Parametric estimation of rational spectra

Roland Badeau,
roland.badeau@telecom-paris.fr

Linear time series (part 2)
TSIA202b

Linear process

- Definition (linear process)
  - $(X_t)_{t\in\mathbb{Z}}$ is a linear process iif there is $\mu_X \in \mathbb{C}$, $Z_t \sim \mathrm{BB}(0,\sigma^2)$ and $(h_n)_{n\in\mathbb{Z}} \in l_1(\mathbb{Z})$ such that $X_t = \mu_X + \sum_{n=-\infty}^{+\infty} h_n Z_{t-n}$ $\forall t \in \mathbb{Z}$
  - $(X_t)_{t\in\mathbb{Z}}$ is causal with respect to $(Z_t)_{t\in\mathbb{Z}}$ iif $h_n = 0$ $\forall n < 0$
  - $(X_t)_{t\in\mathbb{Z}}$ is invertible with respect to $(Z_t)_{t\in\mathbb{Z}}$ iif there is a sequence $(g_n)_{n\geq 0} \in l_1(\mathbb{Z})$ such that $Z_t = \sum_{n=0}^{+\infty} g_n(X_{t-n} - \mu_X)$ $\forall t \in \mathbb{Z}$
- Properties (filtering theorem for WSS processes)
  - X_t is WSS of mean μ_X, autocovariance function $r_{XX}(k) = \mathbb{E}[(X_{t+k} - \mu_X)(\overline{X}_t - \overline{\mu_X})] = \sigma^2 \sum_{n=-\infty}^{+\infty} h_{n+k}\overline{h}_n$, and spectral density $S_{XX}(\nu) = \sigma^2 |H(e^{-2i\pi\nu})|^2$ with $H(e^{-2i\pi\nu}) = \sum_{n\in\mathbb{Z}} h_n e^{-2i\pi\nu n}$

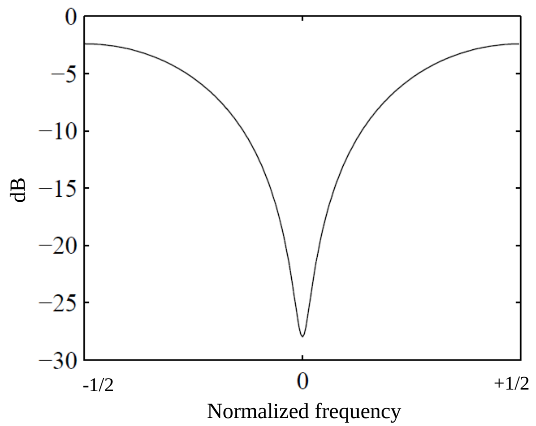
Part I

Reminder: moving average processes

MA(q) process

- Definition
  - The process $(X_t)_{t\in\mathbb{Z}}$ is moving average of order q (or MA(q)) iif $X_t = \sum_{n=0}^{q} b_n Z_{t-n}$ where $Z_t \sim \mathrm{BB}(0,\sigma^2)$, $b_n \in \mathbb{C}$ and $b_0 = 1$.
- Properties (filtering theorem for WSS processes)
  - X_t is WSS of mean 0, of autocovariance function $r_{XX}(k) = \sigma^2 \sum_{n=0}^{q-k} b_{n+k}\overline{b}_n$ for $0 \leq k \leq q$ and $r_{XX}(k) = 0$ if $k > q$, and of spectral density $S_{XX}(\nu) = \sigma^2 \left|\sum_{n=0}^{q} b_n e^{-2i\pi\nu n}\right|^2$.

MA(q) process



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

Characterization of an MA(q) process

- Theorem (characterization of an MA(q) process).
  - Let $(X_t)_{t\in\mathbb{Z}}$ be a centered WSS process of autocovariance function $r_{XX}(k)$, and let $q \geq 1$. Then the two following assertions are equivalent :
    - X_t is an MA process of minimal order q ;
    - $r_{XX}(q) \neq 0$ and $r_{XX}(k) = 0 \ \forall k \geq q+1$.
- Corollary
  - The sum of two decorrelated MA(q) processes is an MA(q) process.

