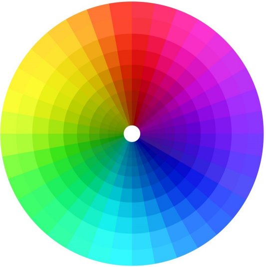
UPPSALA
UNIVERSITET

Nonparametric methods: The Periodogram and the Correlogram



Per Mattsson

Systems and Control
Department of Information Technology
Uppsala University

2019-09-05

Summary

The Spectral Estimation Problem

From a finite dataset $\{y(1), y(2), \ldots, y(N)\}$ of a stationary signal, determine an estimate $\hat{\phi}(\omega)$ of the PSD $\phi(\omega)$ for $\omega \in [-\pi,\ \pi]$

- Power spectral density (PSD)

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t)e^{-i\omega t} \right|^2 \right\}$$

- Autocovariance sequence (ACS)

$$r(k) = E\{y(t)y^*(t-k)\}$$

- DTFT of the ACS gives PSD

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$$

Today

- Estimation theory.
- The Periodogram and Correlogram methods
- Properties and equivalences
- Computing the periodogram in practice.
- Useful Matlab functions.

Time allocation during the course

- Lectures: Present key concepts and methods, Q&A
- Exercise sessions: You/we solve theory problems (prove/show/verify something)
- Labs/homeworks: Simulate things in MATLAB and try the methods
- All over: Discuss properties, interpretations, and intuition

Time allocation during the course

- Lectures: Present key concepts and methods, Q&A
- Exercise sessions: You/we solve theory problems (prove/show/verify something)
- Labs/homeworks: Simulate things in MATLAB and try the methods
- All over: Discuss properties, interpretations, and intuition

Beyond the lectures

- You will need to read more in the book to get the full picture
- The computer labs will let you compare theory and practice in a systematical way
- The homework will require you to code, try different approaches, make engineering decisions, and present your results

UPPSALA UNIVERSITET

Estimation theory

Estimation theory

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.

per.mattsson@it.uu.se

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.
- Often some model connects $\boldsymbol{\theta}$ and $\mathbf{y}$, e.g.,

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{e}$$

per.mattsson@it.uu.se

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.
- Often some model connects $\boldsymbol{\theta}$ and $\mathbf{y}$, e.g.,

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{e}$$

- Uncertainty/noise in measurements: $\mathbf{e} \implies \mathbf{y}$ is stochastic

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.
- Often some model connects $\boldsymbol{\theta}$ and $\mathbf{y}$, e.g.,

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{e}$$

- Uncertainty/noise in measurements: $\mathbf{e} \implies \mathbf{y}$ is stochastic
- Fit model to data.

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.
- Often some model connects $\boldsymbol{\theta}$ and $\mathbf{y}$, e.g.,

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{e}$$

- Uncertainty/noise in measurements: $\mathbf{e} \implies \mathbf{y}$ is stochastic
- Fit model to data. E.g., minimize the least square error:

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmin}_{\boldsymbol{\theta}} \|\mathbf{y} - \mathbf{f}(\boldsymbol{\theta})\|^2$$

UPPSALA UNIVERSITET

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.
- Often some model connects $\boldsymbol{\theta}$ and $\mathbf{y}$, e.g.,

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{e}$$

- Uncertainty/noise in measurements: $\mathbf{e} \implies \mathbf{y}$ is stochastic
- Fit model to data. E.g., minimize the least square error:

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmin}_{\boldsymbol{\theta}} \|\mathbf{y} - \mathbf{f}(\boldsymbol{\theta})\|^2$$

- $\hat{\boldsymbol{\theta}}$ can be seen as stochastic since $\mathbf{y}$ is stochastic.

per.mattsson@it.uu.se

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.
- Often some model connects $\boldsymbol{\theta}$ and $\mathbf{y}$, e.g.,

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{e}$$

- Uncertainty/noise in measurements: $\mathbf{e} \implies \mathbf{y}$ is stochastic
- Fit model to data. E.g., minimize the least square error:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{f}(\boldsymbol{\theta})\|^2$$

- $\hat{\boldsymbol{\theta}}$ can be seen as stochastic since $\mathbf{y}$ is stochastic.
- For any experiment (collected data $\mathbf{y}$) you get one realization $\hat{\boldsymbol{\theta}}$. If you redo the experiment, you will probably get a different $\hat{\boldsymbol{\theta}}$.

per.mattsson@it.uu.se

UPPSALA UNIVERSITET

Estimation theory

- Want to estimate some quantity $\boldsymbol{\theta}$ from data $\mathbf{y}$.
- Often some model connects $\boldsymbol{\theta}$ and $\mathbf{y}$, e.g.,

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{e}$$

- Uncertainty/noise in measurements: $\mathbf{e} \implies \mathbf{y}$ is stochastic
- Fit model to data. E.g., minimize the least square error:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{f}(\boldsymbol{\theta})\|^2$$

- $\hat{\boldsymbol{\theta}}$ can be seen as stochastic since $\mathbf{y}$ is stochastic.
- For any experiment (collected data $\mathbf{y}$) you get one realization $\hat{\boldsymbol{\theta}}$. If you redo the experiment, you will probably get a different $\hat{\boldsymbol{\theta}}$.
- With this view, we can talk about the expected value and variance of $\hat{\boldsymbol{\theta}}$.

