Chapter 2: Non-parametric spectral estimation

PSD definitions and properties

Assume a **wide sense stationary random sequence** $y(n)$.

Denote $r(k) = R_Y(k)$ its **autocorrelation sequence**, which can be defined more generally as

$$r(k) = \mathrm{E}\{y(n)y^*(n-k)\}. \tag{5}$$

Note that $r(k)$ becomes the **autocovariance sequence (ACS)** if the process has zero mean.
Let $S(f)$ be its **PSD**.

Our goal:

Spectral estimation problem

From a finite length record $\{y(n)\}_{n=0}^{N-1}$ of $y(n)$, determine an estimate $\hat{S}(\omega)$ of its power spectral density $S(\omega)$.

Non-parametric spectral estimation is based on the definitions of PSD:

First definition of the PSD: from the Wiener-Khinchin Theorem

$$S(\omega) = \mathcal{F}\{r(k)\} = \sum_{k=-\infty}^{+\infty} r(k)\mathrm{e}^{-\jmath\omega k} \tag{6}$$

cf. eqn. (3) — note that $S(\omega)$ and $S(f)$ are related through the change of variable $\omega = 2\pi f$.

PSD definitions and properties (cont'd)

Second definition of the PSD: exploiting ergodicity

$$S(\omega) = \lim_{N\to\infty} \mathrm{E}\left\{ \frac{1}{N} \left| \sum_{n=0}^{N-1} y(n) e^{-j\omega n} \right|^2 \right\} \tag{7}$$

This second definition of the PSD is based on the fact that, for a wide sense ergodic process:

$$r(k) = \lim_{N\to\infty} \mathrm{E}\left\{ \frac{1}{N} \sum_{n=0}^{N-1} y(n) y^*(n-k) \right\}$$

and is valid under the condition that the sequence $r(k)$ decays sufficiently rapidly:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=-N}^{N} |k||r(k)| = 0.$$

These expressions will be justified in tutorial.

Properties of the PSD

- $S(\omega)$ is real valued and $S(\omega) \geq 0, \forall\omega \in \mathbb{R}$.
- $S(\omega + 2\pi) = S(\omega), \forall\omega \in \mathbb{R} \;\Rightarrow$ we can restrict our attention to $\omega \in [-\pi, \pi[$ or $f \in [-\frac{1}{2}, \frac{1}{2}[$.
- $S(\omega) = S(-\omega)$ if $y(n) \in \mathbb{R}$. Otherwise, $S(\omega) \neq S(-\omega)$ in general.

Periodogram and correlogram methods

By dropping the expectation and the limit in 2nd PSD definition (7), we can naturally define the

Periodogram

$$\hat{S}_{\mathrm{P}}(\omega) = \frac{1}{N}\left|\sum_{n=0}^{N-1} y(n)\mathrm{e}^{-\jmath\omega n}\right|^2 = \frac{1}{N}|Y_N(\omega)|^2 \tag{8}$$

where $Y_N(\omega) = \mathcal{F}\{y_N(n)\}$ is the DTFT of the N-point data sequence:

$$y_N(n) = \begin{cases} y(n) & 0 \leq n \leq (N-1) \\ 0 & \text{otherwise.} \end{cases}$$

By truncating the sum and using an ACS estimate in 1st PSD definition (6), we can define the

Correlogram

$$\hat{S}_{\mathrm{C}}(\omega) = \sum_{k=-(N-1)}^{N-1} \hat{r}(k)\mathrm{e}^{-\jmath\omega k} \tag{9}$$

where $\hat{r}(k)$ is an estimate of the covariance $r(k)$ at lag k, from $y_N(n)$.

Estimation of the autocovariance sequence (ACS)

Unbiased ACS estimator

$$\hat{r}(k) = \frac{1}{N-k} \sum_{n=k}^{N-1} y(n) y^*(n-k), \qquad 0 \le k \le N-1 \tag{10}$$

Unbiased because $\mathrm{E}\{\hat{r}(k)\} = r(k)$.

Biased ACS estimator

$$\hat{r}(k) = \frac{1}{N} \sum_{n=k}^{N-1} y(n) y^*(n-k), \qquad 0 \le k \le N-1 \tag{11}$$

Biased because $\mathrm{E}\{\hat{r}(k)\} = \left(1 - \frac{|k|}{N}\right) r(k)$.

In both cases, we constrain $\hat{r}(k) = \hat{r}^*(-k)$ for $k < 0$.

The **biased ACS estimator is preferred** because:

- it presents reduced variance for large k (take, e.g., $k = N-1$)
- it is guaranteed to be semi-definite positive (important when introducing the covariance matrix; it will be recalled in [Chap. 3])
- the estimation error in $\hat{r}(k)$ is on the order of $\frac{1}{\sqrt{N}}$.

Link between periodogram and correlogram

Link between $\hat{S}_{\mathrm{C}}(\omega)$ and $\hat{S}_{\mathrm{P}}(\omega)$

If the biased ACS estimate of $r(k)$ is used in $\hat{S}_{\mathrm{C}}(\omega)$, then:

$$\hat{S}_{\mathrm{C}}(\omega) = \hat{S}_{\mathrm{P}}(\omega).$$

Proof: in tutorial.

