (0.013)	0.001 (0.018)	– 0.007 (0.017)
4	1024	0.011 (0.014)	– 0.023 (0.018)	– 0.034 (0.02)

[a]Indicate that the difference is significant at 5% level.

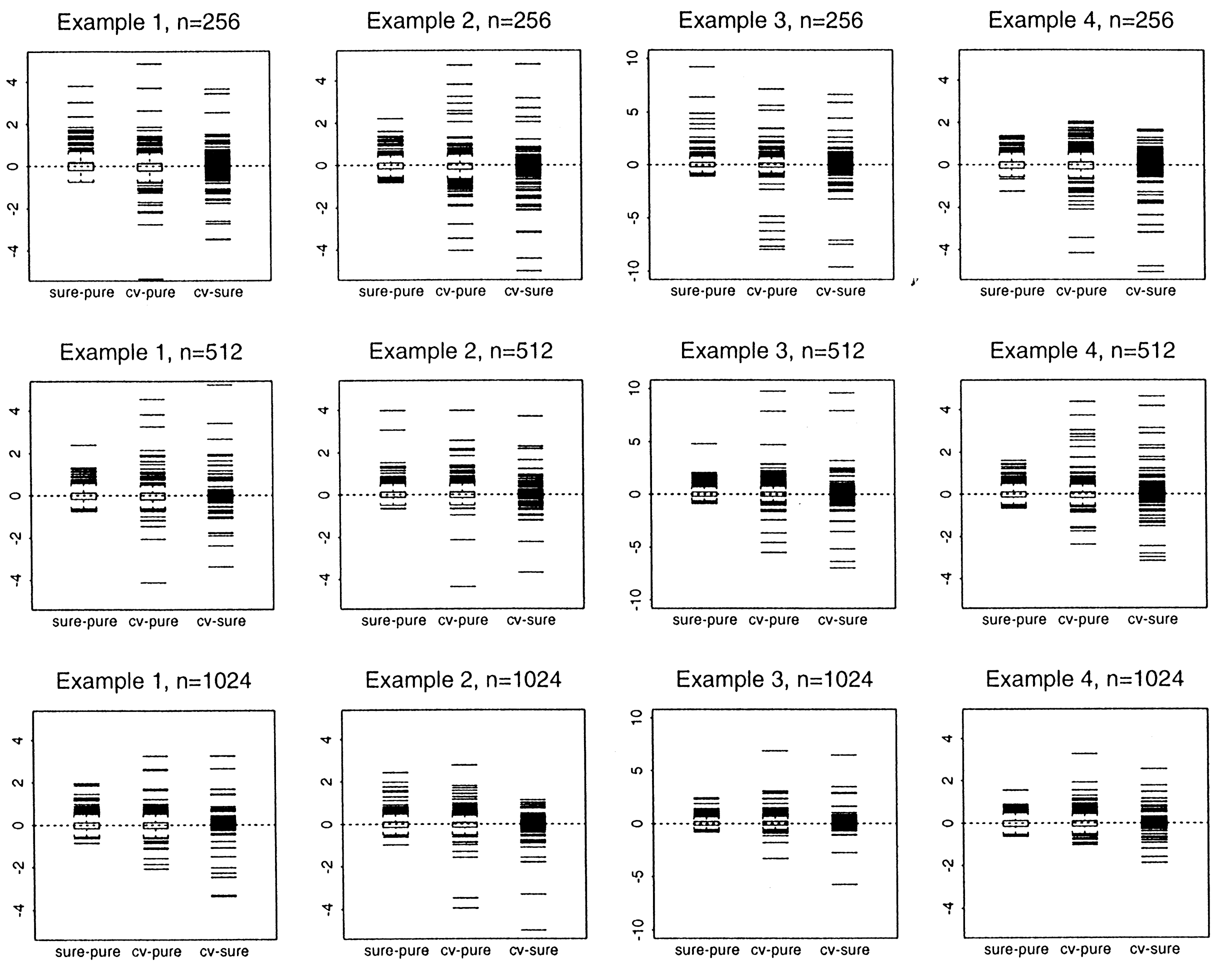

Fig. 6. Similar to Fig. 1 but for log-periodogram smoothing.