Bias, variance, consistency

per.mattsson@it.uu.se

Bias, variance, consistency

- Let $\hat{\theta}$ be the estimate of θ.

per.mattsson@it.uu.se

Bias, variance, consistency

- Let $\hat{\boldsymbol{\theta}}$ be the estimate of $\boldsymbol{\theta}$.
- **Bias:** $E\{\hat{\boldsymbol{\theta}}\} - \boldsymbol{\theta}$

Bias, variance, consistency

- Let $\hat{\boldsymbol{\theta}}$ be the estimate of $\boldsymbol{\theta}$.
- **Bias:** $E\{\hat{\boldsymbol{\theta}}\} - \boldsymbol{\theta}$
- **Variance:** $E\{|\hat{\boldsymbol{\theta}} - E\{\hat{\boldsymbol{\theta}}\}|^2\}$

Bias, variance, consistency

- Let $\hat{\boldsymbol{\theta}}$ be the estimate of $\boldsymbol{\theta}$.
- **Bias:** $E\{\hat{\boldsymbol{\theta}}\} - \boldsymbol{\theta}$
- **Variance:** $E\{|\hat{\boldsymbol{\theta}} - E\{\hat{\boldsymbol{\theta}}\}|^2\}$
- **Mean squared error:**

$$\text{MSE} = E\{|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|^2\} =$$

Bias, variance, consistency

- Let $\hat{\theta}$ be the estimate of θ.
- **Bias:** $E\{\hat{\theta}\} - \theta$
- **Variance:** $E\{|\hat{\theta} - E\{\hat{\theta}\}|^2\}$
- **Mean squared error:**

$$\text{MSE} = E\{|\hat{\theta} - \theta|^2\} = \text{var}\{\hat{\theta}\} + |\text{bias}\{\hat{\theta}\}|^2$$

Bias, variance, consistency

- Let $\hat{\boldsymbol{\theta}}$ be the estimate of $\boldsymbol{\theta}$.
- **Bias:** $E\{\hat{\boldsymbol{\theta}}\} - \boldsymbol{\theta}$
- **Variance:** $E\{|\hat{\boldsymbol{\theta}} - E\{\hat{\boldsymbol{\theta}}\}|^2\}$
- **Mean squared error:**

$$\mathsf{MSE} = E\{|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|^2\} = \mathsf{var}\{\hat{\boldsymbol{\theta}}\} + |\mathsf{bias}\{\hat{\boldsymbol{\theta}}\}|^2$$

- **Consistency:** The way we estimate (our estimator) is consistent if $\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}$ as the number of data points used for estimation goes to infinity.

Parametric vs. non-parametric

Nonparametric (Classical) Methods

No model imposing constraints, or rather, a very general model (i.e. many “parameters”)

Parametric (Modern) Methods

Model the data and estimate the model parameters.

Parametric vs. non-parametric

Nonparametric (Classical) Methods

No model imposing constraints, or rather, a very general model (i.e. many "parameters")

Parametric (Modern) Methods

Model the data and estimate the model parameters.

Trade-offs: (Robustness, and Resolution vs. Variance)

- Parametric methods: better estimates **if** model is good, otherwise the results can be heavily biased.
- Non-parametric methods: works for any dataset, but have relatively high variance, even when $N \to \infty$.
- Resolution – detecting/separating spectral components.

Periodogram and correlogram

Non-parametric PSD estimation

Both

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t) e^{-i\omega t} \right|^2 \right\}$$

and

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k} \quad \text{where} \quad r(k) = E\{y(t)y^*(t-k)\}$$

require an **infinite** number of samples ($N \to \infty$).

Non-parametric PSD estimation

Both

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t) e^{-i\omega t} \right|^2 \right\}$$

and

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k} \quad \text{where} \quad r(k) = E\{y(t)y^*(t-k)\}$$

require an **infinite** number of samples ($N \to \infty$).

- Problem: Finite collection of samples $\{y(1), \dots, y(N)\}$, so cannot compute the sums and expectations.

Non-parametric PSD estimation

Both

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t)e^{-i\omega t} \right|^2 \right\}$$

and

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k} \quad \text{where} \quad r(k) = E\{y(t)y^*(t-k)\}$$

require an **infinite** number of samples ($N \to \infty$).

- Problem: Finite collection of samples $\{y(1), \ldots, y(N)\}$, so cannot compute the sums and expectations.
- Idea:

Non-parametric PSD estimation

Both

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t)e^{-i\omega t} \right|^2 \right\}$$

and

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k} \quad \text{where} \quad r(k) = E\{y(t)y^*(t-k)\}$$

require an **infinite** number of samples ($N \to \infty$).

