

**Chapitre 1 : Rappels sur les variables et vecteurs aléatoires**

1. Variables aléatoires et probabilités
  - Probabilité individuelle et probabilité conjointe
  - Notion d'indépendance
2. Moments statistiques
3. Lois de probabilité importantes (discrètes et continues)
4. Vecteurs aléatoires

**Chapitre 2 : Processus stochastiques et signaux aléatoires**

1. Fonction aléatoire
2. Espérances ou Moments d'une fonction aléatoire
  - Statistiques d'ordre 1 (moyenne, variance, etc.)
  - Statistiques d'ordre 2 (corrélation et covariance)
3. Processus stationnaires
  - Stationnarité au sens large et au sens strict
  - Propriétés de la fonction de corrélation et notion de mémoire
  - Puissance et DSP
4. Processus ergodiques
  - Moyennes temporelles
  - Théorème de Wiener Khintchine

**Chapitre 3 : Filtrage linéaire des signaux aléatoires (SA)**

1. Effet d'un système linéaire sur un SA
  - Moyenne et autocorrélation
  - Densité spectrale et formule des interférences
2. Filtrage adapté
3. Filtrage optimal

**Chapitre 4 : Processus AR, MA et ARMA et filtrage de Wiener**

1. Processus auto-régressif (AR)
2. Processus à moyenne mobile (MA ou FIR)
3. Processus auto-régressif à moyenne mobile (ARMA)
4. Filtre de Wiener

**Chapitre 5 : Théorie de l'estimation**

1. Notion d'estimation (Exemples)
2. Propriétés des estimateurs (biais, variance, efficacité, etc.)
3. Estimation des moyennes statistiques
4. Estimation spectrale (périodogramme, corrélogramme)
5. Estimateurs (Bayes, Maximum de vraisemblance, Moindres carrés)
6. Filtre de Kalman

**TP**

TP n°1 : Variables et Vecteurs aléatoires

TP n°2 : Processus aléatoires (Stationnarité et Ergodisme)

TP n°3 : Filtrage linéaire (optimal et adapté)

TP n°4 : Modélisation AR

TP n°5 : Estimation (moments, spectrale) et Filtrage de Wiener

TP n°6 : Filtrage de Kalman

## Chapitre I : Rappels sur les variables et vecteurs aléatoires

Tout signal physique comporte une composante aléatoire qu'il s'agisse d'une perturbation externe telle une tension aléatoire aux bornes d'une résistance ou une perturbation atmosphérique. Le signal utile est lui même aléatoire si l'on considère le lancé de dés, le loto, les résultats de matchs ou encore les phénomènes météorologiques, etc., ...

Rappelons qu'un signal déterministe possède une formulation mathématique qui permet de connaître sa valeur à tout instant. A contrario, la connaissance de la valeur d'un signal aléatoire à l'instant  $t$  ne nous permet pas de connaître sa valeur à  $t + \tau$ . On ne pourra que faire des prévisions. On dira par exemple qu'il y a 30% de chance que sa valeur reste constante, 20% chance que sa valeur augmente et 50% qu'elle diminue. En fait, on fera appel à des paramètres statistiques définissant les *possibilités* d'évolution du signal. Ainsi, pour étudier l'évolution des phénomènes aléatoires, on a recours à des modèles probabilistes. Ceux-ci tentent de représenter le signal observé par une famille de Variables Aléatoires (VA) indexées par le temps. Chacune de ces VA décrit l'aspect incertain du phénomène à un instant donné.

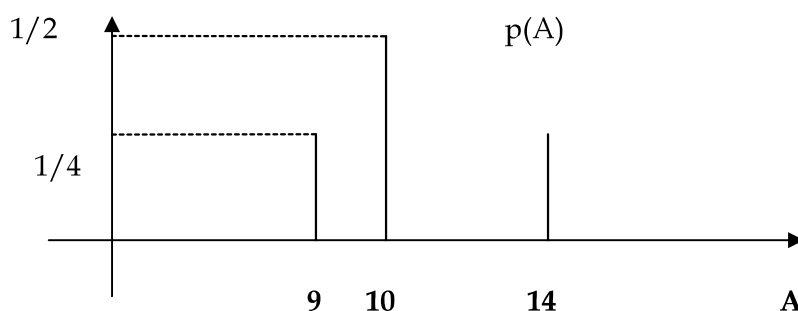
### 1. Variables aléatoires et probabilités

Une VA est discrète ou continue suivant que l'ensemble des valeurs possibles de l'expérience aléatoire est respectivement discret ou continu (rien à voir avec le temps). Un exemple du cas discret est donné par le nombre d'étudiants ayant acquis un module. Par contre, le poids ou la taille d'un individu fournit un exemple de VA. Une variable aléatoire est caractérisée par l'ensemble des valeurs qu'elle peut prendre et par l'expression mathématique de la probabilité de ces valeurs. Cette expression s'appelle *la loi de probabilité* (ou *distribution de probabilité*) de la variable aléatoire. Quant à la fonction de répartition elle représente la probabilité pour qu'une VA  $X$  soit inférieure ou égale à  $x_i$  :  $F(x_i) = \text{Prob}(X \leq x_i)$

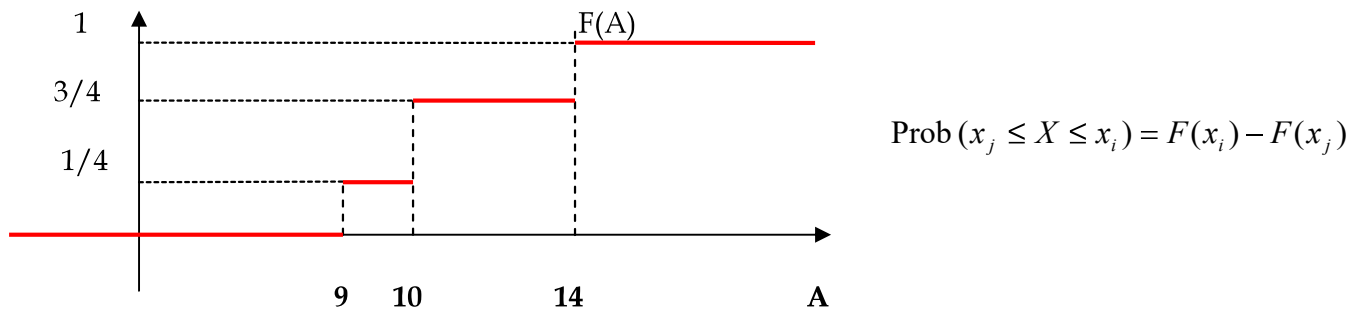
Exemple 1: Le jet de dé.

Il y a six valeurs possibles  $x_i = \{1, 2, 3, 4, 5, 6\}$  avec  $\text{prob}(x_i) = 1/6$  et  $\sum_i \text{prob}(x_i) = 1$

Exemple 2 : Soient 4 étudiants ayant obtenu les notes suivantes à l'examen : 10, 9, 10, 14 dans un module A. Si on cherche la probabilité d'avoir chaque note on trouve :  $\text{prob}(A=9)=1/4$ ,  $\text{prob}(A=10)=1/2$ ,  $\text{prob}(A=14)=1/4$ . On obtient le tracé suivant de  $p(A)$  :



Si on veut connaître la probabilité d'avoir une note inférieure ou égale à 10 ( $\leq 10$ ),  $p(A \leq 10) = 1/4 + 1/2 = 3/4$ . Si on fait ce calcul pour toutes les valeurs A, on obtient la fonction de répartition F(A):



Concrètement la fonction de répartition correspond à la distribution des probabilités cumulées. Le plateau atteint par la fonction de répartition correspond à la valeur de probabilité 1. Ainsi  $F(-\infty) = 0$  et  $F(+\infty) = 1$

### Probabilité simple et probabilité conjointe

Il est évident qu'une probabilité n'a de sens que si celle-ci est calculée sur un grand nombre. Imaginons que l'on lance un dé 6 fois, on n'obtiendra sûrement 6 faces différentes, ni même autant de faces pour 12 ou 60 lancers (voir exo n°1 du TP n°1). Ce n'est qu'au bout d'un nombre de lancers très important ( $N \gg$ ) que les 6 faces apparaîtront avec les mêmes probabilités.

Soit A un résultat possible d'une expérience aléatoire, N le nombre de réalisation de l'expérience et  $n_A$  le nombre de fois que le résultat A est apparu, alors la probabilité de l'événement "le résultat de l'expérience est A" :

$$p(A) \stackrel{\text{déf.}}{=} \lim_{N \rightarrow +\infty} \frac{n_A}{N}$$

De la même façon, on peut définir la probabilité conjointe : Soient A et B deux événements distincts alors la probabilité que A et B aient lieu conjointement est:

$$p(A, B) \stackrel{\text{déf.}}{=} \lim_{N \rightarrow +\infty} \frac{n_{AB}}{N}$$

avec  $n_{AB}$  = nombre de fois que A et B sont apparus

qui peut aussi s'écrire 
$$p(A, B) = \lim_{N \rightarrow +\infty} \left( \frac{n_{AB}}{n_A} \frac{n_A}{N} \right) = \lim_{N \rightarrow +\infty} \left( \frac{n_{AB}}{n_A} \right) \cdot \lim_{N \rightarrow +\infty} \left( \frac{n_A}{N} \right) \stackrel{\text{déf.}}{=} p(B/A) \cdot p(A)$$

On appelle  $p(B/A)$  une probabilité conditionnelle c'est la probabilité que B ait lieu sachant que A a eu déjà lieu. La formulation de la probabilité conditionnelle  $p(A/B)$  permet de retrouver le théorème de Bayes :

$$p(A, B) = p(B/A) \cdot p(A) = p(A/B) \cdot p(B)$$

*Notion d'indépendance:* Supposons que l'on veuille connaître la probabilité qu'il pleuve à Alger (Evènement A) sachant qu'il fait beau / Rio (Evènement B). L'un n'ayant aucune influence sur l'autre (les deux événements étant indépendants) alors :  $p(A/B) = p(A)$  de même  $p(B/A) = p(B)$  d'où  $p(A, B) = p(A) \cdot p(B)$



Remarque: Pour une variable aléatoire continue, la probabilité associée à l'évènement  $\{X = a\}$  est nulle, car il est impossible d'observer exactement cette valeur. On considère alors la probabilité que la variable aléatoire  $X$  prenne des valeurs comprises dans un intervalle  $[a, b]$  tel que  $p(a \leq X \leq b)$ . Lorsque cet intervalle tend vers 0, la valeur prise par  $X$  tend alors vers une fonction que l'on appelle fonction densité de probabilité ou densité de probabilité : [1]

$$\text{Prob}(x \leq x_i \leq x + \Delta x) = \frac{\text{Valeurs entre } x \text{ et } x + \Delta x}{\text{Nombre de valeurs}} \Rightarrow p_X(x) = \lim_{\Delta x \rightarrow 0} \frac{\text{Prob}(x \leq x_i \leq x + \Delta x)}{\Delta x}$$

Lorsque  $\Delta x \rightarrow dx$ , l'histogramme devient une fonction continue de  $x \Rightarrow \text{Prob}[x_1 \leq x_i < x_2] = \int_{x_1}^{x_2} p(x) dx$   
avec  $\int_{-\infty}^{+\infty} p(x) dx = 1$

La fonction de répartition est alors la primitive de la fonction densité de probabilité  $p(x)$ :

$$F(x) = \int_{-\infty}^x p(u) du \quad \Rightarrow \quad p(x) = \frac{dF(x)}{dx}$$

## 2. Moments statistiques

Le tracé de la densité de probabilité nous fournit une première information sur notre VA que l'on complètera par le calcul de différents paramètres statistiques. Deux d'entre eux sont largement utilisés, il s'agit de la moyenne et la variance.

La valeur moyenne équivaut à la composante continue et la variance aux fluctuations autour de cette moyenne. Ainsi, pour le suivi du poids ou de la taille d'un enfant, on se réfère à la moyenne calculée par tranches d'âge pour plusieurs enfants. L'écart par rapport à cette moyenne (supérieur ou inférieur) permettra de savoir si l'enfant a un poids normal pour son âge ou au contraire est en surpoids ou malnutrition



- Pour une VA discrète, la moyenne et la variance sont données par:

$$m = \mu_x = E\{x\} = \sum_i x_i p(x_i) \quad \text{Var} = \sigma_x^2 = E\{(x - \mu_x)^2\} = \sum_i (x_i - \mu_x)^2 p(x_i) = \sum_i x_i^2 p(x_i) - \mu_x^2$$

- Pour une VA continue  $x \in [a, b]$ , la moyenne et la variance sont données par:

$$m = \mu_x = E\{x\} = \int_a^b x \cdot p(x) dx \quad \text{Var} = \sigma_x^2 = E\{(x - \mu_x)^2\} = \int_a^b (x - \mu_x)^2 p(x) dx = \int_a^b x^2 p(x) dx - \mu_x^2$$

De façon plus générale, toute fonction  $f$  d'une variable aléatoire  $x$  est une variable aléatoire  $f$ . La connaissance des lois de probabilités nous permet de déterminer les *valeurs espérées théoriquement* - ou *espérances mathématiques* - ou encore *moyenne (ou moments) statistique* de  $f$  :

$$E\{f(x)\} = \int_{-\infty}^{+\infty} f(x) p(x) dx$$

### Propriétés de l'espérance

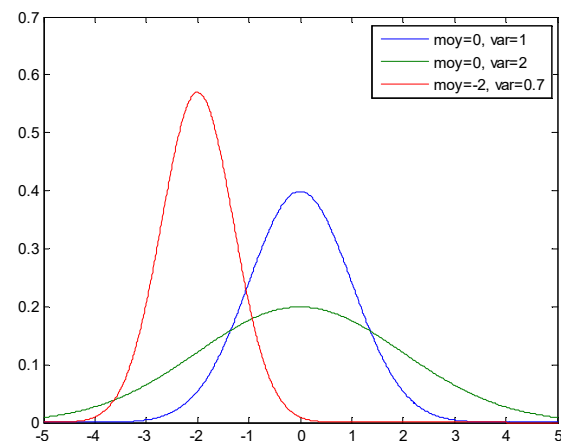
- $E\{X+Y\}=E\{X\}+E\{Y\}$
- $E\{aX\}=aE\{X\} \quad \forall a \in \mathbb{R}$  (a étant une valeur déterministe)
- $E\{a\}=a \quad \forall a \in \mathbb{R}$
- Si  $X \geq 0$  alors  $E\{X\} \geq 0$

Remarque : Lorsque la moyenne est nulle, on dit que la variable aléatoire est *centrée*, lorsque sa variance vaut 1, on dit qu'elle est *réduite*

### 3. Loïs de probabilité usuelles

**Loi Normale ou gaussienne**: Une des plus importantes lois car elle permet de modéliser des phénomènes naturels issus de plusieurs événements aléatoires. Elle est caractérisée par sa moyenne *m* et sa variance  $\sigma^2$ . Plus la variance est élevée, plus la courbe est aplatie.

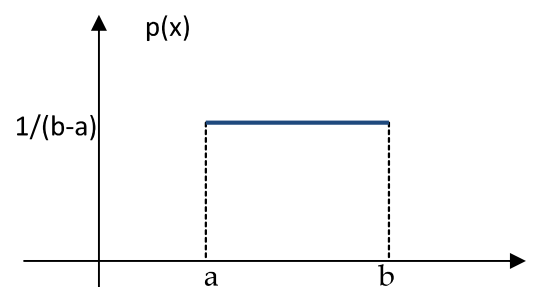
$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}\right)$$



Exemple: Notes obtenues à un examen par un ensemble d'étudiants. La plus part des notes seront comprises entre 7 et 13. Très peu d'étudiants auront une note  $>18$  ou  $<0$ . La loi des notes est gaussienne du fait que la note d'un étudiant dépend de plusieurs facteurs aléatoires (cours donné, révision ou pas, aptitudes, disponibilités de documents, état moral, etc.)

**Loi uniforme** :  $X$  est uniformément répartie entre  $a$  et  $b$  veut dire que la probabilité est identique (ou équiprobable) quelque soit la valeur de  $x$  sur  $[a, b]$  :

$$p_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{ailleurs} \end{cases}, \quad \begin{aligned} m &= (b+a)/2, \\ \sigma^2 &= (b-a)^2/12 \end{aligned}$$

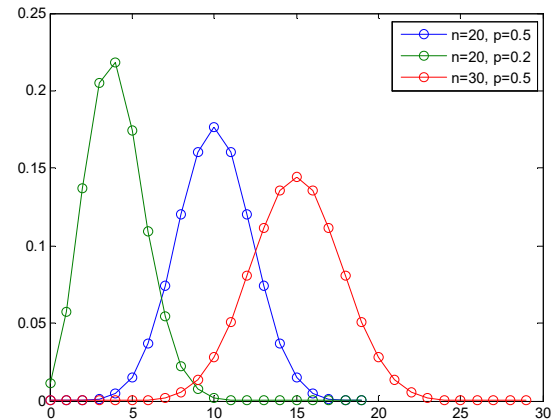


**Loi de Bernoulli** : La loi de Bernoulli, de paramètre  $p$ , est une expérience avec seulement deux issues possibles  $a$  et  $b$  avec les probabilités respectives de  $p$  et  $(1-p)$ . *Exemple*: Résultat pile ou face d'une pièce.

- La densité de probabilité est donnée par :  $p_X(x) = p \cdot \delta(x-a) + (1-p) \cdot \delta(x-b)$
- La fonction de répartition par :  $F(x) = p \cdot U(x-a) + (1-p) \cdot U(x-b)$

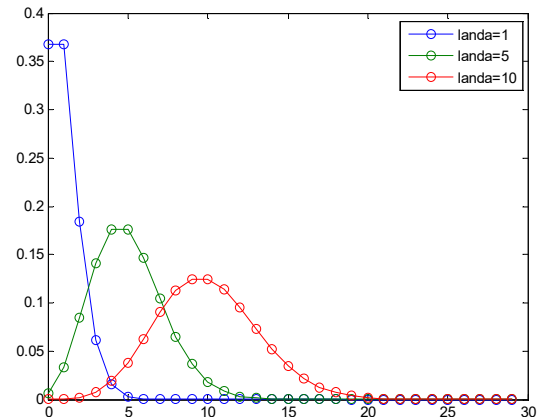
**Loi Binomiale :** de paramètre  $(n, p)$ , c'est la répétition de  $n$  expériences de Bernoulli. Soit un événement  $A$  associé à une expérience. Cet événement a une probabilité  $p$  de survenir. On reproduit  $n$  fois l'expérience et on s'intéresse à la VA  $X$  qui donne le nombre de succès de l'événement  $A$  au cours des  $n$  essais.

$$p_X(x) = C_n^x p^x (1-p)^{(n-x)} C_n^k = \frac{n!}{k!(n-k)!} \quad m = np \quad \sigma^2 = np(1-p)$$



**Loi de Poisson :** La loi de Poisson est appelée la loi des événements rares, elle dépend d'un paramètre  $\lambda$ ; elle trouve son utilité pour la description du comportement d'événements dont les chances de réalisation sont faibles. De nos jours, elle est très utilisée dans les télécommunications (pour compter le nombre de communications dans un intervalle de temps donné), le contrôle de qualité statistique (nombre de pièces défectueuses), etc.

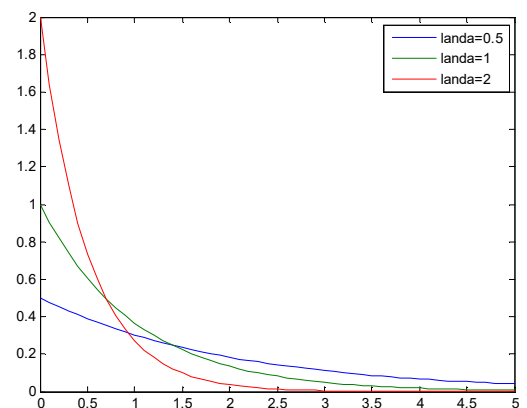
$$p_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad m = \lambda, \quad \sigma^2 = \lambda$$



**Exemple :** Si un événement donné se produit en moyenne 10 fois par minute, pour étudier le nombre d'événements se produisant dans un laps de temps d'une heure, on choisit comme modèle une loi de Poisson de paramètre  $\lambda = 60 \times 10 = 600$ .

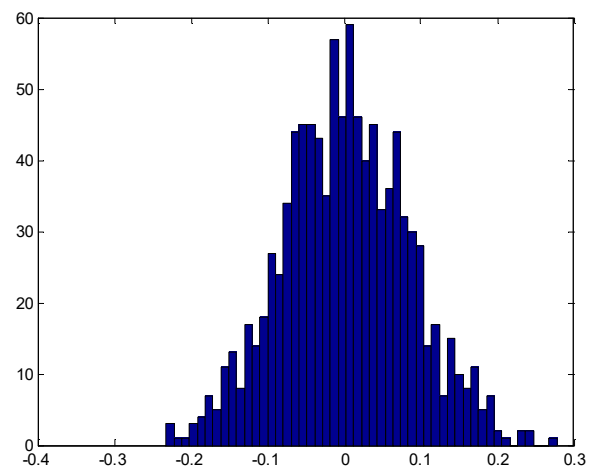
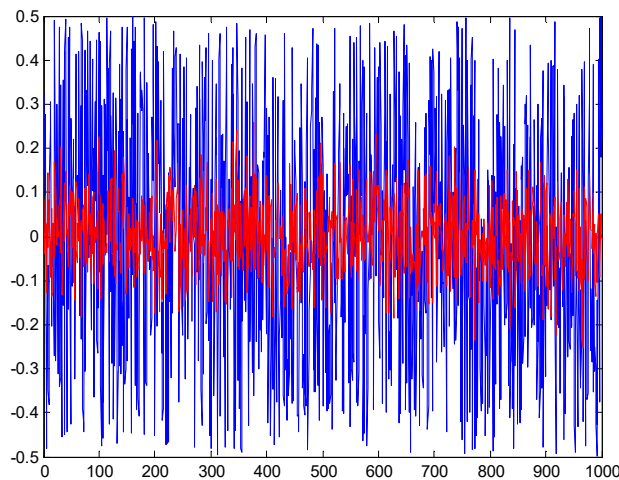
**Loi exponentielle :** Dans ce cas  $X$  est une variable aléatoire définissant la durée de vie d'un phénomène. Si l'espérance de vie du phénomène est  $E(X)$  et si la durée de vie est sans vieillissement (ex: composant électronique (voir exo n°5 du TP n°1).), c'est-à-dire si la durée de vie au-delà de l'instant  $T$  est indépendante de l'instant  $T$ , alors  $X$  a pour densité de probabilité:

$$p(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0 \end{cases} \quad m = 1/\lambda \quad \sigma^2 = 1/\lambda^2$$



En **théorie des files d'attente**, l'arrivée de clients dans une file est souvent modélisée par cette loi.

**Théorème central limite et approximations :** Chaque fois qu'un phénomène peut être considéré comme la résultante d'un grand nombre de causes aléatoires indépendantes (ex: notes obtenues à l'examen, poids d'un enfant), on peut résumer que ce phénomène suit une loi de distribution normale: c'est le théorème central limite. Ainsi, la distribution statistique de la somme de  $n$  VA indépendantes et de même loi tend vers la loi Normale quand  $n$  tend vers l'infini. L'exemple suivant illustre la somme de 12 signaux indépendants et uniformément répartis entre -0.5 et +0.5 et sa densité de probabilité qui suit une loi gaussienne (voir exo n°3 du TP n°1).



Exemple d'application : Soient les notes obtenues dans 2 modules par une section de 10 étudiants .

1. Quel est le nombre d'étudiants ayant le module1
2. Quelle est la probabilité d'avoir la moyenne dans le module1
3. Quelle est la probabilité d'avoir la moyenne dans le module2
4. Quelle est la probabilité d'avoir le module2 quand on eu le module1
5. Quelle est la probabilité d'avoir les 2 modules simultanément
6. Quelle est la probabilité d'avoir une note inférieure à 10 dans le module1

Etudiants	1	2	3	4	5	6	7	8	9	10
Module1	8	10	17	7	9	11	13	6	12	8
Module2	11	12	13	9	10	8	11	8	10	6

#### 4. Vecteurs Aléatoires

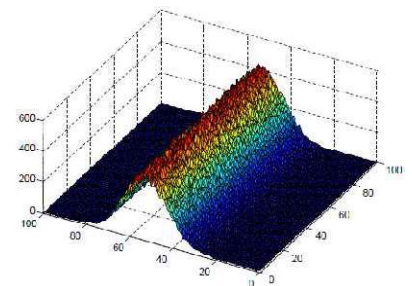
Une variable aléatoire à plusieurs dimensions (encore appelée vecteur de VA) est le résultat dépendant de plusieurs caractères aléatoires. Le tirage de 2 dés à six faces donne une VA discrète à deux dimensions (une matrice 6x6 de probabilités 1/36). Par contre, la taille et le poids forment une VA continue à 2 dimensions.

Exemple:  $p(X,Y)$  ou  $X$  est une Gaussienne et  $Y$  est uniforme

Des variables aléatoires  $X_1, X_2, \dots, X_n$  ont une loi de distribution conjointe si elles sont définies par rapport au même univers  $\Omega$ . Elles peuvent être caractérisées via soit:

- leur fonction de répartition conjointe :  $F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$
- leur fonction de densité de probabilité conjointe :  $p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)$  définie par:

$$F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} p_{X_1, X_2, \dots, X_n}(u_1, u_2, \dots, u_n) du_1 \dots du_n$$



Loi de distribution marginale

Si l'on s'intéresse à un sous-ensemble  $X_1, X_2, \dots, X_m$ ,  $m < n$  des variables aléatoires précédentes distribuées conjointement, on définit la loi de distribution marginale à partir de la distribution conjointe ont une loi de distribution conjointe si elles sont définies par rapport au même univers  $\Omega$ . Elles peuvent être caractérisées via :

$$p_{X_1, X_2, \dots, X_m}(x_1, x_2, \dots, x_m) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p_{X_1, X_2, \dots, X_n}(u_1, \dots, u_m, u_{m+1}, \dots, u_n) du_{m+1} \dots du_n$$

Exemple 1 (VA discrète)

$$\text{Soit Prob}(x,y) = \begin{bmatrix} 1/9 & 1/3 & 0 \\ 2/9 & 1/9 & 2/9 \end{bmatrix}$$

On peut alors déterminer la probabilité de  $x$  soit  $\text{prob}(x)$  et de même celle de  $y$

$$\text{Prob}(x) = [1/9 + 1/3 = 4/9 \quad 2/9 + 1/9 + 2/9 = 5/9]$$

$$\text{Prob}(y) = [1/9 + 2/9 = 1/3 \quad 1/3 + 1/9 = 4/9 + 2/9 = 2/9]$$

Exemple 2 Soient  $X$  et  $Y$  deux variables aléatoires uniformément distribuées sur  $[0, T] \times [0, T]$ , alors :

$$p_{X,Y}(x,y) = \begin{cases} 1/T^2 & \text{pour } x \in [0, T] \text{ et } y \in [0, T] \\ 0 & \text{ailleurs} \end{cases}$$

$$p_X(x) = \int_0^T p_{X,Y}(x,y) dy = \int_0^T 1/T^2 dy = 1/T = p_Y(y)$$

Exemple 3 Soit une distribution bidimensionnelle définie par :

$$p_{X,Y}(x,y) = \begin{cases} 2 & \text{pour } 0 \leq x \leq 1 \quad 0 \leq y \leq x \\ 0 & \text{ailleurs} \end{cases}$$

$$\Rightarrow p_X(x) = \int_0^x p_{X,Y}(x,y) dy = \int_0^x 2 dy = 2x$$

$$p_Y(y) = \int_y^1 p_{X,Y}(x,y) dx = \int_y^1 2 dx = 2(1-y)$$

*Changements de variables* : Connaissant une variable  $X$  et sa loi  $p(x)$  et étant donné une autre variable  $Y$  s'écrivant en fonction de  $X$ . Il existe un moyen de déterminer  $p_Y(y)$  par le biais du Jacobien. Soit  $X$  une VA continue caractérisée par sa densité de probabilité  $p_X(x)$  alors la VA continue  $Y=f(X)$  a une densité de probabilité donnée par :

$$p_Y(y) = |J| [p_X(x)]_{x=f^{-1}(y)} = \left| \frac{\partial X}{\partial Y} \right| [p_X(x)]_{x=f^{-1}(y)}$$

On peut généraliser cela pour n VA. Soit  $(X_1, X_2, \dots, X_n)$  un vecteur de VA continue caractérisé par sa densité de probabilité  $p_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n)$ , alors le vecteur de VA continue  $(Y_1, Y_2, \dots, Y_n) = f(X_1, X_2, \dots, X_n)$  a une densité de probabilité donnée par :

$$p_Y(y_1, \dots, y_n) = |J| \left[ p_{X_1 \dots X_n}(x_1, \dots, x_n) \right]_{x_1, \dots, x_n = f^{-1}(y_1, \dots, y_n)}$$

$$\text{avec } J = \det \begin{bmatrix} \frac{\partial X_1}{\partial Y_1} & \dots & \frac{\partial X_1}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial X_n}{\partial Y_1} & \dots & \frac{\partial X_n}{\partial Y_n} \end{bmatrix}$$

Exemple 1 : Soit  $z = x + y$  où  $x, y$  sont deux variables aléatoires indépendantes. On cherche la densité de probabilité de  $z$  connaissant celles de  $x$  et  $y$  c'est-à-dire  $p(x)$  et  $p(y)$ .

Posons:  $z = x + y$  On a alors  $x = w$

$$\begin{aligned} w &= x & \text{d'où} \\ y &= z - w \end{aligned}$$

$$J = \det \begin{bmatrix} \frac{dx}{dw} & \frac{dx}{dz} \\ \frac{dy}{dw} & \frac{dy}{dz} \end{bmatrix} = \det \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = 1$$

$$\text{alors } p(w, z) = |J| p(x, y) \Big|_{x=f(w,z) \text{ et } y=f(w,z)} = 1 \cdot p_{xy}(w, z - w)$$

$$\text{ainsi, } p(z) = \int p_{xy}(x, z) dx = \int p_{xy}(x, z - x) dx = \int p_x(x) p_y(z - x) dx = p_x(z) * p_y(z)$$

Donc, la densité de probabilité de la somme de deux variables aléatoires indépendantes a pour densité de probabilité le produit de convolution de leurs densités de probabilité respectives (voir exo n°5 du TP n°1).

Remarque : Il faut distinguer le cas où la transformation entre  $X$  et  $Y$  est biunivoque de celui où elle ne l'est pas.

