The Periodogram

Ergodicity

A key but often unspoken assumption in spectral estimation is <u>ergodicity</u>, which essentially says that "ensemble" averages can be replaced by "time" averages. Precisely, we will assume

$$\underbrace{\lim_{M\to\infty} \frac{1}{2M+1} \sum_{m=-M}^{M} x^*[m]\, x[m+k]}_{\boxed{\text{time average}}} = \underbrace{E\left\{x^*[n]\, x[n+k]\right\}}_{\boxed{\text{ensemble average}}} = \gamma_x[k]$$

We will always assume the RPs we study are ergodic. This is what allows spectral estimation from a <u>single</u> realization.

The Periodogram

Suppose we are given a realization $x[0], \dots, x[N-1]$ of a RP and we wish to estimate its PSD.

Recall that if $\sum_{k=-\infty}^{\infty} |\gamma_x[k]| < \infty$ and $\mu[n]=0$ then

$$P_x(f) = \lim_{M\to\infty} E\left\{ \frac{1}{2M+1} \left| \sum_{k=-M}^{M} x[k] e^{-2\pi i f k} \right|^2 \right\}.$$

This motivates the <u>periodogram</u>:

$$\hat{P}_{PER}(f) = \frac{1}{N} \left| \sum_{k=0}^{N-1} (x[k]-\bar{x})\, e^{-2\pi i f k} \right|^2.$$

This essentially computes the magnitude squared of the DTFT of the observed signal (zero padded) and normalizes by N so that $\hat{P}_{PER}$ still approximately integrates to $\gamma_x[0]$.

In practice $\hat{P}_{PER}(f)$ is computed at a finite list of frequencies $f = 0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}$ using the DFT / FFT.

Note that the periodogram can be written

$$\hat{P}_{PER}(f) = \frac{1}{N} \left| \sum_{k=0}^{N-1} (x[k] - \bar{x}) e^{-2\pi i f k} \right|^2$$

$$= \frac{1}{N} \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} (x[p] - \bar{x})^* (x[q] - \bar{x}) e^{-2\pi i f (p-q)}$$

$$= \frac{1}{N} \sum_{k=-(N-1)}^{N-1} \underbrace{\sum_{j=0}^{N-1-|k|} (x[j+|k|] - \bar{x})(x[j] - \bar{x}) e^{-2\pi i f k}}_{\hat{\gamma}[k]}$$

"sample autocovariance"

$$= \frac{1}{N} \sum_{k=-(N-1)}^{N} \hat{\gamma}[k] e^{-2\pi i f k}$$

Where $\hat{\gamma}[k]$ is the estimate of $\gamma[k]$ given by

$$\hat{\gamma}[k] = \begin{cases} \frac{1}{N} \sum_{j=0}^{N-1-k} (x[j+k] - \bar{x})(x[j] - \bar{x}) e^{-2\pi i f k}, & k \geq 0 \\ \hat{\gamma}[-k] & k < 0 . \end{cases}$$

Thus, the periodogram may be thought of as a "plug-in" estimator: estimate $\gamma[k]$ and plug the estimate into the Fourier transform.

Unfortunately, the periodogram is not a very good estimator of the PSD.

Bias/Variance Tradeoff

If θ is a parameter and $\hat{\theta}$ an estimator of θ, then

$$
\begin{aligned}
MSE(\hat{\theta}) &= E\{(\hat{\theta}-\theta)^2\} \\
&= E\{(\hat{\theta} - E\{\hat{\theta}\} + E\{\hat{\theta}\} - \theta)^2\} \\
&= E\{(\hat{\theta} - E\{\hat{\theta}\})^2\} \qquad \rightarrow \text{Var} \\
&\quad + E\{(\hat{\theta} - E\{\hat{\theta}\})\cdot(E\{\hat{\theta}\} - \theta)\} \qquad \rightarrow 0 \\
&\quad + (E\{\hat{\theta}\} - \theta)^2 \qquad \rightarrow \text{Bias}^2 \\
&= \text{Variance}(\hat{\theta}) + \text{Bias}^2(\hat{\theta})
\end{aligned}
$$

In general, there is a tradeoff between bias and variance. If one goes down, the other goes up. Hopefully, as $N \to \infty$, both go to zero.

