

## Processus Stochastiques

**Definition. Caractérisation d'ordre N. Moments. Stationnarité. Ergodicité. Processus Gaussiens. Processus Markoviens. Théorème de Mercer, représentation de Karhunen-Loève. Représentation spectrale. Processus blancs. Modèles paramétriques (AR, MA, ARMA). Modèles d'état. Décomposition de Wold.**

### 1. Définition et propriétés

#### 1.1 Définition

Un processus aléatoire (ou stochastique) (p.s.) peut être défini de deux façons:

- une application de  $S$  -- l'espace des réalisations -- dans un espace de fonctions de variable réelle (temps)

$$x(t): S \rightarrow F$$

$$\omega \mapsto \{x(t, \omega), t \in T\}$$

que, à chaque événement élémentaire fait correspondre une fonction du temps;

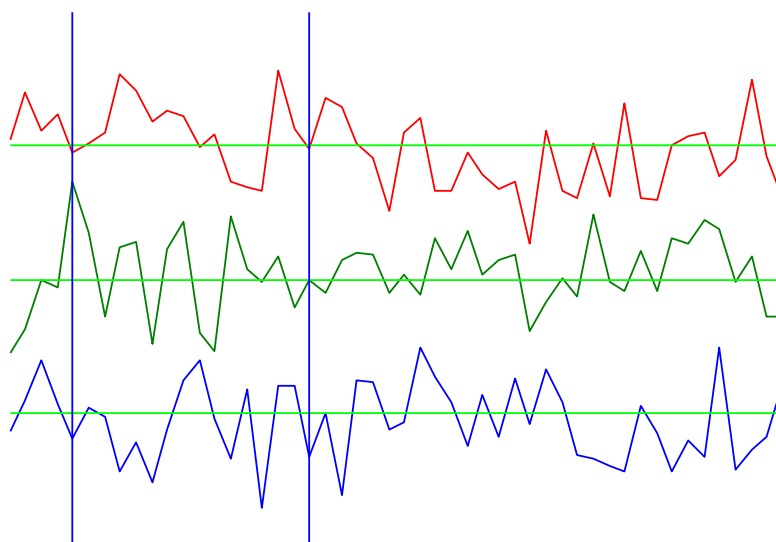
- une collection de variables aléatoires indexées

$$x(t) = \{x(t), t \in T\},$$

où  $x(t)$  est une variable aléatoire pour chaque  $t$ .

Les deux définitions sont équivalentes, et on utilisera une et l'autre selon le problème en étude.

La figure suivante illustre ces deux définitions:



#### Exemple

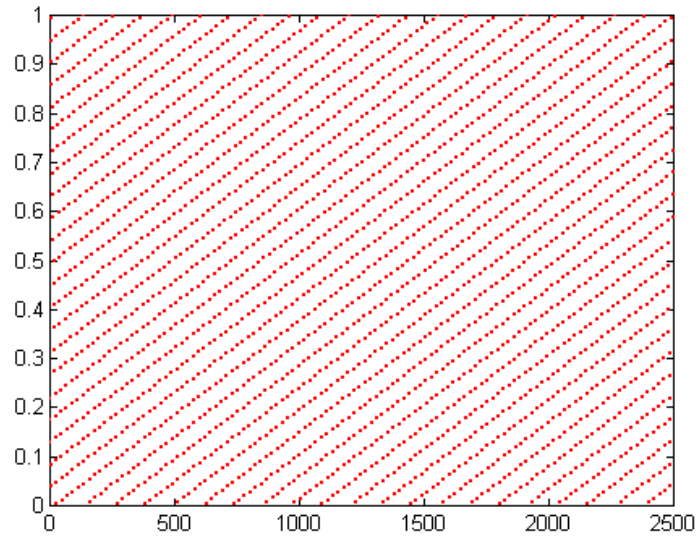
Dans cet exemple on considère un processus aléatoire défini à partir d'une seule variable aléatoire  $\beta$  uniformément distribuée en  $[0,1]$ , de la façon suivante:

$$s(1, \omega) = \beta(\omega)$$

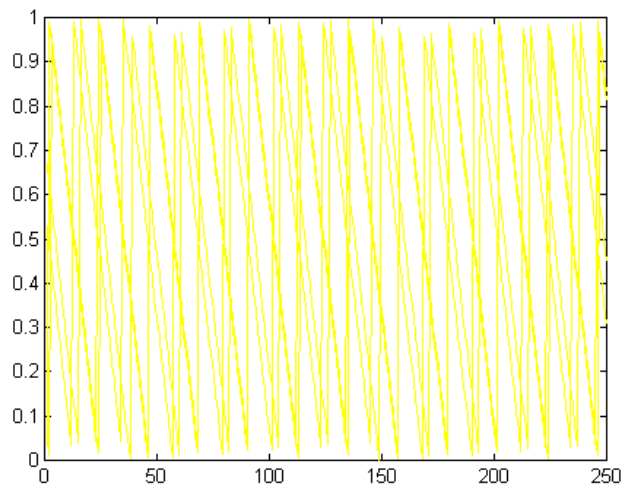
$$s(k+1, \omega) = (s(k, \omega) + \theta) \text{ mod } 1, \quad k = 2, 3, 4, \dots$$

où  $(a) \text{ mod } b$  est le résultat de la division entière de  $a$  par  $b$ , et  $\theta$  est un nombre irrationnel.

La figure suivante illustre une des réalisations possibles de ce processus.



Le caractère aléatoire du processus est plus clairement mis en évidence par la figure suivante, où sont représentées trois réalisations du processus, obtenues à partir de trois conditions initiales différentes:



### Exemple

Dans cet exemple, on présente un processus qui est plus naturellement modélisé comme un ensemble de variables aléatoires indexées.

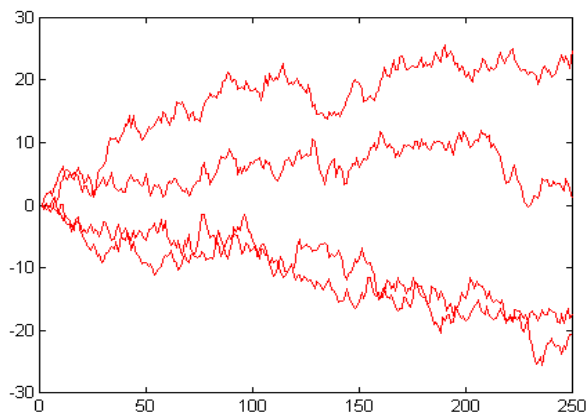
Considérons le processus aléatoire discret  $s(k), k = 1, \dots$

$$s(1) = 0,$$

$$s(k+1) = s(k) + \eta(k), \quad k = 2, 3, \dots$$

où les variables aléatoires  $\eta(k)$  sont iid, gaussiennes, de moyenne nulle et variance  $\sigma^2$ .

La figure suivante illustre quelques réalisations de ce processus.



## 1.2 Caractérisation statistique

### 1.2.1 Caractérisation d'ordre $N$

La caractérisation d'ordre  $N$  d'un processus aléatoire consiste à spécifier la densité de probabilité conjointe de tous les ensembles  $x(t_1), x(t_2), \dots, x(t_N)$ ,  $\forall t_1, t_2, \dots, t_N \in T$ :

$$p_{x(t_1)x(t_2)\dots x(t_N)}(x(t_1), x(t_2), \dots, x(t_N)).$$

**Exercice:** Déterminer la caractérisation d'ordre 1 des processus des exemples 1 et 2.  
Déterminer la caractérisation d'ordre 2 du processus de l'exemple 2.

### 1.2.2 Caractérisation complète

La caractérisation complète d'un processus consiste à spécifier sa caractérisation d'ordre  $N$  pour toutes les valeurs de  $N$  finies ( $N < \infty$ ).

### 1.2.3 Caractérisation partielle

Au lieu de spécifier une densité de probabilité, on définit certaines caractéristiques de la densité conjointe. On utilisera surtout la caractérisation partielle d'ordre deux, qui fait intervenir seulement *deux* variables aléatoires,  $x(t)$  **et**  $x(u)$ ,  $t, u \in T$ .

## 1.3 Moments

### Moment d'ordre 1 (moyenne)

$$m_x(t) = \mathbf{E}[x(t)] = \int_{-\infty}^{\infty} x_t p_{x_t}(x_t) dx_t$$

On notera que la moyenne d'un processus aléatoire, définie comme la valeur moyenne de chacune des variables aléatoires qui constituent le processus, est une fonction *déterministe* de variable réelle.

### Moments d'ordre 2 (auto-corrélation et covariance)

La *fonction d'auto-corrélation* est le moment non-centré d'ordre deux du p.s.:

$$R_x(t, u) = \mathbf{E}[x(t)x(u)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_t x_u p_{x_t x_u}(x_t, x_u) dx_t dx_u$$

## Propriétés

**1.** Une fonction d'auto-corrélation est *définie non-négative*, c'est à dire,

$$\forall \alpha_1 \dots \alpha_n, \forall t_1, \dots, t_n, \sum_i \sum_j \alpha_i \alpha_j^* R(t_i, t_j) \geq 0$$

De la même façon, si  $R(t,s)$  est une fonction définie non-négative, on peut toujours trouver un processus d'ordre deux (c.a.d, de valeur moments d'ordre deux finis), dont la fonction d'auto-corrélation est  $R(t,s)$ .

2. Une fonction d'auto-corrélation est symétrique (pour variables réelles) :

$$R(t, s) = R(s, t)$$

Ceci découle directement de la définition.

3. Inégalité de Schwarz

$$|R(t, s)| \leq \sqrt{R(t, t) R(s, s)}.$$

4. Fermeture

Si  $R(t,s)$  et  $P(t,s)$  sont des fonctions d'auto-corrélation, alors

- $R(t,s) + P(t,s)$  est aussi une fonction d'auto-corrélation
- $R(t,s)P(t,s)$  est aussi une fonction d'auto-corrélation
- Si  $a, b > 0$ , alors  $aR(t,s) + bP(t,s)$  est encore une fonction d'auto-corrélation
- Formes bilinéaires

Pour toute fonction  $f(t)$ ,  $R(t,s) = f(t)f(s)$  est une fonction d'autocorrélation.

La *fonction de covariance* est le moment centré d'ordre deux:

$$\begin{aligned} K_x(t, u) &= \mathbf{E}[(x(t) - m_x(t))(x(u) - m_x(u))] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_t - m_x(t))(x_u - m_x(u)) p_{x_t, x_u}(x_t, x_u) dx_t dx_u \end{aligned}$$

On définit encore le *coefficient de corrélation*, qui est une mesure normalisée de la corrélation statistique entre les variables correspondantes aux deux instants du temps,  $t$  et  $u$ :

$$C_x(t, u) = \frac{K_x(t, u)}{\sqrt{K_x(t, t) K_x(u, u)}}$$

Le module du coefficient de corrélation est toujours inférieur à 1 (*vérifier cette affirmation*).

