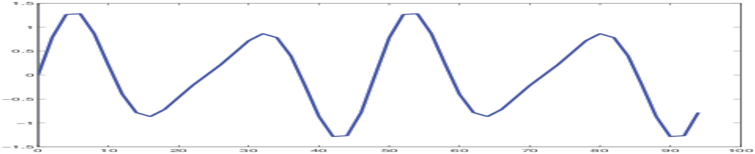
Introduction to Communication, Control and Signal Processing

Ali Mohammad-Djafari
Laboratoire des Signaux et Systèmes (L2S)
UMR8506 CNRS-CentraleSupélec-UNIV PARIS SUD
SUPELEC, 91192 Gif-sur-Yvette, France
http://lss.centralesupelec.fr
Email: djafari@lss.supelec.fr
http://djafari.free.fr
http://publicationslist.org/djafari

Contents

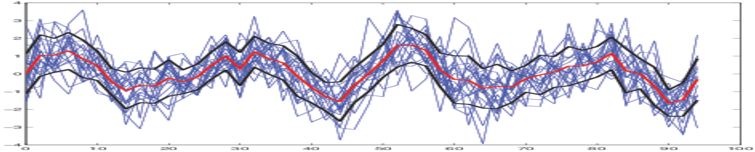
Deterministic and probabilistic modelling

- Deterministic: the value of $X(t)$ at time t is always the same.



$$X(t) = \sin(2\omega t) + .4\sin(3\omega t),\ \omega = \frac{\pi}{12}$$

- Stochastic or Random: the value of $X(t)$ at time t is not always the same.
  $X(t)$ is defined as a random variable with a probability law, mean, variance, ...



$$X(t) = \sin(2\omega t) + .4\sin(3\omega t) + \epsilon(t), \quad \epsilon(t) \sim \mathcal{N}(0,1).$$

Stochastic process

- Formal definition:
  Given a probability space $(\Omega, \mathcal{F}, P)$ and a measurable space (S, Σ), an S-valued stochastic process is a collection of S-valued random variables on Ω, indexed by a totally ordered set T ("time"). That is, a stochastic process X is a collection

$$\{X_t : t \in T\}$$

  where each X_t is an S-valued random variable on Ω. The space S is then called the state space of the process.

- Finite-dimensional distributions:
  Let X be an S-valued stochastic process. For every finite sequence $T = (t_1, \ldots, t_n) \in T^n$, the n-tuple $\mathbf{X}_T = (X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ is a random variable taking values in S^n. The distribution $\mathbb{P}_T(\cdot) = \mathbb{P}(\mathbf{X}_T^{-1}(\cdot))$ of this random variable is a probability measure on S^n. This is called a finite-dimensional distribution of X.

Second order stochastic process

When the stochastic process $\mathbf{X}_T = (X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ can be characterized by its first and second order statistics, we have:

- First order moments (expected values):

$$\mathrm{E}\{\mathbf{X}_T\} = (\bar{X}_{t_1} = \mathrm{E}\{X_{t_1}\}, \bar{X}_{t_2} = \mathrm{E}\{X_{t_2}\}, \cdots, \bar{X}_{t_n} = \mathrm{E}\{X_{t_n}\})$$

- Second order moments (Variances):

$$\mathrm{Var}\{\mathbf{X}_T\} = (\mathrm{Var}\{X_{t_1}\}, \mathrm{Var}\{X_{t_2}\}, \cdots, \mathrm{Var}\{X_{t_n}\})$$

- Covariance matrix:

$$[\mathbf{V}(\mathbf{X}_T)]_{m,n} = \left[\mathrm{E}\left\{(X_{t_m} - \bar{X}_{t_m})(X_{t_n} - \bar{X}_{t_n})\right\}\right]$$

- Correlation matrix:

$$[\mathbf{C}(\mathbf{X}_T)]_{m,n} = [\mathrm{E}\{(X_{t_m} X_{t_n}\}]$$

Second order random signals

For continuous case, we define a random function $X(t)$ and so:

- First order moments (expected values):

$$\bar{X}(t) = \mathrm{E}\{X(t)\}$$

- Second order moments (Variances):

$$\mathrm{Var}\{X(t)\} = \mathrm{E}\left\{(X(t) - \bar{X}(t))^2\right\}$$

- (auto)-correlation function:

$$R_{XX}(t, \tau) = \mathrm{E}\{(X(t)\, X(t+\tau)\}$$

- (inter)-correlation function:

$$R_{XY}(t, \tau) = \mathrm{E}\{(X(t)\, Y(t+\tau)\}$$

Second order stationary random signals

- Strict sense stationary: A random signal $X(t)$ is said to be stationary if the expression of its probability distribution does not depend on time t.
- Wide sense stationary: A random signal $X(t)$ is said to be stationary if the expression of its probability distribution depend only on the two first moments and that these moments do not depend on time t.
- First order moments (expected values):

$$\bar{X}(t) = \mu, \forall t, \quad (\text{Centered signal:} \mu = 0)$$

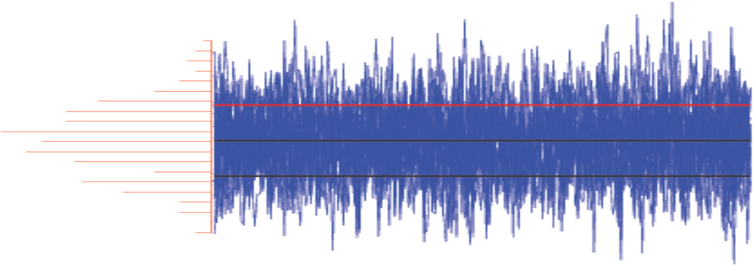
- (auto)-correlation function:

$$R_{XX}(\tau) = \mathrm{E}\left\{X(t)X(t+\tau)\right\}$$

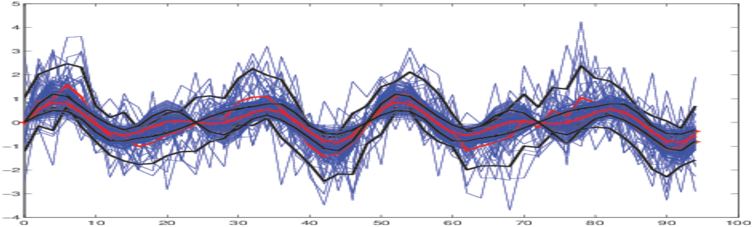
- power spectral density function

$$S_{XX}(\omega) = \int R_{XX}(\tau) \exp\left\{-j\omega\tau\right\} \,\mathrm{d}\tau$$

Stationary/Non Stationary



$$X(t) \sim \mathcal{N}(0,1), \quad \forall t \longrightarrow \mathrm{E}\left\{X(t)\right\} = 0, \mathrm{Var}\left\{X(t)\right\} = 1, \forall t$$



$$X(t) = a_1 \sin(2\omega t) + a_2 \sin(3\omega t),\ \omega = \frac{\pi}{12},\ a_1, a_2 \sim \mathcal{U}(0,1)$$

Second order stationary random signals

For centred random functions $X(t)$ and $Y(t)$ we have:

- First order moments (expected values):

$$\bar{X}(t) = \bar{Y}(t) = 0, \forall t$$

- (auto)-correlation function:

$$R_{XX}(\tau) = \mathrm{E}\left\{X(t)X(t+\tau)\right\}$$

- power (auto)-spectral density function

$$S_{XX}(\omega) = \int R_{XX}(\tau) \exp\left\{-j\omega\tau\right\} \, \mathrm{d}\tau$$

- (inter)-correlation function:

$$R_{XY}(\tau) = \mathrm{E}\left\{X(t)Y(t+\tau)\right\}$$

- power (inter)-spectral density function:

$$S_{XY}(\omega) = \int R_{XY}(\tau) \exp\left\{-j\omega\tau\right\} \, \mathrm{d}\tau$$

Second order stationary discrete time random signals

Replace $X(t)$ with $X(n)$ assuming sampling interval is equal to unity:

- First order moments (expected values):

$$\bar{X}(n) = \mathrm{E}\left\{X(n)\right\} = \mu, \quad \forall n$$

- Second order moments (Variances):

$$\mathrm{Var}\left\{X(n)\right\} = \mathrm{E}\left\{(X(n) - \bar{X}(n))^2\right\} = \sigma^2, \quad \forall n$$

- autocorrelation function:

$$r_{XX}(k) = \mathrm{E}\left\{(X(n)\, X(n+k)\right\}, \quad \forall n$$

- power spectral density function

$$S_{XX}(\omega) = \sum_k r_{XX}(k) \exp\left\{-jk\omega\right\}$$

Second order stationary discrete time random signals

For centred random discrete time signals $X(n)$ and $Y(n)$ we have:

- First order moments (expected values):

$$\bar{X}(n) = \bar{Y}(n) = 0, \quad \forall n$$

- (auto)-correlation function:

$$r_{XX}(k) = \mathrm{E}\left\{X(n)X(n+k)\right\}$$

- power (auto)-spectral density function

$$S_{XX}(\omega) = \sum_{k} r_{XX}(k) \exp\left\{-jk\omega\right\}$$

- (inter)-correlation function:

$$R_{XY}(k) = \mathrm{E}\left\{X(n)Y(n+k)\right\}$$

- power (inter)-spectral density function:

$$S_{XY}(\omega) = \sum_{k} R_{XY}(k) \exp\left\{-jk\omega\right\}$$

Exercises

Compute the autocorrelation function $R(\tau)$ and the power spectral density function $S(\omega)$ for the following signals

1. $U(t)$ is a strictly stationary white Gaussian random signal with zero mean and variance one: $p(U(t)) = \mathcal{N}(0, 1), \forall t$

2. $X(t)$ is obtained by $X(t) = \sum_{k=0}^{K} h(k)U(t-k)$ where $U(t)$ is a strictly stationary white Gaussian random signal with zero mean and variance one: $p(U(t)) = \mathcal{N}(0, 1), \forall t$. Take first $K = 1$ and then extend.

