Topic 8: Power spectral density and LTI systems

- The power spectral density of a WSS random process
- Response of an LTI system to random signals
- Linear MSE estimation

The autocorrelation function and the rate of change

- Consider a WSS random process $X(t)$ with the autocorrelation function $R_X(\tau)$.
- If $R_X(\tau)$ drops quickly with τ, then process $X(t)$ changes quickly with time: its time samples become uncorrelated over a short period of time.
  - Conversely, when $R_X(\tau)$ drops slowly with τ, samples are highly correlated over a long time.
- Thus $R_X(\tau)$ is a measure of the rate of change of $X(t)$ with time and hence is related to the *frequency response* of $X(t)$.
  - For example, a sinusoidal waveform $\sin(2\pi f t)$ will vary rapidly with time if it is at high frequency (large f), and vary slowly at low frequency (small f).
- In fact, the Fourier transform of $R_X(\tau)$ is the average power density over the frequency domain.

The power spectral density of a WSS process

- The *power spectral density* (psd) of a WSS random process $X(t)$ is given by the Fourier transform (FT) of its autocorrelation function

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f\tau} d\tau$$

- For a discrete-time process X_n, the psd is given by the discrete-time FT (DTFT) of its autocorrelation sequence

$$S_x(f) = \sum_{n=-\infty}^{n=\infty} R_x(n) e^{-j2\pi fn}\ , \quad -\frac{1}{2} \leq f \leq \frac{1}{2}$$

  Since the DTFT is periodic in f with period 1, we only need to consider $|f| \leq \frac{1}{2}$.

- $R_X(\tau)$ ($R_x(n)$) can be recovered from $S_x(f)$ by taking the inverse FT

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f\tau} df\ , \quad R_X(n) = \int_{-1/2}^{1/2} S_X(f) e^{j2\pi fn} df$$

Properties of the power spectral density

- $S_X(f)$ is real and even

$$S_X(f) = S_X(-f)$$

- The area under $S_X(f)$ is the average power of $X(t)$

$$\int_{-\infty}^{\infty} S_X(f) df = R_X(0) = E[X(t)^2]$$

- $S_X(f)$ is the average power density, hence the average power of $X(t)$ in the frequency band $[f_1, f_2]$ is

$$\int_{-f_2}^{-f_1} S_X(f)\, df + \int_{f_1}^{f_2} S_X(f)\, df = 2\int_{f_1}^{f_2} S_X(f)\, df$$

- $S_X(f)$ is nonnegative: $S_X(f) \geq 0$ for all f. (shown later)
- In general, any function $S(f)$ that is real, even, nonnegative and has finite area can be a psd function.

White noise

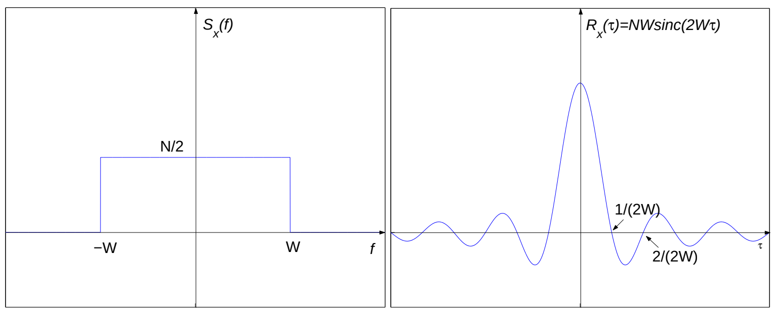
- Band-limited white noise: A zero-mean WSS process $N(t)$ which has the psd as a constant $\frac{N_0}{2}$ within $-W \le f \le W$ and zero elsewhere.
  - Similar to white light containing all frequencies in equal amounts.
  - Its average power is

$$E[X(t)^2] = \int_{-W}^{W} \frac{N_0}{2} df = N_0 W$$

  - Its auto-correlation function is

$$R_X(\tau) = \frac{N_0 \sin(2\pi W\tau)}{2\pi\tau} = N_0 W \text{sinc}(2W\tau)$$

  - For any t, the samples $X(t \pm \frac{n}{2W})$ for $n = 0, 1, 2, \ldots$ are uncorrelated.