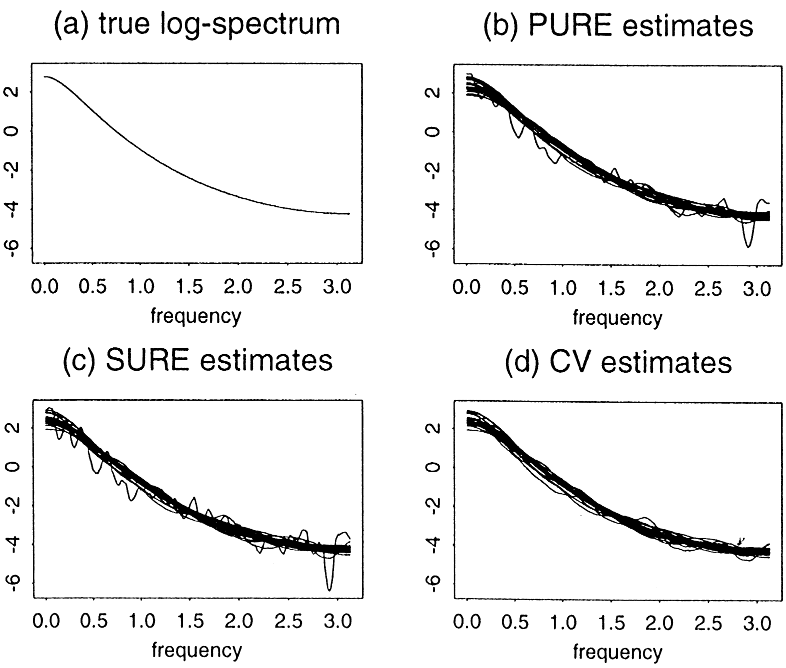

Fig. 7. Similar to Fig. 2 but for log-periodogram smoothing.

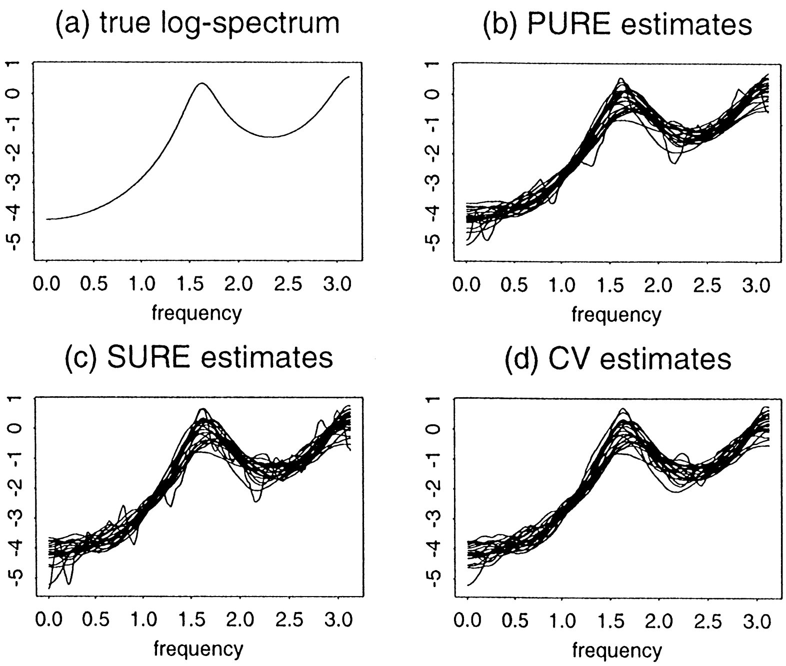

Fig. 8. Similar to Fig. 3 but for log-periodogram smoothing.

