Stochastic Processes

Session 7 — Lecture

Troels Pedersen and Carles Navarro Manchón

Section Wireless Communication Networks,
Department of Electronic Systems, Aalborg University

Fall 2016

Outline for Session 7 — Lecture

Power spectral density

Estimators — definition and associated terms

Estimation of autocorrelation and power spectrum

ILOs for Session 7

After having attended this lecture and solved the exercises you should be able to:

- Explain the definition and the meaning of a PSD to a fellow student.
- Compute the PSD given a particular ACF.
- Use the theoretical properties of any PSD as sanity checks of your derivations.
- Know (without hesitation and computation) the PSD of a white process.
- Know the definition of bias and MSE of an estimator and explain their meaning.

Outline for Session 7 — Lecture

Power spectral density

Estimators — definition and associated terms

Estimation of autocorrelation and power spectrum

Discrete-time versus continuous-time

In this lecture we need to distinguish whether a random process is defined with respect to *discrete-time* or *continuous-time*, i.e. whether $\mathbb{T} \subseteq \mathbb{Z}$ or $\mathbb{T} \subseteq \mathbb{R}$.

Discrete-time notation:
Let $\{X_n\}$ denote a real-valued discrete-time WSS random process with autocorrelation function $R_X(k) = \mathbb{E}\left[X_n X_{n+k}\right]$.

Continuous-time notation:
Let $X(t)$ denote a real-valued continuous-time WSS random process with autocorrelation function $R_X(\tau) = \mathbb{E}\left[X(t)X(t+\tau)\right]$.

Power spectral density (PSD) or power spectrum

Definition: (discrete-time)
The *power spectral density* (PSD) of the process $\{X_n\}$ is the discrete-time Fourier transform of the autocorrelation function $R_X(k)$, i.e.

$$S_X(f) := \mathcal{F}\{R_X\}(f) = \sum_{k=-\infty}^{\infty} R_X(k)\exp(-j2\pi kf), \qquad |f| \leq \frac{1}{2}$$

Definition: (continuous-time)
The *power spectral density* (PSD) of the process $X(t)$ is the continuous-time Fourier transform of the autocorrelation function $R_X(\tau)$, i.e.

$$S_X(f) := \mathcal{F}\{R_X\}(f) = \int_{-\infty}^{\infty} R_X(\tau)\exp(-j2\pi\tau f)\mathrm{d}\tau, \qquad f \in \mathbb{R}$$

Remarks

The PSD $S_x(f)$ of a discrete-time random process is *periodic* in f with unit period since

$$\exp\big(-j2\pi k(f+1)\big) = \exp(-j2\pi kf), \qquad \forall k \in \mathbb{Z},\ \forall f \in \mathbb{R}$$

Hence, $S_x(f)$ is completely specified on any connected interval of unit length.

By convention, we use the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$

The PSD $S_x(f)$ of a continuous-time random process is *not periodic* in f.

Inverse transforms

Discrete-time:

$$R_x(k) = \mathcal{F}^{-1}\{S_x\}(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_x(f)\exp(j2\pi kf)\mathrm{d}f, \qquad k \in \mathbb{Z}$$

Continuous-time:

$$R_x(\tau) = \mathcal{F}^{-1}\{S_x\}(\tau) = \int_{-\infty}^{\infty} S_x(f)\exp(j2\pi\tau f)\mathrm{d}f, \qquad \tau \in \mathbb{R}$$

Interpretation: Distribution of average power (in frequency)

The mean square $\mathbb{E}[X_n^2]$ can be interpreted as the average power of a process.

The PSD can be related to the average power: *(Show this!)*

Discrete-time:

$$R_x(0) = \mathbb{E}\left[X_n^2\right] = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_x(f)\mathrm{d}f$$

Continuous-time:

$$R_x(0) = \mathbb{E}\left[X^2(t)\right] = \int_{-\infty}^{\infty} S_x(f)\mathrm{d}f$$

Properties of $S_x(f)$

Recall that $\{X_n\}$ and $X(t)$ are real-valued WSS random processes.
The properties below are valid no matter if time is discrete or continuous.

