Approximate variances for tapered spectral estimates

Michael Amrein and Hans R. Künsch

Seminar für Statistik
ETH Zentrum
CH-8092 Zürich, Switzerland

September 2010

Abstract

We propose an approximation of the asymptotic variance that removes a certain discontinuity in the formula for the raw and the smoothed periodogram in case a data taper is used. It is based on an approximation of the covariance of the (tapered) periodogram at two arbitrary frequencies.

Key words: Asymptotic variance, data taper, (smoothed) periodogram.

1 Introduction

Spectral estimation is by now a standard topic in time series analysis, and many excellent books are available, e.g. Percival and Walden (1993). The purpose of this short note is to propose an approximation of the asymptotic variance that removes a certain discontinuity in the formula for the raw and the smoothed periodogram in case a taper is used. The standard asymptotic variance of the raw periodogram is independent of the taper chosen, see Formulae (222b) and (223c) in Percival and Walden (1993). However, this changes when the raw periodogram is smoothed over frequencies close by. Then a variance inflation factor C_h, see (4), appears which is equal to one if no taper is used and greater than one otherwise, compare Table 248 in Percival and Walden (1993). The reason for this is that tapering introduces correlations between the raw periodogram at different Fourier frequencies. Because of this, the variance reduction due to smoothing is smaller in the case of no tapering.

The above variance inflation factor is justified asymptotically when the number of Fourier frequencies that are involved in the smoothing tends to infinity (more slowly than the number of observations, otherwise we would have a bias). Hence, if only little smoothing is used, then we except something in between:

some increase in the variance, but less than the asymptotic variance inflation factor C_h. We give here a formula, see (5), which is almost as simple as the inflation factor, but which takes the amount of smoothing into account.

2 Notation and preliminaries

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a real-valued stationary process with observation frequency $\frac{1}{\Delta}$, mean $\mathrm{E}[X_t] = \mu$, autocovariances $s_\tau := \mathrm{Cov}(X_t, X_{t+\tau})$ and spectral density $S(f)$. We assume that $X_1, X_2, \ldots, X_N$ have been observed.

The tapered periodogram (called direct spectral estimator in Percival and Walden (1993)) is

$$\hat{S}^{(tp)}(f) := \frac{\Delta}{\sum_{t=1}^N h_t^2} \left| \sum_{t=1}^N h_t (X_t - \tilde{\mu}) e^{-i2\pi i t f \Delta} \right|^2$$

for $f \in [0, 1/(2\Delta)]$. Here the estimator $\tilde{\mu}$ is usually either the arithmetic mean $\bar{X}$ or the weighted average

$$\frac{\sum_{t=1}^N h_t X_t}{\sum_{t=1}^N h_t}.$$

The latter has the property that $\hat{S}^{(tp)}(0) = 0$. Since the choice is irrelevant for the asymptotics, we can use either version. The taper $(h_1, \ldots, h_N)$ is chosen to reduce the discontinuities of the observation window at the edges $t = 1$ and $t = N$. Usually, it has the form

$$h_t = h\left(\frac{2t-1}{2N}\right)$$

with a function h that is independent of the sample size N. A popular choice is the split cosine taper

$$h^p(x) = \begin{cases} \frac{1}{2}(1 - \cos(2\pi x/p)) & 0 \le x \le \frac{p}{2} \\ 1 & \frac{p}{2} < x < 1 - \frac{p}{2} \\ \frac{1}{2}(1 - \cos(2\pi(1-x)/p)) & 1 - \frac{p}{2} \le x \le 1 \end{cases} \tag{1}$$

The tapered periodogram has the approximate variance

$$\mathrm{Var}[\hat{S}^{(tp)}(f)] \approx S(f)^2, \quad f \notin \{0, 1/(2\Delta)\} \tag{2}$$

