Lecture
Power Spectral Density Estimation
EE 451: Digital Signal Processing

Lecture Notes Update on December 4, 2017

Aly El-Osery, Electrical Engineering Dept., New Mexico Tech

.1

Motivation

Sensors suffer from noise effects that can not be removed through calibration, consquently, we need to

- understand the nature of the noise
- be able to extract parameters from actual data
- develop models to mimic noise in simulation to provide performance capabilities

.2

Purpose

Estimate the distribution of power in a signal. Unfortunately, truth and what is practical cause a problem.

Truth

- Infinitely long.
- Continuous in time and value.
- Provides true distribution of power.

Practice

- Finite length.
- Discrete in time and value.
- Only approximation of distribution of power.

Let's make it more interesting

The signal is stochastic in nature.

.3

1 Review Material

1.1 Signal Classification

Energy and Power

Assume the voltage across a resistor R is $e(t)$ and is producing a current $i(t)$. The instantaneous power per ohm is $p(t) = e(t)i(t)/R = i^2(t)$.

Total Energy

$$E = \lim_{T \to \infty} \int_{-T}^{T} i^2(t)dt \qquad (1)$$

Average Power

$$P = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} i^2(t)dt \tag{2}$$

Arbitrary signal $x(t)$

Total Normalized Energy

$$E \triangleq \lim_{T \to \infty} \int_{-T}^{T} |x(t)|^2 dt = \int_{-\infty}^{\infty} |x(t)|^2 dt \tag{3}$$

Normalized Power

$$P \triangleq \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 dt \tag{4}$$

- $x(t)$ is an *energy signal* iff $0 < E < \infty$, so that $P = 0$.
- $x(t)$ is a *power signal* iff $0 < P < \infty$, so that $E = \infty$.

1.2   Time Averages

Correlation

For Energy Signals

$$\phi(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)dt \tag{5}$$

Provides a measure of similarity or coherence between a signal and a delayed version of itself. Note that $\phi(0) = E$

For Power Signals

$$R(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)x(t+\tau)dt \tag{6}$$

For Periodic Signals

$$R(\tau) = \frac{1}{T_0} \int_{T_0} x(t)x(t+\tau)dt \tag{7}$$

1.3   Frequency Domain

Energy Spectral Density

Rayleigh's Energy Theorem or Parseval's theorem

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(F)|^2 dF \tag{8}$$

Energy Spectral Density

$$G(F) \triangleq |X(F)|^2 \tag{9}$$

with units of *volts^2-sec^2* or, if considered on a per-ohm basis, *watts-sec/Hz=joules/Hz*

Power Spectral Density

$$P = \int_{-\infty}^{\infty} S(F)dF = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 dt \tag{10}$$

where we define $S(F)$ as the power spectral density with units of watts/Hz. Note that $R(0) = \int_{-\infty}^{\infty} S(F)dF$.

2 Random Signals and Noise

Basic Definitions
- Define an *experiment* with random *outcome*.
- Mapping of the outcome to a variable $\Rightarrow$ random variable.
- Mapping of the outcome to a function $\Rightarrow$ random function.

2.1 Statistical Averages

Probability (Cumulative) Distribution Function (cdf)

$$F_X(x) = \texttt{probability that } X \leq x = P(X \leq x) \tag{11}$$

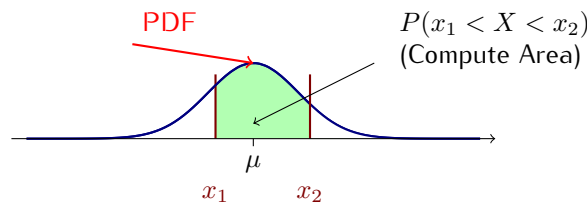Describes the manner random variables take different values.

Probability Density Function (pdf)