- Problem: Finite collection of samples $\{y(1), \ldots, y(N)\}$, so cannot compute the sums and expectations.
- Idea: Skip $\lim_{N\to\infty}$ and $E\{\cdot\}$.

per.mattsson@it.uu.se

Non-parametric PSD estimation

Both

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t) e^{-i\omega t} \right|^2 \right\}$$

and

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k} \quad \text{where} \quad r(k) = E\{y(t)y^*(t-k)\}$$

require an **infinite** number of samples ($N \to \infty$).

- Problem: Finite collection of samples $\{y(1), \ldots, y(N)\}$, so cannot compute the sums and expectations.
- Idea: Skip $\lim\limits_{N\to\infty}$ and $E\{\cdot\}$.
- Idea used by Schuster (~ 1900) to determine "hidden periodicities".

Non-parametric PSD estimation

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t)e^{-i\omega t} \right|^2 \right\} = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$$

Non-parametric PSD estimation

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N} \left| \sum_{t=1}^{N} y(t) e^{-i\omega t} \right|^2 \right\} = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}$$

- Periodogram: The first definition gives

$$\hat{\phi}_p(\omega) = \frac{1}{N} \left| \sum_{t=1}^{N} y(t) e^{-i\omega t} \right|^2 .$$

Non-parametric PSD estimation

$$\phi(\omega) = \lim_{N\to\infty} E\left\{ \frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2 \right\} = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$$

- Periodogram: The first definition gives

$$\hat{\phi}_p(\omega) = \frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2.$$

- Correlogram: The second definition gives

$$\hat{\phi}_c(\omega) = \sum_{k=-(N-1)}^{N-1} \hat{r}(k)e^{-i\omega k},$$

where $\hat{r}(k)$ is an estimate of $r(k) = E\{y(t)y^*(t-k)\}$.

Non-parametric PSD estimation

$$\phi(\omega) = \lim_{N\to\infty} E\left\{\frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2\right\} = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$$

- Periodogram: The first definition gives

$$\hat{\phi}_p(\omega) = \frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2.$$

- Correlogram: The second definition gives

$$\hat{\phi}_c(\omega) = \sum_{k=-(N-1)}^{N-1} \hat{r}(k)e^{-i\omega k},$$

  where $\hat{r}(k)$ is an estimate of $r(k) = E\{y(t)y^*(t-k)\}$.
- Finite sums $\Longrightarrow$ can be computed in practice.

UPPSALA UNIVERSITET

Estimation of ACS

Standard unbiased estimate:

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \leq k \leq N-1.$$

per.mattsson@it.uu.se

Estimation of ACS

Standard unbiased estimate:

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \le k \le N-1.$$

Problems:

- For k large (close to N), $\hat{r}(k)$ has high variance (erratic).

UPPSALA UNIVERSITET

Estimation of ACS

Standard unbiased estimate:

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \le k \le N-1.$$

Problems:

- For k large (close to N), $\hat{r}(k)$ has high variance (erratic).
- May result in negative spectrum.

per.mattsson@it.uu.se

Estimation of ACS

Standard unbiased estimate:

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \leq k \leq N-1.$$

Problems:

- For k large (close to N), $\hat{r}(k)$ has high variance (erratic).
- May result in negative spectrum.

Standard biased estimate:

$$\boxed{\hat{r}(k) = \frac{1}{N} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \leq k \leq N-1.}$$

UPPSALA UNIVERSITET

Estimation of ACS

Standard unbiased estimate:

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \le k \le N-1.$$

Problems:

- For k large (close to N), $\hat{r}(k)$ has high variance (erratic).
- May result in negative spectrum.

Standard biased estimate:

$$\boxed{\hat{r}(k) = \frac{1}{N} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \le k \le N-1.}$$

- $r(k)$ is typically small for large k, and bias is low for small k.

per.mattsson@it.uu.se

Estimation of ACS

Standard unbiased estimate:

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} y(t)y^*(t-k), \quad 0 \le k \le N-1.$$

Problems:

- For k large (close to N), $\hat{r}(k)$ has high variance (erratic).
- May result in negative spectrum.

Standard biased estimate:

$$\boxed{\hat{r}(k) = \frac{1}{N} \sum_{t=k+1}^{N} y(t)y^*(t-k), \quad 0 \le k \le N-1.}$$

- $r(k)$ is typically small for large k, and bias is low for small k.
- $\hat{r}(k)$ positive semidefinite $\Longrightarrow$ non-negative spectrum.

Estimation of ACS

Standard unbiased estimate:

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \le k \le N-1.$$

Problems:

- For k large (close to N), $\hat{r}(k)$ has high variance (erratic).
- May result in negative spectrum.

Standard biased estimate:

$$\boxed{\hat{r}(k) = \frac{1}{N} \sum_{t=k+1}^{N} y(t) y^*(t-k), \quad 0 \le k \le N-1.}$$

- $r(k)$ is typically small for large k, and bias is low for small k.
- $\hat{r}(k)$ positive semidefinite $\implies$ non-negative spectrum.

For both estimates: $\hat{r}(k) = \hat{r}^*(-k)$ for $k < 0$.