Consequence: both can be analyzed simultaneously.

Spectral estimators are random processes

Each N-sample realization of $y(n)$ will provide a (generally) different spectral estimate.

Hence, $\hat{S}_{\mathrm{C}}(\omega)$, $\hat{S}_{\mathrm{P}}(\omega)$ and in general any spectral estimator can be considered as random processes that are functions of frequency ω.

Consequence: spectral estimation quality can be measured by means of performance indices (statistics) such as mean, variance and mean square error.

Performance analysis

Parameter estimation accuracy: mean square error (MSE), bias, variance

Let $\hat{\theta}$ be an estimator of a quantity $\theta \in \mathbb{R}$. The estimator's **mean square error (MSE)** is given by

$$\mathrm{MSE}\{\hat{\theta}\} \stackrel{\mathrm{def}}{=} \mathrm{E}\{(\hat{\theta}-\theta)^2\} = (\mathrm{E}\{\hat{\theta}\}-\theta)^2 + \mathrm{E}\{(\hat{\theta}-\mathrm{E}\{\hat{\theta}\})^2\} = \mathrm{bias}\{\hat{\theta}\}^2 + \mathrm{var}\{\hat{\theta}\}$$

where $\mathrm{bias}\{\hat{\theta}\} = (\mathrm{E}\{\hat{\theta}\}-\theta)$ and $\mathrm{var}\{\hat{\theta}\} = \mathrm{E}\{(\hat{\theta}-\mathrm{E}\{\hat{\theta}\})^2\}$.

Extension to complex-valued case is straightforward by replacing $|\cdot|^2$ for $(\cdot)^2$.

Unbiasedness

Let $\hat{\theta}_N$ be an estimator of a quantity θ computed from a set of N samples of available data. The estimator is said to be **unbiased** if $\mathrm{E}\{\hat{\theta}_N\} = \theta$.

The estimator is **asymptotically unbiased** if $\lim_{N\to\infty} \mathrm{E}\{\hat{\theta}_N\} = \theta$.

Consistency

An estimator $\hat{\theta}$ is said to be **consistent** if $\lim_{N\to\infty} \mathrm{var}\{\hat{\theta}_N\} = 0$.

The estimator is **MSE consistent** if $\lim_{N\to\infty} \mathrm{MSE}\{\hat{\theta}_N\} = 0$.

MSE consistency implies variance consistency and asymptotic unbiasedness.

Performance analysis — bias of the periodogram

From the link between $\hat{S}_{\mathrm{P}}(\omega)$ and $\hat{S}_{\mathrm{C}}(\omega)$, and the definition of correlogram (9):

$$\mathrm{E}\{\hat{S}_{\mathrm{P}}(\omega)\} = \mathrm{E}\{\hat{S}_{\mathrm{C}}(\omega)\} = \sum_{k=-(N-1)}^{N-1} \mathrm{E}\{\hat{r}(k)\}\mathrm{e}^{-\jmath\omega k}.$$

From the definition of the biased ACS estimate (11):

$$\mathrm{E}\{\hat{S}_{\mathrm{P}}(\omega)\} = \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N}\right) r(k)\mathrm{e}^{-\jmath\omega k}.$$

Define the **Bartlett or triangular window**:

$$w_{\mathrm{B}}(k) = \begin{cases} 1 - \frac{|k|}{N} & 0 \leq |k| \leq N-1 \\ 0 & \text{otherwise.} \end{cases}$$

Then:

$$\mathrm{E}\{\hat{S}_{\mathrm{P}}(\omega)\} = \sum_{k=-\infty}^{+\infty} w_{\mathrm{B}}(k) r(k) e^{-j\omega k}.$$

DTFT of a product $\Leftrightarrow$ convolution of the true PSD and Bartlett window in the frequency domain:

$$\mathrm{E}\{\hat{S}_{\mathrm{P}}(\omega)\} = S(\omega) * W_{\mathrm{B}}(\omega) \overset{\text{def}}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi} S(\xi) W_{\mathrm{B}}(\omega - \xi) d\xi.$$

Performance analysis — bias of the periodogram (cont'd)

The mean of the estimator is the convolution of the actual PSD $S(\omega)$ and the Bartlett kernel:

$$W_{\mathrm{B}}(f) = \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N}\right) \mathrm{e}^{-\jmath\omega k} = \frac{1}{N}\left[\frac{\sin(\omega N/2)}{\sin(\omega/2)}\right]^2$$

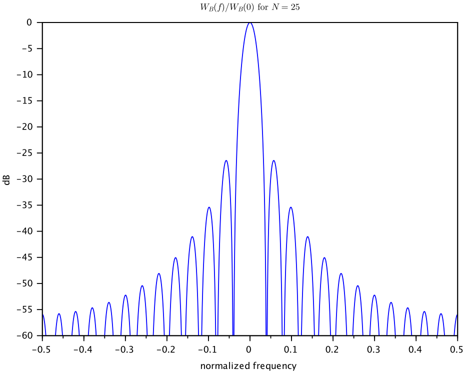

Figure: Normalized DTFT of the Bartlett window, $W_{\mathrm{B}}(f)/W_{\mathrm{B}}(0)$, for $N = 25$ samples.