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{12}$$

and

$$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)dx \tag{13}$$

PDF of Discrete Random Variables
If the random variable X takes a set of discrete values x_i with probability p_i, the pdf of X is expressed in terms of Dirac delta functions, i.e.,
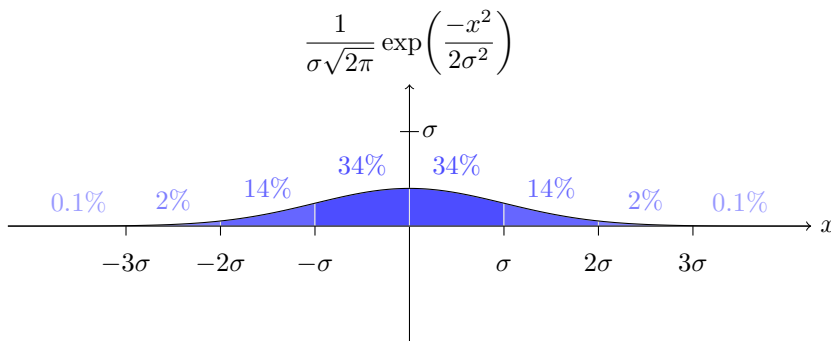
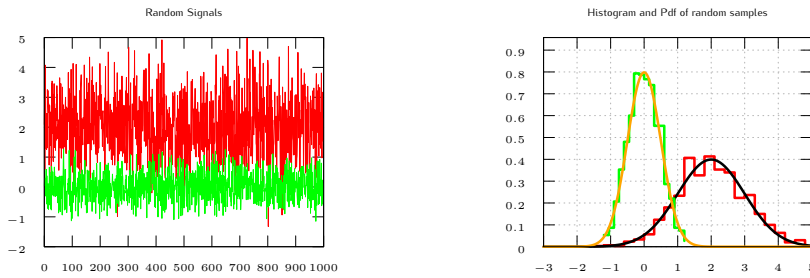$$f_X(x) = \sum_i p_i \delta(x - x_i) \tag{14}$$

Gaussian Distribution

$$f_X(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left[-\frac{x - \mu_x}{2\sigma_x^2}\right] \tag{15}$$

For example if $\sigma_x = \sigma$ and $\mu_x = 0$

3

PDF of White Noise

Mean and Variance

Mean of a Discrete RV

$$\bar{X} = \mathbb{E}[X] = \sum_{j=1}^{M} x_j P_j \tag{16}$$

Mean of a Continuous RV

$$\bar{X} = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx \tag{17}$$

Variance of a RV

$$\sigma_X^2 \triangleq \mathbb{E}\left\{[X - \mathbb{E}(X)]^2\right\} = \mathbb{E}[X^2] - \mathbb{E}^2[X] \tag{18}$$

Covariance and Autocorrelation

Given a two random variables X and Y.

Covariance

$$\mu_{XY} = \mathbb{E}\left\{[X - \bar{x}][Y - \bar{Y}]\right\} = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \tag{19}$$

Correlation Coefficient

$$\rho_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y} \tag{20}$$

Autocorrelation

$$\Gamma_X(\tau) = \mathbb{E}[X(t)X(t+\tau)] \tag{21}$$

2.2 Stochastic Processes

Terminology

See Figure 1

- $X(t, \zeta_i)$: sample function.
- The governing experiment: random or stochastic process.
- All sample functions: ensemble.
- $X(t_j, \zeta)$: random variable.

Strict Sense Stationarity

If the joint pdfs depend only on the time difference regardless of the time origin, then the random process is known as *stationary*.

For stationary process means and variances are independent of time and the covariance depends only on the time difference.

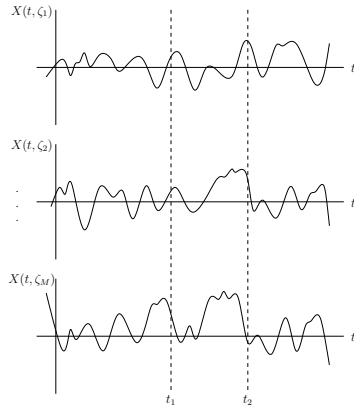Figure 1: Sample functions of a random process

Wide Sense Stationarity

If the joint pdfs depends on the time difference but the mean and variances are time-independent, then the random process is known as *wide-sense-stationary*.

.19

Ergodicity

If the time statistics equals ensemble statistics, then the random process is known as *ergodic*.

Any statistic calculated by averaging of all members of an ergodic ensemble at a fixed time can also be calculated by using a single representative waveform and averaging over all time.

.20

2.3 Correlation and Power Spectral Density

Power Spectral Density

Given a sample function $X(t, \zeta_i)$ of a random process, we obtain the power spectral density by
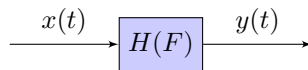
$$S(F) \xleftrightarrow{\mathcal{F}} \Gamma(\tau) \tag{22}$$

i.e., for a wide sense stationary signal, the power spectral density and autocorrelation are Fourier transform pairs.

