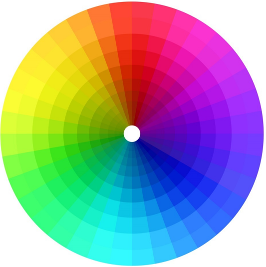
UPPSALA UNIVERSITET

Line spectra, Covariance structure, and NLS



Per Mattsson

Systems and Control
Department of Information Technology
Uppsala University

2019-09-17

Summary from last lecture

- Parametric methods $\Longrightarrow$ fewer parameters to estimate (bias/variance)
- AR and ARMA models for rational (continuous) spectra
- AR – only peaks
- ARMA – peaks and valleys
- 4 approaches:
  - YW (AR)
  - LSAR (AR)
  - MYW (ARMA, two-stage)
  - LSARMA (ARMA, two-stage)
- ARMA is nonlinear, methods have user parameters

Today: Estimating line spectra.

per.mattsson@it.uu.se

UPPSALA UNIVERSITET

Sinusoidal model

Sinusoidal model

- Many applications (e.g. communications, radar/sonar etc) have signals with (near) sinusoidal components.

Sinusoidal model

- Many applications (e.g. communications, radar/sonar etc) have signals with (near) sinusoidal components.
- Stable ARMA-models, $y(t) = \frac{B(z)}{A(z)}e(t)$, are poor approximations.

Sinusoidal model

- Many applications (e.g. communications, radar/sonar etc) have signals with (near) sinusoidal components.
- Stable ARMA-models, $y(t) = \frac{B(z)}{A(z)} e(t)$, are poor approximations.
- Complex sinusoid:

$$x_k(t) = \alpha_k e^{i(\omega_k t + \varphi_k)}$$

per.mattsson@it.uu.se

Sinusoidal model

- Many applications (e.g. communications, radar/sonar etc) have signals with (near) sinusoidal components.
- Stable ARMA-models, $y(t) = \frac{B(z)}{A(z)}e(t)$, are poor approximations.
- Complex sinusoid:

$$x_k(t) = \alpha_k e^{i(\omega_k t + \varphi_k)}$$

- Frequency: ω_k. Amplitude: α_k. Phase: φ_k.

Sinusoidal model

- Many applications (e.g. communications, radar/sonar etc) have signals with (near) sinusoidal components.
- Stable ARMA-models, $y(t) = \frac{B(z)}{A(z)}e(t)$, are poor approximations.
- Complex sinusoid:

$$x_k(t) = \alpha_k e^{i(\omega_k t + \varphi_k)}$$

- Frequency: ω_k. Amplitude: α_k. Phase: φ_k.
- Model with n components and noise,

$$x(t) = \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} = \sum_{k=1}^{n} x_k(t),$$
$$y(t) = x(t) + e(t).$$

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

per.mattsson@it.uu.se

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.
  - Mathematically convenient.

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.
  - Mathematically convenient.
  - Nuisance parameters (we are not interested in their values).

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.
  - Mathematically convenient.
  - Nuisance parameters (we are not interested in their values).
  - Realistic (no reason to assume any particular values).

per.mattsson@it.uu.se

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.
  - Mathematically convenient.
  - Nuisance parameters (we are not interested in their values).
  - Realistic (no reason to assume any particular values).
- A3: $e(t)$ is circular white noise with variance σ^2.

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.
  - Mathematically convenient.
  - Nuisance parameters (we are not interested in their values).
  - Realistic (no reason to assume any particular values).
- A3: $e(t)$ is circular white noise with variance σ^2. That is,

$$E\{e(t)e^*(s)\} = \sigma^2 \delta_{t,s}, \quad E\{e(t)e(s)\} = 0.$$

UPPSALA UNIVERSITET

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.
  - Mathematically convenient.
  - Nuisance parameters (we are not interested in their values).
  - Realistic (no reason to assume any particular values).
- A3: $e(t)$ is circular white noise with variance σ^2. That is,

$$E\{e(t)e^*(s)\} = \sigma^2 \delta_{t,s}, \quad E\{e(t)e(s)\} = 0.$$

  - May be achieved by slow sampling or filtering.

