1 Nonparametric spectral density estimation

Exact spectral density function is obtained from the true autocorrelation sequence by

\[ S_{gg}(\omega) = \Delta x \sum_{l=-\infty}^{\infty} r_{gg}(l)e^{-i2\pi \omega \Delta x} \quad |\omega| \leq 1/(2\Delta x) \]

True autocorrelation sequence is

\[ r_{gg}(l) = E[g(x)g(x + l)] \]

1.1 Empirical estimators of the autocorrelation sequence

An obvious estimator is

\[ \hat{r}_{gg}(l) = \frac{1}{N-|l|} \sum_{x=1}^{N-|l|} g(x)g(x+|l|) \quad |l| \leq (N - 1) \]

If the true mean is known, then \( \hat{r}_{gg}(l) \) is an unbiased estimator:

\[
E[\hat{r}_{gg}(l)] = \frac{1}{N-|l|} \sum_{x=1}^{N-|l|} E[g(x)g(x+|l|)] = \frac{1}{N-|l|} \sum_{x=1}^{N-|l|} r_{gg}(l) = r_{gg}(l)
\]

If, instead, the mean needs to be estimated, then \( \hat{r}_{gg}(l) \) is a biased estimator.

A more popular estimator (with “p” for periodogram) is

\[ \hat{r}_{gg}^p(l) = \frac{1}{N} \sum_{x=1}^{N-|l|} g(x)g(x+|l|) \quad |l| \leq (N - 1) \]

If the true mean is known, then \( \hat{r}_{gg}^p(l) \) is a biased estimator:

\[
E[\hat{r}_{gg}^p(l)] = \frac{1}{N} \sum_{x=1}^{N-|l|} r_{gg}(l) = \frac{N - |l|}{N} \cdot r_{gg}(l)
\]

whose bias increases with \( l \).
1.1.1 Advantages of $\hat{r}_{gg}^p(l)$

In real applications, the bias of $\hat{r}_{gg}^p(l)$ is not necessarily greater than that of $\hat{r}_{gg}^u(l)$, but:

the variance of $\hat{r}_{gg}^p(l)$ is much smaller: at the longest lag, the acvs needs to be estimated from a single pair of observations, and

$$\text{Var} (\hat{r}_{gg}^p(N - 1)) = \frac{1}{N^2} E[g(1)g(N)] = \frac{1}{N^2} \text{Var} (\hat{r}_{gg}^u(N - 1))$$

This generally leads to a lower mean square error

$$\text{mse} = E \left[ (\hat{r} - r)^2 \right] = E \left[ (\hat{r} - E[\hat{r}] + E[\hat{r}] - r)^2 \right]$$

$$= E \left[ (\hat{r} - E[\hat{r}])^2 \right] + E \left[ (E[\hat{r}] - r)^2 \right] + 2E \left[ (\hat{r} - E[\hat{r}]) (E[\hat{r}] - r) \right]$$

$$= E \left[ (\hat{r} - E[\hat{r}])^2 \right] + (E[\hat{r}] - r)^2 + 0$$

1.2 The periodogram

$$S_{gg}(\omega) = \Delta x \sum_{l=-\infty}^{\infty} r_{gg}(l)e^{-i2\pi\omega l \Delta x} \quad |\omega| \leq 1/(2\Delta x) = \omega_N$$

Replace $r_{gg}(l)$ with $\hat{r}_{gg}^p(l)$ and define $\hat{r}_{gg}^p(l) = 0 \forall |l| \geq N$ to obtain the periodogram

$$\hat{S}_{gg}^p(\omega) = \Delta x \sum_{l=-N}^{N-1} \hat{r}_{gg}^p(l)e^{-i2\pi\omega l \Delta x} \quad |\omega| \leq 1/(2\Delta x)$$

$$\hat{r}_{gg}^p(l) = \int_{-\omega_N}^{\omega_N} \hat{S}_{gg}^p(\omega)e^{i2\pi\omega l \Delta x} d\omega \quad |l| \leq N - 1$$
1.2.1 Bias of the periodogram

\[ E \left[ \hat{S}_{gg}(\omega) \right] = E \left[ \Delta x \sum_{l=-(N-1)}^{N-1} \hat{r}_{gg}^p(l) e^{-i2\pi \omega l \Delta x} \right] \]