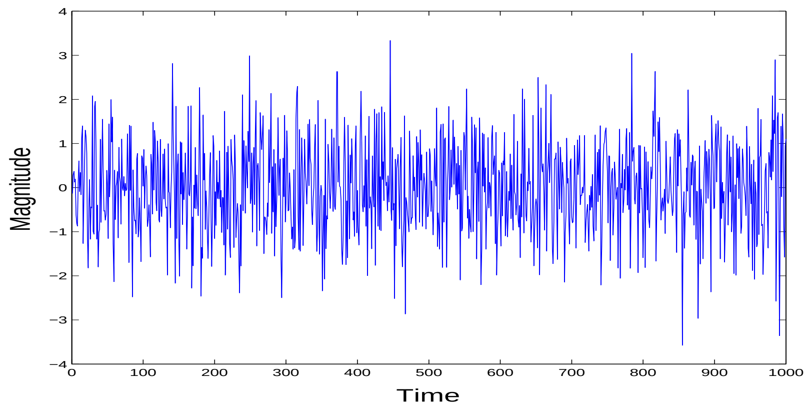


- White-noise process: Letting $W \to \infty$, we obtain a *white noise process*, which has

$$\begin{aligned} S_X(f) &= \frac{N_0}{2} \quad \text{for all } f \\ R_X(\tau) &= \frac{N_0}{2}\delta(\tau) \end{aligned}$$

  - For a white noise process, all samples are uncorrelated.
  - The process has infinite power and hence not physically realizable.
  - It is an idealization of physical noises. Physical systems usually are band-limited and are affected by the noise within this band.
- If the white noise $N(t)$ is a Gaussian random process, then we have

Gaussian white noise (GWN)



– WGN results from taking the derivative of the Brownian motion (or the Wiener process).
– All samples of a GWN process are independent and identically Gaussian distributed.
– Very useful in modeling broadband noise, thermal noise.

Cross-power spectral density

Consider two jointly-WSS random processes $X(t)$ and $Y(t)$:

- Their *cross-correlation* function $R_{XY}(\tau)$ is defined as

$$R_{XY}(\tau) = E[X(t+\tau)Y(t)]$$

  – Unlike the auto-correlation $R_X(\tau)$, the cross-correlation $R_{XY}(\tau)$ is not necessarily even. However

$$R_{XY}(\tau) = R_{Y,X}(-\tau)$$

- The *cross-power spectral density* $S_{XY}(f)$ is defined as

$$S_{XY}(f) = \mathcal{F}\{R_{XY}(\tau)\}$$

  In general, $S_{XY}(f)$ is complex even if the two processes are real-valued.

- Example: Signal plus white noise
  Let the observation be

$$Z(t) = X(t) + N(t)$$

  where $X(t)$ is the wanted signal and $N(t)$ is white noise. $X(t)$ and $N(t)$ are zero-mean uncorrelated WSS processes.

  – $Z(t)$ is also a WSS process

$$\begin{aligned} E[Z(t)] &= 0 \\ E[Z(t)Z(t+\tau)] &= E\left[\{X(t)+N(t)\}\{X(t+\tau)+N(t+\tau)\}\right] \\ &= R_X(\tau) + R_N(\tau) \end{aligned}$$

  – The psd of $Z(t)$ is the sum of the psd of $X(t)$ and $N(t)$

$$S_Z(f) = S_X(f) + S_N(f)$$

  – $Z(t)$ and $X(t)$ are jointly-WSS

$$E[X(t)Z(t+\tau)] = E[X(t+\tau)Z(t)] = R_X(\tau)$$

  Thus $S_{XZ}(f) = S_{ZX}(f) = S_X(f)$.