UPPSALA UNIVERSITET

Equivalence

If the biased ACS estimator $\hat{r}(k)$ is used in $\hat{\phi}_c(\omega)$, then

$$\hat{\phi}_p(\omega) = \frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2 = \sum_{k=-(N-1)}^{N-1} \hat{r}(k)e^{-i\omega k} = \hat{\phi}_c(\omega)$$

$$\boxed{\hat{\phi}_p(\omega) = \hat{\phi}_c(\omega)}$$

- Can analyze the estimates simultaneously!

Analysis

Analysis of the periodogram

- Since $\phi_p(\omega) = \phi_c(\omega)$ we can use both formulations in the analysis.
- Bias-variance tradeoff:

$$\mathsf{MSE} = \mathsf{var}(\hat{\phi}_p(\omega)) + |\mathsf{bias}(\hat{\phi}_p(\omega))|^2.$$

- Wanted: Small bias, small variance, consistent.

The bias of the periodogram

- With the biased ACS estimate we get

$$E\{\hat{\phi}_c(\omega)\} =$$

per.mattsson@it.uu.se

The bias of the periodogram

- With the biased ACS estimate we get

$$E\{\hat{\phi}_c(\omega)\} = \sum_{k=-(N-1)}^{N-1} E\{\hat{r}(k)\}e^{-i\omega k} = \text{[Board]}$$

$$=$$

UPPSALA UNIVERSITET

The bias of the periodogram

- With the biased ACS estimate we get

$$
\begin{aligned}
E\{\hat{\phi}_c(\omega)\} &= \sum_{k=-(N-1)}^{N-1} E\{\hat{r}(k)\}e^{-i\omega k} = \text{[Board]} \\
&= \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N}\right) r(k)e^{-i\omega k} \\
&=
\end{aligned}
$$

per.mattsson@it.uu.se

The bias of the periodogram

- With the biased ACS estimate we get

$$E\{\hat{\phi}_c(\omega)\} = \sum_{k=-(N-1)}^{N-1} E\{\hat{r}(k)\}e^{-i\omega k} = \text{[Board]}$$
$$= \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N}\right) r(k)e^{-i\omega k}$$
$$= \sum_{k=-\infty}^{\infty} w_B(k) r(k) e^{-i\omega k},$$

where the Bartlett (triangular) window is given by

$$w_B(k) = \begin{cases} 1 - \frac{|k|}{N}, & |k| \le N-1 \\ 0, & |k| \ge N \end{cases}.$$

The bias of the periodogram

- With the biased ACS estimate we get

$$E\{\hat{\phi}_c(\omega)\} = \sum_{k=-(N-1)}^{N-1} E\{\hat{r}(k)\}e^{-i\omega k} = \text{[Board]}$$
$$= \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N}\right) r(k)e^{-i\omega k}$$
$$= \sum_{k=-\infty}^{\infty} w_B(k)r(k)e^{-i\omega k},$$

where the Bartlett (triangular) window is given by

$$w_B(k) = \begin{cases} 1 - \frac{|k|}{N}, & |k| \le N-1 \\ 0, & |k| \ge N \end{cases}.$$

- Large bias for small N since $w_B(k) \neq 1$.

The bias of the periodogram

- $E\{\hat{\phi}_c(\omega)\}$ is the DTFT of $w_B(k)r(k)$.

per.mattsson@it.uu.se

The bias of the periodogram

- $E\{\hat{\phi}_c(\omega)\}$ is the DTFT of $w_B(k)r(k)$.
- Time-domain multiplication $\iff$ Frequency-domain convolution

The bias of the periodogram

- $E\{\hat{\phi}_c(\omega)\}$ is the DTFT of $w_B(k)r(k)$.
- Time-domain multiplication $\iff$ Frequency-domain convolution:

$$E\{\hat{\phi}(\omega)\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\phi(\zeta)W_B(\omega-\zeta)d\zeta$$

The bias of the periodogram

- $E\{\hat{\phi}_c(\omega)\}$ is the DTFT of $w_B(k)r(k)$.
- Time-domain multiplication $\iff$ Frequency-domain convolution:

$$E\{\hat{\phi}(\omega)\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(\zeta) W_B(\omega - \zeta) d\zeta,$$

where the DTFT of w_B is given by

$$W_B(\omega) = \frac{1}{N}\left(\frac{\sin(\omega N/2)}{\sin(\omega/2)}\right)^2.$$

The bias of the periodogram

- $E\{\hat{\phi}_c(\omega)\}$ is the DTFT of $w_B(k)r(k)$.
- Time-domain multiplication $\iff$ Frequency-domain convolution:

$$E\{\hat{\phi}(\omega)\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(\zeta)W_B(\omega-\zeta)d\zeta,$$

  where the DTFT of w_B is given by

$$W_B(\omega) = \frac{1}{N}\left(\frac{\sin(\omega N/2)}{\sin(\omega/2)}\right)^2.$$

- Ideally $W_B(\omega)$ should approximate a Dirac impulse $\delta(\omega)$.