.21

2.4 Input-Output Relationship of Linear Systems

Input-Output Relationship of Linear Systems
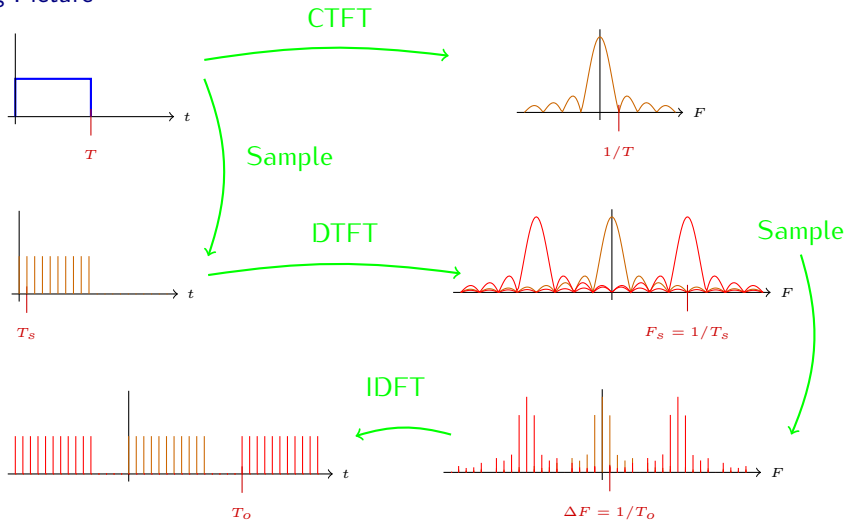

$$S_Y(F) = |H(F)|^2 S_X(F) \tag{23}$$

Noise Shaping

If $x(t)$ is white noise, we can design the filter $h(t)$ to "shape" the noise.

.22

5

3 Discrete Signals and Systems

Big Picture

Sampling Remarks

- Must sample more than twice bandwidth to avoid aliasing.
- FFT represents a periodic version of the time domain signal → could have time domain aliasing.
- Number of points in FFT is the same as number of points in time domain signal.

4 Power Spectral Density

Obtaining PSD for Discrete Signals

What we want is
$$\Gamma_X(\tau) = \mathbb{E}[X(t)X(t+\tau)] \xrightarrow{\mathcal{CTFT}} S_X(F)$$

For infinitely long signals.

What we can compute is
$$\gamma_X(m) = \mathbb{E}[X(n)X(n+m)] \xrightarrow{\mathcal{DFT}} P_X(f)$$

For finite length signals.

What do we need in an estimate

As $N \to \infty$ and in the mean squared sense

Unbiased
Asymptotically the mean of the estimate approaches the true power.

Variance
Variance of the estimate approaches zero.

Resulting in a consistent estimate of the power spectrum.

6

Periodogram
computed using $1/N$ times the magnitude squared of the FFT

$$\lim_{N\to\infty} \mathbb{E}[P_X(f)] = S_X(f)$$

$$\lim_{N\to\infty} var[P_X(f)] = S_X^2(f)$$

Welch Method
computed by segmenting the data (allowing overlaps), windowing the data in each segment then computing the average of the resultant priodogram

$$\mathbb{E}[P_X(f)] = \frac{1}{2\pi MU}S_X(f) \circledast W(f)$$

$$var[P_X(f)] \approx \frac{9}{8L}S_X^2(f)$$

.27

Welch Method

Assuming data length N, segment length M, Bartlett window, and 50% overlap

- FFT length $= M = 1.28/\Delta f = 1.28F_s/\Delta F$
- Resulting number of segments $= L = \frac{2N}{M}$
- Length of data collected in sec. $= \frac{1.28L}{2\Delta F}$
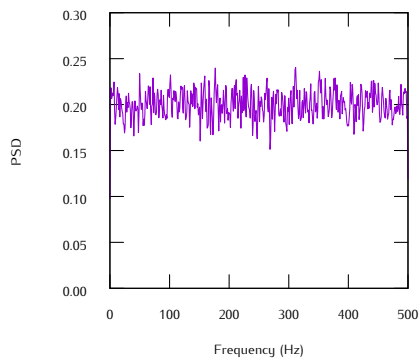
.28

pwelch Function

```
[Pxx,f] = pwelch(x,window,noverlap,...
                 nfft,fs,'range')
```

You can use [] in fields that you want the default to be used.

.29

pwelch Function - WGN signal

```
Fs  = 1000;
x  = sqrt(0.1*Fs)*randn(1,100000);
[Pxx,f]  = pwelch(x,1024,[],[],Fs,'onesided');
```
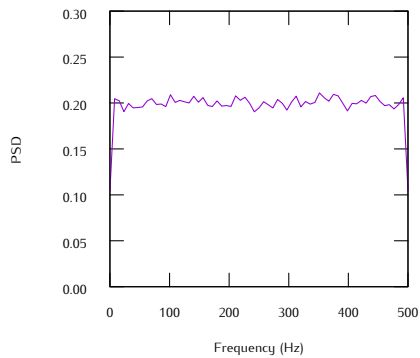

- Variance to high.

