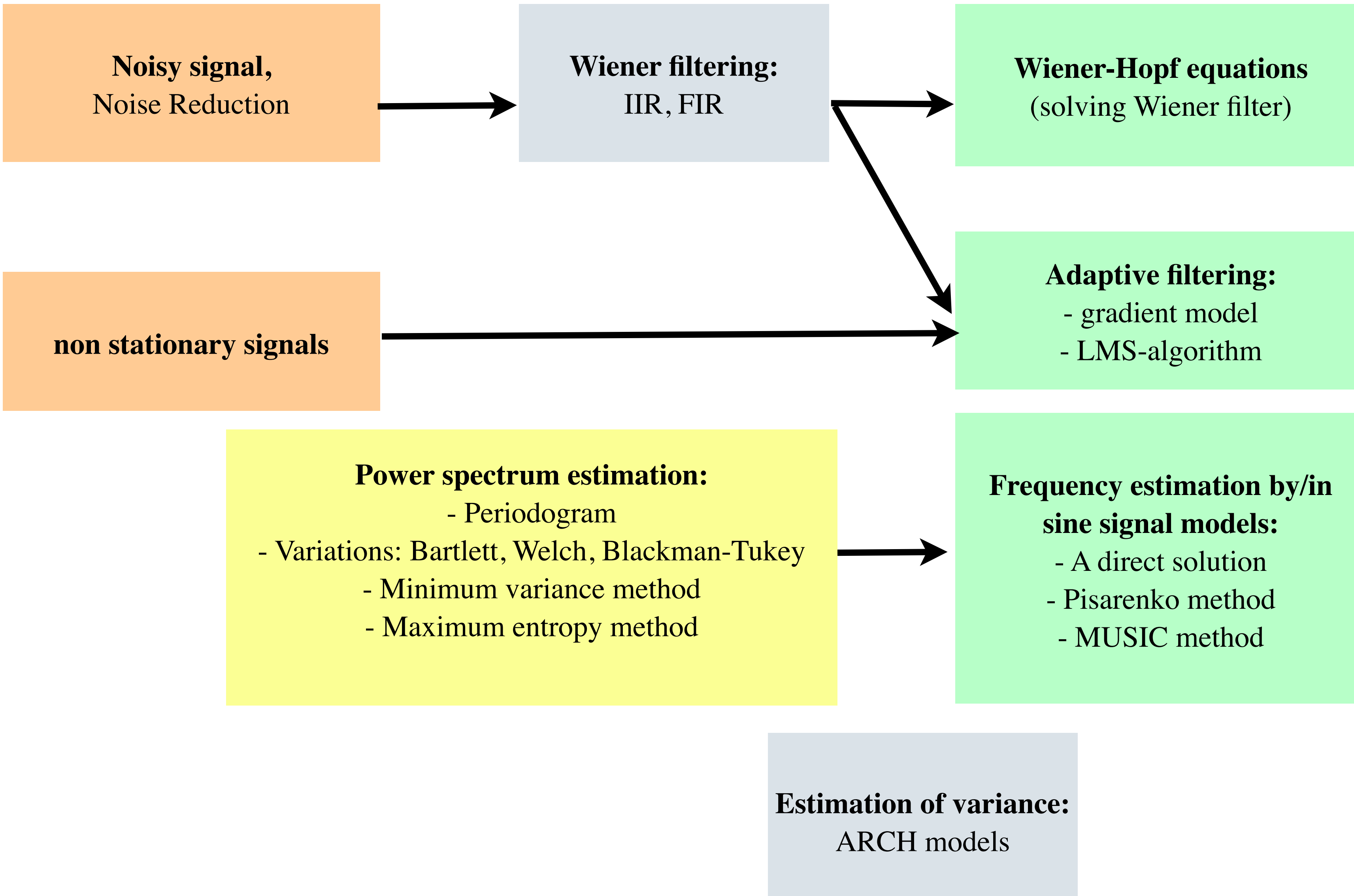
Aalto University

# T61.3040

## Estimating the power spectrum with nonparametric methods

Amaury Lendasse and
Yoan Miche

# Diagram of content of the final part of the course



# Today

**Power spectrum estimation:**
- Periodogram
- Variations: Bartlett, Welch, Blackman-Tukey
- Minimum variance method
- Maximum entropy method