On remarque que les trois fonctions qu'on vient de définir sont des fonctions réelles (déterministes) de deux variables réelles.

**Exercice:** Déterminer la caractérisation partielle d'ordre deux du processus de l'exemple de la page 29.

### 1.3.1 Moments croisés de deux processus

De la même façon qu'on a défini la distribution conjointe deux variables aléatoires, on peut aussi considérer la distribution conjointe de deux processus. Les processus sont alors caractérisés par les distributions conjointes de leurs respectives variables aléatoires. Soient  $x(t)$  et  $y(t)$  deux processus conjointement distribués. Leur densité conjointe d'ordre  $N$  est

$$p_{x_{t_1} \dots x_{t_p} y_{t_{p+1}} \dots y_{t_N}}(x_{t_1}, \dots, x_{t_p}, y_{t_{p+1}}, \dots, y_{t_N}).$$

La *corrélation croisée* des processus  $x(t)$  et  $y(t)$  est le moment croisé (non-centré) d'un pair de variables prises chacune dans un des processus aléatoires:

$$\boxed{R_{xy}(t, u) = \mathbf{E} x(t) y(u)}$$

De la même façon, la *covariance croisée* est le moment centré correspondant:

$$K_{xy}(t, u) = \mathbf{E} \left[ (x(t) - m_x(t)) (y(u) - m_y(u)) \right]$$

### 1.3.2 Processus Orthogonaux

On dit que deux processus sont *orthogonaux* quand leur fonction de corrélation est identiquement nulle:

$$R_{xy}(t, u) = 0, \quad \forall t, u \in T$$

On établit par la suite l'analogie entre cette notion d'orthogonalité (statistique) et la notion géométrique d'orthogonalité de vecteurs dans un espace Euclidien de dimension finie. Soient  $\mathbf{x}$  et  $\mathbf{y}$  deux vecteurs en  $\mathbf{R}^n$ , et représentons leur produit scalaire (ou produit interne) par  $\langle \mathbf{x}, \mathbf{y} \rangle$ . On dit que les vecteurs sont orthogonaux (au sens géométrique, c'est à dire qu'ils forment entre eux un angle de  $90^\circ$ ) si leur produit interne est nul:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^n x_i y_i = 0.$$

Considérons maintenant la condition de orthogonalité statistique,

$$R_{xy}(t, u) = \mathbf{E} x(t) y(u) = 0$$

Admettons, pour simplifier, qu'il s'agit de deux processus qui prennent valeurs dans deux ensembles finis,

$$x(t) \in \{a_1, a_2, \dots, a_{n_x}\} \quad y(u) \in \{b_1, b_2, \dots, b_{n_y}\}$$

et soit  $p(a_i, b_j) = \Pr(x(t) = a_i, y(u) = b_j)$ ,  $i = 1, \dots, n_x; j = 1, \dots, n_y$  la probabilité conjointe des deux processus. Alors la valeur moyenne de l'équation antérieure peut s'écrire

$$R_{xy}(t, u) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} a_i b_j p(a_i, b_j) = 0.$$

Admettons que l'on fait  $N$  tirages au sort des variables aléatoires  $x(t)$  et  $y(u)$  et construisons deux vecteurs aléatoires  $X$  et  $Y$ , de dimension  $N$ :

$$X = x_1 \mathbf{L} x_N, Y = y_1 \mathbf{L} y_N,$$

où  $(x_i, y_i)$  est la paire de valeurs obtenues dans le  $i$ -ième tirage.

Selon la loi des grands nombres, on doit avoir  $x(t) = a_i$  et  $y(u) = b_j$  un nombre de fois égal à  $Np(a_i, b_j)$ . La somme de l'expression antérieure est donc égale à

$$R_{xy}(t, u) = X^T Y = \sum_{i=1}^N x_i y_i$$

qui est bien l'expression du produit interne des deux vecteurs, reliant ainsi les notions de orthogonalité statistique et de orthogonalité géométrique.

### 1.3.3 Processus non-correlés

Deux processus sont non correlés si leur fonction de covariance est identiquement nulle:

$$K_{xy}(t, u) = 0.$$

Notons que non-corrélation n'implique pas, en général, orthogonalité. On peut démontrer que

$$K_{xy}(t, u) = R_{xy}(t, u) - m_x(t) m_y(u).$$

On peut donc avoir corrélation nulle sans que les processus soient orthogonaux. Cependant, pour des processus de moyenne zéro, les deux conditions sont, évidemment, équivalentes.

### 1.4 Processus indépendants

Deux processus aléatoires sont indépendants si n'importe quel ensemble de variables aléatoires prises dans les deux processus sont indépendantes. Cela veut dire que,

$\forall t_1, \dots, t_p, t_{p+1}, \dots, t_N \in T$ , la densité conjointe des variables aléatoires  $x(t_1), \dots, x(t_p)$  de  $x(t)$ , et  $y(t_{p+1}), \dots, y(t_N)$  de  $y(t)$ , factorise de la façon suivante:

$$p_{x_{t_1} \dots x_{t_p} y_{t_{p+1}} \dots y_{t_N}}(x_{t_1}, \dots, x_{t_p}, y_{t_{p+1}}, \dots, y_{t_N}) = p_{x_{t_1} \dots x_{t_p}}(x_{t_1}, \dots, x_{t_p}) p_{y_{t_{p+1}} \dots y_{t_N}}(y_{t_{p+1}}, \dots, y_{t_N})$$

### 1.5 Processus stationnaires

On dit qu'un processus est stationnaire si ses caractéristiques ne varient pas avec la définition de l'origine du temps, ou, encore, si ses caractéristiques statistiques ne varient pas le long du temps. On définit plusieurs types de stationnarité.

#### Stationnarité au sens strict

Un processus est stationnaire au sens strict si pour toute valeur de  $N$ , sa caractérisation d'ordre  $N$  est invariante par rapport à une translation de l'axe des temps:

$$p_{x(t_1) \dots x(t_N)}(x(t_1), \dots, x(t_N)) = p_{x(t_1+\Delta) \dots x(t_N+\Delta)}(x(t_1+\Delta), \dots, x(t_N+\Delta))$$

#### **Exemple.**

Considérer le processus défini dans la page 28. Montrer qu'il s'agit d'un processus stationnaire.

#### Stationnarité au sens large (ou de deuxième ordre)

Un processus aléatoire est stationnaire au sens large si sa moyenne est constante,

$$\boxed{m_x(t) = m_x},$$

et sa fonction d'auto-covariance ne dépend pas de l'origine du temps. Comme  $K_x(t, u)$  ne dépend que de deux instants du temps, la condition précédente est équivalente à

$$\boxed{K_x(t, u) = K_x(t - u)}.$$

#### Processus conjointement stationnaires

Deux processus sont conjointement stationnaires si

a) chacun est stationnaire considéré d'une façon isolée, et

b) leur corrélation croisée  $R_{xy}(t, u)$  dépend seulement de la différence  $\tau = t - u$ :

$$R_{xy}(t, u) = R_{xy}(t - u) = R_{xy}(\tau)$$

On notera que dans ce cas, leur covariance croisée dépend également uniquement de la différence des instants de temps considérés.

#### Processus cyclo-stationnaires

Un processus est cyclo-stationnaire au sens strict si ses statistiques sont des fonctions périodiques du temps (de période  $T$ ).

#### Théorème

Si  $x(t)$  est un processus cyclo-stationnaire au sens strict et  $\theta$  est une variable aléatoire uniforme en  $0, T$ , alors  $\bar{x}(t) = x(t - \theta)$  est un processus stationnaire au sens strict.

Un processus est cyclo-stationnaire au sens large si ses moments d'ordre 1 et 2 sont des fonctions périodiques :

$$m_x(t + kT) = m_x(t)$$

$$R_x(t + kT, u + kT) = R_x(t, u)$$

### Théorème

Si  $x(t)$  est un processus cyclo-stationnaire au sens large et  $\theta$  est une variable aléatoire uniforme en  $[0, T]$ , alors  $\bar{x}(t) = x(t - \theta)$  est un processus stationnaire au sens large, et

$$m_{\bar{x}}(t) = m_x = \frac{1}{T} \int_T m_x(t) dt$$

$$R_{\bar{x}}(\tau) = \frac{1}{T} \int_T R_x(t + \tau, t) dt$$

### 1.6 Ergodicité

La propriété d'ergodicité lie les moyennes statistiques (effectuées sur l'espace des réalisations sous-jacent à la définition des variables aléatoires qui constituent le processus) et les moyennes temporelles (effectuées sur les fonctions du temps qui sont les *réalisations* du processus).

Un processus aléatoire est *ergodique* si ses moments peuvent être obtenus comme des moyennes à partir d'une seule de ses réalisations. Ceci doit être vrai en particulier pour les moments d'ordre 1 et 2:

$$\frac{1}{T} \int_T x(t, \omega) dt \xrightarrow{T \rightarrow \infty} m_x(t)$$

$$\frac{1}{T} \int_T x(t, \omega) x(t + \tau, \omega) dt \xrightarrow{T \rightarrow \infty} R_x(t, t + \tau)$$

Si l'on examine la première équation, on se rend compte que le membre gauche ne dépend pas du temps, et donc que pour que cette équation puisse être vérifiée, il faut que le processus ait une *moyenne constante*. De la même façon, pour que la deuxième équation soit possible le processus doit être *stationnaire au sens large*, on aura alors

$$\frac{1}{T} \int_T x(t, \omega) dt \xrightarrow{T \rightarrow \infty} m_x$$

$$\frac{1}{T} \int_T x(t, \omega) x(t + \tau, \omega) dt \xrightarrow{T \rightarrow \infty} R_x(\tau)$$

On peut donc affirmer que pour qu'un processus soit ergodique, il doit nécessairement être stationnaire:

ergodicité  $\Rightarrow$  stationnarité

L'affirmation contraire est fausse, comme le montre l'exercice suivant.

**Exercice.** Soient  $X$  et  $Y$  deux processus stationnaires et ergodiques. A partir de ces deux processus on construit un troisième processus de la façon suivante. On lance une pièce et si le résultat est face on observe le processus  $X$  :  $z(k) = x(k)$ ; en cas contraire on observe le processus  $Y$ :  $z(k) = y(k)$ . Montrer que le processus  $Z$  est stationnaire, mais pas ergodique.

**Exercice.** Montrer que le processus introduit dans l'exemple 1 est ergodique. Vérifier ceci en comparant (utiliser Matlab) l'histogramme d'une réalisation  $x(1, \omega_1), x(2, \omega_1), \dots, x(N, \omega_1)$  (utiliser plusieurs valeurs de  $N$ ) avec la densité d'ordre 1 obtenue théoriquement.