The Bartlett window

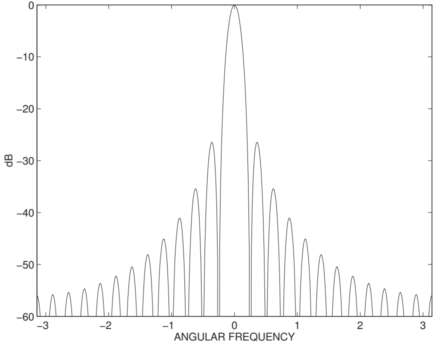


Figure 2.1. $W_B(\omega)/W_B(0)$, for $N = 25$.

The Bartlett window

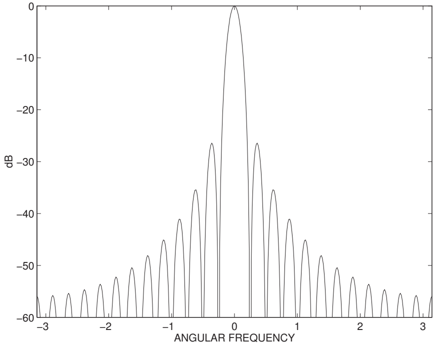


Figure 2.1. $W_B(\omega)/W_B(0)$, for $N = 25$.

- Main lobe 3dB width $\approx 2\pi/N$ radians (or $1/N$ in frequency).

The Bartlett window

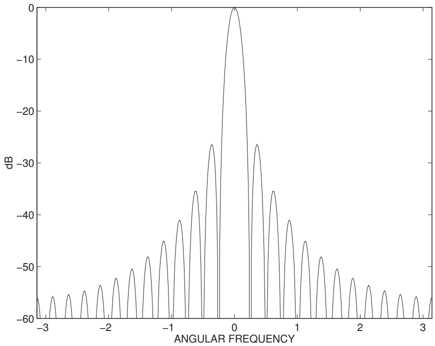


Figure 2.1. $W_B(\omega)/W_B(0)$, for $N = 25$.

- Main lobe 3dB width $\approx 2\pi/N$ radians (or $1/N$ in frequency).
- For small N, quite different from $\delta(\omega)$ (severe bias).

The Bartlett window

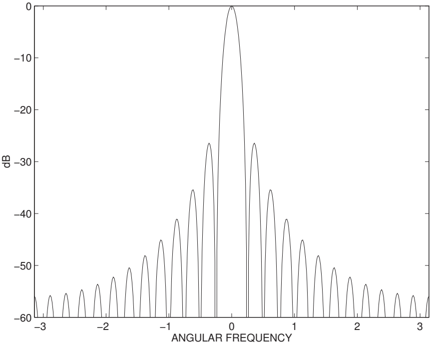


Figure 2.1. $W_B(\omega)/W_B(0)$, for $N = 25$.

- Main lobe 3dB width $\approx 2\pi/N$ radians (or $1/N$ in frequency).
- For small N, quite different from $\delta(\omega)$ (severe bias).
- Non-zero main lobe $\Longrightarrow$ Smearing (smoothing).

The Bartlett window

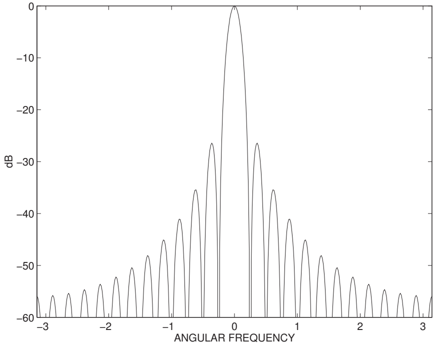


Figure 2.1. $W_B(\omega)/W_B(0)$, for $N = 25$.

- Main lobe 3dB width $\approx 2\pi/N$ radians (or $1/N$ in frequency).
- For small N, quite different from $\delta(\omega)$ (severe bias).
- Non-zero main lobe $\Longrightarrow$ Smearing (smoothing).
- Sidelobes $\Longrightarrow$ Leakage.

Summary of the periodogram bias properties

- For small N, severe bias.

Summary of the periodogram bias properties

- For small N, severe bias.
- However,

$$\lim_{N\to\infty} W_B(\omega) = 2\pi\delta(\omega),$$

Summary of the periodogram bias properties

- For small N, severe bias.
- However,
$$\lim_{N\to\infty} W_B(\omega) = 2\pi\delta(\omega),$$
so,
$$\lim_{N\to\infty} E\{\hat{\phi}_p(\omega)\} = \int_{-\pi}^{\pi} \phi(\zeta)\delta(\omega-\zeta)d\zeta =$$

Summary of the periodogram bias properties

- For small N, severe bias.
- However,

$$\lim_{N\to\infty} W_B(\omega) = 2\pi\delta(\omega),$$

so,

$$\lim_{N\to\infty} E\{\hat{\phi}_p(\omega)\} = \int_{-\pi}^{\pi} \phi(\zeta)\delta(\omega-\zeta)d\zeta = \phi(\omega).$$

UPPSALA UNIVERSITET

Summary of the periodogram bias properties

- For small N, severe bias.
- However,

$$\lim_{N \to \infty} W_B(\omega) = 2\pi\delta(\omega),$$

so,

$$\lim_{N \to \infty} E\{\hat{\phi}_p(\omega)\} = \int_{-\pi}^{\pi} \phi(\zeta)\delta(\omega - \zeta)d\zeta = \phi(\omega).$$

- That is, the periodogram is asymptotically unbiased.