Sinusoidal model, assumptions

$$y(t) = \underbrace{\sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k))}}_{x_k(t)}}_{x(t)} + e(t)$$

- A1: $\alpha_k > 0$ and $\omega_k \in [-\pi, \pi]$. Prevents model ambiguities.
- A2: φ_k are independent random variables uniformly distributed on $[-\pi, \pi]$.
  - Mathematically convenient.
  - Nuisance parameters (we are not interested in their values).
  - Realistic (no reason to assume any particular values).
- A3: $e(t)$ is circular white noise with variance σ^2. That is,

$$E\{e(t)e^*(s)\} = \sigma^2 \delta_{t,s}, \quad E\{e(t)e(s)\} = 0.$$

  - May be achieved by slow sampling or filtering.
  - Some methods are poor for colored noise!

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}$.
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = \text{[Board]} =$$

per.mattsson@it.uu.se

UPPSALA UNIVERSITET

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}$.
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = \text{[Board]} = E\{x(t)x^*(t-k)\} + \sigma^2 \delta_{k,0}.$$

per.mattsson@it.uu.se

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}$.
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = \text{[Board]} = E\{x(t)x^*(t-k)\} + \sigma^2 \delta_{k,0}.$$

- Note that

$$E\{x_p(t)x_j^*(t-k)\} =$$

UPPSALA UNIVERSITET

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}$.
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = \text{[Board]} = E\{x(t)x^*(t-k)\} + \sigma^2\delta_{k,0}.$$

- Note that

$$\begin{aligned} E\{x_p(t)x_j^*(t-k)\} &= \alpha_p\alpha_j e^{i(\omega_p-\omega_j)t} e^{i\omega_j k} E\{e^{i\varphi_p} e^{-i\varphi_j}\} \\ &= \text{[Board]} = \end{aligned}$$

UPPSALA UNIVERSITET

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}$.
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = [\text{Board}] = E\{x(t)x^*(t-k)\} + \sigma^2\delta_{k,0}.$$

- Note that

$$\begin{aligned} E\{x_p(t)x_j^*(t-k)\} &= \alpha_p\alpha_j e^{i(\omega_p-\omega_j)t}e^{i\omega_j k}E\{e^{i\varphi_p}e^{-i\varphi_j}\} \\ &= [\text{Board}] = \alpha_p^2 e^{i\omega_p k}\delta_{p,j}. \end{aligned}$$

UPPSALA UNIVERSITET

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}.$
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = \text{[Board]} = E\{x(t)x^*(t-k)\} + \sigma^2 \delta_{k,0}.$$

- Note that

$$\begin{aligned} E\{x_p(t)x_j^*(t-k)\} &= \alpha_p \alpha_j e^{i(\omega_p - \omega_j)t} e^{i\omega_j k} E\{e^{i\varphi_p} e^{-i\varphi_j}\} \\ &= \text{[Board]} = \alpha_p^2 e^{i\omega_p k} \delta_{p,j}. \end{aligned}$$

- Hence

$$r(k) = \text{[Board]} =$$

per.mattsson@it.uu.se

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}$.
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = \text{[Board]} = E\{x(t)x^*(t-k)\} + \sigma^2 \delta_{k,0}.$$

- Note that

$$\begin{aligned} E\{x_p(t)x_j^*(t-k)\} &= \alpha_p \alpha_j e^{i(\omega_p - \omega_j)t} e^{i\omega_j k} E\{e^{i\varphi_p} e^{-i\varphi_j}\} \\ &= \text{[Board]} = \alpha_p^2 e^{i\omega_p k} \delta_{p,j}. \end{aligned}$$

- Hence

$$r(k) = \text{[Board]} = \sum_{p=1}^{n} \alpha_p^2 e^{i\omega_p k} + \sigma^2 \delta_{k,0}.$$

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}.$
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = \text{[Board]} = E\{x(t)x^*(t-k)\} + \sigma^2 \delta_{k,0}.$$

- Note that

$$\begin{aligned} E\{x_p(t)x_j^*(t-k)\} &= \alpha_p \alpha_j e^{i(\omega_p - \omega_j)t} e^{i\omega_j k} E\{e^{i\varphi_p} e^{-i\varphi_j}\} \\ &= \text{[Board]} = \alpha_p^2 e^{i\omega_p k} \delta_{p,j}. \end{aligned}$$

- Hence

$$r(k) = \text{[Board]} = \sum_{p=1}^{n} \alpha_p^2 e^{i\omega_p k} + \sigma^2 \delta_{k,0}.$$