Bias of Periodogram

First, let's look at $\hat{\gamma}[k]$. We have

$$E\left\{\hat{\gamma}[k]\right\} = E\left\{\frac{1}{N}\sum_{n=0}^{N-1-|k|}\left(x[n+|k|]-\bar{x}\right)\left(x[n]-\bar{x}\right)\right\}$$

$$= \frac{1}{N}\sum_{n=0}^{N-1-|k|} E\left\{\left(x[n+|k|]-\bar{x}\right)\left(x[n]-\bar{x}\right)\right\}$$

$$\approx \frac{1}{N}\sum_{n=0}^{N-1-|k|} E\left\{\left(x[n+|k|]-\mu\right)\left(x[n]-\mu\right)\right\}$$

$$= \frac{1}{N}\sum_{n=0}^{N-1-|k|} \gamma_x[k]$$

$$= \frac{N-|k|}{N}\gamma_x[k].$$

Conclusions:

1.

2.

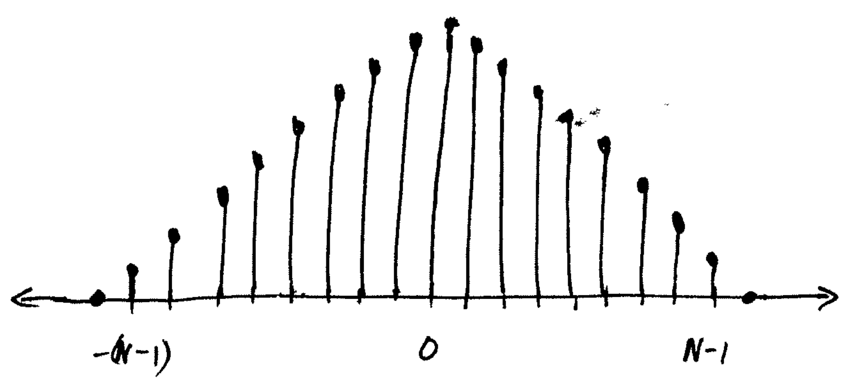
We can write

$$E\{\hat{\gamma}[k]\} = w_B[k] \cdot \gamma_x[k],$$

where

$$w_B[k] = \begin{cases} \dfrac{N-|k|}{N} & -(N-1) \le k \le N-1 \\ 0 & \text{otherwise} \end{cases}$$

is a $2N+1$ point <u>Bartlett window</u>



$\hat{\gamma}[k]$ has large variance for large lags k because little data is used. For example,

$$\hat{\gamma}[N-1] = \frac{1}{N} x[0] \cdot x[N-1]$$

is an average of one term!

We could force the estimate to be unbiased:

$$\check{\gamma}[k] = \frac{1}{N-|k|} \sum_{n=0}^{N-1-|k|} x[n+|k|]\, x[n]$$

Unfortunately, the variance is now even greater.

e.g.,

$$\check{\gamma}[N-1] = 1 \cdot x[0] \cdot x[N-1]$$

N times more variance!

Even worse, $\check{\gamma}[k]$ is usually not non-negative definite, meaning the PSD estimate might not even be a density!

The bias of the periodogram is

$$E\left\{\hat{P}_{PER}(f)\right\} = E\left\{\sum_{k=-(N-1)}^{N} \hat{\gamma}[k]\, e^{-2\pi i f k}\right\}$$

$$= \sum_{k=-(N-1)}^{N-1} E\left\{\hat{\gamma}[k]\right\} e^{-2\pi i f k}$$

$$= \sum_{k=-(N-1)}^{N-1} w_B[k]\, \gamma[k]\, e^{-2\pi i f k}$$

Bartlett window

$$= \sum_{k=-\infty}^{\infty} w_B[k]\, \gamma[k]\, e^{-2\pi i f k}$$

since $0 = w_B[k]$, $|k| \geq N$

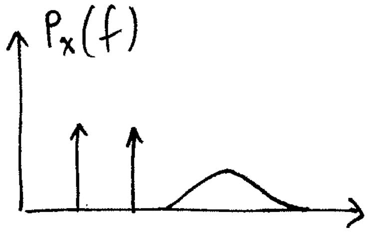
$$= W_B(f) * P_x(f)$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} P_x(f)\, W_B(f-u)\, du$$