Soit la transformation  $Y = X^2$ ,  $X$  est une variable aléatoire de probabilité  $p(x)$  définie tout d'abord sur l'intervalle  $[0, 1]$ . Sur cet intervalle, la transformation est biunivoque.

$$\text{On a alors : } X = \sqrt{Y} \text{ d'où : } p(y) = \frac{1}{2\sqrt{y}} p(x = \sqrt{y})$$

Soit la même transformation définie, cette-fois-ci sur l'intervalle  $[-1, 1]$ . Sur cet intervalle, la transformation n'est pas biunivoque. On a alors :  $X = \sqrt{Y}$  sur  $[0, 1]$  et  $X = -\sqrt{Y}$  sur  $[-1, 0]$  d'où :

$$p(y) = \frac{1}{2\sqrt{y}} p(x = \sqrt{y}) + \frac{1}{2\sqrt{y}} p(x = -\sqrt{y})$$

*Covariance et coefficient de corrélation*

$$C_{x_1x_2} = \sigma_{x_1x_2} = E\{(x_1 - \mu_{x_1})(x_2 - \mu_{x_2})^*\} = E\{x_1 \cdot x_2^*\} - \mu_{x_1} \mu_{x_2}^* = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2^* p(x_1, x_2) dx_1 dx_2 - \mu_{x_1} \mu_{x_2}^* \text{ Si } X_1$$

et  $X_2$  sont indépendants  $\Rightarrow C_{x_1x_2} = 0$ . On dit que les variables sont décorréliées.

Le coefficient de corrélation est donné par :  $\rho_{x_1x_2} = C_{x_1x_2} / (\sigma_{x_1} \sigma_{x_2})$

Ce coefficient mesure le degré de dépendance linéaire entre  $X_1$  et  $X_2$ . Il est compris entre -1 et 1. Un signe positif indique que  $X_1$  et  $X_2$  varient dans le même sens et un signe négatif que les deux variables varient en sens contraire. S'il est nul,  $X_1$  et  $X_2$  sont décorréliées, c'est le cas quand elles sont indépendantes.

Remarque: Si les deux variables sont totalement indépendantes, alors leur corrélation est égale à 0. La réciproque est cependant fautive, car le coefficient de corrélation indique uniquement une dépendance linéaire.

Ainsi, soient deux variables aléatoires  $X = \sin\phi$  et  $Y = \cos\phi$  où  $\phi$  est une phase aléatoire uniformément distribuée dans  $[0, 2\pi]$ . Si on calcule l'espérance de  $X$  et  $Y$ , on trouve une valeur moyenne nulle (voir exo n°6 du TP n°1). De même, le calcul de la covariance nous donne une valeur nulle. Les deux variables sont donc non corrélées. Et pourtant, on a  $X^2 = 1 - Y^2$ , ce qui prouve qu'elles sont dépendantes. En effet, si  $X$  est connue,  $Y$  l'est aussi.

*Emploi de la notation matricielle* : Il est préférable d'utiliser une notation matricielle pour traiter des vecteurs aléatoires de plus de 2 dimensions.

Soit  $\underline{X}$  le vecteur des v.a.,  $\underline{\mu}_X$  le vecteur des moyennes  $\mu_{X_i}$  et  $\underline{X}_C$  le vecteur des v.a. centrées :

$$\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \quad \underline{\mu}_X = \begin{bmatrix} \mu_{X_1} \\ \vdots \\ \mu_{X_n} \end{bmatrix}, \quad \underline{X}_C = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} - \begin{bmatrix} \mu_{X_1} \\ \vdots \\ \mu_{X_n} \end{bmatrix}$$

La matrice de covariance  $C_X$  est donnée par :

$$C_X = E\{\underline{X}_C \cdot \underline{X}_C^T\} = \begin{bmatrix} \sigma_1^2 & C_{12} & \cdot & \cdot & C_{1n} \\ C_{21} & \sigma_2^2 & \cdot & \cdot & C_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ C_{n1} & C_{n2} & \cdot & \cdot & C_{nn} \end{bmatrix}$$

Cas d'un vecteur aléatoire Gaussien : Un vecteur aléatoire réel de dimension  $n$  ( $X_1, \dots, X_n$ ) est un vecteur gaussien si toute combinaison linéaire de  $X_1, \dots, X_n$  est une variable aléatoire gaussienne réelle.

$$p(x_1, x_2, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det(C_X)}} e^{-\frac{1}{2}(\underline{x} - \underline{\mu}_X)^T C_X^{-1} (\underline{x} - \underline{\mu}_X)}$$

Exemple d'application: Soient deux variables aléatoires Gaussiennes  $X$  et  $Y$  décorrélées, montrer que dans ce cas elles sont indépendantes.

Par ailleurs et par définition d'un vecteur aléatoire gaussien, la variable aléatoire :

$$Y = \sum_{k=1}^n \alpha_k X_k = \alpha^T \underline{X}$$

est une variable aléatoire réelle Gaussienne. Par conséquent, sa loi est complètement déterminée par sa moyenne et sa variance qui ont pour expressions respectives :

$$E\{y\} = \sum_{k=1}^n \alpha_k E\{X_k\} = \alpha^T \underline{\mu}_X \text{ et } \sigma_y = \sum_{j,k=1}^n \alpha_j \alpha_k \text{cov}\{X_j, X_k\} = \alpha^T C_X \alpha$$

Exemple :

Si  $Z_1$  et  $Z_2$  sont des variables aléatoires indépendantes de loi normale de paramètres respectifs  $(\mu_1, \sigma_1)$  et  $(\mu_2, \sigma_2)$ , alors, quels que soient les réels  $a$  et  $b$ ,  $aZ_1 + bZ_2$  est de loi normale de paramètres  $(a\mu_1 + b\mu_2, \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2})$



## TD n°1: Variables et Vecteurs aléatoires

1. Une v.a a pour densité de probabilité  $p_X(x)$  définie comme suit :  $p_X(x) = \begin{cases} K & x \in [a, b] \\ 0 & \text{ailleurs} \end{cases}$

- Trouver la valeur de K
- Calculer la fonction de répartition  $F_X(x)$  et la tracer pour  $a=1$  et  $b=2$ .

2. Soit X une v.a discrète pouvant prendre les six valeurs suivantes : 1,2,3,4,5,6 avec la même probabilité  $p(X=i)$  avec  $i=1,2,3,\dots,6$ .

- Calculer  $p(X=i)$
- Donner l'expression de la fonction de répétition  $F_X(x)$  et la tracer.
- Donner l'expression de la densité de probabilité  $p_X(x)$  et la tracer.

3. Soit X une v.a ayant une densité de probabilité  $p_X(x)$  telle que :  $p_X(x) = ae^{-b|x|}$  pour  $-\infty < x < +\infty$

- Trouver la relation entre a et b
- Calculer  $E[X]$  et  $\text{Var}[X]$

4. Soient X une v.a et  $p(x)$  sa densité de probabilité définie comme suit :

$$p(x) = \begin{cases} Ke^{-x} & x \geq 0 \\ 0 & \text{ailleurs} \end{cases} \quad K \text{ est une constante réelle.}$$

- Déterminer la constante K
- Calculer  $E[X]$  et  $\text{Var}[X]$
- Calculer  $p(X<1)$ ,  $p(X<2)$  et  $p(X>2)$ .

5. La densité de probabilité conjointe de deux v.a est donnée par :  $P_{xy}(x,y) = \begin{cases} e^{-x}e^{-y} & x \geq 0 \text{ et } y \geq 0 \\ 0 & \text{ailleurs} \end{cases}$

- Calculer la densité de probabilité marginale de X
- Calculer la densité de probabilité marginale de Y
- Vérifier que  $p_{xy}(x,y)$  est une densité de probabilité
- Confirmer l'indépendance de X et Y par le calcul de densité de probabilité

6. Soit X une v.a ayant une densité de probabilité uniforme pour  $0 \leq x \leq 1$  et nulle ailleurs ; et soit Y une v.a indépendante de X et de densité de probabilité uniforme pour  $0 \leq y \leq 2$  et nulle ailleurs. On considère  $Z=X+Y$ . Calculer  $p_X(x)$ ,  $p_Y(y)$  et  $p_Z(z)$ .

7. Soient deux v.a réels, indépendantes ayant des densités de probabilités de Gauss, centrées et de variance unité. On considère deux v.a W et Z définies par :  $w = \frac{X}{\sqrt{X^2 + Y^2}}$  et  $Z = \sqrt{X^2 + Y^2}$

Déterminer la densité de probabilité conjointe de W et Z

Déterminer les lois de W et Z

Ces variables sont-elles indépendantes ? On donne :  $\int \frac{1}{\sqrt{1-x^2}} dx = \arcsin(x)$

8. Soit  $\phi$  une phase aléatoire équirépartie sur  $[0, 2\pi]$ . Considérons les variables aléatoires x et y définies par  $x = \cos\phi$  et  $y = \sin\phi$ .

Calculer les moyennes  $E[x]$  et  $E[y]$  puis calculer les variances de x et y.

Calculer finalement la covariance de x et y donnée par  $C_{xy} = \sigma_{xy} = E\{(x - \mu_x)(y - \mu_y)\} = E\{xy\} - \mu_x \mu_y$

## Solutions

1.  $K=1/(b-a)$ ;  $F(x)$  est une droite de 1 à 2 de pente 1.

2.  $p(x_i)=1/6$   $F(-\infty)=0$ ,  $F(1)=1/6$ ,  $F(2)=2/6$ ,  $F(3)=3/6$ ,  $F(4)=4/6$ ,  $F(5)=5/6$ ,  $F(6)=1$

3.  $b=2a, \mu_x=0, \sigma_x^2=$       4.  $K=1, \mu_x=1, \sigma_x^2=1, p(X<1)=1-1/e, p(X<2)=1-1/e^2, p(X>2)=1/e^2$   
 5.  $p(x) = e^{-x} \dots x \geq 0 \dots p(y) = e^{-y} \dots y \geq 0 \dots \iint p(x,y) dx dy = 1 \dots p(x,y)=p(x).p(y)$   
 6.  $p(x)=1, p(y)=1/2, p(z)=1/2z$  entre 0 et 1,  $p(z)=1/2$  entre 1 et 2,  $p(z)=1/2(3-z)$  entre 2 et 3  
 7.  $p(z)=2ze^{-z/2}, p(w)=1/(\pi(1-w^2)^{1/2})$  indépendants 8.  $\mu_x=0$  et  $\mu_y=0 \dots \sigma_x^2=0.5$  et  $\sigma_y^2=0.5 \dots C_{x,y}=0$

### Exercices supplémentaires

1. Soit la fonction  $f$  définie comme suit :  $p(x) = \begin{cases} 0 & \dots x < -1 \\ x+1 & \dots x \in [-1,0] \\ -x+1 & \dots x \in [0,1] \\ 0 & \dots x > 1 \end{cases}$

- Montrer que la fonction  $f$  est une densité de probabilité puis déterminer sa fonction de répartition
- Déterminer la moyenne et la variance d'une v.a aléatoire ayant  $f(x)$  pour densité de probabilité

2. Soit la d.d.p conjointe de deux v.a  $X$  et  $Y$  définie par :

$$p_{XY}(x,y) = \begin{cases} \alpha(x+y) & \dots si \dots 0 \leq x \leq 1 \dots et \dots 0 \leq y \leq 1 \\ 0 & \dots ailleurs \end{cases}$$

- Trouver la valeur de  $\alpha$  pour que  $P_{XY}(x,y)$  soit une d.d.p
- Les v.a  $X$  et  $Y$  sont-elles indépendantes ?

3. Considérons la variable aléatoire  $X$  dont la densité de probabilité est :

$$p_x(x) = \begin{cases} A(1+x) & \dots -1 < x \leq 0 \\ A(1-x) & \dots 0 < x < 1 \\ 0 & \dots ailleurs \end{cases}$$

- Déterminer  $A$  et tracer  $p_x(x)$  puis déterminer et tracer  $F_x(x)$
- Déterminer le point  $b$  tel que :  $P[X>b] = \frac{1}{2} P[X \leq b]$

4. Considérons la variable aléatoire  $X$  dont la densité de probabilité :  $p_x(x) = \begin{cases} 0 & \dots x < 0 \\ ce^{-2x} & \dots x \geq 0 \end{cases}$

- Déterminer  $c$ .
- Soit  $a > 0, x > 0$ , déterminer  $P[X \geq x+a]$

5. Soient  $X$  et  $Y$  deux v.a liées par la relation :  $Y = aX+b$  où  $a$  et  $b$  sont deux nombres réels. En supposant que  $F_X(x)$  et  $p_X(x)$  sont connus. Calculer alors  $F_Y(y)$  et  $p_Y(y)$ .

6. Soient  $X$  et  $Y$  deux v.a indépendantes définies par :

$$p_X(x) = \begin{cases} \exp(-x) & \dots si \dots x \geq 0 \\ 0 & \dots sin on \end{cases} \quad p_Y(y) = \begin{cases} 2 \exp(-2y) & \dots si \dots y \geq 0 \\ 0 & \dots sin on \end{cases}$$

On définit la v.a :  $Z=X/Y$ . Calculer  $p_Z(z)$ .

7. Soient deux v.a  $X$  et  $Y$  liée par la relation :  $Y = g(X) = X^2$ . Calculer  $F_Y(y)$  et  $p_Y(y)$  sachant que  $p_X(x)$  a une distribution uniforme sur  $[0,2\pi]$ .

8. La densité de probabilité d'une variable aléatoire  $X$  est donnée par :  $p_x(X) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-(X-\mu_x)^2/2\sigma_x^2}$ ,

soit  $Y = aX + b$  et  $Z = cX^2 + d$ . Calculer les moyennes  $\mu_Y$  et  $\mu_Z$  et les variables  $\sigma_Y^2$  et  $\sigma_Z^2$  puis déterminer les densités de probabilité de  $Y$  et de  $Z$

## TP n°1 : Variables et Vecteurs Aléatoires

### Rappels

Un générateur de nombres aléatoires dans l'intervalle  $[0, 1]$  est une fonction `rand` qui vérifie les deux propriétés suivantes :

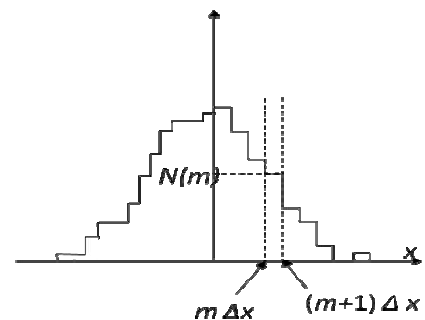
- un appel à la fonction **rand** donne une réalisation d'une variable aléatoire de loi uniforme sur  $[0, 1]$ .

Pour générer une loi uniforme sur  $[a, b]$ , on écrit  $a + (b-a) \cdot \text{rand}(N, 1)$

- les appels successifs à la fonction `rand` fournissent une réalisation d'une suite de variables aléatoires indépendantes.

Pour générer une loi Gaussienne de moyenne  $M$  et variance  $V$ , on écrit :  $M + \sqrt{V} \cdot \text{randn}(N, 1)$

- **s=sum(X)** Calcule la somme des composantes de  $X$ .
- **m=mean(X)** Calcule la moyenne empirique de  $X$
- **v=var(X)** Calcule la variance empirique.
- **sigma=std(X)** Calcule l'écart-type empirique de  $X$
- **hist (X,NC)**: Trace un histogramme non normalisé de  $X$  sur  $NC$  intervalles. Soit une suite finie de points issus de la réalisation d'une VA  $X$  avec  $X_{\min}$ =valeur minimale de la suite de valeurs et  $X_{\max}$ =valeur maximale de la suite de valeurs. On partage le segment  $[X_{\min}, X_{\max}]$  en  $NC$  classes équidistantes et on compte le nombre de valeurs dans chaque classe, on a alors une estimation de la densité de probabilité:



### Remarques :

Sous Matlab au début d'un programme, on peut initialiser le générateur en lui fournissant la première valeur appelée la graine du générateur `rand('state',sum(100*clock))`. Sinon, à chaque fois que l'on démarre Matlab, c'est la même suite de valeurs qui sera donnée.

Matlab dispose d'une boîte à outils Statistics Toolbox qui contient des fonctions prédéfinies permettant de simuler des réalisations de variables de la plupart des lois usuelles. On sert de la fonction dont le nom est obtenu en ajoutant au nom de la distribution le suffixe `rnd` : `distrnd`. Ainsi, pour la loi exponentielle ce sera `exprnd`, pour poisson (`poissrnd`), etc.

### Exercice 1

On veut simuler le lancé d'un dé avec 60 jets. Donner la loi de probabilité puis la fonction de répartition et calculer la moyenne et la variance théoriques.

```
clc; clear all; close all;
N=60; NC=6;
X=1+6*rand(N,1); X=floor(X);
figure; plot(X); %plot(X, '*');
mm=mean(X); vv=var(X);
figure; hist(X,NC)
```

1. Prendre  $N=6$  puis 12 puis 6000 puis 60000 et commenter l'histogramme
2. Quelle est l'effet de l'augmentation de  $N$  sur la moyenne et la variance estimées.
3. Peut-on considérer que l'histogramme représente la densité de probabilité?
4. Utiliser `cumsum` pour tracer la fonction de répartition  $Y = \text{cumsum}(X)$ ; `figure; plot(Y);`

**Exercice 2**

En vous servant de `randn`, écrire un programme permettant de simuler une loi gaussienne représentant les notes de 1000 étudiants. Utiliser différentes valeurs de NC (10, 50 et 100) et commenter.

**Exercice 3**

On veut vérifier le théorème central limite. Pour cela, on crée une matrice A dont chaque ligne représente un échantillon d'une loi uniforme.

1. Créer la matrice  $A = \text{rand}(12, 1000)$ , et un vecteur x de taille 1000 dont les composantes représentent la moyenne de chaque colonne de A.
2. Vérifier que chaque ligne de A suit une loi uniforme en traçant leurs histogrammes sur 50 valeurs (NC=50)
3. Utiliser `hist(x, NC)` pour visualiser la loi de x. A quoi la loi de x ressemble-t-elle?

**Exercice 4**

Ecrire un programme de générer une loi X une v.a ayant une densité de probabilité uniforme pour  $0 \leq x \leq 1$  (N=10000) et nulle ailleurs ; et soit Y une v.a indépendante de X et de densité de probabilité uniforme pour  $0 \leq y \leq 1$  et nulle ailleurs. On considère la variable aléatoire  $Z = X + Y$ ,

1. Sachant que  $p_Z(z) = p_X(x) * p_Y(y)$ , montrer que  $p_Z(z)$  aura l'allure d'un triangle
2. Vérifier par `hist` que c'est le cas.

**Exercice 5**

On modélise fréquemment la durée de vie d'un composant électronique par une loi exponentielle. On suppose que la durée de vie, en jours, d'une ampoule, est une variable aléatoire de loi exponentielle de paramètre  $\lambda = 1/100$ .

1. Donner sa moyenne et sa variance théorique
2. Utiliser `X = exprnd(1/λ, N, 1)` pour générer cette loi. Tracer sa loi.
3. Utiliser `mean` pour déterminer la durée de vie moyenne d'une ampoule.

**Exercice 6**

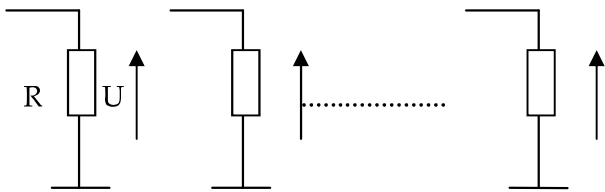
Soi  $\varphi$  une v.a. uniforme entre 0 et  $2\pi$  et soit  $X = \cos(\varphi)$  et  $Y = \sin(\varphi)$

1. Tracer les lois de X et Y et commenter
2. Calculer leurs moyennes et variances théoriques puis empiriques.
2. Sont-ils indépendants?
3. Utiliser `corr2` pour déterminer le coefficient de corrélation. Sont-ils décorrélés? commenter.
4. Reprendre ces questions en considérant  $Y = aX + b$  avec  $a = 2$  et  $b = 3$  puis  $a = -4$ . Conclure.

## Chapitre II : Processus aléatoires

Les processus aléatoires (stochastiques) décrivent l'évolution d'une grandeur aléatoire en fonction du temps (ou de l'espace). On peut définir un processus stochastique comme étant une famille de variables aléatoires indexées par le temps définies sur un même espace de probabilités  $\Omega$ . Un processus aléatoire ne possède pas de représentations temporelles analytiques. Chaque signal aléatoire observé représente une réalisation particulière de ce processus. Le signal de parole, le signal radar, l'électrocardiogramme, l'électro-encéphalogramme, les signaux sismiques en sont des exemples. Rappelons qu'un signal de réception -constitué d'un signal informatif (aléatoire ou déterministe), d'un signal aléatoire d'interférence et de bruit lié au canal de transmission- est aléatoire par nature.

**Exemple 1 :** Si l'on prend plusieurs résistances identiques (de même valeur) et que l'on mesure la tension. On trouvera une valeur non nulle due à l'agitation thermique des électrons libres dans la résistance. Les tensions mesurées au cours du temps sur l'ensemble des résistances fourniront plusieurs signaux aléatoires tous différents qui constituent le processus aléatoires.



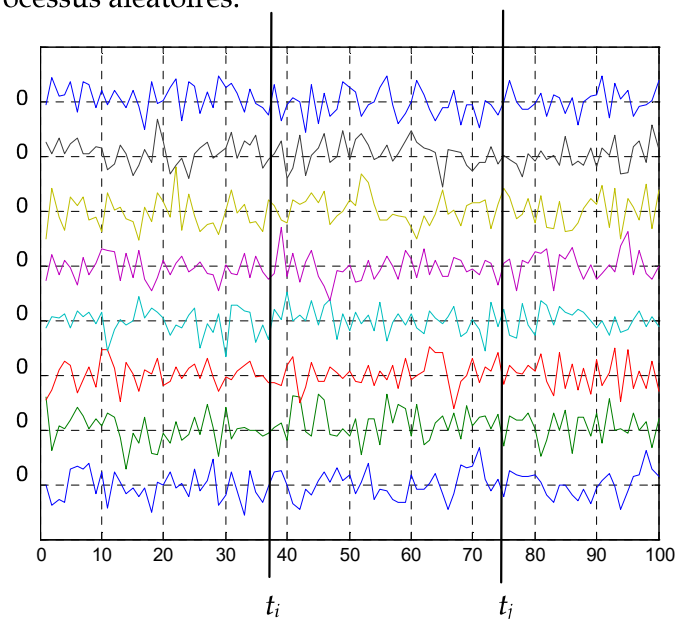
Signaux susceptibles d'être produits en mesurant la tension aux bornes de plusieurs résistances.

- Chaque tracé fournit un signal aléatoire

- à l'instant  $t_i$ , le processus se réduit à une v.a  $x_i$  dont la densité est  $p(x; t_i) = p(x_i)$

- Deux instants  $t_i$  et  $t_j$  permettent de définir deux variables aléatoires

$x_i$  et  $x_j$  on peut définir des densités de probabilités conjointes par:  $p(x; t_i, t_j) = p(x_i, x_j)$

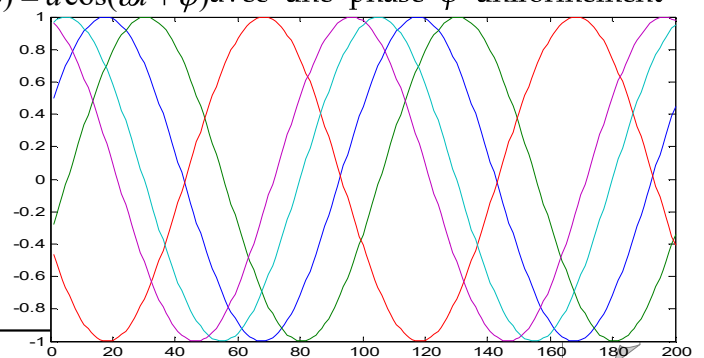


**Exemple 2 :** Signal sinusoïdal à phase aléatoire  $X(t, \varphi) = a \cos(\omega t + \varphi)$  avec une phase  $\varphi$  uniformément répartie entre 0 et  $2\pi$

-  $\varphi$  est à valeur réelle continue

-  $X(t, \varphi)$  est à valeur continue et réelle

- Un signal particulier  $X(t, \varphi_i)$  est déterministe



En pratique, il n'est pas aisé d'obtenir la densité de probabilité  $p(x; t_i) = p(x_i)$ , on se contente alors des moments d'ordre 1 et 2

### 1. Espérances mathématiques

Statistiques de 1<sup>er</sup> ordre: Elles permettent de décrire le signal à un instant donné. Le processus aléatoire devient une simple variable aléatoire que l'on peut décrire à l'aide de moments à condition que sa probabilité soit connue  $\forall t_i$ .

- *moyenne statistique*:  $\mu_x(t_i) = E\{x(t_i)\} = E\{x_i\} = \int_{-\infty}^{+\infty} x \cdot p(x, t_i) dx$

- *variance*:  $\sigma_x^2(t_i) = E\{x_i - \mu_x(t_i)\}^2 = E\{x_i^2\} - \mu_x^2(t_i)$

où  $E\{x_i^2\} = \int_{-\infty}^{+\infty} x^2 \cdot p(x, t_i) dx$  est la valeur quadratique moyenne (on parle aussi de puissance) et  $\sigma_x(t_i)$  est dit écart-type

Exemple 1: Soit le processus stochastique  $x(t)$  défini par  $x(t) = a + b \cdot t$  où  $a$  et  $b$  sont des v.a. dont les probabilités sont connues, alors :

$$\mu_x(t) = E[a + b t_i] = \mu_a + \mu_b t$$

Reprenons l'exemple 2: Pour un instant donné  $t_k$ , on peut calculer des moments statistiques de la variable aléatoire  $X(t_k, u)$

- Moyenne :

$$\mu_x(t) = E[X(t)] = \int_0^{2\pi} f_\varphi(\varphi) a \cos(\omega t + \varphi) d\varphi = \frac{1}{2\pi} a \int_0^{2\pi} \cos(\omega t + \varphi) d\varphi = 0$$

-Variance :

$$\sigma_x^2(t) = E[X^2(t)] - \mu_x^2(t) = E[X^2(t)] = \int_0^{2\pi} f_\varphi(\varphi) a^2 \cos^2(\omega t + \varphi) d\varphi = \frac{a^2}{2}$$

Supposons, maintenant, que  $\varphi$  est déterministe et que  $a$  est aléatoire de densité  $p(a)$ , de moyenne  $\mu_a$  et variance  $\sigma_a^2$ , alors :

$$\mu_x(t) = E[a \cos(\omega t + \varphi)] = \mu_a \cdot \cos(\omega t + \varphi) \quad \text{et} \quad \sigma_x^2(t) = \sigma_a^2 \cdot \cos^2(\omega t + \varphi)$$

Statistiques de 2<sup>ème</sup> ordre: On dit que le processus est connu à 2 instants  $t_1$  et  $t_2$ , si  $\forall t_1, t_2$ , la probabilité conjointe est connue. Dans la mesure où l'espérance fait intervenir le produit de deux variables aléatoires (2 instants), on parlera de statistiques d'ordre 2. on peut estimer les statistiques suivantes:

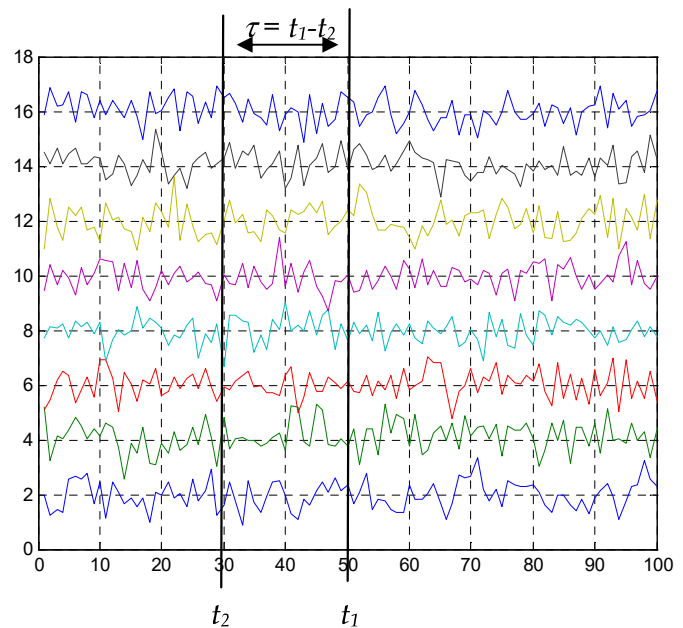
- Fonction d'autocorrélation statistique :

$$R_x(t_1, t_2) = E\{x(t_1) \cdot x^*(t_2)\} = E\{x_1 \cdot x_2^*\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 \cdot x_2^* \cdot p(x_1, x_2; t_1, t_2) \cdot dx_1 \cdot dx_2$$

- Fonction d'autocovariance statistique :

$$C_x(t_1, t_2) = E\{ [x(t_1) - \mu_x(t_1)] \cdot [x(t_2) - \mu_x(t_2)] \} = R_x(t_1, t_2) - \mu_x(t_1) \cdot \mu_x(t_2)$$

Elles illustrent la relation entre les statistiques prises à deux instants  $t_1$  et  $t_2$  différents. La fonction d'autocorrélation mesure la corrélation entre les signaux émis par un même processus à deux instants différents. Lorsqu'il y a corrélation on peut parler d'un "effet mémoire" du processus. On définit la mémoire du processus comme le temps  $t_c$  au delà duquel la corrélation est négligeable,  $t_c$  est aussi appelé temps de corrélation. Par contre, si elle est nulle, le signal est complètement aléatoire et le signal  $x(t)$  à l'instant  $t$  est complètement décorrélé de ce même signal à l'instant passé  $t - \tau$ .



Notons que pour des processus centrés ( $\mu_x(t)=0$ ):  $C_x(t_1, t_2) = R_x(t_1, t_2)$

On peut étendre ces notions pour mesurer ce lien entre deux processus aléatoires par :

- Fonction d'intercorrélation statistique:  $R_{xy}(t_1, t_2) = E\{ x(t_1) \cdot y^*(t_2) \} = E\{ x_1 \cdot y_2^* \}$

- Fonction d'intercovariance statistique :  $C_{xy}(t_1, t_2) = R_{xy}(t_1, t_2) - \mu_x(t_1) \cdot \mu_y^*(t_2)$

- Le coefficient de corrélation est définie par :  $\rho_{xy}(t_1, t_2) = \frac{C_{xy}(t_1, t_2)}{\sigma_x(t_1)\sigma_y(t_2)} \quad -1 \leq \rho_{xy}(t_1, t_2) \leq 1$

Reprenons l'exemple 2 et calculons l'auto-corrélation statistique :

$$R_{xx}(t_1, t_2) = E[a \cos(\omega t_1 + \varphi) a \cos(\omega t_2 + \varphi)] = E\left[\frac{a^2}{2} (\cos(\omega(t_1 + t_2) + 2\varphi) + \cos(\omega(t_1 - t_2)))\right] = \frac{a^2}{2} \cos(\omega(t_1 - t_2))$$

On remarque que c'est une fonction périodique, ne dépendant que de l'écart  $t_1 - t_2$ . Pour  $t_1 = t_2$ , on retrouve la variance du processus aléatoire soit  $a^2/2$ .