Review of LTI systems

- Consider a system that maps $y(t) = T[x(t)]$

  – The system is linear if

$$T[\alpha x_1(t) + \beta x_2(t)] = \alpha T[x_1(t)] + \beta T[x_2(t)]$$

  – The system is time-invariant if

$$y(t) = T[x(t)] \quad \rightarrow \quad y(t-\tau) = T[x(t-\tau)]$$

- An LTI system can be completely characterized by its impulse response

$$h(t) = T[\delta(t)]$$

  – The input-output relation is obtained through convolution

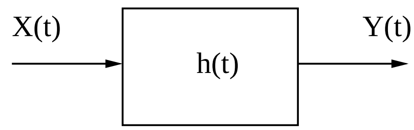
$$y(t) = h(t) * x(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau$$

- In the frequency domain: The system *transfer function* is the Fourier transform of $h(t)$

$$H(f) = \mathcal{F}[h(t)] = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}dt$$

Response of an LTI system to WSS random signals

Consider an LTI system $h(t)$



- Apply an input $X(t)$ which is a WSS random process
  - The output $Y(t)$ then is also WSS

$$\begin{aligned} E[Y(t)] &= m_X \int_{-\infty}^{\infty} h(\tau) d\tau = m_X \, H(0) \\ E\left[Y(t)Y(t+\tau)\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(r)\, h(s)\, R_X(\tau+s-r)\, ds dr \end{aligned}$$

  - Two processes $X(t)$ and $Y(t)$ are jointly WSS

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} h(s) R_X(\tau+s) ds = h(-\tau) * R_X(\tau)$$

  - From these, we also obtain

$$R_Y(\tau) = \int_{-\infty}^{\infty} h(r) R_{XY}(\tau-r) dr = h(\tau) * R_{XY}(\tau)$$

- The results are similar for discrete-time systems. Let the impulse response be h_n
  - The response of the system to a random input process X_n is

$$Y_n = \sum_k h_k X_{n-k}$$

  - The system transfer function is

$$H(f) = \sum_n h_n e^{-j2\pi n f}$$

  - With a WSS input X_n, the output Y_n is also WSS

$$\begin{aligned} m_Y &= m_X H(0) \\ R_Y[k] &= \sum_j \sum_i h_j h_i R_X[k+j-i] \end{aligned}$$

  - X_n and Y_n are jointly WSS

$$R_{XY}[k] = \sum_n h_n R_X[k+n]$$

Frequency domain analysis

- Taking the Fourier transforms of the correlation functions, we have

$$\begin{aligned} S_{XY}(f) &= H^*(f)\, S_X(f) \\ S_Y(f) &= H(f)\, S_{XY}(f) \end{aligned}$$

  where $H^*(f)$ is the complex conjugate of $H(f)$.

- The output-input psd relation

$$S_Y(f) = |H(f)|^2 S_X(f)$$

- Example: White noise as the input.
  Let $X(t)$ have the psd as

$$S_X(f) = \frac{N_0}{2} \quad \text{for all } f$$

  then the psd of the output is

$$S_Y(f) = |H(f)|^2 \frac{N_0}{2}$$

  Thus the transfer function completely determines the shape of the output psd. This also shows that any psd must be *nonnegative*.

Power in a WSS random process

- Some signals, such as $\sin(t)$, may not have finite energy but can have finite average power.

- The average power of a random process $X(t)$ is defined as

$$P_X = E\left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)^2 dt\right]$$

- For a WSS process, this becomes

$$P_X = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} E[X(t)^2] dt = R_X(0)$$

  But $R_X(0)$ is related to the FT of the psd $S(f)$.

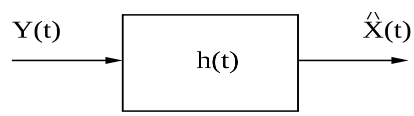
- Thus we have three ways to express the power of a WSS process

$$P_X = E[X(t)^2] = R_X(0) = \int_{-\infty}^{\infty} S(f) df$$

  The area under the psd function is the average power of the process.