→

**Frequency estimation by/in sine signal models:**
- A direct solution
- Pisarenko method
- MUSIC method

- Autocorrelation of WSS process $x(n)$ can be presented in power spectrum in the frequency domain
- Both the presentation containing the same information of the WSS process
- For many applications the frequency representation is more useful
- eg electric motor fault diagnosis: some defects appear at certain frequencies as the power of the frequency increases

- Example: beamforming, direction of arrival
- the signal level is detected at points (0, 0), (1, 0), (2, 0). . . , (M, 0)
- A known sine signal (constant speed and frequency) of long-distance arrives at different moments at these points
- Delay depends on the direction of the incoming signal and change the frequency perceived by the sensors
- by estimating the frequency, one can calculate the arrival direction of the signal

- Let's repeat the definition of power spectrum of WSS process $x(n)$:

$$P_x(\exp(j\omega)) = \sum_{k=-\infty}^{\infty} r_x(k) \exp(-j\omega k)$$

- If the correct autocorrelation is known, then the power spectrum is defined
- In practice, observations $x(0). . . , x(n\text{-}1)$
- Value of the power spectrum $P_x(\exp(jw))$ has to be estimated by means of observations

- An obvious way to estimate the $P_x$ is to use the estimated values of autocorrelation
- We estimate the autocorrelation by the autocorrelation method
- The power spectrum is obtained by the Fourier transform of $\hat{r}_x(k)$
- This estimate is called the periodogram (obtained from transform of the all the autocorrelations of the observations)

# Periodogram

- Based directly from the definition of  the power spectrum
- Nonparametric method: need only WSS and ergodicity
- The method has some *substantial* weaknesses
- These can be mitigated somewhat with variations of the periodogram
- demo: sunspots.R

# Periodogram

- Estimate the autocorrelation, from the observations $x(0), x(1). . . , x(n\text{-}1)$

- Autocorrelation method gives this estimates

$$\hat{r}_x(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} x(n+k)x^*(n),\ k = 0, 1, \ldots, N-1$$

- The number of observations restricts the amount of estimates

# Periodogram

- Since the autocorrelation is obtained as conjugate symmetric

$$\hat{r}_x(-k) = \hat{r}_x^*(k), \quad k = 1, 2, \dots, N-1$$

- Setting the rest to zero, the autocorrelation is

$$\hat{r}_x(k) = \begin{cases} \dfrac{1}{N} \displaystyle\sum_{n=0}^{N-1-k} x(n+k)x^*(n) & k = 0, 1, \dots, N-1 \\ \hat{r}_x^*(-k) & k = -1, -2, \dots, -N+1 \\ 0 & |k| \geq N \end{cases}$$

# Periodogram

- Each value of the autocorrelation is now defined
- There is a finite number of non-zero autocorrelations, so we can in practice calculate an estimate the Fourier transform of the autocorrelation sequence
- It yields to a periodogram

$$\hat{P}_x(\exp(j\omega)) = \sum_{k=-N+1}^{N-1} \hat{r}_x(k)\exp(-j\omega k), \quad \omega \in (-\pi, \pi]$$

# Periodogram

- The periodogram can be easily calculated using the FFT algorithm
- Defining an infinite sequence of observations

$$x_N(n) = \begin{cases} x(n) & 0 \leq n < N \\ 0 & \text{otherwise} \end{cases}$$

- Autocorrelation method is a convolution (to be check as an exercise)

$$\hat{r}_x(k) = \frac{1}{N} \sum_{n=-\infty}^{\infty} x_N(n+k) x_N^*(n) = \frac{1}{N} x_N(k) * x_N^*(-k)$$