### 3. Processus stationnaires

Beaucoup de processus aléatoires observés en pratique ont des propriétés statistiques qui ne dépendent pas du temps ou l'observation est faite. On dit que ce sont des processus stationnaires.

On désigne par *stationnaire* les processus dont les caractéristiques statistiques sont indépendantes de l'origine du temps. Un processus aléatoire est dit stationnaire au sens strict lorsque toutes ces caractéristiques statistiques c'est à dire tous ses moments à tout ordre sont indépendants de l'origine du temps.



$$\begin{aligned} \Leftrightarrow p(x, t_i) &= p(x); \\ \mu_x(t_i) &= \mu_x \\ R_{xy}(t_1, t_2) &= R_{xy}(\tau) \text{ avec } \tau = t_2 - t_1 \\ &\dots \end{aligned}$$

La stationnarité ne signifie pas pour autant que le processus est indépendant du temps mais plutôt que ses propriétés statistiques ne dépendent pas du moment auquel on commence à les estimer. Ainsi, la température est un processus non stationnaire alors que le lancé de dé est un processus stationnaire.

On peut considérer que de nombreux phénomènes sont approximativement stationnaires sur des durées d'observation finies. Comme nous n'avons pas toujours accès à  $p(x, t_i)$ , on se contentera de la *stationnarité au sens large* (SSL) lorsque *seul*  $\mu_x$  et  $R_x$  sont indépendants de  $t$ .

Ainsi :

- Un processus stationnaire est dit stationnaire d'ordre 1 si sa moyenne et sa variance sont constantes et donc indépendantes de tout décalage temporelle :

$$\mu_x(t) = \mu_x = \text{cste} \quad \text{et} \quad \sigma_x^2(t_i) = \sigma_x^2$$

- Un processus stationnaire est dit stationnaire d'ordre 2 si ses statistiques d'ordre 2 ne dépendent que de l'écart temporel entre les deux instants  $t_1$  et  $t_2$ :

$$R_x(t_1, t_2) = R_x(\tau) \text{ avec } \tau = t_2 - t_1$$

Reprenons l'exemple précédent  $X(t, u) = a \cos(\omega t + u)$  et étudions sa stationnarité au sens large lorsque  $u$  est uniforme sur  $[0, 2\pi]$  : la moyenne, la variance sont indépendantes de  $t$  et l'auto-corrélation statistique ne dépend que de  $t_1 - t_2$ . Le processus est donc stationnaire au sens large.

### Exemple d'application:

Montrer que le processus  $x(t) = a \cdot \cos(2\pi t) + b \cdot \sin(2\pi t)$  où  $a, b$  sont des v.a décorréliées de moyenne nulle et de variance unité est stationnaire au sens large.

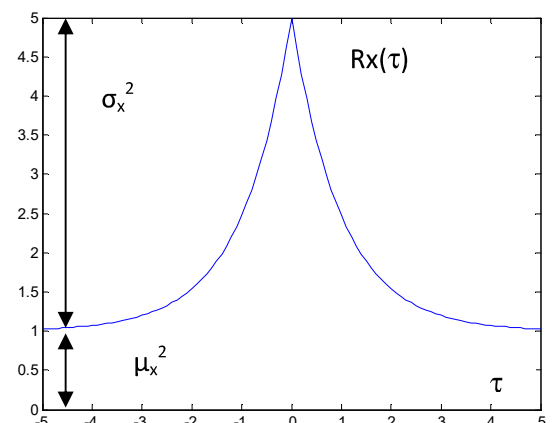
### Propriétés de la fonction d'autocorrélation pour $x$ aléatoire réel SSL

- La fonction de corrélation est paire :  $R_x(\tau) = R_x(-\tau)$  avec  $\tau = t_2 - t_1$

- La fonction est maximum à l'origine  $R_x(0) \geq |R_x(\tau)|$

- Valeurs à l'origine et à l'infini :  $R_x(0) = E(x(t)^2) = \mu_x^2 + \sigma_x^2$

$$\text{et} \quad \lim_{\tau \rightarrow \infty} R_x(\tau) = \lim_{\tau \rightarrow \infty} E\{x(t)x(t+\tau)\} = \mu_x^2$$



En effet, si  $x(t)$  ne contient ni composante périodique ni composante indépendante du temps, lorsque  $t \rightarrow \infty$ , les variables  $x(t)$  et  $x(t + \tau)$  deviennent statistiquement décorréliées.



Cas discret : On remplacera  $t$  par  $n$ , un processus aléatoire  $X$  discret SSL possèdera une moyenne constante et une autocorrélation statistique ne dépendant que de  $k=n_2-n_1$  soient :

$$\mu_x(n) = \mu_x = \text{cste} \quad \text{et} \quad R_x(n_1, n_2) = R_x(k) \quad \text{avec} \quad k = n_2 - n_1$$

$$R_x(n_1, n_2) = \begin{bmatrix} R_x(1,1) & R_x(1,2) & \dots & R_x(1,N) \\ R_x(2,1) & R_x(2,2) & \dots & R_x(2,N) \\ \dots & \dots & \dots & \dots \\ R_x(N,1) & R_x(N,2) & \dots & R_x(N,N) \end{bmatrix}$$

$$\Rightarrow R_x(n_1, n_2) = \begin{bmatrix} R_x(0) & R_x(1) & \dots & R_x(N-1) \\ R_x(1) & R_x(0) & \dots & R_x(N-2) \\ \dots & \dots & \dots & \dots \\ R_x(N-1) & R_x(N-2) & \dots & R_x(0) \end{bmatrix}$$

Ainsi, la matrice de corrélation d'un processus stationnaire discret est une matrice Toeplitz carrée. Une matrice carrée est dite Toeplitz si tous les éléments d'une même diagonale ou sous-diagonale sont égaux. On voit directement que c'est le cas ici. Cette propriété est directement liée à la propriété de stationnarité (au sens large) du processus.

Puissance et DSP : Les personnes qui conversent (dans un café ou en cours) génèrent un signal aléatoire qui selon son volume global possèdera une puissance. Si  $X(t)$  est SSL, on peut calculer l'espérance de la puissance instantanée par :

$$P_x = E(x(t)^2) = R_x(0) = \mu_x^2 + \sigma_x^2$$

L'énergie est alors donnée par :  $E_x = \int P_x dt = \int E(x(t)^2) dt$

Quant au spectre pour un signal aléatoire, il faut noter que chaque réalisation fournira un spectre différent. Or, nous avons déjà un outil statistique qui contient une information unique lorsque le processus est considéré SSL : c'est la fonction de corrélation statistique dont la TF fournira une information sur la distribution fréquentielle de la puissance moyenne du signal.

On définit, alors, la densité spectrale de puissance d'un signal aléatoire  $X(t)$  stationnaire au sens large comme la fonction de la fréquence  $f$  donnée par la TF de la fonction de corrélation statistique du signal :

$$S_x(f) = \int_{-\infty}^{+\infty} R_x(\tau) e^{-2\pi j f \tau} d\tau$$

La puissance moyenne totale du processus est :  $P_x = \int_{-\infty}^{+\infty} S_x(f) df = R_x(0)$

Comme le spectre représente une moyenne sur l'ensemble des réalisations possibles du processus, une réalisation particulière peut toujours avoir un spectre de puissance différent de  $S_x(f)$ .

Bruit blanc : Un bruit blanc est un signal aléatoire stationnaire au second ordre dont la densité spectrale de puissance est constante sur tout l'axe des fréquences.

$$R_x(\tau) = \sigma_x^2 \delta(\tau)$$

Il est défini par : ce qui implique que sa moyenne sera nulle ( $\mu_x=0$ ) et qu'il est décorrélé.

La TF de la corrélations statistique est, alors, donnée par  $S_x(f) = \sigma_x^2$

Dans le cas discret, on aura  $R_x(k) = \sigma_x^2 \delta(k) \Rightarrow R_x(k) = 0$  pour  $k \neq 0$  et  $\mu_x(k) = 0$

On dit alors que le bruit est décorrélé.

Par analogie avec la lumière blanche, on dénomme un tel signal bruit blanc. Il existe donc différents type de bruit blanc en fonction de la variable aléatoire décrivant le bruit. Les plus répandus sont le bruit blanc gaussien où la variable aléatoire suit une loi gaussienne et le bruit blanc uniforme où la variable aléatoire suit une loi uniforme.

#### 4. Processus ergodiques

Il n'est pas toujours possible de réaliser un nombre suffisant de mesures pour établir les propriétés statistiques d'un processus aléatoire. Il est plus facile de lancer un dé 1000 fois que de réquisitionner 1000 personnes pour lancer 1000 dés. Tout comme pour caractériser le bruit thermique, on réaliserait plusieurs mesures en usant de la même résistance au lieu d'en employer 1000. Cela signifie que l'on assimilera les résultats obtenus sur une réalisation à ceux obtenus pour un instant donné  $t_i$  pour différentes réalisations.

Un processus aléatoire stationnaire est dit *ergodique* lorsque les valeurs moyennes statistiques et temporelles sont identiques.  $\mu_x = \bar{x}$ ;  $R_x(\tau) = \varphi_x(\tau) \dots$

$$\text{où } \bar{x} = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \cdot dt \quad \text{et} \quad \varphi_{xy}(\tau) = \overline{x(t) \cdot y(t + \tau)}$$

D'un point de vue pratique lorsqu'un phénomène aléatoire est ergodique et stationnaire, on peut mesurer avec un seul appareil fiable à partir de n'importe quel instant.

Reprenons, une énième fois l'exemple 2 et calculons les moyennes temporelles :

$$E[X(t, u_i)] = \bar{X}(u_i) = \frac{1}{T} \int_{-T/2}^{+T/2} a \cos(\omega t + u_i) dt = 0 \quad E[X^2(t, u_i)] = \bar{X^2}(u_i) = \frac{1}{T} \int_{-T/2}^{+T/2} a^2 \cos^2(\omega t + u_i) dt = \frac{a^2}{2}$$

Il est donc ergodique d'ordre 1

Si le processus est ergodique  $S_x(f)$  représente le spectre de puissance de n'importe quelle réalisation  $x_i(t)$  :

$$P_x = \int_{-\infty}^{+\infty} S_x(f) df = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)^2 dt \quad \text{où } x_i(t) \text{ est une réalisation de } x(t).$$

Dans le cas où le signal  $x(t)$  est ergodique, stationnaire, on détermine  $S_x(f)$  par le calcul de la DSP temporel sur  $T$ :

$$S_x(f) = TF\{R_x(\tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} |X_T(f)|^2 \quad \text{Où } X_T(f) \text{ est la TF de } x(t) \text{ limité à l'intervalle } [0, T]$$

## TD n°2 : Processus aléatoires

1. Soit un processus aléatoire défini par :  $X(t) = a t + b$ ;  $t > 0$ ,  $b$  est une constante et  $a$  est une variable aléatoire qui a pour densité de probabilité :  $f_a(a) = \frac{1}{\sqrt{2\pi}} e^{-a^2/2}$

Calculer la densité de probabilité, la moyenne et la variance de  $X(t)$ .

2. On s'intéresse à la puissance  $W$  du bruit débité à l'instant  $t$ , pour une résistance  $R$  dont la tension  $V$  à ses bornes a un caractère aléatoire.

Déterminer la densité de probabilité de  $W$  dans chacun des cas.

- a)  $V$  est une gaussienne centrée de variance  $\sigma^2$
- b)  $V$  est uniformément distribuée entre  $5V$  et  $10V$

3. Soit le processus aléatoire  $X(t)$  défini par :  $X(t) = A \cos(\omega t + \varphi)$ .

Déterminer pour chacun des cas suivants les statistiques d'ordre 1 :  $p_X(x, t)$ ,  $\mu_X(t)$  et  $\sigma_X^2(t)$

- a) Si  $A$  est une v.a uniformément distribuée sur  $[-1, 1]$ .  $X(t)$  est-il stationnaire d'ordre 1?
- b) Si  $\varphi$  est une v.a uniformément distribuée sur  $[0, 2\pi]$ .  $X(t)$  est-il stationnaire d'ordre 1 et 2?

4. Soient  $X_1(t)$  et  $X_2(t)$  deux signaux aléatoires réels stationnaires du second ordre.

$$E\{X_1(t) X_1(t-\tau)\} = R_{X_1}(\tau) = A_1 e^{-|\tau|} + B_1 \quad \text{et} \quad E\{X_2(t) X_2(t-\tau)\} = R_{X_2}(\tau) = A_2 e^{-|\tau|} + B_2$$

- a) Interpréter  $A_i$  et  $B_i$ ,  $i=1, 2$

On considère  $X_c(t) = X_1(t) + X_2(t)$ , étudier la stationnarité de  $X_c(t)$  au SSL :

- b) Lorsque  $X_1(t)$  et  $X_2(t)$  sont indépendants.
- c) Lorsque  $X_2(t) = K X_1(t)$ .
- d) Mêmes questions pour  $Z(t) = X_1(t) - X_2(t)$ ,

5. Soit le processus aléatoire  $Z(t) = X \cos(2\pi f_0 t) - Y \sin(2\pi f_0 t)$ , où  $X, Y$  sont des variables aléatoires gaussiennes indépendantes à moyennes nulles et de variances  $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ ;  $f_0$  est une constante.

- a) Calculer la valeur moyenne  $E[Z(t)]$  et la variance  $\sigma_Z^2$ .
- b) Calculer la fonction d'autocorrélation  $R_Z(t_1, t_2)$  et examiner si le processus est stationnaire au sens large.

6. Considérons les deux variables aléatoires suivantes :  $A$  est une variable aléatoire gaussienne de moyenne nulle et de variance  $\sigma_A^2$ , et  $\varphi$  une variable aléatoire uniformément répartie entre  $-\pi$  et  $+\pi$ .

- a) Déterminer  $E[A^2]$ . Donner la densité de probabilité de  $\varphi$ . Soit  $B = e^{i\varphi}$  variable aléatoire complexe fonction de  $\varphi$ . Montrer que  $E[B] = 0$  et calculer  $E[B^2]$ .

Considérons la variable aléatoire  $C = \cos(\varphi + \theta)$  où  $\theta$  est une variable certaine.

- b) Montrer que  $E[C] = 0$ ,  $\forall \theta$ .
- c) Calculer  $E[\cos(2\varphi + \theta)]$ .
- d) En déduire  $E[C^2]$ .

On étudie maintenant le signal aléatoire  $X(t) = A \cos(2\pi f_0 t)$

- Donner trois trajectoires différentes du signal  $X(t)$
- Calculer la moyenne  $E[X(t)]$ , la variance  $\sigma_X^2(t)$  et la corrélation  $R_X(t_1, t_2)$ .  $X(t)$  est-il stationnaire d'ordre 1 ?
- Déterminer et tracer la transformée de Fourier de  $X(t)$ .  
Est-elle unique?
- Peut-on calculer la fonction de corrélation temporelle et la DSP correspondante.
- Mêmes questions mais pour le signal aléatoire  $X(t) = \cos(2\pi f_0 t + \varphi)$ .

7. Soit le signal  $z(t)$  résultant de la somme du signal  $x(t) = A \cos(2\pi f_0 t + \varphi)$  et du signal binaire aléatoire  $y(t)$  prenant la valeur 0 et  $V$  sur une durée  $T$  d'une façon équiprobable.  $x(t)$  et  $y(t)$  sont considérés décorrelés.

$\varphi$  : variable aléatoire uniforme entre  $[0, 2\pi]$ .

- Déterminer la fonction de répartition du signal  $x(t)$
- Sachant que la fonction d'autocorrélation de  $y(t)$  est  $A\Lambda_T(\tau) + B$ 
  - Déterminer  $A$  et  $B$
  - Déterminer la fonction d'autocorrélation de  $z(t)$  en fonction de celles de  $x(t)$  et  $y(t)$
  - Déterminer et tracer la fonction d'autocorrélation et la D.S.P du signal  $z(t)$
  - Déterminer la puissance de  $z(t)$ , sa variance et sa moyenne statistique Le processus  $z(t)$  est-il SSL ?

8. Un signal aléatoire  $z(t)$  est le résultat de la somme de deux signaux indépendants  $x(t)$  et  $y(t)$ .

Le signal  $x(t)$  est binaire antipolaire et prend la valeur  $A$  avec une probabilité égale à  $2/3$  et la valeur  $-A$  avec une probabilité égale à  $1/3$ .

Le signal  $y(t)$  est gaussien et possède une densité spectrale de puissance  $S_y(f) = \sigma^2/2 \cdot \text{tri}(f/B)$ .

- Déterminer la densité de probabilité du signal  $z(t)$
- La valeur moyenne et la variance de  $z(t)$  ainsi que le coefficient de corrélation de  $x$  et  $y$

## Solutions

1.  $X(t)$  Gaussienne avec  $\mu_x(t)=b$  et  $\sigma_x^2(t)=t^2$ .  $W=RI^2$  a)  $p(w) = \sqrt{\frac{R}{2\pi w \sigma^2}} e^{-\frac{wR}{2\sigma^2}}$  b)  $p(w) = \frac{1}{5} \sqrt{\frac{R}{w}}$

3.  $p(x;t) = \frac{1}{2|\cos(\omega t + \varphi)|}$  a)  $\mu_x(t) = 0$   $\sigma_x^2(t) = \frac{\cos^2(\omega t + \varphi)}{3}$  Non stat b)  $p(x;t) = \frac{1}{\pi \sqrt{A^2 - x^2(t)}}$

4.  $A_1 = \sigma_{x1}^2$   $B_1 = \mu_{x1}^2$  de même  $A_2 = \sigma_{x2}^2$   $B_2 = \mu_{x2}^2$  b)  $\mu_{xc}(t) = \sqrt{B_1} + \sqrt{B_2}$   $R_{xc}(\tau) = R_{x1}(\tau) + R_{x1}(\tau) + 2\sqrt{B_1 B_2}$   
SSL

c)  $\mu_{xc}(t) = (1+k)\sqrt{B_1}$   $R_{xc}(\tau) = R_{x1}(\tau)(1+2k+k^2)$  SSL d)  $\mu_{xc}(t) = \sqrt{B_1 B_2}$   $R_{xc}(\tau) = R_{x1}(\tau)R_{x2}(\tau)$  SSL

e)  $\mu_{xc}(t) = k(A_1 + B_1)$  pour  $R_{xc}(t, \tau)$  on ne peut conclure

5.  $\mu_z(t) = 0$   $\sigma_z^2(t) = \sigma^2$   $R_z(\tau) = \sigma^2 \cos(2\pi f_0 \tau)$  SSL

6.  $E[A^2] = \sigma_A^2 E[B^2] = 1$

$E[C^2] = 1$

$\mu_x(t) = \mu_A \cos(2\pi f_0 t)$   $\sigma_x^2(t) = \sigma_A^2 \cos(2\pi f_0 t)$   $R_z(\tau) = \sigma^2 \cos(2\pi f_0 t) \cos(2\pi f_0 (t - \tau))$

e)  $X(f) = \frac{A}{2} (\delta(f - f_0) + \delta(f + f_0))$   $R_x(\tau) = \frac{A}{2} \cos(2\pi f_0 \tau)$   $S_x(f) = \frac{A}{4} (\delta(f - f_0) + \delta(f + f_0))$  d) SSL et ergodique

7.  $p(x;t) = \frac{1}{\pi \sqrt{A^2 - x^2(t)}}$   $F(x;t) = \frac{1}{\pi} \arctg(x/A)$   $B=0.5$  V  $A=0.25$  V<sup>2</sup>  $P = R_z(0) = \frac{A^2}{2} + A$

$$R_z(\tau) = \frac{A^2}{2} \cos(2\pi f_0 \tau) + R_y(\tau) \quad S_z(f) = \frac{A^2}{2} (\delta(f - f_0) + \delta(f + f_0)) + AT \sin^2(fT) + B\delta(f)$$

$$8. \mu_x = A/3 \quad \sigma_x^2 = 8A^2/9 \quad R_y(\tau) = \frac{\sigma^2}{2B} \sin^2(B\tau) \quad \mu_z = A/3 \quad \sigma_z^2 = 8A^2/9 + \sigma^2/2B$$

$$p(z) = \frac{2}{3\sqrt{2\pi}\sigma_y} e^{-\frac{1}{2\sigma_y^2}(z-A)^2} + \frac{1}{3\sqrt{2\pi}\sigma_y} e^{-\frac{1}{2\sigma_y^2}(z+A)^2}$$

### Exercices supplémentaires

1. On considère le processus aléatoire  $Z(t) = X(t)Y(t)$  où  $X(t)$  et  $Y(t)$  sont deux processus aléatoires, statistiquement indépendantes qui ont respectivement pour fonction d'autocorrélation :  $R_x(\tau) = e^{-a|\tau|}$  et  $R_y(\tau) = e^{-b|\tau|}$ . Calculer la moyenne et la variance de  $Z(t)$  puis l'autocorrélation de  $Z(t)$

2.  $x(t)$  et  $y(t)$  sont deux signaux aléatoires supposés stationnaires individuellement et conjointement.

- Trouver la fonction d'autocovariance du signal  $z(t) = x(t) + y(t)$
- Trouver la fonction d'autocovariance du signal  $z(t)$  lorsque  $x(t)$  et  $y(t)$  sont non corrélés.
- Trouver la fonction d'autocovariance du signal  $z(t)$  lorsque  $x(t)$  et  $y(t)$  sont non corrélés et centrés.

3. Le signal aléatoire centré  $x(t)$  possède une fonction d'autocorrélation de la forme  $R_x(\tau) = \sigma_x^2 e^{-\beta|\tau|}$ . Un autre signal est lié à  $x(t)$  par l'équation déterministe suivante :  $y(t) = ax(t) + b$ , où  $a$  et  $b$  sont des constantes données.

- Quelle est la fonction d'autocorrélation de  $y(t)$  ?
- Quelle est la fonction d'intercorrrelation de  $x(t)$  et  $y(t)$  ?

4. Soit un processus aléatoire dont chaque réalisation est un signal constant :  $x(t) = a$  où  $a$  est une variable aléatoire (de moyenne  $\mu_a$  et de variance  $\sigma_a^2$ ) dont la densité de probabilité est indépendante du temps.

- Le processus est-il stationnaire ? Est-il ergodique ?

5. Soient  $x(t)$  et  $y(t)$  deux signaux aléatoires définis par :

$x(t) = \alpha$ , où  $\alpha$  est une variable aléatoire uniformément répartie sur l'intervalle  $[0, 1]$ .

$y(t) = e^{\beta t}$ , où  $\beta$  est une variable aléatoire gaussienne de moyenne nulle et de variance  $\sigma^2$ . Les deux signaux sont définis pour  $t \geq 0$  et  $\alpha$  et  $\beta$  sont des v.a. supposées indépendantes.

- Tracer quelques trajectoires des processus  $x(t)$  et  $y(t)$  puis calculer  $E\{x(t)\}$  et  $E\{y(t)\}$

Remarque : Pour  $y(t)$  utiliser les fonctions génératrices des moments d'une variable aléatoire  $\alpha$

gaussienne :  $E\{e^{\alpha t}\} = e^{\mu_\alpha t + \frac{\sigma_\alpha^2 t^2}{2}}$

- Calculer les fonctions d'autocorrélation de  $x(t)$  et de  $y(t)$ . En déduire les variances de  $x(t)$  et  $y(t)$ .
- Conclure sur la stationnarité de  $x(t)$  et  $y(t)$ .
- Calculer la fonction d'intercorrrelation de  $x(t)$  et  $y(t)$ .
- En déduire la fonction d'autocorrélation du signal  $z(t) = x(t) + y(t)$

6. Soit le signal aléatoire défini par :  $x(t) = A_1 e^{j2\pi f_1 t} + A_2 e^{j2\pi f_2 t}$  où  $A_1$  et  $A_2$  sont deux v.a. gaussiennes, décorrelées, centrées et de variance  $\sigma^2$

- Donner la d.d.p conjointe  $f_{A1,A2}(A1,A2)$
- Déterminer la moyenne statistique du signal  $x(t)$  ainsi que sa fonction d'autocorrélation et la D.S.P
- Déterminer la d.d.p de  $x(t)$
- Le processus  $x(t)$  est-il stationnaire au 2<sup>ème</sup> ordre ?

7. Soit le signal  $R(t)$  reçu à partir d'un radar, composé d'un signal utile  $U(t)$  et d'un bruit  $B(t)$  :

$$R(t) = U(t) + B(t)$$

$U(t)$  : signal aléatoire uniformément réparti sur  $[-\pi, \pi]$  et  $B(t)$  : bruit gaussien centré de variance  $\sigma^2 = 1$

- Calculer  $E\{U\}$ ,  $E\{B\}$  et  $E\{R\}$
- Déterminer la fonction d'autocorrélation de  $R(t)$  en fonction de celle de  $U(t)$  et  $B(t)$
- Trouver les deux cas pour lesquels  $R(t)$  est stationnaire au 2<sup>ème</sup> ordre (quelle hypothèses faut-il poser sur  $U(t)$  et  $B(t)$ )
- Pour  $U(t)$  et  $B(t)$  indépendants, déterminer  $S_R(f)$

8. Soient  $X_1(t)$  et  $X_2(t)$  deux signaux aléatoires stationnaires du second ordre. On considère  $Z(t) = X_1(t) + X_2(t)$ , a). Calculer  $R_z(t, \tau)$ . b). Si les deux signaux sont indépendants et qu'au moins l'un deux possède une moyenne nulle, donner l'expression de  $S_z(f)$ . c). Même question pour  $Z(t) = X_1(t) \cdot X_2(t)$

9. Soient  $X(t)$  et  $Y(t)$  deux signaux aléatoires stationnaires et conjointement stationnaires. On considère  $Z(t) = X(t) + Y(t)$ ,

- Calculer  $S_z(f)$ .
- Les deux signaux sont supposés indépendants avec  $R_X(\tau) = e^{-a|\tau|}$  et  $R_Y(\tau) = b \delta(\tau)$ , calculer  $S_z(f)$ .

Si  $X(t) = A \cos(2\pi f_0 t + \varphi)$  avec  $\varphi$  une v.a. équirépartie sur  $[0, 2\pi]$ , calculer  $S_z(f)$

## Solutions

1.

## TP n° 2 : Processus aléatoires (Stationnarité et Ergodisme)

- But :1. Manipuler des signaux aléatoires, les caractériser grâce à leurs moments d'ordre 1 (moyenne, variance) et d'ordre 2 (autocorrélation, covariance).
2. Aborder et acquérir les notions de stationnarité et d'ergodicité.
  3. Comprendre le rôle d'une matrice variance covariance.

### Exercice 1:

```
%% Processus Aléatoire
clc; clear all; close all;
N_realis=1000; N_va=1000; A=1; f0=0.0025;
phi=2*pi*rand(N_realis,1);
for i=1:N_realis
    for j=1:N_va
        X(i,j)=A*cos(2*pi*f0*j+phi(i));
    end
end
t=(1:N_va);
figure (1);
plot(t,X(1,:),t,X(2,:),t,X(3,:),t,X(4,:),t,X(5,:),t,X(6,:))

%% Stationnarité d'ordre 1
moy=mean(X,1);
v=var(X,1,1);
figure (2); subplot (1,2,1); plot(moy);AXIS([1 N_va -1 1])
subplot (1,2,2); plot(v);
AXIS([1 N_va 0 1]);

%% Stationnarité d'ordre 2
RX=cov(X);
figure (3);
plot(t,RX(1,:),t,RX(4,:),t,RX(7,:),t,RX(10,:),t,RX(13,:),t,RX(16,:))

%% Ergodisme d'ordre 1 et 2

% d'ordre 1
moy_t=mean(X(1,:));
v_t=var(X(1,:));

% d'ordre 2
tt=(-N_va+1:N_va-1);
RXt=xcorr(X(1,:));
figure (4);
plot(tt,RXt);
```

1. Quelle est la différence entre les 6 réalisations affichées.
2. A partir d'une visualisation peut-on prévoir la stationnarité ou l'ergodicité?
3. Ce processus est-il stationnaire d'ordre 1 ? Justifier
4. Calculer l'auto-corrélation statistique théorique puis commenter la figure 3.
5. Utiliser le workspace pour visualiser RX, ce processus est-il stationnaire d'ordre 2 ? Justifier.
6. Comparer les moments statistiques et temporeux et conclure. En déduire la DSP de X.

### Exercice 2

On considère le processus aléatoire  $X(t) = A \cos(\omega t + \phi)$  considérant  $\phi$  cste et A uniforme entre -1 et 1.

1. Simuler le processus pour  $N\_realis=100$ ;  $N\_va=1000$
2. Est-il stationnaire d'ordre 1 ? (Vérifier théoriquement puis avec matlab)
3. Est-il stationnaire d'ordre 2 ? (Visualiser  $RX(1,:)$ ,  $RX(20,:)$ ,  $RX(30,:)$ ) et utiliser le workspace.
4. Peut-on calculer sa DSP statistique?



**Exercice 3**

On suppose le processus suivant  $b$  est ergodique d'ordre 2 et on se contente de déduire les propriétés statistiques des propriétés temporelles en étudiant une réalisation unique  $bb$ .

```
clc; clear all; close all;
moy=0; var=1; N=1024; M =100;
b=moy+sqrt(var)*randn(M,N);
bb=b(1,:);
% Afficher la corrélation Sigtab et la dsp du bruit blanc Sbb(f)
RX=xcorr(bb)/N;t=(-N+1:N-1);
subplot(2,1,1); plot(t,RX);
legend('Autocor du bruit blanc'); xlabel('Temps(s)'); ylabel('Amplitude');
Fb=fft(bb);Sbb=abs(Fb).^2/N;f=(-1/2:1/N:1/2-1/N);
subplot(2,1,2);plot(f,abs(Sbb));AXIS([min(f) max(f) 0 20]);
legend('Densité du bruit blanc'); xlabel('Frequence (Hz)');ylabel('Amplitude');
```

1. Commenter les graphes obtenus.
2. Pourquoi n'obtient-on pas une DSP cste?
3. Prendre une moyenne de 2 et une variance de 4 et retrouver ces caractéristiques à partir de l'autocorrélation

**Exercice 4**

Télécharger le fichier 'vous avez du courrier en attente.wav' et le placer dans le même répertoire que votre programme commençant come suit :

```
clc; clear all; close all;
nom_fich = uigetfile('*.wav', 'Selectionner le fichier son');
% Lire, écouter et afficher le son complet
[x,fe]=wavread(nom_fich);sound(x,fe); t=(0:length(x)-1)/fe; subplot(2,1,1);plot(t,x);
legend('Son');xlabel('Temps (s)');ylabel('Amplitude');
% afficher une partie du son qui correspond à une voyelle
N1=21000;N2=21500;x=x(N1:N2)
N=length(y);t=(N1:N2)/fe; subplot(2,1,2);plot(t,x);
legend('Voyelle sur 25 ms');xlabel('Temps (s)');ylabel('Amplitude');
% Calculer et afficher la DSP
Sx= abs(fft(x).^2)/N; f=(0:fe/N:fe/21/N);
figure; plot(f,Sx);
legend('Spectre'); xlabel('Frequence (Hz)'); ylabel('Amplitude (dB)');
```

1. Le signal étudié est un processus aléatoire ou une réalisation d'un signal aléatoire?
2. La DSP calculée est elle d'origine statistique ou temporelle?
3. Calculer et visualiser la DSP pour différentes parties du signal qui se chevauchent pour des durées de  $N=1000, 500, 250, 125$ . Pour quelle valeur de  $N$  peut-on considérer le processus stationnaire?
4. Expliciter l'importance des notions de stationnarité et d'ergodisme.