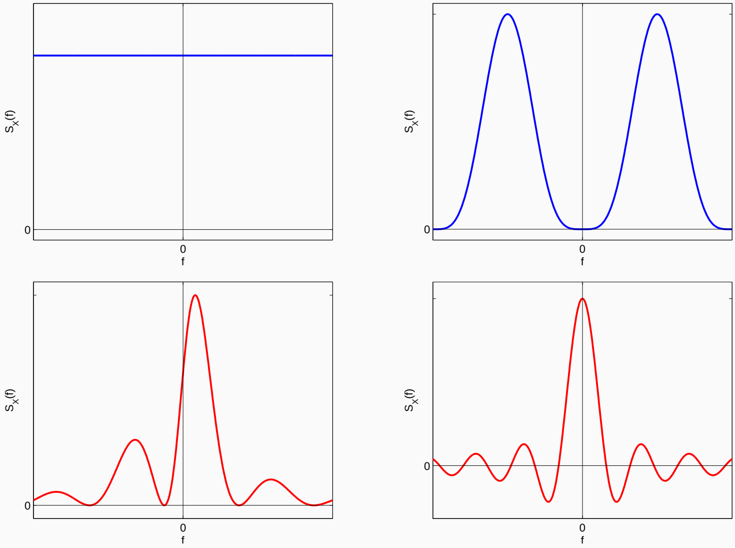
Properties:

1) $S_x(f) \in \mathbb{R}$ for all $f \in \mathbb{R}$
2) $S_x(-f) = S_x(f)$ for all $f \in \mathbb{R}$
3) $S_x(f) \geq 0$ for all $f \in \mathbb{R}$

In plain words this means that $S_x(f)$:

1) is real-valued (even though the Fourier transform is a complex operation)
2) is an even function
3) is not just real-valued but in fact non-negative

Graphical examples (and counterexamples)



White processes

Motivated by the observation that "white light" has constant PSD, we define a white process:

Definition: A discrete-time[1] process $\{X(n)\}$ is *white* if it is WSS, and its PSD is constant, i.e.

$$S_X(f) = \begin{cases} \sigma_X^2, & |f| < \frac{1}{2} \\ 0, & |f| > \frac{1}{2} \end{cases}$$

Correspondingly, the ACF reads

$$R_X(k) = \sigma_X^2 \delta(k)$$

where $\delta(k)$ is the Kronecker delta.

Example:
Let $\{X_n\}$ be an (discrete time) i.i.d. process with zero mean and variance σ_X^2. The autocorrelation function is $R_X(k) = \mathbb{E}[X_n X_{n+k}] = \sigma_X^2 \delta(k)$ and the spectrum is $S_X(f) = \sigma_X^2, |f| \leq \frac{1}{2}$.

Show that the mean of a white process is zero.

[1] A similar definition applies to continuous time processes. Replace the frequency interval by the real line and the $\delta(k)$ by Dirac's delta.

Outline for Session 7 — Lecture

Power spectral density

Estimators — definition and associated terms

Estimation of autocorrelation and power spectrum

Estimation problem

Definition:
An *estimator* is a function $g(\cdot)$ of data X (i.e. a *statistic*) used to infer on the value of an unknown quantity θ. The value $\hat{\theta} = g(X)$ is called an *estimate*.

We mark an estimate by a "hat" over the quantity we estimate.

Example: Mean estimation
Let $\{X_k\}$ be a WSS process with unknown mean μ_X. We estimate μ_X from the observed data vector $\mathbf{X} = [X_1, X_2, \ldots, X_N]^T$ as

$$\hat{\mu}_X = g(\mathbf{X}) := \frac{1}{N}\sum_{n=1}^{N} X_n = \frac{1}{N}\mathbf{1}^T\mathbf{X},$$

where $\mathbf{1} = [\underbrace{1, 1, \ldots, 1}_{N \text{ entries}}]^T$.

Biased and unbiased estimators, mean square error

The estimate $\hat{\theta}$ is a function $g(X)$ of random data X and is therefore random.

Estimator bias: $\text{bias}_{\hat{\theta}} = \mathbb{E}[\hat{\theta} - \theta]$
An estimator is *unbiased* if $\text{bias}_{\hat{\theta}} = 0$, i.e. if $\mathbb{E}[\hat{\theta}] = \mathbb{E}[\theta]$.

Mean square error (MSE): $\text{MSE}_{\hat{\theta}} = \mathbb{E}[(\hat{\theta} - \theta)^2]$

Example: Mean estimation (cont.)

$$\mathbb{E}[\hat{\mu}_X - \mu_X] = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} X_n - \mu_X\right] = \frac{1}{N}\underbrace{\sum_{n=1}^{N}\mathbb{E}[X_n]}_{N\cdot\mu_X} - \mu_X = 0$$

Thus $\hat{\mu}_X$ is unbiased.