Part II

Reminder: autoregressive processes

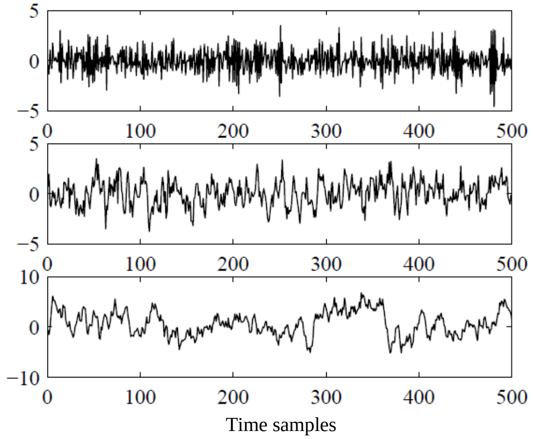
AR(p) process

- Definition
  - The process $(X_t)_{t\in\mathbb{Z}}$ is autoregressive of order p (or AR(p)) iif it is WSS and solution of the equation $X_t = Z_t + \sum_{n=1}^{p} a_n X_{t-n}$ where $Z_t \sim \mathrm{BB}(0, \sigma^2)$, $a_n \in \mathbb{C}$
- The existence and unicity of a WSS solution is a difficult question, which did not exist for the MA process.

AR(1) process, causal case

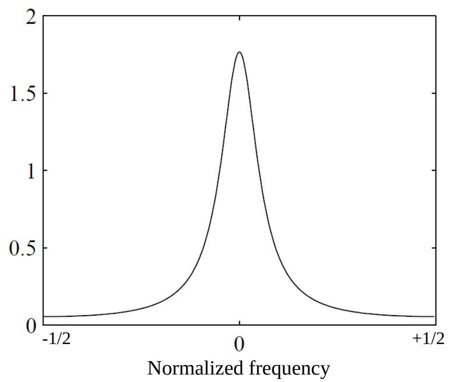
- We apply the recurrence $X_t = Z_t + a_1 X_{t-1}$ with $|a_1| < 1$
- $X_t = \sum_{n=0}^{+\infty} a_1^n Z_{t-n}$ (convergence in L^2 and a.s.)
- Properties (filtering theorem for WSS processes)
  - X_t is WSS of mean 0, of autocovariance function $r_{XX}(k) = \sigma^2 \sum_{n=0}^{+\infty} a_1^{n+k} \overline{a}_1^n = \sigma^2 \frac{a_1^k}{1-|a_1|^2}$ if $k \geq 0$, and of spectral density $S_{XX}(\nu) = \sigma^2 \left| \sum_{n=0}^{+\infty} a_1^n e^{-2i\pi\nu n} \right|^2 = \frac{\sigma^2}{|1-a_1 e^{-2i\pi\nu}|^2}$.

AR(1) process, causal case



Trajectories of a Gaussian AR(1) process, of length 500. Top : $a_1 = -0.7$. Center : $a_1 = 0.5$. Bottom : $a_1 = 0.9$.

AR(1) process, causal case



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

AR(1) process, anti-causal case

- We apply the recurrence $X_t = -a_1^{-1} Z_{t+1} + a_1^{-1} X_{t+1}$ with $|a_1| > 1$
- $X_t = -\sum_{n=1}^{+\infty} a_1^{-n} Z_{t+n}$ (convergence in L^2 and a.s.)
- Properties (filtering theorem for WSS processes)
  - X_t is WSS of mean 0, of autocovariance function $r_{XX}(k) = \sigma^2 \sum_{n=-\infty}^{-1} a_1^n \overline{a}_1^{(n-k)} = \sigma^2 \frac{\overline{a}_1^{-k}}{|a_1|^2-1}$ if $k \geq 0$, and of spectral density $S_{XX}(\nu) = \sigma^2 \left| \sum_{n=0}^{+\infty} a_1^n e^{-2i\pi\nu n} \right|^2 = \frac{\sigma^2}{|1-a_1 e^{-2i\pi\nu}|^2}$.