.30

pwelch Function - WGN signal
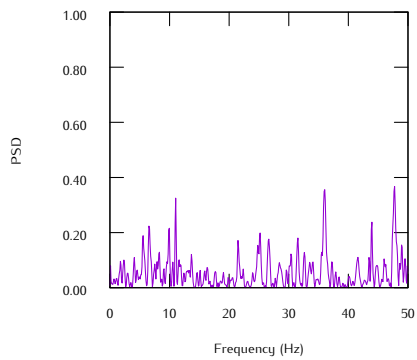
```
    [Pxx,f] = pwelch(x,128,[],[],Fs,'onesided')
```


- Reduced window size.
- Variance is now smaller.

pwelch Function - cos + WGN signal

```
    Fs = 100;    t = 0:1/Fs:5;
    x = cos(2*pi*10*t)+cos(2*pi*11*t)+...
            sqrt(0.1*Fs)*randn(1,length(t));
    [Pxx,f] = pwelch(x,1024,[],[],Fs,'onesided');
```
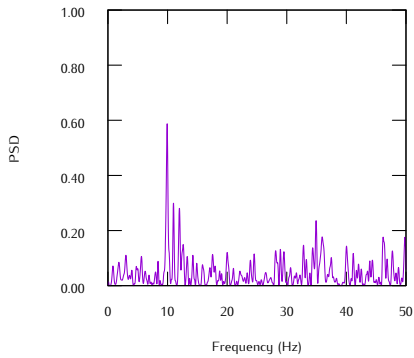

- Window larger than length of data.
- Frequency components can't be resolved.
- Variance high.

pwelch Function - cos + WGN signal

```
    Fs = 100;    t = 0:1/Fs:5;
    x = cos(2*pi*10*t)+cos(2*pi*11*t)+...
            sqrt(0.1*Fs)*randn(1,length(t));
    [Pxx,f] = pwelch(x,1024,[],4096,Fs,'onesided');
```

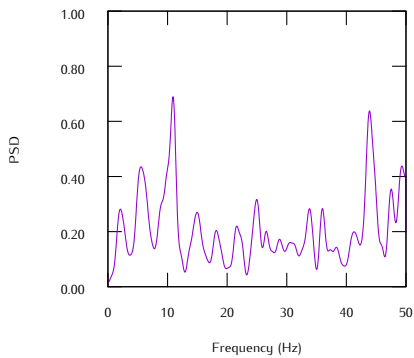- As expected increasing `nFFT` does not help.

pwelch Function - cos + WGN signal

```
Fs = 100;    t = 0:1/Fs:5;
x = cos(2*pi*10*t)+cos(2*pi*11*t)+...
        sqrt(0.1*Fs)*randn(1,length(t));
[Pxx,f] = pwelch(x,128,[],4096,Fs,'onesided');
```


- Decreasing the window size decreases the variance.
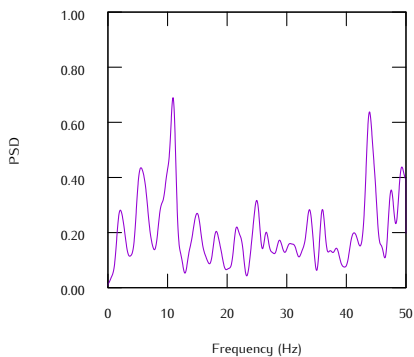- Still can't resolve the two frequencies.

pwelch Function - cos + WGN signal

```
Fs = 100;    t = 0:1/Fs:50;
x = cos(2*pi*10*t)+cos(2*pi*11*t)+...
        sqrt(0.1*Fs)*randn(1,length(t));
[Pxx,f] = pwelch(x,128,[],4096,Fs,'onesided');
```


9

- Length of data sequence must be increased.
- Still can't resolve the two frequencies as the window size is too small.

pwelch Function - cos + WGN signal

```
Fs = 100;    t = 0:1/Fs:50;
x = cos(2*pi*10*t)+cos(2*pi*11*t)+...
        sqrt(0.1*Fs)*randn(1,length(t));
[Pxx,f] = pwelch(x,256,[],4096,Fs,'onesided');
```
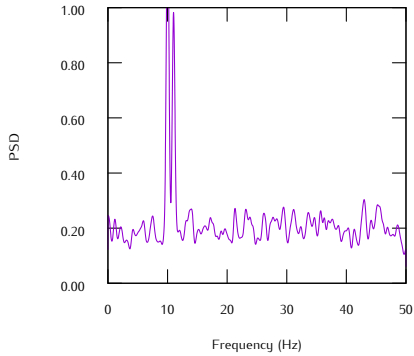

- Now we can resolve the two frequencies.

Spectral Estimation - Remarks

- The length of the data sequence determines the maximum resolution that can be observed.
- Increasing the window length of each segment in the data increases the resolution.
- Decreasing the window length of each segment in the data decreases the variance of the estimate.
- nFFT only affects the amount of details shown and not the resolution.