Mean square error for mean estimation (contd.)

The MSE of μ_X: Observe first that $\hat{\mu}_X - \mu_X = \frac{1}{N}\mathbf{1}^T(\mathbf{X} - \mathbf{1}\mu_X)$, then

$$\begin{aligned}\mathbb{E}[(\hat{\mu}_X - \mu_X)^2] &= \mathbb{E}\left[\frac{1}{N^2}\mathbf{1}^T(\mathbf{X} - \mathbf{1}\mu_X)(\mathbf{X} - \mathbf{1}\mu_X)^T\mathbf{1}\right] \\ &= \frac{1}{N^2}\mathbf{1}^T\mathbb{E}\left[(\mathbf{X} - \mathbf{1}\mu_X)(\mathbf{X} - \mathbf{1}\mu_X)^T\right]\mathbf{1} = \frac{1}{N^2}\mathbf{1}^T\mathbf{C_X}\mathbf{1}.\end{aligned}$$

where $\mathbf{1}^T\mathbf{C_X}\mathbf{1}$ is the sum of all entries of the covariance matrix $\mathbf{C_X}$:

$$\mathbf{C_X} = \begin{bmatrix} C_X(0) & C_X(1) & \dots & C_X(N-1) \\ C_X(-1) & C_X(0) & \dots & C_X(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(-(N-1)) & C_X(-(N-2)) & \dots & C_X(0) \end{bmatrix}$$

Due to the special structure of $\mathbf{C_X}$ we achieve

$$\mathbb{E}[(\hat{\mu}_X - \mu_X)^2] = \frac{1}{N^2}\sum_{k=-(N-1)}^{N-1}(N - |k|) \cdot C_X(k).$$

Special case: For *uncorrelated* $\{X_k\}$ the MSE is $\mathbb{E}[(\hat{\mu}_X - \mu_X)^2] = \frac{\sigma_X^2}{N}$.

Outline for Session 7 — Lecture

Power spectral density

Estimators — definition and associated terms

Estimation of autocorrelation and power spectrum

The spectral estimation problem

Autocorrelation and Spectrum Estimation problem:
Estimate $R_X(k)$ and/or $S_X(f)$ of a discrete-time WSS process $\{X_k\}$ from $\mathbf{X} = [X_0, \ldots, X_{N-1}]^T$.

I.e. find functions $\hat{R}_X(k) = r_k(\mathbf{X})$ and $\hat{S}_X(f) = s_f(\mathbf{X})$.

We will approach this problem by first estimating the autocorrelation and then obtain an estimate of the spectrum.

Here we give only a brief introduction. Details in the lecture notes [LN - Section 6.1].

Sample autocorrelation function

Given a N point sample $X(1), \ldots, X(N)$ of a discrete-time stochastic process $\{X(n)\}$ we can estimate the ACF via the *sample autocorrelation function* defined as:

$$\hat{R}_X(k) = \begin{cases} \dfrac{1}{N} \displaystyle\sum_{n=1}^{N-k} X(n)X(n+k), & k = 0, 1, \ldots, N-1 \\ \hat{R}_X(-k), & k = -(N-1), \ldots, -1 \\ 0, & |k| \geq N. \end{cases}$$

Alternatively, we can express the sample ACF in form of a convolution of a signal $X_{obs}(n)$ with its time-reverse:

$$\hat{R}_X(k) = \frac{1}{N} X_{obs}(n) * X_{obs}(-n) \qquad X_{obs}(n) = \begin{cases} X(n), & n = 1, \ldots, N \\ 0, & \text{otherwise} \end{cases}$$

where $*$ denotes the convolution operator.

Estimation of the PSD using the periodogram

We define the *periodogram* as the Fourier transform of the sample ACF:

$$\hat{S}_X(f) = \mathcal{F}\{\hat{R}_X\}(f) = \sum_{k=-(N-1)}^{N-1} \hat{R}_X(k)\exp(-j2\pi kf), \qquad f \in [-\frac{1}{2}, \frac{1}{2}]$$

The periodogram has a simple form useful for numerical implementation:

$$\begin{aligned}
\hat{S}_X(f) &= \mathcal{F}\{\tfrac{1}{N}X_{obs}(k) * X_{obs}(-k)\}(f) \\
&= \tfrac{1}{N}\mathcal{F}\{X_{obs}(k)\}(f) \cdot \mathcal{F}\{X_{obs}(-k)\}(f) \\
&= \tfrac{1}{N}\mathcal{F}\{X_{obs}(k)\}(f) \cdot \mathcal{F}\{X_{obs}(k)\}(f)^* \\
&= \tfrac{1}{N}\left|\mathcal{F}\{X_{obs}(k)\}(f)\right|^2
\end{aligned}$$

