Outline

- Power Spectrum Estimation: Introduction
- Deterministic Signals: Computation of the Energy Density Spectrum
- Random Signals: Estimation of Autocorrelation and Power Spectrum
- Non-parametric Power Spectrum Estimation Methods
  - Bartlett Method
  - Welch Method
  - Blackman and Tuckey

Power Spectrum Estimation: Introduction

- The goal is to estimate the power density spectrum of a signal given a finite length observation of it
- When the signal is stationary the longer the data the better the estimate is
- When the signal is non-stationary longer data do not guarantee better estimates
- Power spectrum estimation methods can be categorized into two groups
  - Non-parametric estimation: no prior model is assumed, the power density spectrum samples are estimated directly
  - Parametric estimation: prior knowledge is used to model the power density spectrum using a few parameters, these parameters are estimated that yield a final estimate of power density spectrum

Computation of the Energy Density Spectrum

- Remember the definitions of autocorrelation for a deterministic signal

$$R_x(\tau) = \int_{-\infty}^{\infty} x_a^*(t) x_a(t+\tau) \mathrm{d}t$$

  and the energy density spectrum which is the Fourier transform of the autocorrelation function

$$S_x(F) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi F t} \mathrm{d}\tau$$

- Let us consider calculating these from the samples of $x(t)$
- When there is no aliasing there is a one-to-one relation between the samples of a signal and the signal itself

Computation of the Energy Density Spectrum (Cont.)

- For a discrete signal

$$R_x[k] = \sum_{n=-\infty}^{\infty} x^*[n]x[n+k]$$

and

$$S_x(f) = \sum_{k=-\infty}^{\infty} R_x[k]e^{-j2\pi kf}$$

- $S_x(f)$ can also be directly calculated as

$$S_x(f) = |X(f)|^2 = \left| \sum_{n=-\infty}^{\infty} x[n]e^{-j2\pi fn} \right|$$

Computation of the Energy Density Spectrum (Cont.)

- In practice we will have only finite number of samples
- This limitation can mathematicall represented by multiplying the original function with a window function
- Assuming we have N samples, the signal that we will be using is

$$\tilde{x}[n] = x[n]w[n] = \begin{cases} x[n], & 0 \leq n \leq N-1 \\ 0, & otherwise \end{cases}$$

- Using this truncated signal we have

$$\tilde{X}(f) = X(f) * W(f) = \int_{-0.5}^{0.5} X(\alpha)W(f-\alpha)\mathrm{d}\alpha$$

- This windowing results in errenous frequency components, called the leakage, where the actual signal frequency content is zero
- As in filter design, we can use different windows to reduce leakage in the price of a drawback such as increasing the width of the main lobe

Computation of the Energy Density Spectrum (Cont.)

- Using DFT we can calculate the samples of the spectrum

$$S_{\tilde{x}\tilde{x}}\left[\frac{k}{N}\right] = \left|\sum_{n=0}^{N-1} \tilde{x}[n]e^{-j2\pi kn/N}\right|^2$$

- We have distortions due to windowing

Estimation of Autocorrelation and Power Spectrum of Random Signals

- Remember definitions for random signals

$$R_x(\tau) = \mathrm{E}[x * (t)x(t+\tau)]$$

and

$$S_x(F) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi Ft} \mathrm{d}ta$$

- Problem: We do not know true autocorrelation, but need to estimate it from a single realization (what we observe)

$$\hat{R}_x(\tau) = \frac{1}{2T_0} \int_{-T_0}^{T_0} x^*(t)x(t+\tau)\mathrm{d}t \qquad (*)$$

- Assuming ergodicty, we have

$$\lim_{T_0 \to \infty} \hat{R}_x(\tau) = R_x(\tau)$$

- Therefore we will be using Eq. (*) as the estimate of the autocorrelation function

Estimation of Autocorrelation and Power Spectrum of Random Signals (Cont.)

- We obtain an estimate of $S_x(F)$ using the time average of the autocorrelation function

$$
\begin{aligned}
P_x(F) &= \int_{-T_0}^{T_0} R_x(\tau) e^{-j2\pi F\tau} \mathrm{d}\tau \\
&= \frac{1}{2T_0} \int_{-T_0}^{T_0} \left[\int_{-T_0}^{T_0} x^*(t) x(t+\tau) e^{-j2\pi F\tau} \mathrm{d}t \right] \mathrm{d}\tau \\
&= \frac{1}{2T_0} \int_{-T_0}^{T_0} x^*(t) \left[\int_{-T_0}^{T_0} x(t+\tau) \mathrm{d}t \right] e^{-j2\pi F\tau} \mathrm{d}\tau \\
&= \frac{1}{2T_0} \left| \int_{-T_0}^{T_0} x(t) e^{-j2\pi Ft} \mathrm{d}t \right|^2
\end{aligned}
$$

Estimation of Autocorrelation and Power Spectrum of Random Signals (Cont.)