AR(1) process, general case

- If $|a_1| < 1$, $X_t = \sum_{n=0}^{+\infty} a_1^n Z_{t-n}$
- If $|a_1| > 1$, $X_t = -\sum_{n=1}^{+\infty} a_1^{-n} Z_{t+n}$
- Properties (filtering theorem for WSS processes)
  - If $|a_1| \neq 1$, X_t is WSS of mean 0 and of spectral density $S_{XX}(\nu) = \sigma^2 \frac{1}{|1-a_1 e^{-2i\pi\nu}|^2}$.
- If $|a_1| = 1$, the recursive equation does not admit any WSS solution

Properties of the order p AR process

- There is a WSS solution iif $A(z) \neq 0$ for $|z| = 1$
  $\frac{1}{A(z)} = \sum_{n=-\infty}^{+\infty} h_n z^{-n}$ where $\sum_{n\in\mathbb{Z}} |h_n| < +\infty$
  $\Rightarrow X_t = \sum_{n\in\mathbb{Z}} h_n Z_{t-n}$
- If $A(z) = 0 \Rightarrow |z| < 1$, causal solution
- If $A(z) = 0 \Rightarrow |z| > 1$, anti-causal solution
- Otherwise, X is a mixed AR process

Part III

Maximum entropy spectral estimation

Maximum entropy spectral estimation

- Let X_t be a centered WSS process such that $r_{XX} \in l^1(\mathbb{Z})$
- Non-parametric spectral estimation : estimate $S_{XX}(\nu)$ from N samples $X_1 \dots X_N$
- Periodogram, Blackman-Tukey methods : with $M \leq N$,
  - 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 $\hat{r}_{XX}(k)$.
- 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 maximizes the *entropy* of the WSS probability distribution
- *Blind* estimation : no information is available about the WSS process beyond the knowledge of $\hat{r}_{XX}(k)$ for $k \in [-M+1, M-1]$

Entropy of a Gaussian random vector

- Ref. : "Nonlinear Methods of Spectral Analysis", S. Haykin Ed., in "Topics in Applied Physics", Vol. 34, Springer, 1983, chap. 2 p. 67
- For a discrete random variable (r.v.) with M values,
$$H = \sum_{k=1}^{M} p_k \log_2(\tfrac{1}{p_k}) = \tfrac{1}{\ln(2)} \sum_{k=1}^{M} p_k \ln(\tfrac{1}{p_k})$$
- For N continuous variables $x_1, \dots, x_N$,
$H_N = -\int p(x_1, \dots, x_N) \ln(p(x_1, \dots, x_N) c^{\frac{N}{2}}) dx_1 \dots dx_N$
- If the variables are Gaussian,
$p(x_1, \dots, x_N) = \frac{1}{(2\pi)^{\frac{N}{2}} \det(\mathbf{R}_N)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_X)^\top \mathbf{R}_N^{-1} (\mathbf{x} - \mu_X)\right)$
- If we choose c appropriately, then $H_N = \frac{1}{2}\ln(\det(\mathbf{R}_N))$

Entropy of a Gaussian WSS process

- Problem : H_N diverges when $N \to +\infty$
- For a process of infinite length (chap. 2 p. 16), the *entropy rate* is
$$H = \lim_{N\to+\infty} \frac{H_N}{N} = \lim_{N\to+\infty} \frac{1}{2}\ln((\det(\mathbf{R}_N))^{\frac{1}{N}})$$
- Szegö theorem (to be admitted) : if X_t is a WSS process, if $\sigma_0^2 \dots \sigma_{N-1}^2$ are the eigenvalues of $\mathbf{R}_N$, and if g is any continuous function,
$\lim_{N\to+\infty} \frac{1}{N}(g(\sigma_0^2) + \dots + g(\sigma_{N-1}^2)) = \int_{-\frac{1}{2}}^{+\frac{1}{2}} g(S_{XX}(\nu)) d\nu$
- With $g(.) = \ln(.)$, we get $H = \frac{1}{2}\int_{-\frac{1}{2}}^{+\frac{1}{2}} \ln(S_{XX}(\nu)) d\nu$

