Introduction to Time Series Analysis. Lecture 22.

1. Review: The periodogram, the smoothed periodogram.
2. Other smoothed spectral estimators.
3. Consistency.
4. Asymptotic distribution.

Review: Periodogram

The periodogram is defined as

$$I(\nu) = X_c^2(\nu) + X_s^2(\nu).$$

$$X_c(\nu) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \cos(2\pi t\nu) x_t,$$

$$X_s(\nu) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sin(2\pi t\nu_j) x_t.$$

Under general conditions, $X_c(\nu_j)$, $X_s(\nu_j)$ are asymptotically independent and $N(0, f(\nu_j)/2)$. Thus, $\mathbf{E} I(\hat{\nu}^{(n)}) \to f(\nu)$, but $\mathrm{Var}(I(\hat{\nu}^{(n)})) \to f(\nu)^2$.

Review: smoothed periodogram

If $f(\nu)$ is approximately constant in a band of frequencies $[\nu_k - L/(2n), \nu_k + L/(2n)]$, we can average the periodogram over this band:

$$\hat{f}(\nu_k) = \frac{1}{L} \sum_{l=-(L-1)/2}^{(L-1)/2} I(\nu_k - l/n)$$

$$= \frac{1}{L} \sum_{l=-(L-1)/2}^{(L-1)/2} \left(X_c^2(\nu_k - l/n) + X_s^2(\nu_k - l/n) \right).$$

Review: smoothed periodogram

Under general conditions, the $X_c(\nu_k - l/n)$ and $X_s(\nu_k - l/n)$ are asymptotically independent and $N(0, f(\nu_k - l/n)/2)$. Thus, $\mathbf{E}\hat{f}(\nu^{(n)}) \to f(\nu)$ and $\mathrm{Var}\hat{f}(\nu^{(n)}) \to f^2(\nu)/L$.

Notice the *bias-variance trade off*:

1. Our assumption that f is approximately constant on $[\nu - L/(2n), \nu + L/(2n)]$ becomes worse as L increases, so the difference between $\hat{f}(\hat{\nu}^{(n)})$ and $f(\nu)$ (the bias) will increase with L.

2. The variance of our estimate, $\mathrm{Var}\hat{f}(\hat{\nu}^{(n)})$ decreases with L.

Other smoothed spectral estimators

Instead of computing an unweighted average of the periodogram at all nearby frequencies, it is common to consider other weighted averages, typically with a smoother weighting function.

Consider the weighted average

$$\hat{f}(\nu) = \sum_{|j| \leq L_n} W_n(j) I(\hat{\nu}^{(n)} - j/n),$$

where the bandwidth L_n is allowed to vary with n, and W_n is called the *spectral window function*.

Other smoothed spectral estimators

For example, if

$$W_n(j) = \begin{cases} \frac{1}{L} & \text{if } |j| < L/2, \\ 0 & \text{otherwise,} \end{cases}$$

then we have the smoothed spectral estimator we consider earlier,

$$\begin{aligned} \hat{f}(\nu) &= \sum_j W_n(j) I(\hat{\nu}^{(n)} - j/n) \\ &= \frac{1}{L} \sum_{|j|<L/2} I(\hat{\nu}^{(n)} - j/n). \end{aligned}$$

This is *Daniell's estimator* (P. J. Daniell, University of Sheffield, 1946).

Consistency of nonparametric spectral estimation

Suppose L_n and W_n satisfy

$$L_n \to \infty, \qquad \frac{L_n}{n} \to 0,$$

$$W_n(j) \geq 0, \qquad W_n(j) = W_n(-j), \qquad \sum_{|j| \leq L_n} W_n(j) = 1,$$

$$\sum_{|j| \leq L_n} W_n^2(j) \to 0$$

as $n \to \infty$, then...