(see e.g. Percival and Walden (1993), Formula (222b)). In particular, it does not converge to zero. Because of this, one usually smoothes the periodogram over a small band of neighboring frequencies. We smooth discretely over an equidistant grid of frequencies. Let

$$f_{N',k} = \frac{k}{N'\Delta} \quad (0 \le k \le N'/2)$$

for N' an integer greater or equal to N and smaller than $2N$. Then the tapered and smoothed spectral estimate is

$$\hat{S}^{(ts)}(f_{N',k}) = \sum_{j=-M}^{M} g_j \hat{S}^{(tp)}(f_{N',k-j}),$$

where the g_j's are weights with the following properties

$$g_j > 0, \ g_j = g_{-j} \ (-M \leq j \leq M), \quad \sum_{j=-M}^{M} g_j = 1.$$

If $k < M$, the smoothing includes the value $\hat{S}^{(tp)}(0)$ which is equal or very close to zero if the mean μ is estimated. In this case, we should exclude $j = k$ from the sum.

3 Approximations of the variance of spectral estimators

The usual approximation for the variance of $\hat{S}^{(ts)}(f_{N',k})$ is

$$\mathrm{Var}(\hat{S}^{(ts)}(f_{N',k})) \approx S(f_{N',k})^2 \sum_{r=-M}^{M} g_r^2 \frac{N'}{N} C_h \tag{3}$$

for $k \neq 0, N'/2$ where

$$C_h = \frac{\sum_{t=1}^{N} h_t^4 / N}{(\sum_{t=1}^{N} h_t^2 / N)^2}. \tag{4}$$

This formula is given in Bloomfield (1976), p. 195, and it is implemented in the function “spec.pgram” in the language for statistical computing R. In order to see that it is the same as Formula (248a) in Percival and Walden (1993), one has to go back to the definition of W_m in terms of the weights g_j which is given by the formulae (237c), (238d) and (238e).

If we put $M = 0$, this is different from (2). The reason for this difference is that (3) is valid in the limit $M \to \infty$ and $M/N' \to 0$. But in applications M is often small, e.g. $M = 1$ and one wonders how good the approximation is in such a case.

We propose here as alternative the following approximation (for $\mathrm{Var}(\hat{S}^{(ts)}(f_{N',k}))$)

$$S(f_{N',k})^2 \left(\sum_{r=-M}^{M} g_r^2 + 2 \sum_{l=1}^{2M} \frac{\left| H_2^{(N)}(f_{N',l}) \right|^2}{H_2^{(N)}(0)^2} \sum_{r=-M}^{M-l} g_r g_{r+l} \right) \tag{5}$$

(again for $k \neq 0, N'/2$) where

$$H_2^{(N)}(f) = \frac{1}{N} \sum_{t=1}^{N} h_t^2 e^{-i2\pi t f \Delta}.$$

Table 1: New and usual (in parenthesis) variance approximation in the limit $N' = N \to \infty$ relative to the squared spectral density for the split cosine taper with $p = 0.2, 0.5, 1$ and uniform weights $g_j = 1/(2M+1)$ in dependence of M.

p \ M	0	1	2	3
0.2	1.0000 (1.1163)	0.3453 (0.3721)	0.2122 (0.2233)	0.1537 (0.1595)
0.5	1.0000 (1.3471)	0.4017 (0.4490)	0.2521 (0.2694)	0.1836 (0.1924)
1.0	1.0000 (1.9444)	0.5370 (0.6481)	0.3489 (0.3889)	0.2574 (0.2778)

In order to compute this expression, we need to compute the convolution of the weights (g_j) and the discrete Fourier transform of the squared taper. The former is usually not a problem since M is substantially smaller than N'. Using the fast Fourier transform, exact computation of the latter is in most cases also possible. If not, then by the Lemma below we can use

$$H_2^{(N)}(f) \approx \int_0^1 h^2(u) e^{-i2\pi N u f \Delta} du \; e^{-i\pi f \Delta} \frac{\pi f \Delta}{\sin(\pi f \Delta)}.$$

Choosing a simple form for the function h, we can compute the integral on the right exactly. It is obvious that (5) agrees with (2) for $M = 0$. In the next section, we show that it also agrees with (3) for M large.