Maximum entropy method

- Among all WSS processes with fixed $\hat{r}_{XX}(k)$, $|k| < M$ as autocorrelations, which one maximizes the entropy ?
- Response : the AR$(M-1)$ process. Proof :
  - Let $r(k)$ be the autocovariance of a WSS process and $S(\nu)$ its PSD, with $r(k) = \hat{r}_{XX}(k) \forall |k| < M$ and $H = \frac{1}{2}\int_{-\frac{1}{2}}^{+\frac{1}{2}} \ln(S(\nu)) d\nu$ is maximum
  - We thus want $\forall |k| \geq M$, $\frac{\partial H}{\partial r(k)} = 0$
  - However $\frac{\partial H}{\partial r(k)} = \int_{-\frac{1}{2}}^{+\frac{1}{2}} \frac{1}{S(\nu)} e^{-2i\pi\nu k} d\nu$
  - Therefore $r_{YY}(k) \triangleq \int_{-\frac{1}{2}}^{+\frac{1}{2}} \frac{1}{S(\nu)} e^{+2i\pi\nu k} d\nu = 0 \ \forall |k| \geq M$ is the autocovariance function of an MA$(M-1)$ process Y_t
  - Therefore $\frac{1}{S(\nu)} = S_{YY}(\nu) = \sigma^2 \left| \sum_{k=0}^{M-1} b_k e^{-2i\pi\nu k} \right|^2$
  - Finally, $S(\nu) = \frac{1}{S_{YY}(\nu)}$ is the PSD of an AR$(M-1)$ process

Part IV

Reminder: Linear prediction method for AR estimation

Estimation of an autoregressive process

- Ref. : "Spectral analysis of signals", P. Stoica and R. Moses, Prentice Hall, 2005, chap. 3
- Linear predicton of a causal AR process :
  $\widehat{X}_t = \sum_{m=1}^{p} a_m X_{t-m}$ is an estimation of X_t from the past samples
- The estimation error $Z_t = X_t - \widehat{X}_t$ is decorrelated from all the X_{t-k} (i.e. $\operatorname{cov}(Z_t, X_{t-k}) = 0$) for $k > 0$
- We deduce that $\forall k \geq 1$, $r_{XX}(k) = \sum_{j=1}^{p} a_j r_{XX}(k-j)$ 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 $\mathbf{R}_{XX}$:

$$\mathbf{R}_{XX} = \begin{bmatrix} r_{XX}(0) & r_{XX}(-1) & \dots & r_{XX}(-(p-1)) \\ r_{XX}(1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{XX}(-1) \\ r_{XX}(p-1) & \dots & r_{XX}(1) & r_{XX}(0) \end{bmatrix}$$

- We then solve the linear system of equations

$$\mathbf{R}_{XX} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} r_{XX}(1) \\ r_{XX}(2) \\ \vdots \\ r_{XX}(p) \end{bmatrix} \text{ hence } \sigma_Z^2 = r_{XX}(0) - \sum_{k=1}^{p} a_k r_{XX}(k)$$

- The estimated AR filter $\frac{1}{1 - \sum_{m=1}^{p} a_m z^{-m}}$ is always causal and stable
- Fast Levinson-Durbin algorithm, in $O(p^2)$ instead of $O(p^3)$