**Exemple.** Considérer le processus aléatoire stationnaire  $x(t)$  de moyenne nulle et de fonction d'auto-covariance

$$K(\tau) = e^{-\alpha|\tau|}$$

Soit  $\eta_T$  la moyenne temporelle de  $x(t)$  calculée sur l'intervalle  $[-T, T]$ :

$$\eta_T = \frac{1}{2T} \int_{-T}^T x(t) dt.$$

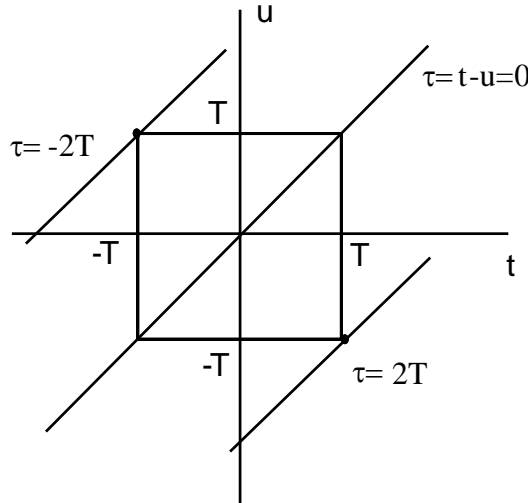
On peut conclure facilement que la valeur moyenne de la variable aléatoire  $\eta_T$  est aussi zéro:

$$E[\eta_T] = E\left[\frac{1}{2T} \int_{-T}^T x(t) dt\right] = \frac{1}{2T} \int_{-T}^T E[x(t)] dt = 0$$

Le calcul de la covariance de  $\eta_T$  peut être fait à partir de la définition de covariance:

$$\begin{aligned} \sigma_{\eta_T}^2 &= E[\eta_T^2] = E\left[\frac{1}{4T^2} \int_{-T}^T x(t) dt \int_{-T}^T x(u) du\right] \\ &= \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T E[x(t)x(u)] dt du \\ &= \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T K(t-u) dt du \end{aligned}$$

L'intégrale de l'équation précédente peut être simplifiée. Si on définit  $\tau = t - u$ , on voit que la fonction qui est intégrée est constante selon les lignes indiquées sur la figure suivante:



On effectue la transformation de variables suivante:

$$\begin{aligned} \tau = t - u &\Leftrightarrow t = \frac{1}{2}(\tau + \xi) \\ \xi = t + u &\Leftrightarrow u = \frac{1}{2}(\xi - \tau) \end{aligned}$$

dont le Jacobien est

$$J = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow |J| = \frac{1}{2}$$

L'intégrale peut donc s'écrire, d'une façon équivalente,

$$\begin{aligned} \sigma_{\eta_T}^2 &= \frac{1}{4T^2} \int_{-2T}^0 \int_{-2T-\tau}^{2T+\tau} \frac{1}{2} d\xi K(\tau) d\tau + \int_0^{2T} \int_{-2T+\tau}^{2T-\tau} \frac{1}{2} d\xi K(\tau) d\tau \\ &= \frac{1}{4T^2} \int_{-2T}^{2T} (2T - |\tau|) K(\tau) d\tau \end{aligned}$$

Si on utilise maintenant la définition de  $K(\tau)$ ,

$$\begin{aligned} \sigma_{\eta_T}^2 &= (1/2T) \int_{-2T}^{2T} (1 - (|\tau|/T)) e^{-\alpha|\tau|} d\tau \\ &= (1/\alpha T) - \frac{1 - e^{-2\alpha T}}{2\alpha^2 T^2} \end{aligned}$$



La limite quand  $T \rightarrow \infty$  de cette expression est 0, et donc la variable aléatoire  $\eta_T$  -- la valeur moyenne temporelle -- tend en probabilité vers la valeur moyenne statistique, ce qui montre que le processus est ergodique pour la moyenne.

Cet exemple montre que l'ergodicité pour la moyenne correspond à une condition sur le moment d'ordre deux du processus. On peut vérifier facilement que l'ergodicité pour l'auto-correlation conduit à une condition sur un moment d'ordre quatre du processus.

On vient d'introduire un ensemble de définitions qui caractérisent statistiquement un processus stochastique. Dans le reste de ce cours, on utilisera souvent la fonction de corrélation, la covariance croisée, etc., et on aura occasion de faire référence à des propriétés comme stationnarité ou ergodicité. Les caractérisations des signaux aléatoires que l'on vient de présenter sont bien différentes de celles qui sont utilisées pour les signaux déterministes, et qui consistent, par exemple, à donner une formule analytique qui décrit son évolution, spécifier leur spectre, ou encore à les décrire comme la sortie d'un système linéaire. Comme on discutera dans une section postérieure, des représentations équivalentes existent aussi pour certains processus aléatoires.

### 1.7 Processus Gaussiens

Il y a plusieurs façons de définir un processus Gaussien. On présente ici deux de ces définitions:

Un p.s. est Gaussien si toute combinaison linéaire (de coefficients qui ne sont pas identiquement nuls) est une variable aléatoire gaussienne.

Un p.s. est Gaussien si ses densités d'ordre  $n$  sont conjointement gaussiennes, pour toutes valeurs de  $n$ .

#### Exercice.

Un p.s. discret,  $x(1), x(2), \dots$ , est Gaussien, de moyenne nulle. Soit  $\Gamma_{ij} = \mathbf{E} x(i)x(j)$ . Montrer que le processus est stationnaire ssi  $\Gamma_{ij}$  dépend uniquement de  $|i - j|$ .

### 1.8 Processus de Markov

Les processus de Markov jouent un rôle important dans la théorie du traitement du signal, parce qu'ils conduisent, dans beaucoup de circonstances, à des filtres de mémoire finie. On présente ici uniquement la définition pour des processus discrets (ensemble  $T$  dénombrable).

#### Définition.

Un processus  $x(t), t \in T$  est de Markov si

$$p(x(k+1) | x(k)x(k-1)x(k-2)\dots) = p(x(k+1) | x(k)).$$

L'équation précédente décrit de façon formelle la propriété suivante: *le futur et le passé sont conditionnellement indépendants, étant donné le présent.*

#### Exemple

Soient  $Y_1, Y_2, \dots$  des variables aléatoires iid, et  $S$  le processus défini par ses sommes partielles

$$S(k) = \sum_{i=1}^k y(i).$$

Montrer que  $S$  est un processus de Markov (noter que ce processus a été introduit dans l'exemple de la page 29).

Quand les variables aléatoires  $x(k)$  sont discrètes, définies dans un même ensemble dénombrable  $\aleph$ , on appelle le processus de Markov une *chaîne de Markov*. Le processus  $S_n$  étudié dans l'appendix du Chapitre 1 est un exemple d'une chaîne de Markov, qui prend des valeurs dans les entiers. Ces processus sont complètement caractérisés par

- la distribution de leur valeur initiale

$$\Pr\{x(0) = a_i\} = p_i^0, \forall a_i \in \aleph$$

- l'ensemble de probabilités conditionnelles

$$\Pr\{x(k) = a_i \mid x(k-1) = a_j\} = p_{ij}(k), \forall a_i, a_j \in \aleph$$

On doit avoir, évidemment,

$$\sum_i p_i^0 = 1 \quad \sum_i p_{ij}(k) = 1, \forall j, \forall k = 1, 2, \dots$$

Si les probabilités  $p_{ij}(k) = p_{ij}$  ne dépendent pas de l'instant  $k$  considéré, on dira qu'il s'agit d'une chaîne de Markov de probabilité de transition stationnaire. Par la suite, on considère uniquement ce type de chaînes.

La probabilité pour que la chaîne prenne la valeur  $x(k) = a_i$  à l'instant  $k$ , sachant que sa valeur à l'instant  $n$  est  $x(n) = a_j$  est donnée par

$$p_{ij}(k, n) = \Pr\{x(k) = a_i \mid x(n) = a_j\} = \sum_q p_{iq}(k, r) p_{qj}(r, n)$$

où  $r$  est un instant quelconque pris entre  $k$  et  $n$ . Cette formule est connue par le nom d'équation de *Chapman Kolmogorov*, et joue un rôle fondamentale dans l'étude des processus Markov.

Pour des chaînes avec des transitions stationnaires, la probabilité dépend uniquement de la distance entre les deux instants considérés :

$$p_{ij}(k - n) = \Pr\{x(k) = a_i \mid x(n) = a_j\} = \sum_q p_{iq}(k - r) p_{qj}(r - n)$$

En définissant  $m = k - r$  et  $s = r - n$  (et donc  $k - n = m + s$ ), on obtient

$$p_{ij}(m + s) = \sum_q p_{iq}(m) p_{qj}(s)$$

Considérons maintenant le cas où l'ensemble de valeurs prises par la chaîne est *fini*, c'est à dire  $\#\aleph = M < \infty$ . Soit  $\mathbf{P}$  la matrice  $M \times M$  dont les entrées sont les valeurs de la probabilité de transition de la chaîne

$$\mathbf{P} = \begin{bmatrix} p_{11} & & p_{M1} \\ & \ddots & \\ p_{1M} & & p_{MM} \end{bmatrix}$$

Alors, on peut montrer que la densité de transition en  $n$  steps,  $\mathbf{P}^{(n)}$ , formée par les valeurs de la probabilité de transition en  $n$  steps, est donnée par la puissance  $n$  de la matrice  $\mathbf{P}$  :

$$\mathbf{P}^{(n)} = \mathbf{P}^n$$

Notons que les équations précédentes permettent le calcul de la loi de probabilité pour n'importe quel instant du temps :

$$p_i^{(k)} = \Pr\{x(k) = a_i\} = P_i^{(n)} p_j^{(0)}$$

et si on définit  $p^{(k)}$  comme le vecteur

$$p^{(k)} = \begin{bmatrix} p_1^{(k)} \\ \vdots \\ p_M^{(k)} \end{bmatrix}$$

on a l'équation

$$p^{(k)} = \mathbf{P}^{(k)} p^{(0)} = \mathbf{P}^n p^{(0)}$$

### 1.8.1 Fermeture et ensembles fermés

On dit que l'état  $a_i$  peut être atteint à partir de l'état  $a_j$  s'il existe un  $n \geq 0$  tel que  $p_{ij}^{(n)} > 0$ .

#### Définition

Un sous-ensemble  $\mathbf{C}$  de l'espace d'états  $\mathbb{S}$  est *fermé* si aucun état en dehors de  $\mathbf{C}$  peut être atteint à partir d'un état dans  $\mathbf{C}$ . Soit  $\mathbf{B}$  un sous-ensemble arbitraire de  $\mathbb{S}$ . On appelle *fermeture de  $\mathbf{B}$*  le plus petit sous-ensemble de  $\mathbb{S}$  contenant  $\mathbf{B}$  qui est fermé. Si la fermeture d'un élément  $a_i$  de  $\mathbb{S}$  coïncide avec  $a_i$ , alors on dit que l'état  $a_i$  est *absorbant*.