UPPSALA UNIVERSITET

Summary of the periodogram bias properties

- For small N, severe bias.
- However,

$$\lim_{N\to\infty} W_B(\omega) = 2\pi\delta(\omega),$$

so,

$$\lim_{N\to\infty} E\{\hat{\phi}_p(\omega)\} = \int_{-\pi}^{\pi} \phi(\zeta)\delta(\omega-\zeta)d\zeta = \phi(\omega).$$

- That is, the periodogram is asymptotically unbiased.
- If bias was the only problem with periodogram, we could solve it by increasing the number of data points used. However....

The asymptotic variance/covariance

- As $N \to \infty$,

$$E\left\{\left[\hat{\phi}_p(\omega_1) - \phi(\omega_1)\right]\left[\hat{\phi}_p(\omega_2) - \phi(\omega_2)\right]\right\}$$

$$=$$

The asymptotic variance/covariance

- As $N \to \infty$,

$$E\left\{\left[\hat{\phi}_p(\omega_1) - \phi(\omega_1)\right]\left[\hat{\phi}_p(\omega_2) - \phi(\omega_2)\right]\right\} = \begin{cases} \phi^2(\omega_1), & \omega_1 = \omega_2 \\ 0, & \omega_1 \neq \omega_2 \end{cases}$$

UPPSALA UNIVERSITET

The asymptotic variance/covariance

- As $N \to \infty$,

$$E\left\{\left[\hat{\phi}_p(\omega_1) - \phi(\omega_1)\right]\left[\hat{\phi}_p(\omega_2) - \phi(\omega_2)\right]\right\}$$
$$= \begin{cases} \phi^2(\omega_1), & \omega_1 = \omega_2 \\ 0, & \omega_1 \neq \omega_2 \end{cases}$$

- Inconsistent estimate. In general $\lim_{N\to\infty} \hat{\phi}_p(\omega) \neq \phi(\omega)$.

UPPSALA UNIVERSITET

The asymptotic variance/covariance

- As $N \to \infty$,

$$E\left\{\left[\hat{\phi}_p(\omega_1) - \phi(\omega_1)\right]\left[\hat{\phi}_p(\omega_2) - \phi(\omega_2)\right]\right\} = \begin{cases} \phi^2(\omega_1), & \omega_1 = \omega_2 \\ 0, & \omega_1 \neq \omega_2 \end{cases}$$

- Inconsistent estimate. In general $\lim_{N\to\infty} \hat{\phi}_p(\omega) \neq \phi(\omega)$.
- Erratic behavior. Estimate at ω_1 uncorrelated (for large N) from estimate at $\omega_2 \neq \omega_1$.

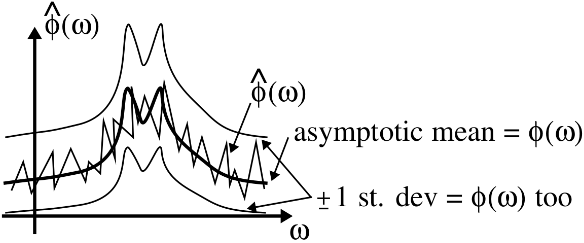
per.mattsson@it.uu.se

The asymptotic variance/covariance

- As $N \to \infty$,

$$E\left\{\left[\hat{\phi}_p(\omega_1) - \phi(\omega_1)\right]\left[\hat{\phi}_p(\omega_2) - \phi(\omega_2)\right]\right\} = \begin{cases} \phi^2(\omega_1), & \omega_1 = \omega_2 \\ 0, & \omega_1 \neq \omega_2 \end{cases}$$

- Inconsistent estimate. In general $\lim_{N\to\infty} \hat{\phi}_p(\omega) \neq \phi(\omega)$.
- Erratic behavior. Estimate at ω_1 uncorrelated (for large N) from estimate at $\omega_2 \neq \omega_1$.



Summary of periodogram properties

$$\hat{\phi}_p = \frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2 = \sum_{k=-(N-1)}^{N-1} \hat{r}(k)e^{-i\omega k} = \phi_c(\omega)$$

- Asymptotically unbiased:

$$\lim_{N\to\infty} E\{\hat{\phi}_p(\omega)\} = \phi(\omega).$$

- Large variance, even as $N \to \infty$.
- Thus, $\hat{\phi}(\omega)$ has bad performance.

Summary of periodogram properties

$$\hat{\phi}_p = \frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2 = \sum_{k=-(N-1)}^{N-1} \hat{r}(k)e^{-i\omega k} = \phi_c(\omega)$$

- Asymptotically unbiased:

$$\lim_{N\to\infty} E\{\hat{\phi}_p(\omega)\} = \phi(\omega).$$

- Large variance, even as $N \to \infty$.
- Thus, $\hat{\phi}(\omega)$ has bad performance.

Intuitive explanation:

- $\hat{r}(k) - r(k)$ may be large for large $|k|$.
- Even if this is not case, the sum of many small errors can become large.