For discrete frequencies $f = \frac{m}{N}, m = 0, \ldots, N-1$, the periodogram may be computed in Matlab by a one-liner:

```
Sx = abs(fft(X)).^2;
```

Bias of the the sample autocorrelation function

The sample ACF is a *biased* estimator of $R_X(k)$ which can be shown by taking the expectation of $\hat{R}_X(k)$ for $k = 0, \ldots, N$:

$$\mathbb{E}[\hat{R}_X(k)] = \frac{1}{N} \sum_{n=1}^{N-k} \mathbb{E}[X(n)X(n+k)] = \frac{N-k}{N} \cdot R_X(k).$$

For general k, we have $\mathbb{E}[\hat{R}_X(k)] = w_B(k) \cdot R_X(k)$ where $w_B(k)$ is called the *Bartlett window*:

$$w_B(k) = \begin{cases} \frac{N-|k|}{N}, & |k| \leq N \\ 0, & \mathit{otherwise} \end{cases}$$

Since $\mathbb{E}[\hat{R}_X(k)] \neq R_X(k)$, we conclude that the sample ACF is biased.

Bias of the periodogram

Also the periodogram is biased:

$$\begin{aligned}\mathbb{E}[\hat{S}_X(k)] &= \mathbb{E}[\mathcal{F}\{\hat{R}_X(k)\}(f)] \\ &= \mathcal{F}\{\mathbb{E}[\hat{R}_X(k)]\}(f) \\ &= \mathcal{F}\{w_B(k) \cdot R_X(k)\}(f) \\ &= \mathcal{F}\{w_B(k)\}(f) * S_X(f)\end{aligned}$$

Thus the spectrum is on average smeared by the Fourier transform of the Bartlett window:

$$\mathcal{F}\{w_B(k)\}(f) = \left(\frac{sin(\pi f N)}{sin(\pi f)}\right)^2$$

which is sometimes called the “Fejér kernel”.

Unbiased sample autocorrelation function

In an attempt to "repair" the sample ACF and the periodogram, we may look at the *unbiased sample ACF*:

$$
\begin{aligned}
\check{R}_X(k) &= \frac{\hat{R}_X(k)}{w_B(k)} \\
&= \begin{cases} \dfrac{1}{N-k} \displaystyle\sum_{n=1}^{N-k} X(n)X(n+k), & k = 0, 1, \ldots, N-1 \\ \check{R}_X(-k), & k = -(N-1), \ldots, -1 \\ 0, & |k| \geq N. \end{cases}
\end{aligned}
$$

It is straightforward to check that $\check{R}_X(k)$ is unbiased for $|k| \leq N-1$:

$$
\mathbb{E}[\check{R}_X(k)] = w_r(k) \cdot R_X(k)
$$

where $w_r(k)$ is the *rectangular window* defined as

$$
w_r(k) = \begin{cases} 1, & |k| \leq N-1 \\ 0, & \text{otherwise.} \end{cases}
$$

Caveat: The unbiased sample ACF is essentially worthless for large time-lags $|k|$ because it exhibits high variance!

PSD estimator for the unbiased sample ACF

Similar to the periodogram, we can define a PSD estimator based un the unbiased ACF as

$$\check{S}_X(f) = \mathcal{F}\{\check{R}_X(k)\}(f)$$

This does not yield an unbiased spectral estimate of the PSD:

$$\begin{aligned}
\mathbb{E}[S_X(f)] &= \mathbb{E}[\mathcal{F}\{\check{R}_X(k)\}(f)] \\
&= \mathcal{F}\{\mathbb{E}[\check{R}_X(k)]\}(f) \\
&= \mathcal{F}\{w_r(k) \cdot R_X(k)\}(f) \\
&= \mathcal{F}\{w_r(k)\}(f) * S_X(f)
\end{aligned}$$

The discrete Fourier transform of a rectangular function (called the Dirichlet kernel) reads

$$\mathcal{F}\{w_r(k)\}(f) = \frac{\sin(2\pi f N)}{\sin(\pi f)}$$

The Dirichlet kernel is negative for some f, and gives rise to negative PSD estimates at some frequencies. What a high price to pay for an unbiased ACF estimate!