Part V

Reminder: ARMA processes

ARMA(p,q) process

- Theorem (existence and unicity of the ARMA$(p,q$ process))
  - Consider the recursive equation :
    $X_t - a_1 X_{t-1} - \ldots - a_p X_{t-p} = Z_t + b_1 Z_{t-1} + \ldots + b_q Z_{t-q}$,
    where $Z_t \sim \mathrm{BB}(0, \sigma^2)$ and $a_j, b_j \in \mathbb{C}$.
  - Let $A(z) = 1 - a_1 z^{-1} - \ldots - a_p z^{-p}$ and
    $B(z) = 1 + b_1 z^{-1} + \ldots + b_q z^{-q}$.
  - We assume that $A(z)$ and $B(z)$ do not have common zeros.
  - Then the equation admits a WSS solution iif $A(z) \neq 0 \ \forall |z| = 1$.
  - This solution is unique and its expression is $X_t = \sum_{n=-\infty}^{+\infty} h_n Z_{t-n}$,
    where the h_n are given by the coefficients of the expansion
    $\frac{B(z)}{A(z)} = \sum_{n=-\infty}^{+\infty} h_n z^{-n}$, converging in the ring
    $\{z \in \mathbb{C}, \delta_1 < |z| < \delta_2\}$, where $\delta_1 < 1$ and $\delta_2 > 1$ are defined by
    $\delta_1 = \max\{z \in \mathbb{C}, |z| < 1, A(z) = 0\}$ and
    $\delta_2 = \min\{z \in \mathbb{C}; |z| > 1; A(z) = 0\}$.

Spectral density of an ARMA process

- Theorem (spectral density of an ARMA(p,q) process).
  - Let (X_t) be an ARMA(p,q) process, i.e. the stationary solution of the equation $X_t - a_1X_{t-1} - \ldots - a_pX_{t-p} = Z_t + b_1Z_{t-1} + \ldots + b_qZ_{t-q}$, where $B(z)$ and $A(z)$ are polynomials of degree q and p which do not have common zeros and $A(z) \neq 0\ \forall |z| = 1$. Then (X_t) has a spectral density whose expression is :

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

Representations of an ARMA(p,q) process

- Let X_t be an ARMA(p,q) process solution of $X_t - a_1X_{t-1} - \ldots - a_pX_{t-p} = Z_t + b_1Z_{t-1} + \ldots + b_qZ_{t-q}$.
- Then X_t admits a linear representation $X_t = \sum_{n=-\infty}^{+\infty} h_n Z_{t-n}$ for a well chosen sequence $h_n \in l^1(\mathbb{Z})$.
- We say that the ARMA(p,q) representation is
  - causal if the filter $H(z)$ is causal ($A(z) \neq 0\ \forall |z| \geq 1$)
  - invertible if the filter $H(z)$ is invertible and if its inverse is causal ($B(z) \neq 0\ \forall |z| \geq 1$)
  - canonical if it is causal and invertible (i.e. $H(z)$ is minimum phase)
- Theorem (canonical representation)
  - Let X_t be an ARMA(p,q) process solution of $X_t - a_1X_{t-1} - \ldots - a_pX_{t-p} = Z_t + b_1Z_{t-1} + \ldots + b_qZ_{t-q}$.
  - We assume that $A(z) \neq 0$ and $B(z) \neq 0\ \forall |z| = 1$
  - Then X_t admits a canonical representation

Covariances of a causal ARMA process

- First method
  - Use the expression $r_{XX}(k) = \sigma^2 \sum_{n=0}^{+\infty} h_{n+k}\overline{h}_n$ where h_n is determined recursively from $H(z)A(z) = B(z)$, by identification of the term in z^{-n}. For the first terms we find :

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

- Second method
  - Use a recursion formula, verified by the autocovariance function of an ARMA(p,q) process, which is obtained by multiplying by $\overline{X}_{t-k}$ the two members of $X_t - a_1X_{t-1} - \ldots - a_pX_{t-p} = Z_t + b_1Z_{t-1} + \ldots + b_qZ_{t-q}$, and by taking the mathematical expectation.

Part VI

Durbin method for ARMA estimation

ARMA estimation

- Let X_t be a causal ARMA(p,q) process such that

$$X_t - a_1 X_{t-1} - \ldots - a_p X_{t-p} = b_0 Z_t + b_1 Z_{t-1} + \ldots + b_q Z_{t-q}$$

  where $b_0 = 1$ and $Z_t \sim BB(0, \sigma_Z^2)$
- How to estimate the ARMA parameters ?

