Fiche - TSIA202b

Partie 1 : Non-parametric spectral estimation :

1. Reminder : WSS processes

déf : (WSS process) is a complex sequence of R.V. $X_t \in \mathbb{C}, \forall t \in \mathbb{Z}$ s.t
$$\begin{cases} \cdot \ \mathbb{E}|X_t|^2 < \infty \\ \cdot \ \mathbb{E}[X_t] = \mu_x \Rightarrow \perp\!\!\!\perp t \\ \cdot \ \forall k \in \mathbb{Z},\ cov[X_{t+k}, X_t] = \mathbb{E}[(X_{t+k}-\mu_x)\overline{(X_t-\mu_x)}] = \mathbb{E}[X_t^c \bar{X}_t^c] \Rightarrow \perp\!\!\!\perp t \end{cases}$$
$X_t^c = X_t - \mu_x$

PROP : (Strict vs Wide sense stationnary)
- si X_t est stationnaire au sens strict tq $\mathbb{E}|X_t|^2 < \infty$, alors : X_t est stationnaire au sens large
- si X_t est un proc. Gaussien, alors : il est stationnaire strict ssi il est stationnaire large

déf : (Autocov func.) Let X_t complex WSS proc., ACF : $\forall k \in \mathbb{Z},\ R_{xx}(k) = cov(X_{t+k}, X_t) = \mathbb{E}[X_{t+k}^c \bar{X}_t^c]$

↳ prop :
- $Var(X) = R_{xx}(0) \geq 0$
- $\forall k \in \mathbb{Z},\ R_{xx}(-k) = \overline{R_{xx}(k)}$ (Hermitian symmetry)
- $\forall k \in \mathbb{Z}, \forall t_1, \dots, t_k, \forall \lambda_1, \dots, \lambda_k \in \mathbb{C},\ \sum_{i=1}^{k}\sum_{j=1}^{k} \lambda_i \bar{\lambda}_j R_{xx}(t_i - t_j) \geq 0$ (Positive semi-definiteness) (car $\mathbb{E}|\sum_i \lambda_i X_{t_i}|^2 \geq 0$)
- $\forall k \in \mathbb{Z},\ |R_{xx}(k)| \leq R_{xx}(0)$ (boundedness) (par C-S)

↳ Remark : power of a WSS proc. : $P_x = \mathbb{E}|X_t|^2 = R_{xx}(0) + |\mu_x|^2$

déf : (PSD) Let X_t be a complex WSS proc., if $R_{xx}(k) \in \ell^1(\mathbb{Z}) \rightarrow$ PSD of X_t : $\forall \nu \in \mathbb{R},\ S_{xx}(\nu) = \sum_{k=-\infty}^{+\infty} R_{xx}(k) e^{-2i\pi\nu k}$ ($= TFtD[R_{xx}]$)

↳ prop :
- (Inversion) $\cdot\ R_{xx}(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{xx}(\nu) e^{+2i\pi\nu k} d\nu$ ($= TFtD^{-1}(S_{xx})$)
- (C^0) $\cdot\ \nu \mapsto S_{xx}$ is continuous
- (≥ 0) $\cdot\ \forall \nu \in \mathbb{R},\ S_{xx}(\nu) \geq 0$ (par Herglotz thm)

↳ Remark : $P_x = R_{xx}(0) + |\mu_x|^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{xx}(\nu) d\nu + |\mu_x|^2$

$= \sum_{k=-\infty}^{+\infty} h_k e^{-2i\pi\nu k}$

theorem: (Filtering thm for WSS proc.) Let h_k the impulse response of a stable filter of frequency response $H(\nu)$, X_t a WSS proc and $Y_t = h * X_t$

then: Y is a WSS proc with
- mean: $\mu_Y = \mu_X H(0) = \mu_X \sum_{-\infty}^{+\infty} h_k$
- ACF: $R_{YY} = h * \tilde{h} * R_{XX}$ where $\tilde{h}_t = \overline{h}_{-t} \overset{TF}{\to} \overline{H(\nu)}$
- PSD: if $R_{XX} \in \ell^1$, $S_{YY}(\nu) = \underbrace{|H(\nu)|^2}_{\geq 0 \text{ et } C^0} \underbrace{S_{XX}(\nu)}_{\geq 0 \text{ et } C^0}$ ≥ 0 et C^0

Estimation of the mean and the ACF

- **Parametric estimation:** Let X r.v. parameterized by θ,
  - $\hat{\theta}$ estimator of θ is a fct° of X
  - Bias: $b(\theta,\hat{\theta}) = \mathbb{E}_\theta(\hat{\theta}(X) - \theta)$
  - MSE: $R(\theta,\hat{\theta}) = \mathbb{E}_\theta[|\hat{\theta}(X) - \theta|^2] = Var(\hat{\theta}(X)) + |b(\theta,\hat{\theta})|^2$

→ Asymptotic approach of estimation: Let $X = [X_1, \dots, X_N]^T$ r.v. (obs vector) whose distrib is parameterized by θ
- An estimator $\hat{\theta}_N$ is asymptotically unbiased iff $\lim_{N\to+\infty} b(\theta,\hat{\theta}_N) = 0$
- Mean square consistency: $\lim_{N\to+\infty} R(\theta,\hat{\theta}_N) = 0$

- **Estimation of the mean:** Let X_t be a WSS proc. of mean μ_X and ACF $R_{XX}(k)$,
  - Empirical mean: $\hat{\mu}_X = \frac{1}{N}\sum_{t=1}^{N} X_t$ → prop:
    - it's an unbiased estimator: $\mathbb{E}[\hat{\mu}_X] = \mu_X$
    - variance: $Var(\hat{\mu}_X) = \frac{1}{N}\sum_{k=-(N-1)}^{N-1} (1 - \frac{|k|}{N}) R_{XX}(k)$
    - If $R_{XX} \in \ell^1(\mathbb{Z})$, $var(\hat{\mu}_X) \sim \frac{S_{XX}(0)}{N} \to 0$: Mean square consistency

- **Estimation of the ACF:** Let X_t a centred WSS proc. of ACF $R_{XX}(k)$, (i.e. the mean is known)
  - Empirical ACF: $\hat{R}_{XX}(k) = \begin{cases} \frac{1}{N}\sum_{t=1}^{N-k} X_{t+k}\overline{X_t} & \text{if } 0 \leq k < N \\ 0 & \text{if } |k| \geq N \end{cases}$ and $\hat{R}_{XX}(-k) = \overline{\hat{R}_{XX}(k)}$

→ il est biaisé et on le garde ainsi pour garder la prop. de positivité

→ **prop:**
- $\hat{R}_{xx}$ is positive semi-definite
- it's an asymptotically unbiased estimator : $\mathbb{E}[\hat{R}_{xx}(k)] = \left(1 - \frac{|k|}{N}\right) R_{xx}(k)$
- if: in addition X_t is a strong linear proc. : $X_t = \sum_{k \in \mathbb{Z}} h_k Z_{t-k}$ where $h_k \in \ell^1(\mathbb{Z})$ and $Z_t \sim IID(0,\sigma^2)$ with $\mathbb{E} Z_t^4 < \infty$

  then : $\hat{R}_{xx}$ is mean square consistent : $\forall k \in \mathbb{Z}$, $\mathrm{var}(\hat{R}_{xx}(k)) = \mathcal{O}\left(\frac{1}{N}\right)$