Linear estimation

- Let $X(t)$ be a zero-mean WSS random process, which we are interested in estimating.
- Let $Y(t)$ be the observation, which is also a zero-mean random process jointly WSS with $X(t)$
  - For example, $Y(t)$ could be a noisy observation of $X(t)$, or the output of a system with $X(t)$ as the input.
- Our goal is to design a linear, time-invariant filter $h(t)$ that processes $Y(t)$ to produce an estimate of $X(t)$, which is denoted as $\hat{X}(t)$



  - Assuming that we know the auto- and cross-correlation functions $R_X(\tau)$, $R_Y(\tau)$, and $R_{XY}(\tau)$.
- To estimate each sample $X(t)$, we use an observation window on $Y(\alpha)$ as $t-a \le \alpha \le t+b$.
  - If $a=\infty$ and $b=0$, this is a (causal) *filtering* problem: estimating

X_t from the past and present observations.
  - If $a=b=\infty$, this is an *infinite smoothing* problem: recovering X_t from the entire set of noisy observations.
- The linear estimate $\hat{X}(t)$ is of the form

$$\hat{X}(t) = \int_{-b}^{a} h(\tau)Y(t-\tau)d\tau$$

- Similarly for discrete-time processing, the goal is to design the filter coefficients h_i to estimate X_k as

$$\hat{X}_k = \sum_{i=-b}^{a} h_i Y_{k-i}$$

- Next we consider the optimum linear filter based on the MMSE criterion.

Optimum linear MMSE estimation

- The MMSE linear estimate of $X(t)$ based on $Y(t)$ is the signal $\hat{X}(t)$ that minimizes the MSE

$$\text{MSE} = E\left[\left(X(t) - \hat{X}(t)\right)^2\right]$$

- By the orthogonality principle, the MMSE estimate must satisfy

$$E\left[e_t Y(t-\tau)\right] = E\left[\left(X(t) - \hat{X}(t)\right) Y(t-\tau)\right] = 0 \quad \forall\, \tau$$

  The error $e_t = X(t) - \hat{X}(t)$ is orthogonal to all observations $Y(t-\tau)$.

  - Thus for $-b \le \tau \le a$

$$\begin{aligned} R_{XY}(\tau) &= E\left[X(t)Y(t-\tau)\right] = E\left[\hat{X}(t)Y(t-\tau)\right] \\ &= \int_{-b}^{a} h(\beta) R_Y(\tau - \beta) d\beta \end{aligned} \tag{1}$$

    * To find $h(\beta)$ we need to solve an infinite set of integral equations. Analytical solution is usually not possible in general.
    * But it can be solved in two important special case: infinite smoothing ($a = b = \infty$), and filtering ($a = \infty, b = 0$).

  - Furthermore, the error is orthogonal to the estimate $\hat{X}_t$

$$E\left[e_t \hat{X}_t\right] = \int_{-b}^{a} h(\tau) E\left[e_t Y(t-\tau)\right] d\tau = 0$$

  - The MSE is then given by

$$\begin{aligned} E\left[e_t^2\right] &= E\left[e_t \left(X(t) - \hat{X}(t)\right)\right] = E\left[e_t X(t)\right] \\ &= E\left[\left(X(t) - \hat{X}(t)\right) X(t)\right] \\ &= R_X(0) - \int_{-b}^{a} h(\beta) R_{XY}(\beta) d\beta \end{aligned} \tag{2}$$

- For the discrete-time case, we have

$$R_{XY}(m) = \sum_{i=-b}^{a} h_i R_X(m-i) \tag{3}$$

$$E\left[e_k^2\right] = R_X(0) - \sum_{i=-b}^{a} h_i R_{XY}(i) \tag{4}$$

  From (3), one can design the filter coefficients h_i.

Infinite smoothing

- When $a, b \to \infty$, we have

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} h(\beta) R_Y(\tau - \beta) d\beta = h(\tau) * R_Y(\tau)$$

  - Taking the Fourier transform gives the transfer function for the optimum filter

$$S_{XY}(f) = H(f) S_Y(f) \quad \Rightarrow \quad H(f) = \frac{S_{XY}(f)}{S_Y(f)}$$