- so,

$$\phi(\omega) = \text{[Board]} =$$

The spectrum

- PSD: $\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}, \quad r(k) = E\{y(t)y^*(t-k)\}$.
- With the sinusoidal model

$$E\{y(t)y^*(t-k)\} = [\text{Board}] = E\{x(t)x^*(t-k)\} + \sigma^2 \delta_{k,0}.$$

- Note that

$$\begin{aligned} E\{x_p(t)x_j^*(t-k)\} &= \alpha_p \alpha_j e^{i(\omega_p-\omega_j)t} e^{i\omega_j k} E\{e^{i\varphi_p} e^{-i\varphi_j}\} \\ &= [\text{Board}] = \alpha_p^2 e^{i\omega_p k} \delta_{p,j}. \end{aligned}$$

- Hence

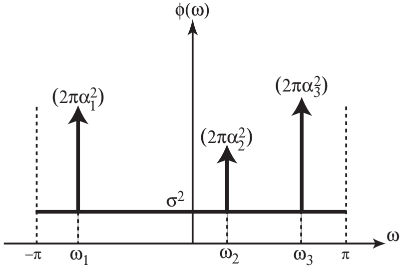
$$r(k) = [\text{Board}] = \sum_{p=1}^{n} \alpha_p^2 e^{i\omega_p k} + \sigma^2 \delta_{k,0}.$$

- so,

$$\phi(\omega) = [\text{Board}] = 2\pi \sum_{p=1}^{n} \alpha_p^2 \delta(\omega - \omega_p) + \sigma^2.$$

The theoretical spectrum

$$\boxed{\phi(\omega) = 2\pi \sum_{p=1}^{n} \alpha_p^2 \delta(\omega - \omega_p) + \sigma^2}$$



Need to estimate $\{\omega_k\}$, $\{\alpha_k\}$ and σ^2.

Parameter estimation

Parameter estimation

- If ω_k is given, then α_k (and φ_k) can be estimated from

$$y(t) = \sum_{k=1}^{n} \beta_k e^{i\omega_k t} + e(t),$$

with $\beta_k = \alpha_k e^{i\varphi_k} \Longrightarrow \alpha_k = |\beta_k|, \; \varphi_k = \arg(\beta_k).$

Parameter estimation

- If ω_k is given, then α_k (and φ_k) can be estimated from

$$y(t) = \sum_{k=1}^{n} \beta_k e^{i\omega_k t} + e(t),$$

with $\beta_k = \alpha_k e^{i\varphi_k} \implies \alpha_k = |\beta_k|,\ \varphi_k = \arg(\beta_k)$.

- Using least squares, we get

$$\min_{\beta} \|y - B\beta\|_2^2,$$

$$y = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}, \quad B = \begin{bmatrix} e^{i\omega_1} & \cdots & e^{i\omega_n} \\ \vdots & & \vdots \\ e^{iN\omega_1} & \cdots & e^{iN\omega_n} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

Parameter estimation

- If ω_k is given, then α_k (and φ_k) can be estimated from

$$y(t) = \sum_{k=1}^{n} \beta_k e^{i\omega_k t} + e(t),$$

with $\beta_k = \alpha_k e^{i\varphi_k} \implies \alpha_k = |\beta_k|$, $\varphi_k = \arg(\beta_k)$.

- Using least squares, we get

$$\min_{\beta} \|y - B\beta\|_2^2,$$

$$y = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}, \quad B = \begin{bmatrix} e^{i\omega_1} & \cdots & e^{i\omega_n} \\ \vdots & & \vdots \\ e^{iN\omega_1} & \cdots & e^{iN\omega_n} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

- However, in general ω_k are not given!

Nonlinear least squares

- What if we optimize over ω_k too? That is, minimize

$$f(\omega, \alpha, \varphi) = \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2 = \|y - B\beta\|_2^2.$$

UPPSALA UNIVERSITET

Nonlinear least squares

- What if we optimize over ω_k too? That is, minimize

$$f(\omega, \alpha, \varphi) = \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2 = \|y - B\beta\|_2^2.$$

- $(+)$ Maximum likelihood estimator if $e(t)$ is Gaussian.

per.mattsson@it.uu.se

Nonlinear least squares

- What if we optimize over ω_k too? That is, minimize

$$f(\omega, \alpha, \varphi) = \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2 = \|y - B\beta\|_2^2.$$

- (+) Maximum likelihood estimator if $e(t)$ is Gaussian.
- (-) Nonlinear in ω_k, hard to solve (local minima).