For the split cosine taper $h^p(.)$, see (1), and uniform weights $g_j = 1/(2M+1)$, we get for the new approximation (5) in the limit $N' = N \to \infty$ and $k/N' \to f$

$$\frac{S(f)^2}{2M+1}\left\{1+\frac{1}{(1-5p/8)^2}\cdot\sum_{l=1}^{2M}\frac{2M+1-l}{2M+1}\left(\frac{(2l^2p^2-5)\sin(lp\pi)}{2l\pi(lp-2)(lp-1)(lp+1)(lp+2)}\right)^2\right\}$$

with an extension by continuity when lp equals 1 or 2, whereas the usual approximation (3) yields

$$\frac{S(f)^2}{2M+1}\left(\frac{1-93p/128}{(1-5p/8)^2}\right)$$

(see Bloomfield (1976), page 194 and 195). In Table 1, we compare the two approximations (relative to $S(f)^2$) for $p = 0.2, 0.5, 1$ and $M = 0, 1, 2, 3$.

4 Justification of the approximation

The idea is simple: We just plug in a suitable approximation for the covariances of the tapered periodogram values into the exact expression for the variance $\mathrm{Var}(\hat{S}^{(ts)}(f_{N',k}))$, that is,

$$\sum_{r=-M}^{M}\sum_{s=-M}^{M} g_r g_s \mathrm{Cov}(\hat{S}^{(tp)}(f_{N',k-r}), \hat{S}^{(tp)}(f_{N',k-s})).$$

The asymptotic behavior of these covariances is well known. Theorem 5.2.8 of Brillinger (1975) shows that, under suitable conditions, we have for two frequencies $0 < f \le g < \frac{1}{2\Delta}$ that $\mathrm{Cov}(\hat{S}^{(tp)}(f), \hat{S}^{(tp)}(g))$ is equal to

$$\frac{S(f)S(g)}{\left|H_2^{(N)}(0)\right|^2} \cdot \left(\left|H_2^{(N)}(f-g)\right|^2 + \left|H_2^{(N)}(f+g)\right|^2\right) + O(N^{-1}). \qquad (6)$$

The statement in Brillinger (1975) is actually asymmetric in f and g since it has $S(f)^2$ instead of $S(f)S(g)$ on the right. Our statement can be proved by the same argument since the covariance is of the order $O(N^{-1})$ unless $|f-g| = O(N^{-1})$.

Using the approximation (6) directly would lead to an approximation which depends on k. Having to compute $N'/2$ different approximate variances is usually too complicated. However, the second term $|H_2^{(N)}(f+g)|^2$ is small unless $\Delta(f+g)$ is close to zero modulo one. This has been pointed out by Thomson (1977), see also the discussion on p. 230–231 of Percival and Walden (1993). If we omit the second term, then we obtain our new approximation (5) by a simple change in the summation indices.

We next give a simple Lemma that justifies the omission of the second term in (6). In addition, it also shows how the usual approximation (3) follows from (5).

Lemma 4.1. *If ψ is twice continuously differentiable on $[0,1]$ then for any $\lambda \in [0, 0.5]$*

$$\frac{1}{N}\sum_{t=1}^{N} \psi\left(\frac{2t-1}{2N}\right) e^{-i2\pi\lambda t} = \int_0^1 \psi(u) e^{-i2\pi N\lambda u} du \; e^{-i\pi\lambda} \frac{\pi\lambda}{\sin(\pi\lambda)} + R$$

where

$$|R| \le \frac{1}{24} \sup |\psi''(x)| \frac{1}{N}.$$

Proof. Put $\epsilon = 1/(2N)$. By a Taylor expansion, we obtain for any $x \in [0,1]$