Une chaîne de Markov est *irréductible* si le seul sous-ensemble fermé de  $\mathbb{S}$  est  $\mathbb{S}$ .

Il est évident que  $\mathbf{C}$  est fermé si et seulement si  $p_{ij} = 0, a_i \notin \mathbf{C}, a_j \in \mathbf{C}$ . Dans ce cas, on peut éliminer toutes les lignes et colonnes correspondantes aux états en dehors de  $\mathbf{C}$ , et la matrice résultante est encore une probabilité de transition pour une chaîne réduite, qui a pour espace d'états  $\mathbb{S}$ .

#### Exemple

Considérons une chaîne avec la matrice de transition suivante

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & .5 & 0 & 0 & 0 & 0 & .7 \\ 0 & .3 & .3 & 0 & .3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1. & 0 \\ 1. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .7 & 0 & 0 & 0 & 0 \\ 0 & .5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .2 & 0 & 0 & 0 & 1. & 1. & 0 & 0 \\ 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 & 0 & 0 & .3 \end{bmatrix}$$

Remarquons d'abord que l'état 7 ne peut conduire qu'à lui-même. Il s'agit donc d'un état *absorbant* : si la chaîne rentre dans cet état, elle y restera toujours.

L'état 6 peut conduire à lui-même ou à l'état 7. Donc,  $C_1 = \{e_6, e_7\}$  est un ensemble fermé.

On remarque aussi que l'état 1 ne peut conduire qu'à l'état 4, et que celui-ci ne peut conduire qu'à 1 de nouveau ou à 9. Finalement, l'état 9 peut de nouveau conduire à 1 ou rester dans le même état. L'ensemble  $C_2 = \{e_1, e_4, e_9\}$  constitue donc un ensemble fermé.

L'état 2 ne peut conduire qu'aux états 6 ou 7.  $C_3 = \{e_2, e_6, e_7\}$  est donc aussi un ensemble fermé.

On reconnaîtra facilement les autres ensembles fermés de cette chaîne :  $C_4 = \{e_2, e_3, e_6, e_7, e_8\}$ ,  $C_5 = \{e_2, e_5, e_6, e_7\}$ .

Notez qu'une renumérotation des états (6,7,2,5,3,8,1,4,9) permet d'écrire la matrice de transition de cette chaîne de la forme suivante :

$$\mathbf{P} = \begin{bmatrix} 1. & 1. & .2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .3 & .3 & .3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1. & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .5 & .7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1. & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .5 & .3 \end{bmatrix}$$

Comme l'exemple précédent montre, si une chaîne a un sous-ensemble fermé de dimension  $r$ , alors on peut toujours re-ordonner les états de façon à écrire sa matrice de transition dans la forme

$$\mathbf{P} = \begin{bmatrix} Q & U \\ 0 & V \end{bmatrix}$$

où la matrice  $Q$  est de dimension  $r \times r$  et  $V$  est une matrice carrée de dimension  $M-r$ . Dans ce cas, on vérifie facilement que

$$\mathbf{P}^{(n)} = \begin{bmatrix} Q^n & U_n \\ 0 & V^n \end{bmatrix}$$

Ceci indique qu'on peut étudier séparément l'évolution des états dans un ensemble fermé et dans son complément.

### Définition

Un état  $a_i$  a *période*  $t > 1$  si  $p_{ii}^{(n)} = 0, n \neq kt$ , et  $t$  est le plus grand entier avec cette propriété.

### Définitions

Soit  $f_{ij}^{(n)}$  la probabilité pour que depuis l'état  $i$  le premier retour à l'état  $j$  soit obtenu après  $n$  steps. Par définition,  $f_{ij}^{(0)} = 0$ . Alors, la probabilité pour que la chaîne retourne à  $i$  après avoir passé par  $j$  est donnée par

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

et le nombre moyen de steps nécessaires (temps moyen de recurrence) pour y retourner est

$$\mu_{ij} = \sum_{n=1}^{\infty} n f_{ij}^{(n)}$$

Quand  $f_{ij}=1$ , les  $f_{ij}^{(n)}$  sont une loi de probabilité, et on l'appelle la distribution du premier temps de passage.

### Définition

Un état  $a_i$  est *persistant* si  $f_{ii}=1$  et *transitoire* si  $f_{ii}<1$ . Un état persistant est dit *nul* si son temps moyen de recurrence est infini. Un état apériodique persistant et non-nul est dit *ergodique*.

## 2. Représentations de Processus Stochastiques

### 2.1 Théorème de Mercer. Représentation de Karhunen Loève.

La première représentation qu'on étudiera est étroitement liée à la fonction d'auto-covariance du processus. Par la suite, et pour simplifier la notation, on considère que le processus a moyenne nulle :  $\forall t \in T, m_x(t) = 0$ .

Avant de présenter cette représentation, on rappelle la notion d'*ensemble complet ortho-normé de fonctions*.

Soit  $\mathbf{F}$  un espace (vectoriel) de fonctions de dimension infinie  $\mathbf{F} = \{f(t)\}$ . Une famille dénombrable de fonctions de  $\mathbf{F}$  :  $\{\phi_i(t)\}_{i=1,\dots}$ ,  $\phi_i(t) \in \mathbf{F}$ , est un ensemble complet ortho-normé si

i) elles sont orthogonales,

$$\phi_i(t) \perp \phi_j(t), i \neq j,$$

et normées

$$\|\phi_i(t)\| = 1, \quad \forall i$$

ii) toutes les fonctions de  $\mathbf{F}$  peuvent être écrites comme une combinaison linéaire (éventuellement infinie) des fonctions  $\phi_i(t)$ :

$$\forall f(t) \in \mathbf{F}, \quad \exists f_i \in \mathbb{R}, \quad i = 1, 2, \dots$$

$$f(t) = \sum_i f_i \phi_i(t) \quad .$$

On rappelle encore que les coefficients  $f_i$  sont donnés par le produit interne

$$f_i = \int_T f(t) \phi_i(t) dt .$$

La *représentation de Karhunen Loève* est une représentation de ce type, mais où les fonctions  $\phi_i(t)$  sont choisies de façon à ce que les coefficients de la représentation soient des *variables aléatoires non-correlées*.

Soit  $x(t)$  un processus de moyenne nulle et de fonction de corrélation  $R_x(t, u)$ , et soit  $\{\phi_i(t)\}_{i=1,\dots}$  un ensemble complet ortho-normé de fonctions. Définissons l'ensemble dénombrable de variables aléatoires que l'on obtient en faisant le produit interne de  $x(t)$  par chacune des ces fonctions:

$$x_i = \int_T x(t) \phi_i(t) dt .$$

La corrélation croisée des variables  $x_i$  est, par définition,

$$\begin{aligned} \mathbf{E}[x_i x_j] &= \mathbf{E}\left[\int_T x(t) \phi_i(t) dt \int_T x(s) \phi_j(s) ds\right] \\ &= \int_T dt \int_T \phi_i(t) \mathbf{E}[x(t) x(s)] \phi_j(s) ds \\ &= \int_T dt \phi_i(t) \int_T R_x(t, s) \phi_j(s) ds \end{aligned}$$

Soit  $\kappa_j(t)$  le résultat du deuxième intégral de l'expression précédente de façon à ce que :

$$\mathbf{E}[x_i x_j] = \int_T dt \phi_i(t) \kappa_j(t)$$

$$\kappa_j(t) = \int_T R_x(t, s) \phi_j(s) ds$$

et exprimons  $\kappa_j(t)$  en fonction de l'ensemble complet ortho-normé original:

$$\kappa_j(t) = \sum_p \kappa_{jp} \phi_p(t).$$

La substitution de cette expression donne

$$\begin{aligned} \mathbf{E}[x_i x_j] &= \int_T dt \phi_i(t) \sum_p \kappa_{jp} \phi_p(t) = \sum_p \kappa_{jp} \int_T \phi_i(t) \phi_p(t) dt \\ &= \sum_p \kappa_{jp} \delta_{ip} = \kappa_{ji} \end{aligned}$$

où  $\delta_{ij}$  est le symbole de Kroenecker:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

On cherche les fonctions de façon à ce que

$$\boxed{\mathbf{E} x_i x_j = \sigma_i^2 \delta_{ij}}$$

ce qui implique donc

$$\kappa_{ji} = \lambda_i^2 \delta_{ij} \Rightarrow \kappa_j(t) = \lambda_i^2 \phi_i(t)$$

ou encore, d'après la définition de  $\kappa_j(t)$ :

$$\int_T R_x(t, s) \phi_i(s) ds = \lambda_i^2 \phi_i(t).$$

L'équation précédente est l'équation aux valeurs et fonctions propres de la fonction  $R_x(t, u)$ . On peut donc maintenant identifier l'ensemble complet ortho-normé qui produit des coefficients non-correlés comme la famille de fonctions propres de la fonction de covariance du processus.

### **Théorème** (de Mercer)

Soit  $K_x(t, u)$  la fonction de covariance du processus  $x(t)$ , et soient  $\phi_i(t)$  et  $\lambda_i^2$  ses fonctions et valeurs propres. Alors,

$$\boxed{K_x(t, u) = \sum_i \lambda_i^2 \phi_i(t) \phi_i(u)}$$

et la série converge uniformément en  $T$ .

Le processus  $x(t)$  peut lui-même être représenté à l'aide des coefficients  $x_i$  ainsi obtenus:

$$x(t) = \text{l.i.m.}_{N \rightarrow \infty} \sum_i x_i \phi_i(t)$$

où la convergence de la série est prise au sens de la convergence en erreur quadratique moyenne:

$$x(t) = \lim_{N \rightarrow \infty} z_N(t) \Leftrightarrow \mathbf{E}[(x(t) - z_N(t))^2] \xrightarrow{N \rightarrow \infty} 0$$

## 2.2 Spectre de Puissance de Signaux Stationnaires

Dans cette section on considère la caractérisation spectrale de processus stationnaires en sens large. On commence par énoncer quelques propriétés des moments d'ordre 1 et 2 de processus stationnaires réels.

La valeur moyenne est une constante:

$$E x(t) = m_x$$

et la fonction d'auto-correlation dépend uniquement de l'intervalle entre les deux instants considérés:

$$E x(t)x(t + \tau) = R(\tau)$$

### Propriétés

$$1. R(\tau) = R(-\tau)$$

$$2. z(t) = x(t) + y(t) \Rightarrow R_z(\tau) = R_x(\tau) + R_y(\tau) + R_{xy}(\tau) + R_{yx}(\tau)$$

$$3. |R(\tau)| \leq |R(0)|$$

**Dém.:**

$$E[(x(t + \tau) + x(t))^2] \geq 0 \Rightarrow 2[R(0) + R(\tau)] \geq 0 \Rightarrow R(\tau) \geq -R(0)$$

$$E[(x(t + \tau) - x(t))^2] \geq 0 \Rightarrow 2[R(0) - R(\tau)] \geq 0 \Rightarrow R(\tau) \leq R(0)$$

et ces deux inégalités sont équivalentes à l'inégalité énoncée.

$$4. R_{xy}^2(\tau) \leq R_x(0)R_y(0)$$

**Dém.:** considérer la valeur quadratique moyenne de la somme  $x(t + \tau) + ay(t)$ , pour des valeurs arbitraires de  $a$ .