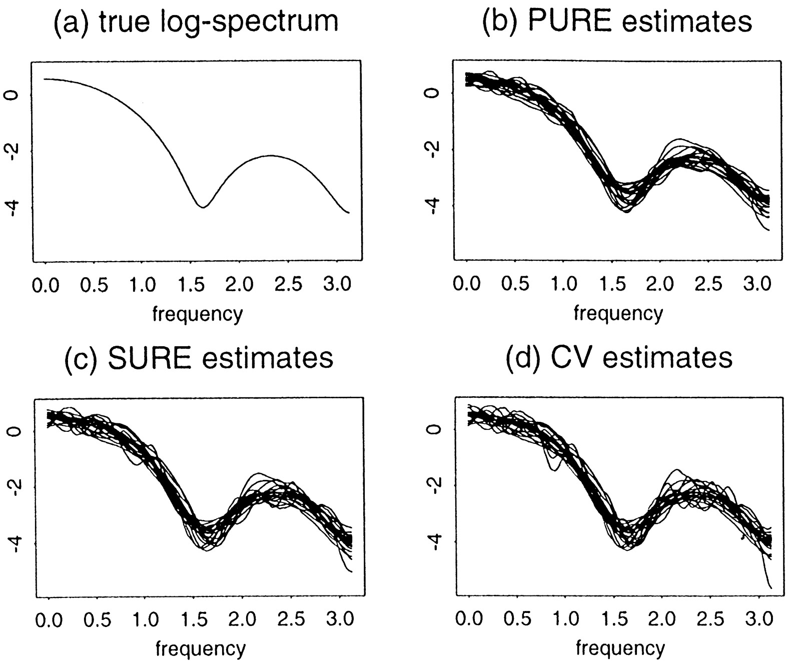

Fig. 9. Similar to Fig. 4 but for log-periodogram smoothing.

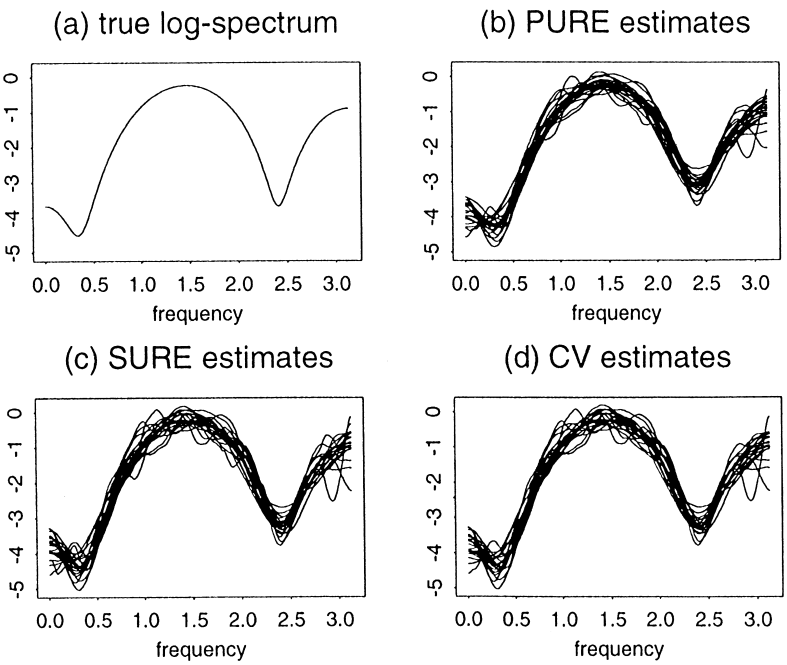

Fig. 10. Similar to Fig. 5 but for log-periodogram smoothing.

6. Concluding remarks

In this paper a new bandwidth selection methodology for periodogram and log-periodogram kernel smoothing is proposed. The proposed methodology is based on a combining of both the "plug-in" and the unbiased risk estimation techniques. The new methodology is easy to describe, simple to implement, and its good performance has been demonstrated by simulation. It is also anticipated that the performance can be further improved, at an expense of increasing computational time, by using the "kernel-selection" technique proposed in [19].