$$\int_{x-\epsilon}^{x+\epsilon} \psi(u) e^{-i2\pi N\lambda u} du$$
$$= \psi(x) e^{-i2\pi N\lambda x} \int_{-\epsilon}^{\epsilon} e^{-i2\pi N\lambda u} du + \psi'(x) e^{-i2\pi N\lambda x} \int_{-\epsilon}^{\epsilon} u e^{-i2\pi N\lambda u} du + R$$

where the remainder satisfies

$$|R| \le \frac{\epsilon^3}{3} \sup |\psi''(x)| = \frac{1}{24} \sup |\psi''(x)| \frac{1}{N^3}.$$

Next, observe that

$$\int_{-\epsilon}^{\epsilon} e^{-i2\pi N\lambda u} du = \frac{\sin(\pi\lambda)}{\pi\lambda} \frac{1}{N}$$

and

$$\int_{-\epsilon}^{\epsilon} u e^{-i2\pi N\lambda u} du = \frac{i}{2\pi\lambda}\left(\cos(\pi\lambda) - \frac{\sin(\pi\lambda)}{\pi\lambda}\right) \frac{1}{N^2}.$$

From this the lemma follows by taking $x = (2t-1)/(2N)$ for $t = 1, \ldots, N$ and summing up all terms. □

If $\psi(0) = \psi(1) = 0$, then by partial integration

$$\left| \int_0^1 \psi(u) e^{-i2\pi N\lambda u} du \right| \leq \frac{2 \sup |\psi'(x)| + \sup |\psi''(x)|}{4\pi^2} \frac{1}{N^2 \lambda^2}.$$

Hence by setting $\psi(u) = h^2(u)$, we obtain

$$H_2^{(N)}(f) \leq \text{const.} \frac{1}{N^2 (f\Delta)^2} + \text{const.} \frac{1}{N} \quad (7)$$

Therefore the second term in (6) is negligible unless $f+g$ is of the order $O(N^{-1})$.

Finally, we derive the usual variance approximation (3) as follows. If the weights g_j change smoothly as a function of the lag j, then for any fixed k

$$\sum_{r=-M}^{M-l} g_r g_{r+l} \sim \sum_{r=-M}^{M} g_r^2.$$

Hence by dominated convergence

$$\mathrm{Var}(\hat{S}^{(ts)}(f_{N',k})) \approx S(f_{N',k})^2 \frac{1}{H_2^{(N)}(0)^2} \sum_{r=-M}^{M} g_r^2 \sum_{l=-M}^{M} \left| H_2^{(N)}(f_{N',l}) \right|^2.$$

Note that $H_2^{(N)}(0) = 1/N \cdot \sum_{t=1}^{N} h_t^2$ and because of (7),

$$\sum_{l=-M}^{M} \left| H_2^{(N)}(f_{N',l}) \right|^2 = \sum_{l=-N'/2}^{N'/2} \left| H_2^{(N)}(f_{N',l}) \right|^2 + O(N^{-1}).$$

Furthermore

$$\sum_{l=-N'/2}^{N'/2} \left| H_2^{(N)}(f_{N',l}) \right|^2 \approx N' \int_{-0.5}^{0.5} |H_2^{(N)}(f/\Delta)|^2 df = \frac{N'}{N} \frac{1}{N} \sum_{t=1}^{N} h_t^4.$$

References

Bloomfield, P. (1976). *Fourier Analysis of Time Series: An Introduction.* Wiley, NY.

Brillinger, D. R. (1975). *Time Series, Data Analysis and Theory.* International Series in Decision Processes. Holt, Rinehart and Winston, Inc., New York.

Percival, D. B. and Walden, A. T. (1993). *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques.* Cambridge University Press.

Thomson (1977). Spectrum estimation techniques for characterization and development of WT4 waveguide - I. *Bell System Technical Journal*, 56:1769–1815.