3. $X(t)$ is obtained by $X(t) = aX(t-1) + U(t)$ where $U(t)$ is a strictly stationary white Gaussian random function with zero mean and variance one: $p(U(t)) = \mathcal{N}(0, 1), \forall t$

Answers

Case 1: white strict stationary Gaussian process:

$U(t)$ is a strictly stationary white Gaussian random signal with zero mean and variance one: $p(U(t)) = \mathcal{N}(0,1), \forall t$

$$R(\tau) = \mathrm{E}\left\{U(t)U(t+\tau)\right\} = \begin{cases} 1 & \tau = 0 \\ 0 & else \end{cases}$$

$$R(\tau) = \delta(\tau) \longrightarrow S(\omega) = 1, \forall \omega$$

Matlab:

$N = 200; t = [0:N-1]; u = randn(N,1); figure(1), plot(t,x)$

Answers

Case 2 MA Gaussian Process:

$X(t)$ is obtained by $X(t) = \sum_{k=0}^{K} h(k)U(t-k)$ where $U(t)$ is a stationary Gaussian random function with zero mean and variance one: $p(U(t)) = \mathcal{N}(0,1), \forall t$. For numerical computation take $h(k) = \exp\{-\gamma k\}$ with $\gamma = .1$ and $K = 7$.
First take $K = 1$:

$$
\begin{aligned}
R(\tau) =& \mathrm{E}\left\{X(t)X(t+\tau)\right\} \\
=& \mathrm{E}\left\{[h(0)U(t) + h(1)U(t-1)][h(0)U(t+\tau) + h(1)U(t+\tau-1)\right\} \\
=& \mathrm{E}\left\{[h^2(0)U(t)U(t+\tau)\right\} \\
&+ \mathrm{E}\left\{h(0)h(1)U(t)U(t+\tau-1)\right\} \\
&+ \mathrm{E}\left\{h(1)h(0)U(t-1)U(t+\tau)\right\} \\
&+ \mathrm{E}\left\{h^2(1)U(t-1)U(t+\tau-1)\right\}
\end{aligned}
$$

$$
\begin{aligned}
\tau = 0 &: R(0) = h^2(0) + 0 + 0 + h^2(1) + 0 \\
\tau = 1 &: R(1) = 0 + h(0)h(1) + 0 \\
\tau > 1 &: R(\tau) = 0 + 0 + 0 + 0
\end{aligned}
$$

Answers

Case 2 MA Gaussian Process:

$$
\begin{aligned}
R(\tau) &= \mathrm{E}\left\{X(t)X(t+\tau)\right\} \\
&= \mathrm{E}\left\{[\sum_{k=0}^{K} h(k)U(t-k)][\sum_{k'=0}^{K} h(k')U(t-k'+\tau)]\right\} \\
&= \mathrm{E}\left\{[\sum_{k=0}^{K}\sum_{k'=0}^{K} h(k)h(k')U(t-k)U(t-k'+\tau)]\right\} \\
&= \sum_{k=0}^{K}\sum_{k'=0}^{K} h(k)h(k')\mathrm{E}\left\{U(t-k)U(t-k'+\tau)\right\} \\
&= \sum_{k=0}^{K}\sum_{k'=0}^{K} h(k)h(k'), \quad \text{if } k'-\tau=k, \quad 0, \text{ else}
\end{aligned}
$$

$$
\begin{aligned}
&\tau = 0 : R(0) = \sum_{k=0}^{K} h^2(k) \\
&\tau = 1 : R(1) = \sum_{k=0}^{K} h(k)h(k+1) \\
&\tau = 2 : R(2) = \sum_{k=0}^{K} h(k)h(k+2)
\end{aligned}
$$

...

$$
R(\tau) = \sum_{k=0}^{K} h(k)h(k+\tau) \longrightarrow S(\omega) = \left|\sum_{k=0}^{K} h(k)\exp\left\{-jk\omega\right\}\right|^2
$$

Answers

Case 3 First order AR Gaussian Process:

$X(t)$ is obtained by $X(t) = aX(t-1) + U(t)$ where $U(t)$ is a stationary Gaussian random function with zero mean and variance one: $p(U(t)) = \mathcal{N}(0,1), \forall t$.

$$
\begin{aligned}
R(\tau) &= \mathrm{E}\left\{X(t)X(t+\tau)\right\} \\
&= \mathrm{E}\left\{(aX(t-1)+U(t))(aX(t+\tau-1)+U(t+\tau))\right\}
\end{aligned}
$$

$$
\begin{aligned}
R(0) &= \mathrm{E}\left\{X(t)X(t)\right\} \\
&= \mathrm{E}\left\{(aX(t-1)+U(t))(aX(t-1)+U(t))\right\} \\
&= a^2\mathrm{E}\left\{X(t-1)X(t-1)\right\} + \sigma^2 \\
&= a^2 R(0) + \sigma^2 \longrightarrow R(0)(1-a^2) = \sigma^2 \longrightarrow R(0) = \frac{\sigma^2}{1-a^2} \\
R(1) &= aR(0) \\
R(2) &= aR(1) \\
\ldots & \\
R(\tau) &= \frac{\sigma^2}{1-a^2}(a)^{|\tau|} \longrightarrow S(\omega) = \frac{1}{\sqrt{2\pi}} \frac{\sigma^2}{1-a^2} \frac{\gamma}{\pi(\gamma^2+\omega^2)}
\end{aligned}
$$

Correlation matrix and its properties

For a centred wide sense stationary random discrete time signals $X(n)$:

- First order moments (expected values): $\bar{X}(n) = 0, \quad \forall n$
- (auto)-correlation function: $r_{XX}(k) = \mathrm{E}\left\{X(n)X(n+k)\right\}$
- If we define a $M \times \ 1$ vector
  $\mathbf{x}(n) = [X(n), X(n-1), \cdots, X(n-M+1)]'$,
  then the $M \times \ M$ correlation matrix $\mathbf{R}$ is defined by

$$\mathbf{R} = \mathrm{E}\left\{\mathbf{x}(n)\mathbf{x}^H(n)\right\}$$

  where the superscript H denotes Hermitian transposition.
  This matrix has the form:

$$\mathbf{R} = \begin{bmatrix} r(0) & r(1) & \cdots & r(-M+1) \\ r(-1) & r(0) & \cdots & r(-M+2) \\ \vdots & \vdots & \ddots & \vdots \\ r(M-1) & r(M-2) & \vdots & r(0) \end{bmatrix}$$

Properties of the correlation matrix of a stationary discrete-time stochastic process

- $\mathbf{R}$ is Hermitian: $r(-k) = r^*(k) \longrightarrow \mathbf{R}^H = \mathbf{R}$
- $\mathbf{R}$ is Toeplitz: All the diagonal elements are the same.
- $\mathbf{R}$ is definite positive: For any arbitrary $M \times 1$ vector $\mathbf{a}$ we have $\mathbf{a}^H \mathbf{R} \mathbf{a} \geq 0$
- $\mathbf{R}$ is nonsingular: $\det(\mathbf{R}) \neq 0$. This is due to $|r(l)| < r(0), \forall l \neq 0$.
  This property is important for computational implication
  $$\mathbf{R}^{-1} = \frac{\operatorname{adj}(\mathbf{R})}{\det(\mathbf{R})}$$
- If we define a *Backward* $M \times 1$ vector $\mathbf{x}^B(n) = [X(n-M+1), X(n-M+2), \cdots, X(1)]'$ of $\mathbf{x}(n) = [X(n), X(n-1), \cdots, X(n-M+1)]'$, then
  $$\mathbf{R} = \mathrm{E}\left\{\mathbf{x}(n)\mathbf{x}^H(n)\right\} \text{ and } \mathbf{R} = \mathrm{E}\left\{\mathbf{x}^B(n)\mathbf{x}^{B^H}(n)\right\} = \mathbf{R}^T$$

Properties of the correlation matrix of a stationary discrete-time stochastic process

- If we define the *Backward* $(M+1)\times\ 1$ vectors
  $\mathbf{x}(n)=[X(n),X(n-1),\cdots,X(n-M+1),X(n-M+2)]'$
  and
  $\mathbf{x}^B(n)=[X(n-M+2),X(n-M+1),X(n-M+2),\cdots,X(1)]'$,
  then

$$\mathbf{R}_{M+1}=\left[\begin{array}{c:c} r(0) & \mathbf{r}^H \\ \cdots & \cdots \\ \mathbf{r} & \mathbf{R}_M \end{array}\right]=\left[\begin{array}{c:c} \mathbf{R}_M & \mathbf{r}^{B^*} \\ \cdots & \cdots \\ \mathbf{r}^{B^T} & r(0) \end{array}\right]$$

  where $\mathbf{r}=[r(1),r(2),\cdots,r(M)]'$
  and $\mathbf{r}^{B^T}=[r(-M),X(-M+1),\cdots,r(-1)]'$.

Exercises

1. Consider the vector $\mathbf{x}(n) = [X(n), X(n-1)]'$ where $X(n)$ is a stationary discrete time Gaussian process: $p(x(n)) = \mathcal{N}(0, 1)$. Write down the expression of correlation matrix $\mathbf{R}_2$.

2. Now consider the vector $\mathbf{x}^B(n) = [X(n-1), X(n)]'$. Write down the expression of correlation matrix $\mathbf{R}_2^B$.

3. Now, consider the vectors $\mathbf{x}(n) = [X(n), X(n-1), X(n-2)]'$ and $\mathbf{x}^B(n) = [X(n-2), X(n-1), X(n)]'$. Write down the expression of correlation matrix $\mathbf{R}_3$ and $\mathbf{R}_3^B$.

