Time Series lecture 8

Frequency domain estimation

Sofia Olhede

April 9, 2025

Lecture outline

1. The periodogram
2. Tapering
3. Multitapering

The periodogram

Spectral estimation

In this lecture, we will develop methodology to estimate the spectral density function from some observed time series. In other words, say we have a finite set of equally spaced time points

$$T = \{1, \ldots, n\},$$

and we observe X_t for all $t \in T$, then how do we estimate $S(f)$?

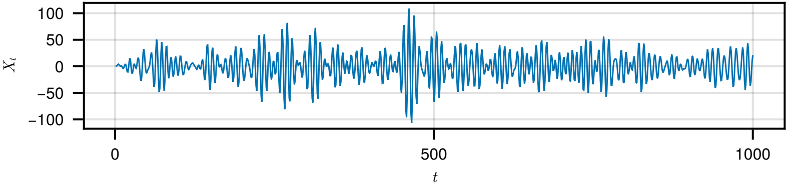

Figure: Simulated AR(4) time series with $n = 1000$

We set $\sigma^2 = 1$, $\phi = [2.7607, -3.8106, 2.6535, -0.9238]$, replicating an example from Percival and Walden (1993).

Recall the relationship

$$S(f) = \sum_{\tau \in \mathbb{Z}} \gamma_\tau e^{-2\pi i f \tau}.$$

We can therefore produce an estimator of $S(f)$ from $\hat{\gamma}_\tau$ by a plug–in estimator.

Definition 8.1 (The periodogram)

The periodogram is defined as

$$\widehat{S}^{(p)}(f) = \sum_{\tau \in \mathbb{Z}} \hat{\gamma}_\tau e^{-2\pi i f \tau} \tag{8.1}$$

where $\{\hat{\gamma}_\tau\}_{\tau \in \mathbb{Z}}$ is given by

$$\hat{\gamma}_\tau = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-|\tau|} (X_t - \bar{X})(X_{t+|\tau|} - \bar{X}) & \text{if } |\tau| \leq n-1, \\ 0 & \text{otherwise.} \end{cases}$$

Relation to the finite Fourier transform of the data

Proposition 8.2

The periodogram can be written in terms of the finite Fourier transform of the observed data, in that

$$\widehat{S}^{(p)}(f) = |J(f)|^2, \tag{8.2}$$

where recalling $T = \{1, \ldots, n\}$, *we define*

$$J(f) = \sqrt{\frac{1}{n}} \sum_{t \in T} \left(X_t - \bar{X}\right) e^{-2\pi i t f}. \tag{8.3}$$

- See the exercise sheet for a proof.
- $J(f)$ can be computed efficiently using the Fast Fourier Transform.
- In fact, one can compute the sample autocovariance efficiently by simply inverse transforming the periodogram!

Properties of the periodogram

Ideally as an estimator we would have

1. $\mathbb{E}\left[\widehat{S}^{(p)}(f)\right] = S(f)$,
2. $\text{Var}\left(\widehat{S}^{(p)}(f)\right) \to 0$ as $n \to \infty$,
3. $\text{Cov}\left(\widehat{S}^{(p)}(f), \widehat{S}^{(p)}(f')\right) = 0$ for $f \neq f'$.

However instead we find that

1. is approximately valid.
2. is false.
3. holds approximately if f and f' have a particular form (the Fourier frequencies).

We will deal with each of these in turn.

Expectation

Since the periodogram is a finite sum of the sample autocovariance, (and assuming zero-mean) we see

$$\begin{aligned}\mathbb{E}\left[\widehat{S}^{(p)}(f)\right] &= \sum_{\tau\in\mathbb{Z}} \mathbb{E}\left[\hat{\gamma}_\tau\right] e^{-2\pi i\tau f} \\ &= \sum_{\tau\in\mathbb{Z}} \left(1-\frac{|\tau|}{n}\right) \mathbb{1}_{[-n,n]}(\tau)\gamma_\tau e^{-2\pi i\tau f} \\ &= \sum_{\tau\in\mathbb{Z}} w_\tau \gamma_\tau e^{-2\pi i\tau f}\end{aligned}$$

where

$$w_\tau = \left(1-\frac{|\tau|}{n}\right) \mathbb{1}_{[-n,n]}(\tau)$$

is a weight sequence, which results in bias.