Nonlinear least squares

- What if we optimize over ω_k too? That is, minimize

$$f(\omega, \alpha, \varphi) = \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2 = \|y - B\beta\|_2^2.$$

- (+) Maximum likelihood estimator if $e(t)$ is Gaussian.
- (-) Nonlinear in ω_k, hard to solve (local minima).

Concentrate out β:

- The β that minimizes f is given by

$$\hat{\beta} =$$

UPPSALA UNIVERSITET

Nonlinear least squares

- What if we optimize over ω_k too? That is, minimize

$$f(\omega, \alpha, \varphi) = \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2 = \|y - B\beta\|_2^2.$$

- (+) Maximum likelihood estimator if $e(t)$ is Gaussian.
- (-) Nonlinear in ω_k, hard to solve (local minima).

Concentrate out β:

- The β that minimizes f is given by

$$\hat{\beta} = (B^* B)^{-1} B^* y.$$

per.mattsson@it.uu.se

UPPSALA UNIVERSITET

Nonlinear least squares

- What if we optimize over ω_k too? That is, minimize

$$f(\omega, \alpha, \varphi) = \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2 = \|y - B\beta\|_2^2.$$

- (+) Maximum likelihood estimator if $e(t)$ is Gaussian.
- (-) Nonlinear in ω_k, hard to solve (local minima).

Concentrate out β:

- The β that minimizes f is given by

$$\hat{\beta} = (B^* B)^{-1} B^* y.$$

- Inserting this back into f, we can see that $\hat{\omega}$ can be found by

$$\min_{\omega} \|y - B\hat{\beta}\|_2^2 =$$

per.mattsson@it.uu.se

Nonlinear least squares

- What if we optimize over ω_k too? That is, minimize

$$f(\omega, \alpha, \varphi) = \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2 = \|y - B\beta\|_2^2.$$

- (+) Maximum likelihood estimator if $e(t)$ is Gaussian.
- (-) Nonlinear in ω_k, hard to solve (local minima).

Concentrate out β:

- The β that minimizes f is given by

$$\hat{\beta} = (B^* B)^{-1} B^* y.$$

- Inserting this back into f, we can see that $\hat{\omega}$ can be found by

$$\min_{\omega} \|y - B\hat{\beta}\|_2^2 = \min_{\omega} \|y - B(B^* B)^{-1} B^* y\|_2^2$$

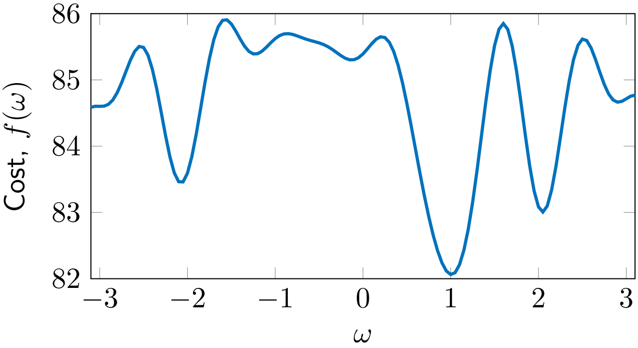
Nonlinear least squares, properties

$$\boxed{\begin{aligned}\hat{\omega} &= \underset{\omega}{\operatorname{argmin}} \|y - B(B^*B)^{-1}B^*y\|_2^2 \\ \hat{\beta} &= (B^*B)^{-1}B^*y\Big|_{\omega=\hat{\omega}}\end{aligned}}$$

- Gives a consistent estimate.
- For $N \gg 1$, $\operatorname{var}(\hat{\omega}_k) = \frac{6\sigma^2}{N^3\alpha_k^2}$.
- Best (minimum variance) unbiased estimator of ω_k.
- Problem: No (general) good method available to carry out the minimization.

Nonlinear least squares, example

The cost function when $y(t) = e^{it} + e(t)$, and $N = 100$:



No general way to find the global minimum!