- Now let us turn back to discrete case
- We have the estimate of the autocorrelation function

$$\hat{R}_x[m] = \frac{1}{N-M} \sum_{n=0}^{N-m-1} x^*[n]x[n+m], \qquad m = 0, 1, \ldots, N-1$$

- Let us have a look at the mean of this estimate

$$\mathrm{E}\left[\hat{R}_x[m]\right] = \frac{1}{N-M} \sum_{n=0}^{N-m-1} \mathrm{E}\left[x^*[n]x[n+m]\right] = R_x[m]$$

- Unbiased estimate
- The variance can be shown to limit to zero for large sample size

Estimation of Autocorrelation and Power Spectrum of Random Signals (Cont.)

- Although the variance is zero for large N, it is high for large values of m given a finite M
- To obtain an estimate with smaller variance, we sacrifice unbiasedness
- Consider the estimate

$$R_x[m] = \frac{1}{N} \sum_{m=0}^{N-m-1} x^*[n][n+m] \qquad (*)$$

- This estimate is clearly biased but can be shown to have smaller variance and is preferrable
- Bias which is equal to $|m|R_x[m]/N$ vanishes and variance aproaches zero as N gets larger

Estimation of Autocorrelation and Power Spectrum of Random Signals (Cont.)

- Using Eq. (*) as the estimate of autocorrelation, we obtain the periodogram

$$
\begin{aligned}
P_x(f) &= \sum_{m=-N+1}^{N-1} \hat{R}_x[m] e^{-j2\pi f m} \\
&= \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j2\pi f n} \right|^2 = \frac{1}{N} |X(f)|^2
\end{aligned}
$$

Estimation of Autocorrelation and Power Spectrum of Random Signals (Cont.)

- Let us have a look at the mean of $P_x(f)$

$$\begin{aligned}
\mathrm{E}[P_x(f)] &= \mathrm{E}\left[\sum_{m=-N+1}^{N-1} \hat{R}_x[m] e^{-j2\pi fm}\right] \\
&= \sum_{m=-N+1}^{N-1} \mathrm{E}\left[\hat{R}_x[m]\right] e^{-j2\pi fm} \\
&= \sum_{m=-N+1}^{N-1} \left(1 - \frac{|m|}{N}\right) R_x[m] e^{-j2\pi fm}
\end{aligned}$$

- That is the mean of the estimated spectrum is the FT of windowed version of the autocorrelation function

$$\mathrm{E}[P_x(f)] = \int_{-0.5}^{0.5} R_x(\alpha) W(f-\alpha) \mathrm{d}\alpha$$

- The bias vanishes for large N, but not the variance
- We lose an important property with this straightforward estimator. Hence we need more complicated methods

Estimation of Autocorrelation and Power Spectrum of Random Signals (Cont.)

- Similar to the deterministic case, we can obtain the samples of the estimate of the density (periodogram) using DFT

$$P_x(\frac{k}{N}) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n]e^{-j2\pi nk/N}\right|^2 \qquad k = 0, 1, \ldots, N-1$$

- In practice, these samples do not provide a good representation, hence we need to sample more densely. This can be performed by zero padding the signal first, upto L samples

$$P_x(\frac{k}{L}) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n]e^{-j2\pi nk/L}\right|^2 \qquad k = 0, 1, \ldots, L-1$$

Nonparametric Power Spectrum Estimation Methods

- We consider three nonparametric methods (no assumption on data, no modeling)
  - The Bartlett Method: Averaging Periodograms
  - The Welch Method: Averaging Modified Periodograms
  - The Blackman and Tuckey Method: Smoothing Periodograms

Nonparametric Power Spectrum Estimation Methods: The Bartlett Method

- Reduce variance by averaging the periodogram of the parts of the original signal

- Group data of length N into smaller segments of length M

$$x_i[n] = x(n + iM) \qquad i = 0, 1, \ldots, K-1 \qquad n = 0, 1, \ldots, M-1$$

- For each of these subgroups we have

$$P_x^{(i)}(f) = \frac{1}{M} \left| \sum_{n=0}^{M-1} x_i[n] e^{-j2\pi f n} \right|^2, \qquad i = 0, 1, \ldots, K-1$$

- Averaging yields an estimate of the periodogram

$$P_x^{\mathrm{B}}(f) = \frac{1}{K} \sum_{i=0}^{K-1} P_x^{(i)}(f)$$

Bartlett Method: Properties

- Let us have a look at the mean

$$\begin{aligned} \mathrm{E}\left[P_x^{\mathrm{B}}(f)\right] &= \tfrac{1}{K}\sum_{i=0}^{K-1}\mathrm{E}\left[P_x^{(i)}(f)\right] \\ &= \mathrm{E}\left[P_x^{(i)}(f)\right] \end{aligned}$$

- The mean of the subgroups is

$$\mathrm{E}\left[P_x^{(i)}(f)\right] = \sum_{-M+1}^{M-1}\left(1-\frac{|m|}{M}\right)R_x[m]e^{-j2\pi fm} = S_x(f) * w(f)$$