  NTM LE Gradient

2. Non-parametric estimation of the PSD

→ Estimation of the PSD :

→ **Periodogram :** Let X_t be a centred WSS proc. s.t. $R_{xx} \in \ell^1(\mathbb{Z})$

estime la DSP

► Periodogram : $\hat{S}_{p,xx}(\nu) = \frac{1}{N} \left| \sum_{t=1}^{N} X_t e^{-2i\pi\nu t} \right|^2$

► Correlogram : $\hat{S}_{c,xx}(\nu) = \sum_{k=-(N-1)}^{N-1} \hat{R}_{xx}(k) e^{-2i\pi\nu k}$ where : $\hat{R}_{xx}(k) = \begin{cases} 0 & \forall |k| \geq N \\ \frac{1}{N-k} \sum_{t=1}^{N-k} X_{t+k} \bar{X}_t & \text{(unbiased)} \\ \frac{1}{N} \sum_{t=1}^{N-k} X_{t+k} \bar{X}_t & \text{(biased + psd)} \end{cases}$ $\forall k \in [0, N[$

RPZ graph corrélations entre séries de data

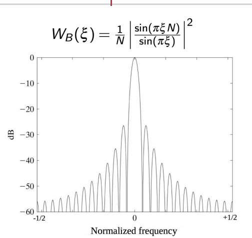
↳ positive semi-definite

↳ **prop:** $\hat{S}_{p,xx} = \hat{S}_{c,xx}$ if $\hat{R}_{xx}$ is the biased estimator

→ **Bias analysis of the periodogram :**

- Mean : $\mathbb{E}[\hat{S}_{p,xx}] = \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N}\right) R_{xx}(k) e^{-2i\pi\nu k}$

We define $w_B(k) = \begin{cases} 1 - \frac{|k|}{N} & \text{if } |k| < N \\ 0 & \text{if } |k| \geq N \end{cases}$ $\Rightarrow \mathbb{E}[\hat{S}_{p,xx}] = \sum_{k \in \mathbb{Z}} w_B(k) R_{xx}(k) e^{-2i\pi\nu k} = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{xx}(\nu - \varepsilon) W_B(\varepsilon)\, d\varepsilon$



Fejer kernel, $W_B(\nu)/W_B(0)$, for $N = 25$

Fejer Kernel : $W_B(\varepsilon) = \frac{1}{N} \left| \frac{\sin(\pi \varepsilon N)}{\sin(\pi \varepsilon)} \right|^2$ (car $w_B(k) = \frac{1}{N} \mathbb{1}_{[0,N-1]}(k) * \mathbb{1}_{[0,N-1]}(k)$

→ Conséquences :
- main lobe → smearing (= étalement) → width $= \frac{2}{N}$
- side lobes → leakage (= energie du main lobe fuit ds les lobes latéraux)
- loss of resolution ($\neq$ precision)

prop: if $R_{xx} \in \ell^1(\mathbb{Z})$, $\hat{S}_{p,xx}$ is asymptotically unbiased : $\lim\limits_{N\to+\infty} \mathbb{E}[\hat{S}_{p,xx}] = S_{xx}$ (par TCD)

→ Variance analysis of the periodogram

▷ **Definition:** complex (or circular) white noise is a WSS proc. s.t. :

$$\begin{cases} \mathbb{E}[Z_t \bar{Z}_s] = \sigma^2 \delta_{t,s} \\ \mathbb{E}[Z_t Z_s] = 0 \quad \forall t,s \end{cases} \iff \begin{cases} \mathbb{E}[Re(Z_t) Re(Z_s)] = \frac{1}{2}\sigma^2 \delta_{t,s} \\ \mathbb{E}[Im(Z_t) Im(Z_s)] = \frac{1}{2}\sigma^2 \delta_{t,s} \\ \mathbb{E}[Re(Z_t) Im(Z_s)] = 0 \quad \forall t,s \end{cases}$$

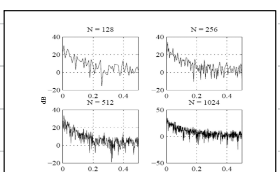
→ prop: (Variance analysis of the periodogram of a cplx white noise)

if Z_t is Gaussian white noise then : $\lim\limits_{N\to+\infty} \mathrm{cov}(\hat{S}_{p,zz}(\nu), \hat{S}_{p,zz}(\epsilon)) = \begin{cases} S_{zz}(\nu)^2 & \forall \nu = \epsilon \\ 0 & \forall \nu \neq \epsilon \end{cases}$ where $S_{zz} = \sigma^2$

→ prop: (Variance analysis of the periodogram of a stationary Gaussian proc.)

let Z_t cplx Gaussian white noise and let : $X_t = \sum\limits_{k\in\mathbb{Z}} h_k Z_{t-k}$ where $(h_k)_k \in \ell^1(\mathbb{Z})$ then $\forall \nu, \epsilon \in]-\frac{1}{2}, \frac{1}{2}]$, the prop holds.

↳ conséquence : $\hat{S}_{p,xx}$ is not even asymptotically mean square consistent (i.e. le risque $\not\to 0$)



Periodogram of a WSS process for $N = 128, 256, 512, 1024$

on veut lisser le periodogramme ⟺ fenêtrer R_{xx} ds le domaine temporel

⟺ reduire la variance

→ Blackman-Tukey method

Dans le but de réduire la variance, BT tronque l'ACF empirique

- let X_t centred WSS s.t. $R_{xx} \in \ell^1(\mathbb{Z})$. With $M < N$, the BT estimator of its PSD is defined as :

$$\boxed{\hat{S}_{BT,xx}(\nu) = \sum_{k=-(M+1)}^{M+1} \hat{R}_{xx}(k)\, e^{-2i\pi\nu k}}$$

ici on fenêtre avec une fenêtre rectangle

prop:
- if $M \to +\infty$, $\hat{S}_{BT,xx}$ is asymptotically unbiased
- if $M/N \to 0$, $Var(\hat{S}_{BT,xx}(\nu)) = O(\frac{M}{N}) \to 0$

↳ eg: $M = N^\alpha$ with $\alpha \in]0,1[$, $\hat{S}_{BT,xx}$ is mean square consistent

- BT offre un compromis entre la résolution spectrale $(O(\frac{1}{M}))$ et la variance $O(\frac{M}{N})$

lobe width : $2/M$

→ avec une fenêtre symmétrique de support $[-(M-1),(M-1)]$ avec $M < N$, $w(-k) = w(k)$ et $w(0) = 1$ ($\Leftrightarrow$ moyenne $\Sigma = 1$)