Periodogram vs NLS

- For $n = 1$ (one complex sinusoid) it can be shown that the NLS estimate is given by

$$\hat{\omega}_1 =$$

Periodogram vs NLS

- For $n = 1$ (one complex sinusoid) it can be shown that the NLS estimate is given by

$$\hat{\omega}_1 = \underset{\omega}{\operatorname{argmax}}\ \hat{\phi}_p(\omega).$$

Periodogram vs NLS

- For $n = 1$ (one complex sinusoid) it can be shown that the NLS estimate is given by

$$\hat{\omega}_1 = \underset{\omega}{\operatorname{argmax}}\ \hat{\phi}_p(\omega).$$

- For $n > 1$ and

$$|\omega_p - \omega_k| > 2\pi/N, \quad \text{(resolution limit)}$$

the NLS estimates is approximately the location of the n largest peaks in $\hat{\phi}_p(\omega)$.

Periodogram vs NLS

- For $n = 1$ (one complex sinusoid) it can be shown that the NLS estimate is given by

$$\hat{\omega}_1 = \underset{\omega}{\operatorname{argmax}}\ \hat{\phi}_p(\omega).$$

- For $n > 1$ and

$$|\omega_p - \omega_k| > 2\pi/N, \quad \text{(resolution limit)}$$

the NLS estimates is approximately the location of the n largest peaks in $\hat{\phi}_p(\omega)$.

- For tightly spaced peaks, super resolution are needed!

Models of sinusoidal signals in noise

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

- Annihilating filter: $(1 - e^{i\omega_k} z^{-1}) x_k(t) = 0.$

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

- Annihilating filter: $(1 - e^{i\omega_k} z^{-1}) x_k(t) = 0.$
- Let

$$A(z) = \prod_{k=1}^{n} (1 - e^{i\omega_k} z^{-1}),$$

  then

$$A(z)y(t) =$$

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

- Annihilating filter: $(1 - e^{i\omega_k} z^{-1}) x_k(t) = 0.$
- Let

$$A(z) = \prod_{k=1}^{n} (1 - e^{i\omega_k} z^{-1}),$$

then

$$A(z)y(t) = A(z)e(t)$$

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

- Annihilating filter: $(1 - e^{i\omega_k} z^{-1}) x_k(t) = 0$.
- Let

  $$A(z) = \prod_{k=1}^{n} (1 - e^{i\omega_k} z^{-1}),$$

  then

  $$A(z)y(t) = A(z)e(t)$$

- Can estimate ω_k by estimating $A(z)$

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

- Annihilating filter: $(1 - e^{i\omega_k} z^{-1}) x_k(t) = 0$.
- Let
$$A(z) = \prod_{k=1}^{n} (1 - e^{i\omega_k} z^{-1}),$$
then
$$A(z)y(t) = A(z)e(t)$$
- Can estimate ω_k by estimating $A(z)$
- Note: Can't divide away $A(z)$ due to zeros on the unit circle!

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

- Annihilating filter: $(1 - e^{i\omega_k} z^{-1}) x_k(t) = 0.$
- Let

$$A(z) = \prod_{k=1}^{n} (1 - e^{i\omega_k} z^{-1}),$$

  then

$$A(z) y(t) = A(z) e(t)$$

- Can estimate ω_k by estimating $A(z)$
- Note: Can't divide away $A(z)$ due to zeros on the unit circle!
- Use Modified Yule-Walker?

ARMA-model

$$y(t) = \sum_{k=1}^{n} \underbrace{\alpha_k e^{i(\omega_k t + \varphi_k)}}_{x_k(t)} + e(t).$$

- Annihilating filter: $(1 - e^{i\omega_k} z^{-1}) x_k(t) = 0.$
- Let
$$A(z) = \prod_{k=1}^{n} (1 - e^{i\omega_k} z^{-1}),$$
then
$$A(z)y(t) = A(z)e(t)$$
- Can estimate ω_k by estimating $A(z)$
- Note: Can't divide away $A(z)$ due to zeros on the unit circle!
- Use Modified Yule-Walker?
- Can be improved by instead estimating $B(z) = A(z)\tilde{A}(z)$ in
$$B(z)y(t) = B(z)e(t).$$

Covariance model (1)

Notation:

$$a(\omega) \triangleq [1,\ e^{-i\omega}, \ldots, e^{-i(m-1)\omega}]^{\mathsf{T}} \qquad (m \times 1)$$

$$A_m = [a(\omega_1), \ldots, a(\omega_n)] \qquad (m \times n)$$

For some $m \geq n$

$$\boxed{\operatorname{rank}(A_m) = n \quad \text{if} \quad \omega_k \neq \omega_p \text{ for } k \neq p}$$

Covariance model (1)