Performance analysis — bias of the periodogram (cont'd)

Ideally: $\mathrm{E}\{\hat{S}_{\mathrm{P}}(\omega)\}$ should be as close as possible to $S(\omega) \Leftrightarrow W_{\mathrm{B}}(\omega)$ should be as close as possible to Dirac's impulse $\delta(\omega)$.

Asymptotically:

$$\lim_{N\to\infty} W_{\mathrm{B}}(\omega) = \delta(\omega) \quad \Rightarrow \quad \lim_{N\to\infty} \mathrm{E}\{\hat{S}_{\mathrm{P}}(\omega)\} = S(\omega).$$

The periodogram is asymptotically unbiased estimator of the PSD (even though implicitly computed from the biased ACS estimate).

In practice, for finite sample size N, **two side effects**:

Spectral smearing (or smoothing): caused by main lobe of $W_{\mathrm{B}}(\omega)$

The half-power bandwidth of $W_{\mathrm{B}}(\omega)$ can be shown to be approximately $f_{3\mathrm{dB}} \simeq 1/N$.

This is the **fundamental resolution of the periodogram**: $\hat{S}_{\mathrm{P}}(\omega)$ cannot resolve frequency components spaced by less than $\Delta f_{\min} \simeq 1/N$ sample^{-1} (or $\Delta\omega_{\min} \simeq 2\pi/N$ rad/sample).

Power leakage: caused by sidelobes of $W_{\mathrm{B}}(\omega)$

A frequency component present at f_0 will leak at frequencies $f_0 + p/N$, $p \in \mathbb{Z}$.

Performance analysis — variance of the periodogram

Variance of the periodogram

For N sufficiently large, we have:

$$\mathrm{var}\{\hat{S}_{\mathrm{P}}(\omega)\} \approx S^2(\omega).$$

Hence, **the periodogram is an inconsistent estimator of the PSD**.

Consequences:

- Variance cannot be improved by increasing the sample size N.
- Inconsistency — as spectral smearing — has an adverse effect on resolvability properties.

These results will be illustrated in a computer lab.

Performance analysis — bias of the correlogram with unbiased ACS

What happens if we use the unbiased ACS estimator (10)**?**

$$\mathrm{E}\{\hat{r}(k)\} = r(k), \qquad k \geq 0.$$

The expected correlogram (9) becomes

$$\mathrm{E}\{\hat{S}_\mathrm{C}(\omega)\} = \sum_{k=-(N-1)}^{N-1} r(k)\mathrm{e}^{-\jmath\omega k} = \sum_{k=-\infty}^{+\infty} w_\mathrm{R}(k) r(k) \mathrm{e}^{-\jmath\omega k}$$

with

$$w_\mathrm{R}(k) = \begin{cases} 1 & k = -(N-1), \dots, 0, \dots, N-1 \\ 0 & \text{otherwise.} \end{cases}$$

$w_\mathrm{R}(k)$ is a **rectangular window**, with DTFT:

$$W_\mathrm{R}(\omega) = \frac{\sin[(2N-1)\omega/2]}{\sin[\omega/2]}.$$

Performance analysis — bias of the correlogram with unbiased ACS

- **Positive effect**: narrower main lobe $\rightarrow$ narrower smearing $\rightarrow$ improved frequency resolution.
- **Negative effect**: more prominent sidelobes $\rightarrow$ increased spectral leakage.

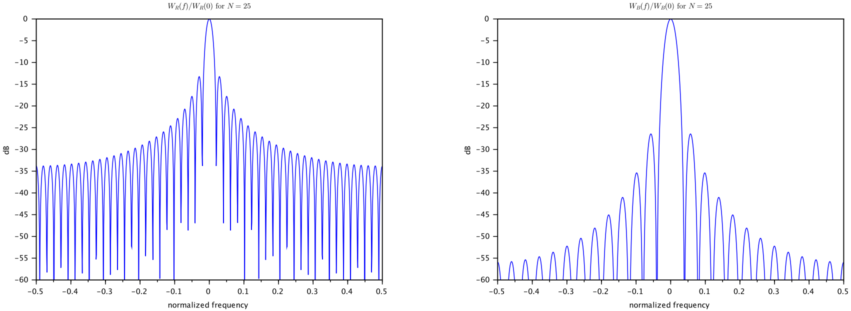

Figure: Normalized DTFT of the rectangular window (left) and Bartlett window (right), for $N = 25$ samples.

Improved periodogram-based methods

Motivation

- Periodogram: based on the FFT $\rightarrow$ simple, computationally efficient PSD estimate.
- Asymptotically unbiased but inconsistent estimator: variance does not improve with sample size N.
- Periodogram variance mainly caused by ACS estimates with large $|k|$, which have large variance due to small number of contributing samples.
- To improve variance, two main families of approaches:
  - **periodogram smoothing**: Blackman-Tuckey, Daniell
  - **periodogram averaging**: Bartlett, Welch.