↳ pour pas modifier
| l'amplitude
| la puissance moyenne

Windowed periodogram : $\hat{S}_{BT,xx}(\nu) = \sum_{k=-(M-1)}^{M-1} w(k)\, \hat{R}_{xx}(k)\, e^{-2i\pi\nu k} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{S}_{P,xx}(\nu - \varepsilon)\, W(\varepsilon)\, d\varepsilon$

where $W(\varepsilon) = \sum_{k=-(M-1)}^{M-1} w_k\, e^{-2i\pi\varepsilon k} \in \mathbb{R}$ ⇒ BT performs a local weighted average

↳ if $w(k)$ positive | semi-definite | symétrique then $W(\varepsilon) \geq 0$, therefore $\hat{S}_{BT,xx}(\nu) \geq 0 \quad \forall \nu \in \mathbb{R}$

Le choix de :
- la largeur de la fenêtre est basé sur le compromis entre la résolution spectrale et la variance
- la forme ______________ entre l'étalement spectral et la fuite spectrale

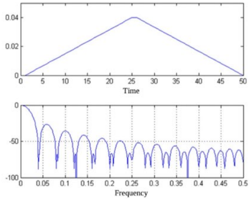
↳ réduit la résolution fréquentielle

↳ contribution des freq non-présentes dans le signal (lobes secondaires)

Bartlett window

$$w(k) = \left(1 - \frac{|k|}{M}\right) \mathbf{1}_{[-(M-1)\dots M-1]}(k)$$

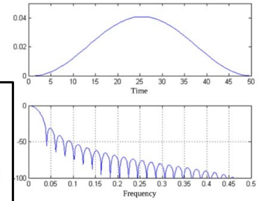
- Width : $2/M$, second lobe : -26 dB, decrease : -12 dB / octave



Hann window

$$w(k) = \left(0.5 + 0.5\cos\left(\frac{\pi k}{M}\right)\right) \mathbf{1}_{[-(M-1)\dots M-1]}(k)$$

- Width : $2/M$, second lobe : -31 dB, decrease : -18 dB / octave

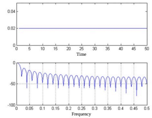


Rectangular window

$$w(k) = \mathbf{1}_{[-(M-1)\dots M-1]}(k)$$

- Width : $1/M$, second lobe : -13 dB, decrease : -6 dB / octave

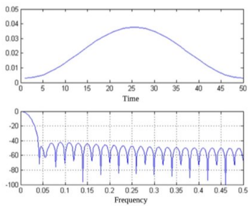
freq x2



Hamming window

$$w(k) = \left(0.54 + 0.46\cos\left(\frac{\pi k}{M-1}\right)\right) \mathbf{1}_{[-(M-1)\dots M-1]}(k)$$

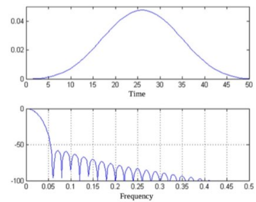
- Width : $2/M$, second lobe : -41 dB, decrease : -6 dB / octave



Blackman window

$$w(k) = \left(0.4266 + 0.4965\cos\left(\frac{\pi k}{M-1}\right) + 0.076\cos\left(\frac{2\pi k}{M-1}\right)\right) \mathbf{1}_{[-(M-1)\dots M-1]}(k)$$

- Width : $3/M$, second lobe : -57 dB, décroissance : -18 dB / octave



Bartlett method

- Idea : compared to the original periodogram, split up the data and average several periodograms in order to reduce the variance
- Segment N samples into L sub-samples of size $M = \frac{N}{L}$
- $\hat{S}_{B,XX}(\nu) = \frac{1}{L}\sum_{i=1}^{L} \frac{1}{M}\left|\sum_{t=1}^{M} \tilde{X}_{i,t} e^{-2i\pi\nu t}\right|^2$ where $\tilde{X}_{i,t} = X_{(i-1)M+t}$ for $t \in [1, M]$ and $i \in [1, L]$
- The spectral resolution is $O(\frac{1}{M})$ and the variance is $O(\frac{M}{N})$
- Same trade-off between spectral resolution and variance as the Blackman-Tukey estimate with a rectangular window

TELECOM Paris

Welch method = Bartlett + Blackman

- Refinement of the Bartlett method :
  - data segments overlap
  - each data segment is windowed
- $\tilde{X}_{i,t} = X_{(i-1)K+t}$ for $t \in [1, M]$ and $i \in [1, S]$
- If $K = M \Rightarrow$ Bartlett : $S = L = \frac{N}{M}$
- Recommended : $K = \frac{M}{2}$, $S \approx \frac{2N}{M}$
- $\hat{S}_{W,XX}(\nu) = \frac{1}{S}\sum_{i=1}^{S} \hat{S}^{(i)}_{P,XX}(\nu)$ and $\hat{S}^{(i)}_{P,XX}(\nu) = \frac{1}{MP}\left|\sum_{t=1}^{M} v(t)\tilde{X}_{i,t} e^{-2i\pi\nu t}\right|^2$ chaque trame est fenêtrée
  where $P = \frac{1}{N}\sum_{t=1}^{M} |v(t)|^2$ in order to normalize every periodogram
- Better control of smearing and leakage, variance similar to Bartlett

TELECOM Paris

Daniell method

- Idea : reduce the variance by smoothing the periodogram :

$$\hat{S}_{D,XX}(\nu) = \frac{1}{2J+1}\sum_{j=-J}^{J} \hat{S}_{P,XX}\left(\nu + \frac{j}{\tilde{N}}\right)$$

  where $\tilde{N} = N$ without *zero-padding*, or $\tilde{N} > N$ with *zero-padding*.
- The continuous version of the Daniell method is $\hat{S}_{D,XX}(\nu) = \frac{1}{\beta}\int_{\nu-\frac{\beta}{2}}^{\nu+\frac{\beta}{2}} \hat{S}_{P,XX}(\nu+\xi)\,d\xi$ with $\beta = \frac{2J}{\tilde{N}}$
- It can be seen as a particular case of the Blackman-Tukey method, with $W(\xi) = 1/\beta$ if $\xi \in [-\frac{\beta}{2}, \frac{\beta}{2}]$, or $W(\xi) = 0$ otherwise.