\[ = \Delta x \sum_{l=-(N-1)}^{N-1} E \left[ \hat{r}_{gg}^p(l) \right] e^{-i2\pi \omega l \Delta x} \]

\[ = \Delta x \sum_{l=-(N-1)}^{N-1} \frac{N - |l|}{N} \cdot r_{gg}(l)e^{-i2\pi \omega l \Delta x} \]

\[ = \Delta x \sum_{l=-(N-1)}^{N-1} (\Pi_N \ast \Pi_N)(l) \cdot r_{gg}(l)e^{-i2\pi \omega l \Delta x} \]

\[ = \Delta x N \int_{-\omega N}^{\omega N} (D_N \cdot D_N) (\omega - \omega') S_{gg}(\omega')d\omega' \]

The expected value of the periodogram is a convolution of the true spectral density with the "Féjer-kernel" $D^2$, i.e. the square of the "Dirichlet kernel" $D$, a periodic version of the sinc.

The main lobe leads to blurring of the spectra, the side lobes to leakage (to prevent leakage of DC component → subtract mean).
1.2.2 Variance of the Periodogram

The variance of the periodogram does not decrease as $N \to \infty$! In essence, the reason is that the number of Fourier frequencies in the periodogram, $N$, grows with the length of data; no averaging takes place.

Example: Gaussian white noise process $g(x)$ with zero mean and variance $\sigma^2$.

$$G^*(\omega) = (\Delta x/N)^{1/2} \sum_{x=1}^{N} g(x) e^{i2\pi \omega x \Delta x}$$

$$\Re(G^*(\omega)) = (\Delta x/N)^{1/2} \sum_{x=1}^{N} g(x) \cos(2\pi \omega x \Delta x)$$

$$\Im(G^*(\omega)) = (\Delta x/N)^{1/2} \sum_{x=1}^{N} g(x) \sin(2\pi \omega x \Delta x)$$

$$\text{Var}(\Re(G^*(\omega))) = \frac{\sigma^2 \Delta x}{N} \sum_{x=1}^{N} \cos^2(2\pi \omega x \Delta x)$$

$$\text{Var}(\Im(G^*(\omega))) = \frac{\sigma^2 \Delta x}{N} \sum_{x=1}^{N} \sin^2(2\pi \omega x \Delta x)$$
\[
\text{Cov}(\Re(G^*(\omega)), \Im(G^*(\omega'))) = \frac{\Delta x}{N} \text{Cov}\left( \sum_{x=1}^{N} g(x) \cos(2\pi \omega x \Delta x), \sum_{x'=1}^{N} g(x') \sin(2\pi \omega' x' \Delta x) \right) \\
= \frac{\sigma^2 \Delta x}{N} \sum_{x=1}^{N} \cos(2\pi \omega x \Delta x) \sin(2\pi \omega' x \Delta x) \\
= 0 \quad \forall \omega, \omega'
\]

\[
\text{Cov}(\Re(G^*(\omega)), \Re(G^*(\omega'))) = 0 \quad \omega \neq \omega'
\]

\[
\text{Cov}(\Im(G^*(\omega)), \Im(G^*(\omega'))) = 0 \quad \omega \neq \omega'
\]

Since \(g(x)\) are Gaussian random variables, so are their linear combinations \(\Re(G^*(\omega))\) and \(\Im(G^*(\omega))\).

\(\Re(G^*(\omega))\) and \(\Im(G^*(\omega))\) are uncorrelated and Gaussian → independent.

Sum of \(n\) squared independent zero mean unit variance Gaussian variables is distributed according to \(\chi^2_n\).

With

\[
\hat{S}^p_{gg}(\omega) = \Re^2(G^*(\omega)) + \Im^2(G^*(\omega))
\]

we have

\[
\hat{S}^p_{gg}(\omega) \begin{cases} \\
\frac{\sigma^2 \Delta x}{2} \chi^2_2 & \omega \neq 0, \omega_N \\
(\sigma^2 \Delta x) \chi^2_1 & \omega = 0, \omega_N
\end{cases}
\]
\( \chi_n^2 \) with \( n \) degrees of freedom gives the distribution of \( \sum_{i=1}^{n} \mathcal{N}^2(0, 1) \)