### Définition (Densité spectrale de puissance)

La densité spectrale de puissance d'un processus stationnaire,  $S(\omega)$ , est la transformée de Fourier de sa fonction d'auto-correlation:

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau} d\tau$$

La symétrie de la fonction d'auto-correlation implique que  $S(\omega)$  est une fonction réelle. Pour des signaux aléatoires réels, la densité spectrale (encore appelée *spectre*) est aussi une fonction *paire*.

De la formule d'inversion de la transformée de Fourier,

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)e^{j\omega\tau} d\omega$$

Si on prend  $\tau = 0$  dans cette équation,

$$R(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) d\omega = E[x^2(t)]$$

c'est à dire, la surface totale du spectre du signal est égale à sa puissance moyenne (à part le facteur  $2\pi$ ).

Des propriétés générales de la transformée de Fourier découlent un certain nombre de propriétés de la densité spectrale d'un signal aléatoire:

$x(t)$	$R(\tau)$	$S(\omega)$
$ax(t)$	$ a ^2 R(\tau)$	$ a ^2 S(\omega)$
$\frac{dx(t)}{dt}$	$-\frac{d^2 R(\tau)}{d\tau^2}$	$\omega^2 S(\omega)$
$\frac{d^n x(t)}{dt^n}$	$(-1)^n \frac{d^{2n} R(\tau)}{d\tau^{2n}}$	$\omega^{2n} S(\omega)$
$x(t)e^{\pm j\omega_i t}$	$R(\tau)e^{\pm j\omega_i \tau}$	$S(\omega \pm \omega_i)$

### Exemple

Soit  $n(t)$  le nombre d'événements d'un processus de Poisson de paramètre  $\lambda$  dans l'intervalle  $0, t$ , de façon que

$$\Pr(n(t) = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!},$$

et le nombre d'événements observés dans des intervalles disjoints sont indépendants.

Basé sur  $n(t)$ , on construit le processus  $x(t)$  de la façon suivante

$$x(t) = \begin{cases} 1, & \text{si } n(t) \text{ est paire} \\ -1, & \text{si } n(t) \text{ est impaire} \end{cases}$$

La probabilité d'avoir un nombre paire d'événements est

$$\Pr[x(t) = 1] = e^{-\lambda t} \left[ 1 + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^4}{4!} + \dots \right] = e^{-\lambda t} \cosh(\lambda t)$$

et, de la même façon, on obtient,

$$\Pr[x(t) = -1] = e^{-\lambda t} \left[ \lambda t + \frac{(\lambda t)^3}{3!} + \frac{(\lambda t)^5}{5!} + \dots \right] = e^{-\lambda t} \sinh(\lambda t)$$

On peut facilement, en ce moment, calculer le premier moment du processus  $x(t)$ :

$$E[x(t)] = \Pr(x(t) = 1) - \Pr(x(t) = -1) = e^{-\lambda t} (\cosh(\lambda t) - \sinh(\lambda t)) = e^{-2\lambda t}.$$

La valeur moyenne n'est pas constante, et donc on peut déjà conclure que  $x(t)$  n'est pas stationnaire.

On calcule maintenant sa fonction d'auto-corrélation:

$$\begin{aligned} R(t_1, t_2) &= E[x(t_1)x(t_2)] \\ &= 1 \cdot (\Pr(x(t_1) = 1, x(t_2) = 1) + \Pr(x(t_1) = -1, x(t_2) = -1)) \\ &\quad - 1 \cdot (\Pr(x(t_1) = -1, x(t_2) = 1) + \Pr(x(t_1) = 1, x(t_2) = -1)) \end{aligned}$$

Le calcul des probabilités conjointes (d'ordre 2) nécessaire pour déterminer la fonction d'auto-corrélation est simple si on considère les probabilités conditionnelles correspondantes. Par exemple, supposons que  $t_1 > t_2 \Leftrightarrow \tau = t_1 - t_2 > 0$  et calculons

$$\Pr(x(t_1) = 1, x(t_2) = 1) = \Pr(x(t_1) = 1 | x(t_2) = 1) \Pr(x(t_2) = 1).$$

Mais,  $\Pr(x(t_1) = 1 | x(t_2) = 1)$  est la probabilité d'occurrence d'un nombre paire de points dans l'intervalle  $[t_2, t_1]$ :

$$\Pr(x(t_1) = 1 | x(t_2) = 1) = e^{-\lambda \tau} \cosh(\lambda \tau)$$

et donc

$$\Pr(x(t_1) = 1, x(t_2) = 1) = e^{-\lambda \tau} \cosh(\lambda \tau) e^{-\lambda t_2} \cosh(\lambda t_2).$$

De la même manière, on obtient



$$\begin{aligned}\Pr(x(t_1) = -1, x(t_2) = -1) &= e^{-\lambda\tau} \cosh(\lambda\tau) e^{-\lambda t_2} \sinh(\lambda t_2) \\ \Pr(x(t_1) = -1, x(t_2) = 1) &= e^{-\lambda\tau} \sinh(\lambda\tau) e^{-\lambda t_2} \cosh(\lambda t_2) \\ \Pr(x(t_1) = 1, x(t_2) = -1) &= e^{-\lambda\tau} \sinh(\lambda\tau) e^{-\lambda t_2} \sinh(\lambda t_2)\end{aligned}$$

et finalement

$$R(t_1, t_2) = e^{-2\lambda(t_1 - t_2)}, \quad t_1 > t_2.$$

Si on échange les rôles de  $t_1$  et de  $t_2$ , on obtient que l'auto-corrélation de  $x(t)$  est

$$R(t_1, t_2) = e^{2\lambda|t_1 - t_2|}.$$

### Exemple

A partir du processus de l'exemple antérieur, on va construire un processus *stationnaire*, et déterminer sa densité spectrale de puissance.

Soit  $a$  une variable aléatoire, indépendante du processus  $x(t)$ , et qui prend les valeurs  $\pm 1$  avec égale probabilité:

$$\Pr(a = 1) = \Pr(a = -1) = 0.5$$

On obtient facilement

$$E a = 0, \quad E a^2 = 1.$$

Soit  $y(t)$  le processus défini par

$$y(t) = ax(t)$$

On a, avec égale probabilité,  $y(t) = x(t)$  ou  $y(t) = -x(t)$ .

De l'indépendance statistique de  $a$  et  $x(t)$ ,

$$E[y(t)] = E[ax(t)] = E[a]E[x(t)] = 0$$

$$R_y(t_1, t_2) = E[a^2 x(t_1)x(t_2)] = R_x(t_1, t_2)$$

et que  $y(t)$  est *stationnaire au sens large*.

Sa densité spectrale de puissance est, par définition, la transformée de Fourier de sa fonction d'auto-corrélation:

$$R_y(\tau) = e^{-\lambda|\tau|} \Rightarrow S_y(\omega) = \int_{-\infty}^{\infty} e^{-\lambda|\tau|} e^{-j\omega\tau} d\tau = \frac{4\lambda}{4\lambda^2 + \omega^2}$$

### Théorème

Etant donnée une densité spectrale de puissance (fonction positive, symétrique et réelle)  $S(\omega)$ , on peut toujours trouver un processus  $x(t)$  dont  $S(\omega)$  est le spectre.

*Démonstration*

Soit  $a$  la constante

$$a = \frac{1}{\pi} \int_{-\infty}^{\infty} S(\omega) d\omega$$

et  $\omega$  une variable aléatoire de densité de probabilité

$$f(\omega) = \frac{S(\omega)}{\pi a^2}.$$

Soit le processus

$$x(t) = a \cos(\omega t + \phi)$$

$\phi$  où est une autre variable aléatoire, indépendante de  $\omega$  et uniforme en  $-\pi, \pi$ . Alors, le spectre de  $x(t)$  est  $S(\omega)$ , comme on le démontrera.

La moyenne de  $x(t)$  est

$$\begin{aligned}E[a \cos(\omega t + \phi)] &= a E[\cos(\omega t + \phi)] = a E_{\omega} [E_{\phi} [\cos(\omega t + \phi) | \omega]] \\ &= a E_{\omega} [E_{\phi} [\cos(\omega t) \cos(\phi) - \sin(\omega t) \sin(\phi) | \omega]] \\ &= a E_{\omega} [\cos(\omega t) E_{\phi} [\cos(\phi) | \omega]] - a E_{\omega} [\sin(\omega t) E_{\phi} [\sin(\phi) | \omega]] \\ &= a E_{\omega} [\cos(\omega t)] E_{\phi} [\cos(\phi)] - a E_{\omega} [\sin(\omega t)] E_{\phi} [\sin(\phi)]\end{aligned}$$

où le passage de la troisième à la quatrième ligne est justifié par l'indépendance statistique des deux variables aléatoires.

On constate facilement que les deux moyennes sur  $\phi$  sont nulles, et donc que

$$E x(t) = 0.$$

On détermine la fonction d'auto-corrélation,

$$\begin{aligned} R_x(t, t + \tau) &= E[a^2 \cos(\omega t + \phi) \cos(\omega t + \omega \tau + \phi)] \\ &= \frac{a^2}{2} \left\{ E_{\omega \phi} [\cos(\omega \tau) + \cos(2\omega t + \omega \tau + 2\phi)] \right\} \\ &= \frac{a^2}{2} E_{\omega} [\cos(\omega \tau)] \\ &= \frac{a^2}{2} \int_{-\infty}^{\infty} f(\omega) \cos(\omega \tau) d\omega \end{aligned}$$

Le spectre de  $x(t)$  est donné par la transformée de Fourier de la fonction d'auto-corrélation

$$\begin{aligned} S_x(\omega') &= \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega' \tau} d\tau \\ &= \int_{-\infty}^{\infty} \frac{a^2}{2} \int_{-\infty}^{\infty} f(\omega) \cos(\omega \tau) d\omega e^{-j\omega' \tau} d\tau \\ &= \frac{a^2}{2} \int_{-\infty}^{\infty} f(\omega) \int_{-\infty}^{\infty} \cos(\omega \tau) e^{-j\omega' \tau} d\tau d\omega \\ &= \frac{a^2}{2} \int_{-\infty}^{\infty} f(\omega) \pi [\delta(\omega - \omega') + \delta(\omega + \omega')] d\omega = \pi a^2 f(\omega') = \pi a^2 \frac{S(\omega')}{\pi a^2} = S(\omega') \end{aligned}$$

où on a utilisé la transformée de Fourier du  $\cos(\cdot)$  et la symétrie de la densité spectrale originale.