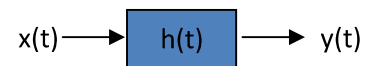


### Chapitre III : Filtrage linéaire des signaux aléatoires

Un signal aléatoire est amené à être transmis, analysé, transformé, etc. Conserve-t-il son caractère aléatoire? sa stationnarité ? Que deviennent sa moyenne et son auto-corrélation statistiques lors d'un filtrage linéaire. On examinera la transformation des caractéristiques du signal dans le domaine fréquentiel qui permettra d'aborder la notion de filtre formeur. Ces connaissances nous permettront d'envisager une application directe qui est le filtrage optimal et adapté.

#### 1. Processus aléatoires et SLIT

Soit un système linéaire et invariant dans le temps (SLIT) défini par sa réponse impulsionnelle  $h(t)$  ou sa fonction de transfert  $H(f)$  :



Rappelons que la réponse du système linéaire et invariant à un signal quelconque déterministe est donnée par :

$$y(t) = x(t) * h(t) = \int x(\tau) h(t - \tau) d\tau$$

Cette expression implique que si  $x(t)$  est un signal aléatoire, le signal en sortie  $y(t)$  est forcément un signal aléatoire puisque la sortie est une somme pondérée de l'entrée. Il faudra donc caractériser  $y(t)$  de manière statistique, de même que pour  $x(t)$ , en étudiant les grandeurs statistiques déjà vues au chapitre précédent.

#### Moyenne et autocorrélation statistiques de $y(t)$

Supposons que  $x(t)$  est un processus aléatoire, la moyenne  $\mu_y(t)$  du signal de sortie (aléatoire aussi) peut se calculer par :

$$\begin{aligned} \mu_y(t) &= E\{y(t)\} = E\{x(t) * h(t)\} = E\left\{\int h(\tau) x(t - \tau) d\tau\right\} \\ &= \int h(\tau) E\{x(t - \tau)\} d\tau = \int h(\tau) \mu_x(t - \tau) d\tau = h(t) * \mu_x(t) \end{aligned}$$

De la même façon, on peut calculer l'auto-corrélation statistique de  $y(t)$  soit  $R_y(t_1, t_2)$ :

$$\begin{aligned} R_y(t_1, t_2) &= E\{y(t_1) y^*(t_2)\} = E\{x(t_1) * h(t_1) \cdot x^*(t_2) * h^*(t_2)\} = E\left\{\int h(\tau_1) x(t_1 - \tau_1) d\tau_1 \int h^*(\tau_2) x^*(t_2 - \tau_2) d\tau_2\right\} \\ &= \iint h(\tau_1) h^*(\tau_2) E\{x(t_1 - \tau_1) x^*(t_2 - \tau_2)\} d\tau_1 d\tau_2 \end{aligned}$$

Si  $x(t)$  est stationnaire au sens large (SSL) alors:

- La moyenne statistique de  $y(t)$  devient :

$$\mu_y(t) = \int h(\tau) \mu_x(t - \tau) d\tau = \int h(\tau) \mu_x d\tau = \mu_x \int h(\tau) d\tau = \mu_x H(0) \quad (f = 0)$$

- La corrélation statistique de  $y$  devient :

$$\begin{aligned} R_y(t_1, t_2) &= \iint h(\tau_1) h^*(\tau_2) E\{x(t_1 - \tau_1) x^*(t_2 - \tau_2)\} d\tau_1 d\tau_2 \\ &= \iint h(\tau_1) h^*(\tau_2) R_x(\tau_2 - \tau_1) d\tau_1 d\tau_2 = \text{fct}(\tau) = R_y(\tau) \end{aligned}$$

Ainsi, si le signal d'entrée est stationnaire au sens large (SSL) alors le signal de sortie sera aussi SSL

### Densité spectrale

En appliquant la transformée de Fourier aux deux membres de l'équation précédente, la densité spectrale de puissance est obtenue :

$$\begin{aligned} S_y(f) &= TF(R_y(\tau)) = TF(E\{x(t_1) * h(t_1) \cdot x^*(t_2) * h^*(t_2)\}) = TF(E\{h(t_1) * x(t_1) \cdot x^*(t_2) * h^*(t_2)\}) \\ &= TF(h(t_1) * E\{x(t_1) \cdot x^*(t_2)\}) = TF(h(t_1) * R_x(t_1 - t_2) * h^*(t_2)) = |H(f)|^2 S_x(f) \end{aligned}$$

Ainsi, la densité spectrale de puissance du signal de sortie est égale à la densité spectrale de puissance du signal d'entrée multipliée par le carré du module de la réponse en fréquences du système (la phase de  $H(f)$  n'intervient pas). C'est une propriété très importante à la base de nombreuses applications dont notamment la notion de filtre formeur et le filtrage optimal.

Par ailleurs, si l'on souhaite calculer la corrélation du signal de sortie en évitant la lourdeur de calcul inhérente au produit de convolution, on calculera la densité spectrale de puissance du signal de sortie et on prendra la transformée de Fourier inverse.

$$R_y(\tau) = TF^{-1}(|H(f)|^2 S_x(f)) = \int |H(f)|^2 S_x(f) e^{2\pi i f \tau} df$$

La puissance moyenne du signal de sortie est alors obtenue par :  $E\{y(t)^2\} = R_y(0) = \int |H(f)|^2 S_x(f) df$

### Formule des interférences

Soient  $y_1(t)$  la sortie d'un système  $h_1(t)$  dont l'entrée est un signal aléatoire  $x_1(t)$  et  $y_2(t)$  la sortie d'un système  $h_2(t)$  dont l'entrée est un signal aléatoire  $x_2(t)$ . La formule de l'interférence permet de relier l'intercorrélation entre les sorties de deux filtres à celles des entrées de ces filtres:

$$S_{y_1 y_2}(f) = H_1(f) S_{x_1 x_2}(f) H_2^*(f)$$

On a aussi :  $R_{xy}(\tau) = R_x(\tau) * h^*(\tau) \Rightarrow S_{xy}(f) = S_x(f) H^*(f)$

et  $R_{yx}(\tau) = R_x(\tau) * h(\tau) \Rightarrow S_{yx}(f) = H(f) S_x(f)$

Exemple 1 : Soit le processus stochastique  $x(t)$  SSL de corrélation statistique :  $R_x(\tau) = \sigma^2 e^{-\frac{|\tau|}{\theta}}$  et le système linéaire et invariant de réponse impulsionnelle  $h(t) = 5e^{-2t}$ . On obtient alors:

$$-\mu_y(t) = \mu_x H(0) = 0 \times \frac{5}{2 + 2\pi j f} \Big|_{f=0} = 0 \times 5/2 = 0$$

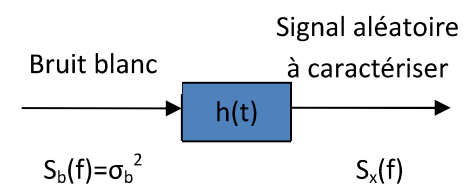
$$-S_y(f) = |H(f)|^2 S_x(f) = \left| \frac{5}{2 + 2\pi j f} \right|^2 \cdot TF \left( \sigma^2 e^{\frac{-|z|}{\theta}} \right) = \frac{5}{4 + (4\pi^2 f^2)} \cdot \frac{2\sigma^2 \theta}{1 + 4\pi^2 f^2 \theta^2}$$

### Exemple 2 :

Considérons par exemple le cas où le signal en entrée est un bruit blanc  $b(t)$ . Sa DSP est donc une fonction constante. Alors, en sortie, on aura un signal tel que :  $S_x(f) = \text{constante } |H(f)|^2$

### Notion de filtre formeur

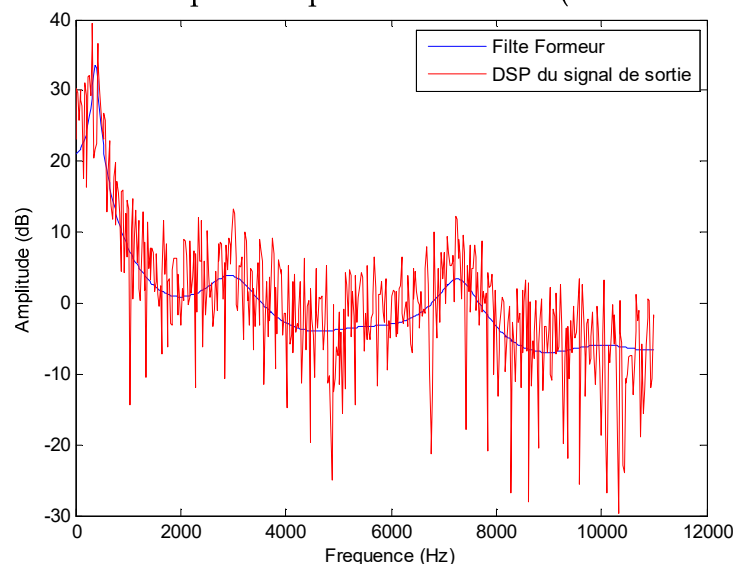
Supposons donné un signal aléatoire  $x(t)$ : On appelle filtre formeur de  $x(t)$ ; le filtre de fonction de transfert  $H(f)$ ; tel que  $x(t)$  est généré par passage d'un bruit blanc  $b(t)$  dans  $H(f)$ .



La détermination du filtre formeur s'effectue en inversant la formule précédente:

$$|H(f)|^2 = S_x(f) / \sigma_b^2$$

On pourra donc décrire le signal aléatoire par les paramètres du filtre et la variance du bruit blanc. En effet, si on fait passer le bruit blanc dans un filtre linéaire et stationnaire à paramètres ajustables et si on obtient le signal désiré (à modéliser) à la sortie du filtre, alors on peut dire que toute l'information spectrale est contenue dans le filtre représenté par ses coefficients (Voir exo 1 du TP n°3)



Remarque : Ainsi le bruit blanc joue pour l'aléatoire l'équivalent de la distribution de Dirac pour le déterministe.

Rappelons que le bruit blanc n'a pas d'existence physique car il serait de puissance infinie. Une approximation du bruit blanc à bande limitée appelé aussi bruit blanc coloré est défini par :

$$S_b(f) = \begin{cases} \sigma_b^2 & |f| < B \\ 0 & \text{ailleurs} \end{cases}$$

## 2. Application : Filtrage adapté et filtrage optimal

La transmission d'un signal s'accompagne de distorsions (dus aux milieux de transmission) qu'il serait souhaitable d'éliminer ou au moins d'atténuer avant tout traitement ultérieur. Si l'on connaît le signal (déterministe) d'origine, on parlera de détection. Dans le cas contraire ou si le signal est aléatoire on emploiera le terme estimation. Il existe de nombreuses approches de détection, on peut, entre autres procéder par filtrage qui permettra de rehausser le signal noyé dans le bruit.

Considérons un signal déterministe  $x(t)$  supposé connu, dont on souhaite tester la présence possible dans une observation  $s(t)$ . Le bruit d'observation est supposé quant à lui stationnaire SSL de densité spectrale  $S_b(f)$ . On cherche un filtre  $H(f)$  qui maximise le SNR à un instant précis  $T_0$ .

On suppose donc que le signal utile  $x(t)$  est noyé dans un bruit  $b(t)$  stationnaire SSL additif, d'où :

$$s(t) = x(t) + b(t)$$

On filtre le signal par un filtre linéaire dont la réponse impulsionnelle est  $h(t)$ . A la sortie du filtre, on obtient un signal :

$$y(t) = s(t) * h(t) = x(t) * h(t) + b(t) * h(t) = x_2(t) + b_2(t)$$

A la sortie du filtre et à l'instant  $T_0$ , le SNR s'écrit :  $SNR(T_0) = \frac{Puis(x_2(T_0))}{Puis(b_2(T_0))}$

Le but étant de trouver  $h(t)$ , on va alors exprimer le SNR en fonction de  $h(t)$  (ou  $H(f)$ ), ainsi :

- Le signal au numérateur  $x_2(t)$  est déterministe donc sa puissance s'exprime comme suit:

$$Puis(x_2(T_0)) = |x_2(T_0)|^2 = |TF^{-1}(X_2(f))|^2 = \left| \int X(f)H(f)e^{2\pi j f T_0} df \right|^2$$

- Le signal  $b_2(t)$  est issu du filtrage d'un signal aléatoire  $b(t)$  SSL donc il est aussi SSL, c'est ainsi que sa puissance pourra se formuler comme suit:

$$Puis(b_2(T_0)) = E\{b_2(T_0)^2\} = R_{b_2}(\tau = 0) = TF^{-1}(S_{b_2}(f))\bigg|_{\tau=0} = \int S_b(f)|H(f)|^2 df$$

Ainsi, on arrive à exprimer le SNR comme suit :

$$SNR(T_0) = \frac{Puis(x_2(T_0))}{Puis(b_2(T_0))} = \frac{\left| \int X(f)H(f)e^{2\pi j f T_0} df \right|^2}{\int S_b(f)|H(f)|^2 df}$$

Que l'on peut écrire aussi, comme étant :

$$SNR(T_0) = \frac{\left| \int a(f)b^*(f) df \right|^2}{\int a(f)a^*(f) df} \quad a(f) = \sqrt{S_b(f)}H(f) \text{ et } b(f) = X^*(f)e^{-2\pi j f T_0} / \sqrt{S_b(f)}$$

Pour trouver  $H(f)$  qui maximise ce rapport, on fait appel à l'inégalité de cauchy-schwartz:

$$\left| \int a(f)b^*(f) df \right|^2 \leq \int a(f)a^*(f) df \int b(f)b^*(f) df$$

Ce qui nous donne :  $SNR(T_0) = \frac{\left| \int a(f)b^*(f)df \right|^2}{\int a(f)a^*(f)df} \leq \int b(f)b^*(f)df$

avec égalité (soit le maximum) lorsque  $a(f)=kb(f) \Rightarrow$  Cette dernière nous fournit l'expression suivante du filtre optimal  $H(f)$  :

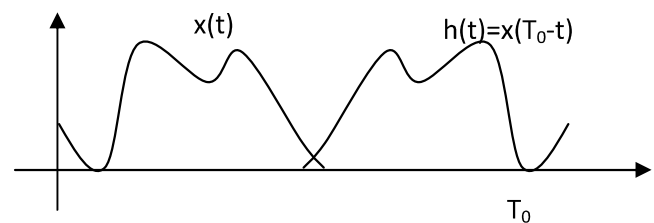
$$H(f)_{optimal} = k.X^*(f)e^{-2\pi jfT_0} / S_b(f)$$

L'expression du SNR devient :  $SNR(T_0)_{max} = \int \frac{|X(f)|^2}{S_b(f)} df$  (Ce maximum est indépendant de  $T_0$ )

Cas particulier : Par ailleurs, si le bruit  $b(n)$  est blanc, on parle de filtre adapté:

$$H(f)_{adapté} = k / \sigma_b^2 X^*(f)e^{-2\pi jfT_0} \text{ donnant } h(t)_{adapté} = k / \sigma_b^2 x^*(T_0 - t)$$

La réponse impulsionnelle du filtre représente le signal utile  $x(t)$  renversé et décalé de  $T_0$ . Le filtrage adapté revient à effectuer l'inter-corrélation entre l'observation et le signal à détecter.



Cette réponse n'est pas causale ce qui ne permet pas de l'appliquer en temps réel. Cependant, ce filtre peut-être appliqué à un signal sauvegardé.

### Exemple1 : Détection d'une impulsion

Considérons un système émettant une impulsion  $x(t)$  rectangulaire de durée  $T_0$  et d'amplitude  $A$ .  $\text{Rec}(t)$  étant la fonction porte égale à un pour  $t$  entre  $-1/2$  et  $1/2$  et nulle ailleurs. Du fait de la parité de cette fonction, le filtre aura la même expression et sera centré en  $T_0$ .

Exemple 2 : En sonar ou en radar, on cherche à localiser une « cible ». Cette cible peut être par exemple un bateau ou un avion. Pour cela, on procède de la façon suivante : on émet un signal  $x(t)$ , qui parcourt la distance  $d$  jusqu'à la cible, sur laquelle il sera réfléchi en direction d'un récepteur. Le récepteur reçoit alors le signal  $y(t)$  bruité, atténué (de  $a$ ) et retardé de  $T_{AR}$  (voir exo 2 du TP n°3)

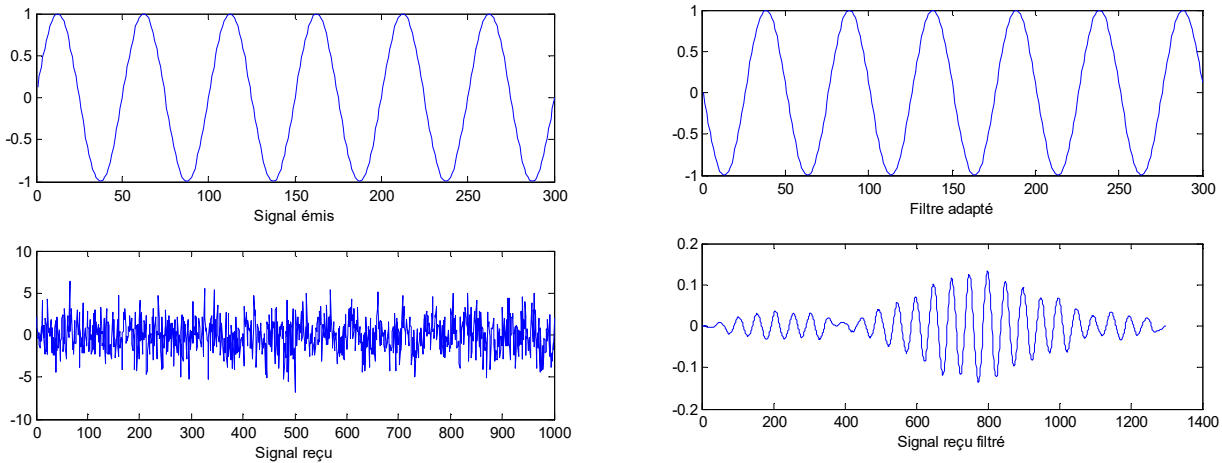
$$y(t) = a.x(t - T_{AR}) + b(t). \quad \text{où } T_{AR} \text{ correspondant au temps d'aller-retour } (T_{AR} = 2d/v)$$

On utilisera alors comme filtre adapté le filtre adapté à  $x(t)$ , soit  $h(t) = x^*(T_0 - t)$

La sortie de ce filtre sera l'intercorrélation des signaux  $y(t)$  et  $x(t)$ . On obtient alors :

$$R_{yx}(t - T_0) = a.R_{xx}(t - T_0 - T_{AR}) + R_{bx}(t - T_0).$$

Sachant que l'autocorrélation est maximale en 0, l'intercorrélation sera maximale pour  $t = T_0 + T_{AR}$  ce qui nous permettra de déterminer  $T_{AR}$ .



### Remarques

- Le choix du signal 'x' est important : Il est souhaitable que le max de sa fonction d'autocorrélation  $R_{xx}$  soit facile à identifier (pic bien visible comme pour la fonction triangle)
- Dans le cas où le signal est sinusoïdal, le filtre adapté est un passe bande idéal centré sur la fréquence du signal (détection synchrone).

### Exercice d'Application

Soit  $x(t) = 1 - t/T$  avec  $t \in [0, T]$

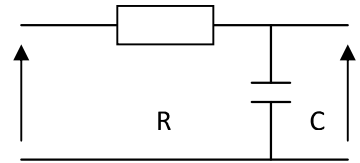
On utilise ce signal pour déterminer la distance d'un objet. Le signal reçu par le récepteur après réflexion sur l'objet est :  $y(t) = \alpha x(t - T_D) + b(t)$  où  $\alpha$  est une constante réelle positive et  $b(t)$  est un bruit blanc de densité spectrale de puissance  $\sigma^2$ . On veut maximiser le rapport signal sur bruit

1. Déterminer la réponse impulsionnelle du filtre  $h(t)$  telle que  $\int h(t)^2 dt = 1$  (prendre  $T_0 = T$ ).
2. Donner le rapport signal sur bruit après filtrage.
3. Exprimer l'inter-corrélation temporelle  $R_{yx}(t - T_0)$  en fonction de l'auto-corrélation de  $x(t)$  et de l'inter-corrélation des signaux  $b(t)$  et  $x(t)$
4. Comment déterminer le temps  $T_D$ ?
5. Pourquoi ne peut-on pas utiliser le filtrage optimal pour un signal déterministe dont on ignore l'expression?

## TD n° 3 : Processus aléatoires et Filtrage linéaire

1. L'entrée  $x(t)$  du circuit donnée ci-dessous est un bruit blanc de fonction d'autocorrélation  $R_x(\tau) = \sigma_x^2 \delta(\tau)$

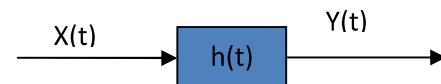
- Déterminer la densité spectrale de la sortie  $y(t)$ , notée  $S_y(f)$
- Déterminer la fonction d'autocorrélation de la sortie  $y(t)$ , notée  $R_y(\tau)$  ainsi que sa puissance



2. Soit  $X(t) = A + b(t)$  un signal réel aléatoire, où  $A$  est une constante réelle et  $b(t)$  est un bruit blanc de densité spectrale de puissance  $\sigma_b^2$ , et soit un filtre moyennneur de réponse impulsionnelle:

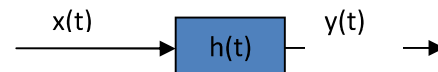
$$h(t) = \frac{1}{T} \Pi(t - T/2)$$

- Exprimer l'autocorrélation statistique du signal d'entrée  $R_{xx}(t, \tau)$
- $X(t)$  est-il SSL ?
- Déterminer la moyenne statistique de  $y(t)$ .
- Montrer que  $R_{yy}(\tau) = \frac{\sigma^2}{T} \Lambda_T(\tau) + A^2$  et en déduire  $\sigma_y$  et  $\mu_y$ .
- Prendre  $A=2$  et tracer  $X(t)$  et  $Y(t)$ .



3. On considère le schéma ci-contre où  $x(t) = s(t) + b(t)$  avec  $s(t) = \sum_i \delta(t - t_i)$  et  $b(t)$  est un bruit blanc Gaussien centré de variance 1

- Déterminer la densité de probabilité de  $x(t)$ .  $x(t)$  est-il SSL ?
- On considère  $b'(t)$  et  $s'(t)$  les sorties correspondant respectivement à  $b(t)$  et  $s(t)$ .
- Calculer  $S_{b'}(f)$  et  $S_{s'}(f)$  en fonction de  $H(f)$
- On veut maximiser le rapport signal sur bruit (SNR):
- Déterminer et tracer  $h(t)$  pour  $k=2$  et  $T_0=2$ . Calculer alors le SNR.
- Quelle loi de probabilité suit  $y(t)$  (justifier).
- Donner un exemple concret de l'utilisation de ce type de filtre
- Si  $s(t)$  est déterministe et inconnu, quel filtre utilise-t-on ?



4. On dispose d'un signal reçu qui est la version bruitée retardée et atténuée d'un signal d'intérêt  $s(n)$ . Le bruit  $b(n)$  est supposé blanc gaussien de variance  $\sigma^2$ . Le problème est de déterminer l'amplitude  $A$  et le retard  $n_0$  dans le signal reçu  $x(n) = A.s(n - n_0) + b(n)$ . Sachant que le rapport signal-à-bruit est maximum en sortie du filtre adapté de réponse  $h(n) = s(-n)$ .

- Vérifier que la sortie  $y(n)$  de ce filtre s'exprime comme la somme de deux fonctions de corrélation.
- Calculer la variance  $\sigma_b^2$  du bruit de sortie.
- On prend pour  $s(n)$  une impulsion rectangulaire de largeur  $L$ , tracer alors un exemple de signal reçu et de sortie du filtre adapté.

## Solutions

- $H(f) = 1/(1 + 2\pi j f RC)$   $S_y(f) = \sigma_x^2 / (1 + 4(\pi f RC)^2)$   $R_y(\tau) = \sigma_x^2 / 2RC e^{-|\tau|/RC}$   $P = R_y(0) = \sigma_x^2 / 2RC$
- $R_{xx}(t, \tau) = A^2 + \sigma_b^2 \delta(\tau) = fct(\tau)$   $\mu_x(t) = A \Rightarrow$  SSL  $H(f) = \text{sinc}(fT) e^{-j\tau f T}$   $\mu_y(t) = A$  filtre moyennneur
- $x(t)$  Gaussien de moyenne  $s(t)$  et de variance 1.  $x(t)$  non stationnaire  $S_{b'}(f) = |H(f)|^2$   $S_{s'}(f) = |H(f)|^2 \text{sinc}^2(f)$   
 $h(t) = 2\Pi(2 - t)$  SNR=1  $y(t)$  Gaussien Radar ou Sonar Passe-bas (moyennneur)



$$4. y(n) = A.R_s(n-n_0) + R_{bs}(n) \sigma_b^2 = R_b(0) \quad S_b(f) = \sigma_b^2 |H(f)|^2 \Rightarrow R_b(0) = \sigma_b^2 \int_{-1/2}^{1/2} |H(f)|^2 df$$

### Exercices supplémentaires

1. Soit  $Y(t)$  un processus aléatoire défini par  $Y(t) = X(t+1) - X(t-1)$ , où  $X(t)$  est un processus aléatoire stationnaire de moyenne nulle. Montrer que  $S_Y(f) = 4 \cdot S_X(f) \cdot \sin^2(2\pi f)$

2. On considère le signal  $x(k) = g(k)\cos(2\pi n_0 k/N + \phi)$  défini pour  $k \in [0, N-1]$ , et où  $g(k)$  est une fonction aléatoire, indépendante de  $\phi$ .

- Calculez l'autocorrélation de  $x(k)$
- Déduisez en sa densité spectrale de puissance, en fonction de la densité spectrale de puissance de  $g$ ,  $\phi_g(f)$ .

Rappels :  $\cos a \cos b = [1/2][\cos(a+b) + \cos(a-b)]$ ,  $\cos x = [(e^{jx} + e^{-jx})/2]$ ,  $\sum_{n=0}^{N-1} q^n = [(1 - q^N)/(1 - q)]$

3. On considère la transmission de deux symboles :

$$s_0(t) = A \quad t \in [0, T]$$

$$s_1(t) = -A \quad t \in [0, T]$$

à travers un canal à bruit blanc additif Gaussien. Le signal reçu s'écrit :  $x(t) = s_i(t) + b(t)$  où  $b(t)$  est un bruit blanc centré de DSP  $\sigma_b^2$

- Déterminer la réponse impulsionnelle du filtre adapté  $h(t)$  correspondant à  $s_0(t)$  telle que  $\int h(t)^2 dt = 1$
- Même question pour  $s_1(t)$ .
- Donner le rapport signal sur bruit, dans chaque cas.

On note  $y_{si}(t)$ , la sortie du filtre correspondant à  $s_i(t)$ ,

- Montrer que pour  $b(t)$  elle s'écrit :  $y_b(T) = \frac{1}{\sqrt{T}} \int_0^T b(t) dt$

- Déterminer la moyenne et la variance de  $y_b(T)$

- Dans l'hypothèse où le bruit blanc est Gaussien, déterminer la loi de probabilité de la variable aléatoire  $Y = y_{si}(t) + y_b(t)$

4. On considère un signal déterministe  $s(t)$  qui modélise une impulsion (par exemple radar/sonar). Cette impulsion est réfléchiée sur une cible et on suppose que le signal reçu en retour (par exemple sur l'antenne de réception) s'écrit :

$$x(t) = s(t - \tau) + b(t)$$

où  $\tau \geq 0$  représente un retard et  $b(t)$  est un bruit blanc centré. Un traitement est appliqué au signal  $x(t)$  en réception afin de maximiser le critère de "rapport signal sur bruit". Plus précisément, on applique à  $x(t)$  un filtre dont la réponse impulsionnelle est notée  $h(t)$  et la réponse en fréquence est notée  $H(f)$ .

- Rappeler comment s'exprime l'énergie  $E_s$  du signal  $s(t)$  ?

- Si  $y(t)$  est la sortie du filtre  $h(t)$  lorsque  $x(t)$  est en entrée, justifier en deux mots que l'on puisse écrire  $x_e(t) = s_e(t - \tau) + b_e(t)$  où  $s_e(t)$  et  $b_e(t)$  sont les sorties du même filtre  $h(t)$  avec respectivement  $s(t)$  et  $b(t)$  en entrée.

- Exprimer  $s_e(t)$  en fonction de  $H(f)$  et  $S(f)$  et en déduire  $s_e(0) = \int H(f)S(f) df$ .

- On note  $N_0/2$  la densité spectrale de puissance de  $b(t)$ .

(a) Que vaut la densité spectrale de puissance de  $b_e(t)$  (notée  $S_{be}(f)$ ) ?

(b) En déduire la puissance de  $b_e(t)$ . Que vaut  $E\{|b_e(\tau)|^2\}$  ?

On désire maximiser le critère "rapport signal sur bruit" en sortie du filtre à l'instant  $\tau$  par un filtrage adapté.

- Donner l'expression de  $h(t)$ ,  $H(f)$  et du RNS.