4. What relations exist between $\mathbf{R}_3$ and $\mathbf{R}_3^B$ and $\mathbf{R}_2$ and $\mathbf{R}_2^B$?

Answers

Consider the real valued vector $\mathbf{x}(n) = [X(n), X(n-1)]'$ where $X(n)$ is a real stationary discrete time Gaussian process: $p(x(n)) = \mathcal{N}(0, 1)$. Write down the expression of correlation matrix $\mathbf{R}_2$.

$$\mathbf{R}_2 = \mathbf{R}_2^t = \mathbf{R}_2 = \begin{bmatrix} r(0 & r(1) \\ r(1) & r(0) \end{bmatrix}$$

Now consider the vector $\mathbf{x}^B(n) = [X(n-1), X(n)]'$. Write down the expression of correlation matrix $\mathbf{R}_2^B$.

$$\mathbf{R}_2^B = \begin{bmatrix} r(0 & r(1) \\ r(1) & r(0) \end{bmatrix}$$

Answers

Now, consider the real vectors $\mathbf{x}(n) = [X(n), X(n-1), X(n-2)]'$ and $\mathbf{x}^B(n) = [X(n-2), X(n-1), X(n)]'$. Write down the expression of correlation matrix $\mathbf{R}_3$ and $\mathbf{R}_3^B$.

$$\mathbf{R}_3 = \mathbf{R}_3^{B'} = \begin{bmatrix} r(0) & r(1) & r(2) \\ r(1) & r(0) & r(2) \\ r(2) & r(1) & r(0) \end{bmatrix}$$

What relations exist between $\mathbf{R}_3$ and $\mathbf{R}_3^B$ and $\mathbf{R}_2$ and $\mathbf{R}_2^B$?

$$\mathbf{R}_3 = \mathbf{R}_3' = \begin{bmatrix} r(0) & | & \mathbf{r}' \\ \cdots & \cdots & \cdots \\ \mathbf{r} & | & \mathbf{R}_2 \end{bmatrix}$$

with $\mathbf{r}' = [r(1), r(2)]$

Exercises: Correlation of a sine wave plus noise

Consider $X(n) = a \exp\{j\omega n\} + \epsilon(n), n = 0, \cdots, N-1$ with $\epsilon(n) \sim \mathcal{N}(0, \sigma^2)$

1. Compute its autocorrelation function $r(k)$
2. Given a set of samples $\mathbf{x} = [x(n), x(n-1), \cdots, x(n-M+1)]'$, write down its correlation matrix $\mathbf{R}_M$
3. Can we determine a and ω from these samples?
4. If now, we consider two sine waves $X(n) = a_1 \exp\{j\omega_1 n\} + a_2 \exp\{j\omega_2 n\} + \epsilon(n), n = 0, \cdots, N-1$. How can we determine a_1, a_2 and ω_1 and ω_2?
5. Extend this result to the general case of K sine waves $X(n) = \sum_{k=1}^{K} a_k \exp\{j\omega_k n\} + \epsilon(n), n = 0, \cdots, N-1$.

Answers: Correlation of a sine wave plus noise

$$x(n) = a \exp\{j\omega n\} + \epsilon(n)$$

$$\begin{aligned} r(k) &= \mathrm{E}\{x(n)x^*(n+k)\} \\ &= \mathrm{E}\{[a\exp\{j\omega n\} + \epsilon(n)][a\exp\{j\omega(n+k)\} + \epsilon(n+k)]^*\} \\ &= \mathrm{E}\{[a\exp\{j\omega n\} + \epsilon(n)][a^*\exp\{-j\omega(n+k)\} + \epsilon^*(n+k)]\} \end{aligned}$$

$$r(k) = \begin{cases} |a|^2 + \sigma^2 = |a|^2(1+\frac{1}{\rho}) & k = 0, \quad \rho = \frac{|a|^2}{\sigma^2}, \\ |a|^2 \exp\{j\omega k\} & k \neq 0 \end{cases}$$

When having M samples, we can make the correlation matrix:

$$\mathbf{R} = |a|^2 \begin{bmatrix} 1+\frac{1}{\rho} & \exp\{j\omega\} & \cdots & \exp\{j\omega(M-1)\} \\ \exp\{j\omega\} & 1+\frac{1}{\rho} & \cdots & \exp\{j\omega(M-2)\} \\ \vdots & \vdots & \ddots & \vdots \\ \exp\{j\omega(M-1)\} & \exp\{j\omega(M-2)\} & \cdots & 1+\frac{1}{\rho} \end{bmatrix}$$

Stochastic models

$$\epsilon(n) \longrightarrow \boxed{\begin{array}{c}\text{Discrete-time}\\ \text{linear filter}\end{array}} \longrightarrow x(n)$$

with $\epsilon(n)$ purely random (stationary, white process):

$$\mathrm{E}\left\{\epsilon(n)\right\} = 0, \; \mathrm{E}\left\{\epsilon(n)\epsilon^*(n+k)\right\} = \begin{cases} \sigma^2 & \text{if } k = 0 \\ 0 & \text{else} \end{cases} .$$

$x(n)$ can be:

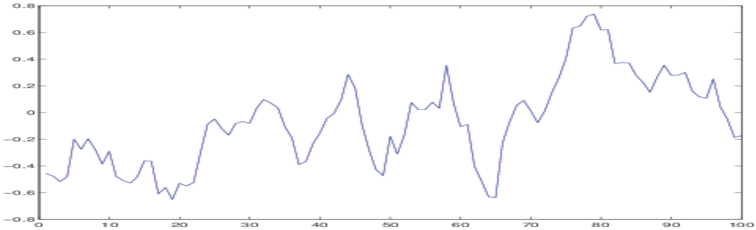
- a combination of past values of $u(n)$
  (Moving Average (MA) model)
- a combination of past values of $x(n)$ and present value of $u(n)$
  (Autoregressive (AR) model)
- a combination of past values of $x(n)$ and past and present values of $u(n)$
  (Autoregressive Moving Average (ARMA) model)

Stochastic models: Moving Average (MA)

$$x(n) = \sum_{k=0}^{q} b(k)\epsilon(n-k), \quad \forall n$$

$$\epsilon(n) \longrightarrow \boxed{B(z) = \sum_{k=0}^{q} b(k) z^{-k}} \longrightarrow x(n)$$

$$S_{xx}(\omega) = \left| \sum_{k=0}^{q} b(k) \exp\{-jk\omega\} \right|^2$$



MA, $q = 15$, $v = 1$, $b_k = \exp\{-.05 * k^2\}$

Stochastic models: AutoRegressive (AR)

$$x(n) = \sum_{k=1}^{p} a(k)\, x(n-k) + \epsilon(n), \quad \forall n$$

$$\sum_{k=0}^{p} b(k)\, x(n-k) = \epsilon(n), \text{ with } b(0) = 1, b(k) = -a(k)$$

$$x(n) \longrightarrow \boxed{B(z) = \sum_{k=0}^{p} b(k)\, z^{-k}} \longrightarrow \epsilon(n)$$

$$\epsilon(n) \longrightarrow \boxed{H(z) = \frac{1}{1 - \sum_{k=1}^{p} a(k)\, z^{-k}}} \longrightarrow x(n)$$

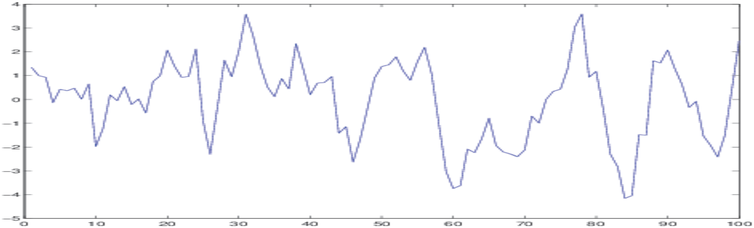
$$S_{xx}(\omega) = \frac{1}{\left|1 - \sum_{k=1}^{p} a(k) \exp\{-jk\omega\}\right|^2}$$

Stochastic models: AutoRegressive (AR)

$$x(n) = \sum_{k=1}^{p} a(k)\, x(n-k) + \epsilon(n), \quad \forall n$$

$$\epsilon(n) \longrightarrow \boxed{H(z) = \frac{1}{1 - \sum_{k=1}^{p} a(k)\, z^{-k}}} \longrightarrow x(n)$$

$$S_{xx}(\omega) = \frac{\sigma^2}{\left|1 - \sum_{k=1}^{p} a(k) \exp\{-jk\omega\}\right|^2}$$



AR1, $a = .7$, $v = 1$.

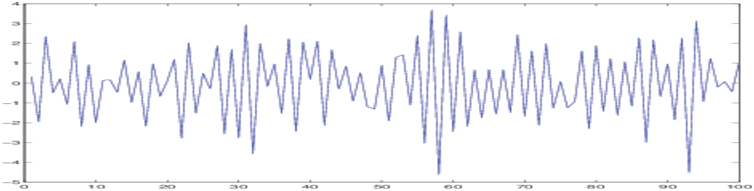
Stochastic models: ARMA

$$x(n)=\sum_{k=1}^{p}a(k)\,x(n-k)+\sum_{l=0}^{q}b(l)\,\epsilon(n-l)$$

$$\epsilon(n)\longrightarrow\boxed{H(z)=\frac{B(z)}{A(z)}=\frac{\sum_{k=0}^{q}b(k)z^{-k}}{1-\sum_{k=1}^{p}a(k)\,z^{-k}}}\longrightarrow x(n)$$

$$\epsilon(n)\longrightarrow\boxed{B_q(z)}\longrightarrow\boxed{\frac{1}{A_p(z)}}\longrightarrow x(n)$$

$$S_{xx}(\omega)=\frac{\left|\sum_{k=0}^{q}b(k)\exp\{-jk\omega\}\right|^2}{\left|1-\sum_{k=1}^{p}a(k)\exp\{-jk\omega\}\right|^2}$$



ARMA(1,2), $q=1,\ v=1,\ b_1=1,\ a_1=1, a_2=.8$

Autocorrelation of a Stationary AR process

$$x(n) = \sum_{k=1}^{p} a_k\, x(n-k) + \epsilon(n) \longrightarrow \sum_{k=0}^{p} b_k\, x(n-k) = \epsilon(n)$$

with $b_0 = 1,\ b_k = -a_k, k = 1, \cdots, p$.

$$\sum_{k=0}^{p} b_k\, x(n-k)x(n-l) = \epsilon(n)x(n-l), l > 0$$

$$\mathrm{E}\left\{\sum_{k=0}^{p} b_k\, x(n-k)x(n-l)\right\} = \mathrm{E}\left\{\epsilon(n)x(n-l)\right\}, l > 0$$

$$\sum_{k=0}^{p} b_k \mathrm{E}\left\{x(n-k)x(n-l)\right\} = 0 \longrightarrow \sum_{k=0}^{p} b_k\, r(l-k) = 0,\ l > 0$$

Autocorrelation of a Stationary AR process

$$\sum_{k=0}^{p} b_k r(l-k) = 0,\ l > 0 \longrightarrow r(l) = \sum_{k=1}^{p} a_k r(l-k),\ l > 0$$

- This difference equation has a general form solution:

$$r(l) = \sum_{k=1}^{p} c_k p_k^l$$

  with c_k are constants and p_k the roots of the Characteristic function

$$\sum_{k=0}^{p} b_k z^{-k} = 0 \quad \text{or} \quad 1 - \sum_{k=1}^{p} a_k z^{-k} = 0$$

- Asymptotic stationarity condition: $|p_k| < 1$
- All the poles of the AR filter lie inside of the unit circle in the z-plane.