# Periodogram

- Fourier-transform of the autocorrelation

$\hat{r}_x(k) = \frac{1}{N} x_N(k) * x_N^*(-k)$:

$$\begin{aligned}\hat{P}_x &= \frac{1}{N} X_N X_N^* \\ &= \frac{1}{N} |X_N|^2\end{aligned}$$

$X_N = \sum_{n=0}^{N-1} x(n) \exp(-jn\omega)$ Fourier transform of sequence $x_N$

- The periodogram FFT algorithm:

$$x(n) \stackrel{FFT}{\rightarrow} X_N(k) \rightarrow \frac{1}{N} |X_N(k)|^2 = \hat{P}_x(\exp(j2\pi k/N))$$

# Periodogram

- Example: the power spectrum of the white noise $v(n)$
- From The definition of power spectrum => $P_V$ is a constant:

$$
\begin{aligned}
P_v(\exp(j\omega)) &= \sum r_v(k)\exp(-j\omega k) \\
&= r_v(0)\exp(-j\omega 0) = \sigma_v^2
\end{aligned}
$$

- Simulate the white noise and calculate the periodogram
- Demo: whitenoise.R

# Periodogram

- periodogram performed poorly at least with a small number of observations
- Is it worth to increase The number of observations like in general for estimation?
- Let’s take as criteria: MSE = variance + bias
- We need the mean and variance of the periodogram

# Periodogram

- Expected value of the estimated autocorrelation is

$$\begin{aligned} \mathrm{E}(\hat{r}_x(k)) &= \frac{N-k}{N} r_x(k) \\ &= w_B(k) r_x(k) \end{aligned}$$

whith Bartlett window

$$w_B(k) = \begin{cases} \frac{N-|k|}{N}, & |k| \leq N \\ 0, & |k| > N \end{cases}$$

# Periodogram

- Since the periodogram is the Fourier transform of $r_x(k)$, convolution theorem gives

$$\mathrm{E}(\hat{P}_x(\exp(j\omega))) = \frac{1}{2\pi} P_x(\exp(j\omega)) * W_B(\exp(j\omega))$$

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

# Periodogram

- $W_B(\exp(j\omega)) \to 2\pi\delta(\omega)$ when *N* tends to infinity. Derived from: Parseval's formula, using a rectangular window, and calculating $\int_{\epsilon}^{\pi} W_B(\exp(j\omega))d\omega$
- $\delta(\omega)$ satisfies $x(\omega) = x(\omega) * \delta(\omega)$ for all $x(\omega)$
- Then $\frac{1}{2\pi} P_x * 2\pi\delta(\omega) = P_x * \delta(\omega) = P_x$
  and periodogram is asymptotically unbiased
- demo: bartdft.R

# Periodogram

- Example: the noisy sine signal

$$x(n) = A\sin(n\omega_0 + \phi) + v(n), \quad A = 5, \quad \omega_0 = 0.4\pi$$

- Correct autocorrelation is

$$r_x(k) = \frac{A^2}{2}\cos(k\omega_0) + \sigma^2\delta(k)$$

- Correct power spectrum is

$$P_x(\exp(j\omega)) = \frac{1}{2}\pi A^2(\delta(\omega - \omega_0) + \delta(\omega + \omega_0)) + \sigma^2$$

# Periodogram

- We simulate 50 realizations (with random phase), and
- We calculate the average of the periodogram when $N = 64$ and $N = 1024$
- We find that a larger amount of samples gives a better estimate of periodogram
- Demo: bartsin1.R

# Periodogram

- We repeat for two sine signals

$$x(n) = A_1 \sin(n\omega_1 + \phi_1) + A_2 \sin(n\omega_2 + \phi_2) + v(n)$$

- When the Bartlett window is wider than sine frequency difference, then the periodogram does not differentiate between the two sine signals
- periodogram resolution is proportional to $1/N$
- Demo: bartsin2.R