\[
\begin{align*}
\chi_n^2 \quad &\text{with} \quad n \quad \text{degrees of freedom gives the distribution of} \quad \sum_{i=1}^{n} \mathcal{N}^2(0, 1) \\
&\text{With} \quad E[\chi_n^2] = n \quad \text{and} \quad \text{Var}(\chi_n^2) = 2n \quad \text{and} \quad \text{Var}(cX) = c^2\text{Var}(X)
\end{align*}
\]

\[
E \left[ \hat{S}_{gg}^p(\omega) \right] = \sigma^2 \Delta x \quad \forall \omega
\]

\[
\text{Var} \left( \hat{S}_{gg}^p(\omega) \right) = \begin{cases} 
\sigma^4(\Delta x)^2 & \omega \neq 0, \omega_N \\
2\sigma^4(\Delta x)^2 & \omega = 0, \omega_N 
\end{cases}
\]

the latter being independent of \( N! \)

Similar results can be proven for more general processes.

The Kolmogorov goodness-of-fit test for the cumulative periodogram of supposedly white noise is based on this result.
1.3 Fighting leakage: prewhitening

- Leakage becomes negligible only if the dynamic range is low.
- White noise has a constant spectrum and thus the lowest dynamic range. Idea: make spectrum look as if it came from white noise.
- Prewhitening filters seek to flatten the spectrum. The problem amounts to finding an inverse system, which requires knowledge of the system under investigation→ somewhat circular problem.
1.4 Fighting leakage: tapering

Multiply $g(x)$ with a taper $w(x)$, $w(x) = 0 \ \forall x < 1, x > N$ smoother than $\Pi$ to convolve the true spectrum with a kernel $W$ better than $D^2$ to give the “direct” spectral estimator

$$\hat{S}_{gg}(\omega) = \Delta x \left| \sum_{x=1}^{N} w(x)g(x)e^{-i2\pi \omega \Delta x} \right|^2$$

$$= \Delta x \sum_{l=-(N-1)}^{N-1} \sum_{x=1}^{N-|l|} g(x)w(x)g(x+|l|)w(x+|l|)e^{-i2\pi \omega l \Delta x}$$

$$E[\hat{S}_{gg}(\omega)] = \Delta x N \int_{-\omega_N}^{\omega_N} (W \cdot \hat{S}_{gg}(\omega'))d\omega'$$

Signal energy is conserved if $\sum_{x=1}^{N} w^2(x) = 1$.

---

*a taper – German Keil, Kegel; to taper off – German langsam auslaufen lassen

*b one with lower side lobes

---

[Graph showing box, bartlett, blackman, hamming, hann, kaiser 2.5 spectral estimators with tradeoff: size of side lobes vs. width of main lobe]
1.5 Fighting variance: smoothing the periodogram

If there is enough data such that the periodogram is approximately unbiased and if it is approximately constant in a local neighborhood, then the estimator

\[
\tilde{S}_{gg} (\omega) = \frac{1}{2M + 1} \sum_{\omega' = -M}^{M} \hat{S}_{gg}^p (\omega - \omega')
\]

has only variance

\[
\text{Var} \left( \tilde{S}_{gg} (\omega) \right) \approx \frac{1}{(2M + 1)^2} \sum_{\omega' = -M}^{M} \text{Var} \left( \hat{S}_{gg}^p (\omega - \omega') \right) = \frac{1}{2M + 1} \text{Var} \left( \hat{S}_{gg}^p (\omega) \right)
\]

Using the more general weights of a “smoothing window” \( W \) of width \( m \) in the summation above gives a “lag window” direct spectral estimator

\[
\hat{S}_{lw} (\omega) = \int_{-\omega_N}^{\omega_N} W_m (\omega - \omega') \tilde{S}_{gg} (\omega') d\omega'
\]

which can be implemented by a finite summation in the spatial domain:

\[
\hat{S}_{lw} (\omega) = \int_{-\omega_N}^{\omega_N} W_m (\omega - \omega') \left( \Delta x \sum_{x = - (N-1)}^{N-1} \hat{r}^d_{gg} (x) e^{-i2\pi \omega' x \Delta x} \right) d\omega'
\]

\[
= \Delta x \sum_{x = - (N-1)}^{N-1} \left( \int_{-\omega_N}^{\omega_N} W_m (\omega - \omega') e^{i2\pi (\omega - \omega') x \Delta x} d\omega' \right) \hat{r}^d_{gg} (x) e^{-i2\pi \omega x \Delta x}
\]