Wold decomposition

Wold (1938): Any stationary discrete-time stochastic process $x(n)$ may be decomposed into the sum of a general linear process and a predictable process, with these two process being uncorrelated.

$$x(n) = u(n) + s(n)$$

where

- $u(n)$ and $s(n)$ are uncorrelated;
- $u(n)$ is a general linear MA process:

$$u(n) = \sum_{k=0}^{\infty} b_k^* \epsilon(n-k)$$

  with $b_0 = 1$, $\sum_{k=0}^{\infty} |b_k|^2 < \infty$ and $\mathrm{E}\{\epsilon(n) s^*(k)\} = 0, \forall (n,k)$.
- $s(n)$ is a predictable process, i.e. it can be predicted from its own pqst with zero prediction error.

Parameter estimation: Yule-Walker equations

$$\sum_{k=0}^{p} b_k r(l-k) = 0, l > 0 \text{ with } b_0 = 1 \longrightarrow r(l) = \sum_{k=1}^{p} a_k r(l-k)$$

$$r(l) = \sum_{k=1}^{p} a_k r(l-k), \quad l = 1, 2, \cdots, p \longrightarrow$$

$$\begin{bmatrix} r(1) = & r(0)a_1 & +r(-1)a_2 & + \cdots & +r(-p+1)a_p \\ r(2) = & r(1)a_1 & +r(0)a_2 & + \cdots & +r(-p+2)a_p \\ \vdots & & & & \\ r(p) = & r(p)a_1 & +r(p-1)a_2 & + \cdots & +r(0)a_p \end{bmatrix}$$

$$\begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(p) \end{bmatrix} = \begin{bmatrix} r(0) & r(-1) & \cdots & r(-p+1) \\ r(1) & r(0) & \cdots & r(-p+2) \\ \vdots & & & \\ r(p) & r(p-1) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

$$\mathbf{r} = \mathbf{R}\mathbf{a} \longrightarrow \mathbf{R}\mathbf{a} = \mathbf{r}$$

Parameter estimation: Yule-Walker equations

$$\mathbf{Ra} = \mathbf{r} \longrightarrow \mathbf{a} = \mathbf{R}^{-1}\mathbf{r}$$

If $\mathbf{R}$ is non singular, we have a unique relationship between

$$\mathbf{a} = \{a_1, a_2, \cdots, a_p\}$$

and the normalized correlation coefficients

$$\boldsymbol{\rho} = \{\rho_1, \rho_2, \cdots, \rho_p\}$$

with $\rho_k = r(k)/r(0)$.
Conclusion: Given $r(0), r(1), \cdots, r(p)$, we can compute $\{\rho_1, \rho_2, \cdots, \rho_p\}$, then compute $\{a_1, a_2, \cdots, a_p\}$ and also the variance of the noise

$$\sigma^2 = \sum_{k=0}^{p} a_k r(k)$$

with $a_0 = 0$.

Exercise

1. Consider a first order AR model $x(n) = a_1 x(n-1) + \epsilon(n)$ with $\epsilon(n) \sim \mathcal{N}(0, \sigma^2)$. First compute $r(0)$ and $r(1)$. Then construct the YW equation and find the solution for a_1 and σ^2.

2. Consider now a second order AR model $x(n) = a_1 x(n-1) + a_2 x(n-2) + \epsilon(n)$ with $\epsilon(n) \sim \mathcal{N}(0, \sigma^2)$. First compute $r(0)$, $r(1)$ and $r(2)$. Then construct the YW equation and find the solution for a_1, a_2 and σ^2.

Answers

First order AR model $x(n) = a_1 x(n-1) + \epsilon(n)$

$$\begin{aligned}
r(0) &= \mathrm{E}\{x(n)x(n)\} \\
&= \mathrm{E}\{(a_1 x(n-1) + \epsilon(n))(a_1 x(n-1) + \epsilon(n))\} \\
&= a_1^2 r(0) + \sigma^2 \\
r(1) &= \mathrm{E}\{x(n)x(n+1)\} \\
&= \mathrm{E}\{x(n)(a_1 x(n) + \epsilon(n+1))\} \\
&= a_1 r(0)
\end{aligned}$$

Yule-Walker:

$$\left[r(0)\right]\left[a_1\right] = \left[r(1)\right] \longrightarrow a_1 = \frac{r(1)}{r(0)}$$

$$r(0) = a_1^2 r(0) + \sigma^2 \longrightarrow \sigma^2 = (1 - a_1^2) r(0) = r(0) - a_1 r(1)$$

Numerical example: $r(0) = 1, r(1) = .9$:
$a_1 = .9, \sigma^2 = (1 - .9^2) = 1 - .81 = .19$

Answers

Second order $x(n) = a_1 x(n-1) + a_2 x(n-2) + \epsilon(n)$

$$\begin{aligned} r(0) &= \mathrm{E}\left\{x(n)x(n)\right\} \\ &= \mathrm{E}\left\{x(n)(a_1 x(n-1) + a_2 x(n-2) + \epsilon(n))\right\} \\ &= a_1 r(-1) + a_2 r(-2) \\ r(1) &= \mathrm{E}\left\{x(n)x(n+1)\right\} \\ &= \mathrm{E}\left\{x(n)(a_1 x(n) + a_2 x(n-1) + \epsilon(n+1))\right\} \\ &= a_1 r(0) + a_2 r(-1) \\ r(2) &= \mathrm{E}\left\{x(n)x(n+2)\right\} \\ &= \mathrm{E}\left\{x(n)(a_2 x(n) + a_1 x(n+1) + \epsilon(n+2))\right\} \\ &= a_2 r(0) + a_1 r(1) \end{aligned}$$

$$\begin{bmatrix} r(0) & r(1) \\ r(1) & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} r(1) \\ r(2) \end{bmatrix}$$

When $r(0), r(1)$ and $r(2)$ computed, we can compute

$$\sigma^2 = \sum_{k=0}^{2} a_k r(k) = r(0) + a_1 r(1) + a_2 r(2)$$

Answers: Recursive computation

$$r(l) = \sum_{k=1}^{p} a_k r(l-k), \quad r(0) = \sum_{k=1}^{p} a_k r(-k) + \sigma^2$$

By dividing both sides of the first one by $r(0)$ we obtain:

$$\rho_l = \sum_{k=1}^{p} a_k \rho_{(l-k)}$$

$p = 1$:

$$r(1) = a_1 r(0) \longrightarrow a_1 = \rho_1 \qquad \rho_1 = a_1$$

$p = 2$:

$$r(1) = a_1 r(0) + a_2 r(-1) \qquad \rho_1 = \frac{a_1}{1-a_2}$$

$$r(2) = a_2 r(0) + a_1 r(0)$$

Using recursion:

$$\rho_2 = a_1 \rho_1 + a_2 \rho_0 = a_1 \frac{a_1}{1-a_2} + a_2$$

$$\rho_2 = \frac{a_1^2 - a_2^2 + a_2}{1-a_2}$$

...

Power spectral density

Given a discrete time stationary process $u(n)$ and a symmetric window observation
$\mathbf{u}_N = \{u(-N), \cdots, u(-1), u(0), u(1), \cdots, u(N)\}$
and defining its DFT:

$$U_N(\omega) = \sum_{n=-N}^{N} u(n) \exp\{-j\omega n\}$$

and

$$\begin{aligned}
\mathrm{E}\left\{|U_N(\omega)|^2\right\} &= \sum_{n=-N}^{N} \sum_{m=-N}^{N} \mathrm{E}\{u(n)u^*(m)\} \exp\{-j\omega(n-m)\} \\
&= \sum_{n=-N}^{N} \sum_{m=-N}^{N} r(m-n) \exp\{-j\omega(n-m)\}
\end{aligned}$$

it can be shown that

$$\lim_{N\mapsto\infty} \frac{1}{N} \mathrm{E}\left\{|U_N(\omega)|^2\right\} \mapsto \sum_{k=-\infty}^{\infty} r(k) \exp\{-j\omega k\} \stackrel{\triangle}{=} S(\omega)$$

Power spectral density

Direct definition:

$$S(\omega) = \lim_{N \mapsto \infty} \frac{1}{N} \mathrm{E}\left\{ |U_N(\omega)|^2 \right\}$$

Definition through autocorrelation coefficients:

$$r(k) = \mathrm{E}\left\{ u(n) u^*(n+k) \right\}$$

$$S(\omega) = \sum_{k=-\infty}^{\infty} r(k) \exp\left\{ -j\omega m \right\}, \quad -\pi \leq \omega \leq \pi$$

$$r(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) \exp\left\{ -j\omega m \right\} \, \mathrm{d}\omega, \quad m = 0, \pm 1, \pm 2, \cdots$$

Properties:

- $S(\omega)$ is periodic
- $S(\omega)$ for a stationary discrete-time process is real.
- $S(\omega)$ for a real stationary discrete-time process is symmetric.
- $r(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) \, \mathrm{d}\omega$
- $S(\omega) \geq 0, \forall \omega$.

Power spectral density through a linear filter

$$x(n) \longrightarrow \boxed{\begin{array}{c}\text{Linear}\\ \text{Filter}\end{array}} \longrightarrow y(n) = h(n) * x(n)$$

Deterministic signals: Ordinary DFT:

$$y(n) = h(n) * x(n) = \sum_k h(k)x(n-k) \longrightarrow Y(\omega) = H(\omega)X(\omega)$$

Stochastic signals:

$$\begin{array}{c} r_{XX}(k) \\ S_{YY}(\omega) \end{array} \longrightarrow \boxed{\begin{array}{c}\text{Linear}\\ \text{Filter}\end{array}} \longrightarrow \begin{array}{c} r_{YY}(k) \\ S_{YY}(\omega) \end{array}$$

$$r_{YY}(m) = \sum_l \sum_k h(l)h^*(k)r_{XX}(k-l+m)$$

$$S_{YY}(\omega) = |H(\omega)|^2 S_{XX}(\omega) \qquad S_{YX}(\omega) = H^*(\omega)S_{XX}(\omega)$$

Power spectral estimation

Given a set of samples from a stochastic process $x(n)$, estimate its power spectral density function (called also power spectrum) $S_{XX}(\omega)$

Periodogram-based (Direct computation)

$$S(\omega) = \lim_{N \mapsto \infty} \frac{1}{N} \mathrm{E}\left\{|U_N(\omega)|^2\right\}$$

- Take a very large window of the data, compute its DFT, look at the amplitude power 2 as the power spectrum.
- If a great number of samples are available, cut them in M blocs of each N samples. For each bloc compute $|U_N(\omega)|^2$ and then average them to obtain the power spectrum.