# Periodogram variance

- We have seen that $\mathrm{E}(\hat{P}_x) \rightarrow P_x$ when *N* tends to infinity
- What happens to the variance $\mathrm{var}(\hat{P}_x)$, when *N* tends to infinity?
- It can be shown that for normally distributed white noise (variance $\sigma^2$) periodogram variance is

$$\mathrm{var}[\hat{P}_v(\exp(j\omega))] = \sigma^4 = P_v^2(\exp(j\omega))$$

With normal distributed process *x*(*n*), it holds approximately that $\mathrm{var}[\hat{P}_x] \approx P_x^2$

# Periodogram variance

- The variance of the periodogram does not decrease when $N$ increases!
- This is a surprising feature, which leads to problems
- The periodogram can be improved to reduce the variance, but then other properties will deteriorate
- Demo: periovar.R

# Improved nonparametric methods

- The periodogram is based directly on the definition of power spectrum
- Method can be changed by estimation of the autocorrelation, as well as the observations can be divided into shorter intervals
- We get variations, which decreases the variance of the periodogram when the number of observations grows

# Improved nonparametric methods

- because of these variations are based on the Fourier transform of the autocorrelation, then the properties of these methods can be assessed using the properties of the periodogram:
  - The variance does not depend on $N$
  - The resolution is proportional to $1/N$
- For the modified methods, the Fourier transform of the autocorrelations, which is usually estimated from only part of the observations, so $N$ is replaced by the number of observations used in this case: $L<N$

# Improved nonparametric methods

- In periodogram the resolution improves when $N$ increases
- Variance is not reduced when $N$ increases, but it also does not increase when $N$ decreases
- If the observation is enough, we can calculate several periodograms
- In general, for such as estimation, the variance can be reduced by averaging the independent (or uncorrelated) observations

# Bartlett method

- Dividing *N* observation in *K* parts, where *L*=*N*/*K* observations
- eg $N = 10, K = 2$:

$$x_1(n) = x(n)$$
$$x_2(n) = x(n+5), \quad n = 0, 1, \ldots, 4$$

- For each part is calculated own periodogram

# Bartlett method

- Estimate the power spectrum averaged over the subsequence periodogram:

$$\hat{P}_x(\exp(j\omega)) = \frac{1}{K}\sum_{i=1}^{K} \hat{P}_x^{(i)}(\exp(j\omega))$$

- This is called the Bartlett method

# Bartlett method

- Bartlett periodogram, the expectation is $\frac{1}{2\pi} P_x * W_B$ where $W_B$ is Bartlett-window for *L* observation
- Assuming that the subsequences obtained did not correlate

$$\begin{aligned}\mathrm{var}\{\hat{P}_x(\exp(j\omega))\} &= \frac{1}{K}\mathrm{var}\{\hat{P}_x^{(i)}(\exp(j\omega))\} \\ &\approx \frac{1}{K} P_x^2(\exp(j\omega))\end{aligned}$$

- Approximate value, since *x*(*n*) is not white noise and subsequences are not uncorrelated
- Demo: bwnoise.R, bwsin.R

# Welch method

- Bartlett's method is a compromise between resolution (subsequence length) and variance (subsequence number)
- Welch method is added to the method of Bartlett
- two heuristics:

  - Subsequence may overlap

  - Subsequence windowing by window function

# Welch method

- $N$ observations are divided in $L$-length partially overlapping subsequences $x_i(0)$, $x_i(1)$, ..., $x_i(L-1)$, $i = 1, . . . , K$
- $K>N/L$ due to the overlap
- Each subsequence is windowing function $w(n)$, $n = 0, 1, ..., L-1$
- From subsequence $x_i(n)w(n)$, periodogram is calculated

# Welch method

- This periodogram is

$$\hat{P}_x^{(i)}(\exp(j\omega)) = \frac{1}{LU}\left|\sum_{k=0}^{L-1} w(k)x_i(k)\exp(-jk\omega)\right|^2$$

