Introduction to Time Series Analysis. Lecture 23.

1. Review: The smoothed periodogram.
2. Examples.
3. Parametric spectral density estimation.

Review: Periodogram

The periodogram is defined as

$$\begin{aligned} I(\nu) &= |X(\nu)|^2 \\ &= X_c^2(\nu) + X_s^2(\nu). \end{aligned}$$

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

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

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 spectral estimators

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

where the *spectral window function* satisfies $L_n \to \infty$, $L_n/n \to 0$, $W_n(j) \geq 0$, $W_n(j) = W_n(-j)$, $\sum W_n(j) = 1$, and $\sum W_n^2(j) \to 0$.

Then $\hat{f}(\nu) \to f(\nu)$ (in the mean square sense), and asymptotically

$$\hat{f}(\nu_k) \sim f(\nu_k) \frac{\chi_d^2}{d},$$

where $d = 2/\sum W_n^2(j)$.

Example: Southern Oscillation Index

Figure 4.4 in the text shows the periodogram of the SOI time series. The SOI is the scaled, standardized, mean-adjusted, difference between monthly average air pressure at sea level in Tahiti and Darwin:

$$SOI = 10 \frac{P_{\text{Tahiti}} - P_{\text{Darwin}}}{\sigma} - \bar{x}.$$

For the time series in the text, $n = 453$ months.

The periodogram has a large peak at $\nu = 0.084$ cycles/sample. This corresponds to 0.084 cycles per month, or a period of $1/0.084 = 11.9$ months.

There are smaller peaks at $\nu \approx 0.02$: $I(0.02) \approx 1.0$. The frequency $\nu = 0.02$ corresponds to a period of 50 months, or 4.2 years.

Example: Southern Oscillation Index

Consider the hypothesized El Niño effect, at a period of around four years. The approximate 95% confidence interval at this frequency, $\nu = 1/(4 \times 12)$, is

$$\frac{2I(\nu)}{\chi_2^2(0.025)} \leq f(\nu) \leq \frac{2I(\nu)}{\chi_2^2(0.975)}$$
$$\frac{2 \times 0.64}{7.3778} \leq f(\nu) \leq \frac{2 \times 0.64}{0.0506}$$
$$0.17 \leq f(\nu) \leq 25.5.$$

The lower extreme of this confidence interval is around the noise baseline, so it is difficult to conclude much about the hypothesized El Niño effect.

Example: Southern Oscillation Index

Figure 4.5 in the text shows the smoothed periodogram, with $L = 9$.

Again, there is a large peak at $\nu = 0.08$ cycles/month (period ≈ 12 months). There are smaller peaks at integer multiples of this frequency (harmonics).

There is also a peak at $\nu = 0.0215$ (period 46.5 months): $\hat{f}(0.0215) \approx 0.62$.

Example: Southern Oscillation Index

The approximate 95% confidence interval at the hypothesized El Niño frequency is

$$\frac{2L\hat{f}(\nu)}{\chi^2_{2L}(0.025)} \leq f(\nu) \leq \frac{2L\hat{f}(\nu)}{\chi^2_{2L}(0.975)}$$
$$\frac{18 \times 0.62}{31.526} \leq f(\nu) \leq \frac{18 \times 0.62}{8.231}$$
$$0.354 \leq f(\nu) \leq 1.36.$$

The lower extreme of this confidence interval is well above the noise baseline (the level of the spectral density if the signal were white and the energy were uniformly spread across frequencies).

The text modifies the number of degrees of freedom slightly, to account for the fact that the signal is padded with zeros to make n a highly composite number, which simplifies the computation of the periodogram.

Choosing the bandwidth

A common approach is to start with a large bandwidth, and look at the effect on the spectral estimates as it is reduced ('closing the window'). As the bandwidth becomes too small, the variance gets large and the spectral estimate becomes more jagged, with spurious peaks introduced. But if it is too small, the spectral estimate is excessively smoothed, and details of the shape of the spectrum are lost.

The value of $L = 9$ chosen in the text for Figure 4.5 corresponds to a bandwidth of $B = L/n = 9/480 = 0.01875$ cycles per month. This means we are averaging over frequencies in a band of this width, so we are treating the spectral density as approximately constant over this bandwidth. Equivalently, we are not hoping to resolve frequencies more finely than about half of this bandwidth.