Power spectral estimation

Use the autocorrelation coefficients $r(k)$:

$$S(\omega) = \sum_{k=-\infty}^{\infty} r(k) \exp\{-j\omega m\}, \quad -\pi \le \omega \le \pi$$

- Try first to estimate $r(k)$ for $k = 0, 1, \cdots, K$ with K as great as possible, then use this approximation

$$S(\omega) = \sum_{k=-K}^{K} r(k) \exp\{-j\omega m\}, \quad -\pi \le \omega \le \pi$$

which is good if $r(k) = 0, k > K$.

Power spectral estimation

Model based (AR process)

Choose a model order p, estimate $r(k), k = 0, 1, \cdots, p$, deduce the parameters of the model $\{a_1, a_2, \cdots, a_p\}$ and the noise variance σ^2 using Yule-Walker relation and then compute

$$S_{xx}(\omega) = \frac{\sigma^2}{\left|1 - \sum_{k=1}^{p} a(k) \exp\{-jk\omega\}\right|^2}$$

Examples:

- AR0:

$$S_{xx}(\omega) = \sigma^2$$

- AR1:

$$S_{xx}(\omega) = \frac{\sigma^2}{\left|1 - a_1 \exp\{-j\omega\}\right|^2} = \frac{\sigma^2}{1 + a_1^2 - 2a_1 \cos(\omega)}$$

Power spectral estimation

AR2:

$$S_{xx}(\omega) = \frac{\sigma^2}{\left|1 - a_1 \exp\{-j\omega\} - a_2 \exp\{-j2\omega\}\right|^2}$$

- $a_1 > 0$ Low pass filter around 0
- $a_1 < 0$ High pass filter around $\omega = \pi$
- $-1 < a_2 < 1 - |a_1|$ the process is stable

$$S_{xx}(f) = \frac{\sigma^2}{1 + a_1^2 + a_2^2 - 2a_1(1 - a_2)\cos(2\pi f) - 2a_2\cos(4\pi f)}$$

Power spectral estimation

Model based (AR process) with direct estimation of the model parameters

Estimate directly the parameters of the model from the data using a Least Square (LS) criterion

$$LS(\mathbf{a}) = \frac{1}{2} \sum_n \left| x(n) - \sum_{k=1}^{p} a_k \, x(n-k) \right|^2$$

(or any other criteria as we will see later) and then use it.

$$\frac{\partial LS(\mathbf{a})}{\partial a_k} = -\sum_n x(n-k) \left(x(n) - \sum_{k=1}^{p} a(k) \, x(n-k) \right) = 0$$

$$\sum_n x(n-k) x(n) = \sum_n x(n-k) \sum_{k=1}^{p} a(k) \, x(n-k)$$

Solve these equations either simultaneously or recursively.
When **a** obtained use the theoretical expressions.

Linear prediction and AR modelling

- $\{x(1), \cdots, x(n-1)\}$ observed samples of a signal. Predict $x(n)$
- Prediction or innovation Erreur: $\epsilon_n = x(n) - \widehat{x}(n)$
- The linear predictor:

$$\widehat{x}(n) = \sum_{k=1}^{p} a(k)\, x(n-k), \quad \forall n$$

- Mean Square Errors (MSE): $\mathrm{MSE} = \sum_n |\epsilon_n|^2 = \sum_n |x(n) - \widehat{x}(n)|^2$
- Least Mean Squares (LMS) Error

$$\widehat{x}(n) = \arg\min_{x(n)} \{\mathrm{MSE}\}$$

$$\mathrm{MSE} = \sum_n |x(n) - \widehat{x}(n)|^2 = \sum_n \left| x(n) - \sum_{k=1}^{p} a(k)\, x(n-k) \right|^2$$

Minimum Variance Estimation

- $x(n)$, An n sample of a signal
- AR model:

$$\widehat{x}(n) = \sum_k a(k)\, x(n-k)$$

- modelling errror

$$\epsilon_n = x(n) - \widehat{x}(n)$$

- Criterium

$$\beta^2 = \min \mathrm{E}\left\{|\epsilon_n|^2\right\} = \min \mathrm{E}\left\{[x(n) - \widehat{x}(n)]^2\right\}$$

- Orthogonality Condition

$$\mathrm{E}\left\{[x(n) - \sum_{k'} a(k')\, x(n-k')]\, x(n-k)\right\} = \beta^2\, \delta(k),\ k = 1, \ldots, p$$

$$r(k) - \sum_{k'} a(k')\, r(k'-k) = \beta^2\, \delta(k)$$

Minimum Variance Estimation

- Correlation matrix

$$\mathbf{R} = \begin{bmatrix} r(0) & r(1) & r(2) & \cdots & \cdots & r(p-1) \\ r(1) & \ddots & \ddots & \ddots & & \vdots \\ r(2) & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & r(2) \\ \vdots & & \ddots & \ddots & \ddots & r(1) \\ r(p-1) & \cdots & \cdots & r(2) & r(1) & r(0) \end{bmatrix}$$

$$\mathbf{r} = [r(1), \ldots r(p)]^t, \quad \mathbf{a} = [a(1), \ldots, a(p)]^t,$$

- Normal equations

$$\begin{aligned} \mathbf{Ra} &= \mathbf{r} \\ r(0) - \mathbf{a}^t\mathbf{r} &= \beta^2 \end{aligned}$$

Levinson algorithm

The reverse Levinson-Durbin recursion implements the step-down algorithm for solving the following symmetric Toeplitz system of linear equations for $\mathbf{r}$, where $\mathbf{r} = [r(0), \cdots, r(p)]'$.

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(p-1) \\ r(1) & r(0) & \cdots & r(p-2) \\ \vdots & \ddots & \ddots & \vdots \\ r(p-1) & r(p) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} \begin{bmatrix} r(0) \\ r(1) \\ \vdots \\ r(p) \end{bmatrix}$$

Great Exercise: Generation, Characterization, Modelling, Transmission and input Estimation by Wiener Filter

Source generator $\rightarrow$ Transmitter Characterization $\rightarrow$ Transmission Channel $\rightarrow$ Receiver Characterization

Step 1: Generation of a signal:

1. We want to generate a signal using a first order AR Gaussian process. $x(n) = ax(n-1) + u(n)$, where $x(0) \sim \mathcal{N}(0,1)$ and $u(n) \sim \mathcal{N}(0,1)$, $n = 1, \cdots, N$.
2. Take the numerical example: $N = 200$ and two different values $a = 0.1$ and $a = 0.9$ and call them X and Y. Plot these two signals.
3. plot $X(t)$, $Y(t)$. What do you remark?

Great Exercise: Generation, Characterization, Modelling, Transmission and input Estimation by Wiener Filter

Source generator $\rightarrow$ Transmitter Characterization $\rightarrow$ Transmission Channel $\rightarrow$ Receiver Characterization

Step 2: Characterization of generated signals:

1. Compute the theoretical expressions of the autocorrelation functions $r_{XX}(k)$ and $r_{YY}(k)$ and their corresponding power spectral density functions $S_{xx}(\omega)$ and $S_{yy}(\omega)$.
2. plot all these quantities and interpret them. What do you remark?

Great Exercise: Generation, Characterization, Modelling, Transmission and input Estimation by Wiener Filter

Step 3: Modelling

1. Given the generated signals $X(n)$ and $(Y(n)$, assuming that they can be modelled with a first order AR processes, estimate their corresponding parameters. Compare the parameters with those used to generate them. Give your conclusion.
2. Now, assume that these signals can be modelled wit fifth order MA processes: $X(n) = \sum_{k=0}^{5} b(k)u(n-k)$ where $u(n) \sim \mathcal{N}(0,1)$. Estimate then the corresponding parameters for X and for Y.
3. Compute now the correlation functions and the power spectral density functions. Compare with the original and the AR model ones.

Great Exercise: Generation, Characterization, Modelling, Transmission and input Estimation by Wiener Filter

Source generator → Transmitter Characterization → Transmission Channel → Receiver Characterization

Step 4: Transmission through a FIR channel

1. Now, we want to transmit $X(n)$ through a channel. We model the channel as a FIR model with $h(k) = \exp(-.1k), k = 0, \cdots, 5$. If we call $Z(n)$ the received signal, write the relation between X and Z.
2. Give the relations which existent between $X(n)$ and $Z(n)$, between $r_{ZZ}(k)$, $r_{XX}(k)$ and $R_{ZX}(k)$, between $R_{ZZ}(\omega)$, $R_{XX}(\omega)$ and $R_{ZX}(\omega)$.
3. Do the same with $Y(n)$.

Great Exercise: Generation, Characterization, Modelling, Transmission and input Estimation by Wiener Filter

Source generator	$\rightarrow$	Transmitter Characterization	$\rightarrow$	Transmission Channel	$\rightarrow$	Receiver Characterization

Step 5: Receiver (No noise channel)

1. We want to retrieve the transmitted signal X from the received signal Z. Is it possible?
2. First assume that the channel does not add any noise. Use the relations between the quantities $R_{ZZ}(\omega)$, $R_{XX}(\omega)$ and $R_{ZX}(\omega)$ to design an Inverse Filter to retrieve the original signal.