- In order for the expectation of the periodogram to be properly normalized, we need the term

$$U = \frac{1}{L}\sum_{n=0}^{L-1} |\, w(n)\,|^2 = \frac{1}{2\pi L}\int |\, \mathrm{W}(\exp(j\omega))\,|^2\, d\omega$$

- Welch periodogram estimate is the average

$$\hat{P}_{\mathrm{W}} = \frac{1}{K}\sum_{i=1}^{K} \hat{P}_x^{(i)}$$

# Welch method

- Its expectation is a convolution

$$\mathrm{E}(\hat{P}_W(\exp(j\omega))) = \frac{1}{2\pi LU} P_x(\exp(j\omega)) * |W(\exp(j\omega))|^2$$

- Welch method is
  - Asymptotically unbiased when *L* is growing
  - the variance is Proportional to *1/K* (*K>N/L*, so in this respect, the variance decreases more than the Bartlett method. Subsequence overlap on the other hand, increases the variance)

# Blackman-Tukey method

- Periodogram estimation of all the possible autocorrelation
- When $k$ is large, so $r_x(k)$ are estimated on a small number of terms $x(n+k)x^*(n)$
- The method reduce the impact of the estimates on the power spectra estimation
- This can be done by autocorrelation windowing

# Blackman-Tukey method

- In the Blackman-Tukey method, we selected a conjugate symmetric window function *w*(*k*) for the autocorrelation:

$$\hat{P}_{BT}(\exp(j\omega)) = \sum_{k=-M}^{M} \hat{r}_x(k) w(k) \exp(-j\omega k)$$

- In BT the Fourier-transform of the product $\hat{r}_x(k)w(k)$ so convolution theorem gives

$$\hat{P}_{BT}(\exp(j\omega)) = \frac{1}{2\pi} \hat{P}_x(\exp(j\omega)) * W(\exp(j\omega))$$

# Blackman-Tukey method

- $P_{BT}$ is a windowed by $W$ periodogram, and not the correct power spectrum
- Periodogram "smoothes" in the convolution, because the function $W(\exp(jw))$ is not an impulse
- For the window function $w(k)$, in addition to conjugate symmetricity, we require

$$W(\exp(j\omega)) \geq 0 \quad \text{for all} \quad \omega$$

# Blackman-Tukey method

- If *N>M* then we can roughly estimate the expected value of the Blackman-Tukeyn method

$$\mathrm{E}(\hat{P}_{BT}) \approx \frac{1}{2\pi} P_x(\exp(j\omega)) * W(\exp(j\omega))$$

- If *N* is large and *M* is also large enough so the variance is approximately

$$\mathrm{var}(\hat{P}_{BT}) \approx P_x^2 \frac{1}{N} \sum_{k=-M}^{M} w^2(k)$$

# Nonparametric methods Comparison

- Let's examines the different methods in respect of two criteria:

$$\mathcal{V} = \frac{\mathrm{var}\{\hat{P}_x(\exp(j\omega))\}}{E^2\{\hat{P}_x(\exp(j\omega))\}}$$

$$\Delta\omega = \text{ Resolution}$$

- $\mathcal{V}$ does not change when $P_x$ multiplied by an arbitrary constant
- Defining $\mathcal{M} = \mathcal{V}\Delta\omega$, which is a compromise between resolution and variance

| | Normalized Variance $\mathcal{V}$ | Resolution $\Delta\omega$ | $\mathcal{M}$ |
|---|---|---|---|
| Periodogram $\hat{P}_x$ | $1$ | $1/N$ | $1/N$ |
| Bartlett $\hat{P}_B$ | $1/K$ | $K/N$ | $1/N$ |
| Welch $\hat{P}_W$ | $1/K$ | $1/L$ | $1/N$ |
| Blackman-Tukey $\hat{P}_{BT}$ | $M/N$ | $1/M$ | $1/N$ |