### TP n° 3 : Filtrage des signaux aléatoires

Ce TP a pour objectif :

- D'étudier comment sont transformés les signaux aléatoires, ou plus exactement leurs caractéristiques, lors d'un filtrage linéaire. Assimiler la notion de filtre formeur.
- Aborder une application directe qui est le filtrage adapté pour la détection d'un signal connu noyé dans du bruit blanc en augmentant le SNR

**Exercice 1** Reprenons le programme 3 du TP n°2 et en y rajoutant ce qui suit, on obtient :

```
clc; clear all; close all;
sigma=0.0023; fe=22050; N=1000;
bb = sqrt(sigma)*randn(N,1);

% Afficher la corrélation Sigtab et la dsp du bruit blanc Sbb(f)
Sigtab=xcorr(bb)/N;t=(-N+1:N-1);
figure; subplot(2,1,1); plot(t,Sigtab)
legend('Autocor du bruit blanc'); xlabel('Temps(s)'); ylabel('Amplitude');
Fb=fft(bb);
Sbb=abs(Fb).^2/N;
f=(0:fe/N:fe/2-1/N);
subplot(2,1,2); plot(f,abs(Sbb(1:N/2)));
legend('Densité du bruit blanc'); xlabel('Frequence (Hz)'); ylabel('Amplitude');
% Reponse fréquentielle du filtre formeur
a=[1.0 -0.839 -0.015 -0.320 0.197 0.055 -0.285 0.067 0.044 0.003 0.178];
b=[1 0];
y= filter(b,a,bb);
[H,f]=freqz(b,a,length(f),fe);
figure; plot(f,20*log10(abs(H(1:length(f)))+eps),'r');
Ys=fft(y); plot(f,20*log10(abs(H(1:length(f)))+eps),'r',
f,20*log10(abs(Ys(1:length(f)))),'b');
legend('Filtre Formeur','DSP du signal de sortie');
xlabel('Hz'); ylabel('Amplitude (dB)')
```

1. Quel est le but des lignes rajoutées ? Commenter chaque figure.
2. Quelle est la nature du filtre formeur ? A quoi peut servir ce filtre ?
3. Rajouter les instructions nécessaires pour vérifier les formules suivantes :

$$- \mu_y = \mu_x \quad H(f) \Big|_{f=0}$$

$$- S_y(f) = TF(R_y(\tau)) = |H(f)|^2 S_x(f)$$

Remarque :  $f=0$  correspond sous matlab au 1er élément de H

4. Inverser a et b et commenter.
5. Donner une application cette notion de filtre formeur.

**Exercice 2** On veut simuler l'émission et la réception d'un signal radar. On suppose que le signal émis  $s(n)$  est un signal sinusoïdal et que le signal reçu  $x(n)$  est une version bruitée (bruit blanc) atténuée et retardée du signal émis.

```
clc; clear all; close all;
%Signal émis de longueur L
L = 300; t = (1:L)'; f0=0.02; s=sin(2*pi*f0*t);
%Signal reçu
N=1000; var = 4; bruit = sqrt(var)*randn(N,1);
Retard = 500; A=0.8;
x= bruit;
x(Retard+t) = A*x(Retard+t)+s;
figure; subplot(211); plot(s); xlabel('Signal émis');
subplot(212); plot(x); xlabel('Signal reçu');
% Création du Filtre adapté
for i=0:L-1;
    h(i+1)=s(L-i);
end

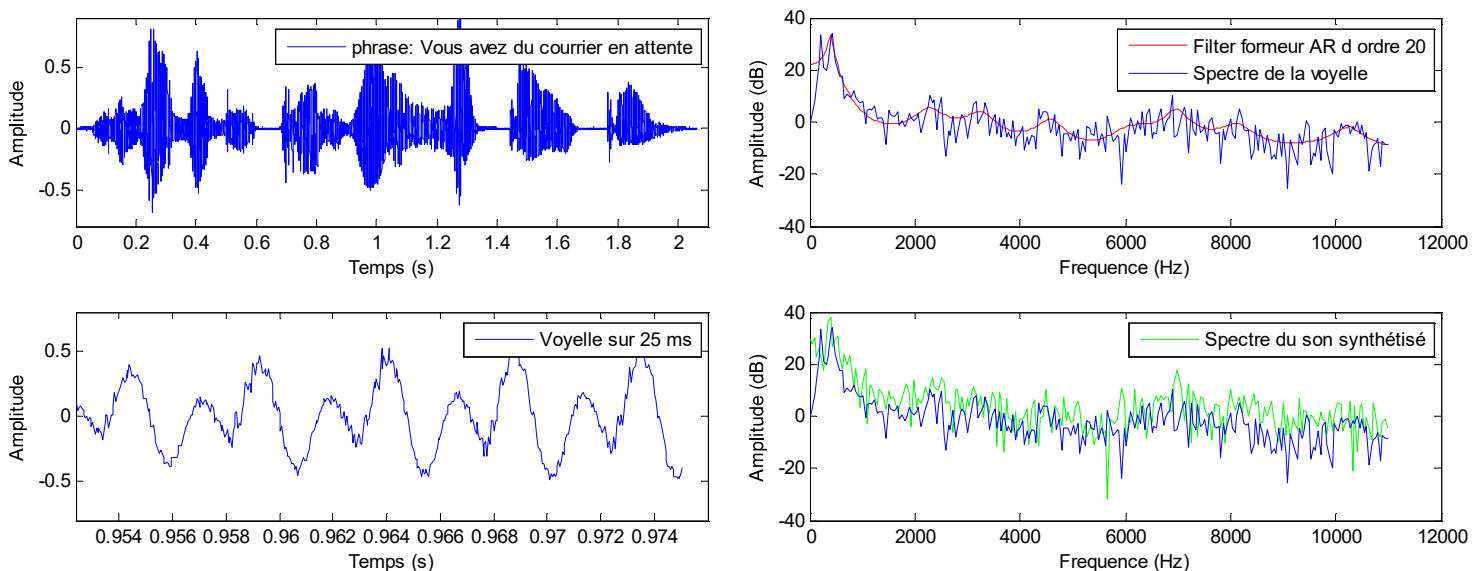
figure; subplot(211); plot(s); xlabel('Sig émis');
subplot(212); plot(h); xlabel('Filtre adapté');
% Filtrage
y = conv(h,x); y = y/(length(y));
figure; subplot(2,1,1); plot(x); xlabel('Signal reçu');
subplot(2,1,2); plot(y); xlabel('Signal reçu filtré');
tems_aller_retour=find(y>=max(y))-L
```

1. Montrer que pour  $y(t) = a.x(t - T_{AR}) + b(t)$  où  $b(t)$  est un bb, on obtient lors d'un filtrage adapté ( $h(t) = x^*(T_0 - t)$ ) en sortie l'expression suivante  $R_{yx}(t - T_0) = a.R_{xx}(t - T_0 - T_{AR}) + R_{bx}(t - T_0)$ .
2. Que représentent A, s et x?
3. Arrive-t-on à identifier la partie comportant la sinusoïde dans le signal reçu? pourquoi?
4. Expliciter la boucle. Quel est, alors, le lien entre x et h?
5. Expliquer l'instruction `tems_aller_retour=find(y>=max(y))-L`
6. Commenter le programme plus particulièrement les figures. Quel est le but de ce programme?
7. Tester d'autres signaux utiles (par ex : `x = ones(L,1)`) et commenter. Pour détecter le temps d'aller-retour, lequel des 2 signaux vous paraît préférable?
8. Faire varier la puissance du bruit (`var=2` puis `4` en gardant une `moy=0`) et commenter.
9. Simuler le processus de l'aide au stationnement en mettant ce programme dans une boucle pour un Retard de 700 à 0 avec un pas de 50 et rajouter à la fin les instructions :  
`pause(tems_aller_retour/500);beep;` dans la boucle.

### Chapitre 4 : Processus générateurs AR, MA et ARMA

Nous avons vu au chapitre précédent qu'un signal aléatoire peut être modélisé (synthétisé) comme la réponse d'un filtre linéaire à une excitation sous forme de bruit blanc tel que  $|H(f)|^2 = S_x(f) / \sigma_b^2$ . Ce filtre formeur  $H(f)$  est aussi dit processus générateur. Ces paramètres associés à la variance du bruit  $\sigma_b^2$  constituent le modèle mathématique correspondant au signal aléatoire. Le concept de processus générateur de signal a été particulièrement développé et appliqué avec des filtres numériques. Ce sont les modèles de signaux les plus utilisés en traitement statistique du signal (estimation, prédiction...). Selon la nature du filtre, on peut obtenir différents modèles de signaux (AR, MA, ARMA, etc.)

Exemple : Lors d'un appel par GSM (téléphone portable), le portable qui fait office d'un micro-ordinateur regroupant différentes fonctionnalités dont l'analyse, la synthèse, le codage, etc. va nous permettre de modéliser la parole en opérant un codage LPC par tranches de dizaines de ms. Elle consiste à retrouver les paramètres du filtre formeur pour chaque tranche enregistrée et analysée. Ceux sont ces paramètres qui seront transmis pour produire un signal de synthèse approchant le signal original. Ci-dessous la phrase "vous avez du courrier en attente" échantillonnée à une fréquence  $f_e = 22050$  (45531 échantillons) et un zoom sur une voyelle de durée 25 ms (500 échantillons).



Il vaut mieux transmettre 21 coefficients que 500 valeurs. C'est ainsi que les modèles autorégressifs sont d'un emploi de plus en plus répandu en traitement du signal : codage et transmission par prédiction linéaire, synthèse de parole, reconnaissance, etc.

Remarque : Le signal de parole est un processus aléatoire non-stationnaire à long terme, mais il est considéré comme stationnaire dans des fenêtres temporelles d'analyse de l'ordre de 20 à 30ms. Cette propriété de stationnarité à court terme permet donc une analyse et modélisation progressive du signal de parole. Pour éviter toutes pertes d'information, on veillera à prendre des fenêtres.

### 1. Modèle auto-régressif (AR)

Les signaux autorégressifs sont obtenus par passage d'un bruit blanc dans un filtre purement récursif. Ce filtre est donc de réponse impulsionnelle infinie.

$$H(z) = 1 / \left( 1 + \sum_{i=1}^N a_i z^{-i} \right)$$

A partir de  $H(z)$ , on peut déterminer l'équation aux différences :  $y(n) = x(n) - \sum_{i=1}^N a_i y(n-i)$

Cela signifie que le signal  $y(n)$  est supposé être prédictible en fonction d'un certain nombre de ses valeurs antérieures.

Sachant que  $x(n)$  est un bruit blanc alors d'où  $R_x(k) = \sigma^2 \delta(k)$

$$- \mu_x = E\{x(n)\} = 0$$

$$- R_{xx}(0) = E\{x(n)^2\} = \sigma^2$$

$$- R_{xx}(k) = E\{x(n)x(n-k)\} = 0 \text{ pour } k \neq 0$$

Calculons alors  $R_{yy}(k)$

$$\begin{aligned} R_{yy}(k) &= E\{y(n)y(n-k)\} = E\{x(n)y(n-k)\} - \sum_{i=1}^N a_i E\{y(n-i)y(n-k)\} \\ &= E\left\{x(n) \sum_{k'} x(n-k-k') h(k')\right\} - \sum_{i=1}^N a_i R_{yy}(k-i) = \sum_{k'} E\{x(n)x(n-k-k')\} h(k') - \sum_{i=1}^N a_i R_{yy}(k-i) \\ &= \sum_{k'} R_{xx}(k+k') h(k') - \sum_{i=1}^N a_i R_{yy}(k-i) = R_{xx}(k)h(0) + R_{xx}(k+1)h(1) + \dots - \sum_{i=1}^N a_i R_{yy}(k-i) \end{aligned}$$

$$\text{- Si } k=0, R_{yy}(0) = R_{xx}(0)h(0) - \sum_{i=1}^N a_i R_{yy}(-i) = \sigma^2 \cdot 1 - \sum_{i=1}^N a_i R_{yy}(i)$$

$$\text{- Pour } k=1 \text{ à } N, R_{yy}(k) = 0 - \sum_{i=1}^N a_i R_{yy}(k-i) = - \sum_{i=1}^N a_i R_{yy}(k-i)$$

On peut utiliser une forme matricielle :  $R \cdot \underline{a} = \underline{S}$  avec

$$\begin{bmatrix} R_{yy}(0) & R_{yy}(1) & \dots & R_{yy}(N) \\ R_{yy}(1) & R_{yy}(0) & \dots & R_{yy}(N-1) \\ & & \dots & \\ R_{yy}(N) & R_{yy}(N-1) & \dots & R_{yy}(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

R est la matrice d'autocorrélation dont le terme général  $r_{ij}$  ne dépend que de la différence  $i-j$  (Matrice de Toeplitz). La résolution de cette équation permet de connaître les paramètres du filtre et la variance du bruit blanc. Il existe divers algorithmes (Burg, Levinson) qui permettent d'estimer assez rapidement les  $a_i$  et  $\sigma$  sans passer par l'inversion matricielle. Tout comme il est possible de déterminer l'ordre N adéquat (critère AIC).

### Exemple d'application

1. On considère le modèle auto-régressif (AR) d'ordre 1 tel que :  $x(n) = -a_1 x(n-1) + b(n)$

- Déterminer les équations de Yule-Walker pour ce modèle
- En supposant que  $x(n)$  est connu, déterminer les paramètres du modèle.
- Déterminer les  $R_x(k)$  (les  $a_i$  sont supposés connus)

Réponses :

- $(a_1 = -R_x(1)/R_x(0) \quad \sigma^2 = R_x(0)(1-a_1^2))$
- $(R_x(0) = \sigma^2/(1-a_1^2) \quad R_x(1) = -a_1 \sigma^2/(1-a_1^2) \quad R_x(k) = (-a_1)^k \sigma^2/(1-a_1^2))$

2. Refaire le même travail pour un modèle d'ordre 2 tel que :  $x(n) = -a_1 x(n-1) - a_2 x(n-2) + b(n)$

Réponses :

- $a_1 = R_x(1)[R_x(2) - R_x(1)^2] / [R_x(0)^2 - R_x(1)^2] \quad a_2 = [R_x(1)^2 - R_x(0)R_x(2)] / [R_x(0)^2 - R_x(1)^2]$
- $\sigma^2 = R_x(0) + R_x(1)^2[R_x(2) - R_x(0)] / [R_x(0)^2 - R_x(1)^2] + R_x(2)[R_x(1)^2 - R_x(0)R_x(2)] / [R_x(0)^2 - R_x(1)^2]$
- $R_x(1) = -a_1 R_x(0) / (a_2 + 1) \quad R_x(2) = (-a_2 + a_1^2 / (1 + a_2)) R_x(0) \quad R_x(0) = (1 + a_2) \sigma^2 / (1 + a_2 - a_1^2 - a_1^2 - a_2^3 + a_2 a_1^2)$

## 2. Modèle à moyenne ajustée (MA)

Les signaux à moyenne mobile sont obtenus par passage d'un bruit blanc dans un filtre purement transverse.

Ce filtre est aussi appelé filtre à réponse impulsionnelle finie :  $H(z) = \sum_{i=0}^M b_i z^{-i}$

Le signal  $y(n)$  est supposé pouvoir s'écrire comme une combinaison linéaire d'échantillons décorrélés entre eux, ce qui peut se formaliser comme une combinaison linéaire d'échantillons d'un bruit blanc  $x(n)$ .

$$\text{On a donc : } y(n) = \sum_{i=0}^M b_i x(n-i)$$

$$\text{et } \mu_y = E\{y(n)\} = \sum_{i=0}^M b_i \mu_x = \mu_x \sum_{i=0}^M b_i = \mu_x \cdot H_f(0)$$

On cherche les paramètres du filtre qui génèrent  $y(t)$  à partir de  $x(t)$ , bruit blanc centré :

$$R_{yy}(k) = E\{y(n)y(n-k)\} = E\left\{\sum_{i=0}^M b_i x(n-i) \cdot \sum_{j=0}^M b_j x(n-j-k)\right\}.$$

$$R_{yy}(k) = \sum_{i=0}^M b_i \cdot \sum_{j=0}^M b_j E\{x(n-i) \cdot x(n-j-k)\} = \sum_{i=0}^M b_i \cdot \sum_{j=0}^M b_j R_{xx}(j+k-i)$$

- Si  $j+k \neq i \Rightarrow R_{yy}(k) = 0$
- Sinon  $\Rightarrow R_{yy}(k) = \sigma^2 \sum_{j=0}^{M-k} b_{j+k} \cdot b_j$

Le problème est non linéaire en fonction des coefficients, il faut un algorithme de programmation non linéaire pour obtenir  $b_i$  à partir des  $R_{yy}(k)$ . Cependant, l'algorithme de Durbin permet d'approcher la solution optimale avec de bons résultats. Le principe de cet algorithme consiste à identifier le modèle MA d'ordre M avec un modèle AR d'ordre  $N \gg M$ .

En effet, tout modèle MA peut être identifié à un modèle AR d'ordre infini:  $\sum_{i=0}^M b_i z^{-i} = 1 / \sum_{i=0}^{\infty} a_i z^{-i}$

### Exemple

On considère le modèle à moyenne ajustée (MA) :

A. d'ordre 1 tel que  $x(n) = e(n) + b_1 \cdot e(n-1)$

- Calculer  $\mu_x$ .
- En supposant que  $x(n)$  est connu, déterminer les paramètres du modèle.
- Connaissant les paramètres du modèle, déterminer les  $R_x(k)$

B. d'ordre 2 tel que :  $x(n) = e(n) + b_1 \cdot e(n-1) + b_2 \cdot e(n-2)$

- Calculer  $\mu_x$ .
- Connaissant les paramètres du modèle, déterminer les  $R_x(k)$

### Réponses

$$\begin{aligned} \mu_x &= 0 & R_x(0) &= (1 + b_1^2) \cdot \sigma^2 & R_x(1) &= b_1 \cdot \sigma^2 & R_x(k) &= 0 \text{ pour } k \geq 2 \\ b_1 &= (R_x(0) \pm \sqrt{R_x(0)^2 - 4R_x(1)^2}) / 2R_x(1) & \sigma^2 &= 2 / (R_x(0) \pm \sqrt{R_x(0)^2 - 4R_x(1)^2}) \\ R_x(0) &= (1 + b_1^2 + b_2^2) \cdot \sigma^2 & R_x(1) &= (b_1 + b_1 b_2) \cdot \sigma^2 & R_x(2) &= b_2 \cdot \sigma^2 & R_x(k) &= 0 \text{ pour } k \geq 3 \end{aligned}$$

Remarque : Il est très important de remarquer que nous ne disposons que d'une seule réalisation du signal aléatoire à modéliser  $y(n)$ , de ce fait l'auto-corrélation statistique  $R_{yy}(k)$  est obtenu de l'auto-corrélation temporelle en considérant le processus ergodique.

### 3. Modèle ARMA

Les signaux ARMA sont obtenus par passage d'un bruit blanc dans un filtre récursif appelé aussi filtre à réponse impulsionnelle infinie (R.I.I). Ces signaux sont une combinaison des signaux AR et MA. La fonction de transfert du filtre présente un numérateur et un dénominateur:

$$\text{Soit } y(n) = \sum_{i=0}^M b_i x(n-i) - \sum_{i=1}^N a_i y(n-i) \quad H(z) = \sum_{i=0}^M b_i z^{-i} / \left( 1 + \sum_{i=1}^N a_i z^{-i} \right)$$

La corrélation statistique de  $y(n)$  s'écrit alors : 
$$R_{yy}(k) = -\sum_{i=1}^N a_i R_{yy}(k-i) + \sigma^2 \sum_{j=0}^{M-k} b_{j+k} \cdot b_j$$

C'est une équation non linéaire en  $a_i$  et  $b_j$ .

La modélisation ARMA peut se décomposer en une modélisation AR suivie d'une modélisation MA. Le modèle AR présente une simplicité de calcul par rapport aux modèles MA et ARMA du fait où les coefficients AR sont solutions d'un système linéaire d'équations. Alors que la détermination des coefficients MA et ARMA requiert la résolution d'équations non linéaires. Cependant, le modèle ARMA permet de modéliser aussi bien les minima que les maxima de la DSP et est donc moins restrictif que le modèle AR.

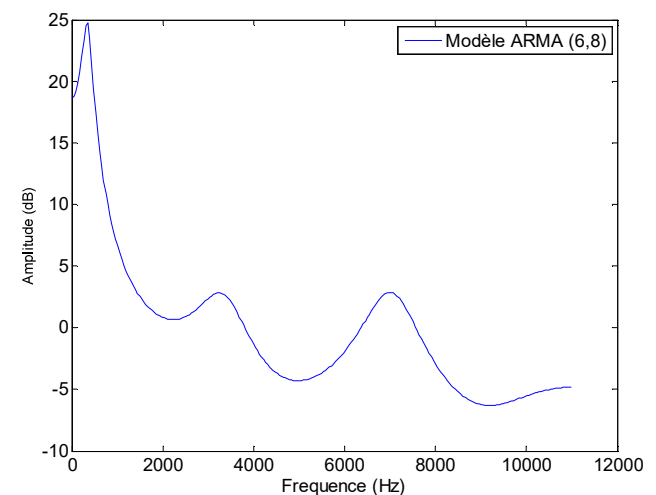
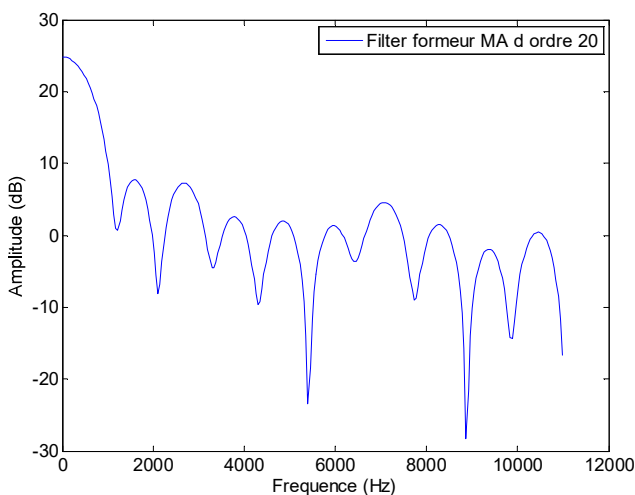
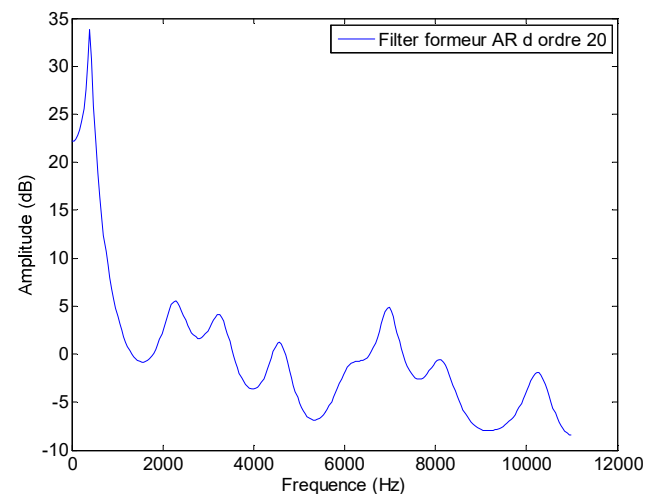
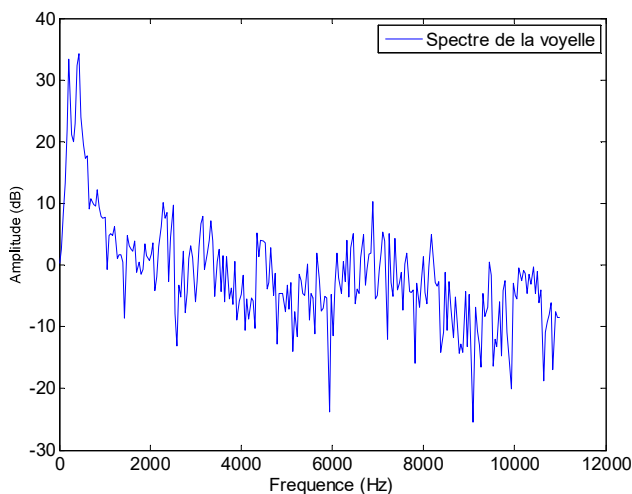
Les application des modèles AR, MA, ARMA sont nombreuses, entre autres :

- la modélisation et la prédiction de série temporelle,
- l'estimation du spectre d'un signal aléatoire, etc.

Cette dernière application est basée sur l'identification des paramètres du modèle considéré:

- Le modèle AR est bien adapté aux signaux composés de raies pures dans du bruit blanc.
- Le modèle MA est bien adapté aux signaux dont la puissance est nulle dans certaines bandes de fréquences.

Voir exemple ci-dessous .





### TD n°4 : Modèles AR, MA, ARMA

1. Soit un signal aléatoire  $y(n)$  SSL et ergodique dont l'autocorrélation temporelle  $\overline{R}_y(k) = \alpha^{|k|}$  avec  $0 < \alpha < 1$ 
  - Identifier le modèle linéaire adéquat (AR ou MA) pour  $y(n)$ .
  - En supposant que le système  $h(n)$  est filtre purement récursif, donner le schéma du modèle en définissant l'entrée, le système, et la sortie.
  - Rappeler les hypothèses nécessaires liées à l'emploi de ce modèle.
  - On considère que le modèle est d'ordre 1, déterminer ses paramètres.
2. On considère le filtre linéaire à temps discret défini par  $y(n) = x(n) + b_1 x(n-1) + b_2 x(n-2)$ .  
où  $X(n)$  et  $Y(n)$  désignent respectivement les processus aléatoires réels d'entrée et de sortie du filtre où  $b_1$  et  $b_2$  sont 2 coefficients réels. On suppose que  $x(n)$  est une suite de variables aléatoires centrées, indépendantes et de variance  $\sigma^2$ .
  - Donner l'expression de  $R_x(k)$  et  $S_x(f)$  puis donner l'expression de  $R_y(k)$  et tracer la pour  $b_1=1$  et  $b_2=-1$ .
  - Connaissant la DSP du signal  $y(n)$ , sur quoi se base-t-on pour le choix du modèle ?
3. Soit un processus AR défini par :  $y(n) = -a_1 y(n-1) - a_2 y(n-2) + x(n)$  où  $x(n)$  bb décorrélé de variance 1
  - Calculer  $\mu_y(n)$  puis sans calcul, expliquer pourquoi  $y(n)$  est SSL.
  - Montrer que pour  $k > 0$ ,  $R_y(k) = -a_1 R_y(k-1) - a_2 R_y(k-2)$
  - Déterminer  $a_1$  et  $a_2$
4. Soit un filtre formeur dont l'équation aux différences est  $y(n) = 0.25 y(n-1) - 0.25 y(n-2) + x(n)$ 
  - Identifier ce modèle linéaire AR ou MA (Justifier)
  - Déterminer la moyenne de  $y(n)$  et donner l'expression de  $R_{yy}(k)$ .
  - Calculer et tracer  $R_{yy}(k)$  (Prendre  $R_{xx}(0) = \sigma^2 = 1$ ).

#### Solutions

1. Modèle AR ( $R_y(k) \neq 0$ )  
 $x(n)$  entrée bb,  $y(n)$  signal aléatoire à modéliser  $h(n)$  filtre formeur (modèle math)  
 entrée bb + ergodisme  $a_1 = -\alpha$   $\sigma_x^2 = 1 - \alpha^2$
2.  $R_x(k) = \sigma_x^2 \delta(k)$   $S_x(f) = \sigma_x^2$  MA d'ordre 2  $R_y(0) = 1 + b_1^2 + b_2^2$   $R_y(1) = b_1(1 + b_2)$   $R_y(2) = b_2 R_y(1)$   $R_y(k) = 0$  pour  $k \geq 0$
3.  $\mu_y(n) = 0$  entrée bb SSL  $\Rightarrow$  sortie SSL voir cours (page 41)

#### Exercices supplémentaires

1. Les signaux  $x(n)$  et  $y(n)$  ont été obtenus en filtrant, au moyen d'un filtre à réponse impulsionnelle finie, un bruit blanc  $b(n)$  gaussien centré de variance  $\sigma^2$ .

A] L'équation de filtrage est la suivante :  $x(n) = 2.b(n) + 0.5.b(n-1) - 0.2.b(n-2) + 0.1.b(n-3)$

- Calculez, en fonction de  $\sigma^2$ , les coefficients d'autocorrélation d'ordre 0,1,2,3 du signal  $x(n)$ .
- On notera  $R_{xx}(0)$ ,  $R_{xx}(1)$ ,  $R_{xx}(2)$ ,  $R_{xx}(3)$  ces coefficients.
- La répartition des niveaux d'amplitude du signal  $x(n)$  est-elle gaussienne (sans justifier)?

B]  $y(n) = b(n) - b(n-2)$

- Donnez la TZ de la réponse impulsionnelle du filtre qui a permis d'obtenir  $y(n)$  à partir de  $b(n)$ .
- Placez les zéros de ce filtre sur un cercle unité, quelles sont la ou les fréquence(s) coupées par ce filtre ?

- Tracez approximativement sa réponse en fréquence.

C] On considère maintenant le coefficient d'intercorrélation  $R_{xy}(k)$ , entre les signaux  $x(n)$  et  $y(n)$  donné par  $R_{xy}(k) = E[x(n) y(n-k)^*]$ .

Calculez  $R_{xy}(0)$ ,  $R_{xy}(1)$ ,  $R_{xy}(2)$ .

2. On considère un signal aléatoire stationnaire  $x(n)$  et l'on suppose connu ses coefficients d'autocorrélation :

$$R(0) = 3\sigma^2, R(1) = 2\sigma^2, R(2) = \sigma^2, R(3) = 0$$

- On cherche le filtre MA d'ordre 3 de ce signal. Identifiez les paramètres du filtre.

- Si le signal  $x(n)$  a été obtenu par filtrage d'un bruit blanc gaussien de variance  $\sigma^2$  par un filtre à réponse impulsionnelle finie de fonction de transfert  $H(z) = 1 + a.z^{-1} + b.z^{-2}$ , en déduire les valeurs de  $a$  et  $b$ .

3. Soit un filtre formeur dont l'équation aux différences est  $y(n) = 0.5(x(n) + x(n-1) + x(n-2) + x(n-3))$

- Expliquer la notion de filtre formeur
- Identifier ce modèle linéaire AR ou MA
- Donner la moyenne, l'autocorrélation et la DSP de son entrée  $x(n)$ .
- Calculer et tracer  $R_{yy}(k)$  puis déterminer  $S_{yy}(f)$

### Solutions

$$1. R_x(0) = (b_0^2 + b_1^2 + b_2^2 + b_3^2) \cdot \sigma^2 \quad R_x(1) = (b_0 b_1 + b_1 b_2 + b_2 b_3) \cdot \sigma^2 \quad R_x(2) = (b_0 b_2 + b_1 b_3) \cdot \sigma^2 \quad R_x(3) = (b_0 b_3) \cdot \sigma^2 \quad R_x(k \geq 4) = 0$$

$$H(z) = (z^2 - 1) / z^2 \quad R_{xy}(0) = 2.2 \quad R_{xy}(1) = 0.4 \quad R_{xy}(2) = -0.2 \quad R_{xy}(3) = 0.1 \quad R_{xy}(-1) = -0.5 \quad R_{xy}(-2) = -2$$

$$2. a_1 = -3/4, a_2 = 0 \text{ et } a_3 = 1/4 \text{ et } \sigma_b^2 = \sigma^2(3 + 2a_1 + a_2) \text{ D'après A (exo1) par identification, on trouve } a = 1 \text{ et } b = 1$$

## TP n° 4 : Modèles AR et MA

Ce TP a pour objectif de modéliser un signal aléatoire par les modèles AR et MA (par approximation AR) et d'en extraire les informations utiles. Puis, une application d'identification sera envisagée.