Notation:

$$a(\omega) \triangleq [1,\ e^{-i\omega}, \dots, e^{-i(m-1)\omega}]^\mathsf{T} \qquad (m \times 1)$$

$$A_m = [a(\omega_1), \dots, a(\omega_n)] \qquad (m \times n)$$

For some $m \geq n$

$$\boxed{\operatorname{rank}(A_m) = n \quad \text{if} \quad \omega_k \neq \omega_p \text{ for } k \neq p}$$

We can write

$$\tilde{y}_m(t) \triangleq \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-m+1) \end{bmatrix} = A_m \tilde{x}_m(t) + \tilde{e}_m(t)$$

where

$$\tilde{x}_m(t) = [x_1(t), \dots, x_n(t)]^\mathsf{T}, \quad \tilde{e}_m(t) = [e(t), \dots, e(t-m+1)]^\mathsf{T}.$$

Covariance model (2)

This gives the covariance matrix model

$$\boxed{R \triangleq E\{\tilde{y}_m(t)\tilde{y}_m^*(t)\} = A_m P A_m^* + \sigma^2 I, \qquad P = \begin{bmatrix} \alpha_1^2 & & 0 \\ & \ddots & \\ 0 & & \alpha_n^2 \end{bmatrix}}$$

The eigenstructure of R contains complete frequency information.

per.mattsson@it.uu.se

Covariance model (2)

This gives the covariance matrix model

$$\boxed{R \triangleq E\{\tilde{y}_m(t)\tilde{y}_m^*(t)\} = A_m P A_m^* + \sigma^2 I, \qquad P = \begin{bmatrix} \alpha_1^2 & & 0 \\ & \ddots & \\ 0 & & \alpha_n^2 \end{bmatrix}}$$

The eigenstructure of R contains complete frequency information.

Furthermore, using non-overlapping parts of $\{y(t)\}$ (uncorrelated)

$$\boxed{\Gamma \triangleq E\{\tilde{y}_M(t-L-1)\tilde{y}_{L+1}^*(t)\} = A_M P_{L+1} A_{L+1}^*}$$

where

$$P_K = \begin{bmatrix} \alpha_1^2 e^{-i\omega_1 K} & & 0 \\ & \ddots & \\ 0 & & \alpha_n^2 e^{-i\omega_n K} \end{bmatrix}$$

Null space of Γ ($L, M \geq n$) contains complete frequency info.

UPPSALA UNIVERSITET

Singular value decomposition (SVD)

Definition

Factor a matrix $A \in \mathbb{C}^{m \times n}$ as

$$A = U \Sigma V^*$$

where $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary matrices, and $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal with non-negative diagonal values σ_i (singular values) ordered in decreasing order

(See A.4 in the book for more details)

- Low-rank approximations
- Useful to find the best fit subspaces of A
- Projection onto subspace – perpendicular distance
- LS fit – vertical distance

SVD (2)

$A \in \mathbb{C}^{m\times n}$ is of rank $r < \min(m,n)$ then

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} = U_1 \Sigma_1 V_1^*$$

where $\Sigma_1 \in \mathbb{R}^{r\times r}$, then

- U_1 is an orthonormal basis in $\mathcal{R}(A)$
- U_2 is an orthonormal basis in $\mathcal{N}(A^*)$
- V_1 is an orthonormal basis in $\mathcal{R}(A^*)$
- V_2 is an orthonormal basis in $\mathcal{N}(A)$

Moreover, $\mathcal{R}(A)$ and $\mathcal{N}(A^*)$ are orthogonal to each other and they span $\mathbb{C}^m$
($\mathcal{N}(A^*)$ is the orthogonal complement to $\mathcal{R}(A)$ and vice versa)

Summary

- Signal model for line spectra
- Estimating the spectrum $\implies$ frequency estimation
- Models of the covariance matrix and ARMA-type model
- Nonlinear least squares

Summary

- Signal model for line spectra
- Estimating the spectrum $\Longrightarrow$ frequency estimation
- Models of the covariance matrix and ARMA-type model
- Nonlinear least squares

Next lecture:

- Methods for estimating line spectra (super-resolution): HOYW, MUSIC, Min-norm, ESPRIT
- All exploit eigen-decomposition of the covariance matrix and corresponding subspaces

Interesting complements: 4.9.5 Minimization methods, Appendix B Cramér-Rao bound