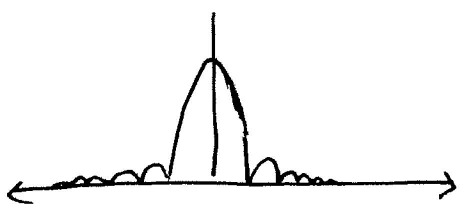
$\Rightarrow$ The ensemble average of many periodograms is smeared / blurred by the frequency response $W_B(f)$ of the $2N+1$ point Bartlett window

$\Rightarrow$ Biased

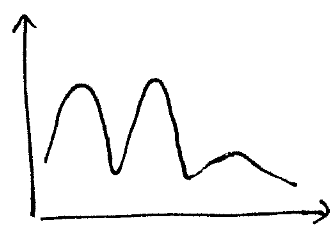
$$W_B(f) = \frac{1}{N}\left(\frac{\sin \pi f N}{\sin \pi f}\right)^2$$



TRUE PSD



DTFT of Bartlett WINDOW



EXPECTED VALUE OF PERIODOGRAM

<u>Asymptotically:</u>

$$\lim_{N\to\infty} W_B(f) = \delta(f) \qquad \text{(DIRAC FUNCTION)}$$

so $\hat{P}_{PER}$ is *asymptotically unbiased.*

Variance of the Periodogram

The main problem with the periodogram is its variance. Consider the case of $x[n]$ being white Gaussian noise.

Recall

$$\text{cov}\{\hat{P}_{PER}(f_1) \cdot \hat{P}_{PER}(f_2)\} = E\{\hat{P}_{PER}(f_1) \cdot \hat{P}_{PER}(f_2)\}$$

$$- E\{\hat{P}_{PER}(f_1)\} E\{\hat{P}_{PER}(f_2)\}.$$

For WGN is can be shown (see Kay, p. 97) that

$$E\{\hat{P}_{PER}(f_1)\hat{P}_{PER}(f_2)\} =$$

$$\sigma_x^4 \left[1 + \left(\frac{\sin N\pi(f_1+f_2)}{N\sin\pi(f_1+f_2)}\right)^2 + \left(\frac{\sin N\pi(f_1-f_2)}{N\sin\pi(f_1-f_2)}\right)^2\right].$$

We also know

$$E\{\hat{P}_{PER}(f_1)\} = E\{\hat{P}_{PER}(f_2)\} = \sigma_x^2.$$

Therefore,

$$\operatorname{cov}\left\{\hat{P}_{PER}(f_1)\ \hat{P}_{PER}(f_2)\right\} = \sigma_x^4\left[\left(\frac{\sin N\pi(f_1+f_2)}{N\sin\pi(f_1+f_2)}\right)^2 + \left(\frac{\sin N\pi(f_1-f_2)}{N\sin\pi(f_1-f_2)}\right)^2\right].$$

Setting $f_1 = f_2$ yields

$$\operatorname{var}\left\{\hat{P}_{PER}(f)\right\} = P_x^2(f)\left[1 + \left(\frac{\sin 2\pi N f}{N\sin 2\pi f}\right)^2\right]$$

For f not near $0, \pm\frac{1}{2}$,

$$\operatorname{var}\left\{\hat{P}_{PER}(f)\right\} \approx P_x^2(f) = \sigma_x^2.$$

Conclusions:

1. The variance of the periodogram
   (a) does not tend to zero as $N \to \infty$
   (b) is on the order of the mean
2. $\hat{P}_{PER}(f)$ is inconsistent!

Similar but more difficult analysis applies to processes other than WGN.

Special Case: White Noise

If $w[k]$ is white noise then

$$\gamma_w[k] = \sigma_w^2 \delta[k]$$

so

$$E\{\hat{P}_{PER}(f)\} = \sum_{k=-\infty}^{\infty} w_B[k] \cdot \sigma_w^2 \delta[k] e^{-i2\pi f k}$$

$$= \sigma_w^2$$

$$= P_w(f)$$

$\Rightarrow$ for white noise, periodogram is unbiased for all N

Interpretation: Convolving a flat line with $W_B(f)$ doesn't change the flat line - it can't be blurred any more.