**Exercice 1 :** Télécharger le fichier 'vous avez du courrier en attente.wav' et le placer dans le même répertoire que votre programme commençant come suit :

```
clc; clear all; close all;
nom_fich = uigetfile('*.wav', 'Selectionner le fichier son');
% Lire, écouter et afficher le son complet
[x,fe]=wavread(nom_fich);sound(x,fe); t=(0:length(x)-1)/fe; subplot(2,1,1);plot(t,x);
legend('Son');xlabel('Temps (s)');ylabel('Amplitude');
% lire, écouter et afficher une partie du son qui correspond à une voyelle
N1=21000;N2=21500;[y,fe]=wavread(nom_fich,[N1 N2]);sound(y,fe)
N=length(y);t=(N1:N2)/fe; subplot(2,1,2);plot(t,y);
legend('Voyelle sur 25 ms');xlabel('Temps (s)');ylabel('Amplitude');
% Afficher le spectre
Sy=fft(y); f=(0:fe/N:fe/2-1/N);figure;plot(f,20*log10(abs(Sy(1:length(f)))+eps));
legend('Spectre'); xlabel('Frequence (Hz)'); ylabel('Amplitude (dB)');
```

Nous allons commencer par déterminer les paramètres du modèle soit la modélisation

1. Ecrire les instruction nécessaires pour résoudre les équations de Yule-Walker en suivant ces étapes:

a. Montrer qu'on peut les formuler comme suit :

$$\begin{bmatrix} R_{yy}(0) & R_{yy}(1) & \dots & R_{yy}(P-1) \\ R_{yy}(1) & R_{yy}(0) & \dots & R_{yy}(P-2) \\ & & \dots & \\ R_{yy}(P-1) & R_{yy}(P-2) & \dots & R_{yy}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} -R_{yy}(1) \\ -R_{yy}(2) \\ \vdots \\ -R_{yy}(P) \end{bmatrix} \quad \sigma^2 = R_{yy}(0) + \sum_{i=1}^P a_i R_{yy}(i)$$

b. Pour l'ordre P=20, utiliser la fonction autocorr pour calculer les valeurs de  $R_{yy}$  sur P valeurs.

Pour la matrice C, utiliser la fonction de Toeplitz (Faire attention aux indices).

Par inversion matricielle ( $\text{inv}(C)*B$ ), déterminer les coefficients 'a' (rajouter  $a_0=1$  aux coefficients trouvés) puis déterminer  $\sigma^2$ .

2. Rajouter les instructions permettant de visualiser le filtre formeur en db superposé au spectre de 'y'.

3. L'autocorrélation  $R_{yy}$  intervenant dans les équations de Yule-Walker est-elle statistique ou temporelle ? Qu'en est-il de celle calculée dans ce programme? Commenter.

Entamons la partie synthèse en ayant comme données de départ les paramètres du modèle ( $a_i$  et sigma)

3. A partir de  $\sigma^2$  et des coefficients AR 'a' trouvés, synthétiser et écouter le son 'zz' (utiliser randn et filter).

4. Visualiser le son modélisé et le son synthétisé ainsi que leurs spectres en db sur le même graphe.

5. Modifier N et commenter les graphes précédents.

6. Rajouter les lignes suivantes et donner une application possible.

```

r=roots(a); % Racines de du dénominateur
r=r(imag(r)>0.01); % Recherche des pôles positifs (freq>0)
freq_form=sort(atan2(imag(r),real(r))*fe/(2*pi)); % convert en Hz et trier
for i=1:length(freq_form)
fprintf('Formant %d Frequence %.1f\n',i,freq_form(i));
end

```

7. Pourquoi utilise-t-on le modèle AR pour trouver les fréquences principales au lieu du signal 'y'?

### Exercice 2 :

Sachant qu'un modèle MA peut être obtenu en identifiant le modèle MA d'ordre M avec un modèle AR d'ordre  $P \gg M$  :

$$\sum_{i=0}^M b_i z^{-i} = 1 / \sum_{i=0}^{\infty} a_i z^{-i}$$

1. Modifier le programme précédent en rajoutant l'instruction suivante au bon endroit: `bb=impz(1,a,P);`
2. Expliquer comment on s'est servi de cette instruction pour obtenir les  $b_i$ .
3. Visualiser les  $b_i$  pour  $P=60$  puis en prendre la moitié puis le tiers et reprendre les questions de 2 à 4 de l'exercice 1 en modifiant les instructions nécessaires.

### Exercice 3

Vous avez deux fichiers audio contenant l'enregistrement des 5 voyelles.

1. Lire le premier fichier audio et identifier le début et la fin de chaque voyelle et mettre les signaux obtenus dans une matrice de 5 lignes.
2. Ecrire une boucle, qui pour chaque voyelle, permet de :
  - écouter la voyelle, la centrer et la visualiser
  - visualiser son spectre
  - déterminer les paramètres du modèle AR d'ordre 20 par la commande `[a,sigma]=lpc(x,20);`
  - Déterminer les formants à partir du modèle
3. Prendre le deuxième fichier son et lui faire subir les mêmes traitements.
4. Comparer par visualisation les deux signaux correspondant à la même voyelle ainsi que leur spectre et commenter
5. Comparer les formants de la voyelle "a" du premier fichier à ceux de toutes les voyelles du deuxième fichier son. L'identification est-elle concluante?

## Chapitre 5 : Notions d'estimation

Supposons que l'on veuille connaître le poids 'p' ou la taille 'T' que devrait avoir un enfant de 4 ans. Pour ce faire, on prendra un échantillon d'enfants de 4 que l'on pèsera et dont on mesurera la taille. Une moyenne sur ces mesures nous fournira une estimation  $\hat{p}$  du poids une estimation  $\hat{T}$  de la taille. Cet échantillon devra bien sûr être le plus grand possible et le plus hétérogène possible (plusieurs ethnies). L'estimation de la durée moyenne d'une communication téléphonique est très utile aux opérateurs de téléphonie pour ajuster les prix et offres en conséquence ou pour la gestion du trafic. Autre exemple : les intentions de vote. Dans chacun des cas, un échantillon est considéré pour représenter la population.

### 1. Estimateur

Dans de nombreuses situations, on ne dispose pas directement de mesures sur la variable d'intérêt mais que d'une observation liée à notre variable inconnue. Le but des techniques d'estimation est d'utiliser les observations pour extraire de l'information sur la grandeur d'intérêt. Ceci doit être fait de la meilleure façon qui soit et implique plusieurs choix dont la relation mathématique entre l'observation et la variable à estimer.

Dans les problèmes d'estimation, on a, en général, affaire à deux catégories de variables:

$\underline{X}$ : Un vecteur de variables inconnues que l'on cherche à déterminer

$\underline{Y}$ : Un vecteur de variables mesurées (observation ou échantillon) liées à  $\underline{x}$ .

Déterminer  $\underline{X}$  à partir de  $\underline{Y}$  (et donc en fonction de  $\underline{Y}$ ) consiste à traiter de manière déterministe les mesures  $\underline{Y}$  pour obtenir une grandeur  $\hat{X}$ , proche de  $\underline{X}$ , soit :  $\hat{X} = fct(\underline{Y})$ .

On peut distinguer les deux cas suivants :

- $\underline{X}$  et  $\underline{Y}$  sont aléatoire (estimateurs de Bayes).
- Si  $\underline{X}$  est déterministe et  $\underline{Y}$  aléatoire (estimateurs de Fisher)

Un estimateur est donc une valeur  $\hat{X}$  calculée sur un échantillon  $\underline{Y}$  tiré au hasard (dit aussi observations), ce qui fait de  $\hat{X}$  une variable aléatoire possédant une moyenne et une variance.

### Exemple

Considérons que l'on veuille estimer la valeur d'une résistance  $R$  (voir Exo 1 du TP n°5). On procède comme suit : On réalise  $N$  expériences indépendantes où l'on mesure le courant  $I(k)$  passant par la résistance et la tension  $U(k)$  à ses bornes. Ces mesures de courant et de tension sont toutes bruitées par un bruit dont on ignore la densité de probabilité.

Observations :  $\underline{Y} = (I(1), U(1), I(2), U(2), \dots, I(N), U(N))$

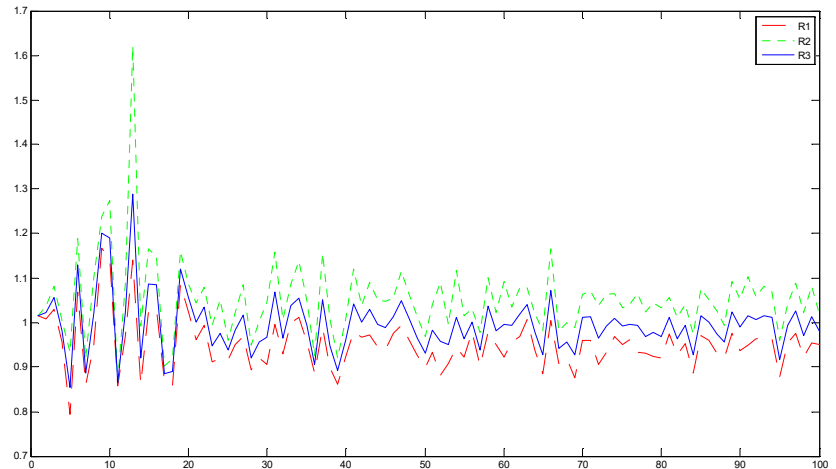
Variable à estimer :  $\underline{X} = R$

Bien que  $X$  et  $Y$  ne soient pas supposés avoir une densité de probabilité, on peut construire différents estimateurs de  $R$  en fonctions des observations  $U(k)$  et  $I(k)$

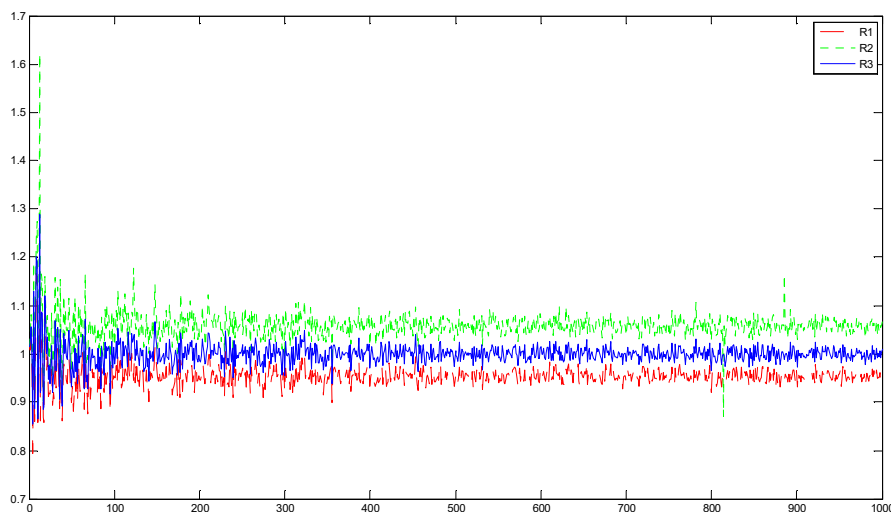
$$\text{Estimateur 1 : } \hat{R}_1 = \frac{\sum_{k=1}^N U(k)I(k)}{\sum_{k=1}^N I^2(k)}$$

$$\text{Estimateur 2 : } \hat{R}_2 = \frac{1}{N} \sum_{k=1}^N \frac{U(k)}{I(k)}$$

$$\text{Estimateur 3 : } \hat{R}_3 = \frac{\sum_{k=1}^N U(k)}{\sum_{k=1}^N I(k)}$$



Ces figures montrent l'évolution de l'estimée d'une résistance  $R=1$  en fonction du nombre de mesures pour  $U=I=1$ . Ci-contre une estimation pour  $N=100$  et ci-bas pour  $N=1000$



On remarque que le premier estimateur donne des valeurs comprises entre 0.9 et 1. Le deuxième fournit une estimation proche de 1 et 1.2 Ils sont donc tous deux fortement biaisés. Le troisième estimateur est quant à lui non biaisé.

Ainsi, pour pouvoir déterminer d'une façon constructive une règle d'estimation, il faut définir un critère qui évalue la qualité des résultats, et définir l'estimée comme l'application de  $Y$  en  $X$  qui optimise ce critère (minimiser une erreur ou une distance, dite aussi fonction de coût). Deux mesures de qualité de l'estimation sont largement utilisées, il s'agit du biais et de la variance.

## 2. Propriétés des estimateurs

Le biais est la moyenne de l'écart et la variance est la puissance de l'écart (mesure les fluctuations de l'estimateur autour de la valeur souhaité).

L'estimateur  $\hat{x}$  a pour biais  $b = E\{\hat{x} - x\}$  et pour variance:  $\sigma^2 = E\{(\hat{x} - E\{\hat{x}\})^2\} = E\{\hat{x}^2\} - E\{\hat{x}\}^2$

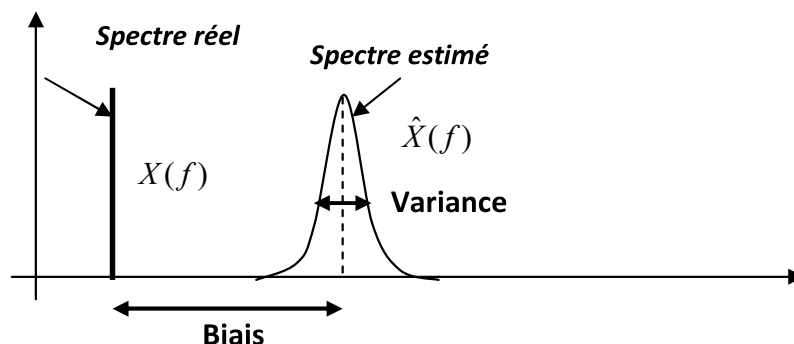
- Si  $x$  est déterministe, alors le biais est  $b = E\{\hat{x}\} - x$
- Si  $x$  est aléatoire, alors le biais vaut  $b = E\{\hat{x}\} - \mu_x$

□ Le biais indique la valeur moyenne de l'erreur d'estimation, trois cas sont possibles:

- $E\{\hat{x}\} = x$  ou  $\mu_x$  pour toutes les valeurs possibles du paramètre. On dit alors que l'estimée est non-biaisée ;
- $E\{\hat{x}\} = x + b$  ou  $\mu_x + b$ . Si  $b$  est indépendant de  $x$ , dans ce cas l'estimateur a un biais constant et connu, qui peut toujours être éliminé ;
- $E\{\hat{x}\} = x + b(x)$ , c'est-à-dire, on a un biais qui dépend de  $x$  (qui est inconnu).

□ La variance doit être aussi petite que possible, de façon à que l'estimée soit concentrée autour de la vraie valeur du paramètre.

□ Plus  $b$  et  $\sigma^2$  sont faibles, meilleur est l'estimateur (estimateur non biaisée à variance minimale). Malheureusement, la diminution de l'un provoque l'augmentation de l'autre. Dans ces conditions, un estimateur biaisé pourrait être préférable, si le biais est faible, à un estimateur non biaisé mais de très grande variance. C'est le compromis biais-variance, on peut démontrer que  $b^2 = E\{\hat{x} - x\}^2 - \sigma^2$  où  $E\{\hat{x} - x\}^2$  est l'erreur quadratique moyenne.



□ On dit qu'un estimateur est *consistant* s'il tend vers la vraie valeur du paramètre quand le nombre d'observations tend vers infini (comportement asymptotique) :

$$\lim_{N \rightarrow +\infty} b_N = \lim_{N \rightarrow +\infty} \sigma_N = 0$$

□ Si l'estimateur est non biaisé, nous avons une borne *inférieure de la variance* (borne de Cramer Rao). Si la variance atteint la borne de CR, on dit que l'estimateur est *efficace*. Autrement dit, un estimateur non-biaisé est appelé efficace si sa variance dans est plus petite que celle de n'importe quel autre estimateur non-biaisé.

**Remarque:** Dans ce qui suit, on parlera de variables  $x_i$  iid (indépendantes et identiquement distribuées). L'exemple classique est celui du lancé de dé : Les variables aléatoires représentent chaque résultat des lancers (0 pour face et 1 pour pile) suivent toutes la même loi de Bernoulli. Bien que les lancers soient successifs, le résultats d'un lancé donné ne dépend pas des résultats précédents et n'aura aucune influence sur les prochains lancers (pas de lien de dépendance)

Ce sont des variables aléatoires qui suivent toutes la même loi de probabilité (donc même moyenne et même variance) et sont indépendantes.



On en déduit les propriétés suivantes:

- $E\{x_i\} = m \quad \forall i$
- $E\{(x_i - m)^2\} = \sigma^2$
- $E\{(x_i - m)(x_k - m)\} = 0 \quad \text{pour } k \neq i \text{ (indépendants et centrés)}$

### Exemple d'application

Supposons que l'on observe  $N$  échantillons indépendants  $Y_1, Y_2, \dots, Y_n$  d'une variable aléatoire  $Y$  de moyenne  $m = E\{Y_i\}$  et de variance  $\sigma^2 = E\{Y_i^2\} - m^2$ .

1. On désire estimer  $X = m$  (déterministe) en supposant  $\sigma^2$  connue. Pour cela, on étudie les deux estimateurs suivants :

$$X_1 = \frac{1}{n} \sum_{i=1}^n Y_i \qquad X_2 = \frac{2}{n(n+1)} \sum_{i=1}^n i \cdot Y_i$$

- Biais de  $X_1 = E\{\hat{X}_1\} - X = 0$
- Variance de  $X_1 = \sigma^2/n$
- Biais de  $X_2 = E\{\hat{X}_2\} - X = 0$
- Variance de  $X_2 = 2\sigma^2(2n+1)/3n(n+1)$

2. On désire estimer  $X = \sigma^2$  en supposant  $m$  connue. Pour cela, on étudie les deux estimateurs suivants :

$$X_1 = \frac{1}{n} \sum_{i=1}^n (Y_i - m)^2 \text{ et } X_2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \mu_Y)^2$$

- Biais de  $X_1 = E\{\hat{X}_1\} - X = 0$
- Biais de  $X_2 = E\{\hat{X}_2\} - X = (n-1) \cdot \sigma^2/n$

### **3. Estimateurs des moyennes statistiques**

Les quantités  $E\{x\}$ ,  $E\{x^2\}$ ,  $S_{xx}(f)$ ,... sont impossibles à calculer sur un ordinateur car elles nécessiteraient un nombre de points infini. Sur calculateur, on ne dispose que d'une séquence discrète et finie de  $N$  points. En réalité, on calcule des estimées de ces grandeurs en supposant généralement le signal stationnaire et ergodique. On remplace alors le calcul des moyennes statistiques par des moyennes temporelles.

### Estimation de Moyenne

Soit  $N$  échantillons  $(x_0, x_1, \dots, x_{N-1})$  indépendants et identiquement distribuées (même loi avec même paramètre) d'un signal aléatoire stationnaire. On définit l'estimée  $\hat{m}$  de  $E\{x\}$

$$\hat{m} = \frac{1}{N} \sum_{k=0}^{N-1} x_k$$

On remarque donc qu'on remplace la moyenne statistique par la moyenne temporelle sur la séquence finie. Cet estimateur est non biaisé. En effet, du fait que le signal soit iid,  $\forall k; E\{x_k\} = m$ , alors :

$$E\{\hat{m}\} = \frac{1}{N} \sum_{k=0}^{N-1} E\{x_k\} = \frac{1}{N} \cdot N \cdot m = m \Rightarrow b\{\hat{m}\} = 0$$

Pour la variance d'estimation, on obtient :

$$\begin{aligned}\sigma_m^2 &= \text{var}\{\hat{m}\} = E\{(\hat{m} - E\{\hat{m}\})^2\} = E\{\hat{m}^2\} - E\{\hat{m}\}^2 \\ &= E\left\{\left(\frac{1}{N} \sum_{k=0}^{N-1} x_k - m\right)^2\right\} = E\left\{\left(\frac{1}{N} \sum_{k=0}^{N-1} x_k - \frac{1}{N} \sum_{k=0}^{N-1} m\right)^2\right\} = \frac{1}{N^2} E\left\{\left(\sum_{k=0}^{N-1} (x_k - m)\right)^2\right\}\end{aligned}$$

Sachant que les  $x_k$  sont indépendants et que  $E\{x_k - m\} = 0$ , alors :  $\sigma_m^2 = \frac{1}{N^2} \sum_{k=0}^{N-1} (x_k - m)^2 = \sigma^2 / N$

Cet estimateur est consistant, puisque pour  $N \rightarrow \infty$ , biais et variance tendent vers 0.

### Estimateurs de variance

Selon que l'on connaisse ou non la moyenne, on utilisera l'un des deux estimateurs suivants :

$$\hat{\sigma}_1^2 = \frac{1}{N} \sum_{k=0}^{N-1} (x_k - m)^2 \quad \hat{\sigma}_2^2 = \frac{1}{N-1} \sum_{k=0}^{N-1} (x_k - \hat{m})^2$$

Calculons le biais des deux estimateurs

- Pour le premier, on a :  $E\{\hat{\sigma}_1^2\} = \frac{1}{N} \sum_{k=0}^{N-1} E\{(x_k - m)^2\} = \sigma^2 \quad \Rightarrow \quad b_{\hat{\sigma}_1^2} = 0$

- Pour le deuxième, on a :  $E\{\hat{\sigma}_2^2\} = \frac{1}{(N-1)} \sum_{k=0}^{N-1} E\{(x_k - \hat{m})^2\} = \frac{1}{(N-1)} \sum_{k=0}^{N-1} E\{(x_k - m + m - \hat{m})^2\}$

$$E\{\hat{\sigma}_2^2\} = \frac{1}{(N-1)} \sum_{k=0}^{N-1} E\left\{\left(x_k - m + m - \frac{1}{N} \sum_{j=0}^{N-1} x_j\right)^2\right\} = \frac{1}{(N-1)} \sum_{k=0}^{N-1} E\left\{\left((x_k - m) - \frac{1}{N} \sum_{j=0}^{N-1} (x_j - m)\right)^2\right\}$$

Sachant que  $E\{(x_k - m)\} = 0$ , on a :

$$\begin{aligned}E\{\hat{\sigma}_2^2\} &= \frac{1}{(N-1)} \sum_{k=0}^{N-1} \left( E\{(x_k - m)^2\} + \left(\frac{1}{N} \sum_{j=0}^{N-1} (x_j - m)\right)^2 - \frac{2}{N} E\{(x_k - m)^2\} \right) \\ &= \frac{1}{(N-1)} \sum_{k=0}^{N-1} \left( \sigma^2 + \frac{1}{N^2} \sum_{j=0}^{N-1} E\{(x_j - m)^2\} - \frac{2\sigma^2}{N} \right) = \frac{1}{(N-1)} \sum_{k=0}^{N-1} \left( \sigma^2 + \frac{\sigma^2}{N} - \frac{2\sigma^2}{N} \right) = \frac{N\sigma^2}{(N-1)} \left( \frac{N-1}{N} \right) = \sigma^2\end{aligned}$$

Les deux estimateurs ne sont pas biaisés puisque on obtient la vraie valeur à chaque fois.

### Estimateurs de la corrélation

Dans le cas discret, la fonction d'autocorrélation d'un signal aléatoire  $x$ ; supposé ergodique, est définie

par :  $R_{xx}(k) = E\{x(n)x^*(n-k)\} = E\{x(n)x^*(n-k)\} = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^{n=N} x(n)x^*(n-k)$

A partir d'un échantillon  $(x_0, x_1, \dots, x_{N-1})$  indépendants et identiquement distribués, plusieurs estimateurs sont alors possibles :

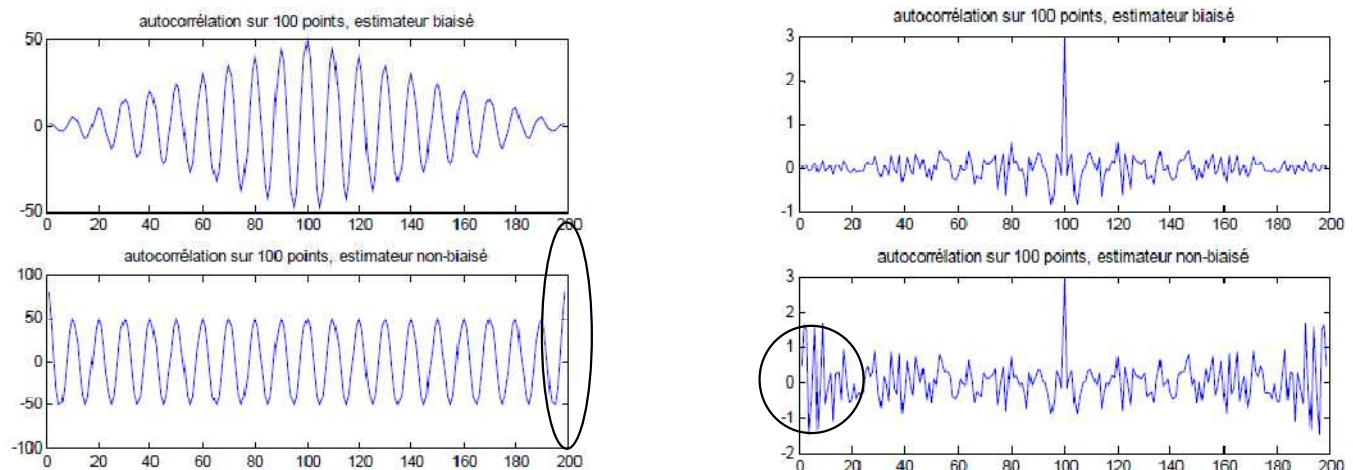
$$\hat{R}_{xx}(k) = \frac{1}{N} \sum_{n=k}^{N-1} x(n)x^*(n-k) \quad \hat{R}_{xx}(k) = \frac{1}{N-k} \sum_{n=k}^{N-1} x(n)x^*(n-k)$$

Le premier estimateur est biaisé car on ne tient pas compte du nombre d'échantillons disponibles qui varie avec le pas  $k$ . Le deuxième en tient compte, il est non biaisé. En effet :

$$E\{\hat{R}_{xx}(k)\} = \frac{1}{N-k} \sum_{n=k}^{N-1} E\{x(n)x^*(n-k)\} = \frac{1}{N-k} \sum_{n=k}^{N-1} R_{xx}(k) = R_{xx}(k)$$

On peut démontrer que la variance de cet estimateur tend vers 0 quand  $N$  tend vers l'infini. Cet estimateur est donc consistant.

On observe par contre que lorsque  $k$  approche du nombre d'échantillons  $N$ , la variance de l'estimateur non biaisé devient excessive (voir exemple pour un bruit blanc et une sinusoïde). Alors que celle de l'estimateur biaisé varie beaucoup moins. C'est une des raisons pour lesquelles cet estimateur est souvent utilisé par la suite, malgré son biais.



### Estimateurs de la densité spectrale

#### a. Méthode du périodogramme

Supposons que l'on dispose d'une séquence de  $N$  points du signal aléatoire :  $x [x_0, x_1, \dots, x_{N-1}]$ .

Soit donc la séquence  $y_k$  obtenue à partir de la séquence  $x_k$  pondérée par la fenêtre  $f_k$  :  $y_k = x_k f_k$

La transformée de Fourier calculée sur la séquence finie sera donc convoluée avec le spectre de la fenêtre rectangulaire c'est-à-dire un sinus cardinal. Rappelons que les propriétés spectrales du sinus cardinal ne sont pas bien adaptées à l'analyse spectrale du signal (lobes secondaires importants). Pour y remédier, on emploie des fenêtres de pondérations (Hamming, Hanning, Kaiser, etc.)

La méthode du périodogramme à prendre la TF d'une réalisation :

$$\hat{S}_{xx}(f) = \frac{1}{N} |Y(f)|^2 \quad \text{avec} \quad Y(f) = \sum_{k=0}^{N-1} y_k e^{-2\pi j f k}$$

$$\begin{aligned}
 \text{D'où } \hat{S}_{xx}(f) &= \frac{1}{N} \sum_{l=0}^{N-1} y_l e^{-2\pi j f l} \sum_{m=0}^{N-1} y_m e^{2\pi j f m} \frac{1}{N} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} y_l y_m e^{-2\pi j f (l-m)} \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=k}^{N-1} y_l y_{l-k} e^{-2\pi j f k} = \sum_{k=0}^{N-1} \hat{R}_{yy}(k) e^{-2\pi j f k} = TF(\hat{R}_{yy}(k))
 \end{aligned}$$

Le calcul de l'espérance de  $\hat{S}_{xx}(f)$  donne :

$$\begin{aligned}
 E\{\hat{S}_{xx}(f)\} &= E\left\{\sum_{k=0}^{N-1} \hat{R}_{yy}(k) e^{-2\pi j f k}\right\} = E\left\{\frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=k}^{N-1} y_l y_{l-k} e^{-2\pi j f k}\right\} \\
 E\{\hat{S}_{xx}(f)\} &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=k}^{N-1} E\{y_l y_{l-k}\} e^{-2\pi j f k} = \sum_{k=0}^{N-1} \frac{N-k}{N} R_{yy}(k) e^{-2\pi j f k} = S_{yy}(f) * N \text{sinc}(fN)^2
 \end{aligned}$$

Le périodogramme est un estimateur biaisé. Le périodogramme est donc en moyenne la convolution du véritable spectre avec la transformée de Fourier de la fenêtre triangulaire. Néanmoins, lorsque  $N \rightarrow \infty$ , le biais devient nul. La variance est pratiquement indépendante de  $N$  et proportionnelle au spectre  $\approx S_x(f)^2$

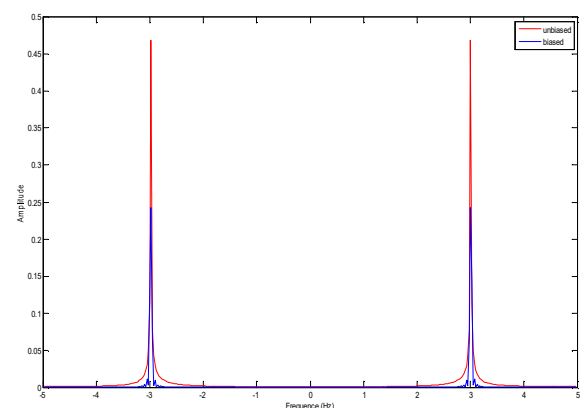
Afin de diminuer la variance de cet estimateur, on peut utiliser un périodogramme moyenné. Cela consiste à séparer le signal en  $K$  tranches (de longueur  $N/K$ ), à calculer le périodogramme sur chaque tranche et à faire la moyenne.