Great Exercise: Generation, Characterization, Modelling, Transmission and input Estimation by Wiener Filter

Source generator	$\rightarrow$	Transmitter Characterization	$\rightarrow$	Transmission Channel	$\rightarrow$	Receiver Characterization

Step 5: Receiver (Noisy channel)

1. Now assume that the channel adds an additive noise $\epsilon(n)$ with $p(\epsilon(n)) = \mathcal{N}(0, \sigma_\epsilon^2)$ with $\sigma_\epsilon^2 = 0.01$.
2. Design a Wiener filter to do this operation. Again write and use the relations between the quantities $R_{ZZ}(\omega)$, $R_{XX}(\omega)$ and $R_{ZX}(\omega)$ to design a Wiener Filter to retrieve the original signal.
3. Discuss the implementation issues.
4. Other possible solutions?
5. Recursive methods
6. Kalman filtering

Wiener Filtering

Objective: a signal $f(t)$ is transmitted through a media which is assumed to act as a linear and invariant system. The transmission canal is also noisy. We receive the noisy signal $g(t)$. We want to estimate the transmitted signal $\widehat{f}(t)$.

$$f(t) \longrightarrow \boxed{H(\omega)} \longrightarrow \overset{\epsilon(t)\,\downarrow}{\oplus} \longrightarrow g(t) \qquad\qquad g(t) \longrightarrow \boxed{\text{Wiener filter}} \rightarrow \widehat{f}(t)$$

$$g(t) = h(t) * f(t) + \epsilon(t)$$

$$\mathrm{E}\left\{\epsilon(t)\right\} = 0, \mathrm{E}\left\{f(t)\right\} = 0 \rightarrow \mathrm{E}\left\{g(t)\right\} = h(t) * \mathrm{E}\left\{f(t)\right\} + \mathrm{E}\left\{\epsilon(t)\right\} = 0$$

$$\begin{array}{l} R_{gg}(\tau) = \mathrm{E}\left\{g(t)\, g(t+\tau)\right\} \\ R_{ff}(\tau) = \mathrm{E}\left\{f(t)\, f(t+\tau)\right\} \\ R_{gf}(\tau) = R_{fg}(-\tau) = \mathrm{E}\left\{g(t)\, f(t+\tau)\right\} \end{array}$$

$\epsilon(t)$ is assumed to be centred and independent of $f(t)$.

$$\mathrm{E}\left\{\epsilon(t)\right\} = 0, \mathrm{E}\left\{\epsilon(t)\epsilon(t+\tau)\right\} = \begin{cases} \sigma_\epsilon^2 & \text{if } \tau = 0 \\ 0 & \text{else} \end{cases}$$

Wiener Filtering

$$f(t) \longrightarrow \boxed{H(\omega)} \longrightarrow \bigoplus \longrightarrow g(t) \qquad \text{with } \epsilon(t) \text{ added at } \bigoplus$$

$$g(t) \longrightarrow \boxed{\text{Wiener filter}} \longrightarrow \widehat{f}(t)$$

$$\begin{aligned} R_{gg}(\tau) &= h(t) * h(t) * R_{ff}(\tau) + R_{\epsilon\epsilon}(\tau) \\ R_{gf}(\tau) &= h(t) * R_{ff}(\tau) \end{aligned}$$

$$\begin{aligned} S_{gg}(\omega) &= |H(\omega)|^2 S_{ff}(\omega) + R_{\epsilon\epsilon}(\omega) \\ S_{gf}(\omega) &= H(\omega) S_{ff}(\omega) \\ S_{fg}(\omega) &= H^*(\omega) S_{ff}(\omega) \end{aligned}$$

$$g(t) \rightarrow \boxed{W(\omega)} \rightarrow \widehat{f}(t) \quad \text{or} \quad g(t) \rightarrow \boxed{w(t)} \rightarrow \widehat{f}(t)$$

$$\widehat{f}(t) = w(t) * g(t)$$

Wiener Filtering

$$MSE = \mathrm{E}\left\{[f(t) - \widehat{f}(t)]^2\right\} = \mathrm{E}\left\{[f(t) - w(t) * g(t)]^2\right\}$$

$$\frac{\partial MSE}{\partial w(t)} = -2\mathrm{E}\left\{[f(t) - w(t) * g(t)] g(t+\tau)\right\} = 0$$

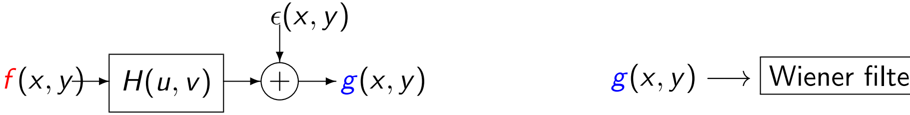
$$\mathrm{E}\left\{[f(t) - w(t) * g(t)]\, g(t+\tau)\right\} = 0 \quad \forall t, \tau \longrightarrow$$

$$R_{fg}(\tau) = w(t) * R_{gg}(\tau) \longrightarrow S_{fg}(\omega) = W(\omega) S_{gg}(\omega)$$

$$W(\omega) = \frac{S_{fg}(\omega)}{S_{gg}(\omega)} = \frac{H^*(\omega)\, S_{ff}(\omega)}{|H(\omega)|^2\, S_{ff}(\omega) + S_{\epsilon\epsilon}(\omega)}$$

$$\boxed{W(\omega) = \frac{H^*(\omega) S_{ff}(\omega)}{|H(\omega)|^2 S_{ff}(\omega) + S_{\epsilon\epsilon}(\omega)} = \frac{1}{H(\omega)} \frac{|H(\omega)|^2}{|H(\omega)|^2 + \frac{S_{\epsilon\epsilon}(\omega)}{S_{ff}(\omega)}}}$$

Wiener filtering



Signal	Image
$W(\omega) = \dfrac{S_{fg}(\omega)}{S_{gg}(\omega)}$	$W(u,v) = \dfrac{S_{fg}(u,v)}{S_{gg}(u,v)}$

$f(x,y)$ and $\epsilon(x,y)$ are assumed to be centred and non correlated

$$S_{fg}(u,v) = H'(u,v)\, S_{ff}(u,v)$$

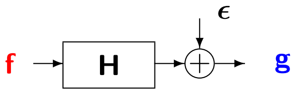
$$S_{gg}(u,v) = |H(u,v)|^2\, S_{ff}(u,v) + S_{\epsilon\epsilon}(u,v)$$

$$W(u,v) = \frac{H'(u,v) S_{ff}(u,v)}{|H(u,v)|^2 S_{ff}(u,v) + S_{\epsilon\epsilon}(u,v)}$$

Signal	Image
$W(\omega) = \dfrac{1}{H(\omega)} \dfrac{\vert H(\omega)\vert^2}{\vert H(\omega)\vert^2 + \frac{S_{\epsilon\epsilon}(\omega)}{S_{ff}(\omega)}}$	$W(u,v) = \dfrac{1}{H(u,v)} \dfrac{\vert H(u,v)\vert^2}{\vert H(u,v)\vert^2 + \frac{S_{\epsilon\epsilon}(u,v)}{S_{ff}(u,v)}}$

Wiener Filtering: Discrete version

Objective: a signal $\mathbf{f}$ is transmitted through a media which is assumed to act as a linear system. The transmission canal is also noisy. We receive the noisy signal $\mathbf{g}$. We want to estimate the transmitted signal $\widehat{\mathbf{f}}$.



$$\mathbf{g} \longrightarrow \boxed{\mathbf{W}} \longrightarrow \widehat{\mathbf{f}}$$

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}$$

$$\mathrm{E}\{\boldsymbol{\epsilon}\} = 0, \mathrm{E}\{\mathbf{f}\} = 0 \rightarrow \mathrm{E}\{\mathbf{g}\} = \mathbf{H}\mathrm{E}\{\mathbf{f}\} + \mathrm{E}\{\boldsymbol{\epsilon}\} = 0$$

$$\mathbf{R}_{\epsilon\epsilon} = \mathrm{E}\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'\}$$
$$\mathbf{R}_{\epsilon f} = \mathbf{R}'_{f\epsilon} = \mathrm{E}\{\boldsymbol{\epsilon}\,\mathbf{f}\} = 0$$
$$\mathbf{R}_{gg} = \mathrm{E}\{\mathbf{g}\,\mathbf{g}'\} = [\mathbf{H}\mathbf{R}_{ff}\mathbf{H}' + \mathbf{R}_{\epsilon\epsilon}]$$

$$\mathbf{R}_{ff} = \mathrm{E}\{\mathbf{f}\,\mathbf{f}'\}$$
$$\mathbf{R}_{gf} = \mathbf{R}'_{fg} = \mathrm{E}\{\mathbf{g}\,\mathbf{f}\} = \mathbf{H}'\mathbf{R}_{ff}$$

Wiener Filtering: Discrete version

Mean Square Error (MSE):

$$MSE = \mathrm{E}\left\{[\widehat{\mathbf{f}} - \mathbf{f}]'[\widehat{\mathbf{f}} - \mathbf{f}]'\right\} = \mathrm{E}\left\{[\mathbf{W}\mathbf{g} - \mathbf{f}]'[\mathbf{W}\mathbf{g} - \mathbf{f}]\right\} = \mathrm{E}\left\{\|\mathbf{W}\mathbf{g} - \mathbf{f}\|_2^2\right\}$$

Orthogonality: $\frac{\partial MSQE}{\partial \mathbf{W}} = 0$

$$\mathrm{E}\left\{[\widehat{\mathbf{f}} - \mathbf{f}]\mathbf{g}'\right\} = \mathrm{E}\left\{[\mathbf{W}\mathbf{g} - \mathbf{f}]\mathbf{g}'\right\} = 0 \rightarrow \mathrm{E}\left\{\mathbf{g}\mathbf{g}'\right\}\mathbf{W} = \mathrm{E}\left\{\mathbf{f}\mathbf{g}'\right\}$$

$$\mathbf{W} = \mathrm{E}\left\{\mathbf{f}\mathbf{g}'\right\}\left[\mathrm{E}\left\{\mathbf{g}\mathbf{g}'\right\}\right]^{-1} = \mathbf{R}_{fg}[\mathbf{R}_{gg}]^{-1}$$

$$\mathbf{R}_{fg} = \mathbf{R}_{ff}\mathbf{H}', \quad \mathbf{R}_{gg} = \mathbf{H}\mathbf{R}_{ff}\mathbf{H}' + \mathbf{R}_{\epsilon\epsilon}$$

$$\mathbf{W} = \mathbf{R}_{fg}[\mathbf{R}_{gg}]^{-1} = \mathbf{R}_{ff}\mathbf{H}'[\mathbf{H}\mathbf{R}_{ff}\mathbf{H}' + \mathbf{R}_{\epsilon\epsilon}]^{-1}$$

Wiener Filtering: Discrete version

- Matrix Inversion Lemma:

$$\mathbf{W} = \mathbf{R}_{ff}\mathbf{H}'[\mathbf{H}\mathbf{R}_{ff}\mathbf{H}' + \mathbf{R}_{\epsilon\epsilon}]^{-1} = [\mathbf{H}'\mathbf{R}_{\epsilon\epsilon}^{-1}\mathbf{H} + \mathbf{R}_{ff}^{-1}]^{-1}\mathbf{H}'\mathbf{R}_{\epsilon\epsilon}^{-1}$$

- $\mathbf{R}_{\epsilon\epsilon} = \sigma_\epsilon^2\mathbf{I}$ (white noise)

$$\mathbf{W} = \mathbf{R}_{ff}\mathbf{H}'[\mathbf{H}\mathbf{R}_{ff}\mathbf{H} + \sigma_\epsilon^2\mathbf{I}]^{-1} = [\mathbf{H}'\mathbf{H} + \sigma_\epsilon^2\mathbf{R}_{ff}^{-1}]^{-1}\mathbf{H}'$$