### Exercice

Dire si les fonctions suivantes sont définies non-négatives

$$(a) R(t, s) = \begin{cases} 1 - |t - s|, & |t - s| \leq 1 \\ 0, & |t - s| > 1 \end{cases}$$

$$(b) R(t, s) = e^{-|t-s|}, \quad -\infty < t, s < \infty$$

$$(c) R(t, s) = e^{|t-s|}$$

$$(d) R(t, s) = \begin{cases} 1, & |t - s| \leq 1 \\ 0, & |t - s| > 1 \end{cases}$$

### Bruit Blanc

On appelle *bruit blanc* un processus qui a une densité spectrale de puissance constante, ou, d'une façon équivalente, une fonction d'auto-corrélation qui est nulle en dehors de l'origine.

$$S_n(\omega) = \sigma_n^2 \Leftrightarrow R_n(\tau) = \sigma_n^2 \delta(\tau).$$

Cela veut dire que toutes les variables aléatoires distinctes qui peuvent être définies sont *non-correlées*.

## 2.3 Représentation Entrée - Sortie

Un signal est souvent modélisé comme la réponse d'un système à un autre signal, de caractéristiques plus simples:



Un modèle dynamique discret général (linéaire et invariant) peut être décrit par

$$y(k) = \sum_{n=-\infty}^{\infty} h(n)u(k-n)$$

où  $y(k)$  est la *sortie* à l'instant  $k$ ,  $u(k)$  est l'*entrée* à l'instant  $k$ , et  $h(k)$  est la réponse impulsionnelle du système. Pour des systèmes *causaux*, la sortie  $y(k)$  ne dépendent que des valeurs de l'entrée  $u(n)$  pour des instants de temps  $n \leq k$ , c'est à dire,

$$h(n) = 0, n < 0.$$

### Exercices.

Démontrer les relations suivantes entre les moments du signal d'entrée et de sortie:

$$(a) m_y(k) = \sum_{n=-\infty}^{\infty} h(n)m_u(k-n)$$

$$(b) R_y(\tau) = \sum_{n=-\infty}^{\infty} h(n) \sum_{m=-\infty}^{\infty} h(m)R_u(\tau+n-m)$$

$$(c) S_y(\omega) = S_x(\omega)|H(\omega)|^2$$

où  $H(\omega)$  est la fonction de transfert du système, c.a.d., la transformée de Fourier de la réponse impulsionnelle

Soit  $q^{-1}$  l'opérateur de retard:

$$q^{-1}u(k) = u(k-1).$$

Alors, le système antérieur peut aussi être décrit par l'équation

$$y(k) = H(q^{-1})u(k),$$

où on a défini

$$\begin{aligned} H(q^{-1}) &= \sum_{n=0}^{\infty} h(n)q^{-n} \\ &= h(0) + h(1)q^{-1} + h(2)q^{-2} + \dots \end{aligned}$$

Si le système à une mémoire finie (*réponse impulsionnelle finie*), le polynôme  $H(q^{-1})$  est de la forme

$$H(q^{-1}) = B(q^{-1}) = h(0) + h(1)q^{-1} + h(2)q^{-2} + \dots + h(n)q^{-n}$$

Ceci montre que, à chaque instant, le signal  $y(k)$  est une moyenne pondérée des échantillons du signal d'entrée. On appelle ce type de systèmes (et des signaux de sortie correspondants) *Moyenne Régressive*, et on les représente par  $\mathbf{MA}(n)$ , où  $n$  indique l'ordre du modèle.

Admettons maintenant que  $H(q^{-1})$  peut être l'inverse d'un polynôme d'ordre finie  $q$ :

$$H(q^{-1}) = \frac{b_0}{1 + a_1q^{-1} + \dots + a_pq^{-p}}.$$

Ceci implique que chaque échantillon du signal  $y(k)$  peut être exprimé en fonction de l'entrée et des valeurs de  $y(n)$  pour  $n < k$ :

$$y(k) = -a_1y(k-1) - \dots - a_py(k-p) + b_0u(k).$$

On dit dans ce cas qu'il s'agit d'un signal du type *Auto-Regressive*, et on représente cette classe de signaux par  $\mathbf{AR}(p)$ , où  $p$  indique l'ordre de la régression.

Plus généralement,  $H(q^{-1})$  est une fonction rationnelle, égale au quotient de deux polynômes:

$$H(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})}$$

ce qui conduit à une équation aux différences du type

$$y(k) + a_1 x(k-1) + \dots + a_p x(k-p) = b_0 u(k) + b_1 u(k-1) + \dots + b_n u(k-n).$$

Ce type de signaux est représenté par **ARMA**( $p, n$ ).

Usuellement, le signal que l'on souhaite modéliser est soumis à des perturbations (typiquement, un bruit). Pour modéliser ces perturbations au modèle nominal entrée-sortie, on ajoute à l'équation un deuxième terme:

$$y(k) = H(q^{-1})u(k) + v(k)$$

## 2.4 Modèles d'état

Les modèles d'état sont une façon alternative de décrire une relation entre un signal d'entrée  $u(k)$  et un signal de sortie,  $y(k)$ . Ce type de modèles utilise un signal auxiliaire, qui est l'*état* du système (qui représente la mémoire du système):

$$x(k+1) = F(k)x(k) + G(k)u(k)$$

$$y(k) = H(k)x(k)$$

Cette classe de signaux est plus générale que la précédente (elle décrit des systèmes que ne sont pas invariants, ce qui est indiqué par la dépendance en  $k$  de  $F, G$  et  $H$ ) et est bien adaptée à la modélisation de signaux non stationnaires.

On peut, de la même façon que pour le modèle entrée-sortie, introduire des perturbations aléatoires:

$$x(k+1) = F(k)x(k) + G(k)u(k) + w(k)$$

$$y(k) = H(k)x(k) + e(k)$$

On appelle à la perturbation aléatoire de l'équation qui décrit l'évolution de l'état *bruit du processus*, et à la perturbation qui affecte directement le signal de sortie (le signal observé), le *bruit de mesure*.

En général, ces bruits sont modélisés comme des séquences blanches et non-correlées entre elles. On reviendra sur ce modèle quand on étudiera le *Filtre de Kalman*.

## Contrôlabilité et Observabilité

Soit le modèle d'état

$$x_{k+1} = Ax_k + Bu_k$$

$$y_k = Cx_k + n_k$$

où le vecteur d'état a dimension  $n$ .

Un modèle d'état est *contrôlable à partir de*  $x_0$  en si, pour tout  $x \in \mathbb{R}^n$  il existe un  $t_{x,t_0} \geq t_0$  et une séquence d'entrées  $\{u_i\}_{i=t_0}^{t_{x,t_0}}$  qui mènent à l'état  $x$  en  $t_{x,t_0}$ .

Si le système est contrôlable à partir de tout  $x_0 \in \mathbb{R}^n$ , alors on dit qu'il est *contrôlable* en  $t_0$ , et s'il est contrôlable pour tout  $t_0$ , on dit simplement qu'il est *controllable*.

Un modèle d'état est *observable* en  $t_0$  si pour tout pair de conditions initiales

$x_1(t_0), x_2(t_0) \in \mathfrak{R}^n$ , l'égalité de la réponse du système à une séquence d'entrée  $\{u_i\}, i > t_0$ , implique  $x_1(t_0) = x_2(t_0)$ . Si ceci est vrai pour tout  $t_0$ , on dit que le système est *observable*. Soient

$$O = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^n \\ \vdots \end{bmatrix} \quad C = [B \quad AB \quad \dots \quad A^n B \quad \dots]$$

Le modèle est *observable* ssi le rang de la matrice  $O$  est égal à  $n$ . Le modèle est *contrôlable* ssi le rang de  $C$  est égal à  $n$ .

Ces deux propriétés garantissent que le système a un comportement « normal ». En particulier, si le système est observable, il existent des estimateurs *consistents* de l'état étant données les observations de sa sortie.<sup>1</sup>

En fait, on peut écrire les sorties jusqu'à l'instant  $k$  en fonction d'un sous-bloc de la matrice d'observabilité  $O$ , et de la valeur initiale de l'état

$$Y_k = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^k \end{bmatrix} x_0 + \begin{bmatrix} H_0 & 0 & 0 & 0 \\ H_1 & H_0 & 0 & 0 \\ \vdots & & \ddots & 0 \\ H_k & H_{k-1} & \dots & H_0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_k \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_k \end{bmatrix} = O_k x_0 + N_k$$

où  $H_i = CA^i B$  et la définition du vecteur  $N_k$  qui ne dépend que des  $u_k$  et des  $n_k$  est évidente.

On verra par la suite que pour des modèles d'observation de cette forme, on peut définir des estimateurs consistents, non-biaisés et efficaces du vecteur  $x_0$  étant données les observations  $Y_k$ , ssi l'opérateur d'observation  $O_k$  a un rang égal à la dimension du vecteur à estimer. Donc, si le système est observable, on pourra toujours trouver des estimateurs consistents de l'état.

Si un système est *contrôlable*, pour tout point de  $\mathfrak{R}^n$ , on peut toujours trouver une séquence d'entrées qui mène l'état dans ce point. En particulier, cela veut dire que les ensembles de probabilité nulle de l'espace d'états ont une mesure (de Lebesgue) nulle.

Ceci est facilement vérifié à partir de l'équation

$$x_k = A^k x_0 + \begin{bmatrix} A^{k-1} B & \dots & AB & B \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_{k-1} \\ u_k \end{bmatrix} = A^k x_0 + C_k U_k$$

où les définitions de  $C_k$  et de  $U_k$  sont évidentes.

On peut donc voir que le si  $C_k$  n'est pas de rang complet, il peuvent toujours exister des valeurs de l'état qui ne seront pas atteints.

<sup>1</sup> Note : Comme on le verra dans un Chapitre ultérieur, on dit qu'un estimateur est consistant s'il converge (avec probabilité un) vers la vraie valeur du paramètre estimé quand le nombre d'observations tend vers  $\infty$ .

Ces deux propriétés ensemble garantissent la minimalité de la représentation d'état : il n'existe pas de système équivalent avec un vecteur d'état de dimension inférieure à  $n$ .

## 2.5 Relation entre modèles entrée-sortie et modèles d'état

Un modèle d'état stationnaire (où les paramètres du modèle ne sont pas fonction du temps) peut facilement être mis sous la forme d'un modèle entrée sortie.