TELECOM Paris

Partie 2 : Parametric estimation of rationnal spectra :

1. Reminder : (linear process)

déf : (linear proc.) • $(X_t)_{t\in\mathbb{Z}}$ is a linear proc. iif there is $\mu_X \in \mathbb{C}$, $Z_t \sim BB(0,\sigma^2)$ and $(h_m)_m \in \ell^1(\mathbb{Z})$ s.t. :
$$X_t = \mu_X + \sum_{m=-\infty}^{+\infty} h_m Z_{t-m}, \quad \forall t \in \mathbb{Z}$$

- $(X_t)_{t\in\mathbb{Z}}$ is causal w.r.t to $(Z_t)_{t\in\mathbb{Z}}$ iif $h_m = 0 \quad \forall m < 0$
- $(X_t)_{t\in\mathbb{Z}}$ is invertible w.r.t. to $(Z_t)_{t\in\mathbb{Z}}$ iif there is $(g_m)_{m\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$ s.t. $Z_t = \sum_{m=0}^{+\infty} g_m (X_{t-m} - \mu_X) \ \forall t\in\mathbb{Z}$

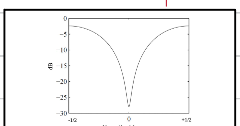
  ↳ causal

prop : (filtering thm for wss proc.) let X_t a linear proc., then X_t is WSS proc. with $\mathbb{E} = \mu_X$ and ACF : $R_{XX}(\ell) = \sigma^2 + \sum_{m=-\infty}^{+\infty} h_{m+\ell}\overline{h_m}$
and spectral density : $S_{XX}(\nu) = \sigma^2 |H(e^{-2i\pi\nu})|^2$ with $H(e^{-2i\pi\nu}) = \sum_{m\in\mathbb{Z}} h_m e^{-2i\pi\nu m}$

→ MA(q) proc. : déf : $(X_t)_{t\in\mathbb{Z}}$ is MA(q) iff $X_t = \sum_{m=0}^{q} b_m Z_{t-m}$ where • $Z_t \sim BB(0,\sigma^2)$
• $b_m \in \mathbb{C}$ and $b_0 = 1$ i.e. b_m is RIF

prop : (filtering thm for wss proc.) Let $(X_t)_{t\in\mathbb{Z}}$ a MA(q) proc., then : X_t is a wss proc. with $\mathbb{E} = 0$, $\underbrace{\text{ACF}}_{\text{à supp fini}}$: $R_{XX}(\ell) = \begin{cases} \sigma^2 \sum_{m=0}^{q-\ell} b_{m+\ell}\overline{b_m} & \text{if } 0 \le \ell \le q \\ 0 & \text{if } \ell > q \end{cases}$

and a spectra density $S_{XX}(\nu) = \sigma^2 \left| \sum_{m=0}^{q} b_m e^{-2i\pi\nu m} \right|^2$



PSD (in dB) of an MA(1) process with $\sigma = 1$ and $b_1 = -0.9$

Theorem : (Characterization of a MA(q) proc.) let $(X_t)_{t\in\mathbb{Z}}$ be a centred WSS of ACF R_{XX}, and let $q \geq 1$, then :

X_t is a MA proc. of minimal order $q \iff R_{XX}(q) \neq 0$ and $R_{XX}(\ell \geq q+1) = 0$ (i.e. à support fini)

↳ corrolary : the sum of $\underbrace{\text{2 decorrelated}}_{\gamma_{X+Y} = \gamma_X + \gamma_Y}$ MA(q) is a MA(q) proc.

→ **AR(p) proc.:** **déf:** $(X_t)_{t\in\mathbb{Z}}$ is AR(p) iff • it's WSS and solution of: $X_t = Z_t + \sum_{m=1}^{p} a_m X_{t-m}$ where $Z_t \sim BB(0,\sigma^2)$, $a_m \in \mathbb{C}$

cas p=1: $X_t = Z_t + a_1 X_{t-1}$ avec $|a_1|<1$ → l'implémentation stable est causale

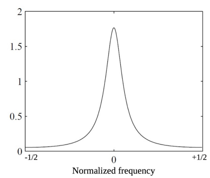
↳ $X_t = \sum_{m=0}^{+\infty} a_1^m Z_{t-m}$ (cv in L^2 and a.s.)

↳ if $|a_1|>1$, → anti-causal case: $|a_1|>1$, $X_t = -\frac{1}{a_1} Z_{t+1} + \frac{1}{a_1} X_{t+1} \Rightarrow X_t = -\sum_{m=0}^{+\infty} a_1^{-m} Z_{t+m}$

↳ **prop:** by the filtering theorem for WSS proc., if $|a_1|\neq 1$, X_t is WSS with: • $\mathbb{E}=0$

• $S_{XX}(\nu) = \sigma^2 \left|\sum_{m=0}^{+\infty} a_1^m e^{-2i\pi\nu m}\right| = \frac{\sigma^2}{|1-a_1 e^{-2i\pi\nu}|^2}$

AR(1), causal case



PSD of a Gaussian AR(1) process, with $\sigma = 1$ and $a_1 = 0.7$.

↳ pour $|a_1|<1$, si $k\geq 0$, $R_{XX}(k) = \sigma^2 \sum_{m=0}^{+\infty} a_1^{m+k} \bar{a}_1^{-m} = \sigma^2 \frac{a_1^k}{1-|a_1|^2}$

↳ pour $|a_1|>1$, si $k\geq 0$, $R_{XX}(k) = \sigma^2 \sum_{m=0}^{+\infty} a_1^m \bar{a}_1^{(m-k)} = \sigma^2 \frac{\bar{a}_1^{-k}}{|a_1|^2-1}$

→ **properties of an AR(p):**

- There is a WSS solution iff $A(z)\neq 0$ for $|z|=1$ → $\frac{1}{A(z)} = \sum_{-\infty}^{+\infty} h_m z^{-m}$ where $\sum_{m\in\mathbb{Z}} |h_m| < +\infty \Rightarrow X_t = \sum_{m\in\mathbb{Z}} h_m Z_{t-m}$
- if $A(z)=0 \Rightarrow |z|<1$ ⇒ causal solution
- if $A(z)=0 \Rightarrow |z|>1$ ⇒ anti-causal solution
- else X is a mixed AR proc. ⇒ ni causal ni anti-causal

2. Maximum entropy spectral estimation: let X_t be a centred WSS proc. s.t. $R_{XX} \in \ell^1(\mathbb{Z})$,

Non-parametric spectral estimation = estimate $S_{XX}(\nu)$ from N samples $X_1,\dots,X_N$

↳ Periodogram and Black-Tukey methods = for $M\leq N$, • first, compute estimates $\hat{R}_{XX}(k)$ of $R_{XX}(k)$ for $k\in [\![-(M-1),(M-1)]\!]$

• then, estimate $\hat{S}_{XX}(\nu)$ via a (windowed) DTFT of $\hat{R}_{XX}(k)$

↳ MA(M)

▷ New idea: with fixed $\hat{R}_{xx}(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{S}_{xx}(\nu) e^{+2i\pi\nu k} d\nu$ for $k \in [\![-(M-1),(M-1)]\!]$, compute the estimate $\hat{S}_{xx}(\nu)$ that that maximizes the entropy of the WSS probability distribution ⇒ ça devient un problème d'opti

on ne suppose plus que $\hat{R}_{xx}(k)$ est nul en dehors de $[\![-(M-1),(M-1)]\!]$

i.e. on maximise l'information

→ blind estimation: no information available about the WSS proc. beyond the knowledge of $\hat{R}_{xx}(k)$ for $k \in [\![-(M-1),(M-1)]\!]$

