Kernel spectral estimation

Basic spectral estimator: *Periodogram*

![Graph of log I(T)(l) vs. l]

Problems:
- asymptotically unbiased but inconsistent estimator
- fluctuates wildly (low correlation between neighboring frequencies)
- difficult to analyze (e.g. location of significant peaks)

Idea: Local smoothing of periodogram

![Graph of log \( \hat{f}^T(\lambda) \) vs. \( \lambda \)]

Kernel (regression) spectral estimator:

\[
\hat{f}^{(T)}(\lambda) = \int_{-\pi}^{\pi} w^{(T)}(\mu) I^{(T)}(\lambda - \mu) d\mu
\]

with kernel

\[
w^{(T)}(\lambda) = \frac{1}{b_T} w\left(\frac{\lambda}{b_T}\right)
\]

and bandwidth \( b_T \).
Kernel spectral estimator: the mean

Assumptions on kernel function:

○ $w$ is an even function of bounded variation

○ $\int \Pi w(\lambda) d\lambda = 1$

○ $\int \Pi w(\lambda)^2 d\lambda < \infty$

Mean of $\hat{f}^{(T)}(\lambda)$:

$$
\mathbb{E}(\hat{f}^{(T)}(\lambda)) = \int \Pi w^{(T)}(\lambda - \mu) \mathbb{E}(I^{(T)}(\mu)) d\mu = \int \Pi \int \Pi w^{(T)}(\lambda - \mu) \Phi^{(T)}_2(\mu - \eta) f(\eta) d\eta d\mu = \int \Pi K^{(T)}(\lambda - \eta) f(\eta) d\eta 
$$

where

$$
\Phi^{(T)}_2(\lambda) = \frac{|H^{(T)}(\lambda)|^2}{2\pi H^{(T)}_2(0)}
$$

$$
K^{(T)}(\lambda) = \int \Pi w^{(T)}(\lambda - \mu) \Phi^{(T)}_2(\mu) d\mu
$$

Implication:

If $b_T \to 0$ as $T \to \infty$ then $\hat{f}^{(T)}(\lambda)$ is asymptotically unbiased.

Question: How fast should $b_T$ converge to 0?

if $b_T \to 0$ too fast (e.g. in finite time), we obtain the periodogram $\sim$ inconsistency
Kernel spectral estimator: the variance

Variance of $\hat{f}(T)(\lambda)$:

$$\text{var}(\hat{f}(T)(\lambda)) = \frac{1}{T b_T} \cdot \frac{2\pi H_4}{H_2^2} \int_{\Pi} w(\mu)^2 d\mu \cdot f(\lambda)^2 + o\left(\frac{1}{T b_T}\right)$$

Implications:

○ The variance depends on the kernel through the factor $\int_{\Pi} w(\mu)^2 d\mu$.

○ The variance depends on the taper through the factor $H_4/H_2^2$.

○ Cauchy-Schwarz inequality implies

$$H_2^2 = \left(\int_0^1 h(t)^2 dt\right)^2 \leq \int_0^1 h(t)^4 dt \int_0^1 dt = H_4.$$ 

Thus tapering always increases the variance (due to loss of information)

Tradeoff between bias and variance

○ Tapering:

  Increase amount of tapered data ⇒ reduce bias, increase variance

○ Smoothing:

  Decrease bandwidth $b_T$ ⇒ reduce bias, increase variance
Asymptotic distribution and confidence intervals

Use discrete approximation
\[
\hat{f}(T)\lambda_j \approx \frac{2\pi}{T} \sum_{-\pi < \lambda_k \leq \pi} w(T)\lambda_k I(T)(\lambda_j - \lambda_k)
\]

The periodogram ordinates \( I(T)(\lambda_j - \lambda_k) \) are approximately independently \( \chi^2 \) distributed.