\[
= \Delta x \sum_{x = - (N-1)}^{N-1} \omega' (x) \hat{r}^d_{gg} (x) e^{-i2\pi \omega x \Delta x}
\]

with lag window \( w' = \mathcal{F}(W) \).
Since the last summation does not involve \( w'(x), \ |x| \geq N \), we can define

\[
    w_m(x) = \begin{cases} 
    w_m(x) & |x| < N \\
    0 & |x| \geq N 
\end{cases}
\]

and write

\[
\hat{S}^{lw}(\omega) = \Delta x \sum_{x=-(N-1)}^{N-1} w_m(x) \hat{i}^{d}_{gg}(x) e^{-i 2 \pi x \Delta x}
\]

1.5.1 Bias of lag window estimators

\[
E \left[ \hat{S}^{lw}_{gg}(\omega) \right] = \int_{-\omega_N}^{\omega_N} W_m(\omega - \omega') E \left[ \hat{S}^{d}_{gg}(\omega') \right] d\omega'
\]

\[
= \int_{-\omega_N}^{\omega_N} W_m(\omega - \omega') \int_{-\omega_N}^{\omega_N} \mathcal{W}(\omega' - \omega'') S_{gg}(\omega'') d\omega'' d\omega'
\]

\[
= \int_{-\omega_N}^{\omega_N} \int_{-\omega_N}^{\omega_N} W_m(\omega - \omega') \mathcal{W}(\omega' - \omega'') d\omega' S_{gg}(\omega'') d\omega''
\]

\[
= \int_{-\omega_N}^{\omega_N} \mathcal{U}_m(\omega - \omega'') S_{gg}(\omega'') d\omega''
\]

with the “spectral window” \( \mathcal{U}_m(\omega) \)

The wider the smoothing window \( W(\omega) \), the more spectral resolution is lost.
1.5.2 Variance of lag window estimators

Large sample approximation, i.e. $N$ large, bandwidth small such that $W(\omega) \approx 0 \forall |\omega| > \omega'(J) := J/(N'\Delta x)$ where $N' < N$ is chosen such that $\hat{S}^d_{gg}(\omega)$ is almost uncorrelated at a distance $\omega'(k) := k/(N'\Delta x)$

\[
\hat{S}^\text{lw}(\omega) = \int_{-\omega_N}^{\omega_N} W_m(\omega')\hat{S}^d_{gg}(\omega + \omega')d\omega' \\
\approx \int_{-\omega'(J)}^{\omega'(J)} W_m(\omega')\hat{S}^d_{gg}(\omega' + \omega'(k))d\omega' \\
\approx \sum_{j=-J}^{J} W_m(\omega'(j))\hat{S}^d_{gg}(\omega'(j) + \omega'(k)) \frac{1}{N'\Delta x}
\]

where the integral is approximated by a Riemann sum. Expression now depends on $N!$

\[
\text{Var}\left(\hat{S}^\text{lw}(\omega'(k))\right) \approx \frac{1}{(N'\Delta x)^2} \text{Var}\left(\sum_{j=-J}^{J} W_m(\omega'(j))\hat{S}^d_{gg}(\omega'(j + k))\right) \\
\approx \frac{1}{(N'\Delta x)^2} \sum_{j=-J}^{J} W_m^2(\omega'(j))S^2_{gg}(\omega'(j + k))
\]

by assuming that the spectral density is uncorrelated at the frequencies $\omega'(k), k \in \mathbb{Z}$ and using a large-sample result

\[
\approx \frac{1}{(N'\Delta x)^2} S^2_{gg}(\omega'(k)) \sum_{j=-J}^{J} W_m^2(\omega'(j))
\]

based on a smoothness assumption

\[
\approx \frac{1}{(N'\Delta x)^2} S^2_{gg}(\omega'(k)) \sum_{j=-N'/2}^{N'/2} W_m^2(\omega'(j))
\]

again assuming that the window is approximately zero beyond $\omega'(J)$
Overall,

\[
\text{Var} \left( \hat{S}'w(\omega'(k)) \right) \simeq \frac{1}{(N' \Delta x)^2} \text{Var}\left( S^{2}_{gg}(\omega'(k)) \right) \int_{-\omega_{m}}^{\omega_{N}} W_{m}^{2}(\omega) d\omega
\]

variance decreases with growing window width.