Example

Consider an example of an AR(4) model. Below is a comparison between the true spectral density and the periodogram from a simulated series with $n = 1000$ points. We see a substantial discrepancy here.

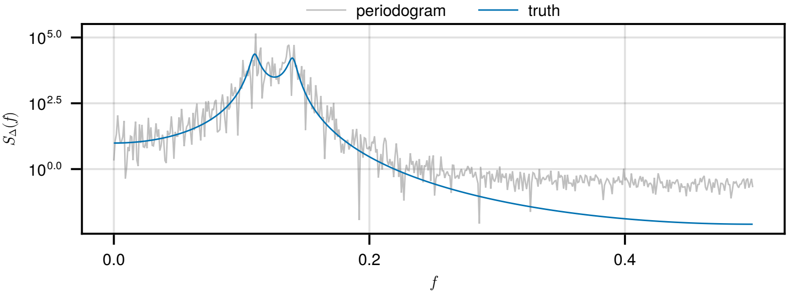

Figure: The periodogram of a simulated AR(4) model vs the true spectral density function, recreated from Percival and Walden (1993).

Expectation as a convolution

Since we are interested in the spectral density function, it may be useful to rewrite this expectation in the frequency domain. Since we have a multiplication in time, this becomes a convolution in frequency so that

$$\mathbb{E}\left[\widehat{S}^{(p)}(f)\right] = \int_{-\frac{1}{2}}^{\frac{1}{2}} S\left(f'\right) \mathcal{F}_n\left(f - f'\right) \mathrm{d}f' \tag{8.4}$$

where

$$\begin{aligned} \mathcal{F}_n(f) &= \sum_{\tau \in \mathbb{Z}} w_\tau e^{-2\pi i \tau f} \\ &= \begin{cases} \frac{1}{n} \frac{\sin^2(n\pi f)}{\sin^2(\pi f)} & \text{if } f \neq 0, \\ n & \text{if } f = 0. \end{cases} \end{aligned}$$

The function $\mathcal{F}_n$ called the Fejér kernel.

Blurring

Examining the Fejér kernel, we see that it does not decay quickly in frequency. This causes the bias seen earlier, an effect typically called blurring.

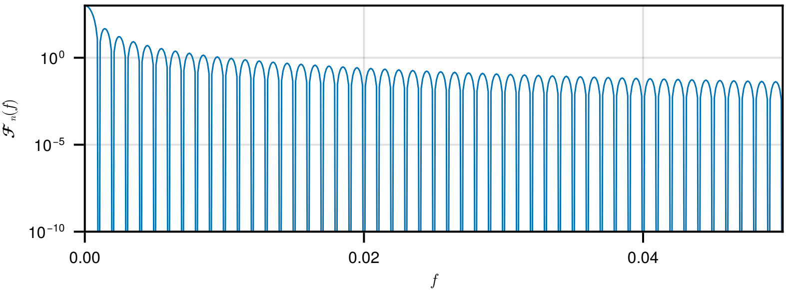

Figure: The Fejér kernel for $n = 1000$.

Variance

The second issue is the variance. In particular, one can show that (for $f \neq 0 \mod \pi$)

$$\operatorname{Var}\left(\widehat{S}^{(p)}(f)\right) \to S(f)^2 \tag{8.5}$$

as $n \to \infty$.

- So the variance does not decrease with increasing sample size.
- We cover solutions to both the bias and variance problems in the following sections.

Tapering

Tapering

One way to fix blurring is known as tapering. Here we replace the discrete Fourier transform by a tapered version

$$J_h(f) = \sum_{t \in T} h_t \cdot \left(X_t - \bar{X}\right) e^{-2\pi i t f} \tag{8.6}$$

where $\{h_t\}_{t \in T}$ is called a data taper, and satisfies

$$\|h\|_2^2 = 1.$$

Clearly setting $h_t = \sqrt{1/n}$ we recover the standard discrete Fourier transform. Now the tapered periodogram is

$$\widehat{S}_h^{(p)}(f) = |J_h(f)|^2 . \tag{8.7}$$

Effect of tapering

One can show that (see exercises)

$$\mathbb{E}\left[\widehat{S}_h^{(p)}(f)\right] = \int_{-1/2}^{1/2} S\left(f'\right)\left|H(f-f')\right|^2 \mathrm{d}f' \tag{8.8}$$

where H is the discrete Fourier transform of h.
We therefore want to choose h so that $|H(f)|^2$ is focused around zero frequency.