UPPSALA UNIVERSITET

Computing the periodogram

Computing the periodogram

$$\phi_p(\omega) = \frac{1}{N}|Y_N(\omega)|^2, \quad Y_N(\omega) = \sum_{t=1}^{N} y(t)e^{-i\omega t}.$$

Computing the periodogram

$$\phi_p(\omega) = \frac{1}{N}|Y_N(\omega)|^2, \quad Y_N(\omega) = \sum_{t=1}^{N} y(t)e^{-i\omega t}.$$

- Can only compute for a finite number of ω.

Computing the periodogram

$$\phi_p(\omega) = \frac{1}{N}|Y_N(\omega)|^2, \quad Y_N(\omega) = \sum_{t=1}^{N} y(t)e^{-i\omega t}.$$

- Can only compute for a finite number of ω.
- For large N, heavy computations (sum of N complex multiplications for each frequency).

Computing the periodogram

$$\phi_p(\omega) = \frac{1}{N}|Y_N(\omega)|^2, \quad Y_N(\omega) = \sum_{t=1}^{N} y(t)e^{-i\omega t}.$$

- Can only compute for a finite number of ω.
- For large N, heavy computations (sum of N complex multiplications for each frequency).

Common solution:

- Compute for $\omega = \frac{2\pi}{N}k$, $k = 0, 1, \ldots, N-1$.

Computing the periodogram

$$\phi_p(\omega) = \frac{1}{N}|Y_N(\omega)|^2, \quad Y_N(\omega) = \sum_{t=1}^{N} y(t)e^{-i\omega t}.$$

- Can only compute for a finite number of ω.
- For large N, heavy computations (sum of N complex multiplications for each frequency).

Common solution:

- Compute for $\omega = \frac{2\pi}{N}k$, $k = 0, 1, \ldots, N-1$.
- Then we obtain the DFT,

$$Y(k) = \sum_{t=1}^{N} y(t)e^{-i2\pi kt/N}.$$

UPPSALA UNIVERSITET

Computing the periodogram

$$\phi_p(\omega) = \frac{1}{N}|Y_N(\omega)|^2, \quad Y_N(\omega) = \sum_{t=1}^{N} y(t)e^{-i\omega t}.$$

- Can only compute for a finite number of ω.
- For large N, heavy computations (sum of N complex multiplications for each frequency).

Common solution:

- Compute for $\omega = \frac{2\pi}{N}k$, $k = 0, 1, \ldots, N-1$.
- Then we obtain the DFT,

$$Y(k) = \sum_{t=1}^{N} y(t)e^{-i2\pi kt/N}.$$

- Can be computed using Fast Fourier Transform (FFT) (computations in the order of $N \log_2 N$ vs N^2 for direct approach). Often assumes $N = 2^m$.

Zeropadding

- Append the data vector given to FFT with zeros,

$$\{\underbrace{y(1), \ldots, y(N), 0, \ldots 0}_{L}\}$$

Zeropadding

- Append the data vector given to FFT with zeros,

$$\{\underbrace{y(1),\ldots,y(N),0,\ldots 0}_{L}\}$$

- Does not change $\hat{\phi}_p(\omega)$,

$$\hat{\phi}_p(\omega) = \frac{1}{N}\left|\sum_{t=1}^{L} y(t)e^{-i\omega t}\right|^2, \quad \text{with } y(t) = 0 \text{ for } t > N.$$

Zeropadding

- Append the data vector given to FFT with zeros,

$$\{\underbrace{y(1),\ldots,y(N),0,\ldots 0}_{L}\}$$

- Does not change $\hat{\phi}_p(\omega)$,

$$\hat{\phi}_p(\omega) = \frac{1}{N}\left|\sum_{t=1}^{L} y(t)e^{-i\omega t}\right|^2, \quad \text{with } y(t) = 0 \text{ for } t > N.$$

- Can ensure that $L = 2^m$ (not necessary for `fft` in Matlab).

Zeropadding

- Append the data vector given to FFT with zeros,

$$\underbrace{\{y(1), \ldots, y(N), 0, \ldots 0\}}_{L}$$

- Does not change $\hat{\phi}_p(\omega)$,

$$\hat{\phi}_p(\omega) = \frac{1}{N} \left| \sum_{t=1}^{L} y(t) e^{-i\omega t} \right|^2, \quad \text{with } y(t) = 0 \text{ for } t > N.$$