Simultaneous confidence intervals

We derived the confidence intervals for $f(\nu)$ assuming that ν was fixed. But in examining peaks, we might wish to choose ν after we've seen the data. If we want to make statements about the probability that k unlikely events $E_1, \ldots, E_k$ occur, we can use the Bonferroni inequality (also called the union bound):

$$\Pr\left\{\bigcup_{i=1}^{k} E_i\right\} \leq \sum_{i=1}^{k} \Pr\{E_i\},$$

and this probability is no more than $k\alpha$ if $\Pr\{E_i\} = \alpha$. For example, if E_i represents the event that $f(\nu_i)$ falls outside some confidence interval at level α, then we can bound the probability that the spectral density is far from our estimates at any of the frequencies $\nu_1, \ldots, \nu_k$.

Parametric versus nonparametric estimation

Parametric estimation = estimate a model that is specified by a fixed number of parameters.

Nonparametric estimation = estimate a model that is specified by a number of parameters that can grow as the sample grows.

Thus, the smoothed periodogram estimates we have considered are *nonparametric*: the estimates of the spectral density can be parameterized by estimated values at each of the Fourier frequencies. As the sample size grows, the number of distinct frequency values increases.

The time domain models we considered (linear processes) are *parametric*. For example, and ARMA(p,q) process can be completely specified with $p + q + 1$ parameters.

Parametric spectral estimation

In *parametric* spectral estimation, we consider the class of spectral densities corresponding to ARMA models.

Recall that, for a linear process $Y_t = \psi(B)W_t$, $f_y(\nu) = \left|\psi\left(e^{2\pi i\nu}\right)\right|^2 \sigma_w^2$.

For an AR model, $\psi(B) = 1/\phi(B)$, so $\{Y_t\}$ has the rational spectrum

$$\begin{aligned} f_y(\nu) &= \frac{\sigma_w^2}{\left|\phi\left(e^{-2\pi i\nu}\right)\right|^2} \\ &= \frac{\sigma_w^2}{\phi_p^2 \prod_{j=1}^{p} \left|e^{-2\pi i\nu} - p_j\right|^2}, \end{aligned}$$

where p_j are the poles, or roots of the polynomial ϕ.

Parametric spectral estimation

The typical approach to parametric spectral estimation is to use the maximum likelihood parameter estimates $(\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\sigma}_w^2)$ for the parameters of an AR(p) model for the process, and then compute the spectral density for this estimated AR model:

$$\hat{f}_y(\nu) = \frac{\hat{\sigma}_w^2}{\left|\hat{\phi}\left(e^{-2\pi i\nu}\right)\right|^2}.$$

Parametric spectral estimation

For large n,

$$\mathrm{Var}(\hat{f}(\nu)) \approx \frac{2p}{n} f^2(\nu).$$

(There are results for the asymptotic distribution, but they are rather weak.)

Notice the bias-variance trade-off in the parametric case: as we increase the number of parameters, p:

• The bias decreases; we can model more complex spectra. For example, with an AR(p), we cannot have more than $\lfloor p/2 \rfloor$ spectral peaks in the interval $(0, 1)$. (This is because each pair of complex conjugate poles contributes one factor and hence peak to the product.)

• The variance increases linearly with p.

ARMA spectral estimation

Sometimes ARMA models are used instead: estimate the parameters of an ARMA(p,q) model and compute its spectral density (recall that $\psi(B) = \theta(B)/\phi(B)$):

$$\hat{f}(\nu) = \hat{\sigma}_w^2 \left| \frac{\hat{\theta}(e^{-2\pi i\nu})}{\hat{\phi}\left(e^{-2\pi i\nu}\right)} \right|^2 .$$

However, it is more common to use large AR models, rather than ARMA models.

Parametric versus nonparametric spectral estimation

The main advantage of parametric spectral estimation over nonparametric is that it often gives better *frequency resolution* of a small number of peaks: To keep the variance down with a parametric estimate, we need to make sure that we do not try to estimate too many parameters. While this may affect the bias, even $p = 2$ allows a sharp peak at one frequency. In contrast, to keep the variance down with a nonparametric estimate, we need to make sure that the bandwidth is not too small. This corresponds to having a smooth spectral density estimate, so the frequency resolution is limited.

This is especially important if there is more than one peak at nearby frequencies.

The disadvantage is the inflexibility (bias) due to the use of the restricted class of ARMA models.

Parametric spectral estimation: Summary

Given data $x_1, x_2, \ldots, x_n$,

1. Estimate the AR parameters $\phi_1, \ldots, \phi_p, \sigma_w^2$ (for example, using Yule-Walker/least squares or maximum likelihood),
   and choose a suitable model order p (for example, using $\mathrm{AIC}_c = (n+p)/(n-p-2)$ or $\mathrm{BIC} = p \log n/n$).

2. Use the estimates $\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\sigma}_w^2$ to compute the estimated spectral density:

$$\hat{f}_y(\nu) = \frac{\hat{\sigma}_w^2}{\left|\hat{\phi}\left(e^{-2\pi i \nu}\right)\right|^2}.$$