Acknowledgements

The author is grateful to the reviewers for their constructive comments.

References

[1] S.-T. Chiu, Some stabilized bandwidth selectors for nonparametric regression, Ann. Stat. 19 (1991) 1528–1546.
[2] S.-T. Chiu, An automatic bandwidth selector for kernel density estimation, Biometrika 79 (1992) 771–782.
[3] J. Fan, E. Kreutzberger, Automatic local smoothing for spectral density estimation, Scand. J. Stat. 25 (1998) 359–369.
[4] H.-Y. Gao, Choice of thresholds for wavelet shrinkage estimate of the spectrum, J. Time Ser. Anal. 18 (1997) 231–251.
[5] T. Gasser, H.-G. Műller, V. Mammitzsch, Kernels for nonparametric curve estimation, J. Roy. Stat. Soc. Ser. B 47 (1985) 238–252.
[6] P. Hall, J.S. Marron, B.U. Park, Smoothed cross-validation, Probab. Theory Related Fields 92 (1992) 1–20.
[7] W. Hardle, P. Hall, J.S. Marron, How far are automatically chosen regression smoothing parameters from their optimum? J. Am. Stat. Assoc. 83 (1988) 86–95.
[8] W. Härdle, P. Hall, J.S. Marron, Regression smoothing parameters that are not far from their optimum, J. Am. Stat. Assoc. 87 (1992) 227–233.
[9] C.M. Hurvich, Data-driven choice of a spectrum estimate: extending the applicability of cross-validation methods, J. Am. Stat. Assoc. 80 (1985) 933–940.
[10] C. Kooperberg, C.J. Stone, Y.K. Troung, Logspline estimation of a possibly mixed spectral distribution, J. Time Ser. Anal. 16 (1995) 359–388.
[11] T.C.M. Lee, A simple span selector for periodogram smoothing, Biometrika 84 (1997) 965–969.
[12] T.C.M. Lee, V. Solo, Bandwidth selection for local linear regression: a simulation study, Comput. Stat. 14 (1999) 515–532.
[13] P. Moulin, Wavelet thresholding techniques for power spectrum estimation, IEEE Trans. Signal Process. 42 (1994) 3126–3136.
[14] Y. Pawitan, F. O'Sullivan, Nonparametric spectral density estimation using penalized Whittle likelihood, J. Am. Stat. Assoc. 89 (1994) 600–610.
[15] K.S. Riedel, A. Sidorenko, Adaptive smoothing of the log-spectrum with multiple tapering, IEEE Trans. Signal Process. 44 (1996) 1794–1800.
[16] B. Silverman, Density Estimation for Statistics and Data Analysis, Chapman & Hall, London, 1986.
[17] V. Solo, A sure fired way to choose smoothing parameters in ill-conditioned inverse problems, Proceedings of IEEE 1996 International Conference on Image Processing, Vol. 3, 1996, pp. 89–92.
[18] C.M. Stein, Estimation of the mean of a multivariate normal distribution, Ann. Stat. 9 (1981) 1135–1151.
[19] P. Stoica, T. Sundin, Optimally smoothed periodogram, Signal Processing 78 (1999) 253–264.
[20] G. Wahba, Automatic smoothing of the log periodogram, J. Am. Stat. Assoc. 75 (1980) 122–132.
[21] A.T. Walden, D.B. Percival, E.J. McCoy, Spectrum estimation by wavelet thresholding of multitaper estimators, IEEE Trans. Signal Process. 46 (1998) 3153–3165.
[22] M. Wand, R.G. Gutierrez, Exact risk approaches to smoothing parameter selection, J. Nonparametric Stat. 8 (1997) 337–354.
[23] M.P. Wand, M.C. Jones, Kernel Smoothing, Chapman & Hall, London, 1995.