- Particular Case: $\sigma_b^2 = 0$ (No noise channel)

$$\mathbf{W} = \begin{cases} [\mathbf{H}'\mathbf{H}]^{-1}\mathbf{H}' & \longrightarrow \widehat{\mathbf{f}} = [\mathbf{H}'\mathbf{H}]^{-1}\mathbf{H}'\mathbf{g} \\ \mathbf{R}_{ff}\mathbf{H}'[\mathbf{H}\mathbf{R}_{ff}\mathbf{H}']^{-1} & \longrightarrow \widehat{\mathbf{f}} = \mathbf{R}_{ff}\mathbf{H}'[\mathbf{H}\mathbf{R}_{ff}\mathbf{H}']^{-1}\mathbf{g} \end{cases}$$

Generalized Wiener filter using unitary transforms

$$\mathbf{W} = [\mathbf{H}'\mathbf{H} + \sigma_b^2 \mathbf{R}_{ff}^{-1}]^{-1}\mathbf{H}'$$

$$\mathbf{P} \stackrel{\triangle}{=} [\mathbf{H}'\mathbf{H} + \sigma_b^2 \mathbf{R}_{ff}^{-1}]^{-1}$$

Consider a unitary transform $\mathbf{F}$ such that $\quad \mathbf{F}'\mathbf{F} = \mathbf{F}\mathbf{F}' = \mathbf{I}$

$$\widehat{\mathbf{f}} = \mathbf{F}'[\mathbf{F}\mathbf{P}\mathbf{F}']\mathbf{F}\mathbf{H}'\mathbf{g} \stackrel{\triangle}{=} \mathbf{F}'\bar{\mathbf{P}}\mathbf{z}$$

$$\bar{\mathbf{P}} \stackrel{\triangle}{=} [\mathbf{F}\mathbf{P}\mathbf{F}'], \quad \mathbf{z} \stackrel{\triangle}{=} \mathbf{F}\mathbf{H}'\mathbf{g}$$

$$\mathbf{g} \longrightarrow \boxed{\mathbf{H}'} \longrightarrow \boxed{\mathbf{F}} \longrightarrow \mathbf{z} \longrightarrow \boxed{\bar{\mathbf{P}} = [\mathbf{F}\mathbf{P}\mathbf{F}']} \longrightarrow \widehat{\mathbf{z}} \longrightarrow \boxed{\mathbf{F}'} \longrightarrow \widehat{\mathbf{f}}$$

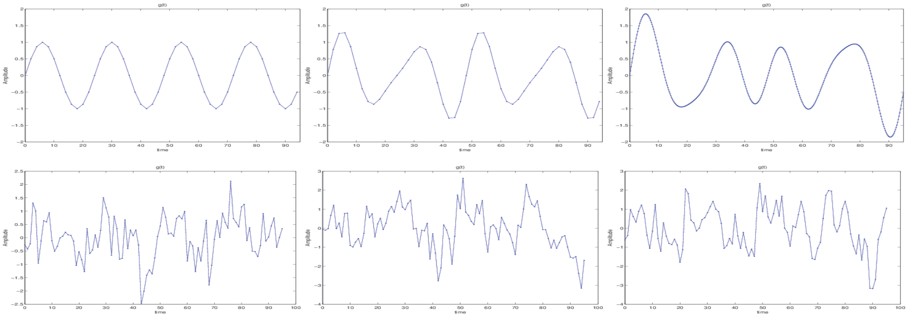
For an appropriate unitary transforms $\bar{\mathbf{P}}$ becomes an almost diagonal matrix

$$\widehat{\mathbf{z}} = \bar{\mathbf{P}}\mathbf{z} \quad \Longrightarrow \quad \widehat{z}(k) \simeq \bar{p}(k)\, z(k)$$

2. modelling: parametric and non-parametric, MA, AR and ARMA models

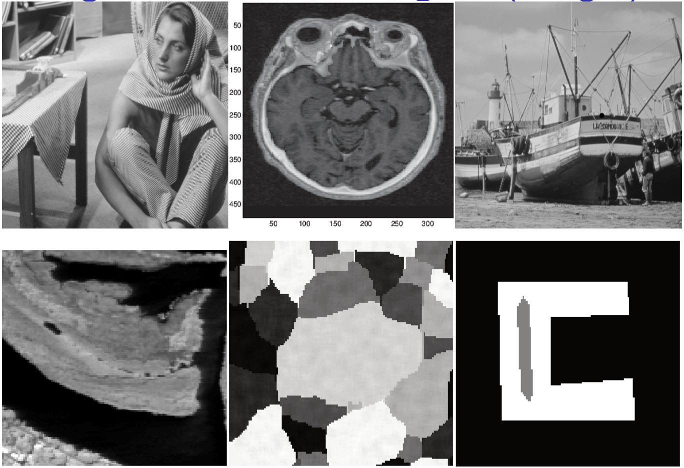
- modelling ? for what ?
- Deterministic / Probalistic modelling
- Parametric / Non Parametric
- Moving Average (MA)
- Autoregressive (AR)
- Autoregressive Moving Average (ARMA)
- Classical methods for parameter estimation (LS, WLS)

modelling ? What for ? 1D signals



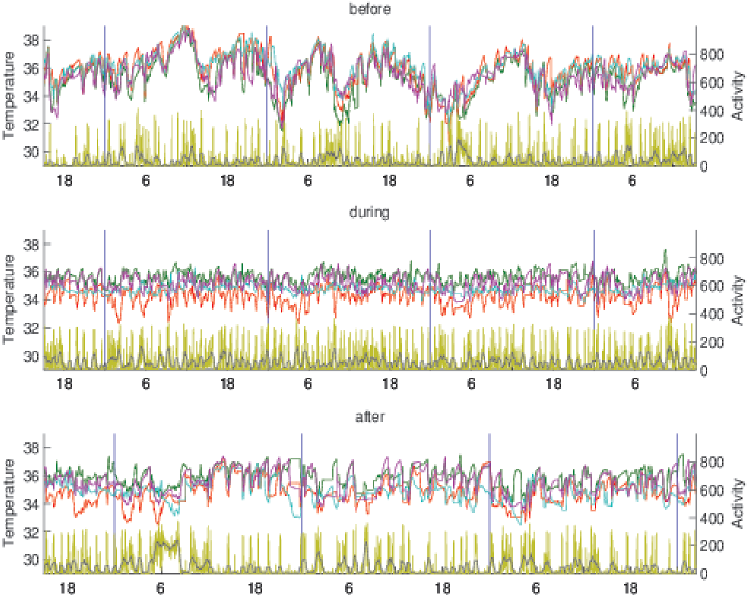
- 1D signals:
  - Is it periodic? What is the period?
  - Is there any structure?
  - Has something changed before, during and after some traitement
  - Can we compress it? How? How much?

modelling ? What for ? 2D signals (Images)

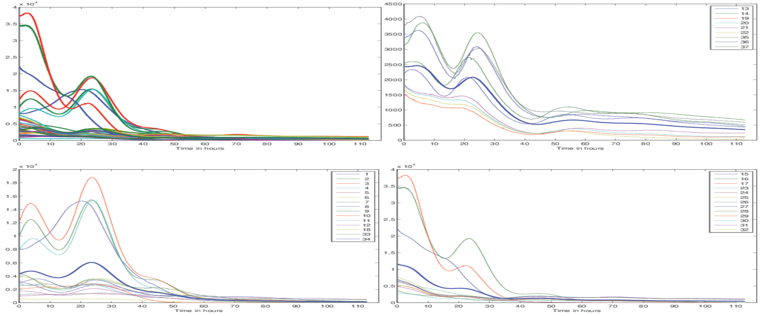


- Images:
  - Is there any structure?
  - Contours? Regions?
  - Can we compress it? How? How much?

modelling ? What for ? multi dimensional time series



modelling ? What for ? multi dimensional time series

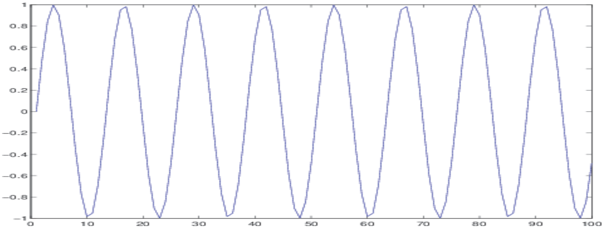


- Multi Dimentionsional signals $g_1(t), \cdots, g_n(t)$
  - Dependancy: Are they all independent?
    If not, which ones are related?
  - Dimensionality reduction: Can we reduce the dimensionality?
  - Principal Components Analysis (PCA):
    What are the principal components?
  - Independent Components Analysis (ICA):
    What are the independent components?
  - Factor Analysis (FA): What are the principal factors?

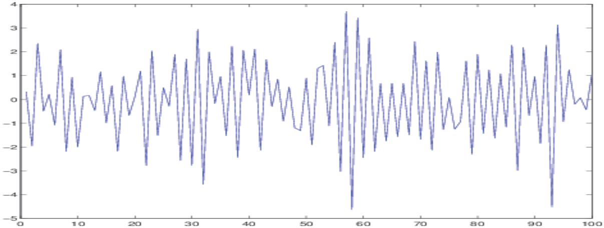
Deterministic / Probalistic modelling

- **Deterministic:**
  - The signal is a sinusoid $X(t) = a \sin(\omega t + \phi)$. We need just to determine the three parameters a, ω, ϕ.
  - The signal is periodic
    $X(t) = \sum_{k=1}^{K} a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t)$.
    If we know ω_0, then, we need just to determine the parameters $(a_k, b_k), k = 1, \cdots, K$.
  - The signal represents a Gaussian form spectra
    $X(t) = \sum_{k=1}^{K} a_k \mathcal{N}(m_k, v_k)$. We need just to determine the parameters $(a_k, m_k, v_k), k = 1, \cdots, K$.
  - In the last two cases, one great difficulty is determining K
- **Probabilistic:**
  - The shape of the signal is more sophisticated.
  - 1 Sinusoid + noise $X(t) = a \sin(\omega t + \phi) + \epsilon(t)$
  - K Sinusoids + noise $X(t) = \sum_{k=1}^{K} a_k \sin(\omega_k t + \phi_k) + \epsilon(t)$
  - No specific shapes: MA, AR, ARMA, ...