Du fait des  $K$  moyennages, la variance est presque divisée par  $K$  : néanmoins, les tranches étant plus courtes, la résolution diminue. En pratique, un taux de recouvrement de 40% donne de bons résultats. Au delà, l'indépendance entre les tranches n'est plus respectée.

b. La méthode du corrélogramme consiste à calculer d'abord l'estimée de  $R_{xx}(k)$  de la fonction d'autocorrélation puis à prendre pour estimée de la densité spectrale la TF de cette estimée.

$$\hat{S}_{xx}(f) = TF[\hat{R}_{xx}(k)]$$

Les corrélogrammes présentent chacun un extrema pour la fréquence normalisée 3 qui est bien la fréquence de la sinusoïde étudiée. Il ressort également de ces courbes que le corrélogramme non-biaisé permet de mieux faire ressortir la fréquence de la sinusoïde mais au dépend d'une plus grande variance sur les bords.



c) Dans cette méthode, on recherche un modèle AR, MA ou ARMA pour la séquence  $x_k$ . Dans le cas d'un modèle auto-régressif  $x_k = a_1 x_{k-1} + \dots + a_r x_{k-r} + u_k$

$$\hat{S}_{xx}(f) = \left| \frac{1}{1 + \sum_{k=1}^r a_k e^{-i2\pi f k}} \right|^2 u_0$$

## TD n°5 : Notions d'estimation

1. Les éléments d'une population possèdent un caractère  $X$  qui suit une loi de probabilité dont la densité est donnée par :

$$p_{\theta}(x) = \begin{cases} \frac{kx^3}{\theta^4} & \text{si } 0 < x < \theta \\ 0 & \text{sinon} \end{cases} \quad \text{Où } k = \text{cste et } \theta > 0 \text{ est le paramètre inconnu.}$$

- Déterminer la constante  $k$ .

- Montrer que l'estimateur  $\frac{1}{N} \sum_{i=1}^N X_i$  de  $\theta$  est biaisé et en déduire un non biaisé.

2. Soit  $(X_1, X_2, \dots, X_N)$  avec  $N > 10$  un échantillon aléatoire issu d'une population suivant une loi de Bernoulli de paramètre  $p$ .  $p(x) = p \cdot \delta(x-1) + (1-p) \cdot \delta(x)$ . Considérons les trois estimateurs du paramètre  $p$  :

$$\hat{p}_1 = \left( \sum_{i=1}^N X_i - 1 \right) / N \quad \hat{p}_2 = \frac{1}{N/2} \sum_{i=1}^{N/2} X_{2i} \quad \hat{p}_3 = \frac{1}{10} \sum_{i=1}^{10} X_i$$

- Comparer les trois estimateurs (consistance)

3. Soit  $(X_1, X_2, \dots, X_N)$  une population de moyenne  $\mu$  et de variance  $\sigma^2$ . On considère les deux estimateurs de la moyenne :

$$\mu_1 = \frac{1}{N} \sum_{i=1}^N X_i \quad \mu_2 = \sum_{i=1}^N a_i X_i / \sum_{i=1}^N a_i \quad a_i \geq 1$$

- Montrer que ces deux estimateurs sont non biaisés.

- Lequel sera le plus favorable à un étudiant ?

4. Soient  $(X_1, X_2, \dots, X_N)$ ,  $N$  variables aléatoires indépendantes suivant des lois de poisson de paramètre  $\lambda$ . Soient les estimateurs suivants du paramètre  $\lambda$  suivant :

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^N x_i}{N} \quad \hat{\lambda}_2 = \frac{x_1 + x_N}{2}$$

- Que signifient le biais et la variance pour un estimateur ?
- Calculer biais et variance des 2 estimateurs.
- Lequel est le plus efficace et pourquoi ?

Solutions

$$1. k=4 \quad b=(4\theta/5)-\theta \quad \theta = \frac{5}{4N} \sum_{i=1}^N x_i$$

$$2. p(x) = p\delta(x-1) + (1-p)\delta(x) \quad \mu_x = p \quad \sigma_x^2 = p(1-p)$$

- |         |         |                               |                                       |                |
|---------|---------|-------------------------------|---------------------------------------|----------------|
| - $p_1$ | $b=1/N$ | $\sigma^2 = p(1-p)/N$         | $(=0 \quad N \rightarrow \infty)$     | consistant     |
| - $p_2$ | $b=0$   | $\sigma^2 = 2 \cdot p(1-p)/N$ | $(=0 \quad N \rightarrow \infty)$     | consistant     |
| - $p_3$ | $b=0$   | $\sigma^2 = p(1-p)/10$        | $(\neq 0 \quad N \rightarrow \infty)$ | non consistant |
- $p_1$  meilleur estimateur

$$3. \mu_1 \text{ voir cours} \quad \mu_2: b=0, \sigma^2 = \sum_{i=1}^n a_i^2 \sigma^2 / \left( \sum_{i=1}^n a_i \right)^2$$

4.  $b_{\lambda 1}=0 \quad b_{\lambda 2}=0 \quad \sigma_{\lambda 1}^2 = \sigma^2/N \quad \sigma_{\lambda 2}^2 = \sigma^2/2$  Le premier (variance plus petite pour  $N > 2$ )

## Exercices supplémentaires

1. Soit  $X$  une variable aléatoire qui suit une loi Binomiale  $B(16; p)$ , où  $p$  est un paramètre inconnu. On se propose d'estimer  $p$  à l'aide d'un échantillon de taille  $n$  de  $X$  :  $(x_1; \dots; x_n)$ . On se propose d'étudier les estimateurs suivants :

$$T_1 = \frac{X_1 + \dots + X_n}{n} ; T_2 = \frac{X_1 + \dots + X_n}{16n} ; T_3 = \frac{1}{16} \left[ X_1 + \frac{X_2 + \dots + X_n}{n-1} \right]$$

- Dire lesquels sont sans biais, convergents.
- Lequel doit-on employer de préférence?

2. Soit  $D$  la variable aléatoire représentant la durée d'une communication téléphonique. On suppose que la loi de  $D$  est la loi Uniforme  $U$  sur  $[0, \theta]$ , pour  $\theta > 0$ . On veut estimer  $\theta$ , pour cela, on observe un  $n$  échantillon : soit les durées  $D_i$  de  $N$  communications. Etudier l'estimateur suivant :  $\theta = \frac{2}{N} \sum_{i=1}^N D_i$

- Calculer son biais, sa variance, est-il consistant?

3. Soient  $s_1^2, s_2^2, \dots, s_k^2$  les  $k$  variances-échantillons obtenues à partir de  $k$  échantillons aléatoires simples indépendants de tailles respectives  $n_1, n_2, \dots, n_k$ .

1.  $U = \frac{n_1 s_1^2 + n_2 s_2^2 + \dots + n_k s_k^2}{n_1 + n_2 + \dots + n_k}$  est-il un estimateur non biaisé de  $\sigma^2$  ?
2. Si ce n'est pas le cas, comment le biais peut-il être enlevé ?

4. Soient les variables aléatoires  $X_i (i = 1; 2)$  i.i.d. de moyenne  $\mu$  et de variance  $\sigma^2$ . Lequel des deux estimateurs de  $\mu$  non biaisés suivants conseilleriez-vous :

$$M_1 = \frac{X_1 + X_2}{2} \text{ ou } M_2 = \frac{aX_1 + bX_2}{a + b},$$

où  $a, b \in \mathbb{R}$  ?

## TP n°5 : Notions d'estimation

But : Aborder les notions d'estimation ainsi que les propriétés des estimateurs (biais, variance) à travers un exemple puis étudier les estimateurs de la moyenne, de la variance, de la corrélation et de la DSP.

**Exercice 1:** On reprend les 3 estimateurs de la valeur d'une résistance R donnés comme suit :

$$\hat{R}_1 = \frac{\sum_{k=1}^N U(k)I(k)}{\sum_{k=1}^N I^2(k)} \quad \hat{R}_2 = \frac{1}{N} \sum_{k=1}^N \frac{U(k)}{I(k)} \quad \hat{R}_3 = \frac{\sum_{k=1}^N U(k)}{\sum_{k=1}^N I(k)}$$

Le programme suivant permet de simuler des courants et tensions bruitées et d'étudier les 3 estimateurs.

```
clc;clear all; close all;
randn('state', sum(100*clock));N=1000; it=1:N;R=1;
for k=2:N
    U=1+0.22*randn(1,k); I=1+0.22*randn(1,k);
    R1(k)=sum(U.*I)/sum(I.^2);    R2(k)=sum(U./I)/k;    R3(k)=sum(U)/sum(I);
end
figure; plot(it(1:100),R1(1:100),'-r',it(1:100),R2(1:100),'g',it(1:100),R3(1:100),'-b');
legend('R1','R2','R3');
figure; plot(it,R1,'-r',it,R2,'g',it,R3,'-b');legend('R1','R2','R3');
for k=2:N
    biais_R1(k)=mean(R1(1:k))-R;biais_R2(k)=mean(R2(1:k))-R;biais_R3(k)=mean(R3(1:k))-R;
end
figure ;subplot(311); plot(biais_R1); grid; legend('Biais de R1');
subplot(312); plot(biais_R2); grid; legend('Biais de R2');
subplot(313); plot(biais_R3);grid; legend('Biais de R3');
for k=2:N
    var_R1(k)=var(R1(1:k));var_R2(k)=var(R2(1:k));var_R3(k)=var(R3(1:k));
end
figure ;subplot(311); plot(var_R1); legend('Var de R1');
subplot(312); plot(var_R2); legend('Var de R2');subplot(313); plot(var_R3);legend('Var de R3');
```

1. Expliquer toutes les boucles.
2. Etudier le biais et la variance de chaque estimateur. Conclure.
3. Lesquels sont consistants? Lequel est le plus efficace?

**Exercice 2:** Utiliser le programme ci-dessus pour construire les estimateurs de la moyenne et de la variance

$$\hat{m} = \frac{1}{N} \sum_{k=0}^{N-1} x_k \quad \hat{\sigma}_1^2 = \frac{1}{N} \sum_{k=0}^{N-1} (x_k - m)^2 \quad \hat{\sigma}_2^2 = \frac{1}{N-1} \sum_{k=0}^{N-1} (x_k - \hat{m})^2$$

1. Prendre  $m=1$  et  $\sigma^2=2$ . Etudier le biais, la variance et la consistance de l'estimateur de la moyenne
2. Quelle est la différence entre les deux estimateurs de la variance ?
3. Comparer les 2 estimateurs de la variance sur les 100èmes valeurs. Quel est le biais de chacun d'eux ?
4. Que remarque-t-on lorsque N augmente ? Les 3 estimateurs sont-ils efficaces?

**Exercice 3 :** Les deux estimateurs classiques de la fonction d'autocorrélation sont donnés par :

$$\hat{R}_{kk}(k) = \hat{R}_{kk}(-k) = \frac{1}{N} \sum_{i=1}^{N-k} x(n)x(n+k) \quad \text{et} \quad \hat{R}_{kk}(k) = \hat{R}_{kk}(-k) = \frac{1}{N-k} \sum_{i=1}^{N-k} x(n)x(n+k)$$

```
clc; clear all; close all;
randn('state', sum(100*clock));
N = 200; f0=1; Fe=5*f0; Te=1/Fe; t=(1:N)*Te; m=0; v=2;tt=(-N+1:N-1)*Te;ff=Fe/2*(-N+1:N-1)/N;
bb=m+sqrt(v)*randn(1,N);Rbb1=xcorr(bb,'unbiased');Rbb2=xcorr(bb,'biased');
ss=cos(2*pi*f0*t); Rss1=xcorr(ss,'unbiased');Rss2=xcorr(ss,'biased');
figure ; subplot(2,1,1); plot(tt,Rbb1); subplot(2,1,2) ; plot(tt,Rbb2);
figure ; subplot(2,1,1); plot(tt,Rss1); subplot(2,1,2) ; plot(tt,Rss2);
figure;plot(ff,abs(fftshift(fft(Rss1)/N)), 'r',ff,abs(fftshift(fft(Rss2)/N)), 'b');
```



```
legend('unbiased','biased'); xlabel('Frequence (Hz)'); ylabel('Amplitude');
```

1. Des 2 estimations classiques de l'autocorrélation pour le bruit, laquelle paraît la plus satisfaisante ?
2. Même question pour l'autocorrélation de la sinusoïde
3. Observer la DSP dans chaque et commenter en justifiant les différences (amplitude, lobes secondaires, etc.)

**Exercice 4 :** Il existe différentes méthodes d'estimation de la densité spectrale de puissance. Les méthodes les plus couramment utilisées sont :

- Périodogramme simple:  $\hat{S}_{xx}(f) = \frac{1}{N} |X_T(f)|^2$
- Corrélogramme: TF de la corrélation  $\hat{S}_{xx}(f) = TF\{\hat{R}_{xx}(k)\}$
- Périodogramme moyenné: Une manière de réduire la variance de l'estimateur est de subdiviser l'intervalle de définition du signal  $x(n)$  en un certain nombre de sous-intervalles. On calcule alors la moyenne des périodogrammes sur chacun des sous-intervalles (Méthode de Welch).
- Périodogramme moyenné avec recouvrement (overlapping): Cette fois-ci, on considère des sous intervalles de mêmes longueurs que précédemment mais avec des recouvrements d'une demi-longueur.

Le programme suivant permet d'estimer la DSP par les deux première méthodes

```
clc; close all; clear all;
N = 1024; Te = 1/N; Fe=N; t = (0:N-1)*Te; f1 = 100; f2 = 150; A1 = 2; A2=1.2; A3=4;
x = A1*cos(2*pi*f1*t) + A2*cos(2*pi*f2*t)+A3*randn(1,N); figure;plot(t,x);

% periodogramme
Sx1=abs(fft(x).^2)/N; ff1= Fe*(0:N/2-1)/N;

% spectre de l'autocorrélation
Rx=xcorr(x,'biased'); Sx2=fft(Rx); N2=length(Rx);
ff2=Fe*(0:N2/2-1)/N2;

% Méthode de welch
T=256; SXXF=[];k=1;
for ii=1:T:N-T
    xt=x(ii+1:ii+T); Sx3(k,:)=abs(fft(xt,N)).^2/T; k=k+1;
end
Moy_Sx3=mean(Sx3);

% Périodogramme moyenné avec recouvrement
T=256; Sx4=[];k=1;
for ii=1:T/2:N-T
    xt=x(ii+1:ii+T); Sx4(k,:)=abs(fft(xt,N)).^2/T; k=k+1;
end
Moy_Sx4=mean(Sx4);

%affichage
figure; plot(ff1,Sx1(1:N/2),'b',ff2,abs(Sx2(1:N2/2)),'r');
legend('Périodogramme','Correlogramme'); xlabel('Frequence (Hz)'); ylabel('Amplitude');
AXIS([min(ff2) max(ff2) min(Sx1) max(Sx1)]);
figure; plot(ff1,Sx1(1:N/2),'b',ff1,Moy_Sx3(1:N/2),'r');
legend('Périodogramme','Périodogramme moyenné'); xlabel('Frequence (Hz)');
ylabel('Amplitude');
figure; plot(ff1,Sx1(1:N/2),'b',ff1,Moy_Sx4(1:N/2),'r');
legend('Périodogramme','Périodogramme moyenné avec recouvrement'); xlabel('Frequence (Hz)');
ylabel('Amplitude');
figure; plot(ff1,Moy_Sx3(1:N/2),'b',ff1,Moy_Sx4(1:N/2),'r');
legend('Périodogramme moyenné','Périodogramme moyenné avec recouvrement'); xlabel('Frequence (Hz)'); ylabel('Amplitude');
```

1. Comparer les 4 techniques en faisant varier A1, A2 et A3 et commenter.
2. Rapprocher les fréquences et observer.
3. Commenter les quatre techniques d'estimation de la DSP en faisant varier N.
4. Quelle est l'avantage et l'inconvénient du moyennage?

## Chapitre 6 : Les estimateurs usuels

Nous avons au chapitre précédent que le but des techniques d'estimation est d'utiliser les observations  $X$  (aléatoires) pour extraire de l'information sur la grandeur d'intérêt  $Y$ . Suivant que  $Y$  soit déterministe ou aléatoire, on fera appel à différentes techniques. Ainsi :

- Pour estimer une variable certaine, dans le cas le plus général, on souhaite que l'estimateur soit non biaisé, on cherche donc l'estimateur à minimum de variance dont le plus connue est le maximum de vraisemblance (MV) basée sur l'emploi de la probabilité des valeurs observées  $p(X/Y)$ . Si on autorise un biais, on cherche alors à minimiser l'erreur quadratique moyenne.
- Lorsque  $Y$  la variable aléatoire à estimer est aléatoire, on aura recours aux estimateurs Bayésiens. Plusieurs cas de figures sont alors possibles :
  - Lorsqu'on suppose  $p(Y)$  connue, on utilisera le maximum à postériori (MAP).
  - Si seules sont connues les moments d'ordre 1 et 2 de  $p(Y)$  et  $p(X)$ , on utilise alors souvent l'estimateur linéaire non-biaisé à variance minimale.
- Dans le cas où on ne possède aucune information statistique sur  $X$  et  $Y$  et que la seule dont on dispose est que  $X$  est une mesure bruitée de  $fct(Y)$ , on adoptera l'estimateur des moindres carrés pour estimer  $Y$  qui s'affranchit de tout cadre probabiliste.

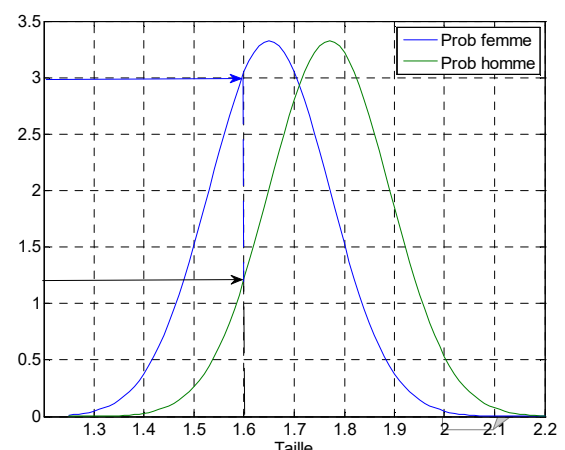
### 1. Estimateur du maximum de vraisemblance

On utilise cette estimateur lorsque que l'on cherche à déterminer une variable  $Y$  déterministe mais inconnue dépendant de mesures aléatoires (observations)  $X : (x_1, \dots, x_n)$ .

Nous avons vu, au chapitre précédent qu'il était possible de déterminer les moments statistiques en se basant sur la loi des grands nombres. Le maximum de vraisemblance est une technique plus générale pour construire un estimateur (dont la variance est la plus petite possible) des ses moments lorsque l'on dispose de la probabilité de la loi ayant engendré les mesures aléatoires.

Exemple : Supposons que lors d'une enquête policière, un suspect inconnu (femme ou homme) mesure environ 1.60 m, on aura plutôt tendance à rechercher une femme tandis que s'il mesure environ 1.80 m, on recherchera plutôt un homme. La notion de maximum de vraisemblance permet de formaliser cette intuition. On peut modéliser la distribution des tailles (en mètres) féminines par une loi gaussienne d'espérance  $\mu_F = 1,65$  et d'écart type  $\sigma_F = 0,12$  et celle des tailles masculines par une loi gaussienne d'espérance  $\mu_H = 1,75$  et d'écart type  $\sigma_H = 0,13$ .

Les densités de ces deux lois sont représentées sur la figure ci-contre. Lorsque l'on connaît la taille  $x$  d'un suspect, on pourra supposer que ce suspect est une femme si la densité des tailles féminines prise en  $x$  est supérieure à celle des tailles masculines et vice et versa.



On peut estimer le genre de cet individu en choisissant  $X \in \{F, H\}$  qui maximise la vraisemblance

$$X \rightarrow \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(Y-\mu_X)^2}{2\sigma_X^2}}$$

Définition : On suppose que pour toute réalisation  $x = (x_1, \dots, x_n)$  de l'échantillon  $X$  de densité  $p(x/y)$ , il existe une unique valeur  $y$  qui maximise la vraisemblance de la réalisation  $x$ . La vraisemblance s'écrit :

$$L(x/y) = \prod_{i=1}^n p(x_i/y)$$

Ainsi, le principe de la vraisemblance revient à déterminer la valeur du paramètre  $y$  et ce en fonction des observations  $(x_1, \dots, x_n)$ ; qui assure la plus grande probabilité d'apparition de ces observations  $(x_1, \dots, x_n)$ .

Comme la fonction logarithme est strictement croissante, il revient au même de maximiser la vraisemblance  $Y \rightarrow L(x/y)$  que maximiser la log-vraisemblance  $Y \rightarrow \ln(L(x/y))$

Exemple 1 On a observé un échantillon indépendant de taille  $n$  :  $x = (x_1, \dots, x_n)$  issu d'une loi de Poisson de moyenne  $\mu$ . Trouvons l'estimateur du maximum de vraisemblance de  $\mu$ .

Rappelons que pour une loi de Poisson  $p_X(x) = \frac{\mu^x e^{-\mu}}{x!}$ ,  $m = \mu$ ,  $\sigma^2 = \mu$

D'abord, on applique le log à  $p(x_1, \dots, x_n, \mu)$ , puis on dérive par rapport à  $\mu$  pour trouver le maximum.

$$\ln(L(x_1, \dots, x_n | \mu)) = \ln(p(x_1, \dots, x_n | \mu)) = -n\mu + \ln(\mu) \cdot \sum_{i=1}^n x_i - \sum_{i=1}^n \ln(x_i!)$$

$$\frac{d\ln(L(x_1, \dots, x_n, \mu))}{d\mu} = -n + \frac{1}{\mu} \sum_{i=1}^n x_i \text{ d'où } \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \text{ soit la moyenne arithmétique de l'échantillon.}$$

Exemple 2 On réalise un ensemble de  $k$  mesures regroupées dans le vecteur  $r$ . Chaque mesure est de la forme  $r_i = a + n_i$ , où  $a$  est un paramètre déterministe inconnu et  $n_i$  sont des échantillons indépendants et identiquement distribués (iid) selon une loi gaussienne  $N(0, \sigma^2)$

Puisque les bruits  $n_i$  sont indépendants, on peut écrire :

$$L(r|a) = \prod_{i=1}^k p(r_i|a) = \prod_{i=1}^k \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(r_i-a)^2}{2\sigma^2}} = \frac{1}{(\sigma \sqrt{2\pi})^k} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^k (r_i - a)^2\right)$$

$$\text{Alors : } \ln(L(r|a)) = -k \ln(\sigma \sqrt{2\pi}) - \frac{1}{2\sigma^2} \sum_{i=1}^k (r_i - a)^2$$

$$\text{On dérive par rapport à } a, \text{ on trouve : } \frac{d\ln(L(r|a))}{da} = -\frac{1}{\sigma^2} \sum_{i=1}^k (r_i - a) = -\frac{1}{\sigma^2} \sum_{i=1}^k r_i + \frac{1}{\sigma^2} k.a$$

$$\Rightarrow a = \frac{1}{k} \sum_{i=1}^k r_i$$

## 2. Estimateurs Bayésiens

On fait appel aux estimateurs Bayésiens pour estimer une variable aléatoire notée  $y$  aléatoire possédant une loi à priori et dépendant des observations  $x$  de dimension  $n$  :  $(x_1, \dots, x_n)$ . La stratégie bayésienne consiste à estimer  $y$  de façon à minimiser l'écart entre la valeur estimée  $\hat{y}$  et la vraie valeur  $y$  à travers une fonction de coût  $C(y; \hat{y})$  associée à chaque estimation, l'estimation optimale étant alors celle assurant la minimisation de ce coût. Les choix les plus classiques de fonctions de coût sont alors les suivants :

1. Coût quadratique :  $C(y; \hat{y}) = (y - \hat{y})^2 \Rightarrow$  (estimateur moyenne conditionnelle a posteriori MMSE : Minimisation de l'erreur quadratique moyenne)
2. Coût en valeur absolue :  $C(y; \hat{y}) = |y - \hat{y}| \Rightarrow$  (estimateur médiane à posteriori)
3. Coût uniforme :  $C(y; \hat{y}) = 1 - 1_{\Delta}(y - \hat{y}) \Rightarrow$  (estimateur maximum à posteriori MAP)

Estimateur MAP L'estimateur du maximum a posteriori est obtenu en considérant les fonctions de coût égales à un, partout sauf sur un petit voisinage des valeurs de  $y$  où  $p(y/x)$  est maximale et où la fonction de coût vaut zéro. On obtient l'estimateur du maximum a posteriori en faisant tendre vers 0 la taille de ce voisinage. Minimiser le coût, revient à minimiser le risque de Bayes  $B(\hat{y}/x) = 1 - p(\hat{y}/x)$ .

Minimiser  $B(\hat{y}/x)$  par rapport à  $\hat{y}$  revient donc à maximiser la densité de probabilité  $p(\hat{y}/x)$ . L'estimateur qui maximise  $p(\hat{y}/x)$  est appelé l'estimateur du maximum à posteriori. Grâce à la formule de Bayes, la densité de probabilité conditionnelle peut se formuler :

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)}$$

Ainsi, maximiser  $p(y/x)$  revient à maximiser  $p(x/y)p(y)$  par rapport à  $y$ . On ne considère pas  $p(x)$  puisque cette probabilité ne dépend pas de  $y$ .

En outre, la fonction logarithme est monotone, cela revient à maximiser  $\ln(p(x/y)) + \ln(p(y))$

Exemple : Soit une séquence de  $N$  observations supposées indépendantes :  $y_i = a + n_i$  dans laquelle  $a$  suit une loi normale  $N(m; \sigma_n)$  et  $n_i \sim N(0; \sigma_n)$  pour tout  $i = 1, \dots, N$ .

L'observation  $y$  étant constituée par la collections des échantillons  $y_i$ . L'hypothèse d'indépendance et de normalité de ces derniers garantit que la densité  $p(y/a)$  se met sous la forme :

$$p(y|a) = \prod_{i=1}^N p(y_i|a) = \frac{1}{(\sigma_n \sqrt{2\pi})^N} \exp\left(-\frac{1}{2\sigma_n^2} \sum_{i=1}^N (y_i - a)^2\right) \text{ et } p(a) = \frac{1}{\sigma_a \sqrt{2\pi}} \exp\left(-\frac{(a-m)^2}{2\sigma_a^2}\right)$$

On en déduit que :

$$\ln(p(y|a)) + \ln(p(a)) = -N \ln(\sigma_n \sqrt{2\pi}) - \frac{1}{2\sigma_n^2} \sum_{i=1}^N (y_i - a)^2 - \ln(\sigma_a \sqrt{2\pi}) - \frac{1}{2\sigma_a^2} (m - a)^2$$

En dérivant cette expression par rapport à  $a$ , on obtient :

$$-\frac{1}{\sigma_n^2} \sum_{i=1}^N (y_i - a) - \frac{(m - a)}{\sigma_a^2} = -\frac{1}{\sigma_n^2} \sum_{i=1}^N y_i + \frac{1}{\sigma_n^2} N a - \frac{(m - a)}{\sigma_a^2}$$

Si on pose  $b = N\sigma_a^2 / \sigma_n^2$  alors  $\hat{a} = \frac{1}{1+1/b} \hat{\mu}_y + \frac{1}{1+b} m$

Cet estimateur  $\hat{a}$  au sens du MAP réalise une moyenne pondérée entre cette moyenne empirique  $\mu_y$  et la moyenne a priori  $m$ , sa valeur tendant vers l'une ou l'autre selon que le nombre des observations et/ou le rapport signal-sur-bruit est plus fort ou plus faible :

$$\lim_{b \rightarrow \infty} \hat{a} = \hat{\mu}_y \quad \lim_{b \rightarrow 0} \hat{a} = m$$

Remarque : L'estimateur de Bayes est en principe le plus intéressant car il utilise l'information maximale. Mais celle-ci (en particulier la densité a priori  $p(a)$ ) n'est pas toujours disponible, et il faut alors recourir à des estimateurs sous-optimaux. Remarquons en particulier que s'il n'y a pas d'information a priori sur  $a$  et que tous les  $a$  sont donc équiprobables (ce qui revient à dire que  $p(a)$  est uniforme), alors on revient à l'estimateur du maximum de vraisemblance.

### 3. Estimateur linéaire à variance minimale

Lorsque les probabilités précédentes  $p(y/x)$  et  $p(y)$  ne sont pas connues mais que l'on dispose des moments d'ordre 1 et 2 (moyennes supposées centrées des covariances  $R_{xy}$  et  $R_{xx}$ ), on utilise alors souvent l'estimateur linéaire non-biaisé à variance minimale. On recherche une estimée  $\hat{y}$  de  $y$  qui soit une fonction linéaire des observations, c'est-à-dire

$$\hat{y} = \sum_{i=1}^n h_i x_i = h^T x$$

Dans le cas de l'estimation en moyenne quadratique, le filtre  $h$  doit alors être déterminé de telle sorte que la variance de l'erreur d'estimation  $E\{(y - \hat{y})^2\}$  soit minimale :

$$E\{(y - \hat{y})^2\} = E\{(y - h^T x)^T (y - h^T x)\} = E\{y y^T\} - E\{y^T h^T x\} - E\{h x^T y\} + E\{h x^T h^T x\}$$

On commence par dériver par rapport à  $h$  puis on annule la dérivée, ce qui nous donne

$$-E\{y^T x\} - E\{x^T y\} + 2h^T E\{x^T x\} = 0 \Rightarrow h = E\{x^T y\} / E\{x^T x\} = R_{xy} / R_{xx}$$

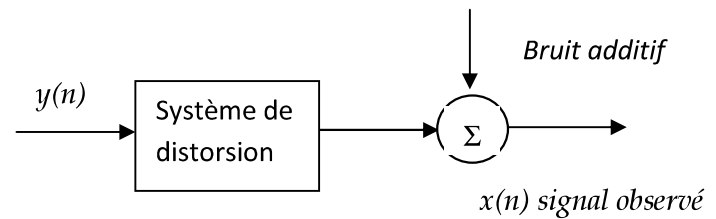
#### Application : Filtre de Wiener

Dans de nombreuses applications, les signaux temporels sont entachées d'un bruit que l'on souhaite supprimer ou du moins réduire. Comme le signal utile aléatoire occupe les mêmes bandes de fréquences

que le signal parasite, on ne peut recourir au filtrage classique. Le filtre de Wiener apporte une solution à ce problème lorsque le processus est stationnaire.

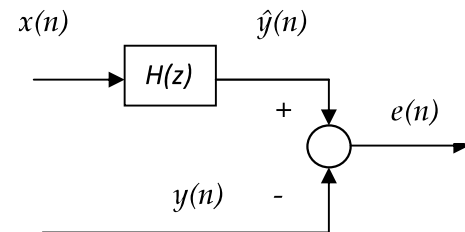
Les filtres de Wiener développés à partir de concepts temporels sont conçus pour minimiser l'erreur quadratique moyenne entre leur sortie et une sortie désirée. Ils sont dits optimum au sens du critère de l'erreur quadratique moyenne. Dans ce cas les coefficients du filtre sont liés à la fonction d'autocorrélation du signal d'entrée et à l'intercorrrelation entre les signaux d'entrée et de sortie désirée. Cela signifie ces fonctions sont considérées comme des fonctions aléatoires centrées stationnaires au second ordre.