Consistency of nonparametric spectral estimation

. . . for a large class of stationary processes, $\hat{f}(\nu) \to f(\nu)$ in the mean square sense. In particular, $\mathbf{E}\hat{f}(\nu) \to f(\nu)$ and

$$\left(\sum_{|j| \le L_n} W_n^2(j) \right)^{-1} \text{Cov}(\hat{f}(\nu_1), \hat{f}(\nu_2)) \to \begin{cases} f^2(\nu_1) & \text{if } \nu_1 = \nu_2 \in (0, 1/2) \\ 0 & \text{if } \nu_1 \neq \nu_2. \end{cases}$$

The conditions on the bandwidth parameter L_n ensure that, as the sample size grows, the window width goes to zero, but includes an infinite number of terms. The conditions on the spectral window function W_n ensure that the expectation of $\hat{f}(\nu)$ converges to $f(\nu)$ and its variance converges to zero.

Consistency of nonparametric spectral estimation

$$\mathbf{E}\left(\hat{f}(\nu)\right) = \sum_{|j|\le L_n} W_n(j)\mathbf{E}\left(I(\hat{\nu}^{(n)} - j/n)\right)$$
$$\approx \sum_{|j|\le L_n} W_n(j) f(\nu) = f(\nu).$$

$$\text{Var}\left(\hat{f}(\nu)\right) = \sum_{j,k} W_n(j) W_n(k) \text{Cov}\left(I(\hat{\nu}^{(n)} - j/n), I(\hat{\nu}^{(n)} - k/n)\right)$$
$$\approx \sum_{j} W_n^2(j) \text{Var}\left(I(\hat{\nu}^{(n)} - j/n)\right)$$
$$\approx f^2(\nu) \sum_{j} W_n^2(j) \to 0.$$

Nonparametric spectral estimation: asymptotics

Recall that for Daniell's estimator we have

$$\hat{f}(\nu_k) = \frac{1}{L} \sum_{l=-(L-1)/2}^{(L-1)/2} \left(X_c^2(\nu_k - l/n) + X_s^2(\nu_k - l/n) \right),$$

which is (asymptotically) a sum of $2L$ independent χ_1^2 random variables, so

$$\hat{f}(\nu_k) \sim f(\nu_k) \frac{\chi_{2L}^2}{2L}.$$

But for a non-uniform weighting, we form a weighted sum of these χ_1^2 random variables, so we cannot count up the degrees of freedom in the same way. But we can still approximate a general smoothed spectrum by $\hat{f}(\nu_k) \sim c_k \chi_d^2$ for some c_k and d.

Nonparametric spectral estimation: asymptotics

Suppose that $\hat{f}(\nu_k) \sim c_k \chi^2_d$. What values should c_k and d take? We have, for a suitable spectral window W_n,

$$f(\nu_k) \approx \mathbf{E}\hat{f}(\nu_k) = c_k d,$$

$$f^2(\nu_k) \sum_{|j| \le L_n} W_n^2(j) \approx \mathrm{Var}\hat{f}(\nu_k) = 2c_k^2 d.$$

Thus, we get

$$c_k = \frac{f(\nu_k)}{d},$$

$$2c_k = f(\nu_k) \sum_{|j| \le L_n} W_n^2(j),$$

$$d = \frac{2}{\sum_{|j| \le L_n} W_n^2(j)}.$$

Nonparametric spectral estimation: asymptotics

$$c_k = \frac{f(\nu_k)}{d},$$
$$d = \frac{2}{\sum_{|j|\le L_n} W_n^2(j)}.$$

This d is often referred to as the *equivalent degrees of freedom* for a smoothed spectrum. Under suitable conditions (and for a slightly different definition of d), it can be shown that, asymptotically,

$$\hat{f}(\nu^{(n)}) \sim f(\nu)\frac{\chi_d^2}{d}.$$

Nonparametric spectral estimation: the lag window

We can also view smoothing the spectrum in the frequency domain as smoothing in the time domain, via

$$\hat{f}(\nu) = \sum_{|j| \le L_n} w_n(j) \hat{\gamma}(j) e^{-2\pi i \nu j},$$

where w_n is the inverse Fourier transform of the spectral window. This is known as the *lag window*.

Tapering techniques are also popular: Define $y_t = h_t x_t$ for some weighting function h_t. Then the tapered estimator is the smoothed spectral estimator for the tapered series y_t. For a weighting function h_t that smoothly diminishes values near the ends of the time series, we see less leakage.