Determinist/Probabilist

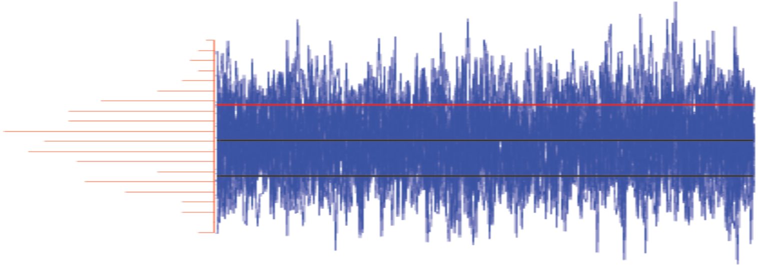


$X(t) = a \sin(\omega t + \phi),\ a = 1, \omega = 2\pi, \phi = 0$

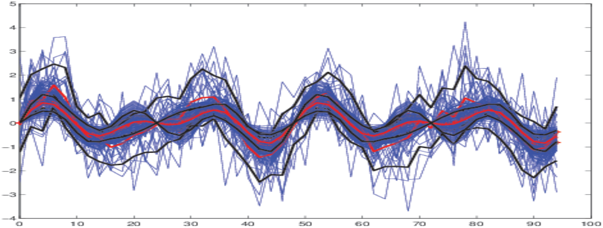


ARMA(1,2), $q = 1,\ v = 1,\ b_1 = 1,\ a_1 = 1, a_2 = .8$

Stationary/Non Stationary



$$X(t) \sim \mathcal{N}(0,1), \quad \forall t \longrightarrow \mathrm{E}\{X(t)\} = 0, \mathrm{Var}\{X(t)\} = 1, \forall t$$



$$f_3(t) = a_1 \sin(2\omega t) + a_2 \sim (3\omega t),\ \omega = \pi/12,\ a_1, a_2 \sim \mathcal{U}(0,1)$$

Parametric / Non Parametric

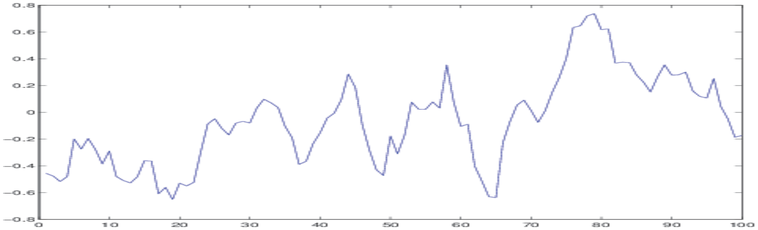
- **Parametric**
  - K Sinusoids + noise $X(t) = \sum_{k=1}^{K} a_k \sin(\omega_k t + \phi_k) + \epsilon(t)$. The parameters are $(a_k, \omega_k, \phi_k), k = 1, \cdots, K$ and v_ϵ.
  - K Complex exponentials + noise $X(t) = \sum_{k=1}^{K} c_k \exp\{-j\omega_k t\} + \epsilon(t)$. The parameters are $(c_k, \omega_k), k = 1, \cdots, K$ and v_ϵ.
  - Sum of K Gaussian shapes: $X(t) = \sum_{k=1}^{K} a_k \mathcal{N}(m_k, v_k)$. The parameters are $(a_k, m_k, v_k), k = 1, \cdots, K$ and v_ϵ.
- **Non-Parametric**
  - The shape of the signal is more sophisticated.
  - The shape is composed of as much as the number of data of Complex exponentials + noise $X(t) = \sum_{n=1}^{N} c_n \exp\{-jn\omega_0 t\} + \epsilon(t)$. If we know ω_0, then, the parameters are $c_n, n = 1, \cdots, N$ and v_ϵ.
  - Sum of the Gaussian shapes: $X(t) = \sum_{k=1}^{K} a_n \mathcal{N}(m_n, v_n)$. The parameters are $(a_n, m_n, v_n), n = 1, \cdots, N$ and v_ϵ.

Moving Average (MA)

$$\text{Continuous:} \quad x(t) = b(t) * \epsilon(t) = \int b(\tau)\epsilon(t-\tau)\,\mathrm{d}\tau$$

$$\text{Discrete:} \quad x(n) = \sum_{k=0}^{q} b(k)\epsilon(n-k), \quad \forall n$$

$$\epsilon(n) \longrightarrow \boxed{B(z) = \sum_{k=0}^{q} b(k) z^{-k}} \longrightarrow x(n)$$



MA, $q = 15$, $v = 1$, $b_k = \exp\left\{-.05 * k^2\right\}$

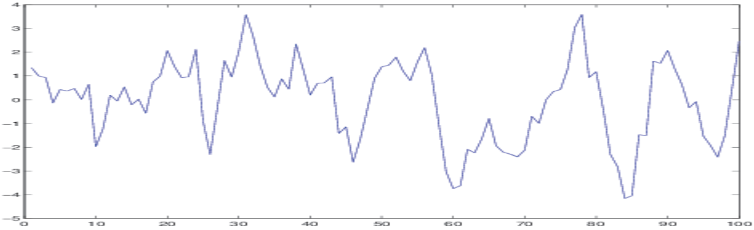
Autoregressive (AR)

$$x(t) = \sum_{k=1}^{p} a(k)\, x(t-k\Delta t) + \epsilon(t) \longrightarrow x(n) = \sum_{k=1}^{p} a(k)\, x(n-k) + \epsilon(n), \quad \forall n$$

$$\mathrm{E}\left\{\epsilon(n)\right\} = 0, \quad \mathrm{E}\left\{|\epsilon(n)|^2\right\} = \beta^2,$$

$$\mathrm{E}\left\{\epsilon(n)\, x(m)\right\} = 0, \quad m \neq n$$

$$\epsilon(n) \longrightarrow \boxed{H(z) = \frac{1}{A(z)} = \frac{1}{1 - \sum_{k=1}^{p} a(k)\, z^{-k}}} \longrightarrow x(n)$$



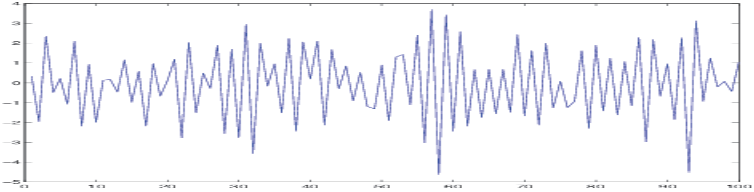
AR1, $a = .7$, $v = 1$.

Autoregressive Moving Average (ARMA)

$$x(n) = \sum_{k=1}^{p} a(k)\, x(n-k) + \sum_{l=0}^{q} b(l)\, \epsilon(n-l)$$

$$\epsilon_n \longrightarrow \boxed{H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^{q} b(k) z^{-k}}{1 - \sum_{k=1}^{p} a(k)\, z^{-k}}} \longrightarrow x(n)$$

$$\epsilon_n \longrightarrow \boxed{B_q(z)} \longrightarrow \boxed{\frac{1}{A_p(z)}} \longrightarrow x(n)$$



ARMA(1,2), $q = 1$, $v = 1$, $b_1 = 1$, $a_1 = 1$, $a_2 = .8$

Causal ou Non-Causal AR models

- **Causal :**

$$x(n) = \sum_{k=1}^{q} a(k)\, x(n-k) + \epsilon_n, \quad \forall n$$

$$A(z) = 1 - \sum_{k=1}^{q} a(k)\, z^{-k} \longrightarrow \epsilon_n \longrightarrow \boxed{H(z) = \frac{1}{A(z)}} \longrightarrow x(n)$$

- **Non–causal :**

$$x(n) = \sum_{\substack{k=-p \\ k\neq 0}}^{+q} a(k)\, x(n-k) + \epsilon_n, \quad \forall n$$

$$A(z) = 1 - \sum_{\substack{k=-p \\ k\neq 0}}^{+q} a(k)\, z^{-k} \longrightarrow \epsilon_n \longrightarrow \boxed{H(z) = \frac{1}{A(z)}} \longrightarrow x(n)$$

2D AR Models

$$A(z_1, z_2) = 1 - \sum_{(k,l)} \sum_{\in S} a(k,l) z_1^{-k} z_2^{-k}$$

$$f(m,n) = \sum_{(k,l)} \sum_{\in S} a(k,l)\, f(m-k, n-l) + \epsilon(m,n)$$

- **Non–causal**
  $S = \{l \geq 1, \forall k\} \cup \{l = 0, k \neq 0\}$
- **Semi–ausal**
  $S = \{l \geq 1, \forall k\} \cup \{l = 0, k \geq 1\}$
- **Causal**
  $S = \{(k,l) \neq (0,0)\}$

2D AR Models

- **Causal**
  - $S = \{l \geq 1, \forall k\} \cup \{l = 0, k \neq 0\}$
  - Recursive Filtre
  - Finite Differential Equations with initial conditions
  - Hyperbolic Partial Differential Equations
- **Semi–causal**
  - $S = \{l \geq 1, \forall k\} \cup \{l = 0, k \geq 1\}$
  - Semi–recursif Filters
  - Finite Differential Equations with initial conditions in one dimention and limit conditions in other dimension
  - Parabolic Partial Differential Equations
- **Non–causal**
  - $S = \{(k, l) \neq (0, 0)\}$
  - Non-recursive Filtre
  - Finite Differential Equations with limit conditions in both dimensions
  - Elliptic Partial Differential Equations

Causal/Non-Causal Prediction

- **Causal :**

$$\widehat{x}(n) = \sum_{k} a(k)\, x(n-k)$$

- **Non-Causal :**

$$\widehat{x}(n) = \sum_{\substack{k=-p \\ k \neq 0}}^{+q} a(k)\, x(n-k)$$

2D AR models and 2D prediction

$$A(z_1, z_2) = 1 - \sum_{(k,l)} \sum_{\in S} a(k, l) z_1^{-k} z_2^{-k}$$

$$f(m, n) = \sum_{(k,l)} \sum_{\in S} a(k, l)\, f(m - k, n - l) + \epsilon(m, n)$$

- **Non–causal**
- **Semi–ausal**
- **Causal**

2D AR models and 2D prediction

- **Causal**
  - $S = \{l \geq 1, \forall k\} \cup \{l = 0, k \neq 0\}$
  - Recursive Filtering
  - Finite Difference Equation (FDE) with initial conditions
  - Partial Differential Equations (Hyperbolic)
- **Semi–causal**
  - $S = \{l \geq 1, \forall k\} \cup \{l = 0, k \geq 1\}$
  - Filtre semi–récursif
  - FDE with initial conditions in one direction and limit conditions in other direction.
  - Partial Differential Equations (Parabolic)
- **Non–causal**
  - $S = \{(k, l) \neq (0, 0)\}$
  - Non-Recursive Filtering
  - FDE with limit conditions in both directions
  - Partial Differential Equations (Elliptic)