Estimation of the AR part

Modified Yule-Walker Method :

$$\mathbb{E}[\overline{X_{t-n}} Z_{t-q}] = 0 \ \forall n > q \Rightarrow \begin{bmatrix} r_{XX}(q+1) \\ \vdots \\ r_{XX}(q+p) \end{bmatrix} =$$

$$\begin{bmatrix} r_{XX}(q) & r_{XX}(q-1) & \ldots & r_{XX}(q-p+1) \\ r_{XX}(q+1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{XX}(q-1) \\ r_{XX}(q+p-1) & \ldots & r_{XX}(q+1) & r_{XX}(q) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}$$

Estimation of the MA part : first approach

- Let $Y_t = X_t - a_1 X_{t-1} - \ldots - a_p X_{t-p} = b_0 Z_t + b_1 Z_{t-1} + \ldots + b_q Z_{t-q}$
- $r_{YY}(k) = \begin{cases} \sigma_Z^2 (b_0 b_k + b_1 b_{k+1} + \ldots + b_{q-k} b_q) \text{ if } k \leq q \\ 0 \text{ if } k > q \end{cases}$
- $S_{YY}(\nu) = S_{XX}(\nu) |A(e^{2i\pi\nu})|^2 = \sum_{k=-q}^{q} r_{YY}(k) e^{-2i\pi\nu k}$
- First ARMA PSD estimate : $\hat{S}_{XX}(\nu) = \frac{\sum_{k=-q}^{q} \hat{r}_{YY}(k) e^{-2i\pi\nu k}}{|\hat{A}(e^{2i\pi\nu})|^2}$
- Problem : the numerator is not necessarily non-negative

Estimation of the MA part : Durbin method

- Let $\hat{r}_{YY}(k)$. We want to find $\hat{b}_0 \ldots \hat{b}_q, \hat{\sigma}_Z^2$ such that $\hat{S}_{YY}(\nu) = \hat{\sigma}_Z^2 |\hat{B}(e^{2i\pi\nu})|^2$
- Solve Yule-Walker equations to find an AR(L) on the $\hat{r}_{YY}(k)$ for $k = 0 \ldots L \gg q$
- $\rightarrow [\hat{a}_{1,L}, \ldots, \hat{a}_{L,L}]$ such that $\hat{S}_{YY}(\nu) = \frac{\sigma_1^2}{|\hat{A}_L(e^{2i\pi\nu})|^2}$ then let $\hat{r}_L(k)$ be the sequence such that $|\hat{A}_L(e^{2i\pi\nu})|^2 = \sum_{k=-L}^{L} \hat{r}_L(k) e^{-2i\pi\nu k}$
- On the $\hat{r}_L(k)$, estimate an AR(q) by solving Yule-Walker equations
- $\rightarrow [\hat{b}_1, \ldots, \hat{b}_q]$ such that $|\hat{A}_L(e^{2i\pi\nu})|^2 = \frac{\sigma_2^2}{|\hat{B}(e^{2i\pi\nu})|^2}$ with $\hat{B}(z) = 1 + \sum_{n=1}^{q} \hat{b}_n z^{-q}$, hence $\hat{S}_{YY}(\nu) = \frac{\sigma_1^2}{|\hat{A}_L(e^{2i\pi\nu})|^2} = \frac{\sigma_1^2}{\sigma_2^2} |\hat{B}(e^{2i\pi\nu})|^2$
- ARMA PSD estimate : $\hat{S}_{XX}(\nu) = \hat{\sigma}_Z^2 \frac{|\hat{B}(e^{2i\pi\nu})|^2}{|\hat{A}(e^{2i\pi\nu})|^2}$ with $\hat{\sigma}_Z^2 = \frac{\sigma_1^2}{\sigma_2^2}$