→ Entropy of a Gaussian Random vector: For a discrete R.V. with M values, entropy: $H := \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) = \frac{1}{\ln(2)} \sum_{k=1}^{M} p_m \ln(1/p_k)$

for N continuous variables $x_1, \dots, x_N$, $H_N = -\int p(x_1,\dots,x_N) \ln\left(p(x_1,\dots,x_N) c^{N/2}\right) dx_1 \dots dx_N = \mathbb{E}\left(\ln\left(p(x_1,\dots,x_N) c^{N/2}\right)\right)$

▷ if the variable are Gaussian, $p(x_1,\dots,x_N) = \frac{1}{(2\pi)^{N/2} \det(R_N)^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_x)^T R_N^{-1} (x-\mu_x)\right)$, if we choose c appropriately: $\boxed{H_N = \frac{1}{2} \ln(\det(R_N))}$

celle qui maximise l'entropie

$c = 2\pi e$

Proof: $H_N = -\mathbb{E}\left[\ln p(x_1,\dots,x_N) c^{N/2}\right] = -\mathbb{E}\left[-\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(\det R_N) - \frac{1}{2}(x-\mu_x)^T R_N^{-1}(x-\mu_x)\right]$

$= \frac{N}{2}\ln(2\pi) + \frac{1}{2}\ln(\det R_N) + \frac{1}{2}\,\mathbb{E}\left[\mathrm{trace}\left((x-\mu_x)^T R_N^{-1}(x-\mu_x)\right)\right]$

$= \mathrm{trace}\left(R_N^{-1}(x-\mu_x)(x-\mu_x)^T\right)$

$= \mathrm{trace}\left(R_N^{-1} \underbrace{\mathbb{E}\left((x-\mu_x)(x-\mu_x)^T\right)}_{R_N}\right)$

$= N$

$\Rightarrow H_N = \frac{1}{2}\ln(\det R_N) + \frac{N}{2}\ln(2\pi e) - \frac{N}{2}\ln(c)$

$c = 2\pi e$ $\Rightarrow = \frac{1}{2}\ln(\det R_N)$

$= \frac{1}{2}\sum_{m=0}^{N-1} \ln(\sigma_m^2)$

vp de R_N

$\Rightarrow H = \lim_{N\to+\infty} \frac{H_N}{N} = \frac{1}{2}\int_{-\frac{1}{2}}^{1/2} \ln(S_{xx}(\nu))\,d\nu$

- Problm: H_N diverges when $N \to +\infty$
- for a proc. of ∞-length, the entropy rate is: $H = \lim_{N\to+\infty} \frac{H_N}{N} = \lim_{N\to+\infty} \frac{1}{2}\ln\left((\det R_N)^{1/N}\right)$

→ **Maximum entropy method:** Among all the WSS proc. with fixed ACF $\hat{R}_{xx}(k)$ for $|k|<M$, the one which maximizes the entropy rate is the AR(M-1)

Proof: let $R(k)$ be the ACF of a WSS proc. and $S(\nu)$ it's PSD with: • $R(k) = \hat{R}_{xx}(k) \ \forall |k|<M$

We want to maximize the entropy rate of $R(k)$: $H = \int_{-\frac{1}{2}}^{1/2} \ln(S(\nu))\, d\nu$ w.r.t $\hat{R}_{xx}(k)$ for $|k|\geq M$

We thus want $\forall |k|\geq M$, $\frac{\partial H}{\partial R(k)} = 0$: $\frac{\partial H}{\partial R(k)} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{S(\nu)} e^{-2i\pi\nu k} d\nu$ $\left(\text{car } S(\nu) = \sum_{k\in\mathbb{Z}} \underbrace{r(k)}_{\text{fixé}} e^{-2i\pi\nu k} \Rightarrow \frac{\partial H}{\partial r(k)} = \frac{1}{2}\int_{-\frac{1}{2}}^{1/2} \underbrace{\frac{\partial \ln(S(\nu))}{\partial r(k)}}_{= \frac{e^{-2i\pi\nu k}}{S(\nu)}} d\nu \right)$

Let $R_{YY}(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \underbrace{\frac{1}{S(\nu)}}_{\geq 0} e^{+2i\pi\nu k}$ ⇒ R_{YY} is the ACF of a WSS proc. Y. But we want $R_{YY}(k) = 0 \ \forall |k|\geq M$, it's the case if Y is an MA(M-1)

⇒ Therefore, there are coeff $b_0, \dots, b_{M-1}$ s.t. $\frac{1}{S(\nu)} = S_{YY}(\nu) = \sigma^2 \left| \sum_{k=0}^{M-1} b_k e^{-2i\pi\nu k} \right|^2$ ⇒ $S(\nu) = \frac{1}{S_{YY}(\nu)}$ is the PSD

3. Linear prediction method for AR estimation :

1. Reminder:

◦ Estimation of an AR proc. The most popular way of estimating the parameters of a causal AR proc. consits in predicting the value of the current sample X_t as a linear combinaison of the p previous samples: $\hat{X}_t = \sum_{m=1}^{p} a_m X_{t-m}$

The prediction coeff $a_1, \dots, a_p$ are choosen so as to minimize the MSE $\mathbb{E}|Z_t|^2$ where $Z_t = X_t - \hat{X}_t$ = prediction error

Find $a^* = \arg\min_{a_1,\dots,a_p} \mathbb{E}|X_t - \hat{X}_t|^2 \Leftrightarrow$ decorrelate Z_t from the past samples X_{t-k} for $0<k\leq p$

This approach leads to Yule-Walker equations: $\boxed{\forall k\in\{1,\dots,p\},\ R_{xx}(k) = \sum_{j=1}^{p} a_j R_{xx}(k-j) \ \text{ and } \ R_{xx}(0) = \sigma_Z^2 + \sum_{k=1}^{p} a_k R_{xx}(k)}$

→ Yule-Walker equations: In order to estimate a_j and σ^2, we first estimate R_{xx} (cov matrix) which can be defined indifferently from $X_t = [X_t, X_{t+1}, \dots, X_{t+p-1}]^T$ or from $X_t = [\bar{X}_t, \bar{X}_{t-1}, \dots, \bar{X}_{t-p+1}]^T$:

$\underbrace{X_t, X_{t+1}, \dots, X_{t+p-1}}$ forward data vector — $\underbrace{\bar{X}_t, \bar{X}_{t-1}, \dots, \bar{X}_{t-p+1}}$ backward data vector

$$R_{xx} = \begin{bmatrix} R_{xx}(0) & R_{xx}(-1) & \cdots & R_{xx}(-(p-1)) \\ R_{xx}(1) & \ddots & & \vdots \\ \vdots & & \ddots & R_{xx}(-1) \\ R_{xx}(p-1) & \cdots & R_{xx}(1) & R_{xx}(0) \end{bmatrix}$$