Si on utilise la définition de l'opérateur de retard dans l'équation d'état, on obtient

$$qx(k) = Fx(k) + Gu(k) + w(k) \Leftrightarrow x(k) = q^{-1}(I - q^{-1}F)^{-1}Gu(k) + q^{-1}(I - q^{-1}F)^{-1}w(k)$$

Si on utilise ce résultat dans l'équation de sortie,

$$\begin{aligned} y(k) &= q^{-1}H(I - q^{-1}F)^{-1}Gu(k) + q^{-1}H(I - q^{-1}F)^{-1}w(k) + e(k) \\ &= q^{-1}H(I - q^{-1}F)^{-1}[Gu(k) + w(k)] + e(k) \\ &= H^*(q^{-1})u'(k) + e(k) \end{aligned}$$

qui est une description entrée-sortie, qui donne le signal  $y(k)$  comme réponse au signal d'entrée

$$u'(k) = Gu(k) + w(k).$$

Le passage d'un modèle entrée sortie à la forme d'un modèle d'état peut être fait de plusieurs manières (il n'existe pas un modèle d'état unique qui correspond à un modèle entrée-sortie donné). Une des représentations possibles est connue par la *forme canonique observable*, et fait correspondre à un modèle entrée-sortie

$$A(q^{-1})y(k) = B(q^{-1})u(k) + e(k)$$

l'équation d'état suivante:

$$\begin{aligned} x(k+1) &= \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix} x(k) + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} u(k) \\ y(k) &= [1 \quad 0 \quad \cdots \quad 0]x(k) + e(k) \end{aligned}$$

Dans cette équation, les paramètres  $a_i$  et  $b_i$  sont les coefficients des polynômes qui définissent le modèle entrée-sortie:

$$\begin{aligned} A(q^{-1}) &= 1 + a_1q^{-1} + \dots + a_nq^{-n} \\ B(q^{-1}) &= b_1q^{-1} + \dots + b_nq^{-n} \end{aligned}$$

Le nom de ce modèle est dû au fait que le signal observé est la première composante du vecteur d'état.

### Exemple

Cet exemple illustre le fait que la représentation d'état n'est pas unique.  
Considérons le model AR(2)

$$y_k + 1.2y_{k-1} + 0.35y_{k-2} = u_k$$

Etudions sa conversion en modèle d'état.

Soit

$$\xi_k = \begin{bmatrix} y_{k-1} \\ y_k \end{bmatrix}$$

On peut vérifier que le modèle

$$\begin{aligned}\xi_{k+1} &= A\xi_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_{k+1} \\ y_k &= \begin{bmatrix} 0 & 1 \end{bmatrix} \xi_k\end{aligned}$$

où la matrice  $A$  est définie comme

$$A = \begin{bmatrix} 0 & 1 \\ -0.35 & -1.2 \end{bmatrix}$$

est équivalent au modèle original.

Le choix

$$\psi_k = \begin{bmatrix} y_{k-1} \\ y_{k-u} \end{bmatrix}$$

conduit à un modèle où la définition de la matrice  $A$  est la même, mais

$$\begin{aligned}\psi_{k+1} &= A\psi_k + \begin{bmatrix} 1 \\ -5 \end{bmatrix} u_{k+1} \\ y_k &= \begin{bmatrix} 0 & 1 \end{bmatrix} \psi_k + u_k\end{aligned}$$

## 2.6 Décomposition de Wold : innovations

Dans cette section, on reprend la formulation d'espace de Hilbert de variables aléatoires, présentée à la fin du Chapitre précédent.

### 2.6.1 Espace de Hilbert engendré par un processus. Projection orthogonale

Soit  $x = \{x_t, t \in T\}$  un processus stochastique défini dans l'espace de probabilité  $(\Omega, \mathcal{B}, \mathcal{P})$ , et considérons l'espace  $\mathbf{H}$  de toutes les combinaisons linéaires de coefficients réels

$$\sum_{i=1}^N \alpha_i x_{t_i}$$

de valeur quadratique moyenne finie :

$$\mathbf{E} \left[ \sum_{i=1}^N \alpha_i x_{t_i} \right]^2 < \infty$$

$\mathbf{H}$  est un espace pré-Hilbertien, avec produit interne

$$\langle a, b \rangle_{\mathbf{H}} = \mathbf{E}(ab)$$

Mais, comme on l'a vu à la fin du Chapitre 1, l'espace des variables aléatoires de valeur quadratique moyenne est complet, c'est dire, les limites de ces séquences existent en  $L^2(\Omega, \mathcal{B}, \mathcal{P})$ . On appelle l'espace obtenu en complétant  $\mathbf{H}$  avec les limites de ces sommes, *l'espace de Hilbert engendré par le processus  $x$* , et on le représente par  $\mathbf{H}^x$ .

#### Exemple

Soit  $x_n$  le processus défini par

$$x_n = \alpha x_{n-1} + e_n, \quad n = 1, 2, \dots$$

où  $e_n$  est une séquence blanche stationnaire,  $\mathbf{E}e_n^2 = \sigma^2 < \infty$

Alors

$$x_n = \sum_{i=0}^{\infty} \alpha^i e_{n-i}$$

Soit

$$x_n^N = \sum_{i=0}^N \alpha^i e_{n-i} \in \mathbf{H}$$

Pour tout  $M < N$

$$\|x_n^N - x_n^M\|^2 = \mathbf{E} \left( \sum_{i=M+1}^N \alpha^i e_{n-i} \right)^2 = \sigma^2 \sum_{i=M+1}^N \alpha^{2i} < \frac{\sigma^2 \alpha^{2(M+1)}}{1 - \alpha^2}$$

et est donc une séquence de Cauchy, qui converge vers la variable aléatoire  $x_n \in \mathbf{H}^x$ .

Soit  $\mathbf{H}_n^x$  l'espace engendré par  $\{x_s, s \leq t, t \in T\}$  et  $u < t$ . Alors,  $\mathbf{H}_u^x \subseteq \mathbf{H}_t^x \subseteq \mathbf{H}^x \subseteq L^2(\Omega, \mathcal{B}, \mathcal{P})$ . Soit  $\mathbf{H}$  l'espace engendré par un ensemble de variables aléatoires,  $v$  en élément de  $\mathbf{H}$ , et  $\mathbf{M} \subset \mathbf{H}$  un sous-espace de  $\mathbf{H}$ . On représente par  $(v | \mathbf{M})$  la projection orthogonale de  $v$  en  $\mathbf{M}$ , qui est définie par la propriété suivante :

$$\inf_{m \in \mathbf{M}} \|v - m\| = \|v - (v | \mathbf{M})\|$$

### Exercice

Soit  $x$  un processus stationnaire au sens large de moyenne nulle, et  $\{\mathbf{H}_n^x, n \in \mathbb{Z}\}$  la séquence emboîtée de sous-espaces de  $\mathbf{H}_n^x$  associée. Montrer que  $\{(x_n | \mathbf{H}_{n-1}^x), n \in \mathbb{Z}\}$  est aussi un processus stationnaire au sens large.

*Suggestion*

Soit  $\{\hat{x}_{n+k,p}, p \in \mathbb{Z}\}$  la projection orthogonale de  $x_{n+k}$  dans l'espace engendré par  $\{x_{n+k-p}, \dots, x_{n+k-2}, x_{n+k-1}\}$ . Montrer que  $\{\hat{x}_{n+k,p}\}$  converge vers  $(x_{n+k} | \mathbf{H}_{n+k-1}^x)$  quand  $p \rightarrow \infty$ . Alors, par la stationnarité de la séquence initiale, et la continuité du produit interne,

$$\langle (x_{n+k+\tau} | \mathbf{H}_{n+k+\tau-1}^x), (x_{n+\tau} | \mathbf{H}_{n+\tau-1}^x) \rangle = \langle (x_{n+\tau} | \mathbf{H}_{n+k-1}^x), (x_n | \mathbf{H}_{n-1}^x) \rangle$$

## 2.6.2 Décomposition de Wold

Un des résultats fondamentaux de la théorie des processus stationnaires au sens large est la décomposition de Wold, qui permet d'interpréter n'importe quel processus purement aléatoire comme la réponse d'un système linéaire à un bruit blanc.

L'idée de base est la suivante. Admettons que  $x_n$  appartient à l'espace  $\mathbf{H}_n^x$  et qu'il n'est pas contenu dans un sous-espace propre  $\mathbf{H}_{n-1}^x$  de  $\mathbf{H}_n^x$ . Soit  $\hat{x}_{n-1} = (x_n | \mathbf{H}_{n-1}^x)$ . Alors,

$$x_n = \hat{x}_{n-1} + e_n^x, \quad \hat{x}_{n-1} \in \mathbf{H}_{n-1}^x, \quad e_n^x \perp \mathbf{H}_{n-1}^x$$

La même procédure peut être appliquée  $\hat{x}_{n-1}$  par projection dans  $\mathbf{H}_{n-2}^x \subset \mathbf{H}_{n-1}^x$ , pour obtenir

$$x_n = \hat{x}_{n-2} + e_{n-2}^x + e_n^x, \quad \hat{x}_{n-2} \in \mathbf{H}_{n-2}^x, \quad e_n^x \perp e_{n-1}^x \perp \mathbf{H}_{n-2}^x.$$

Par application répétée de cette procédure on écrit  $x_n$  comme la somme de variables aléatoires  $\{e_{n-p}^x, p \in \mathbb{Z}\}$  non-correlés (séquence blanche).

Soit  $\mathbf{H}_n = Sp\{\dots, x_{-n}, x_n\}$ , et  $\mathbf{H}_{-\infty} = \bigcap_{n \in \mathbb{Z}} \mathbf{H}_n$ ,  $\mathbf{H}_{\infty} = \bigcup_{n \in \mathbb{Z}} \mathbf{H}_n$ . On dit que  $x$  est un processus purement déterministe ssi  $x_n \in \mathbf{H}_{n-1}^x, \forall n \in \mathbb{Z}$ . Si  $x$  n'est pas purement déterministe, on dit qu'il est *non-déterministe*. Si  $\mathbf{H}_{-\infty} = \{0\}$  on dit que  $x$  est *purement aléatoire*.



On énonce quelques propriétés des espaces de Hilbert qui nous seront utiles par la suite.

1. Si  $\mathbf{M}_{-\infty} = \bigcap_{n \in \mathbb{Z}} \mathbf{M}_n$  où les  $\mathbf{M}_n$  forment une famille emboîtée de sous-espaces,

$$\mathbf{M}_{n-1} \subset \mathbf{M}_n, \forall n, \text{ alors, } \forall h \in \mathbf{H}$$

$$(h | \mathbf{M}_{-\infty}) = \lim_{n \rightarrow -\infty} (h | \mathbf{M}_n).$$

2. Soit  $\mathbf{M}_{\infty} = \bigcup_{n \in \mathbb{Z}} \mathbf{M}_n$ . Alors

$$(h | \mathbf{M}_{\infty}) = \lim_{n \rightarrow \infty} (h | \mathbf{M}_n)$$

3. Soit  $\{e_k, k \in \mathbb{Z}\}$  une séquence stationnaire au sens large, blanche, dans l'espace de Hilbert  $\mathbf{H}$  et soit  $\mathbf{E}_n = Sp\{e_0, \dots, e_{n-1}, e_n\}$ ,  $\mathbf{E}_{\infty} = \bigcup_{n \in \mathbb{Z}} \mathbf{E}_n$ . Soit  $h$  une autre variable dans  $\mathbf{H}$ . Alors,

$$(h | \mathbf{E}_{\infty}) = \sum_{k=0}^{\infty} A_k e_k$$

où  $A_k$  satisfait

$$A_k e_k = \langle h, e_k \rangle$$

La représentation est unique si

$$\Sigma = E(e_n^2) > 0.$$

### Définition

Le processus défini par

$$e_n = x_n - (x_n | \mathbf{H}_{n-1}), \quad n \in \mathbb{Z}$$

est appelé les *innovations (linéaires)* du processus  $x$ . Si  $E e_n^2 > 0, \forall n$ , on dit que le processus est *de rang complet*.