- That is the expected value is equal to the convolved version of the original density speectrum as before
- However, now the convolving window is more narrow (M samples instead of N)
- That is we lose frequency resolution by a factor of K
- Advantage: decreased variance, variance is reduced by a factor of K

Nonparametric Power Spectrum Estimation Methods: The Welch Method

- Reduce variance by averaging the periodograms as in Bartlett method
- Now, the groups are allowed to have overlapping samples, and the avreaging is done using some modified version of the periodogram
- Let us group the data into overlapping subgroups

$$x_i[n] = x(n + iD) \qquad i = 0, 1, \ldots, L-1 \qquad n = 0, 1, \ldots, M-1$$

- Now let us also modify the segments with a window function to obtain

$$\tilde{P}^{(i)}(f) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_i[n] w[n] e^{-j2\pi f n} \right|^2$$

where $U = \frac{1}{M} \sum_{n=0}^{M-1} w^2[n]$ is a normalization factor

Nonparametric Power Spectrum Estimation Methods: The Welch Method (Cont.)

- The Welch method then results in the following estimate

$$P_x^{\mathrm{W}}(f) = \frac{1}{L}\sum_{i=0}^{L-1}\tilde{P}_x^{(i)}(f)$$

Welch Method: Properties

- Let us have a look at the expected value of the estimate

$$\begin{aligned} \mathrm{E}\left[P_x^{\mathrm{W}}(f)\right] &= \tfrac{1}{K}\sum_{i=0}^{K-1}\mathrm{E}\left[\tilde{P}_x^{(i)}(f)\right] \\ &= \mathrm{E}\left[\tilde{P}_x^{(i)}(f)\right] \end{aligned}$$

- The expected value for the segments is

$$\begin{aligned} \mathrm{E}\left[\tilde{P}_x^{(i)}(f)\right] &= \tfrac{1}{MU}\sum_{n=0}^{M-1}\sum_{m=0}^{M-1} w[n]w[m]\mathrm{E}\left[x_i[n]x_i^*[m]\right]e^{-j2\pi f(n-m)} \\ &= \tfrac{1}{MU}\sum_{n=0}^{M-1}\sum_{m=0}^{M-1} w[n]w[m]R_x(n-m)e^{-j2\pi f(n-m)} \\ &= \tfrac{1}{MU}\sum_{n=0}^{M-1}\sum_{m=0}^{M-1} w[n]w[m] \\ &\quad \times \int_{-0.5}^{0.5} S_x(\alpha)e^{-j2\pi(n-m)(f-\alpha)}\mathrm{d}\alpha \\ &= S_x(f) * W(f) \end{aligned}$$

where

$$W(f) = \frac{1}{MU}\left|\sum_{n=0}^{M-1} w[n]e^{-j2\pi fn}\right|^2$$

Welch Method: Properties

- The variance has a more flexible expression (since overlapping is allowed) than the Bartlett method, allowing for better tradeof between frequency resolution and variance

Nonparametric Power Spectrum Estimation Methods: The Blackman and Tuckey Method

- The estimated autocorrelation function is windowed first before the Fourier transform yielding the estimate for the spectrum
- The windowing helps to give less weight (or eliminate) the large lag samples. These large lag samples produce poor results since less samples are used in the estimation
- Therefore, our estimate is

$$P_x^{\mathrm{BT}}(f) = \sum_{m=-M+1}^{M-1} R_x[m]w[m]e^{-j2\pi fm}$$

- Since we now have a window function we can write

$$P_x^{\mathrm{BT}}(f) = \sum_{m=-\infty}^{\infty} R_x[m]w[m]e^{-j2\pi fm}$$

- Then

$$P_x^{\mathrm{BT}}(f) = P_x(f) * W(f)$$

Blackman and Tuckey Method: Properties

- The expected value of the estimate is

$$\mathrm{E}[P_x^{\mathrm{BT}}(f)] = \int_{-0.5}^{0.5} \mathrm{E}[P_x(\alpha)] W(f-\alpha) \mathrm{d}\alpha$$

- Substituting

$$\mathrm{E}[P_x(\alpha)] = \int_{-0.5}^{0.5} [S_x(\theta)] W_{\mathrm{B}}(\alpha-\theta) \mathrm{d}\theta$$

we obtain the mean as

$$\mathrm{E}[P_x^{\mathrm{BT}}(f)] = \int_{-0.5}^{0.5}\int_{-0.5}^{0.5} S_x(\theta) W_{\mathrm{B}}(\alpha-\theta) W(f-\alpha) \mathrm{d}\alpha \mathrm{d}\theta$$

- Under certain assumptions the variance of BT spectrum estimate is approximately

$$S_x^2(f)\left[\frac{1}{N}\sum_{m=-M+1}^{M-1} w^2[m]\right]$$