One choice of such tapers are the dpss tapers, which try to optimally achieve this. This is beyond the scope of this course, but see Percival and Walden (1993) for more details.

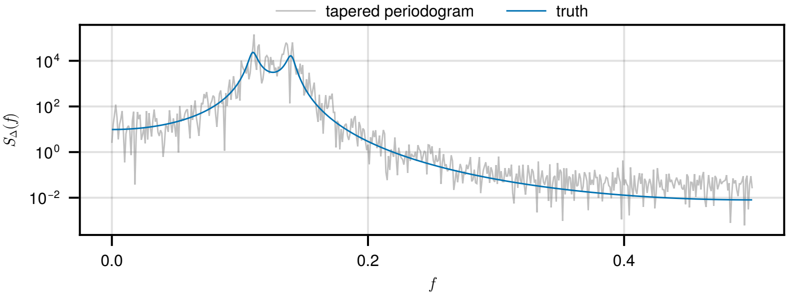

Figure: The same AR(4) example but applying a taper.

Taper kernel

Here the $|H|^2$ decays faster than Fejér kernel, resulting in the decrease in bias we previously noted.

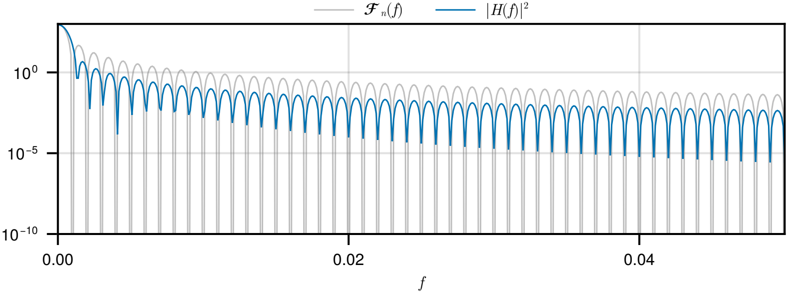

Figure: A comparison of the Fejér kernel and $|H|^2$ for a choice of taper.

Multitapering

A family of tapers

In order to reduce variance, we need to average. Define a family of tapers for $k = 1, \ldots, K$

$$h_k = \{h_{t,k}\}_{t \in T} \tag{8.9}$$

which are orthogonal, i.e. for $1 \leq k, k' \leq K$

$$\sum_{t \in T} h_{t,k} h_{t,k'} = \delta_{k,k'}.$$

Note that these tapers do depend on n, but we suppress this for notational convenience.

Multitaper

The simplest multitaper estimate is

$$\widehat{S}^{(mt)}(f) = \frac{1}{K}\sum_{k=1}^{K} \widehat{S}^{(p)}_{h_k}(f)\,. \tag{8.10}$$

We can determine its expectation:

$$\begin{aligned}\mathbb{E}\left[\widehat{S}^{(mt)}(f)\right] &= \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}\left[\widehat{S}^{(p)}_{h_k}(f)\right] \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \overline{\mathcal{H}}\left(f - f'\right) S\left(f'\right) \mathrm{d}f'.\end{aligned}$$

We call $\overline{\mathcal{H}}(f) = \frac{1}{K}\sum_{k=1}^{K} |H(f)|^2$ the average kernel.

One can show that asymptotically (see exercises)

$$\mathrm{Cov}\left(\widehat{S}^{(p)}_{h_k}(f), \widehat{S}^{(p)}_{h_{k'}}(f)\right) \to S(f)^2\,\delta_{k,k'}$$

as $n \to \infty$ and therefore

$$\mathrm{Var}\left(\widehat{S}^{(mt)}(f)\right) \approx \frac{S(f)^2}{K}.$$

Bibliography

Percival, D. B. and Walden, A. T. (1993). *Spectral analysis for physical applications*. Cambridge University Press.