### Relation avec estimation d'erreur quadratique moyenne minimale

Remarquons que  $(x_n | \mathbf{H}_{n-1})$  est la variable aléatoire qui peut s'exprimer comme combinaison linéaires des valeurs précédentes du processus et qui minimise la valeur quadratique de la différence par rapport à la valeur du processus (par définition de norme en  $L^2(\Omega, \mathcal{B}, \mathcal{P})$ ). Pour cette raison, on appelle aussi à  $e_n$  **l'erreur de prédiction**. Notons cependant, que cette projection est l'estimateur optimal uniquement si on se restreint à des estimateurs **linéaires**.

Soit  $\mathbf{N}_n = Sp\{e_k, k \leq n\}$ , c'est à dire, l'espace de Hilbert engendré par les innovations jusqu'à l'instant  $n$ . Alors

1.  $\mathbf{H}_n = \mathbf{H}_m + Sp\{e_{m+1}, \dots, e_n\}, n > m$
2.  $\mathbf{H}_m \perp Sp\{e_{m+1}, \dots, e_n\}, n > m$
3.  $\mathbf{H}_n = \mathbf{H}_{-\infty} + \mathbf{N}_n, \mathbf{H}_{-\infty} \perp \mathbf{N}_{\infty}, \mathbf{N}_{-\infty} = \{0\}$

### Théorème de Wold

Soit  $x$  un processus stationnaire au sens large, et  $e_n$  son processus d'innovations. Alors,

- (i)  $x_n = u_n + v_n$  où  $u_n = (x_n | \mathbf{N}_n)$ ,  $v_n = (x_n | \mathbf{H}_{-\infty})$  et  $u_n \perp v_n, \forall n$ .
- (ii) le processus  $u$  est stationnaire au sens large et a la représentation

$$u_n = \sum_{k=0}^{\infty} A_k e_k$$

où  $e$  est le processus d'innovations de  $x$

$$E(e_k e_j) = \Sigma \delta_{kj}$$

Et

$$0 \leq \sum A_k \Sigma A_k \leq \infty, \quad E(x_0 e_{-k}) = E(u_0 e_{-k}) = A_k e_k$$

$$A_0 \Sigma = \Sigma = \Sigma A_0$$

Si le processus  $x$  a range complet, alors la séquence  $A$  est unique, et  $A_0 = 1$ .

(iii)  $v$  est un processus purement déterministe stationnaire et  $\mathbf{H}_{-\infty} = Sp\{\dots, v_{n-1}, v_n\}$ .

En conclusion, ce théorème nous permet de écrire un processus stationnaire au sens large dans la forme

$$x_n = \sum_{k=0}^{\infty} A_k e_{n-k} + v_n$$

où  $e$  est un processus stationnaire blanc et orthogonal au processus déterministe  $v$ . Si  $x$  est purement aléatoire, alors  $v_n = 0, \forall n$ , et on a simplement

$$x_n = \sum_{k=0}^{\infty} A_k e_{n-k}$$

Soit  $x$  un processus de rang complet et  $\Sigma^{1/2}$  une racine carrée symétrique de  $\Sigma$ . Définissons

$\tilde{A}_k = A_k \Sigma^{1/2}$ , et  $\varepsilon_n = \Sigma^{1/2} e_n$ , on peut vérifier facilement qu'une représentation équivalente est

$$x_n = \sum_{k=0}^{\infty} \tilde{A}_k \varepsilon_{n-k} + v_n, \quad E(\varepsilon_k \varepsilon_p) = \delta_{kp}$$

On appelle cette représentation la *décomposition de Wold normalisée*.

## Exercices

1. Soit  $w$  une variable aléatoire uniforme en  $[0,1]$ , et  $X(t)$  le signal aléatoire défini par

$$X_t(w) = tw, \quad 0 \leq t \leq 1$$

(a) Déterminer les caractérisations d'ordre 1 et 2 de ce processus.

(b) Déterminer la moyenne et la fonction de covariance de ce processus.

2. L'intégrale d'un processus aléatoire dans un intervalle fini :

$$s = \int_a^b x(t) dt$$

est une variable aléatoire, et sa valeur dans une expérience particulière est la surface sous la courbe  $x(t)$  dans l'intervalle  $a, b$ . Déterminer la moyenne et la variance de  $s$ .

3. Déterminer la fonction d'auto-corrélation du processus

$$x(t) = r \cos(\omega t + \theta),$$

où  $r$  et  $\theta$  sont des variables aléatoires indépendantes et  $\theta$  est uniformément distribuée en  $-\pi, \pi$ .

4. Soit  $x(t)$  un processus stationnaire de moyenne nulle et fonction d'auto-corrélation  $R(\tau)$ . Déterminer la fonction d'auto-corrélation de

$$y(t) = x(t) + f(t)$$

où  $f(t)$  est une fonction déterministe (connue).

5. Démontrer les deux théorèmes sur les processus cyclo-stationnaires énoncés dans ce chapitre.

### *Théorème*

Si  $x(t)$  est un processus cyclostationnaire au sens strict et  $\theta$  est une variable aléatoire uniforme en  $0, T$ , alors  $\bar{x}(t) = x(t - \theta)$  est un processus stationnaire au sens strict.

### *Théorème*

Si  $x(t)$  est un processus cyclo-stationnaire au sens large et  $\theta$  est une variable aléatoire uniforme en  $0, T$ , alors  $\bar{x}(t) = x(t - \theta)$  est un processus stationnaire au sens large, et

$$m_{\bar{x}}(t) = m_{\bar{x}} = \frac{1}{T} \int_T m_x(t) dt$$

$$R_{\bar{x}}(\tau) = \frac{1}{T} \int_T R_x(t + \tau, t) dt$$

6. Admettre qu'un processus prend avec égale probabilité les valeurs  $\pm 1$  dans chaque intervalle  $(n-1)T, nT$  de longueur  $T$ , et que ses valeurs dans des intervalles disjoints sont indépendantes.

i) déterminer la fonction d'auto-corrélation de  $x(t)$ .

ii) déterminer l'auto-corrélation du processus décalé  $x(t - \theta)$  où  $\theta$  est une variable aléatoire uniformément distribuée en  $-T/2, T/2$ . Déterminer son spectre.

7. Montrer que si  $x(t)$  est strictement stationnaire et  $\varepsilon$  est une variable aléatoire indépendante de  $x(t)$  alors  $x(t - \varepsilon)$  est strictement stationnaire.

8. Montrer que si l'entrée d'un système causal est blanche, de spectre  $S(\omega) = \kappa$ , la valeur quadratique moyenne de la sortie est

$$E[y^2(t)] = \kappa \int_0^\infty h^2(t) dt$$

où  $h(t)$  est la réponse impulsionnelle du système.

9. Considerer un processus défini dans l'intervalle  $[0, T]$  avec fonction de covariance

$$R(t, s) = \min(t, s)$$

Montrer que les fonctions propres de  $R(t, s)$  et les valeurs propres, c'est à dire, les solutions de l'équation intégrale

$$\int_0^T R(t, s) \phi(s) ds = \lambda \phi(t), \quad 0 \leq t \leq T$$

sont données par

$$\lambda_n = \frac{T^2}{(n + \frac{1}{2})^2 \pi^2}, \quad n=0, 1, 2, \dots$$

$$\phi(t) = A \sin \frac{1}{\sqrt{\lambda_n}} t, \quad t \in [0, T]$$

10. Soit

$$R(t, s) = \begin{cases} 1 - |t - s|, & 0 \leq |t - s| \leq 1 \\ 0, & |t - s| > 1 \end{cases}$$

Déterminer les valeurs propres et un ensemble ortho-normé de fonctions pour

$$\int_0^{1/2} R(t, s) \phi(s) ds = \lambda \phi(t), \quad 0 \leq t \leq 1/2$$

Suggestion :  $\frac{\partial^2 R(t, s)}{\partial t^2} = -2\delta(t - s), \quad 0 < t, s < 1/2$

11. Un système de communications numérique (binaire) utilise un codage PAM (*Pulse Amplitude Modulation*). Le signal transmit pendant la période de symbole  $T$

$$x(t) = a_n \cos(\omega_o t + \theta), \quad t \in T,$$

où  $a_n$  est le  $n$ -ème symbole transmit :

$$a_n = \begin{cases} A, & \text{avec probabilité } p \\ 0, & \text{avec probabilité } 1 - p \end{cases}$$

Considérez le vecteur obtenu par échantillonnage de  $x(t)$  pendant l'intervalle correspondant à un même symbole :

$$\mathbf{r} = \begin{bmatrix} x(t_1) \\ \vdots \\ x(t_N) \end{bmatrix}$$

(i) Montrez que le vecteur aléatoire  $\mathbf{r}$  peut s'écrire comme

$$\mathbf{r} = a_n \mathbf{m},$$

où  $\mathbf{m}$  est un vecteur déterministe.

(ii) Quelle est la dimension de l'espace de Hilbert engendré par  $\mathbf{r}$ ,  $\mathbf{H}^{\mathbf{r}}$ , comme sous-ensemble de  $L^2(\Omega, \mathcal{B}, \mathcal{P})$  ?

(iii) Considérez que l'on observe la superposition de  $\mathbf{r}$  avec un bruit blanc :

$$\mathbf{z} = \mathbf{r} + \mathbf{n}$$

où  $\mathbf{n}$  est un vecteur aléatoire indépendant de  $\mathbf{r}$ . Quelle est la dimension de  $\mathbf{H}^{\mathbf{z}}$ , l'ensemble de toutes les statistiques linéaires scalaires de  $\mathbf{z}$  ?

$$s = \mathbf{h}^T \mathbf{z}, \mathbf{h} \in \mathfrak{R}^N$$

(iv) Admettez que l'on ne connaît pas le modèle du signal de transmission, sachant uniquement qu'il s'agit d'une modulation PAM, et que l'on observe  $M$  symboles (observations bruitées). Sans rien savoir de plus, comment peut-on reconstruire la séquence transmise ?

(v) Simulez le comportement de l'alinéa précédente, en considérant du bruit blanc.