[Yule-Walker]: $R_{xx} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} R_{xx}(1) \\ \vdots \\ R_{xx}(p) \end{bmatrix}$ hence $\sigma_z^2 = R_{xx}(0) - \sum_{k=1}^{p} a_k R_{xx}(k)$

permet de connaître les a_i

Si on estime avec R_{xx} biaisé alors :

[YK] → The estimated AR filter $\dfrac{1}{1 - \sum_{m=1}^{p} a_m z^{-m}}$ is always causal and stable

Fast Levison Durbin ⇒ $O(p^2)$ instead of $O(p^3)$

4. Reminder - ARMA processes:

→ ARMA(p,q) proc. déf: Consider the recursive equation $\boxed{X_t - a_1 X_{t-1} - \dots - a_p X_{t-p} = Z_t + b_1 Z_{t-1} + \dots + b_q Z_{t-q}}$ [ARMA] where $Z_t \sim BB(0,\sigma^2)$, $a_j, b_j \in \mathbb{C}$

→ let $A(z) = 1 - a_1 z^{-1} - \dots - a_p z^{-p}$ and $B(z) = 1 + b_1 z^{-1} + \dots + b_q z^{-q}$

↳ Hyp: $A(z)$ and $B(z)$ don't have common zeros

⇒ Then: [ARMA] eq° admits a WSS solution iff $A(z) \neq 0 \ \forall |z| = 1$ (i.e. pas de pôle sur le cercle unité)

↳ this solution exists and is unique : $X_t = \sum_{m=-\infty}^{+\infty} h_m Z_{t-m}$

where h_m are given by : $\dfrac{B(z)}{A(z)} = \sum_{m=-\infty}^{+\infty} h_m z^{-m}$ converging in the ring $\{z \in \mathbb{C} / \delta_1 < |z| < \delta_2\}$ where $\delta_1 < 1$ and $\delta_2 > 1$

couronne de cv qui contient le cercle unité

Thm (Spectral density of an ARMA(p,q) proc.) Let $(X_t)_t$ be an ARMA(p,q) proc., i.e. the stationary solution of the equation [ARMA], then:

$$S_{XX}(\nu) = \sigma^2 \frac{\left|1+\sum_{m=1}^{q} b_m e^{-2i\pi\nu m}\right|^2}{\left|1-\sum_{m=1}^{p} a_m e^{-2i\pi\nu m}\right|^2}$$

(Rep en freq)²

DSP du BB

→ **Rpz of an ARMA(p,q) proc.** Let X_t be an ARMA(p,q) proc. solut° of [ARMA] ⇒ $X_t = \sum_{m=-\infty}^{+\infty} h_m Z_{t-m}$ for a well chosen $h_m \in \ell^1(\mathbb{Z})$

We say that the ARMA(p,q) Rpz is:
- causal if $H(z)$ is causal ($A(z) \neq 0$ $\forall |z| \geq 1$)
- invertible if $H(z)$ is invertible and $\frac{1}{H(z)}$ is causal ($B(z) \neq 0$ $\forall |z| \geq 1$)
- canonical if it's causal + invertible (i.e. $H(z)$ is minimum phase)

→ **Thm (Canonical Rpz):** Let X_t be an ARMA(p,q) proc. solution of [ARMA],
if $\forall |z|=1$, $A(z) \neq 0$ and $B(z) \neq 0$ then X_t admits a canonical Rpz

→ **Covariance of a Causal ARMA proc.** (On veut calculer l'ACF sachant les param d'un ARMA causal)

- **First method:** use $R_{XX}(k) = \sigma^2 \sum_{m=0}^{+\infty} h_{m+k} \bar{h}_m$ where h_m is determined recursively from $H(z)A(z) = B(z)$ by id. of the term in z^{-m}. For the first terms we find:

$$\begin{cases} h_0 = 1 \\ h_1 = b_1 + h_0 a_1 \\ h_2 = b_2 + h_0 a_2 + h_1 a_1 \end{cases}$$

5. Durbin method for ARMA estimation:

(Mtn on veut trouver les param. de l'ARMA sachant l'ACF)

Let X_t be a causal ARMA(p,q) proc. defined by [ARMA]

→ **Estimation of the AR part** (modified Yule-Walker Method) By multiplying the equation [ARMA] with $\bar{X}_{t-m}$ for $m > q$ and by applying the mathematical expectation, we get:

$$R_{XX}(m) - \sum_{k=1}^{p} a_k R_{XX}(m-k) = \sum_{k=0}^{q} b_k \underbrace{\mathbb{E}[\bar{X}_{t-m} Z_{t-k}]}_{=0 \ \forall m>q \text{ bc } X_t \text{ is causal}}$$

$\Rightarrow R_{xx}(m) = \sum_{k=1}^{p} a_k R_{xx}(m-k) \Rightarrow \forall m \in \{q+1, \dots, q+p\},$ $\begin{bmatrix} R_{xx}(q+1) \\ \vdots \\ R_{xx}(q+p) \end{bmatrix} = \begin{bmatrix} R_{xx}(q) & R_{xx}(q-1) & \cdots & R_{xx}(q-p+1) \\ R_{xx}(q+1) & \ddots & & R_{xx}(q-1) \\ \vdots & & \ddots & \\ R_{xx}(q+p-1) & \cdots & R_{xx}(q+1) & R_{xx}(q) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}$ [Modified YW]

$\Rightarrow$ solve to find estimates $\hat{a}_1, \dots, \hat{a}_p$

$\rightarrow$ Estimation of the MA part:

- First approach: estimation of the PSD:

Consider the WSS proc.: $Y_t = X_t - a_1 X_{t-1} - \dots - a_p X_{t-p} \Rightarrow S_{YY}(\nu) = S_{XX}(\nu) |A(e^{2i\pi\nu})|^2$ where $A(\nu) = 1 - \sum_{m=1}^{p} a_m e^{-2i\pi\nu m}$

However, [ARMA] $\Rightarrow Y_t = Z_t + b_1 Z_{t-1} + \dots + b_q Z_{t-q} = MA(q) \Rightarrow \hat{S}_{YY}(\nu) = \sum_{k=-q}^{q} \hat{R}_{YY}(k) e^{-2i\pi\nu k}$

then we get a first ARMA PSD estimate: $\hat{S}_{XX}(\nu) = \dfrac{\sum_{k=-q}^{q} \hat{R}_{YY}(k) e^{-2i\pi\nu k}}{|\hat{A}(e^{2i\pi\nu})|^2}$ where $\hat{A}(e^{2i\pi\nu}) = \sum_{m=1}^{p} \hat{a}_m e^{-2i\pi\nu m}$

Problem: the numerator is not necessarily ≥ 0

- Second approach: Durbin method: let $R_{YY}(k)$ the ACF of the MA proc. Y_t. We want to find $\hat{b}_0, \dots, \hat{b}_q, \hat{\sigma}_Z^2$ s.t.

$\hat{S}_{YY}(\nu) = \hat{\sigma}_Z^2 |\hat{B}(e^{2i\pi\nu})|^2$ with $\hat{B}(z) = 1 + \sum_{m=1}^{q} \hat{b}_m z^{-m}$ $\rightarrow$ ds le cours c'est un q...