- Can ensure that $L = 2^m$ (not necessary for `fft` in Matlab).
- Increases the number of frequencies computed by FFT,

$$\omega = \frac{2\pi}{N} k, \quad k = 0, 1, \ldots, L-1.$$

Zeropadding

- Append the data vector given to FFT with zeros,

$$\{\underbrace{y(1), \ldots, y(N), 0, \ldots 0}_{L}\}$$

- Does not change $\hat{\phi}_p(\omega)$,

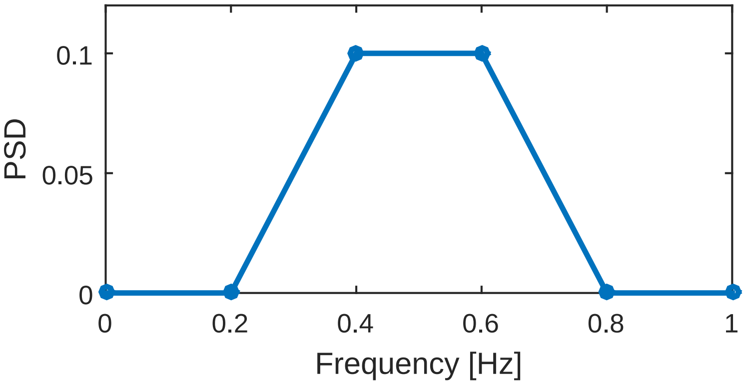
$$\hat{\phi}_p(\omega) = \frac{1}{N} \left| \sum_{t=1}^{L} y(t) e^{-i\omega t} \right|^2, \quad \text{with } y(t) = 0 \text{ for } t > N.$$

- Can ensure that $L = 2^m$ (not necessary for `fft` in Matlab).
- Increases the number of frequencies computed by FFT,

$$\omega = \frac{2\pi}{N} k, \quad k = 0, 1, \ldots, L-1.$$

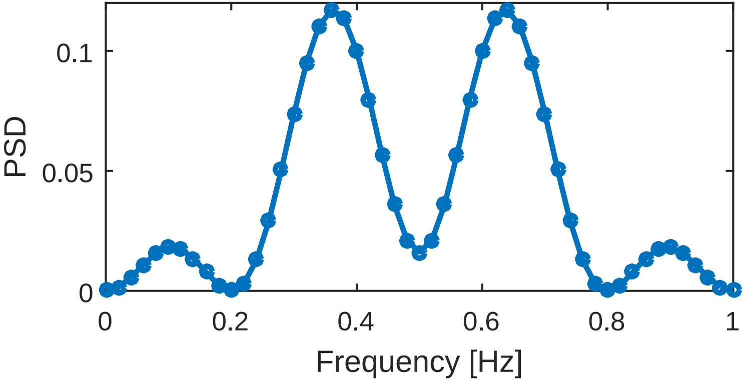
- There are FFT algorithms that eliminate operations on zeros.

Example: Zeropadding



6 samples (no zeropadding)

Example: Zeropadding



60 samples ($L = 10N$ zeropadding)

Summary

- The periodogram and correlogram.
- Smearing and leakage.
- Asymptotically unbiased.
- "High" variance even when $N \to \infty$.
- Using the FFT for computing the periodogram.

Summary

- The periodogram and correlogram.
- Smearing and leakage.
- Asymptotically unbiased.
- "High" variance even when $N \to \infty$.
- Using the FFT for computing the periodogram.

Try the following:

- Simulate a sinusoid/complex exponential in noise and plot the Periodogram (compute the periodogram manually and with FFT!)
- Add another periodic component, study the effect of smearing.
- Exercise 1.5, 1.7, 1.10, 2.3, 2.6

UPPSALA UNIVERSITET

MATLAB

Useful functions:

- `randn(N,1)` – Generates Gaussian white noise.
- `fft(y,L)` – Performs FFT padded with zeros.
- `fftshift(phi)` – Shifts frequencies from $[0, 2\pi]$ to $[-\pi, \pi]$.
- `periodogramse(y,v,L)` – (Windowed) periodogram.
- `filter(B,A,e)` – Generate $y(t) = \frac{B(q)}{A(q)}e(t)$. `A` and `B` are vectors with the coefficients of the polynomials $A(q), B(q)$.

MATLAB

Useful functions:

- `randn(N,1)` – Generates Gaussian white noise.
- `fft(y,L)` – Performs FFT padded with zeros.
- `fftshift(phi)` – Shifts frequencies from $[0, 2\pi]$ to $[-\pi, \pi]$.
- `periodogramse(y,v,L)` – (Windowed) periodogram.
- `filter(B,A,e)` – Generate $y(t) = \frac{B(q)}{A(q)}e(t)$. `A` and `B` are vectors with the coefficients of the polynomials $A(q), B(q)$.

With `fft`, the corresponding normalized frequencies are

```
» omega = 0:2*pi/L:2*pi(L-1)/L; %or
» omega = (0:L-1)*2*pi/L;
```

MATLAB

Useful functions:

- `randn(N,1)` – Generates Gaussian white noise.
- `fft(y,L)` – Performs FFT padded with zeros.
- `fftshift(phi)` – Shifts frequencies from $[0, 2\pi]$ to $[-\pi, \pi]$.
- `periodogramse(y,v,L)` – (Windowed) periodogram.
- `filter(B,A,e)` – Generate $y(t) = \frac{B(q)}{A(q)} e(t)$. `A` and `B` are vectors with the coefficients of the polynomials $A(q), B(q)$.

With `fft`, the corresponding normalized frequencies are

```
» omega = 0:2*pi/L:2*pi(L-1)/L; %or
» omega = (0:L-1)*2*pi/L;
```

If you have also applied `fftshift`, the normalized frequencies are

```
» omega = (0:L-1)*2*pi/L - pi;
```