**Approximation for the distribution of \( \hat{f}(T)\lambda_j \):**

\[
\hat{f}(T)\lambda_j \overset{D}{\approx} a \chi^2 \nu
\]

where \( a\nu = \mathbb{E}(\hat{f}(T)\lambda_j) \) and \( 2a^2\nu = \text{var}(\hat{f}(T)\lambda_j) \)

\( \nu \) is called equivalent degrees of freedom

\[
\nu = \frac{2 \mathbb{E}(\hat{f}(T)\lambda_j)}{\text{var}(\hat{f}(T)\lambda_j)} \overset{T \to \infty}{\approx} \frac{2}{\pi^2 H_4^2 \int w(\mu)^2 d\mu} \frac{H_2^2 T b_T}{\mathbb{E}(\hat{f}(T)\lambda_j)}
\]

For \( Tb_T \to \infty \), we have

\[
\sqrt{Tb_T}(\hat{f}(T)(\lambda) - f(\lambda)) \overset{D}{\to} \mathcal{N}\left(0, \frac{2\pi H_4}{H_2^2} \int_{\pi} w(\mu)^2 d\mu f(\lambda)^2\right)
\]

or equivalently

\[
\sqrt{Tb_T}(\log \hat{f}(T)(\lambda) - \log f(\lambda)) \overset{D}{\to} \mathcal{N}\left(0, \frac{2\pi H_4}{H_2^2} \int_{\pi} w(\mu)^2 d\mu\right)
\]
Lag-window estimators

**Motivation:** The periodogram can be written as

\[ I^{(T)}(\lambda) = \frac{1}{2\pi} \sum_{|u|<T} \hat{\gamma}(u) \exp(-i\lambda u) \]

**Heuristic:** \( \text{var}(\hat{\gamma}(u)) = O(1/T) \Rightarrow \text{var}(I^{(T)}(\lambda)) = O(1) \)

**Idea:** Take only \( \hat{\gamma}(u) \) for \( |u| \leq M \):

\[
\hat{f}^{(T)}(\lambda) = \frac{1}{2\pi} \sum_{|u|\leq M} \hat{\gamma}(u) \exp(-i\lambda u)
= \frac{1}{2\pi} \sum_{u \in \mathbb{Z}} W^{(M)}(u) \hat{\gamma}(u) \exp(-i\lambda u)
\]

where

\[
W^{(M)}(u) = \begin{cases} 
1 & \text{if } |u| \leq M \\
0 & \text{if } |u| > M 
\end{cases}
\]

\( \Rightarrow \text{var}(\hat{f}^{(T)}(\lambda)) = O(M/T). \)

Choose \( M = o(1/T) \) then \( \text{var}(\hat{\gamma}(\lambda)) = o(1). \)

**Relation to kernel spectral estimation:**

Define

\[
w^{(M)}(\lambda) = \frac{1}{2\pi} \int_{\Pi} W^{(M)}(u) e^{-i\lambda u} = \frac{\sin(\lambda(M + \frac{1}{2}))}{2\pi \sin(\lambda/2)}
\]

then

\[
\hat{f}^{(T)}(\lambda) = \int_{\Pi} w^{(M)}(\lambda - \mu) I^{(T)}(\mu) d\mu
\]

with \( \int w^{(M)}(\lambda) d\lambda = 1. \)

- \( W^{(M)} \) lag window
- \( w^{(M)} \) spectral window
Choice of window

- Daniell window (rectangular kernel)
  \[ w(\lambda) = \frac{1}{2M}1_{[-M,M]}(\lambda) \quad W(u) = \frac{\sin(u)}{u} \]

- Bartlett-Priestley window (quadratic kernel)
  \[ w(\lambda) = \frac{3}{4}(1 - \lambda^2)1_{[-1,1]}(\lambda) \quad W(u) = \frac{3}{u^2}\left(\frac{\sin(u)}{u} - \cos(u)\right) \]

- Dirichlet window (truncated periodogram)
  \[ w(\lambda) = \frac{1}{\pi} \frac{\sin(\lambda)}{\lambda} \quad W(u) = 1_{[-1,1]}(u) \]

- Bartlett window
  \[ w(\lambda) = \frac{1}{2\pi}\left(\frac{\sin(\lambda/2)}{\lambda/2}\right)^2 \quad W(u) = (1 - |u|)1_{[-1,1]}(u) \]

- Parzen window
  \[ w(\lambda) = \frac{1}{\pi}\left(\frac{\sin(\lambda/4)}{\lambda/4}\right)^4 \quad W(u) = \begin{cases} 1 - 6u^2 + 6|u|^3 & \text{if } |u| \leq 1/2 \\ 2(1 - |u|)^3 & \text{if } 1/2 < |u| \leq 1 \end{cases} \]

- Papoulis window
  \[ W(u) = \left[\frac{1}{\pi} |\sin(\pi u)| + (1 - |u|)\cos(\pi u)\right]1_{[-1,1]}(u) \]


Note: Much more important than the choice of the window or kernel is the choice of the bandwidth parameter.