$\Rightarrow$ To do so, we solve Yule-Walker equations to find an AR(L) model that fits $\hat{R}_{YY}(k)$ for $k = 0, \dots, L \gg q$

Thus: [YK] $\Rightarrow [\hat{a}_{1,L}, \dots, \hat{a}_{L,L}]$ s.t. $\hat{S}_{YY}(\nu) = \dfrac{\hat{\sigma}_1^2}{|\hat{A}_L(\nu)|^2}$ with $\hat{A}_L(\nu) = 1 - \sum_{m=1}^{L} \hat{a}_{m,L} z^{-m}$

then: we consider the sequence $\hat{R}_L(k)$ s.t. $|\hat{A}_L(\nu)|^2 = \sum_{k=-L}^{L} \hat{R}_L(k) e^{-2i\pi\nu k}$

We now solve a another set of YK-eq° to find a AR(q) model that fits $\hat{R}_L(k)$

thus: [YK] $\Rightarrow [\hat{b}_1, \dots, \hat{b}_q]$ s.t. $|\hat{A}_L(\nu)|^2 = \dfrac{\hat{\sigma}_2^2}{|\hat{B}(\nu)|^2}$ with $\hat{B}(z) = 1 + \sum_{m=1}^{q} \hat{b}_m z^{-m}$ $\longrightarrow$ ds le cours c'est un q hence : $\hat{S}_{YY}(\nu) = \dfrac{\sigma_1^2}{|\hat{A}_L(e^{2i\pi\nu})|^2} = \dfrac{\sigma_1^2}{\sigma_2^2}|\hat{B}(e^{2i\pi\nu})|^2 \Rightarrow$ ARMA PSD estimate : $\hat{S}_{XX}(\nu) = \sigma_Z^2 \dfrac{|\hat{B}(e^{2i\pi\nu})|^2}{|\hat{A}(e^{2i\pi\nu})|^2}$ with $\sigma_Z^2 = \dfrac{\sigma_1^2}{\sigma_2^2}$

Partie 3 : Filter bank methods

1. The periodogram as a filter bank

We know that the periodogram $\hat{S}_{XX}(\nu)$ of a centred COSS X_t is : $\hat{S}_{XX}(\nu) = \dfrac{1}{N}\left|\sum_{t=0}^{N-1} X_t e^{-2i\pi\nu t}\right|^2$

prop (filter bank interpretation of the periodogram) $\hat{S}_{XX}(\nu)$ can also be expressed as : $\hat{S}_{XX}(\nu) = N\left|\sum_{k\in\mathbb{Z}} w_k^\nu X_{N-1-k}\right|^2 = \dfrac{|Y_{N-1}|^2}{1/N}$

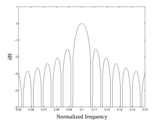
where : $Y_t = w^\nu * X_t$ and $w_k^\nu = \begin{cases} \frac{1}{N}e^{2i\pi\nu k} & \text{if } k \in [\![0, N-1]\!] \\ 0 & \text{if } k \notin [\![0, N-1]\!] \end{cases}$ RIF

$X_t \longrightarrow \boxed{w^\nu} \longrightarrow Y_t$

$\left(\Rightarrow Y_n = \sum_{k\in\mathbb{Z}} w_n^\nu X_{k-n} = \sum_{k=0}^{N-1} w_n^\nu X_{n-k} \Rightarrow Y_{N-1} = \sum_{k=0}^{N-1} w_k^\nu X_{N-1-k} = \sum_{t=0}^{N-1} X_t \underbrace{w_{N-t}^\nu}_{\frac{1}{N}e^{2i\pi\nu(N-t)}}\right)$

$t = N-1-k$

▶ w^ν is such that $|W^\nu(\xi)| = \left|\sum_{k\in\mathbb{Z}} w_k^\nu e^{-i2\pi\xi k}\right| = \frac{1}{N}\left|\frac{\sin(\pi(\xi-\nu)N)}{\sin(\pi(\xi-\nu))}\right|$



Frequency response $\xi \mapsto |W^\nu(\xi)|$ with $\nu = 0.1$

on veut minimiser $\mathbb{E}|Y_{N-1}|^2 = \int |W^\nu(\varepsilon)|^2 S_{XX}(\varepsilon)\, d\varepsilon$

Perceval

↳ baisser là où il y'a trop d'energie en dehors de ν

- The periodogram is $\Leftrightarrow$ to filtering the signal X_t by a FIR filter, and computing the energy of the $(N-1)^{th}$ output sample
- The resulting frequency response W^ν has a main lobe with a small width $(2/N)$ but side lobes with high amplitudes
- Y has a non-zero power on a large frequency band arround ν, and the estimation of $\hat{S}_{XX}(\nu)$ is biased

2. Capon's method :

consists in determining a filter w^ν s.t. $W^\nu(\nu) = 1$, which minimize the energy of the output signal at frequencies other than ν (en fait, on veut un dirac mais propre $\Rightarrow W^\nu(\nu)=1$ et tue le reste)

let $Y = w^\nu * X$ where :
- w^ν is a FIR filter of support $[\![0, N-1]\!]$ $\Rightarrow$ $w^\nu = \{w_0^\nu, \dots, w_{N-1}^\nu\}$
- X_t a centred WSS proc. $\Rightarrow$ $X = \{\overline{X_{N-1}}, \dots, \overline{X_0}\}$

and let
- $R_{XX} = \mathbb{E}[XX^H]$ the covariance matrix, we assume that R_{XX} is positive semi-definite, thus non-singular
- $e(\varepsilon) = [1, e^{2i\pi\varepsilon}, \dots, e^{2i\pi\varepsilon(N-1)}]^T$ $\Rightarrow$ $W^\nu(\varepsilon) = e(\varepsilon)^H w^\nu$ and we assume that $e(\nu)^H w^\nu = W^\nu(\nu) = 1$
- $w_{opt}^\nu = \dfrac{R_{XX}^{-1}\, e(\nu)}{e(\nu)^H R_{XX}^{-1}\, e(\nu)}$