La figure suivante illustre un problème courant d'estimation linéaire.  $y(n)$  correspond au signal aléatoire stationnaire qui nous intéresse mais n'est pas directement accessible. Seul  $x(n)$  l'est, il est obtenu après passage de  $y(n)$  dans un système



linéaire suivi de l'addition d'un bruit aléatoire stationnaire  $bb(n)$ .

Le problème qui se pose est comment retrouver  $y(n)$  à partir de  $x(n)$ . Une solution consiste à filtrer  $x(n)$  de telle sorte que la sortie  $\hat{y}(n)$  soit la plus proche possible de  $y(n)$ . On peut mesurer la qualité de l'estimation par  $e(n)$  défini par :  $e(n) = y(n) - \hat{y}(n)$



On cherche donc un filtre qui minimisera l'erreur. Il est pratique de chercher à minimiser  $e^2(n)$  car c'est une fonction quadratique facilement dérivable. Par ailleurs, étant donné que les signaux sont aléatoires, la fonction coût à minimiser est l'erreur quadratique moyenne (MSE) définie par :  $\xi(n) = E(e^2(n))$

Si on suppose que le filtre recherché  $H$  est un filtre RIF de longueur  $N$ , on peut en calculer les coefficients par résolution d'un système linéaire d'équations.  $h = [b_0 \ b_1 \ \dots \ b_{N-1}]^T$

Le signal estimé  $\hat{y}(n)$  peut alors s'écrire :  $\hat{y}(n) = \sum_{i=0}^{N-1} b_i x(n-i)$

Rappelons que l'on cherche à minimiser  $\xi(n) = E\{e^2(n)\} = E\left\{\left(y(n) - \sum_{i=0}^{N-1} b_i x(n-i)\right)^2\right\}$

Pour en obtenir le minimum, il suffit de chercher de dériver et d'annuler la fonction coût par rapport aux variables  $b_i$  de la réponse impulsionnelle du filtre. La dérivée de la fonction coût par rapport au  $j^{\text{ème}}$  point de la réponse impulsionnelle est donnée par :

$$\frac{\partial \xi}{\partial b_j} = E\left\{\frac{\partial}{\partial b_j}\{e^2(n)\}\right\} = E\left\{2e(n)\frac{\partial e(n)}{\partial b_j}\right\} = E\left\{2e(n)\frac{\partial}{\partial b_j}\{-b_j x(n-j)\}\right\}$$

$$\frac{\partial \xi}{\partial b_j} = -E\{2e(n)x(n-j)\} = -E\left\{2\left(y(n) - \sum_{i=0}^{N-1} b_i x(n-i)\right)x(n-j)\right\}$$

En faisant l'hypothèse que les signaux  $x(n)$  et  $y(n)$  sont stationnaires, et en annulant la dérivée, on

$$\text{trouve : } R_{yx}(j) = \sum_{i=0}^{N-1} b_i R_{xx}(j-i)$$

Ce qui pour les différentes valeurs de  $j$ , nous donne le système d'équations suivantes à résoudre :

$$\begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(N-1) \\ R_{xx}(1) & R_{xx}(0) & \dots & R_{xx}(N-2) \\ & & \dots & \\ R_{xx}(N-1) & R_{xx}(N-2) & \dots & R_{xx}(0) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{N-1} \end{bmatrix} = \begin{bmatrix} R_{xy}(0) \\ R_{xy}(1) \\ \vdots \\ R_{xy}(N-1) \end{bmatrix}$$

Application : Réduction de bruit (pour la transmission vocale, les images, etc.)

Hypothèses supplémentaires : On suppose que l'observation  $x(n)=y(n)+bb(n)$  et que le bruit additif  $bb(n)$  est centré et non corrélé au signal. On peut alors simplifier les équations de Wiener-Hopf :

$$R_{xy}(k) = E\{x(n)(y(n-k)+bb(n-k))\} = R_{yy}(k)$$

$$R_{xx}(k) = E\{(y(n)+bb(n))(y(n-k)+bb(n-k))\} = R_{yy}(k) + R_{bb}(k)$$

Soit le système à résoudre suivant :

$$\begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(N-1) \\ R_{xx}(1) & R_{xx}(0) & \dots & R_{xx}(N-2) \\ & & \dots & \\ R_{xx}(N-1) & R_{xx}(N-2) & \dots & R_{xx}(0) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{N-1} \end{bmatrix} = \begin{bmatrix} R_{xx}(0) - R_{bb}(0) \\ R_{xx}(1) - R_{bb}(1) \\ \vdots \\ R_{xx}(N-1) - R_{bb}(N-1) \end{bmatrix}$$

Exemple 1 Le signal  $x(n)$  est un processus AR dont la fonction d'autocorrélation est  $R_y(k) = \alpha^{|k|}$   $0 < \alpha < 1$

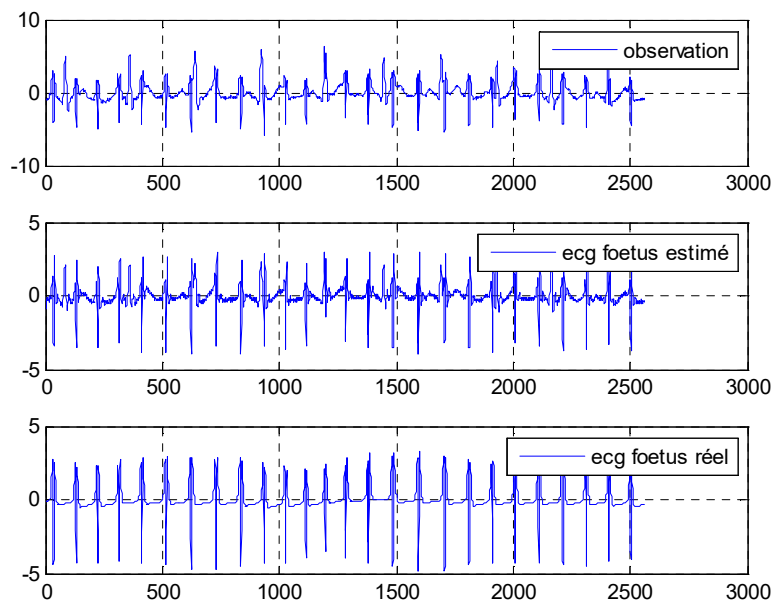
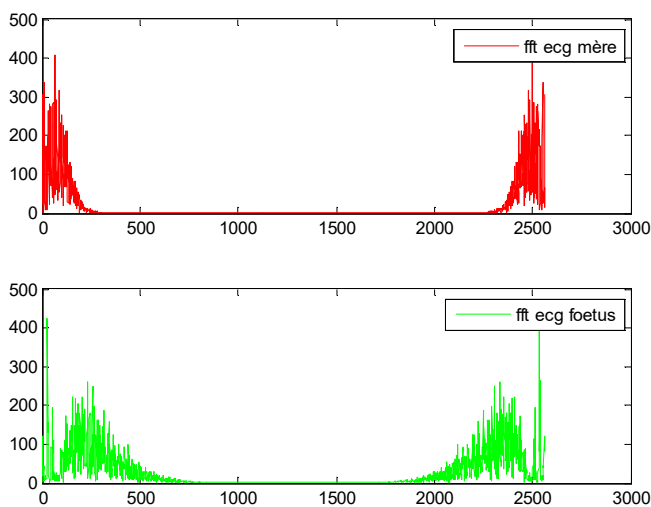
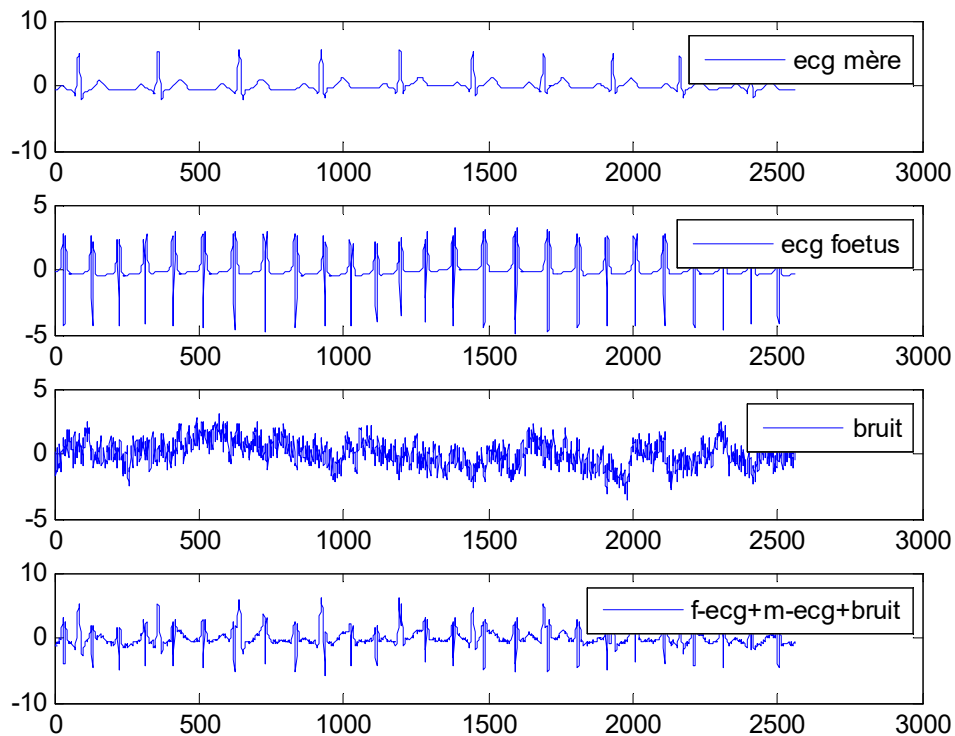
avec  $y(n) = x(n) + bb(n)$  où  $bb(n)$  est un bruit blanc

non corrélé de variance  $\sigma_b^2$  Cherchons  $h[n]$  tel que  $H(z) = b_0 + b_1 z^{-1}$ .

$$\begin{bmatrix} 1 + \sigma_b^2 & \alpha \\ \alpha & 1 + \sigma_b^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha \end{bmatrix} \quad H(z) = \frac{1}{(1 + \sigma_b^2) - \alpha^2} [(1 + \sigma_b^2 - \alpha^2) + \alpha \sigma_b^2 z^{-1}]$$

Exemple 2: Comme illustration du filtrage de Wiener, on prend souvent la mesure de l'activité cardiaque d'un fœtus à l'aide d'un électrocardiogramme (ECG) pris au niveau de l'abdomen de la mère (le signal  $x(n)$ ) et qui va naturellement être perturbé par l'ECG de celle-ci auquel se rajoute le bruit thermique des électrodes et des équipements électroniques. Pour retrouver l'ecg du fœtus, on réalise une deuxième mesure fournissant l'ECG de la mère (le signal  $bb$ ). On peut employer alors le filtre de Wiener pour estimer le signal  $y(n)$  représentant l'ECG du fœtus.





On peut observer que les TF des deux signaux ecg (mère et foetus) occupent la même plage de fréquences. A partir des auto-corrélations de l'ECG mesuré sur l'abdomen  $x(n)$  et celui de la mère  $bb(n)$ , on retrouve les paramètres  $b_i$  du filtre qu'on appliquera à  $x(n)$  pour obtenir une estimée de  $y(n)$  soit l'ECG foetal.

*Remarque* Quand les fonctions d'auto et d'intercorrélation ne sont pas connues (cas le plus courant), alors on va approcher le filtre optimal de Wiener en utilisant une boucle de retour et un algorithme de minimisation: c'est ce que l'on appelle **le filtrage adaptatif**. Dans ce cas, on remplacera la connaissance

des fonctions de corrélation par une phase d'apprentissage permettant de modifier itérativement la réponse impulsionnelle du filtre.

#### 4. Estimateurs au sens des moindres carrés

Dans le filtrage de Wiener, le critère d'optimalité est stochastique : on désire minimiser la moyenne de l'erreur au carré. Cela requiert des statistiques de second ordre des processus (moyennes et covariances). Une autre approche consiste à minimiser, non plus la moyenne stochastique, mais temporelle de cette erreur, c'est la méthode des moindres carrés. C'est la moins contraignante, aucun aspect probabiliste n'entre en jeu.

$\xi = \sum_{n=L}^N e^2(n)$  où L et N sont des indices qui représentent l'intervalle dans lequel la minimisation se fait.

On suppose connues les données  $\{x(0), x(1), \dots, x(N-1)\}$  et en adoptant à nouveau un filtre FIR de longueur L, on a

$$e(n) = y(n) - \sum_{i=0}^{L-1} b_i x(n-i), \text{ on cherche à minimiser } \xi = \sum_{n=L}^N e^2(n) = \sum_{n=L}^N \left( y(n) - \sum_{i=0}^{L-1} b_i x(n-i) \right)^2$$

Le signal d'entrée prend alors la forme suivante :

$$\begin{bmatrix} y(L) & y(L+1) & \dots & y(N) \\ y(L-1) & y(L) & \dots & y(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ y(1) & y(2) & \dots & y(N-L+1) \end{bmatrix}$$

On procède comme pour le filtre de Wiener, on dérive et annule les dérivées par rapport à  $b_i$  on obtient:

$$\frac{\partial \xi}{\partial b_j} = 2 \sum_{n=L}^N \frac{\partial e(n)}{\partial b_j} e(n) = -2 \sum_{n=L}^N x(n-i) e(n) = 0 \Rightarrow \sum_{n=L}^N x(n-i) \left[ y(n) - \sum_{k=0}^{L-1} b_k x(n-k) \right] = 0, \quad i = 0, 2, \dots, L-1$$

$$\text{d'où } \sum_{k=0}^{L-1} b_k \left[ \sum_{n=L}^N x(n-i)x(n-k) \right] = \sum_{n=L}^N x(n-i)y(n) \Rightarrow \sum_{k=0}^{L-1} b_k \overline{R_{xx}(k-i)} = \overline{R_{yx}(i)} \quad i = 0, 2, \dots, L-1$$

donc un système d'équations

$$\begin{bmatrix} \overline{R_{xx}(0)} & \overline{R_{xx}(1)} & \dots & \overline{R_{xx}(L-1)} \\ \overline{R_{xx}(1)} & \overline{R_{xx}(0)} & \dots & \overline{R_{xx}(L-2)} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{R_{xx}(L-1)} & \overline{R_{xx}(L-2)} & \dots & \overline{R_{xx}(0)} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{N-1} \end{bmatrix} = \begin{bmatrix} \overline{R_{yx}(0)} \\ \overline{R_{yx}(1)} \\ \vdots \\ \overline{R_{yx}(L-1)} \end{bmatrix}$$

Où  $R_{xx}$  et  $R_{yx}$  sont respectivement des autocorrélations et intercorrélations temporelles.

Remarque : Le filtrage des moindres carrés se place en déterministe. Alors que le filtre de Wiener utilise la méthode des moindres carrés en stochastique.

## TD n°6 : Estimateurs

1. Soient  $(X_1, X_2, \dots, X_N)$ ,  $N$  variables aléatoires indépendantes de même loi de densité :

$$p(x) = \beta e^{\beta(\alpha-x)} U(x-\alpha)$$

- $\alpha$  étant connu, déterminer l'estimateur du maximum de vraisemblance  $w=1/\beta$  notée  $\hat{w}$ .
- Vérifier qu'il est sans biais et convergent

2. Soient  $(X_1, X_2, \dots, X_N)$ ,  $N$  variables aléatoires indépendantes suivant une loi de probabilité définie par :

$$p(x; \theta) = \frac{1}{(r-1)!} \frac{x^{r-1}}{\theta^r} e^{-x/\theta} \quad \text{avec} \quad \text{moy} = r\theta \quad \text{et} \quad \text{var} = r\theta^2$$

- Dans quel cas utilise-t-on le maximum de vraisemblance ?
- Déterminer  $\hat{\theta}$  l'estimateur du maximum de vraisemblance de  $\theta$ .
- Calculer le biais
- Calculer la variance de  $\hat{\theta}$ .
- Cet estimateur est-il consistant ?

3. La même information binaire  $\theta \in \{0,1\}$  est transmise 2 fois consécutives vers un récepteur à travers un canal de transmission. Ces 2 informations sont perturbées par un bruit supposé Gaussien centré de variance  $\sigma^2$ . Le message reçu s'écrit alors  $z=(z_1, z_2)$  où  $z_i = \theta + e_i$  et  $e_i \in N(0, \sigma^2)$ . Le problème consiste à trouver le symbole émis à partir du message reçu  $z$ .

- Déterminer l'estimateur du maximum de vraisemblance du paramètre  $\theta$ .

On suppose que l'on dispose d'une information à priori sur les bits '0' et '1' qui se traduit par  $p(0)=p(1)=0.5$ .

- Déterminer l'estimateur du maximum à posteriori de  $\theta$ .
- Même question pour  $p(0)=q$  et  $p(1)=1-q$

4. On considère un processus de la forme  $X(n) = \theta + W(n)$  où  $W(n)$  est un processus gaussien de moyenne 0 et de variance 1 tel que  $W(n)$  et  $W(j)$  sont indépendants si  $n \neq j$ . On suppose que  $\theta$  suit une loi  $N(0, \sigma^2)$  indépendante de  $W(n)$   $n \in \mathbb{Z}$ .

- Caractériser le filtre de Wiener permettant d'estimer  $\theta$  à partir de  $X_n, X_{n-1}, \dots, X_{n-p+1}$ .
- Calculer les coefficients du filtre.

5. On considère un processus de la forme  $X(n) = b(n) + ab(n-1) + W(n)$  où  $W(n)$  est un processus gaussien de moyenne 0 et de variance  $\sigma^2$  tel que  $W(n)$  et  $W(j)$  sont indépendants si  $n \neq j$ . On suppose que  $b(n)$  est une variable aléatoire uniforme à valeurs dans  $\{-1, 1\}$ , indépendante de  $W(n)$   $n \in \mathbb{Z}$  et que de même,  $b(n)$  et  $b(j)$  sont indépendants si  $n \neq j$  ( $P(b(n) = 1) = P(b(n) = -1) = 1/2$ ).

- Construire un filtre qui estime  $b(n)$ .

6. On désire dé-bruiter un signal de parole  $z(n)$  corrompu par du bruit additif  $b(n)$  indépendant du signal sonore et ce, par filtrage de Wiener. On suppose connues quelques valeurs d'autocorrélation pour les deux signaux telles que :

$$R_{zz}[0] = 1.5; R_{zz}[1] = 0.5; R_{zz}[2] = 0.25; R_{zz}[3] = 0.125; R_{zz}[4] = 0.0625$$

$$R_{bb}[0] = 1; R_{bb}[1] = 0.25; R_{bb}[2] = 0.0625; R_{bb}[3] = 0.015625$$

- Pourquoi ne peut-on pas employer un filtre adapté ou un filtre moyennneur ?
- Donner les équations de Wiener-Hopf permettant d'estimer  $z(n)$ .
- Déterminer le filtre de Wiener d'ordre 2 permettant de retrouver le signal utile  $\hat{z}(n)$ .
- Exprimer  $H(z)$ .

7. On considère un problème d'estimation de bruit  $b(n)$ .

Le signal observé est  $x(n) = s(n) + b(n) - b(n-1)$ .

On suppose que le signal  $s(n)$  est centré avec  $R_{ss}(n) = 0.8^{|n|}$  et qu'il est décorrélé du bruit dont l'autocorrélation est  $R_{bb}(n) = 0.8 \delta(n)$ .

- Déterminer les moyennes statistiques de  $s(n)$  et  $b(n)$ .
- Quand à-t-on recours au filtre de Wiener?
- Donner les équations de Wiener-Hopf permettant d'estimer  $b(n)$
- Déterminer le filtre de Wiener d'ordre 2 permettant de retrouver le signal utile  $\hat{b}(n)$ .
- Exprimer  $\hat{b}(n)$

8. Dans une usine de fabrication de composants électroniques, on prélève au hasard des composants jusqu'à tomber sur un qui serait défectueux. Cette expérience suit une loi géométrique de paramètre  $p$  :

$$p(x/\theta) = p(1-p)^x \quad \text{avec} \quad \mu_x = \frac{(1-p)}{p} \quad \sigma_x^2 = \frac{(1-p)}{p^2}$$

- Déterminer  $\hat{\theta}$  l'estimateur du maximum de vraisemblance de  $\theta = (1-p)/p$
- Montrer que le biais est nul et calculer la variance de cet estimateur
- Cet estimateur est-il consistant ?
- Dans quelle situation utilise-t-on l'estimateur linéaire à variance minimale?

### Solutions

$$1. \hat{\omega} = \frac{1}{n} \sum_{i=1}^n x_i - \alpha = \bar{x}_i - \alpha \quad p(x) = \frac{1}{\omega} e^{\frac{1}{\omega}(\alpha-x)} \Rightarrow \text{loi exponentielle avec } y = \alpha - x \text{ et } \mu_y = \omega \quad \sigma_y^2 = \omega^2, \quad b=0 \quad \sigma^2 =$$

$$\omega^2/n \quad 2. \text{Variable déterministe, } \theta = \frac{1}{rN} \sum_{i=1}^N x_i, \quad b=0, \quad \sigma^2 = \theta^2/Nr, \text{ consistant.}$$

$$3. \hat{\theta} = \frac{1}{2}(z_1 + z_2) \text{ pour que } \theta=0, \quad z_1 + z_2 < 1,$$

$\theta$  uniforme MAP  $\Leftrightarrow$  MV,  $\theta=0$  si  $z_1 + z_2 < 1 + 2\sigma^2 \ln((1-q)/q)$

$$4. R_{xx}(0) = 1 + \sigma^2 \quad R_{xx}(k > 0) = \sigma^2 \quad R_{\theta x}(k) = \sigma^2 \quad b_i = \sigma^2 / (N\sigma^2 + 1)$$

$$5. R_{xx}(1:3) = (1 + \alpha^2 + \sigma^2, \alpha, 0) \quad R_{bx}(1:3) = (1, \alpha, 0)$$

6.  $\mu_z = 0, \mu_b = 0$ , signal (parole) aléatoire,

$$7. \mu_s = 0, \mu_b = 0, S_s(f) = \frac{1}{1 - 0.8e^{-2\pi j f T_e}} + \frac{0.8e^{-2\pi j f T_e}}{1 - 0.8e^{-2\pi j f T_e}}, S_b(f) = 0.8. \text{ Quand signal utile et bruit occupent même plage de fréquences. } b_1 = 1/3 \text{ et } b_2 = -1/3$$

$$8. \hat{\theta} = \frac{\sum_{i=1}^N X_i}{N} \quad b_{\hat{\theta}} = 0 \quad \sigma_{b_{\hat{\theta}}}^2 = 0$$

### Exercices supplémentaires

1. Déterminer l'estimateur à vraisemblance maximale des paramètres  $\theta$  des distributions :

$$- p(x; \theta) = (1 + \theta) x^\theta \quad 0 \leq x \leq 1 \text{ et } \theta > -1$$

$$- p(x; \theta) = \theta e^{-\theta x} \quad x \geq 0 \text{ et } \theta > 0$$

$$- p(x; \theta) = \alpha \theta x^{\alpha-1} \exp(-\theta x^\alpha) \quad x \geq 0 \text{ et } \theta > 0$$

2. La durée  $T$  séparant deux arrivées successives de requêtes à un serveur suit une loi exponentielle de paramètre  $\theta$  et de densité de probabilité :  $p(t/\theta) = \theta e^{-\theta t}$  avec  $t > 0$ . On suppose que le paramètre  $\theta$  est stochastique et suit une loi exponentielle de paramètre  $\lambda$  connu, de densité  $p(\theta) = \lambda e^{-\lambda \theta}$ .

- Montrer que la densité a posteriori de  $\theta$  est proportionnelle à  $\theta e^{-\theta(t+\lambda)}$

- On observe un échantillon de  $n$  durée  $t_1, \dots, t_N$ .

Déterminer à partir des observations  $t_i$  les estimateurs du maximum a posteriori de  $\theta$ .

3. Soient  $(X_1, X_2, \dots, X_N)$ ,  $N$  variables aléatoires indépendantes suivant des lois de poisson de paramètre  $\lambda$ .

- Déterminer l'estimateur du maximum de vraisemblance du paramètre  $\lambda$ .

- Est-il sans biais, efficace ?

3. On considère un processus de la forme  $X(n) = S(n) - 2S(n-1) + S(n-2) + W(n)$  où  $W(n)$  est un processus gaussien de moyenne 0 et de variance  $\sigma^2$  tel que  $W(n)$  et  $W(j)$  sont indépendants si  $n \neq j$ . On suppose que  $S(n)$  suit une loi  $N(0,1)$  indépendante de  $W(n)$   $n \in \mathbb{Z}$  et que de même,  $S(n)$  et  $S(j)$  sont indépendants si  $n \neq j$ .

- Donner l'équation de Wiener-Hopf permettant de calculer les coefficients du filtre de Wiener d'ordre 3 permettant d'estimer  $S(n)$  à partir de  $X(n)$ ,  $X(n-1)$  et  $X(n-2)$ .

- Vérifier que si  $\sigma = 0$ , la solution est :  $S(n-1) = -(X(n) + 3X(n-1) + X(n-2))/5$

4. On considère une observation  $x(n) = s(n) + w(n)$  où  $w(n)$  est un bruit blanc de variance  $\sigma^2$  indépendant de  $x(n)$ . On suppose aussi que le signal utile  $x(n)$  est centrée et réduit et que les  $x(n)$  sont indépendants.

- Dans quel cas, utilise-t-on le filtre de Wiener ?

- Déterminer le filtre de Wiener d'ordre 3 permettant de retrouver le signal utile  $\hat{s}(n)$ .

- Donner l'expression de  $\hat{s}(n)$  en fonction de  $x(n)$ .

5. La distribution de Pareto est un type particulier de loi de puissance qui a des applications en sciences physiques et sociales. elle est définie par :

$$p(x/\alpha; \theta) = \theta \alpha^\theta x^{-\theta-1} \quad x > \alpha \text{ avec} \quad \mu_x = \frac{\alpha \theta}{\theta - 1} \quad \sigma_x^2 = \frac{\theta \alpha^2}{(\theta - 1)^2 (\theta - 2)}$$

• On suppose  $\alpha$  connue, déterminer alors  $\hat{\theta}$  l'estimateur du maximum de vraisemblance de  $\theta$

• Dans quelle situation utilise-t-on l'estimateur MAP?

**Solutions**

$$1. \hat{\theta} = \frac{-N}{\sum_{i=1}^n \ln(x_i)} - 1 \quad \hat{\theta} = N / \sum_{i=1}^n x_i \quad \hat{\theta} = N / \sum_{i=1}^n x_i^\alpha \quad 2. p(t/\theta)p(\theta) = \lambda \theta e^{-\theta(t+\lambda)} \quad \hat{\theta}_{MAP} = N / \left( \lambda + \sum_{i=1}^N t_i \right)$$

$$3. b=0, \quad \sigma^2 = \lambda/n \quad (=0 \quad n \rightarrow \infty) \quad \text{constant}$$

4. Estimer un signal aléatoire stationnaire (le débruiter).

$$R_{xx}(0) = 1 + \sigma^2, \quad R_{xx}(k > 0) = 0, \quad R_{sx}(k) = 1, \quad R_{sx}(k > 0) = 0 \quad \hat{s}(n) = \frac{1}{1 + \sigma^2} x(n)$$

## TP n°6 : Estimation par MV et Filtrage de Wiener

But : A travers deux exemples concrets, étudier deux techniques d'estimation : le maximum de vraisemblance et l'estimateur linéaire à variance minimale

**Exercice 1:** La même information binaire  $X \in \{0,1\}$  est transmise 2 fois consécutives vers un récepteur à travers un canal de transmission. Ces 2 informations sont perturbées par un bruit  $B$  supposé Gaussien centré de variance  $\sigma^2$ . Le message reçu s'écrit alors  $R=(R_1,R_2)$  où  $R_i=X+B_i$ . Le problème consiste à trouver le symbole émis à partir du message reçu  $R$ . Pour considérer que le signal émis est 0, il faudrait que  $p(R_1,R_2;0) < p(R_1,R_2;1)$ , soit après développement (voir TD), on aboutit à la condition  $R_1+R_2 < 1$  pour 0.

```
N=10; var=0.25; M=2*N; B=sqrt(var)*randn(M,1); X=2*rand(N,1); X=floor(X);
%Signal émis
for i=1:N
    E(2*i-1,1)=X(i);    E(2*i,1)=X(i);
end
%Signal Reçu
R=E+B;
%Estimation par maximum de vraisemblance
for i=1:N
    if (R(2*i-1,1)+R(2*i,1))>1    Z(i)=1 ;    else Z(i)= 0;
end
end
figure; subplot(311);stem(X);legend('Signal original');
subplot(312);stem(R);legend('Signal reçu');
subplot(313);stem(Z);legend('Signal estimé par MV')
```

1. Expliquer brièvement le programme.
2. Augmenter la variance du bruit et commenter pour différentes valeurs de  $N$ .

**Exercice 2:** L'extraction d'ECG du fœtus se fait à partir de plusieurs électrodes placées en différents endroits du ventre de la mère ; mais les enregistrements obtenus sont des mélanges de l'ECG du fœtus noté (FECG) et celui de la mère noté (MECG) auquel se rajoute le bruit thermique des électrodes et des équipements électroniques d'où l'emploi du filtre de Wiener.

```
load mecg1.dat; load fecg1.dat; load noise1.dat; fe=256; obs=mecg1+fecg1+0.15*noise1;
figure; subplot(411);plot(mecg1);legend('ecg mère');
subplot(412);plot(fecg1);legend('ecg fœtus');
subplot(413);plot(noise1);legend('bruit');
subplot(414);plot(obs);legend('observation=fecg+mecg+bruit');
figure; subplot(211); plot(abs(fft(mecg1)),'r'); legend('fft ecg mère');
subplot(212); plot(abs(fft(fecg1)),'g'); legend('fft ecg fœtus');
% Détermination de l'ECG du fœtus avec comme donnée l'observation et l'ECG de la mère
xm=mecg1-mean(mecg1); obs=obs-mean(obs);
%==== Détermination des coefficients du filtre
Ncoeff= 20; N=length(obs); Rxx= xcorr(obs);Rxx=Rxx(N:N+Ncoeff-1);
Rbb= xcorr(xm); Rbb=Rbb(N:N+Ncoeff-1); C = toeplitz(Rxx); D= Rxx-Rbb; b = C\D;
%====estimation du Signal
fecg1_est=filter(b,1,obs);
%==== Affichages
figure; subplot(311); plot(obs); legend('observation');grid
subplot(312); plot(fecg1_est);legend('ecg fœtus estimé'); grid
subplot(313); plot(fecg1);legend('ecg fœtus réel'); grid
```

1. Pourquoi a-t-on utilisé l'instruction  $D= Rxx-Rbb$  ; ?
2. Quelles lignes du programmes faut-il changer pour retrouver l'ECG de la mère.
3. Au final, est ce une estimation au sens des moindres carrés?