- $Y_{N-1} = \sum_{k=0}^{N-1} w_k^\nu \, \overline{(X_{N-1-k})} = X^H w^\nu$ $\Rightarrow$ $\mathbb{E}|Y_{N-1}|^2 = \mathbb{E}\, Y_{N-1}\overline{Y_{N-1}} = \mathbb{E}\, w^{\nu H} X X^H w^\nu = w^{\nu H} \underbrace{\mathbb{E}[XX^H]}_{R_{XX}} w^\nu = w^{\nu H} R_{XX}\, w^\nu$
- $e(\varepsilon)^H w^\nu = \sum_{k=0}^{N-1} e^{-2i\pi\varepsilon k} w_k^\nu = W^\nu(\varepsilon)$
- $(w^\nu - w_{opt}^\nu)^H R_{XX} (w^\nu - w_{opt}^\nu) + \dfrac{1}{e(\nu)^H R_{XX}^{-1} e(\nu)} = w^{\nu H} R_{XX} w^\nu - \dfrac{\overset{W^\nu(\nu)=1}{\cancel{e(\nu)^H w^\nu}}}{e(\nu)^H R_{XX}^{-1} e(\nu)} - \dfrac{\overset{W^\nu(\nu)=1}{\cancel{w^{\nu H} e(\nu)}}}{e(\nu)^H R_{XX}^{-1} e(\nu)} + \dfrac{\cancel{e(\nu)^H R_{XX}^{-1} R_{XX} R_{XX}^{-1} e(\nu)}\ 1}{(e(\nu)^H R_{XX}^{-1} e(\nu))^{\cancel{2}}} + \dfrac{1}{e(\nu)^H R_{XX}^{-1} e(\nu)} = w^{\nu H} R_{XX}\, w^\nu = \mathbb{E}|Y_{N-1}|^2$

(MVDR filter)
$\Rightarrow$ The MVDR filter which minimizes $\mathbb{E}|Y_{N-1}|^2$ (Minimum Variance) subject to $\underbrace{e(\nu)^H w^\nu}_{W^\nu(\nu)} = 1$ (Distortionless Response) is $w^\nu = w_{opt}^\nu$

$\Rightarrow$ The energy at the output is $\mathbb{E}[|Y_{N-1}|^2] = \dfrac{1}{e(\nu)^H R_{XX}^{-1} e(\nu)}$

$\Rightarrow$ Capon's spectral estimator is : $\hat{S}_{CAP,XX}(\nu) = \dfrac{\mathbb{E}|Y_{N-1}|^2}{1/N} \overset{w^\nu = w_{opt}^\nu}{=} \dfrac{w_{opt}^{\nu H} R_{XX}\, w_{opt}^\nu}{1/N} = \dfrac{N}{e(\nu)^H R_{XX}^{-1} e(\nu)}$

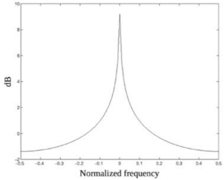
↳ In practice R_{XX} is unknown and has to be estimated

eg: if X_t is a white noise, then $R_{XX} = \sigma_X^2 I_N$ $\Rightarrow$ $w_{opt}^\nu = \dfrac{R_{XX}^{-1} e(\nu)}{e(\nu)^H R_{XX}^{-1} e(\nu)} = \dfrac{1}{N} e(\nu)$ = Rep en freq du periodogramm

Example : AR(1) process

- $X_t = a_1 X_{t-1} + Z_t$ where $a_1 \in]0,1[$ and $Z_t \sim \mathrm{WN}(0, \sigma_Z^2)$
- Then $S_{XX}(\nu) = \frac{\sigma_Z^2}{|1 - a_1 e^{-i2\pi\nu}|^2}$
- $\mathbf{R}_{XX} = \frac{\sigma_Z^2}{1-a_1^2} \begin{bmatrix} 1 & a_1 & \dots & a_1^{N-1} \\ a_1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ a_1^{N-1} & \dots & a_1 & 1 \end{bmatrix}$

Example : AR(1) process



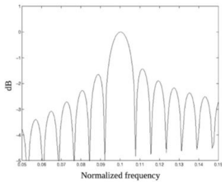
PSD of the AR process : $\nu \mapsto |S_{XX}(\nu)|$ with $a_1 = 0.99$ and $\sigma_Z^2 = 1$

Example : AR(1) process

- $\mathbf{R}_{XX}^{-1} = \frac{1}{\sigma_Z^2} \begin{bmatrix} 1 & -a_1 & 0 & \dots & 0 \\ -a_1 & 1+a_1^2 & -a_1 & \ddots & \vdots \\ 0 & -a_1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1+a_1^2 & -a_1 \\ 0 & \dots & 0 & -a_1 & 1 \end{bmatrix}$
- Therefore $\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu) = \frac{\left|1 - a_1 e^{-i2\pi\nu}\right|^2}{\sigma_Z^2} (\mathbf{e}(\nu) + \mathbf{v}(\nu))$
  where $\mathbf{v}(\nu) = a_1 [\frac{e^{-i2\pi\nu}}{1-a_1 e^{-i2\pi\nu}}, 0, \dots, 0, \frac{e^{+i2\pi N\nu}}{1-a_1 e^{+i2\pi\nu}}]^\top$
  and $\mathbf{w}_{\mathrm{opt}}^{\nu} = \frac{\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)}{\mathbf{e}(\nu)^H \mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)} = \frac{\mathbf{e}(\nu) + \mathbf{v}(\nu)}{N + 2a_1 \frac{\cos(2\pi\nu) - a_1}{1 - 2a_1\cos(2\pi\nu) + a_1^2}}$

Example : AR(1) process

- The *shape* of the filter depends on the estimated frequency ν :



Frequency response $\xi \mapsto |W^{\nu}(\xi)|$ with $\nu = 0.1$ and $a_1 = 0.99$

→ Statiscal properties: • This estimator has a better resolution than the periodogram

• it's variance is lower than that of autoregressive methods, but its spectral resolution is worse

prop: if $\hat{S}_{CAP,XX}$ is calculated with a $N \times N$ matrix R_{XX}^{-1}, then $\hat{S}_{CAP,XX}$ is related to AR estimators through:

$$\frac{1}{\hat{S}_{CAP,XX}(\nu)} = \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{\hat{S}^{(k)}_{AR,XX}(\nu)}$$

where $\hat{S}^{(k)}_{AR,XX}(\nu)$ is the AR estimator of order k, with cov matrix defined as the $k \times k$ lower right block of $\hat{R}_{XX}$

Variant (Lagunas)

- Compute the ratio between the obtained power and that of white noise filtered by $\mathbf{w}_{\text{opt}}^{\nu}$
- $\mathbf{w}_{\text{opt}}^{\nu}{}^{H}\mathbf{I}\mathbf{w}_{\text{opt}}^{\nu} = \frac{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}}{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)}\frac{\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)}{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)} = \frac{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-2}\mathbf{e}(\nu)}{(\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu))^2}$ and $\frac{\mathbf{w}_{\text{opt}}^{\nu}{}^{H}\mathbf{R}_{XX}^{-1}\mathbf{w}_{\text{opt}}^{\nu}}{\mathbf{w}_{\text{opt}}^{\nu}{}^{H}\mathbf{I}\mathbf{w}_{\text{opt}}^{\nu}} = \frac{1}{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)}\frac{(\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu))^2}{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-2}\mathbf{e}(\nu)} = \frac{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)}{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-2}\mathbf{e}(\nu)}$
- Hence

$$\hat{S}_{\text{LAG},XX}(\nu) = \frac{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-1}\mathbf{e}(\nu)}{\mathbf{e}(\nu)^H\mathbf{R}_{XX}^{-2}\mathbf{e}(\nu)}$$

TELECOM Paris